



Étude théorique et numérique d'équations cinétiques stochastiques multi-échelles

Shmuel Rakotonirina-Ricquebourg

► To cite this version:

Shmuel Rakotonirina-Ricquebourg. Étude théorique et numérique d'équations cinétiques stochastiques multi-échelles. Equations aux dérivées partielles [math.AP]. Université de Lyon, 2021. Français. NNT : 2021LYSE1142 . tel-03301561

HAL Id: tel-03301561

<https://theses.hal.science/tel-03301561>

Submitted on 27 Jul 2021

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.



N d'ordre NNT : 2021LYSE1142

THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON

Opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale 512
InfoMaths

Spécialité de doctorat : Mathématiques

Soutenue publiquement le 05/07/2021, par :

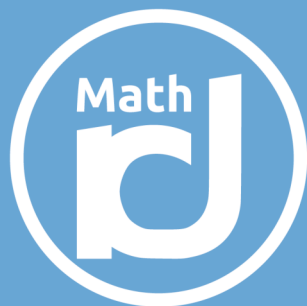
RAKOTONIRINA--RICQUEBOURG
Shmuel

Étude théorique et numérique d'équations cinétiques stochastiques multi-échelles

Devant le jury composé de :

DE BOUARD, Anne
CERRAI, Sand
GOUDENÈGE, Ludovi
TEXIER, Benjamin
BRÉHIER Charles-Édouard
VOVELLE Julien

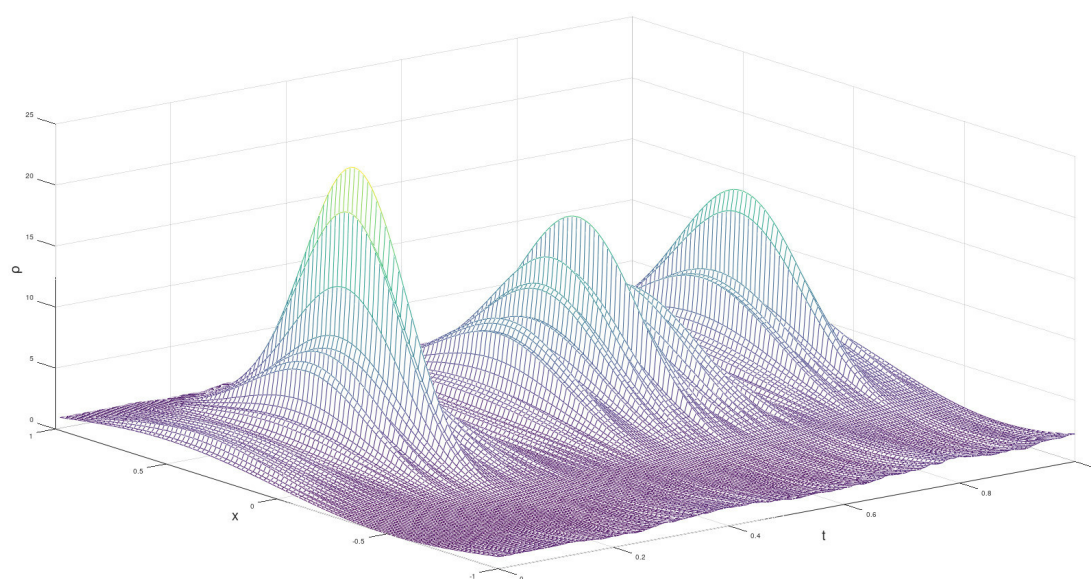
Présidente, examinatrice Directrice de Recherche, École Polytechnique
Rapporteure Professor, University of Maryland
Rapporteur Chargé de Recherche, CentraleSupélec
Examineur Maître de Conférences, Université Lyon 1
Directeur de thèse Chargé de Recherche, Université Lyon 1
Codirecteur de thèse Directeur de Recherche, ENS Lyon



Institut
Camille
Jordan

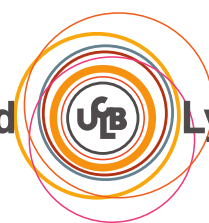
Laboratoire de recherche en mathématiques Lyon/Saint-Étienne

— Étude théorique et numérique d'équations —
— cinétiques stochastiques multi-échelles —



Shmuel Rakotonirina-Ricquebourg

Thèse de doctorat



Université Claude Bernard – LYON 1

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Jean-François MORNEX
Directeur Général des Services	M. Pierre ROLLAND

COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI

Résumé - Abstract

Étude théorique et numérique d'équations cinétiques stochastiques multi-échelles

Cette thèse est dédiée à l'étude d'une classe de systèmes multi-échelles modélisés par une Équation aux Dérivées Partielles Stochastique (EDPS) linéaire cinétique ou une Équation Différentielle Stochastique (EDS). On étudie ces systèmes d'un point de vue théorique et numérique, dans deux régimes asymptotiques : le régime de moyennisation et le régime d'approximation-diffusion.

Les deux premiers chapitres énoncent les principaux résultats théoriques de cette thèse. On montre à chaque fois la convergence de la composante lente du système d'EDPS considéré vers la solution d'une équation de diffusion munie d'un terme source qui dépend du régime asymptotique. Dans le premier chapitre, on considère le régime d'approximation-diffusion, dans lequel le terme source de l'équation limite est un terme diffusif au sens probabiliste (processus de Wiener). Dans le deuxième, on considère le régime de moyennisation, dans lequel le terme source de l'équation limite est la moyenne du terme source de l'EDPS originale.

Les deux derniers chapitres constituent la partie numérique de cette thèse. De manière générale, un schéma numérique peut être consistant avec un système multi-échelle à un paramètre $\varepsilon > 0$ fixé mais se révéler inefficace dans le régime asymptotique $\varepsilon \rightarrow 0$, à cause d'un terme raide dans le modèle. À l'opposé, certains schémas préservent l'asymptotique : ils sont consistants à $\varepsilon > 0$ fixé, convergent vers un schéma limite quand $\varepsilon \rightarrow 0$ et ce schéma limite est consistant avec l'équation limite. Le but des deux derniers chapitres est de proposer, respectivement pour les EDS et les EDPS considérées, des schémas préservant l'asymptotique, de les étudier et d'illustrer numériquement leur efficacité.

Mots-clefs : équations cinétiques, Équations Différentielles Stochastiques, EDP Stochastiques, approximation-diffusion, principe de moyennisation, méthodes multi-échelles, schéma préservant l'asymptotique

Theoretical and numerical study of multiscale stochastic kinetic equations

In this thesis, we study a class of slow-fast systems modeled by kinetic linear Stochastic Partial Differential Equations (SPDEs) or Stochastic Differential Equations (SDEs). We study these systems from theoretical and a numerical points of view in two asymptotic regimes: the averaging regime and the diffusion approximation regime.

The first two chapters state the main theoretical contributions of this work. We prove the convergence of the slow component of the considered SPDEs to the solution of a diffusion equation with a source term depending on the asymptotic regime. The first chapter focuses on the diffusion approximation regime, where the source term of the limiting equation is a stochastic diffusive term (Wiener process). The second chapter focuses on the averaging regime, where the limiting source term is the average of the original source term.

The last two chapters are devoted to the numerical part of this work. In general, a numerical scheme which is consistent with a multiscale system for a fixed parameter $\varepsilon > 0$ can perform badly in the asymptotic regime $\varepsilon \rightarrow 0$ due to the presence of stiff terms in the model. On the contrary, some schemes are asymptotic preserving: they are consistent for fixed $\varepsilon > 0$, converge to some limiting schemes when $\varepsilon \rightarrow 0$ and the limiting scheme is consistent with the limiting equation. The goal of the last two chapters is to design asymptotic preserving schemes, respectively for the class of SDEs and SPDEs we consider. We also analyze these schemes and illustrate numerically their efficiency.

Keywords: kinetic equations, slow-fast systems, Stochastic Differential Equations, Stochastic PDEs, diffusion approximation, averaging principle, multiscale methods, Asymptotic Preserving schemes

Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, F-69622 Villeurbanne Cedex

Contents

Résumé - Abstract	7
Introduction (français)	15
1 Description du modèle	15
2 Convergence de (1) dans le régime d'approximation-diffusion	17
2.1 Contexte	17
2.2 Stratégie de la preuve dans [DV12]	18
2.3 Stratégie de la preuve dans le chapitre I	19
3 Convergence de (1) dans le régime de moyennisation	20
4 Schémas préservant l'asymptotique pour des EDS	21
4.1 Contexte	21
4.2 Principales contributions du chapitre III	22
4.3 Estimées d'erreurs dans le régime de moyennisation	24
5 Schémas préservant l'asymptotique pour des EDP stochastiques cinétiques	25
5.1 Contexte dans le cas déterministe	25
5.2 Contexte dans le cas probabiliste en approximation-diffusion	26
5.3 Principales contributions du chapitre IV	27
6 Perspectives	27
6.1 Résultats à prouver	27
6.2 Vers d'autres modèles	28
Introduction (English version)	29
1 Description of the model	29
2 Convergence of (1) in the diffusion approximation regime	31
2.1 Context	31
2.2 Strategy for the proof in [DV12]	32
2.3 Strategy for the proof in Chapter I	33
3 Convergence of (1) in the averaging regime	34
4 Asymptotic preserving schemes for SDEs	35
4.1 Context	35
4.2 Main contributions of Chapter III	36
4.3 Error estimates in the averaging regime	38
5 Asymptotic preserving schemes for kinetic stochastic PDEs	39
5.1 Context in the deterministic case	39
5.2 Context in the stochastic case, in the diffusion approximation regime	40
5.3 Main contributions of Chapter IV	41
6 Perspectives	41

6.1	Results to investigate	41
6.2	Towards other models	42
I	Convergence in the diffusion approximation regime with an unbounded driving process	43
I.1	Introduction	43
I.2	Assumptions and main result	45
I.2.1	Driving random term	46
I.2.2	Main result	48
I.2.3	Strategy of the proof of Theorem I.2.1	48
I.3	Preliminary results	50
I.3.1	Resolvent operator	50
I.3.2	Properties of the covariance operator	53
I.3.3	Behavior of the stopping time for the driving process	54
I.3.4	Pathwise solutions	55
I.3.5	Estimate in $L^2(\mathcal{M}^{-1})$	57
I.4	Martingale problems and perturbed test functions	60
I.4.1	Generator and martingales	60
I.4.2	The perturbed test functions method	67
I.5	Dynamics associated with the limiting equation	73
I.6	Tightness of the coupled stopped process	75
I.6.1	Proof of the first claim Item (i)	76
I.6.2	Proof of the second claim Item (ii)	79
I.7	Identification of the limit points	82
I.7.1	Convergence of the auxiliary process	82
I.7.2	Convergence of the stopped martingale problems	83
I.7.3	Identification of the limit point	85
I.7.4	Convergence of the unstopped process	88
I.8	Strong convergence	89
II	Convergence in the averaging regime with an unbounded driving process	91
II.1	Introduction	91
II.2	Assumptions and main results	93
II.2.1	Functional spaces	93
II.2.2	Assumptions	93
II.2.3	Statement of the main results	95
II.2.4	Strategy of the proof	96
II.3	Preliminary results	97
II.3.1	Well-posedness of the equation for fixed ε	97
II.3.2	Estimate on the solution in $L^2(\mathcal{M}^{-1})$	99
II.3.3	Uniqueness of the solution for the limiting equation	102
II.4	Description of the perturbed test functions method	103
II.4.1	Poisson equation and auxiliary tools	105
II.4.2	Construction of the first corrector	106
II.4.3	Construction of the second corrector	106
II.4.4	Verification of (II.4.5)	108
II.5	Martingale problems	108

II.5.1	Martingale problem for the equation at fixed ε	108
II.5.2	An equivalent formulation for the limit equation	109
II.6	Convergence in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$	110
II.6.1	Asymptotic behavior of the stopping time	110
II.6.2	Tightness	112
II.6.3	Identification of the limit	114
II.7	Convergence in $L^2([0, T], L^2(\mathbb{T}^d))$	116
II.7.1	Convergence of ρ^ε	116
II.7.2	Convergence of f^ε	117
III	Asymptotic preserving schemes for Stochastic Differential Equations	119
III.1	Introduction	119
III.2	Slow-fast SDE models and their limits	122
III.2.1	The averaging regime	123
III.2.2	The diffusion approximation regime	125
III.3	Numerical discretization and asymptotic preserving schemes	127
III.3.1	Asymptotic preserving schemes: definition and properties	128
III.3.2	An example of AP scheme in the averaging regime	131
III.3.3	An example of AP scheme in the diffusion approximation regime	134
III.4	Uniform accuracy and error analysis	139
III.4.1	Uniform accuracy in the averaging regime (proof of Theorem III.3.4)	139
III.4.2	Auxiliary error estimates	141
III.4.3	Estimates on solutions of Kolmogorov equations	151
III.4.4	On the diffusion approximation regime	154
III.5	Qualitative numerical experiments	159
III.5.1	Illustration in the averaging regime	160
III.5.2	Illustration in the diffusion approximation regime	160
III.6	Quantitative illustration in the averaging regime	164
III.6.1	Total error	164
III.6.2	Averaging error at the discrete time level	167
III.7	Quantitative illustration in the diffusion approximation regime	170
III.7.1	Total error	170
III.7.2	Diffusion approximation error at the discrete time level	171
III.8	Conclusion	173
III.A	Appendix - Derivation of the limiting models	174
III.A.1	Sketch of proof of Proposition III.2.1 (averaging regime)	174
III.A.2	Sketch of proof of Proposition III.2.2 (diffusion approximation regime)	175
IV	Asymptotic preserving schemes for kinetic stochastic PDEs	177
IV.1	Introduction	177
IV.2	Models	179
IV.2.1	Averaging regime	181
IV.2.2	Diffusion approximation regime	182
IV.3	Numerical scheme: splitting procedure	182
IV.3.1	Asymptotic preserving scheme for the deterministic part	183
IV.3.2	Numerical schemes for the stochastic part	185
IV.4	Numerical experiments	187
IV.4.1	Illustration in the averaging regime	188

IV.4.2 Illustration in the diffusion approximation regime	192
IV.5 Conclusion	200

List of Figures

III.1	Evolution of the AP scheme (III.3.4) (left), the crude scheme (III.3.7) (right), and the reference scheme (III.5.1) (averaging regime), with $\Delta t = 0.004$.	160
III.2	Evolution of the AP scheme (III.3.4), the crude scheme (III.3.7) and the reference scheme (III.5.1) (averaging regime), with $\Delta t = 0.004$ and $\varepsilon = 0.001$.	161
III.3	Evolution of the AP scheme (III.3.19) (left), the crude scheme (III.3.14) (right), and the reference scheme (III.5.2) (diffusion approximation regime regime, first example), with $\Delta t = 0.004$.	161
III.4	Evolution of the AP scheme (III.3.19), the crude scheme (III.3.14) and the reference scheme (III.5.2) (diffusion approximation regime, first example), with $\Delta t = 0.004$ and $\varepsilon = 0.01$.	162
III.5	Evolution of the AP schemes (III.3.23) and (III.3.19) (left), the crude scheme (III.5.4) with $\theta \neq \theta'$ (right), and the reference scheme (III.5.3) (diffusion approximation regime, first example with $\sigma(x) = x$), $\Delta t = 0.004$ and $\varepsilon = 0.01$.	163
III.6	Evolution of the AP scheme (III.3.21), the crude scheme (III.3.14), and the reference scheme (III.5.5) for the second example (III.2.22) (diffusion approximation regime, second example), with $\Delta t = 0.004$ and $\varepsilon = 0.01$.	163
III.7	Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^9$ Monte-Carlo samples.	165
III.8	Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = \cos(2\pi x)$, $\varphi(x) = x^2$ and $M = 10^6$ Monte Carlo samples.	165
III.9	Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = x$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^6$ Monte Carlo samples.	166
III.10	Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^6$ Monte Carlo samples.	166
III.11	Total error (III.6.1) of the non AP scheme (III.3.7) with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^6$ Monte Carlo samples.	167
III.12	Averaging error at the discrete time level (III.6.2) of the AP scheme (III.3.4), with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^8$ Monte Carlo samples.	168
III.13	Averaging error at the discrete time level (III.6.2) of the AP scheme (III.3.4), with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^6$ Monte Carlo samples.	169
III.14	Averaging error at the discrete time level (III.6.2) of the non AP scheme (III.3.7), with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^6$ Monte Carlo samples.	169
III.15	Total error (III.7.1) of the AP scheme (III.3.12), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^9$ Monte Carlo samples.	171

III.16	Total error (III.7.1) of the non AP scheme (III.3.14), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^7$ Monte Carlo samples.	172
III.17	Diffusion approximation error at the discrete time level (III.7.2) for the AP scheme (III.3.12), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$, $M = 10^6$ Monte Carlo samples, $\Delta t \in \{10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$ and $\varepsilon \in \{4^i 10^{-4} \mid i \in \llbracket 0, 6 \rrbracket\}$	172
III.18	Diffusion approximation error at the discrete time level (III.7.2) for the non AP scheme (III.3.14), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$, $M = 10^6$ Monte Carlo samples, $\Delta t \in \{10^{-4}, 10^{-5}, 10^{-6}\}$ and $\varepsilon \in \{2^i \mid i \in \llbracket -10, -2 \rrbracket\}$	173
III.19	Diffusion approximation error at the discrete time level (III.7.2) for the non AP scheme (III.3.14), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$, $M = 10^6$ Monte Carlo samples, $\Delta t \in \{10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$ and $\varepsilon \in \{2^i \mid i \in \llbracket -10, -2 \rrbracket\}$	174
IV.1	AP Scheme 1 and its limiting Scheme 2 in the averaging regime when $\varepsilon \rightarrow 0$ for fixed δ , with initial condition $\tilde{m}_0 = 10$	189
IV.2	AP Scheme 1 and its limiting Scheme 3 in the averaging regime when $\delta \rightarrow 0$ for fixed ε , with initial condition $\tilde{m}_0 = 10$	190
IV.3	AP Scheme 1 and its limiting Scheme 4 in the averaging regime when $\varepsilon = \delta \rightarrow 0$, with initial condition $\tilde{m}_0 = 10$	191
IV.4	Weak error when $\rho_N^{\varepsilon, \delta}$ is computed with AP Scheme 1 in the averaging regime, with initial condition $\tilde{m}_0 = 1$ and mesh size $\frac{1}{64}$. The curves have slope 1.	192
IV.5	AP Scheme 5 and its limiting Scheme 7 in the diffusion approximation regime when $\varepsilon \rightarrow 0$ for fixed δ , with initial condition $\tilde{m}_0 = 10$	194
IV.6	Evolution of $\ \rho_n^{\varepsilon, \delta}\ _{L^2}$ and $\ \rho_n^\varepsilon\ _{L^2}$ for AP and non AP schemes in the diffusion approximation regime as a function of t_n when $\delta \rightarrow 0$ for fixed ε , with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{32}$	196
IV.7	Final value $\rho_N^{\varepsilon, \delta}(x)$ and $\rho_N^\varepsilon(x)$ for AP and non AP schemes in the diffusion approximation regime as a function of x when $\delta \rightarrow 0$ for fixed ε , with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{64}$	196
IV.8	Evolution of $\ \rho_n^{\varepsilon, \delta}\ _{L^2}$ and $\ \rho_n\ _{L^2}$ for AP and non AP schemes in the diffusion approximation regime as a function of t_n when $\varepsilon = \delta \rightarrow 0$, with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{32}$	197
IV.9	Final value $\rho_N^{\varepsilon, \delta}(x)$ and $\rho_N(x)$ for AP and non AP schemes in the diffusion approximation regime as a function of x when $\varepsilon = \delta \rightarrow 0$, with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{64}$	198
IV.10	Weak error when $\rho_N^{\varepsilon, \delta}$ is computed with AP Scheme 5 or non AP Scheme 6 in the diffusion approximation regime, with initial condition $\tilde{m}_0 = 1$ and mesh size $\frac{1}{64}$	199

Introduction (français)

Les modèles cinétiques avec des paramètres microscopiques apparaissent dans de nombreuses situations, par exemple dans l'étude des semi-conducteurs [GP92] et des modèles à vitesse discrète [LT97], ou encore dans la limite de systèmes de particules, que ce soit d'une seule particule [GR09] ou de plusieurs [PV03]. Ces modèles peuvent être déterministes ou probabilistes. Il est important de comprendre leurs équations limites, qui sont en général plus faciles à simuler numériquement. Les questions numériques soulevées par de tels modèles multi-échelles sont déjà étudiées en profondeur dans la littérature, mais la plupart de ces travaux se concentrent sur la simulation soit du modèle seul, soit de la limite seule. Il est naturel de chercher à écrire des schémas numériques qui soient efficaces simultanément pour le modèle et sa limite, ce qui conduit à la notion de schémas AP (*Asymptotic Preserving*, préservant l'asymptotique).

1 Description du modèle

On considère la classe d'Équations aux Dérivées Partielles Stochastiques (EDPS) linéaires cinétiques suivante :

$$\partial_t f^{\varepsilon,\delta} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^{\varepsilon,\delta} = \frac{1}{\varepsilon^2} L f^{\varepsilon,\delta} + \frac{1}{\delta^\kappa} f^{\varepsilon,\delta} m^\delta, \quad f^{\varepsilon,\delta} = f_0^{\varepsilon,\delta}, \quad (1)$$

ainsi que la classe d'Équations Différentielles Stochastiques (EDS) suivante :

$$dX_t^\delta = b(X_t^\delta, m_t^\delta) dt + \sigma(X_t^\delta, m_t^\delta) dB_t. \quad (2)$$

Quand on prend $\varepsilon = \delta$, $b = 0$ et $\sigma(x, m) = x$ dans le second modèle, on obtient une version en dimension finie du premier.

Ce premier modèle (1) est une EDPS cinétique avec un terme source. Son inconnue $f^{\varepsilon,\delta}$ est définie sur $[0, T] \times \mathbb{T}^d \times V$ pour un temps final $T \in (0, \infty)$. Il modélise l'évolution de l'état $f^{\varepsilon,\delta}(t, x, v)$ d'un système de particules ayant pour position x et pour vitesse $a(v)$ au temps t , où $v \in V$ et où $(V, d\mu)$ est un espace mesuré. On suppose que $a \in L^\infty(V, d\mu)$. Cette condition est satisfaite par exemple dans le cas discret ($V = \{\pm 1\}^d$ et $a(v) = v$) ainsi que dans le cas de particules relativistes ($V = \mathbb{R}^d$ et $a(v) = v/\sqrt{1 + \|v\|^2}$ dans de bonnes unités), mais pas par le cas standard $V = \mathbb{R}^d$ et $a(v) = v$. Cette hypothèse est uniquement technique, et les résultats devraient rester valable sans elle, au prix de complications techniques non triviales non étudiées dans ce travail.

L'opérateur de Bhatnagar-Gross-Krook (BGK) L est défini sur $L^1(V, d\mu)$ par

$$Lf = \rho \mathcal{M} - f,$$

où $\mathcal{M} \in L^1(V, d\mu)$ est une fonction positive d'intégrale 1 et où $\rho \doteq \int_V f d\mu$. Le terme $Lf^{\varepsilon, \delta}$ vient d'un processus de Poisson qui redistribue les vitesses selon la distribution $\mathcal{M}d\mu$ avec des intervalles de temps suivant des lois exponentielles. Nos hypothèses sont vérifiées par le cas classique de la distribution normale où $\mathcal{M}(v)d\mu(v) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{\|v\|^2}{2})dv$.

Le système dépend de deux paramètres ε et δ tendant vers 0 : ε est le chemin moyen entre deux collisions de particules dans le domaine \mathbb{T}^d (de taille 1), tandis que δ un paramètre séparant les échelles de temps. On se concentre principalement sur le régime $\varepsilon = \delta \rightarrow 0$, mais on considère aussi les régimes $\delta > 0$, $\varepsilon \rightarrow 0$ et $\varepsilon > 0$, $\delta \rightarrow 0$ dans les simulations numériques du chapitre IV.

Le terme source $m^\delta(t, x)$ introduit de l'aléatoire dans le système. C'est un processus stochastique ergodique défini sur $\mathbb{R}^+ \times \mathbb{T}^d$. On note que ce processus ne dépend pas de la variable de vitesse. L'évolution de m^δ en fonction du temps est à l'échelle t/δ^2 et le comportement asymptotique du modèle dépend de la valeur de κ . Dans notre travail, on considère les cas $\kappa = 1$ et $\kappa = 0$. Quand $\delta \rightarrow 0$, par ergodicité, m^δ converge en distribution vers son unique loi invariante. Dans le cas $\kappa = 1$, il ne peut donc pas y avoir de limite si la moyenne de cette loi invariante est non nulle. Ainsi, quand $\kappa = 1$, on suppose que la loi invariante de m^δ est centrée.

Comme m^δ évolue à la vitesse t/δ^2 avec $\delta \rightarrow 0$, on l'appelle la composante rapide, tandis que $\rho^{\varepsilon, \delta} = \int_V f^{\varepsilon, \delta} d\mu$ est la composante lente. La composante lente est $\rho^{\varepsilon, \delta}$ au lieu de $\int_V f^{\varepsilon, \delta} d\mu$ car le terme le plus raide de (1) est $\frac{1}{\varepsilon^2} Lf^{\varepsilon, \delta}$, on veut donc avoir asymptotiquement $Lf = 0$, c'est-à-dire $f = \rho\mathcal{M}$. Les modèles avec de telles séparations des échelles de temps apparaissent dans de nombreuses situations, dans le cas déterministe sous forme d'Équations Différentielles Ordinaires (EDO) et d'Équations aux Dérivées Partielles, ainsi que dans le cas probabiliste sous forme d'EDS et d'EDPS. C'est le cas par exemple en physique, dans l'étude de systèmes de particules [PV03, GR09] ou en astrophysique dans l'étude de grain de poussière dans une structure macroscopique [LBL20]. On peut aussi mentionner l'étude des semi-conducteurs [GP92]. Parmi ces exemples, l'étude des semi-conducteurs et des systèmes de particules peuvent tous les deux être modélisés par des modèles cinétiques.

Dans cette thèse, on étudie principalement le modèle (1) d'un point de vue théorique et numérique. L'étude théorique (chapitres I et II) consiste à prouver la convergence de la composante lente vers la solution d'une équation limite. L'étude numérique (chapitres III et IV) consiste à créer des schémas numériques qui soient à la fois consistant pour $\varepsilon > 0$ et $\delta > 0$ fixés, et dans les régimes asymptotiques $\varepsilon \rightarrow 0$ avec $\delta > 0$ fixé, $\delta \rightarrow 0$ avec $\varepsilon > 0$ fixé et $\varepsilon = \delta \rightarrow 0$. Le schéma doit admettre une limite pour tout pas de temps fixé dans chacun de ces régimes, et le schéma limite doit être consistant avec l'équation limite. Un tel schéma est dit AP (*Asymptotic Preserving*, préservant l'asymptotique).

Le manuscrit suit l'organisation suivante.

Dans le chapitre I, on considère le cas $\varepsilon = \delta \rightarrow 0$ et $\kappa = 1$. On prouve un résultat d'approximation-diffusion, au sens des EDP et au sens probabiliste : quand $\varepsilon \rightarrow 0$, la composante lente ρ^ε converge en loi vers la solution ρ d'une équation parabolique de diffusion linéaire avec un terme source diffusif (bruit blanc en temps). Ce travail est une amélioration de [DV12], au sens où on prouve le même résultat en affaiblissant une hypothèse de borne sur le processus m^ε . Cette généralisation est non triviale et requiert l'utilisation d'outils complexes comme des temps d'arrêts et des problèmes de martingale arrêtés. Ce chapitre est basé sur la prépublication [RR20].

Dans le chapitre II, on considère le cas $\varepsilon = \delta \rightarrow 0$ et $\kappa = 0$. On prouve un résultat d'approximation-diffusion au sens EDP et de moyennisation au sens probabiliste : quand

$\varepsilon \rightarrow 0$, la composante lente ρ^ε converge en loi vers la solution ρ d'une équation parabolique de diffusion avec un terme source étant la moyenne de la distribution invariante du terme source ergodique. Le chapitre II est techniquement plus simple que le chapitre I mais ce manuscrit est écrit dans l'ordre chronologique du travail de la thèse. Une prépublication associée est en cours de préparation.

Dans le chapitre III, on construit des schémas numériques pour (2), dans les régimes d'approximation-diffusion et de moyennisation. On définit une notion de schéma AP en loi adaptée aux problèmes stochastiques multi-échelles considérés et prouve que les schémas sont AP. On donne aussi des exemples de schémas non AP et on illustre numériquement l'efficacité ou non de chaque schéma. On observe clairement la supériorité des schémas AP, qualitativement et quantitativement. De plus, on prouve une propriété de précision uniforme pour le schéma AP dans le régime de moyennisation, c'est-à-dire une estimée d'erreur uniforme en $\delta > 0$ quand le pas de temps $\Delta t \rightarrow 0$. Ce chapitre est basé sur la prépublication [BRR20], à laquelle on ajoute des simulations numériques quantitatives.

Dans le chapitre IV, on combine les schémas du chapitre III avec des schémas AP pour des EDP en utilisant une procédure de *splitting*. On illustre l'efficacité des schémas proposés dans les régimes de moyennisation et d'approximation-diffusion. On considère aussi tous les régimes $\varepsilon = \delta$ et $\varepsilon \neq \delta$, de sorte qu'on puisse fixer un des paramètres et faire tendre l'autre vers 0.

2 Convergence de (1) dans le régime d'approximation-diffusion

Le chapitre I est basé sur la prépublication [RR20] et est dédié à prouver la convergence suivante pour $\rho^\varepsilon = \int_V f^\varepsilon d\mu$, où f^ε est la solution de (1) avec $\varepsilon = \delta$ et $\kappa = 1$.

Théorème. *Soit W un processus de Wiener cylindrique sur $L^2(\mathbb{T}^d)$ et ρ la solution faible de l'équation parabolique de diffusion*

$$d\rho = \operatorname{div}(K\nabla\rho)dt + \rho Q^{1/2} \circ dW(t), \quad (3)$$

avec la condition initiale $\rho(0) = \lim_{\varepsilon \rightarrow 0} \rho_0^\varepsilon$. La matrice K et l'opérateur Q sont respectivement définis par (1.2.2) et (1.2.5)

Sous de bonnes hypothèses, pour tout $\sigma > 0$ et $T > 0$, la densité ρ^ε converge en loi vers ρ dans $C^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ et dans $L^2([0, T], L^2(\mathbb{T}^d))$.

C'est un résultat d'approximation-diffusion, aux sens EDP et probabilistes. Ce théorème est une généralisation de [DV12]. D'une part, on affaiblit la condition de borne presque sûre sur le processus m^ε en une condition sur ses moments, condition vérifiée par exemple par un processus d'Ornstein-Uhlenbeck. D'autre part, on ajoute une convergence plus forte, dans $L^2([0, T], L_x^2)$, sous une hypothèse supplémentaire qui permet d'utiliser un lemme de moyenne [BD99, Theorem 2.3]. Il faut noter que cette convergence n'a lieu qu'en loi en général.

2.1 Contexte

Les résultats d'approximation-diffusion pour des équations cinétiques déterministes sont communs dans la littérature, en commençant par [LK74, BLP79]. Dans ces deux articles,

le paramètre $\varepsilon \rightarrow 0$ est le ratio entre le chemin moyen entre deux collisions de particules par la taille du domaine où elles évoluent. La méthode est de construire un développement asymptotique en ε de la solution de l'équation. Dans le cas de (1) avec $m^\varepsilon = 0$, [DGP00] montre la convergence (déterministe) en utilisant un développement similaire $f^\varepsilon = \rho^\varepsilon \mathcal{M} + \varepsilon g^\varepsilon$. Le terme d'ordre 0 ne dépend de f^ε qu'à travers ρ^ε , ce qui justifie qu'on ne considère que la convergence de ρ^ε .

Obtenir une limite diffusive au sens stochastique est aussi standard : quand un terme aléatoire est à la bonne échelle, cela induit un bruit blanc en temps dans l'équation limite, et la convergence vers cette limite n'a plus lieu qu'en loi en général. L'étude de telles convergences a commencé avec Khasminskii [Kha66a, Kha66b]. Dans ces articles, les résultats sont prouvés en établissant la convergence des distributions de dimension finie.

Dans [PSV77], les auteurs introduisent une autre approche basée sur une formulation martingale et des fonctions tests perturbées. Cette approche est fortement utilisée dans les chapitres I et II, nous allons donc l'expliquer ici. Soit X un processus stochastique et \mathcal{L} un opérateur. Soit

$$M_\varphi(t) = \varphi(X(t)) - \varphi(X(0)) - \int_0^t \mathcal{L}\varphi(X(s))ds$$

où φ est une fonction test dans une classe Θ . Le processus X résout le problème de martingale associé à \mathcal{L} si, pour tout $\varphi \in \Theta$, M_φ est une martingale. Alors, sous de bonnes hypothèses sur la classe Θ , X est un processus de Markov de générateur \mathcal{L} si et seulement si il résout le problème de martingale associé à \mathcal{L} . Ainsi, si on considère la solution X^ε d'une EDPS de générateur \mathcal{L}^ε , on obtient une famille de martingale au niveau $\varepsilon > 0$. Supposons qu'on ait montré que la famille $(X^\varepsilon)_{\varepsilon>0}$ est tendue et considérons une valeur d'adhérence X pour la convergence en loi. Si on peut passer à la limite quand $\varepsilon \rightarrow 0$ dans le problème de martingale, on obtient un problème de martingale limite qui permet d'identifier X sous réserve d'unicité de la solution de ce problème. On obtient ainsi la convergence de X^ε vers X en loi. Cependant, pour des problèmes multi-échelles singuliers, le terme $\mathcal{L}^\varepsilon \varphi(X^\varepsilon(t))$ ne converge pas, on ne peut donc pas prendre la limite directement. Pour remédier à cela, l'idée est de perturber la fonction test φ en une fonction test $\varphi^\varepsilon = \varphi + \varepsilon \varphi_1 + \varepsilon^2 \varphi_2 + \dots$, telle que $\mathcal{L}^\varepsilon \varphi^\varepsilon \rightarrow \mathcal{L}\varphi$ pour un opérateur \mathcal{L} . En supposant que la fonction test perturbée φ^ε est dans la classe Θ , on peut alors considérer le problème de martingale avec φ^ε et passer à la limite pour obtenir que X résout le problème de martingale associé à \mathcal{L} .

Il est à noter que l'usage des fonctions tests perturbées dans le contexte d'EDP (déterministes ou stochastiques) avec des limites diffuses n'est pas spécifique aux équations cinétiques : [Eva89] les utilise pour des solutions de viscosité, [PP03] pour une équation parabolique et [Mar06, dBD10, DT11, dBG12] pour des équations de Schrödinger non linéaires.

2.2 Stratégie de la preuve dans [DV12]

Le schéma de la preuve dans [DV12], basée sur une approche martingale et sur des fonctions tests perturbées, est la suivante :

- On montre la tension de la famille de solution $(\rho^\varepsilon)_{\varepsilon>0}$. Ceci nécessite des estimées sur les moments de ρ^ε , uniformément en ε .
- On construit une perturbation φ^ε d'une fonction test φ . Dans (1), le générateur peut être décomposé en $\mathcal{L}^\varepsilon = \varepsilon^{-1} \mathcal{L}_1 + \varepsilon^{-2} \mathcal{L}_2$. En utilisant le développement asymptotique

$\varphi^\varepsilon = \varphi + \sum \varepsilon^i \varphi_i$, la convergence $\mathcal{L}^\varepsilon \varphi^\varepsilon \rightarrow \mathcal{L} \varphi$ n'est vérifiée que si chaque correcteur φ_i résout l'équation de Poisson associée au générateur \mathcal{L}_2 suivante

$$\mathcal{L}_1 \varphi_{i-1} + \mathcal{L}_2 \varphi_i = \vartheta_{i-1},$$

où le membre de droite ϑ_{i-1} ne dépend que de φ_j pour $j < i$. Sous de bonnes hypothèses, la solution de cette équation est bien défini et on peut construire successivement les correcteurs φ_j .

- Par tension, il existe une valeur d'adhérence pour la convergence en loi, ρ et une suite $\varepsilon_i \xrightarrow{i \rightarrow \infty} 0$ telle que ρ^{ε_i} converge vers X en loi. Alors, on peut prendre la limite dans la propriété de martingale

$$\mathbb{E} \left[\varphi^{\varepsilon_i}(\rho^{\varepsilon_i}(t)) - \varphi^{\varepsilon_i}(\rho^{\varepsilon_i}(0)) - \int_0^t \mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(\rho^{\varepsilon_i}(s)) ds \mid \mathcal{F}_s^{\rho^{\varepsilon_i}}, s \leq t \right] = 0,$$

pour obtenir que ρ résout le problème de martingale associé à \mathcal{L} . En montrant l'unicité d'une telle solution, cela conclut la preuve de la convergence en loi.

2.3 Stratégie de la preuve dans le chapitre I

On décrit ici les nouveautés de notre travail. Dans [DV12], les auteurs supposent que m^ε est majoré par une borne presque sûre. C'est une hypothèse forte, qui n'est pas vérifiée par exemple par un processus Gaussien. Dans ce chapitre, l'objectif est d'affaiblir cette hypothèse. Une conséquence directe est qu'on perd les estimées de moments uniformes en ε . Pour palier ce problème, on introduit des temps d'arrêts dépendant d'un nombre réel arbitrairement grand $\Lambda \in (0, \infty)$, tel que, pour Λ fixé, on a une estimation uniforme en ε du processus stoppé $\rho^{\varepsilon, \tau_\Lambda^\varepsilon}$. Plus précisément, le temps d'arrêt est défini comme

$$\tau_\Lambda^\varepsilon \doteq \tau^\varepsilon \wedge \tau_\Lambda(\zeta^\varepsilon),$$

où

$$\tau^\varepsilon \doteq \inf \{ t \in \mathbb{R}^+ \mid \|m^\varepsilon(t)\|_E > \varepsilon^{-\alpha} \},$$

pour un certain paramètre α , et où

$$\tau_\Lambda(z) \doteq \inf \{ t \in \mathbb{R}^+ \mid \|z(t)\|_{\mathcal{C}^1(\mathbb{T}^d)} \geq \Lambda \},$$

et

$$\zeta^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t m^\varepsilon(s) ds = \varepsilon \int_0^{t/\varepsilon^2} m(s) ds.$$

On peut alors prouver la tension de la famille $(\rho^{\varepsilon, \tau_\Lambda^\varepsilon})_{\varepsilon > 0}$ et suivre la preuve standard avec des martingales arrêtées au lieu de martingales. Une difficulté majeure vient alors du comportement asymptotique de ces temps d'arrêt. D'une part, quand α est bien choisi, $\tau^\varepsilon \rightarrow \infty$ en probabilité. D'autre part, on montre un Théorème Central Limite pour m^ε , prouvant que ζ^ε converge en loi vers un processus de Wiener ζ et que $\tau_\Lambda(\zeta^\varepsilon) \rightarrow \tau_\Lambda(\zeta)$ en loi pour tout Λ hors d'un ensemble dénombrable. Ainsi, τ_Λ^ε converge en loi vers $\tau_\Lambda(\zeta)$ quand $\varepsilon \rightarrow 0$. Pour que ce temps d'arrêt soit adapté à la filtration associée au processus limite, il faut considérer la convergence du couple $(\rho^\varepsilon, \zeta^\varepsilon)$ au lieu de la convergence seule de ρ^ε .

On montre que le processus couplé stoppé $((\rho^\varepsilon, \tau_\Lambda^\varepsilon, \zeta^\varepsilon, \tau_\Lambda^\varepsilon))_{\varepsilon>0}$ est tendu et qu'une valeur d'adhérence pour la convergence en loi satisfait un problème de martingale arrêtée au temps d'arrêt $\tau_\Lambda(\zeta)$, ce qui donne une famille de problèmes de martingales indexée par Λ . Comme chaque problème de martingale ne caractérise sa solution $X_\Lambda = (\rho_\Lambda, \zeta_\Lambda)$ que jusqu'au temps $\tau_\Lambda(\zeta)$, cette solution peut ne pas être unique. Cependant, on prouve que la famille de problèmes de martingales arrêtées est consistante, au sens où X_Λ est égale en loi à $X^{\tau_\Lambda(\zeta)}$ pour un certain processus X . Comme $\tau_\Lambda(\zeta) \rightarrow \infty$ quand $\Lambda \rightarrow \infty$, on utilise ces résultats pour déduire la convergence du processus non stoppé ρ^ε .

3 Convergence de (1) dans le régime de moyennisation

Le chapitre II est dédié à prouver la convergence suivante pour $\rho^\varepsilon = \int_V f^\varepsilon d\mu$ où f^ε est solution de (1) avec $\varepsilon = \delta$ et $\kappa = 0$.

Théorèmes. *Soit \bar{m} la moyenne de l'unique loi invariante de m^ε et soit ρ la solution faible de l'équation linéaire de diffusion*

$$\partial_t \rho = \operatorname{div}(K \nabla \rho) + \bar{m} \rho, \quad (4)$$

avec $\rho(0) = \lim_{\varepsilon \rightarrow 0} \rho_0^\varepsilon$.

Sous de bonnes hypothèses, pour tout $\sigma > 0$ et $T > 0$, la fonction ρ^ε converge en loi vers ρ dans $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ et dans $L^2([0, T], L^2(\mathbb{T}^d))$ et f^ε converge en loi vers $\rho \mathcal{M}$ dans $L^2([0, T], L^2(\mathbb{T}^d \times V, \frac{dx d\mu(v)}{\mathcal{M}(v)}))$. De plus, si $\rho(0)$ est déterministe, alors toutes ces convergences ont lieu en probabilité.

C'est un résultat d'approximation-diffusion au sens EDP comme dans le chapitre I, mais un résultat de moyennisation au sens probabiliste. Le principe de moyennisation pour des systèmes stochastiques a été introduit dans l'article classique [Kha68]. De nombreux résultats sont associés à ce principe, utilisant différentes preuves et différents critères de convergence, voir par exemple [PS08, Cer09, CF09, Bré12, Bré20]. Certaines de ces articles étudient aussi la vitesse de convergence de la principe de moyennisation dans des EDPS (voir [Bré12, Bré20]). Dans le chapitre II, on ne montre que la convergence. L'analyse de la vitesse de convergence est laissée pour des travaux futurs.

La stratégie de la preuve est similaire au chapitre I : on suit une approche basée sur les problèmes de martingales et sur les fonctions tests perturbées, et on obtient des estimées sur les moments en arrêtant les processus à des temps d'arrêts. La principale différence avec le chapitre précédent est qu'ici, le temps d'arrêt $\tau^\varepsilon \rightarrow \infty$ en probabilité. Ainsi, le problème de martingale limite n'est pas un problème de martingale arrêté, et on retrouve le cadre classique où la solution est unique et où la convergence suit d'arguments simples. Plus précisément, le processus auxiliaire est cette fois défini par

$$\zeta^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t (m^\varepsilon(s) - \bar{m}) ds,$$

et les temps d'arrêt par

$$\begin{aligned} \tau^\varepsilon &= \tau_m^\varepsilon \wedge \tau_\zeta^\varepsilon, \\ \tau_m^\varepsilon &= \inf \left\{ t \in \mathbb{R}^+ \mid \|m^\varepsilon(t)\|_E \geq \varepsilon^{-\alpha} \right\}, \\ \tau_\zeta^\varepsilon &= \inf \left\{ t \in \mathbb{R}^+ \mid \|\zeta^\varepsilon(t)\|_{C_x^1} \geq \varepsilon^{-1} \right\}. \end{aligned}$$

Comme dans le chapitre I, pour un bon choix de α , on a $\tau_m^\varepsilon \rightarrow \infty$ en probabilité. Cependant, dans le temps d'atteinte τ_ζ^ε , le seuil fixe Λ est remplacé par ε^{-1} , ce qui permet d'obtenir $\tau_\zeta^\varepsilon \rightarrow \infty$ directement.

4 Schémas préservant l'asymptotique pour des EDS

Le chapitre III est basé sur la prépublication [BRR20] et est dédié à construire des schémas AP pour la classe d'EDS (2) dans les régimes d'approximation-diffusion et de moyennisation. On travaille d'abord sur des EDS, avant de revenir au cas EDPS dans le chapitre IV, car nous avons remarqué que dans ce cas plus simple, notre travail apportait déjà des contributions non triviales. On renvoie à [PS08] pour les résultats de convergence sur des EDS multi-échelles : [PS08, Chapter 11] pour l'approximation-diffusion et [PS08, Chapter 10] pour la moyennisation.

4.1 Contexte

Commençons par motiver ce travail. Dans le cas déterministe, le champ des schémas numériques pour EDO est très bien étudié [HW96]. La présence de termes raides dans le système, typiquement à cause de la composante rapide, peut induire des conditions de stabilités fortes de la forme $\Delta t = o(\varepsilon^\alpha)$, puisqu'on veut que le pas de temps mis à l'échelle $\Delta t/\varepsilon^\alpha$ converge vers 0. Ainsi, le coût en calculs explose quand $\varepsilon \rightarrow 0$. Quand le système est probabiliste, la simulation numérique peut être encore plus difficile.

[FG18] donne des exemples où l'EDS satisfaite par certaines quantités statistiques d'un schéma (par exemple la densité de probabilité) est différent dans l'équation limite et dans le système en temps continu. Un de ces exemples est l'interprétation d'Itô du bruit quand l'EDS limite a une interprétation de Stratonovich de ce bruit. On détaille cet exemple plus bas.

[LAE08] illustre comme les schémas pour des EDO raides peuvent produire de mauvaises solutions quand la composante rapide n'est pas résolue. Les auteurs montrent même qu'une discrétisation peut converger vers un schéma X^c quand $\Delta t \rightarrow 0$ et $\varepsilon \rightarrow 0$ avec un ratio $c = \Delta t/\varepsilon^\alpha$ fixé, où le schéma X^c dépend fortement de c .

Un moyen bien étudié d'éviter ces problèmes est d'approcher le modèle limite pour la composante lente d'intérêt, tout en ne conservant que l'information pertinente de la composante rapide. On détaille ici plusieurs des méthodes qui suivent cette approche.

La HMM (*Heterogeneous Multiscale Method*, méthode hétérogène multi-échelle, voir [AEEVE12, Bré13, ELVE05]) est basée sur un schéma macroscopique qui effectue les calculs jusqu'à ce qu'un schéma microscopique est nécessaire. Plus précisément, on commence par effectuer les calculs au niveau macroscopique. Quand on a besoin d'informations qui ne peuvent être obtenues qu'au niveau microscopique, on effectue des appels à un schémas microscopiques, avec des contraintes pour être consistant avec les calculs déjà effectués au niveau macroscopique. Cela nous permet d'estimer les informations manquantes et de les réinjecter dans le schéma macroscopique. De manière analogue, la méthode d'intégration projective [GKK06] est basée sur un schéma macroscopique et sur des appels en temps courts au schéma microscopique, afin d'éviter d'augmenter le coût en calcul, tout en prenant en compte la dynamique microscopique. Ces deux méthodes sont des méthodes de basées sur de une moyennisation de la dynamique microscopique.

La méthode *equation-free coarse-grained* [KGH⁺03] approche à la place l'évolution de la composante macroscopique via des appels correctement initialisés au schéma microscopique pour des temps courts et de petits domaines spatiaux, par exemple en intégrant la dynamique sur des temps long ou de grands domaines spatiaux. En particulier, cette approche est sans équation (*equation-free*) au sens où on n'a pas besoin de dériver l'équation d'évolution macroscopique à partir de la dynamique microscopique.

Plus récemment, [APV17] propose une méthode spectrale. les coefficients de l'équation homogénéisée sont calculés en utilisant les solutions de bonnes équations de Poisson. Ces équations de Poisson sont résolues en utilisant une approche spectrale basée sur un développement en fonctions de Hermite.

Dans [VZS20], les auteurs décrivent une méthode d'accélération micro-macro. L'algorithme alterne entre un schéma microscopique pour des temps courts et une extrapolation de l'état macroscopique sur des temps longs. Ensuite, ils calculent la loi qui serait consistante avec cette extrapolation, tout en modifiant le moins possible la distribution microscopique précédemment calculée. Cela leur permet de reconstruire l'état microscopique.

Enfin, [LLMS20] présente une version micro-macro de l'algorithme pararéel, qui parallélise les calculs en fonction de la variable de temps, ce qui permet une réduction drastique du temps de calcul.

4.2 Principales contributions du chapitre III

Pour la classe d'EDS sur lesquelles on se concentre, les schémas standard se comportent mal dans le régime asymptotique. Commençons par considérer le régime d'approximation-diffusion. Le système suivant est un cas particulier de (2)

$$\begin{cases} dX_t^\varepsilon = \frac{\sigma(X_t^\varepsilon)m_t^\varepsilon}{\varepsilon}dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon^2}dt + \frac{1}{\varepsilon}d\beta_t. \end{cases}$$

En prenant $\sigma(x) = x$, on obtient une version en dimension finie de (1) avec $\varepsilon = \delta$, $\kappa = 1$ et où m^ε est un processus d'Ornstein-Uhlenbeck à l'échelle t/ε^2 . Prenons $d = 1$. Quand $\varepsilon \rightarrow 0$, le résultat d'approximation-diffusion (voir [PS08, Chapter 11]) dit que X^ε converge en loi vers la solution X de l'EDS

$$dX_t = \sigma(X_t) \circ d\beta_t,$$

où le bruit est interprété avec la convention de Stratonovich, la forme d'Itô s'écrivant

$$dX_t = \frac{1}{2}\sigma'(X_t)\sigma(X_t)dt + \sigma(X_t)d\beta_t.$$

Pour $\varepsilon > 0$ fixé, le schéma numérique

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \sigma(X_n^\varepsilon)\frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2}m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon}\gamma_n, \end{cases}$$

avec $(\gamma_n)_{n \in \mathbb{N}_0}$ une suite i.i.d. de variable normale centrée réduites, est consistant avec le système. Cependant, quand $\varepsilon \rightarrow 0$, on a la convergence $X_n^\varepsilon \rightarrow X_n$ où

$$X_{n+1} = X_n + \sqrt{\Delta t}\sigma(X_n)\gamma_n.$$

Ce schéma limite est consistant avec l'interprétation d'Itô du bruit, au lieu de l'interprétation Stratonovich, il échoue donc à capturer le bon comportement asymptotique du système. Dans le chapitre III, on construit un schéma AP, c'est-à-dire un schéma tel que le diagramme suivant est commutatif :

$$\begin{array}{ccc} X_N^\varepsilon & \xrightarrow[\Delta t \rightarrow 0]{} & X^\varepsilon(T) \\ \downarrow \varepsilon \rightarrow 0 & & \varepsilon \rightarrow 0 \downarrow \\ X_N & \xrightarrow[\Delta t \rightarrow 0]{} & X(T) \end{array}$$

Ce schéma AP est basé sur une procédure de prédiction-correction :

$$\left\{ \begin{array}{l} \hat{m}_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(X_n^\varepsilon) \hat{m}_{n+1}^\varepsilon}{\varepsilon^2} + \frac{\Delta t f(X_n^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon}, \\ \hat{X}_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \sigma(X_n^\varepsilon) \frac{\Delta t \hat{m}_{n+1}^\varepsilon}{\varepsilon}, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon) m_{n+1}^\varepsilon}{\varepsilon^2} + \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon}, \\ Y_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \sigma(X_n^\varepsilon) \frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}, \\ X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \frac{\sigma(X_n^\varepsilon) + \sigma(Y_{n+1}^\varepsilon)}{2} \frac{\Delta t}{\varepsilon} \frac{\hat{m}_{n+1}^\varepsilon + m_{n+1}^\varepsilon}{2}. \end{array} \right.$$

Ce schéma est AP pour une large classe d'EDS, y compris pour l'exemple précédent où il capture l'interprétation correcte du bruit, mais aussi dans les systèmes considérés dans l'article d'astrophysique [LBL20], où l'équation limite contient un terme de dérive induit par le bruit quand f est non constant.

Dans le régime de moyennisation, le même type de problème apparaît. Considérons

$$\left\{ \begin{array}{l} dX_t^\varepsilon = b(X_t^\varepsilon, m_t^\varepsilon) dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon} dt + \frac{\sqrt{2}}{\sqrt{\varepsilon}} d\beta_t. \end{array} \right.$$

En prenant $b(x, m) = xm$, on obtient une version en dimension finie de (1) avec $\varepsilon = \delta$, $\kappa = 0$ et où m^ε est un processus d'Ornstein-Uhlenbeck à l'échelle t/ε . Quand $\varepsilon \rightarrow 0$, le principe de moyennisation (voir [PS08, Chapter 10]) affirme que X^ε converge en loi vers la solution X de l'EDS

$$\dot{X} = \bar{b}(X),$$

où $\bar{b}(x) = \int b(x, m) d\nu(m)$ et $\nu = \mathcal{N}(0, 1)$. Pour $\varepsilon > 0$ fixé, le schéma

$$\left\{ \begin{array}{l} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon), \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon + \sqrt{\frac{2\Delta t}{\varepsilon}} \gamma_n, \end{array} \right.$$

est consistant avec le système, mais sa limite est donnée par

$$X_{n+1} = X_n + \Delta t b(X_n, 0),$$

qui est en général inconsistant avec l'équation limite. Encore une fois, on construit un schéma AP tel que le diagramme commutatif plus haut est de nouveau commutatif, cette fois basé sur une discrétisation exacte en loi de la composante rapide :

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon) + \sqrt{\Delta t} \sigma(X_n^\varepsilon, m_{n+1}^\varepsilon) \Gamma_n \\ m_{n+1}^\varepsilon = e^{-\Delta t/\varepsilon} m_n^\varepsilon + \sqrt{1 - e^{-2\Delta t/\varepsilon}} h(X_n^\varepsilon) \gamma_n. \end{cases} \quad (5)$$

Après avoir construit ces schémas AP, on présente trois types de simulations. Chacun des trois illustre la supériorité du schéma AP sur les schémas naïfs :

- Le premier type est qualitatif : on trace la trajectoire de la solution en fonction du temps t , pour une valeur donnée du pas de temps Δt et pour différentes valeurs de ε . Les trajectoires sont comparées à un schéma de référence.
- Le deuxième type est quantitatif : on trace une estimation de l'erreur faible totale

$$|\mathbb{E} [\varphi(X_N^\varepsilon)] - \mathbb{E} [\varphi(X^\varepsilon(T))]|,$$

pour une fonction test φ , en fonction de Δt pour différentes valeurs de ε .

- Le troisième type est aussi quantitatif : on trace une estimation de l'erreur faible d'approximation-diffusion ou de moyennisation au niveau discret

$$|\mathbb{E} [\varphi(X_N^\varepsilon)] - \mathbb{E} [\varphi(X_N)]|,$$

en fonction de ε pour différentes valeurs de Δt .

Dans les deux types de simulations quantitatives, on observe des résultats pour le schéma AP cohérents avec l'analyse d'erreur ci-dessous, et on obtient même des ordres de convergences plus grands que ceux qu'on réussit à montrer. Le schéma non AP lui devient mauvais quand ε devient petit, les constantes d'erreur en ε (respectivement Δt) étant non uniforme en Δt (respectivement ε).

4.3 Estimées d'erreurs dans le régime de moyennisation

Dans le régime de moyennisation, on prouve un résultat de précision uniforme pour le schéma AP (5). Ici, le processus m^ε est à l'échelle t/ε .

Théorème. *Sous de bonnes hypothèses, pour tout $T \in (0, \infty)$ et toute fonction test $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$ de classe \mathcal{C}^4 , il existe $C(T, \varphi) \in (0, \infty)$ tel que pour tout $\Delta t \in (0, \Delta t_0]$ et $\varepsilon \in (0, 1]$, on a*

$$|\mathbb{E} [\varphi(X_N^\varepsilon)] - \mathbb{E} [\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \min \left(\frac{\Delta t}{\varepsilon}, \Delta t + \varepsilon \right), \quad (6)$$

et le schéma est uniformément précis avec l'estimée d'erreur suivante : pour tout $\Delta t \in (0, \Delta t_0]$, on a

$$\sup_{\varepsilon \in (0, 1]} |\mathbb{E} [\varphi(X_N^\varepsilon)] - \mathbb{E} [\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \sqrt{\Delta t}. \quad (7)$$

À notre connaissance, c'est le premier résultat de précision uniforme pour un schéma AP dans des équations stochastiques.

L'idée derrière la preuve de l'uniforme précision est d'estimer le membre de gauche de (6) en passant par deux chemins sur le diagramme commutatif

$$\begin{array}{ccc} X_N^\varepsilon & \xrightarrow[\Delta t \rightarrow 0]{A} & X^\varepsilon(T) \\ B \downarrow \varepsilon \rightarrow 0 & & \varepsilon \rightarrow 0 \downarrow D \\ X_N & \xrightarrow[C]{\Delta t \rightarrow 0} & X(T) \end{array}$$

où A , B , C et D sont les erreurs dans chacune des convergences :

- L'estimée $O(\frac{\Delta t}{\varepsilon})$ vient du chemin direct $|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq A$. Comme on a utilisé un schéma classique avec un pas de temps $\Delta t/\varepsilon$, on peut montrer que $A = O(\Delta t/\varepsilon)$ en utilisant des arguments standards (équation de Kolmogorov).
- L'estimée $O(\Delta t + \varepsilon)$ vient du chemin indirect $|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq B + C + D$. Comme D est l'erreur dans le principe de moyennisation, on a $D = O(\varepsilon)$. L'erreur C est obtenue avec des équations de Poisson et de Kolmogorov, et on obtient $C = O(\Delta t)$. Montrer une estimée d'erreur satisfaisante pour B est le terme le plus difficile. On adapte la preuve de D au cas discret, et on réussit à obtenir $B = O(\max(\Delta t, \varepsilon))$.

On arrive à (6) en écrivant $|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq \min(A, B + C + D)$ et à (7) en utilisant la première ou la deuxième estimée selon que $\sqrt{\Delta t} \leq \varepsilon$ ou $\varepsilon \leq \sqrt{\Delta t}$. Cet ordre $\frac{1}{2}$ est peut-être non optimal, les simulations montrant un ordre 1.

Dans le régime d'approximation-diffusion, on n'a pas de résultat général d'erreur uniforme, si ce n'est dans des cas particuliers.

5 Schémas préservant l'asymptotique pour des EDP stochastiques cinétiques

Fort du travail du chapitre III sur le cas EDS, on peut à présent revenir au cas EDPS : le chapitre IV est dédié à construire des schémas AP pour (1) dans les régimes d'approximation-diffusion et de moyennisation. Contrairement aux chapitres I et II, on ne considère par seulement le cas $\varepsilon = \delta$, mais on autorise les deux paramètres à converger vers 0 indépendamment l'un de l'autre.

5.1 Contexte dans le cas déterministe

La notion de schéma AP dans le cas déterministe est introduite dans [Jin99], avec applications à des EDP cinétiques multi-échelles convergeant vers des EDP paraboliques de diffusion. Le système (1) qu'on considère est un cas particulier de ce cadre si $m^\delta = 0$. Depuis, de nombreux schémas AP ont été construits dans ce cadre. Les paragraphes suivants décrivent certains de ces schémas.

Pour les modèles BGK avec un échelle hyperbolique

$$\partial_t f^\varepsilon + a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon} L f^\varepsilon,$$

les articles de revue [DP14, Section 7] et [Jin12] mentionnent tous les deux un schéma de type *splitting*. Le *splitting* en temps est effectué en résolvant d’abord le terme de collision Lf^ε et puis le terme de transport $a(v) \cdot \nabla_x f^\varepsilon$. On peut alors se concentrer sur le premier et utiliser un schéma explicite *upwind* pour le terme de transport. Le schéma obtenu sera AP si le schéma utilisé pour le terme de collision l’est. [DP14, Section 7] donne une condition suffisante pour que pour qu’un schéma de Runge-Kutta pour le terme de collision soit AP, tandis que [Jin12] propose un schéma AP implicite.

Ces deux articles mentionnent aussi une famille de schémas sans *splitting* dits implicites-explicites (IMEX). Par exemple, un schéma IMEX peut être obtenu en utilisant une discrétisation implicite du terme de collision et explicite du terme de transport. Dans ces deux méthodes (*splitting* et IMEX), même si le schéma est partiellement implicite, il peut être exprimé explicitement et ne nécessite donc pas d’inverser un système non linéaire d’équations.

Dans [Pup19, Section 4], l’auteur mentionne un schéma AP basé sur une discrétisation *upwind* de la variable d’espace avec un schéma d’ordre élevé en espace.

Pour les modèles BGK avec une échelle parabolique comme dans (1), c’est-à-dire

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon,$$

des schémas AP sont proposés par [Hiv18, LM08] par exemple.

Un premier schéma AP peut être construit en utilisant la décomposition micro-macro $f^\varepsilon = \rho^\varepsilon \mathcal{M} + \varepsilon g^\varepsilon$ et puis en construisant des schémas basés sur les équations vérifiées par ρ^ε et g^ε . Un deuxième schéma AP est obtenu par une discrétisation entièrement implicite de l’équation vérifiée par la transformée de Fourier en espace \hat{f}^ε de f^ε . Ce second schéma est détaillé dans le chapitre IV. Bien que ces schémas soient AP, ils ne sont pas uniformément précis, au sens où l’erreur n’est pas uniforme en ε .

Un troisième schéma AP, basé sur une formulation intégrale de l’équation, est uniformément précis. Plus précisément, on écrit l’équation vérifiée par \hat{f}^ε , on l’écrit sous forme Duhamel puis on intègre cette expression en la variable de vitesse. En discrétisant l’intégrale en temps due à la forme Duhamel, on utilise une règle de quadrature d’ordre 2, ce qui permet à [Hiv18] de montrer que le schéma obtenu pour $\hat{\rho}^\varepsilon$ est uniformément précis. On n’utilise pas ce schéma ici pour considérer une implémentation plus simple. L’application de ce schéma uniformément précis dans le cadre stochastique sera étudiée dans des travaux futurs.

5.2 Contexte dans le cas probabiliste en approximation-diffusion

Dans un contexte probabiliste, une difficulté supplémentaire s’ajoute : en général, la convergence n’a plus lieu qu’en loi. Ainsi, la propriété AP doit être formulée pour la convergence en loi. On reprend donc la définition introduite dans le chapitre III.

Les travaux mentionnés plus hauts sur des modèles déterministes ont été étendus à des modèles probabilistes. Dans [HJ17, Jin18, JLP18], les auteurs étudient des schémas AP pour des EDP avec des coefficients aléatoires. Dans [DPS18, RLJ14], les auteurs utilisent des méthodes de Monte-Carlo pour des problèmes déterministes. Cependant, ces deux situations ne s’appliquent pas aux EDPS.

Dans [AF19], les auteurs considèrent l’EDP stochastique cinétique

$$df^\varepsilon + \frac{1}{\varepsilon} v \partial_x f^\varepsilon dt = \frac{\sigma}{\varepsilon^2} \mathcal{L} f^\varepsilon dt + f \circ Q dW_t,$$

où le terme source est un processus de Wiener, ce qui est assez proche de notre modèle (1). Plus précisément, on retrouve cette équation en prenant la limite quand $\delta \rightarrow 0$ dans (1) dans le régime d'approximation-diffusion, bien que notre opérateur L soit différent du leur. Ils construisent un schéma AP utilisant une décomposition micro-macro $f^\varepsilon = \rho^\varepsilon \mathcal{M} + \varepsilon g^\varepsilon$ et une discrétisation *upwind* des dérivées en espace. Une conséquence est que le schéma limite est explicite, et introduit donc une condition de stabilité CFL. Dans le chapitre IV, on construit un schéma AP sans condition CFL et avec un bruit non nécessairement Wiener.

5.3 Principales contributions du chapitre IV

Notre construction de schémas AP est basée sur une méthode de *splitting* en temps, via la décomposition de (1) en une partie déterministe

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon, \quad (8)$$

et une partie probabiliste

$$\partial_t f^\delta = \frac{1}{\delta^\kappa} m^\delta f^\delta. \quad (9)$$

Ainsi, le schéma AP est une combinaison des schémas mentionnés plus hauts pour (8) et d'un schéma pour (9). Il est à noter que (9) est plus simple que le cas général (2) considéré dans le chapitre III, puisque linéaire en m^δ . Ainsi, on peut soit utiliser les schémas généraux du chapitre III, soit d'autres schémas particuliers.

On illustre les qualités des schémas AP de trois manières :

- Les deux premières sont qualitatives : on trace la trajectoire de $\|\rho(t)\|_{L^2}$ en fonction de t ou la valeur au temps final $\rho^{\varepsilon,\delta}(T, x)$ en fonction de x , pour une valeur fixée du pas de temps τ et pour différentes valeurs de ε et δ .
- La troisième manière est quantitative : on trace une estimation de l'erreur faible

$$|\mathbb{E}[\varphi(\rho^{\varepsilon,\delta}(T))] - \mathbb{E}[\varphi(\rho(T))]|$$

en fonction du pas de temps τ , pour différentes valeurs de ε et δ .

6 Perspectives

6.1 Résultats à prouver

Nous avons laissé certains résultats pour des travaux futurs. Dans la partie théorique de cette thèse, l'approximation-diffusion peut être vue comme un régime de Théorème Central Limite du point de vue de la moyennisation. Ce point de vue pourrait permettre d'obtenir des informations supplémentaires sur le régime de moyennisation, comme la vitesse de convergence de $f^{\varepsilon,\delta}$ vers $\rho \mathcal{M}$.

Certaines pistes restent aussi à explorer dans les parties numériques de cette thèse. Par exemple, on ne prouve la précision uniforme que dans le régime de moyennisation des EDS. Le même type de résultat mériterait d'être obtenu dans le régime d'approximation-diffusion.

Dans le cas EDPS, on pourrait aussi étudier la stabilité numérique de nos schémas AP plus précisément, par exemple en montrant des bornes uniformes en le pas de temps sur les

moments L^2 des schémas. Ensuite, on pourrait aussi étudier la vitesse de convergence, à la fois d'un point de vue théorique et numérique. Il est à noter que les simulations numériques pour estimer la vitesse de convergence sont lourdes en calcul, il faudrait donc auparavant implémenter des versions plus efficaces de nos schémas.

Les schémas AP ont en l'état peu de chance d'être uniformément précis, puisque la partie déterministe utilise un schéma non uniformément précis. Cependant, l'avantage de la méthode de *splitting* est qu'il est aisé de remplacer cette partie par un autre schéma, par exemple par le schéma basé sur la formulation intégrale donné dans [Hiv18], qui est à la fois AP et uniformément précis. Ce nouveau schéma de *splitting* a des chances d'être uniformément précis lui aussi.

6.2 Vers d'autres modèles

Une autre perspective est d'étendre les résultats théoriques et numériques à d'autres modèles ou vers des hypothèses plus faibles.

La principale hypothèse qui mériterait d'être affaiblie est celle demandant que la vitesse $a(v)$ soit bornée. Affaiblir cette hypothèse est nécessaire pour inclure le cas classique $a(v) = v$ avec $V = \mathbb{R}^d$. Cependant, affaiblir cette hypothèse n'est pas chose aisée et pose des problèmes techniques non triviaux. Une autre hypothèse à affaiblir est $\mu(V) < \infty$. Il faut faire cependant attention aux espaces dans lesquels on travail, puisque par exemple on a montré que $f^{\varepsilon, \delta} \rightarrow \rho\mathcal{M}$ dans $L_T^2 L^2(\mathcal{M}^{-1}) \doteq L^2([0, T], L^2(\mathbb{T}^d \times V, \frac{dx d\mu(v)}{\mathcal{M}(v)}))$ quand $\mu(V) < \infty$ dans le régime de moyennisation, mais si $\mu(V) = \infty$, alors $\rho\mathcal{M} \notin L_T^2 L^2(\mathcal{M}^{-1})$.

Une autre perspective importante est d'adapter la généralité du cas EDS (chapitre III) au cas EDPS. Par exemple, on pourrait vouloir obtenir les résultats théoriques des chapitres I et II en remplaçant le terme source linéaire $m^\delta f^{\varepsilon, \delta}$ par un terme plus général $b(m^\delta) f^{\varepsilon, \delta}$. On pourrait aussi vouloir étudier des modèles couplés, où la composante rapide dépend de la composante lente. Dans le cas EDS, on voit dans le chapitre III que quand ce couplage est non trivial, il peut induire un terme de dérive induit par le bruit dans l'équation limite. Une prépublication récente [CX20] étudie le cas parabolique et confirme l'ajout d'un terme de dérive induit par le bruit dans l'équation limite.

Nos schémas numériques dans les chapitres III et IV sont écrit spécifiquement pour une composante rapide sous la forme d'un processus d'Ornstein-Uhlenbeck (quand la composante lente est gelée). Une perspective naturelle mais non triviale est de remplacer cette composante rapide par la solution ergodique d'une EDS générale avec des coefficients non linéaires, comme dans les chapitres I et II.

Enfin, on pourrait étudier des modèles proches, par exemple en remplaçant le terme source $\delta^{-\kappa} m^\delta f^{\varepsilon, \delta}$ par un terme de force $\delta^{-\kappa'} m^\delta \cdot \nabla_v f^{\varepsilon, \delta}$, où on s'attend à $\kappa' = 2$ pour être dans le régime d'approximation-diffusion et à $\kappa' = 1$ pour être dans le régime de moyennisation. Ce modèle a déjà été étudié d'un point de vue théorique [DV20], mais il n'existe à notre connaissance pas de schémas AP pour ce modèle. L'approche par une méthode de *splitting* devrait s'avérer efficace.

Introduction (English version)

Kinetic models with small parameters appear in various situations, for example when studying semi-conductors [GP92] and discrete velocity models [LT97] or as a limit of a particle system, either with a single particle [GR09] or multiple ones [PV03]. These models can either be deterministic or stochastic. It is important to understand the limiting equations, which are in general much easier to simulate numerically. This numerical aspect is a thoroughly studied field in the literature, but most works focus on simulating either only the limiting equation, or only the initial model. It is natural to try to design numerical schemes that perform well both in the model and in its limit, leading to the notion of Asymptotic Preserving schemes.

1 Description of the model

We consider a class of linear kinetic Stochastic Partial Differential Equation (SPDE) with the following form:

$$\partial_t f^{\varepsilon, \delta} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^{\varepsilon, \delta} = \frac{1}{\varepsilon^2} L f^{\varepsilon, \delta} + \frac{1}{\delta \kappa} f^{\varepsilon, \delta} m^\delta, \quad f^{\varepsilon, \delta} = f_0^{\varepsilon, \delta}, \quad (1)$$

as well as a class of Stochastic Differential Equation (SDE)

$$dX_t^\delta = b(X_t^\delta, m_t^\delta) dt + \sigma(X_t^\delta, m_t^\delta) dB_t. \quad (2)$$

The interest of the second model is that when taking $\varepsilon = \delta$, $b = 0$ and $\sigma(x, m) = x$, one gets a finite dimensional version of (1).

The first model is a kinetic SPDE with a source term and its unknown $f^{\varepsilon, \delta}$ is defined on $[0, T] \times \mathbb{T}^d \times V$ for some final time $T \in (0, \infty)$. This models the evolution of the state $f^{\varepsilon, \delta}(t, x, v)$ of a system of particles with position x and velocity $a(v)$ at time t , where $v \in V$ for some measured space $(V, d\mu)$. We assume that $a \in L^\infty(V, d\mu)$. This assumption is satisfied in the discrete case ($V = \{\pm 1\}^d$ and $a(v) = v$) as well as in the case of relativistic particles ($V = \mathbb{R}^d$ and $a(v) = v/\sqrt{1 + \|v\|^2}$ in convenient units), but excludes the natural case $V = \mathbb{R}^d$ and $a(v) = v$. This assumption is done for technical reasons and we think it could be relaxed at the cost of non trivial technical complications that are not studied in this work.

The Bhatnagar-Gross-Krook (BGK) operator L is defined on $L^1(V, d\mu)$ by

$$Lf = \rho \mathcal{M} - f,$$

where $\mathcal{M} \in L^1(V, d\mu)$ is a non-negative function of mass 1 and $\rho \doteq \int_V f d\mu$. This term comes from a Poisson process redistributing the velocity following the distribution $\mathcal{M} d\mu$ in

exponentially distributed random interval of times. Our framework includes the classical Gaussian framework where $\mathcal{M}(v)d\mu(v) = \frac{1}{\sqrt{2\pi}^d} \exp(-\frac{\|v\|^2}{2})dv$.

The system depends on two small parameters ε and δ : ε is the mean path between collisions of the particles in the domain \mathbb{T}^d of size 1, while δ is a time scale separation parameter. We mainly focus on the regime $\varepsilon = \delta \rightarrow 0$, but also consider the regimes $\delta > 0$, $\varepsilon \rightarrow 0$ and $\varepsilon > 0$, $\delta \rightarrow 0$ in the numerical experiments in Chapter IV.

The source term $m^\delta(t, x)$ introduces randomness in the system. It is an ergodic stochastic process defined on $\mathbb{R}^+ \times \mathbb{T}^d$. Note that this process does not depend on the velocity variable. The evolution of m^δ with respect to the time is scaled as t/δ^2 . The asymptotic behavior depends on κ . In our work, we consider the two cases $\kappa = 1$ and $\kappa = 0$. When $\delta \rightarrow 0$, the ergodic process m^δ will converge in distribution to its unique invariant distribution. Thus, when $\kappa = 1$, there cannot be a limit if the average of this invariant distribution is non zero. Therefore, when $\kappa = 1$, we assume that this invariant distribution is centered.

Since m^δ evolves at speed t/δ^2 with $\delta \rightarrow 0$, it is referred to as the fast component, while $\rho^{\varepsilon, \delta} = \int_V f^{\varepsilon, \delta} d\mu$ is referred to as the slow component. The slow component is $\rho^{\varepsilon, \delta}$ instead of $\int_V f^{\varepsilon, \delta} d\mu$ because the stiffest term of (1) is $\frac{1}{\varepsilon^2} Lf^{\varepsilon, \delta}$, we must thus have asymptotically $Lf = 0$, namely $f = \rho\mathcal{M}$. Slow-fast systems with separated time-scales appear in various situations, in a deterministic setting both as ordinary differential equations (ODEs) and partial differential equations (PDEs), but also in a stochastic setting, both as stochastic differential equations (SDEs) and stochastic partial differential equations (SPDEs). For instance they appear in physics when studying a limit of a system of particles [PV03, GR09] or in astrophysics when studying dust grain in a macroscopic structure [LBL20]. We can also mention the study of semi-conductors [GP92]. Among these examples, semi-conductors and systems of particles can both be modeled using kinetic slow-fast models.

In this work, we study the model (1) both from theoretical and numerical points of view. The theoretical study (Chapters I and II) consists in proving the convergence of the solution to the solution of a limiting equation. The numerical study (Chapters III and IV) consists in designing numerical schemes for the slow-fast system that are both consistent for fixed $\varepsilon > 0$, $\delta > 0$ and in the asymptotic regimes $\varepsilon \rightarrow 0$ for fixed $\delta > 0$, $\delta \rightarrow 0$ for fixed $\varepsilon > 0$ and $\varepsilon = \delta \rightarrow 0$. The scheme must admit a limiting scheme for any fixed time-step size in these regimes, and this limiting scheme must be consistent with the limiting equation derived in the theoretical study. Such schemes are called Asymptotic Preserving (AP).

This manuscript is organized as follows.

In Chapter I, we consider the model with $\varepsilon = \delta \rightarrow 0$ and $\kappa = 1$. We prove a diffusion approximation result, both in the PDE and in the stochastic sense: when $\varepsilon \rightarrow 0$, the slow component ρ^ε converges in distribution to the solution ρ of a linear parabolic diffusion equation with a diffusive source term (namely a white noise in time). This work is an improvement of [DV12] in the sense that we prove the same result while relaxing the assumption of boundedness on the source term m^ε . This generalization is non trivial as it requires the use of more intricate tools, including stopping times and stopped martingale processes. This chapter is based on the preprint [RR20].

In Chapter II, we consider the model with $\varepsilon = \delta \rightarrow 0$ and $\kappa = 0$. We prove a diffusion approximation in the PDE sense and an averaging principle in the stochastic sense: when $\varepsilon \rightarrow 0$, the slow component ρ^ε converges in distribution to the solution ρ of a parabolic diffusion equation governed by the stationary average value of the invariant distribution

of the ergodic driving process. Chapter II is technically simpler than Chapter I but this manuscript is organized in the chronological order of the work of the thesis. A preprint based on this chapter is in preparation.

In Chapter III, we design numerical schemes for (2), both in the diffusion approximation regime and in the averaging regime. We define a notion of AP scheme in distribution, adapted to the multiscale stochastic problems we consider, and prove this AP property for our schemes, give examples of non AP schemes and illustrate numerically the performance of the different schemes. The superiority of the AP schemes can be clearly seen in these experiments, both qualitatively and quantitatively. Moreover, we prove a uniform accuracy property for the AP scheme in the averaging regime, namely an error estimate when the time-step size $\Delta t \rightarrow 0$ which is uniform in $\delta > 0$. This chapter is based on the preprint [BRR20], to which we add some quantitative numerical experiments.

In Chapter IV, we combine the schemes of Chapter III and some already known AP schemes for PDEs to design AP schemes for our SPDE, using a splitting procedure. We illustrate the efficiency of the proposed schemes in both the averaging regime $\kappa = 0$ and the diffusion approximation regime $\kappa = 1$. We consider both cases $\varepsilon = \delta$ and $\varepsilon \neq \delta$, so that we can fix one of these two parameters and let the other one converge to 0.

2 Convergence of (1) in the diffusion approximation regime

Chapter I is based on the preprint [RR20] and is devoted to proving the following convergence for $\rho^\varepsilon = \int_V f^\varepsilon d\mu$ where f^ε is the solution of (1) with $\varepsilon = \delta$ and $\kappa = 1$.

Theorem. *Let W be a cylindrical Wiener process on $L^2(\mathbb{T}^d)$, ρ_0 be a random variable in $L^2(\mathbb{T}^d)$ and ρ be the weak solution of the linear stochastic diffusion equation*

$$d\rho = \operatorname{div}(K\nabla\rho)dt + \rho Q^{1/2} \circ dW(t), \quad (3)$$

with initial condition $\rho(0) = \lim_{\varepsilon \rightarrow 0} \rho_0^\varepsilon$. The matrix K and the operator Q are defined by (I.2.2) and (I.2.5)

Under appropriate assumptions, for all $\sigma > 0$ and $T > 0$, the density ρ^ε converges in distribution to ρ in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ and in $L^2([0, T], L^2(\mathbb{T}^d))$.

This is a diffusion approximation result, both in the PDE and stochastic sense. This theorem is a generalization of the main result of [DV12]. First, we relax the condition of boundedness of the driving process m^ε into a condition on its moments that is satisfied for instance by an Ornstein-Uhlenbeck process. Second, we add a stronger convergence (in $L^2([0, T], L_x^2)$) under an additional assumption which allows to use an averaging lemma [BD99, Theorem 2.3]. Note that the convergence only holds in distribution in general.

2.1 Context

Results of diffusion approximation for deterministic kinetic equations (hence in the PDE sense) are common in the literature, starting with [LK74, BLP79]. In both these articles, the small parameter ε is the ratio between the mean path between collisions of particles in a domain to the size of the domain. The method is to construct an asymptotic expansion of the solution of the kinetic equation with respect to ε . In the case of (1) with $m^\varepsilon = 0$,

[DGP00] proves the (deterministic) convergence using a similar expansion $f^\varepsilon = \rho^\varepsilon \mathcal{M} + \varepsilon g^\varepsilon$. Note that the term of order 0 depends on f^ε only through ρ^ε , justifying why we consider the convergence of ρ^ε instead of f^ε .

The diffusive limit in the stochastic sense is also standard: when a correctly scaled random term is introduced in a differential equation, it induces a white noise in time in the limiting equation, and the convergence towards the limiting equation holds only in distribution in general. The study of such convergence results has been initiated by Khasminskii [Kha66a, Kha66b]. In these papers, the results are proved by establishing the convergence of the finite dimensional distributions.

In [PSV77], the authors introduce another approach based on martingale formulation and perturbed test functions. Let us explain the idea of this approach, which is heavily used in Chapters I and II. Let X be a stochastic process and \mathcal{L} be an operator. Set

$$M_\varphi(t) = \varphi(X(t)) - \varphi(X(0)) - \int_0^t \mathcal{L}\varphi(X(s))ds$$

for all test function φ in a class of functions Θ . The process X solves the martingale problem associated to \mathcal{L} if for all φ in Θ , M_φ is a martingale. Then, under some assumptions on the class Θ , the process X is a Markov process of infinitesimal generator \mathcal{L} if and only if it solves the martingale problem associated to \mathcal{L} . Therefore, if we consider a solution X^ε of a SPDE of generator \mathcal{L}^ε , we get a family of martingale at level $\varepsilon > 0$. Assume the family $(X^\varepsilon)_{\varepsilon>0}$ is tight and consider X a limit point. If one can take the limit $\varepsilon \rightarrow 0$, one can hope to get a limiting martingale problem that allows to identify X using a uniqueness result, and thus to get the convergence in distribution. However, for multiscale singular problems, $\mathcal{L}^\varepsilon \varphi(X^\varepsilon(t))$ does not converge, hence one cannot take the limit. To remediate this, the idea is to perturb the test function φ into a test function $\varphi^\varepsilon = \varphi + \varepsilon \varphi_1 + \varepsilon^2 \varphi_2 + \dots$, such that $\mathcal{L}^\varepsilon \varphi^\varepsilon \rightarrow \mathcal{L}\varphi$ for some operator \mathcal{L} . Assuming the perturbed test function φ^ε belongs to the class Θ , one can use the martingale problem on φ^ε , take the limit to retrieve a martingale problem associated with \mathcal{L} and get the convergence in distribution of X^ε to a Markov process X of infinitesimal generator \mathcal{L} .

Note that perturbed test functions method in the context of PDEs (either deterministic or stochastic) with diffusive limits is not specific to kinetic equations: [Eva89] use them for viscosity solutions, [PP03] for a parabolic PDE and [Mar06, dBD10, DT11, dBG12] for nonlinear Schrödinger equations.

2.2 Strategy for the proof in [DV12]

The scheme of the proof in [DV12], using a martingale approach and perturbed test functions, is as follows:

- Prove the tightness of the family of solution $(\rho^\varepsilon)_{\varepsilon>0}$. This requires some estimates on the moments of ρ^ε , uniformly in ε .
- Construct a perturbation φ^ε of a test function φ . In (1), the infinitesimal generator can be decomposed as $\mathcal{L}^\varepsilon = \varepsilon^{-1} \mathcal{L}_1 + \varepsilon^{-2} \mathcal{L}_2$. Owing to the asymptotic expansion $\varphi^\varepsilon = \varphi + \sum \varepsilon^i \varphi_i$, the convergence $\mathcal{L}^\varepsilon \varphi^\varepsilon \rightarrow \mathcal{L}\varphi$ only holds if each corrector φ_i solves the solution of the Poisson equation associated with the operator \mathcal{L}_2

$$\mathcal{L}_1 \varphi_{i-1} + \mathcal{L}_2 \varphi_i = \vartheta_{i-1},$$

for some right-hand side ϑ_{i-1} that only depends on φ_j for $j < i$. Under appropriate assumptions, the solution of this equation is well-defined and one can construct successively the correctors φ_j .

- Since the family $(\rho^\varepsilon)_{\varepsilon>0}$ is tight, there exists a limit point ρ for the convergence in distribution and a sequence $\varepsilon_i \xrightarrow{i \rightarrow \infty} 0$ such that ρ^{ε_i} converges to X in distribution. Then, one can take the limit of the martingale property

$$\mathbb{E} \left[\varphi^{\varepsilon_i}(\rho^{\varepsilon_i}(t)) - \varphi^{\varepsilon_i}(\rho^{\varepsilon_i}(0)) - \int_0^t \mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(\rho^{\varepsilon_i}(s)) ds \mid \mathcal{F}_s^{\rho^{\varepsilon_i}}, s \leq t \right] = 0$$

to get that ρ is the solution of the martingale problem associated to \mathcal{L} . Assuming the uniqueness of such a solution, this concludes the proof of the convergence in distribution.

2.3 Strategy for the proof in Chapter I

Let us now describe the main novelties of our work. In [DV12], the driving process m^ε is assumed to be bounded. This is a strong assumption, which is not satisfied by a Gaussian process for instance. In Chapter I, the objective is to relax this assumption. As a consequence, uniform moment estimates are missing. To overcome this issue, we introduce stopping times τ_Λ^ε depending on a arbitrarily large real number $\Lambda \in (0, \infty)$, such that, for fixed Λ , we have an estimate on the stopped process $\rho^{\varepsilon, \tau_\Lambda^\varepsilon}$ uniformly in ε . More precisely, the stopping time is defined as

$$\tau_\Lambda^\varepsilon \doteq \tau^\varepsilon \wedge \tau_\Lambda(\zeta^\varepsilon),$$

where

$$\tau^\varepsilon \doteq \inf \{ t \in \mathbb{R}^+ \mid \|m^\varepsilon(t)\|_E > \varepsilon^{-\alpha} \},$$

for some parameter α , where

$$\tau_\Lambda(z) \doteq \inf \{ t \in \mathbb{R}^+ \mid \|z(t)\|_{\mathcal{C}^1(\mathbb{T}^d)} \geq \Lambda \},$$

and

$$\zeta^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t m^\varepsilon(s) ds = \varepsilon \int_0^{t/\varepsilon^2} m(s) ds.$$

We are able to prove the tightness of the family $(\rho^{\varepsilon, \tau_\Lambda^\varepsilon})_{\varepsilon>0}$ and to follow the standard proof with stopped martingales instead of martingales. A major difficulty of this framework comes from the asymptotic behavior of the stopping times. On the one hand, we prove a Central Limit Theorem on m^ε stating that ζ^ε converges in distribution to a Wiener process ζ , hence $\tau_\Lambda(\zeta^\varepsilon) \rightarrow \tau_\Lambda(\zeta)$ in distribution for all Λ outside of a countable set. On the other hand, for a well-chosen parameter α , $\tau^\varepsilon \rightarrow \infty$ in probability. Hence, the stopping time τ_Λ^ε converges in distribution to $\tau_\Lambda(\zeta)$ when $\varepsilon \rightarrow 0$. For this stopping time to be adapted to the filtration associated with the limiting process, we need to consider the convergence of the coupled process $(\rho^\varepsilon, \zeta^\varepsilon)$ instead of just ρ^ε .

The stopped coupled process $((\rho^{\varepsilon, \tau_\Lambda^\varepsilon}, \zeta^{\varepsilon, \tau_\Lambda^\varepsilon}))_{\varepsilon>0}$ still form a tight family. Then, a limit point for the convergence in distribution satisfies a martingale problem stopped at time $\tau_\Lambda(\zeta)$, hence a family of martingale problems indexed by the parameter Λ . Since this stopped martingale problem only characterizes the dynamics of a limit point $X_\Lambda = (\rho_\Lambda, \zeta_\Lambda)$

up to time $\tau_\Lambda(\zeta)$, there can be multiple solutions of the stopped martingale problem. However, we are able to prove that the family of stopped martingale problems are consistent, in the sense that X_Λ has the same distribution than $X^{\tau_\Lambda(\zeta)}$ for some process X . Since $\tau_\Lambda(\zeta) \rightarrow \infty$ when $\Lambda \rightarrow \infty$, we use these results to deduce the convergence of the unstopped process ρ^ε .

3 Convergence of (1) in the averaging regime

Chapter II is devoted to proving the following convergence results for $\rho^\varepsilon = \int_V f^\varepsilon d\mu$ where f^ε is the solution of (1) with $\varepsilon = \delta$ and $\kappa = 0$.

Theorems. *Let \bar{m} be the average of the unique invariant distribution of m^ε , let ρ_0 be a random variable in $L^2(\mathbb{T}^d)$ and let ρ be the weak solution of the linear stochastic diffusion equation*

$$\partial_t \rho = \operatorname{div}(K \nabla \rho) + \bar{m} \rho, \quad (4)$$

with initial condition $\rho(0) = \lim_{\varepsilon \rightarrow 0} \rho_0^\varepsilon$.

Under appropriate assumptions, for all $\sigma > 0$ and $T > 0$, the density ρ^ε converges in distribution to ρ in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ and in $L^2([0, T], L^2(\mathbb{T}^d))$ and f^ε converges in distribution to $\rho \mathcal{M}$ in $L^2([0, T], L^2(\mathbb{T}^d \times V, \frac{dx d\mu(v)}{\mathcal{M}(v)}))$. Moreover, if ρ_0 is deterministic, then every convergence holds in probability.

This is a diffusion approximation result in the PDE sense as in Chapter I, but an averaging result in the stochastic sense. The averaging principle for stochastic system has been introduced in the classical paper [Kha68]. Numerous results are associated to this principle, using different proofs and different criterion of convergence, see for instance [PS08, Cer09, CF09, Bré12, Bré20]. Some of these papers also state the convergence speed in the averaging principle in SPDEs (see [Bré12, Bré20]). In Chapter II, we only prove convergence. The analysis of convergence rates is left for future works.

The strategy of the proof is similar to Chapter I: we follow the approach based on martingale problems and perturbed test functions, and cope with the unboundedness of the driving process m^ε by stopping the processes and the martingale problems at some stopping time τ^ε . The main difference with Chapter I is that, in this framework, $\tau^\varepsilon \rightarrow \infty$ in probability. Hence, the limiting martingale problem is not stopped at some stopping time: it is a standard martingale problem, and since its solution is unique, we deduce the desired convergence with simpler arguments. More precisely, the auxiliary process is this time defined as

$$\zeta^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t (m^\varepsilon(s) - \bar{m}) ds,$$

where the average \bar{m} of the unique invariant distribution of m^ε is non zero. Then, the stopping times are defined by

$$\begin{aligned} \tau^\varepsilon &= \tau_m^\varepsilon \wedge \tau_\zeta^\varepsilon, \\ \tau_m^\varepsilon &= \inf \left\{ t \in \mathbb{R}^+ \mid \|m^\varepsilon(t)\|_E \geq \varepsilon^{-\alpha} \right\}, \\ \tau_\zeta^\varepsilon &= \inf \left\{ t \in \mathbb{R}^+ \mid \|\zeta^\varepsilon(t)\|_{C_x^1} \geq \varepsilon^{-1} \right\}. \end{aligned}$$

As in Chapter I, for suitable α , we have $\tau_m^\varepsilon \rightarrow \infty$ in probability when $\varepsilon \rightarrow 0$. However, in the hitting time τ_ζ^ε of the auxiliary process, the fixed threshold Λ can be replaced by ε^{-1} , yielding $\tau_\zeta^\varepsilon \rightarrow \infty$ in probability when $\varepsilon \rightarrow 0$.

4 Asymptotic preserving schemes for SDEs

Chapter III is based on the preprint [BRR20] and is devoted to designing AP schemes for (2) in the diffusion approximation and averaging regimes. We work on SDEs first because we noticed that our work already brought non-trivial new contributions in this simpler case. We come back to the SPDE case in Chapter IV. We refer to [PS08] for convergence results on multiscale SDEs: [PS08, Chapter 11] states the convergence in the diffusion approximation regime and [PS08, Chapter 10] states the convergence in the averaging regime.

4.1 Context

Let us motivate this work. In the deterministic case, the field of stiff ODE solvers is well-studied [HW96]. The presence of a stiff term in a system, typically related to the fast component, can induce strong stability conditions of the form $\Delta t = o(\varepsilon^\alpha)$, since one needs the scaled time step size $\Delta t/\varepsilon^\alpha$ to converge to 0. A consequence is an explosion of the computational cost when simulating the system for small ε . When the system follows a stochastic dynamics, the simulation can prove to be even more challenging.

In [FG18], the authors give examples where the SDE satisfied by some statistical quantities of a scheme (for instance the probability density function) differs from the limiting SDE of the continuous-time system. One of these examples is the Itô interpretation of the noise when the limiting SDE uses a Stratonovich interpretation. We investigate this example further below.

In [LAE08], the authors also illustrate how a stiff ODE solver may not be effective in the case of stiff SDEs, when the invariant measure is non-Dirac, and how these solvers may produce wrong solutions when the fast scale dynamics are not resolved. Even worse, they prove how a discretization X_n^ε may converge to a scheme X^c when $\Delta t \rightarrow 0$ and $\varepsilon \rightarrow 0$ with fixed ratio $c = \Delta t/\varepsilon^\alpha$, where the scheme X^c heavily depends on c .

A well-studied way of avoiding these issues is to approximate the limiting model for the slow variables of interest, while only partial but relevant information coming from the fast dynamics is taken into account. Let us mention some of the many methods that follow this approach.

The Heterogeneous Multiscale Method (HMM) [AEEVE12, Bré13, ELVE05] is based on a macroscopic solver that performs computations for the microscopic dynamics when needed. More precisely, one computes the macroscopic dynamics using a macroscopic solver. When the macroscopic solver needs some data that cannot be computed using this solver, typically data that can only be read at the microscopic level, one may perform a series of microscopic computations constrained to be consistent with the value of the macroscopic dynamics at that time. Then, using these computations, one may compute an estimation of the missing data and inject it back to the macroscopic solver. Similarly, the projective integration method [GKK06] is based on a solver for the macroscopic component and on short runs of a solver for the microscopic dynamics, to avoid high computational cost while estimating the effect of the microscopic component on the macroscopic one. Due to their approximation of the microscopic dynamics, the projective integration method and the HMM can be called averaging-based methods.

