



HAL
open science

Ergodicity of fractional stochastic differential equations and related problems

Maylis Varvenne

► **To cite this version:**

Maylis Varvenne. Ergodicity of fractional stochastic differential equations and related problems. Analysis of PDEs [math.AP]. Université Paul Sabatier - Toulouse III, 2019. English. NNT : 2019TOU30046 . tel-02737000

HAL Id: tel-02737000

<https://theses.hal.science/tel-02737000>

Submitted on 2 Jun 2020

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

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *24 Juin 2019* par :

MAYLIS VARVENNE

**Ergodicité des équations différentielles stochastiques fractionnaires
et problèmes liés**

JURY

PATRICK CATTIAUX	Université de Toulouse 3	Examineur
LAURE COUTIN	Université de Toulouse 3	Directrice de thèse
ARNAUD GUILLIN	Université de Clermont Auvergne	Rapporteur
EULALIA NUALART	Université Pompeu Fabra de Barcelone	Examinatrice
FABIEN PANLOUP	Université d'Angers	Directeur de thèse
SEBASTIAN RIEDEL	Université de Berlin	Rapporteur

École doctorale et spécialité :

MITT : Domaine Mathématiques : Mathématiques appliquées

Unité de Recherche :

Institut de Mathématiques de Toulouse (UMR 5219)

Directeur(s) de Thèse :

Laure Coutin et Fabien Panloup

Rapporteurs :

Arnaud Guillin et Sebastian Riedel

Remerciements

Mes remerciements s'adressent bien sûr en tout premier lieu à mes directeurs de thèse Laure Coutin et Fabien Panloup. Fabien, qui m'a guidée pour mes premiers pas dans la recherche dès le stage de M2 et Laure qui nous a rejoint pour la grande aventure qu'est la thèse. Malgré la distance lorsqu'il est parti à Angers, Fabien a toujours été très présent tout au long de ma thèse y compris dans les moments où le stress et les doutes m'envahissaient. En particulier, lors de l'ascenseur émotionnel le plus spectaculaire de cette thèse juste avant l'envoi du manuscrit aux rapporteurs : erreur fatale dans BDG (panique totale), résolution en deux jours (soulagement intense) et modification de dernière minute du manuscrit (ouf !). Laure, toute aussi attentionnée, m'a bien sûr épaulée dans mes questionnements de recherche, mais a aussi toujours pris du temps pour savoir si mon moral allait bien, savoir comment j'envisageais la suite de la thèse et m'aider à trouver des solutions. Laure et Fabien, je vous remercie du fond du cœur pour tout ce que vous m'avez apporté.

J'exprime ensuite mes plus sincères remerciements à Arnaud Guillin et Sebastian Riedel d'avoir pris de leur temps pour lire et évaluer mon travail. Je remercie également Patrick Cattiaux et Eulalia Nualart d'avoir accepté d'être membres de mon jury de thèse.

Je tiens à exprimer ma gratitude à Joseph Najnudel et Ju-Yi Yen qui nous ont très bien accueillis avec Valentin lors de notre visite de deux mois à Cincinnati : d'une part au sein de l'Université de Cincinnati où l'on a pu travailler dans d'excellentes conditions et d'autre part pour la visite de la ville ! Bien sûr, je voudrais remercier Samy Tindel que j'ai eu la chance d'aller voir deux fois à West Lafayette au cours de ce séjour américain. C'était le début d'une collaboration scientifique très enrichissante et la découverte d'une belle personne !

Durant ces trois années, je me suis épanouie au sein de l'IMT et je tiens à remercier quelques personnes en particulier. La quotidienne pause café du midi dans le bâtiment 1R1 m'aura permis de passer de chouettes moments de convivialité avec bon nombre de membres permanents du laboratoire : en particulier Reda pour qui la salle café semble être le bureau côté ESP ; Guillaume qui pourra vous expliquer ce que sont les œufs en meurette, spécialité bourguignonne (si comme moi vous êtes ignorants à ce sujet) ; Sébastien Gerchinovitz, François B., Clément P., Sébastien Gadat, Manon, Tristan et encore bien d'autres... Je suis obligée de citer aussi Pierre Petit, notre « dealer » attitré en café sans qui cette salle n'aurait plus beaucoup de sens, merci ! Côté enseignement, je

souhaite remercier François Costantino avec qui j'ai eu la chance de collaborer pendant trois ans, il est toujours à l'écoute de nos avis et soucieux d'améliorer chaque année avec nous la qualité de nos enseignements. De nouveau, je pense à Patrick Cattiaux qui m'a très bien accueillie dès mon M2 et a accepté d'être mon tuteur pédagogique pour ces trois années de thèse. Merci également à Marie-Line qui serait prête à réserver l'amphi Schwartz pendant tout le mois de juin pour ma soutenance si je le lui demandais... Elle est vraiment aux petits soins pour nous, merci beaucoup ! Je ne peux parler de Marie-Line sans parler de Monique, toujours avec le sourire et un mot gentil le matin, merci à vous. Pour tout ce qui est logistique de mission, on peut compter sur Françoise qui prend le temps de tout nous expliquer avec patience. Je remercie aussi Agnès et Martine de l'école doctorale qui nous guident dans les démarches administratives tout au long de la thèse et c'est loin d'être trivial. Merci à tout le personnel administratif de l'IMT.

Il est temps maintenant de m'attarder sur la grande équipe des doctorants de l'IMT, merci à tous. Je commence évidemment par mes co-bureaux : An et Mickael (ou Kiki pour les intimes), président de l'UT4 et organisateur de soirées : « ça vous dit une raclette cette semaine ? - Mais on est en plein mois d'août ! - Et alors ??? ». Merci à tous les participants du célèbre championnat de coinche de l'IMT qui a malheureusement disparu suite à vos départs successifs : Anton, Aurore, Kevin, Théo, Ioana et Guillaume (#teamzérodech). Merci à l'ensemble du « grand cru 2016 » : Baptiste, ESP par adoption, notre Ch'ti roux « mouque tin nez ta le brin al botine » (c'est bien ça?) ; Trang, heureusement que tu as appris le français pour que notre communication ne se limite plus à se sourire ; William, le Booba des maths comme dit Valentin, ça te définit tellement bien ; Fabien, étudiant sans argent obligé de vivre au tripode qui ramène des petits beurrés et sa bonne humeur lors des soirées (« c'est un tétard qui croyait qu'il était tôt mais en fait... »). Merci au « cru 2015 » : Pière et le Rio Loco mémorable (KENYYYYY!!!) ; Hugo, le « concierge » de l'IMT (adieu les potins depuis que tu es parti) et sa fourgonnette de soirées ; Antoine, mon grand frère de thèse, expert dans l'art de me faire flipper par un simple « Salut ! » lancé dans mon dos après une entrée sans bruit dans mon bureau. Je remercie les plus « anciens » également : Fanny, Claire, Sofiane (mon autre grand frère de thèse), Kevin, Malika. Merci bien sûr à la jeune génération ESP : Eva, Camille, José, Clément (attention je te confie la survie du séminaire étudiant côté ESP!), et la toute dernière arrivée : Laetitia. Enfin, je n'oublie pas les MIP (grâce à Baptiste qui est un peu le ciment entre nos deux familles) : Anthony, Alexis, Joachim, Susely, Mathias, Michèle (grande joueuse au sein de l'IMTFC créé par Sebastien D. et Marcello, merci à eux !).

Bien sûr je ne serais pas arrivée jusque là sans l'aide de mes nombreux professeurs de mathématiques depuis le collège, aussi j'exprime mes sincères remerciements à mesdames et messieurs : Courteville, Toulouse, Nadau, Rignanese, Pagot, Dehevels, Breton, Cadre, Beauchard. En particulier, je tiens à exprimer ma profonde gratitude à Monsieur La Hargue qui lors de ma 5/2 n'a pas hésité à prendre des heures sur son temps libre pour m'aider à préparer le concours de l'ENS et avec qui j'ai depuis lors toujours gardé contact.

Je remercie profondément mes amis Ker-Lannais : Laulau, Caro, Cam's, Ninon, Fred, Maud, Florian, Adrien, Kevin, Agathe, Claire, Estelle. Les trois années passées à Rennes m'ont fait vous rencontrer et de très belles amitiés, qui dureront longtemps j'en suis sûre, sont nées.

Je remercie chaleureusement mes amis de prépa et plus particulièrement : Lisa et Sarah (j'adore nos moments de complicités et nos rires pas discrets du tout !) mais aussi Moïse et Antoine que j'ai eu le plaisir de revoir il n'y a pas si longtemps.

Je tiens à remercier aussi Axelle, que je connais depuis le CE2, j'espère que tu ne m'en veux plus pour mon mauvais accueil à ton arrivée dans notre petite école de Sinzos ;-)

Un immense merci ensuite à mes trois amies de toujours. Je veux bien sûr parler de Lola (rencontrée en seconde), Clémentine et Marion (rencontrées en sixième). Membres de notre groupe légendaire des « Black Moon », elles sont pour moi bien plus que des amies, elles sont une famille de plus, née de l'amitié et de la musique. Je sais que je peux compter sur elles quoiqu'il arrive, je mesure à quel point j'ai de la chance de vous avoir rencontrées. Je ne peux parler de Clémentine et Marion sans évoquer leurs parents : Jean, Anne, Thierry et Françoise qui font aussi partie de ma famille. Ils sont pour moi le deuxième parent que je n'ai jamais eu.

Je remercie du fond du cœur mes frères et sœurs (sans oublier leurs moitiés respectives), ma grande fratrie adorée, vénérée, adulée : Damien et Stine, Céline et Jean, David et Claire (qui reste dans notre cœur à tous), Thomas et Lætitia, Jòan et Pauline, Marion. Je ne pourrais me passer du joyeux bazar lorsqu'on se retrouve tous à la maison : nos petits-déjeuners à rallonge pour attendre que tout le monde soit levé, nos après-midis au Pouquet et nos pauses « petit café » ! La famille s'agrandit de jour en jour, je remercie mes trois neveux et ma petite nièce qui me comblent de bonheur : Jules, Olive, Theo et le petit dernier Jack !

Mes pensées se tournent également vers mes grands-parents maternels, Mamy et Grand-père, ainsi que mes cousins et cousines avec qui j'ai régulièrement partagé de très bons moments.

Mes sentiments émus s'adressent maintenant à ma mère, qui a su nous transmettre de belles valeurs et a toujours eu une confiance inconditionnelle en nos capacités. Tu nous donnes une force exceptionnelle, j'espère réussir à faire aussi bien si j'ai des enfants un jour.

Je remercie bien sûr ma toute nouvelle famille (depuis 6 ans déjà), les Bahier : Véronique, Florent et Maéva. Vous m'avez accueilli très chaleureusement et le calvaire de la belle-famille pour certains est un réel bonheur pour moi !

Enfin, merci à Valentin, mon moussin, j'ai tellement de chance de t'avoir à mes côtés, je t'aime de tout mon cœur.

Table des matières

1	Introduction générale	1
1.1	Le mouvement brownien fractionnaire	2
1.1.1	Définition et propriétés	2
1.1.2	Représentations	2
1.2	Ergodicité des équations différentielles stochastiques à mémoire	3
1.2.1	Bruit fractionnaire	4
1.2.2	Bruit Gaussien général	7
1.3	Principaux résultats de la thèse	8
1.3.1	Ergodicité de dynamiques discrètes à mémoire	8
1.3.2	Inégalités de concentration en temps long pour des EDS fractionnaires	15
1.3.3	Estimation paramétrique du terme de drift pour des EDS fractionnaires	20
2	Discrete-time stochastic dynamics with memory	29
2.1	Introduction	29
2.2	Setting and main results	32
2.2.1	Notations	32
2.2.2	Dynamical system and Markovian structure	32
2.2.3	Preliminary tool: a Toeplitz type operator	34
2.2.4	Assumptions and general theorem	35
2.2.5	The Euler Scheme	38
2.2.6	Two explicit cases which satisfy $(\mathbf{H}_{\text{exp}})$	39
2.2.7	Polynomial case: from a general class of examples to the fractional case	40
2.3	Existence of invariant distribution	43
2.4	General coupling procedure	43
2.4.1	Scheme of coupling	44
2.4.2	Coupling lemmas to achieve Step 1 and 2	45
2.5	Coupling under $(\mathbf{H}_{\text{poly}})$ or $(\mathbf{H}_{\text{exp}})$	50
2.5.1	Admissibility condition	50
2.5.2	Lower-bound for the successful-coupling probability	51
2.6	(K, α) -admissibility	57
2.6.1	On condition (2.5.1)	57
2.6.2	Compact return condition (2.5.2)	62
2.7	Proof of Theorem 2.2.1	64

2.A	Proof of Theorem 2.2.2	67
2.B	Explicit formula for the sequence $(b_k)_{k \geq 0}$	70
2.C	Particular case: when the sequence $(a_k)_{k \geq 0}$ is log-convex	71
2.D	Proof of Proposition 2.2.3	72
2.E	Proof of Theorem 2.3.1	76
2.F	Proof of Lemma 2.5.3	78
2.G	Proof of Hypothesis (\mathbf{H}'_1)	80
3	Concentration inequalities for additive fractional SDE	91
3.1	Introduction	91
3.2	Setting and main results	93
3.2.1	Notations	93
3.2.2	Assumptions and general result	94
3.2.3	Long time concentration inequalities for occupation measures	96
3.3	Sketch of proof	96
3.4	Control of the martingale increments	98
3.4.1	First case : $u \geq 2$	100
3.4.2	Second case : $u \in [0, 2]$	104
3.5	Conditional exponential moments of the martingale increments	105
3.5.1	Conditional moments of the martingale increments	105
3.5.2	Proof of Proposition 3.2.1	108
3.A	Sub-Gaussianity of the supremum of the Browian motion	110
3.B	Uniform sub-Gaussianity of $\ G^{(k)}\ _{\alpha, [0, 2]}$	111
4	Drift estimation procedure for additive fractional SDE	117
4.1	Introduction	117
4.2	Main results	119
4.2.1	Notations	119
4.2.2	Assumptions	120
4.2.3	Statistical setting and construction of the estimator	121
4.2.4	Main consistency results	123
4.2.5	Rate of convergence	124
4.3	Preliminary Results	126
4.3.1	Ergodic properties of the SDE and of the Euler scheme	126
4.3.2	Continuity of $\vartheta \mapsto d(\nu_\vartheta, \nu_{\vartheta_0})$	129
4.3.3	Further controls under (\mathbf{C}_s)	131
4.4	Proof of the consistency theorems	134
4.4.1	Uniform convergence of the contrast	134
4.4.2	A general convergence result	142
4.4.3	Proofs of the convergence theorems	143
4.5	Rate of convergence: proof of Theorem 4.2.4	144
4.5.1	L^q bound on $D_n^{(1)}$	145
4.5.2	L^q bound on $D_{N, \gamma}^{(2)}$	147
4.5.3	L^q bound on $D_{N, \gamma}^{(3)}$	147
4.6	Identifiability assumption	151

4.6.1	Case of a diffusion process	151
4.6.2	Fractional Brownian motion case	152
4.7	Numerical Discussions and Illustrations	156
4.A	Tightness and convergence of the occupation measures	160
4.A.1	Moment estimates	160
4.A.2	Proof of Proposition 4.3.3	166
4.B	Proof of Proposition 4.3.8	168
5	Perspectives	171
5.1	Sur l'ergodicité de dynamiques discrètes fractionnaires	171
5.2	Sur la concentration pour des EDS fractionnaires	172
5.3	Sur l'estimation du drift pour des EDS fractionnaires	172

Notations de l'introduction

La liste suivante présente les principales notations utilisées dans l'introduction générale de la thèse (chapitre 1). Les notations spécifiques aux autres chapitres sont intégrées à ceux-ci.

$\langle \cdot, \cdot \rangle$	Produit scalaire usuel sur \mathbb{R}^d .
$ \cdot $	Norme euclidienne sur \mathbb{R}^d ou valeur absolue sur \mathbb{R} .
$B(a, R)$	Boule de centre a et de rayon R sur \mathbb{R}^d pour la norme euclidienne.
$\ \cdot\ _{\text{Lip}}$	Semi-norme Lipschitz : soit $f : (E, d_E) \rightarrow (F, d_F)$ où (E, d_E) et (F, d_F) désignent deux espaces métriques, alors

$$\|f\|_{\text{Lip}} := \sup_{x \neq y} \frac{d_F(f(x), f(y))}{d_E(x, y)}.$$

$\ \cdot\ _{\alpha, [a,b]}$	Semi-norme α -Hölder : soit $f : [a, b] \rightarrow \mathbb{R}^d$ alors,
-----------------------------	---

$$\|f\|_{\alpha, [a,b]} := \sup_{a \leq s < t \leq b} \frac{|f(t) - f(s)|}{|t - s|^\alpha}.$$

$\mathbb{M}_{d \times d}$	Matrices à coefficients réels de taille $d \times d$.
\mathbb{M}_d	Matrices diagonales à coefficients réels de taille $d \times d$.
$\mathcal{M}_1(E)$	Mesures de probabilités sur un espace donné E .
$\ \cdot\ _{\text{TV}}$	Distance en variation totale : soient $\mu, \nu \in \mathcal{M}_1(E)$ pour un espace E donné, alors

$$\|\nu - \mu\|_{\text{TV}} := \sup_{A \subset E} |\nu(A) - \mu(A)|.$$

$\Pi(\cdot, \cdot)$	Ensemble des couplages de mesures de probabilités : soient $\mu, \nu \in \mathcal{M}_1(E)$ pour un espace E donné, alors
---------------------	--

$$\Pi(\nu, \mu) := \{\pi \in \mathcal{M}_1(E \times E) \text{ telles que } \pi(\cdot, E) = \nu(\cdot) \text{ et } \pi(E, \cdot) = \mu(\cdot)\}.$$

$\mathcal{W}_p(\cdot, \cdot)$	Distance de Wasserstein d'ordre p : soient $\mu, \nu \in \mathcal{M}_1(E)$ pour un espace métrique (E, d) donné, alors
-------------------------------	--

$$\mathcal{W}_p(\nu, \mu) := \inf_{\pi \in \Pi(\nu, \mu)} \left(\int_{E \times E} d(x, y)^p d\pi(x, y) \right)^{1/p}.$$

1. Introduction générale

En théorie des probabilités, l'étude du régime stationnaire de systèmes dynamiques aléatoires est un domaine de recherche très actif. Plusieurs aspects sont abordés : l'existence et l'unicité d'un tel régime, la convergence et la vitesse de convergence vers ce régime stationnaire. L'intérêt porté à ce sujet vient notamment du fait que l'évolution de nombreux phénomènes réels peut être modélisée à l'aide de systèmes dynamiques aléatoires. En particulier, les systèmes markoviens ont fait l'objet et font encore l'objet de nombreux travaux. Cependant, ce type de modèles apparaît dans certaines situations comme peu adapté ou trop restrictif. C'est donc naturellement que des dynamiques aléatoires plus générales dirigées par des bruits à mémoire ont commencé à être étudiées.

Dans un cadre continu, l'étude du régime stationnaire des Équations Différentielles Stochastiques (EDS) dirigées par un processus gaussien à accroissements stationnaires suscite beaucoup d'intérêt depuis une dizaine d'années. Un exemple de tels processus est le mouvement brownien fractionnaire, qui est une généralisation du mouvement brownien standard. En particulier, les EDS fractionnaires sont utilisées dans de nombreux domaines : par exemple pour modéliser l'évolution de phénomènes physiques, biologiques ou financiers (voir [32, 42, 45, 57]). Par conséquent, ces équations apparaissent dans de nombreux travaux : sur l'existence et l'unicité d'une mesure invariante dans ce cadre a priori non markovien (voir [35, 34, 38]), sur la vitesse de convergence à l'équilibre (voir [36, 26, 22]) ou encore sur l'approximation de la mesure invariante (voir [15, 16]). Nous détaillerons ces derniers résultats dans la Section 1.2.1 après avoir introduit de façon précise le mouvement brownien fractionnaire dans la Section 1.1. Pour un bruit gaussien à accroissements stationnaires plus général, nous verrons dans la Section 1.2.2 que de récents résultats [59] nous donnent une vitesse de convergence à l'équilibre sous certaines hypothèses bien choisies.

Dans un cadre discret, Hairer dans [37] propose des conditions générales donnant l'existence et l'unicité d'une mesure invariante pour une grande classe de processus non-markoviens. De plus, l'étude en temps long de systèmes discrets particuliers comme le schéma d'Euler associé à une EDS dirigée par un bruit gaussien général (dans [15]) ou par le mouvement brownien fractionnaire (dans [16]) permet d'approcher la loi stationnaire de l'EDS, ce qui est essentiel car il est difficile de la calculer explicitement dans la plupart des cas.

Dans cette thèse, nous nous intéressons à différents problèmes en lien avec l'ergodicité de tels systèmes dynamiques à mémoire, en temps continu ou discret, et tout particulièrement pour un bruit fractionnaire. Dans la suite de ce chapitre, nous allons détailler les

résultats existants sur lesquels cette thèse s'appuie avant d'exposer nos résultats principaux.

1.1 Le mouvement brownien fractionnaire

Le mouvement brownien fractionnaire est l'un des principaux bruits gaussiens considérés dans cette thèse. Nous allons donc nous attacher tout au long de cette section à le définir et à donner ses principales propriétés.

1.1.1 Définition et propriétés

Définition 1.1.1. Soit $H \in (0, 1)$. Le *mouvement brownien fractionnaire* (mBf) d -dimensionnel de paramètre de Hurst H , noté $(B_t)_{t \geq 0}$, est un processus gaussien centré dont la fonction de covariance est donnée par

$$\mathbb{E} [B_t^i B_s^j] = \frac{\delta_{ij}}{2} [t^{2H} + s^{2H} - |t - s|^{2H}] \quad \text{pour tout } t, s \geq 0.$$

En particulier, les accroissements du mBf sont stationnaires :

$$\mathbb{E} [(B_t^i - B_s^i)(B_t^j - B_s^j)] = \delta_{ij} |t - s|^{2H} \quad \text{pour tout } t, s \geq 0,$$

et le mBf est un processus H -autosimilaire, c'est-à-dire tel que

$$\mathcal{L}((B_{ct})_{t \geq 0}) = \mathcal{L}(c^H(B_t)_{t \geq 0}) \quad \text{pour tout } c > 0.$$

Remarque 1.1.1. Le processus B n'est pas une semimartingale excepté lorsque $H = 1/2$ (voir Liptser et Shiryaev [49]). Dans ce cas, B n'est autre que le mouvement brownien standard et ses accroissements sont indépendants. Notons que la théorie de l'intégration d'Itô ne s'applique donc pas au mBf.

Le paramètre de Hurst H gouverne le signe de la corrélation des accroissements du mouvement brownien fractionnaire (suivant si $H < 1/2$ ou $H > 1/2$) mais prescrit aussi la régularité de ses trajectoires comme expliqué dans la proposition suivante.

Proposition 1.1.1. Soit $H \in (0, 1)$. Les trajectoires du mouvement brownien fractionnaire de paramètre de Hurst H sont presque sûrement β -Hölder pour tout $\beta < H$. Autrement dit, pour tout $T > 0$ et pour $0 < \beta < H$,

$$\|B\|_{\beta, [0, T]} := \sup_{0 \leq s < t \leq T} \frac{|B_t - B_s|}{|t - s|^\beta} < +\infty \quad p.s.$$

1.1.2 Représentations

Dans [50], Mandelbrot et Van Ness montrent que le mouvement brownien fractionnaire peut s'obtenir en intégrant un noyau déterministe contre une mesure gaussienne. En effet, soit $(W_t)_{t \in \mathbb{R}}$ un mouvement brownien standard d -dimensionnel, alors le processus

B défini par

$$B_t = \kappa(H) \int_{\mathbb{R}} \left[(t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right] dW_s, \quad t \in \mathbb{R}. \quad (1.1.1)$$

où $\kappa(H)$ est une constante de normalisation qui ne dépend que de H , est un mBf sur \mathbb{R}^d . Cette *représentation en moyenne mobile* est primordiale pour étudier l'ergodicité des EDS fractionnaires. Cela permet de plonger la solution d'une telle EDS dans une structure markovienne homogène et donc de donner un sens à la notion de mesure invariante dans ce cadre. Cet aspect sera développé dans la Section 1.2.1.

Une deuxième représentation du mBf, indexé sur \mathbb{R}_+ (voir par exemple [18] ou [20]), appelée *représentation de Volterra* est la suivante : soit $(W_t)_{t \geq 0}$ un mouvement Brownien d -dimensionnel alors le processus B défini par

$$B_t = \int_0^t K_H(t,s) dW_s, \quad t \geq 0, \quad (1.1.2)$$

où K_H est un noyau déterministe donné par

$$K_H(t,s) = c_H \left[\frac{t^{H-\frac{1}{2}}}{s^{H-\frac{1}{2}}} (t-s)^{H-\frac{1}{2}} - \left(H - \frac{1}{2} \right) \int_s^t \frac{u^{H-\frac{3}{2}}}{s^{H-\frac{1}{2}}} (u-s)^{H-\frac{1}{2}} du \right] \quad (1.1.3)$$

est un mBf sur \mathbb{R}^d . Cette représentation est utilisée dans les travaux décrits dans la Section 1.3.2.

1.2 Ergodicité des équations différentielles stochastiques à mémoire

Dans cette partie, nous nous focalisons sur les EDS du type suivant : soit $(Y_t)_{t \geq 0}$ un processus à valeurs dans \mathbb{R}^d tel que

$$dY_t = b(Y_t)dt + \sigma(Y_t)dG_t \quad (1.2.1)$$

où $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $\sigma : \mathbb{R}^d \rightarrow \mathbb{M}_{d \times d}$ sont à minima des fonctions continues et G est un processus gaussien d -dimensionnel centré à accroissements stationnaires. Dans les deux prochaines sous-parties, nous précisons les hypothèses supplémentaires pour que cette EDS ait du sens en fonction de la nature exacte de G .

Remarque 1.2.1. Excepté dans le cas où les accroissements de G seraient indépendants (cas qui ne nous concernera pas ici), ce type de dynamique est clairement non-markovien. Ceci explique la terminologie “à mémoire” employée dans le titre de la section.

Nous aurons également besoin du schéma d'Euler à pas fixé $\gamma > 0$ associé à (1.2.1) défini par une donnée initiale $Z_0 := Y_0$ et la relation de récurrence suivante

$$\forall n \geq 0, \quad Z_{(n+1)\gamma} = Z_{n\gamma} + \gamma b(Z_{n\gamma}) + \sigma(Z_{n\gamma})\Delta_{n+1} \quad (1.2.2)$$

où $\Delta_{n+1} := G_{(n+1)\gamma} - G_{n\gamma}$.

1.2.1 Bruit fractionnaire

On considère dans toute cette section le cas de l'équation (1.2.1) lorsque G n'est autre que le mouvement brownien fractionnaire B de paramètre $H \in (0, 1)$ (voir Définition 1.1.1). La mémoire relative aux accroissements du mBf implique que les solutions de (1.2.1), quand elles existent, sont non-markoviennes. Sans cette propriété, il paraît difficile de définir une mesure invariante, d'étudier son existence et la convergence éventuelle de la solution de l'équation vers celle-ci. L'idée fondatrice introduite par Hairer dans [36] est de "grossir" l'espace en incluant tout le passé du bruit fractionnaire pour se ramener à un cadre markovien homogène. Ce principe repose sur la représentation de Mandelbrot-Van Ness du mBf définie par (1.1.1), l'objet d'intérêt n'étant plus $(Y_t)_{t \geq 0}$ mais $(Y_t, (W_s)_{s \leq t})_{t \geq 0}$ où W est le mouvement brownien standard sous-jacent (voir (1.1.1)). Dans son article, Hairer construit le semi-groupe de Feller associé à $(Y_t, (W_s)_{s \leq t})_{t \geq 0}$ en introduisant l'espace de bruit \mathcal{W} comme la fermeture de $\mathcal{C}_0^\infty(\mathbb{R}_-)$ (fonctions \mathcal{C}^∞ à support compact qui s'annulent en 0) pour la norme

$$\|w\|_H := \sup_{t, s \in \mathbb{R}_-} \frac{|w(t) - w(s)|}{|t - s|^{\frac{1-H}{2}} (1 + |t| + |s|)^{\frac{1}{2}}}.$$

Ce semi-groupe, noté \mathcal{Q} , est alors défini par : pour tout $f : \mathbb{R}^d \times \mathcal{W} \rightarrow \mathbb{R}_+$ mesurable,

$$\mathcal{Q}_t f(x, w) := \mathbb{E}_{(x, w)} [f(Y_t, (W_s)_{s \leq t})]. \quad (1.2.3)$$

La notion de mesure invariante peut ainsi être définie comme suit :

Définition 1.2.1. Une *mesure invariante* pour (1.2.1) est une mesure de probabilité μ sur $\mathbb{R}^d \times \mathcal{W}$ invariante pour \mathcal{Q} , c'est-à-dire vérifiant

$$\mathcal{Q}_t \mu = \mu, \quad \forall t \geq 0.$$

Remarque 1.2.2. Notons que dans tous les résultats qui vont suivre, lorsqu'on parlera d'*unicité de la mesure invariante*, cela sera à la relation d'équivalence près suivante. Soient μ, ν deux mesures de probabilités sur $\mathbb{R}^d \times \mathcal{W}$. On note $\mathcal{S}\mu := \mathcal{L}((Y_t^\mu)_{t \geq 0})$, c'est-à-dire la loi de la solution de l'EDS (1.2.1) partant de la mesure initiale μ . Alors la relation d'équivalence \sim est donnée par : $\mu \sim \nu \Leftrightarrow \mathcal{S}\mu = \mathcal{S}\nu$.

Vitesse de convergence à l'équilibre

Commençons par le cas *additif* considéré par Hairer dans [36], c'est-à-dire lorsque σ est une matrice constante inversible dans (1.2.1). Dans ce cadre, sous des hypothèses de croissance polynômiale sur le drift b et sa dérivée et sous la condition de stabilité suivante :

(\mathbf{H}_b) : il existe $\alpha, \alpha', \beta > 0$ tels que

$$\langle b(x) - b(y), x - y \rangle \leq \min\{ \beta - \alpha|x - y|^2, \alpha'|x - y|^2 \}, \quad \forall x, y \in \mathbb{R}^d,$$

Hairer donne une borne sur la vitesse de convergence de la solution vers la mesure

invariante pour la distance en variation totale. L'hypothèse (\mathbf{H}_b) est une condition de contraction en dehors d'un compact.

Théorème 1.2.1 (Hairer [36]). *Sous les hypothèses décrites ci-dessus, l'équation (1.2.1) lorsque G est un mBf admet une unique mesure invariante μ_* . De plus, pour toute mesure initiale μ et pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$ telle que*

$$\|\mathcal{L}((Y_{s+t}^\mu)_{s \geq 0}) - \mathcal{S}\mu_*\|_{TV} \leq C_\varepsilon t^{-(\alpha_H - \varepsilon)} \quad (1.2.4)$$

où

$$\alpha_H = \begin{cases} \frac{1}{8} & \text{si } H \in (\frac{1}{4}, 1) \setminus \{\frac{1}{2}\} \\ H(1 - 2H) & \text{si } H \in (0, \frac{1}{4}). \end{cases} \quad (1.2.5)$$

Remarque 1.2.3. \triangleright Notons ici qu'il y a un léger abus, les conditions initiales doivent être *admissibles*, c'est-à-dire que leur projection sur \mathcal{W} doit être la loi de $(W_t)_{t \leq 0}$.

\triangleright L'optimalité de cette vitesse est encore une question ouverte.

La preuve de ce théorème est basée sur une méthode de couplage qui est plus complexe que dans le cadre markovien classique. En effet l'idée est d'utiliser le fait que

$$\|\mathcal{L}((Y_{s+t}^\mu)_{s \geq 0}) - \mathcal{S}\mu_*\|_{TV} \leq \mathbb{P}(\tau > t)$$

où $\tau := \inf\{t \geq 0 \mid Y_s^1 = Y_s^2, \forall s \geq t\}$ et $(Y^1, Y^2) = ((Y_t^\mu)_{t \geq 0}, (Y_t^{\mu_*})_{t \geq 0})$. Il est important de noter que la méthode de couplage va comporter une étape supplémentaire par rapport au couplage classique markovien : tout d'abord l'étape 1 consiste à réussir à coller les trajectoires Y^1 et Y^2 a un instant donné, ensuite on doit essayer de les maintenir collées (étape 2 spécifique au non-markovien) et enfin en cas d'échec de l'une des deux premières étapes, on doit imposer $W^1 = W^2$ jusqu'à ce que les deux trajectoires reviennent dans un compact donné (où W^1 et W^2 sont les mouvements browniens sous-jacents à la représentation (1.1.1) pour les mBf associés respectivement à Y^1 et Y^2). Cette dernière étape permet de tenter de nouveau l'étape 1 en contrôlant le coût d'un éventuel échec de cette même étape.

L'existence et l'unicité d'une probabilité invariante ont été également étudiées dans un cadre *multiplicatif*, c'est à dire lorsque σ n'est plus constant, par Hairer et Ohashi dans [34] pour $H > 1/2$ puis par Hairer et Pillai dans [35] pour $H \in (1/3, 1/2)$. Concernant la vitesse de convergence, le Théorème 1.2.1 a aussi été prolongé. Tout d'abord lorsque $H > 1/2$, Fontbona et Panloup dans [26] prouvent ce même résultat lorsque σ est bornée lipschitz et inversible avec $\sigma^{-1} = \nabla h$ où h est un \mathcal{C}^1 -difféomorphisme de \mathbb{R}^d . Ces hypothèses sur σ permettent de pallier au fait que les deux propriétés suivantes ne sont plus vraies dans le cadre multiplicatif :

- Lorsque σ est constant, si deux mBf B^1 et B^2 diffèrent d'un drift alors les solutions Y^1 et Y^2 aussi, ce qui facilite l'étape 1 du couplage.
- Sous l'hypothèse (\mathbf{H}_b) et lorsque σ est constant, deux trajectoires dirigées par le même mBf se rapprochent automatiquement et la distance entre les deux trajectoires peut être contrôlée par une constante déterministe.

Remarque 1.2.4. Notons que la restriction à $H > 1/2$ permet d'utiliser la théorie des intégrales de Young qui donne un sens aux intégrales du type $\int_0^t f_s dg_s$ pour deux fonctions f et g Hölder d'indices respectifs α_1, α_2 avec $\alpha_1 + \alpha_2 > 1$: pour toute subdivision Π de l'intervalle $[0, t]$,

$$\int_0^t f_s dg_s := \lim_{|\Pi| \rightarrow 0} \sum_{[u,v] \in \Pi} f_u (g_v - g_u).$$

Toujours dans le cadre multiplicatif, le Théorème 1.2.1 est enfin démontré dans le cas $H > 1/3$ par Deya, Panloup et Tindel dans [22]. Ils supposent en particulier que σ est régulière bornée à dérivées bornées et inversible d'inverse borné. Dans le cas $H > 1/2$, cela prolonge un peu le résultat donné par [26] étant donné que l'on ne suppose plus que σ est une matrice jacobienne. Le cas $H \in (1/3, 1/2)$ apporte quant à lui une contribution plus signifiante qui fait appel à la théorie des trajectoires rugueuses afin de donner un sens à l'intégrale contre le mBf (voir [28] ou [29] pour plus de détails sur ce sujet).

Approximation de la loi stationnaire

Une fois la convergence vers la mesure invariante établie, une question naturelle est de savoir comment approcher cette mesure invariante, qui est en général difficile à calculer explicitement. Rappelons que la mesure invariante μ_\star donnée dans le Théorème 1.2.1 est une mesure de probabilité sur $\mathbb{R}^d \times \mathcal{W}$, nous nous intéressons ici à sa projection sur la première marginale, que l'on va noter $\bar{\mu}_\star \in \mathcal{M}_1(\mathbb{R}^d)$ dans la suite.

Dans le but d'approcher $\bar{\mu}_\star$, l'idée est d'utiliser le schéma d'Euler défini en (1.2.1) toujours dans le cas où G est le mouvement brownien fractionnaire. Dans [15] pour le cas additif avec $H \in (0, 1)$ et dans [16] pour le cas multiplicatif avec $H \in (1/2, 1)$, les auteurs utilisent la mesure d'occupation associée au schéma d'Euler (1.2.2) pour approcher $\bar{\mu}_\star$. Plus précisément, commençons par le cas additif [15]. En supposant que le drift b est Lipschitz et satisfait la condition de stabilité (\mathbf{H}_b) énoncée dans la section précédente avec $\beta = 0$, les auteurs obtiennent le résultat suivant (énoncé dans le cas particulier du mBf ici) :

Théorème 1.2.2 (Cohen, Panloup [15]). *Soit Z défini par (1.2.2) lorsque $G = B$ de paramètre de Hurst $H \in (0, 1)$ et lorsque σ est constant et inversible. Sous les hypothèses énoncées ci-dessus sur le drift b ,*

$$\lim_{\gamma \rightarrow 0} \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{Z_{k\gamma}} = \bar{\mu}_\star \quad p.s. \quad (1.2.6)$$

au sens de la convergence étroite sur \mathbb{R}^d .

Remarque 1.2.5. Le fait de supposer $\beta = 0$ dans l'hypothèse (\mathbf{H}_b) permet ici de montrer l'unicité de la solution stationnaire associée à (1.2.2) (pour $\gamma > 0$ suffisamment petit) et à (1.2.1). On peut noter ici que la définition de solution stationnaire diffère légèrement de celle de Hairer [36], il s'agit ici d'un processus solution de (1.2.1) ou (1.2.2) dont la loi est stationnaire. Il n'y a pas de condition d'adaptation du processus à la filtration naturelle engendrée par le bruit fractionnaire $(B_t)_{t \in \mathbb{R}}$. Cependant, lorsqu'il s'agit de solutions stationnaires adaptées, il y a en fait unicité même si (\mathbf{H}_b) est vérifiée pour un

β quelconque : pour le schéma d'Euler ceci est implicitement démontré dans [16] (en adaptant ce qui est fait au cadre additif et pour un $H \in (0, 1)$ quelconque) qui utilise les résultats de [37] et enfin Hairer le démontre dans [36] pour l'EDS (1.2.1) lorsque G est le mBf (comme on l'a vu précédemment).

Le Théorème 1.2.2 est adapté au cadre multiplicatif dans [16] lorsque $H \in (1/2, 1)$ pour un drift b sous linéaire de classe \mathcal{C}^2 satisfaisant toujours l'hypothèse de contraction en dehors d'un compact (\mathbf{H}_b) et pour un σ borné de classe \mathcal{C}^2 inversible d'inverse borné. Le cadre multiplicatif avec un paramètre de Hurst $H < 1/2$ n'a pas encore été étudié mais comme il est expliqué dans [16, Remarque 4], il semble que tous les outils pour le faire sont disponibles. Il faudrait évidemment utiliser les trajectoires rugueuses ne serait-ce que pour donner du sens à l'équation (1.2.1), mais aussi considérer un schéma d'Euler plus complexe faisant intervenir par exemple des produits d'incrémentes du mBf (voir [21, 19]).

1.2.2 Bruit Gaussien général

Dans toute cette section, nous considérons l'EDS (1.2.1) pour un bruit gaussien général d -dimensionnel G centré, à accroissements stationnaires et dont les composantes sont indépendantes. De plus on se place dans le cadre additif, c'est-à-dire avec σ constant et on suppose que σ est inversible.

Un travail récent de Panloup et Richard [59] considère le cas d'un bruit gaussien *purement non déterministe*¹ à valeurs dans \mathbb{R}^d . Cette dernière propriété est équivalente dans le cadre d'un bruit gaussien centré à accroissements stationnaires au fait de posséder une représentation en moyenne mobile (voir [59, Annexe B]). Autrement dit, il existe $\{\mathcal{G}(t)\}_{t \in \mathbb{R}}$ à valeurs dans \mathbb{M}_d telle que $\mathcal{G}(t) = 0$ pour tout $t > 0$ et

$$G_t = \int_{\mathbb{R}} \{\mathcal{G}(s-t) - \mathcal{G}(s)\} dW_s, \quad \forall t \geq 0 \quad (1.2.7)$$

où $(W_t)_{t \in \mathbb{R}}$ est un mouvement brownien standard sur \mathbb{R}^d . De plus, \mathcal{G} est telle que $\int_{\mathbb{R}} \|\mathcal{G}(s-t) - \mathcal{G}(s)\|^2 ds < +\infty$. Grâce à cette représentation, ils construisent comme Hairer [36] dans le cas du mBf une structure markovienne au-dessus de (1.2.1) et la notion de mesure invariante. Ils introduisent ensuite l'hypothèse suivante sur le drift b :

(\mathbf{H}'_b) : il existe $\alpha, R > 0$ tels que

- (i) $\langle b(x) - b(y), x - y \rangle \leq 0, \quad \forall x, y \in \mathbb{R}^d$
- (ii) $\langle b(x) - b(y), x - y \rangle \leq -\alpha|x - y|^2, \quad \forall x, y \in \mathbb{R}^d \setminus B(0, R).$

Remarque 1.2.6. Cette hypothèse suppose en fait que le drift b est contractant sur \mathbb{R}^d mais fortement contractant seulement en dehors d'un compact aussi grand que l'on veut.

Sous des hypothèses de contrôle sur la mémoire du processus G (reliées au contrôle de $\|\mathcal{G}''(u)\|$ par des fonctions polynômiales) et sous (\mathbf{H}'_b) , ils montrent l'existence et

¹Un processus $(X_s)_{s \in \mathbb{R}}$ est dit *purement non déterministe* si $\bigcap_{t \in \mathbb{R}} \overline{\text{Vect}}\{X_s \mid s \in (-\infty, t]\} = \{0\}$ où $\overline{\text{Vect}}(A)$ est la fermeture dans $L^2(\Omega)$ de l'espace engendré par $A \subset L^2(\Omega)$.

l'unicité de la mesure invariante associée à (1.2.1), notée $\mu_\star \in \mathcal{M}_1(\mathbb{R}^d \times \mathcal{W})$. De plus, si Y est une solution de (1.2.1) satisfaisant $\mathbb{E}[|Y_0|^{2(1+v)}] < +\infty$ pour un certain $v > 0$ alors ils montrent que $\mathcal{L}(Y_t)$ converge vers $\bar{\mu}_\star$ (projection de μ_\star sur \mathbb{R}^d) pour la distance de Wasserstein d'ordre 2 à vitesse sous-exponentielle.

Enfin, sous des hypothèses *ad hoc* reliées à l'inversibilité de la relation (1.2.7), ils prolongent le résultat qui vient d'être énoncé en remplaçant la distance de Wasserstein d'ordre 2 par la distance en variation totale. En particulier, toutes les hypothèses sur G s'appliquent au mBf quelque soit son paramètre de Hurst $H \in (0, 1)$ et donc ce résultat montre que la vitesse obtenue par Hairer dans le Théorème 1.2.1 est améliorée sous l'hypothèse un peu plus forte ($\mathbf{H}'_{\mathfrak{p}}$) sur le drift. Notons que la question de l'optimalité de la vitesse obtenue dans le Théorème 1.2.1 par Hairer [36] est encore une question ouverte.

Concernant l'approximation de la mesure invariante dans ce cadre additif, Cohen et Panloup [15] montrent en fait le Théorème 1.2.2 pour un bruit gaussien général G à accroissements stationnaires et sous des conditions sur la fonction de covariance de la suite Δ des accroissements de G .

Les travaux que nous venons de détailler sont ceux qui ont initié les différents travaux de recherche de cette thèse et donc les résultats associés. Nous décrivons ceux-ci dans la prochaine section et donnons une esquisse des principaux outils déployés pour prouver ces résultats.

1.3 Principaux résultats de la thèse

Cette thèse se découpe en trois parties que nous allons maintenant présenter. La première concerne la vitesse de convergence à l'équilibre (*i.e.* vers une mesure invariante) pour des dynamiques discrètes générales à mémoire. La deuxième est une étude de la concentration de la loi de la solution d'une EDS fractionnaire (dans le cas additif) en temps long. Enfin, la dernière est un travail commun avec Panloup et Tindel sur l'estimation du terme de drift dans une EDS fractionnaire additive, utilisant les travaux [15, 16] précédemment cités sur l'approximation de la mesure invariante.

1.3.1 Ergodicité de dynamiques discrètes à mémoire

Cette première partie présente les résultats qui seront entièrement détaillés ensuite dans le Chapitre 2 qui correspond à l'article [71]. La principale motivation à l'origine de ce travail était d'étudier la convergence à l'équilibre de dynamiques discrètes à mémoire de type fractionnaire de sorte que l'on s'affranchisse du caractère rugueux du mBf spécifique au cadre continu. Plus précisément, l'objectif initial était d'étudier comment la vitesse serait affectée par ce changement de cadre et ainsi d'évaluer la qualité des bornes obtenues dans le Théorème 1.2.1 par Hairer. Nous avons ensuite été naturellement amenés à travailler dans un cadre plus large en considérant des bruits gaussiens à mémoire quelconques.

Nous nous intéressons ici à des dynamiques de la forme suivante : soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires sur \mathbb{R}^d définie par sa valeur initiale X_0 et la relation de récurrence

$$X_{n+1} = F(X_n, \Delta_{n+1}) \quad (1.3.1)$$

où $F : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une fonction générale à minima continue et $(\Delta_n)_{n \in \mathbb{Z}}$ est une suite stationnaire gaussienne d -dimensionnelle purement non-déterministe dont les composantes sont indépendantes. Les conditions sur la suite Δ nous permettent d'utiliser le théorème de décomposition de Wold (voir par exemple [7]) qui nous donne

$$\forall n \in \mathbb{Z}, \quad \Delta_n = \sum_{k=0}^{+\infty} a_k \xi_{n-k} \quad (1.3.2)$$

avec

$$\begin{cases} (a_k)_{k \geq 0} \in \mathbb{R}^{\mathbb{N}} \text{ telle que } a_0 \neq 0 \text{ et } \sum_{k=0}^{+\infty} a_k^2 < +\infty \\ (\xi_k)_{k \in \mathbb{Z}} \text{ une suite i.i.d telle que } \xi_1 \sim \mathcal{N}(0, I_d). \end{cases} \quad (1.3.3)$$

Remarque 1.3.1. Nous pouvons déjà noter ici le rôle de la suite déterministe $(a_k)_{k \geq 0}$ qui encode en quelque sorte la mémoire du bruit gaussien Δ puisque sa fonction de covariance

$$c(k) := \mathbb{E}[\Delta_n \Delta_{n+k}] = \sum_{i=0}^{+\infty} a_i a_{k+i},$$

ne dépend que de ces coefficients $(a_k)_{k \geq 0}$. Plus ils décroissent vite plus la “mémoire” de Δ est faible. Il est donc évident que le comportement asymptotique de cette suite va jouer un rôle essentiel dans l'étude de la vitesse de convergence à l'équilibre du système dynamique stochastique (1.3.1).

Pour étudier le comportement en temps long de ce type de dynamiques, l'idée est comme dans [36, 26, 22] de construire une structure markovienne autour de (1.3.1).

Structure markovienne homogène

L'idée est la suivante, au lieu de s'intéresser seulement à $(X_n)_{n \geq 0}$ nous nous concentrons sur l'évolution du couple $(X_n, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$. Nous nous plaçons donc sur un espace plus grand noté $\mathcal{X} \times \mathcal{W}$ où $\mathcal{X} := \mathbb{R}^d$ est l'espace d'état de la suite X et $\mathcal{W} := (\mathbb{R}^d)^{\mathbb{Z}^-}$ est l'espace du bruit associé à Δ . La dynamique (1.3.1) est ainsi équivalente à

$$(X_{n+1}, (\Delta_{n+1+k})_{k \leq 0}) = \varphi((X_n, (\Delta_{n+k})_{k \leq 0}), \Delta_{n+1}) \quad (1.3.4)$$

où

$$\begin{aligned} \varphi : (\mathcal{X} \times \mathcal{W}) \times \mathbb{R}^d &\rightarrow \mathcal{X} \times \mathcal{W} \\ ((x, w), w') &\mapsto (F(x, w'), w \sqcup w') \end{aligned}$$

avec $(w \sqcup w')_0 = w'$ et $\forall k < 0, (w \sqcup w')_k = w_{k+1}$. On peut alors définir le noyau de transition \mathcal{Q} associé à ce nouveau système (1.3.4) comme suit : pour toute fonction

mesurable $g : \mathcal{X} \times \mathcal{W} \rightarrow \mathbb{R}$,

$$\int_{\mathcal{X} \times \mathcal{W}} g(x', w') \mathcal{Q}((x, w), (dx', dw')) = \int_{\mathbb{R}^d} g(\varphi((x, w), \delta)) \mathcal{P}(w, d\delta) \quad (1.3.5)$$

où $\mathcal{P}(w, d\delta) := \mathcal{L}(\Delta_{n+1} | (\Delta_{n+k})_{k \leq 0} = w)$ ne dépend pas de n car $(\Delta_n)_{n \in \mathbb{Z}}$ est une suite stationnaire. On peut donc réaliser la suite $(X_n, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ à partir d'une mesure initiale $\mu \in \mathcal{M}_1(\mathcal{X} \times \mathcal{W})$ et du noyau de transition \mathcal{Q} . Comme dans la Section 1.2.1, on obtient alors la définition de mesure invariante dans ce cadre, à savoir

Définition 1.3.1. Une *mesure invariante* pour (1.3.1) ou de façon équivalente pour (1.3.4) est une mesure de probabilité μ sur $\mathcal{X} \times \mathcal{W}$ invariante pour \mathcal{Q} , c'est-à-dire telle que

$$\mathcal{Q}\mu = \mu.$$

D'autre part, dans toute la suite, lorsqu'on parlera d'unicité de la mesure invariante, ce sera à la relation d'équivalence près définie ci-après. Pour toute mesure $\mu \in \mathcal{M}_1(\mathcal{X} \times \mathcal{W})$, on note $\mathcal{S}\mu := \mathcal{L}((X_n^\mu)_{n \geq 0})$ la loi de la projection sur \mathcal{X} de la suite $(X_n, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ partant de la mesure initiale μ . Ainsi on dit que $\mu, \nu \in \mathcal{M}_1(\mathcal{X} \times \mathcal{W})$ sont *équivalentes* et on note $\mu \sim \nu$ si et seulement si

$$\mathcal{S}\mu = \mathcal{S}\nu. \quad (1.3.6)$$

Autrement dit, on a bien unicité sur la marginale d'intérêt \mathcal{X} .

Avant d'énoncer le théorème principal au sujet de l'existence, l'unicité d'une mesure invariante et de la vitesse de convergence vers cette mesure pour la distance en variation totale, il nous faut donner quelques précisions sur la représentation en moyenne mobile (1.3.2) associée à Δ .

Inversion de la représentation en moyenne mobile

On peut associer à (1.3.2) de façon naturelle l'opérateur noté \mathbf{T}_a suivant

$$\mathbf{T}_a(w) := \left(\sum_{l=0}^{+\infty} a_l w_{-k-l} \right)_{k \geq 0}$$

pour toute suite $w \in (\mathbb{R}^d)^{\mathbb{Z}^-}$ telle que la quantité ci-dessus soit finie quelque soit $k \geq 0$. Nous montrons que cet opérateur est inversible et que l'inverse est tout simplement \mathbf{T}_b où $(b_k)_{k \geq 0}$ est la suite définie récursivement par

$$b_0 = \frac{1}{a_0} \quad \text{et} \quad \forall k \geq 1, \quad b_k = -\frac{1}{a_0} \sum_{l=1}^k a_l b_{k-l}. \quad (1.3.7)$$

La relation qui relie les suites (a_k) et (b_k) est complexe, l'expression explicite de (b_k) en fonction seulement de (a_k) est donnée dans le Chapitre 2 et est assez technique. En particulier, il n'existe pas de règle générale permettant de déduire le comportement asymptotique de la suite (b_k) en fonction de celui de la suite (a_k) . Nous étudions quelques

exemples qui illustrent ce phénomène dans les Sections 2.2.6 et 2.2.7. Nous aurons donc besoin d'hypothèses sur les deux suites pour le théorème principal.

Hypothèses et théorème principal

Nous allons considérer deux jeux d'hypothèses sur les coefficients (a_k) et (b_k) : le cas dit *polynômial* et le cas dit *exponentiel*.

Hypothèse (\mathbf{H}_{poly}): Les conditions suivantes sont satisfaites,

- il existe $\rho, \beta > 1/2$ et $C_\rho, C_\beta > 0$ tels que

$$\forall k \geq 0, \quad |a_k| \leq C_\rho (k+1)^{-\rho} \quad \text{et} \quad \forall k \geq 0, \quad |b_k| \leq C_\beta (k+1)^{-\beta}.$$

- il existe $\kappa \geq \rho + 1$ et $C_\kappa > 0$ tels que

$$\forall k \geq 0, \quad |a_k - a_{k+1}| \leq C_\kappa (k+1)^{-\kappa}.$$

Hypothèse (\mathbf{H}_{exp}): Il existe $\lambda, \zeta > 0$ et $C_\lambda, C_\zeta > 0$ tels que,

$$\forall k \geq 0, \quad |a_k| \leq C_\lambda e^{-\lambda k} \quad \text{et} \quad \forall k \geq 0, \quad |b_k| \leq C_\zeta e^{-\zeta k}.$$

Nous pouvons voir assez facilement que sous ses conditions, la fonction de covariance $c(k)$ associée à la suite Δ décroît respectivement à vitesse polynômiale et à vitesse exponentielle lorsque k tend vers $+\infty$. Cela donne donc en quelque sorte le poids de la mémoire de la suite Δ . Un cas mélangeant les deux types de décroissances asymptotiques pourrait également être envisagé mais pour des raisons de simplicité nous avons choisi de les traiter séparément. Notons qu'un tel cas satisferait en particulier l'hypothèse (\mathbf{H}_{poly}) puisque la décroissance (asymptotique) exponentielle implique la décroissance polynômiale.

Il reste maintenant à donner quelques précisions sur la fonction F qui dirige la dynamique (1.3.1).

Hypothèse (\mathbf{H}_1): Il existe une fonction continue $V : \mathcal{X} \rightarrow \mathbb{R}_+^*$ satisfaisant $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ et $\exists \gamma \in (0, 1)$ et $C > 0$ tels que pour tout $(x, w) \in \mathcal{X} \times \mathbb{R}^d$,

$$V(F(x, w)) \leq \gamma V(x) + C(1 + |w|). \tag{1.3.8}$$

Remarque 1.3.2. \triangleright Cette condition assure l'existence d'une fonction de Lyapunov (au sens de la Définition 2.3.1) et donc l'existence d'une mesure invariante, nous verrons cela en détail dans la Section 2.3. Ce type d'hypothèse existe aussi dans la littérature des chaînes de Markov discrètes (voir par exemple l'équation (7) dans [24]) mais sous une forme intégrale.

\triangleright Cette condition est aussi vraie si on a l'existence d'une telle fonction V avec $V(F(x, w)) \leq \gamma V(x) + C(1 + |w|^p)$ et $p > 1$. Il suffit de prendre alors $V^{1/p}$ à la place de V pour vérifier (\mathbf{H}_1).

Une deuxième hypothèse sur la fonction F , notée (\mathbf{H}_2) dans la suite, est nécessaire pour la preuve du théorème principal. Elle est indispensable pour la première étape de la procédure de couplage (stratégie de preuve déjà évoquée dans un cadre continu dans la Section 1.2.1 et qui est utilisée ici aussi) qui amène avec probabilité strictement positive à la coalescence des deux trajectoires à un instant donné. Cette hypothèse est à mettre en parallèle avec l'hypothèse de minoration du noyau de transition dans un cadre purement markovien. Nous avons choisi de ne pas la détailler ici pour ne pas gêner la fluidité de lecture, nous donnerons cependant dans la suite un des exemples principaux de fonctions satisfaisant cette condition (une grande classe d'exemples sera également abordée dans la remarque 2.2.5).

Nous pouvons maintenant énoncer le résultat principal de cette section.

Théorème 1.3.1. *On suppose (\mathbf{H}_1) et (\mathbf{H}_2) . Alors,*

- (i) *il existe une mesure invariante μ_\star associée à (1.3.1).*
- (ii) *En supposant de plus que $(\mathbf{H}_{\text{poly}})$ est vérifiée pour $\rho, \beta > 1/2$ et $\rho + \beta > 3/2$, il y a unicité de la mesure invariante μ_\star . De plus, pour toute mesure initiale μ_0 telle que $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_0(dx) < +\infty$ et pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$ telle que*

$$\|\mathcal{L}((X_{n+k}^{\mu_0})_{k \geq 0}) - \mathcal{S}\mu_\star\|_{TV} \leq C_\varepsilon n^{-(v(\beta, \rho) - \varepsilon)}.$$

où la fonction v est donnée par

$$v(\beta, \rho) = \sup_{\alpha \in (\frac{1}{2}v(\frac{3}{2} - \beta), \rho)} \min\{1, 2(\rho - \alpha)\}(\min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2).$$

- (iii) *En supposant de plus que $(\mathbf{H}_{\text{exp}})$ est vérifiée, il y a unicité de la mesure invariante μ_\star . De plus, pour toute mesure initiale μ_0 telle que $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_0(dx) < +\infty$ et pour tout $p > 0$, il existe $C_p > 0$ telle que*

$$\|\mathcal{L}((X_{n+k}^{\mu_0})_{k \geq 0}) - \mathcal{S}\mu_\star\|_{TV} \leq C_p n^{-p}.$$

Remarque 1.3.3. La vitesse polynômiale de tout ordre pour le cas *exponentiel* semble suggérer que l'on pourrait obtenir mieux. Pourquoi pas une vitesse exponentielle ? Cette question est difficile à traiter sans avoir détaillé la stratégie de preuve, elle sera donc abordée un peu plus loin.

Ce théorème englobe une classe de dynamiques assez générale par deux aspects : le fait que l'on ne donne pas une forme explicite à la fonction F et le fait que les hypothèses faites sur le bruit gaussien sont seulement relatives à sa mémoire. Dans la Section 2.2, nous étudions d'une part le cas où F correspond au schéma d'Euler associé à l'équation (1.2.1) et d'autre part des cas particuliers de suite (a_k) et (b_k) pour lesquelles on est capable de vérifier les hypothèses $(\mathbf{H}_{\text{poly}})$ ou $(\mathbf{H}_{\text{exp}})$. En particulier, dans la section qui suit nous avons choisi de développer le cas du schéma d'Euler associé à l'EDS (1.2.1) dans le cas fractionnaire.

Schéma d'Euler d'une EDS fractionnaire

Nous décrivons maintenant l'un des principaux exemples d'application du Théorème 1.3.1. On pose

$$F(x, w) = x + hb(x) + \sigma(x)w \quad \text{et} \quad \Delta_n = B_{(n+1)h} - B_{nh} \quad (1.3.9)$$

où $h > 0$, $b : \mathcal{X} \rightarrow \mathcal{X}$, $\sigma : \mathcal{X} \rightarrow \mathbb{M}_{d \times d}$ et B est le mouvement brownien fractionnaire de paramètre $H \in (0, 1)$. Comme nous l'avons dit précédemment, il s'agit du schéma d'Euler à pas fixé $h > 0$ associé à (1.2.1) lorsque $G = B$.

On pose **(L1)** et **(L2)** les deux conditions suivantes sur b :

(L1) $\exists C > 0$ telle que

$$\forall x \in \mathcal{X}, \quad |b(x)| \leq C(1 + |x|).$$

(L2) $\exists \tilde{\beta} \in \mathbb{R}$ et $\tilde{\alpha} > 0$ tels que

$$\forall x \in \mathcal{X}, \quad \langle x, b(x) \rangle \leq \tilde{\beta} - \tilde{\alpha}|x|^2.$$

Théorème 1.3.2. *On suppose que σ est continue, inversible d'inverse continu et telle que $|\sigma(x)| \leq C(1 + |x|^\kappa)$ avec $\kappa \in (0, 1)$. On suppose que b est continue et satisfait les conditions **(L1)** et **(L2)** ci-dessus. Alors pour $h > 0$ suffisamment petit, F satisfait les deux hypothèses **(H1)** et **(H2)**.*

D'autre part, les coefficients de la représentation en moyenne mobile de la suite $(\Delta_n) = (B_{(n+1)h} - B_{nh})$ nous sont donnés par la suite $(a_k^H)_{k \geq 0}$ définie par (voir [58])

$$\begin{cases} a_0^H = h^H \kappa(H) 2^{1/2-H} \\ a_k^H = h^H \kappa(H) \left(\left(k + \frac{1}{2}\right)^{H-1/2} - \left(k - \frac{1}{2}\right)^{H-1/2} \right) \text{ pour } k \geq 1. \end{cases} \quad (1.3.10)$$

où

$$\kappa(H) = \frac{\sqrt{\sin(\pi H) \Gamma(2H + 1)}}{\Gamma(H + 1/2)}.$$

Nous pouvons facilement vérifier que $a_k^H \underset{k \rightarrow +\infty}{\sim} C_{h,H}(k+1)^{-(3/2-H)}$ et $|a_k^H - a_{k+1}^H| \leq C'_{h,H}(k+1)^{-(5/2-H)}$. Nous sommes donc a priori dans un cadre polynômial et nous obtenons le résultat suivant :

Proposition 1.3.1. *Il existe $C''_{h,H} > 0$ telle que pour tout $H \in (0, 1/2)$*

$$\forall k \geq 0, \quad |b_k^H| \leq C''_{h,H}(k+1)^{-(H+1/2)}. \quad (1.3.11)$$

Ainsi, dans le cas de la fonction F considérée ci-dessus, nous déduisons du Théorème 1.3.2 et de la Proposition 1.3.1 que le Théorème 1.3.1 est vérifié lorsque $H \in (0, 1/2)$ avec le taux

$$v(3/2 - H, H + 1/2) = \begin{cases} H(1 - 2H) & \text{si } H \in (0, 1/4] \\ \frac{1}{8} & \text{si } H \in (1/4, 1/2). \end{cases}$$

Nous retrouvons le taux donné dans un cadre continu par Hairer dans [36]. La preuve de la Proposition 1.3.1 est très technique et demande des majorations très fines. Nous n'avons pour l'instant qu'une conjecture pour le comportement asymptotique de la suite (b_k^H) dans le cas où $H \in (1/2, 1)$ et le taux conjecturé nous donne une vitesse plus faible que celle de Hairer dans le cas continu. Une démarche spécifique tenant compte de l'expression exacte de la suite (a_k^H) est une piste éventuelle pour améliorer ce résultat.

D'autres exemples de suites (a_k) sont également étudiées dans les Sections 2.2.6 et 2.2.7 satisfaisant l'une ou l'autre des hypothèses $(\mathbf{H}_{\text{poly}})$ ou $(\mathbf{H}_{\text{exp}})$

Stratégie de preuve par couplage

On suppose ici que l'existence d'une mesure invariante est déjà actée. La preuve de la vitesse de convergence donnée dans le Théorème 1.3.1 repose sur une méthode par couplage à laquelle nous avons déjà fait allusion dans le cadre continu des EDS fractionnaires. On considère deux couples $(X_n^1, (\Delta_{n+k}^1)_{k \leq 0})_{n \geq 0}$ et $(X_n^2, (\Delta_{n+k}^2)_{k \leq 0})_{n \geq 0}$ suivant la dynamique (1.3.4) et avec pour conditions initiales respectives μ_0 et μ_* (une mesure invariante). On note $(\xi_n^1)_{n \in \mathbb{Z}}$ (respectivement $(\xi_n^2)_{n \in \mathbb{Z}}$) la suite de gaussiennes sous-jacente à la représentation en moyenne mobile de Δ^1 (respectivement de Δ^2), définie en (1.3.2).

On a la majoration suivante

$$\|\mathcal{L}((X_{n+k}^1)_{k \geq 0}) - \mathcal{S}\mu_*\|_{TV} \leq \mathbb{P}(\tau_\infty > n) \quad \forall n \geq 0 \quad (1.3.12)$$

où $\mathcal{S}\mu_* = \mathcal{L}((X_{n+k}^2)_{k \geq 0})$ quelque soit $n \geq 0$ car μ_* est invariante et où τ_∞ est défini par

$$\tau_\infty = \inf\{n \geq 0 \mid X_k^1 = X_k^2, \forall k \geq n\}.$$

Dans un cadre purement markovien, ce temps d'arrêt serait réduit au premier instant où les deux trajectoires se collent. En effet, la position suivante ne dépendant que de la position courante, elles resteraient alors automatiquement collées. On a donc ici une étape supplémentaire dans une tentative de couplage. Ainsi, une procédure de couplage se compose d'une série d'essais qui comprennent chacun trois étapes :

- * **Étape 1:** Essayer de coller les deux trajectoires à un instant donné avec un "coût contrôlé".
- * **Étape 2:** (spécifique au cadre non markovien) Essayer de maintenir les trajectoires collées.
- * **Étape 3:** Si l'une des deux étapes précédentes échoue, on attend suffisamment longtemps de façon à pouvoir réaliser l'étape 1 avec un "coût contrôlé". Pendant cette étape, on impose $\xi_n^1 = \xi_n^2$.

Dans l'étape 1, lorsqu'on parle de coût contrôlé, il s'agit en fait d'être en mesure de contrôler la probabilité de réussite de cette étape et en même temps de contrôler l'écart entre les trajectoires en cas d'échec. Plus précisément, on ne va tenter l'étape 1 à un instant donné τ que si en cas d'échec, on est capable d'assurer que X_τ^1 et X_τ^2 sont dans

un compact fixé au préalable et que le *drift* entre les bruits Δ_τ^1 et Δ_τ^2 est suffisamment petit (voir condition (2.5.1)). Ceci est réalisable en calibrant la durée de l'étape 3 qui impose $\xi_n^1 = \xi_n^2$. Cette étape permet d'alléger petit à petit le poids du passé et donc d'obtenir le retour des trajectoires dans un compact donné.

Toute la difficulté de la preuve repose sur la construction des suites ξ^1 et ξ^2 au fur et à mesure des tentatives de couplage. Cela fait l'objet de la Section 2.4.2. Bien que ce soit deux suites i.i.d., elles vont par construction être fortement dépendantes l'une par rapport à l'autre.

La preuve du Théorème 1.3.1 repose ensuite sur l'inégalité de Markov

$$\mathbb{P}(\tau_\infty > n) \leq \frac{\mathbb{E}[\tau_\infty^p]}{n^p}$$

pour $p > 0$. L'idée est de découper τ_∞ suivant les différentes tentatives de couplage et à la fin de maximiser en p tout en s'assurant que $\mathbb{E}[\tau_\infty^p] < +\infty$.

Remarque 1.3.4. Le fait de considérer l'inégalité de Markov avec des fonctions polynômiales permet de séparer complètement les contributions des différentes tentatives (via l'inégalité de Minkowski pour $p > 1$ ou l'inégalité élémentaire $|u + v|^p \leq |u|^p + |v|^p$ pour $0 < p < 1$), ce qui n'est pas possible si on prend une fonction exponentielle par exemple. Et dans ce cadre non purement markovien, c'est indispensable. Ceci explique en particulier pourquoi dans le cas de l'hypothèse (\mathbf{H}_{exp}), nous avons seulement pu prouver une vitesse polynômiale de tout ordre (cet aspect est d'avantage développé dans le Chapitre 2 en remarque 2.6.1).

1.3.2 Inégalités de concentration en temps long pour des EDS fractionnaires

Dans cette section nous présentons les résultats obtenus dans l'article [70] et qui seront développés dans le Chapitre 3. L'idée de ce travail est née au cours du projet sur l'estimation du drift pour des EDS fractionnaires dont les résultats seront présentés à la prochaine section. Pour obtenir une vitesse de convergence de l'estimateur que l'on considèrerait, il est apparu indispensable de démontrer des résultats de concentration pour des mesures d'occupation discrètes associées à l'EDS (1.2.1) dans un cadre additif et fractionnaire. Plus précisément, nous considérons ici l'EDS suivante : soit $(Y_t)_{t \geq 0}$ un processus à valeurs dans \mathbb{R}^d tel que $Y_0 = x \in \mathbb{R}^d$ et

$$dY_t = b(Y_t)dt + \sigma dB_t \tag{1.3.13}$$

où $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une fonction au minimum continue, $\sigma \in \mathbb{M}_{d \times d}$ et B est un mouvement brownien fractionnaire d -dimensionnel de paramètre de Hurst $H \in (0, 1)$.

La question principale qui a motivé ce travail était donc d'établir des inégalités de concentration en temps long pour des mesures du type $\frac{1}{n} \sum_{k=1}^n \delta_{Y_{k\Delta}}$ où $\Delta > 0$ est un pas fixé. En statistique, le processus est observé uniquement à des temps discrets, c'est pourquoi cette question est essentielle. Un moyen classique d'obtenir des inégalités de concentration est d'utiliser les inégalités de transport. Précisons de quoi il s'agit. Une mesure de probabilité μ sur un espace métrique (E, d) satisfait une *inégalité de transport*

L^p si et seulement si pour toute mesure de probabilité ν sur E , il existe une constante $C > 0$ telle que

$$\mathcal{W}_p(\mu, \nu) \leq \sqrt{2C\mathbf{H}(\nu|\mu)} \quad (1.3.14)$$

où

$$\mathcal{W}_p(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{E \times E} d(x, y)^p d\pi(x, y) \right)^{1/p} \quad (1.3.15)$$

est la distance de Wasserstein d'ordre p et

$$\mathbf{H}(\nu|\mu) := \begin{cases} \int \log \left(\frac{d\nu}{d\mu} \right) d\nu, & \text{si } \nu \ll \mu, \\ +\infty & \text{sinon,} \end{cases}$$

est l'entropie. On note alors $\mu \in T_p(C)$. Dans le cas où $p = 1$, cette inégalité est très fortement reliée à la concentration. Cet aspect a d'abord été abordé par Marton [52, 51], Talagrand [66], Bobkov et Götze [5] et amplement exploré par Ledoux [48, 47]. On peut montrer que l'inégalité (1.3.14) pour $p = 1$ est en fait équivalente à la propriété suivante : pour toute fonction F à valeurs réelles μ -intégrable et α -Lipschitz, et pour tout $\lambda \in \mathbb{R}$, on a

$$\mathbb{E}(\exp(\lambda(F(X) - \mathbb{E}[F(X)]))) \leq \exp\left(C\alpha^2 \frac{\lambda^2}{2}\right) \quad (1.3.16)$$

avec $\mathcal{L}(X) = \mu$. L'inégalité de Markov permet alors de conclure à de la concentration. Dans un cadre continu, les inégalités de transport L^1 et L^2 pour la loi de la solution d'une EDS fractionnaire ont été étudiées notamment par Guendouzi [33], Sausseureau [65] et Riedel [62] (ce dernier étudie même le cas de bruits gaussiens plus généraux que le mBf). En particulier, Sausseureau obtient des inégalités de concentration en temps long pour la mesure d'occupation continue associée à une EDS fractionnaire de paramètre $H \in (1/2, 1)$ sous une condition de contractivité forte sur le terme de drift. Ces résultats sont dus au fait que la constante qu'il obtient dans les inégalités de transport L^2 (et donc L^1 puisque $T_2(C)$ implique $T_1(C)$) est suffisamment bien contrôlée en fonction de l'horizon T sur lequel il considère la solution de l'EDS. L'approche développée par Sausseureau ne semble pas s'adapter à notre cadre discret. Nous avons par exemple essayé d'adapter les résultats en changeant la métrique pour l'inégalité $T_1(C)$ puis de montrer que l'application : $(y_t)_{t \in [0, T]} \mapsto (y_{k\Delta})_{k \in \{1, \dots, n\}}$ est Lipschitz pour cette nouvelle métrique. Dans toutes nos tentatives pour adapter la preuve de Sausseureau à une version discrétisée, nous obtenions une constante C dans l'inégalité $T_1(C)$ dont la dépendance en n ne nous permettait pas de déduire de la concentration en temps long.

Nous avons finalement évolué vers une approche fortement inspirée de Djellout, Guillin et Wu [23] dans un cadre diffusif. Au lieu de démontrer une inégalité de transport L^1 directement sous la forme (1.3.14), nous nous sommes attachés à démontrer sa formulation équivalente (1.3.16). Le point de départ de [23] est de décomposer $F(X) - \mathbb{E}[F(X)]$ en accroissements de martingale comme suit :

$$F(X) - \mathbb{E}[F(X)] = \sum_{k=1}^n \mathbb{E}[F(X)|\mathcal{F}_k] - \mathbb{E}[F(X)|\mathcal{F}_{k-1}] \quad (1.3.17)$$

où dans notre cas $X = (Y_{k\Delta})_{1 \leq k \leq n}$ et Y est la solution de (1.3.13). Ensuite en contrôlant

“suffisamment bien” tous les moments conditionnels de ces accroissements de martingale puis en utilisant le développement en série de la fonction exponentielle, nous avons pu montrer une inégalité du type (1.3.16). Cette démonstration repose sur la représentation de Volterra du mBf donnée par (1.1.2) et la filtration naturelle associée au mouvement brownien standard $(W_t)_{t \geq 0}$ sous-jacent. L’utilisation de cette représentation ici n’est pas anodine. Elle permet en effet d’avoir $\mathbb{E}[F(X)|\mathcal{F}_0] = \mathbb{E}[F(X)]$, ce qui est cohérent avec l’égalité (1.3.17), contrairement à la représentation de Mandelbrot Van-Ness (1.1.1) qui fait intervenir un mouvement brownien indexé sur \mathbb{R} .

Cette méthode nous a finalement permis de montrer des résultats de concentration en temps long pour la mesure d’occupation discrète mais aussi continue associée à (1.3.13). De plus, les résultats exposés sont vrais quel que soit le paramètre de Hurst $H \in (0, 1)$. Ils sont décrits ci-après.

Hypothèses et résultats principaux

Nous considérons l’équation (1.3.13) sous les hypothèses suivantes sur le terme de drift b :

Hypothèse 1.3.3. *On suppose que b est continue et qu’il existe $\alpha, L > 0$ telles que :*

(i) *pour tout $x, y \in \mathbb{R}^d$,*

$$\langle b(x) - b(y), x - y \rangle \leq -\alpha|x - y|^2,$$

(ii) *pour tout $x, y \in \mathbb{R}^d$,*

$$|b(x) - b(y)| \leq L|x - y|.$$

Avant de donner le résultat principal, nous devons introduire quelques notations spécifiques à ce résultat. Pour un $n \in \mathbb{N}^*$ et $(x, y) \in (\mathbb{R}^d)^n \times (\mathbb{R}^d)^n$, on note d_n la distance L^1 suivante :

$$d_n(x, y) := \sum_{k=1}^n |x_k - y_k|. \quad (1.3.18)$$

De même, pour un $T > 0$ donné et $(x, y) \in \mathcal{C}([0, T], \mathbb{R}^d) \times \mathcal{C}([0, T], \mathbb{R}^d)$, on note d_T la distance L^1 classique :

$$d_T(x, y) := \int_0^T |x_t - y_t| dt. \quad (1.3.19)$$

Enfin, pour $F : ((\mathbb{R}^d)^n, d_n) \rightarrow (\mathbb{R}, |\cdot|)$ et $\tilde{F} : (\mathcal{C}([0, T], \mathbb{R}^d), d_T) \rightarrow (\mathbb{R}, |\cdot|)$ deux fonctions Lipschitz, on note

$$F_Y := F(Y_{t_1}, \dots, Y_{t_n}) \quad \text{et} \quad \tilde{F}_Y = \tilde{F}((Y_t)_{t \in [0, T]}) \quad (1.3.20)$$

avec $0 < \Delta = t_1 < \dots < t_n$ et $t_{k+1} - t_k = \Delta$ pour un $\Delta > 0$ donné.

Nous sommes maintenant en mesure d’énoncer un résultat de contrôle des moments exponentiels :

Proposition 1.3.2. *Soit $H \in (0, 1)$ et $\Delta > 0$. Soient $n \in \mathbb{N}^*$, $T \geq 1$ et d_n, d_T les distances définies respectivement par (3.2.1) et (3.2.2). Alors,*

- (i) il existe $C_{H,\Delta} > 0$ telle que pour toute fonction Lipschitz $F : ((\mathbb{R}^d)^n, d_n) \rightarrow (\mathbb{R}, |\cdot|)$ et pour tout $\lambda > 0$,

$$\mathbb{E}[\exp(\lambda(F_Y - \mathbb{E}[F_Y]))] \leq \exp(C_{H,\Delta} \|F\|_{\text{Lip}}^2 \lambda^2 n^{2H \vee 1}). \quad (1.3.21)$$

- (ii) il existe $\tilde{C}_H > 0$ telle que pour toute fonction Lipschitz $\tilde{F} : (\mathcal{C}([0, T], \mathbb{R}^d), d_T) \rightarrow (\mathbb{R}, |\cdot|)$ et pour tout $\lambda > 0$,

$$\mathbb{E}[\exp(\lambda(\tilde{F}_Y - \mathbb{E}[\tilde{F}_Y]))] \leq \exp(\tilde{C}_H \|\tilde{F}\|_{\text{Lip}}^2 \lambda^2 T^{2H \vee 1}). \quad (1.3.22)$$

Remarque 1.3.5. Il est important de noter que ces inégalités sont du type de (1.3.16). Comme mentionné dans l'introduction de cette section, cette proposition est donc équivalente à ce que $\mathcal{L}((Y_{k\Delta})_{1 \leq k \leq n}) \in T_1(2C_{H,\Delta} n^{2H \vee 1})$ pour la distance d_n et $\mathcal{L}((Y_t)_{t \in [0, T]}) \in T_1(2\tilde{C}_H T^{2H \vee 1})$ pour la distance d_T .

