



**HAL**  
open science

# Non-asymptotic estimates of invariant measures and regularisation by a degenerate noise for a chain of ordinary differential equations

Igor Honoré

► **To cite this version:**

Igor Honoré. Non-asymptotic estimates of invariant measures and regularisation by a degenerate noise for a chain of ordinary differential equations. Analysis of PDEs [math.AP]. Université Paris-Saclay, 2018. English. NNT: . tel-01961770

**HAL Id: tel-01961770**

**<https://theses.hal.science/tel-01961770v1>**

Submitted on 20 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT

de

L'UNIVERSITÉ PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH, ED 574)

*Établissement d'inscription* : Université d'Évry-Val d'Essonne

*Laboratoire d'accueil* : Laboratoire de mathématiques et modélisation d'Évry, UMR 8071  
CNRS-INRA

*Spécialité de doctorat* : Mathématiques appliquées

**Igor HONORÉ**

Estimations non-asymptotiques de mesures invariantes et  
régularisation par un bruit dégénéré de chaînes d'Équations  
Différentielles Ordinaires

*Date de soutenance* : 05/11/18

*Après avis des rapporteurs* : DENIS TALAY (Inria Sophia-Antipolis)  
ALEXANDER YU. VERETENNIKOV (University of Leeds)

*Jury de soutenance* :

FRANÇOIS BOLLEY	(LPSM, Jussieu) Examineur
ARNAUD GLOTER	(LaMME, Évry) Examineur
ALESSANDRA LUNARDI	(Università di Parma) Présidente du jury
STÉPHANE MENOZZI	(LaMME, Évry) Directeur de thèse
GILLES PAGÈS	(LPSM, Jussieu) Examineur
DENIS TALAY	(Inria Sophia-Antipolis) Rapporteur



# Remerciements

---

Ce manuscrit étant relativement long et dense, je profite de ces premières pages, alors que le lecteur a encore l'esprit frais, pour remercier les personnes sans qui ce travail n'aurait pas été possible.

En premier lieu, je tiens à remercier chaleureusement mon directeur de thèse, Stéphane Menozzi, avec qui j'ai eu la chance de me lancer dans cette aventure doctorale. Ce brillant mathématicien s'est énormément investi dans ma thèse. Depuis l'autre bout du monde, nous avons travaillé d'arrache-pied à des heures indues (quel que soit le fuseau horaire). Lors de mon stage de master sous sa direction, j'ai eu le plaisir de travailler avec Gilles Pagès. Ce fut mon premier contact extérieur avec la recherche en mathématiques. Cette collaboration, très intéressante et fructueuse (cf. Chapitre 3), m'a donné le goût de la constante optimale, cf. Chapitre 4. Ce chapitre assez technique, que j'ai signé seul, a été rendu lisible (non sans souffrance pour la première version) grâce à l'assistance appuyée de Stéphane.

Lors de la thèse, j'ai aussi travaillé avec Arnaud Gloter et Dasha Loukianova, afin d'adapter les techniques établies au Chapitre 3 à des sauts. Cette collaboration de chercheurs a été très enrichissante.

Enfin, ma dernière année de thèse a été marquée par ma collaboration avec Paul-Éric Chaudru de Raynal et Stéphane Menozzi. Ce fut une année intense qui a abouti aux Chapitres 6 et 7. Nous avons reculé (*backward*) pour mieux avancer (*forward*).

Ces trois belles années de doctorat m'ont ouvert quantité de sujets de recherche que j'ai hâte d'explorer.

Mes années doctorales ont aussi été l'occasion de diriger des TD d'Algèbre. Je remercie Abdelmejid Bayad avec qui j'ai eu le plaisir de m'occuper des L3 pendant trois ans. Je remercie également Charlotte Scribot et Shiqi Song, avec qui j'ai préparé d'autres TD, de m'avoir aidé d'un point de vue pédagogique.

Je suis très reconnaissant aux rapporteurs, Denis Talay et Alexander Yu. Veretennikov, d'avoir consacré un temps certain à la lecture de mon manuscrit (j'avoue avoir eu quelques scrupules à envoyer ma thèse de plus de 300 pages au lendemain de la victoire des bleus).

Je souhaite par ailleurs exprimer ma gratitude envers le jury de soutenance, François Bolley, Arnaud Gloter, Alessandra Lunardi, Stéphane Menozzi, Gilles Pagès et Denis Talay, pour avoir accepté d'assister à la soutenance. J'ai une pensée particulière pour ceux venus de loin. Je pense notamment à Enrico Priola

Lors de ce doctorat, j'ai fait beaucoup de belles rencontres :

Je pense notamment à Pierre-Gilles Lemarié-Rieusset, avec qui Stéphane Menozzi et moi-même faisons partie des «matheux du 18<sup>ème</sup>» ; à Valérie Gontier-Picot qui a toujours brillé par sa disponibilité et sa gentillesse ; mais aussi à El Maouloud Ould Baba qui me transmettait sa bonne humeur chaque matin. J'ai beaucoup apprécié le périple avec Diego

Chamoro pour le colloque à Modène. Je remercie Etienne Chevalier grâce à qui j'ai pu effectuer le master à l'UEVE.

Merci à Oscar Jarrin, Mhamed Gaïgi, Kawther Mayoufi, Florian Rasamoely, Ricardo Romo et Mohamed Seddik de m'avoir accueilli dans l'équipe des doctorants dès mon stage.

J'ai eu également beaucoup de plaisir à échanger avec Chiara Amorino, Halaleh Kamari, Vincent Runge, Wissal Sabbagh, Elizabeth Zuniga, Yannick Armenti, Quentin Clairon, Virginie Stanislas.

Je n'oublie pas bien sûr Christophe Profeta, Stéphane Crépey, Monique Jeanblanc, Gilles Lacombe, Julia Matos, Sergio Pulido, Abass Sagna, Vincent Torri, Alexandre Vidal ainsi que les autres chercheurs et personnels de l'université.

Je remercie mon «grand frère de thèse», Lorick Huang pour ses conseils avisés, et souhaite bon courage à Lorenzo Marino qui me succède auprès de Stéphane Menozzi.

Enfin et surtout, je remercie ma famille et mes amis dont le soutien sans faille a été déterminant :

Ilona, que je connais depuis peu, et qui me donne le sourire chaque jour.

Cécilia, sans qui je n'aurais jamais pu réussir, ma première supportrice, celle qui est toujours là pour moi.

Mes parents, pour leurs encouragements et leur implication dans mon parcours scolaire.

Mamie Lène, Mamie Thé, Alaric, Coralie, Basile, Louis, Siegrid, Inaya, Stan, Raoul d'être à mes côtés.

Arlette et Pierre pour leur soutien et leurs importants coups de main, notamment lors de mes déplacements.

Ma marraine, Jean-Marc, Victor et Léa.

Delphine, Christian, Martin et Axelle.

Mes oncles, tantes, cousins, cousines.

Les Phoques pour leur soutien.

Stéphane qui prend la peine de venir me voir régulièrement depuis le Canada.

L'U.S Bondues Judo. La chorale Championnet.

Guimlie, et Ulysse...

Enfin, je serai éternellement reconnaissant aux équipes des hôpitaux Delafontaine à Saint-Denis et Lariboisière à Paris qui ont rendu possible ce qui semblait impossible.

À la mémoire de Papi Cel, Mamie Leau, Papi Bog et Babcia.

# Organisation du mémoire

---

Afin de faciliter la lecture du document, les principaux symboles sont regroupés à la page 14 avec une brève explication, ainsi que la page où ledit symbole apparaît pour la première fois (dans les deux dernières parties du mémoire).

La première partie de la thèse, rédigée en français, est liminaire et permet de contextualiser les travaux. Dans cette partie, nous présentons également les principaux outils utilisés dans la suite du mémoire. Cette introduction se divise en deux parties comme le reste du manuscrit : - Estimations non-asymptotiques de mesures invariantes - Régularisation par un bruit dégénéré.

La suite de la thèse est rédigée en anglais. Chaque chapitre contient une petite introduction, ainsi qu'une présentation des notations utilisées. La première partie est composée de trois chapitres :

- Le Chapitre 3, **A first non-asymptotic concentration result**, est un travail effectué en collaboration avec Stéphane Menozzi et Gilles Pagès. Ce chapitre est disponible en prépublication [HMP18]. Il traite d'inégalités de concentration non-asymptotiques des déviations de la mesure invariante avec une mesure empirique approchant cette dernière.
- Le Chapitre 4, **A sharp non-asymptotic concentration**, est issu d'un article, écrit seul, disponible en prépublication [Hon18]. Il présente des inégalités de concentration du même type que ceux du Chapitre 1 mais avec la constante de concentration optimale.
- Le Chapitre 5, **Extension to SDEs driven by a Poisson compound process**, a été réalisé en collaboration avec Arnaud Gloter et Dasha Loukianova et a donné lieu à une prépublication [GHL18]. Dans ce chapitre, on adapte l'algorithme présenté dans les chapitres précédents à une certaine classe d'EDS à sauts.

La seconde partie de la thèse comporte deux chapitres :

- Le Chapitre 6, **Sharp Schauder Estimates for some Degenerate Kolmogorov Equations**, en coopération avec Paul Éric Chaudru de Raynal et Stéphane Menozzi. Nous établissons des estimées de Schauder associées au problème de Cauchy pour l'équation de Kolmogorov.
- Le Chapitre 7, **Strong uniqueness of degenerate stochastic system**, également écrit avec Paul Éric Chaudru de Raynal et Stéphane Menozzi. Nous démontrons l'unicité forte à régularité minimale pour une chaîne d'oscillateurs bruités en la première composante (cadre dégénéré).



# Table des matières

Liste des Symboles	11
Table des figures	15
<b>I Introduction</b>	<b>17</b>
<b>1 Estimations de mesures invariantes</b>	<b>19</b>
1 Processus Ergodiques . . . . .	19
1.1 Mesure invariante . . . . .	19
1.2 Théorèmes Ergodiques . . . . .	21
1.3 Discrétisation du processus . . . . .	23
1.4 Inégalités de concentration . . . . .	26
1.5 De $\nu_n(\mathcal{A}\varphi)$ à l'estimation de $\nu(f)$ pour une source $f$ donnée . . . . .	27
2 Premier résultat de concentration . . . . .	27
2.1 Améliorations du Théorème 2 . . . . .	30
2.2 Sur le problème de Poisson . . . . .	31
2.3 Résultat de type Slutsky . . . . .	33
2.4 Source Lipschitz . . . . .	33
3 Inégalité de concentration optimale . . . . .	34
3.1 Source Lipschitz . . . . .	35
4 Extension aux EDS dirigées par un Lévy . . . . .	36
<b>2 Régularisation par un bruit dégénéré</b>	<b>39</b>
1 Régularisation par un bruit dégénéré . . . . .	39
1.1 Équation de Kolmogorov et chaîne dégénérée . . . . .	39
1.2 Présentation du modèle . . . . .	41
1.3 Propriété d'échelles . . . . .	42
1.4 Résultats existants . . . . .	45
1.5 Estimées de Schauder établies . . . . .	45
1.6 Grandes lignes de l'analyse . . . . .	46
1.7 Contrôle de la norme uniforme . . . . .	49
1.8 Conclusion . . . . .	53
1.9 Unicité forte . . . . .	54



<b>II</b>	<b>Non-asymptotic concentration inequalities</b>	<b>61</b>
<b>3</b>	<b>A first non-asymptotic concentration result</b>	<b>63</b>
1	Introduction . . . . .	63
1.1	Setting . . . . .	63
1.2	Assumptions and Related Asymptotic Results . . . . .	66
1.3	Notations . . . . .	71
2	Main results . . . . .	72
2.1	Result of non-asymptotic Gaussian concentration . . . . .	72
2.2	Uniqueness of the invariant distribution and Regularity issues for the Poisson problem . . . . .	75
2.3	Practical Deviation Bounds . . . . .	79
3	Proof of the concentration results (Theorem 2) . . . . .	81
3.1	Strategy . . . . .	81
3.2	Explicit controls on the remainders . . . . .	83
3.3	Proof of the Technical Lemmas . . . . .	88
4	A refinement when $\ \sigma\ ^2 - \nu(\ \sigma\ ^2)$ is a Coboundary . . . . .	97
5	Smoothness Results for the Poisson Problems (Proof of Theorem 4) . . . . .	103
5.1	Proof of Theorem 4 . . . . .	103
5.2	Proof of the Practical Results of Section 2.3 . . . . .	111
5.3	Regularization of Lipschitz Sources . . . . .	113
6	Applications . . . . .	116
6.1	Non-Asymptotic Deviation Bounds in the Almost Sure CLT . . . . .	116
6.2	Numerical Results . . . . .	119
<b>4</b>	<b>A sharp non-asymptotic concentration</b>	<b>125</b>
1	Introduction . . . . .	125
1.1	Statement of the problem . . . . .	125
2	Assumptions and Existing Results . . . . .	129
2.1	General notations . . . . .	129
2.2	Hypotheses . . . . .	130
2.3	On some Related Existing Result . . . . .	132
3	Main results . . . . .	135
3.1	User's guide to the proof . . . . .	135
3.2	Technical lemmas and Proof of the Main Results . . . . .	139
4	Proofs of technical lemmas . . . . .	146
4.1	Remainders from the Taylor decomposition . . . . .	146
4.2	Asymptotics in the parameter $\rho$ . . . . .	147
4.3	Technical Lemmas for Unbounded Innovations . . . . .	149
5	Regularity Results and Consequences . . . . .	154
5.1	Assumptions and Regularity Results . . . . .	155
5.2	Concentration inequalities for a Lipschitz source in a non-degenerate setting . . . . .	156
6	Optimisation over $\rho$ under <i>Gaussian</i> and <i>super Gaussian deviations</i> . . . . .	158
7	Numerical Results . . . . .	164

8	Appendix: Computation of asymptotic analysis . . . . .	166
<b>5</b>	<b>Extension for a Poisson compound process</b>	<b>169</b>
1	Introduction . . . . .	169
1.1	Setting . . . . .	169
1.2	General notations . . . . .	172
1.3	Hypotheses . . . . .	173
1.4	Existing results . . . . .	176
2	Main results . . . . .	177
2.1	Result of non-asymptotic Gaussian concentration . . . . .	177
2.2	Strategy . . . . .	178
2.3	Proof of our main result . . . . .	181
2.4	Proof of the Gaussian property of the jump innovation . . . . .	184
3	Exponential integrability of the square root of Lyapunov function. . . . .	185
4	Proof of the Technical Lemmas . . . . .	190
5	Numerical Results . . . . .	198
<b>III</b>	<b>Degenerate Kolmogorov chains</b>	<b>201</b>
<b>6</b>	<b>Sharp Schauder Estimates</b>	<b>203</b>
1	Introduction and Main Results . . . . .	203
1.1	Intrinsic scales of the system and associated distance . . . . .	208
1.2	Associated Hölder spaces . . . . .	208
1.3	Assumptions and main result . . . . .	210
2	Detailed Guide to the proof . . . . .	212
3	Gaussian proxy and associated controls . . . . .	224
3.1	Controls for the frozen density . . . . .	225
3.2	Additional controls associated with the (mollified) flow and its linearization . . . . .	229
4	Control of the supremum of the derivatives in the non-degenerate variables	230
4.1	Control of the non-degenerate part of the perturbative term . . . . .	230
4.2	Control of the degenerate part of the perturbative term . . . . .	231
4.3	Non-degenerate derivatives for the frozen semi-group : terminal condition and source . . . . .	237
5	Hölder controls . . . . .	239
5.1	Hölder norms for the frozen semi-group . . . . .	241
5.2	Hölder norms associated with the Green kernel . . . . .	246
5.3	Hölder norms of the perturbative contribution . . . . .	249
5.4	Controls of the discontinuity terms arising from the change of freezing point . . . . .	253
6	Scaling issues and final proof of Theorem 1 . . . . .	257
6.1	Scaling settings and controls . . . . .	257
6.2	Scaling properties . . . . .	259
6.3	Conclusion: final proof of Theorem 1. . . . .	263
7	Appendix: Proof of technical results . . . . .	267

7.1	Technical results associated with the flow . . . . .	267
7.2	Sensitivities for the scaled flows . . . . .	274
8	Appendix: Sensitivity results for the resolvent and covariance . . . . .	274
8.1	Sensitivity Lemma for the resolvent . . . . .	274
8.2	Sensitivity Lemma for the covariances . . . . .	276
8.3	Reverse Taylor formula . . . . .	278
9	Scaling Control of the degenerate part of the perturbative term . . . . .	278
<b>7</b>	<b>Strong Uniqueness</b>	<b>285</b>
1	Introduction . . . . .	285
1.1	Statement of the problem and related results . . . . .	285
1.2	Notations, assumptions and main result . . . . .	289
1.3	Proof of the main result: Zvonkin Transform and smoothing properties of the PDE associated with (1.1) . . . . .	291
1.4	Regularization properties of the underlying PDE (1.5): strategy of proof and primer . . . . .	293
2	Perturbation techniques for the PDE : proof of Theorem 2 . . . . .	299
2.1	Gaussian proxy and associated controls . . . . .	300
2.2	Control of the sensitivities: proof of Theorem 3 . . . . .	305
3	Estimates in Besov norm . . . . .	316
3.1	Proof of Lemma 3 . . . . .	316
3.2	Proof of Lemma 4 . . . . .	324
3.3	Proof of the technical lemmas . . . . .	335
4	Proof of technical results of Lemma 6 . . . . .	337
4.1	A first sensitivity result for the flow . . . . .	337
5	Appendix: Some reminders about Besov spaces . . . . .	343
5.1	Thermic characteristic of the Besov space . . . . .	343
5.2	Equivalence of Besov norms . . . . .	343

# Liste des Symboles

## Partie II

$(\gamma_k)_{k \geq 1}$	the step sequence of the scheme . . . . .	60
$(U_k)_{k \geq 1}$	the discretized Wiener increment . . . . .	60
$(X_k)_{k \geq 0}$	the discretized process . . . . .	60
$[f]_\beta$	the usual Hölder modulus, $[f]_\beta := \sup_{x \neq x'} \frac{ f(x) - f(x') }{ x - x' ^\beta}$ . . . . .	68
$\Gamma_n$	the current time of the scheme, $\Gamma_n := \sum_{k=1}^n \gamma_k$ . . . . .	60
$\Gamma_n^{(\ell)}$	$\Gamma_n^{(\ell)} := \sum_{k=1}^n \gamma_k^\ell$ . . . . .	62
$\mathcal{A}$	the infinitesimal generator of the diffusion . . . . .	59
$\mathcal{C}^\beta$	the usual (homogeneous) Hölder space . . . . .	68
$\mathcal{C}^{k,\beta}$	the usual (homogeneous) Hölder space with $k + \beta$ as parameter . . . . .	68
$\mathcal{C}_b^{k,\beta}$	the usual (inhomogeneous) Hölder space with $k + \beta$ as parameter . . . . .	68
$\mathcal{F}_k$	the filtration associated with the scheme, $\mathcal{F}_k := \sigma((X_j)_{j \in \llbracket 0, k \rrbracket})$ . . . . .	68
$\mathcal{R}_n$	a generic sequence s.t. $\mathcal{R}_n = \mathcal{R}_n(\mathbf{A}) \rightarrow_n 1$ that may change from line to line . . . . .	125
$\  \sigma(x) \ _F$	the Fröbenius norm of $\sigma(x)$ . . . . .	71
$\  B_{p,q}^{\alpha}(\mathbb{R}^d) \ $	the Besov norm . . . . .	213
$\  \cdot \ _\infty$	the uniform norm associated with Euclidian or operator norm . . . . .	62
$\  \sigma(x) \ $	the operator norm of $\sigma(x)$ , $\  \sigma(x) \  = \sup_{z \in \mathbb{R}^r,  z  \leq 1}   \sigma(x)z  $ . . . . .	63
$  \cdot  $	the Euclidian norm . . . . .	62
$\nu$	the invariant measure of the diffusion . . . . .	59
$\nu(  \sigma^* \nabla \varphi(x)  ^2)$	the “carré du champ ” pour la diffusion Brownienne . . . . .	64
$\nu_n$	the empirical measure of the scheme, $\nu_n(\cdot) := \frac{\sum_{k=1}^n \gamma_k \delta_{X_{k-1}(\omega)}(\cdot)}{\sum_{k=1}^n \gamma_k}$ . . . . .	60
$\Sigma$	$\Sigma(x) := \sigma \sigma^*(x)$ . . . . .	63
$\sigma_\varphi^2$	the “carré du champ ” associated with the diffusion with jumps, $\sigma_\varphi^2 := \int_{\mathbb{R}^d} (   \sigma^* \nabla \varphi  ^2 (x) + \int_{\mathbb{R}^d}   \varphi(x + \kappa(x)y) - \varphi(x)  ^2 \pi(dy) ) \nu(dx)$ . . . . .	172

$\ \cdot\ $	any matrix norm dominating the operator norm.....	71
$B_{p,q}^{\tilde{\alpha}}(\mathbb{R}^d)$	the Besov space.....	213
$e_n$	a generic sequence s.t. $e_n = e_n(\mathbf{A}) \rightarrow_n 0$ that may change from line to line.....	125
$h_v(z)$	the usual heat kernel space.....	213
$V$	Lyapunov function.....	59
$(\mathcal{L}_V)$	Assumption : Lyapunov assumption.....	63
$(\mathbf{C}_R)$	Assumption : (strong) confluence and regularity, $(\mathbf{D}_\alpha^p)$ , $(\mathbf{R}_{3,\beta})$ , and $\ D\sigma\ _\infty^2 \leq \frac{2\alpha}{2(3+\beta)-p}$ .....	72
$(\mathbf{C}_{UE})$	Assumption : (mild) confluence and non-degeneracy, $(\mathbf{D}_\alpha^p)$ , $(\mathbf{R}_{1,\beta})$ , $\ D\sigma\ _\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$ , and $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$ .....	73
$(\mathbf{D}_\alpha^p)$	Assumption : confluence condition.....	72
$(\mathbf{R}_{k,\beta})$	Assumption : smoothness of the coefficients and the source.....	72
$(\mathbf{UE})$	Assumption : uniform ellipticity.....	73
$(\mathbf{A})$	Assumption : Chapitre 1 $(\mathbf{C1})$ , $(\mathbf{GC})$ , $(\mathbf{C2})$ , $(\mathcal{L}_V)$ , $(\mathbf{U})$ , $(\mathbf{S})$ .....	77
$(\mathbf{A})$	Assumption : Chapitre 2 $(\mathbf{C1})$ , $(\mathbf{GC})$ , $(\mathbf{C2})$ , $(\mathcal{L}_V)$ , $(\mathbf{U})$ , $(\mathbf{S})$ , $(\mathbf{T}_\beta)$ .....	128
$(\mathbf{A})$	Assumption : Chapitre 3 $(\mathbf{C0})$ , $(\mathbf{C1})$ , $(\mathbf{GC})$ , $(\mathbf{C2})$ , $(\mathcal{L}_V)$ , $(\mathbf{U})$ , $(\mathbf{S})$ , $(\mathbf{T}_\beta)$ .....	171
$(\mathbf{C0})$	Assumption : diffusion coefficients are Lipschitz continuous.....	169
$(\mathbf{C1})$	Assumption : $X_0$ is exponentially integrable.....	62
$(\mathbf{C2})$	Assumption : Boundness of $\sigma$ .....	63
$(\mathbf{C2})$	Assumption : boundedness of $\sigma$ and $\kappa$ .....	169
$(\mathbf{GC})$	Assumption : Gaussian Concentration.....	62
$(\mathbf{G}_V)$	Assumption : $ \varphi(x)  \leq C_{V,\varphi}(1 + \sqrt{V(x)})$ .....	69
$(\mathbf{P}_\beta)$	Assumption : $(\mathbf{UE})$ , $(\mathbf{D}_\alpha^p)$ , $\ D\sigma\ _\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$ , $(\mathbf{R}_{1,\beta})$ , and $(\Sigma)$ .....	151
$(\mathbf{S})$	Assumption : small steps are small enough.....	63
$(\mathbf{T}_\beta)$	Assumption : regularity of the test function $\varphi$ .....	127
$(\mathbf{U})$	Assumption : uniqueness of the invariant measure of the diffusion ..	63
$(\Sigma)$	Assumption : $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$ .....	151

### Partie III

$\Pi_i^x(\psi)(\mathbf{z})$	$\Pi_i^x(\psi)(\mathbf{z}) = \psi(\mathbf{z}_1, \dots, \mathbf{z}_i + x, \dots, \mathbf{z}_n)$ .....	205
$\psi_i(\mathbf{z}, x)$	$\psi_i(\mathbf{z}, x) = \Pi_i^x(\psi)(\mathbf{z})$ .....	205

$[\psi]_{\mathbf{d}}^\gamma$	$\sup_{\mathbf{x} \neq \mathbf{x}', (\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2} \frac{ \psi(\mathbf{x}) - \psi(\mathbf{x}') }{\mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')} \dots\dots\dots$	206
$\bar{p}_{C^{-1}}^\xi(t, s, \mathbf{x}, \mathbf{y})$	$\bar{p}_{C^{-1}}^\xi(t, s, \mathbf{x}, \mathbf{y}) = \frac{C}{(s-t)^{\frac{n^2 d}{2}}} \exp(-C^{-1}(s-t)  \mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^\xi(\mathbf{x}) - \mathbf{y}) ^2)$ , also noted $\bar{p}_{C^{-1}}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) \dots\dots\dots$	209
$\mathbb{T}_t$	intrinsic scale matrix. $\dots\dots\dots$	209
$\theta_{v, \tau}(\xi)$	the flow associated with the drift, $\dot{\theta}_{v, \tau}(\xi) = \mathbf{F}(v, \theta_{v, \tau}(\xi))$ , $\theta_{\tau, \tau}(\xi) = \xi$	209
$\xi$	the spatial freezing parameter. $\dots\dots\dots$	209
$\mathbf{D}$	the full spatial gradient, $\mathbf{D} = (D_{\mathbf{x}_1}, \dots, D_{\mathbf{x}_n}) \dots\dots\dots$	199
$\mathbf{d}$	the parabolic pseudo-distance, $\mathbf{d}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^d  \mathbf{y}_i - \mathbf{x}_i ^{\frac{1}{2i-1}} \dots\dots\dots$	204
$\mathbf{F}_i(t, \mathbf{x})$	the drift at the $i^{\text{th}}$ level, $\mathbf{F}_i(t, \mathbf{x}) = \mathbf{F}_i(t, \mathbf{x}^{i-1:n}) \dots\dots\dots$	199
$\mathbf{m}_{s,t}^\xi(\mathbf{x})$	the mean of the proxy process, also noted $\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) \dots\dots\dots$	209
$\mathbf{x}^{i:n}$	$x^{i-1:n} := (\mathbf{x}_{i-1}, \dots, \mathbf{x}_n) \dots\dots\dots$	199
$\mathbf{X}_t$	the diffusion associated with the Kolmogorov operator $\dots\dots\dots$	200
$ \cdot _{C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)}$	the Euclidean norm. $\dots\dots\dots$	205
$\ \psi\ _{C_{b, \mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)}$	$\ \psi\ _{C_{b, \mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)} = \sum_{i=1}^n \sup_{\mathbf{z} \in \mathbb{R}^{nd}} \sup_{\mathbf{z} \in \mathbb{R}^{nd}} \ \psi(\mathbf{z}, \cdot)\ _{C_b^{\frac{k+\gamma}{2i-1}}(\mathbb{R}^d, \mathbb{R}^\ell)} \dots\dots\dots$	206
$\ \cdot\ _{C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)}$	the usual Hölder norm, $\ \psi\ _{C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)} = \sum_{i=1}^k \sup_{ \alpha =i} \ D^\alpha \psi\ _{L^\infty(\mathbb{R}^m, \mathbb{R}^\ell)} + \sup_{ \alpha =k} [D^\alpha \psi]_\beta \dots\dots\dots$	205
$\ \cdot\ _{C_b^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)}$	the non-homogeneous Hölder norm, $\ \cdot\ _{C_b^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)} = \ \cdot\ _{C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)} + \ \cdot\ _{L^\infty(\mathbb{R}^m, \mathbb{R}^\ell)} \dots\dots\dots$	205
$\ \cdot\ _{C_{\mathbf{d}}^{k+\gamma}}$	$\ \cdot\ _{C_{\mathbf{d}}^{k+\gamma}} = \ \cdot\ _{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)} \dots\dots\dots$	205
$\ \cdot\ _{C_{b, \mathbf{d}}^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)}$	$\ \cdot\ _{C_{b, \mathbf{d}}^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)} = \sum_{i=1}^m \sup_{\mathbf{z} \in \mathbb{R}^{nd}} \ \psi_i(\mathbf{z}, \cdot)\  \dots\dots\dots$	205
$\ \cdot\ _{L^\infty}$	$\ \cdot\ _{L^\infty} = \ \cdot\ _{L^\infty(\mathbb{R}^{nd}, \mathbb{R}^\ell)} \dots\dots\dots$	205
$\tau$	the time freezing parameter. $\dots\dots\dots$	209
$\tilde{G}^{m, \xi}$	the Green kernel associated with the proxy, also noted $\tilde{G}^{m, \xi} \dots\dots\dots$	210
$\tilde{p}^\xi(t, s, \mathbf{x}, \cdot)$	the multi-scale density of the proxy, also noted $\tilde{p}^{m, (\tau, \xi)}(t, s, \mathbf{x}, \cdot) \dots\dots\dots$	209
$\tilde{P}_{T, t}^\xi$	the semi-group associated with the proxy, also noted $\tilde{P}_{T, t}^{m, \xi} \dots\dots\dots$	210
$\tilde{\mathbf{K}}_{v, t}^{(\tau, \xi)}$	the covariance matrix of the proxy. $\dots\dots\dots$	221
$\tilde{\mathbf{R}}^{(\tau, \xi)}(v, t)$	the resolvent associated with $(D\mathbf{F}(v, \theta_{v, \tau}(\xi)))_{v \in [t, s]} \dots\dots\dots$	220
$\tilde{\mathbf{X}}_v$	the linearized processus. $\dots\dots\dots$	209
$a$	the matrix diffusion. $\dots\dots\dots$	199
$C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)$	the non-homogeneous Hölder space. $\dots\dots\dots$	205

$C_b^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)$	the homogeneous Hölder space . . . . .	205
$D\mathbf{F}$	the subdiagonal of the Jacobian matrix $D\mathbf{F}$ . . . . .	209
$d$	the dimension of each block of the degenerate chain . . . . .	199
$f$	the source of the Kolmogorov equation . . . . .	199
$g$	the terminal condition of the Kolmogorov equation . . . . .	199
$L_t$	the spatial Kolmogorov operator . . . . .	200
$n$	the number of levels in the degenerate chain . . . . .	199
$T$	the time horizon . . . . .	199
$(\mathbf{T}_\beta)$	Assumption : $(\mathbf{F}_i)_{i \in \llbracket 1, (j+1) \wedge n \rrbracket}$ are uniformly $\beta_j$ -Hölder continuous in the $j^{\text{th}}$ spatial variable with $\beta_j \in (\frac{2j-2}{2j-1}, 1]$ . . . . .	284
$(\mathbf{H}_\eta)$	Assumption : $D_{\mathbf{x}_{i-1}}\mathbf{F}_i$ is non-degenerate and $\frac{\eta}{2i-3}$ Hölder continuous w.r.t. $\mathbf{x}_{i-1}$ . . . . .	284
$(\mathbf{A})$	Assumption : Chapter 6 $(\mathbf{UE}), (\mathbf{H}), (\mathbf{S})$ . . . . .	207
$(\mathbf{A})$	Assumption : Chapter 7 $(\mathbf{ML}), (\mathbf{UE}), (\mathbf{T}_\beta), (\mathbf{H}_\eta)$ . . . . .	284
$(\mathbf{H})$	Assumption : Weak Hörmander like condition . . . . .	206
$(\mathbf{ML})$	Assumption : $\mathbf{F}$ and $\sigma$ are measurable in time and $\sigma$ is uniformly Lipschitz . . . . .	284
$(\mathbf{S})$	Assumption : Smoothness of the Coefficient, $a \in L^\infty(C_{b,\mathbf{d}}^\gamma), \mathbf{F}_1 \in L^\infty(C_{\mathbf{d}}^\gamma), \mathbf{F}_i \in L^\infty(C_{b,\mathbf{d}}^{2i-3+\gamma})$ . . . . .	206
$(\mathbf{UE})$	Assumption : Uniform Ellipticity of the diffusion Coefficient . . . . .	206

# Table des figures

2.1	Dessin venant de [DM10]. . . . .	40
3.1	Unbiased Case. Plot of $a \mapsto g_{n,\theta}(a)$ , for $(\theta_k)_{k \in \llbracket 1,5 \rrbracket}$ , with $\varphi(x) = \sigma(x) = x + \varepsilon \cos(x)$ , $\varepsilon = 0.01$ . . . . .	121
3.2	Biased Case. Plot of $a \mapsto g_{n,\theta}(a)$ , for $\theta_0 = \frac{1}{3}$ , with $\varphi(x) = \sigma(x) = \cos(x)$ . . . . .	122
3.3	Plot of $a \mapsto g_{n,\theta}(a)$ , for $(\theta_k)_{k \in \llbracket 1,5 \rrbracket}$ , with $\varphi(x) = \sigma(x) = \cos(x)$ . . . . .	122
3.4	Plot of $a \mapsto g_n(a)$ with $f(x) = \frac{ x ^\beta}{1+ x ^\beta}$ , $\beta = .5$ . . . . .	124
4.1	Plot of $a \mapsto g_n(a)$ with $\varphi(x) = \sigma(x) = \cos(x)$ . . . . .	166
5.1	Plot of $a \mapsto g_n(a)$ , for $\theta = \frac{1}{3}$ , with $\varphi(x) = \sigma(x) = \cos(x)$ . . . . .	199





---

# Première partie

## Introduction



# Chapitre 1

## Estimations non-asymptotiques de mesures invariantes

**Résumé :** Dans cette première partie de thèse, nous chercherons à estimer la mesure invariante d'un processus ergodique dirigé par une Équation Différentielle Stochastique. Le Théorème ergodique nous suggère de considérer la mesure empirique associée à un schéma d'approximation du processus sous-jacent qui peut se voir comme le pendant discret de la mesure d'occupation dudit processus. Dans [LP02], Lamberton et Pagès ont introduit un algorithme de discrétisation à pas décroissant qui assure la convergence de la mesure empirique du schéma vers la mesure invariante du processus considéré ainsi qu'un théorème central limite (TCL) quantifiant asymptotiquement l'écart entre ces deux mesures. Nous établissons des inégalités de concentration non-asymptotiques pour les déviations de la mesure empirique (cohérentes avec le TCL mentionné ci-avant), ainsi que des contrôles sur la solution de l'équation de Poisson associée, utiles pour ces inégalités.

### 1 Processus Ergodiques

#### 1.1 Mesure invariante

L'étude du comportement de certaines EDS en temps long et l'approximation de leurs états stationnaires interviennent dans de nombreux champs d'applications. Mentionnons par exemple la mécanique Hamiltonienne. On peut penser en particulier à des matériaux soumis à des séismes et des intempéries dont le comportement en temps long doit être estimé, c.f. par exemple le livre de Soize [Soi94], ou encore à des particules Browniennes interagissant en vitesse/position, voir le Brownien physique de Nelson [Nel67]. Un autre champ d'application est fourni par la finance, le cadre ergodique apparaît naturellement pour la modélisation de dynamiques de taux d'intérêt ou de volatilités stochastiques, c.f. Fouque *et al.* [FPS00]

Dans une perspective d'approximations numériques concrètes, nous nous concentrerons sur des estimations ergodiques non-asymptotiques. Nous considérons dans un premier temps que la dynamique du processus sous-jacent est donnée par l'équation différentielle stochastique (EDS) suivante :

$$d\mathbf{X}_t = b(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)dW_t, \quad t > 0, \tag{1.1}$$

où  $(W_t)_{t \geq 0}$  est un mouvement Brownien de dimension  $r$  sur un espace de probabilité filtré  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ,  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  seront des fonctions Lipschitz. \*Tout au long de la Partie II de la thèse, sauf mention explicite, nous supposons que  $\sigma$  peut éventuellement dégénérer (i.e. être une matrice non-inversible). Nous nous focaliserons sur l'étude de processus  $(\mathbf{X}_t)_{t \geq 0}$  de dynamique (1.1) qui admettent une unique mesure  $\nu$  stationnaire et sont ergodiques. Dans la suite de l'introduction, nous réservons la notation en caractères gras pour des objets ayant une dynamique à temps continu. Dans le cas d'un processus stationnaire, on a en particulier pour toute fonction  $f$  suffisamment régulière et  $t \geq 0$ ,

$$\int P_t f(x) \nu(dx) = \nu(f) =: \int f(x) \nu(dx), \quad (1.2)$$

où  $P_t$  est le semi-groupe associé au processus de Markov  $\mathbf{X}_t$  défini par

$$P_t f(x) := \mathbb{E}[f(\mathbf{X}_t) | \mathbf{X}_0 = x].$$

L'identité (1.2) signifie que si le processus part avec la loi  $\nu$  alors celle-ci est conservée au cours du temps.

On dit d'un processus stationnaire de mesure invariante  $\nu$  qu'il est ergodique s'il vérifie le **théorème ergodique**. Introduisons pour tout  $t > 0$  la mesure d'occupation suivante :

$$\nu_t := \frac{1}{t} \int_0^t \delta_{\mathbf{X}_s} ds. \quad (1.3)$$

Le théorème ergodique est vérifié si pour toute fonction  $f$  continue bornée presque sûrement (p.s.) :

$$\lim_{t \rightarrow +\infty} \nu_t(f) = \lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t f(\mathbf{X}_s) ds = \nu(f) = \int f(x) \nu(dx). \quad (1.4)$$

L'exemple typique vérifiant les propriétés précédentes est fourni par le processus d'Ornstein-Uhlenbeck, associé à  $b(x) = -\frac{x}{2}$  et  $\sigma(x) = I_d$  pour la dynamique (1.1), qui s'écrit alors :

$$d\mathbf{X}_t = -\frac{1}{2} \mathbf{X}_t dt + dW_t, \quad (1.5)$$

et s'intègre explicitement en

$$\mathbf{X}_t = e^{-\frac{t}{2}} \mathbf{X}_0 + e^{-\frac{t}{2}} \int_0^t e^{\frac{s}{2}} dW_s. \quad (1.6)$$

On déduit de (1.6) que pour un  $\mathbf{X}_0 \in L^2(\mathbb{P})$  indépendant de  $(W_t)_{t \geq 0}$ ,  $\mathbf{X}_t \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, I_d)$  et que si  $\mathbf{X}_0 \stackrel{\mathcal{L}}{=} \mathcal{N}(0, I_d)$  alors pour tout  $t \geq 0$ ,  $\mathbf{X}_t \stackrel{\mathcal{L}}{=} \mathcal{N}(0, I_d)$ . Les propriétés ergodiques de l'Ornstein-Uhlenbeck se sont aisément déduites des conditions garantissant la validité du théorème ergodique données dans la section suivante.

---

\*. Nous renvoyons à la monographie [KP11] pour d'avantage d'applications faisant intervenir des EDS.

## 1.2 Théorèmes Ergodiques

Une propriété fondamentale du théorème ergodique est que les moyennes temporelles et les moyennes spatiales sont égales. La quantité aléatoire  $\nu_t(f)$  (intégrale temporelle) converge p.s. vers  $\nu(f)$  (intégrale spatiale par rapport à la mesure invariante). Cette remarquable propriété permet de se focaliser sur l'étude d'une trajectoire du processus pour estimer la mesure invariante à partir d'une seule réalisation observée en temps infini. Cet aspect se révèle très intéressant d'un point de vue numérique lorsque l'on peut observer le processus sur un temps suffisamment long.

Une question naturelle consiste à se demander quels types de critères conduisent au théorème ergodique et à l'unicité de la mesure invariante. Dans le cadre général d'un processus de Markov Fellerien (semi-groupe continu en 0 et contractant), deux grands types d'hypothèses le permettent : *l'irréductibilité* et *la confluence*.

Dans ces deux cas, on suppose l'existence d'une fonction  $V$  dite de Lyapunov assurant la non-explosion du processus ainsi qu'un retour à la moyenne. Ces propriétés s'expriment en terme du générateur infinitésimal  $\mathcal{A}$  du processus défini comme suit. Pour tout  $\varphi \in C_0^2$ ,

$$\lim_{t \rightarrow 0} \frac{P_t \varphi - \varphi}{t} =: \mathcal{A}\varphi. \quad (1.7)$$

Dans le cas d'une diffusion de type (1.1), nous avons en particulier par la formule d'Itô :

$$\mathcal{A}\varphi = \left. \frac{dP_t \varphi}{dt} \right|_{t=0} = \langle b, \nabla \varphi \rangle + \frac{1}{2} \text{Tr}(\sigma \sigma^* D^2 \varphi). \quad (1.8)$$

Introduisons le Critère d'Has'minskii (voir e.g. [KM11]) :

( $\mathcal{H}$ ) Il existe une fonction de Lyapunov  $V : \mathbb{R}^d \rightarrow ]0, +\infty[$  satisfaisant les conditions suivantes :

- i) *Coercivité.*  $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ .
- ii) *Non-explosion.* Il existe  $C_V > 0$  tel que  $\mathcal{A}V \leq C_V$ .
- iii) *Retour à la moyenne.*  $\limsup_{|x| \rightarrow \infty} \mathcal{A}V(x) < 0$ .

Présentons maintenant brièvement les deux hypothèses indiquées ci-dessus et renvoyons pour plus d'informations à Ethier et Kurtz [EK86] ainsi qu'à l'article de synthèse de Pagès [Pag01].

- *Hypothèse d'irréductibilité* : on suppose que ( $\mathcal{H}$ ) est vérifiée, et que le processus admet une densité strictement positive. Cette hypothèse assure aussi l'unicité de la mesure invariante.
- *Hypothèse de confluence* : sous des conditions de type Lyapunov *plus fortes* que ( $\mathcal{H}$ ) (voir ( $\mathcal{L}_V$ ) ci-après et également [LP03] et [PP14] pour des conditions moins fortes), le théorème ergodique est vérifié. Si de plus les coefficients de diffusion satisfont une hypothèse de confluence (on parle aussi de diffusion asymptotiquement plate, i.e. avec les notations usuelles pour les processus de Markov, en un certain sens, on a pour tous points initiaux  $x, y \in \mathbb{R}^d$ ,  $|X_t^x - X_t^y| \underset{t \rightarrow +\infty}{\rightsquigarrow} 0$ ), alors on a également l'unicité de la mesure invariante.

En particulier, le premier jeu d'hypothèses est valable dans le cadre de diffusions uniformément elliptiques vérifiant  $(\mathcal{H})$  à coefficients réguliers. L'hypothèse de positivité de la densité est clairement liée à la loi du processus et impliquée par une condition de type non-dégénérescence (aspect régularisant du semi-groupe sous-jacent). L'hypothèse de confluence n'est, quant à elle, pas *a priori* liée à des conditions de non-dégénérescence. Elle est de nature beaucoup plus trajectorielle dès que la dimension est supérieure à un  $\dagger$ .

Notons que le processus d'Ornstein-Uhlenbeck, vérifie ces deux types d'hypothèses et remarquons que  $\mathcal{A}\varphi(x) = -\langle \frac{x}{2}, \nabla\varphi(x) \rangle + \frac{1}{2}\Delta\varphi(x)$ . En effet, en prenant  $V$  une forme quadratique, i.e.  $V(x) = 1 + |x|^2$  les hypothèses de  $(\mathcal{H})$  sont vraies. On a clairement de (1.6) que la densité de ce processus  $\mathbf{X}_t^x \sim \mathcal{N}(e^{-\frac{t}{2}}x, 1 - e^{-t})$  est séparée de 0 (irréductible), et que la confluence est vérifiée :  $\mathbf{X}_t^x - \mathbf{X}_t^y = e^{-\frac{t}{2}}(x - y) \xrightarrow[t \rightarrow \infty]{p.s.} 0$ .

Il est impossible d'observer en temps infini un processus et donc d'avoir exactement la limite de la mesure d'occupation. Nous cherchons donc à l'estimer et à contrôler l'erreur d'approximation. La vitesse de convergence vers la mesure invariante est donnée par le TCL. Pour la diffusion (1.1), Bhattacharya [Bha82], sous des hypothèses de type irréductibilité (diffusions non-dégénérées), a montré que pour toute fonction  $f$  à croissance polynomiale on a le théorème central limite :

$$\sqrt{t}(\nu_t(f - \nu(f))) \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi(x)|^2 \nu(dx)\right), \quad (1.9)$$

avec  $\varphi$  solution de l'équation de Poisson

$$\mathcal{A}\varphi = f - \nu(f). \quad (1.10)$$

La force du TCL est qu'il permet de rendre compte de l'ordre typique des fluctuations de la mesure d'occupation  $\nu_t$  à l'échelle  $\sqrt{t}$ . La variance asymptotique  $\nu(|\sigma^* \nabla \varphi|^2)$  est l'intégrale de ce que l'on appelle le **carré du champ**. Pour toutes fonctions  $\varphi, \psi \in \mathcal{C}_0^2$ , on définit le *carré du champ* par  $\Gamma(\varphi, \psi) := \mathcal{A}(\varphi \cdot \psi) - \mathcal{A}\varphi \cdot \psi - \varphi \cdot \mathcal{A}\psi$ . Ce qui conduit à

$$\nu(\Gamma(\varphi, \varphi)) = \int \Gamma(\varphi, \varphi) \nu(dx) = -2 \int \mathcal{A}\varphi \cdot \varphi \nu(dx) = \int |\sigma^* \nabla \varphi|^2 \nu(dx).$$

Nous renvoyons pour plus de précisions, à Bakry *et al.* [BGL14] et Ledoux [Led99]. Nous dirons par la suite que  $f$  est un **cobord** du générateur infinitésimal s'il existe une solution régulière  $\varphi$  au problème de Poisson (1.10) (voir la Section 2.2). Cette équation est difficile à résoudre, surtout dans notre approche qui nécessite d'avoir des contrôles ponctuels. Un bon cadre est fourni par les hypothèses précédentes. Notons par ailleurs que  $\nu$  vérifie l'équation de Fokker-Planck au sens des distributions :  $\mathcal{A}^*\nu = 0$ , où  $\mathcal{A}^*$  est l'opérateur adjoint de  $\mathcal{A}$ . En fait,  $\nu \in \mathcal{I}_P$ , où  $\mathcal{I}_P$  est l'ensemble des mesures invariantes pour le processus  $\mathbf{X}_t$ , si et seulement si pour tout  $\varphi \in \mathcal{D}(\mathcal{A})$  (le domaine de définition du générateur  $\mathcal{A}$ ) on a  $\nu(\mathcal{A}\varphi) = 0$ . Si  $\mathcal{I}_P$  est non vide et compact (au sens de la convergence étroite) alors le théorème de Krein-Millman assure qu'il existe au moins une mesure extrémale  $\nu^*$  de  $\mathcal{I}_P$  (i.e. qui ne s'écrit pas comme une combinaison convexe de deux autres mesures de

---

$\dagger$ . rappelons en effet, c.f. Has'minskii [KM11] et Lemaire *et al.* [LPP15] qu'en dimension 1 toutes les diffusions ergodiques sont confluentes en un certain sens.

$\mathcal{I}_P$ ). Notons au passage que l'existence de  $\nu^*$  implique que  $(\mathbf{X}_t)_{t \geq 0}$  est ergodique. Pour plus de renseignements, nous renvoyons à [Pag01] et à [DS58].

Cette caractérisation d'une mesure invariante se comprend bien formellement. En effet, pour tout  $\varphi \in \mathcal{C}_0^2$ , d'après la définition du générateur infinitésimal dans (1.7) :

$$\nu(\mathcal{A}\varphi) = \int \mathcal{A}\varphi \nu(dx) = \int \lim_{t \rightarrow 0} \frac{P_t \varphi(x) - \varphi(x)}{t} \nu(dx) = \lim_{t \rightarrow 0} \frac{\int P_t \varphi(x) \nu(dx) - \nu(\varphi)}{t} = 0, \quad (1.11)$$

l'avant-dernière égalité se justifie par un argument d'interversion limite intégrale, et la dernière identité par la définition (1.2) d'une mesure invariante. Enfin, remarquer que l'identité  $\nu(\mathcal{A}\varphi) = 0$  se vérifie pour tout  $\varphi$  suffisamment régulière telle que l'équation (1.10) est vérifiée.

L'hypothèse de régularité Lipschitz des coefficients  $b$  et  $\sigma$  ainsi que l'hypothèse de Lyapunov *ad hoc* assure que  $\mathcal{I}_V \neq \emptyset$ . Notamment, l'*Hypothèse de confluence* ou l'*Hypothèse d'irréductibilité* précédemment introduits assurent également l'unicité de la mesure invariante. Pour plus d'informations, nous renvoyons à la monographie de Khasminskii [Kha80] (ainsi qu'à la seconde édition [KM11]), ou bien à celle de Villani [Vil09], à [Pag01] ainsi qu'à [DPG01] et à [MPW02]. Enfin la compacité de  $\mathcal{I}_P$  se montre à partir des hypothèse de Lyapunov, c.f. [LP02]

### 1.3 Discrétisation du processus

En pratique, nous ne pouvons pas simuler directement le processus continu  $(\mathbf{X}_t)_{t \geq 0}$ . Il nous faut donc passer par une méthode de discrétisation. Une approche classique est fourni par le schéma d'Euler à pas constant  $\gamma > 0$  associé à (1.1) :

$$X_{n+1}^\gamma = X_n^\gamma + \gamma b(X_n^\gamma) + \sqrt{\gamma} \sigma(X_n^\gamma) U_{n+1}, \quad (1.12)$$

où  $(U_n)_{n \geq 2}$  est une suite de variables aléatoires de  $\mathbb{R}^r$  indépendantes et identiquement distribuées (i.i.d.) avec les trois premiers moments similaires à ceux de la Gaussienne, et indépendantes de  $\mathbf{X}_0$ . Ainsi  $(X_n^\gamma)_{n \geq 1}$  est une chaîne de Markov homogène, ce qui nous permet de bénéficier d'une riche littérature (voir par exemple Meyn et Tweedie [MT09] pour l'aspect temps long et stabilisation). Une abondante littérature existe pour la convergence de schéma d'Euler à pas constant. Dans le cadre de l'erreur faible à temps fini, i.e. pour des quantités du type  $\mathcal{E} := \mathbb{E}[f(\mathbf{X}_T) - f(X_n)]$  avec  $n\gamma = T > 0$  pour une bonne «classe» de fonctions  $f$ , mentionnons le papier fondateur de Talay et Tubaro [TT90], qui les premiers établirent un contrôle de l'erreur  $\mathcal{E}$  à l'ordre  $\gamma$  sous des hypothèses de régularité sur  $b, \sigma$  et  $f$ . Indiquons également les extensions proposées par Bally et Talay dans un cas hypoelliptique [BT96a] et [BT96b] ainsi que les approches parametrix considérées par Konakov et Mammen [KM00] et [KM02].

Mentionnons par ailleurs qu'un cas ergodique elliptique est déjà considéré dans [TT90]. Des cadres dégénérés liés aux applications précédemment évoquées associées au cadre de la mécanique Hamiltonienne ou problèmes cinétiques ont aussi été abordés par Talay [Tal02] et par Mattingly *et al.* [MSH02]. Pour simplifier, les dynamiques considérées dans les articles indiqués sont de la forme suivante :

$$d\mathbf{X}_t = \begin{pmatrix} b_1(\mathbf{X}_t) \\ \mathbf{X}_t^1 \end{pmatrix} dt + \begin{pmatrix} \sigma_1(\mathbf{X}_t) \\ \mathbf{0}_{d,d} \end{pmatrix} dW_t, \quad (1.13)$$



où  $W_t$  est un mouvement Brownien de  $\mathbb{R}^d$ ,  $\mathbf{X}_t = (\mathbf{X}_t^1, \mathbf{X}_t^2)^*$  avec  $\mathbf{X}_t^1, \mathbf{X}_t^2 \in \mathbb{R}^d$ . Avec les notations de (1.1), la dynamique de (1.13) correspond à  $b(x) = (b_1(x), x_1)^*$ ,  $\sigma(x) = (\sigma_1(x), \mathbf{0}_{d,d})$ . Dans les deux cas, les auteurs obtiennent sous des hypothèses de type irréductibilité, en exhibant une fonction de Lyapunov adéquate et pour un schéma implicite chez [Tal02], des contrôles sur l'erreur faible  $\nu(f) - \nu^\gamma(f)$  en fonction du pas  $\gamma$ .

De plus, sous les conditions considérées, le théorème ergodique donne que pour  $f$  à croissance polynomiale,  $\nu_n^\gamma(f) \xrightarrow[n \rightarrow \infty]{p.s.} \nu^\gamma(f)$  la mesure invariante du schéma et non celle de la diffusion (1.1), i.e. la mesure d'occupation converge vers la mesure invariante du schéma et non celle de la diffusion. On pourrait espérer établir un théorème central limite (TCL) pour les déviations de la mesure d'occupation du schéma à pas constant mais cette approche induirait à considérer deux erreurs. Quoiqu'il en soit l'erreur globale se décompose en ce cadre en deux contributions, l'erreur statistique et l'erreur de discrétisation, i.e.

$$\nu_n^\gamma(f) - \nu(f) = \nu_n^\gamma(f) - \nu^\gamma(f) + \nu^\gamma(f) - \nu(f) =: \mathcal{E}_S + \mathcal{E}_D.$$

Notons que l'étude de l'erreur de discrétisation  $\mathcal{E}_D$  conduit à résoudre deux fois l'équation de Poisson. Celle (1.10) associée à  $\mathcal{A}$  et celle associée à  $\mathcal{A}^\gamma$  (générateur du schéma) à savoir  $\mathcal{A}^\gamma \varphi = f - \nu(f)$ , avec  $\mathcal{A}^\gamma \varphi := \gamma^{-1}(P_\gamma \varphi - \varphi)$ , où  $P_\gamma$  est le semi-groupe associé au schéma. L'étude de l'équation de Poisson dans le cadre continu (1.10) est déjà non triviale (voir Section 2.2), ajouter l'étude d'une équation de Poisson discrète rend l'analyse d'autant plus difficile. En outre, d'un point de vue algorithmique, à un pas de temps donné  $\gamma$ , on ne peut espérer atteindre la bonne mesure (sans biais).

La force de l'algorithme à pas décroissant que nous allons utiliser est de supprimer le terme d'erreur de discrétisation  $\mathcal{E}_S$  (la moyenne ergodique du schéma convergera vers la mesure cible  $\nu$ ). Il est également suffisamment robuste pour prendre en compte des dynamiques dégénérées comme celles considérées dans l'équation (1.13) dès lors que le problème associé à la diffusion est, en un certain sens, bien posé.

Les premiers travaux en ce sens remontent à Basak *et al.* [BHW97]. L'algorithme à pas décroissant considéré s'écrit pour  $n \geq 0$  :

$$X_{n+1} = X_n + \gamma_{n+1} b(X_n) + \sqrt{\gamma_{n+1}} \sigma(X_n) U_{n+1}, X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}), \quad (1.14)$$

où  $(\gamma_k)_{k \geq 1}$  désigne une suite de pas décroissants. Leur résultat principal a été d'établir que la loi de  $X_n$  converge faiblement (au sens des mesures) vers la mesure invariante  $\nu$  de la diffusion. Le pas décroissant permet d'une certaine manière d'être de plus en plus précis quand le temps devient grand. Le pas tend vers 0 au fur et à mesure que le temps augmente et, par rapport à un schéma à pas constants comme en (1.12), dans le cadre d'une simulation ergodique, on fait d'un coup tendre  $n$  vers l'infini et  $\gamma_n$  vers 0.

Mentionnons également [PS94], qui ont montré un théorème ergodique, et une convergence  $L^2$  associée, pour la mesure d'occupation empirique d'un schéma à pas décroissant de type (1.14) approchant un processus de diffusion à valeurs dans des groupes de Lie compacts. Lambertson et Pagès [LP02] ont ensuite poussé beaucoup plus loin l'étude de ce type de schémas en établissant dans le cadre de diffusions ergodiques Browniennes générales un théorème ergodique pour la mesure d'occupation empirique du schéma (1.14) ainsi qu'un théorème central limite (TCL) pour les déviations associées.

Pour  $n \geq 0$  fixé, la mesure d'occupation empirique se définit précisément de la façon suivante : pour tout  $A \in \mathcal{B}(\mathbb{R}^d)$  (sous-ensemble Borélien de  $\mathbb{R}^d$ ),

$$\nu_n(A) := \nu_n(\omega, A) := \frac{\sum_{k=1}^n \gamma_k \delta_{X_{k-1}(\omega)}(A)}{\sum_{k=1}^n \gamma_k}. \quad (1.15)$$

Nous noterons par la suite  $\Gamma_n := \sum_{k=1}^n \gamma_k$  qui représente le temps courant associé à (1.14) à l'étape  $n$ . Indiquons que  $\Gamma_n$  est l'analogie discret du  $t$  dans la mesure d'occupation  $\nu_t$  associée à  $(\mathbf{X}_s)_{s \in [0, t]}$  introduite en (1.3). Dans une perspective d'étude en temps long, la suite des pas de temps  $(\gamma_k)_{k \geq 1}$  sera choisie telle que  $\Gamma_n \xrightarrow[n]{n} +\infty$ .

Les hypothèses de Lyapunov *renforcées* (par rapport au critère  $(\mathcal{H})$ ) considérées par Lamberton et Pagès sont les suivantes :

$(\mathcal{L}_V)$  Il existe une fonction de Lyapunov  $V : \mathbb{R}^d \rightarrow [v^*, +\infty[$ , avec  $v^* > 0$  satisfaisant les conditions suivantes :

- i) Régularité-Coercivité.  $V$  est une fonction de classe  $\mathcal{C}^2$ ,  $\|D^2V\|_\infty < +\infty$ ,  
 $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ .
- ii) Contrôle de la croissance. Il existe  $C_V \in (0, +\infty)$  tel que pour tout  $x \in \mathbb{R}^d$  :

$$|\nabla V(x)|^2 + |b(x)|^2 \leq C_V V(x).$$

- iii) Stabilité. Il existe  $\alpha_V > 0$ ,  $\beta_V \in \mathbb{R}^+$  tel que pour tout  $x \in \mathbb{R}^d$ ,

$$\mathcal{A}V(x) \leq -\alpha_V V(x) + \beta_V.$$

Ces hypothèses signifient que  $V$  est sous-quadratique et  $b$  sous-linéaire. Un exemple typique de processus vérifiant ces hypothèses est de nouveau l'Ornstein-Uhlenbeck. Il faut voir la condition iii) comme une contrainte sur la dérive qui doit se comporter comme une force de rappel sous-linéaire afin de faire revenir le processus à la moyenne.

Sous  $(\mathcal{L}_V)$ , Lamberton et Pagès montrent d'abord le théorème ergodique associé à l'algorithme (1.14), i.e. pour toute fonction  $\nu - p.s.$  continue  $f$  à croissance polynomiale,

$$\nu_n(f) \xrightarrow[n]{p.s.} \nu(f) = \int_{\mathbb{R}^d} f(x) \nu(dx), \quad (1.16)$$

qui est le pendant discrétisé du Théorème ergodique (1.4). De ce fait, le principal intérêt de ce schéma vis-à-vis de celui à pas constant est qu'il fait converger la mesure empirique directement vers la mesure invariante de la diffusion. La démonstration de ce théorème ergodique repose sur le caractère tendu de la suite  $(\nu_n(V))_{n \geq 1}$  où  $V$  est une fonction de Lyapunov et sur le théorème d'Echeverria Weiss (identification de la limite en terme de problème de martingale, c.f. [EK86]).

Une fois le résultat ergodique (1.16) à disposition, la question naturelle est de nouveau d'établir un TCL, en regard du résultat de Bhattacharya rappelé en (1.9) dans le cas continu. Notons  $\Gamma_n^{(\ell)} := \sum_{k=1}^n \gamma_k^\ell$ ,  $\ell > 0$ , de sorte qu'en particulier  $\Gamma_n^{(1)} = \Gamma_n$ . En se concentrant sur des fonctions  $f$  de la forme  $f - \nu(f) = \mathcal{A}\varphi$  où  $\varphi$  est une fonction régulière, et pour lesquelles, d'après (1.11),  $\nu(\mathcal{A}\varphi) = 0$  et  $\nu_n(f) - \nu(f) = \nu_n(f - \nu(f)) = \nu_n(\mathcal{A}\varphi)$ , Lamberton et Pagès ont montré le TCL suivant.

**Theorem 1 (TCL de [LP02])** *Sous les conditions de Lyapunov ( $\mathcal{L}_V$ ), s'il existe une unique mesure invariante à la diffusion (1.1), et si  $\lim_n \Gamma_n^{(2)}/\sqrt{\Gamma_n} = 0$ , alors pour toute fonction Lipschitz  $\varphi \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$  telle que  $D^2\varphi$  soit bornée,  $D^3\varphi$  soit bornée et Lipschitzienne, et  $\sup_{x \in \mathbb{R}^d} \frac{|\sigma^* \nabla \varphi(x)|^2}{V(x)} < +\infty$  on a :*

$$\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right).$$

L'algorithme à pas décroissant permet effectivement d'avoir le pendant discret du TCL pour la diffusion énoncé en (1.9). La condition  $\lim_n \Gamma_n^{(2)}/\sqrt{\Gamma_n} = 0$  est une condition sur les pas de temps. Précisément, pour une suite de pas  $(\gamma_k)_{k \geq 1}$  telle que  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in (0, 1]$ , où l'on note  $\gamma_k \asymp k^{-\theta}$  pour indiquer qu'il existe une constante  $C \geq 1$  telle que pour tout  $k \geq 0$ ,  $C^{-1}k^{-\theta} \leq \gamma_k \leq Ck^{-\theta}$ , le critère  $\lim_n \Gamma_n^{(2)}/\sqrt{\Gamma_n} = 0$  est vérifié pour  $\theta \in (\frac{1}{3}, 1]$ . Le TCL précédent est remarquable au sens où, pour la plage de pas considérée, la discrétisation par le schéma (1.14) n'est pas visible asymptotiquement. En effet, la loi limite est la même que dans (1.9) et la variance asymptotique correspond bien au *carré du champ* de la mesure invariante.

Indiquons également que plus le pas est grand plus la convergence sera rapide. Le TCL précédent peut s'étendre au cas  $\theta = \frac{1}{3}$  (meilleure vitesse) au prix de l'apparition d'un biais. L'effet de la discrétisation se fait alors sentir. Le seuil  $\theta = \frac{1}{3}$  est doublement critique au sens où si  $\theta < \frac{1}{3}$  le TCL n'est plus valable. Il est en quelque sorte *caché* par l'erreur de discrétisation. Seule subsiste une convergence en probabilité après renormalisation. Nous renvoyons au Chapitre 3 pour de plus amples discussions sur les biais. Indiquons toutefois que celui-ci n'est pas aisé à approcher numériquement dans la mesure où il fait intervenir les dérivées de la fonction  $\varphi$  considérée jusqu'à l'ordre 4 ainsi que des intégrales par rapport à la mesure invariante et la loi de l'innovation.

La démonstration du Théorème 1 repose sur le TCL pour des accroissements de martingales de carré intégrable (voir e.g. [HH80] et [Duf90]). Il faut ainsi décomposer  $\nu_n(\mathcal{A}\varphi) = \text{Martingale} + \text{Reste}$  (comme nous le ferons dans l'équation (2.17) ci-après), puis vérifier que le terme *Martingale* satisfasse à la condition de Lindeberg (voir e.g. [HH80]).

En pratique, il est important de connaître les intervalles de confiance associés à l'estimation de la mesure invariante. Or le TCL n'est pas l'outil le plus adapté. *A priori* notre estimation renormalisée  $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$  **ne suit pas** une loi normale. Le TCL fournit un résultat **asymptotique** et ne permet pas, à  $n \geq 0$  donné de quantifier l'écart à la loi normale limite.

Le propos de cette première partie de thèse sera d'établir, sous des hypothèses similaires à celles du TCL précédent, une estimation de l'intervalle de confiance **non-asymptotique**. Nous le ferons grâce à des inégalités de concentration.

## 1.4 Inégalités de concentration

Le but de la première partie de cette thèse est d'établir une inégalité de concentration (sous)-Gaussienne non-asymptotique pour la mesure empirique renormalisée  $\sqrt{\Gamma_n} \nu_n$  associée au schéma à pas décroissants. Avec des pas constants, Malrieu et Talay dans

[MT06] ont montré une inégalité de concentration pour la mesure empirique en établissant une inégalité de type Log-Sobolev sur le schéma. On rappelle que l'inégalité de Log-Sobolev implique la concentration Gaussienne (voir par exemple l'argument de Herbst dans [BGL14]). Cette méthode est néanmoins très rigide pour un processus discrétisé.

L'approche développée dans cette thèse, privilégie l'utilisation directe de la propriété de concentration Gaussienne du terme d'innovation  $(U_k)_{k \geq 1}$  dans (1.14) qui se transmettra au schéma. Dorénavant, nous supposons qu'une variable aléatoire  $U$  de même loi que les  $(U_k)_{k \geq 1}$  dans (1.14) vérifie la propriété de concentration Gaussienne suivante :

**(GC)** La variable aléatoire  $U$  admet les mêmes trois premiers moments que la loi normale  $\mathcal{N}(0, I_r)$ , et vérifie pour toute fonction  $g : \mathbb{R}^r \rightarrow \mathbb{R}$ , 1-Lipschitz et pour tout  $\lambda > 0$  :

$$\mathbb{E}[\exp(\lambda g(U))] \leq \exp\left(\lambda \mathbb{E}[g(U)] + \frac{\lambda^2}{2}\right).$$

Mentionnons que la propriété **(GC)** implique directement, via l'inégalité de Bienaymé – Chebyshev – Markov exponentielle et une optimisation en  $\lambda$ , que si l'on note  $\mu$  la loi de  $U$ , alors, pour toute fonction  $g : \mathbb{R}^r \rightarrow \mathbb{R}$ , 1-Lipschitz et pour tout  $a > 0$  :

$$\mu(g - \mu(g) \geq a) \leq \exp(-\frac{a^2}{2}).$$

La propriété **(GC)** est entre autres vérifiée par la loi normale et la loi de Rademacher.

## 1.5 De $\nu_n(\mathcal{A}\varphi)$ à l'estimation de $\nu(f)$ pour une source $f$ donnée

Nous avons vu que le TCL du Théorème 1 s'énonçait pour des fonctions de la forme  $f - \nu(f) = \mathcal{A}\varphi$  pour  $\varphi$  suffisamment régulière. Dans la pratique, la problématique est inverse, on a une source  $f$  à disposition dont on souhaite estimer la moyenne  $\nu(f)$ . Ceci pose donc, à source  $f$  donnée, la question de trouver  $\varphi$  suffisamment régulière telle que  $\mathcal{A}\varphi = f - \nu(f)$ . En d'autres termes, il faut résoudre l'équation de Poisson pour la source  $f$  et obtenir une régularité suffisante pour la solution. De façon similaire à ce qui a été évoqué pour le théorème ergodique, deux grands jeux d'hypothèses (la non-dégénérescence et la confluence) permettront d'aboutir au résultat attendu. Nous renvoyons à la Section 2.2 ci-après et au Chapitre 3 pour une présentation plus détaillée. Jusqu'à cette section nous énoncerons nos résultats, de façon similaire au Théorème 1, pour les déviations de  $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$ .

## 2 Premier résultat de concentration

Le premier résultat du Chapitre 3 est le résultat de concentration suivant qui peut se voir comme un premier pendant non-asymptotique du Théorème 1.

**Theorem 2** *Supposons qu'il existe une unique mesure invariante associée à (1.1), que  $\sigma$  est bornée, que  $X_0$  et  $(U_k)_{k \in \llbracket 1, n \rrbracket}$  vérifient **(GC)**. Si  $(\mathcal{L}_\mathbf{v})$  est vérifiée alors pour toute fonction  $\varphi$  de classe  $\mathcal{C}^3$ , telle que  $\nabla\varphi, D^2\varphi, D^3\varphi$  sont bornés et que  $D^3\varphi$  est  $\beta$ -Hölder,*

$\beta \in (0, 1)$  vérifiant pour une certaine constante  $C_{V,\varphi} > 0$  ( $|\varphi| \leq C_{V,\varphi}(1 + \sqrt{V})$ ), et que  $x \mapsto \langle b, \nabla\varphi \rangle(x)$  est Lipschitz si  $\Gamma_n^{(\frac{3+\beta}{2})}/\sqrt{\Gamma_n} \rightarrow_n 0$ , alors il existe  $c_n \leq 1 \leq C_n$ ,  $n \geq 1$ , avec  $\lim_n C_n = \lim_n c_n = 1$ , telles que pour tout  $n \geq 1$  et  $a > 0$  (qui peut dépendre de  $n$ ) :

$$\mathbb{P}\left[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a\right] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right),$$

où  $C_n = \exp\left(\frac{[D^3\varphi]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)} \frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} + p_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}\right)$  avec  $p_n$  une suite  $\geq 1$  telle que  $p_n \rightarrow_n +\infty$  et  $p_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow_n 0$ .

Remarquons que la variance de la borne supérieure de la déviation est une majoration du carré du champ. En effet,  $\nu(|\sigma^*\nabla\varphi|^2) = \int |\sigma^*\nabla\varphi|^2(x)\nu(dx) \leq \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2$ , avec  $\|\sigma\|_\infty := \sup_{x \in \mathbb{R}^d} \|\sigma(x)\|$  où  $\|\cdot\|$  désigne une norme matricielle. On peut considérer typiquement la norme d'opérateur, i.e.  $\|\sigma(x)\|_O := \sup_{\xi \in \mathbb{R}^d} \frac{|\sigma(x)\xi|}{|\xi|}$ , en notant par  $|\cdot|$  la norme Euclidienne. Celle-ci est indépendante de la dimension mais potentiellement délicate à estimer. Un autre choix possible est la norme de Fröbenius,  $\|\sigma(x)\|_F = \sum_{1 \leq i \leq d, 1 \leq j \leq r} \sigma_{ij}^2(x)$ , qui est facile à calculer mais dépend de la dimension. De façon similaire, on note  $\|\nabla\varphi\|_\infty := \sup_{x \in \mathbb{R}^d} |\nabla\varphi(x)|$ .

Ainsi, la majoration du Théorème 2 n'est pas optimale vis-à-vis de son pendant asymptotique : le Théorème 1. Nous fournirons plus loin une inégalité de concentration non-asymptotique similaire mais avec le carré du champ dans la borne de droite (voir la Section 3). Contrairement au Théorème 1, le cadre sans biais est pour  $\Gamma_n^{(\frac{3+\beta}{2})}/\sqrt{\Gamma_n} \rightarrow_n 0$  et non  $\Gamma_n^{(2)}/\sqrt{\Gamma_n} \rightarrow_n 0$ . Pour le choix naturel de pas de temps  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in (0, 1]$ , ceci signifie que le théorème précédent est valable pour  $\theta \in (\frac{1}{2+\beta}, 1]$ . Il y a donc bien continuité avec le Théorème 1 pour  $\beta = 1$ .

Nous avons réussi à demander moins de régularité à  $D^3\varphi$  dans le Théorème 2 (où  $D^3\varphi$  est  $\beta$ -Hölder) par rapport au Théorème 1 (où  $D^3\varphi$  est Lipschitz). La moindre régularité de  $\varphi$  se paye en terme de vitesse de convergence. Mentionnons que l'affaiblissement de la régularité est motivé par la possible prise en considération de sources  $f$  Lipschitz pour notre estimation ergodique (cadre usuel des inégalités fonctionnelles). Ce cadre est détaillé en Section 2.4.

Notons que si  $\varphi$  est solution de l'équation de Poisson (1.10), sous les hypothèses du Théorème 2,  $\nabla\varphi$  et  $D^2\varphi$  sont Lipschitz. Ainsi, pour  $\sigma$  borné et Lipschitz, dès lors que  $f$  est aussi Lipschitz, on a clairement de la relation  $\langle b, \nabla\varphi \rangle = f - \frac{1}{2}\text{Tr}(\sigma\sigma^*D^2\varphi)$ , que  $x \mapsto \langle b(x), \nabla\varphi(x) \rangle$  est également Lipschitz.

L'hypothèse  $|\varphi| \leq C_{V,\varphi}(1 + \sqrt{V})$  est essentiellement technique et utilisée pour contrôler des termes de restes dans la décomposition (2.17) ci-dessous (voir également Chapitre 3). Dans le cas particulier où  $V(x) \asymp 1 + |x|^2$ , elle est automatiquement vérifiée de par le caractère uniformément Lipschitzien de  $\varphi$ .

Enfin, comme corollaire important du Théorème 2, lorsque  $\varphi$  est solution de  $\mathcal{A}\varphi = f - \nu(f)$ , nous déduisons comme annoncé précédemment, l'estimation non-asymptotique suivante pour l'intervalle de confiance associé à la simulation ergodique :

$$\mathbb{P}\left[\nu(f) \in \left[\nu_n(f) - \frac{a}{\sqrt{\Gamma_n}}, \nu_n(f) + \frac{a}{\sqrt{\Gamma_n}}\right]\right] \geq 1 - 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right).$$

Soulignons que  $f$  n'est pas forcément borné, et qu'à aucun moment de la thèse la source  $f$  ne sera supposée bornée, elle sera toutefois au moins Lipschitzienne.

**Idée de la preuve.** La démonstration s'appuie sur une méthode de martingale (dite approche d'Azuma, voir [Azu67]), qui est déjà employée pour prouver la convergence de l'algorithme et le TCL de [LP02]. Cette approche a également été utilisée dans l'article [FM12] afin de contrôler la concentration des déviations du schéma d'Euler à horizon de temps fini dans l'algorithme de Monte-Carlo. Nous nous intéresserons aux déviations de  $\nu_n(\mathcal{A}\varphi)$ , l'étude de  $|\nu_n(\mathcal{A}\varphi)|$  se déduit par symétrie. Pour toute fonction  $\varphi$  de classe  $C^2$ , on effectue un développement de Taylor entre  $\varphi(X_k)$  et  $\varphi(X_{k-1})$  pour faire apparaître le générateur défini en (1.8). Ce qui nous donne schématiquement :

$$\varphi(X_k) - \varphi(X_{k-1}) = \gamma_k \mathcal{A}\varphi(X_{k-1}) + \Delta_k(X_{k-1}, U_k) + R_{k,k-1}, \quad (2.17)$$

où  $R_{k,k-1}$  est un terme de reste et  $\Delta_k(X_{k-1}, U_k)$  est un accroissement de martingale. L'écriture explicite de  $\Delta_k(X_{k-1}, U_k)$  fait apparaître un terme quadratique en  $U_k$ , mais il se trouve que  $u \mapsto \Delta_k(X_{k-1}, u)$  est Lipschitz. En effet, dans l'identité (2.17), à droite de l'égalité seul  $\psi_k(X_{k-1}, U_k)$  dépend de  $U_k$  alors qu'à gauche de l'égalité  $\varphi(X_k)$  est bien Lipschitz en  $U_k$ . En particulier, on a conditionnellement à  $X_{k-1}$  :

$$[\Delta_k(X_{k-1}, \cdot)]_1 \leq \sqrt{\gamma_k} \|\sigma(X_{k-1})\| \|\nabla\varphi\|_\infty \leq \sqrt{\gamma_k} \|\sigma\|_\infty \|\nabla\varphi\|_\infty. \quad (2.18)$$

Notons qu'à ce stade nous choisissons de majorer le module Lipschitz de  $\Delta_k(X_{k-1}, \cdot)$  par la dernière inégalité afin d'avoir une borne déterministe, ce qui simplifie l'analyse. Toutefois, nous verrons dans la Section 2.1 que ce contrôle de module de Lipschitz est crucial pour avoir la meilleure constante de concentration possible. On somme le développement de Taylor (2.17) pour aboutir à :

$$\Gamma_n \nu_n(\mathcal{A}\varphi) = -M_n - R_n, \quad (2.19)$$

avec  $R_n = \varphi(X_0) - \varphi(X_n) + \sum_{k=1}^n R_{k,k-1}$  comme terme de reste et  $M_n = \sum_{k=1}^n \Delta_k(X_{k-1}, U_k)$  comme terme martingale qui induit la concentration sous-Gaussienne. Le terme de reste  $R_n$  se traite de façon assez technique avec des arguments usuels (inégalité de Hölder, Jensen) en exploitant l'intégrabilité exponentielle de la fonction de Lyapunov, i.e. sous les hypothèses du Théorème 2, il existe une constante  $c_V := c_V(\mathbf{A}) > 0$  telle que :  $I_V^1 := \sup_{n \geq 0} \mathbb{E}[\exp(c_V V(X_n))] < +\infty$ .

Pour le terme martingale, l'idée est d'utiliser, en conditionnant itérativement, la propriété **(GC)** de concentration Gaussienne de l'innovation, couplée au contrôle du module de Lipschitz (2.18). Précisément, l'inégalité de Bienaymé – Chebyshev – Markov exponentielle donne pour tout  $a > 0$ ,  $\lambda > 0$  :

$$\mathbb{P}\left[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a\right] \leq \exp\left(-\frac{\lambda a}{\sqrt{\Gamma_n}}\right) \mathbb{E}[\exp(\lambda \nu_n(\mathcal{A}\varphi))].$$

On isole ensuite par l'inégalité d'Hölder le terme martingale et le terme de restes. Pour  $p, q > 1$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ , on obtient :

$$\mathbb{P}\left[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a\right] \leq \exp\left(-\frac{\lambda a}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(-\frac{\lambda q M_n}{\Gamma_n}\right)\right]^{1/q} \mathbb{E}\left[\exp\left(\frac{\lambda p R_n}{\Gamma_n}\right)\right]^{1/p}.$$

On peut choisir  $p = p(n) \rightarrow_n +\infty$  tel que  $\mathbb{E}[\exp(\frac{\lambda p R_n}{\Gamma_n})]^{1/p} = \mathcal{R}_n \rightarrow_n 1$ . Dans l'inégalité précédente, en écrivant  $\mathbb{E}[e^{-\frac{\lambda q M_n}{\Gamma_n}}] = \mathbb{E}[e^{-\frac{\lambda q M_{n-1}}{\Gamma_n}} \mathbb{E}[e^{\frac{\lambda q \Delta_n (X_{n-1}, U_n)}{\Gamma_n}} | \mathcal{F}_{n-1}]]$  et en utilisant **(GC)**, il vient :

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] &\leq \mathcal{R}_n \exp(-\frac{\lambda a}{\sqrt{\Gamma_n}}) \mathbb{E}[\exp(-\frac{\lambda q M_n}{\Gamma_n})]^{1/q} \\ \stackrel{\text{(GC)}}{\leq} &\mathcal{R}_n \exp(-\frac{\lambda a}{\sqrt{\Gamma_n}} + \frac{\lambda^2 q \gamma_n \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}{2\Gamma_n^2}) \mathbb{E}[\exp(-\frac{\lambda q M_{n-1}}{\Gamma_n})]^{1/q}. \end{aligned}$$

Après itération, on obtient facilement que :

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq \exp(-\frac{\lambda a}{\sqrt{\Gamma_n}} + \frac{q \lambda^2 \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}{2\Gamma_n}). \quad (2.20)$$

Une optimisation en  $\lambda$  conduit au résultat.

## 2.1 Améliorations du Théorème 2

Pour l'instant, le Théorème 2 nous donne une majoration dudit *carré du champ* :  $\nu(|\sigma^* \nabla \varphi|^2) \leq \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2$ . Par la suite, nous réussissons à améliorer la constante de concentration du Théorème 2 et à s'approcher du *carré du champ*.

**Theorem 3** *Sous les mêmes hypothèses et notations que le Théorème 2, s'il existe  $\vartheta$  de même régularité que  $\varphi$  tel que  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$ , pour tout  $n \geq 1$ ,  $0 < a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$  où  $\chi_n > 0$  tel que  $\chi_n \rightarrow_n 0$  (arbitrairement lentement)*

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 \tilde{C}_n \exp\left(-\tilde{C}_n \frac{a^2}{2\nu(\|\sigma\|^2) \|\nabla\varphi\|_\infty^2}\right).$$

Nous voyons toujours l'importance d'étudier le problème de Poisson afin de vérifier que  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  est cobord. Indiquons que pour cette étape il est appréciable de bien connaître  $\|\cdot\|$ , i.e. calculable, en pratique on choisit la norme de Fröbenius. Dans le cas général où  $\|\cdot\|$  est une norme d'opérateur, l'étude de l'équation de Poisson  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$  est plus complexe. On a besoin de savoir que  $\|\sigma\|$  est suffisamment régulière. En outre, rappelons que  $\sqrt{\Gamma_n}/\Gamma_n^{(2)} \rightarrow_n +\infty$  pour tout  $\theta \in (\frac{1}{2+\beta}, 1]$ , la condition sur  $a$  est valable pour les intervalles de confiance à  $a$  fixé. Ce théorème représente un véritable intérêt d'un point de vue numérique notamment lorsque  $\nu(\|\sigma\|^2)$  est bien plus petit que  $\|\sigma\|_\infty^2$  (voir la partie numérique du Chapitre 3).

**Idée de la preuve** La principale idée derrière ce théorème est de constater que l'on peut estimer plus précisément la norme Lipschitz que cela n'a été fait dans l'inégalité (2.18), i.e. on écrit  $[\Delta_k(X_{k-1}, \cdot)]_1 \leq \gamma_k \|\nabla\varphi\|_\infty \|\sigma(X_{k-1})\|$ . Par souci de simplicité, nous écrirons  $\|\cdot\|$  pour la norme matricielle  $\|\cdot\|$ . Ainsi, l'idée est de soustraire la somme des carrés de la norme Lipschitz précédente afin de faire une inégalité de Hölder globale avec

$\rho > 1$  :

$$\begin{aligned} \mathbb{E}[\exp(\frac{\lambda M_n}{\Gamma_n})] &\leq \mathbb{E}\left[\exp\left(-\frac{\rho\lambda M_n}{\Gamma_n} - \frac{\lambda^2\rho^2}{2\Gamma_n^2} \sum_{k=1}^n \gamma_k \|\nabla\varphi\|_\infty^2 \|\sigma(X_{k-1})\|^2\right)\right]^{1/\rho} \\ &\quad \times \mathbb{E}\left[\exp\left(\frac{\lambda^2\rho^2}{2(\rho-1)\Gamma_n^2} \sum_{k=1}^n \gamma_k \|\nabla\varphi\|_\infty^2 \|\sigma(X_{k-1})\|^2\right)\right]^{1-\frac{1}{\rho}} \end{aligned} \quad (2.21)$$

Le premier terme de l'inégalité s'apparente à une martingale exponentielle. En fait, par des inégalités de concentration itérées, on voit que ce terme est une sur-martingale majorée par 1. Tandis que pour le second terme, on a par hypothèse que  $\|\nabla\varphi\|_\infty^2 = \nu(\|\sigma\|^2) + \mathcal{A}\vartheta$  on réitère ensuite l'analyse du Théorème 2 pour le terme  $\mathcal{A}\vartheta$ . En revanche, le coefficient devant  $\nu_n(\mathcal{A}\vartheta)$  est en  $\lambda^2$ . On obtient *in fine* une majoration de la déviation qui est en exponentielle d'un polynôme d'ordre 4 en  $\lambda$  :

$$\mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a] \leq C_n \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}} + c_n\left(\frac{\lambda^2}{\Gamma_n} \frac{\rho[\varphi]_1^2 \nu(\|\sigma\|^2)}{2} + \frac{\lambda^4 \rho \|\nabla\varphi\|_\infty^4 \|\sigma\|_\infty^2 \|\nabla\vartheta\|_\infty^2}{8\Gamma_n^3 (\rho-1)^3}\right)\right),$$

avec  $c_n, C_n > 0$  tels que  $\lim_n C_n = \lim_n c_n = 1$ . Pour optimiser en  $\lambda$ , il nous faut donc trouver les racines d'un polynôme de degré 3. On utilise à cet effet la formule de Cardan-Tartaglia selon laquelle il existe une unique solution  $\lambda_\rho > 0$  dépendante de  $\rho$ . Il faut donc encore optimiser suivant  $\rho$ . Une courte étude de cette optimisation au Chapitre 3 montre qu'il y a plusieurs régimes possibles. En effet, si  $a/\sqrt{\Gamma_n} \rightarrow_n 0$  alors on obtient le résultat, dans le cas contraire on obtient un autre régime strictement sous-Gaussien :

$$\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^{4/3} \Gamma_n^{1/3}}{2\|\sigma\|_\infty^{2/3} [\vartheta]_1^{2/3}}\right). \quad (2.22)$$

Enfin, cette distinction de régimes permet d'obtenir des résultats numériques intéressants. Même si théoriquement dans le cas où  $a \gg \sqrt{\Gamma_n}/\Gamma_n^{(2)}$  le Théorème 2 est plus fort (une vraie borne Gaussienne), en pratique ce régime n'est jamais atteint. On voit *in fine* que le Théorème 3 avec la déviation (2.22) rend numériquement mieux compte de la transition du régime  $a \ll \sqrt{\Gamma_n}/\Gamma_n^{(2)}$  à celui  $a \asymp \sqrt{\Gamma_n}/\Gamma_n^{(2)}$ . Pour plus de détails voir le Chapitre 4. Dans la prochaine section, nous verrons les grandes lignes de notre étude sur l'équation de Poisson. Ainsi, grâce au contrôle des dérivées de  $\varphi$ , solution de ladite équation de Poisson, on pourra s'intéresser directement à la déviation de  $\nu_n(f - \nu(f))$  pour tout type de fonction  $f$  de classe  $\mathcal{C}^{1,\beta}$  (ou plus régulière).

## 2.2 Sur le problème de Poisson

Formellement, la solution au problème de Poisson s'écrit par la formule de Feynman-Kac :

$$\varphi(x) = \int_0^{+\infty} \mathbb{E}[f(\mathbf{X}_t^{0,x}) - \nu(f)] dt. \quad (2.23)$$

Nous cherchons à établir un contrôle ponctuel de la solution ainsi que de ses dérivées. Néanmoins, la littérature à disposition se concentre surtout sur des contrôle de type Sobolev, voir par exemple Pardoux et Veretennikov [PV01] pour un cadre elliptique. Dans un



second papier, les auteurs [PV03] établissent également un contrôle ponctuel du gradient mais à dérive bornée et à coefficient de diffusions réguliers. Avec des conditions locales de Doebelin, [PV05] montrent l'unicité au problème de martingale associé à l'équation de Poisson (possiblement dégénérée). Mentionnons par ailleurs, Pagès et Panloup [PP14], qui contrôlent, sous hypothèses de confluence, le cas d'un processus d'Ornstein-Uhlenbeck dégénéré. Une condition classique est l'hypothèse de Bakry-Emery. Une propriété importante de cette condition est qu'elle implique une relation de commutation entre le semi-groupe et le carré du champ associés à la diffusion, pour plus de détails voir e.g. [BE85], [BGL14], [BG10]. Cette hypothèse du type courbure-dimension (voir e.g. [BB06]) permet d'avoir une borne du gradient. Néanmoins, sous ce jeu d'hypothèses il semble compliqué d'avoir des contrôles plus fins, en premier lieu le module Hölder du gradient. En outre, cette condition est difficile à vérifier en pratique, même pour le cas scalaire  $d = 1$ . Nous utiliserons par la suite une hypothèse de type confluence :

$(\mathbf{D}_p^\alpha)$  Supposons qu'il existe  $\alpha > 0$  et  $p \in (1, 2]$  tel que pour tout  $x \in \mathbb{R}^d$ ,  $\xi \in \mathbb{R}^d$

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r \left( (p-2) \frac{|\langle D\sigma_{\cdot j}(x) \xi, \xi \rangle|^2}{|\xi|^2} + |D\sigma_{\cdot j} \xi|^2 \right) \leq -\alpha |\xi|^2,$$

où  $Db$  est à la Jacobienne de  $b$  et  $D\sigma_{\cdot j}$  est la  $j$ -ième colonne de  $D\sigma$ . Autrement dit, cette condition signifie que la dérive  $b$  permet un rappel «suffisamment fort». Nous proposons deux jeux d'hypothèses afin de contrôler la solution au problème de Poisson. Soit on suppose que les coefficients sont assez réguliers pour dériver le flot à la Kunita (voir e.g. [Kun97]), soit  $\sigma$  est non-dégénéré alors on peut bénéficier du bootstrap elliptique (e.g. [KP10], dans le Chapitre 6 nous étudierons le bootstrap dans un cadre dégénéré proche des hypothèses d'Hörmander).

$(\mathbf{C}_R)$  *Confluence et Régularité* : on suppose que  $(\mathbf{D}_p^\alpha)$  est vérifiée, que  $f$  et  $\sigma, b$  sont de classe  $C^{3+\beta}$ , et que l'hypothèse de confluence forte  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{d(2(3+\beta)-p)}$  est vraie.

$(\mathbf{C}_{UE})$  *Confluence et Uniforme Ellipticité* : on suppose que  $(\mathbf{D}_p^\alpha)$  est vérifiée, que  $f$  et  $\sigma, b$  sont de classe  $C^{1+\beta}$ , que l'hypothèse de confluence  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{d(2(1+\beta)-p)}$  est vraie ainsi que la structure autonome  $(\sigma\sigma^*)_{i,j}(x) = (\sigma\sigma^*)_{i,j}(x_{i \wedge j}, \dots, x_d)$ .

**Theorem 4 (Contrôle du gradient)** *Supposons que  $(\mathcal{L}_\nu)$  et  $(\mathbf{C}_R)$  ou  $(\mathbf{C}_{UE})$  sont vérifiés alors il existe une unique fonction  $\varphi \in C^{3+\beta}$  telle que  $\mathcal{A}\varphi = f - \nu(f)$ , et :*

$$\|\nabla\varphi\|_\infty \leq \frac{[f]_1}{\alpha}.$$

**Idée de la preuve** Tout d'abord, on applique la formule d'Itô à la fonction  $y \in \mathbb{R}^d \mapsto |y|^p$ . On exploite ainsi l'hypothèse de confluence  $(\mathbf{D}_p^\alpha)$ , que l'on utilise pour contrôler le terme de type  $|\partial_{x_i} Y_t^{0,x}|^p$  apparaissant dans la formule de Feynman-Kac (2.23) associée à l'équation de Poisson. On conclut ainsi :

$$\forall i \in \llbracket 1, d \rrbracket, \forall x \in \mathbb{R}^d, |\partial_{x_i} \varphi(x)| \leq \frac{[f]_1}{\alpha}. \quad (2.24)$$

On refait le même type de calcul pour montrer que  $\nabla\varphi$  est  $\beta$ -Hölder en utilisant la condition  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{d(2(1+\beta)-p)}$  ou  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{d(2(3+\beta)-p)}$ . Dans le cas où  $(\mathbf{C}_{UE})$  est vérifié, on peut dériver l'équation de Poisson (1.10) afin d'avoir un système d'équations de Poisson avec potentiel et comme source une fonction  $\nabla f$  de continuité  $\beta$ -Hölder. Le critère de système autonome  $(\sigma\sigma^*)_{i,j}(x) = (\sigma\sigma^*)_{i,j}(x_{i \wedge j}, \dots, x_d)$  permet d'avoir un système vérifiant les conditions de [KP10] et dont on peut appliquer le bootstrap elliptique terme à terme. Remarquons également que le potentiel  $Db$  dans l'équation de Poisson est bien strictement négatif (en termes de valeurs propres) par l'hypothèse  $(\mathbf{D}_\alpha^p)$ . Dans le cas où  $(\mathbf{C}_R)$  est vérifié, on continue de dériver le flot  $Y_t^{0,x}$  jusqu'à l'ordre 3 puis on vérifie que la dérivée troisième est bien  $\beta$ -Hölder. À chaque dérivée, la condition de confluence doit être plus restrictive, i.e.  $\|D\sigma\|_\infty$  doit être de plus en plus petit avec le nombre de dérivées d'où le terme en  $3 + \beta$  dans la borne supérieure dans la condition  $(\mathbf{C}_R)$ . En d'autres termes, dans les deux jeux d'hypothèses, on obtient le contrôle de la norme  $\|\varphi\|_{C^{3,\beta}}$  en fonction de la norme associée à la régularité de  $f$  ( $\|f\|_{C^{3,\beta}}$  sous  $(\mathbf{C}_R)$ ,  $\|f\|_{C^{1,\beta}}$  sous  $(\mathbf{C}_{UE})$ ). Ce contrôle permet en particulier de gérer les termes de reste  $R_n$  dans (2.19).

### 2.3 Résultat de type Slutsky

Le défaut de l'approche du Théorème 3 est que la constante de concentration, i.e.  $\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2$ , dépend de la mesure invariante que nous approchons. Pour éviter cette difficulté, nous établissons également un résultat de type Slutsky. En effet, dans une perspective numérique, on cherche à remplacer  $\nu(\|\sigma\|^2)$  par une estimation simulable, ici  $\nu_n(\|\sigma\|^2)$ . Sous les mêmes conditions et notations que le Théorème 3 et le Théorème 4, on a :

$$\mathbb{P}\left[|\sqrt{\Gamma_n}\frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}}| \geq a\right] \leq 2C_n \exp\left(-c_n \frac{a^2\alpha^2}{2[f]_1^2}\right).$$

La démonstration est une conséquence directe du Théorème 3. Remarquons que la renormalisation par  $\sqrt{\nu_n(\|\sigma\|^2)}$  ne change pas la vitesse et permet de connaître parfaitement la grandeur de déviation. Dans la section suivante, nous montrerons que l'on peut prendre une source  $f$  qui soit seulement de régularité Lipschitz.

### 2.4 Source Lipschitz

Dans la plupart des inégalités fonctionnelles, la source est Lipschitz. On pense notamment à la distance de  $p$ -Wasserstein<sup>‡</sup> (ou appelée aussi distance de Kantorovich qui fut le premier à introduire cette distance) entre deux mesures de probabilité  $\mu_1$  et  $\mu_2$  qui est un fondamental pour le transport optimal. Pour  $p = 1$ , on a la formule de dualité de Kantorovich qui associe la distance de Wasserstein au sup sur les fonctions Lipschitz. C'est pourquoi nous avons étendu nos résultats à une source  $f$  Lipschitz. En effet, on montre que sous les mêmes hypothèses et notations que les Théorèmes 2 et 4 mais en supposant que la fonction  $f$  est seulement Lipschitz, si  $\theta \in (1/2, 1]$  alors

$$\mathbb{P}\left[|\sqrt{\Gamma_n}(\nu_n(f) - \nu(f))| \geq a\right] \leq 2C_n \exp\left(-c_n \frac{a^2\alpha^2}{2\|\sigma\|_\infty^2 [f]_1^2}\right), \quad (2.25)$$

---

‡. orthographié également Vaserstein

où  $\alpha$  est la constante issue de l'hypothèse de confluence ( $\mathbf{D}_\alpha^p$ ).

**Idée de la preuve** Il s'agit seulement de se ramener aux cas précédents, en régularisant la source  $f$ . Écrivons :  $f = f_\delta + [f - f_\delta]$ , où  $f_\delta = f \star \rho_\delta$  correspond à la convolution de  $f$  avec une fonction régularisante (ou "mollifier")  $\eta_\delta$  telle que  $f_\delta$  est de classe  $C^\infty$  et  $\eta_\delta$  a pour limite la distribution de Dirac quand  $\delta$  tend vers 0. Il faut dès lors étudier l'équation de Poisson  $\mathcal{A}\varphi_\delta = f_\delta - \nu(f_\delta)$ . Or le prix à payer de cette régularisation se trouve dans les termes de reste qui explosent en  $\delta$ . On écrit dès lors :

$$\mathbb{P}[\sqrt{\Gamma_n}(\nu_n(f) - \nu(f)) \geq a] = \mathbb{P}\left[(\nu_n(f_\delta) - \nu(f_\delta)) + R_{n,\delta}(f) \geq \frac{a}{\sqrt{\Gamma_n}}\right],$$

avec  $R_{n,\delta}(f)$  un reste tel que  $|R_{n,\delta}(f)| \leq 2C_\eta\delta[f]_1$ . Ainsi, il nous faut trouver un équilibre sur  $\delta$  qui nous conduit à considérer une nouvelle plage de vitesse, à savoir  $\theta \in (\frac{1}{2}, 1]$ . Remarquons d'ores et déjà qu'il y a une forme de continuité entre la vitesse de convergence de l'algorithme et la régularité de la source  $f$ . C'est-à-dire, pour une source  $f \in \mathcal{C}^{1,\beta}$  dans le Théorème 2 il faut que  $\theta \in (\frac{1}{2+\beta}, 1]$ . Si  $\beta = 1$  i.e.  $\nabla f$  est Lipschitz alors on retrouve bien la condition  $\theta \in (\frac{1}{3}, 1]$  du Théorème 1. Enfin si  $f$  est juste Lipschitz, cela correspond à la condition  $\theta \in (\frac{1}{2}, 1]$ , c'est le même critère que pour le Théorème 2 avec  $\beta = 0$ . En résumé, de (2.25) on obtient :

$$\sup_{f, [f]_1 \leq 1} \mathbb{P}[|\sqrt{\Gamma_n}(\nu_n(f) - \nu(f))| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2 \alpha^2}{2\|\sigma\|_\infty^2}\right).$$

Un futur travail sera de montrer que l'on peut faire « entrer » le sup dans la probabilité et de bénéficier de la formule de dualité de Kantorovich pour avoir une inégalité de concentration non-asymptotique de la distance de Wasserstein. On pourrait imaginer utiliser des méthodes de localisation ou de couplage subtiles pour y parvenir (voir e.g. [Boi11], [BGV07]). Dans la prochaine section, nous présentons l'inégalité de concentration optimale qui correspond asymptotiquement au Théorème 1.

### 3 Inégalité de concentration optimale

Comme nous le verrons dans le Chapitre 4, la constante de concentration a une grande importance pour la précision des simulations numériques. D'où l'intérêt d'avoir l'inégalité de concentration non-asymptotique avec comme constante de concentration le *carré du champ*, c'est-à-dire la variance apparaissant dans le TCL du Théorème 1.

**Theorem 5 (Inégalité de concentration non-asymptotique optimale)** *Sous les mêmes hypothèses et notations que le Théorème 2, s'il existe une fonction  $\vartheta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  de même régularité que  $\varphi$  et telle que  $\mathcal{A}\vartheta = |\sigma^*\nabla\varphi|^2 - \nu(|\sigma^*\nabla\varphi|^2)$ , alors, pour tout  $\theta \in (\frac{1}{2+\beta}, 1]$ ,  $n \geq 1$ ,  $a > 0$  vérifiant  $a/\sqrt{\Gamma_n} \rightarrow 0$  (déviations Gaussiennes) :*

$$\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\nu(|\sigma^*\nabla\varphi|^2)}\right).$$

L'hypothèse sur le caractère cobord de  $|\sigma^*\nabla\varphi|^2 - \nu(|\sigma^*\nabla\varphi|^2)$  est vraie sous les hypothèses du Théorème 4. Le cas où  $a/\sqrt{\Gamma_n} \rightarrow +\infty$  conduit à l'inégalité (2.22).

**Idée de la preuve** Il s'agit d'améliorer encore une fois l'estimation de  $[\Delta_k(X_{k-1}, \cdot)]_1$  faite dans (2.18) et dans l'idée de la preuve du Théorème 3. Nous chercherons à approcher  $[\Delta_k(X_{k-1}, \cdot)]_1$  par  $\sqrt{\gamma_k} \sigma^* \nabla \varphi(X_{k-1})$ , à savoir la racine carré du *carré du champ*. Pour cette estimation du module Lipschitz, par souci de simplification, nous nous placerons dans le cadre unidimensionnel et où l'innovation  $(U_k)_{k \geq 1}$  est bornée. Remarquons dans un premier temps que pour  $d = 1$ , pour  $\varphi$  solution de l'équation de Poisson  $\mathcal{A}\varphi = f - \nu(f)$ , il existe une constante  $C_\varphi > 0$  telle que pour tout  $x \in \mathbb{R} : |D^2\varphi(x)| \leq C_\varphi/(1 + |x|)$ . On utilise dès lors cette inégalité dans l'estimation de la norme Lipschitz avec l'hypothèse  $\|U_k\|_\infty < \infty$  :

$$|\nabla_u \Delta_k(X_{k-1}, u)|_{u=U_k} \leq |\sqrt{\gamma_k} \sigma_{k-1}^* \nabla \varphi(X_{k-1})| + C\gamma_k,$$

où  $C = C(\|U_k\|, C_\varphi, \sigma, \nabla \varphi) > 0$ . Le terme  $C\gamma_k$  est un terme de reste. De (2.21), on a :

$$\mathbb{E}[\exp(\frac{\lambda M_n}{\Gamma_n})] \leq \mathbb{E}[\exp(\frac{\lambda^2 q}{2\Gamma_n^2} \sum_{k=1}^n \gamma_k |\sigma^* \nabla \varphi(X_{k-1})|^2 + C\gamma_k^{3/2})]^{1/q} \quad (3.26)$$

Or par hypothèse, il existe  $\vartheta$  de classe  $C^3$  telle que  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$ . Ainsi le terme  $\nu(\|\sigma\|^2)$  est celui qui va donner la constante de concentration, tandis que celui en  $\mathcal{A}\vartheta$  se traite de la même manière que précédemment  $\mathcal{A}\varphi$ . Dans le cas de la dimension supérieure, l'inégalité  $|D^2\varphi(x)| \leq C_\varphi/(1 + |x|)$  n'est plus valable. Il faut utiliser le fait que  $x \mapsto \langle b(x), \nabla \varphi(x) \rangle$  est Lipschitz. Dans le cas où l'innovation est non bornée, il faut procéder à une méthode de localisation subtile. Pour plus de précisions se référer au Chapitre 4.

Dans la section suivante, nous montrons que dans ce contexte on peut encore se ramener à une source  $f$  seulement Lipschitz.

### 3.1 Source Lipschitz

Dans le même cadre que la Section 2.4, nous établissons le résultat suivant : sous les mêmes hypothèses et notations que les Théorèmes 5 et 4 mais en supposant que la fonction  $f$  est seulement Lipschitz, on a :

$$\mathbb{P}[|\sqrt{\Gamma_n}(\nu_n(f) - \nu(f))| \geq a] \leq 2C_n \exp(-c_n \frac{a^2}{2\nu(|\sigma^* \nabla \varphi|^2)}), \quad (3.27)$$

où  $\varphi \in \mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R}) \cap W_{2,loc}^2(\mathbb{R}^d, \mathbb{R})$  est une solution au sens faible de l'équation de Poisson  $\mathcal{A}\varphi = f - \nu(f)$ . Nous avons le pendant du résultat de la Section 2.4 mais avec la constante de concentration optimale,  $\nu(|\sigma^* \nabla \varphi|^2)$ .

**Idée de la preuve** La preuve de ce théorème, comme pour le résultat de la Section 2.4, s'articule aussi autour d'une régularisation. Cependant, ici  $\|\varphi_\delta\|$  n'est pas forcément borné quand  $\delta \rightarrow 0$ . En effet, on a seulement besoin de régularité Sobolev, i.e. que  $\varphi \in W_{2,loc}^2(\mathbb{R}^d, \mathbb{R}) \implies \nu(|\sigma^* \nabla \varphi|^2) < +\infty$ . Pour se faire, nous utilisons le Theorem 1 de [PV01] qui nous donne cette régularité de la solution de Poisson (1.10).

Il serait tentant d'établir un théorème de type Slutsky comme dans la Section 2.3 avec le *carré du champ*. Il ne serait pas compliqué d'établir une inégalité de concentration de la déviation de  $\frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(|\sigma^* \nabla \varphi|^2)}}$ . En revanche, à  $f$  donné, nous ne connaissons pas  $\varphi$  ni

ses dérivées. Donc ce qu'il resterait potentiellement à faire dans cette direction serait de trouver une suite (aléatoire ou non) dépendant de  $f$  qui converge presque sûrement vers  $\nu(|\sigma^*\nabla\varphi|^2)$ . Remarquons aussi que  $Var_\nu(f) = \nu(f^2) - \nu(f)^2$  n'a aucune raison d'être égale à celle du TCL  $\nu(|\sigma^*\nabla\varphi|^2)$ . Dans la partie suivante, nous chercherons à tester la robustesse de notre méthode en considérant des EDS dirigées par des processus à accroissements indépendants et stationnaires plus généraux, des processus de Lévy.

## 4 Extension aux EDS dirigées par un Lévy

On considère maintenant, au lieu de la diffusion (1.1), une EDS avec un processus de Lévy :

$$d\mathbf{X}_t = b(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)dW_t + \kappa(\mathbf{X}_{t-})d\mathbf{Z}_t, \quad (4.28)$$

où  $(Z_t)_{t \geq 0}$  un processus de Lévy de carré intégrable à valeur dans  $\mathbb{R}^r$  (par souci de simplicité), et  $\kappa : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  une fonction Lipschitz. L'existence de mesure invariante a été montrée par Panloup dans [Pan08b]. L'existence et l'unicité ont été établies dans [Mas07] sous hypothèse d'irréductibilité. Les premiers à avoir proposé un schéma de discrétisation d'une EDS à saut furent Protter et Talay dans l'article [PT97], où ils établirent l'erreur faible associée à un schéma d'Euler dit «exact» (avec les vrais accroissements de processus de Lévy) pour un algorithme de Monte Carlo. Par la suite, Panloup a proposé différents schémas qui permettent la simulation du processus dirigé par tout Lévy de carré intégrable.

Pour utiliser nos techniques, notamment la transformation exponentielle, il nous faut considérer des intensités de sauts sous-Gaussiens, nous nous sommes ainsi concentrés sur des sauts de type Poisson à sauts sous-Gaussiens. Dans un cadre plus général, un schéma de discrétisation à pas décroissant pour cette EDS, (4.28) a été introduit dans [Pan08b] ainsi que la preuve du Théorème ergodique (1.16) associé. Il s'agit de l'adaptation du schéma et des arguments pour les EDS Browniennes de Lambert Pagès [LP02]. Cependant, pour la partie saut, le pas décroissant du schéma se retrouve dans le temps et non pas en facteur de l'innovation. Autrement dit, on considère que la probabilité de saut du processus discrétisé de  $Z_t$  au  $n$ -ième pas de temps est linéaire au pas de temps  $\gamma_n$ . Cette approche différente liée à la présence de sauts induit de nouvelles difficultés. En effet, un des outils fondamentaux utilisé jusque-là se résume en cette phrase : «plus on fait du développement de Taylor, plus on fait sortir du  $\gamma_k$ ». Donc pour pouvoir appliquer nos techniques, on ne peut prendre directement un processus de Poisson composé dans notre schéma, puisqu'il ne vérifie pas *a priori* l'inégalité de concentration **(GC)** même avec sauts sous-Gaussien. C'est pourquoi, nous avons choisi un cas particulier du schéma qui permet d'approcher un processus de Poisson composé à sauts sous-Gaussien et qui vérifie **(GC)** :

$$X_{n+1} = X_n + \gamma_{n+1}b(X_n) + \sqrt{\gamma_{n+1}}\sigma(X_n)U_{n+1} + \kappa(X_n)Z_{n+1}, \quad (4.29)$$

avec  $U_{n+1}$  et  $\gamma_{n-1}$  définis comme dans (1.14) et pour tout  $n \geq 1$  :

$$Z_n := B_n Y_n, \quad (4.30)$$

avec  $(B_n)_{n \geq 1}$  est une suite i.i.d. unidimensionnelle et suivant une loi de Bernoulli, i.e.  $B_n \stackrel{\mathcal{L}}{=} \text{Bern}(\gamma_n)$ , indépendantes de  $X_0$ , de  $(U_n)_{n \geq 1}$  et de  $(Y_n)_{n \geq 1}$ . Enfin,  $(Y_n)_{n \geq 1}$  est aussi

une suite i.i.d. telle que  $\mathbb{E}[Y_n^{\otimes 2}] = I_r$  et vérifie **(GC)**. Ce choix d'innovation rend compte du fait que la probabilité d'avoir au moins un saut pour un processus de Poisson composé tend vers 0 avec le temps. En fait, Panloup a établi ces résultats dans un cadre d'hypothèse de Lyapunov plus large que celle présentée en  $(\mathcal{L}_V)$ , qui permet entre autres un retour à la moyenne moins fort, i.e. en particulier il suppose qu'il existe  $\mathbf{a} \in (0, 1]$  tel que  $\langle b, \nabla V \rangle \leq -\alpha_V V^{\mathbf{a}} + \beta_V$  (voir aussi [Lem07] pour  $\sigma$  non borné). L'étude d'inégalité de concentration non-asymptotique pour des rappels plus faibles que ceux considérés dans  $(\mathcal{L}_V)$  sera l'objet de futurs travaux. Notons de plus que l'hypothèse de Lyapunov supposée jusque-là n'a pas besoin d'être changée. En effet, écrivons le générateur infinitésimal associé à la diffusion (4.28) pour toute fonction  $\varphi \in \mathcal{C}^2(\mathbb{R}^d)$  et  $x \in \mathbb{R}^d$  :

$$\mathcal{A}\varphi(x) = b(x)\nabla\varphi(x) + \frac{1}{2}\text{Tr}(\sigma\sigma^*(x)D^2\varphi(x)) + \int_{\mathbb{R}^d} (\varphi(x + \kappa(x)y) - \varphi(x))\pi(dy) \quad (4.31)$$

où  $\pi$  est la distribution de  $Y_1$ . On pose  $\tilde{\mathcal{A}}\varphi(x) = b(x)\nabla\varphi(x) + \frac{1}{2}\text{Tr}(\sigma\sigma^*(x)D^2\varphi(x))$  correspondant au générateur infinitésimal de la partie continue de la diffusion. Il se trouve que pour toute fonction  $\varphi$  de classe  $C^2$  avec  $\nabla\varphi, D^2\varphi$  bornées, on obtient par nos hypothèses que  $|\mathcal{A}\varphi - \tilde{\mathcal{A}}\varphi|$  est bornée par  $\frac{1}{2}\|D^2\varphi\|_{\infty}\|\kappa\|_{\infty}^2(|\cdot|^2)$ . Sous de bonnes hypothèses de stabilité, dans [Pan08a] l'auteur établit le TCL correspondant.

**Theorem 6 (TCL de [Pan08a])** *Supposons  $(\mathcal{L}_V)$ , s'il existe  $p > 0$  t.q  $\mathbb{E}|Z_t|^{2p} < +\infty$  alors pour tout  $\theta \in (\frac{1}{3}, 1]$  et  $\varphi$  avec la même régularité que pour le Théorème 1 :*

$$\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int_{\mathbb{R}^d} (|\sigma^*\nabla\varphi|^2(x) + \int_{\mathbb{R}^d} |\varphi(x + \kappa(x)y) - \varphi(x)|^2\pi(dy))\nu(dx)\right).$$

Remarquons que la variance asymptotique est différente ici par rapport au cas continu. En effet, on y retrouve le *carré du champ* continu plus un terme dû aux sauts :

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\varphi(x + \kappa(x)y) - \varphi(x)|^2\pi(dy)\nu(dx).$$

La présence de sauts modifie le générateur associé à la diffusion, et donc le *carré du champ*. Pour une étude complète et plus pratique, il faudrait étudier le comportement non pas de  $\nu_n(\mathcal{A}\varphi)$  mais de  $\nu_n(f)$  en toute généralité, i.e. étudier l'équation intégro-différentielle de Poisson, ce qui fera l'objet de futurs travaux.

**Theorem 7 (Concentration pour une EDS à sauts)** *Sous les mêmes hypothèses et notations que le Théorème 2, pour  $\forall n \geq 1, 0 < a \ll \sqrt{\Gamma_n}/\Gamma_n^{(2)}$  :*

$$\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2(2\|\kappa\|_{\infty}^2\|\nabla\varphi\|_{\infty}^2 + \|\sigma\|_{\infty}^2\|\nabla\varphi\|_{\infty}^2)}\right).$$

La constante de concentration est «polluée» par une constante multiplicative devant le terme de saut. Contrairement au Théorème 2, nous n'avons pas obtenu une simple borne supérieure du *carré du champ* mais deux fois la borne supérieure du terme de sauts. Cela est dû à la difficulté de vérifier que  $(Z_n)_{n \geq 1}$  satisfasse **(GC)**. Le principal obstacle dans ce

travail est la plage de validité de la concentration de  $(Z_n)_{n \geq 1}$ . En effet, on établit l'inégalité de concentration de type **(GC)** mais avec  $\varepsilon \in (0, 1)$ , et tel que  $0 < \lambda < \frac{\varepsilon}{6[g]_1}$  on a

$$\mathbb{E}e^{\lambda g(Z_n)} \leq e^{\lambda \mathbb{E}g(Z_n) + \frac{\lambda^2 \gamma_n(2+\varepsilon)[g]_1^2}{2}}. \quad (4.32)$$

Remarquons qu'il y a maintenant une contrainte sur la borne de validité du  $\lambda$ , et qu'on voit apparaître un terme en  $(2 + \varepsilon)$  qui vient «perturber» l'inégalité de concentration. C'est cette perturbation qui permet d'avoir cette inégalité, pour un  $\lambda$  suffisamment petit, en faisant un développement en série entière. Ces contraintes sur la concentration impliquent une plus petite intégrabilité de la fonction de Lyapunov, i.e.  $(\sqrt{V}(X_n))_{n \geq 0}$  est exponentiellement intégrable. Pour la suite de la preuve, il s'agit dans les grandes lignes de la même chose que pour le Théorème 2. La principale difficulté est de réussir à intégrer les contraintes de la concentration (4.32).

Une ouverture possible pour la suite serait de s'intéresser à des Lévy plus généraux avec des contrôles de type Berry-Esseen.

# Chapitre 2

## Régularisation par un bruit dégénéré

### 1 Régularisation par un bruit dégénéré

**Résumé :** Dans une seconde partie, nous établissons des estimées de Schauder liées à des équations paraboliques associées à un système stochastique dégénéré, où la dérive est un champ de vecteurs vérifiant une condition de type Hörmander (faible) mais en cherchant la régularité Hölder minimale sur les coefficients. Ce travail fait suite à l'article de Delarue et Menozzi [DM10]. Enfin, notre approche nous permet de montrer l'unicité forte du système stochastique considéré dans le cadre de coefficients Hölder, étendant ainsi le résultat obtenu en dimension 2 par Chaudru de Raynal [CdR17].

#### 1.1 Équation de Kolmogorov et chaîne dégénérée

L'étude des chaînes dégénérées est prégnante dans différents domaines. Par exemple, en finance, la dynamique d'une option asiatique forme une chaîne dégénérée de dimension 2. Ce sont des options dont le résultat à son échéance (ce que gagne le détenteur, appelé aussi *payoff*) dépend du prix moyen de l'actif sous-jacent. En d'autres termes, le *payoff* d'un call est  $(\frac{1}{T} \int S_t dt - K)_+$ , où  $T$  est l'échéance,  $K$  le strike, et  $S_t$  le sous-jacent à l'instant  $t$ . Pour plus d'applications en finance, mentionnons par exemple Jeanblanc *et al.* [JYC09]. Ces équations sont dites de type vitesse/position. Cela peut se voir comme un système dont les inconnues sont  $\mathbf{X}_t$  (la position) et  $\dot{\mathbf{X}}_t$  (la vitesse), où  $\mathbf{X}_t$  ne dépend que de  $t$  et de  $\dot{\mathbf{X}}_t$ . Autrement dit, le bruit n'affecte directement que  $\dot{\mathbf{X}}_t$  et se propage en  $\mathbf{X}_t$  via sa dépendance en  $\dot{\mathbf{X}}_t$ . En particulier, on retrouve ce système vitesse/position dans la mécanique Hamiltonienne. Dans un cadre plus général, nous nous intéresserons à des chaînes d'oscillateurs (voir par exemple les oscillateurs élasto-plastiques considérés dans [BT08] et [BMPT09]) qui sont un modèle typique en sismologie par exemple, où l'on considère la propagation d'un aléa sur plusieurs structures qui se transmettent la secousse. Le modèle souvent utilisé pour représenter ce phénomène est un système de ressorts attachés les uns aux autres.

Enfin mentionnons également les modèles microscopiques associés à la diffusion de la chaleur (voir par exemple Heckmann et Hairer [EH00]).



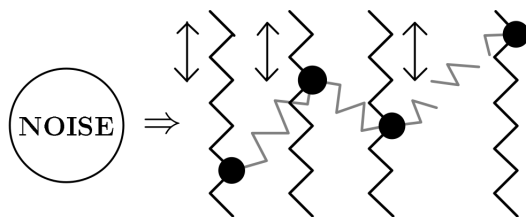


FIGURE 2.1 – Dessin venant de [DM10].

Dans cette partie, nous nous intéresserons à l'équation de Kolmogorov qui s'écrit pour  $T > 0$  :

$$\begin{cases} (\partial_t + L)u(t, \cdot) = -f(t, \cdot), & t \in [0, T), \\ u(T, \cdot) = 0, \end{cases} \quad (1.1)$$

où  $L$  est un opérateur différentiel qui peut se voir comme le générateur infinitésimal associé à un processus de diffusion  $(X_t)_{t \geq 0}$ . Nous étudierons cette équation à temps fini, i.e. pour tout  $t \in [0, T]$  à  $T > 0$  donné. Pour se ramener à l'équation de Poisson considérée dans la partie précédente, il faut considérer un potentiel qui permet d'avoir des estimées en temps long apparaissant dans la représentation de Feynman-Kac (2.23). Ce terme de potentiel apparaît notamment quand on dérive l'équation de Kolmogorov et sous la condition de Lyapunov  $(\mathcal{L}_V)$ . L'approche que nous développerons après permet cette procédure.

Dans un premier temps, nous établirons des estimées de Schauder associées à l'équation (1.1) pour des chaînes dégénérées. Ces estimées rendent compte du *bootstrap* parabolique. Illustrons tout d'abord cette notion de bootstrap parabolique dans un cadre uniformément elliptique à coefficients bornés. Dans ce cas, pour des coefficients suffisamment réguliers en temps et en espace, à l'aide de méthode de contrôles *a priori*, Friedmann [Fri64] et Krylov [Kry96] ont montré le contrôle suivant :

$$\|u\|_{C_b^{\frac{2+\gamma}{2}, 2+\gamma}} \leq C \|f\|_{C_b^{\frac{\gamma}{2}, \gamma}},$$

où  $\|\cdot\|_{C_b^{\frac{2+\gamma}{2}, 2+\gamma}}$ ,  $\|\cdot\|_{C_b^{\frac{\gamma}{2}, \gamma}}$  sont les normes Hölder classiques. Plus précisément, si  $f$  est  $\frac{\gamma}{2}$ -Hölder en temps et  $\gamma$ -Hölder en espace alors la solution est  $\frac{2+\gamma}{2}$ -Hölder en temps et  $2 + \gamma$ -Hölder en espace. Le caractère borné des coefficients en espace et leur régularité Hölder en temps permet de considérer le gain de régularité sur la variable temporelle.

Néanmoins, sans cette hypothèse de bornitude, à notre connaissance il n'existe pas de contrôles Schauder pour des problèmes paraboliques généraux. Le cas où les coefficients sont bornés en temps et à croissance linéaire en espace

$$\|u\|_{L^\infty(C_b^{2+\gamma})} \leq C \|f\|_{L^\infty(C^\gamma)},$$

où  $\|\cdot\|_{L^\infty(C^{2+\gamma})}$ ,  $\|\cdot\|_{L^\infty(C^\gamma)}$  sont les normes uniformes en temps et respectivement  $2 + \gamma$  et  $\gamma$  Hölder homogène en espace (la non-bornitude de  $f$  est gérée par la présence d'un

potentiel). Cette estimée, dans un cadre plus proche du nôtre mais avec un potentiel séparé de 0, a été montré par Krylov et Priola [KP10]. La méthode usuelle utilisée pour les estimées de Schauder et appliquée dans cet article est également une méthode de contrôle *a priori*. Nous procéderons par une méthode dite *perturbative*.

Ce *bootstrap* parabolique à coefficient non-bornés signifie que si  $f$  est bornée et  $\gamma$ -Hölder,  $\gamma \in (0, 1)$ , alors la solution  $u$  est bornée et  $\mathbf{2} + \gamma$ -Hölder, c'est-à-dire  $u$  est 2 fois dérivable, chaque dérivée est bornée et  $D^2u$  est  $\gamma$ -Hölder.

Mentionnons également, Kruzhkov et al. [KCL75] qui ont établi des estimés de Schauder dans un cadre elliptique pour des coefficients bornés en temps et en espace mais sans supposer de régularité en temps. Par la suite, Lorenzi [Lor11] a étendu ces résultats au cadre non-borné.

De plus, cette estimée montre également, en un certain sens, l'unicité de l'équation (1.1). En effet, si  $u_1$  et  $u_2$  sont deux solutions de (1.1) alors on a :

$$(\partial_t + L)(u_1 - u_2) = 0,$$

l'estimée de Schauder donne ainsi :

$$\|u_1 - u_2\|_{C^{2+\gamma}} \leq 0.$$

Ce qui conduit directement à  $u_1 = u_2$ .

L'étude de cette équation de Kolmogorov (1.1) est également centrale dans la transformation de Zvonkin pour établir l'unicité forte de l'EDS associée.

Nous disons que la solution d'une EDS dirigée par un Brownien est *forte* si elle est l'unique processus vérifiant l'EDS et étant adapté à la filtration Brownienne. Nous montrons l'unicité trajectorielle, et par la propriété de Yamada Watanabe [YW71], nous savons que l'unicité forte (ou unicité dans l'espace des réalisations) est déduite de l'unicité trajectorielle et de l'existence faible (ou dans l'espace des probabilités).

## 1.2 Présentation du modèle

Considérons le problème de Cauchy associé à l'équation de Kolmogorov, pour tout  $T > 0$ ,  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ ,

$$\begin{cases} \partial_t u(t, \mathbf{x}) + \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}u(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u(t, \mathbf{x}) a(t, \mathbf{x})) = -f(t, \mathbf{x}), \\ u(T, \mathbf{x}) = g(\mathbf{x}), \mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{nd}, \end{cases} \quad (1.2)$$

où la dérive a la structure suivante  $\mathbf{F}(t, \mathbf{x}) := (\mathbf{F}_1(t, \mathbf{x}), \dots, \mathbf{F}_n(t, \mathbf{x}))$  telle que

$$\forall i \in \llbracket 2, n \rrbracket, \mathbf{F}_i(t, \mathbf{x}) := \mathbf{F}_i(t, \mathbf{x}^{i-1:n}), \mathbf{x}^{i-1:n} := (\mathbf{x}_{i-1}, \dots, \mathbf{x}_n). \quad (1.3)$$

Autrement dit, chaque composante  $\mathbf{F}_i$  a une seule variable qui transmet le bruit, à savoir  $\mathbf{x}_{i-1}$ . Cette équation de Kolmogorov (1.2) est dégénérée au sens où l'opérateur différentielle d'ordre 2,  $D_{\mathbf{x}_1}^2$  n'est pris que sur la première variable, la variable dite *non-dégénérée*.

Formellement l'équation de Kolmogorov (1.2), est étroitement liée au processus stochastique suivant :

$$\begin{aligned}
 d\mathbf{X}_t^1 &= \mathbf{F}_1(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt + \sigma(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dW_t, \\
 d\mathbf{X}_t^2 &= \mathbf{F}_2(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt, \\
 d\mathbf{X}_t^3 &= \mathbf{F}_3(t, \mathbf{X}_t^2, \dots, \mathbf{X}_t^n)dt, \\
 &\vdots \\
 d\mathbf{X}_t^n &= \mathbf{F}_n(t, \mathbf{X}_t^{n-1}, \mathbf{X}_t^n)dt,
 \end{aligned}
 \tag{1.4}$$

avec  $\sigma$  tel que  $a = \sigma\sigma^*$ . Le caractère dégénéré de la diffusion associée est d'autant plus évident sous cette forme. Nous noterons également le générateur infinitésimal associé :  $(L_t)_{t \in [0, T]}$ , pour toute fonction  $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$  :

$$L_t\varphi(\mathbf{x}) = \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}\varphi(\mathbf{x}) \rangle + \frac{1}{2}\text{Tr}(D_{\mathbf{x}_1}^2\varphi(\mathbf{x})a(t, \mathbf{x})).
 \tag{1.5}$$

Par la suite, nous supposerons des hypothèses de régularité (Hölder appropriée) et de non-dégénérescence (qui conduit à des hypothèses de type Hörmander faible) sur les composants  $a, \mathbf{F}$  ce qui implique que le problème martingale associé à (1.5) est bien posé, voir par exemple Menozzi [Men11] (dérive Lipschitz et  $\sigma$  Hölder), [Men18] ( $\sigma$  continu et dérive Lipschitz), et Chaudru de Raynal et Menozzi [CdRM17] (dérive et  $\sigma$  Hölder à coefficient «sharp») et montre l'unicité faible de (1.33).

### 1.3 Propriété d'échelles

Donnons tout d'abord un exemple classique du système d'équation stochastique (1.33), l'exemple de Kolmogorov (Ann. Math. 1934 [Kol34]) :

$$\begin{aligned}
 X_t^1 &= x_0^1 + W_t, \\
 X_t^2 &= x_0^2 + \int_0^t X_s^1 ds,
 \end{aligned}$$

où l'on a pris  $n = 2$ ,  $\mathbf{F}_1(\mathbf{x}) = \mathbf{x}_1$ ,  $\mathbf{F}_2(\mathbf{x}) = \mathbf{x}_1$ , et  $\sigma(\mathbf{x}) = 1_{d,d}$ . Il s'agit d'un processus Gaussien  $(X_t^1, X_t^2) \sim \mathcal{N}(\theta_t(x_0), \mathbf{K}_t)$  tel que sa densité est :

$$p(t, x_0, y) = \frac{(2\sqrt{3})^d}{(2\pi)^{d t^{(\frac{1}{2} + \frac{3}{2})d}}} \exp\left(-\frac{1}{2}\langle \mathbf{K}_t^{-1}(\theta_t(x_0) - y), \theta_t(x_0) - y \rangle\right),$$

avec  $\mathbf{K}_t$  la matrice de covariance qui vérifie la propriété de la «bonne échelle» (voir la Définition 3.2 dans [DM10]) :

$$\mathbf{K}_t^{1/2} = \begin{pmatrix} tI_d & \frac{t^2}{2}I_d \\ \frac{t^2}{2}I_d & \frac{t^3}{3}I_d \end{pmatrix}^{1/2} \asymp \begin{pmatrix} t^{1/2}I_d & 0_d \\ 0_d & t^{3/2}I_d \end{pmatrix}.
 \tag{1.6}$$

Pour des matrices  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{nd} \otimes \mathbb{R}^{nd}$ , la notation  $\mathbf{A} \asymp \mathbf{B}$  signifie qu'il existe une constante  $C \geq 1$  telle que pour tout  $\zeta \in \mathbb{R}^{nd}$ , on a  $C^{-1}|\mathbf{A}\zeta|^2 \leq |\mathbf{B}\zeta|^2 \leq C|\mathbf{A}\zeta|^2$ . Cette échelle

rend compte de l'homogénéité qu'il y a entre  $X_t^1$  et le mouvement Brownien, et  $X_t^2$  et l'intégrale du Brownien.

Dans un cadre plus général, Lunardi [Lun97] a montré des estimées de Schauder associées à (1.2), lorsque  $a$  a une régularité Hölder appropriée (notée  $(\mathbf{S})$  plus loin) et est uniformément elliptique. Notamment, elle a considéré une chaîne dégénérée complète avec pour tout  $n \in \mathbb{N}$  donné et une dérive linéaire  $\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ , avec

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_{1,1} & \cdots & \cdots & \cdots & \mathbf{a}_{1,n} \\ \mathbf{a}_{2,1} & \cdots & \cdots & \cdots & \mathbf{a}_{2,n} \\ \mathbf{0}_{d,d} & \mathbf{a}_{3,2} & \cdots & \cdots & \mathbf{a}_{3,n} \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & \mathbf{a}_{n,n-1} & \mathbf{a}_{n,n} \end{pmatrix},$$

où  $(\mathbf{a}_{i,j})_{i,j \in \llbracket 1,n \rrbracket^2} \in \mathbb{R}^d \otimes \mathbb{R}^d$  sont tels que  $(\mathbf{a}_{i,i-1}) \in GL_d(\mathbb{R})$ ,  $i \in \llbracket 2,n \rrbracket^2$ . En particulier, remarquons que la matrice  $\mathbf{A}$  est compatible avec la structure de la dérive dans (1.3), ce qui implique que le système vérifie des hypothèses de type Hörmander faibles. Enfin, le processus

$$\mathbf{X}_t = \mathbf{x} + \int_0^t \mathbf{A}\mathbf{X}_s ds + \int_0^t \begin{pmatrix} \sigma(\mathbf{X}_s) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} dW_s$$

est contrôlé par un processus Gaussien en termes de densité pour les marginales (cf [DM10], [Men11]), avec comme matrice de covariance  $\mathbf{K}_t$  vérifiant également la propriété de «bonne échelle». Autrement dit, on a :

$$\mathbf{K}_t^{1/2} \asymp t^{1/2} \mathbb{T}_t = \begin{pmatrix} t^{1/2} \mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & t^{3/2} \mathbf{I}_{d,d} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & t^{(2n-1)/2} \mathbf{I}_{d,d} \end{pmatrix}. \quad (1.7)$$

Insistons encore une fois sur le fait que cette matrice rend compte des échelles intrinsèques des composants du processus, tout comme pour l'exemple de Kolmogorov. Les coefficients diagonaux reflètent l'homogénéité en temps des intégrales itérées du mouvement Brownien.

Un modèle typique du cas considéré par Lunardi a pour opérateur correspondant dans l'équation (1.2) :

$$\mathcal{L}_0 := \partial_t + \langle \mathbf{A}_0 \mathbf{x}, D \rangle + \frac{1}{2} \Delta_{\mathbf{x}_1},$$

avec

$$\mathbf{A}_0 = \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{I}_{d,d} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \mathbf{0}_{d,d} \end{pmatrix},$$

Considérons dès lors l'opérateur de dilatation adapté à l'échelle du système stochastique, à savoir pour tout  $\lambda > 0$ ,

$$\delta_\lambda : (t, \mathbf{x}) \in \mathbb{R}^+ \times \mathbb{R}^{nd} \mapsto \delta_\lambda(t, \mathbf{x}) = (\lambda^2 t, \lambda \mathbf{x}_1, \lambda^3 \mathbf{x}_2, \dots, \lambda^{2n-1} \mathbf{x}_n).$$

Maintenant, remarquons que l'opérateur sus-défini  $\mathcal{L}_0$  est invariant par la dilatation, c'est-à-dire :  $\mathcal{L}_0 v = 0 \implies \mathcal{L}_0(v \circ \delta_\lambda) = 0$ . Nous voyons ainsi l'importance des différentes échelles associées à chaque variable  $(\mathbf{x}_i)_{i \in \llbracket 1, n \rrbracket}$ . Il est dès lors naturel de considérer la pseudo-distance homogène associée :

$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^d |\mathbf{y}_i - \mathbf{x}_i|^{\frac{1}{2i-1}}. \quad (1.8)$$

Une fois cette pseudo-distance en main, nous introduisons les différentes normes Hölder aux bonnes échelles. Pour toute fonction  $u : [0, T] \times \mathbb{R}^{nd} \mapsto \mathbb{R}$ ,  $T > 0$  fixé, et pour tout  $\gamma \in (0, 1)$ , on note pour tout  $i \in \llbracket 1, n \rrbracket$  le module Hölder associé à la variable  $\mathbf{x}_i$  :

$$[u_i]_{\mathbf{d}}^\gamma := \sup_{t, \mathbf{x}, \mathbf{x}'_i} \frac{|u(t, \mathbf{x}) - u(t, \mathbf{x}_{1:i-1}, \mathbf{x}'_i, \mathbf{x}_{i+1})|}{|\mathbf{x}_i - \mathbf{x}'_i|^{\frac{\gamma}{2i-1}}}.$$

Nous notons le module Hölder homogène complet comme la somme de tous les modules Hölder suivant les variables à l'échelle correspondante,

$$\|u\|_{L^\infty(C_{\mathbf{d}}^\gamma)} := \sum_{i=1}^n [u_i]_{\mathbf{d}}^\gamma \asymp \sup_{(t, \mathbf{x}, \mathbf{x}')} \frac{|u(t, \mathbf{x}) - u(t, \mathbf{x}')|}{\mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')} ,$$

remarquons à ce stade que  $L^\infty$  signifie la norme uniforme suivant la composante de temps. Notons également la norme Hölder homogène totale quand l'indice est de type  $2 + \gamma$  :

$$\|u\|_{L^\infty(C_{\mathbf{d}}^{2+\gamma})} := \|D_{\mathbf{x}_1} u\|_{L^\infty} + \|D_{\mathbf{x}_1}^2 u\|_{L^\infty} + \sum_{i=1}^n [D_{\mathbf{x}_1}^2 u_i]_{\mathbf{d}}^\gamma + \sum_{i=2}^n [u_i]_{\mathbf{d}}^{2+\gamma}.$$

Pour établir les estimées de Schauder, il conviendra en particulier de montrer que la norme précédente associée à la solution  $u$  de (1.2) est finie. C'est-à-dire que suivant la  $i$ -ème variable,  $u$  est  $\frac{2+\gamma}{2i-1}$ -Hölder si  $i > 1$  (sinon on doit considérer des indices  $> 1$ ) et que suivant la  $i$ -ème variable,  $D_{\mathbf{x}_1}^2 u$  est  $\frac{\gamma}{2i-1}$ -Hölder. En fait, nous supposons que la source  $f$  et que la condition terminale sont bornées, on écrit ainsi pour tout  $\tilde{\gamma} \in (0, 3)$  (en pratique  $\tilde{\gamma} \in \{\gamma, 2 + \gamma\}$ ) les pseudo-normes précédemment introduites avec un «b» pour préciser que la fonction considérée est bornée.

$$\|u\|_{L^\infty(C_{b, \mathbf{d}}^{\tilde{\gamma}})} := \|u\|_{L^\infty} + \|u\|_{L^\infty(C_{\mathbf{d}}^{\tilde{\gamma}})}.$$

Nous écrivons de ce fait l'espace de Hölder associé aux normes introduites :

$$L^\infty(C_{\mathbf{d}}^{\tilde{\gamma}}) := L^\infty([0, T], C_{\mathbf{d}}^{\tilde{\gamma}}) = \{u \in \mathcal{B}([0, T] \times \mathbb{R}^{nd}, \mathbb{R}) : \|u\|_{L^\infty(C_{\mathbf{d}}^{\tilde{\gamma}})} < \infty\}$$

où  $\mathcal{B}([0, T] \times \mathbb{R}^{nd}, \mathbb{R})$  correspond à l'ensemble des fonctions Boréliennes de  $[0, T]$  dans  $\mathbb{R}$ . Pareillement pour  $L^\infty(C_{b, \mathbf{d}}^{\tilde{\gamma}}) := L^\infty([0, T], C_{b, \mathbf{d}}^{\tilde{\gamma}}) = \{u \in \mathcal{B}([0, T] \times \mathbb{R}^{nd}, \mathbb{R}) : \|u\|_{L^\infty(C_{b, \mathbf{d}}^{\tilde{\gamma}})} < \infty\}$  et  $C_{b, \mathbf{d}}^{\tilde{\gamma}}$  si la fonction ne dépend pas du temps (la condition terminale  $g$  de (1.2) par exemple).

## 1.4 Résultats existants

Après avoir introduit ces notations, nous pouvons maintenant énoncer un premier contrôle Schauder associé au problème de Cauchy (1.2).

**Theorem 1 (Estimées de Schauder de [Lun97])** *Pour  $\sigma \in L^\infty(C_{b,d})$ , Uniformément Elliptique et telle qu'il existe  $\sigma_0 \in GL_{nd,d}$  t.q.  $\sigma(\mathbf{x}) \xrightarrow{|\mathbf{x}| \rightarrow \infty} \sigma_0$ ,  $\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ , et si  $(f, g) \in L^\infty(C_{b,d}^\gamma) \times C_{b,d}^{2+\gamma}$  alors il existe une unique solution mild et faible  $u \in L^\infty(C_{b,d}^{2+\gamma})$  à (1.2) et il existe une constante  $C_1 > 0$  telle que*

$$\|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq C_1(\|f\|_{L^\infty(C_{b,d}^\gamma)} + \|g\|_{C_{b,d}^{2+\gamma}}) \quad (1.9)$$

Précisons que nous appelons solution mild de (1.2) au sens de Stroock et Varadhan [SV79] (voir aussi e.g. [Kol11] ) toute solution  $u$  pour laquelle la limite de la régularisation  $(\partial_t + L_t^m)u_m = -f_m$  quand  $m \rightarrow +\infty$  existe et est égale à  $f$  mais dont les dérivées partielles de  $u$  intervenant dans  $(\partial_t + L_t)u$  ne sont pas forcément définies. Notons que dans l'article de [Lun97], la notion de solution forte correspond à notre version de solution *mild*. La solution faible correspond à la solution au sens des distributions, i.e. que pour toute fonction  $\varphi \in C_0^\infty$  (fonction de classe  $C^\infty$  à support compact), on a

$$\int \varphi(x)(\partial_t + L_t)\varphi(x)dx = \int f(x)\varphi(x)dx.$$

Mentionnons l'article de Priola [Pri09] qui établit également des estimées de Schauder comme pour le Théorème 1 avec le même critère asymptotique sur  $\sigma$  et la dérive s'écrivant

$$\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \begin{pmatrix} \tilde{\mathbf{F}}_1(\mathbf{x}) \\ \mathbf{0}_{(n-1)d,d} \end{pmatrix},$$

et  $\tilde{\mathbf{F}}_1$  est  $C_d^\gamma$ . Intuitivement, on utilise la formule de Girsanov avec une version régularisée de  $\tilde{\mathbf{F}}_1$  pour se ramener au cas considéré par Lunardi. La régularité de  $\tilde{\mathbf{F}}_1$  intervient lorsqu'on dérive le flot dans martingales exponentielles. Enfin on injecte l'estimée de Schauder avec comme source la différence la version régularisée et non-régularisée de  $\tilde{\mathbf{F}}_1$ .

Dans la suite, nous établirons des estimés de Schauder pour des solutions mild et faibles. Nous établirons des contrôles de la solution classique dans le cadre de l'unicité forte. Pour plus de détails sur la solution au sens classique de l'équation de Kolmogorov, nous renvoyons par exemple à [KLL08], [Lun95].

Dans la section suivante, nous présentons des estimées de Schauder dans un cadre plus général que celui considéré par [Pri09].

## 1.5 Estimées de Schauder établies

### Hypothèses

Tout d'abord, présentons le cadre d'hypothèses qui nous permet d'avoir le pendant du Théorème 1 dans le cas d'une dérive plus générale présenté dans (1.2).

**(UE) Uniforme Ellipticité du coefficient de diffusion.** Il existe  $\kappa \geq 1$  tel que pour tout  $(t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^{nd}$ ,  $z \in \mathbb{R}^d$ ,  $\kappa^{-1}|z|^2 \leq \langle a(t, \mathbf{x})z, z \rangle \leq \kappa|z|^2$ .

**(H) Condition du type Hörmander faible.** Pour tout  $i \in \llbracket 2, n \rrbracket$ , et  $(t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^{nd}$ ,  $D_{\mathbf{x}_{i-1}} \mathbf{F}_i(t, \mathbf{x}_{i-1}, \dots, \mathbf{x}_n)$  est une matrice de  $\mathbb{R}^d \otimes \mathbb{R}^d$  inversible.

**(S) Régularité des coefficients.** Soit  $\gamma \in (0, 1)$ .

- **Régularité du coefficient de diffusion.** On suppose que  $a \in L^\infty(C_{b,d}^\gamma)$ .

- **Régularité de la dérive.** On suppose que  $\mathbf{F}$  est mesurable et borné en temps.

• On suppose que  $\mathbf{F}_1 \in L^\infty(C_{\mathbf{d}}^\gamma)$ .

• Pour tout  $i \in \llbracket 2, n \rrbracket$ , on suppose que  $\mathbf{F}_i \in L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})$ . En particulier  $\mathbf{F}_i(t, \cdot, \mathbf{x}^{i:n}) \in C^{1+\frac{\gamma}{2(i-1)-1}}$ .

Quand on fait tendre  $\gamma$  vers 0, on retrouve les exposants de régularité critiques présents dans [CdRM17] pour obtenir l'unicité faible.

## Résultat principal

Sous ces conditions de régularités minimales, nous établissons des Estimées Schauder.

**Theorem 2 (Estimées de Schauder.)** *Soit  $\gamma \in (0, 1)$  donné, on suppose (UE), (H) et (S), si  $(f, g) \in L^\infty(C_{b,d}^\gamma) \times C_{b,d}^{2+\gamma}$  alors il existe une unique solution mild et faible  $u \in L^\infty(C_{b,d}^{2+\gamma})$  à l'équation (1.2), et il existe une constante  $C_2 := C_2((\mathbf{UE}), (\mathbf{H}), (\mathbf{S}), T)$  telle que*

$$\|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq C_2 (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)}). \quad (1.10)$$

## 1.6 Grandes lignes de l'analyse

Rappelons que nous avons régularisé l'équation (1.2). Il nous faut donc de (1.12), contrôler  $\|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}$  uniformément par rapport au coefficient de régularisation  $m$ . À cette fin, nous utilisons une méthode de type perturbative en développant la solution  $u$  autour d'un proxy  $\tilde{\mathbf{X}}_t$  de type Ornstein-Uhlenbeck.

### Méthode de type perturbative

Soit le générateur infinitésimal  $\tilde{L}$  associé au proxy  $\tilde{\mathbf{X}}_t$  qui a pour densité  $\tilde{p}(t, s, \mathbf{x}, \mathbf{y})$ . On réécrit le problème de Cauchy (1.2) :

$$\begin{cases} \partial_t u(t, \mathbf{x}) + \tilde{L}u = -f(t, \mathbf{x}) - (L - \tilde{L})u(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \\ u(T, \mathbf{x}) = g(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^{nd}. \end{cases} \quad (1.11)$$

Maintenant, nous régularisons (1.2) pour avoir un problème de Cauchy *bien posé*.

$$\begin{cases} \partial_t u_m(t, \mathbf{x}) + \langle \mathbf{F}^m(t, \mathbf{x}), \mathbf{D}u_m(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}) a_m(t, \mathbf{x})) = -f_m(t, \mathbf{x}), \\ u_m(T, \mathbf{x}) = g_m(\mathbf{x}), & \mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{nd}, \end{cases} \quad (1.12)$$

avec  $u_m, \mathbf{F}_m, f_m, g_m \xrightarrow{m \rightarrow +\infty} u, \mathbf{F}, f, g$ . En effet, avec une méthode de dérivation du flot à la Kunita (voir la partie précédente et e.g. [Kun97]), on obtient la représentation de Feynman-Kac :

$$u_m(t, \mathbf{x}) := \mathbb{E}[g_m(\mathbf{X}_T^m) | \mathbf{X}_t = \mathbf{x}] + \int_t^T \mathbb{E}[f_m(s, \mathbf{X}_s^m) | \mathbf{X}_t = \mathbf{x}] ds, \quad (1.13)$$

où  $\sigma_m$  est une racine carrée  $a_m$  et

$$\begin{aligned} d\mathbf{X}_t^{m,1} &= \mathbf{F}_1^m(t, \mathbf{X}_t^{m,1}, \dots, \mathbf{X}_t^{m,n}) dt + \sigma(t, \mathbf{X}_t^{m,1}, \dots, \mathbf{X}_t^{m,n}) dW_t, \\ d\mathbf{X}_t^{m,2} &= \mathbf{F}_2^m(t, \mathbf{X}_t^{m,1}, \dots, \mathbf{X}_t^{m,n}) dt, \\ d\mathbf{X}_t^m &= d\mathbf{X}_t^{m,3} = \mathbf{F}_3^m(t, \mathbf{X}_t^{m,2}, \dots, \mathbf{X}_t^{m,n}) dt, \\ &\vdots \\ d\mathbf{X}_t^{m,n} &= \mathbf{F}_n^m(t, \mathbf{X}_t^{m,n-1}, \mathbf{X}_t^{m,n}) dt. \end{aligned} \quad (1.14)$$

Dans la suite de cette introduction, par souci de simplicité, nous omettrons l'indice de régularisation  $m$ . On utilise dès lors la formule de Duhamel, ce qui correspond au développement paramétrix du premier ordre :

$$u(t, \mathbf{x}) = \tilde{P}_{T,t} g(\mathbf{x}) + \tilde{G} f(t, \mathbf{x}) + \int_t^T ds \int_{\mathbb{R}^{nd}} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) (L - \tilde{L}) u(s, \mathbf{y}) d\mathbf{y}, \quad (1.15)$$

où nous définissons le semi-groupe :

$$\tilde{P}_{T,t} g(\mathbf{x}) := \int_{\mathbb{R}^{nd}} \tilde{p}(t, T, \mathbf{x}, \mathbf{y}) g(\mathbf{y}) d\mathbf{y},$$

et le noyau de Green associés au proxy :

$$\tilde{G} f(t, \mathbf{x}) := \int_t^T ds \tilde{P}_{s,t} f(s, \mathbf{x}).$$

Il reste à savoir quel proxy utiliser pour que le dernier terme de droite dans (1.15),  $\int_t^T ds \int_{\mathbb{R}^{nd}} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) (L - \tilde{L}) u(s, \mathbf{y}) d\mathbf{y}$ , soit vu comme un reste. Dans le paragraphe suivant, nous expliciterons le choix de linéarisation Gaussien.

### Linéarisation du proxy

L'idée est de linéariser l'EDS (1.14), pour cela considérons le flot, solution de l'EDO associée à la dérive :

$$\dot{\boldsymbol{\theta}}_{v,\tau}(\boldsymbol{\xi}) = \mathbf{F}(v, \boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi})), \quad v \in [\tau, T], \quad \boldsymbol{\theta}_{\tau,\tau}(\boldsymbol{\xi}) = \boldsymbol{\xi}, \quad (1.16)$$

où  $(\tau, \boldsymbol{\xi}) \in [0, T] \times \mathbb{R}^{nd}$  sont *des paramètres de gel*, respectivement en temps et en espace qui seront explicités plus loin. Cette ODE admet une unique solution, comme nous avons régularisé  $\mathbf{F} = \mathbf{F}^m$  est Lipschitz.

Nous avons choisi une approche dite *forward*, dans le sens où le flot (1.16) part de  $\boldsymbol{\xi}$  (qui sera choisi égal à  $\mathbf{x}$ ) pour aller vers  $\mathbf{y}$  (variable d'intégration en (1.15)). Cette technique



a notamment été utilisée par Chaudru de Raynald [CdR17] pour montrer l'unicité forte dans le cas où  $n = 2$ . Il existe aussi une approche dite *backward* introduite par [MS67], et exploitée par la suite par [DM10], [Men18]. Cette méthode peut sembler naturelle pour écrire formellement si  $g = 0$ ,  $u = \tilde{G} \circ (I - R)^{-1}f$  où  $R = \tilde{G}(L - \tilde{L})$  et où l'inverse de l'opérateur  $(I - R)$  est pris dans l'espace de Hölder approprié. Cette inversion d'opérateur est également la technique employée dans [CdRM17] et [SV79]. Cette approche permet d'estimer la densité, néanmoins des difficultés apparaissent lorsqu'il faut contrôler ses dérivées. En effet, dans le cas *backward*  $\tilde{p}(t, T, \mathbf{x}, \mathbf{y}) = \tilde{p}^{\mathbf{y}}(t, T, \mathbf{x}, \mathbf{y})$ , où  $\tilde{p}^{\mathbf{y}}$  exprime que le paramètre de gel dépend de la variable d'intégration  $\mathbf{y}$ , n'est pas une densité de probabilité. En particulier, il n'est pas évident que  $\tilde{P}_{T,t}g$  conserve la même régularité de  $g$ .

Maintenant, écrivons le processus gelé autour du flot défini en (1.16) pour tout  $v \in [t, s]$

$$\begin{aligned} d\tilde{\mathbf{X}}_v^{(\tau, \xi)} &= [\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\xi)) + D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\xi))(\tilde{\mathbf{X}}_v^{(\tau, \xi)} - \boldsymbol{\theta}_{v, \tau}(\xi))]dv + B\sigma(v, \boldsymbol{\theta}_{v, \tau}(\xi))dW_v, \\ \tilde{\mathbf{X}}_t^{(\tau, \xi)} &= \mathbf{x}, \end{aligned} \tag{1.17}$$

avec  $B = \begin{pmatrix} I_d & \mathbf{0}_{(n-1)d, d} \\ \mathbf{0}_{d, d} & \mathbf{0}_{(n-1)d, d} \end{pmatrix}$  et pour tout  $\mathbf{z} \in \mathbb{R}^{nd}$  :

$$D\mathbf{F}(v, \mathbf{z}) = \begin{pmatrix} \mathbf{0}_{d, d} & \cdots & \cdots & \mathbf{0}_{d, d} \\ D_{\mathbf{z}_1}\mathbf{F}_2(v, \mathbf{z}) & \mathbf{0}_{d, d} & \cdots & \mathbf{0}_{d, d} \\ \vdots & \ddots & \cdots & \vdots \\ \mathbf{0}_{d, d} & \cdots & D_{\mathbf{z}_{n-1}}\mathbf{F}_n(v, \mathbf{z}_{n-1}, \mathbf{z}_n) & \mathbf{0}_{d, d} \end{pmatrix}.$$

Notons que  $D\mathbf{F}$  est du même type que  $\mathbf{A}_0$ , avec pour tout  $i \in \llbracket 1, n \rrbracket$ ,  $D_{\mathbf{x}_{i-1}}\mathbf{F}_i$  est une matrice  $d \times d$  inversible, il faut donc s'attendre à obtenir la même propriété de «bonne échelle».

### Propriété du proxy

Nous pouvons écrire explicitement la dynamique du proxy.

$$\begin{aligned} \tilde{\mathbf{X}}_v^{(\tau, \xi)} &= \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t)\mathbf{x} + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) \left( \mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\xi)) - D\mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\xi))\boldsymbol{\theta}_{u, \tau}(\xi) \right) du \\ &\quad + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\xi)) dW_u \\ &=: \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\xi)) dW_u, \end{aligned}$$

où  $\tilde{\mathbf{R}}^{(\tau, \xi)}(v, u)$  est la résolvante associée à  $D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\xi))$ , c'est-à-dire :

$$\partial_v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t) = D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\xi))\tilde{\mathbf{R}}^{(\tau, \xi)}(v, t), \quad \tilde{\mathbf{R}}^{(\tau, \xi)}(t, t) = \mathbf{I}_{nd \times nd},$$

et  $\mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x})$  correspond à la moyenne de  $\tilde{\mathbf{X}}_v^{(\tau, \xi)}$  qui est affine au point de départ  $\mathbf{x}$ .

Remarquons que finalement ce choix de proxy est naturel pour l'étude du terme de reste de (1.15). En effet, le flot,  $\boldsymbol{\theta}_{v, \tau}(\xi)$  rend compte du transport de  $\xi$  (qui sera pris égal à  $\mathbf{x}$ ) à  $\mathbf{y}$ .

En particulier,  $\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})}$  est un processus Gaussien dont la matrice de covariance est :

$$\tilde{\mathbf{K}}_{v,t}^{(\tau, \boldsymbol{\xi})} := \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) Ba(u, \boldsymbol{\theta}_{u,\tau}(\boldsymbol{\xi})) B^* \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u)^* du.$$

La covariance satisfait la «bonne échelle»  $\tilde{\mathbf{K}}_{v,t}^{(\tau, \boldsymbol{\xi})} \asymp (v-t)^{-1} \mathbb{T}_{v-t}^2$  où  $\mathbb{T}_u$  est définie en (1.7).

Le processus gelé,  $\tilde{\mathbf{X}}_s^{(\tau, \boldsymbol{\xi})}$  a une densité Gaussienne (multi-échelle) qui s'écrit :

$$\begin{aligned} & \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \\ & := \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s,t}^{(\tau, \boldsymbol{\xi})})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{s,t}^{(\tau, \boldsymbol{\xi})})^{-1} (\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) - \mathbf{y}), \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) - \mathbf{y} \right\rangle\right). \end{aligned}$$

Nous montrons qu'il existe  $C > 0$  tel que pour tout multi-indice  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ ,  $|\alpha| \leq 3$ , avec la dérivée multi-indice  $D_{\mathbf{x}}^\alpha := D_{\mathbf{x}_1}^{\alpha_1} \dots D_{\mathbf{x}_n}^{\alpha_n}$ , telle que :

$$\begin{aligned} & |D_{\mathbf{x}}^\alpha \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})| \\ & \leq \frac{C}{(s-t)^{\sum_{i=1}^n \alpha_i (i-\frac{1}{2}) + \frac{n^2 d}{2}}} \exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) - \mathbf{y})|^2\right) \\ & =: \frac{C}{(s-t)^{\sum_{i=1}^n \alpha_i (i-1/2)}} \tilde{p}_{C^{-1}}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned} \tag{1.18}$$

Autrement dit, dériver la densité gelée induit des singularités en temps correspondant aux échelles intrinsèques des indices de dérivation. Le travail qui suit consiste principalement à rattraper ces singularités temporelles par la régularité spatiale.

## 1.7 Contrôle de la norme uniforme

### Contrôle de la norme uniforme des premiers termes

Après de potentielles dérivées suivant  $\mathbf{x}$ , nous choisissons le point de gel  $(\tau, \boldsymbol{\xi}) = (t, \mathbf{x})$ , ainsi ce choix ne change pas les dérivées intervenant dans (1.15). En particulier, nous pouvons voir grâce à la dynamique (1.17) qu'avec ce choix de gel  $\mathbf{m}_{v,t}^{(t, \mathbf{x})}(\mathbf{x}) = \boldsymbol{\theta}_{v,t}(\boldsymbol{\xi})$ .

Le contrôle des normes associées au semi-groupe  $\tilde{P}_{T,t}^{(\tau, \boldsymbol{\xi})} : \|\tilde{P}_{T,t}^{(\tau, \boldsymbol{\xi})} g\|_{L^\infty}$ ,  $\|D_{\mathbf{x}_1} \tilde{P}_{T,t}^{(\tau, \boldsymbol{\xi})} g\|_{L^\infty}$ ,  $\|D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{(\tau, \boldsymbol{\xi})} g\|_{L^\infty}$ , et associées au noyau de Green  $\tilde{G}^{(\tau, \boldsymbol{\xi})} : \|\tilde{G}^{(\tau, \boldsymbol{\xi})} f\|_{L^\infty}$  et  $\|D_{\mathbf{x}_1} \tilde{G}^{(\tau, \boldsymbol{\xi})} f\|_{L^\infty}$ ,  $\|D_{\mathbf{x}_1}^2 \tilde{G}^{(\tau, \boldsymbol{\xi})} f\|_{L^\infty}$  est établi par des techniques de type cancellation et en utilisant la régularité des paramètres intervenant dans le proxy (1.18). Pour illustrer le principe de cancellation, établissons un contrôle en norme uniforme du noyau de Green :

$$\begin{aligned} & \|D_{\mathbf{x}_1}^2 \tilde{G}^{(\tau, \boldsymbol{\xi})} f\|_{L^\infty} \\ & = \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \left| \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) d\mathbf{y} \right|_{(\tau, \boldsymbol{\xi}) = (t, \mathbf{x})} \\ & = \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \left| \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) [f(s, \mathbf{y}) - f(s, \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))] d\mathbf{y} \right|_{(\tau, \boldsymbol{\xi}) = (t, \mathbf{x})}. \end{aligned}$$

La deuxième égalité vient du fait que  $\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} = 0$ , c'est ce qu'on appelle en effet le principe de cancellation. Or comme  $f \in C_{\mathbf{d}}^\gamma(\mathbb{R}^{nd}, \mathbb{R})$  et par (1.18), on déduit :

$$\begin{aligned}
 & \|D_{\mathbf{x}_1}^2 \tilde{G}^{(\tau, \boldsymbol{\xi})} f\|_{L^\infty} \\
 & \leq C[f]_{\mathbf{d}}^\gamma \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \left| \int_t^T ds \int_{\mathbb{R}^{nd}} \frac{\tilde{p}_{C^{-1}}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})}{s-t} \mathbf{d}^\gamma(\mathbf{y}, \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})) d\mathbf{y} \right|_{(\tau, \boldsymbol{\xi})=(t, \mathbf{x})} \\
 & \leq C[f]_{\mathbf{d}}^\gamma \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \left| \int_t^T ds \int_{\mathbb{R}^{nd}} \frac{\tilde{p}_{C^{-1}}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{1-\frac{\gamma}{2}}} d\mathbf{y} \right|_{(\tau, \boldsymbol{\xi})=(t, \mathbf{x})} \\
 & = C[f]_{\mathbf{d}}^\gamma (T-t)^{\frac{\gamma}{2}}. \tag{1.19}
 \end{aligned}$$

La dernière inégalité vient du fait que pour tout  $\zeta \in \mathbb{R}_+$  il existe  $C_\zeta \geq 1$  tel que pour tout  $0 \leq \beta \leq \zeta$  et  $x \in \mathbb{R}_+$  :

$$x^\beta \exp(-x^2) \leq C_\zeta \exp(-C_\zeta^{-1} x^2).$$

Nous contrôlons de manière similaire les premiers termes de reste  $\int_t^T ds \int_{\mathbb{R}^{nd}} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) (L - \tilde{L})u(s, \mathbf{y}) d\mathbf{y}$  intervenant dans (1.15), c'est-à-dire les termes faisant intervenir les dérivées suivant les composantes non-dégénérées :

$$\begin{aligned}
 & \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \left( \langle \mathbf{F}_1(s, \mathbf{y}) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,\tau}(\boldsymbol{\xi})), D_{\mathbf{y}_1} u(s, \mathbf{y}) \rangle \right. \\
 & \quad \left. + \frac{1}{2} \left( (a(s, \mathbf{y}) - a(s, \boldsymbol{\theta}_{s,\tau}(\boldsymbol{\xi}))) D_{\mathbf{y}_1}^2 u(s, \mathbf{y}) \right) \right) d\mathbf{y}, \tag{1.20}
 \end{aligned}$$

se contrôle en norme uniforme.

### Contrôle de la norme uniforme des derniers termes

Néanmoins, pour les derniers coefficients du terme de reste dans (1.15) faisant intervenir les dérivées suivant les variables dégénérées, i.e.  $(D_{\mathbf{y}_i} u(s, \mathbf{y}))_{i \in \llbracket 2, n \rrbracket}$ , le contrôle en norme uniforme est plus délicat. En effet, la version non régularisée de  $u$  n'est *a priori* pas dérivable suivant les variables dégénérées  $(\mathbf{x}_i)_{i \in \llbracket 2, n \rrbracket}$ , mais on s'attend des résultats précédents à ce qu'elle soit seulement  $(2 + \gamma)/(2i - 1)$ -Hölder uniformément en  $m$ . Pour cela, définissons :

$$\Delta_{i, \mathbf{F}}(\tau, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}) := \left( \mathbf{F}_i(s, \mathbf{y}) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,\tau}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,\tau}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,\tau}(\boldsymbol{\xi}))_{i-1} \right), \tag{1.21}$$

et par intégrations par parties on a :

$$\begin{aligned}
 & \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i, \mathbf{F}}(\tau, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \rangle \\
 & = - \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{y}_i} \cdot [D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \Delta_{i, \mathbf{F}}(\tau, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y})] u(s, \mathbf{y}). \tag{1.22}
 \end{aligned}$$

L'idée, ensuite, sera de contrôler  $D_{\mathbf{y}_i} [D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \Delta_{i, \mathbf{F}}(\tau, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y})]$  suivant la norme adaptée en dualité à la norme  $L^\infty(C_{b, \mathbf{d}}^{2+\gamma})$  de  $u(s, \mathbf{y})$ . C'est *a priori* un contrôle non direct étant donné que le terme  $D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \Delta_{i, \mathbf{F}}(\tau, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y})$  dont on dérive en  $D_{\mathbf{y}_i}$  n'est pas dérivable en  $\mathbf{y}_i$  à cause de l'irrégularité des coefficients.

