Ergodicité de dynamiques de McKean-Vlasov : Mesures Invariantes & Distributions Quasi-stationnaires

Mohamed, Alfaki Aboubacrine Assadek

▶ **To cite this version:**

Mohamed, Alfaki Aboubacrine Assadek. Ergodicité de dynamiques de McKean-Vlasov : Mesures Invariantes & Distributions Quasi-stationnaires. Statistiques [math.ST]. Université d'Angers, 2025. Français. ⟨NNT : 2025ANGE0025⟩. ⟨tel-05460067⟩

HAL Id: tel-05460067

https://theses.hal.science/tel-05460067v1

Submitted on 15 Jan 2026

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.

HAL Authorization

COLLEGE ED MATHEMATIQUES
DOCTORAL STIC
PAYS DE LA LOIRE

université angers

THÈSE DE DOCTORAT

DE

L'UNIVERSITÉ D'ANGERS

SOUS LE SCEAU DE

LA COMUE ANGERS-LE MANS

ÉCOLE DOCTORALE N° 641
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Mathématiques & Interactions*

Par

Mohamed, Alfaki ABOUBACRINE ASSADEK

Ergodicité de dynamiques de McKean-Vlasov

Mesures Invariantes & Distributions Quasi-stationnaires

Thèse présentée et soutenue à Angers, le 29 septembre 2025
Unité de recherche : LAREMA, UMR CNRS 6093

Rapporteurs avant soutenance :

François	BOLLEY	Professeur des universités	ENS de Rennes
Zhenjie	REN	Professeur des universités	Université d'Évry

Composition du Jury :

Président :	Loïc	CHAUMONT	Professeur des universités	Université d'Angers
Examinateurs :	Arnaud	GUILLIN	Professeur des universités	Université Clermont-Auvergne
	Eva	LÖCHERBACH	Professeur des universités	Université Paris 1
	François	BOLLEY	Professeur des universités	ENS de Rennes
	Zhenjie	REN	Professeur des universités	Université d'Évry
Directeur de thèse :	Fabien	PANLOUP	Professeur des universités	Université d'Angers

Remerciements

Une âme délicate est gênée de savoir qu'on lui doit des remerciements, une âme grossière, de savoir qu'elle en doit.

FRIEDRICH NIETZSCHE

Dédicaces

JE DÉDIE MA THÈSE

▷ *À ma famille, à ma mère qui se sacrifie chaque jour et qui a fait de moi la personne que je suis aujourd'hui.*

▷ *À la mémoire de toutes les personnes qui nous ont quittés en particulier, mon père parti trop tôt : j'espère du tréfonds du cœur que, du monde qui est sien maintenant, il apprécie cet humble geste comme preuve de reconnaissance éternelle de la part d'un fils qui a toujours prié pour le salut de son âme. Puisse Dieu, le tout puissant, l'avoir en sa sainte miséricorde !*

La plus grande probabilité de voir nos vœux accomplis laisse encore un doute ; c'est pourquoi, quand nos espérances se réalisent, nous sommes toujours surpris.

JOHANN WOLFGANG VON GOETHE

MES REMERCIEMENTS marquent la fin de neuf belles années à l'université d'Angers. Ces années ont été fructueuses sur plusieurs plans.

EN CE QUI CONCERNE MA THÈSE, en premier lieu, je tiens du tréfonds du cœur à exprimer mon éternelle reconnaissance à mon directeur FABIEN PANLOUP, pour avoir accepté d'encadrer ma thèse et pour la qualité de sa direction tant humainement qu'intellectuellement : sa disponibilité associée à ses nombreuses suggestions et relectures. Je le remercie spécialement de m'avoir fait découvrir le monde fascinant des équations différentielles stochastiques de type McKean-Vlasov pendant mon stage de fin de master recherche [1] qui a débouché sur ma thèse. FRANÇOIS BOLLEY et ZHENJIE REN ont accepté d'être les rapporteurs de ma thèse. Je les en remercie très sincèrement. Je suis très reconnaissant pour le temps qu'ils ont consacré à la lecture de mes travaux, et pour toutes leurs remarques. Je remercie également EVA LÖCHERBACH, LOÏC CHAUMONT et ARNAUD GUILLIN de me faire l'honneur de participer à mon jury de thèse.

RÉALISER MA THÈSE EN BONNE ET DUE FORME a été un travail collectif acharné : le mathématicien que je suis aujourd'hui ne s'est évidemment pas construit tout seul et rien n'aurait été possible sans les belles âmes que j'ai rencontrées tout au long de mon parcours.

CES ÂMES DÉLICATES qui ont contribué de près ou de loin à la réalisation de ma thèse, je vous remercie infiniment. En particulier, je remercie du tréfonds du cœur toutes les personnes avec lesquelles j'ai partagé ces neuf belles années à l'université d'Angers : les étudiant(e)s, les membres du LAREMA, etc.

INFINIMENT RECONNAISSANT envers mes ami(e)s et ma famille en particulier, mon défunt père et ma mère qui se sacrifie chaque jour pour ses enfants et qui a fait de moi la personne que je suis aujourd'hui.

1. Merci au CENTRE HENRI LEBESGUE (**ANR-11-LABX-0020-01**) et à RAWABRANCH (**ANR-23-CE40-0008**)

Table des matières

Introduction **9**

0.1 Mise en contexte et organisation du manuscrit . . . 10

0.1.1 Limites de champ moyen . . . 10

0.1.2 Fonctionnelles de Lyapunov . . . 24

0.1.3 Distributions quasi-stationnaires . . . 39

0.2 Résumés des chapitres et contributions . . . 57

0.2.1 Contenu du chapitre 1 . . . 59

0.2.2 Contenu du chapitre 2 . . . 66

0.2.3 Contenu du chapitre 3 . . . 70

1 Ergodicité exponentielle de limites de champ moyen **79**

1.1 Introduction . . . 80

1.2 Notations and Definitions . . . 91

1.3 Mean-Field Systems and Assumptions . . . 94

1.3.1 Our Systems . . . 94

1.3.2 Our Assumptions . . . 98

1.4 Main Theorems . . . 100

1.4.1 First-order case . . . 100

1.4.2 Kinetic case . . . 101

1.4.3 Examples . . . 101

1.5 Sketch of proofs and preliminaries . . . 106

1.5.1 Sketch of proofs . . . 106

1.5.2 Preliminaries . . . 110
1.6 Proofs of Main Theorems . . . 129
1.7 Appendix and additional proofs . . . 142
1.7.1 Propagation of chaos . . . 142
1.7.2 McKean-Vlasov theory . . . 152
1.7.3 Gibbs-Laplace Variational Principle . . . 152
1.7.4 Principle of contraction and tensorization . . . 155
1.7.5 Entropy and Chaos . . . 155
1.7.6 Proofs . . . 157

2 Dynamique de Langevin de champ moyen et transition de phase 165
2.1 Introduction . . . 166
2.2 Settings and assumptions . . . 171
2.2.1 Settings . . . 172
2.2.2 Assumptions . . . 173
2.3 Main theorems . . . 174
2.3.1 First order case . . . 174
2.3.2 Kinetic case . . . 175
2.3.3 Comments on drift and assumptions . . . 176
2.4 Examples . . . 177
2.4.1 Two-body interaction: granular media flat energy . . . 177
2.4.2 Many-body interaction: polynomial flat energy . . . 179
2.4.3 Internal flat energy . . . 180
2.5 Connection with partial differential equations . . . 181
2.6 Proof of main theorems . . . 182
2.6.1 Proof of main results of the first order case . . . 182
2.6.2 Proof of main results of the kinetic case . . . 190
2.6.3 Proof related to the examples . . . 199

3 QSDs de McKean-Vlasov et schémas d'Euler 203
3.1 Introduction . . . 204
3.2 Setting and Main Results . . . 207
3.2.1 Notation and Setting . . . 207
3.2.2 Applications . . . 215
3.2.3 Simulations . . . 219

3.3 Preliminaries, Poisson Equation and limiting dynamics . . . 220
3.4 Long-time behavior of the limiting dynamics . . . 229
3.5 Proof of Theorem 3.2.8 . . . 238
3.6 Tightness . . . 240
3.7 Euler schemes of McKean-Vlasov SDEs . . . 246
3.7.1 About ($\mathbf{A_V}$) . . . 246
3.7.2 Proof of Proposition 3.2.13(i) . . . 248
3.7.3 Proof of Proposition 3.2.15 . . . 250
3.7.4 Proof of Proposition 3.2.13(ii) . . . 251
3.8 Postponed proofs . . . 260

Conclusion & Perspectives **263**
Perspectives du chapitre 1 . . . 264
Perspectives du chapitre 2 . . . 264
Perspectives du chapitre 3 . . . 264
Publications . . . 269

Bibliographie **271**

Introduction

La mathématique est l'art de donner le même nom à des choses différentes.

HENRI POINCARÉ

Résumé de l'introduction

Ce chapitre, comme son nom l'indique, a vocation à introduire le sujet de cette thèse sur articles. D'abord, il est question d'une part, de la mise en contexte des problèmes étudiés dans ce manuscrit et d'autre part, de l'organisation du document. Ensuite, l'intérêt est porté sur l'arsenal nécessaire à la résolution de ces problèmes. Enfin, des résumés des chapitres du manuscrit dans lesquels sont expliquées les contributions par rapport à la littérature font l'objet de la dernière section de l'introduction.

0.1 Mise en contexte et organisation du manuscrit

L'interprétation des solutions d'équations aux dérivées partielles compte parmi les champs d'application les plus motivants de la théorie des probabilités : elle apporte parfois une meilleure compréhension de la dynamique du système étudié et une façon très intuitive de résoudre numériquement des problèmes dans lesquels les méthodes déterministes ne peuvent plus s'appliquer-principalement en grande dimension. Par exemple, la formule de Feynman-Kac, d'après Richard Feynman et Mark Kac, établit un lien entre les équations aux dérivées partielles paraboliques et les processus stochastiques : on parle aussi d'équations backward et forward de Kolmogorov. Les équations forward de Kolmogorov sont aussi appelées équations de Fokker-Planck.

Cette thèse porte principalement sur l'étude probabiliste du comportement en temps long de deux types d'équations aux dérivées partielles d'évolution décrites dans les prochaines sous-sections.

0.1.1 Limites de champ moyen

Dans un premier temps, l'intérêt est porté aux équations d'évolution appelées limites de champ moyen ou équations de McKean-Vlasov : pour décrire le comportement moyen d'un nombre infini de particules, on est amené à étudier une classe particulière et particulièrement intéressante d'équations différentielles stochastiques dite de McKean-Vlasov. Par exemple, ces équations peuvent servir à modéliser un gaz à forte densité de particules comme l'équation de Boltzmann rend compte du comportement des gaz raréfiés. Ces deux équations peuvent être obtenues par un schéma analogue à celui des équations des milieux granulaires : les systèmes de particules se choquant de manière inélastique ont été étudiés comme un modèle de l'évolution des milieux granulaires. En effet, l'une des caractéristiques de ces systèmes est leur propension à former des amas comme les grains de sable sur une feuille de papier que l'on secoue. McNamara and Young ont montré qu'une équation des milieux granulaires peut être obtenue formellement comme distribution limite d'une particule dans un système de particules déterministes se choquant inélastiquement (en renormalisant convenablement l'amortissement dans les chocs). Les équations des milieux granulaires sont un exemple d'équations de McKean-Vlasov. La différence entre les équations de McKean-Vlasov et l'équation de Boltzmann réside dans le fait de privilégier la fréquence des chocs ou leur intensité. Dans le premier cas, les collisions sont si fréquentes qu'à la limite, les termes de collision sont remplacés par des termes d'interaction de type champ moyen

traduisant l'attraction ou la répulsion des particules. Au contraire, pour les gaz raréfiés, le faible nombre de collisions donne, à la limite, un terme de transport libre et un noyau de collision.

En plus des équations des milieux granulaires, les équations de McKean-Vlasov regroupent quantité d'équations célèbres comme celles de Burgers ou de Navier-Stokes (en dimension deux) si l'on autorise des coefficients irréguliers.

Pour étudier une équation aux dérivées partielles, le probabiliste lui associe un processus stochastique dont la loi est solution, a priori au sens faible, de l'équation. Les processus dits de McKean-Vlasov sont régis par l'équation différentielle stochastique (EDS) :

$$\begin{cases} dX_t = b(t, X_t, \mathbb{P}_{X_t})\,dt + \sigma(t, X_t, \mathbb{P}_{X_t})\,dB_t \\ (B_t)_{t\geq 0} \quad \text{mouvement brownien standard } p\text{-dimensionnel} \\ b : (0,+\infty) \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \longrightarrow \mathbb{R}^d \quad \text{coefficient de drift} \\ \sigma : (0,+\infty) \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \longrightarrow \mathbb{M}_{d,p}(\mathbb{R}) \quad \text{coefficient de diffusion} \\ L_t^X := \mathbb{P}_{X_t} \quad \text{loi de} \quad X_t \\ X_0 \sim \mu \in \mathscr{P}(\mathbb{R}^d) \end{cases} \tag{0.1.1}$$

La différence entre une EDS classique et une EDS de McKean-Vlasov réside dans le fait que les coefficients de l'équation dépendent non seulement de la position de la solution, mais également de sa loi. Les EDS de McKean-Vlasov constituent des exemples d'EDS non-linéaires.

FORMULE D'ITÔ LE LONG D'UN FLOT DE MESURES DE PROBABILITÉ

La formule d'Itô le long d'un flot de mesures permet d'obtenir l'équation d'évolution dite équation maîtresse satisfaite par le flot des lois d'une EDS de McKean-Vlasov.

La formule d'Itô le long d'un flot de mesures de probabilité a été développée sous l'impulsion de la théorie des jeux à champ moyen et des EDS de McKean-Vlasov. Les jeux à champ moyen ont été introduits indépendamment par Huang et al. et par Lasry et Lions. La notion de master equation ou équation maîtresse a été introduite par Lions dans son cours au Collège de France dans le but de décrire les jeux à champ moyen. Les équations maîtresses sont des EDPs définies sur un espace de mesures de probabilité. La formule d'Itô le long d'un flot de mesures est l'un des outils qui permet d'établir ce genre d'équations. On renvoie au cours de Lions, aux notes écrites par Cardaliaguet sur le sujet et les deux livres écrits par Carmona and Delarue ([51, 50]) pour plus de détails sur les jeux à champ moyen et les équations

maîtresses.

La formule d'Itô le long d'un flot de mesures permet d'associer naturellement une EDP sur l'espace des mesures de probabilité à une EDS de McKean-Vlasov, plus précisément à son semi-groupe qui agit sur les fonctions définies sur un espace de mesures : le but de la formule d'Itô le long d'un flot de mesures de probabilité est de décrire la dynamique de $t \in (0,+\infty) \mapsto \Psi(\mathrm{L}_t^{\mathrm{X}})$ avec $\Psi : \mathscr{P}_{p\geq 0}(\mathbb{R}^d) \to (-\infty,+\infty)$, comme le fait la formule d'Itô classique pour $t \mapsto f(\mathrm{X}_t)$ avec $f : \mathbb{R}^d \to (-\infty,+\infty)$. Dans le cadre classique, la formule d'Itô requiert de la régularité sur la fonction f. Il en est de même pour la formule d'Itô pour un flot de mesures : régularité sur la fonctionnelle Ψ. En effet, supposons que Ψ est $\mathscr{C}^1$ en l'argument mesure au sens de la dérivée plate (dérivée au sens de Gateaux sur l'espace des mesures signées finies) donnée par $\frac{\delta}{\delta m} : \Psi \mapsto \frac{\delta \Psi}{\delta m} : \mathscr{P}_{p\geq 0}(\mathbb{R}^d) \times \mathbb{R}^d \to (-\infty,+\infty)$ l'opérateur de dérivation plate vérifiant

$$\Psi(\mu_1) - \Psi(\mu_0) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta \Psi}{\delta m}((1-t)\mu_0 + t\mu_1, x)(\mu_1 - \mu_0)(dx)dt. \tag{0.1.2}$$

Dans ce cas, si pour tout μ, $\frac{\delta \Psi}{\delta m}(\mu,\cdot)$ est $\mathscr{C}^2$ et $(\mu,x) \mapsto \nabla^2 \frac{\delta \Psi}{\delta m}(\mu,x)$ est continue, de [51, Théorème 5.99] (dans lequel il y a un énoncé général avec la dérivée de Lions), on a la formule d'Itô le long d'un flot de mesures de probabilité :

$$\begin{aligned}\Psi(\mathrm{L}_t^{\mathrm{X}}) = \Psi(\mathrm{L}_0^{\mathrm{X}}) &+ \int_0^t \mathbb{E}\Big[b(s,\mathrm{X}_s,\mathrm{L}_s^{\mathrm{X}}) \cdot \nabla \frac{\delta \Psi}{\delta m}(\mathrm{L}_s^{\mathrm{X}},\mathrm{X}_s)\Big] ds \\ &+ \frac{1}{2}\int_0^t \mathbb{E}\Big[\mathbf{trace}\Big(\sigma\sigma^*(s,\mathrm{X}_s,\mathrm{L}_s^{\mathrm{X}})\nabla^2 \frac{\delta \Psi}{\delta m}(\mathrm{L}_s^{\mathrm{X}},\mathrm{X}_s)\Big)\Big] ds.\end{aligned} \tag{0.1.3}$$

On peut l'appeler également équation maîtresse. Cela s'avère être un outil important pour étudier le flot de mesures associé à une EDS de McKean-Vlasov, comme expliqué par Carmona and Delarue ([51, Chapitre 5]). Enfin, la formule d'Itô le long d'un flot de mesures est un outil important quand on étudie des problèmes de contrôles d'EDS de McKean-Vlasov. Elle permet d'obtenir un principe de programmation dynamique qui décrit la fonction valeur du problème, comme présenté par Carmona and Delarue ([51, Chapitre 6]).

De la formule d'Itô le long d'un flot de mesures de probabilité, en prenant pour toute fonction test φ, $\Psi_\varphi : \mu \mapsto \int \varphi d\mu$, on déduit l'équation aux dérivées partielles satisfaite par le flot $(\mathrm{L}_t^{\mathrm{X}})_{t\geq 0}$ des lois du processus de McKean-Vlasov aussi appelée équation de Fokker-Planck non linéaire : (au sens des distributions)

$$\partial_t \mathrm{L}_t^{\mathrm{X}} = \mathrm{L}_t^{\mathrm{X}} \mathscr{L}^{\mathrm{L}_t^{\mathrm{X}}} \tag{0.1.4}$$

vu que plus généralement [2], on a

$$\partial_t \Psi(\mathrm{L}_t^{\mathrm{X}}) = \int_{\mathbb{R}^d} \mathscr{L}^{\mathrm{L}_t^{\mathrm{X}}} \frac{\delta \Psi}{\delta m}(\mathrm{L}_t^{\mathrm{X}}, \cdot)(x) \mathrm{L}_t^{\mathrm{X}}(dx) \tag{0.1.5}$$

avec

$$\mathscr{L}^{\mu} := b(\cdot, \cdot, \mu) \cdot \nabla + \frac{1}{2} \mathbf{trace}(\sigma\sigma^*(\cdot, \cdot, \mu) \nabla^2) \tag{0.1.6}$$

le générateur infinitésimal du processus de diffusion solution de

$$d\mathrm{X}_t^{\mu} = b(t, \mathrm{X}_t^{\mu}, \mu) dt + \sigma(t, \mathrm{X}_t^{\mu}, \mu) d\mathrm{B}_t^{\mu}. \tag{0.1.7}$$

Le flot des lois du processus $(\mathrm{X}_t^{\mu})_{t \geq 0}$ est décrit par son semi-groupe $(\mathrm{P}_t^{\mu})_{t \geq 0}$ défini par

$$\mathrm{P}_t^{\mu} \varphi(x) := \mathbb{E}[\varphi(\mathrm{X}_t^{\mu}) | \mathrm{X}_0^{\mu} = x]. \tag{0.1.8}$$

En effet, le flot des lois de la solution de l'équation 0.1.7 partant de $\mathrm{X}_0^{\mu} \sim \nu \in \mathscr{P}(\mathbb{R}^d)$ est donné par $(\nu \mathrm{P}_t^{\mu})_{t \geq 0}$:

$$\langle \nu \mathrm{P}_t^{\mu}, \varphi \rangle := \int_{\mathbb{R}^d} \mathrm{P}_t^{\mu} \varphi(x) \nu(dx). \tag{0.1.9}$$

De plus, $(\nu \mathrm{P}_t^{\mu})_{t \geq 0}$ est solution de l'équation de Fokker-Planck :

$$\partial_t \nu \mathrm{P}_t^{\mu} = \nu \mathrm{P}_t^{\mu} \mathscr{L}^{\mu}. \tag{0.1.10}$$

Pour plus de détails à ce sujet, le livre de Bakry et al. sur l'analyse et la géométrie des processus de diffusion markoviens est une excellente monographie. Notons qu'il y a équivalence entre la donnée de l'équation 0.1.1 et la donnée de l'équation 0.1.4 : on parle de problème de martingales non linéaire. La formulation forte de l'équation de Fokker-Planck non-linéaire est donnée par

$$\partial_t f_t = \frac{1}{2} \sum_{i,j} \partial_{i,j}^2 \Big(\sigma\sigma^*_{i,j}(t, \cdot, \mathrm{L}_t^{\mathrm{X}}) f_t \Big) - \nabla \cdot \Big(b(t, \cdot, \mathrm{L}_t^{\mathrm{X}}) f_t \Big). \tag{0.1.11}$$

EXISTENCE ET UNICITÉ

La première question qu'on peut se poser concerne le caractère bien posé de l'équation de McKean-Vlasov. Par bien posé, on entend l'existence et l'unicité des solutions, qui peuvent être considérées au sens fort ou au sens faible.

2. Sous les hypothèses sur Ψ introduites plus haut.

EDS linéaires. Dans le cas des EDS linéaires, c'est-à-dire quand les coefficients ne dépendent pas de la loi, il est bien connu que le bruit permet de montrer le caractère bien posé, sous des hypothèses de régularité en espace plus faible que lipschitzienne. Il s'agit du phénomène de régularisation par le bruit : voir, par exemple, le livre de Stroock and Varadhan pour plus de détails à ce sujet.

EDS non-linéaires. La dépendance des coefficients par rapport à la loi est plus complexe à traiter. Le choix de la distance utilisée sur l'espace des mesures est crucial, comme illustré dans ce qui suit, puisque les distances ne sont pas équivalentes.

Cadre Lipschitz. Si les coefficients sont globalement lipschitziens sur $\mathbb{R}^d \times \mathscr{P}_{p\geq 2}(\mathbb{R}^d)$, localement uniformément en temps, alors il existe une unique solution forte. Ici, on travaille avec la distance de Wasserstein $\mathscr{W}_p$, qui est contrôlée par la distance L^p sur les trajectoires. C'est ce qui permet de prouver le caractère bien posé dans ce cadre. Sznitman a fait la preuve dans le cas $p = 2$.

Choix de la métrique. Le contre-exemple de Scheutzow montre que si on considère cette fois la distance en variation totale sur $\mathscr{P}(\mathbb{R}^d)$, alors on perd le caractère bien posé sous des hypothèses lipschitziennes et sans bruit. On n'a donc pas de théorème de Cauchy-Lipschitz pour la distance en variation totale. Cela montre que le choix de la distance sur $\mathscr{P}(\mathbb{R}^d)$ fait dans les hypothèses lipschitziennes est crucial car il influence le caractère bien posé.

Régularisation par le bruit. On peut se demander ce qui se passe dans le même cadre que le contre-exemple précédent, mais cette fois en présence de bruit. Shiga and Tanaka ont prouvé l'unicité trajectorielle. Le bruit permet donc de restaurer l'unicité qui fait défaut dans le contre-exemple de Scheutzow. Il s'agit là encore d'un phénomène de régularisation par le bruit. De plus, ce résultat a été étendu par Jourdain dans le cas où la dérive est bornée, lipschitzienne en variation totale par rapport à la mesure, et où la diffusion est constante. Le bruit permet donc ici d'établir le caractère bien posé, sous des hypothèses lipschitziennes relatives à une plus grande classe de distances utilisées, par rapport à la variable mesure.

Cadre non Lipschitz. Contrairement à la variable spatiale, se passer du caractère lipschitzien par rapport à l'argument mesure, relativement à une certaine distance, semble très compliqué même en présence de bruit. On peut l'illustrer à travers le contre-exemple suivant, dû à François Delarue. Soit $(\mathrm{Y}_t)_{t\geq 0}$ le processus de McKean-Vlasov tel que

$$\forall\, t \geq 0, \quad \mathrm{Y}_t = \mathrm{B}_t + \int_0^t \widetilde{b}(\mathbb{E}[\mathrm{Y}_s])\, ds \tag{0.1.12}$$

avec $\widetilde{b} : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ hölderienne et bornée. Notons que le drift[3] de l'équation de McKean-Vlasov dont $(Y_t)_{t\geq 0}$ est solution a une régularité hölderienne par rapport à la métrique de Wasserstein $\mathcal{W}_1$. En prenant l'espérance, on remarque que l'application $t \in (0,+\infty) \mapsto \mathbb{E}[Y_t]$ est solution de $y' = \widetilde{b}(y)$, qui peut avoir plusieurs solutions puisque $\widetilde{b}$ est seulement höldérienne.

Quant aux équations maîtresses, la question de l'existence et l'unicité de solutions classiques a été étudiée notamment par Cardaliaguet, Delarue, Lasry et Lions.

PROPAGATION DU CHAOS

Le fait que les processus de McKean-Vlasov décrivent la dynamique moyenne d'un nombre infini de particules se traduit par la notion de propagation du chaos.

On se donne un système de $n \geq 1$ particules telles que chaque particule est régie par l'équation 0.1.1 dans laquelle on remplace l'interaction avec la loi par une interaction avec la mesure empirique aléatoire du système définie par

$$\forall \omega \in \Omega, \quad \Pi_t^{(n)}(\omega) := \frac{1}{n}\sum_{j=1}^{n} \delta_{X_t^{(n),j}(\omega)} \in \mathscr{P}(\mathscr{P}(\mathbb{R}^d)) \tag{0.1.13}$$

et on prend des copies indépendantes du mouvement brownien :

$$\forall i \in \{1,\dots,n\}, \quad dX_t^{(n),i} = b(t,X_t^{(n),i},\Pi_t^{(n)})dt + \sigma(t,X_t^{(n),i},\Pi_t^{(n)})dB_t^{(n),i}. \tag{0.1.14}$$

Ce type d'interaction est appelé *champ moyen* puisque chaque particule interagit avec la distribution statistique de tout le système. Les systèmes de particules en interaction de type champ moyen ont de nombreuses applications. En physique, ils ont un rôle crucial en théorie cinétique, comme mentionné précédemment et dans la suite. Ils ont également de nombreuses applications en biologie, notamment pour étudier la dynamique d'une population de cellules, en neurosciences pour modéliser les interactions entre neurones, en sciences sociales pour décrire des mouvements d'auto-organisation et aussi dans la théorie des jeux à champ moyen ([17, 18, 192, 193, 177, 211, 90, 91, 94, 202, 123, 53, 140, 49, 50, 51]).

Quand on part à l'instant initial ($t = 0$) avec des particules indépendantes et identiquement distribuées selon μ, à t fixé, en faisant tendre le nombre de particules vers l'infini ($n \to +\infty$), on peut espérer que k (fixé!) particules choisies au hasard parmi les n seront in-

3. $\mu \longmapsto \widetilde{b}\Big(\int_{\mathbb{R}^d} x\mu(dx)\Big)$.

dépendantes et identiquement distribuées selon L_t^X : on parle de *phénomène de propagation du chaos*. Les particules sont identiquement distribuées grâce à la symétrie de l'interaction et le caractère i.i.d. des bruits et des données initiales. En revanche, elles ne sont pas indépendantes puisque chacune interagit avec les autres mais elles sont échangeables : la loi d'un nombre fini de particules est invariante par permutations.

L'origine de la notion de propagation du chaos remonte à Boltzmann et l'hypothèse de chaos moléculaire à l'origine de la théorie cinétique. La justification mathématique de ses travaux et de cette hypothèse est restée longtemps sans réponse. C'est Kac qui donne le premier une définition rigoureuse du chaos et l'illustre sur un système simplifié en s'intéressant au lien unissant les versions détaillées et réduites des équations décrivant l'évolution d'un système de n particules en interaction. Le problème initial concernait l'équation de Boltzmann pour des gaz dilués de sphères dures. Peu après Kac, en s'intéressant à l'équation de Vlasov, qui décrit la distribution des particules chargées dans un plasma, McKean a introduit une classe de modèles de diffusion qui ne faisaient pas initialement partie de la théorie de Boltzmann mais qui satisfont la propriété de propagation du chaos. Ensuite, Sznitman donne une definition dans laquelle il suppose que les conditions initiales sont seulement chaotiques au lieu d'indépendantes et identiquement distribuées.

Donnons des définitions plus précises de la propagation du chaos au sens faible et au sens fort.

Définition 0.1.1 (Suite chaotique au sens de Kac)**.** Soit $(L_t^{(n)})_{t\geq 0}$ le flot des lois du système de particules et $s \geq 0$. On dit que $(L_s^{(n)})_{n\geq 1}$ est L_s^X-chaotique si, et seulement si, pour tout k, la marginale $L_s^{(n),k} \in \mathscr{P}((\mathbb{R}^d)^k)$ converge en loi vers $(L_s^X)^{\otimes k}$.

Cela signifie qu'à un tel instant, asymptotiquement, pour tout k, k particules sont indépendantes et identiquement distribuées selon la loi du processus de McKean-Vlasov.

Définition 0.1.2 (Propagation du chaos faible)**.** On dit qu'il y a propagation du chaos faible si, et seulement si, pour tout $t \geq 0$, $(L_t^{(n)})_{n\geq 1}$ est L_t^X-chaotique.

Cela signifie qu'on a le chaos au sens de Kac à tout instant. Sznitman ([217, Proposition 2.2]) a montré que la propagation du chaos faible est équivalente à la convergence en loi pour tout $t \geq 0$ de la suite $(\Pi_t^{(n)})_{n\geq 1}$ des mesures empiriques aléatoires du système de particules vers la variable aléatoire constante L_t^X. Puisque ces deux notions sont équivalentes, on s'autorisera dans la suite à parler de propagation du chaos faible dès lors qu'on prouve une des deux propriétés équivalentes.

On peut également s'intéresser à la propagation du chaos au niveau des trajectoires. Introduisons des copies i.i.d. $(X^{(n),\infty})_{n\geq 1}$ de $(X_t)_{t\geq 0}$, où les données initiales et les bruits sont identiques à ceux du système de particules : on parle de *couplage synchrone*.

Définition 0.1.3 ($L^{p\geq 1}$-Propagation du chaos forte (trajectorielle))**.** On dit qu'il y a $L^{p\geq 1}$-propagation du chaos forte (ou trajectorielle) pour le système de particules, si et seulement si, pour tout $t \geq 0$,

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[|X_t^{(n),i} - X_t^{(i),\infty}|^p] \overset{n\to+\infty}{\longrightarrow} 0. \tag{0.1.15}$$

Mentionnons qu'on peut utiliser d'autres couplages : par exemple, Eberle a popularisé le *couplage par reflexion*. La terminologie de propagation du chaos faible ou forte utilisée ici provient des propriétés correspondantes pour les schémas numériques d'EDS. De plus, la propagation du chaos forte entraîne la propagation du chaos faible. En effet, s'il y a propagation du chaos au sens fort, alors, pour tous k et t, $(L_t^{(n),k})_{n\geq 1}$ converge en métrique $\mathscr{W}_p$ de Wasserstein vers $(L_t^X)^{\otimes k}$ (donc en loi).

Ces propagations du chaos sont dites *qualitatives* car il s'agit de résultats de convergence sans vitesse explicite par rapport au nombre de particules. Quand des taux de convergence explicites par rapport au nombre de particules sont prouvés, on dit que la propagation du chaos est *quantitative*. Précisons ce que cela signifie pour la propagation du chaos forte. On se place désormais sur un intervalle de temps borné de la forme $[0,T]$ avec $T > 0$. On peut sans ambiguïté dire qu'il s'agit de trouver une suite $(\varepsilon_n^T)_{n\geq 1}$ explicite par rapport au nombre de particules, tendant vers 0 et telle que

$$\forall n \geq 1 \quad \forall t \in [0,T], \quad \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[|X_t^{(n),i} - X_t^{(i),\infty}|^p] \leq \varepsilon_n^T. \tag{0.1.16}$$

On dit que la propagation du chaos quantitative a une borne uniforme en temps si on peut choisir $(\varepsilon_n^T)_{n\geq 1}$ indépendemment de T. Le couplage par reflexion doit sa popularité aux reponses qu'a apportées Eberle sur les questions de propagation du chaos quantitative *uniforme en temps* et de comportement en temps long quantitatif (voir entre autres [106, 109]).

La classe de McKean-Vlasov n'est pas markovienne. Néanmoins, le couple $(X_t, L_t^X) \in \mathbb{R}^d \times \mathscr{P}_p(\mathbb{R}^d)$ définit un processus de Markov sur un espace de dimension infinie, ce qui génère naturellement de nombreuses difficultés. Pour pallier ces problèmes, on préfère plutôt étudier le système de particules sous-jacent. En effet, cette approche particulaire permet de restaurer le bon vieux caractère markovien en dimension finie pour lequel les questions d'existence, d'unicité et de comportement en temps long sont largement étudiées dans la

littérature : voir, entre autres, l'excellente monographie de Bakry et al.

ÉQUATIONS DE LANGEVIN DE CHAMP MOYEN

Dans cette thèse, on considère l'équation de Langevin de champ moyen définie par

$$dX_t = \sqrt{2}dB_t - \nabla\frac{\delta H}{\delta m}(L_t^X, X_t)dt \tag{0.1.17}$$

avec $H : \mathscr{P}_p(\mathbb{R}^d) \longrightarrow (-\infty, +\infty)$ une fonctionnelle appelée *hamiltonien macroscopique* et $\nabla\frac{\delta H}{\delta m}$ sa dérivée de Fréchet [4]. On considère aussi l'*équation cinétique* associée :

$$\begin{cases} dP_t &= V_t dt; \\ dV_t &= \sqrt{2}dB_t - V_t dt - \nabla\frac{\delta H}{\delta m}(L_t^P, P_t)dt. \end{cases} \tag{0.1.18}$$

Notons que les équations cinétiques sont des équations de McKean-Vlasov dites du second ordre.

Soit $\mu \in \mathscr{P}(\mathbb{R}^d)$. Il est bien connu que sous des hypothèses adéquates, le processus de Langevin $(X_t^\mu)_{t\geq 0}$ admet une unique mesure de probabilité invariante :

$$\Phi(\mu)(dx) := \frac{1}{Z_\mu} e^{-\frac{\delta H}{\delta m}(\mu,x)} dx. \tag{0.1.19}$$

De plus, les mesures de probabilité invariantes du processus de Langevin de champ moyen sont les éventuels *points fixes* de $\mu \mapsto \Phi(\mu)$ appelée *application mesure invariante*. Quant à l'équation cinétique, on a les mêmes conclusions avec

$$\begin{aligned} \Psi(\mu)(dpdv) &:= \frac{1}{C_\mu} e^{-\frac{1}{2}|v|^2 - \frac{\delta H}{\delta m}(\mu,p)} dpdv \\ &= \Phi(\mu) \bigotimes \mathcal{N}(0, \mathbf{Id}_d)(dpdv). \end{aligned} \tag{0.1.20}$$

Les systèmes de particules sous-jacents sont respectivement solutions des équations de Lan-

4. Gradient d'une dérivée plate définie plus haut.

gevin sur-amortie et amortie :

$$dX_t^{(n)} = \sqrt{2}dB_t^{(n)} - \nabla H_n(X_t^{(n)})dt; \tag{0.1.21}$$

$$\begin{cases} dP_t^{(n),i} &= V_t^{(n),i}dt; \\ dV_t^{(n),i} &= \sqrt{2}dB_t^{(n),i} - V_t^{(n),i}dt - \nabla_{x_i}H_n(P_t^{(n)})dt. \end{cases} \tag{0.1.22}$$

avec $H_n : (\mathbb{R}^d)^n \longrightarrow (-\infty, +\infty)$ une fonction appelée *hamiltonien microscopique* et définie par

$$H_n(\mathbf{x}) := nH\Big(\frac{1}{n}\sum_{i=1}^n \delta_{x_i}\Big). \tag{0.1.23}$$

De plus, les mesures de probabilité invariantes associées sont respectivement les mesures de Gibbs et de Boltzmann-Gibbs suivantes :

$$\mu_\infty^{(n)}(d\mathbf{x}) := \frac{1}{Z_n}e^{-H_n(\mathbf{x})}d\mathbf{x} \tag{0.1.24}$$

$$\nu_\infty^{(n)} := \mu_\infty^{(n)} \bigotimes \mathcal{N}(0, \mathbf{Id}_{nd})(d\mathbf{p}d\mathbf{v}). \tag{0.1.25}$$

Approches du temps long. Pour l'étude du comportement en temps long, il y a deux approches : une directe (par exemple, par *couplage* ou par *dissipation d'énergie*) et l'autre indirecte (via la propagation du chaos).

La première approche consiste à étudier directement l'équation d'évolution de Langevin de champ moyen : du fait de la non-linéarité, l'étude est en général très difficile voire impossible et nécessite beaucoup de régularité.

Quant à la seconde approche, à partir de l'étude de l'équation d'évolution *linéaire* du système de particules sous-jacent, on déduit le comportement en temps long de l'équation d'évolution *non-linéaire* : via la propagation du chaos, on étudie l'équation d'évolution de la limite de champ moyen, ce qui est plus raisonnable que l'étudie directe et nécessite moins de régularité.

Milieux granulaires. Un exemple d'une telle dynamique est donné par l'équation des milieux granulaires :

$$dX_t = \sqrt{2}dB_t - \nabla(V + W * L_t^X)(X_t)dt \tag{0.1.26}$$

avec $V : \mathbb{R}^d \longrightarrow (-\infty, +\infty)$ une fonction appelée *potentiel de confinement*, $W : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow (-\infty, +\infty)$ une fonction symétrique appelée *potentiel d'interaction* et $W * L_t^X$ le *produit de*

convolution entre le potentiel d'interaction et la loi à l'instant t du processus :

$$\mathrm{W} * \mathrm{L}_t^{\mathrm{X}}(x) := \int \mathrm{W}(x,y)\mathrm{L}_t^{\mathrm{X}}(dy). \tag{0.1.27}$$

Littérature. Concernant l'équation des milieux granulaires, les questions sur la propagation du chaos et le comportement en temps long ont été largement étudiées dans la littérature : on peut citer Carrillo et al. (approche directe du temps long par dissipation d'entropie), Malrieu (propagation du chaos uniforme en temps et comportement en temps long quantitatifs), Cattiaux et al. (cas non uniformément convexe), Bolley et al. (temps long par dissipation de la métrique de Wasserstein), Bolley et al. (temps long par couplage pour l'équation cinétique associée), Guillin et al. (inégalités fonctionnelles uniformes et temps long), Monmarché (cas cinétique), Guillin and Monmarché (cas cinétique avec non convexité), etc.

Quant à l'équation de Langevin de champ moyen, depuis que Hu et al. ont établi le lien avec l'optimisation des réseaux de neurones, ces questions ont été largement étudiées dans la littérature : on peut citer Chen et al. (propagation du chaos uniforme en temps et convergence en temps long), Liu et al. (convergence exponentielle en temps long et algorithme de descente de gradient), Wang (inégalité fonctionnelle uniforme avec énergie plate convexe), etc.

Quelques fondamentaux de la thèse. Illustrons les deux approches sur l'équation des milieux granulaires. Supposons que $\mathrm{W} := \mathbf{W}(\cdot - \cdot)$ avec $\mathbf{W} : \mathbb{R}^d \longrightarrow (-\infty, +\infty)$ un potentiel pair.

Concernant l'approche directe, supposons que les potentiels sont deux fois différentiables, la hessienne de $\mathbf{W}$ est bornée et les hypothèses de convexité suivantes sont satisfaites :

$$\exists \lambda_{\mathrm{V}} > 0 \quad \forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \langle \nabla \mathrm{V}(x) - \nabla \mathrm{V}(y), x - y \rangle \geq \lambda_{\mathrm{V}} |x-y|^2; \tag{0.1.28}$$

$$\langle \nabla \mathbf{W}(x) - \nabla \mathbf{W}(y), x - y \rangle \geq 0. \tag{0.1.29}$$

En faisant un couplage synchrone de deux solutions $(\mathrm{X}_t)_{t\geq 0}$ et $(\mathrm{Y}_t)_{t\geq 0}$, on a

$$\begin{aligned} \frac{d}{dt}|\mathrm{X}_t - \mathrm{Y}_t|^2 &= -2\langle \nabla(\mathrm{V} + \mathrm{W} * \mathrm{L}_t^{\mathrm{X}})(\mathrm{X}_t) - \nabla(\mathrm{V} + \mathrm{W} * \mathrm{L}_t^{\mathrm{Y}})(\mathrm{Y}_t), \mathrm{X}_t - \mathrm{Y}_t \rangle \\ &\leq -2\lambda_{\mathrm{V}}|\mathrm{X}_t - \mathrm{Y}_t|^2 - 2\langle \nabla \mathrm{W} * \mathrm{L}_t^{\mathrm{X}}(\mathrm{X}_t) - \nabla \mathrm{W} * \mathrm{L}_t^{\mathrm{Y}}(\mathrm{X}_t), \mathrm{X}_t - \mathrm{Y}_t \rangle \\ &\leq -2\lambda_{\mathrm{V}}|\mathrm{X}_t - \mathrm{Y}_t|^2 + 2\|\nabla^2 \mathrm{W}\|_{\mathbf{op},\infty} \mathbb{E}[|\mathrm{X}_t - \mathrm{Y}_t|] \cdot |\mathrm{X}_t - \mathrm{Y}_t|. \end{aligned} \tag{0.1.30}$$

En prenant l'espérance et en appliquant l'*inégalité de Jensen*, on a

$$\frac{d}{dt}\mathbb{E}[|X_t - Y_t|^2] \leq 2(\|\nabla^2 W\|_{\mathbf{op},\infty} - \lambda_V)\mathbb{E}[|X_t - Y_t|^2]. \tag{0.1.31}$$

En appliquant le *lemme de Gronwall*, on en déduit que

$$\mathbb{E}[|X_t - Y_t|^2] \leq \mathbb{E}[|X_0 - Y_0|^2]\, e^{2(\|\nabla^2 W\|_{\mathbf{op},\infty} - \lambda_V)t}. \tag{0.1.32}$$

Par définition de la métrique de Wasserstein et l'*existence d'un couplage optimal*, on a

$$\mathscr{W}_2(L_t^X, L_t^Y) \leq e^{(\|\nabla^2 W\|_{\mathbf{op},\infty} - \lambda_V)t}\mathscr{W}_2(L_0^X, L_0^Y). \tag{0.1.33}$$

En particulier, si $\|\nabla^2 W\|_{\mathbf{op},\infty} - \lambda_V < 0$ (interaction petite par rapport au confinement), alors on a contraction. Il en découle que $\mu \mapsto \Phi(\mu)$ admet au plus un point fixe et si μ^* est un point fixe, alors il y a convergence exponentielle :

$$\mathscr{W}_2(L_t^X, \mu^*) \leq e^{(\|\nabla^2 W\|_{\mathbf{op},\infty} - \lambda_V)t}\mathscr{W}_2(L_0^X, \mu^*). \tag{0.1.34}$$

Quant au système de particules, on fait la même chose. En effet, le système de particules est solution de l'équation de Langevin standard :

$$dX_t^{(n)} = \sqrt{2}dB_t^{(n)} - \nabla H_n(X_t^{(n)})\,dt \tag{0.1.35}$$

avec $H_n : (\mathbb{R}^d)^n \longrightarrow (-\infty, +\infty)$ une fonction appelée *hamiltonien microscopique* et définie par

$$H_n(\mathbf{x}) := \sum_{i=1}^n V(x_i) + \frac{1}{2n}\sum_{i,j=1}^n \mathbf{W}(x_i - x_j). \tag{0.1.36}$$

Comme précédemment, si

$$\exists \lambda_n > 0 \quad \forall (\mathbf{x},\mathbf{y}) \in (\mathbb{R}^d)^n \times (\mathbb{R}^d)^n, \quad \langle \nabla H_n(\mathbf{x}) - \nabla H_n(\mathbf{y}), \mathbf{x} - \mathbf{y}\rangle \geq \lambda_n |\mathbf{x} - \mathbf{y}|^2, \tag{0.1.37}$$

en faisant un couplage synchrone de deux solutions $(X_t^{(n)})_{t\geq 0}$ et $(Y_t^{(n)})_{t\geq 0}$, on a

$$\frac{d}{dt}|X_t^{(n)} - Y_t^{(n)}|^2 \leq -2\lambda_n |X_t^{(n)} - Y_t^{(n)}|^2. \tag{0.1.38}$$

Par le lemme de Gronwall, on a

$$|X_t^{(n)} - Y_t^{(n)}|^2 \le |X_0^{(n)} - Y_0^{(n)}|^2 e^{-2\lambda_n t}. \tag{0.1.39}$$

On dit qu'on a *confluence des trajectoires* à l'infini. En prenant l'espérance, par définition de la métrique de Wasserstein et l'*existence d'un couplage optimal*, on en déduit qu'on a contraction exponentielle en métrique de Wasserstein :

$$\mathscr{W}_2(\mathbb{P}_{X_t^{(n)}}, \mathbb{P}_{Y_t^{(n)}}) \le e^{-\lambda_n t} \mathscr{W}_2(\mathbb{P}_{X_0^{(n)}}, \mathbb{P}_{Y_0^{(n)}}). \tag{0.1.40}$$

Il en découle que

$$\mathscr{W}_2(\mathbb{P}_{X_t^{(n)}}, \mu_\infty^{(n)}) \le e^{-\lambda_n t} \mathscr{W}_2(\mathbb{P}_{X_0^{(n)}}, \mu_\infty^{(n)}). \tag{0.1.41}$$

Exemple 0.1.1 (Conditions sur les potentiels). Comme précédemment (eq. (0.1.28),eq. (0.1.29)), on va donner des conditions explicites sur les potentiels assurant la propriété de convexité (eq. (0.1.37)).

Vu que

$$\mathrm{H}(\mu) = \int \mathrm{V} d\mu + \frac{1}{2} \int \mathbf{W}(x-y)\mu(dx)\mu(dy); \tag{0.1.42}$$

$$\frac{\delta \mathrm{H}}{\delta m}(\mu, \cdot) = \mathrm{V} + \mathbf{W} * \mu; \tag{0.1.43}$$

$$\nabla_{x_i} \mathrm{H}_n(\mathbf{x}) = \nabla \frac{\delta \mathrm{H}}{\delta m}\Big(\frac{1}{n}\sum_{j=1}^n \delta_{x_j}, x_i\Big); \tag{0.1.44}$$

on a

$$\begin{aligned} \langle \nabla \mathrm{H}_n(\mathbf{y}) - \nabla \mathrm{H}_n(\mathbf{z}), \mathbf{y}-\mathbf{z}\rangle &= \sum_{i=1}^n \langle \nabla_{x_i} \mathrm{H}_n(\mathbf{y}) - \nabla_{x_i} \mathrm{H}_n(\mathbf{z}), y_i - z_i\rangle \\ &= \sum_{i=1}^n \langle \nabla \mathrm{V}(y_i) - \nabla \mathrm{V}(z_i), y_i - z_i\rangle \\ &\quad + \frac{1}{n}\sum_{i,j=1}^n \langle \nabla \mathbf{W}(y_i - y_j) - \nabla \mathbf{W}(z_i - z_j), y_i - z_i\rangle. \end{aligned} \tag{0.1.45}$$

Si V vérifie eq. (0.1.28), on a

$$\sum_{i=1}^n \langle \nabla \mathrm{V}(y_i) - \nabla \mathrm{V}(z_i), y_i - z_i\rangle \ge \lambda_{\mathrm{V}} |\mathbf{y}-\mathbf{z}|^2. \tag{0.1.46}$$

D'autre part, si $\mathbf{W}$ vérifie eq. (0.1.29), on a

$$\begin{aligned}\frac{1}{n}\sum_{i,j=1}^{n}\langle\nabla\mathbf{W}(y_i-y_j)-\nabla\mathbf{W}(z_i-z_j),y_i-z_i\rangle &= \frac{1}{n}\sum_{i,j=1}^{n}\langle\nabla\mathbf{W}(y_i-y_j)-\nabla\mathbf{W}(z_i-y_j),y_i-z_i\rangle \\ &+\frac{1}{n}\sum_{i,j=1}^{n}\langle\nabla\mathbf{W}(z_i-y_j)-\nabla\mathbf{W}(z_i-z_j),y_i-z_i\rangle \\ &\geq \frac{1}{n}\sum_{i,j=1}^{n}\langle\nabla\mathbf{W}(z_i-y_j)-\nabla\mathbf{W}(z_i-z_j),y_i-z_i\rangle.\end{aligned} \tag{0.1.47}$$

De plus, si $\nabla\mathbf{W}$ est lipschitzien, par l'inégalité de Cauchy-Schwarz, on a

$$\frac{1}{n}\sum_{i,j=1}^{n}\langle\nabla\mathbf{W}(z_i-y_j)-\nabla\mathbf{W}(z_i-z_j),y_i-z_i\rangle \geq -\frac{[\nabla\mathbf{W}]_1}{n}|\mathbf{y}-\mathbf{z}|^2. \tag{0.1.48}$$

On en déduit que

$$\langle\nabla\mathrm{H}_n(\mathbf{y})-\nabla\mathrm{H}_n(\mathbf{z}),\mathbf{y}-\mathbf{z}\rangle \geq \left(\lambda_{\mathrm{V}}-\frac{[\nabla\mathbf{W}]_1}{n}\right)|\mathbf{y}-\mathbf{z}|^2. \tag{0.1.49}$$

Si $\lambda_{\mathrm{V}}-\frac{[\nabla\mathbf{W}]_1}{n}>0$, on peut poser $\lambda_n=\lambda_{\mathrm{V}}-\frac{[\nabla\mathbf{W}]_1}{n}$. Notons que

$$\exists n_0\in\mathbb{N}\quad \forall n\geq n_0,\quad \lambda_{\mathrm{V}}>\frac{[\nabla\mathbf{W}]_1}{n}. \tag{0.1.50}$$

En effet, il suffit de prendre la partie entière supérieure de $\frac{[\nabla\mathbf{W}]_1}{\lambda_{\mathrm{V}}}$. Ainsi quand le nombre de particules est suffisamment grand, on n'a pas besoin de comparer le confinement et l'interaction.

Concernant le cas cinétique, Bolley et al. ont étudié la dynamique

$$\begin{cases} d\mathrm{P}_t &= \mathrm{V}_t\,dt;\\ d\mathrm{V}_t &= \sqrt{2}d\mathrm{B}_t-\mathrm{A}(\mathrm{V}_t)\,dt-\mathrm{B}(\mathrm{P}_t)\,dt-\nabla\mathbf{W}*\mathrm{L}^{\mathrm{P}}_t(\mathrm{P}_t)\,dt;\end{cases} \tag{0.1.51}$$

avec A un champ de frictions (vitesses) et B un champ de confinement extérieur tels que

$$|\mathrm{A}(v_1)-\mathrm{A}(v_2)| \le [\mathrm{A}]_1 |v_1-v_2|; \tag{0.1.52}$$

$$\langle \mathrm{A}(v_1)-\mathrm{A}(v_2), v_1-v_2\rangle \ge \lambda_{\mathrm{A}} |v_1-v_2|^2; \tag{0.1.53}$$

$$\mathrm{B}(p) := \lambda_{\mathrm{B}} p + \mathrm{D}(p); \tag{0.1.54}$$

$$|\mathrm{D}(p_1)-\mathrm{D}(p_2)| \le [\mathrm{D}]_1 |p_1-p_2|; \tag{0.1.55}$$

$$|\nabla \mathbf{W}(p_1)-\nabla \mathbf{W}(p_2)| \le [\nabla \mathbf{W}]_1 |p_1-p_2|. \tag{0.1.56}$$

Sous ces conditions, les auteurs ont montré qu'on a ([41, Théorème 1])

$$\mathscr{W}_2(\mathbb{P}_{(\mathrm{P}_t,\mathrm{V}_t)}, \mathbb{P}_{(\overline{\mathrm{P}}_t,\overline{\mathrm{V}}_t)}) \le \mathrm{C}_1 e^{-\mathrm{C}_2 t} \mathscr{W}_2(\mathbb{P}_{(\mathrm{P}_0,\mathrm{V}_0)}, \mathbb{P}_{(\overline{\mathrm{P}}_0,\overline{\mathrm{V}}_0)}). \tag{0.1.57}$$

avec C_1 et C_2 des constantes positives. La preuve de cette contraction consiste en un changement de métrique sur l'espace des configurations : on établit un lemme de Gronwall avec une métrique induite par une forme quadratique puis on conclut par équivalence des normes en dimension finie. Cette idée fondamentale est expliquée dans la démonstration de [41, Proposition 8]. D'après [41, Lemme 10], il existe un unique équilibre et on a donc contraction vers ce dernier. Via cette convergence exponentielle, les auteurs ont prouvé un résultat de propagation du chaos forte quantitative uniforme en temps pour le système de particules sous-jacent ([41, Théorème 2]). Ils terminent en montrant une inégalité de Talagrand uniforme en le nombre de particules ([41, Théorème 5]) pour la distribution du système : inégalité qui permet de quantifier exponentiellement les déviations de la distribution empirique du système par rapport à l'équilibre de l'équation d'évolution non linéaire à un instant donné.

Dans les deux premiers chapitres de cette thèse, on établit des contractions exponentielles comme dans eq. (0.1.34), eq. (0.1.41) et eq. (0.1.57) pour des dynamiques de Langevin de champ moyen. L'approche directe utilisée ci-dessus (couplage synchrone) est le point de vue considéré dans le chapitre 2. En revanche, le chapitre 1 est connecté à l'approche indirecte qui est détaillée dans la sous-section suivante : via une *inégalité fonctionnelle uniforme* (en le nombre de particules) pour la mesure invariante du système de particules, on établit les contractions exponentielles.

0.1.2 Fonctionnelles de Lyapunov

Soit $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ une fonction et $y' = f(y)$ un *système dynamique*, avec y^* un *point d'équilibre* de ce système, c'est-à-dire que $f(y^*) = 0$. Par le changement de variable $z = y - y^*$,

on peut se ramener au cas où l'origine est un point d'équilibre ($f(0) = 0$).

Une fonction $V : \mathbb{R}^d \longrightarrow \mathbb{R}$ est appelée fonction *candidate de Lyapunov*[5] si, et seulement si, $V(0) = 0$ et il existe $\mathcal{V}$ un *voisinage épointé de l'origine* tel que pour tout $x \in \mathcal{V}$, $V(x) > 0$. On dit qu'une telle fonction est une *fonction de Lyapunov* si, et seulement si, $V(y)$ est une fonction décroissante du temps :

$$V(y)' := \nabla V(y) \cdot f(y) \leq 0. \tag{0.1.58}$$

Une fonction de Lyapunov est une fonction qui permet d'estimer la *stabilité d'un point d'équilibre* (ou, plus généralement, d'un mouvement, c'est-à-dire d'une solution maximale) d'une équation différentielle.

Dans cette thèse, on s'intéresse à des fonctionnelles qui permettent, comme dans le cas des équations différentielles, de mesurer la stabilité de $(L_t)_{t\geq 0}$ une solution de l'équation aux dérivées partielles d'évolution (au sens des distributions)

$$\partial_t L_t = L_t \mathscr{L}^{L_t}. \tag{0.1.59}$$

Définition 0.1.4 (Fonctionnelle de Lyapunov)**.** Une fonctionnelle $\mathbf{E} : \mathscr{E} \subset \mathscr{P}(\mathbb{R}^d) \to (-\infty, +\infty)$ est une fonctionnelle de Lyapunov pour une solution de l'équation d'évolution si, et seulement si,

1. $\forall\, t \geq 0, L_t \in \mathscr{E}$;
2. $\frac{d}{dt}\mathbf{E}[L_t] \leq 0$.

FLOT STATIONNAIRE ET LEMME DE GRONWALL

Définition 0.1.5 (État stationnaire)**.** $(L_t)_{t\geq 0}$ est un flot stationnaire si, et seulement si,

$$\forall\, t \geq 0, \quad L_t \mathscr{L}^{L_t} = 0. \tag{0.1.60}$$

Dans ce cas, L_0 est appelée mesure invariante.

Définition 0.1.6 (Bonne fonctionnelle de Lyapunov)**.** S'il existe une fonctionnelle de Lyapunov $\mathbf{E}$ et $\lambda > 0$ telles que

$$\forall\, t \geq 0, \quad \frac{d}{dt}\mathbf{E}[L_t] \leq -\lambda \mathbf{E}[L_t], \tag{0.1.61}$$

5. Terminologie issue de la théorie des systèmes dynamiques.

on dit que **E** est une bonne fonctionnelle de Lyapunov.

Par le lemme de Gronwall, pour une bonne fonctionnelle de Lyapunov, on a décroissance exponentielle de la fonctionnelle le long d'un flot solution de l'équation d'évolution :

$$\mathbf{E}[\mathrm{L}_t] \leq \mathbf{E}[\mathrm{L}_0] e^{-\lambda t}. \tag{0.1.62}$$

La recherche de telles fonctionnelles de Lyapunov pour les équations d'évolution est un problème fondamental pour l'étude du comportement en temps long. En effet, ces fonctionnelles peuvent être utilisées pour décrire la tendance d'un état de non-équilibre vers un état d'équilibre : elles sont donc d'une grande importance pour les équations aux dérivées partielles issues de la physique statistique (en particulier, la théorie cinétique). À ce sujet, une excellente référence est le livre de Villani sur l'hypocoercivité. De plus, il n'est généralement pas facile de trouver une bonne fonctionnelle de Lyapunov pour les équations non linéaires. Dans le contexte de cette thèse, les bonnes fonctionnelles de Lyapunov sont des fonctionnelles d'entropie.

FONCTIONNELLES D'ENTROPIE DE PROCESSUS DE DIFFUSION MARKOVIENS

On va partir de l'exemple fondamental des diffusions linéaires pour progressivement comprendre ce qu'il faut pour définir de bonnes fonctionnelles de Lyapunov pour les équations de Langevin de champ moyen. Concernant le comportement en temps long des équations paraboliques linéaires, la monographie de Bakry et al. sur le sujet est une excellence référence.

Soit $(\mathrm{P}_t)_{t\geq 0}$ le semi-groupe d'un processus de diffusion homogène $(\mathrm{Z}_t)_{t\geq 0}$:

$$d\mathrm{Z}_t = \sigma(\mathrm{Z}_t)\, d\mathrm{B}_t + b(\mathrm{Z}_t)\, dt. \tag{0.1.63}$$

En posant

$$a_{i,j}(\cdot) := \frac{1}{2}\sigma\sigma^*_{i,j}(\cdot), \tag{0.1.64}$$

par la formule d'Itô, son générateur infinitésimal est donné par

$$\mathscr{L} := \sum_{i,j} a_{i,j}(\cdot)\partial^2_{i,j} + \sum_i b_i(\cdot)\partial_i. \tag{0.1.65}$$

On suppose que ce générateur admet une mesure de probabilité invariante μ telle que $\mathscr{L}$ est auto-adjoint dans $\mathrm{L}^2(\mu)$: on dit que μ est une mesure invariante réversible. On suppose de

plus que le semi-groupe à tout instant strictement positif admet une densité par rapport à l'équilibre appartenant au domaine du générateur.

Théorème 0.1.1 (Variance et fonctionnelle de Lyapunov)**.** *La fonctionnelle*

$$\mathbf{E}_\mu : \nu \mapsto \begin{cases} \mathbb{V}_\mu[\frac{d\nu}{d\mu}] & \text{si} \quad \nu \ll \mu \quad \text{et} \quad \frac{d\nu}{d\mu} \in \mathrm{L}^2(\mu); \\ +\infty & \text{sinon.} \end{cases} \tag{0.1.66}$$

est une fonctionnelle de Lyapunov pour $(\nu \mathrm{P}_t)_{t\geq 0}$.

Démonstration. En effet, comme le générateur est auto-adjoint et

$$\mathscr{L}(fg) = f\mathscr{L}g + g\mathscr{L}f + 2\mathrm{A}(\cdot)\nabla f \cdot \nabla g, \quad \mathrm{A}(\cdot) := (a_{i,j}(\cdot))_{1\leq i,j\leq d}, \tag{0.1.67}$$

on a

$$\begin{aligned} \frac{d}{dt}\mathbf{E}_\mu[\nu \mathrm{P}_t] &= 2\mathbb{E}_\mu\Big[\frac{d\nu \mathrm{P}_t}{d\mu}\mathscr{L}\frac{d\nu \mathrm{P}_t}{d\mu}\Big] \\ &= -2\mathbb{E}_\mu\Big[\Big\|\nabla\frac{d\nu \mathrm{P}_t}{d\mu}\Big\|_{\mathrm{A}}^2\Big]. \end{aligned} \tag{0.1.68}$$

■

Définition 0.1.7 (Inégalité de Poincaré)**.** On dit que μ satisfait à une inégalité de Poincaré si, et seulement si,

$$\exists \lambda > 0 \quad \forall \varphi \in \mathrm{H}^1(\mu), \quad \lambda \mathbb{V}_\mu[\varphi] \leq \mathbb{E}_\mu[\|\nabla\varphi\|_{\mathrm{A}}^2]. \tag{0.1.69}$$

La constante optimale $\lambda(\mu)$ pour laquelle on a une telle inégalité est appelée la constante de Poincaré ou la constante de trou spectral : λ minore l'ensemble des valeurs propres non nulles de $-\mathscr{L}$.

Preuve de la minoration des valeurs propres non nulles. Vu que $-\mathscr{L}$ est auto-adjoint positif, son spectre est positif. Soit f une fonction propre de $-\mathscr{L}$ associée à la valeur propre $\lambda_f > 0$:

$$-\mathscr{L}f = \lambda_f f. \tag{0.1.70}$$

Comme

$$\begin{aligned} \int f d\mu &= -\frac{1}{\lambda_f}\int \mathscr{L}f d\mu \\ &= 0, \end{aligned} \tag{0.1.71}$$

on a

$$\begin{aligned}\int f^2 d\mu &= \mathbb{V}_\mu[f] \\ &\leq \frac{1}{\lambda}\int (-\mathscr{L} f) f\, d\mu \\ &= \frac{\lambda_f}{\lambda}\int f^2 d\mu.\end{aligned} \tag{0.1.72}$$

■

Proposition 0.1.1 (Inégalité de Poincaré et bonne fonctionnelle de Lyapunov)**.** *Si la mesure invariante satisfait à une inégalité de Poincaré, la fonctionnelle* $\mathbf{E}_\mu$ *est une bonne fonctionnelle de Lyapunov.*

Démonstration. On a par le lemme de Gronwall :

$$\mathbf{E}_\mu[\nu \mathrm{P}_t] \leq \mathbf{E}_\mu[\nu] e^{-2\lambda t}. \tag{0.1.73}$$

■

Théorème 0.1.2 (Énergie et fonctionnelle de Lyapunov)**.** *La fonctionnelle*

$$\mathbf{EE}_\mu : \nu \mapsto \begin{cases} \mathbb{E}_\mu[\|\nabla \frac{d\nu}{d\mu}\|_{\mathrm{A}}^2] & \text{si} \quad \nu \ll \mu \quad \text{et} \quad \frac{d\nu}{d\mu} \in \mathrm{H}^1(\mu); \\ +\infty & \text{sinon.} \end{cases} \tag{0.1.74}$$

est une fonctionnelle de Lyapunov pour $(\nu \mathrm{P}_t)_{t\geq 0}$.

Démonstration.

$$\begin{aligned}\frac{d}{dt}\mathbf{EE}_\mu[\nu \mathrm{P}_t] &= 2\mathbb{E}_\mu\Big[\mathrm{A}(\cdot)\nabla \frac{d\nu \mathrm{P}_t}{d\mu} \cdot \nabla \mathscr{L} \frac{d\nu \mathrm{P}_t}{d\mu}\Big] \\ &= -2\mathbb{E}_\mu\Big[\Big|\mathscr{L} \frac{d\nu \mathrm{P}_t}{d\mu}\Big|^2\Big].\end{aligned} \tag{0.1.75}$$

■

Proposition 0.1.2 (Inégalité de Poincaré duale et bonne fonctionnelle de Lyapunov)**.** *Si la mesure invariante satisfait à une inégalité de Poincaré, la fonctionnelle* $\mathbf{EE}_\mu$ *est une bonne fonctionnelle de Lyapunov.*

Démonstration. De l'analyse spectrale du générateur, on a

$$\lambda \mathbb{E}_\mu[\|\nabla\varphi\|_A^2] \le \mathbb{E}_\mu[|\mathscr{L}\varphi|^2]. \tag{0.1.76}$$

On en déduit par le lemme de Gronwall :

$$\mathbf{EE}_\mu[\nu P_t] \le \mathbf{EE}_\mu[\nu] e^{-2\lambda t}. \tag{0.1.77}$$

■

L'inégalité de Poincaré permet donc d'avoir des convergences exponentielles dans $L^2(\mu)$ et $H^1(\mu)$.

Théorème 0.1.3 (Entropie relative et fonctionnelle de Lyapunov)**.** *La fonctionnelle*

$$\mathbf{H}[\cdot|\mu] : \nu \mapsto \begin{cases} \mathbb{E}_\nu[\ln \frac{d\nu}{d\mu}] & \text{si} \quad \nu \ll \mu \quad \text{et} \quad \frac{d\nu}{d\mu} \ln \frac{d\nu}{d\mu} \in L^1(\mu); \\ +\infty & \text{sinon.} \end{cases} \tag{0.1.78}$$

est une fonctionnelle de Lyapunov pour $(\nu P_t)_{t\ge 0}$.

Démonstration. En effet, comme le générateur est auto-adjoint, la densité par rapport à l'équilibre

$$h_t := \frac{d\nu P_t}{d\mu} \tag{0.1.79}$$

est donnée par

$$\partial_t h_t = \mathscr{L} h_t. \tag{0.1.80}$$

Comme

$$\mathscr{L}(fg) = f\mathscr{L}g + g\mathscr{L}f + 2A(\cdot)\nabla f \cdot \nabla g, \tag{0.1.81}$$

on en déduit que

$$\begin{aligned} \frac{d}{dt}\mathbf{H}[\mu_t|\mu] &= \frac{d}{dt}\int h_t \ln h_t d\mu \\ &= \int (\mathscr{L} h_t) \ln h_t d\mu \\ &= -\int \frac{\|\nabla h_t\|_A^2}{h_t} d\mu. \end{aligned} \tag{0.1.82}$$

■

Définition 0.1.8 (Inégalité de Sobolev logarithmique)**.** On dit que μ satisfait à une inégalité de Sobolev logarithmique si, et seulement si,

$$\exists \rho_{LS} > 0, \quad \rho_{LS}\mathbf{H}[\cdot|\mu] \leq 2\mathbf{I}[\cdot|\mu] \tag{0.1.83}$$

avec

$$\mathbf{I}[\cdot|\mu] : \nu \mapsto \begin{cases} \mathbb{E}_\mu[\|\nabla\sqrt{\frac{d\nu}{d\mu}}\|_A^2] & \text{si} \quad \nu \ll \mu \quad \text{et} \quad \sqrt{\frac{d\nu}{d\mu}} \in H^1(\mu); \\ +\infty & \text{sinon.} \end{cases} \tag{0.1.84}$$

l'information de Fisher relative. La constante optimale $\rho_{LS}(\mu)$ pour laquelle on a une telle inégalité est appelée constante de Sobolev logarithmique.

Proposition 0.1.3 (Inégalité de Sobolev logarithmique et bonne fonctionnelle de Lyapunov)**.** *Si la mesure invariante satisfait à une inégalité de Sobolev logarithmique, la fonctionnelle* $\mathbf{H}$ *est une bonne fonctionnelle de Lyapunov.*

Démonstration. On a par le lemme de Gronwall :

$$\mathbf{H}[\nu P_t|\mu] \leq \mathbf{H}[\nu|\mu]e^{-2\rho_{LS}t}. \tag{0.1.85}$$

■

L'inégalité de Sobolev logarithmique implique une inégalité de Poincaré : la convergence exponentielle en entropie implique celle en variance. De plus, l'inégalité de Sobolev logarithmique implique une T2-inégalité de Talagrand :

$$\rho_{LS}\mathcal{W}_2^2(\cdot,\mu) \leq 2\mathbf{H}[\cdot|\mu]. \tag{0.1.86}$$

Cette inégalité de Talagrand implique une convergence exponentielle en métrique de Wasserstein : pour plus de détails à ce sujet, le livre de Villani sur le transport optimal est une excellente référence.

L'une des propriétés les plus importantes des inégalités fonctionnelles de Poincaré et de Sobolev logarithmique est leur *stabilité par produit tensoriel* : si μ_a et μ_b satisfont à une inégalité fonctionnelle, alors $\mu_a \otimes \mu_b$ satisfait à une inégalité fonctionnelle de constante la plus grande des deux constantes. Cette propriété est très utilisée en *mécanique statistique* : par exemple, dans l'étude du comportement en temps long des systèmes de particules de champ moyen.

Exemple 0.1.2 (Équation de Langevin standard). On considère l'équation de Fokker-Planck standard d'hamiltonien $V : \mathbb{R}^d \to (-\infty, +\infty)$:

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot (\mu_t \nabla V). \tag{0.1.87}$$

Il est bien connu que sous des hypothèses adéquates de régularité sur V, cette équation d'évolution admet une unique mesure invariante : la mesure de Gibbs $\mu_V(dx) := \frac{1}{Z_V} e^{-V(x)} dx$.

— Si μ_V est log-concave (V convexe), alors elle satisfait à une inégalité de Poincaré.

— S'il existe $\alpha > 0$ et $R \geq 0$ telles que pour tout $|x| \geq R$, $x \cdot \nabla V(x) \geq \alpha |x|$, alors μ_V satisfait à une inégalité de Poincaré.

— S'il existe $a \in (0,1)$, $c > 0$ et $R \geq 0$ telles que pour tout $|x| \geq R$, $a|\nabla V(x)|^2 - \Delta V(x) \geq c$, alors μ_V satisfait à une inégalité de Poincaré.

— S'il existe deux constantes positives c_1 et c_2 telles que $x \cdot \nabla V(x) \geq c_1 |x|^2 - c_2$ (condition de stabilité), et la hessienne $\nabla^2 V$ de V est minorée, alors μ_V satisfait à une inégalité de Sobolev logarithmique.

— Si V vérifie le critère de courbure dimension de Bakry and Émery ($\nabla^2 V \geq \rho \mathbf{Id}$ au sens des formes quadratiques), μ_V satisfait à une inégalité de Sobolev logarithmique. L'idée de Bakry and Émery est de comparer la dissipation d'entropie à la dissipation de dissipation d'entropie pour obtenir l'inégalité de Sobolev logarithmique, c'est-à-dire un contrôle de l'entropie par sa dissipation. Nous verrons dans la suite que cette idée est très fructueuse puisqu'elle fournit encore le même type de résultat pour des équations non linéaires.

— (Holley-Stroock) Si $V = V_1 + V_2$ avec μ_{V_1} satisfaisant à une inégalité de Sobolev logarithmique et V_2 à *oscillations bornées*, alors μ_V satisfait à une inégalité de Sobolev logarithmique.

— Si $V = V_1 + V_2$ avec μ_{V_1} satisfaisant à une inégalité de Sobolev logarithmique et V_2 est lipschitzienne, alors d'après le théorème de perturbation de Aida and Shigekawa, μ_V satisfait à une inégalité de Sobolev logarithmique.

FONCTIONNELLES D'ENTROPIE DE TYPE LANGEVIN DE CHAMP MOYEN

On arrive à l'un des deux objectifs principaux de cette sous-section : un procédé de construction de bonnes fonctionnelles de Lyapunov pour l'équation de Langevin de champ moyen.

L'idée fondamentale est d'associer à une équation d'évolution non linéaire une métrique sur la variété riemaniennne des mesures de probabilité et une fonctionnelle d'entropie qui caractérise l'équilibre. Plus précisément, l'évolution définit un gradient flot relativement à la fonctionnelle de Lyapunov : le livre de Ambrosio et al. sur les gradients flots sur des espaces de mesures et celui de Villani sur le transport optimal sont d'excellentes références.

Pour faire apparaître la métrique et la fonctionnelle appropriée, il faut réécrire l'équation de Langevin de champ moyen sous la forme

$$\partial_t \rho_t = -\nabla \cdot \Big(\rho_t \nabla p\Big); \tag{0.1.88}$$

$$p := -\ln \rho_t - \frac{\delta \mathrm{H}}{\delta m}(\mu_t, \cdot). \tag{0.1.89}$$

On peut alors suivre l'approche adoptée par Otto pour l'équation des milieux poreux. Il s'agit de considérer l'ensemble $\mathscr{P}$ des mesures de probabilité absolument continues par rapport à la mesure de Lebesgue comme une variété de dimension infinie. L'espace tangent à $\mathscr{P}$ en ρ est

$$\mathrm{T}_\rho \mathscr{P} := \Big\{ s \in \mathrm{L}^1(\mathbb{R}^d), \quad \int s = 0 \Big\}. \tag{0.1.90}$$

On identifie à présent $\mathrm{T}_\rho \mathscr{P}$ à l'ensemble des fonctions intégrables quotienté par la relation d'équivalence "égales à une constante près" à l'équation elliptique

$$-\nabla \cdot \Big(\rho \nabla p\Big) = s. \tag{0.1.91}$$

La fonction p est bien définie vu qu'on résout

$$\Delta p + \nabla p \cdot \nabla \ln \rho = -\frac{s}{\rho} \tag{0.1.92}$$

avec $\frac{s}{\rho}$ d'intégrale nulle par rapport à ρ. On définit alors le tenseur g et la fonctionnelle $\mathbf{E}_\mathrm{H}$:

$$\begin{aligned} g_\rho(s_1, s_2) &= \int \rho \nabla p_1 \cdot \nabla p_2 \\ &= -\int p_1 \nabla \cdot \Big(\rho \nabla p_2\Big) \\ &= \int p_1 s_2. \end{aligned} \tag{0.1.93}$$

$$\mathbf{E}_\mathrm{H}[\rho] := \mathrm{H}(\rho) + \int \rho \ln \rho. \tag{0.1.94}$$

Le tenseur g transforme la différentielle $\mathbf{DiffE}_{\mathrm{H}}$ de $\mathbf{E}_{\mathrm{H}}$, qui est un champ de vecteurs cotangents en son gradient $\mathbf{GradE}_{\mathrm{H}}$, qui est un champ de vecteurs tangents :

$$g_\rho(\mathbf{GradE}_{\mathrm{H}}, s) = \mathbf{DiffE}_{\mathrm{H}} \cdot s. \tag{0.1.95}$$

Or, on a, pour tout vecteur tangent s_1,

$$\begin{aligned}\mathbf{DiffE}_{\mathrm{H}} \cdot s_1 &= \int \frac{\delta \mathbf{E}_{\mathrm{H}}}{\delta m}(\rho, \cdot) s_1 \\ &= \int \left(\ln \rho + \frac{\delta \mathrm{H}}{\delta m}(\rho, \cdot)\right) s_1 \\ &= -\int p s_1 \\ &= g_\rho\left(\nabla \cdot \left(\rho \nabla p\right), s_1\right) \\ &= -g_\rho(\partial_t \rho, s_1).\end{aligned} \tag{0.1.96}$$

L'équation de Langevin de champ moyen s'interprète donc comme la solution d'un flot gradient :

$$\partial_t \rho = -\mathbf{GradE}_{\mathrm{H}}(\rho). \tag{0.1.97}$$

Théorème 0.1.4 (Fonctionnelle de Lyapunov de type Langevin de champ moyen)**.** $\mathbf{E}_{\mathrm{H}}$ *est une fonctionnelle de Lyapunov pour l'évolution de Langevin de champ moyen.*

Démonstration.

$$\begin{aligned}\frac{d}{dt}\mathbf{E}_{\mathrm{H}}[\rho_t] &= -g_\rho(\partial_t \rho, \partial_t \rho) \\ &= -\int \rho |\nabla p|^2 \\ &= -\int \rho \left| \frac{\nabla \rho}{\rho} + \nabla \frac{\delta \mathrm{H}}{\delta m}(\rho, \cdot) \right|^2.\end{aligned} \tag{0.1.98}$$

■

La fonctionnelle

$$\mathbf{I}_{\mathrm{H}} : \rho \mapsto \int \rho \left| \frac{\nabla \rho}{\rho} + \nabla \frac{\delta \mathrm{H}}{\delta m}(\rho, \cdot) \right|^2 \tag{0.1.99}$$

est appelée *dissipation d'entropie* ou à une constante près [6] l'*information de Fisher de champ moyen.*

6. Quatre fois l'information de Fisher de champ moyen.

Proposition 0.1.4 (Fonctionnelle de Lyapunov de champ moyen et mesures invariantes)**.** *Les éventuels minimiseurs de la fonctionnelle de Lyapunov de champ moyen sont nécessairement des mesures invariantes de l'évolution de Langevin de champ moyen.*

Démonstration. Le tenseur métrique peut être défini par polarisation à partir de

$$g_\rho(s,s) = \inf_{\substack{u \\ s=-\nabla\cdot(\rho u)}} \int \rho |u|^2. \tag{0.1.100}$$

Cette remarque est très importante : du point de vue physique, elle traduit la minimisation de l'énergie. De plus, elle indique que le tenseur g munit la variété $\mathscr{P}$ d'une structure riemannienne mathématiquement satisfaisante : il est possible d'identifier les géodésiques et la distance riemanienne associées en terme de transport de mesures et de distance de Wasserstein. ■

En particulier, comme $\rho \mapsto \int \rho \ln \rho$ est strictement convexe, si l'hamiltonien macroscopique H est convexe, il y a au plus une mesure invariante : pas de transition de phase.

Définition 0.1.9 (Inégalité log-Sobolev non linéaire de type Langevin de champ moyen)**.** On dit qu'on a une inégalité log-Sobolev non linéaire de type Langevin de champ moyen si, et seulement si,

$$\exists \varrho > 0, \quad 2\varrho \mathbf{E}_{\mathrm{H}} \leq \mathbf{I}_{\mathrm{H}}. \tag{0.1.101}$$

Proposition 0.1.5 (Log-Sob et bonne fonctionnelle de Lyapunov de champ moyen)**.** *Si* $\mathbf{E}_{\mathrm{H}}$ *satisfait à une inégalité log-Sobolev non linéaire, alors c'est une bonne fonctionnelle de Lyapunov.*

Démonstration. La preuve de cette proposition est une conséquence directe du travail effectué en amont. En effet, comme

$$\frac{d}{dt}\mathbf{E}_{\mathrm{H}}[\rho_t] = -\mathbf{I}_{\mathrm{H}}[\rho_t], \tag{0.1.102}$$

par le lemme de Gronwall, on a

$$\mathbf{E}_{\mathrm{H}}[\rho_t] \leq \mathbf{E}_{\mathrm{H}}[\rho_0] e^{-2\varrho t}. \tag{0.1.103}$$

■

Exemple 0.1.3 (Équation des milieux granulaires). Carrillo et al. ont montré que pour

$$\mathrm{H} : \mu \mapsto \int \mathrm{V} d\mu + \frac{1}{2} \iint \mathbf{W}(x-y)\mu(dx)\mu(dy), \tag{0.1.104}$$

$\mathbf{E}_H - \inf \mathbf{E}_H$ est une bonne fonctionnelle de Lyapunov pour l'équation des milieux granulaires. Les auteurs se sont inspirés de l'idée fondamentale de Bakry and Émery : comparer la dissipation d'entropie $\mathbf{I}_H$ avec sa dissipation.

Concernant le cas cinétique, l'article de Villani [7] sur l'information de Fisher en théorie cinétique est une excellente référence. Le cas cinétique est plus compliqué pour deux raisons décrites ci-dessous.

— On a une diffusion dégénérée dans laquelle le laplacien n'agit que sur la variable de vitesse.

— Il n'y a pas de structure de flot gradient mais on a simultanément une dynamique hamiltonienne et un flot gradient.

FONCTIONNELLES MICROSCOPIQUES ET MACROSCOPIQUES

On arrive enfin à l'objectif final de cette sous-section : la description détaillée de l'approche indirecte.

Soit un système physique constitué d'un très grand nombre de petites particules dont le libre parcours moyen est grand par rapport à leur taille. Il existe alors plusieurs niveaux de description. L'*échelle macroscopique* constitue l'échelle accessible facilement expérimentalement et pour laquelle on peut faire des mesures de quantités caractéristiques telles que le volume, la température, la pression, etc. L'*échelle microscopique* correspond à l'échelle des composants élémentaires de la matière (atomes, molécules). On peut aussi introduire une description intermédiaire correspondant à une description statistique. Il s'agit alors de l'*échelle mésoscopique*. À ce niveau, on ne s'intéresse plus à l'évolution des particules individuellement mais à l'évolution de la densité de particules (propagation du chaos).

Par passage à la limite sur le nombre de particules, sous des hypothèses adéquates, la limite de champ moyen peut espérer hériter la propriété d'ergodicité du système de particules : plus précisément, il est question de conditions sur les systèmes de particules assurant l'existence d'une suite de bonnes fonctionnelles de Lyapunov associée aux systèmes de particules convergeant vers une bonne fonctionnelle de Lyapunov pour la limite de champ

7. Villani : In several respects these lectures are the sequel of my course at Institut Henri Poincaré in 2001, Entropy production and convergence to equilibrium [175], but this is also my first research memoir after ten years of leave. So I consider it a key step bridging my past and future research. Awkward as this may seem, on this occasion I consider it fair to thank the french politicians (EM, PM, JLM) who ensured, willingly or unwillingly, my defeat at the 2022 French Parliamentary elections, even though by a statistically nonsignificant margin, making way for my comeback in research.

moyen. L'intérêt de cette approche réside dans le fait qu'elle demande moins de régularité à l'équation d'évolution non linéaire : on fait tous les calculs sur l'équation d'évolution du système de particules, ce qui coûte moins cher en régularité.

Supposons que la mesure invariante du système de particules satisfait à une inégalité de Sobolev logarithmique uniforme (en le nombre de particules) :

$$\exists \varrho > 0 \quad \forall n \geq 1, \quad \varrho \mathbf{H}[\cdot|\mu_\infty^{(n)}] \leq 2\mathbf{I}[\cdot|\mu_\infty^{(n)}]. \tag{0.1.105}$$

Autrement dit, on a convergence exponentielle en entropie uniforme :

$$\mathbf{H}[\mathbb{P}_{X_t^{(n)}}|\mu_\infty^{(n)}] \leq \mathbf{H}[\mathbb{P}_{X_0^{(n)}}|\mu_\infty^{(n)}]e^{-2\varrho t}. \tag{0.1.106}$$

On a une T2-inégalité de Talagrand uniforme :

$$\varrho \mathcal{W}_2^2(\cdot, \mu_\infty^{(n)}) \leq 2\mathbf{H}[\cdot|\mu_\infty^{(n)}]. \tag{0.1.107}$$

Par propriété de tensorisation de la métrique de Wasserstein, on a

$$\sum_{i=1}^n \mathcal{W}_2^2(\mathbb{P}_{X_t^{(n),i}}, \mu_\infty^{(n),i}) \leq \mathcal{W}_2^2(\mathbb{P}_{X_t^{(n)}}, \mu_\infty^{(n)}). \tag{0.1.108}$$

Par symétries du système de particules et de sa mesure invariante, on a

$$\mathcal{W}_2^2(\mathbb{P}_{X_t^{(n),1}}, \mu_\infty^{(n),1}) \leq \frac{1}{n}\mathcal{W}_2^2(\mathbb{P}_{X_t^{(n)}}, \mu_\infty^{(n)}). \tag{0.1.109}$$

FIGURE 1 – Échelles en physique statistique

On en déduit que

$$\mathscr{W}_2^2(\mathbb{P}_{\mathrm{X}_t^{(n),1}}, \mu_\infty^{(n),1}) \le \frac{2}{\varrho}\frac{1}{n}\mathbf{H}[\mathbb{P}_{\mathrm{X}_0^{(n)}}|\mu_\infty^{(n)}]e^{-2\varrho t}. \tag{0.1.110}$$

La propagation du chaos assure la convergence de $((\mathbb{P}_{\mathrm{X}_t^{(n),1}})_{t\ge 0})_{n\ge 1}$ vers $(\mathrm{L}_t^{\mathrm{X}})_{t\ge 0}$. Quant aux autres quantités mises en jeu dans l'inégalité, elles sont plus délicates à étudier [8]. Dans cette thèse, on utilise la *théorie des grandes déviations* en montrant que la fonctionnelle de Lyapunov $\mathbf{E}_{\mathrm{H}}$ est une *bonne fonction de taux* de convergence pour un principe de grandes déviations de type Sanov pour la mesure empirique du système : ses ensembles de niveau sont compacts (existence d'un minimiseur) et quand $n \to +\infty$,

$$\mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^{n}\delta_{\mathrm{X}^{(n),i}} \in \mathscr{O}\Big) \simeq e^{-n\inf_{\mathscr{O}}\mathbf{E}_{\mathrm{H}}}. \tag{0.1.111}$$

Ce principe de grandes déviations permet de montrer que

$$\frac{1}{n}\mathbf{H}[\mathbb{P}_{\mathrm{X}_0^{(n)}}|\mu_\infty^{(n)}] \overset{n\to+\infty}{\longrightarrow} \mathbf{E}_{\mathrm{H}}[\mathrm{L}_0^{\mathrm{X}}] - \inf\mathbf{E}_{\mathrm{H}}. \tag{0.1.112}$$

Et si la mesure invariante est unique, on a

$$\mu_\infty^{(n),1} \overset{n\to+\infty}{\longrightarrow} \mathrm{L}_\infty^{\mathrm{X}}. \tag{0.1.113}$$

Par semi-continuité de la métrique de Wasserstein, on en déduit que

$$\mathscr{W}_2^2(\mathrm{L}_t^{\mathrm{X}}, \mathrm{L}_\infty^{\mathrm{X}}) \le \frac{2}{\varrho}\Big(\mathbf{E}_{\mathrm{H}}[\mathrm{L}_0^{\mathrm{X}}] - \inf\mathbf{E}_{\mathrm{H}}\Big)e^{-2\varrho t} \tag{0.1.114}$$

Par la loi des grands nombres, on montre que

$$\frac{1}{n}\mathbf{I}[\mathbb{P}_{\mathrm{X}_0^{(n)}}|\mu_\infty^{(n)}] \overset{n\to+\infty}{\longrightarrow} \frac{1}{4}\mathbf{I}_{\mathrm{H}}[\mathrm{L}_0^{\mathrm{X}}]. \tag{0.1.115}$$

On en déduit qu'on a des inégalités de Sobolev logarithmique et de Talagrand non linéaires :

$$2\varrho(\mathbf{E}_{\mathrm{H}} - \inf\mathbf{E}_{\mathrm{H}}) \le \mathbf{I}_{\mathrm{H}}; \tag{0.1.116}$$

$$\varrho\mathscr{W}_2^2(\cdot, \mathrm{L}_\infty^{\mathrm{X}}) \le 2(\mathbf{E}_{\mathrm{H}} - \inf\mathbf{E}_{\mathrm{H}}), \tag{0.1.117}$$

ce qui assure que la fonctionnelle de Lyapunov $\mathbf{E}_{\mathrm{H}} - \inf\mathbf{E}_{\mathrm{H}}$ est une bonne fonctionnelle de

8. Au sens de la figure 3 introduite plus loin.

Lyapunov :

$$\mathbf{E}_{\mathrm{H}}[\mathrm{L}_t^{\mathrm{X}}] - \mathbf{E}_{\mathrm{H}}[\mathrm{L}_\infty^{\mathrm{X}}] \leq \Big(\mathbf{E}_{\mathrm{H}}[\mathrm{L}_0^{\mathrm{X}}] - \mathbf{E}_{\mathrm{H}}[\mathrm{L}_\infty^{\mathrm{X}}]\Big) e^{-2\varrho t}. \tag{0.1.118}$$

Notons que dans le cas cinétique, les résultats se déduisent du cas non cinétique par les structures particulières du générateur (*forme de Hormander*), de la mesure invariante (produit tensoriel de la mesure invariante du non cinétique avec une gaussienne centrée réduite), et la théorie de l'*hypocoercivité* de Villani (condition de bornitude).

Exemple 0.1.4 (Équation des milieux granulaires). Malrieu a montré la convergence exponentielle en métrique de Wasserstein en utilisant la propagation du chaos et une inégalité fonctionnelle uniformes. Guillin et al. ont montré cette convergence en utilisant une inégalité fonctionnelle uniforme, la propagation du chaos et les grandes déviations.

0.1.3 Distributions quasi-stationnaires

Dans cette sous-section, il est question de la mise en contexte de l'objet du chapitre 3.

GÉNÉRALITÉS SUR LES QSDS DE PROCESSUS DE MARKOV ABSORBÉS

Dans le cadre de cette théorie, on considère un processus de Markov évoluant dans un domaine où il y a un ensemble d'états interdits qui constitue un piège. On dit que le processus est tué lorsqu'il atteint un ensemble de tels états et on suppose que cela se produit presque sûrement. Nous étudions le comportement du processus avant d'être tué, plus précisément nous étudions ce qui se passe lorsqu'on conditionne le processus à survivre longtemps. Fixons quelques notations. On considère un espace d'états $\mathscr{E}$ muni d'une structure d'espace mesurable (par exemple, un espace métrique), $(Y_t)_{t\geq 0}$ un processus de Markov à valeurs dans $\mathscr{E}$ et $(\mathbb{P}_x)_{x\in\mathscr{E}}$ la famille des distributions induites par la condition initiale. Comme dit, il existe un ensemble mesurable d'états interdits que nous désignons par $\partial\mathscr{E}$ que nous noterons parfois $\mathscr{E}^{tr}$ car il ne coïncide pas forcément avec la frontière de $\mathscr{E}$). Son complémentaire $\mathscr{E}^a := \mathscr{E} - \partial\mathscr{E}$ est l'ensemble des états autorisés. Soit $\tau := \tau_{\partial\mathscr{E}}$ le temps d'atteinte de $\partial\mathscr{E}$. On suppose que le processus est presque sûrement tué, c'est-à-dire pour tout $x \in \mathscr{E}^a$,

$$\mathbb{P}_x(\tau < +\infty) = 1. \tag{0.1.119}$$

Nous nous intéresserons principalement à $(Y_t)_{t<\tau}$ la trajectoire avant extinction, donc sans perte de généralité, on peut supposer $Y_t = Y_\tau$ lorsque $t \geq \tau$, c'est-à-dire que l'ensemble des états interdits est absorbant (ou un piège). Parmi les principaux concepts de notre étude figurent les distributions quasi-stationnaires (QSDs abréviation en anglais pour faire court), ce sont des distributions qui sont invariantes dans le temps lorsque le processus est conditionné à survivre. Plus précisément,

Définition 0.1.10. Une mesure de probabilité ν sur $\mathscr{E}^a$ est appelée une QSD (pour le processus tué) si pour tout ensemble mesurable $B \subset \mathscr{E}^a$,

$$\forall\, t \geq 0, \quad \mathbb{P}_\nu(Y_t \in B | \tau > t) = \nu(B) \quad \text{avec} \quad \mathbb{P}_\nu(\cdot) := \int \mathbb{P}_x(\cdot)\nu(dx). \tag{0.1.120}$$

Le fait que le processus est presque sûrement tué (eq. (0.1.119)) est aussi équivalent à

$$\forall\, \mu \in \mathscr{P}(\mathscr{E}^a), \quad \mathbb{P}_\mu(\tau < +\infty) = 1. \tag{0.1.121}$$

Alors, ν est une QSD si

$$\forall (B,t) \in \mathscr{B}(\mathscr{E}^a) \times [0,+\infty[, \quad \mathbb{P}_\nu(Y_t \in B, \tau > t) = \mathbb{P}_\nu(\tau > t)\nu(B). \tag{0.1.122}$$

Vu que Y est arrêté à τ et que d'après eq. (0.1.119),

$$\forall \nu \in \mathscr{P}(\mathscr{E}^a) \quad \exists t(\mu) \geq 0 \quad \forall t \geq t(\mu), \quad \mathbb{P}_\mu(\tau > t) < 1, \tag{0.1.123}$$

on en déduit respectivement que $\mathbb{P}_\nu(Y_t \in B, \tau > t) = \mathbb{P}_\nu(Y_t \in B)$,

$$\forall (B,t) \in \mathscr{B}(\mathscr{E}^a) \times [0,+\infty[, \quad \mathbb{P}_\nu(Y_t \in B) = \mathbb{P}_\nu(\tau > t)\nu(B), \tag{0.1.124}$$

et donc, il ne peut y avoir existence de distribution stationnaire supportée par $\mathscr{E}^a$ qui soit une QSD pour un tel processus...

Loi de τ sous ν Soit ν une QSD. La première remarque est qu'en partant de ν, le temps de mise à mort τ est distribué de manière exponentielle :

Théorème 0.1.5 (Loi du temps de mise à mort sous la QSD)**.**

$$\exists \lambda(\nu) > 0 \quad \forall t \geq 0, \quad \mathbb{P}_\nu(\tau > t) = e^{-\lambda(\nu)t}. \tag{0.1.125}$$

Autrement dit, partant de ν, $\tau \sim \mathscr{E}(\lambda(\nu))$.

Démonstration. Vu que ν est une QSD, eq. (0.1.122) assure que

$$\forall g \in \mathscr{M}_+(\mathscr{E}^a) \bigcap \mathscr{M}_b(\mathscr{E}^a) \quad \forall t \geq 0, \quad \mathbb{E}_\nu[g(Y_t)\mathbb{1}_{\tau > t}] = \mathbb{E}_\nu[g]\mathbb{P}_\nu(\tau > t). \tag{0.1.126}$$

En particulier, pour $g : x \in \mathscr{E}^a \mapsto \mathbb{P}_x(\tau > s)$, comme $\int g(x)\nu(dx) = \mathbb{P}_\nu(\tau > s)$, on en déduit que

$$\forall t \geq 0, \quad \mathbb{E}_\nu[\mathbb{P}_{Y_t}(\tau > s)\mathbb{1}_{\tau > t}] = \mathbb{P}_\nu(\tau > s)\mathbb{P}_\nu(\tau > t). \tag{0.1.127}$$

Par la PROPRIÉTÉ DE MARKOV, on a

$$\begin{aligned}\forall (t,s)\in[0,+\infty[\times[0,+\infty[,\quad \mathbb{P}_\nu(\tau>t+s) &:= \mathbb{E}_\nu[\mathbb{I}_{\tau>t+s}] \\ &= \mathbb{E}_\nu[\mathbb{I}_{\tau>t}\mathbb{E}[\mathbb{I}_{\tau>t+s}|\mathscr{F}_t]] \\ &= \mathbb{E}_\nu[\mathbb{I}_{\tau>t}\mathbb{E}_{Y_t}[\mathbb{I}_{\tau>s}]] \\ &= \mathbb{P}_\nu(\tau>t)\mathbb{P}_\nu(\tau>s).\end{aligned} \tag{0.1.128}$$

En notant $\psi : t\mapsto \mathbb{P}_\nu(\tau>t)$ la fonction de survie du temps de mise à mort τ sous la QSD ν, on en déduit la relation fonctionnelle

$$\begin{cases}\psi(t+s) & =\psi(t)\psi(s) \quad \forall (t,s)\in[0,\infty[\times[0,+\infty[;\\ \psi(0) & =1;\\ \lim_{t\to+\infty}\psi(t) & =0.\end{cases} \tag{0.1.129}$$

Toute fonction f continue non identiquement nulle telle que $f(t+s)=f(t)f(s)$ est de la forme $x\mapsto e^{\theta x}$ avec $\theta\in\mathbb{R}$ une constante. Comme celle qu'on cherche admet une limite nulle à l'infini, on en déduit que $\theta := -\lambda(\nu)<0$. Ce qui achève la preuve. ■

L'analogue discret de ce théorème est donné par l'équation

$$\exists \kappa(\nu)\in]0,1[,\quad \mathbb{P}_\nu(\tau>n)=\kappa(\nu)^n \quad \text{et de plus} \quad \kappa(\nu)=\mathbb{P}_\nu(\tau>1). \tag{0.1.130}$$

Quant à $\lambda(\nu)$, on a par inversion

$$\forall\, t>0,\quad \lambda(\nu)=-\frac{1}{t}\ln\mathbb{P}_\nu(\tau>t)\in]0,+\infty[. \tag{0.1.131}$$

Alors, ν est une QSD si, et seulement si

$$\exists\theta\in(0+\infty)\quad \forall(B,t)\in\mathscr{B}(\mathscr{E}^a)\times[0,+\infty[,\quad \mathbb{P}_\nu(Y_t\in B,\tau>t)=e^{-\theta t}\nu(B). \tag{0.1.132}$$

Dans ce cas, $\theta=\lambda(\nu)$.

Moments exponentiels de τ sous ν

$$\forall\theta<\lambda(\nu),\quad \mathbb{E}_\nu[e^{\theta\tau}]<+\infty. \tag{0.1.133}$$

Ce qui implique que

$$\forall \theta < \lambda(\nu), \quad \text{pour } \nu\text{-presque tout } x, \quad \mathbb{E}_x[e^{\theta\tau}] < +\infty. \tag{0.1.134}$$

La condition sur les moments exponentiels se réécrit

Théorème 0.1.6. *Nous avons l'égalité*

$$\theta_x^* := \sup\left\{\theta, \quad \mathbb{E}_x[e^{\theta\tau}] < +\infty\right\} = \liminf_{t\to+\infty} -\frac{1}{t}\ln \mathbb{P}_x(\tau > t), \tag{0.1.135}$$

et une condition nécessaire à l'existence d'une QSD est l'existence d'un moment exponentiel, ou de manière équivalente, un taux de survie exponentiel positif :

$$\exists x \in \mathscr{E}^a, \quad \theta_x^* > 0. \tag{0.1.136}$$

Notons que si ν est une QSD, alors $\mathbb{E}_\nu[e^{\lambda(\nu)\tau}] = +\infty$. Donc, si $\theta > 0$ satisfait la condition

$$\sup_{x\in\mathscr{E}^a} \mathbb{E}_x[e^{\theta\tau}] < +\infty, \tag{0.1.137}$$

il n'existe pas de QSD ν telle que $\lambda(\nu) = \theta$ sinon, on aurait

$$+\infty = \mathbb{E}_\nu[e^{\lambda(\nu)\tau}] \le \sup_{x\in\mathscr{E}^a} \mathbb{E}_x[e^{\lambda(\nu)\tau}] < +\infty. \tag{0.1.138}$$

Indépendance entre le temps de sortie et l'état de sortie Soit ν une QSD (pour le processus tué). Alors τ et Y_τ sont des variables aléatoires $\mathbb{P}_\nu$-indépendantes, c'est-à-dire

$$\forall (a, B) \in (0,+\infty) \times \partial\mathscr{E}, \quad \mathbb{P}_\nu(\tau \le a, Y_\tau \in B) = \mathbb{P}_\nu(\tau \le a)\mathbb{P}_\nu(Y_\tau \in B). \tag{0.1.139}$$

De cette indépendance, découle

$$\frac{d}{dt}\mathbb{P}_\nu(Y_t \in B)\Big|_{t=0} = \lambda(\nu)\mathbb{P}_\nu(Y_\tau \in B). \tag{0.1.140}$$

Caractérisation de QSDs par les semi-groupes Soit $(T_t)_{t\ge0}$ le semi-groupe du processus avant absorption, c'est-à-dire

$$\forall \varphi \in \mathscr{M}_b(\mathscr{E}^a) \quad \forall t \ge 0 \quad \forall x \in \mathscr{E}^a, \quad T_t\varphi(x) := \mathbb{E}_x[\varphi(Y_t), \tau > t] = \mathbb{E}_x[\varphi(Y_t)\mathbb{I}_{\tau>t}]. \tag{0.1.141}$$

La propriété de Markov implique la relation de semi-groupe, c'est-à-dire

$$\forall (t,s) \in [0,+\infty[^2, \quad \mathrm{T}_{t+s} = \mathrm{T}_t \circ \mathrm{T}_s = \mathrm{T}_s \circ \mathrm{T}_t. \tag{0.1.142}$$

Notons que, pour tout $x \in \mathscr{E}^a$ et $t \geq 0$, $\mathrm{T}_t \mathbb{1}_{\mathscr{E}^a}(x) = \mathbb{P}_x(\tau > t)$. Notons que $(\mathrm{T}_t)_{t\geq 0}$ est une famille de contractions linéaires de $(\mathscr{M}_b(\mathscr{E}^a), ||\cdot||_\infty)$, c'est-à-dire une famille d'endomorphismes bornés de $(\mathscr{M}_b(\mathscr{E}^a), ||\cdot||_\infty)$ telle que

$$\forall t \geq 0, \quad |||\mathrm{T}_t|||_\infty \leq 1. \tag{0.1.143}$$

Notons que, pour toute fonction $g \in \mathscr{M}_b(\mathscr{E})$ et tout $x \in \mathscr{E}^a$, on a

$$\mathbb{E}_x[g(\mathrm{Y}_t)] = \mathbb{E}_x[g(\mathrm{Y}_t)\mathbb{1}_{\tau>t}] + g(\mathrm{Y}_\tau)\mathbb{P}_x(\tau \leq t). \tag{0.1.144}$$

En particulier, si $g(x) = 0$ pour tout $x \in \partial\mathscr{E}$, alors pour tout $x \in \mathscr{E}^a$, $\mathbb{E}_x[g(\mathrm{Y}_t)] = \mathbb{E}_x[g(\mathrm{Y}_t)\mathbb{1}_{\tau>t}]$. Pour toute mesure ν sur $\mathscr{E}^a$, $(\mathbb{E}_\nu[\mathrm{T}_t\mathbb{1}_\mathrm{B}], \mathrm{B} \in \mathscr{B}(\mathscr{E}^a))$ définit une mesure sur $\mathscr{E}^a$ qu'on note $\nu\mathrm{T}_t$: mesure induite par l'action faible (au sens des distributions) de T_t. Ainsi, par linéarisation, cette action peut être également définie pour toutes les mesures signées finies. D'où l'action du semi-groupe sur l'espace des mesures sur $\mathscr{E}^a$, ou l'espace des mesures finies signées sur $\mathscr{E}^a$, est définie par

$$\forall \varphi \in \mathscr{M}_b(\mathscr{E}^a), \quad \langle \nu\mathrm{T}_t, \varphi\rangle = \langle \nu, \mathrm{T}_t\varphi\rangle = \mathbb{E}_\nu[\varphi(\mathrm{Y}_t)\mathbb{1}_{\tau>t}]. \tag{0.1.145}$$

Notons $\mathrm{T}_t^\dagger \nu := \nu\mathrm{T}_t$. $(\mathrm{T}_t^\dagger)_{t\geq 0}$ est aussi un semi-groupe agissant sur ces espaces de mesures.