De la Proposition 1.3.2 découle les inégalités de concentration suivantes :

Théorème 1.3.4. Soit $H \in (0, 1)$ et $\Delta > 0$. Soit $n \in \mathbb{N}^*$, $T \geq 1$ et d_n, d_T les distances définies respectivement par (1.3.18) et (1.3.19). Alors,

- (i) il existe $C_{H,\Delta} > 0$ telle que pour toute fonction Lipschitz $F : ((\mathbb{R}^d)^n, d_n) \rightarrow (\mathbb{R}, |\cdot|)$ et pour tout $r \geq 0$,

$$\mathbb{P}(F_Y - \mathbb{E}[F_Y] > r) \leq \exp\left(-\frac{r^2}{4C_{H,\Delta} \|F\|_{\text{Lip}}^2 n^{2H \vee 1}}\right). \quad (1.3.23)$$

- (ii) il existe $\tilde{C}_H > 0$ telle que pour toute fonction Lipschitz $\tilde{F} : (\mathcal{C}([0, T], \mathbb{R}^d), d_T) \rightarrow (\mathbb{R}, |\cdot|)$ et pour tout $r \geq 0$,

$$\mathbb{P}(\tilde{F}_Y - \mathbb{E}[\tilde{F}_Y]) > r) \leq \exp\left(-\frac{r^2}{4\tilde{C}_H \|\tilde{F}\|_{\text{Lip}}^2 T^{2H \vee 1}}\right). \quad (1.3.24)$$

Remarque 1.3.6. \triangleright Il est intéressant de voir que l'on retrouve dans les cas $H \in (1/2, 1)$ et $H = 1/2$ les mêmes ordres dans les puissances de T ou n que Sausseureau [65] et [23]. Cependant, l'optimalité de ces ordres n'est pas démontrée.

\triangleright Noter ici qu'évidemment on n'obtient de la concentration en temps long que pour des fonctionnelles F (respectivement \tilde{F}) dont la norme Lipschitz décroît plus vite que $n^{-H \vee \frac{1}{2}}$ (respectivement $T^{-H \vee \frac{1}{2}}$).

Application aux mesures d'occupation continues et discrètes

Du Théorème 1.3.4, nous obtenons le résultat recherché sur les mesures d'occupation discrètes mais également sur les mesures d'occupation continues :

Théorème 1.3.5. Soit $H \in (0, 1)$ et $\Delta > 0$. Soit $n \in \mathbb{N}^*$ et $T \geq 1$. Alors,

(i) il existe $C_{H,\Delta} > 0$ telle que pour toute fonction Lipschitz $f : (\mathbb{R}^d, |\cdot|) \rightarrow (\mathbb{R}, |\cdot|)$ et pour tout $r \geq 0$,

$$\mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n (f(Y_{t_k}) - \mathbb{E}[f(Y_{t_k})]) > r \right) \leq \exp \left(-\frac{r^2 n^{2-(2H \vee 1)}}{4C_{H,\Delta} \|f\|_{\text{Lip}}^2} \right). \quad (1.3.25)$$

(ii) il existe $\tilde{C}_H > 0$ telle que pour toute fonction Lipschitz $f : (\mathbb{R}^d, |\cdot|) \rightarrow (\mathbb{R}, |\cdot|)$ et pour tout $r \geq 0$,

$$\mathbb{P} \left(\frac{1}{T} \int_0^T (f(Y_t) - \mathbb{E}[f(Y_t)]) dt > r \right) \leq \exp \left(-\frac{r^2 T^{2-(2H \vee 1)}}{4\tilde{C}_H \|f\|_{\text{Lip}}^2} \right). \quad (1.3.26)$$

Ce dernier résultat repose sur l'application du Théorème 1.3.4 aux fonctionnelles définies ci-après :

$$\forall x \in (\mathbb{R}^d)^n, \quad F(x) := \frac{1}{n} \sum_{k=1}^n f(x_k) \quad \text{et} \quad \forall x \in \mathcal{C}([0, T], \mathbb{R}^d), \quad \tilde{F}(x) := \frac{1}{T} \int_0^T f(x_t) dt$$

qui sont respectivement $\frac{\|f\|_{\text{Lip}}}{n}$ -Lipschitz par rapport à d_n et $\frac{\|f\|_{\text{Lip}}}{T}$ -Lipschitz par rapport à d_T .

Esquisse de preuve de la Proposition 1.3.2 dans le cas discret

Supposons le pas Δ égal à 1. Comme nous l'avons mentionné précédemment, la preuve repose sur la décomposition suivante :

$$F_Y - \mathbb{E}[F_Y] = \sum_{k=1}^n \mathbb{E}[F_Y | \mathcal{F}_k] - \mathbb{E}[F_Y | \mathcal{F}_{k-1}] =: \sum_{k=1}^n M_k - M_{k-1}$$

où la filtration considérée est la filtration brownienne naturelle associée à W dans la représentation de Volterra du mBf (1.1.2).

La première étape de la preuve, qui est la plus longue, consiste à montrer que l'on peut majorer tous les moments d'ordre p des accroissements $M_k - M_{k-1}$ conditionnés à \mathcal{F}_{k-1} de la façon suivante :

$$\mathbb{E}[|M_k - M_{k-1}|^p | \mathcal{F}_{k-1}] \leq C \|F\|_{\text{Lip}}^p \psi_{n,k}^p \zeta^{p/2} p \Gamma \left(\frac{p}{2} \right) \quad (1.3.27)$$

où $C, \zeta > 0$, $\psi_{n,k}$ est une suite de coefficients fortement liés au noyau K_H de la représentation de Volterra du mBf et Γ est la fonction d'Euler. Conditionnellement à \mathcal{F}_{k-1} , M_k et M_{k-1} peuvent être vues comme des fonctionnelles de processus dont le bruit associé ne diffère que sur l'intervalle $[0, 1]$. Comme les fonctionnelles en jeu sont lipschitziennes, contrôler $|M_k - M_{k-1}|$ revient en réalité à contrôler l'impact de cette perturbation du bruit sur la distance entre les trajectoires associées. L'hypothèse de forte contraction joue ici un rôle majeur.

La deuxième étape de la preuve a pour objectif de majorer les moments exponentiels

des accroissements $M_k - M_{k-1}$ conditionnés à \mathcal{F}_{k-1} . Ceci s'obtient à l'aide du lemme suivant (tiré du Chapitre 1 de [63] et dont une démonstration est donnée dans la Section 3.3) qui permet pour une variable aléatoire réelle X centrée de passer d'une majoration des moments à une majoration de $\mathbb{E}[e^{\lambda X}]$ pour tout $\lambda > 0$.

Lemme 1.3.1. *Soit X une variable aléatoire réelle centrée telle que pour tout $p \geq 2$, il existe $C, \zeta > 0$ telles que*

$$\mathbb{E}[|X|^p] \leq C\zeta^{p/2} p\Gamma\left(\frac{p}{2}\right).$$

Alors pour tout $\lambda > 0$,

$$\mathbb{E}[e^{\lambda X}] \leq e^{2C'\zeta\lambda^2}.$$

La dernière étape combine la majoration obtenue des moments exponentiels conditionnés avec l'implication suivante : si il existe une suite déterministe (u_k) telle que

$$\mathbb{E}\left[e^{\lambda(M_k - M_{k-1})} \middle| \mathcal{F}_{k-1}\right] \leq e^{\lambda^2 u_k},$$

alors

$$\mathbb{E}\left[e^{\lambda M_n}\right] = \mathbb{E}\left[e^{\lambda M_{n-1}} \mathbb{E}\left[e^{\lambda(M_n - M_{n-1})} \middle| \mathcal{F}_{n-1}\right]\right] \leq \exp(\lambda^2 u_n) \mathbb{E}\left[e^{\lambda M_{n-1}}\right]$$

et donc

$$\mathbb{E}\left[e^{\lambda M_n}\right] \leq \exp\left(\lambda^2 \sum_{k=1}^n u_k\right).$$

À partir de ce schéma de preuve, on obtient finalement la majoration voulue dans la Proposition 1.3.2.

1.3.3 Estimation paramétrique du terme de drift pour des EDS fractionnaires

Dans cette partie, nous présentons les principaux résultats obtenus dans un travail effectué en collaboration avec Panloup et Tindel [60]. Ils seront développés dans le Chapitre 4. Il s'agit d'un travail sur l'estimation paramétrique du terme de drift pour une EDS fractionnaire dans le cas additif. Ce travail repose sur l'idée que, sous des hypothèses raisonnables, le coefficient de drift de l'EDS est bien déterminé par son régime stationnaire, *i.e.* que la mesure invariante (lorsqu'elle existe et qu'elle est unique) caractérise le terme de drift².

Introduisons précisément le modèle auquel on s'intéresse. On considère le processus Y à valeurs dans \mathbb{R}^d satisfaisant l'EDS :

$$Y_t = y_0 + \int_0^t b_{\vartheta_0}(Y_s) ds + \sigma B_t, \quad t \in [0, T]. \quad (1.3.28)$$

Ici $y_0 \in \mathbb{R}^d$ est une condition initiale donnée et $B = (B^1, \dots, B^d)$ est un mouvement brownien fractionnaire d -dimensionnel de paramètre de Hurst $H \in (0, 1)$, le paramètre

²Si cette hypothèse est difficile à vérifier pour les EDS fractionnaires, elle est vraie pour des diffusions dans un cadre assez général (voir Proposition 4.6.1).

inconnu ϑ_0 appartient à un certain ensemble noté $\Theta \subset \mathbb{R}^q$ ($q \geq 1$), $\{b_\vartheta(\cdot), \vartheta \in \Theta\}$ est une famille connue de coefficients de drift avec $b_\vartheta(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$, et $\sigma \in \mathbb{M}_{d \times d}$ est une matrice supposée connue.

De nombreux travaux sur l'estimation du terme de drift pour des EDS fractionnaires du type (1.3.28) existent dans la littérature. Une partie d'entre eux traitent de coefficients $b_{\vartheta_0}(Y_s)$ de la forme $\vartheta_0 b(Y_s)$ ou même $\vartheta_0 Y_s$ (voir par exemple [3, 39, 40, 43, 46, 61, 67]). Dans la plupart de ces travaux, on suppose que l'on observe le processus de façon continue ou que l'on discrétise des observations continues du processus. Neuenkirch et Tindel dans [55] améliorent ces deux aspects en proposant une estimation basée sur des observations discrètes et pour un drift $b_{\vartheta_0}(Y_s)$ s'exprimant sous la forme d'un gradient (ce qui reste une certaine restriction lorsque la dimension d est plus grande que 1). Dans ce papier, la consistance de l'estimateur est cependant obtenue sous l'hypothèse d'identifiabilité suivante : si on note \bar{Y} la solution stationnaire associée à (1.3.28) alors

$$\mathbb{E} [|b_{\vartheta_0}(\bar{Y}_0)|^2] = \mathbb{E} [|b_\vartheta(\bar{Y}_0)|^2] \quad \text{si et seulement si} \quad \vartheta = \vartheta_0.$$

Cette hypothèse est assez contraignante et difficile à vérifier en pratique.

Notons également quelques travaux qui utilisent des méthodes non-paramétriques pour estimer le terme de drift de (1.3.28) : Mishra et Prakasa Rao [54], Comte et Marie [17]. Ils considèrent des estimateurs à noyau qui reposent sur une observation continue du processus ou même l'observation de plusieurs trajectoires de (1.3.28). Dans ces deux travaux, (1.3.28) est considérée en dimension 1 et pour un paramètre de Hurst $H > 1/2$.

Dans ce travail, nous proposons une méthode d'estimation paramétrique qui ne suppose aucune restriction sur $H \in (0, 1)$ ou sur la dimension $d \geq 1$. Comme mentionné plus haut, l'hypothèse d'identifiabilité que l'on fait repose sur la mesure invariante associée à (1.3.28), elle est donc plus générale que celle faite dans [55] évoquée plus haut. Nous pouvons maintenant exposer les hypothèses considérées.

Hypothèses du modèle

Énonçons tout d'abord une hypothèse classique de compacité sur l'ensemble de paramètre Θ :

(H₀) : L'ensemble Θ est un compact de \mathbb{R}^q pour un certain $q \geq 1$.

Nous rappelons que les estimateurs considérés se basent sur la notion de mesure invariante associée à (1.3.28). Comme nous l'avons vu plus haut, l'existence et l'unicité de celle-ci repose sur des hypothèses de coercivité sur le terme de drift b . Dans ce travail, nous allons naviguer entre deux types d'hypothèses (faible et forte) notées respectivement **(C_w)** et **(C_s)**. L'hypothèse faible est donnée par :

(C_w) : $b \in \mathcal{C}^{1,1}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$ et il existe $\alpha, \beta, C, L > 0$ et $r \in \mathbb{N}$ telles que
(i) Pour tout $x, y \in \mathbb{R}^d$ et $\vartheta \in \Theta$,

$$\langle b_\vartheta(x) - b_\vartheta(y), x - y \rangle \leq \beta - \alpha |x - y|^2 \quad \text{et} \quad |b_\vartheta(x) - b_\vartheta(y)| \leq L |x - y| \quad (1.3.29)$$

(ii) Pour tout $x \in \mathbb{R}^d$ et $\vartheta \in \Theta$ la majoration suivante est satisfaite :

$$|\partial_{\vartheta} b_{\vartheta}(x)| \leq C(1 + |x|^r). \quad (1.3.30)$$

L'hypothèse forte est alors définie comme suit :

(**C_s**) : L'hypothèse (**C_w**) est satisfaite pour $\beta = 0$.

Sous l'hypothèse (**C_w**) (ou sous (**C_s**) qui est plus forte) et à condition que σ soit inversible, d'après le Théorème 1.2.1 énoncé plus haut (dû à Hairer [36]), il y a existence et unicité de la mesure invariante associée à l'équation suivante quelque soit $\vartheta \in \Theta$:

$$dY_t^{\vartheta} = b_{\vartheta}(Y_t^{\vartheta}) dt + \sigma dB_t, \quad t \in [0, T]. \quad (1.3.31)$$

On note alors ν_{ϑ} la projection sur la première marginale \mathbb{R}^d . De façon équivalente, si l'on note $(\bar{Y}_t^{\vartheta})_{t \geq 0}$ la solution stationnaire associée à (1.3.31), alors

$$\nu_{\vartheta} = \mathcal{L}(\bar{Y}_0^{\vartheta}).$$

Dans la suite on appellera ν_{ϑ} la mesure invariante marginale. On peut maintenant introduire les hypothèses d'identifiabilité (faible et forte) liées à cette mesure invariante marginale.

(**I_w**) : Pour tout $\vartheta \in \Theta$, on a $\nu_{\vartheta} = \nu_{\vartheta_0}$ si et seulement si $\vartheta = \vartheta_0$.

Notons que si d est une distance sur $\mathcal{M}_1(\mathbb{R}^d)$, alors (**I_w**) se réexprime de la façon suivante :

$$d(\nu_{\vartheta}, \nu_{\vartheta_0}) = 0 \quad \text{si et seulement si} \quad \vartheta = \vartheta_0. \quad (1.3.32)$$

Soit d une distance sur $\mathcal{M}_1(\mathbb{R}^d)$. L'hypothèse d'identifiabilité forte relative à d est

(**I_s**) Il existe une constante $C > 0$ et un paramètre $\varsigma \in (0, 1]$ tels que $\forall \vartheta \in \Theta$,

$$d(\nu_{\vartheta}, \nu_{\vartheta_0}) \geq C|\vartheta - \vartheta_0|^{\varsigma}. \quad (1.3.33)$$

Remarque 1.3.7. \triangleright Sous l'hypothèse d'identifiabilité faible, nous obtenons des résultats de consistance pour nos estimateurs tandis que l'hypothèse forte nous permet de donner une vitesse de convergence dans certains cas. Plus précisément, sans cette hypothèse, nous sommes capables d'établir des vitesses de convergence entre mesures invariantes. Pour en déduire des vitesses sur l'estimateur, ce type d'hypothèse devient alors nécessaire.

\triangleright Ces hypothèses ont été vérifiées pour une certaine classe d'équations construites comme des perturbations du processus d'Ornstein-Uhlenbeck fractionnaire (voir Section 4.6).

Construction de l'estimateur

La construction des estimateurs considérés ici repose sur l'approximation de la mesure invariante marginale donnée par le Théorème 1.2.2. On suppose ici que l'on observe la solution de (1.3.28) $(Y_t)_{t \geq 0}$ à des instants discrets $\{t_k; 0 \leq k < n\}$, avec $t_{k+1} - t_k = \kappa$ et

$\kappa > 0$. Sous l'hypothèse (\mathbf{C}_w) il peut être montré que

$$\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}} \xrightarrow{n \rightarrow +\infty} \nu_{\vartheta_0} \quad p.s.,$$

où “ \implies ” correspond ici à la convergence étroite pour des mesures de probabilités sur \mathbb{R}^d . Cette convergence est démontrée dans le Chapitre 4. En conséquence de cette convergence et de l'hypothèse d'identifiabilité (\mathbf{I}_w) introduite plus haut, il est naturel de considérer un estimateur de la forme :

$$\hat{\vartheta} := \operatorname{argmin}_{\vartheta \in \Theta} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \nu_{\vartheta} \right)$$

où d est une distance sur $\mathcal{M}_1(\mathbb{R}^d)$. Cependant, comme nous l'avons déjà mentionné, il est difficile en général d'obtenir une expression explicite de ν_{ϑ} (une exception étant le processus d'Ornstein-Uhlenbeck fractionnaire). On va donc considérer une approximation numérique de la mesure invariante marginale ν_{ϑ} et ce, à l'aide des travaux précédemment cités sur ce sujet ([15, 16]).

Plus précisément, nous construisons l'approximation discrète du processus Y^{ϑ} solution de (1.3.31) de la façon suivante : soit $(s_k)_{k \geq 0}$ une suite croissante et positive telle que $s_0 = 0$ et $\lim_{k \rightarrow +\infty} s_k = +\infty$, on définit alors le schéma d'Euler Z^{ϑ} par $Z_0^{\vartheta} = y_0$ et

$$\forall k \geq 0, \quad Z_{s_{k+1}}^{\vartheta} = Z_{s_k}^{\vartheta} + (s_{k+1} - s_k) b_{\vartheta}(Z_{s_k}^{\vartheta}) + \sigma (\tilde{B}_{s_{k+1}} - \tilde{B}_{s_k}). \quad (1.3.34)$$

où \tilde{B} est un mBf d -dimensionnel simulé a priori distinct de celui de (1.3.31).

Remarque 1.3.8. En pratique dans le Chapitre 4, nous ne ferons plus de distinction entre B et \tilde{B} mais le lecteur doit garder en tête qu'ils sont différents. En particulier, le lecteur pourra remarquer que nous ne ferons jamais de comparaison trajectorielles entre le processus observé Y dont on veut estimer le paramètre ϑ_0 et le schéma Z^{ϑ} .

Dans la suite, nous distinguerons deux types de schémas : le *schéma à pas constant* pour lequel $s_k = k\gamma$ avec $\gamma > 0$ que l'on notera $Z^{\vartheta, \gamma}$ et le *schéma à pas décroissant* pour lequel la suite de pas définie par $\gamma_k = s_k - s_{k-1}$ est décroissante et tend vers 0. À partir de ces deux schémas, on construit les deux estimateurs suivants :

$$\hat{\vartheta}_{N, n, \gamma} = \operatorname{argmin}_{\vartheta \in \Theta} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta, \gamma}} \right), \quad (1.3.35)$$

et

$$\hat{\vartheta}_{N, n} = \operatorname{argmin}_{\vartheta \in \Theta} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^{\vartheta}} \right), \quad (1.3.36)$$

où d est une distance sur $\mathcal{M}_1(\mathbb{R}^d)$. Les deux mesures d'occupation associées aux schémas d'Euler représentent notre approximation de ν_{ϑ} . En particulier, nous montrerons que les deux quantités à minimiser constituent des approximations de $d(\nu_{\vartheta_0}, \nu_{\vartheta})$ (ceci est démontré dans les Propositions 4.4.1 et 4.4.2) ce qui était notre but.

Résultats de consistance

Avant d'énoncer les résultats, précisons les distances que l'on considère pour nos estimateurs. Soit $p > 0$, on définit

$$\mathcal{D}_p := \{ \text{distances } d \text{ sur } \mathcal{M}_1(\mathbb{R}^d); \exists c > 0 \text{ tel que } \forall \nu, \mu \in \mathcal{M}_1(\mathbb{R}^d), d(\nu, \mu) \leq c\mathcal{W}_p(\nu, \mu) \} \quad (1.3.37)$$

où \mathcal{W}_p est la distance de Wasserstein d'ordre p dont la définition est rappelée en (1.3.15) (où ici on choisit l'espace métrique $(\mathbb{R}^d, |\cdot|)$).

Le premier résultat de consistance que nous allons décrire, qui fait intervenir l'hypothèse faible (\mathbf{C}_w) , nécessite une discrétisation de l'espace des paramètres Θ dans le sens suivant. L'hypothèse de compacité (\mathbf{H}_0) et la propriété de Borel-Lebesgue nous donnent l'existence pour tout $\varepsilon > 0$ de $M_\varepsilon \in \mathbb{N}$ et $\Theta^{(\varepsilon)} := \{\vartheta_i^{(\varepsilon)}; 1 \leq i \leq M_\varepsilon\} \subset \Theta^{M_\varepsilon}$ tels que $\Theta \subset \bigcup_{i=1}^{M_\varepsilon} B(\vartheta_i^{(\varepsilon)}, \varepsilon)$. De cette façon, on définit la discrétisation de Θ

$$\vartheta^{(\varepsilon)} := \operatorname{argmin}_{\vartheta' \in \Theta^{(\varepsilon)}} |\vartheta' - \vartheta|. \quad (1.3.38)$$

Ceci nous amène à notre premier résultat de consistance

Théorème 1.3.6. *On suppose (\mathbf{H}_0) , (\mathbf{C}_w) et (\mathbf{I}_w) satisfaites. Soit $p > 0$ et soit d une distance sur $\mathcal{M}_1(\mathbb{R}^d)$ appartenant à \mathcal{D}_p (où \mathcal{D}_p est défini par (1.3.37)). Alors la famille $\{\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)}; N \geq 1, n \geq 1, \gamma > 0, \varepsilon > 0\}$ définie par*

$$\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)} = \operatorname{argmin}_{\vartheta \in \Theta^{(\varepsilon)}} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{s_k}^{\vartheta,\gamma}} \right), \quad N, n \in \mathbb{N}, \gamma > 0, \varepsilon > 0 \quad (1.3.39)$$

est un estimateur fortement consistant de ϑ_0 . Plus précisément, on a

$$\lim_{\varepsilon \rightarrow 0} \lim_{\gamma \rightarrow 0} \lim_{N,n \rightarrow +\infty} \hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)} = \vartheta_0 \quad p.s.$$

Remarque 1.3.9. La discrétisation de Θ est indispensable sous (\mathbf{C}_w) car sous cette hypothèse plus faible, nous perdons certains résultats de convergence uniforme en ϑ valables sous (\mathbf{C}_s) (typiquement dans les Propositions 4.4.1 et 4.4.2). Ces résultats sont essentiels pour démontrer les deux théorèmes de consistance qui suivent.

Nous pouvons maintenant énoncer les deux résultats principaux de consistance. Le premier concerne l'estimateur défini en (1.3.35) lié au schéma d'Euler à pas constant.

Théorème 1.3.7. *On suppose (\mathbf{H}_0) , (\mathbf{C}_s) et (\mathbf{I}_w) satisfaites. Comme dans le Théorème 1.3.6, soit $p > 0$ et d une distance sur $\mathcal{M}_1(\mathbb{R}^d)$ appartenant à \mathcal{D}_p . Alors la famille $\{\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)}; N \geq 1, n \geq 1, \gamma > 0\}$ définie par (1.3.35) est un estimateur fortement consistant de ϑ_0 au sens suivant :*

$$\lim_{\gamma \rightarrow 0} \lim_{N,n \rightarrow +\infty} \hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)} = \vartheta_0 \quad p.s.$$

Enfin, un dernier résultat de consistance de l'estimateur (1.3.36) utilisant le schéma à pas décroissant est le suivant.

Théorème 1.3.8. *On suppose (\mathbf{H}_0) , (\mathbf{C}_s) et (\mathbf{I}_w) satisfaites. Soit $p \geq 2$ et soit d une distance sur $\mathcal{M}_1(\mathbb{R}^d)$ appartenant \mathcal{D}_p . On considère le schéma d'Euler à pas décroissant dont la suite de pas $(\gamma_k)_{k \geq 1}$ satisfait la condition technique (4.4.3) (en plus des conditions déjà évoquées plus haut). Alors la famille $\{\hat{\vartheta}_{N,n}; N \geq 1, n \geq 1\}$ définie par (1.3.36) est un estimateur fortement consistant de ϑ_0 au sens suivant :*

$$\lim_{N,n \rightarrow +\infty} \hat{\vartheta}_{N,n} = \vartheta_0 \quad p.s.$$

Remarque 1.3.10. La condition technique (4.4.3) ($\sum_{k \geq 1} \gamma_k^{p'H+1} s_k^{-1} < +\infty$ pour un certain $p' \geq p$) est vérifiée pour un très grand ensemble de suites. Elle est liée à la variance des accroissements du mBf. En particulier, elle est satisfaite par toutes les suites de la forme $\gamma_k = \gamma k^{-\rho}$ avec $\rho \in (0, 1]$ ou encore les suites de la forme $\gamma_k = \gamma(\ln(k))^{-\beta}$ avec $\beta > 0$. En revanche, une suite de la forme $\gamma_k = \gamma(\ln(\ln(k)))^{-1}$ ne satisfait pas cette condition.

Les preuves de ces trois résultats de consistance reposent en grande partie sur les propriétés d'ergodicité des EDS (1.3.31) et du schéma d'Euler associé, résultats que l'on a décrits dans la Section 1.2. Le schéma de preuve de chacun de ces théorèmes est basé sur une propriété permettant de caractériser la convergence du minimum d'une famille de fonctions aléatoires $(\vartheta \mapsto L_r(\vartheta))_r$ vers le minimum de la limite de cette suite de fonctions : $\vartheta \mapsto L(\vartheta)$. Autrement dit, sous des hypothèses de convergence uniforme de $L_r(\vartheta)$ vers sa limite $L(\vartheta)$, on peut déduire la convergence du minimum de $L_r(\vartheta)$ vers le minimum de $L(\vartheta)$. Tout ceci est décrit précisément par la Proposition 4.4.3 tandis que la Section 4.4.3 explique comment ce résultat général est appliqué pour la preuve des Théorèmes de consistance 1.3.6, 1.3.7 et 1.3.8.

Vitesse de convergence

Pour une classe de distances sur $\mathcal{M}_1(\mathbb{R}^d)$ plus restreinte, il est possible d'obtenir des vitesses de convergence. Avant d'aller plus loin, précisons de quelles distances il s'agit.

Définition 1.3.2. On définit sur $\mathcal{M}_1(\mathbb{R}^d)$ les deux distances suivantes : soient ν et μ de mesures de probabilités sur \mathbb{R}^d ,

- (i) On considère le noyau intégrable $g_p(\xi) := c_p(1 + |\xi|^2)^{-p}$ où $c_p := \left(\int_{\mathbb{R}^d} (1 + |\xi|^2)^{-p} d\xi\right)^{-1}$. Alors, la distance $d_{CF,p}$ entre ν et μ est définie par

$$d_{CF,p}(\nu, \mu) := \left(\int_{\mathbb{R}^d} \left[\nu \left(e^{i\langle \xi, \cdot \rangle} \right) - \mu \left(e^{i\langle \xi, \cdot \rangle} \right) \right]^2 g_p(\xi) d\xi \right)^{1/2}. \quad (1.3.40)$$

- (ii) Soit $\{f_i; i \geq 1\}$ une famille de \mathcal{C}_b^1 , supposée dense dans l'espace \mathcal{C}_b^0 des fonctions continues, bornées et décroissant vers 0 à l'infini. Alors, la distance d_s entre ν et μ est définie par

$$d_s(\nu, \mu) := \sum_{i=0}^{+\infty} 2^{-i} (|\nu(f_i) - \mu(f_i)| \wedge 1). \quad (1.3.41)$$

Remarque 1.3.11. Ces deux distances métrisent la convergence étroite et appartiennent à \mathcal{D}_1 (défini par (1.3.37)). La première représente la distance L^2 pour un certain noyau

g_p entre les fonctions caractéristiques de ν et μ . La deuxième, appelée “weak- \star distance” dans [72], est utilisée notamment dans [76] pour des problèmes de filtrage. Ces distances sont sous la forme d’une intégrale ou d’une somme et ne font pas intervenir de supremum sur un ensemble de fonctions. C’est cet aspect qui permet de traiter le problème de la vitesse de convergence pour notre estimateur (1.3.35) lorsqu’il est défini avec l’une de ces deux distances.

Nous sommes maintenant en mesure d’énoncer le résultat relatif à la vitesse de convergence qui repose notamment sur l’hypothèse de minoration (\mathbf{I}_s).

Théorème 1.3.9. *On suppose (\mathbf{H}_0), (\mathbf{C}_s) et (\mathbf{I}_s) satisfaites, où (1.3.33) dans l’hypothèse (\mathbf{I}_s) est considérée pour les distances d_s ou $d_{CF,p}$ avec $p > (q + d)/2$ et $\varsigma = 2/q$ pour un certain $q \geq 2$. Soit $\hat{\vartheta}_{N,n,\gamma}$ l’estimateur défini par (1.3.35). Alors, il existe $C_q > 0$ telle que*

$$\mathbb{E} \left[|\hat{\vartheta}_{N,n,\gamma} - \vartheta_0|^2 \right] \leq C_q \left(n^{-\frac{q}{2}(2-(2H \vee 1))} + \gamma^{qH} + T^{-\tilde{\eta}} \right) \quad (1.3.42)$$

avec $\tilde{\eta} := \frac{q^2}{2(q+d)}(2 - (2H \vee 1))$ et $T := N\gamma$.

Remarque 1.3.12. En fait, on peut même montrer ce type de majoration pour tous les moments de $|\hat{\vartheta}_{N,n,\gamma} - \vartheta_0|$, la preuve s’adapte facilement. Cela revient juste à remplacer le 2 dans $\varsigma = 2/q$ par l’ordre du moment que l’on souhaite majorer. Nous avons choisi de nous restreindre à l’erreur quadratique pour simplifier la lecture.

Le Théorème 1.3.5 sur la concentration est un élément clé dans la preuve du Théorème 1.3.9. Afin de comprendre pourquoi, nous donnons maintenant quelques éléments de la preuve qui illustrent cet aspect. D’après l’inégalité (1.3.33) dans l’hypothèse (\mathbf{I}_s) appliquée à une distance donnée d (correspondant à d_s ou $d_{CF,p}$), majorer $\mathbb{E} \left[|\hat{\vartheta}_{N,n,\gamma} - \vartheta_0|^2 \right]$ revient à majorer

$$\mathbb{E} \left[d \left(\nu_{\hat{\vartheta}_{N,n,\gamma}}, \nu_{\vartheta_0} \right)^q \right]$$

avec $q = 2/\varsigma$. Ensuite, l’idée consiste à décomposer la quantité $d \left(\nu_{\hat{\vartheta}_{N,n,\gamma}}, \nu_{\vartheta_0} \right)$ en des termes faisant notamment intervenir la distance des mesures d’occupation (discrete et continue) à la mesure invariante, c’est-à-dire de la forme $d \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}} \right)$ et $d \left(\nu_{\vartheta}, \frac{1}{T} \int_0^T \delta_{Y_t} dt \right)$. C’est à ce stade que le Théorème 1.3.5 est primordial. Il permet de gérer la déviation des mesures d’occupation par rapport à leur moyenne (un résultat intermédiaire qui montre assez bien comment ce théorème est utilisé est la Proposition 4.5.1).

Les hypothèses d’identifiabilité (\mathbf{I}_w) et (\mathbf{I}_s) ne sont pas évidentes à vérifier en toute généralité. Cependant, le Chapitre 4 et plus spécifiquement les Sections 4.6 et 4.7 donnent des éléments de justifications théoriques et des illustrations numériques pour ces hypothèses d’identifiabilité. Plus précisément, la partie 4.6 montre que l’hypothèse (\mathbf{I}_s) pour la distance $d_{CF,p}$ est vérifiée pour des petites perturbations du processus d’Ornstein-Uhlenbeck fractionnaire. La partie 4.7 quant à elle, illustre par des simulations numériques (en dimension 1) l’hypothèse (\mathbf{I}_s) pour les distances $d_{CF,p}$ et \mathcal{W}_p pour un drift de la forme $b_{\vartheta}(x) = -\vartheta x$ ou $b_{\vartheta}(x) = -x(1 + \cos(\vartheta x))$.

Les trois prochains chapitres de cette thèse développent les résultats que nous venons d'énoncer et donnent les preuves de ces résultats. Chacun correspond à un article. Le Chapitre 2 contient les résultats d'ergodicité pour des dynamiques aléatoires discrètes à mémoire [71], le Chapitre 3 donnent les résultats sur la concentration pour la solution d'une EDS fractionnaire [70], tandis que le Chapitre 4 concerne les résultats d'estimation du drift pour des EDS fractionnaires [60], effectués en collaboration avec Panloup et Tindel.

2. Rate of convergence to equilibrium for discrete-time stochastic dynamics with memory

This chapter consists in the paper [71] accepted for publication to *Bernoulli Journal*.

Contents

2.1	Introduction	29
2.2	Setting and main results	32
2.3	Existence of invariant distribution	43
2.4	General coupling procedure	43
2.5	Coupling under $(\mathbf{H}_{\text{poly}})$ or $(\mathbf{H}_{\text{exp}})$	50
2.6	(K, α) -admissibility	57
2.7	Proof of Theorem 2.2.1	64
2.A	Proof of Theorem 2.2.2	67
2.B	Explicit formula for the sequence $(b_k)_{k \geq 0}$	70
2.C	Particular case: when the sequence $(a_k)_{k \geq 0}$ is log-convex	71
2.D	Proof of Proposition 2.2.3	72
2.E	Proof of Theorem 2.3.1	76
2.F	Proof of Lemma 2.5.3	78
2.G	Proof of Hypothesis (\mathbf{H}'_1)	80

2.1 Introduction

Convergence to equilibrium for Stochastic dynamics is one of the most natural and most studied problems in probability theory. Regarding Markov processes, this topic has been deeply undertaken through various approaches: spectral analysis, functional inequalities or coupling methods. However, in many applications (Physics, Biology, Finance...) the future evolution of a quantity may depend on its own history, and thus, noise with independent increments does not accurately reflect reality. A classical way to overcome this problem is to consider dynamical systems driven by processes with stationary increments like fractional Brownian motion (fBm) for instance which is widely used in applications (see e.g [32, 42, 45, 57]). In a continuous time framework,

Stochastic Differential Equations (SDEs) driven by Gaussian processes with stationary increments have been introduced to model random evolution phenomena with long range dependence properties. Consider SDEs of the following form

$$dX_t = b(X_t)dt + \sigma(X_t)dZ_t \quad (2.1.1)$$

where $(Z_t)_{t \geq 0}$ is a Gaussian process with ergodic stationary increments and $\sigma : \mathbb{R}^d \rightarrow \mathbb{M}_{d \times d}$, $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ are functions defined in a such a way that existence of a solution holds. The ergodic properties of such processes have been a topic of great interest over the last decade. For general Gaussian processes, existence and approximation of stationary solutions are provided in [15] in the additive case (i.e. when σ is constant). The specific situation where $(Z_t)_{t \geq 0}$ is a fractional Brownian motion has received significant attention since in the seminal paper [36] by Hairer, a definition of invariant distribution is given in the additive case through the embedding of the solution to an infinite-dimensional markovian structure. This point of view led to some probabilistic uniqueness criteria (on this topic, see *e.g.* [34, 38]) and to some coupling methods in view of the study of the rate of convergence to equilibrium. More precisely, in [36], some coalescent coupling arguments are also developed and lead to the convergence of the process in total variation to the stationary regime with a rate upper-bounded by $C_\varepsilon t^{-(\alpha_H - \varepsilon)}$ for any $\varepsilon > 0$, with

$$\alpha_H = \begin{cases} \frac{1}{8} & \text{if } H \in (\frac{1}{4}, 1) \setminus \{\frac{1}{2}\} \\ H(1 - 2H) & \text{if } H \in (0, \frac{1}{4}). \end{cases} \quad (2.1.2)$$

In the multiplicative noise setting (i.e. when σ is not constant), Fontbona and Panloup in [26] extended those results under selected assumptions on σ to the case where $H \in (\frac{1}{2}, 1)$ and finally Deya, Panloup and Tindel obtained in [22] this type of results in the rough setting $H \in (\frac{1}{3}, \frac{1}{2})$.

In this chapter, we focus on a general class of recursive discrete dynamics of the following form: we consider \mathbb{R}^d -valued sequences $(X_n)_{n \geq 0}$ satisfying

$$X_{n+1} = F(X_n, \Delta_{n+1}) \quad (2.1.3)$$

where $(\Delta_n)_{n \in \mathbb{Z}}$ is an ergodic stationary Gaussian sequence and $F : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a deterministic function. As a typical example for F , we can think about Euler discretization of (2.1.1) (see Subsection 2.2.5 for a detailed study) or to autoregressive processes in the particular case where F is linear. Note that such dynamics can be written as (2.1.3) through the so-called Wold's decomposition theorem which implies that we can see $(\Delta_n)_{n \in \mathbb{Z}}$ as a moving-average of infinite order (see [7] to get more details). To the best of our knowledge, in this linear setting the litterature mainly focuses on the statistical analysis of the model (see for instance [8, 9]) and on mixing properties of such Gaussian processes when it is in addition stationary (on this topic see *e.g.* [6, 44, 64]). Back to the main topic, namely ergodic properties of (2.1.3) for general functions F , Hairer in [37] has provided technical criteria using ergodic theory to get existence and uniqueness of invariant distribution for this kind of a priori non-Markovian processes. Here, the objective is to investigate the problem of the long-time behavior of

(2.1.3). To this end, we first explain how it is possible to define invariant distributions in this non-Markovian setting and to obtain existence results. More precisely, with the help of the moving average representation of the noise process $(\Delta_n)_{n \in \mathbb{Z}}$, we prove that $(X_n, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ can be realized through a Feller transformation \mathcal{Q} . In particular, an initial distribution of the dynamical system (X, Δ) is a distribution μ on $\mathbb{R}^d \times (\mathbb{R}^d)^{\mathbb{Z}^-}$. Rephrased in more probabilistic terms, an initial distribution is the distribution of the couple $(X_0, (\Delta_k)_{k \leq 0})_{n \geq 0}$. Then, such an initial distribution is called an invariant distribution if it is invariant by the transformation \mathcal{Q} . The first part of the main theorem establishes the existence of such an invariant distribution.

Then, our main contribution is to state a general result about the rate of convergence to equilibrium in terms of the covariance structure of the Gaussian process. To this end, we use a coalescent coupling strategy.

Let us briefly explain how this coupling method is organized in this discrete-time framework. First, one considers two processes $(X_n^1, (\Delta_{n+k}^1)_{k \leq 0})_{n \geq 0}$ and $(X_n^2, (\Delta_{n+k}^2)_{k \leq 0})_{n \geq 0}$ following (2.1.3) starting respectively from μ_0 and μ_\star (an invariant distribution). As a preliminary step, one waits that the two paths get close. Then, at each trial, the coupling attempt is divided in two steps. First, one tries in Step 1 to stick the positions together at a given time. Then, in Step 2, one attempts to ensure that the paths stay clustered until $+\infty$. Actually, oppositely to the Markovian setting where the paths remain naturally fastened together (by putting the same innovation on each marginal), the main difficulty here is that, staying together has a cost. In other words, this property can be ensured only with a non trivial coupling of the noises. Finally, if one of the two previous steps fails, one begins Step 3 by putting the same noise on each coordinate until the “cost” to attempt Step 1 is not too big. More precisely, during this step one waits for the memory of the coupling cost to decrease sufficiently and for the two trajectories to be in a compact set with lower-bounded probability.

In the main theorem previously announced, as a result of this strategy, we are able to prove that the law of the process $(X_{n+k})_{k \geq 0}$ following (2.1.3) converges in total variation to the stationary regime with a rate upper-bounded by Cn^{-v} . The quantity v is directly linked to the assumed exponential or polynomial asymptotic decay of the sequence involved in the moving-average representation of $(\Delta_n)_{n \in \mathbb{Z}}$, see (2.2.2) (or equivalently in its covariance function, see Remark 2.2.1 and 2.2.3).

Then, we apply our main theorem to fractional memory (including the Euler Scheme associated to fractional Stochastic Differential Equations). We first emphasize that with covariance structures with the same order of memory but different local behavior, we can get distinct rates of convergence to equilibrium. Secondly, we highlight that the computation of the asymptotic decay of the sequence involved in the inversion of the Toeplitz operator (related to the moving-average representation) can be a very technical task (see proof of Proposition 2.2.3).

Now, let us discuss about specific contributions of this discrete-time approach. Above all, our result is quite general since, for instance, it includes discretization of (2.1.1) for a large class of Gaussian noise processes. Then, in several ways, we get a further understanding of arguments used in the coupling procedure. We better target the impact of the memory through the sequence both appearing in the moving-average representation and the covariance function of the Gaussian noise. Regarding Step 1 of the coupling

strategy, the “admissibility condition” (which means that we are able to attempt Step 1 with a controlled cost) is rather more explicit than in the continuous-time setting. Finally, this chapter, by deconstructing the coupling method through this explicit discrete-time framework, may weigh in favour of the sharpness of Hairer’s approach.

The following section gives more details on the studied dynamics, describes the assumptions required to get the main result, namely Theorem 2.2.1 and discuss about the application of our main result to the case of fractional memory in Subsection 2.2.7. Then, the proof of Theorem 2.2.1 is achieved in Sections 2.3, 2.4, 2.5, 2.6 and 2.7, which are outlined at the end of Section 2.2.

2.2 Setting and main results

2.2.1 Notations

The usual scalar product on \mathbb{R}^d is denoted by $\langle \cdot, \cdot \rangle$ and $|\cdot|$ stands either for the Euclidean norm on \mathbb{R}^d or the absolute value on \mathbb{R} . We denote by $\mathbb{M}_{d \times d}$ the space of real matrices of size $d \times d$. For a given $K > 0$ we denote by $B(0, K)$ the \mathbb{R}^d closed ball centered in 0 with radius K . Then, the *state space* of the process X and the *noise space* associated to $((\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ are respectively denoted by $\mathcal{X} := \mathbb{R}^d$ and $\mathcal{W} := (\mathbb{R}^d)^{\mathbb{Z}^-}$. For a given measurable space $(\mathcal{X}_1, \mathcal{A}_1)$, $\mathcal{M}_1(\mathcal{X}_1)$ will denote the set of all probability measures on \mathcal{X}_1 . Let $(\mathcal{X}_2, \mathcal{A}_2)$ be an other measurable space and $f : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ be a measurable mapping. Let $\mu \in \mathcal{M}_1(\mathcal{X}_1)$, we denote by $f^* \mu$ the pushforward measure given by :

$$\forall B \in \mathcal{A}_2, \quad f^* \mu(B) := \mu(f^{-1}(B)).$$

$\Pi_{\mathcal{X}} : \mathcal{X} \times \mathcal{W} \rightarrow \mathcal{X}$ and $\Pi_{\mathcal{W}} : \mathcal{X} \times \mathcal{W} \rightarrow \mathcal{W}$ stand respectively for the projection on the marginal \mathcal{X} and \mathcal{W} . For a given differentiable function $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and for all $x \in \mathbb{R}^d$ we will denote by $J_f(x)$ the Jacobian matrix of f valued at point x . Finally, we denote by $\|\cdot\|_{TV}$ the classical total variation norm: let $\nu, \mu \in \mathcal{M}_1(\mathcal{X}_1)$,

$$\|\mu - \nu\|_{TV} := \sup_{A \subset \mathcal{X}_1} |\mu(A) - \nu(A)|.$$

2.2.2 Dynamical system and Markovian structure

Let $X := (X_n)_{n \geq 0}$ denote an \mathbb{R}^d -valued sequence defined by: X_0 is a random variable with a given distribution and

$$\forall n \geq 0, \quad X_{n+1} = F(X_n, \Delta_{n+1}), \tag{2.2.1}$$

where $F : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is continuous and $(\Delta_n)_{n \in \mathbb{Z}}$ is a stationary and purely non-deterministic Gaussian sequence with d independent components. Hence, by Wold’s decomposition theorem [7] it has a moving average representation

$$\forall n \in \mathbb{Z}, \quad \Delta_n = \sum_{k=0}^{+\infty} a_k \xi_{n-k} \tag{2.2.2}$$

with

$$\begin{cases} (a_k)_{k \geq 0} \in \mathbb{R}^{\mathbb{N}} \text{ such that } a_0 \neq 0 \text{ and } \sum_{k=0}^{+\infty} a_k^2 < +\infty \\ (\xi_k)_{k \in \mathbb{Z}} \text{ an i.i.d sequence such that } \xi_1 \sim \mathcal{N}(0, I_d). \end{cases} \quad (2.2.3)$$

Without loss of generality, we assume that $a_0 = 1$. Actually, if $a_0 \neq 1$, we can come back to this case by setting

$$\tilde{\Delta}_n = \sum_{k=0}^{+\infty} \tilde{a}_k \xi_{n-k}$$

with $\tilde{a}_k := \frac{a_k}{a_0}$.

Remark 2.2.1. The asymptotic behavior of the sequence $(a_k)_{k \geq 0}$ certainly plays a key role to compute the rate of convergence to equilibrium of the process $(X_n)_{n \geq 0}$. Actually, the memory induced by the noise process is quantified by the sequence $(a_k)_{k \geq 0}$ through the identity

$$\forall n \in \mathbb{Z}, \forall k \geq 0, \quad c(k) := \mathbb{E}[\Delta_n \Delta_{n+k}] = \sum_{i=0}^{+\infty} a_i a_{k+i}.$$

The stochastic dynamics described in (2.2.1) is clearly non-Markovian. Let us see how it is possible to introduce a Markovian structure and how to define invariant distribution. This method is inspired by [37]. The first idea is to look at $(X_n, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ instead of $(X_n)_{n \geq 0}$. Let us introduce the following concatenation operator

$$\begin{aligned} \sqcup : \mathcal{W} \times \mathbb{R}^d &\rightarrow \mathcal{W} \\ (w, w') &\mapsto w \sqcup w' \end{aligned} \quad (2.2.4)$$

where $(w \sqcup w')_0 = w'$ and $\forall k < 0, (w \sqcup w')_k = w_{k+1}$. Then (2.2.1) is equivalent to the system

$$(X_{n+1}, (\Delta_{n+1+k})_{k \leq 0}) = \varphi((X_n, (\Delta_{n+k})_{k \leq 0}), \Delta_{n+1}) \quad (2.2.5)$$

where

$$\begin{aligned} \varphi : (\mathcal{X} \times \mathcal{W}) \times \mathbb{R}^d &\rightarrow \mathcal{X} \times \mathcal{W} \\ ((x, w), w') &\mapsto (F(x, w'), w \sqcup w'). \end{aligned}$$

Therefore, $(X_n, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ can be realized through the Feller Markov transition kernel \mathcal{Q} defined by

$$\int_{\mathcal{X} \times \mathcal{W}} g(x', w') \mathcal{Q}((x, w), (dx', dw')) = \int_{\mathbb{R}^d} g(\varphi((x, w), \delta)) \mathcal{P}(w, d\delta) \quad (2.2.6)$$

where $\mathcal{P}(w, d\delta) := \mathcal{L}(\Delta_{n+1} | (\Delta_{n+k})_{k \leq 0} = w)$ does not depend on n since (Δ_n) is a stationary sequence, and $g : \mathcal{X} \times \mathcal{W} \rightarrow \mathbb{R}$ is a measurable function.

Definition 2.2.1. A measure $\mu \in \mathcal{M}_1(\mathcal{X} \times \mathcal{W})$ is said to be an invariant distribution for (2.2.5) and then for (2.2.1) if it is invariant by \mathcal{Q} , i.e.

$$\mathcal{Q}\mu = \mu.$$

However, the concept of uniqueness will be slightly different from the classical setting. Indeed, denote by $\mathcal{S}\mu$ the distribution of $(X_n^\mu)_{n \geq 0}$ when we realize $(X_n^\mu, (\Delta_{n+k})_{k \leq 0})_{n \geq 0}$ through the transition \mathcal{Q} and with initial distribution μ . Then, we will speak of uniqueness of the invariant distribution up to the equivalence relation: $\mu \sim \nu \iff \mathcal{S}\mu = \mathcal{S}\nu$. Moreover, here uniqueness will be deduced from the coupling procedure. There exist some results about uniqueness using ergodic theory, like in [37], but they will be not outlined here.

2.2.3 Preliminary tool: a Toeplitz type operator

The moving-average representation of the Gaussian sequence $(\Delta_n)_{n \in \mathbb{Z}}$ naturally leads us to define an operator related to the coefficients $(a_k)_{k \geq 0}$. First, set

$$\ell_a(\mathbb{Z}^-, \mathbb{R}^d) := \left\{ w \in (\mathbb{R}^d)^{\mathbb{Z}^-} \mid \forall k \geq 0, \left| \sum_{l=0}^{+\infty} a_l w_{-k-l} \right| < +\infty \right\}$$

and define \mathbf{T}_a on $\ell_a(\mathbb{Z}^-, \mathbb{R}^d)$ by

$$\mathbf{T}_a(w) = \left(\sum_{l=0}^{+\infty} a_l w_{-k-l} \right)_{k \geq 0}. \quad (2.2.7)$$

Due to the Cauchy-Schwarz inequality, we can check that for instance $\ell^2(\mathbb{Z}^-, \mathbb{R}^d) := \{w \in (\mathbb{R}^d)^{\mathbb{Z}^-} \mid \sum_{k=0}^{+\infty} |w_{-k}|^2 < +\infty\}$ is included in $\ell_a(\mathbb{Z}^-, \mathbb{R}^d)$ due to the assumption $\sum_{k \geq 0} a_k^2 < +\infty$. This Toeplitz type operator \mathbf{T}_a links $(\Delta_n)_{n \in \mathbb{Z}}$ to $(\xi_n)_{n \in \mathbb{Z}}$. The following proposition spells out the reverse operator.

Proposition 2.2.1. *Let \mathbf{T}_b be the operator defined on $\ell_b(\mathbb{Z}^-, \mathbb{R}^d)$ in the same way as \mathbf{T}_a but with the following sequence $(b_k)_{k \geq 0}$*

$$b_0 = \frac{1}{a_0} \quad \text{and} \quad \forall k \geq 1, \quad b_k = -\frac{1}{a_0} \sum_{l=1}^k a_l b_{k-l}. \quad (2.2.8)$$

Then,

$$\forall w \in \ell_a(\mathbb{Z}^-, \mathbb{R}^d), \quad \mathbf{T}_b(\mathbf{T}_a(w)) = w \quad \text{and} \quad \forall w \in \ell_b(\mathbb{Z}^-, \mathbb{R}^d), \quad \mathbf{T}_a(\mathbf{T}_b(w)) = w$$

that is $\mathbf{T}_b = \mathbf{T}_a^{-1}$ and $\ell_b(\mathbb{Z}^-, \mathbb{R}^d) = \mathbf{T}_a(\ell_a(\mathbb{Z}^-, \mathbb{R}^d))$.

Proof. Let $w \in \ell_a(\mathbb{Z}^-, \mathbb{R}^d)$. Then let $n \geq 0$,

$$\begin{aligned} (\mathbf{T}_b(\mathbf{T}_a(w)))_{-n} &= \sum_{k=0}^{+\infty} b_k \sum_{l=0}^{+\infty} a_l w_{-n-k-l} = \sum_{k=0}^{+\infty} \sum_{i=k}^{+\infty} b_k a_{i-k} w_{-n-i} \\ &= \sum_{i=0}^{+\infty} \underbrace{\left(\sum_{k=0}^i b_k a_{i-k} \right)}_{=0 \text{ except for } i=0} w_{-n-i} = w_{-n}. \end{aligned}$$

We show in the same way that for $w \in \ell_b(\mathbb{Z}^-, \mathbb{R}^d)$, we have $\mathbf{T}_a(\mathbf{T}_b(w)) = w$. \square

Remark 2.2.2. The sequence $(b_k)_{k \geq 0}$ is of first importance in the sequel. The sketch of the proof of Theorem 2.2.1 will use an important property linked to the sequence $(b_k)_{k \geq 0}$: if two sequences u and v are such that

$$\forall n \geq 1, \quad u_n = \sum_{k=0}^{n-1} a_k v_{n-k}$$

then,

$$\forall n \geq 1, \quad v_n = \sum_{k=0}^{n-1} b_k u_{n-k}.$$

This reverse identity and the asymptotic behavior of $(b_k)_{k \geq 0}$ play a significant role in the computation of the rate of convergence.

The following section is devoted to outline assumptions on $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ and then on F to get the main result.

2.2.4 Assumptions and general theorem

First of all, let us introduce assumptions on $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$. All along the chapter, we will switch between two types of assumptions called respectively the *polynomial case* and the *exponential case*.

Hypothesis (\mathbf{H}_{poly}): The following conditions hold,

- there exist $\rho, \beta > 1/2$ and $C_\rho, C_\beta > 0$ such that

$$\forall k \geq 0, \quad |a_k| \leq C_\rho (k+1)^{-\rho} \quad \text{and} \quad \forall k \geq 0, \quad |b_k| \leq C_\beta (k+1)^{-\beta}.$$

- there exist $\kappa \geq \rho + 1$ and $C_\kappa > 0$ such that

$$\forall k \geq 0, \quad |a_k - a_{k+1}| \leq C_\kappa (k+1)^{-\kappa}.$$

Hypothesis (\mathbf{H}_{exp}): There exist $\lambda, \zeta > 0$ and $C_\lambda, C_\zeta > 0$ such that,

$$\forall k \geq 0, \quad |a_k| \leq C_\lambda e^{-\lambda k} \quad \text{and} \quad \forall k \geq 0, \quad |b_k| \leq C_\zeta e^{-\zeta k}.$$

Remark 2.2.3. \triangleright (\mathbf{H}_{poly}) and (\mathbf{H}_{exp}) are general parametric hypothesis which apply to a large class of Gaussian driven dynamics. These assumptions implicitly involve the covariance function of the noise process $(c(k))_{k \in \mathbb{N}}$ (see Remark 2.2.1) : there exists $\tilde{C}_\lambda > 0$ and for all $\varepsilon \in [0, \rho]$ such that $\rho + \varepsilon > 1$, there exists $\tilde{C}_{\rho, \varepsilon} > 0$ such that

$$\forall k \geq 0, \quad |c(k)| \leq \begin{cases} \tilde{C}_{\rho, \varepsilon} (k+1)^{-\rho + \varepsilon} & \text{under } (\mathbf{H}_{\text{poly}}) \\ \tilde{C}_\lambda e^{-\lambda k} & \text{under } (\mathbf{H}_{\text{exp}}) \end{cases}.$$

\triangleright (\mathbf{H}_{poly}) and (\mathbf{H}_{exp}) also involve the coefficients appearing in the reverse Toeplitz operator \mathbf{T}_a^{-1} (see Proposition 2.2.1). Even though $(a_k)_{k \geq 0}$ and $(b_k)_{k \geq 0}$ are related by

(2.2.8), there is no general rule which connects ρ and β . This fact will be highlighted in Subsection 2.2.7. Moreover, for the sake of clarity, we have chosen to state our main result when (a_k) and (b_k) belong to the same family of asymptotic decay rate.

▷ Due to the strategy of the proof (coalescent coupling in a non Markovian setting) we also need a bound on the discrete derivative of $(a_k)_{k \geq 0}$.

Let us now introduce some assumptions on the function F which defines the dynamics (2.2.1). Throughout this chapter $F : \mathcal{X} \times \mathbb{R}^d \rightarrow \mathcal{X}$ is a continuous function and the following hypothesis **(H₁)** and **(H₂)** are satisfied.

Hypothesis (H₁): There exists a continuous function $V : \mathcal{X} \rightarrow \mathbb{R}_+^*$ satisfying $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$ and $\exists \gamma \in (0, 1)$ and $C > 0$ such that for all $(x, w) \in \mathcal{X} \times \mathbb{R}^d$,

$$V(F(x, w)) \leq \gamma V(x) + C(1 + |w|). \quad (2.2.9)$$

Remark 2.2.4. ▷ As we will see in Section 2.3, this condition ensures the existence of a Lyapunov function V and then of an invariant distribution. Such a type of assumption also appears in the litterature of discrete Markov Chains (see e.g. equation (7) in [24]) but in an integrated form. More precisely, in our non-Markovian setting, a pathwise control is required to ensure some control on the moments of the trajectories before the successful coupling procedure (this fact is detailed in Subsection 2.6.2).

▷ This assumption is also fulfilled if we have a function V with $V(F(x, w)) \leq \gamma V(x) + C(1 + |w|^p)$ for a given $p > 1$ instead of (2.2.9) since the function $V^{1/p}$ satisfies **(H₁)** (with the help of the elementary inequality $|u + v|^{1/p} \leq |u|^{1/p} + |v|^{1/p}$ when $p > 1$).

We define $\tilde{F} : \mathcal{X} \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathcal{X}$ by $\tilde{F}(x, u, y) = F(x, u + y)$. We assume that \tilde{F} satisfies the following conditions:

Hypothesis (H₂): Let $K > 0$. We assume that there exists $\tilde{K} > 0$ such that for every $\mathbf{x} := (x, x', y, y')$ in $B(0, K)^4$, there exist $\Lambda_{\mathbf{x}} : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $M_K > 0$ and $C_{\tilde{K}}$ such that the following holds

- $\Lambda_{\mathbf{x}}$ is a bijection from \mathbb{R}^d to \mathbb{R}^d . Moreover, it is a \mathcal{C}^1 -diffeomorphism between two open sets U and D such that $\mathbb{R}^d \setminus U$ and $\mathbb{R}^d \setminus D$ are negligible sets.
- for all $u \in B(0, \tilde{K})$,

$$\tilde{F}(x, u, y) = \tilde{F}(x', \Lambda_{\mathbf{x}}(u), y') \quad (2.2.10)$$

$$\text{and } |\det(J_{\Lambda_{\mathbf{x}}}(u))| \geq C_{\tilde{K}}. \quad (2.2.11)$$

- for all $u \in \mathbb{R}^d$,

$$|\Lambda_{\mathbf{x}}(u) - u| \leq M_K. \quad (2.2.12)$$

Remark 2.2.5. ▷ Let us make a few precisions on the arguments of \tilde{F} : x is the position of the process, u the increment of the innovation process and y is related to the past of the process (see next item for more details). The boundary $C_{\tilde{K}}$ and M_K are independent from x, x', y and y' . This assumption can be viewed as a kind of controllability assumption in the following sense: the existence of $\Lambda_{\mathbf{x}}$ leads to the coa-

lescence of the positions by (2.2.10). Rephrased in terms of our coalescent coupling strategy, this *ad hoc* assumption is required to achieve the first step. More precisely, as announced in the introduction, we take two trajectories (X^1, Δ^1) and (X^2, Δ^2) following (2.2.1) and we want to stick X^1 and X^2 at a given time $n + 1$. Through the function $\Lambda_{\mathbf{x}}$ in (\mathbf{H}_2) , we can build a couple of Gaussian innovations $(\xi_{n+1}^1, \xi_{n+1}^2)$ with marginal distribution $\mathcal{N}(0, I_d)$ to achieve this goal (with lower-bounded probability), so that: $X_{n+1}^1 = X_{n+1}^2$ which is equivalent to $\tilde{F}(x, \xi_{n+1}^1, y) = \tilde{F}(x', \xi_{n+1}^2, y')$ with $(x, x', y, y') = (X_n^1, X_n^2, \sum_{k=1}^{+\infty} a_k \xi_{n+1-k}^1, \sum_{k=1}^{+\infty} a_k \xi_{n+1-k}^2)$ (see Subsection 2.4.2).

▷ Assumption (\mathbf{H}_2) can be applied to a large class of functions F , as for example: $F(x, w) = f(b(x) + \sigma(x)w)$ where σ is continuously invertible and σ^{-1} and b are continuous functions (we do not need any assumption on f as we will see in the appendix Remark 2.A.2). Actually, Condition (2.2.10) can be obtained through an application of the implicit function theorem: if we assume that there exists a point $(0, u'_x)$ such that $\tilde{F}(x, 0, y) = \tilde{F}(x', u'_x, y')$ and denote by $G_{\mathbf{x}} : (u, u') \mapsto \tilde{F}(x, u, y) - \tilde{F}(x', u' + u'_x, y')$, then if $(\partial_{u'} G_{\mathbf{x}})(0, 0) \neq 0$, the implicit function theorem yields (2.2.10). As we will see in Subsection 2.2.5, Condition (2.2.12) can be also easily fulfilled (see proof of Theorem 2.2.2).

We are now in position to state our main result.

Theorem 2.2.1. *Assume (\mathbf{H}_1) and (\mathbf{H}_2) . Then,*

- (i) *There exists an invariant distribution μ_{\star} associated to (2.2.1).*
- (ii) *Assume that $(\mathbf{H}_{\text{poly}})$ is true with $\rho, \beta > 1/2$ and $\rho + \beta > 3/2$. Then, uniqueness holds for the invariant distribution μ_{\star} . Furthermore, for every initial distribution μ_0 for which $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_0(dx) < +\infty$ and for all $\varepsilon > 0$, there exists $C_{\varepsilon} > 0$ such that*

$$\|\mathcal{L}((X_{n+k}^{\mu_0})_{k \geq 0}) - \mathcal{S}\mu_{\star}\|_{TV} \leq C_{\varepsilon} n^{-(v(\beta, \rho) - \varepsilon)}.$$

where the function v is defined by

$$v(\beta, \rho) = \sup_{\alpha \in (\frac{1}{2} \vee (\frac{3}{2} - \beta), \rho)} \min\{1, 2(\rho - \alpha)\} (\min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2).$$

- (iii) *Assume that $(\mathbf{H}_{\text{exp}})$ is true, then uniqueness holds for the invariant distribution μ_{\star} . Furthermore, for every initial distribution μ_0 for which $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_0(dx) < +\infty$ and for all $p > 0$, there exists $C_p > 0$ such that*

$$\|\mathcal{L}((X_{n+k}^{\mu_0})_{k \geq 0}) - \mathcal{S}\mu_{\star}\|_{TV} \leq C_p n^{-p}.$$

Remark 2.2.6. ▷ In view of Theorem 2.2.1 (iii), one can wonder if we could obtain exponential or subexponential rates of convergence in this case. We focus on this question in Remark 2.6.1.

▷ The rates obtained in Theorem 2.2.1 hold for a large class of dynamics. This generality implies that the rates are not optimal in all situations. In particular, when F have “nice” properties an adapted method could lead to better rates. For example, let us

mention the particular case where the dynamical system (2.2.1) is reduced to: $X_{n+1} = AX_n + \sigma\Delta_{n+1}$ where A and σ are some given matrices. As for (fractional) Ornstein-Uhlenbeck processes in a continuous setting, the study of linear dynamics can be achieved with specific methods. Here, the sequence X benefits of a Gaussian structure so that the convergence in distribution could be studied through the covariance of the process. One can also remark that since for two paths X and \tilde{X} built with the same noise, we have: $\tilde{X}_{n+1} - X_{n+1} = A(\tilde{X}_n - X_n)$, a simple induction leads to $\mathbb{E}[|\tilde{X}_n - X_n|^2] \leq |A^n|^2 \mathbb{E}[|\tilde{X}_0 - X_0|^2]$. So without going into the details, if $\rho(A) := \lim_{n \rightarrow +\infty} |A^n|^{1/n} < 1$, such bounds lead to geometric rates of convergence in Wasserstein distance and also in total variation distance (on this topic, see e.g. [59]).

In the following subsection, we test the assumptions of our main result Theorem 2.2.1 (especially (\mathbf{H}_1) and (\mathbf{H}_2)) on the Euler scheme of SDEs like (2.1.1).

2.2.5 The Euler Scheme

Recall that $\mathcal{X} = \mathbb{R}^d$. In this subsection, set

$$\begin{aligned} F_h : \quad \mathcal{X} \times \mathbb{R}^d &\rightarrow \mathcal{X} \\ (x, w) &\mapsto x + hb(x) + \sigma(x)w. \end{aligned} \quad (2.2.13)$$

where $h > 0$, $b : \mathcal{X} \rightarrow \mathcal{X}$ is continuous and $\sigma : \mathcal{X} \rightarrow \mathbb{M}_{d \times d}$ is a continuous and bounded function on \mathcal{X} . For all $x \in \mathcal{X}$ we suppose that $\sigma(x)$ is invertible and we denote by $\sigma^{-1}(x)$ the inverse. Moreover, we assume that σ^{-1} is a continuous function and that b satisfies a Lyapunov type assumption that is:

(L1) $\exists C > 0$ such that

$$\forall x \in \mathcal{X}, \quad |b(x)| \leq C(1 + |x|). \quad (2.2.14)$$

(L2) $\exists \tilde{\beta} \in \mathbb{R}$ and $\tilde{\alpha} > 0$ such that

$$\forall x \in \mathcal{X}, \quad \langle x, b(x) \rangle \leq \tilde{\beta} - \tilde{\alpha}|x|^2. \quad (2.2.15)$$

Remark 2.2.7. This function F_h corresponds to the Euler scheme associated to SDEs like (2.1.1). The conditions on the function b are classical to get existence of invariant distribution.

In this setting the function \tilde{F}_h (introduced in Hypothesis (\mathbf{H}_2)) is given by

$$\begin{aligned} \tilde{F}_h : \quad \mathcal{X} \times \mathbb{R}^d \times \mathbb{R}^d &\rightarrow \mathcal{X} \\ (x, u, y) &\mapsto x + hb(x) + \sigma(x)(u + y). \end{aligned}$$

Theorem 2.2.2. *Let $h > 0$. Let F_h be the function defined above. Assume that $b : \mathcal{X} \rightarrow \mathcal{X}$ is a continuous function satisfying **(L1)** and **(L2)** and $\sigma : \mathcal{X} \rightarrow \mathbb{M}_{d \times d}$ is a continuous and bounded function such that for all $x \in \mathcal{X}$, $\sigma(x)$ is invertible and $x \mapsto \sigma^{-1}(x)$ is a continuous function. Then, **(H1)** and **(H2)** hold for F_h as soon as $0 < h < \min \left\{ \sqrt{1 + \frac{\tilde{\alpha}}{2C^2}} - 1, \frac{1}{\tilde{\alpha}} \right\}$ where $\tilde{\alpha}$ and C are given by **(L1)** and **(L2)**.*

Proof. For the sake of conciseness, the proof is detailed in Appendix 2.A. Regarding (\mathbf{H}_1) , it makes use of ideas developed in [15, 16]. For (\mathbf{H}_2) , the construction of $\Lambda_{\mathbf{x}}$ is explicit. The idea is to take $\Lambda_{\mathbf{x}}(u) := \sigma^{-1}(x')\sigma(x)u + \sigma^{-1}(x')(x - x' + h(b(x) - b(x'))) + \sigma^{-1}(x')\sigma(x)y - y'$ inside $B(0, K)$ (which ensures (2.2.10) with $\tilde{K} = K$), to set $\Lambda_{\mathbf{x}}(u) := u$ outside an other ball $B(0, K_1)$ with a well chosen K_1 (which almost gives (2.2.12)) and to extend $\Lambda_{\mathbf{x}}$ into \mathbb{R}^d by taking into account the various hypothesis on $\Lambda_{\mathbf{x}}$. \square

Remark 2.2.8. The coefficient σ is assumed to be bounded but we can relax a bit this assumption: if $|\sigma(x)| \leq C(1 + |x|^\kappa)$ for some $\kappa \in (0, 1)$, then Theorem 2.2.2 is true again. We provide in Remark 2.A.2 the necessary adjustments of the proof to get this result.

The two following subsections are devoted to outline examples of sequences which satisfy hypothesis $(\mathbf{H}_{\text{exp}})$ or $(\mathbf{H}_{\text{poly}})$. In particular, Subsection 2.2.7 includes the case where the process $(\Delta_n)_{n \in \mathbb{Z}}$ corresponds to fractional Brownian motion increments.

2.2.6 Two explicit cases which satisfy $(\mathbf{H}_{\text{exp}})$

First, let us mention the explicit exponential case with the following definition for the sequence $(a_k)_{k \geq 0}$

$$a_0 = 1 \quad \text{and} \quad \forall k \in \mathbb{N}^*, \quad a_k = C_a e^{-\lambda k} \quad (2.2.16)$$

with $C_a \in \mathbb{R}$. Let us recall that $b_0 = 1$ (since $a_0 = 1$) and for all $k \geq 1$, we can get the following general expression of b_k (see Appendix 2.B for more details):

$$b_k = \sum_{p=1}^k \frac{(-1)^p}{a_0^{p+1}} \left(\sum_{\substack{k_1, \dots, k_p \geq 1 \\ k_1 + \dots + k_p = k}} \prod_{i=1}^p a_{k_i} \right). \quad (2.2.17)$$

A classical combinatorial argument shows that $\#\{(k_1, \dots, k_p) \in \mathbb{N}^* \mid k_1 + \dots + k_p = k\} = \binom{k-1}{p-1}$. As a consequence, when the sequence $(a_k)_{k \geq 0}$ is defined by (2.2.16), we can easily prove that for $k \geq 1$,

$$b_k = -C_a(1 - C_a)^{k-1} e^{-\lambda k}.$$

Hence, to satisfy $(\mathbf{H}_{\text{exp}})$, we only need C_a to be such that $\zeta := \lambda - \ln|1 - C_a| > 0$ and then for all $k \in \mathbb{N}^*$, we get

$$|b_k| \leq C_b e^{-\zeta k} \quad (2.2.18)$$

with $C_b > 0$ a constant depending on C_a .

Remark 2.2.9. \triangleright In this setting where everything is computable, it's interesting to see that the asymptotic decrease of the sequence $(|b_k|)$ is not only related to the one of the sequence $(|a_k|)$. For instance, if we take $C_a < 0$, the simple fact that $a_0 > 0$ and $a_k < 0$ for all $k > 0$ makes (b_k) diverge to $+\infty$ and nevertheless, $(|a_k|)$ decreases to 0 at an exponential rate.

\triangleright If we take $C_a = 1$, we can reduce $(\Delta_n)_{n \in \mathbb{Z}}$ to the following induction: $\forall n \in \mathbb{Z}, \Delta_{n+1} = \xi_{n+1} + e^{-\lambda} \Delta_n$.

Let us finally consider finite moving averages, i.e. when $a_k = 0$ for all $k > m$ (for some $m \geq 1$). In this setting, one can expect $(\mathbf{H}_{\text{exp}})$ to be satisfied since the memory is finite. This is actually the case when the finite moving average is invertible, namely: $P(\lambda) := 1 + \sum_{k=1}^m a_k \lambda^k$ has all its roots outside the unit circle (see [7] Theorem 3.1.2). In that case, there exists $\lambda \in \mathbb{C}$ such that $|\lambda| > 1$ and $\frac{1}{P(\lambda)} = \sum_{k=0}^{+\infty} b_k \lambda^k < +\infty$ (to get more details on this equality, see Appendix 2.B). Then, there exists $C > 0$ such that for all $k \geq 0$,

$$|b_k| \leq C \left(\frac{1}{|\lambda|} \right)^k$$

and finally $(\mathbf{H}_{\text{exp}})$ holds true.

When the invertibility is not fulfilled, the situation is more involved but $(\mathbf{H}_{\text{exp}})$ is still true up to another Wold decomposition. More precisely, one can find another white noise ξ and another set of coefficients a_k such that the invertibility holds true, on this topic see e.g. [7] Proposition 4.4.2.

2.2.7 Polynomial case: from a general class of examples to the fractional case

A natural example of Gaussian sequence $(\Delta_n)_{n \in \mathbb{Z}}$ which leads to polynomial memory is to choose $a_k = (k+1)^{-\rho}$ for a given $\rho > 1/2$. In that case, we have the following result.

Proposition 2.2.2. *Assume (\mathbf{H}_1) and (\mathbf{H}_2) . Let $\rho > 1/2$. If for all $k \geq 0$, $a_k = (k+1)^{-\rho}$, then we have $|b_k| \leq (k+1)^{-\rho}$. Moreover, if $\rho > 3/4$ Theorem 2.2.1 (ii) holds with the rate*

$$v(\rho, \rho) = \begin{cases} 2(\rho - 3/4)^2 & \text{if } \rho \in (3/4, 1] \\ \frac{1}{2}(\rho - 1/2)^2 & \text{if } \rho > 1. \end{cases}$$

Remark 2.2.10. The main novelty here comes from the proof of the inequality $|b_k| \leq (k+1)^{-\rho}$ for all $k \geq 0$ which is outlined in Appendix 2.C and which is based on results of [27]. The key argument in this proof is the log-convexity property of the sequence $(a_k)_{k \in \mathbb{N}}$, which means that for all $k \in \mathbb{N}$, $a_k \geq 0$ and for $k \geq 1$, $a_k^2 - a_{k-1}a_{k+1} \leq 0$.

With Proposition 2.2.2 in hand, the purpose of the remainder of this section is to focus on Gaussian sequences of *fractional type*, i.e. when the sequence $(a_k)_{k \geq 0}$ satisfies:

$$\forall k \geq 0, \quad |a_k| \leq C_\rho (k+1)^{-\rho} \quad \text{and} \quad |a_k - a_{k+1}| \leq \tilde{C}_\rho (k+1)^{-(\rho+1)} \quad (2.2.19)$$

with $\rho := 3/2 - H$ and $H \in (0, 1)$ is the so-called Hurst parameter. In particular, through this class of sequences, we provide an explicit example which shows that computing the rate of convergence of the sequence $(b_k)_{k \geq 0}$ is a hard task and strongly depends on the variations of $(a_k)_{k \geq 0}$. Condition (2.2.19) includes both cases of Proposition 2.2.2 with $\rho := 3/2 - H$ and when $(\Delta_n)_{n \in \mathbb{Z}}$ corresponds to the fractional Brownian motion (fBm) increments (as we will see below), we therefore decided to use the terminology “*fractional type*”. Recall that a d -dimensional fBm with Hurst parameter $H \in (0, 1)$ is a centered Gaussian process $(B_t)_{t \geq 0}$ with stationary increments satisfying

$$\forall t, s \geq 0, \quad \forall i, j \in \{1, \dots, d\}, \quad \mathbb{E} \left[(B_t^i - B_s^i)(B_t^j - B_s^j) \right] = \delta_{ij} |t - s|^{2H}.$$

In our discrete-time setting, we are thus concerned by the long time behavior of (2.2.1) if we take for $h > 0$

$$(\Delta_n)_{n \in \mathbb{Z}} = (B_{nh} - B_{(n-1)h})_{n \in \mathbb{Z}} \quad (2.2.20)$$

which is a stationary Gaussian sequence. It can be realized through a moving average representation with coefficients $(a_k^H)_{k \geq 0}$ defined by (see [58]):

$$a_0^H = h^H \kappa(H) 2^{1/2-H} \quad \text{and for } k \geq 1, \quad a_k^H = h^H \kappa(H) \left(\left(k + \frac{1}{2}\right)^{H-1/2} - \left(k - \frac{1}{2}\right)^{H-1/2} \right) \quad (2.2.21)$$

where

$$\kappa(H) = \frac{\sqrt{\sin(\pi H) \Gamma(2H + 1)}}{\Gamma(H + 1/2)}.$$

One can easily check that $a_k^H \underset{k \rightarrow +\infty}{\sim} C_{h,H} (k+1)^{-(3/2-H)}$ and $|a_k^H - a_{k+1}^H| \leq C'_{h,H} (k+1)^{-(5/2-H)}$.

Hence $(a_k^H)_{k \geq 0}$ is of *fractional type* in the sense of (2.2.19). Now, the question is: how does the corresponding (b_k^H) behave? When H belongs to $(0, 1/2)$, only a_0^H is positive and then (a_k^H) is not log-convex. Therefore, we cannot use this property to get the asymptotic behavior of (b_k^H) as we did in Proposition 2.2.2. However, thanks to simulations (see Figure 1a and 1b), we conjectured and we proved Proposition 2.2.3.