Pour se faire, remarquons que pour  $\tilde{\alpha} \in \mathbb{R}$ ,  $C_b^{\tilde{\alpha}}(\mathbb{R}^d, \mathbb{R}) = B_{\infty, \infty}^{\tilde{\alpha}}(\mathbb{R}^d, \mathbb{R})$ , où pour tout  $(p, q, s) \in [1, +\infty]^2 \times \mathbb{R}$ ,  $B_{p, q}^s$  est un espace de Besov. Il existe plusieurs façons de définir les espaces de Besov (module de continuité, décomposition de Littlewood-Paley, représentation thermique). La représentation thermique est la formulation la plus naturelle dans notre contexte Gaussien. On définit ainsi  $B_{p, q}^s(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{\mathcal{H}_{p, q}^s} < +\infty\}$  pour tout  $m \in \mathbb{N}_0$  tel que  $m > \frac{s}{2}$  :

$$\|f\|_{\mathcal{H}_{p, q}^{\tilde{\alpha}}} := \|\varphi(D)f\|_{L^p(\mathbb{R}^d)} + \left( \int_0^1 v^{(m - \frac{\tilde{\alpha}}{2})q} \|\partial_v^m h_v \star f\|_{L^p(\mathbb{R}^d)}^q \frac{dv}{v} \right)^{\frac{1}{q}}, \quad (1.23)$$

où  $\varphi \in C_0^\infty(\mathbb{R}^d)$  telle que  $\varphi(0) \neq 0$ ,  $\varphi(D)f := (\varphi \hat{f})^\vee$  où on définit la transformée de Fourier  $\hat{f}$  et la transformée de Fourier inverse  $(\varphi \hat{f})^\vee$ , et pour tout  $v > 0$ ,  $z \in \mathbb{R}^d$ ,

$$h_v(z) := \frac{1}{(2\pi v)^{\frac{d}{2}}} \exp\left(-\frac{|z|^2}{2v}\right)$$

est le noyau de la chaleur usuel. Pour plus d'informations nous pouvons nous référer à Triebel [Tri83].

La propriété qui nous intéresse dans les espaces de Besov est en fait la relation de dualité entre  $B_{\infty, \infty}^{\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R})$  et  $B_{1, 1}^{-\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R})$  (voir par exemple [LR02]). Ainsi par des méthodes de cancellation, nous établissons la norme Besov de chaque terme de (1.22). Autrement dit, pour tout  $i \in \llbracket 2, n \rrbracket$ , on montre que la norme Besov prise pour la  $i$ -ème variable :

$$\begin{aligned} & \left\| \mathbf{y}_i \mapsto D_{\mathbf{y}_i} (D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \Delta_{i, \mathbf{F}}(\tau, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y})) \right\|_{B_{1, 1}^{-\frac{2+\gamma}{2i-1}}(\mathbb{R}^d)} \\ & \leq C(s-t)^{\frac{\gamma}{2}-1} \int_{\mathbb{R}^d} \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y}_i. \end{aligned}$$

En effet, la caractérisation thermique permet de faire passer la dérivée  $\mapsto D_{\mathbf{y}_i}$  sur le noyau de la chaleur  $h_v$ .

### Contrôle des modules Hölder

Pour étudier les différents modules Hölder, pour  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$  donné et  $0 < c_0 < 1$  supposé «petit», nous procédons classiquement à une distinction entre le régime *hors-diagonal* :  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') > (s-t)$  et le régime *diagonal* :  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \leq (s-t)$ .

- *hors-diagonal*, dans ce cas, les points sont loin par rapport au temps courant. Donc il n'y a pas d'intérêt à utiliser des développements des termes à contrôler entre ces deux points  $\mathbf{x}, \mathbf{x}'$ . Nous utilisons ainsi le contrôle en norme sup précédemment établi avec comme

choix de points de gel  $(\tau, \boldsymbol{\xi}, \boldsymbol{\xi}') = (t, \mathbf{x}, \mathbf{x}')$ . Par exemple, nous écrivons pour le contrôle Hölder du noyau de Green :

$$\begin{aligned}
 & \left| \tilde{G}^{m, \boldsymbol{\xi}} f(t, \mathbf{x}) - \tilde{G}^{m, \boldsymbol{\xi}} f(t, \mathbf{x}') \right|_{(\tau, \boldsymbol{\xi}, \boldsymbol{\xi}') = (t, \mathbf{x}, \mathbf{x}')} \\
 = & \left| \int_t^{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')} ds \int_{\mathbb{R}^{nd}} [D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi}')} (t, s, \mathbf{x}', \mathbf{y})] f(s, \mathbf{y}) d\mathbf{y} \right|_{(\tau, \boldsymbol{\xi}, \boldsymbol{\xi}') = (t, \mathbf{x}, \mathbf{x}')} \\
 \leq & 2 \sup_{(\tau, \boldsymbol{\xi}, \boldsymbol{\xi}') = (t, \mathbf{x}, \mathbf{x}')} \left| \int_t^{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) f(s, \mathbf{y}) \right| \\
 \leq & C \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{1.24}
 \end{aligned}$$

La première inégalité est une conséquence de (1.19).

• *diagonal*, dans le cas où  $\mathbf{x}, \mathbf{x}'$  sont proches vis-à-vis du temps courant, nous développons la densité gelée entre  $\mathbf{x}$  et  $\mathbf{x}'$ . Pour se faire, nous choisissons les points de gel  $(\tau, \boldsymbol{\xi}, \boldsymbol{\xi}') = (t, \mathbf{x}, \mathbf{x})$ , de sorte que les densités à contrôler soient les mêmes.

Autrement dit on écrit :

$$\begin{aligned}
 & \left| \int_{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} [D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi}')} (t, s, \mathbf{x}', \mathbf{y})] f(s, \mathbf{y}) \right| \\
 \leq & \left| \int_{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\lambda (\mathbf{x} - \mathbf{x}') \mathbf{D}_x D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}' + \lambda(\mathbf{x} - \mathbf{x}'), \mathbf{y}) f(s, \mathbf{y}) \right| \\
 \leq & C \|f\|_{L^\infty(C_{b, \mathbf{d}}^\gamma)} \sum_{i=1}^n |\mathbf{x}_i - \mathbf{x}'_i| \int_{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (s-t)^{-1-(i-\frac{1}{2})} \tilde{p}_{C^{-1}}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \\
 \leq & C \|f\|_{L^\infty(C_{b, \mathbf{d}}^\gamma)} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{1.25}
 \end{aligned}$$

L'avant dernière inégalité est une conséquence du contrôle des dérivées de la densité gelée (1.18) et de la cancellation  $\int_{\mathbb{R}^{nd}} d\mathbf{y} \mathbf{D}_x D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}' + \lambda(\mathbf{x} - \mathbf{x}'), \mathbf{y}) f(s, \boldsymbol{\theta}_{s,t}(\mathbf{x})) = \mathbf{0}_{d,d,d}$ , la dernière identité vient du fait que  $|\mathbf{x}_i - \mathbf{x}'_i| \leq \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}')$ .

Cependant, pour un tel choix de paramètres de gel  $(\tau, \boldsymbol{\xi}, \boldsymbol{\xi}')$  dépendant de la variable d'intégration  $s$ , il nous faut modifier par la formule de Duhamel l'équation (1.15). Autrement dit pour tout  $(t, \mathbf{x}') \in [0, T] \times \mathbb{R}^{nd}$  et  $r \in (t, T]$ ,  $\boldsymbol{\xi}' \in \mathbb{R}^{nd}$  on a :

$$\begin{aligned}
 u_m(t, \mathbf{x}') &= \tilde{P}_{r,t}^{m, (\tau, \boldsymbol{\xi}')} u_m(r, \mathbf{x}') + \tilde{G}_{r,t}^{m, (\tau, \boldsymbol{\xi}')} f_m(t, \mathbf{x}') \\
 &+ \int_t^r ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m, (\tau, \boldsymbol{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m, (\tau, \boldsymbol{\xi}')} ) u_m(s, \mathbf{y}),
 \end{aligned}$$

avec

$$\forall 0 \leq v < r \leq T, \quad \tilde{G}_{r,v}^{m, (\tau, \boldsymbol{\xi}')} f_m(t, \mathbf{x}) = \int_v^r ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m, (\tau, \boldsymbol{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) f_m(s, \mathbf{y}). \tag{1.26}$$

Dériver cette expression suivant  $r$ , engendre pour tout  $\boldsymbol{\xi}' \in \mathbb{R}^{nd}$  :

$$\begin{aligned}
 0 &= \partial_r \tilde{P}_{r,t}^{m, (\tau, \boldsymbol{\xi}')} u_m(r, \mathbf{x}') + \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m, (\tau, \boldsymbol{\xi}')} (t, r, \mathbf{x}', \mathbf{y}) f_m(r, \mathbf{y}) \\
 &+ \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m, (\tau, \boldsymbol{\xi}')} (t, r, \mathbf{x}', \mathbf{y}) (L_r^m - \tilde{L}_r^{m, (\tau, \boldsymbol{\xi}')} ) u_m(r, \mathbf{y}) d\mathbf{y}. \tag{1.27}
 \end{aligned}$$

Ensuite, nous intégrons cette équation entre  $t$  et  $t_0 \in (t, T]$  pour un premier  $\xi'$  et entre  $t_0$  et  $T$  avec un  $\tilde{\xi}'$  éventuellement différent :

$$\begin{aligned}
 0 &= \tilde{P}_{t_0,t}^{m,(\tau,\xi')} u_m(t_0, \mathbf{x}') - u_m(t, \mathbf{x}') + \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y}) f_m(s, \mathbf{y}) \\
 &+ \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\xi')}) u_m(s, \mathbf{y}) \\
 &\tilde{P}_{T,t}^{m,(\tau,\tilde{\xi}')} u_m(T, \mathbf{x}') - \tilde{P}_{t_0,t}^{m,(\tau,\tilde{\xi}')} u_m(t_0, \mathbf{x}') + \int_{t_0}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\tilde{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) f_m(s, \mathbf{y}) \\
 &+ \int_{t_0}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau,\tilde{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\tilde{\xi}')} ) u_m(s, \mathbf{y}).
 \end{aligned}$$

Rappelons que  $u_m(T, \mathbf{x}') = g_m(\mathbf{x}')$  (condition terminale), et avec les notations de (1.26) l'équation précédente se réécrit :

$$\begin{aligned}
 u_m(t, \mathbf{x}') &= \tilde{P}_{T,t}^{m,(\tau,\tilde{\xi}')} g_m(\mathbf{x}') + \tilde{G}_{t_0,t}^{m,(\tau,\xi')} f_m(t, \mathbf{x}') + \tilde{G}_{T,t_0}^{m,(\tau,\tilde{\xi}')} f_m(t, \mathbf{x}') \\
 &+ \tilde{P}_{t_0,t}^{m,(\tau,\xi')} u_m(t_0, \mathbf{x}') - \tilde{P}_{t_0,t}^{m,(\tau,\tilde{\xi}')} u_m(t_0, \mathbf{x}') \\
 &+ \int_t^{T} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left( \mathbb{I}_{s \leq t_0} \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\xi')}) \right. \\
 &\left. + \mathbb{I}_{s > t_0} \tilde{p}^{m,(\tau,\tilde{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\tilde{\xi}')} ) \right) u_m(s, \mathbf{y}). \tag{1.28}
 \end{aligned}$$

Nous voyons dès lors que pour  $\xi' \neq \tilde{\xi}'$  apparaît un nouveau terme dû à la discontinuité du changement de point de gel. Notons toutefois, que le rôle de  $t_0$  est semblable à celui de  $(\xi, \xi')$ , dans le sens où après de potentielles dérivations en espace nous posons

$$t_0 = t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}'). \tag{1.29}$$

Autrement dit, le temps  $t_0$  correspond exactement au temps critique de changement de régime. Le terme de discontinuité  $\tilde{P}_{t_0,t}^{(\tau,\xi')} u(t_0, \mathbf{x}') - \tilde{P}_{t_0,t}^{(\tau,\tilde{\xi}')} u_m(t_0, \mathbf{x}')$  se traite identiquement à ceux précédemment contrôlés.

## 1.8 Conclusion

A ce stade, nous pouvons énoncer le contrôle de la solution de (1.15) uniformément à l'indice de régularisation  $m$  :

$$\|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq C \left( \|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)} + \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \left[ 1 + c_0^{\frac{\gamma}{2}} + T^{\frac{\gamma}{2}} \right] \right). \tag{1.30}$$

La constante  $C > 0$  dépend des normes Hölder de  $a$ , et de  $\mathbf{F}$ . Pour des normes Hölder de  $a$  et de  $\mathbf{F}$ , pour  $c_0$  et pour  $T > 0$  suffisamment petits, on conclut facilement l'estimée de Schauder. Pour des coefficients  $a$  et  $\mathbf{F}$  quelconques, nous procédons par une méthode de *scaling*. C'est-à-dire pour tout  $\lambda > 0$  petit, on écrit le problème de Cauchy équivalent à

(1.2) :

$$\begin{cases} \partial_t u^\lambda(t, \mathbf{x}) + \langle \mathbf{F}(t, \lambda^{-1/2} \mathbb{T}_\lambda \mathbf{x}), \lambda^{1/2} \mathbb{T}_\lambda^{-1} \mathbf{D} u^\lambda(t, \mathbf{x}) \rangle \\ + \frac{\lambda^{-1}}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u^\lambda(t, \mathbf{x}) a(t, \lambda^{-1/2} \mathbb{T}_\lambda \mathbf{x})) = -f(t, \lambda^{-1/2} \mathbb{T}_\lambda \mathbf{x}), \\ u^\lambda(T, \mathbf{x}) = g(\lambda^{-1/2} \mathbb{T}_\lambda \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^{nd}. \end{cases} \quad (1.31)$$

En reprenant les contrôles précédents on obtient dès lors :

$$\|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq C \left( \|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)} + \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} [\lambda^{\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}} + T^{\frac{\gamma}{2}}] \right). \quad (1.32)$$

Pour  $T$ ,  $c_0$  et  $\lambda$  suffisamment petits devant les normes Hölder de  $a$  et de  $\mathbf{F}$ , on obtient  $C[\lambda^{\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}} + T^{\frac{\gamma}{2}}] < 1$ . Nous avons donc le contrôle Schauder voulu en passant le terme  $C\|u\|_{L^\infty(C_{b,d}^{2+\gamma})}[\lambda^{\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}} + T^{\frac{\gamma}{2}}]$  à gauche de l'inégalité (1.32). Pour passer à temps quelconque, nous itérons le raisonnement précédent en remplaçant  $g(\mathbf{x})$  par  $u(T_n, \mathbf{x})$  pour  $T_n = \frac{N-k+1}{N}T$ , avec  $N$  suffisamment grand,  $k \in \llbracket 0, N \rrbracket$ . Autrement dit, on écrit pour  $t \in [0, (1 - \frac{k-1}{N})T)$  :

$$\begin{cases} \partial_t u_k(t, \mathbf{x}) + \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D} u_k(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u_k(t, \mathbf{x}) a(t, \mathbf{x})) = -f(t, \mathbf{x}), \\ u_k((1 - \frac{k-1}{N})T, \mathbf{x}) = u_{k-1}((1 - \frac{k-1}{N})T, \mathbf{x}). \end{cases}$$

Enfin, par un argument de compacité (le Théorème d'Ascoli) et l'unicité au problème de la martingale associé établie par [CdRM17], on obtient qu'il existe une unique solution mild vérifiant l'estimée de Schauder (1.10).

Pour passer de la solution mild à la solution faible, nous exploitons de nouveau la dualité des espaces de Besov.

## 1.9 Unicité forte

### Présentation du problème

Pour un système d'Équations Différentielles Ordinaires (EDO), le caractère bien posé s'avère être un véritable enjeu quand on quitte le cadre Lipschitz, voir par exemple [PL89]. En effet l'exemple de Peano montre que l'unicité de la solution peut ne pas être garantie quand la dérive du système n'est que Hölder. Néanmoins, la présence d'un bruit permet de restaurer ladite unicité. Ce phénomène s'appelle la régularisation par le bruit, voir e.g. Flandoli [Fla11]. Zvonkin [Zvo74] a été le premier à considérer ce cadre en dimension 1. Puis une généralisation multidimensionnelle et elliptique a été établie par Veretennikov [Ver80]. Mentionnons également Krylov et Röckner [KR05] qui ont vérifié l'unicité forte pour un système non-dégénéré avec dérive  $L_q - L_p$ , ainsi que Zhang [Zha10] pour une matrice de diffusion faiblement Lipschitz. Enfin Fedrizzi et Flandoli [FF11] ont établi l'unicité forte dans un cadre elliptique et avec une dérive  $L_q - L_p$  et Hölder.

Le cadre, dégénéré avec  $n = 2$  a été étudié par Chaudru de Raynal [CdR17], Wang et Zhang [WZ16], Fedrizzi, Flandoli, Priola et Vovelle [FFPV17], Zhang [Zha18].

Nous considérerons comme précédemment dans l'étude des estimées de Schauder, le système suivant :

$$\begin{aligned}
 d\mathbf{X}_t^1 &= \mathbf{F}_1(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt + \sigma(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dW_t, \\
 d\mathbf{X}_t^2 &= \mathbf{F}_2(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt, \\
 d\mathbf{X}_t^3 &= \mathbf{F}_3(t, \mathbf{X}_t^2, \dots, \mathbf{X}_t^n)dt, \\
 &\vdots \\
 d\mathbf{X}_t^n &= \mathbf{F}_n(t, \mathbf{X}_t^{n-1}, \mathbf{X}_t^n)dt,
 \end{aligned}
 \quad t \geq 0. \quad (1.33)$$

L'unicité forte est triviale lorsque les coefficients sont Lipschitz. En effet, on note :

$$\begin{aligned}
 \mathbf{X}_t &= \mathbf{x} + \int_0^t \mathbf{F}(s, \mathbf{X}_s)ds + \int_0^t \sigma(s, \mathbf{X}_s)dW_s, \\
 \mathbf{X}'_t &= \mathbf{x} + \int_0^t \mathbf{F}(s, \mathbf{X}'_s)ds + \int_0^t \sigma(s, \mathbf{X}'_s)dW_s.
 \end{aligned}
 \quad (1.34)$$

Si  $\sigma$  et  $\mathbf{F}$  sont Lipschitz, on déduit qu'il existe  $C = C(\sigma, \mathbf{F}) > 0$  tel que :

$$\mathbb{E}|\mathbf{X}_t - \mathbf{X}'_t|^2 \leq C \int_0^t \mathbb{E}|\mathbf{X}_s - \mathbf{X}'_s|^2 ds. \quad (1.35)$$

Par l'inégalité de Gronwall on a directement que  $X_t = X'_t$  pour tout  $t \geq 0$  presque sûrement.

Dans le cas où la dérive n'est pas Lipschitz, l'étude de l'unicité forte s'avère nettement plus difficile. Nous supposons que :

- **Régularité du coefficient de diffusion.**  $\sigma$  est Lipschitz (une régularité plus faible reste un problème ouvert),
- **Régularité de la dérive.** pour tout  $(i \in \llbracket 1, n \rrbracket)$ , on a  $[(\mathbf{F}_i)]_{\mathbf{d}}^{\beta_j} < \infty$  avec  $\beta_j \in (\frac{2j-2}{2j-1}, 1]$  et  $[(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)]_{\mathbf{d}}^{\eta} < \infty$ , pour un «petit»  $\eta$ .

La régularité de la dérive est optimale pour notre approche, ces seuils de régularité sont cohérents à ceux pour l'estimée de Schauder. La différence vient du fait que les composants de la dérive ont tous la même régularité, c'est-à-dire que la régularité de  $\mathbf{F}_i$  ne dépend pas de  $i$ . Cela s'explique par le fait que la dérive  $\mathbf{F}_i$ , à chaque étage va être amené à jouer le rôle de source dans l'équation de Kolmogorov associé à la transformation de Zvonkin (voir (1.36) plus loin) et que les contrôles que l'on souhaite obtenir soient sans échelle, i.e. du type  $|\mathbf{X}_s - \mathbf{X}'_s|$  comme dans (1.35). Autrement dit, l'étage de la dérive doit fournir la même régularité. Enfin, remarquons que l'on retrouve le même type de régularité supposée dans le cas des estimées de Schauder pour  $\gamma = 1 + \eta$ .

**Theorem 3** *Pour  $\sigma$  non-dégénéré et  $\mathbf{F}$  avec les régularités précédemment supposées, il existe une unique solution forte au système (1.33).*



### Transformation de Zvonkin

Soit  $\mathbf{U}$  la solution au problème de Cauchy suivant :

$$\begin{cases} \partial_t \mathbf{U}(t, \mathbf{x}) + \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}\mathbf{U}(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 \mathbf{U}(t, \mathbf{x}) a(t, \mathbf{x})) = -\mathbf{F}(t, \mathbf{x}), \\ \mathbf{U}(T, \mathbf{x}) = 0, \quad \mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{nd}. \end{cases}$$

L'idée principale de la transformation de Zvonkin est de supprimer le terme de dérive en écrivant la formule d'Itô :

$$\int_0^t \mathbf{F}(s, \mathbf{X}_s) ds = -\mathbf{U}^m(0, \mathbf{x}) + \mathbf{U}^m(t, \mathbf{X}_t) - \int_0^t \mathbf{D}\mathbf{U}^m(s, \mathbf{X}_s) B \sigma(t, \mathbf{X}_t) dW_s + \mathcal{R}_t^m(\mathbf{X}), \quad (1.36)$$

avec

$$\mathcal{R}_t^m(\mathbf{X}) := \int_0^t \mathbf{F}(s, \mathbf{X}_s) - \mathbf{F}^m(s, \mathbf{X}_s) ds - \int_0^t (L_s - L_s^m) \mathbf{U}^m(s, \mathbf{X}_s) ds.$$

Rappelons que l'indice  $m$  signifie que l'on a régularisé, de sorte que le terme de reste  $\mathcal{R}_t^m(\mathbf{X}) \xrightarrow[m \rightarrow +\infty]{p.s.} 0$ . Ainsi, écrivons :

$$\mathbf{X}_s - \mathbf{U}^m(s, \mathbf{X}_s) = -\mathbf{U}^m(0, \mathbf{x}) - \int_0^t \mathbf{D}\mathbf{U}^m(s, \mathbf{X}_s) B \sigma(s, \mathbf{X}_s) dW_s + \mathcal{R}_t^m(\mathbf{X}),$$

et l'étude de l'équation de Kolmogorov (1.36), avec une étude similaire à celle menée pour (1.2) et détaillée dans la partie 1.9, permet d'établir le contrôle

$$\|\mathbf{D}\mathbf{U}^m\|_\infty + \|\mathbf{D}(\mathbf{D}\mathbf{U}^m B)\|_\infty \leq C_T. \quad (1.37)$$

Notons toutefois que contrairement à (1.2) la source  $\mathbf{F}$  n'est pas bornée. Ce qui ne gêne pas ici étant donné que nous n'avons besoin que des bornes sup des dérivées de  $\mathbf{U}$ .

Ensuite, en reprenant  $\mathbf{X}$  et  $\mathbf{X}'$  satisfaisant (1.34) et par (1.36), on écrit

$$\begin{aligned} & \mathbb{E} \left[ \sup_{s \leq t} |\mathbf{X}_s - \mathbf{X}'_s|^2 \right] \\ & \leq 5 \left( \mathbb{E} \left[ \sup_{s \leq t} |\mathbf{U}^m(s, \mathbf{X}_s) - \mathbf{U}^m(s, \mathbf{X}'_s)|^2 \right] \right. \\ & \quad + \mathbb{E} \left[ \int_0^t |[\mathbf{D}\mathbf{U}^m B](s, \mathbf{X}_s) - [\mathbf{D}\mathbf{U}^m B](s, \mathbf{X}'_s)|^2 \|\sigma\|_\infty^2 ds \right] \\ & \quad \left. + \mathbb{E} \left[ \int_0^t (\|\mathbf{D}\mathbf{U}^m B\|_\infty + 1) |\sigma(s, \mathbf{X}_s) - \sigma(s, \mathbf{X}'_s)|^2 ds \right] + 2\|\mathcal{R}^m(\cdot)\|_\infty \right). \end{aligned}$$

L'identité (1.39) implique ainsi qu'il existe  $\mathcal{T} > 0$  tel que pour tout  $0 < T < \mathcal{T}$  et une constante  $C_T = C_T(T) \xrightarrow{T \rightarrow 0} 0$  telle que

$$\mathbb{E} \left[ \sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] \leq C_T \mathbb{E} \left[ \sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] + 10\|\mathcal{R}^m(\cdot)\|_\infty. \quad (1.38)$$

Pour  $\mathcal{T}$  suffisamment petit, on conclut donc en faisant tendre  $m$  vers  $+\infty$  et par argument circulaire on a que  $\mathbb{E} [\sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2] = 0$ . Autrement dit, on vient de montrer l'unicité forte de la dynamique (1.33). Pour un temps quelconque  $T$ , on itère le raisonnement sur  $[\frac{k}{N}T, \frac{k+1}{N}T]$  avec  $k \in \llbracket 0, N-1 \rrbracket$  et  $N$  suffisamment grand.

### Contrôles de gradient

Le cœur de l'analyse est de montrer les contrôles de gradients suivant.

**Theorem 4** *Il existe  $\mathcal{T} > 0$  tel que pour tout  $T \leq \mathcal{T}$  il existe une constante  $C_T := C((\mathbf{A}), T) > 0$  dépendant seulement des coefficients de diffusion et tel que  $C_T \rightarrow 0$  quand  $T \rightarrow 0$  tel que pour tout  $m \geq 0$  la solution  $\mathbf{U}^m$  satisfait à l'inégalité suivante :*

$$\|\mathbf{DU}^m\|_\infty + \|\mathbf{D}(\mathbf{DU}^m B)\|_\infty \leq C_T. \quad (1.39)$$

Notons que les inégalités voulues sont pour les dérivées de  $U^m$ , ainsi le caractère non borné de la source  $\mathbf{F}$  ne posera pas de problème à l'analyse. Notamment, les normes Besov qui seront considérées seront des normes non-homogènes. Ce qui rend les propriétés de dualité bien plus simple.

Chaque composante  $\mathbf{u}_i^m$ ,  $i \in \llbracket 1, n \rrbracket$  de la solution  $\mathbf{U}^m$  de (1.36) s'écrit pour tout  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$

$$\mathbf{u}_i^m(t, \mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\{ -\mathbf{F}_i^m(s, \mathbf{y}) + (L_s^m - \tilde{L}_s^m) \mathbf{u}_i^m(s, \mathbf{y}) \right\} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}), \quad (1.40)$$

avec la même méthode de linéarisation que pour le chapitre précédent sur les estimées de Schauder.

Dit autrement, la formule de Duhamel (ou le développement paramétrix du premier ordre) donne :

$$\mathbf{u}_i(t, \mathbf{x}) = - \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \xi)} \mathbf{F}_i(s, \cdot) \right] (\mathbf{x}) ds + \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \xi)} \left( (L_s - \tilde{L}_s^\xi) \mathbf{u}_i \right) (s, \cdot) \right] (\mathbf{x}),$$

et par définition des opérateurs  $\tilde{L}_s$  et  $L_s$  on a :

$$\begin{aligned} \mathbf{u}_i(t, \mathbf{x}) &= - \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \xi)} \mathbf{F}_i(s, \cdot) \right] (\mathbf{x}) ds \\ &+ \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \xi)} \left( (\mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\xi))) D_1 \mathbf{u}_i(s, \cdot) \right) \right] (\mathbf{x}) \\ &+ \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \xi)} \left( \frac{1}{2} \text{Tr} [a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\xi))] D_1^2 \mathbf{u}_i(s, \cdot) \right) \right] (\mathbf{x}) ds \quad (1.41) \\ &+ \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \xi)} \left( \sum_{k=2}^n (\mathbf{F}_k(s, \cdot) - \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\xi))) \right. \right. \\ &\left. \left. - D_{\mathbf{x}_{k-1}} \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\xi)) (\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{k-1} D_k \mathbf{u}_i(s, \cdot) \right) \right] (\mathbf{x}) ds, \end{aligned}$$

On montre par des méthodes de cancellation (partielles) et par un choix de point de gel convenable (voir les estimées de Schauder) que

$$\left\| \mathbf{D} D_{\mathbf{x}_1} \int_t^T ds \int_{\mathbb{R}^{nd}} \mathbf{F}_k^m(s, \mathbf{y}) \tilde{p}^{(\tau, \xi)}(t, s, \cdot, \mathbf{y}) d\mathbf{y} \right\|_\infty \leq C_T.$$

Pour les autres termes de l'équation (1.41) faisant intervenir la dérivé de  $\mathbf{u}_i$  suivant la composante non-dégénérée  $\mathbf{x}_1$ , on écrit aussi par cancellation (partielle) :

$$\begin{aligned} & \sum_{l=1}^n \left\| D_l D_{\mathbf{x}_1} \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \boldsymbol{\xi})} \left( (\mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) D_1 \mathbf{u}_i(s, \cdot) \right) \right] (\cdot) ds \right\|_{\infty} \\ & + \sum_{l=1}^n \left\| D_l D_{\mathbf{x}_1} \int_t^T \left[ \tilde{P}_{s,t}^{(\tau, \boldsymbol{\xi})} \left( \frac{1}{2} \text{Tr} \left[ a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) D_1^2 \mathbf{u}_i(s, \cdot) \right] \right) \right] (\mathbf{x}) ds \right\|_{\infty} \leq C_T. \end{aligned}$$

En revanche, pour le terme correspondant aux dérivées  $D_k \mathbf{u}_i$ ,  $k \in \llbracket 2, n \rrbracket$ , l'analyse doit être plus subtile.

On développe d'abord les  $l-1$  premières composantes des termes de dérive autour du point de gel  $\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})$ . C'est-à-dire, on écrit :

$$\begin{aligned} & \mathbf{F}_k(s, \cdot) - \mathbf{F}_k(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{k-1} \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k-1} \\ & = \left[ \mathbf{F}_k(s, \cdot) - \mathbf{F}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}) \right] \\ & \quad + \left[ \mathbf{F}_k(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}) - \mathbf{F}_k(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{k-1} \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k-1} \right] \end{aligned}$$

Le premier terme permet de rattraper la singularité de  $D_l D_1 \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y})$ . Puis, formellement, on développe on fait un développement de Taylor à l'ordre 2 de  $u$  à l'aide d'une cancellation partielle :

$$\begin{aligned} & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \sum_{k=2}^l (\mathbf{F}_k(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_k(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \right. \right. \\ & \quad \left. \left. \left. - D_{k-1} \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k-1} \right) D_k \mathbf{u}_i(s, \cdot) \right] (\mathbf{x}) \right|_{\boldsymbol{\xi}=\mathbf{x}} \\ & = \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \sum_{k=2}^l (\mathbf{F}_k(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{k-1}} \mathbf{F}_k(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k-1} \right) \right. \right. \\ & \quad \left. \left. \times \sum_{j=l}^n \int_0^1 d\lambda D_j D_k \mathbf{u}_i(s, \cdot)_{1:l-1}, \cdot)_{l:n} + \lambda (\cdot)_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_j \right] (\mathbf{x}) \right|. \end{aligned} \quad (1.42)$$

Notons en effet, par la structure triangulaire de  $D\mathbf{F}$  on a pour tout  $l \in \llbracket 2, n \rrbracket$ ,  $k \in \llbracket 2, l \rrbracket$ , et  $j \in \llbracket l, n \rrbracket$  :

$$\begin{aligned} & D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( (\mathbf{F}_k(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{k-1}} \mathbf{F}_k(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k-1} \right) \right. \\ & \quad \left. \times D_j D_k \mathbf{u}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) \right] (\mathbf{x}) = \mathbf{0}_{d,d,d,d}. \end{aligned}$$

On ne s'attend pas à ce que toutes les dérivées de  $\mathbf{u}_i$  existent, néanmoins l'identité (1.42) a un sens d'un point de vue de la dualité Besov. C'est-à-dire, que l'on montre maintenant

que :

$$\begin{aligned}
 & \left| \sum_{k=2}^l \sum_{j=l}^n \int_0^1 d\lambda D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \left( \mathbf{F}_k(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right) \right. \right. \right. \\
 & \quad \left. \left. \left. - D_{\mathbf{x}_{k-1}} \mathbf{F}_k(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k-1} \right) \right. \right. \\
 & \quad \left. \left. \times D_k D_j \mathbf{u}_i(s, \cdot_{1:l-1}, \cdot_{l:n} + \lambda(\cdot_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})))(\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_j \right) \right] (\mathbf{x}) \Big| \\
 & \leq \sum_{k=2}^l \sum_{j=l}^n \left| \int_0^1 d\lambda \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \right. \\
 & \quad \left. \times \int_{\mathbb{R}^d} d\mathbf{y}_k \left\{ \left[ \Psi_{k,(l,1),j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), (t, \mathbf{x})}(\mathbf{y}_k) \right] \varphi_{k,l,j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), t, \boldsymbol{\xi}, \lambda}(\mathbf{y}_k) \right\} \right| \\
 & \leq \int_t^T ds \left\| \left[ \Psi_{k,(l,1),j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_k^j}} \sup_{\mathbf{y}_j, j \in \llbracket 1, n \rrbracket, j \neq k, \lambda \in (0,1)} \left\| \varphi_{k,l,j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), t, \boldsymbol{\xi}, \lambda}(\cdot) \right\|_{B_{\infty, \infty}^{\alpha_k^j - 1}} \\
 & \leq CT^\gamma (\|\mathbf{D}\mathbf{u}_i\|_\infty + \|\mathbf{D}\mathbf{D}_1 \mathbf{u}_i\|_\infty), \tag{1.43}
 \end{aligned}$$

avec

$$\begin{aligned}
 \Psi_{k,(l,1),j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), (t, \mathbf{x})} : \mathbf{y}_k \in \mathbb{R}^d & \mapsto \Psi_{k,(l,1),j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), (t, \mathbf{x})}(\mathbf{y}_k) = \\
 & \left[ D_{\mathbf{x}_l} D_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y}) \left( \left( \mathbf{F}_k(\mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_k(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right) \right. \right. \\
 & \quad \left. \left. - D_{\mathbf{x}_{k-1}} \mathbf{F}_k(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k-1} \right) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_j \right] \\
 \varphi_{k,l,j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), t, \boldsymbol{\xi}, \lambda} : \mathbf{y}_k \in \mathbb{R}^d & \mapsto \varphi_{k,l,j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), t, \boldsymbol{\xi}, \lambda}(\mathbf{y}_k) = \\
 & D_{\mathbf{y}_k} D_{\mathbf{y}_j} \mathbf{u}_i(s, \mathbf{y}_{1:l-1}, \mathbf{y}_{l:n} + \lambda(\mathbf{y}_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))). \tag{1.44}
 \end{aligned}$$

Notons que la norme  $\sup_{\mathbf{y}_j, j \in \llbracket 1, n \rrbracket, j \neq k, \lambda \in (0,1)} \left\| \varphi_{k,l,j}^{(s, \mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n}), t, \boldsymbol{\xi}, \lambda}(\cdot) \right\|_{B_{\infty, \infty}^{\alpha_k^j - 1}}$  est équivalent à la norme  $\alpha_k^j$ -Hölder suivant la  $k$ -ème variable de  $D_j \mathbf{u}_i$  que l'on note  $\|D_j(\mathbf{u}_i)_k\|_{\alpha_k^j}$ . Pour établir le contrôle de cette norme, nous procédons par argument circulaire, c'est-à-dire qu'on montre une inégalité du type

$$\|D_j(\mathbf{u}_i)_k\|_{\alpha_k^j} \leq (1 + T^\gamma (\|D_1 \mathbf{u}_i\|_\infty + \|D_j(\mathbf{u}_i)_k\|_{\alpha_k^j})),$$

en injectant la formule de Duhamel (1.41). Le principal intérêt de cette formulation est que contrairement à la méthode adoptée pour les estimées de Schauder, il n'y a plus besoin de méthode de scaling. En effet, l'inégalité (1.43) associée aux autres contrôles implique

$$\|\mathbf{D}\mathbf{D}_1 \mathbf{u}_i\|_\infty \leq C(1 + T^\gamma (\|D_1 \mathbf{u}_i\|_\infty + \|\mathbf{D}\mathbf{D}_1 \mathbf{u}_i\|_\infty)),$$

on déduit dès lors pour  $T$  suffisamment petit que

$$\|\mathbf{D}\mathbf{D}_1 \mathbf{u}_i\|_\infty + \|\mathbf{D}\mathbf{D}_1 \mathbf{u}_i\|_\infty \leq C.$$



---

## Part II

# Non-asymptotic concentration inequalities



# Chapter 3

## A first non-asymptotic concentration result

**Abstract :** We obtain non-asymptotic Gaussian concentration bounds for the difference between the invariant distribution  $\nu$  of an ergodic Brownian diffusion process and the empirical distribution of an approximating scheme with decreasing time step along a suitable class of (smooth enough) test functions  $f$  such that  $f - \nu(f)$  is a coboundary of the infinitesimal generator. We show that these bounds can still be improved when some suitable squared-norms of the diffusion coefficient also lie in this class. We apply these estimates to design computable non-asymptotic confidence intervals for the approximating scheme. As a theoretical application, we finally derive non-asymptotic deviation bounds for the almost sure Central Limit Theorem.

### 1 Introduction

#### 1.1 Setting

The aim of this chapter is to approach the invariant distribution of the solution of the diffusion equation:

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t, \tag{1.1}$$

where  $(W_t)_{t \geq 0}$  is a Wiener process of dimension  $r$  on a given filtered probability space  $(\Omega, \mathbf{G}, (\mathbf{G}_t)_{t \geq 0}, \mathbb{P})$ ,  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , and  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  are assumed to be Lipschitz continuous functions and to satisfy a mean-reverting assumption in the following sense. If  $\mathcal{A}$  denotes the infinitesimal generator of the diffusion (1.1), there exists a twice continuously differentiable Lyapunov function  $V : \mathbb{R}^d \rightarrow (0, +\infty)$  such that  $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$  and  $\mathcal{A}V \leq \beta - \alpha V$  where  $\beta \in \mathbb{R}$  and  $\alpha > 0$ . Such a condition ensures the existence of an invariant distribution. We will also assume uniqueness of the invariant distribution, denoted from now by  $\nu$ . We refer to the monographs by Khasminskii [Kha80] (see also its augmented second edition [KM11]), or Villani [Vil09] and to the survey paper [Pag01], for in-depth discussions on the conditions yielding such existence and uniqueness results.

We introduce an approximation algorithm based on an Euler like discretization with



decreasing time step, which may use more general innovations than the Brownian increments. Namely, for the step sequence  $(\gamma_k)_{k \geq 1}$  and  $n \geq 0$ , we define:

$$X_{n+1} = X_n + \gamma_{n+1}b(X_n) + \sqrt{\gamma_{n+1}}\sigma(X_n)U_{n+1}, \quad (1.2)$$

where  $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$  and  $(U_n)_{n \geq 1}$  is an i.i.d. sequence of centered random variables matching the moments of the Gaussian law on  $\mathbb{R}^r$  up to order three, independent of  $X_0$ .

We define the empirical (random) occupation measure of the scheme in the following way. For all  $A \in \mathcal{B}(\mathbb{R}^d)$  (where  $\mathcal{B}(\mathbb{R}^d)$  denotes the Borel  $\sigma$ -field on  $\mathbb{R}^d$ ):

$$\nu_n(A) := \nu_n(\omega, A) := \frac{\sum_{k=1}^n \gamma_k \delta_{X_{k-1}(\omega)}(A)}{\sum_{k=1}^n \gamma_k}. \quad (1.3)$$

The measure  $\nu_n$  is here defined accordingly to the *intrinsic* time scale of the scheme. Namely,  $\Gamma_n = \sum_{k=1}^n \gamma_k$  represents the current time associated with the Euler scheme (1.2) after  $n$  iterations. Since we are interested in long time approximation, we consider steps  $(\gamma_k)_{k \geq 1}$  such that  $\Gamma_n := \sum_{k=1}^n \gamma_k \xrightarrow[n]{n} +\infty$ . We also assume  $\gamma_k \xrightarrow[k]{k} 0$ . Observe that, for a bounded  $\nu$ -a.s. continuous function  $f$ , it is proved in [LP02] (see *e.g.* Theorem 1), that:

$$\nu_n(f) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(X_{k-1}) \xrightarrow[n]{a.s.} \nu(f) = \int_{\mathbb{R}^d} f(x) \nu(dx), \quad (1.4)$$

or equivalently that  $\nu_n(\omega, \cdot) \xrightarrow[n]{w} \nu$ ,  $\mathbb{P}(d\omega) - a.s.$  The above result can be seen as an inhomogeneous counterpart of *stability results* discussed for homogeneous Markov chains in Duflo [Duf90]. Intuitively, the decreasing steps make the approximation more and more accurate in long time and, therefore, the ergodic empirical mean of the scheme converges to the quantity of interest. Put it differently, there is no bias. This is a significant advantage w.r.t. a more *naïve* discretization method that would rely on a constant step scheme. Indeed, even if this latter approach gains in simplicity, taking  $\gamma_k = h > 0$  in (1.2) would lead to replace the r.h.s. of (1.4) by the quantity  $\nu^h(f) := \int_{\mathbb{R}^d} f(x) \nu^h(dx)$ , with  $\nu^h$  denoting the invariant distribution of the Euler scheme with step  $h$ . In such a case, for the analysis to be complete, one needs to investigate the difference  $\nu - \nu^h$  through the corresponding continuous and discrete Poisson problems. We refer to Talay *et al.* [TT90], [Tal02] for a precise presentation of this approach.

Now, once (1.4) is available, the next question naturally concerns the rate of that convergence. This was originally investigated by Lamberton and Pagès [LP02] for functions  $f$  of the form  $f - \nu(f) = \mathcal{A}\varphi$ , i.e.  $f - \nu(f)$  is a *coboundary* for  $\mathcal{A}$ . The specific reason for focusing on such a class of functions is that an invariant distribution  $\nu$  is characterized as a solution in the distribution sense of the stationary Fokker-Planck equation  $\mathcal{A}^*\nu = 0$  (where  $\mathcal{A}^*$  stands for the adjoint of  $\mathcal{A}$ ). Thus, for smooth enough functions  $\varphi$  (at least  $\mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ ), one has  $\nu(\mathcal{A}\varphi) = \int_{\mathbb{R}^d} \mathcal{A}\varphi(x) \nu(dx) = 0$ . The authors then investigate the weak convergence of  $\nu_n(f) - \nu(f)$  once suitably renormalized. However, in these results, the assumptions are made on the function  $\varphi$  itself rather than on  $f$ . To overcome this limitation and exploit directly some assumptions on the function  $f$  requires to solve the Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$ . This is precisely for this step that some structure conditions are needed, namely (hypo)ellipticity or confluence conditions. We refer for

instance to the work of Pardoux and Veretennikov [PV01], Rothschild and Stein [RS76] or Villani [Vil09] who discuss the solvability of the Poisson problem under some ellipticity or hypoellipticity assumptions. We also mention the work of Pagès and Panloup [PP14] who exploit some confluence conditions allowing to handle for instance the case of an Ornstein-Uhlenbeck process with degenerate covariance matrix. We refer to Sections 2.2 and 2.3 for precise assumptions giving the uniqueness of the invariant distribution of (1.1) and the expected smoothness properties for the associated Poisson problem.

In the current chapter, our goal is to establish for this recursive procedure a non-asymptotic Gaussian control for the deviations of the quantity  $\nu_n(f) - \nu(f)$  for *possibly unbounded* Lipschitz continuous functions  $f$ . Such non-asymptotic bounds are crucial in many applicative fields. Indeed, for specific practical simulations, it is not always possible to run ergodic means for very large values of  $n$ . It will be direct to derive, as a by-product of our deviations estimates, some *computable* non-asymptotic confidence intervals. A specific feature of such non-asymptotic deviation inequalities is that their accuracy depends again on the status of the diffusion coefficient  $\sigma$  with respect to the Poisson equation. Thus, if  $\|\sigma\|^2 - \nu(\|\sigma\|^2) = \mathcal{A}\vartheta$  is a coboundary (where  $\|\cdot\|$  denotes a matrix norm), we manage to improve our analysis, to derive better concentration bounds in a certain deviation range as well as some additional deviation regimes. Also, this additional study seems rather efficient to capture the numerical behavior of the empirical deviations. We refer to Section 4 and 6.2 for details about these points. Eventually, our main deviation results allow to provide deviation inequalities for plain Lipschitz continuous sources  $f$  in the ergodic approximation, by using a suitable regularization procedure, as established in Theorem 7. As expected, dealing with this general class of functions requires more stringent constraints on the time steps, that must be *small* enough, and prevents from obtaining the fastest convergence rates (see again Theorem 7 and Section 5.3).

The main feature of the sequence (1.4) of weighted empirical measures is that it targets the *true* invariant distribution  $\nu$  of the continuous time diffusion. The price to pay is the use of an Euler scheme with decreasing step which is a non-homogeneous Markov chain. This induces new difficulties compared to the extensive literature on deviation inequalities for ergodic homogeneous Markov chains. In particular, our approximation procedure produces some remainder terms that need to be controlled accurately enough in a non-asymptotic way to produce tractable deviation inequalities asymptotically close to their counterparts for the diffusion itself. This a major difficulty compared to a *CLT* where these remainder terms are simply requested to go to 0 fast enough.

As mentioned above and like for the *CLT* (see [Bha82] for the diffusion or [LP02] for the weighted empirical measures  $\nu_n$ ), these deviation inequalities are naturally established for coboundaries  $f - \nu(f) = \mathcal{A}(\varphi)$ , the assumptions being made on  $\varphi$ . Our second objective in this chapter is to state our results so that all assumptions could be read on the source function  $f$  itself. This first requires to solve the Poisson equation in that spirit, that means deriving pointwise regularity results on  $\varphi$  from those made on  $f$ . Again, for Lipschitz sources, this step will require an appropriate regularization procedure.

In particular, we will not rely on the Sobolev regularity (see e.g. Pardoux and Veretennikov [PV01]) but rather on some Schauder estimates in line with the works by Krylov and Priola [KP10], which allow to benefit from the elliptic regularity for operators with unbounded coefficients. For more details, we refer to the introduction of Section 5.

## 1.2 Assumptions and Related Asymptotic Results

From now on, we will extensively use the following notations.

For a given step sequence  $(\gamma_n)_{n \geq 1}$ , we denote:

$$\forall \ell \in \mathbb{R}, \Gamma_n^{(\ell)} := \sum_{k=1}^n \gamma_k^\ell, \quad \Gamma_n := \sum_{k=1}^n \gamma_k = \Gamma_n^{(1)}.$$

In practice, we will consider time step sequences:  $\gamma_n \asymp \frac{1}{n^\theta}$  with  $\theta \in (0, 1]$ , where for two sequences  $(u_n)_{n \in \mathbb{N}}$ ,  $(v_n)_{n \in \mathbb{N}}$  the notation  $u_n \asymp v_n$  means that  $\exists n_0 \in \mathbb{N}$ ,  $\exists C \geq 1$  such that  $\forall n \geq n_0$ ,  $C^{-1}v_n \leq u_n \leq Cv_n$ .

For a vector  $v \in \mathbb{R}^k$ ,  $k \in \{d, r\}$ , we denote by  $|v| := (\sum_{j=1}^k v_j^2)^{\frac{1}{2}}$  its (canonical) Euclidean norm. Also, for a function  $\psi : \mathbb{R}^q \rightarrow \mathbb{R}^d$ , we set  $\|\psi\|_\infty := \sup_{x \in \mathbb{R}^q} |\psi(x)|$ .

### Hypotheses

**(C1)** The random variable  $X_0$  is supposed to be sub-Gaussian, *i.e.* its square is exponentially integrable up to some threshold. Namely, there exists  $\lambda_0 \in \mathbb{R}_+^*$  such that:

$$\forall \lambda < \lambda_0, \quad \mathbb{E}[\exp(\lambda|X_0|^2)] < +\infty.$$

**(GC)** The  $\mu$ -distributed i.i.d. innovation sequence  $(U_n)_{n \geq 1}$  is such that  $\mathbb{E}[U_1] = 0$  and for all  $(i, j, k) \in \{1, \dots, r\}^3$ ,  $\mathbb{E}[U_1^i U_1^j] = \delta_{ij}$ ,  $\mathbb{E}[U_1^i U_1^j U_1^k] = 0$ . Also,  $(U_n)_{n \geq 1}$  and  $X_0$  are independent. Eventually,  $U_1$  satisfies the following Gaussian concentration property, *i.e.* for every 1-Lipschitz continuous function  $g : \mathbb{R}^r \rightarrow \mathbb{R}$  and every  $\lambda > 0$ :

$$\mathbb{E}[\exp(\lambda g(U_1))] \leq \exp\left(\lambda \mathbb{E}[g(U_1)] + \frac{\lambda^2}{2}\right).$$

Observe that if  $U_1 \stackrel{(\text{law})}{=} \mathcal{N}(0, I_r)$  or  $U_1 \stackrel{(\text{law})}{=} (\frac{1}{2}(\delta_1 + \delta_{-1}))^{\otimes r}$ , *i.e.* for Gaussian or symmetric Bernoulli increments which are the most commonly used sequences for the innovations, the above identity holds. On the other hand, what follows can be adapted almost straightforwardly for a wider class of sub-Gaussian distributions satisfying that for some  $\varpi > 0$  and for all  $\lambda > 0$ :

$$\mathbb{E}[\exp(\lambda g(U_1))] \leq \exp\left(\lambda \mathbb{E}[g(U_1)] + \frac{\varpi \lambda^2}{4}\right), \quad (1.5)$$

which yields that for all  $r > 0$ ,  $\mathbb{P}[|U_1| \geq r] \leq 2 \exp(-\frac{r^2}{\varpi})$  (sub-Gaussian concentration of the innovation). The case  $\varpi = 2$  corresponds to the standard Gaussian concentration. This is also the constant in the logarithmic Sobolev inequality fulfilled by the standard Gaussian measure.

**(C2)** There exists a positive constant  $\kappa$  such that,

$$\sup_{x \in \mathbb{R}^d} \|\sigma(x)\|^2 \leq \kappa,$$

where  $\|\sigma(x)\|$  stands for the operator norm of  $\sigma(x)$ , i.e.  $\|\sigma(x)\| = \sup_{z \in \mathbb{R}^r, |z| \leq 1} |\sigma(x)z|$  (keep in mind that  $\|\sigma(x)\| = \|\sigma^*(x)\| = \|\sigma\sigma^*(x)\|^{\frac{1}{2}}$ ). We then set  $\|\sigma\|_\infty := \sup_{x \in \mathbb{R}^d} \|\sigma(x)\|$ .

( $\mathcal{L}_V$ ) There exists a Lyapunov function  $V : \mathbb{R}^d \longrightarrow [v^*, +\infty[$ , with  $v^* > 0$ , satisfying the following conditions:

- i) Regularity-Coercivity.  $V$  is a  $\mathcal{C}^2$  function,  $\|D^2V\|_\infty < +\infty$ , and  $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ .
- ii) Growth control. There exists  $C_V \in (0, +\infty)$  such that for all  $x \in \mathbb{R}^d$ :

$$|\nabla V(x)|^2 + |b(x)|^2 \leq C_V V(x).$$

- iii) Stability. Let  $\mathcal{A}$  be the infinitesimal generator associated with the diffusion Equation (1.1), defined for all  $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R})$  and for all  $x \in \mathbb{R}^d$  by:

$$\mathcal{A}\varphi(x) = \langle b(x), \nabla\varphi(x) \rangle + \frac{1}{2} \text{Tr}(\Sigma(x)D^2\varphi(x)), \quad \Sigma(x) := \sigma\sigma^*(x),$$

where, for two vectors  $v_1, v_2 \in \mathbb{R}^d$ , the symbol  $\langle v_1, v_2 \rangle$  stands for the canonical inner product of  $v_1$  and  $v_2$  and for  $M \in \mathbb{R}^d \otimes \mathbb{R}^d$ ,  $\text{Tr}(M)$  denotes the trace of the matrix  $M$ .

There exist  $\alpha_V > 0$ ,  $\beta_V \in \mathbb{R}^+$  such that for all  $x \in \mathbb{R}^d$ ,

$$\mathcal{A}V(x) \leq -\alpha_V V(x) + \beta_V.$$

As a consequence of ( $\mathcal{L}_V$ ) i), there exist constants  $K$  and  $\bar{c}$  such that for  $|x| \geq K$ ,  $|V(x)| \leq \bar{c}|x|^2$ , which in turn implies, from ( $\mathcal{L}_V$ ) ii), that  $|b(x)| \leq \sqrt{C_V \bar{c}}|x|$ .

( $\mathbf{U}$ ) There exists a unique invariant distribution, denoted from now on by  $\nu$ , for Equation (1.1).

( $\mathbf{S}$ ) For a Lyapunov function  $V$  satisfying ( $\mathcal{L}_V$ ), we assume that the step sequence  $(\gamma_k)_{k \geq 1}$  satisfies for all  $k \geq 1$ ,  $\gamma_k \leq \frac{1}{2} \min(\frac{1}{\sqrt{C_V \bar{c}}}, \frac{\alpha_V}{C_V \|D^2V\|_\infty})$ .

Condition ( $\mathbf{S}$ ) means that we assume the time steps are sufficiently *small* w.r.t. the upper bounds of the coefficients and the Lyapunov function.

**Remark 1** *We have assumed ( $\mathbf{U}$ ) without imposing some specific non-degeneracy conditions. Observe that ( $\mathcal{L}_V$ ) yields existence of an invariant distribution (see e.g. Chapter 4.9 in [EK86]). Additional structure conditions on the coefficients ((hypo-)ellipticity [KM11], [PV01], [PV03], [PV05], [Vil09] or confluence [PP14]) then yield uniqueness.*

*Assumption ( $\mathbf{S}$ ) is a technical condition which is exploited in order to derive the non-asymptotic controls of Theorem 2 (see especially the proof of Proposition 1 below).*

Observe that, as soon as conditions ( $\mathbf{C2}$ ), ( $\mathcal{L}_V$ ), ( $\mathbf{U}$ ) are satisfied and  $\mathbb{E}[U_1] = 0$ ,  $\mathbb{E}[U_1^{\otimes 3}] = 0$ , the following Central Limit Theorem (CLT) holds (see Theorems 9, 10 in [LP02]).

**Theorem 1 (CLT)** Under **(C2)**,  $(\mathcal{L}_\nu)$ , **(U)**, if  $\mathbb{E}[U_1] = 0$ ,  $\mathbb{E}[U_1^{\otimes 3}] = 0$  and  $\mathbb{E}[V(X_0)] < +\infty$ , we have the following results.

- (a) *Fast decreasing step.* If  $\lim_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} = 0$  and  $\mathbb{E}[|U_1|^6] < +\infty$ , then, for any Lipschitz continuous function  $\varphi$  in  $\mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$  with  $D^2\varphi$  and  $D^3\varphi$  bounded, one has (with  $(\mathcal{L})$  denoting weak convergence)

$$\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \xrightarrow{(\mathcal{L})} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right).$$

- (b) *Critical and slowly decreasing step.* If  $\lim_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} = \tilde{\gamma} \in ]0, +\infty]$  and if  $\mathbb{E}[|U_1|^8] < +\infty$ , then for every Lipschitz continuous function  $\varphi \in \mathcal{C}^4(\mathbb{R}^d, \mathbb{R})$  with  $(D^i\varphi)_{i \in \{2,3,4\}}$  bounded:

$$\begin{aligned} \sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) &\xrightarrow{(\mathcal{L})} \mathcal{N}\left(\tilde{\gamma}m, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right) \quad \text{if } \tilde{\gamma} < +\infty, \text{ (critical decreasing step)} \\ \frac{\Gamma_n}{\Gamma_n^{(2)}}\nu_n(\mathcal{A}\varphi) &\xrightarrow{\mathbb{P}} m \quad \text{if } \tilde{\gamma} = +\infty, \text{ (slowly decreasing step),} \end{aligned}$$

$$\text{where} \quad m := - \int_{\mathbb{R}^d} \left( \frac{1}{2} D^2\varphi(x)b(x)^{\otimes 2} + \Phi_4(x) \right) \nu(dx),$$

$$\text{with} \quad \Phi_4(x) := \int_{\mathbb{R}^r} \left( \frac{1}{2} \langle D^3\varphi(x)b(x), (\sigma(x)u)^{\otimes 2} \rangle + \frac{1}{24} D^4\varphi(x)(\sigma(x)u)^{\otimes 4} \right) \mu(du)$$

and  $\mu$  denotes the distribution of the innovations  $(U_k)_{k \geq 1}$ . In the above definition of  $\Phi_4$ , the term  $D^3\varphi$  stands for the order 3 tensor  $(\partial_{x_i, x_j, x_k}^3 \varphi)_{(i,j,k) \in \llbracket 1, d \rrbracket^3}$  and we denote, for all  $x \in \mathbb{R}^d$ , by  $D^3\varphi(x)b(x)$  the  $\mathbb{R}^d \otimes \mathbb{R}^d$  matrix with entries  $(D^3\varphi(x)b(x))_{ij} = \sum_{k=1}^d (D^3\varphi(x))_{ijk} b_k(x)$ ,  $(i, j) \in \llbracket 1, d \rrbracket^2$ .

**Remark 2** Let us specify that for a step sequence  $(\gamma_n)_{n \in \mathbb{N}}$  such that  $\gamma_n \asymp n^{-\theta}$ ,  $\theta \in (0, 1]$ , it is easily checked that case (a) occurs for  $\theta \in (\frac{1}{3}, 1]$  for which  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow 0$ . In case (b), that is for  $\theta \in (0, \frac{1}{3}]$ ,  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow \tilde{\gamma}$ , with  $\tilde{\gamma} < +\infty$  for  $\theta = \frac{1}{3}$  and  $\tilde{\gamma} = +\infty$  for  $\theta \in (0, \frac{1}{3})$ .

Let us mention that, when  $\lim_n \frac{\Gamma_n^{(3/2)}}{\sqrt{\Gamma_n}} < +\infty$ , i.e.  $\gamma_n \asymp n^{-\theta}$ ,  $\theta \in (1/2, 1]$ , the CLT of point (a) holds without the condition  $\mathbb{E}[U_1^{\otimes 3}] = 0$  provided  $\mathbb{E}[|U_1|^4] < +\infty$  (see Theorem 9 in [LP02]). Moreover, the boundedness condition **(C2)** can be relaxed to derive the CLT, which holds provided  $\lim_{|x| \rightarrow +\infty} \frac{|\sigma^* \nabla \varphi(x)|^2}{V(x)} = 0$  (strictly sublinear diffusion) in case (a) and  $\sup_{x \in \mathbb{R}^d} \frac{|\sigma^* \nabla \varphi(x)|^2}{V(x)} < +\infty$  (sublinear diffusion) in case (b). We refer again to Theorems 9 and 10 in [LP02] for further considerations.

**Remark 3** The reader should have in mind that an ergodic result similar to the one stated in the fast decreasing step setting holds for the diffusion itself under the same structure

assumptions, i.e. **(C2)**, **(L<sub>v</sub>)**, **(U)** (see Bhattacharya [Bha82]). In fact **(C2)** can be partially relaxed as well, as mentioned above. Precisely,

$$\frac{1}{\sqrt{t}} \int_0^t \mathcal{A}\varphi(Y_s) ds \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right) \quad \text{as } t \rightarrow +\infty.$$

Note that the asymptotic variance corresponds to the usual integral of the “carré du champ” w.r.t. to the invariant distribution (see again Bhattacharya [Bha82] or the monograph by Bakry et al. [BGL14]), i.e.:

$$\int_{\mathbb{R}^d} |\sigma^* \nabla \varphi(x)|^2 \nu(dx) = -2 \int_{\mathbb{R}^d} \langle \mathcal{A}\varphi, \varphi \rangle(x) \nu(dx).$$

In both settings, the normalization is the same:  $\sqrt{t}$  for the diffusion and  $\sqrt{\Gamma_n}$  for the scheme. Except that, as emphasized by Theorem 1, for slowly decreasing step – when  $\theta < 1/3$  – the time discretization effect becomes prominent and “hides” the CLT so that  $\theta = 1/3$  (critical value between “fast” and “slow” settings) yields the fastest rate with a biased CLT.

**Remark 4** We would like to mention that, in the biased case (b), for steps of the form  $\gamma_k = \gamma_0 k^{-1/3}$ ,  $k \geq 1$ , it is important for a practical implementation to choose  $\gamma_0$  in an appropriate way, namely by minimizing the function  $\gamma_0 \mapsto c_1 \gamma_0 + c_2 \gamma_0^{-1/2}$ ,  $c_1 = \lim_n \frac{\sum_{k=1}^n k^{-2/3}}{(\sum_{k=1}^n k^{-1/3})^{1/2}}$ ,  $c_2 = \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu$ , which corresponds to the mean-variance contribution deriving from the biased limit Theorem. Of course,  $c_2$  is usually unknown, and the concrete optimization has to be performed replacing  $c_2$  by a computable estimate, like for instance upper bounds, i.e.  $c_2 \leq \|\sigma\|_\infty \|\nabla \varphi\|_\infty$ .

The purpose of this work is to obtain non-asymptotic deviation results which match with the above CLT. In the current ergodic framework, the very first non-asymptotic results were established for the Euler scheme with constant time step by Malrieu and Talay in [MT06] when the diffusion coefficient  $\sigma$  in (1.1) is constant. The key tool in their approach consists in establishing a Log Sobolev inequality, which implies Gaussian concentration, for the Euler scheme. This approach allows to easily control the invariant distribution associated with the diffusion process (1.1), see *e.g.* Ledoux [Led99] or Bakry *et al.* [BGL14] in a general framework. However Log Sobolev, and even Poincaré, inequalities turn out to be rather *rigid* tools and are not very well adapted for discretization schemes like (1.2) with or without decreasing steps.

Our approach relies on martingale techniques, which were already a crucial tool to establish the asymptotic results of [LP02] and have been successfully used in Frikha and Menozzi [FM12] as well to establish non-asymptotic bounds for the regular Monte Carlo error associated with the Euler discretization of a diffusion over a finite time interval  $[0, T]$  and a class of stochastic algorithms of Robbins-Monro type. Roughly speaking, for a given  $n$ , we decompose the quantity  $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$  as  $M_n + R_n$  where  $(M_k)_{k \geq 0}$  is a martingale which has Gaussian concentration and  $R_n$  is a remainder term to be controlled in a non-asymptotic way.

We can as well refer to the recent work by Dedecker and Gouëzel [DG15] who also use a martingale approach to derive non-asymptotic deviation bounds for separately bounded functionals of geometrically ergodic Markov chains on a general state space.

Let us also mention that many non-asymptotic results have been obtained based on functional inequalities. Bolley, Guillin and Villani [BGV07] derived non-asymptotic controls for the deviations of the Wasserstein distance between a reference measure and its empirical counterpart, establishing a non-asymptotic version of the Sanov theorem. Deviation estimates for sums of weakly dependent random variables (with sub exponential mixing rates) have been considered in Merlevède *et al.* [MPR11]. From a more dynamical viewpoint, let us mention the work of Joulin and Ollivier [JO10], who introduced for rather general homogeneous Markov chains a kind of *curvature* condition to derive a spectral gap for the chain, and therefore an exponential convergence of the marginal laws towards the stationary distribution. We also mention a work of Blower and Bolley [BB06], who obtain Gaussian concentration properties for deviations of functional of the path for metric space valued homogeneous Markov chains or Boissard [Boi11] who established non-asymptotic deviation bounds for the Wasserstein distance between the marginal distributions and the stationary law, still in the homogeneous case. The common idea of these works is to prove some contraction properties of the transition kernel of the Markov chain in Wasserstein metric. However, this usually requires to have some continuity in Wasserstein metric for the transition law involved, see *e.g.* condition (ii) in Theorems 1.2 and 2.1 of [BB06]. Checking such continuity conditions can be difficult in practice. Sufficient conditions, which require absolute continuity and smoothness of the transition laws are given in Proposition 2.2 of [BB06].

Though potentially less sharp for the derivation of constants, the adopted martingale-based approach in this work turns out to be rather simple, robust and can be very naturally adapted to both discrete innovations and inhomogeneous time steps dynamics like the one we currently consider.

It should as well allow to control deviations for functionals of the path, in the spirit of those considered in [PP12], [PP14]. Also, the approach could possibly extend to diffusions with less stringent Lyapunov conditions, like the weakly mean reverting drifts considered in [LP03], or even to more general ergodic Markov processes, see *e.g.* Pagès and Rey [PR17]. These aspects will concern further research.

As an application of our non-asymptotic concentration results, we will discuss two important topics:

- The first one is of numerical interest and deals with non-asymptotic confidence intervals associated with the estimation of the ergodic mean. Such results can be very useful in practice when the computational resources are constrained (by time, by the model itself, ...). If we assume that  $\varphi \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$ , Lipschitz continuous with  $(D^i \varphi)_{i \in \{2,3\}}$  bounded, such that the mapping  $x \in \mathbb{R}^d \mapsto \langle b(x), \nabla \varphi(x) \rangle$  and  $D^3 \varphi$  are Lipschitz continuous, we then establish that there are explicit sequences  $c_n \leq 1 \leq C_n$  converging to 1 such that for all  $n \in \mathbb{N}$ , for all  $a > 0$  and  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in (\frac{1}{3}, 1]$ ,

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}\right). \quad (1.6)$$

When the diffusion coefficient  $\sigma$  is such that  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  is itself a coboundary (or its counterpart for any other norm dominating  $\|\cdot\|$ ), the previous bound improves in a certain deviation range for  $a$ . Namely, we are able to replace  $\|\sigma\|_\infty^2$  by  $\nu(\|\sigma\|^2)$  in (1.6),

going thus closer to the theoretical limit variance involving the “*carré du champ*”. Moreover, a mixed regime appears in the non-asymptotic deviation bounds which dramatically improves, from the numerical viewpoint, the general case for a certain deviation range. In particular, the corresponding variance is closer to the asymptotic one given by the “*carré du champ*” (see Theorem 8 below). In accordance with the limit results of Theorem 1, the drifts associated with the fastest convergence rates can be handled as well. We obtain in full generality, results of type (1.6) under slightly weaker smoothness assumptions, considering *e.g.*  $D^3\varphi$  being  $\beta \in (0, 1]$ -Hölder continuous. Eventually, under suitable ellipticity conditions on  $\sigma$ , we are able to give non-asymptotic deviation bounds for a Lipschitz source  $f$  as well as explicit gradient bounds for the solution  $\varphi$  of the corresponding Poisson problem.

– The second one is mainly theoretical and concerns non-asymptotic deviation bounds for the *celebrated* almost-sure CLT first established by Brosamler and Schatte (see [Bro88] and [Sch88]) and revisited through the ergodic discretization schemes viewpoint in [LP02].

Both applications require a careful investigation of the corresponding Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$ . We will in particular prove that some pointwise regularity properties can be transferred from  $f$  to  $\varphi$ .

The chapter is organized as follows. We conclude this section by introducing some notations. Our main results are presented in Section 2. We first state therein the specific concentration results for functions  $f$  writing  $f = \mathcal{A}\varphi + \nu(f)$  (see Section 2.1). We then proceed with some suitable controls on the Poisson problem associated with  $\mathcal{A}$  and  $f$  in a confluent framework under the two main cases considered: namely a possibly degenerate setting, which requires a strong confluence condition and *smooth* source and coefficients, and a non degenerate setting, which allows to weaken the confluence condition as well as the smoothness assumptions on the source and the coefficients since in that case we manage to benefit from an elliptic bootstrap property (see Section 2.2). We eventually give in Section 2.3 some practical and tractable deviation bounds and non-asymptotic confidence intervals, including a Slutsky like result, for a given specific source  $f$  under the afore mentioned conditions on the coefficients of (1.1).

We prove our main concentration result in Section 3. Section 4 is devoted to the case where  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  is a coboundary. We then prove in Section 5 the required controls on the Poisson equation for our deviation result to hold as well as the practical controls of Section 2.3. Section 6.1 is dedicated to the non-asymptotic deviation bounds for the almost-sure CLT and Section 6.2 to the numerical illustration of our non-asymptotic confidence intervals.

### 1.3 Notations

In the following, we will denote by  $C$  a constant that may change from line to line and depend, uniformly in time, on known parameters appearing in **(C1)**, **(GC)**, **(C2)**, **(L<sub>V</sub>)**, **(S)**. Other possible dependencies will be explicitly specified. We will also denote by  $\mathcal{R}_n$  and  $e_n$  deterministic remainder terms that respectively converge to 1 and 0 with  $n$ . The explicit dependencies of those sequences again appear in the proofs.



For a function  $f \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R})$ ,  $\beta \in (0, 1]$ , we denote

$$[f]_\beta := \sup_{x \neq x'} \frac{|f(x) - f(x')|}{|x - x'|^\beta} < +\infty$$

its Hölder modulus of continuity. Observe carefully that, when  $f$  is additionally bounded, we have that for all  $0 < \beta' < \beta$ :

$$[f]_{\beta'} \leq [f]_\beta^{\frac{\beta}{\beta'}} (2\|f\|_\infty)^{1-\frac{\beta}{\beta'}}. \quad (1.7)$$

Additionally, for  $f \in \mathcal{C}^p(\mathbb{R}^d, \mathbb{R})$ ,  $p \in \mathbb{N}$ , we set for  $\beta \in (0, 1]$ :

$$[f^{(p)}]_\beta := \sup_{x \neq x', |\alpha|=p} \frac{|D^\alpha f(x) - D^\alpha f(x')|}{|x - x'|^\beta} \leq +\infty,$$

where  $\alpha$  (viewed as an element of  $\mathbb{N}_0^d \setminus \{0\}$  with  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ ) is a multi-index of length  $p$ , *i.e.*  $|\alpha| := \sum_{i=1}^d \alpha_i = p$ .

For notational convenience, we also introduce for  $k \in \mathbb{N}_0$ ,  $\beta \in (0, 1]$  and  $m \in \{1, d, d \times r\}$  the Hölder space

$$\mathcal{C}^{k,\beta}(\mathbb{R}^d, \mathbb{R}^m) := \left\{ f \in \mathcal{C}^k(\mathbb{R}^d, \mathbb{R}^m) : \forall \alpha, |\alpha| \in \llbracket 1, k \rrbracket, \sup_{x \in \mathbb{R}^d} |D^\alpha f(x)| < +\infty, [f^{(k)}]_\beta < +\infty \right\}.$$

We also denote by  $\mathcal{C}_b^{k,\beta}$  the subset of  $\mathcal{C}^{k,\beta}$  for which the functions themselves are bounded. In particular,  $\mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R}^m)$  is the space of Lipschitz continuous functions from  $\mathbb{R}^d$  to  $\mathbb{R}^m$  and  $\mathcal{C}_b^{0,\beta}(\mathbb{R}^d, \mathbb{R}^m)$  denotes the space of bounded  $\beta$ -Hölder continuous functions. Observe as well that, if  $f \in \mathcal{C}^{k,\beta}$ ,  $k \geq 1$  then  $f$  is Lipschitz continuous.

We will as well use the notation  $\llbracket n, p \rrbracket$ ,  $(n, p) \in (\mathbb{N}_0)^2$ ,  $n \leq p$ , for the set of integers being between  $n$  and  $p$ . Also, for a given Borel function  $f : \mathbb{R}^d \rightarrow E$ , where  $E$  can be  $\mathbb{R}, \mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^q$ ,  $q \in \{r, d\}$ , we set for  $k \in \mathbb{N}_0$ :

$$f_k := f(X_k).$$

Eventually, for  $k \in \mathbb{N}_0$ , we denote by  $\mathcal{F}_k := \sigma((X_j)_{j \in \llbracket 0, k \rrbracket})$ .

## 2 Main results

### 2.1 Result of non-asymptotic Gaussian concentration

Our main concentration result is given by the following theorem. In this theorem, we consider a slightly more general situation than for the CLT recalled in Theorem 1. We only assume  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ ,  $\beta \in (0, 1]$  instead of  $\varphi \in \mathcal{C}^4(\mathbb{R}^d, \mathbb{R})$  with existing bounded partial derivatives up to order four (which in particular implies that in Theorem 1  $\varphi \in \mathcal{C}^{3,1}(\mathbb{R}^d, \mathbb{R})$ ).

**Theorem 2** Assume **(C1)**, **(GC)**, **(C2)**, **(L<sub>V</sub>)**, **(U)**, **(S)** hold. Consider a Lipschitz continuous (possibly unbounded) function  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  for some  $\beta \in (0, 1]$ . Let us furthermore suppose that:

$$\exists C_{V,\varphi} > 0, \forall x \in \mathbb{R}^d, |\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)}). \quad (\mathbf{G}_V)$$

Let  $\theta \in [1/(2 + \beta), 1]$  and assume the step sequence  $(\gamma_k)_{k \geq 1}$  is of the form  $\gamma_k = k^{-\theta}$ .

(a) **Unbiased Case (sub-optimal convergence rate)**: Let  $\theta \in (\frac{1}{2+\beta}, 1]$ .

(i) Assume that the mapping  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  is Lipschitz continuous.

Then, there exist two explicit monotonic sequences  $c_n \leq 1 \leq C_n$ ,  $n \geq 1$ , with  $\lim_n C_n = \lim_n c_n = 1$ , such that for all  $n \geq 1$  and for every  $a > 0$ :

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}\right).$$

(ii) Suppose that the mapping  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  **is not** Lipschitz continuous. The above result still holds for  $0 < a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$  for a positive sequence  $\chi_n \xrightarrow{n} 0$  arbitrarily slowly, so that  $\chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \xrightarrow{n} +\infty$ . In particular, for a fixed  $a > 0$ , the above concentration inequality holds for  $n$  large enough.

(b) **Biased Case (Optimal Convergence Rate)**: Let  $\theta = \frac{1}{2+\beta}$ . We set for all  $(k, t, u, x) \in \llbracket 1, n \rrbracket \times [0, 1]^2 \times \mathbb{R}^d$ :

$$\Lambda_{k-1}^\beta(t, u, x) := \mathbb{E}\left[\text{Tr}\left(\left(D^3\varphi(x + \gamma_k b(x) + ut\sqrt{\gamma_k}\sigma(x)U_k)\sigma(x)U_k\right)(\sigma(x)U_k \otimes U_k\sigma(x)^*)\right)\right] \quad (2.1)$$

keeping in mind that, since  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ ,  $[D^3\varphi]_\beta < +\infty$ . We define subsequently:

$$E_n^\beta := \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^{3/2} \int_0^1 dt (1-t)t \int_0^1 du \Lambda_{k-1}^\beta(t, u, X_{k-1}). \quad (2.2)$$

Set now

$$\mathcal{B}_{n,\beta} := E_n^\beta, \text{ if } \beta \in (0, 1),$$

$$\begin{aligned} \mathcal{B}_{n,\beta} := & E_n^\beta + \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t) \text{Tr}\left(D^2\varphi(X_{k-1} + t\gamma_k b_{k-1})b_{k-1} \otimes b_{k-1}\right) dt \\ & + \frac{1}{2\sqrt{\Gamma_n}} \sum_{k=1}^n \gamma_k \text{Tr}\left(\left(D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})\right)\Sigma_{k-1}\right), \text{ if } \beta = 1. \end{aligned} \quad (2.3)$$

There exist two explicit monotonic sequences  $c_n \leq 1 \leq C_n$ ,  $n \geq 1$ , with  $\lim_n C_n = \lim_n c_n = 1$  such that for all  $n \geq 1$  and for every  $a > 0$ :

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta}| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}\right).$$

For  $\beta \in (0, 1)$ , the random variables  $|\mathcal{B}_{n,\beta}| = |E_n^\beta| \leq \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} \xrightarrow[n]{a_{\beta,\infty} > 0 \text{ a.s.}}$  Also, for  $\beta = 1$ , the  $(\mathcal{B}_{n,1})_{n \geq 1}$  are exponentially integrable and if, furthermore,  $D^3\varphi$  is  $\mathcal{C}^1$ ,  $\mathcal{B}_{n,1} \xrightarrow[n]{} -\tilde{\gamma}m$  a.s. where  $\tilde{\gamma}m$  is as in Theorem 1. In any case, a bias appears in our deviation controls when we consider, for a given smoothness of order  $\beta \in (0, 1]$  for  $D^3\varphi$ , the fastest associated time steps  $\gamma_k \asymp k^{-\theta}$ ,  $\theta = \frac{1}{2+\beta}$ .

**Remark 5** Observe that, when  $\beta = 1$ , the above result provides a non-asymptotic counterpart of the limit Theorem 1. In particular, the concentration constants appearing in Theorem 2 asymptotically match those of the centered CLT recalled in Theorem 1, up to a substitution of the asymptotic variance  $\int_{\mathbb{R}^d} |\sigma^* \nabla \varphi(x)|^2 \nu(dx)$  by its natural upper bound  $\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2$ .

Importantly, these bounds do not require “a priori” non-degeneracy conditions and only depend on the diffusion coefficient through the sup-norm of the diffusion matrix  $\Sigma$ , assumption **(C2)**. It will anyhow be very natural to consider a non-degeneracy condition ([PV01], [RS76], [Vil09]), or a confluence condition ([PP14]), when investigating the deviations for a given function  $f$ , in order to ensure the solvability of the corresponding Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$  and to derive explicit upper bounds for  $\|\nabla \varphi\|_\infty$  in terms of the coefficients  $b, \sigma$  and the source  $f$  which turn out to be crucial to design computable non-asymptotic confidence intervals. These aspects are discussed in Section 2.2 below.

The alternative form of the asymptotic variance (see Remark 3)  $\int_{\mathbb{R}^d} |\sigma^* \nabla \varphi(x)|^2 \nu(dx) = -2 \int_{\mathbb{R}^d} f(x) \varphi(x) \nu(dx)$  suggests that for bounded source terms  $f$ , an associated natural variance bound would be  $2\|f\|_\infty \|\varphi\|_\infty$ . Such a control would a priori require less regularity on  $\varphi$  than assumed in Theorem 2. One could for instance try to exploit suitable regularization procedures, like for instance the one proposed in Section 5.3 for the proof of Theorem 7 below, to establish non-asymptotic deviation results under weaker assumptions. Our main objective being to capture unbounded Lipschitz functions  $f$ , these aspects will concern further research.

**Remark 6 (Smoothness and Convergence Rate)** Observe that, in coherence with the asymptotic setting of the CLT recalled in Theorem 1, for a given  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ ,  $\beta \in (0, 1]$ , the fastest convergence rate for the deviations is attained for  $\theta = \frac{1}{2+\beta}$ . A bias appears, which can be difficult to estimate in practice since  $\varphi$  is usually unknown.

**Remark 7 (On the smoothness property of  $x \mapsto \langle b(x), \nabla \varphi(x) \rangle$ )** The Lipschitz continuity assumption on the above mapping appearing in case (i) might seem awkward at first sight. It is non-intrinsic in the sense that it involves both the drift  $b$  of the model and the test function  $\varphi$ . However, this condition naturally appears when  $\varphi$  is a smooth solution to the Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$ . Indeed, recalling the definition of  $\mathcal{A}$  in  $(\mathcal{L}_\nu)$ , iii), we can rewrite:

$$\langle \nabla \varphi(x), b(x) \rangle = f(x) - \nu(f) - \frac{1}{2} \text{Tr} \left( \Sigma(x) D^2 \varphi(x) \right).$$

Hence, the Lipschitz continuity of the function in the above left hand side readily follows as soon as the source  $f$  is Lipschitz and if  $D^2\varphi$  is bounded and Lipschitz continuous (since  $\sigma$  is also bounded and Lipschitz). Note that with the previous notations for function spaces

the previous conditions are implied if  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}) \subset \mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R})$ ,  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R}) \Rightarrow D^2\varphi \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d) \subset \mathcal{C}_b^{0,1}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ . We refer to Section 5.1 for details.

We now state an improvement of the previous concentration bound when  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  is itself a coboundary, i.e. when the Poisson problem  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$  can be solved with  $\vartheta$  satisfying the assumptions required for  $\varphi$  in Theorem 2. Precisely, we have the following result.

**Theorem 3** (a) *Under the assumptions of Theorem 2 and with the notations introduced therein, provided that  $\vartheta$  solution to the Poisson equation  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$  satisfies the same smoothness and growth conditions as  $\varphi$ , for  $\beta \in (0, 1]$  and  $\theta \in (\frac{1}{2+\beta}, 1]$  (unbiased case), there exist two explicit monotonic sequences  $\tilde{c}_n \leq 1 \leq \tilde{C}_n$ ,  $n \geq 1$ , with  $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$  such that for all  $n \geq 1$  for all  $0 < a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$  for a positive sequence  $\chi_n \xrightarrow[n]{\infty} 0$  arbitrarily slowly, so that  $\chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \rightarrow +\infty$ :*

$$\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2\tilde{C}_n \exp\left(-\tilde{c}_n \frac{a^2}{2\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2}\right). \quad (2.4)$$

(b) *If  $\vartheta$  solve the Poisson equation  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$  mutatis mutandis for a matrix norm dominating the operator norm ( $\|\sigma(x)\| \geq \|\sigma(x)\|$ ), then the above bound (2.4) still holds with  $\nu(\|\sigma\|^2)$  instead of  $\nu(\|\sigma\|^2)$ .*

Importantly, the above result allows to improve the natural variance bound  $\|\nabla\varphi\|_\infty^2\|\sigma\|_\infty^2$  of Theorem 2 by a more refined, namely  $\|\nabla\varphi\|_\infty^2\nu(\|\sigma\|^2)$ . Such a bound can be particularly interesting when the supremum norm of  $\sigma$  is *high* but its average w.r.t. the invariant distribution  $\nu$  significantly lower. We refer to Section 4, Theorem 8 (general form of Theorem 3) and 6.2 (numerical results) for further discussions on that topic.

Of course Claim (b) is less sharp than (a) stated with the operator norm  $\|\cdot\|$  but solving the Poisson equation for  $\|\sigma(x)\|$  seems highly non trivial. By contrast, if  $\|\sigma(x)\| = \|\sigma(x)\|_F := [\text{Tr}(\sigma\sigma^*(x))]^{1/2}$  stands for the Fröbenius norm, Theorem 4 below yields the expected smoothness properties on  $\|\sigma\|_F^2 - \nu(\|\sigma\|_F^2)$  that ensure the existence of a solution to  $\mathcal{A}\vartheta = \|\sigma\|_F^2 - \nu(\|\sigma\|_F^2)$  meeting the required smoothness conditions. The price to pay with such computable norms being that they usually induce some dependence on the dimension  $d$  on the estimates (observe e.g. for the identity matrix  $I_d$  of  $\mathbb{R}^d \otimes \mathbb{R}^d$ ,  $\|I_d\|_F = d^{1/2}$ ).

## 2.2 Uniqueness of the invariant distribution and Regularity issues for the Poisson problem

For our deviation analysis to work, we need to have the uniqueness of the invariant distribution  $\nu$  and to establish some pointwise controls on the solution of the associated Poisson equation. Namely, we need to have quantitative bounds on its derivatives and the associated Hölder continuity modulus up to order 3.

To do so, additionally to our main assumptions introduced for Theorem 2, we will work in the confluent setting. In dimension one, any ergodic diffusion is in *some sense* confluent (see [Kha80], Appendix of the English translation, Theorem 2.2 p. 308 and its alternative proof in [LPP15] Theorem 2). Here, we will suppose that the following condition holds:

• **Confluence Conditions**

$(\mathbf{D}_\alpha^p)$  We assume that there exists  $\alpha > 0$  and  $p \in (1, 2]$  such that for all  $x \in \mathbb{R}^d$ ,  $\xi \in \mathbb{R}^d$

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r \left( (p-2) \frac{|\langle D\sigma_{\cdot j}(x) \xi, \xi \rangle|^2}{|\xi|^2} + |D\sigma_{\cdot j} \xi|^2 \right) \leq -\alpha |\xi|^2,$$

where  $Db$  stands here for the Jacobian of  $b$ ,  $\sigma_{\cdot j}$  stands for the  $j^{\text{th}}$  column of the diffusion matrix  $\sigma$  and  $D\sigma_{\cdot j}$  for its Jacobian matrix.

Within the confluent framework, we will consider from now on two kinds of assumptions which first give the uniqueness of  $\nu$  and that can lead to the required smoothness and to computable gradient bounds, which are crucial since they are precisely the quantities appearing in the non-asymptotic Gaussian deviation controls as emphasized in the statement of Theorem 2.

- Strong Confluence condition and regularity of the coefficients, which means that the drift is sufficiently dominant in the dynamics and the coefficients are *smooth* (see assumption  $(\mathbf{C}_R)$  below). Note that these conditions may hold for degenerate diffusion coefficients.
- Non-degeneracy of the diffusion coefficient and mild confluence condition and smoothness on the coefficients (see assumption  $(\mathbf{C}_{UE})$  below).

Under a sufficiently strong confluence condition, i.e. when  $\alpha$  is large enough in  $(\mathbf{D}_\alpha^p)$ , and provided that the coefficients  $b, \sigma, f$  are sufficiently smooth, it is quite direct to derive, through stochastic flow techniques *à la Kunita*, the required pointwise bounds for the derivatives of the Feynman-Kac representation of the solution to the Poisson equation (see [PP14] and Section 5.1).

In the non-degenerate case, the main advantage is that we can alleviate some restrictions on  $\alpha$  and the smoothness assumptions on  $b, \sigma, f$  to benefit from an elliptic regularity bootstrap deriving from suitable Schauder estimates available in the current setting from the work by Krylov and Priola [KP10].

We now introduce a smoothness assumption on  $b, \sigma, f$  that will be useful in both the considered cases.

- **Smoothness of the coefficients and the source.** For  $k \in \{1, 3\}$  and  $\beta \in (0, 1)$  define  $(\mathbf{R}_{k,\beta})$  The coefficients in equation (1.1) are s.t.  $b \in \mathcal{C}^{k,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ ,  $\sigma \in \mathcal{C}_b^{k,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ . Also, the source  $f$  for which we want to estimate  $\nu(f)$  belong to  $\mathcal{C}^{k,\beta}(\mathbb{R}^d, \mathbb{R})$ .

With these assumptions at hand, we now introduce the first setting we consider.

- ◇ **The confluent and regular assumption  $(\mathbf{C}_R)$** , holds if  $(\mathbf{D}_\alpha^p)$ ,  $(\mathbf{R}_{3,\beta})$ , for some  $\beta \in (0, 1]$ , are in force and  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(3+\beta)-p}$  where  $\|D\sigma\|_\infty := \sup_{x \in \mathbb{R}^d} \left( \sum_{j=1}^d \|D\sigma_{\cdot j}(x)\|^2 \right)^{\frac{1}{2}}$  recalling that, for every  $j \in \llbracket 1, d \rrbracket$ ,  $\|D\sigma_{\cdot j}(x)\|$  stands for the operator norm of  $D\sigma_{\cdot j}(x)$ .

In particular, we do not impose in this case any additional structure condition on  $\sigma$  which can degenerate.

In our second main framework, we will assume some uniform ellipticity conditions.

• **Non-degeneracy Conditions.**

**(UE)** Uniform ellipticity. We assume that w.l.o.g. that  $r = d$  ( $r \geq d$  could also be considered) in (1.1) and that the diffusion coefficient  $\sigma$  is such that

$$\exists \underline{\sigma} > 0, \forall \xi \in \mathbb{R}^d, \langle \sigma \sigma^*(x) \xi, \xi \rangle \geq \underline{\sigma} |\xi|^2.$$

We now introduce our second main setting:

◊ **The confluent and non-degenerate assumption ( $\mathbf{C}_{\mathbf{UE}}$ )**, holds if  $(\mathbf{D}_\alpha^p)$ ,  $(\mathbf{R}_{1,\beta})$ , for some  $\beta \in (0, 1]$ , are in force. If  $d > 1$ , we also assume that  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$  and that the diffusion matrix  $\Sigma$  is such that, for all  $(i, j) \in \llbracket 1, d \rrbracket^2$ ,  $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$ .

**Theorem 4** *Assume that  $(\mathcal{L}_\nu)$  and either  $(\mathbf{C}_\mathbf{R})$  or  $(\mathbf{C}_{\mathbf{UE}})$  are in force. Then there exists a unique invariant distribution for the solution of (1.1), i.e. assumption  $(\mathbf{U})$  holds.*

*The associated Poisson equation*

$$\forall x \in \mathbb{R}^d, \mathcal{A}\varphi(x) = f(x) - \nu(f), \quad (2.5)$$

*admits a unique solution  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ ,  $\beta \in (0, 1)$  centered w.r.t.  $\nu$ . Furthermore, the following gradient bound holds*

$$\|\nabla\varphi\|_\infty \leq \frac{[f]_1}{\alpha},$$

*and the mapping  $x \mapsto \langle \nabla\varphi(x), b(x) \rangle$  is Lipschitz continuous.*

**About the regularity of the coefficients**

Under  $(\mathbf{C}_\mathbf{R})$ , the derivatives can be expressed using iterated tangent processes and we cannot hope, without *a priori* any non-degeneracy condition, for a smoothing effect to hold. To have  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ , we need to consider a source  $f \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  and the same smoothness on  $b, \sigma$  (Assumption  $(\mathbf{R}_{3,\beta})$ ). We refer to Section 5.1 for the proof of Theorem 4 under  $(\mathbf{C}_\mathbf{R})$ .

In the non-degenerate case, the solvability of the Poisson problem is usually studied in a Sobolev setting, see *e.g.* [PV01]. Let us also indicate that pointwise gradient bounds have been obtained by the same authors in [PV03] for bounded drifts and diffusion coefficients which are additionally supposed to be *smooth*, i.e. at least  $\mathcal{C}_b^{2,\gamma}$  with the notations introduced in paragraph 1.3. We point out that these estimates do not apply in our current setting in which the drift has typically linear growth.

We eventually mention the last paper by these authors, namely [PV05]. They derive therein the uniqueness of the martingale solution to the Poisson equation in a potentially degenerate setting under suitable local Doeblin conditions. In that framework, pointwise controls are obtained as well for the solution itself but not for its *derivatives*.

To obtain the required smoothness, we use here in the non-degenerate framework of  $(\mathbf{C}_{\mathbf{UE}})$  some Schauder estimates, deriving from the work of Krylov and Priola [KP10], which allow to benefit from the elliptic regularity. Namely, to obtain the mentioned

smoothness on  $\varphi$  solving  $\mathcal{A}\varphi = f - \nu(f)$ , that we expect to be in  $\mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ ,  $\beta \in (0, 1)$ , we can take a source  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$  and  $b \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ ,  $\sigma \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ .

We would eventually like to emphasize that the structure condition on  $\Sigma$  might seem weird at first sight. It is actually needed to decouple the PDEs formally satisfied by  $(\partial_{x_i}\varphi)_{i \in \llbracket 1, d \rrbracket}$  in order to exploit the *a priori* estimates of [KP10] established for scalar valued PDEs. We refer to Section 5.1 for a proof and details.

### About the confluence condition and the restrictions on $\sigma$

We work here in the confluent setting of  $(\mathbf{D}_\alpha^p)$ . This assumption will allow, through a pathwise analysis associated with the tangent flow, to derive a pointwise gradient bound. Another possibility to obtain such a bound is to assume a so-called Bakry and Émery curvature criterion, see [BE85, BGL14]. Under this condition, the gradient and semi-group commute up to an exponential multiplicative factor (see equation (2.7) below).

▷ *Bakry and Émery curvature criterion.* First, we recall that the “*carré du champ*” operator  $\Gamma$  of a Markov process with generator  $\mathcal{A}$  reads, for every  $f, g$  in its domain  $\mathcal{D}(\mathcal{A})$

$$\Gamma(f, g) := \frac{1}{2} \left( \mathcal{A}(fg) - f\mathcal{A}g - g\mathcal{A}f \right) \quad \text{and} \quad \Gamma(f) := \Gamma(f, f).$$

We also need to define the  $\Gamma_2$  operator

$$\Gamma_2(f) = \frac{1}{2} \left( \mathcal{A}\Gamma(f) - 2\Gamma(f, \mathcal{A}f) \right).$$

In our Brownian diffusion setting, we have

$$\forall x \in \mathbb{R}^d, \quad \Gamma(f)(x) = |\sigma^* \nabla f(x)|^2.$$

whereas the computation of  $\Gamma_2$  is significantly more involved. However, if the diffusion matrix  $\Sigma = \sigma\sigma^*$  is constant then:

$$\Gamma_2(f)(x) := \text{Tr}((D^2 f(x)\Sigma)^2) - \langle \nabla f, Db\Sigma\nabla f \rangle(x).$$

With these notations at hand, we say that *the semi-group  $(P_t)_{t \geq 0}$  of  $\mathcal{A}$  satisfies the Bakry and Émery curvature criterion with parameter  $\rho > 0$*  if

$$(\mathbf{BE}_\rho) \quad \forall f \in \mathcal{D}(\mathcal{A}), \quad \Gamma_2(f) \geq \rho \Gamma(f).$$

Observe that for  $\Sigma = I_d$  the condition  $(\mathbf{BE}_\rho)$  is actually equivalent to  $(\mathbf{D}_\alpha^p)$  with  $\alpha = \rho$  (and any  $p \in (1, 2]$  since  $D\sigma = 0$ ) and reads

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle \leq -\rho |\xi|^2.$$

The computation of the  $\Gamma_2$  for a general non-degenerate diffusion of the form (1.1) is not easy and is discussed in [ACJ08]. In particular, in whole generality, the computation of the  $\Gamma_2$  requires the coefficients of the operator itself to be *smooth* (i.e. at least  $\mathcal{C}^2$ ).

We also refer to [AMTU01] if the diffusion matrix is scalar diagonal, *i.e.*  $\Sigma(x) = \varsigma(x)I_d$ ,  $x \in \mathbb{R}^d$ , where  $\varsigma$  is real valued. In that case, it is then shown that  $(\mathbf{BE}_\rho)$  holds if and only if:

$$-\frac{1}{2}\langle (M(x) + M^*(x))\xi, \xi \rangle \leq -\rho\varsigma(x)|\xi|^2, \quad (2.6)$$

where

$$M(x) = \frac{1}{2}\left(\varsigma(x)\Delta\varsigma(x) + \langle b(x), \nabla\varsigma(x) \rangle - \|\nabla\varsigma(x)\|^2\right)I_d + \left(\frac{1}{2} - \frac{d}{4}\right)\nabla\varsigma \otimes \nabla\varsigma(x) - \varsigma(x)^2 Db(x).$$

An important property when  $(\mathbf{BE}_\rho)$  holds, see again [BE85], [BGL14], is that the following commutation inequality holds:

$$\forall t \geq 0, \forall x \in \mathbb{R}^d, \Gamma(P_t f)(x) \leq \exp(-2\rho t)P_t \Gamma(f). \quad (2.7)$$

To conclude, let us say that the Bakry-Emery curvature condition is a very powerful tool to derive pointwise gradient bounds. In our framework, this is unfortunately not enough as soon as  $d > 1$ , because additionally to this kind of bounds we also need, to enter in the framework of Schauder estimates under  $(\mathbf{C}_{\mathbf{UE}})$ , a control of the  $\beta$ -Hölder modulus of the gradient (see Section 5.1). It does not seem that the condition  $(\mathbf{BE}_\rho)$  helps to get such controls. The restrictions on the variations of  $D\sigma$  appearing in both assumptions  $(\mathbf{C}_{\mathbf{UE}})$  and  $(\mathbf{C}_{\mathbf{R}})$  are precisely needed to derive in the first case the bounds on  $[D\varphi]_\beta$  and in the second one to prove that the derivatives exist up to order 3 and that  $[D^3\varphi]_\beta$  is controlled as well. This explains why the conditions on  $D\sigma$  are more stringent in the potentially degenerate setting  $(\mathbf{C}_{\mathbf{R}})$ . In each case, those bounds are obtained through pathwise analysis and the restrictions on  $D\sigma$  ensure the time integrability of the iterated tangent flows, see again Section 5.1 and Appendix A in [PP14] for details.

## 2.3 Practical Deviation Bounds

**A first Non-asymptotic confidence interval result.**

**Theorem 5 (Non-asymptotic confidence intervals without bias)** *Let the assumptions of Theorem 4 be in force. Then, there exists a unique invariant distribution  $\nu$  for (1.1), *i.e.*  $(\mathbf{U})$  holds. Also,  $\varphi$  satisfies  $(\mathbf{G}_{\mathbf{V}})$  introduced in Theorem 2 for  $V(x) = 1 + |x|^2$ .*

*Assume that  $(\mathbf{C1})$  (sub-gaussian tails of the innovation) holds and that the step sequence  $(\gamma_k)_{k \geq 1}$  is such that  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in (\frac{1}{2+\beta}, 1]$ . Then, for  $(c_n)_{n \geq 1}, (C_n)_{n \geq 1}$  like in Theorem 2 with  $\lim_n c_n = \lim_n C_n = 1$ , we have that for all  $n \geq 1$  and  $a > 0$  and for any matrix norm  $\|\cdot\|$  dominating  $\|\cdot\|$ :*

$$\mathbb{P}\left[\sqrt{\Gamma_n}|\nu_n(f) - \nu(f)| > a\right] \leq 2C_n \exp\left(-c_n \frac{a^2 \alpha^2}{2\|\sigma\|_\infty^2 [f]_1^2}\right) \quad \text{with } \|\sigma\|_\infty := \sup_{x \in \mathbb{R}^r} \|\sigma(x)\|, \quad (2.8)$$

$$\mathbb{P}\left[\nu(f) \in \left[\nu_n(f) - \frac{a\|\sigma\|_\infty [f]_1}{\alpha\sqrt{\Gamma_n}}, \nu_n(f) + \frac{a\|\sigma\|_\infty [f]_1}{\alpha\sqrt{\Gamma_n}}\right]\right] \geq 1 - 2C_n \exp\left(-c_n \frac{a^2}{2}\right), \quad (2.9)$$

where the parameter  $\alpha$  is the same as in the pointwise gradient bound of Theorem 4.



**Proof:** Equation (2.8) is a direct consequence of Theorem 2 and the gradient bound in Theorem 4. Indeed, the mean-value Theorem readily yields that  $(\mathbf{G}_\nu)$  holds. It then suffice to observe that  $\nu_n(f) - \nu(f) = \nu_n(\mathcal{A}\varphi)$ . To prove (2.9), setting  $a_{\sigma,f,\alpha} := a \|\sigma\|_\infty \frac{[f]_1}{\alpha}$ , it suffices to write:

$$\mathbb{P} \left[ \nu(f) \in \left[ \nu_n(f) - \frac{a_{\sigma,f,\alpha}}{\sqrt{\Gamma_n}}, \nu_n(f) + \frac{a_{\sigma,f,\alpha}}{\sqrt{\Gamma_n}} \right] \right] = 1 - \mathbb{P}[\sqrt{\Gamma_n} |\nu_n(f) - \nu(f)| \geq a_{\sigma,f,\alpha}]$$

and conclude by (2.8).  $\square$

**A more refined non-asymptotic confidence interval when  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  is a coboundary.**

We provide in Theorem 6 below a kind of *Slutsky's Lemma* when, for a matrix norm  $\|\cdot\|$  dominating  $\|\sigma(x)\| \leq \|\sigma(x)\|$ , s.t.  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$  is a coboundary.

**Theorem 6 (Slutsky type concentration result for the coboundary case)** *Under the assumptions of Theorem 5, for  $\beta \in (0, 1]$  and  $\theta \in (\frac{1}{2+\beta}, 1]$  (unbiased case), assuming as well that there is a unique solution  $\vartheta$  to  $\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2)$  satisfying the same assumptions as  $\varphi$  in Theorem 5, there exist two explicit monotonic sequences  $c_n \leq 1 \leq C_n$ ,  $n \geq 1$ , with  $\lim_n C_n = \lim_n c_n = 1$  such that for all  $n \geq 1$ , for all  $a > 0$ , the following bounds hold: if  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$  (Gaussian deviations) then,*

$$\mathbb{P} \left[ \left| \sqrt{\Gamma_n} \frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}} \right| \geq a \right] \leq 2C_n \exp \left( -c_n \frac{a^2 \alpha^2}{2[f]_1^2} \right), \quad (2.10)$$

$$\mathbb{P} \left[ \nu(f) \in \left[ \nu_n(f) - \frac{a \sqrt{\nu_n(\|\sigma\|^2)} [f]_1}{\alpha \sqrt{\Gamma_n}}, \nu_n(f) + \frac{a \sqrt{\nu_n(\|\sigma\|^2)} [f]_1}{\alpha \sqrt{\Gamma_n}} \right] \right] \geq 1 - 2C_n \exp \left( -c_n \frac{a^2}{2} \right). \quad (2.11)$$

Again, the non-asymptotic confidence interval is explicitly computable in function of the given source  $f$ , the coefficients in the dynamics and the chosen (computable) matrix norm  $\|\cdot\|$ . It is also sharper than the one in (2.9).

**Towards Lipschitz sources in the non-degenerate case**

We conclude this section stating a non-asymptotic deviation result for Lipschitz sources under some non-degeneracy conditions (assumption  $(\mathbf{C}_{\text{UE}})$  of Theorem 4 replacing the condition stated there for  $f$  by a Lipschitz condition).

**Theorem 7 (Non-asymptotic concentration bounds for Lipschitz functions)**

*Let the assumptions of Theorem 4 with  $(\mathbf{C}_{\text{UE}})$  hold except that  $f$  is here solely a Lipschitz continuous function. For a time step sequence  $(\gamma_k)_{k \geq 1}$  of the form  $\gamma_k = k^{-\theta}$ ,  $\theta \in (1/2, 1]$ , we have that, there exist two explicit monotonic sequences  $c_n \leq 1 \leq C_n$ ,  $n \geq 1$ , with  $\lim_n C_n = \lim_n c_n = 1$  such that for all  $n \geq 1$  and for every  $a > 0$ :*

$$\mathbb{P} \left[ \left| \sqrt{\Gamma_n} (\nu_n(f) - \nu(f)) \right| \geq a \right] \leq 2C_n \exp \left( -c_n \frac{a^2 \alpha^2}{2 \|\sigma\|_\infty^2 [f]_1^2} \right) \quad (2.12)$$

where  $\alpha$  is as in Theorem 4.

Such estimates are important since they allow to get rather close to the natural framework which appear in functional inequalities (that mainly deal with Wasserstein distances and their possible deviations). Indeed, through the Monge-Kantorovich formulation, the Wasserstein distance involves Lipschitz functions, since it is precisely achieved taking the minimum over Lipschitz functions for all possible coupling with marginal corresponding to the arguments of the distance (see [BGL14]).

In the literature, some non-asymptotic bounds can be found for the deviations from its mean for the Wasserstein distance between the empirical measure of a homogeneous Markov chain and its stationary distribution (see Boissard [Boi11]). Here, we manage to get directly the non-asymptotic deviation bounds over all possible Lipschitz functions for the empirical measure of the scheme aiming directly to approximate the target stationary distribution of the diffusion. Handling the Wasserstein distance in our framework would amount to consider the supremum over the Lipschitz functions in the probability in (2.12). This will concern further research.

We eventually point out that Theorem 7 is obtained through regularization arguments of the source  $f$  exploiting the previous results of Theorems 2 and 4 (see Section 5.3 for details). This leads to a constraint on the steps, i.e.  $\gamma_n \asymp n^{-\theta}$ ,  $\theta \in (\frac{1}{2}, 1]$ . This is the price to pay, indeed a bigger  $\theta$  yields a lower convergence rate, to handle less regular Lipschitz sources. Also, to perform the approximation procedure we precisely need a kind of elliptic bootstrap (like in Theorem 4 under  $(\mathbf{C}_{UE})$ ). This is why we impose the non-degeneracy assumptions.

### 3 Proof of the concentration results (Theorem 2)

For notational convenience, we say that assumption  $(\mathbf{A})$  holds whenever  $(\mathbf{C1})$ ,  $(\mathbf{GC})$ ,  $(\mathbf{C2})$ ,  $(\mathcal{L}_V)$ ,  $(\mathbf{U})$  and  $(\mathbf{S})$  are fulfilled. We assume throughout this section that  $(\mathbf{A})$  is in force and that the function  $\varphi$  appearing in the lemmas satisfies the smoothness assumptions of Theorem 2.

#### 3.1 Strategy

To control the deviations of  $\nu_n(\mathcal{A}\varphi)$  we first give a decomposition lemma, obtained by a standard Taylor expansion. The idea is to perform a kind of *splitting* between the deterministic contributions in the transitions and the random innovations. Doing so, we manage to prove that the contributions involving the innovations can be gathered into conditionally Lipschitz continuous functions of the noise, with small Lipschitz constant (functions  $(\psi_k(X_{k-1}, \cdot))_{k \in \llbracket 1, n \rrbracket}$  below). These functions precisely give the Gaussian concentration, see Lemma 2. The other terms, that we will call from now on “remainders”, will be shown to be uniformly controlled w.r.t.  $n$  and do not give any asymptotic contribution in the “fast decreasing” case  $\theta > 1/(2 + \beta)$  (with the terminology of Theorem 2), see Lemmas 3, 4 and 5.

**Lemma 1 (Local Decomposition of the empirical measure )** *For all  $n \geq 1$  and*

$k \in \llbracket 0, n-1 \rrbracket$ :

$$\begin{aligned}
 & \varphi(X_k) - \varphi(X_{k-1}) \\
 = & \gamma_k \mathcal{A}\varphi(X_{k-1}) + \left[ \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \right. \\
 & \left. + \frac{1}{2} \gamma_k \operatorname{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1}^2 \right) + \psi_k(X_{k-1}, U_k) \right] \\
 =: & \gamma_k \mathcal{A}(X_{k-1}) + \left( \psi_k(X_{k-1}, U_k) + R_{n,k}^1(X_{k-1}) \right), \tag{3.13}
 \end{aligned}$$

where for all  $k \in \llbracket 1, n \rrbracket$ , conditionally to  $\mathcal{F}_{k-1}$ , the mapping  $u \mapsto \psi_k(X_{k-1}, u)$  is Lipschitz continuous in  $u$  with constant  $\sqrt{\gamma_k} \|\sigma_{k-1}\| \|\nabla \varphi\|_\infty$ .

Introducing for a given  $k$ , the mapping

$$u \mapsto \Delta_k(X_{k-1}, u) := \psi_k(X_{k-1}, u) - \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}],$$

we then rewrite:

$$\varphi(X_k) - \varphi(X_{k-1}) = \gamma_k \mathcal{A}\varphi(X_{k-1}) + \Delta_k(X_{k-1}, U_k) + R_{n,k}(X_{k-1}),$$

with  $R_{n,k}(X_{k-1}) := R_{n,k}^1(X_{k-1}) + \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]$ . The contribution  $\Delta_k(X_{k-1}, U_k)$  can be viewed as a martingale increment. Introduce now the associated (true) martingale

$$M_n := \sum_{k=1}^n \Delta_k(X_{k-1}, U_k). \tag{3.14}$$

Summing over  $k$  yields:

$$\varphi(X_n) - \varphi(X_0) = \Gamma_n \nu_n(\mathcal{A}\varphi) + M_n + \sum_{k=1}^n R_{n,k}(X_{k-1}). \tag{3.15}$$

Defining  $R_n := \sum_{k=1}^n R_{n,k}(X_{k-1}) + \varphi(X_0) - \varphi(X_n)$  we obtain the following decomposition of the empirical measure:

$$\nu_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n} (M_n + R_n). \tag{3.16}$$

- *Unbiased Case (Sub-Optimal Convergence Rate)*. This case corresponds to *fast* decreasing steps of the form  $\gamma_k \asymp k^{-\theta}$ ,  $\theta > 1/(2 + \beta)$ . To investigate the non-asymptotic deviations of the empirical measure, the idea is now to write for  $a, \lambda > 0$ :

$$\begin{aligned}
 \mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] & \leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(-\frac{\lambda}{\Gamma_n}(M_n + R_n)\right)\right] \\
 & \leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right]^{1/q} \mathbb{E}\left[\exp\left(\frac{p\lambda}{\Gamma_n} |R_n|\right)\right]^{1/p}, \tag{3.17}
 \end{aligned}$$

where  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $p, q > 1$ . We actually aim to choose  $q := q(n) \xrightarrow[n]{} 1$ . For a suitable choice of  $q$  satisfying the previous condition, we manage, in the fast decreasing case, to show that  $\mathcal{R}_n := \mathbb{E}[\exp(\frac{p\lambda}{\Gamma_n}|R_n|)]^{1/p} \xrightarrow[n]{} 1$ . For the term involving the martingale  $M_n$  we actually use the Gaussian concentration property (**GC**) of the innovation on its increments  $(\Delta_k(X_{k-1}, U_k))_{k \in \llbracket 1, n \rrbracket}$ . Namely, using the control of the Lipschitz constant of  $\Delta_k(X_{k-1}, \cdot)$  stated in Lemma 1, we derive:

$$\begin{aligned} \mathbb{E} \left[ \exp \left( - \frac{q\lambda}{\Gamma_n} M_n \right) \right] &= \mathbb{E} \left[ \exp \left( - \frac{q\lambda}{\Gamma_n} M_{n-1} \right) \mathbb{E} \left[ \exp \left( - \frac{q\lambda}{\Gamma_n} \Delta_{n-1}(X_{n-1}, U_n) \right) \mid \mathcal{F}_{n-1} \right] \right] \\ &\leq \mathbb{E} \left[ \exp \left( - \frac{q\lambda}{\Gamma_n} M_{n-1} \right) \right] \exp \left( \frac{\lambda^2 q^2}{2\Gamma_n^2} \gamma_n \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2 \right) \\ &\leq \exp \left( \frac{\lambda^2 q^2}{2\Gamma_n} \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2 \right), \end{aligned} \quad (3.18)$$

iterating the procedure to derive the last identity. From (3.17), we thus get:

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq \mathcal{R}_n \exp \left( - \frac{a\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2 q}{2\Gamma_n} \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2 \right).$$

Keeping in mind that we manage to find  $q := q(n) \downarrow_n 1$  such that the remainder  $\mathcal{R}_n \downarrow_n 1$ , the result of Theorem 2 in the considered case then follows from a quadratic optimization over the parameter  $\lambda$ .

- *Biased Case (Optimal Convergence Rate)*. This case corresponds to *slow* decreasing steps of the form  $\gamma_k \asymp k^{-\theta}$ ,  $\theta = 1/(2 + \beta)$ . In this setting, some terms of the remainder  $R_n$  in (3.16) give a non trivial asymptotic contribution. We choose to substract them before studying the deviation (term  $\mathcal{B}_{n,\beta}$  in (2.3)).

### 3.2 Explicit controls on the remainders

Summing the increments appearing in (3.13), we now choose for the analysis to write for a given  $n \in \mathbb{N}$  the remainder  $R_n$  defined after (3.15) as

$$R_n = \sum_{k=1}^n R_{n,k}(X_{k-1}) + \varphi(X_0) - \varphi(X_n) = (D_{2,b,n} + D_{2,\Sigma,n}) + \bar{G}_n - L_n,$$

where:

$$\begin{aligned} D_{2,b,n} &:= \sum_{k=1}^n \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\ D_{2,\Sigma,n} &:= \frac{1}{2} \sum_{k=1}^n \gamma_k \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1}^2 \right), \\ \bar{G}_n &:= \sum_{k=1}^n \mathbb{E}[\psi_k(X_{k-1}, U_k) \mid \mathcal{F}_{k-1}], \\ L_n &:= \varphi(X_n) - \varphi(X_0). \end{aligned} \quad (3.19)$$

We refer to the proof of Lemma 1 to check that the above definition of  $\bar{G}_n$  actually matches the term  $\sqrt{\Gamma_n}E_n^\beta$  introduced in equation (2.2) of Theorem 2. We rewrite from (3.16)

$$\nu_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n}(M_n + R_n) = -\frac{1}{\Gamma_n}(M_n + (D_{2,b,n} + D_{2,\Sigma,n}) + \bar{G}_n - L_n). \quad (3.20)$$

We now split the analysis according to the cases (a) and (b) introduced in Theorem 2.

(a)  $\theta \in (1/(2 + \beta), 1]$ ,  $\beta \in (0, 1]$ . From (3.20), the exponential Tchebychev and Hölder inequalities yield that, for all  $\lambda \in \mathbb{R}_+$  and all  $p, q \in (1, +\infty)$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ ,

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a] &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n}M_n\right)\right)^{\frac{1}{q}} \\ &\times \left(\mathbb{E} \exp\left(\frac{2p\lambda}{\Gamma_n}(|L_n| + |\bar{G}_n|)\right)\right)^{\frac{1}{2p}} \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n}|D_{2,b,n}|\right)\right)^{\frac{1}{4p}} \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n}|D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p}}. \end{aligned} \quad (3.21)$$

(b)  $\theta = \frac{1}{2+\beta}$ ,  $\beta \in (0, 1]$ . If  $\beta = 1$ , denoting,  $D_{2,n} := D_{2,b,n} + D_{2,\Sigma,n}$ , we have from (3.19) and with the notations of (2.3),  $(\bar{G}_n + D_{2,n}) = \sqrt{\Gamma_n}\mathcal{B}_{n,1}$ . We study the deviations of:

$$\begin{aligned} &\mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta} \geq a] \\ &= \mathbb{P}\left[\nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n + D_{2,n}}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n}M_n\right)\right)^{\frac{1}{q}} \left(\mathbb{E} \exp\left(\frac{p\lambda}{\Gamma_n}|L_n|\right)\right)^{\frac{1}{p}}. \end{aligned} \quad (3.22)$$

For  $\beta \in (0, 1)$ , the contributions of  $D_{2,n}$  do not yield any asymptotic bias. Recalling from (2.3) that  $\mathcal{B}_{n,\beta} = E_n^\beta = \frac{\bar{G}_n}{\sqrt{\Gamma_n}}$ , we write:

$$\begin{aligned} &\mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta} \geq a] \\ &= \mathbb{P}\left[\nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}}\right] \leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n}M_n\right)\right)^{\frac{1}{q}} \\ &\times \left(\mathbb{E} \exp\left(\frac{2p\lambda}{\Gamma_n}|L_n|\right)\right)^{\frac{1}{2p}} \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n}|D_{2,b,n}|\right)\right)^{\frac{1}{4p}} \left(\mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n}|D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p}}. \end{aligned} \quad (3.23)$$

**Remark 8** Observe that in case (a), the “small steps” and the corresponding sufficient smoothness of  $\varphi$  prevent from the appearance of a bias. As a result, the concentration bound is, at the non-asymptotic level, the same as in Theorem 1, up to the additional upper-bound for the variance. In case (b), we subtract the terms  $\mathcal{B}_{n,\beta}$  that asymptotically give a bias. When  $\beta = 1$ , this is the case for both terms  $\frac{\bar{G}_n}{\Gamma_n}, \frac{D_{2,n}}{\Gamma_n}$ . Also, for  $D^3\varphi \in \mathcal{C}^1$ ,  $\mathcal{B}_{n,1} = \frac{\bar{G}_n + D_{2,n}}{\sqrt{\Gamma_n}} \xrightarrow[n]{-}$   $-\tilde{\gamma}m$  introduced in Theorem 1. For  $\beta \in (0, 1)$  and  $\varphi \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$ ,  $[\varphi^{(3)}]_\beta < +\infty$ , the only term giving a bias is  $\mathcal{B}_{n,\beta} = E_n^\beta = \frac{\bar{G}_n}{\sqrt{\Gamma_n}}$ .

The lemma below provides the Gaussian contribution to be exploited in inequalities (3.21) - (3.23).

**Lemma 2 (Gaussian concentration)** For  $a > 0$ ,  $q \in (1, +\infty)$ , setting

$$\lambda_n := \frac{a}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\sqrt{\Gamma_n}, \quad (3.24)$$

we derive:

$$\exp\left(-\lambda_n\frac{a}{\sqrt{\Gamma_n}}\right)\left(\mathbb{E}\exp\left(-\frac{q\lambda_n}{\Gamma_n}M_n\right)\right)^{\frac{1}{q}} \leq \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2}\right).$$

**Lemma 3 (Bounds for the Conditional Expectations)** With the above notations, we have that for  $\beta \in (0, 1]$ ,  $\theta \in [\frac{1}{2+\beta}, 1]$ :

$$|E_n^\beta| = \frac{|\bar{G}_n|}{\sqrt{\Gamma_n}} \leq a_n := \frac{[\varphi^{(3)}]_\beta\|\sigma\|_\infty^{(3+\beta)}\mathbb{E}[|U_1|^{3+\beta}]\Gamma_n^{\frac{(3+\beta)}{2}}}{(1+\beta)(2+\beta)(3+\beta)\sqrt{\Gamma_n}}, \text{ a.s.}$$

Moreover,  $a_n \xrightarrow{n} a_\infty$ , with  $a_\infty = 0$  if  $\theta \in (\frac{1}{2+\beta}, 1]$  and  $a_\infty > 0$  if  $\theta = \frac{1}{2+\beta}$ . Also, for  $\beta \in (0, 1]$ ,  $\theta \in (\frac{1}{2+\beta}, 1]$ :

$$\left(\mathbb{E}\exp\left(\frac{2p\lambda_n}{\Gamma_n}|\bar{G}_n|\right)\right)^{\frac{1}{2p}} \leq \exp\left(\frac{\lambda_n}{\sqrt{\Gamma_n}}a_n\right) \leq \exp\left(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{a_n^2 p}{2}\right), \quad \forall p > 1. \quad (3.25)$$

As indicated before, we now aim at controlling the remainders. In particular, from (3.17) and (3.19), we are led to handle terms of the form

$$\mathbb{E}\exp\left(c\sum_{k=1}^n\gamma_k^2|b(X_{k-1})|^2\right) \stackrel{(\mathcal{L}_V)}{\leq} \mathbb{E}\exp\left(cC_V\sum_{k=1}^n\gamma_k^2|V(X_{k-1})|\right)$$

for small enough real constants  $c > 0$ .

To this end, we will thoroughly rely on the following important integrability result for the Lyapunov function.

**Proposition 1** Under **(A)** there is a constant  $c_V := c_V((\mathbf{A})) > 0$  such that for all  $\lambda \in [0, c_V]$ ,  $\xi \in [0; 1]$ :

$$I_V^\xi := \sup_{n \geq 0} \mathbb{E}[\exp(\lambda V_n^\xi)] < +\infty.$$

We now have the following results for the terms appearing in (3.19).

**Lemma 4 (Initial term)** Let  $q \in (1, +\infty)$  be fixed and  $\lambda_n$  be as in (3.24) in Lemma 2. For functions  $\varphi$  satisfying **(G<sub>V</sub>)**, i.e. there exists  $C_{V,\varphi} > 0$  such that for all  $x \in \mathbb{R}^d$ ,  $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$ , for  $p := \frac{q}{q-1}$  and  $j \in \{1, 2\}$ :

$$\begin{aligned} \left(\mathbb{E}\exp\left(jp\lambda_n\frac{|L_n|}{\Gamma_n}\right)\right)^{\frac{1}{jp}} &\leq (I_V^1)^{\frac{1}{jp}} \exp\left(\frac{(j+1)pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p}\right) \\ &= (I_V^1)^{\frac{1}{jp}} \exp\left(\frac{(j+1)pC_{V,\varphi}^2a^2}{c_Vq^2\|\sigma\|_\infty^4\|\nabla\varphi\|_\infty^4\Gamma_n} + \frac{c_V}{p}\right), \end{aligned}$$

with  $c_V$ ,  $I_V^1$  like in Proposition 1.

**Lemma 5 (Remainders)** *Let  $q \in (1, +\infty)$  be fixed and  $\lambda_n$  be as in Lemma 2. Then, there exists  $C_{3.26} := C_{3.26}((\mathbf{A}), \varphi)$  such that for  $p = \frac{q}{q-1}$ :*

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}| \right) \right)^{\frac{1}{4p}} \leq \exp \left( C_{3.26} \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}}. \quad (3.26)$$

We also have:

- *If the mapping  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  is Lipschitz continuous, then there exists  $C_{3.27} := C((\mathbf{A}), \varphi) > 0$  such that*

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} \leq \exp \left( C_{3.27} \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}}. \quad (3.27)$$

- *For  $a \leq \frac{c_v q}{4C_V p} \frac{\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2 \sqrt{\Gamma_n}}{\|D^2 \varphi\|_\infty^2 \Gamma_n^{(2)}}$ , there exists an  $\mathbb{R}^+$ -valued sequence  $(v_n)_{n \geq 1}$  such that  $|v_n| \leq C_{3.28} := C_{3.28}((\mathbf{A}), \varphi)$  and*

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} \leq (I_V^1)^{v_n}. \quad (3.28)$$

Also,  $v_n \rightarrow v_\infty$  where  $v_\infty = 0$  if  $\theta > 1/3$  and  $v_\infty > 0$  for  $\theta = 1/3$ .

### Proof of Theorem 2.

From Lemma 2 we get:

$$\left( \mathbb{E} \exp \left( -q\lambda_n \frac{M_n}{\Gamma_n} \right) \right)^{\frac{1}{q}} \exp \left( -\frac{a\lambda_n}{\sqrt{\Gamma_n}} \right) \leq \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2} \right). \quad (3.29)$$

(a) We deal with the case  $\beta \in (0, 1], \theta \in (\frac{1}{2+\beta}, 1]$ .

(i) We suppose that the mapping  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  is Lipschitz continuous. Plugging in (3.21) the controls from (3.29), Lemma 3 equation (3.25), Lemma 4 (with  $j = 2$ ) and Lemma 5 (equations (3.26), (3.27)), we get:

$$\begin{aligned} & \mathbb{P} \left[ \nu_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}} \right] \\ & \leq \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2} \right) \exp \left( \frac{\lambda_n^2}{2\Gamma_n p} + \frac{pa_n^2}{2} \right) \exp \left( \frac{3pC_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2} + \frac{c_V}{p} \right) \\ & \quad \times (I_V^1)^{\frac{1}{2p}} \exp \left( C_{3.26} \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}} \times \exp \left( C_{3.27} \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}} \\ & \leq \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2} \left( 1 - \frac{1}{q\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2} \left\{ \frac{p}{\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + 2[C_{3.26} + C_{3.27}] (\Gamma_n^{(2)})^2 \right) + \frac{1}{p} \right\} \right) \right) \\ & \quad \times (I_V^1)^{\frac{1}{p}} \exp \left( \frac{c_V}{p} + \frac{pa_n^2}{2} \right). \end{aligned} \quad (3.30)$$

Recall now that for  $\theta > \frac{1}{2+\beta} \geq 1/3$ ,  $\Gamma_n^{(\frac{3+\beta}{2})}/\sqrt{\Gamma_n} \xrightarrow[n]{n} 0$ ,  $\Gamma_n^{(2)}/\sqrt{\Gamma_n} \xrightarrow[n]{n} 0$  (see Lemma 3 and Remark 2). We now take  $p := p_n \xrightarrow[n]{n} +\infty$ , and therefore  $q := q_n \xrightarrow[n]{n} 1$ , such that  $p_n^{1/2} \frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \xrightarrow[n]{n} 0$  so that from Lemma 3,  $p_n a_n^2 \xrightarrow[n]{n} 0$ . Since  $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \geq \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}$  this in turn implies:

$$d_n := \frac{1}{q_n \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} \left\{ \frac{p_n}{\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + [2C_{3.26} + 3C_{3.27}](\Gamma_n^{(2)})^2 \right) + \frac{1}{p_n} \right\} \xrightarrow[n]{n} 0. \quad (3.31)$$

We conclude from (3.30) setting  $c_n = q_n^{-1}(1-d_n)$ ,  $C_n := (I_V^1)^{\frac{1}{p_n}} \exp\left(\frac{1}{p_n} [c_V + \frac{C_{3.27}}{2}] + \frac{p_n a_n^2}{2}\right) \xrightarrow[n]{n} 1$ . Observe that taking an increasing sequence  $(p_n)_{n \geq 1}$  readily yields  $C_n \downarrow_n 1$ , and  $q_n \downarrow_n 1$ . Also, the sequence  $(p_n)_{n \geq 1}$  can be chosen in order to have, for  $n$  large enough,  $d_n \downarrow_n 0$  so that  $c_n \uparrow_n 1$ .

(ii) Assume  $a \leq \frac{c_V q}{4C_V p} \frac{\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 \sqrt{\Gamma_n}}{\|D^2\varphi\|_\infty^2 \Gamma_n^{(2)}}$ . Plugging in (3.21) the controls from (3.29), Lemma 3, equation (3.25), Lemmas 4 (with  $j = 2$ ), 5 (equations (3.26), (3.28)) we then derive:

$$\begin{aligned} & \mathbb{P}\left[\nu_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ & \leq \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right) \exp\left(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{p a_n^2}{2}\right) \exp\left(\frac{3p C_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2} + \frac{c_V}{p}\right) \\ & \quad \times (I_V^1)^{\frac{1}{2p}} \exp\left(C_{3.26} \frac{p \lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}} (I_V^1)^{v_n} \\ & \leq (I_V^1)^{v_n + \frac{3}{4p}} \exp\left(\frac{c_V}{p} + \frac{p a_n^2}{2}\right) \\ & \quad \exp\left(-\frac{a^2}{2q\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} \left(1 - \frac{1}{q\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} \left\{ \frac{p}{\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + 2C_{3.26}(\Gamma_n^{(2)})^2 \right) + \frac{1}{p} \right\}\right)\right). \end{aligned} \quad (3.32)$$

Since  $\theta > \frac{1}{2+\beta} \geq 1/3$  (see Remark 2), we again take  $p := p_n \uparrow_n +\infty$  so that  $p_n^{1/2} a_n \xrightarrow[n]{n} 0$  which also guarantees:

$$d_n := \frac{1}{q_n \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} \left\{ p_n \left( \frac{6C_{V,\varphi}^2}{c_V \Gamma_n} + \frac{2C_{3.26}(\Gamma_n^{(2)})^2}{\Gamma_n} \right) + \frac{1}{p_n} \right\} \xrightarrow[n]{n} 0. \quad (3.33)$$

In this case, we derive the result by setting  $c_n := q_n^{-1}(1-d_n) \xrightarrow[n]{n} 1$ , and  $C_n := (I_V^1)^{v_n + \frac{3}{4p_n}} \exp\left(\frac{c_V}{p_n} + \frac{p_n a_n^2}{2}\right) \xrightarrow[n]{n} 1$  (see the limits of  $v_n$  following equation (3.28) and (3.46)). Again,  $(p_n)_{n \geq 1}$  can be chosen in order to have the stated monotonicity for  $n$  large enough. Set now

$$\chi_n := \frac{c_V \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 q_n}{4C_V \|D^2\varphi\|_\infty^2 p_n}, \quad (3.34)$$

so that  $a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$ . Thus, the slower  $p_n$  goes to infinity, the wider the domain of validity for the estimate in the parameter  $a$ .



(b) It remains to analyze the case  $\beta \in (0, 1]$ ,  $\theta = \frac{1}{2+\beta}$ . Let us deal with  $\beta = 1$ . From (3.22), the controls of (3.29) and Lemma 4 (with  $j = 1$ ) we get:

$$\mathbb{P} \left[ \nu_n(\mathcal{A}\varphi) + \frac{\bar{G}_n + D_{2,n}}{\Gamma_n} \geq \frac{a}{\sqrt{\Gamma_n}} \right] \leq \exp \left( -\frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \right) \exp \left( \frac{2pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p} \right) (I_V^1)^{\frac{1}{p}}.$$

Recalling the definition of  $\lambda_n$  in (3.24), we conclude as previously with obvious modifications of  $(c_n)_{n \geq 1}$ ,  $(C_n)_{n \geq 1}$ . The case  $\beta \in (0, 1)$  is handled similarly starting from (3.23).

Also, when  $D^3\varphi \in \mathcal{C}^1$ , we derive similarly to the proof of Theorem 10 in [LP02] that  $\mathcal{B}_{n,1} \xrightarrow[n]{} -\tilde{\gamma}m$ .

Eventually, the final control involving the two sided deviation is derived by symmetry.

□

### 3.3 Proof of the Technical Lemmas

This section is devoted to the proof of the previously used Lemmas 1–5 and Proposition 1 which were the key ingredients to derive Theorem 2.

#### Proof of Lemma 1.

For  $k \in \llbracket 1, n \rrbracket$ , we first write:

$$\begin{aligned} \varphi(X_k) - \varphi(X_{k-1}) &= (\varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1})) + (\varphi(X_{k-1} + \gamma_k b_{k-1}) - \varphi(X_{k-1})) \\ &=: T_{k-1,r}(\varphi) + T_{k-1,d}(\varphi), \end{aligned} \tag{3.35}$$

in order to split the random and deterministic contributions in the transitions of the scheme (1.2).

We then perform a Taylor expansion with integral remainder at order 2 for the function  $\varphi$  in the two terms of the r.h.s. of (3.35). Namely, with the above notations:

$$\begin{aligned} T_{k-1,d}(\varphi) &= \gamma_k b_{k-1} \cdot \nabla\varphi(X_{k-1}) + \gamma_k \int_0^1 \langle \nabla\varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla\varphi(X_{k-1}), b_{k-1} \rangle dt, \\ T_{k-1,r}(\varphi) &= \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &\quad + \gamma_k \int_0^1 (1-t) \text{Tr} \left( D^2\varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right) dt. \end{aligned}$$

Hence,

$$\begin{aligned}
 & \varphi(X_k) - \varphi(X_{k-1}) \\
 = & \gamma_k \mathcal{A} \varphi(X_{k-1}) \\
 & + \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\
 & + \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\
 & + \gamma_k \int_0^1 (1-t) \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* - D^2 \varphi(X_{k-1}) \Sigma_{k-1} \right) dt \\
 = & \gamma_k \mathcal{A} \varphi(X_{k-1}) + \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\
 & + \gamma_k \int_0^1 (1-t) \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1} \right) dt + \psi_k(X_{k-1}, U_k) \\
 =: & \gamma_k \mathcal{A} \varphi(X_{k-1}) + D_{2,b}^k + D_{2,\Sigma}^k + \psi_k(X_{k-1}, U_k), \tag{3.36}
 \end{aligned}$$

where

$$\begin{aligned}
 \psi_k(X_{k-1}, U_k) & = \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\
 & + \gamma_k \int_0^1 (1-t) \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right. \\
 & \left. - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right) dt. \tag{3.37}
 \end{aligned}$$

Observe now that, conditionally to  $\mathcal{F}_{k-1}$ , the mapping  $u \mapsto \psi_k(X_{k-1}, u)$  is Lipschitz continuous: indeed, the innovation  $U_k$  does not appear in the other contributions of the right side of (3.36). Consequently, as  $\varphi$  is Lipschitz continuous we derive, for all  $(u, u') \in (\mathbb{R}^d)^2$ :

$$|\psi_k(X_{k-1}, u) - \psi_k(X_{k-1}, u')| \leq \sqrt{\gamma_k} \|\sigma_{k-1}\| \|\nabla \varphi\|_\infty |u - u'|.$$

The result is obtained by summing up the previous identities from  $k = 1$  to  $n$ , observing, with the notations of (3.19), that  $L_n = \sum_{k=1}^n \varphi(X_k) - \varphi(X_{k-1})$ ,  $D_{2,b,n} = \sum_{k=1}^n D_{2,b}^k$ ,  $D_{2,\Sigma,n} = \sum_{k=1}^n D_{2,\Sigma}^k$ ,  $G_n := \sum_{k=1}^n \psi_k(X_{k-1}, U_k)$ .  $\square$

### Proof of Lemma 2.

The idea is to use conditionally and iteratively the Gaussian concentration property **(GC)** of the innovation. Let us note that this strategy was already the key ingredient in [FM12]. In the current framework, we exploit that the functions  $u \mapsto \Delta_k(X_{k-1}, u) := \psi_k(X_{k-1}, u) - \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]$  are conditionally independent w.r.t.  $\mathcal{F}_{k-1}$  and Lips-

chitz continuous with constant  $\sqrt{\gamma_k}\|\sigma\|_\infty\|\nabla\varphi\|_\infty$  by Lemma 1. We thus write:

$$\begin{aligned}
 \mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right) &= \mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^n \Delta_k(X_{k-1}, U_k)\right) \\
 &= \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^{n-1} \Delta_k(X_{k-1}, U_k)\right) \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} \Delta_n(X_{n-1}, U_n)\right) \middle| \mathcal{F}_{n-1}\right]\right] \\
 &\leq \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^{n-1} \Delta_k(X_{k-1}, U_k)\right) \exp\left(\frac{q^2\lambda^2}{2\Gamma_n^2} \gamma_n \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2\right)\right], \quad (3.38)
 \end{aligned}$$

where we used **(GC)** in the third line recalling as well that  $\mathbb{E}[\Delta_n(X_{n-1}, U_n) | \mathcal{F}_{n-1}] = 0$ .

Iterating the process over  $k$ , we obtain:

$$\left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} = \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} \sum_{k=1}^n \Delta_k(X_{k-1}, U_k)\right)\right)^{\frac{1}{q}} \leq \exp\left(\frac{q\lambda^2 \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}{2\Gamma_n}\right). \quad (3.39)$$

Finally,

$$\exp\left(-\frac{\lambda a}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} \leq \exp\left(\frac{g(\lambda)}{\sqrt{\Gamma_n}}\right),$$

where  $g : \mathbb{R}^+ \rightarrow \mathbb{R}$  is defined by  $g(\lambda) = -\frac{a}{\sqrt{\Gamma_n}}\lambda + \frac{q\lambda^2}{2\Gamma_n} \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2$ . As  $a > 0$ , the function attains its minimum at  $\lambda_n$  given in (3.24). This eventually yields the expected bound.  $\square$

### Proof of Lemma 3.

From the definition in (3.37) and the Fubini theorem, we have that for all  $k \in \llbracket 1, n \rrbracket$ :

$$\begin{aligned}
 &\mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}] \\
 &= \gamma_k \int_0^1 (1-t) \text{Tr}\left(\mathbb{E}[D^2\varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k}\sigma_{k-1}U_k)\sigma_{k-1}U_k \otimes U_k\sigma_{k-1}^* \right. \\
 &\quad \left. - D^2\varphi(X_{k-1} + \gamma_k b_{k-1})\Sigma_{k-1} | \mathcal{F}_{k-1}]\right) dt. \quad (3.40)
 \end{aligned}$$

Recalling that  $U_k$  has the same moments as the standard Gaussian random variable up

to order three (see **(GC)**) and is independent of  $\mathcal{F}_{k-1}$ , a Taylor expansion yields:

$$\begin{aligned}
 & \mathbb{E} \left[ \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right. \right. \\
 & \quad \left. \left. - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right) \middle| \mathcal{F}_{k-1} \right] \\
 &= \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \sigma_{k-1} \mathbb{E} [U_k \otimes U_k] \sigma_{k-1}^* \right) \\
 & \quad + \int_0^1 \mathbb{E} \left[ \text{Tr} \left( (D^3 \varphi(X_{k-1} + \gamma_k b_{k-1} + ut\sqrt{\gamma_k} \sigma_{k-1} U_k) t\sqrt{\gamma_k} \sigma_{k-1} U_k) (\sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^*) \right) \middle| \mathcal{F}_{k-1} \right] du \\
 & \quad - \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right) \\
 &= \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \sigma_{k-1} \underbrace{(\mathbb{E} [U_k \otimes U_k] - I)}_{=0} \sigma_{k-1}^* \right) \\
 & \quad + t\sqrt{\gamma_k} \int_0^1 \mathbb{E} \left[ \text{Tr} \left( ([D^3 \varphi(X_{k-1} + \gamma_k b_{k-1} + ut\sqrt{\gamma_k} \sigma_{k-1} U_k) - D^3 \varphi(X_{k-1} + \gamma_k b_{k-1})] \sigma_{k-1} U_k) \right. \right. \\
 & \quad \left. \left. \times (\sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^*) \right) \middle| \mathcal{F}_{k-1} \right] du,
 \end{aligned}$$

recalling from **(GC)** that for all  $(i, j, l) \in \llbracket 1, r \rrbracket$ ,  $\mathbb{E} [U_k^i U_k^j U_k^l | \mathcal{F}_{k-1}] = \mathbb{E} [U_1^i U_1^j U_1^l] = 0$  (cancellation argument). Hence,

$$\begin{aligned}
 |\mathbb{E} [\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]| &\leq \gamma_k \int_0^1 (1-t) t^{1+\beta} [\varphi^{(3)}]_\beta \mathbb{E} \left[ \gamma_k^{\frac{1+\beta}{2}} \|\sigma_{k-1}\|^{3+\beta} |U_k|^{3+\beta} \int_0^1 u^\beta du \middle| \mathcal{F}_{k-1} \right] dt \\
 &= \frac{[\varphi^{(3)}]_\beta \gamma_k^{\frac{3+\beta}{2}} \|\sigma_{k-1}\|^{3+\beta} \mathbb{E} [|U_k|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)},
 \end{aligned}$$

recalling that the third derivatives of  $\varphi$  are  $\beta$ -Hölder continuous for the first inequality. We thus derive:

$$|E_n^\beta| = \frac{|\bar{G}_n|}{\sqrt{\Gamma_n}} \leq \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^{n-1} |\mathbb{E} [\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]| \leq \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E} [|U_1|^{3+\beta}] \Gamma_n^{\left(\frac{3+\beta}{2}\right)}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} =: a_n.$$

### Proof of Proposition 1.

First of all, let us decompose the Lyapunov function  $V$  with a Taylor expansion like in Lemma 1. We again use a splitting between the deterministic contributions and those

involving the innovation. We write for all  $n \in \mathbb{N}$ :

$$\begin{aligned}
 & V(X_n) - V(X_{n-1}) \\
 = & \gamma_n \mathcal{A}V(X_{n-1}) + \gamma_n^2 \int_0^1 (1-t) \text{Tr} \left( D^2V(X_{n-1} + t\gamma_n b_{n-1}) b_{n-1} \otimes b_{n-1} \right) dt \\
 & - \frac{\gamma_n}{2} \text{Tr} (D^2V(X_{n-1})) \Sigma_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) \\
 & + \gamma_n \int_0^1 (1-t) \text{Tr} \left( D^2V(X_{n-1} + \gamma_n b_{n-1} + t\sqrt{\gamma_n} \sigma_{n-1} U_n) \sigma_{n-1} U_n \otimes U_n \sigma_{n-1}^* \right) dt \\
 \leq & -\gamma_n \alpha_V V(X_{n-1}) + \gamma_n \beta_V + C_V \frac{\gamma_n^2}{2} \|D^2V\|_\infty V(X_{n-1}) \\
 & + \frac{\gamma_n}{2} \|D^2V\|_\infty \|\sigma\|_\infty^2 + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) + \frac{\gamma_n}{2} \|D^2V\|_\infty \|\sigma\|_\infty^2 |U_n|^2 \\
 \leq & \gamma_n \left( -\frac{\alpha_V}{2} V(X_{n-1}) + \tilde{c} \right) + \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla V(X_{n-1} + \gamma_n b_{n-1}) + \frac{\gamma_n}{2} \|D^2V\|_\infty \|\sigma\|_\infty^2 |U_n|^2
 \end{aligned} \tag{3.41}$$

for a constant  $\tilde{c} := \tilde{c}(V, \sigma, \beta_V)$ . We have in fact considered the time steps sufficiently small (in **(S)**), we have chosen for all  $n \in \mathbb{N}$ ,  $\gamma_n < \min(\frac{1}{2\sqrt{C_V \tilde{c}}}, \frac{\alpha_V}{2C_V \|D^2V\|_\infty})$ . The two terms involving the innovation  $U_n$  in the above decomposition can be controlled thanks to the Gaussian concentration hypothesis **(GC)**. Let us define for all  $x \in \mathbb{R}^d$  and all  $\gamma, \lambda > 0$  the quantities:

$$\begin{aligned}
 I_1(\gamma, \lambda, x) & := \mathbb{E} \left[ \exp \left( \lambda \sqrt{\gamma} \sigma(x) U_1 \cdot \nabla V(x + \gamma b(x)) \right) \right], \\
 I_2(\gamma, \lambda) & := \mathbb{E} \left[ \exp \left( \lambda \frac{\gamma}{2} \|D^2V\|_\infty \|\sigma\|_\infty^2 |U_1|^2 \right) \right].
 \end{aligned}$$

The first one is directly controlled owing to hypothesis **(GC)**:

$$I_1(\gamma_n, \lambda, x) \leq \exp \left( \frac{\lambda^2 \gamma_n |\sigma^*(x) \nabla V(x + \gamma_n b(x))|^2}{2} \right) \underset{(\mathcal{L}_V)}{\leq} \exp \left( \frac{\lambda^2 \gamma_n C_V \|\sigma\|_\infty^2 V(x + \gamma_n b(x))}{2} \right). \tag{3.42}$$

Furthermore, under **(GC)**, for all  $c < \frac{1}{2}$ ,  $I_c := \mathbb{E} [\exp(c|U_n|^2)] < +\infty$ . Hence, for all  $\lambda < \frac{2c}{\|D^2V\|_\infty \|\sigma\|_\infty^2 \gamma_1}$ , Jensen's inequality yields:

$$I_2(\gamma_n, \lambda) \leq \left[ \mathbb{E} \exp(c|U_n|^2) \right]^{\frac{\lambda \gamma_n \|D^2V\|_\infty \|\sigma\|_\infty^2}{2c}} = \exp \left( \gamma_n \ln(I_c) \frac{\lambda \|D^2V\|_\infty \|\sigma\|_\infty^2}{2c} \right). \tag{3.43}$$

These controls allow to prove the integrability statement of the proposition by induction. For  $n = 0$ , recalling from assumption **(C1)** that for all  $\lambda < \lambda_0$ ,  $\mathbb{E} \exp(\lambda|X_0|^2) < +\infty$  and from **(L<sub>V</sub>)**, i) that  $V(x) \leq \bar{c}|x|^2$  outside of a compact set, we derive that for all  $\lambda \in (0, \frac{\lambda_0}{\bar{c}})$ , there exists  $C_{V,\lambda}^0 \in (1, +\infty)$  such that

$$\mathbb{E} \exp(\lambda V(X_0)) \leq C_{V,\lambda}^0.$$

Set now  $\tilde{\beta}_V := \tilde{c} + \ln(I_c) \frac{\|D^2V\|_\infty \|\sigma\|_\infty^2}{2c}$  and  $\tilde{\alpha}_V := \min\left(\frac{1}{\gamma_1}, \frac{\alpha_V}{2} - \lambda C_V \|\sigma\|_\infty^2 (1 + \gamma_1 C_V [1 + \frac{\gamma_1 \|D^2V\|_\infty}{2}])\right) \in (0, \frac{1}{\gamma_1}]$ , for  $\lambda < \frac{\alpha_V}{2C_V \|\sigma\|_\infty^2 (1 + \gamma_1 C_V [1 + \frac{\gamma_1 \|D^2V\|_\infty}{2}])}$ .

Let us assume that for all  $\lambda < \lambda_V := \min\left(\frac{\lambda_0}{2\tilde{c}}, \frac{\alpha_V}{2C_V \|\sigma\|_\infty^2 (1 + \gamma_1 C_V [1 + \frac{\gamma_1 \|D^2V\|_\infty}{2}])}, \frac{c}{\|D^2V\|_\infty \|\sigma\|_\infty^2 \gamma_1}\right)$ , the property

$$\forall k \in \llbracket 0, n-1 \rrbracket, \mathbb{E} \exp(\lambda V(X_k)) \leq C_{V,\lambda} := C_{V,\lambda}^0 \vee \exp\left(\frac{\lambda \tilde{\beta}_V}{\tilde{\alpha}_V}\right), \quad (\mathcal{P}_{n-1})$$

holds for a fixed  $n-1 \in \mathbb{N}_0$  and let us prove  $(\mathcal{P}_n)$ . By inequalities (3.41), (3.42) and (3.43) and the Cauchy-Schwarz inequality, we derive that for all  $\lambda < \lambda_V$ ,

$$\begin{aligned} & \mathbb{E} \exp(\lambda V(X_n)) \\ &= \mathbb{E} \left[ \exp(\lambda V(X_{n-1})) \mathbb{E} \left[ \exp(\lambda(V(X_n) - V(X_{n-1}))) \middle| \mathcal{F}_{n-1} \right] \right] \\ &\leq \mathbb{E} \left[ \exp\left(\lambda \left[V(X_{n-1}) \left(1 - \frac{\alpha_V}{2} \gamma_n\right) + \tilde{c} \gamma_n\right] I_1(\gamma_n, 2\lambda, X_{n-1})^{1/2} I_2(\gamma_n, 2\lambda)^{1/2}\right) \right] \\ &= \exp(\lambda \gamma_n \tilde{\beta}_V) \mathbb{E} \left[ \exp\left(\lambda \left(1 - \frac{\alpha_V}{2} \gamma_n\right) V(X_{n-1}) + \lambda^2 \gamma_n C_V \|\sigma\|_\infty^2 V(X_{n-1} + \gamma_n b_{n-1})\right) \right]. \end{aligned}$$

Recall now that

$$\begin{aligned} V(X_{n-1} + \gamma_n b_{n-1}) &\leq V(X_{n-1}) + \gamma_n \|\nabla V(X_{n-1})\| |b_{n-1}| + \frac{\gamma_n^2}{2} \|D^2V\|_\infty |b_{n-1}|^2 \\ &\stackrel{(\mathcal{L}_V), ii)}{\leq} V(X_{n-1}) \left(1 + \gamma_n C_V \left[1 + \frac{\gamma_n \|D^2V\|_\infty}{2}\right]\right). \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E} \left[ \exp(\lambda V(X_n)) \right] &\leq \exp(\lambda \gamma_n \tilde{\beta}_V) \mathbb{E} \left[ \exp\left(\lambda \underbrace{\left(1 - \gamma_n \tilde{\alpha}_V\right)}_{\in [0,1]} V(X_{n-1})\right) \right] \\ &\stackrel{(\text{Jensen})}{\leq} \exp(\lambda \gamma_n \tilde{\beta}_V) \mathbb{E} \left[ \exp(\lambda V(X_{n-1})) \right]^{(1 - \gamma_n \tilde{\alpha}_V)} \\ &\leq \exp(\lambda \gamma_n \tilde{\beta}_V) C_{V,\lambda}^{(1 - \gamma_n \tilde{\alpha}_V)} \end{aligned}$$

using  $(\mathcal{P}_{n-1})$  for the last inequality. From the above equation and the previous definition of  $C_{V,\lambda}$  we have:

$$\exp(\lambda \gamma_n \tilde{\beta}_V) C_{V,\lambda}^{(1 - \gamma_n \tilde{\alpha}_V)} \leq C_{V,\lambda} \iff C_{V,\lambda} \geq \exp\left(\frac{\lambda \tilde{\beta}_V}{\tilde{\alpha}_V}\right).$$

Hence,  $(\mathcal{P}_n)$  holds. Taking  $c_V < \lambda_V$  completes the proof.  $\square$

**Remark 9** Noting that  $v^* := \inf_{x \in \mathbb{R}^d} V(x) > 0$ , we get that for all  $(n, \xi) \in \mathbb{N} \times [0, 1]$ , and for all  $\lambda < \lambda_V (v^*)^{1-\xi}$ :

$$\mathbb{E} \exp(\lambda V_n^\xi) = \mathbb{E} \exp\left(\lambda (v^*)^\xi \underbrace{\left(\frac{V_n}{v^*}\right)^\xi}_{\geq 1}\right) \leq \mathbb{E} \exp(\lambda (v^*)^{\xi-1} V_n) \leq C_{V,\lambda (v^*)^{\xi-1}} < +\infty.$$

Thus, we readily get as a by-product of Proposition 1 that, for all  $\xi \in [0, 1]$ ,  $\lambda < \lambda_V(v^*)^{1-\xi}$ ,  $\sup_{n \in \mathbb{N}} \mathbb{E} \exp(\lambda V_n^\xi) < +\infty$ . We refer to Lemaire (see e.g. Theorem 17 in [Lem07]) for additional results in that direction.

**Proof of Lemma 4.**

Recalling from  $(\mathbf{G}_V)$  that there exists  $C_{V,\varphi} > 0$  such that for all  $x \in \mathbb{R}^d$ ,  $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$ , we get for  $j \in \{1, 2\}$ :

$$\begin{aligned} & \left[ \mathbb{E} \exp \left( jp \lambda_n \frac{|\varphi(X_0) - \varphi(X_n)|}{\Gamma_n} \right) \right]^{\frac{1}{jp}} \leq \left[ \mathbb{E} \exp \left( jp \lambda_n \frac{C_{V,\varphi}(2 + \sqrt{V(X_0)} + \sqrt{V(X_n)})}{\Gamma_n} \right) \right]^{\frac{1}{jp}} \\ & \leq \exp \left( 2C_{V,\varphi} \frac{\lambda_n}{\Gamma_n} \right) \left[ \mathbb{E} \exp \left( 2jp C_{V,\varphi} \lambda_n \frac{\sqrt{V(X_0)}}{\Gamma_n} \right) \right]^{\frac{1}{2jp}} \left[ \mathbb{E} \exp \left( 2jp C_{V,\varphi} \lambda_n \frac{\sqrt{V(X_n)}}{\Gamma_n} \right) \right]^{\frac{1}{2jp}}. \end{aligned}$$

Write now for  $i \in \{0, n\}$  by the Young inequality:

$$2jp C_{V,\varphi} \lambda_n \frac{\sqrt{V(X_i)}}{\Gamma_n} \leq c_V V(X_i) + \frac{(jp)^2 C_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2},$$

where  $c_V$  is the positive real constant such that  $I_V^1 = \sup_{n \geq 0} \mathbb{E} [\exp(c_V V(X_n))] < +\infty$  (see Proposition 1). We then get

$$\begin{aligned} & \left[ \mathbb{E} \exp \left( jp \lambda_n \frac{|\varphi(X_0) - \varphi(X_n)|}{\Gamma_n} \right) \right]^{\frac{1}{jp}} \\ & \leq \exp \left( 2C_{V,\varphi} \frac{\lambda_n}{\Gamma_n} \right) \exp \left( \frac{jp C_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2} \right) \left( \mathbb{E} \exp(c_V V(X_0)) \right)^{\frac{1}{2jp}} \left( \mathbb{E} \exp(c_V V(X_n)) \right)^{\frac{1}{2jp}} \\ & \leq \exp \left( \frac{(j+1)p C_{V,\varphi}^2 \lambda_n^2}{c_V \Gamma_n^2} \right) \exp \left( \frac{c_V}{p} \right) (I_V^1)^{\frac{1}{jp}}. \quad \square \end{aligned}$$

**Proof of Lemma 5.**

- *Proof of inequalities (3.27) and (3.28).*

- If  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  is Lipschitz continuous. We first rewrite from the definition of  $D_{2,b,n}$  in (3.19):

$$\begin{aligned} D_{2,b,n} &= \sum_{k=1}^n \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\ &= \sum_{k=1}^n \gamma_k \left[ \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}), b_{k-1} - b(X_{k-1} + t\gamma_k b_{k-1}) \rangle dt \right. \\ & \quad \left. + \int_0^1 (\langle \nabla \varphi, b \rangle(X_{k-1} + t\gamma_k b_{k-1}) - \langle \nabla \varphi, b \rangle(X_{k-1})) dt \right]. \end{aligned}$$

From the boundedness of  $\nabla \varphi$ , and the Lipschitz property of the mappings  $x \mapsto b(x)$  (which has been assumed from the very beginning) and  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  (assumed for

the current inequality), recalling that  $b_{k-1} = b(X_{k-1})$ , one derives that :

$$|D_{2,b,n}| \leq \sum_{k=1}^n \gamma_k^2 \left( \|\nabla\varphi\|_\infty [b]_1 + [\langle \nabla\varphi, b \rangle]_1 \right) \frac{|b_{k-1}|}{2} \leq C \sum_{k=1}^n \gamma_k^2 |b_{k-1}|, \quad C := C(b, \varphi). \quad (3.44)$$

From this inequality, assumption  $(\mathcal{L}_V)$ , ii) and the Jensen inequality (applied to the exponential function for the measure  $\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \delta_k$ ), we derive:

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} \leq \left( \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp \left( \frac{4p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} C \sqrt{C_V} \sqrt{V_{k-1}} \right) \right)^{\frac{1}{4p}}.$$

From the Young inequality we obtain:

$$\mathbb{E} \exp \left( \frac{4p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} C \sqrt{C_V} \sqrt{V_{k-1}} \right) \leq \exp \left( \left( \frac{2\sqrt{2}p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \frac{C \sqrt{C_V}}{\sqrt{c_V}} \right)^2 \right) \mathbb{E} [\exp(c_V V_{k-1})].$$

We finally derive with the notations of Proposition 1:

$$\begin{aligned} \left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} &\leq \exp \left( \frac{2p\lambda_n^2 (\Gamma_n^{(2)})^2 (C \sqrt{C_V})^2}{\Gamma_n^2 c_V} \right) (I_V^1)^{\frac{1}{4p}} \\ &\leq \exp \left( C_{3.27} \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}}, \end{aligned}$$

setting  $C_{3.27} := 2 \frac{(C \sqrt{C_V})^2}{c_V}$  with  $C = \frac{1}{2} (\|\nabla\varphi\|_\infty [b]_1 + [\langle \nabla\varphi, b \rangle]_1)$  as in (3.44). □

- If  $a \leq \frac{c_V q \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 \sqrt{\Gamma_n}}{4C_V p \|\nabla^2\varphi\|_\infty^2 \Gamma_n^{(2)}} = \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$  with the notation introduced in (3.34). Write first from (3.19) (definition of  $D_{2,b,n}$ ), using a Taylor expansion on  $\nabla\varphi$ :

$$\begin{aligned} &\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} \\ &\leq \left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} \sum_{k=1}^n \gamma_k^2 \int_0^1 (1-t) \left| \text{Tr} \left( D^2\varphi(X_{k-1} + t\gamma_k b_{k-1}) b_{k-1} \otimes b_{k-1} \right) \right| dt \right) \right)^{\frac{1}{4p}}. \end{aligned} \quad (3.45)$$

We first easily get from the assumptions on  $\varphi$  and point ii) of  $(\mathcal{L}_V)$  that:

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} \leq \left( \mathbb{E} \exp \left( \frac{2p\lambda_n}{\Gamma_n} \sum_{k=1}^n \gamma_k^2 C_V V_{k-1} \|D^2\varphi\|_\infty \right) \right)^{\frac{1}{4p}}.$$

From the Jensen inequality, we derive:

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} \leq \left( \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp \left( \frac{2p\lambda_n \Gamma_n^{(2)}}{\Gamma_n} \|D^2\varphi\|_\infty C_V V_{k-1} \right) \right)^{\frac{1}{4p}}.$$



We then have from the definition of  $\lambda_n$  in (3.24) that:

$$\bar{v}_n := \frac{2p\lambda_n\Gamma_n^{(2)}}{\Gamma_n} \|D^2\varphi\|_\infty \frac{C_V}{c_V} = \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \frac{2C_V p}{c_V q} \frac{\|D^2\varphi\|_\infty}{\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} a \leq 1.$$

The Jensen inequality for concave functions yields for all  $k \in \llbracket 1, n \rrbracket$ :

$$\mathbb{E} \exp\left(\frac{2p\lambda_n\Gamma_n^{(2)}}{\Gamma_n} \|D^2\varphi\|_\infty C_V V_{k-1}\right) = \mathbb{E} \exp\left(\bar{v}_n c_V V_{k-1}\right) \leq \left(\mathbb{E} \exp\left(c_V V_{k-1}\right)\right)^{\bar{v}_n}.$$

Thus, setting

$$v_n := \frac{\bar{v}_n}{4p} = \frac{\lambda_n\Gamma_n^{(2)}}{2\Gamma_n} \|D^2\varphi\|_\infty \frac{C_V}{c_V}, \quad (3.46)$$

we finally derive,

$$\left[\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}|\right)\right]^{\frac{1}{4p}} \leq \left[\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \left(\sup_{l \geq 1} \mathbb{E} [\exp(c_V V_{l-1})]\right)^{\bar{v}_n}\right]^{\frac{1}{4p}} = (I_V^1)^{v_n} =: C_n,$$

using again the notations of Proposition 1. This gives (3.28).

- *Proof of inequality (3.26).* We proceed as for the proof of (3.28) and (3.27). Write:

$$\begin{aligned} & \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p}} \\ & \leq \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} \sum_{k=1}^n \frac{\gamma_k}{2} \left|\text{Tr}\left((D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1}))\Sigma_{k-1}\right)\right|\right)\right)^{\frac{1}{4p}} \\ & \leq \left(\mathbb{E} \exp\left(\frac{2p\lambda_n}{\Gamma_n} \|\sigma\|_\infty^2 [\varphi^{(2)}]_1 \sum_{k=1}^n \gamma_k^2 |b_{k-1}|\right)\right)^{\frac{1}{4p}} \\ & \leq \left(\mathbb{E} \exp\left(\frac{2p\lambda_n}{\Gamma_n} \|\sigma\|_\infty^2 [\varphi^{(2)}]_1 C_V^{\frac{1}{2}} \sum_{k=1}^n \gamma_k^2 |V_{k-1}|^{\frac{1}{2}}\right)\right)^{\frac{1}{4p}} \\ & \leq \left(\frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp\left(\frac{2p\lambda_n\Gamma_n^{(2)}}{\Gamma_n} \|\sigma\|_\infty^2 [\varphi^{(2)}]_1 C_V^{\frac{1}{2}} |V_{k-1}|^{\frac{1}{2}}\right)\right)^{\frac{1}{4p}}. \end{aligned}$$

Using once again the Young inequality and setting  $C_{3.26} := \frac{\|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2 C_V}{4 c_V}$ , we obtain:

$$\begin{aligned} \left(\mathbb{E} \exp\left(\frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p}} & \leq \exp\left(\frac{p\lambda_n^2}{4} \left(\frac{\Gamma_n^{(2)}}{\Gamma_n}\right)^2 \|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2 \frac{C_V}{c_V}\right) (I_V^1)^{\frac{1}{4p}} \\ & \leq \exp\left(C_{3.26} p \lambda_n^2 \left(\frac{\Gamma_n^{(2)}}{\Gamma_n}\right)^2\right) (I_V^1)^{\frac{1}{4p}}. \quad \square \end{aligned}$$

## 4 A refinement when $\|\|\sigma\|\|^2 - \nu(\|\|\sigma\|\|^2)$ is a Coboundary

We will assume in this section that there exists a solution  $\vartheta$  of the Poisson problem  $\mathcal{A}\vartheta = \|\|\sigma\|\|^2 - \nu(\|\|\sigma\|\|^2)$ , where  $\|\|\cdot\|\|$  is a matrix norm such that  $\|\cdot\| \leq \|\|\cdot\|\|$ , satisfying the assumptions stated for  $\varphi$  in Theorem 2. This is in particular the case for the Fröbenius norm  $\|\cdot\|_F$  under the assumptions of the previous Theorem 4.

In this special case, we have a slightly different concentration result improving our previous ones for a certain deviation range.

**Theorem 8** *Under the assumptions of Theorem 2 and with the notations introduced therein, we have that:*

- (a) For  $(\beta \in (0, 1]$  and  $\theta \in (\frac{1}{2+\beta}, 1])$ , there exist two explicit monotonic sequences  $\tilde{c}_n \leq 1 \leq \tilde{C}_n$ ,  $n \geq 1$ , with  $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$  such that for all  $n \geq 1$  for all  $a > 0$ :

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 \tilde{C}_n \exp\left(-\frac{\tilde{c}_n}{2\nu(\|\|\sigma\|\|^2)\|\nabla\varphi\|_\infty^2} \Phi_n(a)\right),$$

$$\Phi_n(a) := \left[ \left( a^2 \left( 1 - \frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \right) \right) \vee \left( a^{\frac{4}{3}} \Gamma_n^{\frac{1}{3}} \tilde{c}_n \left( 1 - \frac{2}{3} \tilde{c}_n \left( \frac{\Gamma_n}{a^2} \right)^{\frac{1}{3}} \right)_+ \right) \right],$$

where  $x_+ = \max(x, 0)$  and  $\tilde{c}_n := \left( \frac{[\varphi]_1}{[\vartheta]_1} \right)^{2/3} \nu(\|\|\sigma\|\|^2) \|\|\sigma\|\|_\infty^{-2/3} \tilde{c}_n$  with  $\tilde{c}_n$  being an explicit positive sequence s.t.  $\tilde{c}_n \downarrow 1$ .