Théorème 0.1.7 (QSD et vecteur propre pour le semi-groupe faible)**.** *ν est une QSD si, et seulement si $\exists \theta > 0$ telle que*

$$\left(\forall \varphi \in \mathscr{M}_b(\mathscr{E}^a) \quad \forall t \geq 0, \quad \langle \mathrm{T}_t^\dagger \nu, \varphi\rangle = e^{-\theta t}\langle \nu, \varphi\rangle\right) \Longleftrightarrow \forall t \geq 0, \quad \mathrm{T}_t^\dagger \nu = e^{-\theta t}\nu. \tag{0.1.146}$$

Le paramètre θ s'avère égal à $\lambda(\nu)$, le taux exponentiel de survie de ν.

Existence de QSDs On peut montrer qu'il existe une QSD lorsque l'équation de mesure propre est satisfaite par une mesure finie mais pour un temps fixe strictement positif.

Théorème 0.1.8 (Existence QSD via le semi-groupe faible)**.** *Soient $\widetilde{\nu} \in \mathscr{P}(\mathscr{E}^a)$ et $\beta > 0$ tels que $\mathrm{T}_1^\dagger \widetilde{\nu} = \beta\widetilde{\nu}$. Alors $\beta < 1$ et il existe une QSD ν dont le taux de survie exponentiel est $\lambda(\nu) = -\ln\beta >$*

0 *et telle que*

$$\forall B \in \mathscr{B}(\mathscr{E}^a), \quad \nu(B) := \int_0^1 e^{\lambda(\nu)t}\langle \tilde{\nu}, T_t \mathbb{1}_B\rangle dt = \int_0^1 e^{\lambda(\nu)t} T_t^\dagger \tilde{\nu}(B) dt. \tag{0.1.147}$$

Rappelons un des théorèmes de Krein dans Oikhberg et Troitsky (2005), nous nous référons au théorème 4 de celle-ci. Soit $(B, \leq, ||\cdot||)$ un espace normé ordonné et supposons qu'il existe un élément $e \in B$ qui vérifie $||e|| = 1$ et $e \geq x$ pour tout $x \in B$ avec $||x|| \leq 1$. Alors, tout opérateur positif $R : B \to B$ est tel que son adjoint (transposition) R^* admet un vecteur propre positif v. La propriété de positivité de $v \in B^*$ signifie que $v(b) \geq 0$ pour tout $b \geq 0$. Ce théorème implique le résultat suivant sur l'existence de QSDs

Théorème 0.1.9 (Existence QSD via continuité)**.** *Supposons que $\mathscr{E}^a$ est un espace de Hausdorff compact et que T_1 préserve l'ensemble des fonctions continues : $T_1(\mathscr{C}(\mathscr{E}^a)) \subset \mathscr{C}(\mathscr{E}^a)$. Alors il existe une QSD.*

Démonstration. L'ensemble $\mathscr{C}(\mathscr{E}^a)$ est un espace normé ordonné, muni de l'ordre usuel entre les fonctions et la norme supremum $||\cdot||_\infty$. La fonction 1 vérifie la condition $f \leq 1$ pour tout $f \in \mathscr{C}(\mathscr{E}^a)$ avec $||f||_\infty \leq 1$. Puisque $T_1 : \mathscr{C}(\mathscr{E}^a) \to \mathscr{C}(\mathscr{E}^a)$ est un opérateur positif, le théorème de Krein ci-dessus stipule que $T_1' : \mathscr{C}(\mathscr{E}^a)' \to \mathscr{C}(\mathscr{E}^a)'$ a une fonction propre positive. Le THÉORÈME DE REPRÉSENTATION DE RIESZ implique que cette fonction propre est une mesure finie positive η. Donc $T_1'\eta = \beta\eta$ pour un certain β, et on a nécessairement $\beta > 0$. La mesure de probabilité $\tilde{\nu} = \frac{\eta}{\langle \eta, 1\rangle}$ vérifie $T_1'\tilde{\nu} = \beta\tilde{\nu}$. Puisque $T_1' = T_1^\dagger$ sur le dual topologique $\mathscr{C}(\mathscr{E}^a)'$, du theorem 0.1.8, on obtient que $\nu = \int_0^1 e^{-t\ln\beta} T_t^\dagger \tilde{\nu} dt$ est une QSD dont le taux exponentiel de survie est $\lambda(\nu) = -\ln\beta > 0$. ■

Nous énonçons maintenant un résultat d'existence d'une QSD lorsque $\mathscr{E}^a$ n'est pas nécessairement compact. On le retrouve dans Collet et al. (2011). On rappelle que l'ensemble des fonctions continues bornées $\mathscr{C}_b(\mathscr{E}^a)$ est un sous-espace fermé de l'espace de Banach des fonctions bornées muni de la norme $||\cdot||_\infty$: c'est aussi un espace de Banach pour la même norme.

Théorème 0.1.10 (Existence QSD via continuité dans le cas non compact)**.** *Supposons que T_1 envoie $\mathscr{C}_b(\mathscr{E}^a)$ sur lui-même de sorte que $T_1 : \mathscr{C}_b(\mathscr{E}^a) \to \mathscr{C}_b(\mathscr{E}^a)$. Supposons également que l'hypothèse suivante soit vraie : $\exists \psi \in \mathscr{C}_b(\mathscr{E}^a)$ telle que $\psi \geq 1$ et vérifie*

1. *Pour tout $u \geq 0$, l'ensemble $\psi^{-1}([0,u])$ est compact : on dit que ψ est propre.*
2. *Il existe trois constantes $c_1 > \gamma > 0$ et $\beta > 0$ telles que $T_1\mathbb{1}_{\mathscr{E}^a} \geq c_1\mathbb{1}_{\mathscr{E}^a}$ et pour tout $h \in \mathscr{C}_b(\mathscr{E}^a)$ avec $0 \leq h \leq \psi$, $T_1 h \leq \gamma\psi + \beta$.*

Alors, il existe une QSD.

Unicité de la QSD Nous relions ce problème au comportement à l'infini du processus Y. Dans le problème de l'unicité, il est naturel d'étudier la notion de distribution quasi-limite (QLD).

PROCESSUS DE LANGEVIN SUR-AMORTI ABSORBÉ

Soient $d \geq 1$, $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ un espace probabilisé filtré, $(B_t)_{t\geq 0}$ un $\mathscr{F}_t$-mouvement brownien dans $\mathbb{R}^d$, $F : \mathbb{R}^d \to \mathbb{R}^d \in \mathscr{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ et $\mathscr{O} \subset \mathbb{R}^d$ un ouvert CONNEXE $\mathscr{C}^2$ BORNÉ. On considère le processus de Langevin sur-amorti $(q_t)_{t\geq 0}$ défini par

$$dq_t = \sqrt{\frac{2}{\beta}} dB_t + F(q_t)dt \tag{0.1.148}$$

où F ne s'écrit pas nécessairement sous forme gradient. Son générateur infinitésimal est donné par

$$\mathscr{L}_F := \frac{1}{\beta}\Delta + F \cdot \nabla. \tag{0.1.149}$$

Pour $A \in \mathscr{F}$, notons $\mathbb{P}_x(A)$ la probabilité de l'événement A sous la condition $q_0 = x$. Si θ est une mesure de probabilité sur $\mathscr{O}$, nous noterons $\mathbb{P}_\theta(A) := \mathbb{E}_\theta[\mathbb{P}_\cdot(A)] = \int \mathbb{P}_x(A)\theta(dx)$, τ_∂ le premier temps de sortie du processus $(q_t)_{t\geq 0}$ du domaine $\mathscr{O}$ donné par

$$\tau_\partial := \inf\Big\{t > 0, \quad q_t \notin \mathscr{O}\Big\}. \tag{0.1.150}$$

Définition 0.1.11 (Processus absorbé (tué))**.** Par régularité de F, de l'ouvert $\mathscr{O}$ et bornitude de ce domaine, le processus $(q_t)_{0\leq t\leq \tau_\partial}$ est bien défini et appelé processus sur-amorti absorbé.

Nous rappelons brièvement dans cette sous-section certains résultats bien connus sur le processus sur-amorti absorbé dans un domaine borné régulier. Ces résultats seront en particulier utiles pour introduire la QSD associée au processus absorbé. Nous commençons par rappeler qu'il existe une unique solution classique [9] au problème suivant avec une condition

9. élément de $\mathscr{C}^{1,2}(\mathbb{R}_+^* \times \mathscr{O}) \bigcap \mathscr{C}_b(\mathbb{R}_+ \times \overline{\mathscr{O}} / (\{0\} \times \partial\mathscr{O}))$

initiale $f \in \mathscr{C}_b(\mathscr{O})$ et une condition au bord du domaine, $g \in \mathscr{C}_b(\partial\mathscr{O})$:

$$\begin{cases} \partial_t u(t,x) = \mathscr{L}_{\mathrm{F}} u(t,x), & t > 0, \quad x \in \mathscr{O}, \\ u(0,x) = f(x), & x \in \mathscr{O}, \\ u(t,x) = g(x), & t > 0, \quad x \in \partial\mathscr{O}. \end{cases} \tag{0.1.151}$$

Ce résultat est classiquement obtenu de la manière suivante :

— construire une solution faible u à partir d'une approche variationnelle ;

— déduire de la régularisation parabolique (l'opérateur $\mathscr{L}_{\mathrm{F}}$ est elliptique) que les solutions faibles sont en réalité régulières ;

— utiliser la régularité de u pour appliquer la formule d'Itô afin d'obtenir la représentation probabiliste suivante

$$u(t,x) = \mathbb{E}_x[f(q_t)\mathbb{1}_{\tau_\partial > t} + g(q_{\tau_\partial})\mathbb{1}_{\tau_\partial \le t}]. \tag{0.1.152}$$

La représentation probabiliste de la solution implique en particulier l'unicité de solution classique au problème au bord.

Les propriétés suivantes sont également bien connues pour le processus absorbé :

— Pour tout $x \in \mathscr{O}$, la mesure $\mathrm{P}_t^{\mathscr{O}}(x,\cdot) := \mathbb{P}_x(q_t \in \cdot, t < \tau_\partial)$ admet une densité (dite de transition) régulière par rapport à la mesure de Lebesgue sur $\mathscr{O}$;

— cette densité de transition satisfait les équations de Kolmogorov rétrograde et progressive sur $\mathbb{R}_+^* \times \mathscr{O} \times \mathscr{O}$;

— pour tout $t > 0$, la fonction de transition est strictement positive sur $\mathscr{O} \times \mathscr{O}$ et admet une extension continue sur $\overline{\mathscr{O}} \times \overline{\mathscr{O}}$ qui s'annule au bord $\partial(\mathscr{O} \times \mathscr{O})$;

— pour tout $\mathrm{T} > 0$, la densité de transition admet un majorant gaussien de la forme $e^{\frac{-\alpha}{4t}||\cdot - \cdot||^2}$, avec α une constante indépendante du temps, sur $(0,\mathrm{T}] \times \mathscr{O} \times \mathscr{O}$ (basé sur la méthode parametrix par exemple).

Il convient de se demander ce qu'il se passe par exemple lorsque le processus de Langevin sur-amorti demeure longtemps piégé dans un état : est ce que le processus atteint un état d'équilibre "local" sur le domaine considéré ? Nous consacrons cette sous-section aux réponses qui ont été apportées dans la littérature à cette question dans le cas du processus sur-amorti, absorbé au bord d'un domaine $\mathscr{O}$ régulier borné.

MÉTASTABILITÉ

La différence d'échelles de temps entre les comportements microscopique et macroscopique est liée à un phénomène bien connu appelé la métastabilité.

Définition 0.1.12 (Métastabilité)**.** On dit qu'un processus est métastable si il reste piégé dans des zones de l'espace pendant des temps très longs. Ces zones sont appelées des états métastables.

Un exemple très simple d'une telle situation est donné dans la fig. 1.1 ci-dessous et expliqué ci-dessous. Considérons une particule suivant le processus de Langevin donné par

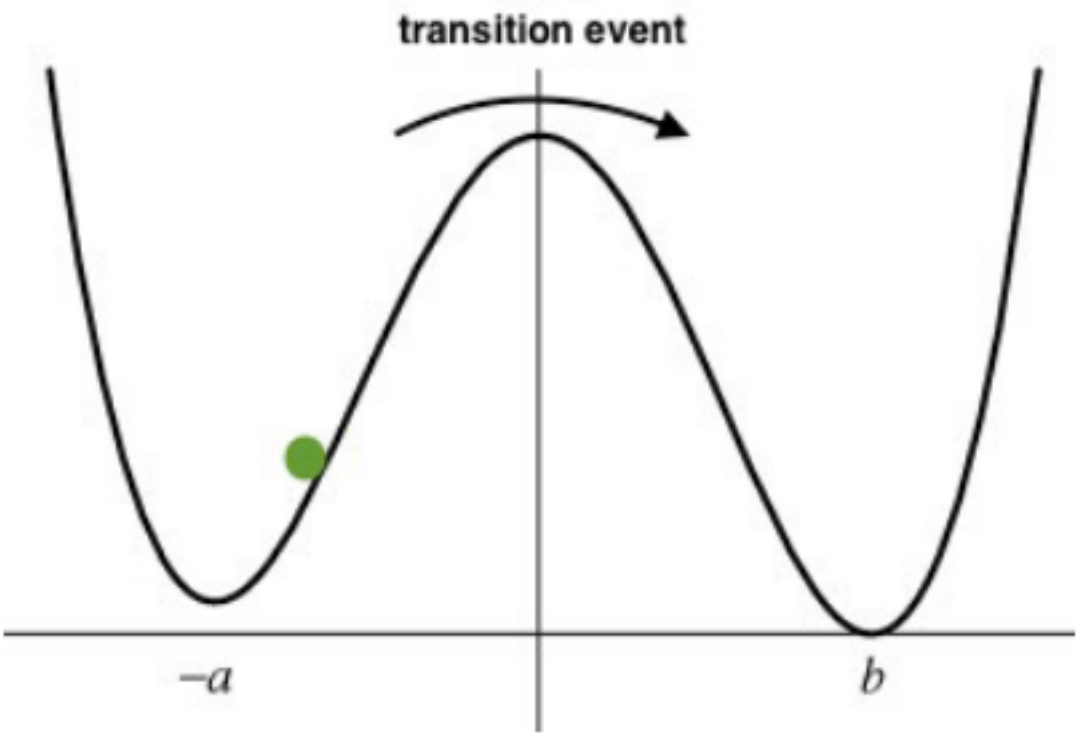

FIGURE 2 – Particule dans un potentiel à double-puits

l'équation de Fokker-Planck cinétique en une dimension ($d = 1$) avec le potentiel H représenté sur la fig. 1.1. Si la position de la particule correspond à la boule verte, alors celle-ci va osciller pendant un temps très long autour du minimum local $x = -a$ avant d'effectuer un saut vers le puits de potentiel centré en $x = b$. Dans cette situation l'état métastable correspond à un puits de potentiel. Le système oscille très longtemps avant de franchir la barrière énergétique nécessaire pour quitter son état présent. Cependant, d'autres types de métastabilité existent, par exemple la métastabilité "entropique" où la transition ne nécessite pas

de franchir une barrière énergétique mais de s'aventurer par un passage très étroit. L'étude des processus métastables basée sur la distribution quasi-stationnaire s'applique à la fois à des métastabilités énergétique et entropique. Cette étude concerne l'évènement de sortie : le couple temps de sortie et état de sortie. Intuitivement, cette mesure d'équilibre local, qu'on notera $\overline{\mu}$, est à support sur l'état que nous considérons et elle est telle que si le processus $(q_t)_{t\geq 0}$ est initialement distribué suivant $\overline{\mu}$, alors il continue de suivre la même loi $\overline{\mu}$ tant qu'il n'est pas sorti du domaine. Considérons un domaine $\mathscr{O} \subset \mathbb{R}^d$, la définition rigoureuse est la suivante :

Définition 0.1.13 (Distribution quasi-stationnaire (QSD))**.** Une mesure de probabilité $\overline{\mu}$ sur $\mathscr{O} \subset \mathbb{R}^d$ est une QSD sur $\mathscr{O}$ pour le processus $(q_t)_{t\geq 0}$, si pour tout borélien A de $\mathscr{O}$,

$$\forall\, t \geq 0, \quad \mathbb{P}_{\overline{\mu}}(q_t \in \mathrm{A}, \tau_\partial > t) = \overline{\mu}(\mathrm{A})\mathbb{P}_{\overline{\mu}}(\tau_\partial > t) \tag{0.1.153}$$

Si $\mathbb{P}_{\overline{\mu}}(\tau_\partial > t) > 0$, on a

$$\mathbb{P}_{\overline{\mu}}(q_t \in \mathrm{A} | \tau_\partial > t) = \overline{\mu}(\mathrm{A}). \tag{0.1.154}$$

On voit donc qu'une QSD peut être vue comme une mesure invariante pour le processus conditionné à rester dans le domaine.

CARACTÉRISATIONS DES QSDS DU SUR-AMORTI & CONVERGENCES

En plus des questions d'existence et d'unicité d'une QSD, il convient également de se poser la question de la convergence de la loi du processus conditionné à rester dans le domaine $\mathscr{O}$, soit la loi $\mathscr{L}(q_t | \tau_\partial > t)$, lorsque le temps t tend vers l'infini. En particulier, nous souhaiterions une convergence vers la QSD sur $\mathscr{O}$ du processus $(q_t)_{t\geq 0}$. Les résultats obtenus dans la littérature vont dans ce sens et traitent de ces trois questions d'existence, d'unicité et de convergence, avec des outils relativement diversifiés. Par exemple, Le Bris et al. prennent $\mathrm{F} = -\nabla \mathrm{V}$ et l'approche utilisée se base sur le fait que le processus $(q_t)_{t\geq 0}$ est réversible par rapport à la mesure de Gibbs $\nu(dx) := \frac{1}{Z} e^{-\beta \mathrm{V}(x)} dx$ avec $\mathrm{Z} := \int e^{-\beta \mathrm{V}(x)} dx$ supposément finie. L'inverse de l'opérateur $\mathscr{L}_\mathrm{F}$ avec condition de Dirichlet au bord $\partial\mathscr{O}$ étant compact de $\mathrm{L}^2(d\nu)$ vers $\mathrm{L}^2(d\nu)$ et auto-adjoint, il est possible d'obtenir une décomposition spectrale de $\mathscr{L}_\mathrm{F}$ avec cette condition au bord. On peut alors montrer que le vecteur propre u_1 associé à la plus grande valeur propre de $\mathscr{L}_\mathrm{F}$ avec condition de Dirichlet au bord $\partial\mathscr{O}$ est tel que $u_1 \frac{d\nu}{dx}$ est la densité de la QSD. On en déduit l'existence et l'unicité de la QSD ainsi que la convergence de la loi conditionnelle vers la QSD pour le processus $(q_t)_{t\geq 0}$ en utilisant la décomposition

spectrale. Dans le cas plus général où F ne s'écrit pas comme gradient d'une fonctionnelle, l'argument de réversibilité et donc de symétrie sur $\mathscr{L}_F$ ne fonctionne pas. Cependant ce cas a été traité par Gong et al. par exemple, en appliquant le théorème de Krein-Rutman au dual du semi-groupe du processus absorbé sur l'espace de Banach

$$\left\{ f \in \mathscr{C}_b(\mathscr{O}), \quad \frac{f}{d(\cdot,\partial\mathscr{O})} \quad \text{est uniformément continue sur} \quad \mathscr{O} \right\} \tag{0.1.155}$$

avec $d_\partial := d(\cdot,\partial\mathscr{O})$ désignant la fonction distance euclidienne au bord $\partial\mathscr{O}$, permettant d'obtenir "le premier" vecteur propre de ce semi-groupe et donc la densité de la QSD. Une autre approche plus probabiliste a été invoquée par Champagnat et al. qui ont prouvé que le semi-groupe du processus $(q_t)_{t\geq 0}$ absorbé satisfait certaines estimées de gradient, une condition d'irréductibilité ainsi qu'un contrôle de la probabilité d'absorption près du bord $\partial\mathscr{O}$, à partir desquelles ils ont déduit l'existence d'une unique QSD ainsi que la convergence attendue de la loi conditionnelle. Nous n'avons mentionné que trois approches et d'autres preuves peuvent être trouvées par exemple dans la littérature mentionnée plus bas. Au final, les résultats obtenus dans la littérature peuvent être résumés dans le théorème ci-dessous dans le cas du processus sur-amorti $(q_t)_{t\geq 0}$.

Théorème 0.1.11 (Existence/unicité/convergence QSD)**.** *Sous les hypothèses de régularité de* F *et* $\mathscr{O}$*, de bornitude et de connexité de* $\mathscr{O}$*, le processus* $(q_t)_{t\geq 0}$ *a une unique QSD* $\overline{\mu}$ *telle que*

—

$$\exists \psi \in \mathscr{C}^2(\mathscr{O}) \bigcap \mathscr{C}_b(\overline{\mathscr{O}}), \quad \overline{\mu}(dx) = \psi(x)dx. \tag{0.1.156}$$

Autrement dit, la QSD admet une densité avec une certaine régularité par rapport à la mesure de Lebesgue.

— *il existe un unique couple* (λ,η)*, à normalisation près de* η*, tel que* η *est une solution positive classique dans* $\mathscr{C}^2(\mathscr{O})\bigcap\mathscr{C}_b(\overline{\mathscr{O}})$ *du problème aux valeurs propres suivant*

$$\begin{cases} \mathscr{L}_F^*\eta(x) = -\lambda\eta(x), & x \in \mathscr{O}, \\ \eta(x) = 0, & x \in \partial\mathscr{O}. \end{cases} \tag{0.1.157}$$

De plus, $\frac{\eta}{\int_{\mathscr{O}}\eta} = \psi$ *et* λ *est la plus petite valeur propre de* $-\mathscr{L}_F^*$ *l'adjoint Lebesgue de* $\mathscr{L}_F$*, qui satisfait de plus* $\lambda > 0$*.*

— *il existe* $C > 0, \alpha > 0$ *tels que pour toute mesure de probabilité* θ *(*$q_0 \sim \theta$*) sur* $\mathscr{O}$*, pour tout*

$t \geq 0$,

$$||\mathbb{P}_\theta(q_t \in \cdot | \tau_\partial > t) - \overline{\mu}||_{\mathrm{TV}} \leq Ce^{-\alpha t} \tag{0.1.158}$$

où $||\cdot||_{\mathrm{TV}}$ *correspond à la norme en variation totale.*

Q-PROCESSUS & CONVERGENCE EXPONENTIELLE (CNS)

Champagnat and Villemonais ont montré qu'il y a équivalence entre les assertions suivantes

— $\exists \nu \in \mathscr{P}(E)\ \exists C > 0\ \exists c > 0$,

$$\forall \mu \in \mathscr{P}(E) \quad \forall t \geq 0, \quad |\mathbb{P}_\mu(X_t \in \cdot | t < \tau_\partial) - \nu|_{\mathrm{TV}} \leq Ce^{-ct}. \tag{0.1.159}$$

Ce qui implique immédiatement que ν est l'unique QSD de X.

— $\exists \nu \in \mathscr{P}(E)\ \exists t_0 > 0\ \exists c_1 > 0\ \exists c_2 > 0$,

$$\forall x \in E, \quad \mathbb{P}_x(X_{t_0} \in \cdot | t_0 < \tau_\partial) \geq c_1 \nu; \tag{0.1.160}$$

$$\forall x \in E \quad \forall t \geq 0, \quad \mathbb{P}_\nu(t < \tau_\partial) \geq c_2 \mathbb{P}_x(t < \tau_\partial). \tag{0.1.161}$$

De plus, si l'une ou l'autre de ces assertions est satisfaite, il existe $\psi_\nu : E \cup \{\partial\} \to [0, +\infty)$ une fonction bornée et nulle sur ∂ telle que

$$\forall \mu \in \mathscr{P}(E), \quad \langle \mu, \psi_\nu \rangle = \lim_{t \to +\infty} \frac{\mathbb{P}_\mu(t < \tau_\partial)}{\mathbb{P}_\nu(t < \tau_\partial)} = \lim_{t \to +\infty} e^{\lambda(\nu)t} \mathbb{P}_\mu(t < \tau_\partial); \tag{0.1.162}$$

où la convergence est uniforme sur l'ensemble des mesures de probabilité et en particulier il existe $C_1 > 0$ telle que

$$\forall x \in E \quad \forall t \geq 0, \quad |e^{\lambda(\nu)t} \mathbb{P}_x(t < \tau_\partial) - \psi_\nu(x)| \leq C_1 e^{-ct}. \tag{0.1.163}$$

Dans le cas des processus de diffusion, ψ_ν est une solution positive bornée de $\mathscr{L}\psi_\nu = -\lambda(\nu)\psi_\nu$ (problème aux valeurs propres) avec $\mathscr{L}$ le générateur infinitésimal de X sur l'espace des fonctions mesurables bornées muni de la norme L^∞. En particulier, il existe $C_2 > 0$ telle que

$$\forall x \in E \quad \forall t \geq 0, \quad \mathbb{P}_x(t < \tau_\partial) \leq C_2 e^{-\lambda(\nu)t}. \tag{0.1.164}$$

Champagnat and Villemonais ont mieux en montrant qu'il existe $t_1 > 0$ et $C_3 \geq 0$ telles que

$$\forall x \in E \quad \forall t \geq t_1, \quad |e^{\lambda(v)t}\mathbb{P}_x(t < \tau_\partial) - \psi_v(x)| \leq C_3\psi_v(x)e^{-ct}. \tag{0.1.165}$$

On peut supposer sans perte de généralité que $C_3 e^{-ct_1} \leq \frac{1}{2}$. Dans le cas irréductible, l'ergodicité exponentielle est connue pour être équivalente à une propriété de trou spectral. Quant aux QSDs, on passe par la notion de Q-processus.

Théorème 0.1.12 (Existence et convergence exponentielle de Q-processus)**.** *Sous les conditions nécessaires et suffisantes de convergence exponentielle, on a*

— ***Existence de Q-processus.*** *Il existe* $(\mathbb{Q}_x)_{x\in E}$ *une famille de mesures de probabilité sur* Ω *définie par*

$$\mathbb{Q}_x(A) := \lim_{t\to+\infty} \mathbb{P}_x(A|t < \tau_\partial) \tag{0.1.166}$$

pour tout A $\mathscr{F}_s$*-mesurable. Le processus* $(\Omega, (\mathscr{F}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (\mathbb{Q}_x)_{x\in E})$ *est un processus de Markov homogène à valeurs dans* E*. De plus, si* X *est un Markov fort sous* $\mathbb{P}$*, il l'est aussi sous* $\mathbb{Q}$*.*

— ***Noyau de transition.***

$$p_t^{\mathbb{Q}}(x, dy) := e^{\lambda(v)t}\frac{\psi_v(y)}{\psi_v(x)}p_t^{\mathbb{P}}(x, dy). \tag{0.1.167}$$

Autrement dit (semi-groupes),

$$\forall \varphi \in \mathscr{B}(E, \mathbb{R}) \quad \forall t \geq 0, \quad P_t^{\mathbb{Q}}\varphi(x) = \frac{e^{\lambda(v)t}}{\psi_v(x)}P_t^{\mathbb{P}}(\psi_v\varphi)(x). \tag{0.1.168}$$

— ***Convergence exponentielle.*** *La mesure de probabilité sur* E *définie par*

$$\beta(dx) := \psi_v(x)\nu(dx) \tag{0.1.169}$$

est l'unique mesure de probabilité invariante de X *sous* $\mathbb{Q}$*. De plus, on a pour tout* $(\mu_0, \mu_1) \in \mathscr{P}(E) \times \mathscr{P}(E)$*,*

$$||\mathbb{Q}_{\mu_0}(X_t \in \cdot) - \mathbb{Q}_{\mu_1}(X_t \in \cdot)||_{TV} \leq (1 - c_1' c_2')^{\mathbf{E}(\frac{t}{t_0})}||\mu_0 - \mu_1||_{TV} \tag{0.1.170}$$

avec $\mathbb{Q}_\mu := \int_E \mathbb{Q}_x\mu(dx)$*.*

APPROXIMATION DE LA QSD D'UNE DIFFUSION HUE

Soit $(\Omega, ((\mathscr{F}_t)_{t\in[0,+\infty)}, (\mathbb{P}_x)_{x\in E\cup\{\partial\}}, (X_t)_{t\in[0+\infty)})$ un procesus de Markov homogène (en temps) d'espace d'états $E\cup\{\partial\}$ avec E un espace mesurable et $\partial\notin E$ un état absorbant pour le processus : cela signifie que si $X_s=\partial$, alors pour tout $t\geq s$, $X_t=\partial$, $\mathbb{P}_x$-presque sûrement pour tout $x\in E$ et, en particulier,

$$\tau_\partial := \inf\{t\geq 0,\quad X_t=\partial\}\quad \text{est un temps d'arrêt.} \tag{0.1.171}$$

On suppose aussi que pour tout $(t,x)\in[0,+\infty)\times E$,

$$\mathbb{P}_x(\tau_\partial<+\infty)=1\quad\text{and}\quad \mathbb{P}_x(\tau_\partial>t)>0. \tag{0.1.172}$$

On considère le processus aléatoire $(Y_t)_{t\geq 0}$ qui évolue suivant la dynamique de X lorsqu'il appartient à E et qui est ré-échantillonné selon sa mesure d'occupation lorsqu'il atteint ∂. Plus précisément, étant donnée une mesure de probabilité μ sur E, pour tout $t\geq 0$, on pose

$$Y_t := \sum_{k=1}^{+\infty} \mathbb{I}_{[\theta_{k-1},\theta_k)}(t) X^{(k)}_{t-\theta_{k-1}} \tag{0.1.173}$$

avec $\theta_0=0$,

— $(X^{(1)}_t)_{t\geq 0}$ est une réalisation du processus $(X_t)_{t\geq 0}$ telle que $X^{(1)}_0\sim\mu$ (c'est-à-dire sous $\mathbb{P}_\mu$) et le temps d'arrêt θ_1 est égal à $\tau^{(1)}_\partial$ le premier temps d'atteinte de ∂ par $X^{(1)}$;

— $X^{(1)}$ donné, $(X^{(2)}_t)_{t\geq 0}$ est une réalisation de $(X_t)_{t\geq 0}$ telle que $X^{(2)}_0\sim\mu_{\theta_1}$ avec

$$\mu_{\theta_1} := \frac{1}{\theta_1}\int_0^{\theta_1}\delta_{Y_s}ds \tag{0.1.174}$$

et $\theta_2-\theta_1=\tau^{(2)}_\partial$ le premier temps d'atteinte de ∂ par $X^{(2)}$;

— Pour tout $k\geq 1$, étant donnés $(X^{(1)},\dots,X^{(k)})$, $(X^{(k+1)}_t)_{t\geq 0}$ est une réalisation de $(X_t)_{t\geq 0}$ telle que $X^{(k+1)}_0\sim\mu_{\theta_k}$ avec

$$\mu_{\theta_k} := \frac{1}{\theta_k}\int_0^{\theta_k}\delta_{Y_s}ds \tag{0.1.175}$$

et $\theta_{k+1}-\theta_k=\tau^{(k+1)}_\partial$ le premier temps d'atteinte de ∂ par $X^{(k+1)}$ de sorte que

$$\forall k\geq 0,\quad \theta_k=\sum_{j=1}^{k}\tau^{(j)}_\partial. \tag{0.1.176}$$

On définit pour tout $t>0$, la mesure d'occupation

$$\mu_t := \frac{1}{t}\int_0^t \delta_{Y_s} ds \Longleftrightarrow \forall f \in \mathscr{M}_b(E), \quad \langle \mu_t, f\rangle = \frac{1}{t}\int_0^t f(Y_s) ds. \tag{0.1.177}$$

On considère désormais un processus de diffusion $(X_t)_{t\geq 0}$ évoluant dans un ouvert $D \subset \mathbb{R}^{d\geq 2}$ borné, connexe, de bord ∂D $\mathscr{C}^2$ et absorbé au bord. On suppose que X est solution de l'équation différentielle stochastique

$$dX_t = \sigma(X_t) dB_t + b(X_t) dt \tag{0.1.178}$$

avec $(B_t)_{t\geq 0}$ un mouvement brownien r-dimensionnel, $b : D \to \mathbb{R}^d$ un drift continu borné et $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ continue telle que $\sigma\sigma^*$ est uniformément elliptique et pour tout $\rho > 0$,

$$\sup_{(x,y)\in D^2, |x-y|=\rho} \frac{|\sigma(x)-\sigma(y)|^2}{\rho} \leq g(\rho) \tag{0.1.179}$$

avec $g : (0,+\infty) \to (0,+\infty)$ une fonction telle que $\int_0^1 g(r)dr < +\infty$. Notons que dans ce cas, le processus $(Y_t, \mu_t)_{t\geq 0}$ est bien défini puisqu'on peut montrer que $\theta_k \overset{k\to+\infty}{\longrightarrow} +\infty$ presque sûrement (voir [29, Lemma 8.1]). Dans [72, Section 5.3], il a été prouvé que, sous l'hypothèse de régularité ci-dessus que le processus de diffusion tué X admet une distribution quasi-stationnaire unique, c'est-à-dire une mesure de probabilité ν sur D telle que pour tout $t \geq 0$,

$$\nu = \mathbb{P}_\nu(X_t \in \cdot | t < \tau_\partial) \tag{0.1.180}$$

avec τ_∂ le premier temps d'atteinte de ∂D par X. De plus, il est bien connu que

$$\exists \lambda_0 > 0 \quad \forall t \geq 0, \quad \mathbb{P}_\alpha(\tau_\partial > t) = e^{-\lambda_0 t}. \tag{0.1.181}$$

Benaïm et al. ont montré le résultat ci-dessous

Théorème 0.1.13 (Convergence de la mesure d'occupation vers la QSD et approximation du taux de survie)**.**

$$\forall f \in \mathscr{M}_b(D), \quad \langle \mu_t, f\rangle \overset{t\to+\infty}{\longrightarrow} \langle \nu, f\rangle \quad \text{presque sûrement.} \tag{0.1.182}$$

$$\frac{\theta_n}{n} \overset{n\to+\infty}{\longrightarrow} \frac{1}{\lambda_0} \quad \text{presque sûrement.} \tag{0.1.183}$$

L'outil central est l'étude des pseudo-trajectoires asymptotiques associées à une dynamique d'équation différentielle ordinaire sur un espace de mesures. Soit $(P_t^\partial)_{t\geq 0}$ le semi-groupe de Markov (non conservatif) associé à X :

$$\forall \varphi \in \mathscr{B}(E,\mathbb{R}) \quad \forall x \in E, \quad P_t^\partial \varphi(x) := \mathbb{E}_x[\varphi(X_t)\mathbb{I}_{t<\tau_\partial}] \tag{0.1.184}$$

À noter que nous avons fait ici le léger abus de notation que $\varphi(\partial)\cdot 0 = 0$. D'après le theorem 0.1.12, on peut définir l'opérateur de Green-Poisson associé sur $\mathscr{B}(E,\mathbb{R})$ par

$$A\varphi(x) := \mathbb{E}_x \int_0^{\tau_\partial} \varphi(X_s)\,ds = \int_0^{+\infty} P_t^\partial \varphi(x)\,dt. \tag{0.1.185}$$

C'est un opérateur (endomorphisme) borné de $\mathscr{B}(E,\mathbb{R})$ muni de la norme L^∞ :

$$\begin{aligned} \forall \varphi \in \mathscr{B}(E,\mathbb{R}) \quad \forall x \in E, \quad |A\varphi(x)| &\leq ||\varphi||_\infty \mathbb{E}_x[\tau_\partial] \\ &= ||\varphi||_\infty \int_0^{+\infty} \mathbb{P}_x(\tau_\partial > t)\,dt \\ &\leq C_2 ||\varphi||_\infty \int_0^{+\infty} e^{-\lambda(\nu)t}\,dt \\ &= \frac{C_2}{\lambda(\nu)} ||\varphi||_\infty. \end{aligned} \tag{0.1.186}$$

En conclusion, on a

$$|||A|||_{\mathscr{B}(E,\mathbb{R})\to\mathscr{B}(E,\mathbb{R})} = \sup_{x\in E} A\mathbb{I}_E(x) = \sup_{x\in E} \mathbb{E}_x[\tau_\partial] \leq \frac{C_2}{\lambda(\nu)} < +\infty. \tag{0.1.187}$$

On étend l'action de A aux mesures par transposition :

$$\langle \mu A, \varphi\rangle := \langle \mu, A\varphi\rangle = \int_E A\varphi(x)\,\mu(dx) = \mathbb{E}_\mu \int_0^{\tau_\partial} \varphi(X_s)\,ds = \int_0^{+\infty} \langle \mu P_t^\partial, \varphi\rangle\,dt. \tag{0.1.188}$$

En particulier, on a

$$A\varphi(x) = \langle \delta_x A, \varphi\rangle \quad \text{et} \quad \langle \nu A, \varphi\rangle = \langle \nu, \varphi\rangle \int_0^{+\infty} e^{-\lambda(\nu)t}\,dt = \frac{1}{\lambda(\nu)}\langle \nu, \varphi\rangle. \tag{0.1.189}$$

Comme A est borné, pour tout $t \geq 0$, e^{tA} est bien défini et borné : la série d'opérateurs bornés converge normalement et

$$|||e^{tA}||| \leq e^{t|||A|||}. \tag{0.1.190}$$

De plus, pour tout $n \geq 1$ on a (par induction)

$$\mu A^n = \frac{1}{\Gamma(n)} \int_0^{+\infty} t^{n-1} \mu P_t^{\partial} dt \tag{0.1.191}$$

avec $\Gamma : x \in]0,+\infty[\mapsto \int_0^{+\infty} u^{x-1} e^{-u} du$ la trace sur les réels positifs stricts de la fonction gamma d'Euler. Benaïm et al. ont montré les estimées ci-dessous.

Proposition 0.1.6 (Estimations). *Sous les conditions nécessaires et suffisantes de convergence exponentielle du theorem 0.1.12, on a*

$$\forall \mu \in \mathscr{P}(E) \quad \forall \varphi \in \mathscr{B}(E,\mathbb{R}) \quad \forall n \geq 1, \quad \left| \langle \mu A^n, \varphi \rangle - \frac{1}{\lambda(\mu)^n} \langle \nu, \varphi \rangle \langle \mu, \psi_\nu \rangle \right| \leq \frac{CC_2 + C_1}{(\lambda(\nu)+c)^n} ||\varphi||_\infty. \tag{0.1.192}$$

De plus, il existe $C_4 > 0$ *telle que*

$$\left|\left| \frac{\mu A^n}{\langle \mu A^n, \mathbb{I}_E \rangle} - \nu \right|\right|_{TV} \leq \frac{C_4}{\langle \mu, \psi_\nu \rangle} \left(\frac{\lambda(\nu)}{\lambda(\nu)+c} \right)^n ; \tag{0.1.193}$$

$$\forall t \geq 0, \quad \left|\left| \frac{\mu e^{tA}}{\langle \mu e^{tA}, \mathbb{I}_E \rangle} - \nu \right|\right|_{TV} \leq \frac{C_4}{\langle \mu, \psi_\nu \rangle} e^{-t \frac{c}{\lambda(\nu)(\lambda(\nu)+c)}}. \tag{0.1.194}$$

De ces estimées, Benaïm et al. ont montré les résultats de régularité ci-dessous.

Proposition 0.1.7 (Régularisation). *Pour toute fonction mesurable bornée* $\varphi : D \to \mathbb{R}$, *sous les conditions nécessaires et suffisantes de convergence exponentielle du theorem 0.1.12,* $x \in D \mapsto \langle \delta_x A, \varphi \rangle$ *est lipschitzienne de rapport proportionnel à* $||\varphi||_\infty$. *En particulier, pour toute fonction continue bornée* φ, $x \in D \mapsto \langle \delta_x A, \varphi \rangle$ *est continue et bornée. cette propriété de Feller assure que* $\mu \in \mathscr{M}^{+<\infty}(D) \mapsto \mu A$ *et* $(t,\mu) \in [0+,\infty) \times \mathscr{M}^{+<\infty}(D) \mapsto \mu e^{tA} \in \mathscr{M}^{+<\infty}(D)$ *sont continues pour les topologies faibles sur les espaces* $\mathscr{M}^{+<\infty}(D)$ *des mesures positives finie sur* D *et* $[0+,\infty) \times \mathscr{M}^{+<\infty}(D)$.

De ces résultats, Benaïm et al. ont défini la dynamique de l'EDO ci-dessous.

Théorème 0.1.14 (EDO). *Pour toute mesure de probabilité* μ *sur* D, *On considère le problème de Cauchy sur* $\mathscr{M}^{+<\infty}(D)$

$$\begin{cases} \frac{d}{dt}\nu_t &= \mathscr{F}(\mu_t) \\ \nu_0 &= \mu \end{cases} \tag{0.1.195}$$

avec $\mathscr{F} : \mu \mapsto \mu A - \langle \mu A, \mathbb{I}_D \rangle \mu$. *Ce problème de Cauchy admet une unique solution faible élément de* $\mathscr{C}([0,+\infty), \mathscr{M}^{+<\infty}(D))$ *avec* $\mathscr{M}^{+<\infty}(D)$ *muni de sa topologie faible en ce sens que pour toute*

fonction continue bornée $\varphi : \mathrm{D} \to \mathbb{R}$,

$$\forall\, t \geq 0, \quad \langle \nu_t, \varphi \rangle = \langle \mu, \varphi \rangle + \int_0^t \langle \mathscr{F}(\nu_s), \varphi \rangle \, ds. \tag{0.1.196}$$

De plus, cette unique solution faible prend ses valeurs dans $\mathscr{P}(\mathrm{D})$ *et est donnée par*

$$\forall\, t \geq 0, \quad \nu_t = \frac{\mu e^{t\mathrm{A}}}{\langle \mu e^{t\mathrm{A}}, \mathbb{I}_{\mathrm{D}} \rangle}. \tag{0.1.197}$$

De cette dynamique, Benaïm et al. ont défini un algorithme stochastique satisfait par la mesure empirique aléatoire du processus redistribué $(\mathrm{Y}_t)_{t\geq 0}$ pris aux instants $(\theta_k)_{k\geq 0}$ de redistribution. D'autre part, Benaïm et al. ont montré la tension du flot $(\mu_t)_{t\geq 0}$ des mesures d'occupation en montrant qu'il n'y a pas concentration de la masse au bord via l'équation eikonale pour la fonction distance au bord assurant l'application de la formule d'Itô sur un voisinage du bord et une inversion du temps permettant de se ramener à une dynamique de brownien réfléchi drifté ergodique de mesure invariante une mesure de Gibbs. Ensuite, Benaïm et al. ont montré que la mesure empirique aléatoire converge presque sûrement vers l'unique QSD du processus absorbé en montrant qu'une version continue est une pseudo-trajectoire asymptotique relativement compacte de l'EDO dont l'unique QSD est un attracteur global compact. Enfin, Benaïm et al. ont montré par des arguments de martingales et la convergence des moyennes des instants de redistribution vers l'espérance du temps d'atteinte du bord sous la QSD la convergence de la mesure d'occupation du processus redistribué vers la QSD.

0.2 Résumés des chapitres et contributions

Dans cette section, il est expliqué brièvement le contenu de chaque chapitre en mettant en avant les contributions par rapport à la littérature.

▷ Guillin et al. ont établi des inégalités fonctionnelles non linéaires de type Poincaré et Sobolev logarithmiques pour l'équation des milieux granulaires. Le chapitre 1 est consacré à l'article [6] qui généralise ces résultats à l'équation de Langevin de champ moyen. Plus précisément, comme expliqué précédemment dans la deuxième sous-section sur les fonctionnelles de Lyapunov, le chapitre 1 est connecté à l'"approche indirecte" : via une inégalité fonctionnelle uniforme (en le nombre de particules) pour la mesure invariante du système de particules, on établit une dissipation d'entropie qui implique une contraction exponentielle en métrique de Wasserstein. Les principaux résultats du chapitre sont présentés dans les théorèmes 1.4.1, 1.4.2, 1.4.3 et 1.4.4. Des exemples concrets sont proposés dans la sous-section 1.4.3.

▷ Bolley et al. ont montré par couplage synchrone une contraction exponentielle en métrique de Wasserstein pour l'équation de Vlasov-Fokker-Planck. Ils en ont déduit une propagation du chaos uniforme en temps. Le chapitre 2 (fruit de l'article [1]) étend ces résultats aux équations de Langevin de champ moyen. Plus précisément, comme expliqué précédemment dans la première sous-section sur les limites de champ moyen, le chapitre 2 est connecté à l'"approche directe" : contraction exponentielle en métrique de Wasserstein par couplage synchrone et propagation du chaos uniforme en temps. Les principaux résultats du chapitre sont présentés dans les théorèmes 2.3.1, 2.3.2, 2.3.3, 2.3.4 et 2.3.5. Des exemples concrets sont proposés dans la section 2.4.

▷ Benaim et al. ont construit et étudié un algorithme basé sur l'(auto-)interaction d'une dynamique avec sa mesure d'occupation pour approximer des distributions quasi-stationnaires (QSDs) de chaînes de Markov générales conditionnées à rester dans un ensemble compact : pour plus de détails, voir la troisième sous-section précédente sur les distributions quasi-stationnaires. Le chapitre 3 (fruit de l'article [2] en collaboration avec Fabien PANLOUP) propose une extension au cas des dynamiques de McKean-Vlasov absorbées : convergence d'algorithmes auto-attractifs vers la QSD pour des dynamiques avec interaction de type McKean-Vlasov. Les QSDs de McKean-Vlasov, contrairement au cadre markovien, ont été très peu étu-

diées : à notre connaissance, le seul article sur le sujet est celui de Tough and Nolen sur les systèmes de particules de Fleming-Viot pour les dynamiques absorbées de type McKean-Vlasov. Ici, on s'intéressera principalement à des dynamiques en temps discret et on commencera par rappeler/étendre des propriétés standard des QSDs au cadre McKean-Vlasov puis on se focalisera sur l'étude de notre algorithme "auto-attractif". Comme indiqué plus haut, l'objectif est ici de généraliser un résultat de convergence de dynamique auto-attractive vers la QSD. Par dynamique auto-attractive, l'idée est ici de considérer la dynamique d'une particule qui, lorsqu'elle meurt, est régénérée selon sa mesure d'occupation, c'est-à-dire en tirant un point au hasard dans le passé de la trajectoire. Cette approche se distingue des algorithmes dits de Fleming-Viot où l'on considère un système de particules et lorsqu'une d'elles meurt, elle est regénérée selon la mesure empirique du système, c'est-à-dire en tirant au hasard parmi les autres particules. Après avoir considéré le problème "trop difficile" consistant à regarder le problème dit de *hard-killing* en temps continu dont subsistent quelques résultats dans la partie Conclusion & Perspectives, l'idée a été de considérer des dynamiques à temps discret, typiquement des schémas d'Euler d'équations de McKean-Vlasov absorbées. Dans ce cadre, un résultat général de convergence (Théorème 3.2.8) est obtenu de type "tension + identification de la limite" à l'aide d'outils issus des algorithmes stochastiques consistant à prouver que la suite peut être vue comme une *pseudo-trajectoire* d'une EDO dont les points fixes sont des QSDs. Ce travail contient plusieurs résultats vraiment nouveaux tels que la tension de la suite (dans le cas non compact!) sous des hypothèses de Lyapounov ainsi que l'identification de la limite qui requiert des outils de couplage fins permettant de pallier la non-unicité des objets limites dans ce cadre non-linéaire. Des applications sont proposées dans lesquelles sont prouvées la convergence de l'algorithme pour le schéma d'Euler dans le cas compact et dans un cas non-compact (pas tout à fait standard). Une étude du comportement limite lorsque le pas tend vers 0 montre que l'on retrouve alors les QSDs d'EDS de McKean-Vlasov absorbées dans le cas où l'absorption a lieu en sortant d'un compact, ce qui permet de produire une réponse "faible" au problème posé initialement associé au hard-killing en temps continu. L'adjectif "faible" fait référence à la double limite : faire tendre le temps vers l'infini puis le pas vers zéro (s'opposant à une approche forte dans laquelle la dynamique à temps continu serait directement considérée). Enfin, nous illustrons nos résultats théoriques par plusieurs simulations (voir les figures 3.1 et 3.2) : des exemples concrets sont proposés dans la sous-section 3.2.2 et nous proposons des simulations dans la sous-section 3.2.3.

0.2.1 Contenu du chapitre 1

Dans le premier chapitre, nous sommes intéressés par l'ergodicité exponentielle de limites de champ moyen [10]. Plus précisément, nous établissons le comportement en temps long à taux exponentiel d'équations de McKean-Vlasov et de Vlasov-Fokker-Planck avec dépendance en la loi dérivant d'une dérivée plate d'un hamiltonien macroscopique : polynôme sur l'espace des probabilités.

Soit $\mathrm{H} : \mathscr{P}(\mathbb{R}^d) \to \overline{\mathbb{R}}$, $\mu \mapsto \int \mathrm{V} d\mu + \sum_{k=2}^{\mathrm{N}} \int \mathrm{W}^{(k)} d\mu^{\otimes k}$ un polynôme homogène (*hamiltonien macroscopique*) de coefficients $\mathrm{V} : \mathbb{R}^d \to \mathbb{R}$ un *potentiel de confinement* et pour tout $k \in \{2, \dots, \mathrm{N}\}$, $\mathrm{W}^{(k)} : (\mathbb{R}^d)^k \to \mathbb{R}$ un *potentiel d'interaction* entre k particules. On considère l'*équation de McKean-Vlasov*

$$d\mathrm{X}_t = \sqrt{2} d\mathrm{B}_t - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t, \mathrm{X}_t) dt \tag{0.2.1}$$

avec $\mu_t := \mathbb{P}_{\mathrm{X}_t}$. $(\mu_t)_{t \geq 0}$ est solution de

$$\partial_t \mu_t = \mu_t \mathscr{L}_{\mu_t} \Big(\Longleftrightarrow \partial_t \mu_t = \Delta \mu_t + \nabla \cdot \Big(\mu_t \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t, \cdot) \Big) \Big) \tag{0.2.2}$$

avec $\frac{\delta \mathrm{H}}{\delta m}$ une dérivée plate de H définie par

$$\begin{aligned} \forall\, t \in [0,1] \quad & \forall \gamma_t := t\mu_1 + (1-t)\mu_0 \in \mathscr{P}(\mathbb{R}^d), \\ & \frac{d}{dt} \mathrm{H}(\gamma_t) = \int \frac{\delta \mathrm{H}}{\delta m}(\gamma_t, \cdot) d\dot{\gamma}_t \end{aligned} \tag{0.2.3}$$

et $\mathscr{L}_\mu := \Delta - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, \cdot) \cdot \nabla$. On considère aussi l'analogue cinétique.

Soit l'*hamiltonien microscopique*

$$\begin{aligned} \mathrm{H}_n : (\mathbb{R}^d)^n &\longrightarrow \mathbb{R} \\ x &\longmapsto n\mathrm{H}\Big(\frac{1}{n} \sum_{p=1}^{n} \delta_{x_p} \Big) = \sum_{p=1}^{n} \mathrm{V}(x_p) + \sum_{k=2}^{\mathrm{N}} n \mathrm{U}_n(\mathrm{W}^{(k)}). \end{aligned} \tag{0.2.4}$$

10. Guillin et al. ont établi des inégalités fonctionnelles non linéaires de type Poincaré et Sobolev logarithmiques pour l'équation des milieux granulaires. Le chapitre 1 est consacré à l'article [6] qui généralise ces résultats à l'équation de Langevin de champ moyen. Plus précisément, comme expliqué précédemment dans la deuxième sous-section sur les fonctionnelles de Lyapunov, le chapitre 1 est connecté à l'approche indirecte : via une inégalité fonctionnelle uniforme (en le nombre de particules) pour la mesure invariante du système de particules, on établit une dissipation d'entropie qui implique une contraction exponentielle en métrique de Wasserstein. Les principaux résultats du chapitre sont présentés dans les théorèmes 1.4.1, 1.4.2, 1.4.3 et 1.4.4. Des exemples concrets sont proposés dans la sous-section 1.4.3.

On considère le système de particules défini par l'*équation de Langevin*

$$dX_t^{(n)} = \sqrt{2}dB_t^{(n)} - \nabla H_n(X_t^{(n)})dt. \tag{0.2.5}$$

On considère aussi l'analogue cinétique donné pour tout $z := (x_1,\dots,x_n,v_1,\dots,v_n) \in \mathbb{R}^{2nd}$,

$$H_n^Z(z) = \frac{1}{2}\sum_{j=1}^{n}|v_j|^2 + 2V(x_j) + n\sum_{k=2}^{N} U_n(W^{(k)}) \tag{0.2.6}$$

et $Z^n := (X^{n,1},\dots,X^{n,n},V^{n,1},\dots,V^{n,n}) \in (\mathbb{R}^d\times\mathbb{R}^d)^n$ par

$$\begin{cases} dX_t^{n,i} = \nabla_{v_i} H_n^Z(Z_t^n)dt \\ dV_t^{n,i} = -\Big(\nabla_{x_i} H_n^Z(Z_t^n) + \nabla_{v_i} H_n^Z(Z_t^n)\Big)dt + \sqrt{2}dB_t^i. \end{cases} \tag{0.2.7}$$

La figure 3 résume les principaux résultats du chapitre 1 : établir les quatres flèches du diagramme.

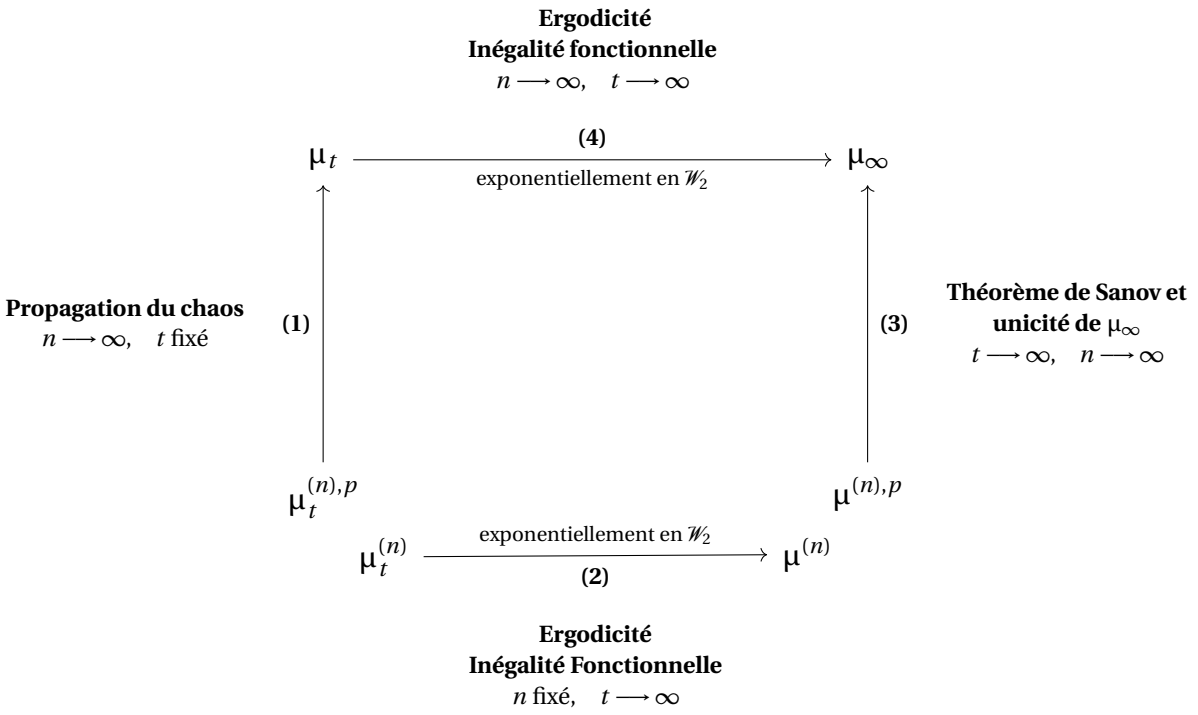

FIGURE 3 – Diagramme des convergences

▷ Quelques notations et définitions :

— $\mathscr{W}_2$ la L^2-*métrique de Wasserstein*

— $\mathrm{U}_n(\mathrm{W}^{(k)})(x_1,\dots,x_n) := \frac{1}{\binom{n}{k}}\sum_{1\le i_1<i_2<\dots<i_k\le n}\mathrm{W}^{(k)}(x_{i_1},\dots,x_{i_k})$ la *U-statisque* d'*ordre* k et de *noyau* $\mathrm{W}^{(k)}$

— $(\mathrm{P}_t^{(n)})_{t\ge0}$ le *semi-groupe* associé à $\mathrm{X}^{(n)}$ et $\mathscr{L}^{(n)} = \Delta - \nabla \mathrm{H}_n\cdot\nabla$ son *générateur infinitésimal*

— $\mu_t^{(n)} := \mathbb{P}_{\mathrm{X}_t^{(n)}} = \mu_0^{\otimes n}\mathrm{P}_t^{(n)}$ la *loi* au temps t du *système*

— $\mu_t^{(n),p}$ la *loi* au temps t d'une *particule du système*

— $\mu^{(n)}(dx) = \frac{1}{\mathrm{Z}_n}e^{-\mathrm{H}_n(x)}dx$ la *mesure de Gibbs* associée à H_n.

— $\mathbf{H}[\nu|\mu] = \begin{cases}\mathbb{E}_\nu[\log\frac{d\nu}{d\mu}] =: \mathbf{Ent}_\mu[\frac{d\nu}{d\mu}] & \text{si } \nu\ll\mu,\\ +\infty & \text{sinon.}\end{cases}$ l'*entropie relative* entre deux mesures.

— $\mathbf{I}[\nu|\mu] = \frac14\int|\nabla\frac{\delta\mathbf{H}[\cdot|\mu]}{\delta m}(\nu,y)|^2\nu(dy) = \frac14\int|\nabla\log\frac{d\nu}{d\mu}|^2d\nu = \int|\nabla\sqrt{\frac{d\nu}{d\mu}}|^2d\mu$ l'*information de Fisher relative.*

— $(\mathrm{P}_t^\mu)_{t\ge0}$ le *semi-groupe* associé à X^μ le processus de *diffusion* de *générateur* $\mathscr{L}_\mu = \Delta - \nabla\frac{\delta\mathrm{H}}{\delta m}(\mu,\cdot)\cdot\nabla$

— $\Phi(\mu)(dx) = \frac{1}{\mathrm{Z}_\mu}e^{-\frac{\delta\mathrm{H}}{\delta m}(\mu,\cdot)}dx$ la *mesure de Gibbs* associée à $\frac{\delta\mathrm{H}}{\delta m}(\mu,\cdot)$

— $\mathbf{E}_\mathrm{W}[\nu] = \mathbf{H}[\nu|\alpha] + \mathrm{H}(\nu) - \int\mathrm{V}d\nu = \mathbf{H}[\nu|\Phi(\nu)] + c$ l'*énergie de champ moyen*

— $\mathbf{H}_\mathrm{W} = \mathbf{E}_\mathrm{W} - \inf\mathbf{E}_\mathrm{W}$ l'*entropie de champ moyen*

— $\mathbf{I}_\mathrm{W}[\nu] = \frac14\int|\nabla\frac{\delta\mathbf{H}_\mathrm{W}}{\delta m}(\nu,y)|^2\nu(dy) = \frac14\int|\nabla\log\frac{d\nu}{d\alpha} + \nabla\frac{\delta\mathrm{H}}{\delta m}(\nu,\cdot) - \nabla\mathrm{V}|^2d\nu$ l'*information de Fisher de champ moyen*

▷ **Flèche 1.** Dans un premier temps, nous abordons la notion de propagation du chaos : nous considérons les systèmes de particules sous-jacents régis respectivement par des équations de Langevin standard et cinétique d'hamiltoniens microscopiques fonctions de U-statistiques définies ci-dessus. On construit sur le même espace de probabilité (c'est-à-dire avec le même caractère aléatoire) le système de particules et n processus non linéaires indépendants (corrélés à lui) : n copies de $(\mathrm{X}_t)_{t\ge0}$ telles que $(\mathrm{X}_t^{(p)})_{t\ge0}$ et $(\mathrm{X}_t^{(n),p})_{t\ge0}$ ont la même condition initiale $\mu_0\in\mathscr{P}_2(\mathbb{R}^d)$ et sont dirigés par le même mouvement brownien.
Quand le nombre de particules tend vers l'infini, une particule du système décrit la dynamique non linéaire. Plus précisément, si on part initialement avec des particules indépendantes et identiquement distribuées, asymptotiquement elles restent indépendantes et identiquement distribuées selon la loi de la dynamique non linéaire. Nous proposons des conditions donnant des estimées uniformes en temps.

Théorème 1.5.1 (L^2-Propagation du chaos)

Supposons que $\exists \beta \in \mathbb{R}\ \forall k \in \{2,\dots,N\}\ \exists \beta_k \in \mathbb{R}\ \forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$\langle \nabla V(x) - \nabla V(y), x-y\rangle \geq -\beta |x-y|^2; \tag{0.2.8}$$

$$\langle \nabla_{x_1} W^{(k)}(x,\cdot) - \nabla_{x_1} W^{(k)}(y,\cdot), x-y\rangle \geq -\beta_k |x-y|^2. \tag{0.2.9}$$

Alors, par *symétrie*, pour tout $T > 0$, il existe une constante K_T telle que

$$\forall \mu_0 \in \mathscr{P}_2(\mathbb{R}^d), \quad \sup_{0\leq t\leq T} \mathbb{E}|X_t^{(n),1} - X_t^{(1)}|^2 \leq \frac{K_T}{n}. \tag{0.2.10}$$

De plus, si $\omega := \beta + \sum_{k=2}^{N} k\beta_k < 0$, cette borne supérieure est **uniforme en temps**, *i.e.*,

$$\sup_{t\geq 0} \mathbb{E}|X_t^{(n),1} - X_t^{(1)}|^2 \leq \frac{K_\infty}{n}, \quad K_\infty < +\infty. \tag{0.2.11}$$

▷ **Flèche 2.** Dans un second temps, nous étudions le comportement en temps long du système de particules via une inégalité fonctionnelle uniforme en le nombre de particules pour la mesure de Gibbs invariante du système.

Évolution du flot $(\mu_t^{(n)})_{t\geq 0}$ et maxwelliennes. De

$$\partial_t \mu_t^{(n)} = \mu_t^{(n)} \mathscr{L}^{(n)}, \tag{0.2.12}$$

avec les *hypothèses* de *régularité* et de *stabilité* sur les *potentiels*, pour tout $t > 0$, $\mu_t^{(n)}$ a une *densité* $f_t^{(n)}$ solution de l'équation de Fokker-Planck

$$\partial_t f_t^{(n)} = \mathbf{div}(f_t^{(n)} \nabla H_n + \nabla f_t^{(n)}). \tag{0.2.13}$$

Une *probabilité invariante* $\mu_\infty^{(n)}$ vérifie $\mu_\infty^{(n)} \mathscr{L}^{(n)} = 0$ ou de façon équivalente (existence de densité) $\mathbf{div}(f_\infty^{(n)} \nabla H_n + \nabla f_\infty^{(n)}) = 0$. Il en découle que $\mu_\infty^{(n)} = \mu^{(n)}$. De plus, $\mu^{(n)}$ *minimise* l'entropie relative $\mathbf{H}[\cdot|\mu^{(n)}]$. C'est un *point critique* de l'*entropie relative* : $\frac{\delta \mathbf{H}[\cdot|\mu^{(n)}]}{\delta m}(\mu^{(n)},\cdot) = 0$.

Condition sur la probabilité invariante $\mu^{(n)}$. $\exists \rho_n > 0$ telle que $\limsup_n \rho_n > 0$ et $\rho_n \mathbf{H}[\cdot|\mu^{(n)}] \leq 2\mathbf{I}[\cdot|\mu^{(n)}]$ (*inégalité de Sobolev logarithmique uniforme* en le nombre de particules). Notons $\rho_{LS}(\mu^{(n)})$ la *constante optimale* pour laquelle $\mu^{(n)}$ vérifie une inégalité de Sobolev logarith-

mique. Cas satisfaisants : *critère de courbure* de *Bakry-Emery* ou *bornitude* des *oscillations* de l'*hamiltonien* de $\frac{d\mu^{(n)}}{d\alpha^{\otimes n}}$ (α mesure de Gibbs associée au confinement).

Inégalité fonctionnelle et (hypo)coercivité. Comme

$$\frac{d}{dt}\mathbf{H}[\mu_t^{(n)}|\mu^{(n)}] = -4\mathbf{I}[\mu_t^{(n)}|\mu^{(n)}] \leq -2\rho_n\mathbf{H}[\mu_t^{(n)}|\mu^{(n)}],$$

par le *lemme de Gronwall*, on a

$$\mathbf{H}[\mu_t^{(n)}|\mu^{(n)}] \leq \mathbf{H}[\mu_0^{(n)}|\mu^{(n)}]e^{-2\rho_n t}.$$

Comme l'inégalité de Sobolev logarithmique implique la T_2-*inégalité de Talagrand*, cette *dissipation d'entropie* implique une *contraction en métrique de Wasserstein* :

$$\rho_n\mathcal{W}_2^2(\mu_t^{(n)},\mu^{(n)}) \leq 2\mathbf{H}[\mu_t^{(n)}|\mu^{(n)}].$$

▷ **Flèche 3.** Ensuite, via des résultats de grandes déviations de type Sanov sur les U-statistiques, nous établissons des relations entre la mesure de probabilité invariante du système de particules et la mesure de probabilité invariante du processus de McKean-Vlasov.

Liu and Wu ont montré un théorème de *grandes déviations* de type *Sanov* sur les *U-statistiques* dans lequel $\mathbf{E}_W$ est une *bonne fonction taux* de convergence, ce qui implique que $\mathbf{H}_W$ admet au moins un *minimiseur* μ_∞ sur $\mathscr{P}(\mathbb{R}^d)$ par *inf-compacité*. Nous montrons ainsi l'existence d'une probabilité invariante.

Proposition 1.5.9 (Existence d'une probabilité invariante)

Il y a équivalence entre

— **Points critiques de l'énergie.** $\mu_\infty \in \{\nu, \quad \frac{\delta \mathbf{E}_W}{\delta m}(\nu,\cdot) = 0\} \neq \emptyset$.

— **Points fixes.** $\mu_\infty \in \{\nu, \quad \Phi(\nu) = \nu\} \neq \emptyset$.

— **Maxwelliennes.** μ_∞ est une *probabilité invariante* du processus de McKean-Vlasov : $\mu_\infty\mathscr{L}_{\mu_\infty} = 0$ ou de façon équivalente (existence de *densité*) $\mathbf{div}(f_\infty\nabla\frac{\delta H}{\delta m}(\mu_\infty,\cdot) + \nabla f_\infty) = 0$.

Quant à l'*unicité*, elle n'est pas assurée : en *physique statistique*, on parle de *transition de phase*. Nous donnons une condition assurant l'unicité et une conséquence de cette dernière.

L'*unicité* est assurée via une *propriété de contraction* sur l'*application mesure invariante* Φ. En cas d'unicité, on a $\mu^{(n),p} \overset{n\to+\infty}{\longrightarrow} \mu_\infty$ avec $\mu^{(n),p}$ une *marginale* de $\mu^{(n)}$.

Le *théorème de Sanov* implique $\frac{1}{n}\ln Z_n - C \overset{n\to+\infty}{\longrightarrow} -\inf \mathbf{E}_W$ grâce à laquelle on montre la relation entre les entropies.

Proposition 1.5.10 (H-Tensorisation)

Pour toute mesure de probabilité ν sur $\mathbb{R}^d$ telle que $\mathbf{H}[\nu|\alpha] < +\infty$, on a

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu^{(n)}] \overset{n\to+\infty}{\longrightarrow} \mathbf{H}_W[\nu]. \tag{0.2.14}$$

La *loi des grands nombres* nous permet d'avoir la relation entre les informations de Fisher.

Proposition 1.5.11 (I-Tensorisation)

Pour toute mesure de probabilité ν sur $\mathbb{R}^d$ telle que $\mathbf{I}[\nu|\alpha] < +\infty$, on a

$$\frac{1}{n}\mathbf{I}[\nu^{\otimes n}|\mu^{(n)}] \overset{n\to+\infty}{\longrightarrow} \mathbf{I}_W[\nu]. \tag{0.2.15}$$

▷ **Flèche 4.** Pour finir, nous obtenons une inégalité fonctionnelle non linéaire pour la mesure de probabilité invariante du processus de McKean-Vlasov qui, comme dans le cas markovien, encode le comportement en temps long à taux exponentiel du flot des distributions de l'équation de McKean-Vlasov.

Théorèmes 1.4.1 et 1.4.2 (Convergence exponentielle)

Les *conditions sur les potentiels*, l'*inégalité de Sobolev uniforme* et la *condition de contraction* impliquent que $\exists!\ \mu_\infty \in \mathscr{P}_2(\mathbb{R}^d)$ telle que $\mu_\infty = \mathbf{argmin}\left\{\mathbf{H}_W[\nu], \nu \in \mathscr{P}_1(\mathbb{R}^d)\right\}$, avec $\mathbf{H}_W$ l'entropie de champ moyen. $\rho_{LS} := \limsup_{n\to+\infty} \rho_{LS}(\mu^{(n)}) > 0$ vérifie : $\forall \nu \in \mathscr{P}_2(\mathbb{R}^d)$, $\rho_{LS}\mathbf{H}_W[\nu] \le 2\mathbf{I}_W[\nu]$ et $\rho_{LS}\mathscr{W}_2^2(\nu,\mu_\infty) \le 2\mathbf{H}_W[\nu]$. On parle respectivement d'*inégalités de Sobolev et de Talagrand non linéaires*. Il en découle une *dissipation d'entropie* : $\forall\, t \ge 0$, $\mathbf{H}_W[\mu_t] \le \mathbf{H}_W[\mu_0]e^{-2\rho_{LS}t}$ et une *contraction en métrique de Wasserstein* :

$$\forall\, t \ge 0, \quad \mathscr{W}_2^2(\mu_t,\mu_\infty) \le \frac{2}{\rho_{LS}}\mathbf{H}_W[\mu_0]e^{-2\rho_{LS}t}. \tag{0.2.16}$$

Dans le cas cinétique, on a les résultats qui suivent.

Théorème 1.4.3 (H^1-Convergence exponentielle)

$$\exists\alpha>0 \quad \exists\beta>0 \quad \forall n \ge 2, \quad \left|\left|\left|P_t^{Z,(n)} - \mu_Z^n\right|\right|\right|_{H^1\to H^1} \le \alpha e^{-\beta t}. \tag{0.2.17}$$

Théorème 1.4.4 (Convergence exponentielle)

Il existe des constantes $C>0$, $\xi>0$ et $\kappa>0$ telles que $\forall \mu \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, $\forall n \ge 2$ et $\forall\, t>0$,

$$\mathbf{H}[\mu_Z^n(t)|\mu_Z^n] \le C\mathbf{H}[\mu_Z^n(0)|\mu_Z^n]e^{-\xi t}, \tag{0.2.18}$$

$$\mathscr{W}_2^2(\mu_t^{\mathbf{VFP}},\mu_\infty^Z) \le \kappa C\mathscr{S}[\mu]e^{-\xi t}, \tag{0.2.19}$$

où μ est la condition initiale et $\mathscr{S}$ (défini dans eq. (1.3.24)) est l'entropie de champ moyen associée à notre système du second ordre donné par eq. (1.3.18).

0.2.2 Contenu du chapitre 2

Le second chapitre étudie à la fois le comportement quantitatif à long terme de la dynamique de Langevin de champ moyen du premier ordre (sur-amortie) et du second ordre (sous-amortie) et leurs propagations du chaos uniformes en temps [11]. Dans un premier temps, on donne des conditions suffisantes sur l'hamiltonien macroscopique H assurant la contraction de l'application mesure invariante du processus de Langevin de champ moyen.

Théorème 2.3.1 (Contraction de Φ)

Supposons que la fonctionnelle H admet des dérivées plates du premier et du second ordre

$$\frac{\delta \mathrm{H}}{\delta m} : \mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \longrightarrow \mathbb{R}; \quad \frac{\delta^2 \mathrm{H}}{\delta m^2} : \mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R};$$

qui sont jointement continues et sont $\mathscr{C}^2$ dans les variables spatiales. Supposons de plus que $(\mu, x, y) \mapsto \|\mathscr{D}_m^2 \mathrm{H}(\mu, x, y)\|_2$ est bornée ($\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} < +\infty$) et $(\mu, x) \mapsto \lambda_{\nabla \mathscr{D}_m \mathrm{H}(\mu,x)}$ est strictement bornée par le bas par $\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}$. Dans ce cas,

$$\forall p \geq 0, \quad \Phi : \quad \mathscr{P}_p(\mathbb{R}^d) \longrightarrow \mathscr{P}_p(\mathbb{R}^d), \tag{0.2.20}$$

et pour

$$\underline{\lambda} := \inf_{\mu,x} \lambda_{\nabla \mathscr{D}_m \mathrm{H}(\mu,x)} \quad (> \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}), \tag{0.2.21}$$

pour tout $p \geq 1$, Φ définit une contraction stricte de $(\mathscr{P}_p(\mathbb{R}^d), \mathscr{W}_1)$: $\forall (\mu, \nu) \in \mathscr{P}_p(\mathbb{R}^d) \times \mathscr{P}_p(\mathbb{R}^d)$,

$$\mathscr{W}_1(\Phi(\mu), \Phi(\nu)) \leq \frac{\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}}{\underline{\lambda}} \mathscr{W}_1(\mu, \nu). \tag{0.2.22}$$

Dans ce cas, Φ admet au plus un point fixe sur $\mathscr{P}_p(\mathbb{R}^d)$: si $p = 1$, il existe un unique point fixe.

11. Bolley et al. ont montré par couplage synchrone une contraction exponentielle en métrique de Wasserstein pour l'équation de Vlasov-Fokker-Planck. Ils en ont déduit une propagation du chaos uniforme en temps. Le chapitre 2 (fruit de l'article [1]) étend ces résultats aux équations de Langevin de champ moyen. Des exemples concrets sont proposés dans la section 2.4.

Dans un second temps, on montre par couplage synchrone une contraction exponentielle en métrique $\mathscr{W}_2$ de Wasserstein.

Théorème 2.3.2 ($\mathscr{W}_2$-Contraction par couplage synchrone)

Supposons que la fonctionnelle H admette des dérivées plates du premier et du second ordre qui sont jointement continues et sont $\mathscr{C}^2$ dans les variables spatiales. Supposons que

$$\sup_{\mu\in\mathscr{P}(\mathbb{R}^d)} \left\| \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, 0) \right\|_2 < +\infty \tag{0.2.23}$$

et qu'il existe $\lambda > 0$ tel que

1. pour tout $(\mu, x, y) \in \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d$,

$$\left\langle \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, x) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, y), x - y \right\rangle \geq \lambda \|x - y\|_2^2; \tag{0.2.24}$$

2. l'application $(\mu, x, y) \mapsto \|\mathscr{D}_m^2 \mathrm{H}(\mu, x, y)\|_2$ est bornée ($\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} < +\infty$) et $\lambda > \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}$.

Dans ce cas, pour toutes solutions $(\mathrm{X}_t)_{t\geq 0}$ et $(\mathrm{Y}_t)_{t\geq 0}$ de l'équation de Langevin de champ moyen eq. (2.1.1) de conditions initiales $\mu_0^{\mathrm{X}} \in \mathscr{P}_2(\mathbb{R}^d)$ et $\mu_0^{\mathrm{Y}} \in \mathscr{P}_2(\mathbb{R}^d)$, on a

— pour tout $t \geq 0$, $\mu_t^{\mathrm{X}} \in \mathscr{P}_2(\mathbb{R}^d)$;

— pour tout $t \geq 0$,

$$\mathscr{W}_2(\mu_t^{\mathrm{X}}, \mu_t^{\mathrm{Y}}) \leq e^{(\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - \lambda)t} \mathscr{W}_2(\mu_0^{\mathrm{X}}, \mu_0^{\mathrm{Y}}). \tag{0.2.25}$$

En particulier, il existe une unique mesure invariante $\mu_\infty \in \mathscr{P}_2(\mathbb{R}^d)$ et

$$\forall\, t \geq 0, \quad \mathscr{W}_2(\mu_t^{\mathrm{X}}, \mu_\infty) \leq e^{(\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - \lambda)t} \mathscr{W}_2(\mu_0^{\mathrm{X}}, \mu_\infty). \tag{0.2.26}$$

Ensuite, on en déduit une L^2-propagation du chaos uniforme en temps avec une borne de Fournier and Guillin. Soit $(\widetilde{\mathrm{X}}^{(n)})_{n\geq 1}$ une suite de copies indépendantes du processus de Langevin de champ moyen $(\mathrm{X}_t)_{t\geq 0}$ couplées de manière synchrone aux particules $(\mathrm{X}^{(n)})_{n\geq 1}$ du système de Langevin de champ moyen. Si la fonctionnelle d'énergie H vérifie les hypothèses

du théorème 2.3.2, on a

$$\forall\, t \geq 0, \quad \left\langle \mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle \leq \begin{cases} \left\langle \mu_0^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle e^{\beta t} + \frac{\alpha}{\beta}(e^{\beta t} - 1) & \text{si} \quad \beta \neq 0; \\ \left\langle \mu_0^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle + \alpha t & \text{si} \quad \beta = 0. \end{cases} \tag{0.2.27}$$

$$\alpha := 2d + 2\sup_{\mu} \left\| \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, 0) \right\|_2; \tag{0.2.28}$$

$$\beta := -2\Big(\lambda - \sup_{\mu} \left\| \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, 0) \right\|_2\Big); \tag{0.2.29}$$

De plus, pour tout $\mathrm{T} \geq 0$, on a

$$\forall\, t \in [0, \mathrm{T}], \quad \mathbb{E}[\|\mathrm{X}_t^{(1)} - \widetilde{\mathrm{X}}_t^{(1)}\|^2] \leq \begin{cases} \frac{\alpha_1(\mathrm{T})}{\beta_1}(e^{\beta_1 t} - 1) & \text{si} \quad \beta_1 \neq 0; \\ \alpha_1(\mathrm{T})\, t & \text{si} \quad \beta_1 = 0. \end{cases} \tag{0.2.30}$$

$$\alpha_1(\mathrm{T}) := \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} \mathrm{C}(d) \Big(\sup_{t \in [0,\mathrm{T}]} \left\langle \mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle \Big) \delta_d(n); \tag{0.2.31}$$

$$\delta_d(n) := \begin{cases} \frac{1}{\sqrt{n}} & \text{si} \quad d < 4; \\ \frac{\ln(n+1)}{\sqrt{n}} & \text{si} \quad d = 4; \\ \frac{1}{n^{\frac{2}{d}}} & \text{si} \quad d > 4; \end{cases} \tag{0.2.32}$$

$$\beta_1 := 3\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - 2\lambda. \tag{0.2.33}$$

Théorème 2.3.3 (L^2-Propagation du chaos uniforme)

La borne est uniforme en temps si, et seulement si, $\beta < 0$ et $\beta_1 < 0$:

$$\sup_{t \geq 0} \mathbb{E}[\|\mathrm{X}_t^{(1)} - \widetilde{\mathrm{X}}_t^{(1)}\|^2] \leq \frac{\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} \mathrm{C}(d)}{2\lambda - 3\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}} \Big(\left\langle \mu_0^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle - \frac{\alpha}{\beta} \Big) \delta_d(n). \tag{0.2.34}$$

Quant au cas cinétique, on a les résultats qui suivent.

Théorème 2.3.4 ($\mathcal{W}_2$-Contraction par couplage synchrone)

Supposons que

— $[\mathrm{A}]_1 < +\infty$, $\mathrm{B} : p \mapsto \lambda_{\mathrm{B}} p + \mathrm{D}(p)$ avec $\lambda_{\mathrm{B}} \geq 0$ et $[\mathrm{D}]_1 < +\infty$.

— Il existe $\lambda_{\mathrm{A}} \geq 0$ tel que pour tout $(v, w) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$\langle \mathrm{A}(v) - \mathrm{A}(w), v - w \rangle \geq \lambda_{\mathrm{A}} |v - w|^2. \tag{0.2.35}$$

— $[\mathscr{D}_m \mathrm{H}]_{1,\infty} < +\infty$ et $\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} < +\infty$.

Dans ce cas, il existe deux constantes positives C et C′ telles que

$$\forall\, t \geq 0, \quad \mathcal{W}_2(\mathbb{P}_{(\mathrm{P}_t, \mathrm{V}_t)}, \mathbb{P}_{(\overline{\mathrm{P}}_t, \overline{\mathrm{V}}_t)}) \leq \mathrm{C}' e^{-\mathrm{C}t} \mathcal{W}_2(\mathbb{P}_{(\mathrm{P}_0, \mathrm{V}_0)}, \mathbb{P}_{(\overline{\mathrm{P}}_0, \overline{\mathrm{V}}_0)}). \tag{0.2.36}$$

De plus, il existe une unique solution stationnaire ξ_∞ et

$$\forall\, t \geq 0, \quad \mathcal{W}_2(\mathbb{P}_{(\mathrm{P}_t, \mathrm{V}_t)}, \xi_\infty) \leq \mathrm{C}' e^{-\mathrm{C}t} \mathcal{W}_2(\mathbb{P}_{(\mathrm{P}_0, \mathrm{V}_0)}, \xi_\infty). \tag{0.2.37}$$

Théorème 2.3.5 (L^2-Propagation du chaos uniforme)

Sous les hypothèses du théorème 2.3.4, il existe une constante positive C indépendante du nombre de particules et telle que

$$\sup_{t \geq 0} \mathbb{E}\left[|\mathrm{P}_t^{(i)} - \overline{\mathrm{P}}_t^{(i)}|^2 + |\mathrm{V}_t^{(i)} - \overline{\mathrm{V}}_t^{(i)}|^2 \right] \leq \mathrm{C} \begin{cases} \frac{1}{\sqrt{n}} & \text{si} \quad d < 4; \\ \frac{\ln(n+1)}{\sqrt{n}} & \text{si} \quad d = 4; \\ \frac{1}{n^{\frac{2}{d}}} & \text{si} \quad d > 4. \end{cases} \tag{0.2.38}$$

C ne dépend que des coefficients de l'équation et du second moment de la condition initiale.

0.2.3 Contenu du chapitre 3

Dans le troisième chapitre, on s'intéresse à l'approximation des QSDs des processus de McKean-Vlasov [12].

Soit $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$ une famille de noyaux sous-markoviens sur $\mathscr{E}$ et soit $(Y_n)_{n\geq 0}$ la suite définie sur $\mathscr{E}\cup\{\partial\}$ par : $Y_0 \sim \nu_0$,

$$Y_{n+1} = \begin{cases} Z_{n+1} & \text{si } \tau > n+1 \\ \partial & \text{si } \tau \leq n+1, \end{cases} \tag{0.2.39}$$

avec

12. Benaim et al. ont construit et étudié un algorithme basé sur l'(auto-)interaction d'une dynamique avec sa mesure d'occupation pour approximer des distributions quasi-stationnaires (QSDs) de chaînes de Markov générales conditionnées à rester dans un ensemble compact : pour plus de détails, voir la troisième sous-section précédente sur les distributions quasi-stationnaires. Le chapitre 3 (fruit de l'article [2] en collaboration avec Fabien PANLOUP) propose une extension au cas des dynamiques de McKean-Vlasov absorbées : convergence d'algorithmes auto-attractifs vers la QSD pour des dynamiques avec interaction de type McKean-Vlasov. Les QSDs de McKean-Vlasov, contrairement au cadre markovien, ont été très peu étudiées : à notre connaissance, le seul article sur le sujet est celui de Tough and Nolen sur les systèmes de particules de Fleming-Viot pour les dynamiques absorbées de type McKean-Vlasov. Ici, on s'intéressera principalement à des dynamiques en temps discret et on commencera par rappeler/étendre des propriétés standard des QSDs au cadre McKean-Vlasov puis on se focalisera sur l'étude de notre algorithme "auto-attractif". Comme indiqué plus haut, l'objectif est ici de généraliser un résultat de convergence de dynamique auto-attractive vers la QSD. Par dynamique auto-attractive, l'idée est ici de considérer la dynamique d'une particule qui, lorsqu'elle meurt, est régénérée selon sa mesure d'occupation, c'est-à-dire en tirant un point au hasard dans le passé de la trajectoire. Cette approche se distingue des algorithmes dits de Fleming-Viot où l'on considère un système de particules et lorsqu'une d'elles meurt, elle est regénérée selon la mesure empirique du système, c'est-à-dire en tirant au hasard parmi les autres particules. Après avoir considéré le problème "trop difficile" consistant à regarder le problème dit de *hard-killing* en temps continu dont subsistent quelques résultats dans la partie Conclusion & Perspectives, l'idée a été de considérer des dynamiques à temps discret, typiquement des schémas d'Euler d'équations de McKean-Vlasov absorbées. Dans ce cadre, un résultat général de convergence (Théorème 3.2.8) est obtenu de type "tension + identification de la limite" à l'aide d'outils issus des algorithmes stochastiques consistant à prouver que la suite peut être vue comme une *pseudo-trajectoire* d'une EDO dont les points fixes sont des QSDs. Ce travail contient plusieurs résultats vraiment nouveaux tels que la tension de la suite (dans le cas non compact!) sous des hypothèses de Lyapounov ainsi que l'identification de la limite qui requiert des outils de couplage fins permettant de pallier la non-unicité des objets limites dans ce cadre non-linéaire. Des applications sont proposées dans lesquelles sont prouvées la convergence de l'algorithme pour le schéma d'Euler dans le cas compact et dans un cas non-compact (pas tout à fait standard). Une étude du comportement limite lorsque le pas tend vers 0 montre que l'on retrouve alors les QSDs d'EDS de McKean-Vlasov absorbées dans le cas où l'absorption a lieu en sortant d'un compact, ce qui permet de produire une réponse "faible" au problème posé initialement associé au hard-killing en temps continu. L'adjectif "faible" fait référence à la double limite : faire tendre le temps vers l'infini puis le pas vers zéro (s'opposant à une approche forte dans laquelle la dynamique à temps continu serait directement considérée). Enfin, nous illustrons nos résultats théoriques par plusieurs simulations (voir les figures 3.1 et 3.2) : des exemples concrets sont proposés dans la sous-section 3.2.2 et nous proposons des simulations dans la sous-section 3.2.3.

— τ un $(\mathscr{F}_n)_{n\geq 0}$ temps d'arrêt défini par

$$\tau := \inf\{n \geq 1, \Theta_n = 1\},$$

avec $(\Theta_n)_{n\geq 1}$ une suite de variables aléatoires de Bernoulli telle que

$$\mathbb{P}(\Theta_{n+1} = 1 | \mathscr{F}_n, Y_n \in \mathscr{E}) = \delta_{\nu_n}(Y_n) \quad \text{avec} \quad \delta_\mu(x) = 1 - K_{\mu,\partial}\mathbf{1}(x),$$

$(\mathscr{F}_n)_{n\geq 0}$ est définie par $\mathscr{F}_0 = \sigma(Y_0)$ et pour $n \geq 1$,

$$\mathscr{F}_n = \sigma(Y_0, \ldots, Y_n, \Theta_1, \ldots, \Theta_n),$$

et

$$\nu_n = \mathscr{L}(Y_n | \tau > n).$$

— $(Z_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathscr{E}$ telle que

$$\mathscr{L}(Z_{n+1} | \mathscr{F}_n, Y_n \in \mathscr{E}) = K_{\nu_n,\partial}(Y_n, .).$$

La définition ci-dessus implique implicitement que $\mathbb{P}(\tau > n) > 0$. Néanmoins, la définition reste cohérente si $\mathbb{P}(\tau > n) = 0$, car dans ce cas, $Y_n = \partial$ est presque certain.

Définition 0.2.1. Une mesure de probabilité $\nu \in \mathscr{P}(\mathscr{E})$ est appelée une distribution quasi-stationnaire (QSD abréviation en anglais) pour $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$ si $Y_0 \sim \nu$ implique que

$$\forall n \in \mathbb{N}, \quad \mathbb{P}(Y_n \in ., \tau > n) = \nu(.)\mathbb{P}(\tau > n). \tag{0.2.40}$$

Par une induction, il est aisé de vérifier que sous la QSD, τ suit une loi exponentielle de paramètre $\lambda = \nu K_{\nu,\partial}(\mathscr{E})$. En particulier, $\mathbb{P}_\nu(\tau > n) > 0$ pour tout n et la définition implique donc

$$\forall n \in \mathbb{N}, \quad \mathscr{L}(Y_n | \tau > n) = \nu.$$

On retrouve ainsi les propriétés classiques du cadre markovien. Il est néanmoins important de garder à l'esprit que cette propriété implique également que, sous le régime quasi-stationnaire, la dynamique du processus tué est markovienne (puisqu'à chaque instant n, la distribution d'interaction est également égale à ν). Ceci est similaire au régime stationnaire (classique) de la dynamique avec interaction, où l'on retrouve la propriété de Markov lorsque le processus part de sa distribution invariante.

Avant d'aller plus loin, introduisons deux premières hypothèses :

($\mathbf{A_0}$) : Il existe $\ell \geq 1$ tel que $\sup_{x\in\mathscr{E},\mu\in\mathscr{P}(\mathscr{E})} \mathrm{K}^{\ell}_{\mu,\partial}\mathbf{1}(x) < 1$.
Ici et dans la suite, on écrit $\mathbf{1}$ au lieu de $\mathbf{1}_{\mathscr{E}}$ (de sorte qu'en particulier, $\mathrm{K}^{\ell}_{\mu,\partial}\mathbf{1}(x) = \mathrm{K}^{\ell}_{\mu,\partial}(x,\mathscr{E})$).

($\mathbf{A_1}$) : Pour tout $\mu \in \mathscr{P}(\mathscr{E})$, $\mathrm{K}_{\mu,\partial}$ est Feller et pour tout $x \in \mathscr{E}$, $\mu \mapsto \mu\mathrm{K}_{\mu,\partial}$ est continue pour la topologie de la convergence faible sur $\mathscr{P}(\mathscr{E})$.

Ces hypothèses seront utiles pour une première caractérisation des QSDs basée sur la dynamique avec renouvellement : soit (X^{μ}_n) la chaîne de Markov de noyau

$$\mathbb{K}_{\mu}(x,.) = \mathrm{K}_{\mu,\partial}(x,.) + \delta_{\mu}(x)\mu(.).$$

Cela signifie que lorsque la particule est tuée, elle est régénérée par la distribution μ. Posons

$$\mathscr{A}_{\mu}(x,.) = \sum_{n\geq 0} \mathrm{K}^{n}_{\mu,\partial}(x,.), \tag{0.2.41}$$

Sous **($\mathbf{A_0}$)** et **($\mathbf{A_1}$)**, on peut montrer (voir lemme 3.3.1) que $\mathscr{A}_{\mu}(x,.)$ est bien définie pour tout $(x,\mu) \in \mathscr{E}\times\mathscr{P}(\mathscr{E})$ et que

$$\Pi_{\mu} = \frac{\mu\mathscr{A}_{\mu}}{\mu\mathscr{A}_{\mu}\mathbf{1}}$$

est l'unique distribution invariante pour $(\mathrm{X}^{\mu}_n)_{n\geq 1}$. Nous introduisons une hypothèse de type Lyapunov qui sera utilisée pour l'existence de QSD dans le cas non compact (et la tension dans la suite).

($\mathbf{A_V}$) : Pour tout $\varepsilon > 0$, il existe $\beta_{\varepsilon} \in \mathbb{R}$ et une fonction inf-compacte [13] $\mathrm{V}_{\varepsilon} : \mathscr{E} \to \mathbb{R}_+$ telle que

$$\forall\, \mu \in \mathscr{P}(\mathscr{E}), \quad \mathrm{K}_{\mu,\partial}\mathrm{V}_{\varepsilon} \leq \varepsilon\mathrm{V}_{\varepsilon} + \beta_{\varepsilon}.$$

On se réfère au lemme 3.7.1 pour des conditions concrètes sur la dynamique afin de garantir **($\mathbf{A_V}$)**.

Sous **($\mathbf{A_0}$)** et **($\mathbf{A_1}$)**, on a (voir lemme 3.2.7) :

13. Une fonction $\mathrm{V} : \mathscr{E} \to \mathbb{R}$ est inf-compacte si pour tout $r > 0$, $\mathrm{K}_r = \{x \in \mathscr{E}, \mathrm{V}(x) \leq r\}$ est compacte.

(i) $\mu^\star$ est une QSD pour $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$ si, et seulement si,

$$\mu^\star = \Pi_{\mu^\star}.$$

(ii) De plus, si $\mathscr{E}$ est compact ou si $(\mathbf{A_V})$ est vrai et $\inf_{\mu\in\mathscr{P}(\mathscr{E})} \mu K_{\mu,\partial}\mathbf{1} > 0$, alors l'existence est vraie pour la QSD associée à $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$.

Le fait que $\mu^\star$ puisse être considérée comme un point fixe de $\mu \mapsto \Pi_\mu$ correspondra à la cible naturelle de notre algorithme. Nous entendons par là que notre algorithme est naturellement connecté à l'application $\mu \mapsto \Pi_\mu$ et c'est pourquoi nous avons choisi de rappeler cette caractérisation dans le lemme 3.2.7. Comme mentionné dans la remarque 3.2.12, la condition $\inf_{\mu\in\mathscr{P}(\mathscr{E})} \mu K_{\mu,\partial}\mathbf{1} > 0$ n'est pas très contraignante. Notez que s'il existe $x \in \mathscr{E}$ et μ tels que $K_{\mu,\partial}\mathbf{1}(x) = 0$, alors l'existence est également vérifiée puisque δ_x est une QSD triviale (mais peu intéressante).

Enfin, précisons que nous n'aborderons pas le problème d'unicité dans ce chapitre, qui semble être une question difficile dans ce contexte non linéaire.

Dynamique auto-attractive. Pour approximer les QSDs de $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$, on introduit le processus auto-attractif $(X_n, \mu_n)_{n\geq 0}$ défini récursivement sur $\mathscr{E}\times\mathscr{P}(\mathscr{E})$ par :

$$X_0 = x \in \mathscr{E}, \quad \mu_0 = \delta_x$$

et pour tout $n \geq 0$,

$$\mathbb{P}(X_{n+1} \in .|\mathscr{F}_n) = \mathbb{K}_{\mu_n}(X_n, .) \tag{0.2.42}$$

où $\mathscr{F}_n = \sigma(X_0, \ldots, X_n)$ et

$$\mu_n = \frac{1}{H_n}\sum_{k=1}^{n} \eta_k \delta_{X_{k-1}}$$

est une mesure d'occupation pondérée connexe : les η_k sont supposés être des nombres positifs. Tout au long du chapitre, nous supposons également que

$$H_n = \sum_{k=1}^{n} \eta_k \to +\infty \quad \text{et} \quad \gamma_n := \frac{\eta_n}{H_n} \quad \text{décroît vers 0 lorsque } n \to +\infty.$$

Notons que le choix le plus naturel est $\eta_n = 1$, ce qui conduit à $\gamma_n = \frac{1}{n}$ (voir remarque 3.2.12 pour plus d'informations sur les pondérations). La suite $(\mu_n)_{n\geq 1}$ satisfait la formule récursive

suivante :

$$\mu_{n+1} = \mu_n(1-\gamma_{n+1}) + \gamma_{n+1}\delta_{X_{n+1}}. \tag{0.2.43}$$

L'objectif est de démontrer que les points limites de $(\mu_n)_{n\geq 0}$ sont des QSD de $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$. En suivant [28], nous écrivons artificiellement :

$$\mu_{n+1} = \mu_n + \gamma_{n+1}F(\mu_n) + \gamma_{n+1}\varepsilon_{n+1},$$

avec

$$F(\mu) = -\mu + \Pi_\mu \quad \text{et} \quad \varepsilon_{n+1} = \delta_{X_{n+1}} - \Pi_{\mu_n}.$$

Avec cette forme, une QSD est un zéro de F et l'enjeu est de montrer que (μ_n) se comporte asymptotiquement comme une discrétisation de l'EDO (ce qui signifie que ε_n est asymptotiquement un bruit) et que l'EDO $\dot{\mu} = F(\mu)$ possède des propriétés suffisantes pour assurer les convergences vers ses zéros. Nous verrons que la non-linéarité induira de nombreuses difficultés pour ce point.

Pour énoncer notre résultat principal, nous devons introduire une version continue interpolée de (μ_n). Soit $(t_k)_{k\geq 0}$ la suite croissante de nombres positifs définie par : $t_0 = 0$ et $t_k = \sum_{m=1}^{k}\gamma_m$. Ensuite, nous définissons $(\tilde{\mu}_t)_{t\geq 0}$ par

$$\forall s \in [0,\gamma_k], \qquad \tilde{\mu}_{t_{k-1}+s} = \mu_k + \frac{s}{\gamma_k}(\mu_{k+1} - \mu_k). \tag{0.2.44}$$

$(\tilde{\mu}_t)_{t\geq 0}$ est donc presque sûrement un élément de l'espace $\mathscr{C}(\mathbb{R}_+, \mathscr{P}(\mathscr{E}))$ des processus continus sur $\mathbb{R}_+$ de valeurs $\mathscr{P}(\mathscr{E})$ (pour la topologie faible) et nous considérons la suite décalée associée $(\tilde{\mu}^{(n)})_{n\geq 1}$ définie par :

$$\forall t \geq 0, \quad \tilde{\mu}^{(n)}_t = \tilde{\mu}_{t_n+t}. \tag{0.2.45}$$

Nous introduisons maintenant des hypothèses supplémentaires :

($\mathbf{A_2}$) : $\mu \mapsto K_{\mu,\partial}(x,.)$ est lipschitzien pour la norme TV, uniformément en $x \in \mathscr{E}$: il existe $C > 0$ tel que

$$\sup_{x\in\mathscr{E}} \|K_{\mu,\partial}(x,.) - K_{\nu,\partial}(x,.)\|_{TV} \leq C\|\mu - \nu\|_{TV}.$$

($\mathbf{A_3}$) : Il existe $\ell \in \mathbb{N}$, $\ell \geq 1$, $\varepsilon > 0$ tels qu'il existe $\Psi \in \mathscr{P}(\mathscr{E})$ tel que pour tout $\mu \in \mathscr{P}(\mathscr{E})$,

$$K^{\ell}_{\mu,\partial}(.,dy) \geq \varepsilon\Psi(dy).$$

La dernière hypothèse est liée à la même probabilité Ψ que dans $(\mathbf{A_3})$.

$(\mathbf{A_4})$: Il existe $c>0$ tel que pour tout $k\in\mathbb{N}$, pour tout $\mu_1,\dots,\mu_k\in\mathscr{P}(\mathscr{E})$, pour tout $n_1,\dots,n_k\in\mathbb{N}$,

$$\frac{\Psi(K_{\mu_1,\partial}\circ\dots\circ K_{\mu_k,\partial}\mathbf{1})}{K_{\mu_1,\partial}\circ\dots\circ K_{\mu_k,\partial}\mathbf{1}}\geq c.$$

Cette hypothèse [14] est discutée dans la remarque 3.2.9 et dans le lemme 3.2.10.

Théorème 3.2.8 (Tension, QSDs et EDO)

Supposons $(\mathbf{A_0})$, $(\mathbf{A_1})$, $(\mathbf{A_2})$ et $(\mathbf{A_3})$. Supposons que $(\mu_n)_{n\geq1}$ est *p.s.* tendue sur $\mathscr{P}(\mathscr{E})$.

(i) La suite $(\tilde{\mu}^{(n)})_{n\geq1}$ est presque sûrement tendue sur $\mathscr{C}(\mathbb{R}_+,\mathscr{P}(\mathscr{E}))$ et toute limite faible de $(\tilde{\mu}^{(n)})_{n\geq1}$ est une solution de l'EDO

$$\frac{d\nu_t}{dt}=-\nu_t+\Pi_{\nu_t}. \tag{0.2.46}$$

(ii) Si de plus $(\mathbf{A_4})$ est vraie, alors toute limite faible de $(\tilde{\mu}^{(n)})_{n\geq1}$ est stationnaire, ce qui implique que toute limite faible de $(\mu_n)_{n\geq1}$ est une QSD $\mu^\star$ reliée à $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$.

▷ Dans la sous-section 3.2.2, nous appliquons notre théorème général à deux modèles :

— **Modèle 1 :** (Schéma d'Euler d'une diffusion de McKean-Vlasov absorbée à un instant de sortie). Supposons que D est un ouvert borné de $\mathbb{R}^d$. Soit $h>0$. Soit $(\zeta_t)_{t\geq0}$ un mouvement brownien sur $\mathbb{R}^d$ ou, plus généralement, un processus de Lévy centré donné. On pose également

$$\Delta\zeta_{s,t}=\zeta_t-\zeta_s,\quad 0\leq s\leq t.$$

Soit ∂ un point cimetière. Définissons $\mathscr{E}=\bar{D}$, $Y_0\sim\mu_0\in\mathscr{P}(\mathscr{E})$ et

$$Y_{n+1}=\begin{cases}Y_n+hb(Y_n,\nu_n)+\sigma(Y_n,\nu_n)\Delta\zeta_{nh,(n+1)h} & \text{tant que } \tau>n+1,\\ \partial & \text{si } \tau\leq n+1,\end{cases}$$

où

$$\tau:=\inf\{n\in\mathbb{N},Y_n+hb(Y_n,\nu_n)+\sigma(Y_n,\nu_n)\Delta\zeta_{nh,(n+1)h}\in\mathscr{E}^c\},$$

14. Des conditions suffisantes de minoration et majoration de type Doeblin sont données.

et

$$\nu_n = \mathscr{L}(Y_n | \tau > n).$$

Si μ_0 a un support dans D et que σ n'est pas dégénéré (typiquement uniformément elliptique), τ est presque sûrement égal au temps où le schéma quitte D (et non $\mathscr{E} = \bar{D}$).

Notons que dans cet exemple, $K_{\mu,\partial}$ (noté $K^{(h)}_{\mu,\partial}$) est défini par : pour toute fonction mesurable $f : \mathscr{E} \to \mathbb{R}$,

$$K^{(h)}_{\mu,\partial} f(x) = \mathbb{E}[f(\Upsilon^h_x)\mathbf{1}_{\Upsilon^h_x \in D}] \quad \text{avec} \quad \Upsilon^h_x = x + hb(x,\mu) + \sigma(x,\mu)\Delta\zeta_{0,h}. \tag{0.2.47}$$

Ce schéma d'Euler est relié à la diffusion de McKean-Vlasov :

$$\begin{cases} dY_t = b(Y_t, \nu_t)dt + \sigma(Y_t, \nu_t)dW_t \\ \nu_t = \mathscr{L}(Y_t | \tau > t) \end{cases}$$

où $\tau := \inf\{t \geq 0, Y_t \in D^c\}$.

Dans la proposition 3.2.13, nous fournissons un ensemble d'hypothèses qui garantissent la convergence [15] vers les QSDs de ces dynamiques discrètes et montrons également que lorsque le pas h tend vers 0, nous approximons également les QSDs de la dynamique de McKean-Vlasov avec *hard killing*. Ce résultat montre notamment que l'algorithme peut générer un moyen de simuler des QSDs d'EDS de McKean-Vlasov en temps continu. Dans la sous-section 3.2.3, nous fournissons plusieurs illustrations numériques.