Proposition 2.2.3. *There exists $C''_{h,H} > 0$ such that for all $H \in (0, 1/2)$*

$$\forall k \geq 0, \quad |b_k^H| \leq C''_{h,H} (k+1)^{-(H+1/2)}. \quad (2.2.22)$$

Then, if we assume (\mathbf{H}_1) and (\mathbf{H}_2) , Theorem 2.2.1 (ii) holds with the rate

$$v(\rho, 2 - \rho) = v(3/2 - H, H + 1/2) = \begin{cases} H(1 - 2H) & \text{if } H \in (0, 1/4] \\ \frac{1}{8} & \text{if } H \in (1/4, 1/2). \end{cases}$$

Remark 2.2.11. For the sake of conciseness, we provided the details of the proof in Appendix 2.D. The proof of (2.2.22) is based on a technical induction which involves very sharp inequalities.

In Proposition 2.2.2 (with $\rho = 3/2 - H$ and $H \in (0, 1/2)$) and in the above proposition, dealing with the same order of memory, we get really different orders of rate of convergence: one easily checks that for all $H \in (0, 1/2)$, $v(3/2 - H, H + 1/2) < v(3/2 - H, 3/2 - H)$. Finally, we have seen that managing the asymptotic behavior of (b_k) is both essential and a difficult task.

To end this section, let us briefly discuss on the specific statements on fBm increments and compare with the continuous time setting (see [36, 26, 22]). For this purpose, we introduce the following conjecture (based on simulations, see Figure 1c and 1d) when H belongs to $(1/2, 1)$:

Conjecture: There exists $C''_{h,H} > 0$ and $\beta_H > 1$ such that

$$\forall k \geq 0, \quad |b_k^H| \leq C''_{h,H} (k+1)^{-\beta_H}. \quad (2.2.23)$$

Remark 2.2.12. We do not have a precise idea of the expression of β_H with respect to H . But, we can note that if $\rho < 1$ and $\beta > 1$ in $(\mathbf{H}_{\text{poly}})$, then the rate of convergence in Theorem 2.2.1 is $v(\rho, \beta) = \frac{(2\rho-1)^2}{8}$ and does not depend on β . Hence, if $H \in (1/2, 1)$, $\rho = 3/2 - H$ and $\beta_H > 1$, we fall into this case and then the dependence of β_H in terms of H does not matter.

If the previous conjecture is true we get the following rate of convergence for $H \in (1/2, 1)$ in Theorem 2.2.1:

$$v(\rho, \beta_H) = v(3/2 - H, \beta_H) = \frac{(1-H)^2}{2}.$$

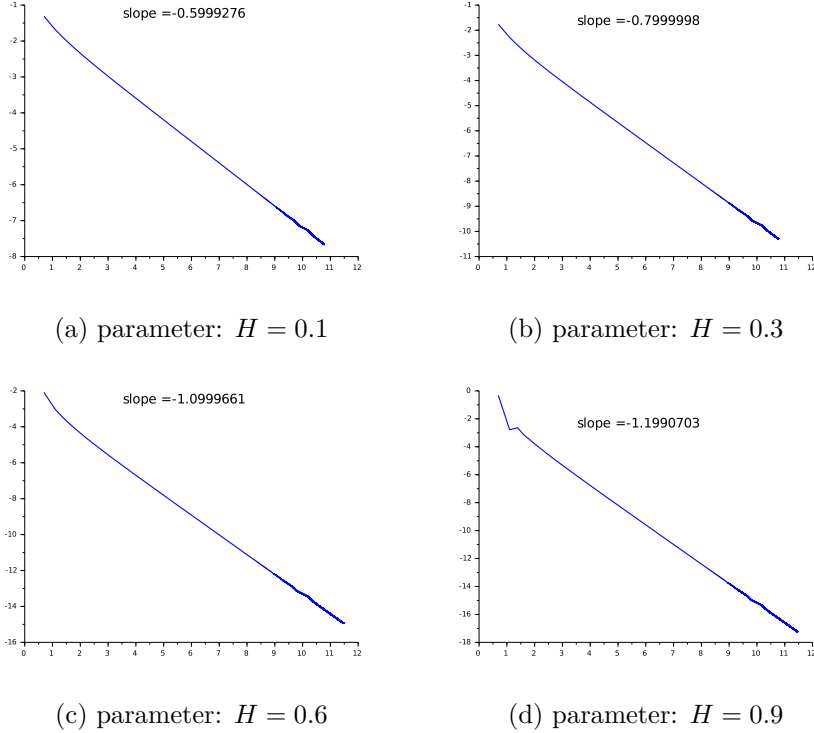


Figure 1 – $(\log |b_k^H|)$ according to $(\log(k+1))$ with different Hurst parameters H .

Then, when H belongs to $(0, 1/2)$ Proposition 2.2.3 gives exactly the same rate of convergence obtained in [36, 22]. However, when $H > 1/2$ it seems that we will get a smaller rate than in a continuous time setting. The reason for this may be that Theorem 2.2.1 is a result with quite general hypothesis on the Gaussian noise process $(\Delta_n)_{n \in \mathbb{Z}}$. In the case of fBm increments, the moving average representation is explicit. Hence, we may use a more specific approach and significantly closer to Hairer's, especially with

regard to Step 2 in the coupling method (see Subsection 2.5.2) by not exploiting the technical lemma 2.5.3 for instance. This seems to be a right track in order to improve our results on this precise example.

We are now ready to begin the proof of Theorem 2.2.1. In Section 2.3, we establish the first part of the theorem, i.e. (i). Then, in Section 2.4 we explain the scheme of coupling before implementing this strategy in Sections 2.5 and 2.6. Finally, in Section 2.7, we achieve the proof of (ii) and (iii) of Theorem 2.2.1.

2.3 Existence of invariant distribution

Denote by \mathbb{P}_w the law of $(\Delta_k)_{k \leq 0}$. Since $(\Delta_n)_{n \in \mathbb{Z}}$ is stationary we immediately get the following property:

Property 2.3.1. *If a measure $\mu \in \mathcal{M}_1(\mathcal{X} \times \mathcal{W})$ is such that $\Pi_{\mathcal{W}}^* \mu = \mathbb{P}_w$ then $\Pi_{\mathcal{W}}^* \mathcal{Q} \mu = \mathbb{P}_w$.*

We can now define the notion of Lyapunov function.

Definition 2.3.1. A function $\psi : \mathcal{X} \rightarrow [0, +\infty)$ is called a Lyapunov function for \mathcal{Q} if ψ is continuous and if the following holds:

(i) $\psi^{-1}([0, a])$ is compact for all $a \in [0, +\infty)$.

(ii) $\exists \beta > 0$ and $\alpha \in (0, 1)$ such that:

$$\int_{\mathcal{X} \times \mathcal{W}} \psi(x) \mathcal{Q} \mu(dx, dw) \leq \beta + \alpha \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mu)(dx)$$

for all $\mu \in \mathcal{M}_1(\mathcal{X} \times \mathcal{W})$ such that $\Pi_{\mathcal{W}}^* \mu = \mathbb{P}_w$ and $\int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mu)(dx) < +\infty$.

The following result ensures the existence of invariant distribution for \mathcal{Q} .

Theorem 2.3.1. *If there exists a Lyapunov function ψ for \mathcal{Q} , then \mathcal{Q} has at least one invariant distribution μ_\star , in other words $\mathcal{Q} \mu_\star = \mu_\star$.*

A detailed proof of this result is given in Appendix 2.E. Finally, we get the first part (i) of Theorem 2.2.1 about the existence of an invariant distribution by setting $\psi := V$ (with V the function appearing in (\mathbf{H}_1)) and by saying that ψ is a Lyapunov function for \mathcal{Q} .

2.4 General coupling procedure

We now turn to the proof of the main result of the chapter, i.e. Theorem 2.2.1 (ii) and (iii) about the convergence in total variation. This result is based on a coupling method first introduced in [36], but also used in [26] and [22], in a continuous time framework. The coupling strategy is slightly different in our discrete context, the following part is devoted to explain this procedure.

2.4.1 Scheme of coupling

Let $(\Delta_n^1)_{n \in \mathbb{Z}}$ and $(\Delta_n^2)_{n \in \mathbb{Z}}$ be two stationary and purely non-deterministic Gaussian sequences with the following moving average representations

$$\begin{cases} \Delta_n^1 = \sum_{k=0}^{+\infty} a_k \xi_{n-k}^1 \\ \Delta_n^2 = \sum_{k=0}^{+\infty} a_k \xi_{n-k}^2 \end{cases}$$

with

$$\begin{cases} (a_k)_{k \geq 0} \in \mathbb{R}^{\mathbb{N}} \text{ such that } a_0 = 1 \text{ and } \sum_{k=0}^{+\infty} a_k^2 < +\infty \\ \xi^i := (\xi_k^i)_{k \in \mathbb{Z}} \text{ an i.i.d sequence such that } \xi_1^i \sim \mathcal{N}(0, I_d) \text{ for } i = 1, 2. \end{cases} \quad (2.4.1)$$

We denote by (X^1, X^2) the solution of the system:

$$\begin{cases} X_{n+1}^1 = F(X_n^1, \Delta_{n+1}^1) \\ X_{n+1}^2 = F(X_n^2, \Delta_{n+1}^2) \end{cases} \quad (2.4.2)$$

with initial conditions $(X_0^1, (\Delta_k^1)_{k \leq 0})$ and $(X_0^2, (\Delta_k^2)_{k \leq 0})$. We assume that $(X_0^2, (\Delta_k^2)_{k \leq 0}) \sim \mu_\star$ where μ_\star denotes a fixed invariant distribution associated to (2.2.1). The previous section ensures that such a measure exists. We define the natural filtration associated to (2.4.2) by

$$(\mathcal{F}_n)_{n \in \mathbb{N}} = (\sigma((\xi_k^1)_{k \leq n}, (\xi_k^2)_{k \leq n}, X_0^1, X_0^2))_{n \geq 0}.$$

To lower the ‘‘weight of the past’’ at the beginning of the coupling procedure, we assume that a.s,

$$(\Delta_k^1)_{k \leq 0} = (\Delta_k^2)_{k \leq 0}$$

which is actually equivalent to assume that a.s $(\xi_k^1)_{k \leq 0} = (\xi_k^2)_{k \leq 0}$ since the invertible Toeplitz operator defined in Subsection 2.2.3 links $(\Delta_k^i)_{k \leq 0}$ to $(\xi_k^i)_{k \leq 0}$ for $i = 1, 2$. Lastly, we denote by $(g_n)_{n \in \mathbb{Z}}$ and $(f_n)_{n \in \mathbb{Z}}$ the random variable sequences defined by

$$\xi_{n+1}^1 = \xi_{n+1}^2 + g_n \quad \text{and} \quad \Delta_{n+1}^1 = \Delta_{n+1}^2 + f_n. \quad (2.4.3)$$

They respectively represent the ‘‘drift’’ between the underlying noises (ξ_k^i) and the real noises (Δ_k^i) . By assumption, we have $g_n = f_n = 0$ for $n < 0$.

Remark 2.4.1. From the moving average representations, we deduce immediately the following relation for all $n \geq 0$,

$$f_n = \sum_{k=0}^{+\infty} a_k g_{n-k} = \sum_{k=0}^n a_k g_{n-k}. \quad (2.4.4)$$

The aim is now to build $(g_n)_{n \geq 0}$ and $(f_n)_{n \geq 0}$ in order to stick X^1 and X^2 . We set

$$\tau_\infty = \inf\{n \geq 0 \mid X_k^1 = X_k^2, \forall k \geq n\}.$$

In a purely Markovian setting, when the paths coincide at time n then they remain stuck for all $k \geq n$ by putting the same innovation into both processes. Due to the memory this phenomenon cannot happen here. Hence, this involves a new step in the coupling scheme: try to keep the paths fastened together (see below).

Recall that $\mathcal{L}((X_k^2)_{k \geq n}) = \mathcal{S}\mu_\star$. The purpose of the coupling procedure is to bound the quantity $\mathbb{P}(\tau_\infty > n)$ since by a classical result we have

$$\|\mathcal{L}((X_k^1)_{k \geq n}) - \mathcal{S}\mu_\star\|_{TV} \leq \mathbb{P}(\tau_\infty > n). \quad (2.4.5)$$

Hence, we realize the coupling after a series of trials which follows three steps:

- * **Step 1:** Try to stick the positions at a given time with a “controlled cost”.
- * **Step 2:** (specific to non-Markov processes) Try to keep the paths fastened together.
- * **Step 3:** If Step 2 fails, we wait long enough so as to allow Step 1 to be realized with a “controlled cost” and with a positive probability. During this step, we assume that $g_n = 0$.

More precisely, let us introduce some notations,

- Let $\tau_0 \geq 0$. We begin the first trial at time $\tau_0 + 1$, in other words we try to stick $X_{\tau_0+1}^1$ and $X_{\tau_0+1}^2$. Hence, we assume that

$$\forall n < \tau_0, \quad g_n = f_n = 0. \quad (2.4.6)$$

- For $j \geq 1$, let τ_j denote the end of trial j . More specifically,
 - ▷ If $\tau_j = +\infty$ for some $j \geq 1$, it means that the coupling tentative has been successful.
 - ▷ Else, τ_j corresponds to the end of Step 3, that is $\tau_j + 1$ is the beginning of Step 1 of trial $j + 1$.

The real meaning of “controlled cost” will be clarified on Subsection 2.5.1. But the main idea is that at Step 1 of trial j , the “cost” is represented by the quantity $g_{\tau_{j-1}}$ that we need to build to get $X_{\tau_{j-1}+1}^1 = X_{\tau_{j-1}+1}^2$ with positive probability. Here the cost does not only depend on the positions at time τ_{j-1} but also on all the past of the underlying noises ξ^1 and ξ^2 . Hence, we must have a control on $g_{\tau_{j-1}}$ in case of failure and to this end we have to wait enough during Step 3 before beginning a new attempt of coupling.

2.4.2 Coupling lemmas to achieve Step 1 and 2

This section is devoted to establish coupling lemmas in order to build (ξ^1, ξ^2) during Step 1 and Step 2.

Hitting step

If we want to stick X^1 and X^2 at time $n + 1$, we need to build $(\xi_{n+1}^1, \xi_{n+1}^2)$ in order to get $F(X_n^1, \Delta_{n+1}^1) = F(X_n^2, \Delta_{n+1}^2)$ with positive probability, that is to get

$$\begin{aligned} F\left(X_n^1, \xi_{n+1}^1 + \sum_{k=1}^{+\infty} a_k \xi_{n+1-k}^1\right) &= F\left(X_n^2, \xi_{n+1}^2 + \sum_{k=1}^{+\infty} a_k \xi_{n+1-k}^2\right) \\ \iff \tilde{F}\left(X_n^1, \xi_{n+1}^1, \sum_{k=1}^{+\infty} a_k \xi_{n+1-k}^1\right) &= \tilde{F}\left(X_n^2, \xi_{n+1}^2, \sum_{k=1}^{+\infty} a_k \xi_{n+1-k}^2\right). \end{aligned} \quad (2.4.7)$$

The following lemma will be the main tool to achieve this goal.

Lemma 2.4.1. *Let $K > 0$ and $\mu := \mathcal{N}(0, I_d)$. Under the controllability assumption (\mathbf{H}_2) , there exists $\tilde{K} > 0$ (given by (\mathbf{H}_2)), such that for every $\mathbf{x} := (x, x', y, y')$ in $B(0, K)^4$, we can build a random variable (Z_1, Z_2) with values in $(\mathbb{R}^d)^2$ such that*

$$(i) \quad \mathcal{L}(Z_1) = \mathcal{L}(Z_2) = \mu,$$

(ii) *there exists $\delta_{\tilde{K}} > 0$ depending only on \tilde{K} such that*

$$\mathbb{P}(\tilde{F}(x, Z_1, y) = \tilde{F}(x', Z_2, y')) \geq \mathbb{P}(Z_2 = \Lambda_{\mathbf{x}}(Z_1), |Z_1| \leq \tilde{K}) \geq \delta_{\tilde{K}} > 0 \quad (2.4.8)$$

where $\Lambda_{\mathbf{x}}$ is the function given by hypothesis (\mathbf{H}_2) ,

(iii) *there exists $M_K > 0$ given by (\mathbf{H}_2) depending only on K such that*

$$\mathbb{P}(|Z_2 - Z_1| \leq M_K) = 1. \quad (2.4.9)$$

Proof. Let $\mathbf{x} := (x, x', y, y') \in B(0, K)^4$. First, let us denote by π_1 (resp. π_2) the projection from $\mathbb{R}^d \times \mathbb{R}^d$ to \mathbb{R}^d of the first (resp. the second) coordinate. Introduce the two following functions defined on \mathbb{R}^d

$$\begin{aligned} \Lambda_1 : u_1 &\mapsto (u_1, \Lambda_{\mathbf{x}}(u_1)) \\ \Lambda_2 : u_2 &\mapsto (\Lambda_{\mathbf{x}}^{-1}(u_2), u_2) \end{aligned}$$

where $\Lambda_{\mathbf{x}}$ is the function given by (\mathbf{H}_2) . Now, we set

$$\mathbf{P}_1 = \frac{1}{2}(\Lambda_1^* \mu \wedge \Lambda_2^* \mu).$$

Let us find a simplest expression for \mathbf{P}_1 . For every measurable function $f : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_+$, we have

$$\Lambda_1^* \mu(f) = \int_{\mathbb{R}^d} f(u_1, \Lambda_{\mathbf{x}}(u_1)) \mu(du_1) = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(u_1, u_2) \delta_{\Lambda_{\mathbf{x}}(u_1)}(du_2) \mu(du_1)$$

and

$$\begin{aligned}
 \Lambda_2^* \mu(f) &= \int_{\mathbb{R}^d} f(\Lambda_{\mathbf{x}}^{-1}(u_2), u_2) \mu(du_2) \\
 &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(\Lambda_{\mathbf{x}}^{-1}(u_2), u_2) \exp\left(-\frac{|u_2|^2}{2}\right) du_2 \\
 &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(u_1, \Lambda_{\mathbf{x}}(u_1)) \exp\left(-\frac{|\Lambda_{\mathbf{x}}(u_1)|^2}{2}\right) |\det(J_{\Lambda_{\mathbf{x}}}(u_1))| du_1 \quad (\text{by setting } u_1 = \Lambda_{\mathbf{x}}^{-1}(u_2)) \\
 &= \int_{\mathbb{R}^d} f(u_1, \Lambda_{\mathbf{x}}(u_1)) \underbrace{\exp\left(\frac{|u_1|^2}{2} - \frac{|\Lambda_{\mathbf{x}}(u_1)|^2}{2}\right) |\det(J_{\Lambda_{\mathbf{x}}}(u_1))|}_{=: D_{\Lambda_{\mathbf{x}}}(u_1)} \mu(du_1) \\
 &= \int_{\mathbb{R}^d \times \mathbb{R}^d} f(u_1, u_2) \delta_{\Lambda_{\mathbf{x}}(u_1)}(du_2) D_{\Lambda_{\mathbf{x}}}(u_1) \mu(du_1).
 \end{aligned}$$

By construction, we then have

$$\mathbf{P}_1(du_1, du_2) = \frac{1}{2} \delta_{\Lambda_{\mathbf{x}}(u_1)}(du_2) (D_{\Lambda_{\mathbf{x}}}(u_1) \wedge 1) \mu(du_1). \quad (2.4.10)$$

Write $S(u_1, u_2) = (u_2, u_1)$ and denote by $\tilde{\mathbf{P}}_1$ the ‘‘symmetrized’’ non-negative measure induced by \mathbf{P}_1 ,

$$\tilde{\mathbf{P}}_1 = \mathbf{P}_1 + S^* \mathbf{P}_1. \quad (2.4.11)$$

We then define (Z_1, Z_2) as follows:

$$\mathcal{L}(Z_1, Z_2) = \tilde{\mathbf{P}}_1 + \Delta^*(\mu - \pi_1^* \tilde{\mathbf{P}}_1) = \mathbf{P}_1 + \mathbf{P}_2 \quad (2.4.12)$$

with $\Delta(u) = (u, u)$ and $\mathbf{P}_2 = S^* \mathbf{P}_1 + \Delta^*(\mu - \pi_1^* \tilde{\mathbf{P}}_1)$. It remains to prove that $\mathcal{L}(Z_1, Z_2)$ is well defined and satisfies all the properties required by the lemma.

First step: Prove that \mathbf{P}_2 is the sum of two non-negative measures.

Using (2.4.10), we can check that for all non-negative function f ,

$$\pi_1^* \mathbf{P}_1(f) \leq \frac{1}{2} \mu(f)$$

and

$$\pi_2^* \mathbf{P}_1(f) = \pi_1^*(S^* \mathbf{P}_1)(f) \leq \frac{1}{2} \mu(f).$$

By adding the two previous inequalities, we deduce that the measure $\mu - \pi_1^* \tilde{\mathbf{P}}_1$ is non-negative. This concludes the first step.

Second step: Prove that $\pi_1^*(\mathbf{P}_1 + \mathbf{P}_2) = \pi_2^*(\mathbf{P}_1 + \mathbf{P}_2) = \mu$.

This fact is almost obvious. We just need to use the fact that

$$\pi_1 \circ \Delta = \pi_2 \circ \Delta = \text{Id}$$

and the symmetry property of $\tilde{\mathbf{P}}_1$,

$$\text{i.e. } \pi_1^* \tilde{\mathbf{P}}_1 = \pi_2^* \tilde{\mathbf{P}}_1.$$

Third step: Prove (2.4.8) and (2.4.9).

Let us first remark that the support of $\mathbf{P}_1 + \mathbf{P}_2$ is included in

$$\{(u, v) \in \mathbb{R}^d \times \mathbb{R}^d \mid v = \Lambda_{\mathbf{x}}(u)\} \cup \{(u, v) \in \mathbb{R}^d \times \mathbb{R}^d \mid v = \Lambda_{\mathbf{x}}^{-1}(u)\} \cup \{(u, v) \in \mathbb{R}^d \times \mathbb{R}^d \mid v = u\}.$$

Therefore, by (2.2.12) in (\mathbf{H}_2) and the fact that

$$(\forall u \in \mathbb{R}^d, |\Lambda_{\mathbf{x}}(u) - u| \leq M_K) \iff (\forall u \in \mathbb{R}^d, |\Lambda_{\mathbf{x}}^{-1}(u) - u| \leq M_K)$$

since $\Lambda_{\mathbf{x}}$ is invertible on \mathbb{R}^d , we finally get (2.4.9).

Then, using again (\mathbf{H}_2) where \tilde{K} is defined and the definition of the subprobability \mathbf{P}_1 we get

$$\mathbb{P}(\tilde{F}(x, Z_1, y) = \tilde{F}(x', Z_2, y')) \geq \underbrace{\mathbf{P}_1(B(0, \tilde{K}) \times \Lambda(B(0, \tilde{K})))}_{=\mathbb{P}(Z_2 = \Lambda_{\mathbf{x}}(Z_1), |Z_1| \leq \tilde{K})} \quad (2.4.13)$$

and

$$\mathbb{P}(Z_2 = \Lambda_{\mathbf{x}}(Z_1), |Z_1| \leq \tilde{K}) = \frac{1}{2} \int_{B(0, \tilde{K})} (D_{\Lambda_{\mathbf{x}}}(u) \wedge 1) \mu(du).$$

It just remains to use (2.2.11) and (2.2.12) in (\mathbf{H}_2) to conclude. Indeed,

$$\begin{aligned} \mathbb{P}(Z_2 = \Lambda_{\mathbf{x}}(Z_1), |Z_1| \leq \tilde{K}) &= \frac{1}{2} \int_{B(0, \tilde{K})} \left(\exp\left(\frac{|u|^2}{2} - \frac{|\Lambda_{\mathbf{x}}(u)|^2}{2}\right) |\det(J_{\Lambda_{\mathbf{x}}}(u))| \right) \wedge 1 \mu(du) \\ &\geq \frac{1}{2} \mu(B(0, \tilde{K})) \left[\left(\exp\left(-\frac{(M_{\tilde{K}} + \tilde{K})^2}{2}\right) C_{\tilde{K}} \right) \wedge 1 \right] =: \delta_{\tilde{K}} > 0 \end{aligned}$$

which concludes the proof. \square

Sticking step

Now, if the positions X_{n+1}^1 and X_{n+1}^2 are stuck together, we want that they remain fastened together for all $k > n + 1$ which means that:

$$\begin{aligned} \forall k \geq n + 1, \quad F(X_k^1, \Delta_{k+1}^1) &= F(X_k^2, \Delta_{k+1}^2) \\ \iff \forall k \geq n + 1, \quad F(X_k^1, \xi_{k+1}^1 + \sum_{l=1}^{+\infty} a_l \xi_{k+1-l}^1) &= F(X_k^1, \xi_{k+1}^2 + \sum_{l=1}^{+\infty} a_l \xi_{k+1-l}^2) \quad (2.4.14) \end{aligned}$$

since $X_k^1 = X_k^2$. Recall that for all $k \in \mathbb{Z}$, $g_k = \xi_{k+1}^1 - \xi_{k+1}^2$ is the drift between the underlying noises. Then, if we have

$$\begin{aligned} \forall k \geq n+1, \quad \xi_{k+1}^1 + \sum_{l=1}^{+\infty} a_l \xi_{k+1-l}^1 &= \xi_{k+1}^2 + \sum_{l=1}^{+\infty} a_l \xi_{k+1-l}^2 \\ \iff \forall k \geq n+1, \quad g_k &= - \sum_{l=1}^{+\infty} a_l g_{k-l}, \end{aligned} \quad (2.4.15)$$

the identity (2.4.14) is automatically satisfied.

Remark 2.4.2. The successful g_k defined by relation (2.4.15) is \mathcal{F}_k -measurable. This explains why we chose to index it by k even if it represents the drift between ξ_{k+1}^1 and ξ_{k+1}^2 .

Hence, we will try to get (2.4.15) on successive finite intervals to finally get a bound on the successful-coupling probability. The size choice of those intervals will be important according to the hypothesis (\mathbf{H}_{poly}) or (\mathbf{H}_{exp}) that we made. The two next results will be our tools to get (2.4.15) on Subsection 2.5.2. For the sake of simplicity we set out these results on \mathbb{R} . On \mathbb{R}^d we just have to apply them on every marginal. Lemma 2.4.2 is almost the statement of Lemma 5.13 of [36] or Lemma 3.2 of [26].

Lemma 2.4.2. *Let $\mu := \mathcal{N}(0, 1)$. Let $a \in \mathbb{R}$, $b \geq |a|$ and $M_b := \max(4b, -2 \log(b/8))$.*

(i) *For all $b \geq |a|$, there exist δ_b^1 and $\delta_b^2 \in (0, 1)$, such that we can build a probability measure $\mathcal{N}_{a,b}^2$ on \mathbb{R}^2 with every marginal equal to μ and such that:*

$$\mathcal{N}_{a,b}^2(\{(x, y) \mid y = x + a\}) \geq \delta_b^1 \quad \text{and} \quad \mathcal{N}_{a,b}^2(\{(x, y) \mid |y - x| \leq M_b\}) = 1.$$

(ii) *Moreover, if $b \in (0, 1)$, the previous statement holds with $\delta_b^1 = 1 - b$.*

The following corollary is an adapted version of Lemma 3.3 of [26] to our discrete context.

Corollary 2.4.1. *Let $T > 0$ be an integer, $b > 0$, $g = (g_0, g_1, \dots, g_T) \in \mathbb{R}^{T+1}$ such that $\|g\| \leq b$ where $\|\cdot\|$ is the euclidian norm on \mathbb{R}^{T+1} and set $M_b := \max(4b, -2 \log(b/8))$.*

(i) *Then, there exists $\delta_b^1 \in (0, 1)$, for which we can build a random variable $((\xi_{k+1}^1)_{k \in \llbracket 0, T \rrbracket}, (\xi_{k+1}^2)_{k \in \llbracket 0, T \rrbracket})$ with values in $(\mathbb{R}^{T+1})^2$, with marginal distribution $\mathcal{N}(0, I_{T+1})$ and satisfying:*

$$\mathbb{P}(\xi_{k+1}^1 = \xi_{k+1}^2 + g_k \quad \forall k \in \llbracket 0, T \rrbracket) \geq \delta_b^1$$

and

$$\mathbb{P}(\|\xi^1 - \xi^2\| \leq M_b) = 1.$$

(ii) *Moreover, if $b \in (0, 1)$, the previous statement holds with $\delta_b^1 = 1 - b$.*

Proof. Let $(u_k)_{k \in \llbracket 0, T \rrbracket}$ be an orthonormal basis of \mathbb{R}^{T+1} with $u_0 = \frac{g}{\|g\|}$. We denote by (U_1, U_2) a random variable which has distribution $\mathcal{N}_{a,b}^2$ (with $a = \|g\|$) given in the

lemma 2.4.2. Let $(\varepsilon_k)_{k \in \llbracket 1, T \rrbracket}$ be an iid random variable sequence with $\varepsilon_1 \sim \mathcal{N}(0, 1)$ and independent from (U_1, U_2) . Then, for $i = 1, 2$ we define the isometry:

$$\begin{aligned} \mathcal{W}^i : \mathbb{R}^{T+1} &\rightarrow \mathcal{W}^i(\mathbb{R}^{T+1}) \subset L^2(\Omega, \mathcal{F}, \mathbb{P}) \\ u_0 &\mapsto U_i \\ u_k &\mapsto \varepsilon_k \text{ for } k \in \llbracket 1, T \rrbracket. \end{aligned} \quad (2.4.16)$$

And we set for all $n \in \llbracket 0, T \rrbracket$, $\xi_{n+1}^i := \mathcal{W}^i(e_n)$ where e_n is the vector of \mathbb{R}^{T+1} for which every coordinate is 0 except the $(n+1)^{\text{th}}$ which is 1. Since $(u_k)_{k \in \llbracket 0, T \rrbracket}$ is an orthonormal basis of \mathbb{R}^{T+1} , we then have:

$$e_n = \sum_{k=0}^T \langle e_n, u_k \rangle u_k.$$

Hence,

$$\xi_{n+1}^i = \mathcal{W}^i \left(\sum_{k=0}^T \langle e_n, u_k \rangle u_k \right) = U_i \frac{g_n}{\|g\|} + \sum_{k=1}^T \langle e_n, u_k \rangle \varepsilon_k.$$

ξ_{n+1}^i is clearly centered and Gaussian as a linear combination of independent centered Gaussian random variables and using that \mathcal{W}^i is an isometry, we get that $(\xi_{k+1}^i)_{k \in \llbracket 0, T \rrbracket}$ has distribution $\mathcal{N}(0, I_{T+1})$ for $i = 1, 2$. Therefore, we built ξ^1 and ξ^2 as announced. Indeed, by Lemma 2.4.2

$$\mathbb{P}(\xi_{n+1}^1 = \xi_{n+1}^2 + g_n \quad \forall n \in \llbracket 0, T \rrbracket) = \mathbb{P}(U_1 = U_2 + \|g\|) \geq \delta_b^1$$

and

$$\mathbb{P}(\|\xi^1 - \xi^2\| \leq M_b) = \mathbb{P}(|U_1 - U_2| \leq M_b) = 1.$$

(ii) also follows immediately from Lemma 2.4.2. □

2.5 Coupling under $(\mathbf{H}_{\text{poly}})$ or $(\mathbf{H}_{\text{exp}})$

We can now move on the real coupling procedure to finally get a lower-bound for the successful-coupling probability. In a first subsection, we explain exactly what we called “controlled cost” and in a second subsection we spell out our bound.

2.5.1 Admissibility condition

The “controlled cost” is called “admissibility” in [36]. Here, we will talk about (K, α) -admissibility, as in [26], but in the following sense:

Definition 2.5.1. Let $K > 0$ and $\alpha > 0$ be two constants and τ a random variable with values in \mathbb{N} . We say that the system is (K, α) -admissible at time τ if $\tau(\omega) < +\infty$ and if $(X_\tau^1(\omega), X_\tau^2(\omega), (\xi_n^1(\omega), \xi_n^2(\omega))_{n \leq \tau})$ satisfies

$$\forall n \geq 0, \left| \sum_{k=n+1}^{+\infty} a_k g_{\tau+n-k}(\omega) \right| \leq v_n \quad (2.5.1)$$

and

$$|X_\tau^i(\omega)| \leq K, \quad \left| \sum_{k=1}^{+\infty} a_k \xi_{\tau+1-k}^i(\omega) \right| \leq K \quad \text{for } i = 1, 2 \quad (2.5.2)$$

with

$$v_n = (n+1)^{-\alpha} \text{ under } (\mathbf{H}_{\text{poly}}) \quad \text{and} \quad v_n = e^{-\alpha n} \text{ under } (\mathbf{H}_{\text{exp}}). \quad (2.5.3)$$

Remark 2.5.1. On the one hand, condition (2.5.1) measures the distance between the past of the noises (before time τ). On the other hand, condition (2.5.2) has two parts: the first one ensures that at time τ both processes are not far from each other and the second part is a constraint on the memory part of the Gaussian noise $\Delta_{\tau+1}^i$.

The aim is to prove that under those two conditions, the coupling will be successful with a probability lower-bounded by a positive constant. To this end, we will need to ensure that at every time τ_j , the system will be (K, α) -admissible with positive probability. We set:

$$\Omega_{\alpha, \tau}^1 := \left\{ \omega, \tau(\omega) < +\infty, \left| \sum_{k=n+1}^{+\infty} a_k g_{\tau+n-k}(\omega) \right| \leq v_n \quad \forall n \in \mathbb{N} \right\} \quad (2.5.4)$$

and

$$\Omega_{K, \tau}^2 := \left\{ \omega, \tau(\omega) < +\infty, |X_\tau^i(\omega)| \leq K \text{ and } \left| \sum_{k=1}^{+\infty} a_k \xi_{\tau+1-k}^i(\omega) \right| \leq K \quad \text{for } i = 1, 2 \right\}. \quad (2.5.5)$$

We define

$$\Omega_{K, \alpha, \tau} = \Omega_{\alpha, \tau}^1 \cap \Omega_{K, \tau}^2. \quad (2.5.6)$$

If $\omega \in \Omega_{K, \alpha, \tau}$, we will try to couple at time $\tau + 1$. Otherwise, we say that Step 1 fails and one begins Step 3. Hence, Step 1 of trial j has two ways to fail: either ω belongs to $\Omega_{K, \alpha, \tau_{j-1}}^c$ and one moves directly to Step 3 or ω belongs to $\Omega_{K, \alpha, \tau_{j-1}}$, one tries to couple and it fails.

2.5.2 Lower-bound for the successful-coupling probability

The main purpose of this subsection is to get a positive lower-bound for the successful-coupling probability which will be independent of j (the number of the tentative), in other words we want to prove the following proposition

Proposition 2.5.1. *Assume (\mathbf{H}_1) and (\mathbf{H}_2) . Let $K > 0$, $\alpha > \frac{1}{2} \vee (\frac{3}{2} - \beta)$ if we are under $(\mathbf{H}_{\text{poly}})$ and $\alpha > 0$ different from ζ if we are under $(\mathbf{H}_{\text{exp}})$. In both cases, there exists δ_0 in $(0, 1)$ such that for all $j \geq 1$,*

$$\delta_0 \leq \mathbb{P}(\Delta\tau_j = +\infty | \Omega_{K, \alpha, \tau_{j-1}}) \quad (2.5.7)$$

where $\Delta\tau_j := \tau_j - \tau_{j-1}$ and τ_j is defined in Subsection 2.4.1 as the end of trial j .

Moreover, we can choose $\delta_1 \in (0, 1)$ such that

$$\forall j \geq 1, \quad \delta_1 \leq \mathbb{P}(\tau_j < \infty | \tau_{j-1} < \infty). \quad (2.5.8)$$

The second part of Proposition 2.5.1 may appear of weak interest but will be of first importance in Subsection 2.6.2.

Step 1 (hitting step)

Lemma 2.5.1. *Let $K > 0$ and $\alpha > 0$. Assume (\mathbf{H}_1) and (\mathbf{H}_2) . Let $\tilde{K} > 0$ be the constant appearing in (\mathbf{H}_2) , $\delta_1 \in (0, 1)$ and τ be a stopping time with respect to $(\mathcal{F}_n)_{n \in \mathbb{Z}}$ such that $\mathbb{P}(\Omega_{K, \alpha, \tau}) > 0$.*

We can build $(\xi_{\tau+1}^1, \xi_{\tau+1}^2)$ with $\xi_{\tau+1}^1 \sim \mathcal{N}(0, I_d)$ and $\xi_{\tau+1}^2 \sim \mathcal{N}(0, I_d)$ such that

(i) *There exist $K_1 \in (0, \tilde{K}]$ and $\delta_{K_1} \in (0, 1)$ such that*

$$\mathbb{P}(X_{\tau+1}^1 = X_{\tau+1}^2 | \Omega_{K, \alpha, \tau}) \geq \mathbb{P}(\xi_{\tau+1}^2 = \Lambda_{\mathbf{x}}(\xi_{\tau+1}^1), |\xi_{\tau+1}^1| \leq K_1 | \Omega_{K, \alpha, \tau}) \geq \delta_{K_1} > 0 \quad (2.5.9)$$

and

$$\mathbb{P}(\Omega_{K, \alpha, \tau}^c \cup (\{\xi_{\tau+1}^2 \neq \Lambda_{\mathbf{x}}(\xi_{\tau+1}^1) \text{ or } |\xi_{\tau+1}^1| > K_1\} \cap \Omega_{K, \alpha, \tau})) \geq \delta_1 \quad (2.5.10)$$

where $\mathbf{x} := (X_{\tau}^1, X_{\tau}^2, \sum_{k=0}^{+\infty} a_k \xi_{\tau+1-k}^1, \sum_{k=0}^{+\infty} a_k \xi_{\tau+1-k}^2)$ and $\Lambda_{\mathbf{x}}$ comes from (\mathbf{H}_2) .

(ii) *There exists $M_K > 0$ such that*

$$|g_{\tau}| = |\xi_{\tau+1}^1 - \xi_{\tau+1}^2| \leq M_K \quad a.s.$$

Remark 2.5.2. The constant δ_1 is chosen independently from K and α .

Before proving this result, let us explain a bit why we add the lower-bound (2.5.10). As we already said, we will see further (in Subsection 2.6.2) that we need the (uniform) bound on the failure-coupling probability given in (2.5.8). Therefore, for every $j \geq 1$, we will consider that Step 1 of trial j fails if and only if $\omega \in \Omega_{K, \alpha, \tau_{j-1}}^c \cup (\{\xi_{\tau_{j-1}+1}^2 \neq \Lambda_{\mathbf{x}}(\xi_{\tau_{j-1}+1}^1) \text{ or } |\xi_{\tau_{j-1}+1}^1| > K_1\} \cap \Omega_{K, \alpha, \tau_{j-1}})$ and in this case one immediately begins Step 3. Hence, for all $j \geq 1$, thanks to Lemma 2.5.1 we get the existence of K_1 such that:

$$\begin{aligned} & \mathbb{P}(\tau_j < \infty | \tau_{j-1} < \infty) \\ & \geq \mathbb{P}\left(\Omega_{K, \alpha, \tau_{j-1}}^c \cup \left(\{\xi_{\tau_{j-1}+1}^2 \neq \Lambda_{\mathbf{x}}(\xi_{\tau_{j-1}+1}^1) \text{ or } |\xi_{\tau_{j-1}+1}^1| > K_1\} \cap \Omega_{K, \alpha, \tau_{j-1}}\right)\right) \geq \delta_1 \end{aligned}$$

and then (2.5.8) derives from Lemma 2.5.1. This construction may seem artificial but it is necessary to prove Proposition 2.6.2. Moreover, this has no impact on the computation of the rate of convergence to equilibrium since it only affects Step 1. We can now move on the proof of Lemma 2.5.1.

Proof. (i) Set $\mathbf{x} := (X_{\tau}^1, X_{\tau}^2, \sum_{k=1}^{+\infty} a_k \xi_{\tau+1-k}^1, \sum_{k=1}^{+\infty} a_k \xi_{\tau+1-k}^2)$. Conditionnally to $\Omega_{K, \alpha, \tau}$ we have $\mathbf{x} \in B(0, K)^4$ and we can build (Z_1, Z_2) as in Lemma 2.4.1. Let

$\xi \sim \mathcal{N}(0, 1)$ be independent from (Z_1, Z_2) and set

$$(\xi_{\tau+1}^1, \xi_{\tau+1}^2) = (\mathbb{1}_{\Omega_{K,\alpha,\tau}} Z_1 + \mathbb{1}_{\Omega_{K,\alpha,\tau}^c} \xi, \mathbb{1}_{\Omega_{K,\alpha,\tau}} Z_2 + \mathbb{1}_{\Omega_{K,\alpha,\tau}^c} \xi). \quad (2.5.11)$$

Therefore, we deduce by Lemma 2.4.1 and its proof that for all $K_1 \in (0, \tilde{K}]$,

$$\mathbb{P}(X_{\tau+1}^1 = X_{\tau+1}^2 | \Omega_{K,\alpha,\tau}) \geq \underbrace{\mathbb{P}(Z_2 = \Lambda_{\mathbf{x}}(Z_1), |Z_1| \leq K_1 | \Omega_{K,\alpha,\tau})}_{=\mathbb{P}(\xi_{\tau+1}^2 = \Lambda_{\mathbf{x}}(\xi_{\tau+1}^1), |\xi_{\tau+1}^1| \leq K_1 | \Omega_{K,\alpha,\tau})} \geq \delta_{K_1} > 0. \quad (2.5.12)$$

And the first part of (i) is proven. It remains to choose the good $K_1 \in (0, \tilde{K}]$ to get the second part. Set $p_K := \mathbb{P}(\Omega_{K,\alpha,\tau})$ and $\mu := \mathcal{N}(0, I_d)$, then

$$\begin{aligned} & \mathbb{P}(\Omega_{K,\alpha,\tau}^c \cup (\{\xi_{\tau+1}^2 \neq \Lambda_{\mathbf{x}}(\xi_{\tau+1}^1) \text{ or } |\xi_{\tau+1}^1| > K_1\} \cap \Omega_{K,\alpha,\tau})) \\ &= 1 - p_K + p_K \mathbb{P}(\{\xi_{\tau+1}^2 \neq \Lambda_{\mathbf{x}}(\xi_{\tau+1}^1) \text{ or } |\xi_{\tau+1}^1| > K_1\} | \Omega_{K,\alpha,\tau}) \\ &\geq 1 - p_K + p_K \mathbb{P}(|\xi_{\tau+1}^1| > K_1 | \Omega_{K,\alpha,\tau}) \\ &\geq 1 - p_K + p_K \mu(B(0, K_1)^c) = 1 - p_K + p_K(1 - \mu(B(0, K_1))) \end{aligned}$$

where the last inequality is due to Lemma 2.4.1 one more time. Finally, it remains to choose $K_1 \in (0, \tilde{K}]$ small enough in order to get $1 - p_K + p_K(1 - \mu(B(0, K_1))) \geq \delta_1$.

(ii) If $\omega \in \Omega_{K,\alpha,\tau}$, by the previous construction and Lemma 2.4.1, we have $|g_\tau(\omega)| = |Z_1(\omega) - Z_2(\omega)| \leq M_K$. And if $\omega \in \Omega_{K,\alpha,\tau}^c$ then $|g_\tau(\omega)| = |\xi(\omega) - \xi(\omega)| = 0$ which concludes the proof of (ii). \square

To fix the ideas let us recall what we mean by “success of Step 1” and “failure of Step 1” of trial j ($j \geq 1$):

$$\{\text{success of Step 1}\} = \Omega_{K,\alpha,\tau_{j-1}} \cap \{\xi_{\tau_{j-1}+1}^2 = \Lambda_{\mathbf{x}}(\xi_{\tau_{j-1}+1}^1), |\xi_{\tau_{j-1}+1}^1| \leq K_1\} \quad (2.5.13)$$

$$\{\text{failure of Step 1}\} = \Omega_{K,\alpha,\tau_{j-1}}^c \cup \left(\{\xi_{\tau_{j-1}+1}^2 \neq \Lambda_{\mathbf{x}}(\xi_{\tau_{j-1}+1}^1) \text{ or } |\xi_{\tau_{j-1}+1}^1| > K_1\} \cap \Omega_{K,\alpha,\tau_{j-1}} \right) \quad (2.5.14)$$

where $\mathbf{x} := \left(X_{\tau_{j-1}}^1, X_{\tau_{j-1}}^2, \sum_{k=0}^{+\infty} a_k \xi_{\tau_{j-1}+1-k}^1, \sum_{k=0}^{+\infty} a_k \xi_{\tau_{j-1}+1-k}^2 \right)$.

Step 2 (sticking step)

Step 2 of trial j consists in trying to keep the paths fastened together on successive intervals $I_{j,\ell}$. More precisely, during trial j , we set

$$\begin{aligned} & I_{j,0} := \{\tau_{j-1} + 1\}, \quad I_{j,1} := \llbracket \tau_{j-1} + 2, \tau_{j-1} + 2c_2 - 1 \rrbracket \\ & \text{and } \forall \ell \geq 2, \quad I_{j,\ell} := \llbracket \tau_{j-1} + c_2 s_\ell, \tau_{j-1} + c_2 s_{\ell+1} - 1 \rrbracket \end{aligned} \quad (2.5.15)$$

where $c_2 \geq 2$ will be chosen further and with

$$\forall \ell \geq 2, \quad s_\ell = \begin{cases} 2^\ell & \text{under } (\mathbf{H}_{\text{poly}}) \\ \ell & \text{under } (\mathbf{H}_{\text{exp}}). \end{cases} \quad (2.5.16)$$

We denote

$$\ell_j^* := \sup\{\ell \geq 1 \mid \forall n \in I_{j,\ell-1}, g_{n-1} = g_{n-1}^{(s)}\} \quad (2.5.17)$$

where $g_{n-1}^{(s)}$ is the successful-coupling drift defined by (2.4.15), i.e. $g_{n-1}^{(s)} = -\sum_{l=1}^{+\infty} a_l g_{n-1-l}$. In other words, I_{j,ℓ_j^*} is the interval where the failure occurs. If $\{\ell \geq 1 \mid \forall n \in I_{j,\ell-1}, g_{n-1} = g_{n-1}^{(s)}\} = \emptyset$, we adopt the convention $\ell_j^* = 0$, it corresponds to the case where the failure occurs at Step 1. When $\ell_j^* = +\infty$, trial j is successful. For a given positive α and $K > 0$, we set

$$\mathcal{B}_{j,\ell} := \Omega_{K,\alpha,\tau_{j-1}} \cap \{\ell_j^* > \ell\} \quad \forall j \geq 1, \ell \geq 0, \quad (2.5.18)$$

which means that failure of Step 2 may occur at most after ℓ trials. With this notations we get

$$\mathbb{P}(\Delta\tau_j = +\infty \mid \Omega_{K,\alpha,\tau_{j-1}}) = \mathbb{P}(\text{success of Step 1} \mid \Omega_{K,\alpha,\tau_{j-1}}) \prod_{\ell=1}^{+\infty} \mathbb{P}(\mathcal{B}_{j,\ell} \mid \mathcal{B}_{j,\ell-1}) \quad (2.5.19)$$

where the event $\{\text{success of Step 1}\}$ is defined by (2.5.13).

Remark 2.5.3. There is an infinite product in this expression of the successful-coupling probability. Hence, the size choice of the intervals $I_{j,\ell}$ defined in (2.5.16) will play a significant role in the convergence of the product to a positive limit.

In the following lemma, similarly to the above definitions, we consider for a stopping time τ the intervals $(I_{\tau,\ell})_{\ell \geq 1}$, the integer ℓ_τ^* and the events $\mathcal{B}_{\tau,\ell}$, replacing τ_{j-1} by τ .

Lemma 2.5.2. *Let $K > 0$, assume (\mathbf{H}_1) and (\mathbf{H}_2) . Let $\alpha > \frac{1}{2} \vee (\frac{3}{2} - \beta)$ under $(\mathbf{H}_{\text{poly}})$ or $\alpha > 0$ different from ζ under $(\mathbf{H}_{\text{exp}})$. Let τ be a stopping time with respect to $(\mathcal{F}_n)_{n \in \mathbb{Z}}$ (defined in Subsection 2.4.1) and assume that the system is (K, α) -admissible at time τ , then there exists $C_K > 0$ such that for $c_2 \geq 2$ large enough the successful drift $g^{(s)}$ satisfies*

$$\text{for } \ell = 1, \quad \|g^{(s)}\|_{I_{\tau,1}} \leq C_K$$

and

$$\forall \ell \geq 2, \quad \|g^{(s)}\|_{I_{\tau,\ell}} = \left(\sum_{k=\tau+c_2s_\ell}^{\tau+c_2s_{\ell+1}-1} |g_{k-1}^{(s)}|^2 \right)^{1/2} \leq 2^{-\tilde{\alpha}\ell}$$

where $\tilde{\alpha} := \begin{cases} \min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2 - \varepsilon & \text{for all } \varepsilon > 0 \text{ under } (\mathbf{H}_{\text{poly}}) \\ \min(\alpha, \zeta) & \text{under } (\mathbf{H}_{\text{exp}}) \end{cases}$.

Therefore, for all $\ell \geq 1$, we can build thanks to Corollary 2.4.1 $(\xi_k^1)_{k \in I_{\tau,\ell}}, (\xi_k^2)_{k \in I_{\tau,\ell}}$ during Step 2 in such a way that

$$\mathbb{P}(\mathcal{B}_{\tau,1} \mid \mathcal{B}_{\tau,0}) \geq \delta_K^1 \quad \text{and} \quad \forall \ell \geq 2, \quad \mathbb{P}(\mathcal{B}_{\tau,\ell} \mid \mathcal{B}_{\tau,\ell-1}) \geq 1 - 2^{-\tilde{\alpha}\ell}$$

where $\delta_K^1 \in (0, 1)$.

Moreover, if $2 \leq \ell_\tau^* < +\infty$, there exists $C_\alpha > 0$ independent from K such that

$$\left(\sum_{k=\tau+c_2s_{\ell_\tau^*}}^{\tau+c_2s_{\ell_\tau^*+1}-1} |g_{k-1}|^2 \right)^{1/2} \leq C_\alpha(\ell_\tau^* + 1)$$

and if $\ell_\tau^* = 1$, $\|g\|_{I_{\tau,1}} \leq C'_K$ for some constant $C'_K > 0$.

Remark 2.5.4. \triangleright Under hypothesis $(\mathbf{H}_{\text{poly}})$ the condition $\alpha > \frac{1}{2} \vee (\frac{3}{2} - \beta)$ will ensure that

$$\min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2 > 0.$$

\triangleright In the polynomial case, for technical reasons $\tilde{\alpha}$ depends on $\varepsilon > 0$. This expression allows us to put together different cases and simplify the lemma. Indeed, if $(\alpha, \beta) \notin \{1\} \times (0, 1] \cup (0, 1] \times \{1\}$, we can take $\varepsilon = 0$.

To prove this lemma we will use in the *polynomial case* the following technical result which a more precise statement and a proof are given in Appendix 2.F.

Lemma 2.5.3 (Technical lemma). *Let $\alpha > 0$ and $\beta > 0$ such that $\alpha + \beta > 1$. Then, there exists $C(\alpha, \beta) > 0$ such that for every $\varepsilon > 0$,*

$$\forall n \geq 0, \quad \sum_{k=0}^n (k+1)^{-\beta} (n+1-k)^{-\alpha} \leq C(\alpha, \beta) (n+1)^{-\min\{\alpha, \beta, \alpha+\beta-1\}+\varepsilon}.$$

When $(\alpha, \beta) \notin \{1\} \times (0, 1] \cup (0, 1] \times \{1\}$, we can take $\varepsilon = 0$ in the previous inequality.

We can now move on the proof of Lemma 2.5.2.

Proof. Let us prove the first part of the lemma, namely the upper-bound of the ℓ^2 norm for the successful-coupling drift term on the intervals I_ℓ^τ . Indeed, the second part is just an application of Corollary 2.4.1. Since the system is (K, α) -admissible at time τ , we get by (2.5.1)

$$\forall n \geq 0, \quad \left| \sum_{k=n+1}^{+\infty} a_k g_{\tau+n-k} \right| \leq v_n.$$

But, if Step 2 is successful, we recall that by (2.4.15) the successful drift satisfies $g_{\tau+n}^{(s)} = -\sum_{k=1}^{+\infty} a_k g_{\tau+n-k}$ for all $n \geq 1$, hence

$$\forall n \geq 1, \quad \sum_{k=0}^n a_k g_{\tau+n-k}^{(s)} = - \underbrace{\sum_{k=n+1}^{+\infty} a_k g_{\tau+n-k}}_{=: u_n} \quad \text{and we set} \quad u_0 := g_\tau^{(s)}.$$

Therefore thanks to Remark 2.2.2, this is equivalent to $g_{\tau+n}^{(s)} = \sum_{k=0}^n b_k u_{n-k}$ for all $n \geq 0$. By the admissibility assumption, we have $\forall k \in \{0, \dots, n-1\}$, $|u_{n-k}| \leq v_{n-k}$ and by Lemma 2.5.1 (ii), $|u_0| \leq M_K$. Hence, we get

$$\forall n \geq 0, \quad |g_{\tau+n}^{(s)}| \leq M_K \sum_{k=0}^n |b_k| v_{n-k}. \quad (2.5.20)$$

• Polynomial case: Assume $(\mathbf{H}_{\text{poly}})$. Then for all $n \geq 0$, $v_n = (n+1)^{-\alpha}$ with $\alpha > \frac{1}{2} \vee \left(\frac{3}{2} - \beta\right)$.

Here, (2.5.20) is equivalent to

$$\begin{aligned} \forall n \geq 0, |g_{\tau+n}^{(s)}| &\leq M_K C_\beta \sum_{k=0}^n (k+1)^{-\beta} (n+1-k)^{-\alpha} \\ &\leq M'_K (n+1)^{-\min\{\alpha, \beta, \alpha+\beta-1\}+\varepsilon}, \end{aligned}$$

for all $\varepsilon > 0$ by applying the technical lemma 2.5.3 and setting $M'_K := C(\alpha, \beta) M_K C_\beta$. We then set $\tilde{\alpha} := \min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2 - \varepsilon$.

Hence, for all $\ell \geq 2$,

$$\begin{aligned} \|g^{(s)}\|_{I_{\tau, \ell}} &= \left(\sum_{k=\tau+c_2 2^\ell}^{\tau+c_2 2^{\ell+1}-1} |g_{k-1}^{(s)}|^2 \right)^{1/2} \\ &\leq M'_K \left(\sum_{k=c_2 2^\ell}^{c_2 2^{\ell+1}-1} k^{-2\tilde{\alpha}-1} \right)^{1/2} \\ &\leq M'_K \times \left(c_2 2^\ell \times (c_2 2^\ell)^{-2\tilde{\alpha}-1} \right)^{1/2} = M'_K c_2^{-\tilde{\alpha}} 2^{-\tilde{\alpha}\ell}. \end{aligned}$$

It remains to choose $c_2 \geq 2 \vee (M'_K)^{1/\tilde{\alpha}}$ to get the desired bound.

• Exponential case: Assume $(\mathbf{H}_{\text{exp}})$. Then for all $n \geq 0$, $v_n = e^{-\alpha n}$ with $\alpha > 0$ and $\alpha \neq \zeta$.

Here, (2.5.20) is equivalent to

$$\begin{aligned} \forall n \geq 0, |g_{\tau+n}^{(s)}| &\leq M_K C_\zeta \sum_{k=0}^n e^{-\zeta k} e^{-\alpha(n-k)} \\ &\leq M'_K e^{-\min\{\alpha, \zeta\}n}, \end{aligned}$$

where we have set $M'_K := M_K C_\zeta$. We then define $\tilde{\alpha} := \min(\alpha, \zeta)$.

Hence, for all $\ell \geq 2$,

$$\begin{aligned} \|g^{(s)}\|_{I_\ell^{(\tau)}} &= \left(\sum_{k=\tau+c_2 \ell}^{\tau+c_2(\ell+1)-1} |g_{k-1}^{(s)}|^2 \right)^{1/2} \\ &\leq M'_K \left(\sum_{k=c_2 \ell}^{c_2(\ell+1)-1} e^{-2\tilde{\alpha}(k-1)} \right)^{1/2} \\ &\leq M'_K \times \left(\frac{e^{-2\tilde{\alpha}(c_2 \ell - 2)}}{2\tilde{\alpha}} [1 - e^{-2\tilde{\alpha}c_2}] \right)^{1/2} \quad (\text{by integral upper-bound}) \\ &\leq e^{-\tilde{\alpha}\ell} \quad (\text{by choosing } c_2 \geq 2 \text{ large enough}) \\ &\leq 2^{-\tilde{\alpha}\ell}. \end{aligned}$$

For $\ell = 1$, in both polynomial and exponential cases, the same approach gives us the existence of $C_K > 0$ such that

$$\|g^{(s)}\|_{I_{\tau,1}} \leq C_K.$$

□

By combining Lemma 2.5.1, Lemma 2.5.2 and the expression (2.5.19) we finally get Proposition 2.5.1.

2.6 (K, α) -admissibility

2.6.1 On condition (2.5.1)

Let $\Delta t_3^{(j)}$ be the duration of Step 3 of trial j for $j \geq 1$. The purpose of the next proposition is to prove that thanks to a calibration of $\Delta t_3^{(j)}$, one satisfies almost surely condition (2.5.1) at time τ_j .

Proposition 2.6.1. *Assume (\mathbf{H}_1) and (\mathbf{H}_2) . Let $\alpha \in (\frac{1}{2} \vee (\frac{3}{2} - \beta), \rho)$ if we are under $(\mathbf{H}_{\text{poly}})$ and $\alpha \in (0, \lambda)$ different from ζ if we are under $(\mathbf{H}_{\text{exp}})$. Assume that for all $j \geq 1$,*

$$\Delta t_3^{(j)} = \begin{cases} t_* \varsigma^j 2^{\theta \ell_j^*} & \text{with } \theta > (2(\rho - \alpha))^{-1} \quad \text{under } (\mathbf{H}_{\text{poly}}) \\ t_* + \varsigma^j + \theta \ell_j^* & \text{with } \theta > 0 \quad \text{under } (\mathbf{H}_{\text{exp}}) \end{cases}$$

where ℓ_j^* is defined in (2.5.17) and with $\varsigma > 1$ arbitrary. Then for every $K > 0$, there exists a choice of t_* such that, for all $j \geq 0$, condition (2.5.1) is a.s. true at time τ_j on the event $\{\tau_j < +\infty\}$. In other words, for all $j \geq 0$,

$$\mathbb{P}(\Omega_{\alpha, \tau_j}^1 | \{\tau_j < +\infty\}) = 1.$$

Remark 2.6.1. ▷ With Proposition 2.6.1 in hand, we are now in position to discuss the statement of Theorem 2.2.1 (iii) as mentioned in Remark 2.2.6. Adapting the proof of (2.7.3) (by taking an exponential Markov inequality), it appears that a fundamental tool to get exponential rate of convergence to equilibrium would be: for all $j \geq 1$, there exist $\lambda, C > 0$ such that

$$\mathbb{E}[e^{\lambda \Delta \tau_j} | \mathcal{F}_{\tau_{j-1}}] \leq C \tag{2.6.1}$$

where $\Delta \tau_j := \tau_j - \tau_{j-1}$. But, under $(\mathbf{H}_{\text{exp}})$, Proposition 2.6.1 shows that $\Delta \tau_j \geq \varsigma^j$ with $\varsigma > 1$. This dependency on j conflicts with the necessary control of the conditional expectation previously cited.

▷ Now, let us focus on the particular case of a finite memory: there exists $m \geq 1$ such that $a_k = 0$ for all $k > m$. As we saw in the second part of Subsection 2.2.6, $(|b_k|)$ decays exponentially fast. Then, an adjustment¹ of the proof of Proposition 2.6.1 would lead to (2.6.1) and perhaps in such case, our robust but general approach could be simplified.

¹In this case, we can fix the duration of Step 3 equal to m since the memory only involves the m previous times.

Proof. Let us begin by the first coupling trial, in other words for $j = 0$. We recall that $g_k = 0$ for all $k < \tau_0$ (see (2.4.6)), therefore

$$\forall n \geq 0, \left| \sum_{k=n+1}^{+\infty} a_k g_{\tau_0+n-k} \right| = 0 \leq v_n$$

and then condition (2.5.1) is a.s. true at time τ_0 . Now, we assume $j \geq 1$ and we work on the event

$$\{\tau_j < +\infty\} \quad (\supset \{\tau_m < +\infty\} \text{ for all } 0 \leq m \leq j-1).$$

Let us prove that on this event we have for all $n \geq 0$, $\left| \sum_{k=n+1}^{+\infty} a_k g_{\tau_j+n-k} \right| \leq v_n$. Set $u_n := \sum_{k=n+1}^{+\infty} a_k g_{\tau_j+n-k}$. Since $g_k = 0$ for all $k < \tau_0$, we get $u_n = \sum_{k=n+1}^{n+\tau_j-\tau_0} a_k g_{\tau_j+n-k}$. Let us now separate the right term into the contributions of the different coupling trials. We get

$$u_n = \sum_{m=1}^j \left(\sum_{k=n+\tau_j-\tau_{m-1}+1}^{n+\tau_j-\tau_{m-1}} a_k g_{\tau_j+n-k} \right) = \sum_{m=1}^j \underbrace{\left(\sum_{k=\tau_{m-1}}^{\tau_m-1} a_{n+\tau_j-k} g_k \right)}_{(\star)_m}$$

and $(\star)_m$ corresponds to the contribution of trial m , divided into two parts: success and failure. We have now to distinguish two cases:

First case: $\ell_m^* \geq 1$, in other words the failure occurs during Step 2. We recall that in this case the system was automatically (K, α) -admissible at time τ_{m-1} , which will allow us to use Lemma 2.5.2 on τ_{m-1} .

Then, since $g_k = 0$ on $[\tau_{m-1} + c_2 s_{\ell_m^*+1}, \tau_m - 1]$ by definition of Step 3 of the coupling procedure,

$$\begin{aligned} (\star)_m &= \sum_{k=\tau_{j m-1}}^{\tau_{m-1} + c_2 s_{\ell_m^*+1} - 1} a_{n+\tau_j-k} g_k \\ &= \underbrace{\sum_{k=0}^{c_2 s_{\ell_m^*} - 1} a_{n+\tau_j-\tau_{m-1}-k} g_{\tau_{m-1}+k}}_{\text{success}} + \underbrace{\sum_{k=c_2 s_{\ell_m^*}}^{c_2 s_{\ell_m^*+1} - 1} a_{n+\tau_j-\tau_{m-1}-k} g_{\tau_{m-1}+k}}_{\text{failure}}. \end{aligned}$$

We have now to make the distinction between the polynomial and the exponential case.

▷ Under $(\mathbf{H}_{\text{poly}})$: $s_\ell = 2^\ell$, $|a_k| \leq C_\rho (k+1)^{-\rho}$ and then $a_k^2 \leq C_\rho^2 (k+1)^{-2\rho}$.

Using Cauchy-Schwarz inequality, the domination assumption on (a_k) , and the fact

that

$$n + \tau_j - \tau_{m-1} - k + 1 \geq n + \tau_m - \tau_{m-1} - k + 1 = \Delta t_3^{(m)} + c_2 2^{\ell_m^* + 1} + n - k$$

we get,

$$\begin{aligned} |\text{success}| &\leq \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} a_{n+\tau_j-\tau_{m-1}-k}^2 \right)^{1/2} \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} |g_{\tau_{m-1}+k}|^2 \right)^{1/2} \\ &\leq C_\rho \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} (n + \tau_j - \tau_{m-1} - k + 1)^{-2\rho} \right)^{1/2} \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} |g_{\tau_{m-1}+k}|^2 \right)^{1/2} \\ &\leq C_\rho \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} (\Delta t_3^{(m)} + c_2 2^{\ell_m^* + 1} + n - k)^{-2\rho} \right)^{1/2} \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} |g_{\tau_{m-1}+k}|^2 \right)^{1/2} \\ &= C_\rho \left(\sum_{k=\Delta t_3^{(m)} + c_2 2^{\ell_m^* + n + 1}}^{\Delta t_3^{(m)} + c_2 2^{\ell_m^* + 1} + n} k^{-2\rho} \right)^{1/2} \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} |g_{\tau_{m-1}+k}|^2 \right)^{1/2} \\ &\leq C_\rho \left(c_2 2^{\ell_m^*} (n + \Delta t_3^{(m)})^{-2\rho} \right)^{1/2} \left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} |g_{\tau_{m-1}+k}|^2 \right)^{1/2}. \end{aligned}$$

By the same arguments, we obtain

$$|\text{failure}| \leq C_\rho \left(c_2 2^{\ell_m^*} (n + \Delta t_3^{(m)})^{-2\rho} \right)^{1/2} \left(\sum_{k=c_2 2^{\ell_m^*}}^{c_2 2^{\ell_m^* + 1} - 1} |g_{\tau_{m-1}+k}|^2 \right)^{1/2}.$$

Hence, by the triangular inequality, we have

$$|(\star)_m| \leq C_\rho \sqrt{c_2} 2^{\ell_m^*/2} (n + \Delta t_3^{(m)})^{-\rho} \left[\underbrace{\left(\sum_{k=0}^{c_2 2^{\ell_m^* - 1}} |g_{\tau_{m-1}+k}|^2 \right)^{1/2}}_{(1)} + \underbrace{\left(\sum_{k=c_2 2^{\ell_m^*}}^{c_2 2^{\ell_m^* + 1} - 1} |g_{\tau_{m-1}+k}|^2 \right)^{1/2}}_{(2)} \right].$$

Since $\ell_j^* \geq 1$, by Lemma 2.5.1 (ii) and Lemma 2.5.2, we have

$$(1) \leq M_K + \sum_{\ell=1}^{\ell_m^* - 1} \|g^{(s)}\|_{I_\ell} \leq M_K + C_K + \sum_{\ell=2}^{+\infty} 2^{-\tilde{\alpha}\ell} =: \tilde{C}_K$$

and

$$(2) \leq \max(C_\alpha, C'_K) (\ell_m^* + 1).$$

Therefore

$$|(\star)_m| \leq C_K^{(2)} 2^{\ell_m^*/2} (\ell_m^* + 1) \left(n + \Delta t_3^{(m)} \right)^{-\rho} \quad (2.6.2)$$

where $C_K^{(2)} := \max(C_\alpha, C'_K, \tilde{C}_K)$.

Moreover, recall that under $(\mathbf{H}_{\text{poly}})$

$$\Delta t_3^{(m)} = t_* \varsigma^m 2^{\theta \ell_m^*} \quad \text{with } \theta > (2(\rho - \alpha))^{-1} \text{ and } \varsigma > 1. \quad (2.6.3)$$

Plugging the definition of $\Delta t_3^{(m)}$ into (2.6.2) and using that for all $x, y > 0$, $(x + y)^{-\rho} \leq x^{-(\rho-\alpha)} y^{-\alpha}$

$$\begin{aligned} |(\star)_m| &\leq C_K^{(2)} 2^{\ell_m^*/2} (\ell_m^* + 1) (n + \Delta t_3^{(m)})^{-\rho} \\ &\leq C_K^{(2)} (\ell_m^* + 1) 2^{(1/2-\theta(\rho-\alpha))\ell_m^*} (t_* \varsigma^m)^{-(\rho-\alpha)} (n + 1)^{-\alpha}. \end{aligned} \quad (2.6.4)$$

Since $\theta > (2(\rho - \alpha))^{-1}$ we have

$$C_{\alpha, K} = \sup_{\ell^* > 0} C_K^{(2)} (\ell^* + 1) 2^{(1/2-\theta(\rho-\alpha))\ell^*} < +\infty,$$

and (2.6.4) yields

$$|(\star)_m| \leq C_{\alpha, K} (t_* \varsigma^m)^{-(\rho-\alpha)} (n + 1)^{-\alpha} \quad \text{under } (\mathbf{H}_{\text{poly}}). \quad (2.6.5)$$

▷ Under $(\mathbf{H}_{\text{exp}})$: $s_\ell = \ell$, $|a_k| \leq C_\lambda e^{-\lambda k}$ and then $a_k^2 \leq C_\lambda^2 e^{-2\lambda k}$.

Since the proof is almost the same in the exponential case, we will go faster and skip some details.

Using again Cauchy-Schwarz inequality, the domination assumption on (a_k) , and the fact that

$$n + \tau_j - \tau_{m-1} - k \geq n + \tau_m - \tau_{m-1} - k = \Delta t_3^{(m)} + c_2 (\ell_m^* + 1) + n - k - 1$$

we get

$$|\text{success}| \leq \frac{C_\lambda e^\lambda}{\sqrt{2\lambda}} e^{-\lambda(n + \Delta t_3^{(m)})} \left(\sum_{k=0}^{c_2 2^{\ell_m^*} - 1} |g_{\tau_{m-1} + k}|^2 \right)^{1/2}$$

and by the same arguments,

$$|\text{failure}| \leq \frac{C_\lambda e^\lambda}{\sqrt{2\lambda}} e^{-\lambda(n + \Delta t_3^{(m)})} \left(\sum_{k=c_2 \ell_m^*}^{c_2 (\ell_m^* + 1) - 1} |g_{\tau_{m-1} + k}|^2 \right)^{1/2}.$$

As in the polynomial case, by using Lemma 2.5.1 and 2.5.2 we get the existence of $C_K^{(3)} > 0$ such that

$$|(\star)_m| \leq C_K^{(3)}(\ell_m^* + 1)e^{-\lambda(n + \Delta t_3^{(m)})}. \quad (2.6.6)$$

Moreover, recall that under $(\mathbf{H}_{\text{exp}})$

$$\Delta t_3^{(m)} = t_* + \varsigma^m + \theta \ell_m^* \quad \text{with } \theta > 0 \text{ and } \varsigma > 1. \quad (2.6.7)$$

Plugging the definition of $\Delta t_3^{(m)}$ into (2.6.6), we get

$$\begin{aligned} |(\star)_m| &\leq C_K^{(3)}(\ell_m^* + 1)e^{-\lambda(n + \Delta t_3^{(m)})} \\ &\leq C_K^{(3)}(\ell_m^* + 1)e^{-(\lambda - \alpha)\Delta t_3^{(m)}} e^{-\alpha n} \\ &= C_K^{(3)}(\ell_m^* + 1)e^{-(\lambda - \alpha)(t_* + \varsigma^m + \theta \ell_m^*)} e^{-\alpha n}. \end{aligned}$$

We set

$$C'_{\alpha, K} = \sup_{\ell^* > 0} C_K^{(3)}(\ell^* + 1)e^{-\theta(\lambda - \alpha)\ell^*} < +\infty.$$

And this gives us

$$|(\star)_m| \leq C'_{\alpha, K} e^{-(\lambda - \alpha)(t_* + \varsigma^m)} e^{-\alpha n} \quad \text{under } (\mathbf{H}_{\text{exp}}). \quad (2.6.8)$$

Second case: $\ell_m^* = 0$, in other words, failure occurs during Step 1. This includes the case when the system is not (K, α) -admissible at time τ_{m-1} .

We have

$$(\star)_m = a_{n + \tau_j - \tau_{m-1}} g_{\tau_{m-1}}.$$

By Lemma 2.5.1 (ii), $|g_{\tau_{m-1}}| \leq M_K$.

Moreover, since $n + \tau_j - \tau_{m-1} \geq n + \tau_m - \tau_{m-1} = n + \Delta t_3^{(m)}$, we obtain by using the same method as in the first case,

$$|(\star)_m| \leq \begin{cases} M_K (t_* \varsigma^m)^{-(\rho - \alpha)} (n + 1)^{-\alpha} & \text{under } (\mathbf{H}_{\text{poly}}) \\ M_K e^{-(\lambda - \alpha)(t_* + \varsigma^m)} e^{-\alpha n} & \text{under } (\mathbf{H}_{\text{exp}}) \end{cases}. \quad (2.6.9)$$

By putting (2.6.5), (2.6.8) and (2.6.9) together, we finally get

$$|(\star)_m| \leq \begin{cases} \max(M_K, C_{\alpha, K})(t_* \varsigma^m)^{-(\rho - \alpha)} (n + 1)^{-\alpha} & \text{under } (\mathbf{H}_{\text{poly}}) \\ \max(M_K, C'_{\alpha, K}) e^{-(\lambda - \alpha)(t_* + \varsigma^m)} e^{-\alpha n} & \text{under } (\mathbf{H}_{\text{exp}}) \end{cases}.$$

Set $S_1 = \sum_{m=1}^{+\infty} \varsigma^{-(\rho - \alpha)m}$ and $S_2 = \sum_{m=1}^{+\infty} e^{-(\lambda - \alpha)\varsigma^m}$. By choosing t_* large enough,

we obtain for all $1 \leq m \leq j$:

$$|(\star)_m| \leq \begin{cases} \frac{1}{S_1} \zeta^{-(\rho-\alpha)m} (n+1)^{-\alpha} & \text{under } (\mathbf{H}_{\text{poly}}) \\ \frac{1}{S_2} e^{-(\lambda-\alpha)\zeta^m} e^{-\alpha n} & \text{under } (\mathbf{H}_{\text{exp}}) \end{cases}. \quad (2.6.10)$$

Finally, by adding (2.6.10) for $m = 1, \dots, j$ we have

$$\forall n \geq 0, \quad |u_n| \leq \begin{cases} (n+1)^{-\alpha} & \text{under } (\mathbf{H}_{\text{poly}}) \\ e^{-\alpha n} & \text{under } (\mathbf{H}_{\text{exp}}) \end{cases}$$

which concludes the proof. □

2.6.2 Compact return condition (2.5.2)

In the sequel, we set

$$\mathcal{E}_j := \{\tau_j < \infty\} (= \{\tau_1 < \infty, \dots, \tau_j < \infty\}). \quad (2.6.11)$$

The aim of this subsection is to prove the following proposition:

Proposition 2.6.2. *Assume (\mathbf{H}_1) and (\mathbf{H}_2) . For all $\varepsilon > 0$, there exists $K_\varepsilon > 0$ such that*

$$\mathbb{P}(\Omega_{K_\varepsilon, \tau_j}^2 | \mathcal{E}_j) \geq 1 - \varepsilon. \quad (2.6.12)$$

At this stage, we assume that $(\mathbf{H}_{\text{poly}})$ is true. Indeed, the exponential case will immediately follow from the polynomial one since $(\mathbf{H}_{\text{exp}})$ implies $(\mathbf{H}_{\text{poly}})$.

Since for every events A_1, A_2, A_3 and A_4 , we have $\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) \geq \sum_{i=1}^4 \mathbb{P}(A_i) - 3$, it is enough to prove that for all $\varepsilon > 0$, there exists $K_\varepsilon > 0$ such that

$$\mathbb{P}(|X_{\tau_j}^i| \leq K_\varepsilon | \mathcal{E}_j) \geq 1 - \varepsilon \quad \text{and} \quad \mathbb{P}\left(\left|\sum_{k=1}^{+\infty} a_k \xi_{\tau_j+1-k}^i\right| \leq K_\varepsilon \mid \mathcal{E}_j\right) \geq 1 - \varepsilon \quad \text{for } i = 1, 2 \quad (2.6.13)$$

to get (2.6.12). Let us first focus on the first part of (2.6.13) concerning $|X_{\tau_j}^i|$ for $i = 1, 2$. Recall that the function $V : \mathbb{R}^d \rightarrow \mathbb{R}_+^*$ appearing in (\mathbf{H}_1) is such that $\lim_{|x| \rightarrow +\infty} V(x) = +\infty$.

For $K > 0$ large enough we then have: $|x| \geq K \Rightarrow V(x) \geq K$. Therefore, for $i = 1, 2$ and K_ε large enough, using Markov inequality we get

$$\mathbb{P}(|X_{\tau_j}^i| \geq K_\varepsilon | \mathcal{E}_j) \leq \mathbb{P}(V(X_{\tau_j}^i) \geq K_\varepsilon | \mathcal{E}_j) \leq \frac{\mathbb{E}(V(X_{\tau_j}^i) | \mathcal{E}_j)}{K_\varepsilon}. \quad (2.6.14)$$

Hence, the first part of (2.6.13) is true if there exists a constant C such that for every $j \in \mathbb{N}$ and for every $K > 0$,

$$\mathbb{E}(V(X_{\tau_j}^i) | \mathcal{E}_j) \leq C \quad \text{for } i = 1, 2. \quad (2.6.15)$$

Indeed, plugging (2.6.15) into (2.6.14) and taking $K_\varepsilon \geq \frac{C}{\varepsilon}$ yield the desired inequality. We see here that the independence of C with respect to K is essential. For the sake of simplicity, we will first use the following hypothesis to prove (2.6.15):

(\mathbf{H}'_1): Let $\gamma \in (0, 1)$. There exists $C_\gamma > 0$ such that for all $j \in \mathbb{N}$, for every $K > 0$ and for $i = 1, 2$,

$$\mathbb{E} \left[\sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j-l} |\Delta_{\tau_{j-1}+l}^i| \mid \mathcal{E}_j \right] < C_\gamma$$

where $\Delta\tau_j := \tau_j - \tau_{j-1}$ and Δ^i is the stationary Gaussian sequence defined by (2.2.2).