- (b) For  $\beta \in (0, 1]$ ,  $\theta = \frac{1}{2+\beta}$ , there exist two explicit monotonic sequences  $\tilde{c}_n \leq 1 \leq \tilde{C}_n$ ,  $n \geq 1$ , with  $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$  such that for all  $n \geq 1$  for all  $a > 0$ :

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) + \mathcal{B}_{n,\beta}| \geq a] \leq 2 \tilde{C}_n \exp\left(-\frac{\tilde{c}_n}{2\nu(\|\|\sigma\|\|^2)\|\nabla\varphi\|_\infty^2} \Phi_n(a)\right).$$

**Remark 10 (About deviation rates)** *Observe that in order to derive **global** deviation bounds (valid for every  $a > 0$ ) two concentration regimes appear in the previous bounds. For an arbitrary fixed  $a > 0$ , we have that for  $n$  large enough (depending on  $a$ ), the Gaussian concentration regime will give the fastest decay, since  $\frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \rightarrow 0$ . Also, when  $a \asymp \sqrt{\Gamma_n}$  the two above contributions give a Gaussian bound, with suboptimal constants. Eventually, when  $a \gg \sqrt{\Gamma_n}$ , for a fixed  $n$ , we have that the first term is “stuck” at the threshold  $\Gamma_n$  whatever level  $a$  is considered, i.e.  $a^2 \left( 1 - \frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \right) \xrightarrow{a \rightarrow \infty} \tilde{c}_n^3 \Gamma_n$  whereas the second clearly becomes bigger.*

To summarize, when the Gaussian regime prevails (i.e. when  $\frac{a}{\sqrt{\Gamma_n}}$  is small), the results of Theorem 2 have been improved in the sense that the variance in the deviations is a sharper upper bound of the “carré du champ”  $\int_{\mathbb{R}^d} |\sigma^* \nabla \varphi(x)|^2 \nu(dx)$  appearing in the asymptotic Theorem 1. Indeed, we managed to replace the supremum norm  $\|\|\sigma\|\|_\infty^2$  deriving from Theorem 2 and the domination condition on the matrix norms by  $\nu(\|\|\sigma\|\|^2)$ . However, our martingale approach naturally leads to a bound in  $\|\nabla\varphi\|_\infty^2$ .

On the other hand, the global double regime seems to be the price to pay to benefit from the better approximation of the “carré du champ” in the Gaussian regime.

Eventually, Theorem 3 is a direct consequence of the previous theorem in the Gaussian regime.

**Proof:** We focus on case (a) for  $\beta \in (0, 1)$ ,  $\theta \in (1/(2 + \beta), 1]$ . Case (b) could be derived similarly following the proof of Theorem 2. We restart from the computations of Section 3.1 that give for all  $\lambda > 0$  the control in equation (3.21). Let us now deal with the term giving the concentration and write for all  $\rho > 1$ :

$$\begin{aligned} \mathbb{E} \exp \left( -\frac{q\lambda}{\Gamma_n} M_n \right) &\leq \left( \mathbb{E} \exp \left( -\rho \frac{q\lambda}{\Gamma_n} M_n - \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2}{2\Gamma_n^2} \sum_{k=1}^n \gamma_k \mathcal{A}^\vartheta(X_{k-1}) \right) \right)^{\frac{1}{\rho}} \\ &\quad \times \left( \mathbb{E} \exp \left( \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2}{2(\rho-1)\Gamma_n^2} \sum_{k=1}^n \gamma_k \mathcal{A}^\vartheta(X_{k-1}) \right) \right)^{1-\frac{1}{\rho}} \\ &=: \mathcal{T}_1^{\frac{1}{\rho}} \mathcal{T}_2^{1-\frac{1}{\rho}}. \end{aligned} \quad (4.1)$$

Since for all  $x \in \mathbb{R}^d$ ,  $\mathcal{A}^\vartheta(x) = \|\sigma(x)\|^2 - \nu(\|\sigma\|^2)$ , we obtain:

$$\mathcal{T}_1 = \exp \left( \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n} \right) \mathbb{E} \exp \left( -\rho \frac{q\lambda}{\Gamma_n} M_n - \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2}{2\Gamma_n^2} \sum_{k=1}^n \gamma_k \|\sigma(X_{k-1})\|^2 \right).$$

The key idea is that we have exploited the Poisson equation solved by  $\vartheta$  to replace the previous *rough* control  $\exp \left( \frac{(q\lambda)^2 [\varphi]_1^2 \|\sigma\|_\infty^2}{2\Gamma_n} \right)$ , coming from the martingale increment obtained in equation (3.18) and the domination condition on the matrix norms, by the above term  $\exp \left( \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n} \right)$ . This last contribution will be part of the optimization procedure over  $\lambda$ . This improvement will be all the more significant that neighborhoods of the points where the norm of the diffusion coefficient  $\sigma$  attains its supremum are not *very much* charged by the invariant distribution. The point for  $\mathcal{T}_1$  is then to prove that the remaining expectation is less than 1. It will be shown by exhibiting an appropriate underlying supermartingale.

Set to this end  $\widetilde{\mathcal{T}}_1 := \exp \left( -\frac{\rho^2 (q\lambda)^2 [\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n} \right) \mathcal{T}_1$ . Define now, for a given  $n \in \mathbb{N}$  and  $m \in \mathbb{N}_0$ ,  $S_m := \exp \left( -\rho \frac{q\lambda}{\Gamma_n} M_m - \frac{\rho^2 (q\lambda)^2 [\varphi]_1^2}{2\Gamma_n^2} \sum_{k=1}^m \gamma_k \|\sigma(X_{k-1})\|^2 \right)$ . From the definition of the martingale  $(M_k)_{k \geq 1}$  in (3.14) and the controls of the Lipschitz constants of the functions  $(\psi_k(X_{k-1}, \cdot))_{k \in \llbracket 1, n \rrbracket}$  in Lemma 1, we get by iterated conditioning:

$$\begin{aligned} \widetilde{\mathcal{T}}_1 &\leq \mathbb{E} \left[ S_{n-1} \exp \left( -\frac{\rho^2 (q\lambda)^2 [\varphi]_1^2}{2\Gamma_n^2} \gamma_n \|\sigma(X_{n-1})\|^2 \right) \mathbb{E} \left[ \exp \left( -\rho \frac{q\lambda}{\Gamma_n} (M_n - M_{n-1}) \right) \middle| \mathcal{F}_{n-1} \right] \right] \\ &\stackrel{(\text{GC})}{\leq} \mathbb{E} \left[ S_{n-1} \exp \left( -\frac{\rho^2 (q\lambda)^2 [\varphi]_1^2}{2\Gamma_n^2} \gamma_n \|\sigma(X_{n-1})\|^2 \right) \exp \left( \frac{\rho^2 (q\lambda)^2}{2\Gamma_n^2} \gamma_n [\varphi]_1^2 \|\sigma(X_{n-1})\|^2 \right) \right] \\ &\leq \mathbb{E} [S_{n-1}] \leq 1. \end{aligned}$$

In other words,  $(S_m)_{m \geq 0}$  is a positive supermartingale. We finally get that, for all  $\rho > 1$ :

$$\mathcal{I}_1^{\frac{1}{\rho}} \leq \exp\left(\frac{\rho(q\lambda)^2[\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n}\right). \quad (4.2)$$

For the term  $\mathcal{I}_2$ , we have that setting  $\mu := \mu(q, n, \rho, \lambda) = \frac{(q\lambda)^2 \rho^2 [\varphi]_1^2}{2(\rho-1)\Gamma_n}$ ,

$$\mathcal{I}_2 = \mathbb{E} \exp\left(\frac{\mu}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathcal{A}^\vartheta(X_{k-1})\right),$$

so that this contribution can be controlled from the previous expansion of Lemma 1 exploiting the technical lemmas of Section 3.1 replacing  $\lambda$  by  $\mu$  and  $\varphi$  by  $\vartheta$ .

In case (a), for  $\theta \in (1/(2 + \beta), 1]$ ,  $\beta \in (0, 1]$ , the Hölder inequalities yield that for all  $\mu \in \mathbb{R}_+$  and all  $\bar{p}, \bar{q} \in (1, +\infty)$ ,  $\frac{1}{\bar{p}} + \frac{1}{\bar{q}} = 1$ , similarly to (3.21),

$$\begin{aligned} \mathcal{I}_2 &= \mathbb{E} \exp\left(\frac{\mu}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathcal{A}^\vartheta(X_{k-1})\right) \leq \left(\mathbb{E} \exp\left(-\frac{\bar{q}\mu}{\Gamma_n} M_n^\vartheta\right)\right)^{\frac{1}{\bar{q}}} \\ &\times \left(\mathbb{E} \exp\left(\frac{2\bar{p}\mu}{\Gamma_n} |L_n^\vartheta|\right)\right)^{\frac{1}{2\bar{p}}} \left(\mathbb{E} \exp\left(\frac{4\bar{p}\mu}{\Gamma_n} |D_{2,b,n}^\vartheta|\right)\right)^{\frac{1}{4\bar{p}}} \left(\mathbb{E} \exp\left(\frac{4\bar{p}\mu}{\Gamma_n} |D_{2,\Sigma,n}^\vartheta|\right)\right)^{\frac{1}{4\bar{p}}}, \end{aligned} \quad (4.3)$$

where the superscripts in  $\vartheta$  emphasize that the contributions to be analyzed are those associated with the solution  $\vartheta$  of the Poisson problem with source  $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ .

Still for simplicity, we assume as well (case (i)) that the mapping  $x \mapsto \langle b(x), \nabla \vartheta(x) \rangle$  is Lipschitz continuous. Plugging in (4.3) the controls established in Lemma 4 (with  $j = 2$ ), Lemma 5 (equations (3.26) and (3.27)) and (3.39), then replacing  $\lambda_n$  by  $\mu$ , we get similarly to the first inequality of (3.30) and with the notations of Lemma 3:

$$\begin{aligned} \mathcal{I}_2 &\leq \exp\left(\frac{\bar{q}\mu^2 \|\sigma\|_\infty^2 [\vartheta]_1^2}{2\Gamma_n}\right) \exp\left(\frac{\mu^2}{2\Gamma_n \bar{p}} + \frac{\bar{p}a_n^2}{2}\right) \exp\left(\frac{3\bar{p}C_{V,\vartheta}^2 \mu^2}{c_V \Gamma_n^2} + \frac{c_V}{\bar{p}}\right) (I_V^1)^{\frac{1}{2\bar{p}}} \\ &\times \exp\left(C_{3.26} \frac{\bar{p}\mu^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4\bar{p}}} \times \exp\left(C_{3.27} \left(\frac{3\bar{p}\mu^2 (\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2\bar{p}}\right)\right) (I_V^1)^{\frac{1}{4\bar{p}}}. \\ &\leq \exp\left(\frac{\mu^2}{\Gamma_n} \left(\frac{\bar{q} \|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{p} \left(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{3.26} + \frac{3}{2}C_{3.27}] + \frac{3C_{V,\vartheta}^2}{c_V \Gamma_n}\right) + \frac{1}{2\bar{p}}\right)\right) \\ &\times \exp\left(\frac{1}{\bar{p}} \left(c_V + \frac{C_{3.27}}{2}\right) + \frac{\bar{p}a_n^2}{2}\right) (I_V^1)^{\frac{1}{\bar{p}}}. \end{aligned}$$

Set now

$$\begin{aligned} \bar{C}_n &:= \exp\left(\frac{1}{\bar{p}} \left(c_V + \frac{C_{3.27}}{2}\right) + \frac{\bar{p}a_n^2}{2}\right) (I_V^1)^{\frac{1}{\bar{p}}}, \\ \bar{e}_n &:= \bar{p} \left(\frac{(\Gamma_n^{(2)})^2}{\Gamma_n} [C_{3.26} + \frac{3}{2}C_{3.27}] + \frac{3C_{V,\vartheta}^2}{c_V \Gamma_n}\right) + \frac{1}{2\bar{p}}. \end{aligned} \quad (4.4)$$

In the considered case, the exponent  $\bar{p} := \bar{p}_n$  can again be taken such that  $\bar{p}_n \rightarrow +\infty$  and  $\bar{p}_n \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} \rightarrow 0$  in order to have,  $\bar{e}_n \rightarrow 0$ ,  $\bar{C}_n \rightarrow 1$  with the indicated monotonicity for large enough  $n$ .

We derive from the above control and (4.2) that for all  $q, \rho > 1$ :

$$\begin{aligned} & \left( \mathbb{E} \exp \left( -\frac{\lambda q}{\Gamma_n} M_n \right) \right)^{\frac{1}{q}} \leq \left( \mathcal{T}_1^{\frac{1}{\rho}} \mathcal{T}_2^{1-\frac{1}{\rho}} \right)^{\frac{1}{q}} \\ & \leq \exp \left( \frac{\rho q \lambda^2 [\varphi]_1^2 \nu(\|\sigma\|^2)}{2\Gamma_n} \right) \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp \left( \frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left( \frac{\bar{q} \|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right) \right). \end{aligned}$$

Plugging this bound in (3.21), using again the controls of Lemmas 4 and 5, eventually yields:

$$\begin{aligned} & \mathbb{P} \left[ \sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a \right] \\ & \leq \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \right) \exp \left( \frac{\lambda^2}{2\Gamma_n} (\rho q [\varphi]_1^2 \nu(\|\sigma\|^2) + \frac{1}{p}) \right) \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp \left( \frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left( \frac{\bar{q} \|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right) \right) \\ & \quad \times \exp \left( \frac{\lambda^2}{\Gamma_n} p \left( \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} (C_{3.26} + \frac{3}{2} C_{3.27}) + \frac{3C_{V,\varphi}^2}{c_V \Gamma_n} \right) \right) \exp \left( \frac{1}{p} (c_V + \frac{C_{3.27}}{2}) + \frac{pa_n^2}{2} \right) (I_V^1)^{\frac{1}{p}}. \end{aligned}$$

Choosing  $p := p_n \xrightarrow{n} +\infty$  and such that  $p_n \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} \xrightarrow{n} 0$ , we get by a standard symmetry and with the notations introduced in the proof of Theorem 2:

$$\begin{aligned} \mathbb{P} \left[ \left| \sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \right| \geq a \right] & \leq 2 C_n \bar{C}_n^{\frac{\rho-1}{\rho q}} \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \right) \exp \left( \frac{\lambda^2}{\Gamma_n} \left( \frac{\rho q [\varphi]_1^2 \nu(\|\sigma\|^2)}{2} + e_n \right) \right) \\ & \quad \times \exp \left( \frac{\rho-1}{\rho q} \frac{\mu^2}{\Gamma_n} \left( \frac{\bar{q} \|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right) \right), \end{aligned}$$

where  $e_n$  is defined similarly to  $\bar{e}_n$  in (4.4) replacing  $\bar{p}$  by  $p$ . In particular  $e_n \xrightarrow{n} 0$ . Note that for the previous choices of  $p, \bar{p}$ , we have that  $\tilde{C}_n := C_n \bar{C}_n^{\frac{\rho-1}{\rho q}} \xrightarrow{n} 1$  uniformly in  $\rho > 1$ . Recalling that  $\mu = \frac{(q\lambda)^2 \rho^2 [\varphi]_1^2}{2(\rho-1)\Gamma_n}$ , we are thus led to minimize the polynomial function

$$P : \lambda \mapsto -\frac{a\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2}{\Gamma_n} A_n + \frac{\lambda^4}{\Gamma_n^3} B_n,$$

where  $A_n = A_n(\rho) = \rho \tilde{A}_n$  and  $B_n = B_n(\rho) = \frac{\rho^3}{\rho-1} \tilde{B}_n$  with

$$\tilde{A}_n := \frac{q[\varphi]_1^2 \nu(\|\sigma\|^2)}{2} + e_n \quad \text{and} \quad \tilde{B}_n := \frac{q^3 [\varphi]_1^4}{4} \left( \frac{\bar{q} \|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right). \quad (4.5)$$

Note that both sequences  $(\tilde{A}_n)_{n \geq 1}$  and  $(\tilde{B}_n)_{n \geq 1}$  are bounded and bounded away from zero sequences (and do not depend on  $\rho$ ). The function  $P$  is clearly convex and coercive so it attains its minimum at  $\lambda_{\min}$ , unique zero of the equation  $P'(\lambda_{\min}) = 0$ . This equation reads

$$\lambda^3 + \frac{A_n \Gamma_n^2}{2B_n} \lambda - \frac{a \Gamma_n^{\frac{5}{2}}}{4B_n} = 0 \quad (4.6)$$

which is the canonical form of this third degree equation to apply the Cardan-Tartaglia formula (\*) so that

$$\lambda_{\min}(\rho) = \frac{\Gamma_n}{2} \left[ \left( \frac{a}{\sqrt{\Gamma_n B_n}} + \sqrt{\left(\frac{2A_n}{3B_n}\right)^3 + \frac{a^2}{\Gamma_n B_n^2}} \right)^{\frac{1}{3}} + \left( \frac{a}{\sqrt{\Gamma_n B_n}} - \sqrt{\left(\frac{2A_n}{3B_n}\right)^3 + \frac{a^2}{\Gamma_n B_n^2}} \right)^{\frac{1}{3}} \right]. \quad (4.7)$$

In order to derive our non-asymptotic bound, we select two “regimes” based on a first order expansion of  $\lambda_{\min}$  in two cases  $\frac{a}{B_n \sqrt{\Gamma_n}} \rightarrow 0$  and  $\frac{B_n \sqrt{\Gamma_n}}{a} \rightarrow 0$ , assuming that the free parameter  $\rho = \rho_n$  to be specified later on remains bounded, *e.g.*  $\rho \in (1, 3]$  (which implies that both quantities  $\frac{A_n}{B_n}$  and  $\frac{1}{B_n}$  remain bounded as well). Also, note that if  $\rho \rightarrow 1$ , then  $\frac{1}{B_n}$  and  $\frac{A_n}{B_n} \rightarrow 0$ . First, one easily checks that if  $(x_n)_{n \geq 1}$  and  $(a_n)_{n \geq 1}$  are two sequences of positive real numbers where  $(a_n)_{n \geq 1}$  is bounded, then

$$\left( x_n + \sqrt{a_n^3 + x_n^2} \right)^{\frac{1}{3}} + \left( x_n - \sqrt{a_n^3 + x_n^2} \right)^{\frac{1}{3}} \sim \begin{cases} \frac{2}{3} \frac{x_n}{a_n} & \text{if } x_n = o\left(a_n^{\frac{3}{2}}\right) \text{ (then } x_n \rightarrow 0), \\ (2x_n)^{\frac{1}{3}} & \text{if } a_n = o\left(x_n^{\frac{2}{3}}\right) \text{ (then } x_n \rightarrow +\infty). \end{cases} \quad (4.8)$$

• If  $\frac{a}{B_n \sqrt{\Gamma_n}} = o\left(\left(\frac{A_n}{B_n}\right)^{\frac{3}{2}}\right)$  (hence goes to 0), setting then  $x_n = \frac{a}{B_n \sqrt{\Gamma_n}}$  and  $a_n = \frac{2A_n}{3B_n}$  yields

$$\lambda_{\min}(\rho) \sim \lambda^*(\rho) := \frac{a \sqrt{\Gamma_n}}{2A_n} \quad \text{as } n \rightarrow +\infty.$$

Note that  $\lambda^* := \lambda^*(\rho)$  corresponds to the optimization of the quadratic part of  $P$ . Then

$$P(\lambda^*) = -\frac{a^2}{4A_n} \left( 1 - \frac{a^2}{4A_n^3} \frac{B_n}{\Gamma_n} \right) = -\frac{a^2}{4\tilde{A}_n \rho} \left( 1 - \frac{a^2}{4\tilde{A}_n^3 (\rho - 1)} \frac{\tilde{B}_n}{\Gamma_n} \right).$$

Set now  $\xi_n := \frac{\alpha_n(a)}{\rho - 1}$  with  $\alpha_n(a) = \frac{\tilde{B}_n}{4\tilde{A}_n^3} \frac{a^2}{\Gamma_n}$ . Then

$$P(\lambda^*) = -\frac{a^2}{4\tilde{A}_n} \frac{1 - \xi_n}{1 + \frac{\alpha_n(a)}{\xi_n}}.$$

It remains to maximize the mapping  $\xi \mapsto \frac{1 - \xi}{1 + \alpha_n(a) \xi^{-1}}$  over  $(0, 1)$ . Its optimum is attained for  $\xi_n^* = \frac{1}{1 + \sqrt{1 + \frac{1}{\alpha_n(a)}}}$ , which in turn yields

$$P(\lambda^*) = -\frac{a^2}{4\tilde{A}_n} \left( 1 - \frac{2}{1 + \sqrt{1 + 4 \frac{\tilde{A}_n^3 \Gamma_n}{B_n a^2}}} \right). \quad (4.9)$$

Note that, with the resulting specification of  $\rho = \rho_n^* := 1 + \frac{\alpha_n(a)}{\xi_n^*} \in (1, 3]$  (at least for large enough  $n$ ), the above condition  $x_n = o\left(a_n^{\frac{3}{2}}\right)$  in (4.8) is satisfied *a posteriori*.

---

\*. If the equation  $z^3 + pz + q = 0$  has a unique real zero  $z_*$  then its discriminant  $\Delta = 4p^3 + 27q^2 > 0$  and  $z_* = \left(\frac{1}{2}(-q + \sqrt{\frac{\Delta}{27}})\right)^{\frac{1}{3}} + \left(\frac{1}{2}(-q - \sqrt{\frac{\Delta}{27}})\right)^{\frac{1}{3}}$ .

- If  $\frac{a}{B_n\sqrt{\Gamma_n}} \rightarrow +\infty$ , then, still owing to (4.8),

$$\lambda_{\min}(\rho) \sim \bar{\lambda}^*(\rho) = \frac{\Gamma_n}{2} \left( \frac{2a}{B_n\sqrt{\Gamma_n}} \right)^{\frac{1}{3}} = \left( \frac{a\Gamma_n}{4B_n} \right)^{\frac{1}{3}} \sqrt{\Gamma_n} \quad \text{as } n \rightarrow +\infty.$$

The value  $\bar{\lambda}^*(\rho)$  corresponds to the *quartic* pseudo-optimum of  $P$  (*i.e.* obtained by neglecting the quadratic term). This yields, when reintroducing the parameter  $\rho$ ,

$$P(\bar{\lambda}^*(\rho)) = -a^{\frac{4}{3}} \Gamma_n^{\frac{1}{3}} \frac{(\rho-1)^{\frac{1}{3}}}{\rho(4\tilde{B}_n)^{\frac{1}{3}}} \left( \frac{3}{4} - \frac{\tilde{A}_n}{(4\tilde{B}_n)^{\frac{1}{3}}} \frac{\Gamma_n^{\frac{1}{3}}}{a^{\frac{2}{3}}} (\rho-1)^{\frac{1}{3}} \right).$$

The right hand side of this equality is a function of  $\rho \in (1, +\infty)$ . Its analysis yields that the optimum is attained in  $(1, 3/2]$  and that it tends asymptotically in  $n$  to  $3/2$  in our considered regime. Taking as suboptimal  $\rho = 3/2$  gives:

$$P(\bar{\lambda}^*(\rho)) \leq -\frac{a^{\frac{4}{3}}}{4} \left( \frac{\Gamma_n}{\tilde{B}_n} \right)^{\frac{1}{3}} \left( 1 - \frac{2\tilde{A}_n}{3\tilde{B}_n} \left( \frac{\Gamma_n}{a^2} \right)^{\frac{1}{3}} \right). \quad (4.10)$$

From (4.10), (4.9) and (4.5), we conclude the proof of case (a) by setting  $\bar{c}_n := \tilde{A}_n \tilde{B}_n^{-\frac{1}{3}}$  which matches with the definition in the statement of the theorem.

In the biased case, the result follows similarly from the corresponding analysis performed in Section 2 taking  $A_n(\rho) = \frac{\rho q[\varphi]_{1\nu}^2(\|\sigma\|^2)}{2}$ .  $\square$

**Remark 11** When  $a \asymp \sqrt{\Gamma_n}$ , one checks that  $\lambda_{\min}(\rho) \asymp \Gamma_n$  and  $P(\lambda_{\min}(\rho)) \asymp -\Gamma_n$ . This behavior is consistent with our non-asymptotic bound. However, for practical and numerical purposes observe that the optimum can be estimated. Namely, plugging the identity (4.6) satisfied by  $\lambda_{\min}(\rho)$  in (4.7) into the definition of  $P$ , yields

$$\begin{aligned} P(\lambda_{\min}(\rho)) &= -\frac{\lambda_{\min}(\rho)}{2\sqrt{\Gamma_n}} \left( \frac{3a}{2} - \frac{\lambda_{\min}(\rho)\rho\tilde{A}_n}{\sqrt{\Gamma_n}} \right) \\ &= -\frac{\sqrt{\Gamma_n}(\rho-1)^{\frac{1}{3}}}{4} \frac{\Phi_n(a, \rho)}{\rho} \left( \frac{3a}{2} - \frac{\sqrt{\Gamma_n}}{2} (\rho-1)^{\frac{1}{3}} \tilde{A}_n \Phi_n(a, \rho) \right), \end{aligned}$$

where

$$\Phi_n(a, \rho) = \left( \frac{a}{\sqrt{\Gamma_n}\tilde{B}_n} + \left( (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 + \frac{a^2}{\tilde{B}_n^2\Gamma_n} \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} + \left( \frac{a}{\sqrt{\Gamma_n}\tilde{B}_n} - \left( (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 + \frac{a^2}{\tilde{B}_n^2\Gamma_n} \right)^{\frac{1}{2}} \right)^{\frac{1}{3}}.$$

Then, an optimization in  $\rho \in (1, +\infty)$  for given  $a, \Gamma_n$  can be performed (noting that  $\rho \mapsto (\rho-1)^{i/3}\rho^{-1}$ ,  $i \in \{1, 2\}$  are bounded functions over  $(1, +\infty)$ ).

## 5 Smoothness Results for the Poisson Problems (Proof of Theorem 4)

We first prove here Theorem 4 which allows to derive from the deviation results of Theorems 2 and 3 the practical deviation bounds of Section 2.3 (i.e. Theorems 5, 6 and 7). We recall that we work in the confluent setting of  $(\mathbf{D}_\alpha^p)$  and that we additionally consider two main types of assumptions:

- Strong confluence conditions and smoothness  $(\mathbf{C}_R)$ . Namely, assumptions  $(\mathcal{L}_V)$ ,  $(\mathbf{D}_\alpha^p)$  and  $(\mathbf{R}_{3,\beta})$  introduced in Sections 1.2 and 2.2 with the condition  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(3+\beta)-p}$ .
- Mild confluence conditions and non-degeneracy  $(\mathbf{C}_{UE})$ . Namely, assumptions  $(\mathcal{L}_V)$ ,  $(\mathbf{D}_\alpha^p)$ ,  $(\mathbf{R}_{1,\beta})$  and  $(\mathbf{UE})$  introduced in Sections 1.2 and 2.2 together, when  $d > 1$ , with the condition  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$  and the technical structure assumption on the diffusion coefficient that for all  $(i, j) \in \llbracket 1, d \rrbracket^2$ ,  $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$ .

It is well known that when  $(\mathbf{C}_R)$  or  $(\mathbf{C}_{UE})$  are in force, there exists a unique invariant distribution for (1.1), i.e. assumption  $(\mathbf{U})$  holds. We refer to [Kha80], [PP14], [Pag01], [PV01] for proofs of this assertion. The next step consists precisely in investigating the smoothness of the corresponding Poisson problem as well as some associated quantitative pointwise bounds on the gradient of its solution, which is one of the key terms appearing in the deviation bounds of Theorems 2 and 3.

Let us indicate that the conditions appearing in  $(\mathbf{C}_R)$  depend on pure pathwise properties, whereas the case  $(\mathbf{C}_{UE})$  takes advantage of the regularity of the underlying semi-group which allows to alleviate some smoothness assumptions on the coefficients and some restrictions on the variations of  $\sigma$ . When the dimension increases, it becomes useful to benefit from the smoothing effects of a non-degenerate semi-group, especially if we keep in mind that one of our goals is to handle Lipschitz continuous sources.

### 5.1 Proof of Theorem 4

Under  $(\mathbf{C}_{UE})$  or  $(\mathbf{C}_R)$ , it is well known that the Poisson equation (2.5) that we now recall:

$$\forall x \in \mathbb{R}^d, \mathcal{A}\varphi(x) = f(x) - \nu(f),$$

admits a unique solution centered w.r.t.  $\nu$  and with linear growth, in  $W_{p,\text{loc}}^2(\mathbb{R}^d, \mathbb{R})$  for any  $p > 1$  under  $(\mathbf{C}_{UE})$  (see [PV01]), or in  $\mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  under  $(\mathbf{C}_R)$  (see Proposition A.8 in [PP14]). In both cases, we have the following representation:

$$\varphi(x) = - \int_{\mathbb{R}_+} (P_t f(x) - \nu(f)) dt \quad \text{where} \quad P_t f(x) := \mathbb{E}[f(Y_t^{0,x})] \quad (5.1)$$

and  $Y_t^{0,x}$  solves (1.1) with  $Y_0^{0,x} = x$ . To comply with the framework of the above Theorems 5 and 6, the first step is to establish a pointwise gradient control.



### Gradient Control

Under  $(\mathbf{C}_{\mathbf{UE}})$  or  $(\mathbf{C}_{\mathbf{R}})$  we manage to obtain pointwise gradient bounds for  $\varphi$ . In our current confluent setting, these estimates are obtained through controls on the tangent flow, again without any *a priori* uniform ellipticity condition of type  $(\mathbf{UE})$ .

**Lemma 6 (Pointwise Gradient Bounds)** *Assume that  $(\mathbf{C}_{\mathbf{UE}})$  or  $(\mathbf{C}_{\mathbf{R}})$  holds. Then*

$$\|\nabla\varphi\|_{\infty} \leq \frac{[f]_1}{\alpha},$$

with  $\alpha$  as in  $(\mathbf{D}_{\alpha}^p)$ .

#### Proof: Gradient Control in the Confluent framework

Assume now that  $(\mathbf{D}_{\alpha}^p)$  holds. Observe that, as soon as  $(\mathbf{R}_{1,\beta})$  holds, it is well known that  $\nabla_x Y_t^{0,x}$  is well defined and belongs to  $L^2(\mathbb{P})$ , see [IW80]. Hence, for  $t > 0$ ,  $i \in \llbracket 1, d \rrbracket$ :

$$\begin{aligned} \partial_{x_i} \mathbb{E}[f(Y_t^{0,x})] &= \mathbb{E}[\langle \nabla f(Y_t^{0,x}), \partial_{x_i} Y_t^{0,x} \rangle], \\ \partial_{x_i} Y_t^{0,x} &= e_i + \int_0^t Db(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} ds + \sum_{j=1}^d \int_0^t D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} dW_s^j, \end{aligned}$$

where  $e_i$  stands for the  $i^{\text{th}}$  canonical vector and  $Db, D\sigma_{\cdot j} \in \mathbb{R}^d \otimes \mathbb{R}^d$ .

Let  $p \in (1, 2]$  be given such that  $(\mathbf{D}_{\alpha}^p)$  holds. Considering the mapping  $y \in \mathbb{R}^d \mapsto |y|^p$ , where  $|\cdot|$  stands for the Euclidean norm of  $\mathbb{R}^d$ , it is easily seen from Itô's formula that:

$$\begin{aligned} &|\partial_{x_i} Y_t^{0,x}|^p \\ &= 1 + p \int_0^t \left\langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, Db(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \right\rangle |\partial_{x_i} Y_s^{0,x}|^p ds \\ &\quad + p \sum_{j=1}^d \int_0^t \left\langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, D\sigma_{\cdot j}(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \right\rangle |\partial_{x_i} Y_s^{0,x}|^p dW_s^j \\ &\quad + \frac{p}{2} \sum_{j=1}^d \int_0^t \left( \frac{|D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x}|^2}{|\partial_{x_i} Y_s^{0,x}|^2} + (p-2) \frac{|\langle \partial_{x_i} Y_s^{0,x}, D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} \rangle|^2}{|\partial_{x_i} Y_s^{0,x}|^4} \right) |\partial_{x_i} Y_s^{0,x}|^p ds \\ &= \exp \left( p \int_0^t \left\langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, Db(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \right\rangle ds \right) \times \mathcal{E}(M)_t \\ &\quad \times \exp \left( \frac{p}{2} \sum_{j=1}^d \int_0^t \left( \frac{|D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x}|^2}{|\partial_{x_i} Y_s^{0,x}|^2} + (p-2) \frac{|\langle \partial_{x_i} Y_s^{0,x}, D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} \rangle|^2}{|\partial_{x_i} Y_s^{0,x}|^4} \right) ds \right), \end{aligned} \tag{5.2}$$

where  $(M_t)_{t \geq 0} := \left( p \sum_{j=1}^d \int_0^t \left\langle \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|}, D\sigma_{\cdot j}(Y_s^{0,x}) \frac{\partial_{x_i} Y_s^{0,x}}{|\partial_{x_i} Y_s^{0,x}|} \right\rangle dW_s^j \right)_{t \geq 0}$  is a square integrable martingale with bounded integrand and  $\mathcal{E}(M)_t := \exp(M_t - \frac{1}{2} \langle M \rangle_t)$  denotes the associated Doléans exponential martingale. From condition  $(\mathbf{D}_{\alpha}^p)$ , we thus get:

$$|\partial_{x_i} Y_t^{0,x}|^p \leq \exp(-\alpha p t) \times \mathcal{E}(M_t). \tag{5.3}$$

We eventually derive:

$$\int_0^{+\infty} |\mathbb{E}[\langle \nabla f(Y_t^{0,x}), \partial_{x_i} Y_t^{0,x} \rangle]| dt \leq [f]_1 \int_0^{+\infty} \mathbb{E}[|\partial_{x_i} Y_t^{0,x}|^p]^{1/p} dt \leq [f]_1 \int_0^{+\infty} \exp(-\alpha t) dt = \frac{[f]_1}{\alpha}.$$

From the above control and equation (5.1), we thus derive:

$$\forall i \in \llbracket 1, d \rrbracket, \forall x \in \mathbb{R}^d, |\partial_{x_i} \varphi(x)| \leq \frac{[f]_1}{\alpha}. \quad (5.4)$$

Similarly, for all  $x \in \mathbb{R}^d$ ,  $\nabla \varphi(x) = \int_0^{+\infty} \mathbb{E}[(\nabla Y_t^{0,x})^* \nabla f(Y_t^{0,x})] dt$  where we have  $\nabla Y_t^{0,x} = (\partial_{x_1} Y_t^{0,x} \ \dots \ \partial_{x_d} Y_t^{0,x})$  so that  $(\nabla Y_t^{0,x})^* = \begin{pmatrix} (\partial_{x_1} Y_t^{0,x})^* \\ \vdots \\ (\partial_{x_d} Y_t^{0,x})^* \end{pmatrix}$ . Hence, recalling that  $|\cdot|$

stands for the Euclidean norm,  $|\nabla \varphi(x)| \leq \int_0^{+\infty} \mathbb{E}[\|(\nabla Y_t^{0,x})^*\| \|\nabla f(Y_t^{0,x})\|] dt$  where we recall that for  $A \in \mathbb{R}^d \otimes \mathbb{R}^d$ ,  $\|A\| := \sup_{|z| \leq 1, z \in \mathbb{R}^d} |Az|$  denotes the operator (or spectral) matrix norm. Thus,  $|\nabla \varphi(x)| \leq \|\nabla f\|_\infty \int_0^{+\infty} \mathbb{E}[\|(\nabla Y_t^{0,x})^*\|^p]^{1/p} dt = \|\nabla f\|_\infty \int_0^{+\infty} \mathbb{E}[\|\nabla Y_t^{0,x}\|^p]^{1/p} dt$ . Now,

$$\|\nabla Y_t^{0,x}\| = \sup_{|z| \leq 1} |\nabla Y_t^{0,x} z|.$$

For any  $z \in \mathbb{R}^d$ ,  $|z| \leq 1$ , setting  $Z_t^{0,x,z} := \nabla Y_t^{0,x} z$ , one has the following dynamics for the  $\mathbb{R}^d$ -valued process  $(Z_s^{0,x,z})_{s \in [0,t]}$ :

$$Z_t^{0,x,z} := z + \int_0^t Db(Y_s^{0,x}) Z_s^{0,x,z} ds + \sum_{j=1}^d \int_0^t D\sigma_{\cdot j}(Y_s^{0,x}) Z_s^{0,x,z} dW_s^j.$$

Hence, we derive similarly to (5.3) that  $|Z_t^{0,x,z}|^p \leq |z|^p \exp(-p\alpha t) \mathcal{E}(M_t)$ , where  $\mathcal{E}(M_t)$  does not depend on  $z$ . Write now,

$$\mathbb{E}[\|\nabla Y_t^{0,x}\|^p]^{1/p} = \mathbb{E}[\sup_{|z| \leq 1} |Z_t^{0,x,z}|^p]^{1/p} \leq \mathbb{E}[\sup_{|z| \leq 1} |z|^p \exp(-p\alpha t) \mathcal{E}(M_t)]^{1/p} \leq \exp(-\alpha t). \quad (5.5)$$

This eventually proves the claim  $\|\nabla \varphi\|_\infty := \sup_{x \in \mathbb{R}^d} |\nabla \varphi(x)| \leq \frac{\|\nabla f\|_\infty}{\alpha}$ .  $\square$

### Additional smoothness

- Theorem 4 can be derived under  $(\mathbf{C}_R)$ , by iterating computations similar to the ones performed in Lemma 6. On the other hand, to have the required smoothness, since we cannot expect some smoothing effect from a non-degenerate diffusion coefficient, we have to impose that  $b, \sigma, f$  themselves lie in  $\mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  and the restriction on the variations of  $\sigma$  which ensures exponential integrability in time for the expectations of the iterated tangent flows, see Lemma A.8 in [PP14] for details (see the parallel between the above condition on  $D\sigma$  and assumption  $(\mathbf{AC}_p)$  appearing p. 559 in [PP14]).
- Proving Theorem 4 under  $(\mathbf{C}_{UE})$  requires more sophisticated tools (Schauder estimates for operators with unbounded coefficients).

**Proof: Theorem 4 under (C<sub>UE</sub>)**

Let us begin with the scalar case. For  $d = 1$ , set for all  $x \in \mathbb{R}$ ,

$$\begin{aligned} v(x) &:= - \int_0^{+\infty} dt \mathbb{E}[\Psi(Y_t^{0,x}) \partial_x Y_t^{0,x}] \\ &= - \int_0^{+\infty} dt \mathbb{E} \left[ \Psi(Y_t^{0,x}) \exp \left( \int_0^t b'(Y_s^{0,x}) ds \right) \mathcal{E} \left( \int_0^t \sigma'(Y_s^{0,x}) dW_s \right) \right], \end{aligned} \quad (5.6)$$

where for all  $y \in \mathbb{R}$ ,  $\Psi(y) := \partial_y f(y)$ . We observe that  $\partial_x \varphi(x) = v(x)$ . Also, from our assumptions on  $f, b, \sigma$ , we have that  $\Psi, b', \sigma' \in \mathcal{C}_b^{0,\beta}(\mathbb{R}^d, \mathbb{R})$ . Theorems 2.4-2.6 in Krylov and Priola, [KP10] then yield the existence of a unique solution to the PDE:

$$\tilde{\mathcal{A}}w(x) + b'(x)w(x) = \Psi(x), \quad \text{where} \quad \tilde{\mathcal{A}}w(x) = \mathcal{A}w(x) + \sigma\sigma'(x)w'(x), \quad (5.7)$$

belonging to  $\mathcal{C}_b^{2,\beta}(\mathbb{R}^d, \mathbb{R})$  and such that the following Schauder estimate holds:

$$\exists C \geq 1, \quad \|w\|_{2,\beta} \leq C(1 + \|\Psi\|_\beta). \quad (5.8)$$

Indeed, from  $(\mathbf{D}_\alpha^p)$ , we get that  $b'(x) \leq -\alpha < 0$  and the potential in (5.7) has the good sign. From (5.6) and the Girsanov theorem, we also get:

$$v(x) = - \int_0^{+\infty} dt \mathbb{E} \left[ \Psi(\tilde{Y}_t^{0,x}) \exp \left( \int_0^t b'(\tilde{Y}_s^{0,x}) ds \right) \right],$$

where  $d\tilde{Y}_s^{0,x} = (b(\tilde{Y}_s^{0,x}) + \sigma\sigma'(\tilde{Y}_s^{0,x}))ds + \sigma(\tilde{Y}_s^{0,x})dW_s$ . Note that  $\tilde{Y}$  has generator  $\tilde{\mathcal{A}}$ . A simple identification procedure, similar to the proof of Theorem II.1.1 in Bass [Bas97] then gives  $v = w$ . The result follows from (5.8). Let us emphasize that this is a quite deep and involved result for unbounded coefficients.

In the multi-dimensional setting, recalling the technical condition that for all  $i \in \llbracket 1, d \rrbracket$ ,  $j \geq i$ ,  $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_i, \dots, x_d)$ , we have that differentiating formally the PDE (2.5) in the space variable  $x_i$ ,  $i \in \llbracket 1, d \rrbracket$  yields that  $\partial_{x_i} \varphi = v_i$  should satisfy:

$$\begin{aligned} \tilde{\mathcal{A}}w_i(x) + \partial_{x_i} b_i(x)w_i(x) &= \Psi_i(x) - \sum_{j \in \llbracket 1, d \rrbracket \setminus \{i\}} \partial_{x_i} b_j(x)v_j(x) \\ &\quad - \frac{1}{2} \sum_{j \in \llbracket 1, i-1 \rrbracket} \partial_{x_i} \Sigma_{j,j}(x) \partial_{x_j} v_j(x) - \sum_{j \in \llbracket 1, i-1 \rrbracket} \sum_{k \in \llbracket j+1, d \rrbracket \setminus \{i\}} \partial_{x_i} \Sigma_{j,k}(x) \partial_{x_j} v_k(x), \end{aligned} \quad (5.9)$$

with  $\Psi_i(x) := \partial_{x_i} f(x)$  and

$$\tilde{\mathcal{A}}w_i(x) := \mathcal{A}w_i(x) + \frac{1}{2} \partial_{x_i} \Sigma_{i,i}(x) \partial_{x_i} w_i(x) + \sum_{j \in \llbracket 1, d \rrbracket \setminus \{i\}} \partial_{x_i} \Sigma_{i,j}(x) \partial_{x_j} w_i(x).$$

We would now like to enter the previous framework of Schauder estimates. To do so, we first observe from  $(\mathbf{D}_\alpha^p)$  and the Cauchy-Schwarz inequality that  $\partial_{x_i} b_i(x) \leq -\alpha < 0$ . Consider now  $i = 1$  in (5.9). From our current assumptions on  $f, b$  and the previous computations on the gradient for the multi-dimensional case, it remains to prove  $\tilde{\Psi}_1(x) := \Psi_1(x) - \sum_{j \neq 1} \partial_{x_1} b_j(x)v_j(x) \in \mathcal{C}_b^{0,\beta}(\mathbb{R}^d, \mathbb{R})$ . This will be the case, once we will

have proved that  $\nabla\varphi$  is  $\beta$ -Hölder continuous, which is a priori not direct. This property is assumed for the remaining of the proof and shown below. In particular, it leads to the restriction concerning the variations of  $\sigma$  when  $d > 1$ . Hence, Theorems 2.4-2.6 in Krylov and Priola, [KP10] still apply and give that there exists a unique solution  $w_1 \in \mathcal{C}_b^{2,\beta}(\mathbb{R}^d, \mathbb{R})$  to (5.9) which also satisfies:

$$\exists C \geq 1, \|w_1\|_{2,\beta} := \sum_{\alpha, |\alpha| \in \llbracket 0, 2 \rrbracket} \|D^\alpha w_1\|_\infty + [D^{(2)}w_1]_\beta \leq C(1 + \|\tilde{\Psi}_1\|_\beta) =: \bar{C}((\mathcal{L}_\mathbf{v}), (\mathbf{R}_{1,\beta}), (\mathbf{UE})). \quad (5.10)$$

The identification  $w_1 = \partial_{x_1}\varphi = v_1$  is standard. The control (5.10) allows to iterate, since it gives that  $\nabla w_1 = (\partial_{x_1}v_1, \dots, \partial_{x_d}v_1) = (\partial_{x_1, x_1}\varphi_1, \dots, \partial_{x_d, x_1}\varphi)$  is  $\beta$ -Hölder. We thus get by induction, from the specific chosen structure on  $\sigma$  and by Theorems 2.4-2.6 in Krylov and Priola, [KP10], that for all  $i \in \llbracket 1, d \rrbracket$  there exists a unique solution  $w_i \in \mathcal{C}_b^{2,\beta}(\mathbb{R}^d, \mathbb{R})$  to (5.9) such that:

$$\begin{aligned} \exists C \geq 1, \|w_i\|_{2,\beta} &\leq C(1 + \|\tilde{\Psi}_i\|_\beta) =: \bar{C}((\mathcal{L}_\mathbf{v}), (\mathbf{R}_{1,\beta}), (\mathbf{UE})), \\ \tilde{\Psi}_i(x) &:= \Psi_i(x) - \sum_{j \in \llbracket 1, d \rrbracket \setminus \{i\}} \partial_{x_i} b_j(x) v_j(x) - \frac{1}{2} \sum_{\substack{1 \leq j < i, \\ k \in \llbracket 1, d \rrbracket \setminus \{i\}}} \partial_{x_i} \Sigma_{j,k}(x) \partial_{x_j} v_k(x). \end{aligned} \quad (5.11)$$

The Lipschitz property of the mapping  $x \mapsto \langle \nabla\varphi(x), b(x) \rangle$  is eventually derived following the procedure described in Remark 7. □

**Remark 12 (Structure of  $\sigma$ )** *We emphasize that the structure condition on  $\sigma$  assumed in Theorem 4 under  $(\mathbf{C}_{\mathbf{UE}})$  is mainly technical. It is of course always verified in dimension  $d = 1$ . For  $d > 1$  it is motivated by the fact that, differentiating (2.5) without this assumption yields to consider a system of coupled linear PDEs with growing coefficients for which the Schauder estimates have not been established yet. Following the existing literature for Schauder estimates for systems (see e.g. Boccia [Boc13]), we think that the results of Krylov and Priola should extend to this case. This would allow to get rid of the indicated condition. Here, the condition simply allows to decouple the system.*

*Let us mention too that the results by Priola [Pri09] could also be a starting point to investigate the smoothness of the Poisson problem for degenerate kinetic models.*

*These aspects will concern further research.*

*Additional Smoothness continued:  $\beta$ -Hölder continuity of the gradient through pathwise analysis.* We control here, under  $(\mathbf{D}_\alpha^p)$ ,  $p \in (1, 2]$  and  $(\mathbf{R}_{1,\beta})$ ,  $\beta \in (0, 1]$ , the  $\beta$ -Hölder modulus of continuity of the gradient. We will progressively see how the restrictions on  $D\sigma$  come out. For  $(x, x') \in \mathbb{R}^{2d}$ , write for all  $i \in \llbracket 1, d \rrbracket$ :

$$\begin{aligned} & \left| \partial_{x_i} \varphi(x) - \partial_{x_i} \varphi(x') \right| \\ &= \left| \int_0^{+\infty} \left( \mathbb{E} [\langle \nabla f(Y_t^{0,x}), \partial_{x_i} Y_t^{0,x} \rangle] - \mathbb{E} [\langle \nabla f(Y_t^{0,x'}), \partial_{x_i} Y_t^{0,x'} \rangle] \right) dt \right| \\ &\leq \left| \int_0^{+\infty} \left( [\nabla f]_\beta \mathbb{E} [|Y_t^{0,x} - Y_t^{0,x'}|^\beta |\partial_{x_i} Y_t^{0,x}|] + \|\nabla f\|_\infty \mathbb{E} [|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|] \right) dt \right| \\ &=: (G_1^\beta + G_2^\beta)(x, x'). \end{aligned} \quad (5.12)$$

Let us first deal with the expectation in  $G_1^\beta$ . Namely, write

$$\mathbb{E}[|Y_t^{0,x} - Y_t^{0,x'}|^\beta |\partial_{x_i} Y_t^{0,x}|] \leq \mathbb{E}[|Y_t^{0,x} - Y_t^{0,x'}|^{\bar{p}\beta}]^{\frac{1}{\bar{p}}} \mathbb{E}[|\partial_{x_i} Y_t^{0,x}|^{\bar{q}}]^{\frac{1}{\bar{q}}}, \quad \bar{p}, \bar{q} > 1, \quad \bar{p}^{-1} + \bar{q}^{-1} = 1.$$

Take now  $\bar{p}\beta = \bar{q} \iff \bar{p} = \frac{1+\beta}{\beta}$ ,  $\bar{q} = 1 + \beta$  which leads to the same integrability constraints on the flows.

If  $\beta + 1 \leq p$  in  $(\mathbf{D}_\alpha^p)$ , then we readily get similarly to (5.3) that  $\mathbb{E}[|\partial_{x_i} Y_t^{0,x}|^{\bar{q}}]^{\frac{1}{\bar{q}}} \leq \exp(-\alpha t)$ .

If now  $\beta + 1 > p$ , as soon as  $(\mathbf{D}_{\bar{\alpha}}^{1+\beta})$  holds for some  $\bar{\alpha} > 0$ , which is actually the case provided that

$$\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{1 + \beta - p}, \quad (5.13)$$

for  $\bar{q} = 1 + \beta$ , we again get similarly to (5.3) that  $\mathbb{E}[|\partial_{x_i} Y_t^{0,x}|^{\bar{q}}]^{\frac{1}{\bar{q}}} \leq \exp(-\bar{\alpha} t)$ . On the other hand, the mean value theorem yields:

$$\begin{aligned} \mathbb{E}[|Y_t^{0,x} - Y_t^{0,x'}|^{\bar{p}\beta}]^{\frac{1}{\bar{p}}} &\leq |x - x'|^\beta \mathbb{E}\left[\int_0^1 d\lambda \|\nabla Y_t^{0,x'+\lambda(x-x')}\|^{\bar{p}\beta}\right]^{\frac{1}{\bar{p}}} \\ &\leq |x - x'|^\beta \left(\int_0^1 d\lambda \mathbb{E}[\|\nabla Y_t^{0,x'+\lambda(x-x')}\|^{\bar{p}\beta}]\right)^{\frac{1}{\bar{p}}} \\ &\leq |x - x'|^\beta [\exp(-\alpha\beta t)\mathbb{I}_{1+\beta \leq p} + \exp(-\bar{\alpha}\beta t)\mathbb{I}_{1+\beta > p}], \end{aligned}$$

exploiting (5.5) for the last inequality provided that (5.13), which in turn implies that  $(\mathbf{D}_{\bar{\alpha}}^{1+\beta})$  for some  $\bar{\alpha} > 0$ , holds if  $1 + \beta > p$ . Plugging these bounds in (5.12) gives that:

$$\forall (x, x') \in (\mathbb{R}^d)^2, \quad |G_1^\beta(x, x')| \leq \frac{[\nabla f]_\beta}{(1 + \beta)} \left[ \frac{\mathbb{I}_{1+\beta \leq p}}{\alpha} + \frac{\mathbb{I}_{1+\beta > p}}{\bar{\alpha}} \right] |x - x'|^\beta. \quad (5.14)$$

We already see that, when  $1 + \beta > p$ , for the parameter  $p$  of the initial confluence condition  $(\mathbf{D}_\alpha^p)$ , a first constraint on the variations of  $\sigma$ , namely (5.13) appears.

Let us now turn to  $G_2^\beta$ . Following the expansion of (5.2) write:

$$\begin{aligned} |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 &= 2 \int_0^t \left\langle \partial_{x_i} Y_s^{0,x} - \partial_{x_i} Y_s^{0,x'}, Db(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} - Db(Y_s^{0,x'}) \partial_{x_i} Y_s^{0,x'} \right\rangle ds \\ &\quad + 2 \sum_{j=1}^d \int_0^t \left\langle \partial_{x_i} Y_s^{0,x} - \partial_{x_i} Y_s^{0,x'}, D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} - D\sigma_{\cdot j}(Y_s^{0,x'}) \partial_{x_i} Y_s^{0,x'} \right\rangle dW_s^j \\ &\quad + \sum_{j=1}^d \int_0^t |D\sigma_{\cdot j}(Y_s^{0,x}) \partial_{x_i} Y_s^{0,x} - D\sigma_{\cdot j}(Y_s^{0,x'}) \partial_{x_i} Y_s^{0,x'}|^2 ds. \end{aligned}$$

Let  $u(t) := \mathbb{E} |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2$ ,  $t \geq 0$ . First note that  $u(0) = 0$ . Taking now the expectation and interchanging expectation and time integration yields

$$u(t) = \int_0^t \mathbb{E} \Xi_s ds$$

where  $(\Xi_t)_{t \geq 0}$  is a pathwise continuous process clearly determined by the terms inside the above time integrals. One readily checks that,  $t \mapsto \mathbb{E} \Xi_s$  is continuous so that  $u$  is continuously differentiable and satisfies

$$\begin{aligned} u'(t) &= 2 \mathbb{E} \left\langle \partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}, Db(Y_t^{0,x}) \partial_{x_i} Y_t^{0,x} - Db(Y_t^{0,x'}) \partial_{x_i} Y_t^{0,x'} \right\rangle \\ &\quad + \sum_{j=1}^d \mathbb{E} |D\sigma_{\cdot j}(Y_t^{0,x}) \partial_{x_i} Y_t^{0,x} - D\sigma_{\cdot j}(Y_t^{0,x'}) \partial_{x_i} Y_t^{0,x'}|^2. \end{aligned}$$

Using the Young inequality for a parameter  $\varepsilon \in (0, 1]$ , small enough and to be chosen further, we derive:

$$\begin{aligned} u'(t) &\leq 2 \mathbb{E} \left[ \left\langle \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|}, Db(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right\rangle |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right. \\ &\quad \left. + \int_0^t \|Db(Y_t^{0,x}) - Db(Y_t^{0,x'})\| |\partial_{x_i} Y_t^{0,x'}| |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}| \right] \\ &\quad + \mathbb{E} \left[ (1 + \varepsilon) \sum_{j=1}^d \left| D\sigma_{\cdot j}(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right|^2 |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right. \\ &\quad \left. + (1 + \varepsilon^{-1}) \sum_{j=1}^d \|D\sigma_{\cdot j}(Y_t^{0,x}) - D\sigma_{\cdot j}(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2 \right]. \end{aligned}$$

From this computation, the point is now to make the confluence condition  $(\mathbf{D}_\alpha^p)$  appear and to separate the components for which we will exploit the  $\beta$ -Hölder continuity, namely  $Db, (D\sigma_{\cdot j})_{j \in \llbracket 1, n \rrbracket}$ . To do so we first observe that:

$$\begin{aligned} \left\langle \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|}, Db(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right\rangle + \frac{1}{2} \sum_{j=1}^d \left| D\sigma_{\cdot j}(Y_t^{0,x}) \frac{\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}}{|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|} \right|^2 \\ \leq -\alpha + (2-p) \frac{1}{2} \|D\sigma\|_\infty^2 = -\tilde{\alpha}, \end{aligned}$$

where we suppose from now on that

$$-\tilde{\alpha} := -\alpha + (2-p) \frac{1}{2} \|D\sigma\|_\infty^2 < 0 \iff \|D\sigma\|_\infty^2 < \frac{2\alpha}{2-p}. \quad (5.15)$$

Hence,

$$\begin{aligned} u'(t) &\leq 2 \mathbb{E} \left[ (-\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2) |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right] \\ &\quad + 2 \mathbb{E} \left[ \|Db(Y_t^{0,x}) - Db(Y_t^{0,x'})\| |\partial_{x_i} Y_t^{0,x'}| |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}| \right] \\ &\quad + \mathbb{E} \left[ (1 + \varepsilon^{-1}) \sum_{j=1}^d \|D\sigma_{\cdot j}(Y_t^{0,x}) - D\sigma_{\cdot j}(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2 \right]. \end{aligned}$$

Using now again the Young inequality, with  $\eta \in (0, 1]$  small enough, for the middle term of the above r.h.s., we obtain:

$$\begin{aligned}
 u'(t) &\leq 2 \left( -\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2 + \frac{\eta}{2} \right) \mathbb{E} \left[ |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right] \\
 &\quad + \eta^{-1} \mathbb{E} \left[ \|Db(Y_t^{0,x}) - Db(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2 \right] \\
 &\quad + (1 + \varepsilon^{-1}) \sum_{j=1}^d \mathbb{E} \left[ \|D\sigma_{\cdot j}(Y_t^{0,x}) - D\sigma_{\cdot j}(Y_t^{0,x'})\|^2 |\partial_{x_i} Y_t^{0,x'}|^2 \right] \\
 &\leq 2 \left( -\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2 + \frac{\eta}{2} \right) \mathbb{E} \left[ |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right] \\
 &\quad + \eta^{-1} [Db]_\beta^2 \mathbb{E} \left[ |Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2 \right] \\
 &\quad + (1 + \varepsilon^{-1}) [D\sigma]_\beta^2 \mathbb{E} \left[ |Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2 \right]. \tag{5.16}
 \end{aligned}$$

Denote:

$$-\tilde{\alpha}_{\varepsilon,\eta,\sigma} := -\tilde{\alpha} + \frac{\varepsilon}{2} \|D\sigma\|_\infty^2 + \frac{\eta}{2} < 0,$$

for  $\varepsilon, \eta$  small enough. Setting for every  $t \geq 0$ ,

$$r(t) := \eta^{-1} [Db]_\beta^2 \mathbb{E} \left[ |Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2 \right] + (1 + \varepsilon^{-1}) [D\sigma]_\beta^2 \mathbb{E} \left[ |Y_t^{0,x} - Y_t^{0,x'}|^{2\beta} |\partial_{x_i} Y_t^{0,x'}|^2 \right],$$

equation (5.16) reads an ordinary differential inequation:

$$u'(t) \leq -2\tilde{\alpha}_{\varepsilon,\eta,\sigma} u(t) + r(t), \quad u(0) = 0.$$

We derive from the Gronwall lemma that

$$u(t) = \mathbb{E} \left[ |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right] \leq \exp(-2\tilde{\alpha}_{\varepsilon,\eta,\sigma} t) \int_0^t \exp(2\tilde{\alpha}_{\varepsilon,\eta,\sigma} s) r(s) ds.$$

Reproducing as well the computations that led to (5.14), we derive:

$$\begin{aligned}
 u(t) &\leq C_{\eta,\varepsilon,\beta} |x - x'|^{2\beta} \int_0^t \exp(-2\tilde{\alpha}_{\varepsilon,\eta,\sigma}(t-s)) \left( \mathbb{E} \left[ |\partial_{x_i} Y_s^{0,x'}|^{2(1+\beta)} \right]^{\frac{1}{1+\beta}} \right. \\
 &\quad \left. + \int_0^1 d\lambda \mathbb{E} \left[ \|\nabla Y_s^{0,x'+\lambda(x-x')}\|^{2(1+\beta)} \right]^{\frac{\beta}{1+\beta}} \right) ds.
 \end{aligned}$$

From the analysis leading to (5.3), (5.5) we now derive:

$$\begin{aligned}
 &\mathbb{E} \left[ |\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2 \right] \\
 &\leq \frac{C_{\eta,\varepsilon,\beta}}{2} |x - x'|^{2\beta} \exp(-2\tilde{\alpha}_{\varepsilon,\eta,\sigma} t) \left[ \frac{\exp(2(\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \tilde{\alpha}_{2(1+\beta)})t)}{\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \tilde{\alpha}_{2(1+\beta)}} + \frac{\exp(2(\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \beta\tilde{\alpha}_{2(1+\beta)})t)}{\tilde{\alpha}_{\varepsilon,\eta,\sigma} - \beta\tilde{\alpha}_{2(1+\beta)}} \right],
 \end{aligned}$$

and

$$-\tilde{\alpha}_{2(1+\beta)} \leq -\alpha + (2(1+\beta) - p) \frac{1}{2} \|D\sigma\|_\infty^2.$$

Thus,  $\tilde{\alpha}_{2(1+\beta)} < 0$  as soon as

$$\|D\sigma\|_\infty^2 < \frac{2\alpha}{2(1+\beta) - p}, \quad (5.17)$$

which is precisely the restriction on the variations of  $\sigma$  appearing in  $(\mathbf{C}_{\text{UE}})$  when  $d > 1$ , then  $\tilde{\alpha}_{2(1+\beta)} > 0$  and:

$$\mathbb{E}[|\partial_{x_i} Y_t^{0,x} - \partial_{x_i} Y_t^{0,x'}|^2] \leq \bar{C}_{\eta,\varepsilon,\beta} |x - x'|^{2\beta} \exp(-\tilde{\alpha}_{2(1+\beta)} t).$$

This last control then gives the expected bound for the  $\beta$ -Hölder modulus of the gradient. Namely, from (5.12), (5.14),

$$[\partial_{x_i} \varphi]_\beta < \frac{[\nabla f]_\beta}{(1+\beta)} \left[ \frac{\mathbb{I}_{1+\beta \leq p}}{\alpha} + \frac{\mathbb{I}_{1+\beta > p}}{\bar{\alpha}} \right] + \frac{\|\nabla f\|_\infty \bar{C}_{\eta,\varepsilon,\beta}}{\tilde{\alpha}_{2(1+\beta)}}.$$

## 5.2 Proof of the Practical Results of Section 2.3

We first begin with the proof of the

### Slutsky like Theorem 6

We keep here for simplicity the generic notation  $\|\cdot\|$  for any admissible matrix norm according to the assumptions of the theorem. We first write:

$$\mathbb{P} \left[ \sqrt{\Gamma_n} \frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}} \geq a \right] = \mathbb{P} \left[ \nu_n(\mathcal{A}\varphi) \geq \frac{a}{\sqrt{\Gamma_n}} \sqrt{\nu_n(\|\sigma\|^2)} \right]. \quad (5.18)$$

We then proceed similarly to Theorem 8, with an exponential Bienaymé-Tchebychev inequality, for all  $\lambda > 0$  we have:

$$\begin{aligned} & \mathbb{P} \left[ \sqrt{\Gamma_n} \frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}} \geq a \right] \leq \mathbb{E} \left[ \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{\nu_n(\|\sigma\|^2)} \right) \exp(\lambda \nu_n(\mathcal{A}\varphi)) \right] \\ &= \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{\nu(\|\sigma\|^2)} \right) \mathbb{E} \left[ \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \left[ \sqrt{\nu_n(\|\sigma\|^2)} - \sqrt{\nu(\|\sigma\|^2)} \right] \right) \exp(\lambda \nu_n(\mathcal{A}\varphi)) \right] \\ &= \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{\nu(\|\sigma\|^2)} \right) \mathbb{E} \left[ \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\|\sigma\|^2) - \nu(\|\sigma\|^2)}{\sqrt{\nu_n(\|\sigma\|^2)} + \sqrt{\nu(\|\sigma\|^2)}} \right) \exp(\lambda \nu_n(\mathcal{A}\varphi)) \right] \\ &= \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{\nu(\|\sigma\|^2)} \right) \mathbb{E} \left[ \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)} + \sqrt{\nu(\|\sigma\|^2)}} \right) \exp(\lambda \nu_n(\mathcal{A}\varphi)) \right]. \end{aligned}$$

By the Hölder inequality, for  $\tilde{p}, \tilde{q} > 1$ , such that  $\frac{1}{\tilde{p}} + \frac{1}{\tilde{q}} = 1$ :

$$\begin{aligned} & \mathbb{P} \left[ \sqrt{\Gamma_n} \frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}} \geq a \right] \\ & \leq \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{\nu(\|\sigma\|^2)} \right) \left[ \mathbb{E} \exp \left( -\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2)} + \sqrt{\nu(\|\sigma\|^2)}} \right) \right]^{1/\tilde{p}} \left[ \mathbb{E} \exp(\lambda \tilde{q} \nu_n(\mathcal{A}\varphi)) \right]^{1/\tilde{q}}. \end{aligned}$$



The proof of Theorem 8 yields:

$$\begin{aligned} \mathbb{P} \left[ \sqrt{\Gamma_n} \frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}} \geq a \right] &\leq \mathcal{R}_n \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{\nu(\|\sigma\|^2)} \right) \exp \left( \frac{\rho \tilde{q} \lambda^2}{\Gamma_n} \tilde{A}_n + \frac{\rho^3 \tilde{q}^3 \lambda^4}{(\rho-1)\Gamma_n} \tilde{B}_n \right) \\ &\quad \times \left( \mathbb{E} \exp \left( -\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2) + \sqrt{\nu(\|\sigma\|^2)}}} \right) \right)^{1/\tilde{p}}, \end{aligned} \quad (5.19)$$

where we recall from identity (4.5):

$$\tilde{A}_n = \frac{q[\varphi]_1^2 \nu(\|\sigma\|^2)}{2} + e_n \quad \text{and} \quad \tilde{B}_n = \frac{q^3[\varphi]_1^4}{4} \left( \frac{\bar{q}\|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + \bar{e}_n \right).$$

Also,  $\mathcal{R}_n \xrightarrow[n]{n} 1$  denotes a "generic" remainder. Observe that thanks to the bounds of Theorem 4 (stated in the above Lemma 6), we get:

$$\tilde{A}_n \leq \frac{q[f]_1^2 \nu(\|\sigma\|^2)}{2\alpha} + e_n. \quad (5.20)$$

Let us now handle the remainder  $\left[ \mathbb{E} \exp \left( -\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2) + \sqrt{\nu(\|\sigma\|^2)}}} \right) \right]^{1/\tilde{p}}$ :

$$\begin{aligned} &\left[ \mathbb{E} \exp \left( -\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2) + \sqrt{\nu(\|\sigma\|^2)}}} \right) \right]^{1/\tilde{p}} \\ &= \left[ \mathbb{E} \left( \exp \left[ -\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2) + \sqrt{\nu(\|\sigma\|^2)}}} \right] (\mathbf{1}_{\nu_n(\mathcal{A}\vartheta) \geq 0} + \mathbf{1}_{\nu_n(\mathcal{A}\vartheta) < 0}) \right) \right]^{1/\tilde{p}} \\ &\leq \left( \left[ \mathbb{E} \exp \left( \frac{a\tilde{p}^2\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu(\|\sigma\|^2)}} \right) \right]^{1/\tilde{p}} \mathbb{P}[\nu_n(\mathcal{A}\vartheta) \geq 0]^{1/\tilde{q}} \right. \\ &\quad \left. + \left[ \mathbb{E} \exp \left( -\frac{a\tilde{p}^2\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu(\|\sigma\|^2)}} \right) \right]^{1/\tilde{p}} \mathbb{P}[\nu_n(\mathcal{A}\vartheta) < 0]^{1/\tilde{q}} \right)^{1/\tilde{p}}. \end{aligned}$$

Let us mention that we introduced the above partition in order to get a sharper constant in the final inequality, 1 below instead of 2, which would follow getting rid of the indicator functions. Now, by Theorem 2 we easily get:

$$\begin{aligned} &\mathbb{E} \left[ \exp \left( -\frac{a\tilde{p}\lambda}{\sqrt{\Gamma_n}} \frac{\nu_n(\mathcal{A}\vartheta)}{\sqrt{\nu_n(\|\sigma\|^2) + \sqrt{\nu(\|\sigma\|^2)}}} \right) \right]^{1/\tilde{p}} \\ &\leq \mathcal{R}_n \exp \left( \frac{a^2 \tilde{p}^2}{\Gamma_n \nu(\|\sigma\|^2)} \frac{\lambda^2}{\Gamma_n} \left( \frac{\|\sigma\|_\infty^2 \|\nabla \vartheta\|_\infty^2}{2} + e_n \right) \right) \\ &\quad \times \left[ \mathbb{P}[\nu_n(\mathcal{A}\vartheta) \geq 0]^{1/\tilde{q}} + \mathbb{P}[\nu_n(\mathcal{A}\vartheta) < 0]^{1/\tilde{q}} \right]^{1/\tilde{p}}. \end{aligned} \quad (5.21)$$

We choose  $\tilde{p} := \tilde{p}(n) \rightarrow +\infty$ , such that  $\frac{\tilde{p}^2 a^2}{\Gamma_n} \rightarrow 0$ , and so  $(\mathbb{P}[\nu_n(\mathcal{A}\vartheta) \geq 0]^{1/\tilde{q}} + \mathbb{P}[\nu_n(\mathcal{A}\vartheta) < 0]^{1/\tilde{q}})^{1/\tilde{p}} \leq 2^{1/\tilde{p}} \rightarrow 1$ . Moreover, exploiting again that for the *Gaussian regime*,  $\frac{\tilde{p}^2 a^2}{\Gamma_n} \rightarrow 0$ , we obtain by (5.21) and (5.19):

$$\begin{aligned} & \mathbb{P} \left[ \sqrt{\Gamma_n} \frac{\nu_n(f) - \nu(f)}{\sqrt{\nu_n(\|\sigma\|^2)}} \geq a \right] \\ & \leq \mathcal{R}_n \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \sqrt{\nu(\|\sigma\|^2)} \right) \exp \left( \frac{\rho\tilde{q}\lambda^2}{\Gamma_n} (\tilde{A}_n + e_n) + \frac{\rho^3\tilde{q}^3\lambda^4}{(\rho-1)\Gamma_n} \tilde{B}_n \right). \end{aligned} \quad (5.22)$$

From identity (5.22), the optimization over  $\lambda$  is similar to the one performed in the proof of Theorem 8. This yields the deviation bound (2.10). The non-asymptotic confidence interval in (2.11) is derived as for Theorem 5 from the gradient bounds of Theorem 4 and (2.10).

### 5.3 Regularization of Lipschitz Sources

We assume here that assumptions **(C2)**, **(L<sub>v</sub>)**, **(UE)** are in force. We suppose as well that the following smoothness holds for  $b, \sigma$ :

**(R<sub>b,σ</sub>)** Regularity and Structure. We assume that there exists  $\beta \in (0, 1)$  such that  $b, \sigma$  in (1.1) belong to  $\mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$  and  $\mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$  respectively. Also, for all  $(i, j) \in \llbracket 1, d \rrbracket^2$ ,  $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d)$ .

Importantly, we are interested, under assumptions **(C2)**, **(L<sub>v</sub>)**, **(UE)**, **(R<sub>b,σ</sub>)**, in giving controls for the estimation of  $\nu(f)$  when the source  $f$  is simply Lipschitz continuous. This is indeed the natural framework for the source which can be handled through functional inequality techniques, see [MT06], [Boi11].

To comply with our previous framework, namely to exploit the smoothness result of Theorem 4 under **(C<sub>UE</sub>)**, we need to regularize the source. Let  $\eta$  be a mollifier (i.e. a non-negative compactly supported function such that  $\int_{\mathbb{R}^d} \eta(x) dx = 1$ ). Define for  $\delta > 0$ ,  $\eta_\delta(x) = \frac{1}{\delta^d} \eta(\frac{x}{\delta})$ . We regularize  $f$  introducing  $f_\delta := f \star \eta_\delta$  where  $\star$  stands for the convolution on  $\mathbb{R}^d$ . From usual estimates, we obtain:

$$\begin{aligned} \exists C_\eta > 0, \quad \forall x \in \mathbb{R}^d, \quad & |f_\delta(x) - f(x)| \leq C_\eta \delta [f]_1, \\ \forall \beta \in (0, 1), \quad & [\nabla f_\delta]_\beta \leq C_\eta [f]_1 \delta^{-\beta}. \end{aligned} \quad (5.23)$$

We emphasize here that we will choose  $\beta$  later in order to be compatible with a certain range of step sequences. We assume for simplicity that  $\theta \in (1/3, 1]$  (no bias). Recall that we want to investigate:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}(\nu_n(f) - \nu(f)) \geq a] &= \mathbb{P}\left[(\nu_n(f_\delta) - \nu(f_\delta)) + R_{n,\delta}(f) \geq \frac{a}{\sqrt{\Gamma_n}}\right], \\ R_{n,\delta}(f) &:= [(\nu_n(f) - \nu(f)) - (\nu_n(f_\delta) - \nu(f_\delta))]. \end{aligned} \quad (5.24)$$

From (5.23), one readily gets:

$$|R_{n,\delta}(f)| \leq 2C_\eta \delta [f]_1. \quad (5.25)$$

On the other hand, the coefficients  $b, \sigma$  and the source  $f_\delta$  satisfy assumption  $(\mathbf{R}_{1,\beta})$  (observe indeed that the mollified function  $f_\delta \in C^{1,\beta}(\mathbb{R}^d, \mathbb{R})$ ). Hence, Theorem 4 yields that there exists a unique solution  $\varphi_\delta \in C^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  to the equation:

$$\mathcal{A}\varphi_\delta = f_\delta - \nu(f_\delta). \quad (5.26)$$

Observe from the proof of Theorem 4 under  $(\mathbf{C}_{\mathbf{UE}})$  (see equations (5.4) and (5.11)) and (5.23) that:

$$\begin{aligned} \|\nabla\varphi_\delta\|_\infty &\leq \alpha^{-1}[f]_1, \\ \forall \beta \in (0, 1), \exists C_\beta > 0, \forall i \in \{1, 2\}, [\varphi_\delta^{(i)}]_1 &\leq C_\beta(1 + \|\nabla f_\delta\|_{C^\beta}) \leq C_\beta\delta^{-\beta}, [\varphi_\delta^{(3)}]_\beta \leq C_\beta\delta^{-\beta}, \\ [\langle \nabla\varphi_\delta, b \rangle]_1 &\leq C_\beta\delta^{-\beta}. \end{aligned} \quad (5.27)$$

Now, from (5.26) the deviation in (5.24) rewrites:

$$\mathbb{P}\left[\sqrt{\Gamma_n}(\nu_n(f) - \nu(f)) \geq a\right] = \mathbb{P}\left[\nu_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f) \geq \frac{a}{\sqrt{\Gamma_n}}\right]. \quad (5.28)$$

From (5.25), the term  $R_{n,\delta}(f)$  can be seen as a remainder as soon as  $\frac{a}{\sqrt{\Gamma_n}} \gg 2C_\eta\delta[f]_1 \geq |R_{n,\delta}(f)|$ . On the other hand, the deviations of  $\nu_n(\mathcal{A}\varphi_\delta)$  can be analyzed as above, reproducing the proofs of Theorems 2 and 8, replacing the bounds on  $([\varphi^{(i)}]_1)_{i \in \{1,2\}}, [\varphi^{(3)}]_\beta$  appearing therein by those of equation (5.27). Precisely, we get from (5.25), similarly to (3.30) (replacing the controls on  $\varphi$  by those on  $\varphi_\delta$  in the proofs of Lemmas 3 and 5):

$$\begin{aligned} &\mathbb{P}\left[\left|\nu_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f)\right| \geq \frac{a}{\sqrt{\Gamma_n}}\right] \\ &\leq 2\left[\mathbb{E}\exp\left(-\frac{q\lambda_n}{\Gamma_n}M_n\right)\right]^{\frac{1}{q}} \exp\left(-\frac{a\lambda_n}{\sqrt{\Gamma_n}}\left(1 - \frac{\sqrt{\Gamma_n}2C_\eta[f]_1\delta}{a}\right)\right) \\ &\quad \times \exp\left(\frac{\lambda_n^2}{2\Gamma_n p} + \frac{p(a_n^\delta)^2}{2}\right) \exp\left(\frac{3pC_{V,\varphi}^2\lambda_n^2}{c_V\Gamma_n^2} + \frac{c_V}{p}\right) (I_V^1)^{\frac{1}{2p}} \\ &\quad \times \exp\left(C_{3.26}^\delta \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}} \times \exp\left(C_{3.27}^\delta \frac{p\lambda_n^2(\Gamma_n^{(2)})^2}{\Gamma_n^2}\right) (I_V^1)^{\frac{1}{4p}}, \end{aligned} \quad (5.29)$$

where  $C_{3.27}^\delta := C_\beta\delta^{-\beta}\frac{(C\sqrt{C_V})^2}{c_V}$  and  $C_{3.26}^\delta := \frac{\|\sigma\|_\infty^4 C_\beta^2 \delta^{-2\beta} C_V}{4 c_V}$  precisely correspond to the modifications of the constants  $C_{3.27}$  and  $C_{3.26} := \frac{\|\sigma\|_\infty^4 [\varphi^{(2)}]_1^2 C_V}{4 c_V}$  introduced in the proof of Lemma 5 when replacing  $\|D^2\varphi\|_\infty$  by  $\|D^2\varphi_\delta\|_\infty \leq C_\beta\delta^{-\beta}$  and  $[\langle \nabla\varphi, b \rangle]_1 \leq C$  by  $[\langle \nabla\varphi_\delta, b \rangle]_1 \leq C_\beta\delta^{-\beta}C$ . Similarly,

$$a_n^\delta := \frac{[\varphi_\delta^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{\frac{(3+\beta)}{2}}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} \leq \frac{C_\beta\delta^{-\beta} \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{\frac{(3+\beta)}{2}}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}},$$

is obtained from the definition of  $a_n$  in Lemma 3 replacing  $[\varphi^{(3)}]_\beta$  by  $[\varphi_\delta^{(3)}]_\beta$ . From the above equation and Lemma 2 we get:

$$\begin{aligned} & \mathbb{P} \left[ |\nu_n(f) - \nu(f)| \geq \frac{a}{\sqrt{\Gamma_n}} \right] = \mathbb{P} \left[ \left| \nu_n(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f) \right| \geq \frac{a}{\sqrt{\Gamma_n}} \right] \\ & \leq 2(I_V^1)^{\frac{1}{p}} \exp \left( \frac{c_V}{p} + \frac{p(a_n^\delta)^2}{2} \right) \\ & \quad \times \exp \left( - \frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \left( 1 - \frac{\sqrt{\Gamma_n}4C_\eta[f]_1\delta}{a} \right. \right. \\ & \quad \left. \left. - \frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \left\{ \frac{p}{\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + 2[C_{3.26}^\delta + C_{3.27}^\delta](\Gamma_n^{(2)})^2 \right) + \frac{1}{p} \right\} \right) \right). \end{aligned}$$

The Young inequality yields that for all  $\varepsilon_n > 0$ :

$$\begin{aligned} & \mathbb{P} \left[ |\nu_n(f) - \nu(f)| \geq \frac{a}{\sqrt{\Gamma_n}} \right] \\ & \leq 2(I_V^1)^{\frac{1}{p}} \exp \left( \frac{c_V}{p} + \frac{p(a_n^\delta)^2}{2} + \varepsilon_n^{-1}\Gamma_n\delta^2 \right) \\ & \quad \times \exp \left( - \frac{a^2}{2q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \right. \\ & \quad \left. \left( 1 - \frac{1}{q\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2} \underbrace{\left\{ 2\varepsilon_n C_\eta^2[f]_1^2 + \frac{p}{\Gamma_n} \left( \frac{6C_{V,\varphi}^2}{c_V} + 2[C_{3.26}^\delta + C_{3.27}^\delta](\Gamma_n^{(2)})^2 \right) + \frac{1}{p} \right\}}_{=:d_n^\delta} \right) \right) \right). \end{aligned} \tag{5.30}$$

We now want to let  $p := p(n) \xrightarrow{n} +\infty$ ,  $\varepsilon_n \xrightarrow{n} 0$  so that the associated contributions in the above equation can be viewed as remainders. From the previous definitions of  $C_{3.27}^\delta$ ,  $C_{3.26}^\delta$ , we see that, to achieve this goal, two constraints need to be fulfilled: namely, we must choose  $\delta$ ,  $p$  such that

$$\varepsilon_n^{-1}\Gamma_n\delta^2 \xrightarrow{n} 0 \quad \text{and} \quad p(a_n^\delta)^2 \xrightarrow{n} 0.$$

Now, if  $\theta \in (1/2, 1]$  there exists  $\beta \in (0, 1)$  such that  $\Gamma^{(\frac{3+\beta}{2})} \leq C$ . In that case:  $a_n^\delta \leq \frac{C}{\sqrt{\Gamma_n}}\delta^{-\beta} = \Gamma_n^{-(\frac{1}{2}(1-\beta)-\beta\varepsilon)} \xrightarrow{n} 0$  for  $\delta = \Gamma_n^{-(\frac{1}{2}+\varepsilon)}$  and  $\varepsilon < \frac{1-\beta}{2\beta}$ . Taking  $p := p(n) = \Gamma_n^{(\frac{1}{2}(1-\beta)-\beta\varepsilon)}$  yields  $p(a_n^\delta)^2 \xrightarrow{n} 0$ . On the other hand,  $\varepsilon_n = \Gamma_n^{-\varepsilon}$  also yields  $\varepsilon_n^{-1}\Gamma_n\delta^2 = \Gamma_n^{-\varepsilon} \xrightarrow{n} 0$ .

For  $\theta \in (1/3, 1/2)$ ,  $\Gamma_n^{(\frac{3+\beta}{2})}$  diverges for all  $\beta \in (0, 1)$ , we then have  $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \leq Cn^{\frac{1}{2}-\theta(1+\frac{\beta}{2})}$ .

Hence, there exists  $\beta \in (0, 1)$  such that  $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \leq Cn^{\frac{1}{2}-\theta(1+\frac{\beta}{2})} \xrightarrow{n} 0$ . However, taking  $\delta = \Gamma_n^{-(\frac{1}{2}+\varepsilon)}$ , which seems to be an almost ‘‘necessary’’ choice to satisfy the first constraint  $\varepsilon_n^{-1}\Gamma_n\delta^2 \xrightarrow{n} 0$ , yields:

$$a_n^\delta = \delta^{-\beta} \frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} = n^{(1+\beta)(\frac{1}{2}-\theta)+\varepsilon\beta(1-\theta)} \xrightarrow{n} +\infty,$$

so that the second constraint cannot be fulfilled. This means that the regularization induces a constraint on the time steps which must not be too large. In other words, under the sole Lipschitz assumption on the source  $f$ , the fastest convergence regime is out of reach.

Summing up the previous computations, we complete the proof of Theorem 7.

## 6 Applications

### 6.1 Non-Asymptotic Deviation Bounds in the Almost Sure CLT

Let  $(U_n)_{n \geq 1}$  be an i.i.d sequence of centered  $d$ -dimensional random variables with unit covariance matrix. We define the sequence of normalized partial sums by  $Z_0 = 0$  and

$$Z_n := \frac{\sum_{k=1}^n U_k}{\sqrt{n}}, n \geq 1.$$

The almost sure Central Limit Theorem (denoted from now on *a.s.* CLT) describes how the weighted sum of the renormalized sums  $Z_n$  which appear in the *usual asymptotic* CLT, behaves viewed as a random measure. Precisely, it states that setting for  $k \geq 1$ ,  $\gamma_k = 1/k$ :

$$\nu_n^Z := \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \delta_{Z_k} \xrightarrow[n]{w, a.s.} G, \quad G(dx) := \exp\left(-\frac{|x|^2}{2}\right) \frac{dx}{(2\pi)^{d/2}}. \quad (6.1)$$

The above convergence had been established in [LP02], as a by-product of their results concerning the approximation of invariant distributions, under the minimal moment condition  $U_i \in L^2(\mathbb{P})$ , thus weakening the initial assumptions by Brosamler and Schatte (see [Bro88] and [Sch88]). The underlying idea is to use a reformulation of the dynamics of  $(Z_n)_{n \geq 0}$  in terms of a discretization scheme appearing as a perturbation of (1.2). One indeed easily checks that, for  $n \geq 0$ :

$$Z_{n+1} = Z_n - \frac{\gamma_{n+1}}{2} Z_n + \sqrt{\gamma_{n+1}} U_{n+1} + r_n Z_n, \quad r_n := \sqrt{1 - \frac{1}{n+1}} - 1 + \frac{1}{2(n+1)} = O\left(\frac{1}{n^2}\right). \quad (6.2)$$

Thus, the sequence  $(Z_n)_{n \geq 0}$  appears as a perturbed Euler scheme with decreasing step  $\gamma_n = \frac{1}{n}$  of the Ornstein-Uhlenbeck process  $dX_t = -\frac{1}{2}X_t dt + dW_t$  whose invariant distribution is  $G$ . Then the regular Euler scheme

$$X_{n+1} = X_n - \frac{\gamma_{n+1}}{2} X_n + \sqrt{\gamma_{n+1}} U_{n+1}, \quad (6.3)$$

satisfies (1.4) with  $\nu = G$ . The *a.s.* weak convergence (6.1) established in [LP02] follows as a consequence of the (fast enough) convergence of  $Z_n$  towards  $X_n$  as  $n$  goes to infinity.

Moreover, this rate is fast enough to guarantee that the sequence  $\nu_n^Z$  satisfies the conclusion of Theorem 1 point (a) (when  $\gamma_n = \frac{1}{n}$ ,  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow 0$ ), *i.e.* its convergence rate is ruled by a CLT at rate  $\sqrt{\log(n)}$ . In fact this holds under a lower moment assumption  $U_1 \in L^3(\mathbb{P})$ .

Let us mention that the convergence rates related to the *a.s.* CLT had already been investigated by several authors. Let us quote among relevant works, Csorgo and Horváth [CH92], for real valued i.i.d. random variables, Chaâbane and Maâouia [CM00], who investigate the convergence rate of the strong quadratic law of large numbers for some extensions to vector-valued martingales, and Heck [Hec98], for large deviation results. As an application of our previous results, we will derive some new non-asymptotic Gaussian deviation bounds for the *a.s.* CLT, when the involved random variables  $(U_n)_{n \geq 1}$  satisfy **(GC)**. We insist here that the sub-Gaussianity of the innovations is crucial to get a non-asymptotic Gaussian deviation bound. The result readily extends to the wider class of innovations satisfying the general sub-Gaussian exponential deviation inequality (1.5). Also, we slightly weaken the regularity assumptions needed on the function  $f$  in [LP02] for the associated *a.s.* CLT to hold.

### Non-Asymptotic Deviation Bounds.

**Theorem 9** *Assume the innovation sequence  $(U_n)_{n \geq 1}$  satisfies **(GC)** and let  $f$  be a Lipschitz continuous function such that  $G(f) = \int_{\mathbb{R}^d} f(x)G(dx) = 0$ . Then, there exist two explicit monotonic sequences  $c_n \leq 1 \leq C_n$ ,  $n \geq 1$ , with  $\lim_n C_n = \lim_n c_n = 1$  such that for all  $a > 0$  and  $n \geq 1$ :*

$$\mathbb{P}\left[\sqrt{\log(n) + 1}|\nu_n^Z(f)| \geq a\right] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\nabla\varphi\|_\infty^2}\right), \quad (6.4)$$

where  $\varphi$  denotes the solution of the Poisson equation:

$$\forall x \in \mathbb{R}^d, \quad \frac{1}{2}\Delta\varphi(x) - \frac{1}{2}x \cdot \nabla\varphi(x) = f(x), \quad (6.5)$$

which, under the current assumptions, is unique and belongs to  $W_{p,\text{loc}}^2(\mathbb{R}^d, \mathbb{R})$ , for any  $p > 1$ , with  $\|\nabla\varphi\|_\infty \leq 2[f]_1$ .

*Proof.* For  $(Z_n)_{n \geq 0}$  as in (6.2), and  $(X_n)_{n \geq 0}$  as in (6.3) we introduce:

$$\Delta_n := Z_n - X_n.$$

With the definition of  $\nu_n^Z$  in (6.1), write  $\nu_n^Z(f) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(Z_{k-1})$ . We also have similarly  $\nu_n^X(f) := \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k f(X_{k-1})$ . For all  $\lambda > 0$ , we derive similarly to (3.21) (see as well (5.29)) and with the notations of (5.24):

$$\begin{aligned} \mathbb{P}\left[\sqrt{\Gamma_n}|\nu_n^Z(f)| \geq a\right] &= \mathbb{P}\left[\sqrt{\Gamma_n}\left|\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k (f(Z_{k-1}) - f(X_{k-1})) + \nu_n^X(f)\right| \geq a\right] \\ &\leq \mathbb{P}\left[\sqrt{\Gamma_n}\left|\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k (f(Z_{k-1}) - f(X_{k-1})) + \nu_n^X(\mathcal{A}\varphi_\delta) + R_{n,\delta}(f)\right| \geq a\right] \\ &\leq 2 \exp\left(-\frac{\lambda a}{\sqrt{\Gamma_n}}\left(1 - \frac{2\sqrt{\Gamma_n}C_\eta[f]_1\delta}{a}\right)\right) \left(\mathbb{E} \exp\left(\bar{p}[f]_1 \lambda \nu_n^\Delta(|\cdot|)\right)\right)^{\frac{1}{\bar{p}}} \\ &\quad \times \left(\mathbb{E} \exp\left(-\frac{q\bar{q}\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q\bar{q}}} \left(\mathbb{E} \exp\left(\frac{2p\bar{q}\lambda}{\Gamma_n} (|L_n| + |\bar{G}_n|)\right)\right)^{\frac{1}{2p\bar{q}}} \left(\mathbb{E} \exp\left(\frac{4p\bar{q}\lambda}{\Gamma_n} |D_{2,b,n}|\right)\right)^{\frac{1}{4p\bar{q}}} \\ &\quad \times \left(\mathbb{E} \exp\left(\frac{4p\bar{q}\lambda}{\Gamma_n} |D_{2,\Sigma,n}|\right)\right)^{\frac{1}{4p\bar{q}}} \end{aligned} \quad (6.6)$$

for  $\bar{q}, q \in (1, +\infty)$ ,  $\bar{p} = \frac{\bar{q}}{q-1}$ ,  $p = \frac{q}{q-1}$ . Also,  $\varphi_\delta$  corresponds to the solution of the Poisson equation (6.5) obtained replacing  $f$  by its mollified version  $f_\delta$ . Now, we need the following lemma to control  $\nu_n^\Delta(|\cdot|) := \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k |\Delta_{k-1}|$ .

**Lemma 7** *There is a non-negative constant  $C_{6.7}$  such that for all  $\lambda > 0$ :*

$$\mathbb{E} \exp \left( \lambda \nu_n^\Delta(|\cdot|) \right) = \mathbb{E} \exp \left( \frac{\lambda}{\Gamma_n} \sum_{k=1}^n \gamma_k |\Delta_{k-1}| \right) \leq \exp \left( \frac{C_{6.7} \lambda \mathbb{E} [|U_1|] \Gamma_n^{(\frac{3}{2})}}{\Gamma_n} + \frac{C_{6.7}^2 \lambda^2 \Gamma_n^{(3)}}{2\Gamma_n^2} \right). \quad (6.7)$$

For clarity, we postpone the proof to the end of the current section. On the other hand, from Section 5.3 we have that  $\varphi_\delta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  for all  $\beta \in (0, 1)$ . We derive from (6.6), (6.7) similarly to the proof of Theorem 7 by setting  $\bar{\lambda}_n := \frac{\alpha \sqrt{\Gamma_n}}{q\bar{q} \|\nabla \varphi\|_\infty^2}$ :

$$\begin{aligned} & \mathbb{P}[\sqrt{\Gamma_n} |\nu_n^Z(f)| \geq a] \\ & \leq 2 \exp \left( - \frac{a^2}{2q\bar{q} \|\nabla \varphi\|_\infty^2} \left( 1 - \frac{4\sqrt{\Gamma_n} C_\eta [f]_1 \delta}{a} \right) \right) \exp \left( \frac{C_{6.7} \bar{\lambda}_n [f]_1 \mathbb{E} [|U_1|] \Gamma_n^{(\frac{3}{2})}}{\Gamma_n} \right) \\ & \quad \exp \left( \frac{C_{6.7}^2 \bar{p} [f]_1^2 \bar{\lambda}_n^2 \Gamma_n^{(3)}}{2\Gamma_n^2} \right) (I_V^1)^{\frac{1}{p\bar{q}}} \exp \left( \frac{1}{p\bar{q}} \left( c_V + \frac{C_{3.27}^\delta}{2} \right) + \frac{(pa_n^\delta)^2}{2\bar{q}} \right) \\ & \quad \times \exp \left( \bar{\lambda}_n^2 \left( p\bar{q} \left( \frac{3C_{V,\varphi}^2}{c_V \Gamma_n^2} + [C_{3.26}^\delta + \frac{3}{2} C_{3.27}^\delta] \frac{(\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) + \frac{1}{2p\bar{q}} \right) \right) \\ & \leq 2(I_V^1)^{\frac{1}{p\bar{q}}} \exp \left( \frac{1}{p\bar{q}} \left( c_V + \frac{C_{3.27}^\delta}{2} \right) + \frac{p(a_n^\delta)^2}{2\bar{q}} + \varepsilon_n^{-1} \Gamma_n \delta^2 \right) \\ & \quad \times \exp \left( - \frac{a^2}{2q\bar{q} \|\nabla \varphi\|_\infty^2} \left( 1 - d_n^\delta - \frac{\bar{p}}{q\bar{q} \|\nabla \varphi\|_\infty^2} \frac{[f]_1^2 C_{6.7}^2 \left( \Gamma_n^{(3)} + \mathbb{E} [|U_1|]^2 (\Gamma_n^{(\frac{3}{2})})^2 \right)}{\Gamma_n} \right) \right), \end{aligned}$$

for  $\varepsilon_n > 0$  and  $d_n^\delta$  as in (5.30). Choose again  $(p_n)_{n \geq 1}$  and  $\delta$  as in Section 5.3 so that  $q_n \xrightarrow[n]{n} 1$ ,  $d_n^\delta \xrightarrow[n]{n} 0$  with the indicated monotonicity for  $n$  large enough. We can now take  $\bar{p} := \bar{p}_n \xrightarrow[n]{n} +\infty$  such that  $\frac{\bar{p}}{\Gamma_n} \xrightarrow[n]{n} 0$ . The above inequality then gives the result up to a direct modification of the sequences  $(C_n)_{n \geq 1}$ ,  $(c_n)_{n \geq 1}$ .  $\square$

### Proof of Lemma 7

The definition of  $\Delta_n$  implies:

$$\Delta_{n+1} = \Delta_n \left( 1 - \frac{\gamma_{n+1}}{2} \right) + r_n Z_n,$$

where we recall from (6.2) that  $r_n := \sqrt{1 - \frac{1}{n+1}} - 1 + \frac{1}{2(n+1)} = O(\frac{1}{n^2})$ . In particular, there exists  $\tilde{C}_1 > 0$  such that for all  $n \geq 1$ ,

$$|r_n| \leq \frac{\tilde{C}_1}{n^2}. \quad (6.8)$$

Setting now  $\rho_0 = 1$  and for  $n \geq 1$ :

$$\rho_n := \left[ \prod_{k=1}^n \left(1 - \frac{\gamma_k}{2}\right) \right]^{-1} = \prod_{k=1}^n \frac{2k}{2k-1},$$

a direct induction on  $\Delta_n$  yields:

$$\Delta_n = \frac{1}{\rho_n} \sum_{k=1}^n r_k \rho_k Z_k = \frac{1}{\rho_n} \sum_{k=1}^n r_k \rho_k \left( \sum_{l=1}^k \frac{U_l}{\sqrt{k}} \right) = \frac{1}{\rho_n} \sum_{l=1}^n \left( \sum_{k=l}^n \frac{r_k \rho_k}{\sqrt{k}} \right) U_l. \quad (6.9)$$

Also, from the Wallis formula  $\rho_n \sim_n \sqrt{\pi n}$ , which implies that there exists  $\bar{C}_2 \geq 1$  such that for all  $n \geq 1$ :

$$\bar{C}_2^{-1} \sqrt{n} \leq \rho_n \leq \bar{C}_2 \sqrt{n}. \quad (6.10)$$

We now get from (6.9) and the Fubini theorem:

$$\begin{aligned} \Gamma_n \nu_n^\Delta(|\cdot|) &= \sum_{k=1}^n \gamma_k |\Delta_{k-1}| \leq \sum_{k=1}^n \frac{\gamma_k}{\rho_{k-1}} \sum_{l=1}^{k-1} \left( \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \right) |U_l| \\ &= \sum_{l=1}^{n-1} \left[ \sum_{k=l+1}^n \frac{\gamma_k}{\rho_{k-1}} \left( \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \right) \right] |U_l|. \end{aligned} \quad (6.11)$$

Combining (6.8) and (6.10), we get that there exist constants  $\bar{C}_3, \bar{C}_4 > 0$  such that for all  $k \in \llbracket l+1, n \rrbracket$ .

$$\frac{\gamma_k}{\rho_{k-1}} \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \leq \frac{\bar{C}_3}{k^{3/2} l}, \quad \sum_{k=l+1}^n \frac{\gamma_k}{\rho_{k-1}} \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \leq \frac{\bar{C}_4}{l^{3/2}}. \quad (6.12)$$

Plugging this inequality in (6.11), we derive:

$$\nu_n^\Delta(|\cdot|) \leq \frac{1}{\Gamma_n} \sum_{l=1}^{n-1} \left[ \sum_{k=l+1}^n \frac{\gamma_k}{\rho_{k-1}} \sum_{m=l}^{k-1} \frac{|r_m| \rho_m}{\sqrt{m}} \right] |U_l| \leq \frac{\bar{C}_4}{\Gamma_n} \sum_{l=1}^{n-1} \frac{|U_l|}{l^{3/2}}. \quad (6.13)$$

For any  $\lambda > 0$ , Equation (6.13) and the Gaussian concentration property **(GC)** of the innovation entail:

$$\begin{aligned} \mathbb{E} \exp \left( \lambda \nu_n^\Delta(|\cdot|) \right) &\leq \prod_{k=1}^{n-1} \mathbb{E} \exp \left( \frac{C_4 \lambda}{\Gamma_n k^{3/2}} |U_k| \right) \leq \prod_{k=1}^{n-1} \exp \left( \frac{\bar{C}_4 \lambda}{\Gamma_n k^{3/2}} \mathbb{E} [|U_1|] + \frac{1}{2} \left( \frac{\bar{C}_4 \lambda}{\Gamma_n k^{3/2}} \right)^2 \right) \\ &= \exp \left( \frac{\bar{C}_4 \lambda \mathbb{E} [|U_1|] \Gamma_n^{(3/2)}}{\Gamma_n} + \frac{\bar{C}_4^2 \lambda^2 \Gamma_n^{(3)}}{2 \Gamma_n^2} \right). \end{aligned}$$

This completes the proof.  $\square$

## 6.2 Numerical Results

We present in this section numerical results associated with the computation of the empirical measure  $\nu_n$  illustrating our previous theorems.



### Sub-Gaussian tails

We first consider  $d = r = 1$ . Also, for simplicity, the innovations  $(U_i)_{i \geq 1}$  and  $X_0$  are Bernoulli variables with  $\mathbb{P}(U_1 = -1) = \mathbb{P}(U_1 = 1) = \frac{1}{2}$ . We illustrate here Theorem 2 taking  $b(x) = -\frac{x}{2}$ , and  $\sigma(x) = \cos(x)$  in (1.1). This is a (weakly) hypoelliptic example. Indeed, setting for  $x \in \mathbb{R}$ ,  $X_1(x) = \cos(x)\partial_x$  and  $X_0(x) = -\frac{x}{2}\partial_x$ , we have  $\text{span}\{X_1, [X_1, X_0]\} = \mathbb{R}$ . We choose as well to compute  $\nu_n(\mathcal{A}\varphi)$  for  $\varphi(x) = x + \varepsilon \cos(x)$  for  $\varepsilon = 0.01$ , and  $\varphi(x) = \cos(x)$ . The function  $\varphi$  is here given. The assumptions of Theorem 2 follow from Theorem 18 in Rotschild and Stein [RS76] (up to the introduction of a suitable partition of unity). From Theorem 2, for steps of the form  $(\gamma_k)_{k \geq 1} = (k^{-\theta})_{k \geq 1}$ ,  $\theta \in [1/3, 1]$  (corresponding to  $\beta = 1$  in Theorem 2), the function  $a \in \mathbb{R}^+ \mapsto g_{n,\theta}(a) :=$

$$\begin{cases} \log(\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a]), & \theta \in (1/3, 1], \\ \log(\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + (\mathcal{B}_{n,1} - E_n^1)| \geq a]), & \theta = 1/3, \end{cases}$$

is such that for  $a > a_n := a_n(\theta)$

where for  $\theta \in (1/3, 1]$ ,  $a_n(\theta) = 0$  and for  $\theta = 1/3$ ,  $a_n(\theta) = \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)} \frac{\Gamma_n^{(3+\frac{\beta}{2})}}{\sqrt{\Gamma_n}}$ :

$$g_{n,\theta}(a) \leq -c_n \frac{(a - a_n)^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2} + \log(2C_n).$$

We plot in Figure 3.1 the curves of  $g_{n,\theta}$  for  $\theta$  varying as  $\theta_j = \frac{1}{3} + (1 - \frac{1}{3})\frac{j}{5}$ , for  $j \in \llbracket 1, 5 \rrbracket$ ,  $\varphi(x) = x + \varepsilon \cos(x)$  and in Figure 3.2 the curve of  $g_{n,\theta}$  for  $\theta = \theta_0 = \frac{1}{3}$  and  $\varphi(x) = \cos(x)$ . The simulations have been performed for  $n = 5 \times 10^4$  in Figure 3.1,  $n = 5 \times 10^6$  in Figure 3.2, and the probability estimated by Monte Carlo simulation for  $MC = 10^4$  realizations of the random variable  $|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)|$  in the unbiased case and in the biased case of the random variable  $|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) + (\mathcal{B}_{n,1} - E_n^1)^M|$ , where  $(\mathcal{B}_{n,1} - E_n^1)^M$  is obtained from  $\mathcal{B}_{n,1} - E_n^1$  replacing the integral over  $[0,1]$ , that needs to be evaluated at every time step, by a quantization of the uniform law on  $[0,1]$  with  $M = 10$  points. We refer to [GL00] or [Pag97] for details on quantization. We point out that this is one drawback that appears to obtain the fastest convergence rate, the bias needs to be estimated and therefore the function  $\varphi$  in some sense known (since the approximation of the bias requires to compute its derivatives). The corresponding 95% confidence intervals have size at most of order 0.016. To compare with, we also introduce the functions  $S_{n,\theta}(a) := -\frac{(a - a_n(\theta))^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}$ ,  $S_{n,\theta,c}(a) := -\frac{(a - a_n(\theta))^2}{2\nu_{n_c}(\sigma^2) \|\nabla\varphi\|_\infty^2}$ ,  $S_{n,\theta,A}(a) := -\frac{(a - a_n(\theta))^2}{2\nu_{n_c}(|\sigma\nabla\varphi|^2)}$  and the optimal concentration  $P(\lambda_{\min})(n, \theta, a, \rho)$ , obtained in Remark 11, optimizing numerically in  $\rho$ . The quantities  $\nu_{n_c}(\sigma^2), \nu_{n_c}(|\sigma\nabla\varphi|^2)$  in the previous expressions actually correspond to the numerical estimation, for  $n_c = 10^4$  and  $(\gamma_k^c)_{k \geq 1} = (k^{-\theta^c})_{k \geq 1}$  with  $\theta^c = \frac{1}{3} + 10^{-3}$ , of  $\nu(\sigma^2), \nu(|\sigma\nabla\varphi|^2)$  appearing respectively in the sharper concentration bound of Theorem 8 when  $\sigma^2 - \nu(\sigma^2)$  is a coboundary and in the asymptotic Theorem 1. In the unbiased case of Figure 3.1, we plot the maximum in  $j$  of the  $(S_{n,\theta_j})_{j \in \llbracket 1, 5 \rrbracket}, (S_{n,\theta_j,c})_{j \in \llbracket 1, 5 \rrbracket}, (S_{n,\theta_j,A})_{j \in \llbracket 1, 5 \rrbracket}, (P(\lambda_{\min})(n, \theta_j, a, \rho))_{j \in \llbracket 1, 5 \rrbracket}$  corresponding to  $j = 1$ . The associated curves are denoted by  $S_n, S_{n,c}, S_{n,A}$  and  $P(\lambda_{\min})(n)$ .

The Figures 3.1 and 3.2 correspond to the unbiased and biased cases respectively. In the unbiased case, we observe that the curves almost overlay, the optimal deviation rate  $P(\lambda_{\min})$  is very close to the empirical data. It is also below the numerical estimation of the asymptotic threshold given by  $S_{n,\theta,A}$  which is, for our considered example, almost indistinguishable from the coboundary  $S_{n,\theta,c}$  (indeed, since  $\varepsilon = 0.01$ ,  $\|\nabla\varphi\|_\infty^2 \leq 1 + \varepsilon^2$

and  $\nu(\sigma^2)\|\nabla\varphi\|_\infty^2 \simeq \nu(|\sigma\nabla\varphi|^2)$  and far below from the bounds of  $S_{n,\theta}$ . In the biased case,  $P(\lambda_{\min})$  stays very close to the theoretical asymptotic bound given by  $S_{n,\theta,A}$  up to a certain deviation level  $a$ , namely for  $a \in [0, 0.5]$ . It then remains the best bound provided by our results. In this example, the improvement associated with  $S_{n,\theta,c}$  is also notable. It is precisely because the source term has a more oscillating gradient that we have also considered a larger running time, corresponding to  $n = 10^6$ , for the empirical curves. For this choice, we see relatively good agreement w.r.t. to the asymptotic deviation bounds of  $S_{n,\theta_0,A}$ .

The figures below thus illustrate that the explicit *optimal rate* of Remark 11 seems rather appropriate to capture the deviations of the empirical random measures.

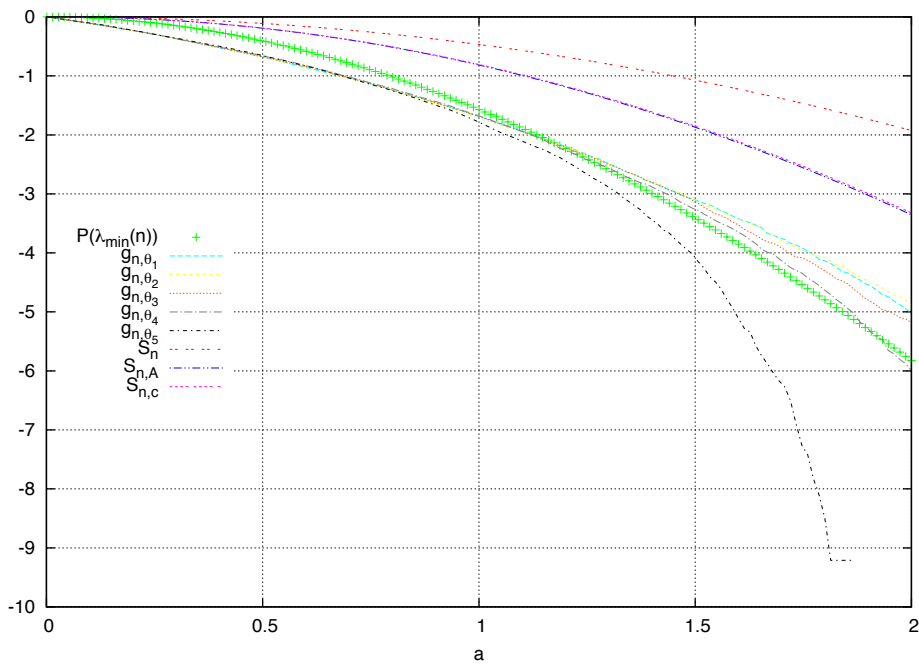


Figure 3.1 – Unbiased Case. Plot of  $a \mapsto g_{n,\theta}(a)$ , for  $(\theta_k)_{k \in \llbracket 1,5 \rrbracket}$ , with  $\varphi(x) = \sigma(x) = x + \varepsilon \cos(x)$ ,  $\varepsilon = 0.01$ .

We eventually plot below the deviation curves with source  $\varphi(x) = \cos(x)$  adding a last curve obtained replacing in the formula for  $P(\lambda_{\min})$  of Remark 11 the  $\|\nabla\varphi\|_\infty^2 \nu(\sigma^2)$  by  $\nu(|\sigma\nabla\varphi|^2)$ . For practical purposes, this last quantity is again estimated numerically with the same previous parameters. Even if the analysis of Theorem 8 cannot be extended to justify such a choice, the empirical evidence is rather striking.

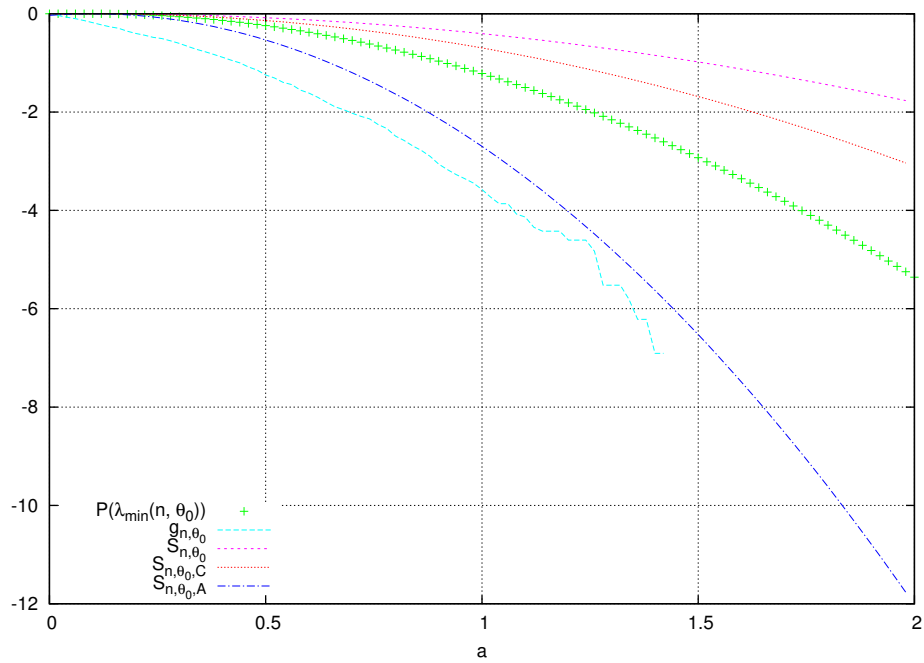


Figure 3.2 – Biased Case. Plot of  $a \mapsto g_{n,\theta}(a)$ , for  $\theta_0 = \frac{1}{3}$ , with  $\varphi(x) = \sigma(x) = \cos(x)$ .

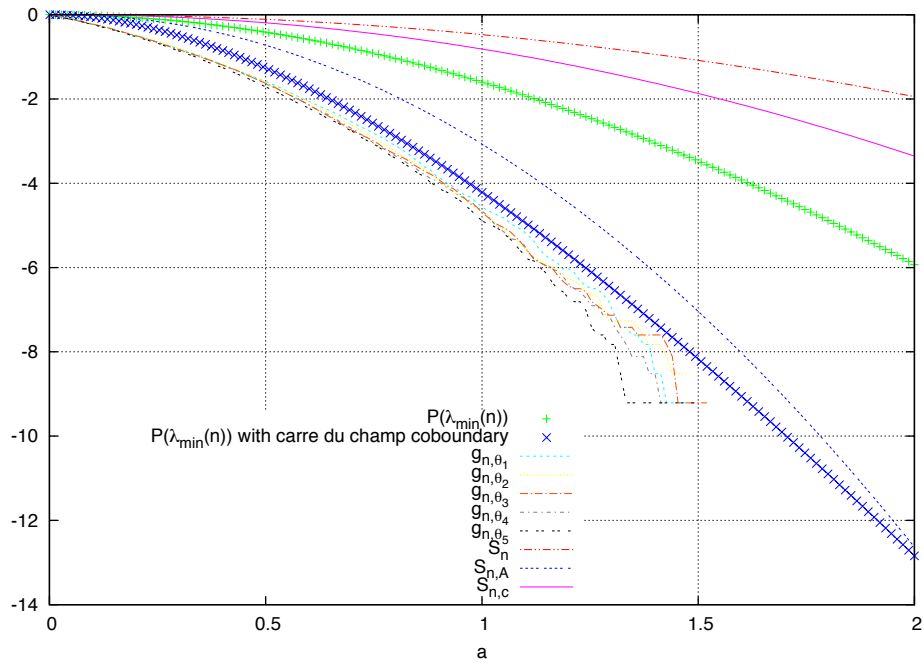


Figure 3.3 – Plot of  $a \mapsto g_{n,\theta}(a)$ , for  $(\theta_k)_{k \in \llbracket 1,5 \rrbracket}$ , with  $\varphi(x) = \sigma(x) = \cos(x)$ .

### Slutsky like result

In this paragraph, we illustrate our results from Theorem 6, which can be viewed as an extension of the usual Slutsky's Lemma to our current framework, for a multidimensional process, precisely for  $r = d = 2$  in the case  $\beta \in (0, 1)$ . In order to converge as fast as possible without bias, we take  $\theta = \frac{1}{2+\beta} + \frac{1}{1000}$ . We also choose a model which satisfies the assumptions of Theorem 4 under  $(\mathbf{C}_{\mathbf{UE}})$  and Lemma 6. We consider:

$$f(x) = \frac{|x|^{1+\beta}}{1 + |x|^\beta}, \quad b(x) = \begin{pmatrix} -4x_1 + 6x_2 \\ -5x_1 - 5x_2 \end{pmatrix}, \quad \sigma\sigma^*(x) = \begin{pmatrix} \frac{\cos(x_1+x_2)}{2} + 1 & \frac{\sin(x_1)\sin(x_2)}{4} \\ \frac{\sin(x_1)\sin(x_2)}{4} & 1 - \frac{\sin(x_2)}{2} \end{pmatrix}.$$

Remark that the non-degeneracy condition  $(\mathbf{UE})$  is fulfilled by  $\Sigma = \sigma\sigma^*$ , as well as the condition set in Theorem 4 under  $(\mathbf{C}_{\mathbf{UE}})$ ,  $\Sigma_{i,j}(x) = \Sigma_{i,j}(x_i, \dots, x_d)$ , for all  $1 \leq i \leq j \leq d$ . Furthermore, from the Cholesky decomposition, we write:

$$\sigma(x) = \begin{pmatrix} \sqrt{\frac{\cos(x_1+x_2)}{2} + 1} & 0 \\ \frac{\sin(x_1)\sin(x_2)}{4\sqrt{\frac{\cos(x_1+x_2)}{2} + 1}} & \sqrt{-\frac{\sin(x_1)^2\sin(x_2)^2}{16(\frac{\cos(x_1+x_2)}{2} + 1)} + 1 - \frac{\sin(x_2)}{2}} \end{pmatrix}.$$

Let us check that  $(\mathbf{D}_\alpha^p)$  is satisfied. Firstly, remark that  $\frac{Db+Db^*}{2}$  is a constant matrix whose eigenvalues are  $\{-\frac{\sqrt{2+9}}{2}, \frac{\sqrt{2-9}}{2}\}$ . Direct computations yield that, for all  $x \in \mathbb{R}^d$ ,  $\xi \in \mathbb{R}^d$ :

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r |D\sigma_{\cdot j}(x)\xi|^2 \leq -3.085|\xi|^2.$$

It can be checked similarly that the condition  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$  is satisfied for  $\alpha = 3.085$  and  $\beta = .5$  which we consider below. Also, the condition  $(\mathbf{R}_{1,\beta})$  clearly holds. In other words, all assumptions of Theorem 6 are in force. We set for the following plot:

$$g_n^\sigma(a) = \log \mathbb{P}[\sqrt{\Gamma_n} |\nu_n f| \geq a], \quad S_\sigma(a) = -\frac{a^2 \alpha^2}{2[f]_1^2},$$

with  $\alpha = 3.085$ , and  $[f]_1 = 1$ .

Unlike in the previous simulations, we do not know here the value of  $\nu(f)$ . In fact, in paragraph 6.2 we had chosen to compute the deviation of  $\mathcal{A}\varphi$  from  $0 = \nu(\mathcal{A}\varphi)$ . Here, we estimate from the ergodic theorem  $\nu(f)$ , taking  $\beta = .5$ , by  $\nu_{n^c}(f) \approx 0.71308$  for  $n^c = 5 \cdot 10^5$ . Running  $MC = 10^2$  samples, we find that the size of the associated 95% confidence interval is  $3.208 \cdot 10^{-4}$ . Finally, the simulations are performed for  $n = 5 \times 10^4$ , and the probability is calculated by Monte Carlo algorithm for  $MC = 10^3$  realizations. The maximum size of the associated 95% confidence interval is  $4.75054 \cdot 10^{-5}$ . The innovations are Gaussian random variables.

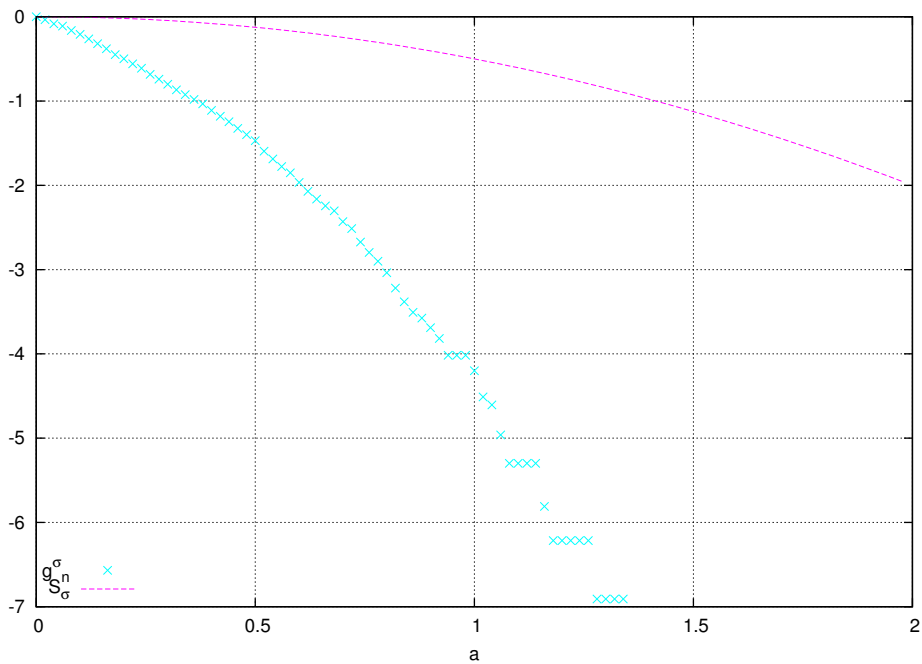


Figure 3.4 – Plot of  $a \mapsto g_n(a)$  with  $f(x) = \frac{|x|^\beta}{1+|x|^\beta}$ ,  $\beta = .5$ .

In Figure 3.4, we observe that the curve  $S_\sigma$  stays above  $g_n^\sigma$  as proved in Theorem 6. However, remark that the graphs are quite spaced. This can be explained, among other things, by the difference between  $\nu(\|\sigma\|^2) \frac{\lfloor f \rfloor_1}{\alpha}$  and the asymptotic variance  $\nu(|\sigma^* \nabla \varphi|^2)$ . Furthermore we have represented  $S_\sigma$  which is a kind of asymptotic version of  $P(\lambda_{\min}(n))$  in the previous plots.

# Chapter 4

## A sharp non-asymptotic concentration

**Abstract :** For an ergodic Brownian diffusion with invariant distribution  $\nu$ , we consider a sequence of empirical distributions  $(\nu_n)_{n \geq 1}$  associated with an approximation scheme with decreasing time step  $(\gamma_n)_{n \geq 1}$  along an adapted regular enough class of test functions  $f$  such that  $f - \nu(f)$  is a coboundary of the infinitesimal generator  $\mathcal{A}$ . Denote by  $\sigma$  the diffusion coefficient and  $\varphi$  the solution of the Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$ . When the square norm  $|\sigma^* \nabla \varphi|^2$  lies in the same coboundary class as  $f$ , we establish sharp non-asymptotic concentration bounds for suitable normalizations of  $\nu_n(f) - \nu(f)$ . Our bounds are optimal in the sense that they match the asymptotic limit obtained by Lamberton and Pagès in [LP02], for a certain large deviation regime. In particular, this allows us to derive sharp non-asymptotic confidence intervals. Eventually, we are able to handle, up to an additional constraint on the time steps, Lipschitz sources  $f$  in an appropriate non-degenerate setting.

### 1 Introduction

#### 1.1 Statement of the problem

Consider the stochastic differential equation

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t, \quad (1.1)$$

where  $(W_t)_{t \geq 0}$  stands for a Wiener process of dimension  $r \in \mathbb{N}$  on a given filtered probability space  $(\Omega, \mathbf{G}, (\mathbf{G}_t)_{t \geq 0}, \mathbb{P})$ ,  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , and  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  are Lipschitz continuous functions and satisfy a Lyapunov condition (see further Assumption  $\mathcal{L}_\nu$ ) which provides the existence of an invariant distribution  $\nu$ . Throughout the chapter, uniqueness of the invariant distribution  $\nu$  is assumed. The purpose of this work is to estimate the invariant distribution of the diffusion equation (1.1).

In order to make a clear parallel with the objects we will introduce for the approximation of  $\nu$ , let us first recall some basic facts on  $(Y_t)_{t \geq 0}$  and  $\nu$ .

Introduce for a bounded continuous function  $f$  and  $t \in \mathbb{R}_+$  the average occupation measure:

$$\nu_t(f) := \frac{1}{t} \int_0^t f(Y_s)ds. \quad (1.2)$$

Foremost, bear in mind the *usual* ergodic theorem which holds under appropriate Lyapunov conditions (see e.g. [KM11]):

$$\nu_t(f) \xrightarrow[t \rightarrow +\infty]{a.s.} \nu(f) := \int f d\nu. \quad (1.3)$$

Under suitable stability and regularity conditions, Bhattacharya [Bha82] then established a corresponding Central Limit Theorem (CLT). Namely, for all smooth enough function  $f$ ,

$$\sqrt{t}(\nu_t(f) - \nu(f)) \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi(x)|^2 \nu(dx)\right), \quad (1.4)$$

where  $\varphi$  is the solution of the Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$  and  $\mathcal{A}$  stands for the infinitesimal operator of the diffusion (1.1) (see (2.3) below for more details). In the following, we say that  $f$  is **coboundary** when there is a smooth solution  $\varphi$  to the Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$ .

Identity (1.4) is a Central Limit Theorem (CLT) whose asymptotic variance is the integral of the well known **carré du champ** (called energy but we will say, from now on, by abuse of terminology, *carré du champ*), for more precision, see [BGL14] and [Led99]. The *carré du champ* is actually a bilinear operator defined for any smooth functions  $\varphi, \psi$  by  $\Gamma(\varphi, \psi) := \mathcal{A}(\varphi \cdot \psi) - \mathcal{A}\varphi \cdot \psi - \varphi \cdot \mathcal{A}\psi$ , and so:

$$\nu(\Gamma(\varphi, \varphi)) = \int_{\mathbb{R}^d} \Gamma(\varphi, \varphi) \nu(dx) = -2 \int_{\mathbb{R}^d} \mathcal{A}\varphi \cdot \varphi \nu(dx) = \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 \nu(dx).$$

Indeed, observe that  $\nu(\mathcal{A}(\varphi \cdot \psi)) = 0$ . This is a consequence of the fact that  $\nu$  solves in the distributional sense the Fokker-Planck equation  $\mathcal{A}^* \nu = 0$ . Also, this observation yields that, in order to bypass solving a Poisson equation, a common trick consists in dealing with smooth functions of the form  $\mathcal{A}\varphi$ .

From a practical point of view, several questions appear: how to approach the process  $(Y_t)_{t \geq 0}$ , the integral  $\nu_t$ , and the deviation from the asymptotic measure appearing in (1.4)? Here, the first question is addressed by considering a suitable discretization scheme with decreasing time steps,  $(\gamma_k)_{k \geq 1}$ . The integral  $\nu_t$  can then be approximated by the associated empirical measure, whose deviations will be controlled in our main results. In particular, we take advantage of the discrete analogue to (1.4) established by Lambertson and Pagès in [LP02] for the current approximation scheme, to derive sharp non-asymptotic bounds for the empirical measure.

We propose an approximation algorithm based on an Euler like discretization with decreasing time step, first introduced by Lambertson and Pagès in [LP02] who derived related asymptotic limit theorem in the spirit of (1.4), and exploited as well in Chapter 3 where some corresponding non-asymptotic bounds are obtained.

For the decreasing step sequence  $(\gamma_k)_{k \geq 1}$  and  $n \geq 0$ , the scheme deriving from (1.1) is defined by:

$$\begin{cases} X_{n+1} = X_n + \gamma_{n+1} b(X_n) + \sqrt{\gamma_{n+1}} \sigma(X_n) U_{n+1}, \\ X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}), \end{cases} \quad (1.5)$$

where  $(U_n)_{n \geq 1}$  is an i.i.d. sequence of random variables on  $\mathbb{R}^r$ , independent of  $X_0$ , and whose moments match with the Gaussian ones up to order 3. In particular, more general innovations than the Brownian increments can be used.

Intuitively, the decreasing steps in (1.5) allow to be more and more precise when time grows.

The empirical (random) occupation measure of the scheme is defined for all  $A \in \mathcal{B}(\mathbb{R}^d)$  (where  $\mathcal{B}(\mathbb{R}^d)$  denotes the Borel  $\sigma$ -field on  $\mathbb{R}^d$ ) by:

$$\nu_n(A) := \nu_n(\omega, A) := \frac{\sum_{k=1}^n \gamma_k \delta_{X_{k-1}(\omega)}(A)}{\sum_{k=1}^n \gamma_k}. \quad (1.6)$$

We are interested in the long time approximation, so we need to consider steps  $(\gamma_k)_{k \geq 1}$  such that  $\Gamma_n := \sum_{k=1}^n \gamma_k \xrightarrow[n]{} +\infty$ .

Under suitable Lyapunov like assumptions, Lamberton and Pagès in [LP02] first proved the following ergodic result: for any  $\nu$ -a.s. continuous function  $f$  with polynomial growth,  $\nu_n(f) \xrightarrow[n]{a.s.} \nu(f) = \int_{\mathbb{R}^d} f(x) \nu(dx)$ , which is the discrete analogue of (1.3).

The main benefit of decreasing steps instead of constant ones is thus that the empirical measure directly converges towards the invariant one. Otherwise, taking  $\gamma_k = h > 0$  in (1.5), the previous ergodic theorem must be changed into:  $\nu_n(f) \xrightarrow[n]{a.s.} \nu^h(f) = \int_{\mathbb{R}^d} f(x) \nu^h(dx)$ , where  $\nu^h$  is the invariant distribution of the scheme. So, an extra study must be carried out, namely the difference  $\nu - \nu^h$  should be estimated. For more details about this approach we refer to [TT90], [Tal02] and the work of Malrieu and Talay [MT06]. This work first addressed the issue of deriving non-asymptotic controls for the deviations of empirical measure of type (1.6) when  $\gamma_k = \gamma > 0$  (constant step). The backbone of their approach consisted in establishing a Log Sobolev inequality, which implies Gaussian concentration, for the Euler scheme. In whole generality, functional inequalities (such as the Log Sobolev one) are a powerful tools to get simple controls on the invariant distribution associated with the diffusion process (1.1), see e.g. Ledoux [Led99] or Bakry *et al.* [BGL14]. Withal Log Sobolev, and Poincaré inequalities turn out to be quite *rigid* in the framework of discretization schemes like (1.5) with or without decreasing steps.

For the CLT associated with stationary Markov chains, we refer to Gordin's Theorem (see [GL78]). Note as well that the variance of the limit Gaussian law is also the *carré du champ* for discrete Poisson equation associated with the generator of the chain.

Let us mention as well some related works. In [BB06], Blower and Bolley establish Gaussian concentration properties for deviations of functional of the path in the case of metric space valued homogeneous Markov chains. Non-asymptotic deviation bounds for the Wasserstein distance between the marginal distributions and the stationary law, in the homogeneous case can be found in [Boi11] (see also Boissard and Le Gouic in [BLG14] for controls on the expectations of this Wasserstein distance). The key point of these works is to demonstrate contraction properties of the transition kernel of the homogeneous Markov chain for a Wasserstein metric, which requires some continuity in this metric for the transition law involved, see e.g. [BB06].

In the current work, we aim to establish an optimal non-asymptotic concentration inequality for  $\nu_n(f) - \nu(f)$ . When  $|\sigma^* \nabla \varphi|^2 - \nu(|\sigma^* \nabla \varphi|^2)$  is a coboundary, we manage to improve the estimates in Chapter 3. Insofar, we better the variance in the upper-bound obtained in Theorem 3 therein: when the time step is such that  $\gamma_n \asymp n^{-\theta}$ , for  $\theta > \frac{1}{3}$ , for all  $n \in \mathbb{N}$ , for a smooth enough function  $\varphi$  s.t.  $\mathcal{A}\varphi = f - \nu(f)$ , under suitable assumptions (further called Assumptions **(A)**), and if  $\|\sigma\|^2$  is coboundary then there exist explicit non-



negative sequences  $(\tilde{c}_n)_{n \geq 1}$  and  $(\tilde{C}_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$  s.t. for all  $n \geq 1$  and  $0 < a = o(\sqrt{\Gamma_n})$ :

$$\mathbb{P}[\sqrt{\Gamma_n}|\nu_n(f) - \nu(f)| \geq a] = \mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2\tilde{C}_n \exp\left(-\frac{\tilde{c}_n a^2}{2\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2}\right).$$

In fact, we get below the optimal variance bound, namely the *carré du champ*  $\nu(|\sigma^*\nabla\varphi|^2)$ , instead of the expression  $\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2$  as in the previous inequality. Up to the same previously indicated deviation threshold,  $a = o(\sqrt{\Gamma_n})$ , we derive the optimal Gaussian concentration. Consequently, we are able to derive directly some sharp non-asymptotic confidence intervals.

To establish our non-asymptotic results, we use martingale increment techniques which turn out to be very robust in a rather large range of application fields. Let us for instance mention the work of Frikha and Menozzi [FM12] which establishes non-asymptotic bounds for the regular Monte Carlo error associated with the Euler discretization of a diffusion until a finite time interval  $[0, T]$  and for a class of stochastic algorithms of Robbins-Monro type. Still with martingale approach, Dedecker and Gouëzel [DG15] have obtained non-asymptotic deviation bounds for separately bounded functionals of geometrically ergodic Markov chains on a general state space. Eventually, we can refer again to the work in Chapter 3 in the current setting.

The chapter is organized as follows. In Section 2, we state our notations and assumptions as well as some known and useful results related to our approximation scheme. Section 3 is devoted to our main concentration results (for a certain deviation regime that we will call *Gaussian deviations*), we also state therein several technical lemmas whose proofs are postponed to Section 4. Importantly, we also provide a user's guide to the proof which emphasizes the key steps in our approach. Section 4 is the technical core of the chapter. We then discuss in Section 5 some regularity issues for the considered test functions. Namely, we recall some assumptions introduced in Chapter 3 which yield appropriate regularity concerning the solution of Poisson equation. We also extend there our main results to test functions  $f$  that are Lipschitz continuous, up to some constraints on the step sequence.

We proceed in Section 6 to the explicit optimization of the constants appearing in the concentration bound deriving from our approach. Intrinsically, this procedure conducts to two deviation regimes, the *Gaussian* one up to  $a = o(\sqrt{\Gamma_n})$  and the *super Gaussian* one for  $a \gg \sqrt{\Gamma_n}$  which deteriorates the concentration rate. Even though awkward at first sight (see e.g. Remark 4 below), this refinement turns out to be useful for some numerical purposes, as it emphasized in Section 7. We conclude there with some numerical results associated with a degenerate diffusion. Some additional technical details needed in Section 6 are gathered in Appendix 8.

## 2 Assumptions and Existing Results

### 2.1 General notations

For all step sequence  $(\gamma_n)_{n \geq 1}$ , we denote:

$$\forall \ell \in \mathbb{R}, \Gamma_n^{(\ell)} := \sum_{k=1}^n \gamma_k^\ell, \Gamma_n := \sum_{k=1}^n \gamma_k = \Gamma_n^{(1)}.$$

Practically, the time step sequence is assumed to have the form:  $\gamma_n \asymp \frac{1}{n^\theta}$  with  $\theta \in (0, 1]$ , where for two sequences  $(u_n)_{n \in \mathbb{N}}, (v_n)_{n \in \mathbb{N}}$  the notation  $u_n \asymp v_n$  means that  $\exists n_0 \in \mathbb{N}, \exists C \geq 1$  s.t.  $\forall n \geq n_0, C^{-1}v_n \leq u_n \leq Cv_n$ .

We will denote by  $C$  a non negative constant, and by  $(e_n)_{n \geq 1}, (\mathcal{R}_n)_{n \geq 1}$  deterministic generic sequences s.t.  $e_n \xrightarrow[n]{} 0$  and  $\mathcal{R}_n \xrightarrow[n]{} 1$ , that may change from line to line. The constant  $C$  depends, uniformly in time, as well as the sequences  $(e_n)_{n \geq 1}, (\mathcal{R}_n)_{n \geq 1}$ , on known parameters appearing in the assumptions introduced in Section 2.2 (called **(A)** throughout the document). Other possible dependencies will be explicitly specified.

In the following, for any smooth enough function  $f$ , for  $k \in \mathbb{N}$  we will denote  $D^k f$  the tensor of the  $k^{\text{th}}$  derivatives of  $f$ . Namely  $D^k f = (\partial_{i_1} \dots \partial_{i_k} f)_{1 \leq i_1, \dots, i_k \leq d}$ . However, for a multi-index  $\alpha \in \mathbb{N}_0^d := (\mathbb{N} \cup \{0\})^d$ , we set  $D^\alpha f = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} f : \mathbb{R}^d \rightarrow \mathbb{R}$ .

For a  $\beta$ -Hölder continuous function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , we introduce the notation

$$[f]_\beta := \sup_{x \neq x'} \frac{|f(x) - f(x')|}{|x - x'|^\beta} < +\infty,$$

for its Hölder modulus of continuity. Here,  $|x - x'|$  stands for the Euclidean norm of  $x - x' \in \mathbb{R}^d$ .

We denote, for  $(p, m) \in \mathbb{N}^2$ , by  $\mathcal{C}^p(\mathbb{R}^d, \mathbb{R}^m)$  the space of  $p$ -times continuously differentiable functions from  $\mathbb{R}^d$  to  $\mathbb{R}^m$ . Besides, for  $f \in \mathcal{C}^p(\mathbb{R}^d, \mathbb{R}^m)$ ,  $p \in \mathbb{N}$ , we define for  $\beta \in (0, 1]$  the Hölder modulus:

$$[f^{(p)}]_\beta := \sup_{x \neq x', |\alpha|=p} \frac{|D^\alpha f(x) - D^\alpha f(x')|}{|x - x'|^\beta} \leq +\infty,$$

where  $\alpha$  (viewed as an element of  $\mathbb{N}^d$ ) is a multi-index of length  $p$ , i.e.  $|\alpha| := \sum_{i=1}^d \alpha_i = p$ . Hence, in the above definition, the  $|\cdot|$  in the numerator is the usual absolute value. We will as well use the notation  $\llbracket n, p \rrbracket$ ,  $(n, p) \in (\mathbb{N}_0)^2, n \leq p$ , for the set of integers being between  $n$  and  $p$ .

From now on, we introduce for  $k \in \mathbb{N}_0, \beta \in (0, 1]$  and  $m \in \{1, d, d \times r\}$  the Hölder spaces

$$\begin{aligned} \mathcal{C}^{k, \beta}(\mathbb{R}^d, \mathbb{R}^m) &:= \{f \in \mathcal{C}^k(\mathbb{R}^d, \mathbb{R}^m) : \forall \alpha \in \mathbb{N}^d, |\alpha| \in \llbracket 1, k \rrbracket, \sup_{x \in \mathbb{R}^d} |D^\alpha f(x)| < +\infty, [f^{(k)}]_\beta < +\infty\}, \\ \mathcal{C}_b^{k, \beta}(\mathbb{R}^d, \mathbb{R}^m) &:= \{f \in \mathcal{C}^{k, \beta}(\mathbb{R}^d, \mathbb{R}^m) : \|f\|_\infty < +\infty\}. \end{aligned} \quad (2.1)$$

In the above definition, for a bounded mapping  $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}^m$ ,  $m \in \{1, d, d \times r\}$ , we write  $\|\zeta\|_\infty := \sup_{x \in \mathbb{R}^d} \|\zeta \zeta^*(x)\|$  with  $\|\zeta(x)\| = \text{Tr}(\zeta \zeta^*(x))^{1/2}$ , where for  $M \in \mathbb{R}^m \otimes \mathbb{R}^m$ ,  $\text{Tr}(M)$  stands for the trace of  $M$ . Hence  $\|\cdot\|$  is the Fröbenius norm\*.

With these notations,  $\mathcal{C}^{k,\beta}(\mathbb{R}^d, \mathbb{R}^m)$  stands for the subset of  $\mathcal{C}^k(\mathbb{R}^d, \mathbb{R}^m)$  whose elements have bounded derivatives up to order  $k$  and  $\beta$ -Hölder continuous  $k^{\text{th}}$  derivatives. For instance, the space of Lipschitz continuous functions from  $\mathbb{R}^d$  to  $\mathbb{R}^m$  is denoted by  $\mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R}^m)$ .

Eventually, for a given Borel function  $f : \mathbb{R}^d \rightarrow E$ , where  $E$  can be  $\mathbb{R}$ ,  $\mathbb{R}^d$ ,  $\mathbb{R}^d \otimes \mathbb{R}^r$ ,  $\mathbb{R}^d \otimes \mathbb{R}^d$ , we set for  $k \in \mathbb{N}_0$ :

$$f_k := f(X_k).$$

For  $k \in \mathbb{N}_0$ , we denote by  $\mathcal{F}_k := \sigma((X_j)_{j \in \llbracket 0, k \rrbracket})$  the  $\sigma$ -algebra generated by the  $(X_j)_{j \in \llbracket 0, k \rrbracket}$ .

## 2.2 Hypotheses

**(C1)** The first term of the random sequence  $X_0$  is supposed to be sub-Gaussian, i.e. there is a threshold  $\lambda_0 > 0$  such that:

$$\forall \lambda < \lambda_0, \quad \mathbb{E}[\exp(\lambda |X_0|^2)] < +\infty.$$

**(GC)** The innovations  $(U_n)_{n \geq 1}$  form an i.i.d. sequence with law  $\mu$ , we also assume that  $\mathbb{E}[U_1] = 0$  and for all  $(i, j, k) \in \{1, \dots, r\}^3$ ,  $\mathbb{E}[U_1^i U_1^j] = \delta_{ij}$ ,  $\mathbb{E}[U_1^i U_1^j U_1^k] = 0$ . Moreover,  $(U_n)_{n \geq 1}$  and  $X_0$  are independent. Eventually,  $U_1$  satisfies the following standard Gaussian concentration property, i.e. for every 1-Lipschitz continuous function  $g : \mathbb{R}^r \rightarrow \mathbb{R}$  and every  $\lambda > 0$ :

$$\mathbb{E}[\exp(\lambda g(U_1))] \leq \exp\left(\lambda \mathbb{E}[g(U_1)] + \frac{\lambda^2}{2}\right).$$

In particular, Gaussian and symmetrized Bernoulli random variables (in short r.v.) satisfy this inequality.

Pay attention that a wider class of sub-Gaussian distributions could be considered. Namely, random variables for which there exists  $\varpi > 0$  s.t. for all  $\lambda > 0$ :

$$\mathbb{E}[\exp(\lambda g(U_1))] \leq \exp\left(\lambda \mathbb{E}[g(U_1)] + \frac{\varpi \lambda^2}{4}\right). \quad (2.2)$$

It is well know that this assumption yields that for all  $r \geq 0$ ,  $\mathbb{P}[|U_1| \geq r] \leq 2 \exp(-\frac{r^2}{\varpi})$ .

**(C2)** There is a positive constant  $\kappa$  s.t., defining for all  $x \in \mathbb{R}^d$ ,  $\Sigma(x) := \sigma \sigma^*(x)$ :

$$\sup_{x \in \mathbb{R}^d} \text{Tr}(\Sigma(x)) = \sup_{x \in \mathbb{R}^d} \|\sigma(x)\|^2 \leq \kappa.$$

---

\*. This notation allows to define similarly vector and matrix norms. In fact,  $\mathbb{R}^d$  vectors can be regarded as line vectors. Then we define similarly for both cases the uniform norm  $\|\cdot\|_\infty$

( $\mathcal{L}_V$ ) We consider the following Lyapunov like stability condition:

There exists  $V : \mathbb{R}^d \longrightarrow [v^*, +\infty[$  with  $v^* > 0$  s.t.

i)  $V \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ ,  $\|D^2V\|_\infty < \infty$ , and  $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ .

ii) There exists  $C_V \in (0, +\infty)$  s.t. for all  $x \in \mathbb{R}^d$ :

$$|\nabla V(x)|^2 + |b(x)|^2 \leq C_V V(x).$$

iii) Let  $\mathcal{A}$  be the infinitesimal generator associated with the diffusion equation (1.1), defined for all  $\varphi \in \mathcal{C}_0^2(\mathbb{R}^d, \mathbb{R})$  and for all  $x \in \mathbb{R}^d$  by:

$$\mathcal{A}\varphi(x) = b(x) \cdot \nabla\varphi(x) + \frac{1}{2}\text{Tr}(\Sigma(x)D^2\varphi(x)), \quad (2.3)$$

where, for two vectors  $v_1, v_2 \in \mathbb{R}^d$ , the symbol  $v_1 \cdot v_2$  stands for the canonical inner product of  $v_1$  and  $v_2$ . There exist  $\alpha_V > 0$ ,  $\beta_V \in \mathbb{R}^+$  s.t. for all  $x \in \mathbb{R}^d$ ,

$$\mathcal{A}V(x) \leq -\alpha_V V(x) + \beta_V.$$

( $\mathbf{U}$ ) There is a unique invariant distribution  $\nu$  to equation (1.1).

For  $\beta \in (0, 1]$ , we introduce:

( $\mathbf{T}_\beta$ ) We choose a test function  $\varphi$  for which

i)  $\varphi$  smooth enough, i.e.  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ ,

We further assume that:

ii) the mapping  $x \mapsto \langle b(x), \nabla\varphi(x) \rangle$  is Lipschitz continuous.

iii) there exists  $C_{V,\varphi} > 0$  s.t. for all  $x \in \mathbb{R}^d$   $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$ .

( $\mathbf{S}$ ) We assume that the sequence  $(\gamma_k)_{k \geq 1}$  is *small enough*, Namely, we suppose that for all  $k \geq 1$ :

$$\gamma_k \leq \min \left( \frac{1}{2\sqrt{C_V \bar{c}}}, \frac{\alpha_V}{2C_V \|D^2V\|_\infty} \right).$$

The constraint in ( $\mathbf{S}$ ) means that the time steps have to be sufficiently *small* w.r.t. the diffusion coefficients and the Lyapunov function.

**Remark 1** *The above condition ( $\mathcal{L}_V$ ) actually implies that the drift coefficient  $b$  lies, out of a compact set, between two hyperplanes separated from 0. Also, the Lyapunov function is lower than the square norm. In other words, there exist constants  $K, \bar{c} > 0$  such that for all  $|x| \geq K$ ,*

$$|V(x)| \leq \bar{c}|x|^2, \quad |b(x)| \leq \sqrt{C_V \bar{c}}|x|. \quad (2.4)$$

*Actually, for  $\sigma$  bounded,  $b$  Lipschitz continuous, and if there is  $c \geq 1$  s.t. for all  $x \in \mathbb{R}^d$ ,  $-c^{-1}|x|^2 < \langle b(x), x \rangle < -c|x|^2$  then the assumption ( $\mathcal{L}_V$ ) is satisfied.*

*Observe that we have supposed ( $\mathbf{U}$ ) without imposing any non-degeneracy conditions. Existence of invariant distribution follows from ( $\mathcal{L}_V$ ) (see [EK86]). For uniqueness, additional conditions need to be considered ((hypo)ellipticity [KM11], [PV01], [Vil09] or confluence [PP12]).*

**Remark 2** In  $(\mathbf{T}_\beta)$ , condition ii) is direct if we consider Lyapunov function  $V(x) = 1 + |x|^2$ . Indeed,  $\varphi$  is supposed, in condition  $(\mathbf{T}_\beta)$  i), to be Lipschitz continuous and so under a linear map. Hypothesis ii) is natural when there is a function  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$  with  $\nu(f) = 0$  s.t.

$$\mathcal{A}\varphi = f. \quad (2.5)$$

From the definition of  $\mathcal{A}$  in (2.3), we rewrite:

$$\langle \nabla\varphi(x), b(x) \rangle = f(x) - \nu(f) - \frac{1}{2} \text{Tr} \left( \Sigma(x) D_x^2 \varphi(x) \right). \quad (2.6)$$

Since the source  $f$  is Lipschitz continuous, and  $\sigma, D_x^2 \varphi$  are bounded and Lipschitz continuous, the left hand side of the equation (2.6) is also Lipschitz continuous.

We say that assumption **(A)** holds whenever **(C1)**, **(GC)**, **(C2)**, **(L<sub>v</sub>)**, **(U)**, **(T<sub>β</sub>)** for some  $\beta \in (0, 1]$  and **(S)** are fulfilled. Except when explicitly indicated, we assume throughout the chapter that assumption **(A)** is in force.

Assume the step sequence  $(\gamma_k)_{k \geq 1}$  is chosen s.t.  $\gamma_k = k^{-\theta}$ ,  $\theta \in (0, 1]$ . In particular, this implies that, for any  $\varepsilon \geq 0$ ,  $\Gamma_n^{(\varepsilon)} = n^{1-\varepsilon\theta}$  if  $\varepsilon\theta < 1$ ,  $\Gamma_n^{(\varepsilon)} = \ln(n)$  if  $\varepsilon\theta = 1$  and  $\Gamma_n^{(\varepsilon)} = 1$  if  $\varepsilon\theta > 1$ .

## 2.3 On some Related Existing Result

In [LP02], Lamberton and Pagès, proved an asymptotic result with the decreasing step scheme (1.5). Precisely, they obtain the discrete counterpart of (1.4) established in [Bha82], emphasizing as well some discretization effects, leading to a bias in the limit law, when the time step becomes too coarse. This last case is however the one leading to the highest convergence rates in the CLT. We recall here their main results, Theorem 10 of the above reference, for the sake of completeness.

**Theorem 1 (Asymptotic Limit Results in [LP02])** *Assume **(C2)**, **(L<sub>v</sub>)**, **(U)** hold. If  $\mathbb{E}[U_1] = 0$ ,  $\mathbb{E}[U_1^{\otimes 3}] = 0$ , we get the following limit results where  $\nu_n$  stands for the empirical measure defined in (1.6).*

- (a) *Fast decreasing step. If  $\theta \in (\frac{1}{3}, 1]$  and  $\mathbb{E}[|U_1|^6] < +\infty$ , then, for all function  $\varphi \in \mathcal{C}^{2,1}(\mathbb{R}^d, \mathbb{R}) \cap \mathcal{C}^3(\mathbb{R}^d, \mathbb{R})$ , one has:*

$$\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(0, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right).$$

- (b) *Critical decreasing step. If  $\theta = \frac{1}{3}$  and if  $\mathbb{E}[|U_1|^8] < +\infty$ , then for all function  $\varphi \in \mathcal{C}^{3,1}(\mathbb{R}^d, \mathbb{R}) \cap \mathcal{C}^4(\mathbb{R}^d, \mathbb{R})$ , one gets:*

$$\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}\left(\tilde{\gamma}m, \int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu\right),$$

where

$$\begin{aligned}\tilde{\gamma} &:= \lim_{n \rightarrow +\infty} \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}, \quad m := - \int_{\mathbb{R}^d} \left( \text{Tr} \left( \frac{1}{2} D^2 \varphi(x) b(x)^{\otimes 2} \right) + \Phi_4(x) \right) \nu(dx), \\ \Phi_4(x) &:= \int_{\mathbb{R}^r} \text{Tr} \left( \frac{1}{2} D^3 \varphi(x) b(x) (\sigma(x)u)^{\otimes 2} + \frac{1}{24} D^4 \varphi(x) (\sigma(x)u)^{\otimes 4} \right) \mu(du),\end{aligned}$$

recalling that  $\mu$  denotes the law of the i.i.d. innovations  $(U_k)_{k \geq 1}^\dagger$ .

(c) *Slowly decreasing step.* If  $\theta \in (0, \frac{1}{3})$  and if  $\mathbb{E}[|U_1|^8] < +\infty$ , then for all globally Lipschitz function  $\varphi \in \mathcal{C}^{3,1}(\mathbb{R}^d, \mathbb{R}) \cap \mathcal{C}^4(\mathbb{R}^d, \mathbb{R})$ , one gets:

$$\frac{\Gamma_n}{\Gamma_n^{(2)}} \nu_n(\mathcal{A}\varphi) \xrightarrow{\mathbb{P}} m.$$

It is possible to relax the boundedness condition on  $\sigma$  in **(C2)**, considering  $\frac{|\sigma^* \nabla \varphi(x)|^2}{V(x)} \xrightarrow{|x| \rightarrow +\infty} 0$  (strictly sublinear diffusion) in case (a) and  $\sup_{x \in \mathbb{R}^d} \frac{|\sigma^* \nabla \varphi(x)|^2}{V(x)} < +\infty$  (sublinear diffusion) in case (b). We refer to Theorems 9 and 10 in [LP02] for additional details.

**Remark 3** *First of all, observe that the normalization is the same as for (1.4). It is the square root of the considered running time, namely  $t$  for the diffusion and  $\Gamma_n$  for the scheme. In other words, a CLT is still available for the discretization procedure. However, by choosing a critical time step, i.e. for the fast convergence  $\theta = \frac{1}{3}$ , a bias is begot. It can be regarded as a discretization effect. Note that, for all  $\theta \geq \frac{1}{3}$ , any step leads to the same asymptotic variance, namely the carré du champ,  $\int_{\mathbb{R}^d} |\sigma^* \nabla \varphi|^2 d\nu$ , like in [Bha82]. However, for the slow decreasing step,  $\theta < \frac{1}{3}$ , the discretization effect is prominent and “hides” the CLT.*

Let us also mention the work of Panloup [Pan08b], where under similar assumptions for stochastic equation driven by a Lévy process, the convergence of the decreasing time step algorithm towards the invariant distribution of the stochastic process is established (see also [Pan08a] for the CLT associated with square integrable Lévy innovations).

In the current diffusive context, i.e. under **(A)**, some non-asymptotic results were successfully established in Chapter 3. It was as well observed there that if we slacken the regularity of the test function  $\varphi$ , a new bias looms. For  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  with  $\beta \in (0, 1)$ , if  $\theta = \frac{1}{2+\beta}$  then  $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$  exhibits deviations similar to the ones of a biased normal law with a different bias than in Theorem 1, c.f. Theorems 2, 3, 5, 7 and 9 in Chapter 3. When  $\beta = 1$ , the two biases correspond. We willingly shirk any discussion about bias appearance, which is discussed in the formerly mentioned chapter. Our target is to refine Theorem 4 in Chapter 3 that we recall:

**Theorem 2 (Non-asymptotic concentration inequalities in Chapter 3)**

Assume **(A)** holds, if there is  $\vartheta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  satisfying **(T $_\beta$ )** and s.t.

$$\mathcal{A}\vartheta = \|\sigma\|^2 - \nu(\|\sigma\|^2).$$

---

$\dagger$ . With our tensor notations,  $D^2 \varphi(x) b(x)^{\otimes 2} \in (\mathbb{R}^d)^{\otimes 2}$ ,  $D^3 \varphi(x) b(x) (\sigma(x)u)^{\otimes 2} \in (\mathbb{R}^d)^{\otimes 3}$ , and  $D^4 \varphi(x) (\sigma(x)u)^{\otimes 4} \in (\mathbb{R}^d)^{\otimes 4}$ .

For  $\beta \in (0, 1]$  and  $\theta \in (\frac{1}{2+\beta}, 1]$ , there exist two explicit monotonic sequences  $\tilde{c}_n \leq 1 \leq \tilde{C}_n$ ,  $n \geq 1$ , with  $\lim_n \tilde{C}_n = \lim_n \tilde{c}_n = 1$  such that for all  $n \geq 1$  and  $a > 0$ :

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 \tilde{C}_n \exp\left(-\frac{\tilde{c}_n}{2\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2} \Phi_n(a)\right),$$

$$\Phi_n(a) := \left[ \left( a^2 \left( 1 - \frac{2}{1 + \sqrt{1 + 4\tilde{c}_n^3 \frac{\Gamma_n}{a^2}}} \right) \right) \vee \left( a^{4/3} \Gamma_n^{1/3} \tilde{c}_n \left( 1 - \frac{2}{3} \tilde{c}_n \left( \frac{\Gamma_n}{a^2} \right)^{1/3} \right)_+ \right) \right], \quad (2.7)$$

where  $x_+ = \max(x, 0)$  and  $\tilde{c}_n := \left( \frac{[\varphi]_1}{[\vartheta]_1} \right)^{2/3} \nu(\|\sigma\|^2) \|\sigma\|_\infty^{-2/3} \check{c}_n$  with  $\check{c}_n$  being an explicit non-negative sequence s.t.  $\check{c}_n \downarrow_n 1$ .

**Remark 4** Observe that two regimes compete in the above bound. From now on, we refer to **Gaussian deviations** when  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ . In this case, asymptotically the right hand side of the inequality (2.7) is  $2 \exp\left(-\frac{a^2}{2\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2}\right)$ . In other words, the empirical measure is sub-Gaussian with asymptotic variance equals to  $\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2$  which is an upper-bound of the carré du champ,  $\nu(|\sigma^* \nabla\varphi|^2)$  (asymptotic variance in the limit theorem). Thus, this is not fully satisfactory. Throughout the chapter, we refer to **super Gaussian deviations** when  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$ . In this case, a subtle phenomenon appears: the right hand side gives a super Gaussian regime. In particular, the term in the exponential of the r.h.s. of (2.7) is bounded from above and below by  $a^{4/3} \Gamma_n^{1/3}$ . We anyhow emphasize that Theorem 2 in Chapter 3 provides a non-asymptotic Gaussian concentration for all deviation regimes:

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 \tilde{C}_n \exp\left(-\frac{\tilde{c}_n a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right). \quad (2.8)$$

In particular, for super Gaussian deviations, the deviation (2.8) is asymptotically better. However, it had already been observed in Chapter 3 that the bound (2.7) turned out to be useful for numerical purposes as it led to bounds closer to the empirical realizations. We will derive in Theorem 6 of Section 6 a deviation bound similar to (2.7) with an improved variance bound. Namely, we succeed to replace  $\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2$  by the carré du champ. We then observe in the numerical results of Section 7 that the associated deviation bounds match rather precisely those of the empirical realizations.

We will also employ the terminology of **intermediate Gaussian deviations** when  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow C > 0$ . For this regime, we keep a Gaussian regime with deteriorated constants. Again, we first deal with Gaussian deviations, and we postpone the study of super Gaussian deviations to Section 6.

**Remark 5** Actually, in the proof of Theorem 2, we can only use a map  $\vartheta$  satisfying assumption  $(\mathbf{T}_\beta)$  s.t.  $\mathcal{A}\vartheta \geq |\sigma^* \nabla\varphi|^2 - \nu(|\sigma^* \nabla\varphi|^2)$ . However, this inequality is equivalent to the coboundary condition  $\nu - a.s.$ . In fact, we set the function  $f := \mathcal{A}\vartheta - |\sigma^* \nabla\varphi|^2 + \nu(|\sigma^* \nabla\varphi|^2) \geq 0$ , and

$$\nu(f) = \nu(\mathcal{A}\vartheta) = 0. \quad (2.9)$$

As  $f$  is continuous and non negative,  $f = 0$   $\nu - a.s.$  Hence  $\mathcal{A}\vartheta = |\sigma^* \nabla\varphi|^2 - \nu(|\sigma^* \nabla\varphi|^2)$   $\nu - a.s.$

### 3 Main results

Our main contribution consists in establishing a concentration inequality whose variance matches asymptotically the *carré du champ*, see (1.4) and Theorem 1 in what we called the regime of *Gaussian deviations*. In the numerical part of Chapter 3, we see that changing the bound  $\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2$  by the *carré du champ*, leads to bounds much closer to the realizations. Here, we state a simple and “sharp” inequality.

**Theorem 3 (Sharp non-asymptotic deviation results)** *Assume (A) is in force. Suppose that there exists  $\vartheta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  satisfying  $(\mathbf{T}_\beta)$  for some  $\beta \in (0, 1]$  s.t.*

$$\mathcal{A}\vartheta = |\sigma^*\nabla\varphi|^2 - \nu(|\sigma^*\nabla\varphi|^2). \quad (3.1)$$

Then, for  $\theta \in (\frac{1}{2+\beta}, 1]$ , there exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$  s.t. for all  $n \geq 1$ ,  $a > 0$  satisfying  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$  (Gaussian deviations), the following bound holds:

$$\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\nu(|\sigma^*\nabla\varphi|^2)}\right),$$

with  $C_n = \exp\left(\frac{[D^3\varphi]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{(\frac{3+\beta}{2})}}{(1+\beta)(2+\beta)(3+\beta)\sqrt{\Gamma_n}} + p_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}\right)$  for  $p_n \geq 1$  such that  $p_n \rightarrow_n +\infty$  and  $p_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow_n 0$ .

**Remark 6** *We obtain the optimal Gaussian bound with the carré du champ as variance which corresponds to CLT. This is asymptotically the sharpest result that we can expect. This inequality is very important for confidence intervals, as in this context  $a$  is supposed to be “small”, i.e. bounded. Under suitable regularity assumptions on  $f$ , which guarantee that the function  $\varphi$  solving  $\mathcal{A}\varphi = f - \nu(f)$  satisfies  $(\mathbf{T}_\beta)$  for some  $\beta \in (0, 1]$ , it readily follows from Theorem 3 that:*

$$\mathbb{P}\left[\nu(f) \in \left[\nu_n(f) - \frac{a}{\sqrt{\Gamma_n}}, \nu_n(f) + \frac{a}{\sqrt{\Gamma_n}}\right]\right] \geq 1 - 2C_n \exp\left(-c_n \frac{a^2}{2\nu(|\sigma^*\nabla\varphi|^2)}\right).$$

The conditions on  $f$  that lead to the required smoothness on  $\varphi$  and  $\vartheta$  are discussed in Section 5 (see in particular Theorem 4 and Corollary 1). Briefly, it suffices to consider that, additionally to  $(\mathbf{C2})$  and  $(\mathcal{L}_\nu)$ ,  $\Sigma$  is also uniformly elliptic,  $b \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ ,  $\sigma \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$  and that the source  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ . This last assumption on  $f$  can be weakened to Lipschitz continuous (see Theorem 5) with some restriction on the steps.

#### 3.1 User’s guide to the proof

Recall that, for a fixed given  $n \in \mathbb{N}$  and  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ , we want to estimate the quantity

$$\mathbb{P}[\sqrt{\Gamma_n}|\nu_n(\mathcal{A}\varphi)| \geq a], \quad \forall a > 0,$$

where  $\nu_n(\mathcal{A}\varphi) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \mathcal{A}\varphi(X_{k-1})$ . We focus below on the term  $\mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a]$ . Indeed, the contribution  $\mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \leq -a]$  can be handled by symmetry.



The first step of the proof consists in writing  $(\mathcal{A}\varphi(X_{k-1}))_{k \in \llbracket 1, n \rrbracket}$  with a splitting method to isolate the terms depending on the current innovation  $U_k$  for  $\mathcal{A}\varphi(X_{k-1})$ . This is done in Lemma 1 below. Precisely, for all  $k \in \llbracket 1, n \rrbracket$  and  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  we prove that:

$$\begin{aligned} & \varphi(X_k) - \varphi(X_{k-1}) \\ = & \gamma_k \mathcal{A}\varphi(X_{k-1}) + \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\ & + \frac{1}{2} \gamma_k \operatorname{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1} \right) + \psi_k(X_{k-1}, U_k), \end{aligned} \quad (3.2)$$

where

$$\begin{aligned} \psi_k(X_{k-1}, U_k) &= \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &+ \gamma_k \int_0^1 (1-t) \operatorname{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right. \\ &\quad \left. - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right) dt. \end{aligned} \quad (3.3)$$

Observe that  $\psi_k(X_{k-1}, U_k)$  in the r.h.s. of (3.2) is the only term containing the current innovation  $U_k$ . Thus, the mapping  $u \mapsto \psi_k(X_{k-1}, u)$  is Lipschitz continuous, because  $\varphi$  is.

This property is crucial to proceed with a martingale increment technique. Indeed, introducing the compensated increment  $\Delta_k(X_{k-1}, U_k) := \psi_k(X_{k-1}, U_k) - \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]$ , assumption **(GC)** allows to derive:

$$\forall \lambda > 0, \quad \mathbb{E}[\exp(-\lambda \Delta_k(X_{k-1}, U_k)) | \mathcal{F}_{k-1}] \leq \exp\left(\frac{\lambda^2 [\psi(X_{k-1}, \cdot)]_1^2}{2}\right). \quad (3.4)$$

The corner stone of the proof is then to apply recursively this control to the martingale  $M_m := \sum_{k=1}^m \Delta_k(X_{k-1}, U_k)$ ,  $m \in \llbracket 1, n \rrbracket$ .