The equation-free coarse-grained method [KGH⁺03] instead approximate the evolution of the macroscopic component through appropriately initialized calls to a microscopic solved

for short times and on small spatial domains, for example by integrating the dynamics over long times or large spatial domains. In particular, this approach is equation-free in the sense that one does not always need to derive the macroscopic evolution equation from the microscopic one.

More recently, in [APV17], the authors design a spectral method. The coefficients of the homogenized equation are computed using solutions of appropriate Poisson equations. These Poisson equations are solved using spectral Hermite methods, namely an expansion in Hermite functions.

In [VZS20], the authors describe a micro-macro acceleration method. The algorithm intersperse alternatively a microscopic solver for short times and an extrapolation of the macroscopic state over a longer time interval. Then, they compute the probability distribution that is consistent with this extrapolation, while minimally modifying the computed microscopic distribution. This allows them to reconstruct the microscopic state.

Finally, [LLMS20] present a micro-macro version of the parareal algorithm, which parallelize the computations with respect to the time variable, allowing to significantly reduce the computation time.

4.2 Main contributions of Chapter III

For the class of SDEs we focus on in this manuscript, standard schemes also perform badly in the asymptotic regime. Let us first consider the diffusion approximation regime. Consider the particular case of (2)

$$\begin{cases} dX_t^\varepsilon = \frac{\sigma(X_t^\varepsilon)m_t^\varepsilon}{\varepsilon}dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon^2}dt + \frac{1}{\varepsilon}d\beta_t. \end{cases}$$

If one takes $\sigma(x) = x$, it is a finite dimensional version of (1) with $\varepsilon = \delta$, $\kappa = 1$ and where m^ε is an Ornstein-Uhlenbeck process scaled as t/ε^2 . For simplicity, assume $d = 1$. When $\varepsilon \rightarrow 0$, the diffusion approximation result (see [PS08, Chapter 11]) states that X^ε converges in distribution to the solution X of the SDE

$$dX_t = \sigma(X_t) \circ d\beta_t,$$

where the noise is interpreted in the Stratonovich sense, the Itô form of this SDE being

$$dX_t = \frac{1}{2}\sigma'(X_t)\sigma(X_t)dt + \sigma(X_t)d\beta_t.$$

For fixed $\varepsilon > 0$, the numerical scheme

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \sigma(X_n^\varepsilon)\frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2}m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon}\gamma_n, \end{cases}$$

where $(\gamma_n)_{n \in \mathbb{N}_0}$ is an i.i.d. sequence of standard Gaussian random variables, is consistent with the system. However, when $\varepsilon \rightarrow 0$, X_n^ε converges to X_n given by

$$X_{n+1} = X_n + \sqrt{\Delta t}\sigma(X_n)\gamma_n.$$

This limiting scheme is consistent with the Itô interpretation of the noise, instead of the correct Stratonovich one, and therefore fails to capture the asymptotic behavior of the system. In Chapter III, we design an AP scheme, namely a scheme for which the following diagram is commutative

$$\begin{array}{ccc} X_N^\varepsilon & \xrightarrow[\Delta t \rightarrow 0]{} & X^\varepsilon(T) \\ \downarrow \varepsilon \rightarrow 0 & & \varepsilon \rightarrow 0 \downarrow \\ X_N & \xrightarrow[\Delta t \rightarrow 0]{} & X(T) \end{array}$$

The AP scheme is based on a prediction-correction procedure:

$$\left\{ \begin{array}{l} \hat{m}_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(X_n^\varepsilon) \hat{m}_{n+1}^\varepsilon}{\varepsilon^2} + \frac{\Delta t f(X_n^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon}, \\ \hat{X}_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \sigma(X_n^\varepsilon) \frac{\Delta t \hat{m}_{n+1}^\varepsilon}{\varepsilon}, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon) m_{n+1}^\varepsilon}{\varepsilon^2} + \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon}, \\ Y_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \sigma(X_n^\varepsilon) \frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}, \\ X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \frac{\sigma(X_n^\varepsilon) + \sigma(Y_{n+1}^\varepsilon)}{2} \frac{\Delta t}{\varepsilon} \frac{\hat{m}_{n+1}^\varepsilon + m_{n+1}^\varepsilon}{2}. \end{array} \right.$$

This scheme is efficient for a large class of SDEs, including the previous example where it captures the Stratonovich interpretation of the noise, but also in the system considered in [LBL20], where the limiting equation contains a noise-induced drift term that is captured by our AP scheme and not by a standard scheme.

In the averaging regime, the same kind of problems occurs. Consider the system

$$\left\{ \begin{array}{l} dX_t^\varepsilon = b(X_t^\varepsilon, m_t^\varepsilon) dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon} dt + \frac{\sqrt{2}}{\sqrt{\varepsilon}} d\beta_t. \end{array} \right.$$

If one takes $b(x, m) = xm$, it is a finite dimensional version of (1) with $\varepsilon = \delta$, $\kappa = 0$ and where m^ε is an Ornstein-Uhlenbeck process scaled as t/ε . When $\varepsilon \rightarrow 0$, the averaging principle (see [PS08, Chapter 10]) states that X^ε converges in distribution to the solution X of the ODE

$$\dot{X} = \bar{b}(X),$$

where $\bar{b}(x) = \int b(x, m) d\nu(m)$ and $\nu = \mathcal{N}(0, 1)$ is the standard Gaussian random variable. For fixed $\varepsilon > 0$, the numerical scheme

$$\left\{ \begin{array}{l} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon), \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon + \sqrt{\frac{2\Delta t}{\varepsilon}} \gamma_n, \end{array} \right.$$

is consistent with the system. However, $X_n^\varepsilon \xrightarrow[\varepsilon \rightarrow 0]{} X_n$ where X_n is given by

$$X_{n+1} = X_n + \Delta t b(X_n, 0),$$

which is in general inconsistent with the limiting equation. Once again, we design an AP scheme such that the aforementioned diagram is commutative, based on an exact (in distribution) discretization of the fast component:

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon) + \sqrt{\Delta t} \sigma(X_n^\varepsilon, m_{n+1}^\varepsilon) \Gamma_n \\ m_{n+1}^\varepsilon = e^{-\Delta t/\varepsilon} m_n^\varepsilon + \sqrt{1 - e^{-2\Delta t/\varepsilon}} h(X_n^\varepsilon) \gamma_n. \end{cases} \quad (5)$$

After constructing these AP schemes, we present three types of numerical experiments. Each of them illustrate the superiority of the AP schemes compared to crude schemes:

- The first one is qualitative: we plot the trajectories of the discretized solution as a function of the time t_n , for a fixed value of the time-step size Δt and for different values of ε . The trajectories are compared to a reference scheme.
- The second one is quantitative: we plot an estimation of the weak total error

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]|,$$

for some test functions φ , as a function of Δt for different values of ε .

- The third one is also quantitative: we plot an estimation of the weak diffusion approximation or averaging error at the discrete time level

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)]|,$$

for some test functions φ , as a function of ε for different values of Δt .

In both quantitative experiments, the observed results for the AP scheme (5) are consistent with the error estimates below and gives higher order of convergence than what we manage to prove. Meanwhile, the non AP scheme does not perform well for small ε , the error with respect with ε (respectively Δt) being non uniform in Δt (respectively ε). The order $\frac{1}{2}$ may not be optimal, since the numerical experiments show an order 1.

In the diffusion approximation regime, we do not have a general uniform accuracy result, except in some particular cases.

4.3 Error estimates in the averaging regime

In the averaging regime, prove a uniform accuracy theorem for the AP Scheme (5). In this case, the fast component m^ε scales as t/ε .

Theorem. *Under appropriate assumptions, for any $T \in (0, \infty)$ and any function $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$ of class \mathcal{C}^4 , there exists $C(T, \varphi) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ and $\varepsilon \in (0, 1]$ one has*

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \min\left(\frac{\Delta t}{\varepsilon}, \Delta t + \varepsilon\right), \quad (6)$$

and the AP scheme is uniformly accurate with the following error estimate: for all $\Delta t \in (0, \Delta t_0]$, one has

$$\sup_{\varepsilon \in (0, 1]} |\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \sqrt{\Delta t}. \quad (7)$$

To the best of our knowledge, this is the first uniform accuracy result for an AP scheme for stochastic equations.

The idea of the proof of uniform accuracy is to estimate the left-hand side of (6) using two different paths on the commutative diagram

$$\begin{array}{ccc} X_N^\varepsilon & \xrightarrow[\Delta t \rightarrow 0]{A} & X^\varepsilon(T) \\ B \downarrow \varepsilon \rightarrow 0 & & \varepsilon \rightarrow 0 \downarrow D \\ X_N & \xrightarrow[C]{\Delta t \rightarrow 0} & X(T) \end{array}$$

where A , B , C and D denotes the errors in the different convergences:

- The estimate $O(\frac{\Delta t}{\varepsilon})$ comes from the direct path $|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq A$. Since we use a classical scheme with time-step size $\Delta t/\varepsilon$, we are able to prove that $A = O(\Delta t/\varepsilon)$ using standard arguments (Kolmogorov equation).
- The estimate $O(\Delta t + \varepsilon)$ comes from the indirect path $|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq B + C + D$. Since D is the error is the averaging principle, we have $D = O(\varepsilon)$. Proving a nice error estimate for the error term B is the most challenging part: we adapt the proof for D to the discrete case and are able to get $B = O(\max(\Delta t, \varepsilon))$. The last error C is estimated using Kolmogorov and Poisson equations, and we get $C = O(\Delta t)$.

This leads to (6) by writing $|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq \min(A, B + C + D)$ and to (7) by using the first or the second estimate depending on whether $\sqrt{\Delta t} \leq \varepsilon$ or $\varepsilon \leq \sqrt{\Delta t}$.

5 Asymptotic preserving schemes for kinetic stochastic PDEs

Using the work of Chapter III on SDEs, we go back to our SPDE framework: Chapter IV is devoted to designing asymptotic preserving schemes for (1) in the diffusion approximation and averaging regimes. Unlike Chapters I and II, we do not only consider the case $\varepsilon = \delta$, but allow the two parameters to converge to 0 independently from one another.

5.1 Context in the deterministic case

The notion of AP schemes in the deterministic case has been introduced in [Jin99], for applications to multiscale kinetic Partial Differential Equations (PDEs), which converge to parabolic diffusion PDEs. Our system (1) is a particular case of this setting if $m^\delta = 0$. By now, various AP schemes have been designed. Let us describe some of them.

For kinetic BGK models with the hyperbolic scaling

$$\partial_t f^\varepsilon + a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon} L f^\varepsilon,$$

the review articles [DP14, Section 7] and [Jin12] both mention a splitting scheme. The splitting in time is done by solving first the collision term $L f^\varepsilon$ and then the convection term $a(v) \cdot \nabla_x f^\varepsilon$. One can then focus on the collision term and use a upwind explicit

scheme for the convection term: the splitting scheme is AP if the scheme for the collision term is also AP. For the collision term, [DP14, Section 7] gives a sufficient condition for a Runge-Kutta method to be AP, while [Jin12] gives an AP implicit scheme.

The two review articles also mention a family of non splitting Implicit-Explicit scheme (IMEX). For instance, an IMEX scheme can be obtained by using an implicit discretization of the collision term while the discretization of the convection term is explicit. In both those methods (splitting and IMEX), even though the scheme is at least partially implicit, the scheme can be expressed explicitly and one does not need to invert a system of nonlinear equations.

In [Pup19, Section 4], the author mentions an AP scheme based on a upwind space discretization with high order scheme in space for the equation.

For kinetic BGK models with the parabolic scaling as in (1), namely

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon,$$

some AP schemes are given in [Hiv18, LM08] for instance.

A first AP scheme can be designed using a micro-macro decomposition of f^ε as $f^\varepsilon = \rho^\varepsilon \mathcal{M} + \varepsilon g^\varepsilon$ and designing a scheme based on the equations satisfied by ρ^ε and g^ε . A second AP scheme is obtained using an fully implicit discretization of the equation satisfied by \hat{f}^ε the Fourier transform in space of f^ε . This second scheme is detailed in Chapter IV. Even though they are asymptotic preserving, none of these schemes are uniformly accurate, in the sense that the precision of the scheme is not uniform in ε .

One can design a uniformly accurate AP scheme as follows: using an integral formulation of the equation. More precisely, one writes the equation solved by \hat{f}^ε the Fourier transform in space of f^ε , expresses the Duhamel form of \hat{f}^ε and integrate this expression with respect to the velocity variable. By discretizing the integral in time (due to the Duhamel form) using a quadrature rule of order 2, [Hiv18] shows that one obtains a uniformly accurate AP scheme on $\hat{\rho}^\varepsilon$. We do not use this scheme in the numerical experiments, preferring a simpler implementation. Application of this uniformly accurate in a stochastic setting will be investigated in future work.

5.2 Context in the stochastic case, in the diffusion approximation regime

In a stochastic setting, an additional difficulty arises: in general, the convergence only holds in distribution. Thus, the AP property must be formulated for the convergence in distribution. This corresponds exactly to the definition of AP scheme introduced in Chapter III.

The aforementioned works on deterministic dynamics have been extended to a stochastic framework. In [HJ17, Jin18, JLP18], the authors study AP schemes for PDEs with random coefficients. In [DPS18, RLJ14], the authors use Monte-Carlo methods for deterministic problems. However, these methods do not apply to SPDEs.

In [AF19], the authors deal with the kinetic SPDE

$$df^\varepsilon + \frac{1}{\varepsilon} v \partial_x f^\varepsilon dt = \frac{\sigma}{\varepsilon^2} \mathcal{L} f^\varepsilon dt + f \circ Q dW_t,$$

driven by a Q -Wiener process, which is this time very close to our framework (1). More precisely, we recover this setting after taking the limit $\delta \rightarrow 0$ in (1) in the diffusion

approximation regime, albeit with a different operator L than ours. They design an AP scheme using the micro-macro decomposition $f^\varepsilon = \rho^\varepsilon \mathcal{M} + \varepsilon g^\varepsilon$ and an upwind discretization of the space derivatives. As a consequence, the limiting scheme is explicit and requires a CFL condition for stability. In Chapter IV, we construct an AP scheme with no CFL condition and when the noise is not necessarily a Wiener process.

5.3 Main contributions of Chapter IV

Our construction of AP schemes is based on splitting (1) into a deterministic part

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon, \quad (8)$$

and a stochastic part

$$\partial_t f^\delta = \frac{1}{\delta^\kappa} m^\delta f^\delta. \quad (9)$$

Therefore, we design an AP scheme using a splitting procedure by combining one of the aforementioned AP schemes for (8) and an AP scheme for (9). Note that (9) is simpler than the general system (2) considered in Chapter III, since it is linear in m^δ . As a consequence, one can either use the general schemes designed in Chapter III or some specific schemes.

We present three types of experiments to illustrate the efficiency of our AP schemes compared to crude schemes:

- The first two types are qualitative: we plot the trajectory of $\|\rho^{\varepsilon,\delta}(t)\|_{L^2}$ as a function of t or the value at final time $\rho^{\varepsilon,\delta}(T, x)$ as a function of x , for a fixed value of the time-step size τ and for different values of ε and δ .
- The third type of experiments is quantitative: we plot an estimation of the weak error

$$|\mathbb{E}[\varphi(\rho^{\varepsilon,\delta}(T))] - \mathbb{E}[\varphi(\rho(T))]|$$

depending on the time-step size τ , for different values of ε and δ .

6 Perspectives

6.1 Results to investigate

We left some results on our models for future works. In the theoretical part of this manuscript, the diffusion approximation regime (Chapter I) can be seen as a Central Limit Theorem regime for the averaging regime (Chapter II). The former may help getting information on the latter for example to get the convergence rate of $f^{\varepsilon,\delta}$ towards $\rho \mathcal{M}$.

In the numerical part of this manuscript (Chapters III and IV). For instance, we only proved uniform accuracy in the averaging regime for SDEs. It is worth investigating the same type of results in the diffusion approximation regime.

In the SPDE case, one should study numerical stability for our AP schemes more precisely, for instance by getting L^2 moments uniformly with respect to the time-step size. Then, one could study the convergence rate, both from a theoretical and a numerical point of view. It should be noted that the numerical study of the convergence rate is computationally heavy and requires more efficient implementations of our schemes.

There is little hope for our AP schemes to be uniformly accurate since the deterministic part of the splitting is computed using a non uniformly accurate scheme. However, the splitting procedure allows to replace this part by any AP scheme, for instance the integral formulation-based scheme given in [Hiv18], which is both AP and uniformly accurate. The splitting scheme based on two uniformly accurate scheme will hopefully be uniformly accurate.

6.2 Towards other models

Another possible prolongation of this work is to extend both the theoretical and the numerical results, to weaker assumptions or to related models.

For instance, we believe that our results still holds without the boundedness assumption on the velocity $a(v)$, and removing this assumption is essential to retrieve the classical framework $a(v) = v$ with $V = \mathbb{R}^d$. However, relaxing this assumption brings non trivial technical complications. Similarly, relaxing the assumption $\mu(V) < \infty$ is not straightforward. For instance, when $\mu(V) = \infty$ in the averaging regime, the functional space $L_T^2 L^2(\mathcal{M}^{-1}) \doteq L^2([0, T], L^2(\mathbb{T}^d \times V, \frac{dx d\mu(v)}{\mathcal{M}(v)}))$ may not be suitable to study $f^{\varepsilon, \delta}$, since $\rho \mathcal{M} \notin L_T^2 L^2(\mathcal{M}^{-1})$ and $f^{\varepsilon, \delta} \rightarrow \rho \mathcal{M}$ when $\mu(V) < \infty$.

Another perspective is to adapt the generality of our SDE model (chapitre III) to our SPDE model. For instance, one could try to get the results of Chapters I et II after replacing the linear source term $m^\delta f^{\varepsilon, \delta}$ by a general term $b(m^\delta) f^{\varepsilon, \delta}$. One could also investigate coupled models, where the fast component depends on the slow component. We only studied coupled SDE, but coupled SPDEs are a natural extension of our work. Note that in the SDE case, a non trivial coupling can induce a noise-induced drift term in the diffusion approximation regime. A recent preprint [CX20] studies the parabolic case and confirms the addition of a noise-induced drift term in the limiting equation.

Another interesting perspective comes from the fact that our numerical schemes heavily relies on the fast component being an Ornstein-Uhlenbeck process when the slow component is frozen. Designing AP schemes when the fast component solves a general ergodic SDE with nonlinear coefficients is not straightforward.

Finally, our results could be proved for different related models, for instance when the source term $\delta^{-\kappa} m^\delta f^{\varepsilon, \delta}$ is replaced by a forcing term $\delta^{-\kappa'} m^\delta \cdot \nabla_v f^{\varepsilon, \delta}$, where we expect $\kappa' = 2$ in the diffusion approximation regime and $\kappa' = 1$ in the averaging regime. This model has already been studied from a theoretical point of view [DV20], but to the best of our knowledge, there exists no AP schemes for such kinetic stochastic PDEs. A splitting-based approach should reveal efficient.

Chapter I

Convergence in the diffusion approximation regime with an unbounded driving process

This chapter is a detailed version of the article *Diffusion limit for a stochastic kinetic problem with unbounded driving process* [RR20]. The preprint can be found at <https://arxiv.org/abs/2009.10406>.

I.1 Introduction

Our aim in this chapter is to study the scaling limit of a stochastic kinetic equation in the diffusion approximation regime, both in Partial Differential Equation (PDE) and probabilistic senses. For deterministic problems, this is a thoroughly studied field in the literature, starting historically with [LK74, BLP79]. Kinetic models with small parameters appear in various situations, for example when studying semi-conductors [GP92] and discrete velocity models [LT97] or as a limit of a particle system, either with a single particle [GR09] or multiple ones [PV03]. It is important to understand the limiting equations, which are in general much easier to simulate numerically. For instance, in the asymptotic regime we study, the velocity variable disappears at the limit.

When a random term with the correct scaling (here t/ε^2) is added to a differential equation, it is classical that, when $\varepsilon \rightarrow 0$, the solution may converge in distribution to a diffusion process, which solves a Stochastic Differential Equation (SDE) driven by a white noise in time. This is a diffusive limit in the probabilistic sense. Such convergence has been proved initially by Khasminskii [Kha66a, Kha66b] and then, using the martingale approach and perturbed test functions, in the classical article [PSV77] (see also [Kus84, EK86, FGPS107, PS08, dBG12]). The use of perturbed test functions in the context of PDEs with diffusive limits also concerns various situations, for instance in the context of viscosity solutions [Eva89], nonlinear Schrödinger equations [Mar06, dBD10, DT11, dBG12], a parabolic PDE [PP03] or, as in this chapter, kinetic SPDEs [DV12, DV20, DRV20].

In this chapter, we consider the following equation

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon + \frac{1}{\varepsilon} f^\varepsilon \bar{m}^\varepsilon, \quad (\text{I.1.1})$$

$$f^\varepsilon(0) = f_0^\varepsilon. \quad (\text{I.1.2})$$

where f^ε is defined on $\mathbb{R}^+ \times \mathbb{T}^d \times V$, L is a linear operator (see (I.1.3) below) and the source term \bar{m}^ε is a random process defined on $\mathbb{R}^+ \times \mathbb{T}^d$ (satisfying assumptions given in Section I.2.1). The goal of this chapter is to study the limit $\varepsilon \rightarrow 0$ of its solution f^ε , and to generalize previous results of [DV12].

The solution $f^\varepsilon(t, x, v)$ is interpreted as a probability distribution function of particles, having position x and velocity $a(v)$ at time t . The variable v belongs to a measure space (V, μ) , where μ is a probability measure. The function a models the velocity.

The Bhatnagar-Gross-Krook operator L expresses the particle interactions, defined on $L^1(V, \mu)$ by

$$L f = \rho \mathcal{M} - f, \quad (\text{I.1.3})$$

where $\rho \doteq \int_V f d\mu$ and $\mathcal{M} \in L^1(V)$.

The source term \bar{m}^ε is defined as

$$\bar{m}^\varepsilon(t, x) = \bar{m}(t/\varepsilon^2, x), \quad (\text{I.1.4})$$

where \bar{m} is a random process, not depending on ε .

In the deterministic case $\bar{m}^\varepsilon = 0$, such a problem occurs in various physical situations [DGP00]. The density $\rho^\varepsilon \doteq \int_V f^\varepsilon d\mu$ converges to the solution of the linear parabolic equation

$$\partial_t \rho - \operatorname{div}(K \nabla \rho) = 0, \quad (\text{I.1.5})$$

on $\mathbb{R}^+ \times \mathbb{T}^d$. This is a diffusive limit in the PDE sense, since the limit equation is a diffusion equation.

In this chapter, the diffusion limit of (I.1.1) is considered simultaneously in the PDE and in the probabilistic sense. The main result, Theorem I.2.1, establishes that, under appropriate assumptions, the density $\rho^\varepsilon = \int_V f^\varepsilon d\mu$ converges in distribution in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ for any $\sigma > 0$ and in $L^2([0, T], L^2(\mathbb{T}^d))$ to the solution of the stochastic linear diffusion equation

$$d\rho = \operatorname{div}(K \nabla \rho) dt + \rho \circ Q^{1/2} dW(t),$$

with K as in (I.1.5). The equation is written in Stratonovich form and is driven by a cylindrical Wiener process W , the covariance operator Q being trace-class. As usual in the context of diffusion limit, the stochastic equation involves a Stratonovich product. The diffusive limit in the stochastic case has been first proved in [DV12], under a restrictive condition on the driving random term: \bar{m} is bounded almost surely. The boundedness of \bar{m} is a strong assumption, which is not satisfied by an Ornstein-Uhlenbeck process for instance. The contribution of this chapter is to relax this assumption: we prove the convergence under a moment bound assumption for the driving process.

The main tools of [DV12] are the perturbed test function method and the concept of solution in the martingale sense. Our general strategy for the proof is similar, therefore those tools are also used here. The main novelty is the introduction of stopping times to obtain the estimates required to establish tightness and convergence. Indeed, relaxing the

conditions on \overline{m} implies that moments of the solutions are not controlled (exponential moments for \overline{m} would be necessary). The strategy from [DV12] needs to be substantially modified: the martingale problem approach is combined with the use of stopping times. At the limit, the stopping times persist, thus the limit processes solves the limit martingale problem only up to a stopping time. We manage to identify them nonetheless as a stopped version of the global solution.

This chapter is organized as follows: in Section I.2, we set some notation, the assumptions on the driving random term and the main result, Theorem I.2.1. Section I.3 states some auxiliary results that are used in the later sections. In Section I.4, we introduce the notion of martingale problem and the perturbed test function method that are used to prove the convergence. In Section I.6, we prove the tightness of the family of processes $((\rho^\varepsilon, \tau_\Lambda^\varepsilon, \zeta^\varepsilon, \tau_\Lambda^\varepsilon))_\varepsilon$ stopped at the random time τ_Λ^ε . Section I.7 takes the limit when $\varepsilon \rightarrow 0$ in the martingale problems and establishes the convergence of ρ^ε in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$. In Section I.8, we prove the convergence in a stronger sense, namely in $L^2([0, T], L^2(\mathbb{T}^d))$, using an additional assumption and an averaging lemma.

I.2 Assumptions and main result

Assumption I.1. The operator L is defined on $L^1(V, \mu)$ by (I.1.3), with $\mathcal{M} \in L^1(V, \mu)$ such that $\inf_V \mathcal{M} > 0$ and $\int_V \mathcal{M} d\mu = 1$.

Let us define the spaces $L^2(\mathcal{M}^{-1})$ and L_x^2 and the associated inner products:

$$L^2(\mathcal{M}^{-1}) \doteq L^2(\mathbb{T}^d \times V, dx \mathcal{M}^{-1}(v) d\mu(v)), (f, g)_{L^2(\mathcal{M}^{-1})} \doteq \int_{\mathbb{T}^d} \int_V \frac{f(x, v) g(x, v)}{\mathcal{M}(v)} d\mu(v) dx,$$

$$L_x^2 \doteq L^2(\mathbb{T}^d, dx), (f, g)_{L_x^2} \doteq \int_{\mathbb{T}^d} f(x) g(x) dx.$$

We also define the norms $\|\cdot\|_{L^2(\mathcal{M}^{-1})}$ and $\|\cdot\|_{L_x^2}$ associated with these inner products.

Note that L is an orthogonal projection in $L^2(\mathcal{M}^{-1})$, hence

$$\forall f \in L^2(\mathcal{M}^{-1}), \|Lf\|_{L^2(\mathcal{M}^{-1})} \leq \|f\|_{L^2(\mathcal{M}^{-1})}.$$

Assumption I.2. The function a is bounded ($a \in L^\infty(V, \mu; \mathbb{R}^d)$), centered for $\mathcal{M} d\mu$, namely

$$\int_V a(v) \mathcal{M}(v) d\mu(v) = 0, \tag{I.2.1}$$

and the following matrix is symmetric and positive definite

$$K \doteq \int_V a(v) \otimes a(v) \mathcal{M}(v) d\mu(v) > 0. \tag{I.2.2}$$

The following assumption is not required to get the convergence in $\mathcal{C}^0([0, T], H_x^{-\sigma})$ but is used in Section I.8 to retrieve a stronger convergence (in $L^2([0, T], L_x^2)$). It is exactly the assumption of [BD99, Theorem 2.3].

Assumption I.3. We have $(V, d\mu) = (\mathbb{R}^n, \psi(v) dv)$ for some function $\psi \in H^1(\mathbb{R}^n)$, $a \in \text{Lip}_{loc}(\mathbb{R}^n; \mathbb{R}^d)$ and there exists $C \geq 0$ and $\sigma^* \in (0, 1]$ such that

$$\forall u \in S^{d-1}, \forall \lambda \in \mathbb{R}, \forall \delta > 0, \int_{\lambda < a(v) \cdot u < \lambda + \delta} (|\psi(v)|^2 + |\nabla \psi(v)|^2) dv \leq C \delta^{\sigma^*}.$$

Assumption I.4. We have

$$\sup_{\varepsilon \in (0,1)} \mathbb{E} \left[\|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^{24} \right] < \infty. \quad (\text{I.2.3})$$

and ρ_0^ε converges in distribution in L_x^2 to ρ_0 .

Remark I.2.1. The moments of order 24 in Assumption I.4 are useful in Sections I.4.1 and I.4.2.

I.2.1 Driving random term

Consider the normed space

$$E \doteq \mathcal{C}^{2[d/2]+4}(\mathbb{T}^d),$$

where the norm is given by

$$\|\cdot\|_E = \sum_{|\beta| \leq 2[d/2]+4} \sup_{x \in \mathbb{T}^d} \left| \frac{\partial^{|\beta|} \cdot}{\partial x^\beta} \right|,$$

where $\beta \in \mathbb{N}^d$, $|\beta| = \sum_{i=1}^d \beta_i$ and

$$\frac{\partial^{|\beta|}}{\partial x^\beta} = \frac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \dots \partial x_d^{\beta_d}}.$$

Assumption I.5. The family of process $(m(\cdot, n))_{n \in E}$ is a E -valued, càdlàg, stochastically continuous and homogeneous Markov process with initial condition $m(0, n) = n$. It admits a unique centered stationary distribution ν

$$\int_E \|n\|_E d\nu(n) < \infty \text{ and } \int_E n d\nu(n) = 0.$$

The driving process \overline{m} is the stationary Markov process associated with $(m(\cdot, n))_{n \in E}$, meaning that for all $t \in \mathbb{R}^+$, the distribution of $\overline{m}(t)$ is ν . It is adapted to a filtration $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ satisfying the usual conditions (complete and right-continuous).

For $\theta \in L^1(E) \doteq L^1(E, \nu)$, set

$$\langle \theta \rangle \doteq \int_E \theta d\nu.$$

Note that most of the arguments below only require $\overline{m}(t) \in \mathcal{C}^1(\mathbb{T}^d)$. However, in Section I.6, we use the compact embedding $H^{[d/2]+2}(\mathbb{T}^d) \subset \mathcal{C}^1(\mathbb{T}^d)$ and in Section I.3.2, we need $\overline{m}(t) \in \mathcal{C}^{2s}(\mathbb{T}^d)$ with $H^s(\mathbb{T}^d) \subset \mathcal{C}^1(\mathbb{T}^d)$, hence $s = [d/2] + 2$.

Definition I.2.1. For $\varepsilon > 0$, the random process \overline{m}^ε is defined by (I.1.4) where \overline{m} is defined by Assumption I.5. Let $\mathcal{F}_t^\varepsilon = \mathcal{F}_{t/\varepsilon^2}$ so that \overline{m}^ε is adapted to the filtration $(\mathcal{F}_t^\varepsilon)_{t \in \mathbb{R}^+}$.

Assumption on moments

From now on, we depart from the setting of [DV12]. In the previous works [DV12, DV20], it is assumed that there exists $C_* \in \mathbb{R}^+$ such that, almost surely,

$$\forall t \in \mathbb{R}^+, \|\overline{m}(t)\|_E \leq C_*.$$

The main novelty of this chapter is that we relax this assumption into Assumptions I.7 and I.6 concerning moments.

Assumption I.6. There exists $\gamma \in (4, \infty)$ such that

$$\mathbb{E} \left[\sup_{t \in [0,1]} \|\overline{m}(t)\|_E^\gamma \right] < \infty.$$

The condition $\gamma > 4$ is required below in Assumption I.7, where we also assume that the moments on $m(t, n)$ depend polynomially on n .

Assumption I.7. There exists $b \in [0, \frac{\gamma}{2} - 2)$ such that

$$\sup_{n \in E} \sup_{t \in \mathbb{R}^+} \frac{\mathbb{E} [\|m(t, n)\|_E^2]^{\frac{1}{2}}}{1 + \|n\|_E^b} < \infty,$$

and such that ν has a finite $8(b+2)$ -order moment, namely

$$\int_E \|n\|_E^{8(b+2)} d\nu(n) < \infty.$$

For instance, if m is an Ornstein-Uhlenbeck process

$$dm(t) = -\theta m(t)dt + \sigma dW(t),$$

with W a E -valued Wiener process, then m satisfies Assumptions I.6 and I.7.

Moreover, any process satisfying the boundedness assumption in [DV12] also satisfies Assumptions I.6 and I.7.

Mixing property

Assumption I.8 (Mixing property). There exists a nonnegative integrable function $\gamma_{mix} \in L^1(\mathbb{R}^+)$ such that, for all $n_1, n_2 \in E$, there exists a coupling $(m^*(\cdot, n_1), m^*(\cdot, n_2))$ of the processes $m(\cdot, n_1)$ and $m(\cdot, n_2)$ such that

$$\forall t \in \mathbb{R}^+, \mathbb{E} [\|m^*(t, n_1) - m^*(t, n_2)\|_E^2]^{1/2} \leq \gamma_{mix}(t) \|n_1 - n_2\|_E.$$

Typically, γ_{mix} is expected to be of the form $\gamma_{mix}(t) = C_{mix} e^{-\beta_{mix} t}$ for some $\beta_{mix} > 0$. In the example where m is an Ornstein-Uhlenbeck process, consider $m^*(\cdot, n_1)$ and $m^*(\cdot, n_2)$ driven by the same Wiener process W . Owing to Gronwall's Lemma, it is straightforward to prove that this coupling satisfies Assumption I.8 and that γ_{mix} decays exponentially fast.

We also need Assumptions I.9 and I.10 concerning the transition semi-group associated to the homogeneous Markov process $(m(\cdot, n))_{n \in E}$. Since those assumptions are quite technical, we postpone their statement in Section I.3.1.

I.2.2 Main result

For $x, y \in \mathbb{T}^d$, define the kernel

$$k(x, y) = \mathbb{E} \left[\int_{\mathbb{R}} \overline{m}(0)(x) \overline{m}(t)(y) dt \right], \quad (\text{I.2.4})$$

and for $f \in L_x^2$ and $x \in \mathbb{T}^d$, let us recall from [DV12]

$$Qf(x) = \int_{\mathbb{T}^d} k(x, y) f(y) dy. \quad (\text{I.2.5})$$

Theorem I.2.1. *Let W be a cylindrical Wiener process on L_x^2 , ρ_0 be a random variable in L_x^2 and ρ be the weak solution of the linear stochastic diffusion equation*

$$d\rho = \operatorname{div}(K \nabla \rho) dt + \rho Q^{1/2} \circ dW(t), \quad (\text{I.2.6})$$

with initial condition $\rho(0) = \rho_0$, in the sense of Definition I.5.1.

Let Assumptions I.1, I.2, I.4, I.5, I.6, I.7, I.8, I.9 and I.10 be satisfied. Also assume that $\rho^\varepsilon(0)$ converges in distribution to ρ_0 in L_x^2 . Then, for all $\sigma > 0$ and $T > 0$, the density ρ^ε converges in distribution in $\mathcal{C}^0([0, T], H_x^{-\sigma})$ to ρ .

If moreover Assumption I.3 is satisfied, then the convergence also holds in $L^2([0, T], L_x^2)$.

The noise in (I.2.6) involves a Stratonovich product, which is usual in the context of diffusion limit. Written with a Itô product, the limit becomes

$$d\rho = \operatorname{div}(K \nabla \rho) dt + \frac{1}{2} F \rho dt + \rho Q^{1/2} dW(t), \quad (\text{I.2.7})$$

where F is the trace of Q , namely

$$F(x) = k(x, x). \quad (\text{I.2.8})$$

This equation is well-posed, as discussed after Definition I.5.1.

This is the same limit than in [DV12]. Compared with [DV12], we obtain a stronger convergence result, namely a convergence in $L^2([0, T], L_x^2)$ under additional assumptions.

I.2.3 Strategy of the proof of Theorem I.2.1

A standard strategy to prove the convergence of ρ^ε when $\varepsilon \rightarrow 0$ (see [DV12, DV20, DRV20]) is first to establish the tightness of the family $(\rho^\varepsilon)_{\varepsilon>0}$, and then the uniqueness of the limit point of this family and solves (I.2.13). The tightness usually comes from estimates on moments of trajectories. It is the case in [DV12], where the boundedness of \overline{m} is used to get an estimate on $\mathbb{E} [\sup_{t \in [0, T]} \|f^\varepsilon(t)\|_{L^2}^p]$ for all $T > 0$ and $p \geq 1$. However, without an almost sure bound on \overline{m} , we do not manage to get this estimate. Instead, we introduce a stopping time τ_Λ^ε depending on a parameter Λ such that the estimate holds for $f^{\varepsilon, \tau_\Lambda^\varepsilon} \doteq f^\varepsilon(t \wedge \tau_\Lambda^\varepsilon)$.

More precisely, define a first stopping time

$$\tau^\varepsilon \doteq \inf \{t \in \mathbb{R}^+ \mid \|\overline{m}^\varepsilon(t)\|_E > \varepsilon^{-\alpha}\}, \quad (\text{I.2.9})$$

for some parameter α . Let $\mathcal{C}_x^1 \doteq \mathcal{C}^1(\mathbb{T}^d)$ and define the hitting time of a threshold Λ by $z \in \mathcal{C}^0([0, T], \mathcal{C}_x^1)$

$$\tau_\Lambda(z) \doteq \inf \left\{ t \in \mathbb{R}^+ \mid \|z(t)\|_{\mathcal{C}_x^1} \geq \Lambda \right\}. \quad (\text{I.2.10})$$

Then, define the auxiliary process

$$\zeta^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t \overline{m}^\varepsilon(s) ds = \varepsilon \int_0^{t/\varepsilon^2} \overline{m}(s) ds \in E \subset \mathcal{C}_x^1. \quad (\text{I.2.11})$$

Observe that $\frac{1}{\varepsilon} \overline{m}^\varepsilon = \partial_t \zeta^\varepsilon$.

We can now define

$$\tau_\Lambda^\varepsilon \doteq \tau^\varepsilon \wedge \tau_\Lambda(\zeta^\varepsilon). \quad (\text{I.2.12})$$

The times τ^ε and $\tau_\Lambda(\zeta^\varepsilon)$ have different asymptotic behaviors. On the one hand, Lemma I.3.4 states that $\tau^\varepsilon \rightarrow \infty$ in probability. On the other hand, Section I.7.1 establishes that ζ^ε converges in distribution, when $\varepsilon \rightarrow 0$, to a Wiener process ζ . Thus, we prove that, for all Λ outside of a countable set, $\tau_\Lambda(\zeta^\varepsilon)$ converges in distribution to $\tau_\Lambda(\zeta)$. Hence, τ_Λ^ε converges in distribution to $\tau_\Lambda(\zeta)$.

In Section I.3.5, we prove an estimate on $f^{\varepsilon, \tau_\Lambda^\varepsilon}$ depending only on T , Λ and f_0^ε . This estimate leads to prove the tightness of the family of stopped processes $(\rho^{\varepsilon, \tau_\Lambda^\varepsilon})_{\varepsilon > 0}$. Then, we identify the limit points of this family using the notions of martingale problems and perturbed test functions, and we deduce the convergence of the stopped process to a limit ρ_Λ .

Since $\tau^\varepsilon \rightarrow \tau_\Lambda(\zeta)$ and we expect $\rho^\varepsilon \rightarrow \rho$, it is convenient to study the process $(\rho^\varepsilon, \zeta^\varepsilon)$ to be able to write the limit of $\rho^{\varepsilon, \tau_\Lambda^\varepsilon}$ as $\rho^{\tau_\Lambda(\zeta)}$. Moreover, to prove that ρ^ε indeed converges to ρ and that ρ satisfies (I.2.6), we need $\tau_\Lambda(\zeta)$ to be a stopping time for the limit process. Thus, we need to consider the convergence in distribution of the couple $(\rho^\varepsilon, \zeta^\varepsilon)$ in $\mathcal{C}^0([0, T], H_x^{-\sigma}) \times \mathcal{C}^0([0, T], \mathcal{C}_x^1)$ to the solution (ρ, ζ) of

$$\begin{cases} d\rho = \operatorname{div}(K \nabla \rho) dt + \frac{1}{2} F \rho dt + \rho Q^{1/2} dW(t) \\ d\zeta = Q^{1/2} dW(t), \end{cases} \quad (\text{I.2.13})$$

with initial condition $\rho(0) = \rho_0$ and $\zeta(0) = 0$. In this framework, $\tau_\Lambda(\zeta^\varepsilon)$ is a stopping time for $(\rho^\varepsilon, \zeta^\varepsilon)$ and $\tau_\Lambda(\zeta)$ is a stopping time for the limit (ρ, ζ) .

We first state in Section I.3 some consequences of our assumptions in Section I.2.1 and introduce the stopping times. In Section I.4, we define the martingale problem solved by the process $(\rho^\varepsilon, \zeta^\varepsilon)$ and set up the perturbed test functions strategy.

In Section I.6, we prove the tightness of the stopped process in the space $\mathcal{C}^0([0, T], H_x^{-\sigma}) \times \mathcal{C}^0([0, T], \mathcal{C}_x^1)$, using the perturbed test functions of Section I.4. Then, in Section I.7, we establish the convergence of the martingale problems when $\varepsilon \rightarrow 0$ to identify the limit as a solution of a stopped martingale problem, and deduce the convergence of the original process $(\rho^\varepsilon, \zeta^\varepsilon)$ in $\mathcal{C}^0([0, T], H_x^{-\sigma}) \times \mathcal{C}^0([0, T], \mathcal{C}_x^1)$.

In Section I.8, we prove the tightness of the stopped process in $L^2([0, T], L_x^2)$ under the assumptions of Theorem I.2.1, using an averaging lemma. Combined with the previous results, we deduce the convergence of the original process ρ^ε in $L^2([0, T], L_x^2)$.

I.3 Preliminary results

I.3.1 Resolvent operator

Additional assumptions

Denote by $(P_t)_{t \in \mathbb{R}^+}$ the transition semi-group on E associated to the homogeneous Markov process \overline{m} and let B denote its infinitesimal generator

$$\forall n \in E, B\theta(n) = \lim_{t \rightarrow 0} \frac{P_t\theta(n) - \theta(n)}{t}.$$

The usual framework for Markov processes and their transition semi-groups is to consider continuous bounded test functions $\theta \in \mathcal{C}_b^0(E)$, so that $P_t\theta$ is a contraction semi-group (see [EK86]). Here, we need unbounded test functions (see Section I.4.2), thus consider the action of the semi-group on $\mathcal{C}^0(E) \cap L^1(E)$. We also consider the domain of B

$$D(B) \doteq \left\{ \theta \in \mathcal{C}^0(E) \cap L^1(E) \mid \forall n \in E, B\theta(n) \text{ exists and } B\theta \in \mathcal{C}^0(E) \cap L^1(E) \right\}.$$

We need a continuity property for the semi-group $(P_t)_{t \in \mathbb{R}^+}$. Define first the resolvent operator.

Definition I.3.1. For $\lambda \in [0, \infty)$ and $\theta \in \mathcal{C}^0(E) \cap L^1(E)$ such that $\int_0^\infty |P_t\theta(n)| dt < \infty$ for all $n \in E$, define the resolvent: for all $n \in E$

$$R_\lambda\theta(n) \doteq \int_0^\infty e^{-\lambda t} P_t\theta(n) dt.$$

Assumption I.9. The family $(P_t)_{t \in \mathbb{R}^+}$ is a semi-group on $\mathcal{C}^0(E) \cap L^1(E)$. Moreover, for all $(\lambda_i)_{1 \leq i \leq 4} \in [0, \infty)^4$ and $(\theta_i)_{1 \leq i \leq 4} \in (\mathcal{C}^0(E) \cap L^1(E))^4$ such that $R_{\lambda_i}\theta_i$ are well-defined by Definition I.3.1, we have

$$\forall j \in \llbracket 1, 4 \rrbracket, \Pi_{i=1}^j R_{\lambda_i}\theta_i \in D(B).$$

In addition, we assume that for $\lambda \in [0, \infty)$ and θ such that $R_\lambda\theta \in D(B)$, the commutation formula holds

$$B \int_0^\infty e^{-\lambda t} P_t\theta(\cdot) dt = \int_0^\infty e^{-\lambda t} B P_t\theta(\cdot) dt.$$

The first part of Assumption I.9 is satisfied for instance by an Ornstein-Uhlenbeck process. The second part of Assumption I.9 is satisfied under a continuity property for the semi-group $(P_t)_{t \in \mathbb{R}^+}$. Indeed, consider the following computations

$$\lim_{s \rightarrow 0} \frac{P_s - \text{id}}{s} \int_0^\infty e^{-\lambda t} P_t\theta(\cdot) dt = \lim_{s \rightarrow 0} \int_0^\infty e^{-\lambda t} \frac{P_{s+t} - P_t}{s} \theta(\cdot) dt = \int_0^\infty e^{-\lambda t} \lim_{s \rightarrow 0} \frac{P_{s+t} - P_t}{s} \theta(\cdot) dt$$

To justify the first equality, it is sufficient to assume point-wise continuity of P_t for all t on the space $\mathcal{C}^0(E) \cap L^1(E)$. The second equality is a consequence of the bounded convergence theorem.

Note that by means of Assumption I.9, $-R_0$ is the inverse of B . Indeed, for θ such that $R_0\theta \in D(B)$, we have

$$B \int_0^\infty P_t\theta(\cdot) dt = \int_0^\infty B P_t\theta(\cdot) dt = \int_0^\infty \partial_t P_t\theta(\cdot) dt = -\theta.$$

We sometimes use functions having at most polynomial growth. Our last assumption is that B preserves this property.

Assumption I.10. If $\theta \in D(B)$ has at most polynomial growth, then $B\theta$ has at most polynomial growth with the same degree. Namely, there exists $C_B \in (0, \infty)$ such that, for any $\theta \in D(B)$ and $k \in \mathbb{N}$,

$$\sup_{n \in E} \frac{|B\theta(n)|}{1 + \|n\|_E^k} \leq C_B \sup_{n \in E} \frac{|\theta(n)|}{1 + \|n\|_E^k}$$

Results on the resolvent operator

We introduce a class of pseudo-linear (respectively pseudo-quadratic) functions, which behave like linear (respectively quadratic) functions for our purposes.

Definition I.3.2. A function $\theta \in \text{Lip}(E)$ such that $\langle \theta \rangle = 0$, is called pseudo-linear. Denote by $[\theta]_{\text{Lip}}$ its Lipschitz constant.

A function $\theta : E \rightarrow \mathbb{R}$ is called pseudo-quadratic if there exists a function $b_\theta : E^2 \rightarrow \mathbb{R}$ satisfying

- for all $n \in E$, $\theta(n) = b_\theta(n, n)$,
- for all $n \in E$, $b_\theta(n, \cdot)$ and $b_\theta(\cdot, n)$ are Lipschitz continuous,
- the mappings $n \mapsto [b_\theta(n, \cdot)]_{\text{Lip}}$ and $n \mapsto [b_\theta(\cdot, n)]_{\text{Lip}}$ have at most linear growth.

If θ is a pseudo-quadratic function, then let

$$[\theta]_{\text{quad}} \doteq \sup_{n_1 \neq n_2 \in E} \frac{|\theta(n_2) - \theta(n_1)|}{(1 + \|n_1\|_E + \|n_2\|_E) \|n_2 - n_1\|_E} < \infty.$$

Let E^* denote the dual space of E . Any element $\theta \in E^*$ is pseudo-linear.

A consequence of the mixing property (Assumption I.8) is that the pseudo-linear and the pseudo-quadratic functions introduced in Definition I.3.2 satisfy the conditions of Definition I.3.1.

Lemma I.3.1. *Let θ be a pseudo-linear function. Then, for all $\lambda \geq 0$, $R_\lambda \theta$ is well-defined and is pseudo-linear. Moreover, let*

$$C_\lambda = \int_0^\infty e^{-\lambda t} \gamma_{\text{mix}}(t) dt \text{ and } C'_\lambda = \left(1 \vee \int \|n_2\|_E d\nu(n_2) \right) C_\lambda.$$

Then, we have

$$[R_\lambda \theta]_{\text{Lip}} \leq [\theta]_{\text{Lip}} C_\lambda,$$

and for $n \in E$,

$$|R_\lambda \theta(n)| \leq C'_\lambda [\theta]_{\text{Lip}} (1 + \|n\|_E). \quad (\text{I.3.1})$$

Let θ be a pseudo-quadratic function. Then, for $\lambda \geq 0$, $R_\lambda [\theta - \langle \theta \rangle]$ is well-defined. Moreover, there exists $C''_\lambda \in (0, \infty)$ depending only on C_λ and b such that, for $n \in E$,

$$|R_\lambda [\theta - \langle \theta \rangle](n)| \leq C''_\lambda [\theta]_{\text{quad}} (1 + \|n\|_E^{b+1})$$

where b is defined in Assumption I.7.

Proof. Let $n_1, n_2 \in E$ and denote by $(m^*(\cdot, n_1), m^*(\cdot, n_2))$ the coupling introduced in Assumption I.8. If θ is Lipschitz continuous, then for all $t \in \mathbb{R}^+$, Assumption I.8 leads to

$$|P_t\theta(n_1) - P_t\theta(n_2)| \leq [\theta]_{\text{Lip}} \gamma_{\text{mix}}(t) \|n_1 - n_2\|_E. \quad (\text{I.3.2})$$

Recall that P_t is ν -invariant, i.e. $\nu P_t = \nu$, hence we have

$$\begin{aligned} |P_t\theta(n_1) - \langle \theta \rangle| &= \left| \int_E (P_t\theta(n_1) - \theta(n_2)) d\nu(n_2) \right| \\ &= \left| \int_E (P_t\theta(n_1) - P_t\theta(n_2)) d\nu(n_2) \right| \\ &\leq \gamma_{\text{mix}}(t) [\theta]_{\text{Lip}} \int_E \|n_1 - n_2\|_E d\nu(n_2) \\ &\leq \left(1 \vee \int \|n_2\|_E d\nu(n_2) \right) \gamma_{\text{mix}}(t) [\theta]_{\text{Lip}} (1 + \|n_1\|_E). \end{aligned} \quad (\text{I.3.3})$$

Assume that θ is pseudo-linear. Since $\langle \theta \rangle = 0$, (I.3.3) implies that $R_\lambda\theta$ is well-defined for all $\lambda \in [0, \infty)$, and (I.3.2) implies that $R_\lambda\theta$ is $([\theta]_{\text{Lip}} C_\lambda)$ -Lipschitz continuous. Moreover, by means of Fubini's Theorem,

$$\begin{aligned} \int_E R_\lambda\theta(n) d\nu(n) &= \int_E \int_0^\infty e^{-\lambda t} P_t\theta(n) dt d\nu(n) \\ &= \int_0^\infty e^{-\lambda t} \int_E P_t\theta(n) d\nu(n) dt \\ &= \int_0^\infty e^{-\lambda t} \int_E \theta(n) d\nu(n) dt \\ &= 0. \end{aligned}$$

This concludes the proof that $R_\lambda\theta$ is pseudo-linear. Finally, (I.3.1) is a straightforward consequence of (I.3.3).

Now assume that θ is a pseudo-quadratic function: using Assumptions I.7 and I.8 and Cauchy-Schwarz inequality, we have for all $n_1, n_2 \in E$

$$|P_t\theta(n_1) - P_t\theta(n_2)| \leq C [\theta]_{\text{quad}} \left(1 + \|n_1\|_E^b + \|n_2\|_E^b \right) \gamma_{\text{mix}}(t) \|n_1 - n_2\|_E,$$

for some constant C depending on b . Since P_t is ν -invariant and ν has a finite moment of order $b+1$ by Assumption I.7, we get

$$|P_t\theta(n) - \langle \theta \rangle| \leq C' [\theta]_{\text{quad}} \gamma_{\text{mix}}(t) (1 + \|n\|_E^{b+1}),$$

for some constant C' depending on b . Integrating with respect to t gives the announced result. \square

Let us recall notation from [DV12]. For $\rho, \rho' \in L_x^2$, denote by $\psi_{\rho, \rho'} \in E^*$ the continuous linear form

$$\forall n \in E, \psi_{\rho, \rho'}(n) \doteq (\rho n, \rho')_{L_x^2}$$

The linear form $\psi_{\rho, \rho'}$ is pseudo-linear and

$$[\psi_{\rho, \rho'}]_{\text{Lip}} = \|\psi_{\rho, \rho'}\|_{E^*} \leq \|\rho\|_{L_x^2} \|\rho'\|_{L_x^2}.$$

Hence, by Lemma I.3.1, we have for $\rho, \rho' \in L_x^2$, $\lambda \geq 0$ and $n \in E$

$$|R_\lambda \psi_{\rho, \rho'}(n)| \leq C'_\lambda \|\rho\|_{L_x^2} \|\rho'\|_{L_x^2} (1 + \|n\|_E).$$

Thus, for all $n \in E$, $(\rho, \rho') \mapsto R_\lambda \psi_{\rho, \rho'}(n)$ is a continuous bilinear form on L_x^2 . By means of Riesz Representation Theorem, there exists a continuous linear map $\tilde{R}_\lambda(n) : L_x^2 \rightarrow L_x^2$ such that

$$\forall \rho, \rho' \in L_x^2, \forall n \in E, R_\lambda \psi_{\rho, \rho'}(n) = (\tilde{R}_\lambda(n) \rho, \rho')_{L_x^2}.$$

By a slight abuse of notation, denote $R_\lambda(n) = \tilde{R}_\lambda(n)$. For $\varphi \in \mathcal{C}^1(L_x^2)$, the linear mapping $D\varphi(\rho)$ can also be identified as an element of L_x^2 :

$$\forall \rho, h \in L_x^2, D\varphi(\rho)(h) = (h, D\varphi(\rho))_{L_x^2},$$

so that we can define $D\varphi(\rho)(R_\lambda(n)h)$ for $\rho, h \in L_x^2$.

Now consider ρ and ρ' in dual Sobolev spaces H_x^k and H_x^{-k} (for $k \in \mathbb{N}$ such that $E \subset \mathcal{C}_x^k$, namely $k \leq 2 \lfloor d/2 \rfloor + 2$). We also may define $R_\lambda(n) : H_x^k \rightarrow H_x^k$ in a compatible way.

I.3.2 Properties of the covariance operator

Recall that k , Q and F are defined by equations (I.2.4), (I.2.5) and (I.2.8).

Lemma I.3.2. *The kernel k is symmetric and in $L^\infty(\mathbb{T}^d \times \mathbb{T}^d)$. Moreover, Q is a bounded, self-adjoint, compact and non-negative operator on L_x^2 .*

Proof. Since \bar{m} is stationary, we have the identity

$$k(x, y) = \int_E \psi_x(n) R_0 \psi_y(n) d\nu(n) + \int_E R_0 \psi_x(n) \psi_y(n) d\nu(n), \quad (\text{I.3.4})$$

with $\psi_x(n) = n(x)$ for all $n \in E$ and $x \in \mathbb{T}^d$. The functions ψ_x and ψ_y are continuous linear forms, thus we have

$$\sup_{x, y \in \mathbb{T}^d, n \in E} \frac{|\psi_x R_0 \psi_y(n)|}{1 + \|n\|_E^2} < \infty.$$

Owing to Assumption I.6, $\int_E \|n\|_E^2 d\nu(n) < \infty$, thus k is well-defined, bounded and symmetric. It implies that Q is a bounded operator on L_x^2 and is self-adjoint and compact (see for instance [DS64, XI.6]).

The proof of non-negativity of Q is given in [DV12]. □

By means of Lemma I.3.2, the operator $Q^{1/2}$ can be defined ($L_x^2 \rightarrow L_x^2$). Note that Q is trace-class, that $Q^{1/2}$ is Hilbert-Schmidt on L_x^2 and that

$$\|Q^{1/2}\|_{\mathcal{L}_2}^2 = \text{Tr } Q = \int_{\mathbb{T}^d} F(x) dx.$$

Let $(F_i)_i$ be a orthonormal and complete system of eigenvectors of $Q^{1/2}$, and $(\sqrt{q_i})_i$ their associated eigenvalues.

Lemma I.3.3. *For all i , $F_i \in \mathcal{C}_x^1$ and*

$$\sum_i q_i \|F_i\|_{\mathcal{C}_x^1}^2 < \infty.$$

Proof. Let $s = \lfloor \frac{d}{2} \rfloor + 2$, so that we have the continuous embeddings $H_x^s \subset \mathcal{C}_x^1$. It is thus sufficient to prove that $\sum_i q_i \|F_i\|_{H_x^s}^2 < \infty$.

Since $\overline{m}(t) \in E = \mathcal{C}_x^{2s}$ and is mixing, is it straightforward to prove that $k \in H_x^{2s}$ using a differentiation under the integral sign in (I.2.4).

For $|\beta| \leq s$, we multiply the identity $q_i F_i = Q F_i$ by $\frac{\partial^{2|\beta|} F_i}{\partial x^{2\beta}}$ and integrate by parts both sides of the equality to get

$$(-1)^{|\beta|} q_i \left\| \frac{\partial^{|\beta|} F_i}{\partial x^\beta} \right\|_{L_x^2}^2 = \int \int \frac{\partial^{2|\beta|} k(x, y)}{\partial x^{2\beta}} F_i(x) F_i(y) dx dy.$$

Using (I.3.4), we have

$$\begin{aligned} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\partial^{2|\beta|} k(x, y)}{\partial x^{2\beta}} F_i(x) F_i(y) dx dy &= \int_E \int_{\mathbb{T}^d} \frac{\partial^{2|\beta|} n}{\partial x^{2\beta}}(x) F_i(x) dx \int_{\mathbb{T}^d} R_0 \psi_y(n) F_i(y) dy d\nu(n) \\ &\quad + \int_E \int_{\mathbb{T}^d} \frac{\partial^{2|\beta|} (R_0 \psi_x(n))}{\partial x^{2\beta}} F_i(x) dx \int_{\mathbb{T}^d} n(y) F_i(y) dy d\nu(n). \end{aligned}$$

We sum with respect to i , use the Parseval identity and the Cauchy-Schwarz inequality to get

$$\begin{aligned} \sum_i q_i \left\| \frac{\partial^{|\beta|} F_i}{\partial x^\beta} \right\|_{L_x^2}^2 &\leq 2 \int_E \|R_0 \psi_x(n)\|_{H_x^{2|\beta|}} \|n\|_{H_x^{2|\beta|}} d\nu(n) \\ &\leq C \int_E (1 + \|n\|_E^2) d\nu(n), \end{aligned}$$

for some constant C , using Lemma I.3.1. This upper bound is finite by Assumption I.7. Summing with respect to β concludes the proof. \square

I.3.3 Behavior of the stopping time for the driving process

Recall that τ^ε is defined by (I.2.9): $\tau^\varepsilon \doteq \inf \{t \in \mathbb{R}^+ \mid \|\overline{m}^\varepsilon(t)\|_E > \varepsilon^{-\alpha}\}$.

In this section, we establish Lemma I.3.4 below. Its objectives are twofold. On the one hand, it shows that \overline{m}^ε is almost surely bounded on any interval $[0, T]$, which is useful to justify the well-posedness of (I.1.1). On the other hand, it gives us an estimate for $\varepsilon^\alpha \|\overline{m}^{\varepsilon, \tau^\varepsilon}(t)\|_E$ uniform in t and ε for some small α . This estimate will prove useful for Sections I.4.1, I.6 and I.7. Therefore, it is a key result of this chapter.

Lemma I.3.4. *Let Assumptions I.5 and I.6 be satisfied and let $T > 0$. Then almost surely*

$$\sup_{t \in [0, T]} \|\overline{m}^\varepsilon(t)\|_E < \infty.$$

Moreover, let $\alpha > \frac{2}{\gamma}$ and define the $(\mathcal{F}_t^\varepsilon)_t$ -stopping time τ^ε by (I.2.9). Then, we have

$$\forall T > 0, \mathbb{P}(\tau^\varepsilon < T) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Remark I.3.1. Equation (I.2.9) implies that for $t \in \mathbb{R}^+$,

$$\|\overline{m}^{\varepsilon, \tau^\varepsilon}(t)\|_E = \|\overline{m}^\varepsilon(t \wedge \tau^\varepsilon)\|_E \leq \varepsilon^{-\alpha} \vee \|\overline{m}(0)\|_E. \quad (\text{I.3.5})$$

In particular, on the event $\{\|\overline{m}(0)\|_E > \varepsilon^{-\alpha}\}$, we have $\tau^\varepsilon = 0$ and $\overline{m}^{\varepsilon, \tau^\varepsilon}(t) = \overline{m}(0)$. Thus, one does not necessarily have the estimate $\|\overline{m}^{\varepsilon, \tau^\varepsilon}(t)\|_E \leq \varepsilon^{-\alpha}$.

Proof. Let $(S_k)_k$ be the identically distributed random variables defined by

$$\forall k \in \mathbb{N}_0, S_k \doteq \sup_{t \in [k, k+1]} \|\bar{m}(t)\|_E.$$

By means of Assumptions [I.5](#), [I.6](#) and [I.7](#), for all $k \in \mathbb{N}$, $\mathbb{E}[S_k^\gamma] = \mathbb{E}[S_0^\gamma] < \infty$. Thus, almost surely, $S_k < \infty$ for all $k \in \mathbb{N}_0$. This yields the first result:

$$\mathbb{P}\text{-a.s.}, \forall T > 0, \sup_{t \in [0, T]} \|\bar{m}^\varepsilon(t)\|_E \leq \sup_{k \leq T\varepsilon^{-2}+1} S_k < \infty.$$

Since $\alpha > \frac{2}{\gamma}$, there exists δ such that $\frac{\alpha}{2} > \delta > \frac{1}{\gamma}$. Then, the Markov inequality yields

$$\sum_{k \in \mathbb{N}} \mathbb{P}(S_k \geq k^\delta) \leq \sum_{k \in \mathbb{N}} \frac{\mathbb{E}[S_k^\gamma]}{k^{\delta\gamma}} = \mathbb{E}[S_0^\gamma] \sum_{k \in \mathbb{N}} \frac{1}{k^{\delta\gamma}} < \infty.$$

By means of the Borel-Cantelli Lemma, almost surely, there exists a random integer $k_0 \in \mathbb{N}$ such that

$$\mathbb{P}\text{-a.s.}, \forall k > k_0, S_k < k^\delta.$$

Define the random variable $Z \doteq \sup_{k \leq k_0} S_k$. Then Z is almost surely finite and

$$\mathbb{P}\text{-a.s.}, \forall t \in \mathbb{R}^+, \|\bar{m}(t)\|_E \leq S_{[t]} \leq Z + [t]^\delta \leq Z + t^\delta.$$

Finally, for $T > 0$, using that $\alpha > 2\delta$, we get

$$\mathbb{P}(\tau^\varepsilon < T) = \mathbb{P}\left(\sup_{t \in [0, T]} \|\bar{m}^\varepsilon(t)\|_E > \varepsilon^{-\alpha}\right) \leq \mathbb{P}(Z + (T\varepsilon^{-2})^\delta > \varepsilon^{-\alpha}) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

□

In the sequel, α will be required to satisfy the constraint

$$\alpha < \frac{1}{b+2}. \tag{I.3.6}$$

Combined with the condition $\alpha > \frac{2}{\gamma}$ appearing in Lemma [I.3.4](#), this motivates the condition $\gamma > 2(b+2)$ in Assumption [I.7](#).

I.3.4 Pathwise solutions

By means of Lemma [I.3.4](#), we are in position to prove the existence and uniqueness of pathwise solutions of [\(I.1.1\)](#) and [\(I.1.2\)](#) (namely solutions when ω is fixed).

Proposition I.3.5. *Let Assumptions [I.5](#) and [I.6](#) be satisfied. Let $T > 0$ and $\varepsilon > 0$. Then for any $f_0^\varepsilon \in L^2(\mathcal{M}^{-1})$, there exists, almost surely, a unique solution f^ε of [\(I.1.1\)](#) in $\mathcal{C}^0([0, T]; L^2(\mathcal{M}^{-1}))$, in the sense that*

$$\mathbb{P}\text{-a.s.}, \forall t \in [0, T], f^\varepsilon(t) = e^{-\frac{t}{\varepsilon}A} f_0^\varepsilon + \int_0^t e^{-\frac{t-s}{\varepsilon}A} \left(\frac{1}{\varepsilon^2} L f^\varepsilon(s) + \frac{1}{\varepsilon} f^\varepsilon(s) \bar{m}^\varepsilon(s) \right) ds$$

where A is the operator defined by

$$\begin{aligned} D(A) &\doteq \{f \in L^2(\mathcal{M}^{-1}) \mid (x, v) \mapsto a(v) \cdot \nabla_x f(x, v) \in L^2(\mathcal{M}^{-1})\} \\ Af(x, v) &\doteq a(v) \cdot \nabla_x f(x, v). \end{aligned}$$

Note that here ε is fixed. Thus, the proof is standard, based on a fixed-point theorem.

Proof. Let $\omega \in \Omega$ and $\varepsilon > 0$. For $f \in \mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1}))$, let

$$\Phi(f) = e^{-\frac{t}{\varepsilon}A} f_0^\varepsilon + \int_0^t e^{-\frac{t-s}{\varepsilon}A} \left(\frac{1}{\varepsilon^2} Lf(s) + \frac{1}{\varepsilon} f(s) \overline{m}^\varepsilon(s) \right) ds.$$

Owing to the Banach fixed-point theorem, it is sufficient to prove that Φ is a contraction for some Banach norm on $\mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1}))$. For $r \in [0, \infty)$, we consider the following Banach norm

$$\forall f \in \mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1})), \|f\|_r = \sup_{t \in [0, T]} e^{-rt} \|f(t)\|_{L^2(\mathcal{M}^{-1})}.$$

Since the semi-group associated to A is given by

$$\forall f \in L^2(\mathcal{M}^{-1}), \forall x \in \mathbb{T}^d, \forall v \in V, e^{tA} f(x, v) = f(x + ta(v), v),$$

for $f \in L^2(\mathcal{M}^{-1})$, we have for all $t \in \mathbb{R}$ and $f \in L^2(\mathcal{M}^{-1})$,

$$\|e^{tA} f\|_{L^2(\mathcal{M}^{-1})} = \|f\|_{L^2(\mathcal{M}^{-1})}.$$

Thus, for $t \in [0, T]$, and $f, g \in \mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1}))$, we get

$$\begin{aligned} \|\Phi(f)(t) - \Phi(g)(t)\|_{L^2(\mathcal{M}^{-1})} &\leq \frac{1}{\varepsilon^2} \int_0^t \|L(f - g)(s)\|_{L^2(\mathcal{M}^{-1})} ds \\ &\quad + \frac{1}{\varepsilon} \int_0^t \|(f - g)(s) \overline{m}^\varepsilon(s)\|_{L^2(\mathcal{M}^{-1})} ds. \end{aligned}$$

By Lemma I.3.4, since ω is fixed, \overline{m}^ε is bounded in E on $[0, T]$. Since $\|Lh\|_{L^2(\mathcal{M}^{-1})} \leq \|h\|_{L^2(\mathcal{M}^{-1})}$ for $h \in L^2(\mathcal{M}^{-1})$, we get

$$\|\Phi(f)(t) - \Phi(g)(t)\|_{L^2(\mathcal{M}^{-1})} \leq \left(\frac{1}{\varepsilon^2} + \frac{1}{\varepsilon} \sup_{t \in [0, T]} \|\overline{m}^\varepsilon(t)\|_E \right) \int_0^t e^{rs} \|f - g\|_r ds.$$

Hence, we have

$$e^{-rt} \|\Phi(f)(t) - \Phi(g)(t)\|_{L^2(\mathcal{M}^{-1})} \leq \left(\frac{1}{\varepsilon^2} + \frac{1}{\varepsilon} \sup_{t \in [0, T]} \|\overline{m}^\varepsilon(t)\|_E \right) \frac{1 - e^{-rt}}{r} \|f - g\|_r,$$

and

$$\|\Phi(f) - \Phi(g)\|_r \leq \frac{1}{r} \left(\frac{1}{\varepsilon^2} + \frac{1}{\varepsilon} \sup_{t \in [0, T]} \|\overline{m}^\varepsilon(t)\|_E \right) \|f - g\|_r.$$

By taking r large enough, we get that Φ is contracting, which concludes the proof. \square

I.3.5 Estimate in $L^2(\mathcal{M}^{-1})$

In this section, we obtain an upper bound on $\|f^{\varepsilon, \tau_\Lambda^\varepsilon}\|_{L^2(\mathcal{M}^{-1})}$. Note that in the case where the driving process \overline{m} is bounded, [DV12] establishes a similar upper bound without introducing a stopping time. Here, the unboundedness of the stopped process $\overline{m}^{\varepsilon, \tau_\Lambda^\varepsilon}$ requires more intricate arguments and an additional stopping time $\tau_\Lambda(\zeta^\varepsilon)$. One of these arguments is the introduction of a weight \mathcal{M}^ε that depends on ε .

Proposition I.3.6. *Assume that $f_0^\varepsilon \in L^2(\mathcal{M}^{-1})$. For $\Lambda > 0$ and $\varepsilon > 0$, define ζ^ε by (I.2.11) and $\tau_\Lambda(\zeta^\varepsilon)$ by (I.2.10).*

Then almost surely, for all $t \in [0, T]$ and $\varepsilon \in (0, (4\|a\|_{L^\infty}\Lambda)^{-1}]$,

$$\|f^{\varepsilon, \tau_\Lambda(\zeta^\varepsilon)}(t)\|_{L^2(\mathcal{M}^{-1})}^2 + \frac{1}{\varepsilon^2} \int_0^{t \wedge \tau_\Lambda(\zeta^\varepsilon)} \|Lf^{\varepsilon, \tau_\Lambda(\zeta^\varepsilon)}(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \leq C_\Lambda(T) \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2, \quad (\text{I.3.7})$$

for some $C_\Lambda(T) > 0$ depending only on Λ , $\|a\|_{L^\infty}$ and T .

Note that $\tau_\Lambda(\zeta^\varepsilon) > 0$ almost surely since $\zeta^\varepsilon(0) = 0$.

Remark I.3.2. The condition $\varepsilon \in (0, (4\|a\|_{L^\infty}\Lambda)^{-1}]$ only reads: we fix Λ , then take ε small enough ($\varepsilon \rightarrow 0$) depending on the fixed Λ . From now on, we always assume

$$\varepsilon \leq (4\|a\|_{L^\infty}\Lambda)^{-1} < 1.$$

In particular, we denote by \sup_ε the supremum with respect to $\varepsilon \in (0, (4\|a\|_{L^\infty}\Lambda)^{-1}]$.

In most of the chapter, we neglect the integral term of the left-hand side of (I.3.7) and we only use

$$\|f^{\varepsilon, \tau_\Lambda(\zeta^\varepsilon)}(t)\|_{L^2(\mathcal{M}^{-1})}^2 \leq C_\Lambda(T) \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2.$$

Equation (I.3.7) will prove useful in Section I.8.

Let us introduce some notation. For any variable u , $x \lesssim_u y$ means that there exists C such that $x \leq Cy$ where C depends only on u , a , \mathcal{M} , B , ν , γ , α , γ_{mix} , b and $\mathbb{P}_{f_0^\varepsilon}$ the distribution of f_0^ε . With this notation, (I.3.7) yields

$$\|f^{\varepsilon, \tau_\Lambda(\zeta^\varepsilon)}(t)\|_{L^2(\mathcal{M}^{-1})}^2 \lesssim_\Lambda \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2.$$

and

$$\frac{1}{\varepsilon^2} \int_0^{t \wedge \tau_\Lambda(\zeta^\varepsilon)} \|Lf^{\varepsilon, \tau_\Lambda(\zeta^\varepsilon)}(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \lesssim_\Lambda \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2.$$

Proof. Define the time-dependent weight

$$\mathcal{M}^\varepsilon(t, x, v) = e^{2\zeta^\varepsilon(t, x)} \mathcal{M}(v),$$

and the associate weighted norm

$$\|f\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})} \doteq \left(\int \int \frac{|f^\varepsilon(x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} dx d\mu(v) \right)^{\frac{1}{2}}.$$

We have, for $t \in \mathbb{R}^+$,

$$\begin{aligned} \frac{1}{2} \partial_t \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 &= \int \int \left(\frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} \left(-\frac{1}{\varepsilon} A f^\varepsilon + \frac{1}{\varepsilon^2} L f^\varepsilon + \frac{1}{\varepsilon} f^\varepsilon \overline{m}^\varepsilon \right) (t, x, v) \right. \\ &\quad \left. - \frac{|f^\varepsilon(t, x, v)|^2}{2 |\mathcal{M}^\varepsilon(t, x, v)|^2} \partial_t \mathcal{M}^\varepsilon(t, x, v) \right) dx d\mu(v) \\ &= \mathcal{A}_\varepsilon + \mathcal{B}_\varepsilon \end{aligned}$$

with

$$\begin{aligned} \mathcal{A}_\varepsilon &= \frac{1}{\varepsilon^2} \int \int \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} \left(L f^\varepsilon + \varepsilon f^\varepsilon \overline{m}^\varepsilon - \frac{\varepsilon^2}{2} \frac{f^\varepsilon}{\mathcal{M}^\varepsilon} \partial_t \mathcal{M}^\varepsilon \right) (t, x, v) dx d\mu(v) \\ \mathcal{B}_\varepsilon &= -\frac{1}{\varepsilon} \int \int \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} A f^\varepsilon(t, x, v) dx d\mu(v). \end{aligned}$$

On the one hand, the weight \mathcal{M}^ε has been chosen in order to satisfy $\varepsilon \overline{m}^\varepsilon - \frac{\varepsilon^2}{2} \frac{\partial_t \mathcal{M}^\varepsilon}{\mathcal{M}^\varepsilon} = 0$. Moreover, since $f^\varepsilon = \rho^\varepsilon \mathcal{M} - L f^\varepsilon$ and $\int_V L f^\varepsilon d\mu = 0$, we get

$$\begin{aligned} \mathcal{A}_\varepsilon &= \frac{1}{\varepsilon^2} \int \int \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} L f^\varepsilon(t, x, v) dx d\mu(v) \\ &= \frac{1}{\varepsilon^2} \int_{\mathbb{T}^d} e^{-2\zeta^\varepsilon(t, x)} \rho^\varepsilon(t, x) \int_V L f^\varepsilon(t, x, v) d\mu(v) dx \\ &\quad - \frac{1}{\varepsilon^2} \int \int \frac{|L f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} d\mu(v) dx \\ &= -\frac{1}{\varepsilon^2} \int \int \frac{|L f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} d\mu(v) dx = -\frac{1}{\varepsilon^2} \|L f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2. \end{aligned}$$

On the other hand, by means of an integration by parts (we take a primitive of $f^\varepsilon \partial_{x_i} f^\varepsilon$ and a derivative of $\frac{1}{\mathcal{M}^\varepsilon}$), we write

$$\begin{aligned} \mathcal{B}_\varepsilon &= -\frac{1}{\varepsilon} \int \int a(v) \cdot \frac{f^\varepsilon(t, x, v) \nabla_x f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} dx d\mu(v) \\ &= -\frac{1}{\varepsilon} \int \int a(v) \cdot \frac{\frac{1}{2} |f^\varepsilon(t, x, v)|^2 \nabla_x \mathcal{M}^\varepsilon(t, x, v)}{|\mathcal{M}^\varepsilon(t, x, v)|^2} dx d\mu(v) \\ &= -\frac{1}{2\varepsilon} \int \int \left| \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} \right|^2 A \mathcal{M}^\varepsilon(t, x, v) dx d\mu(v). \end{aligned}$$

Then, by definition of \mathcal{M}^ε and A , we have

$$\mathcal{B}_\varepsilon = -\frac{1}{\varepsilon} \int_{\mathbb{T}^d} \nabla_x \zeta^\varepsilon(t, x) \cdot \int_V \frac{|f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} a(v) d\mu(v) dx.$$

Using once again the identity $f^\varepsilon = \rho^\varepsilon \mathcal{M} - L f^\varepsilon$, we get

$$\begin{aligned} \mathcal{B}_\varepsilon &= -\frac{1}{\varepsilon} \int_{\mathbb{T}^d} e^{-2\zeta^\varepsilon(t, x)} |\rho^\varepsilon(t, x)|^2 \nabla_x \zeta^\varepsilon(t, x) \cdot \int_V a(v) \mathcal{M}(v) d\mu(v) dx \\ &\quad - \frac{1}{\varepsilon} \int_{\mathbb{T}^d} \nabla_x \zeta^\varepsilon(t, x) \cdot \int_V \frac{|L f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} a(v) d\mu(v) dx \\ &\quad + \frac{2}{\varepsilon} \int \int e^{-2\zeta^\varepsilon(t, x)} \nabla_x \zeta^\varepsilon(t, x) \cdot a(v) \rho^\varepsilon(t, x) L f^\varepsilon(t, x, v) d\mu(v) dx \\ &= \mathcal{B}_\varepsilon^1 + \mathcal{B}_\varepsilon^2 + \mathcal{B}_\varepsilon^3. \end{aligned}$$

- Since a is centered for $\mathcal{M}\mu$, $\mathcal{B}_\varepsilon^1 = 0$.
- For $t \leq \tau_\Lambda(\zeta^\varepsilon)$, we have $\|\zeta^\varepsilon(t)\|_{C_x^1} \leq \Lambda$ and we assumed $\Lambda \leq (4\|a\|_{L^\infty} \varepsilon)^{-1}$. Thus, we get

$$\forall t \leq \tau_\Lambda(\zeta^\varepsilon), |\mathcal{B}_\varepsilon^2| \leq \frac{1}{4\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2.$$

- Using the Young inequality, we have

$$|\mathcal{B}_\varepsilon^3| \leq \frac{1}{4\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 + 4\|a\|_{L^2(\mathcal{M})}^2 \|\nabla_x \zeta^\varepsilon(t)\|_{C^0(\mathbb{T}^d)}^2 \int_{\mathbb{T}^d} e^{-2\zeta^\varepsilon(t,x)} |\rho^\varepsilon(t,x)|^2 dx,$$

with $\|a\|_{L^2(\mathcal{M})}^2 \doteq \int_V |a(v)|^2 \mathcal{M}(v) d\mu(v)$. Now using the Cauchy-Schwarz inequality and the identity $\int_V \mathcal{M}(v) d\mu(v) = 1$, we have

$$|\rho^\varepsilon(t,x)|^2 \leq \int_V \frac{|f^\varepsilon(t,x,v)|^2}{\mathcal{M}^\varepsilon(t,x,v)} d\mu(v) \int_V \mathcal{M}^\varepsilon(t,x,v) d\mu(v) = \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 e^{2\zeta^\varepsilon(t,x)},$$

hence

$$|\mathcal{B}_\varepsilon^3| \leq \frac{1}{4\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 + 4\|a\|_{L^2(\mathcal{M})}^2 \|\nabla_x \zeta^\varepsilon(t)\|_{C^0(\mathbb{T}^d)}^2 \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2.$$

We finally get, for $t \leq \tau_\Lambda(\zeta^\varepsilon)$,

$$\begin{aligned} \partial_t \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 &\leq -\frac{1}{2\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 \\ &\quad + 4\|a\|_{L^2(\mathcal{M})}^2 \|\nabla_x \zeta^\varepsilon(t)\|_{C^0(\mathbb{T}^d)}^2 \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2. \end{aligned}$$

For $t \leq \tau_\Lambda(\zeta^\varepsilon)$, Gronwall's Lemma implies

$$\|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 + \int_0^t \frac{1}{2\varepsilon^2} \|Lf^\varepsilon(s)\|_{L^2(\mathcal{M}^\varepsilon(s)^{-1})}^2 ds \leq \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2 e^{4\|a\|_{L^2(\mathcal{M})}^2 \int_0^t \|\nabla_x \zeta^\varepsilon(s)\|_{C^0(\mathbb{T}^d)}^2 ds}.$$

Since, for $t \in \mathbb{R}^+$, we have

$$\|\cdot\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 \geq \|\cdot\|_{L^2(\mathcal{M}^{-1})}^2 e^{-2\|\zeta^\varepsilon(t)\|_{C^0(\mathbb{T}^d)}},$$

we get, for $t \leq \tau_\Lambda(\zeta^\varepsilon)$,

$$\begin{aligned} \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^{-1})}^2 &+ \int_0^t \frac{1}{2\varepsilon^2} \|Lf^\varepsilon(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \\ &\leq \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2 \exp \left(2 \sup_{s \in [0,t]} \|\zeta^\varepsilon(s)\|_{C^0(\mathbb{T}^d)} + 4\|a\|_{L^2(\mathcal{M})}^2 \int_0^t \|\nabla_x \zeta^\varepsilon(s)\|_{C^0(\mathbb{T}^d)}^2 ds \right). \end{aligned}$$

To conclude, it is sufficient to recall that for $t \leq \tau_\Lambda(\zeta^\varepsilon)$, we have $\|\zeta^\varepsilon(t)\|_{C_x^1} \leq \Lambda$. \square

I.4 Martingale problems and perturbed test functions

The proof of Theorem I.2.1 heavily relies on the notion of martingale problems as introduced in [SV06]. To identify a limit point of $(\mathbb{P}_{\rho^\varepsilon})_{\varepsilon>0}$, we characterize it by a family of martingales and take the limit when $\varepsilon \rightarrow 0$ in their martingale properties.

The characterization of the distribution of a solution of a SPDE in terms of martingales is based on the Markov property satisfied by this solution. However, we expect a limit point ρ_Λ of the stopped process $\rho^{\varepsilon, \tau_\Lambda^\varepsilon}$ to be stopped at some $\tau_\Lambda(\zeta)$, as mentioned in Section I.3.5. Since $\tau_\Lambda(\zeta)$ is not a stopping time for the filtration generated by ρ_Λ , this latter process should not be Markov. Thus, we need to consider the convergence of the couple $(\rho^\varepsilon, \zeta^\varepsilon)$ instead of just ρ^ε . We will see more precisely in Section I.7 at which point this matter occurs.

I.4.1 Generator and martingales

Also note that $(f^\varepsilon, \zeta^\varepsilon)$ is not a Markov process. As in [DV12], we consider the coupled process with \bar{m}^ε and thus consider the $L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E$ -valued Markov process $X^\varepsilon \doteq (f^\varepsilon, \zeta^\varepsilon, \bar{m}^\varepsilon)$.

Denote by \mathcal{L}^ε the infinitesimal generator of X^ε . Since f^ε is solution of (I.1.1) and since $\partial_t \zeta^\varepsilon = \frac{1}{\varepsilon} \bar{m}^\varepsilon$, the infinitesimal generator has an expression of the type

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon} \mathcal{L}_1 + \frac{1}{\varepsilon^2} \mathcal{L}_2 \quad (\text{I.4.1})$$

with

$$\begin{aligned} \mathcal{L}_1 \varphi(f, z, n) &= D_f \varphi(f, z, n)(-Af + nf) + D_z \varphi(f, z, n)(n) \\ \mathcal{L}_2 \varphi(f, z, n) &= D_f \varphi(f, z, n)(Lf) + B \varphi(f, z, n), \end{aligned}$$

where B is the infinitesimal generator of \bar{m} . The domain of this generator contains the class of good test functions defined below. The terminology of "good test function" is inherited from [DV12], although our definition is a little more restrictive.

Definition I.4.1. A function $\varphi : L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E \rightarrow \mathbb{R}$ is called a good test function if

- It is continuously differentiable on $L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E$ with respect to the first and second variables.
- For $\ell \in \{1, 2\}$, $B(\varphi(f, z, \cdot)^\ell)$ is defined for all $(f, z) \in L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1$, and

$$B(\varphi^\ell) : L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E \rightarrow \mathbb{R}$$

is continuous.

- If we identify the differential D_f with the gradient, then for $f \in L^2(\mathcal{M}^{-1})$, $z \in \mathcal{C}_x^1$ and $n \in E$, we have

$$D_f \varphi(f, z, n) \in H^1(\mathbb{T}^d \times V, dx \mathcal{M}^{-1}(v) d\mu(v)). \quad (\text{I.4.2})$$

- The functions φ , $D_z\varphi$, $D_f\varphi$ and $AD_f\varphi$ have at most polynomial growth in the following sense: there exists $C_\varphi > 0$ such that for $f, h \in L^2(\mathcal{M}^{-1})$, $z \in \mathcal{C}_x^1$ and $n_1, n_2 \in E$, we have

$$\begin{aligned}
|\varphi(f, z, n_1)| &\leq C_\varphi (1 + S_1^3) (1 + S_2^{b+2}) \\
|D_f\varphi(f, z, n_1)(Ah)| &\leq C_\varphi (1 + S_1^3) (1 + S_2^{b+2}) \\
|D_f\varphi(f, z, n_1)(n_2h)| &\leq C_\varphi (1 + S_1^3) (1 + S_2^{b+2}) \\
|D_f\varphi(f, z, n_1)(Lh)| &\leq C_\varphi (1 + S_1^3) (1 + S_2^{b+2}) \\
|D_z\varphi(f, z, n_1)(n_2)| &\leq C_\varphi (1 + S_1^3) (1 + S_2^{b+2}),
\end{aligned} \tag{I.4.3}$$

where $S_1 = \|f\|_{L^2(\mathcal{M}^{-1})} \vee \|h\|_{L^2(\mathcal{M}^{-1})}$ and $S_2 = \|n_1\|_E \vee \|n_2\|_E$.

See Section I.4.2 for a justification of the need to consider growth as appearing in (I.4.3).

A consequence of (I.4.2) is that $AD_f\varphi$ is well-defined. Thus, for $f, h \in L^2(\mathcal{M}^{-1})$, $z \in \mathcal{C}_x^1$ and $n \in E$, we can define

$$D_f\varphi(f, z, n)(Ah) \doteq -(AD_f\varphi(f, z, n), h)_{L^2(\mathcal{M}^{-1})},$$

even though Ah is not necessarily defined in $L^2(\mathcal{M}^{-1})$.

The class of test-function introduced in Definition I.4.1 is chosen such that the Proposition I.4.1 holds.

Proposition I.4.1. *Let φ be a good test function in the sense of Definition I.4.1. Define for all $t \geq 0$*

$$M_\varphi^\varepsilon(t) \doteq \varphi(X^\varepsilon(t)) - \varphi(X^\varepsilon(0)) - \int_0^t \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds, \tag{I.4.4}$$

and consider the stopping time τ_Λ^ε defined by (I.2.12).

Then $M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}$ is a càdlàg $(\mathcal{F}_t^\varepsilon)_t$ -martingale and

$$\begin{aligned}
\forall t \in \mathbb{R}^+, \mathbb{E} \left[\left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \right|^2 \right] &= \mathbb{E} \left[\int_0^{t \wedge \tau_\Lambda^\varepsilon} (\mathcal{L}^\varepsilon(\varphi^2) - 2\varphi \mathcal{L}^\varepsilon \varphi)(X^\varepsilon(s)) ds \right] \\
&= \frac{1}{\varepsilon^2} \mathbb{E} \left[\int_0^{t \wedge \tau_\Lambda^\varepsilon} (B(\varphi^2) - 2\varphi B\varphi)(X^\varepsilon(s)) ds \right].
\end{aligned}$$

This result is expected to hold as in the standard framework [DV20]. However, due to the presence of stopping times, the proof is very technical.

Proof. Note that in this section, ε is fixed, it is therefore not required to prove bounds which are uniform with respect to ε .

Let φ be a good test function. Observe that φ and φ^2 are in the domain of \mathcal{L}^ε , by means of Definition I.4.1.

Let $s, t \in \mathbb{R}^+$, $\delta > 0$ and let $s = t_1 < \dots < t_n = t$ be a subdivision of $[s, t]$ such that $\max_i |t_{i+1} - t_i| = \delta$. Let g be a $\mathcal{F}_s^\varepsilon$ -measurable and bounded function. To simplify notation, let

$$f_i \doteq f^{\varepsilon, \tau_\Lambda^\varepsilon}(t_i), \zeta_i \doteq \zeta^{\varepsilon, \tau_\Lambda^\varepsilon}(t_i), m_i \doteq \overline{m}^{\varepsilon, \tau_\Lambda^\varepsilon}(t_i).$$

Then, we have

$$\begin{aligned}\mathbb{E} \left[\left(M_{\varphi}^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(t) - M_{\varphi}^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(s) \right) g \right] &= \mathbb{E} \left[\left(\varphi(X^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(t)) - \varphi(X^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(s)) - \int_{s \wedge \tau_{\Lambda}^{\varepsilon}}^{t \wedge \tau_{\Lambda}^{\varepsilon}} \mathcal{L}^{\varepsilon} \varphi(X^{\varepsilon}(u)) du \right) g \right] \\ &= r_f + r_z + r_n,\end{aligned}$$

where

$$\begin{aligned}r_f &= \sum_{i=1}^{n-1} \mathbb{E} \left[\left(\varphi(f_{i+1}, \zeta_{i+1}, m_{i+1}) - \varphi(f_i, \zeta_{i+1}, m_{i+1}) \right. \right. \\ &\quad \left. \left. - \int_{t_i \wedge \tau_{\Lambda}^{\varepsilon}}^{t_{i+1} \wedge \tau_{\Lambda}^{\varepsilon}} D_f \varphi(X^{\varepsilon}(u)) \left(-\frac{1}{\varepsilon} A f^{\varepsilon}(u) + \frac{1}{\varepsilon^2} L f^{\varepsilon}(u) + \frac{1}{\varepsilon} f^{\varepsilon}(u) \overline{m}^{\varepsilon}(u) \right) du \right) g \right], \\ r_z &= \sum_{i=1}^{n-1} \mathbb{E} \left[\left(\varphi(f_i, \zeta_{i+1}, m_{i+1}) - \varphi(f_i, \zeta_i, m_{i+1}) - \int_{t_i \wedge \tau_{\Lambda}^{\varepsilon}}^{t_{i+1} \wedge \tau_{\Lambda}^{\varepsilon}} D_z \varphi(X^{\varepsilon}(u)) \left(\frac{1}{\varepsilon} \overline{m}^{\varepsilon}(u) \right) du \right) g \right],\end{aligned}$$

and

$$r_n = \sum_{i=1}^{n-1} \mathbb{E} \left[\left(\varphi(f_i, \zeta_i, m_{i+1}) - \varphi(f_i, \zeta_i, m_i) - \int_{t_i \wedge \tau_{\Lambda}^{\varepsilon}}^{t_{i+1} \wedge \tau_{\Lambda}^{\varepsilon}} \frac{1}{\varepsilon^2} B \varphi(X^{\varepsilon}(u)) du \right) g \right].$$

Straightforward computations lead to

$$r_f = \mathbb{E} \left[\left(\int_s^t r'_f(u) du \right) g \right], r_z = \mathbb{E} \left[\left(\int_s^t r'_z(u) du \right) g \right]$$

with

$$\begin{aligned}r'_f(u) &= \sum_{i=1}^{n-1} \mathbb{1}_{[t_i \wedge \tau_{\Lambda}^{\varepsilon}, t_{i+1} \wedge \tau_{\Lambda}^{\varepsilon}]}(u) \left[D_f \varphi(f^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u), \zeta_{i+1}, m_{i+1}) - D_f \varphi(X^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u)) \right] (\partial_t f^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u)), \\ r'_z(u) &= \sum_{i=1}^{n-1} \mathbb{1}_{[t_i \wedge \tau_{\Lambda}^{\varepsilon}, t_{i+1} \wedge \tau_{\Lambda}^{\varepsilon}]}(u) \left[D_z \varphi(f_i, \zeta^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u), m_{i+1}) - D_z \varphi(X^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u)) \right] (\partial_t \zeta^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u)).\end{aligned}$$

Let us now check that $r_n = \mathbb{E} \left[\left(\int_s^t r'_n(u) du \right) g \right]$ with

$$r'_n(u) = \frac{1}{\varepsilon^2} \sum_{i=1}^{n-1} \mathbb{1}_{\{[t_i \wedge \tau_{\Lambda}^{\varepsilon}, t_{i+1} \wedge \tau_{\Lambda}^{\varepsilon}]\}}(u) \left[B \varphi(f_i, \zeta_i, \overline{m}^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u)) - B \varphi(X^{\varepsilon, \tau_{\Lambda}^{\varepsilon}}(u)) \right].$$

For $\theta \in \mathcal{C}^0(E) \cap L^1(E)$, the Markov property for \overline{m} yields

$$\mathbb{E} [\theta(\overline{m}(t)) \mid \mathcal{F}_s] = P_{t-s} \theta(\overline{m}(s)).$$

Usually, this property is written for θ deterministic, continuous and bounded, but it is straightforward to check that it is still satisfied when $\theta \in \mathcal{C}^0(E) \cap L^1(E)$ \mathcal{F}_s -measurable. The standard proof to show that \overline{m} solves the martingale problem associated to B (see for example [DV20, Theorem B.3]) can be applied, and we get that, for $\theta \in D(B)$,

$$t \mapsto \theta(\overline{m}(t)) - \theta(\overline{m}(0)) - \int_0^t B \theta(\overline{m}(u)) du$$

is an integrable \mathcal{F}_t -martingale. By rescaling the time to retrieve \bar{m}^ε , stopping the martingale at τ_Λ^ε and using a conditioning argument (g, f_i and ζ_i are \mathcal{F}_{t_i} -measurable), we get

$$\mathbb{E}[(\varphi(f_i, \zeta_i, m_{i+1}) - \varphi(f_i, \zeta_i, m_i))g] = \mathbb{E}\left[g \int_{t_i \wedge \tau_\Lambda^\varepsilon}^{t_{i+1} \wedge \tau_\Lambda^\varepsilon} \frac{1}{\varepsilon^2} B\varphi(X^\varepsilon(u))du\right].$$

Hence, we can write $r_n = \mathbb{E}\left[\left(\int_s^t r'_n(u)du\right)g\right]$ as claimed above.