Proposition 2.6.3. *Assume (\mathbf{H}_1) , (\mathbf{H}_2) and (\mathbf{H}'_1) . Let $(X_n^1, X_n^2)_{n \in \mathbb{N}}$ be a solution of (2.4.2) with initial condition (X_0^1, X_0^2) satisfying $\mathbb{E}(V(X_0^i)) < \infty$ for $i = 1, 2$. Moreover, assume that $\tau_0 = 0$ and that $(\tau_j)_{j \geq 1}$ is built in such a way that for all $j \geq 1$, $\mathbb{P}(\mathcal{E}_j | \mathcal{E}_{j-1}) \geq \delta_1 > 0$ (where δ_1 is not depending on j) and $\Delta\tau_j \geq \frac{\log(\delta_1/2)}{\log(\gamma)}$. Then, there exists a constant C such that for all $j \in \mathbb{N}$ and for every $K > 0$,*

$$\mathbb{E}(V(X_{\tau_j}^i) \mid \mathcal{E}_j) \leq C \quad \text{for } i = 1, 2.$$

Remark 2.6.2. \triangleright Actually, hypothesis (\mathbf{H}'_1) is true under $(\mathbf{H}_{\text{poly}})$ and will be proven Appendix 2.G.

\triangleright The existence of $\delta_1 > 0$ independent from j is proven in Subsection 2.5.2.

\triangleright To get $\Delta\tau_j \geq \frac{\log(\delta_1/2)}{\log(\gamma)}$, it is sufficient to choose t_* large enough in the expression of $\Delta t_3^{(j)}$ (see Proposition 2.6.1).

\triangleright Since in Theorem 2.2.1 we made the assumption $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_0(dx) < +\infty$ and since an invariant distribution (extracted thanks to Theorem 2.3.1) also satisfies $\int_{\mathcal{X}} V(x) \Pi_{\mathcal{X}}^* \mu_*(dx) < +\infty$, we get that $\mathbb{E}(V(X_0^i)) < \infty$ for $i = 1, 2$. Hence, we can set $\tau_0 = 0$.

Proof. By (\mathbf{H}_1) , there exist $\gamma \in (0, 1)$ and $C > 0$ such that for all $n \geq 0$ and for $i = 1, 2$ we have

$$V(X_{n+1}^i) \leq \gamma V(X_n^i) + C(1 + |\Delta_{n+1}^i|).$$

By applying this inequality at time $n = \tau_j - 1$, and by induction, we immediately get:

$$V(X_{\tau_j}^i) \leq \gamma^{\Delta\tau_j} V(X_{\tau_{j-1}}^i) + C \sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j-l} (1 + |\Delta_{\tau_{j-1}+l}^i|). \quad (2.6.16)$$

By assumption $\Delta\tau_j \geq \frac{\log(\delta_1/2)}{\log(\gamma)}$ then $\gamma^{\Delta\tau_j} \leq \frac{\delta_1}{2}$. Moreover, since $\mathcal{E}_j \subset \mathcal{E}_{j-1}$ and

$\mathbb{P}(\mathcal{E}_j | \mathcal{E}_{j-1}) \geq \delta_1$, we get

$$\begin{aligned} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_j] &= \frac{1}{\mathbb{P}(\mathcal{E}_j)} \mathbb{E}[V(X_{\tau_{j-1}}^i) \mathbb{1}_{\mathcal{E}_j}] \leq \underbrace{\frac{\mathbb{P}(\mathcal{E}_{j-1})}{\mathbb{P}(\mathcal{E}_j)}}_{=\mathbb{P}(\mathcal{E}_j | \mathcal{E}_{j-1})^{-1}} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_{j-1}] \\ &\leq \delta_1^{-1} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_{j-1}]. \end{aligned}$$

Therefore

$$\mathbb{E}[\gamma^{\Delta\tau_j} V(X_{\tau_{j-1}}^i) | \mathcal{E}_j] \leq \frac{\delta_1}{2} \delta_1^{-1} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_{j-1}] = \frac{1}{2} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_{j-1}].$$

Hence, by taking (2.6.16), we have

$$\begin{aligned} \mathbb{E}[V(X_{\tau_j}^i) | \mathcal{E}_j] &\leq \frac{1}{2} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_{j-1}] + C \sum_{l=0}^{+\infty} \gamma^l + C \mathbb{E} \left[\sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j - l} |\Delta_{\tau_{j-1}+l}^i| \middle| \mathcal{E}_j \right] \\ &= \frac{1}{2} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_{j-1}] + \frac{C}{1-\gamma} + C \mathbb{E} \left[\sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j - l} |\Delta_{\tau_{j-1}+l}^i| \middle| \mathcal{E}_j \right]. \end{aligned}$$

Hypothesis (\mathbf{H}'_1) allows us to say

$$\mathbb{E}[V(X_{\tau_j}^i) | \mathcal{E}_j] \leq \frac{1}{2} \mathbb{E}[V(X_{\tau_{j-1}}^i) | \mathcal{E}_{j-1}] + C \left(\frac{1}{1-\gamma} + C_\gamma \right).$$

By induction, we get the existence of a constant $\tilde{C}_\gamma > 0$ such that

$$\mathbb{E}[V(X_{\tau_j}^i) | \mathcal{E}_j] \leq \left(\frac{1}{2} \right)^j \mathbb{E}[V(X_{\tau_0}^i) | \mathcal{E}_0] + \tilde{C}_\gamma.$$

Since $\mathbb{P}(\mathcal{E}_0) = 1$, it comes

$$\mathbb{E}[V(X_{\tau_j}^i) | \mathcal{E}_j] \leq \left(\frac{1}{2} \right)^j \mathbb{E}[V(X_{\tau_0}^i)] + \tilde{C}_\gamma.$$

Since $\tau_0 = 0$ and we assumed that $\mathbb{E}[V(X_0^i)] < \infty$, the proof is over. \square

As a result of the proof of (\mathbf{H}'_1) (see Appendix 2.G), we get the first part of (2.6.13). The second part can be also deduced from the proof of (\mathbf{H}'_1) thanks to Remark 2.G.1.

2.7 Proof of Theorem 2.2.1

Now we have all the necessary elements to prove the second part of the main theorem 2.2.1 concerning the convergence in total variation to the unique invariant distribution (where the uniqueness will immediately follow from this convergence).

We recall that $\Delta\tau_j$ denotes the duration of coupling trial j and we set

$$j^{(s)} := \inf\{j > 0, \Delta\tau_j = +\infty\}. \quad (2.7.1)$$

$j^{(s)}$ corresponds to the trial where the coupling procedure is successful. The aim of this section is to bound $\mathbb{P}(\tau_\infty > n)$ where $\tau_\infty = \inf\{n \geq 0 \mid X_k^1 = X_k^2, \forall k \geq n\}$, since $\|\mathcal{L}((X_k^1)_{k \geq n}) - \mathcal{S}\mu_\star\|_{TV} \leq \mathbb{P}(\tau_\infty > n)$. But, we have

$$\mathbb{P}(\tau_\infty > n) = \mathbb{P}\left(\sum_{k=1}^{+\infty} \Delta\tau_k \mathbf{1}_{\{j^{(s)} > k\}} > n\right)$$

where $j^{(s)}$ is defined in (2.7.1). It remains to bound the right term. Let $p \in (0, +\infty)$. If $p \in (0, 1)$, then by the Markov inequality and the simple inequality $|a+b|^p \leq |a|^p + |b|^p$, we get

$$\begin{aligned} \mathbb{P}\left(\sum_{k=1}^{+\infty} \Delta\tau_k \mathbf{1}_{\{j^{(s)} > k\}} > n\right) &\leq \frac{1}{n^p} \sum_{k=1}^{+\infty} \mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{j^{(s)} > k\}}] \\ &\leq \frac{1}{n^p} \sum_{k=1}^{+\infty} \mathbb{E}[\mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}]] \mathbf{1}_{\{\tau_{k-1} < +\infty\}}. \end{aligned} \quad (2.7.2)$$

Else, if $p \geq 1$, by Markov inequality and Minkowski inequality, we have

$$\begin{aligned} \mathbb{P}\left(\sum_{k=1}^{+\infty} \Delta\tau_k \mathbf{1}_{\{j^{(s)} > k\}} > n\right) &\leq \frac{1}{n^p} \mathbb{E}\left[\left(\sum_{k=1}^{+\infty} \Delta\tau_k \mathbf{1}_{\{j^{(s)} > k\}}\right)^p\right] \\ &\leq \frac{1}{n^p} \left(\sum_{k=1}^{+\infty} \mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{j^{(s)} > k\}}]^{1/p}\right)^p \\ &\leq \frac{1}{n^p} \left(\sum_{k=1}^{+\infty} \mathbb{E}[\mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}]] \mathbf{1}_{\{\tau_{k-1} < +\infty\}}\right)^p. \end{aligned} \quad (2.7.3)$$

We define the event $\mathcal{A}_{k,\ell} := \mathcal{B}_{k,\ell}^c \cap \mathcal{B}_{k,\ell-1}$ which corresponds to the failure of Step 2 after ℓ attempts at trial k . Both in (2.7.2) and (2.7.3), we separate the term $\mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}]$ through the events $\mathcal{A}_{k,\ell}$ which gives

$$\mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}] = \sum_{\ell=1}^{+\infty} \mathbb{E}[\mathbf{1}_{\mathcal{A}_{k,\ell}} |\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}]. \quad (2.7.4)$$

Moreover, thanks to Lemma 2.5.2 and the definition of the events $\mathcal{A}_{k,\ell}$, we deduce that for $\ell \geq 2$,

$$\mathbb{P}(\mathcal{A}_{k,\ell} \mid \{\tau_{k-1} < +\infty\}) \leq 2^{-\tilde{\alpha}\ell} \quad (2.7.5)$$

where $\tilde{\alpha} := \begin{cases} \min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2 - \varepsilon & \text{for all } \varepsilon > 0 \text{ under } (\mathbf{H}_{\text{poly}}) \\ \min(\alpha, \zeta) & \text{under } (\mathbf{H}_{\text{exp}}). \end{cases}$

We have now to distinguish the polynomial case from the exponential one.

▷ Under $(\mathbf{H}_{\text{poly}})$:

We have a bound of type $\Delta\tau_k \leq C_1\varsigma^k 2^{\theta\ell}$ (due to the value of $\Delta t_3^{(k)}$, see Proposition 2.6.1) on the event $\mathcal{A}_{k,\ell}$ where $\varsigma > 1$ is arbitrary. Indeed, on $\mathcal{A}_{k,\ell}$, we have

$$\begin{aligned} \Delta\tau_k = \tau_k - \tau_{k-1} &\leq c_2 2^{\ell+1} + \Delta t_3^{(k)} \\ &= c_2 2^{\ell+1} + t_* \varsigma^k 2^{\theta\ell} \\ &\leq C_1 \varsigma^k 2^{(\theta\vee 1)\ell} \quad (\text{for } C_1 \text{ large enough}). \end{aligned}$$

Hence, in (2.7.4) we get

$$\begin{aligned} \mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}] &\leq C_1^p \varsigma^{kp} \left(\sum_{\ell=1}^{+\infty} 2^{(\theta\vee 1)p\ell} \mathbb{P}(\mathcal{A}_{k,\ell} \mid \{\tau_{k-1} < +\infty\}) \right) \\ &\leq C_1^p \varsigma^{kp} \left(2^{(\theta\vee 1)p} + \sum_{\ell=2}^{+\infty} 2^{((\theta\vee 1)p - \tilde{\alpha})\ell} \right) \quad \text{using (2.7.5)} \\ &\leq C \varsigma^{kp} \quad \text{if } p \in \left(0, \frac{\tilde{\alpha}}{\theta \vee 1} \right). \end{aligned}$$

Then for $p \in (0, \frac{\tilde{\alpha}}{\theta \vee 1})$,

$$\mathbb{E}[\mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}] \mathbf{1}_{\{\tau_{k-1} < +\infty\}}] \leq C \varsigma^{kp} \mathbb{P}(j^{(s)} > k-1) \quad (2.7.6)$$

and it remains to control $\mathbb{P}(j^{(s)} > k-1)$. We have

$$\mathbb{P}(j^{(s)} > k-1) = \prod_{j=1}^{k-1} \mathbb{P}(\mathcal{E}_j \mid \mathcal{E}_{j-1}) = \prod_{j=1}^{k-1} (1 - \mathbb{P}(\mathcal{E}_j^c \mid \mathcal{E}_{j-1}))$$

where \mathcal{E}_j is defined in (2.6.11). By Proposition 2.5.1 and 2.6.2 applied for $\varepsilon = 1/2$, we get for every $j \geq 2$,

$$\mathbb{P}(\mathcal{E}_j^c \mid \mathcal{E}_{j-1}) \geq \mathbb{P}(\Delta\tau_j = +\infty \mid \Omega_{K_{1/2}, \alpha, \tau_{j-1}}) \mathbb{P}(\Omega_{K_{1/2}, \alpha, \tau_{j-1}} \mid \mathcal{E}_{j-1}) \geq \frac{\delta_0}{2}$$

where $\delta_0 > 0$ depends on $K_{1/2}$. Therefore, $\mathbb{P}(j^{(s)} > k-1) \leq \left(1 - \frac{\delta_0}{2}\right)^{k-1}$ and by (2.7.6)

$$\mathbb{E}[\mathbb{E}[|\Delta\tau_k|^p \mathbf{1}_{\{\Delta\tau_k < +\infty\}} \mid \{\tau_{k-1} < +\infty\}] \mathbf{1}_{\{\tau_{k-1} < +\infty\}}] \leq C \varsigma^{kp} \left(1 - \frac{\delta_0}{2}\right)^{k-1}. \quad (2.7.7)$$

Finally, by choosing $1 < \varsigma < \left(1 - \frac{\delta_0}{2}\right)^{-1/p}$, we get using (2.7.2) or (2.7.3) that for all

$p \in (0, \frac{\tilde{\alpha}}{\theta\sqrt{1}})$, there exists $C_p > 0$ such that

$$\mathbb{P}(\tau_\infty > n) \leq \mathbb{P}\left(\sum_{k=1}^{+\infty} \Delta\tau_k \mathbf{1}_{\{j^{(s)} > k\}} > n\right) \leq C_p n^{-p}. \quad (2.7.8)$$

It remains to optimize the upper-bound $\frac{\tilde{\alpha}}{\theta\sqrt{1}}$ for p . Since $\tilde{\alpha} := \min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2 - \varepsilon$ with $\varepsilon > 0$ as small as necessary and since by Proposition 2.6.1: $\theta > (2(\rho - \alpha))^{-1}$ and $\alpha \in (\frac{1}{2} \vee (\frac{3}{2} - \beta), \rho)$, we finally get (2.7.8) for all $p \in (0, v(\beta, \rho))$ where

$$v(\beta, \rho) = \sup_{\alpha \in (\frac{1}{2} \vee (\frac{3}{2} - \beta), \rho)} \min\{1, 2(\rho - \alpha)\}(\min\{\alpha, \beta, \alpha + \beta - 1\} - 1/2).$$

This concludes the proof of Theorem 2.2.1 in the polynomial case.

▷ Under $(\mathbf{H}_{\text{exp}})$:

The proof is almost the same. The only differences are that we use the following bound

$$\begin{aligned} \Delta\tau_k &= \tau_k - \tau_{k-1} \leq c_2(\ell + 1) + \Delta t_3^{(k)} \\ &= c_2(\ell + 1) + t_* + \zeta^k + \theta\ell \\ &\leq C_1 \zeta^k \theta\ell \quad (\text{for } C_1 \text{ large enough}) \end{aligned}$$

on the events $\mathcal{A}_{k,\ell}$ and the upperbound $\mathbb{P}(\mathcal{A}_{k,\ell} \mid \{\tau_{k-1} < +\infty\}) \leq 2^{-\tilde{\alpha}\ell}$ given in (2.7.5). And then we get for all $p > 0$ the existence of $C_p > 0$ such that

$$\mathbb{P}(\tau_\infty > n) \leq \mathbb{P}\left(\sum_{k=1}^{+\infty} \Delta\tau_k \mathbf{1}_{\{j^{(s)} > k\}} > n\right) \leq C_p n^{-p}. \quad (2.7.9)$$

by choosing $1 < \zeta < \left(1 - \frac{\delta_0}{2}\right)^{-1/p}$ and the proof of Theorem 2.2.1 is over.

2.A Proof of Theorem 2.2.2

The beginning of the following proof makes use of ideas developed in [15, 16].

Let us recall that $F_h(x, w) := x + hb(x) + \sigma(x)w$.

Proof. Set $V(x) = |x|$. Let us begin by proving that (\mathbf{H}_1) holds with V for F_h with $h > 0$ small enough. We have:

$$|F_h(x, w)|^2 = |x|^2 + h^2|b(x)|^2 + 2h\langle x, b(x) \rangle + 2\langle x, \sigma(x)w \rangle + 2h\langle b(x), \sigma(x)w \rangle + |\sigma(x)w|^2.$$

Then, using the inequality $|\langle a, b \rangle| \leq \frac{1}{2}(\varepsilon|a|^2 + \frac{1}{\varepsilon}|b|^2)$ for all $\varepsilon > 0$, we get

$$|\langle x, \sigma(x)w \rangle| \leq \frac{1}{2}(\varepsilon|x|^2 + \frac{1}{\varepsilon}|\sigma(x)w|^2) \quad \text{and} \quad |\langle b(x), \sigma(x)w \rangle| \leq \frac{1}{2}(\varepsilon|b(x)|^2 + \frac{1}{\varepsilon}|\sigma(x)w|^2).$$

Moreover, assumptions **(L1)** and **(L2)** on b give the existence of $\tilde{\beta} \in \mathbb{R}$, $\alpha > 0$ and $\tilde{C} > 0$ such that

$$|\langle b(x), x \rangle| \leq \tilde{\beta} - \tilde{\alpha}|x|^2 \quad \text{et} \quad |b(x)|^2 \leq \tilde{C}(1 + |x|^2).$$

Hence, we finally have

$$\begin{aligned} |F_h(x, w)|^2 &\leq |x|^2 + \tilde{C}h^2(1 + |x|^2) + 2h(\tilde{\beta} - \tilde{\alpha}|x|^2) + \varepsilon|x|^2 + \frac{1}{\varepsilon}|\sigma(x)w|^2 \\ &\quad + \tilde{C}h\varepsilon(1 + |x|^2) + \frac{h}{\varepsilon}|\sigma(x)w|^2 + |\sigma(x)w|^2 \\ &\leq |x|^2 + 2h(\tilde{\beta} - \tilde{\alpha}|x|^2) + \tilde{C}(\varepsilon + h\varepsilon + h^2)(1 + |x|^2) + \left(1 + \frac{h+1}{\varepsilon}\right)|\sigma(x)w|^2. \end{aligned}$$

Now, set $\varepsilon = h^2$ and choose $0 < h < \min\left\{\sqrt{1 + \frac{\tilde{\alpha}}{\tilde{C}}} - 1, \frac{1}{\tilde{\alpha}}\right\}$. Then, we have $\tilde{C}(\varepsilon + h\varepsilon + h^2) \leq \tilde{\alpha}h$ and $0 < 1 - \tilde{\alpha}h < 1$. Therefore,

$$|F_h(x, w)|^2 \leq |x|^2 + h(\tilde{\gamma} - \tilde{\alpha}|x|^2) + \left(1 + \frac{h+1}{\varepsilon}\right)|\sigma(x)w|^2$$

where $\tilde{\gamma} = 2\tilde{\beta} + \tilde{\alpha}$. Then

$$|F_h(x, w)|^2 \leq (1 - \tilde{\alpha}h)|x|^2 + h\tilde{\gamma} + \left(1 + \frac{h+1}{\varepsilon}\right)|\sigma(x)w|^2. \quad (2.A.1)$$

By assumption σ is a bounded function on \mathbb{R}^d . Then, there exists $C' > 0$ depending on h and σ such that

$$|F_h(x, w)|^2 \leq (1 - \tilde{\alpha}h)|x|^2 + C'(1 + |w|^2).$$

Using the classical inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, we finally get the existence of $\gamma \in (0, 1)$ and $C > 0$ such that for all $(x, w) \in \mathbb{R}^d \times \mathbb{R}^d$

$$|F_h(x, w)| \leq \gamma|x| + C(1 + |w|) \quad (2.A.2)$$

which achieves the proof of **(H1)**.

We now turn to the proof of **(H2)**. Let $K > 0$ and take $\mathbf{x} = (x, x', y, y') \in B(0, K)^4$. Here we take $\tilde{K} = K$. Hence, let us now define $\Lambda_{\mathbf{x}}$. For all $u \in B(0, K)$, we set

$$\Lambda_{\mathbf{x}}(u) = A(\mathbf{x})u + B(\mathbf{x}) \quad (2.A.3)$$

with $A(\mathbf{x}) := \sigma^{-1}(x')\sigma(x)$ and $B(\mathbf{x}) := \sigma^{-1}(x')(x - x' + h(b(x) - b(x'))) + \sigma^{-1}(x')\sigma(x)y - y'$.

Then, (2.A.3) is equivalent to $\tilde{F}_h(x, u, y) = \tilde{F}_h(x', \Lambda_{\mathbf{x}}(u), y')$ for all $u \in B(0, K)$. Hence, for all $u \in B(0, K)$,

$$J_{\Lambda_{\mathbf{x}}}(u) = \sigma^{-1}(x')\sigma(x). \quad (2.A.4)$$

Since σ, σ^{-1} and b are continuous, there exist $C_K > 0$ and $m_K > 0$ independent from \mathbf{x} such that for all $u \in B(0, K)$,

$$\begin{aligned} |\det(J_{\Lambda_{\mathbf{x}}}(u))| &\geq C_K \\ |\Lambda_{\mathbf{x}}(u) - u| &\leq m_K. \end{aligned}$$

For the sake of simplicity, let us set $\tilde{\Lambda}_{\mathbf{x}}(u) = \Lambda_{\mathbf{x}}(u) - B(\mathbf{x})$ and extend $\tilde{\Lambda}_{\mathbf{x}}$ to \mathbb{R}^d . Now, let $K_1 > 0$ be independent of \mathbf{x} such that $\sup_{u \in B(0, K)} |A(\mathbf{x})u| < K_1$ and set for all

$u \in \mathbb{R}^d \setminus B(0, K_1)$, $\tilde{\Lambda}_{\mathbf{x}}(u) = u$. Hence, $\tilde{\Lambda}_{\mathbf{x}}$ is a \mathcal{C}^1 -diffeomorphism from $B(0, K)$ to $\mathcal{E}_K := \{A(\mathbf{x})u \mid u \in B(0, K)\}$ and from $\mathbb{R}^d \setminus B(0, K_1)$ to itself. It remains to extend it with a \mathcal{C}^1 -diffeomorphism from $B(0, K_1) \setminus B(0, K)$ to $B(0, K_1) \setminus \mathcal{E}_K$. To this end, we consider q the positive definite quadratic form associated to the ellipsoid \mathcal{E}_K and we denote by $\mathcal{B}' := (e'_1, \dots, e'_d)$ the orthonormal basis which diagonalizes q , so that if $x = (x'_1, \dots, x'_d)$ in \mathcal{B}' we have $q(x) = \sum_{i=1}^d \lambda_i (x'_i)^2$ with $\lambda_i > 0$ for all $i \in \{1, \dots, d\}$. Let $\mathcal{B} := (e_1, \dots, e_d)$ be the canonical basis and $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be the linear application such that $\varphi(e_i) = \frac{1}{\sqrt{\lambda_i}} e'_i$ for all $i \in \{1, \dots, d\}$.

Remark 2.A.1. This application φ gives also a \mathcal{C}^1 -diffeomorphism from $B(0, K)$ to \mathcal{E}_K by construction and $|u| = K \iff \sqrt{|q(\varphi(u))|} = K$.

Now, set for all $u \in B(0, K_1) \setminus B(0, K)$,

$$\tilde{\Lambda}_{\mathbf{x}}(u) := \left[\left(1 - \frac{|u| - K}{K_1 - K} \right) \frac{K}{|u|} + \left(\frac{|u| - K}{K_1 - K} \frac{K_1}{|\varphi(u)|} \right) \right] \varphi(u).$$

This is just an interpolation between $\frac{K}{|u|}\varphi(u)$ and $\frac{K_1}{|\varphi(u)|}\varphi(u)$. It is a \mathcal{C}^1 -diffeomorphism from $B(0, K_1) \setminus B(0, K)$ to $B(0, K_1) \setminus \mathcal{E}_K$ and the inverse is given by, for all $v \in B(0, K_1) \setminus \mathcal{E}_K$:

$$\tilde{\Lambda}_{\mathbf{x}}^{-1}(v) := \left[\left(1 - \frac{|v| - \alpha(|v|)}{K_1 - \alpha(|v|)} \right) \frac{K}{\sqrt{q(v)}} + \left(\frac{|v| - \alpha(|v|)}{K_1 - \alpha(|v|)} \frac{K_1}{|\varphi^{-1}(v)|} \right) \right] \varphi^{-1}(v)$$

with $\alpha(|v|) := |v|/\sqrt{q(v)}$. Finally, one can check that we have all the elements to conclude that \tilde{F}_h satisfies **(H₂)**. \square

Remark 2.A.2. \triangleright If we relax the boundedness assumption on σ and assume that $|\sigma(x)| \leq C(1 + |x|^\kappa)$ for some $\kappa \in (0, 1)$, only the proof of **(H₁)** is changed. The beginning of the proof is exactly the same. From (2.A.1), we use the classical Young inequality $|ab| \leq \frac{1}{p} \frac{|a|^p}{\varepsilon^p} + \frac{1}{q} |b|^q \varepsilon^q$ with $p = \frac{1}{\kappa}$ and $q = \frac{1}{1-\kappa}$ for the term $|\sigma(x)w|^2$. Then, it suffices to use $|\sigma(x)| \leq C(1 + |x|^\kappa)$ and to calibrate ε to get an inequality of the type: $|F(x, w)|^2 \leq \gamma|x|^2 + C(1 + |w|^{\frac{2}{1-\kappa}})$. We conclude that **(H₁)** holds with $V(x) := |x|^{1-\kappa}$.

\triangleright Let us consider the family of functions given by $F(x, w) = f(b(x) + \sigma(x)w)$. Provided that f is well defined on \mathbb{R}^d , σ is continuously invertible and σ^{-1} and b are continuous

on \mathbb{R}^d , we can build a function $\Lambda_{\mathbf{x}}$ which satisfies **(H₂)** exactly in the same way as in the preceding proof.

2.B Explicit formula for the sequence $(b_k)_{k \geq 0}$

Theorem 2.B.1. *Let $(u_n)_{n \in \mathbb{N}}$ and $(v_n)_{n \in \mathbb{N}}$ be two sequences such that for $n \in \mathbb{N}$,*

$$u_n = \sum_{k=0}^n a_k v_{n-k} \quad (2.B.1)$$

then we have:

$$v_n = \sum_{k=0}^n b_k u_{n-k} \quad (2.B.2)$$

where

$$b_0 := \frac{1}{a_0} \quad \text{and} \quad \forall k \geq 1, \quad b_k := \sum_{p=1}^k \frac{(-1)^p}{a_0^{p+1}} \left(\sum_{\substack{k_1, \dots, k_p \geq 1 \\ k_1 + \dots + k_p = k}} \prod_{i=1}^p a_{k_i} \right).$$

Proof. It suffices to reverse a triangular Toeplitz matrix. Indeed, equation (2.B.1) is equivalent to:

$$\forall n \in \mathbb{N}, \quad \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{pmatrix} = \begin{pmatrix} a_0 & 0 & \dots & 0 \\ a_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n-1} & \dots & a_1 & a_0 \end{pmatrix} \begin{pmatrix} g_0 \\ g_1 \\ \vdots \\ g_n \end{pmatrix}. \quad (2.B.3)$$

Denote by A the matrix associated to the system. Denote by N the following nilpotent matrix:

$$N = \begin{pmatrix} 0 & \dots & \dots & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ (0) & & 1 & 0 \end{pmatrix}.$$

Then, $A = a_0 I_n + a_1 N + \dots + a_{n-1} N^{n-1}$ and we are looking for B such that $B = b_0 I_n + b_1 N + \dots + b_{n-1} N^{n-1}$ and $AB = I_n$. Denote by

$$S(z) = \sum_{k \geq 0} a_k z^k \quad \text{and} \quad S^{-1}(z) = \sum_{k \geq 0} b_k z^k,$$

we are interested in the $(n-1)$ first coefficients of $S^{-1}(z)$.

And formally,

$$\begin{aligned}
 S^{-1}(z) &= \frac{1}{S(z)} = \frac{1}{a_0} \left(\frac{1}{1 + \sum_{k \geq 1} \frac{a_k}{a_0} z^k} \right) \\
 &= \frac{1}{a_0} \sum_{p \geq 0} \frac{(-1)^p}{a_0^p} \left(\sum_{k \geq 1} a_k z^k \right)^p \\
 &= \frac{1}{a_0} + \sum_{p \geq 1} \frac{(-1)^p}{a_0^{p+1}} \sum_{k \geq p} \left(\sum_{\substack{k_1, \dots, k_p \geq 1 \\ k_1 + \dots + k_p = k}} a_{k_1} a_{k_2} \dots a_{k_p} \right) z^k \\
 &= \frac{1}{a_0} + \sum_{k \geq 1} \sum_{p=1}^k \left(\frac{(-1)^p}{a_0^{p+1}} \sum_{\substack{k_1, \dots, k_p \geq 1 \\ k_1 + \dots + k_p = k}} a_{k_1} a_{k_2} \dots a_{k_p} \right) z^k.
 \end{aligned}$$

Finally, we identify the desired coefficients. □

2.C Particular case: when the sequence $(a_k)_{k \geq 0}$ is log-convex

This section is based on a work made by N.Ford, D.V.Savostyanov and N.L.Zamarashkin in [27].

Lemma 2.C.1. *Let $(a_n)_{n \in \mathbb{N}}$ be a log-convex sequence in the following sense*

$$a_n \geq 0 \quad \text{for } n \geq 0 \quad \text{and} \quad a_n^2 \leq a_{n-1} a_{n+1} \quad \text{for } n \geq 1.$$

If $a_0 > 0$, then the sequence $(b_n)_{n \in \mathbb{N}}$ defined by

$$b_0 = \frac{1}{a_0} \quad \text{and} \quad \forall n \geq 1, \quad b_n = -\frac{1}{a_0} \sum_{k=1}^n a_k b_{n-k}$$

satisfies

$$\forall n \geq 1, \quad b_n \leq 0 \quad \text{and} \quad |b_n| \leq b_0 a_n \tag{2.C.1}$$

Remark 2.C.1. The sequence $a_n = (n+1)^{-\rho}$ is log-convex for all $\rho > 0$, then the corresponding $(b_n)_{n \in \mathbb{N}}$ is such that $\forall n \in \mathbb{N}, \quad |b_n| \leq (n+1)^{-\rho}$.

Proof. Without loss of generality, we assume that $a_0 = 1$.

- First, following Theorem 4 of [27] we can prove by strong induction that for all $n \geq 1$, $b_n \leq 0$.
- The second property satisfied by (b_n) directly follows from the first one. Let $n \geq 1$, as

we just saw $b_n \leq 0$ therefore $|b_n| = -b_n$. But,

$$\begin{aligned} -b_n &= \sum_{k=1}^{n-1} a_k \underbrace{(b_{n-k})}_{\leq 0} + a_n b_0 \\ &\leq b_0 a_n \end{aligned}$$

and the lemma is proven. \square

2.D Proof of Proposition 2.2.3

We recall that $\rho = 3/2 - H$ where $H \in (0, 1/2)$ is the Hurst parameter and $(b_n)_{n \in \mathbb{N}}$ is defined by

$$b_0 = \frac{1}{a_0^H} \quad \text{and for } n \geq 1, \quad b_n = -\frac{1}{a_0^H} \sum_{k=1}^n a_k^H b_{n-k}. \quad (2.D.1)$$

and for all $k \geq 1$,

$$\frac{a_k^H}{a_0^H} = \left((2k+1)^{1-\rho} - (2k-1)^{1-\rho} \right).$$

We want to show that $|b_n| \leq C_b(n+1)^{-(2-\rho)}$ by induction. To this end we only need to prove that for n large enough,

$$S_n := \sum_{k=1}^n \left((2k-1)^{1-\rho} - (2k+1)^{1-\rho} \right) (n+1-k)^{-(2-\rho)} \leq (n+1)^{-(2-\rho)}. \quad (2.D.2)$$

For the sake of simplicity we assume that n is even.

$$\begin{aligned} S_n &= \sum_{k=1}^{n/2} \left((2k-1)^{1-\rho} - (2k+1)^{1-\rho} \right) (n+1-k)^{-(2-\rho)} \\ &\quad + \sum_{k=n/2+1}^n \left((2k-1)^{1-\rho} - (2k+1)^{1-\rho} \right) (n+1-k)^{-(2-\rho)} \\ S_n &=: S_n^{(1)} + S_n^{(2)}. \end{aligned} \quad (2.D.3)$$

▷ We begin with $S_n^{(1)}$. Summation by parts:

$$\begin{aligned}
 S_n^{(1)} &= \left(\frac{n}{2} + 1\right)^{-(2-\rho)} (1 - (n+1)^{1-\rho}) - \sum_{k=1}^{n/2-1} (1 - (2k+1)^{1-\rho}) \left((n-k)^{-(2-\rho)} - (n+1-k)^{-(2-\rho)} \right) \\
 &= \left(\frac{n}{2} + 1\right)^{-(2-\rho)} (1 - (n+1)^{1-\rho}) - \left(\left(\frac{n}{2} + 1\right)^{-(2-\rho)} - n^{-(2-\rho)} \right) \\
 &\quad + \sum_{k=1}^{n/2-1} (2k+1)^{1-\rho} \left((n-k)^{-(2-\rho)} - (n+1-k)^{-(2-\rho)} \right) \\
 &= n^{-(2-\rho)} - \left(\frac{n}{2} + 1\right)^{-(2-\rho)} (n+1)^{1-\rho} + \sum_{k=1}^{n/2-1} (2k+1)^{1-\rho} \left((n-k)^{-(2-\rho)} - (n+1-k)^{-(2-\rho)} \right) \\
 S_n^{(1)} &= n^{-(2-\rho)} - \left(\frac{n}{2}\right)^{-(2-\rho)} (n+1)^{1-\rho} + \sum_{k=1}^{n/2} (2k+1)^{1-\rho} \left((n-k)^{-(2-\rho)} - (n+1-k)^{-(2-\rho)} \right).
 \end{aligned} \tag{2.D.4}$$

We set $\tilde{S}_n := \sum_{k=1}^{n/2} (2k+1)^{1-\rho} \left((n-k)^{-(2-\rho)} - (n+1-k)^{-(2-\rho)} \right)$. Then,

$$\begin{aligned}
 \tilde{S}_n &= \frac{1}{n} \sum_{k=1}^{n/2} \left(\frac{2k+1}{n}\right)^{1-\rho} \left(\left(1 - \frac{k}{n}\right)^{-(2-\rho)} - \left(1 - \frac{k-1}{n}\right)^{-(2-\rho)} \right) \\
 &= \frac{2^{1-\rho}}{n} \sum_{k=1}^{n/2} \left(\frac{k+1/2}{n}\right)^{1-\rho} \left(\left(1 - \frac{k}{n}\right)^{-(2-\rho)} - \left(1 - \frac{k-1}{n}\right)^{-(2-\rho)} \right).
 \end{aligned}$$

Moreover,

$$\begin{aligned}
 &\sum_{k=1}^{n/2} \left(\frac{k+1/2}{n}\right)^{1-\rho} \left(\left(1 - \frac{k}{n}\right)^{-(2-\rho)} - \left(1 - \frac{k-1}{n}\right)^{-(2-\rho)} \right) \\
 &= (2-\rho) \left(\int_0^{1/2} \left(x + \frac{1}{2n}\right)^{1-\rho} (1-x)^{-(3-\rho)} dx \right. \\
 &\quad \left. - \sum_{k=1}^{n/2} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\left(x + \frac{1}{2n}\right)^{1-\rho} - \left(\frac{k+1/2}{n}\right)^{1-\rho} \right) (1-x)^{-(3-\rho)} dx \right)
 \end{aligned} \tag{2.D.5}$$

and

$$\begin{aligned}
 \int_0^{1/2} \left(x + \frac{1}{2n}\right)^{1-\rho} (1-x)^{-(3-\rho)} dx &= \left[\frac{(1-x)^{\rho-2} \left(x + \frac{1}{2n}\right)^{2-\rho}}{(2-\rho) \left(1 + \frac{1}{2n}\right)} \right]_0^{1/2} \\
 &= \frac{1}{2-\rho} \left(1 + \frac{1}{2n}\right)^{-1} \left[\left(1 + \frac{1}{n}\right)^{2-\rho} - \left(\frac{1}{2n}\right)^{2-\rho} \right].
 \end{aligned} \tag{2.D.6}$$

Hence by putting together (2.D.5) and (2.D.6) we get

$$\begin{aligned} \tilde{S}_n = & \frac{2^{1-\rho}}{n} \left(\left(1 + \frac{1}{2n}\right)^{-1} \left[\left(1 + \frac{1}{n}\right)^{2-\rho} - \left(\frac{1}{2n}\right)^{2-\rho} \right] \right. \\ & \left. - (2-\rho) \sum_{k=1}^{n/2} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\left(x + \frac{1}{2n}\right)^{1-\rho} - \left(\frac{k+1/2}{n}\right)^{1-\rho} \right) (1-x)^{-(3-\rho)} dx \right). \end{aligned} \quad (2.D.7)$$

We deduce from (2.D.4) and (2.D.7) that

$$S_n^{(1)} \leq n^{-(2-\rho)} - \left(\frac{n}{2}\right)^{-(2-\rho)} (n+1)^{1-\rho} + \frac{2^{1-\rho}}{n} \left(1 + \frac{1}{2n}\right)^{-1} \left[\left(1 + \frac{1}{n}\right)^{2-\rho} - \left(\frac{1}{2n}\right)^{2-\rho} \right]. \quad (2.D.8)$$

▷ Now, we look after $S_n^{(2)}$:

As before, using the fact that

$$\begin{aligned} \int_0^{1/2} \left(1 - x + \frac{1}{2n}\right)^{-\rho} x^{-(2-\rho)} dx &= \left[\frac{\left(1 - x + \frac{1}{2n}\right)^{1-\rho} x^{\rho-1}}{(\rho-1) \left(1 + \frac{1}{2n}\right)} \right]_0^{1/2} \\ &= \frac{1}{\rho-1} \left(1 + \frac{1}{2n}\right)^{-1} \left(1 + \frac{1}{n}\right)^{1-\rho} \end{aligned} \quad (2.D.9)$$

we get

$$\begin{aligned} S_n^{(2)} = & \frac{2^{1-\rho}}{n} \left(\left(1 + \frac{1}{2n}\right)^{-1} \left(1 + \frac{1}{n}\right)^{1-\rho} \right. \\ & \left. - (\rho-1) \sum_{k=1}^{n/2} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(x^{-(2-\rho)} - \left(\frac{k}{n}\right)^{-(2-\rho)} \right) \left(1 - x + \frac{1}{2n}\right)^{-\rho} dx \right). \end{aligned} \quad (2.D.10)$$

Now, for all $k \in \{1, \dots, n/2\}$, we set

$$I_k := \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(x^{-(2-\rho)} - \left(\frac{k}{n}\right)^{-(2-\rho)} \right) \left(1 - x + \frac{1}{2n}\right)^{-\rho} dx.$$

Thanks to the substitution $t = x - \frac{k-1}{n}$, we have

$$I_k = \int_0^{1/n} \left(\left(t + \frac{k-1}{n}\right)^{-(2-\rho)} - \left(\frac{k}{n}\right)^{-(2-\rho)} \right) \left(1 + \frac{1}{2n} - t - \frac{k-1}{n}\right)^{-\rho} dt$$

Taylor-Lagrange expansion:

$$\bullet \left(t + \frac{k-1}{n}\right)^{-(2-\rho)} - \left(\frac{k}{n}\right)^{-(2-\rho)} = \left(\frac{1}{n} - t\right) (2-\rho) \left(\frac{k}{n}\right)^{-(3-\rho)} + \frac{1}{2} \left(\frac{1}{n} - t\right)^2 (2-\rho)(3-\rho) \xi^{-(4-\rho)}$$

with $\xi \in]t + (k-1)/n, k/n[$.

$$\bullet \left(1 + \frac{1}{2n} - t - \frac{k-1}{n}\right)^{-\rho} = \left(1 + \frac{1}{2n} - \frac{k-1}{n}\right)^{-\rho} + t\rho \left(1 + \frac{1}{2n} - c\right)^{-\rho-1}$$

with $c \in](k-1)/n, t + (k-1)/n[$.

Therefore, we deduce that

$$I_k \geq \int_0^{1/n} \left(\frac{1}{n} - t\right) (2-\rho) \left(\frac{k}{n}\right)^{-(3-\rho)} \left(1 + \frac{1}{2n} - \frac{k-1}{n}\right)^{-\rho} dt = \frac{2-\rho}{2n^2} \left(\frac{k}{n}\right)^{-(3-\rho)} \left(1 + \frac{1}{2n} - \frac{k-1}{n}\right)^{-\rho}.$$

Then we add the inequality for k from 1 to $n/2$,

$$\sum_{k=1}^{n/2} I_k \geq \frac{2-\rho}{2n} \times \underbrace{\frac{1}{n} \sum_{k=1}^{n/2} \left(\frac{k}{n}\right)^{-(3-\rho)} \left(1 + \frac{1}{2n} - \frac{k-1}{n}\right)^{-\rho}}_{U_n}. \quad (2.D.11)$$

We easily show that

$$U_n \geq \int_0^{1/2} \left(y + \frac{1}{n}\right)^{-(3-\rho)} \left(1 + \frac{3}{2n} - y\right)^{-\rho} dy =: J_n. \quad (2.D.12)$$

By integration by parts on J_n we get:

$$\begin{aligned} J_n &= \left[\frac{-(y + \frac{1}{n})^{-(2-\rho)}}{2-\rho} \left(1 + \frac{3}{2n} - y\right)^{-\rho} \right]_0^{1/2} \\ &\quad + \frac{\rho}{2-\rho} \int_0^{1/2} \left(y + \frac{1}{n}\right)^{-(2-\rho)} \left(1 + \frac{3}{2n} - y\right)^{-\rho-1} dy \\ &= \frac{1}{2-\rho} \underbrace{\left[\left(\frac{1}{n}\right)^{-(2-\rho)} \left(1 + \frac{3}{2n}\right)^{-\rho} - \left(\frac{1}{2} + \frac{1}{n}\right)^{-(2-\rho)} \left(\frac{1}{2} + \frac{3}{2n}\right)^{-\rho} \right]}_{\underset{n \rightarrow +\infty}{\sim} n^{2-\rho}} \\ &\quad + \frac{\rho}{2-\rho} \underbrace{\int_0^{1/2} \left(y + \frac{1}{n}\right)^{-(2-\rho)} \left(1 + \frac{3}{2n} - y\right)^{-\rho-1} dy}_{\underset{n \rightarrow +\infty}{\rightarrow} \int_0^{1/2} y^{-(2-\rho)} (1-y)^{-\rho-1} dy}. \end{aligned}$$

Hence, for n large enough, we have

$$U_n \geq J_n \geq \frac{1}{2(2-\rho)} n^{2-\rho} \quad (2.D.13)$$

By combining (2.D.11) and (2.D.13) we get for n large enough

$$\sum_{k=1}^{n/2} I_k \geq \frac{1}{4} n^{1-\rho}.$$

Finally we get for $S_n^{(2)}$ the following upper-bound for n large enough,

$$S_n^{(2)} \leq \frac{2^{1-\rho}}{n} \left(\left(1 + \frac{1}{2n}\right)^{-1} \left(1 + \frac{1}{n}\right)^{1-\rho} - \frac{\rho-1}{4} n^{1-\rho} \right). \quad (2.D.14)$$

By putting together (2.D.8) and (2.D.14) and when factoring by $(n+1)^{-(2-\rho)}$, we get for S_n

$$S_n \leq (n+1)^{-(2-\rho)} \left(1 - \frac{1}{n}\right)^{-(2-\rho)} \times u_n \quad (2.D.15)$$

with

$$\begin{aligned} u_n = & 1 - 2^{2-\rho} (n+1)^{1-\rho} \\ & + 2^{1-\rho} n^{1-\rho} \left[\left(1 + \frac{1}{2n}\right)^{-1} \left(\left(1 + \frac{1}{n}\right)^{2-\rho} - \left(\frac{1}{2n}\right)^{2-\rho} \right) \right. \\ & \left. + \left(1 + \frac{1}{2n}\right)^{-1} \left(1 + \frac{1}{n}\right)^{1-\rho} - \frac{\rho-1}{4} n^{1-\rho} \right] \end{aligned}$$

Lastly, we have the following asymptotic expansion:

$$\left(1 - \frac{1}{n}\right)^{-(2-\rho)} \times u_n = 1 - \frac{2^{1-\rho}(\rho-1)}{4} n^{2-2\rho} + o\left(\frac{1}{n}\right)$$

Since $\rho \in (1, 3/2)$ we have $2 - 2\rho \in (-1, 0)$ therefore for n large enough we conclude that

$$S_n \leq (n+1)^{-(2-\rho)}.$$

2.E Proof of Theorem 2.3.1

Let $x_0 \in \mathcal{X}$ and $\mu = \delta_{x_0} \times \mathbb{P}_w$. We have $\Pi_{\mathcal{W}}^* \mu = \mathbb{P}_w$ therefore by Property 2.3.1 we get $\forall k \in \mathbb{N}$, $\Pi_{\mathcal{W}}^*(\mathcal{Q}^k \mu) = \mathbb{P}_w$. Moreover, we clearly have $\int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mu)(dx) = \psi(x_0) < +\infty$. We now set for all $n \in \mathbb{N}^*$,

$$R_n \mu = \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{Q}^k \mu.$$

The aim is to prove that the sequence $(R_n \mu)_{n \in \mathbb{N}^*}$ is tight. First, let us prove that $(\Pi_{\mathcal{X}}^* R_n \mu)_{n \in \mathbb{N}^*}$ is tight.

By Definition 2.3.1 (ii), we have $\forall k \geq 0$:

$$\int_{\mathcal{X} \times \mathcal{W}} \psi(x) \mathcal{Q}^{k+1} \mu(dx, dw) - \alpha \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mathcal{Q}^k \mu)(dx) \leq \beta.$$

By adding for k from 0 to $n - 1$, dividing by n and reordering the terms, we get:

$$\begin{aligned}
 (1 - \alpha) \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* R_n \mu)(dx) \\
 + \frac{1}{n} \sum_{k=0}^{n-1} \left(\int_{\mathcal{X} \times \mathcal{W}} \psi(x) \mathcal{Q}^k \mu(dx, dw) - \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mathcal{Q}^k \mu)(dx) \right) \\
 + \frac{1}{n} \int_{\mathcal{X} \times \mathcal{W}} \psi(x) \mathcal{Q}^{n+1} \mu(dx, dw) - \frac{1}{n} \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mu)(dx) \leq \beta. \quad (2.E.1)
 \end{aligned}$$

Since we are in a Polish space (here $\mathcal{X} \times \mathcal{W}$) we can “disintegrate” $\mathcal{Q}^k \mu$ for all $k \in \{0, \dots, n - 1\}$ (see [1] for background):

$$\mathcal{Q}^k \mu(dx, dw) = (\mathcal{Q}^k \mu)^x(dw) (\Pi_{\mathcal{X}}^* \mathcal{Q}^k \mu)(dx).$$

By integrating first with respect to w and then with respect to x , we get:

$$\int_{\mathcal{X} \times \mathcal{W}} \psi(x) \mathcal{Q}^k \mu(dx, dw) - \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mathcal{Q}^k \mu)(dx) = 0.$$

Let us return to (2.E.1),

$$\begin{aligned}
 (1 - \alpha) \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* R_n \mu)(dx) &\leq \beta + \frac{1}{n} \left(\int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mu)(dx) - \int_{\mathcal{X} \times \mathcal{W}} \psi(x) \mathcal{Q}^{n+1} \mu(dx, dw) \right) \\
 &= \beta + \frac{1}{n} \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mu - \Pi_{\mathcal{X}}^* (\mathcal{Q}^{n+1} \mu))(dx).
 \end{aligned}$$

Set $A_{n+1} = \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* \mathcal{Q}^{n+1} \mu)(dx)$. By Definition 2.3.1 (ii) and by induction, we have

$$0 \leq \frac{A_{n+1}}{n} \leq \frac{\beta}{n} \sum_{k=0}^n \alpha^k + \frac{\alpha^{n+1}}{n} A_0 = \frac{\beta}{n} \frac{1 - \alpha^{n+1}}{1 - \alpha} + \frac{\alpha^{n+1}}{n} \psi(x_0).$$

Hence we deduce that $\lim_{n \rightarrow +\infty} \frac{A_{n+1}}{n} = 0$. Then, there exists $C > 0$ such that $\forall n \in \mathbb{N}^*$:

$$(1 - \alpha) \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* R_n \mu)(dx) \leq C$$

and then

$$\sup_{n \geq 1} \int_{\mathcal{X}} \psi(x) (\Pi_{\mathcal{X}}^* R_n \mu)(dx) \leq \frac{C}{1 - \alpha}.$$

Let $\delta > 0$ and $K_\delta := \{x \in \mathcal{X} \mid \psi(x) \leq \delta\} = \psi^{-1}([0, \delta])$. By Definition 2.3.1 (ii), K_δ is a compact set. For all $x \in \mathcal{X}$, we have $\mathbb{1}_{K_\delta^c}(x) \leq \frac{\psi(x)}{\delta}$, so

$$\forall n \in \mathbb{N}^*, \quad (\Pi_{\mathcal{X}}^* R_n \mu)(K_\delta^c) \leq \frac{C}{\delta(1 - \alpha)}.$$

By setting $\frac{\varepsilon}{2} = \frac{C}{\delta(1 - \alpha)}$, we deduce that $(\Pi_{\mathcal{X}}^* R_n \mu)_{n \in \mathbb{N}^*}$ is tight.

Let us now go back to the tightness of $(R_n \mu)_{n \in \mathbb{N}^*}$. Let K be a compact set of \mathcal{W} such

that $\mathbb{P}_w(K^c) < \frac{\varepsilon}{2}$, this is possible since \mathcal{W} is Polish. We then get

$$\begin{aligned} R_n\mu((K_\delta \times K)^c) &\leq R_n\mu(K_\delta^c \times \mathcal{W}) + R_n\mu(\mathcal{X} \times K^c) \\ &= (\Pi_{\mathcal{X}}^* R_n\mu)(K_\delta^c) + (\Pi_{\mathcal{W}}^* R_n\mu)(K^c) \\ &= (\Pi_{\mathcal{X}}^* R_n\mu)(K_\delta^c) + (\mathbb{P}_w)(K^c) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Finally, $(R_n\mu)_{n \in \mathbb{N}^*}$ is tight. Let μ_\star be one of its accumulation points. By the Krylov-Bogoliubov criterium we deduce that μ_\star is an invariant distribution for \mathcal{Q} .

2.F Proof of Lemma 2.5.3

In this section we will prove a slightly more precise result than Lemma 2.5.3 which is the following: for all $\alpha, \beta > 0$ such that $\alpha + \beta > 1$, there exists $C(\alpha, \beta) > 0$ such that for all $n \geq 0$,

$$\sum_{k=0}^n (k+1)^{-\beta} (n+1-k)^{-\alpha} \leq C(\alpha, \beta) \begin{cases} (n+1)^{-\beta} \ln(n) & \text{if } \alpha = 1 \text{ and } \beta \leq 1 \\ (n+1)^{-\alpha} \ln(n) & \text{if } \beta = 1 \text{ and } \alpha \leq 1 \\ (n+1)^{-\min\{\alpha, \beta, \alpha+\beta-1\}} & \text{else} \end{cases}. \quad (2.F.1)$$

For the sake of simplicity, we will prove this result when n is odd. If n is even, the proof is almost the same. Set $N := \frac{n+1}{2}$. Then, we get

$$\begin{aligned} \sum_{k=0}^n (k+1)^{-\beta} (n+1-k)^{-\alpha} &= \sum_{k=1}^{n+1} k^{-\beta} (n+2-k)^{-\alpha} \\ &= \sum_{k=1}^N k^{-\beta} (n+2-k)^{-\alpha} + \sum_{k=N+1}^{n+1} k^{-\beta} (n+2-k)^{-\alpha} \\ &= \sum_{k=1}^N k^{-\beta} (n+2-k)^{-\alpha} + \sum_{k=1}^N k^{-\alpha} (n+2-k)^{-\beta} \\ \sum_{k=0}^n (k+1)^{-\beta} (n+1-k)^{-\alpha} &= S_N(\beta, \alpha) + S_N(\alpha, \beta) \end{aligned} \quad (2.F.2)$$

by setting $S_N(\alpha, \beta) := \sum_{k=1}^N k^{-\alpha} (n+2-k)^{-\beta} = \sum_{k=1}^N k^{-\alpha} (2N - (k-1))^{-\beta}$.

▷ If $\alpha \in (0, 1)$, we have $\alpha + \beta - 1 \leq \beta$ and $S_N(\alpha, \beta) \leq \tilde{C}(\alpha, \beta) (n+1)^{-(\alpha+\beta-1)}$.

Indeed,

$$\begin{aligned}
 S_N(\alpha, \beta) &= \sum_{k=1}^N k^{-\alpha} (2N - (k-1))^{-\beta} \\
 &= N^{-(\alpha+\beta-1)} \times \frac{1}{N} \sum_{k=1}^N \left(\frac{k}{N}\right)^{-\alpha} \left(2 - \frac{k-1}{N}\right)^{-\beta} \\
 &\leq N^{-(\alpha+\beta-1)} \times \frac{1}{N} \sum_{k=1}^N \left(\frac{k}{N}\right)^{-\alpha} \left(2 - \frac{k}{N}\right)^{-\beta}
 \end{aligned}$$

and

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{k=1}^N \left(\frac{k}{N}\right)^{-\alpha} \left(2 - \frac{k}{N}\right)^{-\beta} = \int_0^1 x^{-\alpha} (2-x)^{-\beta} dx$$

where the integral is well defined since $\alpha \in (0, 1)$. Therefore, since $N = \frac{n+1}{2}$, we deduce that there exists $\tilde{C}(\alpha, \beta) > 0$ such that $S_N(\alpha, \beta) \leq \tilde{C}(\alpha, \beta)(n+1)^{-(\alpha+\beta-1)}$.

▷ If $\alpha > 1$, we have $\alpha + \beta - 1 > \beta$ and $S_N(\alpha, \beta) \leq \tilde{C}(\alpha, \beta)(n+1)^{-\beta}$.

Indeed,

$$\begin{aligned}
 S_N(\alpha, \beta) &= \sum_{k=1}^N k^{-\alpha} (2N - (k-1))^{-\beta} \\
 &\leq (2N - (N-1))^{-\beta} \sum_{k=1}^N k^{-\alpha} \\
 &\leq (N+1)^{-\beta} \sum_{k=1}^{+\infty} k^{-\alpha}.
 \end{aligned}$$

Therefore, as before we deduce that there exists $\tilde{C}(\alpha, \beta) > 0$ such that $S_N(\alpha, \beta) \leq \tilde{C}(\alpha, \beta)(n+1)^{-\beta}$.

▷ If $\alpha = 1$, in the same way as in the case $\alpha > 1$, we get

$$\begin{aligned}
 S_N(\alpha, \beta) &\leq (N+1)^{-\beta} \sum_{k=1}^N \frac{1}{k} \\
 &\leq \tilde{C}(N+1)^{-\beta} \ln(N)
 \end{aligned}$$

Therefore, there exists $\tilde{C}(\alpha, \beta) > 0$ such that $S_N(\alpha, \beta) \leq \tilde{C}(\alpha, \beta)(n+1)^{-\beta} \ln(n)$.

Finally, we get that for all $\alpha > 0$ and $\beta > 0$ such that $\alpha + \beta > 1$,

$$S_N(\alpha, \beta) \leq \tilde{C}(\alpha, \beta) \begin{cases} (n+1)^{-\min\{\alpha, \beta, \alpha+\beta-1\}} & \text{if } \alpha \neq 1 \\ (n+1)^{-\beta} \ln(n) & \text{if } \alpha = 1 \end{cases} \quad (2.F.3)$$

Putting this inequality into (2.F.2) we finally get the desired inequality and the proof is finished.

2.G Proof of Hypothesis (\mathbf{H}'_1)

We recall that we want to prove that under (\mathbf{H}_{poly}), the following hypothesis is true:

(\mathbf{H}'_1): Let $\gamma \in (0, 1)$. There exists $C_\gamma > 0$ such that for all $j \in \mathbb{N}$, for every $K > 0$ and for $i = 1, 2$,

$$\mathbb{E} \left[\sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j-l} |\Delta_{\tau_{j-1}+l}^i| \middle| \mathcal{E}_j \right] < C_\gamma$$

where $\Delta\tau_j := \tau_j - \tau_{j-1}$ and Δ^i is the stationary Gaussian sequence defined in Equation (2.2.2).

Remark 2.G.1. Since the proof of this assumption will exclusively use the domination assumption on $(a_k)_{k \geq 0}$ and since $(\tilde{a}_k)_{k \geq 0} := (a_{k+1})_{k \geq 0}$ satisfies the same domination assumption, we will also get that for $i = 1, 2$,

$$\mathbb{E} \left[\sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j-l} |\tilde{\Delta}_{\tau_{j-1}+l}^i| \middle| \mathcal{E}_j \right] < C_\gamma$$

where $\tilde{\Delta}_{\tau_{j-1}+l}^i = \sum_{k=0}^{+\infty} a_{k+1} \xi_{\tau_{j-1}+l-k}^i$. Hence, we will get that for $i = 1, 2$

$$\mathbb{E} \left[|\tilde{\Delta}_{\tau_j}^i| \middle| \mathcal{E}_j \right] = \mathbb{E} \left[\left| \sum_{k=0}^{+\infty} a_{k+1} \xi_{\tau_j-k}^i \right| \middle| \mathcal{E}_j \right] < C_\gamma.$$

Then, by the Markov inequality we finally get the second part of Equation (2.6.13).

We now turn to the proof of (\mathbf{H}'_1). We work on the set $\mathcal{E}_j = \{\tau_j < +\infty\}$. We have

$$\sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j-l} |\Delta_{\tau_{j-1}+l}^i| = \sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j-u} |\Delta_u^i|.$$

But,

$$|\Delta_u^i| = \left| \sum_{k=-\infty}^{k=u} a_{u-k} \xi_k^i \right| = \left| \sum_{m=0}^j \Lambda_m^i(u) \right|.$$

where

$$\Lambda_m^i(u) = \sum_{k=\tau_{m-1}+1}^{\tau_m} a_{u-k} \xi_k^i \quad \text{pour } m \in \{1, \dots, j-1\}, \quad (2.G.1)$$

$$\Lambda_0^i(u) = \sum_{k=-\infty}^{k=\tau_0} a_{u-k} \xi_k^i \quad \text{et} \quad \Lambda_j^i(u) = \sum_{k=\tau_{j-1}+1}^u a_{u-k} \xi_k^i. \quad (2.G.2)$$

With these notations, we get the following upper-bound

$$\sum_{l=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j-l} |\Delta_{\tau_{j-1}+l}^i| \leq \sum_{m=0}^j \sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j-u} |\Lambda_m^i(u)|. \quad (2.G.3)$$

The goal of the following lemmas is to get an upper-bound of the quantity $\mathbb{E}[\sup_{u \in \llbracket \tau_{j-1}+1, \tau_j \rrbracket} |\Lambda_m^i(u)| | \mathcal{E}_j]$ when $m \in \{0, \dots, j-1\}$.

Lemma 2.G.1. *Assume (\mathbf{H}_{poly}). Let $t_0, t_1 \in \mathbb{Z}$ and $u \in \mathbb{N}$ such that $t_0 < t_1 < u$. Let $(\xi_k)_{k \in \mathbb{Z}}$ be a sequence with values in \mathbb{R}^d . Then,*

$$\begin{aligned} \left| \sum_{k=t_0}^{t_1} a_{u-k} \xi_k \right| &\leq C_\rho (u+1-t_0)^{-\rho} \left| \sum_{k=t_0}^{t_1} \xi_k \right| + C_\kappa \sum_{k=1}^{t_1-t_0} \left| \sum_{l=k+t_0}^{t_1} \xi_l \right| (u+1-t_0-k)^{-\kappa} \\ &\leq C_\rho (u+1-t_0)^{-\rho} \left| \sum_{k=t_0}^{t_1} \xi_k \right| + C_\kappa \sum_{k=1}^{t_1-t_0} \left| \sum_{l=k+t_0}^{t_1} \xi_l \right| (u+1-t_0-k)^{-(\rho+1)}. \end{aligned}$$

Remark 2.G.2. The last inequality just follows from the fact that $\kappa \geq \rho + 1$ by assumption.

Proof. The proof is essentially based on a summation by parts argument. We set

$$\sum_{k=t_0}^{t_1} a_{u-k} \xi_k = \sum_{k=0}^{t_1-t_0} \underbrace{a_{u-t_0-k}}_{=: a'_k} \underbrace{\xi_{t_0+k}}_{=: \xi'_k}$$

and

$$\tilde{B}_k := \sum_{l=k}^{t_1-t_0} \xi'_l \quad \text{for } k \in \llbracket 0, t_1 - t_0 \rrbracket.$$

We then have

$$\begin{aligned} \sum_{k=0}^{t_1-t_0} a'_k \xi'_k &= \sum_{k=0}^{t_1-t_0-1} a'_k (\tilde{B}_k - \tilde{B}_{k+1}) + a'_{t_1-t_0} \xi'_{t_1-t_0} \\ &= \sum_{k=0}^{t_1-t_0} a'_k \tilde{B}_k - \sum_{k=1}^{t_1-t_0} a'_{k-1} \tilde{B}_k \\ &= a'_0 \tilde{B}_0 + \sum_{k=1}^{t_1-t_0} (a'_k - a'_{k-1}) \tilde{B}_k \\ \sum_{k=0}^{t_1-t_0} a'_k \xi'_k &= a_{u-t_0} \left(\sum_{k=t_0}^{t_1} \xi_k \right) + \sum_{k=1}^{t_1-t_0} \left(\sum_{l=k+t_0}^{t_1} \xi_l \right) [a_{u-t_0-k} - a_{u-t_0-(k-1)}]. \end{aligned}$$

Finally, by using triangular inequality and (\mathbf{H}_{poly}) we deduce that

$$\left| \sum_{k=t_0}^{t_1} a_{u-k} \xi_k \right| \leq C_\rho (u+1-t_0)^{-\rho} \left| \sum_{k=t_0}^{t_1} \xi_k \right| + C_\kappa \sum_{k=1}^{t_1-t_0} \left| \sum_{l=k+t_0}^{t_1} \xi_l \right| (u+1-t_0-k)^{-\kappa}.$$

□

In the next lemma we adopt the convention $\sum_{\emptyset} = 1$. Moreover, recall that by Proposition 2.6.1, we have for every $j \in \mathbb{N}^*$, $\Delta\tau_j \geq \varsigma^j$ for an arbitrary $\varsigma > 1$.

Lemma 2.G.2. *Assume $(\mathbf{H}_{\text{poly}})$. We suppose that $\tau_0 = 0$ and that there exists $\delta_1 \in (0, 1)$ such that for all $m \geq 1$ and $K > 0$, $\mathbb{P}(\mathcal{E}_m | \mathcal{E}_{m-1}) \geq \delta_1$. Then, for $i = 1, 2$, for all $p > 1$ and for every $\varepsilon \in (0, \rho - 1/2)$, there exists $C > 0$ such that for all $j \geq 1$, $m \in \{0, \dots, j-1\}$ and $K > 0$,*

$$\mathbb{E}\left[\sup_{u \in [\tau_{j-1}+1, \tau_j]} |\Lambda_m^i(u)| \mid \mathcal{E}_j\right] \leq C \frac{\left(\sum_{l=m+1}^{j-1} \varsigma^l\right)^{1/2-\rho+\varepsilon}}{\delta_1^{\frac{j-m}{p}}}. \quad (2.G.4)$$

Consequently, there exist $\eta \in (0, 1)$ and $C > 0$ such that for all $j \geq 1$ and $m \in \{0, \dots, j-1\}$,

$$\mathbb{E}\left[\sup_{u \in [\tau_{j-1}+1, \tau_j]} |\Lambda_m^i(u)| \mid \mathcal{E}_j\right] \leq C\eta^{j-m}. \quad (2.G.5)$$

Proof. First of all, let us prove that (2.G.4) induces (2.G.5). Let $\alpha_1 \in (0, +\infty)$ such that $\varsigma = \delta_1^{-\alpha_1}$. One just have to remark that for $j \geq 2$ and $m \in \{1, \dots, j-2\}$,

$$\left(\sum_{l=m+1}^{j-1} \varsigma^l\right)^{1/2-\rho+\varepsilon} \delta_1^{\frac{m-j}{p}} \leq \delta_1^{(\alpha_1(\rho-1/2-\varepsilon)-1/p)(j-m)}$$

We choose for instance $\varepsilon = \frac{1}{2}(\rho - 1/2)$ and $p > \frac{2}{\alpha_1(\rho-1/2)}$ in such a way that

$$\alpha_1(\rho - 1/2 - \varepsilon) - 1/p > 0.$$

We then deduce (2.G.5).

Now, it remains to show (2.G.4). For clarity, we set

$$E_j := [\tau_{j-1} + 1, \tau_j].$$

Using that for $m \geq 1$, $\mathcal{E}_m \subset \mathcal{E}_{m-1}$ and $\mathbb{P}(\mathcal{E}_m | \mathcal{E}_{m-1}) \geq \delta_1 \in (0, 1)$ and Hölder inequality we deduce the following inequalities,

$$\begin{aligned} \mathbb{E}\left[\sup_{u \in E_j} |\Lambda_m^i(u)| \mid \mathcal{E}_j\right] &\leq \mathbb{E}\left[\sup_{u \in E_j} |\Lambda_m^i(u)|^p \mid \mathcal{E}_j\right]^{1/p} = \left(\mathbb{E}\left[\sup_{u \in E_j} |\Lambda_m^i(u)|^p \mathbf{1}_{\mathcal{E}_j} \frac{1}{\mathbb{P}(\mathcal{E}_j)}\right]\right)^{1/p} \\ &\leq \left(\mathbb{E}\left[\sup_{u \in E_j} |\Lambda_m^i(u)|^p \mathbf{1}_{\mathcal{E}_{j-1}} \frac{1}{\mathbb{P}(\mathcal{E}_j)}\right]\right)^{1/p} = \left(\mathbb{E}\left[\sup_{u \in E_j} |\Lambda_m^i(u)|^p \mid \mathcal{E}_{j-1}\right] \frac{\mathbb{P}(\mathcal{E}_{j-1})}{\mathbb{P}(\mathcal{E}_j)}\right)^{1/p} \\ &\leq \delta_1^{-1/p} \mathbb{E}\left[\sup_{u \in E_j} |\Lambda_m^i(u)|^p \mid \mathcal{E}_{j-1}\right]^{1/p} \\ &\leq (\delta_1^{-1})^{\frac{j-m}{p}} \mathbb{E}\left[\sup_{u \in E_j} |\Lambda_m^i(u)|^p \mid \mathcal{E}_m\right]^{1/p} \quad (\text{by induction}). \end{aligned}$$

It remains to prove the existence of C such that for all $j \geq 1$, $m \in \{0, \dots, j-1\}$ and

$K > 0$,

$$\mathbb{E}[\sup_{u \in E_j} |\Lambda_m^i(u)|^p | \mathcal{E}_m]^{1/p} \leq C \left(\sum_{l=m+1}^{j-1} \varsigma^l \right)^{1/2-\rho+\varepsilon}$$

with again the convention $\sum_{\emptyset} = 1$. We separate the end of the proof into three cases.

Case 1: $j \geq 3$ and $m \in \{1, \dots, j-2\}$.

By Lemma 2.G.1, applied with $t_0 = \tau_{m-1} + 1$ and $t_1 = \tau_m$

$$|\Lambda_m^i(u)| \leq C_\rho (u - \tau_{m-1})^{-\rho} \left| \sum_{k=\tau_{m-1}+1}^{\tau_m} \xi_k^i \right| + C_\kappa \sum_{k=1}^{\tau_m - \tau_{m-1} - 1} \left| \sum_{l=k+\tau_{m-1}+1}^{\tau_m} \xi_l^i \right| (u - \tau_{m-1} - k)^{-(\rho+1)}.$$

But, $u - \tau_{m-1} \geq \tau_{j-1} - \tau_{m-1} \geq \sum_{l=m+1}^{j-1} \varsigma^l$ and $u - \tau_{m-1} - k \geq \tau_{j-1} - \tau_m + 1 \geq \sum_{l=m+1}^{j-1} \varsigma^l$. Let $\varepsilon \in (0, \rho - 1/2)$, we then have

$$|\Lambda_m^i(u)| \leq \left(\sum_{l=m+1}^{j-1} \varsigma^l \right)^{1/2-\rho+\varepsilon} \left[C_\rho (u - \tau_{m-1})^{-(1/2+\varepsilon)} \left| \sum_{k=\tau_{m-1}+1}^{\tau_m} \xi_k^i \right| + C_\kappa \sum_{k=1}^{\tau_m - \tau_{m-1} - 1} \left| \sum_{l=k+\tau_{m-1}+1}^{\tau_m} \xi_l^i \right| (u - \tau_{m-1} - k)^{-(3/2+\varepsilon)} \right].$$

We denote by $\tilde{\Lambda}_m^i(u)$ the above quantity between brackets. Hence

$$\mathbb{E}[\sup_{u \in E_j} |\Lambda_m^i(u)|^p | \mathcal{E}_m]^{1/p} \leq \left(\sum_{l=m+1}^{j-1} \varsigma^l \right)^{1/2-\rho+\varepsilon} \mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{E}_m]^{1/p}.$$

We now have to prove the existence of C such that

$$\mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{E}_m]^{1/p} \leq C \quad \text{for all } p \in (1, +\infty).$$

We write $\mathcal{E}_m = \cup_{\ell \geq 0} \mathcal{A}_{m,\ell}$ with

$$\mathcal{A}_{m,\ell} = \mathcal{B}_{m,\ell}^c \cap \mathcal{B}_{m,\ell-1} \tag{2.G.6}$$

where $\mathcal{B}_{m,\ell}$ is defined in Equation (2.5.18). In other words, $\mathcal{A}_{m,0}$ is the failure of Step 1 of tentative m and for $\ell \geq 1$, $\mathcal{A}_{m,\ell}$ is the event ‘‘Step 2 of trial m fails after exactly ℓ attempts’’.

Let $\ell \in \mathbb{N}$, we begin by studying $\mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{A}_{m,\ell}]^{1/p}$. Since $u > \tau_m$,

$$|\tilde{\Lambda}_m^i(u)| \leq C_\rho (\Delta\tau_m)^{-(1/2+\varepsilon)} \left| \sum_{k=\tau_{m-1}+1}^{\tau_{m-1}+\Delta\tau_m} \xi_k^i \right| + C_\kappa \sum_{k=1}^{\Delta\tau_m-1} \left| \sum_{l=k+\tau_{m-1}+1}^{\tau_{m-1}+\Delta\tau_m} \xi_l^i \right| (\Delta\tau_m - k)^{-(3/2+\varepsilon)}.$$

By Minkowski inequality and the fact that $\Delta\tau_m =: \Delta(m, \ell)$ is constant on $\mathcal{A}_{m, \ell}$, we get

$$\begin{aligned} \mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{A}_{m, \ell}]^{1/p} &\leq C_\rho (\Delta(m, \ell))^{-(1/2+\varepsilon)} \mathbb{E} \left[\left| \sum_{k=\tau_{m-1}+1}^{\tau_{m-1}+\Delta(m, \ell)} \xi_k^i \right|^p | \mathcal{A}_{m, \ell} \right]^{1/p} \\ &\quad + C_\kappa \sum_{k=1}^{\Delta(m, \ell)-1} (\Delta(m, \ell) - k)^{-(3/2+\varepsilon)} \mathbb{E} \left[\left| \sum_{l=k+\tau_{m-1}+1}^{\tau_{m-1}+\Delta(m, \ell)} \xi_l^i \right|^p | \mathcal{A}_{m, \ell} \right]^{1/p}. \end{aligned} \quad (2.G.7)$$

Moreover, using Cauchy-Schwarz inequality,

$$\begin{aligned} \mathbb{E} \left[\left| \sum_{k=\tau_{m-1}+1}^{\tau_{m-1}+\Delta(m, \ell)} \xi_k^i \right|^p | \mathcal{A}_{m, \ell} \right]^{1/p} &= \mathbb{E} \left[\left| \sum_{k=\tau_{m-1}+1}^{\tau_{m-1}+\Delta(m, \ell)} \xi_k^i \right|^p \mathbb{1}_{\mathcal{A}_{m, \ell}} | \mathcal{E}_{m-1} \right]^{1/p} \mathbb{P}(\mathcal{A}_{m, \ell} | \mathcal{E}_{m-1})^{-1/p} \\ &\leq \mathbb{E} \left[\left(\sum_{k=\tau_{m-1}+1}^{\tau_{m-1}+\Delta(m, \ell)} \xi_k^i \right)^{2p} | \mathcal{E}_{m-1} \right]^{1/2p} \mathbb{P}(\mathcal{A}_{m, \ell} | \mathcal{E}_{m-1})^{-1/2p} \\ &\leq c_p \sqrt{\Delta(m, \ell)} \mathbb{P}(\mathcal{A}_{m, \ell} | \mathcal{E}_{m-1})^{-1/2p}. \end{aligned} \quad (2.G.8)$$

In the last inequality we use the fact that $\sum_{k=\tau_{m-1}+1}^{\tau_{m-1}+\Delta(m, \ell)} \xi_k^i$ is independent from \mathcal{E}_{m-1} and that its law is $\mathcal{N}(0, \Delta(m, \ell))$. In the same way, we obtain

$$\mathbb{E} \left[\left| \sum_{l=k+\tau_{m-1}+1}^{\tau_{m-1}+\Delta(m, \ell)} \xi_l^i \right|^p | \mathcal{A}_{m, \ell} \right]^{1/p} \leq c_p \sqrt{\Delta(m, \ell) - k} \mathbb{P}(\mathcal{A}_{m, \ell} | \mathcal{E}_{m-1})^{-1/2p}. \quad (2.G.9)$$

We deduce from (2.G.8) and (2.G.9) that in (2.G.7)

$$\begin{aligned} \mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{A}_{m, \ell}]^{1/p} &\leq \left(c_p C_\rho (\Delta(m, \ell))^{-\varepsilon} + c_p C_\kappa \sum_{k=1}^{\Delta(m, \ell)-1} (\Delta(m, \ell) - k)^{-(1+\varepsilon)} \right) \mathbb{P}(\mathcal{A}_{m, \ell} | \mathcal{E}_{m-1})^{-1/2p} \\ &\leq C_{p, \varepsilon} \mathbb{P}(\mathcal{A}_{m, \ell} | \mathcal{E}_{m-1})^{-1/2p}. \end{aligned} \quad (2.G.10)$$

Then by using the inequality $(a + b)^{1/p} \leq a^{1/p} + b^{1/p}$ for $p > 1$ and (2.G.10) for ℓ from

0 to $+\infty$, we get

$$\begin{aligned}
 \mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{E}_m]^{1/p} &= \left(\sum_{\ell \geq 0} \mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{A}_{m,\ell}] \mathbb{P}(\mathcal{A}_{m,\ell} | \mathcal{E}_m) \right)^{1/p} \\
 &\leq \sum_{\ell \geq 0} \mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{A}_{m,\ell}]^{1/p} \mathbb{P}(\mathcal{A}_{m,\ell} | \mathcal{E}_m)^{1/p} \\
 &\leq C_{p,\varepsilon} \sum_{\ell \geq 0} \mathbb{P}(\mathcal{A}_{m,\ell} | \mathcal{E}_{m-1})^{-1/2p} \left(\frac{\mathbb{P}(\mathcal{A}_{m,\ell} | \mathcal{E}_{m-1})}{\mathbb{P}(\mathcal{E}_m | \mathcal{E}_{m-1})} \right)^{1/p} \\
 &\leq C_{p,\varepsilon} \delta_1^{-1/p} \sum_{\ell \geq 0} \mathbb{P}(\mathcal{A}_{m,\ell} | \mathcal{E}_{m-1})^{1/2p}.
 \end{aligned}$$

But $\mathcal{A}_{m,\ell} \subset \mathcal{E}_m \subset \mathcal{E}_{m-1}$ hence for $\ell > 0$,

$$\mathbb{P}(\mathcal{A}_{m,\ell} | \mathcal{E}_{m-1}) = \mathbb{P}(\mathcal{B}_{m,\ell-1} | \mathcal{E}_{m-1}) \mathbb{P}(\mathcal{B}_{m,\ell}^c | \mathcal{B}_{m,\ell-1}) \leq \mathbb{P}(\mathcal{B}_{m,\ell}^c | \mathcal{B}_{m,\ell-1}),$$

Therefore, for all $\varepsilon \in (0, \rho - 1/2)$ and $p \in (1, +\infty)$, by applying Lemma 2.5.2, this gives us the existence of C such that

$$\mathbb{E}[\sup_{u \in E_j} |\tilde{\Lambda}_m^i(u)|^p | \mathcal{E}_m]^{1/p} \leq C_{p,\varepsilon} \delta_1^{-1/p} (2 + \sum_{\ell \geq 2} 2^{-\tilde{\alpha}\ell/2p}) < C.$$

The first case is now achieved.