To control the deviation, the first step is an exponential inequality which combined to (3.2) yields:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp(\lambda \nu_n(\mathcal{A}\varphi))\right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(-\frac{\lambda q M_n}{\Gamma_n}\right)\right]^{1/q} \mathcal{R}_n, \end{aligned} \quad (3.5)$$

for  $\lambda > 0$ ,  $q > 1$  and  $\mathcal{R}_n$  is a remainder (whose behaviour is investigated in Lemma 6). The main contribution in the above equation is the one involving  $M_n$  which can be analyzed thanks to (3.4). Namely:

$$\begin{aligned} \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right] &= \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} M_{n-1}\right) \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} \Delta_n(X_{n-1}, U_n)\right) \middle| \mathcal{F}_{n-1}\right]\right] \\ &\leq \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n} M_{n-1}\right) \exp\left(\frac{q^2 \lambda^2 [\psi(X_{n-1}, \cdot)]_1^2}{2\Gamma_n^2}\right)\right]. \end{aligned} \quad (3.6)$$

A first approach in Chapter 3 in order to iterate the estimates involving the conditional expectations, consisted in bounding uniformly  $[\psi(X_{n-1}, \cdot)]_1 \leq \sqrt{\gamma_n} \|\sigma\|_\infty \|\nabla\varphi\|_\infty$  (which is easily deduced from (3.2)). Iterating the procedure led to the estimate

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \exp\left(\frac{q\lambda^2}{2\Gamma_n} \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2\right) \mathcal{R}_n. \quad (3.7)$$

Optimizing over  $\lambda$ , letting as well  $q \downarrow 1$  in a suitable way, gives the deviation upper-bound  $C_n \exp(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2})$ , with  $C_n, c_n > 0$  respectively increasing and decreasing to 1 with  $n$  (see Theorem 2 of Chapter 3 for details).

To obtain the expected variance corresponding to the *carré du champ*  $\nu(|\sigma^* \nabla\varphi|^2)$ , the key point is to control finely the Lipschitz modulus of  $\psi_k(X_{k-1}, \cdot)$  in (3.4), (5.4).

From (3.2), we get the following simple expression of the derivative  $\nabla_u \psi_k(X_{k-1}, u)|_{u=U_k} = \sqrt{\gamma_k} \sigma_{k-1}^* \nabla\varphi(X_k)$ . Hence, there is a remainder term  $\mathcal{R}(\gamma_k, X_{k-1}, U_k)$  and a constant  $C_{(3.8)} = C_{(3.8)}((\mathbf{A})) > 0$  s.t.

$$|\nabla_u \psi_k(X_{k-1}, u)|^2|_{u=U_k} = \gamma_k |\sigma_{k-1}^* \nabla\varphi_{k-1}|^2 + C_{(3.8)} \gamma_k^2 \sqrt{V_{k-1}} + \mathcal{R}(\gamma_k, X_{k-1}, U_k), \quad (3.8)$$

for more details see (3.29) below.

In order to exhibit for each evaluation of the conditional expectations in (3.4), the contribution  $\nu(|\sigma^* \nabla\varphi|^2)$ , we use the auxiliary Poisson problem:

$$\mathcal{A}\vartheta = |\sigma^* \nabla\varphi|^2 - \nu(|\sigma^* \nabla\varphi|^2). \quad (3.9)$$

We then write for the main term to control in (3.5),

$$\mathbb{E}[\exp(-\frac{\lambda q M_n}{\Gamma_n})] \leq \mathcal{I}_1^{\frac{1}{\rho}} \mathcal{I}_2^{\frac{\rho-1}{q\rho}} \mathcal{I}_3^{\frac{\rho-1}{\rho\rho}}, \quad (3.10)$$

for  $\rho > 1$ ,  $\hat{p}, \hat{q} > 1$  s.t.  $\frac{1}{\hat{p}} + \frac{1}{\hat{q}} = 1$  where:

$$\begin{aligned} \mathcal{I}_1 &:= \mathbb{E} \exp\left(-\rho \frac{q\lambda}{\Gamma_n} M_n - \frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} \sum_{k=1}^n \gamma_k \mathcal{A}\vartheta(X_{k-1}) - C_{(3.8)} \gamma_k^2 \sqrt{V_{k-1}}\right), \\ \mathcal{I}_2 &:= \mathbb{E} \exp\left(\frac{\lambda^2 q^2 \rho^2 \hat{q}}{2(\rho-1)\Gamma_n^2} \sum_{k=1}^n \gamma_k \mathcal{A}\vartheta(X_{k-1})\right), \\ \mathcal{I}_3 &:= \mathbb{E} \exp\left(\frac{\lambda^2 q^2 \rho^2 \hat{p}}{2(\rho-1)\Gamma_n^2} \sum_{k=1}^n C_{(3.8)} \gamma_k^2 \sqrt{V_{k-1}}\right). \end{aligned} \quad (3.11)$$

Exploiting (3.9), we can now rewrite

$$\begin{aligned} &\mathcal{I}_1 \\ &= \mathbb{E} \exp\left(-\rho \frac{q\lambda}{\Gamma_n} M_n - \frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} \sum_{k=1}^n \left(\gamma_k [|\sigma^* \nabla\varphi(X_{k-1})|^2 - \nu(|\sigma^* \nabla\varphi|^2)] + C_{(3.8)} \gamma_k^2 \sqrt{V_{k-1}}\right)\right) \\ &= \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n} \nu(|\sigma^* \nabla\varphi|^2)\right) \\ &\quad \times \mathbb{E} \exp\left(-\rho \frac{q\lambda}{\Gamma_n} M_n - \frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} \sum_{k=1}^n \left(\gamma_k |\sigma^* \nabla\varphi(X_{k-1})|^2 + C_{(3.8)} \gamma_k^2 \sqrt{V_{k-1}}\right)\right). \end{aligned} \quad (3.12)$$

The first term in the above r.h.s. yields the expected variance when we optimize over  $\lambda$  for  $q$  and  $\rho$  going to 1, which is the case in the regime of so called *Gaussian deviations* in Theorem 3. It improves the previous bound (3.7). Introduce now for  $m \in \llbracket 1, n \rrbracket$ ,

$$S_m := \exp \left( -\frac{\rho q \lambda}{\Gamma_n} M_m - \frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} \sum_{k=1}^m (\gamma_k |\sigma^* \nabla \varphi(X_{k-1})|^2 + C \gamma_k^2 \sqrt{V_{k-1}}) \right). \quad (3.13)$$

Bringing to mind that  $M_m = \sum_{k=1}^m \Delta_k(X_{k-1}, U_k)$ , where  $\mathbb{E}[\Delta_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}] = 0$  and  $[\Delta_k(X_{k-1}, \cdot)]_1 = [\psi_k(X_{k-1}, \cdot)]_1$ , we get from (3.8), that, up to the remainder term  $(\mathcal{R}(\gamma_k, X_{k-1}, U_k))_{k \in \llbracket 1, n \rrbracket}$ ,  $S_m$  can be viewed as a super martingale (see Lemma 5 for details). We actually rigorously show that, in the Gaussian regime (i.e. for  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ ), for  $\theta \in (1/3, 1)$

$$\mathbb{E}[S_n]^{\frac{1}{\rho q}} \leq \mathcal{R}_n \xrightarrow{n \rightarrow +\infty} 1.$$

For  $\theta = 1$ , or for *super Gaussian deviations* (i.e. for  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$ , see Section 6) with  $\theta \in (1/3, 1)$  we get:

$$\mathbb{E}[S_n]^{\frac{1}{\rho q}} \leq \mathcal{R}_n \exp \left( \left( \frac{\rho q \lambda^2}{\Gamma_n} + \frac{\rho^3 q^3 \lambda^4}{(\rho - 1)\Gamma_n^3} \right) e_n \right),$$

where  $e_n > 0$  decreases to 0 with  $n$  and  $\mathcal{R}_n > 0$  is still going to 1 with  $n$ . The difficulty in the above control is that the optimized  $\lambda$  also depends on  $n$  and  $\rho$  (see (3.33) below).

The second term  $\mathcal{T}_2$  is estimated directly repeating the arguments of the proof of Theorem 2 in Chapter 3 which are recalled above (see equations (3.5) to (3.7)). We apply the previous martingale increment technique that previously led to (3.7). Denoting by  $M_n^\vartheta$  the martingale associated with the  $(\psi_k^\vartheta(X_{k-1}, U_k))_{k \in \llbracket 1, n \rrbracket}$  deriving from the expansion of  $\mathcal{A}^\vartheta$  similarly to (3.2), we obtain:

$$\mathcal{T}_2 \leq \mathbb{E} \left[ \exp \left( -\frac{\lambda^2 q^2 \rho^2 \bar{q} M_n^\vartheta}{2(\rho - 1)\Gamma_n^2} \right) \right]^{1/\bar{q}} \mathcal{R}_n^\vartheta \leq \exp \left( \frac{\lambda^4 q^4 \rho^4 \bar{q}}{8(\rho - 1)^2 \Gamma_n^3} \|\sigma\|_\infty^2 \|\nabla \vartheta\|_\infty^2 \right) \mathcal{R}_n^\vartheta, \quad (3.14)$$

for  $\bar{q} > 1$  and where the superscript  $\vartheta$  means that we only need to replace  $\varphi$  by  $\vartheta$  in the previous definitions. Like in (3.5),  $\mathcal{R}_n^\vartheta$  is here a remainder.

The third component  $\mathcal{T}_3$  is first controlled by Jensen inequality (over the exponential function and the measure is  $\frac{1}{\Gamma_n^{(3)}} \sum_{k=1}^n \gamma_k^2 \delta_k$ ):

$$\mathcal{T}_3 \leq \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp \left( \frac{\lambda^2 q^2 \rho^2 \hat{p} \Gamma_n^{(2)}}{2(\rho - 1)\Gamma_n^2} C \sqrt{V_{k-1}} \right). \quad (3.15)$$

For the control of this term (as well for remainders from Taylor expansion in (3.2) and in Lemma 1), we recall a useful result from Chapter 3 (see Proposition 1 therein). Under **(A)**, there is a constant  $c_V := c_V(\mathbf{A}) > 0$  such that for all  $\lambda \in [0, c_V]$ ,  $\xi \in [0, 1]$ :

$$I_V^\xi := \sup_{n \geq 0} \mathbb{E}[\exp(\lambda V_n^\xi)] < +\infty. \quad (3.16)$$

We also refer to Lemaire (see [Lem07]) for additional integrability results of the Lyapunov functions in a more general framework. The identity (3.15) is handled by Young inequality

$$\mathcal{F}_3 \leq \frac{\exp\left(\frac{1}{2c_V} \left(\frac{\lambda^2 q^2 \rho^2 \hat{p} \Gamma_n^{(2)}}{2(\rho-1)\Gamma_n^2} C\right)^2\right)}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp(c_V V_{k-1}) = \frac{\exp\left(\frac{\lambda^4}{\Gamma_n^3} e_n\right)}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp(c_V V_{k-1}),$$

with  $\hat{p} \rightarrow_n +\infty$  s.t. for fixed  $\rho, q > 1$ ,  $e_n = \frac{1}{2c_V} \left(\frac{q^2 \rho^2 \hat{p}}{16c_V^2 (\rho-1)^2} \frac{\Gamma_n^{(2)}}{\Gamma_n^2} C\right)^2 \rightarrow_n 0$ , note that for all  $\theta \in (\frac{1}{3}, 1]$ ,  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow_n 0$ . We obtain then by (3.16):

$$\mathcal{F}_3^{\frac{\rho-1}{\hat{p}\rho}} \leq \exp\left(\frac{\lambda^4}{\Gamma_n^3} e_n\right) (I_V^1)^{\frac{\rho-1}{\hat{p}\rho}} = \mathcal{R}_n \exp\left(\frac{\lambda^4}{\Gamma_n^3} e_n\right), \quad (3.17)$$

for  $\hat{p} = \hat{p}(n) \rightarrow_n +\infty$ .

Eventually, by (3.5), (3.12), (3.14) and (3.17) with the different controls of  $\mathbb{E}[S_n]$  (see Lemma 5):

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2}{\Gamma_n} A_n(\rho) + \frac{\lambda^4}{\Gamma_n^3} B_n(\rho)\right) \mathcal{R}_n, \quad (3.18)$$

with  $\mathcal{R}_n \rightarrow 1$ ,  $A_n(\rho) := \rho \left(\frac{q\nu(|\sigma^* \nabla \vartheta|^2)}{2} + e_n\right)$ ,  $B_n := \frac{\rho^3}{\rho-1} \frac{q^3 \hat{q}}{4} \left(\frac{\bar{q} \|\sigma\|_\infty^2 \|\nabla \vartheta\|_\infty^2}{2} + e_n\right)$  for  $e_n > 0$  decreasing to 0 with  $n$ .

We perform an optimization over  $\lambda$  with the Cardan method. However, the optimal choice of  $\lambda$  depends on  $\rho$ . So an optimization can be done for  $\rho$  too. In Lemma 4 below, we choose  $\rho$  for the regime of *Gaussian deviations* (i.e.  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ ) which yields:

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 C_n \exp\left(-c_n \frac{a^2}{2\nu(|\sigma^* \nabla \varphi|^2)}\right),$$

for  $c_n, C_n > 0$  respectively decreasing and increasing (for  $n$  big enough) to 1 with  $n$ .

The optimal choices of  $\lambda$  and  $\rho$  for the regime of *super Gaussian deviations* (i.e.  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$ ) is eventually discussed in Section 6. This leads to

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 C_n \exp\left(-c_n \frac{a^{4/3} \Gamma_n^{1/3}}{2 \|\sigma\|_\infty^{2/3} \|\nabla \vartheta\|_\infty^{2/3}}\right).$$

## 3.2 Technical lemmas and Proof of the Main Results

We first give a decomposition lemma of  $\nu_n(\mathcal{A}\varphi)$  which is the starting point of our analysis. Its proof can be found in Chapter 3 (see Lemma 1 therein).

**Lemma 1 (Decomposition of the empirical measure)** *For all  $n \geq 1$ ,  $k \in \llbracket 1, n \rrbracket$  and  $\varphi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ , the identity (3.2) holds and we have:*

$$\begin{aligned} & \Gamma_n \nu_n(\mathcal{A}\varphi) \\ &= \varphi(X_n) - \varphi(X_0) - \left[ \sum_{k=1}^n \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \right. \\ & \quad \left. + \frac{1}{2} \sum_{k=1}^n \gamma_k \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1}^2 \right) + \sum_{k=1}^n \psi_k(X_{k-1}, U_k) \right], \end{aligned} \quad (3.19)$$

where  $\psi_k(X_{k-1}, U_k)$  is defined in (3.3).

**Remark 7** *In spite of the square terms in  $U_k$  appearing in the r.h.s. of (3.3), we have that, conditionally to  $\mathcal{F}_{k-1}$ ,  $u \mapsto \psi_k(X_{k-1}, u)$  is Lipschitz continuous. Indeed, on the r.h.s.  $U_k$  only appears in  $\psi_k$  and on the l.h.s. we know that  $\varphi$  is Lipschitz. Hence, for all  $(u, u') \in (\mathbb{R}^d)^2$ :*

$$|\psi_k(X_{k-1}, u) - \psi_k(X_{k-1}, u')| \leq \sqrt{\gamma_k} \|\sigma_{k-1}\| \|\nabla \varphi\|_\infty |u - u'|. \quad (3.20)$$

*Our strategy consists in controlling how far the Lipschitz modulus in (3.20) is from  $|\sigma_{k-1}^* \nabla \varphi_{k-1}|$ . The first step is to obtain an explicit derivative of  $\psi_k(X_{k-1}, \cdot)$ , see (3.24) below.*

For notational convenience we introduce, for a given  $n \in \mathbb{N}^*$  the following quantities:

$$\begin{aligned} R_n &:= \varphi(X_n) - \varphi(X_0) - \sum_{k=1}^n \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\ & \quad - \frac{1}{2} \sum_{k=1}^n \gamma_k \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1}^2 \right), \\ M_n &:= \sum_{k=1}^n \Delta_k(X_{k-1}, U_k), \quad \tilde{R}_n := R_n - \sum_{k=1}^n \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}], \end{aligned} \quad (3.21)$$

where for all  $k \in \llbracket 1, n \rrbracket$ :

$$\Delta_k(X_{k-1}, U_k) := \psi_k(X_{k-1}, U_j) - \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}]. \quad (3.22)$$

From these definitions, Lemma 1 can be rewritten:

$$\nu_n(\mathcal{A}\varphi) = \frac{1}{\Gamma_n} (\tilde{R}_n - M_n), \quad (3.23)$$

where  $M_n$  is a martingale. The key idea of the proof is to control more precisely the Lipschitz modulus of  $\psi_n(X_{n-1}, \cdot)$  than it was done in Chapter 3. From the definition in (3.2), let us write for all  $k \in \llbracket 1, n \rrbracket$ :

$$\psi_k(X_{k-1}, U_k) = \varphi_k - \varphi_{k-1} + R_{k-1,k},$$

where

$$\begin{aligned} R_{k-1,k} &:= -\gamma_k \mathcal{A}\varphi(X_{k-1}) - \gamma_k \int_0^1 \langle \nabla\varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla\varphi(X_{k-1}), b_{k-1} \rangle dt \\ &\quad - \frac{1}{2} \gamma_k \text{Tr} \left( (D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})) \Sigma_{k-1}^2 \right). \end{aligned}$$

Hence, by derivation

$$\nabla_u \psi_k(X_{k-1}, u)|_{u=U_k} = \sqrt{\gamma_k} \sigma_{k-1}^* \nabla\varphi(X_k). \quad (3.24)$$

We will establish that  $\nabla_u \psi_k(X_{k-1}, u)|_{u=U_k}$  is not “too far” from  $\sqrt{\gamma_k} \sigma_{k-1}^* \nabla\varphi(X_{k-1})$ .

### Proof of Theorem 3 for bounded innovations

We first give the complete proof in this particular case. We will specify the additional required controls for possibly unbounded innovations in the next subsection.

A key tool in the derivation of our main results is the following lemma whose proof is postponed to Section 4 for the sake of clarity.

**Lemma 2 (Remainders from Taylor decomposition)** *Under (A), for all  $q \geq 1$  and  $\lambda > 0$ , we have:*

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E} \exp\left(-\frac{q\lambda}{\Gamma_n} M_n\right)\right)^{\frac{1}{q}} \exp\left(\frac{\lambda^2}{\Gamma_n} e_n\right) \mathcal{R}_n. \quad (3.25)$$

We will now sharply control the Lipschitz constant of  $\psi_k(X_{k-1}, \cdot)$ , or equivalently  $\Delta_k(X_{k-1}, \cdot)$ , which appears iteratively to handle the martingale term in (3.25).

In case of bounded innovations, we see by assumption **(T<sub>β</sub>)** i) (smoothness of  $\varphi$ ) that:

$$\begin{aligned} |\nabla_u \Delta_k(X_{k-1}, u)|_{u=U_k} &= |\sqrt{\gamma_k} \sigma_{k-1}^* \nabla\varphi(X_k)| \\ &\leq |\sqrt{\gamma_k} \sigma_{k-1}^* \nabla\varphi(X_{k-1})| + |\sqrt{\gamma_k} \sigma_{k-1}^* [\nabla\varphi(X_k) - \nabla\varphi(X_{k-1} + \gamma_k b_{k-1})]| \\ &\quad + |\sqrt{\gamma_k} \sigma_{k-1}^* [\nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) - \nabla\varphi(X_{k-1})]| \\ &\leq \sqrt{\gamma_k} |\sigma_{k-1}^* \nabla\varphi(X_{k-1})| + \gamma_k \|\sigma_{k-1}\|^2 \|D^2\varphi\|_\infty \|U_k\|_\infty \\ &\quad + |\sqrt{\gamma_k} \sigma_{k-1}^* [\nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) - \nabla\varphi(X_{k-1})]|. \end{aligned} \quad (3.26)$$

Remark that we have both controls

$$\begin{aligned} |[\nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) - \nabla\varphi(X_{k-1})]| &\leq \gamma_k \|D^2\varphi\|_\infty |b_{k-1}| \stackrel{(\mathcal{L}_V), \text{ii)}}{\leq} \gamma_k \sqrt{C_V} \|D^2\varphi\|_\infty \sqrt{V_{k-1}}, \\ |[\nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) - \nabla\varphi(X_{k-1})]| &\leq (2\|\nabla\varphi\|_\infty)^{\frac{1}{2}} |[\nabla\varphi(X_{k-1} + \gamma_k b_{k-1}) - \nabla\varphi(X_{k-1})]|^{\frac{1}{2}} \\ &\stackrel{(\mathcal{L}_V), \text{ii)}}{\leq} (2\|\nabla\varphi\|_\infty)^{\frac{1}{2}} \gamma_k^{1/2} / 2C_V^{1/4} \|D^2\varphi\|_\infty^{1/2} V_{k-1}^{1/4}, \end{aligned} \quad (3.27)$$

in order to keep integrable powers of the Lyapunov function. We therefore eventually get from (3.28) and inequalities in (3.27):

$$\begin{aligned} &|\nabla_u \Delta_k(X_{k-1}, u)|_{u=U_k}^2 \\ &\leq \gamma_k |\sigma_{k-1}^* \nabla\varphi(X_{k-1})|^2 \\ &\quad + 2\sqrt{\gamma_k} |\sigma_{k-1}^* \nabla\varphi(X_{k-1})| (\gamma_k \|\sigma_{k-1}\|^2 \|D^2\varphi\|_\infty \|U_k\|_\infty + \gamma_k^{3/2} \|\sigma_{k-1}\|_\infty \sqrt{C_V} \|D^2\varphi\|_\infty \sqrt{V_{k-1}}) \\ &\quad + (\gamma_k \|\sigma_{k-1}\|^2 \|D^2\varphi\|_\infty \|U_k\|_\infty + \|\sigma_{k-1}\|_\infty (2\|\nabla\varphi\|_\infty)^{\frac{1}{2}} \gamma_k C_V^{1/4} \|D^2\varphi\|_\infty^{1/2} V_{k-1}^{1/4})^2 \\ &\leq \gamma_k |\sigma_{k-1}^* \nabla\varphi(X_{k-1})|^2 + C_{1,(3.28)} \gamma_k^{3/2} \|U_k\|_\infty + C_{2,(3.28)} \gamma_k^2 \|U_k\|_\infty^2 + C_{(3.8)} \gamma_k^2 \sqrt{V_{k-1}}, \end{aligned} \quad (3.28)$$

with the constants  $C_{1,(3.28)} := 2\|\sigma\|_\infty^3\|\nabla\varphi\|_\infty\|D^2\varphi\|_\infty$ ,  $C_{2,(3.28)} := 2\|\sigma\|_\infty^4\|D^2\varphi\|_\infty^2$ , and  $C_{(3.8)} := 6\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty\|D^2\varphi\|_\infty$ . The last inequality above is a consequence of convexity inequality (i.e. for all  $(x, y) \in \mathbb{R}^2$ ,  $(x + y)^2 \leq 2x^2 + 2y^2$ ).

Recalling that we consider first  $\|U_k\|_\infty \leq C_\infty$ , we then derive:

$$[\Delta_k(X_{k-1}, \cdot)]_1^2 \leq \gamma_k |\sigma^* \nabla \varphi|^2(X_{k-1}) + C\gamma_k^{3/2} + C_{(3.8)}\gamma_k^2 \sqrt{V_{k-1}}, \quad (3.29)$$

where in the above identity  $C = C_{1,(3.28)}C_\infty + C_{2,(3.28)}\gamma_1^{1/2}C_\infty^2$ . Let us introduce for all  $(m, n) \in \mathbb{N}_0^2$ ,  $m \leq n$  and  $(\rho, q) \in (1, +\infty)^2$ :

$$T_m := \exp\left(-\frac{\rho q \lambda}{\Gamma_n} \Delta_m(X_{m-1}, U_m) - \frac{\rho^2 (q \lambda)^2}{2\Gamma_n^2} \gamma_m |\sigma^* \nabla \varphi(X_{m-1})|^2 - \sum_{k=1}^m C_{(3.8)} \gamma_k^2 V_{k-1}\right). \quad (3.30)$$

From the definition of  $S_m$  in (3.13), we write  $S_m := \prod_{k=1}^m T_k$ . The coefficients  $(T_m)_{m \geq 1}$  can be viewed as multiplicative increments of  $(\tilde{S}_m)_{m \geq 0}$ . Inequality (3.29) precisely allows to quantify the martingality default for  $(\tilde{S}_m)_{m \geq 0}$ . These factors appear when we exploit the auxiliary Poisson problem (3.9) in the definition of  $\mathcal{F}_1$  in (3.11).

$$\mathcal{F}_1 = \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n} \nu(|\sigma^* \nabla \varphi|^2)\right) \mathbb{E}[S_{n-1} \mathbb{E}[T_n | \mathcal{F}_{n-1}]].$$

Thereby, from the upper-bound (3.29) of the Lipschitz modulus, we directly obtain from (3.30) and **(GC)**

$$\begin{aligned} & \mathbb{E}[T_n | \mathcal{F}_{n-1}] \\ &= \exp\left(-\frac{\rho^2 (q \lambda)^2}{2\Gamma_n^2} \gamma_n |\sigma^* \nabla \varphi(X_{n-1})|^2\right) \mathbb{E}\left[\exp\left(\frac{\rho q \lambda}{\Gamma_n} \Delta_n(X_{n-1}, U_n) - \sum_{k=1}^n C_{(3.8)} \gamma_k^2 V_{k-1}\right) \middle| \mathcal{F}_{n-1}\right] \\ &\leq \exp\left(\frac{\rho^2 (q \lambda)^2}{2\Gamma_n^2} C \gamma_n^{3/2} - \sum_{k=1}^{n-1} C_{(3.8)} \gamma_k^2 V_{k-1}\right). \end{aligned}$$

Hence, iterating:

$$\begin{aligned} \mathcal{F}_1 &\leq \exp\left(\frac{\rho^2 (q \lambda)^2}{2\Gamma_n} \nu(|\sigma^* \nabla \varphi|^2)\right) \mathbb{E}[S_{n-1}] \exp\left(\frac{\rho^2 (q \lambda)^2}{2\Gamma_n^2} \gamma_n^{3/2} C\right) \\ &\leq \exp\left(\frac{\rho^2 (q \lambda)^2}{2\Gamma_n} \nu(|\sigma^* \nabla \varphi|^2)\right) \exp\left(\frac{\rho^2 (q \lambda)^2}{2\Gamma_n} \underbrace{\frac{\Gamma_n^{(3/2)}}{\Gamma_n}}_{=e_n} C\right) = \exp\left(\frac{\rho^2 (q \lambda)^2}{2\Gamma_n} (\nu(|\sigma^* \nabla \varphi|^2) + e_n)\right), \end{aligned}$$

where  $e_n \rightarrow 0$ . The controls for  $\mathcal{F}_2$  are deduced from (3.14), and  $\mathcal{F}_3$  from (3.17). We now gather the previous estimates into (3.10) (we recall  $\mathbb{E}[\exp(-\frac{\lambda q M_n}{\Gamma_n})] \leq \mathcal{F}_1^{\frac{1}{\rho}} \mathcal{F}_2^{\frac{\rho-1}{q\rho}} \mathcal{F}_3^{\frac{\rho-1}{p\rho}}$ ) in the following lemma. Note also that the term  $\mathcal{R}_n^\theta$  appearing in (3.14) is controlled similarly to remainders in Lemma 2.

**Lemma 3 (Gaussian concentration term)** *With notations of (3.21), under **(A)**, for a bounded  $\rho > 1$ , we have:*

$$\mathbb{E} \exp\left(-\frac{\lambda q}{\Gamma_n} M_n\right)^{\frac{1}{q}} \leq \exp\left(\frac{\lambda^2}{\Gamma_n} A_n + \frac{\lambda^4}{\Gamma_n^3} B_n\right) \mathcal{R}_n,$$

where

$$A_n := \rho \left( \frac{q\nu(|\sigma^* \nabla \varphi|^2)}{2} + e_n \right) \quad \text{and} \quad B_n := \frac{\rho^3}{\rho-1} \frac{q^3 \hat{q}}{4} \left( \frac{\bar{q} \|\sigma\|_\infty^2 [\vartheta]_1^2}{2} + e_n \right), \quad (3.31)$$

for some  $1 < \bar{q} := \bar{q}(n) \rightarrow 1$ , and with:  $e_n \xrightarrow[n \rightarrow +\infty]{} 0$ ,  $\mathcal{R}_n \xrightarrow[n \rightarrow +\infty]{} 1$  uniformly in  $\lambda$ .

As a consequence of the previous Lemmas 2 and 3, we obtain (3.18), namely:

$$\mathbb{P}(\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a) \leq C_n \exp(P(\lambda)), \quad (3.32)$$

with  $P(\lambda) := -\frac{a\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2}{\Gamma_n} A_n + \frac{\lambda^4}{\Gamma_n^3} B_n$ , where  $A_n = A_n(\rho) := \rho \tilde{A}_n$  and  $B_n = B_n(\rho) := \frac{\rho^3}{\rho-1} \tilde{B}_n$  with

$$\tilde{A}_n = \frac{q\nu(|\sigma^* \nabla \varphi|^2)}{2} + e_n \quad \text{and} \quad \tilde{B}_n = \frac{q^3 \hat{q}}{4} \left( \frac{\bar{q} \|\sigma\|_\infty^2 \|\nabla \vartheta\|_\infty^2}{2} + e_n \right).$$

Next, like enunciated at the end of the *User's guide to the proof*, we optimize a fourth order polynomial by the Cardan method, see (3.33) below (and Section 4 in Chapter 3). If  $\lambda_n = \arg \min_\lambda P(\lambda)$ , then

$$P'(\lambda_n) = -\frac{a}{\sqrt{\Gamma_n}} + \frac{2\lambda_n}{\Gamma_n} A_n + \frac{4\lambda_n^3}{\Gamma_n^3} B_n = 0.$$

The Cardan-Tartaglia formula yields only one positive real root. Namely, setting

$$\Phi_n(a, \rho) = \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} + \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} + \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} - \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}},$$

this conducts to take  $\lambda = \lambda_n$  with:

$$\lambda_n := \frac{\Gamma_n (\rho-1)^{\frac{1}{3}}}{2\rho} \Phi_n(a, \rho). \quad (3.33)$$

Moreover, remark from the binomial Newton expansion that:

$$\Phi_n(a, \rho)^3 = \frac{2a}{\sqrt{\Gamma_n} \tilde{B}_n} - \frac{2(\rho-1)^{1/3} \tilde{A}_n}{\tilde{B}_n} \Phi_n(a, \rho).$$

From (3.33) and the above expression,  $P(\lambda_n) = P_{\min}(a, \Gamma_n, \rho) = \lambda_n \left( -\frac{a}{\sqrt{\Gamma_n}} + \frac{\lambda_n}{\Gamma_n} A_n + \frac{\lambda_n^3}{\Gamma_n^3} B_n \right)$ , we thus obtain

$$P_{\min}(a, \Gamma_n, \rho) := -\frac{\sqrt{\Gamma_n} (\rho-1)^{1/3} \Phi_n(a, \rho)}{2^3 \rho} (3a - \sqrt{\Gamma_n} (\rho-1)^{1/3} \tilde{A}_n \Phi_n(a, \rho)). \quad (3.34)$$

Then, from (3.32):

$$\mathbb{P}(\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a) \leq C_n \exp(P_{\min}(a, \Gamma_n, \rho)), \quad (3.35)$$

which is exactly the same bound appearing in Remark 11 in Chapter 3, up to a modification of  $\tilde{A}_n$ , containing here the expected *carré du champ*.

The optimization over  $\lambda$  leads to study how  $\rho$  should asymptotically behave. The following lemma indicates that, when  $a = o(\sqrt{\Gamma_n})$ , taking  $\rho-1 \asymp \frac{a}{\sqrt{\Gamma_n}}$  yields a Gaussian concentration inequality in (3.32) with the optimal constant.



**Lemma 4 (Choice of  $\rho$  for the Gaussian concentration regime)**

For  $P_{\min}(a, \Gamma_n, \rho)$  as in (3.34), there is  $\rho := \rho(n, a) > 1$  s.t. If  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ , taking  $\rho - 1 \asymp \frac{a}{\sqrt{\Gamma_n}}$

$$P_{\min}(a, \Gamma_n, \rho) \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0}{=} -\frac{a^2}{2\nu(|\sigma^* \nabla \varphi|^2)}(1 + o(1)).$$

For the sake of clarity, the proof of Lemma 4 is postponed to Section 4.2. From (3.35) and Lemma 4, we conclude the proof of Theorems 3 and 7 for bounded innovations.

**Proof of Theorems 3 for unbounded innovations**

Switching to unbounded innovations requires additional technicalities. Our strategy consists in considering a truncation argument writing

$$\Delta_k(X_{k-1}, U_k) = \Delta_k(X_{k-1}, U_k)[\mathbb{1}_{|U_k| \leq \frac{r_{k,n}}{2}} + \mathbb{1}_{|U_k| > \frac{r_{k,n}}{2}}],$$

to control the Lipschitz modulus of  $\psi_{k-1}$  where  $(r_{k,n})_{n \geq 1, k \leq n}$  is a suitable sequence specified in (3.40) below. In particular,  $r_{k,n} := r_{k,n}((\mathbf{A}), \lambda, \rho)$  where  $\lambda > 0, \rho > 1$  are as in the *User's Guide to the Proof*.

For our choice below, we will have that, for all  $k \in \llbracket 1, n \rrbracket$ ,  $r_{k,n} \uparrow_n +\infty$ . That choice for  $r_{k,n}$  also yields that when  $|U_k| \leq r_{k,n}$ , our controls behave like for the bounded case. But when  $|U_k| > r_{k,n}$ , we will handle this large deviation regime by assumption **(GC)**. We indeed know that for all  $K > 0$ :

$$\mu(\{|x| > K\}) \leq 2 \exp\left(-\frac{K^2}{2}\right). \quad (3.36)$$

Let us recall from the definition of  $\Delta_k(X_{k-1}, U_k)$  in (3.3) and (3.21) that:

$$\Delta_k(X_{k-1}, U_k) = \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) + \Xi_k(X_{k-1}, U_k), \quad (3.37)$$

where for all  $(k, u) \in \llbracket 1, n \rrbracket \times \mathbb{R}^r$ :

$$\begin{aligned} \Xi_k(X_{k-1}, u) := & \gamma_k \int_0^1 (1-t) \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t \sqrt{\gamma_k} \sigma_{k-1} u) \sigma_{k-1} u \otimes u \sigma_{k-1}^* \right. \\ & \left. - \mathbb{E} \left[ D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t \sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* | \mathcal{F}_{k-1} \right] \right) dt. \end{aligned} \quad (3.38)$$

For the terms  $(T_k)_{k \leq n}$  defined in (3.30), the lemma below controls the “*super martingality*” default of  $S_n = \prod_{k=1}^n T_k$ .

**Lemma 5** For all  $k \in \llbracket 1, n \rrbracket$

$$\mathbb{E}[T_k | \mathcal{F}_{k-1}] \leq \aleph_{k,n}(\lambda, \gamma_k, r_{k,n}) \exp\left(-\sum_{i=1}^{k-1} C_{(3.8)} \gamma_i^2 V_{i-1}\right),$$

with, as in (3.28),  $C_{(3.8)} := 6\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty\|D^2\varphi\|_\infty$ ,

$$\aleph_{k,n}(\lambda, \gamma_k, r_{k,n}) := \left(1 + 2\exp(-Cr_{k,n}^2)\right) \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right)\right) \exp\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} C \gamma_k^{3/2} r_{k,n}^2\right), \quad (3.39)$$

and

$$r_{k,n} = r_{k,n}(\mathbf{A}, \lambda, \rho) := \begin{cases} r_n = C\left(1 + \frac{\rho q \lambda}{\Gamma_n}\right)\left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/4}, & \text{for } \theta \in \left(\frac{1}{3}, 1\right), \\ C\left(1 + \frac{\rho q \lambda}{\Gamma_n}\right) \ln(n+1)^{1/4} \ln(k+1)^{1/2}, & \text{for } \theta = 1, \end{cases} \quad (3.40)$$

for  $C > 0$  s.t.  $r_{1,1} > \bar{c}\sqrt{\gamma_1}\|\sigma\|_\infty\|\nabla\varphi\|_\infty$  for  $\bar{c}$  large enough (every  $\bar{c} > 8$  works, see the proof of Lemma 5, and equation (4.21)). This choice is briefly explained in Remark 8 below.

**Remark 8** *The specific form of the truncation and of the time steps chosen yields*

$$r_{k,n} = \begin{cases} O\left(\left(1 + \frac{\rho q \lambda}{\Gamma_n}\right)n^{\frac{1-\theta}{4}}\right), & \text{if } \theta \in (2/3, 1), \\ O\left(\left(1 + \frac{\rho q \lambda}{\Gamma_n}\right)\ln(n)^{-1/4}n^{\frac{1}{12}}\right), & \text{if } \theta = 2/3, \\ O\left(\left(1 + \frac{\rho q \lambda}{\Gamma_n}\right)n^{\frac{\theta}{8}}\right), & \text{if } \theta \in (1/3, 2/3). \end{cases}$$

Anyhow, we always have for each  $k \leq n$ ,  $r_{k,n} \xrightarrow[n]{+} +\infty$ . Identity (3.41) in the following lemma can give an intuition of our choice in (3.40). This result ensures that the terms  $\aleph_{k,n}(\lambda, \gamma_k, r_{k,n})$  can be viewed as remainders (observe indeed that some  $e_n \rightarrow 0$  appear in the exponential (3.41) below). From the optimization in  $\lambda$  performed in the proof Lemma 4 (see equation (4.10)), the contribution  $\frac{\rho\lambda}{\Gamma_n}$  appearing in (3.39) will be large in the regime of super Gaussian deviations (see as well Remark 15). The above choice of  $r_{k,n}$  actually permits to control the remainders in all the considered regimes.

For  $\theta \in (\frac{1}{3}, 1)$ , the choice in (3.40) can seem natural in order to absorb the term

$\prod_{k=1}^n \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{16}\right)\right) = \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \Gamma_n^{(1/2)} \exp\left(-\frac{r_n^2}{16}\right)\right)$  coming from the iteration of (3.39). For  $\theta = 1$ , the choice is a bit different due to the associated logarithmic explosion rates (i.e.  $\Gamma_n \asymp \ln(n)$ ). Actually, for the Gaussian deviations ( $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ ), we have  $\frac{\rho q \lambda}{\Gamma_n} \rightarrow_n 0$ , see again (4.10) and Remark 15 below, and the term  $\frac{\rho q \lambda}{\Gamma_n}$  could be removed in (3.40). On the other hand, the contribution  $\exp\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} C \gamma_k^{3/2} r_{k,n}^2\right)$  will eventually yield a negligible contribution in the polynomial appearing in Lemma 3. We refer to the proof of Lemma 5 for details.

**Lemma 6 (Control of “super martingality default” of  $S_n$ )** *There exist non negative sequences  $(\mathcal{R}_n)_{n \geq 1}$ ,  $(e_n)_{n \geq 1}$  s.t.  $\mathcal{R}_n \xrightarrow[n]{+} 1$ ,  $e_n \xrightarrow[n]{+} 0$ , and for all  $n \geq 1$ :*

$$\mathbb{E}[S_n] \leq \prod_{k=1}^n \aleph_{k,n}(\lambda, \gamma_k, r_{k,n}) = \mathcal{R}_n \exp\left(\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n} + \frac{\rho^4 q^4 \lambda^4}{\Gamma_n^3}\right)e_n\right). \quad (3.41)$$

Observe that a term in  $\lambda^4$  appears here for the control of  $\mathbb{E}[S_n]$ . This is specifically due to the unbounded contributions. Namely, the exponential term in (3.41) comes from  $\exp\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} C \gamma_k^{3/2} r_{k,n}^2\right)$  in (3.39), the definition of  $r_{k,n}$  in (3.40), and using as well that  $e_n \asymp$

$\frac{\Gamma_n^{(3/2)}}{\Gamma_n} \xrightarrow{n \rightarrow \infty} 0$ . The other terms in (3.39), corresponding to sub-Gaussian tails, give the remainder  $\mathcal{R}_n$ .

Note now carefully that, reproducing the arguments of the bounded case and using as well Lemma 6 to control  $\mathbb{E}[S_n]$  yields that Lemma 3 remains valid, up to a modification of the remainders  $e_n$ . The proof then follows similarly to the bounded case. To sum up, the specificity of the unbounded innovations was to precisely control the Lipschitz constants, considering a suitable truncation, as well as the ‘‘martingality default’’ of  $S_n$  appearing in  $\mathcal{T}_1$ .

## 4 Proofs of technical lemmas

### 4.1 Remainders from the Taylor decomposition

#### Proof of Lemma 2

From the notations in (3.21), we recall (3.23):  $\nu_n(\mathcal{A}\varphi) = \frac{1}{\Gamma_n}(\tilde{R}_n - M_n)$ . The idea is now to write for  $a, \lambda > 0$ :

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a] &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(\frac{\lambda}{\Gamma_n}(\tilde{R}_n - M_n)\right)\right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E}\left[\exp\left(-\frac{q\lambda}{\Gamma_n}M_n\right)\right]^{1/q} \mathbb{E}\left[\exp\left(\frac{p\lambda}{\Gamma_n}|\tilde{R}_n|\right)\right]^{1/p}, \end{aligned} \quad (4.1)$$

for  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $p, q > 1$ . We rewrite the Taylor expansion with the same notations as in Chapter 3:  $\tilde{R}_n = L_n - (D_{2,b,n} + D_{2,\Sigma,n} + \bar{G}_n)$  where:

$$\begin{aligned} D_{2,b,n} &:= \sum_{k=1}^n \gamma_k \int_0^1 \langle \nabla\varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla\varphi(X_{k-1}), b_{k-1} \rangle dt, \\ D_{2,\Sigma,n} &:= \frac{1}{2} \sum_{k=1}^n \gamma_k \text{Tr}\left(\left(D^2\varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2\varphi(X_{k-1})\right)\Sigma_{k-1}^2\right), \\ \bar{G}_n &:= \sum_{k=1}^n \mathbb{E}[\psi_k(X_{k-1}, U_k) | \mathcal{F}_{k-1}], \\ L_n &:= \varphi(X_n) - \varphi(X_0). \end{aligned} \quad (4.2)$$

From (4.2), (4.1) and the Cauchy-Schwarz inequality, we get:

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a] &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \left(\mathbb{E}\exp\left(-\frac{q\lambda}{\Gamma_n}M_n\right)\right)^{\frac{1}{q}} \times \left(\mathbb{E}\exp\left(\frac{2p\lambda}{\Gamma_n}|L_n|\right)\right)^{\frac{1}{2p}} \\ &\quad \times \left(\mathbb{E}\exp\left(\frac{4p\lambda}{\Gamma_n}|\bar{G}_n|\right)\right)^{\frac{1}{4p}} \left(\mathbb{E}\exp\left(\frac{8p\lambda}{\Gamma_n}|D_{2,\Sigma,n}|\right)\right)^{\frac{1}{8p}} \left(\mathbb{E}\exp\left(\frac{8p\lambda}{\Gamma_n}|D_{2,b,n}|\right)\right)^{\frac{1}{8p}}. \end{aligned} \quad (4.3)$$

The term  $L_n$  in (4.3) is controlled in Lemma 4 in Chapter 3 (for  $j = 2$  therein):

$$\left(\mathbb{E}\exp\left(4p\lambda\frac{|L_n|}{\Gamma_n}\right)\right)^{\frac{1}{4p}} \leq (I_V^1)^{\frac{1}{4p}} \exp\left(\frac{3pC_{V,\varphi}^2\lambda^2}{c_V\Gamma_n^2} + \frac{c_V}{p}\right) = \mathcal{R}_n \exp\left(\frac{\lambda_n^2}{\Gamma_n}e_n\right), \quad (4.4)$$

for  $p = p_n \rightarrow_n +\infty$  s.t.  $\frac{p}{\Gamma_n} \rightarrow_n 0$ .

Thanks to Lemma 3 in Chapter 3, we obtain:

$$\frac{|\bar{G}_n|}{\sqrt{\Gamma_n}} \leq a_n := \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{\frac{3+\beta}{2}}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}}, \text{ a.s. .}$$

Moreover,  $a_n \rightarrow_n a_\infty = 0$  for  $\theta \in (\frac{1}{2+\beta}, 1]$ . Hence, for all  $p > 1$ :

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda}{\Gamma_n} |\bar{G}_n| \right) \right)^{\frac{1}{4p}} \leq \exp \left( \frac{\lambda}{\sqrt{\Gamma_n}} a_n \right) \leq \exp \left( \frac{\lambda^2}{2\Gamma_n p} + \frac{a_n^2 p}{2} \right) = \mathcal{R}_n \exp \left( \frac{\lambda^2}{\Gamma_n} e_n \right). \quad (4.5)$$

for  $p = p_n \rightarrow_n +\infty$  s.t.  $a_n^2 p \rightarrow_n 0$ .

We handle the term  $D_{2,\Sigma,n}$  in (4.3) by Lemma 5 in Chapter 3: there exists  $C_1 := C_1((\mathbf{A}), \varphi) > 0$  such that

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,\Sigma,n}| \right) \right)^{\frac{1}{4p}} \leq \exp \left( C_1 \frac{p\lambda_n^2 (\Gamma_n^{(2)})^2}{\Gamma_n^2} \right) (I_V^1)^{\frac{1}{4p}} = \mathcal{R}_n \exp \left( \frac{\lambda_n^2}{\Gamma_n} e_n \right), \quad (4.6)$$

for  $p = p_n \rightarrow_n +\infty$  s.t.  $p \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} \rightarrow_n 0$  (we recall that for all  $\theta \in (\frac{1}{3}, 1]$ ,  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow_n 0$ ).

We deal with the term  $D_{2,b,n}$  from (4.3). Because  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  is Lipschitz continuous, thanks to Lemma 5 in Chapter 3, we know that there exists  $C_2 := C_2((\mathbf{A}), \varphi) > 0$  such that:

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda_n}{\Gamma_n} |D_{2,b,n}| \right) \right)^{\frac{1}{4p}} \leq \exp \left( C_2 \left( \frac{3p\lambda_n^2 (\Gamma_n^{(2)})^2}{2\Gamma_n^2} + \frac{1}{2p} \right) \right) (I_V^1)^{\frac{1}{4p}} = \mathcal{R}_n \exp \left( \frac{\lambda_n^2}{\Gamma_n} e_n \right), \quad (4.7)$$

also for  $p = p_n \rightarrow_n +\infty$  s.t.  $p \frac{(\Gamma_n^{(2)})^2}{\Gamma_n} \rightarrow_n 0$ .

We gather (4.4), (4.5), (4.6) and (4.7) into (4.3), which allows us to control the remainder involving  $\tilde{R}_n$  previously decomposed in (4.2). There are non-negative sequences  $(\mathcal{R}_n)_{n \geq 1}$ ,  $(e_n)_{n \geq 1}$  s.t.  $\lim_{n \rightarrow +\infty} \mathcal{R}_n = 1$ ,  $\lim_{n \rightarrow +\infty} e_n = 0$  and:

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \right) \left( \mathbb{E} \exp \left( -\frac{q\lambda}{\Gamma_n} M_n \right) \right)^{\frac{1}{q}} \exp \left( \frac{\lambda_n^2}{\Gamma_n} e_n \right) \mathcal{R}_n. \quad (4.8)$$

## 4.2 Asymptotics in the parameter $\rho$

We first begin with the proof of Lemma 4 which is purely analytical and rather independent of our probabilistic setting. We recall that we use it for both bounded and unbounded innovations.

### Proof of Lemma 4

From the expression of  $\Phi_n(a, \rho)$  in Theorem 7, remember that

$$\Phi_n(a, \rho) = \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} + \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} + \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} - \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}}.$$

Here,  $\rho > 1$  is a *free* parameter. Let us set:

$$\rho - 1 := \xi \frac{27 \tilde{B}_n a^2}{8 \tilde{A}_n^3 \Gamma_n}, \quad (4.9)$$

for a parameter  $\xi := \xi(a, n) > 0$  to optimize. This choice yields

$$\Phi_n(a, \rho) = \frac{a^{1/3}}{\tilde{B}_n^{1/3} \Gamma_n^{1/6}} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right).$$

Hence, from the definition of  $\lambda_n$  in (3.33):

$$\frac{\rho \lambda_n}{\Gamma_n} = \frac{1}{2} (\rho - 1)^{1/3} \Phi_n(a, \rho) = \frac{3}{2^2} \frac{a}{\tilde{A}_n \sqrt{\Gamma_n}} \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right). \quad (4.10)$$

We point out that  $\xi \mapsto \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right)$  is a bounded function from  $[0, +\infty)$  to  $[0, \frac{2}{3})$ . In fact,

$$\begin{aligned} & \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \\ & \stackrel{\xi \rightarrow +\infty}{=} \xi^{1/3} (1 + \xi)^{1/6} \left( 1 + \frac{1}{3\sqrt{1 + \xi}} - \left( 1 - \frac{1}{3\sqrt{1 + \xi}} \right) + o\left(\frac{1}{\sqrt{1 + \xi}}\right) \right) \underset{\xi \rightarrow +\infty}{\sim} \frac{2}{3}. \end{aligned} \quad (4.11)$$

From the definition of  $P_{\min}(a, \rho, \Gamma_n) = P(\lambda_n) = \min_{x>0} P(x)$  in (3.34) and from the definition of  $P(\lambda)$  in (3.32),

$$\begin{aligned} P(\lambda_n) &= -\frac{(\rho - 1)^{1/3} \sqrt{\Gamma_n} \Phi_n(a, \rho)}{\rho} (3a - \sqrt{\Gamma_n} (\rho - 1)^{1/3} \tilde{A}_n \Phi_n(a, \rho)) \\ &= -\frac{(\rho - 1)^{1/3} \Phi_n(a, \rho)}{2} \times \frac{\sqrt{\Gamma_n}}{4\rho} \times (3a - \tilde{A}_n \sqrt{\Gamma_n} (\rho - 1)^{1/3} \Phi_n(a, \rho)) \\ &\stackrel{(4.9), (4.10)}{=} -\frac{3^2}{2^4} \frac{a^2}{\tilde{A}_n \left( \xi \frac{27 \tilde{B}_n a^2}{8 \tilde{A}_n^3 \Gamma_n} + 1 \right)} \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \\ &\quad \times \left( 1 - \frac{\xi^{1/3}}{2} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \right) \\ &=: -\frac{3^2}{2^4 \tilde{A}_n} a^2 f_\Psi(\xi), \end{aligned} \quad (4.12)$$

for

$$f_\Psi : \xi \in \mathbb{R}_+ \mapsto \frac{g(\xi)}{\Psi \xi + 1}, \quad (4.13)$$

where

$$g : \xi \mapsto \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \left( 1 - \frac{\xi^{1/3}}{2} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \right),$$

and

$$\Psi := \frac{27 \tilde{B}_n a^2}{8 \tilde{A}_n^3 \Gamma_n} \left( \stackrel{(4.9)}{=} \frac{\rho - 1}{\xi} \right). \quad (4.14)$$

We bring to mind that we consider ‘‘Gaussian deviations’’, namely  $\frac{a}{\sqrt{\Gamma_n}} \xrightarrow[n]{n} 0$ .

From (6.11) and the asymptotic of  $\tilde{A}_n$  defined in (6.24), i.e.  $4\tilde{A}_n \xrightarrow[n]{n} 2\nu(|\sigma^* \nabla \varphi|^2)$ , we want to choose  $\xi := \xi(a, n)$  s.t.  $\Lambda(\xi) := \frac{3^2}{2^4} f_\Psi(\xi) \xrightarrow[\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0]{n} \frac{1}{4}$ . This would indeed yield

$$P(\lambda_n) \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0}{\sim} -\frac{a^2}{4\tilde{A}_n}. \text{ From the definition of } \Psi \text{ in (6.13), we have } \Psi = \frac{27\tilde{B}_n a^2}{8\tilde{A}_n \Gamma_n} \xrightarrow[\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0]{n} 0.$$

Observe then, that, taking  $\xi$  going to infinity such that

$$\xi \Psi \xrightarrow[\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0]{n} 0 \quad (4.15)$$

yields  $\Lambda(\xi) = \frac{3^2}{2^4} \frac{g(\xi)}{\Psi \xi + 1} \xrightarrow[\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0]{n} \frac{1}{4}$ , noting from (4.11) and the above definition of  $g$  that  $g(\xi) \xrightarrow[\xi \rightarrow +\infty]{n} \left(\frac{2}{3}\right)^2$ .

**Remark 9 (Controls of the optimized parameters for *Gaussian deviations*)** *We give here some useful estimates to control the remainder terms in the truncation procedure associated with unbounded innovations, see proof of Lemma 6. They specify the behaviour of the quantity  $\frac{\rho \lambda_n}{\Gamma_n}$ .*

*In the regime of Gaussian deviations, we get from (4.10) and (4.11):*

$$\frac{\rho \lambda_n}{\Gamma_n} \xrightarrow[\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0]{n} 0, \lambda_n \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0}{\asymp} a \sqrt{\Gamma_n}. \quad (4.16)$$

### 4.3 Technical Lemmas for Unbounded Innovations

We proceed with the proof of Lemmas 5 and 6 which are specifically needed for unbounded innovations.

**Proof of Lemma 5** We use a partition at the threshold  $r_{k,n}$  on the variable  $U_k$ , to control finely the term  $\Delta_k(X_{k-1}, U_k)$ :

$$\begin{aligned} \mathbb{E} \left[ \exp \left( -\frac{\rho q \lambda}{\Gamma_n} \Delta_k(X_{k-1}, U_k) \right) \middle| \mathcal{F}_{k-1} \right] &= \mathbb{E} \left[ \exp \left( -\frac{\rho q \lambda}{\Gamma_n} \Delta_n(X_{k-1}, U_k) \right) \mathbb{1}_{|U_k| \leq r_{k,n}} \middle| \mathcal{F}_{k-1} \right] \\ &+ \mathbb{E} \left[ \exp \left( -\frac{\rho q \lambda}{\Gamma_n} \Delta_k(X_{k-1}, U_k) \right) \mathbb{1}_{|U_k| > r_{k,n}} \middle| \mathcal{F}_{k-1} \right] =: T_{k,s}^M + T_{k,l}^M, \end{aligned} \quad (4.17)$$

where  $T_{k,s}^M$  and  $T_{k,l}^M$  stand for the contributions in  $\mathbb{E}[T_k | \mathcal{F}_{k-1}] \exp(\sum_{i=1}^{k-1} C_{(3.8)} \gamma_i^2 V_{i-1})$  associated with the martingale increment  $\Delta(X_{k-1}, U_k)$  for which the innovation is respectively *small* and *large*.

Let us first write:

$$T_{k,s}^M = \mathbb{E} \left[ \exp \left( -\frac{\rho q \lambda}{\Gamma_n} \Delta_n(X_{k-1}, U_k) \right) \mathbb{1}_{|U_k| \leq r_{k,n}} \middle| \mathcal{F}_{k-1} \right] = \int_{|u| \leq r_{k,n}} \exp \left( -\frac{\rho q \lambda}{\Gamma_n} \Delta_n(X_{k-1}, U_k) \right) \mu(du).$$

Observe now that, similarly to the computations of Section 3.2 for bounded innovations, see equation (3.29), if  $|u| \leq r_{k,n}$ ,  $u \in \mathbb{R}^r \mapsto \Delta_n(X_{k-1}, u)$  is s.t.

$$[\Delta_n(X_{k-1}, \cdot)]_{1,B(0,r_{k,n})}^2 \leq \gamma_k |\sigma_{k-1}^* \nabla \varphi(X_{k-1})|^2 + C \gamma_k^{3/2} r_{k,n}^2 + C_{(3.8)} \gamma_k^2 V_{k-1}, \quad (4.18)$$

where  $[\Delta_n(X_{k-1}, \cdot)]_{1, B(0, r_{k,n})}^2$  denotes the Lipschitz modulus of  $\Delta_n(X_{k-1}, \cdot)$  restricted on the ball  $B(0, r_{k,n})$  of  $\mathbb{R}^r$  with radius  $r_{k,n}$ .

Let us now extend  $u \in B(0, r_{k,n}) \mapsto \Delta_n(X_{k-1}, u)$  into a Lipschitz function on the whole set  $\mathbb{R}^r$  which globally verifies the bound of equation (4.18). The easiest way to do so is to consider:

$$u \in \mathbb{R}^r \mapsto \bar{\Delta}(X_{k-1}, u) = \Delta(X_{k-1}, \Pi_{\bar{B}(0, r_{k,n})}(u)),$$

where  $\Pi_{\bar{B}(0, r_{k,n})}(\cdot)$  denotes the projection on  $\bar{B}(0, r_{k,n})$ , namely for  $u \in \bar{B}(0, r_{k,n})$ , we have  $\Pi_{\bar{B}(0, r_{k,n})}(u) = u$ , for  $u \notin B(0, r_{k,n})$ ,  $\Pi_{\bar{B}(0, r_{k,n})}(u) = \frac{u}{|u|}r_{k,n}$ . It is readily seen that:

$$[\bar{\Delta}(X_{k-1}, \cdot)]_1^2 \leq \gamma_k |\sigma_{k-1}^* \nabla \varphi(X_{k-1})|^2 + C \gamma_k^{3/2} r_{k,n}^2 + C_{(3.8)} \gamma_k^2 V_{k-1}.$$

Hence:

$$\begin{aligned} T_{k,s}^M &= \int_{|u| \leq r_{k,n}} \exp\left(-\frac{\rho q \lambda}{\Gamma_n} \bar{\Delta}_n(X_{k-1}, U_k)\right) \mu(du) \leq \mathbb{E}\left[\exp\left(-\frac{\rho q \lambda}{\Gamma_n} \bar{\Delta}_n(X_{k-1}, U_k)\right) \middle| \mathcal{F}_{k-1}\right] \\ &\stackrel{(\mathbf{GC})}{\leq} \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} [\bar{\Delta}(X_{k-1}, \cdot)]_1^2 - \frac{\rho q \lambda}{\Gamma_n} \mathbb{E}[\bar{\Delta}(X_{k-1}, U_k)]\right) \\ &\leq \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} (\gamma_k |\sigma_{k-1}^* \nabla \varphi(X_{k-1})|^2 + C \gamma_k^{3/2} r_{k,n}^2 + C_{(3.8)} \gamma_k^2 V_{k-1}) - \frac{\rho q \lambda}{\Gamma_n} \mathbb{E}[\bar{\Delta}(X_{k-1}, U_k)]\right). \end{aligned} \quad (4.19)$$

Bearing in mind that  $\mathbb{E}[\Delta(X_{k-1}, U_k)] = 0$ , and observe now:

$$\begin{aligned} &\mathbb{E}[\bar{\Delta}(X_{k-1}, U_k)] \\ &= \mathbb{E}[\Delta(X_{k-1}, U_k) \mathbf{1}_{|U_k| < r_{k,n}}] + \mathbb{E}[\bar{\Delta}(X_{k-1}, U_k) \mathbf{1}_{|U_k| \geq r_{k,n}}] \\ &= -\mathbb{E}[\Delta(X_{k-1}, U_k) \mathbf{1}_{|U_k| \geq r_{k,n}}] + \mathbb{E}[\Delta(X_{k-1}, \Pi_{B(0, r_{k,n})}(U_k)) \mathbf{1}_{|U_k| \geq r_{k,n}}] \\ &\leq \left(\mathbb{E}[|\Delta(X_{k-1}, U_k)|^2]^{1/2} + \mathbb{E}[|\Delta(X_{k-1}, \Pi_{B(0, r_{k,n})}(U_k))|^2]^{1/2}\right) \mathbb{P}[|U_k| \geq r_{k,n}]^{1/2} \\ &\leq C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right), \end{aligned}$$

exploiting (3.36), (3.37) and (3.38) for the last inequality. Plugging this bound into (4.19) yields:

$$T_{k,s}^M \leq \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} (\gamma_k |\sigma_{k-1}^* \nabla \varphi(X_{k-1})|^2 + C \gamma_k^{3/2} r_{k,n}^2 + C_{(3.8)} \gamma_k^2 V_{k-1}) + C \gamma_k^{1/2} \frac{\rho q \lambda}{\Gamma_n} \exp\left(-\frac{r_{k,n}^2}{4}\right)\right). \quad (4.20)$$

The remaining term in (4.17), involving also the large deviations of the innovation, can be controlled as follows:

$$\begin{aligned} T_{k,l}^M &= \mathbb{E}\left[\exp\left(-\frac{\rho q \lambda}{\Gamma_n} \Delta_k(X_{k-1}, U_k)\right) \mathbf{1}_{|U_k| > r_{k,n}} \middle| \mathcal{F}_{k-1}\right] \\ &\leq \mathbb{E}\left[\exp\left(-\frac{2\rho q \lambda}{\Gamma_n} \Delta_k(X_{k-1}, U_k)\right) \middle| \mathcal{F}_{k-1}\right]^{1/2} \mathbb{E}\left[\mathbf{1}_{|U_k| > r_{k,n}}\right]^{1/2} \\ &\leq 2\mathbb{E}\left[\exp\left(-\frac{2\rho q \lambda}{\Gamma_n} \Delta_k(X_{k-1}, U_k)\right) \middle| \mathcal{F}_{k-1}\right]^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right) \\ &\stackrel{(\mathbf{GC})}{\leq} 2 \exp\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} \gamma_k \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2 - \frac{r_{k,n}^2}{4}\right). \end{aligned}$$

Let us proceed from our definition of  $r_{k,n} = r_{k,n}((\mathbf{A}), \lambda, \rho)$  in (3.40) and write  $r_{k,n} = (1 + \frac{\rho q \lambda}{\Gamma_n}) u_{k,n}$ . One has that for all  $k \leq n$ ,  $u_{k,n} \xrightarrow[n]{+} +\infty$ . In other words, with the previous inequality, recalling  $r_{k,n}^2 \geq u_{k,n}^2 (1 + \frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2})$ :

$$\begin{aligned} T_{k,l}^M &\leq 2 \exp\left(\frac{q^2 \rho^2 \lambda^2}{\Gamma_n^2} (\gamma_k \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2 - \frac{u_{k,n}^2}{4}) - \frac{u_{k,n}^2}{4}\right) \\ &\leq 2 \exp\left(-C \frac{q^2 \rho^2 \lambda^2}{\Gamma_n^2} \frac{u_{k,n}^2}{8} - \frac{u_{k,n}^2}{4}\right) \leq 2 \exp(-C r_{k,n}^2), \end{aligned} \quad (4.21)$$

since  $u_{k,n}^2 > 16 \gamma_1 \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2$  (which explains our choice for the constant  $r_{k,n}$  in (3.40)). Plugging (4.20) and (4.21) into (4.17) yields:

$$\begin{aligned} &\mathbb{E}\left[\exp\left(-\frac{\rho q \lambda}{\Gamma_n} \Delta_k(X_{k-1}, U_k)\right) \middle| \mathcal{F}_{k-1}\right] \\ &\leq \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} (\gamma_k |\sigma_{k-1}^* \nabla \varphi_{k-1}|^2 + C \gamma_k^{3/2} r_{k,n}^2 + C_{(3.8)} \gamma_k^2 V_{k-1}) + \frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right)\right) \\ &\quad + 2 \exp(-C r_{k,n}^2) \\ &\leq \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} (\gamma_k |\sigma_{k-1}^* \nabla \varphi_{k-1}|^2 + C \gamma_k^{3/2} r_{k,n}^2 + C_{(3.8)} \gamma_k^2 V_{k-1}) + \frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right)\right) \\ &\quad \times (1 + 2 \exp(-C r_{k,n}^2)) \\ &=: \exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} \gamma_k |\sigma_{k-1}^* \nabla \varphi_{k-1}|^2 + C_{(3.8)} \gamma_k^2 V_{k-1}\right) \times \mathfrak{N}_{k,n}(\lambda, \gamma_k, r_{k,n}), \end{aligned}$$

where

$$\mathfrak{N}_{k,n}(\lambda, \gamma_k, r_{k,n}) := \left(1 + 2 \exp(-C r_{k,n}^2)\right) \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right)\right) \exp\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} C \gamma_k^{3/2} r_{k,n}^2\right).$$

We have thus isolated the ‘‘significant’’ term  $\exp\left(\frac{\rho^2 q^2 \lambda^2}{2\Gamma_n^2} \gamma_k |\sigma_{k-1}^* \nabla \varphi_{k-1}|^2\right)$ . The result follows from the definition of  $T_k$  in (3.30).

**Proof of Lemma 6** We recall for convenience the definition of  $r_{k,n}$  in (3.40):

$$r_{k,n} = \begin{cases} r_n = C(1 + \frac{\rho q \lambda}{\Gamma_n}) (\frac{\Gamma_n}{\Gamma_n^{(3/2)}})^{1/4}, & \text{for } \theta \in (\frac{1}{3}, 1), \\ C((1 + \frac{\rho q \lambda}{\Gamma_n}) \ln(n+1))^{1/4} \ln(k+1)^{1/2}, & \text{for } \theta = 1. \end{cases}$$

From the above definition of  $\mathfrak{N}_{k,n}(\lambda, \gamma_k, r_{k,n})$ , we introduce two remainders:

$$\begin{aligned} \mathcal{R}_n^1 &:= \prod_{k=1}^n \left(1 + 2 \exp(-C r_{k,n}^2)\right) \times \prod_{k=1}^n \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right)\right) \\ &=: \mathcal{R}_n^{11} \times \mathcal{R}_n^{12}, \\ \mathcal{R}_n^2 &:= \exp\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} C \sum_{k=1}^n \gamma_k^{3/2} r_{k,n}^2\right), \end{aligned} \quad (4.22)$$



that naturally appear when we iterate Lemma 5. Precisely:

$$\mathbb{E}[S_n] \leq \mathbb{E}[S_{n-1}] \mathfrak{N}_{n,n}(\lambda, \gamma_n, r_{n,n}) \leq \prod_{k=1}^n \mathfrak{N}_{k,n}(\lambda, \gamma_k, r_{k,n}) = \mathcal{R}_n^1 \times \mathcal{R}_n^2. \quad (4.23)$$

- For  $\theta \in (1/3, 1)$ :

We have chosen  $r_{k,n} = r_n$  in (3.40), so,

$$\mathcal{R}_n^1 = \left(1 + 2 \exp(-Cr_n^2)\right)^n \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \Gamma_n^{(1/2)} \exp\left(-\frac{r_n^2}{4}\right)\right) = \mathcal{R}_n^{11} \mathcal{R}_n^{12},$$

and:

$$\begin{aligned} \mathcal{R}_n^{11} &= \left(1 + 2 \exp(-Cr_n^2)\right)^n \leq \exp\left(n \ln\left(1 + 2 \exp\left(-C\left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2}\right)\right)\right) \\ &\leq \exp\left(2n \exp\left(-C\left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2}\right)\right) \xrightarrow{n \rightarrow +\infty} 1, \end{aligned}$$

as, for  $\theta \in (1/3, 1)$ ,  $\left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2} \geq \eta(n)$  where

$$\eta(n) := C\left(n^{(1-\theta)/2} \mathbf{1}_{\theta \in (2/3, 1)} + (n^{1/3} \ln(n)^{-1})^{1/2} \mathbf{1}_{\theta = 2/3} + n^{\theta/4} \mathbf{1}_{\theta \in (1/3, 2/3)}\right), \quad (4.24)$$

see also Remark 8. For the remaining of the proof, we will thoroughly exploit this kind of arguments. Precisely, we recall that:

$$\forall \zeta \in \mathbb{R}_+, \exists C_\zeta \geq 1, \text{ s.t. } \forall 0 \leq \beta \leq \zeta, \forall x \in \mathbb{R}_+, x^\beta \exp(-x^2) \leq C_\zeta \exp(-C_\zeta^{-1} x^2). \quad (4.25)$$

Thus, for the remaining term  $\mathcal{R}_n^{12}$  in  $\mathcal{R}_n^1$ , exploiting (4.25), up to a modification of the constant  $C > 0$  from line to line:

$$\begin{aligned} \mathcal{R}_n^{12} &\leq \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \Gamma_n^{(1/2)} \exp\left(-\frac{r_n^2}{4}\right)\right) \\ &\leq \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \Gamma_n^{(1/2)} \exp\left(-\frac{\rho^2 q^2 \lambda^2 C}{\Gamma_n^2} \left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2}\right) \exp\left(-\frac{C}{4} \left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2}\right)\right) \\ &\leq \exp\left(C \Gamma_n^{(1/2)} \left(\frac{\Gamma_n^{(3/2)}}{\Gamma_n}\right)^{1/4} \exp\left(-\frac{\rho^2 q^2 \lambda^2 C}{\Gamma_n^2} \left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2}\right) \exp\left(-\frac{C}{4} \left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2}\right)\right) \\ &\leq \exp\left(C \Gamma_n^{(1/2)} \exp\left(-\frac{C}{8} \left(\frac{\Gamma_n}{\Gamma_n^{(3/2)}}\right)^{1/2}\right)\right) \leq \exp\left(C n^{1-\theta/2} e^{-\eta(n)}\right) \xrightarrow[n \rightarrow +\infty]{(4.24), (4.25)} 1. \end{aligned}$$

Hence  $\mathcal{R}_n^1 = \mathcal{R}_n^{11} \mathcal{R}_n^{12} \xrightarrow[n]{} 1$ .

Introducing

$$e_n^{\theta < 1} := C \sqrt{\frac{\Gamma_n^{(3/2)}}{\Gamma_n}} \xrightarrow[n \rightarrow +\infty]{} 0,$$

we get from the definition of  $r_{k,n}$  in (3.40) and (4.22)

$$\mathcal{R}_n^2 \leq \exp\left(\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n} + \frac{\rho^4 q^4 \lambda^4}{\Gamma_n^3}\right) e_n^{\theta < 1}\right).$$

Recalling, from (4.23), that  $\mathbb{E}[S_n] \leq \mathcal{R}_n^1 \mathcal{R}_n^2$ , we thus get from the above computations:

$$\mathbb{E}[S_n] \leq \mathcal{R}_n \exp\left(\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n} + \frac{\rho^4 q^4 \lambda^4}{\Gamma_n^3}\right)e_n\right),$$

which gives the result for  $\theta \in (\frac{1}{3}, 1)$ .

- For  $\theta = 1$ , using the previous notations:

$$\begin{aligned} 0 &\leq \ln(\mathcal{R}_n^{11}) \leq \ln\left(\prod_{k=1}^n (1 + 2 \exp(-Cr_{k,n}^2))\right) = \sum_{k=1}^n \ln\left(1 + 2 \exp(-Cr_{k,n}^2)\right) \\ &\leq C \sum_{k=1}^n \exp(-Cr_{k,n}^2) \leq C \sum_{k=1}^n \exp(-C \ln(n+1)^{1/2} \ln(k+1)) \\ &= C \sum_{k=1}^n (k+1)^{-C \ln(n+1)^{1/2}} \leq C 2^{-C \ln(n+1)^{1/2}} + C \int_2^n x^{-C \ln(n+1)^{1/2}} dx \\ &= C 2^{-C \ln(n+1)^{1/2}} + C \frac{2^{1-C \ln(n+1)^{1/2}} - n^{1-C \ln(n+1)^{1/2}}}{C \ln(n+1)^{1/2} - 1} \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

We deduce that:

$$\mathcal{R}_n^{11} = \prod_{k=1}^n \left(1 + 2 \exp(-Cr_{k,n}^2)\right) \xrightarrow{n \rightarrow +\infty} 1.$$

Like for the case  $\theta \in (1/3, 1)$ , we get the control:

$$\begin{aligned} \mathcal{R}_n^{12} &\leq \prod_{k=1}^n \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{r_{k,n}^2}{4}\right)\right) \\ &\leq \prod_{k=1}^n \exp\left(\frac{\rho q \lambda}{\Gamma_n} C \gamma_k^{1/2} \exp\left(-\frac{C}{4} \left[1 + \left(\frac{\rho q \lambda}{\Gamma_n}\right)^2\right] \ln(n+1)^{1/2} \ln(k+1)\right)\right) \\ &\stackrel{(4.25)}{\leq} \prod_{k=1}^n \exp\left(C \gamma_k^{1/2} \exp\left(-\frac{C}{4} \left[1 + C \left(\frac{\rho q \lambda}{\Gamma_n}\right)^2\right] \ln(n+1)^{1/2} \ln(k+1)\right)\right) \\ &\leq \prod_{k=1}^n \exp\left(C \frac{1}{k^{1/2}} (k+1)^{-C \ln(n+1)^{1/2}}\right) \xrightarrow{n \rightarrow +\infty} 1, \end{aligned}$$

in fact, we have already established for the control of  $\mathcal{R}_n^{11}$  that  $\sum_{k=1}^n (k+1)^{-C \ln(n+1)^{1/2}} \xrightarrow{n} 0$ .

Thus, we proved that  $\mathcal{R}_n^1 = \mathcal{R}_n^{11} \mathcal{R}_n^{12} \xrightarrow{n} 1$ .

Let us now turn to the other contribution in (4.22):

$$\begin{aligned} \mathcal{R}_n^2 &= \exp\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} C \sum_{k=1}^n \gamma_k^{3/2} r_{k,n}^2\right) \\ &\leq \exp\left(2C \left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n^2} + \frac{\rho^4 q^4 \lambda^4}{\Gamma_n^4}\right) \ln(n+1)^{1/2} \sum_{k=1}^n \gamma_k^{3/2} \ln(k+1)\right) \\ &\leq \exp\left(2C \left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n} + \frac{\rho^4 q^4 \lambda^4}{\Gamma_n^3}\right) \frac{\sum_{k=1}^n \gamma_k^{3/2} \ln(k+1)}{\sqrt{\Gamma_n}}\right) \\ &= \exp\left(\left(\frac{\rho^2 q^2 \lambda^2}{\Gamma_n} + \frac{\rho^4 q^4 \lambda^4}{\Gamma_n^3}\right) e_n^{\theta=1}\right), \end{aligned}$$

with  $e_{1,n}^{\theta=1} := 2C \frac{1}{\sqrt{\Gamma_n}} \sum_{k=1}^n \frac{\ln(k+1)}{k^{3/2}} \xrightarrow{n \rightarrow +\infty} 0$ , and recalling that  $\sqrt{\Gamma_n} \asymp \ln(n+1)^{1/2}$  for the last inequality.

To sum up, for all  $\theta \in (\frac{1}{3}, 1]$ , thanks to inequality (4.23), we know that there exist non negative sequences  $(\mathcal{R}_n)_{n \geq 1}$ ,  $(e_n)_{n \geq 1}$ , s.t.  $\mathcal{R}_n \xrightarrow[n]{} 1$ ,  $e_n \xrightarrow[n]{} 0$ , and

$$\mathbb{E}[S_n] \leq \mathcal{R}_n \exp \left( \left( \frac{\rho^2 q^2 \lambda^2}{\Gamma_n} + \frac{\rho^4 q^4 \lambda_n^4}{\Gamma_n^3} \right) e_n \right).$$

## 5 Regularity Results and Consequences

This section is devoted to some regularity results for the Poisson problem

$$\mathcal{A}\varphi = f - \nu(f). \quad (5.1)$$

In particular, we state below some Schauder like controls, which are, because of our methodology that requires pointwise controls of the derivatives, more adapted than the standard Sobolev estimates (see e.g. Pardoux and Veretennikov [PV01]).

There are two kinds of assumptions that guarantee the solution  $\varphi$  of (5.1) enjoys the required smoothness of Theorems 3.

- (a) If  $b$ ,  $\sigma$  and  $f$  are *smooth*, under suitable confluence like conditions stated below in  $(\mathbf{D}_\alpha^p)$  with the condition on  $D\sigma$ :  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$  for given  $(\alpha, p) \in (0, +\infty) \times [1, 2)$ , the probabilistic representation  $\varphi(x) = -\int_0^{+\infty} \mathbb{E}[(f(Y_t^{0,x}) - \nu(f))] dt$  of the solution of (5.1) can be differentiated using iterated tangent flows to establish that  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ , for some  $\beta \in (0, 1]$ . It suffices for that to have  $f \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  and  $b \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ ,  $\sigma \in \mathcal{C}_b^{3,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ . We refer to Section 2.2 and 5 of Chapter 3 for additional details. Importantly, under such assumptions, specific non-degeneracy conditions are not needed.
- (b) If we do not assume such an *a priori* smoothness on  $f, b, \sigma$ , we need to make an *extra* non-degeneracy assumption which will allow some regularity gain (elliptic bootstrap) and a stronger confluence like condition  $(\mathbf{D}_\alpha^p)$  with  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(3+\beta)-p}$  for given  $(\alpha, p) \in (0, +\infty) \times [1, 2)$ . Precisely, assuming that the bounded diffusion coefficient  $\Sigma$  is also uniformly elliptic, we exploit the results of Krylov and Priola [KP10] (Theorems 2.4-2.6) to derive that, up to an additional technical condition on  $\Sigma$  when  $d > 1$ , we actually have  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  for some  $\beta \in (0, 1)$ , as soon as  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$  and  $b \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ ,  $\sigma \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ .

We point out that the second set of assumptions is very important in order to go towards the *usual* setting of functional/transport inequalities which typically involves Lipschitz continuous test functions. This is for instance the case for the controls of the Wasserstein distances between the law of  $Y_t$  in (1.1) and the invariant distribution  $\nu$  (see e.g. [BGL14]). We are thus able, in the non-degenerate framework (b), to consider directly sources  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$ . Observe that, when  $\beta \rightarrow 0$ , we are *almost* Lipschitz. Actually, we can, in this setting, handle functions  $f \in \mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R})$  up to a spatial regularization which leads to a constraint on the steps (see Theorem 5 below).

## 5.1 Assumptions and Regularity Results

The regularity results stated here can be found in Section 5 of Chapter 3. Let us now recall the useful assumptions needed.

**(UE)** Uniform Ellipticity. We assume that  $r \geq d$  in (1.1) and that there is  $\underline{\sigma} > 0$  such that

$$\forall \xi \in \mathbb{R}^d, \langle \sigma \sigma^*(x) \xi, \xi \rangle \geq \underline{\sigma} |\xi|^2.$$

For  $\beta \in (0, 1)$ , we introduce the following condition.

**(R<sub>1,β</sub>)** Regularity Condition. From equation (1.1), we suppose  $b \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$ , and  $\sigma \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^r)$ .

**(D<sub>α</sub><sup>p</sup>)** Confluence Conditions. We assume that there exists  $\alpha > 0$  and  $p \in [1, 2)$  such that for all  $x \in \mathbb{R}^d, \xi \in \mathbb{R}^d$

$$\left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r \left( (p-2) \frac{|\langle D\sigma_{\cdot j}(x) \xi, \xi \rangle|^2}{|\xi|^2} + |D\sigma_{\cdot j} \xi|^2 \right) \leq -\alpha |\xi|^2, \quad (5.2)$$

where  $Db$  stands here for the Jacobian of  $b$ ,  $\sigma_{\cdot j}$  stands for the  $j^{\text{th}}$  column of the diffusion matrix  $\sigma$  and  $D\sigma_{\cdot j}$  for its Jacobian matrix.

There are others assumptions than **(D<sub>α</sub><sup>p</sup>)** which yield, in the non-degenerate setting, gradient control. This is the case for the so-called Bakry and Émery curvature criterion ([BE85, BGL14]) which is however pretty hard to check for general multidimensional diffusion coefficients. However for Hölder control of the gradient, this criterion seems to be not adapted, see Chapter 3 Section 2.2 for more details.

We eventually introduce, as in Chapter 3, a technical condition on the diffusion coefficient  $\sigma$ . It allows to prove that each partial derivative  $\partial_{x_i} \varphi$  of the solution of (5.1) satisfies an autonomous scalar Poisson problem. We suppose:

$$(\Sigma) \text{ for every } (i, j) \in \llbracket 1, d \rrbracket^2 \text{ and } x = (x_1, \dots, x_d) \in \mathbb{R}^d, \Sigma_{i,j}(x) = \Sigma_{i,j}(x_{i \wedge j}, \dots, x_d).$$

We say that assumption **(P<sub>β</sub>)** is satisfied if **(UE)**, **(D<sub>α</sub><sup>p</sup>)** with  $\|D\sigma\|_\infty^2 \leq \frac{2\alpha}{2(1+\beta)-p}$ , **(R<sub>1,β</sub>)**, and **(Σ)** are in force. From Section 5.3 of Chapter 3 we have the following result.

**Theorem 4 (Elliptic Bootstrap in a non-degenerate setting)** *Assume **(P<sub>β</sub>)** holds for some  $\beta \in (0, 1)$  and that  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$ . Then, there is a unique  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  solving (5.1).*

Note as well from Remark 2, that there exists  $C > 0$ ,  $|D^2\varphi(x)| \leq C(1 + |x|)^{-1}$ . In other words, the solution  $\varphi$  of (5.1) satisfies **(T<sub>β</sub>)**.

**Remark 10 (On Schauder estimates for  $\beta = 1$ )** *We insist on the fact that  $\beta \in (0, 1)$  in the above theorem. Indeed, it is well known that the Hölder exponent  $\beta$  cannot go to 1 in the Schauder estimates. Note that, for the particular case  $f \in \mathcal{C}^{1,1}(\mathbb{R}^d, \mathbb{R})$ , we also have  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$  for all  $\beta \in (0, 1)$ . This means that, for such  $f$ , the elliptic bootstrap works up to an arbitrarily small correction.*

From Theorem 4 we readily have:

**Corollary 1 (Regularity for the Poisson problem with *carré du champ* source)**

Assume  $(\mathbf{P}_\beta)$  holds for some  $\beta \in (0, 1)$  and that  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$ . Then, there is a unique  $\vartheta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  solving

$$\mathcal{A}\vartheta = |\sigma^*\nabla\varphi|^2 - \nu(|\sigma^*\nabla\varphi|^2)$$

and satisfying  $(\mathbf{T}_\beta)$ .

Indeed, it suffices to observe from Theorem 4 and the assumption  $(\mathbf{R}_{1,\beta})$  in  $(\mathbf{P}_\beta)$  that  $\tilde{f} := |\sigma^*\nabla\varphi|^2 \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$  and to apply again Theorem 4 for this source.

## 5.2 Concentration inequalities for a Lipschitz source in a non-degenerate setting

As indicated in the introduction of the Section, we aim at controlling deviations for Lipschitz sources. In the current Lipschitz framework we aim to address, we need a slightly different set of assumptions. Namely, we will assume that  $(\mathbf{C1})$ ,  $(\mathbf{GC})$ ,  $(\mathbf{C2})$ ,  $(\mathcal{L}_\nu)$ ,  $(\mathbf{U})$ ,  $(\mathbf{S})$  and  $(\mathbf{P}_\beta)$  are in force and we will say that  $(\mathbf{L}_\beta)$  holds. Under this new assumption, we have the following result.

**Theorem 5 (Non-asymptotic concentration bounds in a Lipschitz framework)**

Assume that  $(\mathbf{L}_\beta)$  is in force. Let  $f$  be a Lipschitz continuous function. For a time step sequence  $(\gamma_k)_{k \geq 1}$  of the form  $\gamma_k = k^{-\theta}$ ,  $\theta \in (1/2, 1]$ , we have that, there exist two explicit monotonic sequences  $c_n \leq 1 \leq C_n$ ,  $n \geq 1$ , with  $\lim_n C_n = \lim_n c_n = 1$  such that for all  $n \geq 1$  and for every  $a > 0$ :

$$\mathbb{P}[|\sqrt{\Gamma_n}(\nu_n(f) - \nu(f))| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\nu(|\sigma^*\nabla\varphi|^2)}\right), \quad (5.3)$$

where  $\varphi \in \mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R}) \cap W_{2,loc}^2(\mathbb{R}^d, \mathbb{R})$  is a weak solution of the Poisson equation  $\mathcal{A}\varphi = f - \nu(f)$ .

**Sketch of the proof** To prove the above result, the starting point consists in regularizing the source  $f$  by mollification. Namely, we consider  $f_\delta = f \star \eta_\delta$ , where  $\star$  denotes the usual convolution, for a suitable mollifier  $\eta_\delta(\cdot) := \frac{1}{\delta^d} \eta(\frac{\cdot}{\delta})$ ,  $\delta > 0$ , where  $\eta$  is a compactly supported non-negative function s.t.  $\int_{\mathbb{R}^d} \eta(x) dx = 1$ . We then write  $\nu_n(f) - \nu(f) = \nu_n(f_\delta) - \nu(f_\delta) + (\nu_n - \nu)(f - f_\delta) =: \nu_n(f_\delta - \nu(f_\delta)) + R_{n,\delta}$ . We aim at letting  $\delta$  go to 0 so that  $R_{n,\delta}$  can be viewed as a remainder. On the other hand, we will apply the same strategy as in the proof of Theorem 3 to analyze the deviations of  $\nu_n(f_\delta - \nu(f_\delta)) = \nu_n(\mathcal{A}\varphi_\delta)$ . Precisely, reproducing the arguments of Section 5.3 of Chapter 3 to equilibrate the explosions of the derivatives of  $\varphi_\delta$  in the proof of Theorem 3 yields that there exists two explicit monotonic sequences  $\tilde{c}_n \leq 1 \leq C_n$ ,  $n \geq 1$ , with  $\lim_n C_n = \lim_n \tilde{c}_n = 1$  s.t.

$$\mathbb{P}[|\sqrt{\Gamma_n}(\nu_n(f) - \nu(f))| \geq a] \leq 2C_n \exp\left(-\tilde{c}_n \frac{a^2}{2\nu(|\sigma^*\nabla\varphi_\delta|^2)}\right). \quad (5.4)$$

From the previous Schauder estimates, we know that  $\varphi_\delta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  for all  $\delta > 0$  with explosive  $\mathcal{C}^{3,\beta}$  norm in  $\delta$  but with bounded gradient. Recall indeed that, for all  $\beta \in (0, 1)$ ,

$$\|f_\delta\|_{\mathcal{C}^{1,\beta}} \leq C\delta^{-\beta}, \quad |\nabla\varphi_\delta| \leq \frac{[f_\delta]_1}{\alpha} = \frac{[f]_1}{\alpha}. \quad (5.5)$$

We again refer to Lemma 6 and Section 5.3 in Chapter 3 for details. On the other hand, it is well known, see e.g. [PV01], that  $\varphi_\delta(x) = -\int_0^{+\infty} \mathbb{E}[f_\delta(Y_t^{0,x}) - \nu(f_\delta)]dt$ . From their Proposition 1, we have in our case that, denoting by  $\nu_{Y_t^{0,x}}$  the law of  $Y_t^{0,x}$  we have the following control for the total variation between  $\nu_{Y_t^{0,x}}$  and  $\nu$ . There exists constants  $(C, c) := (C, c)((\mathbf{L}_\beta))$  s.t.

$$\|\nu_{Y_t^{0,x}} - \nu\|_{T.V.} \leq C \exp(c|x|) \exp(-\alpha t). \quad (5.6)$$

Introducing now  $\bar{f}_\delta = f_\delta - f - \nu(f_\delta - f)$ , we rewrite that, for all  $x \in \mathbb{R}^d$ :

$$\begin{aligned} (\varphi_\delta - \varphi)(x) &= -\int_0^{+\infty} \mathbb{E}[\bar{f}_\delta(Y_t^{0,x})]dt, \\ |(\varphi_\delta - \varphi)(x)| &\leq \int_0^{+\infty} \left( \int_{\mathbb{R}^d} |\bar{f}_\delta(y)|^2 (\nu_{Y_t^{0,x}} + \nu)(dy) \right)^{1/2} \|\nu_{Y_t^{0,x}} - \nu\|_{T.V.}^{1/2} dt \\ &\stackrel{(5.6)}{\leq} C^{1/2} \exp\left(\frac{c}{2}|x|\right) \|\bar{f}_\delta\|_\infty \int_0^{+\infty} \exp\left(-\frac{\alpha}{2}t\right) dt. \end{aligned}$$

Recalling that  $f$  is Lipschitz and that  $(f - f_\delta)(x) = \int_{\mathbb{R}^d} (f(x-y) - f(x))\eta_\delta(y)dy$ , we actually have  $\|\bar{f}_\delta\|_\infty \leq C[f]_1\delta$ , which establishes the pointwise convergence  $\varphi_\delta(x) \xrightarrow{\delta \rightarrow 0} -\int_0^{+\infty} \mathbb{E}[f(Y_t^{0,x}) - \nu(f)]dt =: \varphi(x)$  which is the only weak solution of (5.1) in  $W_{p,loc}^2(\mathbb{R}^d, \mathbb{R})$ ,  $p > 1$  (see Theorem 1 in [PV01]).

Let us now prove that

$$\lim_{\delta \rightarrow 0} \nu(|\sigma^* \nabla \varphi_\delta|^2) = \nu(|\sigma^* \nabla \varphi|^2). \quad (5.7)$$

For all  $\varepsilon > 0$ , there is a compact  $K := K(\varepsilon)$  such that, denoting by  $K^c := \mathbb{R}^d \setminus K$ , we have:

$$\int_{\mathbb{R}^d} |\nabla \varphi_\delta - \nabla \varphi|^2(x) \mathbf{1}_{K^c}(x) \nu(dx) \leq 4 \|\nabla \varphi\|_\infty^2 \int_{\mathbb{R}^d} \mathbf{1}_{K^c}(x) \nu(dx) \leq \frac{\varepsilon}{2}.$$

Also, since  $\varphi \in W_{2,loc}^2(\mathbb{R}^d, \mathbb{R})$ , we write:

$$\begin{aligned} \int_{\mathbb{R}^d} |\nabla \varphi_\delta - \nabla \varphi|^2(x) \mathbf{1}_K(x) \nu(dx) &\leq \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} (\nabla \varphi(x-z) - \nabla \varphi(x)) \rho_\delta(z) dz \right|^2 \mathbf{1}_K(x) \nu(dx) \\ &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \left( \int_0^1 D^2 \varphi(x - \lambda z) z d\lambda \right) \rho_\delta(z) dz \right|^2 \mathbf{1}_K(x) \nu(dx) \\ &\leq \int_0^1 d\lambda \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} D^2 \varphi(x - \lambda z) z \rho_\delta(z) dz \right|^2 \mathbf{1}_K(x) \nu(dx) \\ &\leq \int_0^1 d\lambda \left( \int_{\mathbb{R}^d} dz |z|^2 \rho_\delta(z) \int_{\mathbb{R}^d} |D^2 \varphi(x - \lambda z)|^2 \mathbf{1}_K(x) \nu(dx) \right) \leq C\delta^2 \|\varphi\|_{W_2^2(\bar{K}, \mathbb{R})}^2 < \frac{\varepsilon}{2}, \end{aligned}$$

for  $\delta$  small enough, using the Cauchy-Schwarz inequality for the penultimate control and denoting by  $\bar{K}$  a compact set such that for all  $z \in B(0, C\delta) \supset \text{supp}(\eta_\delta)$ ,  $x \in K$ ,  $x - z \in \bar{K}$ . This in particular gives (5.7). Hence, setting

$$c_n := \tilde{c}_n \frac{\nu(|\sigma^* \nabla \varphi|^2)}{\nu(|\sigma^* \nabla \varphi_\delta|^2)} \rightarrow_n 1,$$

and recalling from Section 5.3 in Chapter 3 that  $\delta := \delta(n) \rightarrow_n 0^\ddagger$ , we derive that (5.3) follows from (5.4) up to a modification of  $\tilde{c}_n$ .

Furthermore, let us point out that, the result can alternatively be stated replacing the *carré du champ* in (5.3) by the variance of the Lipschitz source under the invariant law. In fact, we can write by the dominated convergence theorem:

$$\lim_{\delta \rightarrow 0} \nu(|\sigma^* \nabla \varphi_\delta|^2) = \nu(|\sigma^* \nabla \varphi|^2). \quad (5.8)$$

**Remark 11** *The new threshold  $\theta > \frac{1}{2}$  comes from the specific Lipschitz regularity of the test function  $f$ . Intuitively, this threshold naturally appears when we consider  $\beta \rightarrow 0$  in the previous condition  $\theta \in (\frac{1}{2+\beta}, 1]$  induced by the regularity of  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$  which holds, under  $(\mathbf{P}_\beta)$ , when  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$ . We underline anyhow that, for  $\beta = 0$ , the Schauder estimates do not directly apply.*

## 6 Optimisation over $\rho$ under *Gaussian* and *super Gaussian* deviations

In Lemma 4, we performed an asymptotic estimation of the upper-bound for *Gaussian deviations*. However, from a numerical point of view, it appears to be more significant to optimize over  $\rho$  in whole generality, i.e. not only when  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ . This procedure conducts to deviations bounds that are much closer to the realizations.

In particular, in *super Gaussian deviations* framework (i.e.  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ ), we provide here a “weaker” concentration inequality than the Gaussian one, which precisely comes from the optimization over  $\rho$  for this regime. This loss of concentration, with the terminology of Remark 4, is intrinsic to our method, as it will be shown in the proofs of Theorem 6 and Lemma 7 below, see also Remark 17.

**Theorem 6 (Deviations in the *super Gaussian* regime)** *Assume  $(\mathbf{A})$  is in force. If there exists  $\vartheta \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ ,  $\beta \in (0, 1]$  satisfying  $(\mathbf{T}_\beta)$  s.t.*

$$\mathcal{A}\vartheta = |\sigma^* \nabla \varphi|^2 - \nu(|\sigma^* \nabla \varphi|^2), \quad (6.9)$$

*then, for  $\theta \in (\frac{1}{2+\beta}, 1]$ , there exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$  s.t.*

---

‡. which was anyhow constrained to go to 0 sufficiently *slowly* in order to balance the explosions in the derivatives coming from the Schauder estimates, see (5.5). It is specifically this feature that led to the condition  $\gamma_n \asymp n^{-\theta}$ ,  $\theta \in (1/2, 1]$ .

for all  $n \geq 1$ ,  $a > 0$ , the following bounds hold. When  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$  (Super Gaussian deviations):

$$\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^{4/3}\Gamma_n^{1/3}}{2\|\sigma\|_\infty^{2/3}[\vartheta]_1^{2/3}}\right).$$

**Remark 12** Observe from Corollary 1, that, the function  $\vartheta$  enjoys the required smoothness as soon as assumption  $(\mathbf{P}_\beta)$  introduced in Section 5 holds.

**Remark 13** For super Gaussian deviations, we obtain a sharper bound than in Theorem 2. Nonetheless, asymptotically, this regime is less sharp than Theorem 2 in Chapter 3 which provides a Gaussian bound with deteriorated constants (see also the User's guide to the proof in Section 3.1 below). Even if, from a numerical point of view, the deviation bounds in Theorem 7 below yield sharper controls with respect to simulated empirical measures (see Figure 4.1 in the numerical Section below).

For “intermediate Gaussian deviations”, i.e. for  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow_n C > 0$ , the constants in the Gaussian bound deteriorate. So, it seems reasonable to see this situation like for the first regime  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ , namely where there are constants  $C_\infty > 1, c_\infty < 1$  such that  $\lim_n C_n = C_\infty$ ,  $\lim_n c_n = c_\infty$  and

$$\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\nu(|\sigma^*\nabla\varphi|^2)}\right).$$

Observe that for such regimes, there is an equivalence, up to multiplicative constants, between the bounds in Theorem 3 and in Theorem 6.

The idea of the proof of Theorem 6 follows the same lines as for Theorem 3, except for the optimization over  $\rho$  which is more fussy, see Lemma 7 below.

We recall that the analysis in the proof of Theorem 3 leaving open a possible optimization over the parameter  $\rho$  which we now perform. The next lemma indicates that this optimization implies a Gaussian regime for  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$  and a super Gaussian one for  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$ .

**Lemma 7 (Choice of  $\rho$  for the concentration regime)** For  $P_{\min}(a, \Gamma_n, \rho)$  as in (3.34), we have

(a) If  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ , taking  $\rho := \rho(a, n)$  s.t.  $\rho - 1 = \frac{1}{2} \frac{\tilde{B}_n^{1/2} a}{\tilde{A}_n^{3/2} \sqrt{\Gamma_n}} (1 + o(1))$

$$P_{\min}(a, \Gamma_n, \rho) \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0}{=} -\frac{a^2}{2\nu(|\sigma^*\nabla\varphi|^2)} (1 + o(1)).$$

(b) If  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$ , taking  $\rho := \rho(a, n)$  s.t.  $\rho - 1 = \frac{1}{2} + o(1)$

$$P_{\min}(a, \Gamma_n, \rho) \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty}{=} -\frac{a^{4/3}\Gamma_n^{1/3}}{2\|\sigma\|_\infty^{2/3}[\vartheta]_1^{2/3}} (1 + o(1)).$$



**Remark 14** For Gaussian deviations,  $\rho - 1 = \frac{1}{2} \frac{\tilde{B}_n^{1/2} a}{\tilde{A}_n^{3/2} \sqrt{\Gamma_n}} (1 + o(1)) \asymp \frac{a}{\sqrt{\Gamma_n}}$  which corresponds to our choice in Lemma 4. We then retrieve the Gaussian regime. In the super Gaussian deviations framework, the optimization over  $\rho$  leads to consider  $\rho - 1 = \frac{1}{2} + o(1)$  which yields the loss in the concentration inequality.

Let us first continue with the proof of Lemma 7 which is purely analytical and rather independent of our probabilistic setting.

### Proof of Lemma 7

We keep the notations of Lemma 4, that we bring to mind.

$$\begin{aligned} \Phi_n(a, \rho) &= \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} + \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho - 1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} + \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} - \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho - 1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} \\ &= \frac{a^{1/3}}{\tilde{B}_n^{1/3} \Gamma_n^{1/6}} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right), \end{aligned}$$

for

$$\rho - 1 := \xi \frac{27 \tilde{B}_n a^2}{8 \tilde{A}_n^3 \Gamma_n}, \quad (6.10)$$

where  $\xi := \xi(a, n) > 0$  is a parameter that we are going to optimize. Furthermore:

$$P(\lambda_n) = -\frac{3^2}{2^4 \tilde{A}_n} a^2 f_\Psi(\xi), \quad (6.11)$$

for

$$f_\Psi : \xi \in \mathbb{R}_+ \mapsto \frac{g(\xi)}{\Psi \xi + 1}, \quad (6.12)$$

where

$$g : \xi \mapsto \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \left( 1 - \frac{\xi^{1/3}}{2} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \right),$$

and

$$\Psi := \frac{27 \tilde{B}_n a^2}{8 \tilde{A}_n^3 \Gamma_n} \left( \stackrel{(4.9)}{=} \frac{\rho - 1}{\xi} \right). \quad (6.13)$$

• Let us first focus on case (a), ‘‘Gaussian deviations’’ ( $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ ). We, now anyhow, want to maximize  $\Lambda$  in  $\xi$  to obtain the best possible concentration bound. Let  $\mathcal{A} := \{\xi \in [0, +\infty] : f_\Psi(\xi) = \|f_\Psi\|_\infty\}$  be the set of points where  $f_\Psi$  reaches its maximum. Observe that for a fixed  $\Psi$ ,  $f_\Psi(\xi) \xrightarrow{\xi \rightarrow \infty} 0$ . Thus,  $+\infty \notin \mathcal{A}$ . Let now  $\xi_*$  be an arbitrary point in  $\mathcal{A}$ .

From the smoothness of  $f_\Psi$ , the optimality condition writes:

$$f'_\Psi(\xi_*) = \frac{g'(\xi_*)}{(\Psi \xi_* + 1)} - \frac{\Psi g(\xi_*)}{(\Psi \xi_* + 1)^2} = 0 \Leftrightarrow \frac{g'(\xi_*)}{(\Psi \xi_* + 1)} = \frac{\Psi f_\Psi(\xi_*)}{(\Psi \xi_* + 1)} \Leftrightarrow f_\Psi(\xi_*) = \frac{g'(\xi_*)}{\Psi}. \quad (6.14)$$

Recall now that we want to maximize over the  $\xi$  s.t.  $\xi \rightarrow +\infty$ ,  $\xi\Psi \xrightarrow{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0} 0$ . Indeed, from the proof of Lemma 4, we saw that for such a choice, we obtain the expected Gaussian concentration, namely  $P(\lambda_n) \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0}{\sim} -\frac{a^2}{4\tilde{A}_n}$ .

From the computations of Lemma 8 in Appendix 8, we have:

$$g'(\xi) \underset{\xi \rightarrow +\infty}{=} \frac{8}{3^5 \xi^2} + o\left(\frac{1}{\xi^2}\right). \quad (6.15)$$

So,

$$\Lambda(\xi_*) \underset{\xi_* \rightarrow +\infty}{=} \frac{1}{2 \cdot 3^3 \xi_*^2 \Psi} (1 + o(1)) \xrightarrow{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0} \frac{1}{4}, \quad (6.16)$$

where  $o(1)$  denotes here a quantity going to 0 as  $\xi_* \rightarrow +\infty$  and  $\xi_*\Psi \xrightarrow{\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0} 0$ . Inspired by the identity (6.16), and motivated the numerical simulations (see Section 7), we set

$$\bar{\xi}_* := \frac{2^{1/2}}{3^{3/2} \sqrt{\Psi}} = \frac{2^2 \tilde{A}_n^{3/2} \sqrt{\Gamma_n}}{3^9 \tilde{B}_n^{1/2} a}, \quad (6.17)$$

which indeed satisfies (4.15), so that

$$\min_{\rho > 1} P(\lambda_n) \leq -\frac{a^2}{\tilde{A}_n} \Lambda(\bar{\xi}_*) (1 + o(1)) = \frac{a^2}{4\tilde{A}_n} (1 + o(1)) \underset{n \rightarrow +\infty}{\sim} -\frac{a^2}{2\nu(|\sigma^* \nabla \varphi|^2)}.$$

Observe as well that this choice yields:

$$\rho - 1 = \frac{2^{1/2} \Psi}{3^{3/2} \sqrt{\Psi}} = \frac{1}{2} \frac{\tilde{B}_n^{1/2} a}{\tilde{A}_n^{3/2} \sqrt{\Gamma_n}}. \quad (6.18)$$

- Now we will study the the case (b) “super Gaussian deviations” for which  $\frac{a}{\sqrt{\Gamma_n}} \xrightarrow{n} +\infty$ .

Note that we cannot expect a Gaussian regime in this case. In fact, for  $\Psi$  going to infinity (see definition (6.13)), by (6.14), to get a Gaussian regime at a maximizer  $\xi_*$  of  $\Lambda$ , we have from (6.11) that  $f_\Psi(\xi_*) = \frac{g'(\xi_*)}{\Psi}$  has to remain separated from 0 when  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$ . Since, in this case  $\Psi \xrightarrow{n} +\infty$ , this imposes to consider points  $\xi \rightarrow 0$  in order to exploit the asymptotic behaviour (see again Lemma 8 for more details):

$$g'(\xi) \underset{\xi \rightarrow 0}{=} \frac{2^{1/3}}{3\xi^{2/3}} (1 + o(1)). \quad (6.19)$$

In other words, from (6.19), we expect that there is a constant  $K_* > 0$  s.t.

$$f_\Psi(\xi_*) \underset{\xi_* \rightarrow 0}{\sim} \frac{2^{1/3}}{3\xi_*^{2/3} \Psi} \geq K_*.$$

So  $\xi_* \leq \frac{C}{\Psi^{3/2}} \xrightarrow{\Psi \rightarrow +\infty} 0$ . Now,  $|f_\Psi(\xi_*)| = \left| \frac{g'(\xi_*)}{\Psi} \right| \leq |g'(\xi_*)| \xrightarrow{\xi_* \rightarrow 0} 0$ . This means that it is impossible to stay in a Gaussian regime.

We now still look at the optimal  $\xi_* \rightarrow 0$  which allows to stay at “the biggest possible regime”. Thenceforth, we will estimate  $\xi_*$  directly from the map  $f_\Psi$  defined in (6.12):

$$f_\Psi(\xi) \underset{\xi \rightarrow 0}{=} \frac{2^{1/3}\xi^{1/3}}{\Psi\xi + 1}(1 + o(1)) =: f_{\Psi,0}(\xi)(1 + o(1)). \quad (6.20)$$

It can be directly checked that  $\arg \max_{\xi \in \mathbb{R}_+} f_{\Psi,0}(\xi) = \frac{1}{2\Psi}$ . We therefore get:

$$\xi_* \underset{\Psi \rightarrow +\infty}{=} \frac{1}{2\Psi} + o\left(\frac{1}{\Psi}\right) \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty}{=} \frac{4\tilde{A}_n^3\Gamma_n}{27\tilde{B}_n a^2} + o\left(\frac{\Gamma_n}{a^2}\right) \rightarrow 0. \quad (6.21)$$

From (6.11) and (6.20), we get:

$$\begin{aligned} \min_{\rho > 1} P(\lambda_n) &\underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty}{\leq} -\frac{3^2}{2^4\tilde{A}_n} a^2 \frac{2^{1/3}\xi_*^{1/3}}{\Psi\xi_* + 1}(1 + o(1)) \underset{(6.21)}{=} -\frac{3^2}{2^4} \frac{a^2}{\tilde{A}_n(\frac{1}{2} + 1)} \Psi^{-1/3}(1 + o(1)) \\ &\underset{(4.15)}{\leq} -\frac{3^2}{2^4} \frac{a^2}{\tilde{A}_n(\frac{1}{2} + 1)} \left(\frac{2^3\tilde{A}_n^3\Gamma_n}{3^3\tilde{B}_n a^2}\right)^{1/3}(1 + o(1)) \\ &\underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty}{=} -\frac{a^{4/3}\Gamma_n^{1/3}}{2^2\tilde{B}_n^{1/3}}(1 + o(1)) \underset{(3.31)}{\underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty}{=}} -\frac{a^{4/3}\Gamma_n^{1/3}}{2\|\sigma\|_\infty^{2/3}[\vartheta]_1^{2/3}}(1 + o(1)). \end{aligned}$$

From equations (4.9) and (6.13), the choice (6.21) yields:

$$\rho - 1 = \xi\Psi = \frac{1}{2} + o(1). \quad (6.22)$$

**Remark 15 (Controls of the parameters  $(\lambda, \rho)$  for *super Gaussian deviations*)**

We give here some useful estimates to control the remainder terms in the truncation procedure associated with unbounded innovations, see proof of Lemma 6. They specify the behaviour of the quantity  $\frac{\rho\lambda_n}{\Gamma_n}$ .

In the regime of *super Gaussian deviations*, identities (4.10), (6.13) and the choice (6.21) (i.e.  $\xi_* = \frac{1}{2\Psi}$ ) yields:

$$\frac{\rho^2\lambda_n^2}{\Gamma_n^2} \asymp C\Psi \underbrace{\xi_*^{2/3} \left( (1 + \sqrt{1 + \xi_*})^{1/3} + (1 - \sqrt{1 + \xi_*})^{1/3} \right)^2}_{\asymp \frac{2^{1/3}}{\Psi^{2/3}}} \underset{\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty}{\rightarrow} +\infty.$$

**Remark 16** Theorems 3 and 6 are actually a consequence of the more general following result which has a real importance for numerical applications. Indeed, for a given  $n \in \mathbb{N}$ , we have to control the non-asymptotic error in Theorems 3 and 6. Furthermore, in Section 7 (see Remark 18) we will see that for  $\theta \in (\frac{1}{3}, 1]$ , for “reasonable”  $n$  (e.g.  $n = 5 \cdot 10^4$  in the following Section 7) and a  $(\approx 1)$  we are already “out” of the Gaussian deviations regime, namely  $\frac{a}{\sqrt{\Gamma_n}} \asymp 1$ . This illustration justifies the interest of optimizing over  $\rho$  for Gaussian deviations and *super Gaussian deviations*.

**Theorem 7** *Let the assumptions of Theorem 3 be in force. For  $\theta \in (\frac{1}{3}, 1]$ , there exist explicit non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$ , respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$  s.t. for all  $n \geq 1$  for all  $a > 0$ ,*

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 C_n \exp(c_n P_{\min}(a, \Gamma_n, \rho)),$$

where  $\rho > 1$  and

$$P_{\min}(a, \Gamma_n, \rho) = -\frac{(\rho-1)^{1/3} \sqrt{\Gamma_n} \Phi_n(a, \rho)}{\rho} (3a - \sqrt{\Gamma_n} (\rho-1)^{1/3} \tilde{A}_n \Phi_n(a, \rho)),$$

with

$$\Phi_n(a, \rho) := \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} + \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{1/2} \right)^{1/3} + \left( \frac{a}{\sqrt{\Gamma_n} \tilde{B}_n} - \left( \frac{a^2}{\tilde{B}_n^2 \Gamma_n} + (\rho-1) \left( \frac{2\tilde{A}_n}{3\tilde{B}_n} \right)^3 \right)^{1/2} \right)^{1/3} \quad (6.23)$$

and

$$\tilde{A}_n := \frac{q\nu(|\sigma^* \nabla \varphi|^2)}{2} + e_n \quad \text{and} \quad \tilde{B}_n = \frac{q^3 \hat{q}}{4} \left( \frac{\bar{q} \|\sigma\|_\infty^2 \|\nabla \vartheta\|_\infty^2}{2} + e_n \right), \quad (6.24)$$

where  $q := q(n) > 1$ ,  $\bar{q} := \bar{q}(n) > 1$ ,  $\hat{q} := \hat{q}(n)$  with  $q, \bar{q}, \hat{q} \rightarrow 1$  and  $e_n$  is an explicit sequence going to 0.

The results of Theorem 3 explicitly follow taking (see for more details the proof of Lemma 7):

$$\begin{aligned} \rho - 1 &= \frac{1}{2} \frac{\tilde{B}_n^{1/2} a}{\tilde{A}_n^{3/2} \sqrt{\Gamma_n}}, \text{ for } \frac{a}{\sqrt{\Gamma_n}} \rightarrow 0, \\ \rho - 1 &= \frac{1}{2}, \text{ for } \frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty. \end{aligned}$$

**Remark 17** *It is natural to wonder if it is possible to get a sharper variance for super Gaussian deviations. Thereby it would be tempting to bootstrap Lemma 3. Such an iteration would lead to optimize polynomials of higher degrees. Recall from Lemma 4, that we already have to handle a polynomial of order 4 in the current setting. This illustrates that for very large deviations the highest term dominates and deteriorates the concentration. This is intrinsic to our approach. Such a phenomenon would even more pregnant when iterating the procedure the polynomial of higher degree yields, for a certain very large deviation regime, concentration bounds that become closer and closer to the exponential. However, bootstrapping might allow to improve the constants in the successive deteriorated concentration regimes. As indicated in Remark 16, this could be useful for some numerical purposes.*

## 7 Numerical Results

Here, we have chosen to highlight the possible absence of non-degeneracy assumption for our results. To oversimplify, simulations are done with  $r = d = 1$ ,  $X_0$  and  $U_1$  follow the standard normal distribution. Naturally, for a better convergence speed, we take  $\theta \approx \frac{1}{3}$ , precisely  $\theta = \frac{1}{3} + \frac{1}{1000}$ .

For this first example, we choose  $\varphi$  (solution of the Poisson equation). We take  $\varphi = \sigma = \cos$  and for all  $x \in \mathbb{R}$ ,  $b(x) = -\frac{x}{2}$ . By this pick, we compute numerically  $\nu(|\sigma^* \nabla \varphi|^2) \approx 0.1515$  and  $\nu(|\sigma|^2) \approx 0.4171$  (that we provide here for comparison with the previous results in Chapter 3), with the same parameters ( $\theta = \frac{1}{3} + \frac{1}{1000}$ ,  $n = 5 \cdot 10^4$  and  $MC = 10^4$ ).

Heed, for a non trivial test function  $\varphi$ , with our method, we cannot choose functions  $b$  and  $\sigma \neq 0$  canceling at the same point (0 here). Otherwise, the Poisson equation associated with the *carré du champ* source,  $\mathcal{A}\vartheta = |\sigma^* \nabla \varphi|^2 - \nu(|\sigma^* \nabla \varphi|^2)$ , would imply that  $-\nu(|\sigma^* \nabla \varphi|^2) = 0$ , then  $\nabla \varphi = 0$ ,  $\nu$  almost surely.

Let us now check that the Confluence Conditions ( $\mathbf{D}_\alpha^p$ ) are satisfied. For  $p \in [1, 2)$ , we have for all  $x \in \mathbb{R}^d$ ,  $\xi \in \mathbb{R}^d$

$$\begin{aligned} & \left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r \left( (p-2) \frac{|\langle D\sigma_{\cdot j}(x) \xi, \xi \rangle|^2}{|\xi|^2} + |D\sigma_{\cdot j} \xi|^2 \right) \\ &= -\frac{1}{2} \xi^2 + \frac{1}{2} \sin^2(x) \xi^2 (p-1). \end{aligned} \quad (7.1)$$

So, for  $p = \frac{3}{2}$ , we directly obtain:

$$\begin{aligned} & \left\langle \frac{Db(x) + Db(x)^*}{2} \xi, \xi \right\rangle + \frac{1}{2} \sum_{j=1}^r \left( \left(\frac{3}{2} - 2\right) \frac{|\langle D\sigma_{\cdot j}(x) \xi, \xi \rangle|^2}{|\xi|^2} + |D\sigma_{\cdot j} \xi|^2 \right) \\ &= -\frac{1}{2} \xi^2 + \frac{1}{4} \sin^2(x) \xi^2 \leq -\frac{1}{4} \xi^2 =: -\alpha \xi^2. \end{aligned} \quad (7.2)$$

Note that, we have chosen a diffusion coefficient  $\sigma$  which degenerates on  $\{k\pi, k \in \mathbb{Z}\}$ . However, thanks to the smoothness of the diffusion parameters, we can still here apply Lemma 6 in Chapter 3 which gives us a pointwise gradient bound of the solution of the Poisson problem in the current degenerate context. In other words:

$$[\vartheta]_1 \leq \frac{[|\sigma^* \nabla \varphi|^2]_1}{\alpha} = 4[\cos^2 \sin^2]_1 = 4 \sup_{x \in \mathbb{R}} (\cos(2x) \sin(2x)) = 2.$$

Hence, this inequality leads us to approximate  $[\vartheta]_1$  by 2. Pay attention that the control of the Lipschitz constant  $[\vartheta]_1$  is important for the *super Gaussian deviations*. Like illustrated in Remarks 16 and 18, this regime appears “sooner” than we might expect.

From Theorem 3, the function  $a \mapsto g_n(a) := \ln(\mathbb{P}[\sqrt{\Gamma_n} |\nu_n(\mathcal{A}\varphi)| \geq a])$  is s.t. for  $a > 0$ :

$$g_n(a) \leq -c_n \frac{a^2}{2\nu(|\sigma^* \nabla \varphi|^2)} + \ln(2C_n),$$

where  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$  are sequences respectively increasing and decreasing for  $n$  large enough, with  $\lim_n C_n = \lim_n c_n = 1$ .

For Figure 4.1, the simulations have been performed for  $n = 5 \cdot 10^4$  and the probability estimated by Monte Carlo simulation for  $MC = 10^4$  realizations of the random variable  $\sqrt{\Gamma_n} |\nu_n(\mathcal{A}\varphi)|$ . The corresponding 95% confidence intervals have size at most of order 0.0016. We introduce the functions:

$$S(a) := -\frac{a^2}{2\nu(|\sigma^* \nabla \varphi|^2)}, \quad S_{\text{sup}}(a) := -\frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2}.$$

Like in Theorem 7, we take

$$P_{\min}(a, \Gamma_n, \rho) = -\frac{(\rho - 1)^{1/3} \sqrt{\Gamma_n} \Phi_n(a, \rho)}{\rho} \frac{\sqrt{\Gamma_n} \Phi_n(a, \rho)}{8} (3a - (\rho - 1)^{1/3} \tilde{A}_n \Phi_n(a, \rho) \sqrt{\Gamma_n}),$$

where  $\tilde{A}_n, \tilde{B}_n$  and  $\Phi_n(a, \rho)$  are defined in (6.24). Through our numerical results, we take  $e_n = 0$ .

We set also:

$$\rho_0 := 1 + \frac{1}{2} \frac{\tilde{B}_n^{1/2} a}{\tilde{A}_n^{3/2} \sqrt{\Gamma_n}}, \quad \rho_\infty := \frac{3}{2}.$$

We recall here that  $\rho$  and  $\rho_\infty$  respectively correspond to the optimal values of  $\rho$  in the *Gaussian deviations* and *super Gaussian deviations* (see Lemma 7). Eventually, we introduce:

$$\begin{aligned} P_{n,0,\infty}(a) &:= \min \left( P_{\min}(a, \Gamma_n, \rho_0), P_{\min}(a, \Gamma_n, \rho_\infty) \right), \\ P_n(a) &:= \min_{\rho > 1} P_{\min}(a, \Gamma_n, \rho). \end{aligned}$$

Note that, the function  $P_{n,0,\infty}$  takes into account the multi-regime competition. We have estimated  $P_n(a)$  by a mesh method for  $\rho \in (1, 2)$  and for a grid with  $5 \cdot 10^5$  steps.

From the above notations, we add the subscript  $\sigma$  to mean that we change  $\nu(|\sigma^* \nabla \varphi|^2)$  into  $\|\nabla \varphi\|_\infty^2 \nu(\|\sigma\|^2)$ , i.e.

$$S_\sigma(a) = -\frac{a^2}{2\|\nabla \varphi\|_\infty^2 \nu(\|\sigma\|^2)}, \quad P_{n,\sigma}(a) := \min_{\rho > 1} P_{\min,\sigma}(a, \Gamma_n, \rho),$$

and we have changed  $\tilde{A}_n$  into

$$\tilde{A}_{\sigma,n} := \frac{\|\nabla \varphi\|_\infty^2 \nu(\|\sigma\|^2)}{2}.$$

The quantities with subscript  $\sigma$  are those associated with the results in Chapter 3, recalled in the previous Theorem 2, where the variance is less sharp than the constants appearing in Theorems 3, 6 and 7. Thus, we can compare our main results with Remark 10 of Chapter 3 which is a weakened form of Theorem 7 where the *carré du champ* is changed into  $\|\nabla \varphi\|_\infty^2 \nu(\|\sigma\|^2)$  like in Theorem 2.

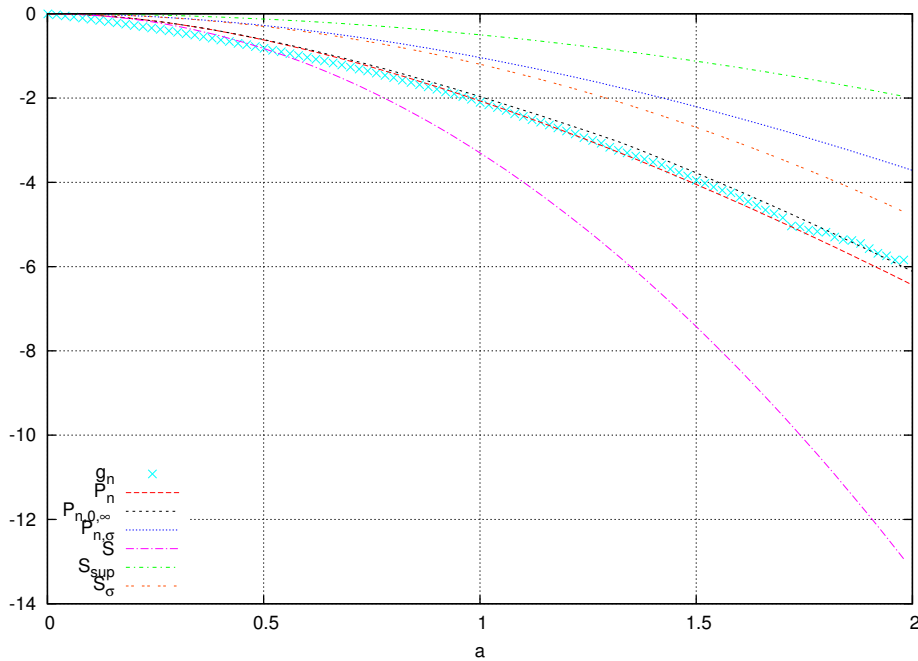


Figure 4.1 – Plot of  $a \mapsto g_n(a)$  with  $\varphi(x) = \sigma(x) = \cos(x)$ .

Figure 4.1 reveals that the asymptotic curve  $S$  is much less sharp with respect to the realizations  $g_n$  than our main estimations  $P_n$  and  $P_{n,0,\infty}$ . In fact, these latter are very close to the realization  $g_n$ . This claim enhances the significance of controlling finely, non-asymptotically, the deviation of the empirical measure.

In this plot, we can see that our pick of  $\rho$  for  $P_{n,0,\infty}$ , set in Lemma 7, is very close to the numerical optimization of  $P_n$  over  $\rho$ . Nevertheless, observe that for  $a > 0.5$ ,  $P_{n,0,\infty}(a)$  and  $P_n(a)$  slightly differ. It means that progressively the regime goes from *Gaussian deviations* (i.e.  $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ ) to *intermediate Gaussian deviations* (i.e.  $\frac{a}{\sqrt{\Gamma_n}} = O(1)$ ). Hence, the importance of optimizing globally the function  $\rho \mapsto P_{\min}(a, \Gamma_n, \rho)$  (appearing in (3.35)) in all regimes.

**Remark 18** Remark that for the graphic 4.1, we chose  $n = 5 \cdot 10^4$ , but for  $\theta \approx \frac{1}{3}$ ,  $\sqrt{\Gamma_n} \approx 37$  and for  $\theta \approx \frac{1}{2+0.5}$ ,  $\sqrt{\Gamma_n} \approx 26$ . In other words, for  $a \approx 1$  we have intermediate Gaussian deviations as emphasized by the graphic. Hence the importance of the study of both regimes, Gaussian deviations ( $\frac{a}{\sqrt{\Gamma_n}} \rightarrow 0$ ) and super Gaussian deviations ( $\frac{a}{\sqrt{\Gamma_n}} \rightarrow +\infty$ ).

## 8 Appendix: Computation of asymptotic analysis

In this section, we perform asymptotic analysis for the map  $g'$  defined in (6.12) in proof of Lemma 4. We recall that for all  $\xi \in \mathbb{R}$ :

$$g(\xi) = \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \left( 1 - \frac{\xi^{1/3}}{2} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right) \right).$$

**Lemma 8**

$$g'(\xi) \underset{\xi \rightarrow 0}{=} \frac{2^{1/3}}{3\xi^{2/3}}(1 + o(1)), \quad g'(\xi) \underset{\xi \rightarrow +\infty}{=} \frac{8}{3^5\xi^2} + o\left(\frac{1}{\xi^2}\right).$$

**Proof:** Denote  $h(\xi) := \xi^{1/3} \left( (1 + \sqrt{1 + \xi})^{1/3} + (1 - \sqrt{1 + \xi})^{1/3} \right)$ , so  $g(\xi) = h(\xi)(1 - \frac{h(\xi)}{2})$ .

Differentiating, we get:

$$\begin{aligned} h'(\xi) &= \frac{(1 - \sqrt{1 + \xi})^{1/3} + (1 + \sqrt{1 + \xi})^{1/3}}{3\xi^{2/3}} \\ &\quad + \xi^{1/3} \left( \frac{1}{6\sqrt{1 + \xi}(1 + \sqrt{1 + \xi})^{2/3}} - \frac{1}{6\sqrt{1 + \xi}(1 - \sqrt{1 + \xi})^{2/3}} \right) \\ &= \frac{(1 - \sqrt{1 + \xi})^{1/3} + (1 + \sqrt{1 + \xi})^{1/3}}{3\xi^{2/3}} + \frac{1}{6\sqrt{1 + \xi}} \frac{(1 - \sqrt{1 + \xi})^{2/3} - (1 + \sqrt{1 + \xi})^{2/3}}{\xi^{2/3}}. \end{aligned}$$

(a) For  $\xi \rightarrow 0$ ,

$$h(\xi) \underset{\xi \rightarrow 0}{=} 2^{1/3}\xi^{1/3} + o(\xi^{1/3}),$$

and

$$h'(\xi) \underset{\xi \rightarrow 0}{=} \frac{2^{1/3}}{3\xi^{2/3}} + \xi^{1/3} \left( \frac{1}{6 \times 2^{2/3}} - \frac{2^{2/3}}{\xi^{2/3}} \right) + o\left(\frac{1}{\xi^{2/3}}\right) = \frac{2^{1/3}}{3\xi^{2/3}} + o\left(\frac{1}{\xi^{2/3}}\right),$$

which yields that

$$g'(\xi) = h'(\xi)(1 - h(\xi)) \underset{\xi \rightarrow 0}{=} \frac{2^{1/3}}{3\xi^{2/3}} - o\left(\frac{1}{\xi^{2/3}}\right).$$

(b) For  $\xi \rightarrow +\infty$ ,

In order to estimate  $g'$  we need to do a Taylor expansion up to the third order:

$$\begin{aligned} h(\xi) &\underset{\xi \rightarrow +\infty}{=} \xi^{1/3}(1 + \xi)^{1/6} \left( \left(1 + \frac{1}{\sqrt{1 + \xi}}\right)^{1/3} - \left(1 - \frac{1}{\sqrt{1 + \xi}}\right)^{1/3} \right) \\ &= \xi^{1/3}(1 + \xi)^{1/6} \left( 1 + \frac{1}{3\sqrt{1 + \xi}} - \frac{1}{3^2(1 + \xi)} + \frac{2 \times 5}{3^3 3!(1 + \xi)^{3/2}} \right. \\ &\quad \left. - \left( 1 - \frac{1}{3\sqrt{1 + \xi}} - \frac{1}{3^2(1 + \xi)} - \frac{2 \times 5}{3^3 3!(1 + \xi)^{3/2}} \right) + o\left(\frac{1}{\xi^{3/2}}\right) \right) \\ &= \xi^{1/2} \left( 1 + \frac{1}{6\xi} + o\left(\frac{1}{\xi}\right) \right) \left( \frac{2}{3\sqrt{1 + \xi}} + \frac{10}{3^4(1 + \xi)^{3/2}} + o\left(\frac{1}{\xi^{3/2}}\right) \right) \\ &= \xi^{1/2} \left( 1 + \frac{1}{6\xi} + o\left(\frac{1}{\xi}\right) \right) \left( \frac{2}{3\sqrt{\xi}} - \frac{1}{3\xi^{3/2}} + \frac{10}{3^4(1 + \xi)^{3/2}} + o\left(\frac{1}{\xi^{3/2}}\right) \right) \\ &= \xi^{1/2} \left( \frac{2}{3\sqrt{\xi}} + \frac{1}{3^2\xi^{3/2}} - \frac{1}{3\xi^{3/2}} + \frac{10}{3^4(1 + \xi)^{3/2}} + o\left(\frac{1}{\xi^{3/2}}\right) \right) \\ &= \frac{2}{3} - \frac{8}{81\xi} + o\left(\frac{1}{\xi}\right). \end{aligned}$$

Differentiating the above expression, we get:

$$h'(\xi) \underset{\xi \rightarrow +\infty}{=} \frac{8}{81\xi^2} + o\left(\frac{1}{\xi^2}\right),$$



which yields that

$$\begin{aligned}g'(\xi) &= h'(\xi)(1 - h(\xi)) \underset{\xi \rightarrow 0}{=} \left(\frac{8}{81\xi^2} + o\left(\frac{1}{\xi^2}\right)\right) \left(1 - \frac{2}{3} + \frac{8}{81\xi} + o\left(\frac{1}{\xi}\right)\right) \\ &= \frac{16}{3^5\xi^2} - \frac{8}{3^5\xi^2} + o\left(\frac{1}{\xi^2}\right) = \frac{8}{3^5\xi^2} + o\left(\frac{1}{\xi^2}\right).\end{aligned}$$

□

# Chapter 5

## Extension to SDEs driven by a Poisson compound process

**Abstract:** In this article we approximate the invariant distribution  $\nu$  of an ergodic Jump Diffusion driven by the sum of a Brownian motion and a Compound Poisson process with sub-Gaussian jumps. We first construct an Euler discretization scheme with decreasing time steps, particularly suitable in cases where the driving Lévy process is a Compound Poisson. This scheme is similar to those introduced by Lambertson and Pagès in [LP02] for a Brownian diffusion and extended by Panloup in [Pan08b] to the Jump Diffusion with Lévy jumps. We obtain a non-asymptotic Gaussian concentration bound for the difference between the invariant distribution and the empirical distribution computed with the scheme of decreasing time step along a appropriate test functions  $f$  such that  $f - \nu(f)$  is a coboundary of the infinitesimal generator.

### 1 Introduction

#### 1.1 Setting

Let  $(\mathbf{X}_t)_{t \geq 0}$  be a  $d$ -dimensional càdlàg process solution of the stochastic differential equation:

$$d\mathbf{X}_t = b(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)dW_t + \kappa(\mathbf{X}_{t-})d\mathbf{Z}_t. \quad (1.1)$$

where  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  and  $\kappa : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  are Lipschitz continuous,  $(W_t)_{t \geq 0}$  is a Wiener process of dimension  $r$ , and  $(\mathbf{Z}_t)_{t \geq 0}$  is a  $\mathbb{R}^r$ -valued compound Poisson process (CPP),

$\mathbf{Z}_t = \sum_{k=1}^{N_t} \mathbf{Y}_k$ , where  $(\mathbf{Y}_k)_{k \in \mathbb{N}}$  are i.i.d.  $r$ -dimensional random vectors with common distribution  $\pi$  on  $\mathcal{B}(\mathbb{R}^r)$  and  $(N_t)_{t \geq 0}$  is a Poisson process, independent of  $(\mathbf{Y}_k)_{k \in \mathbb{N}}$ . The processes  $(W_t)_{t \geq 0}$  and  $(\mathbf{Z}_t)_{t \geq 0}$  are assumed to have the same dimension for the sake of simplicity. Moreover,  $(N_t)_{t \geq 0}$ ,  $(\mathbf{Y}_k)_{k \in \mathbb{N}}$  and  $(W_t)_{t \geq 0}$  are independent and defined on a given filtered probability space  $(\Omega, \mathbf{G}, (\mathbf{G}_t)_{t \geq 0}, \mathbb{P})$ . We assume that  $b$ ,  $\sigma$ , and  $\kappa$  satisfy a suitable Lyapunov condition (assumption  $(\mathcal{L}_{\mathbf{v}})$  in Section 1.3) which ensures the existence of an invariant distribution  $\nu$  of  $(\mathbf{X}_t)_{t \geq 0}$  (see [Pan08b]). For the sake of simplicity we also assume

the uniqueness of the invariant distribution. We refer to [Mas07] under irreducibility and Lyapunov conditions for the existence and uniqueness of the invariant distribution for a diffusion driven by Lévy process.

The aim of this paper is to establish a non-asymptotic bound on the probability of the deviation  $\nu_n(f) - \nu(f)$ , where  $\nu_n$  is an appropriate empirical measure such that  $\lim_{n \rightarrow \infty} \nu_n(f) = \nu(f)$  *a.s.* for all suitable test functions  $f$ .

The algorithm that we define in this article is based on an Euler-like discretization scheme with decreasing time step  $(\gamma_n)_{n \geq 1}$  s.t.  $\lim_n \gamma_n = 0$ . Lamberton and Pagès first introduced such a scheme in [LP02] for a Brownian diffusion. They showed that the empirical measure of their scheme converges to the invariant measure of the diffusion and that it satisfies the Central Limit Theorem. The decreasing steps allows to the empirical measure to directly converge towards the invariant one. If we choose a constant time step  $\gamma_k = h > 0$  in the scheme, the expected ergodic theorem is  $\nu_n(f) \xrightarrow[n]{a.s.} \nu^h(f) = \int_{\mathbb{R}^d} f(x) \nu^h(dx)$ , where  $\nu^h$  is the invariant distribution of the scheme which is supposed to converge toward the invariant measure of the diffusion (1.1) when  $h \rightarrow 0$  (see e.g. For more details about this approach we refer to [TT90], [Tal02] and [MT06]).

Next, Panloup in [Pan08a] and [Pan08b] adapted the algorithm of [LP02] to the Jump Diffusion with Lévy jumps [Pan08b] and also showed the convergence and the Central Limit Theorem for the empirical measure in this case. In the same way as the questions of the convergence of the empirical measure  $\nu_n$  or of its limiting distribution, the natural question is that of the nature of the deviations  $\nu_n(f) - \nu(f)$  along appropriate test functions  $f$ . In the case of the Brownian diffusion this question was considered in Chapters 3 and 4. Note that in the Brownian diffusion case the innovations of the Euler scheme are designed in order to “mimic” Brownian increments, hence it is natural to assume that they satisfy some Gaussian Concentration property (assumption **(GC)** in Section 1.3). In particular this Gaussian Concentration property is satisfied by Gaussian or symmetric Bernoulli law. Taken as an assumption on the Brownian innovations of the scheme, it allows to show a non-asymptotic Gaussian Concentration bound for the probability of the deviations of  $\nu_n(f)$  from  $\nu(f)$ , see Chapters 3 and 4 with sharp constants. The deviation  $\nu_n(f) - \nu(f)$  is evaluated along the functions  $f$  such that  $f - \nu(f)$  is a coboundary of the infinitesimal generator of the diffusion.

When the diffusion contains Lévy jumps, it is not generally expected that these deviations will show a Gaussian behaviour. But such a behaviour seems natural if we suppose that the driving Levy process is a Compound Poisson process and the jump size vectors  $(\mathbf{Y}_k)_{k \in \mathbb{N}}$  satisfy a Gaussian Concentration property **(GC)**. In this paper, we focus on this situation. Before giving its precise formulation we need to introduce some notations. First of all, we introduce our discretization scheme. In general, for a Euler scheme corresponding to a Jump Diffusion with Lévy jumps, one has to define a numerically computable jump vectors designed to “mimic” the increments of the driving Lévy process. In most cases, the increments of a Lévy process are not numerically computable, that is why it is important to propose different ways to approximate these increments according to the nature of the driving Lévy process. In this paper we introduce a scheme (S) particularly suitable in the case where a driven Lévy process is a Compound Poisson. Note that our scheme is close to the scheme (C) of [Pan08b]. Like in the previously mentioned articles,

we denote time steps  $(\gamma_k)_{k \geq 1}$ , and for all  $n \geq 0$ , we define:

$$X_{n+1} = X_n + \gamma_{n+1}b(X_n) + \sqrt{\gamma_{n+1}}\sigma(X_n)U_{n+1} + \kappa(X_n)Z_{n+1}, \quad (1.2)$$

where  $X_0$  is an  $\mathbb{R}^d$  valued random variables such that  $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$ ,  $(U_n)_{n \geq 1}$  is an i.i.d. sequence of centered random variables matching the moments of the Gaussian law on  $\mathbb{R}^r$  up to order three, independent of  $X_0$ . Furthermore, for all  $n \geq 1$  we put

$$Z_n := B_n Y_n, \quad (1.3)$$

where  $(B_n)_{n \geq 1}$  are one-dimensional independent Bernoulli random variables, independent of  $X_0$ ,  $(U_n)_{n \geq 1}$  and  $(Y_n)_{n \geq 1}$ , s.t.  $B_n \stackrel{(\text{law})}{=} \mathcal{B}ern(\mu\gamma_n)$ , where  $\mu$  is an intensity of the Poisson process driving the CPP  $(\mathbf{Z}_t)_{t \geq 0}$ . Without loss of generality we can suppose from now on that  $\mu = 1$ . The choice (1.3) of the innovations  $Z_n, n \in \mathbb{N}$  is motivated by the following heuristic reasoning:  $Z_n$  has to "mimic" the increment of the CPP  $\mathbf{Z}_t = \sum_{k=1}^{N_t} Y_k$  on the small time interval of the length  $\gamma_n$ . The probability that the CPP does not jump on this interval is equal to  $\exp(-\gamma_n) = 1 - \gamma_n + o(\gamma_n)$ , and if the CPP jumps on this interval, it will most probably have only one jump. Hence we approximate the increment  $\mathbf{N}_{\gamma_n}$  of the CPP by a  $\{0, 1\}$  random variable with the probability of 1 equal to  $\gamma_n$ .

We also introduce the empirical (random) measure of the scheme: for all  $A \in \mathcal{B}(\mathbb{R}^d)$  (where  $\mathcal{B}(\mathbb{R}^d)$  is the Borel  $\sigma$ -field on  $\mathbb{R}^d$ ):

$$\nu_n(A) := \nu_n(\omega, A) := \frac{\sum_{k=1}^n \gamma_k \delta_{X_{k-1}(\omega)}(A)}{\sum_{k=1}^n \gamma_k}. \quad (1.4)$$

Obviously, to study long time behaviour, we have to consider steps  $(\gamma_k)_{k \geq 1}$  such that the current time of the scheme  $\Gamma_n := \sum_{k=1}^n \gamma_k \xrightarrow[n]{n} +\infty$ . We recall as well that  $\gamma_k \xrightarrow[k]{k} 0$ . We suppose that both jump amplitudes  $(Y_n)_{n \geq 1}$  and Brownian innovations  $(U_n)_{n \geq 1}$  satisfy a Gaussian concentration (see further the assumption **(GC)**). As we already mentioned, the aim of the paper is to show that this assumption implies a non-asymptotic Gaussian Concentration inequality for the probability of the deviations of  $\nu_n(f)$  from  $\nu(f)$  (see Theorem 2, Section 2). The main argument in the proof of Theorem 2 is the fact that the **(GC)** property of jumps sizes  $Y_k, k \in \mathbb{N}$ , permits to show the similar Gaussian Concentration property for the jump innovations  $Z_k, k \in \mathbb{N}$ . This result is given in Proposition 1. However the Gaussian Concentration property of jump innovations depends on the dimension of the jump heights. This dependence survives in the main Theorem 2 giving Gaussian Concentration of the deviation of  $\nu_n$  from  $\nu$ .

The current chapter is organized as follows. In Section 1.2, we introduce some useful notations. The assumptions required for our main results are outlined in Section 1.3. In this part, we formulate a Gaussian concentration property of the jump innovation  $Z_n$ , the proof is given in Section 2.4. We state in Section 1.4 some already known results connected with the approximation scheme. Our main results are in Section 2, and the demonstration is located in Section 2.3. Section 3 is dedicated to the analysis of the exponential integrability of Lyapunov function. Technical lemmas are stated in Section 2.2, but their proofs are postponed to Section 4. Eventually, we propose a numerical illustration of our main result in Section 5.

## 1.2 General notations

We set for any step sequence  $(\gamma_n)_{n \geq 1}$ :

$$\forall \ell \in (0, +\infty), \Gamma_n^{(\ell)} := \sum_{k=1}^n \gamma_k^\ell, \Gamma_n := \sum_{k=1}^n \gamma_k = \Gamma_n^{(1)},$$

where  $\Gamma_n$  corresponds to the current time, hence  $\Gamma_n \xrightarrow{n \rightarrow +\infty} +\infty$ . For the sake of simplicity, from now on, the time step sequence will have the form:  $\gamma_n \asymp \frac{1}{n^\theta}$  with  $\theta \in (0, 1]$ , where for two sequences  $(u_n)_{n \in \mathbb{N}}, (v_n)_{n \in \mathbb{N}}$  the notation  $u_n \asymp v_n$  means that  $\exists n_0 \in \mathbb{N}, \exists C \geq 1$  s.t.  $\forall n \geq n_0, C^{-1}v_n \leq u_n \leq Cv_n$ .

Henceforth,  $C$  will be a non negative constant, and  $(e_n)_{n \geq 1}, (\mathcal{R}_n)_{n \geq 1}$  will be deterministic sequences s.t.  $e_n \rightarrow_n 0$  and  $\mathcal{R}_n \rightarrow_n 1$ , that may change from line to line. The constant  $C$  as well as the sequences  $(e_n)_{n \geq 1}, (\mathcal{R}_n)_{n \geq 1}$  depend on known parameters appearing in the hypotheses set in Section 1.3 (which will be called **(A)** further). Other possible dependencies will be explicitly specified.

We denote by  $I_m, m \in \{d, r\}$  the identity matrix of dimension  $m$ .

Through the article, for any smooth enough function  $f$ , for  $k \in \mathbb{N}$  we will denote  $D^k f$  the tensor of the  $k^{\text{th}}$  derivatives of  $f$ . Namely  $D^k f = (\partial_{i_1} \dots \partial_{i_k} f)_{1 \leq i_1, \dots, i_k \leq d}$ . Yet, for a multi-index  $\alpha \in \mathbb{N}_0^d := (\mathbb{N} \cup \{0\})^d$ , we set  $D^\alpha f = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} f : \mathbb{R}^d \rightarrow \mathbb{R}$ .

For a  $\beta$ -Hölder continuous function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , we denote by

$$[f]_\beta := \sup_{x \neq x'} \frac{|f(x) - f(x')|}{|x - x'|^\beta} < +\infty,$$

its Hölder modulus of continuity. Here,  $|x - x'|$  stands for the Euclidean norm of  $x - x' \in \mathbb{R}^d$ .

We define for  $(p, d, m) \in \mathbb{N}^3, \mathcal{C}^p(\mathbb{R}^d, \mathbb{R}^m)$  the space of  $p$ -times continuously differentiable functions from  $\mathbb{R}^d$  to  $\mathbb{R}^m$ . Furthermore, for  $f \in \mathcal{C}^p(\mathbb{R}^d, \mathbb{R}^m), p \in \mathbb{N}$ , we set for  $\beta \in (0, 1]$  the Hölder modulus:

$$[f^{(p)}]_\beta := \sup_{x \neq x', |\alpha|=p} \frac{|D^\alpha f(x) - D^\alpha f(x')|}{|x - x'|^\beta} \leq +\infty,$$

where  $\alpha \in \mathbb{N}^d$  is a multi-index of length  $p$ , namely  $|\alpha| := \sum_{i=1}^d \alpha_i = p$ . In other words, in the above definition, the  $|\cdot|$  in the numerator is the usual absolute value. We will also use the notation  $\llbracket n, p \rrbracket, (n, p) \in (\mathbb{N}_0)^2, n \leq p$ , for the set of integers being between  $n$  and  $p$ .

Let us introduce for  $k \in \mathbb{N}_0, \beta \in (0, 1]$  and  $m \in \{1, d, d \times r\}$  the Hölder spaces

$$\begin{aligned} \mathcal{C}^{k, \beta}(\mathbb{R}^d, \mathbb{R}^m) &:= \{f \in \mathcal{C}^k(\mathbb{R}^d, \mathbb{R}^m) : \forall \alpha \in \mathbb{N}^d, |\alpha| \in \llbracket 1, k \rrbracket, \sup_{x \in \mathbb{R}^d} |D^\alpha f(x)| < +\infty, [f^{(k)}]_\beta < +\infty\}, \\ \mathcal{C}_b^{k, \beta}(\mathbb{R}^d, \mathbb{R}^m) &:= \{f \in \mathcal{C}^{k, \beta}(\mathbb{R}^d, \mathbb{R}^m) : \|f\|_\infty < +\infty\}. \end{aligned} \tag{1.5}$$

In the above definition, we denote for all bounded mapping  $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}^m$ ,  $m \in \{1, d, d \times r\}$ , the uniform norm  $\|\zeta\|_\infty := \sup_{x \in \mathbb{R}^d} \|\zeta \zeta^*(x)\|$  with  $\|\zeta(x)\| = \text{Tr}(\zeta \zeta^*(x))^{1/2}$ , where for  $M \in \mathbb{R}^m \otimes \mathbb{R}^m$ ,  $\text{Tr}(M)$  is the trace of  $M$ . In particular,  $\|\cdot\|$  is the Fröbenius norm\*.

Practically, with these notations,  $\mathcal{C}^{k,\beta}(\mathbb{R}^d, \mathbb{R}^m)$  stands for the subset of  $\mathcal{C}^k(\mathbb{R}^d, \mathbb{R}^m)$  whose elements have bounded derivatives up to order  $k$  and  $\beta$ -Hölder continuous  $k^{\text{th}}$  derivatives. In particular, for  $k = 0$ , the space of Lipschitz continuous functions from  $\mathbb{R}^d$  to  $\mathbb{R}^m$  is denoted by  $\mathcal{C}^{0,1}(\mathbb{R}^d, \mathbb{R}^m)$ .

For a given Borel function  $f : \mathbb{R}^d \rightarrow E$ , where  $E$  can be  $\mathbb{R}$ ,  $\mathbb{R}^d$ ,  $\mathbb{R}^d \otimes \mathbb{R}^r$ ,  $\mathbb{R}^d \otimes \mathbb{R}^d$ , we set for  $k \in \mathbb{N}_0$ :

$$f_k := f(X_k).$$

Moreover, for  $k \in \mathbb{N}_0$ , we denote

$$\mathcal{F}_k := \sigma(X_0, (U_j, Z_j)_{j \in \llbracket 1, k \rrbracket}) \quad \text{and} \quad \tilde{\mathcal{F}}_k := \sigma(X_0, (U_j, Z_j)_{j \in \llbracket 1, k \rrbracket}, U_{k+1}). \quad (1.6)$$

Eventually, we define the infinitesimal generator associated with the diffusion (1.1) which writes for all  $\varphi \in \mathcal{C}^2(\mathbb{R}^d)$  and  $x \in \mathbb{R}^d$ :

$$\begin{aligned} \mathcal{A}\varphi(x) &= b(x)\nabla\varphi(x) + \frac{1}{2}\text{Tr}(\sigma\sigma^*(x)D^2\varphi(x)) + \int_{\mathbb{R}^d} (\varphi(x + \kappa(x)y) - \varphi(x))\pi(dy) \\ &=: \tilde{\mathcal{A}}\varphi(x) + \int_{\mathbb{R}^d} (\varphi(x + \kappa(x)y) - \varphi(x))\pi(dy), \end{aligned} \quad (1.7)$$

where  $\pi$  stands for the distribution of  $Y_1$ , and  $\tilde{\mathcal{A}}$  is the infinitesimal generator of the continuous part of the diffusion.

### 1.3 Hypotheses

We assume the following set of hypothesis about the coefficients of the SDE (1.1) and the parameters of the scheme (1.2):

**(C0)** The functions  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  and  $\kappa : \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^r$  are globally Lipschitz continuous.

**(C1)** The first value of the scheme  $X_0$  is sub-Gaussian: there exists  $\lambda_0 \in \mathbb{R}_+^*$  such that

$$\forall \lambda < \lambda_0, \quad \mathbb{E}[\exp(\lambda|X_0|^2)] < +\infty.$$

**(C2)** Defining for all  $x \in \mathbb{R}^d$ ,  $\Sigma(x) := \sigma\sigma^*(x)$ ,  $K(x) = \kappa\kappa^*(x)$ , we suppose that

$$\sup_{x \in \mathbb{R}^d} \text{Tr}(\Sigma(x)) = \sup_{x \in \mathbb{R}^d} \|\sigma(x)\|^2 =: \|\sigma\|_\infty^2 < +\infty, \quad \sup_{x \in \mathbb{R}^d} \text{Tr}(K(x)) = \sup_{x \in \mathbb{R}^d} \|\kappa(x)\|^2 =: \|\kappa\|_\infty^2 < +\infty.$$

---

\*. Remark that this notation is also available for vector norms. Indeed, any  $\mathbb{R}^d$  vectors can be regarded as line vectors, and we define similarly for a vector or for a matrix the uniform norm  $\|\cdot\|_\infty$ .

**(GM)** The sequences of random variables  $(U_n)_{n \geq 1}$  and  $(Y_n)_{n \geq 1}$  are respectively i.i.d., such that

$$\begin{aligned} \mathbb{E}[U_1] = \mathbb{E}[Y_1] = 0; \quad \mathbb{E}[(U_1^i U_1^j)_{1 \leq i, j \leq r}] &=: \mathbb{E}[U_1^{\otimes 2}] = I_r, \\ \mathbb{E}[Y_1^{\otimes 2}] = I_r; \quad \mathbb{E}[(U_1^i U_1^j U_1^k)_{1 \leq i, j, k \leq r}] &=: \mathbb{E}[U_1^{\otimes 3}] = 0^\dagger. \end{aligned}$$

Also,  $(U_n)_{n \geq 1}$ ,  $(Y_n)_{n \geq 1}$  and  $X_0$  are independent.

**(GC)** We say that r.v.  $G \in \mathbb{L}^1$  satisfies Gaussian concentration property, if for every Lipschitz continuous function  $g : \mathbb{R}^r \rightarrow \mathbb{R}$  and every  $\lambda > 0$ :

$$\mathbb{E}[\exp(\lambda g(G))] \leq \exp\left(\lambda \mathbb{E}[g(G)] + \frac{\lambda^2 [g]_1^2}{2}\right), \quad (1.8)$$

We assume that  $U_1$  and  $Y_1$  satisfy the Gaussian concentration property.

**(L<sub>V</sub>)** We assume the following Lyapunov like stability condition:

There exists a non-negative function  $V : \mathbb{R}^d \rightarrow [v^*, +\infty)$  with  $v^* > 0$  such that

- i)  $V$  is a  $\mathcal{C}^2$  continuous function s.t.  $\|D^2 V\|_\infty < +\infty$ , and  $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ .
- ii) There is  $C_V \in (0, +\infty)$  such that for all  $x \in \mathbb{R}^d$ :

$$|\nabla V(x)|^2 + |b(x)|^2 \leq C_V V(x).$$

- iii) There exist  $\alpha_V > 0$ ,  $\beta_V \in \mathbb{R}^+$  such that for all  $x \in \mathbb{R}^d$ ,

$$\mathcal{A}V(x) \leq -\alpha_V V(x) + \beta_V.$$

**(U)** There is a unique invariant distribution  $\nu$  to equation (1.1).

**(S)** We assume that the sequence  $(\gamma_k)_{k \geq 1}$  is small enough, namely for all  $k \geq 1$ ,

$$\gamma_k \leq \frac{1}{2} \min\left(2, \frac{\alpha_V}{4\left(\frac{[b]_1 C_V}{8} + \left(\frac{\sqrt{C_V} [b]_1}{2} + \frac{\sqrt{C_V} \|D^2 V\|_\infty}{2} + \frac{C_V}{4}\right) \sqrt{C_V}\right)}\right),$$

where  $C_V$  is given by the assumption **(L<sub>V</sub>)**. For  $\beta \in (0, 1]$ , we introduce:

**(T<sub>β</sub>)** We choose a test function  $\varphi$  such that

- i)  $\varphi \in \mathcal{C}^{3, \beta}(\mathbb{R}^d, \mathbb{R})$ ,
  - ii)  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  is Lipschitz continuous,
- we further assume that there exist  $C_{V, \varphi} > 0$  s.t. for all  $x \in \mathbb{R}^d$ :
- iii)  $|\varphi(x)| \leq C_{V, \varphi} (1 + \sqrt{V(x)})$ .

**Remark 1** Under the assumption **(C0)** the equation (1.1) admits a unique non-explosive solution, cf [App09] (Theorem 6.2.9).

**Remark 2** The assumption **(GC)** is central for this paper. Note that the laws  $\mathcal{N}(0, I_r)$  and  $(\frac{1}{2}(\delta_1 + \delta_{-1}))^{\otimes r}$ , i.e. for Gaussian or symmetrized Bernoulli increments which are the most commonly used sequences for the sub-Gaussian innovations, satisfy **(GC)**. Moreover, inequality (1.8) yields that for all  $r \geq 0$ ,  $\mathbb{P}[|U_1| \geq r] \leq 2 \exp(-\frac{r^2}{2})$  (sub-Gaussian concentration of the innovation, see e.g. [BGL14]).

**Remark 3** The assumption **(L<sub>V</sub>)** together with **(C2)** ensure, following [Pan08a] (Proposition 1) the existence of at least one invariant distribution of the SDE (1.1). Note that this Lyapunov assumption **(L<sub>V</sub>)** is equivalent to the similar Lyapunov assumption for the continuous part of the equation (1.1). Indeed, using second order Taylor expansion, the fact that  $\pi(\cdot) = 0$  and  $\pi(|\cdot|^2) = r < \infty$ , we get that

$$\left| \int_{\mathbb{R}^r} (V(x + \kappa(x)y) - V(x))\pi(dy) \right| \leq \frac{\|\kappa\|_\infty^2 r \|D^2V\|_\infty}{2}.$$

Hence the condition iii) of **(L<sub>V</sub>)** is equivalent that the generator of the diffusion without jumps satisfies

$$\mathcal{A}V(x) \leq -\tilde{\alpha}_V V(x) + \tilde{\beta}_V, \quad (1.9)$$

with  $\tilde{\alpha}_V = \alpha_V$ ,  $\tilde{\beta}_V = \beta_V + \frac{\|\kappa\|_\infty^2 r \|D^2V\|_\infty}{2}$ . For more information about Lyapunov function existence, see e.g. [GP14].

Moreover, it is classic to see that this assumption constraints the drift coefficient  $b$  to be under a linear map. Indeed, this is the consequence of the fact that the Lyapunov function  $V$  has to be lower than the square norm, i.e. there exist constants  $K, \bar{c} > 0$  such that for all  $|x| \geq K$ ,  $|V(x)| \leq \bar{c}|x|^2$  and hence using ii) of **(L<sub>V</sub>)**  $|b(x)| \leq \sqrt{C_V \bar{c}}|x|$ .

**Remark 4** The assumption **(T<sub>β</sub>)** allows to substantially simplify the proof of our results. The condition iii) is natural for  $\varphi$  Lipschitz continuous, which is obviously lower than the square root of a quadratic function (potentially  $V$ ). Whilst the condition ii) is a direct consequence if  $\varphi$  is the solution of the Poisson equation:

$$\mathcal{A}\varphi = f, \quad (1.10)$$

where  $f \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R})$  s.t.  $\nu(f) = 0$ . If  $\sigma, \kappa \in \mathcal{C}_b^{1,\beta}(\mathbb{R}^d, \mathbb{R}^{d \times r})$ ,  $b \in \mathcal{C}^{1,\beta}(\mathbb{R}^d, \mathbb{R}^d)$  and  $\varphi \in \mathcal{C}^{3,\beta}(\mathbb{R}^d, \mathbb{R})$ , then both sides of the following identity:

$$\langle \nabla \varphi, b \rangle = f - \frac{1}{2} \text{Tr}(\Sigma D^2 \varphi) - \int_{\mathbb{R}^d} (\varphi(\cdot + \kappa(\cdot)y) - \varphi(\cdot)) \pi(dy),$$

are Lipschitz continuous.

From now on, we identify assumptions **(C0)**, **(C1)**, **(C2)**, **(GM)**, **(GC)**, **(L<sub>V</sub>)**, **(U)**, **(S)** and **(T<sub>β</sub>)** for some  $\beta \in (0, 1]$  to **(A)**. Except when explicitly indicated, we assume throughout the paper that assumption **(A)** is in force.

We suppose that the step sequence  $(\gamma_k)_{k \geq 1}$  is taken such that  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in (0, 1]$ . This pick yields for any  $\ell \geq 0$ ,  $\Gamma_n^{(\ell)} \asymp n^{1-\ell\theta}$  if  $\ell\theta < 1$ ,  $\Gamma_n^{(\ell)} \asymp \ln(n)$  if  $\ell\theta = 1$  and  $\Gamma_n^{(\ell)} \asymp 1$  if  $\ell\theta > 1$ .

The corner stone of our analysis is the fact that the jumps innovations  $(Z_n)_{n \in \mathbb{N}}$  inherit the Gaussian concentration property of  $(Y_n)_{n \in \mathbb{N}}$ :



**Proposition 1 (Gaussian concentration of the jumps innovation)** *Let  $g : \mathbb{R}^r \rightarrow \mathbb{R}$  uniform Lipschitz continuous function,  $\varepsilon \in (0, 1)$  and*

$$\rho(r) = \sqrt{r}(3 + r) + \frac{1}{8} + 4 \exp(\sqrt{r} + 1 + r/2). \quad (1.11)$$

*Then for all  $0 < \lambda < \frac{\varepsilon}{6[g]_1 \rho(r)}$  the following inequality holds for all  $n \in \mathbb{N}$*

$$\mathbb{E} \exp(\lambda g(Z_n)) \leq \exp(\lambda \mathbb{E} g(Z_n) + \frac{\lambda^2 \gamma_n (1 + r + \varepsilon) [g]_1^2}{2}). \quad (1.12)$$

**Remark 5** *Let us point out that the concentration inequality is only valid for  $\lambda$  on a compact set. This constraint is due to the difficulty to approximate a compound Poisson process which has actually a sub-exponential tail (and not a sub-Gaussian one).*

The proof of this proposition is given in Subsection 2.4.

## 1.4 Existing results

The natural next question concerns the rate of that convergence. In a Brownian diffusion framework, a Central Limit Theorem (CLT) was established by Lambertson and Pagès [LP02] for functions  $f$  of the form  $f - \nu(f) = \tilde{\mathcal{A}}\varphi$ , namely  $f - \nu(f)$  is a *coboundary* for  $\tilde{\mathcal{A}}$ , where  $\tilde{\mathcal{A}}$  denotes the continuous part of the infinitesimal generator (see (1.7) further). This choice of functions class comes from the characterization of the invariant distribution  $\nu$  by a solution in the distribution sense of the stationary Fokker-Planck equation:  $\tilde{\mathcal{A}}^*\nu = 0$  (where  $\tilde{\mathcal{A}}^*$  stands for the adjoint of  $\tilde{\mathcal{A}}$ ). In other words, for all functions  $\varphi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ , we have  $\nu(\tilde{\mathcal{A}}\varphi) = \int_{\mathbb{R}^d} \tilde{\mathcal{A}}\varphi(x)\nu(dx) = 0$ .

In [Pan08a], the author also provided the rate of convergence through a Central Limit Theorem (CLT) for the already mentioned general scheme:

**Theorem 1 (CLT)** *Under (C2), (U) and  $(\mathcal{L}_\nu)$ , if  $\mathbb{E}|Z_t|^{2p} < +\infty$  for  $p > 2$ , if  $\mathbb{E}[U_1^{\otimes 3}] = 0$ ,  $\mathbb{E}[|U_1|^{2p}] < +\infty$  and  $\lim_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} = 0$  then for all function  $\varphi \in \mathcal{C}^{3,1}(\mathbb{R}^d, \mathbb{R})$  we have the following results (with  $(\mathcal{L})$  denoting the weak convergence):*

$$\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \xrightarrow{(\mathcal{L})} \mathcal{N}(0, \sigma_\varphi^2), \quad (1.13)$$

with

$$\sigma_\varphi^2 := \int_{\mathbb{R}^d} (|\sigma^* \nabla \varphi|^2(x) + \int_{\mathbb{R}^d} |\varphi(x + \kappa(x)y) - \varphi(x)|^2 \pi(dy)) \nu(dx). \quad (1.14)$$

In the Brownian diffusion context ( $\kappa = 0$ ), under some confluence and non-degeneracy or regularity assumptions, in Chapter 3, we established suitable derivatives controls for the Poisson problem (e.g. Schauder estimates). With a compound Poisson process, we think that a similar analysis may work. It will be a future research. Let us mention [Pri10] for some Schauder estimates for Poisson equation, with a potential, associated with a SDE purely driven by stable processes but with a constant drift.

In [HMP18], the authors have established a non-asymptotic Gaussian concentration with  $\kappa = 0$  there are explicit sequences  $c_n \leq 1 \leq C_n$  converging to 1 such that for all  $n \in \mathbb{N}$ , for all  $a > 0$  and  $\gamma_k \asymp k^{-\theta}$ ,  $\theta \in (\frac{1}{3}, 1]$ ,

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty}\right), \quad (1.15)$$

which is our goal for a diffusion with jump contributions. Remark that, in Chapter 4 a non-asymptotic Gaussian concentration was established with the asymptotically best constants for a particular large deviation called ‘‘Gaussian deviations’’ therein. In other words, for  $a = o(\sqrt{\Gamma_n})$ :

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq C_n \exp\left(-c_n \frac{a^2}{2\nu(|\sigma^* \nabla\varphi|^2)}\right). \quad (1.16)$$

In this present work, we aim to obtain a Gaussian deviations bound like (1.15) for the scheme (1.2). To do so, we will perform the so-called martingales increments method which was exploited successfully by Frikha and Menozzi [FM12]. It was also the backbone of the analysis in Chapters 3 and 4. Here, we adapt these techniques for the stochastic differential equation (1.1) driven by the compound Poisson with Jump heigh sizes satisfying Gaussian concentration.

## 2 Main results

### 2.1 Result of non-asymptotic Gaussian concentration

Our main result is stated below.