Since the estimates given by (I.3.5) and Proposition I.3.6 are uniform for $t \in [0, T]$, we can use (I.4.3) with $S_1 \lesssim_\Lambda \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}$ and $S_2 \leq \varepsilon^{-\alpha} \vee \|\bar{m}(0)\|_E$. This leads to

$$\sup_{u \in [0, T]} |r'_*(u)|^2 \lesssim_{\varphi, \Lambda, \varepsilon} \left(1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^6\right) \left(1 + \|\bar{m}(0)\|_E^{2(b+2)}\right), \quad (\text{I.4.5})$$

where $r'_* \in \{r'_f, r'_z, r'_n\}$. Hence, the Cauchy-Schwarz inequality, Assumptions I.4 and I.7 yield

$$\mathbb{E}\left[\sup_{u \in [0, T]} |r'_*(u)|^2\right] \lesssim_{\varphi, \Lambda, \varepsilon} \mathbb{E}\left[\left(1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^{12}\right)\right]^{\frac{1}{2}} \mathbb{E}\left[\left(1 + \|\bar{m}(0)\|_E^{4(b+2)}\right)\right]^{\frac{1}{2}} < \infty. \quad (\text{I.4.6})$$

Thus, the terms r'_* are uniformly integrable with respect to (u, ω) . Recall that $f^{\varepsilon, \tau_\Lambda^\varepsilon}$ and $\zeta^{\varepsilon, \tau_\Lambda^\varepsilon}$ are almost surely continuous and that $\bar{m}^{\varepsilon, \tau_\Lambda^\varepsilon}$ is stochastically continuous. Then, the terms r'_* converge to 0 in probability when $\delta \rightarrow 0$. By uniform integrability, the terms r_* converge to 0, which proves that $M_{\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}}^\varepsilon$ is a $(\mathcal{F}_t^\varepsilon)_t$ -martingale. Note that we only used moments of order 12 and $4(b+2)$, instead of 24 and $8(b+2)$ as assumed in Assumptions I.4 and I.7. Hence, this proof can be adapted to establish that $M_{\varphi^2}^{\varepsilon, \tau_\Lambda^\varepsilon}$ is also a $(\mathcal{F}_t^\varepsilon)_t$ -martingale.

It remains to prove the formulas for the variance. This is done in several steps, following [DV20, Appendix B]. Since φ and φ^2 belong to the domain of \mathcal{L}^ε , the process

$$N^\varepsilon(t) = \int_0^t (\mathcal{L}^\varepsilon(\varphi^2) - 2\varphi\mathcal{L}^\varepsilon\varphi)(X^\varepsilon(s))ds = \frac{1}{\varepsilon^2} \int_0^t (B(\varphi^2) - 2\varphi B\varphi)(X^\varepsilon(s))ds,$$

is well-defined.

The proof of the second equality is straightforward: since $D = \mathcal{L}^\varepsilon - \frac{1}{\varepsilon^2}B$ is a first order differential operator, we have $D(\varphi^2) - 2\varphi D\varphi = 0$.

Let $0 = t_0 < t_1 < \dots < t_n = T$ be a subdivision of $[0, T]$ of step $\max |t_{i+1} - t_i| = \delta$.

Step 1: We claim that the following convergence is satisfied in $\mathbb{L}^2 \doteq \mathbb{L}^2(\Omega)$

$$N^{\varepsilon, \tau_\Lambda^\varepsilon}(t) = \lim_{\delta \rightarrow 0} \sum_i \mathbb{E}[N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i) \mid \mathcal{F}_{t_i}^\varepsilon]. \quad (\text{I.4.7})$$

Let $\Delta_i \doteq N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i) - \mathbb{E}[N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i) \mid \mathcal{F}_{t_i}^\varepsilon]$ so that (I.4.7) is equivalent to $\sum_i \Delta_i \xrightarrow{\delta \rightarrow 0} 0$ in \mathbb{L}^2 . Note that $\mathbb{E}[\Delta_i \Delta_j] = 0$ for $i \neq j$. Hence, we have $\mathbb{E}[\left|\sum_i \Delta_i\right|^2] = \mathbb{E}[\sum_i |\Delta_i|^2]$. Using that a conditional expectation is an orthogonal projection in \mathbb{L}^2 , we get

$$\begin{aligned} \mathbb{E}[|\Delta_i|^2] &\leq \mathbb{E}\left[\left|N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - N^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i)\right|^2\right] \\ &\lesssim_\varepsilon \mathbb{E}\left[\left|\int_{t \wedge t_i \wedge \tau_\Lambda^\varepsilon}^{t \wedge t_{i+1} \wedge \tau_\Lambda^\varepsilon} (B(\varphi^2) - 2\varphi B\varphi)(X^\varepsilon(s))ds\right|^2\right], \end{aligned}$$

By means of Assumption I.10 and Definition I.4.1,

$$|(B(\varphi^2) - 2\varphi B\varphi)(X^\varepsilon(s))|^2 \lesssim_\varphi \left(1 + \|f^\varepsilon(s)\|_{L^2(\mathcal{M}^{-1})}^2\right) \left(1 + \|\overline{m}^\varepsilon(s)\|_E^{4(b+2)}\right).$$

As in (I.4.5) and (I.4.6) (using moments of order 24 and $8(b+2)$ instead of 12 and $4(b+2)$), Proposition I.3.6 and (I.3.5) lead to

$$\mathbb{E} \left[\sup_{s \in [0, T]} |(B(\varphi^2) - 2\varphi B\varphi)(X^{\varepsilon, \tau_\Lambda^\varepsilon}(s))|^2 \right] \lesssim_{\varphi, \Lambda, \varepsilon} 1.$$

Since $t \wedge t_i \wedge \tau_\Lambda^\varepsilon - t \wedge t_{i+1} \wedge \tau_\Lambda^\varepsilon \leq t_{i+1} - t_i$, we get

$$\mathbb{E} [|\Delta_i|^2] \lesssim_{\varphi, \Lambda, \varepsilon} (t_{i+1} - t_i)^2,$$

which then yields $\mathbb{E} [|\sum_i \Delta_i|^2] \lesssim_{\varphi, \Lambda, \varepsilon} T\delta \xrightarrow{\delta \rightarrow 0} 0$, which proves (I.4.7).

Step 2: We claim that

$$\mathbb{E} \left[\sum_i |R_{t_i, t_{i+1}}| \right] \lesssim_{\varphi, \Lambda, \varepsilon} \delta^{1/2}, \quad (\text{I.4.8})$$

where, for $0 \leq t < t' \leq T$,

$$\begin{aligned} R_{t, t'} &= \left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t') - M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \right|^2 - \left| \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t')) - \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right|^2 \\ &= \left| \int_{t \wedge \tau_\Lambda^\varepsilon}^{t' \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds \right|^2 - 2 \left(\varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t')) - \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right) \int_{t \wedge \tau_\Lambda^\varepsilon}^{t' \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds. \end{aligned} \quad (\text{I.4.9})$$

We can write

$$\begin{aligned} \left| \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t')) - \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right|^2 &= M_{\varphi^2}^{\varepsilon, \tau_\Lambda^\varepsilon}(t') - M_{\varphi^2}^{\varepsilon, \tau_\Lambda^\varepsilon}(t) - 2\varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \left(M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t') - M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \right) \\ &\quad + \int_{t \wedge \tau_\Lambda^\varepsilon}^{t' \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon(\varphi^2)(X^\varepsilon(s)) ds - 2\varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \int_{t \wedge \tau_\Lambda^\varepsilon}^{t' \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds. \end{aligned} \quad (\text{I.4.10})$$

As established in the first part of the proof, $M_{\varphi^2}^{\varepsilon, \tau_\Lambda^\varepsilon}$ and $M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}$ are $(\mathcal{F}_s^\varepsilon)_s$ -martingales. Moreover, $\varphi(X_t^{\varepsilon, \tau_\Lambda^\varepsilon})$ is $\mathcal{F}_t^\varepsilon$ -measurable. Thus, taking the expectation in (I.4.10) yields

$$\begin{aligned} \mathbb{E} \left[\left| \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t')) - \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right|^2 \right] &= \\ &= \mathbb{E} \left[\int_{t \wedge \tau_\Lambda^\varepsilon}^{t' \wedge \tau_\Lambda^\varepsilon} (\mathcal{L}^\varepsilon(\varphi^2)(X^\varepsilon(s)) ds - 2\varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s))) ds \right]. \end{aligned}$$

As in Step 1, by Assumption I.10, (I.3.5), Proposition I.3.6 and Definition I.4.1, the integrand is bounded by $\left(1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^6\right) \left(1 + \|\overline{m}(0)\|_E^{2(b+2)}\right)$ (up to a constant depending on φ , Λ and ε) and we get

$$\mathbb{E} \left[\left| \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t')) - \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right|^2 \right] \lesssim_{\varphi, \Lambda, \varepsilon} t' - t, \quad (\text{I.4.11})$$

owing to the Cauchy-Schwarz inequality, Assumptions I.4 and I.7. The Young inequality with a parameter $\eta > 0$ yields

$$\mathbb{E}[|R_{t,t'}|] \lesssim_{\varphi,\Lambda,\varepsilon} \left(1 + \frac{1}{\eta}\right) \mathbb{E} \left[\left| \int_{t \wedge \tau_\Lambda^\varepsilon}^{t' \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds \right|^2 \right] + \eta \mathbb{E} \left[|\varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t')) - \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t))|^2 \right].$$

Similarly, we get

$$\mathbb{E} \left[\left| \int_{t \wedge \tau_\Lambda^\varepsilon}^{t' \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds \right|^2 \right] \lesssim_{\varphi,\Lambda,\varepsilon} (t' - t)^2.$$

Choosing $\eta = (t' - t)^{1/2}$ yields $\mathbb{E}[|R_{t,t'}|] \lesssim_{\varphi,\varepsilon} (t' - t)^{3/2}$, which gives (I.4.8).

Step 3: We claim that $\mathbb{E} \left[\left| M_\varphi^{\varepsilon,\tau_\Lambda^\varepsilon}(t) \right|^2 \right] = \mathbb{E} \left[N^{\varepsilon,\tau_\Lambda^\varepsilon}(t) \right]$.

Taking conditional expectation in (I.4.9) leads to

$$\begin{aligned} \sum_i \mathbb{E} \left[\left| M_\varphi^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - M_\varphi^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_i) \right|^2 \mid \mathcal{F}_{t_i}^\varepsilon \right] &= \sum_i \mathbb{E} \left[R_{t \wedge t_i, t \wedge t_{i+1}} \mid \mathcal{F}_{t_i}^\varepsilon \right] \\ &\quad + \sum_i \mathbb{E} \left[|\varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_{i+1})) - \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_i))|^2 \mid \mathcal{F}_{t_i}^\varepsilon \right]. \end{aligned}$$

Using (I.4.10) and the martingale property on $M_{\varphi^2}^{\varepsilon,\tau_\Lambda^\varepsilon}$ and $M_\varphi^{\varepsilon,\tau_\Lambda^\varepsilon}$, the last term can be rewritten as

$$\sum_i \int_{t \wedge t_i \wedge \tau_\Lambda^\varepsilon}^{t \wedge t_{i+1} \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon(\varphi^2)(X^\varepsilon(s)) ds - 2 \sum_i \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_i)) \int_{t \wedge t_i \wedge \tau_\Lambda^\varepsilon}^{t \wedge t_{i+1} \wedge \tau_\Lambda^\varepsilon} \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds$$

Then, (I.4.7) yields

$$\sum_i \mathbb{E} \left[\left| M_\varphi^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - M_\varphi^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_i) \right|^2 \mid \mathcal{F}_{t_i}^\varepsilon \right] = N^{\varepsilon,\tau_\Lambda^\varepsilon}(t) + r_1 + r_2$$

where

$$\begin{aligned} r_1 &= \sum_i \mathbb{E} \left[R_{t \wedge t_i, t \wedge t_{i+1}} \mid \mathcal{F}_{t_i}^\varepsilon \right] \\ r_2 &= 2 \sum_i \mathbb{E} \left[\int_{t \wedge t_i \wedge \tau_\Lambda^\varepsilon}^{t \wedge t_{i+1} \wedge \tau_\Lambda^\varepsilon} (\varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s)) - \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t \wedge t_i))) \mathcal{L}^\varepsilon \varphi(X^\varepsilon(s)) ds \mid \mathcal{F}_{t_i}^\varepsilon \right]. \end{aligned}$$

By means of (I.4.8), $r_1 \rightarrow 0$ in \mathbb{L}^1 . For r_2 , we have

$$\begin{aligned} \mathbb{E}[|r_2|] &\leq 2 \mathbb{E} \left[\sum_i \int_{t \wedge t_i \wedge \tau_\Lambda^\varepsilon}^{t \wedge t_{i+1} \wedge \tau_\Lambda^\varepsilon} |\varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s)) - \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t_i))| |\mathcal{L}^\varepsilon \varphi(X^\varepsilon(s))| ds \right] \\ &\leq 2 \mathbb{E} \left[\sum_i \int_{t_i}^{t_{i+1}} |\varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s)) - \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t_i))| |\mathcal{L}^\varepsilon \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s))| ds \right] \\ &\leq 2 \sum_i \int_{t_i}^{t_{i+1}} \mathbb{E} \left[|\varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s)) - \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t_i))| |\mathcal{L}^\varepsilon \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s))| \right] ds \\ &\leq 2 \sum_i \int_{t_i}^{t_{i+1}} \mathbb{E} \left[|\varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s)) - \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(t_i))|^2 \right]^{1/2} \mathbb{E} \left[|\mathcal{L}^\varepsilon \varphi(X^{\varepsilon,\tau_\Lambda^\varepsilon}(s))|^2 \right]^{1/2} ds. \end{aligned}$$

As above, one can show $\mathbb{E} \left[\sup_s |\mathcal{L}^\varepsilon \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(s))|^2 \right] \lesssim_{\varphi, \Lambda, \varepsilon} 1$. Thus, (I.4.11) yields

$$\begin{aligned} \mathbb{E} [|r_2|] &\lesssim_{\varphi, \Lambda, \varepsilon} \sum_i \int_{t_i}^{t_{i+1}} \mathbb{E} \left[|\varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(s)) - \varphi(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t_i))|^2 \right]^{1/2} ds \\ &\lesssim_{\varphi, \Lambda, \varepsilon} \sum_i \int_{t_i}^{t_{i+1}} (s - t_i)^{1/2} ds \\ &\lesssim_{\varphi, \Lambda, \varepsilon} \sum_i (t_{i+1} - t_i)^{3/2} ds \\ &\lesssim_{\varphi, \Lambda, \varepsilon} T \delta^{1/2} \xrightarrow{\delta \rightarrow 0} 0. \end{aligned}$$

Thus, in \mathbb{L}^1 , we have

$$\lim_{\delta \rightarrow 0} \sum_i \mathbb{E} \left[\left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i) \right|^2 \mid \mathcal{F}_{t_i}^\varepsilon \right] = N^{\varepsilon, \tau_\Lambda^\varepsilon}(t). \quad (\text{I.4.12})$$

In particular, the expectation converges. Then, the martingale property and the tower property $\mathbb{E} [\mathbb{E} [\cdot \mid \mathcal{F}_s]] = \mathbb{E} [\cdot]$ yield

$$\begin{aligned} \mathbb{E} [N^{\varepsilon, \tau_\Lambda^\varepsilon}(t)] &= \lim_{\delta \rightarrow 0} \mathbb{E} \left[\sum_i \mathbb{E} \left[\left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) - M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i) \right|^2 \mid \mathcal{F}_{t_i}^\varepsilon \right] \right] \\ &= \lim_{\delta \rightarrow 0} \mathbb{E} \left[\sum_i \mathbb{E} \left[\left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) \right|^2 - \left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i) \right|^2 \mid \mathcal{F}_{t_i}^\varepsilon \right] \right] \\ &= \lim_{\delta \rightarrow 0} \sum_i \mathbb{E} \left[\left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_{i+1}) \right|^2 - \left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t \wedge t_i) \right|^2 \right] \\ &= \mathbb{E} \left[\left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \right|^2 \right]. \end{aligned}$$

This conclude the proof that for $t \in \mathbb{R}^+$

$$\mathbb{E} \left[\left| M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \right|^2 \right] = \mathbb{E} \left[\int_0^{t \wedge \tau_\Lambda^\varepsilon} (\mathcal{L}^\varepsilon(\varphi^2) - 2\varphi \mathcal{L}^\varepsilon \varphi)(X^\varepsilon(s)) ds \right].$$

and the proof of Proposition I.4.1. □

Remark I.4.1. Note that if \bar{m} had continuous paths, then M_φ^ε would be a continuous martingale and (I.4.12) would mean that $N^{\varepsilon, \tau_\Lambda^\varepsilon}$ is the quadratic variation of $M_\varphi^{\varepsilon, \tau_\Lambda^\varepsilon}$.

A similar proof leads to the following Proposition, where we take weaker stopping times but add some conditions on φ . The proof is omitted.

Proposition I.4.2. *Let φ be a good test function. The conclusion of Proposition I.4.1 holds in the following cases.*

- The function φ does not depend on f and τ_Λ^ε is replaced by τ^ε .
- The function φ is bounded uniformly in n and does not depend on z and τ_Λ^ε is replaced by $\tau_\Lambda(\zeta^\varepsilon)$.
- The function φ is bounded and depends only on n and τ_Λ^ε is replaced by $+\infty$.

I.4.2 The perturbed test functions method

We use the perturbed test functions method as in [PSV77] to exhibit a generator \mathcal{L} such that a possible limit point $(\rho_\Lambda, \zeta_\Lambda)$ of $((\rho^\varepsilon, \tau_\Lambda^\varepsilon, \zeta^\varepsilon, \tau_\Lambda^\varepsilon))_\varepsilon$ solves the martingale problem associated to \mathcal{L} until some limit stopping time depending on Λ . Given a test function φ , two corrector functions φ_1 and φ_2 are constructed, so that

$$\forall (f, z, n) \in L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E, \varphi^\varepsilon(f, z, n) = \varphi(\rho, z) + \varepsilon \varphi_1(f, z, n) + \varepsilon^2 \varphi_2(f, z, n), \quad (\text{I.4.13})$$

satisfies

$$\mathcal{L}^\varepsilon \varphi^\varepsilon = \mathcal{L} \varphi + o(1), \quad (\text{I.4.14})$$

when $\varepsilon \rightarrow 0$. Then, we prove that φ^ε is a good test function and that we can take the limit when $\varepsilon \rightarrow 0$ in the martingale problem associated to \mathcal{L}^ε (Proposition I.4.1) to obtain a stopped martingale problem solved by a limit point.

Based on the decomposition (I.4.1), a sufficient condition to prove (I.4.14) for φ^ε of the form (I.4.13) is to solve the following equations (I.4.15), (I.4.16) and (I.4.17) and to check that (I.4.18) holds when $\varepsilon \rightarrow 0$.

$$\mathcal{L}_2 \varphi = 0 \quad (\text{I.4.15})$$

$$\mathcal{L}_1 \varphi + \mathcal{L}_2 \varphi_1 = 0 \quad (\text{I.4.16})$$

$$\mathcal{L}_1 \varphi_1 + \mathcal{L}_2 \varphi_2 = \mathcal{L} \varphi \quad (\text{I.4.17})$$

$$\mathcal{L}_1 \varphi_2 = O(1). \quad (\text{I.4.18})$$

The properties of the resolvent operators R_λ are employed to invert \mathcal{L}_2 .

Framework for the perturbed test functions method

For a martingale problem to be relevant, it is sufficient that the class of test functions satisfying the martingale problem is separating, namely that if some random variables X and X' satisfy $\mathbb{E}[\varphi(X)] = \mathbb{E}[\varphi(X')]$ for all $\varphi \in \Phi$, then we have $X \stackrel{d}{=} X'$. In this work, we use the following class

$$\Theta = \{(\rho, z) \mapsto \psi((\rho, \xi)_{L_x^2}) \chi(z) \mid \psi \in \mathcal{C}^3(\mathbb{R}), \psi'' \in \mathcal{C}_b^1(\mathbb{R}), \xi \in H_x^3, \chi \in \mathcal{C}_b^3(\mathcal{C}_x^1)\},$$

where $\rho = \int_V f d\mu$. The class Θ is indeed separating because it separates points (see [EK86, Theorem 4.5]).

Note that the test functions depend only on ρ and z , because we expect the limit equation to be satisfied by ρ and z . It is confirmed by Section I.4.2. To simplify the notation, for $\varphi \in \Theta$, we sometimes write $\varphi(f, z, n) \doteq \varphi(\rho, z)$ and $\varphi(\rho, z) = \Psi(\rho) \chi(z)$, where $\Psi(\rho) = \psi((\rho, \xi)_{L_x^2})$.

Proposition I.4.3. *There exists an operator \mathcal{L} whose domain contains Θ and, for all $\varphi \in \Theta$, there exist two good test functions φ_1 and φ_2 such that, for all $(f, z, n) \in L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E$, we have*

$$|\varphi_1(f, z, n)| \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|n\|_E) \quad (\text{I.4.19})$$

$$|\varphi_2(f, z, n)| \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|n\|_E^{b+1}) \quad (\text{I.4.20})$$

$$|\mathcal{L}^\varepsilon \varphi^\varepsilon - \mathcal{L} \varphi|(f, z, n) \lesssim \varepsilon (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|n\|_E^{b+2}). \quad (\text{I.4.21})$$

Moreover, $\varphi^\varepsilon = \varphi + \varepsilon \varphi_1 + \varepsilon^2 \varphi_2$ is a good test function.

Moreover, if φ depends only on z , then φ_1 , φ_2 and φ^ε depend only on z and n .

Consistency result

Since we already expect the limit equation to be satisfied by ρ , equation (I.4.15) will not give us extra information. Hence, this section only present a consistency result, namely that (I.4.15) forces φ to depend on f through ρ .

In fact, let φ depend on f and z but not on n . Since φ does not depend on n , $B\varphi = 0$. Hence, (I.4.15) can be written, for all $f \in L^2(\mathcal{M}^{-1})$ and $z \in \mathcal{C}_x^1$,

$$D_f \varphi(f, z, n)(Lf) = 0. \quad (\text{I.4.22})$$

For $t \in \mathbb{R}^+$ and $f \in L^2(\mathcal{M}^{-1})$, define

$$g(t, f) = \rho\mathcal{M} + e^{-t}(f - \rho\mathcal{M}), \quad (\text{I.4.23})$$

and observe that $\partial_t g(t, f) = Lg(t, f)$ with $g(0, f) = f$. Owing to (I.4.22), the mapping $t \mapsto \varphi(g(t, f), z)$ is constant. Since $g(t, f) \xrightarrow[t \rightarrow \infty]{} \rho\mathcal{M}$, by continuity of φ , we get $\varphi(f, z, n) = \varphi(\rho\mathcal{M}, z, n)$, which depends on f only through ρ .

Construction of the first corrector function φ_1

The first corrector function φ_1 is defined as the solution of (I.4.16): the formal solution to Poisson equation will provide an expression for φ_1 , then we will check that this expression indeed solves (I.4.16).

Let $g(t, f)$ be defined by (I.4.23) and $m(t, n)$ be defined in Section I.2.1 (Markov process of infinitesimal generator B starting from n). The process $((g(t, f), z, m(t, n)))_{t \in \mathbb{R}^+}$ is a $L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E$ -valued Markov process of generator \mathcal{L}_2 starting from (f, z, n) . Denote by $(Q_t)_{t \in \mathbb{R}^+}$ its transition semi-group. Note that this semi-group does not have a unique invariant distribution, since for any ρ fixed, $\delta_{\rho\mathcal{M}} \otimes \delta_z \otimes \nu$ is an invariant distribution. However on every space $\{(f', z', n) \in L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E \mid \int_V f' d\mu = \rho, z' = z\}$, this measure is the unique invariant distribution. Indeed, $\delta_{\rho\mathcal{M}}$, δ_z and ν are respectively the unique invariant distributions of each marginal process (on the corresponding subspaces), and $\delta_{\rho\mathcal{M}} \otimes \delta_z \otimes \nu$ is the only coupling of these three marginal distributions.

For $\Phi : L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E \rightarrow \mathbb{R}$, denote by

$$\langle \Phi \rangle_{\rho, z} \doteq \int_E \Phi(\rho\mathcal{M}, z, n) d\nu(n) = \int_{L^2(\mathcal{M}^{-1}) \times \mathcal{C}_x^1 \times E} \Phi d(\delta_{\rho\mathcal{M}} \otimes \delta_z \otimes \nu)$$

the integral against this invariant distribution. For $\varphi \in \Theta$, $\varphi(\rho, z) = \Psi(\rho)\chi(z)$, let us compute $\mathcal{L}_1\varphi$.

We have

$$\begin{aligned} \mathcal{L}_1\varphi(f, z, n) &= D_f \varphi(f, z, n)(-Af + nf) + D_z \varphi(f, z, n)(n) \\ &= D\Psi(\rho)(-\overline{A}f + n\rho)\chi(z) + \Psi(\rho)D\chi(z)(n), \end{aligned}$$

where $\overline{h} = \int_V h(v) d\mu(v)$.

Owing to (I.2.1), one can write, for all $\rho \in L_x^2$, $\overline{A(\rho\mathcal{M})} = 0$. Moreover, since ν is centered by Assumption I.5, any term linear in n vanishes when integrating with respect to ν . Hence, we have checked that

$$\forall \rho \in L_x^2, \forall z \in \mathcal{C}_x^1, \langle \mathcal{L}_1\varphi \rangle_{\rho, z} = 0.$$

Using the expansion of $\mathcal{L}_1\varphi$, we have for all f, z, n

$$\begin{aligned} \int_0^\infty Q_t \mathcal{L}_1 \varphi(f, z, n) dt &= \int_0^\infty \mathbb{E} [\mathcal{L}_1 \varphi(g(t, f), z, m(t, n))] dt \\ &= \int_0^\infty \left(-D\Psi(\rho)(\overline{Ag(t, f)})\chi(z) + \mathbb{E} [D\Psi(\rho)(\rho m(t, n))]\chi(z) \right. \\ &\quad \left. + \Psi(\rho)D\chi(z)(m(t, n)) \right) dt, \end{aligned}$$

owing to the identity $\overline{g(t, f)} = \rho$. Equation (I.4.23) yields

$$\overline{Ag(t, f)} = e^{-t}\overline{Af},$$

since $\overline{A\rho\mathcal{M}} = 0$. Thus, owing to Definition I.3.1, we define

$$\begin{aligned} \varphi_1(f, z, n) &= \int_0^\infty Q_t \mathcal{L}_1 \varphi(f, z, n) dt \\ &= D\Psi(\rho)(-\overline{Af} + R_0(n)\rho)\chi(z) + \Psi(\rho)R_0[D\chi(z)](n). \end{aligned} \quad (\text{I.4.24})$$

It is straightforward to check that φ_1 defined by (I.4.24) solves (I.4.16). Moreover, it satisfies the condition (I.4.19). It remains to prove that φ_1 is a good test function. Owing to Assumption I.9 and (I.4.24), $\varphi_1 \in D(B)$ and $\varphi_1^2 \in D(B)$. For $h \in L^2(\mathcal{M}^{-1})$, we have

$$\begin{aligned} D_f \varphi_1(f, z, n)(h) &= D^2\Psi(\rho)(-\overline{Af} + R_0(n)\rho, \bar{h})\chi(z) + D\Psi(\rho)(-\overline{Ah} + R_0(n)\bar{h})\chi(z) \\ &\quad + D\Psi(\rho)(\bar{h})R_0[D\chi(z)](n), \end{aligned}$$

hence $D_f \varphi_1(f, z, n)(Ah)$ is well-defined (as in Definition I.4.1) and $\varphi_1, D_f \varphi_1(f, z, n)(h)$ and $D_f \varphi_1(f, z, n)(Ah)$ have at most polynomial growth in the sense of (I.4.3). For $n_2 \in E$, we have

$$\begin{aligned} D_z \varphi_1(f, z, n)(n_2) &= D\Psi(\rho)(-\overline{Af} + R_0(n)\rho)D\chi(z)(n_2) \\ &\quad + \Psi(\rho)D[z' \mapsto R_0[D\chi(z')](n)](z)(n_2). \end{aligned}$$

Using Lemma I.3.1 and the assumption $\chi \in \mathcal{C}_b^3(\mathcal{C}_x^1)$, we write

$$\begin{aligned} D[R_0[D\chi(\cdot)](n)](z)(n_2) &= D\left[z' \mapsto \int_0^\infty P_t D\chi(z')(n) dt\right](z)(n_2) \\ &= \int_0^\infty P_t [D^2\chi(z)(\cdot, n_2)](n) \\ &= R_0[D^2\chi(z)(\cdot, n_2)](n). \end{aligned}$$

This leads to

$$D_z \varphi_1(f, z, n)(n_2) = D\Psi(\rho)(-\overline{Af} + R_0(n)\rho)D\chi(z)(n_2) + \Psi(\rho)R_0[D^2\chi(z)(\cdot, n_2)](n).$$

Once again using Lemma I.3.1 and that $\chi \in \mathcal{C}_b^3(\mathcal{C}_x^1)$, one checks that $D_z \varphi_1$ has at most polynomial growth in the sense of (I.4.3). Thus φ_1 satisfies (I.4.3).

Construction of the second corrector function φ_2

The second corrector φ_2 is defined as a solution of (I.4.17). To solve (I.4.17), we need the centering condition $\langle \mathcal{L}\varphi - \mathcal{L}_1\varphi_1 \rangle_{\rho,z} = 0$. This identity will be the definition of $\mathcal{L}\varphi$.

First, let us compute $\mathcal{L}_1\varphi_1$. Using the derivative calculated in (I.4.24), $\mathcal{L}_1\varphi_1$ can be written as

$$\mathcal{L}_1\varphi_1(f, z, n) = c(f, z) + \ell(f, z, n) + q(f, z, n)$$

where c , ℓ and q are defined by

$$c(f, z) = D^2\Psi(\rho)(\overline{Af}, \overline{Af})\chi(z) + D\Psi(\rho)(\overline{A^2f})\chi(z) \quad (\text{I.4.25})$$

$$\begin{aligned} \ell(f, z, n) = & -D^2\Psi(\rho)(\overline{Af}, R_0(n)\rho + n\rho)\chi(z) \\ & - D\Psi(\rho)(\overline{A(nf)} + R_0(n)(\overline{Af}))\chi(z) \\ & - D\Psi(\rho)(\overline{Af})(R_0[D\chi(z)](n) + D\chi(z)(n)) \end{aligned} \quad (\text{I.4.26})$$

$$\begin{aligned} q(f, z, n) = & D^2\Psi(\rho)(n\rho, R_0(n)\rho)\chi(z) + D\Psi(\rho)(R_0(n)(n\rho))\chi(z) \\ & + D\Psi(\rho)(n\rho)R_0[D\chi(z)](n) + D\Psi(\rho)(R_0(n)\rho)D\chi(z)(n) \\ & + \Psi(\rho)R_0[D^2\chi(z)(\cdot, n)](n). \end{aligned} \quad (\text{I.4.27})$$

Note that, for fixed f and z , c does not depend on n , ℓ is pseudo-linear in n and q is pseudo-quadratic in n as introduced in Definition I.3.2.

The function $\ell(f, z, \cdot)$ is indeed pseudo-linear as a sum of continuous linear and pseudo-linear forms, yielding $\langle \ell \rangle_{\rho,z} = 0$ for all ρ and z . Using also that $\overline{A\rho\mathcal{M}} = 0$, we get an explicit definition of \mathcal{L} :

$$\begin{aligned} \mathcal{L}\varphi(\rho, z) \doteq \langle \mathcal{L}_1\varphi_1 \rangle_{\rho,z} = & D\Psi(\rho)(\overline{A^2\rho\mathcal{M}})\chi(z) \\ & + \int D^2\Psi(\rho)(n\rho, R_0(n)\rho)d\nu(n)\chi(z) \\ & + \int D\Psi(\rho)(R_0(n)(n\rho))d\nu(n)\chi(z) \\ & + \int D\Psi(\rho)(n\rho)R_0[D\chi(z)](n)d\nu(n) \\ & + \int D\Psi(\rho)(R_0(n)\rho)D\chi(z)(n)d\nu(n) \\ & + \Psi(\rho) \int R_0[D^2\chi(z)(\cdot, n)](n)d\nu(n). \end{aligned} \quad (\text{I.4.28})$$

Note that by taking $\chi = 1$, we obtain the same expression of \mathcal{L} as in [DV12].

Since the centering condition for the Poisson equation (I.4.17) is satisfied by construction

of \mathcal{L} , the second corrector function φ_2 can be defined as follows: for all f, z, n ,

$$\begin{aligned}\varphi_2(f, z, n) &\doteq \int_0^\infty Q_t \left(\mathcal{L}_1 \varphi_1 - \langle \mathcal{L}_1 \varphi_1 \rangle_{\rho, z} \right) (f, z, n) dt \\ &= \int_0^\infty Q_t \left(c - \langle c \rangle_{\rho, z} \right) (f, z, n) dt \\ &\quad + \int_0^\infty Q_t \ell(f, z, n) dt \\ &\quad + \int_0^\infty Q_t \left(q - \langle q \rangle_{\rho, z} \right) (f, z, n) dt \\ &\doteq \varphi_2^c(f, z, n) + \varphi_2^\ell(f, z, n) + \varphi_2^q(f, z, n).\end{aligned}$$

Once again, one can check that φ_2 satisfies (I.4.17). It only remains to prove (I.4.20), (I.4.21) and that φ^ε is a good test function. Since

$$\mathcal{L}^\varepsilon \varphi^\varepsilon = \mathcal{L} \varphi + \varepsilon \mathcal{L}_1 \varphi_2, \quad (\text{I.4.29})$$

equation (I.4.21) comes from an estimate on $\mathcal{L}_1 \varphi_2(f, n)$ in terms of f, n , and φ .

Controls on the second corrector function φ_2

The aim of this section is to prove some estimates for $\varphi_2(f, z, n)$ and its derivatives to establish that (I.4.3), (I.4.20) and (I.4.21) are satisfied. Let $f, h \in L^2(\mathcal{M}^{-1})$, $z \in \mathcal{C}_x^1$ and $n, n_2 \in E$ and let $S_1 = \|f\|_{L^2(\mathcal{M}^{-1})} \vee \|h\|_{L^2(\mathcal{M}^{-1})}$ and $S_2 = \|n\|_E \vee \|n_2\|_E$.

Estimates on φ_2^c We have, using $\langle c \rangle_{\rho, z} = c(\rho \mathcal{M}, z)$,

$$c(f, z) - c(\rho \mathcal{M}, z) = D^2 \Psi(\rho)(\overline{A}f, \overline{A}f)\chi(z) + D\Psi(\rho)(\overline{A^2(f - \rho \mathcal{M})})\chi(z).$$

Recall that $\overline{Ag(t, f)} = e^{-t} \overline{A}f$. Hence, using (I.4.23), we get

$$\begin{aligned}Q_t \left(c - \langle c \rangle_{\rho, z} \right) (f, z, n) &= \mathbb{E} [c(g(t, f), z) - c(\rho \mathcal{M}, z)] \\ &= e^{-2t} D^2 \Psi(\rho)(\overline{A}f, \overline{A}f)\chi(z) + e^{-t} D\Psi(\rho)(\overline{A^2(f - \rho \mathcal{M})})\chi(z).\end{aligned}$$

By integration, we get

$$\varphi_2^c(f, z, n) = \frac{1}{2} D^2 \Psi(\rho)(\overline{A}f, \overline{A}f)\chi(z) + D\Psi(\rho)(\overline{A^2(f - \rho \mathcal{M})})\chi(z). \quad (\text{I.4.30})$$

Moreover, we obtain

$$\begin{aligned}D_f \varphi_2^c(f, z, n)(h) &= \frac{1}{2} D^3 \Psi(\rho)(\overline{A}f, \overline{A}f, \overline{h})\chi(z) + D^2 \Psi(\rho)(\overline{A}f, \overline{A}h)\chi(z) \\ &\quad + D^2 \Psi(\rho)(\overline{A^2(f - \rho \mathcal{M})}, \overline{h})\chi(z) + D\Psi(\rho)(\overline{A^2(h - \overline{h} \mathcal{M})})\chi(z), \\ D_z \varphi_2^c(f, z, n)(n_2) &= \frac{1}{2} D\Psi(\rho)(\overline{A}f, \overline{A}f)D\chi(z)(n) + D\Psi(\rho)(\overline{A^2(f - \rho \mathcal{M})})D\chi(z)(n).\end{aligned}$$

Recall that $\|f - \rho \mathcal{M}\|_{L^2(\mathcal{M}^{-1})}^2 + \|\rho\|_{L_x^2}^2 = \|f\|_{L^2(\mathcal{M}^{-1})}^2$, hence $\|f - \rho \mathcal{M}\|_{L^2(\mathcal{M}^{-1})} \leq \|f\|_{L^2(\mathcal{M}^{-1})}$. Then, since $\Psi(\rho) = \psi((\rho, \xi)_{L_x^2})$ and $\psi'' \in \mathcal{C}_b^1(\mathbb{R})$, we get that φ_2^c satisfies (I.4.3). More precisely, the following estimates hold:

$$\begin{aligned}|\varphi_2^c(f, z, n)| &\lesssim_\varphi 1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2 \\ |\mathcal{L}_1 \varphi_2^c(f, z, n)| &\lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|n\|_E).\end{aligned}$$

Estimates on φ_2^ℓ Using (I.4.23), (I.4.26) and that $\overline{A\rho\mathcal{M}} = \overline{A(n\rho)\mathcal{M}} = 0$, we get

$$\forall(f, z, n), \ell(g(t, f), z, n) = e^{-t}\ell(f, z, n).$$

Thus, we have

$$Q_t\ell(f, z, n) = \mathbb{E}[\ell(g(t, f), z, m(t, n))] = e^{-t}\mathbb{E}[\ell(f, z, m(t, n))] = e^{-t}P_t\ell(f, z, n),$$

and by integrating with respect to t , we get

$$\varphi_2^\ell(f, z, n) = R_1\ell(f, z, n).$$

Moreover, from Lemma I.3.1 and (I.4.26), it is straightforward to check that

$$[\ell(f, \cdot)]_{\text{Lip}} \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2).$$

Hence, Lemma I.3.1 yields

$$|\varphi_2^\ell(f, z, n)| \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|n\|_E).$$

Since the operator R_1 acts only on the variable n , it commutes with the derivatives D_f and D_z in the following sense:

$$\begin{aligned} D_f[R_1\ell](f, z, n)(h) &= R_1[D_f\ell(f, z, \cdot)(h)](n) \\ D_z[R_1\ell](f, z, n)(n_2) &= R_1[D_z\ell(f, z, \cdot)(n_2)](n). \end{aligned}$$

Thus, after calculating the derivatives of ℓ , we get estimates on the derivatives of φ_2^ℓ the same way we got estimates on φ_2^ℓ . This leads to

$$\begin{aligned} |D_f\varphi_2^\ell(f, z, n)(Ah)| &\lesssim_\varphi (1 + S_1^3)(1 + S_2) \\ |D_f\varphi_2^\ell(f, z, n)(n_2f)| &\lesssim_\varphi (1 + S_1^3)(1 + S_2^2) \\ |D_z\varphi_2^\ell(f, z, n)(n_2)| &\lesssim_\varphi (1 + S_1^2)(1 + S_2^2), \end{aligned}$$

hence φ_2^ℓ satisfies (I.4.3). Finally, the following estimates hold

$$\begin{aligned} |\varphi_2^\ell(f, z, n)| &\lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|n\|_E) \\ |\mathcal{L}_1\varphi_2^\ell(f, z, n)| &\lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|n\|_E^2). \end{aligned}$$

Estimates on φ_2^q The function q depends of f only through ρ . Since $\overline{g(t, f)} = \rho$ does not depend on t , we get $Q_tq = P_tq$ and

$$\varphi_2^q(f, z, n) = R_0[q - \langle q \rangle_{\rho, z}](f, z, n)$$

It is straightforward to compute the derivatives of q with respect to f and z from (I.4.27). One can deduce estimates for $[q(f, z, \cdot)]_{\text{quad}}$ and for the first order derivatives $[D_fq(f, z, \cdot)(n_2f)]_{\text{quad}}$, $[D_fq(f, z, \cdot)(Af)]_{\text{quad}}$ and $[D_zq(f, z, \cdot)(n_2)]_{\text{quad}}$. Reasoning as for R_1 , the resolvent R_0 acts only on n , and thus commutes with D_f and D_z . Thus, Lemma I.3.1 with $\lambda = 0$ proves that φ_2^q satisfies (I.4.3). Finally, the following estimates hold

$$\begin{aligned} |\varphi_2^q(f, z, n)| &\lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|n\|_E^{b+1}) \\ |\mathcal{L}_1\varphi_2^q(f, z, n)| &\lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|n\|_E^{b+2}). \end{aligned}$$

This concludes the proof that φ_2 satisfies (I.4.3) and the proof of the estimates of Proposition I.4.3 on φ_2 and $\mathcal{L}_1\varphi_2$.

Good test function property

It only remains to prove that φ^ε is a good test function. The estimates (I.4.3) are satisfied by $\varepsilon\varphi_1$ and $\varepsilon^2\varphi_2$, hence by their sum φ^ε . Moreover, using the notation introduced in Section I.3.1, φ^ε can be written as

$$\begin{aligned}\varphi^\varepsilon(f, z, n) &= \varphi(\rho, z) - \varepsilon D\Psi(\rho)(\overline{Af})\chi(z) + \varepsilon R_0 [D\Psi(\rho)(\cdot\rho)](n)\chi(z) \\ &\quad + \varepsilon \Psi(\rho)R_0 [D\chi(z)](n) + \varepsilon^2\varphi_2^\varepsilon(f, z) + \varepsilon^2 R_1 \ell(f, z, n) \\ &\quad + \varepsilon^2 R_0 \left[q - \langle q \rangle_{\rho, z} \right] (f, z, n).\end{aligned}$$

Observe that each term either does not depend on n or can be written $R_\lambda \theta$ with θ as in Definition I.3.1. As a consequence, owing to Assumption I.9, any product of at most two of these terms belongs to $D(B)$. Thus, $\varphi^\varepsilon \in D(B)$ and $(\varphi^\varepsilon)^2 \in D(B)$. This concludes the proof that φ^ε is a good test function, and the proof of Proposition I.4.3.

I.5 Dynamics associated with the limiting equation

In this section, we show that the operator \mathcal{L} is the generator of the limit equation (I.2.13) and that the martingale problem associated to \mathcal{L} characterizes the solution of (I.2.13).

Definition I.5.1. Let $\rho_0 \in L_x^2$ and let $\sigma > 0$. A process (ρ, ζ) is said to be a weak solution to (I.2.13) in L_x^2 if the following assertions are satisfied

- (i) $\rho(0) = \rho_0$,
- (ii) $\rho \in L^\infty([0, T], L_x^2) \cap \mathcal{C}^0([0, T], H_x^{-\sigma})$ a.s. and $\zeta \in \mathcal{C}^0([0, T], \mathcal{C}_x^1)$ a.s.,
- (iii) there exists $(B_i)_i$ a sequence of independent standard Brownian motions such that (ρ, ζ) is adapted to the filtration generated by $(B_i)_i$ and such that, for all $\xi \in L_x^2$ and $t \in [0, T]$, we have a.s.

$$\begin{aligned}(\rho(t), \xi)_{L_x^2} &= (\rho_0, \xi)_{L_x^2} + \int_0^t (\rho(s), \operatorname{div}(K \nabla \xi))_{L_x^2} ds + \int_0^t \left(\frac{1}{2} F \rho(s), \xi \right)_{L_x^2} ds \\ &\quad + \sum_i \sqrt{q_i} \int_0^t (F_i \rho(s), \xi)_{L_x^2} dB_i(s) \quad (\text{I.5.1})\end{aligned}$$

$$\zeta(t) = \sum_i \sqrt{q_i} F_i B_i(t). \quad (\text{I.5.2})$$

Note that the sum in (I.5.2) does converge in $\mathcal{C}^0([0, T], \mathcal{C}_x^1)$ owing to Lemma I.3.3.

The solution to this equation exists and is unique in distribution. The existence can be proved using energy estimates, Itô formula and regularization argument. The uniqueness comes from pathwise uniqueness which derives from the same arguments. We do not give details concerning existence and uniqueness, however, in the proof of the following Proposition, we established the aforementioned energy estimate I.5.4.

Proposition I.5.1. Let $\sigma > 0$ and let $(\rho, \zeta) \in \mathcal{C}^0([0, T], H_x^{-\sigma}) \times \mathcal{C}^0([0, T], \mathcal{C}_x^1)$.

If (ρ, ζ) is the weak solution to (I.2.13) in $H_x^{-\sigma}$, then, for any test function $\varphi \in \Theta$, the process

$$M_\varphi(t) = \varphi(\rho(t), \zeta(t)) - \varphi(\rho_0, 0) - \int_0^t \mathcal{L}\varphi(\rho(s), \zeta(s)) ds$$

is a martingale for the filtration generated by (ρ, ζ) .

Conversely, if for all $\varphi \in \Theta$, M_φ and M_{φ^2} are martingales, then (ρ, ζ) is the weak solution of (I.2.13) in $H_x^{-\sigma}$.

Proof. Let us first prove that \mathcal{L} is the generator associated to (I.2.13). The expression of $\mathcal{L}\varphi$ is given by (I.4.28). First note that $\overline{A^2\rho\mathcal{M}} = \text{div}(K\nabla\rho)$, which is the first term of (I.2.13). The third term of (I.4.28) is associated to the second term of (I.2.13):

$$\begin{aligned} \int D\Psi(\rho)(R_0(n)(n\rho)) d\nu(n) &= \mathbb{E} \left[\int_0^\infty D\Psi(\rho)(\rho\overline{m}(0)\overline{m}(t)) dt \right] \\ &= \frac{1}{2} \mathbb{E} \left[\int_{\mathbb{R}} D\Psi(\rho)(\rho\overline{m}(0)\overline{m}(t)) dt \right] \\ &= \frac{1}{2} D\Psi(\rho)(\rho F). \end{aligned}$$

To rewrite the second term of (I.4.28), assume first that the bilinear form $D^2\Psi(\rho)$ on L_x^2 admits a kernel k_ρ . Then, we have

$$\begin{aligned} \int D^2\Psi(\rho)(n\rho, R_0(n)\rho) d\nu(n) &= \frac{1}{2} \mathbb{E} \left[\int_0^\infty D^2\Psi(\rho)(\rho\overline{m}(0), \rho\overline{m}(t)) dt \right] \\ &= \frac{1}{2} \mathbb{E} \left[\int_0^\infty \int \int k_\rho(x, y) \rho(x) \overline{m}(0)(x) \rho(y) \overline{m}(t)(y) dx dy dt \right] \\ &= \frac{1}{2} \int \int k_\rho(x, y) k(x, y) \rho(x) \rho(y) dx dy. \end{aligned}$$

Owing to Mercer's Theorem (see [FM09]), the kernel k can be expressed in terms of the eigenvectors and eigenvalues of Q :

$$\forall x, y, k(x, y) = \sum_i q_i F_i(x) F_i(y).$$

It is straightforward to check that $k^{(1/2)}(x, y) = \sum_i q_i^{1/2} F_i(x) F_i(y)$, $x, y \in \mathbb{T}^d$, defines a kernel for $Q^{1/2}$ and satisfies $k(x, y) = \int k^{(1/2)}(x, z) k^{(1/2)}(y, z) dz$. Thus, we have

$$\begin{aligned} \int D^2\Psi(\rho)(n\rho, R_0(n)\rho) d\nu(n) &= \frac{1}{2} \int \int \int k_\rho(x, y) k^{(1/2)}(x, z) \rho(x) k^{(1/2)}(y, z) \rho(y) dx dy dz \\ &= \frac{1}{2} \text{Tr} [(\rho Q^{1/2}) D^2\Psi(\rho) (\rho Q^{1/2})^*]. \end{aligned} \tag{I.5.3}$$

By density of the functions whose second derivative admits a kernel k_ρ in \mathcal{C}^2 , this formula holds for all test functions $\varphi \in \Theta$. Using similar reasoning for the three remaining terms, we get

$$\begin{aligned} \mathcal{L}\varphi(\rho, \zeta) &= D\Psi(\rho)(\text{div}(K\nabla\rho) + \frac{1}{2}F\rho)\chi(\zeta) \\ &\quad + \frac{1}{2} \text{Tr} \left[(\rho Q^{1/2}, Q^{1/2}) \begin{pmatrix} D^2\Psi(\rho)\chi(\zeta) & D\Psi(\rho) \otimes D\chi(\zeta) \\ D\Psi(\rho) \otimes D\chi(\zeta) & \Psi(\rho) D^2\chi(\zeta) \end{pmatrix} (\rho Q^{1/2}, Q^{1/2})^* \right], \end{aligned}$$

which is the generator of (I.2.13). Once moment estimates for ρ have been obtained in L_x^2 , integrability of M_φ is ensured. In addition, estimates on $\varphi(\rho(t), \zeta(t))$ and $\mathcal{L}\varphi(\rho(t), \zeta(t))$ (uniformly in $t \in [0, T]$) are also obtained, since φ and $\mathcal{L}\varphi$ have at most quadratic growth. Then, the proof that M_φ is a martingale follows the same strategy as for the proof of Proposition I.4.1. This proof is omitted. It thus remains to prove the moment estimates for ρ .

We apply Itô's formula, equation (I.5.1) and we take the expectation (so that the martingale part vanishes), to get

$$\begin{aligned} \frac{1}{2} \mathbb{E} [(\rho(t), \xi)_{L_x^2}^2] &= \frac{1}{2} \mathbb{E} [(\rho_0, \xi)_{L_x^2}^2] + \mathbb{E} \int_0^t (\rho(s), \operatorname{div}(K \nabla \xi))_{L_x^2} (\rho(s), \xi)_{L_x^2} ds \\ &\quad + \mathbb{E} \int_0^t \left(\frac{1}{2} F \rho(s), \xi \right)_{L_x^2} (\rho(s), \xi)_{L_x^2} ds + \frac{1}{2} \sum_i q_i \mathbb{E} \int_0^t (F_i \rho(s), \xi)_{L_x^2}^2 ds. \end{aligned}$$

Then, we evaluate at $\xi = e_\ell$ with $\ell \in \mathbb{Z}^d$ and e_ℓ the Fourier basis $e_\ell(x) = \exp(2i\pi\ell \cdot x)$. Let $\lambda_\ell = 4\pi^2 \ell \cdot K \ell$ so that $\operatorname{div}(K \nabla e_\ell) = -\lambda_\ell e_\ell$. We sum this formula for $|\ell| \leq L$. Let P_L be the orthogonal projector on the space generated by $\{e_\ell \mid |\ell| \leq L\}$. Since $\lambda_\ell \geq 0$, we get

$$\begin{aligned} \frac{1}{2} \mathbb{E} [\|P_L \rho(t)\|_{L_x^2}^2] &\leq \frac{1}{2} \mathbb{E} [\|P_L \rho_0\|_{L_x^2}^2] + \mathbb{E} \int_0^t \frac{1}{2} \|P_L(F \rho(s))\|_{L_x^2} \|P_L \rho(s)\|_{L_x^2} ds \\ &\quad + \frac{1}{2} \sum_i \mathbb{E} \int_0^t q_i \|P_L(F_i \rho(s))\|_{L_x^2}^2 ds \\ &\leq \frac{1}{2} \mathbb{E} [\|P_L \rho_0\|_{L_x^2}^2] + \frac{1}{2} \left(\|F\|_{L^\infty} + \sum_i q_i \|F_i\|_{L^\infty}^2 \right) \mathbb{E} \int_0^t \|\rho(s)\|_{L_x^2}^2 ds. \end{aligned}$$

Taking $L \rightarrow \infty$, using Lemma I.3.3 and Gronwall's Lemma, we get

$$\mathbb{E} [\|\rho(t)\|_{L_x^2}^2] \lesssim \mathbb{E} [\|\rho_0\|_{L_x^2}^2]. \quad (\text{I.5.4})$$

This concludes the proof of the moment estimates for ρ , hence the proof that M_φ is a martingale.

Conversely, assume that for all $\varphi \in \Theta$, M_φ and M_{φ^2} are martingales. It holds in particular for regular and bounded test functions φ . It is then standard that a solution to this martingale problem is the Markov process of generator \mathcal{L} (see for instance [EK86, Chapter 4]), based on Lévy's martingale representation theorem in Hilbert spaces (see [DPZ14, Theorem 8.2]). This concludes the proof since we already proved that \mathcal{L} is the generator associated to (I.2.13). \square

I.6 Tightness of the coupled stopped process

In this section, we prove the following Proposition.

Proposition I.6.1. *Let $\Lambda \in (0, \infty)$. The family of processes $((\rho^\varepsilon, \tau_\Lambda(\zeta^\varepsilon)), \zeta^\varepsilon, \tau_\Lambda(\zeta^\varepsilon))_\varepsilon$ is tight in the space $\mathcal{C}^0([0, T], H_x^{-\sigma}) \times \mathcal{C}^0([0, T], \mathcal{C}_x^1)$ for any $\sigma > 0$. Moreover, the family $(\zeta^\varepsilon)_\varepsilon$ is tight in $\mathcal{C}^0([0, T], \mathcal{C}_x^1)$.*

To simplify the notation, we write $\mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1$ for $\mathcal{C}^0([0, T], H_x^{-\sigma}) \times \mathcal{C}^0([0, T], \mathcal{C}_x^1)$.

Owing to Slutsky's Lemma (see [Bil99, Theorem 4.1]) and to Lemma I.3.4, Proposition I.6.1 is equivalent to the tightness of $((\rho^{\varepsilon, \tau_\Lambda^\varepsilon}, \zeta^{\varepsilon, \tau_\Lambda^\varepsilon}))_\varepsilon$ and $(\zeta^{\varepsilon, \tau_\Lambda^\varepsilon})_\varepsilon$.

Since these processes are pathwise continuous, we have the following inequality between the modulus of continuity w for continuous functions and the modulus of continuity w' for càdlàg functions (see [Bil99, equation (14.11)]):

$$w_X(\delta) \leq 2w'_X(\delta),$$

with, for a càdlàg function X ,

$$w_X(\delta) \doteq \sup_{0 \leq t \leq s \leq t+\delta \leq T} \|X(s) - X(t)\|$$

$$w'_X(\delta) \doteq \sup_{(t_i)_i} \max_i \sup_{t_i \leq t \leq s < t_{i+1}} \|X(s) - X(t)\|,$$

where $(t_i)_i$ is a subdivision of $[0, T]$. Therefore, the tightness in the Skorokhod space $D_T H_x^{-\sigma} \times D_T \mathcal{C}_x^1$ (respectively $D_T \mathcal{C}_x^1$) implies the tightness in $\mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1$ (respectively in $\mathcal{C}_T^0 \mathcal{C}_x^1$).

Owing to [Jak86, Theorem 3.1], tightness in the Skorokhod space follows from the following claims, which are proved in Sections I.6.1 and I.6.2 respectively.

- (i) For all $\eta > 0$, there exists some compact sets $K_\eta \subset H_x^{-\sigma}$ and $K'_\eta \subset \mathcal{C}_x^1$ such that for all $\varepsilon > 0$,

$$\mathbb{P}(\forall t \in [0, T], \rho^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \in K_\eta) > 1 - \eta \quad (\text{I.6.1})$$

$$\mathbb{P}(\forall t \in [0, T], \zeta^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \in K'_\eta) > 1 - \eta \quad (\text{I.6.2})$$

$$\mathbb{P}(\forall t \in [0, T], \zeta^{\varepsilon, \tau_\Lambda^\varepsilon}(t) \in K'_\eta) > 1 - \eta. \quad (\text{I.6.3})$$

- (ii) If φ is a sum of a finite number of bounded functions $\varphi_i \in \Theta$, then $(\varphi(\rho^{\varepsilon, \tau_\Lambda^\varepsilon}, \zeta^{\varepsilon, \tau_\Lambda^\varepsilon}))_\varepsilon$ is tight in $D([0, T], \mathbb{R})$.

For any $\tilde{\varphi} \in \Theta$ with $\psi = 1$, $(\tilde{\varphi}(\zeta^{\varepsilon, \tau_\Lambda^\varepsilon}))_\varepsilon$ is tight in $D([0, T], \mathbb{R})$.

We ask of φ to be a finite sum of test functions because [Jak86, Theorem 3.1] requires the class of test functions to separate points and to be closed under addition, but Θ does not satisfies the latter condition.

I.6.1 Proof of the first claim Item (i)

Using Proposition I.3.6 and the Markov inequality, we have for $R > 0$,

$$\mathbb{P}\left(\exists t \in [0, T], \|\rho^{\varepsilon, \tau_\Lambda^\varepsilon}(t)\|_{L_x^2} > R\right) \leq \frac{\mathbb{E}\left[\sup_{t \in [0, T]} \|\rho^{\varepsilon, \tau_\Lambda^\varepsilon}(t)\|_{L_x^2}\right]}{R} \lesssim_\Lambda \frac{\mathbb{E}\left[\|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}\right]}{R}.$$

Note that stopping the processes at $\tau_\Lambda(\zeta^\varepsilon)$ is necessary at this point. Owing to the compact embedding $L_x^2 \subset H_x^{-\sigma}$ for $\sigma > 0$, we get (I.6.1).

Since (I.6.2) is a consequence of (I.6.3), it remains to prove (I.6.3). Owing to Ascoli's Theorem, we have a compact embedding of the Hölder space $\mathcal{C}_x^{1, \delta} \subset \mathcal{C}_x^1$ for any $\delta > 0$. Moreover, with $s = \lfloor d/2 \rfloor + 2$, we have a continuous embedding $H_x^s \subset \mathcal{C}_x^{1, \delta}$ for any $\delta \in (0, s - \frac{d}{2} - 1]$. Then (I.6.3) is a consequence of Proposition I.6.2 below and of the Markov inequality.

Proposition I.6.2. Recall that τ^ε is defined by (I.2.9). Then, for all $T > 0$, we have

$$\sup_{\varepsilon} \mathbb{E} \left[\sup_{t \in [0, T]} \|\zeta^{\varepsilon, \tau^\varepsilon}(t)\|_{H_x^{[d/2]+2}}^2 \right] < \infty.$$

Proof. The idea of this proof is to express ζ^ε (and its derivatives) as a sum of a small term and a martingale, and then to estimate the martingale using Doob's Maximal Inequality. This argument is used two times in a row, and the estimates heavily rely on Assumptions I.9 and I.10.

Since $\zeta^\varepsilon(t) \in E = \mathcal{C}_x^{2[d/2]+4} \subset H_x^{[d/2]+2}$, it is sufficient to prove that for all multi-indices β of length $|\beta| \leq [d/2] + 2$, we have

$$\sup_{\varepsilon} \mathbb{E} \left[\sup_{t \in [0, T]} \left\| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t)}{\partial x^\beta} \right\|_{L_x^2}^2 \right] < \infty.$$

Fix such a β and let $\varepsilon > 0$. First note that

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left\| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t)}{\partial x^\beta} \right\|_{L_x^2}^2 \right] \leq \int \mathbb{E} \left[\sup_{t \in [0, T]} \left| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t, x)}{\partial x^\beta} \right|^2 \right] dx. \quad (\text{I.6.4})$$

For $x \in \mathbb{T}^d$, define $\theta_{x, \beta} \in E^*$ by

$$\forall n \in E, \theta_{x, \beta}(n) = \frac{\partial^{|\beta|} n}{\partial x^\beta}(x).$$

Since \bar{m}^ε is almost surely an E -valued càdlàg function, the derivative and the integral commute in the following computation:

$$\frac{\partial^{|\beta|} \zeta^\varepsilon(t, x)}{\partial x^\beta} = \frac{1}{\varepsilon} \int_0^t \frac{\partial^{|\beta|} \bar{m}^\varepsilon(s, x)}{\partial x^\beta} ds = \frac{1}{\varepsilon} \int_0^t \theta_{x, \beta}(\bar{m}^\varepsilon(s)) ds.$$

Owing to the identity $\langle \theta_{x, \beta} \rangle = 0$, Lemma I.3.1 and Assumption I.9, the function $\psi_x \doteq -R_0 \theta_{x, \beta}$ is well-defined, is Lipschitz continuous with $[\psi_x]_{\text{Lip}} \lesssim [\theta_{x, \beta}]_{\text{Lip}} = 1$ and $\psi_x, \psi_x^2 \in D(B)$. Therefore Proposition I.4.2 states that

$$\begin{aligned} M_{\varepsilon \psi_x}^\varepsilon(t) &= \varepsilon \psi_x(\bar{m}^\varepsilon(t)) - \varepsilon \psi_x(\bar{m}(0)) - \frac{1}{\varepsilon^2} \int_0^t \varepsilon B \psi_x(\bar{m}^\varepsilon(s)) ds \\ &= \varepsilon \psi_x(\bar{m}^\varepsilon(t)) - \varepsilon \psi_x(\bar{m}(0)) - \frac{\partial^{|\beta|} \zeta^\varepsilon(t, x)}{\partial x^\beta} \end{aligned}$$

defines a square-integrable martingale such that

$$\mathbb{E} \left[\left| M_{\varepsilon \psi_x}^{\varepsilon, \tau^\varepsilon}(t) \right|^2 \right] = \mathbb{E} \left[\int_0^{t \wedge \tau^\varepsilon} (B(\psi_x^2) - 2\psi_x B \psi_x)(\bar{m}^\varepsilon(s)) ds \right]. \quad (\text{I.6.5})$$

Since $[\psi_x]_{\text{Lip}} \lesssim [\theta_{x, \beta}]_{\text{Lip}} = 1$ and $\alpha < 1$ in (I.3.6), we have

$$\mathbb{E} \left[\sup_{t \in [0, T]} |\varepsilon \psi_x(\bar{m}^{\varepsilon, \tau^\varepsilon}(t))|^2 \right] \lesssim \varepsilon^2 \mathbb{E} [\varepsilon^{-2\alpha} \vee \|\bar{m}(0)\|_E^2] \lesssim 1,$$

and by Doob's Maximal Inequality, we get

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t, x)}{\partial x^\beta} \right|^2 \right]^{\frac{1}{2}} &\lesssim 1 + \mathbb{E} \left[\sup_{t \in [0, T]} \left| M_{\varepsilon \psi_x}^{\varepsilon, \tau^\varepsilon}(t) \right|^2 \right]^{\frac{1}{2}} \\ &\lesssim 1 + \mathbb{E} \left[\left| M_{\varepsilon \psi_x}^{\varepsilon, \tau^\varepsilon}(T) \right|^2 \right]^{\frac{1}{2}}. \end{aligned} \quad (\text{I.6.6})$$

Owing to Proposition I.4.2, we have

$$\mathbb{E} \left[\left| M_{\varepsilon \psi_x}^{\varepsilon, \tau^\varepsilon}(T) \right|^2 \right] = \mathbb{E} \left[\int_0^{T \wedge \tau^\varepsilon} (B(\psi_x^2) - 2\psi_x B\psi_x) (\bar{m}^\varepsilon(s)) ds \right],$$

For now, we only know that the right-hand side is of order $\varepsilon^{-2\alpha}$, by (I.2.9) and (I.6.5). To retrieve an estimate uniform in ε , we use the same martingale argument as before. Let

$$\tilde{\theta}_{x, \beta} \doteq B(\psi_x^2) - 2\psi_x B\psi_x = B((R_0 \theta_{x, \beta})^2) + 2\theta_{x, \beta} R_0 \theta_{x, \beta},$$

so that

$$\mathbb{E} \left[\left| M_{\varepsilon \psi_x}^{\varepsilon, \tau^\varepsilon}(T) \right|^2 \right] = \mathbb{E} \left[\int_0^{T \wedge \tau^\varepsilon} \tilde{\theta}_{x, \beta}(\bar{m}^\varepsilon(s)) ds \right]. \quad (\text{I.6.7})$$

Since $\theta_{x, \beta}$ and $R_0 \theta_{x, \beta}$ are pseudo-linear functions, the function $\theta_{x, \beta} R_0 \theta_{x, \beta}$ is pseudo-quadratic. Thus, by Lemma I.3.1 and Assumption I.9, the function

$$\tilde{\psi}_x = (R_0 \theta_{x, \beta})^2 - 2R_0 [\theta_{x, \beta} R_0 \theta_{x, \beta} - \langle \theta_{x, \beta} R_0 \theta_{x, \beta} \rangle],$$

is well-defined and satisfies $\tilde{\psi}_x, \tilde{\psi}_x^2 \in D(B)$ and $B\tilde{\psi}_x = \tilde{\theta}_{x, \beta} - 2\langle \theta_{x, \beta} R_0 \theta_{x, \beta} \rangle$. As before, introduce the martingale process

$$\begin{aligned} M_{\varepsilon^2 \tilde{\psi}_x}^\varepsilon(t) &= \varepsilon^2 \tilde{\psi}_x(\bar{m}^\varepsilon(t)) - \varepsilon^2 \tilde{\psi}_x(\bar{m}(0)) - \frac{1}{\varepsilon^2} \int_0^t \varepsilon^2 B\tilde{\psi}_x(\bar{m}^\varepsilon(s)) ds \\ &= \varepsilon^2 \tilde{\psi}_x(\bar{m}^\varepsilon(t)) - \varepsilon^2 \tilde{\psi}_x(\bar{m}(0)) - \int_0^t \tilde{\theta}_{x, \beta}(\bar{m}^\varepsilon(s)) ds + 2t \langle \theta_{x, \beta} R_0 \theta_{x, \beta} \rangle. \end{aligned}$$

Owing to Lemma I.3.1, we have for $n \in E$

$$\begin{aligned} \left| \tilde{\psi}_x(n) \right| &\lesssim (1 + \|n\|_E^{b+1} + \|n\|_E^2) \\ \langle \theta_{x, \beta} R_0 \theta_{x, \beta} \rangle &\lesssim 1. \end{aligned}$$

Using the conditions $\alpha(b+1) < 2$ and $\alpha < 1$ in (I.3.6), and using the finiteness of moments of order $2(b+1)$ and 4 of $\bar{m}(0)$ in Assumption I.7, we get

$$\mathbb{E} \left[\int_0^{T \wedge \tau^\varepsilon} \tilde{\theta}_{x, \beta}(\bar{m}^\varepsilon(s)) ds \right] \lesssim_T 1 + \mathbb{E} \left[\left| M_{\varepsilon^2 \tilde{\psi}_x}^{\varepsilon, \tau^\varepsilon}(T) \right| \right], \quad (\text{I.6.8})$$

where, owing to Proposition I.4.2,

$$\mathbb{E} \left[\left| M_{\varepsilon^2 \tilde{\psi}_x}^{\varepsilon, \tau^\varepsilon}(T) \right|^2 \right] = \varepsilon^2 \mathbb{E} \left[\int_0^{T \wedge \tau^\varepsilon} (B(\tilde{\psi}_x^2) - 2\tilde{\psi}_x B\tilde{\psi}_x) (\bar{m}^\varepsilon(s)) ds \right].$$

Owing to Assumption I.10, we have for $n \in E$

$$\left| \left(B(\tilde{\psi}_x^2) - 2\tilde{\psi}_x B\tilde{\psi}_x \right) (n) \right| \lesssim (1 + \|n\|_E^{2(b+1)} + \|n\|_E^4).$$

Since $\alpha(b+1) < 1$ and $2\alpha < 1$ in (I.3.6) and since $\bar{m}(0)$ has finite moments of order $2(b+1)$ and 4 in Assumption I.7, we get

$$\mathbb{E} \left[\left| M_{\varepsilon^2 \tilde{\psi}_x}^{\varepsilon, \tau^\varepsilon}(T) \right|^2 \right] \lesssim_T 1. \quad (\text{I.6.9})$$

Gathering the estimates (I.6.6), (I.6.7), (I.6.8) and (I.6.9), we obtain the required result

$$\sup_{\varepsilon} \sup_{x \in \mathbb{T}^d} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t, x)}{\partial x^\beta} \right|^2 \right] \lesssim_T 1.$$

This concludes the proof by (I.6.4). \square

Proposition I.6.2, together with the compact embedding $H_x^s \subset \mathcal{C}_x^1$ and the Markov inequality, proves that (I.6.3) holds, hence (I.6.2). This concludes the proof of Item (i).

I.6.2 Proof of the second claim Item (ii)

As in [DV20], we prove Item (ii) using the Aldous criterion [JS03, Theorem 4.5 p356].

Let $\varphi = \sum_i \varphi_i$ be the sum of a finite number of bounded functions $\varphi_i \in \Theta$. We set $X^\varepsilon = (f^\varepsilon, \zeta^\varepsilon, \bar{m}^\varepsilon)$ and $\bar{X}^\varepsilon = (\rho^\varepsilon, \zeta^\varepsilon)$. Recall that if $\tilde{\varphi} \in \Theta$ depends only on z , then the perturbed test function $\tilde{\varphi}^\varepsilon$ defined by Proposition I.4.3 depends only on n and z . Using Proposition I.4.2, this allows us to stop the processes only at τ^ε instead of τ_Λ^ε while keeping the same estimates. Therefore, the proof of the tightness of $(\tilde{\varphi}(\zeta^{\varepsilon, \tau^\varepsilon}))_\varepsilon$ is the same as of $(\varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}))_\varepsilon$, and is thus omitted. It only remains to prove $(\varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}))_\varepsilon$ is tight.

The Aldous criterion gives a sufficient condition for the tightness of the family $(\varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}))_\varepsilon$ in $D([0, T], \mathbb{R})$: since φ is bounded, it is sufficient to prove that

$$\forall \eta > 0, \lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \sup_{\substack{\tau_1, \tau_2 \leq T \\ \tau_1 \leq \tau_2 \leq \tau_1 + \delta}} \mathbb{P} \left(\left| \varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2)) - \varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1)) \right| > \eta \right) = 0, \quad (\text{I.6.10})$$

where τ_1, τ_2 are any $(\mathcal{F}_t^\varepsilon)_{t \in \mathbb{R}^+}$ -stopping times.

Define the perturbed test function $\varphi^\varepsilon = \sum_i \varphi_i^\varepsilon$. This sum satisfies the estimates (I.4.19), (I.4.20) and (I.4.21). Then, define

$$\theta^\varepsilon(t) = \varphi(\bar{X}^\varepsilon(0)) + \varphi^\varepsilon(X^\varepsilon(t)) - \varphi^\varepsilon(X^\varepsilon(0)) \quad (\text{I.6.11})$$

$$= \varphi(\bar{X}^\varepsilon(0)) + \int_0^t \mathcal{L}^\varepsilon \varphi^\varepsilon(X^\varepsilon(s)) ds + M_{\varphi^\varepsilon}^\varepsilon(t), \quad (\text{I.6.12})$$

where $M_{\varphi^\varepsilon}^\varepsilon$ is defined by Proposition I.4.1, so that

$$\begin{aligned} \varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2)) - \varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1)) &= (\theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1)) - \left(\varphi^\varepsilon(X^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2)) - \varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2)) \right) \\ &\quad + \left(\varphi^\varepsilon(X^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1)) - \varphi(\bar{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1)) \right). \end{aligned}$$

Using (I.3.5) and Propositions I.3.6 and I.4.3, we get

$$\begin{aligned} & \left| \varphi^\varepsilon(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) - \varphi(\overline{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right| \\ & \lesssim_{\varphi, \Lambda} (1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2)(\varepsilon(1 + \varepsilon^{-\alpha} \vee \|\overline{m}(0)\|) + \varepsilon^2(1 + \varepsilon^{-\alpha(b+1)} \vee \|\overline{m}(0)\|^{b+1})). \end{aligned}$$

Since $\alpha < 1$ and $\alpha(b+1) < 2$ in (I.3.6), we get

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left| \varphi^\varepsilon(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) - \varphi(\overline{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right| \right] \xrightarrow{\varepsilon \rightarrow 0} 0,$$

hence,

$$\sup_{\tau_1, \tau_2} \mathbb{E} \left[\left| \varphi(\overline{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2)) - \varphi(\overline{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1)) \right| \right] \leq \sup_{\tau_1, \tau_2} \mathbb{E} \left[\left| \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1) \right| \right] + o(1),$$

where $o(1)$ denotes any quantity u satisfying $u \xrightarrow{\varepsilon \rightarrow 0} 0$. Using the Markov inequality, we get

$$\sup_{\tau_1, \tau_2} \mathbb{P} \left(\left| \varphi(\overline{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2)) - \varphi(\overline{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1)) \right| > \eta \right) \leq \sup_{\tau_1, \tau_2} \frac{\mathbb{E} \left[\left| \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1) \right| \right]}{\eta} + o(1).$$

Therefore, it is sufficient to prove that

$$\sup_{\tau_1, \tau_2, \varepsilon} \mathbb{E} \left[\left| \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1) \right| \right] \xrightarrow{\delta \rightarrow 0} 0, \quad (\text{I.6.13})$$

to deduce (I.6.10) and then to use Aldous criterion.

Owing to (I.6.12), we have

$$\left| \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1) \right| \leq \int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} |\mathcal{L}^\varepsilon \varphi^\varepsilon(X^\varepsilon(s))| ds + \left| M_{\varphi^\varepsilon}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) - M_{\varphi^\varepsilon}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1) \right|. \quad (\text{I.6.14})$$

Using once again (I.3.5), Propositions I.3.6 and I.4.3, we get

$$\begin{aligned} & \left| \mathcal{L}^\varepsilon \varphi^\varepsilon(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right| \lesssim_{\varphi, \Lambda} \left| \mathcal{L} \varphi(\overline{X}^{\varepsilon, \tau_\Lambda^\varepsilon}(s)) \right| + \varepsilon(1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \varepsilon^{-\alpha(b+2)} \vee \|\overline{m}(0)\|^{b+2}) \\ & \lesssim_{\varphi, \Lambda} 1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2 + \varepsilon(1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \varepsilon^{-\alpha(b+2)} \vee \|\overline{m}(0)\|^{b+2}). \end{aligned}$$

Using the Cauchy-Schwarz inequality, the condition $\alpha(b+2) < 1$ in (I.3.6) and Assumptions I.4 and I.7, we get

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left| \mathcal{L}^\varepsilon \varphi^\varepsilon(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right| \right] \lesssim_{\varphi, \Lambda} 1. \quad (\text{I.6.15})$$

Thus, we get

$$\sup_{\varepsilon} \sup_{\tau_1, \tau_2} \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} |\mathcal{L}^\varepsilon \varphi^\varepsilon(X^\varepsilon(s))| ds \right] \leq \sup_{\varepsilon} \sup_{\tau_1, \tau_2} \delta \mathbb{E} \left[\sup_{t \in [0, T]} \left| \mathcal{L}^\varepsilon \varphi^\varepsilon(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right| \right] \xrightarrow{\delta \rightarrow 0} 0.$$

The last term of (I.6.14) is controlled using martingale arguments. Owing to Proposition I.4.1, $M_{\varphi^\varepsilon}^{\varepsilon, \tau_\Lambda^\varepsilon}$ is indeed a square-integrable martingale and

$$\begin{aligned} \mathbb{E} \left[\left| M_{\varphi^\varepsilon}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) - M_{\varphi^\varepsilon}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1) \right|^2 \right] &= \mathbb{E} \left[\left| M_{\varphi^\varepsilon}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_2) \right|^2 - \left| M_{\varphi^\varepsilon}^{\varepsilon, \tau_\Lambda^\varepsilon}(\tau_1) \right|^2 \right] \\ &= \frac{1}{\varepsilon^2} \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} (B((\varphi^\varepsilon)^2) - 2\varphi^\varepsilon B\varphi^\varepsilon)(X^\varepsilon(s)) ds \right] \\ &= \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} \sum_{i=-2}^2 \varepsilon^i r_i(X^\varepsilon(s)) ds \right] \end{aligned}$$

where the terms r_i are obtained by writing $\varphi^\varepsilon = \varphi + \varepsilon\varphi_1 + \varepsilon^2\varphi_2$ and expanding $B((\varphi^\varepsilon)^2) - 2\varphi^\varepsilon B\varphi^\varepsilon$. The terms containing φ vanish, using $B\varphi = 0$, $B(\varphi^2) = 0$ and $B\varphi\varphi_j = \varphi B\varphi_j$ (since φ does not depend on n). Using Assumption I.10, the remaining terms satisfy

$$\begin{aligned} r_{-2} &= r_{-1} = 0, \\ r_0(f, z, n) &= [B(\varphi_1^2) - 2\varphi_1 B\varphi_1](f, z, n) \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|n\|_E^2), \\ r_1(f, z, n) &= [2B(\varphi_1\varphi_2) - \varphi_1 B\varphi_2 - \varphi_2 B\varphi_1](f, z, n) \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|n\|_E^{b+2}), \\ r_2(f, z, n) &= [B(\varphi_2^2) - 2\varphi_2 B\varphi_2](f, z, n) \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^4)(1 + \|n\|_E^{2(b+1)}). \end{aligned}$$

As for (I.6.15), using that $\alpha(b+2) < 1$ in (I.3.6), we have for $i \in \{1, 2\}$

$$\mathbb{E} \left[\sup_{t \in [0, T]} \varepsilon^i r_i(X^{\varepsilon, \tau_\Lambda^\varepsilon}(t)) \right] \lesssim_{\varphi, \Lambda} 1,$$

and

$$\sup_{\varepsilon} \sup_{\tau_1, \tau_2} \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} \varepsilon^i r_i(X^\varepsilon(s)) ds \right] \xrightarrow{\delta \rightarrow 0} 0.$$

We need to be more cautious when dealing with r_0 , since there are no ε left to compensate the $\varepsilon^{-2\alpha}$ that would appear from bounding $\overline{m}^{\varepsilon, \tau_\Lambda^\varepsilon}$ from above using Proposition I.3.6. The idea is to use estimates for $f^{\varepsilon, \tau_\Lambda^\varepsilon}$ and \overline{m}^ε (instead of $\overline{m}^{\varepsilon, \tau_\Lambda^\varepsilon}$), using that for $s \leq \tau_\Lambda^\varepsilon$, $\overline{m}^{\varepsilon, \tau_\Lambda^\varepsilon}(s) = \overline{m}^\varepsilon(s)$. We write

$$\begin{aligned} \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} r_0(X^\varepsilon(s)) ds \right] &\lesssim_\varphi \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} (1 + \|f^{\varepsilon, \tau_\Lambda^\varepsilon}(s)\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|\overline{m}^\varepsilon(s)\|_E^2) ds \right] \\ &\lesssim_{\varphi, \Lambda} \mathbb{E} \left[\int_{\tau_1}^{\tau_2} (1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|\overline{m}^\varepsilon(s)\|_E^2) ds \right] \\ &\lesssim_{\varphi, \Lambda} \int_0^T \mathbb{E} \left[\mathbb{1}_{[\tau_1, \tau_2]}(s) (1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|\overline{m}^\varepsilon(s)\|_E^2) \right] ds. \end{aligned}$$

Then, we use the Hölder inequality to write

$$\begin{aligned} \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} r_0(X^\varepsilon(s)) ds \right] &\lesssim_{\varphi, \Lambda} \int_0^T \mathbb{E} \left[\mathbb{1}_{[\tau_1, \tau_2]}(s) \right]^{\frac{1}{3}} \mathbb{E} \left[1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^6 \right]^{\frac{1}{3}} \mathbb{E} \left[1 + \|\overline{m}^\varepsilon(s)\|_E^6 \right]^{\frac{1}{3}} ds \\ &\lesssim_{\varphi, \Lambda} \int_0^T \mathbb{E} \left[\mathbb{1}_{[\tau_1, \tau_2]}(s) \right]^{\frac{1}{3}} ds \mathbb{E} \left[1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^6 \right]^{\frac{1}{3}} \mathbb{E} \left[1 + \|\overline{m}^\varepsilon(0)\|_E^6 \right]^{\frac{1}{3}}, \end{aligned}$$

by stationarity of \overline{m} . Using the Cauchy-Schwarz inequality, Assumptions I.4 and I.7, we get

$$\begin{aligned} \mathbb{E} \left[\int_{\tau_1 \wedge \tau_\Lambda^\varepsilon}^{\tau_2 \wedge \tau_\Lambda^\varepsilon} r_0(X^\varepsilon(s)) ds \right] &\lesssim_{\varphi, \Lambda} \int_0^T \mathbb{E} [\mathbb{1}_{[\tau_1, \tau_2]}(s)]^{\frac{1}{2}} ds \\ &\lesssim_{\varphi, \Lambda, T} \left(\int_0^T \mathbb{E} [\mathbb{1}_{[\tau_1, \tau_2]}(s)] ds \right)^{\frac{1}{2}} \\ &\lesssim_{\varphi, \Lambda, T} \delta^{\frac{1}{2}} \rightarrow 0, \end{aligned}$$

uniformly in ε , τ_1 and τ_2 . This concludes the proof of (I.6.13).

We are now in position to apply Aldous' criterion, which proves that $((\rho^{\varepsilon, \tau_\Lambda(\zeta^\varepsilon)}, \zeta^{\varepsilon, \tau_\Lambda(\zeta^\varepsilon)}))_\varepsilon$ is tight in $\mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1$. This concludes the proof of Item (ii), and of Proposition I.6.1.

I.7 Identification of the limit points

In this section, we establish the first convergence result stated in Theorem I.2.1.

We start by proving the convergence of the auxiliary process ζ^ε in Section I.7.1, using the convergence of a simplified martingale problem. Then, in Section I.7.2, we determine the stopped martingale problem solved by a limit point of the stopped process. In Section I.7.3, we use this stopped martingale to identify the limit point of the stopped process. We conclude on the convergence of the unstopped process in Section I.7.4.

I.7.1 Convergence of the auxiliary process

Proving the convergence of ζ^ε is much simpler than for the coupled process \overline{X}^ε . Indeed, as seen in particular in Proposition I.6.1, the only stopping time we need is τ^ε , and $\tau^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} +\infty$. Therefore, the convergence of martingale problems is a little more intricate than the proof used in [DV12], but it remains straightforward.

Proposition I.7.1. *The process ζ^ε converges in distribution in $\mathcal{C}_T^0 \mathcal{C}_x^1$ to a Wiener process of covariance Q when $\varepsilon \rightarrow 0$.*

Proof. Owing to the tightness established in Proposition I.6.1, there exists a sequence $\varepsilon_i \xrightarrow{i \rightarrow \infty} 0$ and $\zeta \in \mathcal{C}_T^0 \mathcal{C}_x^1$ such that ζ^{ε_i} converges in distribution to ζ when $i \rightarrow \infty$. We start by proving that ζ solves the martingale problem associated with the generator \mathcal{L} .

Let $\varphi \in \Theta$ with $\psi = 1$. Let $0 \leq s \leq s_1 \leq \dots \leq s_n \leq t$, let g be a continuous bounded function and for $z \in \mathcal{C}_T^0 \mathcal{C}_x^1$, let $G(z) = g(z(s_1), \dots, z(s_n))$ and

$$\Phi(z) = \left(\varphi(z(t)) - \varphi(z(s)) - \int_s^t \mathcal{L}\varphi(z(u)) du \right) G(z).$$

Note that G and Φ are continuous and bounded on $\mathcal{C}_T^0 \mathcal{C}_x^1$, so $\mathbb{E}[\Phi(\zeta^{\varepsilon_i})] \xrightarrow{i \rightarrow \infty} \mathbb{E}[\Phi(\zeta)]$. Let us establish that $\mathbb{E}[\Phi(\zeta^{\varepsilon_i})]$ also converges to 0.

Let φ^{ε_i} be the perturbed test function introduced in Proposition I.4.3 associated to φ . Since φ^{ε_i} is a good test function, and since $G(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}})$ is $\mathcal{F}_s^{\varepsilon_i}$ -measurable, Proposition I.4.1 yields

$$\mathbb{E} \left[\left(\varphi^{\varepsilon_i}(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(t)) - \varphi^{\varepsilon_i}(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(s)) - \int_{s \wedge \tau^{\varepsilon_i}}^{t \wedge \tau^{\varepsilon_i}} \mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(\zeta^{\varepsilon_i}(u)) du \right) G(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}) \right] = 0.$$

Owing to (I.4.13), this leads to

$$|\mathbb{E}[\Phi(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}})]| \lesssim_g \sum_{j=1}^4 \mathbb{E}[|r_j|],$$

with

$$\begin{aligned} r_1 &= \varepsilon_i(\varphi_1(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(t), \overline{m}^{\varepsilon_i, \tau^{\varepsilon_i}}(t)) - \varphi_1(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(s), \overline{m}^{\varepsilon_i, \tau^{\varepsilon_i}}(s))), \\ r_2 &= \varepsilon_i^2(\varphi_2(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(t), \overline{m}^{\varepsilon_i, \tau^{\varepsilon_i}}(t)) - \varphi_2(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(s), \overline{m}^{\varepsilon_i, \tau^{\varepsilon_i}}(s))), \\ r_3 &= - \int_{s \wedge \tau^{\varepsilon_i}}^{t \wedge \tau^{\varepsilon_i}} (\mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(\zeta^{\varepsilon_i}(u)) - \mathcal{L} \varphi(\zeta^{\varepsilon_i}(u))) du, \\ r_4 &= \int_{t \wedge \tau^{\varepsilon_i}}^t \mathcal{L} \varphi(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(u)) du - \int_{s \wedge \tau^{\varepsilon_i}}^s \mathcal{L} \varphi(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}(u)) du. \end{aligned}$$

Using (I.3.6), (I.4.19), (I.4.20), (I.4.21) and Assumptions I.4 and I.7 we have for $j \in \{1, 2, 3\}$, $\mathbb{E}[|r_j|] \xrightarrow{\varepsilon \rightarrow 0} 0$. It remains to prove that $\mathbb{E}[|r_4|] \rightarrow 0$. The term r_4 does not appear in [DV12], but is simple to manage since $\tau^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \infty$. The Cauchy-Schwarz inequality and Lemma I.3.4 lead to

$$\begin{aligned} \mathbb{E}[|r_4|]^2 &\lesssim_\varphi \mathbb{E}[|t - t \wedge \tau^{\varepsilon_i}|^2 + |s - s \wedge \tau^{\varepsilon_i}|^2] \\ &\lesssim_\varphi T^2 \mathbb{P}(\tau^{\varepsilon_i} < T) \xrightarrow{i \rightarrow \infty} 0 \end{aligned}$$

Thus, we get $\mathbb{E}[\Phi(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}})] \xrightarrow{i \rightarrow \infty} 0$, hence $\mathbb{E}[\Phi(\zeta)] = 0$. The same proof can be adapted when replacing φ by φ^2 . Therefore, the processes M_φ and M_{φ^2} defined in Proposition I.5.1 are martingales. Owing to Proposition I.5.1, ζ satisfies (I.5.2) and is a Q -Wiener process.

This limit point being unique in distribution, ζ^ε converges in distribution to this Wiener process. \square

I.7.2 Convergence of the stopped martingale problems

In this section, we use Proposition I.7.1 to establish the convergence of the stopped martingale problems satisfied by $X^{\varepsilon, \tau^\varepsilon}$. The proof is similar to the proof of Proposition I.7.1, but this time the stopping time persists when $\varepsilon \rightarrow 0$ because of the fixed threshold Λ .

Let us introduce the path space $\underline{\Omega} = \mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1 \times \mathcal{C}_T^0 \mathcal{C}_x^1$, equipped with its Borel σ -algebra. We denote by $(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')$ the canonical process on $\underline{\Omega}$ and by $(\underline{\mathcal{F}}_t)_{t \in \mathbb{R}^+}$ its associated filtration.

Define $\mathbb{P}_{\varepsilon, \Lambda}$ the distribution of $(\rho^{\varepsilon, \tau^\varepsilon}, \zeta^{\varepsilon, \tau^\varepsilon}, \zeta^\varepsilon)$ and $\mathbb{E}_{\varepsilon, \Lambda}$ the expectation under this distribution (on $\underline{\Omega}$). By Proposition I.6.1, the family $(\mathbb{P}_{\varepsilon, \Lambda})_\varepsilon$ is tight. Thus, in this section, we consider a sequence $(\varepsilon_i)_{i \in \mathbb{N}}$ such that $\varepsilon_i \rightarrow 0$ and $\mathbb{P}_{\varepsilon_i, \Lambda} \rightarrow \mathbb{P}_{0, \Lambda}$ weakly when $i \rightarrow \infty$, for some limit point $\mathbb{P}_{0, \Lambda}$. Note that under $\mathbb{P}_{0, \Lambda}$, owing to Proposition I.7.1, $\underline{\zeta}'$ is a Q -Wiener process whose distribution \mathbb{P}_Q does not depend on Λ .

We now state two continuity lemmas.

Lemma I.7.2. *For any fixed $\Lambda \in \mathbb{R}^+$, the mapping $\tau_\Lambda(\cdot)$ defined by (I.2.10) is lower semi-continuous on $\mathcal{C}_T^0 \mathcal{C}_x^1$. Moreover, it is continuous at every z such that $\tau(z)$ is continuous at Λ .*

Lemma I.7.3. *The set $\{\Lambda \geq 0 \mid \mathbb{P}_Q(\tau(\underline{\zeta}'))$ is not continuous at $\Lambda) > 0\}$ is at most countable. Let \mathfrak{L} be its complementary.*

We refer to [HS12, Lemmas 3.5 and 3.6] for the proofs of Lemmas I.7.2 and I.7.3. These results can be applied here since $\|\underline{\zeta}'\|_{\mathcal{C}_x^1}$ is a continuous finite dimensional process and its distribution \mathbb{P}_Q under $\mathbb{P}_{0,\Lambda}$ does not depend on Λ .

Owing to Lemma I.7.3, there exist arbitrarily large numbers $\Lambda \in \mathfrak{L}$ and for all $\Lambda \in \mathfrak{L}$, $\tau(\underline{\zeta}')$ is $\mathbb{P}_{0,\Lambda}$ -almost surely continuous at Λ and by Lemma I.7.2, $\tau_\Lambda(\cdot)$ is $\mathbb{P}_{0,\Lambda}$ -a.s. continuous at $\underline{\zeta}'$. From now on, it is assumed that $\Lambda \in \mathfrak{L}$.

Proposition I.7.4. *Let $\Lambda \in \mathfrak{L}$. For all $\varphi \in \Theta$, the process*

$$t \mapsto \varphi(\underline{\rho}(t), \underline{\zeta}(t)) - \varphi(\underline{\rho}(0), \underline{\zeta}(0)) - \int_0^{t \wedge \tau_\Lambda(\underline{\zeta}')} \mathcal{L}\varphi(\underline{\rho}(u), \underline{\zeta}(u)) du$$

is a $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ -martingale under $\mathbb{P}_{0,\Lambda}$.

Proof. Let $\varphi \in \Theta$. As for Proposition I.7.1, let $0 \leq s \leq s_1 \leq \dots \leq s_n \leq t$, let g be a continuous bounded function, and let

$$G(\underline{\rho}, \underline{\zeta}, \underline{\zeta}') = g(\underline{\rho}(s_1), \underline{\zeta}(s_1), \underline{\zeta}'(s_1), \dots, \underline{\rho}(s_n), \underline{\zeta}(s_n), \underline{\zeta}'(s_n)),$$

and

$$\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}') = \left(\varphi(\underline{\rho}(t), \underline{\zeta}(t)) - \varphi(\underline{\rho}(s), \underline{\zeta}(s)) - \int_{s \wedge \tau_\Lambda(\underline{\zeta}')}^{t \wedge \tau_\Lambda(\underline{\zeta}')} \mathcal{L}\varphi(\underline{\rho}(u), \underline{\zeta}(u)) du \right) G(\underline{\rho}, \underline{\zeta}, \underline{\zeta}').$$

As for Proposition I.7.1, we establish that $\mathbb{E}_{\varepsilon_i, \Lambda}[\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')]$ converges, when $i \rightarrow \infty$, to both $\mathbb{E}_{0, \Lambda}[\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')] and 0.$

On the one hand, since Φ is continuous $\mathbb{P}_{0,\Lambda}$ -almost everywhere, $\mathbb{P}_{\varepsilon_i, \Lambda} \circ \Phi^{-1} \rightarrow \mathbb{P}_{0, \Lambda} \circ \Phi^{-1}$ weakly when $i \rightarrow \infty$ (see [Bou04, Proposition IX.5.7]). Moreover, $(\mathbb{P}_{\varepsilon_i, \Lambda} \circ \Phi^{-1})_{\varepsilon_i}$ is uniformly integrable. Indeed, using (I.4.28), we have

$$\sup_{\varepsilon} \mathbb{E}_{\varepsilon, \Lambda} \left[|\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')|^2 \right] \lesssim_{T, \Lambda, \varphi, g} \sup_{\varepsilon} \mathbb{E} \left[1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^4 \right] < \infty.$$

Owing to [Bil99, Theorem 5.4], uniform integrability and convergence in distribution yield

$$\mathbb{E}_{\varepsilon_i, \Lambda}[\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')] \xrightarrow{i \rightarrow \infty} \mathbb{E}_{0, \Lambda}[\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')] .$$

On the other hand, define the perturbed test function φ^{ε_i} as in Proposition I.4.3. As for Proposition I.7.1, we have

$$\mathbb{E} \left[\left(\varphi^{\varepsilon_i}(X^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(t)) - \varphi^{\varepsilon_i}(X^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(s)) - \int_{s \wedge \tau_\Lambda^{\varepsilon_i}}^{t \wedge \tau_\Lambda^{\varepsilon_i}} \mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(X^{\varepsilon_i}(u)) du \right) G(\rho^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}, \zeta^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}, \zeta^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}) \right] = 0,$$

and

$$|\mathbb{E}_{\varepsilon_i, \Lambda}[\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')]| = \left| \mathbb{E} \left[\Phi(\rho^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}, \zeta^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}, \zeta^{\varepsilon_i}) \right] \right| \lesssim_g \sum_{j=1}^4 \mathbb{E}[|r_j|],$$

with

$$\begin{aligned}
r_1 &= \varepsilon_i (\varphi_1(X^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(t)) - \varphi_1(X^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(s))) \rightarrow 0 \\
r_2 &= \varepsilon_i^2 (\varphi_2(X^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(t)) - \varphi_2(X^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(s))) \rightarrow 0 \\
r_3 &= - \int_{s \wedge \tau_\Lambda^{\varepsilon_i}}^{t \wedge \tau_\Lambda^{\varepsilon_i}} \left(\mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(X^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(u)) - \mathcal{L} \varphi(\overline{X}^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}(u)) \right) du \rightarrow 0 \\
r_4 &= \int_{t \wedge \tau_\Lambda^{\varepsilon_i}}^{t \wedge \tau_\Lambda(\zeta^{\varepsilon_i})} \mathcal{L} \varphi(\overline{X}^{\varepsilon_i, \tau^{\varepsilon_i}}(u)) du - \int_{s \wedge \tau_\Lambda^{\varepsilon_i}}^{s \wedge \tau_\Lambda(\zeta^{\varepsilon_i})} \mathcal{L} \varphi(\overline{X}^{\varepsilon_i, \tau^{\varepsilon_i}}(u)) du.
\end{aligned}$$

For the last term r_4 , we have

$$\begin{aligned}
\mathbb{E}[|r_4|]^2 &\lesssim_{\varphi, \Lambda} \mathbb{E}[|t \wedge \tau_\Lambda(\zeta^{\varepsilon_i}) - t \wedge \tau_\Lambda^{\varepsilon_i}|^2 + |s \wedge \tau_\Lambda(\zeta^{\varepsilon_i}) - s \wedge \tau_\Lambda^{\varepsilon_i}|^2] \\
&\lesssim_{\varphi, \Lambda} T^2 \mathbb{P}(\tau^{\varepsilon_i} < T \wedge \tau_\Lambda(\zeta^{\varepsilon_i})) \text{ using (I.2.12)} \\
&\lesssim_{\varphi, \Lambda} T^2 \mathbb{P}(\tau^{\varepsilon_i} < T) \xrightarrow{i \rightarrow \infty} 0.
\end{aligned}$$

Thus, we get $\mathbb{E}_{\varepsilon_i, \Lambda}[\Phi(\underline{\rho}, \underline{\zeta}, \underline{\zeta}')] \xrightarrow{i \rightarrow \infty} 0$, which concludes the proof of Proposition I.7.4. \square

I.7.3 Identification of the limit point

In Section I.7.1, solving the martingale problem is sufficient to characterize the distribution of the Markov process as a solution of a limit equation, under a uniqueness condition. However, the limit point $\mathbb{P}_{0, \Lambda}$ solves a martingale problem only until a stopping time $\tau_\Lambda(\zeta')$. The goal of this section is to explain how to identify $\mathbb{P}_{0, \Lambda}$ using this stopped martingale problem.