Case 2: Let $j \geq 2$ and $m = j - 1$.

The proof is almost exactly the same as in case 1. We simply use the following controls

$$u - \tau_{j-2} > \Delta\tau_{j-1} \quad \text{and} \quad u - \tau_{j-2} - k > \Delta\tau_{j-1} - k$$

and we do not introduce ε which is useless here since $\sum_{l=m+1}^{j-1} = \sum_{\emptyset}$.

Case 3: Let $j \geq 1$ and $m = 0$. By assumption $\tau_0 = 0$, then

$$\Lambda_0^i(u) = \sum_{k=-\infty}^{k=0} a_{u-k} \xi_k^i.$$

By Lemma 2.G.1, for all $M > 0$,

$$\begin{aligned}
 \left| \sum_{-M}^{k=0} a_{u-k} \xi_k^i \right| &\leq C_\rho (u+1+M)^{-\rho} \left| \sum_{k=-M}^0 \xi_k^i \right| + C_\kappa \sum_{k=1}^M \left| \sum_{l=k-M}^0 \xi_l^i \right| (u+1+M-k)^{-(\rho+1)} \\
 &\leq C_\rho (1+M)^{-\rho} \left| \sum_{k=-M}^0 \xi_k^i \right| + C_\kappa \sum_{k=1}^M \left| \sum_{l=k-M}^0 \xi_l^i \right| (M+1-k)^{-(\rho+1)} \\
 &= C_\rho (1+M)^{-\rho} \left| \sum_{k=-M}^0 \xi_k^i \right| + C_\kappa \sum_{k=0}^{M-1} \left| \sum_{l=-k}^0 \xi_l^i \right| (k+1)^{-(\rho+1)} \\
 &= C_\rho (1+M)^{-\rho} \left| \sum_{k=0}^M \xi_{-k}^i \right| + C_\kappa \sum_{k=0}^{M-1} \left| \sum_{l=0}^k \xi_{-l}^i \right| (k+1)^{-(\rho+1)}.
 \end{aligned}$$

Since $\rho > 1/2$, by means of Borel-Cantelli Lemma and the fact that $\sum_{k=0}^M \xi_{-k}^i \sim \mathcal{N}(0, M+1)$, we can show that $\lim_{M \rightarrow +\infty} (1+M)^{-\rho} \left(\sum_{k=0}^M \xi_{-k}^i \right) = 0$ a.s.

We then get

$$|\Lambda_0^i(u)| \leq C_\kappa \sum_{k=0}^{+\infty} \left| \sum_{l=0}^k \xi_{-l}^i \right| (k+1)^{-(\rho+1)}.$$

Set $W_k^i = \sum_{l=0}^{k-1} \xi_{-l}^i$ for $k > 0$ and $W_0^i = \xi_0^i$. Using Minkowski inequality, we have for all $p \in (1, +\infty)$ and for all $\varepsilon \in (0, \rho - 1/2)$

$$\begin{aligned}
 \mathbb{E}[\sup_{u \in E_j} |\Lambda_0^i(u)|^p | \mathcal{E}_0]^{1/p} &= \mathbb{E}[\sup_{u \in E_j} |\Lambda_0^i(u)|^p]^{1/p} \\
 &\leq C_\kappa \sum_{k=0}^{+\infty} (k+1)^{-(\rho+1/2-\varepsilon)} \mathbb{E} \left[\left(\frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \right)^p \right]^{1/p} \\
 &\leq C_\kappa \mathbb{E} \left[\left(\sup_{k \geq 0} \frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \right)^p \right]^{1/p} \sum_{k=0}^{+\infty} (k+1)^{-(\rho+1/2-\varepsilon)} \\
 &\leq C \mathbb{E} \left[\left(\sup_{k \geq 0} \frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \right)^p \right]^{1/p}
 \end{aligned}$$

because $\rho + 1/2 - \varepsilon > 1$. It remains to prove that for $\varepsilon \in (0, \rho - 1/2)$ and for $p \in (1, +\infty)$

$$\mathbb{E} \left[\left(\sup_{k \geq 0} \frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \right)^p \right]^{1/p} < +\infty.$$

Thanks to a summation by parts, we can show that

$$\frac{W_{k+1}^i}{(k+1)^{1/2+\varepsilon}} = \sum_{l=0}^k \frac{\xi_{-l}^i}{(l+1)^{1/2+\varepsilon}} + \sum_{l=1}^k W_l^i \left(\frac{1}{(l+1)^{1/2+\varepsilon}} - \frac{1}{l^{1/2+\varepsilon}} \right).$$

Hence, using that $\left| \frac{1}{(l+1)^{1/2+\varepsilon}} - \frac{1}{l^{1/2+\varepsilon}} \right| \leq (1/2 + \varepsilon) \frac{1}{l^{3/2+\varepsilon}}$, we get

$$\frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \leq \left| \sum_{l=0}^k \frac{\xi_{-l}^i}{(l+1)^{1/2+\varepsilon}} \right| + (1/2 + \varepsilon) \sum_{l=1}^k \frac{|W_l^i|}{l^{3/2+\varepsilon}}.$$

Therefore

$$\sup_{k \geq 0} \frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \leq \sup_{k \geq 0} \left| \sum_{l=0}^k \frac{\xi_{-l}^i}{(l+1)^{1/2+\varepsilon}} \right| + (1/2 + \varepsilon) \sum_{l=1}^{+\infty} \frac{|W_l^i|}{l^{3/2+\varepsilon}}.$$

We use again Minkowski inequality, which gives

$$\left\| \sup_{k \geq 0} \frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \right\|_p \leq \left\| \sup_{k \geq 0} \left| \sum_{l=0}^k \frac{\xi_{-l}^i}{(l+1)^{1/2+\varepsilon}} \right| \right\|_p + (1/2 + \varepsilon) \sum_{l=1}^{+\infty} \frac{\|W_l^i\|_p}{l^{3/2+\varepsilon}}.$$

On the one hand, we have $\|W_l^i\|_p \leq c_p \sqrt{l}$ because $W_l^i \sim \mathcal{N}(0, l)$. On the other hand, set $N_k := \sum_{l=0}^k \frac{\xi_{-l}^i}{(l+1)^{1/2+\varepsilon}}$. This is a martingale with distribution $\mathcal{N}\left(0, \sum_{l=0}^k (l+1)^{-(1+2\varepsilon)}\right)$ therefore $\|N_k\|_p \leq c_p$ with c_p independent from k . Hence, $(N_k)_{k \in \mathbb{N}}$ converges a.s. and in L^p into $N_\infty \in L^p$. We then deduce by Doob's inequality that

$$\left\| \sup_{k \geq 0} |N_k| \right\|_p \leq \left(\frac{p-1}{p} \right) \|N_\infty\|_p.$$

Finalement,

$$\left\| \sup_{k \geq 0} \frac{|W_{k+1}^i|}{(k+1)^{1/2+\varepsilon}} \right\|_p < +\infty.$$

which achieves the third case. \square

Proposition 2.G.1. *Assume $(\mathbf{H}_{\text{poly}})$. We suppose that $\tau_0 = 0$ and that for all $m \geq 1$ and $K > 0$, $\mathbb{P}(\mathcal{E}_m | \mathcal{E}_{m-1}) \geq \delta_1 \in (0, 1)$. Then (\mathbf{H}'_1) holds true.*

Proof. First, thanks to (2.G.3), we have

$$\sum_{l=1}^{\Delta \tau_j} \gamma^{\Delta \tau_j - l} |\Delta_{\tau_{j-1}+l}^i| \leq \sum_{m=0}^j \sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j - u} |\Lambda_m^i(u)|.$$

The aim is to bound every term in the right-hand side. For $m \in \{0, \dots, j-1\}$ and for all $u \in E_j := \llbracket \tau_{j-1} + 1, \tau_j \rrbracket$,

$$|\Lambda_m^i(u)| \leq \sup_{u \in E_j} |\Lambda_m^i(u)|.$$

Since the right-hand side does not depend on u anymore, we deduce that for all $m \in \{0, \dots, j-1\}$

$$\sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j - u} |\Lambda_m^i(u)| \leq \sup_{u \in E_j} |\Lambda_m^i(u)| \sum_{w=0}^{+\infty} \gamma^w = \sup_{u \in E_j} |\Lambda_m^i(u)| \frac{1}{1-\gamma}.$$

Hence, Lemma 2.G.2, gives that for all $m \in \{0, \dots, j-1\}$

$$\mathbb{E} \left[\sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j-u} |\Lambda_m^i(u)| \mid \mathcal{E}_j \right] \leq \frac{C}{1-\gamma} \eta^{j-m}$$

where $\eta \in (0, 1)$. Consequently,

$$\mathbb{E} \left[\sum_{m=0}^{j-1} \sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j-u} |\Lambda_m^i(u)| \mid \mathcal{E}_j \right] \leq \frac{C}{(1-\gamma)(1-\eta)}. \quad (2.G.11)$$

In inequality (2.G.3), it then remains to bound the term with $\Lambda_j^i(u)$. By substitution, we obtain for $m = j$

$$\sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j-u} |\Lambda_j^i(u)| = \sum_{v=1}^{\Delta\tau_j} \gamma^{\Delta\tau_j-v} |\Lambda_j^i(v + \tau_{j-1})|.$$

As in the proof of Lemma 2.G.2, we use the decomposition of \mathcal{E}_j through the events $\mathcal{A}_{j,\ell}$ and that $\Delta\tau_j =: \Delta(j, \ell)$ is constant on $\mathcal{A}_{j,\ell}$:

$$\mathbb{E} \left[\sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j-u} |\Lambda_j^i(u)| \mid \mathcal{E}_j \right] = \sum_{\ell \geq 0} \sum_{v=1}^{\Delta(j,\ell)} \gamma^{\Delta(j,\ell)-v} \mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})| \mid \mathcal{A}_{j,\ell}] \mathbb{P}(\mathcal{A}_{j,\ell} \mid \mathcal{E}_j). \quad (2.G.12)$$

Using that $\mathcal{A}_{j,\ell} \subset \mathcal{E}_j \subset \mathcal{E}_{j-1}$ and Cauchy-Schwarz inequality, one notes that

$$\begin{aligned} \mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})| \mid \mathcal{A}_{j,\ell}] \mathbb{P}(\mathcal{A}_{j,\ell} \mid \mathcal{E}_j) &\leq \mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})|^2 \mid \mathcal{E}_j]^{1/2} \mathbb{P}(\mathcal{A}_{j,\ell} \mid \mathcal{E}_j)^{1/2} \\ &\leq \frac{\mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})|^4 \mid \mathcal{E}_{j-1}]^{1/4}}{\mathbb{P}(\mathcal{E}_j \mid \mathcal{E}_{j-1})^{1/4}} \mathbb{P}(\mathcal{A}_{j,\ell} \mid \mathcal{E}_j)^{1/2} \\ &\leq \frac{\sup_{v \in \mathbb{N}^*} \mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})|^4 \mid \mathcal{E}_{j-1}]^{1/4}}{\mathbb{P}(\mathcal{E}_j \mid \mathcal{E}_{j-1})^{1/4}} \mathbb{P}(\mathcal{A}_{j,\ell} \mid \mathcal{E}_j)^{1/2}. \end{aligned} \quad (2.G.13)$$

But $\mathbb{P}(\mathcal{E}_j \mid \mathcal{E}_{j-1}) \geq \delta_1 > 0$ and by Lemma 2.5.2, we have for all $\ell \geq 2$,

$$\mathbb{P}(\mathcal{A}_{j,\ell} \mid \mathcal{E}_j) = \frac{\mathbb{P}(\mathcal{A}_{j,\ell} \mid \mathcal{E}_{j-1})}{\mathbb{P}(\mathcal{E}_j \mid \mathcal{E}_{j-1})} \leq \delta_1^{-1} \mathbb{P}(\mathcal{B}_{j,\ell-1} \mid \mathcal{E}_{j-1}) \mathbb{P}(\mathcal{B}_{j,\ell}^c \mid \mathcal{B}_{j,\ell-1}) \leq \delta_1^{-1} 2^{-\tilde{\alpha}\ell}. \quad (2.G.14)$$

We now use (2.G.13) and (2.G.14) into (2.G.12) and this gives the existence of $C_{\delta_1, \gamma}$ such that

$$\mathbb{E} \left[\sum_{u=\tau_{j-1}+1}^{\tau_j} \gamma^{\tau_j-u} |\Lambda_j^i(u)| \mid \mathcal{E}_j \right] \leq C_{\delta_1, \gamma} \sup_{v \in \mathbb{N}^*} \mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})|^4 \mid \mathcal{E}_{j-1}]^{1/4}. \quad (2.G.15)$$

It only remains to show that

$$\sup_{v \in \mathbb{N}^*} \mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})|^4 | \mathcal{E}_{j-1}]^{1/4} < +\infty.$$

By Lemma 2.G.1 and the definition of Λ_j^i in (2.G.2),

$$\begin{aligned} |\Lambda_j^i(v + \tau_{j-1})| &= \left| \sum_{k=\tau_{j-1}+1}^{v+\tau_{j-1}} a_{v+\tau_{j-1}-k} \xi_k^i \right| \\ &\leq C_\rho v^{-\rho} \left| \sum_{k=\tau_{j-1}+1}^{v+\tau_{j-1}} \xi_k^i \right| + C_\kappa \sum_{k=1}^{v-1} \left| \sum_{l=k+\tau_{j-1}+1}^{v+\tau_{j-1}} \xi_l^i \right| (v-k)^{-(\rho+1)}. \end{aligned}$$

We again apply Minkowski inequality

$$\begin{aligned} &\mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})|^4 | \mathcal{E}_{j-1}]^{1/4} \\ &\leq C_\rho v^{-\rho} \mathbb{E} \left[\left| \sum_{k=\tau_{j-1}+1}^{v+\tau_{j-1}} \xi_k^i \right|^4 \middle| \mathcal{E}_{j-1} \right]^{1/4} + C_\kappa \sum_{k=1}^{v-1} (v-k)^{-(\rho+1)} \mathbb{E} \left[\left| \sum_{l=k+\tau_{j-1}+1}^{v+\tau_{j-1}} \xi_l^i \right|^4 \middle| \mathcal{E}_{j-1} \right]^{1/4} \\ &= C_\rho v^{-\rho} \mathbb{E} \left[\left| \sum_{k=1}^v \xi_k^i \right|^4 \right]^{1/4} + C_\kappa \sum_{k=1}^{v-1} (v-k)^{-(\rho+1)} \mathbb{E} \left[\left| \sum_{l=k+1}^v \xi_l^i \right|^4 \right]^{1/4} \\ &\leq c_4 \left(C_\rho v^{-\rho+1/2} + C_\kappa \sum_{k=1}^{v-1} (v-k)^{-(\rho+1/2)} \right) \end{aligned}$$

where c_4 is related to the 4th moment of a centered and reduced Gaussian random variable. Since $\rho + 1/2 > 1$, we immediately deduce that

$$\sup_{v \in \mathbb{N}^*} \mathbb{E}[|\Lambda_j^i(v + \tau_{j-1})|^4 | \mathcal{E}_{j-1}]^{1/4} < +\infty. \quad (2.G.16)$$

We put together (2.G.11), (2.G.15) and (2.G.16) to conclude the proof of (\mathbf{H}'_1) . \square

3. Concentration inequalities for Stochastic Differential Equations with additive fractional noise

This chapter is based on the article [70], submitted in *Electronic Journal of Probability*.

Contents

3.1	Introduction	91
3.2	Setting and main results	93
3.3	Sketch of proof	96
3.4	Control of the martingale increments	98
3.5	Conditional exponential moments of the martingale increments	105
3.A	Sub-Gaussianity of the supremum of the Browian motion	110
3.B	Uniform sub-Gaussianity of $\ G^{(k)}\ _{\alpha, [0,2]}$	111

3.1 Introduction

In this chapter, we consider the solution $(Y_t)_{t \geq 0}$ of the following \mathbb{R}^d -valued Stochastic Differential Equation (SDE) with additive noise:

$$Y_t = x + \int_0^t b(Y_s) ds + \sigma B_t. \tag{3.1.1}$$

with B a d -dimensional fractional Brownian motion (fBm) with Hurst parameter $H \in (0, 1)$. We are interested in questions of long-time concentration phenomenon of the law of the solution Y . A well known way to overcome this type of problem is to prove L^1 -transportation inequalities. Let us precise what it means. Let (E, d) be a metric space equipped with a σ -field \mathcal{B} such that the distance d is $\mathcal{B} \otimes \mathcal{B}$ -measurable. Given $p \geq 1$ and two probability measures μ and ν on E , the Wasserstein distance is defined by

$$\mathcal{W}_p(\mu, \nu) = \inf \left(\int \int d(x, y)^p d\pi(x, y) \right),$$

where the infimum runs over all the probability measures π on $E \times E$ with marginals μ and ν . The entropy of ν with respect to μ is defined by

$$\mathbf{H}(\nu|\mu) = \begin{cases} \int \log\left(\frac{d\nu}{d\mu}\right) d\nu, & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases}$$

Then, we say that μ satisfies an L^p -transportation inequality with constant $C \geq 0$ (noted $\mu \in T_p(C)$) if for any probability measure ν ,

$$\mathcal{W}_p(\mu, \nu) \leq \sqrt{2C\mathbf{H}(\nu|\mu)}. \quad (3.1.2)$$

The concentration of measure is intrinsically linked to the above inequality when $p = 1$. This fact was first emphasized by K.Marton [52, 51], M.Talagrand [66], Bobkov and Götze [5] and amply investigated by M.Ledoux [48, 47]. Indeed, it can be shown (see [48] for a detailed proof) that (3.1.2) for $p = 1$ is actually equivalent to the following: for any μ -integrable α -Lipschitz function F (real valued) we have for all $\lambda \in \mathbb{R}$,

$$\mathbb{E}(\exp(\lambda(F(X) - \mathbb{E}[F(X)]))) \leq \exp\left(C\alpha^2 \frac{\lambda^2}{2}\right) \quad (3.1.3)$$

with $\mathcal{L}(X) = \mu$. This upperbound naturally leads to concentration inequalities through the classical Markov inequality. For several years, L^1 (and L^2 since $T_2(C)$ implies $T_1(C)$) transportation inequalities have then been widely studied and in particular for diffusion processes (see for instance [23, 74, 31]).

For SDE's driven by more general Gaussian processes, S.Riedel established transportation cost inequalities in [62] using Rough Path theory. However, his results do not give long-time concentration, which is our focus here.

In the setting of fractional noise, T.Guendouzi [33] and B.Saussereau [65] have studied transportation inequalities with different metrics in the case where $H \in (1/2, 1)$. In particular, B.Saussereau gave an important contribution: he proved $T_1(C)$ and $T_2(C)$ for the law of $(Y_t)_{t \in [0, T]}$ in various settings and he got a result of large-time asymptotics in the case of a contractive drift. Our first motivation to this work was to get equivalent results in a discrete-time context, i.e. for $\mathcal{L}((Y_{k\Delta})_{1 \leq k \leq n})$ for a given step $\Delta > 0$ and then long-time concentration inequalities for the occupation measure, i.e. for $\frac{1}{n} \sum_{k=1}^n f(Y_{k\Delta})$ (where f is a general Lipschitz function real valued). Indeed, in a statistical framework we only have access to discrete-time observations of the process Y and such a result could be meaningful in such context. To the best of our knowledge, this type of result is unknown in the fractional setting.

We first tried to adapt the methods used in [65] in several ways as for example: find a distance such that $(y_t)_{t \in [0, T]} \mapsto (y_{k\Delta})_{1 \leq k \leq n}$ is Lipschitz and prove $T_1(C)$ with this metric. But the constants obtained in the L^1 -transportation inequalities were not sharp enough, so that we couldn't deduce large-time asymptotic as B.Saussereau.

In [23], H.Djellout, A.Guillin and L.Wu explored transportation inequalities in the diffusive case and both in a continuous and discrete-time setting. In particular, for the discrete-time case, they used a kind of tensorization of the L^1 transportation inequality but the Markovian nature of the process was essential. However, they prove $T_1(C)$

through its equivalent formulation (3.1.3) and to this end, they apply a decomposition of the functional in (3.1.3) into a sum of martingale increments, namely:

$$F(X) - \mathbb{E}[F(X)] = \sum_{k=1}^n \mathbb{E}[F(X)|\mathcal{F}_k] - \mathbb{E}[F(X)|\mathcal{F}_{k-1}]$$

with $X = (Y_{k\Delta})_{1 \leq k \leq n}$ and Y is the solution of (3.1.1) when B is the classical Brownian motion.

This decomposition has inspired the approach described in this chapter: instead of proving an L^1 transportation inequality (3.1.2), we prove its equivalent formulation (3.1.3) by using a similar decomposition and the series expansion of the exponential function. Through this strategy, we prove several results under an assumption of contractivity on the drift term b in (3.1.1). First, in a discrete-time setting, we work in the space $(\mathbb{R}^d)^n$ endowed with the L^1 metric and we show that for any α -Lipschitz functional $F : (\mathbb{R}^d)^n \rightarrow \mathbb{R}$ and for any $\lambda > 0$,

$$\mathbb{E}(\exp(\lambda(F(X) - \mathbb{E}[F(X)]))) \leq \exp(C\alpha^2\lambda^2n^{2H\vee 1})$$

with $X = (Y_{k\Delta})_{1 \leq k \leq n}$. In a similar way, we consider the space of continuous functions $\mathcal{C}([0, T], \mathbb{R}^d)$ endowed with the L^1 metric and we prove that for any α -Lipschitz functional $\tilde{F} : \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$ and for any $\lambda > 0$,

$$\mathbb{E}(\exp(\lambda(\tilde{F}(X) - \mathbb{E}[\tilde{F}(X)]))) \leq \exp(C\alpha^2\lambda^2T^{2H\vee 1})$$

with $X = (Y_t)_{t \in [0, T]}$. From these inequalities, we deduce some general concentration inequalities and large-time asymptotics for occupation measures. Let us note that we have no restriction on the Hurst parameter H and we retrieve the results given by B.Saussereau for $H \in (1/2, 1)$ in a continuous setting and also the result given in [23] for $H = 1/2$, namely for diffusion.

The chapter is organised as follows. In the next section, we describe the assumptions on the drift term and we state the general theorem about concentration, namely Theorem 3.2.2. Then, in Subsection 3.2.3, we apply this result to specific functionals related to the occupation measures (both in a discrete-time and in a continuous-time framework). Section 3.3 outlines our strategy of proof which is fulfilled in Sections 3.4 and 3.5.

3.2 Setting and main results

3.2.1 Notations

The usual scalar product on \mathbb{R}^d is denoted by $\langle \cdot, \cdot \rangle$ and $|\cdot|$ stands either for the Euclidean norm on \mathbb{R}^d or the absolute value on \mathbb{R} . We denote by $\mathcal{M}_d(\mathbb{R})$ the space of real matrices of size $d \times d$. For a given $n \in \mathbb{N}^*$ and $(x, y) \in (\mathbb{R}^d)^n \times (\mathbb{R}^d)^n$, we denote by d_n the following L^1 -distance:

$$d_n(x, y) := \sum_{k=1}^n |x_k - y_k|. \tag{3.2.1}$$

Analogously, for a given $T > 0$ and $(x, y) \in \mathcal{C}([0, T], \mathbb{R}^d) \times \mathcal{C}([0, T], \mathbb{R}^d)$, we denote by d_T the classical L^1 -distance:

$$d_T(x, y) := \int_0^T |x_t - y_t| dt. \quad (3.2.2)$$

Let $F : (E, d_E) \rightarrow (E', d_{E'})$ be a Lipschitz function between two metric spaces, we denote by

$$\|F\|_{\text{Lip}} := \sup_{x \neq y} \frac{d_{E'}(F(x), F(y))}{d_E(x, y)}$$

its Lipschitz norm.

3.2.2 Assumptions and general result

Let B be a d -dimensional fractional Brownian motion (fBm) with Hurst parameter $H \in (0, 1)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and transferred from a d -dimensional Brownian motion W through the Volterra representation (see e.g. [20, 12])

$$\forall t \in \mathbb{R}_+, \quad B_t = \int_0^t K_H(t, s) dW_s, \quad (3.2.3)$$

with

$$K_H(t, s) := c_H \left[\frac{t^{H-\frac{1}{2}}}{s^{H-\frac{1}{2}}} (t-s)^{H-\frac{1}{2}} - \left(H - \frac{1}{2} \right) \int_s^t \frac{u^{H-\frac{3}{2}}}{s^{H-\frac{1}{2}}} (u-s)^{H-\frac{1}{2}} du \right]. \quad (3.2.4)$$

We consider the following \mathbb{R}^d -valued stochastic differential equation driven by B :

$$Y_t = x + \int_0^t b(Y_s) ds + \sigma B_t, \quad t \geq 0. \quad (3.2.5)$$

Here $x \in \mathbb{R}^d$ is a given initial condition, B is the aforementioned fractional Brownian motion and $\sigma \in \mathcal{M}_d(\mathbb{R})$.

We are working under the following assumption :

Hypothesis 3.2.1. *We have $b \in \mathcal{C}(\mathbb{R}^d; \mathbb{R}^d)$ and there exist constants $\alpha, L > 0$ such that:*

(i) *For every $x, y \in \mathbb{R}^d$,*

$$\langle b(x) - b(y), x - y \rangle \leq -\alpha |x - y|^2.$$

(ii) *For every $x, y \in \mathbb{R}^d$,*

$$|b(x) - b(y)| \leq L |x - y|.$$

Remark 3.2.1. This contractivity assumption on the drift term is quite usual to get long-time concentration results (see [23, 65] for instance). At this stage, a more general framework seems elusive.

Let $T > 0$ and $n \in \mathbb{N}^*$. Let $F : ((\mathbb{R}^d)^n, d_n) \rightarrow (\mathbb{R}, |\cdot|)$ and $\tilde{F} : (\mathcal{C}([0, T], \mathbb{R}^d), d_T) \rightarrow$

$(\mathbb{R}, |\cdot|)$ be two Lipschitz functions and set

$$F_Y := F(Y_{t_1}, \dots, Y_{t_n}) \quad \text{and} \quad \tilde{F}_Y = \tilde{F}((Y_t)_{t \in [0, T]}) \quad (3.2.6)$$

with $0 < \Delta = t_1 < \dots < t_n$ and $t_{k+1} - t_k = \Delta$ for a given $\Delta > 0$.

We are now in position to state our results for general functionals F and \tilde{F} . First, we prove a result on the exponential moments of F_Y and \tilde{F}_Y which is crucial to get Theorem 3.2.2.

Proposition 3.2.1. *Let $H \in (0, 1)$ and $\Delta > 0$. Let $n \in \mathbb{N}^*$, $T \geq 1$ and d_n, d_T be the metrics defined respectively by (3.2.1) and (3.2.2). Then,*

- (i) *there exist $C_{H, \Delta} > 0$ such that for all Lipschitz function $F : ((\mathbb{R}^d)^n, d_n) \rightarrow (\mathbb{R}, |\cdot|)$ and for all $\lambda > 0$,*

$$\mathbb{E}[\exp(\lambda(F_Y - \mathbb{E}[F_Y]))] \leq \exp(C_{H, \Delta} \|F\|_{\text{Lip}}^2 \lambda^2 n^{2H \vee 1}). \quad (3.2.7)$$

- (ii) *there exist $\tilde{C}_H > 0$ such that for all Lipschitz function $\tilde{F} : (\mathcal{C}([0, T], \mathbb{R}^d), d_T) \rightarrow (\mathbb{R}, |\cdot|)$ and for all $\lambda > 0$,*

$$\mathbb{E}[\exp(\lambda(\tilde{F}_Y - \mathbb{E}[\tilde{F}_Y]))] \leq \exp(\tilde{C}_H \|\tilde{F}\|_{\text{Lip}}^2 \lambda^2 T^{2H \vee 1}). \quad (3.2.8)$$

Remark 3.2.2. Let us note that this proposition is actually equivalent to L^1 -transportation inequalities as mentioned in the introduction. More precisely, item (i) is equivalent to $\mathcal{L}((Y_{t_k})_{1 \leq k \leq n}) \in T_1(2C_{H, \Delta} n^{2H \vee 1})$ for the metric d_n and item (ii) is equivalent to $\mathcal{L}((Y_t)_{t \in [0, T]}) \in T_1(2\tilde{C}_H T^{2H \vee 1})$ for the metric d_T .

From Proposition 3.2.1, we deduce the following concentration inequalities:

Theorem 3.2.2. *Let $H \in (0, 1)$ and $\Delta > 0$. Let $n \in \mathbb{N}^*$, $T \geq 1$ and d_n, d_T be the metrics defined respectively by (3.2.1) and (3.2.2). Then,*

- (i) *there exist $C_{H, \Delta} > 0$ such that for all Lipschitz function $F : ((\mathbb{R}^d)^n, d_n) \rightarrow (\mathbb{R}, |\cdot|)$ and for all $r \geq 0$,*

$$\mathbb{P}(F_Y - \mathbb{E}[F_Y] > r) \leq \exp\left(-\frac{r^2}{4C_{H, \Delta} \|F\|_{\text{Lip}}^2 n^{2H \vee 1}}\right). \quad (3.2.9)$$

- (ii) *there exist $\tilde{C}_H > 0$ such that for all Lipschitz function $\tilde{F} : (\mathcal{C}([0, T], \mathbb{R}^d), d_T) \rightarrow (\mathbb{R}, |\cdot|)$ and for all $r \geq 0$,*

$$\mathbb{P}(\tilde{F}_Y - \mathbb{E}[\tilde{F}_Y] > r) \leq \exp\left(-\frac{r^2}{4\tilde{C}_H \|\tilde{F}\|_{\text{Lip}}^2 T^{2H \vee 1}}\right). \quad (3.2.10)$$

Proof. We use Markov inequality and Proposition 3.2.1. Then, we optimize in λ to get the result. \square

Remark 3.2.3. The dependency on the Lipschitz constant of F and \tilde{F} is essential since they may depend on n and T . Hence, if they decrease fastest than $n^{-2H\vee 1}$ and $T^{-2H\vee 1}$, we get large time concentration inequalities.

In the following subsection, we outline our main application of Theorem 3.2.2 for which long time concentration holds.

3.2.3 Long time concentration inequalities for occupation measures

We now apply our general result to specific functionals to get the following theorem.

Theorem 3.2.3. *Let $H \in (0, 1)$ and $\Delta > 0$. Let $n \in \mathbb{N}^*$ and $T \geq 1$. Then,*

- (i) *there exist $C_{H,\Delta} > 0$ such that for all Lipschitz function $f : (\mathbb{R}^d, |\cdot|) \rightarrow (\mathbb{R}, |\cdot|)$ and for all $r \geq 0$,*

$$\mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n f(Y_{t_k}) - \mathbb{E}[f(Y_{t_k})] > r \right) \leq \exp \left(- \frac{r^2 n^{2-(2H\vee 1)}}{4C_{H,\Delta} \|f\|_{\text{Lip}}^2} \right). \quad (3.2.11)$$

- (ii) *there exist $\tilde{C}_H > 0$ such that for all Lipschitz function $f : (\mathbb{R}^d, |\cdot|) \rightarrow (\mathbb{R}, |\cdot|)$ and for all $r \geq 0$,*

$$\mathbb{P} \left(\frac{1}{T} \int_0^T (f(Y_t) - \mathbb{E}[f(Y_t)]) dt > r \right) \leq \exp \left(- \frac{r^2 T^{2-(2H\vee 1)}}{4\tilde{C}_H \|f\|_{\text{Lip}}^2} \right). \quad (3.2.12)$$

Proof. We apply Theorem 3.2.2 with the following functions F and \tilde{F} :

$$\forall x \in (\mathbb{R}^d)^n, \quad F(x) = \frac{1}{n} \sum_{k=1}^n f(x_k)$$

and

$$\forall x \in \mathcal{C}([0, T], \mathbb{R}^d), \quad \tilde{F}(x) = \frac{1}{T} \int_0^T f(x_t) dt$$

which are respectively $\frac{\|f\|_{\text{Lip}}}{n}$ -Lipschitz with respect to d_n (defined by (3.2.1)) and $\frac{\|f\|_{\text{Lip}}}{T}$ -Lipschitz with respect to d_T (defined by (3.2.2)). \square

3.3 Sketch of proof

Recall that F_Y and \tilde{F}_Y are defined by (3.2.6). The key element to get the bound (3.2.7) and (3.2.8) is to decompose F_Y and \tilde{F}_Y into a sum of martingale increments as follows. Let $(\mathcal{F}_t)_{t \geq 0}$ be the natural filtration associated to the standard Brownian motion W from which the fBm is derived through (3.2.3). For all $k \in \mathbb{N}$, set

$$M_k := \mathbb{E}[F_Y | \mathcal{F}_{t_k}] \quad \text{and} \quad \tilde{M}_k := \mathbb{E}[\tilde{F}_Y | \mathcal{F}_k]. \quad (3.3.1)$$

With these definitions, we have:

$$F_Y - \mathbb{E}[F_Y] = M_n = \sum_{k=1}^n M_k - M_{k-1} \quad \text{and} \quad \tilde{F}_Y - \mathbb{E}[\tilde{F}_Y] = \tilde{M}_{\lceil T \rceil} = \sum_{k=1}^{\lceil T \rceil} \tilde{M}_k - \tilde{M}_{k-1} \quad (3.3.2)$$

where $\lceil T \rceil$ denotes the least integer greater than or equal to T .

With this decomposition in hand, we first estimate the conditional exponential moments of the martingale increments $M_k - M_{k-1}$ and $\tilde{M}_k - \tilde{M}_{k-1}$ to get Proposition 3.2.1. This is the purpose of Proposition 3.5.2 for which the proof is based on the following lemma:

Lemma 3.3.1. *Let X be a centered real valued random variable such that for all $p \geq 2$, there exist $C, \zeta > 0$ such that*

$$\mathbb{E}[|X|^p] \leq C\zeta^{p/2} p\Gamma\left(\frac{p}{2}\right).$$

Then for all $\lambda > 0$,

$$\mathbb{E}[e^{\lambda X}] \leq e^{2C'\zeta\lambda^2}$$

with $C' = 1 \vee C$.

Proof. Since X is centered, by using the series expansion of the exponential function, we have:

$$\mathbb{E}[\exp(\lambda X)] \leq 1 + C \sum_{p=2}^{+\infty} \frac{\lambda^p \zeta^{\frac{p}{2}} p\Gamma\left(\frac{p}{2}\right)}{p!} \leq 1 + \sum_{p=2}^{+\infty} \frac{\lambda^p (C'\zeta)^{\frac{p}{2}} p\Gamma\left(\frac{p}{2}\right)}{p!} \quad (3.3.3)$$

with $C' = 1 \vee C$. We set $t^2 = \lambda^2 C'\zeta$, then

$$\begin{aligned} 1 + \sum_{p=2}^{+\infty} \frac{(t^2)^{\frac{p}{2}} p\Gamma\left(\frac{p}{2}\right)}{p!} &= 1 + \sum_{p=1}^{+\infty} \frac{(t^2)^p 2p\Gamma(p)}{(2p)!} + \sum_{p=1}^{+\infty} \frac{(t^2)^{p+\frac{1}{2}} (2p+1)\Gamma\left(p+\frac{1}{2}\right)}{(2p+1)!} \\ &= 1 + 2 \sum_{p=1}^{+\infty} \frac{(t^2)^p p\Gamma(p+1)}{(2p)!} + |t| \sum_{p=1}^{+\infty} \frac{(t^2)^p p\Gamma\left(p+\frac{1}{2}\right)}{(2p)!} \\ &\leq 1 + (2 + |t|) \sum_{p=1}^{+\infty} \frac{(t^2)^p p!}{(2p)!} \\ &\leq 1 + \left(1 + \frac{|t|}{2}\right) (e^{t^2} - 1) \quad \text{since} \quad 2(p!)^2 \leq (2p)! . \end{aligned}$$

Since for all $t \in \mathbb{R}$, $\frac{|t|}{2} \leq e^{t^2}$, we get $\frac{|t|}{2}(e^{t^2} - 1) \leq e^{t^2}(e^{t^2} - 1)$ which is equivalent to

$$1 + \left(1 + \frac{|t|}{2}\right) (e^{t^2} - 1) \leq e^{2t^2},$$

so that:

$$1 + \sum_{p=2}^{+\infty} \frac{(t^2)^{\frac{p}{2}} p \Gamma\left(\frac{p}{2}\right)}{p!} \leq e^{2t^2}.$$

Hence, we have in (3.3.3):

$$\mathbb{E}[\exp(\lambda X)] \leq \exp(2\lambda^2 \zeta C')$$

which concludes the proof. □

Remark 3.3.1. The previous proof follows the proof of Lemma 1.5 in Chapter 1 of [63]. We chose to give the details here since this step is crucial to get our main results.

Finally, the end of the proof of Proposition 3.2.1 (i) is based on the following implication: if there exists a deterministic sequence (u_k) such that

$$\mathbb{E}\left[e^{\lambda(M_k - M_{k-1})} \middle| \mathcal{F}_{k-1}\right] \leq e^{\lambda^2 u_k},$$

then

$$\mathbb{E}\left[e^{\lambda M_n}\right] = \mathbb{E}\left[e^{\lambda M_{n-1}} \mathbb{E}\left[e^{\lambda(M_n - M_{n-1})} \middle| \mathcal{F}_{n-1}\right]\right] \leq \exp(\lambda^2 u_n) \mathbb{E}\left[e^{\lambda M_{n-1}}\right]$$

so that

$$\mathbb{E}\left[e^{\lambda M_n}\right] \leq \exp\left(\lambda^2 \sum_{k=1}^n u_k\right).$$

The same arguments are used for item (ii) of Proposition 3.2.1.

Sections 3.4 and 3.5 are devoted to the proof of Proposition 3.2.1. The first step, detailed in Section 3.4, consists in giving a new expression to the martingale increments and to control them. The second step, which is outlined in Section 3.5.1, focuses on managing the conditional moments of these increments to get Proposition 3.5.2. The proof of Proposition 3.2.1 is finally achieved in Section 3.5.2.

Throughout the chapter, constants may change from line to line.

3.4 Control of the martingale increments

For the sake of clarity, we set $\Delta = 1$ in the sequel, so that by (3.2.6) we have $t_k = k$. When $\Delta > 0$ is arbitrary, the arguments are the same, it suffices to apply a rescaling.

Through equation (4.1.2) and the fact that b is Lipschitz continuous, for all $t \geq 0$, Y_t can be seen as a measurable functional of the time t , the initial condition x and the Brownian motion $(W_s)_{s \in [0, t]}$. Denote by $\Phi : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}^d$ this functional, we then have

$$\forall t \geq 0, \quad Y_t := \Phi_t(x, (W_s)_{s \in [0, t]}). \quad (3.4.1)$$

Now, let $k \geq 1$, we have

$$\begin{aligned}
 & |M_k - M_{k-1}| \\
 &= |\mathbb{E}[F_Y | \mathcal{F}_k] - \mathbb{E}[F_Y | \mathcal{F}_{k-1}]| \\
 &= \int_{\Omega} |F(\Phi_1(x, W_{[0,1]}), \dots, \Phi_k(x, W_{[0,k]}), \dots, \Phi_n(x, W_{[0,k]} \sqcup \tilde{w}_{[k,n]})) \\
 &\quad - F(\Phi_1(x, W_{[0,1]}), \dots, \Phi_k(x, W_{[0,k-1]} \sqcup \tilde{w}_{[k-1,k]}), \dots, \Phi_n(x, W_{[0,k-1]} \sqcup \tilde{w}_{[k-1,n]}))| \mathbb{P}_W(d\tilde{w}) \\
 &\leq \|F\|_{\text{Lip}} \int_{\Omega} \sum_{t=k}^n |\Phi_t(x, W_{[0,k]} \sqcup \tilde{w}_{[k,t]}) - \Phi_t(x, W_{[0,k-1]} \sqcup \tilde{w}_{[k-1,t]})| \mathbb{P}_W(d\tilde{w}). \quad (3.4.2)
 \end{aligned}$$

With exactly the same procedure, we get

$$|\tilde{M}_k - \tilde{M}_{k-1}| \leq \|\tilde{F}\|_{\text{Lip}} \int_{\Omega} \int_{k-1}^T |\Phi_t(x, W_{[0,k]} \sqcup \tilde{w}_{[k,t]}) - \Phi_t(x, W_{[0,k-1]} \sqcup \tilde{w}_{[k-1,t]})| dt \mathbb{P}_W(d\tilde{w}). \quad (3.4.3)$$

Let us introduce now some notations. First, for all $t \geq 0$ set $u := t - k + 1$, then for all $u \geq 0$, we define

$$X_u := \begin{cases} \Phi_{u+k-1}(x, (W_s)_{s \in [0,k]} \sqcup (\tilde{w}_s)_{s \in [k, u+k-1]}) & \text{if } u \geq 1 \\ \Phi_{u+k-1}(x, (W_s)_{s \in [0, u+k-1]}) & \text{otherwise,} \end{cases}$$

and

$$\tilde{X}_u := \Phi_{u+k-1}(x, (W_s)_{s \in [0, k-1]} \sqcup (\tilde{w}_s)_{s \in [k-1, u+k-1]}).$$

We then have

$$\begin{aligned}
 X_u &= X_0 + \int_0^u b(X_s) ds + \sigma \int_0^{k-1} (K_H(u+k-1, s) - K_H(k-1, s)) dW_s \\
 &\quad + \sigma \int_{k-1}^{k \wedge (u+k-1)} K_H(u+k-1, s) dW_s + \sigma \int_k^{k \vee (u+k-1)} K_H(u+k-1, s) d\tilde{w}_s
 \end{aligned} \quad (3.4.4)$$

and

$$\begin{aligned}
 \tilde{X}_u &= \tilde{X}_0 + \int_0^u b(\tilde{X}_s) ds + \sigma \int_0^{k-1} (K_H(u+k-1, s) - K_H(k-1, s)) dW_s \\
 &\quad + \sigma \int_{k-1}^{u+k-1} K_H(u+k-1, s) d\tilde{w}_s.
 \end{aligned} \quad (3.4.5)$$

Remark 3.4.1. Let us note that the integrals involving \tilde{w} in (3.4.4) and (3.4.5) and in the sequel have to be seen as Wiener integrals, so that they are defined $\mathbb{P}_W(d\tilde{w})$ almost surely.

Since $X_0 = \tilde{X}_0 = \Phi_{k-1}(x, (W_s)_{s \in [0, k-1]})$, we deduce from (3.4.4) and (3.4.5) that for

all $u \geq 0$

$$\begin{aligned} X_u - \tilde{X}_u &= \int_0^u b(X_s) - b(\tilde{X}_s) ds + \sigma \int_{k-1}^{k \wedge (u+k-1)} K_H(u+k-1, s) d(W - \tilde{w})_s \\ &= \int_0^u b(X_s) - b(\tilde{X}_s) ds + \sigma \int_0^{1 \wedge u} K_H(u+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \end{aligned} \quad (3.4.6)$$

where we have set $(W_s^{(k)})_{s \geq 0} := (W_{s+k-1} - W_{k-1})_{s \geq 0}$ which is a Brownian motion independent from \mathcal{F}_{k-1} and $(\tilde{w}_s^{(k)})_{s \geq 0} := (\tilde{w}_{s+k-1} - \tilde{w}_{k-1})_{s \geq 0}$.

In the remainder of the section, we proceed to a control of the quantity $|X_u - \tilde{X}_u|$. We have the following upperbound on $|X_u - \tilde{X}_u|$:

Proposition 3.4.1. *There exists $C_H > 0$ such that for all $u > 0$ and $k \in \mathbb{N}^*$,*

$$\begin{aligned} &|X_u - \tilde{X}_u| \\ &\leq C_H \sqrt{\Psi_H(u \vee 1, k)} \left(\sup_{v \in [0,1]} |W_v^{(k)} - \tilde{w}_v^{(k)}| + \sup_{v \in [0,2]} |G_v^{(k)}| \right. \\ &\quad \left. + \sup_{v \in [0,1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1-vs)^{H-\frac{3}{2}} d(W^{(k)} - \tilde{w}^{(k)})_s \right| \right) \end{aligned} \quad (3.4.7)$$

where $X_u - \tilde{X}_u$ is defined in (3.4.6), Ψ_H is defined by

$$\Psi_H(u, k) := C'_H \begin{cases} u^{2H-3} & \text{if } H \in (0, 1/2) \\ k^{1-2H} u^{4H-4} + u^{2H-3} & \text{if } H \in (1/2, 1) \end{cases}$$

with $C'_H > 0$ and $G_v^{(k)}$ is given by

$$G_v^{(k)} = \int_0^{1 \wedge v} K_H(v+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s.$$

In Subsections 3.4.1 and 3.4.2, we prove Proposition 3.4.1.

3.4.1 First case : $u \geq 2$

When $k \neq 1$

Lemma 3.4.1. *Let $k \neq 1$. Then, for all $u \geq 2$,*

$$|X_u - \tilde{X}_u|^2 \leq e^{-\alpha(u-2)} |X_2 - \tilde{X}_2|^2 + \Psi_H(u, k) \sup_{s \in [0,1]} |W_s^{(k)} - \tilde{w}_s^{(k)}|^2$$

where Ψ_H is defined in Proposition 3.4.1.

Proof. Let $u \geq 2$. In the following inequalities, we make use of Hypothesis 3.2.1 on the function b and to the elementary Young inequality $\langle a, b \rangle \leq \frac{1}{2} (\varepsilon |a|^2 + \frac{1}{\varepsilon} |b|^2)$ with $\varepsilon = 2\alpha$.

By (3.4.6),

$$\begin{aligned}
 & \frac{d}{du} |X_u - \tilde{X}_u|^2 \\
 &= 2\langle X_u - \tilde{X}_u, b(X_u) - b(\tilde{X}_u) \rangle + \langle X_u - \tilde{X}_u, \sigma \int_0^1 \frac{\partial}{\partial u} K_H(u+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \rangle \\
 &\leq -2\alpha |X_u - \tilde{X}_u|^2 + \alpha |X_u - \tilde{X}_u|^2 + \frac{|\sigma|^2}{2\alpha} \left| \int_0^1 \frac{\partial}{\partial u} K_H(u+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \right|^2 \\
 &\leq -\alpha |X_u - \tilde{X}_u|^2 + \frac{|\sigma|^2}{2\alpha} \left| \int_0^1 \frac{\partial}{\partial u} K_H(u+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \right|^2.
 \end{aligned}$$

We then apply Gronwall's lemma to obtain

$$|X_u - \tilde{X}_u|^2 \leq e^{-\alpha(u-2)} |X_2 - \tilde{X}_2|^2 + \frac{|\sigma|^2}{2\alpha} \int_2^u e^{-\alpha(u-v)} \left| \int_0^1 \frac{\partial}{\partial v} K_H(v+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \right|^2 dv. \quad (3.4.8)$$

Now, we set for all $v \geq 2$,

$$\begin{aligned}
 \varphi_k(v) &:= \int_0^1 \frac{\partial}{\partial v} K_H(v+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \\
 &= c_H \int_0^1 \left(\frac{v+k-1}{s+k-1} \right)^{H-\frac{1}{2}} (v-s)^{H-\frac{3}{2}} d(W^{(k)} - \tilde{w}^{(k)})_s. \quad (3.4.9)
 \end{aligned}$$

We apply an integration by parts to φ_k taking into account that $W_0^{(k)} = \tilde{w}_0^{(k)} = 0$:

$$\begin{aligned}
 \varphi_k(v) &= c_H \left(\frac{v+k-1}{k} \right)^{H-\frac{1}{2}} (v-1)^{H-\frac{3}{2}} (W_1^{(k)} - \tilde{w}_1^{(k)}) \\
 &\quad - c_H (1/2 - H) \int_0^1 (v+k-1)^{H-\frac{1}{2}} (s+k-1)^{-H-\frac{1}{2}} (v-s)^{H-\frac{3}{2}} (W_s^{(k)} - \tilde{w}_s^{(k)}) ds \\
 &\quad - c_H (3/2 - H) \int_0^1 \left(\frac{v+k-1}{s+k-1} \right)^{H-\frac{1}{2}} (v-s)^{H-\frac{5}{2}} (W_s^{(k)} - \tilde{w}_s^{(k)}) ds \\
 &=: c_H (I_1(v) + I_2(v) + I_3(v)). \quad (3.4.10)
 \end{aligned}$$

Recall that by (3.4.8), our goal here is to manage

$$\begin{aligned}
 & \int_2^u e^{-\alpha(u-v)} |\varphi_k(v)|^2 dv \\
 & \leq 3c_H^2 \left(\int_2^u e^{-\alpha(u-v)} |I_1(v)|^2 dv + \int_2^u e^{-\alpha(u-v)} |I_2(v)|^2 dv + \int_2^u e^{-\alpha(u-v)} |I_3(v)|^2 dv \right) \quad (3.4.11)
 \end{aligned}$$

To control each term involving I_1 , I_2 and I_3 in (3.4.11), we will need the following inequality:

$$\begin{aligned} & \int_2^u e^{-\alpha(u-v)} k^{1-2H} (v-1+k)^{2H-1} (v-1)^{2H-3} dv \\ & \leq C_H \begin{cases} k^{1-2H} (u-1)^{4H-4} + (u-1)^{2H-3} & \text{for } H > 1/2 \\ (u-1)^{2H-3} & \text{for } H < 1/2 \end{cases}. \end{aligned} \quad (3.4.12)$$

Inequality (3.4.12) is obtained through Lemma 3.4.2 and the elementary inequalities $(v-1+k)^{2H-1} \leq (v-1)^{2H-1} + k^{2H-1}$ if $H > 1/2$ and $(v-1+k)^{2H-1} \leq k^{2H-1}$ if $H < 1/2$.

Lemma 3.4.2. *Let $\alpha, \beta > 0$. Then, for all $u \geq 2$,*

$$\int_2^u e^{-\alpha(u-v)} (v-1)^{-\beta} dv \leq C_{\alpha, \beta} (u-1)^{-\beta}.$$

Proof. It is enough to apply an integration by parts and then use that

$$\sup_{v \in [2, u]} e^{-\alpha(u-v)} (v-1)^{-\beta-1} = \max \left(e^{-\alpha(u-2)}, (u-1)^{-\beta-1} \right)$$

to conclude the proof. \square

It remains to show how the terms involving I_1 , I_2 and I_3 in (3.4.11) can be reduced to the term (3.4.12). Let us begin with I_1 which is straightforward:

$$\begin{aligned} & \int_2^u e^{-\alpha(u-v)} |I_1(v)|^2 dv \\ & \leq |W_1^{(k)} - w_1^{(k)}|^2 \int_2^u e^{-\alpha(u-v)} k^{1-2H} (v-1+k)^{2H-1} (v-1)^{2H-3} dv \\ & \leq \sup_{s \in [0, 1]} |W_s^{(k)} - w_s^{(k)}|^2 \int_2^u e^{-\alpha(u-v)} k^{1-2H} (v-1+k)^{2H-1} (v-1)^{2H-3} dv. \end{aligned} \quad (3.4.13)$$

Then, using the definition of I_2 ,

$$\begin{aligned} & \int_2^u e^{-\alpha(u-v)} |I_2(v)|^2 dv \\ & \leq (1/2 - H)^2 \int_2^u e^{-\alpha(u-v)} (v-1+k)^{2H-1} (v-1)^{2H-3} (k-1)^{-2H-1} \left(\int_0^1 |W_s^{(k)} - w_s^{(k)}| ds \right)^2 dv \\ & \leq C_H \sup_{s \in [0, 1]} |W_s^{(k)} - w_s^{(k)}|^2 \int_2^u e^{-\alpha(u-v)} k^{1-2H} (v-1+k)^{2H-1} (v-1)^{2H-3} dv. \end{aligned} \quad (3.4.14)$$

Finally,

$$\begin{aligned}
 & \int_2^u e^{-\alpha(u-v)} |I_3(v)|^2 dv \\
 & \leq (3/2 - H)^2 \int_2^u e^{-\alpha(u-v)} (v-1+k)^{2H-1} (v-1)^{2H-5} \left(\int_0^1 (s+k-1)^{\frac{1}{2}-H} |W_s^{(k)} - w_s^{(k)}| ds \right)^2 dv \\
 & \leq C_H \sup_{s \in [0,1]} |W_s^{(k)} - w_s^{(k)}|^2 \int_2^u e^{-\alpha(u-v)} (v-1+k)^{2H-1} (v-1)^{2H-3} \left(\int_0^1 (s+k-1)^{\frac{1}{2}-H} ds \right)^2 dv \\
 & \leq C'_H \sup_{s \in [0,1]} |W_s^{(k)} - w_s^{(k)}|^2 \int_2^u e^{-\alpha(u-v)} k^{1-2H} (v-1+k)^{2H-1} (v-1)^{2H-3} dv
 \end{aligned} \tag{3.4.15}$$

where the last inequality is given by the following fact: there exists $C_H > 0$ such that for all $k \neq 1$, $\sup_{s \in [0,1]} (s+k-1)^{\frac{1}{2}-H} \leq C_H k^{\frac{1}{2}-H}$.

It remains to combine the three above inequalities (3.4.13), (3.4.14) and (3.4.15) with (3.4.12) to get the following in (3.4.11):

$$\begin{aligned}
 & \int_2^u e^{-\alpha(u-v)} |\varphi_k(v)|^2 dv \\
 & \leq C_H \sup_{s \in [0,1]} |W_s^{(k)} - w_s^{(k)}|^2 \begin{cases} k^{1-2H} (u-1)^{4H-4} + (u-1)^{2H-3} & \text{for } H > 1/2 \\ (u-1)^{2H-3} & \text{for } H < 1/2 \end{cases} .
 \end{aligned}$$

Putting this inequality into (3.4.8) gives the result (we can replace $u-1$ by u , the inequality remains true when $u \geq 2$ up to a constant). \square

When $k = 1$

Lemma 3.4.3. *Let $k = 1$. Then, for all $u \geq 2$,*

$$|X_u - \tilde{X}_u|^2 \leq e^{-\alpha(u-2)} |X_2 - \tilde{X}_2|^2 + \Psi_H(u, 1) \sup_{v \in [0, 1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1-vs)^{H-\frac{3}{2}} d(W^{(1)} - \tilde{w}^{(1)})_s \right|^2$$

where Ψ_H is defined in Proposition 3.4.1.

Proof. The proof begins as in the proof of Lemma 3.4.1. We have through inequality (3.4.8):

$$|X_u - \tilde{X}_u|^2 \leq e^{-\alpha(u-2)} |X_2 - \tilde{X}_2|^2 + \frac{|\sigma|^2}{2\alpha} \int_2^u e^{-\alpha(u-v)} |\varphi_1(v)|^2 dv \tag{3.4.16}$$

with

$$\begin{aligned}
 \varphi_1(v) &= c_H v^{H-\frac{1}{2}} \int_0^1 s^{\frac{1}{2}-H} (v-s)^{H-\frac{3}{2}} d(W^{(1)} - \tilde{w}^{(1)})_s \\
 &= c_H v^{H-\frac{1}{2}} v^{H-\frac{3}{2}} \int_0^1 s^{\frac{1}{2}-H} \left(1 - \frac{s}{v}\right)^{H-\frac{3}{2}} d(W^{(1)} - \tilde{w}^{(1)})_s.
 \end{aligned} \tag{3.4.17}$$

Since for $v \geq 2$, $v^{H-\frac{1}{2}}$ is bounded when $H < 1/2$, we have

$$\begin{aligned} & \int_2^u e^{-\alpha(u-v)} |\varphi_1(v)|^2 dv \\ & \leq c_H \int_2^u e^{-\alpha(u-v)} v^{(4H-4)\vee(2H-3)} \left| \int_0^1 s^{\frac{1}{2}-H} \left(1 - \frac{s}{v}\right)^{H-\frac{3}{2}} d(W^{(1)} - \tilde{w}^{(1)})_s \right|^2 dv \\ & \leq C_H \sup_{v' \in [0, 1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1 - v's)^{H-\frac{3}{2}} d(W^{(1)} - \tilde{w}^{(1)})_s \right|^2 \int_2^u e^{-\alpha(u-v)} v^{(4H-4)\vee(2H-3)} dv. \end{aligned}$$

Then, we use Lemma 3.4.2 in the previous inequality, which gives:

$$\begin{aligned} & \int_2^u e^{-\alpha(u-v)} |\varphi_1(v)|^2 dv \\ & \leq C_H \sup_{v' \in [0, 1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1 - v's)^{H-\frac{3}{2}} d(W^{(1)} - \tilde{w}^{(1)})_s \right|^2 (u-1)^{(4H-4)\vee(2H-3)}. \end{aligned}$$

This inequality combined with (3.4.16) concludes the proof (we can replace $u-1$ by u , the inequality remains true when $u \geq 2$ up to a constant). \square

3.4.2 Second case : $u \in [0, 2]$

The idea here is to use Gronwall lemma in its integral form. By Hypothesis 3.2.1, b is L -Lipschitz so that:

$$|X_u - \tilde{X}_u| \leq L \int_0^u |X_s - \tilde{X}_s| ds + \left| \int_0^{1 \wedge u} K_H(u+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \right|.$$

Then, for $u \in [0, 2]$,

$$\begin{aligned} |X_u - \tilde{X}_u| & \leq \left| \int_0^{1 \wedge u} K_H(u+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \right| \\ & \quad + \int_0^u \left| \int_0^{1 \wedge v} K_H(v+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \right| e^{L(u-v)} dv \\ & \leq e^{2L} \sup_{v \in [0, 2]} \left| \int_0^{1 \wedge v} K_H(v+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s \right| \end{aligned} \quad (3.4.18)$$

For all $k \geq 1$ and for all $v \in [0, 2]$, we set

$$G_v^{(k)}(W - \tilde{w}) = \int_0^{1 \wedge v} K_H(v+k-1, s+k-1) d(W^{(k)} - \tilde{w}^{(k)})_s. \quad (3.4.19)$$

The inequality (3.4.18) combined with Lemma 3.4.1 and Lemma 3.4.3 finally prove Proposition 3.4.1.

3.5 Conditional exponential moments of the martingale increments

3.5.1 Conditional moments of the martingale increments

Proposition 3.5.1. (i) *There exists $C, \zeta > 0$ such that for all $k \in \mathbb{N}^*$ and for all $p \geq 2$,*

$$\mathbb{E}[|M_k - M_{k-1}|^p | \mathcal{F}_{k-1}]^{1/p} \leq C \|F\|_{\text{Lip}} \psi_{n,k} \left(\zeta^{p/2} p \Gamma\left(\frac{p}{2}\right) \right)^{1/p} \quad a.s. \quad (3.5.1)$$

(ii) *There exists $C, \zeta > 0$ such that for all $k \in \mathbb{N}^*$ and for all $p \geq 2$,*

$$\mathbb{E}[|\tilde{M}_k - \tilde{M}_{k-1}|^p | \mathcal{F}_{k-1}]^{1/p} \leq C \|\tilde{F}\|_{\text{Lip}} \psi'_{T,k} \left(\zeta^{p/2} p \Gamma\left(\frac{p}{2}\right) \right)^{1/p} \quad a.s. \quad (3.5.2)$$

where $\psi_{n,k} := \sum_{u=1}^{n-k+1} \sqrt{\Psi_H(u, k)}$, $\psi'_{T,k} := \int_0^{T-k+1} \sqrt{\Psi_H(u \vee 1, k)} du$ and Ψ_H is defined in Proposition 3.4.1.

To prove this result, we first need the following intermediate outcome.

Lemma 3.5.1. *For all $k \in \mathbb{N}^*$, let $G^{(k)}$ be defined by (3.4.19). Then, for all $p \geq 2$, there exists $C > 0$ such that*

$$\begin{aligned} & \mathbb{E}[|M_k - M_{k-1}|^p | \mathcal{F}_{k-1}]^{1/p} \\ & \leq 3C \|F\|_{\text{Lip}} \psi_{n,k} \left(\mathbb{E} \left[\sup_{v \in [0,1]} |W_v^{(1)} - \tilde{W}_v^{(1)}|^p \right]^{1/p} + \mathbb{E} \left[\sup_{v \in [0,2]} |G_v^{(k)}(W - \tilde{W})|^p \right]^{1/p} \right. \\ & \quad \left. + \mathbb{E} \left[\sup_{v \in [0,1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1-vs)^{H-\frac{3}{2}} d(W^{(1)} - \tilde{W}^{(1)})_s \right|^p \right]^{1/p} \right) \quad a.s. \end{aligned}$$

Or equivalently, since $W^{(k)}$ and $\tilde{W}^{(k)}$ are iid we can replace $W^{(k)} - \tilde{W}^{(k)}$ by $\sqrt{2}W^{(k)}$:

$$\begin{aligned} & \mathbb{E}[|M_k - M_{k-1}|^p | \mathcal{F}_{k-1}]^{1/p} \\ & \leq 3C\sqrt{2} \|F\|_{\text{Lip}} \psi_{n,k} \left(\mathbb{E} \left[\sup_{v \in [0,1]} |W_v^{(1)}|^p \right]^{1/p} + \mathbb{E} \left[\sup_{v \in [0,2]} |G_v^{(k)}(W)|^p \right]^{1/p} \right. \\ & \quad \left. + \mathbb{E} \left[\sup_{v \in [0,1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1-vs)^{H-\frac{3}{2}} dW_s^{(1)} \right|^p \right]^{1/p} \right) \quad a.s. \end{aligned}$$

where $\psi_{n,k} = \sum_{u=1}^{n-k+1} \sqrt{\Psi_H(u, k)}$ and Ψ_H is defined in Proposition 3.4.1.

The same occurs for \tilde{M} instead of M by replacing F by \tilde{F} and $\psi_{n,k}$ by $\psi'_{T,k} = \int_0^{T-k+1} \sqrt{\Psi_H(u \vee 1, k)} du$.

Proof. For the sake of simplicity, assume that $\|F\|_{\text{Lip}} = 1$. By inequality (3.4.2), we

have for all $p \geq 2$,

$$|M_k - M_{k-1}|^p \leq \left(\int_{\Omega} \sum_{u=1}^{n-k+1} |X_u - \tilde{X}_u| \mathbb{P}_W(d\tilde{w}) \right)^p.$$

Now, we use Proposition 3.4.1 and for the sake of clarity we set $\|W^{(k)} - \tilde{w}^{(k)}\|_{\infty, [0,1]} := \sup_{v \in [0,1]} |W_v^{(k)} - \tilde{w}_v^{(k)}|$, $A(W^{(k)} - \tilde{w}^{(k)}) := \sup_{v \in [0,1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1-vs)^{H-\frac{3}{2}} d(W^{(k)} - \tilde{w}^{(k)})_s \right|$ and

$C_k(W^{(k)} - \tilde{w}^{(k)}) := \sup_{v \in [0,2]} |G_v^{(k)}(W - \tilde{w})|$. Then, by Jensen inequality,

$$\begin{aligned} & |M_k - M_{k-1}|^p \\ & \leq C^p \psi_{n,k}^p \left(\int_{\Omega} \|W^{(k)} - \tilde{w}^{(k)}\|_{\infty, [0,1]} + A(W^{(k)} - \tilde{w}^{(k)}) + C_k(W^{(k)} - \tilde{w}^{(k)}) \mathbb{P}_W(d\tilde{w}) \right)^p \\ & \leq 3^{p-1} C^p \psi_{n,k}^p \left[\int_{\Omega} \|W^{(k)} - \tilde{w}^{(k)}\|_{\infty, [0,1]}^p \mathbb{P}_W(d\tilde{w}) + \int_{\Omega} \left(A(W^{(k)} - \tilde{w}^{(k)}) \right)^p \mathbb{P}_W(d\tilde{w}) \right. \\ & \quad \left. + \int_{\Omega} \left(C_k(W^{(k)} - \tilde{w}^{(k)}) \right)^p \mathbb{P}_W(d\tilde{w}) \right] \end{aligned}$$

Recall that $W^{(k)} = (W_{s+k-1} - W_{k-1})_{s \geq 0}$ and thus $W^{(k)}$ is independent of \mathcal{F}_{k-1} . Then,

$$\begin{aligned} & \mathbb{E}[|M_k - M_{k-1}|^p | \mathcal{F}_{k-1}] \\ & \leq 3^{p-1} C^p \psi_{n,k}^p \mathbb{E} \left[\int_{\Omega} \|W^{(k)} - \tilde{w}^{(k)}\|_{\infty, [0,1]}^p \mathbb{P}_W(d\tilde{w}) + \int_{\Omega} \left(A(W^{(k)} - \tilde{w}^{(k)}) \right)^p \mathbb{P}_W(d\tilde{w}) \right. \\ & \quad \left. + \int_{\Omega} \left(C_k(W^{(k)} - \tilde{w}^{(k)}) \right)^p \mathbb{P}_W(d\tilde{w}) \middle| \mathcal{F}_{k-1} \right] \\ & \leq 3^{p-1} C^p \psi_{n,k}^p \mathbb{E} \left[\int_{\Omega} \|W^{(k)} - \tilde{w}^{(k)}\|_{\infty, [0,1]}^p \mathbb{P}_W(d\tilde{w}) + \int_{\Omega} \left(A(W^{(k)} - \tilde{w}^{(k)}) \right)^p \mathbb{P}_W(d\tilde{w}) \right. \\ & \quad \left. + \int_{\Omega} \left(C_k(W^{(k)} - \tilde{w}^{(k)}) \right)^p \mathbb{P}_W(d\tilde{w}) \right] \end{aligned}$$

We denote by $\mathcal{F}^{(k)}$ the filtration associated to $W^{(k)}$, we rewrite

$$\begin{aligned} & \mathbb{E}[|M_k - M_{k-1}|^p | \mathcal{F}_{k-1}] \\ & \leq 3^{p-1} C^p \psi_{n,k}^p \left(\mathbb{E} \left[\mathbb{E} \left[\|W^{(k)} - \tilde{W}^{(k)}\|_{\infty, [0,1]}^p \middle| \mathcal{F}_1^{(k)} \right] \right] + \mathbb{E} \left[\mathbb{E} \left[\left(A(W^{(k)} - \tilde{W}^{(k)}) \right)^p \middle| \mathcal{F}_1^{(k)} \right] \right] \right. \\ & \quad \left. + \mathbb{E} \left[\mathbb{E} \left[\left(C_k(W^{(k)} - \tilde{W}^{(k)}) \right)^p \middle| \mathcal{F}_1^{(k)} \right] \right] \right) \\ & = 3^{p-1} C^p \psi_{n,k}^p \left(\mathbb{E} \left[\|W^{(k)} - \tilde{W}^{(k)}\|_{\infty, [0,1]}^p \right] + \mathbb{E} \left[\left(A(W^{(k)} - \tilde{W}^{(k)}) \right)^p \right] \right. \\ & \quad \left. + \mathbb{E} \left[\left(C_k(W^{(k)} - \tilde{W}^{(k)}) \right)^p \right] \right) \end{aligned}$$

Using the elementary inequality $(a + b)^{1/p} \leq a^{1/p} + b^{1/p}$, we finally get :

$$\begin{aligned} & \mathbb{E} [|M_k - M_{k-1}|^p | \mathcal{F}_{k-1}]^{1/p} \\ & \leq 3C\psi_{n,k} \left(\mathbb{E} \left[\|W^{(k)} - \tilde{W}^{(k)}\|_{\infty, [0,1]}^p \right]^{1/p} + \mathbb{E} \left[\left(A(W^{(1)} - \tilde{W}^{(1)}) \right)^p \right]^{1/p} + \mathbb{E} \left[\left(C_k(W^{(k)} - \tilde{W}^{(k)}) \right)^p \right]^{1/p} \right) \end{aligned}$$

and the proof is over since $W^{(k)}$ and $\tilde{W}^{(k)}$ have respectively the same distribution as $W^{(1)}$ and $\tilde{W}^{(1)}$.

In the same way, we prove the result for \tilde{M} by using (3.4.3) which gives

$$|\tilde{M}_k - \tilde{M}_{k-1}|^p \leq \left(\int_{\Omega} \int_0^{T-k+1} |X_u - \tilde{X}_u| \mathbb{P}_W(d\tilde{w}) \right)^p$$

and Proposition 3.4.1. □

Proof of Proposition 3.5.1. With Lemma 3.5.1 in hand, we just need to prove that there exist $\zeta > 0$ such that for all $k \in \mathbb{N}^*$ and for all $p \geq 2$

$$\mathbb{E} \left[\sup_{v \in [0,1]} |W_v^{(1)}|^p \right]^{1/p} \leq \left(\zeta^{p/2} p \Gamma \left(\frac{p}{2} \right) \right)^{1/p}, \quad (3.5.3)$$

$$\mathbb{E} \left[\sup_{v \in [0,1/2]} \left| \int_0^1 s^{\frac{1}{2}-H} (1-vs)^{H-\frac{3}{2}} dW_s^{(1)} \right|^p \right]^{1/p} \leq \left(\zeta^{p/2} p \Gamma \left(\frac{p}{2} \right) \right)^{1/p} \quad (3.5.4)$$

$$\text{and } \mathbb{E} \left[\sup_{v \in [0,2]} |G_v^{(k)}(W)|^p \right]^{1/p} \leq \left(\zeta^{p/2} p \Gamma \left(\frac{p}{2} \right) \right)^{1/p}. \quad (3.5.5)$$

Condition (3.5.3) is given in Appendix 3.A and condition (3.5.5) follows from Proposition 3.B.2 since

$$\mathbb{E} \left[\sup_{v \in [0,2]} |G_v^{(k)}(W)|^p \right]^{1/p} \leq 2^{\alpha'_H} \mathbb{E} \left[\|G^{(k)}\|_{\alpha'_H, [0,2]}^p \right]^{1/p}$$

for all $\alpha'_H \in (0, 1)$. Hence, it remains to get (3.5.4). To this end, we set for all $v \in [0, 1/2]$,

$$\tilde{G}_v := \int_0^1 s^{\frac{1}{2}-H} (1-vs)^{H-\frac{3}{2}} dW_s^{(1)}.$$

Let $0 \leq v' < v \leq 1/2$, we have

$$\begin{aligned} \mathbb{E} [|\tilde{G}_v - \tilde{G}_{v'}|^2] &= \int_0^1 s^{1-2H} [(1-vs)^{H-\frac{3}{2}} - (1-v's)^{H-\frac{3}{2}}]^2 ds \\ &= \frac{1}{(3/2-H)^2} \int_0^1 s^{1-2H} \left(\int_{v'}^v (1-us)^{H-\frac{5}{2}} du \right)^2 ds. \end{aligned}$$

Since for all $u \in [0, 1/2]$ and for all $s \in [0, 1]$ we have $\frac{1}{2} \leq 1-us \leq 1$, we deduce that

$$\mathbb{E} [|\tilde{G}_v - \tilde{G}_{v'}|^2] \leq C_H (v-v')^2 \int_0^1 s^{1-2H} ds = \frac{C_H}{2-2H} (v-v')^2.$$

Hence, for all $\alpha \in (0, 1)$,

$$\sup_{0 \leq v' < v \leq \frac{1}{2}} \frac{\mathbb{E}[|\tilde{G}_v - \tilde{G}_{v'}|^2]^{1/2}}{|v - v'|^\alpha} < +\infty \quad (3.5.6)$$

Now, following carefully the proof of Proposition 3.B.2 in Appendix 3.B, one can show that (3.5.6) and the fact that \tilde{G} is a Gaussian process implies (3.5.4) since for all $\alpha \in (0, 1)$

$$\mathbb{E} \left[\sup_{v \in [0, 1/2]} |\tilde{G}_v|^p \right]^{1/p} \leq 2^{-\alpha} \mathbb{E} \left[\|\tilde{G}\|_{\alpha, [0, 1/2]}^p \right]^{1/p} + \mathbb{E}[|\tilde{G}_0|^p]^{1/p}.$$

□

3.5.2 Proof of Proposition 3.2.1

We have the following result:

Proposition 3.5.2. (i) *There exists $C', \zeta > 0$ such that for all $k \in \mathbb{N}^*$ and for all $\lambda > 0$,*

$$\mathbb{E}[\exp(\lambda(M_k - M_{k-1})) | \mathcal{F}_{k-1}] \leq \exp(2\lambda^2 \|F\|_{\text{Lip}}^2 \psi_{n,k}^2 C' \zeta) \quad a.s. \quad (3.5.7)$$

(ii) *There exists $C', \zeta > 0$ such that for all $k \in \mathbb{N}^*$ and for all $\lambda > 0$,*

$$\mathbb{E}[\exp(\lambda(\tilde{M}_k - \tilde{M}_{k-1})) | \mathcal{F}_{k-1}] \leq \exp(2\lambda^2 \|\tilde{F}\|_{\text{Lip}}^2 \psi'_{T,k}{}^2 C' \zeta) \quad a.s. \quad (3.5.8)$$

where $\psi_{n,k} := \sum_{u=1}^{n-k+1} \sqrt{\Psi_H(u, k)}$, $\psi'_{T,k} := \int_0^{T-k+1} \sqrt{\Psi_H(u \vee 1, k)} du$ and Ψ_H is defined in Proposition 3.4.1.

Proof. Let us prove (i). From $\mathbb{E}[M_k - M_{k-1} | \mathcal{F}_{k-1}] = 0$ and Proposition 3.5.1, we immediately get the result by using Lemma 3.3.1. □

Let us now conclude the proof of Proposition 3.2.1 (i). By the decomposition (3.3.2) and Proposition 3.5.2 (i), we have the following recursive inequality :

$$\mathbb{E} \left[e^{\lambda M_n} \right] = \mathbb{E} \left[e^{\lambda M_{n-1}} \mathbb{E} \left[e^{\lambda(M_n - M_{n-1})} \middle| \mathcal{F}_{n-1} \right] \right] \leq \exp(2\lambda^2 \|F\|_{\text{Lip}}^2 \psi_{n,n}^2 C' \zeta) \mathbb{E} \left[e^{\lambda M_{n-1}} \right]$$

which gives

$$\mathbb{E} \left[e^{\lambda M_n} \right] \leq \exp \left(2\lambda^2 \|F\|_{\text{Lip}}^2 C' \zeta \sum_{k=1}^n \psi_{n,k}^2 \right). \quad (3.5.9)$$

Equation (3.5.9) combined with Lemma 3.5.2 (see below) finally proves Proposition 3.2.1 (i). The proof of item (ii) is exactly the same.

Lemma 3.5.2. (i) *Let $n \in \mathbb{N}^*$ and $(\psi_{n,k})$ be defined as in Proposition 3.5.1. There exists $C_H > 0$ such that*

$$\sum_{k=1}^n \psi_{n,k}^2 \leq C_H n^{2(H \vee \frac{1}{2})}.$$

(ii) Let $T \geq 1$ and $(\psi'_{T,k})$ be defined as in Proposition 3.5.1. There exists $C_H > 0$ such that

$$\sum_{k=1}^{[T]} \psi'^2_{T,k} \leq C_H T^{2(H \vee \frac{1}{2})}.$$

Proof. (i) Recall that $\psi_{n,k} = \sum_{u=1}^{n-k+1} \sqrt{\Psi_H(u,k)}$ with

$$\Psi_H(u,k) := C_H \begin{cases} u^{2H-3} & \text{if } H \in (0, 1/2) \\ k^{1-2H} u^{4H-4} + u^{2H-3} & \text{if } H \in (1/2, 1) \end{cases}$$

and $C_H > 0$.

▷ First case: $H \in (0, 1/2)$. We have

$$\sum_{u=1}^{n-k+1} u^{H-\frac{3}{2}} \leq \sum_{u=1}^{+\infty} u^{H-\frac{3}{2}} < +\infty.$$

Then,

$$\sum_{k=1}^n \psi^2_{n,k} \leq C_H n$$

which concludes the proof for $H \in (0, 1/2)$.