**Theorem 2** *For  $\theta \in (\frac{1}{2+\beta}, 1]$ ,  $\beta \in (0, 1]$ , assume that **(A)** is in force. For all positive sequence  $(\chi_n)_{n \geq 1}$  with  $\lim_{n \rightarrow \infty} \chi_n = 0$ , there are two non-negative sequences  $(c_n)_{n \geq 1}$  and  $(C_n)_{n \geq 1}$  satisfying  $c_n \nearrow 1$  and  $C_n \searrow 1$  as  $n \rightarrow \infty$ , such that for all  $n \in \mathbb{N}$ ,  $a > 0$ , satisfying  $a \leq \chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}$ , the following bound holds:*

$$\mathbb{P}[|\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)| \geq a] \leq 2 C_n \exp\left(-c_n \frac{a^2}{2\sigma_\infty^2}\right),$$

where  $\sigma_\infty^2 := (1+r)\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2$  and  $C_n = \exp\left(\frac{[D^3\varphi]_\beta \|\sigma\|_\infty^{3+\beta} \mathbb{E}[|U_1|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)} \frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} + p_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}\right)$  for  $p_n \geq 1$  such that  $p_n \rightarrow_n +\infty$  and  $p_n \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow_n 0$ .

The proof of Theorem 2 is given in Section 2.3.

**Remark 6** *Note that for all  $\theta \in (\frac{1}{2+\beta}, 1]$ ,  $\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} = +\infty$ , then we can choose  $\chi_n$  s.t.  $\chi_n \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \rightarrow +\infty$ . In other words, we can pick  $a = a(n) \rightarrow_n +\infty$ . We have unsurprisingly that  $\sigma_\varphi^2 \leq \sigma_\infty^2$ , where  $\sigma_\varphi^2$  is the asymptotic variance of  $\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi)$  defined in (1.14). Moreover, the difficulty to adapt a Gaussian Concentration result to compound Poisson process yields that the upper-bound variance  $\sigma_\infty^2$  depends on the dimension.*

## 2.2 Strategy

For the analysis of  $\nu_n(\mathcal{A}\varphi)$ , we will first perform an appropriate Taylor expansion (equation (2.3) below). An expansion of this kind is standard in this context, and analogous decompositions were already used in Chapter 3, Chapter 4 and in [Pan08b], [Pan08a] with a jump component. It can be viewed as a kind of Itô formula for Euler scheme, because it permits to write the difference  $\varphi(X_n) - \varphi(X_0)$  as a sum of a martingale, a term involving the generator and a remainder term. Recall that  $\mathcal{F}_k = \sigma(X_0, (U_j, Z_j)_{j \in \llbracket 1, k \rrbracket})$ ,  $k \in \mathbb{N}^*$ . Let us define the contributions of the decomposition of  $\nu_n(\mathcal{A}\varphi)$  in the following lemma.

$$\begin{aligned}
 \psi_k^\varphi(X_{k-1}, U_k) &:= \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\
 &\quad + \gamma_k \int_0^1 (1-t) \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t \sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right. \\
 &\quad \left. - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right) dt, \\
 \Delta_k^\varphi(X_{k-1}, U_k) &:= \psi_k^\varphi(X_{k-1}, U_k) - \mathbb{E}[\psi_k^\varphi(X_{k-1}, U_k) | \mathcal{F}_{k-1}], \\
 \tilde{\Delta}_k^\varphi(X_{k-1}, Z_k) &:= \varphi(X_{k-1} + \kappa_{k-1} Z_k) - \varphi(X_{k-1}) \\
 &\quad - \gamma_k \int_{\mathbb{R}^r} [\varphi(X_{k-1} + \kappa_{k-1} y) - \varphi(X_{k-1})] \pi(dy). \tag{2.1}
 \end{aligned}$$

Moreover, we define the remainder contributions in the decomposition of  $\nu_n(\mathcal{A}\varphi)$ .

$$\begin{aligned}
 D_{2,b}^{k,\varphi}(X_{k-1}) &:= \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t \gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt, \\
 D_{2,\Sigma}^{k,\varphi}(X_{k-1}) &:= \frac{\gamma_k}{2} \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1} \right), \\
 D_j^{k,\varphi}(X_{k-1}, U_k, Z_k) &:= \varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k) \\
 &\quad - (\varphi(X_{k-1} + \kappa_{k-1} Z_k) - \varphi(X_{k-1})). \tag{2.2}
 \end{aligned}$$

**Lemma 1 (Local decomposition of the empirical measure)** *For all  $\varphi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ ,  $k \in \mathbb{N}^*$  the following decomposition holds:*

$$\varphi(X_k) - \varphi(X_{k-1}) = \gamma_k \mathcal{A}\varphi(X_{k-1}) + \Delta_k^\varphi(X_{k-1}, U_k) + \tilde{\Delta}_k^\varphi(X_{k-1}, Z_k) + \mathbf{R}_k^\varphi(X_{k-1}, U_k, Z_k), \tag{2.3}$$

where

$$\mathbf{R}_k^\varphi(X_{k-1}, U_k, Z_k) := D_{2,b}^{k,\varphi}(X_{k-1}) + D_{2,\Sigma}^{k,\varphi}(X_{k-1}) + D_j^{k,\varphi}(X_{k-1}, U_k, Z_k) + \mathbb{E}[\psi_k^\varphi(X_{k-1}, U_k) | \mathcal{F}_{k-1}]. \tag{2.4}$$

Furthermore, we have the following properties:

i) For all  $k \in \mathbb{N}^*$ , the functions  $u \mapsto \Delta_k^\varphi(X_{k-1}, u)$  and  $z \mapsto \tilde{\Delta}_k^\varphi(X_{k-1}, z)$  are Lipschitz, satisfying

$$\begin{aligned}
 [\Delta_k^\varphi(X_{k-1}, \cdot)]_1 &\leq \sqrt{\gamma_k} \|\sigma_{k-1}\| \|\nabla \varphi\|_\infty \leq \sqrt{\gamma_k} \|\sigma\|_\infty \|\nabla \varphi\|_\infty, \\
 [\tilde{\Delta}_k^\varphi(X_{k-1}, \cdot)]_1 &\leq \|\kappa_{k-1}\| \|\nabla \varphi\|_\infty \leq \|\kappa\|_\infty \|\nabla \varphi\|_\infty.
 \end{aligned}$$

ii) For all  $k \in \mathbb{N}^*$ ,  $\Delta_k^\varphi(X_{k-1}, U_k)$  and  $\tilde{\Delta}_k^\varphi(X_{k-1}, Z_k)$  are martingale increments with respect to  $\mathcal{F}_k$ , namely:

$$\mathbb{E}[\Delta_k^\varphi(X_{k-1}, U_k) | \mathcal{F}_{k-1}] = 0, \quad \mathbb{E}[\tilde{\Delta}_k^\varphi(X_{k-1}, Z_k) | \mathcal{F}_{k-1}] = 0.$$

The proof of Lemma 1 is given in Section 4. Now we introduce the martingales associated to these martingale increments:

$$M_n^\varphi := \sum_{k=1}^n \Delta_k^\varphi(X_{k-1}, U_k), \quad \tilde{M}_n^\varphi := \sum_{k=1}^n \tilde{\Delta}_k^\varphi(X_{k-1}, Z_k). \quad (2.5)$$

Summing (2.3) over  $k$  we obtain the following global decomposition of the empirical measure:

$$\nu_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n} (M_n^\varphi + \tilde{M}_n^\varphi + \mathfrak{R}_n^\varphi), \quad (2.6)$$

where we denoted

$$\mathfrak{R}_n^\varphi := \sum_{k=1}^n \mathbf{R}_k^\varphi(X_{k-1}, U_k, Z_k) - (\varphi(X_n) - \varphi(X_0)). \quad (2.7)$$

Using the definition (2.2) we can write  $\mathfrak{R}_n^\varphi = -L_n^\varphi + D_{2,b,n}^\varphi + D_{2,\Sigma,n}^\varphi + D_{j,n}^\varphi + \bar{G}_n^\varphi$ , with

$$\begin{aligned} L_n^\varphi &:= \varphi(X_n) - \varphi(X_0), \quad D_{2,b,n}^\varphi := \sum_{k=1}^n D_{2,b}^{k,\varphi}(X_{k-1}), \quad D_{2,\Sigma,n}^\varphi := \sum_{k=1}^n D_{2,\Sigma}^{k,\varphi}(X_{k-1}), \\ D_{j,n}^\varphi &:= \sum_{k=1}^n D_j^{k,\varphi}(X_{k-1}, U_k, Z_k), \quad \bar{G}_n^\varphi := \sum_{k=1}^n \mathbb{E}[\psi_k^\varphi(X_{k-1}, U_k) | \mathcal{F}_{k-1}]. \end{aligned} \quad (2.8)$$

In the proof of Theorem 2, we need some key results stated below. The proofs of all these statements are postponed to Section 4.

The main contribution in the decomposition (2.6) is given by the martingales  $M_n^\varphi$  and  $\tilde{M}_n^\varphi$ . Their analysis is given with the help of the Gaussian Concentration inequality (1.8) and (1.12), through the following lemma:

**Lemma 2 (Concentration of the martingale increments)** *Let  $\Delta_n^\varphi$  and  $\tilde{\Delta}_n^\varphi$  given by (2.1).*

i) For all  $\Lambda > 0$  we have

$$\mathbb{E} \left[ \exp \left( -\frac{\Lambda}{\Gamma_n} \Delta_n^\varphi(X_{n-1}, U_n) \right) \middle| \mathcal{F}_{n-1} \right] \leq \exp \left( \gamma_n \|\sigma\|_\infty^2 \|\nabla \varphi\|_\infty^2 \frac{\Lambda^2}{2\Gamma_n^2} \right).$$

ii) For all  $0 < \varepsilon < 1$ ,  $n \in \mathbb{N}^*$ , for all  $\Lambda > 0$  s.t.  $\frac{\Lambda}{\Gamma_n} < \frac{\varepsilon}{6\|\kappa\|_\infty \|\nabla \varphi\|_\infty \rho(r)}$ , where  $\rho(r)$  is defined in (1.11), we have

$$\mathbb{E} \left[ \exp \left( -\frac{\Lambda}{\Gamma_n} \tilde{\Delta}_n^\varphi(X_{n-1}, Z_n) \right) \middle| \mathcal{F}_{n-1} \right] \leq \exp \left( \gamma_n \|\kappa\|_\infty^2 \|\nabla \varphi\|_\infty^2 (1 + r + \varepsilon) \frac{\Lambda^2}{2\Gamma_n^2} \right).$$

Now we formulate several propositions and lemmas that are used to control the components of the remainder term  $\mathfrak{R}_n^\varphi$ . The following proposition is the counterpart to the jumps diffusion of the useful Proposition 1 in Chapter 3.

**Proposition 2** *Under (A), there is a constant  $c_V := c_V((\mathbf{A})) > 0$  such that for all  $\lambda \in [0, c_V]$ :*

$$I_V^{\frac{1}{2}} := \sup_{n \geq 0} \mathbb{E} \left[ \exp \left( \lambda \sqrt{V}(X_n) \right) \right] < +\infty. \quad (2.9)$$

**Remark 7** *In particular, we easily see that for all  $\lambda \in [0, c_V]$  and  $\xi \in [0, \frac{1}{2}]$ :*

$$I_V^\xi := \sup_{n \geq 0} \mathbb{E} \left[ \exp \left( \lambda V(X_n)^\xi \right) \right] < +\infty. \quad (2.10)$$

*Note that for  $\kappa = 0$  (purely continuous case), the integrability of  $\exp(\lambda V(X_n)^\xi)$  is available until  $\xi = 1$  (see Proposition 1 in [HMP18]). The loss of integrability is the consequence of the bound condition over  $\lambda$  in the Gaussian Concentration result of Proposition 1.*

We have the following results for the initial term appearing in (2.3) which is handled thanks to the below result.

**Lemma 3 (Initial term)** *For all  $\Lambda > 0$  s.t.  $\frac{\Lambda}{\Gamma_n} < \frac{c_V}{2C_{V,\varphi}}$ :*

$$\mathbb{E} \exp \left( \Lambda \frac{|L_n^\varphi|}{\Gamma_n} \right) \leq \exp \left( 2C_{V,\varphi} \frac{\Lambda}{\Gamma_n} \right) (I_V^{\frac{1}{2}})^{\frac{2C_{V,\varphi}\Lambda}{c_V\Gamma_n}}$$

*with  $c_V, I_V^{\frac{1}{2}}$  given in Proposition 2.*

Next the last remainders are controlled as following:

**Lemma 4 (Remainders)** *For all  $\Lambda > 0$  s.t.  $\frac{\Lambda}{\Gamma_n} < \frac{2c_V}{(\|\nabla\varphi\|_\infty[b]_1 + [\langle \nabla\varphi, b \rangle]_1)\sqrt{C_V}} \frac{1}{\Gamma_n^{(2)}}$ :*

$$\mathbb{E} \exp \left( \frac{\Lambda}{\Gamma_n} |D_{2,b,n}^\varphi| \right) \leq (I_V^{1/2})^{\frac{\Lambda(\|\nabla\varphi\|_\infty[b]_1 + [\langle \nabla\varphi, b \rangle]_1)\sqrt{C_V}\Gamma_n^{(2)}}{2c_V\Gamma_n}}. \quad (2.11)$$

*We also have, for all  $\Lambda > 0$  s.t.  $\frac{\Lambda}{\Gamma_n} \leq \frac{c_V}{\|\sigma\|_\infty^2 \|D^3\varphi\|_\infty \sqrt{C_V}} \frac{1}{\Gamma_n^{(2)}}$ :*

$$\mathbb{E} \exp \left( \frac{\Lambda}{\Gamma_n} |D_{2,\Sigma,n}^\varphi| \right) \leq (I_V^{1/2})^{\frac{\|\sigma\|_\infty^2 \|D^3\varphi\|_\infty C_V^{\frac{1}{2}} \Lambda \Gamma_n^{(2)}}{2c_V\Gamma_n}}. \quad (2.12)$$

**Lemma 5 (Bounds for the Conditional expectations)** *With the notations (2.8), and for  $\theta \in (\frac{1}{2+\beta}, 1]$ , we have that*

$$\frac{|\bar{G}_n^\varphi|}{\sqrt{\Gamma_n}} \stackrel{a.s.}{\leq} \alpha_n := \frac{[\varphi^{(3)}]_\beta \|\sigma\|_\infty^{(3+\beta)} \mathbb{E}[|U_1|^{3+\beta}] \Gamma_n^{\frac{3+\beta}{2}}}{(1+\beta)(2+\beta)(3+\beta) \sqrt{\Gamma_n}} \xrightarrow[n]{n} 0.$$

**Remark 8** *The strongest condition over  $\theta$  comes from this remainder term. Indeed, for  $\theta < \frac{2}{3+\beta}$ ,  $\frac{\Gamma_n^{(\frac{3+\beta}{2})}}{\sqrt{\Gamma_n}} \asymp n^{\frac{1-(2+\beta)\theta}{2}}$  which goes to 0 if and only if  $\theta > \frac{1}{2+\beta}$ . Whilst for the other remainders, for  $\theta < \frac{1}{2}$ , we need to have  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \asymp n^{\frac{1-3\theta}{2}} \rightarrow_n 0$  which is implied by  $\theta > \frac{1}{3}$ .*

Now, let us deal with the remainder term  $D_{j,n}$  due to the jump vector  $(Z_k)_{k \geq 1}$ .

**Lemma 6 (Remainder term due to the jumps)** *If  $0 < \frac{\Lambda}{\Gamma_n} < \frac{1}{12\|\kappa\|_\infty\|\nabla\varphi\|_\infty\rho(r)}$ , then we have:*

$$\mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} D_{j,n}^\varphi \right) \right] \leq \exp \left( \left( \frac{\Lambda}{\sqrt{\Gamma_n}} + \frac{\Lambda^2}{\Gamma_n} \right) e_n \right), \quad (2.13)$$

where we recall that  $e_n = e_n(\mathbf{A})$ ,  $n \geq 1$ , is a sequence such that  $e_n \rightarrow_n 0$ .

The proof of this lemma is one of the most intricate of this article, we the decided to postponed it to the end of Section 4.

## 2.3 Proof of our main result

**Proof: Theorem 2**

Through the following analysis, we deal with  $\mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a]$ . The term  $\mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \leq -a]$  can be handled readily by symmetry.

From notations introduced in (2.6),  $\nu_n(\mathcal{A}\varphi) = -\frac{1}{\Gamma_n}(\mathfrak{R}_n^\varphi + M_n^\varphi + \widetilde{M}_n^\varphi)$ . The idea is now to write for  $a, \lambda > 0$ :

$$\begin{aligned} \mathbb{P}[\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi) \geq a] &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E} \left[ \exp\left(-\frac{\lambda}{\Gamma_n}(\mathfrak{R}_n^\varphi + M_n^\varphi + \widetilde{M}_n^\varphi)\right) \right] \\ &\leq \exp\left(-\frac{a\lambda}{\sqrt{\Gamma_n}}\right) \mathbb{E} \left[ \exp\left(-\frac{q\lambda}{\Gamma_n}(M_n^\varphi + \widetilde{M}_n^\varphi)\right) \right]^{1/q} \mathbb{E} \left[ \exp\left(\frac{p\lambda}{\Gamma_n}|\mathfrak{R}_n^\varphi|\right) \right]^{1/p}, \end{aligned} \quad (2.14)$$

for  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $p, q > 1$ . We will choose later  $p = p(n) \rightarrow_n +\infty$  slowly enough, which implies that  $q = q(n) \rightarrow 1$ . Let  $\lambda > 0$ . Recall that

$$\mathfrak{R}_n^\varphi = -L_n^\varphi + D_{2,b,n}^\varphi + D_{2,\Sigma,n}^\varphi + D_{j,n}^\varphi + \bar{G}_n^\varphi$$

By Cauchy-Schwarz inequality, we obtain:

$$\begin{aligned} \mathbb{E} \left[ \exp\left(\frac{p\lambda}{\Gamma_n}|\mathfrak{R}_n^\varphi|\right) \right]^{1/p} &\leq \left( \mathbb{E} \exp\left(\frac{2p\lambda}{\Gamma_n}|L_n^\varphi|\right) \right)^{\frac{1}{2p}} \left( \mathbb{E} \exp\left(\frac{4p\lambda}{\Gamma_n}|\bar{G}_n^\varphi|\right) \right)^{\frac{1}{4p}} \\ &\left( \mathbb{E} \exp\left(\frac{8p\lambda}{\Gamma_n}|D_{2,b,n}^\varphi|\right) \right)^{\frac{1}{8p}} \left( \mathbb{E} \exp\left(\frac{16p\lambda}{\Gamma_n}|D_{2,\Sigma,n}^\varphi|\right) \right)^{\frac{1}{16p}} \left( \mathbb{E} \exp\left(\frac{16p\lambda}{\Gamma_n}|D_{j,n}^\varphi|\right) \right)^{\frac{1}{16p}}. \end{aligned} \quad (2.15)$$

We recall that all the long of our analysis,  $C > 0$  denotes a generic constant,  $(\mathcal{R}_n)_{n \geq 1}$  and  $(e_n)_{n \geq 1}$  are generic non-negative sequences, depending on coefficients of assumption  $(\mathbf{A})$ ,

which may change from line to line, such that  $\lim_{n \rightarrow \infty} \mathcal{R}_n = 1$ ,  $\lim_{n \rightarrow \infty} e_n = 0$ . For the term associated with  $L_n$  in (2.15), if  $\frac{2p\lambda}{\Gamma_n} < \frac{c_V}{2C_{V,\varphi}}$ , by Lemma 3 we can write:

$$\left( \mathbb{E} \exp \left( 2p\lambda \frac{|L_n^\varphi|}{\Gamma_n} \right) \right)^{\frac{1}{2p}} \leq \exp \left( 2C_{V,\varphi} \frac{\lambda}{\Gamma_n} \right) (I_V^{\frac{1}{2}})^{\frac{2C_{V,\varphi}\lambda}{c_V\Gamma_n}} = \exp \left( C \frac{\lambda}{\Gamma_n} \right). \quad (2.16)$$

From Lemma 5, with  $\alpha_n$  defined there and by Young inequality we obtain:

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda}{\Gamma_n} |\bar{G}_n^\varphi| \right) \right)^{\frac{1}{4p}} \leq \exp \left( \frac{\lambda}{\sqrt{\Gamma_n}} \alpha_n \right) \leq \exp \left( \frac{\lambda^2}{\Gamma_n 2p} + \frac{\alpha_n^2 p}{2} \right) = \mathcal{R}_n \exp \left( \frac{\lambda^2}{\Gamma_n} e_n \right). \quad (2.17)$$

In the last equality,  $\mathcal{R}_n = \exp(\alpha_n^2 p_n / 2)$  and  $e_n = 1/2p_n$ . Recall that  $\alpha_n \rightarrow_n 0$ . We need to choose  $p = p_n \rightarrow_n +\infty$ . We choose  $p_n$  such that  $p_n \alpha_n^2 \rightarrow_n 0$ .

For the term involving  $D_{2,\Sigma,n}^\varphi$  from (2.15), if  $\frac{16p\lambda}{\Gamma_n} \leq \frac{2c_V}{\Gamma_n^{(2)} \|\sigma\|_\infty^2 \|D^3\varphi\|_\infty \sqrt{C_V}}$ , using Lemma 4, we can write

$$\left( \mathbb{E} \exp \left( \frac{16p\lambda}{\Gamma_n} |D_{2,\Sigma,n}^\varphi| \right) \right)^{\frac{1}{16p}} \leq (I_V^{1/2})^{\frac{\|\sigma\|_\infty^2 \|D^3\varphi\|_\infty C_V^{\frac{1}{2}} \lambda \Gamma_n^{(2)}}{c_V \Gamma_n}} = \exp \left( C \frac{\lambda \Gamma_n^{(2)}}{\Gamma_n} \right) = \exp \left( C \frac{\lambda}{\sqrt{\Gamma_n}} e_n \right), \quad (2.18)$$

where in the last equality we take  $e_n = \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}$  and recall that for all  $\theta \in (\frac{1}{3}, 1]$ ,  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow_n 0$ .

For the remainder depending on  $D_{2,b,n}^\varphi$  from (2.15), if  $\frac{4p\lambda}{\Gamma_n} < \frac{2c_V}{\Gamma_n^{(2)} (\|\nabla\varphi\|_\infty [b]_1 + \langle \nabla\varphi, b \rangle_1) \sqrt{C_V}}$ , thanks to Lemma 4 we have again with  $e_n = \frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}}$  :

$$\left( \mathbb{E} \exp \left( \frac{4p\lambda}{\Gamma_n} |D_{2,b,n}^\varphi| \right) \right)^{\frac{1}{4p}} \leq (I_V^{1/2})^{\frac{\lambda (\|\nabla\varphi\|_\infty [b]_1 + \langle \nabla\varphi, b \rangle_1) \sqrt{C_V} \Gamma_n^{(2)}}{2c_V \Gamma_n}} = \exp \left( C \frac{\lambda \Gamma_n^{(2)}}{\Gamma_n} \right) = \exp \left( C \frac{\lambda}{\sqrt{\Gamma_n}} e_n \right). \quad (2.19)$$

Finally, Lemma 6 yields that if  $0 < \frac{16p\lambda}{\Gamma_n} < \frac{1}{12\|\kappa\|_\infty \|\nabla\varphi\|_\infty \rho(r)}$ , then:

$$\mathbb{E} \left[ \exp \left( \frac{16p\lambda}{\Gamma_n} D_{j,n}^\varphi \right) \right] \leq \exp \left( \frac{\lambda}{\sqrt{\Gamma_n}} e_n + \frac{\lambda^2}{\Gamma_n} e_n \right). \quad (2.20)$$

We gather (2.16), (2.17), (2.18), (2.19) and (2.20) into (2.15) and finally from (2.14) we obtain:

$$\mathbb{P} \left[ \sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a \right] \leq \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \right) \mathbb{E} \left[ \exp \left( -\frac{q\lambda}{\Gamma_n} (M_n^\varphi + \widetilde{M}_n^\varphi) \right) \right]^{\frac{1}{q}} \exp \left( \left( \frac{\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2}{\Gamma_n} \right) e_n \right) \mathcal{R}_n. \quad (2.21)$$

Now, let us control the martingale terms thanks to Lemma 2. Let  $0 < \varepsilon < 1$ , and  $\lambda > 0$  s.t.  $q\lambda/\Gamma_n < \frac{C_\varepsilon}{6\|\kappa\|_\infty \|\nabla\varphi\|_\infty \rho(r)}$ , we recall that  $\rho(r)$  is defined in (1.11).

Thanks to Lemma 2 and the independence of  $Z_n$  and  $U_n$  conditionally to  $\mathcal{F}_{n-1}$  we can write

$$\begin{aligned} & \mathbb{E} \exp \left( -\frac{q\lambda}{\Gamma_n} (M_n^\varphi + \widetilde{M}_n^\varphi) \right) \\ &= \mathbb{E} \left[ \exp \left( -\frac{q\lambda}{\Gamma_n} (M_{n-1}^\varphi + \widetilde{M}_{n-1}^\varphi) \right) \mathbb{E} \left[ \exp \left( -\frac{q\lambda}{\Gamma_n} (\Delta_n^\varphi(X_{n-1}, U_n) + \widetilde{\Delta}_n^\varphi(X_{n-1}, Z_n)) \right) \middle| \mathcal{F}_{n-1} \right] \right] \\ &\leq \mathbb{E} \left[ \exp \left( -\frac{q\lambda}{\Gamma_n} (M_{n-1}^\varphi + \widetilde{M}_{n-1}^\varphi) \right) \exp \left( \frac{q^2 \lambda^2 \gamma_n}{2\Gamma_n^2} (\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 + \|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 (1+r+\varepsilon)) \right) \right]. \end{aligned}$$

By induction we obtain

$$\begin{aligned} \mathbb{E} \exp \left( -\frac{q\lambda}{\Gamma_n} (M_n^\varphi + \widetilde{M}_n^\varphi) \right) &\leq \exp \left( \left( \frac{q^2\lambda^2}{2\Gamma_n^2} (\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 (1+r+\varepsilon) + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2) \right) \sum_{k=1}^n \gamma_k \right) \\ &= \exp \left( \frac{q^2\lambda^2}{2\Gamma_n} (\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 (1+r+\varepsilon) + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2) \right). \end{aligned}$$

Plugging this inequality into (2.21) yields:

$$\begin{aligned} &\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \\ &\leq \exp \left( -\frac{a\lambda}{\sqrt{\Gamma_n}} \right) \exp \left( \frac{q\lambda^2}{2\Gamma_n} (\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 (1+r+\varepsilon) + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2) \right) \\ &\quad \times \exp \left( \left( \frac{\lambda}{\sqrt{\Gamma_n}} + \frac{\lambda^2}{\Gamma_n} \right) e_n \right) \mathcal{R}_n. \end{aligned} \quad (2.22)$$

Next, we optimize the polynomial  $-\frac{a\lambda}{\sqrt{\Gamma_n}} + \frac{q\lambda^2}{2\Gamma_n} (\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 (1+r+\varepsilon) + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2)$  over  $\lambda$  which leads to consider:

$$\lambda := \lambda_n := \frac{a\sqrt{\Gamma_n}}{q(\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 (1+r+\varepsilon) + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2)}. \quad (2.23)$$

We check first the assumptions over  $\lambda$  in all lemmas that we used in this proof. In (2.16) and (2.20) we need  $p\lambda_n/\Gamma_n < C$ . And for (2.18) and (2.19) we need  $p\lambda_n/\Gamma_n < \frac{C}{\Gamma_n^{(2)}}$ . Finally  $q\lambda_n/\Gamma_n < C\varepsilon_n$  is required to apply Lemma 2. We recall that we will choose  $p \rightarrow \infty$  and  $q \rightarrow 1$  and finally  $\varepsilon \rightarrow 0$ .

We recall also that from the statement of the theorem  $a = a(n)$  can depends of  $n$  in such a way that  $\frac{a}{\sqrt{\Gamma_n}} \leq \frac{\chi_n}{\Gamma_n^{(2)}} \rightarrow_n 0$ . But if  $q \rightarrow 1$ , for  $n$  big enough  $\frac{\lambda_n}{\Gamma_n} = \frac{a}{\sqrt{\Gamma_n}} < \frac{\chi_n}{\Gamma_n^{(2)}} \rightarrow 0$ . Hence, the condition

$$\frac{p\lambda_n}{\Gamma_n} = \frac{pa}{\sqrt{\Gamma_n}} < \frac{p\chi_n}{\Gamma_n^{(2)}} < C/\Gamma_n^{(2)} \quad (2.24)$$

has to be satisfied. Let us calibrate  $p = p(n) \rightarrow \infty$  depending on  $\chi_n \rightarrow 0$  s.t.  $\limsup_n p\chi_n < C$ . This pick of  $p$  yields (2.24). We can also choose, for  $C > 0$  large enough  $\varepsilon_n = C \frac{\chi_n}{\Gamma_n^{(2)}}$  such that all conditions over  $\lambda_n, p, \varepsilon$  are satisfied with these choices.

The inequality (2.22) yields then for  $\lambda = \lambda_n$ :

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq \mathcal{R}_n \exp \left( -\frac{\tilde{c}_n a^2}{2((1+r)\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2)} + \frac{\lambda_n}{\sqrt{\Gamma_n}} e_n \right),$$

with  $\tilde{c}_n = \frac{(1+r)\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}{q((1+r+\varepsilon)\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2)} \xrightarrow{n \rightarrow +\infty} 1$ , and  $\frac{\lambda_n}{\sqrt{\Gamma_n}} e_n \leq C a e_n$ .

If  $a \leq 1$ , we take  $C_n = \mathcal{R}_n \exp(C a e_n) \xrightarrow{n \rightarrow +\infty} 1$ , otherwise if  $a > 1$  then we set  $c_n = 1 - \frac{C e_n}{a^2} \xrightarrow{n \rightarrow +\infty} 1$ . In any case, we write the result:

$$\mathbb{P}[\sqrt{\Gamma_n} \nu_n(\mathcal{A}\varphi) \geq a] \leq C_n \exp \left( -\frac{c_n a^2}{2((1+r)\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2 + \|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2)} \right).$$

□



## 2.4 Proof of the Gaussian property of the jump innovation

### Proof: Proposition 1

Suppose first that  $g : \mathbb{R}^r \rightarrow \mathbb{R}$  is Lipschitz continuous with  $[g]_1 \leq 1$ . The case of  $[g]_1 > 1$  follows by considering  $\tilde{\lambda} = \lambda[g]_1$  and  $\tilde{g} = \frac{g}{[g]_1}$ . We suppose w.l.o.g. that  $g(0) = 0$ . We recall that thank to the definition (1.3) the law of  $Z_n$  is the same that the law of  $B_n Y$ , where  $B_n$  is a Bernoulli variable with parameter  $\gamma_n$ , independent of the random vector  $Y$  with distribution  $\pi$  on  $\mathcal{B}(\mathbb{R}^r)$ . We will establish first that for all  $\varepsilon \in (0, 1)$  and  $0 < \lambda < \frac{\varepsilon}{\rho(r)}$  (see (1.11)) we have

$$\mathbb{E} \exp(\lambda g(Z_n)) \leq \exp(\lambda \mathbb{E} g(Z_n) + \frac{\lambda^2 \gamma_n (1 + (\mathbb{E} g(Y))^2 + \varepsilon)}{2}). \quad (2.25)$$

Denote for this proof  $m_g := \mathbb{E} g(Y)$ . Using **(GC)** property of  $Y$  we can write

$$\mathbb{E} \exp(\lambda g(Z_n)) = \gamma_n \mathbb{E} \exp(\lambda g(Y)) + (1 - \gamma_n) \leq \gamma_n \exp(\lambda m_g + \lambda^2/2) + (1 - \gamma_n).$$

Denote

$$\Delta_n^{\text{exp}} := \gamma_n \exp(\lambda m_g + \frac{\lambda^2}{2}) + (1 - \gamma_n) - \exp(\lambda \gamma_n m_g + \lambda^2 \frac{\gamma_n (1 + m_g^2 + \varepsilon)}{2}). \quad (2.26)$$

Here, the second exponential corresponds to the right hand side in (2.25). We will show that  $\Delta_n^{\text{exp}} < 0$ . Indeed, let us develop the difference  $\Delta_n^{\text{exp}}$  by power series expansion:

$$\begin{aligned} \Delta_n^{\text{exp}} &= \gamma_n (\lambda m_g + \frac{\lambda^2}{2}) - (\gamma_n \lambda m_g + \lambda^2 \frac{\gamma_n (1 + m_g^2 + \varepsilon)}{2}) \\ &+ \frac{1}{2} \gamma_n (\lambda m_g + \frac{\lambda^2}{2})^2 - \frac{1}{2} (\gamma_n \lambda m_g + \frac{\gamma_n \lambda^2 (1 + m_g^2 + \varepsilon)}{2})^2 + Q(\lambda) \\ &= -\frac{\gamma_n \lambda^2}{2} (\varepsilon + \gamma_n m_g^2) + \frac{\gamma_n \lambda^3}{2} (m_g (1 - \gamma_n (1 + m_g^2 + \varepsilon))) \\ &+ \frac{1}{8} \gamma_n \lambda^4 (1 - \gamma_n (1 + m_g^2 + \varepsilon))^2 + Q(\lambda), \end{aligned}$$

where

$$Q(\lambda) := \gamma_n \sum_{k \geq 3} \frac{1}{k!} (\lambda m_g + \frac{\lambda^2}{2})^k - \sum_{k \geq 3} \frac{1}{k!} (\gamma_n \lambda m_g + \frac{\gamma_n \lambda^2 (1 + m_g^2 + \varepsilon)}{2})^k.$$

In particular, using  $\gamma_n \leq 1$  from **(S)**, and  $\varepsilon < 1$ , we can roughly estimate:

$$\Delta_n^{\text{exp}} \leq -\frac{\gamma_n \lambda^2}{2} \varepsilon + \frac{\gamma_n \lambda^3}{2} (|m_g| (3 + m_g^2)) + \frac{1}{8} \gamma_n \lambda^4 + Q(\lambda), \quad (2.27)$$

Because  $g$  is 1-Lipschitz continuous and from the assumption **(GM)** we obtain:

$$m_g^2 = |\mathbb{E} g(Y_n)|^2 \leq \mathbb{E} |g(Y_n)|^2 = \mathbb{E} |g(Y_n) - g(0)|^2 \leq [g]_1 \mathbb{E} \|Y_n\|^2 = [g]_1 \sum_{k=1}^r |Y^k|^2 \leq r.$$

Using again  $\gamma_n \leq 1$ ,  $\lambda \leq 1$ ,  $\varepsilon \leq 1$  we get

$$Q(\lambda) \leq \gamma_n \sum_{k \geq 3} \frac{\lambda^k}{k!} (\sqrt{r} + \frac{1}{2})^k + \gamma_n \sum_{k \geq 3} \frac{\lambda^k \gamma_n^{k-1}}{k!} (\sqrt{r} + \frac{2+r}{2})^k \leq 2\gamma_n \lambda^3 \exp(\sqrt{r} + 1 + r/2)$$

Thus combined with (2.27) gives :

$$\begin{aligned} \Delta_n^{\text{exp}} &\leq -\frac{\gamma_n \lambda^2}{2} \varepsilon + \frac{\gamma_n \lambda^3}{2} (|m_g|(3 + m_g^2)) + \frac{1}{8} \gamma_n \lambda^4 + 2\gamma_n \lambda^3 \exp(\sqrt{r} + 1 + r/2) \\ &\leq \gamma_n \lambda^2 / 2 \{ -\varepsilon + \lambda \sqrt{r} (3 + r) + \frac{1}{8} \lambda^2 + 4\lambda \exp(\sqrt{r} + 1 + r/2) \} \leq \gamma_n \lambda^2 / 2 \{ -\varepsilon + \lambda \rho(r) \}. \end{aligned}$$

which is negative if  $\lambda \rho(r) < \varepsilon$ , with  $\rho(r)$  defined in (1.11). This proves the (2.25). Together with the inequality  $m_g^2 \leq r$  this proves the concentration inequality in the case  $[g]_1 \leq 1$ .  $\square$

**Remark 9** Note that  $r \rightarrow \rho(r)$  is increasing, hence the condition we need to put on  $\lambda$  in order to propagate the Gaussian concentration from  $Y \in \mathbb{R}^r$  to  $Z$  became stronger if the dimension  $r$  increase.

### 3 Exponential integrability of the square root of Lyapunov function.

In Chapter 3, the exponential moments of the Lyapunov function was used to control the remainder terms of the decomposition of the empirical measure. In this article, we also use the Lyapunov function for this purpose. But our framework yields more constraints over the analysis. Namely, we cannot directly use  $\exp(CV_n)$  which is not *a priori* integrable. Indeed, let us consider the Compound Poisson process  $\tilde{Z}_t := \sum_{k=1}^{N_t} \tilde{Y}_k$  where  $(\tilde{Y}_k)_{k \in \mathbb{N}}$  is an i.i.d. sequence of a standard normal variables independent of  $N_t$  which follows a Poisson law, which is the typical jump random variables that we aim to approximate. Conditionally to  $N_t$ ,  $\tilde{Z}_t \sim \mathcal{N}(0, N_t)$ . So if we choose the Lyapunov function to be the standard quadratic map, i.e. for all  $x \in \mathbb{R}^d$ ,  $V(x) = |x|^2 + 1$ . We obtain *in fine*:

$$\mathbb{E}[\exp(\lambda V(\tilde{Z}_t))] = e^\lambda \mathbb{E}[\mathbb{E}[\exp(\lambda \tilde{Z}_t^2) | N_t]] \geq e^\lambda \mathbb{E} \left[ \int_{\mathbb{R}^d} \exp(\lambda |y_t|^2) \exp\left(-\frac{|y_t|^2}{2N_t}\right) \frac{dy_t}{(2\pi N_t)^{1/2}} \mathbf{1}_{N_t \geq 1} \right],$$

this is integrable if almost surely  $N_t < \frac{1}{2\lambda}$  which is not true for  $\lambda > 0$ .

#### Proof: Proposition 2

Preliminarily to the proof of this proposition, we write some useful controls thanks to assumption  $(\mathcal{L}_V)$ , for all  $x \in \mathbb{R}^d$ ,

$$|\nabla \sqrt{V}(x)| = \left| \frac{\nabla V(x)}{2\sqrt{V}(x)} \right| \leq \frac{\sqrt{C_V}}{2}, \quad (3.1)$$

$$\|D^2 \sqrt{V}(x)\| = \left\| \frac{D^2 V(x)}{2\sqrt{V}(x)} - \frac{\nabla V \nabla V^*(x)}{4V^{3/2}(x)} \right\| \leq \frac{\|D^2 V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}}. \quad (3.2)$$

To begin we check that  $\sqrt{V}$  satisfies assumption  $(\mathcal{L}_V)$  *iii*). We have readily that:

$$\tilde{\mathcal{A}}\sqrt{V} = \frac{1}{2\sqrt{V}}\tilde{\mathcal{A}}V - \frac{1}{8V^{\frac{3}{2}}}|\sigma^*\nabla V|^2 \leq \frac{1}{2\sqrt{V}}(-\tilde{\alpha}_V V + \tilde{\beta}_V) \leq -\frac{\alpha_V}{2}\sqrt{V} + \bar{\beta}_V, \quad (3.3)$$

with  $\bar{\beta}_V := \frac{\tilde{\beta}_V}{2\sqrt{v^*}}$ . The first inequality is a consequence of Remark 3. Furthermore, for the purely jump part of the infinitesimal generator we write:

$$\mathcal{A}\sqrt{V}(x) - \tilde{\mathcal{A}}\sqrt{V}(x) = \int_{\mathbb{R}^r} (\sqrt{V(x + \kappa(x)y)} - \sqrt{V(x)})\pi(dy) \leq \frac{\sqrt{C_V}\|\kappa\|_\infty\pi(|\cdot|)}{2},$$

using (3.1). The previous inequality and (3.3) implies that:

$$\mathcal{A}\sqrt{V} \leq -\frac{\alpha_V}{2}\sqrt{V} + \bar{\beta}'_V, \quad (3.4)$$

where  $\bar{\beta}'_V = \bar{\beta}_V + \frac{\sqrt{C_V}\|\kappa\|_\infty\pi(|\cdot|)}{2}$ .

Next, let us decompose the Lyapunov function  $\sqrt{V}$  with a Taylor expansion similarly to Lemma 1. We again use a splitting between the deterministic contributions and those involving the innovation. We write for all  $n \in \mathbb{N}$ :

$$\begin{aligned} & \sqrt{V}(X_n) - \sqrt{V}(X_{n-1}) = \sqrt{V}(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n) - \sqrt{V}(X_{n-1}) \\ & + \sqrt{V}(X_n) - \sqrt{V}(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n) \\ = & \gamma_n \mathcal{A}\sqrt{V}(X_{n-1}) + \left[ \gamma_n \int_0^1 \langle b_{n-1}, [\nabla\sqrt{V}(X_{n-1} + t\gamma_n b_{n-1}) - \nabla\sqrt{V}(X_{n-1})] \rangle dt \right] \\ & - \left[ \frac{\gamma_n}{2} \text{Tr}(D^2\sqrt{V}(X_{n-1}))\Sigma_{n-1} \right] + \left[ \sqrt{\gamma_n} \sigma_{n-1} U_n \cdot \nabla\sqrt{V}(X_{n-1} + \gamma_n b_{n-1}) \right. \\ & + \gamma_n \int_0^1 (1-t) \text{Tr} \left( D^2\sqrt{V}(X_{n-1} + \gamma_n b_{n-1} + t\sqrt{\gamma_n} \sigma_{n-1} U_n) \sigma_{n-1} U_n \otimes U_n \sigma_{n-1}^* \right) dt \left. \right] \\ & + \left[ \sqrt{V}(X_n) - \sqrt{V}(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n) \right. \\ & \left. - \gamma_n \pi(\sqrt{V}(X_{n-1} + \kappa(X_{n-1})\cdot) - \sqrt{V}(X_{n-1})) \right] \\ =: & \gamma_n \mathcal{A}\sqrt{V}(X_{n-1}) + \mathcal{V}_1(X_{n-1}) + \mathcal{V}_2(X_{n-1}) + \mathcal{V}_3(X_{n-1}, U_n) + \mathcal{V}_4(X_{n-1}, U_n, Z_n), \end{aligned} \quad (3.5)$$

where for all  $x \in \mathbb{R}^d$ , the first term is such that:

$$\begin{aligned} \mathcal{V}_1(x) &= \gamma_n \int_0^1 \langle b(x), \frac{\nabla V}{2\sqrt{V}}(x + t\gamma_n b(x)) - \frac{\nabla V}{2\sqrt{V}}(x) \rangle dt \\ &= \gamma_n \int_0^1 \langle b(x) - b(x + t\gamma_n b(x)), \frac{\nabla V}{2\sqrt{V}}(x + t\gamma_n b(x)) \rangle dt \\ &+ \gamma_n \int_0^1 \langle b, \frac{\nabla V}{2\sqrt{V}} \rangle(x + t\gamma_n b(x)) - \langle b, \frac{\nabla V}{2\sqrt{V}} \rangle(x) dt =: \mathcal{V}_1^1(x) + \mathcal{V}_1^2(x). \end{aligned} \quad (3.6)$$

Because  $b$  is supposed to be Lipschitz continuous and thanks to  $(\mathcal{L}_V)$  *ii*), we readily writes:

$$\mathcal{V}_1^1(x) \leq \frac{\gamma_n^2 [b]_1}{4} |b(x)| \int_0^1 \frac{|\nabla V|}{2\sqrt{V}}(x + t\gamma_n b(x)) dt \leq \frac{\gamma_n^2 [b]_1 C_V}{8} \sqrt{V}(x). \quad (3.7)$$

Whilst the next term is more subtle. Indeed, observe that thanks to  $(\mathcal{L}_V)$  ii) the following term is bounded:

$$\begin{aligned} |\nabla \langle b, \frac{\nabla V}{2\sqrt{V}} \rangle| &\leq |Db \frac{\nabla V}{2\sqrt{V}}| + |b \frac{D^2 V}{2\sqrt{V}}| + |\frac{(\nabla V)(\nabla V)^* b}{4V^{\frac{3}{2}}}| \\ &\leq \frac{\sqrt{C_V}[b]_1}{2} + \frac{\sqrt{C_V}\|D^2 V\|_\infty}{2} + \frac{C_V^{3/2}}{4} =: C_{(3.8)}, \end{aligned} \quad (3.8)$$

which directly yields again thanks to  $(\mathcal{L}_V)$  ii) that

$$\mathcal{V}_1^2(x) \leq \gamma_n^2 C_{(3.8)} \int_0^1 t|b(x)|dt \leq \frac{\gamma_n^2 C_{(3.8)} \sqrt{C_V} \sqrt{V}(x)}{2}. \quad (3.9)$$

Hence plugging (3.7) and (3.9) into (3.6) implies that:

$$\mathcal{V}_1(x) \leq \gamma_n^2 \left( \frac{[b]_1 C_V}{8} + C_{(3.8)} \sqrt{C_V} \right) \sqrt{V}(x). \quad (3.10)$$

The second term is handled by (3.2):

$$\mathcal{V}_2(x) \leq \frac{\gamma_n}{2} \|\sigma\|_\infty^2 \left( \frac{\|D^2 V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}} \right). \quad (3.11)$$

The third term satisfies the following identity:

$$\begin{aligned} \mathcal{V}_3(x, U_n) &= \sqrt{\gamma_n} \sigma(x) U_n \cdot \nabla \sqrt{V}(x + \gamma_n b(x)) \\ &+ \gamma_n \int_0^1 (1-t) \text{Tr} \left( D^2 \sqrt{V}(x + \gamma_n b(x)) + t \sqrt{\gamma_n} \sigma(x) U_n \sigma(x) U_n \otimes U_n \sigma(x)^* \right) dt \\ &\stackrel{(3.2)}{\leq} \sqrt{\gamma_n} \sigma(x) U_n \cdot \nabla \sqrt{V}(x + \gamma_n b(x)) + \frac{\gamma_n}{2} \left( \frac{\|D^2 V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}} \right) \|\sigma\|_\infty^2 |U_n|^2 \\ &=: \mathcal{V}_3^1(x, U_n) + \mathcal{V}_3^2(U_n), \end{aligned} \quad (3.12)$$

and the last term is:

$$\begin{aligned} &\mathcal{V}_4(x, U_n, Z_n) \\ &= \sqrt{V}(x + \gamma_n b(x) + \sqrt{\gamma_n} \sigma(x) U_n + \kappa(x) Z_n) - \sqrt{V}(x + \gamma_n b(x) + \sqrt{\gamma_n} \sigma(x) U_n) \\ &\quad - \gamma_n \pi(\sqrt{V}(x + \kappa(x) \cdot) - \sqrt{V}(x)) \\ &\stackrel{(3.1)}{\leq} \sqrt{V}(x + \gamma_n b(x) + \sqrt{\gamma_n} \sigma(x) U_n + \kappa(x) Z_n) - \sqrt{V}(x + \gamma_n b(x) + \sqrt{\gamma_n} \sigma(x) U_n) \\ &\quad + \frac{\gamma_n \|\kappa\|_\infty \sqrt{C_V} \pi(|\cdot|)}{2} \\ &=: \mathcal{V}_4^1(x, U_n, Z_n) + \frac{\gamma_n \|\kappa\|_\infty \sqrt{C_V} \pi(|\cdot|)}{2}. \end{aligned} \quad (3.13)$$

Hence plugging (3.4), (3.10), (3.11), (3.12) and (3.13) into (3.5):

$$\begin{aligned} &\sqrt{V}(X_n) - \sqrt{V}(X_{n-1}) \\ &\leq -\gamma_n \left[ \frac{\alpha_V}{2} \sqrt{V}(X_{n-1}) - \hat{\beta}'_V \right] + \gamma_n^2 \left( \frac{[b]_1 C_V}{8} + C_{(3.8)} \sqrt{C_V} \right) \sqrt{V}(X_{n-1}) \\ &\quad + \frac{\gamma_n}{2} \|\sigma\|_\infty^2 \left( \frac{\|D^2 V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}} \right) + \mathcal{V}_3(X_{n-1}, U_n) + \frac{\gamma_n \|\kappa\|_\infty \sqrt{C_V} \pi(|\cdot|)}{2} + \mathcal{V}_4^1(X_{n-1}, U_n) \\ &\leq -\gamma_n \frac{\alpha_V}{4} \sqrt{V}(X_{n-1}) + \gamma_n \hat{\beta}_V + \mathcal{V}_3^1(X_{n-1}, U_n) + \mathcal{V}_3^2(U_n) + \mathcal{V}_4^1(X_{n-1}, U_n, Z_n), \end{aligned} \quad (3.14)$$

for

$$\gamma_n \leq \frac{\alpha_V}{4\left(\frac{[b]_1 C_V}{8} + C_{(3.8)} \sqrt{C_V}\right)} = \frac{\alpha_V}{4\left(\frac{[b]_1 C_V}{8} + \left(\frac{\sqrt{C_V} [b]_1}{2} + \frac{\sqrt{C_V} \|D^2 V\|_\infty}{2} + \frac{C_V}{4}\right) \sqrt{C_V}\right)}$$

which corresponds to assumption **(S)** and

$$\hat{\beta}_V := \bar{\beta}'_V + \frac{1}{2} \|\sigma\|_\infty^2 \left( \frac{\|D^2 V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}} \right) + \frac{\|\kappa\|_\infty \sqrt{C_V} \pi(|\cdot|)}{2}.$$

We control the contribution of  $\mathcal{V}_3^1(X_{n-1}, U_n)$  and  $\mathcal{V}_3^2(U_n)$  (defined in (3.12)) in the exponential moment of  $\sqrt{V(X_n)}$  by the Gaussian concentration hypothesis **(GC)** and  $\mathcal{V}_4^1(X_{n-1}, U_n, Z_n)$  (see (3.13)) thanks to Proposition 1. We define for all  $x \in \mathbb{R}^d$  and  $\lambda > 0$ :

$$\begin{aligned} I_1(\lambda, x) &:= \mathbb{E} \left[ \exp(\lambda \mathcal{V}_3^1(x, U_n)) \right], \quad I_2(\lambda) := \mathbb{E} \left[ \exp(\lambda \mathcal{V}_3^2(U_n)) \right], \\ I_3(\lambda, x) &:= \mathbb{E} \left[ \exp(\lambda \mathcal{V}_4^1(x, U_n, Z_n)) \right]. \end{aligned}$$

Indeed, by **(GC)**, we first write:

$$I_1(\lambda, x) \leq \exp \left( \frac{\lambda^2 \gamma_n |\sigma^*(x) \nabla \sqrt{V}(x + \gamma_n b(x))|^2}{2} \right) \stackrel{(3.1)}{\leq} \exp \left( \frac{\lambda^2 \gamma_n C_V \|\sigma\|_\infty^2}{4} \right). \quad (3.15)$$

Next, it is well known that under **(GC)**, for all  $c < \frac{1}{2}$ ,  $I_c := \mathbb{E}[\exp(c|U_n|^2)] < +\infty$ . So we have for all  $\lambda < \frac{2c\sqrt{v^*}}{(C_V/2 + \|D^2 V\|_\infty) \|\sigma\|_\infty^2 \gamma_1}$ , by Jensen's inequality:

$$I_2(\lambda) \leq \left[ \mathbb{E} \exp(c|U_n|^2) \right]^{\frac{\lambda \gamma_n (\frac{\|D^2 V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}}) \|\sigma\|_\infty^2}{2c}} = \exp \left( \gamma_n \ln(I_c) \frac{\lambda}{2c} \left( \frac{\|D^2 V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}} \right) \|\sigma\|_\infty^2 \right). \quad (3.16)$$

Now, let us deal with the third term  $I_3(\lambda, x)$ . First of all, note that from definition in (3.13) and (3.1) the function  $z \mapsto \mathcal{V}_4^1(X_{n-1}, U_n, z)$  is  $\|\kappa\|_\infty \sqrt{C_V}$ -Lipschitz continuous.

Furthermore, we have that

$$|\mathbb{E}[\mathcal{V}_4^1(x, U_n, Z_n) | U_n]| = \gamma_n |\mathbb{E}[\mathcal{V}_4^1(x, U_n, Y_n) | U_n]| \leq \gamma_n \sqrt{C_V} \|\kappa\|_\infty \pi(|\cdot|).$$

Hence, by the Proposition 1, and for all  $0 < \lambda < \frac{1}{6\|\kappa\|_\infty \sqrt{C_V} \rho(r)}$  (see (1.11)), for the corresponding notation of Proposition 1 we take  $\varepsilon = 1$ , and we get:

$$\begin{aligned} I_3(\lambda, x) &\leq \mathbb{E} \left[ \mathbb{E} \left[ \exp(\lambda \mathcal{V}_4^1(x, U_n, Z_n)) \mid U_n \right] \right] \\ &\leq \mathbb{E} \left[ \exp \left( \lambda \mathbb{E}[\mathcal{V}_4^1(x, U_n, Z_n) | U_n] + \frac{(2+r)\gamma_n \lambda^2 [\mathcal{V}_4^1(x, U_n, \cdot)]_1^2}{2} \right) \right] \\ &\leq \exp(\lambda \gamma_n \|\kappa\|_\infty \sqrt{C_V} \pi(|\cdot|)) + (2+r) \|\kappa\|_\infty^2 C_V \frac{\lambda^2 \gamma_n}{2}. \end{aligned} \quad (3.17)$$

From now on, we assume that for all

$$\lambda < \lambda_V := \min \left( 1, \frac{\lambda_0}{2\bar{c}}, \frac{2c\sqrt{v^*}}{(C_V/2 + \|D^2 V\|_\infty) \|\sigma\|_\infty^2 \gamma_1}, \frac{1}{6\|\kappa\|_\infty \sqrt{C_V} \rho(r)} \right).$$

Gathering identities (3.14), (3.15) and (3.16), and by the Cauchy-Schwarz inequality, we obtain that for all  $\lambda < \lambda_V$ ,

$$\begin{aligned} & \mathbb{E} \exp(\lambda \sqrt{V}(X_n)) \\ &= \mathbb{E} \left[ \exp(\lambda \sqrt{V}(X_{n-1})) \mathbb{E} \left[ \exp(\lambda(\sqrt{V}(X_n) - \sqrt{V}(X_{n-1}))) \middle| \mathcal{F}_{n-1} \right] \right] \\ &\leq \mathbb{E} \left[ \exp(\lambda[\sqrt{V}(X_{n-1})(1 - \frac{\alpha_V}{4}\gamma_n) + \hat{\beta}_V \gamma_n]) I_1(2\lambda, X_{n-1})^{1/2} I_2(4\lambda)^{1/4} I_3(4\lambda, X_{n-1})^{1/4} \right] \\ &\leq \exp(\lambda \gamma_n \hat{\beta}'_V) \mathbb{E} \left[ \exp(\lambda(1 - \gamma_n \hat{\alpha}_V) \sqrt{V}(X_{n-1})) \right], \end{aligned}$$

where we have defined:

$$\hat{\beta}'_V := \hat{\beta}_V + \frac{C_V \|\sigma\|_\infty^2}{2} + \ln(I_c) \frac{(\frac{\|D^2V\|_\infty}{2\sqrt{v^*}} + \frac{C_V}{4\sqrt{v^*}}) \|\sigma\|_\infty^2}{2c} + \|\kappa\|_\infty \sqrt{C_V} 2\pi(|\cdot|) + 2(2+r)\|\kappa\|_\infty^2 C_V,$$

and

$$\hat{\alpha}_V := \min\left(\frac{1}{\gamma_1}, \frac{\alpha_V}{4}\right) \in (0, \frac{1}{\gamma_1}].$$

So  $(1 - \gamma_n \hat{\alpha}_V) \in [0, 1)$  and we deduce by Jensen inequality:

$$\mathbb{E} \exp(\lambda \sqrt{V}(X_n)) \leq \exp(\lambda \gamma_n \hat{\beta}'_V) \mathbb{E} \left[ \exp(\lambda \sqrt{V}(X_{n-1})) \right]^{(1 - \gamma_n \hat{\alpha}_V)}. \quad (3.18)$$

For all  $\lambda > 0$ , we introduce

$$C_{V,\lambda} := \max\left(\mathbb{E}[e^{\lambda \sqrt{V}(X_0)}], e^{\lambda \hat{\beta}'_V / \hat{\alpha}_V}\right).$$

In particular, we have  $\mathbb{E}[e^{\lambda \sqrt{V}(X_0)}] \leq C_{V,\lambda}$ . Let us check by induction that for all  $n \in \mathbb{N}$ :

$$\mathbb{E}[e^{\lambda \sqrt{V}(X_n)}] \leq C_{V,\lambda}.$$

We deduce from (3.18) and by induction assumption that:

$$\mathbb{E} \exp(\lambda \sqrt{V}(X_n)) \leq \exp(\lambda \gamma_n \hat{\beta}'_V) C_{V,\lambda}^{(1 - \gamma_n \hat{\alpha}_V)} \leq C_{V,\lambda}.$$

We pick  $c_V < \lambda_V$  and the proof is completed.  $\square$

**Remark 10** Observe also that  $v^* := \inf_{x \in \mathbb{R}^d} V(x) > 0$ , we have that for all  $(n, \xi) \in \mathbb{N} \times [0, \frac{1}{2}]$ ,  $\lambda < \lambda_V (v^*)^{1-\xi}$ :

$$\mathbb{E} \exp(\lambda V_n^\xi) = \mathbb{E} \exp\left(\lambda (v^*)^\xi \underbrace{\left(\frac{V_n}{v^*}\right)^\xi}_{\geq 1}\right) \leq \mathbb{E} \exp(\lambda (v^*)^{\xi-1} V_n) \leq C_{V,\lambda (v^*)^{\xi-1}} < +\infty.$$

Hence,  $\xi \in [0, 1]$ ,  $\lambda < \lambda_V (v^*)^{1-\xi}$ ,  $\sup_{n \in \mathbb{N}} \mathbb{E} \exp(\lambda \sqrt{V}_n^\xi) < +\infty$ .

## 4 Proof of the Technical Lemmas

### Proof: Lemma 1

For  $k \in \llbracket 1, n \rrbracket$ , we first write:

$$\begin{aligned} \varphi(X_k) - \varphi(X_{k-1}) &= (\varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k)) \\ &+ (\varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k) - \varphi(X_{k-1} + \gamma_k b_{k-1})) + (\varphi(X_{k-1} + \gamma_k b_{k-1}) - \varphi(X_{k-1})) \\ &=: T_{k-1,j}(\varphi) + T_{k-1,r}(\varphi) + T_{k-1,d}(\varphi), \end{aligned} \quad (4.1)$$

with

$$\begin{aligned} T_{k-1,j}(\varphi) &= (\varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k)) \\ T_{k-1,d}(\varphi) &= \gamma_k \langle \nabla \varphi(X_{k-1}), b_{k-1} \rangle + \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt, \\ T_{k-1,r}(\varphi) &= \sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &+ \gamma_k \int_0^1 (1-t) \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right) dt. \end{aligned}$$

Thanks to this splitting, we are able to isolate the deterministic, the sub-Gaussian random variable approximating Brownian increments and the jump contributions. Then we proceed by Taylor expansion up to the order 2 for the function  $\varphi$  in the two last terms of the r.h.s. of (4.1),

$$\begin{aligned} &\varphi(X_k) - \varphi(X_{k-1}) \\ &= \gamma_k \mathcal{A} \varphi(X_{k-1}) + \varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k) \\ &\quad - \gamma_k \int_{\mathbb{R}^r} (\varphi(X_{k-1} + \kappa_{k-1} y) - \varphi(X_{k-1})) \pi(dy) \\ &\quad + [\gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt] \\ &\quad + [\frac{\gamma_k}{2} (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1}))] \\ &\quad + [\sqrt{\gamma_k} \sigma_{k-1} U_k \cdot \nabla \varphi(X_{k-1} + \gamma_k b_{k-1}) \\ &\quad + \gamma_k \int_0^1 (1-t) \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right. \\ &\quad \left. - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right) dt] \\ &= \gamma_k \mathcal{A} \varphi(X_{k-1}) + D_{2,b}^{k,\varphi}(X_{k-1}) + D_{2,\Sigma}^{k,\varphi} + \psi_k^\varphi(X_{k-1}, U_k) \\ &\quad + \varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k) \\ &\quad - \gamma_k \int_{\mathbb{R}^r} (\varphi(X_{k-1} + \kappa_{k-1} y) - \varphi(X_{k-1})) \pi(dy), \end{aligned}$$

we then obtain:

$$\begin{aligned}
 & \varphi(X_k) - \varphi(X_{k-1}) \\
 = & \gamma_k \mathcal{A}\varphi(X_{k-1}) + D_{2,b}^{k,\varphi}(X_{k-1}) + D_{2,\Sigma}^{k,\varphi} + \psi_k^\varphi(X_{k-1}, U_k) \\
 & + [\varphi(X_{k-1} + \kappa_{k-1}Z_k) - \varphi(X_{k-1}) - \gamma_k \int_{\mathbb{R}^r} (\varphi(X_{k-1} + \kappa_{k-1}y) - \varphi(X_{k-1})) \pi(dy)] \\
 & + [\varphi(X_k) - \varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k) - (\varphi(X_{k-1} + \kappa_{k-1}Z_k) - \varphi(X_{k-1}))] \\
 = & \gamma_k \mathcal{A}\varphi(X_{k-1}) + D_{2,b}^{k,\varphi}(X_{k-1}) + D_{2,\Sigma}^{k,\varphi}(X_{k-1}) + \psi_k^\varphi(X_{k-1}, U_k) \\
 & + \tilde{\Delta}_k^\varphi(X_{k-1}, Z_k) + D_j^{k,\varphi}(X_{k-1}, U_k, Z_k).
 \end{aligned} \tag{4.2}$$

Note finally, that by definition of  $D_j^{k,\varphi}(X_{k-1}, U_k, Z_k)$  in the previous expansion (4.2):

$$\begin{aligned}
 \psi_k^\varphi(X_{k-1}, U_k) &= \varphi(X_k) - \varphi(X_{k-1}) - \gamma_k \mathcal{A}\varphi(X_{k-1}) - \tilde{\Delta}_k^\varphi(X_{k-1}, Z_k) \\
 &- (D_{2,b}^{k,\varphi}(X_{k-1}) + D_{2,\Sigma}^{k,\varphi}(X_{k-1}) + D_j^{k,\varphi}(X_{k-1}, U_k, Z_k)) \\
 &= \varphi(X_{k-1} + \gamma_k b_{k-1} + \sqrt{\gamma_k} \sigma_{k-1} U_k) + \varphi(X_{k-1} + \kappa_{k-1}Z_k) - \gamma_k \mathcal{A}\varphi(X_{k-1}) - 2\varphi(X_{k-1}) \\
 &- \tilde{\Delta}_k^\varphi(X_{k-1}, Z_k) - (D_{2,b}^{k,\varphi}(X_{k-1}) + D_{2,\Sigma}^{k,\varphi}(X_{k-1})),
 \end{aligned}$$

hence after differentiating, we see that  $u \mapsto \psi_k^\varphi(X_{k-1}, u)$  and hence  $u \mapsto \Delta_k^\varphi(X_{k-1}, u)$  are Lipschitz continuous with a modulus bounded by  $\sqrt{\gamma_{k-1}} \|\sigma_{k-1}\| \|\nabla \varphi\|_\infty \leq \sqrt{\gamma_{k-1}} \|\sigma\|_\infty \|\nabla \varphi\|_\infty$ .

Moreover, from the definition  $\mathbb{E}[\Delta_k^\varphi(X_{k-1}, U_k) | \mathcal{F}_{k-1}] = 0$  and using the definition of  $Z_n$  we get

$$\begin{aligned}
 \mathbb{E}[\tilde{\Delta}_k^\varphi(X_{k-1}, Z_k) | \mathcal{F}_{k-1}] &= \tag{4.3} \\
 \mathbb{E}[\varphi(X_{k-1} + \kappa_{k-1}Z_k) - \varphi(X_{k-1}) | \mathcal{F}_{k-1}] - \gamma_k \int_{\mathbb{R}^r} [\varphi(X_{k-1} + \kappa_{k-1}y) - \varphi(X_{k-1})] \pi(dy) &= 0.
 \end{aligned}$$

□

**Proof: Lemma 2**

We first prove the point **ii**).

For all  $\varepsilon \in (0, 1)$  and  $0 < \frac{\Lambda}{\Gamma_n} < \frac{\varepsilon}{6[\tilde{\Delta}_n(X_{n-1}, \cdot)]_1 \rho(r)}$  ( $\rho(r)$  set in (1.11)), thanks to Proposition 1, we have for all  $n \in \mathbb{N}$ :

$$\begin{aligned}
 & \mathbb{E} \left[ \exp \left( -\frac{\Lambda}{\Gamma_n} \tilde{\Delta}_n^\varphi(X_{n-1}, Z_n) \right) \middle| \mathcal{F}_{n-1} \right] \\
 & \leq \exp \left( -\frac{\Lambda}{\Gamma_n} \mathbb{E}[\tilde{\Delta}_n^\varphi(X_{n-1}, Z_n) | \mathcal{F}_{n-1}] + \gamma_n \frac{\Lambda^2}{2\Gamma_n^2} [\tilde{\Delta}_n^\varphi(X_{n-1}, \cdot)]_1^2 (1 + r + \varepsilon) \right). \tag{4.4}
 \end{aligned}$$

By definition of  $\tilde{\Delta}_n^\varphi(X_{n-1}, Z_n)$  in (2.1), and from (4.3) we have:

$$\mathbb{E} \left[ \frac{\Lambda}{\Gamma_n} \tilde{\Delta}_n^\varphi(X_{n-1}, Z_n) \middle| \mathcal{F}_{n-1} \right] = 0, \text{ and } [\tilde{\Delta}_n^\varphi(X_{n-1}, \cdot)]_1^2 \leq \|\kappa\|_\infty^2 \|\nabla \varphi\|_\infty^2.$$



The previous control with (4.4) directly yield:

$$\mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} \tilde{\Delta}_n^\varphi(X_{n-1}, Z_n) \right) \middle| \mathcal{F}_{n-1} \right] \leq \exp \left( \frac{\gamma_n \Lambda^2}{2\Gamma_n^2} \|\kappa\|_\infty^2 \|\nabla \varphi\|_\infty^2 (1+r+\varepsilon) \right), \quad (4.5)$$

with the constraint  $0 < \frac{\Lambda}{\Gamma_n} < \frac{\varepsilon}{6\|\kappa\|_\infty \|\nabla \varphi\|_\infty \rho(r)}$ .

The demonstration of the point **i**), is a direct consequence of the previous analysis without using Proposition 1, which yields no restriction on  $\lambda$ .  $\square$

**Proof: Lemma 3**

By assumption **(T $_\beta$ )**, we know that there exists  $C_{V,\varphi} > 0$  such that for all  $x \in \mathbb{R}^d$ ,  $|\varphi(x)| \leq C_{V,\varphi}(1 + \sqrt{V(x)})$ , so we obtain:

$$\begin{aligned} \mathbb{E} \exp \left( \Lambda \frac{|\varphi(X_0) - \varphi(X_n)|}{\Gamma_n} \right) &\leq \mathbb{E} \exp \left( \Lambda \frac{C_{V,\varphi}(2 + \sqrt{V(X_0)} + \sqrt{V(X_n)})}{\Gamma_n} \right) \\ &\leq \exp \left( 2C_{V,\varphi} \frac{\Lambda}{\Gamma_n} \right) \left[ \mathbb{E} \exp \left( 2C_{V,\varphi} \Lambda \frac{\sqrt{V(X_0)}}{\Gamma_n} \right) \right]^{\frac{1}{2}} \left[ \mathbb{E} \exp \left( 2C_{V,\varphi} \Lambda \frac{\sqrt{V(X_n)}}{\Gamma_n} \right) \right]^{\frac{1}{2}} \\ &\leq \exp \left( 2C_{V,\varphi} \frac{\Lambda}{\Gamma_n} \right) (I_V^{\frac{1}{2}})^{\frac{2C_{V,\varphi}\Lambda}{c_V\Gamma_n}}. \end{aligned}$$

The last inequality is obtained by Jensen's inequality for  $\frac{\Lambda}{\Gamma_n} < \frac{c_V}{2C_{V,\varphi}}$  and by Proposition 2.  $\square$

**Proof: Lemma 4**

From the definition (2.2) we can write:

$$\begin{aligned} D_{2,b}^{k,\varphi} &= \gamma_k \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}) - \nabla \varphi(X_{k-1}), b_{k-1} \rangle dt \\ &= \gamma_k \left[ \int_0^1 \langle \nabla \varphi(X_{k-1} + t\gamma_k b_{k-1}), b_{k-1} - b(X_{k-1} + t\gamma_k b_{k-1}) \rangle dt \right. \\ &\quad \left. + \int_0^1 (\langle \nabla \varphi, b \rangle(X_{k-1} + t\gamma_k b_{k-1}) - \langle \nabla \varphi, b \rangle(X_{k-1})) dt \right]. \end{aligned}$$

From the boundedness of  $\nabla \varphi$ , Lipschitz property of the mapping  $x \mapsto b(x)$  (assumption **(C $_0$ )**) and Lipschitz property of the mapping  $x \mapsto \langle \nabla \varphi(x), b(x) \rangle$  (assumption **(T $_\beta$ )**), using the assumption **(L $_V$ )**, ii) one derives that:

$$|D_{2,b}^{k,\varphi}| \leq \gamma_k^2 \left( \|\nabla \varphi\|_\infty [b]_1 + [\langle \nabla \varphi, b \rangle]_1 \right) \frac{|b_{k-1}|}{2} \leq C_{(4.6)} \gamma_k^2 \sqrt{V_{k-1}}, \quad (4.6)$$

for  $C_{(4.6)} := (\|\nabla \varphi\|_\infty [b]_1 + [\langle \nabla \varphi, b \rangle]_1) \frac{\sqrt{C_V}}{2}$ . Hence

$$|D_{2,b,n}^\varphi| \leq \sum_{k=1}^n C_{(4.6)} \gamma_k^2 \sqrt{V_{k-1}}.$$

Next, by the Jensen inequality (for the exponential function with  $\frac{1}{\Gamma_n^{(2)}} \sum_{k=1} \gamma_k^2 \delta_k$  as a measure), we deduce that:

$$\begin{aligned} \mathbb{E} \exp \left( \frac{\Lambda}{\Gamma_n} |D_{2,b,n}^\varphi| \right) &\leq \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^{n-1} \gamma_k^2 \mathbb{E} \left[ \exp \left( \Gamma_n^{(2)} \frac{\Lambda}{\Gamma_n} C_{(4.6)} \sqrt{V_{k-1}} \right) \right] \\ &\leq \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^{n-1} \gamma_k^2 \mathbb{E} \left[ \exp \left( c_V \sqrt{V_{k-1}} \right) \right]^{\frac{c_{(4.6)} \Lambda \Gamma_n^{(2)}}{c_V \Gamma_n}}, \end{aligned}$$

for  $\frac{\Lambda}{\Gamma} \leq \frac{c_V}{C_{(4.6)} \Gamma_n^{(2)}} = \frac{2c_V}{(\|\nabla \varphi\|_\infty [b]_1 + \langle \nabla \varphi, b \rangle)_1 \sqrt{C_V} \Gamma_n^{(2)}} < 1$ , and  $c_V$  is introduced in Proposition 2 which readily yields:

$$\mathbb{E} \exp \left( \frac{\Lambda}{\Gamma_n} |D_{2,b,n}^\varphi| \right) \leq \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^{n-1} \gamma_k^2 (I_V^{1/2})^{\frac{c_{(4.6)} \Lambda \Gamma_n^{(2)}}{c_V \Gamma_n}} = (I_V^{1/2})^{\frac{c_{(4.6)} \Lambda \Gamma_n^{(2)}}{c_V \Gamma_n}}.$$

For the second inequality, we first use a Taylor expansion:

$$\begin{aligned} |D_{2,\Sigma}^{k,\varphi}| &= \frac{\gamma_k}{2} \left| \text{Tr} \left( (D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) - D^2 \varphi(X_{k-1})) \Sigma_{k-1} \right) \right| \\ &\leq \frac{1}{2} \|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty \gamma_k^2 |b_{k-1}| \leq \frac{1}{2} \|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty \sqrt{C_V} \gamma_k^2 |V_{k-1}|^{\frac{1}{2}} \end{aligned} \quad (4.7)$$

So

$$|D_{2,\Sigma,n}^\varphi| \leq \frac{1}{2} \|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty \sqrt{C_V} \sum_{k=1}^n \gamma_k^2 |V_{k-1}|^{\frac{1}{2}}.$$

Hence, like previously, by Jensen inequality and Proposition 2 for  $0 < \Lambda$  satisfying  $\frac{\Lambda}{\Gamma_n} \leq \frac{2c_V}{\|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty \sqrt{C_V} \Gamma_n^{(2)}} < 1$  we obtain

$$\begin{aligned} \mathbb{E} \exp \left( \frac{\Lambda}{\Gamma_n} |D_{2,\Sigma,n}^\varphi| \right) &\leq \mathbb{E} \exp \left( \frac{\Lambda}{2\Gamma_n} \|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty \sqrt{C_V} \sum_{k=1}^n \gamma_k^2 |V_{k-1}|^{\frac{1}{2}} \right) \\ &\leq \frac{1}{\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 \mathbb{E} \exp \left( \frac{\Lambda \Gamma_n^{(2)}}{2\Gamma_n} \|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty \sqrt{C_V} |V_{k-1}|^{\frac{1}{2}} \right) \\ &\leq (I_V^{1/2})^{\frac{\Lambda \Gamma_n^{(2)} \|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty \sqrt{C_V}}{2c_V \Gamma_n}}. \end{aligned}$$

□

### Proof: Lemma 5

The proof is similar to the analysis of Lemma 3 in Chapter 3. By the definition (2.1), and because  $U_k$ ,  $k \in \llbracket 1, n \rrbracket$ , has the same moments as the standard Gaussian random variable up to order three (see **(GM)**) we have for all  $k \in \llbracket 1, n \rrbracket$ :

$$\begin{aligned} &\mathbb{E} [\psi_k^\varphi(X_{k-1}, U_k) | \mathcal{F}_{k-1}] \\ &= \gamma_k \int_0^1 (1-t) \text{Tr} \left( \mathbb{E} [D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right. \\ &\quad \left. - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} | \mathcal{F}_{k-1}] \right) dt, \end{aligned}$$

where

$$\begin{aligned}
 & \mathbb{E} \left[ \text{Tr} \left( D^2 \varphi(X_{k-1} + \gamma_k b_{k-1} + t\sqrt{\gamma_k} \sigma_{k-1} U_k) \sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^* \right. \right. \\
 & \quad \left. \left. - D^2 \varphi(X_{k-1} + \gamma_k b_{k-1}) \Sigma_{k-1} \right) \middle| \mathcal{F}_{k-1} \right] \\
 &= t\sqrt{\gamma_k} \int_0^1 \mathbb{E} \left[ \text{Tr} \left( ([D^3 \varphi(X_{k-1} + \gamma_k b_{k-1} + ut\sqrt{\gamma_k} \sigma_{k-1} U_k) - D^3 \varphi(X_{k-1} + \gamma_k b_{k-1})] \sigma_{k-1} U_k \right. \right. \\
 & \quad \left. \left. (\sigma_{k-1} U_k \otimes U_k \sigma_{k-1}^*) \right) \middle| \mathcal{F}_{k-1} \right] du.
 \end{aligned}$$

Then,

$$\begin{aligned}
 & |\mathbb{E} [\psi_k^\varphi(X_{k-1}, U_k) | \mathcal{F}_{k-1}]| \\
 & \leq \gamma_k \int_0^1 (1-t) t^{1+\beta} [\varphi^{(3)}]_\beta \mathbb{E} \left[ \gamma_k^{\frac{1+\beta}{2}} \|\sigma_{k-1}\|^{3+\beta} |U_k|^{3+\beta} \int_0^1 u^\beta du \middle| \mathcal{F}_{k-1} \right] dt \\
 & = \frac{[\varphi^{(3)}]_\beta \gamma_k^{\frac{3+\beta}{2}} \|\sigma_{k-1}\|^{3+\beta} \mathbb{E} [|U_k|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)}. \tag{4.8}
 \end{aligned}$$

We sum over  $k$  to get the result.  $\square$

**Proof: Lemma 6**

Recall that we have denoted for  $n \in \mathbb{N}_0$ ,  $\mathcal{F}_n := \sigma(X_0, (U_j, Z_j)_{j \in \llbracket 1, n \rrbracket})$  and  $\tilde{\mathcal{F}}_n = \mathcal{F}_n \vee \sigma(U_{n+1})$ .

$$\begin{aligned}
 & \mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} \sum_{k=1}^n D_j^{k, \varphi}(X_{k-1}, U_k, Z_k) \right) \right] \tag{4.9} \\
 &= \mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} \sum_{k=1}^{n-1} D_j^{k, \varphi}(X_{k-1}, U_k, Z_k) \right) \mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} D_{j,n}^\varphi(X_{n-1}, U_n, Z_n) \right) \middle| \tilde{\mathcal{F}}_{n-1} \right] \right].
 \end{aligned}$$

The idea is to control the last conditional expectation using Proposition 2. Recall that

$$\begin{aligned}
 & D_{j,n}^\varphi(X_{n-1}, U_n, Z_n) \\
 &= \varphi(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n + \kappa_{n-1} Z_n) - \varphi(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n) \\
 & \quad - [\varphi(X_{n-1} + \kappa_{n-1} Z_n) - \varphi(X_{n-1})].
 \end{aligned}$$

Moreover, we have for all  $z \in \mathbb{R}^r$ :

$$\begin{aligned}
 & |\nabla_z D_{j,n}^\varphi(X_{n-1}, U_n, z)| \\
 &= |\kappa_{n-1}^* (\nabla \varphi(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n + \kappa_{n-1} z) - \nabla \varphi(X_{n-1} + \kappa_{n-1} z))| \\
 & \leq 2 \|\kappa\|_\infty \|\nabla \varphi\|_\infty. \tag{4.10}
 \end{aligned}$$

Hence for all  $X_{n-1}, U_n$  fixed the function  $z \rightarrow D_{j,n}^\varphi(X_{n-1}, U_n, z)$  is Lipschitz continuous satisfying

$$[D_{j,n}^\varphi(X_{n-1}, U_n, z)]_1 \leq 2 \|\kappa\|_\infty \|\nabla \varphi\|_\infty.$$

This estimation is used to bound  $\Lambda$  for which we can apply the Proposition 1. However, we need a more subtle control of the last Lipschitz modulus. Namely, using Taylor expansion we can write

$$\begin{aligned}
 & |\nabla_z D_{j,n}^\varphi(X_{n-1}, U_n, z)| \\
 \leq & |\kappa_{n-1}^* (\nabla\varphi(X_{n-1} + \gamma_n b_{n-1} + \kappa_{n-1} z) - \nabla\varphi(X_{n-1} + \kappa_{n-1} z))| \\
 & + |\kappa_{n-1}^* (\nabla\varphi(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n + \kappa_{n-1} z) - \nabla\varphi(X_{n-1} + \gamma_n b_{n-1} + \kappa_{n-1} z))| \\
 \leq & \|\kappa\|_\infty (\sqrt{2} \|\nabla\varphi\|_\infty^{\frac{1}{2}} \|\nabla\varphi(X_{n-1} + \gamma_n b_{n-1} + \kappa_{n-1} z) - \nabla\varphi(X_{n-1} + \kappa_{n-1} z)\|_{\frac{1}{2}} \\
 & + \sqrt{\gamma_n} \|D^2\varphi\|_\infty \|\sigma\|_\infty |U_n|) \\
 \leq & \sqrt{\gamma_n} \|\kappa\|_\infty (\sqrt{2} C_V^{\frac{1}{4}} \|\nabla\varphi\|_\infty^{\frac{1}{2}} \|D^2\varphi\|_\infty^{\frac{1}{2}} V^{\frac{1}{4}}(X_{n-1}) + \|D^2\varphi\|_\infty \|\sigma\|_\infty |U_n|). \tag{4.11}
 \end{aligned}$$

Now for all  $\Lambda$  satisfying  $0 < \frac{\Lambda}{\Gamma_n} < \frac{1}{12\|\kappa\|_\infty \|\nabla\varphi\|_\infty \rho(r)}$ , we get

$$\begin{aligned}
 & \mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} D_{j,n}^\varphi(X_{n-1}, U_n, Z_n) \right) \middle| \tilde{\mathcal{F}}_{n-1} \right] \\
 & \leq \exp \left( \frac{\Lambda}{\Gamma_n} \mathbb{E}[D_{j,n}^\varphi(X_{n-1}, U_n, Z_n) | \tilde{\mathcal{F}}_{n-1}] + \frac{\gamma_n \Lambda^2}{2\Gamma_n^2} (1+r+\varepsilon) [D_{j,n}^\varphi(X_{n-1}, U_n, \cdot)]_1^2 \right) \\
 & \leq \exp \left( \frac{\Lambda}{\Gamma_n} \gamma_n \mathbb{E}[D_{j,n}^\varphi(X_{n-1}, U_n, Y_n) | \tilde{\mathcal{F}}_{n-1}] + \frac{\gamma_n^2 \Lambda^2}{2\Gamma_n^2} (C_1 \sqrt{V}(X_{n-1}) + C_2 |U_n|^2) \right), \tag{4.12}
 \end{aligned}$$

where we have denoted

$$C_1 := (1+r+\varepsilon) \|\kappa\|_\infty^2 4\sqrt{C_V} \|\nabla\varphi\|_\infty \|D^2\varphi\|_\infty, \quad C_2 := 2(1+r+\varepsilon) \|\kappa\|_\infty^2 \|D^2\varphi\|_\infty^2 \|\sigma\|_\infty^2,$$

and used the following identities:

$$D_{j,n}^\varphi(X_{n-1}, U_n, 0) = 0,$$

and

$$\mathbb{E}[D_{j,n}^\varphi(X_{n-1}, U_n, Z_n) | \tilde{\mathcal{F}}_{n-1}] = \gamma_n \mathbb{E}[D_{j,n}^\varphi(X_{n-1}, U_n, Y_n) | \tilde{\mathcal{F}}_{n-1}],$$

which is a consequence of the definition of  $Z_n$  in (1.3).

To control  $\mathbb{E}[D_{j,n}^\varphi(X_{n-1}, U_n, Y_n) | \tilde{\mathcal{F}}_{n-1}]$  we introduce for all  $(x, y) \in (\mathbb{R}^d)^2$  the function:

$$\bar{\varphi}(x, y) := \mathbb{E}[\varphi(x + \kappa(y)Y_n)] - \varphi(x),$$

which readily implies that:

$$\mathbb{E}[D_{j,n}^\varphi(X_{n-1}, U_n, Y_n) | \tilde{\mathcal{F}}_{n-1}] = \bar{\varphi}(X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n, X_{n-1}) - \bar{\varphi}(X_{n-1}, X_{n-1}).$$

The idea in the following is to apply the expansion of Lemma 1 with  $\kappa = 0$  to the function  $x \rightarrow \bar{\varphi}(x, y)$ , which also corresponds to the expansion of Lemma 1 in Chapter 3 for diffusion without jumps. If  $\kappa = 0$ , then  $X_n = X_{n-1} + \gamma_n b_{n-1} + \sqrt{\gamma_n} \sigma_{n-1} U_n$ , we can write using (2.3) and the definition (2.1) of  $\Delta_k^\varphi$  with  $\tilde{M}_n^\varphi = D_{j,n}^\varphi = 0$ :

$$\bar{\varphi}(X_n, X_{n-1}) - \bar{\varphi}(X_{n-1}, X_{n-1}) = \gamma_n \tilde{\mathcal{A}} \bar{\varphi}(X_{n-1}) + D_{2,b,n}^{\bar{\varphi}}(X_{n-1}) + D_{2,\Sigma,n}^{\bar{\varphi}}(X_{n-1}) + \psi_n^{\bar{\varphi}}(X_{n-1}, U_n). \tag{4.13}$$

All the terms in the right have obviously the same properties as the corresponding terms in the similar expansion of  $\varphi$  given by Lemma 1 with  $\kappa = 0$ . In particular, for all  $y \in \mathbb{R}^d$ , the map  $\bar{\varphi}(\cdot, y)$  is Lipschitz continuous with

$$\|\nabla \bar{\varphi}(\cdot, y)\|_\infty \leq 2\|\nabla \varphi\|_\infty, \quad \|D^2 \bar{\varphi}(\cdot, y)\|_\infty \leq 2\|D^2 \varphi\|_\infty, \quad \|D^3 \bar{\varphi}(\cdot, y)\|_\infty \leq 2\|D^3 \varphi\|_\infty. \quad (4.14)$$

Furthermore,  $D_{2,b,n}^{\bar{\varphi}}$  and  $D_{2,\Sigma,n}^{\bar{\varphi}}$  satisfy similar inequalities as (4.6) and (4.7) where  $\varphi$  is replaced by  $\bar{\varphi}$ . We directly have thanks to the definitions (1.7), (2.2), identities (4.6), (4.7) and (4.14):

$$\begin{aligned} \gamma_n |\tilde{\mathcal{A}}\bar{\varphi}(X_{n-1})| &= \gamma_n |\langle b_{n-1}, \nabla \bar{\varphi}(X_{n-1}) \rangle + \frac{1}{2} \text{Tr}(\sigma \sigma^* D^2 \bar{\varphi}(X_{n-1}))| \\ &\leq \gamma_n (2\sqrt{C_V} \sqrt{V}(X_{n-1}) \|\nabla \varphi\|_\infty + \|\sigma\|_\infty^2 \|D^2 \varphi\|_\infty) \leq \gamma_n C_3 \sqrt{V}(X_{n-1}), \\ |D_{2,b,n}^{\bar{\varphi}}(X_{n-1})| &= \gamma_n \left| \int_0^1 \langle \nabla \bar{\varphi}(X_{n-1} + t\gamma_n b_{n-1}) - \nabla \bar{\varphi}(X_{n-1}), b_{n-1} \rangle dt \right| \\ &\leq \gamma_n^2 \left( \|\nabla \bar{\varphi}\|_\infty [b]_1 + [\langle \nabla \bar{\varphi}, b \rangle]_1 \right) \frac{|b_{n-1}|}{2} \leq \gamma_n^2 2C_{(4.6)} \sqrt{V}(X_{n-1}), \\ |D_{2,\Sigma,n}^{\bar{\varphi}}(X_{n-1})| &= \left| \frac{\gamma_n}{2} \text{Tr}((D^2 \bar{\varphi}(X_{n-1} + \gamma_n b_{n-1}) - D^2 \bar{\varphi}(X_{n-1})) \Sigma_{n-1}) \right| \\ &\leq \gamma_n^2 \|\sigma\|_\infty^2 \|D^3 \varphi\|_\infty C_V^{\frac{1}{2}} \sqrt{V}(X_{n-1}), \end{aligned}$$

where  $C_3 := 2\sqrt{C_V} \|\nabla \varphi\|_\infty + \|\sigma\|_\infty^2 \|D^2 \varphi\|_\infty (v^*)^{-\frac{1}{2}}$ . Therefore, from the previous controls and using (4.12) and (4.13), we get that there is a constant  $C = C(\mathbf{A}) > 0$  such that:

$$\begin{aligned} &\mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} D_{j,n}^\varphi(X_{n-1}, U_n, Z_n) \right) \middle| \tilde{\mathcal{F}}_{n-1} \right] \\ &\leq \exp \left( \frac{C\Lambda}{\Gamma_n} \gamma_n (\gamma_n \sqrt{V}(X_{n-1}) + \psi_n^{\bar{\varphi}}(X_{n-1}, U_n)) + \frac{C\Lambda^2}{\Gamma_n^2} \gamma_n^2 (\sqrt{V}(X_{n-1}) + |U_n|^2) \right). \end{aligned} \quad (4.15)$$

Next, the idea is to separate the unbounded contribution from the terms involving  $(\sqrt{V}(X_k))_{k \in \llbracket 1, n \rrbracket}$ ,  $(|U_k|^2)_{k \in \llbracket 1, n \rrbracket}$  by a global Cauchy-Schwarz inequality:

$$\begin{aligned} &\mathbb{E} \left[ \exp \left( \frac{\Lambda}{\Gamma_n} \sum_{k=1}^n D_j^{k,\varphi}(X_{k-1}, U_k, Z_k) \right) \right] \\ &\leq \mathbb{E} \left[ \exp \left( \frac{2\Lambda}{\Gamma_n} \sum_{k=1}^n \left[ D_j^{k,\varphi}(X_{k-1}, U_k, Z_k) - C\gamma_k^2 \left( 1 + \frac{2\Lambda}{\Gamma_n} \right) \sqrt{V}(X_{k-1}) - \frac{2C\Lambda}{\Gamma_n} \gamma_k^2 |U_k|^2 \right] \right) \right]^{\frac{1}{2}} \\ &\quad \times \mathbb{E} \left[ \exp \left( 2 \sum_{k=1}^n \left[ C\gamma_k^2 \left( \frac{\Lambda}{\Gamma_n} + \frac{2\Lambda^2}{\Gamma_n^2} \right) \sqrt{V}(X_{k-1}) + \frac{2C\Lambda^2}{\Gamma_n^2} \gamma_k^2 |U_k|^2 \right] \right) \right]^{\frac{1}{2}} =: \Upsilon_1^{\frac{1}{2}} \times \Upsilon_2^{\frac{1}{2}}. \end{aligned} \quad (4.16)$$

Again by the Cauchy-Schwarz inequality, we get :

$$\Upsilon_2^{1/2} \leq \mathbb{E} \left[ \exp \left( 4C \left( \frac{\Lambda}{\Gamma_n} + \frac{2\Lambda^2}{\Gamma_n^2} \right) \sum_{k=1}^n \gamma_k^2 \sqrt{V}(X_{k-1}) \right) \right]^{\frac{1}{4}} \mathbb{E} \left[ \exp \left( 4 \frac{C\Lambda^2}{\Gamma_n^2} \sum_{k=1}^n \gamma_k^2 |U_k|^2 \right) \right]^{1/4}. \quad (4.17)$$

We control the second expected value under condition  $\frac{4C\Lambda^2\gamma_1^2}{\Gamma_n^2} < 1$  using Jensen inequality:

$$\begin{aligned} \mathbb{E}\left[\exp\left(4\frac{C\Lambda^2}{\Gamma_n^2}\sum_{k=1}^n\gamma_k^2|U_k|^2\right)\right]^{1/4} &\leq \sum_{k=1}^n\frac{\gamma_k^2}{\Gamma_n^{(2)}}\mathbb{E}\left[\exp\left(\frac{4C\Lambda^2\Gamma_n^{(2)}}{\Gamma_n^2}|U_k|^2\right)\right]^{1/4} \\ &\leq \mathbb{E}\left[\exp\left(\frac{|U_1|^2}{4}\right)\right]^{\frac{16C\Lambda^2\Gamma_n^{(2)}}{\Gamma_n^2}}. \end{aligned} \quad (4.18)$$

Because  $U_1$  satisfies **(GC)**,  $\mathbb{E}[\exp(\frac{|U_1|^2}{4})] < +\infty$ . We handle the first expectation in (4.17) by the same method, using Jensen inequality under condition  $\frac{\Lambda}{\Gamma_n} < \frac{c_V}{2C\Gamma_n^{(2)}}$ ,  $\frac{\Lambda^2}{\Gamma_n^2} < \frac{c_V}{16C\Gamma_n^{(2)}}$  and Proposition 2 and we obtain:

$$\begin{aligned} &\mathbb{E}\left[\exp\left(4C\left(\frac{\Lambda}{\Gamma_n} + \frac{2\Lambda^2}{\Gamma_n^2}\right)\sum_{k=1}^n\gamma_k^2\sqrt{V}(X_{k-1})\right)\right]^{1/4} \\ &\leq \frac{1}{\Gamma_n^{(2)}}\sum_{k=1}^n\gamma_k^2\mathbb{E}\left[\exp\left(4C\Gamma_n^{(2)}\left(\frac{\Lambda}{\Gamma_n} + \frac{2\Lambda^2}{\Gamma_n^2}\right)\sqrt{V}(X_{k-1})\right)\right]^{1/4} \\ &\leq (I_V^{1/2})^{C\Gamma_n^{(2)}\left(\frac{\Lambda}{c_V\Gamma_n} + \frac{2\Lambda^2}{c_V\Gamma_n^2}\right)}. \end{aligned} \quad (4.19)$$

Gathering (4.17), (4.18) and (4.19), recalling that  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow 0$ , we deduce that:

$$\Upsilon_2^{1/2} \leq \exp\left(\left(\frac{\Lambda}{\sqrt{\Gamma_n}} + \frac{\Lambda^2}{\Gamma_n}\right)e_n\right). \quad (4.20)$$

The first term in (4.16),  $\Upsilon_1$ , is handled by identity (4.15).

$$\begin{aligned} \Upsilon_1 &\leq \mathbb{E}\left[\exp\left(\frac{2\Lambda}{\Gamma_n}\gamma_n\psi_n^{\bar{\varphi}}(X_{n-1}, U_n)\right.\right. \\ &\quad \left.\left.+ \frac{2\Lambda}{\Gamma_n}\sum_{k=1}^{n-1}\left[D_j^{k,\varphi}(X_{k-1}, U_k, Z_k) - C\gamma_k^2\left(1 + \frac{2\Lambda}{\Gamma_n}\right)\sqrt{V}(X_{k-1}) - \frac{2C\Lambda^2}{\Gamma_n^2}\gamma_k^2|U_k|^2\right]\right)\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(\frac{2\Lambda}{\Gamma_n}\gamma_n\psi_n^{\bar{\varphi}}(X_{n-1}, U_n)\right)\middle|\mathcal{F}_{n-1}\right]\right. \\ &\quad \left.\times \exp\left(\frac{2\Lambda}{\Gamma_n}\sum_{k=1}^{n-1}\left[D_j^{k,\varphi}(X_{k-1}, U_k, Z_k) - C\gamma_k^2\left(1 + \frac{2\Lambda}{\Gamma_n}\right)\sqrt{V}(X_{k-1}) - \frac{2C\Lambda^2}{\Gamma_n^2}\gamma_k^2|U_k|^2\right]\right)\right] \\ &\stackrel{\text{(GC)}}{\leq} \exp\left(\frac{2\Lambda}{\Gamma_n}\gamma_n^{1+\frac{3+\beta}{2}}\frac{2[\varphi^{(3)}]_\beta\|\sigma\|_\infty^{(3+\beta)}\mathbb{E}[|U_1|^{3+\beta}]}{(1+\beta)(2+\beta)(3+\beta)} + \frac{4\Lambda^2}{\Gamma_n^2}4\gamma_n^3\|\sigma\|_\infty^2\|\nabla\varphi\|_\infty^2\right) \\ &\quad \left.+ \mathbb{E}\left[\exp\left(\frac{2\Lambda}{\Gamma_n}\sum_{k=1}^{n-1}\left[D_j^{k,\varphi}(X_{k-1}, U_k, Z_k) - C\gamma_k^2\left(1 + \frac{2\Lambda}{\Gamma_n}\right)\sqrt{V}(X_{k-1}) - \frac{C2\Lambda^2}{\Gamma_n^2}\gamma_k^2|U_k|^2\right]\right)\right]. \end{aligned}$$

The last inequality is a consequence of the bound (4.8) in the proof of Lemma 5 and the Lipschitz modulus control of  $\psi_n^{\bar{\varphi}}(X_{n-1}, \cdot)$  in Lemma 1. Hence, we iterate this procedure

and with some positive constants  $\tilde{C}_1, \tilde{C}_2$  not depending on  $\Lambda$  neither  $n$  but only on the assumptions we get :

$$\Upsilon_1 \leq \exp\left(\frac{\tilde{C}_1 \Lambda \Gamma_n^{\frac{5+\beta}{2}}}{\Gamma_n} + \frac{\tilde{C}_2 \Lambda^2 \Gamma_n^{(3)}}{\Gamma_n^2}\right) = \exp\left(\left(\frac{\Lambda}{\sqrt{\Gamma_n}} + \frac{\Lambda^2}{\Gamma_n}\right)e_n\right), \quad (4.21)$$

where, using  $\frac{\Gamma_n^{(2)}}{\sqrt{\Gamma_n}} \rightarrow 0$  for  $\theta > \frac{1}{3}$ . Eventually, inequalities (4.20) and (4.21) yields that:

$$\mathbb{E}\left[\exp\left(\frac{\Lambda}{\Gamma_n} \sum_{k=1}^n D_j^k(X_{k-1}, U_k, Z_k)\right)\right] \leq \exp\left(\left(\frac{\Lambda}{\sqrt{\Gamma_n}} + \frac{\Lambda^2}{\Gamma_n}\right)e_n\right).$$

□

## 5 Numerical Results

This section is a numerical illustration of the deviations results of the empirical measure  $\nu_n$  from Theorem 2. We consider the mono-dimensional case,  $d = r = 1$ . The innovations  $(U_i)_{i \geq 1}$  and  $X_0$  are Gaussian variables. Also, a difficulty is to approximate the jump part of generator  $\mathcal{A}\varphi$ , namely  $\pi(\varphi(x + \kappa(x)\cdot) - \varphi(x))$  for  $x \in \mathbb{R}$ . To avoid this problem, we choose  $(Y_k)_{k \geq 0}$  to be Bernoulli variables, hence we directly get  $\pi(\varphi(x + \kappa(x)\cdot) - \varphi(x)) = \frac{1}{2}(\varphi(x + \kappa(x)) + \varphi(x - \kappa(x))) - \varphi(x)$ . We consider for the coefficients and the test function  $b(x) = -\frac{x}{2}$ , and  $\sigma(x) = \kappa(x) = \varphi(x) = \cos(x)$  in (1.1). Note, in particular, that we have picked a degenerate framework. Thanks to Theorem 2, for  $(\gamma_k)_{k \geq 1} = (k^{-\theta})_{k \geq 1}$ ,  $\theta \in [1/3, 1]$  (corresponding to  $\beta = 1$  therein) the function

$$a \in \mathbb{R}^+ \mapsto g_n(a) := \log\left(\mathbb{P}[|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)| \geq a]\right)$$

is such that for

$$\begin{aligned} g_n(a) &\leq -c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2 + 4\|\kappa\|_\infty^2 \|\nabla\varphi\|_\infty^2} + \log(2C_n) \\ &=: c_n S_\nu(a) + \log(2C_n), \end{aligned}$$

In Figure 5.1, we plot the the curves of  $g_n$  for  $\theta = \frac{1}{3} + 10^{-3}$ . We perform the simulations for  $n = 5 \times 10^4$  in Figure 5.1, the probability is estimated by Monte Carlo simulation with  $MC = 10^4$  realizations of the random variable  $|\sqrt{\Gamma_n}\nu_n(\mathcal{A}\varphi)|$  in the unbiased case.

The Figure 5.1 enhance the fact that  $g_n(a)$  is indeed under a quadratic form in  $a$ . Nevertheless, we see s that the result of Theorem 2 is not sharp, to obtain such a result we have to avoid the dimension dependency and a sharp inequality of Proposition 1.

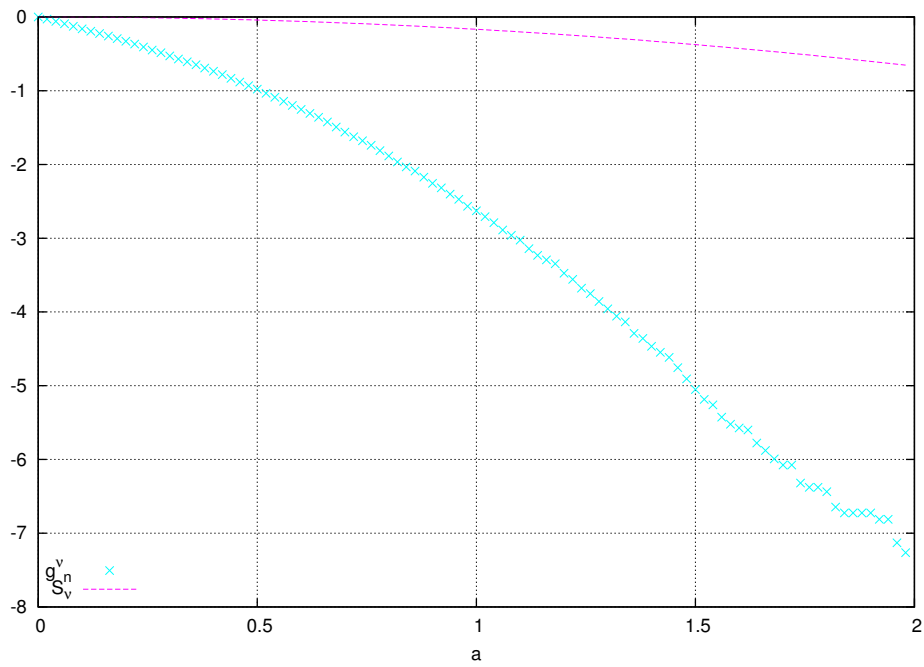


Figure 5.1 – Plot of  $a \mapsto g_n(a)$ , for  $\theta = \frac{1}{3}$ , with  $\varphi(x) = \sigma(x) = \cos(x)$ .







## Part III

# Degenerate Kolmogorov chains





# Chapter 6

## Sharp Schauder Estimates for some Degenerate Kolmogorov Equations

**Abstract :** We provide here some sharp Schauder estimates for degenerate PDEs of Kolmogorov type when the coefficients lie in some suitable anisotropic Hölder spaces and the first order term is non-linear and unbounded. We proceed through a perturbative approach based on forward parametrix expansions.

Due to the low regularizing properties of the degenerate variables, for the procedure to work, we heavily exploit duality results between appropriate Besov spaces.

Our method can be seen as *constructive* and provides, even in the non-degenerate case, an alternative approach to Schauder estimates.

### 1 Introduction and Main Results

We aim at proving, for a fixed time horizon  $T > 0$  and given integers  $n, d \in \mathbb{N}$ , Schauder estimates for degenerate scalar valued Kolmogorov PDEs of the form:

$$\begin{cases} \partial_t u(t, \mathbf{x}) + \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}u(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u(t, \mathbf{x}) a(t, \mathbf{x})) = -f(t, \mathbf{x}), & t \in [0, T], \\ u(T, \mathbf{x}) = g(\mathbf{x}), \end{cases} \quad (1.1)$$

for  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{nd}$  where for all  $i \in \llbracket 1, n \rrbracket$ ,  $\mathbf{x}_i \in \mathbb{R}^d$ . The source  $f$  and the terminal condition  $g$  are scalar. The diffusion matrix  $a$  is  $\mathbb{R}^d \otimes \mathbb{R}^d$ -valued. Also, we denote by  $\mathbf{F}(t, \mathbf{x}) := (\mathbf{F}_1(t, \mathbf{x}), \dots, \mathbf{F}_n(t, \mathbf{x}))$  the vector of the  $\mathbb{R}^d$  valued mappings  $\mathbf{F}_i$  which have, for  $i \in \llbracket 2, n \rrbracket$ , the following structure:

$$\forall (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \mathbf{F}_i(t, \mathbf{x}) := \mathbf{F}_i(t, \mathbf{x}^{i-1:n}), \mathbf{x}^{i-1:n} := (\mathbf{x}_{i-1}, \dots, \mathbf{x}_n). \quad (1.2)$$

In (1.1), the notation  $\mathbf{D} = (D_{\mathbf{x}_1}, \dots, D_{\mathbf{x}_n})$  stands for the full spatial gradient and  $D_{\mathbf{x}_i}$  denotes the partial gradient w.r.t. to  $\mathbf{x}_i$ .

Our goal is to prove that, under suitable non-degeneracy conditions on the diffusion matrix  $a$ , which only acts on the first  $d$  components of the underlying space, and a weak Hörmander like condition on the first order non-linear term  $\mathbf{F}$ , there exists a unique weak

solution to (1.1) which satisfies some appropriate Schauder estimates when the source  $f$ , the terminal condition  $g$  and the coefficients  $a, \mathbf{F}$  lie in appropriate Hölder spaces.

For notational convenience we will denote the spatial operator in (1.1) by  $(L_t)_{t \in [0, T]}$ , i.e. for all  $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$  (space of twice continuously differentiable functions with compact support):

$$L_t \varphi(\mathbf{x}) = \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D} \varphi(\mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 \varphi(\mathbf{x}) a(t, \mathbf{x})). \quad (1.3)$$

Under suitable non-degeneracy and regularity assumptions on  $a, \mathbf{F}$  it can be shown that the martingale problem associated with (1.3) is well posed, see e.g. [Men11], [Men18], [CdRM17]. In that case, there exists a unique weak solution to the stochastic differential equation

$$\begin{aligned} d\mathbf{X}_t^1 &= \mathbf{F}_1(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n) dt + \sigma(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n) dW_t, \\ d\mathbf{X}_t^2 &= \mathbf{F}_2(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n) dt, \\ d\mathbf{X}_t^3 &= \mathbf{F}_3(t, \mathbf{X}_t^2, \dots, \mathbf{X}_t^n) dt, \\ &\vdots \\ d\mathbf{X}_t^n &= \mathbf{F}_n(t, \mathbf{X}_t^{n-1}, \mathbf{X}_t^n) dt, \end{aligned} \quad t \geq 0, \quad (1.4)$$

where  $(W_t)_{t \geq 0}$  is a Brownian motion on some filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . The operator  $(L_t)_{t \geq 0}$  then corresponds to the generator of the process in (1.4) where  $\sigma$  is a square root of  $a$ . The well posedness of the martingale problem in particular implies that (1.1) admits a solution in the mild sense on a suitable function space (see e.g. Kolokoltsov [Kol11]).

The deterministic system (1.1) and its corresponding stochastic counterpart (1.4) appear in many applicative fields from physics to finance (see e.g. [Tal02], [RBT00] or [BPV01]). Another issue concerns the *regularizing* properties of the Brownian motion in the SDE (1.4). In a Hölder framework for the coefficients  $\mathbf{F}, \sigma$ , some minimal thresholds that guarantee weak uniqueness holds where established in [CdRM17]. In the current work, we specifically quantify how adding to these thresholds some suitable regularity, depending on the stage  $i \in \llbracket 1, n \rrbracket$  of the chain and the variable  $j \in \llbracket (i-1) \vee 1, n \rrbracket$ , we get a global similar smoothing effect.

Through the Schauder estimates of Theorem 1, we indeed precisely quantify the *parabolic bootstrap* associated with  $(L_t)_{t \geq 0}$  emphasizing its intrinsic regularization properties. It is known from the seminal work of Lunardi [Lun97] on the topic that the Schauder estimates for degenerate Kolmogorov equations differ from those in the usual non-degenerate setting (see e.g. [Kry96] or [Fri64]). They reflect in some sense the multiple scales in the systems (1.1) and (1.4) (see Section 1.1 below) and are stated in terms of anisotropic Hölder spaces. In particular, those spaces emphasize that the higher is the index of the considered variable in  $\llbracket 1, n \rrbracket$ , the weaker is the associated regularity gain.

Let us now shortly describe some particular cases of dynamics of type (1.1) for which some Schauder estimates have already been proved. We again mention the work by Lunardi [Lun97], who considered the special case of a homogeneous linear drift  $\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x}$

satisfying the condition (1.2). Precisely, the matrix  $\mathbf{A}$  writes in this case:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_{1,1} & \cdots & \cdots & \cdots & \mathbf{a}_{1,n} \\ \mathbf{a}_{2,1} & \cdots & \cdots & \cdots & \mathbf{a}_{2,n} \\ \mathbf{0}_{d,d} & \mathbf{a}_{3,2} & \cdots & \cdots & \mathbf{a}_{3,n} \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & \mathbf{a}_{n,n-1} & \mathbf{a}_{n,n} \end{pmatrix},$$

where the entries  $(\mathbf{a}_{i,j})_{i,j \in \llbracket 1,n \rrbracket^2}$  are in  $\mathbb{R}^d \otimes \mathbb{R}^d$  s.t.  $(\mathbf{a}_{i,i-1})_{i \in \llbracket 2,n \rrbracket}$  are non-degenerate elements of  $\mathbb{R}^d \otimes \mathbb{R}^d$  (which expresses the weak Hörmander condition)\*. Also, the homogeneous diffusion coefficient  $a$  belongs to an appropriate anisotropic Hölder space and asymptotically converges when  $|x| \rightarrow \infty$  to a non-degenerate constant matrix of  $\mathbb{R}^d \otimes \mathbb{R}^d$ . The assumptions on the asymptotic behavior on the diffusion coefficient have then been relaxed by Lorenzi [Lor05b], in the kinetic framework, i.e.  $n = 2$  with the notations of (1.4), up to additional regularity assumptions on  $a$  which could also be unbounded.

Priola established later in [Pri09] Schauder estimates, without dimensional constraints for time homogeneous drifts of the form

$$\mathbf{F}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \begin{pmatrix} \tilde{\mathbf{F}}_1(\mathbf{x}) \\ \mathbf{0}_{(n-1)d,d} \end{pmatrix}, \quad (1.5)$$

for a non-linear drift  $\tilde{\mathbf{F}}_1$  acting on the non-degenerate variable in the expected anisotropic Hölder space. The underlying technique consisted in establishing bounds on the derivatives of the semi-group of the perturbed degenerate Ornstein-Uhlenbeck process (i.e. with  $\tilde{\mathbf{F}}_1$ ) from the usual unperturbed one (with  $(\mathbf{A}\mathbf{x})_1$  only) through the Girsanov theorem assuming first that  $\tilde{\mathbf{F}}_1$  is smooth. This initial smoothness of  $\tilde{\mathbf{F}}_1$  is required in order to compute the associated tangent flows. Through the continuity approach, the author then managed to obtain the estimates for a bounded variable diffusion coefficient lying in the natural *expected* Hölder space similar to the one of [Lun97] with the same asymptotic conditions. The smoothness of  $\tilde{\mathbf{F}}_1$  is then relaxed through an approximation procedure viewing the difference between the Hölder drift and its mollification as a source term and exploiting the estimates established for the *smooth* drift.

Let us also mention the work of Di Francesco and Polidoro [FP06] who derived Schauder estimates for a linear drift of the previous type using an alternative notion of continuity regarding the diffusion coefficient  $a$ , which somehow involves the unbounded transport associated with the drift.

Put it differently, in the current framework of degenerate Kolmogorov equations, focusing on the drift, Schauder estimates hold, to the best of our knowledge, for either linear drifts or Hölder perturbations on the non-degenerate variable of a linear drift.

In the non-degenerate setting, Schauder estimates for unbounded non-linear drift coefficients were obtained under mild smoothness assumptions by Krylov and Priola [KP10]. Their idea was to use the flow associated with the first order vector field in  $L_t$ , i.e.  $\dot{\boldsymbol{\theta}}_t(\mathbf{x}) = \mathbf{F}(t, \boldsymbol{\theta}_t(\mathbf{x}))$ , to precisely get rid of the unbounded terms.

---

\*. Actually the non-zero entries  $(\mathbf{a}_{i,j})$  of  $\mathbf{A}$  can be non-square and simply have full rank in [Lun97]. We restrict here to square matrices for the sake of simplicity.

In the current framework, we will face both difficulties: the degeneracy as well as the non-linearity and unboundedness of the drift. We will prove, in the framework of Hölder spaces for the source  $f$ , the terminal condition  $g$  and the coefficients  $a, \mathbf{F}$ , Schauder estimates similar to those of the previously quoted works. The diffusion coefficient  $a$  and the source term  $f$  will have, as in the non-degenerate case, the same regularity. We mention that, in contrast with the non-degenerate case, this will not be the case for the drift  $\mathbf{F}$  for which some additional smoothness on the degenerate entries  $(\mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$  is needed to guarantee the well posedness of (1.4). In particular, the Hölder indexes of  $\mathbf{F}$  will be above the minimal thresholds appearing in [CdRM17]. The flow associated with the drift term will again play a key role in our setting.

We will here proceed through a perturbative approach. The idea is to perform a first order *parametrix* expansion (or Duhamel expansion) of a solution of (1.1) with mollified coefficients around a suitable linearized Ornstein-Uhlenbeck type semi-group. Since we are interested in controlling derivatives, at least along the non-degenerate variable  $\mathbf{x}_1$ , we will use the so-called *forward parametrix* approach, which had previously been successfully used by Friedman [Fri64] or Il'in *et al.* [IKO62] in the non-degenerate setting to obtain pointwise bounds on the fundamental solution and its derivatives for the corresponding heat-type equation or in [CdR17] to derive strong uniqueness for kinetic SDEs of type (1.4) (i.e.  $n = 2$  with the previous notations). This approach actually allows to exploit cancellation techniques which are crucial when derivatives come in.

On the other hand, for the well posedness of the corresponding martingale problem or density estimates of the fundamental solution of (1.1), the *backward* approach introduced by McKean and Singer [MS67] might be easier. It was extended in the current degenerate setting, which involves unbounded coefficients, and successfully exploited for the two previously indicated applications in [CdRM17] and [DM10]. Unfortunately, this approach does not allow to benefit from the previously mentioned cancellation techniques. This becomes a hard issue in the parabolic setting involving a non-trivial terminal condition.

The perturbative approach is not usual to establish Schauder type estimates. The standard way is to proceed through *a priori* estimates to establish for a given solution of the PDE in a suitable function space, the expected bound. Existence and uniqueness issues, in the considered function space, for the solution of the equation are addressed in a second time. We can refer to [Kry96] for a clear presentation of this approach and to [KP10] for an extension of this method to non-degenerate operators with unbounded drift coefficients. We will here obtain that the solutions of (1.1) with mollified coefficients satisfy, uniformly w.r.t. the mollification parameter, a Schauder type estimate (see Sections 3 to 5 below). From the well posedness of the martingale problem established in [CdRM17] under our current assumptions, we will then derive that the martingale solution to (1.1) actually itself satisfies the Schauder controls. Since we want to be in the *sharpest* possible Hölder setting for the coefficients, source and terminal functions, we will need to establish some subtle controls (in particular we have no true derivatives of the coefficients) which will heavily rely on duality results for Besov spaces (see Section 4.2 below and e.g. Chapter 3 in Lemarié-Rieusset [LR02]).

Let us emphasize that the perturbative approach developed here provides, even in the non-degenerate case, a new alternative to establish Schauder estimates. It can be seen as a constructive one in the sense that, from a sequence of smooth solutions, that uniformly satisfy the expected control, we will extract through convergence in law arguments a limit solution which also satisfies the bound. Uniqueness of the solution in the considered class then again follows from uniqueness in law of the underlying limit process.

The drawback of our approach is that, for the parabolic problem (1.1), we first have to establish our estimates in small time. This is intuitively clear since the perturbative methods (expansions along an *ad hoc proxy*) are precisely designed for small times. To obtain the result for an arbitrary given time, we then have to iterate the estimate. We are therefore *far* from the *optimal* constants for the Schauder estimates established in the non-degenerate setting for time dependent coefficients by Krylov and Priola [KP17]. However, to the best of our knowledge, our approach is the only one allowing to derive Schauder estimates for a parabolic degenerate Kolmogorov equation with fully non-linear drift in Hölder space. Also, we think the strategy developed in the current work should apply for the elliptic degenerate Kolmogorov equation with *good* potential term, i.e. the *negative sign* of the potential would allow to integrate on an infinite time horizon (getting therefore rid of the small time constraint) or for unbounded sources  $f$  that would need in that cases to be somehow controlled by an associated potential as in [KP10].

In connection with the *strong* regularizing properties of Brownian motion, let us mention that the perturbative approach we develop here allows as well to address the problem of strong well posedness for the SDE (1.4) in a Hölder framework for the coefficients. This was done by Chaudru de Raynal in [CdR18] for  $n = 2$ . Let us mention as well in this *kinetic case* the works of Fedrizzi *et al.* [FFPV17] and Zhang [Zha18] who derived strong uniqueness for  $L^p$  drifts on the non-degenerate component and a linear degenerate dynamics in (1.4). We establish strong uniqueness for the full chain for some suitable related Hölder thresholds for the drift in Chapter 7. Let us mention that, from a PDE viewpoint, strong uniqueness for the associated SDE is heavily related to controls of the derivatives of the solution of the PDE in all the directions (including the degenerate ones). In Chapter 7, under some slightly stronger Hölder regularity assumptions on the coefficients (which for strong uniqueness issues then in turn become the source term in the PDE with the Zvonkin approach), we derive some pointwise bounds for the derivatives of the PDEs. Some related issues were also considered under additional smoothness conditions by Lorenzi [Lor05a] in the case of a linear drift.

Before stating our main results, we recall some properties associated with the system (1.1). We first describe in Section 1.1 how the intrinsic multi-scales of the degenerate Kolmogorov like equations appear. We then introduce the appropriate setting of Hölder spaces to consider in Section 1.2. We eventually conclude the introduction stating in Section 1.3 our main results concerning Schauder estimates associated with (1.1).



## 1.1 Intrinsic scales of the system and associated distance

Let us now briefly expose how the system *typically* behaves. To do so, consider the following operator:

$$\mathcal{L}_0 := \partial_t + \langle \mathbf{A}_0 \mathbf{x}, \mathbf{D} \rangle + \frac{1}{2} \Delta_{\mathbf{x}_1}, \quad \mathbf{A}_0 = \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{I}_{d,d} & \ddots & \cdots & \cdots & \vdots \\ \mathbf{0}_{d,d} & \mathbf{I}_{d,d} & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & \mathbf{I}_{d,d} & \mathbf{0}_{d,d} \end{pmatrix}, \quad (1.6)$$

which can be viewed as a typical model for the operator in (1.1). Introducing now, for  $\lambda > 0$ , the dilation operator  $\delta_\lambda : (t, \mathbf{x}) \in \mathbb{R}^+ \times \mathbb{R}^{nd} \mapsto \delta_\lambda(t, \mathbf{x}) = (\lambda^2 t, \lambda \mathbf{x}_1, \lambda^3 \mathbf{x}_2, \dots, \lambda^{2n-1} \mathbf{x}_n) \in \mathbb{R}^+ \times \mathbb{R}^{nd}$ , i.e. with a slight abuse of notation,  $(\delta_\lambda(t, \mathbf{x}))_0 := \lambda^2 t$  and for all  $i \in \llbracket 1, n \rrbracket$ ,  $(\delta_\lambda(t, \mathbf{x}))_i := \lambda^{2i-1} \mathbf{x}_i$ , we have that

$$\mathcal{L}_0 v = 0 \implies \mathcal{L}_0(v \circ \delta_\lambda) = 0. \quad (1.7)$$

This hence lead us to introduce the homogeneous pseudo-norm corresponding to the dilation operator  $\delta_\lambda$ . Precisely, setting for all  $0 \leq s \leq t < +\infty$ ,  $(\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^{nd})^2$ :

$$\mathbf{d}_P((t, \mathbf{x}), (s, \mathbf{y})) = (s - t)^{\frac{1}{2}} + \sum_{i=1}^d |\mathbf{y}_i - \mathbf{x}_i|^{\frac{1}{2i-1}}, \quad (1.8)$$

we indeed have  $\mathbf{d}_P(\delta_\lambda((t, \mathbf{x})), \delta_\lambda((s, \mathbf{y}))) = \lambda \mathbf{d}_P((t, \mathbf{x}), (s, \mathbf{y}))$ . In our current setting we will mainly use the spatial part deriving from the parabolic homogeneous pseudo-distance  $\mathbf{d}_P$  in (1.8). We set accordingly

$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^d |\mathbf{y}_i - \mathbf{x}_i|^{\frac{1}{2i-1}}. \quad (1.9)$$

From a technical point of view, these distances express the spatial homogeneity associated with the intrinsic time scales of the variances for the Gaussian process with generator  $\mathcal{L}_0$ . From a more probabilistic viewpoint the exponents in (1.9) can be related to the characteristic time-scales of the iterated integrals of the Brownian motion (see e.g. [DM10], [CdRM17]).

The non-degeneracy and boundedness assumption on  $a$  as well as the Hörmander condition on the  $(\mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$  that we assume for (1.1)-(1.4) (see assumptions **(UE)** and **(H)** below) will allow us to consider for our analysis the previous pseudo-distances associated with the simplest yet typical equation in the class described by (1.1)-(1.4).

## 1.2 Associated Hölder spaces

We first recall some useful notations and spaces. We denote for  $k \in \mathbb{N}, \beta \in (0, 1)$  by  $\|\cdot\|_{C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)}$ ,  $m \in \{1, d, nd\}$ ,  $\ell \in \{1, d, d^2, nd\}$  the usual homogeneous Hölder space, see

e.g. Krylov [Kry96]. Precisely, for  $\psi \in C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)$ , denoting by  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$  a generic multi-index and  $|\alpha| = \sum_{i=1}^m \alpha_i$ , we define the semi-norm:

$$\begin{aligned} \|\psi\|_{C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)} &:= \sum_{i=1}^k \sup_{|\alpha|=i} \|D^\alpha \psi\|_{L^\infty(\mathbb{R}^m, \mathbb{R}^\ell)} + \sup_{|\alpha|=k} [D^\alpha \psi]_\beta, \\ [D^\alpha \psi]_\beta &:= \sup_{(x,y) \in (\mathbb{R}^m)^2, x \neq y} \frac{|D^\alpha \psi(x) - D^\alpha \psi(y)|}{|x - y|^\beta}, \end{aligned} \quad (1.10)$$

where  $|\cdot|$  denotes the Euclidean norm on the considered space. We will also need to consider the associated subspace with bounded elements. Namely, we set:

$$C_b^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell) := \{\psi \in C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell) : \|\psi\|_{L^\infty(\mathbb{R}^m, \mathbb{R}^\ell)} < +\infty\}.$$

We define correspondingly the Hölder norm:

$$\|\psi\|_{C_b^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)} := \|\psi\|_{C^{k+\beta}(\mathbb{R}^m, \mathbb{R}^\ell)} + \|\psi\|_{L^\infty(\mathbb{R}^m, \mathbb{R}^\ell)}. \quad (1.11)$$

We are now in position to define our inhomogeneous Hölder spaces with multi-index of regularity. Let  $\psi : \mathbb{R}^{nd} \rightarrow \mathbb{R}^\ell$  be a smooth function. We first introduce, for  $i \in \llbracket 1, n \rrbracket$ ,  $x \in \mathbb{R}^d$  the perturbation operator that writes:

$$\forall \mathbf{z} \in \mathbb{R}^{nd}, \Pi_i^x(\psi)(\mathbf{z}) := \psi(\mathbf{z}_1, \dots, \mathbf{z}_i + x, \dots, \mathbf{z}_n). \quad (1.12)$$

We then define for all  $i \in \llbracket 1, n \rrbracket$ , the mapping

$$(\mathbf{z}, x) \in \mathbb{R}^{nd} \times \mathbb{R}^d \longmapsto \psi_i(\mathbf{z}, x) := \Pi_i^x(\psi)(\mathbf{z}). \quad (1.13)$$

Let us introduce the following inhomogeneous Hölder space in  $\mathbf{d}$ -metric: given a parameter  $\gamma \in (0, 1)$ , and  $k \in \mathbb{N}$ , we say that  $\psi$  is in  $C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)$ , if

$$\|\psi\|_{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)} := \sum_{i=1}^n \sup_{\mathbf{z} \in \mathbb{R}^{nd}} \|\psi_i(\mathbf{z}, \cdot)\|_{C_{\mathbf{d}}^{\frac{k+\gamma}{2i-1}}(\mathbb{R}^d, \mathbb{R}^\ell)} < +\infty. \quad (1.14)$$

For the sake of simplicity, we will write:

$$\|\psi\|_{L^\infty} := \|\psi\|_{L^\infty(\mathbb{R}^{nd}, \mathbb{R}^\ell)}, \text{ and } \|\psi\|_{C_{\mathbf{d}}^{k+\gamma}} := \|\psi\|_{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)}.$$

The subscript  $\mathbf{d}$  stands here to indicate the dependence of the Hölder exponents appearing in the r.h.s. on the underlying pseudo-norm  $\mathbf{d}$  reflecting the scale invariance of the system (see equation (1.9) and the comments above for details). Note in particular that, for  $k = 0$ , there exists  $C := C(n, d) \geq 1$  such that:

$$\begin{aligned} C^{-1}[\psi]_{\mathbf{d}}^\gamma &\leq \|\psi\|_{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)} \leq C[\psi]_{\mathbf{d}}^\gamma, \\ [\psi]_{\mathbf{d}}^\gamma &:= \sup_{\mathbf{x} \neq \mathbf{x}', (\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2} \frac{|\psi(\mathbf{x}) - \psi(\mathbf{x}')|}{\mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')}, \end{aligned} \quad (1.15)$$

see also e.g. Lunardi [Lun97].

Also, from (1.11) and (1.14), we write that  $\psi \in C_{b,\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)$  if

$$\|\psi\|_{C_{b,\mathbf{d}}^{k+\gamma}(\mathbb{R}^{nd}, \mathbb{R}^\ell)} := \sum_{i=1}^n \sup_{\mathbf{z} \in \mathbb{R}^{nd}} \|\psi_i(\mathbf{z}, \cdot)\|_{C_b^{\frac{k+\gamma}{2i-1}}(\mathbb{R}^d, \mathbb{R}^\ell)} < +\infty.$$

From now on, we will denote:

$$\|\psi\|_{C_{b,\mathbf{d}}^\gamma} := \|\psi\|_{C_{b,\mathbf{d}}^\gamma(\mathbb{R}^{nd}, \mathbb{R}^\ell)}.$$

Finally, through the article, we use the following notation for all functions  $\varphi_1 \in L^\infty([0, T], C_{b,\mathbf{d}}^{k+\gamma}(\mathbb{R}^m, \mathbb{R}^\ell))$  and  $\varphi_2 \in L^\infty([0, T], C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^m, \mathbb{R}^\ell))$ :

$$\|\varphi_1\|_{L^\infty(C_{b,\mathbf{d}}^{k+\gamma})} := \sup_{t \in [0, T]} \|\varphi_1(t, \cdot)\|_{C_{b,\mathbf{d}}^{k+\gamma}(\mathbb{R}^m, \mathbb{R}^\ell)}, \text{ and } \|\varphi_2\|_{L^\infty(C_{\mathbf{d}}^{k+\gamma})} := \sup_{t \in [0, T]} \|\varphi_2(t, \cdot)\|_{C_{\mathbf{d}}^{k+\gamma}(\mathbb{R}^m, \mathbb{R}^\ell)}.$$

### 1.3 Assumptions and main result

With these notations at hand we can now state our assumptions and main results. In the following, we will assume:

**(UE) Uniform Ellipticity of the diffusion Coefficient.** There exists  $\kappa \geq 1$  s.t. for all  $(t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^{nd}$ ,  $z \in \mathbb{R}^d$ ,

$$\kappa^{-1}|z|^2 \leq \langle a(t, \mathbf{x})z, z \rangle \leq \kappa|z|^2,$$

where  $|\cdot|$  again denotes the Euclidean norm and  $\langle \cdot, \cdot \rangle$  is the inner product.

**(H) Weak Hörmander like condition.** For all  $i \in \llbracket 2, n \rrbracket$ , there exists a closed convex subset  $\mathcal{E}_{i-1} \subset GL_d(\mathbb{R})$  (set of invertible  $d \times d$  matrices) s.t., for all  $t \geq 0$  and  $(\mathbf{x}_{i-1}, \dots, \mathbf{x}_n) \in \mathbb{R}^{(n-i+2)d}$ ,  $D_{\mathbf{x}_{i-1}} F_i(t, \mathbf{x}_{i-1}, \dots, \mathbf{x}_n) \in \mathcal{E}_{i-1}$ . For example,  $\mathcal{E}_{i-1}$  may be a closed ball included in  $GL_d(\mathbb{R})$ , which is an open set.

**(S) Smoothness of the Coefficients.** Fix  $\gamma \in (0, 1)$ . We suppose the following conditions hold.

- (i) **Smoothness of the diffusion coefficient.** We assume that  $a$  is measurable in time and that  $a \in L^\infty([0, T], C_{b,\mathbf{d}}^\gamma(\mathbb{R}^{nd}, \mathbb{R}^d \otimes \mathbb{R}^d))$ .
- (ii) **Smoothness of the drift in time.** We only assume here that the drift is measurable in time and bounded at the origin, i.e. the measurable mapping  $t \mapsto \mathbf{F}(t, \mathbf{0})$  is bounded.
- (iii) **Smoothness of the drift in space.** We now state, for each level  $i \in \llbracket 1, n \rrbracket$  the smoothness assumptions on the drift component  $\mathbf{F}_i$  (see the remark below for more explanations):

$$\mathbf{F}_i \in L^\infty([0, T], C_{\mathbf{d}}^{(2i-3) \vee 0 + \gamma}(\mathbb{R}^{((n-i+2) \wedge n)d}, \mathbb{R}^d)). \quad (1.16)$$

For a fixed parameter  $\gamma \in (0, 1)$ , we will say that **(A)** is in force as soon as **(UE)**, **(H)**, **(S)** hold.

**Remark 1** *Let us come back to assumption **(S)**-(iii), which may seem difficult to understand at first sight. Namely, we here explain a little bit how the particular thresholds appearing in this assumption come from as well as the precise regularity imposed on each component of the drift  $\mathbf{F}$  w.r.t. any space variables.*

- Note first that for  $i = 1$  assumption **(S)**-(iii) readily says, with the previous notations for Hölder spaces, that  $\mathbf{F}_1 \in L^\infty([0, T], C_{\mathbf{d}}^\gamma(\mathbb{R}^{nd}, \mathbb{R}^d))$ .

- For each level  $i \in \llbracket 2, n \rrbracket$  of the chain in (1.2), we shall consider different types of assumptions on  $\mathbf{F}_i$  depending on the variables  $\mathbf{x}_{i-1}$  and  $\mathbf{x}^{i:n} = (\mathbf{x}_i, \dots, \mathbf{x}_n)$  respectively. Let us now fix  $i \in \llbracket 2, n \rrbracket$ .

The component  $\mathbf{x}_{i-1}$  is hence the one which transmits the noise. Coherently with the usual Hörmander setting, we need some differentiability of  $\mathbf{F}_i$  w.r.t.  $\mathbf{x}_{i-1}$ . In order to have a global coherence, in terms of time-space homogeneity, for all the considered variables, the specific smoothness to be considered for that variable is that  $\mathbf{F}_i(t, \cdot, \mathbf{x}^{i:n})$  is in  $C^{1+\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d, \mathbb{R}^d)$ . Recalling now the previous definition of  $\mathbf{d}$  and of the associated Hölder spaces, we have  $C^{1+\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d, \mathbb{R}^d) = C^{\frac{2i-3+\gamma}{2(i-1)-1}}(\mathbb{R}^d, \mathbb{R}^d) = C_{\mathbf{d}}^{2i-3+\gamma}(\mathbb{R}^d, \mathbb{R}^d)$ .

Now, at level  $i$ , the components  $\mathbf{x}^{i:n}$  are above the current characteristic time-scale, i.e. the vector of their associated time rescaling, which writes according to the homogenous quasi-metric  $\mathbf{d}_P$  in (1.8) as  $(t^{i-\frac{1}{2}}, t^{(i+1)-\frac{1}{2}}, \dots, t^{n-\frac{1}{2}})$ , has in small time entries that are actually smaller or equal than the time rescaling of the current variable  $\mathbf{x}_i$  in  $t^{i-1/2}$ . We recall as well that, in order to have the well posedness of the martingale problem associated with the operator  $(L_t)_{t \geq 0}$  in (1.3), some natural minimal thresholds of Hölder continuity appear for these variables. Precisely, at level  $i$ ,  $\mathbf{F}_i$  must be Hölder continuous in  $\mathbf{x}_j$ ,  $j \in \llbracket i, n \rrbracket$ , with index strictly greater than  $\frac{2i-3}{2j-1}$  (see [CdRM17] for details). Here, still to have a global coherence, in terms of time-space homogeneity, for all the considered variables, we assume that  $\mathbf{F}_i$  is  $\frac{2i-3}{2j-1} + \frac{\gamma}{2j-1}$  Hölder continuous in its  $j^{\text{th}}$  variable. This precisely corresponds to the minimal threshold required to which we add the intrinsic  $\gamma$ -Hölder regularity w.r.t to the associated scale appearing in  $\mathbf{d}$  for the considered entry. Thus, with a slight abuse of notations,  $z \mapsto \mathbf{F}_i(t, \mathbf{x}_{i-1}, \mathbf{x}_i, \dots, \mathbf{x}_{j-1}, z, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n)$  is supposed to be in  $C^{\frac{2i-3}{2j-1} + \frac{\gamma}{2j-1}}(\mathbb{R}^d, \mathbb{R}^d) = C_{\mathbf{d}}^{2i-3+\gamma}(\mathbb{R}^d, \mathbb{R}^d)$ .

- “Gathering” the regularity conditions assumed on each variable for each component of  $\mathbf{F}$  hence gives assumption **(S)**-(iii).

**Remark 2** *Concerning the time regularity in the previous assumptions we can refer to the earlier work of Kruzhkov et al. [KCL75] who first consider this type of regularity to establish Schauder estimates in the classical non degenerate framework with bounded coefficients. We can also mention Lorenzi [Lor11] for extensions to unbounded coefficients.*

We are now in position to state our main result.

**Theorem 1 (Schauder Estimates for degenerate Kolmogorov Equations.)**

Let  $\gamma \in (0, 1)$  be given. Suppose that **(A)** is in force and that the terminal condition  $g$  and source term  $f$  of the Cauchy problem (1.1) satisfy:  $g \in C_{b,d}^{2+\gamma}(\mathbb{R}^{nd}, \mathbb{R})$  and  $f \in L^\infty([0, T], C_{b,d}^\gamma(\mathbb{R}^{nd}, \mathbb{R}))$ .

Then, there exists a unique weak solution  $u$  in  $C_{b,d}^{2+\gamma}(\mathbb{R}^{nd}, \mathbb{R})$  to (1.1). Furthermore, there exists a constant  $C_{(1)} := C_{(1)}(\mathbf{(A)}, T)$  s.t.

$$\|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq C_{(1)} (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)}). \quad (1.17)$$

Section 2 below is dedicated to the presentation and description of the various steps that we perform to obtain our main result of Theorem 1. From now on we will denote by  $C$  a generic constant that may change from line to line but only depends on known parameters in **(A)** and the considered fixed final time  $T$ , i.e.  $C := C(\mathbf{(A)}, T)$ . We reserve the notation  $c$  for generic constants that may also change from line to line, depend on **(A)** but are also independent of  $T$ , i.e.  $c := c(\mathbf{(A)})$ .

## 2 Detailed Guide to the proof

We describe in this section the approach we follow to derive our main results. The point is to emphasize the various difficulties arising and to introduce the adapted tools that can be used to circumvent them.

The first step of our strategy is to mollify equation (1.1) in order to get a well-posed Cauchy problem in the classical sense. Precisely, for  $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$ ,  $m \in \mathbb{N}$  and  $t \in [0, T]$  we define the operator:

$$L_t^m \varphi(x) := \langle \mathbf{F}_m(t, \mathbf{x}), \mathbf{D}\varphi(\mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 \varphi(\mathbf{x}) a_m(t, \mathbf{x})), \quad (2.1)$$

where  $\mathbf{F}_m, a_m$  are mollified versions in space of the initial coefficients  $\mathbf{F}, a$  in (1.3), i.e.  $\mathbf{F}_m(t, \mathbf{x}) = \mathbf{F}(t, \cdot) \star \phi_m(\mathbf{x})$ ,  $a_m(t, \mathbf{x}) = a(t, \cdot) \star \phi_m(\mathbf{x})$ , where for all  $\mathbf{z} \in \mathbb{R}^{nd}$ ,  $\phi_m(\mathbf{z}) := m^{nd} \phi(\mathbf{z}m)$  for a smooth, i.e.  $C^\infty$ , non-negative function  $\phi : \mathbb{R}^{nd} \rightarrow \mathbb{R}^+$  s.t.  $\int_{\mathbb{R}^{nd}} \phi(\mathbf{z}) d\mathbf{z} = 1$  and the previous convolutions are to be understood componentwise. We introduce correspondingly the stochastic differential equation with generator  $(L_t^m)_{t \geq 0}$ . Namely, for fixed  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$  and  $s \geq t$ ,

$$\mathbf{X}_s^{m,t,\mathbf{x}} = \mathbf{x} + \int_t^s \mathbf{F}_m(u, \mathbf{X}_u^{m,t,\mathbf{x}}) du + \int_t^s B \sigma_m(u, \mathbf{X}_u^{m,t,\mathbf{x}}) dW_u, \quad (2.2)$$

where  $\sigma_m$  is a square root of  $a_m$ . The dynamics in (2.2) is similar to the one in (1.4) up to the mollification of the coefficients. It can be deduced from the well-posedness of the martingale problem, which holds under our current assumptions from [CdRM17], that  $(\mathbf{X}_t^m)_{t \in [0, T]} \Rightarrow_m (\mathbf{X}_t)_{t \in [0, T]}$  (convergence in law on the path space) where  $(\mathbf{X}_t)_{t \in [0, T]}$  is the unique weak solution of (1.4) (see also [SV79]).

Consider now mollified versions  $f_m, g_m$  of the source  $f$  and the final condition  $g$  in (1.1). It is then rather direct to derive through stochastic flows techniques, see e.g. Kunita [Kun97], that

$$u_m(t, \mathbf{x}) := \mathbb{E}[g_m(\mathbf{X}_T^{m,t,\mathbf{x}})] + \int_t^T \mathbb{E}[f_m(s, \mathbf{X}_s^{m,t,\mathbf{x}})] ds, \quad (2.3)$$

belongs for any given  $m$  to  $C_b^\infty(\mathbb{R}^{nd}, \mathbb{R})$  (space of infinitely differentiable functions with bounded derivatives) and precisely solves for all  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$\begin{cases} \partial_t u_m(t, \mathbf{x}) + \langle \mathbf{F}_m(t, \mathbf{x}), \mathbf{D}u_m(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}) a_m(t, \mathbf{x})) = -f_m(t, \mathbf{x}), & t \in [0, T), \\ \times \mathbb{R}^{nd}, \\ u_m(T, \mathbf{x}) = g_m(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^{nd}. \end{cases} \quad (2.4)$$

The idea is now to obtain controls of the norms  $\|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}$  which are uniform w.r.t. the mollifying parameter  $m$ . To this end, we will use a perturbative method by expanding  $u_m$  around a suitable Ornstein-Uhlenbeck like Gaussian proxy corresponding to an appropriate linearization of the dynamics in (2.2). Consider first the deterministic dynamics deriving from (2.2) obtained setting  $\sigma_m$  to 0, i.e.

$$\dot{\boldsymbol{\theta}}_{v,\tau}^m(\boldsymbol{\xi}) = \mathbf{F}_m(v, \boldsymbol{\theta}_{v,\tau}^m(\boldsymbol{\xi})), \quad v \in [0, T], \quad \boldsymbol{\theta}_{\tau,\tau}^m(\boldsymbol{\xi}) = \boldsymbol{\xi}, \quad (2.5)$$

where  $(\tau, \boldsymbol{\xi}) \in [0, T] \times \mathbb{R}^{nd}$  are *freezing parameters*, respectively in time and space to be specified.

Fix  $0 \leq t < s \leq T$  and  $\mathbf{x} \in \mathbb{R}^{nd}$ . The typical linearization of (2.2) on the time interval  $[t, s]$  around  $(\boldsymbol{\theta}_{v,\tau}^m(\boldsymbol{\xi}))_{v \in [t,s]}$  writes:

$$\begin{aligned} d\tilde{\mathbf{X}}_v^{m,(\tau,\boldsymbol{\xi})} &= [\mathbf{F}_m(v, \boldsymbol{\theta}_{v,\tau}^m(\boldsymbol{\xi})) + D\mathbf{F}_m(v, \boldsymbol{\theta}_{v,\tau}^m(\boldsymbol{\xi}))(\tilde{\mathbf{X}}_v^{m,(\tau,\boldsymbol{\xi})} - \boldsymbol{\theta}_{v,\tau}^m(\boldsymbol{\xi}))]dv + B\sigma_m(v, \boldsymbol{\theta}_{v,\tau}^m(\boldsymbol{\xi}))dW_v, \\ \forall v \in [t, s], \quad \tilde{\mathbf{X}}_t^{m,(\tau,\boldsymbol{\xi})} &= \mathbf{x}, \end{aligned} \quad (2.6)$$

where for all  $\mathbf{z} \in \mathbb{R}^{nd}$ ,

$$D\mathbf{F}_m(v, \mathbf{z}) := \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ D_{\mathbf{z}_1} \mathbf{F}_{m,2}(v, \mathbf{z}) & \mathbf{0}_{d,d} & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & D_{\mathbf{z}_2} \mathbf{F}_{m,3}(v, \mathbf{z}^{2:n}) & \mathbf{0}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & D_{\mathbf{z}_{n-1}} \mathbf{F}_{m,n}(v, \mathbf{z}_{n-1}, \mathbf{z}_n) & \mathbf{0}_{d,d} \end{pmatrix} \quad (2.7)$$

denotes the subdiagonal of the Jacobian matrix  $D_{\mathbf{z}} \mathbf{F}_m(v, \cdot)$  at point  $\mathbf{z}$ . From our previous assumptions (non-degeneracy of  $\sigma$  and Hörmander like condition), the Gaussian process with dynamics (2.6) admits a well controlled multi-scale density  $\tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \cdot)$  (see e.g. Section 3.1 below and for instance [DM10], [CdRM17]). Namely, there exists  $C := C((\mathbf{A}), T) \geq 1$  s.t. for  $j \in \{1, 2\}$  and all  $(i, k) \in \llbracket 1, n \rrbracket^2$ , for all  $0 \leq t < s \leq T$ ,  $(\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^{nd})^2$ :

$$\begin{aligned} & |D_{\mathbf{x}_k} D_{\mathbf{x}_i}^j \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})| \\ & \leq \frac{C}{(s-t)^{\frac{2k-1}{2} + \frac{j(2i-1)}{2} + \frac{n^2 d}{2}}} \exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{m,(\tau,\boldsymbol{\xi})}(\mathbf{x}) - \mathbf{y})|^2\right) \\ & =: \frac{C}{(s-t)^{\frac{2k-1}{2} + \frac{j(2i-1)}{2}}} \tilde{p}_{C^{-1}}^{(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}), \end{aligned} \quad (2.8)$$

where  $\mathbf{m}_{s,t}^{m,(\tau,\xi)}(\mathbf{x})$  stands for the mean of  $\tilde{\mathbf{X}}_s^{m,(\tau,\xi)}$  and for all  $u > 0$ ,  $\mathbb{T}_u$  is the *intrinsic* scale matrix:

$$\mathbb{T}_u = \begin{pmatrix} u\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & u^2\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & u^n\mathbf{I}_{d,d} \end{pmatrix}, \quad (2.9)$$

that is, the  $i^{\text{th}}$  diagonal entry of  $u^{-\frac{1}{2}}\mathbb{T}_u$  reflects the time order of the variances of the  $(i-1)^{\text{th}}$  iterated integral of the standard Brownian motion at time  $u$ . Observe as well that the time singularities in (2.8) precisely reflect the typical scale of the associated variable, i.e. differentiating in  $\mathbf{x}_k$  yields an additional time singularity in  $(s-t)^{-k+\frac{1}{2}}$  where  $(s-t)^{k-\frac{1}{2}}$  is exactly the order of the standard deviation of the  $(k-1)^{\text{th}}$  iterated integral of the Brownian motion.

Denoting by  $(\tilde{L}_v^{m,(\tau,\xi)})_{v \in [t,T]}$  the generator of (2.6), it also holds that:

$$\begin{aligned} (\partial_s - (\tilde{L}_s^{m,(\tau,\xi)})^*)\tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y}) &= 0, \tilde{p}^{m,(\tau,\xi)}(t, s, \cdot, \mathbf{y}) \rightarrow_{s \downarrow t} \delta_{\mathbf{y}}(\cdot), \\ (\partial_t + \tilde{L}_t^{m,(\tau,\xi)})\tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y}) &= 0, \tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \cdot) \rightarrow_{t \uparrow s} \delta_{\mathbf{x}}(\cdot). \end{aligned}$$

The above equations are respectively the *forward* and *backward* Kolmogorov equations. In the first one, the operator  $(\tilde{L}_s^{m,(\tau,\xi)})^*$  acts on the *forward* variable  $\mathbf{y}$  whereas in the second one,  $\tilde{L}_t^{m,(\tau,\xi)}$  acts in the *backward* variable  $\mathbf{x}$ . We will use the notation  $\tilde{P}_{T,t}^{m,(\tau,\xi)}$  for the corresponding semi-group, i.e.

$$\tilde{P}_{T,t}^{m,\xi} g_m(\mathbf{x}) := \int_{\mathbb{R}^{nd}} \tilde{p}^{m,\xi}(t, T, \mathbf{x}, \mathbf{y}) g_m(\mathbf{y}) d\mathbf{y},$$

as well as

$$\tilde{G}^{m,\xi} f_m(t, \mathbf{x}) := \int_t^T ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,\xi}(t, s, \mathbf{x}, \mathbf{y}) f_m(s, \mathbf{y}) d\mathbf{y}, \quad (2.10)$$

for the associated Green kernel (with fixed final time  $T > 0$ ).

For fixed  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$  and the above Gaussian proxy, for which  $(\tau, \xi)$  still remain to be specified, we recall that Duhamel's formula (first order parametrix expansion) yields that:

$$\begin{aligned} u_m(t, \mathbf{x}) &= \tilde{P}_{T,t}^{m,(\tau,\xi)} g_m(\mathbf{x}) + \tilde{G}^{m,(\tau,\xi)} f_m(t, \mathbf{x}) \\ &+ \int_t^T ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\xi)}) u_m(s, \mathbf{y}) d\mathbf{y}. \end{aligned} \quad (2.11)$$

Recall from the statement of our main Theorem 1 that we have to give bounds on  $\|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}$ . For this introduction to the proof, we will focus on the contributions  $D_{\mathbf{x}_1}^2 u_m$ , that already exhibits all the difficulties and for which we want to establish a control in time-space supremum norm and for the  $\gamma$ -Hölder modulus associated with the distance  $\mathbf{d}$ .

Differentiating in  $D_{\mathbf{x}_1}^2$  equation (2.11) gives:

$$\begin{aligned} D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}) &= D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{m,(\tau,\xi)} g_m(\mathbf{x}) + D_{\mathbf{x}_1}^2 \tilde{G}^{m,(\tau,\xi)} f_m(t, \mathbf{x}) \\ &+ \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\xi)}) u_m(s, \mathbf{y}) d\mathbf{y}. \end{aligned}$$

Concentrating on the last term, which turns out to be the most delicate, we see that the choice of  $(\tau, \boldsymbol{\xi})$  must be made in order to equilibrate the time singularities coming from  $D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})$ . Let us first consider the non-degenerate part coming from the difference  $(L_s^m - \tilde{L}_s^{m,(\tau,\boldsymbol{\xi})})u_m(s, \mathbf{y})$  which explicitly writes from (2.1) and (2.6):

$$\begin{aligned} & \Delta_{1, \mathbf{F}_m, \sigma_m}(\tau, s, \mathbf{y}, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}), u_m) \\ &= \langle \mathbf{F}_{m,1}(s, \mathbf{y}) - \mathbf{F}_{m,1}(s, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi})), D_{\mathbf{y}_1} u_m(s, \mathbf{y}) \rangle \\ & \quad + \frac{1}{2} \text{Tr} \left( (a_m(s, \mathbf{y}) - a_m(s, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}))) D_{\mathbf{y}_1}^2 u_m(s, \mathbf{y}) \right), \end{aligned}$$

and can be upper-bounded from the Hölder continuity assumption (w.r.t. the underlying homogeneous metric  $\mathbf{d}$ ) on  $\mathbf{F}_1$  and  $a$  as:

$$\begin{aligned} & |\Delta_{1, \mathbf{F}_m, \sigma_m}(\tau, s, \mathbf{y}, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}), u_m)| \\ & \leq \left( [\mathbf{F}_1(s, \cdot)]_{\mathbf{d}, \gamma} \|D_{\mathbf{y}_1} u_m(s, \cdot)\|_{L^\infty} + \frac{1}{2} [a(s, \cdot)]_{\mathbf{d}, \gamma} \|D_{\mathbf{y}_1}^2 u_m(s, \cdot)\|_{L^\infty} \right) \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi})). \end{aligned} \quad (2.12)$$

The contribution  $\mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}))$  in the above r.h.s. must then equilibrate the time singularity in  $(s-t)^{-1}$  coming from  $D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})$  (see (2.8) and Proposition (2) below). This is possible if  $\mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}))$  is *compatible* with the off-diagonal bound  $(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{m,(\tau,\boldsymbol{\xi})}(\mathbf{x}) - \mathbf{y})|^2$  in (2.8).

This is precisely the case considering  $(\tau, \boldsymbol{\xi}) = (t, \mathbf{x})$  which gives

$$\mathbf{m}_{s,t}^{m,(\tau,\boldsymbol{\xi})}(\mathbf{x})|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} = \boldsymbol{\theta}_{s,t}^m(\mathbf{x}), \quad (2.13)$$

as it can readily be checked from (2.5), (2.6) (taking the expectation) and the Grönwall's Lemma. Therefore, observing precisely from the metric homogeneity (see equations (1.8) and (1.9)) that:

$$\begin{aligned} \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,t}^m(\mathbf{x})) &= (s-t)^{\frac{\gamma}{2}} \mathbf{d}^\gamma \left( (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \mathbf{y}, (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \boldsymbol{\theta}_{s,t}^m(\mathbf{x}) \right) \\ &\leq C(s-t)^{\frac{\gamma}{2}} \left( \sum_{i=1}^n |(s-t)^{\frac{2i-1}{2}} (\mathbf{y} - \boldsymbol{\theta}_{s,t}^m(\mathbf{x}))_i|^{\frac{\gamma}{2i-1}} \right), \end{aligned}$$

we get that the terms of last contribution in the above r.h.s. can precisely be absorbed by the exponential off-diagonal bound in (2.8).

We therefore eventually derive for the non-degenerate contribution with the notation of (2.8):

$$\begin{aligned} & \left| \int_t^T ds \int_{\mathbb{R}^d} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \Delta_{1, \mathbf{F}_m, \sigma_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y}, u_m) d\mathbf{y} \right|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\ & \leq \int_t^T \frac{ds}{(s-t)^{1-\frac{\gamma}{2}}} \int_{\mathbb{R}^{nd}} C \tilde{p}_{C^{-1}}^{(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \left( \|D_{\mathbf{y}_1} u_m(s, \cdot)\|_{L^\infty} + \|D_{\mathbf{y}_1}^2 u_m(s, \cdot)\|_{L^\infty} \right) d\mathbf{y} \\ & \leq \frac{2C}{\gamma} (T-t)^{\frac{\gamma}{2}} (\|D_{\mathbf{y}_1} u_m\|_{L^\infty} + \|D_{\mathbf{y}_1}^2 u_m\|_{L^\infty}) \\ & \leq \frac{2C}{\gamma} (T-t)^{\frac{\gamma}{2}} \|u_m\|_{L^\infty(C_{b,\mathbf{d}}^{2+\gamma})}, \end{aligned} \quad (2.14)$$



yielding precisely a time smoothing effect corresponding exactly to the Hölder continuity exponent  $\gamma$  of the coefficients. The previous choice of  $(\tau, \boldsymbol{\xi})$  is known as the *forward* parametrix and seems adapted as soon as one is led to estimate derivatives of the solution. This was for instance already the choice performed in the non-degenerate case by Il'in *et al* [IKO62] or Friedman [Fri64] in relation with Schauder estimates or by Di Francesco and Polidoro [FP06] in the current degenerate Kolmogorov setting with a linear drift.

Let us mention that, as far as one is concerned with density estimates, which formally amounts to replace  $u_m(s, \mathbf{x})$ ,  $u_m(s, \mathbf{y})$  in (2.11) with  $p^m(t, T, \mathbf{x}, \mathbf{z})$ ,  $p^m(s, T, \mathbf{y}, \mathbf{z})$  (density at some fixed point  $\mathbf{z} \in \mathbb{R}^{nd}$  of  $\mathbf{X}_T^m$  starting from  $\mathbf{x}$  at time  $t$ ), or with the well-posedness of the martingale problem, another choice, consisting in freezing in  $(\tau, \boldsymbol{\xi}) = (s, \mathbf{y})$  in the above equation, could also be considered. Note that the freezing parameters would here depend on the time and spatial integration variables. This *backward* approach was first introduced by McKean and Singer [MS67] and led successfully to density estimates and well-posedness of the martingale problem for the current model (1.4) in the respective works [DM10], [Men18], [CdRM17].

However, when dealing with derivatives, the forward perturbative approach appears more flexible since it allows to exploit cancellation techniques whereas this is much trickier in the backward case for which  $\tilde{p}^{m,(s,\mathbf{y})}(t, s, \mathbf{x}, \mathbf{y})$  is not a density w.r.t.  $\mathbf{y}$ . Some associated errors associated with this approach are thoroughly discussed in [CdRM17].

Let us now turn to the contributions associated with the degenerate variables in the difference  $(L_s^m - \tilde{L}_s^{m,(\tau,\boldsymbol{\xi})})u_m(s, \mathbf{y})$ . They precisely write:

$$\begin{aligned} & \sum_{i=2}^n \left\langle \left( \mathbf{F}_{m,i}(s, \mathbf{y}) - \mathbf{F}_{m,i}(s, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_{m,i}(s, \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,\tau}^m(\boldsymbol{\xi}))_{i-1} \right), D_{\mathbf{y}_i} u_m(s, \mathbf{y}) \right\rangle \\ & =: \sum_{i=2}^n \left\langle \Delta_{i, \mathbf{F}_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u_m(s, \mathbf{y}) \right\rangle. \end{aligned} \quad (2.15)$$

Under the current assumptions, we do not expect to have uniform controls w.r.t. to the smoothing parameter  $m$  for the derivatives  $(D_{\mathbf{y}_i} u_m)_{i \in \llbracket 2, n \rrbracket}$  in the degenerate directions. Our strategy will first consist for those terms in performing an integration by parts leading to:

$$\begin{aligned} & \left| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \left\langle \Delta_{i, \mathbf{F}_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u_m(s, \mathbf{y}) \right\rangle \right| \Bigg|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\ & \leq \sum_{i=2}^n \int_t^T ds \left| \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{y}_i} \cdot \left( (D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \otimes \Delta_{i, \mathbf{F}_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y})) \right) u_m(s, \mathbf{y}) \right| \Bigg|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\ & =: \sum_{i=2}^n \int_t^T ds \left| \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{y}_i} \cdot (\Theta_{i,(t,\mathbf{x})}^{m,\alpha}(s, \mathbf{y})) u_m(s, \mathbf{y}) \right| \Bigg|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})}, \end{aligned} \quad (2.16)$$

where the notation “ $\otimes$ ” between two tensors means the usual tensor product. In particular,  $(D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \otimes \Delta_{i, \mathbf{F}_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y}))$  is a tensor lying in  $(\mathbb{R}^d)^{\otimes 3}$ . Furthermore,  $D_{\mathbf{y}_i} \cdot$  refers to an extended form of the divergence over the  $i^{\text{th}}$  variable ( $\mathbf{y}_i \in \mathbb{R}^d$ ). Precisely,

from (2.16), we rewrite for all  $i \in \llbracket 2, n \rrbracket$ ,  $(s, \mathbf{y}) \in [t, T] \mathbb{R}^{nd}$

$$D_{\mathbf{y}_i} \cdot (\Theta_{i,(t,\mathbf{x})}^{m,\alpha}(s, \mathbf{y})) = \sum_{j=1}^d \partial_{\mathbf{y}_i^j} \left( D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) (\Delta_{i,\mathbf{F}_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y}))_j \right),$$

with  $\mathbf{y}_i = (\mathbf{y}_i^1, \dots, \mathbf{y}_i^n)$ . In other words, this ‘‘enhanced’’ divergence form decreases by one the order of the input tensor. As a particular case, if  $d = 1$ ,  $\Delta_{i,\mathbf{F}_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y})$  is a scalar and the divergence form corresponds to the standard differentiation, i.e.  $D_{\mathbf{y}_i} \cdot = \partial_{\mathbf{y}_i}$ .

Here  $\alpha = (2, 0, \dots, 0)$  is the multi-index which reflects that the above contribution involves the second order derivatives of the frozen heat-kernel w.r.t. to its non-degenerate components.

In view of our main estimates in Theorem 1, we will use the duality between suitable Besov spaces to derive bounds for the spatial integrals in (2.16). Introduce, for all fixed  $i \in \llbracket 2, n \rrbracket$  and any  $(\mathbf{y}_1, \dots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \dots, \mathbf{y}_n) =: (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \in \mathbb{R}^{(n-1)d}$  the mappings

$$\begin{aligned} u_m^{i,(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})} : \mathbf{y}_i &\mapsto u_m(s, \mathbf{y}_{1:i-1}, \mathbf{y}_i, \mathbf{y}_{i+1:n}), \\ \Psi_m^{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})} : \mathbf{y}_i &\mapsto D_{\mathbf{y}_i} \cdot (\Theta_{i,(t,\mathbf{x})}^{m,\alpha}(s, \mathbf{y})). \end{aligned} \quad (2.17)$$

The underlying idea is that we actually want, for the  $i^{\text{th}}$  variable, to control uniformly in  $m$  the Hölder modulus  $[u_m^{i,(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}]_{\frac{2+\gamma}{2i-1}}$  uniformly in  $(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \in [t, T] \times \mathbb{R}^{(n-1)d}$ .

To prove this property we recall that, setting  $\tilde{\alpha}_i := \frac{2+\gamma}{2i-1}$ ,  $C_b^{\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R}) = B_{\infty,\infty}^{\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R})$  with the usual notations for Besov spaces (see e.g. Triebel [Tri83]).

Let us now recall some definitions/characterizations from Section 2.6.4 of Triebel [Tri83]. For  $\tilde{\alpha} \in \mathbb{R}, q \in (0, +\infty], p \in (0, \infty]$ ,  $B_{p,q}^{\tilde{\alpha}}(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{\mathcal{H}_{p,q}^{\tilde{\alpha}}} < +\infty\}$  where  $\mathcal{S}(\mathbb{R}^d)$  stands for the Schwarz class and

$$\|f\|_{\mathcal{H}_{p,q}^{\tilde{\alpha}}} := \|\varphi(D)f\|_{L^p(\mathbb{R}^d)} + \left( \int_0^1 v^{(m-\frac{\tilde{\alpha}}{2})q} \|\partial_v^m h_v \star f\|_{L^p(\mathbb{R}^d)}^q \frac{dv}{v} \right)^{\frac{1}{q}}, \quad (2.18)$$

with  $\varphi \in C_0^\infty(\mathbb{R}^d)$  (smooth function with compact support) is s.t.  $\varphi(0) \neq 0$ ,  $\varphi(D)f := (\varphi \hat{f})^\vee$  where  $\hat{f}$  and  $(\varphi \hat{f})^\vee$  respectively denote the Fourier transform of  $f$  and the inverse Fourier transform of  $\varphi \hat{f}$ . The parameter  $m$  is an integer s.t.  $m > \frac{\tilde{\alpha}}{2}$  and for  $v > 0, z \in \mathbb{R}^d$ ,  $h_v(z) := \frac{1}{(2\pi v)^{\frac{d}{2}}} \exp\left(-\frac{|z|^2}{2v}\right)$  is the usual heat kernel of  $\mathbb{R}^d$ . We point out that the quantities in (2.18) are well defined for  $q < \infty$ . The modifications for  $q = +\infty$  are obvious and can be written passing to the limit.

Observe that the quantity  $\|f\|_{\mathcal{H}_{p,q}^{\tilde{\alpha}}}$ , where the subscript  $\mathcal{H}$  stands to indicate the dependence on the heat-kernel, depends on the considered function  $\varphi$  and the chosen  $m \in \mathbb{N}$ . It also defines a quasi-norm on  $B_{p,q}^s(\mathbb{R}^d)$ . The previous definition of  $B_{p,q}^{\tilde{\alpha}}(\mathbb{R}^d)$  is known as the thermic characterization of Besov spaces and is particularly well adapted to our current framework. By abuse of notation we will write as soon as this quantity is finite  $\|f\|_{\mathcal{H}_{p,q}^{\tilde{\alpha}}} =: \|f\|_{B_{p,q}^{\tilde{\alpha}}}$ .