Let us come back to the space Ω to state more precise results. Recall that the distribution of $(\rho^{\varepsilon, \tau_\Lambda^{\varepsilon}}, \zeta^{\varepsilon, \tau_\Lambda^{\varepsilon}}, \zeta^{\varepsilon})$ is $\mathbb{P}_{\varepsilon, \Lambda}$, and define $(\rho_\Lambda, \zeta_\Lambda, \zeta')$ following the limit distribution $\mathbb{P}_{0, \Lambda}$ (we assume Ω is large enough to define such a process). Recall that $\overline{X}^{\varepsilon, \tau_\Lambda^{\varepsilon}} \doteq (\rho^{\varepsilon, \tau_\Lambda^{\varepsilon}}, \zeta^{\varepsilon, \tau_\Lambda^{\varepsilon}})$. Define $\overline{X}_\Lambda \doteq (\rho_\Lambda, \zeta_\Lambda)$ and \overline{X} a solution of (I.2.13).

In this section, we construct a process Y_Λ that extends \overline{X}_Λ after the stopping time $\tau_\Lambda(\zeta')$ (in distribution) and that solves the martingale problem associated to \mathcal{L} . It is similar to the proof of [SV06, Theorem 6.1.2], but we adapt this proof to see precisely how $\tau_\Lambda(\zeta')$ is linked to the extended process.

Extension after a stopping time We first need a result to assert that $\tau_\Lambda(\zeta')$ is a hitting time for \overline{X}_Λ . Note that until here, we did not use ζ^{ε} when considering the coupled process $(\rho^{\varepsilon}, \zeta^{\varepsilon})$. But had we considered ρ^{ε} alone, the stopping time $\tau_\Lambda(\zeta')$ would not be a hitting time for ρ_Λ (as a matter of fact, $\tau_\Lambda(\zeta')$ is not even a stopping time for the filtration generated by ρ_Λ).

Lemma I.7.5. *Let $\Lambda \in \mathfrak{L}$.*

The processes ζ_Λ and $(\zeta')^{\tau_\Lambda(\zeta')}$ are indistinguishable. In particular, $\tau_\Lambda(\zeta_\Lambda) = \tau_\Lambda(\zeta')$. Moreover, the processes ρ_Λ and $\rho_\Lambda^{\tau_\Lambda(\zeta')}$ are indistinguishable.

This result was expected, given the construction of the stopping times and the fact that ζ_Λ and ζ' are the limit of the same process, respectively with and without a stopping time. The choice $\Lambda \in \mathfrak{L}$ is here necessary to retrieve this result by taking the limit $\varepsilon \rightarrow 0$.

Proof. Since $\tau^{\varepsilon_i} \rightarrow \infty$ in probability by Lemma I.3.4, Slutsky's Lemma yields the following convergence in distribution

$$(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}, \zeta^{\varepsilon_i, \tau^{\varepsilon_i}}, \zeta^{\varepsilon_i}) \xrightarrow{i \rightarrow \infty} (\zeta_\Lambda, \zeta', \zeta')$$

Now, for $z_1, z_2, z_3 \in \mathcal{C}_T^0 \mathcal{C}_x^1$, let

$$\Phi(z_1, z_2, z_3) = \left\| z_1 - z_2^{\tau_\Lambda(z_3)} \right\|_{\mathcal{C}_T^0 \mathcal{C}_x^1}.$$

Owing to Lemma I.7.2, the mapping Φ is almost surely continuous at $(\zeta_\Lambda, \zeta', \zeta')$. Thus $\Phi(\zeta^{\varepsilon_i, \tau^{\varepsilon_i}}, \zeta^{\varepsilon_i, \tau^{\varepsilon_i}}, \zeta^{\varepsilon_i}) = 0$ converges in distribution to $\Phi(\zeta_\Lambda, \zeta', \zeta')$. Hence, we have almost surely $\zeta_\Lambda = (\zeta')^{\tau_\Lambda(\zeta')}$.

The proof for ρ_Λ uses similar arguments with $\Phi(\rho, z) = \left\| \rho - \rho^{\tau_\Lambda(z)} \right\|_{\mathcal{C}_T^0 H_x^{-\sigma}}$. \square

From now on, for any process $Y = (\rho, \zeta)$, we write $\tau_\Lambda(Y) = \tau_\Lambda(\zeta)$ so that $\tau_\Lambda(\overline{X}_\Lambda) = \tau_\Lambda(\zeta') \in [0, \infty]$. We shorten the notation to $\tau_\Lambda \doteq \tau_\Lambda(\overline{X}_\Lambda)$. Introduce the measurable function S_Λ that stops a process at the level Λ , namely $S_\Lambda(Y) = Y^{\tau_\Lambda(Y)}$. Owing to Lemma I.7.5, we have $S_\Lambda(\overline{X}_\Lambda) = \overline{X}_\Lambda$.

This section is devoted to *extending* \overline{X}_Λ after τ_Λ into a solution of the martingale problem associated to \mathcal{L} . Namely, we define a process Y_Λ such that $S_\Lambda(Y_\Lambda) \stackrel{d}{=} \overline{X}_\Lambda$ and such that Y_Λ solves the aforementioned martingale problem.

Fix $\omega' \in \Omega$. Define the process $\overline{X}_{\Lambda, \omega'}$ as follows:

- $\forall \omega \in \Omega, \forall t \leq \tau_\Lambda(\omega'), \overline{X}_{\Lambda, \omega'}(t)(\omega) = \overline{X}_\Lambda(t)(\omega')$. Note that $\tau_\Lambda(\overline{X}_{\Lambda, \omega'}) = \tau_\Lambda(\omega')$ almost surely. In particular, the distribution of $S_\Lambda(\overline{X}_{\Lambda, \omega'})$ is the Dirac distribution at $\overline{X}_\Lambda(\omega')$.
- On $[\tau_\Lambda(\omega'), T]$ (this interval can be empty), $\overline{X}_{\Lambda, \omega'}(\omega)$ is the solution of (I.2.13) starting at time $\tau_\Lambda(\omega')$ from the initial state $\overline{X}_\Lambda(\tau_\Lambda(\omega'))(\omega')$.

It is straightforward to check that

$$\omega' \mapsto \mathbb{P}(\overline{X}_{\Lambda, \omega'} \in \mathcal{C})$$

is measurable for $\mathcal{C} = \{Y \in \mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1 \mid Y(t_1) \in \Gamma_1, \dots, Y(t_n) \in \Gamma_n\}$ with $0 \leq t_1 < \dots < t_n \leq T$ and Γ_i measurable. Since those sets generate the Borel σ -algebra of $\mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1$, and since a pointwise limit of measurable functions is measurable, we can take the limit when the subdivision become thinner to get that the mapping is still measurable for any measurable \mathcal{C} . Thus, we can define a mapping $\mathcal{C} \mapsto \mathbb{E}' \mathbb{P}(\overline{X}_{\Lambda, \omega'} \in \mathcal{C})$, where \mathbb{E}' denotes the integration with respect to ω' . It is also straightforward to check that this mapping is a probability measure, thus we can define on Ω a process Y_Λ following this distribution, namely

$$\mathbb{P}(Y_\Lambda \in \mathcal{C}) = \mathbb{E}' \mathbb{P}(\overline{X}_{\Lambda, \omega'} \in \mathcal{C}).$$

In particular, since $S_\Lambda^{-1}(\mathcal{C})$ is a measurable set, we have

$$\begin{aligned} \mathbb{P}(S_\Lambda(Y_\Lambda) \in \mathcal{C}) &= \mathbb{E}' \mathbb{P}(S_\Lambda(\overline{X}_{\Lambda, \omega'}) \in \mathcal{C}) \\ &= \mathbb{E}' \mathbb{1}_{\{\overline{X}_\Lambda(\omega') \in \mathcal{C}\}} \\ &= \mathbb{P}(\overline{X}_\Lambda \in \mathcal{C}), \end{aligned}$$

hence Y_Λ extends \overline{X}_Λ as announced beforehand, in the sense that $S_\Lambda(Y_\Lambda) \stackrel{d}{=} \overline{X}_\Lambda$. Moreover, for any measurable function Φ such that $\mathbb{E}' \mathbb{E}[|\Phi(\overline{X}_{\Lambda, \omega'})|] < \infty$, we have

$$\mathbb{E}[\Phi(Y_\Lambda)] = \mathbb{E}' \mathbb{E}[\Phi(\overline{X}_{\Lambda, \omega'})]. \quad (\text{I.7.1})$$

Identification of the extended process It remains to prove that Y_Λ solves the martingale problem associated to \mathcal{L} .

For $\varphi \in \Theta$, and a process $Y \in \mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1$, define the process

$$M^Y(t) = \varphi(Y(t)) - \varphi(Y(0)) - \int_0^t \mathcal{L}\varphi(Y(u))du.$$

Let $0 \leq s_1 \leq \dots \leq s_n \leq s < t$ and g be a bounded measurable function. Let $G : Y \mapsto g(Y(s_1), \dots, Y(s_n))$.

Owing to Proposition I.5.1, for almost all $\omega' \in \Omega$, the process

$$N_{\Lambda, \omega'}(t) = M^{\bar{X}_{\Lambda, \omega'}}(t) - M^{\bar{X}_{\Lambda, \omega'}}(t \wedge \tau_\Lambda(\omega'))$$

satisfies the martingale property

$$\mathbb{E} [N_{\Lambda, \omega'}(t)G(\bar{X}_{\Lambda, \omega'})] = \mathbb{E} [N_{\Lambda, \omega'}(s)G(\bar{X}_{\Lambda, \omega'})].$$

Indeed, for $t \in [0, \tau(\omega')]$, $N_{\Lambda, \omega'}(t) = 0$ and after the time $\tau(\omega')$, this process solves the martingale problem starting at time $\tau(\omega')$ by construction. Using (I.5.4) and that φ and $\mathcal{L}\varphi$ have at most quadratic growth, it is straightforward to establish

$$\mathbb{E}' \mathbb{E} [|N_{\Lambda, \omega'}(t)G(\bar{X}_{\Lambda, \omega'})|] < \infty.$$

Thus, (I.7.1) and the identity above yield

$$\mathbb{E} [(M^{Y_\Lambda}(t) - M^{Y_\Lambda}(t \wedge \tau_\Lambda(Y_\Lambda)))G(Y_\Lambda)] = \mathbb{E} [(M^{Y_\Lambda}(s) - M^{Y_\Lambda}(s \wedge \tau_\Lambda(Y_\Lambda)))G(Y_\Lambda)],$$

which can be rewritten as

$$\begin{aligned} \mathbb{E} [M^{Y_\Lambda}(t)G(Y_\Lambda)] &= \mathbb{E} [M^{Y_\Lambda}(s)\mathbb{1}_{\{\tau_\Lambda(Y_\Lambda) \leq s\}}G(Y_\Lambda)] \\ &\quad + \mathbb{E} [M^{Y_\Lambda}(t \wedge \tau_\Lambda(Y_\Lambda))\mathbb{1}_{\{\tau_\Lambda(Y_\Lambda) > s\}}G(Y_\Lambda)]. \end{aligned} \quad (\text{I.7.2})$$

Using that the process Y_Λ and $S_\Lambda(Y_\Lambda)$ are equal until the time $\tau_\Lambda(Y_\Lambda) = \tau_\Lambda(S_\Lambda(Y_\Lambda))$, and that $S_\Lambda(Y_\Lambda)$ and \bar{X}_Λ are equal in distribution, we get for the second term

$$\begin{aligned} \mathbb{E} [M^{Y_\Lambda}(t \wedge \tau_\Lambda(Y_\Lambda))\mathbb{1}_{\{\tau_\Lambda(Y_\Lambda) > s\}}G(Y_\Lambda)] \\ &= \mathbb{E} [M^{S_\Lambda(Y_\Lambda)}(t \wedge \tau_\Lambda(S_\Lambda(Y_\Lambda)))\mathbb{1}_{\{\tau_\Lambda(S_\Lambda(Y_\Lambda)) > s\}}G(S_\Lambda(Y_\Lambda))] \\ &= \mathbb{E} [M^{\bar{X}_\Lambda}(t \wedge \tau_\Lambda)\mathbb{1}_{\{\tau_\Lambda > s\}}G(\bar{X}_\Lambda)]. \end{aligned}$$

Owing to Proposition I.7.4, $t \mapsto M^{\bar{X}_\Lambda}(t \wedge \tau_\Lambda)$ is a martingale for the filtration $\mathcal{F}^{\bar{X}_\Lambda}$ generated by \bar{X}_Λ . Moreover $\mathbb{1}_{\{\tau_\Lambda > s\}}G(\bar{X}_\Lambda)$ is $\mathcal{F}_s^{\bar{X}_\Lambda}$ -measurable, hence the martingale property yields

$$\mathbb{E} [M^{\bar{X}_\Lambda}(t \wedge \tau_\Lambda)\mathbb{1}_{\{\tau_\Lambda > s\}}G(\bar{X}_\Lambda)] = \mathbb{E} [M^{\bar{X}_\Lambda}(s)\mathbb{1}_{\{\tau_\Lambda > s\}}G(\bar{X}_\Lambda)].$$

Using again that $S_\Lambda(Y_\Lambda) \stackrel{d}{=} \bar{X}_\Lambda$, we get

$$\mathbb{E} [M^{\bar{X}_\Lambda}(s)\mathbb{1}_{\{\tau_\Lambda > s\}}G(\bar{X}_\Lambda)] = \mathbb{E} [M^{Y_\Lambda}(s)\mathbb{1}_{\{\tau_\Lambda(Y_\Lambda) > s\}}G(Y_\Lambda)].$$

Finally, owing to (I.7.2), we have

$$\begin{aligned}\mathbb{E} [M^{Y_\Lambda}(t)G(Y_\Lambda)] &= \mathbb{E} [M^{Y_\Lambda}(s)\mathbb{1}_{\{\tau_\Lambda(Y_\Lambda) \leq s\}}G(Y_\Lambda)] + \mathbb{E} [M^{Y_\Lambda}(s)\mathbb{1}_{\{\tau_\Lambda(Y_\Lambda) > s\}}G(Y_\Lambda)] \\ &= \mathbb{E} [M^{Y_\Lambda}(s)G(Y_\Lambda)],\end{aligned}$$

which proves that Y_Λ solves the martingale associated to \mathcal{L} . Owing to Proposition I.5.1, it solves (I.2.13) and since the solution is unique $Y_\Lambda \stackrel{d}{=} \bar{X}$ the solution of (I.2.13). Therefore, the limit point is unique (and does not depend on Λ). This concludes the proof that $X^{\varepsilon, \tau_\Lambda^\varepsilon}$ converges in distribution to \bar{X} .

I.7.4 Convergence of the unstopped process

This section is devoted to the proof that the process $\bar{X}^\varepsilon \doteq (\rho^\varepsilon, \zeta^\varepsilon)$ converges in distribution to $\bar{X} \doteq (\rho, \zeta)$ solution of (I.2.13), in $\mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1$.

Let Φ be a continuous bounded mapping from $\mathcal{C}_T^0 H_x^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1$ to \mathbb{R} . There exists a sequence ε_i such that $\varepsilon_i \rightarrow 0$ when $i \rightarrow \infty$ and

$$\limsup_{\varepsilon \rightarrow 0} |\mathbb{E} [\Phi(\bar{X}^\varepsilon)] - \mathbb{E} [\Phi(\bar{X})]| = \lim_{i \rightarrow \infty} |\mathbb{E} [\Phi(\bar{X}^{\varepsilon_i})] - \mathbb{E} [\Phi(\bar{X})]|.$$

Let $\Lambda \in \mathfrak{L}$. Owing to Proposition I.6.1, up to the extraction of another subsequence, we can assume that $(\bar{X}^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}}, \zeta^{\varepsilon_i})$ converges in distribution to some (\bar{X}_Λ, ζ) in $(\mathcal{C}_T^0 H^{-\sigma} \times \mathcal{C}_T^0 \mathcal{C}_x^1) \times \mathcal{C}_T^0 \mathcal{C}_x^1$. Now we write

$$\begin{aligned}|\mathbb{E} [\Phi(\bar{X}^{\varepsilon_i})] - \mathbb{E} [\Phi(\bar{X})]| &\leq \left| \mathbb{E} [\Phi(\bar{X}^{\varepsilon_i})] - \mathbb{E} [\Phi(\bar{X}^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}(\zeta^{\varepsilon_i})})] \right| \\ &\quad + \left| \mathbb{E} [\Phi(\bar{X}^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}(\zeta^{\varepsilon_i})})] - \mathbb{E} [\Phi(\bar{X})] \right|.\end{aligned}$$

First, we have

$$\left| \mathbb{E} [\Phi(\bar{X}^{\varepsilon_i})] - \mathbb{E} [\Phi(\bar{X}^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}(\zeta^{\varepsilon_i})})] \right| \lesssim_\Phi \mathbb{P}(\tau_\Lambda(\zeta^{\varepsilon_i}) \leq T).$$

By Lemmas I.7.2 and I.7.3, since $\Lambda \in \mathfrak{L}$, $\tau_\Lambda(\zeta^{\varepsilon_i}) \wedge 2T$ converges in distribution to $\tau_\Lambda(\zeta') \wedge 2T$. Then, by Portmanteau's Theorem for closed sets, we have

$$\limsup_i \mathbb{P}(\tau_\Lambda(\zeta^{\varepsilon_i}) \leq T) \leq \mathbb{P}(\tau_\Lambda(\zeta') \leq T).$$

Since Φ is a continuous bounded function, we have

$$\lim_i \left| \mathbb{E} [\Phi(\bar{X}^{\varepsilon_i, \tau_\Lambda^{\varepsilon_i}(\zeta^{\varepsilon_i})})] - \mathbb{E} [\Phi(\bar{X})] \right| = |\mathbb{E} [\Phi(\bar{X}_\Lambda)] - \mathbb{E} [\Phi(\bar{X})]|.$$

Recall that $\bar{X}_\Lambda \stackrel{d}{=} S_\Lambda(Y_\Lambda)$, and that $Y_\Lambda \stackrel{d}{=} \bar{X}$ (by Section I.7.3). Thus, we get

$$\begin{aligned}|\mathbb{E} [\Phi(\bar{X}_\Lambda)] - \mathbb{E} [\Phi(\bar{X})]| &= |\mathbb{E} [\Phi(S_\Lambda(Y_\Lambda))] - \mathbb{E} [\Phi(Y_\Lambda)]| \\ &\lesssim_\Phi \mathbb{P}(\tau_\Lambda(Y_\Lambda) \leq T).\end{aligned}$$

Since $\tau_\Lambda(Y_\Lambda) \stackrel{d}{=} \tau_\Lambda(\bar{X}_\Lambda) = \tau_\Lambda(\zeta')$ by Lemma I.7.5, we finally get for $\Lambda \in \mathfrak{L}$

$$\limsup_{\varepsilon \rightarrow 0} |\mathbb{E} [\Phi(\bar{X}^\varepsilon)] - \mathbb{E} [\Phi(\bar{X})]| \lesssim_\Phi \mathbb{P}(\tau_\Lambda(\zeta') \leq T).$$

Since $\zeta' \in \mathcal{C}_T^0 \mathcal{C}_x^1$, we have $\mathbb{P}(\tau_\Lambda(\zeta') \leq T) \xrightarrow[\Lambda \rightarrow \infty]{} 0$. Recall that we can take this limit since \mathcal{L} contains arbitrarily large Λ 's. Therefore, we have

$$\mathbb{E}[\Phi(\overline{X}^\varepsilon)] \xrightarrow[\varepsilon \rightarrow 0]{} \mathbb{E}[\Phi(\overline{X})].$$

This concludes the proof that \overline{X}^ε converges in distribution to \overline{X} , and in particular that ρ^ε converges in distribution to ρ in $\mathcal{C}_T^0 H_x^{-\sigma}$.

I.8 Strong convergence

In this section, we establish the second convergence result stated in Theorem I.2.1, namely the convergence in $L_T^2 L_x^2$. Given Section I.7 and Proposition I.6.1, it is sufficient to prove that the sequence $(\rho^{\varepsilon, \tau_\Lambda^\varepsilon})_{\varepsilon > 0}$ is tight in $L_T^2 L_x^2$. Indeed, assume this tightness holds. There exists a limit point ρ'_Λ and a subsequence converging to ρ'_Λ in distribution in $L_T^2 L_x^2$, hence in $L_T^2 H_x^{-\sigma}$. Owing to the first part of Theorem I.2.1, this subsequence converges to $\rho^{\tau_\Lambda(\zeta)}$ in distribution in $\mathcal{C}_T^0 H_x^{-\sigma}$, hence in $L_T^2 H_x^{-\sigma}$. As a consequence, ρ_Λ and ρ'_Λ are equal in distribution. As in Section I.7.4, we deduce the convergence in $L_T^2 L_x^2$.

Therefore, it only remains to prove that $(\rho^{\varepsilon, \tau_\Lambda^\varepsilon})_{\varepsilon > 0}$ is tight in $L_T^2 L_x^2$ to get this convergence.

Recall that w_ρ denotes the modulus of continuity of a $H_x^{-\sigma}$ -valued continuous process ρ . Then, using [Sim87, Theorem 5], the set

$$K_R \doteq \left\{ \rho \in L_T^2 L_x^2 \mid \|\rho\|_{L_T^2 H_x^{\sigma'}} \leq R \text{ and } \forall \delta > 0, w_\rho(\delta) < \eta(\delta) \right\}$$

where $R > 0$, $\sigma' > 0$ and $\eta(\delta) \xrightarrow[\delta \rightarrow 0]{} 0$, is compact in $L_T^2 L_x^2$. Using Prokhorov's Theorem, the tightness of $(\rho^{\varepsilon, \tau_\Lambda^\varepsilon})_{\varepsilon > 0}$ in $L_T^2 L_x^2$ will follow if we prove that for all $\eta > 0$, there exists $R > 0$ and $\sigma' > 0$ such that

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \mathbb{P} \left(w_{\rho^{\varepsilon, \tau_\Lambda^\varepsilon}}(\delta) > \eta \right) = 0, \quad (\text{I.8.1})$$

and

$$\sup_{\varepsilon} \mathbb{P} \left(\|\rho^{\varepsilon, \tau_\Lambda^\varepsilon}\|_{L_T^2 H_x^{\sigma'}} > R \right) < \eta. \quad (\text{I.8.2})$$

Equation (I.8.1) is a direct consequence of (I.6.10). It remains to prove (I.8.2). Owing to the Markov Inequality, it is sufficient to prove that, for some $\sigma' > 0$, we have

$$\sup_{\varepsilon} \mathbb{E} \left[\|\rho^{\varepsilon, \tau_\Lambda^\varepsilon}\|_{L_T^2 H_x^{\sigma'}}^2 \right] \lesssim_\Lambda 1. \quad (\text{I.8.3})$$

Let $g^\varepsilon = \varepsilon \partial_t f^\varepsilon + a(v) \cdot \nabla_x f^\varepsilon$. Owing to Assumption I.3, we can use an averaging lemma ([BD99, Theorem 2.3] with $f(t) = f^\varepsilon(\varepsilon t)$, $g(t) = g^\varepsilon(\varepsilon t)$ and $h = 0$ until the time $T \wedge \tau_\Lambda^\varepsilon$) and by rescaling the time, we get

$$\begin{aligned} \|\rho^{\varepsilon, \tau_\Lambda^\varepsilon}\|_{L_T^2 H_x^{\sigma^*/4}}^2 &= \int_0^{T \wedge \tau_\Lambda^\varepsilon} \|\rho^{\varepsilon, \tau_\Lambda^\varepsilon}(t)\|_{H_x^{\sigma^*/4}}^2 dt \\ &\lesssim \varepsilon \|f_0^\varepsilon\|_{L_x^2}^2 + \int_0^{T \wedge \tau_\Lambda^\varepsilon} \|f^{\varepsilon, \tau_\Lambda^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}^2 dt + \int_0^{T \wedge \tau_\Lambda^\varepsilon} \|g^{\varepsilon, \tau_\Lambda^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}^2 dt, \end{aligned}$$

where, using the Cauchy-Schwarz inequality,

$$\begin{aligned} \|g^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}}(t)\|_{L^2(\mathcal{M}^{-1})} &= \left\| f^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}}(t) \bar{m}^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}}(t) + \frac{1}{\varepsilon} L f^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}}(t) \right\|_{L^2(\mathcal{M}^{-1})} \\ &\leq \|f^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}}(t)\|_{L^2(\mathcal{M}^{-1})} \|\bar{m}^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}}(t)\|_{L_x^2} + \frac{1}{\varepsilon} \|L f^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}}(t)\|_{L^2(\mathcal{M}^{-1})}. \end{aligned}$$

Then Assumption [I.4](#), [\(I.3.5\)](#) and Proposition [I.3.6](#) lead to [\(I.8.3\)](#) with $\sigma' = \frac{\sigma^*}{4}$. Since the sets K_R are compacts, Prokhorov's Theorem yields, using [\(I.8.2\)](#) and [\(I.8.1\)](#), that $(\rho^{\varepsilon, \tau_{\bar{\Lambda}}^{\varepsilon}})_{\varepsilon>0}$ is tight in $L_T^2 L_x^2$.

Given Section [I.7](#), this concludes the proof of the convergence in distribution of ρ^{ε} in $L_T^2 L_x^2$ to ρ the solution of [\(I.2.13\)](#), and the proof of Theorem [I.2.1](#).

Chapter II

Convergence in the averaging regime with an unbounded driving process

II.1 Introduction

Our aim in this work is to study the scaling limit of a stochastic kinetic equation towards a diffusion Partial Differential Equation (PDE). For deterministic problems, this is a thoroughly studied field in the literature, starting historically with [LK74, BLP79]. Kinetic models with small parameters appear in various situations, for example when studying semi-conductors [GP92] and discrete velocity models [LT97] or as a limit of a system of particles, either with a single particle [GR09] or multiple ones [PV03]. It is important to understand the limiting equations, which are in general much easier to simulate numerically.

In this chapter, we consider the following equation

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon + m^\varepsilon f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon, \quad f^\varepsilon(0) = f_0^\varepsilon, \quad (\text{II.1.1})$$

where f^ε is defined on $\mathbb{R}^+ \times \mathbb{T}^d \times V$, L is a linear operator (see (II.1.2) below) and the source term m^ε is an ergodic random process defined on $\mathbb{R}^+ \times \mathbb{T}^d$ (satisfying assumptions given in Section II.2.2). The goal of this chapter is to study the limit $\varepsilon \rightarrow 0$ of its solution f^ε .

The solution $f^\varepsilon(t, x, v)$ is interpreted as a distribution function of particles, having position x and velocity $a(v)$ at time t .

The Bhatnagar-Gross-Krook operator L expresses the particle interactions, defined on $L^1(V, d\mu)$ by

$$Lf = \rho \mathcal{M} - f, \quad (\text{II.1.2})$$

where μ is a finite measure on V , where $\rho \doteq \langle f \rangle \doteq \int_V f d\mu$ and where $\mathcal{M} \in L^1(V, d\mu)$ satisfies Assumption II.1.

The source term m^ε is defined as $m^\varepsilon(t, x) = m(t\varepsilon^{-2}, x)$ where m is an ergodic random process defined in Assumption II.2, not depending on ε . When $\varepsilon \rightarrow 0$, $t\varepsilon^{-2} \rightarrow \infty$, hence m^ε converges to \bar{m} the average of the invariant distribution of the ergodic process m , assuming \bar{m} exists.

In the deterministic case $m^\varepsilon = 0$, such a problem occurs in various physical situations [DGP00]. The density $\rho^\varepsilon \doteq \langle f^\varepsilon \rangle$ converges to the solution of the linear parabolic equation

$$\partial_t \rho = \text{div}(K \nabla \rho), \quad (\text{II.1.3})$$

on $\mathbb{R}^+ \times \mathbb{T}^d$, where

$$K \doteq \int_V a(v) \otimes a(v) \mathcal{M}(v) d\mu(v). \quad (\text{II.1.4})$$

This is a diffusive limit in the PDE sense, since the limit equation is a diffusion equation.

Consider now two scaling parameters ε and δ and the equation

$$\partial_t f^{\varepsilon, \delta} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^{\varepsilon, \delta} + m^\delta f^{\varepsilon, \delta} = \frac{1}{\varepsilon^2} L f^{\varepsilon, \delta}.$$

When $\varepsilon > 0$ is fixed and $\delta \rightarrow 0$, the ergodicity of m^δ yields the convergence of $f^{\varepsilon, \delta}$ to the solution of

$$\partial_t f^{\varepsilon, 0} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^{\varepsilon, 0} + \overline{m} f^{\varepsilon, 0} = \frac{1}{\varepsilon^2} L f^{\varepsilon, 0},$$

which converges when $\varepsilon \rightarrow 0$ to a diffusive limit in the PDE sense: as in the deterministic case, $\rho^{\varepsilon, 0}$ converges to the solution of

$$\partial_t \rho^{0, 0} + \overline{m} \rho^{0, 0} = \operatorname{div}(K \nabla \rho^{0, 0}). \quad (\text{II.1.5})$$

Conversely, when $\varepsilon \rightarrow 0$ and $\delta > 0$ is fixed, we get as in the deterministic case the convergence of $\rho^{\varepsilon, \delta}$ to the solution of

$$\partial_t \rho^{0, \delta} + m^\delta \rho^{0, \delta} = \operatorname{div}(K \nabla \rho^{0, \delta}),$$

which also converges to the solution of (II.1.5).

In this chapter, we consider the regime $\varepsilon = \delta \rightarrow 0$. In this case, (II.1.1) is expected to converge both to a diffusive limit in the PDE sense (as in [DGP00]) and to an averaging limit in the probabilistic sense (owing to the averaging principle introduced in [Kha68], see also [PS08, Cer09, CF09, Bré12, Bré20]). The main results, Theorems II.2.1 and II.2.2, establish that, under appropriate assumptions, the density ρ^ε converges in distribution in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ for all arbitrarily small $\sigma \in (0, 1]$ and in $L^2([0, T], L^2(\mathbb{T}^d))$ to the solution of the deterministic linear diffusion equation with a source \overline{m}

$$\partial_t \rho + \overline{m} \rho = \operatorname{div}(K \nabla \rho), \quad \rho(0) = \rho_0, \quad (\text{II.1.6})$$

and converges in probability in the same spaces if ρ_0 is deterministic. The quantities K and \overline{m} are defined in Assumption II.1 and Assumption II.2. Moreover, we prove that f^ε converges in distribution in $L^2([0, T], L^2(\mathbb{T}^d))$ to $\rho \mathcal{M}$, and in probability in the same space if ρ_0 is deterministic.

This chapter focuses on a martingale approach combined with perturbed test functions, as in the classical article [PSV77] (see also [Kus84, EK86, FGPS107, PS08, dBG12]). Perturbed test functions in the context of PDEs with diffusive limits applies in various situations, for instance in the context of viscosity solutions [Eva89], nonlinear Schrödinger equations [dBG12], a parabolic PDE [PP03] or, as in this chapter, kinetic SPDEs [DV12, DV20, DRV20, RR20]. As in Chapter I, we combine this martingale approach with stopping times to get estimates on the processes. However, as opposed to Chapter I, the stopping times all diverge to infinity when $\varepsilon \rightarrow 0$ and thus do not appear in the limit martingale formulation.

This chapter is organized as follows: in Section II.2, we set some notation, the assumptions and the main results, Theorems II.2.1 and II.2.2. Section II.3 proves some estimates on f^ε and m^ε when stopping them at some stopping times and establishes the

well-posedness of equation (II.1.1) and of the limit equation (II.1.6). In Sections II.4 and II.5, we introduce respectively the notion of perturbed test functions and of martingale formulation of equations (II.1.1) and (II.1.6). In Section II.6, we use these tools to prove Theorem II.2.1, namely the convergence in distribution of the family of processes $(\rho^\varepsilon)_\varepsilon$ in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ for all arbitrarily small $\sigma \in (0, 1]$, as well as its convergence in probability. In Section II.7, we prove Theorem II.2.2, which extends the convergence in a stronger sense, namely the convergence of both $(\rho^\varepsilon)_\varepsilon$ and $(f^\varepsilon)_\varepsilon$ in $L^2([0, T], L^2(\mathbb{T}^d))$, using an additional assumption and an averaging lemma.

II.2 Assumptions and main results

II.2.1 Functional spaces

The space V is equipped with a finite measure μ .

In this chapter, we consider the Hilbert spaces L_x^2 , L^2 and $L^2(\mathcal{M}^{-1})$ associated with the following inner products

$$\begin{aligned} (h, k)_{L_x^2} &= \int_{\mathbb{T}^d} h(x)k(x)dx \\ (h, k)_{L^2} &= \int_{\mathbb{T}^d \times V} h(x, v)k(x, v)dx d\mu(v) \\ (h, k)_{L^2(\mathcal{M}^{-1})} &= \int_{\mathbb{T}^d \times V} h(x, v)k(x, v)dx \frac{d\mu(v)}{\mathcal{M}(v)}. \end{aligned}$$

We write $\mathcal{C}_x^i \doteq \mathcal{C}^i(\mathbb{T}^d)$, $\mathcal{C}_T^0 H_x^{-\sigma} \doteq \mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$ and $L_T^p L_x^2 \doteq L^p([0, T], L^2(\mathbb{T}^d))$ for any $T \in (0, \infty)$, $\sigma \in (0, 1]$, $i \in \mathbb{N}_0$ and $p \in [1, \infty]$.

II.2.2 Assumptions

Assumption II.1. The functions $a \in L^\infty(V, d\mu)$ and $\mathcal{M} \in L^1(V, d\mu)$ satisfy

$$\begin{aligned} \forall v \in V, \mathcal{M}(v) &> 0, \\ \int_V \mathcal{M}(v) d\mu(v) &= 1, \\ \int_V a(v) \mathcal{M}(v) d\mu(v) &= 0, \\ K \doteq \int_V a(v) \otimes a(v) \mathcal{M}(v) d\mu(v) &> 0. \end{aligned}$$

Example II.2.1. For any $\sigma \in (0, \infty)$ and any bounded odd function $a \neq 0$, the space $V = \mathbb{R}^d$ satisfy this assumption with $\mathcal{M}(v) d\mu(v) = \frac{1}{\sqrt{2\pi}^d} \exp(-\frac{\|v\|^2}{2}) dv$.

Assumption II.2. The family $(m_\ell(t))_{\ell \in E}$ defines an E -valued Markov process, with $m_\ell(0) = \ell$, where $E = \mathcal{C}_x^{\lfloor d/2 \rfloor + 2}$. Let \mathcal{L}_m be its infinitesimal generator. We also assume the existence and uniqueness of an invariant distribution ν for $(m_\ell(t))_{\ell \in E}$ with $\int_E \|\ell\| d\nu(\ell) < \infty$. Let $\bar{m} \doteq \int_E \ell d\nu(\ell)$. In this chapter, an initial condition $\ell_0 \in E$ is fixed (ℓ_0 does not depend on ε), and we write $m(t) = m_{\ell_0}(t)$.

In this chapter, once ℓ_0 , a , T and \bar{m} are fixed, we always consider $\varepsilon \in (0, \varepsilon_0]$, where

$$\varepsilon_0 < \min \left\{ \|\ell_0\|_E^{-1}, \left(4 \|a\|_{L^\infty} (1 + T \|\bar{m}\|_{C_x^1}) \right)^{-1} \right\}. \quad (\text{II.2.1})$$

This upper bound on ε is justified by Definition II.2.3 and Proposition II.3.3, and is not restrictive, since we consider the regime $\varepsilon \rightarrow 0$.

Assumption II.3. For $\varepsilon \in (0, \varepsilon_0]$, the stochastic process m^ε is defined as $m^\varepsilon(t) = m(t\varepsilon^{-2})$ and is adapted to a filtration $(\mathcal{F}_t^\varepsilon)_{t \in \mathbb{R}^+}$.

Assumption II.4. There exists $\gamma \in (2, \infty)$ such that

$$\sup_{i \in \mathbb{N}_0} \mathbb{E} \left[\sup_{t \in [i, i+1]} \|m(t)\|_E^\gamma \right] < \infty. \quad (\text{II.2.2})$$

Assumption II.5 (Mixing property). There exists a positive function $\gamma_{\text{mix}} \in L^1(\mathbb{R}^+)$ such that, for any $\ell_1, \ell_2 \in E$, there exists a coupling $(m_{\ell_1}^*, m_{\ell_2}^*)$ of (m_{ℓ_1}, m_{ℓ_2}) satisfying

$$\forall t \in \mathbb{R}^+, \mathbb{E} [\|m_{\ell_1}^*(t) - m_{\ell_2}^*(t)\|_E] \leq \gamma_{\text{mix}}(t) \|\ell_1 - \ell_2\|_E.$$

Definition II.2.1 (Resolvent operator). For any continuous linear form $\theta \in E^*$, the resolvent operator is defined as

$$R_0(\theta - \theta(\bar{m}))(\ell) = \int_0^\infty \mathbb{E} [\theta(m_\ell(t)) - \theta(\bar{m})] dt.$$

Owing to the mixing property (Assumption II.5) and the identity $\int_E \theta(\ell) d\nu(\ell) = \theta(\bar{m})$, the function $\psi_\theta = R_0(\theta - \theta(\bar{m}))$ is well-defined and we have for $\ell \in E$

$$|\psi_\theta(\ell)| \leq C \|\theta\|_{E^*} (1 + \|\ell\|_E), \quad (\text{II.2.3})$$

where C depends only on γ_{mix} and ν . Moreover, $\mathcal{L}_m \psi_\theta = -(\theta - \theta(\bar{m}))$.

Assumption II.6. For any $\theta \in E^*$, let $\psi_\theta = R_0(\theta - \theta(\bar{m}))$. Assume that ψ_θ^2 is in the domain $D(\mathcal{L}_m)$ of \mathcal{L}_m and satisfies

$$\sup_{\ell \in E} \frac{|\mathcal{L}_m(\psi_\theta^2)(\ell)|}{1 + \|\ell\|_E^2} < \infty.$$

Example II.2.2. Let us check that the conditions above are satisfied for an Ornstein-Uhlenbeck process. Let $(m_\ell(t))_{\ell \in E}$ be defined by

$$dm_{\ell,t} = -(m_{\ell,t} - \bar{m})dt + dW_t, \quad m_{\ell,0} = \ell, \quad (\text{II.2.4})$$

where W is an E -valued Wiener process. Then, we have

$$m_\ell(t) = \ell e^{-t} + \bar{m} (1 - e^{-t}) + \int_0^t e^{s-t} dW_s. \quad (\text{II.2.5})$$

Assumption II.4 is satisfied for any $\gamma \in (2, \infty)$. The coupling $(m_{\ell_1}^*, m_{\ell_2}^*)$ of Assumption II.5 is obtained by driving both processes by the same Wiener process W . Indeed, (II.2.5) becomes

$$m_{\ell_1}^*(t) - m_{\ell_2}^*(t) = (\ell_1 - \ell_2)e^{-t},$$

and Assumption II.5 is satisfied with $\gamma_{\text{mix}}(t) = e^{-t}$. Moreover, with the notation of Assumption II.6, we have

$$\psi_\theta(\ell) = \int_0^\infty (\theta(\ell) - \theta(\bar{m})) e^{-t} dt = \theta(\bar{m}) - \theta(\ell).$$

Since the infinitesimal generator is given by $\mathcal{L}_m \varphi(\ell) = D\varphi(\ell) \cdot (\bar{m} - \ell) + \frac{1}{2} \text{Tr}(D^2 \varphi(\ell))$, Assumption II.6 is also satisfied.

Assumption II.7. The family of initial conditions $(f_0^\varepsilon)_{\varepsilon \in (0, \varepsilon_0]}$ satisfies

$$\sup_{\varepsilon \in (0, \varepsilon_0]} \mathbb{E} \left[\|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2 \right] < \infty,$$

and the initial density $\rho_0^\varepsilon \in L_x^2$ converges in distribution in L_x^2 to $\rho_0 \in L_x^2$.

II.2.3 Statement of the main results

Definition II.2.2. Let ρ_0 be the L_x^2 -valued random variable defined by Assumption II.7. A stochastic process ρ is a weak solution of (II.1.6) in L_x^2 if $\rho \in L_T^\infty L_x^2$ almost surely and if, for all $\xi \in H_x^2$ and $t \in [0, T]$, almost surely,

$$(\rho(t), \xi)_{L_x^2} = (\rho_0, \xi)_{L_x^2} + \int_0^t (\rho(s), \text{div}(K \nabla \xi) - \bar{m} \xi)_{L_x^2} ds. \quad (\text{II.2.6})$$

Theorem II.2.1. *Let Assumptions II.1 to II.7 be satisfied. Then, ρ^ε converges in distribution in $\mathcal{C}_T^0 H_x^{-\sigma}$ to the unique weak solution ρ of (II.1.6) in the sense of Definition II.2.2. If, moreover, $\rho_0 \in L_x^2$ is deterministic, then the convergence holds in probability.*

When Assumption II.8 below is satisfied, we can prove a stronger convergence. Its formulation corresponds to the assumptions of the averaging lemma [BD99, Theorem 2.3].

Assumption II.8. Consider the case $(V, d\mu) = (\mathbb{R}^n, \frac{d\mu}{dv}(v)dv)$ where $\frac{d\mu}{dv} \in H^1(\mathbb{R}^n)$, and assume that a is locally Lipschitz-continuous and that there exists $C \in (0, \infty)$ and $\sigma' \in (0, 1]$ such that

$$\forall u \in S^{d-1}, \forall \lambda \in \mathbb{R}, \forall \delta \in (0, \infty), \int_{\lambda < a(v) \cdot u < \lambda + \delta} \left(\left| \frac{d\mu}{dv}(v) \right|^2 + \left| \nabla \frac{d\mu}{dv}(v) \right|^2 \right) dv \leq C \delta^{\sigma'}.$$

Theorem II.2.2. *Let the additional Assumption II.8 be satisfied. Then ρ^ε also converges in distribution in $L_T^2 L_x^2$ to ρ and f^ε converges in distribution in $L_T^2 L^2(\mathcal{M}^{-1})$ to $\rho \mathcal{M}$. Moreover, if ρ_0 is deterministic, then those two convergences hold in probability.*

Note that we get convergences in the same spaces as in the deterministic case [DGP00], either in distribution or in probability depending on the initial condition ρ_0 .

Remark II.2.1. In this setting, the initial condition $m^\varepsilon(0) = \ell_0$ does not depend on ε . It may be possible to extend the analysis with some changes in Assumption II.4 and appropriate modifications of the proofs below:

- If there exists a deterministic \mathcal{R} such that $\sup_{\varepsilon \in (0, \varepsilon_0]} \|m^\varepsilon(0)\|_E \leq \mathcal{R}$ almost surely, the result holds if Assumption II.4 is modified as follows: assume that

$$\sup_{i \in \mathbb{N}_0} \mathbb{E} \left[\sup_{\varepsilon \in (0, \varepsilon_0]} \sup_{t \in [i, i+1]} \|m_{m^\varepsilon(0)}(t)\|_E^\gamma \right] < \infty \text{ and } \varepsilon_0 < \mathcal{R}^{-1}.$$

- If $m^\varepsilon(0)$ is a random variable with distribution ν , so that m^ε is a stationary process, then one can keep Assumption II.4. The general results still holds if one assumes higher order moments on ν (hence m) and f_0^ε . The details are omitted.

II.2.4 Strategy of the proof

A standard strategy to prove the convergence of ρ^ε when $\varepsilon \rightarrow 0$ (see [DV12, DV20, DRV20]) is to establish the tightness of the family $(\rho^\varepsilon)_{\varepsilon > 0}$, and to prove that the limit point of this family is unique and solves (II.1.6). The tightness usually comes from estimates on moments of trajectories. In this chapter, these estimates require to introduce an appropriate stopping time (see Definition II.2.3 below). Therefore, as in Chapter I, we prove the tightness and convergence of the family of stopped processes $\rho^{\varepsilon, \tau^\varepsilon}$, and deduce the tightness and convergence of the original process.

Definition II.2.3. Let $\alpha \in (\frac{2}{\gamma}, 1)$ be fixed. The $(\mathcal{F}_t^\varepsilon)_{t \in \mathbb{R}^+}$ -stopping time τ^ε is defined by

$$\tau^\varepsilon = \tau_m^\varepsilon \wedge \tau_\zeta^\varepsilon,$$

where

$$\begin{aligned} \tau_m^\varepsilon &= \inf \left\{ t \in [0, T] \mid \|m^\varepsilon(t)\|_E \geq \varepsilon^{-\alpha} \right\}, \\ \tau_\zeta^\varepsilon &= \inf \left\{ t \in [0, T] \mid \|\zeta^\varepsilon(t)\|_{C_x^1} \geq \varepsilon^{-1} \right\}, \end{aligned}$$

and

$$\zeta^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t (m^\varepsilon(s) - \bar{m}) ds \in E. \quad (\text{II.2.7})$$

The parameter α introduced in Definition II.2.3 is fixed in all this chapter.

For any process X^ε , let $X^{\varepsilon, \tau^\varepsilon}$ be the process stopped at the time τ^ε , namely

$$X^{\varepsilon, \tau^\varepsilon}(t) = X^\varepsilon(t \wedge \tau^\varepsilon),$$

where $s \wedge t \doteq \min(s, t)$.

Remark II.2.2. Since $\alpha < 1$ and $\varepsilon \leq \varepsilon_0$, where ε_0 satisfies (II.2.1), we have $\|m^\varepsilon(0)\|_E = \|\ell_0\|_E \leq \varepsilon^{-\alpha}$. Thus, owing to Definition II.2.3, for $t \in \mathbb{R}^+$, we have

$$\|m^{\varepsilon, \tau^\varepsilon}(t)\|_E = \|m^\varepsilon(t \wedge \tau^\varepsilon)\|_E \leq \varepsilon^{-\alpha}. \quad (\text{II.2.8})$$

Since $\zeta^\varepsilon(0) = 0$, we also have

$$\|\zeta^{\varepsilon, \tau^\varepsilon}(t)\|_E = \|\zeta^\varepsilon(t \wedge \tau^\varepsilon)\|_E \leq \varepsilon^{-1}. \quad (\text{II.2.9})$$

Below, we prove that $\tau^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} +\infty$ in probability (Proposition II.6.1). Owing to Slutsky's Lemma, it follows that the convergence in distribution of $(\rho^\varepsilon)_{\varepsilon \in (0, \varepsilon_0]}$ is equivalent to the convergence in distribution of $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ and any limit point of $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ does not depend on any stopping time. Therefore, we avoid the technical issues of having a stopping time that persists at the limit as in Chapter I.

In Section II.3, we prove an estimate on $f^{\varepsilon, \tau^\varepsilon}$, depending only on T , and f_0^ε . This estimate is a crucial point to prove the tightness of the family of stopped processes $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon > 0}$. In Section II.4, we introduce the framework of perturbed test functions, and Section II.5 make use of this tool to reformulate equation (II.1.1) and the limit equation (II.1.6) in terms of martingale problems. Then, in Section II.6, we use the estimates of Section II.3 and the martingale formulation of (II.1.1) to prove the tightness of $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ in $\mathcal{C}_T^0 H_x^{-\sigma}$. By taking the limit $\varepsilon \rightarrow 0$ in this martingale formulation, we prove that any limit point of this family is the unique solution of the martingale problem associated with the limit equation (II.1.6). We then deduce the convergence in distribution of the stopped process to ρ , and then the convergence of the original process. Portmanteau's Theorem gives the convergence in probability if ρ_0 is deterministic, which concludes the proof of Theorem II.2.1. In Section II.7, we use an averaging lemma to extend the convergence in $\mathcal{C}_T^0 H_x^{-\sigma}$ to a convergence in $L^2([0, T], L_x^2)$ under the assumptions of Theorem II.2.2. Combined with an estimate obtained in Section II.3, we deduce the convergence of the solution f^ε in $L^2([0, T], L_x^2)$.

II.3 Preliminary results

II.3.1 Well-posedness of the equation for fixed ε

In this section, we prove that, when ε is fixed, equation (II.1.1) is well-posed in a almost sure sense. Let us start by proving that, almost surely, the process m^ε is bounded on $[0, T]$ for fixed ε .

Lemma II.3.1. *For all $T \in (0, \infty)$ and $\varepsilon \in (0, \varepsilon_0]$, we have, almost surely,*

$$\sup_{t \in [0, T]} \|m^\varepsilon(t)\|_E < \infty. \quad (\text{II.3.1})$$

Moreover, for any $T \in (0, \infty)$, we have

$$\mathbb{P}(\tau_m^\varepsilon < T) \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (\text{II.3.2})$$

The proof of Lemma II.3.1 is adapted from the proof of Lemma I.3.4.

Proof. For $i \in \mathbb{N}_0$, let

$$S_i \doteq \sup_{t \in [i, i+1]} \|m(t)\|_E.$$

Assumption II.4 can be rewritten $\sup_{i \in \mathbb{N}_0} \mathbb{E}[S_i^\gamma] < \infty$. Thus, we have

$$\mathbb{P}\text{-a.s.}, \forall i \in \mathbb{N}_0, S_i < \infty. \quad (\text{II.3.3})$$

For $\varepsilon \in (0, \varepsilon_0]$ and $T \in (0, \infty)$, (II.3.1) follows:

$$\mathbb{P}\text{-a.s.}, \sup_{t \in [0, T]} \|m^\varepsilon(t)\|_E \leq \sup_{i \leq T\varepsilon^{-2}+1} S_i < \infty.$$

It only remains to prove (II.3.2). Since $\alpha > \frac{2}{\gamma}$ in Definition II.2.3, there exists δ such that $\frac{\alpha}{2} > \delta > \frac{1}{\gamma}$. Then, owing to the Markov inequality,

$$\sum_{i \in \mathbb{N}} \mathbb{P}(S_i \geq i^\delta) \leq \sum_{i \in \mathbb{N}} \frac{\mathbb{E}[S_i^\gamma]}{i^{\delta\gamma}} = \sup_{i \in \mathbb{N}} \mathbb{E}[S_i^\gamma] \sum_{i \in \mathbb{N}} \frac{1}{i^{\delta\gamma}} < \infty,$$

where we used once again Assumption II.4. Then, Borel-Cantelli's Lemma yields the existence of a random variable $i_0 \in \mathbb{N}$ such that

$$\mathbb{P}\text{-a.s.}, \forall i > i_0, S_i < i^\delta.$$

Let $Z \doteq \sup_{i \leq i_0} S_i$. Owing to (II.3.3), the random variable Z is finite almost surely and

$$\mathbb{P}\text{-a.s.}, \forall t \in \mathbb{R}^+, \|m^\varepsilon(t)\|_E \leq S_{\lfloor t\varepsilon^{-2} \rfloor} \leq Z + \lfloor t\varepsilon^{-2} \rfloor^\delta \leq Z + (t\varepsilon^{-2})^\delta.$$

Then, for $T \in (0, \infty)$, we get

$$\mathbb{P}(\tau_m^\varepsilon < T) = \mathbb{P}\left(\sup_{t \in [0, T]} \|m^\varepsilon(t)\|_E > \varepsilon^{-\alpha}\right) \leq \mathbb{P}(Z + (T\varepsilon^{-2})^\delta > \varepsilon^{-\alpha}) \xrightarrow{\varepsilon \rightarrow 0} 0,$$

since $\alpha > 2\delta$ and since Z is almost surely finite. \square

Since the equation is linear in f and a is bounded, the well-posedness of (II.1.1) is a consequence of Lemma II.3.1.

Proposition II.3.2. *Let $T \in (0, \infty)$ and $\varepsilon \in (0, \varepsilon_0]$. Then for any $f_0^\varepsilon \in L^2(\mathcal{M}^{-1})$, there exists, almost surely, a unique solution f^ε of (II.1.1) in $\mathcal{C}^0([0, T]; L^2(\mathcal{M}^{-1}))$, in the sense that*

$$\mathbb{P}\text{-a.s.}, \forall t \in [0, T], f^\varepsilon(t) = e^{-\frac{t}{\varepsilon}A} f_0^\varepsilon + \int_0^t e^{-\frac{t-s}{\varepsilon}A} \left(\frac{1}{\varepsilon^2} Lf^\varepsilon(s) + m^\varepsilon(s)f^\varepsilon(s) \right) ds$$

where A is the operator defined by

$$\begin{aligned} D(A) &\doteq \{f \in L^2(\mathcal{M}^{-1}) \mid (x, v) \mapsto a(v) \cdot \nabla_x f(x, v) \in L^2(\mathcal{M}^{-1})\} \\ Af(x, v) &\doteq a(v) \cdot \nabla_x f(x, v). \end{aligned}$$

Note that in Proposition II.3.2, ε is fixed. Thus, the proof is standard, based on a fixed-point theorem.

Proof. Let $\omega \in \Omega$ and $\varepsilon \in (0, \varepsilon_0]$. For $f \in \mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1}))$, let

$$\Phi(f) = e^{-\frac{t}{\varepsilon}A} f_0^\varepsilon + \int_0^t e^{-\frac{t-s}{\varepsilon}A} \left(\frac{1}{\varepsilon^2} Lf(s) + m^\varepsilon(s)f(s) \right) ds.$$

Owing to the Banach fixed-point theorem, it is sufficient to prove that Φ is a contraction for some Banach norm on $\mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1}))$. For $r \in [0, \infty)$, we consider the following Banach norm

$$\forall f \in \mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1})), \|f\|_r = \sup_{t \in [0, T]} e^{-rt} \|f(t)\|_{L^2(\mathcal{M}^{-1})}.$$

The semi-group associated with A satisfies

$$\forall f \in L^2(\mathcal{M}^{-1}), \forall x \in \mathbb{T}^d, \forall v \in V, e^{tA}f(x, v) = f(x + ta(v), v).$$

As a consequence, e^{tA} is an isometric mapping for all $t \in \mathbb{R}^+$: for $f \in L^2(\mathcal{M}^{-1})$,

$$\|e^{tA}f\|_{L^2(\mathcal{M}^{-1})} = \|f\|_{L^2(\mathcal{M}^{-1})}.$$

Then, for $t \in [0, T]$, and for all $f, g \in \mathcal{C}^0([0, T], L^2(\mathcal{M}^{-1}))$, we get

$$\|\Phi(f)(t) - \Phi(g)(t)\|_{L^2(\mathcal{M}^{-1})} \leq \int_0^t \left(\frac{1}{\varepsilon^2} \|L(f - g)(s)\|_{L^2(\mathcal{M}^{-1})} + \|m^\varepsilon(s)(f - g)(s)\|_{L^2(\mathcal{M}^{-1})} \right) ds.$$

Note that L defined by (II.1.2) is the orthogonal projector on $L^2(\mathcal{M}^{-1})$ of kernel

$$\left\{ \int_V h d\mu_{\mathcal{M}} \mid h \in L^2(\mathcal{M}^{-1}) \right\},$$

hence $\|Lh\|_{L^2(\mathcal{M}^{-1})} \leq \|h\|_{L^2(\mathcal{M}^{-1})}$ for $h \in L^2(\mathcal{M}^{-1})$. Therefore, using (II.3.1) of Lemma II.3.1, we get

$$\|\Phi(f)(t) - \Phi(g)(t)\|_{L^2(\mathcal{M}^{-1})} \leq \left(\frac{1}{\varepsilon^2} + \sup_{t \in [0, T]} \|m^\varepsilon(t)\|_E \right) \int_0^t e^{rs} ds \|f - g\|_r.$$

As a consequence, we have

$$e^{-rt} \|\Phi(f)(t) - \Phi(g)(t)\|_{L^2(\mathcal{M}^{-1})} \leq \left(\frac{1}{\varepsilon^2} + \sup_{t \in [0, T]} \|m^\varepsilon(t)\|_E \right) \frac{1 - e^{-rt}}{r} \|f - g\|_r,$$

and

$$\|\Phi(f) - \Phi(g)\|_r \leq \frac{1}{r} \left(\frac{1}{\varepsilon^2} + \sup_{t \in [0, T]} \|m^\varepsilon(t)\|_E \right) \|f - g\|_r.$$

Owing to Lemma II.3.1, and since ε is fixed, we have almost surely

$$\frac{1}{\varepsilon^2} + \sup_{t \in [0, T]} \|m^\varepsilon(t)\|_E < \infty.$$

Choosing r large enough, Φ is a contraction for the norm $\|\cdot\|_r$, which concludes the proof of Proposition II.3.2. \square

II.3.2 Estimate on the solution in $L^2(\mathcal{M}^{-1})$

Proposition II.3.3. *For $T \in (0, \infty)$, $t \in [0, T]$ and $\varepsilon \in (0, \varepsilon_0]$, we have almost surely*

$$\|f^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}^2 + \frac{1}{\varepsilon^2} \int_0^{t \wedge \tau^\varepsilon} \|Lf^{\varepsilon, \tau^\varepsilon}(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \leq C(T) \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2, \quad (\text{II.3.4})$$

where $C(T) \in (0, \infty)$ only depends on \overline{m} , K (defined by (II.1.4)) and T , and where ε_0 satisfies (II.2.1).

Remark II.3.1. The estimate for the second term in the left-hand side of (II.3.4) is used in Section II.7 to prove Theorem II.2.2.

The proof of Proposition II.3.3 is very similar to its analog in Chapter I (see Proposition I.3.6). The main differences are the different scaling in ε , which requires a different weight \mathcal{M}^ε , and the average value $\bar{m} \neq 0$ which must be taken into account.

Proof. Let us introduce the weight

$$\mathcal{M}^\varepsilon(t, x, v) = \exp(-2\eta^\varepsilon(t, x)) \mathcal{M}(v),$$

and the associated weighted L^2 norm

$$\|f\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 \doteq \int \int \frac{|f(x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} dx d\mu(v),$$

where

$$\eta^\varepsilon(t, x) = \int_0^t m^\varepsilon(s, x) ds = \varepsilon \zeta^\varepsilon(t, x) + t \bar{m}(x) \in E,$$

and ζ^ε is defined by (II.2.7). Note that, owing to (II.2.9), we have for $t \in [0, \tau^\varepsilon]$

$$\|\eta^\varepsilon(t)\|_{C_x^1} \leq 1 + T \|\bar{m}\|_{C_x^1}. \quad (\text{II.3.5})$$

For $t \in \mathbb{R}^+$, we have

$$\begin{aligned} \frac{1}{2} \partial_t \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 &= \int \int \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} \partial_t f^\varepsilon(t, x, v) dx d\mu(v) \\ &\quad - \int \int \frac{|f^\varepsilon(t, x, v)|^2}{2 |\mathcal{M}^\varepsilon(t, x, v)|^2} \partial_t \mathcal{M}^\varepsilon(t, x, v) dx d\mu(v) \\ &= \mathcal{A}_\varepsilon + \mathcal{B}_\varepsilon + \mathcal{C}_\varepsilon \end{aligned}$$

with

$$\begin{aligned} \mathcal{A}_\varepsilon &= \frac{1}{\varepsilon^2} \int \int \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} L f^\varepsilon(t, x, v) dx d\mu(v) \\ \mathcal{B}_\varepsilon &= -\frac{1}{\varepsilon} \int \int \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} a(v) \cdot \nabla_x f^\varepsilon(t, x, v) dx d\mu(v) \\ \mathcal{C}_\varepsilon &= -\int \int \frac{|f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} \left(m^\varepsilon + \frac{\partial_t \mathcal{M}^\varepsilon}{2 \mathcal{M}^\varepsilon} \right) (t, x, v) dx d\mu(v). \end{aligned}$$

The weight \mathcal{M}^ε has been chosen such that $m^\varepsilon + \frac{\partial_t \mathcal{M}^\varepsilon}{2 \mathcal{M}^\varepsilon} = 0$, hence $\mathcal{C}_\varepsilon = 0$ for all $t \in [0, T]$ and $x \in \mathbb{T}^d$.

On the one hand, since $f^\varepsilon = \rho^\varepsilon \mathcal{M} - L f^\varepsilon$ and $\int_V L f^\varepsilon(t, x, v) d\mu(v) = 0$, we have

$$\begin{aligned} \mathcal{A}_\varepsilon &= \frac{1}{\varepsilon^2} \int \int \frac{f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} L f^\varepsilon(t, x, v) dx d\mu(v) \\ &= \frac{1}{\varepsilon^2} \int_{\mathbb{T}^d} e^{2\eta^\varepsilon(t, x)} \rho^\varepsilon(t, x) \int_V L f^\varepsilon(t, x, v) d\mu(v) dx - \frac{1}{\varepsilon^2} \int \int \frac{|L f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} d\mu(v) dx \\ &= -\frac{1}{\varepsilon^2} \|L f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2. \end{aligned}$$

On the other hand, we get by integration by parts

$$\begin{aligned}
\mathcal{B}_\varepsilon &= -\frac{1}{\varepsilon} \int \int a(v) \cdot \frac{f^\varepsilon(t, x, v) \nabla_x f^\varepsilon(t, x, v)}{\mathcal{M}^\varepsilon(t, x, v)} dx d\mu(v) \\
&= -\frac{1}{\varepsilon} \int \int a(v) \cdot \frac{\frac{1}{2} |f^\varepsilon(t, x, v)|^2 \nabla_x \mathcal{M}^\varepsilon(t, x, v)}{|\mathcal{M}^\varepsilon(t, x, v)|^2} dx d\mu(v) \\
&= \frac{1}{\varepsilon} \int \int a(v) \cdot \nabla_x \eta^\varepsilon(t, x) \frac{|f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} d\mu(v) dx.
\end{aligned}$$

Using again the identity $f^\varepsilon = \rho^\varepsilon \mathcal{M} - Lf^\varepsilon$, we get

$$\begin{aligned}
\mathcal{B}_\varepsilon &= \frac{1}{\varepsilon} \int_{\mathbb{T}^d} e^{2\eta^\varepsilon(t, x)} |\rho^\varepsilon(t, x)|^2 \nabla_x \eta^\varepsilon(t, x) \cdot \left(\int_V a(v) \mathcal{M}(v) d\mu(v) \right) dx \\
&\quad + \frac{1}{\varepsilon} \int \int a(v) \cdot \nabla_x \eta^\varepsilon(t, x) \frac{|Lf^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} d\mu(v) dx \\
&\quad - \frac{2}{\varepsilon} \int \int a(v) \cdot \nabla_x \eta^\varepsilon(t, x) \rho^\varepsilon(t, x) Lf^\varepsilon(t, x, v) \frac{\mathcal{M}(v)}{\mathcal{M}^\varepsilon(t, x, v)} d\mu(v) dx \\
&= \mathcal{B}_\varepsilon^1 + \mathcal{B}_\varepsilon^2 + \mathcal{B}_\varepsilon^3.
\end{aligned}$$

These three terms are treated separately:

- By Assumption II.1 $\int_V a(v) \mathcal{M}(v) d\mu(v) = 0$, hence $\mathcal{B}_\varepsilon^1 = 0$.
- Using (II.3.5) and the condition $\varepsilon \leq \varepsilon_0 < \left(4 \|a\|_{L^\infty} (1 + T \|\overline{m}\|_{C_x^1}) \right)^{-1}$, we have for $t \in [0, \tau^\varepsilon]$

$$|\mathcal{B}_\varepsilon^2| \leq \frac{1}{4\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2.$$

- The Young inequality $2|UV| \leq |U|^2 + |V|^2$ with $U = 2a(v) \cdot \nabla_x \eta^\varepsilon(t, x) \rho^\varepsilon(t, x) \mathcal{M}(v)$ and $V = \frac{1}{2\varepsilon} Lf^\varepsilon(t, x, v)$ yields

$$|\mathcal{B}_\varepsilon^3| \leq 4 \operatorname{Tr}(K) \|\nabla_x \eta^\varepsilon(t)\|_{C_x}^2 \int_{\mathbb{T}^d} e^{2\eta^\varepsilon(t, x)} |\rho^\varepsilon(t, x)|^2 dx + \frac{1}{4\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2,$$

where $\operatorname{Tr}(K) \doteq \int_V |a(v)|^2 \mathcal{M}(v) d\mu(v)$. Owing to the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
|\rho^\varepsilon(t, x)|^2 &\leq \int_V \frac{|f^\varepsilon(t, x, v)|^2}{\mathcal{M}^\varepsilon(t, x, v)} d\mu(v) \int_V \mathcal{M}^\varepsilon(t, x, v) d\mu(v) \\
&= \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 e^{-2\eta^\varepsilon(t, x)},
\end{aligned}$$

thus, for $t \in [0, \tau^\varepsilon]$, one has

$$|\mathcal{B}_\varepsilon^3| \leq 4 \operatorname{Tr}(K) (1 + T \|\overline{m}\|_{C_x^1})^2 \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 + \frac{1}{4\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2,$$

where we used again (II.3.5).

Therefore, for $t \in [0, \tau^\varepsilon]$, we get

$$\partial_t \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 \leq -\frac{1}{2\varepsilon^2} \|Lf^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 + 4 \operatorname{Tr}(K)(1 + t \|\overline{m}\|_{C_x^1})^2 \|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2,$$

and owing to Gronwall's Lemma, we obtain the upper estimate

$$\|f^\varepsilon(t)\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 + \int_0^t \frac{1}{2\varepsilon^2} \|Lf^\varepsilon(s)\|_{L^2(\mathcal{M}^\varepsilon(s)^{-1})}^2 ds \leq \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2 e^{4T \operatorname{Tr}(K)(1+T\|\overline{m}\|_{C_x^1})^2}.$$

Note that, for $h \in L^2(\mathcal{M}^{-1})$ and $t \in [0, \tau^\varepsilon]$, one has

$$\begin{aligned} \|h\|_{L^2(\mathcal{M}^\varepsilon(t)^{-1})}^2 &\leq \|h\|_{L^2(\mathcal{M}^{-1})}^2 \exp \left(\sup_{s \in [0, t]} \|2\varepsilon \zeta^\varepsilon(s) + 2t \overline{m}\|_{C_x^0} \right) \\ &\leq \|h\|_{L^2(\mathcal{M}^{-1})}^2 e^{2+2T\|\overline{m}\|_{C_x^0}}, \end{aligned}$$

using (II.2.9). Thus, we get for $t \in [0, \tau^\varepsilon]$

$$\|f^\varepsilon(t)\|_{L^2(\mathcal{M}^{-1})}^2 + \int_0^t \frac{1}{2\varepsilon^2} \|Lf^\varepsilon(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \leq C(T) \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2,$$

with $C(T) = e^{2+2T\|\overline{m}\|_{C_x^0} + 4T \operatorname{Tr}(K)(1+T\|\overline{m}\|_{C_x^1})^2}$. Then, for $t \in [0, T]$, we have $t \wedge \tau^\varepsilon \in [0, \tau^\varepsilon]$ and

$$\begin{aligned} \|f^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}^2 &+ \int_0^{t \wedge \tau^\varepsilon} \frac{1}{2\varepsilon^2} \|Lf^{\varepsilon, \tau^\varepsilon}(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \\ &= \|f^\varepsilon(t \wedge \tau^\varepsilon)\|_{L^2(\mathcal{M}^{-1})}^2 + \int_0^{t \wedge \tau^\varepsilon} \frac{1}{2\varepsilon^2} \|Lf^\varepsilon(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \\ &\leq C(T) \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2, \end{aligned}$$

which concludes the proof. \square

II.3.3 Uniqueness of the solution for the limiting equation

Proposition II.3.4. *For all $T \in (0, \infty)$, the solution of (II.1.6) in the sense of Definition II.2.2 is unique (if it exists).*

Proof. Since equation (II.1.6) is linear, it is sufficient to prove that, if $\rho_0 = 0$, then $\rho(t) = 0$ for all $t \in [0, T]$.

Consider the self-adjoint operator B_K on L_x^2 defined by

$$D(B_K) = H_x^2, \quad B_K \rho = \operatorname{div}(K \nabla \rho) - \overline{m} \rho. \quad (\text{II.3.6})$$

Since K is positive definite (by Assumption II.1), for $\Lambda > \|\overline{m}\|_E$, the operator $B_K - \Lambda \operatorname{id}$ is invertible. Its inverse is a compact operator, owing to the compact embedding $H_x^2 \subset L_x^2$. Therefore, there exists a complete orthonormal system $(e_i)_{i \in \mathbb{N}_0} \in (L_x^2)^{\mathbb{N}_0}$ of eigenvectors for $(B_K - \Lambda \operatorname{id})^{-1}$. Let $(\tilde{\lambda}_i)_{i \in \mathbb{N}_0}$ be the associated (negative) eigenvalues. Thus, $(e_i)_{i \in \mathbb{N}_0}$ are eigenvectors for B_K associated with the eigenvalues $(\lambda_i)_{i \in \mathbb{N}_0}$, where

$$\lambda_i = \Lambda + \tilde{\lambda}_i \leq \Lambda.$$

By Definition II.2.2, $\rho \in L_T^\infty L_x^2$, thus it only remains to prove that, for $i \in \mathbb{N}_0$ and $t \in [0, T]$, $(\rho(t), e_i)_{L_x^2} = 0$.

Fix $i \in \mathbb{N}_0$. Since $B_K^{-1}(L_x^2) \subset H_x^2$, we have $e_i \in H_x^2$. Then (II.2.6) reads

$$(\rho(t), e_i)_{L_x^2} = \lambda_i \int_0^t (\rho(s), e_i)_{L_x^2} ds,$$

hence $(\rho(t), e_i)_{L_x^2} = 0$ for all $t \in [0, T]$. This proves the uniqueness of ρ and concludes the proof of Proposition II.3.4. \square

Note that we do not prove the existence of a solution in this section, since it is a consequence of Theorem II.2.1. More precisely, owing to Proposition II.6.2, the family $(\rho^\varepsilon)_{\varepsilon \in (0, \varepsilon_0]}$ is tight in $\mathcal{C}_T^0 H_x^{-\sigma}$. Thus, this family admits a limit point ρ for the convergence in distribution in $\mathcal{C}_T^0 H_x^{-\sigma}$. Owing to Proposition II.6.3, ρ solves the martingale problem associated with the limiting equation and owing to Proposition II.5.2, it thus solves (II.1.6) in the sense of Definition II.2.2, which concludes to prove the existence of a solution.

II.4 Description of the perturbed test functions method

Let \mathcal{L}^ε be the infinitesimal generator associated with the equation (II.1.1). It is given by

$$\begin{aligned} \mathcal{L}^\varepsilon &= \mathcal{L}_0 + \varepsilon^{-1} \mathcal{L}_1 + \varepsilon^{-2} \mathcal{L}_2, \\ \mathcal{L}_0 \varphi(f, \ell) &= -D_f \varphi(f, \ell) \cdot (\ell f), \\ \mathcal{L}_1 \varphi(f, \ell) &= -D_f \varphi(f, \ell) \cdot (a \cdot \nabla_x f), \\ \mathcal{L}_2 \varphi(f, \ell) &= D_f \varphi(f, \ell) \cdot Lf + \mathcal{L}_m \varphi(f, \ell). \end{aligned}$$

Recall the notation $A = a(v) \cdot \nabla_x$.

First, we introduce two classes of test functions:

- The first class Θ introduced in Definition II.4.1 below, is the class of test functions used to express the martingale problem for fixed $\varepsilon \in (0, \varepsilon_0]$, see Proposition II.5.1. This class satisfies the property: if $\varphi \in \Theta$, then $\varphi \in D(\mathcal{L}^\varepsilon)$ and $\varphi^2 \in D(\mathcal{L}^\varepsilon)$.
- The second class Θ_{lim} introduced in Definition II.4.2 below, is the class of test functions used to characterize the solution of the limit equation by a martingale formulation. Note that $\Theta_{\text{lim}} \subset D(\mathcal{L})$ if \mathcal{L} is the infinitesimal generator of the limit equation.

Second, we construct a perturbed test function $\varphi^\varepsilon \in \Theta$ for any test function $\varphi \in \Theta_{\text{lim}}$, see Proposition II.4.1.

Definition II.4.1. Let Θ be the class of test functions $\varphi : L^2(\mathcal{M}^{-1}) \times E \rightarrow \mathbb{R}$ satisfying the following conditions:

- For $\ell \in E$, $\varphi(\cdot, \ell) \in \mathcal{C}^1(L^2(\mathcal{M}^{-1}))$.
- For $f \in L^2(\mathcal{M}^{-1})$, $\varphi(f, \cdot) \in \mathcal{C}^0(E)$.
- For $i \in \{1, 2\}$ and $f \in L^2(\mathcal{M}^{-1})$, $\varphi(f, \cdot)^i \in D(\mathcal{L}_m)$ and $\mathcal{L}_m(\varphi^i) \in \mathcal{C}^0(L^2(\mathcal{M}^{-1}) \times E)$.

- Consider, for $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, the gradient $\nabla_f \varphi(f, \ell) \in L^2(\mathcal{M}^{-1})$ defined by

$$\forall h \in L^2(\mathcal{M}^{-1}), (\nabla_f \varphi(f, \ell), h)_{L^2(\mathcal{M}^{-1})} = D_f \varphi(f, \ell) \cdot h.$$

Then, for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, we have

$$\int \int \|\nabla_x \nabla_f \varphi(f, \ell)(x, v)\|^2 dx \frac{d\mu(v)}{\mathcal{M}(v)} < \infty.$$

- There exists $C_\varphi(0, \infty)$ such that, for all $f, h \in L^2(\mathcal{M}^{-1})$ and $\ell_1, \ell_2 \in E$,

$$\begin{aligned} & |\varphi(f, \ell_1)| + |D_f \varphi(f, \ell_1)(Ah)| + |D_f \varphi(f, \ell_1)(\ell_2 h)| + |D_f \varphi(f, \ell_1)(Lh)| \\ & \leq C_\varphi \left(1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3 + \|h\|_{L^2(\mathcal{M}^{-1})}^3\right) (1 + \|\ell_1\|_E^2 + \|\ell_2\|_E^2). \end{aligned}$$

Definition II.4.2. Let Θ_{\lim} be the class of test functions φ such that

$$\varphi(f, \ell) = \varphi(\rho) = \chi((\rho, \xi)_{L_x^2}),$$

with $\chi \in \mathcal{C}_b^3(\mathbb{R})$, $\xi \in \mathcal{C}_x^3$ and $\rho = \langle f \rangle = \int_V f d\mu$.

Observe that $\Theta_{\lim} \subset \Theta$.

The main result of this section is the construction of the perturbed test function $\varphi^\varepsilon \in \Theta$ for any $\varphi \in \Theta_{\lim}$, and the identification of the infinitesimal generator \mathcal{L} associated with the limit equation.

To simplify the expression of estimates, let us introduce a notation: for any variable or collection of variables u , and for any non-negative quantities x and y , $x \lesssim_u y$ means that there exists $C \in (0, \infty)$ such that $x \leq Cy$ where C depends only on u and on quantities introduced in Section II.2.2 (namely a , \mathcal{M} , \mathcal{L}_m , ν , \overline{m} , ε_0 , γ , α , γ_{mix} and the moments of f_0^ε uniform in ε). In particular, C does not depend on ε .

Proposition II.4.1. Consider the operator on $\mathcal{C}^0(L_x^2)$

$$\mathcal{L}\varphi(\rho) = D\varphi(\rho) \cdot (\text{div}_x(K\nabla_x \rho) - \overline{m}\rho), \quad D(\mathcal{L}) = \{\varphi \in \mathcal{C}^0(L_x^2) \mid \mathcal{L}\varphi \in \mathcal{C}^0(L_x^2)\}. \quad (\text{II.4.1})$$

Then $\Theta_{\lim} \subset D(\mathcal{L})$ and, for all $\varphi \in \Theta_{\lim}$, there exists functions φ_1 and φ_2 such that, with

$$\varphi^\varepsilon = \varphi + \varepsilon \varphi_1 + \varepsilon^2 \varphi_2, \quad (\text{II.4.2})$$

we have for all $(f, \ell) \in L^2(\mathcal{M}^{-1}) \times E$,

$$|\varphi_1(f)| \lesssim_\varphi \|f\|_{L^2(\mathcal{M}^{-1})} \quad (\text{II.4.3})$$

$$|\varphi_2(f, \ell)| \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \|\ell\|_E) \quad (\text{II.4.4})$$

$$|\mathcal{L}^\varepsilon \varphi^\varepsilon - \mathcal{L}\varphi|(f, \ell) \lesssim_\varphi \varepsilon(1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|\ell\|_E) + \varepsilon^2(1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|\ell\|_E^2). \quad (\text{II.4.5})$$

Moreover, $\varphi^\varepsilon \in \Theta$.

Proof. To prove Proposition II.4.1, we write

$$\mathcal{L}^\varepsilon \varphi^\varepsilon = \varepsilon^{-2} \mathcal{L}_2 \varphi + \varepsilon^{-1} (\mathcal{L}_2 \varphi_1 + \mathcal{L}_1 \varphi) + (\mathcal{L}_2 \varphi_2 + \mathcal{L}_1 \varphi_1 + \mathcal{L}_0 \varphi) + \varepsilon (\mathcal{L}_1 \varphi_2 + \mathcal{L}_0 \varphi_1) + \varepsilon^2 (\mathcal{L}_0 \varphi_2).$$

Therefore, we want for all $(f, \ell) \in L^2(\mathcal{M}^{-1}) \times E$

$$\mathcal{L}_2 \varphi = 0, \quad (\text{II.4.6})$$

$$\mathcal{L}_2 \varphi_1 + \mathcal{L}_1 \varphi = 0, \quad (\text{II.4.7})$$

$$\mathcal{L}_2 \varphi_2 + \mathcal{L}_1 \varphi_1 + \mathcal{L}_0 \varphi = \mathcal{L} \varphi, \quad (\text{II.4.8})$$

$$|\mathcal{L}_1 \varphi_2(f, \ell) + \mathcal{L}_0 \varphi_1(f, \ell)| \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|\ell\|_E), \quad (\text{II.4.9})$$

$$|\mathcal{L}_0 \varphi_2(f, \ell)| \lesssim_\varphi (1 + \|f\|_{L^2(\mathcal{M}^{-1})}^3)(1 + \|\ell\|_E^2). \quad (\text{II.4.10})$$

One can already check that $\mathcal{L}_2 \varphi = 0$. Indeed, $\int_V L \rho d\mu = 0$ and $\mathcal{L}_m \varphi = 0$ since $\varphi \in \Theta_{\text{lim}}$ does not depend on ℓ .

II.4.1 Poisson equation and auxiliary tools

To define φ_1 and φ_2 using (II.4.7) and (II.4.8), we need to solve Poisson equations of the form $\mathcal{L}_2 \psi = -\vartheta$. For this purpose, let us introduce the process $(g_f(t), m_\ell(t))_{t \in \mathbb{R}^+}$ associated with the infinitesimal generator \mathcal{L}_2 , with the initial conditions $g_f(0) = f$ and $m_\ell(0) = \ell$. We have

$$g_f(t) = \rho \mathcal{M} + e^{-t} (f - \rho \mathcal{M}).$$

As a consequence, the density ρ is preserved along the trajectories of g_f (namely, for all $t \in \mathbb{R}^+$, $\langle g_f(t) \rangle = \rho$).

Moreover, owing to Assumption II.1, we have $\langle A g_f(t) \rangle = e^{-t} \langle A f \rangle$ for $f \in L^2(\mathcal{M}^{-1})$.

Under the centering condition

$$\int \vartheta d\mu_\rho = 0 \quad (\text{II.4.11})$$

where $\mu_\rho = \delta_{\rho \mathcal{M}} \otimes \nu$ is the invariant distribution of \mathcal{L}_2 on the space

$$\left\{ (f, \ell) \in L^2(\mathcal{M}^{-1}) \times E \mid \int_V f d\mu = \rho \right\},$$

the solution of the Poisson equation $\mathcal{L}_2 \psi = -\vartheta$ is given by

$$\psi(f, \ell) = \int_0^\infty \mathbb{E} [\vartheta(g_f(t), m_\ell(t))] dt. \quad (\text{II.4.12})$$

Let us introduce a particular class of functions: $\vartheta(f, \ell) = \theta_{h,k}(\ell) - \theta_{h,k}(\overline{m})$, where the linear form $\theta_{h,k} \in E^*$ is defined by $\theta_{h,k}(\ell) = (h\ell, k)_{L^2}$, with $h, k \in L^2$. In this case, one can define $\psi_{h,k} = R_0 \vartheta$ as in Definition II.2.1.

For all $\ell \in E$, $\psi_{h,\cdot}(\ell)$ is a linear form on L^2 . Owing to Riesz Representation Theorem, there exists a bounded operator $R_0(\ell) : L^2 \rightarrow L^2$ such that, for all $\ell \in E$ and $h, k \in L^2$, $\psi_{h,k}(\ell) = (R_0(\ell)h, k)_{L^2}$.

Property II.4.2. *The resolvent operator satisfies the following properties*

- For $\ell \in E$, $R_0(\ell)$ is self-adjoint. Indeed, we have $\theta_{h,k} = \theta_{k,h}$, hence $\psi_{h,k}(\ell) = \psi_{k,h}(\ell)$, namely

$$(R_0(\ell)h, k)_{L^2} = (h, R_0(\ell)k)_{L^2}.$$

- Estimate (II.2.3) with $\theta_{h,k}(\ell) \leq \|h\|_{L^2} \|k\|_{L^2} \|\ell\|_E$ reads $|\psi_{h,k}(\ell)| \lesssim \|h\|_{L^2} \|k\|_{L^2} (1 + \|\ell\|_E)$, hence

$$\|R_0(\ell)h\|_{L^2} \lesssim \|h\|_{L^2} (1 + \|\ell\|_E).$$

- If $h \in H_x^1$ and $k \in L_x^2$, we have $\theta_{h,k}(\ell) \leq \|h\|_{H_x^1} \|k\|_{H_x^{-1}} \|\ell\|_E$ since $E \subset \mathcal{C}_x^1$. Then, estimate (II.2.3) reads $|\psi_{h,k}(\ell)| \lesssim \|h\|_{H_x^1} \|k\|_{H_x^{-1}} (1 + \|\ell\|_E)$. Thus, $R_0(\ell)h \in H_x^1$ and

$$\|R_0(\ell)h\|_{H_x^1} \lesssim \|h\|_{H_x^1} (1 + \|\ell\|_E).$$

The last tool we need is the Cauchy-Schwarz inequality: to obtain estimates depending on the $L^2(\mathcal{M}^{-1})$ -norm, we write for $i \in \{1, 2, 3\}$, $h \in \mathcal{C}^i(\mathbb{T}^d)$ and $f \in L^2(\mathcal{M}^{-1})$,

$$|(f, A^i h)_{L^2}| \leq \|f\|_{L^2(\mathcal{M}^{-1})} \|A^i h\|_{L^2(\mathcal{M})},$$

where $\|A^i h\|_{L^2(\mathcal{M})}^2 \doteq \iint \left((a(v) \cdot \nabla_x)^i h(x) \right)^2 \mathcal{M}(v) d\mu(v) dx < \infty$ by Assumption II.1.

II.4.2 Construction of the first corrector

To shorten the computations, since $\varphi \in \Theta_{\text{lim}}$ is fixed, as well as χ and ξ as in Definition II.4.2, we write $\chi_\rho = \chi((\rho, \xi)_{L_x^2})$, $\chi'_\rho = \chi'((\rho, \xi)_{L_x^2})$, etc...

We want to solve the Poisson equation $\mathcal{L}_2 \varphi_1 = -\mathcal{L}_1 \varphi$ (see (II.4.7)), where

$$\begin{aligned} \mathcal{L}_1 \varphi(f, \ell) &= -\chi'_\rho(\langle Af \rangle, \xi)_{L_x^2} \\ &= \chi'_\rho(f, A\xi)_{L^2}. \end{aligned}$$

The centering condition (II.4.11) is satisfied, we then define φ_1 as follows: for all $(f, \ell) \in L^2(\mathcal{M}^{-1}) \times E$,

$$\varphi_1(f, \ell) = \int_0^\infty \mathbb{E}[\mathcal{L}_1 \varphi(g_f(t), m_\ell(t))] dt.$$

Recall that $\langle Ag_f(t) \rangle = e^{-t} \langle Af \rangle$, hence

$$\begin{aligned} \varphi_1(f) &= - \int_0^\infty e^{-t} \chi'_\rho(\langle Af \rangle, \xi)_{L_x^2} dt \\ &= \chi'_\rho(f, A\xi)_{L^2}, \end{aligned} \tag{II.4.13}$$

where we used that $\chi_\rho^{(i)}$ does not depend on t when evaluated at $g_f(t)$ since for all $t \in \mathbb{R}^+$, $\langle g_f(t) \rangle = \rho$.

It is straightforward to check that estimate (II.4.3) is satisfied. This concludes the construction of the first corrector φ_1 .

II.4.3 Construction of the second corrector

In (II.4.8), the centering condition (II.4.11) is not satisfied. Therefore, we want to solve the Poisson equation $\mathcal{L}_2 \varphi_2 = -(\mathcal{L}_0 \varphi + \mathcal{L}_1 \varphi_1 - \int (\mathcal{L}_0 \varphi + \mathcal{L}_1 \varphi_1) d\mu_\rho)$. To ensure (II.4.8), we need $\mathcal{L} \varphi(\rho) = \int (\mathcal{L}_0 \varphi + \mathcal{L}_1 \varphi_1) d\mu_\rho$. We have for $(f, \ell) \in L^2(\mathcal{M}^{-1}) \times E$

$$\begin{aligned} \mathcal{L}_0 \varphi(f, \ell) + \mathcal{L}_1 \varphi_1(f) &= -\chi'_\rho(\ell \rho, \xi)_{L_x^2} - \chi''_\rho(\langle Af \rangle, \xi)_{L_x^2} (f, A\xi)_{L^2} - \chi'_\rho(Af, A\xi)_{L^2} \\ &= -\chi'_\rho(\ell \rho, \xi)_{L_x^2} + \chi''_\rho(f, A\xi)_{L^2}^2 + \chi'_\rho(f, A^2 \xi)_{L^2}. \end{aligned}$$

Since $\int_V a(v)\mathcal{M}(v)d\mu(v) = 0$ by Assumption II.1, we have $\int (f, A\xi)_{L^2} d\mu_\rho = (\rho\mathcal{M}, A\xi)_{L^2} = 0$. Moreover, owing to (II.1.4), $\int_V \mathcal{M}A^2\xi d\mu = K : \nabla_x^2 \xi = \operatorname{div}_x(K\nabla_x \xi)$. The limit infinitesimal generator \mathcal{L} is then defined for $\rho \in L_x^2$ by

$$\begin{aligned}\mathcal{L}\varphi(\rho) &= -\chi'_\rho(\overline{m}\rho, \xi)_{L_x^2} + \chi'_\rho(\rho\mathcal{M}, A^2\xi)_{L^2} \\ &= -\chi'_\rho(\overline{m}\rho, \xi)_{L_x^2} + \chi'_\rho(\rho, K : \nabla_x^2 \xi)_{L_x^2} \\ &= \chi'_\rho(\operatorname{div}_x(K\nabla_x \rho) - \overline{m}\rho, \xi)_{L_x^2},\end{aligned}\tag{II.4.14}$$

which leads to the definition (II.4.1).

Then, we define φ_2 as the solution of the Poisson equation $\mathcal{L}_2\varphi_2 = -(\mathcal{L}_1\varphi_1 + \mathcal{L}_0\varphi - \mathcal{L}\varphi)$ satisfying the centering condition $\int (\mathcal{L}_1\varphi_1 + \mathcal{L}_0\varphi - \mathcal{L}\varphi) d\mu_\rho = 0$. Equation (II.4.12) reads

$$\varphi_2(f, \ell) = \int_0^\infty \mathbb{E} [\mathcal{L}_0\varphi(g_f(t), m_\ell(t) + \mathcal{L}_1\varphi_1(g_f(t), m_\ell(t)) - \mathcal{L}\varphi(\rho)] dt.$$

Let us obtain the following expression for φ_2 :

$$\varphi_2(f, \ell) = \frac{1}{2}\chi''_\rho(f, A\xi)_{L^2}^2 + \chi'_\rho(f - \rho\mathcal{M}, A^2\xi)_{L^2} - \chi'_\rho(\rho, R_0(\ell)\xi)_{L_x^2}.\tag{II.4.15}$$

The corrector φ_2 can be decomposed as $\varphi_{2,(i)} + \varphi_{2,(ii)} + \varphi_{2,(iii)}$, corresponding to the contributions of the different terms of $\mathcal{L}_0\varphi + \mathcal{L}_1\varphi_1$. For each of these contributions, recall that the quantities $\chi_\rho^{(i)}$ do not depend on t .

(i) The term $-\chi'_\rho(\ell\rho, \xi)_{L_x^2}$ gives

$$\varphi_{2,(i)}(f, \ell) = -\int_0^\infty \mathbb{E} \left[\chi'_\rho((m_\ell(t) - \overline{m})\rho, \xi)_{L_x^2} \right] dt = -\chi'_\rho(R_0(\ell)\rho, \xi)_{L_x^2},$$

where $R_0(\ell)$ is the operator defined in Section II.4.1.

(ii) The term $\chi''_\rho(f, A\xi)_{L^2}^2 = \chi''_\rho(\langle Af \rangle, \xi)_{L_x^2}^2$ gives, using the identity $\langle Ag_f(t) \rangle = e^{-t} \langle Af \rangle$,

$$\begin{aligned}\varphi_{2,(ii)}(f, \ell) &= \int_0^\infty e^{-2t} \chi''_\rho(\langle Af \rangle, \xi)_{L^2}^2 dt \\ &= \int_0^\infty e^{-2t} \chi''_\rho(f, A\xi)_{L^2}^2 dt \\ &= \frac{1}{2} \chi''_\rho(f, A\xi)_{L^2}^2.\end{aligned}$$

(iii) The term $\chi'_\rho(f, A^2\xi)_{L^2}$ gives

$$\begin{aligned}\varphi_{2,(iii)}(f, \ell) &= \int_0^\infty \chi'_\rho(g_f(t) - \rho\mathcal{M}, A^2\xi)_{L^2} dt \\ &= \int_0^\infty e^{-t} \chi'_\rho(f - \rho\mathcal{M}, A^2\xi)_{L^2} dt \\ &= \chi'_\rho(f - \rho\mathcal{M}, A^2\xi)_{L^2}.\end{aligned}$$

It is straightforward to check that φ_2 satisfies (II.4.4). This concludes the construction of the second corrector φ_2 .