▷ Second case: $H \in (1/2, 1)$. We have

$$\sum_{u=1}^{n-k+1} u^{H-\frac{3}{2}} \leq \int_0^{n-k+1} t^{H-\frac{3}{2}} dt = \frac{1}{H-1/2} (n-k+1)^{H-\frac{1}{2}}$$

and

$$\sum_{u=1}^{n-k+1} u^{2H-2} \leq \int_0^{n-k+1} t^{2H-2} dt = \frac{1}{2H-1} (n-k+1)^{2H-1}.$$

Then,

$$\begin{aligned} \sum_{k=1}^n \psi^2_{n,k} &\leq C_{1,H} \sum_{k=1}^n (n-k+1)^{2H-1} + C_{2,H} \sum_{k=1}^n k^{1-2H} (n-k+1)^{4H-2} \\ &\leq C_{1,H} n^{2H} + C_{2,H} (n+1)^{2H} \frac{1}{n+1} \sum_{k=1}^{n+1} \left(\frac{k}{n+1}\right)^{1-2H} \left(1 - \frac{k}{n+1}\right)^{4H-2}. \end{aligned}$$

Since

$$\frac{1}{n+1} \sum_{k=1}^{n+1} \left(\frac{k}{n+1}\right)^{1-2H} \left(1 - \frac{k}{n+1}\right)^{4H-2} \xrightarrow{n \rightarrow +\infty} \int_0^1 x^{1-2H} (1-x)^{4H-2} dx < +\infty$$

we finally get the result when $H \in (1/2, 1)$.

(ii) Recall that $\psi'_{T,k} = \int_0^{T-k+1} \sqrt{\Psi_H(u \vee 1, k)} du$.

▷ First case: $H \in (0, 1/2)$. We have

$$\int_0^{T-k+1} (u \vee 1)^{H-\frac{3}{2}} du \leq 1 + \int_1^{+\infty} u^{H-\frac{3}{2}} du < +\infty.$$

Then,

$$\sum_{k=1}^{\lceil T \rceil} \psi_{T,k}^2 \leq C_H \lceil T \rceil \leq \tilde{C}_H T$$

which concludes the proof for $H \in (0, 1/2)$.

▷ Second case: $H \in (1/2, 1)$. We have

$$\int_0^{T-k+1} (u \vee 1)^{H-\frac{3}{2}} du = 1 + \frac{1}{H-1/2} [(T-k+1)^{H-1/2} - 1] \leq \frac{1}{H-1/2} (T-k+1)^{H-\frac{1}{2}}$$

and

$$\int_0^{T-k+1} (u \vee 1)^{2H-2} du = 1 + \frac{1}{2H-1} [(T-k+1)^{2H-1} - 1] \leq \frac{1}{2H-1} (T-k+1)^{2H-1}.$$

Then,

$$\begin{aligned} \sum_{k=1}^{\lceil T \rceil} \psi_{T,k}^2 &\leq C_{1,H} \sum_{k=1}^{\lceil T \rceil} (T-k+1)^{2H-1} + C_{2,H} \sum_{k=1}^{\lceil T \rceil} k^{1-2H} (T-k+1)^{4H-2} \\ &\leq C_{1,H} \lceil T \rceil T^{2H-1} + C_{2,H} (\lceil T \rceil + 1)^{2H} \frac{1}{\lceil T \rceil + 1} \sum_{k=1}^{\lceil T \rceil + 1} \left(\frac{k}{\lceil T \rceil + 1} \right)^{1-2H} \left(1 - \frac{k}{\lceil T \rceil + 1} \right)^{4H-2}. \end{aligned}$$

Since

$$\frac{1}{\lceil T \rceil + 1} \sum_{k=1}^{\lceil T \rceil + 1} \left(\frac{k}{\lceil T \rceil + 1} \right)^{1-2H} \left(1 - \frac{k}{\lceil T \rceil + 1} \right)^{4H-2} \xrightarrow{T \rightarrow +\infty} \int_0^1 x^{1-2H} (1-x)^{4H-2} dx < +\infty$$

we finally get the result when $H \in (1/2, 1)$. \square

3.A Sub-Gaussianity of the supremum of the Brownian motion

Proposition 3.A.1. *Let $(W_t)_{t \geq 0}$ be a d -dimensional standard Brownian motion. There exist $\eta, \eta' > 0$ such that*

$$\forall x \geq 0, \quad \mathbb{P} \left(\sup_{t \in [0,1]} |W_t| > x \right) \leq \eta' e^{-\eta x^2} \quad (3.A.1)$$

Consequently, for all $p \geq 2$,

$$\mathbb{E} \left[\sup_{t \in [0,1]} |W_t|^p \right] \leq \frac{\eta'}{2} \left(\frac{1}{\eta} \right)^{p/2} p \Gamma \left(\frac{p}{2} \right) \quad (3.A.2)$$

where $\Gamma(x) := \int_0^{+\infty} e^{-u} u^{x-1} du$.

Proof.

$$\sup_{t \in [0,1]} |W_t| = \sup_{t \in [0,1]} \left(\sum_{i=1}^d |W_t^i|^2 \right)^{1/2} = \left(\sup_{t \in [0,1]} \sum_{i=1}^d |W_t^i|^2 \right)^{1/2} \leq \left(\sum_{i=1}^d \sup_{t \in [0,1]} |W_t^i|^2 \right)^{1/2} \leq \sum_{i=1}^d \sup_{t \in [0,1]} |W_t^i|.$$

Therefore for all $x \geq 0$, we have

$$\mathbb{P} \left(\sup_{t \in [0,1]} |W_t| \geq x \right) \leq \mathbb{P} \left(\sum_{i=1}^d \sup_{t \in [0,1]} |W_t^i| \geq x \right) \leq \sum_{i=1}^d \mathbb{P} \left(\sup_{t \in [0,1]} |W_t^i| \geq x \right) = d \times \mathbb{P} \left(\sup_{t \in [0,1]} |W_t^1| \geq x \right).$$

Since $\sup_{t \in [0,1]} |W_t^1| = \max \left(\sup_{t \in [0,1]} (-W_t^1), \sup_{t \in [0,1]} W_t^1 \right)$ and $(W_t^1)_{t \geq 0} \stackrel{\mathcal{L}}{=} (-W_t^1)_{t \geq 0}$, we have

$$\mathbb{P} \left(\sup_{t \in [0,1]} |W_t| \geq x \right) \leq d \left(\mathbb{P} \left(\sup_{t \in [0,1]} (-W_t^1) \geq x \right) + \mathbb{P} \left(\sup_{t \in [0,1]} W_t^1 \geq x \right) \right) = 2d \times \mathbb{P} \left(\sup_{t \in [0,1]} W_t^1 \geq x \right).$$

By the reflection principle, we know that $\mathbb{P} \left(\sup_{t \in [0,1]} W_t^1 \geq x \right) = 2\mathbb{P}(W_1^1 \geq x)$ which induces finally that

$$\mathbb{P} \left(\sup_{t \in [0,1]} |W_t| \geq x \right) \leq 4d \mathbb{P}(W_1^1 \geq x) = \frac{4d}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{1}{2}s^2} ds \leq C_d e^{-\frac{1}{4}x^2}. \quad (3.A.3)$$

Then, (3.A.2) follows from (3.A.1) by using the formula $\mathbb{E}[X] = \int_0^{+\infty} \mathbb{P}(X > x) dx$ for non-negative random variables and a simple change of variable. \square

3.B Uniform sub-Gaussianity of $\|G^{(k)}\|_{\alpha,[0,2]}$

In this section, we consider the following Gaussian processes: for all $k \in \mathbb{N}^*$,

$$\forall v \in [0, 2], \quad G_v^{(k)} := \int_0^{1 \wedge v} K_H(v+k-1, s+k-1) dW_s \quad (3.B.1)$$

where $(W_t)_{t \in [0, T]}$ is a d -dimensional Brownian motion and K_H is defined by (3.2.4).

Remark 3.B.1. Since we are interested in the law of $G^{(k)}$, we have replaced $W^{(k)}$ by W in the expression of $G^{(k)}$ given by (3.4.19).

First, we have the following control on the second moment of $G^{(k)}$ -increments.

Proposition 3.B.1. *There exists $C_H > 0$ such that for all $k \in \mathbb{N}^*$ and for all $0 \leq v' < v \leq 2$,*

$$\mathbb{E} \left[\left| G_v^{(k)} - G_{v'}^{(k)} \right|^2 \right] \leq C_H |v - v'|^{2\alpha_H} \quad (3.B.2)$$

$$\text{with } \alpha_H := \begin{cases} H & \text{if } H < 1/2 \\ \frac{H}{2} & \text{if } H > 1/2 \end{cases}.$$

Proof. Let $0 \leq v' < v \leq 2$. Then,

$$\begin{aligned} & G_v^{(k)} - G_{v'}^{(k)} \\ &= \int_0^{1 \wedge v'} K_H(v + k - 1, s + k - 1) - K_H(v' + k - 1, s + k - 1) dW_s \\ &\quad + \int_{1 \wedge v'}^{1 \wedge v} K_H(v + k - 1, s + k - 1) dW_s \\ &= \int_0^{1 \wedge v'} \left(\int_{v'}^v \frac{\partial}{\partial u} K_H(u + k - 1, s + k - 1) du \right) dW_s + \int_{1 \wedge v'}^{1 \wedge v} K_H(v + k - 1, s + k - 1) dW_s \end{aligned} \quad (3.B.3)$$

with

$$\frac{\partial}{\partial u} K_H(u + k - 1, s + k - 1) = c_H \left(\frac{u + k - 1}{s + k - 1} \right)^{H - \frac{1}{2}} (u - s)^{H - \frac{3}{2}}. \quad (3.B.4)$$

Then, we deduce the following expression for the moment of order 2:

$$\begin{aligned} & \mathbb{E} \left[\left| G_v^{(k)} - G_{v'}^{(k)} \right|^2 \right] \\ &= \int_0^{1 \wedge v'} \left(\int_{v'}^v \frac{\partial}{\partial u} K_H(u + k - 1, s + k - 1) du \right)^2 ds + \int_{1 \wedge v'}^{1 \wedge v} K_H(v + k - 1, s + k - 1)^2 ds \\ &=: I_1(v, v') + I_2(v, v'). \end{aligned} \quad (3.B.5)$$

Now, let us distinguish the two cases: $k > 1$ and $k = 1$:

▷ First case: $k > 1$

We begin with the first integral in (3.B.5), namely $I_1(v, v')$: let us note that in the expression (3.B.4)

$$\sup_{k > 1} \sup_{u, s \in [0, 2]} \left(\frac{u + k - 1}{s + k - 1} \right)^{H - \frac{1}{2}} < +\infty.$$

Hence,

$$\begin{aligned}
 I_1(v, v') &\leq C_H \int_0^{1 \wedge v'} \left(\int_{v'}^v (u-s)^{H-\frac{3}{2}} du \right)^2 ds \\
 &= \frac{C_H}{(H-1/2)^2} \int_0^{1 \wedge v'} \left[(v-s)^{H-\frac{1}{2}} - (v'-s)^{H-\frac{1}{2}} \right]^2 ds \\
 &\leq \frac{C_H}{(H-1/2)^2} \int_0^{v'} \left[(v-s)^{H-\frac{1}{2}} - (v'-s)^{H-\frac{1}{2}} \right]^2 ds \\
 &\leq C'_H \begin{cases} (v-v')^{2H} & \text{if } H < 1/2 \\ (v-v')^H & \text{if } H > 1/2 \end{cases} .
 \end{aligned} \tag{3.B.6}$$

and the last inequality is given by the following estimate:

Lemma 3.B.1. *There exists $\tilde{C}_H > 0$ such that for all $0 \leq v' < v \leq 2$,*

$$\int_0^{v'} \left[(v-s)^{H-\frac{1}{2}} - (v'-s)^{H-\frac{1}{2}} \right]^2 ds \leq \tilde{C}_H \begin{cases} (v-v')^{2H} & \text{if } H < 1/2 \\ (v-v')^H & \text{if } H > 1/2 \end{cases} .$$

Proof. First, we easily have

$$\begin{aligned}
 &\int_0^{v'} \left[(v-s)^{H-\frac{1}{2}} - (v'-s)^{H-\frac{1}{2}} \right]^2 ds \\
 &= \frac{1}{2H} [v^{2H} + v'^{2H} - (v-v')^{2H}] - 2 \int_0^{v'} [(v-s)(v'-s)]^{H-\frac{1}{2}} ds.
 \end{aligned}$$

Now, since

$$[(v-s)(v'-s)]^{H-\frac{1}{2}} \geq \begin{cases} (v-s)^{2H-1} & \text{if } H < 1/2 \\ (v'-s)^{2H-1} & \text{if } H > 1/2 \end{cases}$$

we get after some computations

$$\int_0^{v'} \left[(v-s)^{H-\frac{1}{2}} - (v'-s)^{H-\frac{1}{2}} \right]^2 ds \leq \frac{1}{2H} \begin{cases} v'^{2H} - v^{2H} + (v-v')^{2H} & \text{if } H < 1/2 \\ v^{2H} - v'^{2H} - (v-v')^{2H} & \text{if } H > 1/2 \end{cases} .$$

Moreover, when $H > 1/2$, for all $0 \leq v' < v \leq 2$,

$$\begin{aligned}
 v'^{2H} - v^{2H} + (v-v')^{2H} &= (v^H - v'^H)(v^H + v'^H) - (v-v')^{2H} \\
 &= (v-v')^H (v^H + v'^H - (v-v')^H) \\
 &\leq C_H (v-v')^H
 \end{aligned}$$

and when $H < 1/2$, $v'^{2H} - v^{2H} < 0$. So finally, we have the desired result. \square

We can now move on the second term in (3.B.5), namely $I_2(v, v')$. By Theorem 3.2 in [20], we have the following upperbound

$$I_2(v, v') \leq c_H^2 \int_{1 \wedge v'}^{1 \wedge v} (s+k-1)^{-2|H-\frac{1}{2}|} (v-s)^{-2(\frac{1}{2}-H)} ds$$

where $x_+ = \max(x, 0)$. Then, since $\sup_{k>1} \sup_{s \in [0,2]} (s+k-1)^{-2|H-\frac{1}{2}|} < +\infty$, we have

$$\begin{aligned} I_2(v, v') &\leq C_H \int_{1 \wedge v'}^{1 \wedge v} (v-s)^{-2(\frac{1}{2}-H)_+} ds \\ &= C_H (1 \wedge v - 1 \wedge v')^{2(H \wedge \frac{1}{2})} \\ &\leq C_H (v - v')^{2(H \wedge \frac{1}{2})} \\ &\leq \tilde{C}_H \begin{cases} (v - v')^{2H} & \text{if } H < 1/2 \\ (v - v')^H & \text{if } H > 1/2 \end{cases}. \end{aligned} \quad (3.B.7)$$

By using (3.B.6) and (3.B.7) in (3.B.5), we end the proof of Proposition 3.B.1 for $k > 1$.

▷ Second case: $k = 1$

Let us divide this part of the proof into three new cases:

First, consider $0 \leq v' < v \leq 1$, then $G^{(1)}$ coincides in law with the fractional Brownian motion:

$$\mathbb{E} \left[|G_v^{(1)} - G_{v'}^{(1)}|^2 \right] = (v - v')^{2H}.$$

Secondly, for $1 \leq v' < v \leq 2$, by (3.B.5):

$$\begin{aligned} \mathbb{E} \left[|G_v^{(1)} - G_{v'}^{(1)}|^2 \right] &= \int_0^1 \left(\int_{v'}^v \frac{\partial}{\partial u} K_H(u+k-1, s+k-1) du \right)^2 ds \\ &\leq \int_0^{v'} \left(\int_{v'}^v \frac{\partial}{\partial u} K_H(u+k-1, s+k-1) du \right)^2 ds \\ &\quad + \int_{v'}^v K_H(v+k-1, s+k-1)^2 ds \\ &= (v - v')^{2H}. \end{aligned}$$

Finally, if $0 \leq v' < 1 \leq v \leq 2$, we get the following by using the two previous cases:

$$\begin{aligned} \mathbb{E} \left[|G_v^{(1)} - G_{v'}^{(1)}|^2 \right] &\leq 2\mathbb{E} \left[|G_v^{(1)} - G_1^{(1)}|^2 \right] + 2\mathbb{E} \left[|G_1^{(1)} - G_{v'}^{(1)}|^2 \right] \\ &\leq 2 \left((v-1)^{2H} + (1-v)^{2H} \right) \\ &\leq 4(v - v')^{2H}. \end{aligned}$$

This inequality concludes the proof of Proposition 3.B.1 for $k = 1$. □

We can now state the result of uniform Sub-Gaussianity:

Proposition 3.B.2. *There exist $\eta, \eta' > 0$ such that for all $k \in \mathbb{N}^*$,*

$$\forall x \geq 0, \quad \mathbb{P} \left(\|G^{(k)}\|_{\alpha'_{H,[0,2]}} > x \right) \leq \eta' e^{-\eta x^2} \quad (3.B.8)$$

with $0 < \alpha'_H < \alpha_H$ and α_H is defined in Proposition 3.B.1.

Consequently, for all $k \in \mathbb{N}^*$ and for all $p \geq 2$,

$$\mathbb{E} \left[\|G^{(k)}\|_{\alpha'_H, [0,2]}^p \right] \leq \frac{\eta'}{2} \left(\frac{1}{\eta} \right)^{p/2} p \Gamma \left(\frac{p}{2} \right) \quad (3.B.9)$$

where $\Gamma(x) := \int_0^{+\infty} e^{-u} u^{x-1} du$.

Proof. Let us first note that since $G^{(k)}$ is a centered Gaussian process (for all $k \in \mathbb{N}^*$), there exists $C > 0$ such that for all $p \geq 1$ and for all $0 \leq v' < v \leq 2$:

$$\mathbb{E} \left[\left| G_v^{(k)} - G_{v'}^{(k)} \right|^p \right]^{1/p} \leq C \sqrt{p} \mathbb{E} \left[\left| G_v^{(k)} - G_{v'}^{(k)} \right|^2 \right]^{1/2}.$$

Then, we obtain through Proposition 3.B.1,

$$\forall p \geq 1, \quad \sup_{0 \leq v' < v \leq 2} \frac{\mathbb{E} \left[\left| G_v^{(k)} - G_{v'}^{(k)} \right|^p \right]^{1/p}}{|v - v'|^{\alpha_H}} \leq \tilde{C}_H \sqrt{p}. \quad (3.B.10)$$

Now by Theorem A.19 in [29], (3.B.10) implies that for all $0 < \alpha'_H < \alpha_H$, there exists $\eta_1 > 0$ such that

$$\mathbb{E} \left[\exp \left(\eta_1 \|G^{(k)}\|_{\alpha'_H, [0,2]}^2 \right) \right] < +\infty$$

and by Lemma A.17 in [29] (characterization of Gaussian integrability), this condition is equivalent to the existence of $\eta, \eta' > 0$ (depending only on η_1) such that (3.B.8) is true. Then, (3.B.9) follows from (3.B.8) by using the formula $\mathbb{E}[X] = \int_0^{+\infty} \mathbb{P}(X > x) dx$ for non-negative random variables and a simple change of variable. \square

4. A general drift estimation procedure for Stochastic Differential Equations with additive fractional noise

This chapter is based on the article [60], this is a joint work with Panloup and Tindel submitted in *Electronic Journal of Statistics*.

Contents

4.1	Introduction	117
4.2	Main results	119
4.3	Preliminary Results	126
4.4	Proof of the consistency theorems	134
4.5	Rate of convergence: proof of Theorem 4.2.4	144
4.6	Identifiability assumption	151
4.7	Numerical Discussions and Illustrations	156
4.A	Tightness and convergence of the occupation measures	160
4.B	Proof of Proposition 4.3.8	168

4.1 Introduction

Let B be a d -dimensional fractional Brownian motion with Hurst parameter $H \in (0, 1)$ defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We recall that B is a centered Gaussian process. Its law is thus characterized by its covariance function, which is defined by

$$\mathbb{E}[B_t^i B_s^j] = \frac{1}{2} (t^{2H} + s^{2H} - |t - s|^{2H}) \mathbb{1}_{\{0\}}(i - j), \quad s, t \in \mathbb{R}.$$

The variance of the increments of B is then given by

$$\mathbb{E}[|B_t^i - B_s^i|^2] = |t - s|^{2H}, \quad s, t \in \mathbb{R}, \quad i = 1, \dots, m, \quad (4.1.1)$$

and this implies that almost surely the fBm paths are γ -Hölder continuous for any $\gamma < H$.

In this chapter, we will consider the following \mathbb{R}^d -valued stochastic differential equation driven by B :

$$Y_t = y_0 + \int_0^t b_{\vartheta_0}(Y_s) ds + \sigma B_t, \quad t \in [0, T]. \quad (4.1.2)$$

Here $y_0 \in \mathbb{R}^d$ is a given initial condition, $B = (B^1, \dots, B^d)$ is the aforementioned fractional Brownian motion (fBm), the unknown parameter ϑ_0 lies in a certain set Θ which will be specified later on, $\{b_{\vartheta}(\cdot), \vartheta \in \Theta\}$ is a known family of drift coefficients with $b_{\vartheta}(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$, and σ is a $d \times d$ -matrix which is supposed to be known. More precisely, we do not discuss here the problem of estimation of the diffusion parameter σ and of the Hurst index H (on this topic, see *e.g.* [14], [41] or [68]). For the sake of simplicity, we also assume throughout the chapter that σ is invertible (on this topic, see Remark 4.2.4). Our aim is to get an accurate estimation of ϑ_0 according to some discrete observations of Y .

The estimation problem for the drift term in equation (4.1.2) has received a lot of attention in the recent past (see *e.g.* [3, 39, 40, 43, 46, 61, 67]). However, the following restrictions hold in all those contributions:

- The coefficient $b_{\vartheta_0}(Y_s)$ is of the form $\vartheta_0 b(Y_s)$ or even $\vartheta_0 Y_s$ when Ornstein-Uhlenbeck processes are involved.
- The observation is either in continuous time or a discretized version of continuous observations.
- Rates of convergence of the estimators are not computed, a noticeable exception being the central limit theorems obtained in [39].

Let us also mention the nonparametric method put forward in the interesting recent paper [17]. The context in [17] is much more general than in the aforementioned references, but the estimation procedure is based on the observation of several paths of (4.1.2). This makes its practical implementation delicate.

With these considerations in mind, let us recall that the article [55] proposed an estimator valid for a wide class of functions b in (4.1.2), directly based on discrete observations of the process. This estimator is obtained through a least square procedure which is easily implemented. It is fairly general, but still exhibits some drawbacks:

- For all values of the parameter ϑ , the drift $b_{\vartheta}(\cdot)$ is assumed to be of the form $\nabla F(\cdot; \vartheta)$ for a real valued function F . This is obviously a restriction when $d > 1$.
- The result in [55] is obtained in a quite standard ergodic framework, but it is also supposed that for the stationary solution \bar{Y}_0 to (4.1.2) we have

$$\mathbb{E} [|b_{\vartheta_0}(\bar{Y}_0)|^2] = \mathbb{E} [|b_{\vartheta}(\bar{Y}_0)|^2] \quad \text{iff} \quad \vartheta = \vartheta_0. \quad (4.1.3)$$

This identifiability condition is hard to verify in practice.

The current work proposes thus to improve on the two aspects mentioned above. It introduces a procedure which applies to non gradient drifts and is valid under less restrictive conditions than (4.1.3).

4.2 Main results

In this section we will first give some general notation which will be used throughout the chapter. Then we will specify our assumptions on the coefficients of (4.1.2) and describe the estimator we are considering. Eventually we give our almost sure convergence result as well as the convergence rate we have been able to obtain.

4.2.1 Notations

We consider the set of parameters Θ as a subset of \mathbb{R}^q for $q \geq 1$. Let $f : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}$ be a \mathcal{C}^{p_1, p_2} function, where p_1, p_2 are two integers greater than 1. Then for any $p \leq p_1$ and any tuple $(i_1, \dots, i_p) \in \{1, \dots, d\}^p$, we set $\partial_x^{i_1 \dots i_p} f$ for $\frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}$. Analogously, for $p \leq p_2$ we use the notation $\partial_\vartheta^{i_1 \dots i_p} f$ for $\frac{\partial^p f}{\partial \vartheta_{i_1} \dots \partial \vartheta_{i_p}}$ whenever $(i_1, \dots, i_p) \in \{1, \dots, q\}^p$. Moreover, we will write $\partial_x f$ (resp. $\partial_\vartheta f$) for the Jacobi-matrices $(\partial_{x_1} f, \dots, \partial_{x_d} f)$ (resp. $(\partial_{\vartheta_1} f, \dots, \partial_{\vartheta_q} f)$).

Let $\mathcal{M}_1(\mathbb{R}^d)$ denotes the set of probability measures on \mathbb{R}^d . We say that d is a distance on $\mathcal{M}_1(\mathbb{R}^d)$ if it metrizes its usual topology, namely the weak convergence topology. Among those distances We will consider the p -Wasserstein distance, which is defined as follows: for every $\nu, \mu \in \mathcal{M}_1(\mathbb{R}^d)$, we introduce the set $\mathcal{C}(\nu, \mu)$ of couplings between ν and μ , that is

$$\mathcal{C}(\nu, \mu) = \{(X, Y); \mathcal{L}(X) = \nu, \mathcal{L}(Y) = \mu\}. \quad (4.2.1)$$

Then the p -Wasserstein distance is written as

$$\mathcal{W}_p(\nu, \mu) = \inf \left\{ \mathbb{E}[|X - Y|^p]^{\frac{1}{p}}; (X, Y) \in \mathcal{C}(\nu, \mu) \right\}. \quad (4.2.2)$$

Remark 4.2.1. The distance \mathcal{W}_1 can also be represented as

$$\mathcal{W}_1(\nu, \mu) = \sup \{ |\nu(h) - \mu(h)|; \|h\|_{\text{Lip}} \leq 1 \}. \quad (4.2.3)$$

We will denote by \mathcal{D}_p the set of distances on $\mathcal{M}_1(\mathbb{R}^d)$ dominated by the p -Wasserstein distance for a given $p > 0$. Namely, we set

$$\mathcal{D}_p := \{ \text{distances } d \text{ on } \mathcal{M}_1(\mathbb{R}^d); \exists c > 0 \text{ such that } \forall \nu, \mu \in \mathcal{M}_1(\mathbb{R}^d), d(\nu, \mu) \leq c \mathcal{W}_p(\nu, \mu) \} \quad (4.2.4)$$

In particular, a distance d that belongs to \mathcal{D}_p induces a weaker topology than the p -Wasserstein distance. When necessary in some of the next results (or in the numerical experimentations), we will introduce specific distances which belong to \mathcal{D}_p .

Remark 4.2.2. The family of p -Wasserstein distances obviously provides examples of distances in \mathcal{D}_p . The Fortet-Mourier distance (see e.g [72, Chapter 6]), defined by

$$d_{\text{FM}}(\nu, \mu) = \sup \{ |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|; \text{ where } X \sim \mu, Y \sim \nu, \text{ and } \|h\|_{\text{Lip}} \leq 1, \|h\|_{\infty} \leq 1 \},$$

is also easily seen to be an element of \mathcal{D}_1 thanks to (4.2.3). In this chapter we shall work with the distances $d_{\text{CF}, p}$ and d_s introduced below in Section 4.2.5, which are trivially proved to sit in \mathcal{D}_1 (due to relation (4.2.3)).

For a given function f defined on \mathbb{R}_+ and with values in a Banach space E , we denote by $\delta f_{s,t}$ the increment between some positive s and t : $\delta f_{s,t} = f_t - f_s$.

4.2.2 Assumptions

Before we proceed to a specific statement of our estimator, let us describe the assumptions under which we shall work. We start by a standard hypothesis on the parameter set Θ , which is supposed to be a compact set.

(H₀) : The set Θ is compactly embedded in \mathbb{R}^q for a given $q \geq 1$.

Next we recall that our drift estimators rely on the invariant measure for the solution of equation (4.1.2). The existence and uniqueness of this invariant measure is usually obtained under some coercivity assumptions on the drift b . In the current chapter we will distinguish between two notions of coercivity, respectively named weak and strong and denoted by **(C_w)** and **(C_s)**. The weak assumption can be summarized as follows.

(C_w) : We have $b \in \mathcal{C}^{1,1}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$ and there exist constants $\alpha, \beta, C, L > 0$ and $r \in \mathbb{N}$ such that:

(i) For every $x, y \in \mathbb{R}^d$ and $\vartheta \in \Theta$ we have

$$\langle b_\vartheta(x) - b_\vartheta(y), x - y \rangle \leq \beta - \alpha|x - y|^2 \quad \text{and} \quad |b_\vartheta(x) - b_\vartheta(y)| \leq L|x - y| \quad (4.2.5)$$

(ii) For every $x \in \mathbb{R}^d$ and $\vartheta \in \Theta$ the following growth bound is satisfied:

$$|\partial_\vartheta b_\vartheta(x)| \leq C(1 + |x|^r). \quad (4.2.6)$$

The main part (4.2.5) of assumption **(C_w)** states that the coefficient b is inward looking, except maybe on a compact set which is a neighborhood of 0. We now state the strong assumption **(C_s)**, which specifies that b should be inward looking everywhere, and can be expressed as a particular case of **(C_w)**.

(C_s) : Assumption **(C_w)** holds with $\beta = 0$.

As mentioned above, Hypothesis **(C_w)** combined with the invertibility of σ (and therefore **(C_s)**) classically involves (see *e.g.* [36]) the existence of a unique stationary solution for the solution of the following equation for any $\vartheta \in \Theta$:

$$dY_t^\vartheta = b_\vartheta(Y_t^\vartheta) dt + \sigma dB_t, \quad t \in [0, T]. \quad (4.2.7)$$

Notice that the system (4.2.7) is identical to our original equation (4.1.2). However, let us notice that the fBm is unobserved. Moreover, the uniqueness of the stationary measure must be understood in a weak sense. Namely there exists a unique distribution \mathbb{P}^ϑ on $\mathcal{C}([0, \infty), \mathbb{R}^d)$ such that if (\bar{Y}_t^ϑ) denotes a process with distribution \mathbb{P}^ϑ , then $(\bar{Y}_t^\vartheta)_{t \geq 0}$ is a stationary solution to (4.2.7), *i.e.* shift-invariant (when one considers its canonical version). We denote by ν_ϑ the distribution of \bar{Y}^ϑ at any instant $t \geq 0$, that is

$$\nu_\vartheta = \mathcal{L}(\bar{Y}_0^\vartheta). \quad (4.2.8)$$

Remark 4.2.3. Note that in this non-Markovian setting, ν_ϑ is not exactly the invariant distribution. More precisely, owing to [36], one can embed (4.2.7) into an infinite-

dimensional Markovian structure which allows the construction of an adapted ergodic theory. An invariant distribution $\bar{\nu}_\vartheta$ is then defined on this enlarged structure. Without going into the details, one can just say that in this theory, the probability ν_ϑ can be retrieved as a marginal of the “true” invariant distribution. In the sequel, we will thus talk about *marginal invariant distribution* ν_ϑ .

Remark 4.2.4. As mentioned before, the invertibility assumption on σ combined with (\mathbf{C}_w) ensures uniqueness of the invariant distribution. However, even though this hypothesis is of first importance under (\mathbf{C}_w) (in order to use irreducibility-type arguments), it could be entirely removed under (\mathbf{C}_s) . Actually, in this case, the contraction assumption implies that two solutions of (4.2.7) driven by the same fBm but starting from different initial conditions come together at ∞ , *a.s.* and in L^2 , which classically involves uniqueness (see *e.g.* [15, Lemma 3(ii)] for details). This remark also holds for the Euler scheme (4.2.13) introduced in the next section but for sufficiently small step γ (see again [15, Lemma 3(ii)] for details).

As said previously, we shall obtain our drift estimators through the analysis of the marginal invariant distribution ν_ϑ defined by (4.2.8). If we want this strategy to be successful, it is natural to assume that ν_ϑ characterizes ϑ . We thus label this hypothesis as follows.

(\mathbf{I}_w) : For all $\vartheta \in \Theta$, we have $\nu_\vartheta = \nu_{\vartheta_0}$ iff $\vartheta = \vartheta_0$.

It is worth noticing that if d denotes a distance on $\mathcal{M}_1(\mathbb{R}^d)$, then one can recast (\mathbf{I}_w) as:

$$d(\nu_\vartheta, \nu_{\vartheta_0}) = 0 \quad \text{iff} \quad \vartheta = \vartheta_0. \quad (4.2.9)$$

We shall use this characterization in order to construct the estimator $\hat{\vartheta}$ (see (4.2.12) below). Also notice that (\mathbf{I}_w) refers to a “weak” identifiability condition, which will be resorted to in order to derive the consistency of our estimator $\hat{\vartheta}$. In contrast, the following “strong” identifiability assumption (\mathbf{I}_s) defined for a given distance d on $\mathcal{M}_1(\mathbb{R}^d)$ will be used to get rates of convergence.

(\mathbf{I}_s) There exists a constant $C > 0$ and a parameter $\varsigma \in (0, 1]$ such that $\forall \vartheta \in \Theta$,

$$d(\nu_\vartheta, \nu_{\vartheta_0}) \geq C|\vartheta - \vartheta_0|^\varsigma. \quad (4.2.10)$$

Remark 4.2.5. We will construct a class of equations, basically obtained as perturbations of Langevin type equations, for which our assumptions (\mathbf{I}_w) and (\mathbf{I}_s) are satisfied. See Section 4.6 below.

4.2.3 Statistical setting and construction of the estimator

We wish to construct an estimator based on discrete observations. In this context, the simplest situation (which will mostly prevail in the chapter) is to assume that the solution $(Y_t)_{t \geq 0}$ of (4.2.7) is discretely observed at some instants $\{t_k; 0 \leq k \leq n\}$, with $t_{k+1} - t_k = \kappa$ for a given time step $\kappa > 0$. Under (\mathbf{C}_w) , it can be shown (see Proposition 4.3.3 below) that

$$\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}} \xrightarrow{n \rightarrow +\infty} \nu_{\vartheta_0} \quad a.s., \quad (4.2.11)$$

where \implies stands for the weak convergence of probability measures in \mathbb{R}^d . With this convergence in mind, the heart of our estimation method is then the following observation: under the identifiability assumption (\mathbf{I}_w) , the most natural way to estimate ϑ_0 is to consider

$$\hat{\vartheta} = \operatorname{argmin}_{\vartheta \in \Theta} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \nu_{\vartheta} \right), \quad (4.2.12)$$

where d is a given (arbitrary) distance on $\mathcal{M}_1(\mathbb{R}^d)$. However, in spite of the fact that our formula (4.2.12) is simple enough, it is also easily understood that ν_{ϑ} is far from being explicitly known (except in some very particular cases such as the Ornstein-Uhlenbeck process). In this chapter, we propose to circumvent this difficulty by considering some estimators based on numerical approximations of ν_{ϑ} .

Specifically, the numerical approximations we will resort to are built through an Euler-type discretization of the stochastic process Y^{ϑ} solution to (4.2.7). Namely, let $(s_k)_{k \geq 0}$ be an increasing sequence of numbers such that $s_0 = 0$, and $\lim_{k \rightarrow \infty} s_k = +\infty$. The Euler-Maruyama scheme Z^{ϑ} is then recursively defined by $Z_0^{\vartheta} = z_0 \in \mathbb{R}^d$ and:

$$\text{For all } k \geq 0, \quad \delta Z_{s_k, s_{k+1}}^{\vartheta} = (s_{k+1} - s_k) b_{\vartheta}(Z_{s_k}^{\vartheta}) + \sigma \delta \tilde{B}_{s_k, s_{k+1}}, \quad (4.2.13)$$

where \tilde{B} is a (simulated) m -dimensional fractional Brownian motion which is a priori different from the driving process B in equation (4.2.7) (since B is unobserved). When $s_k = k\gamma$ for a given $\gamma > 0$, we say that the Euler scheme is a *constant step* sequence and we denote it by $Z^{\vartheta, \gamma}$. When $\gamma_k = s_k - s_{k-1}$ is a non-increasing sequence such that $\gamma_k \rightarrow 0$ as $k \rightarrow +\infty$, the Euler scheme will be called *decreasing step* Euler scheme. We will work with these two types of schemes in the sequel.

Remark 4.2.6. In practice it is natural to set $Z_0^{\vartheta} = Y_0$ where Y_0 is the first observation of the process $(Y_t)_{t \geq 0}$. Let us also remark that in the sequel, for notational sake, one usually denotes by B the fBm related to the Euler scheme Z^{ϑ} . However, the reader has to keep in mind that the fact that the fBMs in (4.2.7) and in (4.2.13) are different. This certainly prevents us from any pathwise comparison between the observed process and the simulated one.

Remark 4.2.7. We refer to Section 4.7 for background on the simulation of the increments of the fBm.

Let us now give an explicit expression for the estimator we are considering in this chapter. We will focus on the constant step setting in (4.2.13) for sake of simplicity. Observe that under (\mathbf{C}_w) , it can be shown (see Proposition 4.3.4 below) that

$$\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta, \gamma}} \xrightarrow{n \rightarrow +\infty} \nu_{\vartheta}^{\gamma} \quad a.s. \quad (4.2.14)$$

where ν_{ϑ}^{γ} denotes the unique marginal invariant distribution of the Euler scheme Z^{ϑ} . By marginal, we mean again that $Z^{\vartheta, \gamma}$ can be endowed with a Feller infinite-dimensional Markov structure which admits a unique invariant distribution under (\mathbf{C}_w) (see [71] for details). The first marginal of this invariant distribution is ν_{ϑ}^{γ} . Similarly to what we

proposed in (4.2.12), such a result suggests to define our estimator for ϑ as

$$\hat{\vartheta}_{N,n,\gamma} = \operatorname{argmin}_{\vartheta \in \Theta} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta,\gamma}} \right), \quad (4.2.15)$$

where d is again a distance on $\mathcal{M}_1(\mathbb{R}^d)$. Note that in the decreasing step case analogous constructions may be carried out, and will be introduced later. Let us also remark that relation (4.2.15) only involves one Euler scheme path, which is relevant for numerical implementations.

We are now in a position to state our main results. We divide the presentation in two parts. In the next section, we focus on strong consistency results related to the family $\{\hat{\vartheta}_{N,n,\gamma}; N \geq 1, n \geq 1, \gamma > 0\}$ defined by (4.2.15), as well as its decreasing step counterpart. Then Section 4.2.5 is dedicated to the rate of convergence of the estimator $\hat{\vartheta}_{N,n,\gamma}$. In particular, this second part will involve concentration results related to the SDE and to its Euler discretization.

4.2.4 Main consistency results

We begin with a first result involving the weak assumption (\mathbf{C}_w) , which requires to discretize the set Θ in the following sense. According to our hypothesis (\mathbf{H}_0) , the set Θ is compact in \mathbb{R}^q . Therefore the Borel-Lebesgue property gives us the existence, for every $\varepsilon > 0$, of $M_\varepsilon \in \mathbb{N}$ and $(\vartheta_i^{(\varepsilon)})_{1 \leq i \leq M_\varepsilon} \in \Theta^{M_\varepsilon}$ such that $\Theta \subset \bigcup_{i=1}^{M_\varepsilon} B(\vartheta_i^{(\varepsilon)}, \varepsilon)$. Thanks to this property, we define the following discretization for all $\varepsilon > 0$ and $\vartheta \in \Theta$:

$$\vartheta^{(\varepsilon)} := \operatorname{argmin}_{\vartheta' \in \{\vartheta_i^{(\varepsilon)}\}} |\vartheta' - \vartheta|. \quad (4.2.16)$$

With this notation in hand, we can now state our first consistency theorem.

Theorem 4.2.1. *Assume (\mathbf{H}_0) , (\mathbf{C}_w) and (\mathbf{I}_w) . Let p be a strictly positive real number and consider a distance d on $\mathcal{M}_1(\mathbb{R}^d)$ which belongs to \mathcal{D}_p (recall that \mathcal{D}_p is defined by (4.2.4)). We consider the family $\{\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)}; N \geq 1, n \geq 1, \gamma > 0, \varepsilon > 0\}$ defined by*

$$\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)} = \operatorname{argmin}_{\vartheta \in \Theta^{(\varepsilon)}} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{s_k}^{\vartheta,\gamma}} \right), \quad N, n \in \mathbb{N}, \gamma > 0, \varepsilon > 0 \quad (4.2.17)$$

where $\Theta^{(\varepsilon)} := \{\vartheta_i^{(\varepsilon)}; 1 \leq i \leq M_\varepsilon\}$. Then $\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)}$ is a strong consistent estimator of ϑ_0 . Specifically, we have

$$\lim_{\varepsilon \rightarrow 0} \lim_{\gamma \rightarrow 0} \lim_{N, n \rightarrow +\infty} \hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)} = \vartheta_0 \quad a.s.$$

Let us remark that the discretization of Θ given by (4.2.16) is needed to get strong consistency, due to the fact that under (\mathbf{C}_w) we loose uniformity with respect to ϑ in some of our convergence results. For instance, (\mathbf{C}_w) only warranties the simple convergence of $d(\nu_\vartheta^\gamma, \nu_\vartheta)$ to 0 as $\gamma \rightarrow 0$ (see Proposition 4.3.4). The proof of Theorem 4.2.1 is achieved in Section 4.4.3.

We now turn to our main estimator defined in (4.2.15). The proof of the theorem below is detailed in Section 4.4.3.

Theorem 4.2.2. *Assume (\mathbf{H}_0) , (\mathbf{C}_s) and (\mathbf{I}_w) . As in Theorem 4.2.1, let p be a strictly positive real number and consider a distance d on $\mathcal{M}_1(\mathbb{R}^d)$ which belongs to \mathcal{D}_p . Then the family $\{\hat{\vartheta}_{N,n,\gamma}; N \geq 1, n \geq 1, \gamma > 0\}$ defined by (4.2.15) is a strong consistent estimator of ϑ_0 in the following sense:*

$$\lim_{\gamma \rightarrow 0} \lim_{N,n \rightarrow +\infty} \hat{\vartheta}_{N,n,\gamma} = \vartheta_0 \quad a.s.$$

We close this section with a result concerning the approximation of invariant measures by an Euler scheme with decreasing time step. Namely we consider an approximation scheme denoted also by Z^ϑ , which is defined similarly to $Z^{\vartheta,\gamma}$ in (4.2.13) except for the fact that the sequence $(s_k)_{k \geq 0}$ satisfies:

$$s_{k+1} - s_k = \gamma_{k+1}, \quad k \geq 0, \quad (4.2.18)$$

where $(\gamma_k)_{k \geq 1}$ is a non-increasing sequence of positive numbers such that

$$\sum_{k \geq 1} \gamma_k = +\infty \quad \text{and} \quad \lim_{k \rightarrow +\infty} \gamma_k = 0. \quad (4.2.19)$$

The convergence theorem we obtain in the decreasing step case is the following. As for Theorem 4.2.2, it is achieved under the strong coercivity assumption (\mathbf{C}_s) . Its proof is developed in Section 4.4.3.

Theorem 4.2.3. *Assume (\mathbf{H}_0) , (\mathbf{C}_s) and (\mathbf{I}_w) . Let $p \geq 2$ and consider a distance d on $\mathcal{M}_1(\mathbb{R}^d)$ which belongs to \mathcal{D}_p . Let $(\gamma_k)_{k \geq 1}$ be a non-increasing sequence of positive numbers satisfying condition (4.2.19) and the technical condition (4.4.3). Denote by Z^ϑ the related Euler scheme given by (4.2.13). We consider an estimator $(\hat{\vartheta}_{N,n})_{N,n}$ defined as*

$$\hat{\vartheta}_{N,n} = \operatorname{argmin}_{\vartheta \in \Theta} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^\vartheta} \right), \quad N \in \mathbb{N}. \quad (4.2.20)$$

Then, $(\hat{\vartheta}_{N,n})_{N,n}$ is a strong consistent estimator of ϑ_0 , namely:

$$\lim_{N,n \rightarrow +\infty} \hat{\vartheta}_{N,n} = \vartheta_0 \quad a.s.$$

Remark 4.2.8. The technical condition (4.4.3) ($\sum_{k \geq 1} \gamma_k^{p'H+1} s_k^{-1} < +\infty$ for a given $p' \geq p$) is true in a very large setting. For instance, it can be checked that this is satisfied for any polynomial step sequence : $\gamma_k = \gamma k^{-\rho}$ with $\rho \in (0, 1]$ and $\gamma \in \mathbb{R}_+^*$, but also for less decreasing sequences such as $\gamma_k = \gamma (\log k)^{-1}$. However, this is not true in full generality (the condition does not hold when $\gamma_k = (\log(\log k))^{-1}$ for instance).

4.2.5 Rate of convergence

Under our strong identifiability condition (\mathbf{I}_s) , we will be able to get a rate of convergence for some of our estimators. In order to carry out this task, we shall assume that condition (\mathbf{I}_s) is verified for some specific distances on probability measures called respectively $d_{CF,p}$ and d_s . These distances are defined in the following way:

- (i) Let X and Y be \mathbb{R}^d -valued random variables and $p > (\frac{d}{2} \vee 1)$. We consider the integrable kernel $g_p(\xi) := c_p(1 + |\xi|^2)^{-p}$ where $c_p := (\int_{\mathbb{R}^d} (1 + |\xi|^2)^{-p} d\xi)^{-1}$. Then the $d_{CF,p}$ distance between $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ is defined by:

$$d_{CF,p}(\mathcal{L}(X), \mathcal{L}(Y)) := \left(\int_{\mathbb{R}^d} \left(\mathbb{E}[e^{i\langle \xi, X \rangle}] - \mathbb{E}[e^{i\langle \xi, Y \rangle}] \right)^2 g_p(\xi) d\xi \right)^{1/2}. \quad (4.2.21)$$

- (ii) Let $\{f_i ; i \geq 1\}$ be a family of \mathcal{C}_b^1 , supposed to be dense in the space \mathcal{C}_b^0 of continuous and bounded functions and decreasing to 0 at ∞ . Consider two probability measures ν and μ in $\mathcal{M}_1(\mathbb{R}^d)$. Then the distance d_s between ν and μ is defined as:

$$d_s(\nu, \mu) := \sum_{i=0}^{+\infty} 2^{-i} (|\nu(f_i) - \mu(f_i)| \wedge 1). \quad (4.2.22)$$

Remark 4.2.9. It is readily checked that both $d_{CF,p}$ and d_s metrize the convergence in law (their induced topology are in fact exactly the one induced by the convergence in law). The distance $d_{CF,p}$ is technically convenient for our purpose and close in spirit to the smooth Wasserstein distance invoked in the Stein method literature (see e.g [2]). The distance d_s is called weak- \star distance in [72] and also used in [76] for filtering problems. Also notice that both $d_{CF,p}$ and d_s are elements of \mathcal{D}_1 where \mathcal{D}_1 is defined by (4.2.4).

With the distances $d_{CF,p}$ and d_s in hand, our main result about rates of convergence is the following:

Theorem 4.2.4. *Assume (\mathbf{H}_0) , (\mathbf{C}_s) and (\mathbf{I}_s) hold true, where (4.2.10) in hypothesis (\mathbf{I}_s) is considered for $d = d_s$ or $d = d_{CF,p}$ with $p > (q + d)/2$ and $\varsigma = 2/q$ for a given $q \geq 2$. Let $\hat{\vartheta}_{N,n,\gamma}$ be the estimator given by (4.2.15). Then, we get the following rate of convergence: there exists $C_q > 0$ such that*

$$\mathbb{E} \left[|\hat{\vartheta}_{N,n,\gamma} - \vartheta_0|^2 \right] \leq C_q \left(n^{-\frac{q}{2}(2-(2H \vee 1))} + \gamma^{qH} + T^{-\tilde{\eta}} \right) \quad (4.2.23)$$

with $\tilde{\eta} := \frac{q^2}{2(q+d)}(2 - (2H \vee 1))$ and $T := N\gamma$.

Remark 4.2.10. This non-asymptotical bound theoretically enables to calibrate the “free parameters” γ and N in terms of the number of observations n , which is fixed by the statistical setting. For instance, when $\varsigma = 1$ (i.e. when $q = 2$), the first term is of order $n^{-(2-(2H \vee 1))}$ and hence, in order to preserve this rate order, we have to fix $\gamma \leq n^{-\frac{1-(H \vee \frac{1}{2})}{H}}$ and $N \geq n^{\frac{4+2d}{4}} \gamma^{-1}$. More precisely, for these choices of parameters, the quadratic error induced by this estimator (when $\varsigma = 1$) is of order $n^{-\frac{1}{2}}$ if $H < 1/2$ and $n^{H-\frac{1}{2}}$ if $H > 1/2$. The constant ς , which appears in Assumption (\mathbf{I}_s) , comes from the fact that, the bounds are first established on the distances between the invariant distributions $\nu_{\hat{\vartheta}_{N,n,\gamma}}$ and ν_{ϑ_0} . Nevertheless, except some particular settings such as the Ornstein-Uhlenbeck process, this exponent ς is unfortunately difficult to compute in some general settings. Finally, let us remark that L^p -bounds can be easily deduced from the proof for any $p \geq 2$. However, since they do not modify significantly the results, we chose here to only state the quadratic one.

4.3 Preliminary Results

In this section we label some basic results about equation (4.2.7) and its invariant measure for further use. We first recall some ergodic properties of stochastic differential equations driven by a fBm, then we study the continuity of the invariant measure ν_ϑ with respect to the parameter ϑ . Under the strong coercivity assumption (\mathbf{C}_s) , we quantify the distance between the empirical measures respectively related to the process Y^ϑ and its Euler approximation Z^ϑ . Eventually we give some convergence results for the quantities involved in the right hand side of (4.2.17).

4.3.1 Ergodic properties of the SDE and of the Euler scheme

In this section we review several ergodic properties for equation (4.2.7). These properties are at the heart of our estimation procedure.

Convergence of $\mathcal{L}(Y_t)$

We start by giving the basic convergence in law towards the invariant distribution for our processes Y^ϑ .

Proposition 4.3.1. *Assume (\mathbf{H}_0) and (\mathbf{C}_w) and consider the family of processes $\{Y^\vartheta; \vartheta \in \Theta\}$ defined by (4.2.7). Then the following properties hold true:*

(i) *Existence and uniqueness hold for the invariant distribution related to the dynamical system (4.2.7). Furthermore, having in mind the notations introduced in (4.2.8), for all $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ independent of $\vartheta \in \Theta$ such that*

$$d_{\text{TV}}\left(\mathcal{L}\left(Y_t^\vartheta\right), \nu_\vartheta\right) \leq C_\varepsilon t^{-(\alpha_H - \varepsilon)}, \quad (4.3.1)$$

with an exponent α_H given by

$$\alpha_H = \begin{cases} \frac{1}{8} & \text{if } H \in (\frac{1}{4}, 1) \setminus \{\frac{1}{2}\} \\ H(1 - 2H) & \text{if } H \in (0, \frac{1}{4}). \end{cases}$$

(ii) *For any $p > 0$ and for any distance $d \in \mathcal{D}_p$, the following upper bound holds uniformly in $\vartheta \in \Theta$:*

$$d\left(\mathcal{L}(Y_t^\vartheta), \nu_\vartheta\right) \leq C t^{-\frac{\alpha_H}{4p}}, \quad (4.3.2)$$

for a strictly positive constant C . In particular, for any $p > 0$, we have

$$\sup_{t \geq 0} \mathbb{E}[|Y_t|^p] < +\infty. \quad (4.3.3)$$

Proof. We prove the two statements of our proposition separately.

Proof of item (i). The only difference between our claim and [36, Theorems 1.2 and 1.3] is the uniformity with respect to $\vartheta \in \Theta$ in the convergence in total variation result. However, following carefully the proof of [36], it can be shown that the constants therein do not depend on ϑ if Hypothesis (\mathbf{C}_w) is satisfied. Therefore the constant C_ε in (4.3.1) is uniform in ϑ .

Proof of item (ii). Relation (4.3.3) is proved in Proposition 4.A.1 of the Appendix. In order to prove (4.3.2), consider a couple (X_1, X_2) of random variables such that $X_1 \sim Y_t^\vartheta$ and X_2 has distribution ν_ϑ . By Cauchy-Schwarz inequality,

$$\begin{aligned} \mathbb{E}[|X_1 - X_2|^p]^{\frac{1}{p}} &= \mathbb{E}[|X_1 - X_2|^p \mathbf{1}_{(X_1 \neq X_2)}]^{\frac{1}{p}} \\ &\leq \left(\mathbb{E}[|X_1|^{2p}]^{\frac{1}{2p}} + \mathbb{E}[|X_2|^{2p}]^{\frac{1}{2p}} \right) \mathbb{P}(X_1 \neq X_2)^{\frac{1}{2p}}. \end{aligned} \quad (4.3.4)$$

We now bound separately the two terms on the right hand side of (4.3.4). If we denote by $\mathcal{C}(Y_t^\vartheta, \nu_\vartheta)$ the set of couplings between $\mathcal{L}(Y_t^\vartheta)$ and ν_ϑ (defined as in (4.2.1)), then we have

$$d_{\text{TV}} \left(\mathcal{L}(Y_t^\vartheta), \nu_\vartheta \right) = \inf \left\{ \mathbb{P}(\tilde{X}_1 \neq \tilde{X}_2); (\tilde{X}_1, \tilde{X}_2) \in \mathcal{C}(Y_t^\vartheta, \nu_\vartheta) \right\}.$$

Therefore, owing to (4.3.1) we can choose a coupling $(X_1, X_2) \in \mathcal{C}(Y_t^\vartheta, \nu_\vartheta)$ and a constant $C > 0$ such that

$$\mathbb{P}(X_1 \neq X_2) \leq Ct^{-(\alpha_H - \varepsilon)}. \quad (4.3.5)$$

In addition, according to (4.3.3) and a uniform integrability argument, we easily get the following inequality for the coupling (X_1, X_2) we have chosen in (4.3.5):

$$\mathbb{E}[|X_1|^{2p}]^{\frac{1}{2p}} + \mathbb{E}[|X_2|^{2p}]^{\frac{1}{2p}} \leq 2 \sup_{\vartheta \in \Theta} \sup_{t \geq 0} \mathbb{E}[|Y_t^\vartheta|^{2p}]^{\frac{1}{2p}} < +\infty. \quad (4.3.6)$$

We plug (4.3.5) and (4.3.6), applied to $\varepsilon = \alpha_H/2$, into (4.3.4). Going back to the definition (4.2.2) of the distance \mathcal{W}_p , we obtain that there exists a strictly positive constant C such that for any $t \geq 0$ and any $\vartheta \in \Theta$ we have

$$\mathcal{W}_p(\mathcal{L}(Y_t^\vartheta), \nu_\vartheta) \leq Ct^{-\frac{\alpha_H}{4p}}.$$

The result (4.3.2) follows. \square

Next we observe that whenever (\mathbf{C}_s) is fulfilled, the polynomial convergence in (4.3.2) can be replaced by an exponential rate. This is summarized in the following proposition.

Proposition 4.3.2. *Let $\{Y^\vartheta; \vartheta \in \Theta\}$ be the family of processes defined by (4.2.7). Suppose that Hypothesis (\mathbf{H}_0) and Hypothesis (\mathbf{C}_s) are met. Let d be a distance in \mathcal{D}_p . Then, we have*

$$d(\mathcal{L}(Y_t^\vartheta), \nu_\vartheta) \leq c_1 e^{-c_2 t},$$

with $c_2 = \alpha/2$ where α is the constant featured in equation (4.2.5), and where $c_1 = c_1(H, \alpha)$.

Proof. Let \bar{Y}^ϑ be the stationary solution of equation (4.2.7). One can easily show, by means of the same arguments as in [30], that

$$\mathbb{E} \left[|Y_t^\vartheta - \bar{Y}_t^\vartheta|^p \right]^{1/p} \leq c_1 e^{-c_2 t}. \quad (4.3.7)$$

The result follows trivially. \square

Ergodic Theorems for the SDE

We now summarize the limit theorems obtained for equation (4.2.7) which will be relevant for our purposes, with a special emphasis on the occupation measure $\delta_{Y_s^\vartheta}$.

Proposition 4.3.3. *Let Y^ϑ be the unique solution of (4.2.7) obtained for a parameter $\vartheta \in \Theta$. Assume (\mathbf{H}_0) and (\mathbf{C}_w) hold true and let ν_ϑ be the measure defined by (4.2.8). Then for all $\vartheta \in \Theta$, for any $p > 0$ and for any distance $d \in \mathcal{D}_p$, we have*

(i) *The distance between ν_ϑ and the normalized occupation measure of Y^ϑ converges to 0 as $t \rightarrow \infty$. That is*

$$\lim_{t \rightarrow \infty} d \left(\frac{1}{t} \int_0^t \delta_{Y_s^\vartheta} ds, \nu_\vartheta \right) = 0. \quad (4.3.8)$$

In particular, we have an almost sure uniform bound for the p -th powers of Y_s^ϑ :

$$\sup_{\vartheta \in \Theta} \sup_{t \geq 1} \frac{1}{t} \int_0^t |Y_s^\vartheta|^p ds < \infty \quad \text{a.s.} \quad (4.3.9)$$

(ii) *Some discrete versions of (4.3.8) and (4.3.9) are also available. Specifically, let $\eta > 0$ and set $t_k = k\eta$ for $k \geq 0$. Then for any $\vartheta \in \Theta$ we have*

$$\lim_{n \rightarrow \infty} d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}^\vartheta}, \nu_\vartheta \right) = 0.$$

In particular, the following uniform bound holds true:

$$\sup_{\vartheta \in \Theta} \sup_{n \geq 1} \frac{1}{n} \sum_{k=0}^{n-1} |Y_{t_k}^\vartheta|^p < \infty \quad \text{a.s.} \quad (4.3.10)$$

Proof. Relations (4.3.9) and (4.3.10) are proved in Proposition 4.A.1. As far as the identification of the limit is concerned, the proof follows the lines of [15] and is detailed in Section 4.A.2. \square

Ergodic Theorems for the Euler scheme

Recall that we denote by $(Z_{k\gamma}^{\vartheta,\gamma})_{k \geq 0}$ the Euler scheme with step γ related to $(Y_t^\vartheta)_{t \geq 0}$, as defined in (4.2.13). This section focuses on the asymptotic behavior of $Z_{k\gamma}^{\vartheta,\gamma}$ as $k \rightarrow \infty$ and $\gamma \rightarrow 0$.

Proposition 4.3.4. *Let Y^ϑ be the unique solution of (4.2.7) and consider the Euler scheme $(Z_{k\gamma}^{\vartheta,\gamma})_{k \geq 0}$ with step γ related to $(Y_t^\vartheta)_{t \geq 0}$. Assume (\mathbf{H}_0) and (\mathbf{C}_w) hold true. Then for all $\vartheta \in \Theta$, for any $p > 0$ and for any distance $d \in \mathcal{D}_p$, we have*

(i) *There exists $\gamma_0 > 0$ and a unique family of measures $(\nu_\vartheta^\gamma)_{\gamma \leq \gamma_0}$ such that for all $\vartheta \in \Theta$ and $\gamma \in (0, \gamma_0]$, we have*

$$\lim_{N \rightarrow \infty} d \left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta,\gamma}}, \nu_\vartheta^\gamma \right) = 0.$$

In particular, we get the following uniform bound for the p -th powers of $Z_{k\gamma}^{\vartheta,\gamma}$:

$$\sup_{\vartheta \in \Theta} \sup_{N \geq 1} \frac{1}{N} \sum_{k=0}^{N-1} |Z_{k\gamma}^{\vartheta,\gamma}|^p < \infty \quad a.s. \quad (4.3.11)$$

(ii) The invariant measure ν_ϑ^γ for the Euler scheme converges to the invariant measure ν_ϑ of Y^ϑ as the mesh of the partition goes to 0:

$$\lim_{\gamma \rightarrow 0} d(\nu_\vartheta^\gamma, \nu_\vartheta) = 0.$$

Proof. The weak convergence of $\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta,\gamma}}$ to ν_ϑ^γ as $n \rightarrow +\infty$, as well as the convergence of ν_ϑ^γ to ν_ϑ as $\gamma \rightarrow 0$ are consequences of [16, Theorem 1]. The extension to distances d dominated by \mathcal{W}^p follows from Proposition 4.A.2. More precisely for (ii), we can deduce from Proposition 4.A.2 and from Fatou's lemma that for any $M > 0$, for any $\gamma \in (0, \gamma_0]$,

$$\nu_\vartheta^\gamma(|x|^p \wedge M) \leq \liminf_{N \rightarrow +\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}[|Z_{k\gamma}^{\vartheta,\gamma}|^p] \leq C$$

where C is a positive constant independent of M and γ . Hence, taking limits as M goes to ∞ , this yields

$$\sup_{\gamma \in (0, \gamma_0]} \nu_\vartheta^\gamma(|x|^p) < +\infty,$$

for any $p > 0$. Then one can conclude as in the proof of Proposition 4.3.1 (see relation (4.3.4)), where the distance $d_{\text{TV}}(\nu_\vartheta^\gamma, \nu_\vartheta)$ is upper bounded thanks to [16, Theorem 1]. \square

4.3.2 Continuity of $\vartheta \mapsto d(\nu_\vartheta, \nu_{\vartheta_0})$

The convergence of our estimator $\hat{\vartheta}_{N,n,\gamma}$ defined by (4.2.15) depends crucially on continuity properties of the family $\{\nu_\vartheta; \vartheta \in \Theta\}$. To this aim, we first prove a basic result on the continuity of the map $\vartheta \mapsto Y_t^\vartheta$.

Proposition 4.3.5. *Let ϑ_1 and ϑ_2 be elements of Θ , and consider the respective solutions $(Y_t^{\vartheta_1})_{t \geq 0}$ and $(Y_t^{\vartheta_2})_{t \geq 0}$ of equation (4.2.7). Assume hypothesis (\mathbf{H}_0) and (\mathbf{C}_w) are satisfied. Then for any $p > 0$ and $T > 0$, there exists $C_{T,p} > 0$ independent of ϑ_1 and ϑ_2 such that*

$$\|Y_t^{\vartheta_1} - Y_t^{\vartheta_2}\|_{L^p(\Omega)} \leq C_{T,p} |\vartheta_1 - \vartheta_2|. \quad (4.3.12)$$

Proof. By monotonicity of the norms in $L^p(\Omega)$, it is enough to consider the case $p \geq 2$. Furthermore, it is readily seen from (4.2.7) that we have

$$Y_t^{\vartheta_1} - Y_t^{\vartheta_2} = \int_0^t (b_{\vartheta_1}(Y_s^{\vartheta_1}) - b_{\vartheta_2}(Y_s^{\vartheta_2})) ds. \quad (4.3.13)$$

Starting from (4.3.13), we easily get the following identity for the square of $Y_t^{\vartheta_1} - Y_t^{\vartheta_2}$:

$$\frac{d}{dt} |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 = 2 \langle Y_t^{\vartheta_1} - Y_t^{\vartheta_2}, b_{\vartheta_1}(Y_t^{\vartheta_1}) - b_{\vartheta_2}(Y_t^{\vartheta_2}) \rangle. \quad (4.3.14)$$

We now invoke the fact that b is Lipschitz continuous under (\mathbf{C}_w) plus inequality (4.2.6) on $\partial_\vartheta b_\vartheta$ in order to get

$$\begin{aligned} \frac{d}{dt} |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 &= 2\langle Y_t^{\vartheta_1} - Y_t^{\vartheta_2}, b_{\vartheta_1}(Y_t^{\vartheta_1}) - b_{\vartheta_1}(Y_t^{\vartheta_2}) \rangle + 2\langle Y_t^{\vartheta_1} - Y_t^{\vartheta_2}, b_{\vartheta_1}(Y_t^{\vartheta_2}) - b_{\vartheta_2}(Y_t^{\vartheta_2}) \rangle \\ &\leq c_1 |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 + 2c_2 |\vartheta_1 - \vartheta_2| \left(1 + |Y_t^{\vartheta_2}|^r\right) |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|, \end{aligned}$$

where c_1 and c_2 are two strictly positive constants. Now apply the elementary inequality $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$ with $a = |\vartheta_1 - \vartheta_2|(1 + |Y_t^{\vartheta_2}|^r)$ and $b = |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|$. We deduce the existence of a constant $c > 0$ such that

$$\frac{d}{dt} |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 \leq c \left(|Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 + |\vartheta_1 - \vartheta_2|^2 \left(1 + |Y_t^{\vartheta_2}|^{2r}\right) \right). \quad (4.3.15)$$

With relation (4.3.15) in hand, a standard application of Gronwall's Lemma yields

$$|Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 \leq c |\vartheta_1 - \vartheta_2|^2 \int_0^t e^{c(t-s)} \left(1 + |Y_s^{\vartheta_2}|^{2r}\right) ds. \quad (4.3.16)$$

Let us now get some information about a generic p -th power of $Y_t^{\vartheta_1} - Y_t^{\vartheta_2}$ for $p \geq 2$. To this aim, we resort to Jensen's inequality in relation (4.3.16). This gives the existence of a constant $c(T, p)$ such that for any $t \in [0, T]$ we have

$$|Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^p \leq c(T, p) |\vartheta_1 - \vartheta_2|^p \int_0^t \left(1 + |Y_s^{\vartheta_2}|^{2r}\right)^{\frac{p}{2}} ds.$$

Taking the expectation, we finally get

$$\mathbb{E} \left[|Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^p \right]^{1/p} \leq \tilde{c}(T, p) |\vartheta_1 - \vartheta_2| \left(\int_0^t \mathbb{E} \left[1 + |Y_s^{\vartheta_2}|^{pr} \right] ds \right)^{1/p},$$

where $\tilde{c}(T, p)$ is another finite constant. Hence our result (4.3.12) follows from the bound (4.3.3). \square

We now state the announced continuity property for the family $\{\nu_\vartheta; \vartheta \in \Theta\}$.

Proposition 4.3.6. *Let $\{Y^\vartheta; \vartheta \in \Theta\}$ be the family of processes defined by (4.2.7). Assume (\mathbf{H}_0) and (\mathbf{C}_w) hold true and consider the family $\{\nu_\vartheta; \vartheta \in \Theta\}$ of invariant measures given by Proposition 4.3.1. Let $p > 0$ and pick any distance $d \in \mathcal{D}_p$, where we recall that \mathcal{D}_p is defined by (4.2.4). Then the map $\vartheta \mapsto d(\nu_\vartheta, \nu_{\vartheta_0})$ is continuous on Θ .*

Proof. Owing to the very definition (4.2.4) of \mathcal{D}_p , it is enough to prove the result for $d = \mathcal{W}_p$ and for an arbitrary $p \geq 1$. Next we apply the triangle inequality and the fact that \mathcal{W}_p is defined in (4.2.2) by an infimum over all couplings. This yields the following inequality, valid for any $t \geq 0$:

$$\mathcal{W}_p(\nu_{\vartheta_1}, \nu_{\vartheta_2}) \leq 2 \sup_{\vartheta \in \Theta} \mathcal{W}_p(\mathcal{L}(Y_t^\vartheta), \nu_\vartheta) + \|Y_t^{\vartheta_1} - Y_t^{\vartheta_2}\|_{L^p(\Omega)}. \quad (4.3.17)$$

We now bound the two terms in the right hand side of (4.3.17). In order to handle the term $\mathcal{W}_p(\mathcal{L}(Y_t^\vartheta), \nu_\vartheta)$, we consider a small parameter $\varepsilon > 0$. By Proposition 4.3.1(ii),

there exists t_0 large enough such that

$$2 \sup_{\vartheta \in \Theta} \mathcal{W}_p(\mathcal{L}(Y_{t_0}^\vartheta), \nu_\vartheta) \leq \frac{\varepsilon}{2}. \quad (4.3.18)$$

We will fix this value of t_0 in the right hand side of (4.3.17). Then the difference $Y_{t_0}^{\vartheta_1} - Y_{t_0}^{\vartheta_2}$ is handled thanks to Proposition 4.3.5. Namely consider $\delta > 0$ such that (with the notations of Proposition 4.3.5) we have $C_{t_0} \delta \leq \frac{\varepsilon}{2}$. We get that for all $(\vartheta_1, \vartheta_2) \in \Theta^2$ such that $|\vartheta_1 - \vartheta_2| \leq \delta$, we have

$$\|Y_t^{\vartheta_1} - Y_t^{\vartheta_2}\|_{L^p(\Omega)} \leq \frac{\varepsilon}{2} \quad (4.3.19)$$

We conclude by plugging (4.3.18) and (4.3.19) into (4.3.17). This yields

$$\mathcal{W}_p(\nu_{\vartheta_1}, \nu_{\vartheta_2}) \leq \varepsilon, \quad (4.3.20)$$

for all $(\vartheta_1, \vartheta_2) \in \Theta^2$ such that $|\vartheta_1 - \vartheta_2| \leq \delta$. The continuity of $\vartheta \mapsto \mathcal{W}_p(\nu_\vartheta, \nu_{\vartheta_0})$ on Θ follows. \square

4.3.3 Further controls under (\mathbf{C}_s)

Up to now we have derived properties of the system (4.2.7) under the weak coercive assumption (\mathbf{C}_w) . In this section, we focus on possible additional bounds one can obtain under the stronger hypothesis (\mathbf{C}_s) . We will first see how (\mathbf{C}_s) guarantees a uniform control on the distance between the Euler scheme and the SDE, for a general decreasing sequence of time steps. Then we will show that (\mathbf{C}_s) ensures a some additional uniform continuity in ϑ for the occupation measures of Y^ϑ .

We consider here Euler type approximations in continuous time, with time steps γ_n satisfying (4.2.19). In order to define this Euler approximation $(Z_t^\vartheta)_{t \geq 0}$, we set $s_0 = 0$ and $s_n = \sum_{i=1}^n \gamma_i$ for all $n \geq 1$. Then for any $n \geq 0$, the process $(Z_t^\vartheta)_{t \geq 0}$ is given recursively by

$$Z_{s_n+t}^\vartheta = Z_{s_n}^\vartheta + t b_\vartheta(Z_{s_n}^\vartheta) + \sigma(B_{s_n+t} - B_{s_n}), \quad t \in [0, s_{n+1} - s_n]. \quad (4.3.21)$$

Notice that the fractional Brownian motion B in (4.3.21) is the same as the fBm driving equation (4.2.7). The control we get on Z^ϑ is summarized in the following proposition.

Proposition 4.3.7. *Let Y^ϑ be the solution of equation (4.2.7), and consider the continuous-time Euler scheme $(Z_t^\vartheta)_{t \geq 0}$ with a time steps sequence $(\gamma_n)_{n \geq 1}$ defined by (4.3.21). We assume that (\mathbf{C}_s) holds. Then the following assertions hold true.*

(i) *For any $p \geq 2$, there exist some positive constants ρ and C such that for any $n \geq 1$ we have*

$$|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p \leq e^{-\rho s_n} |Y_0^\vartheta - Z_0^\vartheta|^p + C \sum_{k=0}^{n-1} \phi_{k,p}(Z_{s_k}^\vartheta) e^{-\rho(s_n - s_{k+1})}, \quad (4.3.22)$$

where the function $\phi_{k,p}$ is defined, for any $k \geq 0$, by

$$\phi_{k,p}(z) = \gamma_{k+1}^{p+1} |b_\vartheta(z)|^p + \int_0^{\gamma_{k+1}} |B_{s_k+t} - B_{s_k}|^p dt. \quad (4.3.23)$$

(ii) Assume in addition that $\gamma_n \rightarrow 0$ as $n \rightarrow +\infty$. Then for any $p \geq 2$, there exists $n_0 \in \mathbb{N}$ and some positive constants ρ and C such that for any $n \geq n_0$ we have

$$|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p \leq e^{-\rho(s_n - s_{n_0})} |Y_{s_{n_0}}^\vartheta - Z_{s_{n_0}}^\vartheta|^p + C \sum_{k=n_0}^{n-1} \phi_{k,p}(Y_{s_k}) e^{-\rho(s_n - s_{k+1})}. \quad (4.3.24)$$

Proof. Let $n \geq 0$ and consider the dynamics of $Y^\vartheta - Z^\vartheta$ on $[s_n, s_{n+1})$. That is, set $\varepsilon_t := Y_{s_n+t}^\vartheta - Z_{s_n+t}^\vartheta$ for $t \in [0, \gamma_{n+1})$. Then ε_t verifies the relation

$$\varepsilon_t = Y_{s_n}^\vartheta - Z_{s_n}^\vartheta + \int_{s_n}^{s_n+t} \left(b_\vartheta(Y_s^\vartheta) - b_\vartheta(Z_{s_n}^\vartheta) \right) ds. \quad (4.3.25)$$

Starting from this equation, we divide the proof in several steps.

Step 1: Contracting bound for ε_t . Consider a parameter $\eta > 0$. We wish to use the coercivity assumption $(\mathbf{C}_s)(i)$ in order to get an upper-bound on the following derivative:

$$(e^{\eta t} |\varepsilon_t|^p)' = e^{\eta t} (p |\varepsilon_t|^{p-2} \langle \varepsilon_t, \varepsilon_t' \rangle + \eta |\varepsilon_t|^p). \quad (4.3.26)$$

To this aim, observe that thanks to (4.3.25) the quantity $\langle \varepsilon_t, \varepsilon_t' \rangle$ can be expressed as:

$$\begin{aligned} \langle \varepsilon_t, \varepsilon_t' \rangle &= \langle Y_{s_n+t}^\vartheta - Z_{s_n+t}^\vartheta, b_\vartheta(Y_{s_n+t}^\vartheta) - b_\vartheta(Z_{s_n}^\vartheta) \rangle \\ &= \langle Y_{s_n+t}^\vartheta - Z_{s_n+t}^\vartheta, b_\vartheta(Y_{s_n+t}^\vartheta) - b_\vartheta(Z_{s_n+t}^\vartheta) \rangle + \langle Y_{s_n+t}^\vartheta - Z_{s_n+t}^\vartheta, b_\vartheta(Z_{s_n+t}^\vartheta) - b_\vartheta(Z_{s_n}^\vartheta) \rangle. \end{aligned}$$

Then we invoke $(\mathbf{C}_s)(i)$ and the elementary inequality $ab \leq a^2 + b^2$, valid for all $a, b \geq 0$. We obtain

$$\begin{aligned} \langle \varepsilon_t, \varepsilon_t' \rangle &\leq -\alpha |Y_{s_n+t}^\vartheta - Z_{s_n+t}^\vartheta|^2 + \frac{\alpha}{2} |Y_{s_n+t}^\vartheta - Z_{s_n+t}^\vartheta|^2 + \frac{2}{\alpha} |b_\vartheta(Z_{s_n+t}^\vartheta) - b_\vartheta(Z_{s_n}^\vartheta)|^2 \\ &\leq -\frac{\alpha}{2} |\varepsilon_t|^2 + \frac{2L^2}{\alpha} |Z_{s_n+t}^\vartheta - Z_{s_n}^\vartheta|^2, \end{aligned} \quad (4.3.27)$$

where the second inequality is due to relation (4.2.5) and the definition (4.3.25) of ε_t . We now plug relation (4.3.21) into this inequality in order to get

$$\langle \varepsilon_t, \varepsilon_t' \rangle \leq -\frac{\alpha}{2} |\varepsilon_t|^2 + \frac{2L^2}{\alpha} \left| tb(Z_{s_n}^\vartheta) + \sigma(B_{s_n+t} - B_{s_n}) \right|^2,$$

from which we easily end up with

$$p |\varepsilon_t|^{p-2} \langle \varepsilon_t, \varepsilon_t' \rangle \leq -\frac{p\alpha}{2} |\varepsilon_t|^p + \frac{4pL^2}{\alpha} |\varepsilon_t|^{p-2} \left(t^2 |b_\vartheta(Z_{s_n}^\vartheta)|^2 + |\sigma|^2 |B_{s_n+t} - B_{s_n}|^2 \right). \quad (4.3.28)$$

Eventually we apply Young's inequality with parameters $\bar{p} = p/(p-2)$ and $\bar{q} = p/2$ and some appropriate weights to (4.3.28). This yields the existence of a constant $C = C_{p,L}$

such that

$$p|\varepsilon_t|^{p-2}\langle\varepsilon_t, \varepsilon_t'\rangle \leq -\frac{p\alpha}{4}|\varepsilon_t|^p + C\left(t^p|b_\vartheta(Z_{s_n}^\vartheta)|^p + |B_{s_n+t} - B_{s_n}|^p\right). \quad (4.3.29)$$

We are now ready to give some information about expressions of the form $e^{\eta t}|\varepsilon_t|^p$. Namely we set $\eta = p\alpha/4$, then we apply identity (4.3.26) and inequality (4.3.29). This easily yields

$$(e^{\eta t}|\varepsilon_t|^p)' \leq e^{\eta t} C(t^p|b_\vartheta(Z_{s_n}^\vartheta)|^p + |B_{s_n+t} - B_{s_n}|^p). \quad (4.3.30)$$

Step 2: Inductive procedure. Let us integrate (4.3.30) on the interval $[0, \gamma_{n+1}]$, where we recall that $\gamma_{n+1} = s_{n+1} - s_n$. With the definition (4.3.25) in mind, this gives

$$|Y_{s_{n+1}}^\vartheta - Z_{s_{n+1}}^\vartheta|^p \leq e^{-\eta\gamma_{n+1}}|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p + C \int_0^{\gamma_{n+1}} e^{-\eta(\gamma_{n+1}-t)}(t^p|b_\vartheta(Z_{s_n}^\vartheta)|^p + |B_{s_n+t} - B_{s_n}|^p)dt,$$

which yields

$$|Y_{s_{n+1}}^\vartheta - Z_{s_{n+1}}^\vartheta|^p \leq e^{-\eta\gamma_{n+1}}|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p + C\left(\gamma_{n+1}^{p+1}|b_\vartheta(Z_{s_n}^\vartheta)|^p + \int_0^{\gamma_{n+1}} |B_{s_n+t} - B_{s_n}|^p dt\right). \quad (4.3.31)$$

Thus, setting $\phi_{n,p}(x) = \left(\gamma_{n+1}^{p+1}|b_\vartheta(x)|^p + \int_0^{\gamma_{n+1}} |B_{s_n+t} - B_{s_n}|^p dt\right)$, an elementary induction procedure yields the following relation for every $n \geq 1$:

$$|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p \leq |Y_0^\vartheta - Z_0^\vartheta|^p e^{-\eta s_n} + \sum_{k=0}^{n-1} \phi_{k,p}(Z_{s_k}^\vartheta) e^{-\eta(s_n - s_{k+1})}.$$

This proves our claim (4.3.22).