As indicated above, it is easily seen from (2.18) that  $C_b^{\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R}) = B_{\infty,\infty}^{\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R})$ . It is also well known that  $B_{\infty,\infty}^{\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R})$  and  $B_{1,1}^{-\tilde{\alpha}_i}(\mathbb{R}^d, \mathbb{R})$  are in duality (see e.g. Proposition 3.6

in [LR02]). Namely  $B_{\infty,\infty}^{\tilde{\alpha}_i}$  is the dual of the closure of the Schwartz class  $\mathcal{S}$  in  $B_{1,1}^{-\tilde{\alpha}_i}$ . But  $\mathcal{S}$  is dense in  $B_{1,1}^{-\tilde{\alpha}_i}$  (see for instance Theorem 4.1.3 in [AH96]). We will therefore write from (2.16) and with the notations of (2.17)

$$\begin{aligned} & \sum_{i=2}^n \int_t^T ds \left| \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{y}_i} \cdot (\Theta_{i,(t,\mathbf{x})}^{m,\alpha}(s, \mathbf{y})) u_m(s, \mathbf{y}) \right|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\ & \leq \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \|\Psi_m^{i,(t,\mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}\|_{B_{1,1}^{-\tilde{\alpha}_i}} \|u_m^{i,(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}\|_{B_{\infty,\infty}^{\tilde{\alpha}_i}} \\ & \leq \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \|\Psi_m^{i,(t,\mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}\|_{B_{1,1}^{-\tilde{\alpha}_i}}. \end{aligned} \quad (2.19)$$

Exploiting the thermic characterization of Besov spaces (see again (2.18) and Section 4.2), It will be shown in Lemma 5 below that there exists  $C := C((\mathbf{A}))$  s.t. for all  $i \in \llbracket 2, n \rrbracket$  and all  $m \in \mathbb{N}$ :

$$\int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \|\Psi_m^{i,(t,\mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}\|_{B_{1,1}^{-\tilde{\alpha}_i}} \leq \frac{C}{(s-t)^{1-\frac{\gamma}{2}}}. \quad (2.20)$$

Therefore:

$$\sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \|\Psi_m^{i,(t,\mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}\|_{B_{1,1}^{-\tilde{\alpha}_i}} \leq C(T-t)^{\frac{\gamma}{2}}, \quad (2.21)$$

which plugged into (2.19) eventually gives the following global smoothing effect for the degenerate contributions. Namely,

$$\left| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}_m}(\tau, s, \boldsymbol{\theta}_{s,t}^m(\boldsymbol{\xi}), \mathbf{y}) D_{\mathbf{y}_i} u_m(s, \mathbf{y}) \rangle \right|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \leq C(T-t)^{\frac{\gamma}{2}}, \quad (2.22)$$

which is precisely homogeneous to the bound obtained for the non-degenerate variables in (2.14). In both cases, the contribution  $(T-t)^{\frac{\gamma}{2}}$  derives from the assumed smoothness of the coefficients  $a, \mathbf{F}$  w.r.t.  $\mathbf{d}$  which exactly leads to the same global control for the *a priori* most singular part of expansion (2.12).

From the previous bounds and (2.12) we thus obtain:

$$|D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x})| \leq \left( |D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{m,(\tau,\boldsymbol{\xi})} g_m(\mathbf{x})| + |D_{\mathbf{x}_1}^2 \tilde{G}^{m,(\tau,\boldsymbol{\xi})} f_m(t, \mathbf{x})| \right) \Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} + C(T-t)^{\frac{\gamma}{2}} \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}. \quad (2.23)$$

Since  $\tilde{P}_{T,t}^{m,(\tau,\boldsymbol{\xi})}$  is a *true* semi-group, and  $\tilde{G}^{m,(\tau,\boldsymbol{\xi})}$  the associated Green kernel (precisely because we used a *forward* perturbative expansion), it will be derived in Lemma 6 (thanks to cancellation techniques) that there exists  $C := C((\mathbf{A}))$  s.t. for all  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

$$\begin{aligned} & \left( |D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{m,(\tau,\boldsymbol{\xi})} g_m(\mathbf{x})| + |D_{\mathbf{x}_1}^2 \tilde{G}^{m,(\tau,\boldsymbol{\xi})} f_m(t, \mathbf{x})| \right) \Big|_{(\tau,\boldsymbol{\xi})=(t,\mathbf{x})} \\ & \leq C(\|g_m\|_{C_{b,d}^{2+\gamma}} + (T-t)^{\frac{\gamma}{2}} \|f_m\|_{L^\infty(C_{b,d}^\gamma)}) \\ & \leq C(\|g\|_{C_{b,d}^{2+\gamma}} + (T-t)^{\frac{\gamma}{2}} \|f\|_{L^\infty(C_{b,d}^\gamma)}). \end{aligned} \quad (2.24)$$

Equation (2.24) eventually leads to the following estimate on  $|D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x})|$ :

$$|D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x})| \leq C \left( \|g\|_{C_{b,d}^{2+\gamma}} + (T-t)^{\frac{\gamma}{2}} (\|f\|_{L^\infty(C_{b,d}^\gamma)} + \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}) \right). \quad (2.25)$$

For  $T$  small enough, this equation would be compatible with the estimates of Theorem 1. Equation (2.25) might even seem *too strong* since it also exhibits, additionally to the control of the term associated with the perturbation, a small contribution (in  $(T-t)^{\frac{\gamma}{2}}$  for a small enough  $T$ ) w.r.t. to the source  $f_m$ . This is precisely because  $\|D_{\mathbf{x}_1}^2 u_m(t, \cdot)\|_{L^\infty}$  is not one of the *critical* terms in the Hölder norm  $\|u_m(t, \cdot)\|_{C_{b,d}^{2+\gamma}}$ , i.e. the regularity of the coefficients still gives that it can be viewed as a remainder at first sight. Now, a typical critical term of the Hölder norm, for which we precisely exploit totally the spatial regularity of the coefficients, is  $[D_{\mathbf{x}_1}^2 u_m(t, \cdot)]_{\mathbf{d}}^\gamma$  (see assumption **(S)** and (1.15)). Let us now detail how we can handle it and in which sense it can be viewed as critical.

Of course, if  $g = 0$ , for  $t \in [0, T]$  and given spatial points  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$  we can assume w.l.o.g. that  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq (T-t)^{\frac{1}{2}}$ , i.e. the spatial points are *close* w.r.t. the characteristic time scale  $(T-t)^{\frac{1}{2}}$  for the homogeneous metric  $\mathbf{d}$ . Indeed, if  $\mathbf{d}(\mathbf{x}, \mathbf{x}') > (T-t)^{\frac{1}{2}}$ , equation (2.25) readily gives:

$$\begin{aligned} |D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}')| &\leq |D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x})| + |D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}')| \\ &\leq 2C(T-t)^{\frac{\gamma}{2}} (\|f\|_{L^\infty(C_{b,d}^\gamma)} + \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}) \\ &\leq 2C\mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') (\|f\|_{L^\infty(C_{b,d}^\gamma)} + \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}). \end{aligned} \quad (2.26)$$

Let us now focus, as above, on the Hölder control associated with the perturbative contribution in (2.12) when  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq (T-t)^{\frac{1}{2}}$ . Namely,

$$\begin{aligned} D_{\mathbf{x}_1}^2 \Delta_m^{\tau, \xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') &:= \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{m, (\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) (L_s^m - \tilde{L}_s^{m, (\tau, \xi)}) u_m(s, \mathbf{y}) d\mathbf{y} \\ &\quad - \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{m, (\tau, \xi')}(t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m, (\tau, \xi')}) u_m(s, \mathbf{y}) d\mathbf{y}, \end{aligned} \quad (2.27)$$

where we recall that *a priori* the spatial freezing points  $(\xi, \xi')$  in (2.27) (see also (2.12)) should be different for  $\mathbf{x}$  and  $\mathbf{x}'$  and depend on the position of  $\mathbf{d}(\mathbf{x}, \mathbf{x}')$  w.r.t. the current characteristic time scale in the time integral. Following the terminology of heat kernels, we will say that at time  $s \in [t, T]$  the points  $\mathbf{x}, \mathbf{x}'$  are in the *diagonal regime* if  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \leq s-t$ , i.e. their homogeneous distance is small w.r.t. the characteristic time for a parameter  $c_0$  to be specified later on.

We insist again that we have the usual equivalence between time and space, i.e. time has to be compared with the square of the spatial metric. Similarly, we will say that the *off-diagonal* regime holds when  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') > (s-t)$ . Observing that in the diagonal case  $s \geq t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  (and in the off-diagonal one  $s < t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ ) we split the time integral in (2.27) as:

$$\Delta_m^{\tau, \xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') := \Delta_{m, \text{diag}}^{\tau, \xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') + \Delta_{m, \text{off-diag}}^{\tau, \xi, \xi'}(t, \mathbf{x}, \mathbf{x}'),$$

with

$$\begin{aligned}
 \Delta_{m,\text{off-diag}}^{\tau,\xi,\xi'}(t, \mathbf{x}, \mathbf{x}') &:= \int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\xi)})u_m(s, \mathbf{y})d\mathbf{y}, \\
 &\quad - \int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\xi')})u_m(s, \mathbf{y})d\mathbf{y}, \\
 \Delta_{m,\text{diag}}^{\tau,\xi,\xi'}(t, T, \mathbf{x}, \mathbf{x}') &:= \int_{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\xi)})u_m(s, \mathbf{y})d\mathbf{y} \\
 &\quad - \int_{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\xi')})u_m(s, \mathbf{y})d\mathbf{y}.
 \end{aligned} \tag{2.28}$$

Intuitively, for the term  $D_{\mathbf{x}_1}^2 \Delta_{m,\text{off-diag}}^{\tau,\xi,\xi'}(t, \mathbf{x}, \mathbf{x}')$ , since  $\mathbf{x}, \mathbf{x}'$  are *far* at the characteristic time scale  $(s-t)^{\frac{1}{2}}$ , there is no expectable gain in expanding  $D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y})$ . One therefore writes:

$$\begin{aligned}
 &|D_{\mathbf{x}_1}^2 \Delta_{m,\text{off-diag}}^{\tau,\xi,\xi'}(t, \mathbf{x}, \mathbf{x}')| \\
 &\leq \left| \int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\xi)}(t, s, \mathbf{x}, \mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\xi)})u_m(s, \mathbf{y})d\mathbf{y} \right. \\
 &\quad \left. + \left| \int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y})(L_s^m - \tilde{L}_s^{m,(\tau,\xi')})u_m(s, \mathbf{y})d\mathbf{y} \right| \right|.
 \end{aligned}$$

Now, provided  $\xi = \mathbf{x}, \xi' = \mathbf{x}'$  one derives from the previous equation, similarly to (2.19), (2.20), that

$$\begin{aligned}
 &|D_{\mathbf{x}_1}^2 \Delta_{m,\text{off-diag}}^{\tau,\xi,\xi'}(t, \mathbf{x}, \mathbf{x}')| \\
 &\leq C \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \int_t^{t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} \frac{ds}{(s-t)^{1-\frac{\gamma}{2}}} \leq C \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} c_0^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{2.29}
 \end{aligned}$$

For  $c_0$  small enough, we obtain again an estimate that would be compatible with the global bound on  $\|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})}$  stated in Theorem 1.

Turning now to  $D_{\mathbf{x}_1}^2 \Delta_{m,\text{diag}}^{\tau,\xi,\xi'}(t, T, \mathbf{x}, \mathbf{x}')$  one would therefore be tempted to carry on the analysis with the previous freezing points  $\xi = \mathbf{x}, \xi' = \mathbf{x}'$ . Intuitively, in the diagonal regime this should not have too much impact. This is only partly true, since if we proceed so we will be led to investigate the difference of operators at different freezing spatial points and this leads to compare quantities like  $\theta_{s,t}^m(\mathbf{x}) - \theta_{s,t}^m(\mathbf{x}')$  for which we want a uniform control w.r.t.  $m$ . Since the initial (unmollified) coefficients  $a, \mathbf{F}$  are only Hölder continuous in space, this quantity is typically controlled (see Lemma 3) as:

$$\mathbf{d}(\theta_{s,t}^m(\mathbf{x}), \theta_{s,t}^m(\mathbf{x}')) \leq C(\mathbf{d}(\mathbf{x}, \mathbf{x}') + (s-t)^{\frac{1}{2}}), \tag{2.30}$$

where the time contribution precisely reflects the *roughness* of the coefficients.

Unfortunately, this approach would lead to a final control of order  $(\mathbf{d}(\mathbf{x}, \mathbf{x}') + (s-t)^{\frac{1}{2}})^\gamma \leq C(\mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') + (s-t)^{\frac{\gamma}{2}})$  which is not enough on the considered integration set.

Recall indeed that, in the diagonal regime  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \leq (s-t)$  and the term  $(s-t)^{\frac{\gamma}{2}}$  in the previous r.h.s. is *too big*. This means that for  $\Delta_{m, \text{diag}}^{\tau, \xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}')$ , it would be more appropriate to consider the same spatial freezing point. In that case, taking  $\xi = \xi' = \mathbf{x}$  and expanding the difference of the derivatives of the frozen Gaussian densities yields:

$$\begin{aligned}
 & \Delta_{m, \text{diag}}^{\tau, \xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') \\
 = & - \sum_{j=1}^n \int_{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu D_{\mathbf{x}_j} D_{\mathbf{x}_1}^2 \tilde{p}^{m, (\tau, \xi)}(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x}' - \mathbf{x})_j \\
 & (L_s^m - \tilde{L}_s^{m, (\tau, \xi)}) u_m(s, \mathbf{y}) \\
 = & - \sum_{j=1}^n \int_{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu D_{\mathbf{x}_j} D_{\mathbf{x}_1}^2 \tilde{p}^{m, (\tau, \xi)}(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x}' - \mathbf{x})_j \\
 & \left( \Delta_{1, \mathbf{F}_{m, \sigma_m}}(t, s, \boldsymbol{\theta}_{s,t}^m(\mathbf{x}), \mathbf{y}, u_m) + \sum_{i=2}^n \langle \Delta_{i, \mathbf{F}_{m, \sigma_m}}(t, s, \boldsymbol{\theta}_{s,t}^m(\mathbf{x}), \mathbf{y}) D_{\mathbf{y}_i} u_m(s, \mathbf{y}) \rangle \right), \tag{2.31}
 \end{aligned}$$

using the notations introduced in (2.12) and (2.15) for the last equality. In the previous identities (2.31) and from now on, the symbol “ $\cdot$ ” between two tensors means the usual tensor contraction. In particular  $D_{\mathbf{x}_j} D_{\mathbf{x}_1}^2 \tilde{p}^{m, (\tau, \xi)}(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x}' - \mathbf{x})_j$  is a  $d \times d$  matrix.

In the current diagonal regime, it can be shown from (2.8) and the homogeneity of the distance  $\mathbf{d}$  that there is  $C > 1$  such that for  $\xi = \mathbf{x}$ :

$$\begin{aligned}
 & |D_{\mathbf{x}_j} D_{\mathbf{x}_1}^2 \tilde{p}^{m, (\tau, \xi)}(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y})| \\
 \leq & \frac{C}{(s-t)^{j-\frac{1}{2}+1+\frac{n^2 d}{2}}} \exp(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{m, (t, \xi)}(\mathbf{x} + \mu(\mathbf{x}' - \mathbf{x})) - \mathbf{y})|^2) \\
 \leq & \frac{C}{(s-t)^{j-\frac{1}{2}+1+\frac{n^2 d}{2}}} \exp(C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{m, (t, \xi)}(\mathbf{x} - \mathbf{x}'))|^2) \\
 & \times \exp\left(-\frac{C^{-1}}{2}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{m, (t, \xi)}(\mathbf{x})) - \mathbf{y}|^2\right) \\
 \leq & \frac{C}{(s-t)^{j-\frac{1}{2}+1+\frac{n^2 d}{2}}} \exp\left(-\frac{C^{-1}}{2}(s-t) |\mathbb{T}_{s-t}^{-1}(\boldsymbol{\theta}_{s,t}^m(\mathbf{x})) - \mathbf{y}|^2\right), \tag{2.32}
 \end{aligned}$$

using for the last inequality that  $\mathbf{m}_{s,t}^{m, (t, \xi)}(\mathbf{x})|_{\xi=\mathbf{x}} = \boldsymbol{\theta}_{s,t}^m(\mathbf{x})$  and the fact that, from the linear structure of ODE satisfied by  $\mathbf{m}_{s,t}^{m, (t, \xi)}(\mathbf{x})$  (which can be read from system (2.6) taking the expectation),  $(s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{m, (t, \xi)}(\mathbf{x} - \mathbf{x}'))|^2 \leq C(s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \mathbf{x}')|$ . Since  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \leq s-t \iff c_0 \mathbf{d}^2((s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \mathbf{x}, (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \mathbf{x}') \leq 1$ , we readily derive from the definition of  $\mathbf{d}$  in (1.9) that  $(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{m, (t, \xi)}(\mathbf{x} - \mathbf{x}'))|^2 \leq C$ . These points are thoroughly discussed in Section 3.

From (2.31), (2.32) reproducing the previously described analysis, we finally derive:

$$\begin{aligned}
 & D_{\mathbf{x}_1}^2 \Delta_{m, \mathbf{diag}}^{\tau, \xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') \\
 \leq & C \|u_m\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})} \sum_{j=1}^n \int_{t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')}^T \frac{ds}{(s-t)^{1+(j-\frac{1}{2})-\frac{\gamma}{2}}} |(\mathbf{x} - \mathbf{x}')_j| \\
 \leq & C \|u_m\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})} \sum_{j=1}^n \frac{|(\mathbf{x} - \mathbf{x}')_j|}{(c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}'))^{j-\frac{1}{2}-\frac{\gamma}{2}}} \leq \frac{C}{c_0^{n-\frac{1}{2}-\frac{\gamma}{2}}} \|u_m\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'), \quad (2.33)
 \end{aligned}$$

using again the definition of  $\mathbf{d}$  in (1.9) for the last inequality. We have again globally gained, thanks to the smoothness of the coefficients, a power  $\frac{\gamma}{2}$  in the time singularities of equation (2.32).

From the previous discussion we now have to specify how to modify the freezing parameter depending on the position of the current time variable w.r.t. to the homogeneous spatial distance between the considered points. This can actually be done from the Duhamel formulation up to an additional discontinuity term. Restarting from (2.11) we can indeed rewrite for given  $(t, \mathbf{x}') \in [0, T] \times \mathbb{R}^{nd}$  and any  $r \in (t, T]$ ,  $\xi' \in \mathbb{R}^{nd}$ :

$$\begin{aligned}
 u_m(t, \mathbf{x}') &= \tilde{P}_{r,t}^{m,(\tau, \xi')} u_m(r, \mathbf{x}') + \tilde{G}_{r,t}^{m,(\tau, \xi')} f_m(t, \mathbf{x}') \\
 &\quad + \int_t^r ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \xi')}(t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau, \xi')}) u_m(s, \mathbf{y}), \\
 \forall 0 \leq v < r \leq T, \quad \tilde{G}_{r,v}^{m,(\tau, \xi')} f_m(t, \mathbf{x}') &= \int_v^r ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \xi')}(t, s, \mathbf{x}', \mathbf{y}) f_m(s, \mathbf{y}). \quad (2.34)
 \end{aligned}$$

Differentiating the above expression in  $r \in (t, T]$  yields for any  $\xi' \in \mathbb{R}^{nd}$ :

$$\begin{aligned}
 0 &= \partial_r \tilde{P}_{r,t}^{m,(\tau, \xi')} u_m(r, \mathbf{x}') + \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \xi')}(t, r, \mathbf{x}', \mathbf{y}) f_m(r, \mathbf{y}) \\
 &\quad + \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \xi')}(t, r, \mathbf{x}', \mathbf{y}) (L_r^m - \tilde{L}_r^{m,(\tau, \xi')}) u_m(r, \mathbf{y}) d\mathbf{y}. \quad (2.35)
 \end{aligned}$$

Integrating (2.35) between  $t$  and  $t_0 \in (t, T]$  for a first given  $\xi'$  and between  $t_0$  and  $T$  with a possibly different  $\tilde{\xi}'$  yields:

$$\begin{aligned}
 0 &= \tilde{P}_{t_0,t}^{m,(\tau, \xi')} u_m(t_0, \mathbf{x}') - u_m(t, \mathbf{x}') + \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \xi')}(t, s, \mathbf{x}', \mathbf{y}) f_m(s, \mathbf{y}) \\
 &\quad + \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \xi')}(t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau, \xi')}) u_m(s, \mathbf{y}) \\
 &\quad + \tilde{P}_{T,t}^{m,(\tau, \tilde{\xi}')} u_m(T, \mathbf{x}') - \tilde{P}_{t_0,t}^{m,(\tau, \tilde{\xi}')} u_m(t_0, \mathbf{x}') + \int_{t_0}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \tilde{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) f_m(s, \mathbf{y}) \\
 &\quad + \int_{t_0}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{m,(\tau, \tilde{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau, \tilde{\xi}')} u_m(s, \mathbf{y}).
 \end{aligned}$$

Recalling that  $u_m(T, \mathbf{x}') = g_m(\mathbf{x}')$  (terminal condition), and with the notations of (2.34) the above equation rewrites:

$$\begin{aligned}
 u_m(t, \mathbf{x}') &= \tilde{P}_{T,t}^{m,(\tau,\tilde{\xi}')} g_m(\mathbf{x}') + \tilde{G}_{t_0,t}^{m,(\tau,\xi')} f_m(t, \mathbf{x}') + \tilde{G}_{T,t_0}^{m,(\tau,\tilde{\xi}')} f_m(t, \mathbf{x}') \\
 &\quad + \tilde{P}_{t_0,t}^{m,(\tau,\xi')} u_m(t_0, \mathbf{x}') - \tilde{P}_{t_0,t}^{m,(\tau,\tilde{\xi}')} u_m(t_0, \mathbf{x}') \\
 &\quad + \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left( \mathbb{I}_{s \leq t_0} \tilde{p}^{m,(\tau,\xi')}(t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\xi')}) \right. \\
 &\quad \left. + \mathbb{I}_{s > t_0} \tilde{p}^{m,(\tau,\tilde{\xi}')} (t, s, \mathbf{x}', \mathbf{y}) (L_s^m - \tilde{L}_s^{m,(\tau,\tilde{\xi}')} ) \right) u_m(s, \mathbf{y}). \tag{2.36}
 \end{aligned}$$

We see that for  $\xi' \neq \tilde{\xi}'$  we have an additional discontinuity term deriving from the change of freezing point along the time variable. Of course expression (2.36) can be differentiated in space and taking **then**

$$t_0 = t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}'), \tag{2.37}$$

i.e.  $t_0$  precisely corresponds to the critical time at which a change of regime occurs, and  $\xi' = \mathbf{x}'$ ,  $\tilde{\xi}' = \mathbf{x}$  precisely allows, when expanding  $D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}')$  using (2.12) for the first term and (2.36) for the second one, to exploit the previous analysis that led to (2.29) and (2.33) and which precisely relied on the suitable choice of freezing point. We again insist on the fact that, in the analysis,  $t_0$  is an additional freezing parameter, which is *a posteriori* chosen according to (2.37) as a function of  $(t, \mathbf{x}, \mathbf{x}')$ . In particular the parameter  $t_0$  does not intervene in the various possible differentiations of the considered perturbative expansions.

This approach eventually leads to:

$$\begin{aligned}
 &|D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}')| \\
 &\leq \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \left[ C_1 (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)}) + C (c_0^{-n+\frac{1}{2}+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}}) \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \right] \\
 &\quad + \left| (D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{m,(\tau,\xi')} u_m(t_0, \mathbf{x}') - D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{m,(\tau,\tilde{\xi}')} u_m(t_0, \mathbf{x}')) \Big|_{t_0=t+c_0\mathbf{d}^2(\mathbf{x},\mathbf{x}')} \right|. \tag{2.38}
 \end{aligned}$$

The last contribution can be controlled through cancellation techniques and the key estimate (2.30) on the difference of the flows. The specific choice of  $t_0 = t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  then precisely provides the required order leading to:

$$\begin{aligned}
 &|D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u_m(t, \mathbf{x}')| \\
 &\leq \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \left[ C_1 (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)}) + C (c_0^{-n+\frac{1}{2}+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}} + c_0) \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \right].
 \end{aligned}$$

We refer to Lemma 10 for results associated with the discontinuity term in (2.38).

We have detailed up to now what happens with the second order derivatives w.r.t. the non-degenerate variables. The previous procedure can be applied as well to control the Hölder moduli w.r.t. the degenerate ones. We therefore end up with the following kind of estimate:

$$\|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq C_1 (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)}) + \|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \left[ C (c_0^{-n+\frac{1}{2}+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}}) + T^{\frac{\gamma}{2}} \right]. \tag{2.39}$$



Equation (2.39) would provide exactly the expected control if  $T$  and  $C$  are small enough. On the one hand, the final parameter  $T$  can always be chosen small enough. Indeed, the further extension of the Schauder estimates to an arbitrary fixed final time horizon can then be obtained by iterative applications (chaining) of the main estimates on a sufficiently small time interval. On the other hand, it will appear from the proofs that the constant  $C$  in (2.39) actually depends on the Hölder norms of the considered coefficients. If these quantities are small, i.e. the coefficients do not vary much and the components that transmit the noise are almost linear, then  $C$  will be small. For  $\left[ C(c_0^{-n+\frac{1}{2}+\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2}}) + T^{\frac{\gamma}{2}} \right] \leq k_0 < 1$ , we eventually derive:

$$\|u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq \frac{C_1}{1-k_0} (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)}),$$

which is precisely the expected control. The general case, is proved through a scaling argument which also allows to balance the opposite effects of  $c_0$  (meant to be small) in the above bounds. This last point will be discussed in Section 6.

The remaining part of this article is organized as follows. We prove in Section 3 various properties for the density of the linearized Gaussian proxy: precise pointwise estimates for the density itself and its derivatives (see equation (2.8)) and some useful controls allowing cancellation arguments in our perturbative analysis. We also give therein some properties associated with the rescaled proxy. Section 4 is then devoted to the control of the supremum norms of the non-degenerate derivatives, corresponding to the previous equation (2.25). Section 5 addresses the issues of Hölder controls. Section 6 is concerned with the above mentioned scaling issues and we also conclude there the final proof of Theorem 1. Eventually, some auxiliary, but crucial, technical results are proved in Appendix 7 for the regularity of the flow and the mean, in Appendix 8 for the regularity of the resolvent and the covariance, and in Appendix 9 for a scaling analysis.

### 3 Gaussian proxy and associated controls

We first aim here at proving the control of equation (2.8). We recall that our point is to control the density of  $(\tilde{\mathbf{X}}_s^{m,(\tau,\boldsymbol{\xi})})_{s \in (t,T]}$  satisfying (2.6).

**WARNING:** for notational simplicity, for the rest of the document we drop the sub and superscripts in  $m$  associated with the regularizations. We rewrite, with some notational abuse, for fixed  $(\tau, \boldsymbol{\xi}) \in [0, T] \times \mathbb{R}^{nd}$ , the dynamics in (2.6) as:

$$\begin{aligned} d\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})} &= [\mathbf{F}(v, \boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi})) + D\mathbf{F}(v, \boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))(\tilde{\mathbf{X}}_v^{(\tau,\boldsymbol{\xi})} - \boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))]dv + B\sigma(v, \boldsymbol{\theta}_{v,\tau}(\boldsymbol{\xi}))dW_v, \\ \forall v \in [t, s], \tilde{\mathbf{X}}_t^{(\tau,\boldsymbol{\xi})} &= \mathbf{x}, \end{aligned} \tag{3.1}$$

keeping in mind that  $\mathbf{F}, \boldsymbol{\theta}, \sigma$  in (3.1) are *smooth* coefficients. We will give in the next subsection some key-controls to investigate the terms appearing the perturbative expansions (2.11) and (2.12).

### 3.1 Controls for the frozen density

We explicitly integrate (3.1) to obtain for all  $v \in [t, s]$ :

$$\begin{aligned}\tilde{\mathbf{X}}_v^{(\tau, \xi)} &= \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t)\mathbf{x} + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) \left( \mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\xi)) - D\mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\xi))\boldsymbol{\theta}_{u, \tau}(\xi) \right) du \\ &\quad + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\xi)) dW_u \\ &=: \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\xi)) dW_u,\end{aligned}$$

where  $(\tilde{\mathbf{R}}^{(\tau, \xi)}(v, u))_{t \leq u, v \leq s}$  stands for the resolvent associated with the collection of partial gradients in  $(D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\xi)))_{v \in [t, s]}$ , introduced in (2.7), which satisfies for  $v \in [t, s]$ :

$$\partial_v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t) = D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\xi)) \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t), \quad \tilde{\mathbf{R}}^{(\tau, \xi)}(t, t) = \mathbf{I}_{nd, nd}. \quad (3.2)$$

Note in particular that since the partial gradients are subdiagonal  $\det(\tilde{\mathbf{R}}^{(\tau, \xi)}(v, t)) = 1$ .

Also, for  $v \in [t, s]$ , we recall that  $\mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x})$  stands for the mean of  $\tilde{\mathbf{X}}_v^{(\tau, \xi)}$  and corresponds as well to the solution of (3.1) when  $\sigma = 0$  and the starting point is  $\mathbf{x}$ . We write:

$$\tilde{\mathbf{X}}_v^{(\tau, \xi)} = \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\xi)) dW_u, \quad v \in [t, s]. \quad (3.3)$$

Importantly, we point out that  $\mathbf{x} \in \mathbb{R}^{nd} \mapsto \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x})$  is **affine** w.r.t. the starting point  $\mathbf{x}$ . Precisely, for  $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{nd}$ :

$$\mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x} + \mathbf{x}') = \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t)\mathbf{x}' + \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x}). \quad (3.4)$$

We first give in the next proposition a key estimate on the covariance matrix associated with (3.3) and its properties w.r.t. a suitable scaling of the system.

**Proposition 1 (Good Scaling Properties of the Covariance Matrix)** *The covariance matrix of  $\tilde{\mathbf{X}}_v^{(\tau, \xi)}$  in (3.3) writes:*

$$\tilde{\mathbf{K}}_{v, t}^{(\tau, \xi)} := \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B a(u, \boldsymbol{\theta}_{u, \tau}(\xi)) B^* \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u)^* du.$$

*Uniformly in  $(\tau, \xi) \in [0, T] \times \mathbb{R}^{nd}$  and  $s \in [0, T]$ , it satisfies a good scaling property in the sense of Definition 3.2 in [DM10] (see also Proposition 3.4 of that reference). That is, for all fixed  $T > 0$ , there exists  $C_{(3.5)} := C_{(3.5)}((\mathbf{A}), T) \geq 1$  s.t. for all  $0 \leq t < v \leq s \leq T$ ,  $(\tau, \xi) \in [0, T] \times \mathbb{R}^{nd}$ :*

$$\forall \zeta \in \mathbb{R}^{nd}, \quad C_{(3.5)}^{-1}(v-t)^{-1} |\mathbb{T}_{v-t} \zeta|^2 \leq \langle \tilde{\mathbf{K}}_{v, t}^{(\tau, \xi)} \zeta, \zeta \rangle \leq C_{(3.5)}(v-t)^{-1} |\mathbb{T}_{v-t} \zeta|^2, \quad (3.5)$$

where for all  $u > 0$ , we denote by  $\mathbb{T}_u$  the intrinsic scale matrix introduced in (2.9). Namely:

$$\mathbb{T}_u = \begin{pmatrix} u\mathbf{I}_{d, d} & \mathbf{0}_{d, d} & \cdots & \mathbf{0}_{d, d} \\ \mathbf{0}_{d, d} & u^2\mathbf{I}_{d, d} & \mathbf{0}_{d, d} & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d, d} & \cdots & \mathbf{0}_{d, d} & u^n\mathbf{I}_{d, d} \end{pmatrix}.$$

The proof of the above proposition readily follows from Proposition 3.3 and Lemma 3.6 in [DM10]. We now state some important density bounds for the linearized model.

**Proposition 2 (Density of the linearized dynamics)** *Under (A), we have that, for all  $s \in (t, T]$  the random variable  $\tilde{\mathbf{X}}_s^{(\tau, \xi)}$  in (3.3) admits a Gaussian density  $\tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \cdot)$  which writes for all  $\mathbf{y} \in \mathbb{R}^{nd}$ :*

$$\tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) := \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y}), \mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y} \right\rangle\right), \quad (3.6)$$

with  $\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)}$  as in Proposition 1. Also, there exists  $C := C((\mathbf{A}), T) > 0$  s.t. for all multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ ,  $|\alpha| \leq 3$  and denoting by  $D_{\mathbf{x}}^\alpha := D_{\mathbf{x}_1}^{\alpha_1} \dots D_{\mathbf{x}_n}^{\alpha_n}$ , we have:

$$\begin{aligned} |D_{\mathbf{x}}^\alpha \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| &\leq \frac{C}{(s-t)^{\sum_{i=1}^n \alpha_i (i-\frac{1}{2}) + \frac{n^2 d}{2}}} \exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})|^2\right) \\ &=: \frac{C}{(s-t)^{\sum_{i=1}^n \alpha_i (i-\frac{1}{2})}} \bar{p}_{C^{-1}}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}), \end{aligned} \quad (3.7)$$

with  $\int_{\mathbb{R}^{nd}} d\mathbf{y} \bar{p}_{C^{-1}}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})$ , up to a modification of the constants in (3.7).

**Remark 3 (A slight abuse of notation)** *To ease the reading we denote, when there is no possible ambiguity,  $\bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) := \bar{p}_{C^{-1}}^{(t, \mathbf{x})}(t, s, \mathbf{x}, \mathbf{y})$ .*

**Remark 4 (Regularizing effect of the quasi-distance)** *From equation (3.7), we derive from the definition of  $\mathbf{d}$  in (1.9) that for any given  $\beta > 0$ , there exists  $C_\beta$  s.t.*

$$\begin{aligned} &\mathbf{d}\left((s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}), (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \mathbf{y}\right) |D_{\mathbf{x}}^\alpha \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| \\ &\leq \frac{C_\beta}{(s-t)^{\sum_{i=1}^n \alpha_i (i-\frac{1}{2}) - \frac{\beta}{2}}} \bar{p}_{C_\beta^{-1}}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}), \end{aligned} \quad (3.8)$$

i.e. equation (3.8) quantifies the regularizing effect of the scaled arguments in the quasi-distance.

**Proof:** Expression (3.6) readily follows from (3.2) and (3.3). Differentiating w.r.t.  $\mathbf{x}$  recalling from (3.4) that  $\mathbf{x} \mapsto \mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x})$  is affine yields:

$$D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) = -\left[ [\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* (\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y}) \right]_j \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \quad (3.9)$$

The point is now to use scaling arguments. We can first rewrite

$$[\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* (\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{-1} = (s-t) [\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* \mathbb{T}_{s-t}^{-1} (\widehat{\mathbf{K}}_1^{s,t})^{-1} \mathbb{T}_{s-t}^{-1}, \quad (3.10)$$

where  $\widehat{\mathbf{K}}_1^{s,t}$  is the covariance matrix of the rescaled process  $\left( (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \tilde{\mathbf{X}}_{t+v(s-t)}^{t, \mathbf{x}} \right)_{v \in [0,1]}$  at time 1. From the good-scaling property of Proposition 1, it is plain to derive that  $\widehat{\mathbf{K}}_1^{s,t}$

a non-degenerate bounded matrix, i.e. there exists  $\hat{C} \geq 1$  s.t. for all  $\zeta \in \mathbb{R}^{nd}$ ,  $(\hat{C})^{-1}|\zeta|^2 \leq \langle \widehat{\mathbf{K}}_1^{s,t} \zeta, \zeta \rangle \leq \hat{C}|\zeta|^2$ . A similar rescaling argument yields on the deterministic system (3.2) of the resolvent yields that  $\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)$  can also be written as:

$$[\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* = \mathbb{T}_{s-t}^{-1} \left[ \widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0) \right]^* \mathbb{T}_{s-t}, \quad (3.11)$$

where again  $\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0)$  is the resolvent at time 1 of the rescaled system

$$\left( \mathbb{T}_{s-t} [\tilde{\mathbf{R}}^{(\tau, \xi)}(t + v(s-t), t)]^* \mathbb{T}_{s-t}^{-1} \right)_{v \in [0,1]} = \left( [\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(v, 0)]^* \right)_{v \in [0,1]}$$

associated with (3.2). From the analysis performed in Lemma 5.1 in [HM16] (see also the proof of Proposition 3.7 in [DM10]) one derives that there exists  $\hat{C}_1$  s.t. for all  $\zeta \in \mathbb{R}^{nd}$ ,

$$|[\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0)]^* \zeta| \leq \hat{C}_1 |\zeta|. \quad (3.12)$$

Equations (3.9), (3.10) and (3.11) therefore yield:

$$\begin{aligned} & |D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| \\ & \leq (s-t)^{-j+\frac{1}{2}} \left| \left( [\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0)]^* (\widehat{\mathbf{K}}_1^{s,t})^{-1} ((s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})) \right)_j \right| \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) \\ & \leq C (s-t)^{-j+\frac{1}{2}} (s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})| \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned}$$

From the explicit expression (3.6), Proposition 1 and the above equation, we eventually derive:

$$\begin{aligned} & |D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| \\ & \leq \frac{C}{(s-t)^{j-\frac{1}{2}}} \left( (s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})| \right) \\ & \quad \times \frac{1}{(s-t)^{\frac{n2d}{2}}} \exp \left( -C^{-1} (s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})|^2 \right) \\ & \leq \frac{C}{(s-t)^{j-\frac{1}{2}}} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}), \end{aligned}$$

up to a modification of  $C$ , which gives the statement for one partial derivative. The controls on the higher order derivatives are obtained similarly (see e.g. the proof of Lemma 5.5 of [DM10] for the bounds on  $D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})$ ).  $\square$

We now give some useful controls involving the previous Gaussian kernel which will be used in our perturbative analysis. The main interest of the estimates below is that they precisely allow to exploit cancellation techniques.

**Proposition 3** For all  $0 \leq t \leq s \leq T$ ,  $(\mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$ , the following identities hold:

$$\int_{\mathbb{R}^{nd}} \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2} d\mathbf{y} = [\tilde{\mathbf{K}}_{s,t}^{(\tau, \boldsymbol{\xi})}]_{1,1}, \quad (3.13)$$

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_1 d\mathbf{y} = \mathbf{0}_d, \quad (3.14)$$

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \cdot (\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_1 d\mathbf{y} = \mathbf{0}_{d,d}, \quad k \in \llbracket 1, n \rrbracket, \quad (3.15)$$

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \text{Tr} \left( \mathbf{M}(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2} \right) d\mathbf{y} = 2\mathbf{M}, \quad \mathbf{M} \in \mathbb{R}^d \otimes \mathbb{R}^d, \quad (3.16)$$

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \text{Tr} \left( \mathbf{M}(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2} \right) d\mathbf{y} = \mathbf{0}_{d,d,d}, \quad k \in \llbracket 1, n \rrbracket, \quad \mathbf{M} \in \mathbb{R}^d \otimes \mathbb{R}^d. \quad (3.17)$$

Where, in (3.13), we define for all  $(i, j) \in \llbracket 1, n \rrbracket^2$  and  $M \in (\mathbb{R}^{nd})^{\otimes 2}$ ,  $[M]_{i,j}$  is the  $d \times d$  block matrix corresponding to the entry of  $M$  on the  $i^{\text{th}}$  line and the  $j^{\text{th}}$  column.

**Proof:** First of all remark that equation (3.13) simply follows from a direct covariance computation.

Observe now that from Proposition 2, we have  $\int_{\mathbb{R}^{nd}} \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_1 d\mathbf{y} = \mathbf{0}_d$ . Differentiating twice this expression w.r.t.  $\mathbf{x}_1$  and using the Leibniz formula (recalling as well the identity (3.4) which yields  $D_{\mathbf{x}_1}[\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})]_1 = [\tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(s, t)]_{1,1} = I_{d,d}$ ) gives (3.14). Iterating the differentiation w.r.t.  $D_{\mathbf{x}_k}$  then yields (3.15) (observing again that  $D_{\mathbf{x}_k}[\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})]_1 = [\tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(s, t)]_{1,k}$ , i.e.  $D_{\mathbf{x}_k}[\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})]_1 = I_{d,d}$  if  $k = 1$  and  $\mathbf{0}_{d,d}$  for  $k > 1$ ). Observe that  $D_{\mathbf{x}_1} \int_{\mathbb{R}^{nd}} \text{Tr} \left( \mathbf{M}(\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_1^{\otimes 2} \right) \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y} = D_{\mathbf{x}_1} \text{Tr} \left( \mathbf{M}[\tilde{\mathbf{K}}_{s,t}^{(\tau, \boldsymbol{\xi})}]_{1,1} \right) = \mathbf{0}_d$ . Differentiating again w.r.t.  $D_{\mathbf{x}_1}$ , the Leibniz formula and identity  $D_{\mathbf{x}_1}[\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})]_1 = [\tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(s, t)]_{1,1} = I_{d,d}$  yield (3.16). Eventually, (3.17) can be derived again from derivation or observing that the sum of the length of the multi-derivation index, here 3, and the power integrated, here 2, is an odd number.  $\square$

We conclude this section with a key lemma which will be useful to perform our sensitivity analysis of Section 5.

**Lemma 1 (Sensitivities of the covariance)** *There exists  $C := C((\mathbf{A}))$  s.t. for given  $(\boldsymbol{\xi}, \boldsymbol{\xi}') \in (\mathbb{R}^{nd})^2$  and  $0 \leq t < s \leq T$ ,  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ :*

$$|[\tilde{\mathbf{K}}_{s,t}^{(\tau, \boldsymbol{\xi})}]_{1,1} - [\tilde{\mathbf{K}}_{s,t}^{(\tau, \boldsymbol{\xi}')} ]_{1,1}| \leq C(s-t)(\mathbf{d}^\gamma(\boldsymbol{\xi}, \boldsymbol{\xi}') + (s-t)^{\frac{\gamma}{2}}). \quad (3.18)$$

The proof of Lemma 1 is postponed to Appendix 7 for the sake of clarity. We importantly point out that, in Lemma 1, the constant  $C$  mainly depends on the Hölder norms of the coefficients and is *small* provided the coefficients do not *vary much*. Precisely, it can be shown that  $C$  writes:

$$C := \tilde{C}(\|a\|_{L^\infty(C_d^\gamma)} + \sum_{i=2}^n \|\mathbf{F}_i\|_{L^\infty(C_d^{2i-3+\gamma})}) \quad (3.19)$$

for some *universal* constant  $\tilde{C}$ , where, in the above equation, we write with the notation of Section 1.2:

$$\begin{aligned} \|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})} &:= \sup_{(t,\mathbf{z}) \in [0,T] \times \mathbb{R}^{(n-i+2)d}} \|(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1}(t, \mathbf{z}, \cdot)\|_{C^{\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)} \\ &+ \sum_{j=i}^n \sup_{(t,\mathbf{z}) \in [0,T] \times \mathbb{R}^{(n-i+2)d}} \|(\mathbf{F}_i)_j(t, \mathbf{z}, \cdot)\|_{C^{\frac{2i-3+\gamma}{2j-1}}(\mathbb{R}^d, \mathbb{R}^d)}. \end{aligned} \quad (3.20)$$

Namely, the quantity  $\|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})}$  gathers the Hölder moduli of  $\mathbf{F}_i$  at the intrinsic associated scales according to the distance  $\mathbf{d}$  in the variables  $j \in \llbracket i, n \rrbracket$  as well as the Hölder norm of the gradient w.r.t. the component which transmits the noise (but importantly not its supremum norm). Said differently, the  $L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})$  norm of  $\mathbf{F}_i$  gathers the Hölder norms of the fractional parts of  $(D_{\mathbf{x}_j}^{\lfloor \frac{2i-3+\gamma}{2j-1} \rfloor} \mathbf{F}_i)_{j \in \llbracket i-1, n \rrbracket}$  in the  $j^{\text{th}}$  variable with corresponding Hölder index  $\frac{2i-3+\gamma}{2j-1} - \lfloor \frac{2i-3+\gamma}{2j-1} \rfloor$ , where  $\lfloor \cdot \rfloor$  stands for the integer part.

We again refer to Appendix 7 for a precise statement and proof of this assertion (3.19) (See also Lemma 12 below and its proof for similar properties).

### 3.2 Additional controls associated with the (mollified) flow and its linearization

We now state three important estimates associated with the linearization or sensitivity w.r.t. the initial point for the frozen (mollified) differential system (2.5). For the sake of simplicity their proof is also postponed to Appendix 7. We have:

**Lemma 2** *There exists  $C := C(\mathbf{A})$  s.t. for all  $0 \leq t \leq s \leq T$ ,  $(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$  and  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ :*

$$\mathbf{d}(\boldsymbol{\theta}_{s,t}(\mathbf{x}), \boldsymbol{\theta}_{s,t}(\mathbf{x}')) \leq C \left( \mathbf{d}(\mathbf{x}, \mathbf{x}') + (s-t)^{\frac{1}{2}} \right).$$

The proof of Lemma 2 is postponed to Appendix 7.1.

The second important result concerns the impact of the freezing point in the linearization procedure. Namely,

**Lemma 3 (Sensitivity of the linearized flow w.r.t. the freezing parameter)**

*There exists  $C := C(\mathbf{A})$  s.t. for all  $\tau = t$ ,  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$  at the change of regime time  $t_0$  defined in (2.37):*

$$\mathbf{d}(\mathbf{m}_{t_0,t}^{(\tau,\mathbf{x})}(\mathbf{x}'), \mathbf{m}_{t_0,t}^{(\tau,\mathbf{x}')}(\mathbf{x}')) = \mathbf{d}(\mathbf{m}_{t_0,t}^{(\tau,\mathbf{x})}(\mathbf{x}'), \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')) \leq C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}').$$

Again, the proof of Lemma 3 is postponed to Appendix 7.1.

## 4 Control of the supremum of the derivatives in the non-degenerate variables

**WARNING:** for notational simplicity, we drop from now on the sub and superscripts in  $\tau$  associated with the linearization, since this parameter is chosen to be equal to  $t$ . For example,  $\tilde{p}^{(\tau, \xi)}$ ,  $\mathbf{m}_{s,t}^{(\tau, \xi)}$ ,  $\boldsymbol{\theta}_{u, \tau}$  become respectively  $\tilde{p}^\xi$ ,  $\mathbf{m}_{s,t}^\xi$ ,  $\boldsymbol{\theta}_{u,t}$ .

The point is now to control  $\|u_m\|_{C_{b,d}^{2+\alpha}}$ . From (2.3) it is direct to derive that for all  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

$$|u(t, \mathbf{x})| \leq \|g\|_{L^\infty} + T\|f\|_{L^\infty}.$$

Let us now start from (2.11) to control pointwise the second order derivatives of  $u_m$  in the non-degenerate variables, i.e.  $\|D_{\mathbf{x}_1}^2 u\|_{L^\infty}$ . The first ones can be controlled similarly and more directly. Write for all  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

$$\begin{aligned} & |D_{\mathbf{x}_1}^2 u(t, \mathbf{x})| \\ & \leq \|D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^\xi g\|_\infty + \|D_{\mathbf{x}_1}^2 \tilde{G}^\xi f\|_\infty \\ & \quad + \left| \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \Delta_{1, \mathbf{F}, \sigma}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}, u) d\mathbf{y} \right| \\ & \quad + \left| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{y}_i} \cdot \left( (D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \otimes \Delta_{i, \mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y})) \right) u(s, \mathbf{y}) d\mathbf{y} \right|, \end{aligned} \tag{4.1}$$

we recall from (2.12) and (2.15) that:

$$\begin{aligned} \Delta_{1, \mathbf{F}, \sigma}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}, u) &= \langle (\mathbf{F}_1(s, \mathbf{y}) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))), D_{\mathbf{y}_1} u(s, \mathbf{y}) \rangle \\ & \quad + \frac{1}{2} \text{Tr} \left( (a(s, \mathbf{y}) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) D_{\mathbf{y}_1}^2 u(s, \mathbf{y}) \right), \\ \Delta_{i, \mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}) &= \mathbf{F}_i(s, \mathbf{y}) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}. \end{aligned} \tag{4.2}$$

### 4.1 Control of the non-degenerate part of the perturbative term

The aim of this section is to prove identity (2.14) appearing in the user's guide to the proof.

To this end, we provide a general differentiation result, which will be useful as well in Section 5 to deal with the Hölder norms. Under the current assumptions on  $a, \mathbf{F}$ , the following lemma holds.

**Lemma 4 (First Derivative Control Lemma for the non-degenerate variable)**

For all multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^d$ :

$$\left\| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \Delta_{1, \mathbf{F}, \sigma}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}, u) d\mathbf{y} \right\|_{\xi=\mathbf{x}} \leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} (s-t)^{-\sum_{j=1}^n \alpha_j (j-\frac{1}{2}) + \frac{\gamma}{2}}. \tag{4.3}$$

**Proof: Lemma 4**

We first recall the control (2.12)

$$|\Delta_{1,\mathbf{F},\sigma}(t, s, \mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), u)| \leq \left( [\mathbf{F}_1]_{\mathbf{d},\gamma} \|D_{\mathbf{x}_1} u(s, \cdot)\|_{L^\infty} + \frac{1}{2} [a(s, \cdot)]_{\mathbf{d},\gamma} \|D_{\mathbf{x}_1}^2 u(s, \cdot)\|_{L^\infty} \right) \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})).$$

From this control and Proposition 2, we directly obtain:

$$\begin{aligned} & \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \Delta_{1,\mathbf{F},\sigma}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}, u) d\mathbf{y} \right|_{\boldsymbol{\xi}=\mathbf{x}} \\ & \leq \left( \int_{\mathbb{R}^{nd}} |D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y})| \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \left( \|\mathbf{F}_1\|_{L^\infty(C_{\mathbf{d}}^\gamma)} \|D_{\mathbf{x}_1} u\|_{L^\infty} + \|a\|_{L^\infty(C_{\mathbf{d}}^\gamma)} \|D_{\mathbf{x}_1}^2 u\|_{L^\infty} \right) d\mathbf{y} \right) \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\ & \leq C \|u\|_{L^\infty(C_{\mathbf{b},\mathbf{d}}^{2+\gamma})} (s-t)^{-\sum_{j=1}^n \alpha_j (j-\frac{1}{2})} \int_{\mathbb{R}^{nd}} \bar{p}_{C^{-1}}(s, t, \mathbf{x}, \mathbf{y}) \mathbf{d}^\gamma(\boldsymbol{\theta}_{s,t}(\mathbf{x}), \mathbf{y}) d\mathbf{y} \\ & \leq C \|u\|_{L^\infty(C_{\mathbf{b},\mathbf{d}}^{2+\gamma})} (s-t)^{-\sum_{j=1}^n \alpha_j (j-\frac{1}{2}) + \frac{\gamma}{2}}, \end{aligned}$$

with the notations of Remark 3 for the last but one inequality.  $\square$  Equation (2.14) readily follows from Lemma 4 taking  $\alpha = (2, 0, \dots, 0)$ . Namely:

$$\begin{aligned} & \left| \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \Delta_{1,\mathbf{F},\sigma}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}, u) d\mathbf{y} \right|_{\boldsymbol{\xi}=\mathbf{x}} \\ & \leq C \|u\|_{L^\infty(C_{\mathbf{b},\mathbf{d}}^{2+\gamma})} \int_t^T \frac{ds}{(s-t)^{1-\frac{\gamma}{2}}} \leq C \|u\|_{L^\infty(C_{\mathbf{b},\mathbf{d}}^{2+\gamma})} (T-t)^{\frac{\gamma}{2}}. \end{aligned} \quad (4.4)$$

## 4.2 Control of the degenerate part of the perturbative term

The point is here to control the terms

$$\sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \rangle d\mathbf{y}$$

appearing in equation (2.16) of the user's guide to the proof. We precisely want to derive equation (2.22).

The bound will actually follow from the more general following result, which will again be useful for the Hölder norm in Section 5.

**Lemma 5 (First Besov Control Lemma)** *For all multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^d$ :*

$$\begin{aligned} & \sum_{i=2}^n \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \rangle d\mathbf{y} \right|_{\boldsymbol{\xi}=\mathbf{x}} \\ & \leq C \|u\|_{L^\infty(C_{\mathbf{b},\mathbf{d}}^{2+\gamma})} (s-t)^{-\sum_{j=1}^n \alpha_j (j-\frac{1}{2}) + \frac{\gamma}{2}}. \end{aligned} \quad (4.5)$$

**Proof: Lemma 5**

Similarly to (2.16), we define for all  $i \in \llbracket 2, n \rrbracket$ ,

$$\vartheta_{i,(t,\mathbf{x})}^\alpha(s, \mathbf{y}) := D_{\mathbf{y}_i} \cdot (D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \otimes \Delta_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y})) =: D_{\mathbf{y}_i} \cdot (\Theta_{i,(t,\mathbf{x})}^\alpha(s, \mathbf{y})). \quad (4.6)$$



The contribution of the l.h.s. in (4.5) then rewrites:

$$\begin{aligned} & \sum_{i=2}^n \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}}^{\alpha} \tilde{p}^{(t, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i, \mathbf{F}}(t, s, \boldsymbol{\theta}_{s, t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \rangle d\mathbf{y} \right|_{\boldsymbol{\xi}=\mathbf{x}} \\ &= \sum_{i=2}^n \left| \int_{\mathbb{R}^{nd}} \vartheta_{i, (t, \mathbf{x})}^{\alpha}(s, \mathbf{y}) u(s, \mathbf{y}) d\mathbf{y} \right|. \end{aligned} \quad (4.7)$$

The point now is to observe that for any fixed  $i \in \llbracket 2, n \rrbracket$  and  $\mathbf{z} = (\mathbf{z}_1, \dots, \mathbf{z}_{i-1}, \mathbf{z}_{i+1}, \dots, \mathbf{z}_n) \in \mathbb{R}^{(n-1)d}$  the mapping  $\mathbf{y}_i \mapsto u(s, \mathbf{z}_{1:i-1}, \mathbf{y}_i, \mathbf{z}_{i+1:n})$  is in  $C_b^{\frac{2+\gamma}{2i-1}}(\mathbb{R}^d) = B_{\infty, \infty}^{\frac{2+\gamma}{2i-1}}(\mathbb{R}^d)$  using the Besov space terminology, see e.g. Triebel [Tri83], uniformly in  $s \in [0, T]$ . The point is now to put in duality the mappings  $\mathbf{y}_i \mapsto u_m(s, \mathbf{y}_{1:i-1}, \mathbf{y}_i, \mathbf{y}_{i+1:n})$  and

$$\Psi_{i, (t, \mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha} : \mathbf{y}_i \mapsto \vartheta_{i, (t, \mathbf{x})}^{\alpha}(s, \mathbf{y}), \quad (4.8)$$

see e.g. Proposition 3.6 in [LR02]. This precisely means that we have to prove that  $\Psi_{i, (t, \mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha}$  lies in the suitable Besov space,  $\Psi_{i, (t, \mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha} \in B_{1,1}^{-\frac{2+\gamma}{2i-1}}(\mathbb{R}^d)$  and to control the associated norm. We will actually prove that those norms provide an integrable quantity w.r.t.  $\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}$  as well as an integrable time singularity. This will be done through the thermic characterization of Besov spaces, see e.g. Section 2.6.4 in [Tri83] as well as (2.18) above. Precisely, we recall that for a function  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$  in  $B_{1,1}^{-\tilde{\alpha}_i}(\mathbb{R}^d)$ ,  $\tilde{\alpha}_i := \frac{2+\gamma}{2i-1}$  a quasi-norm is given by:

$$\|\psi\|_{B_{1,1}^{-\tilde{\alpha}_i}(\mathbb{R}^d)} := \|\varphi(D)\psi\|_{L^1(\mathbb{R}^d, \mathbb{R})} + \int_0^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \psi\|_{L^1(\mathbb{R}^d, \mathbb{R})}, \quad (4.9)$$

with

$$\forall z \in \mathbb{R}^d, \quad h_v(z) := \frac{1}{(2\pi v)^{\frac{d}{2}}} \exp\left(-\frac{|z|^2}{v}\right),$$

being the usual *heat kernel* of  $\mathbb{R}^d$ ,  $\star$  standing for the usual convolution on  $\mathbb{R}^d$  for  $\varphi \in C_0^{\infty}(\mathbb{R}^d, \mathbb{R})$  s.t.  $\varphi(0) \neq 0$ .

Taking  $\psi = \Psi_{i, (t, \mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha}$  in the above characterization and from definition (4.6), the main advantage of using (4.9) consists in rebalancing the derivative appearing in the definition (4.6) to the heat kernel or to the smooth compactly supported function  $\varphi$ . Precisely, focusing on the heat kernel part of (4.9), the other contribution being handled more directly, we get:

$$\begin{aligned} & \|h_v \star \Psi_{i, (t, \mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha}\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} h_v(z - \mathbf{y}_i) \vartheta_{i, (t, \mathbf{x})}^{\alpha}(s, \mathbf{y}) d\mathbf{y}_i \right| dz \\ &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} h_v(z - \mathbf{y}_i) D_{\mathbf{y}_i} \cdot (\Theta_{i, (t, \mathbf{x})}^{\alpha}(s, \mathbf{y})) d\mathbf{y}_i \right| dz \\ &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} D_{\mathbf{x}}^{\alpha} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i, \mathbf{F}}(t, s, \boldsymbol{\theta}_{s, t}(\boldsymbol{\xi}), \mathbf{y}), D_z h_v(z - \mathbf{y}_i) \rangle d\mathbf{y}_i \right| dz. \end{aligned} \quad (4.10)$$

We now want to estimate the  $\|\Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{B_{1,1}^{-\hat{\alpha}_i}}$ . To this end, we split the time integral in (4.9) into two parts writing:

$$\begin{aligned} & \int_0^1 \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\ &= \int_0^{(s-t)^{\beta_i}} \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\ & \quad + \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{L^1(\mathbb{R}^d,\mathbb{R})} \end{aligned} \quad (4.11)$$

for a parameter  $\beta_i > 0$  to be specified. Precisely, in order to have a similar smoothing effect in time than for the terms appearing in (4.4), we now want to calibrate  $\beta_i$  to obtain:

$$\begin{aligned} & \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\ & \leq \frac{C}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-\frac{\gamma}{2}}} \hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})), \end{aligned} \quad (4.12)$$

where introducing:

$$\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}) := \prod_{j=1}^n \mathcal{N}_{c(s-t)^{2j-1}}((\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j) = \bar{p}_{c^{-1}}(t, s, \mathbf{x}, \mathbf{y}),$$

where for  $a > 0$ ,  $z \in \mathbb{R}^d$ ,  $\mathcal{N}_a(z) = \frac{1}{(2\pi a)^{\frac{d}{2}}} \exp(-\frac{|z|^2}{2a})$  stands for the standard Gaussian density of  $\mathbb{R}^d$  with covariance matrix  $aI_d$ , we introduce:

$$\hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) = \prod_{j \in \llbracket 1, n \rrbracket, j \neq i} \mathcal{N}_{c(s-t)^{2j-1}}((\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j). \quad (4.13)$$

To choose properly the parameter  $\beta_i$  leading to (4.12), we now write:

$$\begin{aligned} & \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\ & \leq \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_z h_v(z - \mathbf{y}_i) \rangle d\mathbf{y}_i \right|_{\boldsymbol{\xi}=\mathbf{x}} \\ & \leq C \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} \mathbf{d}^{2i-3+\gamma}(\boldsymbol{\theta}_{s,t}(\mathbf{x}), \mathbf{y}) \\ & \leq C \int_{(s-t)^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\hat{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} (s-t)^{\frac{2i-3+\gamma}{2}} \\ & \leq C \hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_{(s-t)^{\beta_i}}^1 dv v^{-\frac{3}{2}+\frac{\hat{\alpha}_i}{2}} (s-t)^{-\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-3+\gamma}{2}} \\ & \leq C \hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) (s-t)^{[-\frac{1}{2}+\frac{\hat{\alpha}_i}{2}]\beta_i - \sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-3+\gamma}{2}} \end{aligned}$$

using Proposition 2 for the third inequality (see also Remark 4, equation (3.8)) recalling as well that  $0 \leq t < s \leq T$  is small.

To obtain (4.12), we then take:

$$\left[-\frac{1}{2} + \frac{\tilde{\alpha}_i}{2}\right]\beta_i - \sum_{j=1}^n \alpha_j \left(j - \frac{1}{2}\right) + \frac{2i-3+\gamma}{2} = - \sum_{j=1}^n \alpha_j \left(j - \frac{1}{2}\right) + \frac{\gamma}{2} \iff \beta_i = \frac{(2i-3)(2i-1)}{2i-3-\gamma}. \quad (4.14)$$

The key point is now to check that the previous choice of  $\beta_i$  also yields a bound similar to (4.12) for the contribution in (4.11) associated with  $v \in [0, (s-t)^{\beta_i}]$ .

To this end, we restart from identity (4.10), which allows to exploit partial cancellations w.r.t. the integration variable  $\mathbf{y}_i$ . Namely, write:

$$\begin{aligned} & \int_{\mathbb{R}^d} h_v(z - \mathbf{y}_i) D_{\mathbf{y}_i} \cdot \Theta_{i,(t,\mathbf{x})}^\alpha(s, \mathbf{y}) d\mathbf{y}_i \\ &= \int_{\mathbb{R}^d} h_v(z - \mathbf{y}_i) D_{\mathbf{y}_i} \cdot \left( \Theta_{i,(t,\mathbf{x})}^\alpha(s, \mathbf{y}) - \Theta_{i,(t,\mathbf{x})}^\alpha(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right) d\mathbf{y}_i \\ &= \int_{\mathbb{R}^d} D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \mathbf{F}_i(s, \mathbf{y}) - \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}), D_z h_v(z - \mathbf{y}_i) \rangle d\mathbf{y}_i \\ & \quad + \int_{\mathbb{R}^d} \left( D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right) \\ & \quad \langle \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \\ & \quad - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}, D_z h_v(z - \mathbf{y}_i) \rangle d\mathbf{y}_i \\ &=: \left( \mathcal{T}_1 + \mathcal{T}_2 \right) (v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n})), \end{aligned} \quad (4.15)$$

using the definition in (4.6) and (4.10) for the last decomposition.

Write now:

$$\begin{aligned} |\mathcal{T}_1(v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}))| &\leq C \int_{\mathbb{R}^d} \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} |z - \mathbf{y}_i|^{\frac{2i-3+\gamma}{2i-1}} d\mathbf{y}_i \\ &\leq C \int_{\mathbb{R}^d} \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{2-\gamma}{4i-2}}} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} d\mathbf{y}_i. \end{aligned} \quad (4.16)$$

We thus derive from (4.16):

$$\begin{aligned} & |\mathcal{T}_1(v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}))| \quad (4.17) \\ &\leq \frac{C}{v^{\frac{2-\gamma}{4i-2}} (s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} \hat{q}_{c,i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \mathcal{N}_{cv+(s-t)^{2i-1}}(z - \boldsymbol{\theta}_{s,t}(\mathbf{x})_i). \end{aligned}$$

Write now from (4.15) and the Taylor formula applied to the  $i^{\text{th}}$  variable for the difference of the derivatives of the densities:

$$\begin{aligned}
 & |\mathcal{T}_2(v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}))| \\
 \leq & C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\mu \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} \\
 & \times |\mathbf{y}_i - z| \left( \left| \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) - \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}(\mathbf{x})_{i:n}) \right| \right. \\
 & \left. + \left| \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}(\mathbf{x})_{i:n}) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\mathbf{x}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1} \right| \right) \\
 \leq & C \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \int_0^1 d\mu \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} \\
 & \times \left( |z - \boldsymbol{\theta}_{s,t}(\mathbf{x})_i|^{\frac{2i-3+\gamma}{2i-1}} + |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_{i-1}|^{1 + \frac{\gamma}{2(i-1)-1}} + \sum_{k=i+1}^n |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_k|^{\frac{2i-3+\gamma}{2k-1}} \right).
 \end{aligned}$$

Writing, for any  $\mu \in [0, 1]$ ,

$$|z - \boldsymbol{\theta}_{s,t}(\mathbf{x})_i| \leq \mu |z - \mathbf{y}_i| + |z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|,$$

we thus derive

$$\begin{aligned}
 & |\mathcal{T}_2(v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}))| \\
 \leq & C \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \int_0^1 d\mu \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} \\
 & \times \left( |\mathbf{y}_i - z|^{\frac{2i-3+\gamma}{2i-1}} + \mathbf{d}^{2i-3+\gamma} \left( \boldsymbol{\theta}_{s,t}(\mathbf{x}), (\mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \right) \right) \\
 \leq & C \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \int_0^1 d\mu \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \\
 & \times \left( \frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \right) \\
 \leq & C \hat{q}_{c|i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \int_0^1 d\mu \int_{\mathbb{R}^d} h_{cv}(z - \mathbf{y}_i) \mathcal{N}_{c(s-t)^{2i-1}}(z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) d\mathbf{y}_i \\
 & \times \left( \frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \right), \tag{4.18}
 \end{aligned}$$

using again (3.8) for the second inequality. From (4.15), (4.17) and (4.18) we derive, with

the notation introduced in (4.13):

$$\begin{aligned}
 & \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1})}^\alpha\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
 \leq & \left( \frac{1}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} + \frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}} \right) \\
 & \times C \hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \\
 & \times \int_0^1 d\mu \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \mathcal{N}_{c(s-t)^{2i-1}}(z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) \\
 \leq & C \hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \\
 & \times \left( \frac{1}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} + \frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}} \right),
 \end{aligned}$$

using the change of variable  $(w_1, w_2) = (z - \mathbf{y}_i, z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)$  for the last inequality. From the above computations and with the notations of (4.11), we derive:

$$\begin{aligned}
 & \int_0^{(s-t)^{\beta_i}} dv v^{\frac{\tilde{\alpha}_i}{2}-1} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1})}^\alpha\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
 \leq & C \hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_0^{(s-t)^{\beta_i}} \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \\
 & \times \left( \frac{1}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} + \frac{v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}} \right) \\
 =: & C \hat{q}_{c\setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \mathcal{B}_{\alpha,\beta_i}(t, s).
 \end{aligned}$$

Let us now prove that for  $\beta_i = \frac{(2i-3)(2i-1)}{2i-3-\gamma}$ , defined in (4.14), we have:

$$\mathcal{B}_{\alpha,\beta_i}(t, s) \leq \frac{C}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-\frac{\gamma}{2}}}. \quad (4.19)$$

To prove (4.19), we now write:

$$\begin{aligned}
 & \mathcal{B}_{\alpha,\beta_i}(t, s) \\
 \leq & C \left[ \frac{v^{\frac{\tilde{\alpha}_i}{2}-\frac{2-\gamma}{4i-2}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} + \frac{v^{\frac{\tilde{\alpha}_i}{2}+\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-1}{2}}} + \frac{v^{\frac{\tilde{\alpha}_i}{2}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}} \right]_{v=0}^{v=(s-t)^{\beta_i}} \\
 \leq & C \left[ (s-t)^{\beta_i(\frac{\tilde{\alpha}_i}{2}-\frac{2-\gamma}{4i-2})-\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} + (s-t)^{\beta_i(\frac{\tilde{\alpha}_i}{2}+\frac{2i-3+\gamma}{2(2i-1)})-(\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-1}{2})} \right. \\
 & \left. + (s-t)^{\beta_i(\frac{\tilde{\alpha}_i}{2}-\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-1+\frac{\gamma}{2})} \right].
 \end{aligned}$$

From the above equation, (4.19) holds as soon as  $\beta_i$  can be chosen so that the three following conditions hold:

$$\beta_i \left( \frac{\tilde{\alpha}_i}{2} - \frac{2-\gamma}{4i-2} \right) - \frac{\gamma}{2} \geq 0, \quad \beta_i \left( \frac{\tilde{\alpha}_i}{2} + \frac{2i-3+\gamma}{2(2i-1)} \right) - \frac{2i-1}{2} - \frac{\gamma}{2} \geq 0, \quad \beta_i \frac{\tilde{\alpha}_i}{2} - 1 \geq 0.$$

Recalling that  $\frac{\tilde{\alpha}_i}{2} = \frac{1+\frac{\gamma}{2}}{2i-1}$  and for the previous choice of  $\beta_i$ , the above conditions rewrite:

$$\begin{aligned} \left( \frac{(2i-3)(2i-1)}{2i-3-\gamma} \right) \left( \frac{1+\frac{\gamma}{2}}{2i-1} - \frac{1-\frac{\gamma}{2}}{2i-1} \right) - \frac{\gamma}{2} \geq 0 &\iff \frac{(2i-3)}{2i-3-\gamma} \gamma - \frac{\gamma}{2} \geq 0, \\ \left( \frac{(2i-3)(2i-1)}{2i-3-\gamma} \right) \left( \frac{2+\gamma}{2(2i-1)} + \frac{2i-3+\gamma}{2(2i-1)} \right) - \frac{2i-1}{2} - \frac{\gamma}{2} \geq 0 \\ &\iff \left( \frac{2i-3}{2i-3-\gamma} \right) (2i-1+2\gamma) - (2i-1+\gamma) \geq 0, \\ \left( \frac{(2i-3)(2i-1)}{2i-3-\gamma} \right) \frac{1+\frac{\gamma}{2}}{2i-1} - 1 \geq 0 &\iff \left( \frac{2i-3}{2i-3-\gamma} \right) \left( 1 + \frac{\gamma}{2} \right) - 1 \geq 0. \end{aligned}$$

All the above conditions are true for  $i \in \llbracket 2, n \rrbracket$ ,  $\gamma \in (0, 1]$ . Note the chosen  $\beta_i$  seems to be rather sharp in the sense that letting  $\gamma$  go to 0 the above constraints become equalities. This proves (4.19). We finally get:

$$\int_0^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{L^1(\mathbb{R}^d, \mathbb{R})} \leq \frac{C}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-\frac{\gamma}{2}}} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})). \quad (4.20)$$

Reproducing the previous computations we also write for a  $C^\infty$  compactly supported function  $\varphi$ :

$$\begin{aligned} &\|\varphi(D)\Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ &\leq \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} D_{\mathbf{y}_i} \hat{\varphi}(z - \mathbf{y}_i) \cdot (D_{\mathbf{x}}^\alpha \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \otimes \Delta_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y})) d\mathbf{y}_i \right| dz \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\ &\leq \frac{C}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} \int_{\mathbb{R}^d} \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}) d^{2i-3+\gamma}(\boldsymbol{\theta}_{s,t}(\mathbf{x}), \mathbf{y}) d\mathbf{y}_i \\ &\leq \frac{C}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-(i-\frac{3}{2}+\frac{\gamma}{2})}} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})). \end{aligned}$$

From (4.9) and (4.20), we finally obtain:

$$\|\Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^\alpha\|_{B_{1,1}^{-\tilde{\alpha}_i}} \leq \frac{C}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-\frac{\gamma}{2}}} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})), \quad (4.21)$$

which together with (4.7) and (4.8) gives the result.  $\square$

Equation (2.23) now follows from Lemma 5 taking  $\alpha = (2, 0, \dots, 0)$ . Namely,

$$\begin{aligned} &\left| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \rangle \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\ &\leq C(T-t)^{\frac{\gamma}{2}} \|u\|_{L^\infty(C_{b,d}^{2+\gamma})}. \end{aligned} \quad (4.22)$$

### 4.3 Non-degenerate derivatives for the frozen semi-group : terminal condition and source

The main result of this section is the following lemma.

**Lemma 6 (Derivatives of frozen semi-group and Green kernel)** *There exists a constant  $C := C((\mathbf{A}))$  s.t. for all  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ ,*

$$\begin{aligned} |D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^\xi g(\mathbf{x})| \Big|_{\xi=\mathbf{x}} &\leq C \|D_{\mathbf{x}_1}^2 g\|_{L^\infty} \leq C \|g\|_{C_{b,d}^{2+\gamma}}, \\ |D_{\mathbf{x}_1}^2 \tilde{G}^\xi f(t, \mathbf{x})| \Big|_{\xi=\mathbf{x}} &\leq C(T-t)^{\frac{\gamma}{2}} \|f\|_{L^\infty(C_{b,d}^\gamma)}. \end{aligned}$$

**Proof: Lemma 6**

Note first that, since  $\tilde{p}^\xi$  is a density one has the following cancellation identity:

$$\begin{aligned} &\left| D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^\xi g(\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \\ &= \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) [g(\mathbf{y}) - g(\mathbf{m}_{T,t}^\xi(\mathbf{x}))] d\mathbf{y} \right| \Big|_{\xi=\mathbf{x}} \\ &\leq \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) [g(\mathbf{y}) - g(\mathbf{y}_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n})] d\mathbf{y} \right| \Big|_{\xi=\mathbf{x}} \\ &\quad + \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) [g(\mathbf{y}_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n}) - g(\mathbf{m}_{T,t}^\xi(\mathbf{x}))] d\mathbf{y} \right| \Big|_{\xi=\mathbf{x}}. \end{aligned} \tag{4.23}$$

The first term in the r.h.s. of the previous identity is readily controlled thanks to Proposition 2

$$\begin{aligned} &\left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) [g(\mathbf{y}) - g(\mathbf{y}_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n})] d\mathbf{y} \right| \Big|_{\xi=\mathbf{x}} \\ &\leq C \|g\|_{C_{b,d}^{2+\gamma}} \int_{\mathbb{R}^{nd}} (T-t)^{-1} \bar{p}_{C^{-1}}(t, T, \mathbf{x}, \mathbf{y}) d^{2+\gamma}(\mathbf{m}_{T,t}^\xi(\mathbf{x}), \mathbf{y}) d\mathbf{y} \Big|_{\xi=\mathbf{x}} \\ &\leq C(T-t)^{\frac{\gamma}{2}} \|g\|_{C_{b,d}^{2+\gamma}}. \end{aligned} \tag{4.24}$$

The second term of (4.23) is more subtle. We need to expand  $g(\mathbf{y}_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n})$  in its non-degenerate variable to take advantage of the corresponding regularity of  $g$ . Namely,:

$$\begin{aligned} &\left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) [g(\mathbf{y}_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n}) - g(\mathbf{m}_{T,t}^\xi(\mathbf{x}))] d\mathbf{y} \right| \Big|_{\xi=\mathbf{x}} \\ &= \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) \left( \langle D_{\mathbf{x}_1} g(\mathbf{m}_{T,t}^\xi(\mathbf{x})), (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1 \rangle \right. \right. \\ &\quad \left. \left. + \int_0^1 d\mu(1-\mu) \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x})_1 + \mu(\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n}) (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1^{\otimes 2} \right) \right) d\mathbf{y} \right| \Big|_{\xi=\mathbf{x}} \\ &\leq C \|D_{\mathbf{x}_1}^2 g\|_{L^\infty} \int_{\mathbb{R}^{nd}} \frac{\bar{p}_{C^{-1}}(t, T, \mathbf{x}, \mathbf{y})}{T-t} |(\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1|^2 d\mathbf{y} \Big|_{\xi=\mathbf{x}} \leq C \|g\|_{C_{b,d}^{2+\gamma}}, \end{aligned} \tag{4.25}$$

recalling from Proposition 3 for the last but one inequality that

$$\int_{\mathbb{R}^d} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{m,\xi}(t, T, \mathbf{x}, \mathbf{y}) \langle D_{\mathbf{x}_1} g(\mathbf{m}_{T,t}^\xi(\mathbf{x})), (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1 \rangle \Big|_{\xi=\mathbf{x}} = \mathbf{0}_{d,d}.$$

Gathering identities (4.24), (4.25) into (4.23), we obtain the stated control for the term  $|D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi} g(\mathbf{x})|_{\xi=\mathbf{x}}$ .

Let us now turn to the Green kernel. We directly get from Proposition 2:

$$\begin{aligned} & |D_{\mathbf{x}_1}^2 \tilde{G}^{\xi} f(t, \mathbf{x})|_{\xi=\mathbf{x}} \\ & \leq \left| \int_t^T ds \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) [f(s, \mathbf{y}) - f(s, \boldsymbol{\theta}_{s,t}(\xi))] d\mathbf{y} \right|_{\xi=\mathbf{x}} \\ & \leq C[f]_{L^\infty(C^\gamma)} \int_t^T ds \int_{\mathbb{R}^{nd}} \frac{1}{s-t} \bar{p}_{C^{-1}}(s, t, \mathbf{x}, \mathbf{y}) \mathbf{d}(\boldsymbol{\theta}_{s,t}(\mathbf{x}), \mathbf{y})^\gamma d\mathbf{y} \\ & \leq C[f]_{L^\infty(C^\gamma)} (T-t)^{\frac{\gamma}{2}}, \end{aligned}$$

which gives the result.  $\square$

Similarly to Lemma 6, we prove from Proposition 2 the following result:

**Lemma 7** For  $\gamma \in (0, 1)$ , under **(A)**, there exists  $C := C((\mathbf{A}), T)$ , s.t. for all function  $\psi \in C_{b,\mathbf{d}}^\gamma(\mathbb{R}^{nd}, \mathbb{R})$ , and any given multi-index  $\alpha$ ,  $|\alpha| \in \llbracket 1, 3 \rrbracket$ , for all  $0 \leq t < s \leq T$ ,  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$|D_{\mathbf{x}}^\alpha \tilde{P}_{s,t}^{\xi} \psi(\mathbf{x})|_{\xi=\mathbf{x}} \leq C \|\psi\|_{C_{b,\mathbf{d}}^\gamma} (s-t)^{-\sum_{i=1}^n |\alpha_i|(i-\frac{1}{2})+\frac{\gamma}{2}}. \quad (4.26)$$

Note carefully that in the above lemma  $|\alpha| \geq 1$ . Indeed, if  $|\alpha| = 0$  we cannot benefit from any regularizing effects which are precisely due to cancellation techniques.

## 5 Hölder controls

The purpose of the section is to give suitable Hölder controls of the terms  $[D_{\mathbf{x}_1}^2 u(t, \cdot)]_\gamma$  and  $\sup_{z \in \mathbb{R}^d} [u(t, z, \cdot)]_{2+\gamma}$  in order to derive our main Schauder estimate of Theorem 1. For a fixed time  $t \in [0, T]$ , the point is to consider for fixed  $\mathbf{x} \in \mathbb{R}^{nd}$  the perturbative expansion (2.11) and for another spatial point  $\mathbf{x}' \in \mathbb{R}^{nd}$ , the possibly more refined version provided by equation (2.36) which precisely allows to take into account the various regimes depending on  $\mathbf{d}(\mathbf{x}, \mathbf{x}')$  and  $(s-t)^{1/2}$ ,  $s \in [t, T]$  detailed in the previous user's guide to the proof.

We will address separately two cases. For a constant  $c_0$  to be specified later on (but formally meant to be *small*), we consider:

- The globally *off-diagonal* regime  $T-t < c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ . In that case, the spatial points  $\mathbf{x}, \mathbf{x}'$  are globally *far* for the corresponding homogeneous distance  $\mathbf{d}$  over the time horizon  $s \in [t, T]$ . Hence, there is no specific need to exploit (2.36). Expanding the quantities  $D_{\mathbf{x}_1}^2 u(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u(t, \mathbf{x}')$  and  $u(t, z, \mathbf{x}_{2:n}) - u(t, z, \mathbf{x}'_{2:n})$  with (2.11) is enough to get the result.
- The *mixed* regime  $T-t \geq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ . In that case, up to the *transition time*  $t_0$  defined in (2.37),  $s-t < c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ , i.e. the *off-diagonal* regime holds for the times  $s$  close to  $t$ . It is then crucial to exploit the more refined perturbative expansion (2.36) to derive suitable bounds for  $D_{\mathbf{x}_1}^2 u(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u(t, \mathbf{x}')$  and  $u(t, (z, \mathbf{x}_{2:n})) - u(t, (z, \mathbf{x}'_{2:n}))$ . In particular, this leads to handle carefully the additional terms appearing from the change of freezing parameter.



Precisely, in the globally off-diagonal regime we first write from (2.11) for all  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$  s.t.  $(T - t) < c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ ,

$$\begin{aligned}
 & D_{\mathbf{x}_1}^2 u(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u(t, \mathbf{x}') \\
 &= \left( D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi} g(\mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi'} g(\mathbf{x}') + D_{\mathbf{x}_1}^2 \tilde{G}^{\xi} f(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{G}^{\xi'} f(t, \mathbf{x}') \right. \\
 & \quad \left. + \int_t^T ds \int_{\mathbb{R}^{nd}} \left( D_{\mathbf{x}_1}^2 \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y})(L_s - \tilde{L}_s^{\xi})u(s, \mathbf{y}) \right. \right. \\
 & \quad \quad \left. \left. - D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\xi'})u(s, \mathbf{y}) \right) d\mathbf{y} \right) \Bigg|_{(\xi, \xi') = (\mathbf{x}, \mathbf{x}')}, \quad (5.1)
 \end{aligned}$$

and similarly, for all  $z \in \mathbb{R}^d$ ,

$$\begin{aligned}
 & u(t, (z, \mathbf{x}_{2:n})) - u(t, (z, \mathbf{x}'_{2:n})) \\
 &= \left( \tilde{P}_{T,t}^{\xi} g(z, \mathbf{x}_{2:n}) - \tilde{P}_{T,t}^{\xi'} g(z, \mathbf{x}'_{2:n}) + \tilde{G}^{\xi} f(t, (z, \mathbf{x}_{2:n})) - \tilde{G}^{\xi'} f(t, (z, \mathbf{x}'_{2:n})) \right. \\
 & \quad \left. + \int_t^T ds \int_{\mathbb{R}^{nd}} \left( \tilde{p}^{\xi}(t, s, (z, \mathbf{x}_{2:n}), \mathbf{y})(L_s - \tilde{L}_s^{\xi})u(s, \mathbf{y}) \right. \right. \\
 & \quad \quad \left. \left. - \tilde{p}^{\xi'}(t, s, (z, \mathbf{x}'_{2:n}), \mathbf{y})(L_s - \tilde{L}_s^{\xi'})u(s, \mathbf{y}) \right) d\mathbf{y} \right) \Bigg|_{(\xi, \xi') = ((z, \mathbf{x}_{2:n}), (z, \mathbf{x}'_{2:n}))}. \quad (5.2)
 \end{aligned}$$

In this case, we derive from Lemmas 4, 5, 6 and equation (4.1):

$$\begin{aligned}
 & |D_{\mathbf{x}_1}^2 u(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u(t, \mathbf{x}')| \\
 &\leq C \left[ |D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi} g(\mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi'} g(\mathbf{x}')| + (\|f\|_{L^\infty(C_{b,d}^\gamma)} + \|u\|_{L^\infty(C_{b,d}^{2+\gamma})})(T - t)^{\frac{\gamma}{2}} \right] \\
 &\leq C \left[ |D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi} g(\mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi'} g(\mathbf{x}')| + (\|f\|_{L^\infty(C_{b,d}^\gamma)} + \|u\|_{L^\infty(C_{b,d}^{2+\gamma})})c_0^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \right], \\
 & |u(t, (z, \mathbf{x}_{2:n})) - u(t, (z, \mathbf{x}'_{2:n}))| \\
 &\leq C \left( |\tilde{P}_{T,t}^{\xi} g(z, \mathbf{x}_{2:n}) - \tilde{P}_{T,t}^{\xi'} g(z, \mathbf{x}'_{2:n})| + |\tilde{G}^{\xi} f(t, (z, \mathbf{x}_{2:n})) - \tilde{G}^{\xi'} f(t, (z, \mathbf{x}'_{2:n}))| \right. \\
 & \quad \left. + \|u\|_{L^\infty(C_{b,d}^{2+\gamma})}c_0^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \right).
 \end{aligned}$$

Therefore, these equations give the expected controls up to appropriate estimates for the Hölder moduli of the frozen semigroup and Green kernel which are obtained below, for the so-called *mixed* regime, on which we can focus.

In the *mixed* regime, we will start from identity (2.36) for the expansion of  $u(t, \mathbf{x}')$ ,

where we have chosen  $\xi' = \mathbf{x}'$  and  $\tilde{\xi}' = \mathbf{x}$ . Namely,

$$\begin{aligned} & u(t, \mathbf{x}') \\ &= \tilde{P}_{T,t}^{\tilde{\xi}'} g(\mathbf{x}') + \tilde{G}_{t_0,t}^{\xi'} f(t, \mathbf{x}') + \tilde{G}_{T,t_0}^{\tilde{\xi}'} f(t, \mathbf{x}') + \tilde{P}_{t_0,t}^{\xi'} u(t_0, \mathbf{x}') - \tilde{P}_{t_0,t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \\ & \quad + \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left( \mathbb{I}_{s \leq t_0} \tilde{p}^{\xi'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\xi'}) + \mathbb{I}_{s > t_0} p^{\tilde{\xi}'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\tilde{\xi}'}) \right) u(s, \mathbf{y}). \end{aligned}$$

Again,  $t_0$  must be here seen as a frozen parameter, which is *a posteriori*, i.e. after possible differentiation, chosen as in (2.37).

The terms to control then write for the Hölder norm of the derivatives w.r.t. the non-degenerate variables:

$$\begin{aligned} & D_{\mathbf{x}_1}^2 u(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 u(t, \mathbf{x}') \\ &= \left( \left( D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi} g(\mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi} g(\mathbf{x}') \right) + \left( D_{\mathbf{x}_1}^2 \tilde{G}_{t_0,t}^{\xi} f(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{G}_{t_0,t}^{\xi} f(t, \mathbf{x}') \right) \right. \\ & \quad + \left( D_{\mathbf{x}_1}^2 \tilde{G}_{T,t_0}^{\xi} f(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{G}_{T,t_0}^{\xi} f(t, \mathbf{x}') \right) + \left( D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\xi} u(t_0, \mathbf{x}') - D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\xi} u(t_0, \mathbf{x}') \right) \\ & \quad + \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left( \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y})(L_s - \tilde{L}_s^{\xi}) \right. \\ & \quad \left. - \left[ \mathbb{I}_{s \leq t_0} \tilde{p}^{\xi'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\xi'}) + \mathbb{I}_{s > t_0} p^{\xi}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\xi}) \right] u(s, \mathbf{y}) \right) \Big|_{(\xi, \xi') = (\mathbf{x}, \mathbf{x}')} \end{aligned} \quad (5.3)$$

Equation (5.2) would be adapted accordingly.

As a consequence of the previous equations, in order to derive the expected bounds we will devote a subsection to the Hölder controls for the frozen semi-group, for the frozen Green kernel, for the discontinuity term coming from the change of freezing point and for the perturbative contribution (remainder).

## 5.1 Hölder norms for the frozen semi-group

We precisely want to establish the following result.

**Lemma 8** *There exists  $C := C((\mathbf{A}))$  s.t. for all  $(t, \mathbf{x}, \mathbf{x}') \in [0, T] \times \mathbb{R}^{nd} \times \mathbb{R}^{nd}$ , taking*

$$(\xi, \xi') = \begin{cases} (\mathbf{x}, \mathbf{x}'), & \text{if } (T-t)^{1/2} < c_0 \mathbf{d}(\mathbf{x}, \mathbf{x}'), \\ (\mathbf{x}, \mathbf{x}), & \text{if } (T-t)^{1/2} \geq c_0 \mathbf{d}(\mathbf{x}, \mathbf{x}'), \end{cases}$$

one has:

$$\begin{aligned} |D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi}(\mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi'}(\mathbf{x}')| &\leq C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'), \\ |\tilde{P}_{T,t}^{\xi}(\mathbf{x}) - \tilde{P}_{T,t}^{\xi'}(\mathbf{x}')| &\leq C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'), \text{ for } \mathbf{x}_1 = \mathbf{x}'_1. \end{aligned}$$

**Hölder norms of the derivatives w.r.t. the non-degenerate variables**

Let us deal with the first inequality of Lemma 8, i.e. the norms over the non-degenerate variables  $\mathbf{x}_1$ . For the frozen semi-group, we say that the off-diagonal regime (resp. diagonal regime) holds when  $T - t \leq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  (resp.  $T - t \geq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ ).

- Off-diagonal regime. If  $T - t \leq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ , like in (4.23), we write:

$$\begin{aligned}
 & D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^\xi g(\mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi'} g(\mathbf{x}') \\
 = & \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) [g(\mathbf{y}) - g(\mathbf{y}_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n})] d\mathbf{y} \right. \\
 & - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, T, \mathbf{x}', \mathbf{y}) [g(\mathbf{y}) - g(\mathbf{y}_1, \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')_{2:n})] d\mathbf{y} \\
 & + \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) [g(\mathbf{y}_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n}) - g(\mathbf{m}_{T,t}^\xi(\mathbf{x}))] d\mathbf{y} \right. \\
 & \left. - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, T, \mathbf{x}', \mathbf{y}) [g(\mathbf{y}_1, \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')_{2:n}) - g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))] d\mathbf{y} \right] \\
 =: & \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_1 g(\mathbf{x}, \mathbf{x}') + \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_2 g(\mathbf{x}, \mathbf{x}'). \tag{5.4}
 \end{aligned}$$

The first term, which is associated with the degenerate variables, is controlled directly thanks to (4.24), which again readily follows from Proposition 2 (see as well Remark 3.8), for  $(\xi, \xi') = (\mathbf{x}, \mathbf{x}')$ . One readily gets:

$$\left| \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_1 g(\mathbf{x}, \mathbf{x}') \right|_{\xi=\mathbf{x}} \leq 2C(T-t)^{\frac{\gamma}{2}} \|g\|_{C_{b,d}^{2+\gamma}} \leq 2Cc_0^{\frac{\gamma}{2}} \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{5.5}$$

The second term is more delicate, we proceed like in (4.25):

$$\begin{aligned}
 & \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_2 g(\mathbf{x}, \mathbf{x}') \\
 = & \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) \int_0^1 d\mu(1-\mu) \right. \\
 & \text{Tr} \left( \left[ D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x})_1 + \mu(\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1, \mathbf{m}_{T,t}^\xi(\mathbf{x})_{2:n}) - D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x})) \right] \right. \\
 & \left. \left. (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1^{\otimes 2} \right) d\mathbf{y} \right. \\
 & - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, T, \mathbf{x}', \mathbf{y}) \int_0^1 d\mu(1-\mu) \\
 & \text{Tr} \left( \left[ D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')_1 + \mu(\mathbf{y} - \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))_1, \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')_{2:n}) - D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')) \right] \right. \\
 & \left. \left. (\mathbf{y} - \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))_1^{\otimes 2} \right) d\mathbf{y} \right] \\
 + & \left[ \frac{1}{2} \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x})) (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1^{\otimes 2} \right) d\mathbf{y} \right. \\
 & \left. - \frac{1}{2} \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, T, \mathbf{x}', \mathbf{y}) \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')) (\mathbf{y} - \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))_1^{\otimes 2} \right) d\mathbf{y} \right] \\
 =: & \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_{21} g(\mathbf{x}, \mathbf{x}') + \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_{22} g(\mathbf{x}, \mathbf{x}'). \tag{5.6}
 \end{aligned}$$

The first contribution of the previous identity is handled exploiting the smoothness of  $D_{\mathbf{x}_1}^2 g$  and Proposition 2. Namely,

$$\begin{aligned} |\Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_{21} g(\mathbf{x}, \mathbf{x}')| &\leq C [D_{\mathbf{x}_1}^2 g]_{\mathbf{d}}^\gamma \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(T-t)} \left( \tilde{p}_{C^{-1}}^\xi(t, T, \mathbf{x}, \mathbf{y}) |(\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1|^{2+\gamma} \right. \\ &\quad \left. + \tilde{p}_{C^{-1}}^{\xi'}(t, T, \mathbf{x}', \mathbf{y}) |(\mathbf{y} - \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))_1|^{2+\gamma} \right) \Big|_{(\xi,\xi')=(\mathbf{x},\mathbf{x}')}. \end{aligned}$$

Hence,

$$\left| \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_{21} g(\mathbf{x}, \mathbf{x}') \right|_{(\xi,\xi')=(\mathbf{x},\mathbf{x}')} \leq 2C(T-t)^{\frac{\gamma}{2}} \|g\|_{C_{b,\mathbf{d}}^{2+\gamma}} \leq 2C c_0^{\frac{\gamma}{2}} \|g\|_{C_{b,\mathbf{d}}^{2+\gamma}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \quad (5.7)$$

Let us now decompose the last contribution of (5.6):

$$\begin{aligned} &\left| \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_{22} g(\mathbf{x}, \mathbf{x}') \right|_{(\xi,\xi')=(\mathbf{x},\mathbf{x}')} \\ &\leq \frac{1}{2} \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(s-t)} \tilde{p}_{C^{-1}}^\xi(t, T, \mathbf{x}, \mathbf{y}) |D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x})) - D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))| |(\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1|^2 \\ &\quad + \frac{1}{2} \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')) (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}))_1^{\otimes 2} \right) \right. \\ &\quad \left. - D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, T, \mathbf{x}', \mathbf{y}) \text{Tr} \left( \langle D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')) (\mathbf{y} - \mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))_1^{\otimes 2} \rangle \right) d\mathbf{y} \right| \\ &\leq C |D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x})) - D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))| = C |D_{\mathbf{x}_1}^2 g(\boldsymbol{\theta}_{T,t}(\mathbf{x})) - D_{\mathbf{x}_1}^2 g(\boldsymbol{\theta}_{T,t}(\mathbf{x}'))|, \end{aligned} \quad (5.8)$$

exploiting Proposition 2 and equation (3.16) in Proposition 3 to observe that the second contribution of the first inequality above vanishes and recalling as well (2.13) for the last equality to identify the linearized flows, respectively frozen in  $\boldsymbol{\xi} = \mathbf{x}, \boldsymbol{\xi}' = \mathbf{x}'$ , with the initial non-linear ones.

From Lemma 2, we derive that for  $T-t \leq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ :

$$\left| \Delta_{t,T,\xi,\xi'} D_{\mathbf{x}_1}^2 \tilde{P}_{22} g(\mathbf{x}, \mathbf{x}') \right|_{(\xi,\xi')=(\mathbf{x},\mathbf{x}')} \leq C \|g\|_{C_{b,\mathbf{d}}^{2+\gamma}} (\mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') + (T-t)^{\frac{\gamma}{2}}). \quad (5.9)$$

Plugging (5.5), (5.7), (5.9) into (5.4) yields the result.

- Diagonal regime. If  $T-t > c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ , we directly write:

$$\begin{aligned} &|D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^\xi g(\mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\xi'} g(\mathbf{x}')| \\ &\leq \left| \int_{\mathbb{R}^{nd}} [D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, T, \mathbf{x}', \mathbf{y})] g(\mathbf{y}) d\mathbf{y} \right| \\ &\leq \sum_{k=1}^n \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}')_k g(\mathbf{y}) d\mathbf{y} \right| \\ &=: \Delta_{t,T,\xi,\xi'} \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{P}_3 g(\mathbf{x}, \mathbf{x}'), \end{aligned} \quad (5.10)$$

recalling the notation  $\mathbf{D} = (D_{\mathbf{x}_1}, \dots, D_{\mathbf{x}_n})$  introduced after equation (1.1). This contribution is dealt through the cancellation tools of Proposition 3 (see equations (3.15), (3.17)). We get:

$$\begin{aligned}
 & \Delta_{t,T,\xi,\xi'} \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{P}_3 g(\mathbf{x}, \mathbf{x}') \\
 & \leq \sum_{k=1}^n \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, T, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'), \mathbf{y}) \cdot (\mathbf{x}_k - \mathbf{x}'_k) \left[ g(\mathbf{y}) - g(\mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \right) \right. \\
 & \quad - \langle D_{\mathbf{x}_1} g(\mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'))), (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \rangle_1 \\
 & \quad \left. - \frac{1}{2} \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \rangle_1^{\otimes 2} \right) \right] d\mathbf{y} \Big|. \tag{5.11}
 \end{aligned}$$

Because  $g \in C_{b,\mathbf{d}}^{2+\gamma}(\mathbb{R}^{nd}, \mathbb{R})$ , we readily deduce, reproducing the Taylor expansion on  $g$  employed for equations (5.4)-(5.6) above, that:

$$\begin{aligned}
 & \left| g(\mathbf{y}) - g(\mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \right. \\
 & \quad - \langle D_{\mathbf{x}_1} g(\mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'))), (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \rangle_1 \\
 & \quad \left. - \frac{1}{2} \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) (\mathbf{y} - \mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \rangle_1^{\otimes 2} \right) \right| \Big|_{\xi=\mathbf{x}} \\
 & \leq \|g\|_{C_{b,\mathbf{d}}^{2+\gamma}} \mathbf{d}^{2+\gamma}(\mathbf{y}, \mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \Big|_{\xi=\mathbf{x}}. \tag{5.12}
 \end{aligned}$$

Plugging this inequality into (5.11) yields:

$$\begin{aligned}
 & \Delta_{t,T,\xi,\xi'} \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{P}_3 g(\mathbf{x}, \mathbf{x}') \Big|_{\xi=\mathbf{x}} \\
 & \leq C \|g\|_{L^\infty(C_{b,\mathbf{d}}^{2+\gamma})} \sum_{k=1}^n \int_{\mathbb{R}^{nd}} (T-t)^{-1-(k-\frac{1}{2})} \tilde{p}_{C-1}^\xi(t, T, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'), \mathbf{y}) |\mathbf{x}_k - \mathbf{x}'_k| \\
 & \quad \times \mathbf{d}^{2+\gamma}(\mathbf{y}, \mathbf{m}_{T,t}^\xi(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \Big|_{\xi=\mathbf{x}} \\
 & \leq C \|g\|_{L^\infty(C_{b,\mathbf{d}}^{2+\gamma})} \sum_{k=1}^n (T-t)^{-(k-\frac{1}{2})+\frac{\gamma}{2}} |(\mathbf{x} - \mathbf{x}')_k| \leq C \|g\|_{C^{2+\gamma}, \mathbf{d}^\gamma}(\mathbf{x}, \mathbf{x}'), \tag{5.13}
 \end{aligned}$$

using (3.8) for the second inequality and recalling that, since  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') < (T-t)$ , we indeed have  $(T-t)^{-(k-\frac{1}{2})+\frac{\gamma}{2}} |(\mathbf{x} - \mathbf{x}')_k| \leq (c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}'))^{-(k-\frac{1}{2})+\frac{\gamma}{2}} \mathbf{d}^{2k-1}(\mathbf{x}, \mathbf{x}') \leq C \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')$ .

### Hölder control for the degenerate variables

Of course, the analysis of the Hölder regularity of the frozen semi-group w.r.t. the degenerate variables is also based on the previous techniques. We still take advantage of cancellation tools. For the whole paragraph we consider two arbitrary given spatial points  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$  s.t.  $\mathbf{x}_1 = \mathbf{x}'_1$ , i.e. their first entry, corresponding to the non-degenerate variable, coincide.

• Off-diagonal regime. If  $T - t \leq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ , we proceed to an expansion similar to (5.4) for  $D_{\mathbf{x}_1}^2 \tilde{P}_{T,t}^{\mathbf{x}} g(\mathbf{x})$ . In particular, with the notations introduced in (5.4), we write:

$$\tilde{P}_{T,t}^{\xi} g(\mathbf{x}) - \tilde{P}_{T,t}^{\xi'} g(\mathbf{x}') =: \Delta_{t,T,\xi,\xi'} \tilde{P}_1 g(\mathbf{x}, \mathbf{x}') + \Delta_{t,T,\xi,\xi'} \tilde{P}_2 g(\mathbf{x}, \mathbf{x}'). \quad (5.14)$$

We directly obtain from the Proposition 2, similarly to (5.5), that:

$$\left| \Delta_{t,T,\xi,\xi'} \tilde{P}_1 g(\mathbf{x}, \mathbf{x}') \right|_{\xi=\mathbf{x}} \leq 2C(T-t)^{\frac{2+\gamma}{2}} \|g\|_{C_{b,d}^{2+\gamma}} \leq 2C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \quad (5.15)$$

We indeed recall that the difference w.r.t. (5.5) is that we do not have anymore the time-singularities coming therein from the spatial derivatives.

With the notations of (5.6), the second contribution of (5.14) writes:

$$\Delta_{t,T,\xi,\xi'} \tilde{P}_2 g(\mathbf{x}, \mathbf{x}') = \Delta_{t,T,\xi,\xi'} \tilde{P}_{21} g(\mathbf{x}, \mathbf{x}') + \Delta_{t,T,\xi,\xi'} \tilde{P}_{22} g(\mathbf{x}, \mathbf{x}').$$

Proposition 2 again yields, similarly to (5.7), that:

$$\left| \Delta_{t,T,\xi,\xi'} \tilde{P}_{21} g(\mathbf{x}, \mathbf{x}') \right|_{(\xi,\xi')=(\mathbf{x},\mathbf{x}')} \leq 2C(T-t)^{\frac{2+\gamma}{2}} \|g\|_{C_{b,d}^{2+\gamma}} \leq 2C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \quad (5.16)$$

On the other hand, we readily get from Proposition 3 that:

$$\Delta_{t,T,\xi,\xi'} \tilde{P}_{22} g(\mathbf{x}, \mathbf{x}') = \frac{1}{2} \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi}(\mathbf{x})) [\tilde{\mathbf{K}}_{T,t}^{\xi}]_{1,1} - D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')) [\tilde{\mathbf{K}}_{T,t}^{\xi'}]_{1,1} \right).$$

Write now:

$$\begin{aligned} \Delta_{t,T,\xi,\xi'} \tilde{P}_{22} g(\mathbf{x}, \mathbf{x}') &= \frac{1}{2} \text{Tr} \left( [D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi}(\mathbf{x})) - D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))] [\tilde{\mathbf{K}}_{T,t}^{\xi}]_{1,1} \right) \\ &\quad + \frac{1}{2} \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')) [[\tilde{\mathbf{K}}_{T,t}^{\xi}]_{1,1} - [\tilde{\mathbf{K}}_{T,t}^{\xi'}]_{1,1}] \right). \end{aligned} \quad (5.17)$$

Since  $T - t \leq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ , recalling as well that for  $\xi = \mathbf{x}, \xi' = \mathbf{x}'$ ,  $\mathbf{m}_{T,t}^{\xi}(\mathbf{x}) = \boldsymbol{\theta}_{T,t}(\mathbf{x})$ ,  $\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}') = \boldsymbol{\theta}_{T,t}(\mathbf{x}')$ , we readily deduce from Proposition 1 and Lemma 2 that:

$$\begin{aligned} &\frac{1}{2} \left| [D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi}(\mathbf{x})) - D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}'))] [\tilde{\mathbf{K}}_{T,t}^{\xi}]_{1,1} \right| \\ &\leq C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^{\gamma}(\boldsymbol{\theta}_{T,t}(\mathbf{x}), \boldsymbol{\theta}_{T,t}(\mathbf{x}')) (T-t) \\ &\leq C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \end{aligned} \quad (5.18)$$

For the last contribution, we directly obtain from Lemma 1 (equation (3.18) for  $j = 1$ ):

$$\begin{aligned} &\left| \frac{1}{2} D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi'}(\mathbf{x}')) ([\tilde{\mathbf{K}}_{T,t}^{\xi}]_{1,1} - [\tilde{\mathbf{K}}_{T,t}^{\xi'}]_{1,1}) \right|_{(\xi,\xi')=(\mathbf{x},\mathbf{x}')} \\ &\leq C \|g\|_{C_{b,d}^{2+\gamma}} \left( (T-t)^{\frac{2+\gamma}{2}} + (T-t) \mathbf{d}^{\gamma}(\mathbf{x}, \mathbf{x}') \right) \\ &\leq C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'), \end{aligned} \quad (5.19)$$

using again that  $T - t \leq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  for the last inequality. Plugging (5.18) and (5.19) into (5.17) gives the expected bound.

- Diagonal regime. If  $T - t \geq c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ , we write:

$$\begin{aligned}
 & |\tilde{P}_{T,t}^{\mathbf{x}} g(\mathbf{x}) - \tilde{P}_{T,t}^{\mathbf{x}'} g(\mathbf{x}')| \\
 & \leq \left| \int_{\mathbb{R}^{nd}} [\tilde{p}^{\xi}(t, T, \mathbf{x}, \mathbf{y}) - \tilde{p}^{\xi}(t, T, \mathbf{x}', \mathbf{y})] g(\mathbf{y}) d\mathbf{y} \right|_{\xi=\mathbf{x}} \\
 & \leq \left| \int_0^1 d\mu \int_{\mathbb{R}^{nd}} \langle \mathbf{D} \tilde{p}^{\xi}(t, T, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'), \mathbf{y}), (\mathbf{x} - \mathbf{x}') \rangle \left[ g(\mathbf{y}) - g(\mathbf{m}_{T,t}^{\xi}(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'))) \right] \right. \\
 & \quad \left. - \langle D_{\mathbf{x}_1} g(\mathbf{m}_{T,t}^{\xi}(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'))), (\mathbf{y} - \mathbf{m}_{T,t}^{\xi}(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')))_1 \rangle \right. \\
 & \quad \left. - \frac{1}{2} \text{Tr} \left( D_{\mathbf{x}_1}^2 g(\mathbf{m}_{T,t}^{\xi}(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')))(\mathbf{y} - \mathbf{m}_{T,t}^{\xi}(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')))_1^{\otimes 2} \right) \right] d\mathbf{y} \Big|_{\xi=\mathbf{x}},
 \end{aligned}$$

with the same cancellation argument as in (5.10)-(5.11). Observe anyhow that the cancellation involving the gradient in the above equation is possible precisely because  $\mathbf{x}_1 = \mathbf{x}'_1$  and therefore  $D \tilde{p}^{\xi}(t, T, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}') = D_{\mathbf{x}_{2:n}} \tilde{p}^{\xi}(t, T, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}')_{2:n}$ . We then obtain thanks to the previous identity and (5.12):

$$\begin{aligned}
 & |\tilde{P}_{T,t}^{\mathbf{x}} g(\mathbf{x}) - \tilde{P}_{T,t}^{\mathbf{x}'} g(\mathbf{x}')| \\
 & \leq \|g\|_{C_{b,d}^{2+\gamma}} \sum_{k=2}^n \int_{\mathbb{R}^{nd}} |D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, T, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'), \mathbf{y})| |\mathbf{x}_k - \mathbf{x}'_k| \mathbf{d}^{2+\gamma}(\mathbf{y}, \mathbf{m}_{T,t}^{\xi}(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}'))) d\mathbf{y} \\
 & \leq \|g\|_{C_{b,d}^{2+\gamma}} \sum_{k=2}^n \int_{\mathbb{R}^{nd}} \frac{C}{(T-t)^{1+(k-\frac{1}{2})-\frac{2+\gamma}{2}}} \tilde{p}_{C^{-1}}^{\mathbf{x}}(t, T, \mathbf{m}_{T,t}^{\mathbf{x}}(\mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')), \mathbf{y}) |\mathbf{x}_k - \mathbf{x}'_k| d\mathbf{y} \\
 & \leq C \|g\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'), \tag{5.20}
 \end{aligned}$$

reproducing the arguments used to establish (5.13) for the last inequality. Lemma 8 is proved.

## 5.2 Hölder norms associated with the Green kernel

Let us recall that in (2.36), for a source  $f \in L^\infty([0, T], C_{b,d}^\gamma(\mathbb{R}^{nd}, \mathbb{R}))$ , we have to control the Hölder norms of the Green kernel which we split into two parts according to the position of the time integration variable w.r.t. the change of regime time  $t_0$  (see (2.37)) *a posteriori* chosen to be  $t_0 := (t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')) \wedge T$ . This is again the splitting according to the off-diagonal and diagonal regime. The point is that for the Green kernel, if  $t_0 < T$  both regimes appear.

**Lemma 9** *Under (A), for fixed spatial points  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ , we have that there exists*

a constant  $C := C((\mathbf{A}), T)$ , s.t. for all  $f \in L^\infty([0, T], C_{b,d}^\gamma(\mathbb{R}^{nd}, \mathbb{R}))$ :

$$\begin{aligned} & \sup_{t \in [0, T]} (|D_{\mathbf{x}_1}^2 \tilde{G}_{t_0, t}^\xi f(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{G}_{t_0, t}^{\xi'} f(t, \mathbf{x}')| + |D_{\mathbf{x}_1}^2 \tilde{G}_{T, t_0}^\xi f(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{G}_{T, t_0}^{\xi'} f(t, \mathbf{x}')|) \\ & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'), \end{aligned}$$

and

$$\begin{aligned} & \sup_{t \in [0, T]} (|\tilde{G}_{t_0, t}^\xi f(t, \mathbf{x}) - \tilde{G}_{t_0, t}^{\xi'} f(t, \mathbf{x}')| + |\tilde{G}_{T, t_0}^\xi f(t, \mathbf{x}) - \tilde{G}_{T, t_0}^{\xi'} f(t, \mathbf{x}')|) \\ & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'), \text{ if } \mathbf{x}_1 = \mathbf{x}'_1, \end{aligned}$$

where  $\xi = \mathbf{x}$ ,  $\xi' = \mathbf{x}'$ ,  $\tilde{\xi}' = \mathbf{x}$ .

We now turn to the proof of Lemma 9.

**Proof: Lemma 9**

Let us begin with the statement concerning the second order derivatives of the frozen Green kernel w.r.t. the non-degenerate variable  $\mathbf{x}_1$ .

For the *off-diagonal regime*, involving the term  $D_{\mathbf{x}_1}^2 \tilde{G}_{t_0, t}^{\xi'} f(t, \mathbf{x}')$ , we readily get from Lemma 7 that

$$\begin{aligned} & \left| D_{\mathbf{x}_1}^2 \tilde{G}_{t_0, t}^\xi f(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{G}_{t_0, t}^{\xi'} f(t, \mathbf{x}') \right|_{(\xi, \xi') = (\mathbf{x}, \mathbf{x}')} \\ & \leq \left| \int_t^{t_0} ds D_{\mathbf{x}_1}^2 \tilde{P}_{s, t}^\xi f(s, \mathbf{x}) \right|_{\xi = \mathbf{x}} + \left| \int_t^{t_0} ds D_{\mathbf{x}_1}^2 \tilde{P}_{s, t}^{\xi'} f(s, \mathbf{x}') \right|_{\xi' = \mathbf{x}'} \\ & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \int_t^{t_0} ds (s-t)^{-1+\frac{\gamma}{2}} \\ & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \end{aligned} \tag{5.21}$$

For the *diagonal regime*, involving the term  $D_{\mathbf{x}_1}^2 \tilde{G}_{T, t_0}^{\tilde{\xi}'} f(t, \mathbf{x}')$ , we have to be more subtle and perform again a Taylor expansion of  $D_{\mathbf{x}_1}^2 \tilde{P}_{s, t}^\xi f(s, \cdot)$ . Namely:

$$\begin{aligned} & \left| D_{\mathbf{x}_1}^2 \tilde{G}_{T, t_0}^\xi f(t, \mathbf{x}) - D_{\mathbf{x}_1}^2 \tilde{G}_{T, t_0}^{\tilde{\xi}'} f(t, \mathbf{x}') \right|_{(\xi, \tilde{\xi}') = (\mathbf{x}, \mathbf{x})} \\ & \leq \int_{t_0}^T ds \left| \int_0^1 d\mu \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{P}_{s, t}^\xi f(s, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')) \cdot (\mathbf{x} - \mathbf{x}') \right|_{\xi = \mathbf{x}} \\ & \leq \sum_{i=1}^n |(\mathbf{x} - \mathbf{x}')_i| \int_{t_0}^T ds \int_0^1 d\mu |D_{\mathbf{x}_i} D_{\mathbf{x}_1}^2 \tilde{P}_{s, t}^\xi f(s, \mathbf{x}' + \lambda(\mathbf{x} - \mathbf{x}'))|_{\xi = \mathbf{x}} \\ & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \sum_{i=1}^n |(\mathbf{x} - \mathbf{x}')_i| \int_{t_0}^T ds (s-t)^{-1-(i-\frac{1}{2})+\frac{\gamma}{2}}, \end{aligned}$$

using again Lemma 7 and the arguments of (5.13) for the last inequality. This finally



yields, recalling the definition of  $\mathbf{d}$  in (1.9):

$$\begin{aligned}
 & \left| D_{\mathbf{x}_1}^2 \tilde{G}_{T,t_0}^{\xi} f(t, \mathbf{x}) - D_{\mathbf{x}'_1}^2 \tilde{G}_{T,t_0}^{\xi'} f(t, \mathbf{x}') \right|_{(\xi, \xi') = (\mathbf{x}, \mathbf{x}')} \\
 & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \sum_{i=1}^n |(\mathbf{x} - \mathbf{x}')_i| (\mathbf{d}^2(\mathbf{x}, \mathbf{x}'))^{-(i-\frac{1}{2})+\frac{\gamma}{2}} \\
 & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{5.22}
 \end{aligned}$$

Gathering (5.21) and (5.22) gives the result.

Let us now turn to the Hölder controls on the degenerate variables. The idea is here again to perform a Taylor expansion at order one for  $\mathbf{x}_1 = \mathbf{x}'_1$ . Namely, from Lemma 7, the diagonal control is direct. We get

$$\begin{aligned}
 & \left| \tilde{G}_{T,t_0}^{\xi} f(t, \mathbf{x}) - \tilde{G}_{T,t_0}^{\xi'} f(t, \mathbf{x}') \right|_{(\xi, \xi') = (\mathbf{x}, \mathbf{x}')} \\
 & \leq \int_{t_0}^T ds \int_0^1 d\mu \left| \langle \mathbf{D} \tilde{P}_{s,t}^{\xi} f(s, \mathbf{x}' + \mu(\mathbf{x} - \mathbf{x}')), (\mathbf{x} - \mathbf{x}') \rangle \right| \\
 & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \int_{(t+c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')) \wedge T}^T ds \sum_{i=2}^n |(\mathbf{x} - \mathbf{x}')_i| (s-t)^{-(i-\frac{1}{2})+\frac{\gamma}{2}} \\
 & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \sum_{i=2}^n \underbrace{|(\mathbf{x} - \mathbf{x}')_i|}_{\leq \mathbf{d}(\mathbf{x}, \mathbf{x}')^{2i-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}')^{2-(2i-1)+\gamma} \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \tag{5.23}
 \end{aligned}$$

Now, for the off-diagonal bound, associated with the term  $\tilde{G}_{t_0,t}^{\xi} f(t, \cdot)$ , we precisely need to exploit the smoothness of  $f$  associated to the fact that the semi-group  $\tilde{P}_{s,t}^{\xi}$  has a density. Indeed, we cannot take advantage of the cancellation tools of Lemma 7, but we have for all  $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{nd}$  s.t.  $\mathbf{x}_1 = \mathbf{x}'_1$ :

$$\begin{aligned}
 & \left| \int_t^{t_0} ds \left( \tilde{P}_{s,t}^{\xi} f(s, \mathbf{x}) - \tilde{P}_{s,t}^{\xi'} f(s, \mathbf{x}') \right) \right|_{\xi = \mathbf{x}} \\
 & \leq \left| \int_t^{t_0} ds \tilde{P}_{s,t}^{\xi} f(s, \mathbf{x}) - f(s, \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x})) \right|_{\xi = \mathbf{x}} + \left| \int_t^{t_0} ds \tilde{P}_{s,t}^{\xi'} f(s, \mathbf{x}') - f(s, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) \right|_{\xi' = \mathbf{x}'} \\
 & \quad + \left| \int_t^{t_0} ds f(s, \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x})) - f(s, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) \right|_{(\xi, \xi') = (\mathbf{x}, \mathbf{x}')} \tag{5.24}
 \end{aligned}$$

Note that the first two terms in the r.h.s. of inequality (5.24) are handled like in the previous section. Precisely, we write for the first contribution:

$$\begin{aligned}
 & \left| \int_t^{t_0} ds \tilde{P}_{s,t}^{\xi} f(s, \mathbf{x}) - f(s, \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x})) \right|_{\xi = \mathbf{x}} \\
 & = \left| \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} \tilde{p}^{\xi}(t, T, \mathbf{x}, \mathbf{y}) [f(s, \mathbf{y}) - f(s, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}))] \right|_{\xi = \mathbf{x}} \\
 & \leq \|f\|_{L^\infty(C_{b,d}^\gamma)} \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} \tilde{p}^{\xi}(t, T, \mathbf{x}, \mathbf{y}) \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{t_0,t}(\mathbf{x})) \right|_{\xi = \mathbf{x}} \\
 & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \int_t^{t_0} ds (s-t)^{\frac{\gamma}{2}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'),
 \end{aligned}$$

by definition of  $t_0$  in (2.37). The second term of (5.24) is handled similarly. We thus obtain:

$$\begin{aligned} & \left| \int_t^{t_0} ds \tilde{P}_{s,t}^{\xi} f(s, \mathbf{x}) - f(s, \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x})) \right|_{\xi=\mathbf{x}} + \left| \int_t^{t_0} ds \tilde{P}_{s,t}^{\xi'} f(s, \mathbf{x}') - f(s, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) \right|_{\xi'=\mathbf{x}'} \\ & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \end{aligned} \quad (5.25)$$

For the last contribution in (5.24), we have directly that:

$$\left| \int_t^{t_0} ds f(s, \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x})) - f(s, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) \right|_{(\xi, \xi')=(\mathbf{x}, \mathbf{x}')} \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \int_t^{t_0} ds \mathbf{d}^\gamma(\boldsymbol{\theta}_{t_0,t}(\mathbf{x}), \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')). \quad (5.26)$$

Lemma 2 and (5.26) eventually yield:

$$\begin{aligned} & \left| \int_t^{t_0} ds f(s, \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x})) - f(s, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) \right|_{(\xi, \xi')=(\mathbf{x}, \mathbf{x}')} \\ & \leq C \|f\|_{L^\infty(C_{b,d}^\gamma)} \left( (t_0 - t)^{1+\frac{\gamma}{2}} + (t_0 - t) \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \right) \\ & = 2C \|f\|_{L^\infty(C_{b,d}^\gamma)} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \end{aligned}$$

□

### 5.3 Hölder norms of the perturbative contribution

To derive the spatial Hölder continuity of the perturbative term in (5.3) (see also the splitting of equation (2.28) of the user's guide) we proceed with the previous regime separation successfully used to establish in the previous sections the regularity of the semi-group and the Green kernel. Namely, we recall that the critical time giving the change of regime is (chosen after potential differentiation)  $t_0 = t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \wedge T$ . We can assume here w.l.o.g. that  $t_0 < T$  (otherwise there is a globally off-diagonal regime and the analysis becomes easier).

The point of the current section is that we also precisely need to adapt the duality analysis on Besov spaces already exploited for the supremum norm of the spatial derivatives w.r.t. the non-degenerate variable (see Section 4.2).

- Control of  $[D_{\mathbf{x}_1}^2 u(t, \cdot)]_\gamma$ . From (2.28) and (2.36), let us recall the following notations:

$$\Delta^{\xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') = \Delta_{\text{diag}}^{\xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') + \Delta_{\text{off-diag}}^{\xi, \xi'}(t, \mathbf{x}, \mathbf{x}'),$$

with

$$\begin{aligned} \Delta_{\text{off-diag}}^{\xi, \xi'}(t, \mathbf{x}, \mathbf{x}') &= \int_t^T ds \int_{\mathbb{R}^{nd}} dy \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \mathbb{I}_{s \leq t_0} (L_s - \tilde{L}_s^\xi) u(s, \mathbf{y}) \\ &\quad - \int_t^T ds \int_{\mathbb{R}^{nd}} dy \tilde{p}^{\xi'}(t, s, \mathbf{x}', \mathbf{y}) \mathbb{I}_{s \leq t_0} (L_s - \tilde{L}_s^{\xi'}) u(s, \mathbf{y}), \\ \Delta_{\text{diag}}^{\xi, \xi'}(t, T, \mathbf{x}, \mathbf{x}') &= \int_t^T ds \int_{\mathbb{R}^{nd}} dy \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \mathbb{I}_{s > t_0} (L_s - \tilde{L}_s^\xi) u(s, \mathbf{y}) \\ &\quad - \int_t^T ds \int_{\mathbb{R}^{nd}} dy \tilde{p}^{\xi'}(t, s, \mathbf{x}', \mathbf{y}) \mathbb{I}_{s > t_0} (L_s - \tilde{L}_s^{\xi'}) u(s, \mathbf{y}). \end{aligned} \quad (5.27)$$

We recall that in the diagonal regime, i.e.  $c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \leq s - t$ , we choose  $(\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}}') = (\mathbf{x}, \mathbf{x}')$ . The idea is to expand the density  $\tilde{p}^\xi(t, s, \cdot, \mathbf{y})$  between  $\mathbf{x}$  and  $\mathbf{x}'$ . The associated time singularity can then be bounded by a space singularity, recall that on the considered regime  $(s - t)^{-1} \leq (c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}'))^{-1}$  which will in turn be exactly compensated by the difference of spatial points induced by the Taylor expansion and the Hölder regularity of the coefficients.

We decompose from definitions (2.28) and (2.36):

$$\begin{aligned}
 & \left| D_{\mathbf{x}_1}^2 \Delta_{\text{diag}}^{\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}}'}(t, T, \mathbf{x}, \mathbf{x}') \right|_{(\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}}') = (\mathbf{x}, \mathbf{x}')} \\
 \leq & \left| \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left[ D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) \right] \left[ \langle \mathbf{F}_1(s, \mathbf{y}) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})), D_{\mathbf{y}_1} \rangle \right. \right. \\
 & \left. \left. + \frac{1}{2} \text{Tr} \left( (a(s, \mathbf{y}) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) D_{\mathbf{y}_1}^2 \right) \right] u(s, \mathbf{y}) \mathbb{I}_{s > t_0} \right|_{\boldsymbol{\xi} = \mathbf{x}} \\
 & + \left| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left[ D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) \mathbb{I}_{s > t_0} \right] \right. \\
 & \left. \times \left\langle \mathbf{F}_i(s, \mathbf{y}) - [\mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) + D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}], D_{\mathbf{y}_i} u(s, \mathbf{y}) \right\rangle \mathbb{I}_{s > t_0} \right|_{\boldsymbol{\xi} = \mathbf{x}},
 \end{aligned}$$

which readily yields with the notations of (4.3) that:

$$\begin{aligned}
 & \left| D_{\mathbf{x}_1}^2 \Delta_{\text{diag}}^{\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}}'}(t, T, \mathbf{x}, \mathbf{x}') \right|_{(\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}}') = (\mathbf{x}, \mathbf{x}')} \\
 \leq & \left| \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}') \Delta_{1, \mathbf{F}, \sigma}(t, s, \mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), u) \mathbb{I}_{s > t_0} \right|_{\boldsymbol{\xi} = \mathbf{x}} \\
 & + \left| \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}') \right. \\
 & \left. \times \left\langle \Delta_{i, \mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \right\rangle \mathbb{I}_{s > t_0} \right|_{\boldsymbol{\xi} = \mathbf{x}} \\
 =: & \left| D_{\mathbf{x}_1}^2 \Delta_{\text{diag}}^1(t, T, \mathbf{x}, \mathbf{x}') \right| + \left| D_{\mathbf{x}_1}^2 \Delta_{\text{diag}}^{2:n}(t, T, \mathbf{x}, \mathbf{x}') \right|. \tag{5.28}
 \end{aligned}$$

We will now control the first term of the above right hand side. In other words, we specify the control of (2.14). We obtain directly thanks to the smoothness assumption **(S)** on the coefficients and Proposition 2 (see also equation (3.8)) that for all  $k \in \llbracket 1, n \rrbracket$ :

$$\begin{aligned}
 & \left| \mathbf{F}_1(s, \mathbf{y}) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right|_{D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y})} \Big|_{\boldsymbol{\xi} = \mathbf{x}} \\
 \leq & C \left( \left\| \mathbf{F}_1(s, \cdot) \right\|_{C_{\mathbf{d}}^\gamma} \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right) \times (s - t)^{-1 - (k - \frac{1}{2})} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) \Big|_{\boldsymbol{\xi} = \mathbf{x}} \\
 \leq & C \left\| \mathbf{F}_1 \right\|_{L^\infty(C_{\mathbf{d}}^\gamma)} (s - t)^{-1 - (k - \frac{1}{2}) + \gamma/2} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}). \tag{5.29}
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 & |a(s, \mathbf{y}) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))| D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & \leq C \left( \|a(s, \cdot)\|_{C_d^\gamma} \mathbf{d}^\gamma(\mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right) \times (s-t)^{-1-(k-\frac{1}{2})} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & \leq C \|a\|_{L^\infty(C_d^\gamma)} (s-t)^{-1-(k-\frac{1}{2})+\gamma/2} \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}). \tag{5.30}
 \end{aligned}$$

We carefully point-out that the indicated bound only depend on the supremum in time of the Hölder modulus of the coefficients (denoted  $\|\mathbf{F}_1\|_{L^\infty(C_d^\gamma)}$ ,  $\|a\|_{L^\infty(C_d^\gamma)}$  respectively) and not on their supremum norm. In particular, we get from (5.29), (5.30):

$$\begin{aligned}
 & |D_{\mathbf{x}_1}^2 \Delta_{\mathbf{diag}}^1(t, T, \mathbf{x}, \mathbf{x}')| \\
 & \leq \sum_{k=1}^n \int_{t_0}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (s-t)^{-1-(k-\frac{1}{2})+\gamma/2} \\
 & \quad \times [\|D_{\mathbf{x}_1} u\|_{L^\infty} \|\mathbf{F}_1\|_{L^\infty(C_d^\gamma)} + \|D_{\mathbf{x}_1}^2 u\|_{L^\infty} \|a\|_{L^\infty(C_d^\gamma)}] \bar{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) |\mathbf{x}_k - \mathbf{x}'_k| \\
 & \leq C (\|D_{\mathbf{x}_1} u\|_{L^\infty} + \|D_{\mathbf{x}_1}^2 u\|_{L^\infty}) \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'), \tag{5.31}
 \end{aligned}$$

where again the constant  $C$  is *small* provided the coefficients *do not vary much*. Thanks to Lemma 5 and the previous analysis of Section 4.2, we directly deduce:

$$\begin{aligned}
 & |D_{\mathbf{x}_1}^2 \Delta_{\mathbf{diag}}^{2:n}(t, T, \mathbf{x}, \mathbf{x}')| \\
 & \leq \left| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu D_{\mathbf{y}_i} \cdot \left( \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{p}^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}') \right) \right. \\
 & \quad \left. \otimes \left( \mathbf{F}_i(s, \mathbf{y}) - [\mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) + D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}] \right) \right\} u(s, \mathbf{y}) \mathbb{I}_{s>t_0} \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & \leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \sum_{k=1}^n \int_{t_0}^T \frac{|\mathbf{x}_k - \mathbf{x}'_k| ds}{(s-t)^{1+(k-\frac{1}{2})-\frac{\gamma}{2}}} \leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^{\gamma-(2k-1)}(\mathbf{x}, \mathbf{x}') |\mathbf{x}_k - \mathbf{x}'_k| \\
 & \leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'), \tag{5.32}
 \end{aligned}$$

and  $C$  is again small provided the coefficients do not vary much<sup>†</sup>.

Plugging (5.31) and (5.32) into (5.28) yields the stated control for the diagonal contribution.

Let us now turn to the control of  $|D_{\mathbf{x}_1}^2 \Delta_{\mathbf{off-diag}}^{\boldsymbol{\xi}, \boldsymbol{\xi}'}(t, \mathbf{x}, \mathbf{x}')|$  in (5.27) (or (2.28) in the user's

---

<sup>†</sup>. Let us point out that in (5.31), (5.31), the constant  $C := C(\mathbf{A}, c_0)$ , but involves negative exponents of  $c_0$  meant to be *small*. This is not a problem and still gives  $C < 1$  provided the Hölder moduli of the coefficients are sufficiently small.

guide). In the off-diagonal case, we choose  $\boldsymbol{\xi} = \mathbf{x}$  and  $\boldsymbol{\xi}' = \mathbf{x}'$  and

$$\begin{aligned}
 & |D_{\mathbf{x}_1}^2 \Delta_{\text{off-diag}}^{\boldsymbol{\xi}, \boldsymbol{\xi}'}(t, \mathbf{x}, \mathbf{x}')| \\
 & \leq \left| \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y})(L_s - \tilde{L}_s^{\boldsymbol{\xi}})u(s, \mathbf{y}) \right|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & \quad + \left| \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_1}^2 \tilde{p}^{\boldsymbol{\xi}'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\boldsymbol{\xi}'})u(s, \mathbf{y}) \right|_{\boldsymbol{\xi}'=\mathbf{x}'}. \tag{5.33}
 \end{aligned}$$

We readily get thanks to Lemmas 4 and 5:

$$\begin{aligned}
 & |D_{\mathbf{x}_1}^2 \Delta_{\text{off-diag}}^{\boldsymbol{\xi}, \boldsymbol{\xi}'}(t, \mathbf{x}, \mathbf{x}')| \\
 & \leq C \sum_{i=2}^n \int_t^{t_0} \frac{ds}{(s-t)^{1-\frac{\gamma}{2}}} \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \\
 & \leq C \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \|u\|_{L^\infty(C_{b,d}^{2+\gamma})}. \tag{5.34}
 \end{aligned}$$

We point out that the constant  $C$  in (5.34) involves positive exponents of  $c_0$ . We have opposite impacts for this thresholds depending on the diagonal and off-diagonal regimes at hand.

- Control of  $\sup_{z \in \mathbb{R}^d} [u(t, z, \cdot)]_{2+\gamma}$ .

We proceed as above from definitions (5.27) considering spatial points  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$  s.t.  $\mathbf{x}_1 = \mathbf{x}'_1$ . In the diagonal case, we also choose  $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}' = \mathbf{x}$  and we write similarly to (5.28):

$$\begin{aligned}
 & |\Delta_{2:n, \text{diag}}^{\boldsymbol{\xi}, \boldsymbol{\xi}'}(t, T, \mathbf{x}, \mathbf{x}')|_{(\boldsymbol{\xi}, \boldsymbol{\xi}')=(\mathbf{x}, \mathbf{x})} \\
 & \leq \left| \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu \langle D_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}), \mathbf{x} - \mathbf{x}' \rangle \Delta_{1, \mathbf{F}, \sigma}(t, s, \mathbf{y}, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \mathbb{I}_{s>t_0} \right|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & \quad + \left| \sum_{i=2}^n \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu \langle D_{\mathbf{x}} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}), \mathbf{x} - \mathbf{x}' \rangle \right. \\
 & \quad \left. \Delta_{i, \mathbf{F}}(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}) D_{\mathbf{y}_i} u(s, \mathbf{y}) \mathbb{I}_{s>t_0} \right|_{\boldsymbol{\xi}=\mathbf{x}}. \tag{5.35}
 \end{aligned}$$

We have an expression rather similar to the one that appeared for the control of  $[D_{\mathbf{x}_1}^2 u(t, \cdot)]_\gamma$  but with a weaker time singularity. In other words, thanks to identities (5.29), (5.30) and Lemma 5 we obtain:

$$\begin{aligned}
 & |\Delta_{2:n, \text{diag}}^{\boldsymbol{\xi}, \boldsymbol{\xi}'}(t, T, \mathbf{x}, \mathbf{x}')| \\
 & \leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \sum_{k=1}^n \int_{t_0}^T ds (s-t)^{-(k-\frac{1}{2})+\frac{\gamma}{2}} |\mathbf{x}_k - \mathbf{x}'_k| \\
 & \leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \sum_{k=1}^n \mathbf{d}^{2(1-(k-\frac{1}{2})+\frac{\gamma}{2})}(\mathbf{x}, \mathbf{x}') |\mathbf{x}_k - \mathbf{x}'_k| \\
 & \leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'), \tag{5.36}
 \end{aligned}$$

where again the constant  $C$  is small if the Hölder moduli of the coefficients are small.

For the off-diagonal contribution, we get for  $\mathbf{x}_1 = \mathbf{x}'$ ,  $\boldsymbol{\xi} = \mathbf{x}$  and  $\boldsymbol{\xi}' = \mathbf{x}'$ :

$$\begin{aligned}
 & |\Delta_{2:n, \text{off-diag}}^{\boldsymbol{\xi}, \boldsymbol{\xi}'}(t, T, \mathbf{x}, \mathbf{x}')| \\
 & \leq \left| \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y})(L_s - \tilde{L}_s^{\boldsymbol{\xi}})u(s, \mathbf{y}) \right|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & \quad + \left| \int_t^{t_0} ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{\boldsymbol{\xi}'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\boldsymbol{\xi}'})u(s, \mathbf{y}) \right|_{\boldsymbol{\xi}'=\mathbf{x}'} \\
 & \leq C \sum_{k=2}^n \int_t^{t_0} ds (s-t)^{\frac{\gamma}{2}} \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \\
 & = C c_0 \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}') \|u\|_{L^\infty(C_{b,d}^{2+\gamma})}. \tag{5.37}
 \end{aligned}$$

The last but one inequality is a consequence of Lemmas 4 and 5.

## 5.4 Controls of the discontinuity terms arising from the change of freezing point

It now remains to control the contribution arising from the change of freezing point in equation (5.3). The main result of this section is the following lemma.

**Lemma 10 (Control of the discontinuity terms)** *There exists  $C := C((\mathbf{A}))$  s.t. for all  $(t, \mathbf{x}, \mathbf{x}') \in [0, T] \times \mathbb{R}^{nd} \times \mathbb{R}^{nd}$  taking  $\boldsymbol{\xi}' = \mathbf{x}'$ ,  $\tilde{\boldsymbol{\xi}}' = \mathbf{x}$ ,*

$$\begin{aligned}
 |D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\boldsymbol{\xi}'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\tilde{\boldsymbol{\xi}}'} u(t_0, \mathbf{x}')| & \leq C c_0^{\frac{\gamma}{2n-1}} \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'), \\
 |\tilde{P}_{t_0,t}^{\boldsymbol{\xi}'} u(t_0, \mathbf{x}') - \tilde{P}_{t_0,t}^{\tilde{\boldsymbol{\xi}}'} u(t_0, \mathbf{x}')| & \leq C c_0 \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'), \text{ for } \mathbf{x}_1 = \mathbf{x}'_1.
 \end{aligned}$$

We prove the above statement in the next paragraphs respectively dedicated to the control of the derivatives w.r.t. the non-degenerate variables and the control of the Hölder moduli associated with the degenerate ones.

**Control of the derivatives w.r.t. the non-degenerate variables**

We can directly write like for (5.4) and (5.6):

$$\begin{aligned}
 & \left. D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \right|_{(\xi', \tilde{\xi}') = (\mathbf{x}', \mathbf{x})} \\
 = & \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_1, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_{2:n})] d\mathbf{y} \right. \\
 & \left. - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_1, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_{2:n})] d\mathbf{y} \right] \\
 & + \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \int_0^1 d\mu(1-\mu) \right. \\
 & \quad \text{Tr} \left( [D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_1 + \mu(\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_1), \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_{2:n}) \right. \\
 & \quad \left. \left. - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) \right] (\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}'))_1^{\otimes 2} \right) d\mathbf{y} \\
 & - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) \int_0^1 d\mu(1-\mu) \\
 & \quad \text{Tr} \left( [D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_1 + \mu(\mathbf{y} - \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_1), \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_{2:n}) \right. \\
 & \quad \left. \left. - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')) \right] (\mathbf{y} - \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}'))_1^{\otimes 2} \right) d\mathbf{y} \left. \right] \\
 & + \frac{1}{2} \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \text{Tr} \left( D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) (\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}'))_1^{\otimes 2} \right) d\mathbf{y} \right. \\
 & \left. - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) \text{Tr} \left( D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')) (\mathbf{y} - \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}'))_1^{\otimes 2} \right) d\mathbf{y} \right] \\
 =: & \Delta_{t,t_0,\xi',\tilde{\xi}'} D_{\mathbf{x}_1}^2 \tilde{P}_1 u(t_0, \mathbf{x}, \mathbf{x}') + \Delta_{t,t_0,\xi',\tilde{\xi}'} D_{\mathbf{x}_1}^2 \tilde{P}_{21} u(t_0, \mathbf{x}, \mathbf{x}') + \Delta_{t,t_0,\xi',\tilde{\xi}'} D_{\mathbf{x}_1}^2 \tilde{P}_{22} u(t_0, \mathbf{x}, \mathbf{x}'). \tag{5.38}
 \end{aligned}$$

We first directly write from (5.5) and (5.7):

$$\begin{aligned}
 & \left| \Delta_{t,t_0,\xi',\tilde{\xi}'} D_{\mathbf{x}_1}^2 \tilde{P}_1 u(t_0, \mathbf{x}, \mathbf{x}') + \Delta_{t,t_0,\xi',\tilde{\xi}'} D_{\mathbf{x}_1}^2 \tilde{P}_{21} u(t_0, \mathbf{x}, \mathbf{x}') \right|_{(\xi', \tilde{\xi}') = (\mathbf{x}', \mathbf{x})} \\
 \leq & 2C(t_0 - t)^{\frac{\gamma}{2}} \|u\|_{C_{b,d}^{2+\gamma}} \leq Cc_0^{\frac{\gamma}{2}} \|u\|_{C_{b,d}^{2+\gamma}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{5.39}
 \end{aligned}$$

Let us now deal with the last term in (5.38). We proceed similarly to equation (5.8) (control of the frozen semigroup). Write:

$$\begin{aligned}
 & \left| \Delta_{t,t_0,\xi',\tilde{\xi}'} D_{\mathbf{x}_1}^2 \tilde{P}_{22} u(t_0, \mathbf{x}, \mathbf{x}') \right|_{(\xi',\tilde{\xi}')=(\mathbf{x}',\mathbf{x})} \\
 & \leq \frac{1}{2} \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(s-t)} \tilde{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) |D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}'))| |(\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}'))_1|^2 \\
 & \quad + \frac{1}{2} \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_1}^2 \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \operatorname{Tr} \left( D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) (\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}'))_1^{\otimes 2} \right) \right. \\
 & \quad \left. - D_{\mathbf{x}_1}^2 \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) \operatorname{Tr} \left( D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')) (\mathbf{y} - \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}'))_1^{\otimes 2} \right) d\mathbf{y} \right| \\
 & \leq C |D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}'))|,
 \end{aligned} \tag{5.40}$$

exploiting Proposition 2 for the first contribution and identity (3.16) in Proposition 3 for second contribution in the last inequality.

From Lemma 3, we derive that for  $t_0 - t = c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ :

$$\left| \Delta_{t,t_0,\xi',\tilde{\xi}'} D_{\mathbf{x}_1}^2 \tilde{P}_{22} u(t_0, \mathbf{x}, \mathbf{x}') \right|_{(\xi',\tilde{\xi}')=(\mathbf{x}',\mathbf{x})} \leq C \|u(t_0, \cdot)\|_{C_{b,d}^{2+\gamma} c_0^{\frac{\gamma}{2n-1}}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}').$$

In a nutshell, from (5.39) and (5.41), we get the following control:

$$\left| D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_1}^2 \tilde{P}_{t_0,t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \right|_{(\xi',\tilde{\xi}')=(\mathbf{x},\mathbf{x}')} \leq C c_0^{\frac{\gamma}{2n-1}} \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'),$$

which gives the statement for the second order derivatives w.r.t. the non-degenerate variables.

### Hölder controls for the degenerate variables

Again, for the Hölder norm w.r.t. the degenerate variables, the difficulty is that we cannot take any advantage of cancellation tools. We adapt here the arguments employed in Section 5.1 for the frozen semigroup. Precisely, for all  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ ,  $\mathbf{x}_1 = \mathbf{x}'_1$ ,  $\xi' = \mathbf{x}'$ ,  $\tilde{\xi}' = \mathbf{x}$ , we have similarly to (5.38) (but without the spatial derivatives  $D_{\mathbf{x}_1}^2$ ):

$$\begin{aligned}
 & \tilde{P}_{t_0,t}^{\xi'} u(t_0, \mathbf{x}') - \tilde{P}_{t_0,t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \\
 & = \Delta_{t,t_0,\xi',\tilde{\xi}'} \tilde{P}_1 u(t_0, \mathbf{x}', \mathbf{x}') + \Delta_{t,t_0,\xi',\tilde{\xi}'} \tilde{P}_{21} u(t_0, \mathbf{x}', \mathbf{x}') \\
 & \quad + \frac{1}{2} \operatorname{Tr} \left( [D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}'))] [\tilde{\mathbf{K}}_{t_0,t}^{\xi'}]_{1,1} \right) \\
 & \quad + \frac{1}{2} \operatorname{Tr} \left( D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')) ([\tilde{\mathbf{K}}_{t_0,t}^{\xi'}]_{1,1} - [\tilde{\mathbf{K}}_{t_0,t}^{\tilde{\xi}'}]_{1,1}) \right),
 \end{aligned} \tag{5.41}$$



where accordingly with (5.38):

$$\begin{aligned} \Delta_{t,t_0,\xi',\tilde{\xi}'}\tilde{P}_1 u(t_0, \mathbf{x}', \mathbf{x}') &:= \left[ \int_{\mathbb{R}^{nd}} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_1, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_{2:n})] d\mathbf{y} \right. \\ &\quad \left. - \int_{\mathbb{R}^{nd}} \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_1, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_{2:n})] d\mathbf{y} \right], \end{aligned}$$

and

$$\begin{aligned} &\Delta_{t,t_0,\xi',\tilde{\xi}'}\tilde{P}_{21} u(t_0, \mathbf{x}', \mathbf{x}') \\ := &\left[ \int_{\mathbb{R}^{nd}} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \int_0^1 d\mu(1-\mu) \right. \\ &\quad \text{Tr} \left( [D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_1 + \mu(\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_1), \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_{2:n}) \right. \\ &\quad \left. \left. - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_1) \right] (\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')_1)^{\otimes 2} \right) d\mathbf{y} \\ &\quad - \int_{\mathbb{R}^{nd}} \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) \int_0^1 d\mu(1-\mu) \\ &\quad \text{Tr} \left( [D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_1 + \mu(\mathbf{y} - \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_1), \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_{2:n}) \right. \\ &\quad \left. \left. - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_1) \right] (\mathbf{y} - \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')_1)^{\otimes 2} \right) d\mathbf{y} \left. \right]. \end{aligned}$$

Reproducing the arguments that led to Equations (5.15) and (5.16), we derive:

$$\begin{aligned} &\left| \Delta_{t,t_0,\xi',\tilde{\xi}'}\tilde{P}_1 u(t_0, \mathbf{x}', \mathbf{x}') + \Delta_{t,t_0,\xi',\tilde{\xi}'}\tilde{P}_{21} u(t_0, \mathbf{x}', \mathbf{x}') \right|_{(\xi',\tilde{\xi}')=(\mathbf{x}',\mathbf{x})} \\ &\leq C(t_0 - t)^{\frac{2+\gamma}{2}} \|u\|_{C_{b,d}^{2+\gamma}} \\ &\leq Cc_0^{\frac{2+\gamma}{2}} \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \end{aligned} \quad (5.42)$$

Let us now turn to the last two contributions in (5.41). Like for (5.18), from Proposition 1 and Lemma 3 we obtain:

$$\begin{aligned} &\frac{1}{2} \left| [D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}')) - D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}'))] [\tilde{\mathbf{K}}_{t_0,t}^{\xi'}]_{1,1} \right|_{(\xi',\tilde{\xi}')=(\mathbf{x}',\mathbf{x})} \\ &\leq C \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} (t_0 - t) \mathbf{d}^\gamma(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}'), \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')) \\ &\leq Cc_0^{1+\frac{\gamma}{2n-1}} \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \end{aligned} \quad (5.43)$$

The last term of (5.41) is handled like in (5.19). Namely, by Lemma 1 (equation (3.18) for  $j = 1$ ), we obtain:

$$\begin{aligned} &\left| \frac{1}{2} D_{\mathbf{x}_1}^2 u(t_0, \mathbf{m}_{t_0,t}^{\tilde{\xi}'}(\mathbf{x}')) ([\tilde{\mathbf{K}}_{t_0,t}^{\xi'}]_{1,1} - [\tilde{\mathbf{K}}_{t_0,t}^{\tilde{\xi}'}]_{1,1}) \right|_{(\xi',\tilde{\xi}')=(\mathbf{x}',\mathbf{x})} \\ &\leq C \|u(t_0, \cdot)\|_{C_{b,d}^{2+\gamma}} \left( (t_0 - t)^{\frac{2+\gamma}{2}} + (t_0 - t) \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \right) \\ &\leq Cc_0 \|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \end{aligned} \quad (5.44)$$

Plugging (5.42), (5.43) and (5.44) into (5.41), recalling as well that  $c_0 < 1$ , we derive the second statement of Lemma 10.

**Remark 5 (Concluding remark concerning the *a priori* estimates)** *Gathering all the controls of Sections 4 and 5 in equations (4.1) and (5.1)-(5.3), we actually derive the following bound. There exist constants  $C := C((\mathbf{T}))$  and  $C_0 := C_0((\mathbf{A}), T)$  s.t.*

$$\|u\|_{L^\infty(C_{b,d}^{2+\gamma})} \leq C(\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty(C_{b,d}^\gamma)}) + C_0\|u\|_{L^\infty(C_{b,d}^{2+\gamma})}. \quad (5.45)$$

*It can be observed from the previous technical lemmas that provided  $T, c_0$ , and the Hölder moduli of the coefficients  $\|a\|_{L^\infty(C_{\mathbf{d}}^\gamma)}$ ,  $\sum_{i=2}^n \|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})}$  (with the definition of (3.20)) are small enough, then  $C_0 < 1$ , which in turn directly yields the expected Schauder estimate for the mollified dynamics.*

*Of course, the remaining delicate part consists in getting rid of the small Hölder moduli constraint. This can be done through suitable scaling arguments that are exposed in Section 6.*

## 6 Scaling issues and final proof of Theorem 1

The purpose of this section is to first introduce a suitable scaling procedure for the system with mollified coefficients for which we will be able to show equation (5.45). This intuitively means that the scaling has to make the Hölder moduli of the considered coefficients small. The expected control is then obtained going back to the initial variables through the *inverse* scaling procedure. Also, once the estimate is established for *small* final time horizon  $T$ , it can be deduced through iteration up to an arbitrary given time precisely because it provides a kind of stability for the solution in the space  $L^\infty([0, T], C_{b,d}^{2+\gamma}(\mathbb{R}^{nd}, \mathbb{R}))$ . We then conclude the proof of our main result, Theorem 1 through compactness arguments.

### 6.1 Scaling settings and controls

We start here from the smooth solution  $u$  to equation (1.1) with mollified coefficients (that we again denote by a slight abuse of notation without the mollifying parameter  $m$ ). For an additional parameter  $\lambda > 0$  to be specified later (but meant to be *small*), introducing the scaled function  $u^\lambda(t, \mathbf{x}) = u(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x})$ , it is then clear that this latter satisfies

$$\begin{cases} \partial_t u^\lambda(t, \mathbf{x}) + \langle \mathbf{F}(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}), \lambda^{1/2}\mathbb{T}_\lambda^{-1} \mathbf{D}u^\lambda(t, \mathbf{x}) \rangle \\ + \frac{\lambda^{-1}}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u^\lambda(t, \mathbf{x}) a(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x})) = -f(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}), & (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \\ u^\lambda(T, \mathbf{x}) = g(\lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}), & \mathbf{x} \in \mathbb{R}^{nd}. \end{cases} \quad (6.1)$$

This choice of rescaling is natural in view of the invariance by dilatation property (1.7), i.e. each single variable  $\mathbf{x}_i$  is scaled by the parameter  $\lambda$  according to its corresponding intrinsic scale.

We rewrite in short form the above equation as for all  $\mathbf{x} \in \mathbb{R}^{nd}$

$$\begin{cases} \partial_t u^\lambda(t, \mathbf{x}) + \langle \mathbf{F}_\lambda(t, \mathbf{x}), \mathbf{D}u^\lambda(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u^\lambda(t, \mathbf{x}) a_\lambda(t, \mathbf{x})) = -f_\lambda(t, \mathbf{x}), & t \in [0, T], \\ u^\lambda(T, \mathbf{x}) = g_\lambda(\mathbf{x}), \end{cases} \quad (6.2)$$

where

$$\begin{aligned} f_\lambda(t, \mathbf{x}) &:= f(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}), \quad g_\lambda(\mathbf{x}) := g(\lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}), \\ a_\lambda(t, \mathbf{x}) &:= \lambda^{-1}a(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}), \\ \mathbf{F}_\lambda(t, \mathbf{x}) &:= \lambda^{1/2}\mathbb{T}_\lambda^{-1}\mathbf{F}(t, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}). \end{aligned} \quad (6.3)$$

Accordingly, we introduce the spatial operator  $(L_s^\lambda)_{s \geq 0}$  appearing in (6.2) which writes explicitly for all  $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$  as:

$$L_t^\lambda \varphi = \langle \mathbf{F}_\lambda(t, \cdot), \mathbf{D}\varphi \rangle + \frac{1}{2}\text{Tr}\left(a_\lambda(t, \cdot)D_{\mathbf{x}_1}^2 \varphi\right), \quad \lambda > 0.$$

The dynamics of the SDE associated with the second order differential operator  $(L_t^\lambda)_{t \in [0, T]}$  appearing in (6.1)-(6.2) writes for a given starting point  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

$$\mathbf{X}_s^{\lambda, t, \mathbf{x}} = \mathbf{x} + \int_t^s \mathbf{F}_\lambda(u, \mathbf{X}_u^{\lambda, t, \mathbf{x}})du + \int_t^s B\sigma_\lambda(u, \mathbf{X}_u^{\lambda, t, \mathbf{x}})dW_u, \quad s \geq t, \quad (6.4)$$

where  $(W_u)_{u \geq 0}$  is a  $d$ -dimensional Brownian motion on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and  $\sigma_\lambda$  is a square root of the diffusion matrix  $a_\lambda$  introduced in (6.3).

Equation (6.4) then naturally leads to consider, for fixed  $(s, \mathbf{y}) \in [t, T] \times \mathbb{R}^{nd}$ , and with the notations of Section 3, the corresponding linearized model

$$\begin{aligned} d\tilde{\mathbf{X}}_v^{\xi, \lambda} &= [\mathbf{F}_\lambda(v, \boldsymbol{\theta}_{v,t}^\lambda(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_\lambda(v, \boldsymbol{\theta}_{v,t}^\lambda(\boldsymbol{\xi}^\lambda))(\tilde{\mathbf{X}}_v^{\xi, \lambda} - \boldsymbol{\theta}_{v,t}^\lambda(\boldsymbol{\xi}^\lambda))]dv + B\sigma_\lambda(v, \boldsymbol{\theta}_{v,t}^\lambda(\boldsymbol{\xi}^\lambda))dW_v, \\ \tilde{\mathbf{X}}_t^{\xi, \lambda} &= \mathbf{x}, \end{aligned} \quad (6.5)$$

where  $\boldsymbol{\theta}_{v,t}^\lambda(\boldsymbol{\xi}^\lambda) = \lambda^{1/2}\mathbb{T}_\lambda^{-1}\boldsymbol{\theta}_{v,t}(\boldsymbol{\xi}^\lambda)$ ,  $\boldsymbol{\xi}^\lambda = \lambda^{-1/2}\mathbb{T}_\lambda \boldsymbol{\xi}$ .

The associated generator writes for  $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$  and  $(s, \mathbf{y}) \in [t, T] \times \mathbb{R}^{nd}$  as:

$$\begin{aligned} \tilde{L}_s^{\lambda, \xi} \varphi(\mathbf{y}) &= \langle \mathbf{F}_\lambda(s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_\lambda(s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))(\mathbf{y} - \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)), \mathbf{D}\varphi(\mathbf{y}) \rangle \\ &\quad + \frac{1}{2}\text{Tr}(a_\lambda(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))D_{\mathbf{y}_1}^2 \varphi(\mathbf{y})). \end{aligned}$$

Observe from (2.6) and (6.5) that the following very important correspondence holds:

$$\forall v \in [t, T], \quad \tilde{\mathbf{X}}_v^{\xi, \lambda} := \lambda^{1/2}\mathbb{T}_\lambda^{-1}\tilde{\mathbf{X}}_v^{\xi^\lambda}. \quad (6.6)$$

We thus derive from Proposition 2 the following important correspondence for the densities. Denoting by  $\tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y})$  the density of  $\tilde{\mathbf{X}}_v^{\xi, \lambda}$  starting from  $\mathbf{x}$  at time  $t$  at point  $\mathbf{y}$  in  $s$ , and  $\mathbf{x}^\lambda = \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}$ , we have

$$\begin{aligned} &\tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) \\ &= \lambda^{\frac{n^2 d}{2}} \tilde{p}^{\xi^\lambda}(t, s, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{y}) \\ &= \frac{\lambda^{\frac{n^2 d}{2}}}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s,t}^{\xi^\lambda})^{\frac{1}{2}}} \exp\left(-\frac{\lambda^{-1}}{2} \left\langle [\mathbb{T}_\lambda(\tilde{\mathbf{K}}_{s,t}^{\xi^\lambda})^{-1}\mathbb{T}_\lambda](\lambda^{1/2}\mathbb{T}_\lambda^{-1}\mathbf{m}_{s,t}^{\xi^\lambda}(\mathbf{x}^\lambda) - \mathbf{y}), \lambda^{1/2}\mathbb{T}_\lambda^{-1}\mathbf{m}_{s,t}^{\xi^\lambda}(\mathbf{x}^\lambda) - \mathbf{y} \right\rangle\right). \end{aligned} \quad (6.7)$$

In particular, for  $\boldsymbol{\xi} = \mathbf{x}$  one derives:

$$\begin{aligned}
 & \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) \\
 &= \lambda^{\frac{n^2 d}{2}} \tilde{p}^{\boldsymbol{\xi}^\lambda}(t, s, \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \mathbf{x}, \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \mathbf{y}) \\
 &= \frac{\lambda^{\frac{n^2 d}{2}}}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s,t}^{\boldsymbol{\xi}^\lambda})^{\frac{1}{2}}} \exp\left(-\frac{\lambda^{-1}}{2} \left\langle [\mathbb{T}_\lambda (\tilde{\mathbf{K}}_{s,t}^{\boldsymbol{\xi}^\lambda})^{-1} \mathbb{T}_\lambda] (\boldsymbol{\theta}_{s,t}^\lambda(\mathbf{x}^\lambda) - \mathbf{y}), \boldsymbol{\theta}_{s,t}^\lambda(\mathbf{x}^\lambda) - \mathbf{y} \right\rangle\right).
 \end{aligned} \tag{6.8}$$

Equation (6.7) in turn yields the following important control:

$$\begin{aligned}
 & |D_{\mathbf{x}}^\alpha \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y})| \\
 &\leq C \left(\frac{\lambda}{(s-t)}\right)^{\sum_{i=1}^n \alpha_i(i-\frac{1}{2}) + \frac{n^2 d}{2}} \exp\left(-C^{-1} \frac{(s-t)}{\lambda} \left| \mathbb{T}_{\frac{s-t}{\lambda}}^{-1} (\lambda^{\frac{1}{2}} \mathbb{T}_\lambda^{-1} \mathbf{m}_{s,t}^{\boldsymbol{\xi}^\lambda}(\mathbf{x}^\lambda) - \mathbf{y}) \right|^2\right) \\
 &=: C \left(\frac{\lambda}{(s-t)}\right)^{\sum_{i=1}^n \alpha_i(i-\frac{1}{2})} \bar{p}_{C^{-1}, \lambda}^\xi(t, s, \mathbf{x}, \mathbf{y}).
 \end{aligned} \tag{6.9}$$

In the following, we will also denote  $\bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}) := \bar{p}_{C^{-1}, \lambda}^\xi(t, s, \mathbf{x}, \mathbf{y}) \Big|_{\boldsymbol{\xi}=\mathbf{x}}$  in (6.9).

**Remark 6** We emphasize that (6.9) gives that, each derivation of the Gaussian kernel  $\tilde{p}_\lambda^\xi$  makes the small parameter  $\lambda$  appear. Hence, up to the additional time singularities, the iterated derivatives become smaller and smaller.

## 6.2 Scaling properties

The point is now to reproduce the previous perturbative approach for the solution of (6.2). We will here focus on the Hölder norm of the remainder term associated with the second order derivatives w.r.t. the non-degenerate variables. It can indeed be seen from the previous computations that the other contributions can be dealt similarly. We aim at proving that the  $L^\infty([0, T], C_{b,d}^\gamma(\mathbb{R}^{nd}, \mathbb{R}))$  norm of the indicated term is small. This will follow from our scaling procedure. Precisely, we introduce for  $\lambda > 0$ ,  $(t, \mathbf{x}, \mathbf{x}', \mathbf{x}') \in [0, T] \times (\mathbb{R}^{nd})^2$  the quantity:

$$\begin{aligned}
 & (D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)(t, T, \mathbf{x}, \mathbf{x}') \\
 &:= \left( \int_t^{(t+c_0\lambda d^2(\mathbf{x}, \mathbf{x}')) \wedge T} ds \int_{\mathbb{R}^{nd}} \left( D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y})(L_s^\lambda - \tilde{L}_s^{\lambda, \boldsymbol{\xi}}) u^\lambda(s, \mathbf{y}) \right. \right. \\
 &\quad \left. \left. - D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^{\boldsymbol{\xi}'}(t, s, \mathbf{x}', \mathbf{y})(L_s^\lambda - \tilde{L}_s^{\lambda, \boldsymbol{\xi}'}) u^\lambda(s, \mathbf{y}) \right) d\mathbf{y} \right) \Big|_{(\boldsymbol{\xi}, \boldsymbol{\xi}')=(\mathbf{x}, \mathbf{x}')} \\
 &\quad + \left( \int_{(t+c_0\lambda d^2(\mathbf{x}, \mathbf{x}')) \wedge T}^T ds \int_{\mathbb{R}^{nd}} \left( D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y})(L_s^\lambda - \tilde{L}_s^{\lambda, \boldsymbol{\xi}}) u^\lambda(s, \mathbf{y}) \right. \right. \\
 &\quad \left. \left. - D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^{\boldsymbol{\xi}'}(t, s, \mathbf{x}', \mathbf{y})(L_s^\lambda - \tilde{L}_s^{\lambda, \boldsymbol{\xi}'}) u^\lambda(s, \mathbf{y}) \right) d\mathbf{y} \right) \Big|_{(\boldsymbol{\xi}, \boldsymbol{\xi}')=(\mathbf{x}, \mathbf{x})} \\
 &=: \left( (D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{off-diag}} + (D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{diag}} \right)(t, T, \mathbf{x}, \mathbf{x}'),
 \end{aligned} \tag{6.10}$$

using again as in (5.3) different freezing point according to the spatial regime w.r.t. integration time  $s$ . Note however carefully that the cutting threshold here depends on the scaling parameter  $\lambda$ . This is very important in order to balance the various scales that will appear. Pay attention as well that the parameter  $c_0$  also remains. Actually a subtle balance between those two parameters will be needed to derive the expected control. Proceeding as in Section 5.3, we aim at showing that there exists  $C := C(\mathbf{A})$  s.t. for all  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ :

$$\sup_{t \in [0, T]} \frac{|(D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)(t, T, \mathbf{x}, \mathbf{x}')|}{\mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')} \leq C(c_0^{-(n-\frac{1}{2})+\frac{\gamma}{2}} \lambda^{\frac{\gamma}{2}} + c_0^{\frac{\gamma}{2n-1}}) \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})}. \quad (6.11)$$

Let us first consider the diagonal term in (6.10) assuming w.l.o.g. that  $t + c_0 \lambda \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \leq T$  (otherwise we only have the off-diagonal contribution). Write:

$$\begin{aligned} & |(D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{diag}}(t, T, \mathbf{x}, \mathbf{x}')| \\ & \leq \left| \int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left[ D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x}', \mathbf{y}) \right] \right. \\ & \quad \left[ \langle \mathbf{F}_{\lambda,1}(t, \mathbf{y}) - \mathbf{F}_{\lambda,1}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)), D_{\mathbf{y}_1} \rangle \right. \\ & \quad \left. + \frac{1}{2} \text{Tr} \left( (a_\lambda(t, \mathbf{y}) - a_\lambda(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))) D_{\mathbf{y}_1}^2 \right) \right] u^\lambda(s, \mathbf{y}) \Big| \\ & \quad + \left| \sum_{i=2}^n \int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left[ D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x}', \mathbf{y}) \right] \right. \\ & \quad \left. \left\langle \mathbf{F}_{\lambda,i}(t, \mathbf{y}) - [\mathbf{F}_{\lambda,i}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_{\lambda,i}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))(\mathbf{y} - \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))_{i-1}], D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \right\rangle \Big|_{(\boldsymbol{\xi}, \boldsymbol{\xi}') = (\mathbf{x}, \mathbf{x})} \\ & = \left| \int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}') \right. \\ & \quad \left( \left\langle \mathbf{F}_{\lambda,1}(t, \mathbf{y}) - \mathbf{F}_{\lambda,1}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)), D_{\mathbf{y}_1} \right\rangle + \frac{1}{2} \text{Tr} \left( (a_\lambda(t, \mathbf{y}) - a_\lambda(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))) D_{\mathbf{y}_1}^2 \right) \right) u^\lambda(s, \mathbf{y}) \Big| \\ & \quad + \left| \sum_{i=2}^n \int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu \mathbf{D} D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \cdot (\mathbf{x} - \mathbf{x}') \right. \\ & \quad \left. \left\langle \mathbf{F}_{\lambda,i}(t, \mathbf{y}) - [\mathbf{F}_{\lambda,i}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_{\lambda,i}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))(\mathbf{y} - \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))_{i-1}], D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \right\rangle \Big|_{(\boldsymbol{\xi}, \boldsymbol{\xi}') = (\mathbf{x}, \mathbf{x})} \\ & =: (D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{diag},1}(t, T, \mathbf{x}, \mathbf{x}') + (D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{diag},2:n}(t, T, \mathbf{x}, \mathbf{x}'). \end{aligned} \quad (6.12)$$

We will now control the first term of the above right hand side. A key point for the analysis, is to observe that, on the considered diagonal regime, we actually have from equations (6.8)-(6.9), recalling that  $\mathbf{z} \mapsto \mathbf{m}_{s,t}^{\boldsymbol{\xi}^\lambda}(\mathbf{z})$  is affine and using the good scaling property in

(3.5) and (3.11), that:

$$\begin{aligned}
 & |D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y})| \\
 \leq & \frac{C\lambda^{1+(k-\frac{1}{2})}}{(s-t)^{1+(k-\frac{1}{2})}} \bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}) \exp\left(\sum_{j=1}^n \frac{|(\mathbf{x} - \mathbf{x}')_j|^2 \lambda^{2j-1}}{(s-t)^{2j-1}}\right) \\
 \leq & \frac{C\lambda^{1+(k-\frac{1}{2})}}{(s-t)^{1+(k-\frac{1}{2})}} \bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}). \tag{6.13}
 \end{aligned}$$

To obtain the last inequality, we indeed observe from the homogeneity of the metric (see (1.9)) that  $\lambda^{1/2} \mathbf{d}(\mathbf{x}, \mathbf{x}') = \mathbf{d}(\lambda^{-1/2} \mathbb{T}_\lambda \mathbf{x}, \lambda^{-1/2} \mathbb{T}_\lambda \mathbf{x}') = \sum_{j=1}^n |(\mathbf{x} - \mathbf{x}')_j| \lambda^{\frac{2j-1}{2}} \left| \frac{1}{(2j-1)} \right|$ . For the diagonal regime  $\lambda \mathbf{d}^2(\mathbf{x}, \mathbf{x}') \leq (s-t)$  in turn implies that for all  $j \in \llbracket 1, n \rrbracket$ ,  $\frac{|(\mathbf{x} - \mathbf{x}')_j|^2 \lambda^{2j-1}}{(s-t)^{2j-1}} \leq 1$ .