II.4.4 Verification of (II.4.5)

Straightforward computations give the following expressions for $\mathcal{L}_1\varphi_2$, $\mathcal{L}_0\varphi_1$ and $\mathcal{L}_0\varphi_2$:

$$\begin{aligned}\mathcal{L}_1\varphi_2(f, \ell) &= \frac{1}{2}\chi_\rho'''(f, A\xi)_{L^2}^3 + \chi_\rho''(f, A\xi)_{L^2}(f, A^2\xi)_{L^2} + \chi_\rho''(f - \rho\mathcal{M}, A^2\xi)_{L^2}(f, A\xi)_{L^2} \\ &\quad + \chi_\rho'(f, A^3\xi)_{L^2} - \chi_\rho'(f, A(K : \nabla_x^2\xi))_{L^2} + \chi_\rho''(\rho, R_0(\ell)\xi)_{L^2}(f, A\xi)_{L^2} \\ &\quad - \chi_\rho'(f, A(R_0(\ell)\xi))_{L^2}, \\ \mathcal{L}_0\varphi_1(f, \ell) &= -\chi_\rho''(f, A\xi)_{L^2}(\rho, \ell\xi)_{L_x^2} - \chi_\rho'(f, \ell A\xi)_{L^2} \\ \mathcal{L}_0\varphi_2(f, \ell) &= -\frac{1}{2}\chi_\rho'''(f, A\xi)_{L^2}^2(\rho, \ell\xi)_{L_x^2} - \chi_\rho''(f, \ell A\xi)_{L^2}^2 - \chi_\rho''(f - \rho\mathcal{M}, A^2\xi)_{L^2}(\rho, \ell\xi)_{L_x^2} \\ &\quad - \chi_\rho'(f - \rho\mathcal{M}, \ell A^2\xi)_{L^2} + \chi_\rho''(\rho, R_0(\ell)\xi)_{L_x^2}(\rho, \ell\xi)_{L_x^2} + \chi_\rho'(\rho, \ell R_0(\ell)\xi)_{L_x^2}.\end{aligned}$$

One can then check that (II.4.9) and (II.4.10) are satisfied, hence so is (II.4.5).

Finally, owing to Assumption II.6, one can also check that $\varphi^\varepsilon = \varphi + \varepsilon\varphi_1 + \varepsilon^2\varphi_2 \in \Theta$. This concludes the proof of Proposition II.4.1. \square

II.5 Martingale problems

The proof of Theorem II.2.1 heavily relies on the notion of martingale problems as introduced in [SV06]. To identify a limit point of $(\rho^\varepsilon)_{\varepsilon \in (0, \varepsilon_0]}$ for the convergence in distribution, we characterize its distribution by a family of martingales (see Proposition II.5.1 and take the limit when $\varepsilon \rightarrow 0$ in their martingale properties. The limit martingale characterizes the distribution of the limit point (see Proposition II.5.2).

The martingale formulation for fixed ε is also used in combination with martingale arguments (for instance Doob Maximal Inequality) to prove Propositions II.6.1 and II.6.2.

In this Section, we state these martingale formulations and prove the martingale formulation for the limit equation. The proof of the martingale formulation for fixed ε is very technical, due to the presence of a stopping time, but is very similar to the classical proof of the martingale characterization of a Markov process (see [EK86, SV06, DRV20]). See the proof of Proposition I.4.1 for details on how to manage the stopping times.

II.5.1 Martingale problem for the equation at fixed ε

Recall that the stopping time τ^ε is introduced in Definition II.2.3.

Proposition II.5.1. *Let $\varphi \in \Theta$ and for $t \in \mathbb{R}^+$, set*

$$M_\varphi^\varepsilon(t) \doteq \varphi(f^\varepsilon(t), m^\varepsilon(t)) - \varphi(f^\varepsilon(0), m^\varepsilon(0)) - \int_0^t \mathcal{L}^\varepsilon\varphi(f^\varepsilon(s), m^\varepsilon(s))ds.$$

Then, $M_\varphi^{\varepsilon, \tau^\varepsilon}$ is a càdlàg $(\mathcal{F}_t^\varepsilon)_{t \in \mathbb{R}^+}$ -martingale and

$$\begin{aligned}\forall t \in \mathbb{R}^+, \mathbb{E} \left[|M_\varphi^{\varepsilon, \tau^\varepsilon}(t)|^2 \right] &= \mathbb{E} \left[\int_0^{t \wedge \tau^\varepsilon} (\mathcal{L}^\varepsilon(\varphi^2) - 2\varphi\mathcal{L}^\varepsilon\varphi)(f^\varepsilon(s), m^\varepsilon(s))ds \right] \\ &= \frac{1}{\varepsilon^2} \mathbb{E} \left[\int_0^{t \wedge \tau^\varepsilon} (\mathcal{L}_m(\varphi^2) - 2\varphi\mathcal{L}_m\varphi)(f^\varepsilon(s), m^\varepsilon(s))ds \right].\end{aligned}$$

II.5.2 An equivalent formulation for the limit equation

Proposition II.5.2. *Let $\sigma \in (0, 1]$ be arbitrarily small and let ρ be a $\mathcal{C}_T^0 H_x^{-\sigma}$ -valued stochastic process.*

Then, ρ is a weak solution of (II.1.6) in L_x^2 (in the sense of Definition II.2.2) if, and only if, for all $\varphi \in \Theta_{\text{lim}}$, we have almost surely

$$\forall t \in [0, T], M_\varphi(t) = \varphi(\rho(t)) - \varphi(\rho_0) - \int_0^t \mathcal{L}\varphi(\rho(s)) ds = 0.$$

Proof. By definition of the infinitesimal generator \mathcal{L} (see (II.4.1)), and using the notation $\varphi = \chi((\cdot, \xi)_{L_x^2})$ as in Definition II.4.2, we have

$$M_\varphi(t) = \chi((\rho(t), \xi)_{L_x^2}) - \chi((\rho_0, \xi)_{L_x^2}) - \int_0^t \chi'((\rho(s), \xi)_{L_x^2}) (\rho(s), \text{div}(K \nabla_x \xi) - \overline{m} \xi)_{L_x^2} ds.$$

If ρ is a weak solution of (II.1.6) in L_x^2 , then for $\varphi \in \Theta_{\text{lim}}$, we have $\mathcal{L}\varphi(\rho(t)) = \partial_t(\varphi(\rho(t)))$, and $\rho(0) = \rho_0$ a.s., hence $M_\varphi = 0$ a.s.

Conversely, assume $M_\varphi = 0$ for all $\varphi \in \Theta_{\text{lim}}$. For $r \in \mathbb{R}^+$, let χ^r be defined by

$$\begin{aligned} \forall u \in [0, r], \chi^r(u) &= u, \\ \forall u \in [r+1, \infty), \chi^r(u) &= r+1, \\ \forall u \in (r, r+1), \chi^r(u) &= P^r(u), \\ \forall u \in (-\infty, 0), \chi^r(u) &= -\chi^r(-u), \end{aligned}$$

where P^r is a polynomial function such that $\chi^r \in \mathcal{C}_b^3(\mathbb{R})$. Let $\xi \in H_x^2$ and consider $\varphi^r \in \Theta_{\text{lim}}$ defined by

$$\varphi_r(\rho) = \chi^r((\rho, \xi)_{L_x^2}).$$

Since $\rho \in \mathcal{C}_T^0 H_x^{-\sigma}$ a.s., the random variable

$$S \doteq \sup_{t \in [0, T]} |(\rho(t), \xi)_{L_x^2}|,$$

is finite a.s. and $\varphi_S(\rho) = (\rho(t), \xi)_{L_x^2}$. Since $M_{\varphi_S}(t) = 0$, we get that ρ satisfies almost surely (II.2.6), for $\xi \in H_x^2$.

To satisfy Definition II.2.2, it only remains to prove that $\rho \in L_T^\infty L_x^2$ almost surely.

Consider the self-adjoint operator B_K on L_x^2 defined by (II.3.6). As proved in Section II.3.3, there exists a constant $\Lambda \in \mathbb{R}^+$ and a complete orthonormal system $(e_i)_{i \in \mathbb{N}_0}$ of L_x^2 such that, for all $i \in \mathbb{N}_0$, $B_K e_i = \lambda_i e_i$ with $\lambda_i \leq \Lambda$. Fix $i \in \mathbb{N}_0$. Since $e_i \in H_x^2$, (II.2.6) reads for $t \in [0, T]$

$$(\rho(t), e_i)_{L_x^2} = (\rho_0, e_i)_{L_x^2} + \lambda_i \int_0^t (\rho(s), e_i)_{L_x^2} ds.$$

Therefore, we have for $t \in [0, T]$

$$(\rho(t), e_i)_{L_x^2} = e^{\lambda_i t} (\rho_0, e_i)_{L_x^2}.$$

Since $\lambda_i t \leq \Lambda T$ and $\rho_0 \in L_x^2$, $|(\rho(t), e_i)_{L_x^2}|^2$ is summable. We thus have $\rho(t) \in L_x^2$ and

$$\|\rho(t)\|_{L_x^2}^2 \leq e^{2\Lambda T} \|\rho_0\|_{L_x^2}^2.$$

This proves that $\rho \in L_T^\infty L_x^2$, and that ρ satisfies Definition II.2.2, which concludes the proof of Proposition II.5.2. \square

II.6 Convergence in $\mathcal{C}^0([0, T], H^{-\sigma}(\mathbb{T}^d))$

II.6.1 Asymptotic behavior of the stopping time

We are now in position to study the asymptotic behavior of the stopping time τ^ε introduced in Definition II.2.3.

Proposition II.6.1. *For all $T > 0$, we have*

$$\mathbb{P}(\tau^\varepsilon < T) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Remark II.6.1 (about the proof). To prove Proposition II.6.1, we establish the following estimate:

$$\sup_\varepsilon \mathbb{E} \left[\varepsilon^{2p} \sup_{t \in [0, T]} \|\zeta^{\varepsilon, \tau^\varepsilon}(t)\|_{\mathcal{C}_x^1}^2 \right] < \infty,$$

for $p \in (0, 1)$. It is possible to prove this result with $p = 0$, but it requires more intricate arguments (see the proof of Proposition I.6.2) and is not necessary to prove Proposition II.6.1.

Proof. Recall that $\tau^\varepsilon = \tau_m^\varepsilon \wedge \tau_\zeta^\varepsilon$ (see Definition II.2.3). Owing to Lemma II.3.1, it only remains to prove that

$$\mathbb{P}(\tau_\zeta^\varepsilon < T) \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (\text{II.6.1})$$

Owing to the Markov inequality and the continuous embedding $H_x^{[d/2]+2} \subset \mathcal{C}_x^1$, we have for $p \in (0, 1)$

$$\begin{aligned} \mathbb{P}(\tau_\zeta^\varepsilon < T) &\leq \mathbb{P}(\tau_\zeta^\varepsilon < T, \tau_\zeta^\varepsilon < \tau_m^\varepsilon) + \mathbb{P}(\tau_\zeta^\varepsilon < T, \tau_\zeta^\varepsilon \geq \tau_m^\varepsilon) \\ &\leq \mathbb{P}\left(\varepsilon^p \sup_{t \in [0, T]} \|\zeta^{\varepsilon, \tau^\varepsilon}(t)\|_{\mathcal{C}_x^1} \geq \frac{1}{\varepsilon^{1-p}}\right) + \mathbb{P}(\tau_m^\varepsilon < T) \\ &\leq \varepsilon^{2(1-p)} \sup_\varepsilon \mathbb{E} \left[\varepsilon^{2p} \sup_{t \in [0, T]} \|\zeta^{\varepsilon, \tau^\varepsilon}(t)\|_{\mathcal{C}_x^1}^2 \right] + \mathbb{P}(\tau_m^\varepsilon < T) \\ &\leq \varepsilon^{2(1-p)} \sup_\varepsilon \mathbb{E} \left[\varepsilon^{2p} \sup_{t \in [0, T]} \|\zeta^{\varepsilon, \tau^\varepsilon}(t)\|_{H_x^{[d/2]+2}}^2 \right] + \mathbb{P}(\tau_m^\varepsilon < T). \end{aligned}$$

Owing to Lemma II.3.1, we have $\mathbb{P}(\tau_m^\varepsilon < T) \xrightarrow{\varepsilon \rightarrow 0} 0$. Since $\zeta^\varepsilon(t) \in E \subset H_x^{[d/2]+2}$, it only remains to prove that for any multi-index β of length $|\beta| \leq [d/2] + 2$, we have

$$\sup_\varepsilon \mathbb{E} \left[\varepsilon^{2p} \sup_{t \in [0, T]} \left\| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t)}{\partial x^\beta} \right\|_{L_x^2}^2 \right] < \infty,$$

for a real number $p \in (0, 1)$ to be chosen.

Fix any such β and $\varepsilon \in (0, \varepsilon_0]$. We have

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left\| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t)}{\partial x^\beta} \right\|_{L_x^2}^2 \right] \leq \int_{\mathbb{T}^d} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \frac{\partial^{|\beta|} \zeta^{\varepsilon, \tau^\varepsilon}(t, x)}{\partial x^\beta} \right|^2 \right] dx. \quad (\text{II.6.2})$$

For $x \in \mathbb{T}^d$, consider the linear form $\theta_x^\beta \in E^*$ defined by

$$\forall \ell \in E, \theta_x^\beta(\ell) = \frac{\partial^{|\beta|} \ell}{\partial x^\beta}(x).$$

Since $m^\varepsilon \in E$ almost surely, one has

$$\frac{\partial^{|\beta|} \zeta^\varepsilon(t, x)}{\partial x^\beta} = \frac{1}{\varepsilon} \int_0^t \frac{\partial^{|\beta|} (m^\varepsilon(s, x) - \bar{m})}{\partial x^\beta} ds = \frac{1}{\varepsilon} \int_0^t (\theta_x^\beta(m^\varepsilon(s) - \bar{m})) ds.$$

Owing to Assumption II.6, the function $\psi_x^\beta \doteq R_0(\theta_x^\beta - \theta_x^\beta(\bar{m}))$ as in Definition II.2.1 satisfies $\psi_x^\beta \in \Theta$ and $\mathcal{L}_m \psi_x^\beta = -(\theta_x^\beta - \theta_x^\beta(\bar{m}))$. Thus, owing to Proposition II.5.1,

$$\begin{aligned} M_{\varepsilon^{1+p}\psi_x^\beta}^\varepsilon(t) &= \varepsilon^{1+p} \psi_x^\beta(m^\varepsilon(t)) - \varepsilon^{1+p} \psi_x^\beta(m(0)) - \frac{1}{\varepsilon^2} \int_0^t \varepsilon^{1+p} \mathcal{L}_m \psi_x^\beta(m^\varepsilon(s)) ds \\ &= \varepsilon^{1+p} \psi_x^\beta(m^\varepsilon(t)) - \varepsilon^{1+p} \psi_x^\beta(m(0)) + \varepsilon^p \frac{\partial^{|\beta|} \zeta^\varepsilon(t, x)}{\partial x^\beta} \end{aligned} \quad (\text{II.6.3})$$

defines a martingale when stopped at τ^ε .

On the one hand, owing to estimates (II.2.3) and (II.2.8), and using that $\alpha < 1$ in Definition II.2.3, we have

$$\mathbb{E} \left[\sup_{t \in [0, T]} |\varepsilon^{1+p} \psi_x^\beta(m^{\varepsilon, \tau^\varepsilon}(t))|^2 \right] \lesssim \varepsilon^{2+2p} (1 + \varepsilon^{-2\alpha}) \lesssim 1.$$

On the other hand, Doob Maximal Inequality yields

$$\mathbb{E} \left[\sup_{t \in [0, T]} |M_{\varepsilon^{1+p}\psi_x^\beta}^{\varepsilon, \tau^\varepsilon}(t)|^2 \right] \leq 4 \mathbb{E} \left[|M_{\varepsilon^{1+p}\psi_x^\beta}^{\varepsilon, \tau^\varepsilon}(T)|^2 \right].$$

Owing to Proposition II.5.1 and Assumption II.6, we get

$$\begin{aligned} \mathbb{E} \left[|M_{\varepsilon^{1+p}\psi_x^\beta}^{\varepsilon, \tau^\varepsilon}(T)|^2 \right] &= \frac{1}{\varepsilon^2} \mathbb{E} \left[(\varepsilon^{1+p})^2 \int_0^{T \wedge \tau^\varepsilon} \left(\mathcal{L}_m \left((\psi_x^\beta)^2 \right) - 2\psi_x^\beta \mathcal{L}_m \psi_x^\beta \right) (m^\varepsilon(s)) ds \right] \\ &\lesssim_T \varepsilon^{2p} (1 + \varepsilon^{-2\alpha}), \end{aligned}$$

using once again the estimate (II.2.3) and (II.2.8). We now choose $p \in (\alpha, 1)$ (recall that $\alpha < 1$ in Definition II.2.3). Then one has $\mathbb{E} \left[|M_{\varepsilon^{1+p}\psi_x^\beta}^{\varepsilon, \tau^\varepsilon}(T)|^2 \right] \lesssim 1$. Therefore, (II.6.3) yields

$$\mathbb{E} \left[\varepsilon^{2p} \sup_{t \in [0, T]} \left| \frac{\partial^{|\beta|} \zeta^\varepsilon(t, x)}{\partial x^\beta} \right|^2 \right]^{1/2} \lesssim_T 1,$$

which concludes the proof by integrating with respect to x and using (II.6.2). \square

Remark II.6.2. The regularity of order $\lfloor d/2 \rfloor + 2$ of $m_\ell(t)$ is used to prove Proposition II.6.1. In all other proofs, we only need $m_\ell(t) \in \mathcal{C}_x^1$.

II.6.2 Tightness

This section is devoted to proving the following result.

Proposition II.6.2. *The family of processes $(\rho^\varepsilon)_{\varepsilon \in (0, \varepsilon_0]}$ is tight in the space $\mathcal{C}_T^0 H_x^{-\sigma}$ for all arbitrarily small $\sigma \in (0, 1]$.*

Proof. Owing to Slutsky's Lemma [Bil99, Theorem 4.1] and Proposition II.6.1, Proposition II.6.2 is equivalent to the tightness of $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ in the same space, and the convergence in distribution of $(\rho^\varepsilon)_{\varepsilon \in (0, \varepsilon_0]}$ is equivalent to the convergence in distribution of $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ to the same limit. Therefore, we only prove the tightness of the family of stopped processes $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$.

Let us introduce auxiliary tools. Consider the Skorokhod space $D_T H_x^{-\sigma}$, namely the space of $H_x^{-\sigma}$ -valued càdlàg functions on $[0, T]$. For $X \in \mathcal{C}_T^0 H_x^{-\sigma}$, the modulus of continuity w_X for continuous functions and the modulus of continuity w'_X for càdlàg functions are defined for $\delta \in [0, T]$ by

$$w_X(\delta) \doteq \sup_{0 \leq t \leq s \leq t + \delta \leq T} \|X(s) - X(t)\|$$

$$w'_X(\delta) \doteq \sup_{(t_i)_i} \max_i \sup_{t_i \leq t \leq s < t_{i+1}} \|X(s) - X(t)\|,$$

where $(t_i)_i$ denotes any finite subdivision of $[0, T]$. For $\delta \in [0, T]$, [Bil99, equation (14.11)] yields

$$w_X(\delta) \leq 2w'_X(\delta).$$

Thus, it suffices to establish the tightness of $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ in $D_T H_x^{-\sigma}$ to prove Proposition II.6.2 (see for instance [Bil99, Theorems 8.2 and 15.2] which gives sufficient and necessary conditions for tightness in $\mathcal{C}_T^0 H_x^{-\sigma}$ and $D_T H_x^{-\sigma}$).

Tightness in $D_T H_x^{-\sigma}$ is easier to prove than in $\mathcal{C}_T^0 H_x^{-\sigma}$, owing to [Jak86, Theorem 3.1]: since Θ_{lim} is closed under addition and separates points, tightness in the Skorokhod space is equivalent to the following claims:

- (i) For all $\eta \in (0, 1]$, there exists a compact set $K_\eta \subset H_x^{-\sigma}$ such that, for all $\varepsilon \in (0, \varepsilon_0]$,

$$\mathbb{P}(\forall t \in [0, T], \rho^{\varepsilon, \tau^\varepsilon}(t) \in K_\eta) > 1 - \eta.$$

- (ii) For all $\varphi \in \Theta_{\text{lim}}$, $(\varphi(\rho^{\varepsilon, \tau^\varepsilon}))_{\varepsilon \in (0, \varepsilon_0]}$ is tight in $D([0, T], \mathbb{R})$.

Claim (i) is a consequence of Proposition II.3.3, of the Markov inequality and of the compact embedding $L_x^2 \subset H_x^{-\sigma}$.

The rest of this section is devoted to proving Claim (ii). Let us fix a test function $\varphi \in \Theta_{\text{lim}}$ and consider the perturbed test function $\varphi^\varepsilon \in \Theta$ constructed in Proposition II.4.1.

Aldous' Criterion [JS03, Theorem 4.5 p356] gives a sufficient condition for Claim (ii): since φ is bounded, one only needs to prove that

$$\forall \eta \in (0, \infty), \lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta} \mathbb{P}(|\varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_2)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_1))| > \eta) = 0, \quad (\text{II.6.4})$$

where $\sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta}$ denotes the supremum with respect to all $(\mathcal{F}_t^\varepsilon)_{t \in \mathbb{R}^+}$ -stopping times τ_1 and τ_2 satisfying a.s. $\tau_1, \tau_2 \in [0, T]$ and $\tau_1 \leq \tau_2 \leq \tau_1 + \delta$.

To prove (II.6.4), introduce the auxiliary process

$$\begin{aligned}\theta^\varepsilon(t) &= \varphi(\rho^\varepsilon(0)) + \varphi^\varepsilon(f^\varepsilon(t), m^\varepsilon(t)) - \varphi^\varepsilon(f^\varepsilon(0), m^\varepsilon(0)) \\ &= \varphi(\rho^\varepsilon(0)) + \int_0^t \mathcal{L}^\varepsilon \varphi^\varepsilon(f^\varepsilon(s), m^\varepsilon(s)) ds + M_{\varphi^\varepsilon}^\varepsilon(t),\end{aligned}\quad (\text{II.6.5})$$

where $M_{\varphi^\varepsilon}^\varepsilon$ is the martingale defined in Proposition II.5.1. One has

$$\begin{aligned}\varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_2)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_1)) &= (\theta^{\varepsilon, \tau^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau^\varepsilon}(\tau_1)) - (\varphi^\varepsilon(f^{\varepsilon, \tau^\varepsilon}(\tau_2), m^{\varepsilon, \tau^\varepsilon}(\tau_2)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_2))) \\ &\quad + (\varphi^\varepsilon(f^{\varepsilon, \tau^\varepsilon}(\tau_1), m^{\varepsilon, \tau^\varepsilon}(\tau_1)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_1))).\end{aligned}$$

Owing to estimates (II.4.3) and (II.4.4) of Proposition II.4.1, and using (II.2.8) and Proposition II.3.3 to estimate $\|m^{\varepsilon, \tau^\varepsilon}(t)\|_E$ and $\|f^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}$, we have

$$|\varphi^\varepsilon(f^{\varepsilon, \tau^\varepsilon}, m^{\varepsilon, \tau^\varepsilon}(t)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(t))| \lesssim_\varphi (1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2)(\varepsilon + \varepsilon^2(1 + \varepsilon^{-\alpha})).$$

Since $\alpha < 1$ in Definition II.2.3, we get

$$\mathbb{E} \left[\sup_{t \in [0, T]} |\varphi^\varepsilon(f^{\varepsilon, \tau^\varepsilon}, m^{\varepsilon, \tau^\varepsilon}(t)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(t))| \right] \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Therefore, one has

$$\sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta} \mathbb{E} [|\varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_2)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_1))|] \leq \sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta} \mathbb{E} [|\theta^{\varepsilon, \tau^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau^\varepsilon}(\tau_1)|] + o(1),$$

where $o(1)$ denotes a quantity converging to 0 when $\varepsilon \rightarrow 0$. The Markov inequality then yields

$$\sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta} \mathbb{P} (|\varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_2)) - \varphi(\rho^{\varepsilon, \tau^\varepsilon}(\tau_1))| > \eta) \leq \sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta} \frac{\mathbb{E} [|\theta^{\varepsilon, \tau^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau^\varepsilon}(\tau_1)|]}{\eta} + o(1).$$

Therefore, it only remains to prove

$$\sup_{\tau_1, \tau_2, \varepsilon} \mathbb{E} [|\theta^{\varepsilon, \tau^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau^\varepsilon}(\tau_1)|] \xrightarrow{\delta \rightarrow 0} 0, \quad (\text{II.6.6})$$

to deduce (II.6.4) and to use Aldous' Criterion.

By (II.6.5), we have

$$|\theta^{\varepsilon, \tau^\varepsilon}(\tau_2) - \theta^{\varepsilon, \tau^\varepsilon}(\tau_1)| \leq \int_{\tau_1 \wedge \tau^\varepsilon}^{\tau_2 \wedge \tau^\varepsilon} |\mathcal{L}^\varepsilon \varphi^\varepsilon(f^\varepsilon(s), m^\varepsilon(s))| ds + |M_{\varphi^\varepsilon}^\varepsilon(\tau_2) - M_{\varphi^\varepsilon}^\varepsilon(\tau_1)|. \quad (\text{II.6.7})$$

Using estimate (II.2.8), Propositions II.3.3 and II.4.1 and the expression (II.4.14) of the infinitesimal generator \mathcal{L} , we get

$$\begin{aligned}|\mathcal{L}^\varepsilon \varphi^\varepsilon(f^{\varepsilon, \tau^\varepsilon}, m^{\varepsilon, \tau^\varepsilon}(t))| &\lesssim_\varphi |\mathcal{L} \varphi(\rho^{\varepsilon, \tau^\varepsilon}(s))| + (1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^3) (\varepsilon(1 + \varepsilon^{-\alpha}) + \varepsilon^2(1 + \varepsilon^{-2\alpha})) \\ &\lesssim_\varphi (1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^3) (1 + \varepsilon(1 + \varepsilon^{-\alpha}) + \varepsilon^2(1 + \varepsilon^{-2\alpha})).\end{aligned}$$

Since $\alpha < 1$ in Definition II.2.3, it gives

$$\mathbb{E} \left[\sup_{t \in [0, T]} |\mathcal{L}^\varepsilon \varphi^\varepsilon(f^{\varepsilon, \tau^\varepsilon}(t), m^{\varepsilon, \tau^\varepsilon}(t))| \right] \lesssim_\varphi 1.$$

Thus, we get

$$\begin{aligned} \sup_{\varepsilon} \sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta} \mathbb{E} \left[\int_{\tau_1 \wedge \tau^\varepsilon}^{\tau_2 \wedge \tau^\varepsilon} |\mathcal{L}^\varepsilon \varphi^\varepsilon(f^\varepsilon(s), m^\varepsilon(s))| ds \right] \\ \leq \sup_{\varepsilon} \sup_{\tau_1 \leq \tau_2 \leq \tau_1 + \delta} \delta \mathbb{E} \left[\sup_{t \in [0, T]} |\mathcal{L}^\varepsilon \varphi^\varepsilon(f^{\varepsilon, \tau^\varepsilon}, m^{\varepsilon, \tau^\varepsilon}(t))| \right] \xrightarrow{\delta \rightarrow 0} 0. \end{aligned}$$

The last term of (II.6.7) is estimated using martingale arguments. Recall that, owing to Proposition II.5.1, $M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}$ is a square-integrable martingale, hence

$$\begin{aligned} \mathbb{E} \left[\left| M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_2) - M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_1) \right|^2 \right] &= \mathbb{E} \left[\left| M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_2) \right|^2 - \left| M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_1) \right|^2 \right] \\ &= \frac{1}{\varepsilon^2} \mathbb{E} \left[\int_{\tau_1 \wedge \tau^\varepsilon}^{\tau_2 \wedge \tau^\varepsilon} (\mathcal{L}_m((\varphi^\varepsilon)^2) - 2\varphi^\varepsilon \mathcal{L}_m \varphi^\varepsilon)(f^\varepsilon(s), m^\varepsilon(s)) ds \right]. \end{aligned} \quad (\text{II.6.8})$$

Using the definitions of the correctors φ_1 and φ_2 (see (II.4.13) and (II.4.15)), one can write

$$\varphi^\varepsilon(f, \ell) = \mathfrak{A}^\varepsilon(f) + \mathfrak{B}^\varepsilon(f, \ell),$$

where \mathfrak{A}^ε does not depend on ℓ and $\mathfrak{B}^\varepsilon(f, \ell) = -\varepsilon^2 \chi'_\rho(\rho, R_0(\ell)\xi)_{L_x^2}$. Then, we have $\mathcal{L}_m \mathfrak{A}^\varepsilon = \mathcal{L}_m((\mathfrak{A}^\varepsilon)^2) = 0$ and $\mathcal{L}_m(\mathfrak{A}^\varepsilon \mathfrak{B}^\varepsilon) = \mathfrak{A}^\varepsilon \mathcal{L}_m \mathfrak{B}^\varepsilon$. As a consequence, one obtains $\mathcal{L}_m((\varphi^\varepsilon)^2) - 2\varphi^\varepsilon \mathcal{L}_m \varphi^\varepsilon = \mathcal{L}_m((\mathfrak{B}^\varepsilon)^2) - 2\mathfrak{B}^\varepsilon \mathcal{L}_m \mathfrak{B}^\varepsilon$ and (II.6.8) becomes

$$\begin{aligned} \mathbb{E} \left[\left| M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_2) - M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_1) \right|^2 \right] &= \varepsilon^2 (\chi'_\rho)^2 \mathbb{E} \left[\int_{\tau_1 \wedge \tau^\varepsilon}^{\tau_2 \wedge \tau^\varepsilon} \mathcal{L}_m((\rho^\varepsilon(s), R_0(\cdot)\xi)_{L_x^2}^2)(m^\varepsilon(s)) ds \right] \\ &\quad + 2\varepsilon^2 (\chi'_\rho)^2 \mathbb{E} \left[\int_{\tau_1 \wedge \tau^\varepsilon}^{\tau_2 \wedge \tau^\varepsilon} (\rho^\varepsilon(s), R_0(m^\varepsilon(s))\xi)_{L_x^2} (\rho^\varepsilon(s), m^\varepsilon(s)\xi)_{L_x^2} ds \right]. \end{aligned}$$

Using Assumption II.6, estimate (II.2.8), Proposition II.3.3 and the condition $|\tau_2 - \tau_1| \leq \delta$, we get

$$\mathbb{E} \left[\left| M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_2) - M_{\varphi^\varepsilon}^{\varepsilon, \tau^\varepsilon}(\tau_1) \right|^2 \right] \lesssim_\varphi \delta \varepsilon^2 \mathbb{E} \left[(1 + \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2)(1 + \varepsilon^{-2\alpha}) \right].$$

Since $\alpha < 1$ in Definition II.2.3, the convergence (II.6.6) follows.

Therefore, we can apply Aldous' Criterion, which concludes the proof of Claim (ii), and of Proposition II.6.2. \square

II.6.3 Identification of the limit

Owing to Proposition II.6.2, there exists a $\mathcal{C}_T^0 H_x^{-\sigma}$ -valued random variable ρ and a sequence $\varepsilon_i \xrightarrow{i \rightarrow \infty} 0, \varepsilon_i \in (0, \varepsilon_0]$ such that $\rho^{\varepsilon_i} \xrightarrow{i \rightarrow \infty} \rho$.

Recall that, owing to Slutsky's Lemma, we also have $\rho^{\varepsilon_i, \tau^{\varepsilon_i}} \xrightarrow{i \rightarrow \infty} \rho$.

To identify the distribution of ρ , we take the limit $\varepsilon_i \rightarrow 0$ in the martingale problem.

Proposition II.6.3. For $\varphi \in \Theta_{\text{lim}}$,

$$t \mapsto \varphi(\rho(t)) - \varphi(\rho(0)) - \int_0^t \mathcal{L}\varphi(\rho(u))du$$

is a martingale (adapted to the filtration $(\mathcal{F}_t^\rho)_{t \in \mathbb{R}^+}$ generated by ρ).

Proof. Fix $\varphi \in \Theta_{\text{lim}}$. Let $0 \leq s \leq s_1 \leq \dots \leq s_j \leq t$, let $g \in \mathcal{C}_b^0((H_x^{-\sigma})^j)$ be a continuous bounded function and define

$$\Phi : \rho \mapsto \left(\varphi(\rho(t)) - \varphi(\rho(s)) - \int_s^t \mathcal{L}\varphi(\rho(u))du \right) g(\rho(s_1), \dots, \rho(s_j)).$$

it suffices to prove that $\mathbb{E}[\Phi(\rho)] = 0$ to get the martingale property and to prove Proposition II.6.3. To reach this goal, we prove that $\mathbb{E}[\Phi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}})]$ converges to both $\mathbb{E}[\Phi(\rho)]$ and 0.

On the one hand, Φ is continuous and $\rho^{\varepsilon_i, \tau^{\varepsilon_i}} \xrightarrow[i \rightarrow \infty]{d} \rho$, hence $\Phi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}}) \xrightarrow[i \rightarrow \infty]{d} \Phi(\rho)$ [Bou04, Proposition IX.5.7]. Moreover, $(\Phi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}}))_{\varepsilon_i}$ is uniformly integrable. Indeed, owing to Definition (II.4.1) of \mathcal{L} and Proposition II.3.3, we have

$$\sup_{\varepsilon} \mathbb{E} \left[|\Phi(\rho^{\varepsilon, \tau^{\varepsilon}})|^2 \right] \lesssim_{T, \varphi, g} \sup_{\varepsilon} \mathbb{E} \left[1 + \|f_0^{\varepsilon}\|_{L^2(\mathcal{M}^{-1})}^4 \right] < \infty.$$

Since $(\Phi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}}))_{\varepsilon_i}$ is uniformly integrable and converges in distribution to $\Phi(\rho)$, we have (see [Bil99, Theorem 5.4])

$$\mathbb{E}[\Phi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}})] \xrightarrow[i \rightarrow \infty]{} \mathbb{E}[\Phi(\rho)].$$

On the other hand, consider the perturbed test function $\varphi^{\varepsilon_i} \in \Theta$ defined in Proposition II.4.1. Owing to Proposition II.5.1, we have

$$\mathbb{E} \left[\left(\varphi^{\varepsilon_i}(f^{\varepsilon_i, \tau^{\varepsilon_i}}(t), m^{\varepsilon_i, \tau^{\varepsilon_i}}(t)) - \varphi^{\varepsilon_i}(f^{\varepsilon_i, \tau^{\varepsilon_i}}(s), m^{\varepsilon_i, \tau^{\varepsilon_i}}(s)) - \int_{s \wedge \tau^{\varepsilon_i}}^{t \wedge \tau^{\varepsilon_i}} \mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(f^{\varepsilon_i}(u), m^{\varepsilon_i}(u)) du \right) g(\rho^{\varepsilon_i, \tau^{\varepsilon_i}}(s_1), \dots, \rho^{\varepsilon_i, \tau^{\varepsilon_i}}(s_j)) \right] = 0.$$

Using the expansion $\varphi^{\varepsilon} = \varphi + \varepsilon \varphi_1 + \varepsilon^2 \varphi_2$ and the boundedness of g , we write

$$|\mathbb{E}[\Phi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}})]| \lesssim_g \sum_{j=1}^4 \mathbb{E}[|r_j|],$$

with

$$\begin{aligned} r_1 &= \varepsilon_i (\varphi_1(f^{\varepsilon_i, \tau^{\varepsilon_i}}(t), m^{\varepsilon_i, \tau^{\varepsilon_i}}(t)) - \varphi_1(f^{\varepsilon_i, \tau^{\varepsilon_i}}(s), m^{\varepsilon_i, \tau^{\varepsilon_i}}(s))) \\ r_2 &= \varepsilon_i^2 (\varphi_2(f^{\varepsilon_i, \tau^{\varepsilon_i}}(t), m^{\varepsilon_i, \tau^{\varepsilon_i}}(t)) - \varphi_2(f^{\varepsilon_i, \tau^{\varepsilon_i}}(s), m^{\varepsilon_i, \tau^{\varepsilon_i}}(s))) \\ r_3 &= \int_{s \wedge \tau^{\varepsilon_i}}^{t \wedge \tau^{\varepsilon_i}} (\mathcal{L}^{\varepsilon_i} \varphi^{\varepsilon_i}(f^{\varepsilon_i, \tau^{\varepsilon_i}}(u), m^{\varepsilon_i, \tau^{\varepsilon_i}}(u)) - \mathcal{L}\varphi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}}(u))) du \\ r_4 &= \int_{t \wedge \tau^{\varepsilon_i}}^t \mathcal{L}\varphi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}}(u)) du - \int_{s \wedge \tau^{\varepsilon_i}}^s \mathcal{L}\varphi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}}(u)) du. \end{aligned}$$

Since $\alpha < 1$ in Definition II.2.3, using estimates (II.4.3), (II.4.4) and (II.4.5) of Proposition II.4.1 and estimates (II.2.8) and (II.3.4) on $m^{\varepsilon_i, \tau^{\varepsilon_i}}$ and $f^{\varepsilon_i, \tau^{\varepsilon_i}}$, we get

$$\begin{aligned} |r_1| &\lesssim \varepsilon_i \|f_0^{\varepsilon_i}\|_{L^2(\mathcal{M}^{-1})}, \\ |r_2| &\lesssim \varepsilon_i^2 (1 + \|f_0^{\varepsilon_i}\|_{L^2(\mathcal{M}^{-1})}^2) (1 + \varepsilon_i^{-\alpha}), \\ |r_3| &\lesssim T(1 + \|f_0^{\varepsilon_i}\|_{L^2(\mathcal{M}^{-1})}^3) (\varepsilon_i(1 + \varepsilon_i^{-\alpha}) + \varepsilon_i^2(1 + \varepsilon_i^{-2\alpha})). \end{aligned}$$

Since $\alpha < 1$ in Definition II.2.3, and owing to Assumption II.7, we get for $k \in \{1, 2, 3\}$

$$\mathbb{E}[|r_k|] \xrightarrow{i \rightarrow \infty} 0.$$

For the last term r_4 , the Cauchy-Schwarz inequality, Propositions II.6.1 and II.3.3 and Assumption II.7 yield

$$\begin{aligned} \mathbb{E}[|r_4|]^2 &\lesssim_{\varphi} \mathbb{E}\left[\|f_0^{\varepsilon_i}\|_{L^2(\mathcal{M}^{-1})}^2\right] \mathbb{E}\left[(|t - t \wedge \tau^{\varepsilon_i}| + |s - s \wedge \tau^{\varepsilon_i}|)^2\right] \\ &\lesssim_{\varphi} T^2 \mathbb{P}(\tau^{\varepsilon_i} < T) \xrightarrow{i \rightarrow \infty} 0. \end{aligned}$$

Therefore, $\mathbb{E}[\Phi(\rho)] = \lim_{i \rightarrow \infty} \mathbb{E}[\Phi(\rho^{\varepsilon_i, \tau^{\varepsilon_i}})] = 0$, which concludes the proof of Proposition II.6.3. \square

Proof of Theorem II.2.1. Recall that uniqueness of the solution of (II.1.6) is stated in Proposition II.3.4. Thus, owing to Proposition II.6.3 and the uniqueness of the limit point, ρ^{ε} converges in distribution $\mathcal{C}_T^0 H_x^{-\sigma}$ to ρ solution of (II.1.6).

If ρ_0 is deterministic, the solution ρ of the limit equation (II.1.6) is also deterministic. A classical argument gives a convergence in probability: owing to Portmanteau's Theorem [Bil99, Theorem 2.1], since ρ^{ε} converges in distribution in $\mathcal{C}_T H_x^{-\sigma}$ to ρ , for any closed set $C \subset \mathcal{C}_T H_x^{-\sigma}$ we have

$$\limsup_{\varepsilon \rightarrow 0} \mathbb{P}(\rho^{\varepsilon} \in C) \leq \mathbb{P}(\rho \in C).$$

Take $C_{\eta} = \{h \in \mathcal{C}_T H_x^{-\sigma} \mid \|h - \rho\|_{\mathcal{C}_T H_x^{-\sigma}} > \eta\}$ for $\eta \in (0, \infty)$. Then $\mathbb{P}(\rho \in C_{\eta}) = 0$, thus ρ^{ε} converges in probability to ρ in $\mathcal{C}_T H_x^{-\sigma}$. \square

II.7 Convergence in $L^2([0, T], L^2(\mathbb{T}^d))$

This section is devoted to proving Theorem II.2.2 if the additional Assumption II.8 is satisfied.

II.7.1 Convergence of ρ^{ε}

Assume that $(\rho^{\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ is tight in $L_T^2 L_x^2$. There exists a limit point ρ' and a subsequence converging to ρ' in distribution in $L_T^2 L_x^2$, hence in $L_T^2 H_x^{-\sigma}$. Owing to Theorem II.2.1, this subsequence converges to ρ in distribution in $\mathcal{C}_T^0 H_x^{-\sigma}$, hence in $L_T^2 H_x^{-\sigma}$. As a consequence, ρ and ρ' are equal in distribution. Therefore, $(\rho^{\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ is tight in $L_T^2 L_x^2$ and ρ is its unique limit point, thus $\rho^{\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} \rho$ in distribution in $L_T^2 L_x^2$.

It only remains to prove that $(\rho^{\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$, or equivalently $(\rho^{\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$, is tight in $L_T^2 L_x^2$ to get this convergence.

We still note w_X the modulus of continuity of a $H_x^{-\sigma}$ -valued continuous process X . Then, [Sim87, Theorem 5] yields the compactness of the set

$$K_{R,\eta} \doteq \left\{ \rho \in L_T^2 L_x^2 \mid \|\rho\|_{L_T^2 H_x^{\sigma''}} \leq R \text{ and } \forall \delta \in (0, \infty), w_\rho(\delta) < \eta(\delta) \right\}$$

in $L_T^2 L_x^2$, for $R > 0$, $\sigma'' > 0$ and $\eta : (0, \infty) \rightarrow [0, \infty)$ such that $\eta(\delta) \xrightarrow{\delta \rightarrow 0} 0$. By Prokhorov's Theorem, to prove the tightness of $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ in $L_T^2 L_x^2$, it is sufficient to prove that, for all $\eta \in (0, \infty)$, there exists $R \in (0, \infty)$ and $\sigma'' \in (0, \infty)$ such that

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \mathbb{P}(w_{\rho^{\varepsilon, \tau^\varepsilon}}(\delta) > \eta) = 0, \quad (\text{II.7.1})$$

and

$$\sup_{\varepsilon} \mathbb{P}\left(\|\rho^{\varepsilon, \tau^\varepsilon}\|_{L_T^2 H_x^{\sigma''}} > R\right) < \eta. \quad (\text{II.7.2})$$

Equation (II.7.1) is a consequence of the tightness in $\mathcal{C}_T^0 H_x^{-\sigma}$ [Bil99, Theorem 8.2]. To prove (II.7.2), it suffices to prove that for some $\sigma'' \in (0, \infty)$, we have

$$\sup_{\varepsilon} \mathbb{E} \left[\|\rho^{\varepsilon, \tau^\varepsilon}\|_{L_T^2 H_x^{\sigma''}}^2 \right] \lesssim 1, \quad (\text{II.7.3})$$

and (II.7.2) follows from the Markov inequality.

Let $g^\varepsilon = \varepsilon \partial_t f^\varepsilon + a(v) \cdot \nabla_x f^\varepsilon$. By Assumption II.8, we can apply an averaging lemma ([BD99, Theorem 2.3] with $f(t) = f^\varepsilon(\varepsilon t)$, $g(t) = g^\varepsilon(\varepsilon t)$ and $h = 0$ until time $T \wedge \tau^\varepsilon$). After rescaling the time $t \mapsto t/\varepsilon$, we get

$$\begin{aligned} \|\rho^{\varepsilon, \tau^\varepsilon}\|_{L_T^2 H_x^{\sigma'/4}}^2 &= \int_0^{T \wedge \tau^\varepsilon} \|\rho^{\varepsilon, \tau^\varepsilon}(t)\|_{H_x^{\sigma'/4}}^2 dt \\ &\leq \varepsilon \|f_0^\varepsilon\|_{L_x^2}^2 + \int_0^{T \wedge \tau^\varepsilon} \|f^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}^2 dt + \int_0^{T \wedge \tau^\varepsilon} \|g^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}^2 dt, \end{aligned}$$

and owing to the Cauchy-Schwarz inequality,

$$\begin{aligned} \|g^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})} &= \left\| \varepsilon f^{\varepsilon, \tau^\varepsilon}(t) m^{\varepsilon, \tau^\varepsilon}(t) + \frac{1}{\varepsilon} L f^{\varepsilon, \tau^\varepsilon}(t) \right\|_{L^2(\mathcal{M}^{-1})} \\ &\leq \varepsilon \|f^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})} \|m^{\varepsilon, \tau^\varepsilon}(t)\|_{\mathcal{C}_x^0} + \frac{1}{\varepsilon} \|L f^{\varepsilon, \tau^\varepsilon}(t)\|_{L^2(\mathcal{M}^{-1})}. \end{aligned}$$

Then, Assumption II.7, estimate (II.2.8) and Proposition II.3.3 yields (II.7.3) with $\sigma'' = \frac{\sigma'}{4}$. Note that here, we used the estimate on the second term of the left-hand side of (II.3.4). Since we proved (II.7.1) and (II.7.2), and since K_R is compact in $L_T^2 L_x^2$, Prokhorov's Theorem gives that the family $(\rho^{\varepsilon, \tau^\varepsilon})_{\varepsilon \in (0, \varepsilon_0]}$ is tight in $L_T^2 L_x^2$.

This concludes the proof of the convergence in distribution of ρ^ε to ρ in $L_T^2 L_x^2$.

If ρ_0 is deterministic, the convergence in probability follows from the same argument as in Section II.6.3.

II.7.2 Convergence of f^ε

The mapping $h \mapsto h\mathcal{M}$ is a continuous mapping from $L_T^2 L_x^2$ to $L_T^2 L^2(\mathcal{M}^{-1})$ (it is a linear mapping of operator norm 1 owing to Assumption II.1). Thus, the convergence of ρ^ε in

distribution or in probability in $L_T^2 L_x^2$ to ρ implies the same convergence for $\rho^\varepsilon \mathcal{M}$ to $\rho \mathcal{M}$ in $L_T^2 L^2(\mathcal{M}^{-1})$.

Since

$$f^\varepsilon = -L f^\varepsilon + \rho^\varepsilon \mathcal{M},$$

it is thus sufficient to prove that $L f^\varepsilon \xrightarrow[\varepsilon \rightarrow 0]{} 0$ in probability in $L_T^2 L^2(\mathcal{M}^{-1})$ to conclude the proof of Theorem II.2.2. Indeed, assume we proved this convergence. Then, if ρ^ε converges in distribution, Slutsky's Lemma yields the convergence in distribution of f^ε [Bil99, Theorem 4.1]. If ρ^ε also converges in probability, then f^ε converges in probability as a sum of two terms converging in probability).

Let us now prove that $L f^\varepsilon \xrightarrow[\varepsilon \rightarrow 0]{} 0$ in probability in $L_T^2 L^2(\mathcal{M}^{-1})$. Owing to Proposition II.3.3, we have almost surely

$$\int_0^{T \wedge \tau^\varepsilon} \|L f^\varepsilon(s)\|_{L^2(\mathcal{M}^{-1})}^2 ds \leq \varepsilon^2 C(T) \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}^2.$$

Therefore, on the event $\tau^\varepsilon \geq T$, we get $\|L f^\varepsilon\|_{L_T^2 L^2(\mathcal{M}^{-1})} \leq \varepsilon \sqrt{C(T)} \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})}$. Thus, we have for $\eta \in (0, \infty)$

$$\begin{aligned} \mathbb{P} \left(\|L f^\varepsilon\|_{L_T^2 L^2(\mathcal{M}^{-1})} > \eta \right) &= \mathbb{P} \left(\tau^\varepsilon < T, \|L f^\varepsilon\|_{L_T^2 L^2(\mathcal{M}^{-1})} > \eta \right) \\ &\quad + \mathbb{P} \left(\tau^\varepsilon \geq T, \|L f^\varepsilon\|_{L_T^2 L^2(\mathcal{M}^{-1})} > \eta \right) \\ &\leq \mathbb{P}(\tau^\varepsilon < T) + \mathbb{P} \left(\varepsilon \sqrt{C(T)} \|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})} > \eta \right) \\ &\leq \mathbb{P}(\tau^\varepsilon < T) + \eta^{-1} \varepsilon \sqrt{C(T)} \mathbb{E} \left[\|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})} \right], \end{aligned}$$

owing to the Markov inequality. Since $\sup_\varepsilon \mathbb{E} \left[\|f_0^\varepsilon\|_{L^2(\mathcal{M}^{-1})} \right] < \infty$ by Assumption II.7 and using Proposition II.6.1, we get that $L f^\varepsilon$ converges to 0 in probability in $L_T^2 L^2(\mathcal{M}^{-1})$, which concludes the proof of Theorem II.2.2.

Chapter III

Asymptotic preserving schemes for Stochastic Differential Equations

This chapter is a detailed version of the paper *On Asymptotic Preserving schemes for a class of Stochastic Differential Equations in averaging and diffusion approximation regimes*, written with Charles-Édouard Bréhier [BRR20]. The preprint can be found at <https://arxiv.org/abs/2011.02341>.

A notable difference between this chapter and [BRR20] is addition of quantitative experiments (Section III.6 and III.7). The code for these additional simulations was written by Benoit Fabrèges.

III.1 Introduction

Deterministic and stochastic systems are ubiquitous in science and engineering. Traditional modeling and numerical methods become ineffective when systems evolve at different time scales: see for instance the monographs [E11, Kue15] for comprehensive treatment of multiscale dynamics. Averaging and homogenization [PS08] are two popular techniques which are employed to rigorously derive macroscopic limiting equations, starting from (stochastic) slow-fast systems with separated time-scales.

In the last two decades, constructing efficient numerical methods for multiscale stochastic systems has been a very active research area: let us mention the Heterogeneous Multiscale Method (see [AEEVE12, Bré13, ELVE05]), projective integration (see [GKK06]), equation-free coarse-graining (see [KGH⁺03]), spectral methods (see [APV17]), micro-macro acceleration methods (see [VZS20]), parareal algorithms (see [LLMS20]). In the methods mentioned above, the objective is to approximate the limiting model for the slow variables of interest, and only partial but relevant information coming from the fast dynamics is taken into account. As a consequence, these methods may not be appropriate if one wants to approximate simultaneously the original multiscale model and its limit. In this chapter, we focus on the notion of asymptotic preserving schemes, in order to overcome this issue.

To motivate and illustrate our work, let us introduce simplified versions of the systems of Stochastic Differential Equations (SDE) considered in this chapter. The time-scale separation parameter is denoted by $\varepsilon \in (0, 1]$. On the one hand, in the averaging regime (see Equation (III.2.1) in Section III.2.1 for the more general version), we consider systems

of the type

$$\begin{cases} dX_t^\varepsilon = b(X_t^\varepsilon, m_t^\varepsilon)dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon}dt + \frac{\sqrt{2}}{\sqrt{\varepsilon}}d\beta_t. \end{cases} \quad (\text{III.1.1})$$

When $\varepsilon \rightarrow 0$, the averaging principle (see [PS08, Chapter 10]) states that X^ε converges (at least in distribution) to the solution X of the Ordinary Differential Equation $\dot{X} = \bar{b}(X)$ where $\bar{b}(x) = \int b(x, m)d\nu(m)$ and $\nu = \mathcal{N}(0, 1)$ is the standard Gaussian random variable. On the other hand, in the diffusion approximation regime (see Equation (III.2.12) in Section III.2.2 for the more general version), we consider systems of the type

$$\begin{cases} dX_t^\varepsilon = \frac{\sigma(X_t^\varepsilon)m_t^\varepsilon}{\varepsilon}dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon^2}dt + \frac{1}{\varepsilon}d\beta_t. \end{cases} \quad (\text{III.1.2})$$

When $\varepsilon \rightarrow 0$, the diffusion approximation result (see [PS08, Chapter 11]) states that X^ε converges (in distribution) to the solution X of the SDE

$$dX_t = \sigma(X_t) \circ d\beta_t,$$

where the noise is interpreted in the Stratonovich sense. This type of results is related to results known as Wong-Zakai approximation and Smoluchowski-Kramers limits in the literature. In the two SDE systems (III.1.1) and (III.1.2), the fast component is an Ornstein-Uhlenbeck process.

In this chapter, we are interested in the behavior when $\varepsilon \rightarrow 0$ of numerical schemes for the SDEs (III.1.1) and (III.1.2). To explain the challenge faced and the solutions proposed in this chapter, we consider the following schemes, which are both consistent for any fixed value of $\varepsilon > 0$. On the one hand, in the averaging regime one defines

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon), \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon + \sqrt{\frac{2\Delta t}{\varepsilon}} \gamma_n. \end{cases} \quad (\text{III.1.3})$$

On the other hand, in the diffusion approximation regime one defines

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \sigma(X_n^\varepsilon) \frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2} m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon} \gamma_n, \end{cases} \quad (\text{III.1.4})$$

In the schemes (III.1.3) and (III.1.4), $(\gamma_n)_{n \in \mathbb{N}_0}$ is a sequence of independent standard Gaussian random variables. One may check that $X_n^\varepsilon \rightarrow X_n$ for all $n \in \mathbb{N}_0$, in probability, when $\varepsilon \rightarrow 0$, where the limiting schemes are given by

$$X_{n+1} = X_n + \Delta t b(X_n, 0)$$

in the averaging regime, and

$$X_{n+1} = X_n + \sqrt{\Delta t} \sigma(X_n) \gamma_n,$$

in the diffusion approximation regime. Note that, in the second case, the limiting scheme is consistent with the Itô interpretation of the noise, instead of the correct Stratonovich one. In the two cases, the limiting scheme is in general not consistent with the limiting equation, and using such a scheme in practice may lead to drawing false conclusions about the limiting system from numerical experiments. We refer to [FG18, LAE08] for other examples of situations where numerical schemes perform badly when applied to multiscale SDE systems.

The objective of this chapter is to design and study Asymptotic Preserving (AP) schemes, such that the following diagram commutes (where convergence is understood in distribution): if $T = N\Delta t$, one has

$$\begin{array}{ccc} X_N^\varepsilon & \xrightarrow{\Delta t \rightarrow 0} & X^\varepsilon(T) \\ \downarrow \varepsilon \rightarrow 0 & & \downarrow \varepsilon \rightarrow 0 \\ X_N & \xrightarrow{\Delta t \rightarrow 0} & X(T) \end{array}$$

The two schemes (III.1.3) and (III.1.4) described above are not AP. The notion of AP schemes has been introduced in [Jin99], for applications to multiscale kinetic Partial Differential Equations (PDEs), which converge to parabolic diffusion PDEs. We refer to [DP14, Section 7], [HJL17], [Jin12] and [Pup19, Section 4] for recent reviews on AP schemes for this type of models. To the best of our knowledge, the design and analysis of asymptotic preserving schemes for slow-fast SDEs of the type (III.1.1) and (III.1.2) has not been considered so far in the literature. Note that a specific feature (compared with the deterministic case) is the need to consider convergence in distribution. Let us mention related works for Stochastic Partial Differential Equations (SPDEs), in the diffusion approximation regime. First, in [DM16, Mar06], the authors consider Schrödinger equations and study an abstract asymptotic preserving property. However, they do not propose implementable schemes. In [AF19], the authors deal with some multiscale stochastic kinetic PDEs, driven by a Wiener process. However, the structure of the model is different from the one of (III.1.2). In Chapter IV, we apply the findings of this chapter to the SPDE models considered in [AF19]. The works mentioned above concerning SPDE models are limited to diffusion coefficients of the type $\sigma(x) = x$, for which specific arguments may give a straightforward construction of AP schemes, for appropriate discretization of the fast component. An AP scheme in the case $\sigma(x) = 1$ for (III.1.2) is proposed in [PSZ09], however the subtlety of the interpretation of the noise at the limit is not relevant in that case. Finally, let us also mention that AP schemes have also been studied for PDEs with random coefficients, see [HJ17, Jin18, JLP18] or in the context of Monte-Carlo methods for deterministic problems, see [DPS18, RLJ14].

We are now in position to describe the contributions of this chapter. In Section III.3.1, we define the appropriate notion of AP schemes for SDE systems, related to convergence in distribution, and study several general properties.

Our first main result is Theorem III.3.3, which exhibits an example of AP scheme in the averaging regime: for the simplified version (III.1.1), the scheme is given by

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon), \\ m_{n+1}^\varepsilon = e^{-\frac{\Delta t}{\varepsilon}} m_n^\varepsilon + \sqrt{1 - e^{-\frac{2\Delta t}{\varepsilon}}} \gamma_n. \end{cases} \quad (\text{III.1.5})$$

The fast component in the scheme above is discretized using a scheme which is exact in distribution.

Our second main result is Theorem III.3.4, which states error estimates of the type

$$\sup_{\varepsilon \in (0,1]} |\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| = O(\sqrt{\Delta t}),$$

for sufficiently smooth real-valued mappings φ . This error estimate means that the scheme is Uniformly Accurate.

Finally, our third main result is Theorem III.3.5, which exhibits an example of AP scheme in the diffusion approximation regime: for the simplified version (III.1.2) (see Corollary III.3.6), the scheme is given by

$$\begin{cases} m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2} m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon} \gamma_n, \\ Y_{n+1}^\varepsilon = X_n^\varepsilon + \sigma(X_n^\varepsilon) \frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}, \\ X_{n+1}^\varepsilon = X_n^\varepsilon + \frac{\sigma(X_n^\varepsilon) + \sigma(Y_{n+1}^\varepsilon)}{2} \frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}. \end{cases} \quad (\text{III.1.6})$$

A prediction-correction method is employed to retrieve the correct interpretation of the noise for the limiting equation: the scheme (III.1.2) is indeed consistent with the Stratonovich interpretation of the noise.

Let us also mention that another situation is considered in Corollary III.3.7: for the model (III.2.22) taken from [LBL20] (with an application in astrophysics), the limiting equation (III.2.23) contains a so-called noise-induced drift-term, which is captured only for well-designed AP schemes.

Some numerical experiments (see Section III.5) show that the AP schemes (III.1.5) and (III.1.6) are effective in all regimes $\varepsilon > 0$ and $\varepsilon \rightarrow 0$, contrary to the schemes (III.1.3) and (III.1.4) which fail to capture the correct limiting behavior when $\varepsilon \rightarrow 0$.

The chapter is organized as follows. The general SDE models in the averaging and diffusion approximation regimes are presented in Sections III.2.1 and III.2.2. The main results of this chapter are stated in Section III.3: the general theory of AP schemes is presented in Section III.3.1, and it is applied in the averaging and diffusion approximation regimes in Section III.3.2 and III.3.3 respectively. Qualitative numerical experiments are reported in Section III.5, while quantitative numerical experiments are presented in the averaging regime and the diffusion approximation regime, respectively in Sections III.6 and III.7. Section III.4 is devoted to the proof of the error estimates stated in Theorem III.3.4. Finally, Section III.8 gives some conclusions and perspectives.

III.2 Slow-fast SDE models and their limits

Without loss of generality, the time-scale separation parameter ε satisfies $\varepsilon \in (0, 1]$. The time-step size of the integrators studied in this work is denoted by Δt . It is assumed that $\Delta t = \frac{T}{N}$ where $T \in (0, \infty)$ is a fixed time and $N \in \mathbb{N}$. Without loss of generality, it is assumed that $\Delta t \in (0, 1]$.

In the slow-fast systems considered in this work, the slow component X^ε takes values in the d -dimensional flat torus \mathbb{T}^d , where $d \in \mathbb{N}$ is an arbitrary integer, whereas the fast component m^ε takes values in \mathbb{R} . The framework and the models considered in this work may be generalized in many ways to more complex situations, however the arguments and

results below are sufficient to illustrate the difficulties of designing asymptotic preserving schemes for stochastic equations.

Let $(\beta_t)_{t \in \mathbb{R}^+}$ and $(B_t)_{t \in \mathbb{R}^+}$ be two independent standard Wiener processes, with values in \mathbb{R} and \mathbb{R}^D respectively, where $D \in \mathbb{N}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfies the usual conditions.

The following notation for derivatives is used below: $\nabla_x = (\partial_{x_i})_{1 \leq i \leq d} \in \mathbb{R}^d$ and ∂_m are the partial gradient and derivative operators with respect to x and m respectively. If σ is a mapping with values in $\mathcal{M}_{d,D}(\mathbb{R})$ (the space of $d \times D$ matrices with real entries), let σ^\star denote the transpose of σ , and set $\sigma\sigma^\star : \nabla_x^2 = \sum_{i,j=1}^d (\sigma\sigma^\star)_{i,j} \partial_{x_i} \partial_{x_j}$. If b is a \mathbb{R}^d -valued mapping, let $b \cdot \nabla_x = \sum_{i=1}^d b_i \partial_{x_i}$.

Assumption III.1. The initial conditions $X_0^\varepsilon \in \mathbb{T}^d$ and $m_0^\varepsilon \in \mathbb{R}$ of the processes are deterministic quantities and they satisfy

$$X_0^\varepsilon = x_0^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} x_0, \quad \sup_{\varepsilon \in (0,1]} |m_0^\varepsilon| < \infty.$$

III.2.1 The averaging regime

In the so-called averaging regime, we consider slow-fast SDE systems of the type

$$\begin{cases} dX_t^\varepsilon = b(X_t^\varepsilon, m_t^\varepsilon)dt + \sigma(X_t^\varepsilon, m_t^\varepsilon)dB_t, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon}dt + \frac{\sqrt{2}h(X_t^\varepsilon)}{\sqrt{\varepsilon}}d\beta_t. \end{cases} \quad (\text{III.2.1})$$

The coefficients appearing in (III.2.1) are assumed to satisfy the following conditions.

Assumption III.2. The functions $b : \mathbb{T}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{T}^d \times \mathbb{R} \rightarrow \mathcal{M}_{d,D}(\mathbb{R})$ are assumed to be of class \mathcal{C}^4 , and $h : \mathbb{T}^d \rightarrow \mathbb{R}$ is assumed to be of class \mathcal{C}^1 . Moreover, they are all assumed to be bounded and to have bounded derivatives.

Owing to Assumption III.2, for all initial conditions $X_0^\varepsilon \in \mathbb{T}^d$ and $m_0^\varepsilon \in \mathbb{R}$, and for every $\varepsilon \in (0, 1]$, there exists a unique global solution $(X^\varepsilon, m^\varepsilon)$ of the SDE system (III.2.1). Since h is bounded, it is straightforward to check that

$$\sup_{\varepsilon \in (0,1]} \sup_{t \in \mathbb{R}^+} \frac{\mathbb{E}[|m^\varepsilon(t)|^2]}{1 + |m_0^\varepsilon|^2} < \infty. \quad (\text{III.2.2})$$

This estimate will prove useful to prove Proposition III.2.1.

The infinitesimal generator \mathcal{L}^ε associated with the SDE (III.2.1) has the following expression:

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon} \mathcal{L}_{OU} + \mathcal{L}_0, \quad (\text{III.2.3})$$

where

$$\begin{aligned} \mathcal{L}_0 &= b(x) \cdot \nabla_x + \frac{1}{2} \sigma\sigma^\star : \nabla_x^2, \\ \mathcal{L}_{OU} &= -m\partial_m + h(x)^2 \partial_m^2, \end{aligned} \quad (\text{III.2.4})$$

Observe that for fixed $x \in \mathbb{T}^d$, \mathcal{L}_{OU} is the generator of an ergodic Ornstein-Uhlenbeck process. The associated invariant distribution is $\nu^x = \mathcal{N}(0, h(x)^2)$.

Define averaged coefficients as follows: for all $x \in \mathbb{T}^d$

$$\bar{b}(x) = \int b(x, m) d\nu^x(m), \quad \bar{a}(x) = \int \sigma(x, m) \sigma(x, m)^* d\nu^x(m). \quad (\text{III.2.5})$$

Note that $\bar{b} : \mathbb{T}^d \rightarrow \mathbb{R}^d$ is of class \mathcal{C}^4 . The averaging principle result stated below requires the following condition to be satisfied.

Assumption III.3. There exists an integer $\bar{D} \in \mathbb{N}$ and a function $\bar{\sigma} : \mathbb{T}^d \rightarrow \mathcal{M}_{d, \bar{D}}(\mathbb{R})$ of class \mathcal{C}^4 such that for all $x \in \mathbb{T}^d$

$$\bar{a}(x) = \bar{\sigma}(x) \bar{\sigma}(x)^*. \quad (\text{III.2.6})$$

Assumption III.3 holds if there exists $c \in (0, \infty)$ such that $\bar{a}(x) \geq cI$ for all $x \in \mathbb{T}^d$ (as symmetric matrices). This condition is satisfied when σ only depends on the slow variable x ($\partial_m \sigma(x, m) = 0$ for all $(x, m) \in \mathbb{T}^d \times \mathbb{R}$), or when $\sigma(x, m) \sigma(x, m)^* \geq cI$ for all $(x, m) \in \mathbb{T}^d \times \mathbb{R}$. In that case, one can choose $\bar{D} = d$. If the diffusion coefficient is of the type $\sigma(x, m) = \sigma^\sharp(m) \sigma^\dagger(x)$, with $\sigma^\dagger(x) \in \mathbb{R}^d$ and $\sigma^\sharp(m) \in \mathbb{R}$, then one can choose $\bar{D} = D$ and $\bar{\sigma}(x) = \sigma^\dagger(x) \sqrt{\int \sigma^\sharp(m)^2 d\nu^x(m)}$ for all $x \in \mathbb{T}^d$.

We are now in position to state the averaging principle result and to define the limiting process X obtained when $\varepsilon \rightarrow 0$.

Proposition III.2.1. *Let Assumptions III.1, III.2 and III.3 be satisfied. Let $T \in (0, \infty)$. When $\varepsilon \rightarrow 0$, the $\mathcal{C}([0, T], \mathbb{T}^d)$ -valued process $(X^\varepsilon(t))_{0 \leq t \leq T}$ converges in distribution to the solution $(X(t))_{0 \leq t \leq T}$ of the limiting SDE*

$$dX_t = \bar{b}(X_t) dt + \bar{\sigma}(X_t) d\bar{B}_t, \quad (\text{III.2.7})$$

with initial condition $X(0) = x_0$, where the coefficients \bar{b} and $\bar{\sigma}$ are defined by (III.2.5)–(III.2.6), and where $(\bar{B}_t)_{t \in \mathbb{R}^+}$ is a standard $\mathbb{R}^{\bar{D}}$ -valued Wiener process.

The infinitesimal generator \mathcal{L} associated with the limiting SDE (III.2.7) is given by

$$\mathcal{L} = \bar{b}(x) \cdot \nabla_x + \frac{1}{2} \bar{\sigma} \bar{\sigma}^* : \nabla_x^2, \quad (\text{III.2.8})$$

and is such that the following property holds: let $\varphi \in \mathcal{C}^4(\mathbb{T}^d)$, then there exists a function $\varphi^1 : \mathbb{T}^d \times \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\varphi^\varepsilon = \varphi + \varepsilon \varphi^1, \quad (\text{III.2.9})$$

$$\mathcal{L}^\varepsilon \varphi^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \mathcal{L} \varphi. \quad (\text{III.2.10})$$

Finally, let $\varphi \in \mathcal{C}^4(\mathbb{T}^d)$, then there exists $C(T, \varphi) \in (0, \infty)$ such that

$$|\mathbb{E}[\varphi(X^\varepsilon(T))] - \mathbb{E}[\varphi(X(T))]| \leq C(T, \varphi) \varepsilon. \quad (\text{III.2.11})$$

The averaging principle stated in Proposition III.2.1 is a standard result, see for instance [PS08, Chapter 16]. In general the convergence stated in Proposition III.2.1 only holds in distribution, however it holds in stronger sense (for instance in mean-square sense) if σ only depends on x .

We refer to Section III.A.1 for a sketch of the construction of the perturbed test function φ^ε which satisfies (III.2.9)–(III.2.10) (see [FGPS107, Chapter 6] for a detailed description of the perturbed test function method). Note that the perturbed test function appears in Proposition III.3.1 below. For the error estimate (III.2.11), see Lemma III.4.2 and its proof below.

III.2.2 The diffusion approximation regime

General model

In the so-called diffusion approximation regime, we consider slow-fast SDE systems of the type

$$\begin{cases} dX_t^\varepsilon = b(X_t^\varepsilon)dt + \frac{\sigma(X_t^\varepsilon)m_t^\varepsilon}{\varepsilon}dt, \\ dm_t^\varepsilon = f(X_t^\varepsilon)\left(-\frac{m_t^\varepsilon}{\varepsilon^2}dt + \frac{g(X_t^\varepsilon)}{\varepsilon}dt + \frac{h(X_t^\varepsilon)}{\varepsilon}d\beta_t\right). \end{cases} \quad (\text{III.2.12})$$

The coefficients appearing in (III.2.12) are assumed to satisfy the following conditions.

Assumption III.4. The functions $b : \mathbb{T}^d \rightarrow \mathbb{R}^d$ and $g, h : \mathbb{T}^d \rightarrow \mathbb{R}$ are assumed to be of class \mathcal{C}^1 . The functions $\sigma : \mathbb{T}^d \rightarrow \mathbb{R}^d$ and $f : \mathbb{T}^d \rightarrow \mathbb{R}$ are assumed to be of class \mathcal{C}^2 . Moreover, f takes values in $(0, \infty)$: we assume that $\min_{x \in \mathbb{T}^d} f(x) > 0$.

Owing to Assumption III.4, for all initial conditions $X_0^\varepsilon \in \mathbb{T}^d$ and $m_0^\varepsilon \in \mathbb{R}$, and for every $\varepsilon \in (0, 1]$, there exists a unique global solution $(X^\varepsilon, m^\varepsilon)$ of the SDE system (III.2.12). The infinitesimal generator \mathcal{L}^ε associated with the SDE (III.2.12) has the following expression:

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon^2}\mathcal{L}_{OU} + \frac{1}{\varepsilon}\mathcal{L}_1 + \mathcal{L}_0, \quad (\text{III.2.13})$$

where

$$\begin{aligned} \mathcal{L}_0 &= b(x) \cdot \nabla_x, \\ \mathcal{L}_1 &= m\sigma(x) \cdot \nabla_x + f(x)g(x)\partial_m, \\ \mathcal{L}_{OU} &= -f(x)m\partial_m + \frac{1}{2}f(x)^2h(x)^2\partial_m^2. \end{aligned} \quad (\text{III.2.14})$$

Observe that for fixed $x \in \mathbb{T}^d$, \mathcal{L}_{OU} is the generator of an ergodic Ornstein-Uhlenbeck process. The associated invariant distribution is $\nu^x = \mathcal{N}(0, \frac{f(x)h(x)^2}{2})$.

We are now in position to state the diffusion approximation result and to define the limiting process X obtained when $\varepsilon \rightarrow 0$.

Proposition III.2.2. *Let Assumptions III.1 and III.4 be satisfied. Let $T \in (0, \infty)$. When $\varepsilon \rightarrow 0$, the $\mathcal{C}([0, T], \mathbb{T}^d)$ -valued process $(X^\varepsilon(t))_{0 \leq t \leq T}$ converges in distribution to the solution $(X(t))_{0 \leq t \leq T}$ of the limiting SDE*

$$dX_t = \left(b + g\sigma + \frac{h^2}{2}(\sigma \cdot \nabla_x)\sigma - \frac{h^2}{2f}\sigma \cdot \nabla_x f\sigma \right)(X_t)dt + h(X_t)\sigma(X_t)dW_t, \quad X(0) = x_0, \quad (\text{III.2.15})$$

driven by a standard one-dimensional Wiener process $(W(t))_{t \in \mathbb{R}^+}$.

The infinitesimal generator \mathcal{L} associated with the limiting SDE (III.2.15) is given by

$$\begin{aligned} \mathcal{L}\varphi &= (b + g\sigma) \cdot \nabla_x \varphi + \frac{h^2 f \sigma}{2} \cdot \nabla_x \left(\frac{\sigma}{f} \cdot \nabla_x \varphi \right) \\ &= (b + g\sigma) \cdot \nabla_x \varphi + \frac{h^2}{2} \sigma \sigma^\star : \nabla_x^2 \varphi + \frac{h^2}{2} (\sigma \cdot \nabla_x) \sigma \cdot \nabla_x \varphi - \frac{h^2}{2f} \sigma \cdot \nabla_x f \sigma \cdot \nabla_x \varphi, \end{aligned} \quad (\text{III.2.16})$$

and is such that the following property holds: let $\varphi \in \mathcal{C}^3(\mathbb{T}^d)$, then one constructs two functions $\varphi_1, \varphi_2 : \mathbb{T}^d \times \mathbb{R} \rightarrow \mathbb{R}$, such that

$$\varphi^\varepsilon = \varphi + \varepsilon\varphi_1 + \varepsilon^2\varphi_2, \quad (\text{III.2.17})$$

$$\mathcal{L}^\varepsilon \varphi^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \mathcal{L}\varphi. \quad (\text{III.2.18})$$

Finally, let $\varphi \in \mathcal{C}^3(\mathbb{T}^d)$, then there exists $C(T, \varphi) \in (0, \infty)$ such that

$$|\mathbb{E}[\varphi(X^\varepsilon(T))] - \mathbb{E}[\varphi(X(T))]| \leq C(T, \varphi)\varepsilon. \quad (\text{III.2.19})$$

The diffusion approximation stated in Proposition III.2.2 is a standard result, see for instance [PS08, Chapter 18]. We refer to Section III.A.2 for a sketch of the construction of the perturbed test function φ^ε which satisfies (III.2.17)–(III.2.18) (see [FGPS107, Chapter 6] for a detailed description of the perturbed test function method). Since the error estimate (III.2.19) plays no role in the sequel, the proof is omitted. We refer to [KY05] for arguments using asymptotic expansions of solutions of Kolmogorov equations leading to (III.2.19), (see also [LBL20] for related computations).

Two examples in the approximation-diffusion regime

The setting described above encompasses several interesting examples of SDE systems. In order to focus on the different possible issues which need to be overcome when constructing asymptotic preserving numerical schemes in the regime $\varepsilon \rightarrow 0$, we deal with two examples described below. In addition, the asymptotic preserving numerical schemes will have simpler formulations for these examples than in the general case. In both examples, dimension is set equal to $d = 1$ to simplify the presentation, and $b = 0$.

Let us present the first example: consider the system

$$\begin{cases} dX_t^\varepsilon = \frac{\sigma(X_t^\varepsilon)m_t^\varepsilon}{\varepsilon}dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon^2}dt + \frac{1}{\varepsilon}d\beta_t, \end{cases} \quad (\text{III.2.20})$$

where the coefficients in the fast equation are constant: $f(x) = h(x) = 1$ and $g(x) = 0$ for all $x \in \mathbb{T}$. Applying Proposition III.2.2 in this example yields the following limiting equation

$$dX_t = \sigma(X_t) \circ dW_t, \quad (\text{III.2.21})$$

where the noise is interpreted using the Stratonovich convention. With the Itô convention, the equation is written as

$$dX_t = \frac{1}{2}\sigma(X_t)\sigma'(X_t)dt + \sigma(X_t)dW_t.$$

Note that the diffusion approximation result (Proposition III.2.2) may be obtained by straightforward arguments in two cases, which will be repeated at the discrete-time levels. Let $\zeta^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t m^\varepsilon(s)ds$ for all $t \in \mathbb{R}^+$. First, if $\sigma(x) = 1$ for all $x \in \mathbb{T}$, then one has $dX_t^\varepsilon = d\zeta_t^\varepsilon$. Therefore passing to the limit yields

$$X^\varepsilon(t) = X_0^\varepsilon + \zeta^\varepsilon(t) \xrightarrow{\varepsilon \rightarrow 0} x_0 + W(t),$$

and the limiting equation is $dX_t = dW_t$. Second, assume that x , $X^\varepsilon(t)$ and $X(t)$ take values in the real line \mathbb{R} (instead of the torus \mathbb{T}) and that $\sigma(x) = x$ for all $x \in \mathbb{R}$. Then (III.2.21) is written as $dX_t^\varepsilon = X_t^\varepsilon d\zeta_t^\varepsilon$. Computing the solution and passing to the limit then yields

$$X^\varepsilon(t) = X_0^\varepsilon \exp(\zeta^\varepsilon(t)) \xrightarrow{\varepsilon \rightarrow 0} x_0 \exp(W(t)) = X(t),$$

and the limiting equation is $dX_t = X_t \circ dW_t$.

Note that when the function σ is not constant, the Itô and Stratonovich interpretations differ. Constructing an asymptotic preserving requires to capture the correction term in a limiting scheme (which will naturally be associated with an Itô interpretation of the noise).

Let us now present the second example, taken from [LBL20]. The coefficients f, g, h are allowed to depend on the slow component x , whereas it is assumed that $\sigma(x) = 1$ for all $x \in \mathbb{T}$. Therefore, the system in the second example has the following expression

$$\begin{cases} dX_t^\varepsilon = \frac{m_t^\varepsilon}{\varepsilon} dt, \\ dm_t^\varepsilon = f(X_t^\varepsilon) \left(-\frac{m_t^\varepsilon}{\varepsilon^2} dt + \frac{g(X_t^\varepsilon)}{\varepsilon} dt + \frac{h(X_t^\varepsilon)}{\varepsilon} d\beta_t \right), \end{cases} \quad (\text{III.2.22})$$

Applying Proposition III.2.2 in this example yields the following limiting equation

$$dX_t = g(X_t)dt - \frac{h(X_t)^2 f'(X_t)}{2f(X_t)} dt + h(X_t) dW_t. \quad (\text{III.2.23})$$

The noise is interpreted in the Itô sense. Observe that when f is not constant, the noise-induced drift term $\frac{h^2 f'}{2f}$ appears. The construction of asymptotic preserving schemes for this problem requires to be careful in order to capture this additional drift term in the limiting scheme.

III.3 Numerical discretization and asymptotic preserving schemes

The objective of this section is to study the notion of Asymptotic Preserving (AP) schemes for the slow-fast SDE system (III.2.1) (averaging regime) or (III.2.12) (diffusion approximation regime) when $\varepsilon \rightarrow 0$. The fundamental requirements to have an AP scheme are the following ones: given a consistent discretization scheme for the SDE system,

- for any fixed time-step size $\Delta t > 0$, there exists a limiting scheme when $\varepsilon \rightarrow 0$,
- this limiting scheme is consistent with the limiting equation (III.2.7) given by Proposition III.2.1 (averaging regime), or the limiting equation (III.2.15) given by Proposition III.2.2 (diffusion approximation regime).

For the SDE considered in this chapter, consistency is understood in the sense of convergence in distribution. As will be clear below, caution is needed in order to satisfy the second requirement, indeed some standard but naive schemes converge to a limiting scheme which is not consistent with the correct limiting equation. Using such schemes would be dangerous since it could lead to wrong conclusions about the behavior of the SDE system

when $\varepsilon \rightarrow 0$, hence the need to develop simultaneously the theoretical and numerical analysis.

After discussing general properties of AP schemes, we will provide example of such schemes both for the system (III.2.1) (averaging regime) and for the system (III.2.12) (diffusion approximation regime). We will also study how this scheme applies to the two examples (III.2.20) and (III.2.22) described above, and provide a few examples of non AP schemes.

III.3.1 Asymptotic preserving schemes: definition and properties

Let $T \in (0, \infty)$, and let $N \in \mathbb{N}$ and $\Delta t = \frac{T}{N}$ denote the time-step size. Let $(\Gamma_n)_{0 \leq n \leq N-1}$ and $(\gamma_n)_{0 \leq n \leq N-1}$ be two independent families of independent standard \mathbb{R}^D and \mathbb{R} -valued Gaussian random variables. The initial conditions X_0^ε and m_0^ε are assumed to satisfy Assumption III.1.

On the one hand, a discretization scheme for the SDE (III.2.1) is defined as

$$\forall n \in \llbracket 0, N-1 \rrbracket, (X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) = \Phi_{\Delta t}^\varepsilon(X_n^\varepsilon, m_n^\varepsilon, \Gamma_n, \gamma_n). \quad (\text{III.3.1})$$

On the other hand, a discretization scheme for the SDE (III.2.12) is defined as

$$\forall n \in \llbracket 0, N-1 \rrbracket, (X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) = \Phi_{\Delta t}^\varepsilon(X_n^\varepsilon, m_n^\varepsilon, \gamma_n). \quad (\text{III.3.2})$$

The presentation is slightly different in the averaging and diffusion approximation regimes. In the remaining of Section III.3.1, only the case of schemes of the type (III.3.1) is considered. This means that if one considers the SDE (III.2.12) and the scheme (III.3.2) (approximation diffusion regime) the variable Γ_n needs to be omitted – this is also the case if $\sigma = 0$ in the SDE (III.2.1) (averaging regime).

The mapping $\Phi_{\Delta t}^\varepsilon$ appearing in the schemes (III.3.1) and (III.3.2) is referred to as the integrator in the sequel.

Let us first discuss stability issues. Due to the presence of factors $\frac{1}{\varepsilon}$ and $\frac{1}{\varepsilon^2}$ in the SDE (III.2.1) and (III.2.12), using the standard Euler-Maruyama scheme would impose strong stability conditions, of the type $\Delta t \leq \Delta t_0(\varepsilon)$ with $\Delta t_0(\varepsilon) \rightarrow 0$ when $\varepsilon \rightarrow 0$. In order to study the behavior of the scheme when $\varepsilon \rightarrow 0$ for any fixed time-step size Δt , it is necessary to avoid such conditions, and we impose the following assumption (which is generally satisfied for some implicit or implicit-explicit methods).

Assumption III.5. The integrator $\Phi_{\Delta t}^\varepsilon$ is defined for all $\varepsilon \in (0, 1]$ and $\Delta t \in (0, \Delta t_0]$, where $\Delta t_0 > 0$ is independent of ε .

We are now in position to study the consistency of the scheme. First, it is assumed that for all $\varepsilon \in (0, 1]$, the scheme (III.3.1) (resp. (III.3.2)) is consistent with the SDE system (III.2.1) (resp. (III.2.12)). When dealing with numerical methods for SDEs, there exist several notions of convergence: in almost sure sense, in probability, in mean-square sense, or in distribution. Since Propositions III.2.1 and III.2.2 state that X^ε converges in distribution to X when $\varepsilon \rightarrow 0$, the relevant notion is consistency in the weak sense, related to convergence in distribution.

Assumption III.6. For all $\varepsilon \in (0, 1]$, the numerical scheme (III.3.1) (resp. (III.3.2)) is consistent in the weak sense with the SDE system (III.2.1) (resp. (III.2.12)): for all bounded continuous functions $\varphi : \mathbb{T}^d \times \mathbb{R}$,

$$\mathbb{E}[\varphi(X_N^\varepsilon, m_N^\varepsilon)] \xrightarrow{N \rightarrow \infty} \mathbb{E}[\varphi(X^\varepsilon(T), m^\varepsilon(T))],$$

where the time-step size is given by $\Delta t = \frac{T}{N}$, for an arbitrary $T \in (0, \infty)$.

Recall that the consistency in the weak sense of the scheme can be verified using the following equivalent criterion, expressed in terms of the integrator and of the infinitesimal generator: for all $\varphi \in \mathcal{C}_b^2(\mathbb{T}^d \times \mathbb{R})$,

$$\lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[\varphi(\Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma))] - \varphi(x, m)}{\Delta t} = \mathcal{L}^\varepsilon \varphi(x, m),$$

for all $(x, m) \in \mathbb{T}^d \times \mathbb{R}$, where Γ and γ are two independent standard \mathbb{R}^D and \mathbb{R} -valued Gaussian random variables.

The requirements above (Assumptions III.5 and III.6) only depend on the behavior of the scheme for fixed $\varepsilon \in (0, 1]$. We are now in position to study the asymptotic behavior as $\varepsilon \rightarrow 0$, with fixed time-step size $\Delta t \in (0, \Delta t_0]$. To introduce the notion of asymptotic preserving scheme, one first needs to assume the existence of a limiting scheme, as follows.

Assumption III.7. For every $\Delta t \in (0, \Delta t_0]$, there exists a mapping $\Phi_{\Delta t} : \mathbb{T}^d \times \mathbb{R}^2 \rightarrow \mathbb{T}^d$, such that for every $(x, m) \in \mathbb{T}^d \times \mathbb{R}$, and every bounded continuous function $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$,

$$\mathbb{E}[\varphi(\Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma))] \xrightarrow{\varepsilon \rightarrow 0} \mathbb{E}[\varphi(\Phi_{\Delta t}(x, \Gamma, \gamma))]$$

where Γ and γ are two independent standard \mathbb{R}^D and \mathbb{R} valued Gaussian random variables.

Let $(X_n)_{0 \leq n \leq N}$ be defined by

$$\begin{aligned} X_{n+1} &= \Phi_{\Delta t}(X_n, \Gamma_n, \gamma_n), \\ X_0 &= x_0 = \lim_{\varepsilon \rightarrow 0} x_0^\varepsilon. \end{aligned} \tag{III.3.3}$$

where $(\Gamma_n)_{0 \leq n \leq N-1}$ and $(\gamma_n)_{0 \leq n \leq N-1}$ are two independent families of independent standard \mathbb{R}^D and \mathbb{R} valued Gaussian random variables. By a recursion argument, it is straightforward to check that if Assumptions III.1 and III.7 are satisfied, then X_n^ε converges in distribution to X_n , when $\varepsilon \rightarrow 0$, for any fixed $\Delta t \in (0, \Delta t_0]$, and $0 \leq n \leq N$.

We are now in position to introduce the notion of asymptotic preserving schemes. As for Assumptions III.6 and III.7 above, the consistency is understood in the sense of convergence in distribution.

Definition III.3.1. Let Assumptions III.5, III.6 and III.7 be satisfied. The scheme (III.3.1) (resp. (III.3.2)) is said to be Asymptotic Preserving (AP) if the limiting scheme given by Assumption III.7 and (III.3.3) is consistent, in the weak sense, with the limiting equation given by Proposition III.2.1 (resp. Proposition III.2.2): for every continuous function $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$, one has

$$\mathbb{E}[\varphi(X_N)] \xrightarrow{N \rightarrow \infty} \mathbb{E}[\varphi(X(T))],$$

where $\Delta t = \frac{T}{N}$, with an arbitrary $T \in (0, \infty)$.

One of the main contributions of this chapter is the design of AP schemes in the averaging and in the diffusion approximation regimes, see Sections III.3.2 and III.3.3 respectively.

To conclude this section, Proposition III.3.1 and Corollary III.3.2 below are general formulations of the AP property in terms of commuting the limits $\varepsilon \rightarrow 0$ and $\Delta t \rightarrow 0$. As explained above, the result is stated only in the averaging regime to simplify the presentation, however the same result holds also in the diffusion approximation regime with straightforward modifications.

Proposition III.3.1. *Let the setting of Definition III.3.1 be satisfied. The following statements are equivalent.*

- (i) *The scheme (III.3.1) is asymptotic preserving.*
- (ii) *For any continuous function $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$, one has*

$$\lim_{\Delta t \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \mathbb{E} [\varphi(X_N^\varepsilon)] = \lim_{\varepsilon \rightarrow 0} \lim_{\Delta t \rightarrow 0} \mathbb{E} [\varphi(X_N^\varepsilon)],$$

where $T = N\Delta t$.

- (iii) *For any $\varphi \in \mathcal{C}^3(\mathbb{T}^d)$, for all $(x, m) \in \mathbb{T}^d \times \mathbb{R}$, one has*

$$\lim_{\Delta t \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \frac{\mathbb{E} [\varphi^\varepsilon(\Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma))] - \varphi(x)}{\Delta t} = \lim_{\varepsilon \rightarrow 0} \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E} [\varphi^\varepsilon(\Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma))] - \varphi(x)}{\Delta t},$$

where $\varphi^\varepsilon = \varphi + \varepsilon\varphi^1$ is the function introduced by the perturbed test function approach (see (III.2.9), Proposition III.2.1 or (III.2.17), Proposition III.2.2), and Γ and γ are independent \mathbb{R}^D and \mathbb{R} valued standard Gaussian random variables.

Note that using the perturbed test function approach (see Propositions III.2.1 and III.2.2) is the relevant point of view for the statement (iii) above.

Proof of Proposition III.3.1. The equivalence of (i) and (ii) is straightforward. Indeed

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \mathbb{E} [\varphi(X_N^\varepsilon)] &= \lim_{\Delta t \rightarrow 0} \mathbb{E} [\varphi(X_N)], \\ \lim_{\varepsilon \rightarrow 0} \lim_{\Delta t \rightarrow 0} \mathbb{E} [\varphi(X_N^\varepsilon)] &= \lim_{\varepsilon \rightarrow 0} \mathbb{E} [\varphi(X^\varepsilon(T))] = \mathbb{E} [\varphi(X(T))], \end{aligned}$$

using Assumptions III.6 and III.7 and Proposition III.2.1. The two quantities coincide if and only if the limiting scheme is consistent with the limiting equation.

It remains to prove that (i) and (iii) are equivalent. On the one hand, note that

$$\lim_{\Delta t \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \frac{\mathbb{E} [\varphi^\varepsilon(\Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma))] - \varphi(x)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E} [\varphi(\Phi_{\Delta t}(x, \Gamma, \gamma))] - \varphi(x)}{\Delta t},$$

using the fact that $\varphi^\varepsilon - \varphi = O(\varepsilon)$ and the definition of the limiting scheme from Assumption III.7.

On the other hand, one has

$$\lim_{\varepsilon \rightarrow 0} \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E} [\varphi^\varepsilon(\Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma))] - \varphi(x)}{\Delta t} = \lim_{\varepsilon \rightarrow 0} \mathcal{L}^\varepsilon \varphi^\varepsilon(x, m) = \mathcal{L}\varphi(x),$$

using the consistency of the scheme for fixed ε (Assumption III.6), and the property (III.2.10), by construction of the perturbed test function φ^ε .

Then (iii) is equivalent to having

$$\lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[\varphi(\Phi_{\Delta t}(x, \gamma))] - \varphi(x)}{\Delta t} = \mathcal{L}\varphi(x),$$

which means consistency in the weak sense of the limiting scheme with the limiting equation (III.2.7).

This concludes the proof of Proposition III.3.1. \square

The following result is a simple criterion to check whether a scheme satisfies the asymptotic preserving property.

Corollary III.3.2. *Assume that for all $\varphi \in \mathcal{C}^2(\mathbb{T}^d)$, one has*

$$\tilde{\mathcal{L}}\varphi(x) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[\varphi(\Phi_{\Delta t}(x, \gamma))] - \varphi(x)}{\Delta t}$$

where $\tilde{\mathcal{L}}$ is a second-order differential operator.

Then the scheme is AP if and only if the property stated in (iii) in Proposition III.3.1 holds with $\varphi(x) = x_i$ and $\varphi(x) = x_i x_j$, with $1 \leq i, j \leq d$.

The proof of Corollary III.3.2 is straightforward and is thus omitted.

III.3.2 An example of AP scheme in the averaging regime

The objective of this section is to propose an example of AP for the SDE model (III.2.1), see Theorem III.3.3, in the averaging regime. The challenge is to capture the averaged coefficients \bar{b} and $\bar{\sigma}$, given by (III.2.5) and (III.2.6).

Theorem III.3.3. *Introduce the following numerical scheme:*

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon) + \sqrt{\Delta t} \sigma(X_n^\varepsilon, m_{n+1}^\varepsilon) \Gamma_n \\ m_{n+1}^\varepsilon = c\left(\frac{\Delta t}{\varepsilon}\right) m_n^\varepsilon + \sqrt{1 - c\left(\frac{\Delta t}{\varepsilon}\right)^2} h(X_n^\varepsilon) \gamma_n, \end{cases} \quad (\text{III.3.4})$$

with $c(\tau) = e^{-\tau}$. This scheme satisfies Assumptions III.5, III.6 and III.7 and is asymptotic preserving in the sense of Definition III.3.1. Moreover the limiting scheme is given by

$$X_{n+1} = X_n + \Delta t b(X_n, h(X_n) \gamma_n) + \sqrt{\Delta t} \sigma(X_n, h(X_n) \gamma_n) \Gamma_n, \quad (\text{III.3.5})$$

Remark III.3.1. The results of this section can be adapted with $c(\tau) = \frac{1}{1+\tau}$. This gives another AP scheme which is used in Section III.6.

Remark III.3.2. As opposed to the AP scheme we construct in the diffusion approximation regime (see Section III.3.3), Scheme (III.3.4) is not based on the Heun method (referred to as the prediction-correction procedure in Section III.3.3). Even though the Heun method is a standard way to increase the order of convergence, it would not give an AP scheme in the averaging regime, due to the identity $\bar{b}\bar{b} \neq \bar{b}'\bar{b}$.

Let us discuss some properties of the AP scheme (III.3.4) and of the limiting scheme (III.3.5). To simplify the discussion, assume that $h(x) = 1$. First, assume that $\sigma = 0$. Note that even if the limiting equation (III.2.7) is a deterministic ordinary differential equation, the scheme (III.3.5) is random. However, in that case, the convergence of X_N to $X(T)$ when $\Delta t \rightarrow 0$ holds in probability, instead of only in distribution; in that case, the averaging principle result stated in Proposition III.2.1 also holds in probability (and even in mean-square sense). The fundamental property to obtain the AP property is that the random quantity appearing in the limiting scheme (III.3.5) satisfies the property

$$\mathbb{E}[b(X_n, h(X_n)\gamma_n)|X_n] = \bar{b}(X_n). \quad (\text{III.3.6})$$

In the AP scheme (III.3.4), the fast component is discretized exactly in distribution (when $h(x) = 1$): for all $n \in \mathbb{N}_0$, the Gaussian random variables m_n^ε and $m^\varepsilon(n\Delta t)$ are equal in distribution. The fundamental property written above cannot be satisfied if one uses for instance the implicit Euler scheme to discretize the fast component: the scheme defined by

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon) \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon + \sqrt{2\frac{\Delta t}{\varepsilon}} \gamma_n, \end{cases} \quad (\text{III.3.7})$$

is not asymptotic preserving, since the associated limiting scheme is

$$X_{n+1} = X_n + \Delta t b(X_n, 0). \quad (\text{III.3.8})$$

using the identity

$$m_{n+1}^\varepsilon = \frac{1}{1 + \frac{\Delta t}{\varepsilon}} m_n^\varepsilon + \frac{\sqrt{2\frac{\Delta t}{\varepsilon}}}{1 + \frac{\Delta t}{\varepsilon}} \gamma_n \xrightarrow{\varepsilon \rightarrow 0} 0,$$

to pass to the limit.