— **Modèle 2 :** (Dynamique absorbée associée aux temps de sortie de sous-ensembles non bornés). Un exemple typique est un schéma de discrétisation de l'équation unidimensionnelle de McKean-Vlasov suivante, absorbée à la sortie de $D =]0, +\infty[$:

$$dY_t = (-\varphi(Y_t) + \int F(Y_t - y)\mu_t(dy))dt + \sqrt{2}d\zeta_t, \tag{0.2.48}$$

où F est supposée être une fonction bornée et

$$\lim_{x \to +\infty} \frac{\varphi(x)}{x} = +\infty.$$

15. Par convergence, nous entendons abusivement la tension combinée au fait que les limites faibles sont des cibles.

Nous considérons donc une dérive avec croissance superlinéaire : cela permettra de partir de l'infini, mais malheureusement, cela générera également de l'instabilité pour la dynamique discrétisée, si l'on considère le schéma d'Euler associé donné par le noyau $K^{(h)}_{\mu,\partial}$, la « partie déterministe » satisfera pour tout $h>0$:

$$\lim_{x\to+\infty} x - h\varphi(x) + \int F(x-y)\mu(dy) = -\infty.$$

Parmi les conséquences, cela impliquera par exemple que

$$\delta_\mu(x) \xrightarrow{x\to+\infty} 1,$$

ce qui implique que la deuxième hypothèse (tension) du théorème 3.2.11 ne sera pas satisfaite. Il existe plusieurs façons de gérer cet effet de discrétisation (par exemple, on peut envisager des séquences d'étapes adaptées, voir *p. ex.* [164]), mais notre objectif étant uniquement de fournir des exemples, nous choisissons d'introduire la modification simple suivante en considérant :

$$\tilde{K}^{(h)}_{\mu,\partial} f(x) = \mathbb{E}[f(\Upsilon^h_x)\mathbf{1}_{\Upsilon^h_x\in D}] \quad \text{avec} \quad \Upsilon^h_x = t(x+hb(x,\mu)) + \sigma(x,\mu)\Delta\zeta_h \tag{0.2.49}$$

avec la condition :

($\mathbf{H_t}$) : la fonction $t:\mathbb{R}^d\mapsto\mathbb{R}^d$ est $\mathscr{C}_1$ et

$$\sup_{x,\mu\in\mathscr{E}\times\mathscr{P}(\mathscr{E})} \{|t(x+hb(x,\mu))| + \|Dt(x+hb(x,\mu))\| < +\infty.$$

Dans la proposition 3.2.15, on montre que (μ_n) est presque sûrement tendue et chaque limite faible est une QSD pour $(\tilde{K}^{(h)}_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$.

▷ Comme indiqué dans la remarque 3.3.4, un argument de poissonisation pourrait éviter de supposer ($\mathbf{A_3}$) dans la première partie du théorème 3.2.8(i) et de remplacer ($\mathbf{A_3}$) par $\mathscr{A}_\mu(.,dy)\geq\varepsilon\Psi(dy)$ dans le théorème 3.2.8(ii). Néanmoins, comme ($\mathbf{A_4}$) requiert en pratique une hypothèse beaucoup plus contraignante, nous avons choisi d'ignorer ce raffinement (qui aurait ajouté des détails techniques).

▷ Pour la dynamique de Markov linéaire, c'est-à-dire lorsque le noyau ne dépend pas de μ, ($\mathbf{A_4}$) est une hypothèse utilisée pour garantir l'unicité de la QSD. Malheureusement,

ce n'est pas le cas dans ce contexte non linéaire (de la même manière que l'ellipticité uniforme peut garantir l'unicité des distributions invariantes pour les diffusions linéaires, mais ce n'est clairement pas le cas pour les EDS de McKean-Vlasov). Néanmoins, elle assure l'uniformité de la convergence vers l'équilibre de l'EDO (3.2.11), ce qui nous aide à déduire la stationnarité des limites faibles de $(\tilde{\mu}^{(n)})_{n\geq 1}$ et à identifier la limite.
Le principal exemple où $(\mathbf{A_4})$ est satisfait est le cas « borne inférieure/supérieure », donné dans le lemme 3.2.10.

▷ Dans le cas non compact, le résultat ci-dessus requiert une tension *p.s.* de $(\mu_n)_{n\geq 1}$. Un critère est donné par le théorème ci-dessous.

Théorème 3.2.11 (Critère de tension)

Si $(\mathbf{A_V})$ est vérifiée, $\sup_{\mu\in\mathscr{P}(\mathscr{E})} \|\delta_\mu\|_\infty < 1$ et $\sum \gamma_n^2 < +\infty$, alors il existe une fonction inf-compacte $\mathcal{V} : \mathscr{E} \to \mathbb{R}_+$ telle que

$$\sup_{n\geq 1} \mu_n(\mathcal{V}) < +\infty \quad p.s.$$

En particulier, (μ_n) est *p.s.* tendue.

▷ Notons que les deux dernières hypothèses ne sont pas contraignantes. La condition sur δ_μ stipule que, uniformément en μ et $x \in \mathscr{E}$, la probabilité de mourir en une étape est positive (ce qui n'est pas contraignant dans notre cadre discret, mais pourrait l'être pour les problèmes de *hard-killing* en temps continu). La condition sur $\sum \gamma_n^2$ est satisfaite dans de nombreux cas. Par exemple, si $\eta_n = n^\alpha$ avec $\alpha \in (-1,\infty)$ alors

$$H_n \sim \frac{n^{1+\alpha}}{1+\alpha} \quad \text{et} \quad \gamma_n \sim \frac{1}{(1+\alpha)n}.$$

On peut vérifier qu'elle est toujours vraie si $\eta_n = e^{n^\alpha}$ avec $\alpha < 1/2$.

Concernant le *hard-killing*, nous y reviendrons plus en détails à la fin de la thèse dans le chapitre Conclusion & Perspectives.

1

Ergodicité exponentielle de limites de champ moyen

Le Hasard est le plus grand romancier du monde ; pour être fécond, il n'y a qu'à l'étudier.

HONORÉ DE BALZAC

Résumé du chapitre

Ce chapitre est le fruit du premier article. Dans l'article [6], il est principalement question de la convergence exponentielle en métrique $\mathscr{W}_2$ de Wasserstein pour le flot des lois de processus de Langevin de champ moyen avec énergie plate dérivant de la dérivée plate d'un polynôme sur l'espace des probabilités : dépendance polynomiale en l'argument mesure.

1.1 Introduction

In the continuity of the recent papers [133] and [132], we establish exponential convergence towards equilibrium for a class of McKean-Vlasov and Vlasov-Fokker-Planck with *polynomial interaction* (macroscopic interaction associated with U-statistics and defined in eq. (1.1.17) and eq. (1.1.18)). Before going further into the details, we recall the general setting related to our problem.

General homogeneous McKean-Vlasov diffusion. The processes studied in this paper belong to the following class of stochastic differential equations:

$$dX_t = b(X_t, \mathbb{P}_{X_t})\,dt + \sigma(X_t, \mathbb{P}_{X_t})\,dB_t, \tag{1.1.1}$$

with respectively $b : \mathbb{R}^D \times \mathscr{P}(\mathbb{R}^D) \longrightarrow \mathbb{R}^D$ the drift coefficient, $\sigma : \mathbb{R}^D \times \mathscr{P}(\mathbb{R}^D) \longrightarrow \mathscr{M}_{D,p}(\mathbb{R})$ the diffusion coefficient and $(B_t)_{t\geq 0}$ a standard $p-$dimensional Brownian motion. More precisely, we are interested in the study of exponential ergodicity of the process defined by

$$dX_t = -(\mathscr{D}_m F(\mathbb{P}_{X_t}, X_t) + \frac{\sigma^2}{2}\nabla V(X_t))\,dt + \sigma\,dB_t, \tag{1.1.2}$$

where $F : \mathscr{P}(\mathbb{R}^D) \to \overline{\mathbb{R}}$, $\mathscr{D}_m F$ is the intrinsic derivative (**L**−derivation or derivation in the sense of Fréchet of F on the probability measure space, see eq. (1.1.26) for precise definition) which is none other than the gradient of a flat derivative (see eq. (1.1.26)) of F: $\mathscr{D}_m F(m,\cdot) := \nabla \frac{\delta F}{\delta m}(m,\cdot)$ (for example, if $F(m) = \int \varphi\, dm$, we have $\frac{\delta F}{\delta m}(m,x) = \varphi(x)$ then, $\mathscr{D}_m F(m,x) = \nabla\varphi(x)$), V is a *confinement potential* and $\sigma > 0$ (in this paper, without loss of generality and for the sake of standardization, we take $\sigma = \sqrt{2}$). Equation (1.1.2) also writes

$$dX_t = -\nabla \frac{\delta H}{\delta m}(\mathbb{P}_{X_t}, X_t)\,dt + \sigma\,dB_t \tag{1.1.3}$$

with the functional H given by

$$H(\mu) := F(\mu) + \frac{\sigma^2}{2}\int V\,d\mu. \tag{1.1.4}$$

With these notations, considering polynomial interactions means that F is a *polynomial on the probability space of degree at least two* (see eq. (1.1.17) for details). The second term being a polynomial of degree 1, the function H is also a polynomial on the probability space

(without constant term).

The related mean-field particle system. The $n-$particle associated with eq. (1.1.2) is given by the following system of SDEs:

$$\forall i \in \{1,\dots,n\}, \quad dX_t^{i,n} = b(X_t^{i,n}, \mu_{X_t^n})dt + \sigma(X_t^{i,n}, \mu_{X_t^n})dB_t^i, \tag{1.1.5}$$

where $B^1,\dots,B^n$ are n independent Brownian motions and $\mu_{X_t^n}$ denotes the empirical measure defined by

$$\mu_x := \frac{1}{n}\sum_{k=1}^n \delta_{x_k}, \quad x = (x_1,\dots,x_n) \in (\mathbb{R}^D)^n.$$

Under standard assumptions, $(X^{i,n})_{i=1}^n$ is a Markov process with infinitesimal generator defined on an appropriate subspace of $\mathscr{C}_b((\mathbb{R}^D)^n)$ by,

$$\mathscr{L}_n\varphi(x) := \sum_{i=1}^n \mathscr{L}_{\mu_x}\blacksquare_i\varphi(x) \tag{1.1.6}$$

where for a given $\mu \in \mathscr{P}(\mathbb{R}^D)$,

$$\mathscr{L}_\mu := b(\cdot,\mu)\cdot\nabla + \frac{1}{2}\mathbf{Tr}(\sigma\sigma^*(\cdot,\mu)\nabla^2), \quad \mu \in \mathscr{P}(\mathbb{R}^D) \tag{1.1.7}$$

and the notation $\mathscr{L}\blacksquare_i\varphi$ denotes the action of an operator $\mathscr{L}$ defined on (a subset of) $\mathscr{C}_b(\mathbb{R}^D)$ against the i-th variable of a function $\varphi \in \mathscr{C}_b((\mathbb{R}^D)^n)$; in other words, $\mathscr{L}\blacksquare_i\varphi$ is defined as the function:

$$x \in (\mathbb{R}^D)^n \longmapsto \mathscr{L}[y \longmapsto \varphi(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_n)](x_i) \in \mathbb{R}.$$

In the family of equations of type eq. (1.1.5), *kinetic* particle systems correspond to the case where $Z_t^{i,n} := (X_t^{i,n}, V_t^{i,n}) \in \mathbb{R}^d \times \mathbb{R}^d$ is a particle defined by two arguments, its position $X_t^{i,n}$ and its velocity $V_t^{i,n}$ defined as the time derivative of the position. The evolution of a system of kinetic particles is usually governed by Newton's laws of motion. In a random setting, the typical system of SDEs is thus the following:

$$\forall i \in \{1,\dots,n\}, \quad \begin{cases} dX_t^{i,n} = V_t^{i,n}dt \\ dV_t^{i,n} = \mathbf{F}(X_t^{i,n}, V_t^{i,n}, \mu_{X_t^n})dt + \sigma(X_t^{i,n}, V_t^{i,n}, \mu_{X_t^n})dB_t^i, \end{cases} \tag{1.1.8}$$

where $\mathbf{F} : \mathbb{R}^d \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathscr{M}_d(\mathbb{R})$. Note that it is often as-

sumed that the force field induced by the interactions between the particles depends only on their positions. Note that in the system eq. (1.1.5) there are actually nD independent one-dimensional Brownian motions. In particular, for kinetic particles defined by their positions and velocities, the noise is often added on the velocity variable only (this case is nevertheless covered by eq. (1.1.5) with a block-diagonal matrix σ with a vanishing block on the position variable). This special case of the McKean-Vlasov diffusion in $\mathbb{R}^{\mathrm{D}} = \mathbb{R}^d \times \mathbb{R}^d$ is also often called a *second order system* by opposition to the *first order systems* when $\mathbb{R}^{\mathrm{D}} = \mathbb{R}^d$. In this paper, we will establish some uniform exponential convergence of the particle systems eq. (1.3.7) and eq. (1.3.18) (defined below) which in turn will allow us to derive the same properties for their mean-field limiting dynamics.

McKean-Vlasov PDE. It is classically assumed that the domain of the generator $\mathscr{L}_\mu$ does not depend on μ. This domain will be denoted by $\mathscr{F} \subset \mathscr{C}_b(\mathbb{R}^{\mathrm{D}})$. In that case, it is easy to guess the form of the associated nonlinear system obtained when $n \to +\infty$. Taking a test function of the form $\varphi(x_1,\dots,x_n) := \psi(x_1)$, where $\psi \in \mathscr{F}$, one obtains the one-particle Kolmogorov equation:

$$\frac{d}{dt}\langle \mathbb{P}_{\mathrm{X}_t^{1,n}}, \psi\rangle = \int_{(\mathbb{R}^{\mathrm{D}})^n} \mathscr{L}_{\mu_x}\varphi(x)\mathbb{P}_{\mathrm{X}_t^n}(dx) = \mathbb{E}[\mathscr{L}_{\mu_{\mathrm{X}_t^n}}\varphi(\mathrm{X}_t^n)]. \tag{1.1.9}$$

Note that the right-hand side depends on the $n-$particle distribution. If the limiting system exists (propagation of chaos) then, its law μ_t at time $t \geq 0$ is typically obtained as the limit of the empirical measure process:

$$\mu_{\mathrm{X}_t^n} \overset{n\to+\infty}{\longrightarrow} \mu_t \tag{1.1.10}$$

This also implies $\mathbb{P}_{\mathrm{X}_t^{1,n}} \overset{n\to+\infty}{\longrightarrow} \mu_t$. Reporting formally in the previous equation, it follows that μ_t should satisfy

$$\Big(\forall \varphi \in \mathscr{F}, \quad \frac{d}{dt}\langle \mu_t, \varphi\rangle = \langle \mu_t, \mathscr{L}_{\mu_t}\varphi\rangle\Big) \Longleftrightarrow \partial_t \mu_t = \mathscr{L}_{\mu_t}^\dagger \mu_t, \tag{1.1.11}$$

where $\mathscr{L}_{\mu_t}^\dagger$ is the weak adjoint of $\mathscr{L}_{\mu_t}$. This is the *weak form* of the so-called the *(nonlinear) evolution equation* induced by (1.1.1). The evolution equation eq. (1.1.11) can be written in a *strong form* (at least formally) and reads:

$$\partial_t \mu_t(x) = -\nabla_x \cdot (b(x,\mu_t)\mu_t) + \frac{1}{2}\sum_{i,j=1}^{\mathrm{D}} \partial_{x_i}\partial_{x_j}\Big((\sigma\sigma^*)_{i,j}(x,\mu_t)\mu_t\Big). \tag{1.1.12}$$

This is a *nonlinear Fokker-Planck equation* which is used in many important modelling prob-

lems. This equation was obtained (formally) previously using only the generators when $n \to +\infty$. Here, there is an alternative way to derive the limiting system: looking at the SDE system eq. (1.1.5), the empirical measure can be formally replaced by its expected limit μ_t. Since all the particles are *exchangeable*, this can be done in any of the n equations. The result is a process $(\overline{X}_t)_{t\geq 0}$ which solves the SDE: (McKean-Vlasov process)

$$d\overline{X}_t = b(\overline{X}_t, \mu_t)dt + \sigma(\overline{X}_t, \mu_t)dB_t, \tag{1.1.13}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion and $\overline{X}_0 \sim \mu_0$. Moreover, since for all i, $X_t^{i,n}$ has law $\mathbb{P}_{X_t^{1,n}}$ and since it is expected that $\mathbb{P}_{X_t^{1,n}} \overset{n\to+\infty}{\longrightarrow} \mu_t$, the process $(\overline{X}_t)_{t\geq 0}$ and the distributions $(\mu_t)_{t\geq 0}$ should be linked by the relation: for all $t \geq 0$, $\overline{X}_t \sim \mu_t$. The dependency of the solution of a SDE on its law is a special case of what is called a nonlinear process in the sense of McKean (eq. (1.1.1) is equivalent to eq. (1.1.11) via mean-field system given by eq. (1.1.6). Under appropriate conditions, the process eq. (1.1.13) is well defined or (equivalently) the PDE eq. (1.1.12) is well-posed (see [62, Proposition.1] or theorem 1.7.2 for details).

Remark 1.1.1. Note that when $\sigma = 0$, the limit equation eq. (1.1.12) is the renowned *Vlasov equation* which is historically one of the first and most important models in plasma physics and celestial mechanics.

Equivalently, our main objective is the study of the long-time behavior of the solution flow of the nonlinear ($\mathscr{D}_m F$ must at least depend on the measure otherwise we find the standard Fokker-Planck PDE) Fokker-Planck equation:

$$\partial_t m = \nabla \cdot \Big((\mathscr{D}_m F(m, \cdot) + \frac{\sigma^2}{2}\nabla V)m + \frac{\sigma^2}{2}\nabla m \Big). \tag{1.1.14}$$

From two-body to many-body interactions. Depending on the form of the drift and diffusion coefficients, the McKean-Vlasov diffusion can be used in a wide range of *modelling problems*. The first case is obtained when b and σ depend linearly on the measure argument. Namely, for $n, m \in \mathbb{N}$, let us consider two functions $K_1 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^n$, $K_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^m$, and let us take $b(x, \mu) = \overline{b}(x, K_1 \star \mu(x))$, $\sigma(x, \mu) = \overline{\sigma}(x, K_2 \star \mu(x))$, where $\overline{b} : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d$, $\overline{\sigma} : \mathbb{R}^d \times \mathbb{R}^m \to \mathscr{M}_d(\mathbb{R})$ and $K_i \star \mu(x) := \int K_i(x, y)\mu(dy)$. When K_1, K_2 and $\overline{b}, \overline{\sigma}$ are Lipschitz and bounded, the propagation of chaos result is the given by McKean's theorem.
In many applications, σ is a constant diffusion matrix, $K_1(x, y) \equiv K(y - x)$ for a (usually symmetric) radial kernel $K : \mathbb{R}^d \to \mathbb{R}^d$ and $b(x, \mu) = K \star \mu(x)$. Note that the case where K has a singularity is much more delicate but contains many important cases (such as the Biot-

Savart kernel or the 2D-incompressible Navier-Stokes model in fluid dynamics).
The case of *gradient systems* is an important sub-case when $\sigma(x,\mu) = \sigma\mathbf{Id}$ for a constant $\sigma > 0$ and

$$b(x,\mu) = -\nabla\mathbf{V}(x) - \int_{\mathbb{R}^d} \nabla W(x-y)\mu(dy) \tag{1.1.15}$$

where $\mathbf{V}, W$ are two potentials on $\mathbb{R}^d$ respectively called the *confinement potential* and the *interaction potential*. The limit Fokker-Planck equation

$$\partial_t \mu_t = \frac{\sigma^2}{2}\Delta\mu_t + \nabla\cdot\Big(\mu_t \nabla(\mathbf{V} + W\star\mu_t)\Big), \tag{1.1.16}$$

is called the *granular-media* equation. The above models are two-body interactions. This is characterized by the fact that K_1 or K_2 depend on only two variables or equivalently by the fact the functional $F : \mu \mapsto \int K_1(x,y)\mu(dx)\mu(dy)$ is a polynomial of degree two. Nevertheless, in some other models, one may find some interactions which involve more than two particles. This is for instance the case of the Skyrme model (see [212]). This is why in this paper, we choose to consider a polynomial dependence in the measure μ induced by *order statistics (many-body interaction)* in order to generalize the results obtained in the case of a linear interaction in the measure μ defined by the convolution via a potential two-body interaction ([132], [133]). More exactly, under adequate assumptions (see $\mathbb{HMV}$ assumptions 1.3.1, $\mathbb{VFP}$ assumptions 1.3.2), we are interested in the exponential return to equilibrium of the solution of eq. (1.1.14) in the case

$$F(\mu) = \sum_{k=2}^{N} \int W^{(k)} d\mu^{\otimes k}, \tag{1.1.17}$$

where $\forall k \in \{2,\dots,N\}$, $W^{(k)}$ is a *symmetric interaction potential* between k particles and N represents the number of such potentials. The intrinsic derivative $\mathscr{D}_m F(\nu, y)$ associated with this functional is given by

$$\nabla\frac{\delta F}{\delta m}(\nu,y) = \sum_{k=2}^{N}\sum_{j=1}^{k}\int \nabla_{x_j} W^{(k)}(x_1,\dots,x_{j-1},y,x_{j+1},\dots,x_k)\nu^{\otimes k-1}(dx_1,\dots,dx_{j-1},dx_{j+1},\dots,dx_k) \tag{1.1.18}$$

The associated microscopic (particle-level) interaction is given by (U−statistic of order k and kernel $\Phi \equiv W^{(k)}$)

$$U_n(W^{(k)}) := \frac{k!(n-k)!}{n!}\sum_{1\le i_1<\dots<i_k\le n} W^{(k)}(X^{i_1,n},\dots,X^{i_k,n}), \quad \text{where } X^n := (X^{1,n},\dots,X^{n,n}) \in (\mathbb{R}^D)^n. \tag{1.1.19}$$

$\mathrm{U}(\mathrm{X}^n) := \mathrm{U}_n(\Phi)$ is called U−statistic of order k and kernel Φ associated with the sample X^n. This statistic corresponds to the arithmetic mean of the kernel Φ over all the parts at k elements of the set of sample values. We will often write $\mathrm{U}_n(\mathrm{W}^{(n)})(\mathrm{X}^n) =: \mathrm{U}(\mathrm{X}^n)$. We generalize this definition to the space of probabilities by the functional

$$\mu \in \mathscr{P}(\mathbb{R}^{\mathrm{D}}) \longmapsto \int_{\mathbb{R}^{k\mathrm{D}}} \Phi d\mu^{\otimes k}, \tag{1.1.20}$$

called monome of degree k and coefficient Φ on the probability space $\mathscr{P}(\mathbb{R}^{\mathrm{D}})$. The link between these two microscopic and macroscopic interactions is given by

$$\sum_{k=2}^{\mathrm{N}} \mathrm{U}_n(\mathrm{W}^{(k)}) = \mathrm{F}(\mu_{\mathrm{X}^n}). \tag{1.1.21}$$

Remark 1.1.2. Note that the *granular-media equation* (1.1.16) is a particular case of eq. (1.1.14) with

$$\mathrm{F}(\mu) = \int \mathrm{W}^{(2)}(x,y)\mu(dx)\mu(dy) \quad \mathrm{W}^{(2)}(x,y) = \frac{1}{2}\mathrm{W}(x-y) \quad \mathrm{V}(x) \equiv \frac{2}{\sigma^2}\mathbf{V}(x). \tag{1.1.22}$$

Indeed, in this case, we have

$$\frac{\delta \mathrm{F}}{\delta m}(\mu,x) = \int \mathrm{W}^{(2)}(x,y)\mu(dy) + \int \mathrm{W}^{(2)}(y,x)\mu(dy) = \int \mathrm{W}(x-y)\mu(dy) =: \mathrm{W}\star\mu(x), \tag{1.1.23}$$

so that

$$\mathscr{D}_m \mathrm{F}(\mu,x) = \nabla \frac{\delta \mathrm{F}}{\delta m}(\mu,x) = \nabla \mathrm{W}\star\mu(x).$$

Energy and Large Deviations. Consider $\mathscr{M}_1^p(\mathbb{R}^{\mathrm{D}})$ defined in section 1.2, $\mathrm{G} : \mathscr{M}_1^p(\mathbb{R}^{\mathrm{D}}) \to \overline{\mathbb{R}}$ (which can be nonlinear) and the probability (Gibbs) measure α related to V, *i.e.* $\alpha(dx) = \mathrm{Z}_{\mathrm{V}}^{-1} e^{-\mathrm{V}(x)}dx$ with $\mathrm{Z}_{\mathrm{V}} = \int e^{-\mathrm{V}(x)}dx$ (where Z_{V} is assumed to finite). For any $\sigma > 0$, we put

$$\mathrm{V}^{\sigma,\mathrm{G}}(m) := \mathrm{G}(m) + \frac{\sigma^2}{2}\mathbf{H}[m|\alpha]. \tag{1.1.24}$$

$\mathrm{V}^{\sigma,\mathrm{G}}$ is an energy function regularised by the **KL**−divergence $\mathbf{H}[m|\alpha]$ which is given by eq. (1.2.4) in section 1.2. It is known (see e.g. [139, Proposition.2.5]) that $\mathrm{V}^{\sigma,\mathrm{G}}$ is minimized by a measure $m^{\sigma,\star}$ satisfying the following fixed point problem (it is noteworthy that the variational form of the invariant measure of the classic Langevin equation is a particular example of this

first order condition)

$$m^{\sigma,\star}(dx) = \frac{1}{Z_\sigma} e^{-\frac{2}{\sigma^2}\left(\frac{\delta G}{\delta m}(m^{\sigma,\star},x)+\frac{\sigma^2}{2}V(x)\right)} dx, \tag{1.1.25}$$

where Z_σ is the normalising constant, and for any $m \in \mathscr{M}_1^p(\mathbb{R}^D)$ and $x \in \mathbb{R}^D$, $\frac{\delta G}{\delta m}(m,x)$ denotes a flat derivative of G with respect to m, in the direction of x, evaluated at m. For any $\Theta_0, \Theta_1 \in \mathscr{M}_1^p(\mathbb{R}^D)$, the function $\frac{\delta G}{\delta m} : \mathscr{M}_1^p(\mathbb{R}^D) \times \mathbb{R}^D \to \mathbb{R}$ satisfies

$$G(\Theta_1) - G(\Theta_0) = \int_0^1 \int_{\mathbb{R}^D} \frac{\delta G}{\delta m}(\Theta_0 + \lambda(\Theta_1 - \Theta_0), x)(\Theta_1 - \Theta_0)(dx)\lambda_{\mathbb{R}}(d\lambda). \tag{1.1.26}$$

This notion of derivative appears in the literature under several different names, including the linear functional derivative (see e.g [51, Section.5.4.1]) or the first variation [5].

Remark 1.1.3. It is important to note that we have uniqueness modulo the choice of a function of the measure. The McKean-Vlasov SDE given by eq. (1.1.2) (therefore the associated PDE given by eq. (1.1.14) and the invariant measure given by eq. (1.1.25)) does not depend on the choice of the function of the measure in the calculation of a flat derivative: we can therefore do the calculations with any flat derivative. By convexity of $\mathscr{M}_1^p(\mathbb{R}^d)$, for all $t \in [0,1]$, $\Theta_t := (1-t)\Theta_0 + t\Theta_1 \in [\Theta_0, \Theta_1] \subset \mathscr{M}_1^p(\mathbb{R}^d)$. eq. (1.1.26) is equivalent to deriving the functional G along the end segment Θ_0 and Θ_1 parameterized by the path $t \in [0,1] \mapsto \Theta_t$:

$$\frac{d}{dt} G(\Theta_t) = \int_{\mathbb{R}^D} \frac{\delta G}{\delta m}(\Theta_t, x)\partial_t \Theta_t(dx). \tag{1.1.27}$$

In practice, this reformulation via derivation along paths lends itself better to calculations. For example, it is easy to check

$$\nabla \frac{\delta F}{\delta m}(\nu, y) = \sum_{k=2}^{N} k \int \nabla_{x_1} W^{(k)}(y,z)\nu^{\otimes k-1}(dz); \tag{1.1.28}$$

$$\mathscr{D}_m \mathbf{H}[\cdot|\alpha](\nu, y) = \nabla \log\left(\frac{d\nu}{d\alpha}\right)(y). \tag{1.1.29}$$

In [168], the authors showed a *Sanov theorem* by *large deviations* on U-statistics in which $V^{\sqrt{2},F}$ is a *good convergence rate function*, which implies that $V^{\sqrt{2},F}$ admits a *minimizer*: inf-compactness of the mean-field energy functional. As for *uniqueness*, it is not assured: in *statistical physics*, we speak of *phase transition*. We ensure uniqueness by a *contraction property* (see $\mathbb{HMV}$ assumptions 1.3.1) on the *invariant measure application* induced by eq. (1.1.25). Large Deviation Principles imply propagation of chaos, but they do not always give a way to quantify it since the related results are often purely asymptotic (for instance, Sanov the-

orem is non-quantitative). Nevertheless, the results of large deviations turn out to be very useful for the technical passages in the macroscopic limits: when one makes tend the number of particles to infinity. In the seminal article [19], the authors improve results from [154] and [42] on Large Deviation Principles (LDP) for Gibbs measures and obtain as a byproduct a pathwise propagation of chaos result for the McKean-Vlasov diffusion. Firstly, [19, Theorem.A] (or theorem 1.7.3) states a large deviation principle for Gibbs measures with a polynomial potential. [19, Theorem.B] quantifies the fluctuations of μ_{X^n} in the non-degenerate case. Analogous results for the degenerate case are given in [19, Theorem.C]. For more details, see also [62, Theorem.4.7, Corollary.3]. We use the large deviations results obtained on the order statistics in [168]: In addition to the fact that the mean-field entropy functional (eq. (1.7.61) or $V^{\sqrt{2},F}$ defined by eq. (1.1.24)) is a rate function (theorem 1.7.3) for the random empirical measure μ_{X^n}, the authors show that it is a good rate function that has good tensorization properties.

Long time behavior. In the present paper, we are concerned by the long-time convergence towards the solution to an optimization problem on the subspace $\mathscr{M}_1^p(\mathbb{R}^D)$ of probability measures $\mathscr{M}_1(\mathbb{R}^D)$: we consider a function $\mathbf{E}:\mathscr{M}_1^p(\mathbb{R}^D)\to\overline{\mathbb{R}}$ and we want to find a minimizing measure $m^\star := \mathbf{arginf}_{\mathscr{M}_1^p(\mathbb{R}^D)}\mathbf{E}$ such that for a *gradient flow* (see e.g. [5] and [194]) $(m_t)_{t\geq 0}$ associated with $\mathbf{E}$, we have an exponential estimate of the deviation $\mathbf{E}(m_t)-\mathbf{E}(m^\star)$ of the form (with $C\geq 1$ and $\rho>0$)

$$\mathbf{E}(m_t)-\mathbf{E}(m^\star)\leq C(\mathbf{E}(m_0)-\mathbf{E}(m^\star))e^{-\rho t}. \tag{1.1.30}$$

eq. (1.1.30)-type Inequalities are called *hypocoercive inequalities*. We call $\mathbf{E}-\mathbf{E}(m^\star)$ the *entropy functional* ([147],[15],[102]) of the system and $-\frac{d}{dt}(\mathbf{E}(m_t)-\mathbf{E}(m^\star))$ the *production of entropy* (usually called energy in mathematical literature). Clausius invents the concept of entropy, Boltzmann proposes to derive entropy along the flow. Generally speaking, an entropy is a *Lyapunov functional* of a specific form. It is however hard (and even somewhat artificial) to give a formal narrow definition of entropies that distinguishes them from, say, energies. An entropy is a quantity calculated from a solution, which decreases over time when the solution obeys an evolution equation, and which is stationary only for the stationary solutions of the equation. In conclusion, the concept of entropy is a tool that adapts to what we want to study. The notion of *hypocoercivity* was proposed by T. Gallay. The objective is typically to control the entropy at time t by the initial entropy multiplied by a constant C

(always greater than 1) and a exponential decay factor, with exponential decay rate as good as possible in big time. This theory is inspired by the *hypoelliptic theory of L. Hormander*, and the terminology hypocoercivity accounts for the relationship between entropy and its derivative with respect to t. There would be *coercivity* if C = 1, which is clearly not possible in most cases considered in kinetic theory. It is well known that, for the standard *Langevin equation* of Hamiltonian V (given by eq. (1.1.2) in the case F ≡ 0), for $\rho > 0$, the following assertions are equivalent:

$$\forall \varphi \in \mathscr{C}_c^\infty(\mathbb{R}^d), \quad \rho \mathbf{Ent}_\alpha[\varphi^2] \le 2 \int ||\nabla \varphi||^2 d\alpha. \tag{1.1.31}$$

$$\rho \mathbf{H}[\cdot|\alpha] \le 2\mathbf{I}[\cdot|\alpha]. \tag{1.1.32}$$

$$\forall t \ge 0, \quad \mathbf{H}[\mu_t^V|\alpha] \le \mathbf{H}[\mu_0^V|\alpha] e^{-\rho t}. \tag{1.1.33}$$

These three equivalent assertions imply the T2−*Talagrand inequality*

$$\rho \mathscr{W}_2^2(\cdot, \alpha) \le 2\mathbf{H}[\cdot|\alpha], \tag{1.1.34}$$

inequality which, in turn, implies an *exponential contraction in wasserstein metric* $\mathscr{W}_2$, i.e. the exponential convergence of the flow $(\mu_t^V)_{t\ge 0}$ (solution of the Fokker-Planck equation associated with the standard Langevin process of Hamiltonian V) to the maxwellian (invariant measure of the Langevin process that can also be seen from equivalently as the unique $\mathbf{argmin}_{\mathscr{P}(\mathbb{R}^d)}\mathbf{H}[\cdot|\alpha])$ α of the *Fokker-Planck PDE* given by eq. (1.1.14) in the case F ≡ 0:

$$\forall t \ge 0, \quad \mathscr{W}_2^2(\mu_t^V, \alpha) \le \frac{2}{\rho} \mathbf{H}[\mu_0^V|\alpha] e^{-\rho t}. \tag{1.1.35}$$

eq. (1.1.31) and eq. (1.1.32) respectively define the *logarithmic Sobolev inequality* ([15]) and its *dual version*. According to the *dimension curvature criterion of Bakry-Emery*, we have

$$\left(\exists \rho > 0 \quad \forall (x,h) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \langle \nabla^2 V(x) h, h\rangle \ge \rho ||h||_2^2\right) \Longrightarrow eq.\ (1.1.31). \tag{1.1.36}$$

Note that in the case of the symmetric Langevin-Kolmogorov process, we have

$$m_t = \mu_t^V, \quad m^\star = \alpha, \quad \mathbf{E} = \mathbf{H}[\cdot|\alpha], \quad \mathbf{E}(m_t) - \mathbf{E}(m^\star) = \mathbf{H}[\mu_t^V|\alpha], \tag{1.1.37}$$

$$-\frac{d}{dt}(\mathbf{E}(m_t) - \mathbf{E}(m^\star)) = -\frac{d}{dt}\mathbf{H}[\mu_t^V|\alpha] = \mathbf{I}[\mu_t^V|\alpha]. \tag{1.1.38}$$

The objective of this work is to identify a flow of measures $(m_t^{\sigma,F})_{t\geq 0}$ (flow solution of eq. (1.1.14)) such that

$$V^{\sigma,F}(m_t^{\sigma,F}) - V^{\sigma,F}(m^{\sigma,\star}) \overset{t\to+\infty}{\longrightarrow} 0, \tag{1.1.39}$$

as well as conditions ($\mathbb{HMV}$ assumptions 1.3.1, $\mathbb{VFP}$ assumptions 1.3.2) that ensure that this convergence is exponential. To this end, we equip the space $\mathscr{M}_1^p(\mathbb{R}^D)$ with a suitable distance function $\mathbf{d} : \mathscr{M}_1^p(\mathbb{R}^D) \times \mathscr{M}_1^p(\mathbb{R}^D) \to \mathbb{R}_+$ and consider a corresponding gradient flow, where the form of the flow is dictated by the choice of $\mathbf{d}$. Such a problem has been dealt with in the case of the Fisher-Rao metric (see [167]): the authors established from a Polyak-Lojasiewicz inequality the exponential convergence of the gradient flow $(m_t^{\sigma,G})_{t\geq 0}$ described by the birth-death equation along $V^{\sigma,G}$ towards $V^{\sigma,G}(m^{\sigma,\star})$. In our case, eq. (1.1.30) implies the exponential decay in $\mathbf{d}$-metric (transport distance):

$$\mathbf{d}(m_t^{\sigma,F}, m^{\sigma,\star}) \leq \gamma(V^{\sigma,F}(m_0^{\sigma,F}) - V^{\sigma,F}(m^{\sigma,\star}))e^{-\rho t}. \tag{1.1.40}$$

eq. (1.1.40) is a consequence of transport inequalities (see [226]). Moreover, given a measure $m^{\sigma,\star}$ satisfying the first order condition eq. (1.1.25), it is formally a stationary solution to eq. (1.1.14) called the *Maxwellian of the McKean-Vlasov PDE*. Therefore, formally, we have already obtained the correspondence between the *minimiser of the free energy function* and the *invariant measure* of eq. (1.1.2). In this paper, the connection is rigorously proved mainly with a probabilistic argument. The study of stationary solutions to nonlocal, diffusive eq. (1.1.14) is classical topic with it roots in statistical physics literature and with strong links to Kac's program in Kinetic theory [187]. We also refer reader to the excellent monographs [5] and [11]. An important issue is the *long-time behaviour of gradient systems* which is often studied under *convexity assumptions* on the potentials. In particular, variational approach has been developed in [52] and [194] where authors studied *dissipation of entropy* for granular media equations eq. (1.1.16) with the symmetric interaction potential of convolution type (interaction potential corresponds to term $\mathscr{D}_m F$ in eq. (1.1.14)). Following on from the work done in [194] and [52] (among others) on the long-time behavior of eq. (1.1.16), in [133], the authors proved via a *uniform logarithmic Sobolev inequality* in the number of particles that

$$\forall\, t \geq 0, \quad H_W[\nu_t] \leq H_W[\nu_0]e^{-\rho_{LS}\frac{t}{2}} \quad \text{and} \quad \mathscr{W}_2^2(\nu_t, \nu_\infty) \leq \frac{2}{\rho_{LS}} H_W[\nu_0]e^{-\rho_{LS}\frac{t}{2}}. \tag{1.1.41}$$

eq. (1.1.41) translates the exponential decrease of the *mean field entropy* H_W (given by eq. (1.1.30) with $\mathbf{E} = V^{\sigma,F}$) and the *contraction* in Wasserstein metric ($\mathbf{d} = \mathscr{W}_2$) of the solution flow of

eq. (1.1.14) in the case

$$\sigma = \sqrt{2} \quad \text{and} \quad F(\mu) = \frac{1}{2}\int W(x,y)\mu(dx)\mu(dy). \tag{1.1.42}$$

The study of the long-time behaviour for the VFP equation is often more difficult than that of the McKean-Vlasov equation because of two reasons:

1. it is a degenerate diffusion process where the Laplacian acts only on the volocity variable and;
2. it is not a gradient flows but simultaneously presents both Hamiltonian and gradient flows effects.

In [127], combining the results of [133] and [188], the trend to equilibrium in large time is studied for a large particle system (given by eq. (1.3.18) in case of a two-body interaction) associated to a *Vlasov-Fokker-Planck equation* by the authors: they showed that under some conditions (that allow *non-convex confining potentials*), the convergence rate is proven to be independent from the number of particles. From this are derived *uniform in time propagation of chaos estimates* and an *exponentially fast convergence for the nonlinear equation itself*.

Contributions. In this paper, we are going to prove in a polynomial interaction setting

1. propagation of chaos in Wasserstein's $\mathscr{W}_2$-metric for our particle systems given by eq. (1.3.7) and eq. (1.3.18).
2. entropic convergence to equilibrium for the nonlinear McKean-Vlasov SDE (mean field limit of the first order system given by eq. (1.3.7)) generalizing results (given in eq. (1.1.41) of [133]).
3. by *Villani's hypocoercivity theorem* (see e.g. [132, Theorem.3] or [227, Theorem.35]) the H^1-convergence for the kinetic Fokker-Planck equation with mean field interaction given by eq. (1.3.18).
4. exponential convergence towards equilibrium in metric $\mathscr{W}_2-$Wasserstein for the flow solution of the Vlasov-Fokker-Planck equation: mean field limit of the second order system given by eq. (1.3.18).

In the spirit of [133], we give *probabilistic proofs* (see section 1.5, fig. 1.1 and section 1.6) based mainly on the *propagation of chaos* (see theorem 1.5.1); the **H**-tensorization (see proposition 1.5.10) and inf-compactness given by *large deviations principle*; the **I**-tensorization (see proposition 1.5.11) given by *law of large numbers*; the *uniform log-Sobolev inequality*

(see theorem 1.5.3); and the *uniqueness argument* (see proposition 1.5.9). In the kinetic case, we need additional results such as Villani's hypocoercivity ([132, Theorem.3] or [227, Theorem.18 and Theorem.35]) theorem (see proposition 1.5.15) and *Hormander's form* (see e.g. respectively Theorem.7 and Theorem.10 in [188, [189]]). The fact that the interaction is polynomial is important in calculations, among other things, for passing to the limit in the number of particles: technical passage to the limit given, among others, by LDP.

Plan of the paper. Let us finish this introduction by the plan of the paper. In the next three sections, we will present our mean field systems (eq. (1.3.7),eq. (1.3.18)), our set of assumptions ($\mathbb{HMV}$ assumptions 1.3.1 and $\mathbb{VFP}$ assumptions 1.3.2), the main results (and examples) (in section 1.4) of the paper concerning logarithmic Sobolev inequality of mean field particles systems as well as exponential convergences to equilibrium for McKean-Vlasov (theorem 1.4.1,theorem 1.4.2), kinetic Fokker-Planck (theorem 1.4.3) and Vlasov-Fokker-Planck (theorem 1.4.4) SDEs. In section 1.5, we sketch a proof of our results and we introduce the pre-proof tools. In section 1.6, we prove our main results. And we end the paper with the appendix, the acknowledgments and the bibliographical references.

1.2 Notations and Definitions

We try to keep coherent definitions and notations throughout the article, but as the various objects and what they represent may become confusing, we list them here for reference :

Notations. For all $(u,v) \in \mathbb{R}^d \times \mathbb{R}^d$, we note $u \otimes v := uv^{\mathrm{T}} = (u_i v_j)_{1\le i,j\le d}$ the tensor product matrix of two vectors and $u \cdot v := u^{\mathrm{T}} v$ the standard Euclidean scalar product of two vectors. We note $||\cdot||_{\mathbf{op}}$ the matrix subordinate norm to the Euclidean norm which we will note indifferently $||\cdot||_2$ or $|\cdot|$. $\langle \cdot,\cdot \rangle$ represents indifferently the scalar product and the duality bracket. We note $|||\cdot|||_{\mathrm{H}^1\to\mathrm{H}^1}$ the operator norm associated with the weighted Sobolev $\mathrm{H}^1(\mu_Z^n)$ space induced by the invariant measure μ_Z^n of our second-order system given by eq. (1.3.18). We have

$$\mathrm{H}^1(\mu_Z^n) := \left\{ \varphi \in \mathrm{L}^2(\mu_Z^n), \quad \nabla\varphi \in (\mathrm{L}^2(\mu_Z^n))^n \right\}; \tag{1.2.1}$$

$$||\varphi||^2_{\mathrm{H}^1} := ||\varphi||^2_{\mathrm{L}^2(\mu_Z^n)} + \int \left(||\nabla_x \varphi||_2^2 + ||\nabla_v \varphi||_2^2 \right) d\mu_Z^n. \tag{1.2.2}$$

$(\mathrm{B}_t)_{t\ge 0}$ represents the standard Brownian motion. We consider $((\mathrm{B}^i)_{t\ge 0})_{i\in\{1,\dots,n\}}$ n indepen-

dent copies of $(B_t)_{t\geq 0}$. For all $n \geq 1$, $\mathfrak{S}_n$ is the n-th symmetric group. For all $p \in [1+\infty)$, the Wasserstein p-distance between two probability measures μ and ν on $\mathbb{R}^D$ with finite p-moments is given by

$$\mathscr{W}_p(\mu,\nu) := \Big(\inf_{\gamma\in\Gamma(\mu,\nu)}\int_{\mathbb{R}^D\times\mathbb{R}^D}|x-y|^p\gamma(dxdy)\Big)^{\frac{1}{p}}, \tag{1.2.3}$$
$$\Gamma(\mu,\nu) := \Big\{\gamma\in\mathscr{P}(\mathbb{R}^D\times\mathbb{R}^D),\quad \pi_1\gamma=\mu \quad\text{and}\quad \pi_2\gamma=\nu\Big\}.$$

We note $\mathscr{M}_1^p(\mathbb{R}^D)$ the space of probability measures with finite $p-$moments.

Definitions.

Relative entropy: Let $\mu\in\mathscr{P}(\mathbb{R}^D)$. We define $\mathbf{H}[\cdot|\mu]:\mathscr{P}(\mathbb{R}^D)\longrightarrow[0,+\infty]$ such that

$$\mathbf{H}[\nu|\mu] = \begin{cases} \mathbb{E}_\nu[\log\frac{d\nu}{d\mu}] =: \mathbf{Ent}_\mu[\frac{d\nu}{d\mu}] & \text{if } \nu\ll\mu,\\ +\infty & \text{otherwise.}\end{cases} \tag{1.2.4}$$

And we recall that in the first case of absolute continuity, $\frac{d\nu}{d\mu}$ is the *Radon-Nikodym density* of ν with respect to μ.

Relative Fisher information: We also define the *Fisher-Donsker-Varadhan information* of ν with respect to μ by:

$$\mathbf{I}[\nu|\mu] = \int\Big|\Big|\nabla\sqrt{\frac{d\nu}{d\mu}}\Big|\Big|^2 d\mu = \frac{1}{4}\int\Big|\Big|\nabla\log\frac{d\nu}{d\mu}\Big|\Big|^2 d\nu = \frac{1}{4}\int\Big|\Big|\nabla\frac{\delta\mathbf{H}[\cdot|\mu]}{\delta m}(\nu,y)\Big|\Big|^2\nu(dy) \tag{1.2.5}$$

if $\nu\ll\mu$ and $\sqrt{\frac{d\nu}{d\mu}}\in\mathbf{H}^1_\mu$, and $\mathbf{I}[\nu|\mu]=+\infty$ otherwise. $\mathbf{H}^1_\mu$ is the domain of the Dirichlet form

$$\mathscr{E}_\mu : g\longmapsto\int||\nabla g||^2 d\mu. \tag{1.2.6}$$

UPI. We say that $\mu(dx) := \frac{1}{Z}e^{-H(x)}dx$ (Gibbs probability measure of hamiltonian $H:\mathbb{R}^{nD}\to\mathbb{R}$) satisfies a uniform Poincaré inequality if

$$\exists\lambda>0\quad\forall n\geq 2\quad\forall\varphi\in\mathscr{C}_c^\infty(\mathbb{R}^{nD}),\quad \lambda\mathbb{V}_\mu[\varphi]\leq\mathbb{E}_\mu[||\nabla\varphi||^2]. \tag{1.2.7}$$

And we call Poincaré constant the best constant $\lambda_1(\mu)$ for which we have such an inequality.

ULSI. We say that μ satisfies a uniform logarithmic Sobolev inequality if

$$\exists \rho > 0 \quad \forall n \geq 2 \quad \forall \varphi \in \mathscr{C}_c^\infty(\mathbb{R}^{nD}), \quad \rho \mathbf{Ent}_\mu[\varphi^2] \leq \mathbb{E}_\mu[||\nabla \varphi||^2]. \tag{1.2.8}$$

And the best constant $\rho_{LS}(\mu)$ for which such an inequality holds is called the logarithmic Sobolev constant.

Remark 1.2.1. We recall that

$$\textbf{ULSI.} \Longrightarrow \textbf{UPI.} \tag{1.2.9}$$

The Poincaré and log-Sobolev inequalities for μ are equivalent to exponential decreases of the semigroup $(P_t)_{t\geq 0}$ respectively in variance and in entropy, i.e.

— **Poincaré**

$$\forall \quad f \in L^2(\mu) \quad t \geq 0, \quad ||P_t f - \langle \mu, f \rangle||_{L^2(\mu)} \leq e^{-\lambda_1(\mu) t} ||f - \langle \mu, f \rangle||_{L^2(\mu)}. \tag{1.2.10}$$

— **Log-Sobolev**

$$\forall \quad f \in L^1(\mu) \log L^1(\mu) \quad t \geq 0, \quad \mathbf{Ent}_\mu[P_t f] \leq e^{-\rho_{LS}(\mu) t} \mathbf{Ent}_\mu[f]. \tag{1.2.11}$$

Here, the notation $L^1(\mu) \log L^1(\mu)$ denotes the entropy definition domain under μ.

We say that μ satisfies a T_p-transport (Talagrand) inequality if there exists $\alpha > 0$ such that $\mathscr{W}_p(\cdot, \mu) \leq \sqrt{\alpha \mathbf{H}[\cdot|\mu]}$.

Remark 1.2.2. Moreover, as with the Poincaré and log-Sobolev inequalities, T_2-inequality implies the T1-inequality : by definition and Cauchy-Schwarz inequality, we have

$$\mathscr{W}_1(\mu, \nu) := \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}[||X - Y||] \leq \inf_{X\sim\mu, Y\sim\nu} \sqrt{\mathbb{E}[||X - Y||^2]} =: \mathscr{W}_2(\mu, \nu). \tag{1.2.12}$$

The class of probabilities verifying T_1-inequality is identical to that having an exponential moment of finite order 2. The T_2-inequality is significantly more structured than the T_1-inequality since it involves a spectral gap inequality.

1.3 Mean-Field Systems and Assumptions

Throughout the paper, we consider a *confinement potential of a particle* $V : \mathbb{R}^d \longrightarrow \mathbb{R} \in \mathscr{C}^2(\mathbb{R}^d)$ and N *interaction potentials* such that

$$\forall k \in \{2,\dots,N\}, \quad W^{(k)} : (\mathbb{R}^d)^k \longrightarrow \mathbb{R} \in \mathscr{C}^2((\mathbb{R}^d)^k). \tag{1.3.1}$$

We recall that $\forall \sigma \in \mathfrak{S}_k$ and $\forall x = (x_1,\dots,x_k)$,

$$W^{(k)}(\sigma \cdot x) = W^{(k)}(x); \tag{1.3.2}$$

$$\alpha(dx) := \frac{1}{C} e^{-V(x)} dx; \tag{1.3.3}$$

$$U_n(W^{(k)}) := \frac{1}{|I_n^k|} \sum_{(i_1,\dots,i_k)\in I_n^k} W^{(k)}(x_{i_1},\dots,x_{i_k}), \tag{1.3.4}$$

where $I_n^k := \{(i_1,\dots,i_k) \in \mathbb{N}^k | i_p \neq i_q, \quad 1 \leq i_p \leq n\}$ is the set of possible arrangements of k integers of the set of n first nonzero integers, which gives $|I_n^k| = A_n^k := \frac{n!}{(n-k)!}$. We define $W^{(k),-} := \max(-W^{(k)}, 0)$ and $W^{(k),+} := \max(W^{(k)}, 0)$ the negative and positive parts of $W^{(k)}$. $\forall \mu$ such that $W^{(k),-} \in L^1(\mu^{\otimes k})$,

$$\mathbf{W}^{(k)}[\mu] := \mathbb{E}_{\mu^{\otimes k}}[W^{(k)}] = \mathbb{E}_{\mu^{\otimes k}}[W^{(k),+}] - \mathbb{E}_{\mu^{\otimes k}}[W^{(k),-}]. \tag{1.3.5}$$

1.3.1 Our Systems

First order case. We consider the *microscopic mean-field many-body interaction energy* given by

$$H_n(x_1,\dots,x_n) := \sum_{j=1}^{n} V(x_j) + n \sum_{k=2}^{N} U_n(W^{(k)}). \tag{1.3.6}$$

The (non-kinetic) McKean-Vlasov process is defined as the mean field limit (under adequate assumptions given below) of the sequence $(X^n)_{n\geq N}$ of Langevin-Kolmogorov process of Hamiltonian H_n, i.e.: (N fixed)

$$\forall n \geq N, \quad dX_t^n = \sqrt{2} dB_t - \nabla H_n(X_t^n) dt. \tag{1.3.7}$$

Let

$$\mathscr{L}_n := \Delta - \nabla H_n \cdot \nabla \tag{1.3.8}$$

be the *infinitesimal generator* and $(P_t^n)_{t\geq 0}$ the associated *semigroup* of unique invariant measure (under $\mathbb{HMV}$ assumptions 1.3.1 below), the *Gibbs measure*

$$\mu_n(dx) := \frac{1}{Z_n} e^{-H_n(x)} dx \quad \text{with} \quad Z_n := \int_{(\mathbb{R}^d)^n} e^{-H_n(x)} dx < +\infty \tag{1.3.9}$$

is the normalization constant (called *partition function*). Note that

$$\mu_n(dx) = \frac{C^n}{Z_n} e^{-n\sum_{k=2}^N U_n(W^{(k)})} \alpha^{\otimes n}(dx). \tag{1.3.10}$$

Without interaction (i.e. $\forall k$, $W^{(k)} \equiv 0$ or constant), $\mu_n = \alpha^{\otimes n}$ (i.e. the particles are independent). We denote

$$L_n(x;\cdot) := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}(\cdot) \tag{1.3.11}$$

the empirical measurement application. We know that under general conditions, by propagation of chaos ([217]), $L_n(X^n;\cdot)$ converges weakly towards the solution of the nonlinear partial differential equation of McKean-Vlasov associated with the system of particles. We define

$$\mu_n^*(dx) := e^{-n\sum_{k=2}^N U_n(W^{(k)})} \alpha^{\otimes n}(dx) = \frac{Z_n}{C^n}\mu_n(dx). \tag{1.3.12}$$

The *macroscopic mean-field energy* is given by

$$\mathbf{E}_W[\mu] := \begin{cases} \mathbf{H}[\mu|\alpha] + \sum_{k=2}^N \mathbf{W}^{(k)}[\mu] & \text{if } \mathbf{H}[\mu|\alpha] < +\infty \text{ and } W^{(k),-} \in L^1(\mu^{\otimes k}), \\ +\infty & \text{otherwise.} \end{cases} \tag{1.3.13}$$

Let

$$\mathbf{dom}(\mathbf{H}_W) := \left\{\mu, \quad \mathbf{H}[\mu|\alpha] < +\infty \quad \text{and} \quad \forall k, \quad W^{(k),-} \in L^1(\mu^{\otimes k})\right\}. \tag{1.3.14}$$

Remark 1.3.1. $\mathbf{H}_W := \mathbf{E}_W - \inf \mathbf{E}_W$ is called the *mean field entropy*. We can prove that $\mathbf{H}_W$ is *inf-compact* (theorem 1.5.2) and that there is at least one minimizer usually called *equilibrium point*. From the point of view of statistical physics, $\mathbf{H}_W$ is an entropy or *free energy* associated to the *nonlinear McKean-Vlasov equation* given by eq. (1.3.7). The uniqueness of the minimizer means that there is no phase transition for the mean-field. Concerning the work on uniqueness in the case of peer interaction, we can cite among others: [133], [173] and [52]. These authors ([173],[52]) showed that $\mathbf{H}_W$ is *strictly displacement convex* (i.e. along the $\mathcal{W}_2$-geodesic) under various sufficient conditions on the convexity of the confinement potential V and the pair interaction potential $W^{(2)}$. In case of a many-body interaction, under

assumptions in $\mathbb{HMV}$ assumptions 1.3.1, we prove in proposition 1.5.9 the uniqueness: then we denote μ_∞ this minimizer.

Analogously, we define the *mean-field Fisher information* by:

$$\mathbf{I}_{\mathrm{W}}[\mu] := \frac{1}{4}\int \left|\left|\nabla \frac{\delta \mathbf{E}_{\mathrm{W}}}{\delta m}(\mu, y)\right|\right|^2 \mu(dy). \tag{1.3.15}$$

Remark 1.3.2. Without interaction ($\forall k$, $\mathrm{W}^{(k)} \equiv w_k$), we find the Lyapunov functionals associated with the standard symmetric Langevin-Kolmogorov process whose Hamiltonian is given by the confinement potential V. More precisely, in this case:

$$\mathbf{E}_{\mathrm{W}} = \mathbf{H}[\cdot|\alpha] + \sum_{k=2}^{\mathrm{N}} w_k, \quad \mathbf{H}_{\mathrm{W}} = \mathbf{H}[\cdot|\alpha] \quad \text{and} \quad \mathbf{I}_{\mathrm{W}} = \mathbf{I}[\cdot|\alpha]. \tag{1.3.16}$$

Kinetic case. Set

$$z := (x_1, \dots, x_n, v_1, \dots, v_n) \in \mathbb{R}^{2nd}, \quad \mathrm{H}_n^{\mathrm{Z}}(z) = \frac{1}{2}\sum_{j=1}^{n} |v_j|^2 + 2\mathrm{V}(x_j) + n\sum_{k=2}^{\mathrm{N}} \mathrm{U}_n(\mathrm{W}^{(k)}) \tag{1.3.17}$$

and $\mathrm{Z}^n := (\mathrm{X}^{n,1}, \dots, \mathrm{X}^{n,n}, \mathrm{V}^{n,1}, \dots, \mathrm{V}^{n,n}) \in (\mathbb{R}^d \times \mathbb{R}^d)^n$ such that

$$\begin{cases} d\mathrm{X}_t^{n,i} = \nabla_{v_i} \mathrm{H}_n^{\mathrm{Z}}(\mathrm{Z}_t^n)\, dt \\ d\mathrm{V}_t^{n,i} = -\Big(\nabla_{x_i} \mathrm{H}_n^{\mathrm{Z}}(\mathrm{Z}_t^n) + \nabla_{v_i} \mathrm{H}_n^{\mathrm{Z}}(\mathrm{Z}_t^n)\Big) dt + \sqrt{2} d\mathrm{B}_t^i. \end{cases} \tag{1.3.18}$$

We are going to study the long-time behavior of the mean-field limit of the Langevin process $(\mathrm{Z}_t^n)_{t\geq 0}$ of Hamiltonian $\mathrm{H}_n^{\mathrm{Z}}(x,v) := \mathrm{S}_{1,n}(x) + \mathrm{S}_{2,n}(v)$ with $\mathrm{S}_{1,n}$ is none other than the Hamiltonian H_n of the McKean-Vlasov case and $\mathrm{S}_{2,n}$ the velocity part ($\mathrm{S}_{2,n} := \mathrm{H}_n^{\mathrm{Z}} - \mathrm{S}_{1,n}$). Invariant measure of the Langevin process is given by

$$\mu_{\mathrm{Z}}^n(dxdv) = \frac{1}{\widetilde{\mathrm{C}}} e^{-\mathrm{H}_n^{\mathrm{Z}}(z)} dxdv = \frac{1}{\mathrm{C}_{1,n}} e^{-\mathrm{S}_{1,n}(x)} dx \frac{1}{\mathrm{C}_{2,n}} e^{-\mathrm{S}_{2,n}(v)} dv = \mu_{1,n} \bigotimes \mu_{2,n}(dxdv). \tag{1.3.19}$$

And the parabolic PDE in the sense of the distributions associated with this Kolmogorov-Fokker-Planck SDE is:

$$\begin{aligned} \partial_t \mu &= \Delta_v \mu + \nabla \mathrm{S}_{2,n} \cdot \nabla_v \mu - \nabla \mathrm{S}_{1,n} \cdot \nabla_v \mu + \nabla \mathrm{S}_{2,n} \cdot \nabla_x \mu \\ &= \Delta_v \mu + v \cdot \nabla_v \mu - \nabla \mathrm{S}_{1,n} \cdot \nabla_v \mu + v \cdot \nabla_x \mu \\ &= \mathscr{L}_{\mathrm{Z},n}^{\dagger} \mu \end{aligned} \tag{1.3.20}$$

with

$$\mathscr{L}_{Z,n} := \Delta_v - v\cdot\nabla_v + \nabla S_{1,n}\cdot\nabla_v - v\cdot\nabla_x \tag{1.3.21}$$

the generator of the strongly continuous semigroup $(P_t^{Z,(n)})_{t\geq 0}$ (if the *hessian* $\nabla^2 S_{1,n}$ is bounded, it is a Markovian semigroup defined by the Kolmogorov-Fokker-Planck SDE) and we note $\mathscr{L}_{Z,n}^\dagger$ adjoint in the sense of distributions. In other words, for any test function $\varphi \in \mathscr{C}_c^\infty((\mathbb{R}^d \times \mathbb{R}^d)^n)$, the function $(t,z) \longmapsto P_t^{Z,(n)}\varphi(z)$ is the unique solution of the Cauchy problem:

$$\begin{cases} \frac{\partial h}{\partial t} = \mathscr{L}_{Z,n} h, \\ h(0,\cdot) = \varphi. \end{cases} \iff \begin{cases} \frac{\partial \mu_t}{\partial t} = \mathscr{L}_{Z,n}^\dagger \mu_t, \\ \mu_0 = \delta_z. \end{cases} \tag{1.3.22}$$

Vlasov Fokker Planck free energy and associated mean field entropy are given by

$$\begin{aligned} \mathscr{E}[\mu] &:= \mathbf{H}[\mu|dxdv] + \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d} ||v||^2 \mu(dxdv) \\ &+ \sum_{k=2}^N \int_{(\mathbb{R}^d\times\mathbb{R}^d)^k} W^{(k)} d\mu^{\otimes k} + \int V(x)\mu(dxdv) \\ &= \mathbf{H}[\mu|\alpha\otimes\mathscr{N}(0,\mathbf{Id}_d)] + \sum_{k=2}^N \int_{(\mathbb{R}^d\times\mathbb{R}^d)^k} W^{(k)} d\mu^{\otimes k} \end{aligned} \tag{1.3.23}$$

and

$$\mathscr{S} := \mathscr{E} - \mathbf{Inf}\mathscr{E} = \mathscr{E} - \mathscr{E}[\mu_\infty^Z]. \tag{1.3.24}$$

They are *Lyapunov functionals for the Vlasov-Fokker-Planck partial differential equation* whose solutions are obtained as mean-field limits of our kinetic Fokker-Planck particle system given by eq. (1.3.18). Mean Field Fisher Information for Vlasov-Fokker-Planck is given by $\left(A := \begin{pmatrix} 0 \\ \mathbf{Id}_d \end{pmatrix} \in \mathscr{M}_{2d,d}(\mathbb{R})\right)$

$$\begin{aligned} \mathscr{I}[\mu] &:= \int \left\langle \nabla_{x,v}\frac{\delta}{\delta m}\mathscr{E}(\mu,x,v), AA^*\nabla_{x,v}\frac{\delta}{\delta m}\mathscr{E}(\mu,x,v)\right\rangle \mu(dxdv) \\ &= \int \left|\left|\nabla_{x,v}\frac{\delta}{\delta m}\mathscr{E}(\mu,x,v)\right|\right|_{AA^*}^2 \mu(dxdv). \end{aligned} \tag{1.3.25}$$

The functional obtained by replacing A by $Z := \begin{pmatrix} z_1\mathbf{Id}_d \\ z_2\mathbf{Id}_d \end{pmatrix} \in \mathscr{M}_{2d,d}(\mathbb{R})$, we will talk about auxil-

iary Fisher information. We have

$$\frac{d}{dt}\mathscr{E}[\mu_t^{\mathbf{VFP}}] = \frac{d}{dt}\mathscr{S}[\mu_t^{\mathbf{VFP}}] = -\mathscr{I}[\mu_t^{\mathbf{VFP}}] \le 0. \tag{1.3.26}$$

1.3.2 Our Assumptions

Assumptions 1.3.1 ($\mathbb{HMV}$)**.** We put the following hypotheses on the potentials which will ensure properties of existence, uniqueness and contraction:

— (**H1**)**(Hessian)** The hessian of the confinement potential V is bounded from below and the hessians of the interaction potentials $W^{(k)}$, $k = 2,\dots,N$, are bounded.

— (**H2**)**(Lyapunov)** There are two positive constants c_1 and c_2 such that

$$\forall x \in \mathbb{R}^d, \quad x\cdot\nabla V(x) \ge c_1|x|^2 - c_2. \tag{1.3.27}$$

This hypothesis is a Lyapunov condition.

Remark 1.3.3. Since the Hessian $\nabla^2 V$ of V is bounded from below and V satisfies a Lyapunov condition (**H2**), $\alpha \propto e^{-V}$ satisfies a *logarithmic Sobolev inequality* (see *e.g.* [59, 60]).

— (**H3**) For all $k \in \{2,\dots,N\}$,

$$\forall \lambda > 0, \quad \int e^{\lambda W^{(k),-}(x) - \sum_{j=1}^k V(x_j)}\, dx < +\infty. \tag{1.3.28}$$

Remark 1.3.4. This assumption is trivially satisfied if the $W^{(k)}$ are bounded from below. If (**H2**) holds, it is also always true if $W^{(k),-}(x_1,\dots,x_k) = o(\sum_{j=1}^k |x_j|^2)$ as $|x_1|^2 + \dots + |x_k|^2 \to +\infty$ (since (**H2**) involves that $\liminf_{|x|\to+\infty} V(x)/|x|^2 > 0$). Under the exponential integrability condition, if $\mathbf{H}[\mu|\alpha] < +\infty$, for all $k \in \{2,\dots,N\}$, by Donsker-Varadhan variational formula (in the bounded case) and Fatou's lemma (by approximating $W^{(k),-}$ with $(\min(W^{(k),-}, L))_{L\ge 0}$), we have $W^{(k),-} \in L^1(\mu^{\otimes k})$: then $\mathbf{dom}(\mathbf{H}_W)$ given by eq. (1.3.14) satisfies

$$\mathbf{dom}(\mathbf{H}_W) = \Big\{\mu, \quad \mathbf{H}[\mu|\alpha] < +\infty\Big\}.$$

— (**H4**)**(Logsob)** The invariant measure μ_n of the system satisfies a logarithmic Sobolev inequality such that

$$\limsup_{n\to+\infty} \rho_{LS}(\mu_n) > 0. \tag{1.3.29}$$

— (**H5**)**(Contraction)** There exists a distance $\mathbf{d}_{Lip}$ on a subset $\mathscr{Z}$ of $\mathscr{P}(\mathbb{R}^d)$ such that we have: $(\mathscr{P}_2(\mathbb{R}^d), \mathscr{W}_2)$ *continuously injects* into $(\mathscr{Z}, \mathbf{d}_{Lip})$ and $\Phi : \mu \in \mathscr{Z} \longmapsto \Phi(\mu)(dx) := \frac{1}{Z_\mu} e^{-\frac{\delta F}{\delta m}(\mu,x)-V(x)} dx \in \mathscr{Z}$ satisfies

$$\exists k \in (0,1[, \quad \forall \mu, \nu \in \mathscr{Z}, \quad \mathbf{d}_{Lip}(\Phi(\mu), \Phi(\nu)) \leq k \mathbf{d}_{Lip}(\mu, \nu). \tag{1.3.30}$$

In others terms, Φ is k-Lipschitz (*contraction*) for $\mathbf{d}_{Lip}$.

Remark 1.3.5. The two above assumptions are not easy to check in practice. In section 1.4.3, we thus provide several many-body interaction examples where these conditions apply. Nevertheless, let us give some first comments below

▷ **About** (**H4**): (**H4**) can be certainly satisfied under Bakry-Emery criterion (see proposition 1.4.1). There also exist some specific conditions called Zegarlinski conditions (see the references [133],[238],[239]): we recall Zegarlinski conditions refer to specific conditions on the Hessian of the interaction potential, which are then used (together with other conditions) to prove a uniform logarithmic Sobolev inequality. Finally, let us note that as μ_n is a Gibbs measure with respect to $\alpha^{\otimes n}$ and its Hamiltonian is $H_{\alpha,n} := n \sum_{k=2}^{N} U_n(W^{(k)})$, if this Hamiltonian has bounded oscillations ($\mathbf{osc}(H_{\alpha,n}) := \sup H_{\alpha,n} - \inf H_{\alpha,n} < +\infty$) uniformly in n, then we can show that by property of tensorization and stability by bounded perturbation, we have (**H4**) seen that according to Royer's book [207, Proposition 3.1.18], $\rho_{LS}(\mu_n) \geq \rho_{LS}(\alpha) e^{-\mathbf{osc}(H_{\alpha,n})}$. In proposition 1.4.1, our examples will be yet built with the help of the simpler Bakry-Emery condition. In proposition 1.4.4, we provide another class of examples which do not require Bakry-Emery condition.

▷ **About** (**H5**): As concerns (**H5**), we also give some explicit conditions in proposition 1.4.1 and proposition 1.4.4 with $\mathscr{Z} = \mathscr{P}_2(\mathbb{R}^d)$ and $\mathbf{d}_{Lip} = \mathscr{W}_1$. (**H5**) ensures uniqueness of the fixed point (invariant measure of the McKean-Vlasov process): in *statistical physics*, we say that we have no *phase transition*. This is the crucial point for the proof: $\mathbf{H}_W = \mathbf{H}[\cdot|\Phi(\cdot)]$ (which justifies the name *mean field entropy*). The contractivity assumptions in eq. (1.3.30) can follow from Eberle conditions (*lipschitzian spectral gap condition for one particle*): see [133]. To obtain uniqueness, some authors also require *displacement-convexity* (see e.g. [194],[52]): assuming that the functional G in $V^{\sigma,G} : \mu \mapsto \frac{\sigma^2}{2}\mathbf{H}[\mu|\alpha] + G(\mu)$ is displacement convex. And as the relative entropy is

strictly displacement-convex, $V^{\sigma,G}$ is also strictly displacement-convex, which implies the existence of an entropy minimizer ensuring its uniqueness.

Assumptions 1.3.2 ($\mathbb{VFP}$)**.** In this case, all the conditions stated in $\mathbb{HMV}$ assumptions 1.3.1 are assumed, together with the two following additional ones

— **VFP1.** Lipschitz interactions:

$$\forall k \in \{2,3,\dots,N\} \quad \exists K > 0, \quad ||\nabla W^{(k)}|| \leq K. \tag{1.3.31}$$

— **VFP2.** Lyapunov condition on confinement:

$$||\nabla^2 V||_{\mathbf{op}} \leq K_1 |\nabla V| + K_2. \tag{1.3.32}$$

Remark 1.3.6. Either of these conditions ensures that the kinetic Fokker-Planck semi-group converges exponentially (as a family of operators of $\mathscr{H}^1(\mu_Z^n)$ indexed by time) towards μ_Z^n and uniformly in the number of particles (see [132] or [227]).

1.4 Main Theorems

1.4.1 First-order case

Under $\mathbb{HMV}$ assumptions 1.3.1, we establish (see section 1.6 for the proof) the following two main results (thus generalizing those of [133]). Let $(\mu_t)_{t\geq 0}$ (given by the arrow (**1**) in fig. 1.1) be the flow of solution distributions of the McKean-Vlasov equation associated with the particle system defined by the U−statistic and the confinement potential. Then for any initial condition admitting a moment of order 2, the mean field entropy $\mathbf{H}_W$ decreases exponentially along the flow, i.e.:

Theorem 1.4.1 (Exponential decreasing of mean-field entropy)**.** *Assume* $\mathbb{HMV}$ *assumptions 1.3.1 and let* $\mu_0 \in \mathscr{M}_1^2(\mathbb{R}^d)$ *be an initial condition. Then*

$$\forall t \geq 0, \quad \mathbf{H}_W[\mu_t] \leq \mathbf{H}_W[\mu_0] e^{-\rho_{LS}\frac{t}{2}}. \tag{1.4.1}$$

From the exponential decrease of the mean field entropy along the flow, we deduce the following exponential convergence in Wassertein metric:

Theorem 1.4.2 (Exponential convergence in Wasserstein metric from flow to equilibrium)**.** *Assume* $\mathbb{HMV}$ *assumptions 1.3.1 give us an initial condition* $\mu_0 \in \mathscr{M}_1^2(\mathbb{R}^d)$*. Then*

$$\forall\, t \geq 0, \quad \mathscr{W}_2^2(\mu_t, \mu_\infty) \leq \frac{2}{\rho_{LS}} \mathbf{H}_W[\mu_0] e^{-\rho_{LS}\frac{t}{2}}. \tag{1.4.2}$$

1.4.2 Kinetic case

For kinetic type models, the extension of the above results relies on applications of hypocoercivity arguments (see e.g. [132] or [227] for background). In this setting, we first obtain an exponential decrease in $|||\cdot|||_{H^1\to H^1}$ norm (defined in section 1.2).

Theorem 1.4.3 (Uniform exponential convergence to equilibrium in the weighted Sobolev space)**.** *Assume* $\mathbb{VFP}$ *assumptions 1.3.2 and give us an initial condition* $\mu \in \mathscr{M}_1^2(\mathbb{R}^d \times \mathbb{R}^d)$*. Then*

$$\exists \alpha > 0 \quad \exists \beta > 0 \quad \forall\, n \geq 2, \quad \left|\left|\left| P_t^{Z,(n)} - \mu_Z^n \right|\right|\right|_{H^1\to H^1} \leq \alpha e^{-\beta t}. \tag{1.4.3}$$

Remark 1.4.1. We still have theorem 1.4.3 if we replace the uniform logarithmic Sobolev inequality given in $\mathbb{HMV}$ assumptions 1.3.1 by a uniform Poincaré inequality. We keep the logarithmic Sobolev inequality to have the following theorem 1.4.4. Note that the constants $\alpha > 0$ and $\beta > 0$ can be made explicit uniform. The originality of the proof relies on functional inequalities and hypocoercivity with Lyapunov type conditions, usually not suitable to provide adimensional results.

Theorem 1.4.4 (Exponential decay in Wasserstein metric)**.** *Assume* $\mathbb{VFP}$ *assumptions 1.3.2 and* $\nabla^2 V$ *is bounded. Then there are constants* $C > 0$, $\xi > 0$ *and* $\kappa > 0$ *such that* $\forall \mu \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, $\forall\, n \geq 2$ *and* $\forall\, t > 0$,

$$\mathbf{H}[\mu_Z^n(t)|\mu_Z^n] \leq C\mathbf{H}[\mu_Z^n(0)|\mu_Z^n] e^{-\xi t}, \tag{1.4.4}$$

$$\mathscr{W}_2^2(\mu_t^{\mathbf{VFP}}, \mu_\infty^Z) \leq \kappa C \mathscr{S}[\mu] e^{-\xi t}, \tag{1.4.5}$$

where μ *is the initial condition and* $\mathscr{S}$ *(defined in eq. (1.3.24)) is the mean-field entropy associated with our second order system given by eq. (1.3.18).*

1.4.3 Examples

Let us begin with results which provide some explicit conditions on V and the $W^{(k)}$ under which our results apply. We only focus on $\mathbb{HMV}$ assumptions 1.3.1 but the extension to $\mathbb{VFP}$

assumptions 1.3.2 only requires to add the constraints on $\nabla W^{(k)}$ and $\nabla^2 V$ introduced in **VFP1** and **VFP2**. We will use the notation $\underline{A}$ for the lowest eigenvalue of a symmetric matrix.

Proposition 1.4.1 (Application in example 1.4.1)**.** *Assume that* (**H1**) *holds, that* $W^{(k),-}(x_1,\dots,x_k) = o(\sum_{j=1}^k V(x_j))$ *as* $|x_1|^2+\dots+|x_k|^2 \to +\infty$ *and that the following assumption is fulfilled:*

$$\underline{\lambda} > \sum_{k=2}^{N} k(k-1)\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op},\infty} \tag{1.4.6}$$

where $\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op},\infty} = \sup_{x\in(\mathbb{R}^d)^k}\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op}}$ *and*

$$\underline{\lambda} = \inf_{x\in\mathbb{R}^d}\left(\underline{\lambda}_{\nabla^2 V(x)} + \inf_{y\in(\mathbb{R}^d)^{k-1}}\sum_{k=2}^{N} k\underline{\lambda}_{\nabla_{11}^2 W^{(k)}(x,y)}\right).$$

Then, $\mathbb{HMV}$ *assumptions 1.3.1 is true.*

Proof. The proof of this result is achieved in section 1.6. ■

Proposition 1.4.2 (Application in example 1.4.2)**.** *Assume that*

1. (**H1**), (**H2**) *and* (**H3**) *hold;*
2. $\forall k\in\{2,\dots,N\}$, $x\mapsto \nabla_{x_1} W^{(k)}(x)$ *is bounded;*
3. $\mathscr{E}_W : \mu \mapsto \sum_{k=2}^N \int W^{(k)} d\mu^{\otimes k}$ *is convex in the flat interpolation sense:*

$$\forall t\in[0,1] \quad \forall(\mu,\nu)\in\mathscr{P}_2(\mathbb{R}^d)^2, \quad \mathscr{E}_W((1-t)\mu+t\nu) \le (1-t)\mathscr{E}_W(\mu)+t\mathscr{E}_W(\nu); \tag{1.4.7}$$

4. *eq.* (1.4.6) *holds.*

Then, $\mathbb{HMV}$ *assumptions 1.3.1 is true.*

Proof. The proof of this result is achieved in section 1.6. ■

Remark 1.4.2. For instance, Assumption (iii) is true in example 1.4.2.

(⋆) **Regularized Skyrme model**

Example 1.4.1 (Regularized Skyrme model). One of the main areas of research in nuclear physics is the study of nuclei under extreme conditions in spin and isospin. Microscopic methods of mean field type, including the Hartree-Fock method based on the independent particle approximation, are one of the most efficient tools for theoretical predictions in this field. Representing the interactions between nucleons in the nucleus, the effective forces

nucleon-nucleon are the main ingredient of these self-consistent microscopic theories. The Skyrme interaction is a zero-range force allowing to construct the mean field in a relatively simple manner: effective phenomenological interaction of zero range which allows the interactions between nucleons in the nucleus to be modeled in a simple manner. Proposed by Skyrme ([212]), this force is limited to the sum of interactions between two and three nucleons. The interaction potential is given by

$$\overrightarrow{\mathrm{F}} := \binom{n}{2}\mathrm{U}_n(\mathrm{W}^{(2)}) + \binom{n}{3}\mathrm{U}_n(\mathrm{W}^{(3)}) \tag{1.4.8}$$

with $\mathrm{W}^{(2)}$ a potential causing two particles to interact and $\mathrm{W}^{(3)}$ a potential causing three particles to interact. In this model, the potentials are functions of Dirac distributions: therefore singular. We will regularize the problem by replacing the Dirac distributions with a smooth approximation: setting $\mathrm{G}_\sigma := \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} e^{-\frac{|\cdot|^2}{2\sigma^2}}$, $\sigma > 0$, we consider the particle system

$$d\mathrm{X}_t^{(n)} = \sqrt{2}d\mathrm{B}_t^{(n)} - \nabla \mathrm{H}_n^{(3)}(\mathrm{X}_t^{(n)})dt \tag{1.4.9}$$

where

$$\mathrm{H}_n^{(3)}(x) := \sum_{j=1}^{n} \mathrm{V}(x_j) + \frac{2}{n-1}\sum_{1\le i<j\le n} \mathrm{W}^{(2)}(x_i,x_j) + \frac{6}{(n-1)(n-2)}\sum_{1\le i<j<k\le n} \mathrm{W}^{(3)}(x_i,x_j,x_k), \tag{1.4.10}$$

with $\mathrm{W}^{(2)}(x,y) = \mathrm{G}_\sigma(x-y)$ and $\mathrm{W}^{(3)}(x,y,z) = \mathrm{G}_\sigma(x-y)\mathrm{G}_\sigma(x-z)\mathrm{G}_\sigma(y-z)$.

Proposition 1.4.3. *Assume that* V *is a* $\mathscr{C}^2$*-function. Then, if* $\inf_{x\in\mathbb{R}^d} \underline{\lambda}_{\nabla^2\mathrm{V}(x)} > c_1\sigma^{-2-d} + c_2\sigma^{-2-3d}$ *with* $c_1 = 4\times(2\pi)^{-\frac{d}{2}}$ *and* $c_2 = 12(1+4e^{-1})\times(2\pi)^{-\frac{3d}{2}}$*, the regularized Skyrme model satisfies* $\mathbb{HMV}$ *assumptions 1.3.1. If furthermore,* ***VFP2*** *holds true for* V*, then,* $\mathbb{VFP}$ *assumptions 1.3.2 is also satisfied.*

Remark 1.4.3. As expected, these conditions become more and more stringent when σ goes to 0. Thus, considering the long-time behavior of such models with singular kernels would probably require to develop specific techniques.

Proof. Since $\mathrm{W}^{(2)}$ and $\mathrm{W}^{(3)}$ are bounded, it is enough to check Assumption (1.4.6) of proposition 1.4.1. First, $\nabla\mathrm{G}_\sigma(x) = -\sigma^{-2}\mathrm{G}_\sigma(x)x$ and $\nabla^2\mathrm{G}_\sigma(x) = \sigma^{-2}(\sigma^{-2}x\otimes x - \mathbb{I}_d)\mathrm{G}_\sigma(x)$. Using that

for all x, $x\otimes x$ is a nonnegative symmetric matrix with $x\otimes x\le |x|^2\mathbb{I}_d$, it follows that

$$-\frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}}\mathbb{I}_d\le -\frac{1}{\sigma^2}G_\sigma(x)\mathbb{I}_d\le \nabla^2 G_\sigma(x)\le \frac{1}{\sigma^4}G_\sigma(x)\Big(|x|^2-\sigma^2\Big)\mathbb{I}_d. \tag{1.4.11}$$

We deduce that

$$\underline{\lambda}_{\nabla^2_{11}W^{(2)}(x,y)}=\underline{\lambda}_{\nabla^2 G_\sigma(x-y)}\ge -\frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}}$$

and that

$$\|\nabla^2_{12}W^{(2)}(x,y)\|_{\mathbf{op},\infty}\le \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}}.$$

Using that

$$\nabla^2_{11}(G_\sigma(x-y)G_\sigma(x-z))=\nabla^2 G_\sigma(x-y)G_\sigma(x-z)+\nabla^2 G_\sigma(x-z)G_\sigma(x-y)+u_{x,z}\otimes u_{x,y}+u_{x,y}\otimes u_{x,z},$$

with $u_{x,y}=\nabla G_\sigma(x-y)$, one also deduces that

$$\underline{\lambda}_{\nabla^2_{11}W^{(3)}(x,y,z)}\ge -\frac{2\|G_\sigma\|^2_\infty}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}}-2\sigma^{-4}\|G_\sigma\|_\infty|x-y|.|x-z|G_\sigma(x-y)G_\sigma(x-z)\ge -\frac{2(1+2e^{-1})}{(2\pi)^{\frac{3d}{2}}\sigma^{2+3d}},$$

where we used that $\sup_{u\in\mathbb{R}^d}\frac{|u|}{\sigma}G_\sigma(u)\le\sqrt{2e^{-1}}\|G_\sigma\|_\infty$. Finally, one similarly obtains that

$$\|\nabla^2_{12}W^{(3)}(x,y,z)\|_{\mathbf{op},\infty}\le\sigma^{-2}(1+6e^{-1})\|G_\sigma\|^3_\infty=\frac{1+6e^{-1}}{(2\pi)^{\frac{3d}{2}}\sigma^{2+3d}}.$$

Plugging these estimates into Assumption (1.4.6) yields the result. ■

(⋆⋆) Elementary Symmetric Polynomial Interaction Model (ESPIM): $\mu\in\mathscr{P}_G(\mathbb{R}^d)\mapsto P(\langle\mu,G\rangle)$ **with** $P\in\mathbb{R}[X]$

Example 1.4.2 (ESPIM model). Let us finish this section with a class of examples with polynomial interaction inspired by elementary symmetric polynomials (ESP). Note that this class will not require Bakry-Emery criterion. The polynomial interaction is built though the $\mathscr{C}^2$-potential $G:\mathbb{R}^d\to\mathbb{R}$: for $N\ge 2$, $k\in\{2,3,\dots,N\}$ and $j\in\{1,\dots,k\}$, let $G^{(j)}:(\mathbb{R}^d)^k\to\mathbb{R}$ be the following symmetric function

$$G^{(j)}(x_1,\dots,x_k):=\sum_{I\subset\mathscr{P}_k(j)}\prod_{i\in I}G(x_i)\quad\text{with}\quad\mathscr{P}_k(j):=\Big\{I\subset\{1,\dots,k\},\quad\mathbf{card}(I)=j\Big\} \tag{1.4.12}$$

and assume that $W^{(k)} \in \mathbf{Vect}_{\mathbb{R}}\{G^{(1)},\dots,G^{(k)}\}$, i.e.

$$W^{(k)} = \sum_{j=1}^{k} \beta_j^{(k)} G^{(j)} \quad \text{with} \quad (\beta_1^{(k)},\dots,\beta_k^{(k)}) \in \mathbb{R}^k. \tag{1.4.13}$$

Note that

$$G^{(j)}(x_1,\dots,x_k) = P_j(G(x_1),\dots,G(x_k)) \tag{1.4.14}$$

where P_j denotes the j^{th} ESP defined by

$$P_j(y_1,\dots,y_k) := \sum_{I\subset\mathscr{P}_k(j)} \prod_{i\in I} y_i. \tag{1.4.15}$$

The homogeneous polynomial $H : \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}$ associated with the McKean-Vlasov equation is defined by

$$H(\mu) = \int V d\mu + \sum_{k=2}^{N} \int W^{(k)} d\mu^{\otimes k}. \tag{1.4.16}$$

By symmetry and the Fubini-Tonelli-Lebesgue Theorem,

$$\int W^{(k)} d\mu^{\otimes k} = \sum_{j=1}^{k} \beta_j^{(k)} \int G^{(j)} d\mu^{\otimes k} = \sum_{j=1}^{k} \beta_j^{(k)} \sum_{I\subset\mathscr{P}_k(j)} \int \prod_{i\in I} G(x_i) \prod_{i\in I} \mu(dx_i) = \sum_{j=1}^{k} \beta_j^{(k)} \binom{k}{j} \Big(\int G d\mu\Big)^j, \tag{1.4.17}$$

and hence,

$$H(\mu) = \int V d\mu + \sum_{k=2}^{N} \sum_{j=1}^{k} \beta_j^{(k)} \binom{k}{j} \Big(\int G d\mu\Big)^j = \int V d\mu + Q\Big(\int G d\mu\Big) \tag{1.4.18}$$

$$\text{with} \quad Q := \sum_{k=2}^{N} \sum_{j=1}^{k} \beta_j^{(k)} \binom{k}{j} X^j \in \mathbb{R}_N[X].$$

Since $H : \mu \mapsto \int V d\mu + Q \circ T(\mu)$ with $T : \mu \mapsto \int G d\mu$, we have

$$\frac{\delta H}{\delta m}(\mu, x) = V(x) + Q'\Big(\int G d\mu\Big) G(x) \quad \text{with} \quad Q' := \sum_{k=2}^{N} \sum_{j=1}^{k} j\beta_j^{(k)} \binom{k}{j} X^{j-1} \in \mathbb{R}_{N-1}[X]; \tag{1.4.19}$$

$$\Longrightarrow \nabla \frac{\delta H}{\delta m}(\mu,\cdot) = \nabla V + Q'\Big(\int G d\mu\Big) \nabla G \Longrightarrow \nabla^2 \frac{\delta H}{\delta m}(\mu,\cdot) = \nabla^2 V + Q'\Big(\int G d\mu\Big) \nabla^2 G. \tag{1.4.20}$$

Furthermore, we have

$$\sup_{\mu\in\mathscr{P}_G(\mathbb{R}^d)}\left|Q'\Big(\int G d\mu\Big)\right| \le \sum_{k=2}^{N}\sum_{j=1}^{k} j|\beta_j^{(k)}|\binom{k}{j}||G||_\infty^{j-1} =: \gamma_1; \tag{1.4.21}$$

$$\sup_{\mu\in\mathscr{P}_G(\mathbb{R}^d)}\left|Q''\Big(\int G d\mu\Big)\right| \le \sum_{k=2}^{N}\sum_{j=1}^{k} j(j-1)|\beta_j^{(k)}|\binom{k}{j}||G||_\infty^{j-2} =: \gamma_2. \tag{1.4.22}$$

Proposition 1.4.4. *Assume that* G, ∇G *and* $\nabla^2 G$ *are bounded. As* $||G||_\infty < +\infty$, *we have for all* $i = 1,2$, $\gamma_i < +\infty$ *and*

$$\mathscr{P}_2(\mathbb{R}^d) \subset \mathscr{P}_G(\mathbb{R}^d) := \{\mu \in \mathscr{P}(\mathbb{R}^d), \quad G \in L^1(\mu)\} = \mathscr{P}(\mathbb{R}^d).$$

Let us further assume that

— $Q'' \ge 0$ *and the confinement potential* V *satisfies* (**H1**) *and* (**H2**)*;*

—
$$\underline{\lambda}^\star := \inf_{x\in\mathbb{R}^d}\left(\underline{\lambda}_{\nabla^2 V(x)} + \gamma_3\underline{\lambda}_{\nabla^2 G(x)}\right) > 0 \quad \text{and} \quad \frac{\gamma_2}{\underline{\lambda}^\star}||\nabla G||_\infty^2 < 1,$$

with $\gamma_3 := \inf_{\mu\in\mathscr{P}_2(\mathbb{R}^d)} Q'(\int G d\mu) \quad (|\gamma_3| \le \gamma_1)$.

Then, **ESPIM** *satisfies* $\mathbb{HMV}$ *assumptions 1.3.1. If furthermore,* ***VFP2*** *holds true for* V, *then,* $\mathbb{VFP}$ *assumptions 1.3.2 is also satisfied.*

Proof. The proof of this result is achieved in section 1.6. ■

Remark 1.4.4. If $\nabla^2\frac{\delta H}{\delta m}(\mu,\cdot) \ge \rho\mathbf{Id}$ with $\rho > 0$, by the Bakry-Emery curvature criterion, $\Phi(\mu)(dx) = \frac{1}{Z_\mu}e^{-\frac{\delta H}{\delta m}(\mu,x)}dx$ satisfies a uniform logarithmic Sobolev inequality in the measure.

1.5 Sketch of proofs and preliminaries

1.5.1 Sketch of proofs

First order case. The diagram given in fig. 1.1 summarizes the strategy of proof: we show (**4**) from (**1**) , (**2**) and (**3**). And in this diagram, the quantities involved are:

— $\mu^{\otimes n}P_t^{(n)} = \mu_n(t)$ the law at time t of the particle system induced by the confinement potential and the U−statistics;

— $\mu_n^{(i)}(t)$ the i−th marginal of $\mu_n(t)$;

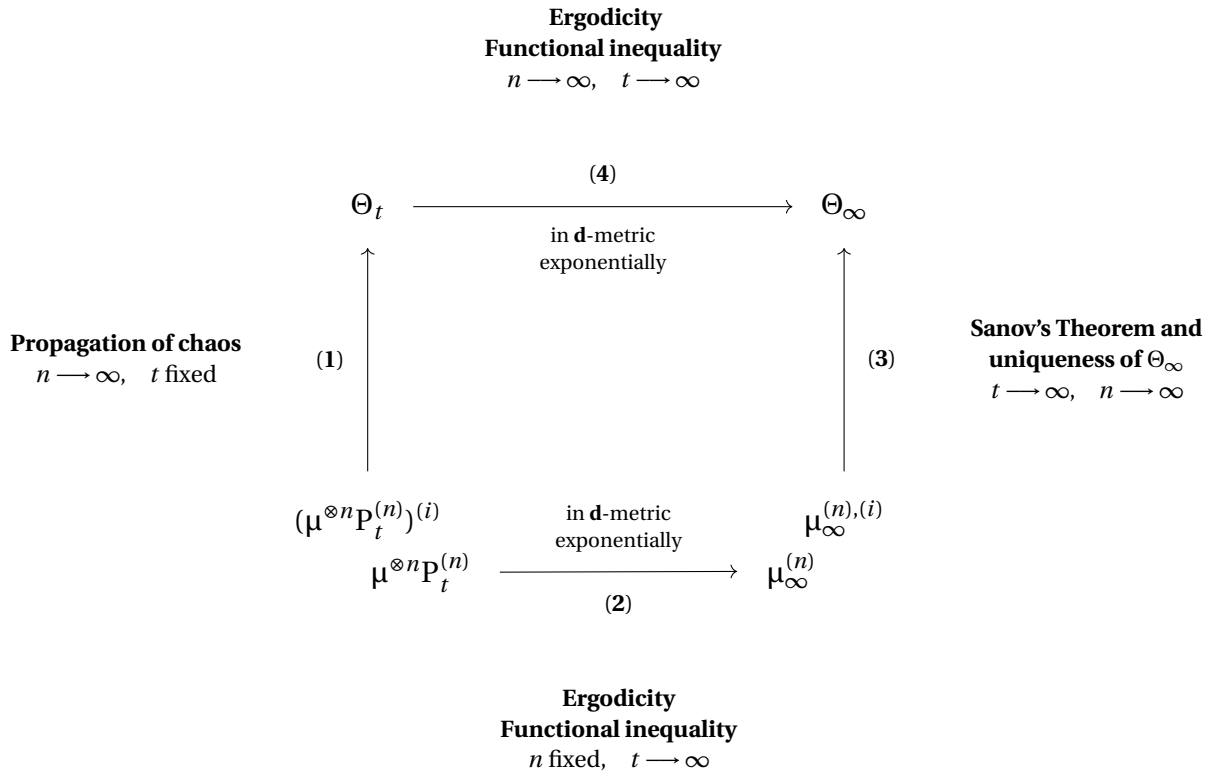

Figure 1.1 – Diagram of convergences

— $\mu_\infty^{(n)} = \mu_n$ the invariant measure of the particle system;

— $\mu_n^{(i)}$ the $i-$th marginal of μ_n;

— $\Theta_t = \mu_t$ the law at time t of the McKean-Vlasov process obtained by propagation of chaos;

— $\Theta_\infty = \mu_\infty$ the invariant measure of the McKean-Vlasov process;

— $\mathbf{d} = \mathscr{W}_2$.

Arrow (1). The McKean-Vlasov process classically appears as the mean-field limit of a particle system. This property is recalled and studied, among others, in [62]. In our case, see theorem 1.5.1.

Arrow (2). The process X^n is a homogeneous diffusion process of the Langevin-Kolmogorov type which is a class of Markov processes. In the literature, the long-time behavior for this class is classically studied (see e.g. [15, [11]]). In order to ensure this property (see section 1.5.2.theorem 1.5.3), exponentially in time and uniformly in number of particle n, we

rely on (**H4**) in $\mathbb{HMV}$ assumptions 1.3.1 and the equivalence between Sobolev's inequality, exponential decay of entropy and Talagrand's second inequality for Gibbs measures.

Arrow (3). This arrow is ensured by (**H1**), (**H2**) and (**H3**) in $\mathbb{HMV}$ assumptions 1.3.1 which allow us to obtain large deviations principle and Sanov-type theorem (see section 1.5.2.theorem 1.5.2.proposition 1.5.8).

Arrow (4). To establish this last arrow, we will use the fact that the nonlinear Sobolev inequality ($\rho_{LS}\mathbf{H}_W \leq 2\mathbf{I}_W$) given in section 1.5.2.theorem 1.5.3 is also equivalent to the exponential decrease of the mean field entropy $\mathbf{H}_W$ along the flow $(\mu_t)_{t\geq 0}$ of the McKean-Vlasov distributions and to the second nonlinear Talagrand inequality ($\rho_{LS}\mathscr{W}_2^2(\cdot,\mu_\infty) \leq 2\mathbf{H}_W$). Note that Talagrand inequalities allow to recover usual Wasserstein convergence (and then convergence in law) from entropic convergence. Note that concentration inequalities could also stem from Talagrand inequalities, although the stronger Logarithmic Sobolev inequality is more often used in this context.

Remark 1.5.1. In the case of the two-body interaction (N = 2) associated with the equation of granular media given by eq. (1.1.16), the exponential convergence in entropy (given in theorem 1.4.1) should be equivalent to the mean field log-Sobolev inequality $\rho_{LS}\mathbf{H}_W \leq 2\mathbf{I}_W$ (in theorem 1.5.3), basing on (gradient flow and Gronwall lemma)

$$-\frac{d}{dt}\mathbf{H}_W[\mu_t] = 4\mathbf{I}_W[\mu_t] \Longrightarrow \frac{d}{dt}\mathbf{H}_W[\mu_t] \leq -2\rho_{LS}\mathbf{H}_W[\mu_t] \Longrightarrow \mathbf{H}_W[\mu_t] \leq \mathbf{H}_W[\mu_0]e^{-2\rho_{LS}t} \quad (1.5.1)$$

noted by Carrillo-McCann-Villani in their convex framework. The proof of $-\frac{d}{dt}\mathbf{H}_W[\mu_t] = 4\mathbf{I}_W[\mu_t]$ demands the *regularity of* $t \longmapsto \mu_t$ (fig. 1.1) which requires the *PDE theory of the McKean-Vlasov equation*. That is why we prefer to give a *rigorously probabilistic proof based directly on the log-Sobolev inequality of* μ_n (fig. 1.1) in $\mathbb{HMV}$ assumptions 1.3.1.(**H4**). As for theorem 1.4.2 on exponential decay in Wasserstein metric, it follows from the previous one (theorem 1.4.1) via Talagrand's T2-inequality.

Second order case. The proof in this case, can also be described by the diagram given in fig. 1.1 but with the following notations:

— $\mu^{\otimes n}P_t^{(n)} = \mu_Z^n(t)$ the law at time t of the kinetic particle system induced by the confinement potential and the U−statistics;

— $(\mu^{\otimes n}P^{(n)})^{(i)} =: \mu_Z^{n,(i)}(t)$ the i−th marginal of $\mu_Z^n(t)$;

— $\mu_\infty^{(n)} = \mu_Z^n$ the invariant measure of the particle system;

— $\mu_Z^{n,(i)}$ the $i-$th marginal of μ_Z^n;

— $\Theta_t = \mu_t^{\mathbf{VFP}}$ the law at time t of the Vlasov-Fokker-Planck process obtained by propagation of chaos;

— $\Theta_\infty = \mu_\infty^Z$ the invariant measure of the Vlasov-Fokker-Planck process;

— $\mathbf{d} = |||\cdot - \cdot|||_{H^1 \to H^1}$ or $\mathbf{d} = \mathscr{W}_2$.

Arrow (1). We first recall the generator $\mathscr{L}_{Z,n}$ defined (in Hormander form) by eq. (1.5.89) is a non-symmetric hypoelliptic operator (see remark 1.5.4). The related n-particle system given by eq. (1.3.18) converges to the Vlasov-Fokker-Planck equation (mean-field limit of eq. (1.3.18)) when $n \to +\infty$ (see theorem 1.5.1).

Arrow (2). The process Z^n is a homogeneous diffusion process of the Langevin type usually called kinetic Fokker-Planck process. The study of the long-time behavior of the particle system requires the help of hypocoercivity tools (see e.g. [132] and [227]). We recall that

$$\forall x \in \mathbb{R}^{nd}, \quad S_{1,n}(x) := \sum_{i=1}^{n} V(x_i) + n \sum_{k=2}^{N} U_n(W^{(k)}). \tag{1.5.2}$$

In particular, assumptions 1.3.2 ensures the following Poincaré and log-Sobolev inequalities

— **UPI.** We say that $\mu_{1,n}$ satisfies a uniform Poincaré inequality if

$$\exists \lambda > 0 \quad \forall n \geq 2 \quad \forall \varphi \in \mathscr{C}_c^\infty(\mathbb{R}^{nd}), \quad \lambda \mathbb{V}_{\mu_{1,n}}[\varphi] \leq \mathbb{E}_{\mu_{1,n}}[||\nabla_x \varphi||^2]. \tag{1.5.3}$$

— **ULSI.** We say that $\mu_{1,n}$ satisfies a uniform logarithmic Sobolev inequality if

$$\exists \rho > 0 \quad \forall n \geq 2 \quad \forall \varphi \in \mathscr{C}_c^\infty(\mathbb{R}^{nd}), \quad \rho \mathbf{Ent}_{\mu_{1,n}}[\varphi^2] \leq \mathbb{E}_{\mu_{1,n}}[||\nabla_x \varphi||^2]. \tag{1.5.4}$$

Under **(UPI)**, we are able to obtain as an application of Villani's theorem the following exponential rate to equilibrium

$$\forall n \geq 2, \quad \left|\left|\left| P_t^{Z,(n)} - \mu_Z^n \right|\right|\right|_{H^1 \to H^1} \leq \alpha e^{-\beta t} \tag{1.5.5}$$

with constants $\alpha > 0$ and $\beta > 0$ make explicit uniform. The idea in Villani's proof of [132, Theorem.3] is as follows: if one could find a Hilbert space such that the operator $\mathscr{L}_{Z,n}$ is coercive with respect to its norm, then one has exponential convergence for the semigroup $(P_t^{Z,(n)})_{t \geq 0}$ under such a norm. If, in addition, this norm is equivalent to some usual norm

(such as $\mathscr{H}^1(\mu_Z^n)-$norm), then one obtains exponential convergence under the usual norm as well. In his statement of [227, Theorem.35], the *boundedness condition* is verified by $||\nabla^2 S_{1,n}||_{\mathbf{op}} \leq C(1+||\nabla S_{1,n}||)$ with a constant M depending unfortunately on the dimension. The L^2 and H^1 norms are not suitable to obtain a result on the non-linear system (such as eq. (1.4.4) and eq. (1.4.5)). On the other hand, thanks to **(ULSI)** playing a fundamental role in the exponential return in Wasserstein metric (see e.g. [188, Theorem.7] or [189, Theorem.10]), we are able to prove eq. (1.4.4) which in turn will allow us to deduce eq. (1.4.5).

Arrow (3). The results of large deviations on the U−statistics in the non-kinetic case in section 1.5.2 and the fact that $\mu_Z^n = \mu_{1,n} \bigotimes \mu_{2,n}$ allow to deduce that the random empire measurements of the kinetic particle system satisfy the principle of large deviations under μ_Z^n of good rate function defined by

$$\forall (\mu_x, \mu_v) \in \mathscr{P}_x(\mathbb{R}^d) \times \mathscr{P}_v(\mathbb{R}^d), \quad \mathbf{I}(\mu_x, \mu_v) := \mathbf{H}_{\mathrm{W}}[\mu_x] + \mathbf{H}[\mu_v | \mathscr{N}(0, \mathbf{Id}_d)]. \tag{1.5.6}$$

Thus there exists by inf-compactness a Maxwellian to the nonlinear Vlasov-Fokker-Planck equation and this equilibrium (invariant measure of the nonlinear Vlasov-Fokker-Planck process) is unique. See section 1.5.2.theorem 1.5.2.proposition 1.5.8.proposition 1.5.9.section 1.7.4.

Arrow (4). This part is obtained by the first-order case by exploiting the uniform logarithmic Sobolev inequality and the Hormander form given by eq. (1.5.89) (see section 1.5.2).

Remark 1.5.2. By applying hypocoercivity tools to the system with n particles given by eq. (1.3.18), we obtain a (uniform in n) convergence rate to equilibrium which in turn extends to the limiting non linear system.

1.5.2 Preliminaries

Propagation of chaos for polynomial interacting particle systems. Below, we recall or extend some conditions on the interaction potentials which guarantee the propagation of chaos for some particle systems with polynomial interaction. Even if our proof only requires such properties in finite horizon, we also provide some properties which lead to propagation of chaos uniform in time.