Another key point is to observe that the contributions  $\frac{\mathbf{d}(\lambda^{-1/2} \mathbb{T}_\lambda \mathbf{y}, \lambda^{-1/2} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}^\lambda(\mathbf{x}^\lambda))}{(s-t)^{1/2}}$  is homogeneous to the argument of the exponential term in  $\bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y})$ . Namely, for any given  $\beta_0 > 0$  and  $\beta \in (0, \beta_0]$ , there exists  $C_{\beta_0}$  s.t.

$$\left( \frac{\mathbf{d}(\lambda^{-1/2} \mathbb{T}_\lambda \mathbf{y}, \lambda^{-1/2} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}^\lambda(\mathbf{x}^\lambda))}{(s-t)^{1/2}} \right)^\beta \bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}) \leq C_{\beta_0} \bar{p}_{\beta_0, \lambda}(t, s, \mathbf{x}, \mathbf{y}). \tag{6.14}$$

Equation (6.14) is a direct consequence of the expression of  $\bar{p}_{C^{-1}, \lambda}^{t, \mathbf{x}}$  in Proposition 2 and the definition of  $\mathbf{d}$  in (1.9).

Hence, from the definition of  $a_\lambda, \mathbf{F}_\lambda$  in (6.3) equations (6.13) and (6.14), we derive that, under **(A)** for all  $k \in \llbracket 1, n \rrbracket$ :

$$\begin{aligned}
 & \left| \mathbf{F}_{\lambda,1}(t, \mathbf{y}) - \mathbf{F}_{\lambda,1}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) \right| \left| D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 \leq & C \left( \lambda^{-1/2} \|\mathbf{F}_1(t, \cdot)\|_{C_d^\gamma} \mathbf{d}^\gamma(\lambda^{-1/2} \mathbb{T}_\lambda \mathbf{y}, \lambda^{-1/2} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) \right) \\
 & \times \lambda^{1+k-\frac{1}{2}} (s-t)^{-1-(k-\frac{1}{2})} \bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}) \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 \leq & C \|\mathbf{F}_1\|_{L^\infty(C_d^\gamma)} \lambda^k (s-t)^{-1-(k-\frac{1}{2})+\gamma/2} \bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}). \tag{6.15}
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 & \left| a_\lambda(t, \mathbf{y}) - a_\lambda(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) \right| \left| D_{\mathbf{x}_k} D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 \leq & C \left( \lambda^{-1} \|a(t, \cdot)\|_{C_d^\gamma} \mathbf{d}^\gamma(\lambda^{-1/2} \mathbb{T}_\lambda \mathbf{y}, \lambda^{-1/2} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) \right) \\
 & \times \lambda^{1+(k-\frac{1}{2})} (s-t)^{-1-(k-\frac{1}{2})} \bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}) \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 \leq & C \|a\|_{L^\infty(C_d^\gamma)} \lambda^{k-\frac{1}{2}} (s-t)^{-1-(k-\frac{1}{2})+\gamma/2} \bar{p}_{C^{-1}, \lambda}(t, s, \mathbf{x}, \mathbf{y}). \tag{6.16}
 \end{aligned}$$

Observe that both r.h.s of (6.15) and (6.16) exhibit a positive power of  $\lambda$ . Hence, those quantities are *small* provided  $\lambda$  is. The key point in the above computations is that the potentially explosive Hölder norms of  $\mathbf{F}_{\lambda,1}, a_\lambda$  (when  $\lambda$  goes to zero) are compensated

by the derivatives of  $\tilde{p}_\lambda^\xi(s, t, \mathbf{x}, \mathbf{y})$  (see equation (6.9)). We again carefully point-out that the previous bounds only depend on the supremum in time of the Hölder moduli of the coefficients (denoted  $\|\mathbf{F}_1\|_{L^\infty(C_d^\gamma)}$ ,  $\|a\|_{L^\infty(C_d^\gamma)}$  respectively) and not on their supremum norm. In particular, we get from (6.15), (6.16) with the notation of (6.12):

$$\begin{aligned}
 & (D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{diag},1}(t, T, \mathbf{x}, \mathbf{x}') \\
 & \leq \sum_{k=1}^n \int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \lambda^{k-\frac{1}{2}} (s-t)^{-1-(k-\frac{1}{2})+\gamma/2} \\
 & \quad \times [\lambda^{\frac{1}{2}} \|D_{\mathbf{x}_1} u^\lambda\|_{L^\infty} \|\mathbf{F}_1\|_{L^\infty(C_d^\gamma)} + \|D_{\mathbf{x}_1}^2 u^\lambda\|_{L^\infty} \|a\|_{L^\infty(C_d^\gamma)}] \bar{p}_{C^{-1},\lambda}(t, s, \mathbf{x}, \mathbf{y}) |\mathbf{x}_k - \mathbf{x}'_k| \\
 & \leq c_0^{-(k-\frac{1}{2})+\gamma/2} C (\|D_{\mathbf{x}_1} u^\lambda\|_{L^\infty} + \|D_{\mathbf{x}_1}^2 u^\lambda\|_{L^\infty}) \lambda^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{6.17}
 \end{aligned}$$

Thanks to the inequality (4.22) and the previous analysis (to be performed according to the current scaling procedure replacing  $(s-t)$  in Sections 4 and 5 by  $(s-t)/\lambda$ ), we deduce:

$$\begin{aligned}
 & (D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{diag},2:n}(t, T, \mathbf{x}, \mathbf{x}') \leq \left| \sum_{i=2}^n \int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_0^1 d\mu \right. \\
 & \quad \left\langle D_{\mathbf{y}_i} \left\{ (\mathbf{F}_{\lambda,i}(t, \mathbf{y}) - [\mathbf{F}_{\lambda,i}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda)) + D\mathbf{F}_{\lambda,i}(t, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))(\mathbf{y} - \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}^\lambda))]) \right\} \right. \\
 & \quad \left. \times \left( \langle \mathbf{x} - \mathbf{x}', D_{\mathbf{x}} D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x} + \mu(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \rangle \right) \right\rangle \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & \leq C \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})} \sum_{k=1}^n \int_{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')}^T \frac{\lambda^{k-\frac{1}{2}} |\mathbf{x}_k - \mathbf{x}'_k| ds}{(s-t)^{1+(k-\frac{1}{2})-\frac{\gamma}{2}}} \\
 & \leq C c_0^{-(k-\frac{1}{2})+\frac{\gamma}{2}} \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})} \lambda^{(k-\frac{1}{2})-(k-\frac{1}{2})+\frac{\gamma}{2}} \mathbf{d}^{\gamma-(2k-1)}(\mathbf{x}, \mathbf{x}') |\mathbf{x}_k - \mathbf{x}'_k| \\
 & \leq C c_0^{-(k-\frac{1}{2})+\frac{\gamma}{2}} \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})} \lambda^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \leq C c_0^{-(k-\frac{1}{2})+\frac{\gamma}{2}} \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})} \lambda^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \tag{6.18}
 \end{aligned}$$

Plugging (6.17) and (6.18) into (6.12) gives a diagonal control which precisely matches the r.h.s. of the expected final bound (6.11).

It therefore remains to handle the off-diagonal contributions. With the notations of (6.10), we deduce from the analysis of Section 4 and the previous arguments that:

$$|(D_{\mathbf{x}_1}^2 \mathcal{R}^\lambda u^\lambda)_{\text{off-diag}}| \leq C \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})} \int_t^{t+c_0\lambda\mathbf{d}^2(\mathbf{x},\mathbf{x}')} \frac{ds}{(s-t)^{1-\frac{\gamma}{2}}} \leq C c_0^{\frac{\gamma}{2}} \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})} \lambda^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \tag{6.19}$$

We emphasize that making the diagonal and off-diagonal sets depend on  $\lambda$  is here crucial. To illustrate this, let us briefly focus on the non-degenerate diffusive contribution. We need to proceed from (6.16) except that we have here no differentiation w.r.t.  $k$  and therefore no positive power in  $\lambda$  anymore. We retrieve the dependence in  $\lambda$  through the time integration. The bound of equation (6.19) completes the proof of (6.11).

In some sense the controls in (6.18) and (6.19) can be seen as a mere consequence of the intrinsic scaling of the system. This is indeed the case, but, in order to avoid any ambiguity, we provide in Appendix 9 a proof of the rescaled key Besov Lemma 5.

We would also derive from (5.41) and the same previous arguments by denoting  $t_0 = t + c_0 \lambda \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  and using the bound of (7.18) for the difference of the scaled flows, that :

$$\left| D_{\mathbf{x}_1}^2 \tilde{P}_{t_0, t}^{\xi'} u^\lambda(t_0, \mathbf{x}') - D_{\mathbf{x}_1}^2 \tilde{P}_{t_0, t}^{\tilde{\xi}'} u^\lambda(t_0, \mathbf{x}') \right|_{(\xi', \tilde{\xi}') = (\mathbf{x}, \mathbf{x}')} \leq C c_0^{\frac{1}{2n-1}} \|u^\lambda\|_{C_{b, \mathbf{d}}^{2+\gamma}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}'). \quad (6.20)$$

Exploiting (6.20) and (6.11), as well as the corresponding control

$$\sup_{t \in [0, T], (\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2, \mathbf{x}_1 = \mathbf{x}'_1} \frac{|(\mathcal{R}^\lambda u^\lambda)(t, T, \mathbf{x}, \mathbf{x}')|}{\mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}')} \leq C c_0^{\frac{1}{2n-1}} \|u^\lambda\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})},$$

that would be established through the same procedure (exploiting the control (8.13) for the difference of the scaled covariances), we eventually derive similarly to (5.45) that:

$$\|u^\lambda\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})} \leq C(\|g_\lambda\|_{C_{b, \mathbf{d}}^{2+\gamma}} + \|f_\lambda\|_{L^\infty(C_{b, \mathbf{d}}^\gamma)}) + C_0(c_0^{-(2n-1)} \lambda^{\frac{\gamma}{2}} + T^{\frac{\gamma}{2}} + c_0^{\frac{1}{2n-1}}) \|u^\lambda\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})},$$

for some constant  $C_0 := C_0(\mathbf{A})$ . Hence, for  $c_0, T$  and  $\lambda$  small enough with  $\lambda \ll c_0$ , i.e. s.t  $\bar{c}_0 := C_0(c_0^{-(2n-1)} \lambda^{\frac{\gamma}{2}} + T^{\frac{\gamma}{2}} + c_0^{\frac{1}{2n-1}}) < 1$  we derive the expect final control

$$\|u^\lambda\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})} \leq \frac{C}{1 - \bar{c}_0} (\|g_\lambda\|_{C_{b, \mathbf{d}}^{2+\gamma}} + \|f_\lambda\|_{L^\infty(C_{b, \mathbf{d}}^\gamma)}). \quad (6.21)$$

### 6.3 Conclusion: final proof of Theorem 1.

Equation (6.21) provide the expected Schauder estimate for the rescaled system with mollified coefficients in *small* time. Recalling that  $u^\lambda(t, \mathbf{x}) = u(t, \lambda^{-1/2} \mathbb{T}_\lambda \mathbf{x})$ , we then derive from (6.21) and (6.3) that for  $T > 0$  and  $\lambda$  small enough there exists  $\bar{C}_0 := \bar{C}_0(\mathbf{A}, T, \lambda) > 1$  s.t.

$$\|u\|_{L^\infty(C_{b, \mathbf{d}}^{2+\gamma})} \leq \bar{C}_0 (\|g\|_{C_{b, \mathbf{d}}^{2+\gamma}} + \|f\|_{L^\infty(C_{b, \mathbf{d}}^\gamma)}), \quad (6.22)$$

which precisely provides the required estimate in small time for the initial system with mollified coefficients.

The point is now to extend the previous bound to an arbitrary fixed time  $T > 0$  not necessarily small. The stability resulting from estimate (6.22) allows to proceed by simple iterative application of the bound changing the final condition.

#### Schauder estimates for the mollified system for a general time

Equation (6.22) is valid for all  $T < T_0$  with  $T_0 \in (0, +\infty)$  sufficiently small. Now, for a given  $T > 0$  (not necessary small), we can solve iteratively  $N = \lceil \frac{T}{T_0} \rceil$  (where  $\lceil \cdot \rceil$  is the ceiling function) Cauchy problems in the following way. Consider first for all  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$\begin{cases} \partial_t u_1(t, \mathbf{x}) + \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}u_1(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u_1(t, \mathbf{x}) a(t, \mathbf{x})) = -f(t, \mathbf{x}), & t \in [T(1 - \frac{1}{N}), T), \\ u_1(T, \mathbf{x}) = g(\mathbf{x}). \end{cases}$$

In other words, our previous analysis, and the previous inequalities, are still available for  $T - (1 - \frac{1}{N})T = \frac{1}{N}T \leq T_0$  small enough. Precisely, from (6.22):

$$\|u_1(T(1 - \frac{1}{N}), \cdot)\|_{C_{b, \mathbf{d}}^{2+\gamma}} \leq \bar{C}_0 (\|g\|_{C_{b, \mathbf{d}}^{2+\gamma}} + \|f\|_{L^\infty([T(1 - \frac{1}{N}), T], C_{b, \mathbf{d}}^{2+\gamma})}). \quad (6.23)$$



Also, for mollified coefficients, it is plain from the Feynman-Kac formula to identify  $u_1$  and  $u$  on  $[T(1 - \frac{1}{N}), T]$  where  $u$  solves (1.1) with mollified coefficients on  $[0, T]$ . Hence, (6.23) gives in particular that  $u_1(T(1 - \frac{1}{N}), \cdot) = u(T(1 - \frac{1}{N}), \cdot) \in C_{b,d}^{2+\gamma}(\mathbb{R}^{nd}, \mathbb{R})$  so that it is in particular natural to consider the following second Cauchy problem for all  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$\begin{cases} \partial_t u_2(t, \mathbf{x}) + \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}u_2(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u_2(t, \mathbf{x}) a(t, \mathbf{x})) = -f(t, \mathbf{x}), \\ t \in [(1 - \frac{2}{N})T, (1 - \frac{1}{N})T], \quad u_2((1 - \frac{1}{N})T, \mathbf{x}) = u((1 - \frac{1}{N})T, \mathbf{x}). \end{cases}$$

Hence  $u_2$  satisfies identity (6.22) for the corresponding time interval and the associated source and final condition. It also coincides with  $u$  on  $[T(1 - \frac{2}{N}), T(1 - \frac{1}{N})]$ . We get:

$$\begin{aligned} \|u_2\|_{L^\infty([T(1-\frac{2}{N}), T(1-\frac{1}{N})], C_{b,d}^{2+\gamma})} &\leq \bar{C}_0(\|u(T(1 - \frac{1}{N}), \cdot)\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty([T(1-\frac{2}{N}), T(1-\frac{1}{N})], C_{b,d}^{2+\gamma})}) \\ &\leq \bar{C}_0\left(\bar{C}_0(\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty([T(1-\frac{1}{N}), T], C_{b,d}^{2+\gamma})})\right. \\ &\quad \left. + \|f\|_{L^\infty([T(1-\frac{2}{N}), T(1-\frac{1}{N})], C_{b,d}^{2+\gamma})}\right) \\ &\leq \bar{C}_0^2(\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty([T(1-\frac{2}{N}), T], C_{b,d}^{2+\gamma})}). \end{aligned}$$

Repeating the analysis  $N$ -times, introducing for  $k \in \llbracket 3, n \rrbracket$  the auxiliary Cauchy problems for all  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$\begin{cases} \partial_t u_k(t, \mathbf{x}) + \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}u_k(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr}(D_{\mathbf{x}_1}^2 u_k(t, \mathbf{x}) a(t, \mathbf{x})) = -f(t, \mathbf{x}), \\ t \in [T(1 - \frac{k}{N}), T(1 - \frac{k-1}{N})], \quad u_k((1 - \frac{k-1}{N})T, \mathbf{x}) = u((1 - \frac{k-1}{N})T, \mathbf{x}), \end{cases}$$

we derive that

$$\|u\|_{L^\infty([0, T], C_{b,d}^{2+\gamma})} \leq \bar{C}_0^N (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty([0, T], C_{b,d}^{2+\gamma})}). \quad (6.24)$$

This precisely gives our main estimate for the system (1.1) with mollified coefficients. Again, even though the coefficients are smooth, all the constants appearing in (6.24) only depend on the Hölder setting of assumption **(A)**.

**Remark 7 (About the constants in the final estimate)** *We could actually have slightly better bounds than those in (6.24). Namely a direct induction shows that the following control holds.*

$$\|u\|_{L^\infty([0, T], C_{b,d}^{2+\gamma})} \leq \bar{C}_0^N \|g\|_{C_{b,d}^{2+\gamma}} + \sum_{i=1}^N C_0^i \|f\|_{L^\infty([T(1-\frac{N-i+1}{N}), T(1-\frac{N-i}{N})], C_{b,d}^{2+\gamma})}. \quad (6.25)$$

*We chose to write the bound in the form of equation (6.24) for simplicity. Note however that in any case, equation (6.24) or (6.25), we still have geometric constants coming from the iterative procedure. This is the main drawback of our approach, which anyhow seems, to the best of our knowledge, the only available one to consider degenerate Kolmogorov systems with non-linear drifts.*

### Compactness arguments

We now make the mollifying parameter appear again using the notations introduced in the user's guide (see equation (2.4)). Equation (6.24) rewrites in the following way. There exists a constant  $C_0$  s.t. for all  $m \in \mathbb{N}$ :

$$\|u_m\|_{L^\infty([0,T],C_{b,d}^{2+\gamma})} \leq \bar{C}_0^N (\|g\|_{C_{b,d}^{2+\gamma}} + \|f\|_{L^\infty([0,T],C_{b,d}^{2+\gamma})}). \quad (6.26)$$

From Ascoli's theorem we deduce that there exists  $u \in L^\infty([0,T],C_{b,d}^{2+\gamma})$  and a sequence of smooth functions  $(u_{m_k})_{k \geq 1}$  solution of (2.4), with  $m_k \rightarrow_k +\infty$ , satisfying (6.26) and s.t.  $u_{m_k} \rightarrow_k u$  where  $u$  also satisfies (6.26). Since we also have from [CdRM17] that  $u_{m_k}(t, \mathbf{x}) \rightarrow_k \mathbb{E}[g(\mathbf{X}_T^{t,\mathbf{x}})] + \int_t^T \mathbb{E}[f(s, \mathbf{X}_s^{t,\mathbf{x}})]ds$  where  $(\mathbf{X}_s^{t,\mathbf{x}})$  denotes the unique in law solution to (1.4), we deduce that  $u(t, \mathbf{x}) = \mathbb{E}[g(\mathbf{X}_T^{t,\mathbf{x}})] + \int_t^T \mathbb{E}[f(s, \mathbf{X}_s^{t,\mathbf{x}})]ds$  corresponds to the martingale, or *mild*, solution of (1.1) which satisfies the stated Schauder estimate (6.24).

### From mild to weak solutions

Let  $\varphi$  be a smooth given function with compact support, i.e.  $\varphi \in C_0^\infty(\mathbb{R}^{nd}, \mathbb{R})$ . It is clear that for the solution of (1.1) with mollified coefficients one indeed has:

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f_m(t, \mathbf{x}) \varphi(t, \mathbf{x}) = \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \varphi(t, \mathbf{x}) (\partial_t + L_t^m) u_m(t, \mathbf{x}). \quad (6.27)$$

Indeed, both the solution and the coefficients are smooth. Integrating by parts yields:

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f_m(t, \mathbf{x}) \varphi(t, \mathbf{x}) = \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left( -\partial_t + (L_t^m)^* \right) \varphi(t, \mathbf{x}) u_m(t, \mathbf{x}),$$

where  $(L_t^m)^*$  denotes the adjoint of  $L_t^m$ . Write now:

$$\begin{aligned} \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f_m(t, \mathbf{x}) \varphi(t, \mathbf{x}) &= \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f(t, \mathbf{x}) \varphi(t, \mathbf{x}) \\ &\quad + \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} (f_m - f)(t, \mathbf{x}) \varphi(t, \mathbf{x}) \\ &=: \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} f(t, \mathbf{x}) \varphi(t, \mathbf{x}) + R_m(T, f). \end{aligned} \quad (6.28)$$

It is clear that under **(A)**, recall that  $f \in L^\infty([0,T], C_{b,d}^\gamma(\mathbb{R}^{nd}, \mathbb{R}))$ ,  $R_m(T, f) \xrightarrow[m]{} 0$ . On the other hand, we now decompose:

$$\begin{aligned} &\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left( -\partial_t + (L_t^m)^* \right) \varphi(t, \mathbf{x}) u_m(t, \mathbf{x}) \\ &= \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left( -\partial_t + (L_t)^* \right) \varphi(t, \mathbf{x}) u(t, \mathbf{x}) + R_m(T, u), \end{aligned} \quad (6.29)$$

with

$$\begin{aligned}
 & R_m(T, u) \\
 := & \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left( (L_t^m)^* - L_t^* \right) \varphi(t, \mathbf{x}) u_m(t, \mathbf{x}) \\
 & + \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left( -\partial_t + (L_t)^* \right) \varphi(t, \mathbf{x}) (u_m(t, \mathbf{x}) - u(t, \mathbf{x})) =: (R_m^1 + R_m^2)(T, u),
 \end{aligned} \tag{6.30}$$

where  $L_t^*$  is the formal adjoint of  $L_t$ . Observe first that:

$$R_{m,0}^2(T, u) := \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \partial_t \varphi(t, \mathbf{x}) (u_m(t, \mathbf{x}) - u(t, \mathbf{x})) \xrightarrow{m} 0,$$

since  $\|u - u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \xrightarrow{m} 0$ . For the terms of  $R_m^2(T, u)$  which involve the adjoint, the point is again to use the Besov duality to control the remainders. Namely, from the previous analysis we get that:

$$\begin{aligned}
 & |R_{m,2:n}^2(T, u)| \\
 := & \left| \sum_{i=2}^n \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} D_{\mathbf{x}_i} (\varphi(t, \mathbf{x}) \mathbf{F}_i(t, \mathbf{x})) (u_m(t, \mathbf{x}) - u(t, \mathbf{x})) \right| \\
 \leq & \sum_{i=2}^n \int_0^T dt \int_{\mathbb{R}^{(n-1)d}} \prod_{j \neq i} d\mathbf{x}_j \|D_{\mathbf{x}_i} (\varphi \mathbf{F}_i(t, \mathbf{x}_{\setminus i}, \cdot))\|_{B_{1,1}^{-\frac{2+\gamma}{2i-1}}(\mathbb{R}^d, \mathbb{R})} \|(u_m - u)(t, \mathbf{x}_{\setminus i}, \cdot)\|_{C_b^{\frac{2+\gamma}{2i-1}}(\mathbb{R}^d, \mathbb{R})},
 \end{aligned}$$

denoting  $\mathbf{x}_{\setminus i} = (\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_n)$  and  $\varphi \mathbf{F}_i(t, \mathbf{x}_{\setminus i}, \cdot) : \mathbf{x}_i \in \mathbb{R}^d \mapsto \varphi \mathbf{F}_i(t, \mathbf{x})$ ,  $(u - u_m)(t, \mathbf{x}_{\setminus i}, \cdot) : \mathbf{x}_i \in \mathbb{R}^d \mapsto (u - u_m)(t, \mathbf{x})$ . Again  $\|u - u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \xrightarrow{m} 0$  and from the arguments of the proof of Lemma 5, there exists  $C$  s.t. for all  $i \in \llbracket 2, n \rrbracket$ ,

$$\|D_{\mathbf{x}_i} (\varphi \mathbf{F}_i(t, \mathbf{x}_{\setminus i}, \cdot))\|_{B_{1,1}^{-\frac{2+\gamma}{2i-1}}(\mathbb{R}^d, \mathbb{R})} \leq C \psi_i(t, \mathbf{x}_{\setminus i}),$$

where  $\psi_i$  has compact support on  $\mathbb{R}^{(n-1)d}$ . We thus readily derive  $|R_{m,2:n}^2(T, u)| \xrightarrow{m} 0$ .

Eventually,

$$\begin{aligned}
 & |R_{m,1}^2(T, u)| \\
 := & \left| \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left[ D_{\mathbf{x}_1} (\varphi(t, \mathbf{x}) \mathbf{F}_1(t, \mathbf{x})) + D_{\mathbf{x}_1}^2 (\varphi(t, \mathbf{x}) a(t, \mathbf{x})) \right] (u_m(t, \mathbf{x}) - u(t, \mathbf{x})) \right| \\
 = & \left| \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left[ (\varphi(t, \mathbf{x}) \mathbf{F}_1(t, \mathbf{x})) D_{\mathbf{x}_1} + (\varphi(t, \mathbf{x}) a(t, \mathbf{x})) D_{\mathbf{x}_1}^2 \right] (u_m(t, \mathbf{x}) - u(t, \mathbf{x})) \right|,
 \end{aligned}$$

which again tends to 0 with  $m$  since  $\|u - u_m\|_{L^\infty(C_{b,d}^{2+\gamma})} \xrightarrow{m} 0$ .

The contributions involving  $((L_t^m)^* - L_t^*) \varphi$  in  $R_m^1(T, u)$  defined in (6.30) can be handled as in the proof of Lemma 5 exploiting that  $\|a - a_m\|_{C_{b,d}^\gamma} + \|\mathbf{F}_1 - \mathbf{F}_{m,1}\|_{C_{b,d}^\gamma} + \sum_{i=2}^n \|\mathbf{F}_i -$

$\mathbf{F}_{m,i} \|_{C_{b,\mathbf{d}}^{2i-3+\gamma}} \rightarrow_m 0$ . We now deduce from (6.29), (6.30) and the previous controls that  $R_m(T, u) \rightarrow_m 0$ . The same computations also give that the term

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left( -\partial_t + (L_t)^* \right) \varphi(t, \mathbf{x}) u(t, \mathbf{x})$$

is well defined under **(A)**. From (6.28), (6.29), we thus finally derive:

$$\int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \varphi(t, \mathbf{x}) f(t, \mathbf{x}) = \int_0^T dt \int_{\mathbb{R}^{nd}} d\mathbf{x} \left( -\partial_t + (L_t)^* \right) \varphi(t, \mathbf{x}) u(t, \mathbf{x}),$$

which gives the statement.

## 7 Appendix: Proof of technical results

### 7.1 Technical results associated with the flow

We begin this paragraph stating and proving a key result for the sensitivity of Hölder flows, i.e. when the coefficients satisfy **(A)**. Those results are of course uniform w.r.t. a mollification procedure of the coefficients as the one previously considered from Section 3 to 5. Also, Lemma 2 is a direct consequence of Lemma 11 below and Young/convexity inequalities.

#### A first sensitivity result for the flow

**Lemma 11** *Under **(A)**, there exists  $C := C(\mathbf{(A)}, T)$  s.t. for all  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ ,  $0 \leq t < s \leq T$  and  $i \in \llbracket 1, n \rrbracket$ :*

$$|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \leq C \left( (s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x} - \mathbf{x}') \right).$$

The previous bound can be interpreted as follows. We somehow have the expected bound involving the spatial points, up to an additional contribution in time, which is precisely due to the pseudo-distance  $\mathbf{d}$ . Indeed, this type of result already appeared (for Lipschitz drifts) in Proposition 4.1 of [Men18]. Through an appropriate mollifying procedure, this result remains unchanged

**Proof:** The main idea to prove this control relies on Grönwall's Lemma. However, under **(A)**, the function  $\mathbf{F}$  is not Lipschitz (solely Hölder continuous). We have then to mollify suitably  $\mathbf{F}$ . Let us denote by  $\delta \in \mathbb{R}^n$ , a vector with positive entries  $\delta_i > 0$  for  $i \in \llbracket 2, n \rrbracket$ . Define as well for all  $v \in [0, T]$ ,  $\mathbf{z} \in \mathbb{R}^{nd}$ ,  $i \in \llbracket 2, n \rrbracket$ ,

$$\mathbf{F}_i^\delta(v, \mathbf{z}^{i-1:n}) := \mathbf{F}_i(v, \cdot) \star \rho_{\delta_i}(\mathbf{z}) = \int_{\mathbb{R}^d} \mathbf{F}_i(v, \mathbf{z}_{i-1}, \mathbf{z}_i - w, \mathbf{z}_{i+1} \cdots, \mathbf{z}_n) \rho_{\delta_i}(w) dw, \quad (7.1)$$

with  $\rho_{\delta_i}(w) := \frac{1}{\delta_i^d} \rho\left(\frac{w}{\delta_i}\right)$  where  $\rho : \mathbb{R}^d \rightarrow \mathbb{R}_+$  is a usual mollifier, namely  $\rho$  has compact support and  $\int_{\mathbb{R}^d} \rho(\mathbf{z}) d\mathbf{z} = 1$ . Finally, we define  $\mathbf{F}^\delta(v, \mathbf{z}) := (\mathbf{F}_1(v, \mathbf{z}), \mathbf{F}_2^\delta(v, \mathbf{z}), \dots, \mathbf{F}_n^\delta(v, \mathbf{z}))$ .

In this definition, we make a slight abuse of notation since the first component  $\mathbf{F}_1$  is not mollified. Due to the final control we want to prove and the intrinsic scale of the first component, the sublinearity of  $\mathbf{F}_1$  (implied by its Hölder property) is enough and it is not needed to mollify this component.

To be at the *good* current time scale for the contributions associated with the mollification, we pick  $\delta_i$  in order to have  $C := C((\mathbf{A}), T) > 0$  s.t. for all  $\mathbf{z} \in \mathbb{R}^{nd}$ ,  $u \in [t, s]$ :

$$\left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}(u, \mathbf{z}) - \mathbf{F}^\delta(u, \mathbf{z}) \right) \right| \leq C(s-t)^{-1}. \quad (7.2)$$

By the previous definition of  $\mathbf{F}^\delta$ , identity (7.2) is equivalent to:

$$\sum_{i=2}^n (s-t)^{\frac{1}{2}-i} \delta_i^{\frac{2i-3+\gamma}{2i-1}} \leq C(s-t)^{-1}. \quad (7.3)$$

Hence, we choose from now on, for all  $i \in \llbracket 2, n \rrbracket$ :

$$\delta_i = (s-t)^{\left(i-\frac{3}{2}\right) \frac{2i-1}{2i-3+\gamma}}. \quad (7.4)$$

Next, let us control the last components of the flow. By the definition of  $\boldsymbol{\theta}_{s,t}$  in (2.5), we get:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\ & \leq |(\mathbf{x} - \mathbf{x}')_n| + \int_t^s \left( |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| \right. \\ & \quad \left. + |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))| + |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}')) - \mathbf{F}_n(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| \right) dv \\ & \leq |(\mathbf{x} - \mathbf{x}')_n| + C \int_t^s \left( |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| + \delta_n^{-1+\frac{2n-3+\gamma}{2n-1}} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_n| \right) dv \\ & \quad + (s-t) \delta_n^{\frac{2n-3+\gamma}{2n-1}}. \end{aligned}$$

Hence by Grönwall's Lemma, we get:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \quad (7.5) \\ & \leq C \exp \left( C(s-t) \delta_n^{-1+\frac{2n-3+\gamma}{2n-1}} \right) \\ & \quad \times \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t) \delta_n^{\frac{2n-3+\gamma}{2n-1}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| dv \right) \\ & \leq C \exp \left( C(s-t)^{\frac{\gamma}{2}} \right) \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| dv \right), \end{aligned}$$

using (7.3) for the last inequality. We proceed similarly for the  $(n-1)^{\text{th}}$  component, but in this case we have to handle the non-Lipschitz continuity of  $\mathbf{F}_{n-1}^\delta$  in its  $n^{\text{th}}$  variable.

For the rescaled flows see e.g. Lemma 2 in [CdRM17] this difficulty could also be circumvented through mollification, the situation is here slightly different and it seems

that Young type controls are more appropriate. Write:

$$\begin{aligned}
 & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_{n-1}| \\
 \leq & C \exp\left(C(s-t)\delta_{n-1}^{-1+\frac{2(n-1)-3+\gamma}{2(n-1)-1}}\right) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)\delta_{n-1}^{\frac{2(n-1)-3+\gamma}{2(n-1)-1}} \right. \\
 & \left. + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| + |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_n| \right\} dv \right) \\
 \leq & C \exp(C(s-t)^{\frac{\gamma}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| \right. \right. \\
 & \left. \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}} + (v-t)^{\frac{2(n-1)-3+\gamma}{2}} \right. \right. \\
 & \left. \left. + \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\frac{2(n-1)-3+\gamma}{2n-1}} \right\} dv \right).
 \end{aligned} \tag{7.6}$$

The last inequality is a consequence of our choice of  $\delta_{n-1}$  in (7.4), identity (7.5) and convexity inequality.

We aim, now, to proceed with Grönwall's Lemma. To do so, first of all we need to use a Young inequality. Namely, we write for all  $\tilde{\delta}_{n-1,n} > 0$  (where the two indexes in the subscript respectively denote the stage of the chain, i.e.  $n-1$ , and the considered variable, i.e.  $n$ ):

$$\begin{aligned}
 & \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\frac{2(n-1)-3+\gamma}{2n-1}} \\
 \leq & C \left( \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right) \tilde{\delta}_{n-1,n}^{-\frac{2n-1}{2(n-1)-3+\gamma}} + \tilde{\delta}_{n-1,n}^{\frac{2n-1}{4-\gamma}} \right).
 \end{aligned}$$

In order to obtain the suitable time scale, we choose  $\tilde{\delta}_n$  s.t.

$$\tilde{\delta}_{n-1,n}^{\frac{2n-1}{4-\gamma}} = (v-t)^{\frac{2(n-1)-3+\gamma}{2}} \iff \tilde{\delta}_{n-1,n} = (v-t)^{\frac{(2(n-1)-3+\gamma)(4-\gamma)}{2(2n-1)}},$$

which also yields that

$$\left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right) \tilde{\delta}_{n-1,n}^{-\frac{2n-1}{2(n-1)-3+\gamma}} \leq \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right) (v-t)^{-\frac{4-\gamma}{2}}.$$

Hence we get from (7.6) and the previous controls that for all  $\tilde{s} \in [t, s]$ :

$$\begin{aligned}
 & |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_{n-1}| \\
 \leq & C \exp(C(\tilde{s}-t)^{\frac{\gamma}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (\tilde{s}-t)^{n-\frac{3}{2}} + \int_t^{\tilde{s}} \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| \right. \right. \\
 & \left. \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}} + (v-t)^{\frac{2(n-1)-3+\gamma}{2}} \right. \right. \\
 & \left. \left. + \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right) (v-t)^{-\frac{4-\gamma}{2}} \right\} dv \right)
 \end{aligned}$$

The point is now to take the supremum in  $\tilde{s} \in [s, t]$  in the above equation. This yields:

$$\begin{aligned} & \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp(C(s-t)^{\frac{\gamma}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}} (s-t) + (s-t)^{1+\frac{2(n-1)-3+\gamma}{2}} \\ & \quad \left. + \left( \int_t^s \sup_{w \in [t, v]} |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| (v-t)^{1-\frac{4-\gamma}{2}} dv \right) \right). \end{aligned}$$

We get then by Grönwall's Lemma:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp(C(s-t)^{\frac{\gamma}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}} (s-t) \right) \end{aligned} \quad (7.7)$$

$$\begin{aligned} & \leq C \exp(C(s-t)^{\frac{\gamma}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} \right), \end{aligned} \quad (7.8)$$

using again the Young inequality  $|(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1}} (s-t) \leq C((s-t)^{n-\frac{3}{2}} + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-3+\gamma}{2n-1} \frac{2n-3}{2n-5}})$  for the last identity, recalling as well that  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$ , and therefore  $|(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} \leq 1$ , for the last identity. The purpose of (7.8) is that each entry of the difference of the starting points appears at its intrinsic scale for the homogeneous distance  $\mathbf{d}$ .

Plugging the above inequality into (7.5) we derive:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\ & \leq C \exp\left(C(s-t)^{\frac{\gamma}{2}}\right) \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + |(\mathbf{x} - \mathbf{x}')_{n-1}|(s-t) \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} (s-t) + \int_t^s \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-2}| dw dv \right) \\ & \leq C \exp\left(C(s-t)^{\frac{\gamma}{2}}\right) \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + |(\mathbf{x} - \mathbf{x}')_{n-1}|^{\frac{2n-1}{2n-3}} \right. \\ & \quad \left. + \int_t^s \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-2}| dw dv \right), \end{aligned}$$

using again the Young inequalities  $|(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} (s-t) \leq C(|(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}})$  and  $|(\mathbf{x} - \mathbf{x}')_{n-1}|(s-t) \leq C(|(\mathbf{x} - \mathbf{x}')_{n-1}|^{\frac{2n-1}{2n-3}} + (s-t)^{n-\frac{1}{2}})$  for the last bound. Iterating these computations, we obtain:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \quad (7.9) \\ & \leq C \left( (s-t)^{n-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x} - \mathbf{x}')_j|^{\frac{2n-1}{2j-1}} + \int_t^{v_n=s} dv_{n-1} \dots \int_t^{v_2} dv_1 |(\boldsymbol{\theta}_{v_1,t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1,t}(\mathbf{x}))_1| \right). \end{aligned}$$

Similarly, for  $i \in \llbracket 2, n \rrbracket$  we derive:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| & (7.10) \\ & \leq C \left( (s-t)^{i-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x} - \mathbf{x}')_j|^{\frac{2i-1}{2j-1}} + \int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_2} dv_1 |(\boldsymbol{\theta}_{v_1,t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1,t}(\mathbf{x}))_1| \right). \end{aligned}$$

**Remark 8** We importantly point out that equations (7.9) and (7.10) actually hold not only for the fixed time  $s$  but also for any  $v \in [t, T]$ .

The term for  $i = 1$  is treated slightly differently. Namely, for all  $\tilde{s} \in [t, s]$ , write:

$$|(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \leq |(\mathbf{x} - \mathbf{x}')_1| + C \sum_{j=1}^n \int_t^{\tilde{s}} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{\gamma}{2j-1}} dv,$$

which in turn implies

$$\begin{aligned} \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| & \leq |(\mathbf{x} - \mathbf{x}')_1| + C \left( (s-t) \left( \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \right)^\gamma \right. \\ & \quad \left. + \sum_{j=2}^n \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{\gamma}{2j-1}} dv \right) \\ & \leq |(\mathbf{x} - \mathbf{x}')_1| + C \left( (s-t) \left( \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \right)^\gamma \right. \\ & \quad \left. + \sum_{j=2}^n C_j (s-t) \left( (s-t)^{j-\frac{1}{2}} + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{2j-1}{2k-1}} \right. \right. \\ & \quad \left. \left. + (s-t)^{j-1} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1| \right)^{\frac{\gamma}{2j-1}} \right) \end{aligned}$$

and yields that

$$\begin{aligned} \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| & \leq |(\mathbf{x} - \mathbf{x}')_1| + C \left( (s-t) \left( \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \right)^\gamma \right. \\ & \quad \left. + \sum_{j=2}^n C_j (s-t) \left( (s-t)^{\frac{\gamma}{2}} + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{\gamma}{2k-1}} \right. \right. \\ & \quad \left. \left. + (s-t)^{(j-1)\frac{\gamma}{2j-1}} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|^{\frac{\gamma}{2j-1}} \right) \right), \end{aligned}$$

using as well (7.10), Remark 8 and convexity inequalities for the last bound. We now write,

$$\begin{aligned} \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| & \leq C \left( |(\mathbf{x} - \mathbf{x}')_1| + (s-t)^{1+\frac{\gamma}{2}} + (s-t) \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{\gamma}{2k-1}} \right. \\ & \quad \left. + \sum_{j=1}^n (s-t)^{1+(j-1)\frac{\gamma}{2j-1}} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|^{\frac{\gamma}{2j-1}} \right) \\ & \leq C \left( |(\mathbf{x} - \mathbf{x}')_1| + (s-t) + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}} \right) \quad (7.11) \end{aligned}$$



recalling that  $(s - t) \leq T$  is small, and using again Young inequalities for the last bound. Namely,

$$(s - t)^{1+(j-1)\frac{\gamma}{2j-1}} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|^{\frac{\gamma}{2j-1}} \leq C(s - t) \left( 1 + \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1| \right),$$

and

$$(s - t)|(\mathbf{x} - \mathbf{x}')_k|^{\frac{\gamma}{2k-1}} \leq C((s - t)^{\frac{1}{1-\gamma}} + |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}}).$$

We eventually derive from (7.11) that:

$$\sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1| \leq C((s - t)^{\frac{1}{2}} + \mathbf{d}(\mathbf{x}, \mathbf{x}')),$$

which gives the stated bound for  $i = 1$ . It now remains to plug this control into (7.10). We obtain for all  $i \in \llbracket 2, n \rrbracket$ :

$$\begin{aligned} |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| &\leq C((s - t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (s - t)^{i-1} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1|) \\ &\leq C\left((s - t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (s - t)^{i-1}((s - t)^{\frac{1}{2}} + \mathbf{d}(\mathbf{x}, \mathbf{x}'))\right) \\ &\leq C((s - t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}')), \end{aligned}$$

using again the Young inequality to derive that  $(s - t)^{i-1} \mathbf{d}(\mathbf{x}, \mathbf{x}') \leq C((s - t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}'))$ . The proof is complete.  $\square$

Again, Lemma 2 is a direct consequence of the previous Lemma 11 and Young/convexity inequalities.

We are now in position to prove the sensitivity results for the linearized system w.r.t. the freezing parameter.

### Sensitivity results for the mean

#### Proof: Technical Lemma 3

We assume w.l.o.g. that  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$ . The idea of the proof is to separate the term to control into two contributions. Namely, we write:

$$\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}') = [\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x})] + [\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}')]. \quad (7.12)$$

The definition of the proxy (3.1) yields that the mean value of  $\tilde{\mathbf{X}}_v^{m, \xi}$ ,  $\mathbf{m}_{v,t}^{\xi}$  is s.t.

$$\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}) = \mathbf{x}' - \mathbf{x} + \int_t^s dv D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{v,t}(\mathbf{x})]. \quad (7.13)$$

The sub-triangular structure of  $D\mathbf{F}$  yields that for all  $i \in \llbracket 2, n \rrbracket$ :

$$(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_i = \mathbf{x}'_i - \mathbf{x}_i + \int_t^s dv D_{i-1} \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_{i-1} - \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1}].$$

Also, since  $\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_1 = \mathbf{x}'_1 + \int_t^s \mathbf{F}_1(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) dv$ , so that  $[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_1 - \boldsymbol{\theta}_{v,t}(\mathbf{x})_1] = \mathbf{x}'_1 - \mathbf{x}_1$ , we then obtain by iteration that:

$$\begin{aligned} & (\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_i \\ &= \mathbf{x}'_i - \mathbf{x}_i + \sum_{k=2}^i \left[ \int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^i D_{j-1} \mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})) \right] [\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}], \end{aligned}$$

with the convention that for  $i = 1$ ,  $\sum_{k=2}^i = 0$ . From the above control, equation (7.12) and the dynamics of the flow, recalling that the starting points are the same, so that the contributions involving differences of the spatial points or flows only appear in iterated time integrals, we derive:

$$\begin{aligned} & |(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \\ &\leq \left| \sum_{k=2}^i \left[ \int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^i D_{j-1} \mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})) \right] [\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}] \right| \\ &\quad + \int_t^s |\mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| dv \tag{7.14} \\ &\leq C \left( \sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| \right. \\ &\quad \left. + \int_t^s \left( \sum_{j=i}^n |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{2i-3+\gamma}{2j-1}} + |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{i-1}| \right) dv \right). \end{aligned}$$

From the previous Lemma 11, we thus obtain:

$$\begin{aligned} & |(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \\ &\leq C \left( \sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| + (s-t)^{\frac{2i-3+\gamma}{2}+1} \right. \\ &\quad \left. + \mathbf{d}^{2i-3+\gamma}(\mathbf{x}, \mathbf{x}') (s-t) + ((s-t)^{(i-1)-\frac{1}{2}} + \mathbf{d}^{2(i-1)-1}(\mathbf{x}, \mathbf{x}')) (s-t) \right). \tag{7.15} \end{aligned}$$

In particular, for  $s = t_0 = t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  with  $c_0 < 1$ , the previous equation yields:

$$\begin{aligned} & |(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_i| \\ &\leq C \left( c_0 \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (c_0^{\frac{i-\frac{1}{2}+\frac{\gamma}{2}}{2}} + c_0) \mathbf{d}^{2i-1+\gamma}(\mathbf{x}, \mathbf{x}') + (c_0^{\frac{i-\frac{1}{2}}{2}} + c_0) \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \right). \end{aligned}$$

So, after summing and by convexity inequalities, for  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$ :

$$\mathbf{d}(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}'), \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')) \leq C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}').$$

□

## 7.2 Sensitivities for the scaled flows

For the scaling analysis of Section 6 we also need the scaled versions of the previous Lemmas. Recalling the notations introduced therein, i.e. for  $\lambda > 0$ ,  $0 \leq t \leq v \leq T$ ,  $\boldsymbol{\theta}_{v,t}^\lambda(\mathbf{x}^\lambda) = \lambda^{\frac{1}{2}} \mathbb{T}_\lambda^{-1} \boldsymbol{\theta}_{v,t}(\mathbf{x}^\lambda)$ ,  $\mathbf{x}^\lambda := \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \mathbf{x}$ , as corollaries of Lemmas 11 and 2, we readily get:

$$\mathbf{d}(\boldsymbol{\theta}_{v,t}^\lambda(\mathbf{x}^\lambda), \boldsymbol{\theta}_{v,t}^\lambda(\mathbf{x}'^\lambda)) \leq C \left( \mathbf{d}(\mathbf{x}, \mathbf{x}') + \left( \frac{v-t}{\lambda} \right)^{\frac{1}{2}} \right). \quad (7.16)$$

Setting now  $v = t + c_0 \lambda \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ , recalling that

$$\mathbf{d}(\lambda^{\frac{1}{2}} \mathbb{T}_\lambda^{-1} \mathbf{m}_{v,t}^{\mathbf{x}^\lambda}(\mathbf{x}'^\lambda), \boldsymbol{\theta}_{v,t}^\lambda(\mathbf{x}'^\lambda)) = \lambda^{-\frac{1}{2}} \mathbf{d}(\mathbf{m}_{v,t}^{\mathbf{x}^\lambda}(\mathbf{x}'^\lambda), \boldsymbol{\theta}_{v,t}^\lambda(\mathbf{x}'^\lambda)), \quad (7.17)$$

and reproducing the above arguments, we derive from (7.16) and (7.15) the bound:

$$\mathbf{d}(\lambda^{\frac{1}{2}} \mathbb{T}_\lambda^{-1} \mathbf{m}_{v,t}^{\mathbf{x}^\lambda}(\mathbf{x}'^\lambda), \boldsymbol{\theta}_{v,t}^\lambda(\mathbf{x}'^\lambda)) \leq C (c_0^{\frac{1}{2n-1}} + \lambda^{\frac{\gamma}{2}}) \mathbf{d}(\mathbf{x}, \mathbf{x}'). \quad (7.18)$$

which follows exploiting the intrinsic scaling properties of system.

Observe that the scaling procedure does not provide for the mean sensitivity a small constant (i.e. proportional to  $\lambda^{\frac{\gamma}{2}}$ ). Indeed, in inequality (7.14), the terms involving  $|\mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))|$  and the products  $(\prod_{j=k}^i D_{j-1} \mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})))_{k \in \llbracket 2, i-1 \rrbracket}$  do not allow us to take advantage from the  $1 + \frac{\gamma}{2i-3}$ -Hölder continuity of  $\mathbf{F}_i$  w.r.t the  $(i-1)^{\text{th}}$  spatial component nor from the  $\frac{\gamma}{2j-1}$ -Hölder continuity of  $D_{j-1} \mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x}))$  w.r.t. the  $(i-1)^{\text{th}}$  spatial component. We can only exploit respectively the Lipschitz continuity and the boundedness of such terms which are precisely those who do not scale. Plugging  $\mathbf{x}^\lambda = \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \mathbf{x}$ ,  $\mathbf{x}'^\lambda = \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \mathbf{x}'$  in (7.15) indeed yields for  $v = t + c_0 \lambda \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ :

$$\begin{aligned} |(\mathbf{m}_{v,t}^{\mathbf{x}^\lambda}(\mathbf{x}'^\lambda) - \boldsymbol{\theta}_{v,t}^\lambda(\mathbf{x}'^\lambda))_i| &\leq \lambda^{i-\frac{1}{2}} \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') C \left( \sum_{k=2}^{i-1} (c_0)^{i-k} + c_0^{i-\frac{1}{2}} (c_0 \lambda \mathbf{d}^2(\mathbf{x}, \mathbf{x}'))^{\frac{\gamma}{2}} \right. \\ &\quad \left. + \lambda^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') c_0 + (c_0^{i-\frac{1}{2}} + c_0) \right), \end{aligned}$$

which from (7.17) yields (7.18) recalling that  $c_0$  and  $\mathbf{d}(\mathbf{x}, \mathbf{x}')$  are both less than 1.

## 8 Appendix: Sensitivity results for the resolvent and covariance

### 8.1 Sensitivity Lemma for the resolvent

**Lemma 12 (Controls of the Sensitivities for the Resolvents)** *There exists  $\tilde{C}$  s.t. for all  $0 \leq t \leq s \leq T$ ,  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ , the following control holds. For all  $1 \leq j < i \leq n$ , with the notation of (3.20):*

$$|(\tilde{\mathbf{R}}^{(\tau, \mathbf{x})}(t, s) - \tilde{\mathbf{R}}^{(\tau, \mathbf{x}')} (t, s))_{i,j}| \leq \tilde{C} (s-t)^{i-j} \left( \sum_{k=2}^n \|\mathbf{F}_k\|_{L^\infty(C_{\mathbf{d}}^{2k-3+\gamma})} \right) ((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')).$$

**Proof: Lemma 12**

From the scaling properties of the resolvent, see e.g. the proof of Proposition 2 or Lemma 6.2 in [Men18], we have that:

$$\tilde{\mathbf{R}}^{\mathbf{x}}(s, t) = \mathbb{T}_{s-t}^{-1} \widehat{\mathbf{R}}_1^{s,t,\mathbf{x}} \mathbb{T}_{s-t}^{-1}, \quad \tilde{\mathbf{R}}^{\mathbf{x}'}(s, t) = \mathbb{T}_{s-t}^{-1} \widehat{\mathbf{R}}_1^{s,t,\mathbf{x}'} \mathbb{T}_{s-t}^{-1}, \quad (8.1)$$

where  $\widehat{\mathbf{R}}_1^{s,t,\mathbf{x}}, \widehat{\mathbf{R}}_1^{s,t,\mathbf{x}'}$  are non-degenerate bounded matrices. We define then:

$$\Delta \widehat{\mathbf{R}}_1^{s,t,\mathbf{x},\mathbf{x}'} := \widehat{\mathbf{R}}_1^{s,t,\mathbf{x}} - \widehat{\mathbf{R}}_1^{s,t,\mathbf{x}'}. \quad (8.2)$$

Hence, from (8.1) and the definitions in (8.2):

$$\begin{aligned} |\Delta \widehat{\mathbf{R}}_1^{s,t,\mathbf{x},\mathbf{x}'}| &= |\widehat{\mathbf{R}}_1^{s,t,\mathbf{x}} - \widehat{\mathbf{R}}_1^{s,t,\mathbf{x}'}| = |\mathbb{T}_{s-t}^{-1}(\tilde{\mathbf{R}}^{\mathbf{x}} - \tilde{\mathbf{R}}^{\mathbf{x}'})\mathbb{T}_{s-t}| \\ &= \left| \mathbb{T}_{s-t}^{-1} \int_t^s \left( D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) \tilde{\mathbf{R}}^{\mathbf{x}}(v, t) - D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}')) \tilde{\mathbf{R}}^{\mathbf{x}'}(v, t) \right) dv \mathbb{T}_{s-t} \right| \\ &\leq \int_t^s |\mathbb{T}_{s-t}^{-1} D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) \mathbb{T}_{s-t}| |\mathbb{T}_{s-t}^{-1}(\tilde{\mathbf{R}}^{\mathbf{x}} - \tilde{\mathbf{R}}^{\mathbf{x}'})\mathbb{T}_{s-t}| dv \\ &\quad + \int_t^s \left| \mathbb{T}_{s-t}^{-1} \left( D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}')) \right) \mathbb{T}_{s-t} \right| |\mathbb{T}_{s-t}^{-1} \tilde{\mathbf{R}}^{\mathbf{x}'}(v, t) \mathbb{T}_{s-t}| dv \\ &\leq C \int_t^s (s-t)^{-1} |D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| dv, \end{aligned}$$

using the Grönwall's Lemma and the structure of the resolvent for the last inequality.

Pay attention that we only know from our smoothness assumption **(S)** that for all  $i \in \llbracket 2, n \rrbracket$ ,  $\forall \mathbf{z}^{i:n} = (\mathbf{z}_i, \dots, \mathbf{z}_n) \in \mathbb{R}^{(n-i+1)d}$ ,  $\mathbf{z}_{i-1} \mapsto D_{\mathbf{x}_{i-1}} \mathbf{F}_i(\mathbf{z}_{i-1}, \mathbf{z}^{i:n})$  is  $C^{\frac{\gamma}{2(i-1)-1}}(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ -Hölder continuous for  $\eta > 0$ . Hence, we proceed carefully like in [CdRM17].

We directly get that:

$$\begin{aligned} |\Delta \widehat{\mathbf{R}}_1^{s,t,\mathbf{x},\mathbf{x}'}| &= |\widehat{\mathbf{R}}_1^{s,t,\mathbf{x}} - \widehat{\mathbf{R}}_1^{s,t,\mathbf{x}'}| = |\mathbb{T}_{s-t}^{-1}(\tilde{\mathbf{R}}^{\mathbf{x}} - \tilde{\mathbf{R}}^{\mathbf{x}'})\mathbb{T}_{s-t}| \\ &\leq C \int_t^s (s-t)^{-1} \sum_{i=2}^n \left( |D_{i-1} \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - D_{i-1} \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))_{i-1}, (\boldsymbol{\theta}_{v,t}(\mathbf{x}'))^{i:n}| \right. \\ &\quad \left. + \|(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1}\|_{L^\infty(C_d^\gamma)} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{i-1}|^{\eta_i} \right) dv \\ &=: (R_1 + R_2)(s, t, \mathbf{x}, \mathbf{x}'), \end{aligned} \quad (8.3)$$

where  $\eta_i := \frac{\gamma}{2(i-1)-1}$  and the notation  $(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1}$  indicates that  $D_{\mathbf{x}_{i-1}} \mathbf{F}_i$  is viewed as a function of its variable  $(i-1)$  and the supremum is taken over the other ones. From Lemma 2 and the definition of  $\mathbf{d}$  (see also Lemma 11), we readily get

$$\begin{aligned} |R_2(s, t, \mathbf{x}, \mathbf{y})| &\leq C \|D\mathbf{F}\|_{L^\infty(C_d^\gamma)} \sum_{i=2}^n \int_t^s (s-t)^{-1} \left( (v-t)^{(i-1-\frac{1}{2})} + \mathbf{d}^{2(i-1)-1}(\mathbf{x}, \mathbf{x}') \right)^{\eta_i} dv \\ &\leq C \|D\mathbf{F}\|_{L^\infty(C_d^\gamma)} \left( (s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}') \right), \end{aligned} \quad (8.4)$$

denoting with a slight abuse of notation  $\|D\mathbf{F}\|_{L^\infty(C_d^\gamma)} := \sum_{i=2}^n \|(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1}\|_{L^\infty(C_d^\gamma)}$ . To control the difference of the gradients terms in  $R_1(s, t, \mathbf{x}, \mathbf{x}')$  in (8.3), we need the following result whose proof is postponed to Appendix 8.3.

**Lemma 13 (Reverse Taylor expansion)** *There is a constant  $C > 0$  s.t, for all  $(\mathbf{z}, \mathbf{z}') \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$ ,  $v \in [0, T]$  :*

$$|D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}')| \leq C \|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})} \mathbf{d}(\mathbf{z}, \mathbf{z}')^\gamma,$$

with the notations of equation (3.20).

From the reverse Taylor expansion of Lemma 13 and the definition in (8.3), we obtain:

$$\begin{aligned} |R_1(s, t, \mathbf{x}, \mathbf{y})| &\leq C(s-t)^{-1} \int_t^s dv \sum_{i=2}^n \|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})} \mathbf{d}^\gamma(\boldsymbol{\theta}_{v,t}(\mathbf{x}), \boldsymbol{\theta}_{v,t}(\mathbf{x}')) \\ &\leq C \sum_{i=2}^n \|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})} ((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')), \end{aligned} \quad (8.5)$$

using again Lemma 2 for the last inequality. Gathering (8.4), (8.5) into (8.3) and recalling the definition in (3.20), we obtain:

$$|\widehat{\Delta \tilde{\mathbf{R}}}_1^{s,t,\mathbf{x},\mathbf{x}'}| \leq C \left( \sum_{i=2}^n \|\mathbf{F}_i\|_{L^\infty(C_{\mathbf{d}}^{2i-3+\gamma})} \right) ((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\mathbf{x}, \mathbf{x}')). \quad (8.6)$$

The result follows from the previous bound, the definition in (8.2) and the scalings of equation (8.1).  $\square$

## 8.2 Sensitivity Lemma for the covariances

### Proof of Lemma 1

Let us first explicitly write the covariance matrices

$$\tilde{\mathbf{K}}_{s,t}^\xi := \int_t^s \tilde{\mathbf{R}}^\xi(s, u) B a(u, \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi})) B^* \tilde{\mathbf{R}}^\xi(s, u)^* du. \quad (8.7)$$

So we have to control the term

$$\begin{aligned} \tilde{\mathbf{K}}_{s,t}^\xi - \tilde{\mathbf{K}}_{s,t}^{\xi'} &= \Delta_1^{\xi, \xi'}(s, t) + \Delta_2^{\xi, \xi'}(s, t), \\ \Delta_1^{\xi, \xi'}(s, t) &:= \int_t^s du \tilde{\mathbf{R}}^\xi(s, u) B \Delta_u a(\boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}), \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}')) B^* \tilde{\mathbf{R}}^\xi(s, u)^*, \\ \Delta_u a(\boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}), \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}')) &:= a(u, \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi})) - a(u, \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}')), \\ \Delta_2^{\xi, \xi'}(s, t) &:= \int_t^s du \Delta \tilde{\mathbf{R}}^{\xi, \xi'}(s, u) B a(u, \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}')) B^* \tilde{\mathbf{R}}^\xi(s, u)^* \\ &\quad + \int_t^s du \tilde{\mathbf{R}}^{\xi'}(s, u) B a(u, \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}')) B^* \Delta \tilde{\mathbf{R}}^{\xi, \xi'}(s, u)^*, \\ \Delta \tilde{\mathbf{R}}^{\xi, \xi'}(s, u) &= \tilde{\mathbf{R}}^\xi(s, u) - \tilde{\mathbf{R}}^{\xi'}(s, u). \end{aligned} \quad (8.8)$$

Hence, from the scalings of (8.1) and the definitions in (8.8), for all  $1 \leq j \leq i \leq n$ :

$$\begin{aligned} |(\Delta_1^{\xi, \xi'}(s, t))_{i,j}| &\leq C(s-t)^{-2} \int_t^s |(\mathbb{T}_{s-t} \Delta_u B a(\boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}), \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}')) B^* \mathbb{T}_{s-t})_{i,j}| du \\ &\leq C(s-t)^{i+j-2} \int_t^s \mathbf{d}(\boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}), \boldsymbol{\theta}_{u,t}(\boldsymbol{\xi}'))^\gamma du. \end{aligned}$$

We deduce by Lemma 2, that

$$|(\Delta_1^{\xi, \xi'}(s, t))_{i,j}| \leq C(s-t)^{i+j-1}((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\xi, \xi')). \quad (8.9)$$

Still from (8.1) and the definitions in (8.8), write now that:

$$\begin{aligned} & |(\Delta_2^{\xi, \xi'}(s, t))_{i,j}| \\ & \leq C(s-t)^{-2} \left( \int_t^s \left( |(\mathbb{T}_{s-t}(\widehat{\mathbf{R}}_1^{s,u,\xi} - \widehat{\mathbf{R}}_1^{s,u,\xi'}) Ba(u, \boldsymbol{\theta}_{u,t}(\xi')) B^* (\widehat{\mathbf{R}}_1^{s,u,\xi'})^* \mathbb{T}_{s-t})_{i,j}| \right. \right. \\ & \quad \left. \left. + |(\mathbb{T}_{s-t} \widehat{\mathbf{R}}_1^{s,u,\xi'} Ba(u, \boldsymbol{\theta}_{u,t}(\xi')) B^* (\widehat{\mathbf{R}}_1^{s,u,\xi'} - \widehat{\mathbf{R}}_1^{s,u,\xi})^* \mathbb{T}_{s-t})_{i,j}| \right) du \right). \end{aligned} \quad (8.10)$$

Thanks to equation (8.6) in the proof of Lemma 12, we thus obtain:

$$|(\Delta_2^{\xi, \xi'}(s, t))_{i,j}| \leq C(s-t)^{i+j-1}((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\xi, \xi')). \quad (8.11)$$

Gathering (8.9) and (8.11) in (8.8) yields:

$$|(\tilde{\mathbf{K}}_{s,t}^\xi - \tilde{\mathbf{K}}_{s,t}^{\xi'})_{i,j}| \leq C(s-t)^{i+j-1}((s-t)^{\frac{\gamma}{2}} + \mathbf{d}^\gamma(\xi, \xi')), \quad (8.12)$$

which precisely gives (3.18) for  $i = j = 1$ .  $\square$

### Sensitivities for the scaled covariance.

In connection with Section 6, we recall the identity in law (6.6), i.e.  $\tilde{\mathbf{X}}_v^{\xi, \lambda} := \lambda^{1/2} \mathbb{T}_\lambda^{-1} \tilde{\mathbf{X}}_v^{\xi^\lambda}$ ,  $v \in [t, T]$ , which readily gives:

$$\tilde{\mathbf{K}}_{v,t}^{\xi, \lambda} := \text{Cov}(\tilde{\mathbf{X}}_v^{\xi, \lambda}) = \lambda \mathbb{T}_\lambda^{-1} \tilde{\mathbf{K}}_{v,t}^{\xi^\lambda} \mathbb{T}_\lambda^{-1}.$$

In particular, we thus derive from the analysis of the previous paragraph:

$$\begin{aligned} [\tilde{\mathbf{K}}_{v,t}^{\xi, \lambda}]_{1,1} & \leq C\lambda^{-1}(v-t), \\ |[\tilde{\mathbf{K}}_{v,t}^{\xi, \lambda}]_{1,1} - [\tilde{\mathbf{K}}_{v,t}^{\xi', \lambda}]_{1,1}| & \leq C\lambda^{-1}(v-t) \left( \mathbf{d}^\gamma(\xi^\lambda, \xi'^\lambda) + (v-t)^{\frac{\gamma}{2}} \right) \\ & \leq C\lambda^{-1}(v-t) \left( \lambda^{\frac{\gamma}{2}} \mathbf{d}^\gamma(\xi, \xi') + (v-t)^{\frac{\gamma}{2}} \right), \end{aligned}$$

recalling that  $\xi^\lambda = \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \xi$  and using the homogeneity properties of  $\mathbf{d}$  for the last inequality. Also, for  $v = t + c_0 \lambda \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  and taking  $\xi = \mathbf{x}$ ,  $\xi' = \mathbf{x}'$ , the above controls rewrite:

$$[\tilde{\mathbf{K}}_{v,t}^{\xi, \lambda}]_{1,1} \leq Cc_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}'), \quad |[\tilde{\mathbf{K}}_{v,t}^{\xi, \lambda}]_{1,1} - [\tilde{\mathbf{K}}_{v,t}^{\xi', \lambda}]_{1,1}| \leq Cc_0 \lambda^{\frac{\gamma}{2}} \mathbf{d}^{2+\gamma}(\mathbf{x}, \mathbf{x}'). \quad (8.13)$$

Note that the sensitivity of the scaled covariance yields a contribution of the scaling coefficient in  $\lambda^{\frac{\gamma}{2}}$ . Unlike for the control of the scaled mean in (7.18), where, as previously noticed, we could not exploit the full regularity of  $\mathbf{F}_i$  w.r.t. the  $(i-1)^{\text{th}}$  variable, we can here precisely take advantage of such a regularity. Indeed, this follows from the expression of the covariance (8.7) which only involves  $D_{i-1} \mathbf{F}_i$  so that one can exploit the associated  $\frac{\gamma}{2i-3}$ -Hölder regularity w.r.t.  $\mathbf{x}_{i-1}$ .

### 8.3 Reverse Taylor formula

**Proof: Lemma 13**

We assume here, for the sake of simplicity and without loss of generality, that  $d = 1$  (scalar case). When  $d > 1$ , the proof below can be reproduced componentwise. Let us decompose the expression around the variables which do/do not transmit the noise. Namely, we write for all  $\delta_i > 0$ :

$$\begin{aligned}
 & D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}') \\
 = & \int_0^1 d\mu \{ D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}_{i-1} + \mu \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i}, \mathbf{z}^{i:n}) \} \\
 & + \{ D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}_{i-1} + \mu \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i}, (\mathbf{z}')^{i:n}) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}') \} \\
 & + \{ D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}_{i-1} + \mu \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i}, \mathbf{z}^{i:n}) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(v, \mathbf{z}_{i-1} + \mu \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i}, (\mathbf{z}')^{i:n}) \} \\
 =: & \sum_{\ell=1}^3 \Delta \mathbf{F}_i^\ell(v, \mathbf{z}, \mathbf{z}'). \tag{8.14}
 \end{aligned}$$

The first two terms can be dealt directly. From (A) we get:

$$|\Delta \mathbf{F}_i^1(v, \mathbf{z}, \mathbf{z}')| \leq \| (D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1} \|_{L^\infty(C_d^\gamma)} \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i \frac{\gamma}{2(i-1)-1}}. \tag{8.15}$$

Similarly,

$$\begin{aligned}
 |\Delta \mathbf{F}_i^2(v, \mathbf{z}, \mathbf{z}')| & \leq \| (D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1} \|_{L^\infty(C_d^\gamma)} (\mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i \frac{\gamma}{2(i-1)-1}} + |(\mathbf{z} - \mathbf{z}')_{i-1}|^{\frac{\gamma}{2(i-1)-1}}) \\
 & \leq C \| (D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i-1} \|_{L^\infty(C_d^\gamma)} \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i \frac{\gamma}{2(i-1)-1}}. \tag{8.16}
 \end{aligned}$$

For  $\Delta \mathbf{F}_i^3(t, \mathbf{z}, \mathbf{z}')$ , we use an explicit reverse Taylor expansion which yields together with the smoothness assumption of  $\mathbf{F}_i$  in (A):

$$\begin{aligned}
 |\Delta \mathbf{F}_i^3(t, \mathbf{z}, \mathbf{z}')| & = \mathbf{d}(\mathbf{z}, \mathbf{z}')^{-\delta_i} \left\| \left[ \mathbf{F}_i(t, \mathbf{z}_{i-1} + \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i}, \mathbf{z}^{i:n}) - \mathbf{F}_i(t, \mathbf{z}_{i-1} + \mathbf{d}(\mathbf{z}, \mathbf{z}')^{\delta_i}, (\mathbf{z}')^{i:n}) \right. \right. \\
 & \quad \left. \left. + \mathbf{F}_i(t, \mathbf{z}_{i-1}, \mathbf{z}^{i:n}) - \mathbf{F}_i(t, \mathbf{z}_{i-1}, (\mathbf{z}')^{i:n}) \right] \right\| \\
 & \leq 2 \| \mathbf{F}_i \|_{L^\infty(C_d^{2i-3+\gamma})} \mathbf{d}(\mathbf{z}, \mathbf{z}')^{2i-3+\gamma-\delta_i}. \tag{8.17}
 \end{aligned}$$

Taking  $\delta_i$  s.t.  $\delta_i \frac{\gamma}{2(i-1)-1} = 2i - 3 + \gamma - \delta_i$ , which implies that  $\delta_i = 2i - 3$ , gives in (8.15), (8.16) and (8.17) a global bound of order  $C \| \mathbf{F}_i \|_{L^\infty(C_d^{2i-3+\gamma})} \mathbf{d}^\gamma(\mathbf{z}, \mathbf{z}')$ . The result then follows from (8.14).  $\square$

## 9 Scaling Control of the degenerate part of the perturbative term

This section is dedicated to the proof of the scaled version of the key Besov Lemma 5. We recall that, with the definitions of Section 6, for all multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^{nd}$ ,  $i \in \llbracket 2, n \rrbracket$ , we aim to control the terms

$$\int_t^T ds \int_{\mathbb{R}^{nd}} D^\alpha D_{\mathbf{x}_1}^2 \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i, \mathbf{F}}^\lambda(t, s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \rangle d\mathbf{y},$$

with

$$\begin{aligned} & \sum_{i=2}^n \langle \mathbf{F}_{\lambda,i}(s, \mathbf{y}) - \mathbf{F}_{\lambda,i}(s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_{\lambda,i}(s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}))_{i-1}, D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \rangle \\ =: & \sum_{i=2}^n \langle \Delta_{i,\mathbf{F}}^\lambda(t, s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u^\lambda(s, \mathbf{y}) \rangle, \end{aligned}$$

which appear in equation (2.16) of the user's guide for the scaled system. We precisely want to specify how the scaling procedure impacts the constants in equation (2.33).

This is exactly what equations (6.18) and (6.19) reflect. Those controls actually follow from the more general following result, which will again be useful for the Hölder norm in Section 5.

**Lemma 14 (Scaled Besov Control Lemma)** *For all multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^d$ :*

$$\begin{aligned} & \sum_{i=2}^n \left| \int_{\mathbb{R}^{nd}} D^\alpha \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}}^\lambda(t, s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \rangle d\mathbf{y} \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\ \leq & C \lambda^{-1 + \sum_{j=1}^n \alpha_j (j - \frac{1}{2})} \|u^\lambda\|_{L^\infty(C_{b,d}^{2+\gamma})} (s-t)^{-\sum_{j=1}^n \alpha_j (j - \frac{1}{2}) + \frac{\gamma}{2}}. \end{aligned} \quad (9.1)$$

With Lemma 14 at hand, we readily derive (6.18) and (6.19) taking  $\alpha = (2, 0, \dots, 0) + e_k$  (where  $e_k$  stands for the  $k^{\text{th}}$  vector of the orthonormal basis) for each  $k \in \llbracket 2, n \rrbracket$  and  $\alpha = (2, 0, \dots, 0)$  respectively. Let us now turn to the proof of the Lemma 14.

**Proof: Lemma 14**

The analysis of singularities is identical to the ones in the proof of Lemma 5. However, here, we have to track the scaling coefficient  $\lambda$  through the identities. Note carefully, we write the upper-script/sub-script  $\lambda$  to mean that we manage the scaled variables. In particular, we write:

$$\vartheta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y}) := D_{\mathbf{y}_i} \cdot \left( D^\alpha \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) \otimes \Delta_{i,\mathbf{F}}^\lambda(t, s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}), \mathbf{y}) \right) =: D_{\mathbf{y}_i} \cdot \Theta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y}), \quad (9.2)$$

and

$$\Psi_{i,(t,\mathbf{x}), (s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha,\lambda} : \mathbf{y}_i \mapsto \vartheta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y}). \quad (9.3)$$

With these notations, we have:

$$\begin{aligned} & \sum_{i=2}^n \left| \int_{\mathbb{R}^{nd}} D^\alpha \tilde{p}_\lambda^{(t,\boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}}^\lambda(t, s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}), \mathbf{y}), D_{\mathbf{y}_i} u(s, \mathbf{y}) \rangle d\mathbf{y} \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\ = & \sum_{i=2}^n \left| \int_{\mathbb{R}^{nd}} \vartheta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y}) u^\lambda(s, \mathbf{y}) d\mathbf{y} \right|. \end{aligned} \quad (9.4)$$

The point is here again to control, for  $i \in \llbracket 2, n \rrbracket$ , the quantity  $\|\vartheta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n})\|_{B_{1,1}^{\tilde{\alpha}_i}}$   $\tilde{\alpha}_i = \frac{2+\gamma}{2i-1}$  with the indicated bounds in the scaling parameter  $\lambda$ . Accordingly with what can be seen e.g. in (6.13), the previous analysis of the proof of the (non-scaled) Lemma



5 can be adapted replacing  $(s - t)$  therein by  $(s - t)/\lambda$  in the computations involving the thermic characterization of Besov spaces.

We also point out that, w.l.o.g., we assume that  $T/\lambda \leq 1$  so that in particular for  $0 \leq t < s \leq T$ ,  $\lambda^{-1}(s - t) \leq 1$ . Indeed, the parameter  $\lambda$  is meant to be *small* (at least  $\lambda \leq 1$ ) but *macro* as well. From the previous analysis and the statement of Lemma 14 it can be seen that the *optimal*  $\lambda$ , i.e. the largest one, actually depends on the Hölder moduli of the coefficients. Hence, the condition  $T/\lambda \leq 1$  is, up to a possible modification of  $T$ , not restrictive.

Let us first introduce some notation:

$$\hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y}) := \prod_{j=1}^n \mathcal{N}_{c\lambda^{\frac{2j-1}{2}}(s-t)^{2j-1}}((\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j) = \bar{p}_{c^{-1}}(t, s, \mathbf{x}, \mathbf{y}),$$

where for  $a > 0$ ,  $z \in \mathbb{R}^d$ , like before  $\mathcal{N}_a(z) = \frac{1}{(2\pi a)^{\frac{d}{2}}} \exp\left(-\frac{|z|^2}{2a}\right)$  is the standard Gaussian density of  $\mathbb{R}^d$  with covariance matrix  $aI_d$ , and:

$$\hat{q}_{c,i,\lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) = \prod_{j \in \llbracket 1, n \rrbracket, j \neq i} \mathcal{N}_{c\lambda^{\frac{2j-1}{2}}(s-t)^{2j-1}}((\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j). \quad (9.5)$$

We recall from (4.14), that the parameter  $\beta_i = \frac{(2i-3)(2i-1)}{2i-3-\gamma}$ . The first contribution of the scaled Besov control is:

$$\begin{aligned} & \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}), (s,\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha,\lambda}\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & \leq \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} D^\alpha \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \Delta_{i,\mathbf{F}}^\lambda(t, s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}), \mathbf{y}), D_z h_v(z - \mathbf{y}_i) \rangle d\mathbf{y}_i \right|_{\boldsymbol{\xi}=\mathbf{x}} \\ & \leq C \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} \\ & \quad \times \lambda^{-i+\frac{1}{2}} \mathbf{d}^{2i-3+\gamma} (\lambda^{-1/2} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}^\lambda(\mathbf{x}), \lambda^{-1/2} \mathbb{T}_\lambda \mathbf{y}) \\ & \leq C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) - (i-\frac{1}{2})} \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \\ & \quad \times \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} (s-t)^{\frac{2i-3+\gamma}{2}}, \end{aligned}$$

exploiting (6.13) for the last inequality. Then

$$\begin{aligned} & \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}), (s,\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})}^{\alpha,\lambda}\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & \leq C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) - (i-\frac{1}{2})} \hat{q}_{c,i,\lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_{[\lambda^{-1}(s-t)]^{\beta_i}}^1 dv v^{-\frac{3}{2} + \frac{\tilde{\alpha}_i}{2}} (s-t)^{-\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-3+\gamma}{2}} \\ & \leq C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) - (i-\frac{1}{2})} \hat{q}_{c,i,\lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \lambda^{\lfloor \frac{1}{2} - \frac{\tilde{\alpha}_i}{2} \rfloor \beta_i} (s-t)^{\lfloor -\frac{1}{2} + \frac{\tilde{\alpha}_i}{2} \rfloor \beta_i - \sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-3+\gamma}{2}} \\ & \leq C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) - 1} \hat{q}_{c,i,\lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) (s-t)^{-\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{\gamma}{2}}, \quad (9.6) \end{aligned}$$

the third inequality is a consequence of Proposition 2, and the last identity comes from the pick of  $\beta_i$  which in particular gives  $-(i - \frac{1}{2}) + [1 - \tilde{\alpha}_i] \frac{\beta_i}{2} = -1$ .

Let us now consider the second contribution of the scaled Besov control, i.e. we now take  $v \in [0, [\lambda^{-1}(s - t)]^{\beta_i}]$ . Write:

$$\begin{aligned}
 & \int_{\mathbb{R}^d} h_v(z - \mathbf{y}_i) D_{\mathbf{y}_i} \cdot (\Theta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y})) d\mathbf{y}_i \\
 = & \int_{\mathbb{R}^d} h_v(z - \mathbf{y}_i) D_{\mathbf{y}_i} \cdot \left( \Theta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y}) - \Theta_{i,(t,\mathbf{x})}^{\alpha,\lambda}(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right) d\mathbf{y}_i \\
 = & \int_{\mathbb{R}^d} D^\alpha \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) \langle \mathbf{F}_{\lambda,i}(s, \mathbf{y}) - \mathbf{F}_{\lambda,i}(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}), D_z h_v(z - \mathbf{y}_i) \rangle d\mathbf{y}_i \\
 & + \int_{\mathbb{R}^d} \left( D^\alpha \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}) - D^\alpha \tilde{p}_\lambda^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right) \\
 & \times \left\langle \left( \mathbf{F}_{\lambda,i}(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) - \mathbf{F}_{\lambda,i}(s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_{\lambda,i}(s, \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}^\lambda(\boldsymbol{\xi}))_{i-1} \right), \right. \\
 & \left. D_z h_v(z - \mathbf{y}_i) \right\rangle d\mathbf{y}_i =: \left( \mathcal{I}_{\lambda,1} + \mathcal{I}_{\lambda,2} \right) (v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n})), \quad (9.7)
 \end{aligned}$$

thanks to the definition in (9.2) for the last identity.

$$\begin{aligned}
 & |\mathcal{I}_{\lambda,1}(v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}))| \quad (9.8) \\
 \leq & C \int_{\mathbb{R}^d} \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{\lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} \hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y}) \lambda^{-i+\frac{1}{2}} (\lambda^{\frac{2i-1}{2}} |z - \mathbf{y}_i|)^{\frac{2i-3+\gamma}{2i-1}} d\mathbf{y}_i \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) - (i-\frac{1}{2}) + \frac{2i-3+\gamma}{2}} \int_{\mathbb{R}^d} \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{2-\gamma}{4i-2}}} \frac{\hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} d\mathbf{y}_i. \\
 \leq & \frac{C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) - 1 + \frac{\gamma}{2}}}{v^{\frac{2-\gamma}{4i-2}} (s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} \hat{q}_{c,i,\lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \mathcal{N}_{cv+\lambda^{\frac{2i-1}{2}}(s-t)^{2i-1}}(z - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_i.
 \end{aligned}$$

Write now from (9.7):

$$\begin{aligned}
 & |\mathcal{I}_{\lambda,2}(v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}))| \\
 \leq & C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\mu \frac{\lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}} \hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} \\
 & \times \left| \mathbf{y}_i - z \right| \lambda^{-i+\frac{1}{2}} \left( \left| \mathbf{F}_i(s, \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda(\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n})) - \mathbf{F}_i(s, \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda(\mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i:n}) \right| \right. \\
 & + \left| \mathbf{F}_i(s, \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda(\mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i:n}) - \mathbf{F}_i(s, \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \\
 & \left. \left. - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \lambda^{-\frac{1}{2}} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}(\mathbf{x})) (\lambda^{-\frac{1}{2}} \mathbb{T}_\lambda(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x})))_{i-1} \right| \right) \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \int_0^1 d\mu \frac{\hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} \\
 & \times \left( (\lambda^{\frac{2i-1}{2}} |z - \boldsymbol{\theta}_{s,t}(\mathbf{x})_i|)^{\frac{2i-3+\gamma}{2i-1}} + (\lambda^{\frac{2(i-1)-1}{2}} |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_{i-1}|)^{1 + \frac{\gamma}{2(i-1)-1}} \right. \\
 & \left. + \sum_{k=i+1}^n (\lambda^{\frac{2j-1}{2}} |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_k|)^{\frac{2i-3+\gamma}{2k-1}} \right).
 \end{aligned}$$

We have for all  $\mu \in [0, 1]$ ,

$$|z - \boldsymbol{\theta}_{s,t}(\mathbf{x})_i| \leq \mu |z - \mathbf{y}_i| + |z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|$$

we thus derive

$$\begin{aligned}
 & |\mathcal{I}_{\lambda,2}(v, t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}))| \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \int_0^1 d\mu \frac{\hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} \\
 & \times \left( \lambda^{\frac{2i-3+\gamma}{2}} |\mathbf{y}_i - z|^{\frac{2i-3+\gamma}{2i-1}} \right. \\
 & \quad \left. + \mathbf{d}^{2i-3+\gamma} \left( \lambda^{-1/2} \mathbb{T}_\lambda \boldsymbol{\theta}_{s,t}(\mathbf{x}), \lambda^{-1/2} \mathbb{T}_\lambda (\mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \right) \right) \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \int_0^1 d\mu \hat{q}_{c,\lambda}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \mu(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \\
 & \times \left( \frac{\lambda^{\frac{2i-3+\gamma}{2}} v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \right) \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \hat{q}_{c \setminus i, \lambda}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \\
 & \times \int_0^1 d\mu \int_{\mathbb{R}^d} h_{cv}(z - \mathbf{y}_i) \mathcal{N}_{c\lambda^{\frac{2i-1}{2}}(s-t)^{2i-1}}(z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) d\mathbf{y}_i \\
 & \times \left( \frac{\lambda^{\frac{2i-3+\gamma}{2}} v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \right). \tag{9.9}
 \end{aligned}$$

From (9.7), (9.9) and (9.9) we deduce, with the notation of (9.5):

$$\begin{aligned}
 & \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^{\alpha,\lambda}\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \hat{q}_{c \setminus i, \lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \\
 & \times \left( \frac{\lambda^{-1+\frac{\gamma}{2}}}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} + \frac{\lambda^{\frac{2i-3+\gamma}{2}} v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \right) \\
 & \times \int_0^1 d\mu \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \mathcal{N}_{c\lambda^{\frac{2i-1}{2}}(s-t)^{2i-1}}(z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \hat{q}_{c \setminus i, \lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \\
 & \times \left( \frac{\lambda^{-1+\frac{\gamma}{2}}}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} + \frac{\lambda^{\frac{2i-3+\gamma}{2}} v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + \frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2}) + 1 - \frac{\gamma}{2}}} \right).
 \end{aligned}$$

The last identity is a again consequence of the change of variable  $(w_1, w_2) = (z - \mathbf{y}_i, z + \mu(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)$ .

We now write:

$$\begin{aligned}
 & \int_0^{[\lambda^{-1}(s-t)]^{\beta_i}} dv v^{\frac{\tilde{\alpha}_i}{2}-1} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^{\alpha,\lambda}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
 \leq & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_0^{[\lambda^{-1}(s-t)]^{\beta_i}} \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \\
 & \times \left( \frac{\lambda^{-1+\frac{\gamma}{2}}}{v^{\frac{2-\gamma}{4i-2}}(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})}} + \frac{\lambda^{\frac{2i-3+\gamma}{2}} v^{\frac{2i-3+\gamma}{2(2i-1)}}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+\frac{2i-1}{2}}} + \frac{1}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})+1-\frac{\gamma}{2}}} \right) \\
 =: & C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \hat{q}_{c \setminus i, \lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \mathcal{B}_{\alpha, \beta_i}^\lambda(t, s). \tag{9.10}
 \end{aligned}$$

Let us now prove that

$$\mathcal{B}_{\alpha, \beta_i}^\lambda(t, s) \leq \frac{C \lambda^{-1}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-\frac{\gamma}{2}}}. \tag{9.11}$$

Integrating in  $v$  in (9.10) we derive:

$$\begin{aligned}
 \mathcal{B}_{\alpha, \beta_i}^\lambda(t, s) \leq & C (s-t)^{-\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \left[ \lambda^{-1+\frac{\gamma}{2}} [\lambda^{-1}(s-t)]^{\beta_i(\frac{\tilde{\alpha}_i}{2}-\frac{2-\gamma}{4i-2})} \right. \\
 & \left. + \lambda^{\frac{2i-3+\gamma}{2}} [\lambda^{-1}(s-t)]^{\beta_i(\frac{\tilde{\alpha}_i}{2}+\frac{2i-3+\gamma}{2(2i-1)})} (s-t)^{i-\frac{1}{2}} + [\lambda^{-1}(s-t)]^{\beta_i \frac{\tilde{\alpha}_i}{2}} (s-t)^{-1+\frac{\gamma}{2}} \right].
 \end{aligned}$$

Recall now from the proof of Lemma 5 that:

$$\beta_i \left( \frac{\tilde{\alpha}_i}{2} - \frac{2-\gamma}{4i-2} \right) - \frac{\gamma}{2} \geq 0, \quad \beta_i \left( \frac{\tilde{\alpha}_i}{2} + \frac{2i-3+\gamma}{2(2i-1)} \right) - \frac{2i-1}{2} - \frac{\gamma}{2} \geq 0, \quad \beta_i \frac{\tilde{\alpha}_i}{2} - 1 \geq 0,$$

with  $\beta_i = \frac{(2i-3)(2i-1)}{2i-3-\gamma}$ ,  $\tilde{\alpha}_i = \frac{2+\gamma}{2i-1}$ . Therefore, since  $(s-t)/\lambda \leq 1$ :

$$\begin{aligned}
 & \mathcal{B}_{\alpha, \beta_i}^\lambda(t, s) \\
 \leq & C (s-t)^{-\sum_{j=1}^n \alpha_j(j-\frac{1}{2})} \left[ \lambda^{-1+\frac{\gamma}{2}} [\lambda^{-1}(s-t)]^{\frac{\gamma}{2}} + \lambda^{-1+\frac{\gamma}{2}} [\lambda^{-1}(s-t)]^{\frac{\gamma}{2}} + [\lambda^{-1}(s-t)] (s-t)^{-1+\frac{\gamma}{2}} \right],
 \end{aligned}$$

which precisely gives (9.11).

Plugging (9.11) into (9.10) and from (9.6) we eventually get:

$$\begin{aligned}
 & \int_0^1 \frac{dv}{v} v^{\frac{\tilde{\alpha}_i}{2}} \|h_v \star \Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^{\alpha,\lambda}\|_{L^1(\mathbb{R}^d,\mathbb{R})} \\
 \leq & \frac{C \lambda^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-1}}{(s-t)^{\sum_{j=1}^n \alpha_j(j-\frac{1}{2})-\frac{\gamma}{2}}} \hat{q}_{c \setminus i, \lambda}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})),
 \end{aligned}$$

which is precisely the stated control. The term  $\|\varphi(D)\Psi_{i,(t,\mathbf{x}),(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})}^{\alpha,\lambda}\|_{L^1(\mathbb{R}^d,\mathbb{R})}$  appearing in the Besov norm could be handled similarly. The result is complete.  $\square$



# Chapter 7

## Strong Uniqueness for a chain of Degenerate Kolmogorov Equations

**Abstract:** We establish strong uniqueness for a class of degenerate Kolmogorov SDEs under suitable Hölder regularity conditions for the associated drift term. In particular, we exhibit some sharp thresholds on the Hölder exponents for the strong uniqueness to hold. Our approach is based on the forward parametrix perturbation technique and also heavily relies on appropriate duality properties on Besov spaces.

### 1 Introduction

#### 1.1 Statement of the problem and related results

In this work, we aim at establishing a strong well posedness result outside the classical Cauchy-Lipschitz framework for the following degenerate Stochastic Differential Equation (SDE) of Kolmogorov type:

$$\begin{aligned}
 d\mathbf{X}_t^1 &= \mathbf{F}_1(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt + \sigma(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dW_t, \\
 d\mathbf{X}_t^2 &= \mathbf{F}_2(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt, \\
 d\mathbf{X}_t^3 &= \mathbf{F}_3(t, \mathbf{X}_t^2, \dots, \mathbf{X}_t^n)dt, \\
 &\vdots \\
 d\mathbf{X}_t^n &= \mathbf{F}_n(t, \mathbf{X}_t^{n-1}, \mathbf{X}_t^n)dt,
 \end{aligned}
 \tag{1.1}$$

where,  $(W_t)_{t \geq 0}$  stands for a  $d$ -dimensional Brownian motion on some filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and for all  $i \in \llbracket 1, n \rrbracket^*$ ,  $t \geq 0$  the component  $\mathbf{X}_t^i$  is  $\mathbb{R}^d$ -valued as well (i.e.  $\mathbf{X}_t \in \mathbb{R}^{nd}$ ). We suppose that the  $(\mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$  satisfy a kind of weak Hörmander condition, i.e. the matrices  $(D_{\mathbf{x}_{i-1}} \mathbf{F}_i(t, \cdot))_{i \in \llbracket 2, n \rrbracket}$  have full rank. However, the coefficients  $(\mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$  can be rather *rough* in their other entries, namely, Hölder continuous. We assume as well that the diffusion coefficient  $\sigma$  is bounded from above and below and spatially Lipschitz continuous.

---

\*. We will use throughout the paper the notation  $\llbracket \cdot, \cdot \rrbracket$  for integer intervals.

For a system of Ordinary Differential Equation (ODE) it may be a real challenge to prove the well posedness outside the Lipschitz framework (see *e.g.* [PL89]) and, as shown by Peano's example, uniqueness may fail as soon as the drift of the system of interest is only Hölder continuous. For an SDE, the story is rather different since the presence of the noise may allow to restore well posedness. Such a phenomenon, called *regularization by noise* (see the Saint Flour Lecture notes of Flandoli [Fla11] and the references therein for an overview of the topic), has been widely studied since the pioneering unidimensional work of Zvonkin [Zvo74] and its generalization to the multi-dimensional setting by Veretennikov [Ver80] where SDEs driven by a *non-degenerate* Brownian noise and a bounded drift are shown to be well-posed. We mean by *non-degenerate* that the noise has the same dimension as the underlying system on which it acts.

Let us mention, among others, and still within the *non-degenerate* setting, the works of Krylov and Röckner [KR05] ( $L_q - L_p$  drift), Zhang [Zha10] ( $L_q - L_p$  drift and weakly Lipschitz diffusion matrix) and also Fedrizzi and Flandoli [FF11] ( $L_q - L_p$  and Hölder drift).

The crucial point in the aforementioned results lies into the non-degeneracy assumption assumed on the noise added in the considered system: we can benefit from the regularization by noise phenomenon only in the directions submitted to the noise and when the noise degenerates it is useless to expect a generalization of the previous results without any additional assumptions.

In our current setting, the non-degeneracy assumption on the Jacobian  $(D_{\mathbf{x}_{i-1}} \mathbf{F}_i(t, \cdot))$ ,  $i \in \llbracket 2, n \rrbracket$  precisely allows the noise to propagate through the chain passing from the  $i^{\text{th}}$  to the  $(i + 1)^{\text{th}}$  level thanks to the drift, hence leading to a propagation of the noise in the whole considered space. The main idea is then to take advantage of this particular propagation, known as *weak Hörmander setting* (in reference to the work of Hörmander on hypoelliptic differential operator [Hör67]), to restore strong well posedness under our current Hölder framework. This feature has already been considered in the literature for the system (1.1) in the particular case  $n = 2$ , see the works of Chaudru de Raynal [CdR17], Wang and Zhang [WZ16], Fedrizzi, Flandoli, Priola and Vovelle [FFPV17], Zhang [Zha18]. In any cases, in addition to the weak Hörmander structure, the regularity of the drift with respect to (w.r.t.) the second space argument is required to be of regularity index superior or equal (depending on the work) to  $2/3$  (critical Hölder index or critical weak differentiation index). As a generalization of these results, we prove in this paper that strong well posedness holds as soon as each drift component  $\mathbf{F}_i$  is  $\beta_j$ -Hölder continuous in the  $j^{\text{th}}$  variable for some  $\beta_j \in \left( (2j - 2)/(2j - 1), 1 \right]$  so that we recover the critical index mentioned above when  $j = 2$ .

In comparison with the results obtained in the non-degenerate framework, the increasing value of the critical Hölder index can be understood as the price to pay to balance the degeneracy of the noise. Indeed, there is a strong competition between the irregularity of the drift and the fluctuations of the noise, which has been highlighted in the work of Delarue and Flandoli [DF14]: as the regularity of the drift decreases the fluctuations of the noise have to increase more and more in order to regularize the system.

In our setting, the way the noise is allowed to propagate implies a loss of the fluctua-

tions at each level of the chain and therefore, a loss of its regularization property. A good manner to understand this phenomenon is the following case. Given a level  $i$  in  $\llbracket 2, n \rrbracket$  of the chain (1.1), assume that  $\mathbf{F}_1 \equiv 0$ ,  $\sigma \equiv \text{Id}$ ,  $\mathbf{F}_\ell(t, \mathbf{x}_{\ell-1}, \dots, \mathbf{x}_n) = \mathbf{x}_{\ell-1}$ ,  $\ell \in \llbracket 2, i-1 \rrbracket$ ,  $\mathbf{F}_i(t, \mathbf{x}_{i-1}, \dots, \mathbf{x}_n) =: \mathbf{x}_{i-1} + \bar{\mathbf{F}}_i(t, \mathbf{x}_i)$ , and  $\mathbf{F}_j = 0$ ,  $j \in \llbracket i+1, n \rrbracket$ . In such a case, the noise added at the  $i^{\text{th}}$  level of the chain is only the  $(i-1)^{\text{th}}$  iterated time integral of a Brownian motion whose fluctuations are of order  $t^{i-1/2}$ , meaning that the Hölder regularity of  $\bar{\mathbf{F}}_i$  has to increase with  $i$ . This simple feature illustrates the loss of fluctuations of the noise from level to level through its propagation to the chain and consequently in the corresponding spatial direction. This is the reason why the index of Hölder regularity increases when considering direction further and further away from the source of noise.

Thanks to Delarue and Flandoli's result [DF14], Chaudru de Raynal [CdR18] and then Chaudru de Raynal and Menozzi [CdRM17] have proposed a class of counter-examples to weak well posedness for the chain (1.1) when  $n = 2$  and in full generality, respectively. This class of counter-examples says that weak well posedness does not hold as soon as for any  $i$  in  $\llbracket 2, n \rrbracket$  the drift  $\mathbf{F}_i$  has an Hölder regularity index with respect to the  $j^{\text{th}}$  variable, denoted by  $\tilde{\beta}_i^j$ , below  $(2i-3)/(2j-1)$  e.g. regarding the regularity of the drift w.r.t. the second space variable we hence obtain  $\tilde{\beta}_2^2 = 1/3$ . Also, in the latest work [CdRM17], weak well posedness is shown to hold as soon as the regularity indexes  $(\tilde{\beta}_i^j)_{2 \leq i, j \leq n}$  satisfy  $(2i-3)/(2j-1) < \tilde{\beta}_i^j \leq 1$  so that the result is *almost sharp*<sup>†</sup>, in the sense that the critical case corresponding to an equality for the thresholds is still open.

In a nutshell, when  $n = 2$  and  $d = 1$ , to exhibit the critical Hölder threshold  $\tilde{\beta}_2^2$  (Hölder continuity index of  $\mathbf{F}_2$  in its second variable) needed for the weak uniqueness to hold, one can go back to the key example of Peano. The point is to compare the maximal solutions of the ODE with dynamics  $\dot{x}_t = \text{sgn}(x_t)|x_t|^{\tilde{\beta}_2^2}$  starting from 0 (which is precisely the singular point of the ODE), which writes as  $\pm ct^{1/(1-\tilde{\beta}_2^2)}$ , and the typical magnitude of the integral of the Brownian motion which is  $t^{3/2}$ . For the noise to dominate in small time, in order to leave the point where uniqueness fails for the deterministic equation, one must take  $3/2 < 1/(1-\tilde{\beta}_2^2) \Leftrightarrow \tilde{\beta}_2^2 > 1/3$  (see [CdR18]). Through a perturbative approach, the threshold can also be deduced from the behavior of the ratio  $|x_2|^{\tilde{\beta}_2^2}/t^{3/2}$  where  $x_2$  is a spatial variable which has the same typical order as the integral of the Brownian motion. Namely  $|x_2| \asymp t^{3/2}$ . Hence,  $|x_2|^{\tilde{\beta}_2^2}/t^{3/2} \asymp t^{-3/2(1-\tilde{\beta}_2^2)}$ . For weak uniqueness to hold, the point is then to impose that the previous ratio gives an integrable singularity in time. This precisely leads to the condition  $-3/2(1-\tilde{\beta}_2^2) > -1 \Leftrightarrow \tilde{\beta}_2^2 > 1/3$ .

The weak uniqueness thresholds for the full chain were derived similarly in [CdRM17] studying for  $i \in \llbracket 2, n \rrbracket$ ,  $|x_j|^{\tilde{\beta}_i^j}/t^{i-1/2}$ ,  $j \in \llbracket i, n \rrbracket$ , where  $x_j$  is a spatial variable which has the same typical order as the  $(j-1)^{\text{th}}$  iterated integral of the Brownian motion. Namely  $|x_j| \asymp t^{j-1/2}$ . The corresponding ratio thus satisfies  $|x_j|^{\tilde{\beta}_i^j}/t^{i-1/2} \asymp t^{-(i-1/2)+\tilde{\beta}_i^j(j-1/2)}$ . In order to obtain an integrable singularity one must take  $-(i-1/2)+\tilde{\beta}_i^j(j-1/2) > -1 \Leftrightarrow \tilde{\beta}_i^j > (2i-3)/(2j-1)$  which is precisely the indicated threshold. Of course, the previous ratios appear through rather lengthy and technical perturbative procedures: the forward parametrix expansion in [CdR18] and the backward one in [CdRM17] (which allows to

---

†. the drift  $\mathbf{F}_1$  may be assumed to be only in  $L_q - L_p$  space for any  $n^2d/p + 2/q < 1$ ,  $p \geq 2$ ,  $q > 2$  or Hölder continuous for any strictly positive Hölder index.



improve the Hölder thresholds on  $\mathbf{F}_1$  w.r.t. [CdR18]).

When comparing the weak well posedness result in [CdRM17] with the strong well posedness proved in this work, one may wonder why in the strong case the critical Hölder exponent are greater and why they do not take into account the level in the chain (recall that we investigate strong well posedness for drifts  $\mathbf{F}_i$  which are  $\beta_j$ -Hölder continuous in the  $j^{\text{th}}$  variable for some  $\beta_j$  independent of  $i$  satisfying  $(2j - 2)/(2j - 1) < \beta_j \leq 1$ ). This in particular gives  $\beta_2 > 2/3$  while we obtained  $\tilde{\beta}_2^2 > 1/3$  for weak well posedness.

One reason for these notable differences relies on Partial Differential Equation (PDE) results. Indeed, it is well known that the family of generators  $(L_t)_{t \geq 0}$  of (1.1) is a family of linear partial differential operators of second order so that any SDE are connected with a PDE (the solution of the SDE being, in fact, the characteristic of the solution of the Cauchy problem associated with  $L$ ). For instance, this connexion has been widely used in the last past five decades allowing to pass from regularization properties of PDE to well posedness result for SDE and conversely, see *e.g.* the book of Bass [Bas97].

Keeping in mind this connexion, one may view the difference of critical Hölder indexes between strong and weak well posedness as a consequence of the different regularization properties required on the underlying PDE to prove the probabilistic result: roughly speaking the weak well posedness requires gradient bounds for the solution of the associated Cauchy problem while the strong well posedness requires in addition to control some of the cross derivatives of the solution. Since the regularization properties of the operator are the same in both cases, this additional control explains the critical Hölder indexes increase for the strong uniqueness to hold.

The reason why the Hölder regularity index is blind to the level of the chain is a bit more involved at this stage of the presentation and we refer to subsections 1.3 and 1.4 below for a more detailed discussion on the subject. Let us only mention that in the weak case, the underlying Cauchy problem is investigated with a global source term  $f$  of independent regularity while in the strong case the source term should be *any* component of the drift of (1.1) itself. This is a consequence of the initial Zvonkin approach. Each of the  $(\mathbf{F}_i)_{i \in \llbracket 1, n \rrbracket}$  has to be considered as a source term for the corresponding PDE.

W.r.t. to the previous heuristic derivation of the thresholds for the weak uniqueness, let us indicate that the Zvonkin approach to strong uniqueness, coupled to a forward perturbative approach actually leads to consider the ratios  $|\mathbf{x}_j|^{\beta_j} / t^{(j-1/2)+1/2}$ ,  $j \in \llbracket 1, n \rrbracket$ , i.e. the time singularity is higher of degree  $1/2$ . We recall that, formally, for the weak uniqueness we are led to control the gradients of the Green kernel associated with a degenerate Ornstein-Uhlenbeck like process and an associated perturbation term involving the difference of the initial generator of (1.1) and the one of the Ornstein-Uhlenbeck proxy. For the strong uniqueness the quantities of interest are the derivatives of the gradient w.r.t. the non-degenerate variables for the Green kernel and the perturbative term. The additional singularity in the previous ratio precisely corresponds to the typical scale induced by the derivative of a Gaussian kernel w.r.t. its non-degenerate entry. Considering again the associated typical scale for the variable  $\mathbf{x}_j$ , which is  $t^{j-1/2}$ , we write  $|\mathbf{x}_j|^{\beta_j} / t^{(j-1/2)+1/2} = t^{-j+\beta_j(j-1/2)}$ . In order to get an integrable singularity in time, we must have  $-j + \beta_j(j - 1/2) > -1 \iff \beta_j > (2j - 2)/(2j - 1)$ , which precisely gives the stated thresholds. These features, Zvonkin transform and perturbative analysis are detailed re-

spectively in Section 1.3 and 1.4 below. One could wonder whether the thresholds we find are linked or not to the Zvonkin approach. A partial answer is that other techniques, based on pathwise arguments, developed by Catellier and Gubinelli [CG16] to address well-posedness of non-degenerate fractional driven SDEs with singular drift, lead to the same thresholds. In the non-degenerate scalar case, let us also mention the works of Gradinaru and Offret [GO13] or Bass and Chen [BC03] who obtained better thresholds than we do. Anyhow, the one dimensional case must be considered apart.

Let us eventually conclude this part by emphasizing that we were unfortunately not able to produce a counter example for strong uniqueness when the Hölder index of regularity of the drift is below the critical value given above. Recall indeed that weak well posedness holds in our considered case so that, if the critical exponent exhibited is sharp (or almost sharp), this means from the Yamada Watanabe Theorem that strong existence fails as well. It hence seems rather difficult to produce a counter example based on the irregularity of the drift in such a case and we are not aware of any result of this type in the literature. Usual counter examples, based on measurability arguments, typically rely on the regularity of the diffusion matrix, as it is the case for the so called Tanaka's example [Tan79] (see also the counter examples in the book of Cherny [Che05]). Nevertheless, as we will discuss in subsection 1.4, these thresholds appear to be almost sharp for the method implemented in this work to prove the strong well posedness result.

## 1.2 Notations, assumptions and main result

**Some notations.** We will denote by a bold letter  $\mathbf{x}, \mathbf{y}$  any element of  $\mathbb{R}^{nd}$ , writing as well  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$  where for  $i \in \llbracket 2, n \rrbracket$ ,  $\mathbf{x}_i \in \mathbb{R}^d$ . For practical purpose we will be led in our analysis to consider *subcomponents* of a vector  $\mathbf{x} \in \mathbb{R}^{nd}$ . Namely, for any  $0 \leq i \leq j \leq n$  and  $\mathbf{x} \in \mathbb{R}^{nd}$ , we introduce the notation  $\mathbf{x}_{i:j} := (\mathbf{x}_i, \dots, \mathbf{x}_j)$ . Accordingly, we write the drift as the mapping

$$\begin{aligned} (s, \mathbf{x}) \in \mathbb{R}^+ \times \mathbb{R}^{nd} \mapsto \mathbf{F}(s, \mathbf{x}) &= (\mathbf{F}_1(s, \mathbf{x}), \dots, \mathbf{F}_n(s, \mathbf{x})) \\ &= (\mathbf{F}_1(s, \mathbf{x}), \mathbf{F}_2(s, \mathbf{x}), \mathbf{F}_3(s, \mathbf{x}_{2:n}) \cdots, \mathbf{F}_n(s, \mathbf{x}_{n-1:n})), \end{aligned}$$

from the specific structure of the drift appearing in (1.1).

For  $f \in C^1(\mathbb{R}^{nd}, \mathbb{R}^k)$ ,  $k \in \{1, d\}$ , we denote for all  $i \in \llbracket 1, n \rrbracket$ , by  $D_{\mathbf{x}_i} f(\mathbf{x})$  the Jacobian matrix of the derivative of  $f$  w.r.t. to its  $\mathbb{R}^d$ -valued variable  $\mathbf{x}_i$ . As shortened form, and when no ambiguity is possible, we also write for all  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{nd}$ ,  $D_{\mathbf{x}_i} f(\mathbf{x}) = D_i f(\mathbf{x})$  and  $D_{\mathbf{y}_i} f(\mathbf{y}) = D_i f(\mathbf{y})$ . Also, if  $k = 1$  we denote by  $\mathbf{D}f(\mathbf{x}) = (D_1 f(\mathbf{x}) \cdots D_n f(\mathbf{x}))^*$  the full gradient of the function  $f$  at point  $\mathbf{x}$ .

Let  $f : \mathbb{R}^{nd} \rightarrow \mathbb{R}^k$  and  $\beta := (\beta_1, \dots, \beta_n) \in (0, 1]^n$  be a multi-index. We say that  $f$  is uniformly  $\beta$ -Hölder continuous if for all  $j \in \llbracket 1, n \rrbracket$

$$[f_j]_{\beta_j} := \sup_{(\mathbf{z}_{1:j-1}, \mathbf{z}_{j+1:n}) \in \mathbb{R}^{(n-1)d}, z \neq z', (z, z') \in (\mathbb{R}^d)^2} \frac{|f(s, \mathbf{z}_{1:j-1}, z, \mathbf{z}_{j+1:n}) - f(s, \mathbf{z}_{1:j-1}, z', \mathbf{z}_{j+1:n})|}{|z - z'|^{\beta_j}} < +\infty. \quad (1.2)$$

For a smooth function  $\Psi : [0, T] \times \mathbb{R}^{nd} \rightarrow \mathbb{R}^{nd}$ , where  $T > 0$  is a fixed given time, writing for  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ ,  $\Psi(t, \mathbf{x}) = (\Psi_1(t, \mathbf{x}), \dots, \Psi_n(t, \mathbf{x}))^*$  where for each  $i \in$

$\llbracket 1, n \rrbracket$ ,  $\Psi_i$  is  $\mathbb{R}^d$  valued, we denote by

$$\|\mathbf{D}\Psi\|_\infty := \sum_{i=1}^n \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \|\mathbf{D}\Psi_i(t, \mathbf{x})\|, \|\mathbf{D}(D_1\Psi)\|_\infty := \sum_{i=1}^n \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \|\mathbf{D}(D_1\Psi_i)(t, \mathbf{x})\|, \quad (1.3)$$

where in the above equation  $\|\cdot\|$  stands for a tensor norm in the appropriate corresponding dimension. Precisely,  $\mathbf{D}\Psi_i(t, \mathbf{x}) \in \mathbb{R}^{nd} \otimes \mathbb{R}^d$  and  $\mathbf{D}(D_1\Psi_i)(t, \mathbf{x}) \in \mathbb{R}^{nd} \otimes \mathbb{R}^d \otimes \mathbb{R}^d$ .

**Assumptions** We will assume throughout the paper that the following conditions hold.

**(ML)** The coefficients  $\mathbf{F}$  and  $\sigma$  are measurable in time. Also, the diffusion coefficient  $\sigma$  is uniformly Lipschitz continuous in space, uniformly in time, i.e. there exists  $\kappa > 0$  s.t. for all  $t \geq 0$ ,  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ :

$$|\sigma(t, \mathbf{x}) - \sigma(t, \mathbf{x}')| \leq \kappa |\mathbf{x} - \mathbf{x}'|.$$

**(UE)** The diffusion matrix  $a := \sigma\sigma^*$  is uniformly elliptic and bounded, uniformly in time, i.e. there exists  $\Lambda \geq 1$  s.t. for all  $t \geq 0$ ,  $(\mathbf{x}, \zeta) \in \mathbb{R}^{nd} \times \mathbb{R}^d$ :

$$\Lambda^{-1}|\zeta|^2 \leq \langle a(t, \mathbf{x})\zeta, \zeta \rangle \leq \Lambda|\zeta|^2.$$

**(T $_\beta$ )** For all  $j \in \llbracket 1, n \rrbracket$ , the functions  $(\mathbf{F}_i)_{i \in \llbracket 1, (j+1) \wedge n \rrbracket}$  are uniformly  $\beta_j$ -Hölder continuous in the  $j^{\text{th}}$  spatial variable with  $\beta_j \in (\frac{2j-2}{2j-1}, 1]$ , uniformly w.r.t. the other spatial variables of  $\mathbf{F}_i$  and in time. In particular, there exists a finite constant  $C_\beta$  s.t.

$$\max_{(i, j) \in \llbracket 1, n \rrbracket^2} \sup_{s \in [0, T]} [(\mathbf{F}_i)_j(s, \cdot)]_{\beta_j} \leq C_\beta.$$

**(H $_\eta$ )** For all  $i \in \llbracket 2, n \rrbracket$ ,  $D_{\mathbf{x}_{i-1}}\mathbf{F}_i$  is non-degenerate (weak type of Hörmander condition) and  $\eta$ -Hölder continuous w.r.t.  $\mathbf{x}_{i-1}$  uniformly in  $\mathbf{x}_{i:n}$  and time. We also assume without loss of generality that  $\eta \in (0, \inf_{j \in \llbracket 2, n \rrbracket} \{\beta_j - \frac{2j-2}{2j-1}\})$ , i.e.  $\eta$  is meant to be *small*.

From now on, we will say that assumption **(A)** is in force provided that **(ML)**, **(UE)**, **(T $_\beta$ )**, **(H $_\eta$ )** hold.

**Main result.** The main result of this work is the following theorem.

**Theorem 1 (Strong uniqueness for the degenerate system (1.1))** *Under **(A)** there exists a unique strong solution to system (1.1).*

**Remark 1** *Still in comparison with the results obtained in the non-degenerate cases, and especially the one of Krylov and Röckner [KR05], we do not tackle the case of drift in  $L_q - L_p$  w.r.t. the first (and then non-degenerate) variable. This is only to keep our result as clear as possible and to concentrate on the novelty of the approach we use here. We are anyhow confident that these specific drifts could be handled. Indeed, all the intermediate results needed to perform the analysis in that setting seem to be already available. We refer to subsection 1.4 for further details.*

### 1.3 Proof of the main result: Zvonkin Transform and smoothing properties of the PDE associated with (1.1)

We emphasize that under our assumptions, it follows from [CdRM17] that (1.1) is well posed in the weak sense. Hence, from Yamada-Watanabe theorem it is sufficient to prove that strong (or pathwise) uniqueness holds to prove strong well posedness. To do so, our main strategy rests upon the Zvonkin transform initiated by Zvonkin in [Zvo74] which has been widely used during the last decade to prove strong well posedness, see *e.g.*, and heavily relies on the connexion between SDE and PDE as we already emphasized.

Eventually, let us mention [Ver83], where the author extended the result in [Ver80] to the considered chain for  $n = 2$ . Note however that this extension only concerns the non-degenerate part, the coefficients being supposed to be twice continuously differentiable functions with bounded derivatives w.r.t the degenerate component.

Introducing the embedding matrix  $B$  from  $\mathbb{R}^d$  into  $\mathbb{R}^{nd}$ , i.e.  $B = (\mathbf{I}_{d,d}, \underbrace{\mathbf{0}_{d,d}, \dots, \mathbf{0}_{d,d}}_{(n-1) \text{ times}})^* = (\mathbf{I}_{d,d}, \mathbf{0}_{d,(n-1)d})^*$ , where “\*” stands for the transpose, we rewrite (1.1) in the shortened form

$$d\mathbf{X}_t = \mathbf{F}(t, \mathbf{X}_t)dt + B\sigma(t, \mathbf{X}_t)dW_t,$$

where  $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_n)$  is an  $\mathbb{R}^{nd}$ -valued function. For all  $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$  and  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$  let

$$L_t\varphi(\mathbf{x}) = \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}\varphi(t, \mathbf{x}) \rangle + \frac{1}{2}\text{Tr}\left(a(t, \mathbf{x})D_{\mathbf{x}_1}^2\varphi(\mathbf{x})\right), \quad (1.4)$$

where  $a = \sigma\sigma^*$  denotes the generator associated with (1.1). We then *formally* associate the SDE (1.1) with the following systems of PDEs:

$$\begin{cases} (\partial_t \mathbf{u}_i + L_t \mathbf{u}_i)(t, \mathbf{x}) = \mathbf{F}_i(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T) \times \mathbb{R}^{nd}, \\ \mathbf{u}_i(T, \mathbf{x}) = \mathbf{0}_d, & i \in \llbracket 1, n \rrbracket, \end{cases} \quad (1.5)$$

and we denote by  $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$  its global solution. Let now  $(\mathbf{F}^m)_{m \geq 0}, (a^m)_{m \geq 0}$  denote two sequences of mollified coefficients satisfying assumption **(A)** uniformly in  $m$  that are infinitely differentiable functions with bounded derivatives of all, order greater than 1 for  $\mathbf{F}^m$ , and converging in supremum norm to  $(\mathbf{F}, a)$  (such sequences are easily obtained from [CdR17]). Then, for each  $m$ , the regularized systems of PDEs associated with (1.5) write:

$$\begin{cases} (\partial_t \mathbf{u}_i^m + L_t^m \mathbf{u}_i^m)(t, \mathbf{x}) = \mathbf{F}_i^m(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T) \times \mathbb{R}^{nd}, \\ \mathbf{u}_i^m(T, \mathbf{x}) = \mathbf{0}_d, & i \in \llbracket 1, n \rrbracket, \end{cases} \quad (1.6)$$

where  $L_t^m$  is obtained from (1.4) replacing  $\mathbf{F}$  by  $\mathbf{F}^m$  and  $a$  by  $a^m$ .

The above system (1.6) is well posed and admits a unique smooth solution  $\mathbf{U}^m = (\mathbf{u}_1^m, \dots, \mathbf{u}_n^m)$ . Hence, applying Itô's Formula, one easily deduces that

$$\int_0^t \mathbf{F}(s, \mathbf{X}_s)ds = -\mathbf{U}^m(0, \mathbf{x}) + \mathbf{U}^m(t, \mathbf{X}_t) - \int_0^t \mathbf{D}\mathbf{U}^m(s, \mathbf{X}_s)B\sigma(s, \mathbf{X}_s)dW_s + \mathcal{R}_t^m(\mathbf{X}), \quad (1.7)$$

where

$$\mathcal{R}_t^m(\mathbf{X}) := \int_0^t \mathbf{F}(s, \mathbf{X}_s) - \mathbf{F}^m(s, \mathbf{X}_s) ds - \int_0^t (L_s - L_s^m) \mathbf{U}^m(s, \mathbf{X}_s) ds.$$

The representation (1.7) is what is usually called the *Zvonkin Transform*. The main thing in this representation is that the *bad drift* can be, up to a remainder, rewritten in term of the solution of the system of PDEs (1.6) for which the source term is precisely a mollified version of the *bad drift* itself.

Then, the main idea consists in taking advantage of the regularization properties of the operator  $L^m$  (uniformly in  $m$ ) and expect that the solutions  $\mathbf{U}^m$ ,  $m \geq 0$  will be smoother than the source term  $\mathbf{F}$  so that the right hand side of (1.7) is smoother than the integrand of the left hand side of the considered equation. In other words, we are looking for a *good regularization theory* for the PDE (1.6) uniformly on the mollification argument. This *good regularization theory* is summarized in the following crucial result whose proof is, in fact, the main subject of this work and is postponed to Section 2.

**Theorem 2** *For  $T > 0$  small enough<sup>‡</sup>, there exists a constant  $C_T := C_T(\mathbf{A}) > 0$  satisfying  $C_T \rightarrow 0$  when  $T \rightarrow 0$  such that for every  $m \geq 0$ , the solution  $\mathbf{U}^m$  satisfies with the notation of (1.3):*

$$\|\mathbf{D}\mathbf{U}^m\|_\infty + \|\mathbf{D}(D_1\mathbf{U}^m)\|_\infty \leq C_T. \quad (1.8)$$

We also point out that, from the uniformity in  $m$  in the previous theorem, we could also derive some regularizing properties for the system (1.5) through appropriate compactness arguments.

Let now  $\mathbf{X}$  and  $\mathbf{X}'$  be two solutions of (1.1). Using the representation (1.7) to express the difference of the *bad drift* in terms of the function  $\mathbf{U}^m$  and its derivative, we write:

$$\begin{aligned} & \mathbf{X}_t - \mathbf{X}'_t \\ = & \mathbf{U}^m(t, \mathbf{X}_t) - \mathbf{U}^m(t, \mathbf{X}'_t) - \int_0^t [\mathbf{D}\mathbf{U}^m(s, \mathbf{X}_s)B\sigma(s, \mathbf{X}_s) - \mathbf{D}\mathbf{U}^m(s, \mathbf{X}'_s)B\sigma(s, \mathbf{X}'_s)] dW_s \\ & + [\mathcal{R}_t^m(\mathbf{X}) - \mathcal{R}_t^m(\mathbf{X}')] + \int_0^t B[\sigma(s, \mathbf{X}_s) - \sigma(s, \mathbf{X}'_s)] dW_s. \end{aligned}$$

Take then the supremum in time of the square of the difference. Passing to the expectation, a convexity inequality then leads to the following estimate:

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] \\ \leq & 5 \left( \mathbb{E} \left[ \sup_{t \leq T} |\mathbf{U}^m(t, \mathbf{X}_t) - \mathbf{U}^m(t, \mathbf{X}'_t)|^2 \right] \right. \\ & + \mathbb{E} \left[ \int_0^T |[\mathbf{D}\mathbf{U}^m B](s, \mathbf{X}_s) - [\mathbf{D}\mathbf{U}^m B](s, \mathbf{X}'_s)|^2 \|\sigma\|_\infty^2 ds \right] \\ & \left. + \mathbb{E} \left[ \int_0^T (\|\mathbf{D}\mathbf{U}^m B\|_\infty + 1) |\sigma(s, \mathbf{X}_s) - \sigma(s, \mathbf{X}'_s)|^2 ds \right] + 2\|\mathcal{R}^m(\cdot)\|_\infty^2 \right). \end{aligned}$$

---

<sup>‡</sup>. By “small enough” we mean that there exists a time  $\mathcal{T} > 0$  depending on known parameters in  $(\mathbf{A})$  s.t. for all  $T \leq \mathcal{T}$  the statement of the theorem holds.

Note that thanks to the particular structure of  $B$  one has  $\mathbf{D}\mathbf{U}^m B = (D_{\mathbf{x}_1} \mathbf{U}^m, \mathbf{0}_{d,d}, \dots, \mathbf{0}_{d,d})^*$ . Hence, thanks to Theorem 2 and Grönwall's lemma, there exists  $\bar{C}_T := \bar{C}_T(C_T, \sigma, n, d, T)$  satisfying  $\bar{C}_T \rightarrow 0$  when  $T$  goes to 0 such that

$$\mathbb{E} \left[ \sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] \leq \bar{C}_T \mathbb{E} \left[ \sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] + 10 \|\mathcal{R}^m(\cdot)\|_\infty^2. \quad (1.9)$$

Letting  $m \rightarrow +\infty$  and choosing  $T$  small enough so that  $\bar{C}_T \leq 1/2$ , we deduce that strong uniqueness holds on a sufficiently small time interval. Iterating this procedure in time gives the result on  $\mathbb{R}^+$  from usual Markov arguments involving the regular versions of conditional expectations, see e.g. [SV79].

#### 1.4 Regularization properties of the underlying PDE (1.5): strategy of proof and primer

As mentioned above, the regularization properties of the PDE (1.6) given by estimate (1.8) in Theorem 2 are the core of this work. Smoothing properties of linear partial differential operators of second order with non-degenerate diffusion matrix have been widely studied in the literature and, in that setting, the estimates of Theorem 2 are well known (see *e.g.* the book of Friedman [Fri64] or of Bass [Bas97]). In our case, the story is rather different since the diffusion matrix  $Ba$  of the system is totally degenerate in the directions 2 to  $n$ . However, as we already emphasized, the non-degeneracy condition assumed on the family of Jacobians  $(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$  allows the noise to propagate in the mentioned directions thanks to the drift. It can be viewed as a weak type Hörmander condition. Under such a condition, the operator  $L^m$  with mollified coefficients is said to be hypoelliptic<sup>§</sup> and it is well known that hypoelliptic differential operators also have some smoothing properties (see the seminal work of Hörmander [Hör67] or, for a probabilistic viewpoint, the ouvrage of Stroock [Str08]). The tricky point in our *weak Hörmander setting* is that the pointwise gradient estimates (1.8) of Theorem 2 had, to the best of our knowledge, not been established yet. Although such a setting has already been considered by several authors (see *e.g.* Delarue and Menozzi [DM10] for density estimates, Menozzi [Men11], [Men18] and Priola [Pri15] for the martingale problem and also Bramanti, Cupini, Lancionelli and Priola [BCLP10], [BCLP13] for related  $L^p$  estimates and Bramanti and Zhu [BZ11] for the VMO framework). We can mention the work of Lorenzi [Lor05a] which gives gradient estimates in the degenerate kinetic like case ( $n = 2$  in our framework) when the diffusion coefficient is sufficiently smooth and the drift linear. We point out that our main estimate in Theorem 2 needs precisely to be uniform w.r.t. the mollification parameter and therefore does not depend on the smoothness of  $\mathbf{F}^m, a^m$ , but only on known parameters appearing in **(A)**. Again, this is what would also allow to transfer those bounds to equation (1.5) from a suitable compactness argument, extending well known results for non-degenerate diffusions with Hölder coefficients to the current degenerate setting.

To prove this result our main strategy rests upon the *parametrix approach* see *e.g.* the work of McKean and Singer [MS67] or the book of Friedman [Fri64]. Roughly speaking,

---

§. Pay attention that this is not the case for  $L$  whose coefficients do not have the required smoothness in  $(\mathbf{T}_\beta)$  to compute the corresponding Lie brackets.

it consists in a perturbation argument of the operator  $L^m$  which is expanded around a *good proxy*, usually denoted by  $\tilde{L}^m$  (we keep here the super-scripts in  $m$  to emphasize that the perturbative technique we perform will concern the system (1.6) with mollified coefficients). The terminology of *good* in our setting relies to the fact that the operator  $\tilde{L}^m$  is the generator of the “closest” Gaussian approximation  $\tilde{\mathbf{X}}^m$  of  $\mathbf{X}^m$  which has generator  $L^m$ . In our case, such a process is well known and is the linearized (with respect to the source of noise) version of (1.1) whose coefficients are frozen along the curve  $(\boldsymbol{\theta}_{s,t}^m)_{s \in [t,T]}$  that solves the deterministic counterpart of (1.1) with mollified coefficients (*i.e.* with  $\sigma^m \equiv \mathbf{0}_{d,d}$ ) namely,  $\dot{\boldsymbol{\theta}}_{s,t}^m = \mathbf{F}^m(s, \boldsymbol{\theta}_{s,t}^m)$ . This process may be seen as a (non-linear) generalization of the so-called Kolmogorov example [Kol34] and we refer the reader to the work of Delarue and Menozzi [DM10] and Menozzi [Men11] for more explanations. Having this proxy at hand, the parametrix procedure consists in deriving the desired estimates for the proxy and control the expansion error.

In [CdRM17], Chaudru de Raynal and Menozzi successfully used this approach in its *backward* form to prove weak well posedness of (1.1) under less restrictive assumptions (the critical thresholds for the Hölder exponents being smaller as indicated above). In that case, the curve along which the system is frozen for the proxy is the solution of the *backward* deterministic counterpart of (1.1)). This *backward* approach is very suitable when investigating the martingale problem associated with our main system since it allows to control subtly the expansion error associating precisely the coefficients  $\mathbf{F}_i$  with their corresponding differentiation operator  $D_i$  and does not require any mollification of the coefficients. Unfortunately, when trying to obtain estimates on the derivatives of the solutions of the PDE, the *backward* approach is not appropriate since the corresponding proxy does not provide an exact density and this fact does not allow to benefit from cancellation techniques which are very helpful in this context (see paragraph below).

Hence, our parametrix approach will be of *forward*¶ form as done in the work of Chaudru de Raynal [CdR17]. This is, in fact, a non-trivial generalization of the approach developed in the aforementioned paper where the strong well posedness of (1.1) is obtained when  $n = 2$ . Indeed, the strategy used in [CdR17] is not adapted to this general case because of some subtle phenomena appearing only when  $n \geq 3$ . In particular, the singularities appearing when considering the remainder term of the parametrix were in [CdR17] equilibrated *at hand* through elementary cancellation arguments, whereas the current approach takes advantage of the full-force duality results between Besov spaces (see Sections 1.4 and 2.2 below). This forward perturbative approach has also been successfully used in Chapter 6 to establish some weaker regularization properties of the PDE (1.5) through appropriate Schauder estimates.

### Regularizing properties of the degenerate Ornstein-Uhlenbeck proxy

When exploiting such a *forward* parametrix approach, a good primer to understand what could be, at best, expected, consists in investigating the regularization property of the proxy operator  $\tilde{L}$ . To be as succinct as possible, let us consider the case where  $(\tilde{L}_t)_{t \geq 0}$

---

¶. Meaning that the freezing curve  $\boldsymbol{\theta}$  solves the corresponding ODE associated with (1.1) in a forward form.

is the generator of a degenerate Ornstein-Uhlenbeck process  $(\tilde{\mathbf{X}}_t)_{t \geq 0}$  with dynamics:

$$d\tilde{\mathbf{X}}_t = \mathbf{A}_t \tilde{\mathbf{X}}_t dt + BdW_t, \quad (1.10)$$

where  $\mathbf{A}_t$  is the  $nd \times nd$  matrix with sub-diagonal block  $\mathbf{a}_{i,i-1}(t)$  of size  $d \times d$  and  $\mathbf{0}_{d,d}$  elsewhere. In particular,

$$\mathbf{A}_t = \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{a}_{2,1}(t) & \mathbf{0}_{d,d} & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & \mathbf{a}_{3,2}(t) & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & \mathbf{a}_{n,n-1}(t) & \mathbf{0}_{d,d} \end{pmatrix}. \quad (1.11)$$

The entries  $(\mathbf{a}_{i,i-1}(t))_{i \in \llbracket 2, n \rrbracket}$  are uniformly in time non-degenerate elements of  $\mathbb{R}^d \otimes \mathbb{R}^d$  (which expresses the weak Hörmander condition). The corresponding generator  $\tilde{L}_t$  writes for all  $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$ :

$$\tilde{L}_t \varphi(\mathbf{x}) = \langle \mathbf{A}_t \mathbf{x}, \mathbf{D}\varphi(\mathbf{x}) \rangle + \frac{1}{2} \Delta_{\mathbf{x}_1} \varphi(\mathbf{x}).$$

In such a case, each component  $\tilde{\mathbf{u}}_i$ ,  $i \in \llbracket 1, n \rrbracket$  of the solution  $\tilde{\mathbf{U}}$  of the corresponding system of PDEs

$$\begin{cases} (\partial_t + \tilde{L}_t) \tilde{\mathbf{U}}(t, \mathbf{x}) = \mathbf{F}(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T) \times \mathbb{R}^{nd}, \\ \tilde{\mathbf{U}}(T, \mathbf{x}) = \mathbf{0}_{nd}, \end{cases} \quad (1.12)$$

where  $\mathbf{F}$  is a non-linear (non-mollified) source satisfying  $(\mathbf{T}_\beta)$ , writes

$$\tilde{\mathbf{u}}_i(t, \mathbf{x}) = - \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \mathbf{F}_i(s, \mathbf{y}) \tilde{p}(t, s, \mathbf{x}, \mathbf{y}). \quad (1.13)$$

Above,  $\tilde{p}$  stands for the transition density of the Gaussian process  $(\mathbf{X}_v)_{v \geq 0}$  with dynamics (1.10). Using the resolvent associated with  $(\mathbf{A}_v)_{v \in [t, s]}$ , i.e.  $\partial_s \tilde{\mathbf{R}}_{s,t} = \mathbf{A}_s \tilde{\mathbf{R}}_{s,t}$ ,  $\tilde{\mathbf{R}}_{t,t} = \mathbf{I}_{nd, nd}$ , the above equation can be explicitly integrated. Precisely, for a fixed starting point  $\mathbf{x}$  at time  $t$ :

$$\tilde{\mathbf{X}}_v = \tilde{\mathbf{R}}_{v,t} \mathbf{x} + \int_t^v \tilde{\mathbf{R}}_{v,u} BdW_u. \quad (1.14)$$

Hence, the covariance between given times  $t < s$  writes  $\tilde{\mathbf{K}}_{s,t} := \int_t^s \tilde{\mathbf{R}}_{s,u} BB^* \tilde{\mathbf{R}}_{s,u}^* du$ . From (1.14), the density at time  $v = s$  and at the spatial point  $\mathbf{y}$  therefore writes:

$$\tilde{p}(t, s, \mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s-t})^{\frac{1}{2}}} \exp \left( -\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{s-t})^{-1} (\tilde{\mathbf{R}}_{s,t} \mathbf{x} - \mathbf{y}), \tilde{\mathbf{R}}_{s,t} \mathbf{x} - \mathbf{y} \right\rangle \right). \quad (1.15)$$

Note that the resolvent also appears in (1.14) and in the density. Since the drift in (1.10) is unbounded, the term  $\tilde{\mathbf{R}}_{s,t} \mathbf{x}$  actually corresponds to the transport of the initial condition  $\mathbf{x}$  through the associated deterministic and linear differential system. It is well known, see e.g. [DM10] and Section 2.1 below, that the covariance  $\tilde{\mathbf{K}}_{s,t}$  enjoys what we will call a



*good scaling property.* Precisely, for a given  $T > 0$  there exists  $C := C((\mathbf{A}_v)_{v \in [0, T]}, T) \geq 1$  s.t. for all  $\boldsymbol{\xi} \in \mathbb{R}^{nd}$ ,

$$C^{-1}(s-t)^{-1}|\mathbb{T}_{s-t}\boldsymbol{\xi}|^2 \leq \langle \tilde{\mathbf{K}}_{s,t}\boldsymbol{\xi}, \boldsymbol{\xi} \rangle \leq C(s-t)^{-1}|\mathbb{T}_{s-t}\boldsymbol{\xi}|^2, \quad (1.16)$$

where for all  $u > 0$ , we denote by  $\mathbb{T}_u$  the *intrinsic* scale matrix:

$$\mathbb{T}_u = \begin{pmatrix} u\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & u^2\mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & u^n\mathbf{I}_{d,d} \end{pmatrix}. \quad (1.17)$$

Importantly, the *good scaling property* stated in (1.16) indicates that, for a given initial time  $t$  and for all  $i \in \llbracket 1, n \rrbracket$ , each  $\mathbb{R}^d$ -valued component  $\tilde{\mathbf{X}}_s^i$  has typical fluctuations of order  $(s-t)^{i-\frac{1}{2}}$  which corresponds to those of the  $(i-1)^{\text{th}}$  iterated integrals of the Brownian motion. Accordingly, we derive that the frozen density  $\tilde{p}$  also satisfies the bound

$$\tilde{p}(t, s, \mathbf{x}, \mathbf{y}) \leq \frac{C}{(s-t)^{\frac{n^2d}{2}}} \exp\left(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\tilde{\mathbf{R}}_{s,t}\mathbf{x} - \mathbf{y})|^2\right) =: C\hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}),$$

where, up to a modification of the constants involved,  $\int_{\mathbb{R}^{nd}} d\mathbf{y} \hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) = 1$ . Similarly, the derivatives of  $\tilde{p}$  will be bounded by a density of the form  $\hat{p}_{C^{-1}}$  up to an additional multiplicative contribution reflecting the time-singularities associated to the derivation index. Precisely, for any multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ , there exists  $C_\alpha$  s.t.

$$\begin{aligned} |D_{\mathbf{x}}^\alpha \tilde{p}(t, s, \mathbf{x}, \mathbf{y})| &\leq \frac{C_\alpha}{(s-t)^{\frac{n^2d}{2} + \sum_{i=1}^n \alpha_i(i-\frac{1}{2})}} \exp\left(-C^{-1}(s-t)|\mathbb{T}_{s-t}^{-1}(\tilde{\mathbf{R}}_{s,t}\mathbf{x} - \mathbf{y})|^2\right) \\ &\leq \frac{\tilde{C}_\alpha}{(s-t)^{\sum_{i=1}^n \alpha_i(i-\frac{1}{2})}} \hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned} \quad (1.18)$$

We refer to the proof of Proposition 2 for a complete version of this statement.

To prove estimate (1.8) of Theorem 2 for the current system (1.12), it follows from the specific structure of the matrix  $B$  that we have to estimate for any  $l \in \llbracket 1, n \rrbracket$  the quantities  $D_{\mathbf{x}_l} D_{\mathbf{x}_1}^r \tilde{\mathbf{u}}_i(t, \mathbf{x})$ ,  $r \in \{0, 1\}$ . From (1.18), we thus have

$$|D_{\mathbf{x}_l} D_{\mathbf{x}_1}^r \tilde{\mathbf{u}}_i(t, \mathbf{x})| \leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |\mathbf{F}_i(s, \mathbf{y})| (s-t)^{-(l-1/2)-r/2} \hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}). \quad (1.19)$$

We now face two problems: first the  $\mathbf{F}_i$  are unbounded, second the above time singularity is, as is, not integrable. Let us consider the worst case *i.e.* when  $r = 1$ . To smoothen the time singularity, the main idea consists in using the regularity of the source term  $\mathbf{F}_i$  by exploiting precisely the fact that, once integrated through the variables  $\mathbf{y}_l$  to  $\mathbf{y}_n$ , the transition density  $\tilde{p}$  does not depend on the variable  $\mathbf{x}_l$  anymore. This is due to the structure of  $\mathbf{A}$  in (1.11), which in particular yields that the resolvent  $(\tilde{\mathbf{R}}_{s,t})_{0 \leq t \leq s \leq T}$  is lower triangular. It is hence equal to 0 when we differentiate it w.r.t.  $\mathbf{x}_l$ : this is what will

be called a *cancellation (or centering) argument* in the following. Precisely, denoting for conciseness by  $\boldsymbol{\theta}_{s,t}(\mathbf{x}) = \tilde{\mathbf{R}}_{s,t}\mathbf{x}$  (which is coherent with the notation below when handling non-linear flows), we write:

$$\int_{\mathbb{R}^{(l-1)d}} d\mathbf{y}_{1:l-1} \mathbf{F}_i(t, \mathbf{y}_1, \dots, \mathbf{y}_{l-1}, \boldsymbol{\theta}_{s,t}^l(\mathbf{x}), \dots, \boldsymbol{\theta}_{s,t}^n(\mathbf{x})) D_{\mathbf{x}_1} D_{\mathbf{x}_l} \int_{\mathbb{R}^{(n-(l-1))d}} d\mathbf{y}_{l:n} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) = 0.$$

When using this property, we obtain that

$$\begin{aligned} & |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{\mathbf{u}}_i(t, \mathbf{x})| \\ &= \left| \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (\mathbf{F}_i(s, \mathbf{y}) - \mathbf{F}_i(t, \mathbf{y}_1, \dots, \mathbf{y}_{l-1}, \boldsymbol{\theta}_{s,t}^l(\mathbf{x}), \dots, \boldsymbol{\theta}_{s,t}^n(\mathbf{x}))) D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) \right|. \end{aligned}$$

We thus obtain from (1.18):

$$\begin{aligned} & |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{\mathbf{u}}_i(t, \mathbf{x})| \\ &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |\mathbf{F}_i(s, \mathbf{y}) - \mathbf{F}_i(t, \mathbf{y}_1, \dots, \mathbf{y}_{l-1}, \boldsymbol{\theta}_{s,t}^l(\mathbf{x}), \dots, \boldsymbol{\theta}_{s,t}^n(\mathbf{x}))| \\ &\quad \times (s-t)^{-(l-1/2)-1/2} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned}$$

Then, using the regularity assumed of  $\mathbf{F}_i$ , which satisfies  $(\mathbf{T}_\beta)$ , we get that for some constant  $C$  (which possibly change from line to line)

$$\begin{aligned} |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{\mathbf{u}}_i(t, \mathbf{x})| &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \sum_{j=l}^n |\mathbf{y}_j - \boldsymbol{\theta}_{s,t}^j(\mathbf{x})|^{\beta_j} (s-t)^{-(l-1/2)-1/2} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}) \\ &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \sum_{j=l}^n \left| \frac{\mathbf{y}_j - \boldsymbol{\theta}_{s,t}^j(\mathbf{x})}{(s-t)^{j-1/2}} \right|^{\beta_j} (s-t)^{\beta_j(j-1/2)} \\ &\quad \times (s-t)^{-(l-1/2)-1/2} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}) \\ &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \sum_{j=l}^n (s-t)^{-(l-1/2)-1/2+\beta_j(j-1/2)} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}), \end{aligned}$$

which is integrable only if for all  $j \in \llbracket l, n \rrbracket$ ,  $-(l-1/2)-1/2+\beta_j(j-1/2) > -1 \iff \beta_j > ((2l-2)/(2j-1))$ . This condition actually holds if for any  $i \in \llbracket 1, n \rrbracket$ ,  $\beta_i > ((2i-2)/(2i-1))$  which is exactly the infimum assumed in  $(\mathbf{T}_\beta)$ . As we can see, there is no hope to obtain better thresholds with such a strategy. This is the reason why we said that these thresholds are *almost sharp* for the approach used here.

## Back to the perturbative analysis

Let us now briefly explain what happens when one wants to control the approximation error in the forward parametrix expansion. Coming back to our general setting and denoting by  $\tilde{p}^m$  the transition density of our proxy, we obtain from the first order parametrix

expansion the following representation for each regularized component  $\mathbf{u}_i^m$ ,  $i \in \llbracket 1, n \rrbracket$  of our solution  $\mathbf{U}^m$ : for all  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$

$$\mathbf{u}_i^m(t, \mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\{ -\mathbf{F}_i^m(s, \mathbf{y}) + (L_s^m - \tilde{L}_s^m) \mathbf{u}_i^m(s, \mathbf{y}) \right\} \tilde{p}^m(t, s, \mathbf{x}, \mathbf{y}). \quad (1.20)$$

Above there is an additional term in the right hand side, in comparison with (1.13), which is precisely the approximation error due to the parametrix expansion. It thus appears that the solution has an implicit representation which makes its derivatives themselves appear. Hence, when differentiating the above representation to derive the estimate (1.8) in Theorem 2, we obtain bounds that depend themselves on the derivatives of the solution. We then have to estimate each derivative appearing in the right hand side and use a *circular argument*. Namely, when differentiating  $\mathbf{u}_i^m(t, \mathbf{x})$ , we will obtain the required estimate provided the multiplicative constants associated with the terms  $\|\mathbf{D}\mathbf{u}_i^m\|_\infty$  and  $\|D_1\mathbf{D}\mathbf{u}_i^m\|_\infty$ , that will appear in the corresponding upper-bound for the above right hand side, are small enough (see also Section 2 of [CdR17] and Section 2.2 below for details).

Moreover, as we have already seen, in order to smoothen the time singularity appearing when we apply a cross differentiation operator in the  $l^{\text{th}}$  and  $1^{\text{st}}$  direction to the term  $\int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_s^m - \tilde{L}_s^m) \mathbf{u}_i^m(s, \mathbf{y}) \tilde{p}^m(t, s, \mathbf{x}, \mathbf{y})$  corresponding to the approximation error, we will have to *center this term around the derivatives of the solution* itself (in the sense given in the above discussion). This procedure allows us, thanks to Taylor expansions, to weaken the singularities and provides integrable (in time) terms. The dramatic point is that, when doing so, our bound involves the cross derivatives  $D_\ell \mathbf{D}\mathbf{u}_i^m$ ,  $\ell \in \llbracket 1, n \rrbracket$  whose control in supremum norm is, as suggested by the discussion done in the explicit case of a simple degenerate Ornstein-Uhlenbeck process, definitely out of reach. In fact, as told by the results in [CdR17], the only thing we could hope is that the gradient in the degenerate directions viewed as a function of the degenerate variables, i.e.  $\mathbf{D}_{2:n} \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot) := (D_2 \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot), \dots, D_n \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot))$  for any  $(t, \mathbf{x}_1) \in [0, T] \times \mathbb{R}^d$ , belongs to an appropriate anisotropic Hölder space<sup>||</sup> with regularity indexes strictly less than 1. Such spaces can as well be viewed as particular cases of anisotropic Besov spaces with corresponding positive regularity indexes. Thus, the generalized derivative of  $\mathbf{D}_{2:n} \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot)$  should belong to some anisotropic Besov space of negative regularity indexes, strictly bigger than  $-1$ .

Here is the main novelty of our approach: to tackle this problem, our main idea, in order to balance the lack of differentiation property of the full gradient, consists in putting precisely in duality the anisotropic Besov norm with negative exponent of  $(D_\ell \mathbf{D}_{2:n} \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot))_{\ell \in \llbracket 2, n \rrbracket}$ , with the corresponding anisotropic Besov norm with positive exponent of the remaining terms coming from the differentiation of (1.20), which in particular involve the coefficients of the operator  $L^m - \tilde{L}^m$  and contain the time singularities coming from the derivatives of the frozen Gaussian kernel  $\tilde{p}^m$ . To the best of our knowledge, this approach is quite new in this parametrix setting and it appears to be very robust. We refer to the proof of the main Theorem 2 in Section 2 for details and to Proposition 3.6 in the book of Lemarié-Rieusset [LR02] for duality results on Besov spaces.

We are thus led to control on the one hand the Besov norm with negative exponent of the cross derivatives of the solution, see Lemma 4, and on the other hand the Besov norm

---

||. In other words, the regularity index depends on the considered variable

with positive exponent of the remaining terms in (1.20) (involving the coefficients of the operator  $L^m - \tilde{L}^m$ ), see Lemma 3. The first control (Besov norm with negative exponent) is crucial and appears to be quite delicate. Indeed, due to the implicit representation (1.20), this estimate also involves supremum norms of the full gradient  $\mathbf{D}u_i^m$  and of the cross derivatives themselves. This again reflects the *circular* nature of the arguments needed to derive the result.

To conclude this discussion on Besov duality, let us mention that a similar strategy has been implemented in Chapter 6 in order to derive sharp Schauder estimates for the PDE (1.5) (with possibly non-trivial final condition). Therein, since we were interested in controlling the Hölder norm of the solution, the duality was anyhow used the other way round: positive regularity indexes for the solution and negative ones for the remaining terms of the perturbative expansion.

Let us close this discussion coming back to Remark 1. As we emphasized, in comparison with the non-degenerate result, Theorem 1 should hold assuming that the drift  $\mathbf{F}_1$  belongs to a suitable  $L_q - L_p$  space w.r.t. time and the non-degenerate variable  $\mathbf{x}_1$ . We are convinced that this is the case but we deliberately decide not to tackle this setting in order to keep this work shorter and more coherent. Indeed, in this case, the difficulty comes from the estimate on the second order derivative in the non-degenerate direction of the first component of the solution  $\mathbf{U}^m$ , namely  $D_{\mathbf{x}_1} D_{\mathbf{x}_1} \mathbf{u}_1^m$  (which is a part of the main estimate (1.8) in Theorem 2). The point is to establish for this quantity an  $L_q - L_p$  control. This cannot be derived from the previously described approach and requires harmonic analysis techniques (see also [KR05]). The main problem to establish the estimate is mainly due to the source term, which is actually  $\mathbf{F}_1$ . To prove it, the main idea consists in exploiting the results of Menozzi [Men18] (where such an estimate is proved under the assumption that the drift is Lipschitz) through the tools developed in [CdRM17] (*backward* parametrix approach for drift  $\mathbf{F}$  whose first component may be in  $L_q - L_p$  and the other ones in Hölder spaces). Then, the Zvonkin Transform should also be tuned a little bit following the strategy developed by Veretennikov (see e.g. [Ver80] and [FFPV17]). Such a program would surely toughen our paper without adding any surprising result and we prefer to focus on the novelty of the approach based on duality results for Besov spaces and the generalization of the strong uniqueness result to the whole chain (*i.e.* to any arbitrary  $n \geq 1$ ) rather than drowning the reader into additional technical considerations.

## 2 Perturbation techniques for the PDE : proof of Theorem 2

In order to keep the notations as clear as possible, we forget the superscript  $m$  standing for the mollifying procedure and we suppose that the following assumptions hold:

**Assumption (AM).** We say that assumptions (AM) hold if the assumptions gathered in (A) hold true and the coefficients  $\mathbf{F}$ ,  $a$  are infinitely differentiable functions with bounded derivatives of all order for  $a$  and greater than 1 for the coefficient  $\mathbf{F}$ .

In the whole section, we consider a fixed final time  $T > 0$  which is meant to be *small*,

i.e.  $T \ll 1$ . Let us consider for this section a generic PDE with generator corresponding to (1.4) and scalar source  $f$  having the same Hölder regularity than the drift terms in (1.1) (i.e. the scalar function  $f$  below can be any of the entries of the  $\mathbb{R}^d$ -valued  $(\mathbf{F}_i)_{i \in \llbracket 1, n \rrbracket}$  in the dynamics (1.1)). Namely, we concentrate on

$$\begin{cases} (\partial_t u + L_t u)(t, \mathbf{x}) = -f(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \\ u(T, \mathbf{x}) = 0, \end{cases} \quad (2.1)$$

where  $(L_t)_{t \geq 0}$  is defined in (1.4) and stands for the generator associated with (1.1) when the coefficients are smooth.

The key result to prove strong uniqueness for the SDE (1.1) is actually the following theorem from which one easily derive Theorem 2 for each component of the solution of the systems (1.5) and then the result in full generality.

**Theorem 3 (Pointwise bounds for the derivatives of the PDE (2.1))** *There exists  $\gamma := \gamma((\mathbf{A})) > 0$  and  $C := C((\mathbf{A})) > 0$  s.t.*

$$\|\mathbf{D}u\|_\infty + \|\mathbf{D}(D_1 u)\|_\infty \leq CT^\gamma, \quad (2.2)$$

with obvious extension of the definition in (1.3) to the current scalar case.

The proof of Theorem 3 is performed in Section 2.2 through the forward parametrix approach consisting in considering a suitable *proxy* semi-group around which the initial solution of (2.1) can be expanded. To this end we first investigate in Section 2.1 below the linearized Gaussian process deriving from the dynamics in (1.1) which will provide the suitable model for the parametrix.

## 2.1 Gaussian proxy and associated controls

### Linearization of the dynamics

Fix some freezing points  $(\tau, \boldsymbol{\xi}) \in [0, T] \times \mathbb{R}^{nd}$ . For fixed initial conditions  $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ , a natural linearization associated with the mollified version of (1.1) writes

$$\begin{aligned} d\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})} &= [\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})) + D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi}))(\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})} - \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi}))]dv + B\sigma(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi}))dW_v, \\ \forall v \in [t, s], \tilde{\mathbf{X}}_t^{(\tau, \boldsymbol{\xi})} &= \mathbf{x}, \end{aligned} \quad (2.3)$$

where

$$\dot{\boldsymbol{\theta}}_{v, \tau}(\boldsymbol{\xi}) = \mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})), \quad v \in [0, T], \quad \boldsymbol{\theta}_{\tau, \tau}(\boldsymbol{\xi}) = \boldsymbol{\xi}, \quad (2.4)$$

and  $D\mathbf{F}(v, \cdot)$  denotes the subdiagonal of the Jacobian matrix  $\mathbf{D}\mathbf{F}(v, \cdot)$ . Namely, for  $\mathbf{z} \in \mathbb{R}^{nd}$ :

$$D\mathbf{F}(v, \mathbf{z}) = \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ D_{\mathbf{z}_1} \mathbf{F}_2(v, \mathbf{z}) & \mathbf{0}_{d,d} & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & D_{\mathbf{z}_2} \mathbf{F}_3(v, \mathbf{z}^{2:n}) & \mathbf{0}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & D_{\mathbf{z}_{n-1}} \mathbf{F}_n(v, \mathbf{z}_{n-1}, \mathbf{z}_n) & \mathbf{0}_{d,d} \end{pmatrix}.$$

From the non-degeneracy of  $\sigma$  and Hörmander like condition, the Gaussian process defined by (2.5) admits a density  $\tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \cdot)$  which is suitably controlled (see Proposition 2 below and for instance [DM10], [CdRM17]).

We explicitly integrate (2.3) to obtain for all  $v \in [t, s]$ :

$$\begin{aligned} \tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})} &= \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, t)\mathbf{x} + \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) \left( \mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) - D\mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi}))\boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi}) \right) du \\ &\quad + \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) dW_u \\ &=: \mathbf{m}_{v, t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) dW_u, \end{aligned}$$

where  $(\tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u))_{t \leq u, v \leq s}$  stands for the resolvent associated with the collection of partial gradients in  $(D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})))_{v \in [t, s]}$  which satisfies for  $v \in [t, s]$ :

$$\partial_v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, t) = D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})) \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, t), \quad \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(t, t) = \mathbf{I}_{nd \times nd}. \quad (2.5)$$

Note in particular that since the partial gradients are subdiagonal  $\det(\tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, t)) = 1$ .

Also, for  $v \in [t, s]$ , we recall that  $\mathbf{m}_{v, t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})$  stands for the mean of  $\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})}$  and corresponds as well to the solution of (2.3) when  $\sigma = 0$  when the starting point is  $\mathbf{x}$ . We write:

$$\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})} = \mathbf{m}_{v, t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) dW_u, \quad v \in [t, s]. \quad (2.6)$$

Importantly, we point out that  $\mathbf{x} \in \mathbb{R}^{nd} \mapsto \mathbf{m}_{v, t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})$  is **affine** w.r.t. the starting point  $\mathbf{x}$ . Precisely, for  $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{nd}$ :

$$\mathbf{m}_{v, t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x} + \mathbf{x}') = \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, t)\mathbf{x}' + \mathbf{m}_{v, t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}). \quad (2.7)$$

We first give in the next proposition a key estimates on the covariance matrix associated with (2.6) and its properties w.r.t. a suitable scaling of the system.

**Proposition 1 (Good Scaling Properties of the Covariance Matrix)** *The covariance matrix of  $\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})}$  in (2.6) writes:*

$$\tilde{\mathbf{K}}_{v, t}^{(\tau, \boldsymbol{\xi})} := \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) B a(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) B^* \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u)^* du.$$

Uniformly in  $(\tau, \boldsymbol{\xi}) \in [0, T] \times \mathbb{R}^{nd}$  and  $s \in [0, T]$ , it satisfies a good scaling property in the sense of Definition 3.2 in [DM10] (see also Proposition 3.4 of that reference). That is, for all fixed  $T > 0$ , there exists  $C_{2.8} := C_{2.8}(\mathbf{A}, T) \geq 1$  s.t. for all  $0 \leq t < v \leq s \leq T$ , for all  $(\tau, \boldsymbol{\xi}) \in [0, T] \times \mathbb{R}^{nd}$ :

$$\forall \boldsymbol{\zeta} \in \mathbb{R}^{nd}, \quad C_{2.8}^{-1}(v-t)^{-1} |\mathbb{T}_{v-t} \boldsymbol{\zeta}|^2 \leq \langle \tilde{\mathbf{K}}_{v, t}^{(\tau, \boldsymbol{\xi})} \boldsymbol{\zeta}, \boldsymbol{\zeta} \rangle \leq C_{2.8}(v-t)^{-1} |\mathbb{T}_{v-t} \boldsymbol{\zeta}|^2, \quad (2.8)$$

where we again use the notation introduced in (1.17) for the scaling matrix  $\mathbb{T}_{v-t}$ .

The proof of the above proposition readily follows from Proposition 3.3 and Lemma 3.6 in [DM10]. We now state some important density bounds for the linearized model.

**Proposition 2 (Density of the linearized dynamics)** *Under (A), we have that, for all  $s \in (t, T]$  the random variable  $\tilde{\mathbf{X}}_s^{(\tau, \xi)}$  in (2.6) admits a Gaussian density  $\tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \cdot)$  which writes for all  $\mathbf{y} \in \mathbb{R}^{nd}$ :*

$$\tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) := \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y}), \mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y} \right\rangle\right), \quad (2.9)$$

with  $\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)}$  as in Proposition 1. Also, there exists  $C := C((\mathbf{A}), T) > 0$  s.t. for all multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ ,  $|\alpha| \leq 3$  and denoting by  $D_{\mathbf{x}}^\alpha := D_{\mathbf{x}_1}^{\alpha_1} \dots D_{\mathbf{x}_n}^{\alpha_n}$ , we have:

$$\begin{aligned} |D_{\mathbf{x}}^\alpha \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| &\leq \frac{C}{(s-t)^{\sum_{i=1}^n \alpha_i (i-\frac{1}{2}) + \frac{n^2 d}{2}}} \exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})|^2\right) \\ &=: \frac{C}{(s-t)^{\sum_{i=1}^n \alpha_i (i-\frac{1}{2})}} \hat{p}_{C^{-1}}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned} \quad (2.10)$$

**Proof:** Expression (2.9) readily follows from (2.5). The control (2.10) is then a direct consequence of Proposition 1 for  $\alpha = 0$ . Differentiating w.r.t.  $\mathbf{x}$  recalling from (2.7) that  $\mathbf{x} \mapsto \mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x})$  is affine yields:

$$D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) = - \left[ [\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* (\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y}) \right]_j \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \quad (2.11)$$

The point is now to use scaling arguments. We can first rewrite

$$[\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* (\tilde{\mathbf{K}}_{s,t}^{(\tau, \xi)})^{-1} = (s-t) [\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* \mathbb{T}_{s-t}^{-1} (\widehat{\mathbf{K}}_1^{s,t})^{-1} \mathbb{T}_{s-t}^{-1}, \quad (2.12)$$

where  $\widehat{\mathbf{K}}_1^{s,t}$  is the covariance matrix of the rescaled process  $((s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \tilde{\mathbf{X}}_{t+v(s-t)}^{t, \mathbf{x}})_{v \in [0,1]}$  at time 1. From the good-scaling property of Proposition 1, it is plain to derive that  $\widehat{\mathbf{K}}_1^{s,t}$  is a non-degenerate bounded matrix, i.e. there exists  $\hat{C} \geq 1$  s.t. for all  $\zeta \in \mathbb{R}^{nd}$ ,  $\hat{C}^{-1} |\zeta|^2 \leq \langle \widehat{\mathbf{K}}_1^{s,t} \zeta, \zeta \rangle \leq \hat{C} |\zeta|^2$ . A similar rescaling argument yields on the deterministic system (2.5) of the resolvent yields that  $\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)$  can also be written as:

$$[\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* = \mathbb{T}_{s-t}^{-1} \left[ \widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0) \right]^* \mathbb{T}_{s-t}, \quad (2.13)$$

where again  $\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0)$  is the resolvent at time 1 of the rescaled system  $\mathbb{T}_{s-t} [\tilde{\mathbf{R}}^{(\tau, \xi)}(t + v(s-t), t)]^* \mathbb{T}_{s-t}^{-1} = \left( [\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(v, 0)]^* \right)_{v \in [0,1]}$  associated with (2.5). From the analysis performed in Lemma 5.1 in [HM16] (see also the proof of Proposition 3.7 in [DM10]) one derives that there exists  $\hat{C}_1$  s.t. for all  $\zeta \in \mathbb{R}^{nd}$ ,  $|\left[ \widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0) \right]^* \zeta| \leq \hat{C}_1 |\zeta|$ . Equations (2.11), (2.12) and (2.13) therefore yield:

$$\begin{aligned} &|D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| \\ &\leq (s-t)^{-j+\frac{1}{2}} \left| \left( \left[ \widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0) \right]^* (\widehat{\mathbf{K}}_1^{s,t})^{-1} ((s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})) \right) \right|_j \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) \\ &\leq C (s-t)^{-j+\frac{1}{2}} (s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})| \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned}$$

From the explicit expression (2.9), Proposition 1 and the above equation, we eventually derive:

$$\begin{aligned}
 |D_{\mathbf{x}_j} \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})| &\leq \frac{C}{(s-t)^{j-\frac{1}{2}}} \left( (s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) - \mathbf{y})| \right) \\
 &\quad \times \frac{1}{(s-t)^{\frac{n^2 d}{2}}} \exp \left( -C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) - \mathbf{y})|^2 \right) \\
 &\leq \frac{C}{(s-t)^{j-\frac{1}{2}}} \hat{p}_{C^{-1}}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}),
 \end{aligned}$$

up to a modification of  $C$ , which gives the statement for one partial derivative. The controls on the higher order derivatives are obtained similarly (see e.g. the proof of Lemma 5.5 of [DM10] for the bounds on  $D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})$ ).  $\square$

Now, let us specify a useful control involving the previous Gaussian kernel which will be exploited in some cancellation techniques.