Second, assume that σ is not equal to 0. Then the convergence of X_n^ε to X_n only holds in distribution in general. It does not hold in mean-square sense some cases, in which the convergence in Proposition III.2.1 also does not hold in the mean-square sense. It is thus natural to consider convergence in distribution in the notion of asymptotic preserving schemes for SDEs.

Finally, note also that, as above, the scheme

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \sqrt{\Delta t} \sigma(X_n^\varepsilon, m_{n+1}^\varepsilon) \Gamma_n \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon + \sqrt{2\frac{\Delta t}{\varepsilon}} \gamma_n, \end{cases}$$

is not asymptotic preserving, since the associated limiting scheme is

$$X_{n+1} = X_n + \sqrt{\Delta t} \sigma(X_n, 0) \Gamma_n.$$

We are now in position to prove Theorem III.3.3.

Proof of Theorem III.3.3. It is straightforward to check that Assumption III.5 is satisfied. Let us prove that Assumption III.7 holds. We have

$$\begin{aligned} \Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma) &= x + \Delta t b(x, m') + \sqrt{\Delta t} \sigma(x, m') \Gamma \\ \Phi_{\Delta t}(x, \Gamma, \gamma) &= x + \Delta t b(x, h(x)\gamma) + \sqrt{\Delta t} \sigma(x, h(x)\gamma) \Gamma, \end{aligned}$$

with $m' = e^{-\frac{\Delta t}{\varepsilon}} m + \sqrt{1 - e^{-\frac{2\Delta t}{\varepsilon}}} h(x)\gamma$. When $\varepsilon \rightarrow 0$, m' converges almost surely to $h(x)\gamma$, thus $\Phi_{\Delta t}^\varepsilon(x, m, \Gamma, \gamma)$ converges in distribution to $\Phi_{\Delta t}(x, \Gamma, \gamma)$, and Assumption III.7 is satisfied.

It remains to prove that the scheme satisfies Assumption III.6 and is asymptotic preserving in the sense of Definition III.3.1, namely that the schemes (III.3.4) and (III.3.5) are consistent (in the weak sense), with (III.2.1) and (III.2.7) respectively.

Let $\varepsilon > 0$ be fixed. Since h is bounded, it is straightforward to check that

$$\sup_{\varepsilon \in (0,1]} \sup_{n \in \mathbb{N}_0} \frac{\mathbb{E}[|m_n^\varepsilon|^2]}{1 + |m_0^\varepsilon|^2} < \infty. \quad (\text{III.3.9})$$

This estimate will prove useful to prove Lemma III.4.4 of Theorem III.3.4. Since b and σ are also bounded, we get, in $\mathbb{L}^1(\Omega)$, when $\Delta t \rightarrow 0$

$$\begin{aligned} m_{n+1}^\varepsilon - m_n^\varepsilon &= -\frac{\Delta t}{\varepsilon} m_n^\varepsilon + \sqrt{\frac{2\Delta t}{\varepsilon}} h(X_n^\varepsilon) \gamma_n + o_\varepsilon(\Delta t), \\ X_{n+1}^\varepsilon - X_n^\varepsilon &= \Delta t b(X_n^\varepsilon) + \sqrt{\Delta t} \sigma(X_n^\varepsilon) \Gamma_n + o_\varepsilon(\Delta t). \end{aligned}$$

Thus, using that Γ_n , γ_n and X_n^ε are independent, we get the second order Taylor expansion of $\varphi \in \mathcal{C}_b^2(\mathbb{T}^d \times \mathbb{R})$,

$$\begin{aligned} \mathbb{E}[\varphi(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon)] - \mathbb{E}[\varphi(X_n^\varepsilon, m_n^\varepsilon)] &= \Delta t \mathbb{E}[b(X_n^\varepsilon, m_{n+1}^\varepsilon) \cdot \nabla_x \varphi(X_n^\varepsilon, m_n^\varepsilon)] \\ &\quad + \frac{1}{2} \Delta t \mathbb{E}[\sigma \sigma^*(X_n^\varepsilon, m_{n+1}^\varepsilon) : \nabla_x^2 \varphi(X_n^\varepsilon, m_n^\varepsilon)] \\ &\quad - \frac{\Delta t}{\varepsilon} \mathbb{E}[m_n^\varepsilon \partial_m \varphi(X_n^\varepsilon, m_n^\varepsilon)] \\ &\quad + \frac{\Delta t}{\varepsilon} \mathbb{E}[h(X_n^\varepsilon)^2 \partial_m^2 \varphi(X_n^\varepsilon, m_n^\varepsilon)] + o_\varepsilon(\Delta t) \\ &= \Delta t \mathbb{E}[\mathcal{L}^\varepsilon \varphi(X_n^\varepsilon, m_n^\varepsilon)] + o_\varepsilon(\Delta t). \end{aligned}$$

From there, it is straightforward to check that Assumption III.6 is satisfied.

Similarly, to prove the consistency of the limiting scheme (III.3.5) with (III.2.7), for $\varphi \in \mathcal{C}^2(\mathbb{T}^d)$, when $\Delta t \rightarrow 0$, observe that one has

$$\begin{aligned} \mathbb{E}[\varphi(X_{n+1})] - \mathbb{E}[\varphi(X_n)] &= \Delta t \mathbb{E}[b(X_n, h(X_n)\gamma_n) \cdot \nabla_x \varphi(X_n)] \\ &\quad + \frac{1}{2} \Delta t \mathbb{E}[\sigma \sigma^*(X_n, h(X_n)\gamma_n) : \nabla_x^2 \varphi(X_n)] + o(\Delta t). \end{aligned}$$

The key argument of this proof is the following: by conditioning with respect to X_n and the definitions (III.2.5)–(III.2.6) of the averaged coefficients, using the fundamental property (III.3.6) for \bar{b} and $\bar{\sigma} \bar{\sigma}^* = \overline{\sigma \sigma^*}$, yields

$$\mathbb{E}[\varphi(X_{n+1})] = \mathbb{E}[\varphi(X_n)] + \Delta t \mathbb{E}[\mathcal{L} \varphi(X_n)] + o(\Delta t).$$

The limiting scheme is thus consistent with the limiting equation. This concludes the proof of Theorem III.3.3. \square

Beyond the asymptotic preserving property, it is possible to obtain error estimate, and to prove that the scheme (III.3.4) given in Theorem III.3.3 is uniformly accurate (in distribution).

Theorem III.3.4 (Uniform accuracy). *Let Assumptions III.1, III.2 and III.3 be satisfied. Assume moreover that there exists $C_0 \in (0, \infty)$ such that, for $\varepsilon \in (0, 1]$, $d_{\mathbb{T}^d}(x_0^\varepsilon, x_0) \leq C_0 \varepsilon$, where $d_{\mathbb{T}^d}$ is the distance on \mathbb{T}^d . Then, for any $T \in (0, \infty)$ and any function $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$ of class \mathcal{C}^4 , there exists $C(T, \varphi) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ and $\varepsilon \in (0, 1]$ one has*

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \min\left(\frac{\Delta t}{\varepsilon}, \Delta t + \varepsilon\right), \quad (\text{III.3.10})$$

and the scheme (III.3.4) is uniformly accurate with the following error estimate: for all $\Delta t \in (0, \Delta t_0]$, one has

$$\sup_{\varepsilon \in (0, 1]} |\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \sqrt{\Delta t}. \quad (\text{III.3.11})$$

The error estimate (III.3.11) implies that the error $|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]|$ goes to 0 when $\Delta t \rightarrow 0$ uniformly with respect to $\varepsilon \rightarrow 0$. Note that (III.3.11) is a straightforward consequence of (III.3.10), considering the cases $\sqrt{\Delta t} \leq \varepsilon$ and $\varepsilon \leq \sqrt{\Delta t}$ separately. This argument implies a reduction in the order of convergence appearing in (III.3.11): it is equal to $\frac{1}{2}$ whereas for fixed $\varepsilon > 0$ (in (III.3.10)) or when $\varepsilon = 0$ the order of convergence is equal to 1.

The proof of Theorem III.3.4 is long, technical and requires several auxiliary results, it is thus postponed to Section III.4.

III.3.3 An example of AP scheme in the diffusion approximation regime

The objective of this section is to propose an example of AP scheme for the SDE model (III.2.12), see Theorem III.3.5, in the diffusion approximation regime. The challenge is to let the limiting scheme capture the additional drift term appearing in the limiting equation (III.2.15) when σ or f is not constant.

Theorem III.3.5. *Let $\theta \in [\frac{1}{2}, 1]$. Introduce the following numerical scheme:*

$$\left\{ \begin{array}{l} \hat{m}_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(X_n^\varepsilon) \hat{m}_{n+\theta}^\varepsilon}{\varepsilon^2} + \frac{\Delta t f(X_n^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon}, \\ \hat{X}_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \sigma(X_n^\varepsilon) \frac{\Delta t \hat{m}_{n+\theta}^\varepsilon}{\varepsilon}, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon) m_{n+\theta}^\varepsilon}{\varepsilon^2} + \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon}, \\ Y_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \sigma(X_n^\varepsilon) \frac{\Delta t m_{n+\theta}^\varepsilon}{\varepsilon}, \\ X_{n+1}^\varepsilon = X_n^\varepsilon + \Delta t b(X_n^\varepsilon) + \frac{\sigma(X_n^\varepsilon) + \sigma(Y_{n+1}^\varepsilon)}{2} \frac{\Delta t}{\varepsilon} \frac{\hat{m}_{n+\theta}^\varepsilon + m_{n+\theta}^\varepsilon}{2}, \end{array} \right. \quad (\text{III.3.12})$$

where

$$\begin{aligned} \hat{m}_{n+\theta}^\varepsilon &= (1 - \theta) m_n^\varepsilon + \theta \hat{m}_{n+1}^\varepsilon \\ m_{n+\theta}^\varepsilon &= (1 - \theta) m_n^\varepsilon + \theta m_{n+1}^\varepsilon. \end{aligned}$$

This scheme satisfies Assumptions III.5, III.6 and III.7 and is asymptotic preserving in the sense of Definition III.3.1. Moreover the limiting scheme is given by

$$\begin{cases} \hat{X}_{n+1} = X_n + \Delta t (b(X_n) + g(X_n)\sigma(X_n)) + \sigma(X_n)h(X_n)\sqrt{\Delta t}\gamma_n, \\ Y_{n+1} = X_n + \Delta t (b(X_n) + g(X_n)\sigma(X_n)) + \sigma(X_n)h(X_n)\frac{f(X_n)}{f(\hat{X}_{n+1})}\sqrt{\Delta t}\gamma_n, \\ X_{n+1} = X_n + \Delta t \left(b(X_n) + g(X_n)\frac{\sigma(X_n) + \sigma(Y_{n+1})}{2} \right) \\ \quad + \frac{\sigma(X_n) + \sigma(Y_{n+1})}{2} \frac{1 + \frac{f(X_n)}{f(\hat{X}_{n+1})}}{2} h(X_n)\sqrt{\Delta t}\gamma_n. \end{cases} \quad (\text{III.3.13})$$

The design of the scheme III.3.12 is based on a carefully chosen prediction-correction procedure. The limiting scheme (III.3.13) then also contains prediction steps which are the key elements to satisfy the consistency with the limiting SDE (III.2.15). The choice of the prediction-correction procedure is made clearer looking at the two examples (III.2.20) and (III.2.22), see below Corollaries III.3.6 and III.3.7 respectively. The prediction-correction procedure is crucial to obtain the AP property for the scheme: the following simpler scheme (with $\theta = 1$ to simplify the presentation)

$$\begin{cases} X_{n+1}^\varepsilon = X_n + \Delta t b(X_n^\varepsilon) + \sigma(X_n^\varepsilon) \frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{f(X_n^\varepsilon)\Delta t}{\varepsilon^2} m_{n+1}^\varepsilon + \frac{f(X_n^\varepsilon)g(X_n^\varepsilon)\Delta t}{\varepsilon} + \frac{f(X_n^\varepsilon)h(X_n^\varepsilon)\sqrt{\Delta t}}{\varepsilon} \gamma_n, \end{cases} \quad (\text{III.3.14})$$

is not asymptotic preserving, since the associated limiting scheme (see the proof of Theorem III.3.5 for the derivation of the limiting scheme) is

$$X_{n+1} = X_n + \Delta t (b(X_n) + g(X_n)\sigma(X_n)) + h(X_n)\sigma(X_n)\sqrt{\Delta t}\gamma_n. \quad (\text{III.3.15})$$

This limiting scheme is consistent with the SDE

$$dX_t = (b(X_t) + g(X_t)\sigma(X_t)) dt + h(X_t)\sigma(X_t)dW_t,$$

which differs in general – when σ or f is non constant – from the correct limiting equation (III.2.15).

Observe that in the AP scheme (III.3.12) the fast component m^ε is discretized using the θ -method. Choosing $\theta \in [\frac{1}{2}, 1]$ ensures the mean-square stability of the scheme (Assumption III.5), uniformly with respect to ε . Note that the same quantity $(1 - \theta)m_n^\varepsilon + \theta\hat{m}_{n+1}^\varepsilon$ appears in the expressions of $\hat{m}_{n+1}^\varepsilon$ and $\hat{X}_{n+1}^\varepsilon$ in (III.3.12). Similarly, the same quantity $(1 - \theta)m_n^\varepsilon + \theta m_{n+1}^\varepsilon$ appears in the expressions of m_{n+1}^ε and Y_{n+1}^ε in (III.3.12): this highlights the fact that in order to get a limiting scheme, it is fundamental to choose the quadrature rules in this consistent way.

Remark III.3.3. There would be no loss of generality to assume that $b = 0$. Another example of AP scheme would be obtained in the case $b \neq 0$, using a splitting technique: combining the scheme (III.3.12) with $b = 0$, with a standard explicit Euler scheme to treat the contribution of b . Writing the expression of the resulting scheme is left to the reader.

Proof of Theorem III.3.5. It is straightforward to check that Assumption III.5 is satisfied.

Let us prove that Assumption III.7 holds, namely that (III.3.12) converges to (III.3.13) when $\varepsilon \rightarrow 0$. Note that for fixed $\Delta t > 0$ and $0 \leq n \leq N$, one has

$$\sup_{\varepsilon > 0} \mathbb{E}[|m_n^\varepsilon| + |\hat{m}_n^\varepsilon|] < +\infty. \quad (\text{III.3.16})$$

This is proved by a straightforward recursion argument. As a consequence, one obtains convergence of the quantity,

$$\frac{\Delta t \hat{m}_{n+\theta}^\varepsilon}{\varepsilon} = \Delta t g(X_n^\varepsilon) + h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n - \frac{\varepsilon}{f(X_n^\varepsilon)} (\hat{m}_{n+1}^\varepsilon - m_n^\varepsilon).$$

Thus one has $Y_{n+1}^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} Y_{n+1}$ and $\hat{X}_{n+1}^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \hat{X}_{n+1}$. Similarly, one obtains the convergence of $\frac{\Delta t m_{n+\theta}^\varepsilon}{\varepsilon}$, which yields $Y_{n+1}^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} Y_{n+1}$ and $X_{n+1}^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} X_{n+1}$.

It remains to prove that the scheme satisfies Assumption III.6 and is asymptotic preserving in the sense of Definition III.3.1, namely that the schemes (III.3.12) and (III.3.13) are consistent (in the weak sense), with (III.2.12) and (III.2.15) respectively.

On the one hand, let $\varepsilon > 0$ be fixed. To prove that (III.3.12) is consistent with (III.2.12), it is sufficient to prove that, for $\varphi \in \mathcal{C}_b^2(\mathbb{T}^d \times \mathbb{R})$, when $\Delta t \rightarrow 0$,

$$\mathbb{E}[\varphi(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon)] = \mathbb{E}[\varphi(X_n^\varepsilon, m_n^\varepsilon)] + \Delta t \mathbb{E}[\mathcal{L}^\varepsilon \varphi(X_n^\varepsilon, m_n^\varepsilon)] + o_\varepsilon(\Delta t). \quad (\text{III.3.17})$$

It is straightforward to check that, in $\mathbb{L}^1(\Omega)$,

$$\left(\hat{m}_{n+1}^\varepsilon, \hat{X}_{n+1}^\varepsilon, m_{n+1}^\varepsilon, Y_{n+1}^\varepsilon \right) = \left(\hat{m}_n^\varepsilon, X_n^\varepsilon, m_n^\varepsilon, X_n^\varepsilon \right) + o_\varepsilon(1),$$

hence

$$\begin{aligned} m_{n+1}^\varepsilon &= m_n^\varepsilon + \sqrt{\Delta t} \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \gamma_n}{\varepsilon} + \Delta t \left(-\frac{f(X_n^\varepsilon) m_n^\varepsilon}{\varepsilon^2} + \frac{f(X_n^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} \right) + o_\varepsilon(\Delta t), \\ X_{n+1}^\varepsilon &= X_n^\varepsilon + \Delta t \left(b(X_n^\varepsilon) + \frac{\sigma(X_n^\varepsilon) m_n^\varepsilon}{\varepsilon} \right) + o_\varepsilon(\Delta t). \end{aligned}$$

Since $\gamma_n \sim \mathcal{N}(0, 1)$ and the random variables γ_n and X_n are independent, one obtains (III.3.17).

On the other hand, it remains to prove that the limiting scheme (III.3.13) is consistent with (III.2.15), i.e. that, for $\varphi \in \mathcal{C}^2(\mathbb{T}^d)$, when $\Delta t \rightarrow 0$,

$$\mathbb{E}[\varphi(X_{n+1})] = \mathbb{E}[\varphi(X_n)] + \Delta t \mathbb{E}[\mathcal{L} \varphi(X_n)] + o(\Delta t). \quad (\text{III.3.18})$$

To simplify the presentation, for any function ψ , the following notation is used below:

$$\psi_n \doteq \psi(X_n), \quad \nabla \psi_n \doteq \nabla \psi(X_n)$$

The key argument of this proof is the analysis of the asymptotic behavior of the quantity $\frac{\sigma(X_n) + \sigma(Y_{n+1})}{2} \frac{1 + \frac{f(X_n)}{f(\hat{X}_{n+1})}}{2}$, which appears in the scheme in order to capture the drift terms in the limiting equation (III.2.15).

First, performing expansions at order $\sqrt{\Delta t}$ for \hat{X}_{n+1} and Y_{n+1} yields

$$\hat{X}_{n+1} = X_n + \sqrt{\Delta t} h_n \gamma_n \sigma_n + o(\sqrt{\Delta t}), \quad Y_{n+1} = X_n + \sqrt{\Delta t} h_n \gamma_n \sigma_n + o(\sqrt{\Delta t}).$$

Second, writing $\frac{\sigma_n + \sigma(Y_{n+1})}{2} = \sigma_n + \frac{\sigma(Y_{n+1}) - \sigma_n}{2}$ and $\frac{f_n}{f(\bar{X}_{n+1})} = \frac{1}{1 + \frac{f(\bar{X}_{n+1}) - f_n}{f_n}}$, one obtains the following expansion at order $\sqrt{\Delta t}$ for the quantity:

$$\begin{aligned} & \frac{\sigma_n + \sigma(Y_{n+1})}{2} \frac{1 + \frac{f_n}{f(\bar{X}_{n+1})}}{2} \\ &= \left(\sigma_n + \frac{1}{2} \sqrt{\Delta t} h_n \gamma_n (\sigma_n \cdot \nabla) \sigma_n + o(\sqrt{\Delta t}) \right) \left(1 - \frac{1}{2 f_n} \sqrt{\Delta t} h_n \gamma_n \sigma_n \cdot \nabla f_n + o(\sqrt{\Delta t}) \right) \\ &= \sigma_n + \sqrt{\Delta t} \left(-\frac{1}{2 f_n} h_n \gamma_n \sigma_n \cdot \nabla f_n \sigma_n + \frac{1}{2} h_n \gamma_n (\sigma_n \cdot \nabla) \sigma_n \right) + o(\sqrt{\Delta t}). \end{aligned}$$

Finally, one obtains the following asymptotic expansion of X_{n+1}

$$X_{n+1} = X_n + \sqrt{\Delta t} h_n \gamma_n \sigma_n + \Delta t \left(b_n + g_n \sigma_n - \frac{1}{2 f_n} h_n^2 \gamma_n^2 \sigma_n \cdot \nabla f_n \sigma_n + \frac{1}{2} h_n^2 \gamma_n^2 (\sigma_n \cdot \nabla) \sigma_n \right) + o(\Delta t).$$

Since γ_n is centered and γ_n and X_n are independent random variables, one obtains the first order expansion (III.3.18).

This concludes the proof of Theorem III.3.5. \square

The proposed AP scheme given by Theorem III.3.5 can be simplified when it is applied to one of the two examples of SDE models introduced in Section III.2.2. These two examples are employed in the numerical experiments below. To simplify the presentation, we only consider the case $\theta = 1$.

Corollary III.3.6. *Consider the SDE (III.2.20). The AP scheme (III.3.12) given by Theorem III.3.5 is written as follows:*

$$\begin{cases} m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2} m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon} \gamma_n, \\ Y_{n+1}^\varepsilon = X_n^\varepsilon + \sigma(X_n^\varepsilon) \frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}, \\ X_{n+1}^\varepsilon = X_n^\varepsilon + \frac{\sigma(X_n^\varepsilon) + \sigma(Y_{n+1}^\varepsilon)}{2} \frac{\Delta t m_{n+1}^\varepsilon}{\varepsilon}. \end{cases} \quad (\text{III.3.19})$$

The scheme (III.3.19) is asymptotic preserving, and the limiting scheme (III.3.13) is written as

$$\begin{cases} Y_{n+1} = X_n + \sqrt{\Delta t} \sigma(X_n) \gamma_n, \\ X_{n+1} = X_n + \sqrt{\Delta t} \frac{\sigma(X_n) + \sigma(Y_{n+1})}{2} \gamma_n. \end{cases} \quad (\text{III.3.20})$$

The limiting scheme (III.3.20) is consistent with the limiting SDE (III.2.21).

The prediction-correction procedure appearing in the limiting scheme (III.3.20) allows to recover the Stratonovich interpretation of the noise in the limiting SDE (III.2.21).

Corollary III.3.7. Consider the SDE (III.2.22). The AP scheme (III.3.12) given by Theorem III.3.5 is written as follows:

$$\begin{cases} \hat{m}_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(X_n^\varepsilon)}{\varepsilon^2} \hat{m}_{n+1}^\varepsilon + \frac{\Delta t f(X_n^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon} \\ \hat{X}_{n+1}^\varepsilon = X_n^\varepsilon + \frac{\Delta t \hat{m}_{n+1}^\varepsilon}{\varepsilon} \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon)}{\varepsilon^2} m_{n+1}^\varepsilon + \frac{\Delta t f(\hat{X}_{n+1}^\varepsilon) g(X_n^\varepsilon)}{\varepsilon} + \frac{f(X_n^\varepsilon) h(X_n^\varepsilon) \sqrt{\Delta t} \gamma_n}{\varepsilon} \\ X_{n+1}^\varepsilon = X_n^\varepsilon + \frac{\Delta t}{\varepsilon} \frac{\hat{m}_{n+1}^\varepsilon + m_{n+1}^\varepsilon}{2} \end{cases} \quad (\text{III.3.21})$$

The scheme (III.3.21) is asymptotic preserving, and the limiting scheme (III.3.13) is written as

$$\begin{cases} \hat{X}_{n+1} = X_n + \Delta t g(X_n) + h(X_n) \sqrt{\Delta t} \gamma_n, \\ X_{n+1} = X_n + \Delta t g(X_n) + \frac{1 + \frac{f(X_n)}{f(\hat{X}_{n+1})}}{2} h(X_n) \sqrt{\Delta t} \gamma_n. \end{cases} \quad (\text{III.3.22})$$

The limiting scheme (III.3.22) is consistent with the limiting SDE (III.2.23).

The prediction-correction procedure appearing in the limiting scheme (III.3.20) allows to recover the noise-induced drift term appearing in the limiting SDE (III.2.23).

Remark III.3.4. Consider the first example (III.2.20), with $\sigma(x) = x$ and assume that x takes values in the real line \mathbb{R} (instead of the torus \mathbb{T}). The following scheme

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon \exp\left(\frac{\Delta t}{\varepsilon} [(1-\theta)m_n^\varepsilon + \theta m_{n+1}^\varepsilon]\right) \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2} [(1-\theta)m_n^\varepsilon + \theta m_{n+1}^\varepsilon] + \frac{\sqrt{\Delta t}}{\varepsilon} \gamma_n, \end{cases} \quad (\text{III.3.23})$$

where $\theta \in [\frac{1}{2}, 1]$, is another example of AP scheme for this problem. The limiting scheme is given by

$$X_{n+1} = X_n \exp(\sqrt{\Delta t} \gamma_n) \quad (\text{III.3.24})$$

which is consistent with the limiting equation $dX_t = X_t \circ dW_t$.

However, the construction is more subtle if the fast component is discretized using an exponential method: let

$$m_{n+1}^\varepsilon = e^{-\frac{\Delta t}{\varepsilon^2}} m_n^\varepsilon + \sqrt{\frac{1 - e^{-2\frac{\Delta t}{\varepsilon^2}}}{2}} \gamma_n,$$

then $m_{n+1}^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \frac{1}{\sqrt{2}} \gamma_n$, and defining

$$X_{n+1}^\varepsilon = X_n^\varepsilon \exp\left(\frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon\right)$$

does not provide an AP scheme, since there exists no limiting scheme when $\varepsilon \rightarrow 0$.

Inspired by the identity $\frac{1}{\varepsilon} \int_0^t m^\varepsilon(s) ds = \varepsilon(m_0^\varepsilon - m^\varepsilon(t)) + \beta_t$ for all $t \in \mathbb{R}^+$, one may set

$$X_{n+1}^\varepsilon = X_n^\varepsilon \exp\left(\varepsilon(m_n^\varepsilon - m_{n+1}^\varepsilon) + \sqrt{\Delta t} \gamma_n\right),$$

so that one has

$$X_N^\varepsilon = x_0^\varepsilon \exp \left(\varepsilon(m_0^\varepsilon - m_N^\varepsilon) + \sqrt{\Delta t} \sum_{n=0}^{N-1} \gamma_n \right).$$

When $\varepsilon \rightarrow 0$, one obtains the limiting scheme (III.3.24) again. However, the generalization of this construction to the case $\sigma(x) \neq x$ is not straightforward, whereas the scheme proposed in Theorem III.3.5 applies directly to the general case.

III.4 Uniform accuracy and error analysis

III.4.1 Uniform accuracy in the averaging regime (proof of Theorem III.3.4)

The objective of this section is to prove the error estimate (III.3.10). The proof follows from the following four auxiliary lemmas. They are proved in Section III.4.2 below. In their statements, let Assumptions III.1, III.2 and III.3 be satisfied. Let $T \in (0, \infty)$ be fixed and assume that $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$ is of class \mathcal{C}^4 . Recall that the identity $T = N\Delta t$ is assumed to hold. In addition, recall that $(X(t))_{t \in \mathbb{R}^+}$ and $(X_n)_{n \in \mathbb{N}_0}$ are defined by the limiting equation (III.2.7) and the limiting scheme (III.3.5) respectively.

Lemma III.4.1. *There exists $C(T, \varphi) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ and $\varepsilon \in (0, 1]$ one has*

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \frac{\Delta t}{\varepsilon}. \quad (\text{III.4.1})$$

Lemma III.4.2. *There exists $C(T, \varphi) \in (0, \infty)$ such that for all $\varepsilon \in (0, 1]$ one has*

$$|\mathbb{E}[\varphi(X(T))] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi)\varepsilon. \quad (\text{III.4.2})$$

Lemma III.4.3. *There exists $C(T, \varphi) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ one has*

$$|\mathbb{E}[\varphi(X_N)] - \mathbb{E}[\varphi(X(T))]| \leq C(T, \varphi)\Delta t. \quad (\text{III.4.3})$$

Lemma III.4.4. *There exists $C(T, \varphi) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ and $\varepsilon \in (0, 1]$ one has*

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)]| \leq C(T, \varphi) \max(\Delta t, \varepsilon). \quad (\text{III.4.4})$$

The first auxiliary result (Lemma III.4.1) states a weak error estimate for the numerical scheme (III.3.4) for fixed $\varepsilon \in (0, 1]$. Due to the stiffness of the fast component m_n^ε , the right-hand side is not uniform with respect to ε , and it is natural to expect that the upper bound depends on $\tau = \frac{\Delta t}{\varepsilon}$.

The second auxiliary result (Lemma III.4.2) gives an error estimate in the averaging principle (see (III.2.11) in Proposition III.2.1), in the weak sense. This is a standard result in the literature, see for instance [KY05] for an approach using asymptotic expansions for solutions of Kolmogorov equations. The strategy of the proof is based on the introduction of the solutions of relevant Poisson equations, in the spirit of [PS08, Chapter 17] where strong convergence is studied, see [Bré20] and [RSX19] for the weak convergence case.

The two remaining auxiliary lemmas and their proofs are more original than the first two. Lemmas III.4.3 and III.4.4 are quantitative statements concerning two fundamental requirements in the notion of AP scheme (see Definition III.3.1). On the one hand, Lemma

III.4.3 is a quantitative statement of the consistency of the limiting scheme (III.3.5) with the limiting equation (III.2.7), since it provides a weak error when $\Delta t \rightarrow 0$. Since the scheme is not classical (it is not a standard Euler-Maruyama type method, in particular recall that the scheme is random even if $X(T)$ is deterministic, when $\sigma = 0$), a proof is required. On the other hand, Lemma III.4.4 is a quantitative statement about the convergence to the limiting scheme, for fixed $\Delta t \in (0, \Delta t_0]$ (see Assumption III.7). In fact, the left-hand side of (III.4.4) goes to 0 when $\varepsilon \rightarrow 0$, however in the right-hand side of (III.4.4) an additional error term Δt appears. Proving Lemma III.4.4 is the most challenging step towards the proof of Theorem III.3.4, whereas a key argument will be identified in the proof of Lemma III.4.3 related to the consistency of the limiting scheme with the limiting equation.

The following auxiliary results concerning solutions of Kolmogorov equations are required in order to prove the four auxiliary results stated above. They are proved in Section III.4.3 below.

Lemma III.4.5. *Define $u^\varepsilon(t, x) = \mathbb{E}_{x,m}[\varphi(X^\varepsilon(t))]$, for all $t \in \mathbb{R}^+$, $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$, where $(X^\varepsilon, m^\varepsilon)$ is the solution of the SDE system (III.2.1), and $\mathbb{E}_{x,m}$ means that $X^\varepsilon(0) = x$ and $m^\varepsilon(0) = m$. For all $\varepsilon \in (0, 1]$, one has $u^\varepsilon \in \mathcal{C}([0, T], \mathcal{C}_b^3(\mathbb{T}^d \times \mathbb{R}))$. In addition, there exists $C(T, \varphi) \in (0, \infty)$ such that for all $j \in \{1, 2, 3\}$, one has*

$$\sup_{\varepsilon \in (0, 1]} \sup_{(t, x, m) \in [0, T] \times \mathbb{T}^d \times \mathbb{R}} \|D_{x,m}^j u^\varepsilon(t, x, m)\| \leq C(T, \varphi). \quad (\text{III.4.5})$$

Lemma III.4.6. *Define $u(t, x) = \mathbb{E}_x[\varphi(X(t))]$, for all $t \in \mathbb{R}^+$ and $x \in \mathbb{T}^d$, where X is the solution of the SDE (III.2.7) and \mathbb{E}_x means that $X(0) = x$. One has $u \in \mathcal{C}([0, T], \mathcal{C}_b^4(\mathbb{T}^d, \mathbb{R}))$. In addition, there exists $C(T, \varphi) \in (0, \infty)$ such that for all $j \in \{1, 2, 3, 4\}$, one has*

$$\sup_{(t, x) \in [0, T] \times \mathbb{T}^d} \|D_x^j u(t, x)\| \leq C(T, \varphi). \quad (\text{III.4.6})$$

Lemma III.4.7. *Let $\Delta t \in (0, \Delta t_0]$.*

Define $u_n(x) = \mathbb{E}_x[\varphi(X_n)]$, for all $n \in \mathbb{N}$ and $x \in \mathbb{T}^d$, where $(X_n)_{n \in \mathbb{N}_0}$ is defined by the limiting scheme (III.3.5) (see Theorem III.3.3), and \mathbb{E}_x means that $X_0 = x$. For all $n \in \mathbb{N}_0$, one has $u_n \in \mathcal{C}^2(\mathbb{T}^d)$. In addition, there exists $C(T, \varphi) \in (0, \infty)$ such that for all $j \in \{1, 2\}$, one has

$$\sup_{0 \leq n \leq N} \sup_{x \in \mathbb{T}^d} \|D_x^j u_n(x)\| \leq C(T, \varphi) \quad (\text{III.4.7})$$

and, for all $\Delta t \in (0, \Delta t_0]$, one has

$$\sup_{0 \leq n \leq N-1} \sup_{x \in \mathbb{T}^d} \|D_x^j u_{n+1}(x) - D_x^j u_n(x)\| \leq C(T, \varphi) \Delta t. \quad (\text{III.4.8})$$

Based on the auxiliary results stated above, the proof of Theorem III.3.4 is straightforward.

Proof of Theorem III.3.4. Note that

$$\begin{aligned} |\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| &\leq |\mathbb{E}[\varphi(X(T))] - \mathbb{E}[\varphi(X^\varepsilon(T))]| + |\mathbb{E}[\varphi(X_N)] - \mathbb{E}[\varphi(X(T))]| \\ &\quad + |\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)]|, \end{aligned}$$

thus combining (III.4.2), (III.4.3) and (III.4.4) (with $\max(\Delta t, \varepsilon) \leq \Delta t + \varepsilon$) yields

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi)(\Delta t + \varepsilon).$$

Combining that error estimate with (III.4.1) then concludes the proof of the error estimate (III.3.10). As already explained above, the error estimate (III.3.11) is a straightforward consequence of (III.3.10) (considering the cases $\sqrt{\Delta t} \leq \varepsilon$ and $\sqrt{\Delta t} \geq \varepsilon$).

This concludes the proof of Theorem III.3.4. \square

III.4.2 Auxiliary error estimates

Let us now give proofs of the auxiliary lemmas III.4.1, III.4.2, III.4.3 and III.4.4, employing the results of Lemmas III.4.5, III.4.6 and III.4.7 (proofs are given in Section III.4.3 below).

The following notation is used below in the proofs of the auxiliary results: for all $\lambda, \mu \in \mathbb{R}^+$, $\lambda \lesssim \mu$ means that there exists $C(T, \varphi) \in (0, +\infty)$, independent of Δt , ε and n , such that $\lambda \leq C(T, \varphi)\mu$. In addition, the following notation is used for the infinitesimal generator of the Ornstein-Uhlenbeck process:

$$\mathcal{L}_{OU}^x = -m\partial_m + h(x)^2\partial_m^2,$$

in order to let the dependence with respect to x be clear.

Proof of Lemma III.4.1. Let us introduce auxiliary continuous-time processes \tilde{X}^ε and \tilde{m}^ε , such that, for all $n \in \llbracket 0, N \rrbracket$, one has $X_n^\varepsilon = \tilde{X}^\varepsilon(t_n)$ and $m_n^\varepsilon = \tilde{m}^\varepsilon(t_n)$ (recall the definition (III.3.4) of the scheme): for $t_n \leq t \leq t_{n+1}$

$$\begin{aligned} \tilde{X}^\varepsilon(t) &= X_n^\varepsilon + (t - t_n)b(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) + \sigma(X_n^\varepsilon, \tilde{m}^\varepsilon(t))(B(t) - B(t_n)) \\ d\tilde{m}_t^\varepsilon &= -\frac{\tilde{m}_t^\varepsilon}{\varepsilon}dt + \frac{\sqrt{2}h(X_n^\varepsilon)}{\sqrt{\varepsilon}}d\beta_t. \end{aligned}$$

Note that \tilde{m}^ε does satisfy $m_n^\varepsilon = \tilde{m}^\varepsilon(t_n)$, since m_n^ε is exact in distribution and \tilde{m}^ε is an Ornstein-Uhlenbeck process with variance $\frac{2h(X_n^\varepsilon)^2}{\varepsilon}$. The process \tilde{X}^ε satisfies on each subinterval $[t_n, t_{n+1}]$ the following stochastic differential equation: for all $t \in [t_n, t_{n+1}]$, one has

$$\begin{aligned} d\tilde{X}_t^\varepsilon &= b(X_n^\varepsilon, \tilde{m}_t^\varepsilon)dt + \sigma(X_n^\varepsilon, \tilde{m}_t^\varepsilon)dB_t + (t - t_n)\partial_m b(X_n^\varepsilon, \tilde{m}_t^\varepsilon)d\tilde{m}_t^\varepsilon \\ &\quad + \partial_m \sigma(X_n^\varepsilon, \tilde{m}_t^\varepsilon)(B_t - B_{t_n})d\tilde{m}_t^\varepsilon + \frac{h(X_n^\varepsilon)^2}{\varepsilon}(t - t_n)\partial_m^2 b(X_n^\varepsilon, \tilde{m}_t^\varepsilon)dt \\ &\quad + \frac{h(X_n^\varepsilon)^2}{\varepsilon}\partial_m^2 \sigma(X_n^\varepsilon, \tilde{m}_t^\varepsilon)(B_t - B_{t_n})dt. \end{aligned}$$

The expressions for \tilde{X}^ε are complicated due to the fact that in the scheme (III.3.4), b and σ are evaluated with $m = m_{n+1}^\varepsilon$, which is required to satisfy the AP property.

Owing to Lemma III.4.5, the auxiliary function u^ε is of class \mathcal{C}^2 and is solution of the Kolmogorov equation $\partial_t u^\varepsilon = \mathcal{L}^\varepsilon u^\varepsilon$. Using a telescoping sum argument and the definition of the auxiliary processes \tilde{X}^ε and \tilde{m}^ε , the application of Itô's formula yields the following

standard expression for the weak error:

$$\begin{aligned}
\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))] &= \mathbb{E}[u^\varepsilon(0, X_N^\varepsilon, \tilde{m}^\varepsilon(t_N))] - \mathbb{E}[u^\varepsilon(T, X_0^\varepsilon, \tilde{m}^\varepsilon(0))] \\
&= \sum_{n=0}^{N-1} \mathbb{E} \left[u^\varepsilon(T - t_{n+1}, \tilde{X}^\varepsilon(t_{n+1}), \tilde{m}^\varepsilon(t_{n+1})) \right. \\
&\quad \left. - u^\varepsilon(T - t_n, \tilde{X}^\varepsilon(t_n), \tilde{m}^\varepsilon(t_n)) \right] \\
&= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[(-\partial_t + \tilde{\mathcal{L}}_n^\varepsilon) u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \right] dt \\
&= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[(\tilde{\mathcal{L}}_n^\varepsilon - \mathcal{L}^\varepsilon) u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \right] dt,
\end{aligned}$$

where the auxiliary differential operator $\tilde{\mathcal{L}}_n^\varepsilon$ is such that

$$\begin{aligned}
\tilde{\mathcal{L}}_n^\varepsilon u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) &= b(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \cdot \nabla_x u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \\
&\quad + \frac{1}{2} \sigma \sigma^*(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) : \nabla_x^2 u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \\
&\quad + \frac{1}{\varepsilon} \mathcal{L}_{OU}^{X_n^\varepsilon} u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) + r_n^\varepsilon(t),
\end{aligned}$$

and where the remainder term $r_n^\varepsilon(t)$ is given by

$$\begin{aligned}
r_n^\varepsilon(t) &= \frac{1}{\varepsilon} (t - t_n) \mathcal{L}_{OU}^{X_n^\varepsilon} b(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \cdot \nabla_x u^\varepsilon \\
&\quad + \frac{1}{\varepsilon} \mathcal{L}_{OU}^{X_n^\varepsilon} \sigma(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) (B(t) - B(t_n)) \cdot \nabla_x u^\varepsilon \\
&\quad + \frac{h(X_n^\varepsilon)^2}{\varepsilon} D_x^2 u^\varepsilon \cdot ((t - t_n) \partial_m b(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) + \partial_m \sigma(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) (B(t) - B(t_n)))^2 \\
&\quad + \frac{2h(X_n^\varepsilon)^2}{\varepsilon} ((t - t_n) \partial_m b(X_n^\varepsilon, \tilde{m}^\varepsilon(t))) \cdot \nabla_x \partial_m u^\varepsilon \\
&\quad + \frac{2h(X_n^\varepsilon)^2}{\varepsilon} (\partial_m \sigma(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) (B(t) - B(t_n))) \cdot \nabla_x \partial_m u^\varepsilon,
\end{aligned}$$

where u^ε and its derivatives are evaluated at $(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t))$, and where we used the notation $D^2 \varphi(x) \cdot (y)^2 = yy^* : \nabla_x^2 \varphi(x)$ to simplify the presentation.

Let us first deal with the remainder term $r_n^\varepsilon(t)$. Only the second and fifth terms need to be considered carefully. Those two terms are of the form

$$\frac{1}{\varepsilon} \Sigma_n^\varepsilon(t) (B(t) - B(t_n)) \cdot \nabla_x v^\varepsilon(t, \tilde{X}^\varepsilon(t)),$$

with $\Sigma_n^\varepsilon(t)$ a $\mathcal{M}_{d,D}(\mathbb{R})$ -valued bounded random variable and $v^\varepsilon(t, x)$ is a random variable given by $v^\varepsilon(t, x) = u^\varepsilon(T - t, x, \tilde{m}^\varepsilon(t))$ or $v^\varepsilon(t, x) = \partial_m u^\varepsilon(T - t, x, \tilde{m}^\varepsilon(t))$. Observe that those two random variables are $\{X_n^\varepsilon, \tilde{m}^\varepsilon(t)\}$ -measurable and that the processes $(B(t) - B(t_n))_{t \in [t_n, t_{n+1}]}$ and $\{X_n^\varepsilon, (\tilde{m}^\varepsilon(t))_{t \in [t_n, t_{n+1}]}\}$ are independent. Thus, using a conditioning argument, one has

$$\mathbb{E} \left[\frac{1}{\varepsilon} \Sigma_n^\varepsilon(t) (B(t) - B(t_n)) \cdot \nabla_x v^\varepsilon(t, X_n^\varepsilon) \right] = 0.$$

Using that identity and Lemma ??, one obtains for all $t \in [t_n, t_{n+1}]$

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{\varepsilon} \Sigma_n^\varepsilon(t) (B(t) - B(t_n)) \cdot \nabla_x v^\varepsilon(t, \tilde{X}^\varepsilon(t)) \right] \right| \\ & \lesssim \left| \mathbb{E} \left[\frac{1}{\varepsilon} \Sigma_n^\varepsilon(t) (B(t) - B(t_n)) \cdot (\nabla_x v^\varepsilon(t, \tilde{X}^\varepsilon(t)) - \nabla_x v^\varepsilon(t, X_n^\varepsilon)) \right] \right| \\ & \lesssim \frac{t - t_n}{\varepsilon}. \end{aligned}$$

As a consequence,

$$|\mathbb{E}[r_n^\varepsilon(t)]| \lesssim \frac{t - t_n}{\varepsilon}.$$

It remains to deal with

$$\begin{aligned} & (\tilde{\mathcal{L}}_n^\varepsilon - \mathcal{L}^\varepsilon) u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) - r_n^\varepsilon(t) \\ & = \left(b(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) - b(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \right) \cdot \nabla_x u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \\ & \quad + \frac{1}{2} \left(\sigma \sigma^*(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) - \sigma \sigma^*(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \right) : \nabla_x^2 u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \quad (\text{III.4.9}) \\ & \quad + \left(h(X_n^\varepsilon)^2 - h(\tilde{X}^\varepsilon(t))^2 \right) \partial_m^2 u^\varepsilon(T - t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)), \end{aligned}$$

where the expressions (III.2.3) and (III.2.4) for the infinitesimal generator \mathcal{L}^ε have been used. The three quantities appearing in the right-hand side of (III.4.9) above are of the type

$$\left(V(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) - V(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \right) U^\varepsilon(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)),$$

for $V = b, \sigma \sigma^*$ or h^2 and $U^\varepsilon = \nabla_x u^\varepsilon(T - t), \nabla_x^2 u^\varepsilon(T - t)$ or $\partial_m^2 u^\varepsilon(T - t)$. Using again the independence of $(B(t) - B(t_n))_{t_n \leq t \leq t_{n+1}}$ and $\{X_n^\varepsilon, (\tilde{m}^\varepsilon(t))_{t \in [t_n, t_{n+1}]}\}$ and regularity properties of u^ε given in Lemma III.4.5, applying Itô's formula and conditioning with respect to $(X_n^\varepsilon, (\tilde{m}^\varepsilon(t))_{t_n \leq t \leq t_{n+1}})$, one obtains

$$\left| \mathbb{E} \left[\left(V(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) - V(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \right) U^\varepsilon(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \right] \right| \lesssim t - t_n.$$

Moreover, $\|V\|_{\mathcal{C}^1(\mathbb{T}^d, C(\mathbb{R}))} \lesssim 1$ and, owing to Lemma III.4.5, $\|U^\varepsilon\|_{\mathcal{C}^1(\mathbb{T}^d, C(\mathbb{R}))} \lesssim 1$. Therefore, we have

$$\begin{aligned} & \mathbb{E} \left[\left| \left(V(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) - V(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \right) \left(U^\varepsilon(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) - U^\varepsilon(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \right) \right| \right] \\ & \leq \|V(\cdot, \tilde{m}^\varepsilon(t))\|_{\mathcal{C}^1(\mathbb{T}^d)} \mathbb{E} \left[\|U^\varepsilon(\cdot, \tilde{m}^\varepsilon(t))\|_{\mathcal{C}^1(\mathbb{T}^d)} \|\tilde{X}^\varepsilon(t) - X_n^\varepsilon\|^2 \right] \\ & \lesssim \mathbb{E} \left[\|\tilde{X}^\varepsilon(t) - X_n^\varepsilon\|^2 \right] \lesssim t - t_n, \end{aligned}$$

where the last inequality comes from the definition of $\tilde{X}^\varepsilon(t)$. Gathering the two estimates above gives, for all $t \in [t_n, t_{n+1}]$,

$$\left| \mathbb{E} \left[\left(V(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) - V(X_n^\varepsilon, \tilde{m}^\varepsilon(t)) \right) U^\varepsilon(\tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \right] \right| \lesssim t - t_n,$$

and using (III.4.9) finally yields, for all $t \in [t_n, t_{n+1}]$,

$$\mathbb{E} \left[\left| (\tilde{\mathcal{L}}_n^\varepsilon - \mathcal{L}^\varepsilon) u^\varepsilon(T-t, \tilde{X}^\varepsilon(t), m^\varepsilon(t)) \right| \right] \lesssim t - t_n + \frac{t - t_n}{\varepsilon} \lesssim \frac{\Delta t}{\varepsilon}.$$

One then obtains

$$\left| \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[(\tilde{\mathcal{L}}_n^\varepsilon - \mathcal{L}^\varepsilon) u^\varepsilon(T-t, \tilde{X}^\varepsilon(t), \tilde{m}^\varepsilon(t)) \right] dt \right| \lesssim \frac{\Delta t}{\varepsilon},$$

which concludes the proof of (III.4.1) and of Lemma III.4.1. \square

Proof of Lemma III.4.2. For all $x \in \mathbb{T}^d$, introduce the auxiliary Ornstein-Uhlenbeck process m^x solving the SDE

$$dm_t^x = -m_t^x dt + \sqrt{2h(x)} d\beta_t$$

Let $m^x(t, m)$ denote the solution at time t , if the initial condition is given by $m^x(0, m)$. The invariant distribution of the process m^x is equal to ν^x . Note that for m and $m' \in \mathbb{R}$, one has $m^x(t, m) - m^x(t, m') = (m - m')e^{-t}$. Consider $V = b$ or $V = \sigma\sigma^*$, and let

$$\delta(t, x, m) = \mathbb{E} [V(x, m^x(t, m)) - \bar{V}(x)].$$

Note that for m and $m' \in \mathbb{R}$, one has $m^x(t, m) - m^x(t, m') = (m - m')e^{-t}$. As a consequence, we have

$$\|\delta(t, x, m) - \delta(t, x, m')\| \lesssim |m - m'| e^{-t}.$$

By integrating with respect to m' and using the equality $\bar{V}(x) = \int V(x, m') d\nu^x(m')$, one obtains

$$\|\delta(t, x, m)\| \lesssim (1 + |m|) e^{-t}. \quad (\text{III.4.10})$$

Using the fact that δ and its derivatives satisfy (III.4.10) with $V = b$, one is able to check that the function ψ_b given by, for all $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$,

$$\psi_b(x, m) = - \int_0^\infty \mathbb{E} [b(x, m^x(t, m)) - \bar{b}(x)] dt$$

is well-defined (by definition of \bar{b} , see (III.2.5)) and is of class $\mathcal{C}^2(\mathbb{T}^d \times \mathbb{R})$. Moreover, ψ_b and its derivative have at most linear growth in m . In addition, $\psi_b(x, \cdot)$ solves the Poisson equation $\mathcal{L}_{OU}^x \psi_b(x, m) = b(x, m) - \bar{b}(x)$, for all $x \in \mathbb{T}^d$ (indeed ν^x is the invariant distribution associated with the generator \mathcal{L}_{OU}^x of the Ornstein-Uhlenbeck process m^x).

Similarly, define, for all $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$,

$$\psi_\sigma(x, m) = - \int_0^\infty \mathbb{E} [\sigma\sigma^*(x, m^x(t, m)) - \bar{\sigma}\bar{\sigma}^*(x)] dt.$$

The function ψ_σ is well-defined: owing to (III.2.6) (Assumption III.2) one has the equality $\int \sigma\sigma^*(x, m) d\nu^x(m) = \bar{\sigma}\bar{\sigma}^*(x)$ for all $x \in \mathbb{T}^d$, and using the same arguments as above, $\psi_\sigma(x, \cdot)$ solves the Poisson equation $\mathcal{L}_{OU}^x \psi_\sigma(x, m) = \sigma\sigma^*(x, m) - \bar{\sigma}\bar{\sigma}^*(x)$, for all $x \in \mathbb{T}^d$.

Now, for all $t \in [0, T]$, $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$, let

$$\Phi(t, x, m) = \psi_b(x, m) \cdot \nabla_x u(T-t, x) + \psi_\sigma(x, m) : \nabla_x^2 u(T-t, x),$$

where u is given by Lemma III.4.6 and solves the Kolmogorov equation $\partial_t u = \mathcal{L}u$.

On the one hand, applying Itô's formula yields the following expression for the error term:

$$\begin{aligned}
\mathbb{E}[\varphi(X^\varepsilon(T))] - \mathbb{E}[\varphi(X(T))] &= \mathbb{E}[u(0, X^\varepsilon(T)) - u(T, X(0))] \\
&= \mathbb{E}[u(0, X^\varepsilon(T)) - u(T, X^\varepsilon(0))] + u(T, X^\varepsilon(0)) - u(T, X(0)) \\
&= \int_0^T \mathbb{E}[(b(X^\varepsilon(t), m^\varepsilon(t)) - \bar{b}(X^\varepsilon(t))) \cdot \nabla_x u(T-t, X^\varepsilon(t))] dt \\
&\quad + \frac{1}{2} \int_0^T \mathbb{E}[(\sigma\sigma^*(X^\varepsilon(t), m^\varepsilon(t)) - \bar{\sigma}\bar{\sigma}^*(X^\varepsilon(t))) : \nabla_x^2 u(T-t, X^\varepsilon(t))] dt \\
&\quad + u(T, x_0^\varepsilon) - u(T, x_0) \\
&= \int_0^T \mathbb{E}[\mathcal{L}_{OU}^x \Phi(t, X^\varepsilon(t), m^\varepsilon(t))] dt + u(T, x_0^\varepsilon) - u(T, x_0),
\end{aligned}$$

by definition of the auxiliary function Φ , since ψ_b and ψ_σ are solutions of Poisson equations.

On the other hand, applying Itô's formula also gives the identity

$$\begin{aligned}
\mathbb{E}[\Phi(T, X^\varepsilon(T), m^\varepsilon(T))] - \mathbb{E}[\Phi(0, X^\varepsilon(0), m^\varepsilon(0))] \\
= \int_0^T \mathbb{E} \left[\left(\partial_t + b \cdot \nabla_x + \frac{1}{2} \sigma\sigma^* : \nabla_x^2 + \frac{1}{\varepsilon} \mathcal{L}_{OU}^x \right) \Phi(t, X^\varepsilon(t), m^\varepsilon(t)) \right] dt.
\end{aligned}$$

Combining the two expressions then gives

$$\begin{aligned}
\mathbb{E}[\varphi(X^\varepsilon(T))] - \mathbb{E}[\varphi(X(T))] &= \int_0^T \mathbb{E}[\mathcal{L}_{OU}^x \Phi(t, X^\varepsilon(t), m^\varepsilon(t))] dt + u(T, x_0^\varepsilon) - u(T, x_0) \\
&= \varepsilon (\mathbb{E}[\Phi(T, X^\varepsilon(T), m^\varepsilon(T))] - \mathbb{E}[\Phi(0, X^\varepsilon(0), m^\varepsilon(0))]) \\
&\quad - \varepsilon \int_0^T \mathbb{E} \left[\left(\partial_t + b \cdot \nabla_x + \frac{1}{2} \sigma\sigma^* : \nabla_x^2 \right) \Phi(t, X^\varepsilon(t), m^\varepsilon(t)) \right] dt \\
&\quad + u(T, x_0^\varepsilon) - u(T, x_0).
\end{aligned}$$

Using the regularity estimates from Lemma III.4.6, the assumption $d_{\mathbb{T}^d}(x_0^\varepsilon, x_0) \lesssim \varepsilon$ and the identity

$$\partial_t u = \bar{b} \cdot \nabla_x u + \bar{\sigma} \bar{\sigma}^* : \nabla_x^2 u,$$

it is then straightforward to obtain (III.4.2). This concludes the proof of Lemma III.4.2. \square

Proof of Lemma III.4.3. Let us introduce the continuous-time auxiliary process \tilde{X} , such that, for $t \in [t_n, t_{n+1}]$, one has

$$\tilde{X}(t) = X_n + (t - t_n)b(X_n, h(X_n)\gamma_n) + \sigma(X_n, h(X_n)\gamma_n)(B(t) - B(t_n)).$$

Introduce also the second-order differential operator

$$\tilde{\mathcal{L}}_n = b(X_n, h(X_n)\gamma_n) \cdot \nabla + \frac{1}{2} \sigma\sigma^*(X_n, h(X_n)\gamma_n) : \nabla^2.$$

With this notation, for any function $\varphi \in \mathcal{C}^2(\mathbb{T}^d)$, Itô's formula gives, for all $t \in [t_n, t_{n+1}]$,

$$\varphi(\tilde{X}(t)) - \varphi(X_n) = \int_{t_n}^t \tilde{\mathcal{L}}_n \varphi(\tilde{X}(s)) ds + \int_{t_n}^t \nabla \varphi(\tilde{X}(s)) \cdot \sigma(X_n, h(X_n)\gamma_n) dB_s. \quad (\text{III.4.11})$$

Using the same (standard) arguments as in the proof of Lemma III.4.1, one obtains the following decomposition of the error:

$$\begin{aligned}
\mathbb{E} [\varphi(X_N)] - \mathbb{E} [\varphi(X(T))] &= \mathbb{E} [u(0, X_N)] - \mathbb{E} [\varphi(T, X_0)] \\
&= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} [u(T - t_{n+1}, X_{n+1}) - u(T - t_n, X_n)] \\
&= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} [(-\partial_t + \tilde{\mathcal{L}}_n)u(T - t, \tilde{X}(t))] dt \\
&= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[\left(b(X_n, h(X_n)\gamma_n) - \bar{b}(\tilde{X}(t)) \right) \cdot \nabla_x u(T - t, \tilde{X}(t)) \right] dt \\
&\quad + \frac{1}{2} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[\left(\sigma\sigma^*(X_n, h(X_n)\gamma_n) - \bar{\sigma}\bar{\sigma}^*(\tilde{X}(t)) \right) : \nabla_x^2 u(T - t, \tilde{X}(t)) \right] dt.
\end{aligned}$$

The error term $\mathbb{E} \left[\left(b(X_n, h(X_n)\gamma_n) - \bar{b}(\tilde{X}(t)) \right) \cdot \nabla_x u(T - t, \tilde{X}(t)) \right]$ (with $V = b$ and $U = \nabla_x u$) and $\mathbb{E} \left[\left(\sigma\sigma^*(X_n, h(X_n)\gamma_n) - \bar{\sigma}\bar{\sigma}^*(\tilde{X}(t)) \right) : \nabla_x^2 u(T - t, \tilde{X}(t)) \right]$ (with $V = \sigma\sigma^*$ and $U = \nabla_x^2 u$) are written as

$$\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(\tilde{X}(t)) \right) U(T - t, \tilde{X}(t)).$$

Note that $\tilde{X}(t_n) = X_n$. As a consequence,

$$\mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(\tilde{X}(t)) \right) U(T - t, \tilde{X}(t)) \right] = \delta_n^1(t) + \delta_n^2(t) + \delta_n^3(t) + \delta_n^4(t),$$

where

$$\begin{aligned}
\delta_n^1(t) &= \mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(X_n) \right) U(T - t, X_n) \right] \\
\delta_n^2(t) &= \mathbb{E} \left[\left(\bar{V}(X_n) - \bar{V}(\tilde{X}(t)) \right) U(T - t, X_n) \right] \\
\delta_n^3(t) &= \mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(X_n) \right) \left(U(T - t, \tilde{X}(t)) - U(T - t, X_n) \right) \right] \\
\delta_n^4(t) &= \mathbb{E} \left[\left(\bar{V}(X_n) - \bar{V}(\tilde{X}(t)) \right) \left(U(T - t, \tilde{X}(t)) - U(T - t, X_n) \right) \right].
\end{aligned}$$

It remains to treat the four error terms $\delta_n^j(t)$, $j = 1, 2, 3, 4$.

Let us start with the most important observation: since $\bar{V}(x) = \int V(x, m) d\nu^x(m)$, the independence of the random variables X_n and γ_n yields the identity

$$\begin{aligned}
\mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(X_n) \right) U(T - t, X_n) \right] \\
= \mathbb{E} \left[\mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(X_n) \right) U(T - t, X_n) \mid X_n \right] \right] = 0.
\end{aligned}$$

The fact that this term vanishes is fundamental since it justifies the consistency of the scheme (III.3.5) with the limiting equation (III.2.7) (see also Theorem III.3.3 and its proof), and the AP property.

To treat the second term, observe that X_n and $(\tilde{X}(s), B(s))_{s > t_n}$ are independent random variables, thus conditioning with respect to X_n and applying Itô's formula (III.4.11) gives

$$\mathbb{E} \left[\left(\bar{V}(X_n) - \bar{V}(\tilde{X}(t)) \right) U(T - t, X_n) \right] = \mathbb{E} \left[\int_{t_n}^t \tilde{\mathcal{L}}_n \bar{V}(\tilde{X}(s)) ds U(T - t, X_n) \right].$$

Since U is bounded (owing to the regularity estimates from Lemma III.4.6) and since $\bar{V} \in \mathcal{C}^2(\mathbb{T}^d)$ (by assumptions on the coefficients b and σ , see Assumption III.2), one obtains

$$\left| \mathbb{E} \left[\left(\bar{V}(X_n) - \bar{V}(\tilde{X}(t)) \right) U(T-t, X_n) \right] \right| \lesssim t - t_n.$$

The treatment of the third term uses a conditioning argument, and Itô's formula (III.4.11): one has

$$\begin{aligned} & \mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(X_n) \right) \left(U(T-t, \tilde{X}(t)) - U(T-t, X_n) \right) \mid X_n, \gamma_n \right] \\ &= \mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(X_n) \right) \int_{t_n}^t \tilde{\mathcal{L}}_n U(T-t, \tilde{X}(s)) ds \mid X_n, \gamma_n \right]. \end{aligned}$$

Using the regularity properties from Lemma III.4.6, one obtains

$$\left| \mathbb{E} \left[\left(V(X_n, h(X_n)\gamma_n) - \bar{V}(X_n) \right) \left(U(T-t, \tilde{X}(t)) - U(T-t, X_n) \right) \right] \right| \lesssim t - t_n.$$

The treatment of the fourth error term is straightforward: since U and \bar{V} are Lipschitz continuous (owing to Lemma III.4.6), one has

$$\begin{aligned} \left| \mathbb{E} \left[\left(U(T-t, \tilde{X}(t)) - U(T-t, X_n) \right) \left(\bar{V}(X_n) - \bar{V}(\tilde{X}(t)) \right) \right] \right| &\lesssim \mathbb{E} \left[\left\| \tilde{X}(t) - X_n \right\|^2 \right] \\ &\lesssim t - t_n. \end{aligned}$$

The estimates above are of the type

$$|\delta_n^j(t)| \lesssim t - t_n,$$

for all $t \in [t_n, t_{n+1}]$ and $j = 1, 2, 3, 4$. Finally, one obtains

$$|\mathbb{E}[\varphi(X_N)] - \varphi(X(T))| \lesssim \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} (t - t_n) dt \lesssim \Delta t,$$

which concludes the proof of Lemma III.4.3. \square

Proof of Lemma III.4.4. The idea is to adapt the proof of Lemma III.4.2 to the discrete-time situation. Let us start with preparatory computations. A telescoping sum argument yields the equality

$$\begin{aligned} \mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)] &= \mathbb{E}[u_0(X_N^\varepsilon)] - \mathbb{E}[u_N(X_0^\varepsilon)] + u_N(X_0^\varepsilon) - u_N(X_0) \\ &= \sum_{n=0}^{N-1} \left(\mathbb{E}[u_{N-n-1}(X_{n+1}^\varepsilon)] - \mathbb{E}[u_{N-n}(X_n^\varepsilon)] \right) + u_N(X_0^\varepsilon) - u_N(X_0), \end{aligned} \tag{III.4.12}$$

where the auxiliary function u_n is defined in Lemma III.4.7. Using the definition of the scheme (III.3.4), and Markov property combined with the expression of the limiting scheme (III.3.5), one obtains

$$\begin{aligned} u_{N-n-1}(X_{n+1}^\varepsilon) &= u_{N-n-1} \left(X_n^\varepsilon + \Delta t b(X_n^\varepsilon, m_{n+1}^\varepsilon) + \sqrt{\Delta t} \sigma(X_n^\varepsilon, m_{n+1}^\varepsilon) \Gamma_n \right), \\ \mathbb{E}[u_{N-n}(X_n^\varepsilon)] &= \mathbb{E} \left[u_{N-n-1} \left(X_n^\varepsilon + \Delta t b(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n) + \sqrt{\Delta t} \sigma(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n) \Gamma_n \right) \right]. \end{aligned}$$

A second order Taylor expansion then gives

$$\begin{aligned}
& \mathbb{E} [u_{N-n-1}(X_{n+1}^\varepsilon)] - \mathbb{E} [u_{N-n}(X_n^\varepsilon)] \\
&= \Delta t \mathbb{E} [(b(X_n^\varepsilon, m_{n+1}^\varepsilon) - b(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) \cdot \nabla_x u_{N-n-1}(X_n^\varepsilon)] \\
&\quad + \frac{\Delta t}{2} \mathbb{E} [(\sigma \Gamma_n \Gamma_n^* \sigma^*(X_n^\varepsilon, m_{n+1}^\varepsilon) - \sigma \Gamma_n \Gamma_n^* \sigma^*(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) : \nabla_x^2 u_{N-n-1}(X_n^\varepsilon)] \\
&\quad + \sqrt{\Delta t} \mathbb{E} [(\sigma(X_n^\varepsilon, m_{n+1}^\varepsilon) - \sigma(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) \Gamma_n \cdot \nabla_x u_{N-n-1}(X_n^\varepsilon)] \\
&\quad + \Delta t^{3/2} \mathbb{E} [(\sigma \Gamma_n b^*(X_n^\varepsilon, m_{n+1}^\varepsilon) - \sigma \Gamma_n b^*(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) : \nabla_x^2 u_{N-n-1}(X_n^\varepsilon)] \\
&\quad + \Delta t^2 \mathbb{E} [R_n(\Delta t)].
\end{aligned}$$

Using the regularity estimates from Lemma III.4.7, one has $|R_n(\Delta t)| \lesssim 1$. Note that the terms of the orders $\Delta t^{\frac{1}{2}}$ and $\Delta t^{\frac{3}{2}}$ in the right-hand side above vanish, since the random variables Γ_n and $(X_n^\varepsilon, m_{n+1}^\varepsilon, \gamma_n)$ are independent, and $\mathbb{E} [\Gamma_n] = 0$. In addition, since $\mathbb{E} [\Gamma_n \Gamma_n^*] = I$, a conditioning argument yields

$$\begin{aligned}
& \mathbb{E} [u_{N-n-1}(X_{n+1}^\varepsilon) - u_{N-n}(X_n^\varepsilon)] \\
&= \Delta t \mathbb{E} [(b(X_n^\varepsilon, m_{n+1}^\varepsilon) - b(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) \cdot \nabla_x u_{N-n-1}(X_n^\varepsilon)] \\
&\quad + \Delta t \mathbb{E} [(\sigma \sigma^*(X_n^\varepsilon, m_{n+1}^\varepsilon) - \sigma \sigma^*(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) : \nabla_x^2 u_{N-n-1}(X_n^\varepsilon)] \quad (\text{III.4.13}) \\
&\quad + \Delta t^2 \mathbb{E} [R_n(\Delta t)].
\end{aligned}$$

Like in the proofs of Theorem III.3.3 and of Lemma III.4.3, a conditioning argument allows us to rewrite the expressions above in terms of the functions \bar{b} and $\bar{\sigma}$: one has

$$\begin{aligned}
& \mathbb{E} [(b(X_n^\varepsilon, m_{n+1}^\varepsilon) - b(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) \cdot \nabla_x u_{N-n-1}(X_n^\varepsilon)] \\
&= \mathbb{E} [(b(X_n^\varepsilon, m_{n+1}^\varepsilon) - \bar{b}(X_n^\varepsilon)) \cdot \nabla_x u_{N-n-1}(X_n^\varepsilon)],
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E} [(\sigma \sigma^*(X_n^\varepsilon, m_{n+1}^\varepsilon) - \sigma \sigma^*(X_n^\varepsilon, h(X_n^\varepsilon)\gamma_n)) : \nabla_x^2 u_{N-n-1}(X_n^\varepsilon)] \\
&= \mathbb{E} [(\sigma \sigma^*(X_n^\varepsilon, m_{n+1}^\varepsilon) - \bar{\sigma} \bar{\sigma}^*(X_n^\varepsilon)) : \nabla_x^2 u_{N-n-1}(X_n^\varepsilon)],
\end{aligned}$$

We are now in position to employ similar arguments as in the proof of Lemma III.4.2, with important modifications due to the discrete-time setting. Introduce the auxiliary parameter $\tau = \frac{\Delta t}{\varepsilon}$. Instead of studying Poisson equations associated with the infinitesimal generator \mathcal{L}_{OU}^x , one needs to consider the generator L_τ^x and the transition semigroup of a Markov chain: let

$$L_\tau^x = \frac{P_\tau^x - I}{\tau} \text{ where } P_\tau^x \varphi(m) \doteq \mathbb{E}_{\gamma \sim \mathcal{N}(0,1)} [\varphi(e^{-\tau}m + \sqrt{1 - e^{-2\tau}}h(x)\gamma)].$$

We claim that the function ψ_b^τ defined by

$$\psi_b^\tau(x, m) = -\tau \sum_{n=0}^{+\infty} ((P_\tau^x)^n b(x, m) - \bar{b}(x)).$$

is well-defined and solves the Poisson equation $L_\tau^x \psi_b^\tau(x, \cdot) = b(x, \cdot) - \bar{b}(x)$. Indeed, let $(m_n^x(m))_n$ be defined by

$$m_{n+1}^x(m) = e^{-\tau} m_n^x(m) + \sqrt{1 - e^{-2\tau}} h(x) \gamma_n, \quad m_0^x(m) = m.$$

Then, for all m and $m' \in \mathbb{R}$, and all $n \in \mathbb{N}$, one has

$$m_n^x(m) - m_n^x(m') = e^{-n\tau}(m - m').$$

Observe that since b is a Lipschitz continuous function, standard arguments give the following upper bound: for all $n \in \mathbb{N}$, $x \in \mathbb{T}^d$ and all $m, m' \in \mathbb{R}$, if $\delta_n(x, m) = (P_\tau^x)^n b(x, m) - \bar{b}(x)$, then one has

$$\begin{aligned} |\delta_n(x, m)| &= |(P_\tau^x)^n b(x, m) - \int (P_\tau^x)^n b(x, m') d\nu^x(m')| \\ &\leq \int |\mathbb{E}b(x, m_n^x(m)) - \mathbb{E}b(x, m_n^x(m'))| d\nu^x(m') \\ &\lesssim e^{-n\tau}(1 + |m|). \end{aligned}$$

Similarly, since the derivatives of m_n^x with respect to x do not depend on m , one can check that the inequality above holds for $D_x \delta$ and $D_x^2 \delta$ then concludes the proof of the claim. Since $\tau \sum_{n=0}^{\infty} e^{-n\tau} = \frac{\tau}{1-e^{-\tau}} \leq \max(\tau, 1)$, for all $\tau \in (0, \infty)$, one obtains inequalities of the type

$$\frac{\|\psi(x, m)\|}{1 + |m|} \lesssim \max(\tau, 1), \quad (\text{III.4.14})$$

for ψ_b^τ , and its derivatives $D_x \psi_b^\tau$ and $D_x^2 \psi_b^\tau$.

Similarly, define for all $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$,

$$\psi_\sigma^\tau(x, m) = -\tau \sum_{n=0}^{+\infty} ((P_\tau^x)^n \sigma \sigma^*(x, m) - \bar{\sigma} \bar{\sigma}^*(x)).$$

Then ψ_σ^τ is well-defined and solves the Poisson equation $L_\tau^x \psi_\sigma(x, \cdot) = \sigma \sigma^*(x, \cdot) - \bar{\sigma}(x) \bar{\sigma}^*(x)$, by definition of $\bar{\sigma}(x)$, see Assumption III.3. In addition, ψ_σ^τ and its derivatives satisfy upper bound of the type (III.4.14).

Like in the proof of Lemma III.4.2, introduce the auxiliary function defined by

$$\Phi_n(x, m) = \psi_b^\tau(x, m) \cdot \nabla_x u_{N-n-1}(x) + \psi_\sigma^\tau(x, m) : \nabla_x^2 u_{N-n-1}(x),$$

for all $n \in \llbracket 0, N-1 \rrbracket$, $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$, where u_n is defined in Lemma III.4.7. Combining the decomposition of the error (III.4.12) and the identity (III.4.13), one then obtains the following new expression for the error:

$$\begin{aligned} \mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)] &= \Delta t^2 \sum_{n=0}^{N-1} \mathbb{E}[R_n(\Delta t)] + \Delta t \sum_{n=0}^{N-1} L_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon) \\ &\quad + u_N(x_0^\varepsilon) - u_N(x_0). \quad (\text{III.4.15}) \end{aligned}$$

On the one hand, a telescoping sum argument yields the following expression:

$$\begin{aligned}
\mathbb{E} [P_\tau^{X_N^\varepsilon} \Phi_N(X_N^\varepsilon, m_N^\varepsilon) - P_\tau^{X_0^\varepsilon} \Phi_0(X_0^\varepsilon, m_0^\varepsilon)] \\
&= \sum_{n=0}^{N-1} \mathbb{E} [P_\tau^{X_{n+1}^\varepsilon} \Phi_{n+1}(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_n^\varepsilon)] \\
&= \sum_{n=0}^{N-1} \mathbb{E} [P_\tau^{X_{n+1}^\varepsilon} \Phi_{n+1}(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_{n+1}^\varepsilon} \Phi_n(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon)] \\
&\quad + \sum_{n=0}^{N-1} \mathbb{E} [P_\tau^{X_{n+1}^\varepsilon} \Phi_n(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_n^\varepsilon)] \\
&\quad + \sum_{n=0}^{N-1} \mathbb{E} [P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_n^\varepsilon)].
\end{aligned}$$

Note that using Markov property, the first sum on the right-hand side above can be written as $\sum_{n=0}^{N-1} \mathbb{E} [\Phi_{n+1}(X_{n+1}^\varepsilon, m_{n+2}^\varepsilon) - \Phi_n(X_{n+1}^\varepsilon, m_{n+2}^\varepsilon)]$.

On the other hand, by definition of the operator L_τ^x with the parameter $\tau = \frac{\Delta t}{\varepsilon}$, one obtains

$$\begin{aligned}
\Delta t \mathbb{E} [L_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon)] &= \varepsilon \mathbb{E} [P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon) - \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon)] \\
&= \varepsilon \mathbb{E} [P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_n^\varepsilon)].
\end{aligned}$$

Finally, combining the two identities above, one obtains the following expression for the error:

$$\begin{aligned}
\mathbb{E} [\varphi(X_N^\varepsilon)] - \mathbb{E} [\varphi(X_N)] &= \Delta t^2 \sum_{n=0}^{N-1} \mathbb{E} [R_n] + \varepsilon (\mathbb{E} [\Phi_N(X_N^\varepsilon, m_{N+1}^\varepsilon) - \Phi_0(X_0^\varepsilon, m_1^\varepsilon)]) \\
&\quad - \varepsilon \sum_{n=0}^{N-1} \mathbb{E} [\Phi_{n+1}(X_{n+1}^\varepsilon, m_{n+2}^\varepsilon) - \Phi_n(X_{n+1}^\varepsilon, m_{n+2}^\varepsilon)] \\
&\quad - \varepsilon \sum_{n=0}^{N-1} \mathbb{E} [P_\tau^{X_{n+1}^\varepsilon} \Phi_n(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon)] \\
&\quad + u_N(x_0^\varepsilon) - u_N(x_0).
\end{aligned}$$

It then remains to use auxiliary upper bounds to deduce the result, in particular using (III.4.14). Note that $\varepsilon \max(\tau, 1) = \max(\Delta t, \varepsilon)$.

- as explained above, $\mathbb{E} [|R_n(\Delta t)|] \lesssim 1$, thus the first term satisfies

$$\mathbb{E} \left[\left| \Delta t^2 \sum_{n=0}^{N-1} R_n \right| \right] \lesssim \Delta t.$$

- Using the upper bound (III.4.7) from Lemma III.4.7 and (III.4.14) with $\psi = \psi_b^\tau$ and $\psi = \psi_\sigma^\tau$, the second term satisfies

$$\varepsilon (\mathbb{E} [\Phi_N(X_N^\varepsilon, m_{N+1}^\varepsilon) - \Phi_0(X_0^\varepsilon, m_1^\varepsilon)]) \lesssim \max(\Delta t, \varepsilon).$$

- Observe that one has for all $x \in \mathbb{T}^d$ and $x \in \mathbb{R}$,

$$\begin{aligned}\Phi_{n+1}(x, m) - \Phi_n(x, m) &= \psi_b^\tau(x, m) \cdot (\nabla_x u_{N-n}(x) - \nabla_x u_{N-n-1}(x)) \\ &\quad + \psi_\sigma^\tau(x, m) : (\nabla_x^2 u_{N-n}(x) - \nabla_x^2 u_{N-n-1}(x)).\end{aligned}$$

Using the upper bound (III.4.8) from Lemma III.4.7 and (III.4.14), one has

$$|\Phi_{n+1}(x, m) - \Phi_n(x, m)| \lesssim \Delta t \max(\tau, 1)(1 + |m|).$$

As a consequence, owing to (III.3.9) and Assumption III.1, the third term satisfies

$$\begin{aligned}\varepsilon \sum_{n=0}^{N-1} \mathbb{E} [\Phi_{n+1}(X_{n+1}^\varepsilon, m_{n+2}^\varepsilon) - \Phi_n(X_{n+1}^\varepsilon, m_{n+2}^\varepsilon)] &\lesssim \varepsilon \sum_{n=0}^{N-1} \Delta t \max(\tau, 1) \\ &\lesssim \max(\Delta t, \varepsilon).\end{aligned}$$

- Note that Φ_n and its derivatives satisfy the upper bound (III.4.14), owing to (III.4.7) from Lemma III.4.7. Let

$$f_n^\tau(x, m) = P_\tau^x \Phi_n(x, m) = \mathbb{E}_{\gamma \sim \mathcal{N}(0,1)} [\Phi_n(x, e^{-\tau} m + \sqrt{1 - e^{-2\tau}} h(x) \gamma)].$$

It is straightforward to check that f_n^τ is twice differentiable. In addition, f_n^τ and its derivatives satisfy (III.4.14). Using a second order Taylor expansion, one obtains

$$\begin{aligned}\left| P_\tau^{X_{n+1}^\varepsilon} \Phi_n(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon) \right| &= |f_n^\tau(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) - f_n^\tau(X_n^\varepsilon, m_n^\varepsilon)| \\ &\lesssim \Delta t \max(\tau, 1)(1 + |m|).\end{aligned}$$

Finally, using (III.3.9) and Assumption III.1, one obtains

$$\varepsilon \sum_{n=0}^{N-1} \mathbb{E} \left[P_\tau^{X_{n+1}^\varepsilon} \Phi_n(X_{n+1}^\varepsilon, m_{n+1}^\varepsilon) - P_\tau^{X_n^\varepsilon} \Phi_n(X_n^\varepsilon, m_{n+1}^\varepsilon) \right] \lesssim \max(\Delta t, \varepsilon).$$

- Owing to the assumption $d_{\mathbb{T}^d}(x_0^\varepsilon, x_0) \lesssim \varepsilon$ and to Lemma III.4.7, we have

$$|u_N(x_0^\varepsilon) - u_N(x_0)| \lesssim \varepsilon.$$

Gathering the estimates then concludes the proof of Lemma III.4.4. \square

III.4.3 Estimates on solutions of Kolmogorov equations

We refer to [Cer01, Theorem 1.3.6] for the proof of Lemma III.4.6. It thus remains to provide the arguments for the proofs of Lemmas III.4.5 and III.4.7. The strategy is standard in the literature, see for instance [Cer01].

Proof of Lemma III.4.5. Owing to [Cer01, Proposition 1.3.5], for all $\underline{k} = (k_x, k_m) \in \mathbb{R}^d \times \mathbb{R}$, one has

$$D_{x,m} u^\varepsilon(t, x, m) \cdot \underline{k} = \mathbb{E}_{X^\varepsilon(0)=x, m^\varepsilon(0)=m} [D\varphi(X^\varepsilon(t)) \cdot \eta_{x,m}^{\varepsilon, \underline{k}}(t)],$$

where the process $\eta^{\varepsilon, \underline{k}} = (\eta_x^{\varepsilon, \underline{k}}, \eta_m^{\varepsilon, \underline{k}})$ is the solution of the first variation equation associated with (III.2.1): for all $t \in \mathbb{R}^+$,

$$\begin{cases} d\eta_{x,t}^{\varepsilon, \underline{k}} = Db(X_t^\varepsilon, m_t^\varepsilon) \cdot \eta_t^{\varepsilon, \underline{k}} dt + D\sigma(X_t^\varepsilon, m_t^\varepsilon) \cdot \eta_t^{\varepsilon, \underline{k}} dB_t \\ d\eta_{m,t}^{\varepsilon, \underline{k}} = -\frac{\eta_{m,t}^{\varepsilon, \underline{k}}}{\varepsilon} dt + \frac{\sqrt{2}Dh(X_t^\varepsilon) \cdot \eta_{x,t}^{\varepsilon, \underline{k}}}{\sqrt{\varepsilon}} d\beta_t, \end{cases}$$

with initial conditions $\eta_x^{\varepsilon, \underline{k}}(0) = k_x$ and $\eta_m^{\varepsilon, \underline{k}}(0) = k_m$.

On the one hand, the component $\eta_m^{\varepsilon, \underline{k}}$ satisfies the following equality,

$$\eta_m^{\varepsilon, \underline{k}}(t) = e^{-\frac{t}{\varepsilon}} k_m + \frac{1}{\sqrt{\varepsilon}} \int_0^t e^{-\frac{t-s}{\varepsilon}} \sqrt{2} Dh(X^\varepsilon(s)) \cdot \eta_x^{\varepsilon, \underline{k}}(s) dB_s,$$

and by means of Itô's isometry formula, one has

$$\mathbb{E} \left[\left\| \eta_m^{\varepsilon, \underline{k}}(t) \right\|^2 \right] \leq k_m^2 + \frac{2 \|h\|_{C^1(\mathbb{T}^d)}}{\varepsilon} \int_0^t e^{-\frac{2(t-s)}{\varepsilon}} \mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(s) \right\|^2 \right] ds.$$

On the other hand, the functions b and σ are globally Lipschitz continuous (see Assumption III.2), and one has $\left\| \eta^{\varepsilon, \underline{k}} \right\|^2 = \left\| \eta_x^{\varepsilon, \underline{k}} \right\|^2 + \left\| \eta_m^{\varepsilon, \underline{k}} \right\|^2$; using Itô's isometry formula, and the Minkowski and Young inequalities, one obtains

$$\mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(t) \right\|^2 \right] \lesssim \|k_x\|^2 + \int_0^t \left(\mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(s) \right\|^2 \right] + \mathbb{E} \left[\left\| \eta_m^{\varepsilon, \underline{k}}(s) \right\|^2 \right] \right) ds.$$

Combining the two estimates above then yields

$$\begin{aligned} \sup_{s \in [0, t]} \mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(s) \right\|^2 \right] &\lesssim \|\underline{k}\|^2 + \int_0^t \left(\mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(s) \right\|^2 \right] + \frac{1}{\varepsilon} \int_0^s e^{-2\frac{s-r}{\varepsilon}} \mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(r) \right\|^2 \right] dr \right) ds \\ &\lesssim \|\underline{k}\|^2 + \int_0^t \sup_{r \in [0, s]} \mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(r) \right\|^2 \right] ds, \end{aligned}$$

where the inequality $\frac{1}{\varepsilon} \int_0^s e^{-2\frac{s-r}{\varepsilon}} dr \leq \frac{1}{2}$ has been used. Applying Gronwall's lemma, then inserting the result in the estimate above, one obtains the upper bounds

$$\sup_{t \in [0, T]} \mathbb{E} \left[\left\| \eta_x^{\varepsilon, \underline{k}}(t) \right\|^2 \right] \lesssim \|\underline{k}\|^2, \quad \sup_{t \in [0, T]} \mathbb{E} \left[\left\| \eta_m^{\varepsilon, \underline{k}}(t) \right\|^2 \right] \lesssim \|\underline{k}\|^2.$$

Since φ is Lipschitz continuous, using the expression for $D_{x,m} u^\varepsilon(t, x, m)$ stated above, one finally obtains (III.4.5) for the first-order derivative: for all $\underline{k} \in \mathbb{R}^{d+1}$, one has

$$\sup_{(t,x,m) \in [0,T] \times \mathbb{T}^d \times \mathbb{R}} |D_{x,m} u^\varepsilon(t, x, m) \cdot \underline{k}| \lesssim \|\underline{k}\|.$$

The treatment of higher-order derivatives requires the expression of $\lambda^{\varepsilon, \underline{k}_1, \underline{k}_2} = (\lambda_x^{\varepsilon, \underline{k}_1, \underline{k}_2}, \lambda_m^{\varepsilon, \underline{k}_1, \underline{k}_2})$ solution of the second variation equation and $\mu^{\varepsilon, \underline{k}_1, \underline{k}_2, \underline{k}_3} = (\mu_x^{\varepsilon, \underline{k}_1, \underline{k}_2, \underline{k}_3}, \mu_m^{\varepsilon, \underline{k}_1, \underline{k}_2, \underline{k}_3})$ solution of the third variation equation associated with (III.2.1): if, for $j \in \{1, 2, 3\}$ and $t \in \mathbb{R}^+$, we write $D^j b_t^\varepsilon = D^j b(X_t^\varepsilon, m_t^\varepsilon)$, $D^j \sigma_t^\varepsilon = D^j \sigma(X_t^\varepsilon, m_t^\varepsilon)$ and $D^j h_t^\varepsilon = D^j h(X_t^\varepsilon)$, then for all $\underline{k}_1, \underline{k}_2, \underline{k}_3 \in \mathbb{R}^d \times \mathbb{R}$, the variation equations are given by

$$\begin{cases} d\lambda_{x,t}^{\varepsilon, \underline{k}_1, \underline{k}_2} = D^2 b_t^\varepsilon \cdot (\eta_t^{\varepsilon, \underline{k}_1}, \eta_t^{\varepsilon, \underline{k}_2}) dt + D^2 \sigma_t^\varepsilon \cdot (\eta_t^{\varepsilon, \underline{k}_1}, \eta_t^{\varepsilon, \underline{k}_2}) dB_t \\ \quad + Db_t^\varepsilon \cdot \lambda_t^{\varepsilon, \underline{k}_1, \underline{k}_2} dt + D\sigma_t^\varepsilon \cdot \lambda_t^{\varepsilon, \underline{k}_1, \underline{k}_2} dB_t \\ d\lambda_{m,t}^{\varepsilon, \underline{k}_1, \underline{k}_2} = -\frac{\lambda_{m,t}^{\varepsilon, \underline{k}_1, \underline{k}_2}}{\varepsilon} dt + \frac{\sqrt{2} D^2 h_t^\varepsilon \cdot (\eta_{x,t}^{\varepsilon, \underline{k}_1}, \eta_{x,t}^{\varepsilon, \underline{k}_2})}{\sqrt{\varepsilon}} d\beta_t + \frac{\sqrt{2} Dh_t^\varepsilon \cdot \lambda_{x,t}^{\varepsilon, \underline{k}_1, \underline{k}_2}}{\sqrt{\varepsilon}} d\beta_t, \end{cases}$$

$$\left\{ \begin{aligned} d\mu_{x,t}^{\varepsilon,k_1,k_2,k_3} &= D^3 b_t^\varepsilon \cdot (\eta_t^{\varepsilon,k_1}, \eta_t^{\varepsilon,k_2}, \eta_t^{\varepsilon,k_3}) dt + D^3 \sigma_t^\varepsilon \cdot (\eta_t^{\varepsilon,k_1}, \eta_t^{\varepsilon,k_2}, \eta_t^{\varepsilon,k_3}) dB_t \\ &\quad + D^2 b_t^\varepsilon \cdot (\lambda_t^{\varepsilon,k_1,k_3}, \eta_t^{\varepsilon,k_2}) dt + D^2 \sigma_t^\varepsilon \cdot (\lambda_t^{\varepsilon,k_1,k_3}, \eta_t^{\varepsilon,k_2}) dB_t \\ &\quad + D^2 b_t^\varepsilon \cdot (\eta_t^{\varepsilon,k_1}, \lambda_t^{\varepsilon,k_2,k_3}) dt + D^2 \sigma_t^\varepsilon \cdot (\eta_t^{\varepsilon,k_1}, \lambda_t^{\varepsilon,k_2,k_3}) dB_t \\ &\quad + D^2 b_t^\varepsilon \cdot (\lambda_t^{\varepsilon,k_1,k_2}, \eta_t^{\varepsilon,k_3}) dt + D^2 \sigma_t^\varepsilon \cdot (\lambda_t^{\varepsilon,k_1,k_2}, \eta_t^{\varepsilon,k_3}) dB_t \\ &\quad + D b_t^\varepsilon \cdot \mu_t^{\varepsilon,k_1,k_2,k_3} dt + D \sigma_t^\varepsilon \cdot \mu_t^{\varepsilon,k_1,k_2,k_3} dB_t \\ d\mu_{m,t}^{\varepsilon,k_1,k_2,k_3} &= -\frac{\mu_{m,t}^{\varepsilon,k_1,k_2,k_3}}{\varepsilon} dt + \frac{\sqrt{2} D^3 h_t^\varepsilon \cdot (\eta_{x,t}^{\varepsilon,k_1}, \eta_{x,t}^{\varepsilon,k_2}, \eta_{x,t}^{\varepsilon,k_3})}{\sqrt{\varepsilon}} d\beta_t + \frac{\sqrt{2} D^2 h_t^\varepsilon \cdot (\lambda_{x,t}^{\varepsilon,k_1,k_3}, \eta_{x,t}^{\varepsilon,k_2})}{\sqrt{\varepsilon}} d\beta_t \\ &\quad + \frac{\sqrt{2} D^2 h_t^\varepsilon \cdot (\eta_{x,t}^{\varepsilon,k_1}, \lambda_{x,t}^{\varepsilon,k_2,k_3})}{\sqrt{\varepsilon}} d\beta_t + \frac{\sqrt{2} D^2 h_t^\varepsilon \cdot (\lambda_{x,t}^{\varepsilon,k_1,k_2}, \eta_{x,t}^{\varepsilon,k_3})}{\sqrt{\varepsilon}} d\beta_t \\ &\quad + \frac{\sqrt{2} D h_t^\varepsilon \cdot \mu_{x,t}^{\varepsilon,k_1,k_2,k_3}}{\sqrt{\varepsilon}} d\beta_t, \end{aligned} \right.$$

with initial conditions $\lambda^{\varepsilon,k_1,k_2} = \kappa^{\varepsilon,k_1,k_2,k_3} = 0$. The estimate for second and third order derivatives then follow from similar arguments than for the first order derivative.

This concludes the proof of Lemma III.4.5. \square

Proof of Lemma III.4.7. For all $k \in \mathbb{R}^d$, one has

$$Du_n(x) \cdot k = \mathbb{E}_x[D\varphi(X_n) \cdot \eta_n^k],$$

where $\eta_0^h = h$, and for all $n \in \llbracket 0, N-1 \rrbracket$, one has

$$\begin{aligned} \eta_{n+1}^k &= \eta_n^k + \Delta t D_x b(X_n, h(X_n) \gamma_n) \cdot \eta_n^k + \Delta t \partial_m b(X_n, h(X_n) \gamma_n) Dh(X_n) \cdot \eta_n^k \gamma_n \\ &\quad + \sqrt{\Delta t} D_x \sigma(X_n, h(X_n) \gamma_n) \cdot \eta_n^k \Gamma_n + \sqrt{\Delta t} \partial_m \sigma(X_n, h(X_n) \gamma_n) Dh(X_n) \cdot \eta_n^k \Gamma_n \gamma_n. \end{aligned}$$

The functions b , h and σ are Lipschitz continuous (see Assumption III.2). Since γ_n and Γ are independent centered Gaussian random variables, it is straightforward to obtain the upper bound

$$\mathbb{E} \left[\|\eta_{n+1}^k\|^2 \right] \lesssim (1 + \Delta t) \mathbb{E} \left[\|\eta_n^k\|^2 \right].$$

A straightforward recursion argument then gives, for all $n \in \llbracket 0, N \rrbracket$,

$$\mathbb{E} \left[\|\eta_n^k\|^2 \right] \lesssim \|k\|^2$$

and one obtains (III.4.7) for the first-order derivative: for all $n \in \llbracket 0, N \rrbracket$,

$$|Du_n(x) \cdot k| \lesssim \|k\|.$$

The treatment of the second order derivative is similar, using that the solution λ^{k_1,k_2} of the second variation equation associated with the limiting scheme satisfies

$$\begin{aligned} \lambda_{n+1}^{k_1,k_2} &= \lambda_n^{k_1,k_2} + \Delta t D_x^2 b_n \cdot (\eta_n^{k_1}, \eta_n^{k_2}) + \Delta t \partial_m D_x b_n \cdot \eta_n^{k_1} Dh_n \cdot \eta_n^{k_2} \gamma_n + \Delta t D_x b_n \cdot \lambda_n^{k_1,k_2} \\ &\quad + \Delta t D_x \partial_m b_n \cdot \eta_n^{k_2} Dh_n \cdot \eta_n^{k_1} \gamma_n + \Delta t \partial_m^2 b_n Dh_n \cdot \eta_n^{k_1} Dh_n \cdot \eta_n^{k_2} \gamma_n^2 \\ &\quad + \Delta t \partial_m b_n D^2 h_n \cdot (\eta_n^{k_1}, \eta_n^{k_2}) \gamma_n + \Delta t \partial_m b_n Dh_n \cdot \lambda_n^{k_1,k_2} \gamma_n + \sqrt{\Delta t} D_x^2 \sigma_n \cdot (\eta_n^{k_1}, \eta_n^{k_2}) \Gamma_n \\ &\quad + \sqrt{\Delta t} \partial_m D_x \sigma_n \cdot \eta_n^{k_1} Dh_n \cdot \eta_n^{k_2} \gamma_n \Gamma_n + \sqrt{\Delta t} D_x \sigma_n \cdot \lambda_n^{k_1,k_2} \Gamma_n \\ &\quad + \sqrt{\Delta t} D_x \partial_m \sigma_n \cdot \eta_n^{k_2} Dh_n \cdot \eta_n^{k_1} \Gamma_n \gamma_n + \sqrt{\Delta t} \partial_m^2 \sigma_n Dh_n \cdot \eta_n^{k_1} Dh_n \cdot \eta_n^{k_2} \Gamma_n \gamma_n^2 \\ &\quad + \sqrt{\Delta t} \partial_m \sigma_n D^2 h_n \cdot (\eta_n^{k_1}, \eta_n^{k_2}) \Gamma_n \gamma_n + \sqrt{\Delta t} \partial_m \sigma_n Dh_n \cdot \lambda_n^{k_1,k_2} \Gamma_n \gamma_n, \end{aligned}$$

where we use the short notation $D^j b_n = D^j b(X_n, h(X_n)\gamma_n)$, $D^j \sigma_n = D^j \sigma(X_n, h(X_n)\gamma_n)$ and $D^j h_n = D^j h(X_n)$.