To this end, we use the classical (synchrounous) coupling strategy: let $(X^{(n,p)})_{p=1}^n$ denote the

particle system and $(X^{(p)})_{p=1}^n$ denote n copies of the limiting Mc-Kean-Vlasov process built with the same Brownian motions than in the particle system. Assume that all the paths have the same initial condition $X_0 \sim \mu_0$. Then, the following proposition holds:

Theorem 1.5.1 (Chaos propagation in Wasserstein $\mathcal{W}_2$ metric)**.** *Assume that* V *and the* $W^{(k)}$ *are* $\mathscr{C}^2$ *and that* $\exists \beta \in \mathbb{R}\ \forall k \in \{2,\dots,N\}\ \exists \beta_k \in \mathbb{R}\ \forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d$

$$\langle \nabla V(x) - \nabla V(y), x-y\rangle \geq -\beta ||x-y||^2; \tag{1.5.7}$$

$$\langle \nabla_{x_1} W^{(k)}(x,\cdot) - \nabla_{x_1} W^{(k)}(y,\cdot), x-y\rangle \geq -\beta_k ||x-y||^2. \tag{1.5.8}$$

Then, for every $T>0$*, a constant* K_T *exists such that for every* $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$*,*

$$\sup_{0\leq t\leq T} \mathbb{E}|X_t^{(n),1} - X_t^{(1)}|^2 \leq \frac{K_T}{n}. \tag{1.5.9}$$

Furthermore, if $\omega := \beta + \sum_{k=2}^N k\beta_k < 0$*, the upper bound is uniform in time, i.e.,*

$$\sup_{t\geq 0} \mathbb{E}|X_t^{(n),1} - X_t^{(1)}|^2 \leq \frac{K_\infty}{n}, K_\infty < +\infty. \tag{1.5.10}$$

The constants are specified in the proof.

Proof. See section 1.7.1 ■

Remark 1.5.3. ▷ For the sake of simplicity, we only provided the result for the classical McKean-Vlasov process. The extension of the result in finite horizon easily extends to the kinetic setting as soon as (**H1**).

▷ As mentioned before, the statement also leads to uniform in time propagation of chaos but it certainly requires that the function V plays a confinement role which is characterized by the fact that ω is assumed to be negative.

▷ If for all $k \in \{2,\dots,N\}$, $\nabla_{x_1} W^{(k)}$ is just uniformly bounded and uniformly Lipschitzian in the first coordinate, then it does not necessarily satisfy the conditions of the theorem 1.5.1 but if the confinement potentiel V satisfies them, we still have the conclusions at least in a short

time: for all $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$ and $\mu\in\mathscr{P}(\mathbb{R}^d)$,

$$x\cdot\nabla V(x)\geq(-\beta-|\nabla V(0)|)|x|^2-|\nabla V(0)|; \tag{1.5.11}$$

$$\begin{aligned}-\langle\nabla_{x_1}W^{(k)}(x,\cdot)-\nabla_{x_1}W^{(k)}(y,\cdot),x-y\rangle&\leq|\nabla_{x_1}W^{(k)}(x,\cdot)-\nabla_{x_1}W^{(k)}(y,\cdot)|\cdot|x-y| \\ &\leq[x_1\mapsto\nabla_{x_1}W^{(k)}(x_1,\cdot)]_1|x-y|^2;\end{aligned} \tag{1.5.12}$$

$$\begin{aligned}x\cdot\nabla\frac{\delta H}{\delta m}(\mu,x)&=x\cdot\nabla V(x)+\sum_{k=2}^N kx\cdot\nabla_{x_1}W^{(k)}*\mu^{\otimes k-1}(x) \\ &\geq(-\beta-|\nabla V(0)|)|x|^2-|\nabla V(0)|-|x|\sum_{k=2}^N k[x_1\mapsto W^{(k)}(x_1,\cdot)]_1 \\ &\geq\Big(-\beta-|\nabla V(0)|-|\nabla V(0)|-\sum_{k=2}^N k[x_1\mapsto W^{(k)}(x_1,\cdot)]_1\Big)|x|^2 \\ &\quad-|\nabla V(0)|-\sum_{k=2}^N k[x_1\mapsto W^{(k)}(x_1,\cdot)]_1.\end{aligned} \tag{1.5.13}$$

And this last inequality comes from a disjunction of cases depending on whether the vector is on the unit ball or not.

Back to U-statistics. The results on the U−statistics (proposition 1.5.1,proposition 1.5.2,proposition 1.5.3,proposition 1.5.4) and the inf-compactness of the entropy functional $\mathbf{H}_W$ (proposition 1.5.5,theorem 1.5.2,proposition 1.5.8) are inspired by [168] in the case $S=\mathbb{R}^d$. We recall that the expectation of $W^{(k)}$ under $\mu^{\otimes k}$ exists if and only if

$$\mathbb{E}_{\mu^{\otimes k}}[W^{(k),+}]<+\infty\quad\text{or}\quad\mathbb{E}_{\mu^{\otimes k}}[W^{(k),-}]<+\infty. \tag{1.5.14}$$

First we present the law of large numbers of the U−statistic (see [[153], Corollary 3.1.1] or [[168], Lemma 3.1]). We recall that U-statistics are defined in eq. (1.1.19).

Proposition 1.5.1 (law of large numbers for U−statistics)**.** *Let $(X_n)_{n\geq1}$ be a sequence of independent and identically distributed random variables with values in a measurable space $(E,\mathscr{B}(E))$ equipped with its Borelian tribe and $\Phi:E^k\longrightarrow\mathbb{R}$ a symmetric measurable function such that*

$$\mathbb{E}[|\Phi(X_1,\dots,X_k)|]<+\infty\Longrightarrow U_n(\Phi)\overset{n\to+\infty}{\longrightarrow}\mathbb{E}[\Phi(X_1,\dots,X_k)]\quad\text{with probability}\quad1. \tag{1.5.15}$$

Proof. See section 1.7.6 or [168]. ■

In terms of integrals, this result means that for any function $\Phi \in \mathscr{M}_{sym}(E^k, \mathbb{R})$ with E^k provided with the tensor tribe (or product) and any measure $\mu \in \mathscr{P}(E)$ such that $\Phi \in L^1(\mu^{\otimes k})$, we almost surely have

$$U_n(\Phi) \overset{n\to+\infty}{\longrightarrow} \mathbb{E}_{\mu^{\otimes k}}[\Phi] := \int_{E^k} \Phi(x)\mu^{\otimes k}(dx). \tag{1.5.16}$$

This result can also be seen as a law of large numbers for U−statistics. From this result, we deduce that $\forall k \in \{2,\dots,N\}$, if $W^{(k)} \in L^1(\mu^{\otimes k})$, then we almost surely have $U_n(W^{(k)})$ tends to $\mathbf{W}^{(k)}[\mu]$. We first recall the decoupling inequality of Victor H. De La Peña (see [[89], 1992]).

Proposition 1.5.2 (Decoupling and Khintchine inequalities for U−statistics)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of random variables with values in a measurable space $(E, \mathscr{B}(E))$, independent and identically distributed. We assume that*

$$(X_1^j,\dots,X_n^j)_{j=1,\dots,k} \tag{1.5.17}$$

are k independent copies of $(X_1,\dots,X_n)$. Then for all increasing convex functions $\Psi : [0,+\infty) \longrightarrow \mathbb{R}$ and measurable symmetric $\Phi : E^k \longrightarrow \mathbb{R}$ such that $\mathbb{E}[|\Phi(X_1,\dots,X_k)|] < +\infty$, we have

$$\mathbb{E}\Big[\Psi\Big(\Big|\sum_{(i_1,\dots,i_k)\in I_n^k} \Phi(X_{i_1},\dots,X_{i_k})\Big|\Big)\Big] \leq \mathbb{E}\Big[\Psi\Big(C_k\Big|\sum_{(i_1,\dots,i_k)\in I_n^k} \Phi(X_{i_1}^1,\dots,X_{i_k}^k)\Big|\Big)\Big] \tag{1.5.18}$$

with

$$C_2 := 8 \quad \text{and} \quad \forall k \geq 3, \quad C_k := 2^k \prod_{j=2}^{k} (j^j - 1). \tag{1.5.19}$$

Proposition 1.5.3. *Let $1 \leq k \leq n$, $(X_i^j)_{1\leq i\leq n, 1\leq j\leq k}$ be independent random variables with values in $(E, \mathscr{B}(E))$. For all $(i_1,\dots,i_k) \in I_n^k$, defining $\Phi_{i_1,\dots,i_k} : E^k \longrightarrow \mathbb{R}$ a measurable function of k variables, we have*

$$\begin{aligned} &\log \mathbb{E}\Big[\exp\Big(\frac{1}{|I_n^k|}\sum_{i\in I_n^k} \Phi_i(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big] \\ &\leq \frac{n-k+1}{|I_n^k|}\sum_{i\in I_n^k} \log \mathbb{E}\Big[\exp\Big(\frac{1}{n-k+1}\Phi_i(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big]. \end{aligned} \tag{1.5.20}$$

Proof. See section 1.7.6 or [168]. ■

Proposition 1.5.4 (Decoupling corollary)**.** *For $(X_i)_{i\geq 1}$ a sequence of independent and identically distributed random variables according to α, we denote $\Lambda_n(\cdot, W^{(k)})$ the log-Laplace trans-*

formation associated with the U−*statistic of order k, i.e. to within a factor, the logarithm of the moment generating function, namely*

$$\forall n \geq k \geq 2, \quad \forall \lambda > 0, \quad \Lambda_n(\lambda, W^{(k)}) := \frac{1}{n} \log \mathbb{E}\Big[e^{n\lambda U_n(W^{(k)})} \Big]. \tag{1.5.21}$$

If $W^{(k)} \in L^1(\alpha^{\otimes k})$, *then*

$$\Lambda_n(\lambda, W^{(k)}) \leq \frac{1}{k} \log \mathbb{E}\Big[\exp\Big(k C_k \lambda |W^{(k)}(X_1, \ldots, X_k)| \Big) \Big]. \tag{1.5.22}$$

Proof. See section 1.7.6 or [168]. ■

Large deviations: inf-compactness of mean-field entropy and existence of an equilibrium point. We will use a *large deviations result* ensuring the *infcompactness of the entropy functional* to show the *existence of an invariant measure* for the nonlinear process studied.

Proposition 1.5.5 (Lower bound of large deviations for L_n under μ_n^*)**.** *Under the integrability assumptions on the interaction potentials* $(W^{(k)})_{2 \leq k \leq N}$, *we have the lower bound of large deviations for* $\{\mu_n^*(L_n \in \cdot)\}_{n \geq N}$, *i.e.*

$$\begin{aligned} \forall \mathcal{O} \subset \mathcal{M}_1(\mathbb{R}^d) \quad \text{open}, \quad l^*(\mathcal{O}) &:= \liminf_{n \to +\infty} \frac{1}{n} \log(\mu_n^*(L_n \in \mathcal{O})) \\ &\geq -\inf\Big\{ \mathbf{E}_W[\mu] \Big| \quad \mu \in \mathcal{O}, \quad \forall 2 \leq k \leq N, \quad W^{(k)} \in L^1(\mu^{\otimes k}) \Big\}. \end{aligned} \tag{1.5.23}$$

In particular, we have

$$\begin{aligned} &\liminf_{n \to +\infty} \Big\{ \frac{1}{n} \log(Z_n) - \log(C) \Big\} \\ &\geq -\inf\Big\{ \mathbf{E}_W[\mu] \Big| \quad \mu \in \mathcal{M}_1(\mathbb{R}^d), \quad \forall 2 \leq k \leq N, \quad W^{(k)} \in L^1(\mu^{\otimes k}) \Big\}. \end{aligned} \tag{1.5.24}$$

Proof. See section 1.7.6 or [168]. ■

Proposition 1.5.6 (Exponential approximation of the U−statistic)**.** *Assuming that for all* $\lambda > 0$,

$$\mathbb{E}[\exp(\lambda |W^{(k)}|(X_1, \ldots, X_k))] < +\infty, \tag{1.5.25}$$

then there exists a sequence $(W_m^{(k)})_{m \geq 1}$ *of bounded continuous functions such that*

$$\forall \delta > 0, \quad \lim_{m \to +\infty} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{P}(|U_n(W_m^{(k)}) - U_n(W^{(k)})| > \delta) = -\infty. \tag{1.5.26}$$

Proof. See section 1.7.6 or [168]. ■

Theorem 1.5.2 (Large deviations principle for U−statistics)**.** *Let* $(X_i)_{i\geq 1}$ *be a sequence of independent and identically distributed random variables with distribution* α*. We assume that we have exponential integrability of the interaction potentials under the tensor products of* α *by itself, i.e.*

$$\forall k \in \{2,\dots,N\}, \quad \forall \lambda > 0, \quad \left(\mathbb{E}\left[e^{\lambda |W^{(k)}(X_1,\dots,X_k)|}\right] < +\infty \Longleftrightarrow e^{\lambda |W^{(k)}|} \in L^1(\alpha^{\otimes k})\right). \tag{1.5.27}$$

Then

$$\left\{\mathbb{P}\left((L_n, U_n(W^{(2)}),\dots,U_n(W^{(N)})) \in \cdot\right)\right\}_{n\geq N} \tag{1.5.28}$$

satisfies a large deviations principle on the product space $\mathscr{M}_1(\mathbb{R}^d)\times\mathbb{R}^{N-1}$ *and good rate function given by*

$$I_U(\mu, x_2,\dots,x_N) := \begin{cases} H[\mu|\alpha], & \text{if} \quad \forall k, \quad x_k = W^{(k)}[\mu], \\ +\infty & \text{otherwise.} \end{cases} \tag{1.5.29}$$

Proof. Let $(W_m^{(k)})_{m\geq 1}$ be the sequence of bounded continuous functions of the proof of proposition 1.5.6 (see section 1.7.6) such that for all $\lambda > 0$,

$$\varepsilon(\lambda, m, k) := \log \int_{(\mathbb{R}^d)^k} e^{\lambda |W^{(k)} - W_m^{(k)}|} d\alpha^{\otimes k} \overset{m\to+\infty}{\longrightarrow} 0. \tag{1.5.30}$$

For all $m \geq 1$, we set

$$f_m(\mu) := \left(\mu, \mathbf{W}_m^{(2)}[\mu],\dots,\mathbf{W}_m^{(N)}[\mu]\right), \quad f(\mu) := \left(\mu, \mathbf{W}^{(2)}[\mu],\dots,\mathbf{W}^{(N)}[\mu]\right). \tag{1.5.31}$$

We consider the following metric on the product space

$$\mathbf{d}\Big((\mu, x_2,\dots,x_N),(\nu, y_2,\dots,y_N)\Big) := d_{LP}(\mu,\nu) + \sum_{k=2}^{N} |x_k - y_k| = d_{LP}(\mu,\nu) + |x-y|_1, \tag{1.5.32}$$

and note that

$$\mathbf{d}(f_m(\mu), f(\mu)) = \sum_{k=2}^{N} \left| \int_{(\mathbb{R}^d)^k} (W_m^{(k)} - W^{(k)}) d\mu^{\otimes k} \right|. \tag{1.5.33}$$

The sequel of the proof is divided in three steps.

Step 1: Continuity of f_m**.** For this step, it suffices to show that for all $k \in \{2,\dots,N\}$, $\mu \in \mathscr{M}_1(\mathbb{R}^d) \longmapsto \mathbf{W}_m^{(k)}[\mu]$ is continuous for the convergence topology weak. Let $\mu_n \overset{n\to+\infty}{\longrightarrow} \mu$ in

$(\mathscr{M}_1(\mathbb{R}^d), d_{\mathrm{LP}})$. By the Skorokhod representation theorem, there exists a sequence $(\mathrm{Y}_n)_n$ of random variables with values in $\mathbb{R}^d$ such that $\mathrm{Y}_n \sim \mu_n$ and almost surely, $\mathrm{Y}_n \overset{n\to+\infty}{\longrightarrow} \mathrm{Y} \sim \mu$. Let $(\mathrm{Y}_n^{(i)}, n \geq 0, \mathrm{Y}^{(i)})_{1\leq i\leq k}$ be independent copies of $(\mathrm{Y}_n, n \geq 0, \mathrm{Y})$. We have for all i, almost surely, $\mathrm{Y}_n^{(i)} \overset{n\to+\infty}{\longrightarrow} \mathrm{Y}^{(i)}$, which implies that almost surely, $(\mathrm{Y}_n^{(1)}, \ldots, \mathrm{Y}_n^{(k)}) \overset{n\to+\infty}{\longrightarrow} (\mathrm{Y}^{(1)}, \ldots, \mathrm{Y}^{(k)})$. In particular, $\mu_n^{\otimes k}$ tends weakly to $\mu^{\otimes k}$, which proves the continuity of the above functional.

Step 2: Good exponential approximation of $(\mathrm{L}_n, \mathrm{U}_n(\mathrm{W}^{(2)}), \ldots, \mathrm{U}_n(\mathrm{W}^{(\mathrm{N})}))$ **by** $f_m(\mathrm{L}_n)$**.** By exponential approximation of the U−statistic, we have for all $\delta > 0$,

$$\lim_{m\to+\infty} \limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}\Big(\mathbf{d}\Big((\mathrm{L}_n, \mathrm{U}_n(\mathrm{W}^{(2)}), \ldots, \mathrm{U}_n(\mathrm{W}^{(\mathrm{N})})), (\mathrm{L}_n, \mathrm{U}_n(\mathrm{W}_m^{(2)}), \ldots, \mathrm{U}_n(\mathrm{W}_m^{(\mathrm{N})}))\Big) > \delta\Big) = -\infty,$$

i.e. $(\mathrm{L}_n, \mathrm{U}_n(\mathrm{W}_m^{(2)}), \ldots, \mathrm{U}_n(\mathrm{W}_m^{(\mathrm{N})}))$ is a good exponential approximation of $(\mathrm{L}_n, \mathrm{U}_n(\mathrm{W}^{(2)}), \ldots, \mathrm{U}_n(\mathrm{W}^{(\mathrm{N})}))$.

Moreover, $(\mathrm{L}_n, \mathrm{U}_n(\mathrm{W}^{(2)}), \ldots, \mathrm{U}_n(\mathrm{W}^{(\mathrm{N})}))$ and $f_m(\mathrm{L}_n)$ are exponentially equivalent because we have the following uniform estimate

$$\begin{aligned} \Big|\mathrm{U}_n(\mathrm{W}_m^{(k)}) - \int \mathrm{W}_m^{(k)} d\mathrm{L}_n^{\otimes k}\Big| &\leq \Big(1 - \frac{|\mathrm{I}_n^k|}{n^k}\Big)\Big(|\mathrm{U}_n(\mathrm{W}_m^{(k)})| + ||\mathrm{W}_m^{(k)}||_\infty\Big) \\ &\leq 2\Big(1 - \frac{|\mathrm{I}_n^k|}{n^k}\Big)||\mathrm{W}_m^{(k)}||_\infty \overset{n\to+\infty}{\longrightarrow} 0. \end{aligned} \tag{1.5.34}$$

We get that when $m \to +\infty$, $f_m(\mathrm{L}_n)$ is a good exponential approximation of

$$(\mathrm{L}_n, \mathrm{U}_n(\mathrm{W}^{(2)}), \ldots, \mathrm{U}_n(\mathrm{W}^{(\mathrm{N})})).$$

Step 3: LDP. By Sanov theorem and the LDP approximation theorems, to get the desired LDP, it suffices to show that for all $\mathrm{L} > 0$,

$$\sup_{\mu,\ \mathbf{H}[\mu|\alpha]\leq \mathrm{L}} \mathbf{d}(f_m(\mu), f(\mu)) \overset{m\to+\infty}{\longrightarrow} 0. \tag{1.5.35}$$

Indeed, for all $\lambda > 0$, $\mathrm{L} > 0$ and μ such that $\mathbf{H}[\mu|\alpha] \leq \mathrm{L}$, by the variational formula of Donsker-Varadhan and Fatou's lemma, we have for all $k \in \{2 \ldots, \mathrm{N}\}$,

$$\begin{aligned} \int |\mathrm{W}_m^{(k)} - \mathrm{W}^{(k)}| d\mu^{\otimes k} &\leq \frac{1}{\lambda}\Big(\mathbf{H}[\mu^{\otimes k}|\alpha^{\otimes k}] + \log \int e^{\lambda|\mathrm{W}_m^{(k)} - \mathrm{W}^{(k)}|} d\alpha^{\otimes k}\Big) \\ &\leq \frac{1}{\lambda}\Big(k\mathrm{L} + \varepsilon(\lambda, m, k)\Big). \end{aligned} \tag{1.5.36}$$

This completes the proof of the theorem because λ is arbitrary and for all $\lambda > 0$, $\varepsilon(\lambda, m, k) \overset{m\to+\infty}{\longrightarrow} 0$. ■

We are now able to prove the inf-compactness of the mean-field entropy functional.

Proposition 1.5.7 (Inf-compactness of the mean-field entropy functional)**.** *The mean-field entropy functional is inf-compact.*

Proof of proposition 1.5.7. We will do the proof in three steps. We recall that if we have a good rate function, then its infimum on any closed nonempty is reached, that is to say that this infimum is a minimum.

First, Assumption (A1) of [168] is clearly satisfied: taking $\mu = h\alpha$ with h continuous with compact support, we deduce from the local boundedness of $W^{(k)}$ that

$$\mathbf{H}[\mu|\alpha] + \mathbb{E}_{\mu^{\otimes k}}[W^{(k),+}] < +\infty.$$

By (**H3**), we have the strong exponential integrability condition of (A1).

Moreover, under (**H2**), we necessarily have that $\lim_{|x|\to+\infty} \frac{V(x)}{|x|^2} > 0$. Actually, by (**H2**), there exists a positive M such that for every x with $|x| > M$, $x.\nabla V(x) \geq \frac{c_1}{2}|x|^2$. Then, for any x with $|x| \geq M$, set $\rho = \frac{M}{|x|}$. We have

$$\begin{aligned} V(x) - V(\rho x) &= (1-\rho)\int_0^1 \nabla V((1+u(1-\rho))x).x\,du \\ &= \int_0^1 \nabla V((1+u(1-\rho))x).(1+u(1-\rho))x \times \frac{1-\rho}{1+u(1-\rho)}\,du \\ &\geq \frac{c_1|x|^2}{2}\int_0^1 (1+u(1-\rho))(1-\rho)\,du \\ &\geq \frac{(1-\rho)c_1|x|^2}{2} \\ &\geq \frac{c_1|x|^2}{4} \end{aligned} \tag{1.5.37}$$

as soon as $|x| \geq 2M$ (since in this case, $\rho = \frac{M}{|x|} \leq \frac{1}{2}$). We deduce that for any $p \in [1,2)$ and any $\lambda > 0$,

$$\int_{\mathbb{R}^d} e^{\lambda|x|^p}\alpha(dx) < +\infty$$

which corresponds to the last assumption of [168]...

Step 1: $W^{(k)}$ bounded from above. In this case, we have for all $\lambda > 0$,

$$\mathbb{E}[e^{\lambda|W^{(k)}|(X_1,\dots,X_k)}] < +\infty. \tag{1.5.38}$$

In principle, large deviations for the U−statistic, under $\mathbb{P} := \alpha^{\otimes N}$, $(L_n, U_n(W^{(2)}),\dots,U_n(W^{(N)}))$ satisfies a large deviations principle on $\mathscr{M}_1(\mathbb{R}^d)\times\mathbb{R}^{N-1}$ of good rate function $\mathbf{I}_U$. As

$$\sum_{k=2}^{N} U_n(W^{(k)}) \quad \text{is continuous in} \quad (L_n, U_n(W^{(2)}),\dots,U_n(W^{(N)})), \tag{1.5.39}$$
$$\forall p>1, \quad \limsup_{n\to+\infty}\frac{1}{n}\log\mathbb{E}\Big[e^{-np\sum_{k=2}^{N}U_n(W^{(k)})}\Big] < +\infty,$$

by what precedes and the theorem of R.Ellis, we deduce that $\mu_n((L_n, U_n(W^{(2)}),\dots,U_n(W^{(N)}))\in\cdot)$ satisfies a large deviations principle with rate function defined by

$$\tilde{I}(\mu,x_2,\dots,x_N) = \mathbf{I}_U(\mu,x_2,\dots,x_N) + \sum_{k=2}^{N}x_k - \inf_{\mu,x_2,\dots,x_N}\Big\{\mathbf{I}_U(\mu,x_2,\dots,x_N)+\sum_{k=2}^{N}x_k\Big\}. \tag{1.5.40}$$

So

$$\tilde{I}(\mu,x_2,\dots,x_N) = \begin{cases}\mathbf{E}_W[\mu]-\inf_\eta \mathbf{E}_W[\eta] & \text{if} \quad \mathbf{H}[\mu|\alpha]<+\infty, \quad \forall k, \quad x_k=\mathbf{W}^{(k)}[\mu],\\ +\infty & \text{otherwise.}\end{cases} \tag{1.5.41}$$

We conclude by the principle of contraction that $\mu_n(L_n\in\cdot)$ satisfies a PGD of rate function $\mathbf{H}_W$. Note in this case that $\mathbf{E}_W$ is inf-compact, so $\mathbf{H}_W$ too.

Step 2: General case. In this case, for all $L>0$, we set $W_L^{(k)} := \min(W^{(k)}, L)$. So

$$\mathbf{E}_{W_L}[\mu] = \begin{cases}\mathbf{H}[\mu|\alpha]+\sum_{k=2}^{N}\mathbf{W}_L^{(k)}[\mu] & \text{if} \quad \mathbf{H}[\mu|\alpha]<+\infty,\\ +\infty & \text{otherwise.}\end{cases} \tag{1.5.42}$$

is inf-compact on $\mathscr{M}_1(\mathbb{R}^d)$ by **step 1**. This proves that $\mathbf{H}_W$ is also inf -compact by passing to

the monotonous limit. For all closed $\mathscr{F} \subset \mathscr{M}_1(\mathbb{R}^d)$ and $L > 0$, we have

$$\begin{aligned}\mu_n^*(L_n \in \mathscr{F}) &= \int \mathbb{1}_{L_n \in \mathscr{F}} \exp\Big(-n \sum_{k=2}^{N} U_n(W^{(k)})\Big) d\alpha^{\otimes n} \\ &\leq \int \mathbb{1}_{L_n \in \mathscr{F}} \exp\Big(-n \sum_{k=2}^{N} U_n(W_L^{(k)})\Big) d\alpha^{\otimes n} \\ &\leq \exp\Big(-n \inf_{\mu \in \mathscr{F}} \mathbf{E}_{W_L}[\mu] + \mathbf{o}(n)\Big)\end{aligned} \tag{1.5.43}$$

and this last inequality is given by the LDP for the U−statistic and the Varadhan-Laplace lemma. It follows

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mu_n^*(L_n \in \mathscr{F}) \leq -\inf_{\mu \in \mathscr{F}} \mathbf{E}_{W_L}[\mu] \Longrightarrow \limsup_{n\to+\infty} \frac{1}{n} \log \mu_n^*(L_n \in \mathscr{F}) \leq -\inf_{\mu \in \mathscr{F}} \mathbf{E}_W[\mu] \tag{1.5.44}$$

by monotone limit and inf-compactness. In particular, for $\mathscr{F} = \mathscr{M}_1(\mathbb{R}^d)$, we deduce that

$$\limsup_{n\to+\infty} \Big\{\frac{1}{n} \log Z_n - \log C\Big\} \leq -\inf_{\mu \in \mathscr{M}_1(\mathbb{R}^d)} \mathbf{E}_W[\mu]. \tag{1.5.45}$$

By the lower bound of the large deviations for L_n under μ_n^* obtained, this upper bound and given that $\mathbf{E}_W[\mu] = +\infty$ if for a $k \in \{2, \dots, N\}$, $W^{(k)} \notin L^1(\mu^{\otimes k})$, we derive that

$$\lim_{n\to+\infty} \Big\{\frac{1}{n} \log Z_n - \log C\Big\} = -\inf_{\mu \in \mathscr{M}_1(\mathbb{R}^d)} \mathbf{E}_W[\mu] \tag{1.5.46}$$

which is a finite quantity by assumptions and inf-compactness. With this equality, we thus obtain upper and lower bounds of large deviations for $\{\mu_n(L_n \in \cdot)\}_{n \geq N}$.

■

Proposition 1.5.8 (Sanov's theorem for the Wasserstein metric by Wang et.al)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of independent random variables, identically distributed, with values in $\mathbb{R}^d$ endowed with one of its norms that we will denote $||\cdot||$ and law α. We have equivalence between the following two assertions*

1. *$(\mathbb{P}(L_n \in \cdot))_{n\geq 1}$ satisfies a principle of large deviations on the Wasserstein space $(\mathscr{M}_1^p(\mathbb{R}^d), \mathscr{W}_p)$ with speed n and good rate function $\mathbf{H}[\cdot|\alpha]$.*
2. $$\forall \lambda > 0 \quad x_0 \in \mathbb{R}^d, \quad \int_{\mathbb{R}^d} e^{\lambda ||x - x_0||^p} \alpha(dx) < +\infty. \tag{1.5.47}$$

Proof. Since we have established a LDP for the random empirical measure L_n under μ_n on

$\mathscr{M}_1(\mathbb{R}^d)$ equipped with the topology of weak convergence, it suffices to prove the exponential tension of $(\mu_n(L_n \in \cdot))_{n\geq N}$ on $(\mathscr{M}_1^p(\mathbb{R}^d), \mathscr{W}_p)$.
Let $K \subset \mathscr{M}_1^p(\mathbb{R}^d)$ be compact and $(a,b) \in [1,+\infty]^2$ a pair of conjugate exponents ($\frac{1}{a}+\frac{1}{b}=1$). By Holder's inequality, we have

$$\begin{aligned}\mu_n(L_n \notin K) &= \frac{C^n}{Z_n}\int \mathbb{1}_{L_n\notin K}\exp\Big(-n\sum_{k=2}^{N} U_n(W^{(k)})\Big)d\alpha^{\otimes n} \\ &\leq \frac{C^n}{Z_n}\Big(\alpha^{\otimes n}(L_n \notin K)\Big)^{\frac{1}{a}}\Big(\int \exp\Big(-nb\sum_{k=2}^{N} U_n(W^{(k)})\Big)d\alpha^{\otimes n}\Big)^{\frac{1}{b}}.\end{aligned} \tag{1.5.48}$$

It is deduced that

$$\begin{aligned}\limsup_{n\to+\infty}\frac{1}{n}\log\mu_n(L_n\notin K) \leq \frac{1}{a}\limsup_{n\to+\infty}\frac{1}{n}\log\alpha^{\otimes n}(L_n\notin K) - \limsup_{n\to+\infty}\frac{1}{n}\log\frac{Z_n}{C^n} \\ + \frac{1}{b}\limsup_{n\to+\infty}\frac{1}{n}\int\exp\Big(-n\sum_{k=2}^{N}U_n(bW^{(k)})\Big)d\alpha^{\otimes n}.\end{aligned} \tag{1.5.49}$$

Now the right-hand side of this inequality is upper bounded by

$$\frac{1}{a}\limsup_{n\to+\infty}\frac{1}{n}\log\alpha^{\otimes n}(L_n\notin K) + \inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu] - \frac{1}{b}\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_{bW}[\mu], \tag{1.5.50}$$

and from the above, $\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu]$ and

$$\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_{bW}[\mu] := \inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\Big\{\mathbf{H}[\mu|\alpha] + \sum_{k=2}^{N}\int bW^{(k)}d\mu^{\otimes k}\Big\}, \quad \text{are finite quantities.} \tag{1.5.51}$$

Under (**H1**), (**H2**) and (**H3**) in $\mathbb{HMV}$ assumptions 1.3.1, the LDP holds for L_n under $\alpha^{\otimes n}$ on the Wasserstein space. So, for all $L>0$, there is a compact $K_L \subset \mathscr{M}_1^p(\mathbb{R}^d)$ such that

$$\limsup_{n\to+\infty}\frac{1}{n}\log\alpha^{\otimes n}(L_n\notin K_L) \leq -aL - a\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu] + \frac{a}{b}\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_{bW}[\mu]. \tag{1.5.52}$$

It follows that

$$\limsup_{n\to+\infty}\frac{1}{n}\log\mu_n(L_n\notin K_L) \leq -L. \tag{1.5.53}$$

■

Uniqueness of invariant measure. The assumptions on the interaction potentials and the confinement potential ensure the existence (via the inf-compactness of the entropy func-

tional proven in section 1.5.2 and [168]) of an invariant measure (global minimum point for the entropy functional) for the McKean-Vlasov process obtained by propagation of chaos. It remains to prove the uniqueness.

Proposition 1.5.9 (Fixed point uniqueness)**.** *Under* (**H1**), (**H2**) *and* (**H3**) *in* $\mathbb{HMV}$ *assumptions 1.3.1, there exists* μ_∞ *a minimizer of* $\mathbf{H}_W$*. We have equivalently*

— *Critical points of energy:* $\mu_\infty \in \{\nu, \quad \frac{\delta \boldsymbol{E}_W}{\delta m}(\nu,\cdot) = 0\} \neq \emptyset$.

— *Fixed points:* $\mu_\infty \in \{\nu, \quad \Phi(\nu) = \nu\} \neq \emptyset$.

— *Maxwellians:* μ_∞ *is also an invariant probability of the McKean-Vlasov process:* $\mu_\infty \mathscr{L}_{\mu_\infty} = 0$ *or equivalently* $f_\infty := \frac{d\mu_\infty}{dx}$ *(existence of density) satisfies*

$$\mathbf{div}\Big(f_\infty \nabla \frac{\delta H}{\delta m}(\mu_\infty,\cdot) + \nabla f_\infty\Big) = 0.$$

The contraction assumption (**H5**) *in* $\mathbb{HMV}$ *assumptions 1.3.1 ensures the uniqueness of* μ_∞*.*

Proof. To do this, we will use the characterization of the local extrema of a differentiable functional in the sense of Fréchet (flat derivation) on an open set. Let

$$\begin{aligned}\mathscr{O} &:= \Big\{\mu \in \mathscr{P}(\mathbb{R}^d), \quad \mathbf{H}[\mu|\alpha] < +\infty, \quad \forall k, \quad \int W^{(k),-} d\mu^{\otimes k} < +\infty\Big\} \\ &= \mathbf{H}[\cdot|\alpha]^{-1}(]-\infty,+\infty[) \bigcap \Psi^{-1}(]-\infty,+\infty[^{N-1}),\end{aligned} \tag{1.5.54}$$

with

$$\Psi : \mu \longmapsto \Big(\int W^{(2),-} d\mu^{\otimes 2}, \ldots, \int W^{(N),-} d\mu^{\otimes N}\Big). \tag{1.5.55}$$

We know that $\mathbf{E}_W \equiv +\infty$ over $\mathscr{O}^C$. By Fréchet differentiability of the relative entropy $\mathbf{H}[\cdot|\alpha]$ and of Ψ on $\mathscr{M}_1(\mathbb{R}^d)$ endowed with its structure of differential Fréchet manifold, $\mathscr{O}$ is open as an intersection of open sets. We deduce that the local extrema (here minimum) of $\mathbf{E}_W$ are critical points on $\mathscr{O}$, i.e. $\mu \in \mathscr{O}$ such that

$$Z_\mu := \int e^{-\frac{\delta F}{\delta m}(\mu,x) - V(x)} dx < +\infty, \quad \frac{\delta \mathbf{E}_W}{\delta m}(\mu,\cdot) \equiv 0 \Longleftrightarrow \mu(dx) = \frac{1}{Z_\mu} e^{-\frac{\delta F}{\delta m}(\mu,x) - V(x)} dx. \tag{1.5.56}$$

According to the hypothesis (**H5**) of $\mathbb{HMV}$ assumptions 1.3.1, we have

$$\mathbf{d}_{Lip}(\Phi(\mu),\Phi(\nu)) \leq k \mathbf{d}_{Lip}(\mu,\nu), \tag{1.5.57}$$

and since there is a fixed point, suppose by absurd that there is more than one, i.e. there is

$\mu_1, \mu_2 \in \mathscr{O}$ such that $\mu_1 \neq \mu_2$ and for all i, $\Phi(\mu_i) = \mu_i$. It follows that $k \geq 1$ which is absurd because $k < 1$. ■

Cesãro tensorial: About entropies and Fisher Informations. We will establish *convergences in entropy and Fisher information* which are useful for the proof of the *exponential decrease of the mean field entropy* and the establishment of the *nonlinear Talagrand inequality*.

Proposition 1.5.10 (**H**-Tensorization)**.** *For any probability measure* ν *on* $\mathbb{R}^d$ *such that* $\mathbf{H}[\nu|\alpha] < +\infty$*, we have:*

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{H}_W[\nu], \quad \text{where} \quad \mu_n \quad \text{is defined in} \quad eq.\ (1.3.10). \tag{1.5.58}$$

Proof. For μ such that $\mu \ll \alpha$ and for all $k \in \{2,\dots,N\}$, $W^{(k),-} \in L^1(\mu^{\otimes k})$, we have

$$\begin{aligned}\frac{1}{n}\mathbf{H}[\mu^{\otimes n}|\mu_n] &= \frac{1}{n}\mathbb{E}_{\mu^{\otimes n}}\Big[\frac{d\mu^{\otimes n}}{d\alpha^{\otimes n}} + n\sum_{k=2}^{N} U_n(W^{(k)}) + \log\frac{Z_n}{C^n}\Big] \\ &= \mathbf{H}[\mu|\alpha] + \sum_{k=2}^{N}\int W^{(k)} d\mu^{\otimes n} + \frac{1}{n}\log Z_n - \log C.\end{aligned} \tag{1.5.59}$$

We recall that $\alpha(dx) := \frac{e^{-V(x)}}{C}dx$. Under the assumption (**H2**) in $\mathbb{HMV}$ assumptions 1.3.1, we know that $\exists\, \lambda_0 > 0$ such that:

$$\int_{\mathbb{R}^d} e^{\lambda_0|x|^2}\alpha(dx) < +\infty. \tag{1.5.60}$$

By asking:

$$\widetilde{Z}_n := \int_{(\mathbb{R}^d)^n} e^{-n\sum_{k=2}^{N} U_n(W^{(k)})}\alpha^{\otimes n}(dx_1,\dots,dx_n), \tag{1.5.61}$$

we get: (by Fubini-Tonelli)

$$\mu_n(dx) = \frac{C^n}{Z_n}e^{-n\sum_{k=2}^{N} U_n(W^{(k)})}\alpha^{\otimes n}(dx). \tag{1.5.62}$$

Let $\nu \in \mathscr{M}_1(\mathbb{R}^d)$ be such that $\mathbf{H}[\nu|\alpha] < +\infty$. Since

$$\mathbf{H}[\nu^{\otimes k}|\alpha^{\otimes k}] = k\mathbf{H}[\nu|\alpha], \tag{1.5.63}$$

$x \longmapsto e^{\lambda_0|x|^2} \in L^1(\alpha)$ and

$$\forall k \in \{2,\dots,N\} \quad \forall x \in \mathbb{R}^{kd}, \quad |W^{(k)}(x)| \leq \beta(1 + \sum_{j=1}^{k}||x_j||^2) \tag{1.5.64}$$

by boundedness of its hessian $\nabla^2 W^{(k)}$ (hypothesis (**H1**) in $\mathbb{HMV}$ assumptions 1.3.1), according to Donsker-Varadhan variational formula of entropy, we have $W^{(k)} \in L^1(\nu^{\otimes k})$. We have successively: (by a direct calculation and application of the Fubini-Tonelli theorem)

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n] = \frac{1}{n}\mathbf{Ent}_{\mu_n}\Big[\frac{d\nu^{\otimes n}}{d\mu_n}\Big] = \frac{1}{n}\int_{(\mathbb{R}^d)^n} \log\Big(\frac{d\nu^{\otimes n}}{d\mu_n}\Big) d\nu^{\otimes n} \tag{1.5.65}$$

We deduce that:

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n] = \frac{1}{n}\int \sum_{i=1}^{n} \log\Big(\frac{d\nu}{d\alpha}(x_i)\Big) d\nu^{\otimes n} + \sum_{k=2}^{N}\int U_n(W^{(k)} d\nu^{\otimes n} + \frac{1}{n}\log(\widetilde{Z}_n) \tag{1.5.66}$$

$\lim_{n\to+\infty} \frac{1}{n}\log(\widetilde{Z}_n) = -\inf_{\eta\in\mathscr{M}_1(\mathbb{R}^d)} \mathbf{E}_W[\eta]$ (see theorem 1.5.2),

$$\frac{1}{n}\int \sum_{i=1}^{n} \log\Big(\frac{d\nu}{d\alpha}(x_i)\Big) d\nu^{\otimes n} = \mathbf{H}[\nu|\alpha] \tag{1.5.67}$$

and finally, we also have: (see proposition 1.5.1)

$$\sum_{k=2}^{N}\int U_n(W^{(k)}) d\nu^{\otimes n} = \sum_{k=2}^{N}\int W^{(k)}(x)\nu^{\otimes k}(dx). \tag{1.5.68}$$

Thereby:

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{H}[\nu|\alpha] + \sum_{k=2}^{N}\int W^{(k)}(x)\nu^{\otimes k}(dx) - \inf_{\eta\in\mathscr{M}_1(\mathbb{R}^d)} \mathbf{E}_W[\eta] = \mathbf{H}_W[\nu]. \tag{1.5.69}$$

What needed to be proven. ■

Proposition 1.5.11 (**I**-Tensorization)**.** *If* $\mathbf{I}[\nu|\alpha] < +\infty$*, we have:*

$$\frac{1}{n}\mathbf{I}[\nu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{I}_W[\nu]. \tag{1.5.70}$$

Proof. For any probability measure ν on $\mathbb{R}^d$ such that $\mathbf{I}[\nu|\alpha] < +\infty$, by the Lyapunov condition (**H2**) in $\mathbb{HMV}$ assumptions 1.3.1 on the potential V, we have:

$$c_1\int |x|^2 d\nu \le c_2 + \mathbf{I}[\nu|\alpha] < +\infty. \tag{1.5.71}$$

As the second order derivatives of $W^{(k)}$ are bounded by the condition (**H1**) in $\mathbb{HMV}$ assump-

tions 1.3.1 on its Hessian, $\nabla_{x_j} W^{(k)}$ has a linear increase. So $\nabla_{x_j} W^{(k)} \in L^2(\nu^{\otimes k})$. By the law of large numbers for independent and identically distributed sequences, we have successively:

$$\begin{aligned}
\frac{1}{n}\mathbf{I}[\nu^{\otimes n}|\mu_n] &= \frac{1}{4n}\int \Big|\Big|\nabla \log\Big(\frac{d\nu^{\otimes n}}{d\mu_n}\Big)\Big|\Big|^2 d\nu^{\otimes n} \\
&= \frac{1}{4n}\int \sum_{i=1}^{n} \Big|\Big|\nabla_{x_i} \log\Big(\frac{d\nu^{\otimes n}}{d\alpha^{\otimes n}}\Big) + n\sum_{k=2}^{N} \nabla_{x_i} U_n(W^{(k)})\Big|\Big|^2 d\nu^{\otimes n} \\
&= \frac{1}{4}\int \Big|\Big|\nabla \log\Big(\frac{d\nu}{d\alpha}\Big)(x_1) + n\sum_{k=2}^{N} \nabla_{x_1} U_n(W^{(k)})\Big|\Big|^2 d\nu^{\otimes n} \\
&\overset{n\to+\infty}{\longrightarrow} \frac{1}{4}\int \Big|\Big|\nabla \log\Big(\frac{d\nu}{d\alpha}\Big)(y) \\
&+ \sum_{k=2}^{N}\sum_{j=1}^{k}\int \nabla_{x_j} W^{(k)}(x_1,\dots,x_{j-1},y,x_{j+1},\dots,x_k)\nu^{\otimes k-1}\Big(\prod_{i=1,i\neq j}^{k} dx_i\Big)\Big|\Big|^2 \nu(dy) \\
&= \mathbf{I}_W[\nu].
\end{aligned} \tag{1.5.72}$$

■

We recall the tensorisation property of relative entropy: The proposition 1.5.12 on entropy and tensor product allows us, in what follows, to show the exponential decreasing of mean-field entropy along the flow of solution distributions of the McKean-Vlasov equation associated with the particle system.

Proposition 1.5.12 (Relative entropy and tensor product)**.** *Let* $\prod_{i=1}^{N}\alpha_i$ *and* Q *respectively be a product probability measure and a probability measure defined on* $E_1\times\cdots\times E_N$ *a product of Polish spaces. Denoting* Q_i *the marginal distribution of* x_i *under* Q*, we have:*

$$\mathbf{H}[Q|\prod_{i=1}^{N}\alpha_i] \geq \sum_{i=1}^{N}\mathbf{H}[Q_i|\alpha_i]. \tag{1.5.73}$$

Proof. See section 1.7.6 or [133]. ■

Proposition 1.5.13 (Relative entropy and Boltzmann measure)**.** *Let* μ *be a probability measure on a Polish space* E *and* $U : E \longrightarrow (-\infty,+\infty]$ *be a measurable potential such that:*

$$\int e^{-pU} d\mu < +\infty \tag{1.5.74}$$

for some $p > 1$*. Considering the Boltzmann probability measure* $\mu_U := \frac{e^{-U}}{C} d\mu$*, if for some measure* ν*,* $\mathbf{H}[\nu|\mu_U] < +\infty$*, we have successively:*

1. $\mathbf{H}[\nu|\mu] < +\infty$ *and* $U \in L^1(\nu)$.
2. $$\mathbf{H}[\nu|\mu_U] = \mathbf{H}[\nu|\mu] + \int U d\nu + \log \int e^{-U} d\mu. \tag{1.5.75}$$

Proof. See section 1.7.6 or [133]. ■

Functional and transportation inequalities. Functional inequalities are powerful tools to quantify the trend to equilibrium of Markov semigroups and have a wide range of important applications to the concentration of measure phenomenon and hypercontractivity. $\forall n$, we recall that $\mu_n(t) := \mathbb{P} \circ (X_t^n)^{-1}$ and $\beta_n := \rho_{LS}(\mu_n)$.

Theorem 1.5.3 (Transportation inequalities)**.** *Under the assumptions in* $\mathbb{HMV}$ *assumptions 1.3.1, we have*

1. $$\mathbf{H}[\mu_n(t)|\mu_n] \le \mathbf{H}[\mu_n(0)|\mu_n] e^{-\beta_n \frac{t}{2}} = \mathbf{H}[\mu_0^{\otimes n}|\mu_n] e^{-\beta_n \frac{t}{2}}; \tag{1.5.76}$$
$$\rho_{LS}(\mu_n)\mathbf{H}[\cdot|\mu_n] \le 2\mathbf{I}[\cdot|\mu_n]; \tag{1.5.77}$$
$$\rho_{LS}(\mu_n)\mathscr{W}_2^2(\cdot,\mu_n) \le 2\mathbf{H}[\cdot|\mu_n]. \tag{1.5.78}$$
2. $\exists!\ \mu_\infty \in \mathscr{M}_1^2(\mathbb{R}^d)$ *such that: (section 1.5.2.proposition 1.5.9)*
$$\mu_\infty = \mathbf{argmin}\Big\{\mathbf{H}_W[\nu], \nu \in \mathscr{M}_1(\mathbb{R}^d)\Big\}, \tag{1.5.79}$$
with $\mathbf{H}_W$ *the mean field entropy.*
3. $\rho_{LS} := \limsup_{n\to+\infty} \rho_{LS}(\mu_n) > 0$ *checks:*
$$\forall \nu \in \mathscr{M}_1^2(\mathbb{R}^d), \quad \rho_{LS}\mathbf{H}_W[\nu] \le 2\mathbf{I}_W[\nu] \quad \text{and} \quad \rho_{LS}\mathscr{W}_2^2(\nu,\mu_\infty) \le 2\mathbf{H}_W[\nu]. \tag{1.5.80}$$
We say that we have a nonlinear log-Sobolev inequality for the first inequality and a Talagrand transport inequality for the second.

Proof of theorem 1.5.3. The logarithmic Sobolev inequality of constant $\beta_n := \rho_{LS}(\mu_n)$ for μ_n given by (**H4**) in $\mathbb{HMV}$ assumptions 1.3.1, the large deviations principle (Sanov's theorem) in section 1.5.2 and the uniqueness of the minimum argument (μ_∞) in proposition 1.5.9 of the mean field entropy ensure that we have successively:

— $\forall \mu$ such as $\mathbf{H}[\mu|\alpha] < +\infty$, (section 1.5.2.proposition 1.5.10.proposition 1.5.11)

$$\frac{1}{n}\mathbf{H}[\mu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{H}_{\mathrm{W}}[\mu] \quad \text{and} \quad \frac{1}{n}\mathbf{I}[\mu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{I}_{\mathrm{W}}[\mu]. \tag{1.5.81}$$

— Equivalence between Sobolev's inequality, exponential decay of entropy and Talagrand's second inequality for Gibbs measures (Otto-Villani,[195],[226])

$$\beta_n \mathbf{H}[\cdot|\mu_n] \le 2\mathbf{I}[\cdot|\mu_n] \quad \text{and} \quad \beta_n \mathscr{W}_2^2(\cdot,\mu_n) \le 2\mathbf{H}[\cdot|\mu_n]. \tag{1.5.82}$$

— *Chaos propagation.* (theorem 1.5.1) Denoting $(\mu_t)_{t\ge 0}$ the flow of solution distributions of the McKean-Vlasov equation associated with the particle system defined by the U−statistic and the confinement potential, if $\mu_0 \in \mathscr{M}_1^2(\mathbb{R}^d)$, then for any non-empty set $\mathrm{I} \subset \mathbb{N}^*$ of finite cardinality, $\mathbb{P}_{(\mathrm{X}_t^n(i))_{i\in\mathrm{I}}}$ converges in metric L^2−Wasserstein to $\mu_t^{\otimes \mathbf{Card}(\mathrm{I})}$ (arrow (**1**) in fig. 1.1).

— Denoting $\mu_n^{(i)}$ the i-th marginal distribution of μ_n, we have by uniqueness and LDP (arrow (**3**) in fig. 1.1)

$$\mu_n^{(i)} \overset{\mathscr{L}}{\longrightarrow} \mu_\infty. \tag{1.5.83}$$

— By symmetry of μ_n, all its marginal distributions are identical and as

$$\mathscr{W}_2^2(\mu^{\otimes n},\mu_n) \ge \sum_{i=1}^n \mathscr{W}_2^2(\mu_n^{(i)},\mu) = n\mathscr{W}_2^2(\mu_n^{(1)},\mu), \tag{1.5.84}$$

we deduce:

$$n\beta_n \mathscr{W}_2^2(\mu_n^{(1)},\mu) \le 2\mathbf{H}[\mu^{\otimes n}|\mu_n]. \tag{1.5.85}$$

By equivalence of the logarithmic Sobolev inequality to the exponential decrease of entropy along the semigroup, we have (arrow (**2**) in fig. 1.1)

$$\mathbf{H}[\mu_n(t)|\mu_n] \le \mathbf{H}[\mu_n(0)|\mu_n]e^{-\beta_n\frac{t}{2}} = \mathbf{H}[\mu_0^{\otimes n}|\mu_n]e^{-\beta_n\frac{t}{2}}, \quad \mu_n(t) := \mathbb{P}\circ(\mathrm{X}_t^n)^{-1}. \tag{1.5.86}$$

And by lower semi-continuity of the Wasserstein metric, we deduce the *nonlinear* T_2*−Talagrand inequality* given by (arrow (**4**) in fig. 1.1)

$$\rho_{\mathrm{LS}}\mathscr{W}_2^2(\mu,\mu_\infty) \le \rho_{\mathrm{LS}} \liminf_{n\to+\infty} \mathscr{W}_2^2(\mu,\mu_n^{(1)}) \le 2\mathbf{H}_{\mathrm{W}}[\mu], \quad \rho_{\mathrm{LS}} := \limsup_{n\to+\infty} \beta_n > 0. \tag{1.5.87}$$

We also have the *nonlinear logarithmic Sobolev inequality* given by (arrow (**4**) in fig. 1.1)

$$\rho_{LS}\mathbf{H}_W[\cdot] \leq 2\mathbf{I}_W[\cdot]. \tag{1.5.88}$$

■

In Kinetic case. We consider $\mathscr{H}_n := \Delta_x - \nabla_x S_{1,n}\cdot\nabla = \mathscr{L}_n$ the elliptical generator associated with $\mu_{1,n} = \mu_n$.

Remark 1.5.4. $\mathscr{L}_{Z,n}$ admits the following Hormander form

$$\mathscr{L}_{Z,n} = X_0 + Y + \sum_{i=1}^{n}\sum_{j=1}^{d} X_{i,j}^2, \quad X_{i,j} = \frac{\partial}{\partial v_{i,j}} \quad X_0 = -v\cdot\nabla_v \quad Y = \nabla S_{1,n}\cdot\nabla_v - v\cdot\nabla_x \tag{1.5.89}$$

The family

$$\Big\{X_{1,1},\dots,X_{i,j},\dots,X_{i,j},\cdots,X_{n,d},[Y,X_{1,1}],\dots,[Y,X_{i,j}],\dots,[Y,X_{n,d}]\Big\} \tag{1.5.90}$$

form a basis of $\mathbb{R}^{2nd}$ at any point. Which implies by Hormander's theorem that $\mathscr{L}_{Z,n}$ is hypoelliptic. Moreover, $\mathscr{L}_{Z,n}$ is non-symmetric, i.e. in $L^2(\mu_Z^n)$, we have :

$$\begin{aligned} &\mathscr{L}_{Z,n}^* = \mathscr{L}_{Z,n} - 2Y \\ &\quad\Longrightarrow (\mathscr{L}_{Z,n}^*, \mathscr{D}(\mathscr{L}_{Z,n}^*)) \quad \text{is not a closed extension of} \quad (\mathscr{L}_{Z,n}, \mathscr{D}(\mathscr{L}_{Z,n})). \end{aligned} \tag{1.5.91}$$

The following known lemma is a key to the Lyapunov type conditions. We include its simple proof for completeness.

Proposition 1.5.14 (Lemma.8 in [132])**.** *For any function $\varphi \in \mathscr{C}^2(\mathbb{R}^{nd})$ strictly positive ($\varphi > 0$), we have*

$$\forall \psi \in \mathscr{H}^1(\mu_{1,n}), \quad \int -\frac{\mathscr{H}_n\varphi}{\varphi}\psi^2 d\mu_{1,n} \leq \int |\nabla\psi|^2 d\mu_{1,n}. \tag{1.5.92}$$

Proof of proposition 1.5.14. Indeed, by integrating by parts, we successively obtain

$$\begin{aligned} \int -\frac{\mathscr{H}_n\varphi}{\varphi}\psi^2 d\mu_{1,n} &\leq \int \Big\langle \nabla\varphi, \nabla\frac{\psi^2}{\varphi}\Big\rangle d\mu_{1,n} \\ &\leq \int \Big\langle \nabla\varphi, \frac{2\psi\nabla\psi}{\varphi} - \frac{\psi^2\nabla\varphi}{\varphi^2}\Big\rangle d\mu_{1,n} \\ &\leq \int |\nabla\psi|^2 d\mu_{1,n}. \end{aligned} \tag{1.5.93}$$

And this last inequality follows from the inequality

$$\left\langle 2\psi\nabla\psi, \frac{\nabla\varphi}{\varphi}\right\rangle \leq \frac{\psi^2|\nabla\varphi|^2}{\varphi^2} + |\nabla\psi|^2. \tag{1.5.94}$$

■

This second proposition 1.5.15 is the heart of the proof of theorem 1.4.3: this proposition is inspired by [132, Lemma.10] for the two-body interaction. It uses Lyapunov conditions, yet well know for being highly dimensional, but at the marginal level, thus providing results independent of the number of particles.

Proposition 1.5.15. *Under the conditions in* $\mathbb{VFP}$ *assumptions 1.3.2 giving* **UPI**, *there are two constants* C_1 *and* C_2 *depending on* N, K, K_1, K_2 *and* d *(dimension of* $\mathbb{R}^d$*) and such that*

$$\forall \psi \in \mathscr{H}^1(\mu_{1,n}), \quad \int ||\nabla^2 V(x_i)||^2_{op}\psi^2 d\mu_{1,n} \leq C_1 \int |\nabla_x \psi|^2 d\mu_{1,n} + C_2 \int \psi^2 d\mu_{1,n}. \tag{1.5.95}$$

Proof of proposition 1.5.15. This lemma follows from the Lyapunov property, from the particular form of the invariant measure generator [1] $\mu_{1,n}$ and from the previous proposition 1.5.14. Indeed, we have:

1.

$$||\nabla^2 V||^2_{op} \leq \eta_1\Big((1-\gamma)||\nabla V||^2 - \Delta V\Big) + \eta_2, \tag{1.5.96}$$

$$\eta_1 := 5K_1^2 \quad \eta_2 := 4K_2^2 + \frac{25K_1^4 d^2}{4} \quad \gamma := \frac{1}{5}. \tag{1.5.97}$$

2. Since the interactions are Lipschitz, we know that $\forall k \in \{2,\dots,N\} \quad \exists K^{(k)}$ such that $||\nabla W^{(k)}|| \leq K^{(k)}$.
   Let $K := \max\{K^{(k)}, \quad k = 2,\dots,N\}$. It follows that

$$-n\sum_{k=2}^{N} \nabla_{x_i} U_n(W^{(k)}) \cdot \nabla V(x_i) \leq (N-1)K|\nabla V|(x_i) \leq (N-1)\Big(\frac{K^2}{2\gamma} + \frac{\gamma}{2}|\nabla V|^2(x_i)\Big). \tag{1.5.98}$$

1. We have

$$\mathscr{H}_n = \sum_{i=1}^{n} \mathscr{T}_i, \quad \mathscr{T}_i := \Delta_{x_i} - \nabla V(x_i)\cdot\nabla_{x_i} - n\sum_{k=2}^{N} \nabla_{x_i} U_n(W^{(k)})\cdot\nabla_{x_i}.$$

But for $\varphi(x) := e^{\frac{\gamma}{2}V(x_i)}$, we have

$$\frac{\mathscr{H}_n\varphi}{\varphi} = \frac{\mathscr{T}_i\varphi}{\varphi} = \frac{\gamma}{2}\Big(\Delta V(x_i) + (\frac{\gamma}{2}-1)|\nabla V|^2(x_i) - n\sum_{k=2}^{N}\nabla_{x_i}U_n(W^{(k)})\cdot\nabla V(x_i)\Big). \tag{1.5.99}$$

Thereby

$$2\frac{\mathscr{H}_n\varphi}{\gamma\varphi} \le \Delta V(x_i) + (\frac{N\gamma}{2}-1)||\nabla V||^2(x_i) + \frac{(N-1)K^2}{2\gamma}. \tag{1.5.100}$$

Moreover, we have

$$(1-\gamma)||\nabla V||^2(x_i) - \Delta V(x_i) \le -2\frac{\mathscr{H}_n\varphi}{\gamma\varphi} + \frac{(N-1)K^2}{2\gamma}. \tag{1.5.101}$$

Therefore, by the inequality obtained in **(i)**,

$$||\nabla^2 V(x_i)||^2_{\mathbf{op}} \le \eta_1\Big(-2\frac{\mathscr{H}_n\varphi}{\gamma\varphi} + \frac{(N-1)K^2}{2\gamma}\Big) + \eta_2 \tag{1.5.102}$$

Integrating with respect to $\psi^2 d\mu_{1,n}$, we obtain

$$\begin{aligned}\int ||\nabla^2 V(x_i)||^2_{\mathbf{op}}\psi^2 d\mu_{1,n} \le \frac{2\eta_1}{\gamma}\int -\frac{\mathscr{H}_n\varphi}{\varphi}\psi^2 d\mu_{1,n} \\ + \Big(\eta_2 + \frac{(N-1)K^2}{2\gamma}\eta_1\Big)\int \psi^2 d\mu_{1,n}.\end{aligned} \tag{1.5.103}$$

And we conclude by the previous proposition 1.5.14 that

$$\int ||\nabla^2 V(x_i)||^2_{\mathbf{op}}\psi^2 d\mu_{1,n} \le C_1\int |\nabla_x\psi|^2 d\mu_{1,n} + C_2\int \psi^2 d\mu_{1,n}, \tag{1.5.104}$$

where $C_1 = \frac{2\eta_1}{\gamma}$ and $C_2 = \eta_2 + \frac{(N-1)K^2}{2\gamma}\eta_1$.

■

1.6 Proofs of Main Theorems

Proof of theorem 1.4.1. Indeed, we have the inequality (proposition 1.5.12)

$$\frac{1}{n}\mathbf{H}[\mu_n(t)|\alpha^{\otimes n}] \ge \mathbf{H}[\mu_n^{(1)}(t)|\alpha] \tag{1.6.1}$$

and by lower semi-continuity of relative entropy and propagation of chaos,

$$\liminf_{n\to+\infty} \quad \mathbf{H}[\mu_n^{(1)}(t)|\alpha] \geq \mathbf{H}[\mu_t|\alpha] \quad . \tag{1.6.2}$$

On the other hand, we have (theorem 1.5.2.theorem 1.5.3.eq. (1.5.59))

$$\frac{1}{n}\mathbf{H}[\mu_n(t)|\mu_n] \leq \frac{1}{n}\mathbf{H}[\mu_0^{\otimes n}|\mu_n]e^{-\beta_n\frac{t}{2}}; \tag{1.6.3}$$

$$\liminf_{n\to+\infty}\frac{1}{n}\mathbf{H}[\mu_0^{\otimes n}|\mu_n]e^{-\beta_n\frac{t}{2}} = \mathbf{H}_W[\mu_0]e^{-\rho_{LS}\frac{t}{2}}. \tag{1.6.4}$$

Also, as

$$\mu_n(dx) = \frac{C^n}{Z_n}e^{-n\sum_{k=2}^N U_n(W^{(k)})}\alpha^{\otimes n}(dx), \tag{1.6.5}$$

we also have

$$\frac{1}{n}\mathbf{H}[\mu_n(t)|\mu_n] = \frac{1}{n}\mathbf{H}[\mu_n(t)|\alpha^{\otimes n}] + \sum_{k=2}^N\int U_n(W^{(k)})d\mu_n(t) + \Big(\frac{1}{n}\log(Z_n) - \log(C)\Big), \tag{1.6.6}$$

and (section 1.5.2.eq. (1.5.59))

$$\sum_{k=2}^N\int U_n(W^{(k)})d\mu_n(t) \overset{n\to+\infty}{\longrightarrow} \sum_{k=2}^N\int W^{(k)}d\mu_t^{\otimes k} = \sum_{k=2}^N \mathbf{W}^{(k)}[\mu_t]; \tag{1.6.7}$$

$$\frac{1}{n}\log(Z_n) - \log(C) \overset{n\to+\infty}{\longrightarrow} -\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu]. \tag{1.6.8}$$

It is deduced that

$$\begin{aligned}\forall\, t\geq 0, \quad \mathbf{H}_W[\mu_0]e^{-\rho_{LS}\frac{t}{2}} &\geq \liminf_{n\to+\infty}\frac{1}{n}\mathbf{H}[\mu_n(t)|\mu_n] \\ &\geq \mathbf{H}[\mu_t|\alpha] + \sum_{k=2}^N \mathbf{W}^{(k)}[\mu_t] - \inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu] \\ &= \mathbf{H}_W[\mu_t].\end{aligned} \tag{1.6.9}$$

This completes the proof of the exponential decrease of entropy along the flow. ■

Proof of theorem 1.4.2. Just use the nonlinear T_2−Talagrand inequality, i.e.: (theorem 1.5.3)

$$\forall\, t\geq 0, \quad \rho_{LS}\mathscr{W}_2^2(\mu_t,\mu_\infty) \leq 2\mathbf{H}_W[\mu_t]. \tag{1.6.10}$$

This completes the proof of the desired inequality: We conclude with the theorem 1.4.1. ■

Proof of theorem 1.4.3. By the Lyapunov condition in the assumptions assumptions 1.3.2, we can apply proposition 1.5.15 and obtain that for any $\psi \in \mathscr{H}^1(\mu_{1,n})$, it holds

$$\int ||\nabla^2 V(x_i)||^2_{\mathbf{op}} \psi^2 d\mu_{1,n} \leq C_1 \int |\nabla_x \psi|^2 d\mu_{1,n} + C_2 \int \psi^2 d\mu_{1,n}, \tag{1.6.11}$$

with $C_1 = \frac{2\eta_1}{\gamma}$ and $C_2 = \eta_2 + \frac{(N-1)K^2}{2\gamma}\eta_1$ for instance which are independent of the number n of particles. It follows that the boundedness condition in Villani's theorem holds. Since the uniform Sobolev inequality implies the uniform Poincaré inequality, we can apply Villani's hypocoercivity theorem ([132, Theorem.3] or [227, Theorem.18 and Theorem.35]), which completes the proof. ■

Proof of theorem 1.4.4. Note that $(\mu_Z^n(t))_{t\geq 0}$ is a solution of a (large dimensional) linear Fokker-Planck equation, for which the exponential decay of the entropy is already known under assumptions including assumptions 1.3.2 (see e.g. [227]). Consider the generator $\mathscr{L}_{Z,n}$ given by eq. (1.3.21). Then $\Psi_n := \frac{d\mu_Z^n(t)}{d\mu_Z^n}$, the density of the law of the particle system given by eq. (1.3.18) with respect to its equilibrium distribution, solves

$$\partial_t \Psi_n = \mathscr{L}^*_{Z,n} \Psi_n. \tag{1.6.12}$$

This is a linear kinetic Fokker-Planck equation, for which convergence to equilibrium has been proven by many ways. All we need to check is that the explicit estimates we obtain do not depend on n (see e.g. respectively Theorem.7 and Theorem.10 in [188, [189]]). The key point in eq. (1.4.4) is that C and ξ do not depend on n: Indeed, as $\mu_Z^n = \mu_{1,n} \otimes \mu_{2,n}$ and these measures satisfy logarithmic Sobolev inequalities of constants $\rho_{LS}(\mu_{1,n}) = \rho$ and $\rho_{LS}(\mu_{2,n}) = 1$, μ_Z^n satisfies an inequality of logarithmic Sobolev of constant $\rho_{LS}(\mu_Z^n) := \max(\rho, 1)$; moreover, the Hessian of its Hamiltonian is uniformly bounded. This enables us to prove the following: (T2−inequality) $\exists \kappa > 0 \quad \forall n \quad \forall t$,

$$\mathscr{W}_2^2(\mu_Z^n(t), \mu_Z^n) \leq \kappa e^{-\xi t} \mathbf{H}[\mu_Z^n(0)|\mu_Z^n], \quad \mu_Z^n(0) = \mu^{\otimes n}, \quad \mu \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d). \tag{1.6.13}$$

By symmetry, propagation of chaos and Sanov's theorem (LDP), we have respectively

$$n\mathscr{W}_2^2(\mu_Z^{n,(i)}(t), \mu_Z^{n,(i)}) \leq \mathscr{W}_2^2(\mu_Z^n(t), \mu_Z^n); \tag{1.6.14}$$

$$\mu_Z^{n,(i)}(t) \overset{n\to+\infty}{\longrightarrow} \mu_t^{\mathbf{VFP}}; \tag{1.6.15}$$

$$\mu_Z^{n,(i)} \overset{n\to+\infty}{\longrightarrow} \mu_\infty^Z; \tag{1.6.16}$$

for all $i \in \{1,\dots,n\}$. We have by lower semi-continuity

$$\forall \mu \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d), \quad \mathscr{W}_2^2(\mu, \mu_\infty^{\mathrm{Z}}) \le \liminf_{n\to+\infty} \mathscr{W}_2^2(\mu, \mu_{\mathrm{Z}}^{n,(i)}) \tag{1.6.17}$$
$$\le \kappa \liminf_{n\to+\infty} \frac{1}{n} \mathbf{H}[\mu_{\mathrm{Z}}^n(0)|\mu_{\mathrm{Z}}^n]$$
$$= \kappa \mathscr{S}[\mu],$$

$$\frac{1}{n}\mathbf{H}[\mu^{\otimes n}|\mu_{\mathrm{Z}}^n] = \mathbf{H}[\mu|\alpha \otimes \mathscr{N}(0, \mathbf{Id}_d)] + \sum_{k=2}^{\mathrm{N}} \int \mathrm{U}_n(\mathrm{W}^{(k)}) d\mu^{\otimes n} \tag{1.6.18}$$
$$+ \frac{1}{n}\log(\mathrm{Z}_n) - \log(\mathrm{C})$$
$$\overset{n\to+\infty}{\longrightarrow} \mathscr{S}[\mu] := \mathscr{E}[\mu] - \mathscr{E}[\mu_\infty^{\mathrm{Z}}].$$

According to eq. (1.4.4), we have

$$\mathscr{S}[\mu_t^{\mathbf{VFP}}] \le \mathrm{C}\mathscr{S}[\mu] e^{-\xi t}. \tag{1.6.19}$$

It follows that

$$\mathscr{W}_2^2(\mu_t^{\mathbf{VFP}}, \mu_\infty^{\mathrm{Z}}) \le \kappa \mathscr{S}[\mu_t^{\mathbf{VFP}}] \le \kappa \mathrm{C} \mathscr{S}[\mu] e^{-\xi t}. \tag{1.6.20}$$

■

Proof of proposition 1.4.1. First, note that (1.4.6) ensures that $\nabla^2 \mathrm{V} \ge \rho \mathrm{I}_d$ with $\rho > 0$, which in turn implies (**H2**). For (**H3**), this follows from the assumption: $\mathrm{W}^{(k),-}(x_1,\dots,x_k) = o(\sum_{j=1}^k \mathrm{V}(x_j))$ as $|x_1|^2 + \dots + |x_k|^2 \to +\infty$.

To prove (**H4**), we use the classical Bakry-Emery criterion. To this end, let us denote by $\mathrm{A}_{ij} = \nabla^2_{ij} \mathrm{H}_n(x_1,\dots,x_n)$ where we abusively omit the dependence in $(x_1,\dots,x_n)$. Note that A_{ij} is a $d \times d$-matrix. Using the symmetry of the $\mathrm{W}^{(k)}$, we have

$$\mathrm{A}_{ii} = \nabla^2 \mathrm{V}(x_i) + \sum_{k=2}^{\mathrm{N}} k \sum_{(i_1,\dots,i_{k-1}) \in \mathrm{I}_n^{-i}} \binom{n-1}{k-1}^{-1} \nabla^2_{11} \mathrm{W}^{(k)}(x_i, x_{i_1}, \dots, x_{i_{k-1}})$$

and if $i \ne j$,

$$\mathrm{A}_{ij} = \sum_{k=2}^{\mathrm{N}} k \sum_{(i_1,\dots,i_{k-2}) \in \mathrm{I}_n^{-ij}} \binom{n-1}{k-1}^{-1} \nabla^2_{12} \mathrm{W}^{(k)}(x_i, x_j, x_{i_1}, \dots, x_{i_{k-2}})$$

where I_n^{-i} (resp. I_n^{-ij}) denotes the set of increasing sequences $i_1 < \dots < i_{k-1}$ of $[\![1,n]\!] \setminus \{i\}$

(resp. $i_1 < \ldots < i_{k-2}$ of $[\![1,n]\!]\backslash\{i,j\}$). With the notations of the proposition, one easily checks one can find a small enough ε and a large enough n_ε such that for any $\lambda \le \varepsilon$, $n \ge n_\varepsilon$ and $(x_1,\ldots,x_n) \in (\mathbb{R}^d)^n$, we have for all $i \in [\![1,n]\!]$,

$$\|\mathrm{A}_{ii}^{-1} - \lambda\|_{\mathbf{op}}^{-1} \ge \underline{\lambda} - 2\varepsilon > \sum_{k=2}^{\mathrm{N}} k(k-1)\|\nabla_{12}^2 \mathrm{W}^{(k)}\|_\infty + \varepsilon \ge \sum_{j\neq i} \|\nabla_{i,j}^2 \mathrm{A}_{ij}\|_{\mathbf{op}}.$$

This implies that for any $\lambda \in (-\infty,\varepsilon)$, the matrix $\nabla^2 \mathrm{H}_n(x_1,\ldots,x_n) - \lambda \mathrm{I}_{nd}$ is block-diagonally dominant and thus invertible. Hence, $\nabla^2 \mathrm{H}_n \ge \varepsilon \mathrm{I}_{nd}$ which in turn implies (**H4**) by the Bakry-Emery criterion.

Let us now prove that (**H5**) holds. Let $\nu_0,\nu_1 \in \mathscr{P}_2(\mathbb{R}^d)$ and set $\nu_t = (1-t)\nu_0 + t\nu_1$, $t \in [0,1]$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a 1-Lipschitz smooth function. From the very definition of Φ,

$$\langle \Phi(\nu_t), f\rangle = \frac{1}{\mathrm{Z}_{\nu_t}} \int_{\mathbb{R}^d} f(x) e^{-\mathrm{V}(x) - \sum_{k=2}^{\mathrm{N}} k \int \mathrm{W}^{(k)}(x,y)\nu_t^{\otimes k-1}(dy)} dx, \tag{1.6.21}$$

so that setting $g_t(x) = -\partial_t(\sum_{k=2}^{\mathrm{N}} k \int \mathrm{W}^{(k)}(x,y)\nu_t^{\otimes k-1}(dy))$, we get

$$\begin{aligned} \frac{d}{dt}\langle \Phi(\nu_t), f\rangle &= -\frac{\partial_t \mathrm{Z}_{\nu_t}}{\mathrm{Z}_{\nu_t}} \langle \Phi(\nu_t), f\rangle + \langle \Phi(\nu_t), f g_t\rangle \\ &= \langle \Phi(\nu_t), f g_t\rangle - \langle \Phi(\nu_t), f\rangle \langle \Phi(\nu_t), g_t\rangle \\ &= \mathrm{Cov}_{\Phi(\nu_t)}(f, g_t). \end{aligned} \tag{1.6.22}$$

For a probability μ, let $\mathscr{L}_\mu$ be the operator defined on $\mathscr{C}^2$-functions by

$$\mathscr{L}_\mu f = -\nabla f . \nabla\left(\frac{\delta \mathrm{F}}{\delta m}(\mu,\cdot) + \mathrm{V}\right) + \Delta f, \tag{1.6.23}$$

with $\frac{\delta \mathrm{F}}{\delta m}$ defined by (1.1.18). Denoting by φ_t be the solution of the Poisson equation

$$f - \Phi(\nu_t)(f) = \mathscr{L}_{\nu_t} \varphi_t \tag{1.6.24}$$

and using that $\mathscr{L}_{\nu_t}$ is self-adjoint in $\mathrm{L}^2(\Phi(\nu_t))$, we get

$$\mathrm{Cov}_{\Phi(\nu_t)}(f, g_t) = \langle \varphi_t, \mathscr{L}_{\nu_t} g_t\rangle_{\mathrm{L}^2(\Phi(\nu_t))} = -\langle \nabla\varphi_t, \nabla g_t\rangle_{\mathrm{L}^2(\Phi(\nu_t))}.$$

Note that for the second equality, we used the fact that for some $\mathscr{C}^2$-functions f and g,

$\mathscr{L}_{\nu_t}(f.g) = f\mathscr{L}_{\nu_t}g + g\mathscr{L}_{\nu_t}f + 2\nabla f.\nabla g$. With the help of Cauchy-Schwarz inequality and lemma 1.6.1, this leads to

$$|\mathrm{Cov}_{\Phi(\nu_t)}(f,g_t)| \le \|\nabla\varphi_t\|_{\mathrm{L}^2(\Phi(\nu_t))}\|\nabla g_t\|_{\mathrm{L}^2(\Phi(\nu_t))} \le \frac{1}{\underline{\lambda}}\|\nabla g_t\|_{\mathrm{L}^2(\Phi(\nu_t))}. \tag{1.6.25}$$

Let us finally focus on ∇g_t. First, one checks that

$$\begin{aligned}\partial_t \int \mathrm{W}^{(k)}(x,y)\nu_t^{\otimes k-1}(dy) &= \int \mathrm{W}^{(k)}(x,y)\sum_{j=1}^{k-1}(\nu_1-\nu_0)(dy_j)\prod_{i\neq j}\nu_t(dy_i) \\ &= (k-1)\int\int \mathrm{W}^{(k)}(x,x_2,y)(\nu_1-\nu_0)(dx_2)\nu_t^{\otimes k-2}(dy).\end{aligned} \tag{1.6.26}$$

Note that we used the symmetry of $\mathrm{W}^{(k)}$ and the fact that $|\mathrm{W}^{(k)}|$ is subquadratic (due to (**H1**)), which ensures sufficient integrability properties for the above equalities. Now, denoting by $(\mathrm{X}_{\nu_0},\mathrm{X}_{\nu_1})$ an optimal coupling of ν_0 and ν_1 for the 1-Wasserstein distance, one obtains

$$\begin{aligned}|\nabla_x\left(\int \mathrm{W}^{(k)}(x,x_2,y)(\nu_1-\nu_0)(dx_2)\right)| &\le |\mathbb{E}[\nabla_1\mathrm{W}^{(k)}(x,\mathrm{X}_{\nu_1},y)-\nabla_1\mathrm{W}^{(k)}(x,\mathrm{X}_{\nu_0},y)]| \\ &\le \|\nabla_{12}^2\mathrm{W}^{(k)}\|_{\mathbf{op},\infty}\mathscr{W}_1(\nu_0,\nu_1).\end{aligned} \tag{1.6.27}$$

Hence,

$$\|\nabla g_t\|_{\mathrm{L}^2(\Phi(\nu_t))} \le \|\nabla g_t\|_\infty \le \left(\sum_{k=2}^{\mathrm{N}} k(k-1)\|\nabla_{12}^2\mathrm{W}^{(k)}\|_{\mathbf{op},\infty}\right)\mathscr{W}_1(\nu_0,\nu_1),$$

and by (1.6.25), we get for any smooth 1-Lipschitz function

$$\begin{aligned}|\langle\Phi(\nu_1),f\rangle-\langle\Phi(\nu_0),f\rangle| &\le \int_0^1 |\frac{d}{dt}\langle\Phi(\nu_t),f\rangle|dt \\ &\le \frac{\sum_{k=2}^{\mathrm{N}} k(k-1)\|\nabla_{12}^2\mathrm{W}^{(k)}\|_{\mathbf{op},\infty}}{\underline{\lambda}}\mathscr{W}_1(\nu_0,\nu_1).\end{aligned} \tag{1.6.28}$$

By (1.4.6), a density argument and the Kantorovitch-Rubinstein duality relation, it follows that Φ is a contraction on $(\mathscr{P}_2(\mathbb{R}^d),\mathscr{W}_1)$. ■

Proof of proposition 1.4.2. We need to show that (**H4**) and (**H5**) are true. Concerning (**H5**), this is a consequence of assumption (1.4.6), exactly as in the proof of proposition 1.4.1. For (**H4**), we apply [234, Theorem 1]. The inspection of the related assumptions is divided in three steps.

Step 1: Regularity properties. For any $\mu \in \mathscr{P}(\mathbb{R}^d)$, we consider the standard Langevin process $(X_t^\mu)_{t\geq 0}$ of Hamiltonian $x_1 \mapsto \frac{\delta H}{\delta m}(\mu, x_1) = V(x_1) + \sum_{k=2}^N k \int W^{(k)}(x_1, y)\mu^{\otimes k-1}(dy)$, i.e.

$$dX_t^\mu = \sqrt{2}dB_t^\mu - \nabla \frac{\delta H}{\delta m}(\mu, X_t^\mu)dt. \tag{1.6.29}$$

This Langevin process admits for invariant distribution $\Phi(\mu)(dx) = \frac{1}{Z_\mu} e^{-\frac{\delta H}{\delta m}(\mu,x)}dx$. The functional H admits first and second-order flat derivatives that are jointly continuous and are $\mathscr{C}^2$ in the spatial variables. H_n the Hamiltonian of the particle system and its derivatives are related to the macroscopic Hamiltonian H and its derivatives by

$$\begin{aligned} H_n(\mathbf{x}) &= nH(\mu_{\mathbf{x}}) \\ &= nH\Big(\frac{1}{n}\sum_{p=1}^n \delta_{x_p}\Big) \\ &= \sum_{p=1}^n V(x_p) + n\sum_{k=2}^N U_n(W^{(k)}); \end{aligned} \tag{1.6.30}$$

$$\nabla_{x_i} H_n(\mathbf{x}) = \mathscr{D}_m H(\mu_{\mathbf{x}}, x_i); \tag{1.6.31}$$

$$\nabla^2_{x_j x_i} H_n(\mathbf{x}) = \nabla \mathscr{D}_m H(\mu_{\mathbf{x}}, x_i)\delta_{ij} + \frac{1}{n}\mathscr{D}^2_m H(\mu_{\mathbf{x}}, x_i, x_j). \tag{1.6.32}$$

As

$$\frac{\delta^2 H}{\delta m^2}(\mu, x_1, x_2) = \sum_{k=2}^N k(k-1)\int W^{(k)}(x_1, x_2, y)\mu^{\otimes k-2}(dy), \tag{1.6.33}$$

and that for all $k \in \{2, \dots, N\}$, $\nabla_{x_1} W^{(k)}$ and $\nabla^2_{x_1 x_2} W^{(k)}$ are bounded, the functions $(\mu, x_1, x_2) \mapsto \nabla_{x_1}\frac{\delta^2 H}{\delta m^2}(\mu, x_1, x_2)$ and $(\mu, x_1, x_2) \mapsto \nabla^2_{x_1 x_2}\frac{\delta^2 H}{\delta m^2}(\mu, x_1, x_2)$ are bounded. These properties imply that Assumptions 1 and 2 of [234, Theorem 1] are true. On the other hand, Assumption 5 exactly corresponds to Assumption (iii) of proposition 1.4.2. It thus remains to establish Assumption 3 concerning uniform logarithmic Sobolev inequality for $\Phi(\mu)$ and Assumption 4 concerning uniform Poincaré inequality for the i-th conditional marginal of the invariant distribution of the particle system given by $\mu_n(d\mathbf{x}) = \frac{1}{Z_n} e^{-H_n(\mathbf{x})}d\mathbf{x}$.

Step 2: Uniform log-Sobolev inequality. For this step, one uses [59, 60] which gives some sufficient conditions on a potential U to ensure a ρ_U-log-Sobolev inequality for $\alpha_U(dx) := \frac{e^{-U(x)}}{Z_U}dx$. More precisely, such a result holds if U is $\mathscr{C}^2$ with lower bounded Hessian and satisfies the following Lyapunov condition: there are two positive constants c_1^U and c_2^U such

that

$$\forall x \in \mathbb{R}^d, \quad x \cdot \nabla U(x) \geq c_1^U |x|^2 - c_2^U. \tag{1.6.34}$$

Thus, in our setting, the uniform log-Sobolev inequality will hold if these assumptions are true uniformly in μ for $x \mapsto \frac{\delta H}{\delta m}(\mu, x)$. Let us check these conditions: since for all $k \in \{2, \dots, N\}$, $x \mapsto \nabla_{x_1} W^{(k)}(x)$ and $\nabla^2 W^{(k)}$ are bounded, for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,

$$\begin{aligned} x \cdot \nabla \frac{\delta H}{\delta m}(\mu, x) &\geq c_1 |x|^2 - c_2 + \sum_{k=2}^{N} k \int x \cdot \nabla_{x_1} W^{(k)}(x, y) \mu^{\otimes k-1}(dy) \\ &\geq c_1 |x|^2 - c_2 - \Big(\sum_{k=2}^{N} k ||\nabla_{x_1} W^{(k)}||_\infty \Big) |x|; \end{aligned} \tag{1.6.35}$$

and

$$\nabla^2 \frac{\delta H}{\delta m}(\mu, \cdot) = \nabla^2 V + \sum_{k=2}^{N} k \int \nabla_{x_1}^2 W^{(k)}(\cdot, y) \mu^{\otimes k-1}(dy) \tag{1.6.36}$$

is bounded from below. As $|x| \leq \varepsilon |x|^2 + C_\varepsilon$ for any $\varepsilon > 0$, there are two positive constants $c_1^\star$ and $c_2^\star$ such that

$$\forall (x, \mu) \in \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d), \quad x \cdot \nabla \frac{\delta H}{\delta m}(\mu, x) \geq c_1^\star |x|^2 - c_2^\star. \tag{1.6.37}$$

From what precedes, one can conclude $\Phi(\mu)$ satisfies a ρ logarithmic Sobolev inequality with constant ρ independent of μ.

Step 3: Uniform Poincaré inequality for $\mu_n^{i|-i}$. First,

$$\frac{\mu_n^{i|-i}(\cdot | \mathbf{x}^{-i})}{dx_i} = \frac{e^{-nH(\mu_{\mathbf{x}})}}{\int e^{-nH(\mu_{\mathbf{x}})} dx_i} \tag{1.6.38}$$

which involves that its log-density satisfies

$$-\nabla_i \ln \mu_n^{i|-i}(x_i | \mathbf{x}^{-i}) = \mathscr{D}_m H(\mu_{\mathbf{x}}, x_i). \tag{1.6.39}$$

Second, let us denote $\mu_{\mathbf{x}^{-i}}^{\Phi}(\cdot)$ the probability measure defined by

$$\frac{\mu_{\mathbf{x}^{-i}}^{\Phi}(\cdot)}{dx_i} := \frac{\Phi(\mu_{\mathbf{x}^{-i}})}{dx_i} = \frac{e^{-\frac{\delta H}{\delta m}(\mu_{\mathbf{x}^{-i}}, x_i)}}{\int e^{-\frac{\delta H}{\delta m}(\mu_{\mathbf{x}^{-i}}, x_i)} dx_i}. \tag{1.6.40}$$

Its log-density satisfies

$$-\nabla_i \ln \mu^{\Phi}_{\mathbf{x}^{-i}}(x_i) = \mathscr{D}_m \mathrm{H}(\mu_{\mathbf{x}^{-i}}, x_i). \tag{1.6.41}$$

We have (definition of the flat derivative)

$$\mathscr{D}_m \mathrm{H}(\mu_1, x) - \mathscr{D}_m \mathrm{H}(\mu_0, x) = \int_0^1 \int_{\mathbb{R}^d} \nabla_x \frac{\delta^2 \mathrm{H}}{\delta m^2}(t\mu_1 + (1-t)\mu_0, x, y)(\mu_1 - \mu_0)(dy)\,dt. \tag{1.6.42}$$

By boundedness property of $(\mu, x, y) \mapsto \nabla_x \frac{\delta^2 \mathrm{H}}{\delta m^2}(\mu, x, y)$, we have

$$|\mathscr{D}_m \mathrm{H}(\mu_1, x) - \mathscr{D}_m \mathrm{H}(\mu_0, x)| \le ||\nabla_x \frac{\delta^2 \mathrm{H}}{\delta m^2}||_\infty ||\mu_1 - \mu_0||_{\mathrm{TV}}. \tag{1.6.43}$$

Then, for some constant M, H verifies

$$|\mathscr{D}_m \mathrm{H}(\mu_{\mathbf{x}}, x_i) - \mathscr{D}_m \mathrm{H}(\mu_{\mathbf{x}^{-i}}, x_i)| \le \frac{\mathrm{M}}{n}; \tag{1.6.44}$$

then the derivatives can at most differ by $\frac{\mathrm{M}}{n}$. This means that, for n large, the conditional measure $\mu_n^{i|-i}(\cdot|\mathbf{x}^{-i})$ is a *weak log-Lipschitz perturbation* of the measure $\mu^{\Phi}_{\mathbf{x}^{-i}}$. Furthermore, $\mu^{\Phi}_{\mathbf{x}^{-i}}$ satisfies a ρ-log-Sobolev, a fortiori, it satisfies a ρ-Poincaré inequality. By Aida and Shigekawa perturbation theorem [3, Theorem 2.7], the uniform Poincaré inequality for $\mu_n^{i|-i}$ follows.

Conclusion. By [234, Theorem 1], we have the desired uniform log-Sobolev inequality for μ_n. ■

Lemma 1.6.1. *Let $\mu \in \mathscr{P}(\mathbb{R}^d)$. Let $f : \mathbb{R}^d \to$ be a $\mathscr{C}^1$ function with bounded derivative. Let φ be a unique solution the Poisson equation $f - \langle \Phi(\mu), f\rangle = \mathscr{L}_\mu \varphi$ where $\mathscr{L}_\mu$ is defined by (1.6.23). Then,*

$$\|\nabla \varphi\|_\infty \le [f]_1 \underline{\lambda}^{-1}$$

where $[f]_1$ denotes the Lipschitz constant of f and $\underline{\lambda}^{-1}$ is defined by (1.4.6).

Proof. It is well-known that $\varphi(x) = \int_0^{+\infty} \mathrm{P}_t^\mu f(x) - \langle \Phi(\mu), f\rangle dt$ where $(\mathrm{P}_t^\mu)_{t\ge 0}$ denotes the semi-group associated with $\mathscr{L}_\mu$ so that under adequate derivation conditions (which will be satisfied in our setting),

$$\nabla \varphi(x) = \int_0^{+\infty} \nabla \mathrm{P}_t^\mu f(x)\,dt.$$

Now,

$$\nabla P_t^\mu f(x) = \mathbb{E}[\nabla f(X_t^{\mu,x})\partial_x X_t^{\mu,x}]$$

where $(X_t^{\mu,x})$ denotes the solution starting from x of the SDE associated with $\mathscr{L}_\mu$ and $(\partial_x X_t^\mu)$ its first variation process solution to

$$dY_t = -\nabla^2\left(\frac{\delta F}{\delta m}(\mu, X_t^{\mu,x}) + V(X_t^{\mu,x})\right)Y_t dt$$

with $Y_0 = I_d$. Under the assumption (1.4.6), one easily deduces from a Gronwall argument that for any $z \in \mathbb{R}^d$,

$$|Y_t z|^2 \le e^{-2\underline{\lambda}t}|z|^2$$

which implies that

$$|\nabla P_t^\mu f(x)| \le [f]_1 \|\partial_x X_t^{\mu,x}\|_{\mathbf{op}} \le [f]_1 e^{-\underline{\lambda}t}.$$

The result follows. ■

Remark 1.6.1. In the proof of the lemma 1.6.1, we can replace $\underline{\lambda}$ by

$$\inf_{x\in\mathbb{R}^d}\left(\underline{\lambda}_{\nabla^2 V(x)} + \inf_{\mu\in\mathscr{P}(\mathbb{R}^d)}\underline{\lambda}_{\nabla^2\frac{\delta F}{\delta m}(\mu,x)}\right) \quad (\ge \underline{\lambda}). \tag{1.6.45}$$

Proof of proposition 1.4.4. As G, ∇G and $\nabla^2 G$ are bounded, given that the interaction potentials are combinations of tensor products of G, it is easy to check (**H1**), (**H3**), **VFP1**, for all $i = 1,2$, $\gamma_i < +\infty$ and $\mathscr{P}_2(\mathbb{R}^d) \subset \mathscr{P}_G(\mathbb{R}^d) = \mathscr{P}(\mathbb{R}^d)$. As for (**H4**) and (**H5**), they are significantly more difficult to establish. Let's go prove them.

For (**H4**), we will proceed as in the proof of proposition 1.4.2: we will establish assumptions of [234, Theorem 1].

Step 1: Regularity properties. For any $\mu \in \mathscr{P}(\mathbb{R}^d)$, consider the standard Langevin process $(X_t^\mu)_{t\ge 0}$ of Hamiltonian $\frac{\delta H}{\delta m}(\mu,\cdot) = V + Q'(\int G d\mu)G$, i.e.

$$dX_t^\mu = \sqrt{2}dB_t^\mu - \nabla\frac{\delta H}{\delta m}(\mu, X_t^\mu)dt. \tag{1.6.46}$$

This Langevin process admits for invariant distribution $\Phi(\mu)(dx) = \frac{1}{Z_\mu}e^{-\frac{\delta H}{\delta m}(\mu,x)}dx$. The functional H admits first and second-order flat derivatives that are jointly continuous and are $\mathscr{C}^2$ in the spatial variables. H_n the Hamiltonian of the particle system and its derivatives are

related to the macroscopic Hamiltonian H and its derivatives by

$$\begin{aligned} \mathrm{H}_n(\mathbf{x}) &= n\mathrm{H}(\mu_{\mathbf{x}}) \\ &= n\mathrm{H}\Big(\frac{1}{n}\sum_{p=1}^{n}\delta_{x_p}\Big) \\ &= \sum_{p=1}^{n}\mathrm{V}(x_p) + n\mathrm{Q}\Big(\frac{1}{n}\sum_{p=1}^{n}\mathrm{G}(x_p)\Big); \end{aligned} \tag{1.6.47}$$

$$\nabla_{x_i}\mathrm{H}_n(\mathbf{x}) = \mathscr{D}_m\mathrm{H}(\mu_{\mathbf{x}}, x_i); \tag{1.6.48}$$

$$\nabla^2_{x_j x_i}\mathrm{H}_n(\mathbf{x}) = \nabla\mathscr{D}_m\mathrm{H}(\mu_{\mathbf{x}}, x_i)\delta_{ij} + \frac{1}{n}\mathscr{D}^2_m\mathrm{H}(\mu_{\mathbf{x}}, x_i, x_j). \tag{1.6.49}$$

As

$$\frac{\delta^2\mathrm{H}}{\delta m^2}(\mu, x_1, x_2) = \mathrm{Q}''\Big(\int \mathrm{G}d\mu\Big)\mathrm{G}(x_1)\mathrm{G}(x_2), \tag{1.6.50}$$

and G, ∇G and ∇^2G are bounded, the functions $(\mu, x_1, x_2) \mapsto \nabla_{x_1}\frac{\delta^2\mathrm{H}}{\delta m^2}(\mu, x_1, x_2)$ and $(\mu, x_1, x_2) \mapsto \nabla^2_{x_1 x_2}\frac{\delta^2\mathrm{H}}{\delta m^2}(\mu, x_1, x_2)$ are bounded. Then, to apply [234, Theorem 1], the boundedness assumption on second order flat derivative and the convexity in flat interpolation sense assumption ($\mathrm{Q}'' \geq 0$) on $\mathscr{E}_{\mathrm{W}}$ hold true. It remains to establish the assumption 3 concerning uniform logarithmic Sobolev inequality for $\Phi(\mu)$ and the assumption 4 concerning uniform Poincaré inequality for the i-th conditional marginal of the invariant distribution of the particle system given by $\mu_n(d\mathbf{x}) = \frac{1}{\mathrm{Z}_n}e^{-\mathrm{H}_n(\mathbf{x})}d\mathbf{x}$.

Step 2: Uniform log-Sobolev inequality. For this step, one uses [59, 60] which gives some sufficient conditions on a potential U to ensure a ρ_{U}-log-Sobolev inequality for $\alpha_{\mathrm{U}}(dx) := \frac{e^{-\mathrm{U}(x)}}{\mathrm{Z}_{\mathrm{U}}}dx$. More precisely, such a result holds if U is $\mathscr{C}^2$ with lower bounded Hessian and satisfies the following Lyapunov condition: there are two positive constants c_1^{U} and c_2^{U} such that

$$\forall x \in \mathbb{R}^d, \quad x\cdot\nabla\mathrm{U}(x) \geq c_1^{\mathrm{U}}|x|^2 - c_2^{\mathrm{U}}. \tag{1.6.51}$$

Thus, in our setting, the uniform log-Sobolev inequality will hold if these assumptions are true uniformly in μ for $x \mapsto \frac{\delta\mathrm{H}}{\delta m}(\mu, x)$. Let us check these conditions: since $\gamma_1 < +\infty$ ($|\gamma_3| \leq \gamma_1$),

for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, we have

$$\begin{aligned} \nabla^2 \frac{\delta H}{\delta m}(\mu, \cdot) &\geq \nabla^2 V + \gamma_3 \nabla^2 G \qquad (1.6.52) \\ &\geq \nabla^2 V - \gamma_1 \nabla^2 G; \\ x \cdot \nabla \frac{\delta H}{\delta m}(\mu, x) &\geq c_1 |x|^2 - c_2 + \gamma_3 x \cdot \nabla G(x) \qquad (1.6.53) \\ &\geq c_1 |x|^2 - c_2 - \gamma_1 ||\nabla G||_\infty |x|. \end{aligned}$$

As $|x| \leq \varepsilon |x|^2 + C_\varepsilon$ for any $\varepsilon > 0$, there are two positive constants $c_1^\star$ and $c_2^\star$ such that

$$\forall (x, \mu) \in \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d), \quad x \cdot \nabla \frac{\delta H}{\delta m}(\mu, x) \geq c_1^\star |x|^2 - c_2^\star. \qquad (1.6.54)$$

From what precedes, one can conclude $\Phi(\mu)$ satisfies a ρ logarithmic Sobolev inequality with constant ρ independent of μ.

Step 3: Uniform Poincaré inequality for $\mu_n^{i|-i}$. First,

$$\frac{\mu_n^{i|-i}(\cdot|\mathbf{x}^{-i})}{dx_i} = \frac{e^{-nH(\mu_\mathbf{x})}}{\int e^{-nH(\mu_\mathbf{x})} dx_i} \qquad (1.6.55)$$

which involves that its log-density satisfies

$$-\nabla_i \ln \mu_n^{i|-i}(x_i|\mathbf{x}^{-i}) = \mathscr{D}_m H(\mu_\mathbf{x}, x_i). \qquad (1.6.56)$$

Second, let us denote $\mu^\Phi_{\mathbf{x}^{-i}}(\cdot)$ the probability measure defined by

$$\frac{\mu^\Phi_{\mathbf{x}^{-i}}(\cdot)}{dx_i} := \frac{\Phi(\mu_{\mathbf{x}^{-i}})}{dx_i} = \frac{e^{-\frac{\delta H}{\delta m}(\mu_{\mathbf{x}^{-i}}, x_i)}}{\int e^{-\frac{\delta H}{\delta m}(\mu_{\mathbf{x}^{-i}}, x_i)} dx_i}. \qquad (1.6.57)$$

Its log-density satisfies

$$-\nabla_i \ln \mu^\Phi_{\mathbf{x}^{-i}}(x_i) = \mathscr{D}_m H(\mu_{\mathbf{x}^{-i}}, x_i). \qquad (1.6.58)$$

We have (definition of the flat derivative)

$$\mathscr{D}_m H(\mu_1, x) - \mathscr{D}_m H(\mu_0, x) = \int_0^1 \int_{\mathbb{R}^d} \nabla_x \frac{\delta^2 H}{\delta m^2}(t\mu_1 + (1-t)\mu_0, x, y)(\mu_1 - \mu_0)(dy)dt. \qquad (1.6.59)$$

By boundedness property of $(\mu, x, y) \mapsto \nabla_x \frac{\delta^2 H}{\delta m^2}(\mu, x, y)$, we have

$$|\mathscr{D}_m H(\mu_1, x) - \mathscr{D}_m H(\mu_0, x)| \le ||\nabla_x \frac{\delta^2 H}{\delta m^2}||_\infty ||\mu_1 - \mu_0||_{TV}. \tag{1.6.60}$$

Then, for some constant M, H verifies

$$|\mathscr{D}_m H(\mu_{\mathbf{x}}, x_i) - \mathscr{D}_m H(\mu_{\mathbf{x}^{-i}}, x_i)| \le \frac{M}{n}; \tag{1.6.61}$$

then the derivatives can at most differ by $\frac{M}{n}$. This means that, for n large, the conditional measure $\mu_n^{i|-i}(\cdot|\mathbf{x}^{-i})$ is a *weak log-Lipschitz perturbation* of the measure $\mu^{\Phi}_{\mathbf{x}^{-i}}$. Furthermore, $\mu^{\Phi}_{\mathbf{x}^{-i}}$ satisfies a ρ-log-Sobolev, a fortiori, it satisfies a ρ-Poincaré inequality. By Aida and Shigekawa perturbation theorem [3, Theorem 2.7], the uniform Poincaré inequality for $\mu_n^{i|-i}$ follows.

Conclusion. By [234, Theorem 1], we have the desired uniform log-Sobolev inequality for μ_n.

For (**H5**), we follow the same strategy as in the proof of proposition 1.4.1 (see (1.6.21)): for $\nu_0, \nu_1 \in \mathscr{P}_2(\mathbb{R}^d)$ and a 1-Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$, we write $\nu_t = (1-t)\nu_0 + t\nu_1$. First,

$$\langle \Phi(\nu_t), f\rangle = \frac{1}{Z_{\nu_t}} \int_{\mathbb{R}^d} f(x) e^{-V(x) - Q'(\int G(y)\nu_t(dy))G(x)} dx \quad \text{and} \quad \frac{d}{dt}\langle \Phi(\nu_t), f\rangle = \mathrm{Cov}_{\Phi(\nu_t)}(f, g_t) \tag{1.6.62}$$

with

$$g_t(x) = -\partial_t\Big(Q'\Big(\int G(y)\nu_t(dy)\Big)G(x)\Big) = -G(x)Q''\left(\int G(y)\nu_t(dy)\right)\int G(y)(\nu_1 - \nu_0)(dy).$$

Following carefully the proof of proposition 1.4.1, we get

$$|\mathrm{Cov}_{\Phi(\nu_t)}(f, g_t)| \le \|\nabla\varphi_t\|_{L^2(\Phi(\nu_t))}\|\nabla g_t\|_{L^2(\Phi(\nu_t))} \le \frac{1}{\underline{\lambda}^\star}\|\nabla g_t\|_{L^2(\Phi(\nu_t))} \tag{1.6.63}$$

with $\underline{\lambda}^\star = \inf_{x\in\mathbb{R}^d}\left(\underline{\lambda}_{\nabla^2 V(x)} + \gamma_3 \underline{\lambda}_{\nabla^2 G(x)}\right) > 0$ under the assumptions of proposition 1.4.4. Now, since

$$\nabla g_t = -Q''\Big(\int G(y)\nu_t(dy)\Big)\int G(y)(\nu_1 - \nu_0)(dy)\nabla G,$$

since $\gamma_2 < +\infty$ and ∇G is bounded, by the Kantorovitch-Rubinstein duality relation, we have

$$\|\nabla g_t\|_{L^2(\Phi(\nu_t))} \le \|\nabla g_t\|_\infty = \left|Q''\Big(\int G(y)\nu_t(dy)\Big)\right| \cdot \left|\int G(y)(\nu_1-\nu_0)(dy)\right| ||\nabla G||_\infty \le \gamma_2 ||\nabla G||_\infty^2 \mathscr{W}_1(\nu_0,\nu_1)$$

and by (1.6.63), we get for any smooth 1-Lipschitz function

$$|\langle \Phi(\nu_1), f\rangle - \langle \Phi(\nu_0), f\rangle| \le \int_0^1 |\frac{d}{dt}\langle \Phi(\nu_t), f\rangle| dt \le \frac{\gamma_2 ||\nabla G||_\infty^2}{\underline{\lambda}^\star} \mathscr{W}_1(\nu_0,\nu_1).$$

As $\frac{\gamma_2}{\underline{\lambda}^\star}||\nabla G||_\infty^2 < 1$, by a density argument and the Kantorovitch-Rubinstein duality relation, it follows that Φ is a contraction on $(\mathscr{P}_2(\mathbb{R}^d), \mathscr{W}_1)$.