Step 3: Proof of (4.3.24). In order to obtain our second statement (4.3.24), one needs to go back to inequality (4.3.31). Then observe that hypothesis (4.2.5) yields

$$|b(Z_{s_n}^\vartheta)|^p \leq 2^{p-1} \left(|b(Y_{s_n}^\vartheta)|^p + L^p|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p\right),$$

which leads for any p to the existence of a constant C such that

$$|Y_{s_{n+1}}^\vartheta - Z_{s_{n+1}}^\vartheta|^p \leq \left(e^{-\eta\gamma_{n+1}} + C\gamma_{n+1}^{p+1}\right)|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p + C\phi_{n,p}(Y_{s_n}). \quad (4.3.32)$$

Since $\lim_{n \rightarrow \infty} \gamma_n = 0$, one checks easily that there exists a n_0 such that for any $n \geq n_0$ we have the following inequality:

$$e^{-\eta\gamma_{n+1}} + C\gamma_{n+1}^{p+1} \leq e^{-\frac{\eta}{2}\gamma_{n+1}}.$$

Plugging this information into (4.3.32), we end up with

$$|Y_{s_{n+1}}^\vartheta - Z_{s_{n+1}}^\vartheta|^p \leq e^{-\frac{\eta}{2}\gamma_{n+1}}|Y_{s_n}^\vartheta - Z_{s_n}^\vartheta|^p + C\phi_{k,p}(Y_{s_n}).$$

Our assertion (4.3.24) then follows by an induction procedure exactly as for Step 2. \square

We now give a continuity results (with respect to the parameter ϑ) for some occupa-

tion measures related to our processes of interest. The proofs are postponed to Appendix 4.B.

Proposition 4.3.8. *As in Proposition 4.3.6, let Y^ϑ be the solution of equation (4.2.7) and consider the Euler scheme $Z^{\vartheta,\gamma}$ defined by (4.2.13). Also consider $p > 0$ and $d \in \mathcal{D}_p$. We assume that (\mathbf{C}_s) holds true. Then,*

(i) *The occupation measures of the process Y^ϑ are Lipschitz with respect to ϑ , that is there exists a positive random variable C_p such that for all $t \geq 1$:*

$$d\left(\frac{1}{t} \int_0^t \delta_{Y_s^{\vartheta_1}} ds, \frac{1}{t} \int_0^t \delta_{Y_s^{\vartheta_2}} ds\right) \leq C_p |\vartheta_1 - \vartheta_2|. \quad (4.3.33)$$

(ii) *The occupation measures of the Euler approximation $Z^{\vartheta,\gamma}$ are also Lipschitz with respect to ϑ . Namely there exists $\gamma_0 > 0$ such that: for any $\gamma \in (0, \gamma_0]$, there exists a positive random variable $C_p(\gamma)$ such that for all $N \geq 1$*

$$d\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta_1,\gamma}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta_2,\gamma}}\right) \leq C_p(\gamma) |\vartheta_1 - \vartheta_2|. \quad (4.3.34)$$

Remark 4.3.1. In the sequel we will analyze several quantities like (4.3.34), where we compare two discrete random measures on \mathbb{R}^d $\nu_1 = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$ and $\nu_2 = \frac{1}{N} \sum_{i=1}^N \delta_{Y_i}$ with $X_i = X_i(\omega)$ and $Y_i = Y_i(\omega)$. In this context we will always upper bound quantities of the form $d(\nu_1, \nu_2)$ for a distance d in \mathcal{D}_p . To this aim, resorting a trivial coupling between ν_1 and ν_2 , it is enough to prove an almost sure upper bound on

$$\frac{1}{N} \sum_{i=1}^N |X_i - Y_i|^p.$$

We will adopt this strategy throughout the chapter, the typical outcome being an a.s. bound on $d(\nu_1, \nu_2)$. Straightforward extensions to a continuous time setting allow to handle quantities of the form (4.3.33).

4.4 Proof of the consistency theorems

The aim of this section is to achieve the proof of Theorems 4.2.1, 4.2.2 and 4.2.3. We first establish a general asymptotic result for a family of contrasts in Sections 4.4.1 and 4.4.2. Then we will combine this general proposition with our preliminary results of Section 4.3 in order to prove our main claims.

4.4.1 Uniform convergence of the contrast

In this section we state some uniform convergence results for the *contrast*, i.e for the function involved in the definition of estimators such as (4.2.15). We should notice at this point that our uniform convergence results hold only under the assumption (\mathbf{C}_s) . In case of a constant time step Euler scheme, we get the following result.

Proposition 4.4.1. *We consider the same setting as in Proposition 4.3.8. In particular, we assume that (\mathbf{C}_s) holds true for the coefficients of equation (4.2.7). Then the following assertions hold true.*

(i) *The invariant measure ν^γ of the Euler scheme converges uniformly to the invariant measure of Y^ϑ . Namely*

$$\limsup_{\gamma \rightarrow 0} \gamma^{-H} \sup_{\vartheta \in \Theta} d(\nu_\vartheta, \nu_\vartheta^\gamma) < +\infty.$$

In particular, we have

$$\lim_{\gamma \rightarrow 0} \sup_{\vartheta \in \Theta} d(\nu_\vartheta, \nu_\vartheta^\gamma) = 0.$$

(ii) *The occupation measure of the Euler scheme converges to the invariant measure ν_ϑ^γ as the number of steps goes to ∞ , that is:*

$$\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} d\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{s_k}^{\vartheta, \gamma}}, \nu_\vartheta^\gamma\right) = 0.$$

(iii) *We have*

$$\lim_{\gamma \rightarrow 0} \lim_{N, n \rightarrow +\infty} \sup_{\vartheta \in \Theta} \left| d\left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}^\vartheta}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{s_k}^{\vartheta, \gamma}}\right) - d(\nu_{\vartheta_0}, \nu_\vartheta) \right| = 0.$$

Proof. We prove the three items separately.

Proof of (i). For sake of simplicity, we only detail the proof for $p = 2$. The extension to a general p does not generate particular difficulties and can be done as in Proposition 4.3.7. We start from the following inequality:

$$d(\nu_\vartheta^\gamma, \nu_\vartheta) \leq d\left(\mathcal{L}\left(Z_{N\gamma}^{\vartheta, \gamma}\right), \mathcal{L}\left(Y_{N\gamma}^\vartheta\right)\right) + d\left(\nu_\vartheta^\gamma, \mathcal{L}\left(Z_{N\gamma}^{\vartheta, \gamma}\right)\right) + d\left(\mathcal{L}\left(Y_{N\gamma}^\vartheta\right), \nu_\vartheta\right). \quad (4.4.1)$$

Let us consider the three terms of the right-hand side of (4.4.1) successively. First, without loss of generality, we can assume that $Y_0^\vartheta = Z_0^{\vartheta, \gamma}$. Furthermore, we have that $Z_{n\gamma}^{\vartheta, \gamma} = Z_{n\gamma}^\vartheta$ for all $n \geq 1$, where Z^ϑ is defined by (4.3.21). Then, by Proposition 4.3.7 (i) applied with $\gamma_n = \gamma$, we have

$$\mathbb{E}[|Y_{n\gamma}^\vartheta - Z_{n\gamma}^{\vartheta, \gamma}|^2] \leq C \sum_{k=0}^{n-1} \mathbb{E}[\phi_{k,2}(Z_{k\gamma}^{\vartheta, \gamma})] e^{-\rho\gamma(n-1-k)},$$

where we recall that $\phi_{k,p}$ is defined by (4.3.23). Using that b_ϑ is sublinear (uniformly in ϑ) and the fact that the increments of B satisfy relation (4.1.1), one obtains

$$\mathbb{E}[|Y_{n\gamma}^\vartheta - Z_{n\gamma}^{\vartheta, \gamma}|^2] \leq C \sum_{k=0}^{n-1} \left(\gamma^3 \left(1 + \mathbb{E}[|Z_{k\gamma}^{\vartheta, \gamma}|^2]\right) + \gamma^{2H+1}\right) e^{-\rho\gamma(n-1-k)}.$$

It follows easily that

$$\sup_{\vartheta \in \Theta, \gamma \in (0, \gamma_0]} \limsup_{n \rightarrow +\infty} \gamma^{-2H} \mathbb{E}[|Y_{n\gamma}^\vartheta - Z_{n\gamma}^{\vartheta, \gamma}|^2] \leq 1 + \sup_{\vartheta \in \Theta, \gamma \in (0, \gamma_0]} \limsup_{n \rightarrow +\infty} \mathbb{E}[|Z_{k\gamma}^{\vartheta, \gamma}|^2].$$

Moreover, owing to Proposition 4.A.2, there exists γ_0 such that

$$\limsup_{n \rightarrow +\infty} \mathbb{E}[|Z_{k\gamma}^{\vartheta, \gamma}|^2] < \infty.$$

Summarizing, we have obtained

$$\sup_{\vartheta \in \Theta, \gamma \in (0, \gamma_0]} \limsup_{n \rightarrow +\infty} \gamma^{-2H} \mathbb{E}[|Y_{n\gamma}^\vartheta - Z_{n\gamma}^{\vartheta, \gamma}|^2] < \infty.$$

We now consider the term $d(\nu_\vartheta^\gamma, \mathcal{L}(Z_{N\gamma}^{\vartheta, \gamma}))$ in the right hand side of (4.4.1). Using that ν_ϑ^γ is invariant, we remark that

$$d\left(\nu_\vartheta^\gamma, \mathcal{L}\left(Z_{N\gamma}^{\vartheta, \gamma}\right)\right) \leq \mathbb{E}[|Z_{N\gamma}^{\vartheta, \gamma} - \bar{Z}_{N\gamma}^{\vartheta, \gamma}|^2]^{\frac{1}{2}}$$

where $\bar{Z}^{\vartheta, \gamma}$ denotes a stationary Euler scheme built with the same noise process as for $Z^{\vartheta, \gamma}$. Thus, thanks to the fact that

$$Z_{k\gamma}^{\vartheta, \gamma} - \bar{Z}_{k\gamma}^{\vartheta, \gamma} = Z_{(k-1)\gamma}^{\vartheta, \gamma} - \bar{Z}_{(k-1)\gamma}^{\vartheta, \gamma} + \gamma \left(b(Z_{(k-1)\gamma}^{\vartheta, \gamma}) - b(\bar{Z}_{(k-1)\gamma}^{\vartheta, \gamma}) \right)$$

a straightforward induction under assumption (\mathbf{C}_s) , similar to (4.3.27), leads to

$$\left| Z_{k\gamma}^{\vartheta, \gamma} - \bar{Z}_{k\gamma}^{\vartheta, \gamma} \right|^2 \leq (1 - 2\gamma\alpha + \gamma^2 L^2)^k \left| Z_0^{\vartheta, \gamma} - \bar{Z}_0^{\vartheta, \gamma} \right|^2. \quad (4.4.2)$$

We choose $\gamma_0 = \alpha/L^2$ in such a way that $2\alpha - \gamma L^2 \geq \alpha$ for any $\gamma \in (0, \gamma_0]$. In addition, recall that $Z_0^{\vartheta, \gamma} = z_0$. Then, by possibly picking a smaller value of γ_0 , we deduce from Proposition 4.A.2 that

$$\limsup_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} d\left(\nu_\vartheta^\gamma, \mathcal{L}\left(Z_{N\gamma}^{\vartheta, \gamma}\right)\right) = 0.$$

Eventually, the last term in the right hand side of (4.4.1) tends to 0 uniformly in ϑ as $N \rightarrow +\infty$ by Proposition 4.3.1. This concludes the proof of (i).

Proof of (ii). By Proposition 4.3.4, the convergence of $d\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta, \gamma}}, \nu_\vartheta^\gamma\right)$ to 0 is true for the simple convergence. In order to extend this result to a uniform convergence in ϑ , we use Proposition 4.3.8(ii) to obtain that the family $\{\vartheta \mapsto d\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta, \gamma}}, \nu_\vartheta^\gamma\right); N \geq 1, \vartheta \in \Theta\}$ is equicontinuous for a fixed $\gamma \in (0, \gamma_0]$. Actually, for some given ϑ_1 and ϑ_2 ,

$$\left| d\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta_1, \gamma}}, \nu_{\vartheta_1}^\gamma\right) - d\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta_2, \gamma}}, \nu_{\vartheta_2}^\gamma\right) \right| \leq d(\nu_{\vartheta_1}^\gamma, \nu_{\vartheta_2}^\gamma) + d\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta_1, \gamma}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta_2, \gamma}}\right).$$

The second term goes to 0 as $\vartheta_1 - \vartheta_2 \rightarrow 0$ by Proposition 4.3.8(ii). This is also the case for the first one by letting N go to ∞ in Proposition 4.3.8(ii) (for instance).

Proof of (iii). This is a simple consequence of the two previous statements and of Proposition 4.3.3(ii). \square

We now generalize Proposition 4.4.1 to the case of a decreasing time step for the Euler scheme (4.2.13).

Proposition 4.4.2. *We consider the same setting as in Proposition 4.3.8. In particular, we assume that (\mathbf{C}_s) holds true for the coefficients of equation (4.2.7). Let $p \geq 2$ and consider $d \in \mathcal{D}_p$. Let $\{s_k; k \geq 0\}$ be the sequence of time steps defined by (4.2.18), which is assumed to verify*

$$\exists p' \geq p \text{ such that } \sum_{k=1}^{+\infty} \frac{\gamma_{k+1}^{p'H+1}}{s_k} < +\infty. \quad (4.4.3)$$

Then we have

$$\lim_{N,n \rightarrow +\infty} \sup_{\vartheta \in \Theta} \left| d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^\vartheta} \right) - d(\nu_{\vartheta_0}, \nu_\vartheta) \right| = 0. \quad (4.4.4)$$

Proof. For notational convenience, the proof will be detailed for the continuous-time Euler approximation $(Z_t^\vartheta)_{t \geq 0}$ defined by (4.3.21), with step sequence $(\gamma_n)_{n \geq 1}$. An application of the triangular inequality allows us to bound the left hand side of (4.4.4) as follows:

$$\left| d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^\vartheta} \right) - d(\nu_{\vartheta_0}, \nu_\vartheta) \right| \leq A_{1,n} + A_{2,N}(\vartheta). \quad (4.4.5)$$

with

$$A_{1,n} := d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \nu_{\vartheta_0} \right), \quad A_{2,N}(\vartheta) := d \left(\frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^\vartheta}, \nu_\vartheta \right).$$

Our claim can thus be reduced to prove that

$$\lim_{n \rightarrow +\infty} A_{1,n} = 0, \quad \text{and} \quad \lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{2,N}(\vartheta) = 0.$$

Furthermore, the fact that $\lim_{n \rightarrow +\infty} A_{1,n} = 0$ is a direct consequence of Proposition 4.3.3-(ii). We thus focus on the asymptotic behavior of $A_{2,N}$ in the remainder of the proof.

In order to bound $A_{2,N}$, let us set $\underline{s} = \max\{s_k, s_k \leq s\}$. Then we observe that

$$\frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^\vartheta} = \frac{1}{s_N} \int_0^{s_N} \delta_{Z_{\underline{s}}^\vartheta} ds.$$

Therefore we can split $A_{2,N}$ into

$$A_{2,N}(\vartheta) \leq A_{21,N}(\vartheta) + A_{22,N}(\vartheta) + A_{23,N}(\vartheta) \quad (4.4.6)$$

where the quantities $A_{2j,N}(\vartheta)$ for $j = 1, 2, 3$ are defined as follows:

$$A_{21,N}(\vartheta) := d \left(\frac{1}{s_N} \int_0^{s_N} \delta_{Z_{\underline{s}}} ds, \frac{1}{s_N} \int_0^{s_N} \delta_{Y_{\underline{s}}} ds \right),$$

$$A_{22,N}(\vartheta) := d \left(\frac{1}{s_N} \int_0^{s_N} \delta_{Y_{\underline{s}}} ds, \frac{1}{s_N} \int_0^{s_N} \delta_{Y_{\underline{s}}} ds \right), \quad (4.4.7)$$

$$A_{23,N}(\vartheta) := d \left(\frac{1}{s_N} \int_0^{s_N} \delta_{Y_{\underline{s}}} ds, \nu_{\vartheta} \right). \quad (4.4.8)$$

We will now treat those three terms separately.

The term $A_{23,N}$ is easily handled by applying Proposition 4.3.3 (i) (simple convergence) and Proposition 4.4.1 (i) (equicontinuity). We thus get

$$\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{23,N}(\vartheta) = 0.$$

Now, let us prove that $\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{22,N}(\vartheta) = 0$. Since $d \in \mathcal{D}_p \subset \mathcal{D}_{p'}$, we can assume without loss of generality that $d = \mathcal{W}_{p'}$. We invoke the strategy outlined in Remark 4.3.1. This means that we are reduced to prove the following limit :

$$\sup_{\vartheta \in \Theta} \frac{1}{s_N} \int_0^{s_N} |Y_t^\vartheta - Y_{\underline{t}}^\vartheta|^{p'} dt \xrightarrow{N \rightarrow +\infty} 0 \quad a.s. \quad (4.4.9)$$

To this aim, we first note that

$$|Y_t^\vartheta - Y_{\underline{t}}^\vartheta| \leq \int_{\underline{t}}^t |b_\vartheta(Y_s^\vartheta)| ds + \|\sigma\| |B_t - B_{\underline{t}}|, \quad (4.4.10)$$

and thus there exists a constant $c_{p',\sigma}$ such that

$$\frac{1}{s_N} \int_0^{s_N} |Y_t^\vartheta - Y_{\underline{t}}^\vartheta|^{p'} dt \leq c_{p',\sigma} \left[\frac{1}{s_N} \int_0^{s_N} \left| \int_{\underline{t}}^t |b_\vartheta(Y_s^\vartheta)| ds \right|^{p'} dt + \frac{1}{s_N} \int_0^{s_N} |B_t - B_{\underline{t}}|^{p'} dt \right]. \quad (4.4.11)$$

Next we upper bound the first term in the right hand side of (4.4.11) invoking successively Jensen's inequality and Fubini's theorem. We obtain that

$$\begin{aligned} \int_0^{s_N} \left| \int_{\underline{t}}^t |b_\vartheta(Y_s^\vartheta)| ds \right|^{p'} dt &\leq \int_0^{s_N} |t - \underline{t}|^{p'-1} \left(\int_{\underline{t}}^t |b_\vartheta(Y_s^\vartheta)|^{p'} ds \right) dt \\ &= \int_0^{s_N} |b_\vartheta(Y_s^\vartheta)|^{p'} \left(\int_s^{\bar{s}} |t - \underline{s}|^{p'-1} dt \right) ds, \end{aligned}$$

where we have introduced the additional notation $\bar{s} = \min\{s_k, s_k > s\}$. Taking into account the fact that $|t - \underline{s}| \leq \gamma_{k+1}$ for any $s, t \in [s_k, s_{k+1})$, we end up with

$$\int_0^{s_N} \left| \int_{\underline{t}}^t |b_\vartheta(Y_s^\vartheta)| ds \right|^{p'} dt \leq \sum_{k=0}^{N-1} \gamma_{k+1}^{p'} \int_{s_k}^{s_{k+1}} |b_\vartheta(Y_s^\vartheta)|^{p'} ds. \quad (4.4.12)$$

Furthermore, since b_ϑ is uniformly sublinear in ϑ and owing to (4.3.9) we know that

$$C_1 := \sup_{\vartheta \in \Theta} \sup_{N \geq 1} \frac{1}{s_N} \int_0^{s_N} |b_\vartheta(Y_s^\vartheta)|^{p'} ds = \sup_{\vartheta \in \Theta} \sup_{N \geq 1} \frac{1}{s_N} \sum_{k=0}^{N-1} \int_{s_k}^{s_{k+1}} |b_\vartheta(Y_s^\vartheta)|^{p'} ds < +\infty \quad a.s.$$

Moreover, $\lim_{k \rightarrow +\infty} \gamma_{k+1}^{p'} = 0$. Hence, for all $\varepsilon > 0$ there exists $k_0 \geq 0$ such that for all $k \geq k_0$, we have $\gamma_{k+1}^{p'} < \varepsilon$. One can thus deduce that

$$\begin{aligned} \sup_{\vartheta \in \Theta} \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1}^{p'} \int_{s_k}^{s_{k+1}} |b_\vartheta(Y_s^\vartheta)|^{p'} ds \\ \leq \sup_{\vartheta \in \Theta} \frac{1}{s_N} \sum_{k=0}^{k_0-1} \gamma_{k+1}^{p'} \int_{s_k}^{s_{k+1}} |b_\vartheta(Y_s^\vartheta)|^{p'} ds + \sup_{\vartheta \in \Theta} \frac{1}{s_N} \sum_{k=k_0}^{N-1} \gamma_{k+1}^{p'} \int_{s_k}^{s_{k+1}} |b_\vartheta(Y_s^\vartheta)|^{p'} ds \\ \leq C_1 \left(\frac{\gamma_1^{p'} s_{k_0}}{s_N} + \varepsilon \right), \end{aligned} \quad (4.4.13)$$

from which it is easily seen that

$$\sup_{\vartheta \in \Theta} \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1}^{p'} \int_{s_k}^{s_{k+1}} |b_\vartheta(Y_s^\vartheta)|^{p'} ds \xrightarrow{N \rightarrow +\infty} 0 \quad a.s. \quad (4.4.14)$$

We can conclude that the first term in the right hand side of (4.4.11) vanishes as $N \rightarrow +\infty$ due to our identity (4.4.12).

In order to prove that the second term in the right hand side of (4.4.11) converges to 0, we invoke Kronecker's lemma (in its continuous version, see [25, Theorem 2.1]). We get that it is sufficient to prove that:

$$\int_0^{+\infty} \frac{|B_t - B_{\underline{t}}|^{p'}}{1+t} dt < +\infty \quad a.s.$$

However, due to the fact that $(t - \underline{t}) \leq \gamma_{k+1}$ if $t \in [s_k, s_{k+1})$, we have,

$$\mathbb{E} \left[\int_0^{+\infty} \frac{|B_t - B_{\underline{t}}|^{p'}}{1+t} dt \right] \leq \int_0^{+\infty} \frac{1}{1+t} (t - \underline{t})^{p'H} dt \leq c \sum_{k=1}^{+\infty} \frac{\gamma_{k+1}^{p'H+1}}{s_k} < +\infty$$

where c is a positive constant and where the last inequality stems from hypothesis (4.4.3). Hence, Kronecker's lemma yields

$$\frac{1}{s_N} \int_0^{s_N} |B_t - B_{\underline{t}}|^{p'} ds \xrightarrow{N \rightarrow +\infty} 0 \quad a.s. \quad (4.4.15)$$

Now inequality (4.4.11) combined with (4.4.14) and (4.4.15) easily yields (4.4.9). We conclude that $\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{22,N}(\vartheta) = 0$.

Going back to our decomposition (4.4.6), we still have to prove that $\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{21,N}(\vartheta) = 0$. To this end, let us write $A_{21,N}(\vartheta)$ in its discrete

form:

$$A_{21,N}(\vartheta) = d \left(\frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^\vartheta}, \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Y_{s_k}^\vartheta} \right). \quad (4.4.16)$$

Then invoking Remark 4.3.1 once more, we are reduced to prove

$$\frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} |Y_{s_k}^\vartheta - Z_{s_k}^\vartheta|^{p'} \xrightarrow{N \rightarrow +\infty} 0 \quad a.s. \quad (4.4.17)$$

In order to achieve (4.4.17) consider the integer n_0 given by Proposition 4.3.7 (ii). For $k \leq n_0$ we trivially bound $|Y_{s_k}^\vartheta - Z_{s_k}^\vartheta|$ using (4.3.22). Since (4.3.21) asserts that $\sup_{\vartheta \in \Theta, k \in \{1, \dots, n_0\}} |Z_{s_k}^\vartheta| < +\infty$, we get that

$$\sup_{\vartheta \in \Theta, k \in \{1, \dots, n_0\}} |Y_{s_k}^\vartheta - Z_{s_k}^\vartheta| =: C(\omega) < +\infty \quad a.s. \quad (4.4.18)$$

We now bound the right hand side of (4.4.17) by means of (4.4.18) whenever $k \leq n_0$ and invoking (4.3.24) when $k \geq n_0 + 1$. This gives

$$\begin{aligned} \frac{1}{s_N} \sum_{k=0}^{N-1} \gamma_{k+1} |Y_{s_k}^\vartheta - Z_{s_k}^\vartheta|^{p'} &\leq C(\omega)^{p'} \frac{1}{s_N} \left(\sum_{k=0}^{n_0} \gamma_{k+1} + \sum_{k=n_0+1}^{N-1} \gamma_{k+1} e^{-\rho(s_k - s_{n_0})} \right) \\ &\quad + \frac{1}{s_N} \sum_{k=n_0+1}^{N-1} \gamma_{k+1} \sum_{\ell=n_0}^{k-1} \phi_{\ell, p'}(Y_{s_\ell}^\vartheta) e^{-\rho(s_k - s_{\ell+1})}. \end{aligned} \quad (4.4.19)$$

The first term in the right-hand side of (4.4.19) is clearly evanescent as $N \rightarrow +\infty$. Let us focus on the second one. By a Fubini type transformation, we get

$$\begin{aligned} \sum_{k=n_0+1}^{N-1} \gamma_{k+1} \sum_{\ell=n_0}^{k-1} \phi_{\ell, p'}(Y_{s_\ell}^\vartheta) e^{-\rho(s_k - s_{\ell+1})} &= \sum_{\ell=n_0}^{N-2} \phi_{\ell, p'}(Y_{s_\ell}^\vartheta) \sum_{k=\ell+1}^{N-1} \gamma_{k+1} e^{-\rho(s_k - s_{\ell+1})} \\ &\leq \sum_{\ell=0}^{N-2} \phi_{\ell, p'}(Y_{s_\ell}^\vartheta) \sum_{k=\ell+1}^{N-1} \gamma_k e^{-\rho(s_k - s_{\ell+1})} \end{aligned}$$

where the last inequality is due to the fact that (γ_k) is non increasing. Since $x \mapsto e^{-\rho x}$ is a non-increasing function, we remark that for $\ell \in \{0, \dots, N-2\}$,

$$\sum_{k=\ell+1}^{N-1} \gamma_k e^{-\rho(s_k - s_{\ell+1})} \leq e^{\rho s_{\ell+1}} \int_{s_\ell}^{s_{N-1}} e^{-\rho x} dx = \frac{1}{\rho} \left(e^{\rho \gamma_{\ell+1}} - e^{-\rho(s_{N-1} - s_{\ell+1})} \right) \leq C_\rho.$$

Thus, in order to see that the right hand side of (4.4.19) vanishes as $N \rightarrow +\infty$, it remains to show that

$$\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} \frac{1}{s_N} \sum_{\ell=0}^{N-1} \phi_{\ell, p'}(Y_{s_\ell}^\vartheta) = 0 \quad (4.4.20)$$

where we recall that $\phi_{\ell,p'}$ is defined by (4.3.23) and thus

$$\frac{1}{s_N} \sum_{\ell=0}^{N-1} \phi_{\ell,p'}(Y_{s_\ell}^\vartheta) = \frac{1}{s_N} \sum_{\ell=0}^{N-1} \gamma_{\ell+1}^{p'+1} |b_\vartheta(Y_{s_\ell}^\vartheta)|^{p'} + \frac{1}{s_N} \sum_{\ell=0}^{N-1} \int_0^{\gamma_{\ell+1}} |B_{s_\ell+t} - B_{s_\ell}|^{p'} dt. \quad (4.4.21)$$

Let us begin by the term involving b_ϑ in (4.4.21). First notice that

$$\mathcal{W}_{p'} \left(\frac{1}{s_N} \int_0^{s_N} \delta_{Y_{\underline{s}}} ds, \nu_\vartheta \right) \leq A_{22,N}(\vartheta) + A_{23,N}(\vartheta),$$

where $A_{22,N}$ and $A_{23,N}$ are respectively defined by (4.4.7) and (4.4.8) with $d = \mathcal{W}_{p'}$. Furthermore, we have seen that $\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{22,N}(\vartheta) = 0$ and $\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{23,N}(\vartheta) = 0$. We immediately deduce that

$$\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} \mathcal{W}_{p'} \left(\frac{1}{s_N} \int_0^{s_N} \delta_{Y_{\underline{s}}} ds, \nu_\vartheta \right) = 0.$$

Moreover, by Proposition 4.A.1 for instance, we have $\sup_{\vartheta \in \Theta} \nu_\vartheta(|\cdot|^{p'}) < \infty$. We thus deduce that

$$\sup_{\vartheta \in \Theta} \sup_{N \geq 1} \frac{1}{s_N} \sum_{\ell=0}^{N-1} \gamma_{\ell+1} |Y_{s_\ell}^\vartheta|^{p'} < +\infty \quad a.s.$$

Furthermore, since b_ϑ is uniformly sublinear in ϑ and $\lim_{\ell \rightarrow +\infty} \gamma_{\ell+1}^{p'} = 0$, it easily follows along the same lines as for (4.4.14) that:

$$\sup_{\vartheta \in \Theta} \frac{1}{s_N} \sum_{\ell=0}^{N-1} \gamma_{\ell+1}^{p'+1} |b_\vartheta(Y_{s_\ell}^\vartheta)|^{p'} \xrightarrow{N \rightarrow +\infty} 0. \quad (4.4.22)$$

We now turn to the term in (4.4.21) involving the fBm, for which we use the classical discrete Kronecker lemma. To this aim, we remark that

$$\mathbb{E} \left[\sum_{k=1}^{+\infty} \frac{1}{s_k} \int_0^{\gamma_{k+1}} |B_{s_k+t} - B_{s_k}|^{p'} dt \right] = \frac{1}{p'H + 1} \sum_{k=1}^{+\infty} \frac{\gamma_{k+1}^{p'H+1}}{s_k} < +\infty, \quad (4.4.23)$$

where the last inequality stems from assumption (4.4.3). From (4.4.23), it is easily seen that

$$\sum_{\ell=0}^{+\infty} \frac{1}{s_\ell} \int_0^{\gamma_{\ell+1}} |B_{s_\ell+t} - B_{s_\ell}|^{p'} dt < +\infty \quad a.s.$$

We are thus in position to apply Kronecker's lemma to the sequence $(\int_0^{\gamma_{\ell+1}} |B_{s_\ell+t} - B_{s_\ell}|^{p'} dt)_{\ell \geq 0}$. This yields

$$\frac{1}{s_N} \sum_{\ell=0}^{N-1} \int_0^{\gamma_{\ell+1}} |B_{s_\ell+t} - B_{s_\ell}|^{p'} dt \xrightarrow{N \rightarrow +\infty} 0. \quad (4.4.24)$$

Let us summarize our computations so far. Gathering (4.4.22), (4.4.24) and (4.4.21), we have obtained relation (4.4.20). This easily implies that (4.4.17) holds true and hence

$A_{21,N}(\vartheta)$ defined by (4.4.16) satisfies

$$\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{21,N}(\vartheta) = 0.$$

Since similar results have also been shown for $A_{22,N}(\vartheta)$ and $A_{23,N}(\vartheta)$, relation (4.4.6) gives

$$\lim_{N \rightarrow +\infty} \sup_{\vartheta \in \Theta} A_{2,N}(\vartheta) = 0.$$

Eventually, plugging this information into (4.4.5) yields our claim (4.4.4). \square

4.4.2 A general convergence result

The consistence of our estimators will rely on the following general proposition about convergence of minimizers for a sequence of random functions. Observe that our sequences below are indexed by a generic r which sits in an unspecified set. This simplifies the subsequent applications of the proposition to our indices N, n, γ in the remainder of the section.

Proposition 4.4.3. *Let Θ be a compact set and $(\vartheta \mapsto L_r(\vartheta))_r$ denote a family of non-negative random functions. Assume that:*

1. *With probability one, $\lim_r L_r(\vartheta) = L(\vartheta)$ uniformly in $\vartheta \in \Theta$.*
2. *The function $\vartheta \mapsto L(\vartheta)$ is non-random and continuous on Θ .*
3. *For any r , the set $\operatorname{argmin}\{L_r(\vartheta), \vartheta \in \Theta\}$ is nonempty.*

Then, for a fixed r , let $\hat{\vartheta}_r \in \operatorname{argmin}\{L_r(\vartheta), \vartheta \in \Theta\}$. Let \mathcal{A} denote the limit points of $(\hat{\vartheta}_r)_r$. Then we have

$$\mathcal{A} \subset \operatorname{argmin}\{L(\vartheta), \vartheta \in \Theta\}.$$

In particular, if L attains its minimum for a unique ϑ^ , then $\lim_r \hat{\vartheta}_r = \vartheta^*$.*

Proof. Let ϑ^* be an element of $\operatorname{argmin}\{L(\vartheta), \vartheta \in \Theta\}$. We consider a generic element $\vartheta^\infty \in \mathcal{A}$ and its related convergent subsequence $(\hat{\vartheta}_{r_n})_{n \geq 0}$. Then we can upper bound $L(\vartheta^\infty)$ as follows:

$$L(\vartheta^\infty) \leq L_{r_n}(\hat{\vartheta}_{r_n}) + |L(\vartheta^\infty) - L_{r_n}(\hat{\vartheta}_{r_n})|. \quad (4.4.25)$$

We now bound the two terms in the right hand side of (4.4.25). On the one hand, by definition of $\hat{\vartheta}_{r_n}$,

$$L_{r_n}(\hat{\vartheta}_{r_n}) \leq L_{r_n}(\vartheta^*).$$

Hence, thanks to the fact that $\lim_r L_r(\vartheta) = L(\vartheta)$ for all $\vartheta \in \Theta$, we get

$$\limsup_{n \rightarrow +\infty} L_{r_n}(\hat{\vartheta}_{r_n}) \leq L(\vartheta^*). \quad (4.4.26)$$

On the other hand, we also have

$$|L(\vartheta^\infty) - L_{r_n}(\hat{\vartheta}_{r_n})| \leq |L(\vartheta^\infty) - L(\hat{\vartheta}_{r_n})| + \sup_{\vartheta \in \Theta} |L(\vartheta) - L_{r_n}(\vartheta)|. \quad (4.4.27)$$

Therefore we can invoke the continuity of L to bound the first term in the right hand side of (4.4.27), plus the uniform convergence of L_r to L in order to handle the second term. This yields

$$\limsup_{n \rightarrow +\infty} |L(\vartheta^\infty) - L_{r_n}(\hat{\vartheta}_{r_n}^\infty)| = 0 \quad (4.4.28)$$

Plugging (4.4.28) and (4.4.26) into (4.4.25), we obtain that $L(\vartheta^\infty) \leq L(\vartheta^*)$ and thus ϑ^∞ belongs to the set $\operatorname{argmin}\{L(\vartheta), \vartheta \in \Theta\}$. This finishes the proof. \square

4.4.3 Proofs of the convergence theorems

With all our preliminary considerations in hand, we are now ready to prove the main convergence results for our estimators. This is briefly outlined below.

Proof of Theorem 4.2.2. Recall that the family $\{\hat{\vartheta}_{N,n,\gamma}, (N, n, \gamma) \in \mathbb{N}^2 \times \mathbb{R}_+^*\}$ is defined by (4.2.15). We wish to apply Proposition 4.4.3 with $r = (N, n, \gamma) \in \mathbb{N}^2 \times \mathbb{R}_+^*$. We set $L_{N,n,\gamma}(\vartheta) = d(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{s_k}^\vartheta})$. By Proposition 4.4.1, we have uniformly in $\vartheta \in \Theta$,

$$\lim_{\gamma \rightarrow 0} \lim_{N,n \rightarrow +\infty} L_{N,n,\gamma}(\vartheta) = d(\nu_{\vartheta_0}, \nu_\vartheta) =: L(\vartheta). \quad (4.4.29)$$

In addition, owing to Proposition 4.3.6 and Assumption (\mathbf{I}_w) , L is continuous and ϑ_0 is the unique minimum of L . We have thus checked that the hypothesis of Proposition 4.4.3 are fulfilled, from which Theorem 4.2.2 is easily deduced. \square

Proof of Theorem 4.2.3. The proof goes along the same lines as for Theorem 4.2.2. Namely we apply Proposition 4.4.3 to the sequence $\{\hat{\vartheta}_{N,n}, (N, n) \in \mathbb{N}^2\}$ defined by (4.2.20). To this aim, we set

$$L_{N,n}(\vartheta) = d\left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \gamma_{k+1} \delta_{Z_{s_k}^\vartheta}\right).$$

Then according to Proposition 4.4.2, the sequence $(L_{N,n})_{N,n}$ converges uniformly to L defined by (4.4.29) when $N, n \rightarrow +\infty$. Furthermore, the continuity of L follows as in the proof of Theorem 4.2.2. Our claim is thus easily deduced. \square

Proof of Theorem 4.2.1. We still wish to apply Proposition 4.4.3 to the family $\{\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)}, (N, n, \gamma) \in \mathbb{N}^2 \times \mathbb{R}_+^*\}$ defined by (4.2.17). However, since we only assume (\mathbf{C}_w) instead of (\mathbf{C}_s) , one is only able to obtain simple convergence properties on Θ . In order to circumvent this problem, we have restricted our analysis to the discretized parameter set $\Theta^{(\varepsilon)}$ introduced in (4.2.17). For a given $\varepsilon > 0$, $\Theta^{(\varepsilon)}$ is finite and hence, one deduces from Propositions 4.3.3(ii) and 4.3.4, that

$$\lim_{\gamma \rightarrow 0} \lim_{N,n \rightarrow +\infty} \sup_{\vartheta \in \Theta^{(\varepsilon)}} |L_{N,n,\gamma}(\vartheta) - L(\vartheta)| = 0,$$

where L is defined by (4.4.29). Now, denote by $\mathcal{A}^{(\varepsilon)}$ the set of limit points of $(\hat{\vartheta}_{N,n,\gamma}^{(\varepsilon)})_{N,n,\gamma}$. From Proposition 4.4.3, one deduces that

$$\mathcal{A}^{(\varepsilon)} \subset \operatorname{argmin}\{L(\vartheta), \vartheta \in \Theta^{(\varepsilon)}\}.$$

Furthermore, L is a continuous function such that $L(\vartheta_0) = 0$. Thus, since $\text{dist}(\vartheta_0, \Theta^{(\varepsilon)}) \rightarrow 0$ as $\varepsilon \rightarrow 0$, one deduces that $\min_{\vartheta \in \Theta^{(\varepsilon)}} L(\vartheta) \rightarrow 0$ as $\varepsilon \rightarrow 0$. Owing to (\mathbf{I}_w) , this implies that any sequence $(\vartheta^{(\varepsilon)})_\varepsilon$ of $\mathcal{A}^{(\varepsilon)}$ converges to ϑ_0 . This concludes the proof. \square

4.5 Rate of convergence: proof of Theorem 4.2.4

All along this section, we assume (\mathbf{C}_s) and (\mathbf{I}_s) . Our aim is to bound the quantity $\mathbb{E}[|\hat{\vartheta}_{N,n,\gamma} - \vartheta_0|^2]$ where $\hat{\vartheta}_{N,n,\gamma}$ is defined by (4.2.15). Owing to (\mathbf{I}_s) , we are reduced to study

$$\mathbb{E} \left[d \left(\nu_{\vartheta_0}, \nu_{\hat{\vartheta}_{N,n,\gamma}} \right)^q \right] \quad (4.5.1)$$

where $q := 2/\varsigma$ and $\varsigma \in (0, 1]$ is given in (\mathbf{I}_s) . Our strategy of proof is based on the following decomposition

Lemma 4.5.1. *Let $\hat{\vartheta}_{N,n,\gamma}$ be the estimator defined by (4.2.15) and recall that ν_ϑ is defined by (4.2.8) for all $\vartheta \in \Theta$. Then $d \left(\nu_{\vartheta_0}, \nu_{\hat{\vartheta}_{N,n,\gamma}} \right)$ can be decomposed as*

$$d \left(\nu_{\vartheta_0}, \nu_{\hat{\vartheta}_{N,n,\gamma}} \right) \leq 2D_n^{(1)} + 2 \sup_{\vartheta \in \Theta} D_{N,\gamma}^{(2)}(\vartheta) + 2 \sup_{\vartheta \in \Theta} D_{N,\gamma}^{(3)}(\vartheta) \quad (4.5.2)$$

where $D_n^{(1)}$, $D_{N,\gamma}^{(2)}(\vartheta)$, and $D_{N,\gamma}^{(3)}(\vartheta)$ are respectively by

$$D_n^{(1)} := d \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}} \right), \quad (4.5.3)$$

$$D_{N,\gamma}^{(2)}(\vartheta) := d \left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta,\gamma}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Y_{k\gamma}^{\vartheta}} \right), \quad (4.5.4)$$

$$D_{N,\gamma}^{(3)}(\vartheta) := d \left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Y_{k\gamma}^{\vartheta}}, \nu_\vartheta \right). \quad (4.5.5)$$

Proof. Let us write $\hat{\vartheta}$ for $\hat{\vartheta}_{N,n,\gamma}$ throughout the proof in order to ease notations. We first apply the triangular inequality, which yields

$$d \left(\nu_{\vartheta_0}, \nu_{\hat{\vartheta}} \right) \leq d \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}} \right) + d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\hat{\vartheta},\gamma}} \right) + d \left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\hat{\vartheta},\gamma}}, \nu_{\hat{\vartheta}} \right)$$

where $(Y_t)_{t \geq 0}$ is the observation process given by (4.1.2). Next, we invoke the fact that $\hat{\vartheta}$ minimizes the quantity $d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta,\gamma}} \right)$ in Θ , which gives

$$d \left(\nu_{\vartheta_0}, \nu_{\hat{\vartheta}} \right) \leq d \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}} \right) + d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta_0,\gamma}} \right) + \sup_{\vartheta \in \Theta} d \left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta,\gamma}}, \nu_\vartheta \right). \quad (4.5.6)$$

We further split the second term in the right hand side of (4.5.6) as follows:

$$\begin{aligned} & d\left(\frac{1}{n}\sum_{k=0}^{n-1}\delta_{Y_{t_k}}, \frac{1}{N}\sum_{k=0}^{N-1}\delta_{Z_{k\gamma}^{\vartheta_0, \gamma}}\right) \\ & \leq d\left(\frac{1}{n}\sum_{k=0}^{n-1}\delta_{Y_{t_k}}, \nu_{\vartheta_0}\right) + d\left(\nu_{\vartheta_0}, \frac{1}{N}\sum_{k=0}^{N-1}\delta_{Y_{k\gamma}^{\vartheta_0}}\right) + d\left(\frac{1}{N}\sum_{k=0}^{N-1}\delta_{Y_{k\gamma}^{\vartheta_0}}, \frac{1}{N}\sum_{k=0}^{N-1}\delta_{Z_{k\gamma}^{\vartheta_0, \gamma}}\right), \end{aligned} \quad (4.5.7)$$

and resort to a similar decomposition for the third term in the right hand side of (4.5.6). It is then readily checked that plugging (4.5.7) into (4.5.6) we end up with our claim (4.5.2). \square

In the remainder of the section, we shall handle the L^q -moments of $D_n^{(1)}$, $\sup_{\vartheta \in \Theta} |D_{N, \gamma}^{(2)}(\vartheta)|$ and $\sup_{\vartheta \in \Theta} |D_{N, \gamma}^{(3)}(\vartheta)|$ separately, respectively in Sections 4.5.1, 4.5.2 and 4.5.3.

4.5.1 L^q bound on $D_n^{(1)}$

We start this section by giving a notation concerning expectations of empirical measures.

Notation 4.5.1. *Let Y be the solution of equation (4.2.7) and $t \geq 0$. As previously, δ_{Y_t} denotes the Dirac measure at Y_t , considered as a random measure. Then $\mathbb{E}[\delta_{Y_t}]$ is the deterministic measure such that for all continuous and bounded $f : \mathbb{R}^d \rightarrow \mathbb{R}$ we have*

$$\mathbb{E}[\delta_{Y_t}](f) = \mathbb{E}[f(Y_t)].$$

With this notation in mind, we can now deliver our L^q estimate for $D_n^{(1)}$.

Lemma 4.5.2. *Let $D_n^{(1)}$ be the random variable defined by (4.5.3). Then, whenever d is given by $d_{CF, p}$ or d_s , defined respectively by (4.2.21) with $p > (q + d)/2$ and (4.2.22), we have:*

$$\mathbb{E}\left[|D_n^{(1)}|^q\right] \leq C_q \left(n^{-q} + n^{-\frac{q}{2}(2-(2H \vee 1))}\right). \quad (4.5.8)$$

Proof. We decompose $D_n^{(1)}$ as follows:

$$D_n^{(1)} \leq d\left(\nu_{\vartheta_0}, \frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}[\delta_{Y_{t_k}}]\right) + d\left(\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}[\delta_{Y_{t_k}}], \frac{1}{n}\sum_{k=0}^{n-1}\delta_{Y_{t_k}}\right) =: D_n^{(11)} + D_n^{(12)}. \quad (4.5.9)$$

For the term $D_n^{(11)}$, we can use the contractivity assumption (\mathbf{C}_s) on the drift b which implies that two solutions of the SDE (4.1.2) with different initial conditions converge exponentially pathwise to each other as $t \rightarrow +\infty$ (see e.g. [30]). More specifically, we have already seen in (4.3.7) that the arguments of [30] entail $\|Y_t - \bar{Y}_t\|_{L^p(\Omega)} \leq c_1 e^{-c_2 t}$ for two positive constants c_1, c_2 , where we recall that \bar{Y} designates the stationary solution of (4.2.7). Hence for a Lipschitz function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ we easily get the existence of a

constant $C > 0$ such that

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[f(Y_{t_k})] - \nu_{\vartheta_0}(f) \right| \leq \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[|f(Y_{t_k}) - f(\bar{Y}_{t_k})|] \leq \frac{C}{n} \|f\|_{\text{Lip}}. \quad (4.5.10)$$

It remains to take into account the distance d . Recall that we only consider the two distances $d_{CF,p}$ and d_s defined in Subsection 4.2.5. We thus easily deduce from (4.5.10) and the definitions of $d_{CF,p}$ and d_s that there exists a positive constant \tilde{C} such that

$$D_n^{(11)} \leq \max \left\{ d_{CF,p} \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[\delta_{Y_{t_k}}] \right), d_s \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[\delta_{Y_{t_k}}] \right) \right\} \leq \frac{\tilde{C}}{n}. \quad (4.5.11)$$

The term $D_n^{(12)}$ is handled in Proposition 4.5.1 below and specifically in relation (4.5.12). Therefore, plugging (4.5.11) and (4.5.12) into (4.5.9), relation (4.5.8) is proved. \square

Proposition 4.5.1. *Let d be one of the two distances d_s and $d_{CF,p}$ with $p > (q + d)/2$. Then,*

$$\mathbb{E} \left[d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[\delta_{Y_{t_k}}] \right)^q \right] \leq C_q n^{-\frac{q}{2}(2-(2H \vee 1))}. \quad (4.5.12)$$

The proof of the proposition is based on the following lemma:

Lemma 4.5.3. *Recall that $(Y_t)_{t \geq 0}$ is given by (4.1.2). Then for all $q \geq 1$ and for all Lipschitz function $f : \mathbb{R}^d \rightarrow \mathbb{R}$,*

$$\mathbb{E} \left[\left| \frac{1}{n} \sum_{k=0}^{n-1} f(Y_{t_k}) - \mathbb{E}[f(Y_{t_k})] \right|^q \right] \leq C_q \|f\|_{\text{Lip}}^q n^{-\frac{q}{2}(2-(2H \vee 1))}. \quad (4.5.13)$$

Proof. We invoke a concentration result for large time borrowed from [70, Theorem 2.3]. This result asserts that: there exists $C > 0$ such that for all Lipschitz functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and for all $r \geq 0$,

$$\mathbb{P} \left(\frac{1}{n} \sum_{k=0}^{n-1} f(Y_{t_k}) - \mathbb{E}[f(Y_{t_k})] \geq r \right) \leq \exp(-C \|f\|_{\text{Lip}}^{-2} r^2 n^{2-(2H \vee 1)}). \quad (4.5.14)$$

Therefore, one can check that (4.5.13) holds true by plugging inequality (4.5.14) into the classical formula

$$\mathbb{E}[X^q] = \int_0^{+\infty} qx^{q-1} \mathbb{P}(X \geq x) dx,$$

which is valid for any positive random variable X . \square

Proof of Proposition 4.5.1. We will only give details about $d_{CF,p}$ since d_s can be treated exactly along the same lines. Furthermore, since our parameter q is greater than 2, by using Jensen inequality and the linearity of \mathbb{E} into the definition (4.2.21) of $d_{CF,p}$, we get:

$$\mathbb{E} \left[d_{CF,p} \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[\delta_{Y_{t_k}}] \right)^q \right] \leq \int_{\mathbb{R}^d} \mathbb{E} \left[\left| \frac{1}{n} \sum_{k=0}^{n-1} f_{\xi}(Y_{t_k}) - \mathbb{E}[f_{\xi}(Y_{t_k})] \right|^q \right] g_p(\xi) d\xi,$$

where $f_\xi(x) = e^{i\langle \xi, x \rangle}$. Since $\|f_\xi\|_{\text{Lip}} \leq |\xi|$, we thus deduce from Lemma 4.5.3 that

$$\mathbb{E} \left[d_{CF,p} \left(\nu_{\vartheta_0}, \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[\delta_{Y_{t_k}}] \right)^q \right] \leq C_q n^{-\frac{q}{2}(2-(2H \vee 1))} \int_{\mathbb{R}^d} |\xi|^q g_p(\xi) d\xi.$$

The integral in the last inequality is finite owing to the fact that we chose $p > (q+d)/2$. Our claim thus follows. \square

4.5.2 L^q bound on $D_{N,\gamma}^{(2)}$

Our aim in this section is to get an equivalent of relation (4.5.8) for the term $D_{N,\gamma}^{(2)}$ defined by (4.5.4), namely

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} |D_{N,\gamma}^{(2)}(\vartheta)|^q \right] \leq C\gamma^{qH}. \quad (4.5.15)$$

where the distance d in the definition of $D_{N,\gamma}^{(2)}$ is either d_s or $d_{CF,p}$. To this end, resorting to the fact that $d_{CF,p}$ and d_s are both elements of \mathcal{D}_1 , the quantity (4.5.4) can be upper-bounded as follows:

$$\sup_{\vartheta \in \Theta} D_{N,\gamma}^{(2)}(\vartheta) \leq \frac{1}{N} \sum_{k=0}^{N-1} \sup_{\vartheta \in \Theta} |Y_{k\gamma}^\vartheta - Z_{k\gamma}^{\vartheta,\gamma}|.$$

We thus deduce that

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} |D_{N,\gamma}^{(2)}(\vartheta)|^q \right] \leq \mathbb{E} \left[\left(\frac{1}{N} \sum_{k=0}^{N-1} \sup_{\vartheta \in \Theta} |Y_{k\gamma}^\vartheta - Z_{k\gamma}^{\vartheta,\gamma}| \right)^q \right].$$

Recall that $q \geq 2$. Hence a direct application of Jensen's inequality gives

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} |D_{N,\gamma}^{(2)}(\vartheta)|^q \right] \leq \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E} \left[\sup_{\vartheta \in \Theta} |Y_{k\gamma}^\vartheta - Z_{k\gamma}^{\vartheta,\gamma}|^q \right]. \quad (4.5.16)$$

Now according to Proposition 4.3.7 (i) and Proposition 4.A.2 (ii) (see also the proof of Proposition 4.4.1 (i)), it is readily checked that

$$\sup_{k \geq 0} \mathbb{E} \left[\sup_{\vartheta \in \Theta} |Y_{k\gamma}^\vartheta - Z_{k\gamma}^{\vartheta,\gamma}|^q \right] \leq C\gamma^{qH}.$$

Gathering this information with (4.5.16), inequality (4.5.15) is easily deduced.

4.5.3 L^q bound on $D_{N,\gamma}^{(3)}$

The quantity (4.5.5) is the hardest to treat among the terms in our decomposition (4.5.2), due to the fact that we wish to achieve a uniform bound in ϑ . We summarize our analysis in the following lemma.

Lemma 4.5.4. *Let $D_{N,\gamma}^{(3)}$ be the random variable defined by (4.5.5), and assume that d*

is either d_s or $d_{CF,p}$ with $p > (q + d)/2$. Then, we have

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} D_{N,\gamma}^{(3)}(\vartheta)^q \right] \leq C_q (\gamma^{qH} + T^{-\tilde{\eta}}) \quad (4.5.17)$$

with $\tilde{\eta} := \frac{q^2}{2(q+d)}(2 - (2H \vee 1))$ and $T = N\gamma$.

Proof. We will further decompose the term $D_{N,\gamma}^{(3)}$ and then divide our analysis in several steps. First, let us introduce some notations: denote by T the quantity $N\gamma$ and for all $t \in [0, T]$, set $\underline{t} := \inf\{k\gamma \mid k\gamma \leq t < (k+1)\gamma\}$ as we did in the proof of Proposition 4.4.2. With this notations in hand, we have $D_{N,\gamma}^{(3)}(\vartheta) = d(\nu_\vartheta, \frac{1}{T} \int_0^T \delta_{Y_t^\vartheta} dt)$ from which we deduce the following decomposition:

$$\sup_{\vartheta \in \Theta} D_{N,\gamma}^{(3)}(\vartheta) \leq \sup_{\vartheta \in \Theta} D_{N,\gamma}^{(31)}(\vartheta) + \sup_{\vartheta \in \Theta} D_{N,\gamma}^{(32)}(\vartheta) \quad (4.5.18)$$

where

$$D_{N,\gamma}^{(31)}(\vartheta) = d \left(\nu_\vartheta, \frac{1}{T} \int_0^T \delta_{Y_t^\vartheta} dt \right) \quad \text{and} \quad D_{N,\gamma}^{(32)}(\vartheta) = d \left(\frac{1}{T} \int_0^T \delta_{Y_t^\vartheta} dt, \frac{1}{T} \int_0^T \delta_{Y_{\underline{t}}^\vartheta} dt \right).$$

We will now handle those two terms separately:

Step 1: Bound on $D_{N,\gamma}^{(32)}$. As in Section 4.5.2, since $d \in \mathcal{D}_1$ whenever $d = d_s$ or $d = d_{CF,p}$, we have

$$\sup_{\vartheta \in \Theta} D_{N,\gamma}^{(32)}(\vartheta) \leq \frac{1}{T} \int_0^T \sup_{\vartheta \in \Theta} |Y_t^\vartheta - Y_{\underline{t}}^\vartheta| dt. \quad (4.5.19)$$

We now proceed as in Section 4.5.2 in order to get the equivalent of (4.5.16) thanks to Jensen's inequality. We get

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} |D_{N,\gamma}^{(32)}(\vartheta)|^q \right] \leq \frac{1}{T} \int_0^T \mathbb{E} \left[\sup_{\vartheta \in \Theta} |Y_t^\vartheta - Y_{\underline{t}}^\vartheta|^q \right] dt. \quad (4.5.20)$$

In order to bound the right hand side of (4.5.20), we start by recalling the bound (4.4.10) for $Y_t^\vartheta - Y_{\underline{t}}^\vartheta$:

$$|Y_t^\vartheta - Y_{\underline{t}}^\vartheta| \leq \int_{\underline{t}}^t |b_\vartheta(Y_s^\vartheta)| ds + \|\sigma\| |B_t - B_{\underline{t}}|.$$

The drift term above is now bounded thanks to the sublinear growth of b_ϑ given by (4.2.5) and the uniform bound on the L^q moments of Y_t^ϑ given by Proposition 4.A.1. As far as the term $|B_t - B_{\underline{t}}|$ is concerned, we obviously have thanks to (4.1.1) and the fact that $|t - \underline{t}| \leq \gamma$:

$$\mathbb{E} [|B_t - B_{\underline{t}}|^q] \leq C_H \gamma^{qH}.$$

From here, it is readily checked that

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} |Y_t^\vartheta - Y_{\underline{t}}^\vartheta|^q \right] \leq C_H \gamma^{qH}.$$

Plugging this information into (4.5.20) we end up with

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} |D_{N,\gamma}^{(32)}(\vartheta)|^q \right] \leq C_q \gamma^{qH}. \quad (4.5.21)$$

Step 2: Bound on $D_{N,\gamma}^{(31)}(\vartheta)$ for a fixed ϑ . For a fixed value of $\vartheta \in \Theta$, the term $D_{N,\gamma}^{(31)}(\vartheta)$ will be handled similarly to Section 4.5.1. Namely, along the same lines as for relation (4.5.9) we write

$$D_{N,\gamma}^{(31)}(\vartheta) \leq d \left(\nu_\vartheta, \frac{1}{T} \int_0^T \mathbb{E}[\delta_{Y_t^\vartheta}] dt \right) + d \left(\frac{1}{T} \int_0^T \mathbb{E}[\delta_{Y_t^\vartheta}] dt, \frac{1}{T} \int_0^T \delta_{Y_t^\vartheta} dt \right). \quad (4.5.22)$$

Then the first term in the right hand side of (4.5.22) is handled exactly as (4.5.10) in Section 4.5.1, which yields

$$\left| \frac{1}{T} \int_0^T \mathbb{E}[f(Y_t^\vartheta)] dt - \nu_\vartheta(f) \right| \leq \frac{C}{T} \|f\|_{\text{Lip}}. \quad (4.5.23)$$

The second term in the right hand side of (4.5.22) can be upper bounded thanks to a continuous time version of Lemma 4.5.3 (also based on [70, Theorem 2.3] and left to the reader for the sake of conciseness). We get

$$\mathbb{E} \left[\left| \frac{1}{T} \int_0^T \left(f(Y_t^\vartheta) - \mathbb{E}[f(Y_t^\vartheta)] \right) dt \right|^q \right] \leq C_q \|f\|_{\text{Lip}}^q T^{-\frac{q}{2}(2-(2H \vee 1))}. \quad (4.5.24)$$

Therefore putting together (4.5.23) and (4.5.24) and arguing as in Section 4.5.1, we get that for $d = d_{CF,p}$ with $p > (q+d)/2$ or $d = d_s$ and for any $\vartheta \in \Theta$,

$$\mathbb{E} \left[|D_{N,\gamma}^{(31)}(\vartheta)|^q \right] = \mathbb{E} \left[d \left(\nu_\vartheta, \frac{1}{T} \int_0^T \delta_{Y_t^\vartheta} dt \right)^q \right] \leq C_q \left(T^{-q} + T^{-\frac{q}{2}(2-(2H \vee 1))} \right) \quad (4.5.25)$$

where C_q is a positive constant which does not depend on ϑ .

Step 3: Bound on $\sup_{\vartheta \in \Theta} D_{N,\gamma}^{(31)}(\vartheta)$. In order to go from (4.5.25) to a bound for the supremum over Θ , we proceed to a discretization of the parameter space Θ as in Section 4.2.4. Towards this aim, we will use the following notation: for any $\vartheta \in \Theta$, we set

$$\varphi(\vartheta) := d \left(\nu_\vartheta, \frac{1}{T} \int_0^T \delta_{Y_t^\vartheta} dt \right). \quad (4.5.26)$$

Let $\varepsilon > 0$ and recall that $\Theta^{(\varepsilon)} := \{\vartheta_i^{(\varepsilon)} \mid 1 \leq i \leq M_\varepsilon\}$ is defined at the beginning of Subsection 4.2.4 in such a way that $\Theta \subset \bigcup_{i=1}^{M_\varepsilon} B(\vartheta_i^{(\varepsilon)}, \varepsilon)$. Then, for any $\vartheta \in \Theta$,

$$\varphi(\vartheta) \leq |\varphi(\vartheta) - \varphi(\vartheta^{(\varepsilon)})| + |\varphi(\vartheta^{(\varepsilon)})|$$

where $\vartheta^{(\varepsilon)}$ is defined by (4.2.16). Therefore

$$\varphi(\vartheta) \leq |\varphi(\vartheta) - \varphi(\vartheta^{(\varepsilon)})| + \max_{1 \leq i \leq M_\varepsilon} |\varphi(\vartheta_i^{(\varepsilon)})|$$

and finally

$$\begin{aligned} \mathbb{E} \left[\sup_{\vartheta \in \Theta} \varphi(\vartheta)^q \right] &\leq c_q \mathbb{E} \left[\sup_{\vartheta \in \Theta} |\varphi(\vartheta) - \varphi(\vartheta^{(\varepsilon)})|^q \right] + c_q \mathbb{E} \left[\max_{1 \leq i \leq M_\varepsilon} |\varphi(\vartheta_i^{(\varepsilon)})|^q \right] \\ &\leq c_q \mathbb{E} \left[\sup_{\vartheta \in \Theta} |\varphi(\vartheta) - \varphi(\vartheta^{(\varepsilon)})|^q \right] + c_q \sum_{i=1}^{M_\varepsilon} \mathbb{E} \left[|\varphi(\vartheta_i^{(\varepsilon)})|^q \right]. \end{aligned} \quad (4.5.27)$$

Owing to inequality (4.5.25) for a fixed $\vartheta \in \Theta$, we can deduce from (4.5.27) that

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} \varphi(\vartheta)^q \right] \leq c_q \mathbb{E} \left[\sup_{\vartheta \in \Theta} |\varphi(\vartheta) - \varphi(\vartheta^{(\varepsilon)})|^q \right] + c'_q M_\varepsilon \left(T^{-q} + T^{-\frac{q}{2}(2-(2H\nu_1))} \right). \quad (4.5.28)$$

In the remainder of the step, we thus focus on the first right hand term in (4.5.28). Namely we will show the existence of an integrable random variable $\zeta_T > 0$ such that for all $\vartheta_1, \vartheta_2 \in \Theta$ we have

$$|\varphi(\vartheta_1) - \varphi(\vartheta_2)| \leq \zeta_T |\vartheta_1 - \vartheta_2| \quad a.s. \quad (4.5.29)$$

For this purpose, let us split the quantity $|\varphi(\vartheta_1) - \varphi(\vartheta_2)|$ in two terms

$$|\varphi(\vartheta_1) - \varphi(\vartheta_2)| \leq d(\nu_{\vartheta_1}, \nu_{\vartheta_2}) + d \left(\frac{1}{T} \int_0^T \delta_{Y_t^{\vartheta_1}} dt, \frac{1}{T} \int_0^T \delta_{Y_t^{\vartheta_2}} dt \right). \quad (4.5.30)$$

Then, one can show the following inequalities for any $d \in \mathcal{D}_1$

$$d(\nu_{\vartheta_1}, \nu_{\vartheta_2}) \leq C |\vartheta_1 - \vartheta_2| \sup_{\vartheta \in \Theta} \nu_\vartheta (|\cdot|^r) \quad (4.5.31)$$

$$d \left(\frac{1}{T} \int_0^T \delta_{Y_t^{\vartheta_1}} dt, \frac{1}{T} \int_0^T \delta_{Y_t^{\vartheta_2}} dt \right) \leq C |\vartheta_1 - \vartheta_2| \left(\frac{1}{T} \int_0^T \sup_{\vartheta \in \Theta} |Y_s^\vartheta|^r ds \right) \quad (4.5.32)$$

where r is given in assumption (\mathbf{C}_s) and where equation (4.5.31) is obtained by following the proof of Proposition 4.3.8 (see appendix 4.B) for the stationary solutions \bar{Y}^{ϑ_1} and \bar{Y}^{ϑ_2} . Plugging (4.5.31) and (4.5.32) into (4.5.30), we obtain that (4.5.29) holds true with

$$\zeta_T := C \max \left\{ \sup_{\vartheta \in \Theta} \nu_\vartheta (|\cdot|^r), \frac{1}{T} \int_0^T \sup_{\vartheta \in \Theta} |Y_s^\vartheta|^r ds \right\}.$$

By Proposition 4.A.1 (i), we get $\sup_{T>0} \mathbb{E}[\zeta_T^q] < +\infty$. Then plugging (4.5.29) into (4.5.28) we end up with

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} \varphi(\vartheta)^q \right] \leq c_q \sup_{T>0} \mathbb{E}[\zeta_T^q] \varepsilon^q + c'_q M_\varepsilon \left(T^{-q} + T^{-\frac{q}{2}(2-(2H\nu_1))} \right). \quad (4.5.33)$$

For all $\varepsilon > 0$, since Θ is a compact of \mathbb{R}^d , we can choose $M_\varepsilon \leq \frac{C_\Theta}{\varepsilon^d}$ and then if we choose $\varepsilon := T^{-\eta}$ for some $\eta > 0$, we finally get

$$\begin{aligned} \mathbb{E} \left[\sup_{\vartheta \in \Theta} \varphi(\vartheta)^q \right] &\leq c_q \sup_{T>0} \mathbb{E}[\zeta_T^q] T^{-q\eta} + c'_q C_\Theta \left(T^{-q+d\eta} + T^{-\frac{q}{2}(2-(2H\nu_1))+d\eta} \right) \\ &\leq c_q \sup_{T>0} \mathbb{E}[\zeta_T^q] T^{-q\eta} + c'_q C_\Theta T^{-\frac{q}{2}(2-(2H\nu_1))+d\eta}. \end{aligned}$$

It just remains to optimize in η to conclude that:

$$\mathbb{E} \left[\sup_{\vartheta \in \Theta} \varphi(\vartheta)^q \right] \leq C_q T^{-\tilde{\eta}} \quad (4.5.34)$$

with $\tilde{\eta} := \frac{q^2}{2(q+d)}(2 - (2H \vee 1))$. By putting together (4.5.18) with both (4.5.21) and (4.5.34), this concludes the proof of (4.5.17). \square

Let us now conclude the section. Through inequality (4.5.2) and the control of the three right hand side terms, namely (4.5.8), (4.5.15) and (4.5.17), we are in position to conclude that Theorem 4.2.4 holds true.

4.6 Identifiability assumption

In this section we will provide some examples of equations of the form (4.1.2) for which the crucial assumptions (4.2.9) and (4.2.10) are satisfied. We first review briefly the diffusion case in Section 4.6.1, and then give a particular example in the fractional Brownian motion case in Section 4.6.2.

4.6.1 Case of a diffusion process

In this section we consider equation (4.1.2) in the case $H = \frac{1}{2}$, that is when the equation is driven by a d -dimensional Wiener process. Our considerations are summarized in the following proposition.

Proposition 4.6.1. *Consider equation (4.1.2) in the case $H = \frac{1}{2}$. We assume Hypothesis (\mathbf{H}_0) and (\mathbf{C}_w) to be met and call ν_ϑ the invariant measure corresponding to the coefficient b_ϑ . We pick $\vartheta_1, \vartheta_2 \in \Theta$ and set*

$$F = \text{Span}\{\nabla f, \text{ for all } f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}) \text{ with compact support}\}$$

where F is considered as a subspace of $L^2(\nu_{\vartheta_1})$. We assume that $b_{\vartheta_1} - b_{\vartheta_2}$ is not an element of F^\perp in $L^2(\nu_{\vartheta_1})$. Then

$$d(\nu_{\vartheta_1}, \nu_{\vartheta_2}) > 0.$$

If this condition is satisfied for all couples $(\vartheta_1, \vartheta_2) \in \Theta^2$, then (\mathbf{I}_w) holds true.

Proof. For $\vartheta \in \Theta$, let L_ϑ denote the linear operator defined on $\mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ by:

$$L_\vartheta f(x) = \langle \nabla f, b_\vartheta \rangle(x) + \frac{1}{2}(\sigma^* D^2 f(x) \sigma),$$

where $D^2 f$ denote the Hessian matrix of f . By a classical criterion, ν_ϑ is invariant for (4.1.2) when $H = \frac{1}{2}$ if and only if $\nu_\vartheta(L_\vartheta f) = 0$ for any $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ with compact support. As a consequence, $\nu_{\vartheta_1} = \nu_{\vartheta_2}$ if and only if $\nu_{\vartheta_1}((L_{\vartheta_1} - L_{\vartheta_2})f) = 0$ for any compactly supported \mathcal{C}^2 -function $f : \mathbb{R}^d \rightarrow \mathbb{R}$. Now observe that

$$\nu_{\vartheta_1}((L_{\vartheta_1} - L_{\vartheta_2})f) = \int \langle \nabla f(x), b_{\vartheta_1}(x) - b_{\vartheta_2}(x) \rangle \nu_{\vartheta_1}(dx).$$

The result follows. \square

In other words, this result says that in the diffusion setting, the identifiability assumption is true if $\text{Proj}_F(b_\vartheta - b_{\vartheta_0})$ is not the null function for any $\vartheta \neq \vartheta_0$. Notice that this is always true in the one-dimensional case or if b_ϑ is a gradient. Unfortunately, the generalization of this simple characterization to SDEs driven by fBm is far from being straightforward.

4.6.2 Fractional Brownian motion case

In this section we wish to check (\mathbf{I}_s) for some specific examples of equation (4.2.7) and for the distance $d_{\text{CF},p}$. Specifically, we shall consider a family $Y^{\vartheta,\lambda}$ of real valued processes defined by

$$dY_t^{\lambda,\vartheta} = \left[-\vartheta Y_t^{\lambda,\vartheta} + \lambda b_\vartheta(Y_t^{\lambda,\vartheta}) \right] dt + \sigma dB_t \quad (4.6.1)$$

where B is a 1-dimensional fractional Brownian motion. In equation (4.6.1) the quantity λ is a small enough parameter, which is assumed to be known. The estimation procedure is still for ϑ only. The coefficient b_ϑ is bounded together with its derivatives with respect to y and ϑ . The process $Y^{\vartheta,\lambda}$ has to be seen as a small perturbation of a fractional Ornstein-Uhlenbeck process with parameter ϑ . We also assume that ϑ is a 1-dimensional parameter and:

$$\vartheta \in [m, M], \quad \text{with } 0 < m < M < \infty. \quad (4.6.2)$$

Let us start our analysis by the case $X^\vartheta \equiv Y^{0,\vartheta}$, that is the fractional Ornstein-Uhlenbeck process itself, solution of the following equation:

$$dX_t^\vartheta = -\vartheta X_t^\vartheta dt + \sigma dB_t. \quad (4.6.3)$$

It is easily seen that X^ϑ is a centered Gaussian process whose variance is given (see e.g. [36, p.724]) by

$$\mathbb{E}[(X_t^\vartheta)^2] = 2\sigma^2 e^{-\vartheta t} \int_0^t s^{2H-1} \cosh(\vartheta(t-s)) ds. \quad (4.6.4)$$

In this case our assumption (\mathbf{I}_s) is easily satisfied, as shown in the following lemma.

Lemma 4.6.1. *Let $\vartheta \in [m, M]$ as in (4.6.2), and consider the fractional Ornstein-Uhlenbeck process X^ϑ defined by (4.6.3). We call μ_ϑ its invariant measure. Then for all $\vartheta_1, \vartheta_2 \in [m, M]$, we have*

$$d_{\text{CF},p}(\mu_{\vartheta_1}, \mu_{\vartheta_2}) \geq c_{m,M,H} |\vartheta_1 - \vartheta_2|. \quad (4.6.5)$$

Proof. It is well-known (see e.g [11]) that for the fractional Ornstein-Uhlenbeck process we have $\mu_\vartheta = \mathcal{N}(0, \sigma_\vartheta^2)$, with

$$\sigma_\vartheta^2 = \frac{c_H}{\vartheta^{2H}}. \quad (4.6.6)$$

Taking expression (4.2.21) into account, this yields

$$d_{\text{CF},p}^2(\mu_{\vartheta_1}, \mu_{\vartheta_2}) = \int_{\mathbb{R}} \left[\exp\left(-\frac{c_H}{2\vartheta_1^{2H}} \xi^2\right) - \exp\left(-\frac{c_H}{2\vartheta_2^{2H}} \xi^2\right) \right]^2 g_p(\xi) d\xi,$$

from which our claim (4.6.5) is easily proved. Notice that the fact that ϑ is bounded away from 0 is crucial here in order to ensure the continuity of $\vartheta \mapsto \sigma_\vartheta$ in (4.6.6) on the interval $[m, M]$. \square

Let us also state an elementary bound on ordinary differential equations for further use.

Lemma 4.6.2. *Let $f, g : \mathbb{R}_+ \rightarrow \mathbb{R}$ be two functions such that there exist some constants $\kappa, M > 0$ satisfying*

$$f_r \geq \kappa, \quad \text{and} \quad |g_r| \leq M, \quad \text{for all } r \in \mathbb{R}_+. \quad (4.6.7)$$

Let y be the solution of the following differential equation:

$$\dot{y}_t + f_t y_t = g_t. \quad (4.6.8)$$

Then y is uniformly bounded in t and verifies

$$|y_t| \leq \frac{M}{\kappa}.$$

Proof. Equation (4.6.8) admits an explicit solution under the form

$$y_t = \int_0^t \exp\left(-\int_s^t f_r dr\right) g_s ds.$$

Plugging the bounds (4.6.7) into the above expression, we easily get

$$|y_t| \leq M \int_0^t e^{-\kappa(t-s)} ds \leq \frac{M}{\kappa},$$

which is our claim. \square

We now wish to extend Lemma 4.6.1 to the model given by equation (4.6.1). Namely we wish to prove the following proposition.

Proposition 4.6.2. *Let $Y^{\lambda, \vartheta}$ be the process defined by (4.6.1) and consider $p > 3/2$. We assume $\vartheta \in [m, M]$ and $\lambda \in (0, \lambda_0)$ with a small enough $\lambda_0 = \lambda_0(m, M, p)$. Also assume (without loss of generality) that $b_\vartheta, \partial_y b_\vartheta, \partial_\vartheta b_\vartheta, \partial_{yy}^2 b_\vartheta$ are all bounded by 1. Then the following lower bound holds true for any $\vartheta_1, \vartheta_2 \in [m, M]$:*

$$d_{\text{CF}, p}(\nu_{\vartheta_1}, \nu_{\vartheta_2}) \geq c_{m, M, H} |\vartheta_1 - \vartheta_2|. \quad (4.6.9)$$

Proof. Owing to the definition (4.2.21) of the distance $d_{\text{CF}, p}$, we have

$$d_{\text{CF}, p}(\nu_{\vartheta_1}, \nu_{\vartheta_2}) = \left\{ \int_{\mathbb{R}} \left(\mathbb{E}[e^{i\xi \bar{Y}^{\lambda, \vartheta_1}}] - \mathbb{E}[e^{i\xi \bar{Y}^{\lambda, \vartheta_2}}] \right)^2 g_p(\xi) d\xi \right\}^{1/2}$$

We will decompose this quantity as follows:

$$d_{\text{CF}, p}(\nu_{\vartheta_1}, \nu_{\vartheta_2}) \geq I_3^{1/2} - \left(I_2^{1/2} + I_{11}^{1/2} + I_{12}^{1/2} \right), \quad (4.6.10)$$

where the quantities I_{1j} and I_3 are defined by

$$\begin{aligned} I_{1j} &= \int_{\mathbb{R}} \left(\mathbb{E} \left[\exp \left(i\xi \bar{Y}^{\lambda, \vartheta_j} \right) \right] - \mathbb{E} \left[\exp \left(i\xi Y_t^{\lambda, \vartheta_j} \right) \right] \right)^2 g_p(\xi) d\xi \\ I_3 &= \int_{\mathbb{R}} \left(\mathbb{E} \left[\exp \left(i\xi X_t^{\vartheta_1} \right) \right] - \mathbb{E} \left[\exp \left(i\xi X_t^{\vartheta_2} \right) \right] \right)^2 g_p(\xi) d\xi, \end{aligned}$$

and where we recall that the Ornstein-Uhlenbeck process X^ϑ is given by (4.6.3). In equation (4.6.10), we also have

$$\begin{aligned} I_2 &= \int_{\mathbb{R}} \left(\mathbb{E} \left[\exp \left(i\xi Y_t^{\lambda, \vartheta_1} \right) \right] - \mathbb{E} \left[\exp \left(i\xi X_t^{\vartheta_1} \right) \right] \right. \\ &\quad \left. - \mathbb{E} \left[\exp \left(i\xi Y_t^{\lambda, \vartheta_2} \right) \right] + \mathbb{E} \left[\exp \left(i\xi X_t^{\vartheta_2} \right) \right] \right)^2 g_p(\xi) d\xi. \end{aligned}$$

In the definitions above, t is an arbitrarily large time, to be determined later on. Our goal is now to lower bound I_3 and upper bound $I_{1,j}$ and I_2 .