**Proposition 3** *For all  $k \in \llbracket 1, n \rrbracket$ ,  $0 \leq t \leq s \leq T$ ,  $(\mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$ , and  $\mathbf{M} \in \mathbb{R}^d$  the following identity hold:*

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) \langle \mathbf{M}, (\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_k \rangle d\mathbf{y} = \langle \mathbf{M}, \mathbf{1}_d \rangle, \quad \mathbf{1}_d = (1, \dots, 1) \in \mathbb{R}^d. \quad (2.14)$$

**Proof:** From Proposition 2, we have  $\int_{\mathbb{R}^{nd}} \tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) (\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}))_k d\mathbf{y} = \mathbf{0}_d$ . Differentiating this expression w.r.t.  $\mathbf{x}_k$  and using the Leibniz formula (recalling as well the identity (2.7) which yields  $D_{\mathbf{x}_k} [\mathbf{m}_{s,t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x})]_k = (\tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(s, t))_{k,k} = \mathbf{I}_{d,d}$ ) gives (2.14).  $\square$

### Associated inhomogeneous semi-group

Fix  $t \in [0, T]$ ,  $\boldsymbol{\xi} \in \mathbb{R}^{nd}$ . With the notations of the previous paragraph, we introduce the following inhomogeneous semi-group associated with (2.3) for  $\tau = t$ . Namely, for all  $s \in (t, T]$ ,  $g \in B_{\text{lin}}(\mathbb{R}^{nd}, \mathbb{R})$  (space of measurable functions with linear growth),  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$\tilde{P}_{s,t}^{\boldsymbol{\xi}} g(\mathbf{x}) := \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{(t, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y}) g(\mathbf{y}). \quad (2.15)$$

From now on, we will write with a slight abuse of notation  $\tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y}) := \tilde{p}^{(t, \boldsymbol{\xi})}(t, s, \mathbf{x}, \mathbf{y})$ , i.e. we omit the freezing parameter in time when it corresponds to the considered starting time. One can derive from Proposition 2 the following important regularization result.

**Lemma 1 (Regularization effects for the inhomogeneous semi-group)** *Let  $0 < \gamma \leq 1$ . There exists  $C := C(\mathbf{A})$  such that for all  $\ell \in \{0, 1\}$ ,  $l \in \llbracket 1, n \rrbracket$ ,  $k \in \llbracket 1, n \rrbracket$ ,  $\mathbf{x} \in \mathbb{R}^{nd}$ :*

$$|D_{\mathbf{x}_1}^\ell D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \left| \left( \cdot - \mathbf{m}_{s,t}^{(t, \boldsymbol{\xi})}(\mathbf{x}) \right)_k \right|^\gamma \right) (\mathbf{x})|_{\boldsymbol{\xi}=\mathbf{x}} \leq C (s-t)^{-\frac{\ell}{2} - (l-\frac{1}{2}) + \gamma(k-\frac{1}{2})}. \quad (2.16)$$

*Let  $f$  be a  $\beta$ -Hölder continuous functions where  $\beta := (\beta_1, \dots, \beta_n) \in (0, 1]^n$  is a multi-index and for  $i \in \llbracket 1, n \rrbracket$ ,  $\beta_i$  stands for the Hölder regularity of  $f$  in the variable  $\mathbf{x}_i$ . The following result then holds.*



- There exists  $C := C(\mathbf{A})$  s.t. for all  $l \in \{0, 1\}$ ,  $l \in \llbracket 1, n \rrbracket$ ,  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$|D_{\mathbf{x}_1}^\ell D_{\mathbf{x}_l} \tilde{P}_{s,t}^\xi \left( f(\cdot) - f(\mathbf{m}_{s,t}^{(t,\xi)}(\mathbf{x})) \right) (\mathbf{x})|_{\xi=\mathbf{x}} \leq C \sum_{j=1}^n [f_j(s, \cdot)]_{\beta_j} (s-t)^{-\frac{\ell}{2} - (l-\frac{1}{2}) + \beta_j(j-\frac{1}{2})}, \quad (2.17)$$

where  $[f_j(s, \cdot)]_{\beta_j} := \sup_{z \neq z' \in \mathbb{R}^d, (\mathbf{z}_{1:j-1}, \mathbf{z}_{j+1:n}) \in \mathbb{R}^{(n-1)d}} \frac{|f(\mathbf{z}_{1:j-1}, z', \mathbf{z}_{j+1:n}) - f(\mathbf{z}_{1:j-1}, z, \mathbf{z}_{j+1:n})|}{|z-z'|^{\beta_j}}$  stands for the Hölder continuity modulus of order  $\beta_j$  of  $f$  in its  $j^{\text{th}}$  variable.

- Centering arguments. For all  $l \in \llbracket 1, n \rrbracket$ ,  $k \leq l$ , it holds that:

$$D_{\mathbf{x}_l} \tilde{P}_{s,t}^\xi \left( f(\cdot) \right) (\mathbf{x}) = D_{\mathbf{x}_l} \tilde{P}_{s,t}^\xi \left( f(\cdot) - f(\cdot_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) \right) (\mathbf{x}). \quad (2.18)$$

- As particular cases of the previous items, we have that there exists  $C := C(\mathbf{A})$  s.t. for all  $l \in \llbracket 1, n \rrbracket^2$ ,  $\mathbf{x} \in \mathbb{R}^{nd}$ :

$$\begin{aligned} |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{P}_{s,t}^\xi f(\mathbf{x})|_{\xi=\mathbf{x}} &\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\beta_j} (s-t)^{-l + \beta_j(j-\frac{1}{2})}, \\ |D_{\mathbf{x}_l} \tilde{P}_{s,t}^\xi f(\mathbf{x})|_{\xi=\mathbf{x}} &\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\beta_j} (s-t)^{-(l-\frac{1}{2}) + \beta_j(j-\frac{1}{2})}. \end{aligned} \quad (2.19)$$

**Proof:** Let us first mention that identities (2.16) and (2.17) are direct consequences of Proposition 2.

Centering arguments like (2.18) will be a crucial tool in the analysis below. To justify such an identity, write:

$$\begin{aligned} D_{\mathbf{x}_l} \tilde{P}_{s,t}^\xi \left( f(\cdot_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) \right) (\mathbf{x}) &= D_{\mathbf{x}_l} \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) f(\mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) \\ &= D_{\mathbf{x}_l} \int_{\mathbb{R}^{nd}} d\mathbf{y}_{1:l-1} f(\mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) \int_{\mathbb{R}^{nd}} d\mathbf{y}_{l:n} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}). \end{aligned}$$

Now, the structure of the linearized dynamics (2.3) yields that the variable  $\mathbf{x}_l$  only appears in  $(\tilde{\mathbf{R}}^{(t,\xi)}(s, t)\mathbf{x})_{l:n}$  from its  $l^{\text{th}}$  to its  $n$   $d$ -dimensional block. Therefore, setting e.g.  $\tilde{\mathbf{y}}_{l:n} = \mathbf{y}_{l:n} - (\tilde{\mathbf{R}}^{(t,\xi)}(s, t)\mathbf{x})_{l:n}$ , the integrated quantity  $\int_{\mathbb{R}^{nd}} d\mathbf{y}_{l:n} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y})$  does not depend on  $\mathbf{x}_l$  anymore. This gives the statement (2.18).

Let us now prove (2.19). The idea is to use first a centering argument w.r.t. to the variables  $l$  to  $n$ . Namely,

$$\begin{aligned} &\left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^\xi f(s, \cdot) \right] (\mathbf{x}) \right| \\ &= \left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^\xi \left( f(s, \cdot) - f(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) \right) \right] (\mathbf{x}) \right| \\ &\leq C (s-t)^{-(l-1/2)-1/2} \left[ \hat{P}_{s,t}^\xi \left( \sum_{j=l}^n [f_j(s, \cdot)]_{\beta_j} |(\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_j|^{\beta_j} \right) \right] (\mathbf{x}), \end{aligned}$$

where for a function  $g \in B_{\text{lin}}(\mathbb{R}^d, \mathbb{R})$  we denote,

$$\hat{P}_{s,t}^\xi g(\mathbf{x}) = \int_{\mathbb{R}^{nd}} d\mathbf{y} \hat{p}_{C^{-1}}^{(t,\xi)}(t, s, \mathbf{x}, \mathbf{y}) g(\mathbf{y}) := \int_{\mathbb{R}^{nd}} d\mathbf{y} \hat{p}_{C^{-1}}^\xi(t, s, \mathbf{x}, \mathbf{y}) g(\mathbf{y}),$$

where  $\hat{p}_{C^{-1}}^{(t,\boldsymbol{\xi})}(t, s, \mathbf{x}, \cdot) := \hat{p}_{C^{-1}}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \cdot)$  is the density appearing in Proposition 2\*\*. That is,  $\hat{P}_{s,t}^{\boldsymbol{\xi}}$  can somehow be viewed as the *pseudo* semi-group associated with the density  $\hat{p}_{C^{-1}}^{(t,\boldsymbol{\xi})}$  appearing in Proposition 2. The dependence in  $C$  in  $\hat{P}_{s,t}^{\boldsymbol{\xi}}$  is then omitted for notational simplicity.

In the current regularized setting, it is plain to observe that taking  $\boldsymbol{\xi} = \mathbf{x}$ ,  $\mathbf{m}_{s,t}^{(t,\mathbf{x})}(\mathbf{x}) = \boldsymbol{\theta}_{s,t}(\mathbf{x})$ . Therefore, from (2.10):

$$\begin{aligned} \left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) \right|_{\boldsymbol{\xi}=\mathbf{x}} &\leq C \sum_{j=l}^n \int_{\mathbb{R}^{nd}} [f_j(s, \cdot)]_{\beta_j} d\mathbf{y} \frac{|(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_j|^{\beta_j}}{(s-t)^l} \hat{p}_{C^{-1}}^{\mathbf{x}}(t, s, \mathbf{x}, \mathbf{y}) \\ &\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\beta_j} \int_{\mathbb{R}^{nd}} d\mathbf{y} \frac{|(\mathbf{y} - \mathbf{m}_{s,t}^{(t,\mathbf{x})}(\mathbf{x}))_j|^{\beta_j}}{(s-t)^l} \hat{p}_{C^{-1}}^{\mathbf{x}}(t, s, \mathbf{x}, \mathbf{y}) \\ &\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\beta_j} (s-t)^{-l+\beta_j(j-\frac{1}{2})}. \end{aligned}$$

The control for  $|D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\boldsymbol{\xi}} f(\mathbf{x})|_{\boldsymbol{\xi}=\mathbf{x}}$  is derived similarly.  $\square$

We state in the lemma below a useful control to obtain through Lemma 1 some smoothing effects for the degenerate part of the operator. The statement readily follow from  $(\mathbf{T}_\beta)$ .

**Lemma 2** *From the smoothness assumption on the drift coefficient in  $(\mathbf{T}_\beta)$ , there exists  $C := C((\mathbf{A}))$  s.t. for all  $\ell \in \llbracket 2, n \rrbracket$ ,  $k \geq \ell$ , and for all  $(s, \mathbf{x}, \boldsymbol{\xi}) \in [0, T] \times (\mathbb{R}^{nd})^2$ :*

$$\begin{aligned} &\left| \left( \mathbf{F}_\ell(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{\ell-1}} \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1} \right) \right| \\ &\leq C \left\{ \sum_{j=\ell}^{k-1} \left\{ [(\mathbf{F}_\ell)_j(s, \cdot)]_{\beta_j} |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_j|^{\beta_j} \right\} + [(D_{\mathbf{x}_{\ell-1}} \mathbf{F}_\ell)_{\ell-1}(s, \cdot)]_\eta |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_{\ell-1}|^{1+\eta} \right\}. \end{aligned}$$

## 2.2 Control of the sensitivities: proof of Theorem 3

To prove Theorem 3, the idea is to expand the solution of the PDE with regularized coefficients around a suitable proxy, as explained in Section 1.4. The proxy used here is the Gaussian process introduced in Section 2.1 for a suitable freezing parameter  $\boldsymbol{\xi}$  to be specified later on. Then, the Duhamel formula (or first order parametrix expansion) yields:

---

\*\* we again use from now on the same notational simplification as for  $\tilde{p}^{(t,\boldsymbol{\xi})}(t, s, \mathbf{x}, \cdot) := \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \cdot)$ , i.e. we omit the time superscript when it corresponds to the first argument of the density.

$$\begin{aligned}
 u(t, \mathbf{x}) &= \int_t^T \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) ds + \int_t^T \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( (L_s - \tilde{L}_s^{\boldsymbol{\xi}}) u \right) (s, \cdot) \right] (\mathbf{x}) \\
 &= \int_t^T \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) ds + \int_t^T \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \langle \mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \\
 &\quad + \int_t^T \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \frac{1}{2} \text{Tr} \left[ a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) ds \\
 &\quad + \int_t^T \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \sum_{i=2}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) \right. \right. \\
 &\quad \left. \left. - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}, D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) ds, \tag{2.20}
 \end{aligned}$$

for any  $\boldsymbol{\xi}$  in  $\mathbb{R}^{nd}$ .

To establish (2.2) we need to differentiate the above expression w.r.t.  $(\mathbf{x}_l)_{l \in \llbracket 1, n \rrbracket}$  and then w.r.t.  $\mathbf{x}_1$  and  $(\mathbf{x}_l)_{l \in \llbracket 1, n \rrbracket}$  in order to obtain estimates depending only on known parameters in (A). Differentiating first this expression w.r.t.  $\mathbf{x}_l, l \in \llbracket 1, n \rrbracket$  we obtain:

$$\begin{aligned}
 &D_{\mathbf{x}_l} u(t, \mathbf{x}) \\
 &= \int_t^T D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) ds + \int_t^T D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( (L_s - \tilde{L}_s^{\boldsymbol{\xi}}) u \right) (s, \cdot) \right] (\mathbf{x}) ds \\
 &= \left\{ \int_t^T D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) ds \right. \\
 &\quad + \int_t^T D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \langle (\mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) ds \\
 &\quad + \int_t^T D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \frac{1}{2} \text{Tr} \left[ \left( a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) ds \left. \right\} \\
 &\quad + \left\{ \int_t^T D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^{\boldsymbol{\xi}} \left( \sum_{i=2}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) \right. \right. \right. \\
 &\quad \left. \left. \left. - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}, D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) ds \right\} \\
 &=: \int_t^T ds H_l^{\boldsymbol{\xi}}(s, \mathbf{x}) + \int_t^T ds I_l^{\boldsymbol{\xi}}(s, \mathbf{x}). \tag{2.21}
 \end{aligned}$$

The term  $H_l^{\boldsymbol{\xi}}(s, \mathbf{x})$  gathers all the derivatives of the solution w.r.t. the non-degenerate variables whereas  $I_l^{\boldsymbol{\xi}}(s, \mathbf{x})$  precisely gathers the derivatives w.r.t. the degenerate ones. We will now start from the representation (2.21) which we will again differentiate w.r.t. the non-degenerate variable  $\mathbf{x}_1$  in order to prove the estimates of Theorem 3 concerning the second order derivatives which are the trickiest ones. Indeed, as it has been succinctly explained in Section 1.4, when differentiating the kernel associated with the frozen semi-group defined by (2.15) we generate an *a priori* not integrable time singularity which

then needs to be smoothen by using, among others, tools developed in Lemma 1 (centering or cancellation arguments). The worst case then corresponds to the higher order of differentiation, namely  $D_{\mathbf{x}_l}(D_{\mathbf{x}_1}u(t, \mathbf{x}))$  which, as suggested by Proposition 2, generates a time singularity of order  $1/2 + (l - 1/2)$  in the time integrand of the r.h.s. of (2.21). We then only concentrate on this term and omit the proof of the statement concerning the boundedness of the gradient  $D_{\mathbf{x}_l}u(t, \mathbf{x})$  which could be shown more directly.

The proof will be divided into two parts: we first handle the non-degenerate part of the operator (*i.e.* the estimate for  $D_{\mathbf{x}_1}H_l(s, \mathbf{x})$ ) and then the degenerate part (*i.e.* the estimate for  $D_{\mathbf{x}_1}I_l(s, \mathbf{x})$ ) which is a bit more involved.

Finally, before entering into the proof, we introduce some notations to ease the reading. We point out that the contributions  $H_l^\xi, I_l^\xi$  do depend on the freezing variable  $\xi$ . We omit this dependence for notational convenience when, as default  $\xi$  coincides with the spatial argument  $\mathbf{x}$  of the term. We write in this case  $H_l^\xi(s, \mathbf{x})|_{\xi=\mathbf{x}} =: H_l(s, \mathbf{x})$  (resp.  $I_l^\xi(s, \mathbf{x})|_{\xi=\mathbf{x}} =: I_l(s, \mathbf{x})$ ). The freezing parameter explicitly appears when a more careful choice is needed (see Section 3.2). Accordingly with the definition in (1.2), we will also use, with a slight abuse of notation recalling from (1.1) that the  $j^{\text{th}}$  variable appears with Hölder regularity  $\beta_j$  up to the level  $j$  of the chain, the notation

$$[(\mathbf{F}_{1:j})_j(s, \cdot)]_{\beta_j} := \max_{i \in \llbracket 1, j \rrbracket} [(\mathbf{F}_i)_j(s, \cdot)]_{\beta_j}. \quad (2.22)$$

**Source term and non-degenerate part of the operator: estimates for  $D_{\mathbf{x}_1}H_l(s, \mathbf{x})$ .**

We first focus on the source term and the derivatives w.r.t. the non-degenerate variable  $\mathbf{x}_1$  in (2.20) (three first terms in the r.h.s denoted from now on accordingly with (2.21) by  $H_{l,1}(s, \mathbf{x}), H_{l,2}(s, \mathbf{x}), H_{l,3}(s, \mathbf{x})$ ).

For all  $l \in \llbracket 1, n \rrbracket$ , one readily derives from Lemma 1 (centering argument from the variables  $l$  to  $n$ ) that for the source term:

$$|D_{\mathbf{x}_1}H_{l,1}(s, \mathbf{x})| := \left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[ \tilde{P}_{s,t}^\xi f(s, \cdot) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\beta_j} (s-t)^{-l+\beta_j(j-1/2)}. \quad (2.23)$$

Those terms are integrable in time as soon as

$$\beta_j(j - \frac{1}{2}) - l > -1, \quad j \in \llbracket l, n \rrbracket \Leftrightarrow \beta_j \in \left( \frac{2j-2}{2j-1}; 1 \right]. \quad (2.24)$$

These thresholds are precisely those appearing in assumption  $(\mathbf{T}_\beta)$  and which lead to strong uniqueness for the associated SDE. Through our perturbative approach we will actually show that they are also precisely those leading to Theorem 3.

Similarly, from (2.20), (2.21), for the drift associated with the non-degenerate part,

we first rewrite from the centering properties of Lemma 1:

$$\begin{aligned}
 |D_{\mathbf{x}_1} H_{l,2}(s, \mathbf{x})| &:= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \langle \mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\xi)), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) ds \right| \Big|_{\xi=\mathbf{x}} \\
 &\leq \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \langle \mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \\
 &\quad + \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \langle \mathbf{F}_1(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\xi)), \right. \right. \right. \\
 &\quad \left. \left. \left. (D_1 u(s, \cdot) - D_1 u(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi))) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}}.
 \end{aligned}$$

Thanks to Lemma 1, this again gives

$$\begin{aligned}
 |D_{\mathbf{x}_1} H_{l,2}(s, \mathbf{x})| &\leq C \left( \sum_{j=l}^n \|D_1 u\|_{\infty} [(\mathbf{F}_1)_j(s, \cdot)]_{\beta_j} (s-t)^{-l+\beta_j(j-\frac{1}{2})} \right. \\
 &\quad \left. + \sum_{j=l}^n \sum_{k=1}^{l-1} \|D_1 D_j u\|_{\infty} [(\mathbf{F}_1)_k(s, \cdot)]_{\beta_k} (s-t)^{-l+\beta_k(k-\frac{1}{2})+(j-\frac{1}{2})} \right), \tag{2.25}
 \end{aligned}$$

leading precisely to the same integrability thresholds of equation (2.24) and assumption  $(\mathbf{T}_{\beta})$  (as for the source term). The idea behind this control is crucial. We first handle, with the sole Hölder properties of drift and the supremum norm of  $D_1 u$ , the variables which are at a good smoothing scale w.r.t. the induced singularity. For the remaining term, which exhibits for the drift non-sufficient smoothing effects, we then additionally exploit a cancellation argument involving the gradient of the solution itself, which consequently makes the cross derivatives appear.

Eventually, we get for the diffusive part:

$$\begin{aligned}
 &|D_{\mathbf{x}_1} H_{l,3}(s, \mathbf{x})| \\
 := &\left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \frac{1}{2} \text{Tr} \left[ (a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\xi))) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \\
 = &\left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \frac{1}{2} \text{Tr} \left[ (a(s, \cdot) - a(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi))) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \\
 &+ \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \frac{1}{2} \text{Tr} \left[ (a(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - a(s, \boldsymbol{\theta}_{s,t}(\xi))) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \\
 =: &|D_{\mathbf{x}_1} H_{l,31}(s, \mathbf{x})| + |D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})|. \tag{2.26}
 \end{aligned}$$

The term  $D_{\mathbf{x}_1} H_{l,31}(s, \mathbf{x})$  is already centered at the *appropriate* scales, i.e. from variables  $l$  to  $n$ . We thus readily derive, similarly to the previous computations thanks to Lemma 1 and recalling that  $a$  is Lipschitz continuous, that:

$$|D_{\mathbf{x}_1} H_{l,31}(s, \mathbf{x})| \leq C \|D_1^2 u\|_{\infty} [a(s, \cdot)]_1 \sum_{j=l}^n (s-t)^{-l+(j-\frac{1}{2})} \leq C \|D_1^2 u\|_{\infty} [a(s, \cdot)]_1 (s-t)^{-\frac{1}{2}}, \tag{2.27}$$

which does not give a *critical* contribution w.r.t. the previously exhibited thresholds in (2.24) and  $(\mathbf{T}_\beta)$ . For the contribution  $|D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})|$  we use, in the same spirit as for  $|D_{\mathbf{x}_1} H_{l,2}(s, \mathbf{x})|$ , a centering argument and an integration by parts to obtain:

$$\begin{aligned}
 & |D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})| \\
 = & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^\xi \left( \frac{1}{2} \text{Tr} \left[ (a(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) \right. \right. \right. \right. \\
 & \left. \left. \left. \times (D_1^2 u(s, \cdot) - D_1^2 u(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))) \right] \right) \right] (\mathbf{x}) \right|_{\boldsymbol{\xi}=\mathbf{x}} \\
 = & \left| \frac{1}{2} D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left( \sum_{j=1}^d \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\langle \partial_{\mathbf{y}_1^j} \left( \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) (a_{j \cdot}(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - a_{j \cdot}(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) \right) \right. \right. \right. \\
 & \left. \left. \left. (D_1 u(s, \mathbf{y}) - D_1 u(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))) \right\rangle \right) \right|_{\boldsymbol{\xi}=\mathbf{x}},
 \end{aligned}$$

where  $\mathbf{y}_1 = (\mathbf{y}_1^1, \dots, \mathbf{y}_1^d)$ ,  $\partial_{\mathbf{y}_1^j}^j$  denotes the derivative w.r.t. to the  $j^{\text{th}}$  scalar entry of the non-degenerate variable  $\mathbf{y}_1$  and  $a_{j \cdot}$  denotes the  $j^{\text{th}}$  row of the diffusion matrix  $a$ . We therefore derive from Lemma 1 and the smoothness of  $a$  (by the Rademacher theorem,  $a$  is differentiable almost everywhere):

$$\begin{aligned}
 |D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})| & \leq C \| \mathbf{D} D_1 u(s, \cdot) \|_\infty \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(s-t)^l} \hat{p}_{C-1}^\mathbf{x}(t, s, \mathbf{x}, \mathbf{y}) \\
 & \quad \times \left( \frac{|(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{1:l-1}|}{(s-t)^{\frac{1}{2}}} + 1 \right) [a(s, \cdot)]_1 |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{l:n}| \\
 & \leq C \| \mathbf{D} D_1 u(s, \cdot) \|_\infty (s-t)^{-\frac{1}{2}}.
 \end{aligned} \tag{2.28}$$

With the notations of (2.21), plugging (2.27), (2.28) into (2.26) and together with (2.25), (2.23), we eventually derive that there exists  $\delta := \delta((\mathbf{A})) > 0$ :

$$\left| \int_t^T ds D_{\mathbf{x}_1} H_l(s, \mathbf{x}) \right| \leq CT^\delta (\|D_1 u\|_\infty + \|\mathbf{D} D_1 u\|_\infty). \tag{2.29}$$

**Degenerate part of the operator: estimates for  $D_{\mathbf{x}_1} I_l(s, \mathbf{x})$ .** These are the most delicate terms to handle. Restarting from (2.21), we first write for all  $l \in \llbracket 2, n \rrbracket$ :

$$\begin{aligned}
 & |D_{\mathbf{x}_1} I_l(s, \mathbf{x})| \\
 = & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^\xi \left( \sum_{i=2}^n \left\langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}), \right. \right. \right. \right. \\
 & \left. \left. \left. D_i u(s, \cdot) \right\rangle \right) \right] (\mathbf{x}) \right|_{\boldsymbol{\xi}=\mathbf{x}},
 \end{aligned}$$

which readily yields that

$$\begin{aligned}
 & |D_{\mathbf{x}_1} I_l(s, \mathbf{x})| \\
 \leq & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^\xi \left( \sum_{i=l+1}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}), \right. \right. \right. \\
 & \left. \left. \left. D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & + \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^\xi \left( \sum_{i=2}^l \langle \mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})), \right. \right. \right. \\
 & \left. \left. \left. D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 & + \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^\xi \left( \sum_{i=2}^l \langle (\mathbf{F}_i(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \right. \right. \\
 & \left. \left. \left. - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
 =: & |D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})| + |D_{\mathbf{x}_1} I_{2,l}(s, \mathbf{x})| + |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})|. \tag{2.30}
 \end{aligned}$$

We emphasize that the integrands  $D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})$  and  $D_{\mathbf{x}_1} I_{2,l}(s, \mathbf{x})$  are already designed to smoothen the time singularities generated by the cross differentiation of the inhomogeneous semi-group w.r.t. the variables  $\mathbf{x}_l$  and  $\mathbf{x}_1$ . Indeed, on the one hand

$$\begin{aligned}
 & |D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})| \\
 = & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^\xi \left( \sum_{i=l+1}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}), \right. \right. \right. \\
 & \left. \left. \left. D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}}.
 \end{aligned}$$

Hence, Lemmas 1 and 2 yield:

$$\begin{aligned}
 & |D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})| \tag{2.31} \\
 \leq & C \|D_{l+1:n} u\|_\infty \sum_{j=l+1}^n \{ [(D_{\mathbf{x}_{j-1}} \mathbf{F}_j)_{j-1}]_\eta (s-t)^{-l+(1+\eta)(j-3/2)} + [(\mathbf{F}_{l+1:n})_j]_{\beta_j} (s-t)^{-l+\beta_j(j-1/2)} \},
 \end{aligned}$$

where  $D_{l+1:n} u := (D_{l+1} u, \dots, D_n u)$  using as well the notation of (2.22) for the last inequality. Observe that the most singular terms in the previous bounds are those associated with the exponents  $-l + \beta_j(j - 1/2)$ . It turns out that they are indeed integrable under assumption  $(\mathbf{T}_\beta)$  giving again precisely the appropriate thresholds.

On the other hand, Lemma 1 also yields

$$\begin{aligned}
 |D_{\mathbf{x}_1} I_{2,l}(s, \mathbf{x})| &= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \sum_{i=2}^l \langle \mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \\
 &\leq C \|D_{i:l-1} u\|_{\infty} \sum_{j=l}^n [(\mathbf{F}_{1:l})_j(s, \cdot)]_{\beta_j} (s-t)^{-l+\beta_j(j-1/2)}, \tag{2.32}
 \end{aligned}$$

and those terms are again integrable as soon as the thresholds of  $(\mathbf{T}_{\beta})$  hold.

It hence remains to control the terms in  $D_{\mathbf{x}_1} I_{3,l}$ . These terms are the tricky ones since they are, *a priori*, not designed to smoothen the time singularities generated by the cross differentiation. Observe indeed that, if one tries to reproduce the above calculations, we obtain from Lemma 1 that

$$\begin{aligned}
 &|D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
 &= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \sum_{i=2}^l \langle (\mathbf{F}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\xi)) \right. \right. \right. \\
 &\quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}, D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \\
 &\leq C \|D_{2:l} u\|_{\infty} (s-t)^{-l} \\
 &\quad \times \left( \sum_{j=2}^l [(\mathbf{F}_{2:j})_j(s, \cdot)]_{\beta_j} (s-t)^{(j-\frac{1}{2})\beta_j} + [(D_{j-1} \mathbf{F}_j)_{j-1}(s, \cdot)]_{\eta} (s-t)^{(1+\eta)(j-\frac{3}{2})} \right), \\
 &\leq C \|D_{2:l} u\|_{\infty} (s-t)^{-l} \left( (s-t)^{\frac{3}{2}\beta_2} + (s-t)^{\frac{1}{2}(1+\eta)} \right),
 \end{aligned}$$

up to a modification of  $C$  and for  $T$  small enough. This leads, as soon as  $l \geq 2$ , to a time singularity which is not integrable. Indeed,  $\frac{1}{2}(1+\eta) < 1$  (recall that  $\eta$  is meant to be small). To overcome this problem, the idea consists in writing, thanks to Lemma 1,

$$\begin{aligned}
 &|D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
 &= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \sum_{i=2}^l \langle (\mathbf{F}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}, \right. \right. \right. \\
 &\quad \left. \left. \left. (D_i u(s, \cdot) - D_i u(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) \rangle \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \tag{2.33}
 \end{aligned}$$

and to take advantage of the additional smoothing effect from the solution of the regularized PDE itself through the above contribution  $D_i u(s, \cdot) - D_i u(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)$ ,  $i \in \llbracket 2, l \rrbracket$ .



To do so we write, by expanding the gradients in (2.33) with the Taylor formula,

$$\begin{aligned}
 & |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
 = & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \sum_{i=2}^l \left\langle \left( \mathbf{F}_i(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) \right. \right. \right. \right. \\
 & \quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\xi))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1} \right. \right. \right. \\
 & \quad \left. \left. \left. \sum_{k=l}^n \int_0^1 d\lambda D_k D_i u(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi) + \lambda(\cdot_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\xi)))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_k \right\rangle \right) \right] (\mathbf{x}) \right| \\
 = & \left| \sum_{i=2}^l \sum_{k=l}^n \int_0^1 d\lambda D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[ \tilde{P}_{s,t}^{\xi} \left( \left\langle \left( \mathbf{F}_i(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) \right. \right. \right. \right. \right. \\
 & \quad \left. \left. \left. - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1} \right. \right. \right. \\
 & \quad \left. \left. \left. \left. D_i D_k u(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi) + \lambda(\cdot_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\xi)))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_k \right\rangle \right) \right] (\mathbf{x}) \right|, \tag{2.34}
 \end{aligned}$$

for any  $l \in \llbracket 2, n \rrbracket$ . The above expansion hence allows to obtain the additional contribution  $(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_k$  which, thanks to Lemma 1, is precisely designed to smoothen the time singularity coming from the cross differentiation over the variables  $\mathbf{x}_l$  and  $\mathbf{x}_1$  of the semi-group (notice indeed that  $k \geq l$ ). Namely, introducing now for all  $i \in \llbracket 2, l \rrbracket, k \in \llbracket l, n \rrbracket, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \in \mathbb{R}^{(n-1)d}, (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ :

$$\begin{aligned}
 & \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} : \mathbf{y}_i \in \mathbb{R}^d \mapsto \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) = \\
 & \left[ D_{\mathbf{x}_l} D_{\mathbf{x}_1} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) \otimes \left( \left( \mathbf{F}_i(\mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\xi)) \right. \right. \right. \\
 & \quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\xi))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\xi))_{i-1} \right) ((\cdot - \boldsymbol{\theta}_{s,t}(\xi))_k)^* \right], \tag{2.35}
 \end{aligned}$$

where the subscript  $(1, l)$  in  $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$  is here to indicate the differentiation w.r.t.  $D_{\mathbf{x}_l} D_{\mathbf{x}_1}$  acting on the frozen density (we will also use for the analysis the notation  $\Psi_{i,l,k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$  which indicates that the sole derivation w.r.t.  $\mathbf{x}_l$  has to be taken into consideration, see Section 3.2 below). Pay attention that the above function is  $(\mathbb{R}^d)^{\otimes 4}$ -valued.

We can expect a time smoothing effect for  $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$  of order  $(s-t)^{-\frac{1}{2} + (i-\frac{1}{2})\beta_i} + (s-t)^{-\frac{1}{2} + ((i-1)-\frac{1}{2})(1+\eta)}$ , where the exponent  $-\frac{1}{2}$  derives from the derivation  $D_{\mathbf{x}_1}$ . This fact can be easily deduced from Lemma 1 and the previous computations when  $\xi = \mathbf{x}$ . With these notations at hand, it thus remains to control

$$\begin{aligned}
 & |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
 \leq & \sum_{i=2}^l \sum_{k=l}^n \left| \int_0^1 d\lambda \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \right. \\
 & \left. \times \int_{\mathbb{R}^d} d\mathbf{y}_i \left\{ \left( \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) : D_{\mathbf{y}_i} D_{\mathbf{y}_k} u(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi) + \lambda(\mathbf{y}_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\xi))) \right) \right\} \right|, \tag{2.36}
 \end{aligned}$$

where “ : ” stands for the double tensor contraction, in terms of known parameters in **(A)**. Unfortunately, as suggested by the computations in (2.23), the control in supremum norm of  $(D_i D_k u)_{1 \leq i \leq k \leq n}$  is definitely out of reach. Recall indeed that in this equation (which corresponds to the case  $i = 1$ ) the thresholds are sharp in order to retrieve an integrable singularity (see (2.24)).

Roughly speaking, what we can expect is that for fixed  $(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \in \mathbb{R}^{(n-1)d}$ , the partial application  $\mathbf{y}_i \mapsto D_k u^i(s, \mathbf{y}_i) := D_{\mathbf{y}_k} u(s, \mathbf{y}), \mathbf{y} := (\mathbf{y}_{1:i-1}, \mathbf{y}_i, \mathbf{y}_{i+1:n})$  is as smooth as the Green kernel involving the source  $f$  in the indicated equation, i.e.  $\tilde{G}^\xi f(s, \mathbf{y}) := \int_s^T dv \tilde{P}_{v,s}^\xi f(v, \mathbf{y})$ . The specific Hölder modulus  $\alpha_i^k$  of  $\mathbf{y}_i \mapsto D_{\mathbf{y}_k} \tilde{G}^{i,\xi} f(s, \mathbf{y}_i) := D_{\mathbf{y}_k} \tilde{G}^\xi f(s, \mathbf{y})$  can be derived from Lemma 1 and we obtain that any  $\alpha_i^k$  satisfying  $\alpha_i^k < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}}$  is attainable. Our guess is hence that the same property should hold for  $\mathbf{y}_i \mapsto D_{\mathbf{y}_k} u^i(s, \mathbf{y}_i)$ .

Keeping in mind this objective, the next step consists in exploiting a Besov duality result on the variable  $\mathbf{y}_i$  recalling that  $C_b^{\alpha_i^k}(\mathbb{R}^d, \mathbb{R}) = B_{\infty, \infty}^{\alpha_i^k}(\mathbb{R}^d, \mathbb{R})$ , where from now on the notation  $B_{p,q}^s$  stands for a Besov space with associated indexes  $p, q, s$  (see Triebel [Tri83] and Appendix 5 below). The indexes  $p, q$  denote the integrability parameters and  $s$  the smoothness one. However, instead of focusing on the Hölder modulus, we will concentrate on  $\|D_i D_k u^i(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}$  (space of the derivatives of Hölder functions). This specific choice allows to naturally deal with the cancellation techniques which require to change the freezing points depending on the current regime of the underlying frozen heat-kernel (see the proof of Lemma 4 in Section 3.2 below). We also carefully mention that the quantity  $\|D_i D_k u^i(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}$  is equivalent to  $\|D_k u^i(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k}}$  (see Proposition 4).

A classical fact is that  $B_{\infty, \infty}^{\alpha_i^k - 1}$  and  $B_{1,1}^{1-\alpha_i^k}$  can be put in duality, see e.g. Proposition 3.6 in Lemarié-Rieusset ([LR02]). Indeed, with the notations therein  $B_{\infty, \infty}^{\alpha_i^k - 1}$  is the dual of  $B_{1,1}^{1-\alpha_i^k} = \tilde{B}_{1,1}^{1-\alpha_i^k}$  where  $\tilde{B}_{1,1}^{1-\alpha_i^k}$  denotes the closure of the Schwartz class  $\mathcal{S}$  in  $B_{1,1}^{1-\alpha_i^k}$  (see also Theorem 4.1.3 in Adams Hedberg [AH96] for the density of  $\mathcal{S}$  in  $B_{1,1}^{1-\alpha_i^k}$ ). Exploiting this fact, we then derive from (2.36) and the multi-linearity of the tensors involved that:

$$\begin{aligned}
 & |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
 & \leq C \sum_{i=2}^l \sum_{k=l}^n \int_0^1 d\lambda \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \\
 & \quad \times \left\{ \left\| \left[ \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \left\| \mathbf{y}_i \mapsto D_i D_k u(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}) + \lambda(\mathbf{y}_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right\} \\
 & \leq C \sum_{i=2}^l \sum_{k=l}^n \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \left\{ \left\| \left[ \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \right. \\
 & \quad \times \left[ \sup_{\mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq i} \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \mathbb{I}_{i \leq l-1} \right. \\
 & \quad \left. \left. + \sup_{\mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq i, \lambda \in [0,1]} \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \mathbb{I}_{i=l} \right] \right\}. \tag{2.37}
 \end{aligned}$$

Observe now that, for any  $\lambda \in [0, 1]$  and the norm equivalence of Proposition 4, we have:

$$\begin{aligned} & \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k - 1}} \\ & \leq C \left\| D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k}} \\ & \leq C \left\| D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k}}, \end{aligned}$$

exploiting the scaling and shift invariance properties of the Hölder modulus. Using again the norm equivalence of Proposition 4, we eventually derive:

$$\begin{aligned} & \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k - 1}} \\ & \leq C^2 \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k - 1}}. \end{aligned} \quad (2.38)$$

In other words, the dilations-translations of the variable  $\mathbf{y}_i$ , that appear when  $i = l$  in (2.37), do not affect the regularity estimate. We therefore obtain:

$$\begin{aligned} & |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\ & \leq C \sum_{i=2}^l \sum_{k=l}^n \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \left\{ \left\| \left[ \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \right. \\ & \quad \left. \times \sup_{\mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq i} \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k - 1}} \right\}. \end{aligned} \quad (2.39)$$

To conclude this proof we now need the following results whose proofs are postponed to the next subsection:

**Lemma 3** *Let  $l \in \llbracket 2, n \rrbracket$ ,  $i \in \llbracket 2, l \rrbracket$  and  $k \in \llbracket l, n \rrbracket$  and let  $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be the function defined by (2.35). There exist  $C := C(\mathbf{A}) > 0$ ,  $\alpha_i^k := \frac{1+\eta}{2i-1} < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}}$ ,  $\gamma_i^k = \gamma_i := \frac{1}{2} + \eta(i - \frac{3}{2}) > \frac{1}{2}$  such that*

$$\left\| \left[ \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \leq C(s-t)^{-\frac{3}{2} + \gamma_i^k} \hat{q}_{c,i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})),$$

where, with the notations of Proposition 2,

$$\hat{q}_{c,i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) := \int_{\mathbb{R}^d} \hat{p}_c^{\mathbf{x}}(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y}_i = \prod_{j \in \llbracket 1, n \rrbracket, j \neq i}^n \mathcal{N}_{c(s-t)^{2j-1}}((\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j), \quad (2.40)$$

denoting as well for  $a > 0$ ,  $z \in \mathbb{R}^d$ , by  $\mathcal{N}_a(z) = \frac{1}{(2\pi a)^{\frac{d}{2}}} \exp\left(-\frac{|z|^2}{2a}\right)$  the standard Gaussian density of  $\mathbb{R}^d$  with covariance matrix  $a\mathbf{I}_{d,d}$ . In the above control  $c = C^{-1}$ .

**Lemma 4** *Let  $u$  be the solution of (2.1).*

*There exists  $C := C((\mathbf{A})) > 0$  such that for all  $i \leq k \in \llbracket 2, n \rrbracket^2$  and  $\alpha_i^k := \frac{1+\frac{\eta}{4}}{2i-1} < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}}$ ,*

$$\sup_{\mathbf{y}_j, j \in \llbracket 1, n \rrbracket, j \neq i} \left\| D_i D_k u(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k-1}} \leq C(\|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty}). \quad (2.41)$$

**Remark 2 (About Lemmas 3 and 4)** *There are some specific points to be emphasized about the indicated lemmas: We point out that the parameters  $\gamma_i^k$ ,  $\alpha_i^k$  actually do not depend on the index  $k$ . We keep the notations to remember that they are associated with the Hölder regularity of  $D_{\mathbf{x}_k} u$  w.r.t. to its  $i^{\text{th}}$  variable which we choose to investigate in the corresponding negative Besov space for technical reasons. We also emphasize that the norm  $\left\| D_{\mathbf{y}_i} D_{\mathbf{y}_k} u(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k-1}}$  is actually equivalent to  $\left\| D_{\mathbf{y}_k} u(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k}}$  which corresponds to the usual Hölder norm on  $C_b^{\alpha_i^k}$  (see Proposition 4 below). This can be seen again as a specific feature appearing when dealing with strong uniqueness through the Zvonkin approach. The regularity is associated with a variable independently of the level in the chain (1.1).*

We can then deduce, using Lemma 3 and Lemma 4 that

$$|D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \leq C \sum_{i=2}^l \sum_{k=l}^n (s-t)^{-\frac{3}{2}+\gamma_i^k} \left( \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty} \right), \quad (2.42)$$

which are integrable terms since  $\gamma_i^k > \frac{1}{2}$ . With the notations of (2.21), (2.30), we eventually derive from (2.42), (2.32), (2.31) that there exists  $\gamma := \gamma((\mathbf{A})) > 0$  such that:

$$\left| \int_t^T ds D_{\mathbf{x}_1} I_l(s, \mathbf{x}) \right| \leq CT^{\gamma} \left( \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty} \right). \quad (2.43)$$

**Conclusion.** Bringing together (2.43) and (2.29) yields for all  $l \in \llbracket 1, n \rrbracket$  and all  $(t, \mathbf{x}) \in [0, T]$ :

$$|D_{\mathbf{x}_l} D_{\mathbf{x}_1} u(t, \mathbf{x})| \leq C(T^{\gamma} + T^{\delta}) \left( \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty} \right). \quad (2.44)$$

It is clear that the previous analysis can be reproduced without differentiating w.r.t.  $\mathbf{x}_1$ , leading to improved singularity exponents (see also the proof of Lemma 4 which somehow exactly explicit these computations). We therefore get:

$$|D_{\mathbf{x}_l} u(t, \mathbf{x})| \leq C(T^{\gamma'} + T^{\delta'}) \left( \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty} \right), \quad (2.45)$$

for some positive exponents  $\gamma', \delta'$  (with  $\gamma' > \gamma$ ,  $\delta' > \delta$ ).

Taking the time-space supremum in the l.h.s of (2.44) and (2.45), recalling as well that  $T$  is meant to be small, i.e. s.t.  $4CT^{\delta \wedge \gamma} \leq \frac{1}{2}$ , we derive:

$$\|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty} \leq 2C(T^{\gamma} + T^{\delta}).$$

This concludes the proof. □

### 3 Estimates in Besov norm

This section is dedicated to the proofs of the main technical results needed to obtain Theorem 3. Namely, we prove the Besov estimates of Lemmas 3 and 4.

#### 3.1 Proof of Lemma 3

We will here exploit the thermic characterization of Besov spaces (see Chapter 2.6.4 in [Tri83]) which is also recalled in Appendix 5.

From (5.1), we are thus led to estimate, for any  $l \in \llbracket 2, n \rrbracket$ ,  $i \in \llbracket 2, l-1 \rrbracket$  and  $k \in \llbracket l, n \rrbracket$ :

$$\int_0^1 \frac{dv}{v} v^{1+\frac{\alpha_i^k-1}{2}} \|\partial_v h_v \star \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})}.$$

We split the time integral in the above equation into two parts writing:

$$\begin{aligned} & \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_i^k-1}{2}} \|\partial_v h_v \star \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & + \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k-1}{2}} \|\partial_v h_v \star \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & =: \mathbf{Lower} \left[ \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] + \mathbf{Upper} \left[ \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right], \end{aligned} \quad (3.1)$$

for a parameter  $\rho_{i,k} > 0$  to be specified. The term **Upper** corresponding to the upper-part of the integral w.r.t.  $v$  does not involve singularities. We will use this fact to calibrate the associated parameter  $\rho_{i,k}$  in order to match the integrability constraint

$$\mathbf{Upper} \left[ \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] \leq \frac{C}{(s-t)^{1+\frac{1}{2}-\gamma_i^k}} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})),$$

where  $\hat{q}_{c \setminus i}$  has been defined in (2.40) and  $\gamma_i^k > 1/2$  in order to obtain a time integrable singularity. For this term, we will only use crude upper-bounds on the derivatives of the heat-kernel and the coefficients satisfying  $(\mathbf{T}_\beta)$ . On the other hand, the contribution **Lower** in (3.1) precisely contains the singularities w.r.t.  $v$ . It is therefore crucial to use there suitable cancellation tools. The point will then be to prove that the associated estimates are compatible with the upper-bound in equation (3.2).

We now write:

$$\begin{aligned} & \mathbf{Upper} \left[ \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] \\ & = \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k-1}{2}} \|\partial_v h_v \star \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & = \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k-1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\mathbf{y}_i) d\mathbf{y}_i \right|. \end{aligned}$$

Recall from the definition in (2.35) that  $\Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\mathbf{y}_i)$  is  $(\mathbb{R}^d)^{\otimes 4}$ -valued. To proceed with the computations we assume w.l.o.g. for the rest of the proof that  $d = 1$  to avoid

tensor notations for simplicity. Writing explicitly the function  $\Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\mathbf{y}_i)$  leads to:

$$\begin{aligned} & \text{Upper} \left[ \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] \\ & \leq \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k-1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \left( D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \left[ ((\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}) - \mathbf{y})_k) \right. \right. \right. \\ & \quad \left. \left. \left. \left( \mathbf{F}_i(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{i-1}} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right] \right) \right|_{\boldsymbol{\xi}=\mathbf{x}} d\mathbf{y}_i \Big|_{\boldsymbol{\xi}=\mathbf{x}}. \end{aligned}$$

From Lemma 2 and Proposition 2, we derive there exists a  $C := C((\mathbf{A}), T) > 0$  such that introducing  $\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}) = \hat{p}_c^\mathbf{x}(t, s, \mathbf{x}, \mathbf{y})$ ,  $c = C^{-1}$ :

$$\begin{aligned} & \text{Upper} \left[ \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] \\ & \leq C \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k-1}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{(l-\frac{1}{2})+\frac{1}{2}}} |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_k| \\ & \quad \times \left\{ \sum_{j=i}^{l-1} \left\{ |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j|^{\beta_j} \right\} + |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_{i-1}|^{1+\eta} \right\} \\ & \leq C \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k-1}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\frac{1}{2}}} \\ & \quad \times \left\{ \sum_{j=i}^{l-1} (s-t)^{\beta_j(j-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right\} \\ & \leq C \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{-\frac{3}{2}+\frac{\alpha_i^k}{2}} (s-t)^{-1/2} \\ & \quad \times \left\{ \sum_{j=i}^{l-1} (s-t)^{\beta_j(j-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right\} \\ & \leq C \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) (s-t)^{[-\frac{1}{2}+\frac{\alpha_i^k}{2}]\rho_{i,k}-\frac{1}{2}} \left( (s-t)^{\beta_i(i-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right), \end{aligned}$$

recalling that the lower bound of  $\beta_j(j-1/2)$  is increasing for the last inequality (recall indeed that we assumed that  $\beta_j \in \left(\frac{2j-2}{2j-1}, 1\right]$ ). We now want to choose the threshold  $\rho_{i,k}$  in order to match the integrability condition in (3.2). This amounts to write:

$$-1 - \frac{1}{2} + \gamma_i^k = \left[-\frac{1}{2} + \frac{\alpha_i^k}{2}\right]\rho_{i,k} - \frac{1}{2} + \left(\beta_i(i-1/2) \wedge (1+\eta)(i-3/2)\right).$$

Since this condition should hold for any  $\beta_i \in \left(\frac{2i-2}{2i-1}, 1\right]$  and since the parameter  $\eta \in (0, 1)$  is small (see  $(\mathbf{H}_\eta)$ ) we have  $\beta_i(i-1/2) \wedge (1+\eta)(i-3/2) = (1+\eta)(i-3/2)$ . The above condition rewrites:

$$-1 - \frac{1}{2} + \gamma_i^k = \left[-\frac{1}{2} + \frac{\alpha_i^k}{2}\right]\rho_{i,k} - \frac{1}{2} + (1+\eta)(i-3/2).$$

Our global integrability constraint associated with the  $i^{\text{th}}$  variable in the  $k^{\text{th}}$  derivative writes:

$$-1 - \frac{1}{2} + \gamma_i^k = \left[-\frac{1}{2} + \frac{\alpha_i^k}{2}\right] \rho_{i,k} - \frac{1}{2} + (1 + \eta)(i - 3/2) > -1, \quad (3.2)$$

which gives

$$\rho_{i,k} < \frac{(1 + \eta)(2i - 3) + 1}{1 - \alpha_i^k}. \quad (3.3)$$

It therefore remains to check that such a choice is compatible with the time integral part for  $v \in [0, (s - t)^{\rho_{i,k}}]$  in the thermic characterisation of the Besov norm. We point out that for this term it is absolutely essential to get rid of the exponent  $v^{-1}$  coming from the upper-bound of the thermic heat-kernel, i.e.  $\partial_v h_v(z - \mathbf{y}_i)$ . In order to get an integrable singularity in  $v$ , we need to decrease the crude upper-bound on  $\partial_v h_v(z - \mathbf{y}_i)$ . This is done through cancellation techniques exploiting the smoothness properties of  $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$ .

To investigate **Lower**  $\left[ \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right]$  let us first recall from the definition in (2.35) that for all  $k \in \llbracket l, n \rrbracket$ :

$$\begin{aligned} & \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) \\ &= D_{\mathbf{x}_l} D_{\mathbf{x}_1} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) \left[ (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \left( \mathbf{F}_i(\mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \right. \\ & \quad \left. \left. - D_{i-1} \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right]. \end{aligned} \quad (3.4)$$

Let us now specify the dependence w.r.t.  $\mathbf{y}_i$  of the previous expression in function of the considered indexes  $i \in \llbracket 1, l \rrbracket, l \in \llbracket 2, n \rrbracket, k \in \llbracket l, n \rrbracket$ . This will be useful to develop corresponding adapted cancellation arguments.

Observe first that the dependence in  $\mathbf{y}_i$  appears in (3.4) for any  $i \in \llbracket 1, l \rrbracket, l \in \llbracket 2, n \rrbracket, k \in \llbracket l, n \rrbracket$  through the term  $D_{\mathbf{x}_1 \mathbf{x}_2} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y})$ .

For the term into brackets we distinguish two cases. If  $i \leq l - 1$ , and then  $k > i$ , the only bracket term containing  $\mathbf{y}_i$  is the one associated with  $\mathbf{F}_i$ . If now  $i = l \leq k$ , for the contribution  $i = l = k$ , the only term into brackets in (3.4) that will also depend on  $\mathbf{y}_i$  is  $(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_i$  and the cancellation arguments need to be slightly modified. Eventually, when  $i = l < k$ , there will be no dependence on  $\mathbf{y}_i$  for the terms into brackets.

- First case  $i < l \leq k$  ( $i \in \llbracket 1, l - 1 \rrbracket, l \in \llbracket 2, n \rrbracket, k \in \llbracket l, n \rrbracket$ ).

With the notations of (3.1), we write:

$$\begin{aligned} & \text{Lower} \left[ \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] \\ &= \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_{cv}(z - \mathbf{y}_i) \right. \\ & \quad \left. (\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) - \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(z)) \right| \\ &=: \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz \left| (\mathcal{F}_{1,i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} + \mathcal{F}_{2,i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})})(v, z) \right|, \end{aligned} \quad (3.5)$$

where:

$$\begin{aligned} & \mathcal{F}_{1,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z) \\ & := \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \left( D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \right. \\ & \quad \left. \left[ (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \left( \mathbf{F}_i(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) \right) \right] \right] d\mathbf{y}_i, \end{aligned} \quad (3.6)$$

with a slight abuse of notation when  $i = l - 1$  and

$$\begin{aligned} & \mathcal{F}_{2,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z) \\ & := \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \left[ D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right] \left[ (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \right. \\ & \quad \left. \left( \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right] d\mathbf{y}_i. \end{aligned} \quad (3.7)$$

Write now from (3.7), Proposition 2 and Lemma 2 (recalling as well that we took  $\boldsymbol{\xi} = \mathbf{x}$  for the current analysis):

$$\begin{aligned} & \left| \mathcal{F}_{2,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z) \right| \\ & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v} \int_0^1 d\lambda \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+(l-\frac{1}{2})+\frac{1}{2}}} |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| |\mathbf{y}_i - z| \\ & \quad \times \left( \left| \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\mathbf{x})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\mathbf{x})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\mathbf{x})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1} \right| \right) \\ & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+\frac{1}{2}}} \\ & \quad \times \left\{ |z - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|^{\beta_i} + \sum_{j=i+1}^{l-1} (s-t)^{\beta_j(j-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right\}, \end{aligned}$$

where for the second inequality, we used that for  $k \geq l > i$ ,  $|(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| (s-t)^{-(l-1/2)} \leq |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| (s-t)^{-(k-1/2)}$  which can be absorbed by the  $k^{\text{th}}$  variables of  $\hat{q}_c$ .

Writing now for any  $\lambda \in [0, 1]$ ,

$$|z - \boldsymbol{\theta}_{s,t}(\mathbf{x})_i| \leq \lambda |z - \mathbf{y}_i| + |z + \lambda(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|,$$

we thus derive

$$\begin{aligned} & \left| \mathcal{F}_{2,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z) \right| \\ & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \\ & \quad \times \left( v^{\frac{\beta_i}{2}} (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}} + (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+\beta_i(i-1/2)} \right. \\ & \quad \left. + \sum_{j=i+1}^{l-1} (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+\beta_j(j-1/2)} + (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+(1+\eta)(i-3/2)} \right). \\ & \leq C \hat{q}_{c \setminus i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \int_0^1 d\lambda \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \mathcal{N}_{c(s-t)^{2i-1}}(z + \lambda(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) \\ & \quad \times v^{-\frac{1}{2}} \left( v^{\frac{\beta_i}{2}} (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}} + (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+(1+\eta)(i-3/2)} \right), \end{aligned} \quad (3.8)$$



recalling for the last inequality that for all  $j$  in  $\llbracket i, l-1 \rrbracket$ ,  $\beta_j(j - \frac{1}{2}) > j - 1 > i - \frac{3}{2}$  and  $\eta$  is supposed to be a small parameter.

Consider now  $\mathcal{F}_1$ ,

$$\begin{aligned} |\mathcal{F}_{1,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v,z)| &\leq C \int_{\mathbb{R}^d} \frac{h_{cv}(z - \mathbf{y}_i)}{v} \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{(l-\frac{1}{2})+\frac{1}{2}}} |z - \mathbf{y}_i|^{\beta_i} |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| d\mathbf{y}_i \\ &\leq C \int_{\mathbb{R}^d} \frac{h_{cv}(z - \mathbf{y}_i)}{v^{1-\frac{\beta_i}{2}}} \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\frac{1}{2}}} d\mathbf{y}_i. \end{aligned} \quad (3.9)$$

From (3.8) and (3.9) we derive, with the notation introduced in (2.40):

$$\begin{aligned} &\|\partial_v h_v \star \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ &\leq C \hat{q}_{c \setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \left\{ \frac{1}{v^{1-\frac{\beta_i}{2}}(s-t)^{\frac{1}{2}}} + \left( \frac{v^{-\frac{1+\beta_i}{2}}}{(s-t)^i} + \frac{v^{-\frac{1}{2}}}{(s-t)^{i-(1+\eta)(i-\frac{3}{2})}} \right) \right. \\ &\quad \times \int_0^1 d\lambda \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \mathcal{N}_{c(s-t)^{2i-1}}(z + \lambda(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) \left. \right\} \\ &\leq C \hat{q}_{c \setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \left( \frac{1}{v^{1-\frac{\beta_i}{2}}(s-t)^{\frac{1}{2}}} + \frac{v^{-\frac{1+\beta_i}{2}}}{(s-t)^i} + \frac{v^{-\frac{1}{2}}}{(s-t)^{i-(1+\eta)(i-\frac{3}{2})}} \right), \end{aligned}$$

using the change of variable  $(w_1, w_2) = (z - \mathbf{y}_i, z + \lambda(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)$  for the last inequality.

For  $\rho_{i,k}$  chosen as in (3.3) one gets from the definition in (3.1):

$$\begin{aligned} &\text{Lower} \left[ \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] \\ &\leq C \hat{q}_{c \setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_0^{(s-t)^{\rho_{i,k}}} \frac{dv}{v} v^{\frac{\alpha_i^k}{2}} \\ &\quad \times \left( \frac{1}{v^{\frac{1-\beta_i}{2}}(s-t)^{\frac{1}{2}}} + \frac{v^{\frac{\beta_i}{2}}}{(s-t)^i} + \frac{1}{(s-t)^{i-(1+\eta)(i-\frac{3}{2})}} \right) \\ &=: C \hat{q}_{c \setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \mathcal{B}_{1,\rho_{i,k}}(t,s). \end{aligned}$$

It therefore remains to prove that, if  $\mathcal{B}_{1,\rho_{i,k}}(t,s) \leq (s-t)^{-1-\frac{1}{2}+\gamma_i^k}$  for  $\gamma_i^k > 0$  then:

$$-1 - \frac{1}{2} + \gamma_i^k > -1 \iff \gamma_i^k > \frac{1}{2}. \quad (3.10)$$

Write:

$$\begin{aligned} \mathcal{B}_{1,\rho_{i,k}}(t,s) &\leq \int_0^{(s-t)^{\rho_{i,k}}} dv \left( \frac{v^{-\frac{3}{2}+\frac{\alpha_i^k+\beta_i}{2}}}{(s-t)^{\frac{1}{2}}} + \frac{v^{-1+\frac{\alpha_i^k+\beta_i}{2}}}{(s-t)^i} + \frac{v^{-1+\frac{\alpha_i^k}{2}}}{(s-t)^{\frac{3}{2}-\eta(i-\frac{3}{2})}} \right) \\ &\leq C \left( (s-t)^{\rho_{i,k} \frac{\alpha_i^k+\beta_i-1}{2}-\frac{1}{2}} + (s-t)^{\rho_{i,k} \frac{\alpha_i^k+\beta_i}{2}-i} + (s-t)^{\rho_{i,k} \frac{\alpha_i^k}{2}-\frac{3}{2}+\eta(i-\frac{3}{2})} \right). \end{aligned} \quad (3.11)$$

Let us check condition (3.10) is satisfied. Actually, the first two terms in (3.11) yield negligible contributions. Recall indeed from the statement of the lemma that the parameter  $\alpha_i^k$  must be chosen so that  $\alpha_i^k < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}} := \bar{\alpha}_i^k$ . Since  $\beta_k \in (\frac{2k-2}{2k-1}, 1]$ ,  $\bar{\alpha}_i^k = \frac{2-(1-\beta_k)(2k-1)}{2i-1} > \frac{3-2k+(2k-2)}{2i-1} = \frac{1}{2i-1}$ . Thus,

$$\alpha_i^k = \frac{1 + \frac{\eta}{4}}{2i - 1}, \quad (3.12)$$

is an admissible choice (recall indeed that we chose  $\eta < \inf_{j \in \llbracket 2, n \rrbracket} \{\beta_j - \frac{2j-2}{2j-1}\}$  in assumption  $(\mathbf{H}_\eta)$ ). It gives in particular that  $\alpha_i^k + \beta_i > \frac{1+\frac{\eta}{4}+(2i-2)}{2i-1} = 1 + \frac{\frac{\eta}{4}}{2i-1}$ . Hence,

$$\rho_{i,k} \frac{\alpha_i^k + \beta_i - 1}{2} - \frac{1}{2} > -\frac{1}{2}$$

which already provides a regularizing term in time for the first term in the r.h.s. of (3.11). Now from (3.3),  $\rho_{i,k} < \frac{(1+\eta)(2i-3)+1}{1-\alpha_i^k} =: \bar{\rho}_{i,k}$ . The previous choice for  $\alpha_i^k$  gives

$$\bar{\rho}_{i,k} = \frac{2i - 2 + \eta(2i - 3)}{\frac{2i-1-(1+\frac{\eta}{4})}{2i-1}} = (2i - 1) \frac{2i - 2 + \eta(2i - 3)}{2i - 2 - \frac{\eta}{4}} > (2i - 1),$$

and

$$\rho_{i,k} = 2i - 1 \quad (3.13)$$

is an admissible choice. We therefore get for the exponent of the second term in (3.11):

$$\rho_{i,k} \frac{\alpha_i^k + \beta_i}{2} - i > \frac{1}{2} \left( \rho_{i,k} \left( 1 + \frac{\eta}{2i-1} \right) - 2i \right) > -\frac{1}{2}.$$

Eventually, for the exponent of the third contribution in (3.11), for the previous choice of  $\rho_{i,k} = 2i - 1$ , we get

$$\rho_{i,k} \frac{\alpha_i^k}{2} - \frac{3}{2} + \eta \left( i - \frac{3}{2} \right) = \frac{1}{2} (2i - 1) \frac{1 + \frac{\eta}{4}}{2i - 1} - \frac{3}{2} + \eta \left( i - \frac{3}{2} \right) = -1 + \frac{\eta}{8} + \eta \left( i - \frac{3}{2} \right) > - \quad (3.14)$$

which means that criterion (3.10) is indeed satisfied, even though if this last contribution is rather critical in order to obtain the required smoothing effect with  $\gamma_i^k = 1/2 + \eta(i-3/2)$ . This concludes the proof of Lemma 3 for the indexes  $i \in \llbracket 1, l-1 \rrbracket$ .

• Second case  $i = l \leq k$ . Let us begin with the case  $i = l < k$  for which the only contribution in  $\mathbf{y}_i$  in (3.4) appears through  $D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}(t, s, \mathbf{x}, \mathbf{y})$ . Hence, for such indexes,

$$\text{Lower} \left[ \Psi_{i,(i,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] = \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz |\mathcal{F}_{2,i,(i,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z)|,$$

where:

$$\begin{aligned} & \mathcal{F}_{2,i,(i,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z) \\ &= \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z - \mathbf{y}_i) \left[ D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right] \left[ (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \right. \\ & \quad \left. \times \left( \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}^{i:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right]. \quad (3.15) \end{aligned}$$

The previous analysis for this term can be reproduced adapting the computations leading to (3.8). Precisely,

$$\begin{aligned}
 & |\mathcal{F}_{2,i,(i,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z)| \\
 & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v} \int_0^1 d\lambda \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+(l-\frac{1}{2})+\frac{1}{2}}} |(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| |\mathbf{y}_i-z| \\
 & \quad \times \left| (\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1} \right|^{1+\eta} \\
 & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+\frac{1}{2}}} (s-t)^{(1+\eta)(i-3/2)}.
 \end{aligned}$$

This contribution has already been analyzed and yields the expected integrable singularity in time.

Let us now focus on the remaining case  $i = l = k$ , for which we write:

$$\begin{aligned}
 & \mathbf{L} \left[ \Psi_{i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})} \right] \\
 & = \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i-1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z-\mathbf{y}_i) \right. \\
 & \quad \left. (\Psi_{i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(\mathbf{y}_i) - \Psi_{i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(z)) \right| \\
 & = \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i-1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z-\mathbf{y}_i) \right. \\
 & \quad \left[ D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t,s,\mathbf{x},\mathbf{y})(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_i - D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n})(z-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_i \right] \\
 & \quad \left. \left( \mathbf{F}_i(s,\mathbf{y}_{1:i-1},\boldsymbol{\theta}_{s,t}^{i:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right| \\
 & =: \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i-1}{2}} \int_{\mathbb{R}^d} dz \left( \mathcal{F}_{2,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})} + \mathcal{F}_{3,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z) \right), \quad (3.16)
 \end{aligned}$$

where  $\mathcal{F}_{2,i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z)$  is obtained from (3.15) taking  $k = i$  and

$$\begin{aligned}
 & \mathcal{F}_{3,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z) \\
 & := \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z-\mathbf{y}_i) \left( D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}) \right. \\
 & \quad \left. \left[ (\mathbf{y}_i-z) \left( \mathbf{F}_i(s,\mathbf{y}_{1:i-1},\boldsymbol{\theta}_{s,t}^{i:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right] \right). \quad (3.17)
 \end{aligned}$$

Write first:

$$\begin{aligned}
 & |\mathcal{F}_{2,i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z)| \\
 & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v} \int_0^1 d\lambda \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_i + \lambda(z - \mathbf{y}_i), \mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+(i-\frac{1}{2})+\frac{1}{2}}} |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_i| |\mathbf{y}_i - z| \\
 & \quad \times \left| (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1} \right|^{1+\eta} \\
 & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_i + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \\
 & \quad \times \left( \frac{|\mathbf{y}_i + \lambda(z - \mathbf{y}_i) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i| + \lambda|z - \mathbf{y}_i|}{(s-t)^{2i-\frac{1}{2}}} \right) (s-t)^{(i-\frac{3}{2})(1+\eta)}.
 \end{aligned}$$

We thus derive:

$$\begin{aligned}
 & |\mathcal{F}_{2,i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z)| \\
 & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_i + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \\
 & \quad \left( \frac{1}{(s-t)^i} + \frac{v^{\frac{1}{2}}}{(s-t)^{2i-\frac{1}{2}}} \right) (s-t)^{(i-\frac{3}{2})(1+\eta)} \\
 & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_i + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n}) \\
 & \quad \left( \frac{1}{(s-t)^{\frac{3}{2}-\eta(i-\frac{3}{2})}} + \frac{v^{\frac{1}{2}}}{(s-t)^{i-\frac{1}{2}+\frac{3}{2}-\eta(i-\frac{3}{2})}} \right).
 \end{aligned}$$

Finally,

$$\begin{aligned}
 & \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_{i,i}^i-1}{2}} \int_{\mathbb{R}^d} dz |(\mathcal{F}_{2,i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}| \\
 & \leq C (s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \int_0^{(s-t)^{\rho_{i,i}}} dv \left( v^{-1+\frac{\alpha_{i,i}^i}{2}} + \frac{v^{-\frac{1}{2}+\frac{\alpha_{i,i}^i}{2}}}{(s-t)^{i-\frac{1}{2}}} \right) \\
 & \leq C (s-t)^{-1+\eta(i-\frac{3}{2})+\frac{\eta}{8}} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \leq C (s-t)^{-\frac{3}{2}+\gamma_i^i} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}),
 \end{aligned}$$

recalling for the last inequality that, since from (3.12) and (3.13),  $\alpha_{i,i}^i = (1 + \eta/4)/(2i - 1)$ ,  $\rho_{i,i} = 2i - 1$ ,  $\alpha_{i,i}^i \rho_{i,i}/2 = 1/2 + \eta/8$ ,  $\rho_{i,i}/2 = i - 1/2$ . We also refer to (3.14) for similar computations.

On the other hand, we readily get, similarly to the previous contributions from Lemmas 1 and 2:

$$\begin{aligned}
 & |\mathcal{F}_{3,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z)| \\
 & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{|(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}|^{1+\eta}}{(s-t)^i} \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \\
 & \leq C (s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})} v^{-\frac{1}{2}} \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z - \mathbf{y}_i) \hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}).
 \end{aligned}$$

Similarly to the contributions appearing in (3.5) this bound needs to be integrated w.r.t.  $v \in [0, (s-t)^{\rho_{i,i}}]$ . One gets:

$$\begin{aligned}
 & \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i-1}{2}} \int_{\mathbb{R}^d} dz |(\mathcal{I}_{3,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})})| \\
 & \leq C(s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \int_0^{(s-t)^{\rho_{i,i}}} dv v^{-1+\frac{\alpha_i}{2}} \\
 & \leq C(s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})+\frac{\alpha_i}{2}\rho_{i,i}} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \\
 & \leq C(s-t)^{-\frac{3}{2}+\gamma_i^i} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}),
 \end{aligned}$$

reproducing the previous computations done on  $\mathcal{I}_{2,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v, z)$  for the last inequality.

### 3.2 Proof of Lemma 4

We now tackle the Besov estimate of the cross derivative of the solution of (1.5). Fix  $t \in [0, T]$  and  $(\mathbf{x}_{1:i-1}, \mathbf{x}_{i+1:n}) \in \mathbb{R}^{(n-1)d}$ . From the thermic characterization of Besov spaces recalled in equation (5.1), we actually have to control:

$$\begin{aligned}
 & \|D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})\|_{B_{\infty, \infty}^{\alpha_i^k-1}} \tag{3.18} \\
 & := \|\varphi(D) D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})\|_{\infty} + \sup_{v \in [0,1]} v^{\frac{1-\alpha_i^k}{2}} \sup_{z \in \mathbb{R}^d} |h_v \star D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)|.
 \end{aligned}$$

For the proof we focus on the second contribution in the above definition. The first one could be handled similarly and more directly <sup>††</sup>.

Using representation (2.20) we first write with the notations of (2.21):

$$\begin{aligned}
 & |h_v \star D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)| \tag{3.19} \\
 & = \left| \int_{\mathbb{R}^d} d\mathbf{x}_i h_v(z - \mathbf{x}_i) D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \mathbf{x}_i, \mathbf{x}_{i+1:n}) \right| \\
 & = \left| \int_{\mathbb{R}^d} d\mathbf{x}_i D_{\mathbf{x}_i} h_v(z - \mathbf{x}_i) \otimes \left( D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \mathbf{x}_i, \mathbf{x}_{i+1:n}) - D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, z, \mathbf{x}_{i+1:n}) \right) \right|,
 \end{aligned}$$

integrating by parts and using a usual cancellation argument in the above equality.

To handle the above expression we are led to consider separately the time integration interval corresponding to the diagonal and the off-diagonal regime for the considered running time. We hence introduce the time set  $\mathcal{S}_i = \{s \in [t, T] : (s-t) \leq c_0 |\mathbf{x}_i - z|^{2/(2i-1)}\}$  for a constant  $c_0$  to be specified later on and meant to be *small*. The time set  $\mathcal{S}_i$  corresponds to the off-diagonal regime (i.e. the distance between the  $i^{\text{th}}$  spatial is larger than the  $i^{\text{th}}$  characteristic time scale).

<sup>††</sup>. We could also observe from Proposition 4 that this term could also be directly bounded from the supremum norm of the gradient.

The complementary set  $\mathcal{S}_i^c = \{s \in [t, T] : (s - t) > c_0 |\mathbf{x}_i - z|^{2/(2i-1)}\}$  corresponds to the diagonal one (i.e. the distance between the  $i^{\text{th}}$  spatial variable is smaller than the  $i^{\text{th}}$  characteristic time).

Then, we need to expand suitably the terms in (3.19) with our perturbative approach. To this end, we proceed as in Chapter 6. Namely, we will expand the term  $D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \mathbf{x}_i, \mathbf{x}_{i+1:n})$  with (2.21) taking  $\boldsymbol{\xi} = \mathbf{x}$ , whereas we will expand differently the contribution  $D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, z, \mathbf{x}_{i+1:n})$ , depending on the considered regime (off-diagonal or diagonal) for the current running time. Setting  $\mathbf{x}' = (\mathbf{x}_{1:i-1}, z, \mathbf{x}_{i+1:n})$ , this term will indeed be expand as in (2.21) around the freezing point  $\boldsymbol{\xi}' = \mathbf{x}'$  in the off-diagonal regime and around  $\tilde{\boldsymbol{\xi}}' = \mathbf{x}$  in the diagonal one. Denoting by  $t_0 = t + c_0 |\mathbf{x}_i - z|^{2/(2i-1)}$  the transition time between the two regimes, we hence write (see also Section 2 in Chapter 6):

$$\begin{aligned} u(t, \mathbf{x}') &= \tilde{G}_{t_0, t}^{\boldsymbol{\xi}'} f(t, \mathbf{x}') + \tilde{G}_{T, t_0}^{\tilde{\boldsymbol{\xi}}'} f(t, \mathbf{x}') \\ &\quad + \tilde{P}_{t_0, t}^{\boldsymbol{\xi}'} u(t_0, \mathbf{x}') - \tilde{P}_{t_0, t}^{\tilde{\boldsymbol{\xi}}'} u(t_0, \mathbf{x}') \\ &\quad + \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left( \mathbb{I}_{s \leq t_0} \tilde{p}^{\boldsymbol{\xi}'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\boldsymbol{\xi}'}) \right. \\ &\quad \left. + \mathbb{I}_{s > t_0} \tilde{p}^{\tilde{\boldsymbol{\xi}}'}(t, s, \mathbf{x}', \mathbf{y})(L_s - \tilde{L}_s^{\tilde{\boldsymbol{\xi}}'}) \right) u(s, \mathbf{y}), \end{aligned} \quad (3.20)$$

where we denoted for all  $\boldsymbol{\xi}' \in \mathbb{R}^{nd}$ :

$$\forall 0 \leq v < r \leq T, \quad \tilde{G}_{r, v}^{\boldsymbol{\xi}'} f(t, \mathbf{x}') = \int_v^r ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{\boldsymbol{\xi}'}(t, s, \mathbf{x}', \mathbf{y}) f(s, \mathbf{y}). \quad (3.21)$$

Differentiating (3.19) w.r.t.  $\mathbf{x}_k$  and setting then  $\boldsymbol{\xi}' = \mathbf{x}'$ ,  $\tilde{\boldsymbol{\xi}}' = \mathbf{x}$ , we rewrite:

$$\begin{aligned} h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z) &=: (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i} \\ &\quad + (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i^c} + (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\partial \mathcal{S}_i}, \end{aligned} \quad (3.22)$$

where, thanks to (2.21), (3.20) and then (3.19):

$$\begin{aligned} &(h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i} \\ &:= \int_{\mathcal{S}_i} ds \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ \left[ H_k^{\boldsymbol{\xi}}(s, \mathbf{x}) - H_k^{\boldsymbol{\xi}'}(s, \mathbf{x}') \right] + \left[ I_k^{\boldsymbol{\xi}}(s, \mathbf{x}) - I_k^{\boldsymbol{\xi}'}(s, \mathbf{x}') \right] \right\}, \end{aligned} \quad (3.23)$$

corresponds to the difference of the previous expansions on the off-diagonal regime,

$$\begin{aligned} &(h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i^c} \\ &:= \int_{\mathcal{S}_i^c} ds \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ \left[ H_k^{\boldsymbol{\xi}}(s, \mathbf{x}) - H_k^{\tilde{\boldsymbol{\xi}}'}(s, \mathbf{x}') \right] + \left[ I_k^{\boldsymbol{\xi}}(s, \mathbf{x}) - I_k^{\tilde{\boldsymbol{\xi}}'}(s, \mathbf{x}') \right] \right\}, \end{aligned} \quad (3.24)$$

is the contribution of the diagonal regime and

$$\begin{aligned} &(h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\partial \mathcal{S}_i} \\ &:= \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\boldsymbol{\xi}'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\boldsymbol{\xi}}'} u(t_0, \mathbf{x}') \right\}, \end{aligned} \quad (3.25)$$

is the resulting boundary term. This last term, arising from the change of freezing point, is particularly delicate to analyze.

**Off-diagonal estimates: control of (3.23).** On the time set  $\mathcal{S}_i$ , we cannot expect some regularization from the difference of the transition densities so that we bluntly estimate the terms appearing in (3.23), writing:

$$\begin{aligned}
 & \left| h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z) \right|_{\mathcal{S}_i} \\
 := & \left| \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left( \int_{\mathcal{S}_i} ds \left[ H_k^\xi(s, \mathbf{x}) - H_k^{\xi'}(s, \mathbf{x}') \right] \right. \right. \\
 & \left. \left. + \int_{\mathcal{S}_i} ds \left[ I_k^\xi(s, \mathbf{x}) - I_k^{\xi'}(s, \mathbf{x}') \right] \right) \right|_{(\xi, \xi') = (\mathbf{x}, \mathbf{x}')} \\
 \leq & \int_{\mathbb{R}^d} d\mathbf{x}_i \frac{h_{cv}(z - \mathbf{x}_i)}{v^{\frac{1}{2}}} \int_{\mathcal{S}_i} ds \left( \left| H_k^\xi(s, \mathbf{x}) \right|_{\xi = \mathbf{x}} + \left| H_k^{\xi'}(s, \mathbf{x}') \right|_{\xi' = \mathbf{x}'} \right. \\
 & \left. + \left| I_k^\xi(s, \mathbf{x}) \right|_{\xi = \mathbf{x}} + \left| I_k^{\xi'}(s, \mathbf{x}') \right|_{\xi' = \mathbf{x}'} \right). \tag{3.26}
 \end{aligned}$$

Those terms can then be handled following the previous analysis performed in Theorem 3 and Lemma 3, observing here that, w.r.t. the previous proofs, the above terms are not differentiated w.r.t.  $\mathbf{x}_1$ . This improves the exponents of the time singularities of  $1/2$ . Similarly to (2.23), (2.25), (2.27) and (2.28) this therefore yields for the terms  $H_k^\xi(s, \mathbf{x})$ :

$$\left| H_k^\xi(s, \mathbf{x}) \right|_{\xi = \mathbf{x}} \leq C((s-t)^{-1+\delta_i^k} + \|\mathbf{D}D_1 u\|_\infty), \tag{3.27}$$

with  $1 > \delta_i^k > 1/2$ .

Reproducing the arguments that led to equations (2.31), (2.32), (2.39) and the statement of Lemma 3, exploiting again that there is now no differentiation w.r.t.  $\mathbf{x}_1$  we get with the notations of (2.35)

$$\left| I_k^\xi(s, \mathbf{x}) \right|_{\xi = \mathbf{x}} \leq C(s-t)^{-1+\gamma_i^k} (\|\mathbf{D}u\|_\infty + \sup_{\bar{s} \in [0, T]} \|D_i D_k u^i(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}).$$

Similar bounds hold for  $\left| H_k^{\xi'}(s, \mathbf{x}') \right|_{\xi' = \mathbf{x}'}$  and  $\left| I_k^{\xi'}(s, \mathbf{x}') \right|_{\xi' = \mathbf{x}'}$ . Hence, from (3.27) and (3.28),

$$\begin{aligned}
 & \int_{\mathcal{S}_i} ds \left( \left| H_k^\xi(s, \mathbf{x}) \right|_{\xi = \mathbf{x}} + \left| H_k^{\xi'}(s, \mathbf{x}') \right|_{\xi' = \mathbf{x}'} + \left| I_k^\xi(s, \mathbf{x}) \right|_{\xi = \mathbf{x}} + \left| I_k^{\xi'}(s, \mathbf{x}') \right|_{\xi' = \mathbf{x}'} \right) \\
 \leq & C \left( \int_t^{(t+c_0|\mathbf{x}_i-z|^{2i-1}) \wedge T} ds \left( (s-t)^{-1+\delta_i^k} \|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty \right) \right. \\
 & \left. + (s-t)^{-1+\gamma_i^k} (\|\mathbf{D}u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}) \right) \\
 \leq & C \left( |\mathbf{x}_i - z|^{\frac{2\delta_i^k}{2i-1}} \wedge (T-t)^{\delta_i^k} (\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty) \right. \\
 & \left. + |\mathbf{x}_i - z|^{\frac{2\gamma_i^k}{2i-1}} \wedge (T-t)^{\gamma_i^k} (\|\mathbf{D}u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}) \right) \\
 \leq & C |\mathbf{x}_i - z|^{\alpha_i^k} T^{\delta'} \left( \|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right), \tag{3.28}
 \end{aligned}$$

for some  $\delta' := \delta'(\mathbf{A}) > 0$ , denoting by

$$\sup_{\bar{s} \in [0, T]} \left\| (D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} := \sup_{\bar{s} \in [0, T], \mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq i} \left\| D_{\mathbf{x}_i} D_{\mathbf{x}_k} u^i(\bar{s}, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}, \quad (3.29)$$

and recalling for the last inequality that  $\gamma_i^k = 1/2 + \eta(i - 3/2)$  so that  $2\gamma_i^k/(2i - 1) = [1 + 2\eta(i - 3/2)]/(2i - 1) > (1 + \eta/4)/(2i - 1) = \alpha_i^k$  (see also the statements of Lemmas 3 and 4) and similarly for the contributions involving  $\delta_i^k > 1/2$ . We eventually get from (3.28) and (3.26):

$$\begin{aligned} & \left| h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z) \right|_{\mathcal{S}_i} \\ & \leq CT^{\delta'} \left( \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1u\|_{\infty} + \sup_{\bar{s} \in [0, T]} \left\| (D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right) \int_{\mathbb{R}^d} d\mathbf{x}_i \frac{h_{cv}(z - \mathbf{x}_i)}{v^{\frac{1}{2}}} |z - \mathbf{x}_i|^{\alpha_i^k} \\ & \leq CT^{\delta'} \left( \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1u\|_{\infty} + \sup_{\bar{s} \in [0, T]} \left\| (D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right) v^{\frac{\alpha_i^k - 1}{2}}. \end{aligned} \quad (3.30)$$

**Diagonal estimates: control of the term (3.24).** We are now going to handle the term (3.24) which correspond to the r.h.s. of (3.19) on  $\mathcal{S}_i^c$ . In that case the points  $\mathbf{x}_i$  and  $\mathbf{x}'_i$  are close w.r.t. the characteristic time scale of the  $i^{\text{th}}$  variable and the main idea consists in controlling the difference between the frozen densities at  $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}' = \mathbf{x}$  with starting points  $\mathbf{x}$  and  $\mathbf{x}'$  respectively. Precisely, recalling that  $\mathbf{x}$  and  $\mathbf{x}'$  only differ in the  $i^{\text{th}}$  component, we can write:

$$D_{\mathbf{x}_k} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}', \mathbf{y}) = - \int_0^1 d\lambda D_{\mathbf{x}_i} \left( D_{\mathbf{x}_k} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \right) \cdot (\mathbf{x}' - \mathbf{x})_i. \quad (3.31)$$

From Lemma 1 we thus derive:

$$|D_{\mathbf{x}_k} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}', \mathbf{y})| \leq \frac{C|(\mathbf{x}' - \mathbf{x})_i|}{(s - t)^{(i - \frac{1}{2}) + (k - \frac{1}{2})}} \int_0^1 d\lambda \hat{p}_{C-1}^{\boldsymbol{\xi}}(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y}). \quad (3.32)$$

Now, from the definition of  $\hat{p}_{C-1}$  in Proposition 2, recalling as well from (2.7) that  $\mathbf{x} \mapsto \mathbf{m}_{s,t}^{\boldsymbol{\xi}}(\mathbf{x})$  is affine, we get:

$$\begin{aligned} & |\hat{p}_{C-1}^{\boldsymbol{\xi}}(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y})| \\ & \leq \frac{C}{(s - t)^{\frac{n^2 d}{2}}} \exp(-c(s - t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{\boldsymbol{\xi}}(\mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x})) - \mathbf{y})|^2) \\ & \leq \frac{C}{(s - t)^{\frac{n^2 d}{2}}} \exp(c(s - t) |\mathbb{T}_{s-t}^{-1} \mathbf{m}_{s,t}^{\boldsymbol{\xi}}(\mathbf{x} - \mathbf{x}')|^2) \exp(-\frac{c}{2}(s - t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{\boldsymbol{\xi}}(\mathbf{x}) - \mathbf{y})|^2). \end{aligned}$$

Using the rescaling arguments of the proof of Proposition 2 on the resolvent (see equation (2.13)), we then get  $(s - t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1} \mathbf{m}_{s,t}^{\boldsymbol{\xi}}(\mathbf{x} - \mathbf{x}')| \leq C(s - t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \mathbf{x}')| = C(s - t)^{-i+1/2} |(\mathbf{x}' - \mathbf{x})_i| \leq C$ , from the very definition of  $\mathcal{S}_i^c$ . Hence,

$$|\hat{p}_{C-1}^{\boldsymbol{\xi}}(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y})| \leq \frac{C}{(s - t)^{\frac{n^2 d}{2}}} \exp(-\frac{c}{2}(s - t) |\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^{\boldsymbol{\xi}}(\mathbf{x}) - \mathbf{y})|^2),$$



so that, from (3.32) and recalling that on  $\mathcal{S}_i^c$ ,  $|(\mathbf{x}' - \mathbf{x})_i|/(s - t) \leq C(|(\mathbf{x}' - \mathbf{x})_i|/(s - t))^{\alpha_i^k}$ , the following important control holds:

$$|D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y})| \leq \frac{C|(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k}}{(s - t)^{\alpha_i^k(i - \frac{1}{2}) + (k - \frac{1}{2})}} \hat{p}_{C^{-1}}^\xi(t, s, \mathbf{x}, \mathbf{y}). \quad (3.33)$$

Write now from (3.24), recalling that  $\tilde{\xi}' = \xi$ ,

$$\begin{aligned} & \left| h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z) \right|_{\mathcal{S}_i^c} \\ := & \left| \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left( \int_{\mathcal{S}_i^c} ds \left[ H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}') \right] + \int_{\mathcal{S}_i^c} ds \left[ I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}') \right] \right) \right|, \end{aligned} \quad (3.34)$$

and let us discuss how the terms  $H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}')$ ,  $I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}')$  in the above equation can be handled.

We first focus on the term  $I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}')$  in (3.34). This contribution, associated with the degenerate components of perturbed operator, is again the most delicate to handle. From the definitions in (2.30) we are led to control the sum  $\sum_{\ell=1}^3 [I_{\ell,k}^\xi(s, \mathbf{x}) - I_{\ell,k}^\xi(s, \mathbf{x}')$ . For the terms  $I_{1,k}^\xi(s, \mathbf{x}) - I_{1,k}^\xi(s, \mathbf{x}')$ ,  $I_{2,k}^\xi(s, \mathbf{x}) - I_{2,k}^\xi(s, \mathbf{x}')$  we are going to reproduce the analysis leading to (2.31), (2.32). Observe first that the above terms do not involve  $D_{\mathbf{x}_1}$ , therefore we gain a singularity of order 1/2 w.r.t. to the indicated equations (2.31), (2.32). On the other hand, the difference of the derivatives of the frozen densities w.r.t.  $\mathbf{x}_k$  can be handled with (3.33). This leads to:

$$\begin{aligned} & \left| \sum_{\ell=1}^2 \int_{t+c_0|(\mathbf{x}' - \mathbf{x})_i|^{\frac{2}{2i-1}}}^T ds \left( I_{\ell,k}^\xi(s, \mathbf{x}) - I_{\ell,k}^\xi(s, \mathbf{x}') \right) \right| \\ & \leq C|(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} \|\mathbf{D}u\|_\infty \sum_{j=k}^n \int_t^T ds (s - t)^{-(k - \frac{1}{2}) - \alpha_i^k(i - \frac{1}{2})} \left( (s - t)^{\beta_j(j - \frac{1}{2})} + (s - t)^{(1+\eta)(j - \frac{1}{2})} \right), \end{aligned}$$

changing the summation variables from (2.31) for notational simplicity.

From the very definition of  $\alpha_i^k = (1 + \eta/4)/(2i - 1)$  in Lemma 3 and the specific choice of  $\eta$  in assumption **(A)** (see **(H $_\eta$ )**) we derive

$$\left| \sum_{\ell=1}^2 \int_{t+c_0|(\mathbf{x}' - \mathbf{x})_i|^{\frac{2}{2i-1}}}^T ds \left( I_{\ell,k}^\xi(s, \mathbf{x}) - I_{\ell,k}^\xi(s, \mathbf{x}') \right) \right| \leq C|(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} \|\mathbf{D}u\|_\infty T^\delta, \quad (3.35)$$

for some  $\delta > 0$ .

From the previous analysis it is therefore sufficient to focus on the tricky term, namely  $I_{3,k}^\xi(s, \mathbf{x})$  introduced in (2.30). We begin the proof considering first  $I_{3,k}^\xi(s, \mathbf{x})$ . Exploiting as well Lemma 1 for a centering argument w.r.t. the  $k^{\text{th}}$  variable, we write:

$$\begin{aligned} & I_{3,k}^\xi(s, \mathbf{x}) - I_{3,k}^\xi(s, \mathbf{x}') \\ = & \sum_{\ell=2}^k \int_{\mathbb{R}^{nd}} d\mathbf{y} \left( \mathbf{F}_\ell(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\ell-1} \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1} \right) \\ & \times \left( D_{\mathbf{y}_\ell} u(s, \mathbf{y}) - D_{\mathbf{y}_\ell} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) \right) \left( D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) \right). \end{aligned}$$

Let us reproduce now the arguments used in Section 2.2 to handle  $I_{3,k}^\xi$  (see e.g. the computations from equation (2.33) to (2.34)). Expanding with the Taylor formula the difference  $\left(D_{\mathbf{y}_\ell} u(s, \mathbf{y}) - D_{\mathbf{y}_\ell} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi}))\right)$ , using the Schwarz theorem to exchange the order of differentiations<sup>‡‡</sup>, we obtain with the notations of (2.35) (see also the comments following this equation):

$$\begin{aligned}
 & I_{3,k}^\xi(s, \mathbf{x}) - I_{3,k}^\xi(s, \mathbf{x}') \\
 &= \sum_{m=k}^n \sum_{\ell=2}^k \int_0^1 d\lambda \int_{\mathbb{R}^{nd}} d\mathbf{y} \\
 & \quad \left[ \left( \mathbf{F}_\ell(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\ell-1} \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1} \right) \right. \\
 & \quad \left. \times (D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_m \right] \\
 & \quad D_{\mathbf{y}_\ell} D_{\mathbf{y}_m} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi}) + \lambda(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k:n}) \\
 &=: \sum_{m=k}^n \sum_{\ell=2}^k \int_0^1 d\lambda \int_{\mathbb{R}^{nd}} d\mathbf{y} \left[ \Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\mathbf{y}_\ell) - \Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\mathbf{y}_\ell) \right] \\
 & \quad D_{\mathbf{y}_\ell} D_{\mathbf{y}_m} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi}) + \lambda(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k:n}).
 \end{aligned}$$

Thus, we derive similarly to (2.39):

$$\begin{aligned}
 |I_{3,k}^\xi(s, \mathbf{x}) - I_{3,k}^\xi(s, \mathbf{x}')| &\leq \sum_{m=k}^n \sum_{\ell=2}^k \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}) \\
 & \quad \left\{ \left\| \Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right\|_{B_{1,1}^{1-\alpha_\ell^m}} \right. \\
 & \quad \left. \times \sup_{\mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq \ell} \left\| D_\ell D_m u(s, \mathbf{z}_{1:\ell-1}, \cdot, \mathbf{z}_{\ell+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_\ell^m - 1}} \right\},
 \end{aligned} \tag{3.36}$$

where  $\int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})$  means that we integrate over  $\mathbf{y}_{1:\ell-1}$  and  $\mathbf{y}_{\ell+1:n}$ . To conclude, we need the following appropriate version of Lemma 3 to handle the Besov norm with negative exponent in the above r.h.s. Its proof is postponed to the next section.

**Lemma 5** *Let  $k \in \llbracket 2, n \rrbracket$ ,  $\ell \in \llbracket 2, k \rrbracket$  and  $m \in \llbracket k, n \rrbracket$  and let  $\Psi_{\ell, (k, i), m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be the function defined by (2.35). There exist  $C := C(\mathbf{A}) > 0$  and  $\gamma_\ell^m := \gamma_\ell^m(\mathbf{A}) := 1/2 + \eta(\ell - 3/2) > 1/2$  such that*

$$\begin{aligned}
 & \left\| \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right\|_{B_{1,1}^{1-\alpha_\ell^m}} \\
 & \leq C \hat{q}_{c\setminus \ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})) (s-t)^{-1-(i-\frac{1}{2})\alpha_i^k + \gamma_\ell^m} |\mathbf{x} - \mathbf{x}'|_i^{\alpha_i^k},
 \end{aligned}$$

with  $\hat{q}_{c\setminus \ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}))$  as in (2.40).

<sup>‡‡</sup>. Recall indeed that what we are able to control is precisely the Hölder moduli of the derivatives  $D_{\mathbf{y}_m} u(s, \cdot)$  w.r.t. variables  $\ell \leq m$ .

Again, for the specific choice of  $\alpha_i^k = (1 + \eta/4)/(2i - 1)$  performed in the proof of Lemma 3, we eventually derive from Lemma 5 and (3.36), with the notation of (3.29), that:

$$\begin{aligned}
 & \int_{\mathcal{S}_i^c} ds |I_{3,k}^\xi(s, \mathbf{x}) - I_{3,k}^\xi(s, \mathbf{x}')| \\
 \leq & C \left( \sum_{m=k}^n \sum_{\ell=2}^k \sup_{s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{-1 + \alpha_\ell^m}} \int_t^T ds (s-t)^{-1 - (i - \frac{1}{2})\alpha_i^k + \gamma_\ell^m} \right) |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k} \\
 \leq & CT^\delta \sup_{m \in \llbracket k, n \rrbracket, \ell \in \llbracket 1, k \rrbracket, s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_\ell^{m-1}}} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k}, \tag{3.37}
 \end{aligned}$$

for some  $\delta := 3\eta/8 > 0$ . Combining this estimate together with (3.35) we eventually derive

$$\begin{aligned}
 & \left| \int_{\mathcal{S}_i^c} ds \left[ I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}') \right] \right| \\
 \leq & T^\delta C |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} \left( \|\mathbf{D}u\|_\infty + \sup_{m \in \llbracket k, n \rrbracket, \ell \in \llbracket 1, k \rrbracket, s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_\ell^{m-1}}} \right). \tag{3.38}
 \end{aligned}$$

Now, the term  $H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}')$  in (3.34) (non-degenerate variables) can be handled reproducing the same previous arguments for  $I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}')$ , exploiting (3.33) and following the computations performed for  $H_k$  in the proof of Theorem 3 (see e.g. (2.29)). From the definition of  $\mathcal{S}_i^c$  we obtain:

$$\begin{aligned}
 & \left| \int_{t+c_0|(\mathbf{x}' - \mathbf{x})_i|^{\frac{2}{2i-1}}}^T ds \left( H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}') \right) \right| \\
 \leq & C |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} (\|\mathbf{D}D_{\mathbf{x}_1} u\|_\infty + \|\mathbf{D}u\|_\infty) \int_t^T ds (s-t)^{-1 - \alpha_i^k(i - \frac{1}{2}) + \delta_i^k} \\
 \leq & C |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} (\|\mathbf{D}D_{\mathbf{x}_1} u\|_\infty + \|\mathbf{D}u\|_\infty) T^\delta, \tag{3.39}
 \end{aligned}$$

for some  $\delta > 0$  recalling that  $\delta_i^k > 1/2$  for the last inequality. The arguments needed to control this term are actually those already exploited in [CdR17] when  $n = 2$ .

Gathering equations (3.38) and (3.39), we finally derive with the notations of (3.34):

$$\begin{aligned}
 & |h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)| \Big|_{\mathcal{S}_i^c} \\
 \leq & CT^\delta \int_{\mathbb{R}^d} d\mathbf{x}_i \frac{h_{cv}(z - \mathbf{x}_i)}{v^{\frac{1}{2}}} |\mathbf{x}_i - z|^{\alpha_i^k} \\
 & \times \left( \|\mathbf{D}D_{\mathbf{x}_1} u\|_\infty + \|\mathbf{D}u\|_\infty + \sup_{2 \leq \ell \leq m \leq n, s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_\ell^{m-1}}} \right) \\
 \leq & CT^\delta v^{\frac{\alpha_i^k - 1}{2}} \left( \|\mathbf{D}D_{\mathbf{x}_1} u\|_\infty + \|\mathbf{D}u\|_\infty + \sup_{2 \leq \ell \leq m \leq n, s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_\ell^{m-1}}} \right). \tag{3.40}
 \end{aligned}$$

**Discontinuity term associated with the regime time change: control of the term (3.25).** Here, we aim at handling

$$\begin{aligned} & (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\partial \mathcal{S}_i} \\ &= \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \right\}, \end{aligned}$$

which we will actually handle like the off-diagonal components. Recall here that the transition time  $t_0 = t + c_0 |\mathbf{x} - \mathbf{x}'|^{2/(2i-1)}$  and that  $(\xi', \tilde{\xi}') = (\mathbf{x}', \mathbf{x})$ , where these values are again plugged in the expansions after the derivations. From Lemma 1 (cancellation argument), we write:

$$\begin{aligned} & D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \tag{3.41} \\ &= \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n})] d\mathbf{y} \\ &\quad - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n})] d\mathbf{y} \\ &= \left\{ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n})] d\mathbf{y} \right. \\ &\quad \left. - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n})] d\mathbf{y} \right\} \tag{3.42} \\ &\quad + \left\{ \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[ u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}) \right. \right. \right. \\ &\quad \left. \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right] \right. \\ &\quad \left. - \left[ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) \left[ u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}) \right. \right. \right. \\ &\quad \left. \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right] \right\} \tag{3.43} \\ &\quad + \left\{ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k \rangle d\mathbf{y} \right. \tag{3.44} \\ &\quad \left. - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k \rangle d\mathbf{y} \right\} \\ &=: (\mathcal{B}_1^{\xi', \tilde{\xi}'} + \mathcal{B}_2^{\xi', \tilde{\xi}'} + \mathcal{B}_3^{\xi', \tilde{\xi}'})(t_0, \mathbf{x}'). \end{aligned}$$

We now exploit the Hölder regularity of  $D_k u$  w.r.t. the  $k^{\text{th}}$  variable to control the terms

in  $\mathcal{B}_2^{\xi', \tilde{\xi}'}$  ( $t_0, \mathbf{x}'$ ) defined in (3.43). Let us first write from the previous decomposition:

$$\begin{aligned} & |\mathcal{B}_{2,1}^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| \\ &:= \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[ u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k:n}) \right. \right. \\ &\quad \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right| \\ &= \int_{\mathbb{R}^{nd}} \int_0^1 d\lambda \left| D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[ D_k u(t_0, \mathbf{y}_{1:k-1}, \cdot, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n}) \right]_{\alpha_k^k} |(\mathbf{y} - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_k|^{1+\alpha_k^k} d\mathbf{y}, \right. \end{aligned}$$

where  $[\cdot]_{\alpha_k^k}$  denotes the Hölder modulus of order  $\alpha_k^k$ . From Proposition 2, we thus derive:

$$\begin{aligned} |\mathcal{B}_{2,1}^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| &\leq C(t_0 - t)^{\alpha_k^k(k-\frac{1}{2})} \sup_{z_j, j \in \llbracket 1, n \rrbracket, j \neq k} [D_k u(t_0, \mathbf{z}_{1:k}, \cdot, \mathbf{z}_{k+1:n})]_{\alpha_k^k} \\ &\quad \times \int_{\mathbb{R}^{(n-1)d}} d\mathbf{y}_{1:k-1} d\mathbf{y}_{k+1:n} \hat{q}_{C^{-1} \setminus k}(t, s, \mathbf{x}, (\mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n})) \\ &= C(t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} [(D_k u)_k(t_0, \cdot)]_{\alpha_k^k}, \end{aligned} \quad (3.45)$$

recalling from (3.12) that  $\alpha_k^k = (1 + \eta/4)/(2k - 1)$  and using the notation of (1.2) for the last inequality. The same arguments readily give:

$$\begin{aligned} & |\mathcal{B}_{2,2}^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| \\ &:= \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[ u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n}) \right. \right. \\ &\quad \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi'}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right| \\ &\leq C(t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} [(D_k u)_k(t_0, \cdot)]_{\alpha_k^k}. \end{aligned} \quad (3.46)$$

Let us now deal with the contribution  $\mathcal{B}_1^{\xi', \tilde{\xi}'}$  ( $t_0, \mathbf{x}'$ ) in (3.42). Observe from this definition that this term is non zero if and only if  $k < n$ . Write then

$$\begin{aligned} & |\mathcal{B}_1^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| \\ &\leq \int_{\mathbb{R}^{nd}} |D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y})| |u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n})| d\mathbf{y} \\ &\quad + \int_{\mathbb{R}^{nd}} |D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y})| |u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n})| d\mathbf{y} \\ &\leq C \|\mathbf{D}u\|_{\infty} \left( \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t_0 - t)^{k-\frac{1}{2}}} \hat{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) |(\boldsymbol{\theta}_{t_0,t}(\mathbf{x}') - \mathbf{y})_{k+1:n}| \right. \\ &\quad \left. + \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t_0 - t)^{k-\frac{1}{2}}} \hat{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) (|(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \mathbf{y})_{k+1:n}| + |(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n}|) \right). \end{aligned} \quad (3.47)$$

To deal with the last contribution in the r.h.s., we will need some auxilliary lemmas already used in Chapter 6 for Schauder estimates. Namely, analogously to Lemmas 1 and 3 therein, we have the following result.

**Lemma 6** *There exists  $\vartheta = \vartheta((\mathbf{A})) \in (0, 1)$  s.t. for all  $j \in \llbracket k, n \rrbracket$ :*

$$|(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_j| \leq c_0^\vartheta |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2j-1}{2i-1}}. \quad (3.48)$$

In particular, recalling that  $t_0 = t + c_0 |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2}{2i-1}}$  with  $|(\mathbf{x} - \mathbf{x}')_i| \leq 1$ , we obtain

$$\begin{aligned} \frac{\|\mathbf{D}u\|_\infty}{(t_0 - t)^{k-\frac{1}{2}}} |(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k+1:n}| &\leq \frac{\|\mathbf{D}u\|_\infty}{c_0^{k-\frac{1}{2}} |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2k-1}{2i-1}}} c_0^\vartheta |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2(k+1)-1}{2i-1}} \\ &\leq \|\mathbf{D}u\|_\infty c_0^{\vartheta-(k-\frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2}{2i-1}} \\ &\leq \|\mathbf{D}u\|_\infty c_0^{\vartheta-(k-\frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k}. \end{aligned}$$

Plugging the above control in (3.47) we obtain

$$|\mathcal{B}_1^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| \leq C \|\mathbf{D}u\|_\infty \left\{ 2(t_0 - t) + c_0^{\vartheta-(k-\frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k} \right\}. \quad (3.49)$$

Let us eventually control the term  $\mathcal{B}_3^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')$  defined in (3.44) which we rewrite in the following way:

$$\begin{aligned} &\mathcal{B}_3^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}') \\ = &\left\{ \int_{\mathbb{R}^{nd}} dy D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle [D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_{k:n}) - D_k u(t_0, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))], \right. \right. \\ &\quad \left. \left. (\mathbf{y} - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_k \right\rangle \right. \\ &- \int_{\mathbb{R}^{nd}} dy D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle [D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0,t}^{\xi}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')_{k+1:n}) \right. \\ &\quad \left. - D_k u(t_0, (\mathbf{m}_{t_0,t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')_{k+1:n})], (\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x}'))_k \right\rangle \left. \right\} \\ + &\left\{ \int_{\mathbb{R}^{nd}} dy D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \right. \\ &\quad \times \left\langle [D_k u(t_0, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')) - D_k u(t_0, (\mathbf{m}_{t_0,t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_k] \right\rangle \left. \right\} \\ - &\left\{ \int_{\mathbb{R}^{nd}} dy D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle D_k u(t_0, (\mathbf{m}_{t_0,t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0,t}^{\xi}(\mathbf{x}'))_k \right\rangle \right. \\ &\quad \left. - \int_{\mathbb{R}^{nd}} dy D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle D_k u(t_0, (\mathbf{m}_{t_0,t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0,t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_k \right\rangle \right\}, \end{aligned}$$

where, thanks to Proposition 3 the last contribution is actually 0. For the first and second contributions in the above r.h.s. we have, thanks to the Hölder regularity of  $\mathbf{x}_{1:k} \mapsto D_k u(\cdot, \mathbf{x}_{1:k}, \cdot)$ , Proposition 2 and Lemma 6:

$$\begin{aligned}
 & \mathcal{B}_3^{\xi'; \tilde{\xi}'}(t_0, \mathbf{x}') \\
 & \leq C \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\{ \hat{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) + \hat{p}_{C^{-1}}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) \right\} \sum_{j=1}^{k-1} [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} (t_0 - t)^{\alpha_j^k(j-\frac{1}{2})} \\
 & \quad + C \int_{\mathbb{R}^{nd}} d\mathbf{y} \hat{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \sum_{j=1}^k [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} (t_0 - t)^{\alpha_j^k(j-\frac{1}{2})} \\
 & \leq C \left( \sum_{j=1}^k [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} \right) (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}}, \tag{3.50}
 \end{aligned}$$

with again the notation of (1.2). Thus, plugging estimates (3.45), (3.46), (3.49) and (3.50) in (3.41) we deduce that

$$\begin{aligned}
 & |D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\xi}'} u(t_0, \mathbf{x}')| \\
 & \leq C(t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} [(D_k u)_k(t_0, \cdot)]_{\alpha_k^k} + C \|\mathbf{D}u\|_{\infty} \left\{ (t_0 - t) + c_0^{\vartheta - (k-\frac{1}{2})} |\mathbf{x} - \mathbf{x}'|_i^{\alpha_i^k} \right\} \\
 & \quad + C \left( \sum_{j=1}^k [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} \right) (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} \\
 & \leq C(t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} \|(D_k^2 u)_k(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_{k-1}^k}} + C \|\mathbf{D}u\|_{\infty} \left\{ (t_0 - t) + c_0^{\vartheta - (k-\frac{1}{2})} |\mathbf{x} - \mathbf{x}'|_i^{\alpha_i^k} \right\} \\
 & \quad + C \left( \sum_{j=1}^k \|(D_j D_k u)_j(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_{j-1}^k}} \right) (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}}. \tag{3.51}
 \end{aligned}$$

The last inequality is a consequence of Proposition 4, stated in Appendix 5.2, which allows to dominate the Hölder norms  $\left( [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} \right)_{2 \leq j \leq k \leq n}$  in terms of the Besov norms  $\left( \|(D_j D_k u)_j(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_{j-1}^k}} \right)_{2 \leq j \leq k \leq n}$ , with the notations of (3.29). This point is actually crucial to complete our circular argument.

From the definition of  $t_0 = t + c_0 |\mathbf{x} - \mathbf{x}'|_i^{2/(2i-1)}$  and  $\alpha_i^k = (1 + \eta/4)/(2i-1)$ , recalling that  $t_0 - t$  is small as well (i.e.  $t_0 - t \leq C(t_0 - t)^{1/2 + \eta/8}$ ), we obtain from (3.51):

$$\begin{aligned}
 & |D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\xi}'} u(t_0, \mathbf{x}')| \\
 & \leq C \left\{ c_0^{\frac{1}{2} + \frac{\eta}{8}} \left( \sum_{j=1}^k \|(D_j D_k u)_j(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_{j-1}^k}} \right) + (c_0^{\vartheta - (k-\frac{1}{2})} + c_0^{\frac{1}{2} + \frac{\eta}{8}}) \|\mathbf{D}u\|_{\infty} \right\} |\mathbf{x} - \mathbf{x}'|_i^{\alpha_i^k}.
 \end{aligned}$$

Thus, from (3.25), there exists  $\tilde{\delta} := \tilde{\delta}(\mathbf{A}) \in (0, 1)$  such that

$$\begin{aligned}
 & \left| (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)) |_{\partial \mathcal{S}_i} \right| \\
 & \leq v^{\frac{\alpha_i^k - 1}{2}} C(c_0^{-\tilde{\delta} - 1} \|\mathbf{D}u\|_{\infty} + c_0^{\tilde{\delta}} \max_{2 \leq \ell \leq m \leq n} \|(D_{\ell} D_m u)_{\ell}(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_{\ell}^{m-1}}}). \tag{3.52}
 \end{aligned}$$

**Conclusion: control of (3.22).** Plugging (3.30), (3.40) and (3.52) into (3.22), (3.19), we eventually derive that for some positive  $\tilde{\delta} := \tilde{\delta}(\mathbf{A}) > 0$ :

$$\begin{aligned} & \|D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \\ & \leq C \left( c_0^{-\tilde{\delta} - 1} \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_{\mathbf{x}_1} u\|_{\infty} + (c_0^{\tilde{\delta}} + T^{\tilde{\delta}}) \sup_{s \in [0, T], 2 \leq \ell \leq m \leq n} \|(D_{\ell} D_m u)_{\ell}(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_{\ell}^m - 1}} \right). \end{aligned} \quad (3.53)$$

The main point to close our circular argument consists then in taking the supremums over w.r.t.  $\mathbf{x}_{1:i-1}, \mathbf{x}_{i+1:n}, i, k$  and  $t \in [0, T]$  on the l.h.s. and to tune the constant  $c_0$  and the terminal time  $T$  in order to obtain  $C(c_0^{\tilde{\delta}} + T^{\tilde{\delta}}) \leq 1/2$ . We then derive that for all  $2 \leq i \leq k \leq n$ :

$$\sup_{t \in [0, T]} \|(D_i D_k u)_i(t, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \leq C \{ \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_{\mathbf{x}_1} u\|_{\infty} \}, \quad (3.54)$$

which concludes the proof of Lemma 4.

### 3.3 Proof of the technical lemmas

**Proof:** [Proof of Lemma 5]

We follow the proof of Lemma 3, concentrating on the case  $\ell \leq k - 1$ , the specific case  $\ell = k$  could be treated similarly considering the slightly different cancellation terms already discussed in Lemma 3. The quantity to estimate is now:

$$\left[ \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right]. \quad (3.55)$$

Splitting the thermic part of the Besov norm as in (3.1), we obtain the same kind estimate for the non-singular in time part. Indeed, we point out that the difference (3.55) does not involve  $D_{\mathbf{x}_1}$ , therefore we gain a singularity of order  $1/2$  w.r.t. to equation (3.2). On the other hand, the difference of the derivatives of the frozen densities w.r.t.  $\mathbf{x}_k$  can be handled with (3.33). Choosing  $\rho_{\ell, m} = 2\ell - 1$  as in the proof of Lemma 3, and recalling that  $\alpha_i^k(i - 12) = (1 + \eta/4)/2$  (see (3.12)), it is plain to check that:

$$\begin{aligned} & \mathbf{U} \left[ \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right] \\ & \leq \frac{\tilde{C} |\mathbf{x} - \mathbf{x}'|_i \alpha_i^k}{(s - t)^{1 + \alpha_i^k(i - \frac{1}{2}) - \gamma_{\ell}^m}} \hat{q}_{c \setminus \ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})), \end{aligned} \quad (3.56)$$

where from  $(\mathbf{H}_{\eta})$ ,  $1 + \alpha_i^k(i - 1/2) - \gamma_{\ell}^m < 1$ . Turning to the singular in time contribution of the thermic part of the Besov norm of (3.55) we decompose with the notations of Lemma 3 (see e.g. (3.6), (3.7) which exhibit an additional spatial derivative):

$$\begin{aligned} & \mathbf{L} \left[ \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right] \\ & =: \int_0^{(s-t)^{\rho_{\ell, m}}} dw w^{\frac{\alpha_{\ell}^m - 1}{2}} \int_{\mathbb{R}^d} d\tilde{z} |(\mathcal{I}_{1, \ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}, \mathbf{x}')} + \mathcal{I}_{2, \ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}, \mathbf{x}')})(w, \tilde{z})|, \end{aligned}$$



where  $\ell < k \leq m$  and:

$$\begin{aligned} & \mathcal{F}_{1,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x},\mathbf{x}')} (w, \tilde{z}) \\ &= \int_{\mathbb{R}^d} \partial_w h_w(\tilde{z} - \mathbf{y}_\ell) (D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y})) \\ & \quad \langle \mathbf{F}_\ell(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})), (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_m \rangle d\mathbf{y}_\ell, \end{aligned} \quad (3.57)$$

with a slight abuse of notation when  $\ell = k - 1$  and

$$\begin{aligned} & \mathcal{F}_{2,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})} (w, \tilde{z}) \\ &= \int_{\mathbb{R}^d} \partial_w h_w(\tilde{z} - \mathbf{y}_\ell) \left[ \left( D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right. \\ & \quad \left. - \left( D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right] \\ & \quad \langle \mathbf{F}_\ell(s, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\mathbf{x}_{\ell-1}} \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1}, \\ & \quad (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_m \rangle d\mathbf{y}_\ell. \end{aligned} \quad (3.58)$$

Note now that when proceeding first as in (3.31), (3.32) and then control the resulting difference as in (3.7) we get, thanks to (3.33), that:

$$\begin{aligned} & \left| \left( D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right. \\ & \quad \left. - \left( D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right| \\ & \leq \frac{\tilde{C}}{(s-t)^{\ell-\frac{1}{2}+i-\frac{1}{2}+k-\frac{1}{2}}} \int_0^1 d\lambda \hat{p}_{C^{-1}}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z} + \lambda(\mathbf{y}_\ell - z), \mathbf{y}_{\ell+1:n}) |(\mathbf{x}' - \mathbf{x})_i| |\tilde{z} - \mathbf{y}_\ell| \\ & \leq \frac{\tilde{C}}{(s-t)^{(\ell-\frac{1}{2})+\alpha_i^k(i-\frac{1}{2})+(k-\frac{1}{2})}} \int_0^1 d\lambda \hat{p}_{C^{-1}}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z} + \lambda(\mathbf{y}_\ell - z), \mathbf{y}_{\ell+1:n}) |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} |\tilde{z} - \mathbf{y}_\ell|, \end{aligned}$$

using the fact we are in the diagonal regime in the last inequality. With this control at hand, together with estimate (3.33), to handle the contributions  $(\mathcal{F}_{1,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})} + \mathcal{F}_{2,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})})(w, \tilde{z})$ , we can mimic the proof of the estimation of the contributions  $(\mathcal{F}_{1,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} + \mathcal{F}_{2,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})})(v, z)$  done in Lemma 3 to obtain

$$\begin{aligned} & \mathbf{L} \left[ \Psi_{\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})}(\cdot) - \Psi_{\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x}')}(\cdot) \right] \\ & \leq \tilde{C} \hat{q}_{c\ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})) |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} \int_0^{(s-t)^{\rho_{\ell,m}}} dw w^{\frac{\alpha_\ell^m - 1}{2}} \\ & \quad \left\{ \frac{1}{w^{1-\frac{\beta_\ell}{2}} (s-t)^{\alpha_i^k(i-\frac{1}{2})}} + \frac{w^{-\frac{1+\beta_\ell}{2}}}{(s-t)^{\ell-\frac{1}{2}+\alpha_i^k(i-\frac{1}{2})}} + \frac{w^{-\frac{1}{2}}}{(s-t)^{\ell-\frac{1}{2}+\alpha_i^k(i-\frac{1}{2})-(1+\eta)(\ell-3/2)}} \right\} \\ & \leq \tilde{C} \hat{q}_{c\ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})) (s-t)^{-1-(i-\frac{1}{2})\alpha_i^k + \gamma_\ell^m} |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k}, \end{aligned}$$

where  $\gamma_\ell^m = 1/2 + \eta(\ell - 3/2)$ . From  $(\mathbf{H}_\eta)$  and the very definition of  $\alpha_i^k$  we hence have  $1 - \alpha_i^k(i - 1/2) + \gamma_\ell^m < 1$  which, together with (3.56), concludes the proof.  $\square$

## 4 Proof of technical results of Lemma 6

In this section, we establish some control of the flows sensitivity. Our results are still available up to any mollification procedure.

First of all, let us state a control of the spatial regularity of the flow. Next, we will provide a proof of Lemma 6.

### 4.1 A first sensitivity result for the flow

**Lemma 7** *Under  $(\mathbf{A})$ , there exists  $C := C((\mathbf{A}), T)$  s.t. for all  $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$ ,  $0 \leq t < s \leq T$  and  $i \in \llbracket 1, n \rrbracket$ :*

$$|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \leq C \left( (s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x} - \mathbf{x}') \right).$$

The flow,  $\boldsymbol{\theta}_{s,t}$  is, somehow, “almost” Lipschitz continuous in space w.r.t. the homogeneous distance  $\mathbf{d}$ , up to a time additive term. This time contribution is a consequence of the non-Lipschitz continuity of the drift  $\mathbf{F}$ . The analysis which was already done for  $\mathbf{F}$  Lipschitz continuous in Proposition 4.1 of [Men18], and in Chapter 6 with different Hölder regularity of  $\mathbf{F}$ . Actually, as we consider a smoother drift than in Chapter 6, the following lemma is a by-product of Lemma 11 therein. For the sake of completeness, we provide the analogous analysis below.

**Proof:** In the current analysis, we will assume that  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \geq 1$ , a general case can be directly deduced.

The analysis mainly come from a Grönwall’s Lemma. Nevertheless, because  $\mathbf{F}$  is not Lipschitz continuous, we have to mollify suitably the function. We denote by  $\delta \in \mathbb{R}^n$ , the vector entries  $\delta_i > 0$  for  $i \in \llbracket 2, n \rrbracket$ . and for all  $v \in [0, T]$ ,  $\mathbf{z} \in \mathbb{R}^{nd}$ ,  $i \in \llbracket 2, n \rrbracket$ ,

$$\mathbf{F}_i^\delta(v, \mathbf{z}^{i-1:n}) := \mathbf{F}_i(v, \cdot) \star \rho_{\delta_i}(\mathbf{z}) = \int_{\mathbb{R}^d} \mathbf{F}_i(v, \mathbf{z}_{i-1}, \mathbf{z}_i - w, \mathbf{z}_{i+1} \cdots, \mathbf{z}_n) \rho_{\delta_i}(w) dw, \quad (4.1)$$

with  $\rho_{\delta_i}(w) := \frac{1}{\delta_i^d} \rho\left(\frac{w}{\delta_i}\right)$  where  $\rho : \mathbb{R}^d \rightarrow \mathbb{R}_+$  is a usual mollifier, namely  $\rho$  has compact support and  $\int_{\mathbb{R}^d} \rho(\mathbf{z}) d\mathbf{z} = 1$ . Eventually, we write  $\mathbf{F}^\delta(v, \mathbf{z}) := (\mathbf{F}_1(v, \mathbf{z}), \mathbf{F}_2^\delta(v, \mathbf{z}), \dots, \mathbf{F}_n^\delta(v, \mathbf{z}))$ . With a slight abuse of notation in the previous definitions, since the first component  $\mathbf{F}_1$  is not mollified. The sublinearity of  $\mathbf{F}_1$  is actually enough to obtain the aimed control.

To be at the *good* current time scale for the contributions associated with the mollification, we pick  $\delta_i$  in order to have  $C := C((\mathbf{A}), T) > 0$  s.t. for all  $\mathbf{z} \in \mathbb{R}^{nd}$ ,  $u \in [t, s]$ :

$$\left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left( \mathbf{F}(u, \mathbf{z}) - \mathbf{F}^\delta(u, \mathbf{z}) \right) \right| \leq C(s-t)^{-1}. \quad (4.2)$$

By the previous definition of  $\mathbf{F}^\delta$ , identity (4.2) is equivalent to:

$$\sum_{i=2}^n (s-t)^{\frac{1}{2}-i} \delta_i^{\frac{2i-2}{2i-1}} \leq C(s-t)^{-1}. \quad (4.3)$$

Hence, we choose from now on, for all  $i \in \llbracket 2, n \rrbracket$ :

$$\delta_i = (s-t)^{(i-\frac{3}{2})\frac{2i-1}{2i-2}}. \quad (4.4)$$

Next, let us control the last components of the flow. By the definition of  $\boldsymbol{\theta}_{s,t}$  in (2.4), we get:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\ \leq & |(\mathbf{x} - \mathbf{x}')_n| + \int_t^s \left( |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| \right. \\ & \left. + |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))| + |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}')) - \mathbf{F}_n(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| \right) dv \\ \leq & |(\mathbf{x} - \mathbf{x}')_n| + C \int_t^s \left( |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| + \delta_n^{-1+\frac{2n-2}{2n-1}} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_n| \right) dv \\ & + (s-t) \delta_n^{\frac{2n-2}{2n-1}}. \end{aligned}$$

Hence by Grönwall's Lemma, we get:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\ \leq & C \exp \left( C(s-t) \delta_n^{-1+\frac{2n-2}{2n-1}} \right) \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t) \delta_n^{\frac{2n-2}{2n-1}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| dv \right) \\ \leq & C \exp \left( C(s-t)^{\frac{1}{2}} \right) \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| dv \right), \quad (4.5) \end{aligned}$$

using (4.3) for the last inequality. For the  $(n-1)$ <sup>th</sup> component, the situation is quite different in the sense that we have to handle the non-Lipschitz continuity of  $\mathbf{F}_{n-1}^\delta$  in its  $n$ <sup>th</sup> variable.

This complication for the rescaled flows could also be avoided through mollification, see e.g. Lemma 2 in [CdRM17]. But in our framework, it seems that Young type controls are more suitable. Indeed, we write:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_{n-1}| \\ \leq & C \exp \left( C(s-t) \delta_{n-1}^{-1+\frac{2(n-1)-2}{2(n-1)-1}} \right) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t) \delta_{n-1}^{\frac{2(n-1)-2}{2(n-1)-1}} \right. \\ & \left. + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| + |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_n|^{\frac{2(n-1)-2}{2(n-1)-1}} \right\} dv \right) \\ \leq & C \exp(C(s-t)^{\frac{1}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| \right. \right. \\ & \left. \left. + \underbrace{|(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-2}{2(n-1)-1}}}_{\leq |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-2}{2n-1}}} + (v-t)^{\frac{2(n-1)-2}{2}} + \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\frac{2(n-1)-2}{2(n-1)-1}} \right\} dv \right). \quad (4.6) \end{aligned}$$

The last inequality is a consequence of the assumption  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$  and by our pick of  $\delta_{n-1}$  in (4.4), by inequality (4.5) and by convexity inequality.

For all  $\tilde{\delta}_{n-1,n} > 0$  (where the two indexes in the subscript respectively denote the level of the chain, i.e.  $n-1$ , and the considered variable, i.e.  $n$ ), we get by Young inequality:

$$\begin{aligned} & \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\frac{2(n-1)-2}{2(n-1)-1}} \\ & \leq C \left( \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\tilde{\delta}_{n-1,n}^{-\frac{2(n-1)-1}{2(n-1)-2}}} + \tilde{\delta}_{n-1,n}^{\frac{2(n-1)-1}{3}} \right). \end{aligned}$$

In order to obtain the suitable time scale, we choose  $\tilde{\delta}_n$  s.t.

$$\tilde{\delta}_{n-1,n}^{\frac{2(n-1)-1}{3}} = (v-t)^{\frac{2(n-1)-2}{2}} \iff \tilde{\delta}_{n-1,n} = (v-t)^{\frac{3(2(n-1)-2)}{2(2(n-1)-1)}},$$

which also implies that

$$\left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\tilde{\delta}_{n-1,n}^{-\frac{2(n-1)-1}{2(n-1)-2}}} \leq \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right) (v-t)^{-\frac{3}{2}}.$$

From (4.6) and the previous controls, we deduce that for all  $\tilde{s} \in [t, s]$ :

$$\begin{aligned} & |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp(C(\tilde{s}-t)^{\frac{1}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (\tilde{s}-t)^{n-\frac{3}{2}} + \int_t^{\tilde{s}} \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| \right. \right. \\ & \quad \left. \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-2}{2n-1}} + (v-t)^{\frac{2(n-1)-2}{2}} \right. \right. \\ & \quad \left. \left. + \left( \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right) (v-t)^{-\frac{3}{2}} \right\} dv \right) \end{aligned}$$

Taking then the supremum in  $\tilde{s} \in [s, t]$  in the above equation, we obtain:

$$\begin{aligned} & \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp(C(s-t)^{\frac{1}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-2}{2n-1}} (s-t) + (s-t)^{1+\frac{2(n-1)-2}{2}} \right. \\ & \quad \left. + \left( \int_t^s \sup_{w \in [t, v]} |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| (v-t)^{1-\frac{3}{2}} dv \right) \right). \end{aligned}$$

Next, Grönwall's Lemma yields:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp(C(s-t)^{\frac{1}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-2}{2n-1}} (s-t) \right) \\ & \leq C \exp(C(s-t)^{\frac{1}{2}}) \left( |(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} \right). \end{aligned} \tag{4.7}$$

Where we used for the last identity the following Young inequality :  $|(\mathbf{x} - \mathbf{x}')_n|^{\frac{2(n-1)-2}{2n-1}}(s-t) \leq C(|(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1} \frac{2n-3}{2n-5}} + (s-t)^{n-\frac{3}{2}})$ .

We explicitly see from (4.7) that each entry of the difference of the starting points appears at its intrinsic scale for the homogeneous distance  $\mathbf{d}$ .

Plugging the above inequality into (4.5) we derive:

$$\begin{aligned} |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| &\leq C \exp\left(C(s-t)^{\frac{1}{2}}\right) \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + |(\mathbf{x} - \mathbf{x}')_{n-1}|(s-t) \right. \\ &\quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}}(s-t) + \int_t^s \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-2}| dw dv \right) \\ &\leq C \exp\left(C(s-t)^{\frac{1}{2}}\right) \left( |(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + |(\mathbf{x} - \mathbf{x}')_{n-1}|^{\frac{2n-1}{2n-3}} \right. \\ &\quad \left. + \int_t^s \int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-2}| dw dv \right), \end{aligned}$$

using again the Young inequalities  $|(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}}(s-t) \leq C(|(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}})$  and  $|(\mathbf{x} - \mathbf{x}')_{n-1}|(s-t) \leq C(|(\mathbf{x} - \mathbf{x}')_{n-1}|^{\frac{2n-1}{2n-3}} + (s-t)^{n-\frac{1}{2}})$  for the last inequality. We iterate the procedure, hence, we get:

$$\begin{aligned} &|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \tag{4.8} \\ &\leq C \left( (s-t)^{n-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x} - \mathbf{x}')_j|^{\frac{2j-1}{2j-1}} + \int_t^{v_n=s} dv_{n-1} \dots \int_t^{v_2} dv_1 |(\boldsymbol{\theta}_{v_1,t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1,t}(\mathbf{x}))_1| \right). \end{aligned}$$

Analogously, for  $i \in \llbracket 2, n \rrbracket$ , we obtain:

$$\begin{aligned} &|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \tag{4.9} \\ &\leq C \left( (s-t)^{i-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x} - \mathbf{x}')_j|^{\frac{2j-1}{2j-1}} + \int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_2} dv_1 |(\boldsymbol{\theta}_{v_1,t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1,t}(\mathbf{x}))_1| \right). \end{aligned}$$

**Remark 3** Observe that equations (4.8) and (4.9) are available for any fixed time  $s$  and  $v \in [t, T]$ .

The first term, i.e. for  $i = 1$  is controled slightly differently. In other words, for all  $\tilde{s} \in [t, s]$ , write:

$$|(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \leq |(\mathbf{x} - \mathbf{x}')_1| + C \sum_{j=1}^n \int_t^{\tilde{s}} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{1}{2j-1}} dv,$$

which in turn implies

$$\begin{aligned} &\sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \\ &\leq |(\mathbf{x} - \mathbf{x}')_1| + C \left( (s-t) \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| + \sum_{j=2}^n \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{1}{2j-1}} dv \right) \\ &\leq |(\mathbf{x} - \mathbf{x}')_1| + C \left( (s-t) \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_1| \right. \\ &\quad \left. + \sum_{j=2}^n C_j (s-t) \left( (s-t)^{j-\frac{1}{2}} + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{2j-1}{2k-1}} + (s-t)^{j-1} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_1| \right)^{\frac{1}{2j-1}} \right), \end{aligned}$$

which yields that

$$\begin{aligned} \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_1| &\leq |(\mathbf{x} - \mathbf{x}')_1| + C \left( (s-t) \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_1| \right. \\ &\quad + \sum_{j=2}^n C_j (s-t) \left( (s-t)^{\frac{1}{2}} + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}} \right. \\ &\quad \left. \left. + (s-t)^{(j-1)\frac{1}{2j-1}} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|^{\frac{1}{2j-1}} \right) \right). \end{aligned}$$

The last inequality is a consequence of (4.9), Remark 3 and convexity inequalities. Next, let us write,

$$\begin{aligned} &\sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_1| \\ &\leq C \left( |(\mathbf{x} - \mathbf{x}')_1| + (s-t)^{1+\frac{1}{2}} + (s-t) \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}} \right. \\ &\quad \left. + \sum_{j=1}^n (s-t)^{1+(j-1)\frac{1}{2j-1}} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|^{\frac{1}{2j-1}} \right) \\ &\leq C (|(\mathbf{x} - \mathbf{x}')_1| + (s-t) + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}}), \end{aligned} \tag{4.10}$$

recalling that  $(s-t) \leq T < 1$ , and using again Young inequalities for the last bound. Namely,

$$(s-t)^{1+(j-1)\frac{1}{2j-1}} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|^{\frac{1}{2j-1}} \leq C(s-t) \left( 1 + \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1| \right),$$

and

$$(s-t) |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}} \leq |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}}.$$

We eventually derive from (4.10) that:

$$\sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1| \leq C \left( (s-t)^{\frac{1}{2}} + \mathbf{d}(\mathbf{x}, \mathbf{x}') \right),$$

which gives the statement for  $i = 1$ . Now, by plugging this inequality into (4.9), we get for all  $i \in \llbracket 2, n \rrbracket$ :

$$\begin{aligned} &|(\boldsymbol{\theta}_{s, t}(\mathbf{x}) - \boldsymbol{\theta}_{s, t}(\mathbf{x}'))_i| \\ &\leq C \left( (s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (s-t)^{i-1} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1| \right) \\ &\leq C \left( (s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (s-t)^{i-1} \left( (s-t)^{\frac{1}{2}} + \mathbf{d}(\mathbf{x}, \mathbf{x}') \right) \right) \\ &\leq C \left( (s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \right), \end{aligned}$$

for the last identity, we use again the Young inequality to derive that  $(s-t)^{i-1} \mathbf{d}(\mathbf{x}, \mathbf{x}') \leq C \left( (s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \right)$ . The proof is complete.  $\square$

Now, we have the tools to control the sensitivity of the mean of the linearized system w.r.t. the freezing parameter.

**Sensitivity results for the mean**
**Proof: Technical Lemma 6**

Again through the analysis, we assume w.l.o.g. that  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$ . The control is done with a distinction of two contributions to handle.

$$\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}') = [\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x})] + [\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}')]. \quad (4.11)$$

By the proxy definition in (2.3), we deduce that the mean value of  $\tilde{\mathbf{X}}_v^{m,\xi}$ ,  $\mathbf{m}_{v,t}^{\xi}$  is s.t.

$$\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}) = \mathbf{x}' - \mathbf{x} + \int_t^s dv D\mathbf{F}(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{v,t}(\mathbf{x})]. \quad (4.12)$$

The sub-triangular structure of  $D\mathbf{F}$  yields that for all  $i \in \llbracket 2, n \rrbracket$ :

$$(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_i = \mathbf{x}'_i - \mathbf{x}_i + \int_t^s dv D_{i-1}\mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_{i-1} - \boldsymbol{\theta}_{v,t}(\mathbf{x})_{i-1}].$$

Also, since  $\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_1 = \mathbf{x}'_1 + \int_t^s \mathbf{F}_1(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))dv$ , so we obtain that  $[\mathbf{m}_{v,t}^{\mathbf{x}}(\mathbf{x}')_1 - \boldsymbol{\theta}_{v,t}(\mathbf{x})_1] = \mathbf{x}'_1 - \mathbf{x}_1$ , we then obtain by iteration that:

$$\begin{aligned} & (\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_i \\ &= \mathbf{x}'_i - \mathbf{x}_i + \sum_{k=2}^i \left[ \int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^i D_{j-1}\mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})) \right] [\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}], \end{aligned}$$

with the convention that for  $i = 1$ ,  $\sum_{k=2}^i = 0$ . From the above control, equation (4.11) and the dynamics of the flow, and because the starting points are the same, the contributions involving differences of the spatial points ( $\mathbf{x}' - \mathbf{x}$ ) or flows only appear in iterated time integrals, we obtain:

$$\begin{aligned} & |(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \\ &\leq \left| \sum_{k=2}^i \left[ \int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^i D_{j-1}\mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})) \right] [\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}] \right| \\ &\quad + \int_t^s |\mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| dv \\ &\leq C \left( \sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| \right. \\ &\quad \left. + \int_t^s \left( \sum_{j=i}^n |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\frac{2j-2}{2j-1}} + |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{i-1}| \right) dv \right). \end{aligned}$$

We derive from the previous Lemma 7:

$$\begin{aligned} & |(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \\ &\leq C \left( \sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| + (s-t)^{\frac{2i-2}{2}+1} \right. \\ &\quad \left. + \mathbf{d}^{2i-2}(\mathbf{x}, \mathbf{x}')(s-t) + ((s-t)^{(i-1)-\frac{1}{2}} + \mathbf{d}^{2(i-1)-1}(\mathbf{x}, \mathbf{x}'))(s-t) \right). \end{aligned}$$

In particular, for  $s = t_0 = t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$  with  $c_0 < 1$ , the previous equation yields:

$$\begin{aligned} & |(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_i| \\ & \leq C \left( c_0 \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (c_0^{2i-1+\gamma} + c_0) \mathbf{d}^{2i-1+\gamma}(\mathbf{x}, \mathbf{x}') + (c_0^{i-\frac{1}{2}} + c_0) \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \right). \end{aligned}$$

Still for  $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$ , after summing and by convexity inequalities, we eventually deduce:

$$\mathbf{d}(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}'), \boldsymbol{\theta}_{s,t}(\mathbf{x}')) \leq C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}').$$

□

## 5 Appendix: Some reminders about Besov spaces

### 5.1 Thermic characteristic of the Besov space

Let us now recall some definitions/characterizations from Section 2.6.4 of Triebel [Tri83]. For  $\alpha \in \mathbb{R}, q \in (0, +\infty], p \in (0, \infty], B_{p,q}^\alpha(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{\mathcal{H}_{p,q}^\alpha} < +\infty\}$  where  $\mathcal{S}(\mathbb{R}^d)$  stands for the Schwartz class and

$$\|f\|_{\mathcal{H}_{p,q}^\alpha} := \|\varphi(D)f\|_{L^p(\mathbb{R}^d)} + \left( \int_0^1 v^{(m-\frac{\alpha}{2})q} \|\partial_v^m h_v \star f\|_{L^p(\mathbb{R}^d)}^q \frac{dv}{v} \right)^{\frac{1}{q}}, \quad (5.1)$$

with  $\varphi \in C_0^\infty(\mathbb{R}^d)$  (smooth function with compact support) is s.t.  $\varphi(0) \neq 0$ ,  $\varphi(D)f := (\varphi \hat{f})^\vee$  where  $\hat{f}$  and  $(\varphi \hat{f})^\vee$  respectively denote the Fourier transform of  $f$  and the inverse Fourier transform of  $\varphi \hat{f}$ . The parameter  $m$  is an integer s.t.  $m > \frac{\alpha}{2}$  and for  $v > 0, z \in \mathbb{R}^d$ ,  $h_v(z) := \frac{1}{(2\pi v)^{\frac{d}{2}}} \exp\left(-\frac{|z|^2}{2v}\right)$  is the usual heat kernel of  $\mathbb{R}^d$ . We point out that the quantities in (5.1) are well defined for  $q < \infty$ . The modifications for  $q = +\infty$  are obvious and can be written passing to the limit.

Observe that the quantity  $\|f\|_{\mathcal{H}_{p,q}^s}$ , where the subscript  $\mathcal{H}$  stands to indicate the dependence on the heat-kernel, depends on the considered function  $\varphi$  and the chosen  $m \in \mathbb{N}$ . It also defines a quasi-norm on  $B_{p,q}^s(\mathbb{R}^d)$ . The previous definition of  $B_{p,q}^\alpha(\mathbb{R}^d)$  is known as the thermic characterization of Besov spaces and is particularly well adapted to our current framework. By abuse of notation we will write as soon as this quantity is finite  $\|f\|_{\mathcal{H}_{p,q}^\alpha} =: \|f\|_{B_{p,q}^\alpha}$ .

### 5.2 Equivalence of Besov norms

**Proposition 4** *There is a constant  $C > 1$  such that for all  $f \in B_{\infty,\infty}^\alpha(\mathbb{R}^d)$ ,  $\alpha \in (0, 1)$  we have:*

$$C^{-1} \|f\|_{B_{\infty,\infty}^\alpha} \leq \|\mathbf{D}f\|_{B_{\infty,\infty}^{\alpha-1}} \leq C \|f\|_{B_{\infty,\infty}^\alpha}.$$

We write also that  $\|f\|_{B_{\infty,\infty}^\alpha} \asymp \|\mathbf{D}f\|_{B_{\infty,\infty}^{\alpha-1}}$ .



**Proof: Proposition 4**

To establish this equivalence of norms we use the local means characterisation of the Besov space, see for instance 2.5.3 in [Tri83]. In other words, for all function  $k, k_0, k^0 \in C^\infty(\mathbb{R}^d)$  (infinitely differentiable) such that  $\text{supp}(k), \text{supp}(k_0), \text{supp}(k^0) \subset \mathbf{B}(\mathbf{0}_{d,d}, 1)$  (the unit ball centered at the origin) and for all  $x \in \mathbb{R}^d$ :

$$k(x) = \sum_{i=1}^d \partial_{x_i}^2 k^0(x).$$

We define furthermore the local means for all  $(x, t) \in \mathbb{R}^d \times (0, +\infty)$ :

$$k(t, f)(x) := \int_{\mathbb{R}^d} k(y) f(x + ty) dy, \quad k_0(t, f)(x) := \int_{\mathbb{R}^d} k_0(y) f(x + ty) dy.$$

It is then known that for all  $f \in B_{p,q}^\alpha(\mathbb{R}^d)$ :

$$\|f\|_{B_{\infty,\infty}^\alpha} \asymp \|k_0(1, f)(\cdot)\|_\infty + \sup_{t \in (0,1)} (t^{-\alpha} \|k(t, f)(\cdot)\|_\infty).$$

We can consider similarly the sequences of functions,  $k^i = \partial_{x_i} k, k_0^i = \partial_{x_i} k_0, i \in \llbracket 1, d \rrbracket$ , which satisfy the same assumptions as  $k$  and  $k_0$ . Hence, we can write:

$$\begin{aligned} \|f\|_{B_{\infty,\infty}^\alpha} &\asymp \|k_0^i(1, f)(\cdot)\|_\infty + \sup_{t \in (0,1)} (t^{-\alpha} \|k^i(t, f)(\cdot)\|_\infty) \\ &\asymp \|k_0(1, \partial_{x_i} f)(\cdot)\|_\infty + \sup_{t \in (0,1)} (t^{1-\alpha} \|k(t, \partial_{x_i} f)(\cdot)\|_\infty) \\ &\asymp \|\partial_{x_i} f\|_{B_{\infty,\infty}^{\alpha-1}}. \end{aligned}$$

The penultimate identity is a consequence of an integration by parts over the convolution products, and the last equivalence is obtained again thanks to Theorem 2.5.3 in [Tri83]. Iterating over each component yields the result. □

## Bibliography

- [ACJ08] A. Arnold, A. Carlen, and Q. Ju. Large-time behavior of non-symmetric Fokker-Planck type equations. *Comm. Stoch. Analysis*, 2-1:153–175, 2008.
- [AH96] D. R. Adams and L. I. Hedberg. *Function spaces and Potential Theory*. Springer, 1996.
- [AMTU01] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Partial Diff. Equations*, 26-(1-2):43–100, 2001.
- [App09] D. Appelbaum. *Lévy processes and stochastic calculus*. Cambridge University Press, 2009.
- [Azu67] K. Azuma. Weighted sums of certain dependent random variables. *Tohoku Math. J. (2)*, 19(3):357–367, 1967.
- [Bas97] R. F. Bass. *Diffusions and Elliptic Operators*. Springer, 1997.
- [BB06] G. Blower and F. Bolley. Concentration inequalities on product spaces with applications to Markov processes. *Studia Mathematica*, 175-1:47–72, 2006.
- [BC03] R. Bass and Z. Q. Chen. Brownian motion with singular drift. *Ann. Probab.*, 31- 2:791–817, 2003.
- [BCLP10] M. Bramanti, G. Cupini, E. Lanconelli, and E Priola. Global  $L^p$  estimates for degenerate Ornstein-Uhlenbeck operators. *Math. Z.*, 266(4):789–816, 2010.
- [BCLP13] M. Bramanti, G. Cupini, E. Lanconelli, and E. Priola. Global  $L^p$  estimates for degenerate Ornstein-Uhlenbeck operators with variable coefficients. *Math. Nachr.*, 286(11-12):1087–1101, 2013.
- [BE85] D. Bakry and M. Émery. Diffusions hypercontractives. *Séminaire de probabilités*, XIX:177–206, 1985.
- [BG10] F. Bolley and I. Gentil. Phi-entropy inequalities for diffusion semigroups. *Journal de Mathématiques Pures et Appliquées*, 93(5):449 – 473, 2010.
- [BGL14] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.
- [BGV07] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Prob. Th. Rel. Fields*, 137:541–593, 2007.

- [Bha82] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrsch. Verw. Gebiete*, 60(2):185–201, 1982.
- [BHW97] G.K. Basak, I. Hu, and C.-Z. Wei. Weak convergence of recursions. *Stochastic Processes and their Applications*, 68(1):65 – 82, 1997.
- [BLG14] E. Boissard and T. Le Gouic. On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 50(2):539–563, 2014.
- [BMPT09] A. Bensoussan, L. Mertz, O. Pironneau, and Janos Turi. An ultra weak finite element method as an alternative to a Monte Carlo method for an elasto-plastic problem with noise. *SIAM J. Numer. Anal.*, 47(5):3374–3396, 2009.
- [Boc13] S. Boccia. Schauder estimates for solutions of higher-order parabolic systems. *Methods Appl. Anal.*, 20(1):47–67, 2013.
- [Boi11] E. Boissard. Simple bounds for the convergence of empirical and occupation measures in 1-Wasserstein distance. *Electronic Journal of Probability*, 16, 2011.
- [BPV01] E. Barucci, S. Polidoro, and V. Vespri. Some results on partial differential equations and asian options. *Math. Models Methods Appl. Sci*, 3:475–497, 2001.
- [Bro88] G. A. Brosamler. An almost everywhere central limit theorem. *Math. Proc. Cambridge Philos. Soc.*, 104(3):561–574, 1988.
- [BT96a] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function. *Prob. Th. Rel. Fields*, 104-1:43–60, 1996.
- [BT96b] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations, II. Convergence rate of the density. *Monte-Carlo methods and Appl.*, 2:93–128, 1996.
- [BT08] A. Bensoussan and J. Turi. Degenerate Dirichlet problems related to the invariant measure of elasto-plastic oscillators. *Appl. Math. Optim.*, 58(1):1–27, 2008.
- [BZ11] M. Bramanti and M. Zhu.  $L^p$  and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift. *Analysis of PDEs*, 2011.
- [CdR17] P.-E. Chaudru de Raynal. Strong existence and uniqueness for degenerate SDE with Hölder drift. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 53(1):259–286, February 2017.
- [CdR18] P.-E. Chaudru de Raynal. Weak regularization by stochastic drift: result and counter example. *Discrete & Continuous Dynamical Systems - A*, 38:1269, 2018.
- [CdRM17] P-E. Chaudru de Raynal and S. Menozzi. Regularization effects of a noise propagating through a chain of differential equations: an almost sharp result. to appear in trans. american math. society, October 2017.

- [CG16] R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of ODEs. *Stoch. Proc and Appl.*, 126–8:2323–2366, 2016.
- [CH92] M. Csorgo and L. Horvath. Invariance principles for logarithmic averages. *Math. Proc. Cambridge Philos. Soc.*, 112(1):195–205, 1992.
- [Che05] A. S. Cherny. *Singular Stochastic Differential Equations*. Springer, January 2005.
- [CM00] F. Chaabane and F. Maaouia. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM Probab. Statist.*, 4:137–189 (electronic), 2000.
- [DF14] F. Delarue and F. Flandoli. The transition point in the zero noise limit for a 1d Peano example. *Discrete and Continuous Dynamical Systems*, 34(10):4071–4083, April 2014.
- [DG15] J. Dedecker and S. Gouëzel. Subgaussian concentration inequalities for geometrically ergodic markov chains. *Electronic Communications in Probability*, 20, Article 64:1–12, 2015.
- [DM10] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *Journal of Functional Analysis*, 259–6:1577–1630, 2010.
- [DPG01] G. Da Prato and B. Goldys. Elliptic operators on rd with unbounded coefficients. *Journal of Differential Equations*, 172(2):333 – 358, 2001.
- [DS58] N. Dunford and J.T. Schwartz. *Linear Operators*. Wiley-Interscience, New-York, 1958.
- [Duf90] M. Duflo. *Méthodes récursives aléatoires*. Techniques Stochastiques. [Stochastic Techniques]. Masson, Paris, 1990.
- [EH00] J.-P. Eckmann and M. Hairer. Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. *Communications in Mathematical Physics*, 212(1):105–164, Jun 2000.
- [EK86] E. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. Wiley, 1986.
- [FF11] E. Fedrizzi and F. Flandoli. Pathwise Uniqueness and Continuous Dependence for SDEs with Nonregular Drift. *Stochastics*, 83-3:241–257, 2011.
- [FFPV17] E. Fedrizzi, F. Flandoli, E. Priola, and J. Vovelle. Regularity of stochastic kinetic equations. *Electronic Journal of Probability*, 22:# 48, 2017.
- [Fla11] F. Flandoli. *Random perturbation of PDEs and fluid dynamic models*, volume 2015 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010.
- [FM12] N. Frikha and S. Menozzi. Concentration bounds for stochastic approximations. *Electron. Commun. Probab.*, 17:no. 47, 15, 2012.
- [FP06] M. Di Francesco and S. Polidoro. Schauder estimates, harnack inequality and gaussian lower bound for kolmogorov-type operators in non-divergence form. *Adv. Differential Equations*, 11:1261–1320, 2006.

- [FPS00] J. P. Fouque, G. Papanicolaou, and R. Sircar. *Derivatives in Financial Markets with Stochastic Volatility*. WILEY, 2000.
- [Fri64] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, 1964.
- [GHL18] A. Gloter, I. Honoré, and D. Loukianova. Non-asymptotic concentration inequality for an approximation of the invariant distribution of a diffusion driven by compound poisson process. working paper or preprint, October 2018.
- [GL78] M.I. Gordin and B.A. Lifsic. On the central limit theorem for stationary markov processes. *Soviet Math. Dokl.*, 19(2):392–394, 1978.
- [GL00] S. Graf and H. Lushgy. *Foundations of quantization for random vectors*. LNM-1730, Springer-Verlag, 2000.
- [GO13] M. Gradinaru and Y. Offret. Existence and asymptotic behaviour of some time-inhomogeneous diffusions. *Annales Instit. H. Poincaré, Probabilités et Statistiques*, 49(1):182–207, 2013.
- [GP14] S. Gadat and F. Panloup. Long time behaviour and stationary regime of memory gradient diffusions. *Ann. Inst. H. Poincaré Probab. Statist.*, 50(2):564–601, 05 2014.
- [Hec98] M. K. Heck. The principle of large deviations for the almost everywhere central limit theorem. *Stochastic Process. Appl.*, 76(1):61–75, 1998.
- [HH80] P. Hall and C.C. Heyde. *Martingale limit theory and its application*. Probability and mathematical statistics. Academic Press, 1980.
- [HM16] L. Huang and S. Menozzi. A Parametrix Approach for some Degenerate Stable Driven SDEs. *Annales Instit. H. Poincaré*, 52(4):1925–1975, 2016.
- [HMP18] I. Honoré, S. Menozzi, and G. Pagès. Non-Asymptotic Gaussian Estimates for the Recursive Approximation of the Invariant Measure of a Diffusion. working paper or preprint, May 2018.
- [Hon18] I. Honoré. Sharp non-asymptotic Concentration Inequalities for the Approximation of the Invariant Measure of a Diffusion. working paper or preprint, October 2018.
- [Hör67] L. Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119:147–171, 1967.
- [IKO62] A. M. Il'in, A. S. Kalashnikov, and O. A. Oleinik. Second-order linear equations of parabolic type. *Uspehi Mat. Nauk*, 17–3(105):3–146, 1962.
- [IW80] N. Ikeda and S. Watanabe. *Stochastic differential equations*. North Holland, 1980.
- [JO10] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.*, 38- 6:2418–2442, 2010.
- [JYC09] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer Finance. Springer London, 2009.
- [KCL75] S. N. Kružkov, A. Castro, and M. Lopes. Schauder type estimates, and theorems on the existence of the solution of fundamental problems for linear and nonlinear parabolic equations. *Dokl. Akad. Nauk SSSR*, 220:277–280, 1975.

- [Kha80] R.Z. Khasminskii. *Stochastic stability of differential equations*. Mechanics–analysis. Sijthoff & Noordhoff, 1980.
- [KLL08] M. Kunze, L. Lorenzi, and A. Lunardi. Nonautonomous Kolmogorov parabolic equations with unbounded coefficients. *Transactions of the American Mathematical Society*, April 2008.
- [KM00] V. Konakov and E. Mammen. Local limit theorems for transition densities of Markov chains converging to diffusions. *Prob. Th. Rel. Fields*, 117:551–587, 2000.
- [KM02] V. Konakov and E. Mammen. Edgeworth type expansions for euler schemes for stochastic differential equations. *Monte Carlo Methods Appl.*, 8–3:271–285, 2002.
- [KM11] R. Khasminskii and G.N. Milstein. *Stochastic Stability of Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2011.
- [Kol34] A. N. Kolmogorov. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math.*, 2-35:116–117, 1934.
- [Kol11] V. Kolokoltsov. *Markov Processes, Semigroups and Generators*. De Gruyter, 2011.
- [KP10] N. V. Krylov and E. Priola. Elliptic and parabolic second-order PDEs with growing coefficients. *Comm. Partial Differential Equations*, 35(1):1–22, 2010.
- [KP11] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2011.
- [KP17] N. V. Krylov and E. Priola. Poisson stochastic process and basic Schauder and Sobolev estimates in the theory of parabolic equations. *Arch. Ration. Mech. Anal.*, 225(3):1089–1126, 2017.
- [KR05] N. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Prob. Theory Rel. Fields*, 131:154–196, 2005.
- [Kry96] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*. Graduate Studies in Mathematics 12. AMS, 1996.
- [Kun97] H. Kunita. *Stochastic Flows and Stochastic Differential Equations*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1997.
- [Led99] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, Berlin, 1999.
- [Lem07] V. Lemaire. An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.*, 117(10):1491–1518, 2007.
- [Lor05a] L. Lorenzi. Estimates of the derivatives for a class of parabolic degenerate operators with unbounded coefficients in  $\mathbb{R}^N$ . *Annali della Scuola Normale Superiore di Pisa. Classe di Scienze. Serie V*, 2, 01 2005.

- [Lor05b] L. Lorenzi. Schauder estimates for degenerate elliptic and parabolic problems with unbounded coefficients in  $\mathbb{R}^n$ . *Differential Integral Equations*, 18–5:531–566, 2005.
- [Lor11] L. Lorenzi. Optimal Hölder regularity for nonautonomous Kolmogorov equations. *Discrete Contin. Dyn. Syst. Ser. S*, 4(1):169–191, 2011.
- [LP02] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8–3:367–405, 2002.
- [LP03] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion: the case of a weakly mean reverting drift. *Stoch. Dyn.*, 3(4):435–451, 2003.
- [LPP15] V. Lemaire, G. Pagès, and F. Panloup. Invariant measure of duplicated diffusions and application to richardson–romberg extrapolation. *Ann. Inst. H. Poincaré Probab. Statist.*, 51(4):1562–1596, 11 2015.
- [LR02] P.-G. Lemarie-Rieusset. *Recent developments in the Navier-Stokes problem*. CRC Press, 2002.
- [Lun95] A. Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Analytic Semigroups and Optimal Regularity in Parabolic Problems. Springer Basel, 1995.
- [Lun97] A. Lunardi. Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in  $\mathbb{R}^n$ . *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 24(1):133–164, 1997.
- [Mas07] H. Masuda. Ergodicity and exponential  $\beta$ -mixing bounds for multidimensional diffusions with jumps. *Stochastic Processes and their Applications*, 117(1):35–56, 2007.
- [Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electronic Communications in Probability*, 17:234–250, 2011.
- [Men18] S. Menozzi. Martingale problems for some degenerate Kolmogorov equations. *Stoc. Proc. Appl.*, 128-3:756–802, 2018.
- [MPR11] F. Merlevède, M. Peligrad, and E. Rio. A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probability Theory and Related Fields*, 151-3:435–474, 2011.
- [MPW02] G. Metafune, D. Pallara, and M. Wacker. Feller semigroups on  $\mathbb{R}^n$ . *Semigroup Forum*, 65:159–205, 07 2002.
- [MS67] H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1:43–69, 1967.
- [MSH02] J. Mattingly, A. Stuart, and D. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, 101–2:185–232, 2002.
- [MT06] F. Malrieu and D. Talay. Concentration inequalities for Euler Schemes. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pages 355–371. Springer Berlin Heidelberg, 2006.

- [MT09] S. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge Mathematical Library. Cambridge University Press, 2009.
- [Nel67] E. Nelson. *Dynamical theories of Brownian motion*. Princeton University Press, Princeton, N.J, 1967.
- [Pag97] G. Pagès. A space vector quantization method for numerical integration. *Journal of Applied and computational Mathematics*, 89:1–38, 1997.
- [Pag01] G. Pagès. Sur quelques algorithmes récurrents pour les probabilités numériques. *ESAIM: PS*, 5:141–170, 2001.
- [Pan08a] F. Panloup. Computation of the invariant measure of a levy driven SDE: Rate of convergence. *Stochastic processes and Applications*, 118–8:1351–1384, 2008.
- [Pan08b] F. Panloup. Recursive computation of the invariant measure of a stochastic differential equation driven by a lévy process. *Ann. Appl. Probab.*, 18(2):379–426, 04 2008.
- [PL89] R. Di Perna and P. L. Lions. Ordinary Differential Equations, transport theory and Sobolev Spaces. *Inv. Math.*, 98:511–547, 1989.
- [PP12] G. Pagès and F. Panloup. Ergodic approximation of the distribution of a stationary diffusion: rate of convergence. *Ann. Appl. Probab.*, 22(3):1059–1100, 2012.
- [PP14] G. Pagès and F. Panloup. A mixed-step algorithm for the approximation of the stationary regime of a diffusion. *Stochastic Processes and their Applications*, 124(1):522 – 565, 2014.
- [PR17] G. Pagès and C. Rey. Recursive computation of the invariant distribution of Markov and Feller processes. working paper or preprint, January 2017.
- [Pri09] E. Priola. Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.*, 194(2):117–153, 2009.
- [Pri10] E. Priola. Pathwise uniqueness for singular sdes driven by stable processes. *Osaka Journal of Mathematics*, 49, 05 2010.
- [Pri15] E. Priola. On weak uniqueness for some degenerate SDEs by global  $L^p$  estimates. *Potential Anal.*, 42(1):247–281, 2015.
- [PS94] M. Piccioni and S. Scarlatti. An iterative monte carlo scheme for generating lie group-valued random variables. *Advances in Applied Probability*, 26, 09 1994.
- [PT97] P. Protter and D. Talay. The Euler scheme for lévy driven stochastic differential equations. *Ann. Probab.*, 1(25):393–423, 1997.
- [PV01] E. Pardoux and A. Veretennikov. On the Poisson Equation and Diffusion Approximation. I. *Ann. Probab.*, 29–3:1061–1085, 2001.
- [PV03] E. Pardoux and A. Yu. Veretennikov. On poisson equation and diffusion approximation 2. *Ann. Probab.*, 31(3):1166–1192, 07 2003.
- [PV05] E. Pardoux and A. Yu. Veretennikov. On the poisson equation and diffusion approximation 3. *Ann. Probab.*, 33(3):1111–1133, 05 2005.



- [RBT00] L. Rey-Bellet and L. Thomas. Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. *Comm. Math. Phys.*, 215–1:1–24, 2000.
- [RS76] L. P. Rothschild and E. M. Stein. Hypocoercive differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.
- [Sch88] P. Schatte. On strong versions of the Central Limit Theorem. *Math. Nachr.*, 137:249–256, 1988.
- [Soi94] C. Soize. *The Fokker-Planck Equation for Stochastic Dynamical Systems and its Explicit Steady State Solutions*. Series on Advances in Mathematics for Applied Sciences. World Scientific Publishing Co Pte Ltd, Singapore, 1994.
- [Str08] D. W. Stroock. *Partial differential equations for probabilists*, volume 112 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008.
- [SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.
- [Tal02] D. Talay. Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8–2:163–198, 2002.
- [Tan79] H. Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Mathematical Journal*, 9(1):163–177, 1979.
- [Tri83] H. Triebel. *Theory of function spaces, II*. Birkhauser, 1983.
- [TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. and App.*, 8-4:94–120, 1990.
- [Ver80] A. Y. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Matematicheski Sbornik. Novaya Seriya*, 111(153)(3):434–452, 480, 1980.
- [Ver83] A. Y. Veretennikov. Stochastic equations with diffusion that degenerates with respect to part of the variables. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 47(1):189–196, 1983.
- [Vil09] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009.
- [WZ16] F.Y. Wang and X. Zhang. Degenerate SDE with Holder-Dini drift and non-Lipschitz noise coefficient. *SIAM J. Math. Anal.*, 48(3):2189–2226, 2016.
- [YW71] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11(1):155–167, 1971.
- [Zha10] X. Zhang. Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Revista Matemática Iberoamericana*, 29(1), 2010.
- [Zha18] X. Zhang. Stochastic hamiltonian flows with singular coefficients. *Science China Mathematics*, 61(8):1353–1384, Aug 2018.
- [Zvo74] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

**Title :** Non-asymptotic estimates of invariant measure and regularisation by a degenerate noise for a chain of Ordinary Differential Equations.

**Keys words :** Invariant distribution, Concentration inequalities, Poisson Equation, Kolmogorov Equation, Schauder estimates, Strong uniqueness.

**Abstract :** In the first part of this thesis, we aim to estimate the invariant distribution of an ergodic process driven by a Stochastic Differential Equation. The ergodic theorem suggests us to consider the empirical measure associated with a discretization scheme of the process which can be regarded as a discretization of the occupation measure of the process. In [LP02], Lamberton and Pagès introduced an algorithm of discretization with decreasing time steps which allows the convergence of the empirical measure toward the invariant distribution of the process, they also provide a central limit theorem (CLT) which asymptotically quantifies the deviations between these both measures. We establish non-asymptotic concentration inequality for the empirical measure deviations (in accordance with the previously mentioned CLT), and also we give some controls of the solution of the associated Poisson equation which is useful for this concentration inequalities.

In a second part, we establish some Schauder controls associated with parabolic equations related with a degenerate stochastic system, where the drift is a vector field satisfying a weak Hörmander condition like. But we aim to suppose only the minimal Hölder regularity. This work is an extension of the estimates given by Delarue and Menozzi [DM10]. Finally, our approach allows us to proof the strong uniqueness of the considered stochastic equation in a Hölder regularity framework. Our results extend the controls of Chaudru de Raynal [CdR17] for the dimension equal to 2.

**Titre :** Estimations non-asymptotiques de mesures invariantes et régularisation par un bruit dégénéré de chaînes d'Équations Différentielles Ordinaires.

**Mots Clefs :** Mesure invariante, Inégalités de concentration, Équation de Poisson, Équation de Kolomogorov, Estimées de Schauder, Unicité forte.

**Résumé :** Dans la première partie de cette thèse, nous chercherons à estimer la mesure invariante d'un processus ergodique dirigé par une Équation Différentielle Stochastique. Le théorème ergodique nous suggère de considérer la mesure empirique associée à un schéma d'approximation du processus sous-jacent qui peut se voir comme le pendant discret de la mesure d'occupation dudit processus. Dans [LP02], Lamberton et Pagès ont introduit un algorithme de discrétisation à pas décroissant qui assure la convergence de la mesure empirique du schéma vers la mesure invariante du processus considéré ainsi qu'un théorème central limite (TCL) quantifiant asymptotiquement l'écart entre ces deux mesures. Nous établissons des inégalités de concentration non-asymptotiques pour les déviations de la mesure empirique (cohérentes avec le TCL mentionné ci-avant), ainsi que des contrôles sur la solution de l'équation de Poisson associée, utiles pour ces inégalités. Dans une seconde partie, nous établissons des estimées de Schauder liées à des équations paraboliques associées à un système stochastique dégénéré, où la dérive est un champ de vecteurs vérifiant une condition de type Hörmander (faible) mais en cherchant la régularité Hölder minimale. Ce travail fait suite à l'article de Delarue et Menozzi [DM10]. Enfin, notre approche nous permet de montrer l'unicité forte du système stochastique considéré dans le cadre de coefficients Hölder, étendant ainsi le résultat obtenu en dimension 2 par Chaudru de Raynal [CdR17].