■

1.7 Appendix and additional proofs

1.7.1 Propagation of chaos

Theorem 1.7.1 (Moment control). *Let $m \in \mathbb{R}_{\ge 1}$, $||\cdot||_2$ be the standard Euclidean norm on $\mathbb{R}^d$ and the continuous function*

$$\Theta_t : \quad \mathbb{R} \longrightarrow [0,+\infty[\qquad (1.7.1)$$
$$\theta \longmapsto \begin{cases} \frac{1-e^{-2\theta t}}{\theta} & \text{if} \quad \theta \neq 0; \\ 2t & \text{if} \quad \theta = 0. \end{cases}$$

Suppose that there exist pairs of constants $(\theta,\vartheta),(\theta_2,\vartheta_2),,\ldots,(\theta_N,\vartheta_N)$ such that for all $(\mu,x) \in \mathscr{P}(\mathbb{R}^d)\times\mathbb{R}^d$,

$$x\cdot\nabla V(x) \ge \theta||x||^2 + \vartheta; \qquad (1.7.2)$$
$$\forall k \in \{2,\ldots,N\}, \quad x\cdot\nabla_{x_1} W^{(k)} * \mu^{\otimes k-1}(x) \ge \theta_k||x||^2 + \vartheta_k. \qquad (1.7.3)$$

Let $\overline{\omega} := \theta + \sum_{k=2}^N k\theta_k$ and $\overline{\vartheta} := \vartheta + \sum_{k=2}^N k\vartheta_k$. We have

$$\forall \mu_0 \in \mathscr{P}_{2m}(\mathbb{R}^d) \quad \forall t \ge 0, \quad \left|\left|||\cdot||_2^2\right|\right|_{L^m(\mu_t)} \le \left|\left|||\cdot||_2^2\right|\right|_{L^m(\mu_0)} e^{-2\overline{\omega}t} \qquad (1.7.4)$$
$$+ \frac{2m+d-2-\overline{\vartheta}+|2m+d-2-\overline{\vartheta}|}{2}\Theta_t(\overline{\omega}).$$

In particular, if $\overline{\omega} > 0$ *or* $\overline{\omega} = 0$ *and* $2m + d - 2 - \overline{\vartheta} \le 0$, *we have uniformity in time:*

$$\sup_{t\ge 0} \left|\left|\,||\cdot||_2^2\,\right|\right|_{\mathrm{L}^m(\mu_t)} \le \left|\left|\,||\cdot||_2^2\,\right|\right|_{\mathrm{L}^m(\mu_0)} + \begin{cases} \frac{2m+d-2-\overline{\vartheta}+|2m+d-2-\overline{\vartheta}|}{2\overline{\omega}} & \text{if} \quad \overline{\omega} > 0; \\ 0 & \text{if} \quad \overline{\omega} = 0 \quad \text{and} \quad 2m+d-2-\overline{\vartheta} \le 0. \end{cases} \tag{1.7.5}$$

Proof of theorem 1.7.1. By the transfer formula, for any function $\psi \in \mathrm{L}^1(\mu_t)$ or of constant sign, we have

$$\mathbb{E}[\psi(\mathrm{X}_t)] = \int_{\mathbb{R}^d} \psi(x)\mu_t(dx). \tag{1.7.6}$$

For all $m \ge 1$, by Itô's formula, we have

$$\frac{d}{dt}\mathbb{E}[||\mathrm{X}_t||^{2m}] = \mathbb{E}[\mathscr{L}_{\mu_t}||\cdot||_2^{2m}(\mathrm{X}_t)]. \tag{1.7.7}$$

On the other hand, we have

$$\nabla||\cdot||_2^{2m}(x) = 2m||x||^{2(m-1)}x; \tag{1.7.8}$$

$$\begin{aligned} \Delta||\cdot||_2^{2m}(x) &= \nabla\cdot\nabla||\cdot||_2^{2m}(x) \\ &= 2m\sum_{j=1}^d \partial_{x_j}\left(|x||^{2(m-1)}x_j\right) \\ &= 2m\sum_{j=1}^d \left(||x||^{2(m-1)} + 2(m-1)||x||^{2(m-2))}x_j^2\right) \\ &= 2md||x||^{2(m-1)} + 4m(m-1)||x||^{2(m-2)}||x||^2 \\ &= 2m(2m+d-2)||x||^{2(m-1)}. \end{aligned} \tag{1.7.9}$$

We deduce that

$$\begin{aligned} \mathscr{L}_\mu||\cdot||_2^{2m}(x) &= 2m(2m+d-2)||x||^{2(m-1)} - 2m||x||^{2(m-1)}x\cdot\left(\nabla\frac{\delta \mathrm{F}}{\delta m}(\mu,x) + \nabla \mathrm{V}\right) \\ &= 2m||x||^{2(m-1)}\left(2m+d-2-x\cdot\nabla \mathrm{V}(x) - \sum_{k=2}^{\mathrm{N}} kx\cdot\nabla_{x_1}\mathrm{W}^{(k)} * \mu^{\otimes k-1}(x)\right). \end{aligned} \tag{1.7.10}$$

By setting $S_m(t) := \mathbb{E}[||X_t||^{2m}]$, we have

$$S'_m(t) \le 2m\Big(2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k\Big)S_{m-1}(t) - 2m\Big(\theta+\sum_{k=2}^{N} k\theta_k\Big)S_m(t). \tag{1.7.11}$$

As for a finite measure μ, we have

$$\forall p \le q \quad \forall f \in L^q(\mu), \quad ||\psi||_{L^p(\mu)} := \Big(\int |\psi|^p d\mu\Big)^{\frac{1}{p}} \le \Big(\int 1 d\mu\Big)^{\frac{1}{p}-\frac{1}{q}} ||\psi||_{L^q(\mu)}, \tag{1.7.12}$$

we have

$$S_{m-1}(t) \le S_m(t)^{\frac{m-1}{m}}. \tag{1.7.13}$$

Case 1: Using L^p injection. If $\mathbf{2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k > 0}$, we have

$$S'_m(t) \le 2m\Big(2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k\Big)S_m(t)^{\frac{m-1}{m}} - 2m\Big(\theta+\sum_{k=2}^{N} k\theta_k\Big)S_m(t). \tag{1.7.14}$$

We deduce that

$$m\Big(S_m(t)^{\frac{1}{m}}\Big)' \le 2m\Big(2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k\Big) - 2m\Big(\theta+\sum_{k=2}^{N} k\theta_k\Big)S_m(t)^{\frac{1}{m}}. \tag{1.7.15}$$

By setting $z_m(t) := S_m(t)^{\frac{1}{m}}$, we have

$$z'_m(t) \le 2\Big(2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k\Big) - 2\Big(\theta+\sum_{k=2}^{N} k\theta_k\Big)z_m(t). \tag{1.7.16}$$

And by Gronwall's lemma, if $\theta+\sum_{k=2}^{N} k\theta_k \neq 0$, we have

$$\begin{aligned} z_m(t) - \frac{2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k}{\theta+\sum_{k=2}^{N} k\theta_k} \le \Big(z_m(0) \\ - \frac{2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k)}{\theta+\sum_{k=2}^{N} k\theta_k}\Big) e^{-2t\Big(\theta+\sum_{k=2}^{N} k\theta_k\Big)}. \end{aligned} \tag{1.7.17}$$

If $\theta+\sum_{k=2}^{N} k\theta_k = 0$, by integration, we deduce that

$$z_m(t) \le z_m(0) + 2\Big(2m+d-2-\vartheta-\sum_{k=2}^{N} k\vartheta_k\Big)t \tag{1.7.18}$$

In conclusion of this disjunction of cases, whether $\theta + \sum_{\mathbf{k=2}}^{\mathbf{N}} \mathbf{k}\theta_{\mathbf{k}} = \mathbf{0}$ is zero or not, we have

$$z_m(t) \le z_m(0) e^{-2t\left(\theta + \sum_{k=2}^{N} k\theta_k\right)} + \left(2m + d - 2 - \vartheta - \sum_{k=2}^{N} k\vartheta_k\right)\Theta_t\left(\theta + \sum_{k=2}^{N} k\theta_k\right); \tag{1.7.19}$$

$$\Theta_t : \quad \mathbb{R} \longrightarrow [0, +\infty[$$
$$\theta \longmapsto \begin{cases} \frac{1-e^{-2\theta t}}{\theta} & \text{if} \quad \theta \neq 0; \\ 2t & \text{if} \quad \theta = 0. \end{cases}$$

Case 2: Direct increase. If $\mathbf{2m + d - 2 - \vartheta - \sum_{k=2}^{N} k\vartheta_k \le 0}$, we have

$$S_m'(t) \le -2m\left(\theta + \sum_{k=2}^{N} k\theta_k\right) S_m(t). \tag{1.7.20}$$

And by Gronwall's lemma, we have

$$S_m(t) \le S_m(0) e^{-2tm\left(\theta + \sum_{k=2}^{N} k\theta_k\right)}. \tag{1.7.21}$$

In conclusion, in all these cases, if $\mu_0 \in \mathscr{P}_{2m}(\mathbb{R}^d)$, we have

$$\forall T \ge 0, \quad \sup_{0 \le t \le T} \mathbb{E}[||X_t||^{2m}] \le M(m, \mu_0, T). \tag{1.7.22}$$

As for uniformity in time, it is ensured if one of the following conditions is verified

— $\theta + \sum_{k=2}^{N} k\theta_k > 0$;

— $\theta + \sum_{k=2}^{N} k\theta_k = 0$ and $2m + d - 2 - \vartheta - \sum_{k=2}^{N} k\vartheta_k \le 0$.

■

Remark 1.7.1. Suppose that there exist pairs of constants $(\theta, \vartheta), (\theta_2, \vartheta_2), , \ldots, (\theta_N, \vartheta_N)$ such that for all $(\mu, x) \in \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d$,

$$x \cdot \nabla V(x) \ge \theta ||x||^2 + \vartheta; \tag{1.7.23}$$
$$\forall k \in \{2, \ldots, N\}, \quad x \cdot \nabla_{x_1} W^{(k)} * \mu^{\otimes k-1}(x) \ge \theta_k ||x||^2 + \vartheta_k. \tag{1.7.24}$$

If we add $\nabla V(0) = 0$ and $\nabla_{x_1} W^{(k)}(0, \cdot) = 0$ to the reduced Hessian conditions, we obtain the following hypotheses above. Indeed, if there exists $\beta \in \mathbb{R}$ such that $\nabla^2 V \ge -\beta \mathbb{I}_d$ and for all

$k \in \{2,\dots,N\}$, there exists $\beta_k \in \mathbb{R}$ such that $\nabla^2_{x_1} W^{(k)} \geq -\beta_k \mathbb{I}_d$, we have in particular

$$\langle \nabla V(x) - \nabla V(y), x-y\rangle \geq -\beta ||x-y||^2; \tag{1.7.25}$$

$$\langle \nabla_{x_1} W^{(k)}(x,\cdot) - \nabla_{x_1} W^{(k)}(y,\cdot), x-y\rangle \geq -\beta_k ||x-y||^2. \tag{1.7.26}$$

In this case, $\theta = -\beta$, $\vartheta = 0$, for all $k \in \{2,\dots,N\}$, $\theta_k = -\beta_k$ and $\vartheta_k = 0$.

Proof of theorem 1.5.1. To show the result and for greater clarity of proof, we proceed in five steps described below.

Step 1: Itô's formula and drift division. By setting $G = \frac{2V}{(N-1)(N+2)}$, we have

$$\begin{aligned} dX_t^{(n),p} &= \sqrt{2} dB_t^{(n),p} - \sum_{k=2}^{N} \frac{k}{\binom{n-1}{k-1}} \sum_{\substack{1\leq i_1 < i_2 < \dots < i_{k-1} \leq n \\ \forall j,\ i_j \neq p}} \Big(\nabla G(X_t^{(n),p}) \\ &\quad + \nabla_{x_1} W^{(k)}(X_t^{(n),p}, X_t^{(n),i_1}, \dots, X_t^{(n),i_{k-1}}) \Big) dt \\ &= \sqrt{2} dB_t^{(n),p} - \sum_{k=2}^{N} \frac{k}{\binom{n-1}{k-1}} \sum_{\substack{1\leq i_1 < i_2 < \dots < i_{k-1} \leq n \\ \forall j,\ i_j \neq p}} \int \Big(\nabla G(X_t^{(n),p}) \\ &\quad + \nabla_{x_1} W^{(k)}(X_t^{(n),p}, X_t^{(n),i_1}, \dots, X_t^{(n),i_{k-1}}) \Big) d\mu^{\otimes(k-1)} dt. \end{aligned} \tag{1.7.27}$$

$$\begin{aligned} dX_t^{(p)} &= \sqrt{2} dB_t^{(n),p} - \sum_{k=2}^{N} k \int_{(\mathbb{R}^d)^{k-1}} \Big(\nabla G(X_t^{(p)}) + \nabla_{x_1} W^{(k)}(X_t^{(p)}, y) \Big) \mathbb{P}_{X_t}^{\otimes(k-1)}(dy) dt \\ &= \sqrt{2} dB_t^{(n),p} - \sum_{k=2}^{N} \frac{k}{\binom{n-1}{k-1}} \sum_{\substack{1\leq i_1 < i_2 < \dots < i_{k-1} \leq n \\ \forall j,\ i_j \neq p}} \int_{(\mathbb{R}^d)^{k-1}} \Big(\nabla G(X_t^{(p)}) \\ &\quad + \nabla_{x_1} W^{(k)}(X_t^{(p)}, y) \Big) \mathbb{P}_{X_t}^{\otimes(k-1)}(dy) dt. \end{aligned} \tag{1.7.28}$$

So

$$\begin{aligned} X_t^{(n),p} - X_t^{(p)} &= X_r^{(n),p} - X_r^{(p)} + \int_r^t \sum_{k=2}^{N} \frac{k}{\binom{n-1}{k-1}} \sum_{\substack{1\leq i_1 < i_2 < \dots < i_{k-1} \leq n \\ \forall j,\ i_j \neq p}} \\ &\quad \int_{(\mathbb{R}^d)^{k-1}} \Big(\nabla G(X_s^{(p)}) - \nabla G(X_s^{(n),p}) + \nabla_{x_1} W^{(k)}(X_s^{(p)}, y) \\ &\quad - \nabla_{x_1} W^{(k)}(X_s^{(n),p}, X_s^{(n),i_1}, \dots, X_s^{(n),i_{k-1}}) \Big) \mathbb{P}_{X_s}^{\otimes(k-1)}(dy) ds. \end{aligned} \tag{1.7.29}$$

By Itô's formula, we have

$$\sum_{p=1}^{n}||X_t^{(n),p}-X_t^{(p)}||^2=\sum_{p=1}^{n}||X_r^{(n),p}-X_r^{(p)}||^2 \tag{1.7.30}$$
$$-2\sum_{p=1}^{n}\int_r^t\sum_{k=2}^{N}\frac{k}{\binom{n-1}{k-1}}\sum_{1\le i_1<\underset{\forall j,}{i_2}<\ldots<\underset{i_j\neq p}{i_{k-1}}\le n}$$
$$\int_{(\mathbb{R}^d)^{k-1}}\Big\langle\nabla G(X_s^{(n),p})-\nabla G(X_s^{(p)})-\nabla_{x_1}W^{(k)}(X_s^{(p)},y)$$
$$+\nabla_{x_1}W^{(k)}(X_s^{(n),p},X_s^{(n),i_1},\ldots,X_s^{(n),i_{k-1}}),X_s^{(n),p}-X_s^{(p)}\Big\rangle\mathbb{P}_{X_s}^{\otimes(k-1)}(dy)ds.$$

By setting $\mu_t=\mathbb{P}_{X_t}$ and $\nabla_{x_1}W^{(k)}*\mu_s^{\otimes(k-1)}:=\int\nabla_{x_1}W^{(k)}(\cdot,y)\mu_s^{\otimes k-1}(dy)$, we have

$$\sum_{p=1}^{n}||X_t^{(n),p}-X_t^{(p)}||^2=\sum_{p=1}^{n}||X_r^{(n),p}-X_r^{(p)}||^2 \tag{1.7.31}$$
$$-2\sum_{p=1}^{n}\int_r^t\Big\langle\nabla V(X_s^{(n),p})-\nabla V(X_s^{(p)}),X_s^{(n),p}-X_s^{(p)}\Big\rangle ds$$
$$-2\sum_{p=1}^{n}\int_r^t\sum_{k=2}^{N}\frac{k}{\binom{n-1}{k-1}}\sum_{1\le i_1<\underset{\forall j,}{i_2}<\ldots<\underset{i_j\neq p}{i_{k-1}}\le n}\Big\langle\nabla_{x_1}W^{(k)}(X_s^{(n),p},X_s^{(n),i_1},\ldots,X_s^{(n),i_{k-1}})$$
$$-\nabla_{x_1}W^{(k)}*\mu_s^{\otimes(k-1)}(X_s^{(p)}),X_s^{(n),p}-X_s^{(p)}\Big\rangle ds.$$

Let

$$\rho^{(1)}_{i_1,\ldots,i_{k-1}}(s):=\Big\langle\nabla_{x_1}W^{(k)}(X_s^{(n),p},X_s^{(n),i_1},\ldots,X_s^{(n),i_{k-1}}) \tag{1.7.32}$$
$$-\nabla_{x_1}W^{(k)}*\mu_s^{\otimes(k-1)}(X_s^{(p)}),X_s^{(n),p}-X_s^{(p)}\Big\rangle$$
$$=\rho^{(2)}_{i_1,\ldots,i_{k-1}}(s)+\rho^{(3)}_{i_1,\ldots,i_{k-1}}(s);$$
$$\rho^{(2)}_{i_1,\ldots,i_{k-1}}(s):=\Big\langle\nabla_{x_1}W^{(k)}(X_s^{(n),p},X_s^{(n),i_1},\ldots,X_s^{(n),i_{k-1}}) \tag{1.7.33}$$
$$-\nabla_{x_1}W^{(k)}(X_s^{(p)},X_s^{(i_1)},\ldots,X_s^{(i_{k-1})}),X_s^{(n),p}-X_s^{(p)}\Big\rangle;$$
$$\rho^{(3)}_{i_1,\ldots,i_{k-1}}(s):=\Big\langle\nabla_{x_1}W^{(k)}(X_s^{(p)},X_s^{(i_1)},\ldots,X_s^{(i_{k-1})})-\nabla_{x_1}W^{(k)}*\mu_s^{\otimes(k-1)}(X_s^{(p)}),X_s^{(n),p}-X_s^{(p)}\Big\rangle. \tag{1.7.34}$$

Step 2: Control of the confinement term. As the Hessian matrix of the confinement poten-

tial is bounded from below, we have

$$-\sum_{p=1}^{n}\int_r^t \Big\langle \nabla V(X_s^{(n),p}) - \nabla V(X_s^{(p)}), X_s^{(n),p} - X_s^{(p)} \Big\rangle ds \le \beta \int_r^t \sum_{p=1}^{n} ||X_s^{(n),p} - X_s^{(p)}||^2 ds. \tag{1.7.35}$$

Step 3: Control of the interaction term in $\rho^{(2)}$. As the Hessian matrices of the interaction potentials in the first coordinate are uniformly bounded from below, we have

$$\forall k \in \{2,\dots,N\}, \quad -\rho^{(2)}_{i_1,\dots,i_{k-1}}(s) \le \beta_k ||X_s^{(n),p} - X_s^{(p)}||^2. \tag{1.7.36}$$

Step 4: Control of the interaction term in $\rho^{(3)}$. By Cauchy-Schwarz inequality, we have

$$-\mathbb{E}\sum_{\substack{1\le i_1<i_2<\dots<i_{k-1}\le n \\ \forall j,\ i_j\neq p}} \rho^{(3)}_{i_1,\dots,i_{k-1}}(s) \le \sqrt{\zeta_{i_1,\dots,i_{k-1}}(s)\mathbb{E}||X_s^{(n),p} - X_s^{(p)}||^2}; \tag{1.7.37}$$

$$\begin{aligned}\zeta_{i_1,\dots,i_{k-1}}(s) := \mathbb{E}\Big|\Big| &\sum_{\substack{1\le i_1<i_2<\dots<i_{k-1}\le n \\ \forall j,\ i_j\neq p}} \nabla_{x_1} W^{(k)}(X_s^{(p)}, X_s^{(i_1)}, \dots, X_s^{(i_{k-1})}) \\ &- \nabla_{x_1} W^{(k)} * \mu_s^{\otimes(k-1)}(X_s^{(p)})\Big|\Big|^2 \\ &= \sum_{\substack{1\le i_1<i_2<\dots<i_{k-1}\le n \\ \forall j,\ i_j\neq p}} \mathbb{E}||\xi^{(p)}_{i_1,\dots,i_{k-1}}(s)||^2 \\ &+ 2\sum \mathbb{E}\langle \xi^{(p))}_{i_1,\dots,i_{k-1}}(s), \xi^{(p)}_{j_1,\dots,j_{k-1}}(s)\rangle;\end{aligned} \tag{1.7.38}$$

$$\begin{aligned}\xi^{(p)}_{i_1,\dots,i_{k-1}}(s) := &\nabla_{x_1} W^{(k)}(X_s^{(p)}, X_s^{(i_1)}, \dots, X_s^{(i_{k-1})}) \\ &- \nabla_{x_1} W^{(k)} * \mu_s^{\otimes(k-1)}(X_s^{(p)}).\end{aligned} \tag{1.7.39}$$

As the $X_s^{(p)}$ are independent copies of X_s with distribution μ_s, we have

$$\mathbb{E}\langle \xi^{(p)}_{i_1,\dots,i_{k-1}}(s), \xi^{(p)}_{j_1,\dots,j_{k-1}}(s)\rangle = 0. \tag{1.7.40}$$

Moreover, as the McKean-Vlasov flow admits bounded moments (see theorem 1.7.1), we

have

$$\begin{aligned}\mathbb{E}||\xi^{(p)}_{i_1,\dots,i_{k-1}}(s)||^2 &\le \mathbb{E}||\nabla_{x_1} W^{(k)}(X^{(p)}_s, X^{(i_1)}_s, \dots, X^{(i_{k-1})}_s)||^2 \\ &\le \Omega(\mathbb{E}||X^{(p)}_s||^{2m} + \mathbb{E}||X^{(i_1)}_s||^{2m} + \dots + \mathbb{E}||X^{(i_{k-1})}_s||^{2m}) \\ &= k\Omega\mathbb{E}[||X_s||^{2m}] \\ &\le k\Omega M_{2m}(T).\end{aligned} \tag{1.7.41}$$

It follows that

$$-\mathbb{E}\sum_{\substack{1\le i_1<i_2<\dots<i_{k-1}\le n \\ \forall j,\ i_j\neq p}} \rho^{(3)}_{i_1,\dots,i_{k-1}}(s) \le \sqrt{\binom{n-1}{k-1} k\Omega M_{2m}(T)}\sqrt{\mathbb{E}||X^{(n),p}_s - X^{(p)}_s||^2}. \tag{1.7.42}$$

Step 5: Gronwall's lemma. For all $q \in \{1,\dots,n\}$, let $\gamma_q(t) = \mathbb{E}||X^{(n),q}_t - X^{(q)}_t||^2$. By taking the expectation in equation eq. (1.7.31), by the previous controls and the exchangeability of the marginals of the particle system, we obtain

$$\begin{aligned}n\gamma_q(t) \le n\gamma_q(r) &+ 2n\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\int_r^t \gamma_q(s)ds \\ &+ 2n\sqrt{\Omega M_{2m}(T)}\Big(\sum_{k=2}^{N}\frac{k^{\frac{3}{2}}}{\sqrt{\binom{n-1}{k-1}}}\Big)\int_r^t \sqrt{\gamma_q(s)}ds.\end{aligned} \tag{1.7.43}$$

As

$$\binom{n-1}{k-1} = \frac{k}{n}\binom{n}{k}; \tag{1.7.44}$$

$$\Big(\frac{n}{k}\Big)^k \le \binom{n}{k} < e^k\Big(\frac{n}{k}\Big)^k; \tag{1.7.45}$$

we have

$$\sum_{k=2}^{N}\frac{k^{\frac{3}{2}}}{\sqrt{\binom{n-1}{k-1}}} \le \sum_{k=2}^{N} k^{\frac{3}{2}}\Big(\frac{k}{n}\Big)^{\frac{k-1}{2}} = \sum_{k=2}^{N}\frac{k^{\frac{k+2}{2}}}{n^{\frac{k-1}{2}}} \le (N-1)\frac{N^{\frac{N+2}{2}}}{\sqrt{n}}. \tag{1.7.46}$$

By setting $\kappa = (N-1)\sqrt{\Omega M_{2m}(T)N^{N+2}}$, we have

$$\gamma_q(t) \le \gamma_q(r) + 2\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\int_r^t \gamma_q(s)ds + 2\frac{\kappa}{\sqrt{n}}\int_r^t \sqrt{\gamma_q(s)}ds. \tag{1.7.47}$$

Or equivalently

$$\frac{\gamma_q(t)-\gamma_q(r)}{t-r} \le 2\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\frac{\int_r^t \gamma_q(s)ds}{t-r} + 2\frac{\kappa}{\sqrt{n}}\frac{\int_r^t \sqrt{\gamma_q(s)}ds}{t-r}. \tag{1.7.48}$$

And by passing to the limit, we obtain

$$\gamma_q'(t) \le 2\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\gamma_q(t) + 2\frac{\kappa}{\sqrt{n}}\sqrt{\gamma_q(t)}. \tag{1.7.49}$$

If $\omega := \beta + \sum_{k=2}^{N} k\beta_k = 0$, we have ($\gamma_q(0) = 0$)

$$\Big(\sqrt{\gamma_q(t)}\Big)' = \frac{\gamma_q'(t)}{2\sqrt{\gamma_q(t)}} \le \frac{\kappa}{\sqrt{n}} \Longrightarrow \sqrt{\gamma_q(t)} \le \frac{\kappa}{\sqrt{n}}t. \tag{1.7.50}$$

By setting $y_q = \sqrt{\gamma_q}$, if $\omega \neq 0$, we have

$$\begin{aligned}\Big(y_q(t) + \frac{\kappa}{\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Big)' &= y_q'(t) \\ &\le \Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)y_q(t) + \frac{\kappa}{\sqrt{n}} \\ &= \Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\Big(y_q(t) + \frac{\kappa}{\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Big).\end{aligned} \tag{1.7.51}$$

And by Gronwall's lemma, we deduce that ($\gamma_q(0) = 0$)

$$\begin{aligned}y_q(t) + \frac{\kappa}{\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}} &\le \Big(y_q(0) + \frac{\kappa}{\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Big)e^{t\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)}; \\ \sqrt{\gamma_q(t)} &\le \frac{\kappa}{\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Big(e^{t\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)} - 1\Big).\end{aligned} \tag{1.7.52}$$

■

Proposition 1.7.1 (Corollary of the theorem theorem 1.5.1)**.** *The result of the theorem theorem 1.5.1 ensure with explicit rates,*

1. *the weak convergence of the law $\mu_t^{(n),p}$ of a particle towards μ_t: in fact, $X_t^{(n),p}$ has the law $\mu_t^{(n),p}$ by definition and $X_t^{(p)}$ has the law μ_t by construction, therefore for all $t \le T$ and all $n \ge 1$,*

$$\mathcal{W}_2^2(\mu_t^{(n),p}, \mu_t) \le \frac{M}{n} \tag{1.7.53}$$

with M not depending on the number of particles. The uniformity in time of M is verified if $\omega < 0$ in theorem theorem 1.5.1;

2. *the propagation of chaos for the particle system: q being a fixed integer- or more generally a $\mathbf{o}(n)$- and for the Wasserstein distance defined on $(\mathbb{R}^d)^q$, for $\{i_1, \dots, i_q\}$ a part of $\{1, \dots, n\}$, we have*

$$\mathcal{W}_2^2\left(\mathbb{P}_{(X_t^{(n),i_1}, \dots, X_t^{(n),i_q})}, \mu_t^{\otimes q}\right) \le \frac{qM}{n}; \tag{1.7.54}$$

3. *the convergence of the empirical measurement of the particle system towards the McKean-Vlasov particle law: for any Lipschitzian function $\varphi : \mathbb{R}^d \to \mathbb{R}$, we have*

$$\mathbb{E}|\langle \Pi_t^{(n)}, \varphi\rangle - \langle \mu_t, \varphi\rangle|^2 \le \frac{M_{M, M_{2m}, \varphi}}{n}. \tag{1.7.55}$$

Proof of proposition 1.7.1. By definition of the Wasserstein metric and empirical measurement, we have

1.
$$\mathcal{W}_2^2(\mu_t^{(n),p}, \mu_t) \le \mathbb{E}||X_t^{(n),p} - X_t^{(p)}||^2 \le \frac{M}{n}; \tag{1.7.56}$$

2.
$$\begin{aligned} \mathcal{W}_2^2\left(\mathbb{P}_{(X_t^{(n),i_1}, \dots, X_t^{(n),i_q})}, \mu_t^{\otimes q}\right) &\le \mathbb{E}\left|\left|(X_t^{(n),i_1}, \dots, X_t^{(n),i_q}) - (X^{(i_1)}, \dots, X_t^{(i_q)})\right|\right|^2 \\ &\le q\mathbb{E}||X_t^{(n),p} - X_t^{(p)}||^2 \\ &\le \frac{qM}{n} \\ &= M\mathbf{o}(1); \end{aligned} \tag{1.7.57}$$

3.

$$\mathbb{E}|\langle \Pi_t^{(n)}, \varphi\rangle - \langle \mu_t, \varphi\rangle|^2 \le \frac{2M + M_{2m}}{n}[\varphi]_1^2. \tag{1.7.58}$$

■

1.7.2 McKean-Vlasov theory

Theorem 1.7.2 (Existence and uniqueness of solutions of eq. (1.1.13))**.** *Let us assume that the functions b and σ are globally Lipschitz:* $\exists K > 0\ \forall (x, y, \mu, \nu) \in \mathbb{R}^D \times \mathbb{R}^D \times \mathscr{P}_2(\mathbb{R}^D) \times \mathscr{P}_2(\mathbb{R}^D)$,

$$||b(x,\mu) - b(y,\nu)|| + |||\sigma(x,\mu) - \sigma(y,\nu)||| \le K\Big(|x-y| + \mathscr{W}_2(\mu,\nu)\Big), \tag{1.7.59}$$

where $||\cdot||$ denotes a vector norm, $|||\cdot|||$ is a matrix norm and $\mathscr{W}_2$ denotes the Wasserstein-2 distance. Assume that $\mu_0 \in \mathscr{P}(\mathbb{R}^D)$. Then for any $T \ge 0$ the SDE eq. (1.1.13) has a unique strong solution on $[0,T]$ and consequently, its law is the unique weak solution to the Fokker-Planck equation eq. (1.1.12) and the unique solution to the associated nonlinear martingale problem.

The proof of this theorem is fairly classical. This proof is based on a *fixed point argument* that is sketched in [62, Proposition.1].

Theorem 1.7.3 (Polynomial Potential)**.** *Let E be a Polish measurable space. Let $\alpha \in \mathscr{P}(E)$. Let us consider a random vector X^n in E^n, distributed according to the Gibbs measure:*

$$\mu_n(dx) := \frac{1}{Z_n} e^{nF(\mu_x)} \alpha^{\otimes n}(dx), \tag{1.7.60}$$

where Z_n is a normalization constant and F is a polynomial function on $\mathscr{P}(E)$ (called the energy functional) of the form given by eq. (1.1.17). Then (for some symmetric continuous bounded functions $W^{(k)}$) the laws of μ_{X^n} satisfy a large deviation principle in $\mathscr{P}(\mathscr{P}(E))$ with speed $\frac{1}{n}$ and rate function

$$\mu \longmapsto \mathbf{H}[\mu|\alpha] - F(\mu) - \inf_{\eta \in \mathscr{P}(E)} \{\mathbf{H}[\eta|\alpha] - F(\eta)\}. \tag{1.7.61}$$

1.7.3 Gibbs-Laplace Variational Principle

Definition 1.7.1 (Distribution support)**.** Let μ be a probability measure on a Polish space E (or even a measure on a topological space!). We call support of μ noted **supp**(μ) the closed

set defined by

$$\bigcap_{F\subset E \text{ closed},\, \mu(F)=1} F = \Big(\bigcup_{O\subset E \text{ open},\, \mu(O)=0} O \Big)^C. \tag{1.7.62}$$

In other words, the support of a distribution is the complement of the largest open set over which it is zero: the smallest closed set of maximum mass!

Definition 1.7.2 (Extremum essential)**.** Let μ be a probability measure on a Polish space E and $V : E \longrightarrow [-\infty, +\infty]$ measurable. We call infimum $\mu-$essential of V the quantity

$$\mu - \mathbf{essinf} V := \inf\{v \in \mathbb{R}, \quad \mu(\{V \le v\}) > 0\} \tag{1.7.63}$$

Theorem 1.7.4 (Variational principle)**.** *For any probability measure μ on a topological space Ω and any measurable function $V : \Omega \longrightarrow \overline{\mathbb{R}}$, we have*

$$\lim_{n\to+\infty} \frac{1}{n} \log \int e^{-nV} d\mu = -\mu - \mathbf{essinf} V. \tag{1.7.64}$$

Moreover, if V is upper semicontinuous, then

$$\inf_{\mathbf{supp}(\mu)} V = \mu - \mathbf{essinf} V. \tag{1.7.65}$$

Proof Sketch: Suppose $\mu - \mathbf{essinf} V$ is finite. Check that we can assume without loss of generality that $V \ge 0$ and $\mu - \mathbf{essinf} V = 0$. Then check $\mathbb{1}_{V\le\varepsilon} e^{-n\varepsilon} \le e^{-nV} \le 1$ and conclude. Show that the limit is $+\infty$ with the lower bound when $\mu - \mathbf{essinf} V = -\infty$.

Proof. — $\mu - \mathbf{essinf} V$ is finished:

$$\frac{1}{n} \log \int e^{-nV} d\mu + \mu - \mathbf{essinf} V = \frac{1}{n} \log \Big(\int e^{-n(V-\mu-\mathbf{essinf} V)} d\mu \Big). \tag{1.7.66}$$

This implies that we can assume without loss of generality that $V \ge 0$ and $\mu - \mathbf{essinf} V = 0$ because $V - \mu - \mathbf{essinf} V \ge 0$ almost surely, its essential infimum under μ is zero and the convergence that interests us is equivalent to

$$\frac{1}{n} \log \Big(\int e^{-n(V-\mu-\mathbf{essinf} V)} d\mu \Big) \overset{n\to+\infty}{\longrightarrow} 0 \tag{1.7.67}$$

But for all $\varepsilon > 0 = \mu - \mathbf{essinf} V$,

$$\mathbb{1}_{V\le\varepsilon} e^{-n\varepsilon} \le e^{-nV} \le 1 \Longleftrightarrow \frac{\log \mu(V \le \varepsilon)}{n} - \varepsilon \le \frac{1}{n} \log \int e^{-nV} d\mu \le 0. \tag{1.7.68}$$

We deduce that by the bounding limit theorem, we have

$$\limsup_n \frac{1}{n}\log\int e^{-n\mathrm{V}}d\mu = \liminf_n \frac{1}{n}\log\int e^{-n\mathrm{V}}d\mu = 0. \tag{1.7.69}$$

— $\mu-\mathbf{essinf}\mathrm{V} = -\infty$: In this case, for all $v\in\mathbb{R}$, we have $\mu(\mathrm{V}\le v)>0$ and

$$\int_\Omega e^{-n\mathrm{V}}d\mu \ge \int_{\{\mathrm{V}\le v\}} e^{-n\mathrm{V}}d\mu \ge e^{-nv}\mu(\{\mathrm{V}\le v\}). \tag{1.7.70}$$

It is deduced that

$$\begin{aligned}&\forall v\in\mathbb{R},\quad \frac{1}{n}\log\int_\Omega e^{-n\mathrm{V}}d\mu \ge -v + + \frac{\log\mu(\mathrm{V}\le v)}{n}\\ &\Longrightarrow \lim_{n\to+\infty}\frac{1}{n}\log\int e^{-n\mathrm{V}}d\mu = +\infty = -\mu-\mathbf{essinf}\mathrm{V}.\end{aligned} \tag{1.7.71}$$

■

Theorem 1.7.5 (Gibbs measures and deviations)**.** *Let* E *be a Polish space,* μ *a probability measure on* E *and* $\mathrm{V}:\mathrm{E}\longrightarrow\overline{\mathbb{R}}$ *a measurable function. We have:*

—

$$\inf_{\mathbf{supp}(\mu)}\mathrm{V} \le \mu-\mathbf{essinf}\mathrm{V}. \tag{1.7.72}$$

— *If* V *is upper semicontinuous, then*

$$\inf_{\mathbf{supp}(\mu)}\mathrm{V} \ge \mu-\mathbf{essinf}\mathrm{V} \Longrightarrow \inf_{\mathbf{supp}(\mu)}\mathrm{V} = \mu-\mathbf{essinf}\mathrm{V}. \tag{1.7.73}$$

In particular, if V *is continuous, then the principle of large deviations holds for*

$$\mu_n(dx) := \frac{1}{\int e^{-n\mathrm{V}}d\mu}e^{-n\mathrm{V}(x)}\mu(dx) \tag{1.7.74}$$

with rate function $\mathrm{I}^{\mathrm{V}} := \mathrm{V}+\mathrm{I}_0-\inf\{\mathrm{V}+\mathrm{I}_0\}$ *with*

$$\mathrm{I}_0(x) := \begin{cases}0 & \text{if } x\in\mathbf{supp}(\mu),\\ +\infty & \text{else.}\end{cases} \tag{1.7.75}$$

Proof sketch:

— Show that

$$\Big\{x, \quad V(x) < \inf_{\mathbf{supp}(\mu)} V\Big\} \bigcap \mathbf{supp}(\mu) = \emptyset. \tag{1.7.76}$$

Then conclude.

— For all $\varepsilon > 0$, show that

$$\Big\{x, \quad V(x) < \inf_{\mathbf{supp}(\mu)} V + \varepsilon\Big\} \quad \text{is an open containing a support element:} \tag{1.7.77}$$

their intersection is non-empty; then conclude.

1.7.4 Principle of contraction and tensorization

Let $f : X \longrightarrow G$ be continuous between two Polish spaces and (X_N) a random variable sequence of X satisfying the principle of large deviations of rate function $I : X \longrightarrow [0, +\infty]$. Then $((f(X_N))$ satisfies the principle of large deviations of rate function $J : G \longrightarrow [0, +\infty]$ such that

$$J(g) := \inf_{f^{-1}(\{g\})} I. \tag{1.7.78}$$

Let $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ be sequences with values respectively in E_1 and E_2, independent $(\mathbb{P}_{(X_n,Y_n)} = \mathbb{P}_{X_n} \bigotimes \mathbb{P}_{Y_n})$ and both satisfying the principle of large deviations of the respective good rate functions I_1 and I_2. Then $((X_n, Y_n))_{n\geq 1}$ satisfies the principle of large deviations on the product space and of good rate function I defined by

$$I(x, y) := I_1(x) + I_2(y). \tag{1.7.79}$$

1.7.5 Entropy and Chaos

Theorem 1.7.6 (Characterization of relative entropy: Sanov's theorem)**.** *Let* μ *and* ν *be probability measures (even finite!) on a Polish space* E *and* $(\varphi_j)_{j\in\mathbb{N}}$ *a dense sequence of functions bounded uniformly continuous. we have*

$$\lim_{k\to\infty}\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log\mu^{\otimes n}\Big(\Big\{y\in E^n; \quad \forall j\in\{1,\dots,k\}, \quad \Big|\int_E \varphi_j d\nu - \frac{1}{n}\sum_{i=1}^n \varphi_j(y_i)\Big| \leq \varepsilon\Big\}\Big) \tag{1.7.80}$$

$$= -\mathbf{H}[\nu|\mu].$$

We interpret n as the number of particles; the φ_j *a sequence of observables whose mean value is measured; and* ε *as the precision of the measurements. This formula concisely summarizes*

the essential information contained in the Boltzmann function **H**.

Theorem 1.7.7 ((strict) convexity of relative entropy)**.** *Let* $\mu \in \mathscr{P}(\Omega)$. $\mathbf{H}[\cdot|\mu]$ *has values in* $\overline{\mathbb{R}_+}$, *convex, strictly convex on* $\{\nu, \quad \mathbf{H}[\nu|\mu] < +\infty\}$ *and is zero only in* μ.

Theorem 1.7.8 (Tensorization property)**.** *Let* $\mu \in \mathscr{P}(\Omega)$, $\nu \in \mathscr{P}(\Omega^n)$ *with* ν_i *its* $i-th$ *marginal. So*

$$\mathbf{H}[\nu|\mu^{\otimes n}] = \mathbf{H}[\nu|\bigotimes_{i=1}^{n}\nu_i] + \sum_{i=1}^{n}\mathbf{H}[\nu_i|\mu] \tag{1.7.81}$$

Theorem 1.7.9 (Villani)**.** *Let* $(X := (X_1,\dots,X_n)$ *be a random variable on* E^n *with* E *a Polish space,* $\mu_n := \mathbb{P}_X \in \mathscr{P}(E^n)$, $\delta_X := \frac{1}{n}\sum \delta_{X_i}$ *and* $\mu \in \mathscr{P}(E)$. *The following assertions are equivalent:*

— δ_X *converges in law to* μ*:*

$$\forall \varphi \in \mathscr{C}_b(E), \quad \int \varphi d\delta_X \overset{n\to+\infty}{\longrightarrow} \int \varphi d\mu \quad \text{almost surely.} \tag{1.7.82}$$

—

$$\forall \varphi \in \boldsymbol{Lip}_b(E), \quad \lim_{n\to+\infty} \mathbb{E}_{\mu_n}\Big[\Big|\int \varphi d(\delta_X - \mu)\Big|\Big] = 0. \tag{1.7.83}$$

Without repeating the proof, we can say that this result is obtained by defining a metric on $\mathscr{P}(E)$ from a dense sequence of Lipschitz functions and then by defining the transport distance Wasserstein's $\mathscr{W}_1$ on $\mathscr{P}(\mathscr{P}(E))$ associated with this metric. Using this result, we can more formally prove the propagation of chaos.

Definition 1.7.3 (U-statistics)**.** Let E be a set, $k \in \mathbb{N}^*$ and $\Phi : E^k \longrightarrow \mathbb{R}$ a symmetric function. Then the application: $(n \geq k)$

$$X := (x_j)_{j=1,\dots,n} \in E^n \longmapsto U(X) := \frac{k!(n-k)!}{n!} \sum_{1\leq i_1<i_2<\dots<i_k\leq n} \Phi(x_{i_1},\dots,x_{i_k}) \tag{1.7.84}$$

is called U−statistic of order k and kernel Φ. $U(X)$ is called U−statistic of order k and kernel Φ associated with the sample X. This statistic corresponds to the arithmetic mean of the kernel over all the parts at k elements of the set of sample values. we often write $U_n(\Phi)(X) := U(X)$. If E is a measurable space, we generalize this definition to the space of probabilities by the functional $\mu \longmapsto \mathbb{E}_{\mu^{\otimes k}}[\Phi]$.

1.7.6 Proofs

Proof of proposition 1.5.1. Let $\mathfrak{G}_n$ be the group of permutations of $\{1,\dots,n\}$ and $\mathfrak{B}_n$ the $\sigma-$algebra defined by

$$\mathfrak{B}_n := \sigma\Big\{B_n \times C_n | C_n \in \mathscr{B}(E^{[n+1,+\infty[}), \quad B_n \in \mathscr{B}(E^n), \quad \forall \tau \in \mathfrak{G}_n, \quad \tau \mathbb{I}_{B_n} = \mathbb{I}_{B_n}\Big\}. \tag{1.7.85}$$

This $\sigma-$algebra is invariant under permutations and verifies for all $n \geq 1$,

$$\mathfrak{B}_{n+1} \subset \mathfrak{B}_n. \tag{1.7.86}$$

By integrability,

$$\forall (i_1,\dots,i_k) \in I_n^k, \quad \mathbb{E}[\Phi(X_{i_1},\dots,X_{i_k})|\mathfrak{B}_n] = \mathbb{E}[\Phi(X_1,\dots,X_k)|\mathfrak{B}_n], \tag{1.7.87}$$

which implies that

$$U_n(\Phi) = \mathbb{E}[\Phi(X_1,\dots,X_k)|\mathfrak{B}_n]. \tag{1.7.88}$$

According to the limit theorems on martingales (closed martingale) and the law of $0-1$ applied to the asymptotic tribe $\mathfrak{B}_\infty := \bigcap_{n\geq 1} \mathfrak{B}_n$, we deduce that we almost surely have

$$U_n(\Phi) \overset{n\to+\infty}{\longrightarrow} \mathbb{E}[\Phi(X_1,\dots,X_k)|\mathfrak{B}_\infty] = \mathbb{E}[\Phi(X_1,\dots,X_k)]. \tag{1.7.89}$$

■

Proof of proposition 1.5.3. We prove this result by induction. Indeed, for $k=1$ the inequality is verified since we have equality of the two members. Suppose that for $k-1$ the inequality holds. Denote by B_k the left side of this inequality. B_k is equal to

$$\log \mathbb{E}^{X^k}\Big[\mathbb{E}\Big[\exp\Big(\frac{1}{|I_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in I_n^{k-1}}\sum_{i_k\notin\{i_1,\dots,i_{k-1}\}}\frac{1}{n-k+1}\Phi_{i_1,\dots,i_k}(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big|X^k\Big]\Big] \tag{1.7.90}$$

with $X^k := (X_1^k,\dots,X_n^k)$. Let us set

$$\tilde{\Phi}_{i_1,\dots,i_{k-1}} := \frac{1}{n-k+1}\sum_{i_k\notin\{i_1,\dots,i_{k-1}\}}\Phi_{i_1,\dots,i_k}(X_{i_1}^1,\dots,X_{i_k}^k). \tag{1.7.91}$$

By induction hypothesis, we deduce that

$$\mathrm{B}_k \le \log \mathbb{E}^{\mathrm{X}^k}\Big[\exp\Big(\frac{n-k+2}{|\mathrm{I}_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in \mathrm{I}_n^{k-1}} \log \mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde{\Phi}_{i_1,\dots,i_{k-1}}\Big)\Big|\mathrm{X}^k\Big]\Big)\Big]. \tag{1.7.92}$$

Since

$$\log \mathbb{E}^{\mathrm{X}^k}\Big[\exp\Big(\frac{n-k+2}{|\mathrm{I}_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in \mathrm{I}_n^{k-1}} \log \mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde{\Phi}_{i_1,\dots,i_{k-1}}\Big)\Big|\mathrm{X}^k\Big]\Big)\Big] \tag{1.7.93}$$

is upper bounded by

$$\log \mathbb{E}^{\mathrm{X}^k}\Big[\exp\Big(\frac{1}{|\mathrm{I}_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in \mathrm{I}_n^{k-1}} \log\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde{\Phi}_{i_1,\dots,i_{k-1}}\Big)\Big|\mathrm{X}^k\Big]\Big)^{n-k+2}\Big)\Big], \tag{1.7.94}$$

by convexity of $\mathrm{X}\longmapsto \log\mathbb{E}[e^{\mathrm{X}}]$ (consequence of Holder's inequality), we have

$$\mathrm{B}_k \le \frac{1}{|\mathrm{I}_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in \mathrm{I}_n^{k-1}} \log \mathbb{E}^{\mathrm{X}^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde{\Phi}_{i_1,\dots,i_{k-1}}\Big)\Big|\mathrm{X}^k\Big]\Big)^{n-k+2}\Big]. \tag{1.7.95}$$

In this last inequality, for all $(i_1,\dots,i_{k-1})$, the logarithmic term verifies

$$\begin{aligned}&\mathbb{E}^{\mathrm{X}^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde{\Phi}_{i_1,\dots,i_{k-1}}\Big)\Big|\mathrm{X}^k\Big]\Big)^{n-k+2}\Big] \\ &= \mathbb{E}^{\mathrm{X}^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{(n-k+2)(n-k+1)}\sum_{i_k\notin\{i_1,\dots,i_{k-1}\}}\Phi_i(\mathrm{X}^1_{i_1},\dots,\mathrm{X}^k_{i_k})\Big)\Big|\mathrm{X}^k\Big]\Big)^{n-k+2}\Big],\end{aligned} \tag{1.7.96}$$

and by Holder's inequality, we have

$$\begin{aligned}&\mathbb{E}^{\mathrm{X}^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde{\Phi}_{i_1,\dots,i_{k-1}}\Big)\Big|\mathrm{X}^k\Big]\Big)^{n-k+2}\Big] \\ &\le \mathbb{E}^{\mathrm{X}^k}\Big[\Big(\prod_{i_k\notin\{i_1,\dots,i_{k-1}\}}\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\Phi_{i_1,\dots,i_k}(\mathrm{X}^1_{i_1},\dots,\mathrm{X}^k_{i_k})\Big)\Big|\mathrm{X}^k\Big]\Big)^{\frac{n-k+2}{n-k+1}}\Big].\end{aligned} \tag{1.7.97}$$

By Jensen's inequality, we also have the upper bound of the right-hand side of this last inequality by

$$\mathbb{E}^{\mathrm{X}^k}\Big[\prod_{i_k\notin\{i_1,\dots,i_{k-1}\}}\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+1}\Phi_{i_1,\dots,i_k}(\mathrm{X}^1_{i_1},\dots,\mathrm{X}^k_{i_k})\Big)\Big|\mathrm{X}^k\Big]\Big],$$

and by independence, this quantity is equal to

$$\prod_{i_k \notin \{i_1,\dots,i_{k-1}\}} \mathbb{E}\Big[\exp\Big(\frac{1}{n-k+1}\Phi_{i_1,\dots,i_k}(X^1_{i_1},\dots,X^k_{i_k})\Big)\Big]. \tag{1.7.98}$$

■

Proof of proposition 1.5.4. Let $((X^j_1,\dots,X^j_n))_{j=1,\dots,k}$ be independent copies of $(X_1,\dots,X_n)$. By the two propositions above, setting for all $i \in I^k_n$, $\Phi_{i_1,\dots,i_k} \equiv W^{(k)}$, we have for all $\lambda > 0$

$$\Lambda_n(\lambda, W^{(k)}) = \frac{1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n}{|I^k_n|}\sum_{i\in I^k_n} W^{(k)}(X_{i_1},\dots,X_{i_k})\Big)\Big], \tag{1.7.99}$$

and it follows that

$$\begin{aligned}\Lambda_n(\lambda, W^{(k)}) &\le \frac{1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{1}{|I^k_n|}\sum_{i\in I^k_n}\lambda n C_k |W^{(k)}|(X^1_{i_1},\dots,X^k_{i_k})\Big)\Big] \\ &\le \frac{1}{n|I^{k-1}_n|}\sum_{i\in I^k_n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X^1_{i_1},\dots,X^k_{i_k})\Big)\Big].\end{aligned} \tag{1.7.100}$$

Gold

$$\begin{aligned}&\frac{1}{n|I^{k-1}_n|}\sum_{i\in I^k_n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X^1_{i_1},\dots,X^k_{i_k})\Big)\Big] \\ &= \frac{n-k+1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X_{i_1},\dots,X_{i_k})\Big)\Big].\end{aligned} \tag{1.7.101}$$

It is deduced that

$$\begin{aligned}\Lambda_n(\lambda, W^{(k)}) &\le \frac{n-k+1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X_{i_1},\dots,X_{i_k})\Big)\Big] \\ &\le \frac{1}{k}\log\mathbb{E}\Big[\exp\Big(kC_k\lambda|W^{(k)}|(X_1,\dots,X_k)\Big)\Big],\end{aligned} \tag{1.7.102}$$

and this last inequality is obtained by growth on $(0,+\infty)$ of $a \longmapsto \frac{1}{a}\log\mathbb{E}[e^{aX}]$ and from the fact that for all n and k such that $n \ge k$, we have $\frac{n}{n-k+1} \le k$. ■

Proof of proposition 1.5.5. To do this, we will show that for any probability measure μ such that $\mathbf{H}[\mu|\alpha] < +\infty$ and for any k, $W^{(k)} \in L^1(\mu^{\otimes k})$, we have $l^*(\mathscr{O}) \ge -\mathbf{E}_W[\mu]$. Let $\mathbb{B}(\mu,\delta)$ be the open ball with center μ and of radius $\delta > 0$ in $\mathscr{M}_1(\mathbb{R}^d)$ endowed with the Lévy-Prokhorov metric d_{LP} such that $\mathbb{B}(\mu,\delta) \subset \mathscr{O}$. Let us introduce the events

—

$$\mathrm{A}_n := \Big\{ x \in (\mathbb{R}^d)^n \Big| \quad \mathrm{L}_n := \mathrm{L}_n(x,\cdot) \in \mathbb{B}(\mu,\delta) \Big\}; \tag{1.7.103}$$

—

$$\mathrm{B}_n := \Big\{ x \in (\mathbb{R}^d)^n \Big| \quad \frac{1}{n}\sum_{i=1}^n \log \frac{d\mu}{d\alpha}(x_i) = \frac{1}{n}\log\Big(\frac{d\mu}{d\alpha}\Big)^{\otimes n}(x) \le \mathbf{H}[\mu|\alpha] + \varepsilon \Big\}; \tag{1.7.104}$$

—

$$\mathrm{C}_n := \Big\{ x \in (\mathbb{R}^d)^n \Big| \sum_{k=2}^{\mathrm{N}} \mathrm{U}_n(\mathrm{W}^{(k)}) \le \sum_{k=2}^{\mathrm{N}} \mathbf{W}^{(k)}[\mu] + \varepsilon \Big\}. \tag{1.7.105}$$

We deduce that for all $\varepsilon > 0$, we have

$$\begin{aligned} \mu_n^*(\mathrm{L}_n \in \mathbb{B}(\mu,\delta)) &\ge \int_{\mathrm{A}_n} \Big(\frac{d\mu^{\otimes n}}{d\mu_n^*}(x)\Big)^{-1} \mu^{\otimes n}(dx) \\ &= \int_{\mathrm{A}_n} e^{-\sum_{i=1}^n \log \frac{d\mu}{d\alpha}(x_i)} e^{-n\sum_{k=2}^{\mathrm{N}} \mathrm{U}_n(\mathrm{W}^{(k)})} \mu^{\otimes n}(dx) \end{aligned} \tag{1.7.106}$$

and

$$\int_{\mathrm{A}_n} e^{-\sum_{i=1}^n \log \frac{d\mu}{d\alpha}(x_i)} e^{-n\sum_{k=2}^{\mathrm{N}} \mathrm{U}_n(\mathrm{W}^{(k)})} \mu^{\otimes n}(dx) \ge \mu^{\otimes n}(\mathrm{A}_n \cap \mathrm{B}_n \cap \mathrm{C}_n) e^{-n(\mathbf{H}[\mu|\alpha]+\varepsilon)-\gamma} \tag{1.7.107}$$

with $\gamma := n(\sum_{k=2}^{\mathrm{N}} \mathbf{W}^{(k)}[\mu] + \varepsilon)$.

Thereby

$$\mu_n^*(\mathrm{L}_n \in \mathbb{B}(\mu,\delta)) \ge \mu^{\otimes n}(\mathrm{A}_n \cap \mathrm{B}_n \cap \mathrm{C}_n) e^{-n\mathbf{E}_{\mathrm{W}}[\mu]-2n\varepsilon}. \tag{1.7.108}$$

We will prove that

$$\mu^{\otimes n}(\mathrm{A}_n \cap \mathrm{B}_n \cap \mathrm{C}_n) \overset{n\to+\infty}{\longrightarrow} 1. \tag{1.7.109}$$

Indeed, by the law of large numbers, we have

$$\mu^{\otimes n}(\mathrm{A}_n) \overset{n\to+\infty}{\longrightarrow} 1, \quad \mu^{\otimes n}(\mathrm{B}_n) \overset{n\to+\infty}{\longrightarrow} 1. \tag{1.7.110}$$

Moreover, by the law of large numbers for U−statistics (section 1.5.2), we also have

$$\mu^{\otimes n}(\mathrm{C}_n) \overset{n\to+\infty}{\longrightarrow} 1. \tag{1.7.111}$$

It is deduced that

$$l^*(\mathscr{O}) \ge \liminf_{n\to+\infty} \frac{1}{n} \mu_n^*(\mathrm{L}_n \in \mathbb{B}(\mu,\delta)) \ge -\mathbf{E}_{\mathrm{W}}[\mu] - 2\varepsilon, \tag{1.7.112}$$

and we conclude by letting ε tend to zero. ■

Proof of proposition 1.5.6. To do this, consider the truncation function

$$\mathrm{W}^{(k),\mathrm{L}} := \max(-\mathrm{L}, \min(\mathrm{W}^{(k)}, \mathrm{L})). \tag{1.7.113}$$

We have by Lebesgue's dominated convergence theorem

$$\log \mathbb{E}[\exp(m|\mathrm{W}^{(k),\mathrm{L}} - \mathrm{W}^{(k)}|(\mathrm{X}_1,\ldots,\mathrm{X}_k))] \overset{\mathrm{L}\to+\infty}{\longrightarrow} 0. \tag{1.7.114}$$

So we can choose $\mathrm{L} = \mathrm{L}(m)$ so that

$$\log \mathbb{E}[\exp(m|\mathrm{W}^{(k),\mathrm{L}(m)} - \mathrm{W}^{(k)}|(\mathrm{X}_1,\ldots,\mathrm{X}_k))] \leq \frac{1}{m}. \tag{1.7.115}$$

For $m \geq 1$ and $\mathrm{L}(m) > 0$ fixed, we can find a sequence $(\mathrm{W}_l^{(k),\mathrm{L}})_{l\geq 1}$ of continuous functions bounded such that

$$\mathrm{W}_l^{(k),\mathrm{L}}(\mathrm{X}_1,\ldots,\mathrm{X}_k) \overset{l\to+\infty,\mathrm{L}^1}{\longrightarrow} \mathrm{W}^{(k),\mathrm{L}}(\mathrm{X}_1,\ldots,\mathrm{X}_k), \quad \forall l \geq 1, \quad |\mathrm{W}_l^{(k),\mathrm{L}}(\mathrm{X}_1,\ldots,\mathrm{X}_k)| \leq \mathrm{L}, \tag{1.7.116}$$

otherwise, we consider the truncation $\max(-\mathrm{L}, \min(\mathrm{W}_l^{(k),\mathrm{L}}, \mathrm{L}))$. Seen that $\forall l \geq 1$,

$$\exp\Big(m\Big(|\mathrm{W}^{(k)} - \mathrm{W}^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k) + |\mathrm{W}^{(k)} - \mathrm{W}_l^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big) \tag{1.7.117}$$

$$\leq \exp\Big(m|\mathrm{W}^{(k)} - \mathrm{W}^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k) + 2m\mathrm{L}\Big), \tag{1.7.118}$$

by dominated convergence, we have

$$\mathbb{E}\Big[\exp\Big(m\Big(|\mathrm{W}^{(k)} - \mathrm{W}^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k) + |\mathrm{W}^{(k)} - \mathrm{W}_l^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big)\Big] \tag{1.7.119}$$

$$\overset{l\to+\infty}{\longrightarrow} \mathbb{E}\Big[\exp\Big(m|\mathrm{W}^{(k)} - \mathrm{W}^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big].$$

For $\mathrm{L} = \mathrm{L}(m)$, we can choose $l = l(m)$ so that

$$\log \mathbb{E}\Big[\exp\Big(m\Big(|\mathrm{W}^{(k)} - \mathrm{W}^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k) + |\mathrm{W}^{(k)} - \mathrm{W}_l^{(k),\mathrm{L}}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big)\Big] \leq \frac{2}{m}. \tag{1.7.120}$$

By setting $\mathrm{W}_m^{(k)} = \mathrm{W}_{l(m)}^{(k),\mathrm{L}(m)}$ bounded continuous function, we have by triangular inequality

$$\log \mathbb{E}\Big[\exp\Big(m|\mathrm{W}^{(k)} - \mathrm{W}_m^{(k)}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big] \leq \frac{2}{m}. \tag{1.7.121}$$

Since by Jensen's inequality, we have for all $\lambda > 0$, $\forall m \geq \lambda$,

$$\log\mathbb{E}\Big[\exp\Big(\lambda|\mathrm{W}^{(k)} - \mathrm{W}_m^{(k)}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big] \leq \frac{\lambda}{m}\mathbb{E}\Big[\exp\Big(m|\mathrm{W}^{(k)} - \mathrm{W}_m^{(k)}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big], \tag{1.7.122}$$

we deduce that

$$\log\mathbb{E}\Big[\exp\Big(\lambda|\mathrm{W}^{(k)} - \mathrm{W}_m^{(k)}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big] \overset{m\to+\infty}{\longrightarrow} 0. \tag{1.7.123}$$

For all $\delta > 0$ and $\lambda > 0$, by the Markov-Tchebychev inequality, we have

$$\mathbb{P}(|\mathrm{U}_n(\mathrm{W}^{(k)}) - \mathrm{U}_n(\mathrm{W}_m^{(k)})| > \delta) \leq e^{-n\lambda\delta}\mathbb{E}\Big[\exp\Big(n\lambda\mathrm{U}_n(|\mathrm{W}^{(k)} - \mathrm{W}_m^{(k)}|)\Big)\Big]. \tag{1.7.124}$$

From the above (section 1.5.2), we deduce that

$$\begin{aligned}\frac{1}{n}\log\mathbb{P}(|\mathrm{U}_n(\mathrm{W}^{(k)}) - \mathrm{U}_n(\mathrm{W}_m^{(k)})| > \delta) &\leq -\lambda\delta \\ &+ \frac{1}{k}\log\mathbb{E}\Big[\exp\Big(k\mathrm{C}_k\lambda|\mathrm{W}^{(k)} - \mathrm{W}_m^{(k)}|(\mathrm{X}_1,\ldots,\mathrm{X}_k)\Big)\Big].\end{aligned} \tag{1.7.125}$$

We conclude that we have the expected result when $m \to +\infty$ since λ is arbitrary. ■

Proof of proposition 1.5.12. Let $\mathrm{Q}^i(\cdot|x_{[1,i-1]})$ be the conditional distribution of x_i knowing

$$x_{[1,i-1]} := (x_1, \cdots, x_{i-1})$$

(not knowing if $i = 1$). We have:

$$\begin{aligned}\mathbf{H}[\mathrm{Q}|\prod_{i=1}^{\mathrm{N}}\alpha_i] &= \mathbb{E}_{\mathrm{Q}}\Big[\log\Big(\frac{d\mathrm{Q}}{d\prod_{i=1}^{\mathrm{N}}\alpha_i}\Big)\Big] \\ &= \mathbb{E}_{\mathrm{Q}}\Big[\sum_{i=1}^{\mathrm{N}}\log\Big(\frac{\mathrm{Q}^i(dx_i|x_{[1,i-1]})}{\alpha_i(dx_i)}\Big)\Big] \\ &= \mathbb{E}_{\mathrm{Q}}\Big[\sum_{i=1}^{\mathrm{N}}\mathbf{H}[\mathrm{Q}^i(\cdot|x_{[1,i-1]})|\alpha_i]\Big].\end{aligned} \tag{1.7.126}$$

Since

$$\mathbb{E}_{\mathrm{Q}}[\mathrm{Q}^i(\cdot|x_{[1,i-1]})] = \mathrm{Q}_i(\cdot), \tag{1.7.127}$$

we obtain by convexity of the relative entropy (Jensen's inequality):

$$\mathbb{E}_{\mathrm{Q}}\Big[\mathbf{H}[\mathrm{Q}^i(\cdot|x_{[1,i-1]})|\alpha_i]\Big] \geq \mathbf{H}[\mathrm{Q}_i|\alpha_i] \tag{1.7.128}$$

Which shows that we have the result of the proposition. ■

Proof of proposition 1.5.13. For f a measurable function on E, we define:

$$\Lambda_\mu(f) := \log(\mathbb{E}_\mu[e^f]) = \log \int e^f d\mu \in (-\infty, +\infty] \tag{1.7.129}$$

the log-Laplace transformation under μ which is convex in f by Holder's inequality. We have:

$$\Lambda_{\mu_U}(f) = \log \int e^f d\mu_U = \Lambda_\mu(f - U) - \Lambda_\mu(-U) \leq \frac{1}{p}\Lambda_\mu(-pU) + \frac{1}{q}\Lambda_\mu(qf) - \Lambda_\mu(-U) \tag{1.7.130}$$

by Holder's inequality considering the conjugate exponent $q := \frac{p}{p-1}$ of p. By the variational formula of Donsker-Varadhan, we deduce that:

$$\begin{aligned} \mathbf{H}[\nu|\mu_U] &= \sup_{f \in \mathscr{M}_b(E)} \left\{ \int f d\nu - \Lambda_{\mu_U}(f) \right\} \\ &\geq \sup_{f \in \mathscr{M}_b(E)} \left\{ \int f d\nu - \frac{1}{q}\Lambda_\mu(qf) \right\} + \Lambda_\mu(-U) - \frac{1}{p}\Lambda_\mu(-pU). \end{aligned} \tag{1.7.131}$$

Gold:

$$\begin{aligned} &\sup_{f \in \mathscr{M}_b(E)} \left\{ \int f d\nu - \frac{1}{q}\Lambda_\mu(qf) \right\} + \Lambda_\mu(-U) - \frac{1}{p}\Lambda_\mu(-pU) \\ &\qquad = \frac{1}{q}\mathbf{H}[\nu|\mu] + \Lambda_\mu(-U) - \frac{1}{p}\Lambda_\mu(-pU). \end{aligned} \tag{1.7.132}$$

So if $\mathbf{H}[\nu|\mu_U] < +\infty$, $\mathbf{H}[\nu|\mu] < +\infty$ or equivalently, $\log(\frac{d\nu}{d\mu}) \in L^1(\nu)$ and:

$$\log\left(\frac{d\nu}{d\mu_U}\right) = \log\left(\frac{d\nu}{d\mu}\right) + U + \Lambda_\mu(-U) \in L^1(\nu). \tag{1.7.133}$$

This proves the proposition. ■

2

Dynamique de Langevin de champ moyen et transition de phase

[Le grand livre de l'Univers] est écrit dans la langue mathématique, et ses caractères sont des triangles, des cercles et autres figures géométriques, sans lesquels il est humainement impossible d'en comprendre un mot ; sans eux, c'est une errance vaine dans un labyrinthe obscur.

GALILÉE

Résumé du chapitre

En tant qu'exemple de l'équation non linéaire de Fokker-Planck, la dynamique de Langevin de champ moyen a attiré récemment l'attention en raison de sa connexion à la descente de gradient stochastique. Dans ce chapitre (fruit de l'article [1]), l'objectif principal est de donner des conditions suffisantes avec une énergie plate non nécessairement convexe évitant le phénomène de transition de phase : existence d'au plus une distribution de probabilité invariante pour la dynamique de Langevin de champ moyen.

2.1 Introduction

The present work studies both the quantitative long-time behavior of the first order (over-damped) and second order (under-damped) mean field Langevin dynamics and their uniform-in-time propagations of chaos properties, under a functional not necessary convex in the flat interpolation sense.

In this paper, we consider the first order McKean-Vlasov equation

$$dX_t = \sigma dB_t - \mathscr{D}_m H(\mu_t, X_t)dt; \tag{2.1.1}$$

and the second order McKean-Vlasov equation

$$\begin{cases} dP_t &= V_t dt; \\ dV_t &= \sigma dB_t - A(V_t)dt - B(P_t)dt - \mathscr{D}_m H(\mathbb{P}_{P_t}, P_t)dt; \end{cases} \tag{2.1.2}$$

where $\mathscr{D}_m H$ denotes the intrinsic derivative of the functional H (a mapping from $\mathscr{P}_p(\mathbb{R}^d)\times\mathbb{R}^d$ into $\mathbb{R}^d$ and will defined in eq. (2.2.3)), $\sigma > 0$, $\mu_t := \mathbb{P}_{X_t}$, A a friction field (velocity field) and B an external confinement field.Without loss of generality, in the remainder of the paper, we will take $\sigma = \sqrt{2}$.

Let $n \geq 1$ be an integer and let $\mathbf{x} := (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n$ be an n-tuple of coordinates in $\mathbb{R}^d$. We denote by $\mu_{\mathbf{x}}$ the empirical measure formed with the n coordinates $x_1, \ldots, x_n$, that is,

$$\mu_{\mathbf{x}} := \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}. \tag{2.1.3}$$

We consider the first order mean field particle system

$$\forall i \in \{1, \ldots, n\}, \quad dX_t^{(i)} = \sigma dB_t^{(i)} - \mathscr{D}_m H(\mu_{\mathbf{X}_t}, X_t^{(i)})dt; \tag{2.1.4}$$

and the second order mean field particle system

$$\begin{cases} dP_t^{(i)} &= V_t^{(i)} dt; \\ dV_t^{(i)} &= \sigma dB_t^{(i)} - A(V_t^{(i)})dt - B(P_t^{(i)})dt - \mathscr{D}_m H(\mu_{\mathbf{P}_t}, P_t^{(i)})dt; \end{cases} \tag{2.1.5}$$

where $B^{(1)},\dots,B^{(n)}$ are i.i.d. standard Brownian motions in d-dimensions.

For $\mu \in \mathscr{P}(\mathbb{R}^d)$, we consider the over-damped Langevin dynamics:

$$dX_t^\mu = \sqrt{2}dB_t^\mu - \mathscr{D}_m H(\mu, X_t^\mu)dt. \tag{2.1.6}$$

By explicit computations, we can find that the Gibbs measure:

$$\Phi(\mu)(dx) := \frac{e^{-\frac{\delta H}{\delta m}(\mu,x)}}{\int_{\mathbb{R}^d} e^{-\frac{\delta H}{\delta m}(\mu,y)}dy}dx \tag{2.1.7}$$

is invariant to the dynamics ($\frac{\delta H}{\delta m}$ is called *flat derivative* of H and is a continuous mapping from $\mathscr{P}_p(\mathbb{R}^d)\times\mathbb{R}^d$ into $\mathbb{R}$ and will defined in eq. (2.2.8)). By explicit computations, we can find that the set of invariant probability distributions of the mean field Langevin dynamics (over-damped) is defined by

$$\{\mu_\infty, \quad \Phi(\mu_\infty) = \mu_\infty\}. \tag{2.1.8}$$

By explicit computations, we can find that the n-particle Gibbs measure $\mu^{(n)}$ for the energy H, defined by

$$\mu^{(n)}(d\mathbf{x}) := \frac{e^{-nH(\mu_\mathbf{x})}}{\int_{\mathbb{R}^{nd}} e^{-nH(\mu_\mathbf{y})}d\mathbf{y}}d\mathbf{x}, \tag{2.1.9}$$

is invariant to the over-damped dynamics.

In the kinetic case, in general, if the velocity and external confinement fields do not derive from potential gradients, the invariant probabilities are not explicit.

We give in this section a short review of the recent progresses in the long-time behavior and the uniform-in-time propagation of chaos property of McKean–Vlasov dynamics, with an emphasis on (over and under damped) mean field Langevin dynamics ones. We refer readers to [63, 64] for a more comprehensive review of propagation of chaos.

As for the long-time behavior, the uniform in time propagation of chaos and the convergence rates, many works have been devoted to the subject and different approaches have been introduced: coupling methods, functional inequalities...

Coupling approaches. Let us briefly describe the coupling method (the major approach of this paper). The basic idea is that an upper bound on the Wasserstein distance between

two probability distributions is given by the construction of any pair of random variables distributed respectively according to those. The goal is thus to construct simultaneously two solutions hat have a trend to get closer with time. Each solution being driven by a Brownian motion, a coupling of two solutions then follows from a coupling of the Brownian motions. The nature of the coupling therefore depends on the choice of Brownian motions. The coupling is said to be synchronous if the Brownian motions are the same: it is a coupling for which the Brownian noise cancels out in the infinitesimal evolution of the difference. In that case the contraction of a distance between the processes can only be induced by the drift component, as in [56, 40]. Such a contraction holds under the strong convexity conditions.

Nevertheless, in more general cases, the calculation of the evolution of the difference shows that there is still some deterministic contraction on a subspace: see, among others, [108, 131, 152]. We can therefore use a synchronous coupling in the vicinity of this subspace. Outside of this subspace, it is necessary to make use of the noise to get the processes closer together, at least in the direction orthogonal to this space. In order to maximize the variance of this noise, we then use a so-called reflection coupling, which consists of taking opposing Brownian motions in the direction of space given by the difference process, and synchronous in the orthogonal direction.

The granular media and the Vlasov-Fokker-Planck dynamics have attracted a lot of attention, usually named McKean-Vlasov diffusions, these past twenty-four years, by means of a stochastic interpretation and synchronous couplings as in [56, 40] or the recent by reflection couplings as in [131] enabling to get rid of convexity conditions, but limited to small interactions. Remark however that small interactions are natural to get uniform in time propagation of chaos as for large interactions the non linear limit equation may have several stationary measures (see [224] for example).

Concerning the mean-field Langevin dynamics, propagation of chaos in finite time for the first and second orders McKean-Vlasov dynamics given by eq. (2.1.1) and eq. (2.1.2) is relatively easy to show using *synchronous coupling*: for example, regarding *polynomial interaction*, see [7, Theorem 5.1]. The bound obtained by this method, however, generally tends to infinity when the time interval extends to infinity. Besides, the dynamics may possess *multiple invariant measures*, so uniform in time convergence can not be expected without additional assumptions, or a more general definition of convergence itself (e.g. *convergence modulo symmetries*). By assuming the functional convexity of the energy, uniform in time propagation of chaos for the mean-field Langevin dynamics and the kinetic mean-field Langevin dynamics have been proven in [74, 75]. In [152], the authors consider a flat energy which

could be non-convex but is essentially with small (nonlinear) dependence, and aim at obtaining an exponential contraction result. The authors mainly borrow the tools developed in Eberle, Guillin and Zimmer [107, 108], where they initiate the reflection-synchronous coupling technique, further validate it in the study of the Langevin dynamics with a general non-convex potential, and make the point that the technique offers significant flexibility for additional development. But note that [108] is not concerned with mean-field interaction and the rate found there is dimension dependent. In [152], the authors design a new metric involving a quadratic form ([152, Section 4.4.2] to obtain the contraction when the coupled particles are far away, and as a result obtain a dimension-free convergence rate. The construction of the quadratic form shares some flavor with the argument in Bolley, Guillin and Malrieu [40]. Notably, their construction helps to capture the optimal rate in the area of interest ([152, Remark 4.16]), so may be more intrinsic. Notice that most of the articles concerning the ergodicity of underdamped Langevin dynamics obtain the convergence rates depending on the dimension, and in particular very few allow both non-convex potential and the mean-field interaction. In [40], the result holds in the noncompact case with small Lipschitz interaction, and can be seen as a first attempt to deal with more general case. Moreover it shows existence and uniqueness of the equilibrium measure, and in particular does not use its explicit expression (which is unknown in our broader situation). It is also not only a result on the convergence to equilibrium, but also a stability result of all solutions. Let us finally note that it is based on the natural stochastic interpretation and a simple coupling argument, and does not need any hypoelliptic regularity property of the solutions.

Functional approaches. Another approach to long-time behavior and uniform-in-time propagation of chaos is the functional one, and this is also the major approach of [74, 75, 190, 191]. In this situation in order to study the long-time behaviors and propagation of chaos properties, we construct appropriate (Lyapunov) functionals and investigate the change of their values along the dynamics. For example, the relative entropy is used as the functional in many works. In fact, for example, if $\mu^{(n)}$ verifies a ρ_n-*log-Sobolev inequality* :

$$\rho_n \mathbf{H}[\cdot|\mu^{(n)}] \leq 2\mathbf{I}[\cdot|\mu^{(n)}], \tag{2.1.10}$$

then for all initial value $\mu_0^{(n)}$ of the flow,

$$\mathbf{H}[\mu_t^{(n)}|\mu^{(n)}] \leq e^{-2\rho_n t}\mathbf{H}[\mu_0^{(n)}|\mu^{(n)}] \tag{2.1.11}$$

where **H** is the *relative entropy* defined by

$$\mathbf{H}[\nu|\mu] = \begin{cases} \mathbb{E}_\nu[\log \frac{d\nu}{d\mu}] =: \mathbf{Ent}_\mu[\frac{d\nu}{d\mu}] & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.1.12}$$

and **I** is the *Fisher information* defined by

$$\mathbf{I}[\nu|\mu] = \int \left|\left|\nabla \sqrt{\frac{d\nu}{d\mu}}\right|\right|^2 d\mu = \frac{1}{4}\int \left|\left|\nabla \log \frac{d\nu}{d\mu}\right|\right|^2 d\nu = \frac{1}{4}\int \left|\left|\mathscr{D}_m \mathbf{H}[\cdot|\mu](\nu, y)\right|\right|^2 \nu(dy) \tag{2.1.13}$$

if $\nu \ll \mu$ and $\sqrt{\frac{d\nu}{d\mu}} \in \mathbf{H}^1_\mu$, and $\mathbf{I}[\nu|\mu] = +\infty$ otherwise, where $\mathbf{H}^1_\mu$ is the domain of the Dirichlet form

$$\mathscr{E}_\mu : g \longmapsto \int ||\nabla g||^2 d\mu. \tag{2.1.14}$$

We refer to Section 5.2 of the Bakry-Gentil-Ledoux monograph [15] for details. Thus the existence of a log-Sobolev constant independent from n implies that the rate of *exponential convergence in entropy* is independent of the number of particles. The interest of such a universal constant lies in the study of the *long-time behavior* of the *mean field limit* (for example, [7]). In [234], assuming a *log-Sobolev inequality* for $\Phi(\mu)$ uniform in μ, a *Poincaré inequality* independent of n for the conditional marginals of $\mu^{(n)}$ and the *convexity in flat interpolation sense* of H, a log-Sobolev inequality for $\mu^{(n)}$ independent of n is proved. Time-uniform log-Sobolev inequalities satisfied by solutions of semi-linear mean-field equations have recently appeared to be a key tool to obtain time-uniform propagation of chaos estimates.

In this paper, the flat energies considered are not necessarily convex. Let the flat energy regularized by relative entropy:

$$H_0 + \frac{\sigma^2}{2}\mathbf{H}[\cdot|\alpha] \tag{2.1.15}$$

where α is the *Gibbs distribution* associated with a *confinement potential* V and energy H_0 is convex in functional sense:

$$\forall (t,\mu,\nu) \in [0,1] \times \mathscr{P}(\mathbb{R}^d) \times \mathscr{P}(\mathbb{R}^d), \quad H_0((1-t)\mu + t\nu) \le (1-t)H_0(\mu) + tH_0(\nu). \tag{2.1.16}$$

In [138], The drift associated with the Hamiltonian of the mean-field Langevin dynamics given by $\mu \mapsto H_0(\mu) + \int V d\mu$ corresponds to the *gradient descent* in L^2-Wasserstein space of

the *free energy*. Since the free energy is strictly convex, there is at most one minimizer μ_∞ of the energy that satisfies

$$\frac{\delta H_0}{\delta m}(\mu_\infty,\cdot)+\frac{\sigma^2}{2}\ln\frac{d\mu_\infty}{d\alpha}=C_{\mu_\infty}\Longleftrightarrow\mu_\infty(dx)=\frac{1}{Z_{\mu_\infty}}e^{-\frac{2}{\sigma^2}\frac{\delta H_0}{\delta m}(\mu_\infty,\cdot)}\alpha(dx). \tag{2.1.17}$$

Contributions. In the continuity of [7, 40], we are interested in problems of contraction and uniform in time propagation of chaos by synchronous coupling for mean-field Langevin dynamics (over- and under-damped).

1. **Over-damped dynamics**
   - — First, as a generalization of the proof of [7, Proposition 4.1], we show that the invariant measure map is a strict contraction of a metric space (theorem 2.3.1): existence of at most one equilibrium on this space.
   - — Second, by synchronous coupling, we show that there exists a unique equilibrium and that we have exponential convergence towards it (theorem 2.3.2).
   - — Finally, we show a result of propagation of chaos L^2 with a bound which can be chosen uniform in time (theorem 2.3.3).
2. **under-damped dynamics**
   - — First, as a generalization of [40, Theorem 1], we show that there exists a unique kinetic equilibrium and that we have exponential convergence towards it (theorem 2.3.4).
   - — Finally, as a generalization of [40, Theorem 2], we show a result of uniform in time propagation of chaos L^2 (theorem 2.3.5).

Organization of paper. Let us finish this introduction by the plan of the paper. The paper is mainly divided into five sections. In section 2.2, we give the settings and assumptions of the paper. The main theorems are stated in section 2.3 and examples are given in section 2.4. We discuss the connection with partial differential equations in section 2.5. Finally, section 2.6 focuses on the proofs of the main results of the paper.

2.2 Settings and assumptions

We give an informal preview of our settings and assumptions in this section.

2.2.1 Settings

Let $d \geq 1$ be an integer. In this paper, we consider mean field energy functionals H that are defined on the space of probability measures on $\mathbb{R}^d$. More precisely, such functionals are mappings ($p \geq 0$)

$$\mathrm{H}: \quad \mathscr{P}_p(\mathbb{R}^d) \longrightarrow \mathbb{R} \tag{2.2.1}$$

where $\mathscr{P}_p(\mathbb{R}^d)$ denotes the set of probability measures of finite p-moment equipped with its weak topology. We denote $\mathscr{W}_p$ the Wasserstein p-metric defined by

$$\forall (\mu,\nu) \in \mathscr{P}_p(\mathbb{R}^d) \times \mathscr{P}_p(\mathbb{R}^d), \quad \mathscr{W}_p(\mu,\nu) := \inf_{X\sim\mu, Y\sim\nu} (\mathbb{E}[\|X-Y\|^p])^{\min(\frac{1}{p},1)}. \tag{2.2.2}$$

We denote by $\mathscr{D}_m\mathrm{H}$ and $\mathscr{D}_m^2\mathrm{H}$ the first and second orders intrinsic derivatives of H: the first order intrinsic derivative is a mapping from $\mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d$ into $\mathbb{R}^d$; and the second order intrinsic derivative is a mapping from $\mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d$ into $\mathbb{R}^{d\times d}$ and will be defined by

$$\mathscr{D}_m\mathrm{H}(\mu,x) := \nabla_x \frac{\delta \mathrm{H}}{\delta m}(\mu,x); \tag{2.2.3}$$

$$[\mathscr{D}_m\mathrm{H}]_{1,\infty} := \sup_{\mu} [\mathscr{D}_m\mathrm{H}(\mu,\cdot)]_1; \tag{2.2.4}$$

$$\mathscr{D}_m^2\mathrm{H}(\mu,x,y) := \nabla^2_{x,y} \frac{\delta^2 \mathrm{H}}{\delta m^2}(\mu,x,y); \tag{2.2.5}$$

$$\|\mathscr{D}_m^2\mathrm{H}\|_{\mathbf{op},\infty} := \sup_{\mu,x,y} \|\mathscr{D}_m^2\mathrm{H}(\mu,x,y)\|_2; \tag{2.2.6}$$

where $\frac{\delta \mathrm{H}}{\delta m}$ is called (first) *flat derivative* of H and is a continuous mapping from $\mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d$ into $\mathbb{R}$ satisfying the two following properties :

— For all $\mathrm{K} \subset \mathscr{P}_p(\mathbb{R}^d)$ compact,

$$\exists \mathrm{C}_\mathrm{K} \geq 0 \quad \forall x \in \mathbb{R}^d, \quad \sup_{\mu\in\mathrm{K}} \left| \frac{\delta \mathrm{H}}{\delta m}(\mu,x) \right| \leq \mathrm{C}_\mathrm{K}(1+\|x\|^p); \tag{2.2.7}$$

—

$$\mathrm{H}(\mu_1) - \mathrm{H}(\mu_0) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta \mathrm{H}}{\delta m}(t\mu_1 + (1-t)\mu_0, x)(\mu_1-\mu_0)(dx)dt \tag{2.2.8}$$

$$\iff \frac{d}{dt}\mathrm{H}(t\mu_1 + (1-t)\mu_0) = \int_{\mathbb{R}^d} \frac{\delta \mathrm{H}}{\delta m}(t\mu_1 + (1-t)\mu_0, x)(\mu_1-\mu_0)(dx). \tag{2.2.9}$$

2.2.2 Assumptions

We impose the following assumptions on the energy H.

Assumptions of over-damped dynamics

Assumptions 2.2.1. The functional H admits first and second order flat derivatives

$$\frac{\delta \mathrm{H}}{\delta m}: \quad \mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \longrightarrow \mathbb{R}; \tag{2.2.10}$$

$$\frac{\delta^2 \mathrm{H}}{\delta m^2}: \quad \mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}; \tag{2.2.11}$$

that are jointly continuous and are $\mathscr{C}^2$ in the spatial variables. we have the following:

1. $(\mu, x, y) \mapsto \|\mathscr{D}_m^2 \mathrm{H}(\mu, x, y)\|_2$ is bounded ($\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} < +\infty$);
2. $(\mu, x) \mapsto \lambda_{\nabla \mathscr{D}_m \mathrm{H}(\mu,x)}$ is strictly bounded from below by $\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}$.

Assumptions 2.2.2. Assume that the functional H admits first and second order flat derivatives that are jointly continuous and are $\mathscr{C}^2$ in the spatial variables. Suppose that

$$\sup_{\mu \in \mathscr{P}(\mathbb{R}^d)} \left\| \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, 0) \right\|_2 < +\infty \tag{2.2.12}$$

and there exists $\lambda > 0$ such that

1. for all $(\mu, x, y) \in \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d$,

$$\left\langle \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, x) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu, y), x - y \right\rangle \geq \lambda \|x - y\|_2^2; \tag{2.2.13}$$

2. the map $(\mu, x, y) \mapsto \|\mathscr{D}_m^2 \mathrm{H}(\mu, x, y)\|_2$ is bounded ($\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} < +\infty$) and $\lambda > \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}$.

Assumptions of under-damped dynamics

Assumptions 2.2.3. Assume that

— $[\mathrm{A}]_1 < +\infty$, $\mathrm{B}: p \mapsto \lambda_{\mathrm{B}} p + \mathrm{D}(p)$ with $\lambda_{\mathrm{B}} \geq 0$ and $[\mathrm{D}]_1 < +\infty$.

— There exists $\lambda_{\mathrm{A}} \geq 0$ such that for all $(v, w) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$\langle \mathrm{A}(v) - \mathrm{A}(w), v - w \rangle \geq \lambda_{\mathrm{A}} |v - w|^2. \tag{2.2.14}$$

— $[\mathscr{D}_m \mathrm{H}]_{1,\infty} < +\infty$ and $\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} < +\infty$.

2.3 Main theorems

2.3.1 First order case

Theorem 2.3.1 (Contraction of Φ)**.** *If the energy functional* H *satisfies assumptions 2.2.1, then*

$$\forall\, p \geq 0, \quad \Phi : \quad \mathscr{P}_p(\mathbb{R}^d) \longrightarrow \mathscr{P}_p(\mathbb{R}^d), \tag{2.3.1}$$

and for

$$\underline{\lambda} := \inf_{\mu,x} \lambda_{\nabla \mathscr{D}_m \mathrm{H}(\mu,x)} \quad (> \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}), \tag{2.3.2}$$

for all $p \geq 1$, Φ defines a strict contraction of $(\mathscr{P}_p(\mathbb{R}^d), \mathscr{W}_1)$:

$$\forall (\mu,\nu) \in \mathscr{P}_p(\mathbb{R}^d) \times \mathscr{P}_p(\mathbb{R}^d), \quad \mathscr{W}_1(\Phi(\mu),\Phi(\nu)) \leq \frac{\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}}{\underline{\lambda}} \mathscr{W}_1(\mu,\nu). \tag{2.3.3}$$

In this case, Φ admits at most one fixed point on $\mathscr{P}_p(\mathbb{R}^d)$: if $p = 1$, there is a unique fixed point.

Proof. See section 2.6.1. ■

Theorem 2.3.2 ($\mathscr{W}_2$-Exponential contraction by coupling)**.** *If the energy functional* H *satisfies assumptions 2.2.2, for all $(\mathrm{X}_t)_{t\geq 0}$ and $(\mathrm{Y}_t)_{t\geq 0}$ solutions of the mean field Langevin equation eq. (2.1.1) with initial conditions $\mu_0^{\mathrm{X}} \in \mathscr{P}_2(\mathbb{R}^d)$ and $\mu_0^{\mathrm{Y}} \in \mathscr{P}_2(\mathbb{R}^d)$, we have*

— *for all $t \geq 0$, $\mu_t^{\mathrm{X}} \in \mathscr{P}_2(\mathbb{R}^d)$;*

— *for all $t \geq 0$,*

$$\mathscr{W}_2(\mu_t^{\mathrm{X}}, \mu_t^{\mathrm{Y}}) \leq e^{(\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - \lambda)t} \mathscr{W}_2(\mu_0^{\mathrm{X}}, \mu_0^{\mathrm{Y}}). \tag{2.3.4}$$

In particular, there exists a unique invariant measure $\mu_\infty \in \mathscr{P}_2(\mathbb{R}^d)$ and

$$\forall\, t \geq 0, \quad \mathscr{W}_2(\mu_t^{\mathrm{X}}, \mu_\infty) \leq e^{(\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - \lambda)t} \mathscr{W}_2(\mu_0^{\mathrm{X}}, \mu_\infty). \tag{2.3.5}$$

Proof. See section 2.6.1. ■

Theorem 2.3.3 (Uniform in time propagation of chaos for mean field Langevin dynamics)**.** *If the energy functional* H *satisfies assumptions 2.2.2, for $(\widetilde{\mathrm{X}}^{(n)})_{n\geq 1}$ a sequence of independent*

copies of the mean field Langevin process $(X_t)_{t\geq 0}$ *synchronously coupled to particles of the over-damped mean field Langevin system, we have*

$$\forall T\geq 0 \quad \forall t\in[0,T], \quad \mathbb{E}[\|X_t^{(1)}-\widetilde{X}_t^{(1)}\|^2]\leq \begin{cases} \frac{\alpha_1(T)}{\beta_1}(e^{\beta_1 t}-1) & \text{if} \quad \beta_1\neq 0; \\ \alpha_1(T)\,t & \text{if} \quad \beta_1=0. \end{cases} \tag{2.3.6}$$

$$\alpha_1(T):=\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} C(d)\Big(\sup_{t\in[0,T]}\Big\langle \mu_t^X, \|\cdot\|_2^2\Big\rangle\Big)\delta_d(n); \tag{2.3.7}$$

$$\delta_d(n):=\begin{cases} \frac{1}{\sqrt{n}} & \text{if} \quad d<4; \\ \frac{\ln(n+1)}{\sqrt{n}} & \text{if} \quad d=4; \\ \frac{1}{n^{\frac{2}{d}}} & \text{if} \quad d>4; \end{cases} \tag{2.3.8}$$

$$\beta_1:=3\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}-2\lambda; \tag{2.3.9}$$

$$\alpha:=2d+2\sup_{\mu}\Big\|\nabla\frac{\delta H}{\delta m}(\mu,0)\Big\|_2; \tag{2.3.10}$$

$$\beta:=-2\Big(\lambda-\sup_{\mu}\Big\|\nabla\frac{\delta H}{\delta m}(\mu,0)\Big\|_2\Big). \tag{2.3.11}$$

Moreover, this upper bound is uniform in time if, and only if, $\beta<0$ *and* $\beta_1<0$ *:*

$$\sup_{t\geq 0}\mathbb{E}[\|X_t^{(1)}-\widetilde{X}_t^{(1)}\|^2]\leq \frac{\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} C(d)}{2\lambda-3\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}}\Big(\Big\langle \mu_0^X, \|\cdot\|_2^2\Big\rangle-\frac{\alpha}{\beta}\Big)\delta_d(n). \tag{2.3.12}$$

Proof. See section 2.6.1. ∎

2.3.2 Kinetic case

Theorem 2.3.4 (Kinetic contraction by coupling)**.** *If the fields* A *and* B *and the energy functional* H *satisfy assumptions 2.2.3, for all positive* $[A]_1$, λ_A *and* λ_B*, there exists a positive constant c such that, if*

$$0\leq \min([D]_1,[\mathscr{D}_m H]_{1,\infty},\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty})\leq \max([D]_1,[\mathscr{D}_m H]_{1,\infty},\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty})<c, \tag{2.3.13}$$

then there exist positive constants C *and* C′ *such that, for all* $(P_t,V_t)_{t\geq 0}$ *and* $(\overline{P}_t,\overline{V}_t)_{t\geq 0}$ *two solutions of the kinetic mean-field Langevin equation eq. (2.1.2) with initial conditions* $\mathbb{P}_{(P_0,V_0)}\in\mathscr{P}_2(\mathbb{R}^d\times\mathbb{R}^d)$ *and* $\mathbb{P}_{(\overline{P}_0,\overline{V}_0)}\in\mathscr{P}_2(\mathbb{R}^d\times\mathbb{R}^d)$*, we have*

$$\forall t\geq 0, \quad \mathscr{W}_2(\mathbb{P}_{(P_t,V_t)},\mathbb{P}_{(\overline{P}_t,\overline{V}_t)})\leq C' e^{-Ct}\mathscr{W}_2(\mathbb{P}_{(P_0,V_0)},\mathbb{P}_{(\overline{P}_0,\overline{V}_0)}). \tag{2.3.14}$$

Moreover, there exists a unique stationary solution ξ_∞ *and*

$$\forall\, t \geq 0, \quad \mathscr{W}_2(\mathbb{P}_{(\mathrm{P}_t,\mathrm{V}_t)}, \xi_\infty) \leq \mathrm{C}' e^{-\mathrm{C}t} \mathscr{W}_2(\mathbb{P}_{(\mathrm{P}_0,\mathrm{V}_0)}, \xi_\infty). \tag{2.3.15}$$

Proof. See section 2.6.2. ■

Theorem 2.3.5 (Uniform in time propagation of chaos for kinetic mean field Langevin dynamics)**.** *Let* $((\overline{\mathrm{P}}^{(n)}, \overline{\mathrm{V}}^{(n)}))_{n\geq 1}$ *a sequence of independent copies of the kinetic mean field Langevin process* $((\mathrm{P}_t, \mathrm{V}_t))_{t\geq 0}$ *synchronously coupled to particles of the under-damped mean field Langevin system with initial datum* $(\mathrm{P}_0^{(i)}, \mathrm{V}_0^{(i)})$ *for* $i \in \{1, \ldots, n\}$ *be* n *independent* $\mathbb{R}^{2d}$*-valued random variables with law* $f_0 \in \mathscr{P}_2(\mathbb{R}^{2d})$*. If the fields* A *and* B *and the energy functional* H *satisfy assumptions 2.2.3, for all positive* $[\mathrm{A}]_1$, λ_A *and* λ_B*, there exists a positive constant* c *such that, if*

$$0 \leq \min([\mathrm{D}]_1, [\mathscr{D}_m \mathrm{H}]_{1,\infty}, \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}) \leq \max([\mathrm{D}]_1, [\mathscr{D}_m \mathrm{H}]_{1,\infty}, \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}) < c, \tag{2.3.16}$$

then there exists a positive constant C*, independent of* n *, such that for* $i \in \{1, \ldots, n\}$

$$\sup_{t\geq 0} \mathbb{E}\Big[|\mathrm{P}_t^{(i)} - \overline{\mathrm{P}}_t^{(i)}|^2 + |\mathrm{V}_t^{(i)} - \overline{\mathrm{V}}_t^{(i)}|^2 \Big] \leq \mathrm{C} \begin{cases} \frac{1}{\sqrt{n}} & \text{if} \quad d < 4; \\ \frac{\ln(n+1)}{\sqrt{n}} & \text{if} \quad d = 4; \\ \frac{1}{n^{\frac{2}{d}}} & \text{if} \quad d > 4. \end{cases} \tag{2.3.17}$$

Here the constant C *depends only on the coefficients of the equation and the second moment of* f_0*.*

Proof. See section 2.6.2. ■

2.3.3 Comments on drift and assumptions

In theorem 2.3.2, theorem 2.3.3, theorem 2.3.4 and theorem 2.3.5, we can replace the intrinsic derivative by a more general drift

$$b: \quad \mathbb{R}^d \times \mathscr{P}_p(\mathbb{R}^d) \longrightarrow \mathbb{R}^d. \tag{2.3.18}$$

In the first order case, we can replace assumptions 2.2.2 by the following conditions on the new drift b

$$\sup_{\mu} b(0,\mu) < +\infty; \tag{2.3.19}$$

$$\exists \lambda > \beta \geq 0 \quad \forall (x,y,\mu,\nu), \quad \langle b(x,\mu) - b(y,\nu), x-y\rangle \geq \lambda |x-y|^2 - \beta |x-y| \mathscr{W}_p(\mu,\nu). \tag{2.3.20}$$

As for the proofs, it is enough to mimic those with the intrinsic derivative (section 2.6.1) which is why we do them without loss of generality with the intrinsic derivative.

In kinetic case, we can replace assumptions 2.2.3 by the same conditions on the velocity field A and the external confinement field B and the following Lipschtz conditions on the new drift b

$$\exists \gamma_1, \gamma_2 \geq 0, \qquad \forall (x,y,\mu,\nu), \quad |b(x,\mu) - b(y,\mu)| \leq \gamma_1 |x-y|; \tag{2.3.21}$$

$$|b(x,\mu) - b(x,\nu)| \leq \gamma_2 \mathscr{W}_p(\mu,\nu). \tag{2.3.22}$$

As in the previous case, we carry out the proofs with the intrinsic derivative without loss of generality.

2.4 Examples

In this section, we give examples of flat energies for which our results are satisfied.

2.4.1 Two-body interaction: granular media flat energy

In this case, we consider the following flat energy

$$\mathrm{H}_{\mathrm{VW}} : \quad \mu \longmapsto \int \mathrm{V}(x)\mu(dx) + \frac{1}{2}\int \mathrm{W}(x-y)\mu(dx)\mu(dy) \tag{2.4.1}$$

with $\mathrm{V} \in \mathscr{C}^2(\mathbb{R}^d, \mathbb{R})$ called confinement potential and $\mathrm{W} \in \mathscr{C}^2(\mathbb{R}^d, \mathbb{R})$ called interaction potential such that for all $x \in \mathbb{R}$, $\mathrm{W}(-x) = \mathrm{W}(x)$.

Proposition 2.4.1 (Two-body interaction and assumptions 2.2.1)**.** *If* $\|\nabla^2 \mathrm{W}\|_{\mathbf{op},\infty} < +\infty$ *and*

$$\inf_{(\mu,x)} \lambda_{\nabla^2 \mathrm{V}(x) + \nabla^2 \mathrm{W} * \mu(x)} > \|\nabla^2 \mathrm{W}\|_{\mathbf{op},\infty}, \tag{2.4.2}$$

H_{VW} *satisfies assumptions 2.2.1.*

Proof. See section 2.6.3. ■

Remark 2.4.1 (About one-dimensional example). In this case, the contraction conditions become

$$\|\mathrm{W}''\|_\infty < +\infty;$$
$$\forall(\mu, x), \quad \mathrm{V}''(x) + \int \mathrm{W}''(x-y)\mu(dy) > \|\mathrm{W}''\|_\infty.$$

We have

$$\mathrm{V}''(x) + \int \mathrm{W}''(x-y)\mu(dy) \geq \mathrm{V}''(x) - \|\mathrm{W}''\|_\infty.$$

In particular, if for all $x \in \mathbb{R}$, $\mathrm{V}''(x) > 2\|\mathrm{W}''\|_\infty$, we deduce that

$$\forall(\mu, x), \quad \mathrm{V}''(x) + \int \mathrm{W}''(x-y)\mu(dy) > \|\mathrm{W}''\|_\infty.$$

For example, we can choose the following polynomial confinement potentials:

▷ $\mathrm{V}(x) := ax^2 + bx + c$ with a, b and c constants such that $a > \|\mathrm{W}''\|_\infty$ and $(b, c) \in \mathbb{R}^2$.

▷ $\mathrm{V}(x) := ax^4 + bx^3 + cx^2 + dx + e$ with a, b, c, d and e constants such that $a > 0$, $c - \frac{3b^2}{8a} > \|\mathrm{W}''\|_\infty$ and $(d, e) \in \mathbb{R}^2$.

Proposition 2.4.2 (Two-body interaction and assumptions 2.2.2)**.** *If* $\|\nabla \mathrm{W}\|_{2,+\infty} < +\infty$, $\|\nabla^2 \mathrm{W}\|_{\mathbf{op},\infty} < +\infty$ *and there exists* $\lambda > 0$ *such that for all* (μ, x, y),

$$\Big\langle \nabla \mathrm{V}(x) - \nabla \mathrm{V}(y) + \nabla \mathrm{W} * \mu(x) - \nabla \mathrm{W} * \mu(y), x - y \Big\rangle \geq \lambda \|x - y\|_2^2; \tag{2.4.3}$$
$$\lambda > \|\nabla^2 \mathrm{W}\|_{\mathbf{op},\infty}; \tag{2.4.4}$$

H_{VW} *satisfies assumptions 2.2.2.*

Proof. See section 2.6.3. ■

Remark 2.4.2. If for all (x,y),

$$\Big\langle \nabla V(x) - \nabla V(y), x-y \Big\rangle \geq \lambda_V \|x-y\|_2^2; \tag{2.4.5}$$

$$\Big\langle \nabla W(x) - \nabla W(y), x-y \Big\rangle \geq 0; \tag{2.4.6}$$

we have

$$\begin{aligned}\Big\langle \nabla V(x) - \nabla V(y) + \nabla W * \mu(x) - \nabla W * \mu(y), x-y \Big\rangle &= \Big\langle \nabla V(x) - \nabla V(y), x-y \Big\rangle \\ &+ \int \Big\langle \nabla W(x-z) - \nabla W(y-z), x-z-(y-z) \Big\rangle \mu(dz) \\ &\geq \lambda_V \|x-y\|_2^2.\end{aligned} \tag{2.4.7}$$

2.4.2 Many-body interaction: polynomial flat energy

In this case, we consider the following flat energy

$$H_{VW}^N : \quad \mu \longmapsto \int V d\mu + \sum_{k=2}^N \int W^{(k)} d\mu^{\otimes k}, \tag{2.4.8}$$

where $\forall k \in \{2,\dots,N\}$, $W^{(k)}$ is a *symmetric interaction potential* between k particles and N represents the number of such potentials.

Proposition 2.4.3 (Many-body interaction and assumptions 2.2.1)**.** *If for all* $k \in \{2,\dots,N\}$, $\|\nabla^2_{x_1,x_2} W^{(k)}\|_{\mathbf{op},\infty} < +\infty$ *and*

$$\inf_{(\mu,x)} \lambda_{\nabla^2 V(x) + \sum_{k=2}^N k \nabla^2_{x_1} W^{(k)} * \mu^{\otimes k-1}(x)} > \sum_{k=2}^N k(k-1) \|\nabla^2_{x_1,x_2} W^{(k)}\|_{\mathbf{op},\infty}, \tag{2.4.9}$$

H_{VW}^N *satisfies assumptions 2.2.1.*

Proof. See section 2.6.3. ■

Proposition 2.4.4 (Many-body interaction and assumptions 2.2.2)**.** *If for all* $k \in \{2,\dots,N\}$, $\|\nabla_{x_1} W^{(k)}(0,\cdot)\|_{2,+\infty} < +\infty$, $\|\nabla^2_{x_1,x_2} W^{(k)}\|_{\mathbf{op},\infty} < +\infty$ *and there exists* $\lambda > 0$ *such that for all*

(μ, x, y),

$$\Big\langle \nabla V(x) - \nabla V(y) + \sum_{k=2}^{N} k\Big(\nabla_{x_1} W^{(k)} * \mu^{\otimes k-1}(x) - \nabla_{x_1} W^{(k)} * \mu^{\otimes k-1}(y)\Big), x-y \Big\rangle \geq \lambda \|x-y\|_2^2; \tag{2.4.10}$$

$$\lambda > \sum_{k=2}^{N} k(k-1)\|\nabla^2_{x_1,x_2} W^{(k)}\|_{\mathbf{op},\infty}; \tag{2.4.11}$$

H^N_{VW} *satisfies assumptions 2.2.2.*

Proof. See section 2.6.3. ■

Remark 2.4.3. Note that the *granular-media flat energy* is a particular case of the polynomial energy:

$$H_{VW}(\mu) = \int V(x)\mu(dx) + \int W^{(2)}(x,y)\mu(dx)\mu(dy); \tag{2.4.12}$$

$$W^{(2)}(x,y) := \frac{1}{2} W(x-y). \tag{2.4.13}$$

We deduce from this that proposition 2.4.3 and proposition 2.4.4 generalize proposition 2.4.1 and proposition 2.4.2.