Lower bound for I_3 . In order to lower bound I_3 , we proceed as in Lemma 4.6.1. Indeed, Lemma 4.6.1 stems from a lower bound on

$$\left| \mathbb{E}[e^{i\xi X_\infty^{\vartheta_1}}] - \mathbb{E}[e^{i\xi X_\infty^{\vartheta_2}}] \right|,$$

while we are interested here in a lower bound on

$$\left| \mathbb{E}[e^{i\xi X_t^{\vartheta_1}}] - \mathbb{E}[e^{i\xi X_t^{\vartheta_2}}] \right|,$$

for a fixed t . However it is readily checked from (4.6.4) that there exists $t_0 > 0$ such that for all $t \geq t_0$ we have

$$I_3 \geq c_1 |\vartheta_1 - \vartheta_2|^2 \tag{4.6.11}$$

with a given constant $c_1 > 0$ depending on m, M .

Upper bound for I_{1j} . Recall that both b_ϑ and $\partial_y b_\vartheta$ are bounded by 1. We also assume that λ is small enough so that $\lambda \leq m(1 - \varepsilon)$ with $\varepsilon > 0$. Then it is readily checked that $x \mapsto -\vartheta x + \lambda b_\vartheta(x)$ satisfies the condition (\mathbf{C}_s) . Hence one can see as in (4.3.7) that

$$I_{1j} \leq C e^{-\frac{m\varepsilon}{2}t}.$$

If we wish to have $I_{1j} \leq c_3 |\vartheta_1 - \vartheta_2|^2$, with c_3 arbitrarily small, it is thus sufficient to pick $t \geq t_1$ with $t_1 = C \log \left(\frac{1}{|\vartheta_1 - \vartheta_2|^2} \right)$. In the sequel we choose this time t_1 such that

$$I_{11}^{1/2} + I_{12}^{1/2} \leq \frac{\sqrt{c_1}}{4} |\vartheta_1 - \vartheta_2|. \tag{4.6.12}$$

Upper bound for I_2 . We start by recalling that $X^\vartheta = Y^{0,\vartheta}$. Next we set $R = [\vartheta_1, \vartheta_2] \times [0, \lambda]$ and for $r, \tau \in [0, 1]$ we define

$$a(r, \tau) = Y_t^{0,\vartheta_2} + r \left(Y_t^{0,\vartheta_1} - Y_t^{0,\vartheta_2} \right) + \tau \left(Y_t^{\lambda,\vartheta_2} - Y_t^{0,\vartheta_2} \right) + r\tau \Delta_R Y_t, \quad (4.6.13)$$

where the rectangular increment $\Delta_R Y_t$ is given by

$$\Delta_R Y_t = Y_t^{\lambda,\vartheta_1} - Y_t^{0,\vartheta_1} - Y_t^{\lambda,\vartheta_2} + Y_t^{0,\vartheta_2}. \quad (4.6.14)$$

Notice that I_2 can be expressed as

$$I_2 = \int_{\mathbb{R}} \left(\mathbb{E} [\Delta_R \psi_\xi(Y_t)] \right)^2 g_p(\xi) d\xi, \quad (4.6.15)$$

where ψ_ξ is the oscillating function $e^{i\xi x}$ and $\Delta_R \psi_\xi(Y_t)$ still denotes a rectangular increment as in (4.6.14). Moreover, resorting to the path a introduced in (4.6.13) we get

$$\Delta_R \psi_\xi(Y_t) = \int_{[0,1]^2} \partial_{r\tau}^2 [\psi_\xi(a(r, \tau))] dr d\tau,$$

and computing the differential $\partial_{r\tau}^2 [\psi_\xi(a(r, \tau))]$ explicitly we get

$$\Delta_R \psi_\xi(Y_t) = \int_{[0,1]^2} \left(\psi'_\xi(a(r, \tau)) \partial_{r\tau}^2 a(r, \tau) + \psi''_\xi(a(r, \tau)) \partial_r a(r, \tau) \partial_\tau a(r, \tau) \right) dr d\tau. \quad (4.6.16)$$

Taking into account (4.6.16) and (4.6.15), plus the fact that ψ and its derivatives ψ' , ψ'' are bounded by $|\xi|^2$ and $|\xi|^2 g_p$ is integrable if $p > 3/2$, we get that

$$\begin{aligned} I_2 &\leq C_p \left\{ \mathbb{E} \left[\left(|Y_t^{0,\vartheta_2} - Y_t^{0,\vartheta_1}| + |\Delta_R(Y_t)| \right) \left(|Y_t^{\lambda,\vartheta_2} - Y_t^{0,\vartheta_2}| + |\Delta_R(Y_t)| \right) \right] + \mathbb{E} [|\Delta_R(Y_t)|] \right\}^2 \\ &\leq C_{p,m,M,\varepsilon} \mathbb{E} \left[\|\partial_\lambda Y^\vartheta\|_\infty^2 + \|\partial_\vartheta Y^\vartheta\|_\infty^2 + \|\partial_{\lambda\vartheta}^2 Y^\vartheta\|_\infty^2 \right]^2 \lambda^2 |\vartheta_1 - \vartheta_2|^2. \end{aligned} \quad (4.6.17)$$

In order to bound the right hand side of (4.6.17), we are now reduced to the estimation of $\partial_\lambda Y^{\lambda,\vartheta}$, $\partial_\vartheta Y^{\lambda,\vartheta}$ and $\partial_{\lambda\vartheta}^2 Y^{\lambda,\vartheta}$. Let us now show how to establish a bound for $\partial_\lambda Y^{\lambda,\vartheta}$. To this aim, differentiating formally equation (4.6.1), the process $\partial_\lambda Y^{\lambda,\vartheta}$ solves a system of the form

$$\frac{d[\partial_\lambda Y^{\lambda,\vartheta}]_t}{dt} = \left[-\vartheta + \lambda \partial_y b_\vartheta(Y_t^{\lambda,\vartheta}) \right] \partial_\lambda Y_t^{\lambda,\vartheta} + b_\vartheta(Y_t^{\lambda,\vartheta}). \quad (4.6.18)$$

Whenever ϑ satisfies (4.6.2) and $\lambda \leq m(1 - \varepsilon)$ with $\varepsilon > 0$, equation (4.6.18) above fulfills the hypothesis of Lemma 4.6.2. Hence we get

$$|\partial_\lambda Y_t^{\vartheta,\lambda}| \leq c_{m,M,\varepsilon},$$

uniformly in $t \geq 0$. We let the reader check that the same kind of inequality holds true for $\partial_\vartheta Y^{\lambda,\vartheta}$ and $\partial_{\lambda\vartheta}^2 Y^{\lambda,\vartheta}$ as well, and thus we get

$$\|\partial_\lambda Y^\vartheta\|_\infty + \|\partial_\vartheta Y^\vartheta\|_\infty + \|\partial_{\lambda\vartheta}^2 Y^\vartheta\|_\infty \leq c_{m,M,\varepsilon}.$$

Plugging this inequality into (4.6.17), we end up with

$$|I_2| \leq C_{p,m,M,\varepsilon} \lambda^2 |\vartheta_1 - \vartheta_2|^2.$$

We now choose λ such that $\lambda \leq \frac{1}{4} \left(\frac{c_1}{C_{p,m,M,\varepsilon}} \right)^{1/2}$, where c_1 is defined by (4.6.11) and $\varepsilon = 1/2$. This yields

$$I_2^{1/2} \leq \frac{\sqrt{c_1}}{4} |\vartheta_1 - \vartheta_2|. \quad (4.6.19)$$

We now gather (4.6.19) and (4.6.12) into (4.6.10), which proves our claim with $\lambda_0 = \frac{1}{4} \left(\frac{c_1}{C_{p,m,M,\varepsilon}} \right)^{1/2}$. \square

4.7 Numerical Discussions and Illustrations

In this section, we provide several numerical examples in order to illustrate our main results. To this end, we first investigate several numerical questions which are related to our theoretical results.

Simulated data and Euler scheme: In order to test our results, we have chosen to simulate our observations. Nevertheless, the fractional SDE (4.1.2) cannot be simulated exactly, except in some particular cases. Therefore we have opted for a discretization procedure thanks to a simple first order Euler scheme with very small step $\underline{\gamma}$ (namely $\underline{\gamma} = 10^{-3}$) in order to get a sharp approximation of the true process.

Let us recall that in the additive setting of equation (4.1.2) the simple Euler scheme converges strongly to the true SDE, while this is not true in general in the multiplicative case (see *e.g.* [56]). The convergence of the scheme can be checked for instance through Proposition 4.3.7(i), applied with constant step $\underline{\gamma}$. Furthermore, taking the expectation in Proposition 4.3.7(i) leads to a marginal control of the L^2 -distance between the Euler scheme and the true SDE (with same fBm) of order $\underline{\gamma}^H$ (independently of the horizon). This confirms that our approximation of the observations is reasonable when H is not too small (getting a control for the uniform distance is more involved).

Let us also recall that the increments of the fBm can be simulated through the Wood-Chan method (see [73]), which is based on the embedding of the covariance matrix of the fractional increments in a symmetric circulant matrix (whose eigenvalues can be computed using the Fast Fourier Transform). Therefore up to the approximation of the true SDE detailed above, we now assume that we are given a sequence $(Y_{k\underline{\gamma}})_{k \geq 0}$, where $(Y_t)_{t \geq 0}$ is a solution to (4.2.7) with a given θ_0 . Then we select from this path a subsequence of observations $(Y_{t_k})_{k=1}^n$ where $t_k = k\underline{\gamma}$, which means in particular that we assume $\underline{\gamma}$ to be of the form $k_0 \underline{\gamma}$ with $k_0 \in \mathbb{N}^*$.

Computation of the distance between empirical measures: The theoretical construction of an estimator like (4.2.15) involves in practice the computation of the distance d between the empirical measures of the observed process and of the Euler scheme, for a distance $d \in \mathcal{D}_p$ as defined in (4.2.4). We briefly describe how to compute this kind of distance.

Whenever d is the p -Wasserstein distance, an explicit computation of the distance d in (4.2.15) is possible if the observation Y is 1-dimensional. To this aim, one can use the following representation (see [69]): if μ and ν are two one-dimensional probabilities with *c.d.f* F and G respectively, then for all $p > 0$ we have

$$\mathcal{W}_p^p(\mu, \nu) = \int_0^1 |F^-(t) - G^-(t)|^p dt,$$

where F^- and G^- denote the (left or right) pseudo-inverse of F and G . Moreover, when $\mu = \sum_{i=1}^{n_1} p_i \delta_{x_i}$ and $\nu = \sum_{j=1}^{n_2} q_j \delta_{y_j}$, the computation of the right hand side above can be made explicit through a reordering (and using the fact that F^{-1} and G^{-1} are stepwise constant). In particular, when $n_1 = n_2 = n$ and $p_i = q_j = 1/n$, the Wasserstein distance between μ and ν simply reads

$$\mathcal{W}_p^p(\mu, \nu) = \frac{1}{n} \sum_{i=1}^n |x_{(i)} - y_{(i)}|^p. \quad (4.7.1)$$

We will use this representation in our simulations.

In higher dimension, the computation of the Wasserstein distance generally requires approximation/optimization methods which are out of the scope of this paper. In this context it seems to be numerically simpler to work with an approximation of the distance $d_{CF,p}$ (defined in (4.2.21)), which is also used for our analysis of the rate of convergence. Such an approximation can be obtained by a standard discretization of the integral which appears in the definition (4.2.21).

Minimization of the distance with respect to ϑ : Eventually the implementation of our estimation procedure relies on an optimization problem in order to compute the argmin in (4.2.15). More specifically, in case the estimator is built with a constant step Euler scheme, this consists in minimizing the function

$$\mathcal{F}_d : \vartheta \mapsto d \left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{Y_{t_k}}, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{Z_{k\gamma}^{\vartheta, \gamma}} \right). \quad (4.7.2)$$

In this paper, we only use the naive approach which consists in evaluating the function on a (finite) grid and then computing the minimum on this finite set. This minimization algorithm is clearly restricted to a low dimensional setting. A rigorous investigation of this question would involve more sophisticated optimization methods, such as gradient-type descents.

Observe that a gradient descent should be reasonably straightforward to implement on a distance like $d_{CF,p}$, for which the computation of the gradient has an explicit expression. However, the additional workload in order to carry out this optimization method would clearly lead us too far. We have thus chosen to postpone this study to a future paper.

Numerical illustrations: Let us now turn to some numerical tests, for which we consider two one-dimensional examples.

▷ We begin with the classical Ornstein-Uhlenbeck (OU) process: we consider the process X^ϑ defined by (4.6.3), where ϑ is assumed to sit in a compact interval of $(0, +\infty)$ like in (4.6.2). Let us recall that this case is a toy example since the Gaussian linear structure of the OU-process allows to develop specific estimation methods (on this topic, see [43] or more recently [75] and [10]). The assumptions (\mathbf{H}_0) and (\mathbf{C}_s) are clearly satisfied, whereas (\mathbf{I}_s) follows from Lemma 4.6.1. Using the strategy described in the first part of this section, we get a discretely observed path of Y with the following parameters:

$$\vartheta_0 = 2, \quad \underline{\gamma} = 10^{-3}, \quad \gamma = 10^{-2}, \quad n = 3 \cdot 10^4,$$

and different values of H . In Figure 2, we depict the function \mathcal{F}_d defined in (4.7.2) with $d = d_{CF,2}$ for $H = 0.3$ and $H = 0.7$ respectively. As in the next examples, the Euler scheme is computed with $N = n$ and $\gamma = 10^{-2}$ as specified above. We remark that the function attains its minimum very close to the true value of ϑ . If one takes into

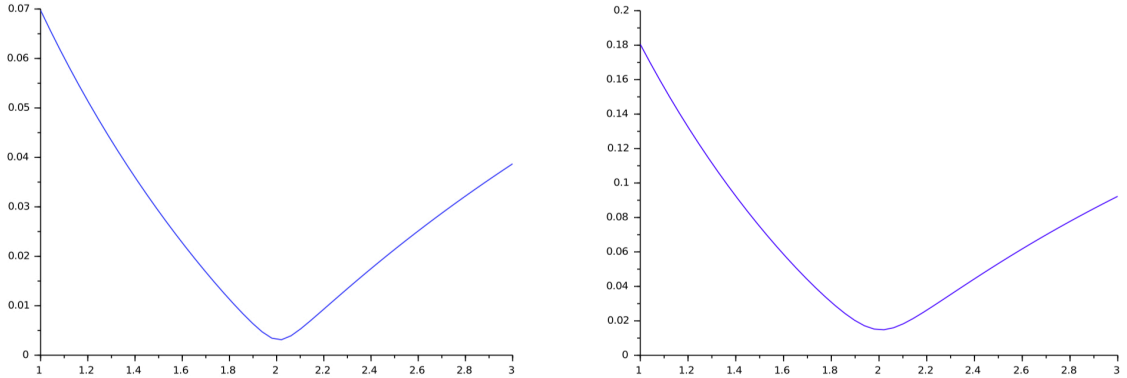


Figure 2 – $\vartheta \mapsto \mathcal{F}_{d_{CF,2}}(\vartheta)$ for $H = 0.3$ (left) and $H = 0.7$ (right).

account the scale difference in the y -axis between the left and right part of Figure 2, it can also be observed that the derivative of \mathcal{F}_d gets smaller for smaller values of H . This is consistent with the fact that $H \mapsto \sigma_\vartheta^2$ in (4.6.6) is an increasing function.

Figure 3 below is devoted to a comparison between the different p -Weisserstein distances as p varies. Namely we fix $H = 0.3$ and we compute the function \mathcal{F}_d defined by (4.7.2) with $d = \mathcal{W}_p$ for different values of p . Notice that in the 1-dimensional case we are considering we can resort to formula (4.7.1), since we have chosen $N = n$. The true parameter is still $\vartheta = 2$. Our distances all perform correctly, although $p = 4$ seems to yields a slightly sharper contrast.

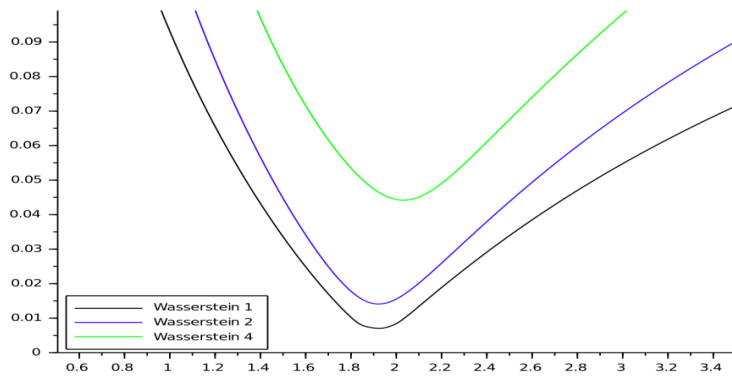


Figure 3 – $\vartheta \mapsto \mathcal{F}_{\mathcal{W}_p}(\vartheta)$, $H = 0.3$.

▷ We now consider a second example with a non linear dependence in ϑ , namely an equation of the form:

$$dY_t^\vartheta = -Y_t^\vartheta(1 + \cos(\vartheta Y_t^\vartheta))dt + dB_t.$$

In this case, we only compute the Wasserstein distance for different values of p with the same choices of parameters. Once again, the minimum of the function $\mathcal{F}_{\mathcal{W}_p}$ is attained close to $\vartheta_0 = 2$. One also observes that the local behavior in the neighborhood of ϑ_0 is similar to the linear case.

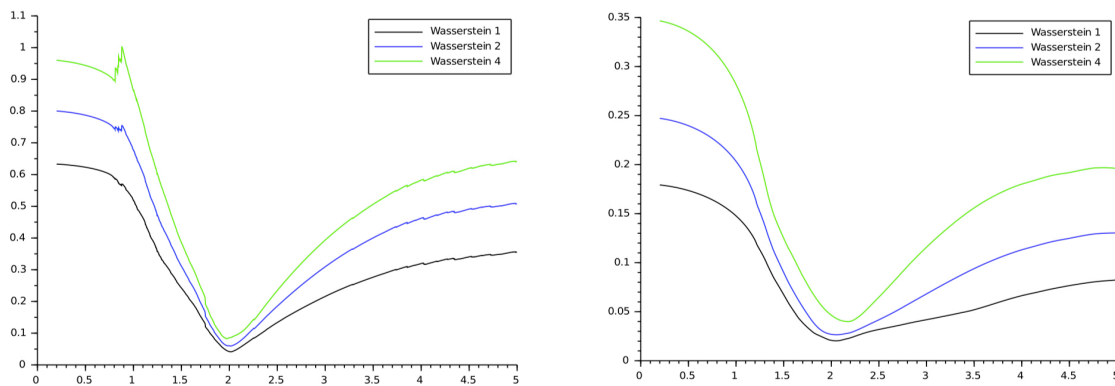


Figure 4 – $\vartheta \mapsto \mathcal{F}_{d_{CF,2}}(\vartheta)$ for $H = 0.3$ (left) and $H = 0.7$ (right).

4.A Tightness and convergence of the occupation measures

In this Appendix, we first show some uniform estimates for Y^ϑ and $Z^{\vartheta,\gamma}$ and their respective occupation measures, as well as some convergence result for the occupation measure of Y^ϑ . Those results are all used in the proof of Proposition 4.3.3 and 4.3.4.

4.A.1 Moment estimates

We start by bounding the moments of Y^ϑ and its occupation measure.

Proposition 4.A.1. *Let Y^ϑ be the unique solution of (4.2.7). Assume (\mathbf{H}_0) and (\mathbf{C}_w) . Then the following inequalities hold true for $p \geq 1$.*

- (i) $\sup_{t \geq 0} \mathbb{E}[\sup_{\vartheta \in \Theta} |Y_t^\vartheta|^p] < +\infty$.
- (ii) $\sup_{\vartheta \in \Theta} \sup_{n \geq 1} \frac{1}{n} \sum_{k=0}^{n-1} |Y_{k\kappa}^\vartheta|^p < +\infty$ a.s.
- (iii) $\sup_{\vartheta \in \Theta} \sup_{t > 0} \frac{1}{t} \int_0^t |Y_t^\vartheta|^p dt < +\infty$ a.s.

Proof. We treat our three items separately.

(i) The proof is done in [36]. It is based on a comparison with the moments of the fractional Ornstein-Uhlenbeck process, similarly to what is done in step (ii) below. Since the constants in (\mathbf{C}_w) are independent from ϑ , we get the uniformity with respect to ϑ .

(ii) Let $p > 0$. By Cauchy-Schwarz inequality it is enough to prove the result with $2p$ instead of p . Let us denote by $(X_t)_{t \geq 0}$ the fractional Ornstein-Uhlenbeck process defined by (4.6.3) with $\vartheta = 1$ and starting from y_0 , i.e.

$$\begin{cases} dX_t = -X_t dt + \sigma dB_t \\ X_0 = y_0. \end{cases} \quad (4.A.1)$$

In this proof, we set $\underline{t} := \inf\{k\kappa \mid k\kappa \leq t < (k+1)\kappa\}$ and $T := n\kappa$. With this notations, similarly to what is done in the proof of Proposition 4.4.2, we have

$$\frac{1}{n} \sum_{k=0}^{n-1} |Y_{k\kappa}^\vartheta|^{2p} = \frac{1}{T} \int_0^T |Y_{\underline{t}}^\vartheta|^{2p} dt.$$

Let us also denote by \tilde{X} the stationary Ornstein-Uhlenbeck process related to (4.A.1). Then the proof of item (ii) is based on the following inequality:

$$\frac{1}{T} \int_0^T |Y_{\underline{t}}^\vartheta|^{2p} dt \leq 3^{2p-1} \left(\frac{1}{T} \int_0^T |Y_{\underline{t}}^\vartheta - X_{\underline{t}}|^{2p} dt + \frac{1}{T} \int_0^T |X_{\underline{t}} - \tilde{X}_{\underline{t}}|^{2p} dt + \frac{1}{T} \int_0^T |\tilde{X}_{\underline{t}}|^{2p} dt \right) \quad (4.A.2)$$

Let us start by bounding the two last term in the right hand side of (4.A.2). First, since $(\tilde{X}_{k\kappa})_{k \geq 0}$ is a stationary Gaussian sequence and since $\mathbb{E}[\tilde{X}_\kappa \tilde{X}_{(n+1)\kappa}] \xrightarrow{n \rightarrow +\infty} 0$ (according

to [13, Theorem 2.3]), a criterion for Gaussian processes (see [53]) gives that $(\tilde{X}_{k\kappa})_{k \geq 0}$ is ergodic. Therefore, we get

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T |\tilde{X}_t|^{2p} dt = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\tilde{X}_{k\kappa}|^{2p} = \mathbb{E}[|\tilde{X}_\kappa|^{2p}] < +\infty, \quad (4.A.3)$$

where the limit holds in the almost sure sense.

In order to treat the second term in the right hand side of (4.A.2), we resort to equation (4.A.1). From there it is readily checked that $\frac{d}{dt}|X_t - \tilde{X}_t|^2 = -2|X_t - \tilde{X}_t|^2$, and thus $|X_t - \tilde{X}_t|^2 = e^{-2t}|y_0 - \tilde{X}_0|^2$. In particular, we have $\lim_{k \rightarrow +\infty} |X_{k\kappa} - \tilde{X}_{k\kappa}|^{2p} = \lim_{k \rightarrow +\infty} e^{-2pk\kappa}|y_0 - \tilde{X}_0|^{2p} = 0$ and the corresponding Cesaro summation tends also to 0, i.e.

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T |X_t - \tilde{X}_t|^{2p} dt = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} |X_{k\kappa} - \tilde{X}_{k\kappa}|^{2p} = 0. \quad (4.A.4)$$

Now it remains to treat the first term in the right hand side of (4.A.2). To this end, we will invoke our hypothesis (\mathbf{C}_w) and a classical argument based on Gronwall's lemma. Namely, note that for all $\varepsilon > 0$,

$$\begin{aligned} \frac{d}{dt}|Y_t^\vartheta - X_t|^2 &= 2\langle Y_t^\vartheta - X_t, b_\vartheta(Y_t^\vartheta) + X_t \rangle \\ &= 2\langle Y_t^\vartheta - X_t, b_\vartheta(Y_t^\vartheta) - b_\vartheta(X_t) \rangle + 2\langle Y_t^\vartheta - X_t, b_\vartheta(X_t) + X_t \rangle \\ &\leq 2\beta - 2\alpha|Y_t^\vartheta - X_t|^2 + \frac{1}{\varepsilon}|Y_t^\vartheta - X_t|^2 + \varepsilon\tilde{C}(1 + |X_t|^{2r}), \end{aligned}$$

where the last inequality stems from (\mathbf{C}_w) , Young's inequality and the linear growth of b_ϑ entailed by (4.2.5). We then set $\varepsilon = 1/\alpha$, which gives

$$\frac{d}{dt}|Y_t^\vartheta - X_t|^2 \leq 2\beta - \alpha|Y_t^\vartheta - X_t|^2 + \frac{\tilde{C}}{\alpha}(1 + |X_t|^{2r}).$$

Hence from Gronwall's lemma, we deduce that

$$|Y_t^\vartheta - X_t|^2 \leq \int_0^t e^{-\alpha(t-s)} \left(2\beta + \frac{\tilde{C}}{\alpha}(1 + |X_s|^{2r}) \right) ds. \quad (4.A.5)$$

Starting from (4.A.5), one can easily get a bound on $|Y_t^\vartheta - X_t|^{2p}$. Namely, due to the fact that $\int_0^t e^{-\alpha(t-s)} ds = \alpha^{-1}(1 - e^{-\alpha t})$, a direct application of Jensen's inequality yields

$$\begin{aligned} |Y_t^\vartheta - X_t|^{2p} &\leq \left(\frac{1 - e^{-\alpha t}}{\alpha} \right)^{p-1} \int_0^t e^{-\alpha(t-s)} \left(2\beta + \frac{\tilde{C}}{\alpha}(1 + |X_s|^{2r}) \right)^p ds \\ &\leq C_p \int_0^t e^{-\alpha(t-s)} (1 + |X_s|^{2pr}) ds. \end{aligned}$$

Then, by using Fubini theorem, it comes

$$\begin{aligned} \frac{1}{T} \int_0^T |Y_{\underline{t}}^\vartheta - X_{\underline{t}}|^{2p} dt &\leq \frac{C_p}{T} \int_0^T \int_0^t e^{-\alpha(t-s)} (1 + |X_s|^{2pr}) ds dt \\ &\leq \frac{C_p}{T} \int_0^T (1 + |X_s|^{2pr}) \int_s^T \mathbf{1}_{[0,\underline{t}]}(s) e^{-\alpha(t-s)} dt ds. \end{aligned} \quad (4.A.6)$$

In addition, for $t \in [0, T]$ we have $\mathbf{1}_{[0,\underline{t}]}(s) e^{-\alpha(t-s)} \leq e^{\alpha\kappa} e^{-\alpha(t-s)}$. Hence integrating the right hand side of (4.A.6) we get

$$\frac{1}{T} \int_0^T |Y_{\underline{t}}^\vartheta - X_{\underline{t}}|^{2p} dt \leq \frac{C_p e^{\alpha\kappa}}{\alpha} \frac{1}{T} \int_0^T (1 + |X_s|^{2pr}) (1 - e^{-\alpha(T-s)}) ds \leq \frac{C_p e^{\alpha\kappa}}{\alpha} \left(1 + \frac{1}{T} \int_0^T |X_s|^{2pr} ds \right).$$

Finally, by the same type of arguments used in the first part of the proof (see (4.A.4)), we can show that $\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T |X_s|^{2pr} ds = \mathbb{E}[|\tilde{X}_0|^{2pr}] < +\infty$. Thus we have obtained

$$\sup_{T > 0} \sup_{\vartheta \in \Theta} \frac{1}{T} \int_0^T |Y_{\underline{t}}^\vartheta - X_{\underline{t}}|^{2p} dt < +\infty. \quad (4.A.7)$$

Plugging (4.A.3), (4.A.4) and (4.A.7) into (4.A.2), our claim (ii) is now proved.

(iii) We let the patient reader check the details for item (iii). It follows the same lines as item (ii), except for the fact that the discretization procedure is avoided. \square

Our next result is an analog of Proposition 4.A.1 for the Euler scheme $Z^{\vartheta, \gamma}$.

Proposition 4.A.2. *Let $Z^{\vartheta, \gamma}$ be the Euler approximation scheme defined by (4.2.13). Assume (\mathbf{H}_0) and (\mathbf{C}_w) . Then, there exists $\gamma_0 > 0$ such that for all $p > 0$,*

- (i) $\sup_{\vartheta \in \Theta} \sup_{\gamma \in (0, \gamma_0]} \limsup_{N \rightarrow +\infty} \mathbb{E}[|Z_{N\gamma}^{\vartheta, \gamma}|^p] < +\infty$.
- (ii) $\sup_{\gamma \in (0, \gamma_0]} \sup_{N \geq 0} \gamma^{p(1-H)} \mathbb{E}[\sup_{\vartheta \in \Theta} |Z_{N\gamma}^{\vartheta, \gamma}|^p] < +\infty$.
- (iii) Fix $\gamma \in (0, 1]$. Then $\sup_{\vartheta \in \Theta} \sup_{N \geq 1} \frac{1}{N} \sum_{k=0}^{N-1} |Z_{k\gamma}^{\vartheta, \gamma}|^p < +\infty$ a.s.

The strategy for the proof of Proposition 4.A.2 is based on a comparison between $Z^{\vartheta, \gamma}$ and the Euler scheme Σ related to the Ornstein-Uhlenbeck process X given by (4.A.1). Namely define Σ recursively by $\Sigma_0 = y_0$ and:

$$\Sigma_{(k+1)\gamma} = (1 - \gamma)\Sigma_{k\gamma} + \sigma((B_{(k+1)\gamma} - B_{k\gamma})), \quad \forall k \geq 0. \quad (4.A.8)$$

We first prove some bounds on Σ itself, which are labeled in the following lemma

Lemma 4.A.1. *Let Σ be the Euler scheme defined by (4.A.8). There exists $\gamma_0 > 0$ such that for all $p > 0$ we have*

- (i) $\sup_{\gamma \in (0, \gamma_0]} \limsup_{N \rightarrow +\infty} \mathbb{E}[|\Sigma_{N\gamma}|^p] < +\infty$.

$$(ii) \sup_{\gamma \in (0, \gamma_0]} \sup_{N \geq 0} \gamma^{p(1-H)} \mathbb{E}[|\Sigma_{N\gamma}|^p] < +\infty.$$

$$(iii) \text{ Fix } \gamma \in (0, \gamma_0]. \text{ Then } \sup_{N \geq 1} \frac{1}{N} \sum_{k=0}^{N-1} |\Sigma_{k\gamma}|^p < +\infty \text{ a.s.}$$

Proof. By Cauchy-Schwarz inequality, it is enough to show the three results for $2p$ instead of p . We treat again the three items separately.

(i) Let $\gamma_0 \in (0, 1)$. By Lemma 2 in [15], we know that

$$\sup_{\gamma \in (0, \gamma_0]} \limsup_{N \rightarrow +\infty} \mathbb{E}[|\Sigma_{N\gamma}|^2] < +\infty. \quad (4.A.9)$$

Since Σ is a Gaussian process, the extension from a second order moment to a moment of order $2p$ is trivial. We thus omit details for sake of conciseness.

(ii) Let $\gamma_0 \in (0, 1)$ and $\gamma \in (0, \gamma_0)$. First note that by induction, we have for all $k \geq 1$:

$$\Sigma_{k\gamma} = (1 - \gamma)^k y_0 + \sigma \sum_{j=0}^{k-1} (1 - \gamma)^j \Delta_{k-j}, \quad (4.A.10)$$

where $\Delta_{k-j} := (B_{(k-j)\gamma} - B_{(k-1-j)\gamma})$. In equation (4.A.10), we apply the triangular inequality for the norm in $L^2(\Omega)$ and we invoke the fact that $\mathbb{E}[(\Delta_k^i)^2]^{1/2} = \gamma^H$ for all $i \in \{1, \dots, d\}$. This yields

$$\begin{aligned} \mathbb{E}[|\Sigma_{k\gamma}|^2]^{1/2} &\leq (1 - \gamma)^k |y_0| + |\sigma| \sum_{j=0}^{k-1} (1 - \gamma)^j \mathbb{E}[|\Delta_{k-j}|^2]^{1/2} \\ &\leq (1 - \gamma)^k |y_0| + |\sigma| \sum_{j=0}^{k-1} (1 - \gamma)^j \sum_{i=1}^d \mathbb{E}[(\Delta_{k-j}^i)^2]^{1/2} \\ &\leq (1 - \gamma)^k |y_0| + d|\sigma| \gamma^H \sum_{j=0}^{k-1} (1 - \gamma)^j \\ &\leq |y_0| + d|\sigma| c \gamma^{H-1}. \end{aligned}$$

Therefore, we easily get that

$$\gamma^{2-2H} \mathbb{E}[|\Sigma_{k\gamma}|^2] \leq 2\gamma^{2-2H} (|y_0|^2 + d^2 |\sigma|^2 \gamma^{2H-2}) \leq 2\gamma_0^{2-2H} |y_0|^2 + 2d^2 |\sigma|^2$$

which gives the following bound

$$\sup_{\gamma \in (0, \gamma_0]} \sup_{N \geq 0} \gamma^{2-2H} \mathbb{E}[|\Sigma_{N\gamma}|^2] < +\infty.$$

Exactly as in item (i), we now resort to the Gaussian nature of Σ in order to conclude that

$$\sup_{\gamma \in (0, \gamma_0]} \sup_{N \geq 0} \gamma^{2p(1-H)} \mathbb{E}[|\Sigma_{N\gamma}|^{2p}] < +\infty, \quad \text{for all } p \geq 1. \quad (4.A.11)$$

(iii) We start from identity (4.A.10). Applying Jensen's inequality we deduce that

$$\begin{aligned} |\Sigma_{k\gamma}|^{2p} &\leq 2^{2p-1}(1-\gamma)^{2kp}|y_0|^{2p} + \left(\frac{2}{\gamma}\right)^{2p-1} |\sigma|^{2p} \sum_{j=0}^{k-1} (1-\gamma)^j |\Delta_{k-j}|^{2p} \\ &\leq 2^{2p-1}|y_0|^{2p} + \left(\frac{2}{\gamma}\right)^{2p-1} |\sigma|^{2p} \sum_{j=0}^{k-1} (1-\gamma)^j |\Delta_{k-j}|^{2p}. \end{aligned} \quad (4.A.12)$$

We have thus obtained that

$$\frac{1}{N} \sum_{k=0}^{N-1} |\Sigma_{k\gamma}|^{2p} \leq (2^{2p-1} + 1)|y_0|^{2p} + \left(\frac{2}{\gamma}\right)^{2p-1} |\sigma|^{2p} \frac{1}{N} \sum_{k=1}^{N-1} \sum_{j=0}^{k-1} (1-\gamma)^j |\Delta_{k-j}|^{2p}.$$

Therefore our claim (iii) is reduced to show that

$$\sup_{N \geq 1} \frac{1}{N} \sum_{k=1}^{N-1} \sum_{j=0}^{k-1} (1-\gamma)^j |\Delta_{k-j}|^{2p} < +\infty. \quad (4.A.13)$$

In order to prove (4.A.13) we make a change of variable $\ell = k - 1 - j$ and apply Fubini's theorem. This gives

$$\begin{aligned} \frac{1}{N} \sum_{k=1}^{N-1} \sum_{j=0}^{k-1} (1-\gamma)^j |\Delta_{k-j}|^{2p} &= \frac{1}{N} \sum_{\ell=0}^{N-2} |\Delta_{\ell+1}|^{2p} \sum_{k=\ell+1}^{N-1} (1-\gamma)^{k-1-\ell} \\ &\leq \frac{1}{\gamma} \frac{1}{N} \sum_{\ell=0}^{N-2} |\Delta_{\ell+1}|^{2p} \leq \frac{d^{p'-1}}{\gamma} \sum_{i=1}^d \frac{1}{N} \sum_{\ell=0}^{N-2} (\Delta_{\ell+1}^i)^{2p}. \end{aligned}$$

Since the sequence $(\Delta_n^i)_{n \geq 1}$ is ergodic for every $i \in \{1, \dots, d\}$, we have

$$\frac{1}{N} \sum_{\ell=0}^{N-2} (\Delta_{\ell+1}^i)^{2p} \xrightarrow{N \rightarrow +\infty} \mathbb{E}[(\Delta_1^i)^{2p}] = c_p \gamma^{2pH}$$

and our claim (4.A.13) follows. This finishes the proof. \square

With those preliminary considerations on Σ in hand, we can now prove our Proposition 4.A.2.

Proof of Proposition 4.A.2. As previously, we will prove the result for $2p$ instead of p . Moreover, for sake of simplicity, we write Z^ϑ instead of $Z^{\vartheta, \gamma}$. According to the dynamics (4.A.8) for Σ and (4.2.13) for Z^ϑ , we have for all $k \geq 1$,

$$|Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^2 = |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^2 + 2\gamma \langle Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}, \Sigma_{(k-1)\gamma} + b_\vartheta(Z_{(k-1)\gamma}^\vartheta) \rangle + \gamma^2 |\Sigma_{(k-1)\gamma} + b_\vartheta(Z_{(k-1)\gamma}^\vartheta)|^2 \quad (4.A.14)$$

In order to treat the second term in (4.A.14), we recast it as

$$\begin{aligned} &\langle Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}, \Sigma_{(k-1)\gamma} + b_\vartheta(Z_{(k-1)\gamma}^\vartheta) \rangle \\ &= \langle Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}, b_\vartheta(Z_{(k-1)\gamma}^\vartheta) - b_\vartheta(\Sigma_{(k-1)\gamma}) \rangle + \langle Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}, \Sigma_{(k-1)\gamma} + b_\vartheta(\Sigma_{(k-1)\gamma}) \rangle \end{aligned}$$

Then, we invoke our condition (\mathbf{C}_w) to bound the first term in the right hand side above and Young's inequality for the second term. Similar manipulations can be performed for the third term in (4.A.14). We let the patient reader check that for some arbitrary parameters $\varepsilon, \varepsilon'$ we get

$$\begin{aligned} & |Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^2 \\ & \leq 2\beta\gamma + (1 - 2\gamma\alpha) |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^2 + \frac{\gamma}{\varepsilon} |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^2 + \gamma\varepsilon |\Sigma_{(k-1)\gamma} + b_\vartheta(\Sigma_{(k-1)\gamma})|^2 \\ & \quad + \frac{\gamma^2}{2\varepsilon'} |\Sigma_{(k-1)\gamma} + b_\vartheta(\Sigma_{(k-1)\gamma})|^2 + \frac{\gamma^2 L^2 \varepsilon'}{2} |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^2. \end{aligned}$$

We now choose $\varepsilon = \frac{1}{\alpha}$ and $\varepsilon' = \frac{\alpha}{\gamma L^2}$, which yields

$$|Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^2 \leq 2\beta\gamma + \left(1 - \frac{\gamma\alpha}{2}\right) |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^2 + \left(\frac{\gamma}{\alpha} + \frac{\gamma^3 L^2}{2\alpha}\right) |\Sigma_{(k-1)\gamma} + b_\vartheta(\Sigma_{(k-1)\gamma})|^2.$$

Observe that under (\mathbf{C}_w) , b_ϑ is sublinear. Hence there exists $C > 0$ depending only on α, β, L such that

$$|Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^2 \leq (1 - \gamma\tilde{\alpha}) |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^2 + C\gamma(1 + \gamma^2)(1 + |\Sigma_{(k-1)\gamma}|^2) \quad (4.A.15)$$

where we have set $\tilde{\alpha} := \alpha/2$ in order to ease our next computations. We now choose $0 < \gamma_0 < 1/\tilde{\alpha}$. By a direct induction, it comes

$$|Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^2 \leq C\gamma(1 + \gamma^2) \sum_{j=0}^{k-1} (1 - \gamma\tilde{\alpha})^{k-1-j} (1 + |\Sigma_{j\gamma}|^2).$$

Finally, applying Jensen's inequality similarly to what is done in (4.A.12), we end up with

$$\begin{aligned} |Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^{2p} & \leq C^p \frac{\gamma}{\tilde{\alpha}^{p'-1}} (1 + \gamma^2)^p \sum_{j=0}^{k-1} (1 - \gamma\tilde{\alpha})^{k-1-j} (1 + |\Sigma_{j\gamma}|^2)^p \\ & \leq C^p \left(\frac{2}{\tilde{\alpha}}\right)^{p-1} \gamma (1 + \gamma^2)^p \sum_{j=0}^{k-1} (1 - \gamma\tilde{\alpha})^{k-1-j} (1 + |\Sigma_{j\gamma}|^{2p}). \end{aligned} \quad (4.A.16)$$

With those preliminary considerations in hand, we now prove the three items in Proposition 4.A.2 separately.

(i) We start from relation (4.A.15). Since $\gamma \in (0, \gamma_0]$, $C(1 + \gamma^2) \leq \tilde{C}$ for a constant $\tilde{C} > 0$

$$|Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^2 \leq (1 - \gamma\tilde{\alpha}) |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^2 + \gamma\tilde{\alpha} \frac{\tilde{C}}{\tilde{\alpha}} (1 + |\Sigma_{(k-1)\gamma}|^2).$$

We now invoke the convexity of $x \mapsto |x|^p$ in order to get

$$\begin{aligned} |Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^{2p} &\leq (1 - \gamma\tilde{\alpha}) |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^{2p} + \gamma\tilde{\alpha} \frac{\tilde{C}_p^p}{\tilde{\alpha}^p} (1 + |\Sigma_{(k-1)\gamma}|^2)^p \\ &\leq (1 - \gamma\tilde{\alpha}) |Z_{(k-1)\gamma}^\vartheta - \Sigma_{(k-1)\gamma}|^{2p} + \tilde{C}_p \gamma (1 + |\Sigma_{(k-1)\gamma}|^{2p}). \end{aligned}$$

Then we take expectations and upper limits in k , which gives

$$\limsup_{k \rightarrow +\infty} \mathbb{E}[|Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^{2p}] \leq \frac{\tilde{C}_p}{\tilde{\alpha}} \left(1 + \limsup_{k \rightarrow +\infty} \mathbb{E}[|\Sigma_{k\gamma}|^{2p}] \right).$$

With this inequality in hand and Proposition 4.A.1 (i), inequality (i) is proved.

(ii) First, we take successively the supremum over Θ and the expectation in (4.A.16). We then multiply by $\gamma^{2p(1-H)}$ and perform the same kind of manipulations as for (i). With the help of Proposition 4.A.1 (ii), we get that

$$\sup_{\gamma \in (0, \gamma_0]} \sup_{k \geq 0} \gamma^{2p(1-H)} \mathbb{E}[\sup_{\vartheta \in \Theta} |Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^{2p}] < +\infty. \quad (4.A.17)$$

Having achieved the upper bound (4.A.17), our claim (ii) now stems from another direct application of Proposition 4.A.1 (ii).

(iii) Here, we just need to sum (4.A.16) for k from 0 to $N - 1$ and divide by N . Then, we use Fubini theorem on the right hand side combined with Proposition 4.A.1 (iii) to conclude that $\sup_{\vartheta \in \Theta} \sup_{N \geq 1} \frac{1}{N} \sum_{k=0}^{N-1} |Z_{k\gamma}^\vartheta - \Sigma_{k\gamma}|^{2p} < +\infty$. Finally, by using again Proposition 4.A.1 (iii), the result follows. \square

4.A.2 Proof of Proposition 4.3.3

We will focus on the proof of (4.3.8) only, the discrete counterpart of Proposition 4.3.3 being obtained by similar argument. We also note that (4.3.8) cannot be obtained as a consequence of Birkhoff's theorem. Indeed, Birkhoff's theorem would yield a $\bar{\nu}_\vartheta$ -a.s. convergence in (4.3.8), where $\bar{\nu}_\vartheta$ is the invariant measure alluded to in Remark 4.2.3. In order to avoid any reference to the support of the invariant measure $\bar{\nu}_\vartheta$, our relation (4.3.8) gives a weak convergence result which does not include any condition on the initial value of Y_0^ϑ .

In order to prove (4.3.8), we first observe that according to (4.3.9), the family of measures $\{\frac{1}{t} \int_0^t \delta_{Y_s^\vartheta} ds; t \geq 0\}$ is *a.s.* tight. Therefore it only remains to prove that any limit is ν_ϑ . We now focus on the convergence of $\frac{1}{t} \int_0^t \delta_{Y_s^\vartheta} ds$. To this aim, we consider an additional family of probability measures on $\mathcal{C}([0, \infty), \mathbb{R}^d)$ in the following way:

$$\pi_t^\vartheta = \frac{1}{t} \int_0^t \delta_{Y_{s+}^\vartheta} ds$$

where $Y_{s+}^\vartheta = (Y_{s+u}^\vartheta)_{u \geq 0}$. We will first prove that $\{\pi_t^\vartheta; t \geq 0\}$ is an *a.s.* tight family in the set $\mathcal{M}_1(\mathcal{C}([0, \infty), \mathbb{R}^d))$, where we recall that the notation \mathcal{M}_1 is introduced in Section 4.2.1. The tightness of $\{\pi_t^\vartheta; t \geq 0\}$ can be handled in the following way: we have seen

that $\{\frac{1}{t} \int_0^t \delta_{Y_s^\vartheta} ds; t \geq 0\}$ is *a.s.* tight in $\mathcal{M}_1(\mathbb{R}^d)$. Therefore a classical criterion (see *e.g.* [4, Theorem 8.3]) ensures that $\{\pi_t^\vartheta; t \geq 0\}$ is *a.s.* tight if for every positive T, η and ε , there exists $\delta > 0$ such that for all $t_0 \in [0, T]$,

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \frac{1}{\delta} \mathbf{1}_{\{\sup_{u \in [t_0, t_0 + \delta]} |Y_{s+u}^\vartheta - Y_{s+t_0}^\vartheta| \geq \varepsilon\}} ds \leq \eta \quad a.s. \quad (4.A.18)$$

Moreover, inequality (4.A.18) holds true as long as there exist some positive r and ρ such that

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \sup_{u \in [t_0, t_0 + \delta]} |Y_{s+u}^\vartheta - Y_{s+t_0}^\vartheta|^r ds \leq C_{r,T} \delta^{1+\rho} \quad a.s. \quad (4.A.19)$$

Let us now prove (4.A.19). On the interval $[t_0, t_0 + \delta]$, we have

$$|Y_{s+u}^\vartheta - Y_{s+t_0}^\vartheta| \leq \int_{t_0}^{t_0 + \delta} |b_\vartheta(Y_{s+u}^\vartheta)| du + \sup_{u \in [t_0, t_0 + \delta]} |B_{s+u} - B_{s+t_0}|. \quad (4.A.20)$$

Using Jensen's inequality and Fubini's Theorem, we get that for any $r > 0$,

$$\frac{1}{t} \int_0^t \left(\int_{t_0}^{t_0 + \delta} |b_\vartheta(Y_{s+u}^\vartheta)| du \right)^r ds \leq \delta^{r-1} \left(\frac{1}{t} \int_0^{t+t_0 + \delta} |b(Y_s^\vartheta)|^r ds \right).$$

Furthermore, since b_ϑ is sublinear according to (\mathbf{C}_w) , it follows from (4.3.10) that

$$\limsup_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \left(\int_{t_0}^{t_0 + \delta} |b_\vartheta(Y_{s+u}^\vartheta)| du \right)^r ds \leq C \delta^{r-1} \quad (4.A.21)$$

for a constant $C > 0$. On the other hand, by the ergodicity of the increments of the fBm and its self-similarity,

$$\frac{1}{t} \int_0^t \sup_{u \in [t_0, t_0 + \delta]} |B_{s+u} - B_{s+t_0}|^r ds \xrightarrow{t \rightarrow +\infty} \mathbb{E} \left[\sup_{u \in [0, \delta]} |B_u|^r \right] = C_r \delta^{Hr}, \quad (4.A.22)$$

where the limit in (4.A.22) holds in the *a.s.* sense. Choosing $r > \sup\{2, 1/H\}$ in (4.A.21) and (4.A.22), then plugging (4.A.21) and (4.A.22) into (4.A.20), we get that (4.A.19) is satisfied. We have thus proved the *a.s.* tightness of $\{\pi_t^\vartheta; t \geq 0\}$.

The second step is then to show that any limiting distribution of $\left(\frac{1}{t} \int_0^t \delta_{Y_{s+}^\vartheta} ds\right)$ is necessarily the law of a stationary solution \bar{Y}^ϑ to SDE (4.2.7). This step in turn implies the result by uniqueness of the stationary solutions.

Let $\left(\frac{1}{t_n} \int_0^{t_n} \delta_{Y_{s+}^\vartheta} ds\right)_n$ be a (pathwise) convergent sequence with limiting distribution μ where $\{t_n; n \geq 1\}$ is an increasing sequence converging to $+\infty$. We first prove that μ is the law of a stationary process. Namely, for any bounded functional $F : \mathcal{C}([0, \infty), \mathbb{R}^d) \rightarrow \mathbb{R}$, we have

$$\mu(F \circ \theta_T) - \mu(F) = \lim_{n \rightarrow +\infty} \frac{1}{t_n} \int_0^{t_n} [F(Y_{s+T+}^\vartheta) - F(Y_{s+}^\vartheta)] ds.$$

However, a simple change of variables reveals that *a.s.* we have

$$\lim_{n \rightarrow +\infty} \frac{1}{t_n} \int_0^{t_n} [F(Y_{s+T+}^\vartheta) - F(Y_{s+}^\vartheta)] ds = \lim_{n \rightarrow +\infty} \frac{1}{t_n} \left(\int_{t_n}^{t_n+T} F(Y_{s+}^\vartheta) ds - \int_0^T F(Y_{s+}^\vartheta) ds \right) = 0.$$

We thus easily get that μ is stationary. Now, let us prove that μ is the law of a solution to (4.2.7). Without loss of generality, we can say that a process $(x_t)_{t \geq 0}$ is a solution to (4.2.7) if $x_t - x_0 - \int_0^t b_\vartheta(x_u) du$ is a fBm. In other words, we have to prove that $\mu \circ G^{-1}$ is the law of a fBm where $G(x) = x_t - x_0 - \int_0^t b_\vartheta(x_u) du$. Since G is continuous for the *u.s.c* topology, it is readily checked that

$$\mu \circ G^{-1} = \lim_{n \rightarrow +\infty} \frac{1}{t_n} \int_0^{t_n} \delta_{G(Y_{s+}^\vartheta)} ds.$$

In addition, by construction $G(Y_{s+}^\vartheta) = B_{s+} - B_s$. Hence the fact that $\mu \circ G^{-1}$ is the law of a fBm follows again from the ergodicity of the increments of the fBm.

Summarizing our considerations, we have proved that μ is a stationary measure related to the system (4.2.7). Therefore we have $\mu = \mathcal{L}((\bar{Y}_t^\vartheta)_{t \geq 0})$, which concludes the proof.

4.B Proof of Proposition 4.3.8

For sake of conciseness, we will focus on the proof of Proposition 4.3.8 (i). The proof of item (ii) relies on the same kind of tools, plus the discrete computations invoked in the proof of Proposition 4.A.2. In order to prove item (i), let us consider $t \geq 1$ and a parameter $\rho > 0$ to be chosen later on. An easy elaboration of (4.3.14) shows that

$$e^{\rho t} |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 = \int_0^t e^{\rho s} \left(\rho |Y_s^{\vartheta_1} - Y_s^{\vartheta_2}|^2 + 2 \langle b_{\vartheta_1}(Y_s^{\vartheta_1}) - b_{\vartheta_2}(Y_s^{\vartheta_2}), Y_s^{\vartheta_1} - Y_s^{\vartheta_2} \rangle \right) ds. \quad (4.B.1)$$

In addition, one can write

$$b_{\vartheta_1}(Y_s^{\vartheta_1}) - b_{\vartheta_2}(Y_s^{\vartheta_2}) = b_{\vartheta_1}(Y_s^{\vartheta_1}) - b_{\vartheta_1}(Y_s^{\vartheta_2}) + b_{\vartheta_1}(Y_s^{\vartheta_2}) - b_{\vartheta_2}(Y_s^{\vartheta_2}).$$

We now combine the assumption (C_s) (including the contraction property, the fact $x \mapsto b_\vartheta(x)$ is uniformly Lipschitz in ϑ and the fact that $\partial_\vartheta b_\vartheta(x)$ has polynomial growth) and Young's inequality $|ab| \leq \frac{1}{2\varepsilon} |a|^2 + \frac{\varepsilon}{2} |b|^2$ for an arbitrary $\varepsilon > 0$. This yields the existence of a constant $L > 0$ such that

$$\langle b_{\vartheta_1}(Y_s^{\vartheta_1}) - b_{\vartheta_2}(Y_s^{\vartheta_2}), Y_s^{\vartheta_1} - Y_s^{\vartheta_2} \rangle \leq \left(-\alpha + \frac{L^2 \varepsilon}{2} \right) |Y_s^{\vartheta_1} - Y_s^{\vartheta_2}|^2 + \frac{C |\vartheta_1 - \vartheta_2|^2 (1 + |Y_s|^r)^2}{2\varepsilon}.$$

Plugging this inequality into (4.B.1) and setting $\varepsilon = L^2/\alpha$ and $\rho = \alpha/2$, we have thus obtained

$$e^{\rho t} |Y_t^{\vartheta_1} - Y_t^{\vartheta_2}|^2 \leq C |\vartheta_1 - \vartheta_2|^2 \int_0^t e^{\rho s} (1 + |Y_s^{\vartheta_2}|^{2r}) ds. \quad (4.B.2)$$

Thus, using Fubini's theorem, one deduces that

$$\begin{aligned} \frac{1}{t} \int_0^t |Y_s^{\vartheta_1} - Y_s^{\vartheta_2}|^2 ds &\leq \frac{C|\vartheta_1 - \vartheta_2|^2}{t} \int_0^t (1 + |Y_s^{\vartheta_2}|^{2r} \int_u^t e^{\rho(s-u)} ds du) \\ &\leq C_\rho |\vartheta_1 - \vartheta_2|^2 \left(1 + \frac{1}{t} \int_0^t (1 + |Y_s^{\vartheta_2}|^{2r} du) \right). \end{aligned}$$

Hence a direct application of inequality (4.3.9) yields the existence of a random variable $C = C(\omega)$ such that for all $(\vartheta_1, \vartheta_2) \in \Theta^2$ we have

$$\sup_{t \geq 1} \frac{1}{t} \int_0^t |Y_s^{\vartheta_1} - Y_s^{\vartheta_2}|^2 ds \leq C |\vartheta_1 - \vartheta_2|^2.$$

This concludes the proof of Proposition 4.3.8 (i) for a distance $d \in \mathcal{D}_2$. To extend the result to any $p \geq 2$, one can apply Jensen's inequality to (4.B.2) and follow the same lines as for $d \in \mathcal{D}_2$.

5. Perspectives

5.1 Sur l'ergodicité de dynamiques discrètes fractionnaires

Rappelons que le premier travail de cette thèse correspondant au chapitre 2 porte sur le comportement en temps long de dynamiques discrètes à mémoire de la forme

$$X_{n+1} = F(X_n, \Delta_{n+1}) \tag{5.1.1}$$

où $(\Delta_n)_{n \in \mathbb{N}}$ est un processus gaussien stationnaire et ergodique (voir l'article [71]). En s'inspirant de Hairer [36], Fontbona-Panloup [26] et Deya-Panloup-Tindel [22] dans un cadre continu et sous de bonnes hypothèses sur F et sur le bruit gaussien, nous avons construit une structure markovienne au dessus de (5.1.1), montré l'existence et l'unicité d'une mesure invariante et donné une borne sur la vitesse de convergence de la loi du processus vers cette mesure (pour la distance en variation totale). La vitesse obtenue dépend de la décroissance asymptotique de la fonction de covariance du bruit gaussien Δ (ou de façon équivalente de celle des coefficients intervenants dans la représentation en moyenne mobile de Δ).

Un premier objectif d'extension de ce travail serait d'établir la vitesse lorsque la dynamique (5.1.1) correspond au schéma d'Euler d'une EDS dirigée par un mouvement brownien fractionnaire (mBf) de paramètre de Hurst $H > 1/2$. En effet, si dans le cas $H < 1/2$, on retrouve les ordres des vitesses du cadre continu (voir Hairer [36]), il ne nous est actuellement pas possible d'obtenir le même type de résultat dans le cas $H > 1/2$. Précisément, le contrôle de l'opérateur d'inversion de la représentation en moyenne mobile (et des b_k associés) n'est pas suffisant pour conclure. À la lecture des arguments du cadre continu (voir Lemme 5.1 de contrôle utilisé par Hairer [36]), il semble qu'une idée pourrait être de tirer profit de la régularité locale du mBf lorsque $H > 1/2$ (ou du noyau de Mandelbrot associé). Cet outil n'est d'apparence pas nécessaire dans le cadre discret où les objets sont bien définis pour tout H . Néanmoins, mimer l'argument (de type intégration par parties) du continu pourrait probablement permettre d'obtenir le résultat souhaité.

Une autre question plus générale en lien avec ce travail serait d'étudier (toujours pour le schéma d'Euler) comment se comporte la vitesse de convergence avec le pas de discrétisation. En l'état, la dépendance en le pas du schéma semble difficile d'accès, cela nécessiterait une étude plus précise et spécifique du schéma.

5.2 Sur la concentration pour des EDS fractionnaires

Le deuxième chapitre de cette thèse porte sur la concentration en temps long à la fois pour des fonctionnelles de la solution d'une EDS fractionnaire additive (de paramètre de Hurst $H \in (0, 1)$) sur un intervalle $[0, T]$ et pour des fonctionnelles d'observations discrètes de ce processus (voir l'article [70]). Le résultat général obtenu est ensuite appliqué à des fonctionnelles spécifiques liées aux mesures d'occupations (discrètes ou continues) de la solution de l'EDS.

Rappelons l'EDS considérée :

$$dX_t = b(X_t)dt + \sigma B_t \quad (5.2.1)$$

où $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une fonction au minimum continue, $\sigma \in \mathbb{M}_{d \times d}$ et B est un mBf de paramètre $H \in (0, 1)$. Un sujet d'intérêt est la relaxation d'une hypothèse sur le drift b . Dans le chapitre 3, on suppose que le drift b est fortement contractant, c'est-à-dire que $\langle x - y, b(x) - b(y) \rangle \leq -\alpha|x - y|^2$. Cette hypothèse induit que deux trajectoires de (5.2.1) partant de deux conditions initiales distinctes convergent l'une vers l'autre à vitesse exponentielle. Cependant, sous l'hypothèse plus faible $\langle x - y, b(x) - b(y) \rangle \leq \beta - \alpha|x - y|^2$, l'ergodicité de l'EDS (5.2.1) est démontrée (voir [36]), mais aucun résultat de concentration en temps long sous cette hypothèse plus faible n'existe à ce jour. À notre connaissance, il semble qu'il existe peu de résultats dans le cadre diffusif (c'est-à-dire lorsque B est le mouvement brownien standard dans (5.2.1)) sans cette hypothèse de forte contractivité. Un premier travail pourrait être d'étudier la relaxation de cette hypothèse dans le cas diffusif pour lequel les solutions sont markoviennes. En particulier, nous pourrions commencer par relaxer l'hypothèse de forte contractivité en une hypothèse de forte contractivité en dehors d'un compact et faible contractivité dans ce compact (correspondant à l'hypothèse (\mathbf{H}'_b) section 1.2.2).

5.3 Sur l'estimation du drift pour des EDS fractionnaires

Le dernier chapitre, qui expose un travail effectué en collaboration avec Panloup et Tindel [60], porte sur l'estimation paramétrique du drift (non linéaire) pour une EDS fractionnaire additive de paramètre de Hurst $H \in (0, 1)$. Nous utilisons une estimation par minimum de contraste basée sur l'identification de la mesure invariante (dont une approximation est construite à partir d'observations discrètes de l'EDS). Nous démontrons la consistance des estimateurs considérés et obtenons des bornes non asymptotiques sur l'erreur quadratique. Nos résultats sont illustrés par des simulations numériques. Enfin, nous montrons sur une classe d'exemples que l'hypothèse d'identifiabilité relative à ce problème d'estimation (intrinsèquement liée à la mesure invariante) est satisfaite.

Nous aimerions maintenant améliorer deux aspects. Premièrement, le récent article de Comte-Marie [17] traite de l'estimation non paramétrique du terme de drift (en dimension 1) lorsque $H > 1/2$ en utilisant les propriétés d'ergodicité de l'EDS. Cependant, les auteurs supposent que le processus est observé continûment ce qui n'est pas le cas en général. Un aspect important à étudier est de voir si une estimation non paramétrique du drift est possible en se basant sur des observations discrètes.

Deuxièmement, d'un point de vue numérique, nous pensons pouvoir développer les résultats numériques exposés dans le chapitre 4. En effet, notre estimateur est défini comme l'argument minimum d'une fonction de contraste. La fonction à minimiser est complexe car elle dépend d'une distance sur l'espace des mesures de probabilités. Pour certains choix de distances (comme $d_{CF,p}$ définie en (1.3.40)), il semble cependant que le calcul du gradient soit possible ce qui permettrait d'obtenir un algorithme d'optimisation de type gradient et d'étudier ses propriétés.

Bibliography

- [1] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.
- [2] Benjamin Arras, Guillaume Mijoule, Guillaume Poly, and Yvik Swan. A new approach to the stein-tikhomirov method: with applications to the second wiener chaos and dickman convergence. *arXiv preprint arXiv:1605.06819*, 2016.
- [3] Rachid Belfadli, Khalifa Es-Sebaiy, and Youssef Ouknine. Parameter estimation for fractional ornstein-uhlenbeck processes: non-ergodic case. *arXiv preprint arXiv:1102.5491*, 2011.
- [4] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [5] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.
- [6] Richard C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2:107–144, 2005. Update of, and a supplement to, the 1986 original.
- [7] Peter J. Brockwell and Richard A. Davis. *Time series: theory and methods*. Springer Series in Statistics. Springer, New York, 2006. Reprint of the second (1991) edition.
- [8] A. Brouste, C. Cai, and M. Kleptsyna. Asymptotic properties of the MLE for the autoregressive process coefficients under stationary Gaussian noise. *Math. Methods Statist.*, 23(2):103–115, 2014.
- [9] Alexandre Brouste and Marina Kleptsyna. Kalman type filter under stationary noises. *Systems Control Lett.*, 61(12):1229–1234, 2012.
- [10] Alexandre Brouste, Marina Kleptsyna, and Alexandre Popier. Design for estimation of the drift parameter in fractional diffusion systems. *Stat. Inference Stoch. Process.*, 15(2):133–149, 2012.
- [11] Boris Buchmann and Claudia Klüppelberg. Maxima of stochastic processes driven by fractional Brownian motion. *Adv. in Appl. Probab.*, 37(3):743–764, 2005.

- [12] Philippe Carmona, Laure Coutin, and Gérard Montseny. Stochastic integration with respect to fractional Brownian motion. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(1):27–68, 2003.
- [13] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima. Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.*, 8:no. 3, 14, 2003.
- [14] Jean-François Coeurjolly. Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Inference Stoch. Process.*, 4(2):199–227, 2001.
- [15] Serge Cohen and Fabien Panloup. Approximation of stationary solutions of Gaussian driven stochastic differential equations. *Stochastic Process. Appl.*, 121(12):2776–2801, 2011.
- [16] Serge Cohen, Fabien Panloup, and Samy Tindel. Approximation of stationary solutions to SDEs driven by multiplicative fractional noise. *Stochastic Process. Appl.*, 124(3):1197–1225, 2014.
- [17] Fabienne Comte and Nicolas Marie. Nonparametric estimation in fractional sde. *Statistical Inference for Stochastic Processes*, pages 1–24, 2018.
- [18] Laure Coutin. An introduction to (stochastic) calculus with respect to fractional brownian motion. In *Séminaire de Probabilités XL*, pages 3–65. Springer, 2007.
- [19] A. M. Davie. Differential equations driven by rough paths: an approach via discrete approximation. *Appl. Math. Res. Express. AMRX*, (2):Art. ID abm009, 40, 2007.
- [20] L. Decreusefond and A. S. Üstünel. Stochastic analysis of the fractional Brownian motion. *Potential Anal.*, 10(2):177–214, 1999.
- [21] A. Deya, A. Neuenkirch, and S. Tindel. A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(2):518–550, 2012.
- [22] Aurélien Deya, Fabien Panloup, and Samy Tindel. Rate of convergence to equilibrium of fractional driven stochastic differential equations with rough multiplicative noise. *The Annals of Probability*, 47(1):464–518, 2019.
- [23] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *The Annals of Probability*, 32(3B):2702–2732, 2004.
- [24] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *The Annals of Probability*, 23(4):1671–1691, 1995.
- [25] Robert J Elliott. A continuous time kronecker’s lemma and martingale convergence. 2001.

- [26] Joaquin Fontbona and Fabien Panloup. Rate of convergence to equilibrium of fractional driven stochastic differential equations with some multiplicative noise. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(2):503–538, 2017.
- [27] Neville J. Ford, Dmitry V. Savostyanov, and Nickolai L. Zamarashkin. On the decay of the elements of inverse triangular Toeplitz matrices. *SIAM J. Matrix Anal. Appl.*, 35(4):1288–1302, 2014.
- [28] Peter K. Friz and Martin Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.
- [29] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.
- [30] María J. Garrido-Atienza, Peter E. Kloeden, and Andreas Neuenkirch. Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion. *Appl. Math. Optim.*, 60(2):151–172, 2009.
- [31] Mathieu Gourcy and Liming Wu. Logarithmic Sobolev inequalities of diffusions for the L^2 metric. *Potential Anal.*, 25(1):77–102, 2006.
- [32] Paolo Guasoni. No arbitrage under transaction costs, with fractional Brownian motion and beyond. *Mathematical Finance*, 16(3):569–582, 2006.
- [33] Toufik Guendouzi. Transportation inequalities for SDEs involving fractional Brownian motion and standard Brownian motion. *Adv. Model. Optim.*, 14(3):615–634, 2012.
- [34] M. Hairer and A. Ohashi. Ergodic theory for SDEs with extrinsic memory. *The Annals of Probability*, 35(5):1950–1977, 2007.
- [35] M. Hairer and N. S. Pillai. Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(2):601–628, 2011.
- [36] Martin Hairer. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *The Annals of Probability*, 33(2):703–758, 2005.
- [37] Martin Hairer. Ergodic properties of a class of non-Markovian processes. In *Trends in stochastic analysis*, volume 353 of *London Math. Soc. Lecture Note Ser.*, pages 65–98. Cambridge Univ. Press, Cambridge, 2009.
- [38] Martin Hairer and Natesh S. Pillai. Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths. *The Annals of Probability*, 41(4):2544–2598, 2013.
- [39] Yaozhong Hu and David Nualart. Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statist. Probab. Lett.*, 80(11-12):1030–1038, 2010.
- [40] Yaozhong Hu, David Nualart, and Hongjuan Zhou. Drift parameter estimation for nonlinear stochastic differential equations driven by fractional brownian motion. *Stochastics*, pages 1–25, 2019.

- [41] Jacques Istas and Gabriel Lang. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(4):407–436, 1997.
- [42] Jae-Hyung Jeon, Vincent Tejedor, Stas Burov, Eli Barkai, Christine Selhuber-Unkel, Kirstine Berg-Sørensen, Lene Oddershede, and Ralf Metzler. In vivo anomalous diffusion and weak ergodicity breaking of lipid granules. *Physical review letters*, 106(4):048103, 2011.
- [43] M. L. Kleptsyna and A. Le Breton. Statistical analysis of the fractional Ornstein-Uhlenbeck type process. *Stat. Inference Stoch. Process.*, 5(3):229–248, 2002.
- [44] Andrei Nikolaevich Kolmogorov and Yu A Rozanov. On strong mixing conditions for stationary gaussian processes. *Theory of Probability & Its Applications*, 5(2):204–208, 1960.
- [45] S. C. Kou. Stochastic modeling in nanoscale biophysics: subdiffusion within proteins. *The Annals of Applied Statistics*, 2(2):501–535, 2008.
- [46] Alain Le Breton. Filtering and parameter estimation in a simple linear system driven by a fractional Brownian motion. *Statist. Probab. Lett.*, 38(3):263–274, 1998.
- [47] M Ledoux. Concentration, transportation and functional inequalities. *Preprint*, 2002.
- [48] Michel Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.
- [49] R. Sh. Liptser and A. N. Shiryaev. *Theory of martingales*, volume 49 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1989. Translated from the Russian by K. Dzjaparidze [Kacha Dzjaparidze].
- [50] Benoit B Mandelbrot and John W Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.
- [51] K. Marton. Bounding \bar{d} -distance by informational divergence: a method to prove measure concentration. *The Annals of Probability*, 24(2):857–866, 1996.
- [52] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3):556–571, 1996.
- [53] Gisiro Maruyama. The harmonic analysis of stationary stochastic processes. *Mem. Fac. Sci. Kyūsyū Univ. A.*, 4:45–106, 1949.
- [54] M. N. Mishra and B. L. S. Prakasa Rao. Nonparametric estimation of trend for stochastic differential equations driven by fractional Brownian motion. *Stat. Inference Stoch. Process.*, 14(2):101–109, 2011.

- [55] Andreas Neuenkirch and Samy Tindel. A least square-type procedure for parameter estimation in stochastic differential equations with additive fractional noise. *Stat. Inference Stoch. Process.*, 17(1):99–120, 2014.
- [56] Ivan Nourdin. Schémas d’approximation associés à une équation différentielle dirigée par une fonction höldérienne; cas du mouvement brownien fractionnaire. *C. R. Math. Acad. Sci. Paris*, 340(8):611–614, 2005.
- [57] David J Odde, Elly M Tanaka, Stacy S Hawkins, and Helen M Buettner. Stochastic dynamics of the nerve growth cone and its microtubules during neurite outgrowth. *Biotechnology and Bioengineering*, 50(4):452–461, 1996.
- [58] Diethelm I. Ostry. Synthesis of accurate fractional Gaussian noise by filtering. *IEEE Trans. Inform. Theory*, 52(4):1609–1623, 2006.
- [59] Fabien Panloup and Alexandre Richard. Sub-exponential convergence to equilibrium for gaussian driven stochastic differential equations with semi-contractive drift. *arXiv preprint arXiv:1804.01348*, 2018.
- [60] Fabien Panloup, Samy Tindel, and Maylis Varvenne. A general drift estimation procedure for stochastic differential equations with additive fractional noise. *arXiv preprint arXiv:1903.10769*, 2019.
- [61] B. L. S. Prakasa Rao. *Statistical inference for fractional diffusion processes*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2010.
- [62] Sebastian Riedel. Transportation-cost inequalities for diffusions driven by Gaussian processes. *Electron. J. Probab.*, 22:Paper No. 24, 26, 2017.
- [63] Philippe Rigollet and Jan-Christian Hütter. High dimensional statistics. *Lecture notes (MIT)*, 2017.
- [64] M. Rosenblatt. Central limit theorem for stationary processes. pages 551–561, 1972.
- [65] Bruno Saussereau. Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion. *Bernoulli*, 18(1):1–23, 2012.
- [66] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.
- [67] Ciprian A. Tudor and Frederi G. Viens. Statistical aspects of the fractional stochastic calculus. *Ann. Statist.*, 35(3):1183–1212, 2007.
- [68] Ciprian A. Tudor and Frederi G. Viens. Variations and estimators for self-similarity parameters via Malliavin calculus. *Ann. Probab.*, 37(6):2093–2134, 2009.
- [69] S. S. Vallander. Calculations of the Vasserštejn distance between probability distributions on the line. *Teor. Veroyatnost. i Primenen.*, 18:824–827, 1973.
- [70] Maylis Varvenne. Concentration inequalities for stochastic differential equations with additive fractional noise. *arXiv preprint arXiv:1901.03502*, 2019.

- [71] Maylis Varvenne. Rate of convergence to equilibrium for discrete-time stochastic dynamics with memory. *arXiv preprint arXiv:1709.06413*, To appear in *Bernoulli*.
- [72] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [73] Andrew T. A. Wood and Grace Chan. Simulation of stationary Gaussian processes in $[0, 1]^d$. *J. Comput. Graph. Statist.*, 3(4):409–432, 1994.
- [74] Li-ming Wu and Zheng-liang Zhang. Talagrand’s T_2 -transportation inequality w.r.t. a uniform metric for diffusions. *Acta Math. Appl. Sin. Engl. Ser.*, 20(3):357–364, 2004.
- [75] Weilin Xiao, Weiguo Zhang, and Weidong Xu. Parameter estimation for fractional Ornstein-Uhlenbeck processes at discrete observation. *Appl. Math. Model.*, 35(9):4196–4207, 2011.
- [76] Jie Xiong. *An introduction to stochastic filtering theory*, volume 18 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2008.

Ergodicité des équations différentielles stochastiques fractionnaires et problèmes liés

RÉSUMÉ

Dans cette thèse, nous nous intéressons à trois problèmes en lien avec l'ergodicité de dynamiques aléatoires à mémoire (discrètes ou continues) et tout particulièrement des Équations Différentielles Stochastiques (EDS) dirigées par un mouvement brownien fractionnaire.

Le premier chapitre porte sur l'étude du comportement en temps long pour une classe générale de dynamiques aléatoires discrètes dirigées par un processus gaussien stationnaire ergodique. En s'inspirant des travaux de Hairer (2005), Fontbona-Panloup (2017), Deya-Panloup-Tindel (2019) sur l'ergodicité des EDS fractionnaires, nous construisons une structure markovienne au dessus de la dynamique considérée, nous démontrons l'existence et l'unicité d'une mesure invariante puis nous donnons une borne sur la vitesse de convergence de la loi du processus vers cette mesure. La vitesse obtenue dépend du comportement asymptotique de la fonction de covariance du processus gaussien qui dirige la dynamique (ou plus précisément de celui des coefficients intervenant dans sa représentation en moyenne mobile).

Le deuxième chapitre expose des résultats sur la concentration en temps long à la fois pour des fonctionnelles de la solution d'une EDS fractionnaire additive sur un intervalle $[0, T]$ et pour des fonctionnelles d'observations discrètes de ce processus. Ce résultat général est ensuite appliqué à des fonctionnelles spécifiques liées aux mesures d'occupations (discrètes ou continues) de la solution de l'EDS.

Le dernier chapitre, dont les résultats utilisent ceux du chapitre 2, est un travail effectué en collaboration avec Panloup et Tindel qui porte sur l'estimation paramétrique du drift (non linéaire) pour une EDS fractionnaire additive. Nous utilisons une estimation par minimum de contraste basée sur l'identification de la mesure invariante (dont une approximation est construite à partir d'observations discrètes de l'EDS). Nous démontrons la consistance des estimateurs considérés et obtenons des bornes non asymptotiques sur l'erreur quadratique. Nos résultats sont illustrés par des simulations numériques. Enfin, nous montrons sur une classe d'exemples que l'hypothèse d'identifiabilité relative à ce problème d'estimation (intrinsèquement liée à la mesure invariante) est satisfaite.

ABSTRACT

In this thesis, we focus on three problems related to the ergodicity of stochastic dynamics with memory (in a discrete-time or continuous-time setting) and especially of Stochastic Differential Equations (SDE) driven by fractional Brownian motion.

In the first chapter, we study the long-time behavior of a general class of discrete-time stochastic dynamics driven by an ergodic and stationary Gaussian noise. Following the seminal paper written by Hairer (2005) on the ergodicity of fractional SDE (see also Fontbona-Panloup (2017) and Deya-Panloup-Tindel (2019)), we first build a Markovian structure above the dynamics, we show existence and uniqueness of the invariant distribution and then we exhibit some upper-bounds on the rate of convergence to equilibrium in terms of the asymptotic behavior of the covariance function of the Gaussian noise (or more precisely, of the asymptotic behavior of the coefficients appearing in its moving average representation).

The second chapter establishes long-time concentration inequalities both for functionals of the whole solution on an interval $[0, T]$ of an additive fractional SDE and for functionals of discrete-time observations of this process. Then, we apply this general result to specific functionals related to discrete and continuous-time occupation measures of the process.

The last chapter, which uses the results developed in Chapter 2, is a joint work with Panloup and Tindel which focuses on the parametric estimation of the (non linear) drift term in an additive fractional SDE. We use a minimum contrast estimation based on the identification of the invariant distribution (for which we build an approximation from discrete-time observations of the SDE). We provide consistency results as well as non-asymptotic estimates of the corresponding quadratic error. Some of our results are illustrating through numerical discussions. We also give some examples for which the identifiability condition related to our estimation procedure (intrinsically linked to the invariant distribution) is fulfilled.