It remains to prove (III.4.8). On the one hand, by definition (III.3.5), a second order Taylor expansion yields, for all $x \in \mathbb{T}^d$ and $n \in \llbracket 0, N-1 \rrbracket$,

$$|u_{n+1}(x) - u_n(x)| = |\mathbb{E}_x[\varphi(X_{n+1}) - \varphi(X_n)]| \lesssim \Delta t,$$

On the other hand, one has

$$\begin{aligned} Du_{n+1}(x) \cdot k - Du_n(x) \cdot k &= \mathbb{E}_x[D\varphi(X_{n+1}) \cdot \eta_{n+1}^k - D\varphi(X_n) \cdot \eta_n^k] \\ &= \mathbb{E}[(D\varphi(X_{n+1}) - D\varphi(X_n)) \cdot (\eta_{n+1}^k - \eta_n^k)] \\ &\quad + \mathbb{E}[D\varphi(X_n) \cdot (\eta_{n+1}^k - \eta_n^k)] + \mathbb{E}[(D\varphi(X_{n+1}) - D\varphi(X_n)) \cdot \eta_n^k]. \end{aligned}$$

It is straightforward to check that one has the inequalities $\mathbb{E}[\|X_{n+1} - X_n\|^2] \lesssim \Delta t$ and $\mathbb{E}[\|\eta_{n+1}^k - \eta_n^k\|^2] \lesssim \Delta t \|k\|^2$. Since φ is Lipschitz continuous, using Cauchy-Schwarz inequality gives

$$|\mathbb{E}[(D\varphi(X_{n+1}(x)) - D\varphi(X_n(x))) \cdot (DX_{n+1}(x) \cdot k - DX_n(x) \cdot k)]| \lesssim \Delta t \|k\|.$$

Using a conditional expectation argument, since γ_n, Γ_n, X_n are independent random variables, one has

$$|\mathbb{E}[D\varphi(X_n) \cdot (\eta_{n+1}^k - \eta_n^k)]| = \Delta t |\mathbb{E}[D\varphi(X_n) \cdot D_x b(X_n, h(X_n)\gamma_n) \cdot \eta_n^k]| \lesssim \Delta t \|k\|.$$

Finally, using a second-order Taylor expansion and conditioning arguments, one obtains

$$|\mathbb{E}[(D\varphi(X_{n+1}) - D\varphi(X_n)) \cdot \eta_n^k]| \lesssim \Delta t \|k\|.$$

As a consequence, one obtains (III.4.8) for $j = 1$. Similar arguments, coupled with the second variation equation satisfied by $\lambda_n^{k_1, k_2}$ yields (III.4.8) for $j = 2$.

This concludes the proof of Lemma III.4.7. \square

III.4.4 On the diffusion approximation regime

In general, we do not know how to prove of uniform accuracy in the diffusion approximation. To prove uniform accuracy, one could try proving Lemmas III.4.1 to III.4.4 for the AP scheme (III.3.12). Lemmas III.4.2 and III.4.3 can be proved in the diffusion approximation regime like in the averaging regime. We also expect Lemma III.4.1 to hold with a different upper bound

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))]| \leq C(T, \varphi) \frac{\Delta t^a}{\varepsilon^b},$$

for some real numbers $a, b \in (0, \infty)$. However, the proof of Lemma III.4.4 relies on the averaging lemma and we did not manage to adapt this argument in the diffusion approximation regime.

This section is devoted to proving uniform accuracy in two particular cases of the diffusion approximation regime (see Propositions III.4.8 and III.4.9 below). We also prove an estimate on the diffusion approximation error at the discrete time level in these case, similar to Lemma III.4.4. Note that there are several differences with the averaging regime: first, in the two particular cases, we prove strong error estimates (instead of weak error, in

distribution). These error estimates implies in particular weak error estimates if the test functions are Lipschitz-continuous (as in Theorem III.3.4 in the averaging regime). Second, the diffusion approximation error at the discrete time level (when $\varepsilon \rightarrow 0$) is uniform in Δt , while the estimate on the averaging error at the discrete time level given by Lemma III.4.4 is $\max(\varepsilon, \Delta t)$ and therefore is not uniform in Δt .

Proposition III.4.8. *Consider the SDE (III.2.12) with $d = 1$, $b = 0$, $\sigma = 1$, $f = 1$, $g = 0$ and $h = 1$, namely*

$$\begin{cases} dX_t^\varepsilon = \frac{m_t^\varepsilon}{\varepsilon} dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon^2} dt + \frac{1}{\varepsilon} d\beta_t. \end{cases}$$

In this case, the AP scheme (III.3.12) given by Theorem III.3.5 with $\theta = 1$ is written as follows:

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + \frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon, \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2} m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon} \gamma_n, \end{cases}$$

and its limiting scheme is given by

$$X_{n+1} = X_n + \sqrt{\Delta t} \gamma_n,$$

where $\sqrt{\Delta t} \gamma_n \doteq \beta(t_{n+1}) - \beta(t_n)$. Assume moreover that there exists $C_0 \in (0, \infty)$ such that, for $\varepsilon \in (0, 1]$, $d_{\mathbb{T}}(x_0^\varepsilon, x_0) \leq C_0 \varepsilon$, where $d_{\mathbb{T}}$ is the distance on \mathbb{T} .

Then, there exists $C(T) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ and $\varepsilon \in (0, 1]$ one has

$$\mathbb{E}[d_{\mathbb{T}}(X_N^\varepsilon, X^\varepsilon(T))] \leq C(T) \sqrt{\Delta t}, \quad (\text{III.4.16})$$

and

$$\mathbb{E}[d_{\mathbb{T}}(X_N^\varepsilon, X_N)] \leq C(T) \varepsilon. \quad (\text{III.4.17})$$

Proposition III.4.9. *Consider the SDE (III.2.20) with $\sigma(x) = x$ and $x \in \mathbb{R}$ as in Remark III.3.4, namely*

$$\begin{cases} dX_t^\varepsilon = \frac{X_t^\varepsilon m_t^\varepsilon}{\varepsilon} dt, \\ dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon^2} dt + \frac{1}{\varepsilon} d\beta_t. \end{cases}$$

Let us recall the AP scheme (III.3.23)

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon \exp\left(\frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon\right) \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2} m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon} \gamma_n, \end{cases}$$

and its limiting scheme (III.3.24)

$$X_{n+1} = X_n \exp(\sqrt{\Delta t} \gamma_n),$$

where $\sqrt{\Delta t} \gamma_n \doteq \beta(t_{n+1}) - \beta(t_n)$. Assume moreover that there exists $C_0 \in (0, \infty)$ such that, for $\varepsilon \in (0, 1]$, $|x_0^\varepsilon - x_0| \leq C_0 \varepsilon$.

Then, there exists $C(T) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ and $\varepsilon \in (0, 1]$ one has

$$\mathbb{E} [|X_N^\varepsilon - X^\varepsilon(T)|] \leq C(T)\sqrt{\Delta t}, \quad (\text{III.4.18})$$

and

$$\mathbb{E} [|X_N^\varepsilon - X_N|] \leq C(T)\varepsilon. \quad (\text{III.4.19})$$

Although we do not know of any proof to generalize these estimates to the general diffusion approximation case, we investigate this generalization from a numerical point of view in Section III.7.

In this section, for all $\lambda, \mu \in \mathbb{R}^+$, $\lambda \lesssim \mu$ means that there exists $C(T) \in (0, +\infty)$, independent of Δt , ε and n , such that $\lambda \leq C(T)\mu$.

In Propositions III.4.8 and III.4.9, the fast component is discretized with an implicit Euler scheme. Let us start by studying this standard scheme.

Lemma III.4.10. *Consider the Ornstein-Uhlenbeck process*

$$dm_t^\varepsilon = -\frac{m_t^\varepsilon}{\varepsilon^2}dt + \frac{1}{\varepsilon}d\beta_t,$$

and its discretization using the implicit Euler Scheme

$$m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2}m_{n+1}^\varepsilon + \frac{\sqrt{\Delta t}}{\varepsilon}\gamma_n,$$

where $\sqrt{\Delta t}\gamma_n \doteq \beta(t_{n+1}) - \beta(t_n)$ and $m^\varepsilon(0) = m_0^\varepsilon$ satisfies Assumption III.1.

Then, there exists $C(T) \in (0, \infty)$ such that for all $\Delta t \in (0, \Delta t_0]$ and $\varepsilon \in (0, 1]$, one has

$$\mathbb{E} [|m_N^\varepsilon|^2] + \mathbb{E} [|m^\varepsilon(T)|^2] \leq C(T), \quad (\text{III.4.20})$$

$$\mathbb{E} [|m_N^\varepsilon - m^\varepsilon(T)|^2] \leq C(T)\frac{\Delta t}{\varepsilon^2}, \quad (\text{III.4.21})$$

with $N = T/\Delta t$.

The proof of Lemma III.4.10 use the following lemma.

Lemma III.4.11. *For all $n \in \mathbb{N}$ and $\tau \in \mathbb{R}_*^+$, we have*

$$\left| \frac{1}{(1+\tau)^n} - e^{-n\tau} \right| \leq \frac{1}{n}.$$

Proof of Lemma III.4.11. Consider the function $f(\tau) = \frac{1}{(1+\tau)^n} - e^{-n\tau}$ on \mathbb{R}_*^+ . Since $e^\tau > 1 + \tau$ for $\tau \in \mathbb{R}_*^+$, f is nonnegative. Moreover, $\lim_{0+} f = 0$ and $\lim_{+\infty} f = 0$, thus there exists $\tau_{\max} \in \mathbb{R}_*^+$ such that $f(\tau_{\max}) = \max_{\mathbb{R}_*^+} f$. The function f is differentiable, thus we have $f'(\tau_{\max}) = 0$ with

$$f'(\tau) = -n\frac{1}{(1+\tau)^{n+1}} + ne^{-n\tau}.$$

As a consequence, $e^{-n\tau_{\max}} = \frac{1}{(1+\tau_{\max})^{n+1}}$ and

$$\begin{aligned} f(\tau_{\max}) &= \frac{1}{(1+\tau_{\max})^n} - \frac{1}{(1+\tau_{\max})^{n+1}} = \frac{\tau_{\max}}{(1+\tau_{\max})^{n+1}} \\ &\leq \frac{\tau_{\max}}{1+(n+1)\tau_{\max}} \leq \frac{1}{n+1}, \end{aligned}$$

where we used that $(1+\tau)^{n+1} \geq (n+1)\tau$ for $\tau \in \mathbb{R}_*^+$. This concludes the proof of Lemma III.4.11. \square

Proof of Lemma III.4.10. Estimate (III.4.20) follows from elementary arguments. Set $\tau = \frac{\Delta t}{\varepsilon^2}$, so that

$$m_{n+1}^\varepsilon = \frac{m_n^\varepsilon}{1 + \tau} + \frac{\beta(t_{n+1}) - \beta(t_n)}{\varepsilon(1 + \tau)}.$$

It follows that

$$m_N^\varepsilon = \frac{m_0^\varepsilon}{(1 + \tau)^N} + \frac{1}{\varepsilon} \sum_{n=0}^{N-1} \frac{\beta(t_{n+1}) - \beta(t_n)}{(1 + \tau)^{N-n}}. \quad (\text{III.4.22})$$

Since the random variables $\beta(t_{n+1}) - \beta(t_n)$ are independent with distribution $\mathcal{N}(0, \Delta t)$, we get

$$\begin{aligned} \mathbb{E}[|m_N^\varepsilon|^2] &= |m_0^\varepsilon|^2 \frac{1}{(1 + \tau)^{2N}} + \frac{1}{\varepsilon^2} \sum_{n=0}^{N-1} \frac{\Delta t}{(1 + \tau)^{2N-2n}} \leq |m_0^\varepsilon|^2 + \tau \sum_{n=1}^{\infty} \frac{1}{(1 + \tau)^n} \\ &\lesssim |m_0^\varepsilon|^2 + 1 \lesssim 1, \end{aligned}$$

owing to Assumption III.1. We also have

$$m^\varepsilon(T) = m_0^\varepsilon e^{-T/\varepsilon^2} + \frac{1}{\varepsilon} \int_0^T e^{-\frac{T-t}{\varepsilon^2}} d\beta_t. \quad (\text{III.4.23})$$

Owing to Itô's isometry formula, we get

$$\begin{aligned} \mathbb{E}[|m^\varepsilon(T)|^2] &= |m_0^\varepsilon|^2 e^{-2T/\varepsilon^2} + \frac{1}{\varepsilon^2} \int_0^T e^{-2\frac{T-t}{\varepsilon^2}} dt = |m_0^\varepsilon|^2 e^{-2T/\varepsilon^2} + \frac{1 - e^{-2T/\varepsilon^2}}{2} \\ &\lesssim |m_0^\varepsilon|^2 + \frac{1}{2} \lesssim 1, \end{aligned}$$

owing to Assumption III.1. This concludes the proof of (III.4.20). Let us now prove (III.4.21). Owing to (III.4.22) and (III.4.23), we have

$$m_N^\varepsilon - m^\varepsilon(T) = m_0^\varepsilon \left(\frac{1}{(1 + \tau)^N} - e^{-T/\varepsilon^2} \right) + \frac{1}{\varepsilon} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \left(\frac{1}{(1 + \tau)^{N-n}} - e^{-\frac{T-t}{\varepsilon^2}} \right) d\beta_t.$$

By means of Itô's isometry formula, we get

$$\mathbb{E}[|m_N^\varepsilon - m^\varepsilon(T)|^2] = |m_0^\varepsilon|^2 \left| \frac{1}{(1 + \tau)^N} - e^{-T/\varepsilon^2} \right|^2 + \frac{1}{\varepsilon^2} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \left| \frac{1}{(1 + \tau)^{N-n}} - e^{-\frac{T-t}{\varepsilon^2}} \right|^2 dt. \quad (\text{III.4.24})$$

Since $e^{-T/\varepsilon^2} = e^{-N\tau}$, Lemma III.4.11 yields

$$\left| \frac{1}{(1 + \tau)^N} - e^{-T/\varepsilon^2} \right|^2 \leq \frac{1}{N^2} = \frac{\Delta t^2}{T^2}. \quad (\text{III.4.25})$$

Moreover, we have for $n \in \llbracket 0, N-1 \rrbracket$ and $t \in [t_n, t_{n+1}]$

$$\left| \frac{1}{(1 + \tau)^{N-n}} - e^{-\frac{T-t}{\varepsilon^2}} \right| \leq \left| \frac{1}{(1 + \tau)^{N-n}} - e^{-\frac{T-t_n}{\varepsilon^2}} \right| + \left| e^{-\frac{T-t_n}{\varepsilon^2}} - e^{-\frac{T-t}{\varepsilon^2}} \right|,$$

thus, owing to Lemma III.4.11, we get

$$\begin{aligned}
\left| \frac{1}{(1+\tau)^{N-n}} - e^{-\frac{T-t}{\varepsilon^2}} \right|^2 &\leq 2 \left| \frac{1}{(1+\tau)^{N-n}} - e^{-(N-n)\tau} \right|^2 + 2 \left| e^{-\frac{T-t_n}{\varepsilon^2}} - e^{-\frac{T-t}{\varepsilon^2}} \right|^2 \\
&\leq \frac{2}{(N-n)^2} + 4 \left| e^{-\frac{T-t_n}{\varepsilon^2}} - e^{-\frac{T-t}{\varepsilon^2}} \right| \\
&\leq \frac{2}{(N-n)^2} + 4 \frac{t-t_n}{\varepsilon^2} e^{-\frac{T-t}{\varepsilon^2}},
\end{aligned}$$

by means of the identities $|e^u - e^v| \leq 2$ and $|e^u - e^v| \leq (u-v)e^v$ for $u \leq v \leq 0$. As a consequence, (III.4.24) and Assumption III.1 give

$$\begin{aligned}
\mathbb{E} [|m_N^\varepsilon - m^\varepsilon(T)|^2] &\lesssim \Delta t^2 + \frac{\Delta t}{\varepsilon^2} \sum_{n=0}^{N-1} \frac{1}{(N-n)^2} + \frac{\Delta t}{\varepsilon^2} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{1}{\varepsilon^2} e^{-\frac{T-t}{\varepsilon^2}} dt \\
&\lesssim \frac{\Delta t}{\varepsilon^2} + \frac{\Delta t}{\varepsilon^2} \int_0^T \frac{1}{\varepsilon^2} e^{-\frac{T-t}{\varepsilon^2}} dt = \frac{\Delta t}{\varepsilon^2} + \frac{\Delta t}{\varepsilon^2} \int_0^{T/\varepsilon^2} e^{-s} ds \\
&\lesssim \frac{\Delta t}{\varepsilon^2}.
\end{aligned}$$

This concludes the proof of (III.4.21) and of Lemma III.4.10. \square

We are now in position to prove Propositions III.4.8 and III.4.9.

Proof of Proposition III.4.8. Observe that $\frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon = \sqrt{\Delta t} \gamma_n + \varepsilon(m_n^\varepsilon - m_{n+1}^\varepsilon)$ and $\sqrt{\Delta t} \gamma_n \doteq \beta(t_{n+1}) - \beta(t_n)$. As a consequence, we have

$$X_{n+1}^\varepsilon = X_n^\varepsilon + \beta(t_{n+1}) - \beta(t_n) + \varepsilon(m_n^\varepsilon - m_{n+1}^\varepsilon),$$

hence

$$X_N^\varepsilon = x_0^\varepsilon + \beta(T) + \varepsilon(m_0^\varepsilon - m_N^\varepsilon).$$

Similarly, we have

$$\begin{aligned}
X^\varepsilon(T) &= x_0^\varepsilon + \int_0^T \frac{m^\varepsilon(t)}{\varepsilon} dt = x_0^\varepsilon + \beta(T) + \varepsilon(m_0^\varepsilon - m^\varepsilon(T)), \\
X_N &= x_0 + \beta(T),
\end{aligned}$$

thus we get

$$X_N^\varepsilon = X^\varepsilon(T) + \varepsilon(m^\varepsilon(T) - m_N^\varepsilon), \quad (\text{III.4.26})$$

$$X_N^\varepsilon = X_N + (x_0^\varepsilon - x_0) + \varepsilon(m_0^\varepsilon - m_N^\varepsilon). \quad (\text{III.4.27})$$

Estimate (III.4.16) follows from (III.4.26) and estimate (III.4.21) of Lemma III.4.10. Estimate (III.4.17) follows from (III.4.27), Assumption III.1, estimate (III.4.20) of Lemma III.4.10 and the assumption $d_{\mathbb{T}}(x_0^\varepsilon, x_0) \lesssim \varepsilon$. This concludes the proof of Proposition III.4.8. \square

Proof of Proposition III.4.9. Using once again the identities $\frac{\Delta t}{\varepsilon} m_{n+1}^\varepsilon = \sqrt{\Delta t} \gamma_n + \varepsilon(m_n^\varepsilon - m_{n+1}^\varepsilon)$ and $\sqrt{\Delta t} \gamma_n \doteq \beta(t_{n+1}) - \beta(t_n)$, we have

$$X_{n+1}^\varepsilon = X_n^\varepsilon \exp(\beta(t_{n+1}) - \beta(t_n) + \varepsilon(m_n^\varepsilon - m_{n+1}^\varepsilon)),$$

hence

$$X_N^\varepsilon = x_0^\varepsilon \exp(\beta(T) + \varepsilon(m_0^\varepsilon - m_N^\varepsilon)).$$

Similarly, we have

$$\begin{aligned} X^\varepsilon(T) &= x_0^\varepsilon \exp\left(\int_0^T \frac{m^\varepsilon(t)}{\varepsilon} dt\right) = x_0^\varepsilon \exp(\beta(T) + \varepsilon(m_0^\varepsilon - m^\varepsilon(T))), \\ X_N &= x_0 \exp(\beta(T)), \end{aligned}$$

thus we get

$$X_N^\varepsilon = X^\varepsilon(T) + X^\varepsilon(T) (e^{\varepsilon(m^\varepsilon(T) - m_N^\varepsilon)} - 1), \quad (\text{III.4.28})$$

$$X_N^\varepsilon = X_N + x_0^\varepsilon e^{\beta(T)} (e^{\varepsilon(m_0^\varepsilon - m_N^\varepsilon)} - 1) + (x_0^\varepsilon - x_0) e^{\beta(T)}. \quad (\text{III.4.29})$$

Let us derive (III.4.18) from (III.4.28). Owing to the identity $|e^u - 1| \leq |u| e^{|u|}$ for $u \in \mathbb{R}$ and to the Hölder inequality,

$$\begin{aligned} \mathbb{E}[|X_N^\varepsilon - X^\varepsilon(T)|] &\leq \varepsilon \mathbb{E}\left[|X^\varepsilon(T)| |m^\varepsilon(T) - m_N^\varepsilon| e^{\varepsilon|m^\varepsilon(T)| + \varepsilon|m_N^\varepsilon|}\right] \\ &\leq \varepsilon \mathbb{E}[|X^\varepsilon(T)|^6]^{1/6} \mathbb{E}[|m^\varepsilon(T) - m_N^\varepsilon|^2]^{1/2} \mathbb{E}[e^{6\varepsilon|m^\varepsilon(T)|}]^{1/6} \mathbb{E}[e^{6\varepsilon|m_N^\varepsilon|}]^{1/6}. \end{aligned}$$

The random variables $m^\varepsilon(T)$ and m_N^ε are centered and normally distributed with variance bounded uniformly in ε , owing to estimate (III.4.20) of Lemma III.4.10. Therefore, we have $\mathbb{E}[e^{6\varepsilon|m^\varepsilon(T)|}] \lesssim 1$ and $\mathbb{E}[e^{6\varepsilon|m_N^\varepsilon|}] \lesssim 1$. Using similar arguments as well as the Hölder inequality and Assumption III.1, we also have $\mathbb{E}[|X^\varepsilon(T)|^6] \lesssim 1$. Then, (III.4.28) follows from estimate (III.4.21) of Lemma III.4.10.

Similar arguments allow us to derive (III.4.19) from (III.4.29): the Hölder inequality yields

$$\mathbb{E}[|X_N^\varepsilon - X_N|] \lesssim \varepsilon |x_0^\varepsilon| \mathbb{E}[e^{4\beta(T)}]^{1/4} \mathbb{E}[|m_0^\varepsilon - m_N^\varepsilon|^2]^{1/2} \mathbb{E}[e^{4\varepsilon(m_0^\varepsilon - m_N^\varepsilon)}]^{1/4} + |x_0^\varepsilon - x_0| \lesssim \varepsilon,$$

owing to Lemma III.4.10 and Assumption III.1. This concludes the proof of Proposition III.4.9. \square

Note once again that in the particular cases of the diffusion approximation regime stated in Propositions III.4.8 and III.4.9, not only do we prove uniform accuracy as in the averaging regime (see Theorem III.3.4) when $\Delta t \rightarrow 0$ but the error estimates (III.4.17) and (III.4.19) (when $\varepsilon \rightarrow 0$ for fixed Δt) are uniform in Δt , as opposed to the corresponding estimate in the averaging regime (see Lemma III.4.4).

III.5 Qualitative numerical experiments

In this section, we illustrate qualitatively the superiority of AP schemes when the parameter ε is small, compared to non AP schemes.

To simplify the presentation, the dimension d is set equal to 1.

III.5.1 Illustration in the averaging regime

The objective of this section is to illustrate qualitatively the superiority of the AP scheme (III.3.4) proposed in Section III.3.2, when the parameter ε is small, compared with the use of crude integrators which are not AP. In particular, the numerical experiments below confirm that the limiting scheme (III.3.5) is consistent with the limiting equation (III.2.7).

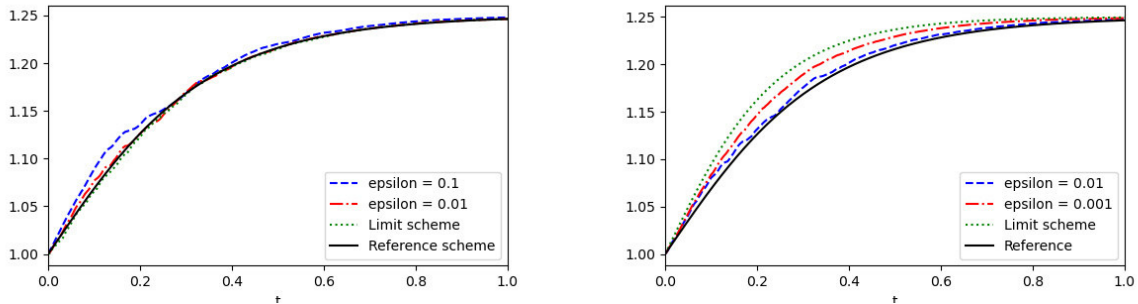
We consider the equation (III.2.1) with a drift given by $b(x, m) = \cos(2\pi x)e^{-\frac{m^2}{2}}$, and diffusion coefficient $\sigma(x, m) = 0$. Let $T = 1$, $x_0^\varepsilon = 1$ and $m_0^\varepsilon = 0$.

Recall that the AP scheme is given by (III.3.4), the limiting scheme is given by (III.3.5) and the limiting equation is given by (III.2.7), with $\bar{b}(x) = \frac{1}{\sqrt{2}} \cos(2\pi x)$ and $\bar{\sigma} = 0$. Let us define X_n^{ref} using the standard Euler scheme applied to this limiting equation:

$$X_{n+1}^{\text{ref}} = X_n^{\text{ref}} + \bar{b}(X_n^{\text{ref}})\Delta t. \quad (\text{III.5.1})$$

The scheme (III.5.1) plays the role of a reference scheme to illustrate the consistency of the limiting scheme (III.3.5) with the limiting equation, and to illustrate the fact that the crude scheme defined by (III.3.7) fails to capture the correct limit and is not AP.

In Figure III.1, we represent the evolution of X_n^ε , X_n and X_n^{ref} as time $t_n = n\Delta t$ evolves, with $\Delta t = 0.004$ and for different values of ε . In Figure III.1a, X_n^ε and X_n are computed using the AP scheme (III.3.4) and the limit scheme (III.3.5), while in Figure III.1b, X_n^ε is computed using the crude scheme (III.3.7). Observe that, in both cases, the scheme converges when $\varepsilon \rightarrow 0$ and that the AP scheme (III.3.4) does capture the correct limiting equation only with AP scheme (III.3.4), as opposed to the crude scheme (III.3.7).



(a) AP scheme (III.3.4) and its limit (III.3.5).

(b) Crude scheme (III.3.7).

Figure III.1 – Evolution of the AP scheme (III.3.4) (left), the crude scheme (III.3.7) (right), and the reference scheme (III.5.1) (averaging regime), with $\Delta t = 0.004$.

In Figure III.2, we represent the evolution of X_n^ε and X_n^{ref} as time $t_n = n\Delta t$ evolves, with $\Delta t = 0.004$ and $\varepsilon = 0.001$, when X_n is computed using the AP scheme (III.3.4) or the crude scheme (III.3.7). It illustrates the superiority of the AP scheme over the crude scheme for a small ε .

III.5.2 Illustration in the diffusion approximation regime

As in the previous section, the objective of this section is to illustrate qualitatively the superiority of the AP scheme (III.3.12) proposed in Section III.3.3, when the parameter ε is small, compared a not AP scheme.

The two examples described in Sections III.2.2 are considered below.

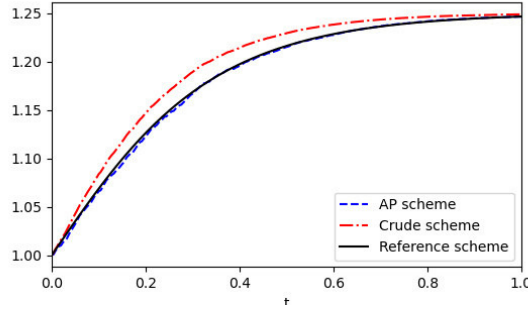


Figure III.2 – Evolution of the AP scheme (III.3.4), the crude scheme (III.3.7) and the reference scheme (III.5.1) (averaging regime), with $\Delta t = 0.004$ and $\varepsilon = 0.001$.

First example

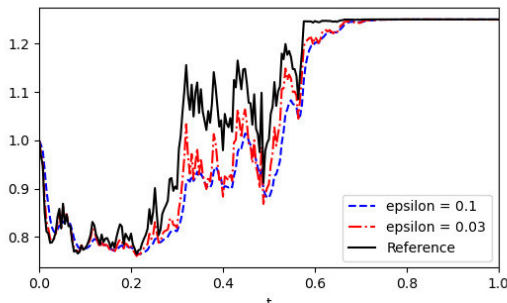
Let us consider the first example, see Equation (III.2.20). The diffusion coefficient is given by $\sigma(x) = \cos(2\pi x)$. Let $T = 1$, $x_0^\varepsilon = 1$ and $m_0^\varepsilon = 0$.

Recall that the AP scheme derived from the general case (III.3.12) in this case is given by (III.3.19), the limiting scheme is given by (III.3.20) and the limiting equation is given by (III.2.21). Let us define X_n^{ref} using the standard Euler-Maruyama scheme applied to this limiting equation (rewritten in Itô form):

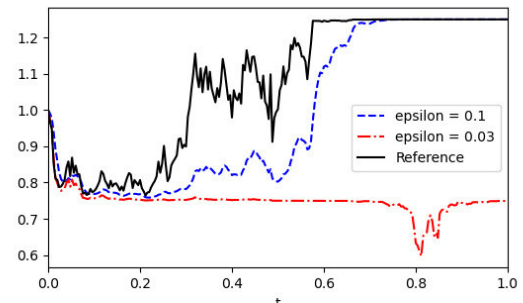
$$X_{n+1}^{\text{ref}} = X_n^{\text{ref}} + \frac{1}{2}\sigma(X_n^{\text{ref}})\sigma'(X_n^{\text{ref}})\Delta t + \sigma(X_n^{\text{ref}})\sqrt{\Delta t}\gamma_n. \quad (\text{III.5.2})$$

The scheme (III.5.2) plays the role of a reference scheme to illustrate the consistency of the limiting scheme (III.3.20) with the limiting equation, and to illustrate the fact that the crude scheme defined by (III.3.14) fails to capture the correct limit and is not AP.

In Figure III.3, we represent the evolution of X_n^ε and X_n^{ref} as time $t_n = n\Delta t$ evolves, with $\Delta t = 0.004$ and for different values of ε . The discretization X_n^ε is computed using the AP scheme (III.3.19) in Figure III.3a and the crude scheme (III.3.14) in Figure III.3b. Observe that, in both cases, the scheme seems to converge when $\varepsilon \rightarrow 0$ but only the AP scheme (III.3.19) captures the correct limiting.



(a) AP scheme (III.3.19).



(b) Crude scheme (III.3.14).

Figure III.3 – Evolution of the AP scheme (III.3.19) (left), the crude scheme (III.3.14) (right), and the reference scheme (III.5.2) (diffusion approximation regime regime, first example), with $\Delta t = 0.004$.

In Figure III.4, we represent the evolution of X_n^ε and X_n^{ref} as time $t_n = n\Delta t$ evolves, with $\Delta t = 0.004$ and $\varepsilon = 0.01$, when X_n^ε is computed using the AP scheme (III.3.19) or the crude scheme (III.3.14). Note that the behavior of the crude scheme differs from the reference. It reveals the superiority of the AP scheme for a small ε .

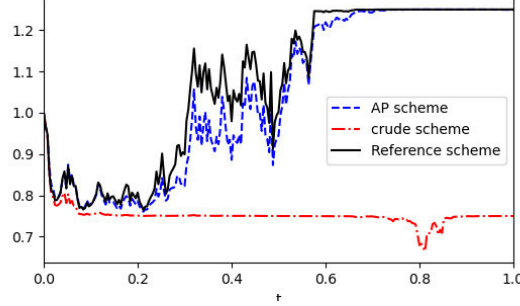


Figure III.4 – Evolution of the AP scheme (III.3.19), the crude scheme (III.3.14) and the reference scheme (III.5.2) (diffusion approximation regime, first example), with $\Delta t = 0.004$ and $\varepsilon = 0.01$.

First example with $\sigma(x) = x$

In this section, we illustrate the performance of the AP scheme presented in Remark III.3.4, and an important feature of all the AP schemes presented in this chapter, concerning the consistency of quadrature rules for discretization of the fast component.

As explained in Remark III.3.4, when $\sigma(x) = x$, where $x \in \mathbb{R}$ belongs to the real line instead of imposing periodic conditions, another type of AP scheme (III.3.23) can be designed. The limiting equation is $dX_t = X_t \circ dW_t$ or, with an Itô convention, $dX_t = \frac{1}{2}X_t dt + X_t dW_t$, and the Euler-Maruyama scheme (used as a reference scheme) for this limiting equation is written as

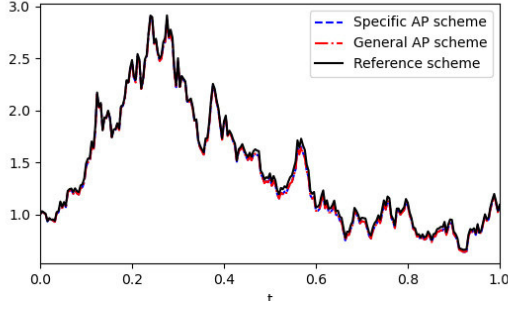
$$X_{n+1}^{\text{ref}} = X_n^{\text{ref}} + \frac{1}{2}X_n^{\text{ref}}\Delta t + X_n^{\text{ref}}\sqrt{\Delta t}\gamma_n. \quad (\text{III.5.3})$$

Recall that in (III.3.23), the quadrature rule used to discretize the integral in the exponential is closely related to the choice of the scheme for the discretization of the fast component. Let us introduce the following scheme where the consistency is not satisfied (scheme (III.3.23) corresponds to $\theta = \theta'$ below):

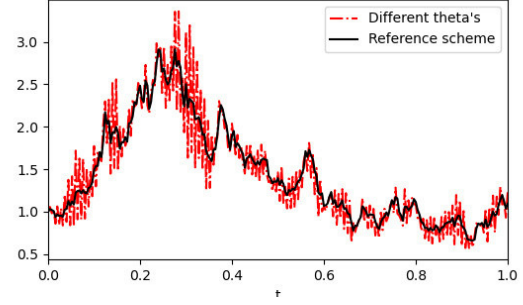
$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon \exp\left(\frac{\Delta t}{\varepsilon}[(1-\theta)m_n^\varepsilon + \theta m_{n+1}^\varepsilon]\right) \\ m_{n+1}^\varepsilon = m_n^\varepsilon - \frac{\Delta t}{\varepsilon^2}[(1-\theta')m_n^\varepsilon + \theta' m_{n+1}^\varepsilon] + \frac{\sqrt{\Delta t}}{\varepsilon}\gamma_n, \end{cases} \quad (\text{III.5.4})$$

with $\theta, \theta' \in [\frac{1}{2}, 1]$.

In Figure III.5, we represent the evolution of X_n^ε and X_n^{ref} as time evolves, with $\Delta t = 0.004$ and $\varepsilon = 0.01$. In Figure III.5a X_n^ε is computed either the specific AP scheme (III.3.23) or the general AP scheme (III.3.12), while in Figure III.5b, it is computed using the scheme (III.5.4) above with $\theta = 1 \neq \theta' = 0.5$. It illustrates the AP property of both schemes (III.3.23) and (III.3.19) and the non convergence when the quadrature rules are not chosen consistently.



(a) AP schemes (III.3.23) and (III.3.19).



(b) Scheme (III.5.4) with $\theta = 1 \neq \theta' = 0.5$.

Figure III.5 – Evolution of the AP schemes (III.3.23) and (III.3.19) (left), the crude scheme (III.5.4) with $\theta \neq \theta'$ (right), and the reference scheme (III.5.3) (diffusion approximation regime, first example with $\sigma(x) = x$), $\Delta t = 0.004$ and $\varepsilon = 0.01$.

Second example

Let us now consider the second example described in Section III.2.2, see Equation (III.2.22). The coefficients are given by $f(x) = \cos(2\pi x) + 1.5$, $g(x) = 0$ and $h(x) = 1$. Let $T = 1$, $x_0^\varepsilon = 1$ and $m_0^\varepsilon = 0$.

The general case (III.3.12) gives in this case the AP scheme (III.3.21) and the limiting scheme (III.3.22), whereas the limiting equation is given by (III.2.23). The reference scheme is obtained by using the standard Euler-Maruyama scheme applied to the limiting equation:

$$X_{n+1}^{\text{ref}} = X_n^{\text{ref}} - \Delta t \frac{f'(X_n^{\text{ref}})}{2f(X_n^{\text{ref}})} + \sqrt{\Delta t} \gamma_n. \quad (\text{III.5.5})$$

We represent in Figure III.6 the evolutions of X_n^ε and X_n^{ref} as time evolves, with $\Delta t = 0.004$ and $\varepsilon = 0.01$, where X_n^ε is computed using the AP scheme (III.3.21) (left) and the crude scheme (III.3.14) (right). Observe that the AP scheme captures the correct limiting equation when $\varepsilon \rightarrow 0$, whereas the crude scheme does not.

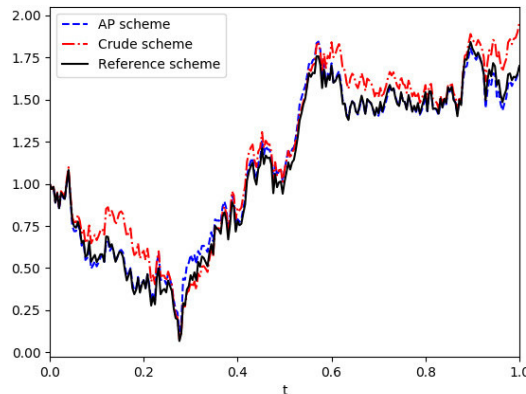


Figure III.6 – Evolution of the AP scheme (III.3.21), the crude scheme (III.3.14), and the reference scheme (III.5.5) for the second example (III.2.22) (diffusion approximation regime, second example), with $\Delta t = 0.004$ and $\varepsilon = 0.01$.

III.6 Quantitative illustration in the averaging regime

In this section, we illustrate quantitatively the superiority of an AP scheme proposed in Section III.3.2 in the averaging regime, compared with a non AP scheme. In particular, we get a better order of convergence when $\Delta t \rightarrow 0$ or $\varepsilon \rightarrow 0$ than the theoretical total error.

We consider the second variant of the AP scheme (III.3.4), introduced in Remark III.3.1, namely with $c(\tau) = \frac{1}{1+\tau}$. We choose this second variant because it becomes hard to estimate the rate of convergence when $\varepsilon \rightarrow 0$ when $c(\tau) = e^{-\tau}$.

For the numerical experiments, we use $b(x, m) = b(x + m)$, with $b(x) = x$ (non-periodic boundary conditions) or $b(x) = \cos(2\pi x)$. For simplicity, we take $\sigma(x, m) = 0$. The initial conditions are $x_0^\varepsilon = 1$ and $m_0^\varepsilon = 0$. The final time is $T = 1$.

The code for these simulations was written by Benoit Fabrèges.

III.6.1 Total error

In this section, we investigate the total error in the averaging regime

$$\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))], \quad (\text{III.6.1})$$

with $N\Delta t = T$, for the AP scheme (III.3.4) and the non AP scheme (III.3.7), where φ is a test function.

We use a Monte-Carlo method to approximate the expectation by the average over M samples, with M varies in $\{10^6, 10^7, 10^8, 10^9\}$. The reference value for $\mathbb{E}[\varphi(X^\varepsilon(T))]$ is computed using the studied scheme with $\Delta t = \Delta t_{\text{ref}} = 2^{-17}$. The test function is given by $\varphi(x) = \sin(2\pi x)^2$, $\varphi(x) = x^2$ or $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$.

In each figure, we plot the total error (III.6.1) in terms of Δt , for different values of ε . We consider $\varepsilon \in \{4^i 10^{-6} \mid i \in \llbracket 1, 6 \rrbracket\}$, $\Delta t_{\text{ref}} = 2^{-17}$ and $\Delta t \in \{2^i \mid i \in \llbracket -16, -6 \rrbracket\}$.

Note that, owing to Theorem III.3.4, the error is of order 1 with respect to Δt for fixed ε , but the error uniformly in ε is of order at least $\frac{1}{2}$ with respect to Δt .

AP scheme

Figures III.7, III.8, III.9 and III.10 illustrate the order of convergence 1 with respect to Δt of the AP scheme for fixed ε . In Figure III.8, the error becomes noisy when Δt is too small, due to the small number of Monte-Carlo samples and maybe to the unboundedness of the test function φ . In Figures III.7, III.9 and III.10, the test function is bounded and a lower bound on the order of convergence appears on the whole range of Δt . For this reason, we only consider bounded test functions in the rest of this paper. The error becomes larger for smaller ε which is consistent with the error estimate $O(\frac{\Delta t}{\varepsilon})$ of Theorem III.3.4 obtained using a direct estimation. However, the error is bounded when $\varepsilon \rightarrow 0$, illustrating the error estimate $O(\varepsilon + \Delta t)$ of Theorem III.3.4 and the uniform accuracy of the scheme. This uniform error of order 1 with respect to Δt is coherent with the order of convergence 1 of the limiting scheme, but is better than the order of convergence $\frac{1}{2}$ predicted by the uniform accuracy analysis (Theorem III.3.4).

In Figures III.7 and III.9, the error curves are slightly concave, illustrating an order of convergence higher than 1 (in particular, an order 2 in Figure III.7) that is not predicted by the error analysis. However, this phenomenon may depend on the choice of the test function, since in Figure III.10 with the test function $\varphi(x) = \sin(2\pi x + \frac{\pi}{4})$, the error curves

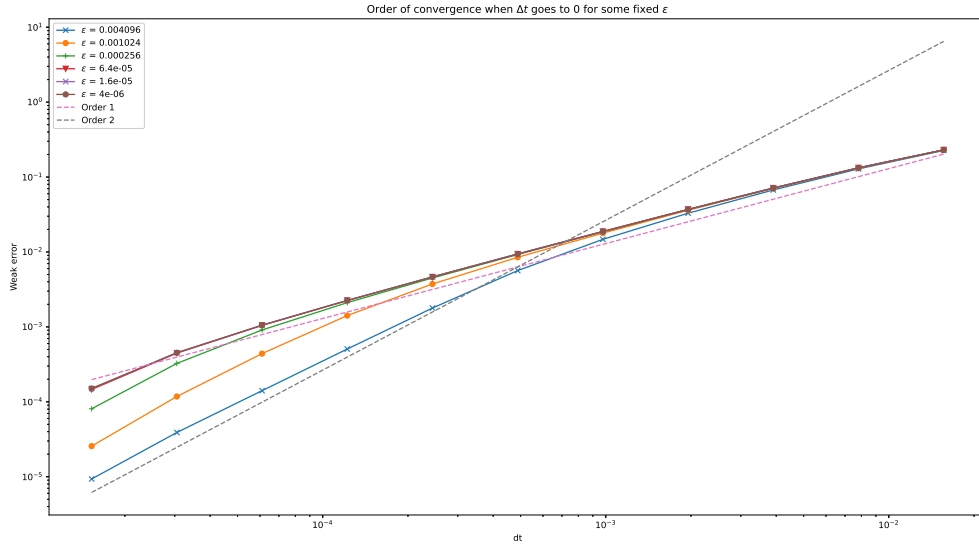


Figure III.7 – Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^9$ Monte-Carlo samples.

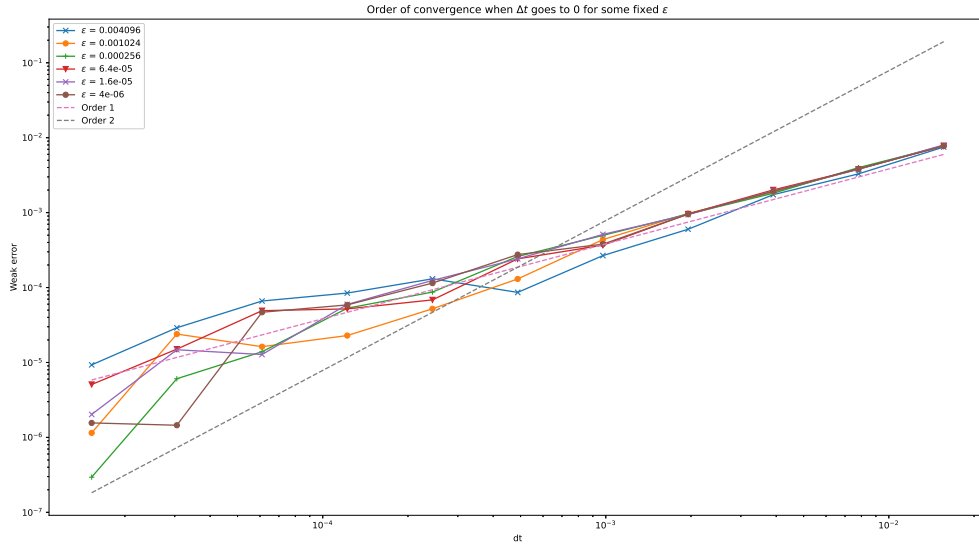


Figure III.8 – Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = \cos(2\pi x)$, $\varphi(x) = x^2$ and $M = 10^6$ Monte Carlo samples.

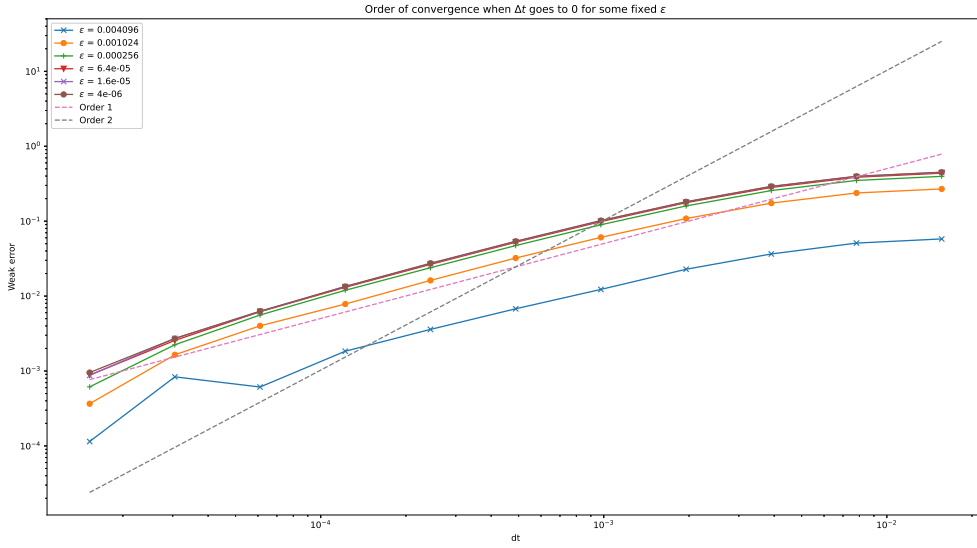


Figure III.9 – Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = x$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^6$ Monte Carlo samples.

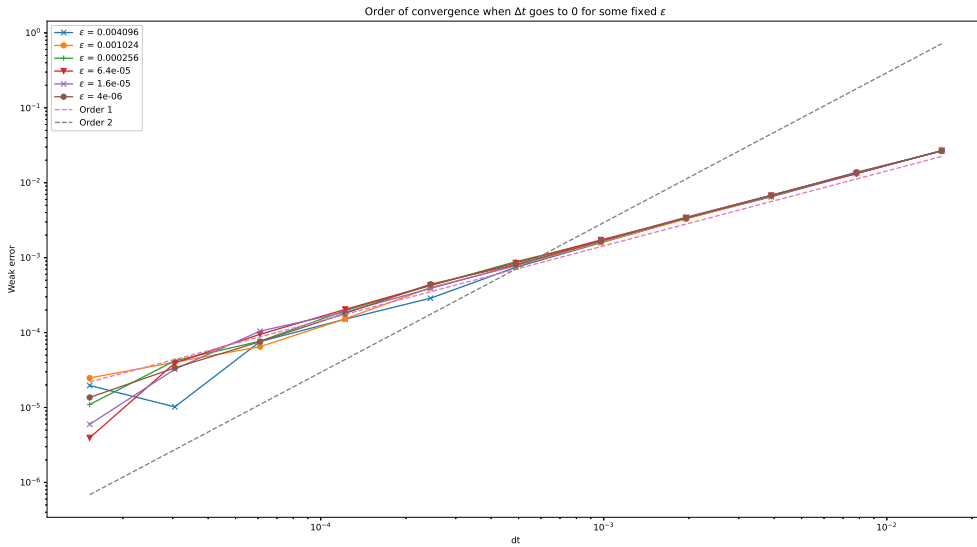


Figure III.10 – Total error (III.6.1) of the AP scheme (III.3.4) with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^6$ Monte Carlo samples.

are straight lines. For instance, the symmetry properties of the distribution with respect to the test function can increase the order of convergence.

Non AP scheme

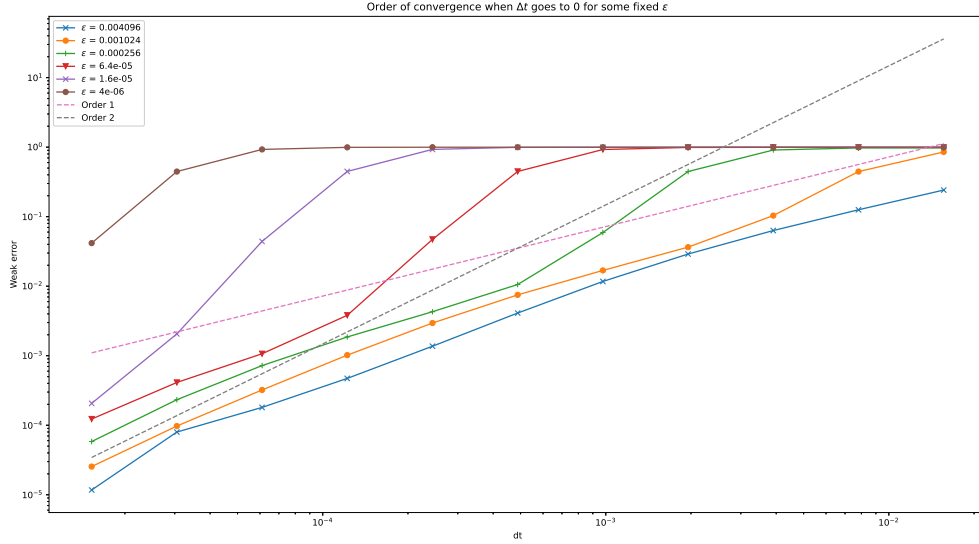


Figure III.11 – Total error (III.6.1) of the non AP scheme (III.3.7) with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^6$ Monte Carlo samples.

In Figure III.11, the error is capped due to the boundedness of the test function φ . However, the slope and the relative position of the curves hints that the error explodes when $\varepsilon \rightarrow 0$ at fixed Δt . This illustrates a case of non AP and non uniform accurate scheme. Indeed, to reduce the error, one needs Δt to be small enough depending on ε , and in particular one needs $\Delta t \xrightarrow{\varepsilon \rightarrow 0} 0$.

III.6.2 Averaging error at the discrete time level

In this section, we investigate the averaging error at the discrete time level when $\varepsilon \rightarrow 0$ at fixed Δt

$$\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)], \quad (\text{III.6.2})$$

with $N\Delta t = T$, where X_N^ε is computed using either the AP scheme (III.3.4) or the non AP scheme (III.3.7), where X_N is computed using the corresponding limiting scheme and where φ is a test function.

Owing to Lemma III.4.4 below (see Section III.4), the AP scheme (III.3.4) and its limiting scheme (III.3.5) satisfy the following estimate: for all $T \in \mathbb{R}^+$ and φ regular enough, there exists $C(T, \varphi) \in (0, \infty)$ such that

$$|\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)]| \leq C(T, \varphi) \max(\Delta t, \varepsilon).$$

We use a Monte-Carlo method to approximate the expectation by the average over M samples, with M varies in $\{10^6, 10^7, 10^8, 10^9\}$. The test function is given by $\varphi(x) = \sin(2\pi x)^2$ or $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$.

Note that for a fixed Δt , the same random variables $(\gamma_n)_{0 \leq n \leq N-1}$ are used for all values of ε . This may be seen as a reduction of variance strategy.

In each figure, we plot the averaging error (III.6.2) in terms of ε , for different values of Δt . We consider $\Delta t \in \{10^{-4}, 10^{-5}, 10^{-6}\}$ and $\varepsilon \in \{4^i 10^{-4} \mid i \in \llbracket -2, 14 \rrbracket\}$.

AP scheme

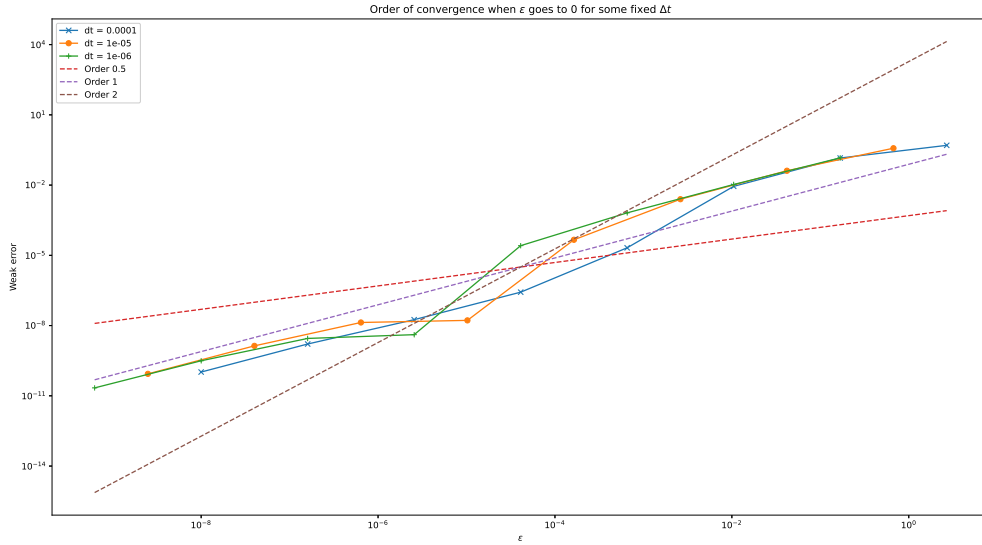


Figure III.12 – Averaging error at the discrete time level (III.6.2) of the AP scheme (III.3.4), with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$ and $M = 10^8$ Monte Carlo samples.

Figures III.12 and III.13 illustrate the order of convergence 1 with respect to ε of the averaging error at the discrete time level. There seems to be a regime $\Delta t \sim \varepsilon$ where the error is significantly reduced, which is not predicted by the error analysis. The anomaly becomes less relevant when choosing another test function (Figure III.13). As a consequence, we primarily focus on the test functions $\varphi(x) = \sin(2\pi x + \frac{\pi}{4})$ from now on.

Note that the observed upper bound on the error is $O(\varepsilon)$ instead of $O(\max(\varepsilon, \Delta t))$ predicted by Theorem III.3.4. Thus, the theoretical upper bound may not be optimal.

Non AP scheme

Figure III.14 illustrates the convergence of the non AP scheme (III.3.7) towards its limiting scheme. The error remains bounded due to the boundedness of the test function φ , but the error is higher when $\Delta t \rightarrow 0$. Therefore, one must satisfy a condition of the form $\Delta t \xrightarrow{\varepsilon \rightarrow 0} 0$ to have a bounded error when using the non AP scheme.

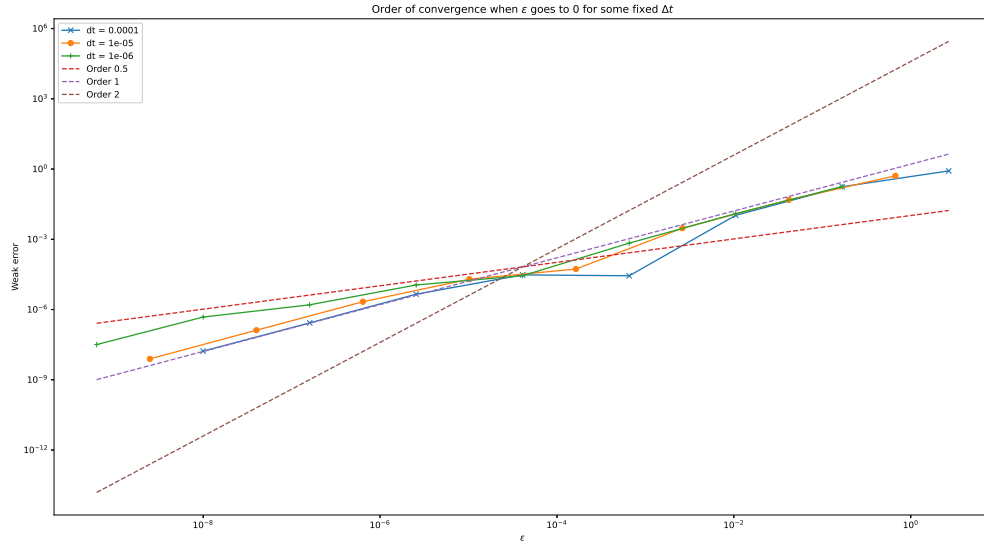


Figure III.13 – Averaging error at the discrete time level (III.6.2) of the AP scheme (III.3.4), with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^6$ Monte Carlo samples.

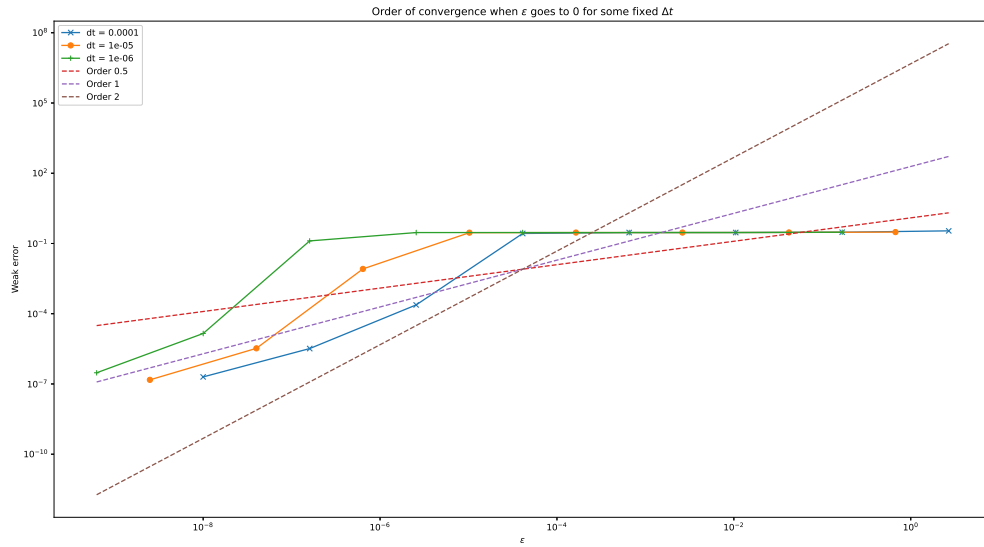


Figure III.14 – Averaging error at the discrete time level (III.6.2) of the non AP scheme (III.3.7), with $b(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^6$ Monte Carlo samples.

III.7 Quantitative illustration in the diffusion approximation regime

In this section, we illustrate quantitatively the superiority of an AP scheme proposed in Section III.3.2, compared with a non AP scheme. It allows us to conjecture the order of convergence in the general case when $\Delta t \rightarrow 0$ or $\varepsilon \rightarrow 0$ and to confirm the orders of convergence in the particular cases of Propositions III.4.8 and III.4.9.

For the numerical experiments, we use $\sigma(x) = \cos(2\pi x)$ or $\sigma(x) = \cos(2\pi x)^2$. The initial conditions are $x_0^\varepsilon = 1$ and $m_0^\varepsilon = 0$. The final time is $T = 1$.

The code for these simulations was written by Benoit Fabrèges.

III.7.1 Total error

In this section, we investigate the total error in the diffusion approximation regime

$$\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X^\varepsilon(T))], \quad (\text{III.7.1})$$

with $N\Delta t = T$, for the AP scheme (III.3.12) and the non AP scheme (III.3.14), where φ is a test function.

We use a Monte-Carlo method to approximate the expectation by the average over M samples, with $M \in \{10^6, 10^7, 10^8, 10^9\}$. The reference value for $\mathbb{E}[\varphi(X^\varepsilon(T))]$ is computed using the studied scheme with $\Delta t = \Delta t_{\text{ref}} = 2^{-17}$. The test function is given by $\varphi(x) = \sin(\pi x + \frac{\pi}{4})^2$.

In each figure, we plot the total error (III.6.1) in terms of Δt , for different values of ε . We consider $\varepsilon \in \{2^i 10^{-3} \mid i \in \llbracket 1, 6 \rrbracket\}$, $\Delta t_{\text{ref}} = 2^{-17}$ and $\Delta t \in \{2^i \mid i \in \llbracket -16, -6 \rrbracket\}$.

AP scheme

We do not know of any estimate for the total error of the AP scheme (III.3.12) uniformly in ε , and use the following numerical experiment to conjecture the order of convergence of the total error. When ε is fixed, (III.3.12) is a variant of the Heun method or the Runge-Kutta-2 scheme. Indeed, if m^ε was exactly known, then it would be a Heun scheme with an order of convergence 2 with respect to Δt [KP92, scheme (1.3) of Section 15.1]. However, m^ε is approximated by an implicit Euler scheme, which converges at order 1. Since we only consider X^ε instead of $(X^\varepsilon, m^\varepsilon)$, we expect an order of convergence 2 for (III.3.12) at fixed ε . The limiting scheme is a classical scheme of order 1 for the limiting SDE with a Stratonovich interpretation of the noise.

Figure III.15 illustrates an order of convergence 2 with respect to Δt for large fixed ε and large Δt for the AP scheme (III.3.12), which is better than the expected order 1. This may be due to the fact that the discretization of the slow component X^ε is of order 2, and that the lower order 1 is caused by the fast component. The error becomes larger for smaller ε but does not explode when $\varepsilon \rightarrow 0$, and instead is of order 1 with respect to Δt when considering the whole range of Δt . We thus conjecture that the scheme is uniformly accurate. For small ε , this is consistent with the order of convergence 1 of the limiting scheme towards the limiting equation with Stratonovich noise. This phenomenon is similar to the averaging regime: we do not observe a reduction of the order of convergence, even though we should expect such reduction if we had an equivalent to Theorem III.3.4.

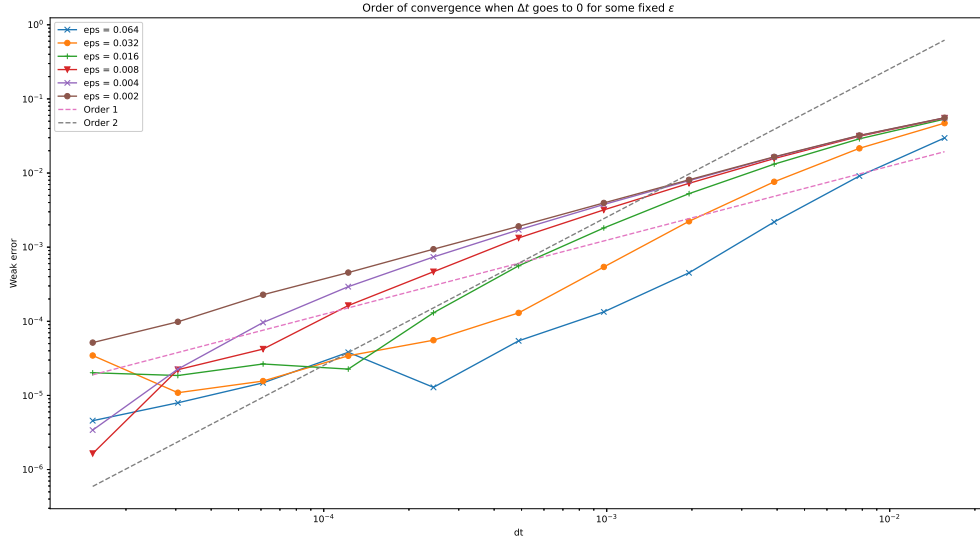


Figure III.15 – Total error (III.7.1) of the AP scheme (III.3.12), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^9$ Monte Carlo samples.

Non AP scheme

In Figure III.16, the error reaches a threshold due to the boundedness of the test function φ , but its behavior for small enough Δt indicates that the error does not vanish when $\varepsilon \rightarrow 0$. The scheme is therefore not uniformly accurate.

III.7.2 Diffusion approximation error at the discrete time level

In this section, we investigate the diffusion approximation error at the discrete time level

$$\mathbb{E}[\varphi(X_N^\varepsilon)] - \mathbb{E}[\varphi(X_N)], \quad (\text{III.7.2})$$

with $N\Delta t = T$, where X_N^ε is computed using either the AP scheme (III.3.12) or the non AP scheme (III.3.14), where X_N is computed using the corresponding limiting scheme and where φ is a test function.

We do not know any estimate for this error and this scheme in the general case, and use the following numerical experiment to conjecture the order of convergence. In the particular cases of Propositions III.4.8 and III.4.9, we proved an order of convergence 1 with respect to ε .

We use a Monte-Carlo method to approximate the expectation by the average over M samples, with $M \in \{10^6, 10^7, 10^8, 10^9\}$. The test function is given by $\varphi(x) = \sin(2\pi x)^2$ or $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$.

Note that for a fixed Δt , the same random variables $(\gamma_n)_{0 \leq n \leq N-1}$ are used for all values of ε .

In each figure, we plot the averaging error (III.6.2) in terms of ε , for different values of Δt .

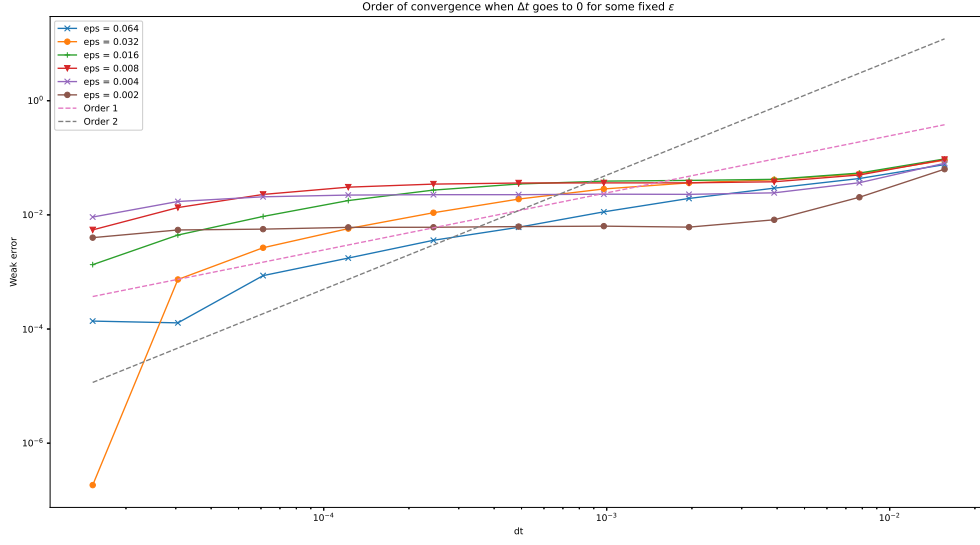


Figure III.16 – Total error (III.7.1) of the non AP scheme (III.3.14), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$ and $M = 10^7$ Monte Carlo samples.

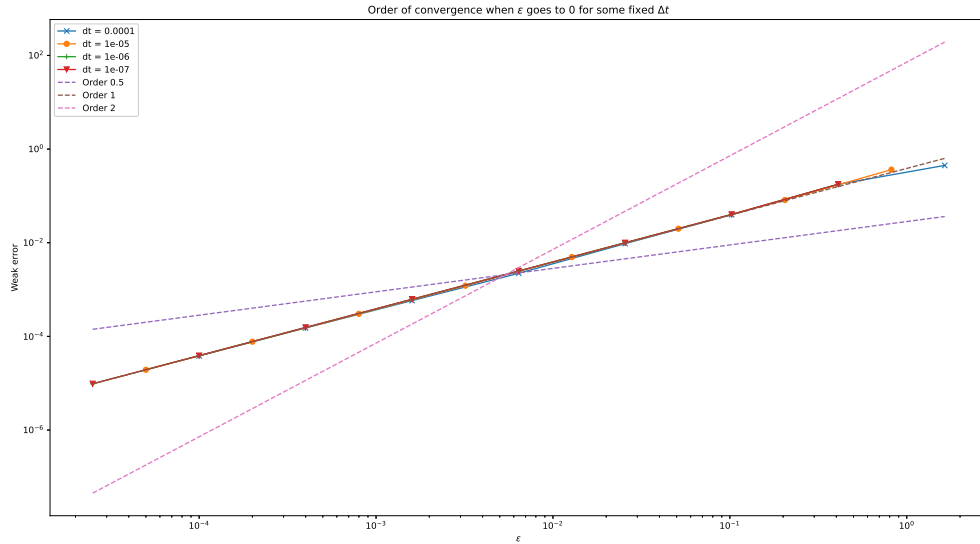


Figure III.17 – Diffusion approximation error at the discrete time level (III.7.2) for the AP scheme (III.3.12), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$, $M = 10^6$ Monte Carlo samples, $\Delta t \in \{10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$ and $\varepsilon \in \{4^i 10^{-4} \mid i \in \llbracket 0, 6 \rrbracket\}$.

AP scheme

Figure III.17 illustrates an order of convergence 1 with respect to ε . This is the same order as in the particular cases of Propositions III.4.8 and III.4.8, suggesting that this is the best order of convergence one could hope for. Moreover, we observe a convergence with respect to ε which is uniform with respect to Δt , and the order of convergence is the same than in the averaging regime.

Non AP scheme

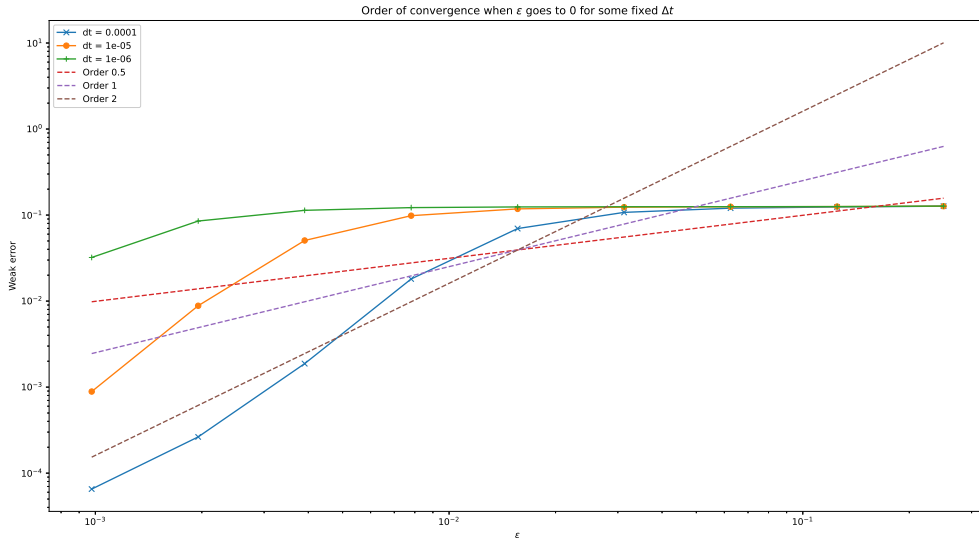


Figure III.18 – Diffusion approximation error at the discrete time level (III.7.2) for the non AP scheme (III.3.14), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(2\pi x)^2$, $M = 10^6$ Monte Carlo samples, $\Delta t \in \{10^{-4}, 10^{-5}, 10^{-6}\}$ and $\varepsilon \in \{2^i \mid i \in \llbracket -10, -2 \rrbracket\}$.

Figures III.18 and III.19 illustrate the convergence of the non AP scheme (III.3.14) towards its limiting scheme. However, in Figure III.18 one must satisfy a condition of the form $\Delta t \xrightarrow{\varepsilon \rightarrow 0} 0$ to have a bounded error, the scheme is thus not AP. Moreover, the limiting scheme is not consistent with the limiting equation. Once again, the threshold is due to the boundedness of the test function φ .

In Figure III.19, the non AP property is more subtle. Indeed, it can only be read in the non uniformity of the error with respect to Δt when $\varepsilon \rightarrow 0$.

III.8 Conclusion

In this chapter, we have studied a general notion of asymptotic preserving schemes, related to convergence in distribution, for a class of SDE systems in averaging and diffusion approximation regimes. Let us mention that some assumptions made to simplify the setting (the slow component takes values in a compact set \mathbb{T}^d and the fast component is one-dimensional) may easily be relaxed. Note that when the slow component takes values

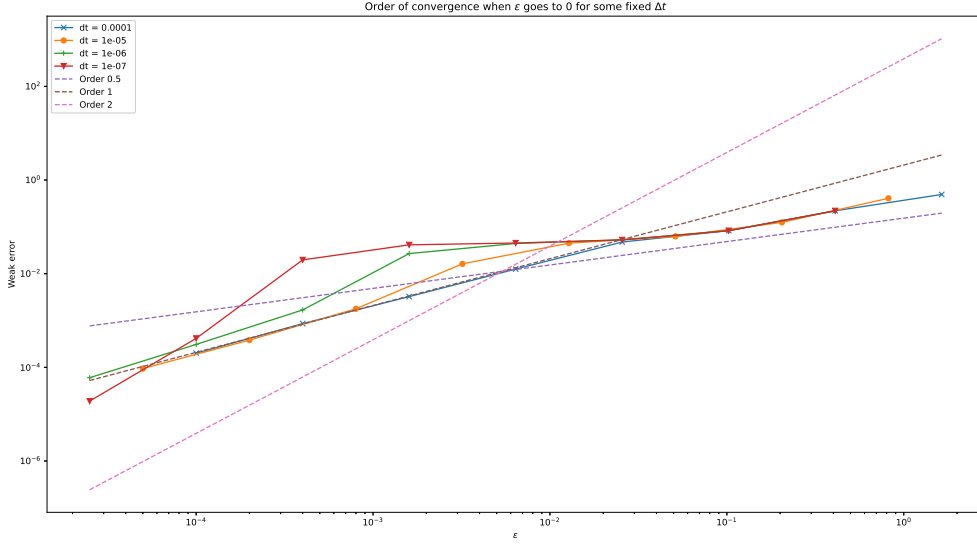


Figure III.19 – Diffusion approximation error at the discrete time level (III.7.2) for the non AP scheme (III.3.14), with $\sigma(x) = \cos(2\pi x)$, $\varphi(x) = \sin(\pi x + \frac{\pi}{4})$, $M = 10^6$ Monte Carlo samples, $\Delta t \in \{10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$ and $\varepsilon \in \{2^i \mid i \in \llbracket -10, -2 \rrbracket\}$.

in \mathbb{R}^d , it is necessary to also study the stability of the numerical schemes, for instance in mean-square sense.

A limitation of our study is the fact that the fast component is an Ornstein-Uhlenbeck process (when the slow component is frozen): even if the general theory of AP schemes described in Section III.3.1 holds in more general settings, the construction of implementable AP schemes (such as the ones described in Sections III.3.2 and III.3.3) is not straightforward if for instance the fast component is solution of a general ergodic SDE with nonlinear coefficients.

We have also left open the question of obtaining a version of the error estimates stated in Theorem III.3.4 in the diffusion approximation case. This question will be studied in future works.

III.A Appendix - Derivation of the limiting models

III.A.1 Sketch of proof of Proposition III.2.1 (averaging regime)

Let us first give details concerning the construction of the perturbed test function φ^ε given by (III.2.9), such that (III.2.10) holds. Recall that this construction is used in the statement of Proposition III.3.1.

Owing to the multiscale expansions (III.2.4) and (III.2.9) of the generator \mathcal{L}^ε and of the perturbed test function $\varphi^\varepsilon = \varphi + \varepsilon\varphi_1$, one has

$$\mathcal{L}^\varepsilon \varphi^\varepsilon = \varepsilon^{-1} \mathcal{L}_{OU} \varphi + (\mathcal{L}_0 \varphi + \mathcal{L}_{OU} \varphi_1) + \varepsilon \mathcal{L}_0 \varphi_1. \quad (\text{III.A.1})$$

Since the test function φ does not depend on m , one has $\mathcal{L}_{OU} \varphi = 0$, thus the term of order ε^{-1} in (III.A.1) vanishes.

Define, for all $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$,

$$\begin{aligned}\mathcal{L}\varphi(x) &= \int_{\mathbb{R}} \mathcal{L}_0\varphi(x, m) d\nu^x(m) \\ &= \bar{b}(x) \cdot \nabla_x \varphi(x) + \bar{\sigma} \bar{\sigma}^*(x) : \nabla_x^2 \varphi(x) \\ \vartheta(x, m) &= \mathcal{L}_0\varphi(x, m) - \mathcal{L}\varphi(x) \\ &= (b(x, m) - \bar{b}(x)) \cdot \nabla_x \varphi(x) + (\sigma\sigma^*(x, m) - \bar{\sigma}\bar{\sigma}^*(x)) : \nabla_x^2 \varphi(x),\end{aligned}$$

where we recall that $\nu^x = \mathcal{N}(0, h(x)^2)$ is the invariant distribution of the ergodic Ornstein-Uhlenbeck process m^x associated to \mathcal{L}_{OU} on \mathbb{R} , for any fixed $x \in \mathbb{T}^d$

$$dm_t^x = -m_t^x dt + \sqrt{2}h(x)d\beta_t.$$

Let $m^x(t, m)$ denote the solution at time t , if the initial condition is given by $m^x(0, m)$. Therefore, the centering condition $\int \vartheta(x, m) d\nu^x(m) = 0$ is satisfied and the Poisson equation $-\mathcal{L}_{OU}\varphi_1(x, \cdot) = \vartheta(x, \cdot)$ admits a solution

$$\varphi_1(x, m) = \int_0^\infty \mathbb{E} [\vartheta(x, m^x(t, m))] dt.$$

The multiscale expansion (III.A.1) becomes

$$\mathcal{L}^\varepsilon \varphi^\varepsilon = \mathcal{L}\varphi + \varepsilon \mathcal{L}_0\varphi_1.$$

To prove (III.2.10), it only remains to get estimates on $\mathcal{L}_0\varphi_1$ uniformly in ε . Consider $V = b$ or $V = \sigma\sigma^*$ and let

$$\delta(t, x, m) \doteq \mathbb{E} [V(x, m^x(t, m)) - \bar{V}(x)].$$

Note that for m and $m' \in \mathbb{R}$, one has $m^x(t, m) - m^x(t, m') = (m - m')e^{-t}$. As a consequence, we have

$$\|\delta(t, x, m) - \delta(t, x, m')\| \lesssim |m - m'| e^{-t}.$$

By integrating with respect to m' and using the equality $\bar{V}(x) = \int V(x, m') d\nu^x(m')$, one obtains

$$\|\delta(t, x, m)\| \lesssim (1 + |m|)e^{-t}. \quad (\text{III.A.2})$$

Since V is of class \mathcal{C}^3 with bounded derivatives, and since the derivatives of $m^x(t, m)$ with respect to x do not depend on m , it is straightforward to generalize III.A.2 to the derivatives of δ . It gives that $\varphi_1 \in \mathcal{C}^2(\mathbb{T}^d \times \mathbb{R})$ and that φ_1 and its derivatives have at most linear growth in m , hence $\mathcal{L}_0\varphi_1$ also does. This leads to (III.2.10) using (III.2.2). This concludes the identification of the limiting generator \mathcal{L} using the perturbed test function method. The remaining ingredients of this strategy to prove the convergence in distribution of the process X^ε to the solution X of the limiting equation associated with the limiting generator \mathcal{L} are standard and are thus omitted.

III.A.2 Sketch of proof of Proposition III.2.2 (diffusion approximation regime)

Let us first give details concerning the construction of the perturbed test function φ^ε given by (III.2.17), such that (III.2.18) holds. Recall that this construction is used in the statement of Proposition III.3.1.

Owing to the multiscale expansions (III.2.14) and (III.2.17) of the generator \mathcal{L}^ε and of the perturbed test function $\varphi^\varepsilon = \varphi + \varepsilon\varphi_1 + \varepsilon^2\varphi_2$, one has

$$\begin{aligned} \mathcal{L}^\varepsilon \varphi^\varepsilon &= \varepsilon^{-2} \mathcal{L}_{OU} \varphi + \varepsilon^{-1} (\mathcal{L}_1 \varphi + \mathcal{L}_{OU} \varphi_1) + (\mathcal{L}_0 \varphi + \mathcal{L}_1 \varphi_1 + \mathcal{L}_{OU} \varphi_2) \\ &\quad + \varepsilon (\mathcal{L}_0 \varphi_1 + \mathcal{L}_1 \varphi_2) + \varepsilon^2 \mathcal{L}_0 \varphi_2. \end{aligned} \quad (\text{III.A.3})$$

Since the test function φ does not depend on m , one has $\mathcal{L}_{OU} \varphi = 0$, thus the term of order ε^{-2} in (III.A.3) vanishes. Define

$$\varphi_1(x, m) \doteq m \frac{\sigma(x)}{f(x)} \cdot \nabla_x \varphi(x). \quad (\text{III.A.4})$$

Then it is straightforward to check that $\mathcal{L}_1 \varphi + \mathcal{L}_{OU} \varphi_1 = 0$, thus the term of order ε^{-1} in (III.A.3) vanishes.

It remains to construct the function φ_2 such that the term of order 1 in (III.A.3) is equal to $\mathcal{L}\varphi$. Define, for all $x \in \mathbb{T}^d$ and $m \in \mathbb{R}$,

$$\begin{aligned} \mathcal{L}\varphi(x) &= \int_{\mathbb{R}} (\mathcal{L}_0 \varphi + \mathcal{L}_1 \varphi_1)(x, m) d\nu^x(m), \\ \vartheta(x, m) &= (\mathcal{L}_0 \varphi + \mathcal{L}_1 \varphi_1)(x, m) - \mathcal{L}\varphi(x), \end{aligned}$$

where we recall that $\nu^x = \mathcal{N}(0, \frac{f(x)h(x)^2}{2})$ is the invariant distribution of the ergodic Ornstein-Uhlenbeck process associated to \mathcal{L}_{OU} on \mathbb{R} , for any fixed $x \in \mathbb{T}^d$.

Let $x \in \mathbb{T}^d$, then the Poisson equation $-\mathcal{L}_{OU} \varphi_2(x, \cdot) = \vartheta(x, \cdot)$ admits a solution φ_2 , since the centering condition $\int \vartheta(x, m) d\nu^x(m) = 0$ is satisfied. Precisely, one has the expressions

$$\begin{aligned} \vartheta(x, m) &= - \left(|m|^2 - \frac{fh^2}{2} \right) \sigma \cdot \nabla_x \left(\frac{\sigma}{f} \cdot \nabla_x \varphi \right), \\ \varphi_2(\cdot, m) &\doteq \frac{|m|^2}{2} \frac{\sigma}{f} \cdot \nabla_x \left(\frac{\sigma}{f} \cdot \nabla_x \varphi \right). \end{aligned} \quad (\text{III.A.5})$$

With the functions φ_1 and φ_2 constructed above, the multiscale expansion (III.A.3) is rewritten as

$$\mathcal{L}^\varepsilon \varphi^\varepsilon = \mathcal{L}\varphi + \varepsilon (\mathcal{L}_0 \varphi_1 + \mathcal{L}_1 \varphi_2) + \varepsilon^2 \mathcal{L}_0 \varphi_2,$$

which gives (III.2.18), more precisely

$$\sup_{x \in \mathbb{T}^d} |\mathcal{L}^\varepsilon \varphi^\varepsilon(x, m) - \mathcal{L}\varphi(x)| \leq C_\varphi (\varepsilon|m| + \varepsilon^2|m|^2),$$

for some constant C_φ depending only on φ and on the coefficients of the SDE.

It remains to check that $\mathcal{L}\varphi(x) = \int_{\mathbb{R}} (\mathcal{L}_0 \varphi + \mathcal{L}_1 \varphi_1)(x, m) d\nu^x(m)$ gives the expression (III.2.16). This concludes the identification of the limiting generator \mathcal{L} using the perturbed test function method. The remaining ingredients of this strategy to prove the convergence in distribution of the process X^ε to the solution X of the limiting equation associated with the limiting generator \mathcal{L} follows from standard arguments which are omitted.

Asymptotic preserving schemes for kinetic stochastic PDEs

IV.1 Introduction

Consider the following kinetic Stochastic Partial Differential Equation (SPDE) with a source term

where $f^{\varepsilon, \delta}$ is defined on $\mathbb{R}^+ \times \mathbb{T}^d \times V$, L is a linear operator (see (IV.1.2) below) and the source term m^δ is an Ornstein-Uhlenbeck process defined on $\mathbb{R}^+ \times \mathbb{T}^d$ (a detailed definition is given in Section IV.2). This chapter aims to propose asymptotic preserving schemes for (IV.1.1) in the regimes $\varepsilon \rightarrow 0$ for fixed δ , $\delta \rightarrow 0$ for fixed ε and $\varepsilon = \delta \rightarrow 0$.

$$Lf = \rho \mathcal{M} - f, \quad (\text{IV.1.2})$$

where μ is a measure on V , where $\rho \doteq \int_V f d\mu$ and where $\mathcal{M} \in L^1(V, d\mu)$ satisfies Assumption IV.1. We consider a general case where (V, μ) is a measured space which includes for instance the standard case $V = \mathbb{R}^d$ and $\mathcal{M}(v)d\mu(v) = \frac{1}{\sqrt{2\pi\sigma^2}^d} \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv$, but we focus in the numerical experiment on the discrete case $V = \{\pm 1\}$.

The fast component m^δ is an \mathbb{R}^J -valued Ornstein-Uhlenbeck process: for $j \in \llbracket 1, J \rrbracket$, consider a function $g_j \in \mathcal{C}^1(\mathbb{T}^d)$ and a one dimensional Ornstein-Uhlenbeck processes m_j^δ defined by

$$dm_j^\delta(t) = -\frac{\theta_j}{\delta^2}(m_j^\delta(t) - \overline{m}_j)dt + \frac{\sigma_j}{\delta}d\beta_j(t), \quad (\text{IV.1.3})$$

where $\theta_j > 0$, $\overline{m}_j \in \mathbb{R}$ and $\sigma_j > 0$. The fast component m^δ is then defined in the averaging regime by

$$m^\delta(t, x) = \sum_{j=1}^J m_j^\delta(t) g_j(x), \quad (\text{IV.1.4})$$

and in the diffusion approximation regime by

$$m^\delta(t, x) = \sum_{j=1}^J \frac{1}{\delta} m_j^\delta(t) g_j(x), \quad \forall j \in \llbracket 1, J \rrbracket, \overline{m}_j = 0. \quad (\text{IV.1.5})$$

Note that in both case, n^δ depends on the time variable t and the space variable x , but not on the velocity variable v .

The slow component $\rho^{\varepsilon, \delta} = \int_V f^{\varepsilon, \delta} d\mu$ converges to different limits in the averaging and diffusion approximation regimes.

As described in Chapter III, a scheme that is consistent for fixed values of $\varepsilon > 0$ and $\delta > 0$ can fail to capture the correct asymptotic regimes: such scheme is not asymptotic preserving.

The notion of AP schemes in the deterministic case has been introduced in [Jin99], for applications to multiscale kinetic Partial Differential Equations (PDEs), which converge to parabolic diffusion PDEs. We refer to [DP14, Section 7], [HJL17], [Jin12] and [Pup19, Section 4] for recent reviews on AP schemes for this type of models. Note that a specific feature in the stochastic case is the need to consider convergence in distribution. Let us mention related works for Stochastic Partial Differential Equations (SPDEs), in the diffusion approximation regime. First, in [DM16, Mar06], the authors consider Schrödinger equations and study an abstract asymptotic preserving property. However, they do not propose implementable schemes. In [AF19], the authors deal with some multiscale stochastic kinetic PDEs, driven by a Wiener process. In this chapter, we recover another AP scheme for this situation after taking the limit $\delta \rightarrow 0$ if m^δ is defined by (IV.1.5), but we present AP schemes in other various situations. Finally, let us also mention that AP schemes have also been studied for PDEs with random coefficients, see [HJ17, Jin18, JLP18] or in the context of Monte-Carlo methods for deterministic problems, see [DPS18, RLJ14].

In this chapter, we aim to define Asymptotic Preserving (AP) schemes for the stochastic kinetic equation (IV.1.1). The detailed definition of AP schemes is defined in Section IV.3 and is equivalent to Definition III.3.1 in the SDE case. Our AP schemes are designed using a splitting procedure, which separates the deterministic and the stochastic parts of (IV.1.1). This leads to Scheme A in the averaging regime (when m^δ is defined by (IV.1.4)) and to Scheme B in the diffusion approximation regime (when m^δ is defined by (IV.1.5)). In both schemes, we write $\langle h \rangle \doteq \int_V h d\mu$ and $\lambda \doteq \frac{\tau}{1 + \frac{\tau}{\varepsilon^2}} \in (0, 1)$, where τ is the time-step

size. Moreover, \hat{h} denotes the Fourier transform of h with respect to the space variable x , computed using a Fast Fourier Transform (FFT).

Scheme A Splitting AP scheme in the averaging regime.

$$\begin{aligned}\widehat{\rho}_{n+1/2,k}^{\varepsilon,\delta} &= \frac{(1-\lambda)\left\langle \frac{\widehat{f}_{n,k}^{\varepsilon,\delta}}{1+i\lambda\varepsilon a \cdot k} \right\rangle}{1-\lambda\left\langle \frac{\mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle} \\ \widehat{f}_{n+1/2,k}^{\varepsilon,\delta} &= \frac{1-\lambda}{1+i\lambda\varepsilon a \cdot k} \widehat{f}_{n,k}^{\varepsilon,\delta} + \frac{\lambda}{1+i\lambda\varepsilon a \cdot k} \mathcal{M} \widehat{\rho}_{n+1/2,k}^{\varepsilon,\delta} \\ m_{n+1}^\delta &= \overline{m} + \frac{\tau/\delta^2}{1+\tau/\delta^2} \theta (m_n^\delta - \overline{m}) + \sqrt{1 - \left(\frac{\tau/\delta^2}{1+\tau/\delta^2} \right)^2} \sigma \gamma_n g. \\ f_{n+1}^{\varepsilon,\delta} &= \exp(\tau m_{n+1}^\delta) f_{n+1/2}^{\varepsilon,\delta}\end{aligned}$$

Scheme B Splitting AP scheme in the diffusion approximation regime.

$$\begin{aligned}\widehat{\rho}_{n+1/2,k}^{\varepsilon,\delta} &= \frac{(1-\lambda)\left\langle \frac{\widehat{f}_{n,k}^{\varepsilon,\delta}}{1+i\lambda\varepsilon a \cdot k} \right\rangle}{1-\lambda\left\langle \frac{\mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle} \\ \widehat{f}_{n+1/2,k}^{\varepsilon,\delta} &= \frac{1-\lambda}{1+i\lambda\varepsilon a \cdot k} \widehat{f}_{n,k}^{\varepsilon,\delta} + \frac{\lambda}{1+i\lambda\varepsilon a \cdot k} \mathcal{M} \widehat{\rho}_{n+1/2,k}^{\varepsilon,\delta} \\ m_{n+1}^\delta &= m_n - \theta \frac{\tau}{\delta^2} m_{n+1}^\delta + \sigma \frac{\sqrt{\tau}}{\delta^2} \gamma_n g \\ f_{n+1}^{\varepsilon,\delta} &= \exp(\tau m_{n+1}^\delta) f_{n+1/2}^{\varepsilon,\delta}\end{aligned}$$

The objective of this chapter is to prove via numerical experiments (see Section IV.4) that the AP Schemes A and B are effective in all regimes, as opposed to the crude schemes which fail to capture the correct limiting behavior when $\delta \rightarrow 0$ and $\varepsilon = \delta \rightarrow 0$.

The chapter is organized as follows. The SPDE models in the averaging and diffusion approximation regimes are presented in Section IV.2, and we state the associated scaling limits. Then, we give examples of AP schemes for the deterministic part and the stochastic part in Section IV.3, and we combine them using a splitting procedure. Numerical experiments are reported in Section IV.4. Finally, Section IV.5 gives some conclusions and perspectives.

IV.2 Models

Assumption IV.1. The velocity space V is embedded in $\mathbb{R}^{d'}$ for some $d' \in \mathbb{N}$ and is equipped with a finite measure μ . The functions $a \in L^\infty(V, d\mu)$ and $\mathcal{M} \in L^1(V, d\mu)$ satisfy

$$\begin{aligned}\forall v \in V, \mathcal{M}(v) &> 0, \\ \int_V \mathcal{M}(v) d\mu(v) &= 1, \\ K \doteq \int_V a(v) \otimes a(v) \mathcal{M}(v) d\mu(v) &> 0.\end{aligned}$$

Moreover, we assume some symmetry on the space $(V, d\mu)$:

- The measure μ is symmetric: for all measurable subset $A \subset V$, $\mu(-A) = \mu(A)$.

- The function a is odd: for all $v \in V$, $a(-v) = -a(v)$.
- The function \mathcal{M} is even: for all $v \in V$, $\mathcal{M}(-v) = \mathcal{M}(v)$.