2.4.3 Internal flat energy

In this case, we consider the following flat energy

$$H_\psi : \quad \mu \longmapsto \psi(\langle \mu, W\rangle) \tag{2.4.14}$$

with $\psi \in \mathscr{C}^2(\mathbb{R},\mathbb{R})$ and $W \in \mathscr{C}^2(\mathbb{R}^d,\mathbb{R})$.

Proposition 2.4.5 (Internal flat energy and assumptions 2.2.1)**.** *If* $\sup_\mu |\psi''(\langle \mu, W\rangle)| < +\infty$, $\|\nabla^2(W\otimes W)\|_{\mathbf{op},\infty} < +\infty$ *and*

$$\inf_{(\mu,x)} \lambda_{\psi'(\langle \mu,W\rangle)\nabla^2 W(x)} > \|\nabla^2(W\otimes W)\|_{\mathbf{op},\infty} \sup_\mu |\psi''(\langle \mu, W\rangle)|, \tag{2.4.15}$$

H_ψ *satisfies assumptions 2.2.1.*

Proof. See section 2.6.3. ■

Proposition 2.4.6 (Internal flat energy and assumptions 2.2.2)**.** *If* $\sup_\mu |\psi'(\langle \mu, W\rangle)| < +\infty$, $\sup_\mu |\psi''(\langle \mu, W\rangle)| < +\infty$, $\|\nabla^2(W\otimes W)\|_{\mathbf{op},\infty} < +\infty$ *and there exists* $\lambda > 0$ *such that for all* (μ, x, y),

$$\psi'(\langle \mu, W\rangle)\Big\langle \nabla W(x) - \nabla W(y), x - y\Big\rangle \geq \lambda \|x-y\|_2^2; \tag{2.4.16}$$

$$\lambda > \|\nabla^2(W\otimes W)\|_{\mathbf{op},\infty} \sup_\mu |\psi''(\langle \mu, W\rangle)|; \tag{2.4.17}$$

H_ψ *satisfies assumptions 2.2.2.*

Proof. See section 2.6.3. ■

2.5 Connection with partial differential equations

The mean field Langevin dynamics corresponds to nonlinear Fokker-Planck equation defined on $[0,+\infty[\times\mathbb{R}^d$ for the flow of probability measures $(\mu_t)_{t\geq 0}$:

$$\frac{d}{dt}\mu_t = \Delta\mu_t + \nabla\cdot(\mu_t \mathscr{D}_m H(\mu_t,\cdot)). \tag{2.5.1}$$

In the kinetic case, we have

$$\frac{d}{dt}f_t + v\cdot\nabla_p f_t - \mathscr{D}_m H(f_t,p)\cdot\nabla_v f_t = \Delta_v f_t + \nabla_v\cdot\Big((A(v)+B(p))f_t\Big). \tag{2.5.2}$$

The particle system corresponds to a Liouville or Fokker-Planck equation defined on $[0,+\infty[\times\mathbb{R}^{nd}$ for the flow of probability measures defined by

$$\forall t\geq 0,\quad \mu_t^{(n)} := \mathbb{P}_{\mathbf{X}_t}; \tag{2.5.3}$$

$$H_n(\mathbf{x}) := nH(\mu_{\mathbf{x}}); \tag{2.5.4}$$

$$\frac{d}{dt}\mu_t^{(n)} = \Delta\mu_t^{(n)} + \nabla\cdot(\mu_t^{(n)}\nabla H_n). \tag{2.5.5}$$

In the kinetic case, we have the kinetic Fokker-Planck equation :

$$\frac{d}{dt}f_t^{(n)} + \mathbf{v}\cdot\nabla_{\mathbf{p}} f_t^{(n)} - \nabla H_n(\mathbf{p})\cdot\nabla_{\mathbf{v}} f_t^{(n)} = \Delta_{\mathbf{v}} f_t^{(n)} + \sum_{i=1}^n \nabla_{v_i}\cdot\Big((A(v_i)+B(p_i))f_t^{(n)}\Big). \tag{2.5.6}$$

Theorem 2.5.1 (Characterization of invariant distributions)**.** *Assume that for all* $\mu\in\mathscr{P}(\mathbb{R}^d)$,

$x \in \mathbb{R}^d \mapsto \frac{\delta \mathrm{H}}{\delta m}(\mu, x)$ *is differentiable and*

$$\mathrm{Z}_\mu := \int_{\mathbb{R}^d} e^{-\frac{\delta \mathrm{H}}{\delta m}(\mu,x)} dx < +\infty. \tag{2.5.7}$$

The following assertions are equivalent:

1. μ_∞ *is a Maxwellian of the McKean-Vlasov PDE:*

$$\mu_\infty \in \{\nu, \quad \Delta \nu + \nabla \cdot (\nu \mathscr{D}_m \mathrm{H}(\nu, \cdot)) = 0\} \tag{2.5.8}$$

2. μ_∞ *is a critical point for the free energy:*

$$\mu_\infty \in \left\{\nu, \quad \mathscr{D}_m \mathrm{H}(\nu, \cdot) + \nabla \ln \frac{d\nu}{dx} = 0\right\}. \tag{2.5.9}$$

3. μ_∞ *is a fixed point of* Φ*:*

$$\mu_\infty \in \{\nu, \quad \Phi(\nu) = \nu\}. \tag{2.5.10}$$

Proof. See section 2.6.3. ■

In the kinetic case, we recall that if the velocity and external confinement fields do not derive from potential gradients, the invariant probabilities are not explicit : in this paper, we will use [40, Lemma 10].

2.6 Proof of main theorems

2.6.1 Proof of main results of the first order case

Proof of theorem 2.3.1. We first check that $\Phi(\mathscr{P}_p(\mathbb{R}^d)) \subset \mathscr{P}_p(\mathbb{R}^d)$. In assumptions 2.2.1, the uniform lower bound of the smallest eigenvalue of the Hessian of the first flat derivative implies that

$$\exists c_1^{\mathrm{H}} > 0 \quad \exists c_2^{\mathrm{H}} \in \mathbb{R} \quad \forall x \in \mathbb{R}^d \quad \forall \mu \in \mathscr{P}_p(\mathbb{R}^d), \quad \frac{\delta \mathrm{H}}{\delta m}(\mu, x) \geq c_1^{\mathrm{H}} \|x\|_2^2 + c_2^{\mathrm{H}}. \tag{2.6.1}$$

We deduce that

$$\begin{aligned}\langle \Phi(\mu), \|\cdot\|^p \rangle &= \frac{1}{Z_\mu} \int_{\mathbb{R}^d} \|x\|^p e^{-\frac{\delta H}{\delta m}(\mu,x)} dx \\ &\leq \frac{e^{-c_2^H}}{Z_\mu} \int_{\mathbb{R}^d} \|x\|^p e^{-c_1^H \|x\|^2} dx,\end{aligned} \tag{2.6.2}$$

and this last integral is finite since

$$\exists M^H > 0 \quad \forall x \in \mathbb{R}^d, \quad \|x\|^p e^{-c_1^H \|x\|^2} \leq M^H e^{-\frac{c_1^H}{2}\|x\|^2}. \tag{2.6.3}$$

Now, let's prove the contraction of Φ. Let $\nu_0, \nu_1 \in \mathscr{P}_p(\mathbb{R}^d)$ and set $\nu_t = (1-t)\nu_0 + t\nu_1$, $t \in [0,1]$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a 1-Lipschitz smooth function. From the very definition of Φ,

$$\langle \Phi(\nu_t), f \rangle = \frac{1}{Z_{\nu_t}} \int_{\mathbb{R}^d} f(x) e^{-\frac{\delta H}{\delta m}(\nu_t, x)} dx, \tag{2.6.4}$$

so that setting

$$g_t(x) := -\frac{d}{dt} \frac{\delta H}{\delta m}(\nu_t, x), \tag{2.6.5}$$

we get

$$\begin{aligned}\frac{d}{dt}\langle \Phi(\nu_t), f \rangle &= -\frac{\partial_t Z_{\nu_t}}{Z_{\nu_t}} \langle \Phi(\nu_t), f \rangle + \langle \Phi(\nu_t), f g_t \rangle \\ &= \langle \Phi(\nu_t), f g_t \rangle - \langle \Phi(\nu_t), f \rangle \langle \Phi(\nu_t), g_t \rangle \\ &= \mathrm{Cov}_{\Phi(\nu_t)}(f, g_t).\end{aligned} \tag{2.6.6}$$

As the functional H admits first and second-order flat derivatives, we have

$$g_t(x) := -\int_{\mathbb{R}^d} \frac{\delta^2 H}{\delta m^2}(\nu_t, x, y)(\nu_1 - \nu_0)(dy). \tag{2.6.7}$$

For a probability μ, let $\mathscr{L}_\mu$ be the operator defined on $\mathscr{C}^2$-functions by

$$\mathscr{L}_\mu f = \Delta f - \mathscr{D}_m H(\mu, \cdot) \cdot \nabla f. \tag{2.6.8}$$

Denoting by φ_t be the solution of the Poisson equation $f - \Phi(\nu_t)(f) = \mathscr{L}_{\nu_t}\varphi_t$ and using that

$\mathscr{L}_{\nu_t}$ is self-adjoint in $L^2(\Phi(\nu_t))$, we get

$$\mathrm{Cov}_{\Phi(\nu_t)}(f,g_t) = \langle \varphi_t, \mathscr{L}_{\nu_t} g_t\rangle_{L^2(\Phi(\nu_t))} = -\langle \nabla\varphi_t, \nabla g_t\rangle_{L^2(\Phi(\nu_t))}.$$

Note that for the second equality, we used the fact that for some $\mathscr{C}^2$-functions f and g, $\mathscr{L}_{\nu_t}(f.g) = f\mathscr{L}_{\nu_t} g + g\mathscr{L}_{\nu_t} f + 2\nabla f.\nabla g$. With the help of Cauchy-Schwarz inequality, this leads to

$$\begin{aligned} |\mathrm{Cov}_{\Phi(\nu_t)}(f,g_t)| &\le \|\nabla\varphi_t\|_{L^2(\Phi(\nu_t))}\|\nabla g_t\|_{L^2(\Phi(\nu_t))} \\ &\le \|\nabla\varphi_t\|_\infty \|\nabla g_t\|_\infty. \end{aligned} \tag{2.6.9}$$

Control of $\|\nabla\varphi_t\|_\infty$. Let $h:\mathbb{R}^d \to \mathbb{R}$ be a $\mathscr{C}^1$ function with bounded derivative. Let φ be a unique solution of the Poisson equation $h - \langle\Phi(\mu), h\rangle = \mathscr{L}_\mu\varphi$. It is well-known that $\varphi(x) = \int_0^{+\infty} P_t^\mu h(x) - \langle\Phi(\mu), h\rangle dt$ where $(P_t^\mu)_{t\ge 0}$ denotes the semi-group associated with $\mathscr{L}_\mu$ so that under adequate derivation conditions (which will be satisfied in our setting),

$$\nabla\varphi(x) = \int_0^{+\infty} \nabla P_t^\mu h(x) dt.$$

Indeed, it is sufficient to verify the conditions of application of the Lebesgue dominated convergence theorem for the derivation under the integral sign, which is done below.

We have

$$\nabla P_t^\mu h(x) = \mathbb{E}[\nabla h(X_t^{\mu,x})\partial_x X_t^{\mu,x}]$$

where $(X_t^{\mu,x})$ denotes the solution starting from x of the SDE associated with $\mathscr{L}_\mu$ and $(\partial_x X_t^\mu)$ its first variation process solution to

$$dY_t = -\nabla\mathscr{D}_m H(\mu, X_t^{\mu,x})Y_t dt$$

with $Y_0 = I_d$. Under the assumption

$$\underline{\lambda} := \inf_{\mu,x} \lambda_{\nabla\mathscr{D}_m H(\mu,x)} > 0, \tag{2.6.10}$$

one easily deduces from a Gronwall argument that for any $z\in\mathbb{R}^d$,

$$|Y_t z|^2 \le e^{-2\underline{\lambda}t}|z|^2$$

which implies that

$$|\nabla P_t^\mu h(x)| \le [h]_1 \|\partial_x X_t^{\mu,x}\|_{\mathbf{op}} \le [h]_1 e^{-\underline{\lambda} t}$$

where $[h]_1$ denotes the Lipschitz constant of h. Then,

$$\|\nabla \varphi\|_\infty \le [h]_1 \underline{\lambda}^{-1}.$$

In particular, we have

$$\|\nabla \varphi_t\|_\infty \le [f]_1 \underline{\lambda}^{-1} = \frac{1}{\underline{\lambda}}. \tag{2.6.11}$$

Control of $\|\nabla g_t\|_\infty$. Let us finally focus on ∇g_t. First, one checks that

$$\nabla g_t(x) := -\int_{\mathbb{R}^d} \nabla_x \frac{\delta^2 H}{\delta m^2}(\nu_t, x, y)(\nu_1 - \nu_0)(dy). \tag{2.6.12}$$

Now, denoting by (X_{ν_0}, X_{ν_1}) an optimal coupling of ν_0 and ν_1 for the 1-Wasserstein distance, one obtains

$$\begin{aligned} \left|\nabla_x \left(\int_{\mathbb{R}^d} \frac{\delta^2 H}{\delta m^2}(\nu_t, x, y)(\nu_1 - \nu_0)(dy)\right)\right| &\le |\mathbb{E}[\nabla_x \frac{\delta^2 H}{\delta m^2}(\nu_t, x, X_{\nu_1}) - \nabla_x \frac{\delta^2 H}{\delta m^2}(\nu_t, x, X_{\nu_0})]| \\ &\le \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} \mathscr{W}_1(\nu_0, \nu_1). \end{aligned} \tag{2.6.13}$$

Hence,

$$\|\nabla g_t\|_\infty \le \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} \mathscr{W}_1(\nu_0, \nu_1),$$

and by (1.6.25), we get for any smooth 1-Lipschitz function

$$|\langle \Phi(\nu_1), f\rangle - \langle \Phi(\nu_0), f\rangle| \le \int_0^1 |\frac{d}{dt}\langle \Phi(\nu_t), f\rangle| dt \le \frac{\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}}{\underline{\lambda}} \mathscr{W}_1(\nu_0, \nu_1).$$

Since

$$\frac{\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}}{\underline{\lambda}} < 1, \tag{2.6.14}$$

by a density argument and the Kantorovitch-Rubinstein duality relation, it follows that Φ is a contraction on $(\mathscr{P}_p(\mathbb{R}^d), \mathscr{W}_1)$.

Concerning the existence of at most one fixed point, we show it by the absurd: assuming

that we have more than one, we would contradict the strict contraction since we would have

$$\frac{\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}}{\underline{\lambda}} \geq 1. \tag{2.6.15}$$

For $p = 1$, existence follows from Picard's fixed point theorem since Φ is a strict contraction of $(\mathscr{P}_1(\mathbb{R}^d), \mathscr{W}_1)$ which is a complete metric space. ■

Proof of theorem 2.3.2. Concerning the moments, we have

$$\begin{aligned}
\left\langle \frac{d}{dt}\mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle &= \left\langle \mu_t^{\mathrm{X}}, \mathscr{L}_{\mu_t^{\mathrm{X}}} \|\cdot\|_2^2 \right\rangle \\
&= \int \left(2d - 2\nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, x)\cdot x\right)\mu_t^{\mathrm{X}}(dx) \\
&\leq 2d - 2\left[\left(\lambda - \left\|\nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, 0)\right\|_2\right)\left\langle \mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle - \left\|\nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, 0)\right\|_2\right] \\
&\leq 2d + 2\left\|\nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, 0)\right\|_2 - 2\left(\lambda - \left\|\nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, 0)\right\|_2\right)\left\langle \mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle.
\end{aligned} \tag{2.6.16}$$

We deduce that there exists $(\alpha, \beta) \in \mathbb{R}^2$ such that

$$\left\langle \frac{d}{dt}\mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle \leq \alpha + \beta \left\langle \mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle. \tag{2.6.17}$$

By Gronwall's lemma, we have

$$\forall\, t \geq 0, \quad \left\langle \mu_t^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle \leq \begin{cases} \left\langle \mu_0^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle e^{\beta t} + \frac{\alpha}{\beta}(e^{\beta t} - 1) & \text{if} \quad \beta \neq 0; \\ \left\langle \mu_0^{\mathrm{X}}, \|\cdot\|_2^2 \right\rangle + \alpha t & \text{if} \quad \beta = 0. \end{cases} \tag{2.6.18}$$

As for the exponential contraction, to do this, we consider the coupling given by

$$\begin{cases} d\mathrm{X}_t &= -\nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, \mathrm{X}_t)\, dt + \sqrt{2} d\mathrm{B}_t \\ d\mathrm{Y}_t &= -\nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{Y}}, \mathrm{Y}_t)\, dt + \sqrt{2} d\mathrm{B}_t. \end{cases} \tag{2.6.19}$$

We have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|X_t-Y_t\|_2^2 &= \left\langle \frac{d}{dt}(X_t-Y_t), X_t-Y_t\right\rangle \\ &= -\left\langle \nabla\frac{\delta H}{\delta m}(\mu_t^X,X_t)-\nabla\frac{\delta H}{\delta m}(\mu_t^Y,Y_t), X_t-Y_t\right\rangle \\ &= -\left\langle \nabla\frac{\delta H}{\delta m}(\mu_t^Y,X_t)-\nabla\frac{\delta H}{\delta m}(\mu_t^Y,Y_t), X_t-Y_t\right\rangle \\ &\quad -\left\langle \nabla\frac{\delta H}{\delta m}(\mu_t^X,X_t)-\nabla\frac{\delta H}{\delta m}(\mu_t^Y,X_t), X_t-Y_t\right\rangle.\end{aligned} \tag{2.6.20}$$

On the one hand, according to (**i**), we have

$$-\left\langle \nabla\frac{\delta H}{\delta m}(\mu_t^Y,X_t)-\nabla\frac{\delta H}{\delta m}(\mu_t^Y,Y_t), X_t-Y_t\right\rangle \le -\lambda\|X_t-Y_t\|_2^2. \tag{2.6.21}$$

On the other hand, by definition of the flat derivative and according to (**ii**), we have

$$\begin{aligned}&-\left\langle \nabla\frac{\delta H}{\delta m}(\mu_t^X,X_t)-\nabla\frac{\delta H}{\delta m}(\mu_t^Y,X_t), X_t-Y_t\right\rangle \\ &= -\int_0^1\int\left\langle \nabla_x\frac{\delta^2 H}{\delta m^2}(s\mu_t^X+(1-s)\mu_t^Y,X_t,y), X_t-Y_t\right\rangle(\mu_t^X-\mu_t^Y)(dy)ds \\ &\le \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}\|X_t-Y_t\|_2\mathscr{W}_1(\mu_t^X,\mu_t^Y) \\ &\le \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}\|X_t-Y_t\|_2\mathbb{E}[\|X_t-Y_t\|_2].\end{aligned} \tag{2.6.22}$$

We deduce that

$$\frac{1}{2}\frac{d}{dt}\|X_t-Y_t\|_2^2 \le -\lambda\|X_t-Y_t\|_2^2+\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}\|X_t-Y_t\|_2\mathbb{E}[\|X_t-Y_t\|_2]; \tag{2.6.23}$$

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\mathbb{E}[\|X_t-Y_t\|_2^2] &\le -\lambda\mathbb{E}[\|X_t-Y_t\|_2^2]+\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}\Big(\mathbb{E}[\|X_t-Y_t\|_2]\Big)^2 \\ &\le (\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}-\lambda)\mathbb{E}[\|X_t-Y_t\|_2^2] \quad \text{Jensen's inequality.}\end{aligned} \tag{2.6.24}$$

According to Gronwall's lemma, we have

$$\forall\, t\ge 0,\quad \mathbb{E}[\|X_t-Y_t\|_2^2]\le \mathbb{E}[\|X_0-Y_0\|_2^2]e^{2(\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}-\lambda)t}. \tag{2.6.25}$$

We deduce that

$$\forall\, t\ge 0,\quad \mathscr{W}_2^2(\mu_t^X,\mu_t^Y)\le \mathbb{E}[\|X_t-Y_t\|_2^2]\le \mathbb{E}[\|X_0-Y_0\|_2^2]e^{2(\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}-\lambda)t}. \tag{2.6.26}$$

We conclude that

$$\forall\, t \geq 0, \quad \mathscr{W}_2(\mu_t^{\mathrm{X}}, \mu_t^{\mathrm{Y}}) \leq e^{(\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - \lambda)t} \mathscr{W}_2(\mu_0^{\mathrm{X}}, \mu_0^{\mathrm{Y}}). \tag{2.6.27}$$

■

Proof of theorem 2.3.3. Consider $(\widetilde{\mathrm{X}}^{(n)})_{n\geq 1}$ a sequence of independent copies of the mean field Langevin process $(\mathrm{X}_t)_{t\geq 0}$. By synchronous coupling of the particle system with these copies, we have

$$\begin{aligned}
\frac{d}{dt}\sum_{i=1}^n \|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|^2 &= -2\sum_{i=1}^n \left\langle \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{X}_t}, \mathrm{X}_t^{(i)}) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, \widetilde{\mathrm{X}}_t^{(i)}), \mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)} \right\rangle \qquad (2.6.28)\\
&\leq -2\lambda \sum_{i=1}^n \|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|^2 - 2\sum_{i=1}^n \left\langle \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{X}_t}, \widetilde{\mathrm{X}}_t^{(i)}) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mu_t^{\mathrm{X}}, \widetilde{\mathrm{X}}_t^{(i)}), \mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)} \right\rangle \\
&\leq -2\lambda \sum_{i=1}^n \|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|^2 + 2\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} \mathscr{W}_1(\mu_{\mathbf{X}_t}, \mu_t^{\mathrm{X}}) \sum_{i=1}^n \|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|.
\end{aligned}$$

On the other hand, we have

$$\mathscr{W}_1(\mu_{\mathbf{X}_t}, \mu_t^{\mathrm{X}}) \leq \mathscr{W}_2(\mu_{\mathbf{X}_t}, \mu_t^{\mathrm{X}}); \tag{2.6.29}$$

$$\mathscr{W}_2(\mu_{\mathbf{X}_t}, \mu_t^{\mathrm{X}}) \leq \mathscr{W}_2(\mu_{\mathbf{X}_t}, \mu_{\widetilde{\mathbf{X}}_t}) + \mathscr{W}_2(\mu_{\widetilde{\mathbf{X}}_t}, \mu_t^{\mathrm{X}}). \tag{2.6.30}$$

By Young's inequality, we have

$$\begin{aligned}
\frac{d}{dt}\sum_{i=1}^n \|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|^2 &\leq 2(\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - \lambda) \sum_{i=1}^n \|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|^2 \qquad (2.6.31)\\
&\quad + n\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} \Big(\mathscr{W}_2^2(\mu_{\mathbf{X}_t}, \mu_{\widetilde{\mathbf{X}}_t}) + \mathscr{W}_2^2(\mu_{\widetilde{\mathbf{X}}_t}, \mu_t^{\mathrm{X}}) \Big).
\end{aligned}$$

We have

$$\mathbb{E}[\mathscr{W}_2^2(\mu_{\mathbf{X}_t}, \mu_{\widetilde{\mathbf{X}}_t})] \leq \frac{1}{n}\sum_{i=1}^n \mathbb{E}[\|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|^2]. \tag{2.6.32}$$

By symmetry, we have

$$\forall\, j \in \{1, \ldots, n\}, \quad \mathbb{E}[\|\mathrm{X}_t^{(j)} - \widetilde{\mathrm{X}}_t^{(j)}\|^2] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[\|\mathrm{X}_t^{(i)} - \widetilde{\mathrm{X}}_t^{(i)}\|^2]. \tag{2.6.33}$$

Setting $\gamma_t := \mathbb{E}[\|\mathrm{X}_t^{(1)} - \widetilde{\mathrm{X}}_t^{(1)}\|^2]$, we have

$$\frac{d}{dt}\gamma_t \leq (3\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} - 2\lambda)\gamma_t + \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty} \mathbb{E}[\mathscr{W}_2^2(\mu_{\widetilde{\mathbf{X}}_t}, \mu_t^{\mathrm{X}})]. \tag{2.6.34}$$

By eq. (2.6.18), we have

$$\forall t \geq 0, \quad \left\langle \mu_t^X, \|\cdot\|_2^2 \right\rangle \leq \begin{cases} \left\langle \mu_0^X, \|\cdot\|_2^2 \right\rangle e^{\beta t} + \frac{\alpha}{\beta}(e^{\beta t} - 1) & \text{if} \quad \beta \neq 0; \\ \left\langle \mu_0^X, \|\cdot\|_2^2 \right\rangle + \alpha t & \text{if} \quad \beta = 0. \end{cases} \tag{2.6.35}$$

$$\alpha := 2d + 2\sup_{\mu} \left\| \nabla \frac{\delta H}{\delta m}(\mu, 0) \right\|_2; \tag{2.6.36}$$

$$\beta := -2\Big(\lambda - \sup_{\mu} \left\| \nabla \frac{\delta H}{\delta m}(\mu, 0) \right\|_2\Big). \tag{2.6.37}$$

If $\beta < 0$, we have

$$\sup_{t \geq 0} \left\langle \mu_t^X, \|\cdot\|_2^2 \right\rangle \leq \left\langle \mu_0^X, \|\cdot\|_2^2 \right\rangle - \frac{\alpha}{\beta}. \tag{2.6.38}$$

By the result of Fournier and Guillin ([114, Theorem 1]), we have

$$\mathbb{E}[\mathscr{W}_2^2(\mu_{\widetilde{\mathbf{X}}_t}, \mu_t^X)] \leq C(d)\Big(\left\langle \mu_0^X, \|\cdot\|_2^2 \right\rangle - \frac{\alpha}{\beta}\Big)\delta_d(n); \tag{2.6.39}$$

$$\delta_d(n) := \begin{cases} \frac{1}{\sqrt{n}} & \text{if} \quad d < 4; \\ \frac{\ln(n+1)}{\sqrt{n}} & \text{if} \quad d = 4; \\ \frac{1}{n^{\frac{2}{d}}} & \text{if} \quad d > 4. \end{cases} \tag{2.6.40}$$

By Gronwall's lemma, we have ($\gamma_0 = 0$)

$$\forall t \geq 0, \quad \gamma_t \leq \begin{cases} \frac{\alpha_1}{\beta_1}(e^{\beta_1 t} - 1) & \text{if} \quad \beta_1 \neq 0; \\ \alpha_1 t & \text{if} \quad \beta_1 = 0. \end{cases} \tag{2.6.41}$$

$$\alpha_1 := \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} C(d)\Big(\left\langle \mu_0^X, \|\cdot\|_2^2 \right\rangle - \frac{\alpha}{\beta}\Big)\delta_d(n); \tag{2.6.42}$$

$$\beta_1 := 3\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} - 2\lambda. \tag{2.6.43}$$

In this case, the bound is uniform in time if, and only if, $\beta_1 < 0$:

$$\sup_{t \geq 0} \gamma_t \leq -\frac{\alpha_1}{\beta_1}. \tag{2.6.44}$$

If $\beta = 0$, we have:

$$\sup_{t \in [0,T]} \left\langle \mu_t^X, \|\cdot\|_2^2 \right\rangle \leq \left\langle \mu_0^X, \|\cdot\|_2^2 \right\rangle + \alpha T. \tag{2.6.45}$$

We deduce that

$$\forall t \in [0,T], \quad \gamma_t \leq \begin{cases} \frac{\alpha_2}{\beta_1}(e^{\beta_1 t}-1) & \text{if} \quad \beta_1 \neq 0; \\ \alpha_2 t & \text{if} \quad \beta_1 = 0. \end{cases} \tag{2.6.46}$$

$$\alpha_2 := \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} C(d)\Big(\Big\langle \mu_0^X, \|\cdot\|_2^2 \Big\rangle + \alpha T\Big)\delta_d(n); \tag{2.6.47}$$

$$\beta_1 := 3\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} - 2\lambda. \tag{2.6.48}$$

In this case, since the moments are only bounded in finite time horizon, the propagation of chaos cannot be uniform in time.

If $\beta > 0$, we have

$$\sup_{t\in[0,T]} \Big\langle \mu_t^X, \|\cdot\|_2^2 \Big\rangle \leq \langle \mu_0^X, \|\cdot\|_2^2 \Big\rangle e^{\beta T} + \frac{\alpha}{\beta}(e^{\beta T}-1). \tag{2.6.49}$$

We deduce that

$$\forall t \in [0,T], \quad \gamma_t \leq \begin{cases} \frac{\alpha_3}{\beta_1}(e^{\beta_1 t}-1) & \text{if} \quad \beta_1 \neq 0; \\ \alpha_3 t & \text{if} \quad \beta_1 = 0. \end{cases} \tag{2.6.50}$$

$$\alpha_3 := \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} C(d)\Big(\Big\langle \mu_0^X, \|\cdot\|_2^2 \Big\rangle e^{\beta T} + \frac{\alpha}{\beta}(e^{\beta T}-1)\Big)\delta_d(n); \tag{2.6.51}$$

$$\beta_1 := 3\|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty} - 2\lambda. \tag{2.6.52}$$

In this case, since the moments are only bounded in finite time horizon, the propagation of chaos cannot be uniform in time.

■

2.6.2 Proof of main results of the kinetic case

Proof of theorem 2.3.4. Let $(P_t, V_t)_{t\geq 0}$ and $(\overline{P}_t, \overline{V}_t)_{t\geq 0}$ two solutions of

$$\begin{cases} dP_t &= V_t dt; \\ dV_t &= \sqrt{2} dB_t - A(V_t)dt - B(P_t)dt - \nabla \frac{\delta H}{\delta m}(\mathbb{P}_{P_t}, P_t)dt; \end{cases} \tag{2.6.53}$$

such that $\mathbb{P}_{(\mathrm{P}_0,\mathrm{V}_0)} \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ and $\mathbb{P}_{(\overline{\mathrm{P}}_0,\overline{\mathrm{V}}_0)} \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$. Then, by difference, $(p_t, v_t) := (\mathrm{P}_t - \overline{\mathrm{P}}_t, \mathrm{V}_t - \overline{\mathrm{V}}_t)$ evolves according to

$$\begin{cases} dp_t &= v_t dt; \\ dv_t &= -\Big(\mathrm{A}(\mathrm{V}_t) - \mathrm{A}(\overline{\mathrm{V}}_t) + \lambda_{\mathrm{B}} p_t + \mathrm{D}(\mathrm{P}_t) - \mathrm{D}(\overline{\mathrm{P}}_t) + \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \mathrm{P}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big) dt. \end{cases} \tag{2.6.54}$$

Then, if a and b are positive constants to be chosen later on,

$$\begin{aligned} \frac{d}{dt}\Big(a|p_t|^2 + 2p_t \cdot v_t + b|v_t|^2\Big) &= 2ap_t \cdot v_t + 2|v_t|^2 \\ &\quad - 2p_t \cdot \Big(\mathrm{A}(\mathrm{V}_t) - \mathrm{A}(\overline{\mathrm{V}}_t) + \lambda_{\mathrm{B}} p_t + \mathrm{D}(\mathrm{P}_t) - \mathrm{D}(\overline{\mathrm{P}}_t)\Big) \\ &\quad - 2bv_t \cdot \Big(\mathrm{A}(\mathrm{V}_t) - \mathrm{A}(\overline{\mathrm{V}}_t) + \lambda_{\mathrm{B}} p_t + \mathrm{D}(\mathrm{P}_t) - \mathrm{D}(\overline{\mathrm{P}}_t)\Big) \\ &\quad - 2(p_t + bv_t) \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \mathrm{P}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big). \end{aligned} \tag{2.6.55}$$

By the Cauchy-Schwarz inequality and assumptions on A and D, the first four terms are bounded by above by

$$2(a - \lambda_{\mathrm{B}} b) p_t \cdot v_t + 2([\mathrm{A}]_1 + [\mathrm{D}]_1 b)|p_t| \cdot |v_t| - 2(\lambda_{\mathrm{A}} b - 1)|v_t|^2 - 2(\lambda_{\mathrm{B}} - [\mathrm{D}]_1)|p_t|^2. \tag{2.6.56}$$

As for the last remaining term, we have

$$\begin{aligned} -2(p_t + bv_t) \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \mathrm{P}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big) &= -2(p_t + bv_t) \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \mathrm{P}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t)\Big) \\ &\quad - 2(p_t + bv_t) \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big). \end{aligned} \tag{2.6.57}$$

Then, we have

$$-2p_t \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \mathrm{P}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t)\Big) \le 2[\mathscr{D}_m \mathrm{H}]_{1,\infty}|p_t|^2; \tag{2.6.58}$$

$$-2p_t \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big) \le 2\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}|p_t|\mathbb{E}[|p_t|]; \tag{2.6.59}$$

$$-2bv_t \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \mathrm{P}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t)\Big) \le 2b[\mathscr{D}_m \mathrm{H}]_{1,\infty}|p_t| \cdot |v_t|; \tag{2.6.60}$$

$$-2bv_t \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big) \le 2b\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}|v_t|\mathbb{E}[|p_t|]. \tag{2.6.61}$$

By the identity $2ab \le a^2 + b^2$, we have

$$-2bv_t \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \mathrm{P}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t)\Big) \le b[\mathscr{D}_m \mathrm{H}]_{1,\infty}(|p_t|^2 + |v_t|^2). \tag{2.6.62}$$

By Jensen's inequality and Young's inequality, we have

$$-2\mathbb{E}\Big[p_t \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big)\Big] \le 2\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}\mathbb{E}[|p_t|^2]; \tag{2.6.63}$$

$$-2b\mathbb{E}\Big[v_t \cdot \Big(\nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t}, \overline{\mathrm{P}}_t) - \nabla \frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\overline{\mathrm{P}}_t}, \overline{\mathrm{P}}_t)\Big)\Big] \le b\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}\mathbb{E}[|p_t|^2 + |v_t|^2]. \tag{2.6.64}$$

Collecting all terms leads to the bound

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\Big[a|p_t|^2 + 2p_t \cdot v_t + b|v_t|^2\Big] \le{}& 2\Big(a - \lambda_{\mathrm{B}} b\Big)\mathbb{E}[p_t \cdot v_t] + 2\Big([\mathrm{A}]_1 + [\mathrm{D}]_1 b\Big)\mathbb{E}[|p_t| \cdot |v_t|] \\ &- 2\Big(\lambda_{\mathrm{B}} - [\mathrm{D}]_1 - ([\mathscr{D}_m \mathrm{H}]_{1,\infty} + \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}) \\ &- \frac{b}{2}([\mathscr{D}_m \mathrm{H}]_{1,\infty} + \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty})\Big)\mathbb{E}[|p_t|^2] \\ &- 2\Big(\lambda_{\mathrm{A}} b - 1 - \frac{b}{2}([\mathscr{D}_m \mathrm{H}]_{1,\infty} + \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty})\Big)\mathbb{E}[|v_t|^2].\end{aligned} \tag{2.6.65}$$

For all $a \ge \lambda_{\mathrm{B}} b$, by the Cauchy-Schwarz inequality and the Young's inequality ($2ab \le a^2 + b^2$), we have

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\Big[a|p_t|^2 + 2p_t \cdot v_t + b|v_t|^2\Big] \le{}& \Big(a - \lambda_{\mathrm{B}} b + [\mathrm{A}]_1 + [\mathrm{D}]_1 b \\ &- 2\Big(\lambda_{\mathrm{B}} - [\mathrm{D}]_1 - ([\mathscr{D}_m \mathrm{H}]_{1,\infty} + \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}) \\ &- \frac{b}{2}([\mathscr{D}_m \mathrm{H}]_{1,\infty} + \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty})\Big)\Big)\mathbb{E}[|p_t|^2] \\ &+ \Big(a - \lambda_{\mathrm{B}} b + [\mathrm{A}]_1 + [\mathrm{D}]_1 b \\ &- 2\Big(\lambda_{\mathrm{A}} b - 1 - \frac{b}{2}([\mathscr{D}_m \mathrm{H}]_{1,\infty} + \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty})\Big)\Big)\mathbb{E}[|v_t|^2].\end{aligned} \tag{2.6.66}$$

For all $b > \frac{1}{\sqrt{\lambda_B}}$, $Q_{\lambda_B b,b} : (p,v) \in \mathbb{R}^{2d} \mapsto \lambda_B b|p|^2 + 2p\cdot v + b|v|^2$ is a positive quadratic form on $\mathbb{R}^{2d}$. Then we look for a and b such that

$$\begin{cases} a - \lambda_B b + [A]_1 + [D]_1 b - 2\Big(\lambda_B - [D]_1 - ([\mathscr{D}_m H]_{1,\infty} + \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}) - \frac{b}{2}([\mathscr{D}_m H]_{1,\infty} + \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty})\Big) & < 0; \\ a - \lambda_B b + [A]_1 + [D]_1 b - 2\Big(\lambda_A b - 1 - \frac{b}{2}([\mathscr{D}_m H]_{1,\infty} + \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty})\Big) & < 0; \\ a \geq \lambda_B b. \end{cases} \tag{2.6.67}$$

in such a way that

$$\exists C > 0 \quad \forall t \geq 0, \quad \frac{d}{dt}\mathbb{E}\Big[Q_{a,b}\Big((p_t, v_t)\Big)\Big] \leq -C\mathbb{E}[|p_t|^2 + |v_t|^2]. \tag{2.6.68}$$

Let us set $\gamma := [\mathscr{D}_m H]_{1,\infty} + \|\mathscr{D}_m^2 H\|_{\mathbf{op},\infty}$. If $a = \lambda_B b$, b satisfies

$$\begin{cases} b < \frac{2\lambda_B - 2[D]_1 - [A]_1 - 2\gamma}{[D]_1 + \gamma}; \\ ([D]_1 + \gamma - 2\lambda_A) b < -[A]_1 - 2. \end{cases} \tag{2.6.69}$$

Such b exists given that

$$2\lambda_B - 2[D]_1 - [A]_1 - 2\gamma > 0; \tag{2.6.70}$$

$$[D]_1 + \gamma - 2\lambda_A < 0; \tag{2.6.71}$$

$$\frac{-[A]_1 - 2}{[D]_1 + \gamma - 2\lambda_A} < \frac{2\lambda_B - 2[D]_1 - [A]_1 - 2\gamma}{[D]_1 + \gamma}. \tag{2.6.72}$$

In fact, for every positive ε, we have

$$\frac{d}{dt}\mathbb{E}\Big[\lambda_B b|p_t|^2 + 2p_t\cdot v_t + b|v_t|^2\Big] \leq -\Big(2\lambda_B - 2\eta - \varepsilon - \eta b\Big)\mathbb{E}[|p_t|^2] - \Big((2\lambda_A - \eta)b - 2 - \frac{[A]_1^2}{\varepsilon}\Big)\mathbb{E}[|v_t|^2] \tag{2.6.73}$$

where $\eta := \gamma + [D]_1$. For us to have

$$\begin{cases} 2\lambda_B - 2\eta - \varepsilon - \eta b & > 0; \\ (2\lambda_A - \eta) b - 2 - \frac{[A]_1^2}{\varepsilon} & > 0; \end{cases} \tag{2.6.74}$$

it is necessary that $\eta < 2\lambda_A$, which is assumed in the sequel. Then, for instance for $\varepsilon = \lambda_B$, the conditions are equivalent to

$$\frac{2+\frac{[A]_1^2}{\lambda_B}}{2\lambda_A-\eta} < b < \frac{\lambda_B-2\eta}{\eta} \quad \text{and} \quad \eta < 2\lambda_A. \tag{2.6.75}$$

We look for η such that

$$\frac{2+\frac{[A]_1^2}{\lambda_B}}{2\lambda_A-\eta} < \frac{\lambda_B-2\eta}{\eta}, \tag{2.6.76}$$

that is,

$$2\eta^2 - \eta\Big(2+\frac{[A]_1^2}{\lambda_B}+\lambda_B+4\lambda_A\Big)+2\lambda_A\lambda_B > 0. \tag{2.6.77}$$

This polynomial takes negative values at $\eta = 2\lambda_A$, so it is positive on an interval $[0,\eta_0[$ for some $\eta_0 < 2\lambda_A$. We further notice that

$$\eta_0 < \frac{\lambda_B\sqrt{\lambda_B}}{1+2\sqrt{\lambda_B}}, \tag{2.6.78}$$

so that there exists b with all the above conditions for any $\eta \in [0,\eta_0[$.

Hence there exists a constant η_0, depending only on $[A]_1$, λ_A and λ_B, such that, if $\gamma+[D]_1 < \eta_0$, then there exist a positive quadratic form $Q_{\lambda_B b,b}$ on $\mathbb{R}^{2d}$ and a constant C depending only on $[A]_1$, λ_A, λ_B, γ and $[D]_1$ such that

$$\forall\, t \geq 0, \quad \frac{d}{dt}\mathbb{E}\Big[Q_{\lambda_B b,b}\Big((p_t,v_t)\Big)\Big] \leq -C\mathbb{E}[|p_t|^2+|v_t|^2]. \tag{2.6.79}$$

In turn, since $\sqrt{Q_{\lambda_B b,b}}$ and $(p,v)\mapsto\sqrt{|p|^2+|v|^2}$ are equivalent on $\mathbb{R}^{2d}$, this is bounded by $-2C\mathbb{E}[Q_{\lambda_B b,b}((p_t,v_t))]$ for a new constant $C>0$, so that,

$$\forall\, t \geq 0, \quad \mathbb{E}\Big[Q_{\lambda_B b,b}\Big((P_t,V_t)-(\overline{P}_t,\overline{V}_t)\Big)\Big] \leq e^{-2Ct}\mathbb{E}\Big[Q_{\lambda_B b,b}\Big((P_0,V_0)-(\overline{P}_0,\overline{V}_0)\Big)\Big] \tag{2.6.80}$$

by integration. We deduce by equivalence of norms on the space of configurations and the existence of an optimal coupling that

$$\forall\, t \geq 0, \quad \mathscr{W}_2(\mathbb{P}_{(P_t,V_t)},\mathbb{P}_{(\overline{P}_t,\overline{V}_t)}) \leq C' e^{-Ct}\mathscr{W}_2(\mathbb{P}_{(P_0,V_0)},\mathbb{P}_{(\overline{P}_0,\overline{V}_0)}). \tag{2.6.81}$$

We conclude that there exists a unique equilibrium ([40, Lemma 10]) and that we contract

exponentially towards the latter. ■

Proof of theorem 2.3.5. The time-uniform propagation of chaos requires a time-uniform bound of the second moment of $\mathbb{P}_{(\mathrm{P}_t,\mathrm{V}_t)}$. Let $(f_t)_{t\geq 0}$ be a solution to eq. (2.5.2) with initial datum $f_0 \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ and let a and b be positive numbers to be chosen later on. Note that for all test function φ,

$$\begin{aligned} &\frac{d}{dt}\int \varphi(p,v) f_t(dpdv) \\ &= \int \Big(\Delta_v \varphi(p,v) + v\cdot\nabla_p\varphi(p,v) - \Big(\mathrm{A}(v)+\mathrm{B}(p)+\mathscr{D}_m\mathrm{H}(f_t,p)\Big)\cdot\nabla_v\varphi(p,v)\Big) f_t(dpdv). \end{aligned} \tag{2.6.82}$$

Then

$$\begin{aligned} &\frac{d}{dt}\int \Big(a|p|^2 + 2p\cdot v + b|v|^2\Big) f_t(dpdv) \\ &= 2bd + 2\int \Big(v\cdot(ap+v) - (p+bv)\cdot\Big(\mathrm{A}(v)+\mathrm{B}(p)+\mathscr{D}_m\mathrm{H}(f_t,p)\Big)\Big) f_t(dpdv) \end{aligned} \tag{2.6.83}$$

where, by Young's inequality and assumption on A, B and H,

$$\begin{aligned} -2p\cdot\mathrm{A}(v) &= -2p\cdot(\mathrm{A}(v)-\mathrm{A}(0)) - 2p\cdot\mathrm{A}(0) && (2.6.84)\\ &\leq 2[\mathrm{A}]_1|p|\cdot|v| - 2p\cdot\mathrm{A}(0); \\ -2p\cdot\mathrm{B}(p) &= -2p\cdot(\lambda_{\mathrm{B}}p + \mathrm{D}(p) - \mathrm{D}(0) + \mathrm{D}(0)) && (2.6.85)\\ &\leq -(2\lambda_{\mathrm{B}} - 2[\mathrm{D}]_1)|p|^2 - 2p\cdot\mathrm{D}(0); \\ -2bv\cdot\mathrm{A}(v) &= -2bv\cdot(\mathrm{A}(v)-\mathrm{A}(0)+\mathrm{A}(0)) && (2.6.86)\\ &\leq -2b\lambda_{\mathrm{A}}|v|^2 - 2bv\cdot\mathrm{A}(0); \\ -2bv\cdot\mathrm{B}(p) &= -2bv\cdot(\lambda_{\mathrm{B}}p + \mathrm{D}(p) - \mathrm{D}(0) + \mathrm{D}(0)) && (2.6.87)\\ &\leq -2b\lambda_{\mathrm{B}}p\cdot v + 2b[\mathrm{D}]_1|p|\cdot|v| - 2bv\cdot\mathrm{D}(0); \\ -2p\cdot\mathscr{D}_m\mathrm{H}(f_t,p) &= -2p\cdot(\mathscr{D}_m\mathrm{H}(f_t,p) - \mathscr{D}_m\mathrm{H}(f_t,0) + \mathscr{D}_m\mathrm{H}(f_t,0)) && (2.6.88)\\ &\leq 2[\mathscr{D}_m\mathrm{H}]_{1,\infty}|p|^2 - 2p\cdot\mathscr{D}_m\mathrm{H}(f_t,0); \\ -2bv\cdot\mathscr{D}_m\mathrm{H}(f_t,p) &= -2bv\cdot(\mathscr{D}_m\mathrm{H}(f_t,p) - \mathscr{D}_m\mathrm{H}(f_t,0) + \mathscr{D}_m\mathrm{H}(f_t,0)) && (2.6.89)\\ &\leq 2b[\mathscr{D}_m\mathrm{H}]_{1,\infty}|p|\cdot|v| - 2bv\cdot\mathscr{D}_m\mathrm{H}(f_t,0) \\ &\leq b[\mathscr{D}_m\mathrm{H}]_{1,\infty}(|p|^2+|v|^2) - 2bv\cdot\mathscr{D}_m\mathrm{H}(f_t,0). \end{aligned}$$

Collecting all terms and using Young's inequality we obtain, with $a = \lambda_B b$ and $\eta = \gamma + [D]_1$,

$$\frac{d}{dt}\int \left(\lambda_B b|p|^2 + 2p\cdot v + b|v|^2\right) f_t(dpdv) \le 2bd + (2[A]_1 + 2[D]_1 b)\int |p|\cdot|v| f_t(dpdv) \tag{2.6.90}$$

$$\begin{aligned}
&+ (b\gamma + 2\gamma - 2\lambda_B + 2[D]_1)\int |p|^2 f_t(dpdv)\\
&+ (2 + \gamma b - 2\lambda_A b)\int |v|^2 f_t(dpdv)\\
&- 2(A(0) + D(0))\cdot\Big(\int p f_t(dpdv) + b\int v f_t(dpdv)\Big)\\
&\le 2bd + \Big(\frac{2}{\varepsilon} + \frac{b^2\varepsilon}{2[A]_1^2}\Big)|A(0) + D(0)|^2\\
&- (2\lambda_B - 2\eta - \varepsilon - \eta b)\int |p|^2 f_t(dpdv)\\
&- \Big((2\lambda_A - \eta)b - 2 - \frac{4[A]_1^2}{\varepsilon}\Big)\int |v|^2 f_t(dpdv)
\end{aligned}$$

for all positive ε. Now, as in the proof of theorem 2.3.4, with $[A]_1$ replaced by $2[A]_1$,we get the existence of a positive constant η_0, depending only on $[A]_1$, λ_A and λ_B, such that for all $0 \le \gamma + [D]_1 < \eta_0$, there exist b (and ε) such that $Q_{\lambda_B b,b} : (p,v) \mapsto \lambda_B b|p|^2 + 2p\cdot v + b|v|^2$ be a positive quadratic form on $\mathbb{R}^{2d}$ and such that

$$\begin{aligned}
\frac{d}{dt}\int Q_{\lambda_B b,b}(p,v) f_t(dpdv) &\le C_1 - C_2\int (|p|^2 + |v|^2) f_t(dpdv) \qquad (2.6.91)\\
&\le C_1 - C_3\int Q_{\lambda_B b,b}(p,v) f_t(dpdv)
\end{aligned}$$

for positive constants C_i. It follows that

$$\sup_{t\ge 0}\int Q_{\lambda_B b,b}(p,v) f_t(dpdv) < +\infty \tag{2.6.92}$$

if initially

$$\int Q_{\lambda_B b,b}(p,v) f_0(dpdv) < +\infty, \tag{2.6.93}$$

that is,

$$\sup_{t\ge 0}\int (|p|^2 + |v|^2) f_t(dpdv) < +\infty \tag{2.6.94}$$

if initially f_0 belongs to $\mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$.

We now turn to the proof of theorem 2.3.5. For each $i \in \{1,\dots,n\}$, the law f_t of $(\overline{\mathrm{P}}_t^{(i)},\overline{\mathrm{V}}_t^{(i)})$ is the solution to eq. (2.5.2) with f_0 as initial datum and the processes $((\overline{\mathrm{P}}_t^{(i)},\overline{\mathrm{V}}_t^{(i)}))_{t\geq 0}$ and $((\mathrm{P}_t^{(i)},\mathrm{V}_t^{(i)}))_{t\geq 0}$ are driven by the same Brownian motion. In particular the differences $p_t^{(i)} := \mathrm{P}_t^{(i)} - \overline{\mathrm{P}}_t^{(i)}$ and $v_t^{(i)} := \mathrm{V}_t^{(i)} - \overline{\mathrm{V}}_t^{(i)}$ evolve according to

$$\begin{cases} dp_t^{(i)} & = v_t^{(i)}\,dt; \\ dv_t^{(i)} & = -\Big(\mathrm{A}(\mathrm{V}_t^{(i)}) - \mathrm{A}(\overline{\mathrm{V}}_t^{(i)}) + \lambda_{\mathrm{B}} p_t^{(i)} + \mathrm{D}(\mathrm{P}_t^{(i)}) - \mathrm{D}(\overline{\mathrm{P}}_t^{(i)}) + \nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\mathrm{P}_t^{(i)}) - \nabla\frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t},\overline{\mathrm{P}}_t^{(i)})\Big)dt; \\ (p_0^{(i)},v_0^{(i)}) & = (0,0). \end{cases} \tag{2.6.95}$$

As in the proof of theorem 2.3.4, by the Young inequality and assumptions on A and D, for all positive ε, we have

$$\begin{aligned} \frac{d}{dt}\Big(\lambda_{\mathrm{B}} b|p_t^{(i)}|^2 + 2p_t^{(i)}\cdot v_t^{(i)} + b|v_t^{(i)}|^2\Big) \leq \Big(\frac{\varepsilon}{2} - 2\lambda_{\mathrm{B}} + 2[\mathrm{D}]_1 + [\mathrm{D}]_1 b\Big)|p_t^{(i)}|^2 \\ + \Big(2 + \frac{2[\mathrm{A}]_1^2}{\varepsilon} - 2\lambda_{\mathrm{A}} b + [\mathrm{D}]_1 b\Big)|v_t^{(i)}|^2 \\ - 2\Big(p_t^{(i)} + bv_t^{(i)}\Big)\cdot\Big(\nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\mathrm{P}_t^{(i)}) - \nabla\frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t},\overline{\mathrm{P}}_t^{(i)})\Big). \end{aligned} \tag{2.6.96}$$

On the other hand, we have

$$\begin{aligned} & -2\Big(p_t^{(i)} + bv_t^{(i)}\Big)\cdot\Big(\nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\mathrm{P}_t^{(i)}) - \nabla\frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t},\overline{\mathrm{P}}_t^{(i)})\Big) \\ & = -2\Big(p_t^{(i)} + bv_t^{(i)}\Big)\cdot\Big(\nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\mathrm{P}_t^{(i)}) - \nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\overline{\mathrm{P}}_t^{(i)})\Big) \\ & -2\Big(p_t^{(i)} + bv_t^{(i)}\Big)\cdot\Big(\nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\overline{\mathrm{P}}_t^{(i)}) - \nabla\frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t},\overline{\mathrm{P}}_t^{(i)})\Big). \end{aligned} \tag{2.6.97}$$

By the Cauchy-Schwarz inequality, the Young inequality and assumptions on H, we have

$$\begin{aligned} & -2\Big(p_t^{(i)} + bv_t^{(i)}\Big)\cdot\Big(\nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\mathrm{P}_t^{(i)}) - \nabla\frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t},\overline{\mathrm{P}}_t^{(i)})\Big) \\ & \leq \Big(2[\mathscr{D}_m \mathrm{H}]_{1,\infty} + b[\mathscr{D}_m \mathrm{H}]_{1,\infty}\Big)|p_t^{(i)}|^2 + b[\mathscr{D}_m \mathrm{H}]_{1,\infty}|v_t^{(i)}|^2 \\ & + 2\Big(|p_t^{(i)}| + b|v_t^{(i)}|\Big)\Big\|\nabla\frac{\delta \mathrm{H}}{\delta m}(\mu_{\mathbf{P}_t},\overline{\mathrm{P}}_t^{(i)}) - \nabla\frac{\delta \mathrm{H}}{\delta m}(\mathbb{P}_{\mathrm{P}_t},\overline{\mathrm{P}}_t^{(i)})\Big\| \\ & \leq \Big(2[\mathscr{D}_m \mathrm{H}]_{1,\infty} + b[\mathscr{D}_m \mathrm{H}]_{1,\infty}\Big)|p_t^{(i)}|^2 + b[\mathscr{D}_m \mathrm{H}]_{1,\infty}|v_t^{(i)}|^2 \\ & + 2\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}\Big(|p_t^{(i)}| + b|v_t^{(i)}|\Big)\mathscr{W}_1(\mu_{\mathbf{P}_t},\mathbb{P}_{\mathrm{P}_t}). \end{aligned} \tag{2.6.98}$$

By the triangle inequality, we have

$$\mathscr{W}_1(\mu_{\mathbf{P}_t},\mathbb{P}_{\mathrm{P}_t}) \le \mathscr{W}_1(\mu_{\mathbf{P}_t},\mu_{\overline{\mathbf{P}}_t}) + \mathscr{W}_1(\mu_{\overline{\mathbf{P}}_t},\mathbb{P}_{\mathrm{P}_t}). \tag{2.6.99}$$

By the Kantorovitch-Rubinstein duality relation, we have

$$\mathscr{W}_1(\mu_{\mathbf{P}_t},\mu_{\overline{\mathbf{P}}_t}) \le \frac{1}{n}\sum_{j=1}^{n} |p_t^{(j)}|. \tag{2.6.100}$$

Moreover, by symmetry, we have

$$\mathbb{E}[\mathscr{W}_1(\mu_{\mathbf{P}_t},\mu_{\overline{\mathbf{P}}_t})] \le \mathbb{E}[|p_t^{(i)}|]. \tag{2.6.101}$$

From the Cauchy-Schwarz, Jensen and Young inequalities, we deduce that

$$\begin{aligned} &2\|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}\Big(|p_t^{(i)}| + b|v_t^{(i)}|\Big)\mathscr{W}_1(\mu_{\mathbf{P}_t},\mathbb{P}_{\mathrm{P}_t}) \qquad (2.6.102)\\ &\le \|\mathscr{D}_m^2 \mathrm{H}\|_{\mathbf{op},\infty}\Big(|p_t^{(i)}|^2 + b|v_t^{(i)}|^2 + (b+1)\mathscr{W}_1(\mu_{\mathbf{P}_t},\mathbb{P}_{\mathrm{P}_t})^2\Big);\\ &\mathbb{E}\Big[2\Big(|p_t^{(i)}| + b|v_t^{(i)}|\Big)\mathscr{W}_1(\mu_{\mathbf{P}_t},\mathbb{P}_{\mathrm{P}_t})\Big] \qquad (2.6.103)\\ &\le (2b+3)\mathbb{E}[|p_t^{(i)}|^2] + b\mathbb{E}[|v_t^{(i)}|^2] + (2b+2)\mathbb{E}[\mathscr{W}_2(\mu_{\overline{\mathbf{P}}_t},\mathbb{P}_{\mathrm{P}_t})^2]. \end{aligned}$$

By the result of Fournier and Guillin ([114, Theorem 1]), we have

$$\sup_{t\ge 0}\mathbb{E}[\mathscr{W}_2(\mu_{\overline{\mathbf{P}}_t},\mathbb{P}_{\mathrm{P}_t})^2] \le \mathrm{C}(d)\Big(\sup_{t\ge 0}\int (|p|^2+|v|^2) f_t(dpdv)\Big)\delta_d(n). \tag{2.6.104}$$

Collecting all terms, it follows that there exists a positive constant c such that for all γ and $[\mathrm{D}]_1$ in $[0,c)$, there exists a constant M such that

$$\begin{aligned} \frac{d}{dt}\mathbb{E}\Big[\Big(\lambda_{\mathrm{B}} b|p_t^{(i)}|^2 + 2p_t^{(i)}\cdot v_t^{(i)} + b|v_t^{(i)}|^2\Big)\Big] \le &-\Big(2\lambda_{\mathrm{B}} - 2\eta - \varepsilon - \eta b\Big)\mathbb{E}[|p_t^{(i)}|^2] \qquad (2.6.105)\\ &-\Big((2\lambda_{\mathrm{A}} - \eta)b - 2 - \frac{4[\mathrm{A}]_1^2}{\varepsilon}\Big)\mathbb{E}[|v_t^{(i)}|^2]\\ &+\mathrm{M}\Big(\frac{2}{\varepsilon} + \frac{b^2\varepsilon}{2[\mathrm{A}]_1^2}\Big)\delta_d(n) \end{aligned}$$

for all positive t, b and ε, where $\eta = \gamma + [\mathrm{D}]_1$.

Now, as in the proof of theorem 2.3.4, with $[\mathrm{A}]_1$ replaced by $2[\mathrm{A}]_1$,we get the existence of a positive constant η_0, depending only on $[\mathrm{A}]_1$, λ_{A} and λ_{B}, such that for all $0 \le \gamma + [\mathrm{D}]_1 < \eta_0$,

there exist b (and ε) such that $Q_{\lambda_B b,b} : (p,v) \mapsto \lambda_B b|p|^2 + 2p\cdot v + b|v|^2$ be a positive quadratic form on $\mathbb{R}^{2d}$ and such that

$$\frac{d}{dt}\mathbb{E}[Q_{\lambda_B b,b}(p_t^{(i)}, v_t^{(i)})] \le -C_1\mathbb{E}[|p_t^{(i)}|^2 + |v_t^{(i)}|^2] + C_2\delta_d(n) \tag{2.6.106}$$

for all $t \ge 0$ and for positive constants C_1 and C_2, also depending on f_0 through its second moment, but not on n. In turn this is bounded by

$$-C_3\mathbb{E}[Q_{\lambda_B b,b}(p_t^{(i)}, v_t^{(i)})] + C_2\delta_d(n), \tag{2.6.107}$$

so that

$$\sup_{t\ge 0}\mathbb{E}[Q_{\lambda_B b,b}(p_t^{(i)}, v_t^{(i)})] \le C_4\delta_d(n) \tag{2.6.108}$$

and finally

$$\sup_{t\ge 0}\mathbb{E}[|p_t^{(i)}|^2 + |v_t^{(i)}|^2] \le C\delta_d(n) \tag{2.6.109}$$

where the constant C depends on the parameters of the equation and on the second moment of f_0, but not on n.

■

2.6.3 Proof related to the examples

Proof of theorem 2.5.1. To do this, simply show the two equivalents (**1**) $\Longleftrightarrow$ (**2**) and (**2**) $\Longleftrightarrow$ (**3**).

(**1**) $\Longleftrightarrow$ (**2**). Let μ_∞ be such that

$$\Delta\mu_\infty + \nabla\cdot(\mu_\infty \mathscr{D}_m H(\mu_\infty,\cdot)) = 0. \tag{2.6.110}$$

For any test function φ, we have

$$\langle \Delta\mu_\infty + \nabla\cdot(\mu_\infty\mathscr{D}_m H(\mu_\infty,\cdot)), \varphi\rangle = \int \Delta\varphi d\mu_\infty - \int \mathscr{D}_m H(\mu_\infty,\cdot)\cdot\nabla\varphi d\mu_\infty. \tag{2.6.111}$$

Since the Laplacian is a self-adjoint operator, we have

$$\int \Delta\varphi d\mu_\infty - \int \mathscr{D}_m \mathrm{H}(\mu_\infty,\cdot)\cdot\nabla\varphi d\mu_\infty = -\int \Big(\nabla\frac{d\mu_\infty}{dx}\nabla\varphi + \frac{d\mu_\infty}{dx}\mathscr{D}_m\mathrm{H}(\mu_\infty,\cdot)\cdot\nabla\varphi\Big)dx \tag{2.6.112}$$
$$= -\int \nabla\Big(\ln\frac{d\mu_\infty}{dx} + \frac{\delta \mathrm{H}}{\delta m}(\mu_\infty,\cdot)\Big)\cdot\nabla\varphi d\mu_\infty.$$

We deduce that

$$(\mathbf{1}) \Longleftrightarrow \forall\varphi,\quad \int \nabla\Big(\ln\frac{d\mu_\infty}{dx} + \frac{\delta \mathrm{H}}{\delta m}(\mu_\infty,\cdot)\Big)\cdot\nabla\varphi d\mu_\infty \Longleftrightarrow \nabla\Big(\ln\frac{d\mu_\infty}{dx} + \frac{\delta \mathrm{H}}{\delta m}(\mu_\infty,\cdot)\Big) = 0, \tag{2.6.113}$$

which proves the expected equivalence.

(**2**) $\Longleftrightarrow$ (**3**). As for this last equivalence, we have

$$\forall (y,z)\in\mathbb{R}^d\times\mathbb{R}^d,\quad \Big|\ln\frac{d\mu_\infty}{dx}(y) + \frac{\delta \mathrm{H}}{\delta m}(\mu_\infty,y) - \ln\frac{d\mu_\infty}{dx}(z) - \frac{\delta \mathrm{H}}{\delta m}(\mu_\infty,z)\Big| \tag{2.6.114}$$
$$\le \Big\|\nabla\Big(\ln\frac{d\mu_\infty}{dx} + \frac{\delta \mathrm{H}}{\delta m}(\mu_\infty,\cdot)\Big)\Big\|_\infty \|y-z\|.$$

We deduce that

$$(\mathbf{2}) \Longleftrightarrow \ln\frac{d\mu_\infty}{dx} + \frac{\delta \mathrm{H}}{\delta m}(\mu_\infty,\cdot) = \mathrm{C}_{\mu_\infty} \Longleftrightarrow (\mathbf{3}). \tag{2.6.115}$$

■

Proof of proposition 2.4.1. We have

$$\mathscr{D}_m \mathrm{H}_{\mathrm{VW}}(\mu,x) = \nabla \mathrm{V}(x) + \int \nabla \mathrm{W}(x-y)\mu(dy); \tag{2.6.116}$$
$$\mathscr{D}_m^2 \mathrm{H}_{\mathrm{VW}}(\mu,x,y) = -\nabla^2 \mathrm{W}(x-y). \tag{2.6.117}$$

■

Proof of proposition 2.4.2. We have

$$\mathscr{D}_m \mathrm{H}_{\mathrm{VW}}(\mu,x) = \nabla \mathrm{V}(x) + \int \nabla \mathrm{W}(x-y)\mu(dy); \tag{2.6.118}$$
$$\mathscr{D}_m^2 \mathrm{H}_{\mathrm{VW}}(\mu,x,y) = -\nabla^2 \mathrm{W}(x-y). \tag{2.6.119}$$

■

Proof of proposition 2.4.3. The intrinsic derivatives associated with this functional are given by

$$\begin{aligned}
&\mathscr{D}_m \mathrm{H}^{\mathrm{N}}_{\mathrm{VW}}(\mu, x) && (2.6.120)\\
&= \nabla \mathrm{V}(x)\\
&+ \sum_{k=2}^{\mathrm{N}} \sum_{j=1}^{k} \int \nabla_{x_j} \mathrm{W}^{(k)}(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_k) \mu^{\otimes k-1}(dx_1, \ldots, dx_{j-1}, dx_{j+1}, \ldots, dx_k)\\
&= \nabla \mathrm{V}(x) + \sum_{k=2}^{\mathrm{N}} k \int \nabla_{x_1} \mathrm{W}^{(k)}(x, y) \mu^{\otimes k-1}(dy)\\
\mathscr{D}^2_m \mathrm{H}^{\mathrm{N}}_{\mathrm{VW}}(\mu, x, y) &= \sum_{k=2}^{\mathrm{N}} k(k-1) \int \nabla^2_{x_1, x_2} \mathrm{W}^{(k)}(x, y, z) \mu^{\otimes k-2}(dz). && (2.6.121)
\end{aligned}$$

■

Proof of proposition 2.4.4. The intrinsic derivatives associated with this functional are given by

$$\begin{aligned}
\mathscr{D}_m \mathrm{H}^{\mathrm{N}}_{\mathrm{VW}}(\mu, x) &= \nabla \mathrm{V}(x) && (2.6.122)\\
&+ \sum_{k=2}^{\mathrm{N}} \sum_{j=1}^{k} \int \nabla_{x_j} \mathrm{W}^{(k)}(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_k) \mu^{\otimes k-1}(dx_1, \ldots, dx_{j-1}, dx_{j+1}, \ldots, dx_k)\\
&= \nabla \mathrm{V}(x) + \sum_{k=2}^{\mathrm{N}} k \int \nabla_{x_1} \mathrm{W}^{(k)}(x, y) \mu^{\otimes k-1}(dy)\\
\mathscr{D}^2_m \mathrm{H}^{\mathrm{N}}_{\mathrm{VW}}(\mu, x, y) &= \sum_{k=2}^{\mathrm{N}} k(k-1) \int \nabla^2_{x_1, x_2} \mathrm{W}^{(k)}(x, y, z) \mu^{\otimes k-2}(dz). && (2.6.123)
\end{aligned}$$

■

Proof of proposition 2.4.5. We have

$$\begin{aligned}
\mathscr{D}_m \mathrm{H}_{\psi}(\mu, x) &= \psi'(\langle \mu, \mathrm{W} \rangle) \nabla \mathrm{W}(x); && (2.6.124)\\
\mathscr{D}^2_m \mathrm{H}_{\mathrm{VW}}(\mu, x, y) &= \psi''(\langle \mu, \mathrm{W} \rangle) \nabla \mathrm{W}(x) \otimes \nabla \mathrm{W}(y). && (2.6.125)
\end{aligned}$$

■

Proof of proposition 2.4.6. We have

$$\mathscr{D}_m \mathrm{H}_\psi(\mu, x) = \psi'(\langle \mu, \mathrm{W} \rangle) \nabla \mathrm{W}(x); \tag{2.6.126}$$

$$\mathscr{D}_m^2 \mathrm{H}_{\mathrm{VW}}(\mu, x, y) = \psi''(\langle \mu, \mathrm{W} \rangle) \nabla \mathrm{W}(x) \otimes \nabla \mathrm{W}(y). \tag{2.6.127}$$

■

3

QSDs de McKean-Vlasov et schémas d'Euler

La théorie des probabilités n'est, au fond, que le bon sens réduit au calcul : elle fait apprécier avec exactitude ce que les esprits justes sentent par une sorte d'instinct, sans qu'ils puissent souvent s'en rendre compte.

PIERRE-SIMON LAPLACE

Résumé du chapitre

Dans ce chapitre qui est le fruit de l'article [2] en collaboration avec Fabien PANLOUP, il est principalement question de l'approximation des distributions quasi-stationnaires de processus de McKean-Vlasov absorbés via le shift associé à la mesure d'occupation empirique de dynamiques auto-attractives associées aux schémas d'Euler.

3.1 Introduction

The study of the long-time behavior of stochastic processes which are absorbed at a given stopping time (usually exit time) is now a widely studied problem in the Markovian setting. Under adequate conditions, Markovian dynamics conditioned to stay alive converge in distribution to the so-called quasi-stationary distribution which can be viewed as an equilibrium related to normalized absorbed dynamics. We refer for instance to the seminal paper [65] for general results which ensure convergence at an exponential rate and to [83] for a survey on recent contributions on the topic. For practical applications, conditioned dynamics attracted much attention for modelling of population dynamics (see *e.g.* [58, 73]) and for speeding up optimization dynamics (see *e.g.* [158] and the references therein).

The question of its approximation thus becomes an important problem and a series of papers investigated mainly two ways of approximations procedures: the Fleming-Viot algorithm (see *e.g.* [48, 82, 97, 231]) where the principle is to simulate a particle system induced by the dynamics and to manage the conditioning by making reborn the killed particles with an offspring being through the empirical measure of the alive particles. Another approach initiated by [24] (see also [34, 28, 31, 233]) is based on self-interacting dynamics, *i.e.*, where the distribution of the reborn particles is given by the occupation measure of the process.

In this paper, we address quasistationary distributions for dynamics with interaction with their distribution that we call *McKean-Vlasov-type* dynamics in the sequel, and especially to their approximation. Our terminology is certainly related to the so-called McKean-Vlasov SDEs whose "standard" convergence to equilibrium, *i.e.* to its "standard" invariant distribution has been widely studied in the last years (among others, see *e.g.* [134] and the references therein). For absorbed McKean-Vlasov-type dynamics, the references seem to be more rare. The topic is adressed in [223] where the authors study the Fleming-Viot dynamics fo McKean-Vlasov SDEs and exhibit some conditions which ensure that for the exit time of a compact set, the Fleming-Viot algorithm converges to the set of QSDs of such dynamics. At this stage, it is important to remark that in order to make sense, absorbed McKean-Vlasov SDEs are certainly considered by interacting with the conditioned distribution:

$$\begin{cases} d\xi_t = b(\xi_t, \mu_t)\,dt + \sigma(\xi_t, \mu_t)\,dW_t \\ \mu_t = \mathscr{L}(\xi_t | \tau > t), \end{cases} \tag{3.1.1}$$

where τ denotes the exit time and $b : \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{M}_{d,d}$ (the space of $d \times d$ real matrices) denote the drift and diffusion coefficients respectively (whose regularity will be precised later).

Initially motivated by the objective to consider the self-interacting approach related to continuous-time dynamics such as (3.1.1), our work focuses in fact to general discrete-time dynamics with interaction [1]. In this sense, it is mainly a generalization to interacting dynamics of [28]. Our prototypical example is the Euler-Maruyama scheme of (3.1.1) but our results may apply to other types of interacting models. Our main contributions are the following:

— For a general Polish space $\mathscr{E}$, for a family of kernels of submarkovian dynamics $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$ and for a stopping time τ, we begin by providing a characterization of QSDs of such interacting dynamics as the fixed point of dynamics with redistribution (whose markovian kernel is denoted by $\mathbb{K}_\mu$ in the sequel).

— We build and study an associated self-interacting sequence of random probabilities denoted by $(\mu_n)_{n\geq 1}$ (where $\mu_n \in \mathscr{P}(\mathscr{E})$ is the occupation measure of the dynamics).

  — We provide Lyapunov criterions which ensure almost sure tightness of $(\mu_n)_{n\geq 1}$.

  — Denoting by $(\tilde{\mu}_t^{(n)})_{t\geq 0}$ a function with values in $\mathscr{P}(\mathscr{E})$ built as an adequate interpolation of the sequence $(\mu_k)_{k\geq n}$, we show that every weak limit of the sequence [2] $(\tilde{\mu}^{(n)})_{n\geq 1}$ is a solution to an ordinary differential equation (ode) on the space of probabilities.

  — We exhibit conditions which ensure that this limiting ode is stationary which in turn implies that the limit points of (μ_n) are almost surely QSDs. This last result (corresponding to Theorem 3.2.8(ii) is probably the part of the paper, where the non-linearity of the dynamics generates the most difficulties. To obtain this property, we show uniform convergence properties of the limiting ode but without uniqueness assumptions (see Proposition 3.4.6). Note that this result is obtained under Assumption **(A$_4$)** which is a typical assumption for ensuring uniqueness of QSDs in the linear case.

1. As mentioned later, the continuous-time setting, especially the *hard-killing* problem, contains additional difficulties that we leave for a future paper. Yet, we tackle the continuous-time by a weak approach considering QSDs of McKean-Vlasov SDEs as limits of QSDs of related Euler schemes.

2. We also show that this sequence is tight for the topology of uniform convergence on compact intervals on the space $\mathscr{C}([0,+\infty),\mathscr{P}(\mathscr{E}))$ of continuous functions with values in the space of probabilities on $\mathscr{E}$ (endowed with the topology of weak convergence).

— We apply our general result to the Euler scheme of a McKean-Vlasov SDE by proving, under ellipticity assumptions and boundedness of the *flat derivatives* of the coefficients, that the procedure converges [3]. In the compact case, we are able to show that when the step h goes to 0, tightness holds (under adequate conditions) and limiting distributions (when $h \to 0$) are QSDs of the McKean-Vlasov SDE. In the non-compact case, we are able to provide a result for Euler schemes of Lévy-driven SDEs *coming down from infinity* [4] when the driving noise has polynomial decrease at infinity. Curiously, our class of examples may contain stable driven SDEs but not the case where the driving noise is the Brownian motion (precisely, we are not able to prove the above assumption ($\mathbf{A_4}$) in this case). In the Brownian case, we still have a non trivial tightness result under a Lyapunov assumption which is for instance satisfied by Ornstein-Uhlenbeck processes with interaction (see Lemma 3.7.1 combined with Lemma 3.2.10).

— We end by some numerical illustrations in the same case as the one considered in [223, Example 2.15] and show that our simulations allow to retrieve the (explicit) QSDs related to this specific case.

To end this description of our main results, let us remark that this contribution could be probably easily adapted to the continuous-time case for soft-killing SDEs but considering the continuous-time for the hard-killing problem (as a generalization of [31]) seems to generate much more difficulties. In some sense, this is why we limit ourselves to a "weak" result which approximates QSDs of McKean-Vlasov SDEs by letting the time go to ∞ and then the step go to 0.

Outline. The paper is organized as follows: In Section 3.2, we detail the framework and the main assumptions, and state our main results about the convergence of the algorithm towards the QSD. We provide examples and numerical illustrations in section 3.2.2.

The sequel of the paper is devoted to the proofs of the results. In section 3.3, we provide preliminary properties and show that the limiting dynamics of the (continuous-time exten-

3. By "converges", we abusively mean that the sequence is tight and every weak limit is a QSD. This certainly leads to convergence when uniqueness holds but we do not address this (probably difficult) question in this paper.

4. This terminology is classical in the literature of QSDs and roughly says that the dynamics can come back in a compact set in a bounded time. Such an assumption is usually required to get uniqueness of QSDs in the Markovian case.

sion of the) algorithm is a solution to a non-linear ode on the space of probabilities. In section 3.4, we exhibit long-time convergence properties of this non-linear ode. As mentioned before, this in some sense in this part, that the non-linearity implies the most difficulties. In section 3.5, we gather the results of the two previous sections to prove theorem 3.2.8. In section 3.6, we prove theorem 3.2.11 which provides sufficient Lyapunov-type conditions to ensure tightness of the sequence of empirical measures. Finally, in section 3.7, we focus on proofs and additional results related to Euler schemes of McKean-Vlasov SDEs. We provide explicit tightness criterions and prove proposition 3.2.13 and proposition 3.2.15.

3.2 Setting and Main Results

3.2.1 Notation and Setting

Notation. Let $\mathscr{E}$ be a Polish space equipped with its Borel σ-field $\mathscr{B}(\mathscr{E})$. Throughout, we let $\mathscr{P}(\mathscr{E})$ denote the space of (Borel) probabilities over $\mathscr{E}$ equipped with the topology of weak convergence. For all $\mu \in \mathscr{P}(\mathscr{E})$ and $f \in L^1(\mu)$ we write $\mu(f)$ for $\int_{\mathscr{E}} f d\mu$. Recall that $\mu_n \to \mu$ in $\mathscr{P}(\mathscr{E})$ provided $\mu_n(f) \to \mu(f)$ for all $f \in C_b(\mathscr{E})$ (denoting the bounded continuous functions on $\mathscr{E}$). The space of signed measures on $\mathscr{E}$ is denoted $\mathscr{M}_s(\mathscr{E})$ and the total variation norm is denoted by $\|.\|_{TV}$.

A *sub-Markovian kernel* on $\mathscr{E}$ is a map $Q : \mathscr{E} \times \mathscr{B}(\mathscr{E}) \mapsto [0,1]$ such that for all $x \in \mathscr{E}$, $A \to Q(x,A)$ is a measure and for all $A \in \mathscr{B}(\mathscr{E})$, $x \to Q(x,A)$ is measurable. If furthermore $Q(x,\mathscr{E}) = 1$ for all $x \in M$, then Q is called a *Markov (or Markovian) kernel*. For every bounded measurable function $f : \mathscr{E} \mapsto \mathbb{R}$, and $\mu \in \mathscr{P}(\mathscr{E})$, we let Qf and μQ respectively denote the map and measure defined by

$$Qf(x) = \int_{\mathscr{E}} f(y)Q(x,dy) \quad \text{and} \quad \mu Q(\cdot) = \int_{\mathscr{E}} \mu(dx)Q(x,\cdot).$$

A (sub)-Markovian kernel is Feller if for all $f \in C_b(\mathscr{E})$, $Qf \in C_b(\mathscr{E})$. For an integer $d \geq 1$, the scalar product and the Euclidean norm on $\mathbb{R}^d$ are denoted by $\langle .,. \rangle$ and $|.|$ whereas $\mathbb{M}_{d,d}$ is the space of $d \times d$ matrices. For $p > 0$, we set $\mathscr{P}_p(\mathbb{R}^d) = \{\mu \in \mathscr{P}(\mathbb{R}^d), \mu(|.|^p) < +\infty\}$. A map $U : \mathscr{P}_p(\mathbb{R}^d) \to \mathbb{R}$ is said to be of class $\mathscr{C}^1$ if there exists a jointly continuous and locally bounded map $\frac{\delta U}{\delta m} : \mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ satisfying: $\sup_{m \in \mathscr{P}_p(\mathbb{R}^d)} |\frac{\delta U}{\delta m}(m,x)| \lesssim (1+|x|^p)$ and for all m_0 and $m_1 \in \mathscr{P}(\mathbb{R}^d)$,

$$U(m_1) - U(m_0) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(tm_1 + (1-t)m_0, y)(m_1 - m_0)(dy)dt.$$

In this case, we set

$$\left\|\frac{\delta U}{\delta m}(m,.)\right\|_\infty = \sup_{x\in\mathbb{R}^d}\left|\frac{\delta U}{\delta m}(m,x)\right| \quad \text{and} \quad \left[\frac{\delta U}{\delta m}(m,.)\right]_1 = \sup_{x,y\in\mathbb{R}^d, x\neq y}\frac{\left|\frac{\delta U}{\delta m}(m,y)-\frac{\delta U}{\delta m}(m,x)\right|}{|y-x|}.$$

Setting. Let $(K_{\mu,\partial})_\mu$ denote a family of subMarkovian kernels on $\mathscr{E}$ and let $(Y_n)_{n\geq 0}$ denote the sequence defined on $\mathscr{E}\cup\{\partial\}$ by: $Y_0 \sim \nu_0$

$$Y_{n+1} = \begin{cases} Z_{n+1} & \text{if } \tau > n+1 \\ \partial & \text{if } \tau \leq n+1, \end{cases} \tag{3.2.1}$$

where

— τ is a $(\mathscr{F}_n)_{n\geq 0}$-stopping time defined by

$$\tau := \inf\{n\geq 1, \Theta_n = 1\},$$

where $(\Theta_n)_{n\geq 1}$ is a sequence of Bernoulli random variables such that

$$\mathbb{P}(\Theta_{n+1}=1|\mathscr{F}_n, Y_n\in\mathscr{E}) = \delta_{\nu_n}(Y_n) \quad \text{with} \quad \delta_\mu(x) = 1 - K_{\mu,\partial}\mathbf{1}(x),$$

$(\mathscr{F}_n)_{n\geq 0}$ is defined by $\mathscr{F}_0 = \sigma(Y_0)$ and for $n\geq 1$,

$$\mathscr{F}_n = \sigma(Y_0,\dots,Y_n,\Theta_1,\dots,\Theta_n),$$

and

$$\nu_n = \mathscr{L}(Y_n|\tau > n).$$

— $(Z_n)_{n\geq 1}$ is a sequence of random variables with values in $\mathscr{E}$ such that

$$\mathscr{L}(Z_{n+1}|\mathscr{F}_n, Y_n\in\mathscr{E}) = K_{\nu_n,\partial}(Y_n,.).$$

Remark 3.2.1. The above definition implies in an underlying way that $\mathbb{P}(\tau > n) > 0$. Nevertheless, the definition is still consistent if $\mathbb{P}(\tau > n) = 0$ since in this case $Y_n = \partial$ almost surely.

Model 1: (Euler Scheme of a Mc-Kean Vlasov diffusion killed at an exit time of a relatively compact set). Assume that D is a bounded open set of $\mathbb{R}^d$. Let $h > 0$. Let $(\zeta_t)_{t\geq 0}$ denote a

Brownian motion on $\mathbb{R}^d$ or more generally a given centered Lévy process. We also set

$$\Delta\zeta_{s,t} = \zeta_t - \zeta_s, \quad 0 \le s \le t.$$

Let ∂ denote a cemetery point. Set $\mathscr{E} = \bar{\mathrm{D}}$. We define $(\mathrm{Y}_n)_{n\ge 0}$ by $\mathrm{Y}_0 \sim \mu_0 \in \mathscr{P}(\mathscr{E})$ and

$$\forall n \ge 0, \quad \mathrm{Y}_{n+1} = \begin{cases} \mathrm{Y}_n + hb(\mathrm{Y}_n, \nu_n) + \sigma(\mathrm{Y}_n, \nu_n)\Delta\zeta_{nh,(n+1)h} & \text{while } \tau > n+1, \\ \partial & \text{if } \tau \le n+1, \end{cases}$$

where $b : \mathscr{E} \times \mathscr{P}(\mathscr{E})$ and $\sigma : \mathscr{E} \times \mathscr{P}(\mathscr{E}) \to \mathbb{M}_{d,d}$ denote (at least jointly continuous) functions,

$$\tau := \inf\{n \in \mathbb{N}, \mathrm{Y}_n + hb(\mathrm{Y}_n, \nu_n) + \sigma(\mathrm{Y}_n, \nu_n)\Delta\zeta_{nh,(n+1)h} \in \mathscr{E}^c\},$$

and

$$\nu_n = \mathscr{L}(\mathrm{Y}_n | \tau > n).$$

Remark 3.2.2. $\triangleright$ If μ_0 has support in D and σ is not degenerated (typically uniformly elliptic), τ is almost surely equal to the time where the scheme leaves D (this will allow us to work on the compact space $\mathscr{E} = \bar{\mathrm{D}}$ in the sequel for the study of the exit time of the open set D).

$\triangleright$ Note that in this example, $\mathrm{K}_{\mu,\partial}$ (denoted $\mathrm{K}^{(h)}_{\mu,\partial}$) is defined by: for every measurable function $f : \mathscr{E} \to \mathbb{R}$,

$$\mathrm{K}^{(h)}_{\mu,\partial} f(x) = \mathbb{E}[f(\Upsilon^h_x)\mathbf{1}_{\Upsilon^h_x \in \mathrm{D}}] \quad \text{with} \quad \Upsilon^h_x = x + hb(x,\mu) + \sigma(x,\mu)\Delta\zeta_{0,h}. \tag{3.2.2}$$

$\triangleright$ This Euler scheme is related to the McKean-Vlasov diffusion:

$$\begin{cases} d\mathrm{Y}_t = b(\mathrm{Y}_t, \nu_t)\,dt + \sigma(\mathrm{Y}_t, \nu_t)\,d\mathrm{W}_t \\ \nu_t = \mathscr{L}(\mathrm{Y}_t | \tau > t) \end{cases}$$

where $\tau := \inf\{t \ge 0, \mathrm{Y}_t \in \mathrm{D}^c\}$.

Model 2: (Euler Scheme of a McKean-Vlasov diffusion killed at an exit time of an unbounded set) Excepted the undboundness, this is exactly the same model. Nevertheless, we distinguish them in order to emphasize the specific difficulties involved by the non-compactness. Furthermore, this numbering will correspond to our two types of examples in the sequel. The exit time of an unbounded set appears naturally in applications. Think for instance to

population dynamics where the exit time corresponds to the extinction of the population: in this case, $\bar{D} = [0, +\infty)$.

Definition 3.2.3. A probability $\nu \in \mathscr{P}(\mathscr{E})$ is called a Quasi-Stationary Distribution (QSD) for $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$ if $Y_0 \sim \nu$ implies that

$$\forall n \in \mathbb{N}, \quad \mathbb{P}(Y_n \in ., \tau > n) = \nu(.)\mathbb{P}(\tau > n). \tag{3.2.3}$$

By an induction, one easily checks that in this case, τ has an exponential distribution with parameter $\lambda = \nu K_{\nu,\partial}(\mathscr{E})$. In particular, $\mathbb{P}_\nu(\tau > n)$ is positive for every n and the definition then implies that

$$\forall n \in \mathbb{N}, \quad \mathscr{L}(Y_n|\tau > n) = \nu.$$

Remark 3.2.4. We thus retrieve the classical properties of the Markov setting. Nevertheless, it is important to keep in mind that this property also involves that under the "quasi-stationary regime", the dynamics of the killed process is markovian (since at each time n, the interaction distribution is also equal to ν). This is similar to the (classical) stationary regime of dynamics with interaction where one retrieves the Markov property when the process starts from its invariant distribution.

Before going further, let us introduce two first assumptions:

($\mathbf{A_0}$): There exists $\ell \geq 1$ such $\sup_{x\in\mathscr{E},\mu\in\mathscr{P}(\mathscr{E})} K^\ell_{\mu,\partial}\mathbf{1}(x) < 1$.

where, here and in the sequel, we write $\mathbf{1}$ instead of $\mathbf{1}_\mathscr{E}$ (so that in particular, $K^\ell_{\mu,\partial}\mathbf{1}(x) = K^\ell_{\mu,\partial}(x,\mathscr{E})$). Before going further, let us remark that (see section 3.8),

Lemma 3.2.5. If $\mathscr{E}$ is compact and if for every $k \geq 1$, $(x,\mu) \mapsto K^k_{\mu,\partial}\mathbf{1}(x)$ is continuous on $\mathscr{E} \times \mathscr{P}(\mathscr{E})$, then ($\mathbf{A_0}$) holds true if for every $(x,\mu) \in \mathscr{E} \times \mathscr{P}(\mathscr{E})$, there exists $\ell \geq 1$ such that $K^\ell_{\mu,\partial}\mathbf{1}(x) < 1$.

($\mathbf{A_1}$): For every $\mu \in \mathscr{P}(\mathscr{E})$, $K_{\mu,\partial}$ is Feller and for all $x \in \mathscr{E}$, $\mu \mapsto \mu K_{\mu,\partial}$ is continuous for the topology of weak convergence on $\mathscr{P}(\mathscr{E})$.

These assumptions will be useful for a first characterization of the QSD based on the *dynamics with renewal*: let (X^μ_n) denote the Markov chain with kernel

$$\mathbb{K}_\mu(x,.) = K_{\mu,\partial}(x,.) + \delta_\mu(x)\mu(.). \tag{3.2.4}$$

This means that when the particle is killed, it is regenerated through the distribution μ. Set

$$\mathscr{A}_\mu(x,.) = \sum_{n\geq 0} \mathrm{K}^n_{\mu,\partial}(x,.), \tag{3.2.5}$$

Under $(\mathbf{A_0})$ and $(\mathbf{A_1})$, it can be shown (see Lemma 3.3.1) that $\mathrm{A}_\mu(x,.)$ is well-defined for every $(x,\mu) \in \mathscr{E} \times \mathscr{P}(\mathscr{E})$ and that

$$\Pi_\mu = \frac{\mu\mathscr{A}_\mu}{\mu\mathscr{A}_\mu \mathbf{1}}$$

is the unique invariant distribution for $(\mathrm{X}^\mu_n)_{n\geq 1}$. We introduce a Lyapunov-type assumption which will be used for existence of QSD in the non-compact case (and tightness in the sequel).

$(\mathbf{A_V})$: For every $\varepsilon > 0$, there exists $\beta_\varepsilon \in \mathbb{R}$ and an inf-compact [5] function $\mathrm{V}_\varepsilon : \mathscr{E} \to \mathbb{R}_+$ such that

$$\forall\, \mu \in \mathscr{P}(\mathscr{E}), \quad \mathrm{K}_{\mu,\partial}\mathrm{V}_\varepsilon \leq \varepsilon \mathrm{V}_\varepsilon + \beta_\varepsilon.$$

Remark 3.2.6. We refer to Lemma 3.7.1 for concrete conditions on dynamics to ensure $(\mathbf{A_V})$.

We have the following property:

Lemma 3.2.7. Assume $(\mathbf{A_0})$ and $(\mathbf{A_1})$.

(i) $\mu^\star$ is a QSD for $(\mathrm{K}_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$ if and only if

$$\mu^\star = \Pi_{\mu^\star}.$$

(ii) Furthermore, if $\mathscr{E}$ is compact or, if $(\mathbf{A_V})$ holds and $\inf_{\mu\in\mathscr{P}(\mathscr{E})} \mu\mathrm{K}_{\mu,\partial}\mathbf{1} > 0$, then existence holds for the QSD related to $(\mathrm{K}_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$.

The fact that $\mu^\star$ can be viewed as a fixed point of $\mu \mapsto \Pi_\mu$ will correspond to the natural target of our algorithm. We mean that our algorithm is naturally connected to the map $\mu \mapsto \Pi_\mu$ and this is why, we chose to recall this characterization in the above lemma. Note that $\inf_{\mu\in\mathscr{P}(\mathscr{E})} \mu\mathrm{K}_{\mu,\partial}\mathbf{1} > 0$ is equivalent to $\sup_\mu \mu\delta_\mu < 1$, see Remark 3.2.12 for comments on this assumption (also remark that if there exists $x \in \mathscr{E}$ and μ such that $\mathrm{K}_{\mu,\partial}\mathbf{1}(x) = 0$, then existence also holds, δ_x being a trivial QSD in this case).

5. A function $\mathrm{V} : \mathscr{E} \to \mathbb{R}$ is inf-compact if for all $r > 0$, $\mathrm{K}_r = \{x \in \mathscr{E}, \mathrm{V}(x) \leq r\}$ is compact.

Finally, let us mention that we will not consider uniqueness problem in this paper, which seems to be a hard question in this non-linear setting. Note that in Section 3.2.3, we consider a very simple example where uniqueness is (already) not true.

Self-Attractive dynamics. To approximate the QSDs of $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$, one introduces the self-attractive process $(X_n,\mu_n)_{n\geq 0}$ recursively defined on $\mathscr{E}\times\mathscr{P}(\mathscr{E})$ by:

$$X_0 = x\in\mathscr{E},\quad \mu_0=\delta_x$$

and for every $n\geq 0$,

$$\mathbb{P}(X_{n+1}\in .|\mathscr{F}_n)=\mathbb{K}_{\mu_n}(X_n,.) \tag{3.2.6}$$

where $\mathscr{F}_n=\sigma(X_0,\dots,X_n)$ and

$$\mu_n=\frac{1}{H_n}\sum_{k=1}^{n}\eta_k\delta_{X_{k-1}}\quad\text{with}\quad H_n=\sum_{k=1}^{n}\eta_k.$$

is a related weighted occupation measure: the η_k are supposed to be positive numbers. Throughout the paper, we also assume that $(\gamma_n)_{n\geq 1}$ defined by $\gamma_n=:=\frac{\eta_n}{H_n}$ is *nonincreasing* and that

$$H_n\xrightarrow{n\to+\infty}+\infty\quad\text{and}\quad \gamma_n\overset{n\to+\infty}{=}o\left(\frac{1}{\log n}\right). \tag{3.2.7}$$

Note that the most natural choice is $\eta_n=1$ which leads to $\gamma_n=\frac{1}{n}$ (see Remark 3.2.12 for additional remarks on, the weights). The sequence $(\mu_n)_{n\geq 1}$ satisfies the following recursive formula:

$$\mu_{n+1}=\mu_n(1-\gamma_{n+1})+\gamma_{n+1}\delta_{X_{n+1}}. \tag{3.2.8}$$

The aim is to prove that the limit points of $(\mu_n)_{n\geq 0}$ are QSDs of $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$. Following [28], we artificially write as follows:

$$\mu_{n+1}=\mu_n+\gamma_{n+1}F(\mu_n)+\gamma_{n+1}\varepsilon_{n+1},$$

with

$$F(\mu)=-\mu+\Pi_\mu\quad\text{and}\quad \varepsilon_{n+1}=\delta_{X_{n+1}}-\Pi_{\mu_n}.$$

With this form, a QSD is a zero of F and the bet is to show that (μ_n) behaves asymptotically as a discretization of the ODE (this means that ε_n is asymptotically a noise) and that the ODE $\dot{\mu}=F(\mu)$ has sufficiently nice properties which ensure convergence properties towards its

zeros. We will see that the non-linearity will induce many difficulties for this point.

To state our main result, we will need to introduce an interpolated continuous version of (μ_n). Let $(t_k)_{k\geq 0}$ be the increasing sequence of positive numbers defined by: $t_0 = 0$ and for all $k \geq 1$, $t_k = \sum_{m=1}^{k} \gamma_m$. Extend the above sequence $(\mu_n)_{n\geq 1}$ to $\mathbb{N}\cup\{0\}$ by assuming that μ_0 is a given probability on $\mathscr{E}$. Then, we define $(\tilde{\mu}_t)_{t\geq 0}$ by: $\tilde{\mu}_0 = \mu_0$ and for all $k \geq 0$,

$$\forall s \in [0, \gamma_{k+1}], \qquad \tilde{\mu}_{t_k+s} = \mu_k + \frac{s}{\gamma_{k+1}}(\mu_{k+1} - \mu_k), \tag{3.2.9}$$

$(\tilde{\mu}_t)_{t\geq 0}$ is thus almost surely an element of the space $\mathscr{C}(\mathbb{R}_+, \mathscr{P}(\mathscr{E}))$ of continuous processes on $\mathbb{R}_+$ with values in $\mathscr{P}(\mathscr{E})$ (for the weak topology) and we consider the related shifted sequence $(\tilde{\mu}^{(n)})_{n\geq 1}$ defined by:

$$\forall t \geq 0, \quad \tilde{\mu}^{(n)}_t = \tilde{\mu}_{t_n+t}. \tag{3.2.10}$$

We now introduce additional assumptions:

(**A$_2$**): $\mu \mapsto K_{\mu,\partial}(x,.)$ is Lipschitz for the TV-norm, uniformly in $x \in \mathscr{E}$: there exists $C > 0$ such that

$$\sup_{x\in\mathscr{E}} \|K_{\mu,\partial}(x,.) - K_{\nu,\partial}(x,.)\|_{TV} \leq C\|\mu - \nu\|_{TV}.$$

(**A$_3$**): There exist $\ell \in \mathbb{N}$, $\ell \geq 1$, $\varepsilon > 0$ such that there exists $\Psi \in \mathscr{P}(\mathscr{E})$ such that for every $\mu \in \mathscr{P}(\mathscr{E})$,

$$K^{\ell}_{\mu,\partial}(., dy) \geq \varepsilon\Psi(dy).$$

The last assumption is related to the same probability Ψ than in (**A$_3$**).

(**A$_4$**): There exists $c > 0$ such that for every $k \in \mathbb{N}$, for every $\mu_1, \ldots, \mu_k \in \mathscr{P}(\mathscr{E})$, for every $n_1, \ldots, n_k \in \mathbb{N}$,

$$\frac{\Psi(K_{\mu_1,\partial} \circ \ldots \circ K_{\mu_k,\partial}\mathbf{1})}{K_{\mu_1,\partial} \circ \ldots \circ K_{\mu_k,\partial}\mathbf{1}} \geq c.$$

This assumption is discussed in Remark 3.2.9 below and in Lemma 3.2.10.

Theorem 3.2.8. Assume (**A$_0$**), (**A$_1$**), (**A$_2$**) and (**A$_3$**). Assume that $(\mu_n)_{n\geq 1}$ is *a.s.* tight on $\mathscr{P}(\mathscr{E})$ and that (γ_n) satisfies (3.2.7). Then,

(i) The sequence $(\tilde{\mu}^{(n)})_{n\geq 1}$ is almost surely tight on $\mathscr{C}(\mathbb{R}_+, \mathscr{P}(\mathscr{E}))$ and every weak limit of

$(\tilde{\mu}^{(n)})_{n\geq 1}$ is a solution of the ODE

$$\frac{d\nu_t}{dt} = -\nu_t + \Pi_{\nu_t}. \tag{3.2.11}$$

(ii) If furthermore (**A4**) holds, then every weak limit of $(\tilde{\mu}^{(n)})_{n\geq 1}$ is stationary which implies that every weak limit of $(\mu_n)_{n\geq 1}$ is a QSD $\mu^\star$ related to $(K_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$.

Remark 3.2.9. ▷ In the non-compact case, the above result requires *a.s.*- tightness of $(\mu_n)_{n\geq 1}$. A criterion is given in Theorem 3.2.11 below.

▷ As quoted in Remark 3.3.4, a Poissonization argument could avoid to assume (**A3**) in the first part of Theorem 3.2.8(i) and to replace (**A3**) by $\mathscr{A}_\mu(.,dy) \geq \varepsilon\Psi(dy)$ in Theorem 3.2.8(ii). Nevertheless, as (**A4**) requires in practice a much more constraining assumption, we chose to skip this refinement (which would had added technicalities).

▷ For (linear) Markov dynamics, *i.e.* when the kernel does not depend on μ, (**A4**) is an assumption which is used to ensure uniqueness of the QSD. Unfortunately, this is not the case in this non linear setting (in the same spirit as uniform ellipticity may ensure uniqueness of invariant distributions for linear diffusions but not -in general- for McKean-Vlasov SDEs). Nevertheless, it provides uniformity in the convergence to equilibrium of the ODE (3.2.11), which in turn helps us to derive stationarity of the weak limits of $(\tilde{\mu}^{(n)})_{n\geq 1}$ and lead to the identification of the limit.
The main example where (**A4**) is satisfied is the "lower/upper-bounded" case which is given in Lemma 3.2.10 below.

Lemma 3.2.10. If there exists a probability Ψ on $\mathscr{E}$ and some positive c_1 and c_2 such that, for every non-negative measurable function f

$$\forall (x,\mu)\in\mathscr{E}\times\mathscr{P}(\mathscr{E}),\quad c_1\Psi(f) \leq K_{\mu,\partial}f(x) \leq c_2\Psi(f), \tag{3.2.12}$$

then (**A4**) holds true.

Proof. Under this assumption,

$$c_1\Psi(K_{\mu_2,\partial}\ldots\circ K_{\mu_k,\partial}\mathbf{1}) \leq K_{\mu_1,\partial}\circ\ldots\circ K_{\mu_k,\partial}\mathbf{1} \leq c_2\Psi(K_{\mu_2,\partial}\circ\ldots\circ K_{\mu_k,\partial}\mathbf{1}).$$

Hence, integrating with respect to Ψ,

$$\Psi(K_{\mu_1,\partial} \circ \ldots \circ K_{\mu_k,\partial}\mathbf{1}) \geq c_1 \Psi(K_{\mu_2,\partial} \ldots \circ K_{\mu_k,\partial}\mathbf{1}) \geq \frac{c_1}{c_2} K_{\mu_1,\partial} \circ \ldots \circ K_{\mu_k,\partial}\mathbf{1}.$$

■

Theorem 3.2.11. Assume **($\mathbf{A_V}$)** and $\sup_{\mu\in\mathscr{P}(\mathscr{E})} \|\delta_\mu\|_\infty < 1$. Assume that $\sum \gamma_n^2 < +\infty$. Then, there exists an inf-compact function $\mathcal{V} : \mathscr{E} \to \mathbb{R}_+$ such that

$$\sup_{n\geq 1} \mu_n(\mathcal{V}) < +\infty \quad a.s.$$

In particular, (μ_n) is *a.s.* tight.

Remark 3.2.12. The assumption on $\sum \gamma_n^2$ is satisfied in many cases. For instance, if $\eta_n = n^\alpha$ with $\alpha \in (-1, \infty)$ then

$$H_n \sim \frac{n^{1+\alpha}}{1+\alpha} \quad \text{and} \quad \gamma_n \sim \frac{1}{(1+\alpha)n}.$$

One can check that it is still true if $\eta_n = e^{n^\alpha}$ with $\alpha < 1/2$. The condition on δ_μ says that, uniformly in μ and $x \in \mathscr{E}$, the probability to die in one step is upper-bounded. In the compact case, this is not constraining for our discrete-time dynamics (it may be for continuous-time *hard-killing* problems). In the non-compact, it generates constraints when the drift is "superlinear" (see 3.2.2). For instance, if $D = (0, +\infty)$ and $b(x) = -x^2$, the probability to die related to the Euler scheme goes to 1 when $x \to +\infty$. This is due to a classical unstability of the Euler scheme when the drift is not sublinear, that we overcome here by introducing a truncation-type function $\mathfrak{t}$ (see (3.2.2)).

3.2.2 Applications

We apply the general theorem to the two models introduced in the previous section.

Euler schemes of McKean-Vlasov on bounded subsets and Approximation of QSDs of continuous-time dynamics

We recall that the first model is an Euler scheme of a McKean Vlasov diffusion in a bounded open set D of $\mathbb{R}^d$. We provide a set of assumptions which guarantee the convergence [6] towards QSDs of these discrete dynamics and also show that when the stepsize h goes to 0, we

6. By convergence, we mean abusively that tightness combined with the fact that weak limits are targets.

also approximate QSDs of the McKean-Vlasov dynamics with *hard killing*. Let us introduce our assumptions:

($\mathbf{H_{MV}}$): $x \mapsto b(x,\mu)$ and $x \mapsto \sigma(x,\mu)$ are Lipschitz continuous on $\mathscr{E}$, uniformly in μ. For all $x \in D$, $b(x,\cdot)$ and $\sigma(x,\cdot)$ admit flat derivatives which are bounded:

$$\sup_{(x,\mu)\in\mathscr{E}\times\mathscr{P}(\mathscr{E})} \left\| \frac{\delta b(x,\cdot)}{\delta m}(\mu,\cdot) \right\|_\infty \vee \left\| \frac{\delta \sigma(x,\cdot)}{\delta m}(\mu,\cdot) \right\|_\infty < +\infty.$$

Finally, $x \mapsto \sigma(x,\mu)$ is uniformly elliptic, uniformly in μ, *i.e.* $\exists \lambda_0 > 0$ such that for all $(x,\mu) \in \mathscr{E}\times\mathscr{P}(\mathscr{E})$, $\sigma\sigma^\star(x,\mu) \geq \lambda_0 \mathrm{Id}$ (in the sense of symmetric matrices).

Under ($\mathbf{H_{MV}}$) and especially under the ellipticity assumption, we can consider the problem on $\bar{D}$ instead of D, *i.e.* we are interested in the QSDs of $(K^{(h)}_{\mu,\partial})_{\mu\in\mathscr{P}(\bar{D})}$, where

$$K^{(h)}_{\mu,\partial} f(x) = \mathbb{E}[f(\Upsilon^h_x)\mathbf{1}_{\Upsilon^h_x\in\bar{D}}], \tag{3.2.13}$$

with Υ^h_x defined in Remark 3.2.2. In order to emphasize the dependence in the step h, we also write (μ^h_n) instead of (μ_n) in the next proposition.

Proposition 3.2.13. Assume ($\mathbf{H_{MV}}$) and D has $\mathscr{C}^2$-boundary. Assume that (γ_n) satisfies (3.2.7). Let $h > 0$ and assume that ζ_h has a positive $\mathscr{C}^1$ density g_h.

(i) Let $h > 0$. Then, (μ^h_n) is tight on $\mathscr{P}(\bar{D})$ and every weak limit of (μ^h_n) is a QSD $\mu^\star_h$ for $(K^{(h)}_{\mu,\partial})_{\mu\in\mathscr{P}(\bar{D})}$. Furthermore, $\mu^\star_h(\partial D) = 0$.

(ii) Furthermore, if $(\zeta_t)_{t\geq 0}$ is a Brownian motion and if

$$\sup_{(x,\mu)\in\mathscr{E}\times\mathscr{P}(\mathscr{E})} \left[\frac{\delta b(x,\cdot)}{\delta m}(\mu,\cdot) \right]_1 \vee \left[\frac{\delta \sigma(x,\cdot)}{\delta m}(\mu,\cdot) \right]_1 < +\infty,$$

such a sequence $(\mu^\star_h)_{h>0}$ is tight on D and every weak limit (when $h \to 0$) is a QSD for the McKean-Vlasov diffusion $(Y_t)_{t\geq 0}$, defined in Remark 3.2.2 and killed at time $\tau := \inf\{t \geq 0, Y_t \in D^c\}$.

This result shows in particular that the algorithm may generate a way to simulate QSDs of continuous-time McKean Vlasov SDEs. In Section 3.2.3, we provide several numerical illustrations.

Dynamics killed with exit times of unbounded subsets

In this section, we aim at considering unbounded examples. A prototypical example is a discretization scheme of the following one-dimensional McKean-Vlasov equation killed when leaving $D =]0, +\infty[$:

$$dY_t = (-\varphi(Y_t) + \int F(Y_t - y)\mu_t(dy))dt + d\zeta_t, \tag{3.2.14}$$

where F is supposed to be a bounded function and

$$\lim_{x\to+\infty} \frac{\varphi(x)}{x} = +\infty.$$

We thus consider drift with superlinear growth: this will help to come from infinity but unfortunately, this will also generate unstability for discretized dynamics: when one considers the genuine related Euler scheme given by the kernel $K^{(h)}_{\mu,\partial}$, the "deterministic part" will satisfy for each $h > 0$:

$$\lim_{x\to+\infty} x - h\varphi(x) + \int F(x-y)\mu(dy) = -\infty.$$

Among the consequences, this will imply for instance that

$$\delta_\mu(x) \xrightarrow{x\to+\infty} 1,$$

which in turn involves that the second assumption of the (tightness) Theorem 3.2.11 will not be satisfied. There are several ways to manage this discretization effect (for instance, one may consider adapted step sequences, see *e.g.* [164]) but in this section where our aim is only to provide examples, we choose to introduce the following simple modification by considering:

$$\tilde{K}^{(h)}_{\mu,\partial} f(x) = \mathbb{E}[f(\tilde{Y}^h_x)\mathbf{1}_{\tilde{Y}^h_x \in D}] \quad \text{with} \quad \tilde{Y}^h_x = \mathfrak{t}(x + hb(x,\mu)) + \sigma(x,\mu)\Delta\zeta_h. \tag{3.2.15}$$

where,

($\mathbf{H}_\mathfrak{t}$): The function $\mathfrak{t} : \mathbb{R}^d \mapsto \mathbb{R}^d$ is $\mathscr{C}_1$ and

$$\sup_{x,\mu\in\mathscr{E}\times\mathscr{P}(\mathscr{E})} \{|\mathfrak{t}(x + hb(x,\mu))| + \|D\mathfrak{t}(x + hb(x,\mu))\|\} < +\infty.$$

Remark 3.2.14. In order that $\tilde{K}^{(h)}_{\mu,\partial}$ remains an approximation of the continuous dynamics, it

is fundamental to choose a function $\mathfrak{t}$ which does very slightly modify the dynamics of the genuine Euler scheme. For instance, in the benchmark example (3.2.14), a possible choice for $\mathfrak{t}$ is to define it by:

$$\mathfrak{t}(x) = \begin{cases} x & \text{if } x \geq -\mathrm{R} \\ -\mathrm{R}-1 & \text{if } x \geq -\mathrm{R}-1, \end{cases}$$

where R is a positive number, and to take a smooth increasing interpolation on $(-\mathrm{R}-1, -\mathrm{R})$. With such a function, one can check that $(\mathbf{H}_{\mathfrak{t}})$ holds true for any $\mathrm{R} > 0$. Such a modification becomes less and less restrictive when $h \to 0$ so that a convergence to the continuous-time dynamics may be true (we do not prove such a property in this part).

The above strategy can be also considered in the multidimensional case but the choice of the function $\mathfrak{t}$ may depend on the model and on the domain. We do not provide additional details.

We here consider the second model (Model 2) where the dynamics live in an unbounded subset $\mathscr{E}$ of $\mathbb{R}^d$. We have the following proposition:

Proposition 3.2.15. Let $h > 0$. Assume $(\mathbf{H_{MV}})$ and $(\mathbf{H}_{\mathfrak{t}})$. Assume that there exists $\alpha > 0$ such that $\mathbb{E}[|\zeta_h|^\alpha] < +\infty$ and that ζ_h has a $\mathscr{C}^1$ density g_h w.r.t. the Lebesgue measure with $\|g_h\|_\infty + \|\nabla g_h\|_\infty < +\infty$. Also assume that the following condition holds: for every $\mathrm{M} > 0$ and $(s_1, s_2) \in (0, +\infty)^2$ with $s_1 \leq s_2$, there exist $c_1, c_2 > 0$, and an integrable function $\mathfrak{p}_h : \mathrm{D} \mapsto (0, +\infty)$ such that for every $m \in \mathrm{B}(0, \mathrm{M})$ and $\sigma \in \mathbb{M}_{d,d}$ with $s_1 \leq |\det(\sigma)| \leq s_2$,

$$\forall z \in \mathrm{D}, \quad c_1 \mathfrak{p}_h(z) \leq g_h\left(\sigma^{-1}(z-m)\right) \leq c_2 \mathfrak{p}_h(z). \tag{3.2.16}$$

Then, if (γ_n) satisfies (3.2.7) and $\sum \gamma_n^2 < +\infty$, (μ_n) is almost surely tight in $\mathscr{P}(\bar{\mathrm{D}})$ and every weak limit $\mu_h^\star$ is a QSD for $(\tilde{\mathrm{K}}_{\mu,\partial}^{(h)})_{\mu \in \mathscr{P}(\bar{\mathrm{D}})}$ which satisfies $\mu_h^\star(\partial \mathrm{D}) = 0$.

Remark 3.2.16. ▷ Let us comment Condition (3.2.16). This assumption is built in order to ensure the conditions of Lemma 3.2.10 which in turn imply the (constraining) assumption $(\mathbf{A_4})$. Condition (3.2.16) is for instance satisfied when g_h is a positive continuous function such that

$$g_h(x) \overset{|x| \to +\infty}{\sim} \frac{1}{|x|^\alpha} \quad \text{with } \alpha > 1.$$

This allows to consider a large family of Lévy processes (such as symmetric stable processes) but unfortunately, it does not allow to consider the Brownian motion. In this case, it seems difficult to obtain the "lower/upper-bound" of the kernel. Nevertheless, we can only lower-

bound and this implies that one preserves the first part of Theorem 3.2.8.

Remark 3.2.17. Let us finally comment about the mean-reverting SDEs with sublinear drift, such as Ornstein-Uhlenbeck processes. We do not consider such dynamics in the non-compact case since Assumptions $(\mathbf{A_3})$ and $(\mathbf{A_4})$ do not hold in this case. Nevertheless, we show in Lemma 3.7.1 that Assumption $(\mathbf{A_V})$ may be true in this setting. Thus, one can obtain tightness of (μ_n). It is interesting to remark that such a property is new with respect to [28] (where only the compact case was studied), even in the case of linear dynamics (by linear, we mean without interaction). Owing to an adaptation of [28, Theorem 6.4] (which does not require $(\mathbf{A_3})$ and $(\mathbf{A_4})$), it may be shown that, in the linear setting, every limit point of (μ_n) is a QSD.

3.2.3 Simulations

We provide several numerical illustrations related to the Euler scheme of the following example

$$d\xi_t = \gamma \mathbb{E}[\xi_t | \tau_D > t]\, dt + dW_t,$$

with $D =]-1,1[$. This example comes from [223] (Example 1.6) where the authors are able to give explicit expressions of the related QSDs, taking advantage of the fact that $b(x,\mu) = \int x\mu(dx)$, so that $x \mapsto b(x,\mu)$ is constant. In this simple example, it can be shown that the QSDs have densities given by

$$\pi_b \propto e^{bx} \cos(\frac{\pi}{2}x),$$

where b is a fixed point of the map

$$b \mapsto \tanh(\gamma b) - \frac{8\gamma b}{4\gamma^2 b^2 + \pi^2}.$$

It can be checked that

— If $\gamma \leq \frac{\pi^2}{\pi^2+8}$, $b = 0$ is the only fixed point so that the unique QSD is $\pi_0 = \frac{4}{\pi}\cos(\frac{\pi}{2}x)\mathbf{1}_{[-1,1]}(x)$.

— If $\gamma > \frac{\pi^2}{\pi^2+8}$, there are three fixed points $b_-, 0, b_+$ with $b_- = -b_+$ which lead to three QSDs denotes π_-, π_0 and π_+.

In Figure 3.1, we compare true (continuous line) and empirical densities (dotted lines). By empirical density, we mean here the convolution of μ_n with a Gaussian kernel. For $\gamma = 0.5$, the convergence holds to the unique QSD whereas for $\gamma = 4$, it is interesting to remark that

the three QSDs are possible limit points of the sequence. We chose the initial points -0.5, 0 and 0.5 in order to increase the probability of approximating π_-, π_0 and π_+ but, of course, starting from 0.5 for instance, there is a positive probability that it falls in the basin of attraction of π_- or π_0. Nevertheless, since the three equilibriums are isolated, it is probably true that the algorithm converges to one of the three QSDs. To finish these short numerical

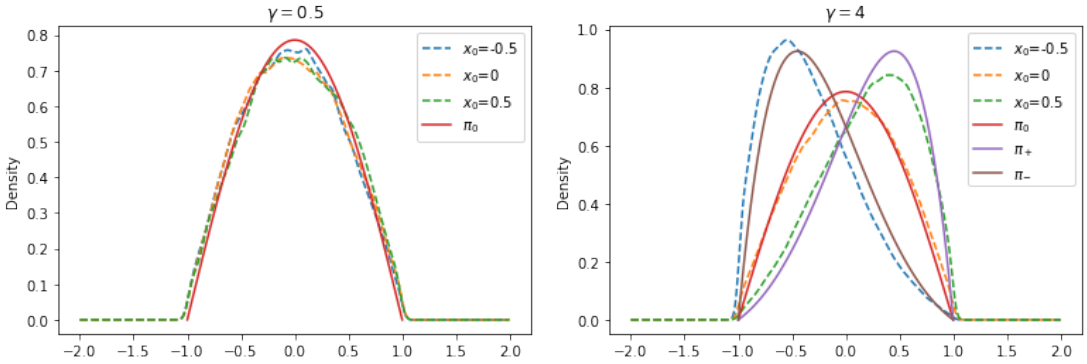

Figure 3.1 – Left. $(\gamma, h, n) = (0.5, 0.01, 10^5)$. Right. $(\gamma, h, n) = (4, 0.01, 5.10^5)$.

illustrations, we consider the case $\gamma = 1$ in Figure 3.2. In this case, there are three QSDs but our simulations suggest that only π_0 is attractive.

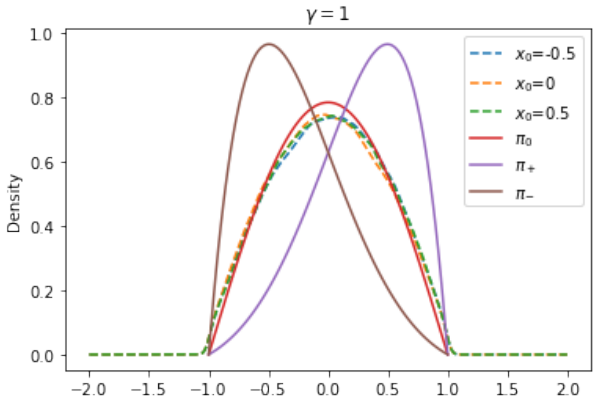

Figure 3.2 – $(\gamma, h, n) = (1, 0.01, 10^5)$.

3.3 Preliminaries, Poisson Equation and limiting dynamics

We begin by giving properties related to Π_μ and $\mathscr{A}_\mu$. Then, we introduce and study the Poisson equation related to the kernel $\mathbb{K}_\mu$. This study is one of the cornerstones of proposi-

tion 3.3.6 which will allow us to characterize the limiting dynamics of $(\tilde{\mu}^{(n)})_{n\geq 1}$.

Explicit form for Π_μ. Recall that $\mathscr{A}_\mu$ is a kernel defined by

$$\mathscr{A}_\mu(x,.) = \sum_{n\geq 0} \mathrm{K}^n_{\mu,\partial}(x,.).$$

Lemma 3.3.1. Assume $(\mathbf{A_0})$ and $(\mathbf{A_1})$. Then,

(i) $\mathscr{A}_\mu$ is well-defined for every $\mu \in \mathscr{P}(\mathscr{E})$ and $\|\mathscr{A}_\mu \mathbf{1}\|_\infty < +\infty$.

(ii) The probability Π_μ defined by

$$\Pi_\mu = \frac{\mu\mathscr{A}_\mu}{(\mu\mathscr{A}_\mu)(\mathbf{1})} \tag{3.3.1}$$

is the unique invariant distribution for $\mathbb{K}_\mu$.

(iii) $\mu^\star$ is a QSD for $(\mathrm{K}_{\mu,\partial})_{\mu\in\mathscr{P}(\mathscr{E})}$ if and only if $\mu^\star = \Pi_{\mu^\star}$. Existence holds for the QSD under the assumptions of Lemma 3.2.7(ii).

Proof of Lemma 3.3.1. (i) By $(\mathbf{A_0})$ and an induction,

$$\mathbb{P}_x(\tau_\mu > n\ell) = \mathrm{K}^{n\ell}_{\mu,\partial}\mathbf{1}(x) \leq \rho^n. \tag{3.3.2}$$

Thus,

$$\mathscr{A}_\mu\mathbf{1}(x) = \sum_{n\geq 0} \mathbb{P}_x(\tau_\mu > n) = \mathbb{E}_x[\tau] \leq \frac{\ell}{1-\rho}.$$

(ii) For any bounded continuous function f,

$$\begin{aligned}\mu\mathscr{A}_\mu\mathbb{K}_\mu(f) &= \mu\left(\sum_{n\geq 0}(\mathrm{K}^{n+1}_{\mu,\partial}f + \mathrm{K}^n_{\mu,\partial}\delta_\mu\mu(f))\right) = \sum_{n\geq 0}\mu\mathrm{K}^{n+1}_{\mu,\partial}f + \mu(f)\mu(\sum_{n\geq 0}\mathrm{K}^n_{\mu,\partial}(\delta_\mu)) \\ &= \mu\mathscr{A}_\mu f\end{aligned}$$

since $\delta_\mu = \mathbf{1} - \mathrm{K}_{\mu,\partial}\mathbf{1}$, which implies that

$$\sum_{n\geq 0}\mathrm{K}^n_{\mu,\partial}(\delta_\mu)) = \sum_{n\geq 0}\mathrm{K}^n_{\mu,\partial}\mathbf{1} - \sum_{n\geq 0}\mathrm{K}^{n+1}_{\mu,\partial}\mathbf{1} = \mathbf{1}.$$

To prove uniqueness, let us set

$$\mathrm{P}_\mu = e^{-1} \sum_{k\geq 0} \frac{\mathbb{K}_\mu^k}{k!}. \tag{3.3.3}$$

One easily checks that if Π_μ is an invariant distribution for $\mathbb{K}_\mu$, Π_μ is also invariant for P_μ and hence, it is enough to prove uniqueness for P_μ. By (3.8.1) and the very definition of $\mathbb{K}_\mu$, we have for every $k \geq 1$,

$$\mathbb{K}_\mu^k \geq (\mathrm{K}_{\mu,\partial}^{k-1}\delta)\mu = (\mathrm{K}_{\mu,\partial}^{k-1}\mathbf{1} - \mathrm{K}_{\mu,\partial}^{k}\mathbf{1})\mu,$$

so that there exists $c > 0$ such that

$$\mathrm{P}_\mu \geq c\sum_{k=1}^{\ell-1} \mathbb{K}_\mu^k \geq c\left(\sum_{k=1}^{\ell-1} \mathrm{K}_{\mu,\partial}^{k-1}\mathbf{1} - \mathrm{K}_{\mu,\partial}^{k}\mathbf{1}\right)\mu = c(1-\mathrm{K}_{\mu,\partial}^{\ell}\mathbf{1})\mu \geq \varepsilon\mu$$

with $\varepsilon = c(1-\rho) > 0$. This property classically implies uniqueness of the invariant distribution of P_μ (by a coupling argument).

(iii) The definition of $\mathbb{K}_\mu$ implies that a probability Π_μ is invariant if and only if:

$$\Pi_\mu \mathrm{K}_{\mu,\partial} = (\Pi_\mu \mathrm{K}_{\mu,\partial}\mathbf{1})\mu.$$

Thus, $\mu = \Pi_\mu$ if and only if

$$\mu\mathrm{K}_{\mu,\partial} = (\mu\mathrm{K}_{\mu,\partial}\mathbf{1})\mu. \tag{3.3.4}$$

Now, if μ is a QSD, then (3.3.4) holds true since it corresponds to the definition at time 1: for all Borel set A of $\mathscr{E}$,

$$\mu(\mathrm{A})(\mu\mathrm{K}_{\mu,\partial}\mathbf{1}) = \mu(\mathrm{A})\mathbb{P}_\mu(\tau > 1) = \mathbb{P}_\mu(\mathrm{Y}_1 \in \mathrm{A}, \tau > 1) = (\mu\mathrm{K}_{\mu,\partial})(\mathrm{A}).$$

Conversely, if (3.3.4) holds, let us first remark that μ is trivially a QSD if $\mu\mathrm{K}_{\mu,\partial}\mathbf{1} = 0$. Otherwise, one shows by induction that (3.2.3) holds true. The initialization follows from the above equality. If the property is true until time n, then $\mathbb{P}_\mu(\tau > n) = (\mu\mathrm{K}_{\mu,\partial}\mathbf{1})^n$ and $\mathscr{L}(\mathrm{Y}_n|\tau > n) = \mu$. Noting that for all Borel set A of $\mathscr{E}$, $\{\mathrm{Y}_{n+1} \in \mathrm{A}\} \subset \{\tau > n+1\}$, we then deduce from the Markov property that

$$\begin{aligned}\mathbb{P}(\mathrm{Y}_{n+1} \in \mathrm{A}, \tau > n+1) &= \mathbb{P}(\mathrm{Y}_{n+1} \in \mathrm{A}, \tau > n) = \mathbb{E}_\mu[\mathrm{K}_{\mu,\partial}(x,\mathrm{A})\mathbf{1}_{\tau>n}] \\ &= \mu\mathrm{K}_{\mu,\partial}(\mathrm{A})\mathbb{P}_\mu(\tau > n),\end{aligned} \tag{3.3.5}$$

where in the last equality, we used the recurrence hypothesis applied to f defined by $f(x) = \mathrm{K}_{\mu,\partial}(x,\mathrm{A})$. By (3.3.4), we deduce that

$$\mathbb{P}(\mathrm{Y}_{n+1} \in \mathrm{A}, \tau > n+1) = \mu(\mathrm{A})(\mu\mathrm{K}_{\mu,\partial}\mathbf{1})\mathbb{P}_\mu(\tau > n) = \mu(\mathrm{A})\mathbb{P}_\mu(\tau > n+1),$$

where in the last inequality, we used (3.3.5) with $\mathrm{A} = \mathscr{E}$. The equivalence is proved.

Let us prove existence of the QSD. First, note that if there exists $\mu \in \mathscr{P}(\mathscr{E})$ such that $\mu\mathrm{K}_{\mu,\partial}\mathbf{1} = 0$, then μ is a trivial QSD (since $\mathbb{P}_\mu(\tau > 1) = 0$ in this case). Thus, we can suppose that $\mu\mathrm{K}_{\mu,\partial}\mathbf{1} > 0$ for every $\mu \in \mathscr{P}(\mathscr{E})$. By $(\mathbf{A_1})$, this involves that the map

$$\Upsilon : \mu \mapsto \frac{\mu\mathrm{K}_{\mu,\partial}}{\mu\mathrm{K}_{\mu,\partial}\mathbf{1}}$$

is continuous on $\mathscr{P}(\mathscr{E})$. From the first part of the proof of (iii) (see (3.3.4)), a fixed point of Υ is a QSD. Thus, it follows from Kakutani fixed point theorem that existence will hold for the QSD as soon as there exist ε and M such that Υ is stable on the convex compact [7] subspace $\mathscr{V}_{\mathrm{M},\varepsilon}$ of $\mathscr{P}(\mathscr{E})$ defined (with the notations of $(\mathbf{A_V})$) by

$$\mathscr{V}_{\mathrm{M},\varepsilon} = \{\mu \in \mathscr{P}(\mathscr{E}), \mu(\mathrm{V}_\varepsilon) \le \mathrm{M}\}.$$

If $\alpha := \inf_{\mu\in\mathscr{P}(\mathscr{E})} \mu\mathrm{K}_{\mu,\partial}\mathbf{1} > 0$, and $(\mathbf{A_V})$ holds, then for all $\varepsilon > 0$,

$$\Upsilon(\mu)(\mathrm{V}_\varepsilon) \le \frac{\varepsilon}{\alpha}\mu(\mathrm{V}_\varepsilon) + \frac{\beta_\varepsilon}{\alpha}.$$

Taking $\varepsilon = \frac{\alpha}{2}$, we get

$$\forall \mu \in \mathscr{P}(\mathscr{E}), \quad \Upsilon(\mu)(\mathrm{V}_\varepsilon) \le \frac{1}{2}\mu(\mathrm{V}_{\frac{\alpha}{2}}) + \frac{\beta_{\alpha/2}}{\alpha}.$$

Thus, setting $\mathrm{M} = \frac{2\beta_{\alpha/2}}{\alpha}$, we get

$$\Upsilon(\mathscr{V}_{\mathrm{M},\frac{\alpha}{2}}) \subset \mathscr{V}_{\mathrm{M},\frac{\alpha}{2}}.$$

This concludes the proof. ■

Lemma 3.3.2. Assume $(\mathbf{A_0})$, $(\mathbf{A_1})$ and $(\mathbf{A_2})$. Then, $\mu \mapsto \mathscr{A}_\mu$ is Lipschitz for the TV-distance, uniformly in $x \in \mathscr{E}$, and $\mu \mapsto \Pi_\mu$ is Lipschitz for the TV-distance.

7. The set $\mathscr{V}_{\mathrm{M},\varepsilon}$ is clearly convex and the inf-compactness of V_ε implies that $\mathscr{V}_{\mathrm{M},\varepsilon}$ is a closed tight subset of $\mathscr{P}(\mathscr{E})$ and thus a compact subset of $\mathscr{P}(\mathscr{E})$.

Proof. Recall that $\mathscr{A}_\mu = \sum_{n\geq 0} \mathrm{K}_{\mu,\partial}^n$. We use the following decomposition:

$$\mathrm{K}_{\mu,\partial}^n - \mathrm{K}_{\nu,\partial}^n = \sum_{k=1}^{n} \mathrm{K}_{\mu,\partial}^{n-k}(\mathrm{K}_{\mu,\partial} - \mathrm{K}_{\nu,\partial})\mathrm{K}_{\nu,\partial}^{k-1}.$$

By $(\mathbf{A_0})$, $\sup_\mu \|\mathrm{K}_{\mu,\partial}^\ell \mathbf{1}\|_\infty < 1$, so that using that $n \mapsto \mathrm{K}_{\mu,\partial}^n \mathbf{1}$ is non-increasing, we deduce that there exists a constant C and a $\rho < 1$ such that

$$\sup_\mu \|\mathrm{K}_{\mu,\partial}^n \mathbf{1}\|_\infty \leq \mathrm{C}\rho^n.$$

Thus, for every bounded measurable function f on $\mathscr{E}$, for every $x \in \mathscr{E}$,

$$\begin{aligned}|\mathrm{K}_{\mu,\partial}^{n-k}(\mathrm{K}_{\mu,\partial} - \mathrm{K}_{\nu,\partial})\mathrm{K}_{\nu,\partial}^{k-1} f(x,.) &\leq \mathrm{C}\rho^{n-k}\|(\mathrm{K}_{\mu,\partial} - \mathrm{K}_{\nu,\partial})\mathrm{K}_{\nu,\partial}^{k-1} f\|_\infty \\ &\leq \mathrm{C}\rho^{n-k}\|\mu-\nu\|_{\mathrm{TV}}\|\mathrm{K}_{\nu,\partial}^{k-1} f\|_\infty \leq \mathrm{C}\|f\|_\infty \rho^n \|\mu-\nu\|_{\mathrm{TV}},\end{aligned}$$

where C may have changed from line to line and where we used $(\mathbf{A_2})$. This means that

$$\sup_{x\in\mathscr{E}} \|\mathrm{K}_{\mu,\partial}^{n-k}(\mathrm{K}_{\mu,\partial} - \mathrm{K}_{\nu,\partial})\mathrm{K}_{\nu,\partial}^{k-1}(x,.)\|_{\mathrm{TV}} \leq \mathrm{C}\rho^n \|\mu-\nu\|_{\mathrm{TV}},$$

and it follows that

$$\sup_{x\in\mathscr{E}} \|\mathscr{A}_\mu(x,.) - \mathscr{A}_\nu(x,.)\|_{\mathrm{TV}} \leq \sum n\rho^n < +\infty.$$

This easily implies that $\mu \mapsto \mu\mathscr{A}_\mu$ and thus $\mu \mapsto \mathscr{A}_\mu \mathbf{1}$ are Lipschitz for the TV-norm. Using that $\mu\mathscr{A}_\mu \mathbf{1} \geq 1$, one then deduces that $\mu \mapsto \Pi_\mu$ is also Lipschitz. ■

Poisson Equation related to $\mathbb{K}_\mu$. Let $(\mathbb{K}_\mu)_{\mu\in\mathscr{P}(\mathscr{E})}$ denote a family of transition kernels satisfying the following assumptions:

$(\mathbf{P_0})$: For every $\mu \in \mathscr{P}(\mathscr{E})$, $\mathbb{K}_\mu$ admits a unique invariant distribution Π_μ and $\mu \mapsto \Pi_\mu$ is Lipschitz continuous on $\mathscr{P}(\mathscr{E})$ for the TV-distance.

$(\mathbf{P_1})$: There exist $\ell \in \mathbb{N}$, $\ell \geq 1$, $\varepsilon > 0$ and $\Psi \in \mathscr{P}(\mathscr{E})$ such that for every $\mu \in \mathscr{P}(\mathscr{E})$,

$$\mathbb{K}_\mu^\ell(., dy) \geq \varepsilon\Psi(dy).$$

$(\mathbf{P_2})$: $\mu \mapsto \mathbb{K}_\mu(x,.)$ is Lipschitz continuous for the TV-distance on $\mathscr{P}(\mathscr{E})$, uniformly in x; namely,

there exists $c_2 > 0$ such that for every $x \in \mathscr{P}(\mathscr{E})$ and every μ and $\nu \in \mathscr{P}(\mathscr{E})$,

$$\|\mathbb{K}_\mu(x,.) - \mathbb{K}_\nu(x,.)\|_{\mathrm{TV}} \le c_2 \|\mu - \nu\|_{\mathrm{TV}}.$$

For a fixed μ and a given function $f : \mathscr{E} \to \mathbb{R}$, let us consider the Poisson equation

$$f - \Pi_\mu f = (\mathrm{Id} - \mathbb{K}_\mu) g. \tag{3.3.6}$$

We have the following result:

Lemma 3.3.3 (Poisson equation). Assume **(P$_0$)**, **(P$_1$)** and **(P$_2$)**. For any $\mu \in \mathscr{P}(\mathscr{E})$, for any measurable function $f : \mathscr{E} \to \mathbb{R}$, the Poisson equation (3.3.6) admits a unique solution denoted $\mathrm{Q}_\mu f$ and defined by

$$\mathrm{Q}_\mu f(x) = \sum_{n=0}^{+\infty} \left(\mathbb{K}_\mu^n f(x) - \Pi_\mu(f) \right). \tag{3.3.7}$$

Furthermore, there exists a constant C such that,

(a) For every $\mu, \alpha \in \mathscr{P}(\mathscr{E})$,

$$|\alpha \mathrm{Q}_\mu f| \le \mathrm{C} \|f\|_\infty \|\alpha - \Pi_\mu\|_{\mathrm{VT}}.$$

(b) For every $\mu \in \mathscr{P}(\mathscr{E})$,

$$\|\mathrm{Q}_\mu f - \mathrm{Q}_\nu f\|_\infty \le \mathrm{C} \|f\|_\infty \|\mu - \nu\|_{\mathrm{TV}}.$$

Remark 3.3.4. $\triangleright$ This result will be applied to $\mathbb{K}_\mu$ defined by (3.2.4) but it applies to every family of kernels satisfying **(P$_0$)**, **(P$_1$)** and **(P$_2$)**.

$\triangleright$ In the proof of this result, one can check that the fact that Ψ does not depend on μ only plays a role in (3.8.2) and that the result is still true if Assumption **(P$_1$)** is satisfied with a probability Ψ_μ where $\mu \mapsto \Psi_\mu$ is Lipschitz continuous for the TV-distance. For instance, if

$$\varepsilon = \inf_{(x,\mu) \in \mathscr{P}(\mathscr{E})} \delta_\mu(x) > 0, \tag{3.3.8}$$

the rebirth part of the dynamics implies that

$$\mathbb{K}_\mu \ge \varepsilon \mu$$

In this case, the result is still true. Such a variation may allow to remove **(A$_3$)** from the assumptions of Theorem 3.2.8(i). One way to definitely remove **(A$_3$)** from Theorem 3.2.8(i)

(without assuming the additional assumption (3.3.8) for instance) is to use a Poissonization argument like in (3.3.3) which allows to lower-bound the (Poissonized) dynamics by $\varepsilon\mu$ with very slight assumptions. For details in this direction, we refer to [28, Lemma 6.3].

Proof. *(a)* We prove (a) and the well-definition of $Q_\mu f$ in the same time. By $(\mathbf{P_1})$,

$$\forall x \in \mathscr{E}, \quad \mathbb{K}_\mu^\ell(x,.) = \varepsilon\Psi(.) + (1-\varepsilon)R_\mu(x,.)$$

where

$$R_\mu = \frac{\mathbb{K}_\mu^\ell - \varepsilon\Psi}{1-\varepsilon},$$

is a transition kernel. Using that for a probability $\alpha \in \mathscr{P}(\mathscr{E})$, $\alpha\Psi = \Psi$, it follows that for any $\alpha,\beta \in \mathscr{P}(\mathscr{E})$.

$$(\alpha-\beta)\mathbb{K}_\mu^\ell = (1-\varepsilon)(\alpha-\beta)R_\mu.$$

Then, an induction leads to

$$(\alpha-\beta)\mathbb{K}_\mu^{n\ell} = (1-\varepsilon)^n(\alpha-\beta)R_\mu^n. \tag{3.3.9}$$

and to

$$\|\alpha\mathbb{K}_\mu^{n\ell} - \beta\mathbb{K}_\mu^{n\ell}\|_{TV} \le (1-\varepsilon)^n\|\alpha-\beta\|_{TV},$$

for any $\alpha,\beta \in \mathscr{P}(\mathscr{E})$. In particular,

$$\|\alpha\mathbb{K}_\mu^{n\ell} - \Pi_\mu\|_{TV} \le (1-\varepsilon)^n\|\alpha-\Pi_\mu\|_{TV},$$

and for every $r \in [\![0,\ell-1]\!]$,

$$\|\alpha\mathbb{K}_\mu^{n\ell+r} - \Pi_\mu\|_{TV} \le (1-\varepsilon)^n\|\alpha\mathbb{K}_\mu^r - \Pi_\mu\mathbb{K}_\mu^r\|_{TV} \le (1-\varepsilon)^n\|\alpha-\Pi_\mu\|_{TV}.$$

In the above inequality, we used that if K is a transition kernel $\|\alpha K - \beta K\|_{TV} \le \|\alpha-\beta\|_{TV}$. These inequalities easily lead to (a) and to the well-definition of Q_μ. Checking that $(\mathrm{Id}-K_\mu)Q_\mu f = Q_\mu(\mathrm{Id}-K_\mu)f$ classically implies that $Q_\mu f$ is the unique solution of the Poisson equation.

(b) Let us prove that $\mu \to Q_\mu f(x)$ is Lipschitz continuous. From the definition of Q_μ and from

(3.3.9), we have

$$\begin{aligned}
Q_\mu f(x) - Q_\nu f(x) &= \sum_{n=0}^{+\infty} \Big((\mathbb{K}_\mu^n f(x) - \Pi_\mu \mathbb{K}_\mu^n(f)) - \big(\mathbb{K}_\nu^n f(x) - \Pi_\nu \mathbb{K}_\nu^n(f)\big) \Big) \\
&= \sum_{n=0}^{+\infty} \sum_{j=0}^{\ell-1} \Big((\mathbb{K}_\mu^{n\ell+j} f(x) - \Pi_\mu \mathbb{K}_\mu^{n\ell+j}(f)) - (\mathbb{K}_\nu^{n\ell+j} f(x) - \Pi_\nu \mathbb{K}_\nu^{n\ell+j}(f)) \Big) \\
&= \sum_{n=0}^{+\infty} (1-\varepsilon)^n \sum_{j=0}^{\ell-1} \Big((\delta_x - \Pi_\mu) R_\mu^n \mathbb{K}_\mu^j f - (\delta_x - \Pi_\nu) R_\nu^n \mathbb{K}_\nu^j f \Big).
\end{aligned}$$

Now, assume without loss of generality that $\|f\|_\infty \le 1$. In this case, one can check that for every $n \ge 0$, $j \in [\![0, \ell-1]\!]$,

$$\begin{aligned}
\left| (\delta_x - \Pi_\mu) R_\mu^n \mathbb{K}_\mu^j f - (\delta_x - \Pi_\nu) R_\nu^n \mathbb{K}_\nu^j f \right| &\le 2 \sup_\alpha \|\alpha R_\mu^n \mathbb{K}_\mu^j - \alpha R_\nu^n \mathbb{K}_\nu^j\|_{TV} + \|\Pi_\mu - \Pi_\nu\|_{TV} \\
&\le 2 \sup_\alpha \|\alpha R_\mu^n - \alpha R_\nu^n\|_{TV} + 2 \sup_\alpha \|\alpha \mathbb{K}_\mu^j - \alpha \mathbb{K}_\nu^j\|_{TV} + \|\Pi_\mu - \Pi_\nu\|_{TV}.
\end{aligned}$$

By **($\mathbf{P_0}$)** and Lemma 3.3.5, one deduces that a constant C exists such that

$$\left| (\delta_x - \Pi_\mu) R_\mu^n \mathbb{K}_\mu^j f - (\delta_x - \Pi_\nu) R_\nu^n \mathbb{K}_\nu^j f \right| \le C(1+n) \|\mu - \nu\|_{TV}.$$

Thus,

$$|Q_\mu f(x) - Q_\nu f(x)| \lesssim \|\mu - \nu\|_{TV} \sum_{n=0}^{+\infty} (1-\varepsilon)^n (1+n) \lesssim \|\mu - \nu\|_{TV}.$$

This concludes the proof.

■

Lemma 3.3.5. Assume **($\mathbf{P_0}$)**, **($\mathbf{P_1}$)** and **($\mathbf{P_2}$)**. Then,

(i) For all $n \in \mathbb{N}$,

$$\sup_{\alpha \in \mathscr{P}(\mathscr{E})} \|\alpha \mathbb{K}_\mu^n - \alpha \mathbb{K}_\nu^n\|_{TV} \le c_2 n \|\mu - \nu\|_{TV}.$$

(ii) Setting $R_\mu = (1-\varepsilon)^{-1}(\mathbb{K}_\mu^\ell - \varepsilon \Psi)$, we have

$$\sup_{\alpha \in \mathscr{P}(\mathscr{E})} \|\alpha R_\mu^n - \alpha R_\nu^n\|_{TV} \le \frac{c_2 n \ell}{1-\varepsilon} \|\mu - \nu\|_{TV}.$$

The proof of lemma 3.3.5 is postponed to section 3.8. In the proposition below, we obtain a result which will lead to the identification of the limit. Note that this result does not directly depend on the dynamics of our algorithm but only on the generic assumptions **($\mathbf{P_0}$)**, **($\mathbf{P_1}$)**

and **(P2)**. For the sake of simplicity, we use the notations of the algorithm but write *ad hoc* assumptions (which are satisfied in this setting) under which the conclusion is true.

Proposition 3.3.6. Assume **(P0)**, **(P1)** and **(P2)**. Let $(X_n, \mu_n)_{n\geq 1}$ denote a random sequence with values in $\mathscr{E} \times \mathscr{P}(\mathscr{E})$ such that μ_n is $\sigma(X_k, k \leq n-1)$-measurable and $\mathscr{L}(X_{n+1}|\sigma(X_k, k \leq n)) = \mathbb{K}_{\mu_n}(X_n, .)$. Then, if $(\gamma_n)_{n\geq 1}$ denotes a non-increasing sequence of positive numbers satisfying (3.2.7) and $\|\mu_{n+1} - \mu_n\|_{TV} \leq C\gamma_{n+1}$ *a.s.* where C does not depend on n, then for any $T > 0$, for any bounded measurable function $f : \mathscr{E} \mapsto \mathbb{R}$

$$\lim_{n\to+\infty} \sup_{m\in[\![n,N(n,T)]\!]} \left| \sum_{k=n}^{m} \gamma_k \left(f(X_k) - \Pi_{\mu_k}(f) \right) \right| = 0,$$

where

$$N(n,T) := \inf\{m \geq n, \sum_{k=n}^{m+1} \gamma_k > T\}.$$

Proof. Without loss of generality, we assume that $\|f\|_\infty \leq 1$. By (3.3.6),

$$\begin{aligned} f(X_{n+1}) - \Pi_{\mu_n}(f) &= Q_{\mu_n} f(X_{n+1}) - \mathbb{K}_{\mu_n} Q_{\mu_n} f(X_{n+1}) \\ &= \gamma_{n+1} \Delta M_{n+1}(f) + \Delta R_{n+1}(f) + \gamma_{n+1} \Delta D_{n+1}(f), \end{aligned}$$

with,

$$\begin{aligned} \Delta M_{n+1}(f) &= Q_{\mu_n} f(X_{n+1}) - \mathbb{K}_{\mu_n} Q_{\mu_n} f(X_n) \\ \Delta R_{n+1}(f) &= (\gamma_{n+1} - \gamma_n) \mathbb{K}_{\mu_n} Q_{\mu_n} f(X_n) + \left(\gamma_n \mathbb{K}_{\mu_n} Q_{\mu_n} f(X_n) - \gamma_{n+1} \mathbb{K}_{\mu_{n+1}} Q_{\mu_{n+1}} f(X_{n+1}) \right) \\ \Delta D_{n+1}(f) &= \left(\mathbb{K}_{\mu_{n+1}} Q_{\mu_{n+1}} f(X_{n+1}) - \mathbb{K}_{\mu_n} Q_{\mu_n} f(X_{n+1}) \right). \end{aligned}$$

First, let us focus on ΔR_{n+1}. Using for the first part that $(\gamma_n)_{n\geq 0}$ is non-increasing and that $(x, \mu) \mapsto \mathbb{K}_\mu Q_\mu f(x)$ is (uniformly) bounded (by Lemma 3.3.3), and a telescoping argument for the second part yields the existence of a constant C such that for any positive integer m:

$$\left| \sum_{k=n}^{m} \Delta R_k(f) \right| \leq C\gamma_n.$$

Second (ΔM_n) is a sequence of $(\mathscr{F}_n)$-martingale increments. From Lemma 3.3.3, $\Delta M_n(f)$ is bounded. As a consequence, by an adaptation of [21, Proposition 4.4] (based on exponential martingales), one is able to obtain that for every $T > 0$, some positive constants C_1 and C_2

exist such that

$$\mathbb{P}\left(\max_{m\in[\![n,\mathrm{N}(n,\mathrm{T})]\!]}\left|\sum_{k=n}^{m}\gamma_k\Delta\mathrm{M}_k(f)\right|\geq\varepsilon\right)\leq \mathrm{C}_1\exp\left(-\mathrm{C}_2\gamma_n^{-1}\varepsilon^2\right),$$

and thus, from the Borel-Cantelli lemma and the assumption $\lim_{n\to+\infty}\gamma_n\log(n)=0$, that for every $\mathrm{T}>0$

$$\limsup_{n\to+\infty}\max_{m\in[\![n,\mathrm{N}(n,\mathrm{T})]\!]}\left|\sum_{k=n}^{m}\gamma_k\Delta\mathrm{M}_k(f)\right|=0\quad a.s.$$

Finally, for the last term, one uses that $\mu\mapsto\mathbb{K}_\mu f$ and $\mu\mapsto\mathrm{Q}_\mu f$ are Lipschitz continuous. More precisely, using Lemma 3.3.3 *(i)*, *(iii)* and the last inequality of Lemma 3.3.5, we see that there exists $\mathrm{C}>0$ such that

$$\begin{aligned}
&\left|\mathbb{K}_{\mu_{n+1}}\mathrm{Q}_{\mu_{n+1}}f(\mathrm{X}_{n+1})-\mathbb{K}_{\mu_n}\mathrm{Q}_{\mu_n}f(\mathrm{X}_{n+1})\right|\\
&\leq\left|\mathbb{K}_{\mu_{n+1}}\mathrm{Q}_{\mu_{n+1}}f(\mathrm{X}_{n+1})-\mathbb{K}_{\mu_{n+1}}\mathrm{Q}_{\mu_n}f(\mathrm{X}_{n+1})\right|+\left|\mathbb{K}_{\mu_{n+1}}\mathrm{Q}_{\mu_n}f(\mathrm{X}_{n+1})-\mathbb{K}_{\mu_n}\mathrm{Q}_{\mu_n}f(\mathrm{X}_{n+1})\right|\\
&\leq\|\mathrm{Q}_{\mu_{n+1}}f-\mathrm{Q}_{\mu_n}f\|_\infty+\left\|(\mathbb{K}_{\mu_{n+1}}-\mathbb{K}_{\mu_n})(\mathrm{Q}_{\mu_n}f)\right\|_\infty\\
&\leq\mathrm{C}\|f\|_\infty\|\mu_{n+1}-\mu_n\|_{\mathrm{TV}}\leq\mathrm{C}\gamma_{n+1},
\end{aligned}$$

This ends the proof (since $\sum_{k=n+1}^{\mathrm{N}(n,\mathrm{T})}|\Delta\mathrm{D}_k(f)|\leq\mathrm{C}\sum_{k=n+1}^{\mathrm{N}(n,\mathrm{T})}\gamma_k\leq\mathrm{CT}$).

■

3.4 Long-time behavior of the limiting dynamics

The aim of this section is to provide long-time properties of the limiting ODE

$$\dot{\nu}=-\nu+\Pi_\nu. \tag{3.4.1}$$

To this end, we use that this equation can be connected to an "almost" linear differential equation through the following property (see remark 3.4.2 for details).

Lemma 3.4.1. Assume $(\mathbf{A_0})$, $(\mathbf{A_1})$ and $(\mathbf{A_2})$.

(i) For any $\mu_0\in\mathscr{P}(\mathscr{E})$, the differential equation

$$\dot{\mu}_t=\mu_t\mathscr{A}_{\mu_t}-(\mu_t\mathscr{A}_{\mu_t}\mathbf{1})\mu_t. \tag{3.4.2}$$

admits a unique solution in $\mathscr{P}(\mathscr{E})$. Furthermore,

$$\forall\, t \geq 0, \quad \mu_t = \frac{\nu_0 \varphi_t}{\nu_0 \varphi_t \mathbf{1}}, \tag{3.4.3}$$

where $(\varphi_t)_{t\geq 0}$ is the unique solution on $(\mathscr{B}(\mathscr{E}), \|.\|_\infty)$, space of bounded linear operators on $\mathscr{C}(\mathscr{E}, \mathbb{R}_+)$ starting from Id to

$$\dot{\varphi} = \varphi_t \mathscr{A}_{\mu_t}, \quad t \geq 0. \tag{3.4.4}$$

(ii) For every $\nu_0 \in \mathscr{P}(\mathscr{E})$, the differential equation (3.4.1) admits a unique solution $(\nu_t)_{t\geq 0}$ in $\mathscr{P}(\mathscr{E})$, starting from ν_0. Furthermore, if $(\mu_t)_{t\geq 0}$ is the solution of (3.4.2), then $\nu_t := \mu_{\tau(t)}$ where τ is a $\mathscr{C}^1$-diffeomorphism from $\mathbb{R}_+$ to $\mathbb{R}_+$ being the unique solution of

$$\tau'(t) = \frac{1}{\mu_{\tau(t)} \mathscr{A}_{\mu_{\tau(t)}} \mathbf{1}} \quad \text{with } \tau(0) = 0. \tag{3.4.5}$$

Remark 3.4.2. For a given $(\mu_t)_{t\geq 0}$, the equation (3.4.4) is an inhomogeneous linear equation. This will allow us to provide a series expansion of its solution (see (3.4.28)) and thus to $(\mu_t)_{t\geq 0}$ by (3.4.3) (which depends on $(\mu_t)_{t\geq 0}$ itself !). This is why we used the terminology "almost linear".

Proof. (i) To prove existence and uniqueness of the solution, we apply the Cauchy-Lipschitz theorem on $\mathscr{M}_s(\mathscr{E})$ which denotes the space of finite signed measures on $\mathscr{E}$ equipped with the total variation norm $\|.\|_{\mathrm{TV}}$. It is well-known that $\mathscr{M}_s(\mathscr{E})$ is a Banach space. Since the operator $\mathscr{A}_\mu$ is not defined on $\mathscr{M}_s(\mathscr{E}) \setminus \mathscr{P}(\mathscr{E})$, we extend the ODE on $\mathscr{M}_s(\mathscr{E})$ by considering

$$\dot{\mu} = \mu \mathscr{A}_{p(\mu)} - (\mu \mathscr{A}_{p(\mu)} \mathbf{1}) \mu, \tag{3.4.6}$$

where p is a Lipschitz continuous function from $\mathscr{M}_s(\mathscr{E})$ to $\mathscr{P}(\mathscr{E})$ which satisfies $p(\mu) = \mu$ on $\mathscr{P}(\mathscr{E})$. For instance, since $\mu \mapsto \mu_+$ is 1-Lipschitz for $\|.\|_{\mathrm{TV}}$, one easily checks that for any $\mu_0 \in \mathscr{P}(\mathscr{E})$

$$p(\mu) = \frac{\mu_+}{\|\mu_+\|_{\mathrm{TV}} \vee 1} + (1 - \|\mu_+\|_{\mathrm{TV}})_+ \mu_0,$$

has the required properties. Now, by Lemma 3.3.2, it is easily seen that

$$\mu \mapsto \mu \mathscr{A}_{p(\mu)} - (\mu \mathscr{A}_{p(\mu)} \mathbf{1}) \mu, \tag{3.4.7}$$

is Lipschitz continuous for $\|.\|_{\mathrm{TV}}$ on $\mathscr{M}_s(\mathscr{E})$ so that the Cauchy-Lipschitz theorem ensures existence and uniqueness of solutions on $\mathscr{M}_s(\mathscr{E})$ of (3.4.6). However, at this stage, we do not know if a solution starting from an initial condition in $\mathscr{P}(\mathscr{E})$ remains in $\mathscr{P}(\mathscr{E})$.

To this end, let us consider the ODE

$$\dot{\varphi} = \varphi_t \mathscr{A}_{p(\mu_t)}, \quad t \geq 0,$$

in $(\mathscr{B}(\mathscr{E}), \|.\|_\infty)$ (which is a Banach space). The function $\varphi \mapsto \varphi \mathscr{A}_{p(\mu_t)}$ is Lipschitz with respect to $\|.\|_\infty$, uniformly in t since by lemma 3.3.1,

$$\sup_{\mu \in \mathscr{P}(\mathscr{E})} \|\mathscr{A}_\mu\|_\infty = \sup_{\mu \in \mathscr{P}(\mathscr{E})} \sup_{\|f\|_\infty \leq 1} \|\mathscr{A}_\mu f\|_\infty = \sup_{\mu \in \mathscr{P}(\mathscr{E})} \|\mathscr{A}_\mu \mathbf{1}\|_\infty < +\infty. \tag{3.4.8}$$

Thus, the equation admits a unique solution $(\varphi_t)_{t\geq 0}$ starting from Id. For a given $\mu_0 \in \mathscr{P}(\mathscr{E})$, set

$$\rho_t := \frac{\mu_0 \varphi_t}{\mu_0 \varphi_t \mathbf{1}}, \tag{3.4.9}$$

Since $\mu_0 \varphi_t \mathbf{1} \geq 1$ for every $t \geq 0$, ρ_t is well-defined, differentiable on $\mathbb{R}_+$ and

$$\dot{\rho}_t = \frac{\mu_0 \varphi_t \mathscr{A}_{p(\mu_t)}}{\mu_0 \varphi_t \mathbf{1}} - \frac{(\mu_0 \varphi_t \mathscr{A}_{p(\mu_t)} \mathbf{1}) \mu_0 \varphi_t}{(\mu_0 \varphi_t \mathbf{1})^2} = \rho_t \mathscr{A}_{p(\mu_t)} - (\rho_t \mathscr{A}_{p(\mu_t)} \mathbf{1}) \rho_t. \tag{3.4.10}$$

The map $\rho \mapsto \rho \mathscr{A}_{p(\mu_t)} - (\rho \mathscr{A}_{p(\mu_t)} \mathbf{1})\rho$ being uniformly (in t) Lipschitz for the TV-norm, $(\rho_t)_{t\geq 0}$ is thus the unique solution (starting from μ_0) of the (linear) ODE $\dot{\rho}_t = \rho_t \mathscr{A}_{p(\mu_t)} - (\rho_t \mathscr{A}_{p(\mu_t)} \mathbf{1})\rho_t$. But, by construction, the solution $(\mu_t)_{t\geq 0}$ of the non linear equation (3.4.7) is also a solution of the linear ODE $\dot{\rho}_t = \rho_t \mathscr{A}_{p(\mu_t)} - (\rho_t \mathscr{A}_{p(\mu_t)} \mathbf{1})\rho_t$. Thus, $\rho_t = \mu_t$ and this implies that $(\mu_t)_{t\geq 0}$ lives in $\mathscr{P}(\mathscr{E})$ since $(\rho_t)_{t\geq 0}$ does (by its very definition (3.4.9)). As a consequence, $p(\mu_t) = \mu_t$ for all $t \geq 0$ and this concludes the proof of this statement.

(ii) Since $\mu \mapsto -\mu + \Pi_\mu$ is Lipschitz for the TV-distance (by lemma 3.3.1), uniqueness follows from a classical Gronwall argument. For the existence, we consider $\nu_t = \mu_{\tau(t)}$ where (μ_t) is a solution of (3.4.3) starting from ν_0 and τ is a solution of (3.4.5). Before going further, let us check that (3.4.5) has a solution. To this end, we remark that the map $\tau \mapsto \mu_\tau \mathscr{A}_{\mu_\tau} \mathbf{1}$ is Lipschitz since $\mu \mapsto \mathscr{A}_\mu \mathbf{1}$ is Lipschitz (by lemma 3.3.1) and $\tau \mapsto \mu_\tau$ also is for the TV-distance. Actually, by (3.4.2),

$$\forall 0 \leq \tau_1 \leq \tau_2, \quad \|\mu_{\tau_2} - \mu_{\tau_1}\|_{\mathrm{TV}} \leq 2\Big(\sup_{\mu \in \mathscr{P}(\mathscr{E})} \mu \mathscr{A}_\mu \mathbf{1}\Big)(\tau_2 - \tau_1) \leq C(\tau_2 - \tau_1),$$

by (3.4.8). Thus, $\tau \mapsto (\mu_\tau \mathscr{A}_{\mu_\tau} \mathbf{1})^{-1}$ is also Lipschitz since $\mu \mathscr{A}_\mu \mathbf{1} \geq 1$. As a consequence, (3.4.5) has a unique solution and by construction,

$$\forall\, t \geq 0, \quad \dot{\nu}_t = \tau'(t)\dot{\mu}_{\tau(t)} = \frac{\mu_{\tau(t)}\mathscr{A}_{\mu_{\tau(t)}}}{\mu_{\tau(t)}\mathscr{A}_{\mu_{\tau(t)}}\mathbf{1}} - \mu_{\tau(t)} = \Pi_{\nu_t} - \nu_t.$$

Finally, since $\mathscr{A}_\mu 1 \geq 1$ and $c = \sup_\mu \|\mathscr{A}_\mu \mathbf{1}\|_\infty < +\infty$,

$$\forall\, t \geq 0, \quad , c^{-1} \leq \tau'(t) \leq 1, \tag{3.4.11}$$

which in turn implies the $\mathscr{C}^1$-diffeomorphism property.

■

The aim is now to be able to study the asymptotic behavior of $(\mu_t)_{t\geq 0}$ defined by (3.4.3) and to deduce some convergence properties for $(\nu_t)_{t\geq 0}$ by time change (given by τ given by (3.4.5)). To this end, we introduce a general continuous function $\alpha = (\alpha_t)_{t\geq 0}$ with values in $\mathscr{P}(\mathscr{E})$ and consider the family of operators $(\mathrm{R}^\alpha_{t,s})_{s\leq t}$

$$\mathrm{R}^\alpha_{t,s} f := \frac{\mathfrak{S}^\alpha_{t,s}(f g^\alpha_s)}{g^\alpha_t} \tag{3.4.12}$$

where $\mathfrak{S}^\alpha_{t,s}$ is the unique solution [8] at time t starting from $\bar{\varphi} = \mathrm{Id}$ at time s to

$$\begin{cases} \dot{\varphi}_v = \varphi_v \mathscr{A}_{\alpha_v}, \quad v \in [s,t] \\ \varphi_s = \bar{\varphi}, \end{cases} \tag{3.4.13}$$

and

$$g^\alpha_t = \mathfrak{S}^\alpha_{t,0}\mathbf{1}.$$

A crucial point for the sequel is the cocycle property for $(\mathfrak{S}^\alpha_{t,s})_{s<t}$:

$$\forall\, 0 \leq u \leq s \leq t, \quad \mathrm{K}_{t,u} = \mathrm{K}_{t,s} \circ \mathrm{K}_{s,u}.$$

This property is a direct consequence of existence and uniqueness of solutions to the ODE (3.4.13). This in turn implies that $(\mathrm{R}^\alpha_{t,s})_{s<t}$ also has the cocycle property:

$$\forall\, 0 \leq u \leq s \leq t, \quad \mathrm{R}_{t,u} = \mathrm{R}_{t,s} \circ \mathrm{R}_{s,u}. \tag{3.4.14}$$

8. By (3.4.8), which holds under ($\mathbf{A_0}$) and ($\mathbf{A_1}$), existence and uniqueness hold.

Also note at this stage that $(\mu_t)_{t\geq 0}$ defined by (3.4.3) and $(\mathfrak{S}^\alpha_{t,s})_{s\leq t}$ are connected by the following relation:

$$\mu_t = \frac{\mu_0 \mathfrak{S}^{\mu_\bullet}_t}{\mu_0 \mathfrak{S}^{\mu_\bullet}_t \mathbf{1}} = \mathrm{R}^{\mu_\bullet}_t \frac{g^{\mu_\bullet}_t}{\mu_0 g^{\mu_\bullet}_t}, \tag{3.4.15}$$

where $\mathfrak{S}^\alpha_t = \mathfrak{S}^\alpha_{t,0}$ and $\mathrm{R}^\alpha_t = \mathrm{R}^\alpha_{t,0}$. In the following, we thus introduce for a given $\alpha := (\alpha_t)_{t\geq 0}$ the map $\mu \mapsto \Phi^\alpha_t(\mu)$ defined by

$$\Phi^\alpha_t(\mu) := \frac{\mu \mathfrak{S}^\alpha_t}{\mu \mathfrak{S}^\alpha_t \mathbf{1}}. \tag{3.4.16}$$

We have the following result:

Proposition 3.4.3. Assume **($\mathbf{A_0}$)** and **($\mathbf{A_1}$)** and that,

$$\|\delta_x \mathrm{R}^\alpha_t - \delta_y \mathrm{R}^\alpha_t\|_{\mathrm{TV}} \overset{t\to+\infty}{\longrightarrow} 0 \quad \text{uniformly in } (x,y,\alpha) \in \mathscr{E}\times\mathscr{E}\times\mathscr{F}(\mathbb{R}_+,\mathscr{P}(\mathrm{E})). \tag{3.4.17}$$

Then,

(i) There exists a map $\alpha \mapsto \mu^\star_\alpha$ from $\mathscr{C}(\mathbb{R}_+,\mathscr{P}(\mathscr{E}))$ to $\mathscr{P}(\mathscr{E})$ such that

$$\sup_{\mu,\alpha} \|\mu \mathrm{R}^\alpha_t - \mu^\star_\alpha\|_{\mathrm{TV}} \overset{t\to+\infty}{\longrightarrow} 0. \tag{3.4.18}$$

(ii) The convergence result also holds for $\Phi^\alpha_t(\mu)$:

$$\sup_{\mu,\alpha} \|\Phi^\alpha_t(\mu) - \mu^\star_\alpha\|_{\mathrm{TV}} \overset{t\to+\infty}{\longrightarrow} 0. \tag{3.4.19}$$

Remark 3.4.4. In this result, the limit does not depend on the starting distribution μ. Unfortunately, when we apply to $\alpha = (\mu_t)_{t\geq 0}$, *i.e.* on the distribution of the solution itself, the limit will *a priori* depend on μ_0.

Proof. (i) Let $f : \mathscr{E} \to \mathbb{R}$ denote a bounded function with $\|f\|_\infty \leq 1$. For all $\mu,\nu \in \mathscr{P}(\mathscr{E})$, we have

$$\begin{aligned} |\mu \mathrm{R}^\alpha_t f - \nu \mathrm{R}^\alpha_t f| &= \left| \int \Big(\mathrm{R}^\alpha_t f(x) - \mathrm{R}^\alpha_t f(y)\Big) \mu(dx)\nu(dy) \right| \\ &\leq \sup_{x,y} |\mathrm{R}^\alpha_t f(x) - \mathrm{R}^\alpha_t f(y)|. \end{aligned} \tag{3.4.20}$$

As $\|\delta_x \mathrm{R}^\alpha_t - \delta_y \mathrm{R}^\alpha_t\|_{\mathrm{TV}} \overset{t\to+\infty}{\longrightarrow} 0$ uniformly in (x,y,α), there exists $(\Delta_t)_{t\geq 0}$ independent of x,y,α,

and f (with $\|f\|_\infty \le 1$) such that

$$|\mu R_t^\alpha f - \nu R_t^\alpha f| \le \Delta_t \overset{t\to+\infty}{\longrightarrow} 0. \tag{3.4.21}$$

Using the cocycle property (3.4.14), we deduce that

$$\|\mu R_{t+s}^\alpha - \mu R_t^\alpha\|_{\mathrm{TV}} = \|(\mu R_{t+s,t}^\alpha) R_t^\alpha - \mu R_t^\alpha\|_{\mathrm{TV}} \le \Delta_t. \tag{3.4.22}$$

and hence, $(\mu R_t^\alpha)_{t\ge 0}$ is a Cauchy sequence for the $\|.\|_{\mathrm{TV}}$-norm. For every (α,μ), this thus converges to a probability $\mu_\alpha^\star$ which does not depend on μ thanks to (3.4.21). The bound in (3.4.22) being uniform in (α,μ), the convergence to this $\mu_\alpha^\star$ is uniform, that is:

$$\sup_{\mu,\alpha} \|\mu R_t^\alpha - \mu_\alpha^\star\|_{\mathrm{TV}} \le \Delta_t \xrightarrow{t\to+\infty} 0. \tag{3.4.23}$$

(ii) For a probability distribution ν and a measurable function $f : \mathscr{E} \to \mathbb{R}$, we have

$$(\Phi_t^\alpha(\mu) - \nu) f = \frac{1}{\mu g_t^\alpha}\left[\mu\left(g_t^\alpha\left(R_t^\alpha f - \nu(f)\right)\right)\right] = \mathbb{E}[R_t^\alpha f(Z) - \nu(f)]$$

where $\mathbb{P}_Z = \frac{\mu(.\times g_t^\alpha)}{\mu g_t^\alpha}$. Thus, using (3.4.23), we deduce that

$$\|\Phi_t^\alpha(\mu) - \mu_\alpha^\star\|_{\mathrm{TV}} \le \sup_{\nu\in\mathscr{P}(\mathscr{E})} \|\nu R_t^\alpha - \mu_\alpha^\star\|_{\mathrm{TV}} \le \Delta_t \xrightarrow{t\to+\infty} 0.$$

■

Now, we want to exhibit some sufficient conditions which ensure (3.4.17).

Proposition 3.4.5. Assume **($\mathbf{A_0}$)** and **($\mathbf{A_1}$)** and that there exist $\varepsilon > 0$, $c > 0$ and a probability distribution Ψ such that for all n and α, $\mathfrak{S}_{n+1,n}^\alpha \ge \varepsilon\Psi$, $\|\mathfrak{S}_{n+1,n}^\alpha\|_\infty \le c$ and

$$\sum_n \frac{\Psi(g_n^\alpha)}{\|g_n^\alpha\|_\infty} = +\infty. \tag{3.4.24}$$

Then, Condition (3.4.17) holds true.

Proof. *Step 1*: We show that under the assumptions,

$$R_{n+1,n}^\alpha \ge \varepsilon_n \Psi_n^\alpha,$$

where $(\varepsilon_n)_{n\geq 0}$ is a sequence of positive numbers with $\sum \varepsilon_n = +\infty$ and $\Psi_n^\alpha \in \mathscr{P}(\mathscr{E})$ (depending on n and α): for every non-negative function f, we have

$$\mathrm{R}_{n+1,n}^\alpha f = \frac{\mathfrak{S}_{n+1,n}^\alpha (f g_n^\alpha)}{g_{n+1}^\alpha}.$$

Since $g_{n+1}^\alpha = \mathfrak{S}_{n+1,n}^\alpha g_n^\alpha$, it follows that

$$\mathrm{R}_{n+1,n}^\alpha f \geq \frac{\varepsilon \Psi(f g_n^\alpha)}{c \| g_n^\alpha \|_\infty} = \varepsilon_n \Psi_n^\alpha,$$

with

$$\varepsilon_n = \frac{\varepsilon \Psi(g_n^\alpha)}{c \| g_n^\alpha \|_\infty} \quad \text{and} \quad \Psi_n^\alpha(f) = \frac{\Psi(f g_n^\alpha)}{\Psi(g_n^\alpha)}.$$

Assumption (3.4.24) finally ensures that $\sum \varepsilon_n = +\infty$.

Step 2: We prove the statement. By Step 1,

$$\delta_x \mathrm{R}_{n+1,n}^\alpha = \varepsilon_n \Psi_n^\alpha + (1-\varepsilon_n) \mathrm{Q}_{n+1}^\alpha(x, \cdot),$$

where Q_{n+1}^α is a Markov kernel. We deduce that

$$\mu \mathrm{R}_{n+1,n}^\alpha - \nu \mathrm{R}_{n+1,n}^\alpha = (1-\varepsilon_n)(\mu \mathrm{Q}_{n+1}^\alpha - \nu \mathrm{Q}_{n+1}^\alpha)$$

Since for a given function f, $\| \mathrm{Q}_{n+1}^\alpha f \|_\infty \leq \| f \|_\infty$, we get

$$\| \mu \mathrm{R}_{n+1,n}^\alpha - \nu \mathrm{R}_{n+1,n}^\alpha \|_{\mathrm{TV}} = (1-\varepsilon_n) \| \mu \mathrm{Q}_{n+1}^\alpha - \nu \mathrm{Q}_{n+1}^\alpha \|_{\mathrm{TV}} \leq (1-\varepsilon_n) \| \mu - \nu \|_{\mathrm{TV}}.$$

Thus, for all $k \in \{0, \ldots, n\}$, we have

$$\| (\mu \mathrm{R}_{n+1,k+1}^\alpha) \mathrm{R}_{k+1,k}^\alpha - (\nu \mathrm{R}_{n+1,k+1}^\alpha) \mathrm{R}_{k+1,k}^\alpha \|_{\mathrm{TV}} \leq (1-\varepsilon_k) \| \mu \mathrm{R}_{n+1,k+1}^\alpha - \nu \mathrm{R}_{n+1,k+1}^\alpha \|_{\mathrm{TV}},$$

and an induction leads to

$$\| \mu \mathrm{R}_n^\alpha - \nu \mathrm{R}_n^\alpha \|_{\mathrm{TV}} \leq \prod_k (1-\varepsilon_k) \| \mu - \nu \|_{\mathrm{TV}} \leq 2 \prod_k (1-\varepsilon_k) \overset{n \to +\infty}{\longrightarrow} 0,$$

since $\sum \varepsilon_n = +\infty$. Furthermore, the convergence is uniform in (ν, μ, α). This uniform conver-

gence extends to $t \in \mathbb{R}_+$ by noting that

$$\|\mu R_t^\alpha - \nu R_t^\alpha\|_{TV} = \|\mu R_{t,\lfloor t\rfloor}^\alpha) R_{\lfloor t\rfloor}^\alpha - (\nu R_{t,\lfloor t\rfloor}^\alpha) R_{\lfloor t\rfloor}^\alpha\|_{TV}.$$

■

We are now ready to state the main result of this section:

Proposition 3.4.6. Assume (**A$_0$**), (**A$_1$**) and (**A$_4$**). Then,

(i)
$$\sup_{\mu,\alpha} \|\Phi_t^\alpha(\mu) - \mu_\alpha^\star\|_{TV} \overset{t\to+\infty}{\longrightarrow} 0. \tag{3.4.25}$$

(ii) For any initial condition $\nu_0 \in \mathscr{P}(\mathscr{E})$, there exists a probability $\mathscr{I}_{\nu_0}$ such that the unique solution to $\dot{\nu} = -\nu + \Pi_\nu$ converges to $\mathscr{I}_{\nu_0}$ as $t \to +\infty$ for the TV-distance. Furthermore, the convergence is uniform in the initial condition:

$$\sup_{\nu_0 \in \mathscr{P}(\mathscr{E})} \|\nu_t^{\nu_0} - \mathscr{I}_{\nu_0}\|_{TV} \xrightarrow{t\to+\infty} 0.$$

Remark 3.4.7. Following carefully the proof, we in fact obtained the existence of some positive C and β such that

$$\sup_{\mu,\alpha} \|\Phi_t^\alpha(\mu) - \mu_\alpha^\star\|_{TV} \le Ce^{-\beta t}. \tag{3.4.26}$$

Proof. *(i)* By Proposition 3.4.3 *(ii)* and Proposition 3.4.5, it is enough to check that (3.4.24) holds true. We begin by giving a series expansion of the solution of

$$\dot{\varphi}_t = \varphi_t \mathscr{A}_{\alpha_t}. \tag{3.4.27}$$

Let $s_0 \in \mathbb{R}_+$:

$$\begin{aligned}\varphi_{s_0} &= \varphi_0 + \int_0^{s_0} \varphi_{s_1} \mathscr{A}_{\alpha_{s_1}} ds_1 = \varphi_0 + \int_0^{s_0} \Big(\varphi_0 + \int_0^{s_1} \varphi_{s_2} \mathscr{A}_{\alpha_{s_2}} ds_2\Big) \mathscr{A}_{\alpha_{s_1}} ds_1 \\ &= \varphi_0 + \int_0^{s_0} \varphi_0 \mathscr{A}_{\alpha_{s_1}} ds_1 + \int_0^{s_0}\int_0^{s_1} \varphi_{s_2} \mathscr{A}_{\alpha_{s_2}} \mathscr{A}_{\alpha_{s_1}} ds_2 ds_1.\end{aligned}$$

An induction then leads to:

$$\forall n \ge 1, \quad \varphi_{s_0} = \varphi_0 + \sum_{k=1}^n \int_0^{s_0} \cdots \int_0^{s_{k-1}} \varphi_0 \mathscr{A}_{\alpha_{s_k}} \cdots \mathscr{A}_{\alpha_{s_1}} ds_k \cdots ds_1 + r_n(s_0),$$

where

$$r_n(s_0) = \int_0^{s_0} \cdots \int_0^{s_n} \varphi_{s_{n+1}} \mathscr{A}_{\alpha_{s_{n+1}}} \cdots \mathscr{A}_{\alpha_{s_1}} \, ds_{n+1} \cdots ds_1.$$

Under the assumptions, $\sup_{\mu \in \mathscr{P}(\mathscr{E})} \|\mathscr{A}_\mu\|_\infty < +\infty$. This implies that

$$\|r_n(s_0)\|_\infty \leq \sup_{s \in [0,s_0]} \|\varphi_s\|_\infty \frac{(cs_0)^{n+1}}{(n+1)!} \leq \|\varphi_0\|_\infty e^{cs_0} \frac{(cs_0)^{n+1}}{(n+1)!},$$

where the second inequality follows from Gronwall's lemma. Thus, for any s_0, $r_n(s_0) \xrightarrow{n \to +\infty} 0$ and hence,

$$\mathrm{K}^\alpha_{s_0,0} = \mathrm{Id} + \sum_{k \geq 1} \int_0^{s_0} \cdots \int_0^{s_{k-1}} \mathscr{A}_{\alpha_{s_k}} \cdots \mathscr{A}_{\alpha_{s_1}} \, ds_k \cdots ds_1. \tag{3.4.28}$$

Note that when $\mu \mapsto \mathscr{A}_\mu$ is constant, one retrieves the classical exponential solution of the system.

We now use the very definition of $\mathscr{A}_\mu$ to get for any $k \geq 1$:

$$\mathscr{A}_{\alpha_{s_k}} \cdots \mathscr{A}_{\alpha_{s_1}} = \sum_{(n_1,\ldots,n_k) \in \mathbb{N}^{k+1}} \mathrm{K}^{n_k}_{\alpha_{s_k},\partial} \cdots \mathrm{K}^{n_1}_{\alpha_{s_1},\partial}.$$

Thus, using **($\mathbf{A_4}$)** in the second line, we get

$$\begin{aligned}
\Psi(g^\alpha_t) = \Psi(\mathrm{K}^\alpha_{t,0}\mathbf{1}) &= 1 + \sum_{k \geq 1} \sum_{(n_1,\ldots,n_k) \in \mathbb{N}^{k+1}} \int_0^{s_0} \cdots \int_0^{s_{k-1}} \Psi(\mathrm{K}^{n_k}_{\alpha_{s_k},\partial} \cdots \mathrm{K}^{n_1}_{\alpha_{s_1},\partial}\mathbf{1}) ds_{s_k} \cdots ds_1 \\
&\geq c \left[1 + \sum_{k \geq 1} \sum_{(n_1,\ldots,n_k) \in \mathbb{N}^{k+1}} \int_0^{s_0} \cdots \int_0^{s_{k-1}} \mathrm{K}^{n_k}_{\alpha_{s_k},\partial} \cdots \mathrm{K}^{n_1}_{\alpha_{s_1},\partial}\mathbf{1} \, ds_k \cdots ds_1 \right] \\
&\geq c g^\alpha_t.
\end{aligned}$$

In particular,

$$\inf_{n,\alpha} \frac{\Psi(g^\alpha_n \mathbf{1})}{\|g^\alpha_n\|_\infty} \geq c > 0,$$

which (trivially) implies (3.4.24).

(ii) To prove this statement, let us recall that by Lemma 3.4.1 (and the related notations), (μ_t) defined by $\mu_t = \nu^{\nu_0}_{\tau^{-1}(t)}$ satisfies (3.4.3) (where τ^{-1} denotes the inverse of the bijective function

τ). By (3.4.15), one deduces that

$$\mu_t = \Phi_t^{\mu_\bullet}(\nu_0).$$

We thus apply *(i)* with $\alpha = \mu_\bullet$. Note that by the uniqueness property given in lemma 3.4.1 *(i)*, $(\mu_t)_{t\geq 0}$ is uniquely determined by its initial condition ν_0. We can thus denote the probability $\mu_\alpha^\star$ of *(i)* by $\mathscr{I}_{\nu_0}$ and deduce that

$$\sup_{\nu_0 \in \mathscr{P}(\mathscr{E})} \|\mu_t^{\nu_0} - \mathscr{I}_{\nu_0}\|_{\mathrm{TV}} \xrightarrow{t\to+\infty} 0.$$

Finally, since the time change τ satisfies (3.4.11) where c is independent of ν_0, the above uniform convergence also applies to $(\nu_t^{\nu_0})_{t\geq 0}$. ■

3.5 Proof of Theorem 3.2.8

(i) We first check that $(\mathbf{A_0})$, $(\mathbf{A_1})$ and $(\mathbf{A_2})$ imply $(\mathbf{P_0})$, $(\mathbf{P_1})$ and $(\mathbf{P_2})$ when

$$\mathbb{K}_\mu = K_{\mu,\partial} + \delta_\mu(.)\mu.$$

The uniqueness of the invariant distribution Π_μ is a direct consequence of $(\mathbf{A_3})$. The fact that $\mu \mapsto \Pi_\mu$ is Lipschitz is established in Lemma 3.3.2. Thus, $(\mathbf{P_0})$ is true. Assumption $(\mathbf{P_1})$ is a direct consequence of $(\mathbf{A_3})$. Finally, by $(\mathbf{A_2})$, $\mu \mapsto K_{\mu,\partial}$ and $\mu \mapsto \delta_\mu$ are clearly Lipschitz for the TV-norm so that $\mu \mapsto \mathbb{K}_\mu$ is finally Lipschitz continuous. Finally, by (3.2.8), $\|\mu_{n+1} - \mu_n\|_{\mathrm{TV}} \leq 2\gamma_{n+1}$, so that (X_n, μ_n) satisfies the assumptions of proposition 3.3.6.

Then, the property (i) is a classical consequence of the conclusion of this proposition (proposition 3.3.6), of the tightness assumption on (μ_n), and of the theory of asymptotic pseudo-trajectories. For the sake of completeness, we recall insights and references. First, using again that $\|\mu_{n+1} - \mu_n\|_{\mathrm{TV}} \leq 2\gamma_{n+1}$, one can check that $(\tilde{\mu}_t)_{t\geq 0}$ is Lipschitz for the TV-distance (which metrizes the weak topology). Then, owing to the tightness of $(\mu_n)_{n\geq 1}$ and to the "shifted" construction of the sequence $(\tilde{\mu}^{(n)})_{n\geq 0}$, one deduces the tightness of $(\tilde{\mu}^{(n)})_{n\geq 0}$ on $\mathscr{C}(\mathbb{R}_+, \mathscr{P}(\mathscr{E}))$ from an Ascoli argument (see [21, Theorem 3.2] or [200, Proof of Proposition 3.5] for similar arguments).

Second, to show that every weak limit is a solution to $\dot{\nu} = F(\nu)$, we first deduce from the

construction that

$$\forall t \geq 0, \quad \tilde{\mu}_t^{(n)} = \tilde{\mu}_0^{(n)} + \int_0^t F(\tilde{\mu}_{N(n,s)}^{(n)}) ds + \int_0^t \varepsilon_{N(n,s)+1} ds.$$

Noting that for a given bounded function f,

$$\int_0^t \varepsilon_{N(n,s)+1}(f) ds = | \sum_{k=n+1}^{N(n,t)+1} \gamma_k \left(f(X_k) - \Pi_{\mu_k}(f) \right) + O(\gamma_n^{-1}),$$

we deduce from Proposition 3.3.6 that for any $T > 0$,

$$\lim_{n \to +\infty} \sup_{t \leq T} \left| \int_0^t \varepsilon_{N(n,s)+1}(f) ds \right| = 0 \quad a.s.$$

Combined with the fact that F is Lipschitz continuous (by lemma 3.3.2), this leads to the following property: for all bounded measurable $f : \mathscr{E} \to \mathbb{R}$, for all $T > 0$,

$$\lim_{n \to +\infty} \sup_{t \leq T} \left| \tilde{\mu}_t^{(n)}(f) - \nu_t^{\mu_n}) \right| = 0 \quad a.s.,$$

where ν_t^{μ} denotes the unique solution of (3.4.1) starting from μ (existence and uniqueness are given by lemma 3.4.1 *(ii)*). This property then implies the result (see [27, Theorem 3.5] for a very similar result).

(ii) From the first statement, we know that *a.s.*, any weak limit of $(\tilde{\mu}^{(n)})_n$ is a solution to $\dot{\nu} = -\nu + \Pi_\nu$. Let $(\tilde{\mu}^{(n_k)})_{k \geq 1}$ denote such a convergent subsequence of $(\tilde{\mu}^{(n)})_{n \geq 1}$ and let ν^∞ denote its limit. The process $(\nu_t^\infty)_{t \geq 0}$ is thus a solution to $\dot{\nu} = -\nu + \Pi_\nu$ starting from ν_0^∞. By Proposition 3.4.6, there exists $\mathscr{I}_{\nu_0^\infty}$ such that

$$\| \nu_t^\infty - \mathscr{I}_{\nu_0^\infty} \|_{TV} \xrightarrow{t \to +\infty} 0.$$

Now, for some given $n \geq 1$ and $\tau \geq 1$, recall that $N(n,\tau) := \max\{k \geq 1, \sum_{\ell=k+1}^n \gamma_\ell \leq \tau\}$. By construction, there exists a sequence $(\tau_n)_{n \geq 1}$ such that

$$\tau_n \leq \tau \leq \tau_n + \gamma_{n+1} \quad \text{and} \quad \mu_n = \tilde{\mu}_{\tau_n}^{(N(n,\tau))}.$$

At the price of extracting again a subsequence, $(\tilde{\mu}^{(N(n_k,\tau))})_{k \geq 1}$ also converges to a solution to $\dot{\nu} = -\nu + \Pi_\nu$. Denote it by $(\nu_t^{\infty,\tau})_{t \geq 0}$. Since $(\tau_n)_{n \geq 1}$ converges to τ and since the limiting

function is continuous, we have:

$$\nu_\tau^{\infty,\tau} = \lim_{n\to+\infty} \mu_n = \nu_0.$$

By uniqueness of solution, this implies that for $t \geq \tau$,

$$\nu_t^{\infty,\tau} = \nu_{t-\tau}^{\infty}.$$

As a consequence, $(\nu_t^{\infty,\tau})_{t\geq 0}$ has the same limit as $t \to +\infty$, that is $\mathscr{I}(\nu_0)$. But since the convergence to $\mathscr{I}(\nu_0)$ is uniform in the initial condition, for any ε, we can fix $\tau_\varepsilon > 0$ such that

$$\|\nu_t^{\infty,\tau_\varepsilon} - \mathscr{I}(\nu_0)\|_{TV} \leq \varepsilon \quad \forall t \geq \tau_\varepsilon.$$

Thus, for any $\varepsilon > 0$,

$$\|\nu_0 - \mathscr{I}(\nu_0)\|_{TV} = \|\nu_{\tau_\varepsilon}^{\infty,\tau_\varepsilon} - \mathscr{I}(\nu_0)\|_{TV} \leq \varepsilon.$$

Finally, $\nu_0 = \mathscr{I}(\nu_0)$ which implies that the process is stationary. As a consequence, $\dot{\nu}_0 = 0$, which in turn implies that $\nu_0 = \Pi_{\nu_0}$. Thus, ν_0 is a fixed point for the map $\mu \mapsto \Pi_\mu$, *i.e.* a QSD by lemma 3.2.7.

3.6 Tightness

We proce Theorem 3.2.11 which provides sufficient conditions for (μ_n) to be *a.s.* tight.

Proof of Theorem 3.2.11. The proof is divided into two steps.

Step 1. We prove that $\exists \varepsilon_0 > 0$ such that $\forall \varepsilon \in (0, \varepsilon_0]$, $\sup_n \{\mathbb{E}[\mu_n(V_\varepsilon)] + \mathbb{E}[V_\varepsilon(X_n)]\} < +\infty$.

By the recursive formula (3.2.8) and the definition of the dynamics of (X_n), we have:

$$\begin{cases} \mu_{n+1}(V_\varepsilon) = \mu_n(V_\varepsilon)(1-\gamma_{n+1}) + \gamma_{n+1} V_\varepsilon(X_{n+1}) \\ \mathbb{E}[V_\varepsilon(X_{n+1})|\mathscr{F}_n] = K_{\mu_n,\delta} V_\varepsilon(X_n) + \delta(X_n)\mu_n(V_\varepsilon). \end{cases} \tag{3.6.1}$$

Thus, setting $u_n = \mathbb{E}[\mu_n(V_\varepsilon)]$ and $v_n = \mathbb{E}[V_\varepsilon(X_n)]$ (depending on ε even if the notation does

not), we deduce from $(\mathbf{A_V})$ that:

$$\begin{cases} u_{n+1} \le u_n(1-\gamma_{n+1}) + \gamma_{n+1} v_{n+1} \\ v_{n+1} \le \varepsilon v_n + \beta_\varepsilon + \|\delta\|_\infty u_n. \end{cases} \tag{3.6.2}$$

Setting $L_n = u_n + \lambda \gamma_n v_n$ where $\lambda > 0$, we obtain:

$$L_{n+1} \le u_n \left(1 - (1-(1+\lambda)\|\delta\|_\infty)\gamma_{n+1}\right) + \lambda \gamma_{n+1} v_n \varepsilon \left(1+\lambda^{-1}\right) + \beta_\varepsilon (1+\lambda)\gamma_{n+1}.$$

Since $\|\delta\|_\infty < 1$, one can fix λ sufficiently small such that

$$\rho_\lambda = 1 - (1+\lambda)\|\delta\|_\infty > 0.$$

Then, if $\varepsilon_0 = \frac{1}{2}(1+\lambda^{-1})^{-1}$, we get for all $\varepsilon \in (0, \varepsilon_0]$ (using that (γ_n) is non-increasing):

$$\begin{aligned} L_{n+1} &\le u_n \left(1 - \rho_\lambda \gamma_{n+1}\right) + \frac{1}{2}\lambda \gamma_n v_n + C\gamma_{n+1} \\ &\le L_n \left(1 - \rho_\lambda \gamma_{n+1}\right) + C\gamma_{n+1} \quad \forall n \ge n_0, \end{aligned}$$

where C is a finite constant and,

$$n_0 := \inf\{n \in \mathbb{N}, 1 - \rho_\lambda \gamma_{n+1} \ge \frac{1}{2}\}.$$

The integer n_0 is certainly finite since (γ_n) tends to 0 as $n \to +\infty$. Then, an induction leads to

$$\begin{aligned} L_n &\le L_{n_0} \prod_{k=n_0+1}^{n} (1 - \rho_\lambda \gamma_k) + C \sum_{k=n_0+1}^{n} \gamma_k \prod_{\ell=k+1}^{n} (1 - \rho_\lambda \gamma_\ell) \\ &\le L_{n_0} + C \sum_{k \ge n_0+1} \gamma_k \exp(-\rho_\lambda \sum_{\ell=k+1}^{n} \gamma_k). \end{aligned} \tag{3.6.3}$$

In the second line, we used that $\rho_\lambda \gamma_n \le 1/2$ for $n \ge n_0 + 1$ and the inequality $\log(1+x) \le x$ for $x > -1$. By an integral-series comparison argument, one checks that (with $t_n = \sum_{k=1}^{n} \gamma_k$):

$$\sum_{k \ge n_0+1} \gamma_k \exp\left(-\rho_\lambda \sum_{\ell=k+1}^{n} \gamma_k\right) \le e^{-\rho_\lambda t_n} \int_0^{t_n} e^{\rho_\lambda s} ds \le \frac{1}{\rho_\lambda}. \tag{3.6.4}$$

Thus, checking by induction that L_{n_0} is finite, it follows that

$$\sup_{n\geq 1} L_n < +\infty. \tag{3.6.5}$$

Thus, $\sup_n u_n < +\infty$ and plugging this control into (3.6.2), we deduce that

$$v_{n+1} \leq \varepsilon v_n + C$$

and $\sup_n v_n < +\infty$.

Step 2. We prove the result.

By Jensen inequality and ($\mathbf{A_V}$), we have: for all $\varepsilon > 0$,

$$\forall \mu \in \mathscr{P}(\mathscr{E}), \quad K_{\mu,\partial}\sqrt{V_\varepsilon} \leq \bar{\varepsilon}\sqrt{V_\varepsilon} + \bar{\beta}_\varepsilon \quad \text{with } \bar{\varepsilon} = \sqrt{\varepsilon} \text{ and } \bar{\beta}_\varepsilon = \sqrt{\beta_\varepsilon}.$$

Now, by (3.6.1) applied with $\mathcal{V} = \sqrt{V_\varepsilon}$,

$$\mu_{n+1}(\mathcal{V}) \leq \mu_n(\mathcal{V})(1 - \gamma_{n+1}(1 - \|\delta\|_\infty)) + \bar{\varepsilon}\gamma_{n+1}\mathcal{V}(X_n) + \gamma_{n+1}\bar{\beta}_\varepsilon + \gamma_{n+1}\Delta M_{n+1}, \tag{3.6.6}$$

with

$$\Delta M_{n+1} = \mathcal{V}(X_{n+1}) - \mathbb{E}[\mathcal{V}(X_{n+1})|\mathscr{F}_n].$$

For the sake of readability, we first provide a simple proof case when the assumption holds with $\varepsilon = 0$ (*i.e.* when there is an inf-compact function V_0 such that $\sup_{\mu,x} K_{\mu,\partial}V_0(x) < +\infty$).

Case $\varepsilon = 0$. Set $\rho = 1 - \|\delta\|_\infty$ and $n_0 = \inf\{n, \rho\gamma_n \leq \frac{1}{2}\}$. An induction leads to

$$\mu_n(\mathcal{V}) \leq \mu_n(\mathcal{V}) \prod_{k=n_0+1}^{n} (1 - \rho\gamma_k) + C \sum_{k=n_0+1}^{n} \gamma_k \Theta_{n,k} + \sum_{k=n_0+1}^{n} \gamma_k \Delta M_k \Theta_{n,k},$$

with $\rho = 1 - \|\delta\|_\infty$ and,

$$\Theta_{n,k} = \prod_{\ell=k+1}^{n} (1 - \rho\gamma_k).$$

As in the first part of the proof, we have

$$\forall\, k \in [\![n_0, n]\!], \quad \Theta_{n,k} \leq e^{-\rho(t_n - t_k)},$$

so that

$$\sup_{n\geq n_0+1}\sum_{k=n_0+1}^{n}\gamma_k\Theta_{n,k}<+\infty.$$

Finally, to prove that

$$\sup_{n\geq n_0+1}\sum_{k=n_0+1}^{n}\gamma_k\Delta M_k\Theta_{n,k}<+\infty\quad a.s.,$$

we apply Lemma 3.6.1 with $b_n=e^{\rho t_n}$ and $u_n=\gamma_n\Delta M_n$ and thus, check that the martingale (N_n) defined by

$$N_n=\sum_{k=1}^{n}\gamma_k\Delta M_k,\quad n\geq 1,$$

is convergent. But,

$$\langle N\rangle_n\leq\sum_{k=1}^{n}\gamma_k^2V_0(X_k),$$

so that by Step 1,

$$\mathbb{E}[\langle N\rangle_\infty]\leq C\sum_{k\geq 1}\gamma_k^2<+\infty.$$

This concludes the proof.

Case $\varepsilon>0$. We now consider the real assumption $(\mathbf{A_V})$. We go back to (3.6.6) and choose to iterate the control of $\mathcal{V}(X_n)$ induced by (3.6.1):

$$\begin{aligned}\bar{\varepsilon}\mathcal{V}(X_n)&\leq\bar{\varepsilon}^2\mathcal{V}(X_{n-1})+\bar{\varepsilon}\Delta M_n+\bar{\varepsilon}\bar{\beta}_\varepsilon+\bar{\varepsilon}\mu_{n-1}(\mathcal{V})\\&\leq\bar{\varepsilon}^n\mathcal{V}(X_0)+\sum_{k=1}^{n-1}\bar{\varepsilon}^{n-k}\left(\Delta M_{k+1}+\bar{\beta}_\varepsilon+\mu_k(\mathcal{V})\right).\end{aligned}$$

Hence, plugging into (3.6.6) and setting

$$\nu_n(\mathcal{V})=\sum_{k=1}^{n-1}\bar{\varepsilon}^{n-1-k}\mu_k(\mathcal{V}),$$

we get:

$$\begin{cases}\mu_{n+1}(\mathcal{V})\leq\mu_n(\mathcal{V})(1-\gamma_{n+1}(1-\|\delta\|_\infty))+\gamma_{n+1}\left[\bar{\varepsilon}\nu_n(\mathcal{V})+C_\varepsilon+\sum_{k=1}^{n}\bar{\varepsilon}^{n-k}\Delta M_{k+1}\right]\\ \nu_{n+1}(\mathcal{V})\leq\bar{\varepsilon}\nu_n(\mathcal{V})+\mu_n(\mathcal{V}),\end{cases}$$

with $C_\varepsilon=\mathcal{V}(X_0)+(1-\bar{\varepsilon})^{-1}\bar{\beta}_\varepsilon$. Setting $\tilde{L}_n=\mu_n(\mathcal{V})+\lambda\gamma_n\nu_n(\mathcal{V})$ (where λ will be calibrated

further), one checks that

$$\tilde{L}_{n+1} \le (1-\gamma_{n+1}(1-\|\delta\|_\infty-\lambda))\mu_n(\mathcal{V}) + \lambda\gamma_{n+1}\left(\frac{\bar{\varepsilon}}{\lambda}+\bar{\varepsilon}\right)\nu_n(\mathcal{V}) + \gamma_{n+1}\left(C_\varepsilon + \sum_{k=1}^{n}\bar{\varepsilon}^{n-k}\Delta M_{k+1}\right).$$

Fix λ small enough in such a way that $\tilde{\rho}_\lambda = 1-\|\delta\|_\infty-\lambda > 0$. Then choosing $\bar{\varepsilon}$ such that $\frac{\bar{\varepsilon}}{\lambda}+\bar{\varepsilon} < 1$, *i.e.* such that ($\bar{\varepsilon} = \sqrt{\varepsilon}$),

$$\varepsilon \le \tilde{\varepsilon}_0 = (1+\lambda^{-1})^{-2}, \tag{3.6.7}$$

leads to the existence of an integer $\tilde{n}_0$ such that for all $n \ge \tilde{n}_0$,

$$\tilde{L}_{n+1} \le (1-\tilde{\rho}_\lambda\gamma_{n+1})\tilde{L}_n + \gamma_{n+1}\left(C_\varepsilon + \sum_{k=1}^{n}\bar{\varepsilon}^{n-k}\Delta M_{k+1}\right).$$

At the price of replacing $\tilde{n}_0$ by a greater integer, we assume that $\tilde{\rho}_\lambda\gamma_n \le 1/2$ for $n \ge \tilde{n}_0$. Then, a similar induction as in (3.6.3) leads to:

$$\tilde{L}_n \le \tilde{L}_{\tilde{n}_0}\exp(-\rho_\lambda(t_n-t_{\tilde{n}_0})) + \sum_{k=\tilde{n}_0+1}^{n}\gamma_k\left(C_\varepsilon + \sum_{\ell=1}^{k}\bar{\varepsilon}^{k-\ell}\Delta M_{\ell+1}\right)\exp(-\rho_\lambda(t_n-t_k)),$$

where $t_n = \sum_{k=1}^{n}\gamma_k$. On the one hand, a similar argument as the one of (3.6.4) leads to

$$\sup_{n\ge\tilde{n}_0+1}\sum_{k=\tilde{n}_0+1}^{n}\gamma_k C_\varepsilon\exp(-\rho_\lambda(t_n-t_k)) < +\infty \quad a.s.$$

One the other, we want to show that,

$$\sup_{n\ge\tilde{n}_0+1} e^{-\rho_\lambda t_n}\sum_{k=\tilde{n}_0+1}^{n}\gamma_k e^{\rho_\lambda t_k}\sum_{\ell=1}^{k}\bar{\varepsilon}^{k-\ell}\Delta M_{\ell+1} < +\infty \quad a.s. \tag{3.6.8}$$

By Lemma 3.6.1 applied with $b_n = e^{\rho t_n}$ and $u_n = \gamma_n\sum_{\ell=1}^{n}\bar{\varepsilon}^{n-\ell}\Delta M_{\ell+1}$, it is enough to show that,

$$\sum\left(\gamma_k\sum_{\ell=1}^{k}\bar{\varepsilon}^{k-\ell}\Delta M_{\ell+1}\right) \quad \text{is } a.s. \text{ a convergent series.}$$

Note that we can assume without loss of generality that $\tilde{n}_0 = 0$. Then, let us remark that

$$\sum_{k=1}^{n} \gamma_k \sum_{\ell=1}^{k} \bar{\varepsilon}^{k-\ell} \Delta M_{\ell+1} = \sum_{\ell=1}^{n} \Delta M_{\ell+1} \Big(\sum_{k=\ell}^{n} \gamma_k \bar{\varepsilon}^{k-\ell} \Big),$$
$$= \underbrace{\sum_{\ell=1}^{n} \left(\sum_{k=\ell}^{+\infty} \gamma_k \bar{\varepsilon}^{k-\ell} \right) \Delta M_{\ell+1}}_{\tilde{N}_n} - \left(\sum_{k=n+1}^{+\infty} \gamma_k \bar{\varepsilon}^{k} \right) \sum_{\ell=1}^{n} \bar{\varepsilon}^{-\ell} \Delta M_{\ell+1}. \tag{3.6.9}$$

Note that in the above equation, we provided a decomposition which allows to replace a triangular array of martingale increments by the difference of a martingale and a simpler triangular array than the previous one. For the martingale term $(\tilde{N}_n)$, recall by Step 1 that if $\varepsilon \le \varepsilon_0$, $\sup_{n\ge 1} \mathbb{E}[V_\varepsilon(X_n)] < +\infty$ so that for every $\varepsilon \in (0, \varepsilon_0]$,

$$\sup_{n\ge 1} \mathbb{E}[(\Delta M_n)^2] \le \sup_{n\ge 1} \mathbb{E}[V_\varepsilon(X_n)] + \infty.$$

Then, since (γ_n) is non-increasing, we obtain

$$\mathbb{E}[\langle \tilde{N} \rangle_\infty] \le C \sum_{\ell=2}^{+\infty} \left(\sum_{k=\ell}^{+\infty} \gamma_k \bar{\varepsilon}^{k-\ell} \right)^2 \le \sum_{\ell \ge 2} \frac{\gamma_\ell^2}{(1-\bar{\varepsilon})^2} < +\infty,$$

which involves that $(\tilde{N}_n)$ *a.s.* converges. Let us finally consider the second term of (3.6.9). Setting $b_n = \left(\sum_{k=n}^{+\infty} \gamma_k \varepsilon^k \right)^{-1}$ (which clearly goes to ∞) and $u_n = \frac{1}{b_n} \varepsilon^{-n} \Delta M_{n+1}$, we remark that

$$\sum_{\ell=2}^{n} u_\ell = \tilde{N}_n,$$

so that the convergence of the series follows from what preceeds. Once again, by Lemma 3.6.1, we derive that

$$\left(\sum_{k=n+1}^{+\infty} \gamma_k \varepsilon^k \right) \sum_{\ell=1}^{n} \varepsilon^{-\ell} \Delta M_{\ell+1} \xrightarrow{n\to+\infty} 0, \quad a.s.$$

As a conclusion, if $\varepsilon \le \varepsilon_0 \wedge \tilde{\varepsilon}_0$ (where ε_0 and $\tilde{\varepsilon}_0$ are respectively defined in Step 1 and in (3.6.7)), the conclusion of the Proposition holds true. ■

For the sake of completeness, we recall below the classical Kronecker Lemma used several times in the above proof.

Lemma 3.6.1 (Kronecker Lemma). Let (u_n) and (b_n) be two real sequences. Assume $\sum u_n$ is convergent and that (b_n) is positive with $\lim_{n\to+\infty} b_n = +\infty$. Then, $b_n^{-1} \sum_{k=1}^{n} b_k u_k \xrightarrow{n\to+\infty} 0$.

3.7 Euler schemes of McKean-Vlasov SDEs

This section is related to Euler schemes of McKean-Vlasov SDEs. We begin by a concrete criterion which ensures **(A_V)** for these dynamics (possibly with jumps) and then focus on the proofs of Proposition 3.2.13 and Proposition 3.2.15.

3.7.1 About **(A_V)**

Lemma 3.7.1. Assume that $\mathscr{E}$ is a closed subset of $\mathbb{R}^d$. Let $(\zeta_t)_{t\geq 0}$ be a centered Lévy process on $\mathbb{R}^d$ and h denote the step of the Euler scheme. Assume that $(x,\mu)\mapsto\sigma(x,\mu)$ is bounded (on $\mathscr{E}\times\mathscr{P}(\mathscr{E})$). Then, **($A_V$)** holds in the two following situations.

(i) (Standard Euler scheme) The dynamics is given by (3.2.2), the step h is small enough, ζ_1 has moments of any order and some positive α and C exist such that

$$\langle x, b(x,\mu)\rangle \leq C-\alpha|x|^2 \quad \text{and} \quad |b(x,\mu)|^2\leq C(1+|x|^2)$$

(ii) (Modified Euler scheme) The dynamics is given by (3.2.15), there exists $\alpha>0$ such that $\mathbb{E}[|\zeta_h|^\alpha]<+\infty$ and Condition **(H_t)** holds.

Remark 3.7.2. ▷ In this discrete-time setting, we prefer to work with closed subsets which are "friendlier" for inf-compactness (for instance, $x\mapsto x$ is not inf-compact on $D=(0,+\infty)$ but is on $\bar{D}=(0,+\infty)$).

▷ The boundedness condition on σ may be alleviated.

▷ The condition in (i) is eventually satisfied when $(\zeta_t)_{t\geq 0}$ is a Brownian motion (and more generally when the large jumps of the Lévy process have moments of any order). Let us also recall that for a Lévy process, ζ_1 has moments of any order is equivalent to: $\forall\, t\geq 0$, ζ_t has moments of any order.

Proof. *(i)* We set $V(x)=|x|^2$ and assume that $Y_0=x$. By the second order Taylor formula and the assumptions on the coefficients, we have

$$V(Y_1)1_{\tau>1}\leq V(x)(1-2\alpha h)+2\langle x,\zeta_h\rangle + c\left(h^2V(x)+|\zeta_h|^2\right),$$

where c is independent of h. By the elementary inequality $|\langle u,v\rangle|\leq\frac{1}{2}(\rho|u|^2+\rho^{-1}|v|^2)$ for

$u, v \in \mathbb{R}^d$ and $\rho > 0$,

$$2\langle x, \zeta_h \rangle \le \frac{\alpha h}{2} V(x) + \frac{2}{\alpha h} |\zeta_h|^2.$$

Hence for $h \le h_0 := \frac{\alpha}{2c}$, a finite C_h exists such that

$$V(Y_1) 1_{\tau > 1} \le V(x)(1 - \alpha h) + C_h (1 + |\zeta_h|^2).$$

Let $\varepsilon > 0$ and set

$$p_\varepsilon = \inf \left\{ p \ge 1, (1 - \alpha h)^p \le \frac{\varepsilon}{2} \right\}.$$

By Lemma 3.7.3 applied with $\delta = 1$, it follows that

$$V^{p_\varepsilon}(Y_1) 1_{\tau > 1} \le \varepsilon V^{p_\varepsilon}(x) + C_{h,p} (1 + |\zeta_h|^{2p_\varepsilon}).$$

The result follows.
(ii) Set $V(x) = |x|^\alpha$. We have

$$\mathbb{E}_x[V(Y_1) 1_{\tau > 1}] \le |t(x + hb(x, \mu))|^\alpha + \mathbb{E}[|\xi_h|^\alpha 1_{\tau > 1}] \le C < +\infty,$$

Thus,

$$\sup_{x \in \mathscr{E}} K_{\mu, \partial} V(x) \le C,$$

which in turn trivially implies $(\mathbf{A_V})$. ■

Lemma 3.7.3. For every $p \ge 1$, for every $\delta > 0$, there exists a finite constant $C_{p,\delta}$ such that for any $u, v \ge 0$,

$$(u + v)^p \le (1 + \delta) u^p + C_{p,\delta} v^p.$$

Proof. By the first order Taylor formula applied to $x \mapsto x^p$ on $\mathbb{R}_+$,

$$(u + v)^p = u^p + p(u + \theta v)^{p-1} v \le u^p + p 2^{p-1} u^{p-1} v + p 2^{p-1} v^p.$$

Then, by the Young inequality applied with $\bar{p} = p$ and $\bar{q} = \frac{p}{p-1}$

$$v u^{p-1} = (\rho^{1-p} v)(\rho u)^{p-1} \le \frac{\rho^{p - p^2}}{p} v^p + \frac{(p-1)\rho^p}{p} u^p, \quad \rho > 0.$$

Taking ρ such that $\frac{(p-1)\rho^p}{p} = \delta$, the result follows.