For the numerical experiments, we consider $V = \{\pm 1\}$, $\mu(\{1\}) = \mu(\{-1\}) = \frac{1}{2}$, $a(v) = v$ and $\mathcal{M}(v) = 1$.

Let us justify the regime corresponding to each definition of m^δ .

- In the averaging regime, m^δ is given by (IV.1.4). If $\bar{m}(x) = \sum_{j=1}^J \bar{m}_j g_j(x)$, then the term $m^\delta f^{\varepsilon, \delta}$ gives, in the limiting equation when $\delta \rightarrow 0$, a averaged source term $\bar{m} f^\varepsilon$ (see Section IV.2.1).
- In the diffusion approximation regime, m^δ is given by (IV.1.5). The term $m^\delta f^{\varepsilon, \delta}$ gives, in the limiting equation when $\delta \rightarrow 0$, a term $f^\varepsilon Q^{1/2} \circ dW(t)$ where W is a cylindrical Wiener process. The definition of the covariance operator Q for a general invariant process m can be found in [DV12] and in Chapter I. When m^δ is defined by (IV.1.5), we get

$$Qf(x) = \sum_{j=1}^J \frac{\sigma_j^2}{\theta_j^2} g_j(x) \int_{\mathbb{T}^d} g_j(y) f(y) dy. \quad (\text{IV.2.1})$$

It is a nonnegative self-adjoint operator and thus admits a square root $Q^{1/2}$.

Assuming the functions g_j are linearly independent and pairwise orthogonal in $L^2(\mathbb{T}^d)$, Q is of rank J , its only non-zero eigenvalues are $\frac{\sigma_j^2}{\theta_j^2} \|g_j\|_{L^2(\mathbb{T}^d)}^2$, associated with the normalized eigenvectors $\|g_j\|_{L^2(\mathbb{T}^d)}^{-1} g_j$, and

$$Q^{1/2} f(x) = \sum_{j=1}^J \frac{\sigma_j}{\theta_j \|g_j\|_{L^2(\mathbb{T}^d)}} g_j(x) \int_{\mathbb{T}^d} g_j(y) f(y) dy.$$

The spectral decomposition yields

$$Q^{1/2} dW(t) = \sum_{j=1}^J \frac{\sigma_j}{\theta_j} g_j d\beta_j(t) \quad (\text{IV.2.2})$$

in distribution. Moreover, the Itô form of $fQ^{1/2} \circ dW(t)$ is

$$\begin{aligned} fQ^{1/2} \circ dW(t) &= \frac{1}{2} F f + fQ^{1/2} dW(t) \\ F(x) &\doteq k(x, x) = \sum_{j=1}^J \frac{\sigma_j^2}{\theta_j^2} g_j(x)^2. \end{aligned} \quad (\text{IV.2.3})$$

Let us detail how to get the expression (IV.2.1) of Q . For $j \in \llbracket 1, J \rrbracket$, consider the Ornstein-Uhlenbeck process $dm_j(t) = -\theta_j m_j(t) dt + \sigma_j d\beta_j(t)$, so that if $m_j(0) = m_j^\delta(0)$, then $m_j(t\delta^{-2}) = m_j^\delta(t)$. The Markov process m_j admits a unique invariant

distribution $\nu_j \sim \mathcal{N}\left(0, \frac{\sigma_j^2}{2\theta_j}\right)$. Let $m_j(0)$ be distributed following ν_j so that m_j is a stationary process, and consider the kernel k defined for $x, y \in \mathbb{T}^d$ by

$$\begin{aligned} k(x, y) &\doteq 2 \sum_{j=1}^J g_j(x) g_j(y) \mathbb{E} \left[\int_{\mathbb{R}^+} m_j(0) m_j(t) dt \right] \\ &= 2 \sum_{j=1}^J g_j(x) g_j(y) \mathbb{E} \left[\int_{\mathbb{R}^+} m_j(0)^2 e^{-\theta_j t} dt \right] + \mathbb{E} \left[\int_{\mathbb{R}^+} m_j(0) \int_0^t \sigma_j e^{-\theta_j(t-s)} d\beta_j(s) dt \right] \\ &= 2 \sum_{j=1}^J g_j(x) g_j(y) \mathbb{E} [m_j(0)^2] \int_{\mathbb{R}^+} e^{-\theta_j t} dt, \end{aligned}$$

since $m_j(0)$ and β_j are independent. Thus, we get $k(x, y) = \sum_{j=1}^J \frac{\sigma_j^2}{\theta_j^2} g_j(x) g_j(y)$. Then, the covariance operator is defined by

$$Qf(x) \doteq \int_{\mathbb{T}^d} k(x, y) f(y) dy = \sum_{j=1}^J \frac{\sigma_j^2}{\theta_j^2} g_j(x) \int_{\mathbb{T}^d} g_j(y) f(y) dy.$$

In the sequel, we consider for simplicity the case $J = 1$ and write $\theta = \theta_1$, $\sigma = \sigma_1$, $g(x) = g_1(x)$ and $\beta = \beta_1$.

The rest of this section is devoted to stating the limiting equations for those two models. Note that when $\varepsilon \rightarrow 0$, the limiting equation is expressed in terms of ρ instead of f . In this case, $\rho^{\varepsilon, \delta} = \int_V f^{\varepsilon, \delta} d\mu$ converges to a function ρ or ρ^δ solution of the given limiting equation.

IV.2.1 Averaging regime

In this section, we consider the averaging regime (IV.1.4). The following results are proved in Chapter II when $\varepsilon = \delta$ and can be deduced by adapting the proofs of Chapter II when $\varepsilon \neq \delta$.

First, when $\varepsilon \rightarrow 0$, the limiting equation is given by

$$\partial_t \rho^\delta = \operatorname{div}(K \nabla \rho^\delta) + m^\delta \rho^\delta, \quad (\text{IV.2.4})$$

where m^δ is defined by (IV.1.4). It is a diffusion approximation regime in the PDE sense, since the limiting equation is a diffusion equation. Its solution ρ^δ converges to the solution of

$$\partial_t \rho = \operatorname{div}(K \nabla \rho) + \overline{m} \rho, \quad (\text{IV.2.5})$$

when $\delta \rightarrow 0$.

Second, when $\delta \rightarrow 0$, the limiting equation is given by

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon + \overline{m} f^\varepsilon. \quad (\text{IV.2.6})$$

The function $\rho^\varepsilon = \int_V f^\varepsilon d\mu$ converges to the solution of (IV.2.5) when $\varepsilon \rightarrow 0$. It is an averaging regime in a stochastic sense, since m^δ converges to its average \overline{m} .

Finally, when $\varepsilon = \delta \rightarrow 0$, the limiting equation is given by (IV.2.5) owing to Chapter II. Note that one obtains the same limiting equation in the three cases.

IV.2.2 Diffusion approximation regime

In this section, we consider the diffusion approximation regime (IV.1.5). The following results are proved in Chapter I and [AF19] when $\varepsilon = \delta$ and can be deduced by adapting the proofs of Chapter I when $\varepsilon \neq \delta$.

First, when $\varepsilon \rightarrow 0$, the limiting equation is given by

$$\partial_t \rho^\delta = \operatorname{div}(K \nabla \rho^\delta) + m^\delta \rho^\delta. \quad (\text{IV.2.7})$$

It is a diffusion approximation regime in the PDE sense, since the limiting equation is a diffusion equation. Its solution ρ^δ converges to the solution of

$$d\rho = \operatorname{div}(K \nabla \rho) dt + \rho Q^{1/2} \circ dW(t), \quad (\text{IV.2.8})$$

when $\delta \rightarrow 0$.

Second, when $\delta \rightarrow 0$, the limiting equation is given by

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon + f^\varepsilon Q^{1/2} \circ dW(t). \quad (\text{IV.2.9})$$

It is a diffusion approximation regime in the stochastic sense, since the source term converge to a stochastic diffusion term. The function $\rho^\varepsilon = \int_V f^\varepsilon d\mu$ converges to the solution of (IV.2.8) when $\varepsilon \rightarrow 0$.

Finally, when $\varepsilon = \delta \rightarrow 0$, the limiting equation is given by (IV.2.8). Note that one obtains the same limiting equation in the three cases.

IV.3 Numerical scheme: splitting procedure

The objective of this section is to give examples of AP schemes in the averaging and the diffusion approximation regime. Given a parameter $\eta = \varepsilon$, $\eta = \delta$ or $\eta = \varepsilon = \delta$, a scheme is asymptotic preserving in the regime $\eta \rightarrow 0$ if

- The scheme is consistent with the equation for fixed $\eta > 0$.
- For any fixed time-step size τ , there exists a limiting scheme when $\eta \rightarrow 0$.
- The limiting scheme is consistent with the limiting equation obtained when $\eta \rightarrow 0$.

Note that the second item already excludes standard explicit Euler or Euler-Maruyama schemes, since these schemes would impose a strong stability condition of the type $\tau \leq \tau_0(\eta)$ with $\tau_0(\eta) \xrightarrow{\eta \rightarrow 0} 0$ due to the stiff terms in (IV.1.1). To be able to choose τ independently of η , we favor implicit schemes, so that we can focus on the third item: capturing the correct limiting equation.

Since equation (IV.1.1) can be split into a deterministic part and a stochastic part, all our schemes are based on a splitting procedure in time. At each time-step:

- We start by solving approximately the deterministic part

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon^2} L f^\varepsilon. \quad (\text{IV.3.1})$$

- Then we solve approximately the stochastic part

$$\partial_t f^\delta = m^\delta f^\delta. \quad (\text{IV.3.2})$$

It is straightforward to check that the splitting scheme obtained using an AP scheme for each step is AP, in each of the three regimes $\varepsilon \rightarrow 0$, $\delta \rightarrow 0$ and $\varepsilon = \delta \rightarrow 0$. We focus on the discretization in time. The stochastic step is adapted to any discretization in space, therefore the discretization in space is chosen with respect to the deterministic step. In our case, the deterministic step is based on the Fourier transform, thus we simply consider a regular mesh of size 2^n for some $n \in \mathbb{N}_0$.

Remark IV.3.1. In the particular case $m(t, x) = m(t)$ (or equivalently $g = 1$), one can change the unknown in (IV.1.1) to solve a PDE and get rid of the stochastic part: the new unknown $\tilde{f}^{\varepsilon, \delta}(t) = \exp\left(-\int_0^t m^\delta(s) ds\right) f^{\varepsilon, \delta}(t)$ solves

$$\partial_t \tilde{f}^{\varepsilon, \delta} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x \tilde{f}^{\varepsilon, \delta} = \frac{1}{\varepsilon^2} L \tilde{f}^{\varepsilon, \delta}.$$

In this particular case, one does not need a splitting scheme, and can instead discretize independently $\tilde{f}^{\varepsilon, \delta}$ and $\exp\left(\int_0^t m^\delta(s) ds\right)$. One can also change the unknown at a discrete time level and would recover the same scheme.

IV.3.1 Asymptotic preserving scheme for the deterministic part

Let us first consider the deterministic part (IV.3.1) of (IV.1.1).

When $\varepsilon \rightarrow 0$, it is standard (see [DGP00]) that $\rho^\varepsilon = \int_V f^\varepsilon d\mu$ converges to ρ solution of the limiting equation

$$\partial_t \rho = \operatorname{div}(K \nabla \rho). \quad (\text{IV.3.3})$$

For this deterministic part, we use a fully implicit scheme written in the Fourier variable, as stated in [Hiv18, Proposition 4]. The rest of this section is devoted to detailing the construction of this scheme and to prove the AP property. Note that there exists other ways to design AP schemes. For instance, [Hiv18, LM08] gives an AP scheme based on a micro-macro decomposition and [Hiv18] an AP scheme based on an integral formulation.

For $h \in L^2(\mathbb{T}^d)$ and $k \in \mathbb{Z}^d$, let $\hat{h}_k = \int_{\mathbb{T}^d} h(x) e^{-ik \cdot x} dx$ denote the k -th coefficient of the Fourier transform of h with respect to the space variable x .

Take the Fourier transform of (IV.3.1) with respect to the space variable x : for $k \in \mathbb{Z}^d$,

$$\partial_t \hat{f}_k^\varepsilon + \frac{1}{\varepsilon} i a \cdot k \hat{f}_k^\varepsilon = \frac{1}{\varepsilon^2} (\mathcal{M} \hat{\rho}_k^\varepsilon - \hat{f}_k^\varepsilon).$$

It is then natural to introduce the implicit scheme

$$\begin{aligned} \hat{f}_{n+1, k}^\varepsilon &= \hat{f}_{n, k}^\varepsilon - \frac{\tau}{\varepsilon} i a \cdot k \hat{f}_{n+1, k}^\varepsilon + \frac{\tau}{\varepsilon^2} \mathcal{M} \hat{\rho}_{n+1, k}^\varepsilon - \frac{\tau}{\varepsilon^2} \hat{f}_{n+1, k}^\varepsilon \\ &= \frac{1 - \lambda}{1 + i \lambda \varepsilon a \cdot k} \hat{f}_{n, k}^\varepsilon + \frac{\lambda}{1 + i \lambda \varepsilon a \cdot k} \mathcal{M} \hat{\rho}_{n+1, k}^\varepsilon, \end{aligned}$$

where the parameter λ is defined by

$$\lambda \doteq \frac{\frac{\tau}{\varepsilon^2}}{1 + \frac{\tau}{\varepsilon^2}} \in (0, 1). \quad (\text{IV.3.4})$$

We deduce the following scheme for $\widehat{\rho}_{n+1,k}^\varepsilon$ by integrating with respect to v :

$$\widehat{\rho}_{n+1,k}^\varepsilon = \left\langle \widehat{f}_{n+1,k}^\varepsilon \right\rangle = (1 - \lambda) \left\langle \frac{\widehat{f}_{n,k}^\varepsilon}{1 + i\lambda\varepsilon a \cdot k} \right\rangle + \lambda \widehat{\rho}_{n+1,k}^\varepsilon \left\langle \frac{\mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle,$$

where $\langle \cdot \rangle = \int_V \cdot d\mu$.

We finally obtain the following scheme

$$\begin{cases} \widehat{f}_{n+1,k}^\varepsilon = \frac{1 - \lambda}{1 + i\lambda\varepsilon a \cdot k} \widehat{f}_{n,k}^\varepsilon + \frac{\lambda}{1 + i\lambda\varepsilon a \cdot k} \mathcal{M} \widehat{\rho}_{n+1,k}^\varepsilon \\ \widehat{\rho}_{n+1,k}^\varepsilon = \frac{(1 - \lambda) \left\langle \frac{\widehat{f}_{n,k}^\varepsilon}{1 + i\lambda\varepsilon a \cdot k} \right\rangle}{1 - \lambda \left\langle \frac{\mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle}, \end{cases} \quad (\text{IV.3.5})$$

Proposition IV.3.1. *Scheme (IV.3.5) is consistent with (IV.3.1) for fixed ε and is asymptotic preserving in the following sense: $(f_{n,k}^\varepsilon, \rho_{n,k}^\varepsilon) \xrightarrow{\varepsilon \rightarrow 0} (f_{n,k}, \rho_{n,k})$ where*

$$\begin{cases} \widehat{f}_{n+1,k} = \mathcal{M} \widehat{\rho}_{n+1,k} \\ \widehat{\rho}_{n+1,k} = \frac{\widehat{\rho}_{n,k}}{1 + \tau \langle (a \cdot k)^2 \mathcal{M} \rangle}, \end{cases} \quad (\text{IV.3.6})$$

which is consistent with (IV.3.3).

This result is not a novelty (see [Hiv18, Proposition 4]) but we prove it here for the convenience of the reader.

Proof. Let us derive the limiting scheme when $\varepsilon \rightarrow 0$. Observe that $\lambda \xrightarrow{\varepsilon \rightarrow 0} 1$, hence $\widehat{f}_{n+1,k}^\varepsilon = \mathcal{M} \widehat{\rho}_{n+1,k}^\varepsilon$. It remains to derive the limiting scheme on ρ . Using the condition $\int_V \mathcal{M}(v) d\mu(v) = 1$, (IV.3.5) can be rewritten as

$$\begin{aligned} \widehat{\rho}_{n+1,k}^\varepsilon &= \frac{(1 - \lambda) \left\langle \frac{\widehat{f}_{n,k}^\varepsilon}{1 + i\lambda\varepsilon a \cdot k} \right\rangle}{(1 - \lambda) \left\langle \frac{\mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle + 1 - \left\langle \frac{\mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle} \\ &= \frac{(1 - \lambda) \left\langle \frac{\widehat{f}_{n,k}^\varepsilon}{1 + i\lambda\varepsilon a \cdot k} \right\rangle}{(1 - \lambda) \left\langle \frac{\mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle + \left\langle \frac{i\lambda\varepsilon a \cdot k \mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle} \\ &= \frac{\left\langle \frac{\widehat{f}_{n,k}^\varepsilon}{1 + i\lambda\varepsilon a \cdot k} \right\rangle}{\left\langle \frac{\mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle + \frac{1}{1 - \lambda} \left\langle \frac{i\lambda\varepsilon a \cdot k \mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle}. \end{aligned}$$

We have

$$\left\langle \frac{i\lambda\varepsilon a \cdot k \mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle = \left\langle \frac{i\lambda\varepsilon a \cdot k \mathcal{M}}{1 + \lambda^2 \varepsilon^2 (a \cdot k)^2} \right\rangle + \left\langle \frac{\lambda^2 \varepsilon^2 (a \cdot k)^2 \mathcal{M}}{1 + \lambda^2 \varepsilon^2 (a \cdot k)^2} \right\rangle.$$

Since we assumed that V is symmetric, a is odd and \mathcal{M} is even, the first integral of the right-hand side vanishes. Owing to the identity $\frac{1}{1 - \lambda} = 1 + \frac{\tau}{\varepsilon^2}$, we get

$$\frac{1}{1 - \lambda} \left\langle \frac{i\lambda\varepsilon a \cdot k \mathcal{M}}{1 + i\lambda\varepsilon a \cdot k} \right\rangle = (\varepsilon^2 + \tau) \lambda^2 \left\langle \frac{(a \cdot k)^2 \mathcal{M}}{1 + \lambda^2 \varepsilon^2 (a \cdot k)^2} \right\rangle.$$

We have $\lambda \xrightarrow{\varepsilon \rightarrow 0} 1$, therefore

$$\frac{1}{1-\lambda} \left\langle \frac{i\lambda\varepsilon a \cdot k \mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle \xrightarrow{\varepsilon \rightarrow 0} \tau \langle (a \cdot)^2 \mathcal{M} \rangle,$$

which gives to the limiting scheme (IV.3.6) when $\varepsilon \rightarrow 0$. This scheme can be rewritten

$$\hat{\rho}_{n+1,k} = \hat{\rho}_{n,k} + \tau \hat{\rho}_{n+1,k} \langle (a \cdot k)^2 \mathcal{M} \rangle.$$

Owing to the definition (IV.1) of K , the limiting equation (IV.3.3) reads in the Fourier variable

$$\partial_t \hat{\rho}_k = \langle (a \cdot k)^2 \mathcal{M} \rangle \hat{\rho}_k.$$

The limiting scheme above is therefore the implicit schemes associated with the parabolic PDE (IV.3.3) in the Fourier variable. As a consequence, it is consistent. \square

Note that the scheme (IV.3.5) and its limiting scheme (IV.3.6) are both implicit Euler scheme for their respective equation and τ can thus be chosen independently of ε .

IV.3.2 Numerical schemes for the stochastic part

In this section, we consider the stochastic part (IV.3.2) of (IV.1.1).

The exact solution of (IV.3.2) is given by $f^\delta(t, x, v) = \exp\left(\int_0^t m^\delta(s, x) ds\right) f^\delta(0, x, v)$. This provides a first step towards an AP scheme: we discretize this exact solution using an implicit scheme to cope with the stiffness of the driving process. Therefore, we consider

$$f_{n+1}^\delta = \exp(\tau m_{n+1}^\delta) f_n^\delta.$$

It only remains to choose a discretization of m^δ such that the AP property is satisfied.

In this section, we consider a sequence $(\gamma_n)_{n \in \mathbb{N}_0}$ of independent standard normal random variables.

Numerical schemes for the stochastic part in the averaging regime

This section gives an example of AP scheme in the averaging regime (IV.1.4). To avoid a CFL condition of the form $\tau \leq \tau_0(\delta)$ where $\tau_0(\delta) \xrightarrow{\delta \rightarrow 0} 0$, we use the scheme (III.3.4) with $c(t) = \frac{1}{1+t}$.

$$\begin{cases} f_{n+1}^\delta = \exp(\tau m_{n+1}^\delta) f_n^\delta \\ m_{n+1}^\delta = \bar{m} + c\left(\frac{\tau}{\delta^2}\right) \theta(m_n^\delta - \bar{m}) + \sqrt{1 - c\left(\frac{\tau}{\delta^2}\right)^2} \sigma \gamma_n g. \end{cases} \quad (\text{IV.3.7})$$

The limiting equation of (IV.1.4) can be deduced from Chapter III and is given by

$$\partial_t f = \bar{m} f, \quad (\text{IV.3.8})$$

which has the exact solution $f(t, x, v) = \exp(t \bar{m}(x)) f(0, x, v)$. The limiting scheme of (IV.3.7) is given by

$$f_{n+1} = \exp(\tau(\bar{m} + \gamma_n)) f_n. \quad (\text{IV.3.9})$$

Note that if $J > 1$, the scheme (IV.3.7) becomes

$$\begin{cases} f_{n+1}^\delta = \exp\left(\tau \sum_{j=1}^J m_{n+1,j}^\delta g_j\right) f_n^\delta \\ m_{n+1,j}^\delta = \bar{m}_j + c \left(\frac{\tau}{\delta^2}\right) \theta_j (m_{n,j}^\delta - \bar{m}_j) + \sqrt{1 - c \left(\frac{\tau}{\delta^2}\right)^2} \sigma_j \gamma_{n,j}. \end{cases}$$

Given the exact solutions of (IV.3.2) and (IV.3.8), it is straightforward to check that the scheme is consistent and AP, since the SDE is discretized using an implicit Euler scheme and the integral in time is discretized with a right-point quadrature rule. Since the scheme is implicit, τ can be chosen independently of δ .

Numerical schemes for the stochastic part in the diffusion approximation regime

This section gives an example of AP scheme in the diffusion approximation regime (IV.1.5). To avoid a CFL condition of the form $\tau \leq \tau_0(\delta)$ where $\tau_0(\delta) \xrightarrow{\delta \rightarrow 0} 0$, we once use an implicit Euler-Maruyama scheme again

$$\begin{cases} f_{n+1}^\delta = \exp(\tau m_{n+1}^\delta) f_n^\delta \\ m_{n+1}^\delta = m_n^\delta - \theta \frac{\tau}{\delta^2} m_{n+1}^\delta + \sigma \frac{\sqrt{\tau}}{\delta^2} \gamma_n g, \end{cases} \quad (\text{IV.3.10})$$

since, owing to (IV.1.5), m^δ solves

$$dm^\delta(t) = -\frac{\theta}{\delta^2} m^\delta(t) dt + \frac{\sigma}{\delta} \frac{g}{\delta} d\beta(t).$$

Let us derive the limiting for any $J \in \mathbb{N}$. Consider $\zeta^\delta(t, x) \doteq \int_0^t m^\delta(s, x) ds$, so that the exact solution of (IV.3.2) is given by

$$f^\delta(t, x, v) = \exp(\zeta^\delta(t, x)) f^\delta(0, x, v).$$

Owing to (IV.1.3) and (IV.1.5), we have

$$\begin{aligned} d\zeta^\delta(t) &= \frac{1}{\delta} \sum_{j=1}^J g_j m_j^\delta(t) dt \\ &= -\sum_{j=1}^J \frac{1}{\theta_j} g_j \delta dm_j^\delta(t) + \sum_{j=1}^J \frac{\sigma_j}{\theta_j} g_j d\beta_j(t) \\ &\xrightarrow{\delta \rightarrow 0} \sum_{j=1}^J \frac{\sigma_j}{\theta_j} g_j d\beta_j(t). \end{aligned}$$

Thus, the solution f^δ converges to

$$\begin{aligned} f(t, x, v) &= \exp\left(\sum_{j=1}^J \frac{\sigma_j}{\theta_j} g_j \beta_j(t)\right) f(0, x, v) \\ &= \exp(Q^{1/2} W(t)) f(0, x, v), \end{aligned}$$

owing to (IV.2.3). This function solves the Stratonovich diffusion equation

$$df = fQ^{1/2} \circ dW(t), \quad (\text{IV.3.11})$$

which is therefore the limiting equation of (IV.3.2).

Let us now go back to the case $J = 1$ to simplify notation and check that (IV.3.10) is AP. By multiplying by δ^2 the second equation of (IV.3.10), we get

$$\theta\tau m_{n+1}^\delta = \delta^2(m_n^\delta - m_{n+1}^\delta) + \sigma\sqrt{\tau}\gamma_n g.$$

Therefore, the limiting scheme is given by

$$f_{n+1} = \exp\left(\frac{\sigma}{\theta} \sqrt{\tau} \gamma_n g\right) f_n, \quad (\text{IV.3.12})$$

which is consistent with the limiting equation (IV.3.11). Since the scheme is implicit, τ can be chosen independently of δ .

IV.4 Numerical experiments

The goal of this section is to illustrate the AP property or the non AP property of the different splitting schemes. We treat separately the averaging regime (IV.1.4) and the diffusion approximation regime (IV.1.5). In each regime, we recall the full expression of the splitting AP scheme (Scheme A and Scheme B, respectively restated in this section as Scheme 1 and Scheme 5). Then, in each of the three regimes $\varepsilon \ll \delta$, $\varepsilon \gg \delta$ and $\varepsilon = \delta$, we detail its limiting scheme and illustrate the asymptotic behavior of the scheme. In the diffusion approximation regime, we compare this behavior to a non AP scheme.

We present two types of illustrations:

- We first propose qualitative results: we plot the L^2 -norm $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ as a function of the time t_n as well as the final value $\rho_N^{\varepsilon,\delta}$ which approximates $\rho^{\varepsilon,\delta}(T)$ as a function of the space variable x , for a fixed value of τ and for different values of ε and δ .
- The third one is quantitative: we plot an estimation of the weak error

$$|\mathbb{E}[\varphi(\rho^{\varepsilon,\delta}(T))] - \mathbb{E}[\varphi(\rho(T))]|$$

depending on the time-step size τ , for different values of ε and δ . The expectations are approximated using a Monte-Carlo method, with 1000 samples. Investigating rates of convergence would require to substantially increase the Monte-Carlo sample size, which we have not done yet. Instead we check that, when $\tau \rightarrow 0$, $\varepsilon \rightarrow 0$ or $\delta \rightarrow 0$, the error of an AP scheme does not explode unlike the error of a non AP scheme.

Let us recall that λ is defined by $\lambda = \frac{\tau/\varepsilon^2}{1+\tau/\varepsilon^2}$ (see (IV.3.4)). Also recall that, since $J = 1$, the expression of Q simplifies (see (IV.2.2) and (IV.2.3)).

For a fixed value of τ for a qualitative result, the sequence of standard normal random variables $(\gamma_n)_{n \in [1, N]}$ used for computation is the same in a given figure.

To simplify the legend, when we use the following convention: when the scheme used is the limiting scheme obtained when $\varepsilon \rightarrow 0$ or $\delta \rightarrow 0$, we write in the legend respectively $\varepsilon = 0$ or $\delta = 0$.

The implementation of the different schemes is based on a code given by H  l  ne Hivert.

IV.4.1 Illustration in the averaging regime

In this section, we illustrate the AP property of Scheme 1 (introduced in Section IV.1 as Scheme A).

In all of our schemes, the deterministic part is still discretized using an AP scheme. Therefore, the non AP scheme is expected to asymptotically preserve in the regime $\varepsilon \rightarrow 0$ for fixed δ .

Scheme 1 AP scheme in the averaging regime.

$$\begin{aligned}
\widehat{f}_n^{\varepsilon, \delta} &= \text{FFT}(f_n^{\varepsilon, \delta}) \\
\widehat{\rho}_{n+1/2, k}^{\varepsilon, \delta} &= \frac{(1-\lambda) \left\langle \frac{\widehat{f}_{n, k}^{\varepsilon, \delta}}{1+i\lambda\varepsilon a \cdot k} \right\rangle}{1-\lambda \left\langle \frac{\mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle} \\
\widehat{f}_{n+1/2, k}^{\varepsilon, \delta} &= \frac{1-\lambda}{1+i\lambda\varepsilon a \cdot k} \widehat{f}_{n, k}^{\varepsilon, \delta} + \frac{\lambda}{1+i\lambda\varepsilon a \cdot k} \mathcal{M} \widehat{\rho}_{n+1/2, k}^{\varepsilon, \delta} \\
f_{n+1/2}^{\varepsilon, \delta} &= \text{FFT}^{-1}(\widehat{f}_{n+1/2}^{\varepsilon, \delta}) \\
m_{n+1}^\delta &= \overline{m} + \frac{1}{1+\tau/\delta^2} \theta (m_n^\delta - \overline{m}) + \sqrt{1 - \left(\frac{1}{1+\tau/\delta^2} \right)^2} \sigma \gamma_n g. \\
f_{n+1}^{\varepsilon, \delta} &= \exp(\tau m_{n+1}^\delta) f_{n+1/2}^{\varepsilon, \delta}
\end{aligned}$$

Scheme 1 is a Lie-Trotter splitting scheme based on the AP schemes (IV.3.5) for the deterministic part and (IV.3.7) for the stochastic part. Its limiting schemes are the splitting schemes obtained with the limiting schemes of (IV.3.5) and (IV.3.7).

Here, we only consider the case where the source term is linear in m^δ . In this particular case, in the averaging regime, most schemes satisfy the AP property, assuming they are stable for τ chosen independently of ε and δ . Therefore, we do not compare the AP scheme to a non AP one. Note that if the source term was non-linear in m^δ , one should be very careful when designing AP schemes (see Chapter III for AP and non AP example in the SDE case).

For the numerical experiments below, the reference is computed using an explicit finite difference method for the limiting equations (IV.2.4), (IV.2.5) and (IV.2.6). In the regime $\varepsilon \rightarrow 0$ with fixed δ , it is combined with an implicit Euler-Maruyama scheme for the stochastic component m^δ .

The coefficients of the fast component m^δ are given by $\theta = 1$ and $\sigma = 2$. The dependence in the space variable x of m^δ is given by $g(x) = 2 + \sin(2\pi x + 1)$. We fix the average value at $\overline{m} = 2g(x)$. The initial conditions are given by $m^\delta(0, x) = \tilde{m}_0 g(x)$ and $f(0, x, v) = (2 + v) \sin(2\pi x)$, with $\tilde{m}_0 = 1$ or $\tilde{m}_0 = 10$. We fix a short final time $T = 0.1$ and a large mesh size, $1/32$ or $1/64$, because of the CFL condition on the reference scheme, to reduce the number of time steps.

Recall that we consider $d = 1$, $V = \{\pm 1\}$, $\mu(\{1\}) = \mu(\{-1\}) = \frac{1}{2}$, $a(v) = v$ and $\mathcal{M}(v) = 1$. As a consequence, (IV.1) yields $K = 1$.

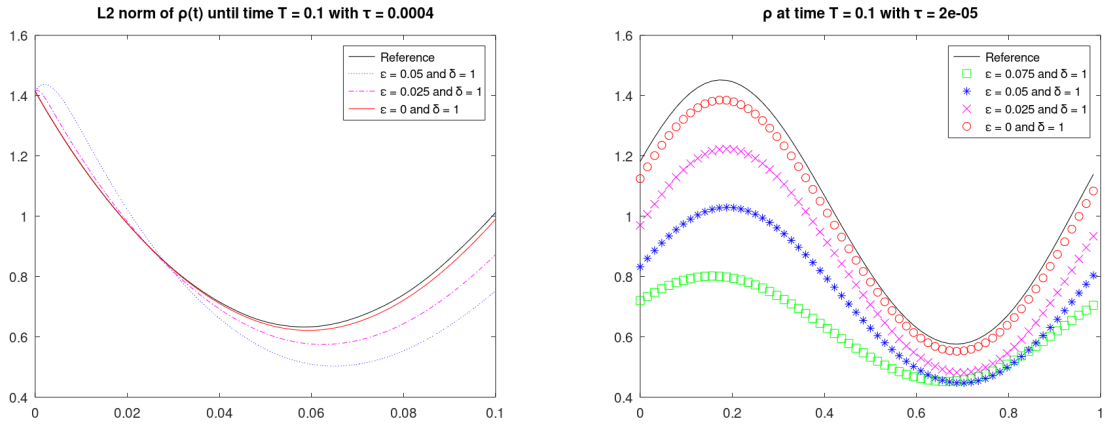
Qualitative experiments in the regime $\varepsilon \rightarrow 0$

When δ is fixed and $\varepsilon \rightarrow 0$, Scheme 1 converges to Scheme 2.

Figure IV.1 represents the evolution of $\|\rho_n^{\varepsilon, \delta}\|_{L^2}$ as a function of the time t_n (Figure IV.1a) as well as the final time $\rho_N^{\varepsilon, \delta}$ as a function of the space variable x (Figure IV.1b), both for the AP Scheme 1 and its limiting Scheme 2. It illustrates the convergence of

Scheme 2 Limiting scheme of Scheme 1 when $\varepsilon \rightarrow 0$.

$$\begin{aligned}
\hat{\rho}_n^\delta &= \text{FFT}(\rho_n) \\
\hat{\rho}_{n+1/2,k}^\delta &= \frac{\hat{\rho}_{n,k}^\delta}{1 + \tau \langle (a \cdot k)^2 \mathcal{M} \rangle} \\
\rho_{n+1/2}^\delta &= \text{FFT}^{-1}(\hat{\rho}_{n+1/2}^\delta) \\
m_{n+1}^\delta &= \bar{m} + \frac{\tau/\delta^2}{1 + \tau/\delta^2} \theta (m_n^\delta - \bar{m}) + \sqrt{1 - \left(\frac{\tau/\delta^2}{1 + \tau/\delta^2} \right)^2} \sigma \gamma_n g. \\
\rho_{n+1}^\delta &= \exp(\tau m_{n+1}^\delta) \rho_{n+1/2}^\delta
\end{aligned}$$



(a) Evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ and $\|\rho_n^\delta\|_{L^2}$ as a function of t_n with mesh size $\frac{1}{32}$. (b) Final value $\rho_N^{\varepsilon,\delta}(x)$ and $\rho_N^\delta(x)$ as a function of x with mesh size $\frac{1}{64}$.

Figure IV.1 – AP Scheme 1 and its limiting Scheme 2 in the averaging regime when $\varepsilon \rightarrow 0$ for fixed δ , with initial condition $\tilde{m}_0 = 10$.

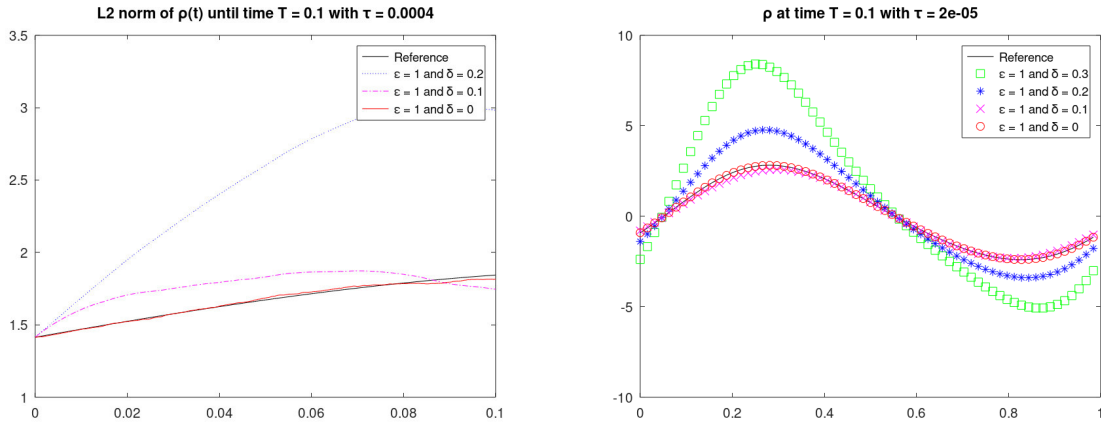
the scheme to its limiting scheme when $\varepsilon \rightarrow 0$, as well as the consistency of this limiting scheme with the solution of the limiting equation. Thus, it illustrates the AP property for this scheme when $\varepsilon \rightarrow 0$.

Qualitative experiments in the regime $\delta \rightarrow 0$

When ε is fixed and $\delta \rightarrow 0$, Scheme 1 converges to Scheme 3.

Scheme 3 Limiting scheme of Scheme 1 when $\delta \rightarrow 0$.

$$\begin{aligned}\hat{f}_n^\varepsilon &= \text{FFT}(f_n^\varepsilon) \\ \hat{\rho}_{n+1/2,k}^\varepsilon &= \frac{(1-\lambda) \left\langle \frac{\hat{f}_{n,k}^\varepsilon}{1+i\lambda\varepsilon a \cdot k} \right\rangle}{1-\lambda \left\langle \frac{\mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle} \\ \hat{f}_{n+1/2,k}^\varepsilon &= \frac{1-\lambda}{1+i\lambda\varepsilon a \cdot k} \hat{f}_{n,k}^\varepsilon + \frac{\lambda}{1+i\lambda\varepsilon a \cdot k} \mathcal{M} \hat{\rho}_{n+1/2,k}^\varepsilon \\ f_{n+1/2}^\varepsilon &= \text{FFT}^{-1}(\hat{f}_{n+1/2}^\varepsilon) \\ f_{n+1}^\varepsilon &= \exp(\tau(\bar{m} + \gamma_n g)) f_{n+1/2}^\varepsilon\end{aligned}$$



(a) Evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ and $\|\rho_n^\varepsilon\|_{L^2}$ as a function of t_n with mesh size $\frac{1}{32}$. (b) Final value $\rho_N^{\varepsilon,\delta}(x)$ and $\rho_N^\varepsilon(x)$ as a function of x with mesh size $\frac{1}{64}$.

Figure IV.2 – AP Scheme 1 and its limiting Scheme 3 in the averaging regime when $\delta \rightarrow 0$ for fixed ε , with initial condition $\tilde{m}_0 = 10$.

Figure IV.2 represents the evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ as a function of the time t_n (Figure IV.2a) as well as the final time $\rho_N^{\varepsilon,\delta}$ as a function of the space variable x (Figure IV.2b), both for the AP Scheme 1 and its limiting Scheme 3. It illustrates the convergence of the scheme to its limiting scheme when $\delta \rightarrow 0$, as well as the consistency of this limiting scheme with the solution of the limiting equation. Thus, it illustrates the AP property for this scheme when $\delta \rightarrow 0$.

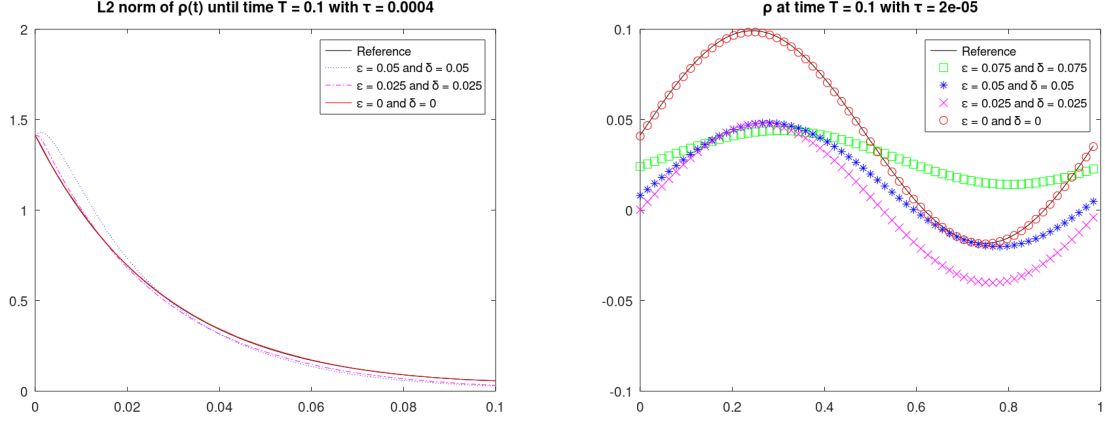
Qualitative experiments in the regime $\varepsilon = \delta \rightarrow 0$

When $\varepsilon = \delta \rightarrow 0$, Scheme 1 converges to Scheme 4.

Figure IV.3 represents the evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ as a function of the time t_n (Figure IV.3a) as well as the final time $\rho_N^{\varepsilon,\delta}$ as a function of the space variable x (IV.3b), both for

Scheme 4 Limiting scheme of Scheme 1 when $\varepsilon = \delta \rightarrow 0$.

$$\begin{aligned}\hat{\rho}_n &= \text{FFT}(\rho_n) \\ \hat{\rho}_{n+1/2,k} &= \frac{\hat{\rho}_{n,k}}{1 + \tau \langle (a \cdot k)^2 \mathcal{M} \rangle} \\ \rho_{n+1/2} &= \text{FFT}^{-1}(\hat{\rho}_{n+1/2}) \\ \rho_{n+1} &= \exp(\tau(\bar{m} + \gamma_n g)) \rho_{n+1/2}\end{aligned}$$



(a) Evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ and $\|\rho_n\|_{L^2}$ as a function of t_n with mesh size $\frac{1}{32}$. (b) Final value $\rho_N^{\varepsilon,\delta}(x)$ and $\rho_N(x)$ as a function of x with mesh size $\frac{1}{64}$.

Figure IV.3 – AP Scheme 1 and its limiting Scheme 4 in the averaging regime when $\varepsilon = \delta \rightarrow 0$, with initial condition $\tilde{m}_0 = 10$.

the AP Scheme 1 and its limiting Scheme 4. It illustrates the convergence of the scheme to its limiting scheme when $\varepsilon = \delta \rightarrow 0$, as well as the consistency of this limiting scheme with the solution of the limiting equation. Thus, it illustrates the AP property for this scheme when $\varepsilon = \delta \rightarrow 0$.

Numerical analysis of the weak error

In this section, we investigate the weak error

$$\left| \mathbb{E} \left[\varphi(\rho_N^{\varepsilon,\delta}) \right] - \mathbb{E} \left[\varphi(\rho^{\varepsilon,\delta}(T)) \right] \right|$$

of Scheme 1. We check that when $\tau \rightarrow 0$, the error goes to 0 uniformly in ε and δ .

We use a Monte-Carlo method to approximate both expectations and the same scheme with $\tau_{\text{ref}} = 2^{-12}$ to approximate $\rho^{\varepsilon,\delta}(T)$. The number of samples used in the Monte-Carlo method is 1000. Since this number is relatively small, we reduce the variance of the Monte-Carlo by using the same noise γ_n with τ and τ_{ref} .

The test function is given by $\varphi = \|\cdot\|_{L_x^2}$.

Figure IV.4 illustrates the weak error when $\tau \rightarrow 0$ for different values of ε and δ . The color and the style of the line represents the considered asymptotic ($\varepsilon \rightarrow 0$ is in blue dashed lines, $\delta \rightarrow 0$ in red dash-dotted lines and $\varepsilon = \delta \rightarrow 0$ is in black continuous lines), while the marker determines the order of magnitude of the parameter that goes to 0 (0.1 is marked by triangles, 0.01 by squares and 0.001 by circles).

One can check that the error goes to 0 when $\tau \rightarrow 0$. An interesting feature of Figure IV.4 is that the errors with $\varepsilon = 0.1$ (namely the triangle-marked black continuous line

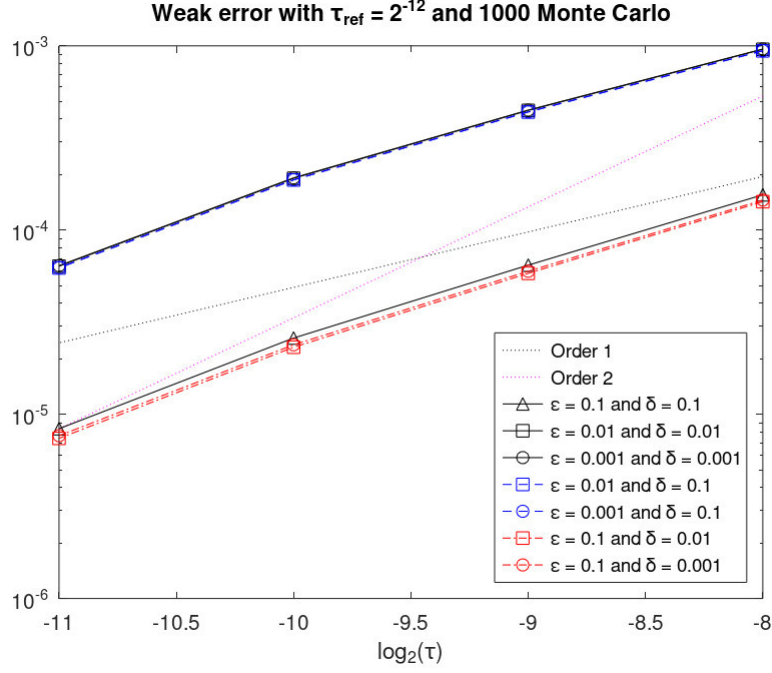


Figure IV.4 – Weak error when $\rho_N^{\epsilon, \delta}$ is computed with AP Scheme 1 in the averaging regime, with initial condition $\tilde{m}_0 = 1$ and mesh size $\frac{1}{64}$. The curves have slope 1.

and the red dash-dotted lines) have a significantly smaller error than with $\epsilon \leq 0.01$. This illustrates how the error increases when $\epsilon \rightarrow 0$, but the error is uniformly bounded in ϵ and δ , since the error does not increase when ϵ is furthermore reduced.

IV.4.2 Illustration in the diffusion approximation regime

In this section, we illustrate the AP property of Scheme 5 (introduced in Section IV.1 as Scheme B).

Scheme 5 AP scheme in the diffusion approximation regime.

$$\begin{aligned}
 \hat{f}_n^{\epsilon, \delta} &= \text{FFT}(f_n^{\epsilon, \delta}) \\
 \hat{\rho}_{n+1/2, k}^{\epsilon, \delta} &= \frac{(1-\lambda) \left\langle \frac{\hat{f}_{n, k}^{\epsilon, \delta}}{1+i\lambda\epsilon a \cdot k} \right\rangle}{1-\lambda \left\langle \frac{\mathcal{M}}{1+i\lambda\epsilon a \cdot k} \right\rangle} \\
 \hat{f}_{n+1/2, k}^{\epsilon, \delta} &= \frac{1-\lambda}{1+i\lambda\epsilon a \cdot k} \hat{f}_{n, k}^{\epsilon, \delta} + \frac{\lambda}{1+i\lambda\epsilon a \cdot k} \mathcal{M} \hat{\rho}_{n+1/2, k}^{\epsilon, \delta} \\
 f_{n+1/2}^{\epsilon, \delta} &= \text{FFT}^{-1}(\hat{f}_{n+1/2}^{\epsilon, \delta}) \\
 m_{n+1}^{\delta} &= m_n - \theta \frac{\tau}{\delta^2} m_{n+1}^{\delta} + \sigma \frac{\sqrt{\tau}}{\delta^2} \gamma_n g \\
 f_{n+1}^{\epsilon, \delta} &= \exp(\tau m \delta_{n+1}) f_{n+1/2}^{\epsilon, \delta}
 \end{aligned}$$

Scheme 5 is a Lie-Trotter splitting scheme based on the AP schemes (IV.3.5) for the deterministic part and (IV.3.10) for the stochastic part. Its limiting schemes are the splitting schemes obtained with the limiting schemes of (IV.3.5) and (IV.3.10).

We can design a non AP scheme for the stochastic part by changing the discretization

of f^δ . As for (IV.3.10), the scheme

$$\begin{cases} f_{n+1}^\delta = f_n^\delta + \tau m_{n+1}^\delta f_n^\delta \\ m_{n+1}^\delta = m_n^\delta - \theta \frac{\tau}{\delta^2} m_{n+1}^\delta + \sigma \frac{\sqrt{\tau}}{\delta^2} \gamma_n g, \end{cases} \quad (\text{IV.4.1})$$

converges to the limiting scheme

$$f^{n+1} = f^n + \frac{\sigma}{\theta} \sqrt{\tau} \gamma_n g f^n, \quad (\text{IV.4.2})$$

which is consistent with $df(t) = fQdW(t)$ with an Itô interpretation of the noise, which differs from the Stratonovich interpretation in the limiting equation (IV.3.11). The scheme (IV.4.1) is thus not AP even though it is consistent at fixed δ . We deduce Scheme 6, a non AP splitting scheme based on (IV.3.5) and (IV.4.1).

Scheme 6 Non AP Scheme in the diffusion approximation regime.

$$\begin{aligned} \widehat{f}_n^{\varepsilon,\delta} &= \text{FFT}(f_n^{\varepsilon,\delta}) \\ \widehat{\rho}_{n+1/2,k}^{\varepsilon,\delta} &= \frac{(1-\lambda) \left\langle \frac{\widehat{f}_{n,k}^{\varepsilon,\delta}}{1+i\lambda\varepsilon a \cdot k} \right\rangle}{1-\lambda \left\langle \frac{\mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle} \\ \widehat{f}_{n+1/2,k}^{\varepsilon,\delta} &= \frac{1-\lambda}{1+i\lambda\varepsilon a \cdot k} \widehat{f}_{n,k}^{\varepsilon,\delta} + \frac{\lambda}{1+i\lambda\varepsilon a \cdot k} \mathcal{M} \widehat{\rho}_{n+1/2,k}^{\varepsilon,\delta} \\ f_{n+1/2}^{\varepsilon,\delta} &= \text{FFT}^{-1}(\widehat{f}_{n+1/2}^{\varepsilon,\delta}) \\ m_{n+1}^\delta &= m_n^\delta - \theta \frac{\tau}{\delta^2} m_{n+1}^\delta + \sigma \frac{\sqrt{\tau}}{\delta^2} \gamma_n g \\ f_{n+1}^{\varepsilon,\delta} &= f_{n+1/2}^{\varepsilon,\delta} + f_{n+1/2}^{\varepsilon,\delta} \tau m_{n+1}^\delta \end{aligned}$$

For the numerical experiments below, the reference scheme for the limiting equations (IV.2.7), (IV.2.8) and (IV.2.9) is obtained using finite differences to discretize $\text{div}(K\nabla\rho) = \partial_x^2 \rho$ and an explicit Euler-Maruyama scheme to discretize $\rho Q^{1/2} \circ dW(t)$ expressed in Itô form (see (IV.2.3)).

For the numerical experiments below, the coefficients of the fast components m^δ are given by $\theta = 1$ and $\sigma = 2$. The dependence in the space variable x of m^δ is given by $g(x) = 2 + \sin(2\pi x + 1)$. The initial conditions are given by $m^\delta(0, x) = \tilde{m}_0 g(x)$ and $f(0, x, v) = (2 + v) \sin(2\pi x)$, with $\tilde{m}_0 = 1$ or $\tilde{m}_0 = 10$. We fix a short final time $T = 0.1$ and a large mesh size, $1/32$ or $1/64$, because of the CFL condition on the reference scheme, to avoid a too large number of time step.

Recall that we consider $d = 1$, $V = \{\pm 1\}$, $\mu(\{1\}) = \mu(\{-1\}) = \frac{1}{2}$, $a(v) = v$ and $\mathcal{M}(v) = 1$. As a consequence, (IV.1) yields $K = 1$.

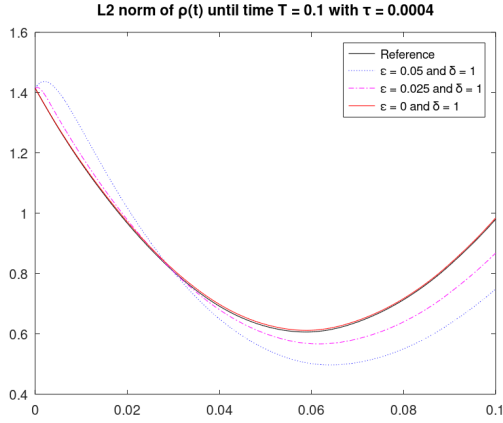
Qualitative experiments in the regime $\varepsilon \rightarrow 0$

When δ is fixed and $\varepsilon \rightarrow 0$, Scheme 5 converges to Scheme 7.

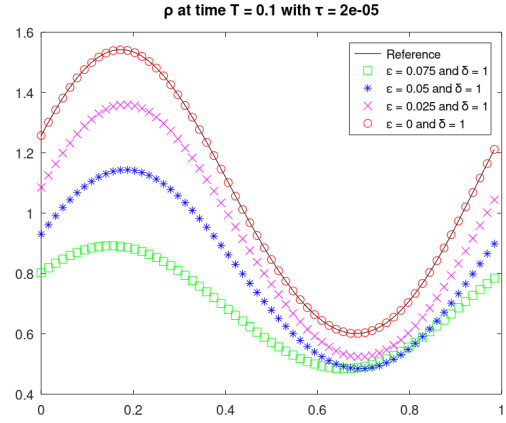
Figure IV.5 represents the evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ as a function of the time t_n (Figure IV.5a) as well as the final time $\rho_N^{\varepsilon,\delta}$ as a function of the space variable x (Figure IV.5b), both for the AP Scheme 5 and its limiting Scheme 7. It illustrates the convergence of the scheme to its limiting scheme when $\varepsilon \rightarrow 0$, as well as the consistency of this limiting scheme with the solution of the limiting equation. Thus, it illustrates the AP property for this scheme when $\varepsilon \rightarrow 0$. Recall that since δ is fixed and the deterministic part is AP,

Scheme 7 Limiting scheme of Scheme 5 when $\varepsilon \rightarrow 0$.

$$\begin{aligned}
\hat{\rho}_n^\delta &= \text{FFT}(\rho_n^\delta) \\
\hat{\rho}_{n+1/2,k}^\delta &= \frac{\hat{\rho}_{n,k}^\delta}{1 + \tau \langle (a \cdot k)^2 \mathcal{M} \rangle} \\
\rho_{n+1/2}^\delta &= \text{FFT}^{-1}(\hat{\rho}_{n+1/2}^\delta) \\
m_{n+1}^\delta &= m_n^\delta - \theta \frac{\tau}{\delta^2} m_{n+1}^\delta + \sigma \frac{\sqrt{\tau}}{\delta^2} \gamma_n g \\
\rho_{n+1}^\delta &= \exp(\tau m_{n+1}^\delta) \rho_{n+1/2}^\delta
\end{aligned}$$



(a) Evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ and $\|\rho_n^\delta\|_{L^2}$ as a function of t_n with mesh size $\frac{1}{32}$.



(b) Final value $\rho_N^{\varepsilon,\delta}(x)$ and $\rho_N^\delta(x)$ as a function of x with mesh size $\frac{1}{64}$.

Figure IV.5 – AP Scheme 5 and its limiting Scheme 7 in the diffusion approximation regime when $\varepsilon \rightarrow 0$ for fixed δ , with initial condition $\tilde{m}_0 = 10$.

we expect this convergence for any consistent scheme for the stochastic part, and the AP (or non AP) property of the stochastic part does not come into play. Hence, we do not illustrate the same property for the non AP Scheme 6.

Qualitative experiments in the regime $\delta \rightarrow 0$

When ε is fixed and $\delta \rightarrow 0$, AP Scheme 5 converges to Scheme 8 and non AP Scheme 6

Scheme 8 Limiting scheme of the AP Scheme 5 when $\delta \rightarrow 0$.

$$\begin{aligned}\hat{f}_n^\varepsilon &= \text{FFT}(f_n^\varepsilon) \\ \hat{\rho}_{n+1/2,k}^\varepsilon &= \frac{(1-\lambda) \left\langle \frac{\hat{f}_{n,k}^\varepsilon}{1+i\lambda\varepsilon a \cdot k} \right\rangle}{1-\lambda \left\langle \frac{\mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle} \\ \hat{f}_{n+1/2,k}^\varepsilon &= \frac{1-\lambda}{1+i\lambda\varepsilon a \cdot k} \hat{f}_{n,k}^\varepsilon + \frac{\lambda}{1+i\lambda\varepsilon a \cdot k} \mathcal{M} \hat{\rho}_{n+1/2,k}^\varepsilon \\ f_{n+1/2}^\varepsilon &= \text{FFT}^{-1}(\hat{f}_{n+1/2}^\varepsilon) \\ f_{n+1}^\varepsilon &= \exp\left(\frac{\sigma}{\theta} \sqrt{\tau} \gamma_n g\right) f_{n+1/2}^\varepsilon\end{aligned}$$

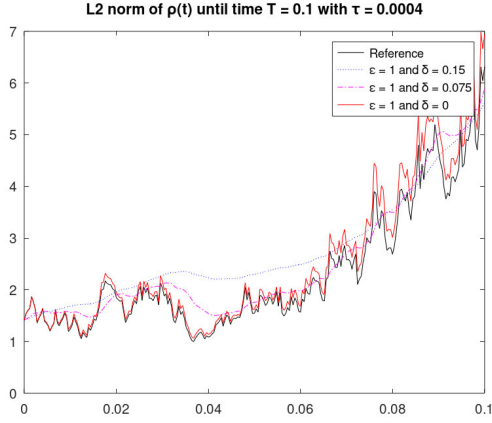
converges to Scheme 9.

Scheme 9 Limiting scheme of the non AP Scheme 6 when $\delta \rightarrow 0$.

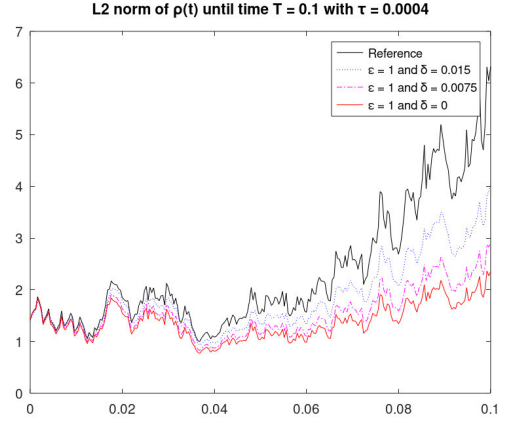
$$\begin{aligned}\hat{f}_n^\varepsilon &= \text{FFT}(f_n^\varepsilon) \\ \hat{\rho}_{n+1/2,k}^\varepsilon &= \frac{(1-\lambda) \left\langle \frac{\hat{f}_{n,k}^\varepsilon}{1+i\lambda\varepsilon a \cdot k} \right\rangle}{1-\lambda \left\langle \frac{\mathcal{M}}{1+i\lambda\varepsilon a \cdot k} \right\rangle} \\ \hat{f}_{n+1/2,k}^\varepsilon &= \frac{1-\lambda}{1+i\lambda\varepsilon a \cdot k} \hat{f}_{n,k}^\varepsilon + \frac{\lambda}{1+i\lambda\varepsilon a \cdot k} \mathcal{M} \hat{\rho}_{n+1/2,k}^\varepsilon \\ f_{n+1/2}^\varepsilon &= \text{FFT}^{-1}(\hat{f}_{n+1/2}^\varepsilon) \\ f_{n+1}^\varepsilon &= f_{n+1/2}^\varepsilon + f_{n+1/2}^\varepsilon \frac{\sigma}{\theta} \sqrt{\tau} \gamma_n g\end{aligned}$$

Figure IV.6 represents the evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ as a function of the time t_n for AP Scheme 5 and its limiting Scheme 8 (Figure IV.6a) and for non AP Scheme 6 and its limiting Scheme 9 (Figure IV.6b). In both cases, it illustrates the convergence of the scheme to its limiting scheme when $\delta \rightarrow 0$. Figure IV.7a also illustrates the consistency of the limiting Scheme 8 with the limiting equation, and therefore the AP property for Scheme 5 in this regime. On Figure IV.7b, the limiting Scheme 9 is not consistent with the limiting equation, illustrating that Scheme 6 fails to capture the correct noise in the limiting equation (IV.2.9), and is therefore not AP.

Figure IV.7 represents the value of the solution at the final time $\rho_N^{\varepsilon,\delta}(x)$ as a function of the space variable x for the AP Scheme 5 and its limiting Scheme 8 (Figure IV.7a) and for non AP Scheme 6 and its limiting Scheme 9 (Figure IV.7b). Like Figure IV.6, it illustrates the convergence of the schemes to their limiting schemes when $\delta \rightarrow 0$, as well as the AP property of Scheme 5 and the non AP property of Scheme 6. In Figure IV.7a, the small difference between the reference scheme and the limiting scheme is due to the step-size τ . Figure IV.7b illustrates once again how Scheme 6 fails to capture the correct interpretation of the noise in the limiting equation (IV.2.9).

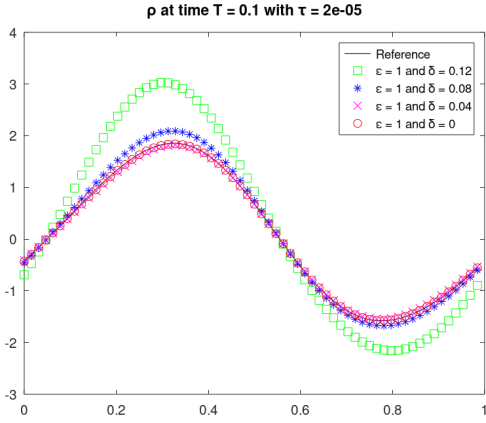


(a) AP Scheme 5.

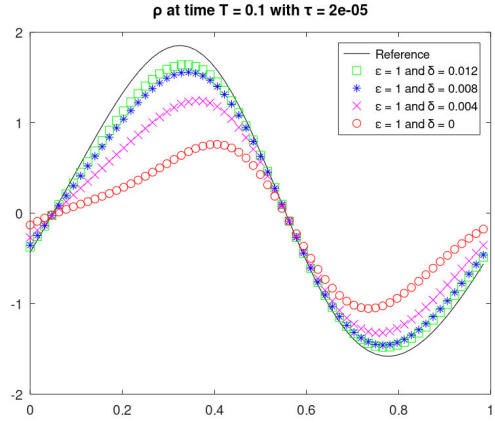


(b) Non AP Scheme 6.

Figure IV.6 – Evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ and $\|\rho_n^\varepsilon\|_{L^2}$ for AP and non AP schemes in the diffusion approximation regime as a function of t_n when $\delta \rightarrow 0$ for fixed ε , with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{32}$.



(a) AP Scheme 5.



(b) Non AP Scheme 6.

Figure IV.7 – Final value $\rho_N^{\varepsilon,\delta}(x)$ and $\rho_N^\varepsilon(x)$ for AP and non AP schemes in the diffusion approximation regime as a function of x when $\delta \rightarrow 0$ for fixed ε , with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{64}$.

Qualitative experiments in the regime $\varepsilon = \delta \rightarrow 0$

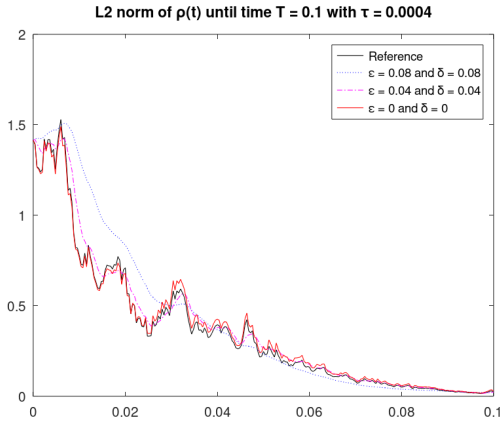
When $\varepsilon = \delta \rightarrow 0$, AP Scheme 5 converges to Scheme 10 and non AP Scheme 6 converges to Scheme 11.

Scheme 10 Limiting scheme of the AP Scheme 5 when $\varepsilon = \delta \rightarrow 0$.

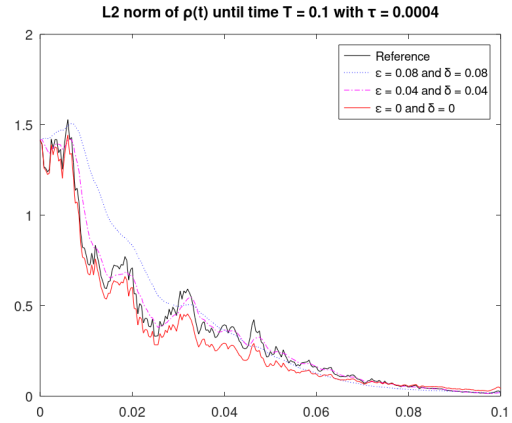
$$\begin{aligned}\hat{\rho}_n &= \text{FFT}(\rho_n) \\ \hat{\rho}_{n+1/2,k} &= \frac{\hat{\rho}_{n,k}}{1 + \tau \langle (a \cdot k)^2 \mathcal{M} \rangle} \\ \rho_{n+1/2} &= \text{FFT}^{-1}(\hat{\rho}_{n+1/2}) \\ \rho_{n+1} &= \exp\left(\frac{\sigma}{\theta} \sqrt{\tau} \gamma_n g\right) \rho_{n+1/2}\end{aligned}$$

Scheme 11 Limiting scheme of the non AP Scheme 6 when $\varepsilon = \delta \rightarrow 0$.

$$\begin{aligned}\hat{\rho}_n &= \text{FFT}(\rho_n) \\ \hat{\rho}_{n+1/2,k} &= \frac{\hat{\rho}_{n,k}}{1 + \tau \langle (a \cdot k)^2 \mathcal{M} \rangle} \\ \rho_{n+1/2} &= \text{FFT}^{-1}(\hat{\rho}_{n+1/2}) \\ \rho_{n+1} &= \rho_{n+1/2} + \rho_{n+1/2} \frac{\sigma}{\theta} \sqrt{\tau} \gamma_n g\end{aligned}$$



(a) AP Scheme 5.

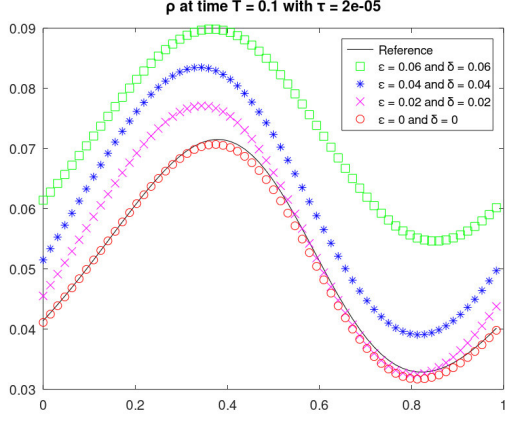


(b) Non AP Scheme 6.

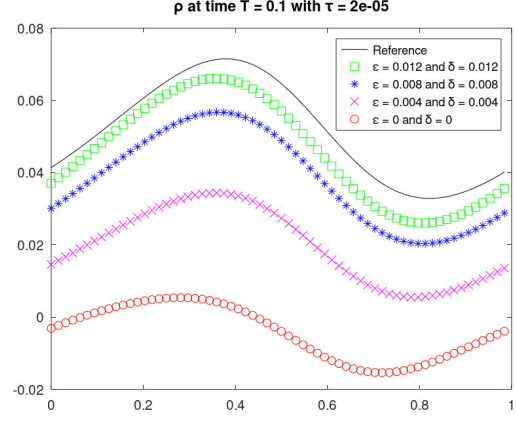
Figure IV.8 – Evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ and $\|\rho_n\|_{L^2}$ for AP and non AP schemes in the diffusion approximation regime as a function of t_n when $\varepsilon = \delta \rightarrow 0$, with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{32}$.

Figure IV.8 represents the evolution of $\|\rho_n^{\varepsilon,\delta}\|_{L^2}$ as a function of the time t_n for AP Scheme 5 and its limiting Scheme 10 (Figure IV.8a) and for non AP Scheme 6 and its limiting Scheme 11 (Figure IV.8b). In both case, it illustrates the convergence of the scheme to its limiting scheme when $\varepsilon = \delta \rightarrow 0$. Figure IV.8a illustrates the consistency of the limiting Scheme 10 with the limiting equation, and therefore the AP property for Scheme 5 in this regime. On Figure IV.8b, the limiting Scheme 11 is not consistent with the limiting equation, illustrating that Scheme 6 fails to capture the correct noise in the limiting equation (IV.2.8).

Figure IV.9 represents the final time $\rho_N^{\varepsilon,\delta}$ as a function of the space variable x for AP Scheme 5 and its limiting Scheme 10 (Figure IV.9a) and for non AP Scheme 6 and



(a) AP Scheme 5.



(b) Non AP Scheme 6.

Figure IV.9 – Final value $\rho_N^{\varepsilon,\delta}(x)$ and $\rho_N(x)$ for AP and non AP schemes in the diffusion approximation regime as a function of x when $\varepsilon = \delta \rightarrow 0$, with initial condition $\tilde{m}_0 = 10$ and mesh size $\frac{1}{64}$.

its limiting Scheme 11 (Figure IV.9b). Once again, it illustrates the convergence of the schemes to their limiting schemes when $\delta \rightarrow 0$, as well as the AP property of Scheme 5 and the non AP property of Scheme 6, since the reference value is computed using a reference scheme with the same time-step size. As in Figure IV.9a, the small difference between the reference scheme and the limiting scheme is due to the step-size τ . Figure IV.9b illustrates how Scheme 6 fails to capture the correct noise in the limiting equation (IV.2.8) and is therefore not AP.

Numerical analysis of the weak error

In this section, we investigate the weak error

$$\left| \mathbb{E} \left[\varphi(\rho_N^{\varepsilon,\delta}) \right] - \mathbb{E} \left[\varphi(\rho^{\varepsilon,\delta}(T)) \right] \right|$$

of Schemes 5 and 6.

We use a Monte-Carlo method to approximate both expectations and the aforementioned reference scheme with $\tau_{\text{ref}} = 2^{-12}$ to approximate $\rho^{\varepsilon,\delta}(T)$. The number of samples used in the Monte-Carlo method is 1000. Since this number is relatively small, we reduce the variance of the Monte-Carlo by using the same noise γ_n with τ and τ_{ref} .

The test function is given by $\varphi = \|\cdot\|_{L_x^2}$.

Figure IV.10 illustrates the weak error when $\tau \rightarrow 0$ for different values of ε and δ for both the AP and the non AP scheme. The color and the style of the line represents the considered asymptotic ($\varepsilon \rightarrow 0$ is in blue dashed lines, $\delta \rightarrow 0$ in red dash-dotted lines and $\varepsilon = \delta \rightarrow 0$ is in black continuous lines), while the marker determines the order of magnitude of the parameter that goes to 0 (0.1 is marked by triangles, 0.01 by squares and 0.001 by circles).

One can check that the error goes to 0 when $\tau \rightarrow 0$. The error in the AP scheme varies with the parameters but is uniform with respect to ε and δ . However, the error in the non AP scheme heavily depends on the parameters, which illustrates the non AP property. Note that in the regime $\varepsilon \rightarrow 0$ for fixed $\delta = 0.1$ (corresponding to the triangle-marked black

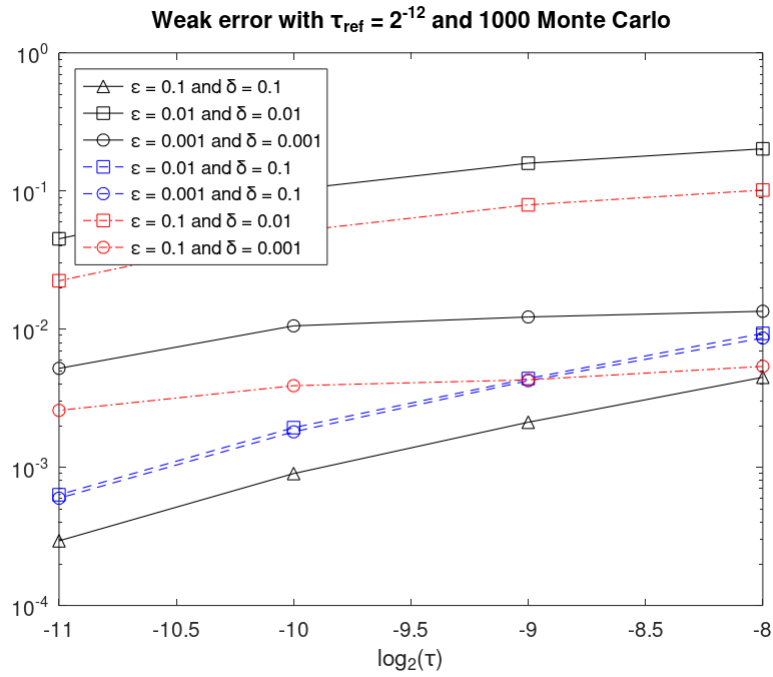
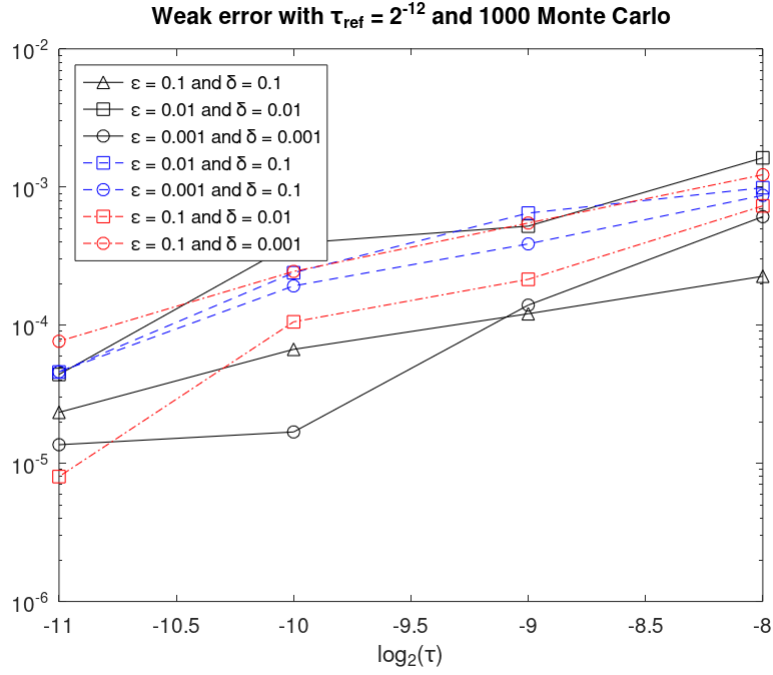


Figure IV.10 – Weak error when $\rho_N^{\varepsilon, \delta}$ is computed with AP Scheme 5 or non AP Scheme 6 in the diffusion approximation regime, with initial condition $\tilde{m}_0 = 1$ and mesh size $\frac{1}{64}$.

continuous line and the blue dashed lines), the error is uniform in ε . This illustrates that in this regime, the scheme is AP. Indeed, only the deterministic part (IV.3.1) converges and the deterministic step in the splitting non AP scheme (6) is AP.

IV.5 Conclusion

In this chapter, we have studied a general notion of asymptotic preserving schemes, related to convergence in distribution or in probability, for a class of SPDE systems in averaging and diffusion approximation regimes. We have illustrated how the splitting procedure allows to design an AP scheme for an SPDE that is sum of a PDE part and a SDE part, starting from AP schemes for each of these parts. Even though we focused on an implicit scheme written in the Fourier space for the PDE part, we are confident that numerical experiment would have the same result if one uses another AP scheme (given for instance in [Hiv18]).

A limitation of our study is the fact that the fast component is an Ornstein-Uhlenbeck process: even if the general theory of AP schemes described in Section III.3.1 holds in more general settings, the construction of implementable AP schemes (such as the ones described in Sections III.3.2 and III.3.3) is not straightforward if for instance the fast component is solution of a general ergodic SDE with nonlinear coefficients.

It would also be worth investigating the case where the fast component depends on the slow component, for example if $g_j = g_j(\rho(t, x))$ or if the parameters θ_j , σ_j and \bar{m}_j depend on $\rho(t, \cdot)$.

We have also left open the question of obtaining the error estimates as in the SDE case presented in Chapter III, as well as a more precise form of stability, for instance by getting L^2 moments uniformly with respect to the time-step size. Studying the convergence rate from a numerical point of view requires more efficient implementations of our schemes.

Finally, one could investigate the case when the source term $\delta^{-\kappa} m^\delta f^{\varepsilon, \delta}$ is replaced by a forcing term $\delta^{-\kappa'} m^\delta \cdot \nabla_v f^{\varepsilon, \delta}$, where we expect $\kappa' = 2$ in the diffusion approximation regime and $\kappa' = 1$ in the averaging regime. This model has already been studied from a theoretical point of view [DV20], but to the best of our knowledge, there exists no AP schemes for such kinetic stochastic PDEs. A splitting-based approach similar to what we did in this chapter should give good results.

Bibliography

- [AEEVE12] Assyr Abdulle, Weinan E, Björn Engquist, and Eric Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012.
- [AF19] Nathalie Ayi and Erwan Faou. Analysis of an asymptotic preserving scheme for stochastic linear kinetic equations in the diffusion limit. *SIAM/ASA J. Uncertain. Quantif.*, 7(2):760–785, 2019.
- [APV17] Assyr Abdulle, Grigorios A. Pavliotis, and Urbain Vaes. Spectral methods for multiscale stochastic differential equations. *SIAM/ASA J. Uncertain. Quantif.*, 5(1):720–761, 2017.
- [BD99] François Bouchut and Laurent Desvillettes. Averaging lemmas without time Fourier transform and application to discretized kinetic equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 129(1):19–36, 1999.
- [Bil99] 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.
- [BLP79] Alain Bensoussan, Jacques-L. Lions, and George C. Papanicolaou. Boundary layers and homogenization of transport processes. *Publ. Res. Inst. Math. Sci.*, 15(1):53–157, 1979.
- [Bou04] Nicolas Bourbaki. *Integration. II. Chapters 7–9*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2004. Translated from the 1963 and 1969 French originals by Sterling K. Berberian.
- [Bré12] Charles-Edouard Bréhier. Strong and weak orders in averaging for SPDEs. *Stochastic Processes and their Applications*, 122(7):2553–2593, 2012.
- [Bré13] Charles-Edouard Bréhier. Analysis of an HMM time-discretization scheme for a system of stochastic PDEs. *SIAM J. Numer. Anal.*, 51(2):1185–1210, 2013.
- [Bré20] Charles-Edouard Bréhier. Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component. *Stochastic Process. Appl.*, 130(6):3325–3368, 2020.
- [BRR20] Charles-Edouard Bréhier and Shmuel Rakotonirina-Ricquebourg. On asymptotic preserving schemes for a class of stochastic differential equations in averaging and diffusion approximation regimes. arXiv preprint arXiv:2011.02341, 2020.

- [Cer01] Sandra Cerrai. *Second order PDE's in finite and infinite dimension*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001. A probabilistic approach.
- [Cer09] Sandra Cerrai. A Khasminskii type averaging principle for stochastic reaction-diffusion equations. *The Annals of Applied Probability*, 19(3):899–948, 2009.
- [CF09] Sandra Cerrai and Mark Freidlin. Averaging principle for a class of stochastic reaction-diffusion equations. *Probability theory and related fields*, 144(1-2):137–177, 2009.
- [CX20] Sandra Cerrai and Guangyu Xi. A Smoluchowski-Kramers approximation for an infinite dimensional system with state-dependent damping. arXiv preprint arXiv:2011.14236, 2020.
- [dBD10] Anne de Bouard and Arnaud Debussche. The nonlinear Schrödinger equation with white noise dispersion. *J. Funct. Anal.*, 259(5):1300–1321, 2010.
- [dBG12] Anne de Bouard and Maxime Gazeau. A diffusion approximation theorem for a nonlinear PDE with application to random birefringent optical fibers. *Ann. Appl. Probab.*, 22(6):2460–2504, 2012.
- [DGP00] Pierre Degond, Thierry Goudon, and Frédéric Poupaud. Diffusion limit for nonhomogeneous and non-micro-reversible processes. *Indiana Univ. Math. J.*, 49(3):1175–1198, 2000.
- [DM16] Romain Duboscq and Renaud Marty. Analysis of a splitting scheme for a class of random nonlinear partial differential equations. *ESAIM Probab. Stat.*, 20:572–589, 2016.
- [DP14] Giacomo Dimarco and Lorenzo Pareschi. Numerical methods for kinetic equations. *Acta Numer.*, 23:369–520, 2014.
- [DPS18] Giacomo Dimarco, Lorenzo Pareschi, and Giovanni Samaey. Asymptotic-preserving Monte Carlo methods for transport equations in the diffusive limit. *SIAM J. Sci. Comput.*, 40(1):A504–A528, 2018.
- [DPZ14] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.
- [DRV20] Arnaud Debussche, Angelo Rosello, and Julien Vovelle. Diffusion-approximation for a kinetic spray-like system with random forcing. arXiv preprint arXiv:2005.09374, 2020.
- [DS64] Nelson Dunford and Jacob T. Schwartz. *Linear Operators: Spectral Theory*, volume 2. Interscience Publishers, 1964.
- [DT11] Arnaud Debussche and Yoshio Tsutsumi. 1D quintic nonlinear Schrödinger equation with white noise dispersion. *J. Math. Pures Appl. (9)*, 96(4):363–376, 2011.

- [DV12] Arnaud Debussche and Julien Vovelle. Diffusion limit for a stochastic kinetic problem. *Commun. Pure Appl. Anal.*, 11(6):2305–2326, 2012.
- [DV20] Arnaud Debussche and Julien Vovelle. Diffusion-approximation in stochastically forced kinetic equations. *Tunisian Journal of Mathematics*, 3(1):1–53, 2020.
- [E11] Weinan E. *Principles of multiscale modeling*. Cambridge University Press, Cambridge, 2011.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [ELVE05] Weinan E, Di Liu, and Eric Vanden-Eijnden. Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.*, 58(11):1544–1585, 2005.
- [Eva89] Lawrence C. Evans. The perturbed test function method for viscosity solutions of nonlinear PDE. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 111(3-4):359–375, 1989.
- [FG18] Jason Frank and Georg A. Gottwald. A note on statistical consistency of numerical integrators for multiscale dynamics. *Multiscale Model. Simul.*, 16(2):1017–1033, 2018.
- [FGPSI07] Jean-Pierre Fouque, Josselin Garnier, George C. Papanicolaou, and Knut Sørensen. *Wave propagation and time reversal in randomly layered media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.
- [FM09] José C. Ferreira and Valdir A. Menegatto. Eigenvalues of integral operators defined by smooth positive definite kernels. *Integral Equations and Operator Theory*, 64(1):61–81, 2009.
- [GKK06] Dror Givon, Ioannis G. Kevrekidis, and Raz Kupferman. Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems. *Commun. Math. Sci.*, 4(4):707–729, 2006.
- [GP92] François Golse and Frédéric Poupaud. Limite fluide des équations de Boltzmann des semi-conducteurs pour une statistique de Fermi-Dirac. *Asymptotic Anal.*, 6(2):135–160, 1992.
- [GR09] Thierry Goudon and Mathias Rousset. Stochastic acceleration in an inhomogeneous time random force field. *Applied Mathematics Research Express*, 2009(1):1–46, 2009.
- [Hiv18] Hélène Hivert. Numerical schemes for kinetic equation with diffusion limit and anomalous time scale. *Kinetic & Related Models*, 11(2):409, 2018.

- [HJ17] Jingwei Hu and Shi Jin. Uncertainty quantification for kinetic equations. In *Uncertainty quantification for hyperbolic and kinetic equations*, volume 14 of *SEMA SIMAI Springer Ser.*, pages 193–229. Springer, Cham, 2017.
- [HJL17] Jinwei Hu, Shi Jin, and Qin Li. Asymptotic-preserving schemes for multiscale hyperbolic and kinetic equations. In *Handbook of numerical methods for hyperbolic problems*, volume 18 of *Handb. Numer. Anal.*, pages 103–129. Elsevier/North-Holland, Amsterdam, 2017.
- [HS12] Martina Hofmanová and Jan Seidler. On weak solutions of stochastic differential equations. *Stoch. Anal. Appl.*, 30(1):100–121, 2012.
- [HW96] Ernst Hairer and Gerhard Wanner. *Stiff and Differential-Algebraic Problems*, volume 14. Springer, 1996.
- [Jak86] Adam Jakubowski. On the Skorokhod topology. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(3):263–285, 1986.
- [Jin99] Shi Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454, 1999.
- [Jin12] Shi Jin. Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review. *Riv. Math. Univ. Parma (N.S.)*, 3(2):177–216, 2012.
- [Jin18] Shi Jin. Mathematical analysis and numerical methods for multiscale kinetic equations with uncertainties. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pages 3611–3639. World Sci. Publ., Hackensack, NJ, 2018.
- [JLP18] Shi Jin, Hanqing Lu, and Lorenzo Pareschi. Efficient stochastic asymptotic-preserving implicit-explicit methods for transport equations with diffusive scalings and random inputs. *SIAM J. Sci. Comput.*, 40(2):A671–A696, 2018.
- [JS03] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.
- [KGH⁺03] Ioannis G. Kevrekidis, C. William Gear, James M. Hyman, Panagiotis G. Kevrekidis, Olof Runborg, and Constantinos Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci.*, 1(4):715–762, 2003.
- [Kha66a] Rafail Z. Khasminskii. A limit theorem for solutions of differential equations with a random right hand part. *Teor. Veroyatnost. i Primenen*, 11:444–462, 1966.
- [Kha66b] Rafail Z. Khasminskii. Stochastic processes defined by differential equations with a small parameter. *Teor. Veroyatnost. i Primenen*, 11:240–259, 1966.

- [Kha68] Rafail Z. Khasminskii. On the principle of averaging the Itô's stochastic differential equations. *Kibernetika* 4, pages 260–279, 1968.
- [KP92] Peter E. Kloeden and Eckhard Platen. *Stochastic Differential Equations*, pages 103–160. Springer Berlin Heidelberg, Berlin, Heidelberg, 1992.
- [Kue15] Christian Kuehn. *Multiple time scale dynamics*, volume 191 of *Applied Mathematical Sciences*. Springer, Cham, 2015.
- [Kus84] Harold J. Kushner. *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory*, volume 6 of *MIT Press Series in Signal Processing, Optimization, and Control*. MIT Press, Cambridge, MA, 1984.
- [KY05] Rafail Z. Khasminskii and George Yin. Limit behavior of two-time-scale diffusions revisited. *J. Differential Equations*, 212(1):85–113, 2005.
- [LAE08] Tiejun Li, Assyr Abdule, and Weinan E. Effectiveness of implicit methods for stiff stochastic differential equations. *Commun. Comput. Phys.*, 3(2):295–307, 2008.
- [LBL20] Guillaume Laibe, Charles-Edouard Bréhier, and Maxime Lombart. On the settling of small grains in dusty discs: analysis and formulae. *Monthly Notices of the Royal Astronomical Society*, 494(4):5134–5147, 05 2020.
- [LK74] Edward W. Larsen and Joseph B. Keller. Asymptotic solution of neutron transport problems for small mean free paths. *J. Mathematical Phys.*, 15:75–81, 1974.
- [LLMS20] Frédéric Legoll, Tony Lelièvre, Keith Myerscough, and Giovanni Samaey. Parareal computation of stochastic differential equations with time-scale separation: a numerical convergence study. *Comput. Vis. Sci.*, 23(1-4):9, 2020.
- [LM08] Mohammed Lemou and Luc Mieussens. A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit. *SIAM Journal on Scientific Computing*, 31(1):334–368, 2008.
- [LT97] Pierre-Louis Lions and Giuseppe Toscani. Diffusive limit for finite velocity Boltzmann kinetic models. *Rev. Mat. Iberoamericana*, 13(3):473–513, 1997.
- [Mar06] Renaud Marty. On a splitting scheme for the nonlinear Schrödinger equation in a random medium. *Commun. Math. Sci.*, 4(4):679–705, 2006.
- [PP03] Étienne Pardoux and Andrey L. Piatnitski. Homogenization of a nonlinear random parabolic partial differential equation. *Stochastic Process. Appl.*, 104(1):1–27, 2003.
- [PS08] Grigorios A. Pavliotis and Andrew M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.

- [PSV77] George C. Papanicolaou, Daniel W. Stroock, and S. R. Srinivasa Varadhan. Martingale approach to some limit theorems. *Duke Univ. Math. Ser., Vol. III*, pages ii+120, 1977.
- [PSZ09] Grigorios A. Pavliotis, Andrew M. Stuart, and Konstantinos C. Zygalakis. Calculating effective diffusiveness in the limit of vanishing molecular diffusion. *J. Comput. Phys.*, 228(4):1030–1055, 2009.
- [Pup19] Gabriella Puppo. Kinetic models of BGK type and their numerical integration. arXiv preprint arXiv:1902.08311, 2019.
- [PV03] Frédéric Poupaud and Alexis Vasseur. Classical and quantum transport in random media. *Journal de mathématiques pures et appliquées*, 82(6):711–748, 2003.
- [RLJ14] Wei Ren, Hong Liu, and Shi Jin. An asymptotic-preserving Monte Carlo method for the Boltzmann equation. *J. Comput. Phys.*, 276:380–404, 2014.
- [RR20] Shmuel Rakotonirina-Ricquebourg. Diffusion limit for a stochastic kinetic problem with unbounded driving process. arXiv preprint arXiv:2009.10406, 2020.
- [RSX19] Michael Röckner, Xiaobin Sun, and Longjie Xie. Strong and weak convergence in the averaging principle for SDEs with Hölder coefficients. arXiv preprint arXiv:1907.09256, 2019.
- [Sim87] Jacques Simon. Compact sets in the space $L^p(0, T; B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.
- [SV06] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
- [VZS20] Hannes Vandecasteele, Przemysław Zieliński, and Giovanni Samaey. Efficiency of a micro-macro acceleration method for scale-separated stochastic differential equations. *Multiscale Model. Simul.*, 18(3):1272–1298, 2020.

Theoretical and numerical study of multiscale stochastic kinetic equations

Abstract: In this thesis, we study a class of slow-fast systems modeled by kinetic linear Stochastic Partial Differential Equations (SPDEs) or Stochastic Differential Equations (SDEs). We study these systems from theoretical and a numerical points of view in two asymptotic regimes: the averaging regime and the diffusion approximation regime.

The first two chapters state the main theoretical contributions of this work. We prove the convergence of the slow component of the considered SPDEs to the solution of a diffusion equation with a source term depending on the asymptotic regime. The first chapter focuses on the diffusion approximation regime, where the source term of the limiting equation is a stochastic diffusive term (Wiener process). The second chapter focuses on the averaging regime, where the limiting source term is the average of the original source term.

The last two chapters are devoted to the numerical part of this work. In general, a numerical scheme which is consistent with a multiscale system for a fixed parameter $\varepsilon > 0$ can perform badly in the asymptotic regime $\varepsilon \rightarrow 0$ due to the presence of stiff terms in the model. On the contrary, some schemes are asymptotic preserving: they are consistent for fixed $\varepsilon > 0$, converge to some limiting schemes when $\varepsilon \rightarrow 0$ and the limiting scheme is consistent with the limiting equation. The goal of the last two chapters is to design asymptotic preserving schemes, respectively for the class of SDEs and SPDEs we consider. We also analyze these schemes and illustrate numerically their efficiency.

Keywords: kinetic equations, slow-fast systems, Stochastic Differential Equations, Stochastic PDEs, diffusion approximation, averaging principle, multiscale methods, Asymptotic Preserving schemes

Étude théorique et numérique d'équations cinétiques stochastiques multi-échelles

Résumé: Cette thèse est dédiée à l'étude d'une classe de systèmes multi-échelles modélisés par une Équation aux Dérivées Partielles Stochastique (EDPS) linéaire cinétique ou une Équation Différentielle Stochastique (EDS). On étudie ces systèmes d'un point de vue théorique et numérique, dans deux régimes asymptotiques : le régime de moyennisation et le régime d'approximation-diffusion.

Les deux premiers chapitres énoncent les principaux résultats théoriques de cette thèse. On montre à chaque fois la convergence de la composante lente du système d'EDPS considéré vers la solution d'une équation de diffusion munie d'un terme source qui dépend du régime asymptotique. Dans le premier chapitre, on considère le régime d'approximation-diffusion, dans lequel le terme source de l'équation limite est un terme diffusif au sens probabiliste (processus de Wiener). Dans le deuxième, on considère le régime de moyennisation, dans lequel le terme source de l'équation limite est la moyenne du terme source de l'EDPS originale.

Les deux derniers chapitres constituent la partie numérique de cette thèse. De manière générale, un schéma numérique peut être consistant avec un système multi-échelle à un paramètre $\varepsilon > 0$ fixé mais se révéler inefficace dans le régime asymptotique $\varepsilon \rightarrow 0$, à cause d'un terme raide dans le modèle. À l'opposé, certains schémas préservent l'asymptotique : ils sont consistants à $\varepsilon > 0$ fixé, convergent vers un schéma limite quand $\varepsilon \rightarrow 0$ et ce schéma limite est consistant avec l'équation limite. Le but des deux derniers chapitres est de proposer, respectivement pour les EDS et les EDPS considérées, des schémas préservant l'asymptotique, de les étudier et d'illustrer numériquement leur efficacité.

Mots clés: équations cinétiques, Équations Différentielles Stochastiques, EDP Stochastiques, approximation-diffusion, principe de moyennisation, méthodes multi-échelles, schéma préservant l'asymptotique

Theoretical and numerical study of multiscale stochastic kinetic equations

Keywords: kinetic equations, slow-fast systems, Stochastic Differential Equations, Stochastic PDEs, diffusion approximation, averaging principle, multiscale methods, Asymptotic Preserving schemes