■

3.7.2 Proof of Proposition 3.2.13(i)

Proof. We apply Theorem 3.2.8. Since $\bar{\mathrm{D}}$ is compact, $(\mu_n^h)_{n\geq 1}$ is certainly almost surely tight and we thus only have to check Assumptions $(\mathbf{A_i})$, $i=0,\dots,4$. In order to alleviate the notations, we write $\mathrm{K}_{\mu,\partial}$ for $\mathrm{K}^{(h)}_{\mu,\partial}$ and g instead of g_h. We have

$$\begin{aligned}\mathrm{K}_{\mu,\partial}f(x) &= \mathbb{E}[f(x+hb(x,\mu)+\sigma(x,\mu)\zeta_h)1_{\{x+hb(x,\mu)+\sigma(x,\mu)\zeta_h\in\bar{\mathrm{D}}\}}]\\ &= \int f(u)\mathbf{1}_{\{u\in\bar{\mathrm{D}}\}}g\left(\varphi_{x,\mu}(u)\right)|\det(\sigma(x,\mu))|^{-1}\lambda_d(du), \end{aligned} \tag{3.7.1}$$

where $\varphi_{x,\mu}(u)=\sigma(x,\mu)^{-1}(u-x-hb(x,\mu))$. Set

$$\mathrm{K}=\overline{\bigcup_{(x,\mu)\in\bar{\mathrm{D}}\times\mathscr{P}(\bar{\mathrm{D}})}\varphi_{x,\mu}(\bar{\mathrm{D}})}. \tag{3.7.2}$$

By $(\mathbf{H_{MV}})$, one can check that the function $(x,\mu)\mapsto\sup_{u\in\bar{\mathrm{D}}}|\varphi_{x,\mu}(u)|$ is bounded so that the set K is a compact subset of $\mathbb{R}^d$. It follows that

$$c_1\lambda_d(\mathrm{D})\Psi(.)\leq\mathrm{K}_{\mu,\partial}(x,.)\leq c_2\lambda_d(\mathrm{D})\Psi(.), \tag{3.7.3}$$

with

$$c_1=\inf_{z\in\mathrm{K}}g(z)\inf_{x,\mu}|\det(\sigma(x,\mu))|^{-1},\quad c_2=\sup_{z\in\mathrm{K}}g(z)\sup_{x,\mu}|\det(\sigma(x,\mu))|^{-1},$$

and Ψ denotes the uniform distribution on D. Under the assumptions, g is continuous and positive on $\mathbb{R}^d$, and $(x,\mu)\mapsto|\det(\sigma(x,\mu))|^{-1}$ is bounded (uniform ellipticity uniformly in the measure argument) so that c_1 and c_2 are positive finite constants. Thus, (3.7.3) corresponds to the assumption of Lemma 3.2.10, which thus involves $(\mathbf{A_4})$. The assumptions $(\mathbf{A_0})$ and $(\mathbf{A_3})$ are also clearly true under (3.7.3). For $(\mathbf{A_1})$, the Feller property is obvious. Since $(\mathbf{A_2})$ implies the second part of $(\mathbf{A_1})$, we now only need to prove $(\mathbf{A_2})$:

Set $\mu_t=(1-t)\mu_0+t\mu_1$ and assume that ζ_h has density g. For every bounded measurable $f:\mathscr{E}\to\mathbb{R}$,

$$\begin{aligned}\mathrm{K}_{\mu_1,\partial}f(x)-\mathrm{K}_{\mu_0,\partial}f(x) &= \int_0^1\partial_t(\mathrm{K}_{\mu_t,\partial}f(x))\,dt\\ &= \int_0^1\int f(u)1_{\{u\in\mathrm{D}\}}\partial_t\left(g(\sigma_t^{-1}(u-m_t))|\det(\sigma_t)|^{-1}\right)\lambda_d(du)dt \end{aligned} \tag{3.7.4}$$

where $\sigma_t=\sigma(x,\mu_t)$ and $m_t=x+hb(x,\mu_t)$. By Lemma 3.7.4 and the fact that K defined in

(3.7.2) is compact (so that g and ∇g are bounded on K), we deduce that a finite constant C exists such that for all x and f,

$$|\mathrm{K}_{\mu_1,\partial} f(x) - \mathrm{K}_{\mu_0,\partial} f(x)| \le \mathrm{C}\|f\|_\infty \|\mu_1 - \mu_0\|_{\mathrm{TV}},$$

which in turn implies $(\mathbf{A_2})$. By Theorem 3.2.8, we can thus conclude that every weak limit $\mu_h^\star$ is a QSD for $(\mathrm{K}^{(h)}_{\mu,\partial})_{\mu\in\mathscr{P}(\bar{\mathrm{D}})}$.

By the very definition of a QSD, for every bounded measurable function $f : \bar{\mathrm{D}} \to \mathbb{R}$,

$$\int \mathrm{K}_{\mu,\partial} f(x)\mu_h^\star(dx) = \rho_h \langle \mu_h^\star, f\rangle. \tag{3.7.5}$$

with $\rho_h = 1 - \mu_h^\star(\delta_{\mu_h^\star}) \in (0,1)$ (the extinction rate). Thus, for $f = \mathbf{1}_{\partial \mathrm{D}}$, this yields:

$$\mu_h^\star(\partial \mathrm{D}) = \rho_h^{-1} \int \mathrm{K}_{\mu,\partial}(x, \partial \mathrm{D})\mu_h^\star(dx) = 0,$$

since $\mathrm{K}_{\mu,\partial}(x,\partial \mathrm{D}) = 0$ for all $x \in \bar{\mathrm{D}}$ (this point is obvious since $\mathrm{K}_{\mu,\partial}(x,.)$ has a density with respect to Lebesgue measure under the assumptions). ■

Lemma 3.7.4. Assume $(\mathbf{H_{MV}})$. Assume that g is $\mathscr{C}^1$. Then,

$$|\partial_t \left(g(\sigma_t^{-1}(u-m_t))|\det(\sigma_t)|^{-1}\right)| \le \mathrm{C}\|\mu_1 - \mu_0\|_{\mathrm{TV}} \left(|\nabla g(\sigma_t^{-1}(u-m_t)| + g(\sigma_t^{-1}(u-m_t))|\right).$$

Proof.

$$\begin{aligned}\partial_t \left(g(\sigma_t^{-1}(u-m_t))\right) &= \nabla g(\sigma_t^{-1}(u-m_t)).\partial_t(\sigma_t^{-1}(u-m_t)) \\ &= -\nabla g(\sigma_t^{-1}(u-m_t)) \left[\sigma_t^{-2}(\partial_t\sigma_t)(u-m_t) + \sigma_t^{-1}(\partial_t m_t)\right].\end{aligned}$$

Now, by definition of a flat derivative,

$$\begin{aligned}\partial_t\sigma_t = \partial_t\sigma(x,\mu_t) &= \int \frac{\delta\sigma(x,\cdot)}{\delta m}(\mu_t,y)(\mu_1-\mu_0)(dy) \quad \text{and,} \\ h^{-1}\partial_t m_t = \partial_t b(x,\mu_t) &= \int \frac{\delta b(x,\cdot)}{\delta m}(\mu_t,y)(\mu_1-\mu_0)(dy).\end{aligned} \tag{3.7.6}$$

By Assumption (**H$_{\mathbf{MV}}$**), it follows that a positive constant C (independent of h) exists such that:

$$\|\partial_t \sigma_t\|_\infty + h^{-1}\|\partial_t m_t\|_\infty \le C\|\mu_1 - \mu_0\|_{TV}. \tag{3.7.7}$$

Under the uniform ellipticity assumption, we get:

$$||\det(\sigma_t)|^{-1}\partial_t\left(g(\sigma_t^{-1}(u-m_t))\right)| \le C|\nabla g(\sigma_t^{-1}(u-m_t)|\|\mu_1-\mu_0\|_{TV}.$$

On the other hand,

$$\partial_t(\det(\sigma_t)) = \mathrm{Tr}({}^t\mathrm{com}(\sigma_t)(\partial_t\sigma_t)).$$

Thus, using (3.7.7), the continuity of $A \mapsto {}^t\mathrm{com}(A)$ and, the boundedness and the uniform ellipticity of σ (under (**H$_{\mathbf{MV}}$**)), we obtain:

$$\|\partial_t(\det(\sigma_t)^{-1})\|_\infty \le C\|\mu_1-\mu_0\|_{TV}.$$

The lemma follows. ■

3.7.3 Proof of Proposition 3.2.15

First, we prove that $(\mu_n)_{n\ge1}$ is tight. By Lemma 3.7.1(ii), Assumption (**A$_{\mathbf{V}}$**) is satisfied. We now check that $\|\delta_\mu\|_\infty < 1$. As in the proof of Proposition 3.2.13, we write $\tilde{K}_{\mu,\partial}$ and g instead of $\tilde{K}^{(h)}_{\mu,\partial}$ and g_h. Similarly as in (3.7.1), we get

$$\tilde{K}_{\mu,\partial}(x,du) = \mathbf{1}_{\{u\in D\}} g\left(\tilde{\varphi}_{x,\mu}(u)\right)|\det(\sigma(x,\mu)|^{-1})\lambda_d(du), \tag{3.7.8}$$

with $\tilde{\varphi}_{x,\mu}(u) = \sigma(x,\mu)^{-1}\left[u - t(x+hb(x,\mu))\right]$. Let $B(a,\rho)\subset D$. The fact that t is bounded combined with the uniform ellipticity assumption implies that the function $(x,\mu)\mapsto \sup_{u\in B(a,\rho)}|\tilde{\varphi}_{x,\mu}(u)|$ is bounded so that the set $\tilde{K}$ defined as the adherence of $\bigcup_{(x,\mu)\in D\times\mathscr{P}(D)}\varphi_{x,\mu}(B(a,\rho))$ is compact. Then, setting $\underline{c} = \inf_{u\in\tilde{K}} g(u)$,

$$\tilde{K}_{\mu,\partial}\mathbf{1}(x) \ge \underline{c}\lambda_d(B(a,r)) \implies \|\delta_\mu\|_\infty < 1.$$

We then deduce tightness from Theorem 3.2.11 (and the fact that $\sum\gamma_n^2 < +\infty$).

Now, one checks the other assumptions. For the same reasons as in the proof of Proposi-

tion 3.2.13(i) (see Section 3.7.2), we only check $(\mathbf{A_2})$ and $(\mathbf{A_4})$. For $(\mathbf{A_2})$, the proof follows the lines of the one of Proposition 3.2.13(i) replacing m_t by $\tilde{m}_t = \mathfrak{t}(x + hb(x, \mu_t))$. Since $\mathfrak{t}$ has a bounded Jacobian, (3.7.6) and $(\mathbf{H_{MV}})$ imply that $\|\partial_t \tilde{m}_t\|_\infty < +\infty$ and one can then check that the conclusions of Lemma 3.7.4 are still true (replacing m_t by $\tilde{m}_t$). As a consequence, since g and ∇g are bounded, one thus deduces that

$$\left| \int_0^1 \int_D f(u) \partial_t \left(g(\sigma_t^{-1}(u - \tilde{m}_t)) |\det(\sigma_t)|^{-1} \right) \lambda_d(du)\, dt \right| \le C \|f\|_\infty \|\mu_1 - \mu_0\|_{TV},$$

which in turn implies $(\mathbf{A_2})$ (by (3.7.4)).

Now, let us consider $(\mathbf{A_4})$. Using that σ is uniformly elliptic (uniformly in μ) and applying (3.2.16) with $m = \mathfrak{t}(x + hb(x, \mu))$, $\sigma = \sigma(x, \mu)$ (and $M = \sup_{x,\mu} |\mathfrak{t}(x + hb(x, \mu))|$), we deduce that there exist some positive c_1 and c_2, and a density $\mathfrak{p}_h$ such that

$$c_1 \mathfrak{p}_h(u) \le g\left(\tilde{\varphi}_{x,\mu}(u)\right) \left(\det(\sigma(x, \mu)^{-1})\right) \le c_2 \mathfrak{p}_h(u),$$

where $\mathfrak{p}_h$ is a density on D. Plugging this inequality into (3.7.8), this implies $(\mathbf{A_4})$ and allows to apply Theorem 3.2.8 to conclude. The fact that $\mu_h^\star(\partial D) = 0$ follows from the same arguments as for Proposition 3.2.13(i).

3.7.4 Proof of Proposition 3.2.13(ii)

This proof is an adaptation of [28, Theorem 3.9] for McKean-Vlasov dynamics.

Let $(W_t)_{t\ge 0}$ denote a Brownian motion. For a given step $h > 0$, we denote by $(\xi_t^{h,\mu})_{t\ge 0}$, the continuous-time Euler scheme defined by $\xi_0^{h,\mu} = y \in D$,

$$\forall n \in \mathbb{N}, \quad \xi_{(n+1)h}^{h,\mu} := \xi_{nh}^{h,\mu} + b(\xi_{nh}^{h,\mu}, \mu)h + \sigma(\xi_{nh}^{h,\mu}, \mu)(W_{(n+1)h} - W_{nh}),$$

and,

$$\forall t \in [nh, (n+1)h), \quad \xi_t^{h,\mu} = \xi_{nh}^{h,\mu} + b(\xi_{nh}^{h,\mu}, \mu)(t - nh) + \sigma(\xi_{nh}^{h,\mu}, \mu)(W_t - W_{nh}).$$

Note that the continuous Euler scheme is an Itô process verifying

$$\xi_t^{h,\mu} = \xi_0^{h,\mu} + \int_0^t b(\xi_{\underline{s}_h}^{h,\mu}, \mu)ds + \int_0^t \sigma(\xi_{\underline{s}_h}^{h,\mu}, \mu)dW_s, \quad \text{with} \quad \underline{s}_h = h\lfloor \frac{s}{h} \rfloor. \tag{3.7.9}$$

With the notations of (3.2.13), the kernel of the killed process related to the Markov chain $(\xi_{nh}^{h,\mu})_{n\geq 0}$ is still denoted by $K_{\mu,\partial}^{(h)}$ but keeping in mind that $\zeta = W$ is now a Brownian motion.

Before going further, we first deduce from (3.7.5) and from an induction that for all $t \geq 0$, for every bounded measurable function $f : \mathbb{R}^d \to \mathbb{R}$,

$$\mathbb{E}_{\mu_h^\star}\left[f(\xi_{\underline{t}_h}^{h,\mu_h^\star}) 1_{\{\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > t\}} \right] = e^{-\lambda_h \underline{t}_h} \mu_h^\star(f), \tag{3.7.10}$$

with $\lambda_h := \frac{\log \rho_h}{h}$ and for a process $(Z_t)_{t\geq 0}$,

$$\bar{\tau}_D^h(Z) = \inf\{nh, Z_{nh} \in D^c\}. \tag{3.7.11}$$

In other words, $\bar{\tau}_D^h$ denotes the exit time related to the discrete-time dynamics (this notation will be still used in the sequel). Note that in order to write (3.7.10), we used that $\{\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > t\} = \{\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > \underline{t}_h\}$.

To prove the result, we now need to make h go to 0 in (3.7.10).

Step 1 (Tightness of the sequence of QSDs): We show that $(\mu_h^\star)_{h\leq h_0}$ is tight on D. For $\delta > 0$, set

$$B_\delta := \Big\{ x \in D, \quad d(x, \partial D) \leq \delta \Big\}. \tag{3.7.12}$$

It is enough to prove that for every $\varepsilon > 0$, there exists $\delta_\varepsilon > 0$ such that for every $h \in (0, h_0]$, $\mu_h^\star(B_{\delta_\varepsilon}) \leq \varepsilon$.

From (3.7.10) applied with $f = \mathbf{1}_{B_\delta}$, we get in particular

$$\mu_h^\star(B_\delta) \leq \mathbb{P}_{\mu_h^\star}(\xi_{\underline{t}_h}^{h,\mu_h^\star} \in B_\delta) \leq \lambda(B_\delta) \sup_{(x,y)\in \bar{D}^2} p_{\underline{t}_h}^{h,\mu_h^\star}(x,y),$$

where $p_t^{h,\mu}(x,.)$ denotes the density of the Euler scheme at time t (which exists under the uniform ellipticity assumption). To bound the density uniformly in h and $\mu \in \mathscr{P}(D)$, we use [181, Theorem 2.1] (which provides Aronson estimates for the density). One checks that the

related assumptions (**UE**) and (**SB**) are satisfied with some parameters L_0 and λ_0 independent of μ under our assumption ($\mathbf{H_{MV}}$). Thus, by this theorem, we deduce that for all $T_0 < T_1$, a constant C exists such that for all $h \in (0, h_0)$, for all $t \in [T_0, T_1]$,

$$\sup_{(x,y)\in\bar{D}^2} p_t^{h,\mu_h^\star}(x,y) \le C. \tag{3.7.13}$$

The tightness follows.

From now, we can thus consider a weak limit $\tilde{\mu}$ of a convergent subsequence $(\mu_{h_n}^\star)_{n\ge1}$ with $h_n \to 0$.

Step 2 (About the left-hand side of (3.7.10)): we show that for all $t \ge 0$ and for all bounded Lipschitz continuous function $f : D \mapsto \mathbb{R}$,

$$\mathbb{E}_{\mu_{h_n}^\star}\left[f(\xi_{\underline{t}_{h_n}}^{h_n,\mu_{h_n}^\star})1_{\bar{\tau}_D^{h_n}(\xi^{h_n,\mu_{h_n}^\star})>\underline{t}_{h_n}}\right] \xrightarrow{n\to+\infty} \mathbb{E}_{\mu^\star}\left[f(\xi_t^{\mu^\star})1_{\tau_D(\xi^{\mu^\star})>t}\right], \tag{3.7.14}$$

where $\xi^{\mu^\star}$ satisfies:

$$\xi_t^{\mu^\star} = \xi_0^{\mu^\star} + \int_0^t b(\xi_s^{\mu^\star}, \mu^\star)ds + \int_0^t \sigma(\xi_s^{\mu^\star}, \mu^\star)dW_s.$$

In order to simplify the notations, we write h instead of h_n in the sequel of this step and assume that $\tilde{\mu}_h^\star \to \mu^\star$ when $h \to 0$. This does not change the proof. We first decompose the error related to (3.7.14) into three parts:

$$\mathbb{E}_{\mu_h^\star}\left[f(\xi_t^{h,\mu_h^\star})1_{\bar{\tau}_D^h(\xi^{h,\mu_h^\star})>t}\right] - \mathbb{E}_{\mu^\star}\left[f(\xi_t^{\mu^\star})1_{\tau_D(\xi^{\mu^\star})>t}\right] = \mu_h^\star(\bar{\mathscr{P}}_t^{h,\mu_h^\star,\bar{\tau}}f - \bar{\mathscr{P}}_t^{h,\mu^\star,\bar{\tau}}f) \tag{3.7.15}$$

$$+ \mu_h^\star(\bar{\mathscr{P}}_t^{h,\mu^\star,\bar{\tau}}f) - \mu_h^\star(P_t^{\mu^\star,\partial}f) \tag{3.7.16}$$

$$+ \mu_h^\star(P_t^{\mu^\star,\partial}f) - \mu^\star(P_t^{\mu^\star,\partial}f), \tag{3.7.17}$$

where,

$$\bar{\mathscr{P}}_t^{h,\mu,\bar{\tau}}f(x) = \mathbb{E}_x\left[f(\xi_t^{h,\mu})1_{\bar{\tau}_D^h(\xi^{h,\mu})>t}\right] \quad \text{and,} \quad P_t^{\mu,\partial}f(x) = \mathbb{E}_x[f(\xi_t^\mu)\mathbf{1}_{\tau_D(\xi^\mu)>t}].$$

To control the right-hand side of (3.7.15), we first use that f is bounded Lipschitz continuous

to obtain that for all $T > 0$, for all $t \in [0,T]$:

$$|\bar{\mathscr{P}}_t^{h,\mu_h^\star,\bar{\tau}} f(x) - \bar{\mathscr{P}}_t^{h,\mu^\star,\bar{\tau}} f(x)| \le [f]_1 \mathbb{E}_x[|\xi_t^{h,\mu_h^\star} - \xi_t^{h,\mu^\star}|] \\ + \|f\|_\infty \mathbb{P}(\{\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > t\} \Delta \{\bar{\tau}_D^h(\xi^{h,\mu^\star}) > t\}) \tag{3.7.18}$$

$$\le C_T \Big(\mathscr{W}_1(\mu_h^\star, \mu^\star) + \mathbb{P}(\{\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > t\} \Delta \{\bar{\tau}_D^h(\xi^{h,\mu^\star}) > t\}) \Big), \tag{3.7.19}$$

where in the second line, we used Lemma 3.7.5 with $x = y$, $\mu_0 = \mu_h^\star$ and $\mu_1 = \mu^\star$. Since in this bounded setting, the weak convergence implies the $\mathscr{W}_1$-convergence, we only focus on the second term. Let us first consider the first part of the symmetric difference:

$$\mathbb{P}(\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > t, \bar{\tau}_D^h(\xi^{h,\mu^\star}) < t). \tag{3.7.20}$$

To manage this term, we apply Lemma 3.7.6 with $X = \xi^{h,\mu_h^\star}$ and $Y = \xi^{h,\mu^\star}$. By Lemma 3.7.5, (iii) is satisfied with $c_3 = C_T \delta^{-1} \mathscr{W}_1(\mu_h^\star, \mu^\star)$. For (ii), we use [46, Theorem 2.4] which states that for a given μ,

$$\mathbb{E}_x[\bar{\tau}_D^h(\xi^{h,\mu})] \le d_{(2.4)}[\tilde{\psi}_D(x) + \sqrt{h}], \tag{3.7.21}$$

where $\tilde{\psi}_D(x)$ is a smooth extension of the algebraic distance to the boundary. More precisely, by [200, Lemma 2.4], it can be shown that $P_t := \tilde{\psi}_D(\xi_t^{h,\mu})$ satisfies the assumption (**P**) of [46] (see also Lemma A.3 of [46] for details). By Lemma A.1 and A.4 of [46] , one can check that the other assumptions of [46, Theorem 2.4] hold true under our assumptions and that the constant denoted by $d_{(2.4)}$ is uniform in $\mu \in \mathscr{P}(D)$, $t \le T$ and $h \le h_0$. We thus deduce from (3.7.21) and Markov inequality that Assumption (ii) of Lemma 3.7.6 holds true with $c_2 = C(\delta + \sqrt{h})$ where C is a constant independent of the parameters.

Finally, let us consider Assumption (i) of Lemma 3.7.6 with $\mu = \mu_h^\star$. Since $(\mu_h^\star)$ is tight, one can find for all $\varepsilon > 0$, a $\delta_\varepsilon > 0$ such that B_{δ_ε} defined by (3.7.12) satisfies: for all $h \in (0, h_0]$, $\mu_h^\star(B_{\delta_\varepsilon})$. Thus,

$$\mathbb{P}_{\mu_h^\star}(\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) \in [t - t_0, t]) \le \varepsilon + \sup_{x, d(x,\partial D) \ge \delta_\varepsilon} \mathbb{P}_x((\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) \le 2t_0) \\ + \sup_{x, d(x,\partial D) \ge \delta_\varepsilon, t \ge 2t_0} \mathbb{P}_x(\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) \in [(t - t_0), t]). \tag{3.7.22}$$

We control the two right-hand terms. First, for all $x \in D$ such that $d(x, \partial D) \geq \delta_\varepsilon$,

$$\begin{aligned}\mathbb{P}_x((\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) \leq 2t_0) &\leq \mathbb{P}_x(\sup_{t\leq 2t_0} |\bar{\tau}_D^h(\xi_t^{h,\mu_h^\star} - x| \geq \delta_\varepsilon) \leq \delta_\varepsilon^{-2}\mathbb{E}_x[\sup_{t\leq 2t_0} |\bar{\tau}_D^h(\xi_t^{h,\mu_h^\star} - x|^2] \\ &\leq 2\delta_\varepsilon^{-2}\left(\mathbb{E}[\int_0^{2t_0} |b(\xi_{\underline{t}_h}^{h,\mu_h^\star}, \mu_h^\star)dt|^2] + \mathbb{E}[\left|\sup_{t\leq 2t_0}\int_0^t \sigma(\xi_{\underline{t}_h}^{h,\mu_h^\star}, \mu_h^\star)dW_s\right|^2]\right) \\ &\leq C\delta_\varepsilon^{-2}(t_0^2 + t_0) \leq C\delta_\varepsilon^{-2}t_0.\end{aligned}$$

where the constant C is universal since the coefficients b and σ are uniformly bounded in x and μ (since they are uniformly Lispschitz on a compact set). For (3.7.22), we decompose into two parts and deduce from the strong Markov property that

$$\mathbb{P}_x(\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) \in [(t-t_0), t]) \leq \mathbb{P}(\xi_{t-t_0}^{h,\mu_h^\star} \in B_{\delta_\varepsilon}) + \sup_{x, d(x,\partial D)\geq\delta_\varepsilon} \mathbb{P}_x((\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) \leq 2t_0).$$

The second term can be managed as previously whereas the first one satisfies the inequality:

$$\mathbb{P}(\xi_{t-t_0}^{h,\mu_h^\star} \in B_{\delta_\varepsilon}) \leq C\lambda(B_{\delta_\varepsilon}),$$

owing to the inequality (3.7.13). This means that we can set

$$c_3 = \varepsilon + C\lambda(B_{\delta_\varepsilon}) + \delta_\varepsilon^{-2}t_0,$$

where C only depends on T. Applying Lemma 3.7.6, we get for all $T > 0$, there exists a constant C such that for all $t \in (0, T]$, for all positive ε and δ,

$$\begin{aligned}&\mathbb{P}(\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > t, \bar{\tau}_D^h(\xi^{h,\mu^\star}) < t) \\ &\qquad \leq C\left[\delta^{-1}\mathcal{W}_1(\mu_h^\star, \mu^\star) + \delta + \sqrt{h} + \varepsilon + \lambda(B_{\delta_\varepsilon}) + \delta_\varepsilon^{-2}t_0\right],\end{aligned}$$

where δ_ε is a positive number which can be taken as small as we want (for a given ε). Taking δ small enough and then small enough h, ε, δ_ε and finally t_0, we thus obtain the expected control of (3.7.20). With the same arguments, one controls the other part of the symmetric difference of (3.7.19) and this ends the study of (3.7.15).

Now, for (3.7.16), this corresponds to the convergence of the killed semi-group of the Euler scheme towards the killed semi-group of the related diffusion with coefficients $b(., \mu^\star)$ and $\sigma(., \mu^\star)$. This can be obtained by a similar strategy as before that is left to the reader.

Finally, for (3.7.17), we need to prove that the map $x \mapsto P_t^{\mu^\star,\partial} f(x)$ is continuous on D. Denote by ξ^x the process solution starting from x to the diffusion, $d\xi_t = b(\xi_t, \mu^\star)dt + \sigma(\xi_t, \mu^\star)dW_t$. Under our assumptions, $\sup_{(x,y)\in D^2} \mathbb{E}[\sup_{t\le T} |\xi_t^x - \xi_t^y|^2] \le C_T |x-y|^2$ so that, keeping in mind that f is Lipschitz continuous, it is thus enough to control $\mathbb{P}(\{\tau_D(\xi^x) > t\} \Delta \{\tau_D(\xi^y) < t\})$. This can be again obtained with an adaptation of Lemma 3.7.6 where for a given $x_0 \in D$, one sets $\delta = \frac{1}{2} d(x_0, \partial D)$. This concludes the proof of this step.

Step 3. (Bounds for the extinction rate) We prove that some positive $\lambda_{\min}$, $\lambda_{\max}$ and h_0 exist such that for any $h \in (0, h_0)$, we have

$$\lambda_{\min} \le \lambda_h \le \lambda_{\max}.$$

For the lower-bound, it is enough to check that for a given $t_0 > 0$

$$\sup_{(\mu,x,h)\in \mathscr{P}(D)\times D\times(0,h_0]} \mathbb{P}_x(\bar{\tau}_D(\xi^{h,\mu}) > \lfloor t_0 \rfloor_h) < 1.$$

For this point, we can for instance again use [181, Theorem 2.1] which provides lower-bounds on the density: taking a given compact set K_0 in D^c, this result involves that under our uniform boundedness, Lipschitz and ellipticity assumptions, for $t_0 \ge 2h_0$, there exists a positive constant C such that for every $(\mu, x, h) \in \mathscr{P}(D) \times D \times (0, h_0]$,

$$\mathbb{P}(\bar{\tau}_D(\xi^{h,\mu}) \le \lfloor t_0 \rfloor_h) \ge \mathbb{P}(\xi^{h,\mu}_{\lfloor t_0 \rfloor_h} \in K_0) \ge C\lambda_d(K_0) > 0,$$

as soon as K_0 non-empty interior. Second, one remark that $(\lambda_h)_{h\le h_0}$ is upper-bounded as soon as there exists $t_0 > 0$ such that

$$\inf_{h>0} \mathbb{P}_{\mu_h^\star}\left(\{\bar{\tau}_D^h(\xi^{h,\mu_h^\star}) > t_0\right) > 0.$$

Denoting by K_1, a compact subset of D such that $\inf_{h\in(0,h_0]} \mu_h^\star(K_1) > 0$ (such a K_1 owing to the tightness property), one can check that the above property is true if

$$\sup_{(\mu,x,h)\in \mathscr{P}(D)\times K_1\times(0,h_0]} \mathbb{P}_x(\tau_D(\xi^{h,\mu}) \le t_0) < 1.$$

But this property is easy to ensure. Actually, with standard Gronwall arguments, it can be

shown that $\mathbb{E}_x[\sup_{0\le t\le t_0}|\xi_t^{h,\mu}-x|^2]\le Ct$ (where C is independent of the parameters under the assumptions of the result) and, setting $\delta = d(K_1,\partial D)$, it follows that

$$\mathbb{P}_x(\tau_D(\xi^{h,\mu})\le t_0)\le \mathbb{P}(\sup_{t\le t_0}|\xi_t^{h,\mu}-x|\ge\delta)\le C\delta^{-2}t_0.$$

For t_0 small enough, $C\delta^{-2}t_0<1$ and the result follows.

Step 4. (Conclusion) Recall that $(\mu^\star_{h_n})_{n\ge1}$ denote a convergent subsequence of $(\mu^\star_h)_{h\le h_0}$ to $\mu^\star$. At the price of an additional extraction, we deduce from Step 3 that $\lambda_{h_n}\to\lambda^\star\in(0,+\infty)$. Then, from Step 2 and (3.7.10), we get for all $t\ge0$ and for all Lipschitz continuous function $f:D\to\mathbb{R}$,

$$\mathbb{E}_{\mu^\star}\left[f(\xi_t^{\mu^\star})\mathbf{1}_{\{\tau_D(\xi^{\mu^\star})>t\}}\right]=e^{-\lambda^\star t}\mu^\star(f). \tag{3.7.23}$$

This concludes the proof.

Lemma 3.7.5. For μ_0 and $\mu_1\in\mathscr{P}(\bar D)$, let $(\xi_t^{h,\mu_0})_{t\ge0}$ and $(\xi_t^{h,\mu_1})_{t\ge0}$ be given by (3.7.9) starting from initial conditions x and y respectively. Then, under the assumptions of Proposition 3.2.13(ii), a universal constant C exists such that for all $T\ge0$,

$$\mathbb{E}\left[\sup_{t\le T}|\xi_t^{h,\mu_1}-\xi_t^{h,\mu_0}|^2\right]\le e^{CT}\left(|x-y|^2+C\mathscr{W}_1(\mu_0,\mu_1)^2\right).$$

By Itô formula, we have

$$\frac12|\xi_t^{h,\mu_1}-\xi_t^{h,\mu_0}|^2=\frac12|\xi_{\underline{t}_h}^{h,\mu_1}-\xi_{\underline{t}_h}^{h,\mu_0}|^2+\int_{\underline{t}_h}^t F(\xi_s^{h,\mu_1},\xi_s^{h,\mu_0},\xi_{\underline{s}_h}^{h,\mu_1},\xi_{\underline{s}_h}^{h,\mu_0})ds+M_t, \tag{3.7.24}$$

where,

$$\begin{aligned}F(x,y,\underline{x},\underline{y})&=\langle x-y,b(\underline{x},\mu_1)-b(\underline{y},\mu_0)\rangle\\&\quad+\frac12\mathrm{Tr}((\sigma(\underline{x},\mu_1)-\sigma(\underline{y},\mu_0))(\sigma(\underline{x},\mu_1)-\sigma(\underline{y},\mu_0))^*),\end{aligned}$$

and,

$$M_t=\int_0^t\langle\xi_s^{h,\mu_1}-\xi_s^{h,\mu_0},(\sigma(\xi_{\underline{s}_h}^{h,\mu_1},\mu_1)-\sigma(\xi_{\underline{s}_h}^{h,\mu_0},\mu_0))dW_s\rangle.$$

On the one hand, the definition of the flat derivative yields:

$$\begin{aligned}\langle x-y, b(\underline{x},\mu_1)-b(\underline{y},\mu_0)\rangle &= \langle x-y, b(\underline{x},\mu_1)-b(\underline{y},\mu_1)\rangle \\ &\quad + \int_0^1 \int \langle x-y, \frac{\delta b_{\underline{y}}}{\delta m}(\mu_t,z)\rangle(\mu_1-\mu_0)(dz)dt,\end{aligned} \tag{3.7.25}$$

with $b_{\underline{y}}$ being the application $\mu \mapsto b(\underline{y},\mu)$ and $\mu_t=(1-t)\mu_0+t\mu_1$. Setting $f(z)=\langle x-y, \frac{\delta b_{\underline{y}}}{\delta m}(\mu_t,z)\rangle$,

$$\begin{aligned}\int \langle x-y, \frac{\delta b_{\underline{y}}}{\delta m}(\mu_t,z)\rangle(\mu_1-\mu_0)(dz) &= \mu_1(f)-\mu_0(f) \\ &\le [f]_1 \mathcal{W}_1(\mu_0,\mu_1) \le |x-y|\left(\sup_{(z,\mu)\in\mathcal{E}\times\mathcal{P}(\mathcal{E})}\left[\frac{\delta b(z,\cdot)}{\delta m}(\mu,\cdot)\right]_1\right)\mathcal{W}_1(\mu_0,\mu_1).\end{aligned}$$

Thus, using the Lipschitz assumption on the flat derivative (and the fact that b is uniformly Lipschitz in the first variable for the first right-hand term of (3.7.25)), we get the existence of a constant C (independent of the parameters) such that

$$\langle x-y, b(\underline{x},\mu_1)-b(\underline{y},\mu_0)\rangle \le C|x-y|\Big(|\underline{x}-\underline{y}|+\mathcal{W}_1(\mu_0,\mu_1)\Big).$$

This leads to

$$\langle x-y, b(\underline{x},\mu_1)-b(\underline{y},\mu_0)\rangle \le C(|x-y|^2+|\underline{x}-\underline{y}|^2+\mathcal{W}_1(\mu_0,\mu_1)^2),$$

by using the elementary inequality $|uv|\le\frac{1}{2}(|u|^2+|v|^2)$ (and by keeping the same notation for the constant C even it may have changed). Similar arguments can be used for the diffusion part to obtain:

$$\mathrm{Tr}((\sigma(\underline{x},\mu_1)-\sigma(\underline{y},\mu_0))(\sigma(\underline{x},\mu_1)-\sigma(\underline{y},\mu_0))^*) \le C\Big(|\underline{x}-\underline{y}|^2+\mathcal{W}_1(\mu_0,\mu_1)^2\Big).$$

Hence, plugging these controls into (3.7.24) and using that $(M_t)_{t\ge0}$ is a true martingale leads to:

$$u_t \le u_{\underline{t}_h}(1+C(t-\underline{t}_h))+C(t-\underline{t}_h)\mathcal{W}_1(\mu_0,\mu_1)^2+C\int_{\underline{t}_h}^t u_s ds,$$

with $u_t=\frac{1}{2}\mathbb{E}[|\xi_t^{h,\mu_1}-\xi_t^{h,\mu_0}|_2]$. The Gronwall lemma then leads to: for all $n\ge0$, for all $t\in[nh,(n+1)h]$,

$$u_t \le \left(u_{nh}(1+Ch)+Ch\mathcal{W}_1(\mu_0,\mu_1)^2\right)e^{Ch} \le u_{nh}e^{Ch}+Ch\mathcal{W}_1(\mu_0,\mu_1)^2.$$

Once again, we do not change the notation for the constant C (even it may change from line to line). By an induction, we obtain for h small enough:

$$\mathbb{E}\left[|\xi_t^{h,\mu_1} - \xi_t^{h,\mu_0}|^2\right] \leq e^{Ct}\left(u_0 + C\mathcal{W}_1(\mu_0,\mu_1)^2\right).$$

We thus finally need to deduce that the supremum can be taken inside the expectation. This is a classical procedure again based on the Gronwall applied to $v_t = \mathbb{E}\left[\sup_{0\leq s\leq t}|\xi_s^{h,\mu_1} - \xi_s^{h,\mu_0}|^2\right]$ using the previous controls on (u_t) and the nice properties on the supremum of a martingale (see *e.g.* the proof of Theorem 7.2. in [197] for a similar argument in a slightly different setting).

Lemma 3.7.6. Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ denote two Markov processes with respect to the same filtration $(\mathscr{F}_t)_{t\geq 0}$ with values in D. Let μ be a probability on D. Let t_0, t be positive numbers such that $t_0 < t$. Denote by τ_X and τ_Y the exit times of D and assume that.

(i) $\mathbb{P}_\mu(\tau_X \in [t-t_0, t]) \leq c_1$.

(ii) $\sup_{x, d(x,\partial D)\leq\delta} \mathbb{P}_x(\tau_Y > t_0) \leq c_2$.

(iii) $\mathbb{P}_\mu(\sup_{s\leq t}|X_s - Y_s| \geq \delta) \leq c_3$.

Then,

$$\mathbb{P}_\mu(\tau_Y > t, \tau_X < t) \leq c_1 + c_2 + c_3.$$

Proof. We have:

$$\begin{aligned}
\mathbb{P}_\mu(\tau_Y > t, \tau_X < t) &\leq \mathbb{P}_\mu(\tau_X \in [t-t_0, t]) + \mathbb{P}_\mu(\tau_Y > t, \tau_X < t - t_0) \\
&\leq c_1 + \mathbb{P}_\mu(|X_{\tau_X} - Y_{\tau_X}| \geq \delta) + \mathbb{P}(\tau_Y > t, d(Y_{\tau_X}, \partial D) \leq \delta, \tau_X \leq t - t_0) \\
&\leq c_1 + c_3 + \mathbb{P}(\tau_Y > t, d(Y_{\tau_X}, \partial D) \leq \delta, \tau_X \leq t - t_0) \\
&\leq c_1 + c_3 + \mathbb{E}[\mathbb{P}(\tau_Y > t - \tau_X | \mathscr{F}_{\tau_X}) \mathbf{1}_{d(Y_{\tau_X}, \partial D)\leq\delta, \tau_X \leq t-t_0)}],
\end{aligned}$$

where in the last line, we used the strong Markov property. The result follows. ∎

3.8 Postponed proofs

Proof of lemma 3.2.5. We prove that there exists $\ell \in \mathbb{N}^*$ such that

$$\rho(\ell) := \sup_{(x,\mu)\in\mathscr{E}\times\mathscr{P}(\mathscr{E})} \mathrm{K}^{\ell}_{\mu,\partial}\mathbf{1}(x) < 1. \tag{3.8.1}$$

Since $(x,\mu) \mapsto \mathrm{K}^{\ell}_{\mu,\partial}\mathbf{1}(x)$ is continuous on the compact space $\mathscr{E}\times\mathscr{P}(\mathscr{E})$, it is enough to check that there exists ℓ such that for every $(x,\mu) \in \mathscr{E}\times\mathscr{P}(\mathscr{E})$, $\mathrm{K}^{\ell}_{\mu,\partial}\mathbf{1}(x) < 1$. If it is not true, one can find a sequence (x_n,μ_n) such that $\mathrm{K}^{n}_{\mu_n,\partial}\mathbf{1}(x_n) = 1$ for every n. Since $n \mapsto \mathrm{K}^{n}_{\mu,\partial}\mathbf{1}(x)$ is non-increasing, this implies that for every (n,k) such that $k \le n$, $\mathrm{K}^{k}_{\mu_n,\partial}\mathbf{1}(x_n) = 1$. Thus, for any limit point $(x^\star,\mu^\star)$ of $(x_n,\mu_n)_{n\ge1}$, $\mathrm{K}^{k}_{\mu^\star,\partial}\mathbf{1}(x^\star) = 1$ for every $k \in \mathbb{N}$, which is a contradiction with the assumption of the lemma. ■

Proof of lemma 3.3.5. (i) For every $n \in \mathbb{N}$, one sets

$$\kappa_n(\mu,\nu) = \sup_{\alpha\in\mathscr{P}(\mathscr{E})} \|\alpha\mathbb{K}^n_\mu - \alpha\mathbb{K}^n_\nu\|_{\mathrm{TV}}.$$

We have $\kappa_0(\mu,\nu) = 0$ and by $(\mathbf{P_2})$,

$$\kappa_1(\mu,\nu) \le c_2\|\mu-\nu\|_{\mathrm{TV}}.$$

Then, for every $n \ge 0$,

$$\begin{aligned}
\|\alpha\mathbb{K}^{n+1}_\mu - \alpha\mathbb{K}^{n+1}_\nu\|_{\mathrm{TV}} &= \|\alpha\mathbb{K}_\mu\mathbb{K}^n_\mu - \alpha\mathbb{K}_\nu\mathbb{K}^n_\nu\|_{\mathrm{TV}} \\
&\le \|\alpha\mathbb{K}_\mu\left[\mathbb{K}^n_\mu - \mathbb{K}^n_\nu\right]\|_{\mathrm{TV}} + \|\alpha\mathbb{K}_\mu\mathbb{K}^n_\nu - \alpha\mathbb{K}_\nu\mathbb{K}^n_\nu\|_{\mathrm{TV}} \\
&\le \kappa_n(\mu,\nu) + \|\alpha\mathbb{K}_\mu - \alpha\mathbb{K}_\nu\|_{\mathrm{TV}} \\
&\le \kappa_n(\mu,\nu) + \kappa_1(\mu,\nu) \\
&\le \kappa_n(\mu,\nu) + c_2\|\mu-\nu\|_{\mathrm{TV}}
\end{aligned}$$

An induction leads to

$$\kappa_n(\mu,\nu) \le c_2 n\|\mu-\nu\|_{\mathrm{TV}}.$$

(ii) First, since

$$R_\mu - R_\nu = \frac{1}{1-\varepsilon}(\mathbb{K}_\mu^\ell - \mathbb{K}_\nu^\ell), \tag{3.8.2}$$

the case $n = 1$ follows from (i). Now, for $n \in \mathbb{N}$,

$$\begin{aligned}
\sup_{\alpha\in\mathscr{P}(\mathscr{E})} \|\alpha R_\mu^{n+1} - \alpha R_\nu^{n+1}\|_{TV} &\le \sup_{\alpha\in\mathscr{P}(\mathscr{E})} \|(\alpha R_\mu^n)R_\mu - (\alpha R_\mu^n)R_\nu\|_{TV} \\
&\quad + \sup_{\alpha\in\mathscr{P}(\mathscr{E})} \|(\alpha R_\mu^n)R_\nu - (\alpha R_\nu^n)R_\nu\|_{TV} \\
&\le \sup_{\alpha\in\mathscr{P}(\mathscr{E})} \|\alpha R_\mu - \alpha R_\nu\|_{TV} + \sup_{\alpha\in\mathscr{P}(\mathscr{E})} \|\alpha R_\mu^n - \alpha R_\nu^n\|_{TV} \\
&\le \frac{c_2\ell}{1-\varepsilon}\|\mu-\nu\|_{TV} + \sup_{\alpha\in\mathscr{P}(\mathscr{E})} \|\alpha R_\mu^n - \alpha R_\nu^n\|_{TV}.
\end{aligned}$$

The result follows by induction. ■

Conclusion & Perspectives

La plus grande probabilité de voir nos vœux accomplis laisse encore un doute ; c'est pourquoi, quand nos espérances se réalisent, nous sommes toujours surpris.

JOHANN WOLFGANG VON GOETHE

Résumé de la conclusion

Dans ce dernier chapitre, nous proposons quelques perspectives autour du travail effectué durant la thèse.

Perspectives du chapitre 1

Dans le premier chapitre, on rappelle que les résultats sont obtenus pour les équations de Langevin de champ moyen lorsque l'interaction est de type polynomial. Il serait intéressant de les étendre à une classe plus large d'équations de Langevin de champ moyen telles que les équations générales présentées dans le chapitre 2 voire à d'autres équations différentielles stochastiques de McKean-Vlasov dont le terme de drift ne dérive pas d'un hamiltonien.

▷ Dans le cas des équations de Langevin de champ moyen, il apparaît qu'un point clé est que l'on conserve une expression explicite de la mesure invariante, ce qui permet d'établir des hypothèses assez générales sous lesquelles les inégalités de log-Sobolev et les conditions sur l'application mesure invariante [9] pourraient être conservées. En revanche, il resterait une difficulté importante dans la preuve du résultat principal liée à l'établissement de résultats de grandes déviations à l'instar de ceux établis par Liu and Wu.

▷ Quant à d'autres EDS de McKean-Vlasov dont le terme de drift ne dérive pas d'un hamiltonien, en plus de la difficulté précédente, il faudra pallier la non explicitation des mesures invariantes, c'est-à-dire qu'on ne peut malheureusement pas caractériser les équilibres par des problèmes de points fixes explicites.

Perspectives du chapitre 2

Dans le deuxième chapitre, on rappelle que les résultats de contraction exponentielle sont établis par couplage synchrone avec la métrique de Wasserstein pour des EDS de McKean-Vlasov avec des coefficients de diffusion constants et des drifts généraux. Il serait intéressant de les obtenir par d'autres arguments de couplage [10] et/ou avec d'autres métriques [11]. De plus, à l'instar de [41, Théorème 1.5], nous pouvons montrer une inégalité de concentration. D'autre part, les résultats de propagation du chaos uniforme en temps étant obtenus via une borne de Fournier and Guillin et cette dernière étant en $\mathbf{O}(\frac{1}{\sqrt{n}})$, il serait intéressant de revoir les preuves pour améliorer la borne en $\mathbf{O}(\frac{1}{n})$ à l'instar de [41, Théorème 1.2].

9. Voir (**H4**) et (**H5**) dans le chapitre 1.
10. Par exemple, les couplages par réflexion.
11. Par exemple, la métrique en variation totale.

Perspectives du chapitre 3

Dans un article en cours [12], on s'intéresse à l'approche continue (mesures d'occupation) de l'approximation de QSDs de McKean-Vlasov, continuité naturelle du chapitre 3.

On définit le *processus de McKean-Vlasov conditionné* à rester dans D par $(X_t)_{0\le t\le\tau}$ un processus continu défini jusqu'au temps d'arrêt $\tau := \inf\{t>0, \quad X_t \in \partial D\}$ et tel que

$$\begin{cases} X_t \in D \quad \forall t<\tau; \\ X_\tau = \lim_{t\uparrow\tau^-} X_t; \\ X_0 \sim \mu_0 \in \mathscr{P}(D); \\ dX_t = \sigma(X_t, \mathscr{L}(X_t|\tau>t))dB_t + b(X_t, \mathscr{L}(X_t|\tau>t))dt \quad \forall t\in[0,\tau]. \end{cases} \tag{3.8.3}$$

Ce type de processus a déjà été étudié dans l'article [222] de Tough and Nolen dans lequel les auteurs étudient la dynamique de Fleming-Viot associée à ce processus et établissent des résultats sur les limites de Yaglom. Le flot $(\mathscr{L}(X_t|\tau>t))_{t\ge0}$ de lois conditionnelles définit un semi-groupe $(G_t)_{t\ge0}$ appelé *semi-groupe conditionné de McKean-Vlasov*. De façon analogue au cas markovien, on définit les QSDs et les QLDs (Quasi-limiting Distributions) de McKean-Vlasov ci-dessous.

$$\pi \in \mathbb{QSD} \Longleftrightarrow \forall t\ge0, \quad G_t(\pi)=\pi. \tag{3.8.4}$$

$$\pi \in \mathbb{QLD} \Longleftrightarrow \exists \mu\in\mathscr{P}(D), \quad \lim_{t\to+\infty} G_t(\mu)=\pi. \tag{3.8.5}$$

Soient $\mu\in\mathscr{P}(D)$, $J_t^\mu := \mathbb{P}_\mu(\tau>t)$ et $\mathscr{L}^\mu$ le *générateur du processus d'interaction figée* μ muni du domaine $\mathscr{C}_0^\infty(\overline{D})$. De plus, on a

$$\frac{d}{dt}G_t(\mu) = G_t(\mu)\mathscr{L}^{G_t(\mu)} - G_t(\mu)\frac{d}{dt}\ln J_t^\mu. \tag{3.8.6}$$

$$\pi\mathscr{L}^\pi = -\lambda(\pi)\pi. \tag{3.8.7}$$

$(\mu\mapsto J_t^\mu G_t(\mu))_{t\ge0}$ définit un semi-groupe $(P_t^\partial)_{t\ge0}$ appelé *semi-groupe de McKean-Vlasov tué* auquel est associé l'opérateur $A := \int_0^{+\infty} P_t^\partial dt$ appelé *opérateur de Green-Poisson-McKean-Vlasov*. Soit A_μ l'*opérateur de Green-Poisson* associé au *processus absorbé d'interaction figée* μ. Dans l'article en cours, on a établi les caractérisations ci-dessous des QSDs de McKean-Vlasov.

12. Nous proposons ci-dessous quelques résultats de l'article (en cours de finalisation)...

Théorème (Caractérisations des QSDs de McKean-Vlasov)

Il y a équivalence entre

— $\pi \in \mathbb{QSD}$.

— $\pi \in \mathbb{QLD}$.

— $\exists \lambda(\pi) > 0$ telle que $\pi \mathscr{L}^{\pi} = -\lambda(\pi)\pi$.

— $\exists \lambda(\pi) > 0$ telle que $\forall t \geq 0$, $\pi \mathrm{P}_t^{\partial} = e^{-\lambda(\pi)t}\pi$.

— $\int_0^{+\infty} \mathrm{J}_t^{\pi} dt < +\infty$ et $\pi \mathrm{A} = \mathbb{E}_{\pi}[\tau]\pi$.

— $\exists \lambda(\pi) > 0$ telle que $\pi \mathrm{A}_{\pi} = \frac{1}{\lambda(\pi)}\pi$.

— π est un point fixe de $\Pi : \mu \mapsto \frac{\mu \mathrm{A}_{\mu}}{\langle \mu \mathrm{A}_{\mu}, \mathbb{I}_{\mathrm{D}} \rangle}$ application mesure invariante (AMI) associée au processus d'interaction et renaissance identiques (PIRI).

Soit $(\mathbf{Y}_t)_{t \geq 0}$ le processus avec interaction et distribution mesure d'occupation (associé à la dynamique auto-attractive) :

— On considère le processus auto-attractif absorbé $(\mathbf{A}_t)_{t \geq 0}$ tel que

$$\forall t < \tau_{\mathbf{A}}, \quad d\mathbf{A}_t = b(\mathbf{A}_t, \mu_t^{\mathbf{A}})dt + \sigma(\mathbf{A}_t, \mu_t^{\mathbf{A}})d\mathrm{B}_t; \tag{3.8.8}$$

$$\mu_t^{\mathbf{A}} := \frac{1}{t}\int_0^t \delta_{\mathrm{A}_s} ds; \tag{3.8.9}$$

$$\tau_{\mathbf{A}} := \inf\{t \geq 0, \quad \mathbf{A}_t \notin \mathrm{D}\}. \tag{3.8.10}$$

— Pour tout $t \geq 0$, on met

$$\mathbf{Y}_t := \sum_{n=1}^{+\infty} \mathbf{A}^{(n)}_{t-\mathbf{s}_{n-1}} \mathbb{I}_{[\mathbf{s}_{n-1}, \mathbf{s}_n)}(t); \tag{3.8.11}$$

$$\mathbf{s}_0 := 0, \quad \mathbf{s}_n := \sum_{k=1}^{n} \tau_{\mathbf{A}}^{(k)}; \tag{3.8.12}$$

avec $(\mathbf{A}^{(n)})_{n \geq 1}$ une suite de copies indépendantes du processus $(\mathbf{A}_t)_{t \geq 0}$ telle que

$$\mathbf{A}_0^{(1)} \sim \mu \in \mathscr{P}(\mathrm{D}); \tag{3.8.13}$$

$$\forall n \geq 2, \quad \mathbf{A}_0^{(n)} \sim \mu^{\mathbf{Y}}_{\mathbf{s}_{n-1}} := \frac{1}{\mathbf{s}_{n-1}} \int_0^{\mathbf{s}_{n-1}} \delta_{\mathbf{Y}_s} ds. \tag{3.8.14}$$

Dans la continuité de l'article du chapitre 3, l'objectif est ici d'étudier la dynamique auto-attractive associée à ce processus et donc en particulier, la convergence de la mesure d'occupation du processus redistribué selon la mesure d'occupation. Il apparaît que dans le cas du *soft killing*, les outils développés dans le chapitre 3 pourraient assez facilement s'adapter. En revanche, dans le cadre du *hard killing*, c'est-à-dire quand on regarde par exemple le temps de sortie d'un domaine, dans ce cadre, il y a des difficultés intrinsèques concernant la tension, la régularité du sémi-groupe tué ainsi que le contrôle de la probabilité de mourir au voisinage du bord.

Même lorsque le domaine est borné, une des premières difficultés du cas continu est que l'espace des probabilités sur le domaine n'est pas compact, et donc la tension même de la suite des mesures d'occupation n'est pas établie. Néanmoins à l'aide d'outils semblables à ceux développés dans les articles [31] et [199] de Benaïm et al. et Panloup and Reygner, nous sommes en mesure d'obtenir ce résultat.

Théorème (Tension du flot des mesures d'occupation)

Supposons que D est un ouvert connexe borné de bord ∂D de classe $\mathscr{C}^2$ et, b et σ sont jointement continus et bornés; il existe $\lambda_\sigma > 0$ et $g : (0,+\infty) \to (0,+\infty)$ tels que pour tout $\mu \in \mathscr{P}(D)$, $\sigma\sigma^*(\cdot,\mu)$ est λ_σ-uniformément elliptique et

$$\forall \rho > 0, \quad \sup_{(x,y)\in D\times D, \|x-y\|=\rho} \frac{\|\sigma(x,\mu)-\sigma(y,\mu)\|}{\rho} \le g(\rho); \tag{3.8.15}$$

$$\int_0^1 g(r)\,dr < +\infty. \tag{3.8.16}$$

Nous avons

$$\forall \varepsilon > 0 \quad \exists \delta > 0, \quad \liminf_{t\to+\infty} \mu_t^{\mathbf{Y}}(\{x \in D, \quad \psi_D(x) < \delta\}) \le \varepsilon. \tag{3.8.17}$$

La tension permet à son tour de montrer l'existence d'une QSD.

Théorème (Existence de QSDs de McKean-Vlasov)

Sous les hypothèses assurant la tension, il existe $\beta > 0$ et $V_\beta : \overline{D} \to (0, +\infty)$ telles que $V_\beta \in \mathscr{C}^0(\overline{D})$, $V_\beta > 0$ sur D et $V_\beta \equiv 0$ sur ∂D ; l'ensemble

$$E_\beta := \left\{ \mu \in \mathscr{P}(D), \quad \int \frac{1}{V_\beta} d\mu \leq \beta \right\} \tag{3.8.18}$$

est un convexe compact non vide de $\mathscr{P}(D)$ muni de la topologie de la convergence faible ; $\mathbb{QSD} \subset \Pi(\mathscr{P}(D)) \subset E_\beta$ avec

$$\Pi : \quad \mathscr{P}(D) \longrightarrow \mathscr{P}(D) \tag{3.8.19}$$

$$\mu \longmapsto \Pi_\mu := \frac{\mu A_\mu}{\langle \mu A_\mu, \mathbb{I}_D \rangle}. \tag{3.8.20}$$

En particulier, si Π est faiblement continue sur E_β, il existe une QSD de McKean-Vlasov :

$$\mathbb{QSD} = \left\{ \mu \in E_\beta, \quad \Pi_\mu = \mu \right\} \neq \emptyset. \tag{3.8.21}$$

Les difficultés restantes résident plutôt dans l'identification de la limite qui implique des contrôles du sémi-groupe tué, de l'opérateur de Green-Poisson et de leurs dérivées que nous n'avons pas encore établis. Ces difficultés palliées, nous aurions le résultat principal ci-dessous.

Conjecture (Tension du shift du flot, EDO et QSDs)

Le shift du flot des mesures d'occupation est tendu et ses valeurs d'adhérence sont solutions stationnaires de

$$\dot{\mu} = \Pi_\mu - \mu. \tag{3.8.22}$$

Afin de prouver cette conjecture, la difficulté principale réside dans la capacité à généraliser les propriétés de convergence de l'EDO telles que celles établies dans la proposition 3.4.6. En effet, des conditions de "minoration/majoration" ne peuvent plus être satisfaites sur le domaine D tout entier et il faudrait donc trouver un moyen de se ramener à un compact à la manière de Champagnat and Villemonais. Pour terminer, notons que des briques complémentaires ont été établies concernant les contrôles des gradients du sémi-groupe tué et de

l'opérateur de Green-Poisson associé. Ceux-ci s'appuient sur des résultats de Malliavin and Thalmaier relatifs au calcul de Malliavin pour les processus tués.

Publications

▷ On the exponential ergodicity of the McKean-Vlasov SDE depending on a polynomial interaction
▷ Exponential ergodicity of mean-field Langevin dynamics by synchronous coupling
▷ Quasi-Stationary Distributions of Interacting Dynamical Systems and their approximation

Bibliographie

[1] M. A. Aboubacrine Assadek. Exponential ergodicity of mean-field Langevin dynamics by synchronous coupling. working paper or preprint, Sept. 2025. URL https://hal.science/hal-05236513.

[2] M. A. Aboubacrine Assadek and F. Panloup. Quasi-Stationary Distributions of Interacting Dynamical Systems and their approximation. working paper or preprint, June 2025. URL https://hal.science/hal-05130267.

[3] S. Aida and I. o. Shigekawa. Logarithmic Sobolev inequalities and spectral gaps : perturbation theory. *J. Funct. Anal.*, 126(2) :448–475, 1994. ISSN 0022-1236,1096-0783. doi : 10.1006/jfan.1994.1154. URL https://doi.org/10.1006/jfan.1994.1154.

[4] D. Aldous, B. Flannery, and J.-L. Palacios. Two applications of urn processes : the fringe analysis of search trees and the simulation of quasi-stationary distributions of markov chains. *Probab. Eng. Inf. Sci.*, 2(3) :293–307, 1988.

[5] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005. ISBN 978-3-7643-2428-5 ; 3-7643-2428-7.

[6] M. A. A. A. Assadeck. On the exponential ergodicity of the mckean-vlasov sde depending on a polynomial interaction, 2023.

[7] M. A. A. A. Assadeck. On the exponential ergodicity of the mckean-vlasov sde depending on a polynomial interaction, 2023. URL https://arxiv.org/abs/2306.01347.

[8] A. Asselah, P. A. Ferrari, and P. Groisman. Quasistationary distributions and Fleming-Viot processes in finite spaces. *J. Appl. Probab.*, 48(2) :322–332, 2011. ISSN 0021-9002. doi : 10.1239/jap/1308662630. URL https://fennec.univ-mlv.fr:443/http/dx.doi.org/10.1239/jap/1308662630.

[9] K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications, Inc., Mineola, NY, 2004. ISBN 0-486-43474-5. Reprint of the 1972 original [Springer, New York; MR0373040].

[10] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994. doi : 10.1007/BFb0073872. URL http://dx.doi.org/10.1007/BFb0073872.

[11] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985. doi : 10.1007/BFb0075847. URL https://doi.org/10.1007/BFb0075847.

[12] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13 :60–66, 2008. doi : 10.1214/ECP.v13-1352. URL https://doi.org/10.1214/ECP.v13-1352.

[13] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes : Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3) :727–759, 2008. ISSN 0022-1236. doi : 10.1016/j.jfa.2007.11.002. URL http://dx.doi.org/10.1016/j.jfa.2007.11.002.

[14] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes : Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3) :727–759, 2008. ISSN 0022-1236. doi : 10.1016/j.jfa.2007.11.002. URL https://doi.org/10.1016/j.jfa.2007.11.002.

[15] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014. ISBN 978-3-319-00226-2 ; 978-3-319-00227-9. doi : 10.1007/978-3-319-00227-9. URL https://doi.org/10.1007/978-3-319-00227-9.

[16] R. Bauerschmidt and T. Bodineau. A very simple proof of the LSI for high temperature spin systems. *J. Funct. Anal.*, 276(8) :2582–2588, 2019. ISSN 0022-1236. doi : 10.1016/j.jfa.2019.01.007. URL https://doi.org/10.1016/j.jfa.2019.01.007.

[17] N. Bellomo, P. Degond, and E. Tadmor, editors. *Active particles. Vol. 1. Advances in theory, models, and applications*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser/Springer, Cham, 2017. ISBN 978-3-319-49994-9; 978-3-319-49996-3.

[18] N. Bellomo, P. Degond, and E. Tadmor, editors. *Active particles. Vol. 2. Advances in theory, models, and applications*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser/Springer, Cham, 2019. ISBN 978-3-030-20296-5; 978-3-030-20297-2. doi : 10.1007/978-3-030-20297-2. URL https://doi.org/10.1007/978-3-030-20297-2.

[19] G. Ben Arous and M. Brunaud. Méthode de Laplace : étude variationnelle des fluctuations de diffusions de type "champ moyen". *Stochastics Stochastics Rep.*, 31(1-4) :79–144, 1990. ISSN 1045-1129. doi : 10.1080/03610919008833649. URL https://doi.org/10.1080/03610919008833649.

[20] M. Benaïm. Vertex-reinforced random walks and a conjecture of Pemantle. *Ann. Probab.*, 25(1) :361–392, 1997. ISSN 0091-1798. doi : 10.1214/aop/1024404292. URL http://dx.doi.org/10.1214/aop/1024404292.

[21] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 1–68. Springer, Berlin, 1999. doi : 10.1007/BFb0096509. URL http://dx.doi.org/10.1007/BFb0096509.

[22] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 1–68. Springer, Berlin, 1999. ISBN 3-540-66342-8. doi : 10.1007/BFb0096509. URL https://doi.org/10.1007/BFb0096509.

[23] M. Benaïm. On Gradient like Properties of Population games, Learning models and Self Reinforced Processes. *ArXiv e-prints*, Sept. 2014. To appear in *Dynamics, Games and Science, DGS II*, Springer-Verlag (J.P Bourguignon, R. Jelstch, A. Pindo and M. Viana, eds).

[24] M. Benaïm and B. Cloez. A stochastic approximation approach to quasi-stationary distributions on finite spaces. *Electron. Commun. Probab.*, 20 :no. 37, 14, 2015. ISSN 1083-589X. doi : 10.1214/ECP.v20-3956. URL http://dx.doi.org/10.1214/ECP.v20-3956.

[25] M. Benaïm and M. W. Hirsch. Asymptotic pseudotrajectories and chain recurrent flows, with applications. *J. Dynam. Differential Equations*, 8(1) :141–176, 1996. ISSN 1040-7294. doi : 10.1007/BF02218617. URL http://dx.doi.org/10.1007/BF02218617.

[26] M. Benaïm and J. W. Weibull. Deterministic approximation of stochastic evolution in games. *Econometrica*, 71(3) :873–903, 2003. ISSN 0012-9682. doi : 10.1111/1468-0262.00429. URL https://dx-doi-org.fennec.u-pem.fr/10.1111/1468-0262.00429.

[27] M. Benaïm, M. Ledoux, and O. Raimond. Self-interacting diffusions. *Probab. Theory Related Fields*, 122(1) :1–41, 2002. ISSN 0178-8051. doi : 10.1007/s004400100161. URL https://dx-doi-org.fennec.u-pem.fr/10.1007/s004400100161.

[28] M. Benaim, B. Cloez, and F. Panloup. Stochastic approximation of quasi-stationary distributions on compact spaces and applications. *Ann. Appl. Probab.*, 28(4) :2370–2416, 2018. ISSN 1050-5164,2168-8737. doi : 10.1214/17-AAP1360. URL https://doi.org/10.1214/17-AAP1360.

[29] M. Benaim, B. Cloez, and F. Panloup. Stochastic approximation of quasi-stationary distributions on compact spaces and applications. *Ann. Appl. Probab.*, 28(4) :2370–2416, 2018. ISSN 1050-5164. doi : 10.1214/17-AAP1360. URL https://doi.org/10.1214/17-AAP1360.

[30] M. Benaïm, N. Champagnat, and D. Villemonais. Stochastic approximation of quasi-stationary distributions for diffusion processes in a bounded domain. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(2) :726–739, 2021. ISSN 0246-0203,1778-7017. doi : 10.1214/20-aihp1093. URL https://doi.org/10.1214/20-aihp1093.

[31] M. Benaïm, N. Champagnat, and D. Villemonais. Stochastic approximation of quasi-stationary distributions for diffusion processes in a bounded domain. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(2) :726–739, 2021. ISSN 0246-0203,1778-7017. doi : 10.1214/20-aihp1093. URL https://doi.org/10.1214/20-aihp1093.

[32] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive algorithms and stochastic approximations*, volume 22 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990. ISBN 3-540-52894-6. Translated from the French by Stephen S. Wilson.

[33] M. Bieniek, K. Burdzy, and S. Pal. Extinction of Fleming-Viot-type particle systems with strong drift. *Electron. J. Probab.*, 17 :no. 11, 15, 2012. ISSN 1083-6489. doi : 10.1214/EJP.v17-1770. URL http://dx.doi.org/10.1214/EJP.v17-1770.

[34] J. Blanchet, P. Glynn, and S. Zheng. Theoretical analysis of a Stochastic Approximation approach for computing Quasi-Stationary distributions. *ArXiv e-prints*, Jan. 2014.

[35] S. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3) :383–400, 1997. ISSN 0178-8051. doi : 10.1007/s004400050090. URL https://doi.org/10.1007/s004400050090.

[36] T. Bodineau and B. Helffer. The log-Sobolev inequality for unbounded spin systems. *J. Funct. Anal.*, 166(1) :168–178, 1999. ISSN 0022-1236. doi : 10.1006/jfan.1999.3419. URL https://doi.org/10.1006/jfan.1999.3419.

[37] T. Bodineau and B. Helffer. Correlations, spectral gap and log-Sobolev inequalities for unbounded spins systems. In *Differential equations and mathematical physics (Birmingham, AL, 1999)*, volume 16 of *AMS/IP Stud. Adv. Math.*, pages 51–66. Amer. Math. Soc., Providence, RI, 2000. doi : 10.1090/amsip/016/06. URL https://doi.org/10.1090/amsip/016/06.

[38] T. Bodineau, I. Gallagher, and L. Saint-Raymond. The Brownian motion as the limit of a deterministic system of hard-spheres. *Invent. Math.*, 203(2) :493–553, 2016. ISSN 0020-9910. doi : 10.1007/s00222-015-0593-9. URL https://doi.org/10.1007/s00222-015-0593-9.

[39] F. Bolley, I. Gentil, and A. Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2) :429–445, 2013. ISSN 0003-9527. doi : 10.1007/s00205-012-0599-z. URL https://doi.org/10.1007/s00205-012-0599-z.

[40] F. c. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model.*

Numer. Anal., 44(5) :867–884, 2010. ISSN 0764-583X,1290-3841. doi : 10.1051/m2an/2010045. URL https://doi.org/10.1051/m2an/2010045.

[41] F. c. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model. Numer. Anal.*, 44(5) :867–884, 2010. ISSN 0764-583X,1290-3841. doi : 10.1051/m2an/2010045. URL https://doi.org/10.1051/m2an/2010045.

[42] E. Bolthausen. Laplace approximations for sums of independent random vectors. *Probab. Theory Relat. Fields*, 72(2) :305–318, 1986. ISSN 0178-8051. doi : 10.1007/BF00699109. URL https://doi.org/10.1007/BF00699109.

[43] L. Boltzmann. *Lectures on gas theory*. University of California Press, Berkeley-Los Angeles, Calif., 1964. Translated by Stephen G. Brush.

[44] L. Boltzmann. *Wissenschaftliche Abhandlungen von Ludwig Boltzmann. I. Band (1865–1874) ; II. Band (1875–1881) ; III. Band (1882–1905)*. Chelsea Publishing Co., New York, 1968. Herausgegeben von Fritz Hasenöhrl.

[45] L. Boltzmann. Journey of a German professor to Eldorado. *Transport Theory Statist. Phys.*, 20(5-6) :499–523, 1991. ISSN 0041-1450. doi : 10.1080/00411459108203917. URL https://doi.org/10.1080/00411459108203917. Translated from the German.

[46] B. Bouchard, S. Geiss, and E. Gobet. First time to exit of a continuous Itô process : general moment estimates and L_1-convergence rate for discrete time approximations. *Bernoulli*, 23(3) :1631–1662, 2017. ISSN 1350-7265,1573-9759. doi : 10.3150/15-BEJ791. URL https://doi.org/10.3150/15-BEJ791.

[47] O. Brandière and M. Duflo. Les algorithmes stochastiques contournent-ils les pièges ? *Ann. Inst. H. Poincaré Probab. Statist.*, 32(3) :395–427, 1996. ISSN 0246-0203. URL http://www.numdam.org/item?id=AIHPB_1996__32_3_395_0.

[48] K. Burdzy, R. Hołyst, and P. March. A Fleming-Viot particle representation of the Dirichlet Laplacian. *Comm. Math. Phys.*, 214(3) :679–703, 2000. ISSN 0010-3616. doi : 10.1007/s002200000294. URL https://fennec.univ-mlv.fr:443/http/dx.doi.org/10.1007/s002200000294.

[49] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The master equation and the convergence problem in mean field games*, volume 201 of *Annals of Mathematics*

Studies. Princeton University Press, Princeton, NJ, 2019. ISBN 978-0-691-19071-6; 978-0-691-19070-9. doi : 10.2307/j.ctvckq7qf. URL https://doi.org/10.2307/j.ctvckq7qf.

[50] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. II*, volume 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. ISBN 978-3-319-56435-7; 978-3-319-56436-4. Mean field games with common noise and master equations.

[51] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. ISBN 978-3-319-56437-1; 978-3-319-58920-6. Mean field FBSDEs, control, and games.

[52] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations : entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3) :971–1018, 2003. ISSN 0213-2230. doi : 10.4171/RMI/376. URL https://doi.org/10.4171/RMI/376.

[53] J. A. Carrillo, S. Jin, L. Li, and Y. Zhu. A consensus-based global optimization method for high dimensional machine learning problems. *ESAIM Control Optim. Calc. Var.*, 27(suppl.) :Paper No. S5, 22, 2021. ISSN 1292-8119. doi : 10.1051/cocv/2020046. URL https://doi.org/10.1051/cocv/2020046.

[54] P. Cattiaux. Stochastic calculus and degenerate boundary value problems. *Ann. Inst. Fourier (Grenoble)*, 42(3) :541–624, 1992. ISSN 0373-0956,1777-5310. doi : 10.5802/aif.1302. URL https://doi.org/10.5802/aif.1302.

[55] P. Cattiaux and A. Guillin. Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity. *J. Funct. Anal.*, 272(6) :2361–2391, 2017. ISSN 0022-1236,1096-0783. doi : 10.1016/j.jfa.2016.10.003. URL https://doi.org/10.1016/j.jfa.2016.10.003.

[56] P. Cattiaux, G. Arnaud, and M. Florent. Probabilistic approach for granular media equations in the non uniformly convex case, 2006. URL https://arxiv.org/abs/math/0603541.

[57] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-

2) :19–40, 2008. ISSN 0178-8051. doi : 10.1007/s00440-007-0056-3. URL https://doi.org/10.1007/s00440-007-0056-3.

[58] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5) :1926–1969, 2009. ISSN 0091-1798,2168-894X. doi : 10.1214/09-AOP451. URL https://doi.org/10.1214/09-AOP451.

[59] P. Cattiaux, A. Guillin, and L.-M. Wu. A note on Talagrand's transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148(1-2) :285–304, 2010. ISSN 0178-8051. doi : 10.1007/s00440-009-0231-9. URL https://doi.org/10.1007/s00440-009-0231-9.

[60] P. Cattiaux, A. Guillin, and L.-M. Wu. Some remarks on weighted logarithmic Sobolev inequality. *Indiana Univ. Math. J.*, 60(6) :1885–1904, 2011. ISSN 0022-2518,1943-5258. doi : 10.1512/iumj.2011.60.4405. URL https://doi.org/10.1512/iumj.2011.60.4405.

[61] D. Chafaï and A. Joulin. Intertwining and commutation relations for birth–death processes. *Bernoulli*, 19(5A) :1855–1879, 2013. ISSN 1350-7265. doi : 10.3150/12-BEJ433. URL http://dx.doi.org/10.3150/12-BEJ433.

[62] L.-P. Chaintron and A. Diez. Propagation of chaos : A review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models*, 15(6) :895–, 2022. ISSN 1937-5093. doi : 10.3934/krm.2022017. URL https://doi.org/10.3934/krm.2022017.

[63] L.-P. Chaintron and A. Diez. Propagation of chaos : a review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models*, 15(6) :895–1015, 2022. ISSN 1937-5093,1937-5077. doi : 10.3934/krm.2022017. URL https://doi.org/10.3934/krm.2022017.

[64] L.-P. Chaintron and A. Diez. Propagation of chaos : a review of models, methods and applications. II. Applications. *Kinet. Relat. Models*, 15(6) :1017–1173, 2022. ISSN 1937-5093,1937-5077. doi : 10.3934/krm.2022018. URL https://doi.org/10.3934/krm.2022018.

[65] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *ArXiv e-prints*, Apr. 2014. accepted for publication in Probability Theory and Related Fields.

[66] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields*, 164(1-2) :243–283, 2016. ISSN 0178-8051,1432-2064. doi : 10.1007/s00440-014-0611-7. URL https://doi.org/10.1007/s00440-014-0611-7.

[67] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1) :177–199, 2017. ISSN 1980-0436.

[68] N. Champagnat and D. Villemonais. Uniform convergence to the Q-process. *Electron. Commun. Probab.*, 22 :Paper No. 33, 7, 2017. ISSN 1083-589X. doi : 10.1214/17-ECP63. URL https://doi.org/10.1214/17-ECP63.

[69] N. Champagnat and D. Villemonais. Uniform convergence of conditional distributions for absorbed one-dimensional diffusions. *Adv. in Appl. Probab.*, 50(1) :178–203, 2018. ISSN 0001-8678,1475-6064. doi : 10.1017/apr.2018.9. URL https://doi.org/10.1017/apr.2018.9.

[70] N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. *Stochastic Process. Appl.*, 135 :51–74, 2021. ISSN 0304-4149. doi : 10.1016/j.spa.2020.12.005. URL https://doi.org/10.1016/j.spa.2020.12.005.

[71] N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. *Electron. J. Probab.*, 28 :Paper No. 22, 84, 2023. doi : 10.1214/22-ejp880. URL https://doi.org/10.1214/22-ejp880.

[72] N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais. Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions. In *Séminaire de Probabilités XLIX*, volume 2215 of *Lecture Notes in Math.*, pages 165–182. Springer, Cham, 2018.

[73] J.-R. Chazottes, P. Collet, and S. Méléard. Sharp asymptotics for the quasi-stationary distribution of birth-and-death processes. *Probab. Theory Related Fields*, 164(1-2) :

285–332, 2016. ISSN 0178-8051,1432-2064. doi : 10.1007/s00440-014-0612-6. URL https://doi.org/10.1007/s00440-014-0612-6.

[74] F. Chen, Z. Ren, and S. Wang. Uniform-in-time propagation of chaos for mean field langevin dynamics, 2023. URL https://arxiv.org/abs/2212.03050.

[75] F. Chen, Y. Lin, Z. Ren, and S. Wang. Uniform-in-time propagation of chaos for kinetic mean field langevin dynamics. *Electronic Journal of Probability*, 29(none), Jan. 2024. ISSN 1083-6489. doi : 10.1214/24-ejp1079. URL http://dx.doi.org/10.1214/24-EJP1079.

[76] M. Chen. Optimal Markovian couplings and applications. *Acta Math. Sinica (N.S.)*, 10(3) :260–275, 1994. ISSN 1000-9574. doi : 10.1007/BF02560717. URL http://dx.doi.org/10.1007/BF02560717. A Chinese summary appears in Acta Math. Sinica **3**8 (1995), no. 4, 575.

[77] M. F. Chen. *From Markov chains to nonequilibrium particle systems*. World Scientific Publishing Co., Inc., River Edge, NJ, 1992. ISBN 981-02-0639-9. doi : 10.1142/1389. URL http://dx.doi.org/10.1142/1389.

[78] M.-F. Chen. *From Markov chains to non-equilibrium particle systems*. World Scientific Publishing Co. Inc., River Edge, NJ, second edition, 2004. ISBN 981-238-811-7. doi : 10.1142/9789812562456. URL http://dx.doi.org/10.1142/9789812562456.

[79] B. Cloez. Limit theorems for some branching measure-valued processes. *ArXiv e-prints*, June 2011.

[80] B. Cloez. *Comportement asymptotique de processus avec sauts et applications pour des modèles avec branchement*. PhD thesis, Université Paris-Est, 2013.

[81] B. Cloez and M. Hairer. Exponential ergodicity for Markov processes with random switching. *ArXiv e-prints*, Mar. 2013.

[82] B. Cloez and M.-N. Thai. Quantitative results for the Fleming-Viot particle system in discrete space. *ArXiv e-prints*, Dec. 2013.

[83] B. Cloez, L. Journel, P. Monmarche, B. Nectoux, and M. Ramil. Recent advances in the long-time analysis of killed degenerate processes and their particle approximation. In *Journées SMAI 2021*, volume 75 of *ESAIM Proc. Surveys*, pages 60–85. EDP Sci., Les Ulis, 2023.

[84] P. Collet, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions for structured birth and death processes with mutations. *Probab. Theory Related Fields*, 151(1-2) :191–231, 2011. ISSN 0178-8051. doi : 10.1007/s00440-010-0297-4. URL http://dx.doi.org/10.1007/s00440-010-0297-4.

[85] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions*. Probability and its Applications (New York). Springer, Heidelberg, 2013. ISBN 978-3-642-33130-5; 978-3-642-33131-2. doi : 10.1007/978-3-642-33131-2. URL http://dx.doi.org/10.1007/978-3-642-33131-2. Markov chains, diffusions and dynamical systems.

[86] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions*. Probability and its Applications (New York). Springer, Heidelberg, 2013. ISBN 978-3-642-33130-5; 978-3-642-33131-2. doi : 10.1007/978-3-642-33131-2. URL https://doi.org/10.1007/978-3-642-33131-2. Markov chains, diffusions and dynamical systems.

[87] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probability*, 2 :88–100, 1965. ISSN 0021-9002.

[88] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Probability*, 4 :192–196, 1967. ISSN 0021-9002.

[89] V. H. de la Peña. Decoupling and Khintchine's inequalities for U-statistics. *Ann. Probab.*, 20(4) :1877–1892, 1992. ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(199210)20:4<1877:DAKIF>2.0.CO;2-O&origin=MSN.

[90] P. Del Moral. Measure-valued processes and interacting particle systems. Application to nonlinear filtering problems. *Ann. Appl. Probab.*, 8(2) :438–495, 1998. ISSN 1050-5164. doi : 10.1214/aoap/1028903535. URL https://doi.org/10.1214/aoap/1028903535.

[91] P. Del Moral. *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. ISBN 0-387-20268-4. doi : 10.1007/978-1-4684-9393-1. URL https://doi.org/10.1007/978-1-4684-9393-1. Genealogical and interacting particle systems with applications.

[92] P. Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013. ISBN 978-1-4665-0405-9.

[93] P. Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013. ISBN 978-1-4665-0405-9.

[94] P. Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013. ISBN 978-1-4665-0405-9.

[95] P. Del Moral and A. Guionnet. On the stability of measure valued processes with applications to filtering. *C. R. Acad. Sci. Paris Sér. I Math.*, 329(5) :429–434, 1999. ISSN 0764-4442. doi : 10.1016/S0764-4442(00)88619-X. URL https://fennec.univ-mlv.fr:443/http/dx.doi.org/10.1016/S0764-4442(00)88619-X.

[96] P. Del Moral and A. Guionnet. On the stability of measure valued processes with applications to filtering. *C. R. Acad. Sci. Paris Sér. I Math.*, 329(5) :429–434, 1999. ISSN 0764-4442. doi : 10.1016/S0764-4442(00)88619-X. URL https://dx-doi-org.fennec.u-pem.fr/10.1016/S0764-4442(00)88619-X.

[97] P. Del Moral and L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic Process. Appl.*, 86(2) :193–216, 2000. ISSN 0304-4149. doi : 10.1016/S0304-4149(99)00094-0. URL https://fennec.univ-mlv.fr:443/http/dx.doi.org/10.1016/S0304-4149(99)00094-0.

[98] P. Del Moral and L. Miclo. On the stability of nonlinear Feynman-Kac semigroups. *Ann. Fac. Sci. Toulouse Math. (6)*, 11(2) :135–175, 2002. ISSN 0240-2963. URL http://www.numdam.org/item?id=AFST_2002_6_11_2_135_0.

[99] P. Del Moral and L. Miclo. On convergence of chains with occupational self-interactions. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041) :325–346, 2004. ISSN 1364-5021. doi : 10.1098/rspa.2003.1245. URL https://dx-doi-org.fennec.u-pem.fr/10.1098/rspa.2003.1245. Stochastic analysis with applications to mathematical finance.

[100] P. Del Moral and L. Miclo. Self-interacting Markov chains. *Stoch. Anal. Appl.*, 24(3) : 615–660, 2006. ISSN 0736-2994. doi : 10.1080/07362990600632029. URL https://dx-doi-org.fennec.u-pem.fr/10.1080/07362990600632029.

[101] B. Delyon. Stochastic approximation with decreasing gain : Convergence and asymptotic theory. 2000.

[102] J. Dolbeault. Functional inequalities : nonlinear flows and entropy methods as a tool for obtaining sharp and constructive results. *Milan J. Math.*, 89(2) :355–386, 2021. ISSN 1424-9286. doi : 10.1007/s00032-021-00341-y. URL https://doi.org/10.1007/s00032-021-00341-y.

[103] M. Duflo. *Algorithmes stochastiques*, volume 23 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1996. ISBN 3-540-60699-8.

[104] A. Durmus, A. Eberle, A. Guillin, and R. Zimmer. An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.*, 148(12) :5387–5398, 2020. ISSN 0002-9939. doi : 10.1090/proc/14612. URL https://doi.org/10.1090/proc/14612.

[105] A. Eberle. Couplings, distances and contractivity for diffusion processes revisited. *ArXiv e-prints*, May 2013.

[106] A. Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, 166(3-4) :851–886, 2016. ISSN 0178-8051. doi : 10.1007/s00440-015-0673-1. URL https://doi.org/10.1007/s00440-015-0673-1.

[107] A. Eberle, A. Guillin, and R. Zimmer. Quantitative harris type theorems for diffusions and mckean-vlasov processes, 2017. URL https://arxiv.org/abs/1606.06012.

[108] A. Eberle, A. Guillin, and R. Zimmer. Couplings and quantitative contraction rates for langevin dynamics, 2018. URL https://arxiv.org/abs/1703.01617.

[109] A. Eberle, A. Guillin, and R. Zimmer. Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. *Trans. Amer. Math. Soc.*, 371(10) :7135–7173, 2019. ISSN 0002-9947. doi : 10.1090/tran/7576. URL https://doi.org/10.1090/tran/7576.

[110] W. L. Erhan Bayraktar, Qi Feng. Exponential entropy dissipation for weakly self-consistent vlasov-fokker-planck equations. *arXiv :2204.12049v1*, 2022.

[111] P. A. Ferrari and N. Marić. Quasi stationary distributions and Fleming-Viot processes in countable spaces. *Electron. J. Probab.*, 12 :no. 24, 684–702, 2007. ISSN 1083-6489. doi : 10.1214/EJP.v12-415. URL https://fennec.univ-mlv.fr:443/http/dx.doi.org/10.1214/EJP.v12-415.

[112] P. A. Ferrari, H. Kesten, S. Martinez, and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.*, 23(2) :501–521, 1995. ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(199504)23:2<501:EOQDAR>2.0.CO;2-J&origin=MSN.

[113] G. Fort. Central Limit Theorems for Stochastic Approximation with controlled Markov chain dynamics. *ArXiv e-prints*, Sept. 2013. Accepted for publication in Esaim Probab. Stat.

[114] N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure, 2013. URL https://arxiv.org/abs/1312.2128.

[115] I. Gallagher, L. Saint-Raymond, and B. Texier. *From Newton to Boltzmann : hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013. ISBN 978-3-03719-129-3.

[116] F. R. Gantmacher. *The theory of matrices. Vols. 1, 2*. Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959.

[117] C. Giardinà, J. Kurchan, and F. Redig. Duality and exact correlations for a model of heat conduction. *J. Math. Phys.*, 48(3) :033301, 15, 2007. ISSN 0022-2488. doi : 10.1063/1.2711373. URL http://dx.doi.org/10.1063/1.2711373.

[118] C. Giardinà, F. Redig, and K. Vafayi. Correlation inequalities for interacting particle systems with duality. *J. Stat. Phys.*, 141(2) :242–263, 2010. ISSN 0022-4715. doi : 10.1007/s10955-010-0055-0. URL http://dx.doi.org/10.1007/s10955-010-0055-0.

[119] E. Gobet. Weak approximation of killed diffusion using Euler schemes. *Stochastic Processes and their Applications*, 87(2) :167–197, June 2000. URL https://ideas.repec.org/a/eee/spapps/v87y2000i2p167-197.html.

[120] F. Golse. On the dynamics of large particle systems in the mean field limit. In *Macroscopic and large scale phenomena : coarse graining, mean field limits and ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 1–144. Springer, [Cham], 2016. doi : 10.1007/978-3-319-26883-5_1. URL https://doi.org/10.1007/978-3-319-26883-5_1.

[121] G. L. Gong, M. P. Qian, and Z. X. Zhao. Killed diffusions and their conditioning. *Probab. Theory Related Fields*, 80(1) :151–167, 1988. ISSN 0178-8051. doi : 10.1007/BF00348757. URL https://doi.org/10.1007/BF00348757.

[122] C. Graham, T. G. Kurtz, S. Méléard, P. E. Protter, M. Pulvirenti, and D. Talay. *Probabilistic models for nonlinear partial differential equations*, volume 1627 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin ; Centro Internazionale Matematico Estivo (C.I.M.E.), Florence, 1996. ISBN 3-540-61397-8. doi : 10.1007/BFb0093175. URL https://doi.org/10.1007/BFb0093175. Lectures given at the 1st Session and Summer School held in Montecatini Terme, May 22–30, 1995, Edited by Talay and L. Tubaro, Fondazione CIME/CIME Foundation Subseries.

[123] S. Grassi and L. Pareschi. From particle swarm optimization to consensus based optimization : stochastic modeling and mean-field limit. *Math. Models Methods Appl. Sci.*, 31(8) :1625–1657, 2021. ISSN 0218-2025. doi : 10.1142/S0218202521500342. URL https://doi.org/10.1142/S0218202521500342.

[124] P. Groisman and M. Jonckheere. Simulation of quasi-stationary distributions on countable spaces. *Markov Process. Related Fields*, 19(3) :521–542, 2013. ISSN 1024-2953.

[125] S. Grosskinsky, F. Redig, and K. Vafayi. Condensation in the inclusion process and related models. *J. Stat. Phys.*, 142(5) :952–974, 2011. ISSN 0022-4715. doi : 10.1007/s10955-011-0151-9. URL http://dx.doi.org/10.1007/s10955-011-0151-9.

[126] S. Grosskinsky, F. Redig, and K. Vafayi. Dynamics of condensation in the symmetric inclusion process. *Electron. J. Probab.*, 18 :no. 66, 23, 2013. ISSN 1083-6489. doi : 10.1214/EJP.v18-2720. URL http://dx.doi.org/10.1214/EJP.v18-2720.

[127] A. Guillin and P. Monmarché. Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes. *J. Stat. Phys.*, 185(2) :Paper No. 15, 20, 2021. ISSN 0022-4715. doi : 10.1007/s10955-021-02839-6. URL https://doi.org/10.1007/s10955-021-02839-6.

[128] A. Guillin and F.-Y. Wang. Degenerate Fokker-Planck equations : Bismut formula, gradient estimate and Harnack inequality. *J. Differential Equations*, 253(1) :20–40, 2012. ISSN 0022-0396. doi : 10.1016/j.jde.2012.03.014. URL https://doi.org/10.1016/j.jde.2012.03.014.

[129] A. Guillin, C. Léonard, L. Wu, and N. Yao. Transportation-information inequalities for Markov processes. *Probab. Theory Related Fields*, 144(3-4) :669–695, 2009. ISSN 0178-8051. doi : 10.1007/s00440-008-0159-5. URL https://doi.org/10.1007/s00440-008-0159-5.

[130] A. Guillin, W. Liu, L. Wu, and C. Zhang. The kinetic fokker-planck equation with mean field interaction, 2019. URL https://arxiv.org/abs/1912.02594.

[131] A. Guillin, P. L. Bris, and P. Monmarché. Convergence rates for the vlasov-fokker-planck equation and uniform in time propagation of chaos in non convex cases, 2021. URL https://arxiv.org/abs/2105.09070.

[132] A. Guillin, W. Liu, L. Wu, and C. Zhang. The kinetic Fokker-Planck equation with mean field interaction. *J. Math. Pures Appl. (9)*, 150 :1–23, 2021. ISSN 0021-7824. doi : 10.1016/j.matpur.2021.04.001. URL https://doi.org/10.1016/j.matpur.2021.04.001.

[133] A. Guillin, W. Liu, L. Wu, and C. Zhang. Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems. *Ann. Appl. Probab.*, 32(3) :1590–1614, 2022. ISSN 1050-5164. doi : 10.1214/21-aap1707. URL https://doi.org/10.1214/21-aap1707.

[134] A. Guillin, W. Liu, L. Wu, and C. Zhang. Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems. *Ann. Appl. Probab.*, 32(3) :1590–1614, 2022. ISSN 1050-5164,2168-8737. doi : 10.1214/21-aap1707. URL https://doi.org/10.1214/21-aap1707.

[135] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003. doi : 10.1007/978-3-540-36107-7_1. URL https://doi.org/10.1007/978-3-540-36107-7_1.

[136] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011. doi : 10.1007/978-3-0348-0021-1_7. URL http://dx.doi.org/10.1007/978-3-0348-0021-1_7.

[137] M. Hauray and S. Mischler. On Kac's chaos and related problems. *J. Funct. Anal.*, 266 (10) :6055–6157, 2014. ISSN 0022-1236. doi : 10.1016/j.jfa.2014.02.030. URL https://doi.org/10.1016/j.jfa.2014.02.030.

[138] K. Hu, Z. Ren, D. Siska, and L. Szpruch. Mean-field langevin dynamics and energy landscape of neural networks, 2020. URL https://arxiv.org/abs/1905.07769.

[139] K. Hu, Z. Ren, D. Šiška, and L. u. Szpruch. Mean-field Langevin dynamics and energy landscape of neural networks. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(4) :2043–2065, 2021. ISSN 0246-0203. doi : 10.1214/20-aihp1140. URL https://doi.org/10.1214/20-aihp1140.

[140] M. Huang, R. Malhame, and P. Caines. Large population stochastic dynamic games : Closed-loop mckean-vlasov systems and the nash certainty equivalence principle. *Commun. Inf. Syst.*, 6, 01 2006. doi : 10.4310/CIS.2006.v6.n3.a5.

[141] P.-E. Jabin. A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models*, 7(4) :661–711, 2014. ISSN 1937-5093. doi : 10.3934/krm.2014.7.661. URL https://doi.org/10.3934/krm.2014.7.661.

[142] P.-E. Jabin and Z. Wang. Mean field limit for stochastic particle systems. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 379–402. Birkhäuser/Springer, Cham, 2017.

[143] P.-E. Jabin and Z. Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1) :523–591, 2018. ISSN 0020-9910. doi : 10.1007/s00222-018-0808-y. URL https://doi.org/10.1007/s00222-018-0808-y.

[144] A. Joulin. Poisson-type deviation inequalities for curved continuous-time Markov chains. *Bernoulli*, 13(3) :782–798, 2007. ISSN 1350-7265. doi : 10.3150/07-BEJ6039. URL http://dx.doi.org/10.3150/07-BEJ6039.

[145] B. Jourdain. Diffusions with a nonlinear irregular drift coefficient and probabilistic interpretation of generalized Burgers' equations. *ESAIM Probab. Statist.*, 1 :339–355, 1995/97. ISSN 1292-8100,1262-3318. doi : 10.1051/ps:1997113. URL https://doi.org/10.1051/ps:1997113.

[146] B. Jourdain, T. Lelièvre, and R. Roux. Existence, uniqueness and convergence of a particle approximation for the adaptive biasing force process. *M2AN Math. Model. Numer. Anal.*, 44(5) :831–865, 2010. ISSN 0764-583X. doi : 10.1051/m2an/2010044. URL https://doi.org/10.1051/m2an/2010044.

[147] A. Jüngel. *Entropy methods for diffusive partial differential equations*. SpringerBriefs in Mathematics. Springer, [Cham], 2016. ISBN 978-3-319-34218-4; 978-3-319-34219-1. doi : 10.1007/978-3-319-34219-1. URL https://doi.org/10.1007/978-3-319-34219-1.

[148] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197. University of California Press, Berkeley-Los Angeles, Calif., 1956.

[149] M. Kac. *Aspects probabilistes de la théorie du potentiel*, volume 1968 of *Séminaire de Mathématiques Supérieures, No. 32 (Été.* Les Presses de l'Université de Montréal, Montreal, Que., 1970.

[150] M. Kac. The Boltzmann equation : theory and applications. In E. G. D. Cohen and W. Thirring, editors, *Proceedings of the International Symposium "100 Years Boltzmann Equation" in Vienna, 4th-8th September 1972*, pages xii+642. Springer-Verlag, Vienna-New York, 1973. Acta Physica Austriaca Supplementum, X.

[151] M. Kac. Fluctuations near and far from equilibrium. In *Statistical mechanics and statistical methods in theory and application (Proc. Sympos., Univ. Rochester, Rochester, N.Y., 1976)*, pages 203–218. Plenum, New York, 1977.

[152] A. Kazeykina, Z. Ren, X. Tan, and J. Yang. Ergodicity of the underdamped mean-field langevin dynamics, 2023. URL https://arxiv.org/abs/2007.14660.

[153] V. S. Koroljuk and Y. V. Borovskich. *Theory of* U*-statistics*, volume 273 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994. ISBN 0-7923-2608-3. doi : 10.1007/978-94-017-3515-5. URL https://doi.org/10.1007/978-94-017-3515-5. Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors.

[154] S. Kusuoka and Y. Tamura. Gibbs measures for mean field potentials. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 31(1) :223–245, 1984. ISSN 0040-8980.

[155] D. Lamberton and G. Pagès. A penalized bandit algorithm. *Electron. J. Probab.*, 13 :no. 13, 341–373, 2008. ISSN 1083-6489. doi : 10.1214/EJP.v13-489. URL http://dx.doi.org/10.1214/EJP.v13-489.

[156] D. Lamberton, G. Pagès, and P. Tarrès. When can the two-armed bandit algorithm be trusted? *Ann. Appl. Probab.*, 14(3) :1424–1454, 2004. ISSN 1050-5164. doi : 10.1214/105051604000000350. URL http://dx.doi.org/10.1214/105051604000000350.

[157] C. Le Bris, T. Lelièvre, M. Luskin, and D. Perez. A mathematical formalization of the parallel replica dynamics. *Monte Carlo Methods Appl.*, 18(2) :119–146, 2012. ISSN 0929-9629. doi : 10.1515/mcma-2012-0003. URL https://doi.org/10.1515/mcma-2012-0003.

[158] C. Le Bris, T. Lelièvre, M. Luskin, and D. Perez. A mathematical formalization of the parallel replica dynamics. *Monte Carlo Methods Appl.*, 18(2) :119–146, 2012. ISSN 0929-9629,1569-3961. doi : 10.1515/mcma-2012-0003. URL https://doi.org/10.1515/mcma-2012-0003.

[159] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 167–194. Springer, Berlin, 2001. doi : 10.1007/978-3-540-44671-2_13. URL https://doi.org/10.1007/978-3-540-44671-2_13.

[160] T. Lelièvre, M. Rousset, and G. Stoltz. Long-time convergence of an adaptive biasing force method. *Nonlinearity*, 21(6) :1155–1181, 2008. ISSN 0951-7715. doi : 10.1088/0951-7715/21/6/001. URL https://doi.org/10.1088/0951-7715/21/6/001.

[161] T. Lelièvre, M. Ramil, and J. Reygner. Quasi-stationary distribution for the Langevin process in cylindrical domains, Part I : Existence, uniqueness and long-time convergence. *Stochastic Process. Appl.*, 144 :173–201, 2022. ISSN 0304-4149. doi : 10.1016/j.spa.2021.11.005. URL https://doi.org/10.1016/j.spa.2021.11.005.

[162] T. Lelièvre, M. Ramil, and J. Reygner. A probabilistic study of the kinetic Fokker-Planck equation in cylindrical domains. *J. Evol. Equ.*, 22(2) :Paper No. 38, 74, 2022. ISSN 1424-3199. doi : 10.1007/s00028-022-00796-5. URL https://doi.org/10.1007/s00028-022-00796-5.

[163] V. Lemaire. An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.*, 117(10) :1491–1518, 2007. ISSN 0304-4149. doi : 10.1016/j.spa.2007.02.004. URL http://dx.doi.org/10.1016/j.spa.2007.02.004.

[164] V. Lemaire. An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.*, 117(10) :1491–1518, 2007. ISSN 0304-4149,1879-209X. doi : 10.1016/j.spa.2007.02.004. URL https://doi.org/10.1016/j.spa.2007.02.004.

[165] C. Léonard. Large deviations and law of large numbers for a mean field type interacting particle system. *Stochastic Process. Appl.*, 25(2) :215–235, 1987. ISSN 0304-4149. doi : 10.1016/0304-4149(87)90199-2. URL https://doi.org/10.1016/0304-4149(87)90199-2.

[166] T. Lindvall. *Lectures on the coupling method*. Dover Publications, Inc., Mineola, NY, 2002. ISBN 0-486-42145-7. Corrected reprint of the 1992 original.

[167] L. Liu, M. B. Majka, and L. Szpruch. Polyak-lojasiewicz inequality on the space of measures and convergence of mean-field birth-death processes. *arXiv*, 2022.

[168] W. Liu and L. Wu. Large deviations for empirical measures of mean-field Gibbs measures. *Stochastic Process. Appl.*, 130(2) :503–520, 2020. ISSN 0304-4149. doi : 10.1016/j.spa.2019.01.008. URL https://doi.org/10.1016/j.spa.2019.01.008.

[169] P. Malliavin and A. Thalmaier. *Stochastic calculus of variations in mathematical finance*. Springer Finance. Springer-Verlag, Berlin, 2006. ISBN 978-3-540-43431-3; 3-540-43431-3.

[170] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Process. Appl.*, 95(1) :109–132, 2001. ISSN 0304-4149. doi : 10.1016/S0304-4149(01)00095-3. URL https://doi.org/10.1016/S0304-4149(01)00095-3.

[171] F. Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2) :540–560, 2003. ISSN 1050-5164. doi : 10.1214/aoap/1050689593. URL https://doi.org/10.1214/aoap/1050689593.

[172] S. Martinez, J. San Martin, and D. Villemonais. Existence and uniqueness of a quasi-stationary distribution for Markov processes with fast return from infinity. *ArXiv e-prints*, Feb. 2012.

[173] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1) :153–179, 1997. ISSN 0001-8708. doi : 10.1006/aima.1997.1634. URL https://doi.org/10.1006/aima.1997.1634.

[174] H. P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56 :1907–1911, 1966. ISSN 0027-8424. doi : 10.1073/pnas.56.6.1907. URL https://doi.org/10.1073/pnas.56.6.1907.

[175] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967.

[176] S. McNamara and W. R. Young. Kinetics of a one-dimensional granular medium in the quasielastic limit. *Phys. Fluids A*, 5(1) :34–45, 1993. ISSN 0899-8213. doi : 10.1063/1.858896. URL https://doi.org/10.1063/1.858896.

[177] S. Mei, A. Montanari, and P.-M. Nguyen. A mean field view of the landscape of two-layer neural networks. *Proc. Natl. Acad. Sci. USA*, 115(33) :E7665–E7671, 2018. ISSN 0027-8424. doi : 10.1073/pnas.1806579115. URL https://doi.org/10.1073/pnas.1806579115.

[178] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996. doi : 10.1007/BFb0093177. URL https://doi.org/10.1007/BFb0093177.

[179] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9 :340–410, 2012. ISSN 1549-5787.

[180] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.*, 9 :340–410, 2012. ISSN 1549-5787. doi : 10.1214/11-PS191. URL https://doi.org/10.1214/11-PS191.

[181] S. Menozzi and V. Lemaire. On some non asymptotic bounds for the euler scheme. *Electronic Journal of Probability*, 15(none), Jan. 2010. ISSN 1083-6489. doi : 10.1214/ejp.v15-814. URL http://dx.doi.org/10.1214/EJP.v15-814.

[182] M. Métivier and P. Priouret. Théorèmes de convergence presque sure pour une classe d'algorithmes stochastiques à pas décroissant. *Probab. Theory Related Fields*, 74(3) : 403–428, 1987. ISSN 0178-8051. doi : 10.1007/BF00699098. URL http://dx.doi.org/10.1007/BF00699098.

[183] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. ISBN 978-0-521-73182-9. doi : 10.1017/CBO9780511626630. URL http://dx.doi.org/10.1017/CBO9780511626630. With a prologue by Peter W. Glynn.

[184] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3) :518–548, 1993. ISSN 0001-8678. doi : 10.2307/1427522. URL http://dx.doi.org/10.2307/1427522.

[185] S. Mischler. Introduction aux limites de champ moyen pour des systèmes de particules. In *DEA*, pp.166. cel-00576329. Université Paris-Dauphine, 2011.

[186] S. Mischler. Kac's chaos and Kac's program. In *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Année 2012–2013*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XXII, 17. École Polytech., Palaiseau, 2014.

[187] S. Mischler and C. Mouhot. Kac's program in kinetic theory. *Invent. Math.*, 193(1) : 1–147, 2013. ISSN 0020-9910. doi : 10.1007/s00222-012-0422-3. URL https://doi.org/10.1007/s00222-012-0422-3.

[188] P. Monmarché. Long-time behaviour and propagation of chaos for mean field kinetic particles. *Stochastic Process. Appl.*, 127(6) :1721–1737, 2017. ISSN 0304-4149. doi : 10.1016/j.spa.2016.10.003. URL https://doi.org/10.1016/j.spa.2016.10.003.

[189] P. Monmarché. Generalized Γ calculus and application to interacting particles on a graph. *Potential Anal.*, 50(3) :439–466, 2019. ISSN 0926-2601. doi : 10.1007/s11118-018-9689-3. URL https://doi.org/10.1007/s11118-018-9689-3.

[190] P. Monmarché and S. Wang. Logarithmic sobolev inequalities for non-equilibrium steady states, 2023. URL https://arxiv.org/abs/2309.07858.

[191] P. Monmarché, Z. Ren, and S. Wang. Time-uniform log-sobolev inequalities and applications to propagation of chaos. *Electronic Journal of Probability*, 29(none), Jan.

2024. ISSN 1083-6489. doi : 10.1214/24-ejp1217. URL http://dx.doi.org/10.1214/24-EJP1217.

[192] A. Muntean and F. Toschi, editors. *Collective dynamics from bacteria to crowds*, volume 553 of *CISM International Centre for Mechanical Sciences. Courses and Lectures*. Springer, Vienna, 2014. ISBN 978-3-7091-1784-2; 978-3-7091-1785-9. doi : 10.1007/978-3-7091-1785-9. URL https://doi.org/10.1007/978-3-7091-1785-9. An excursion through modeling, analysis and simulation.

[193] G. Naldi, L. Pareschi, and G. Toscani, editors. *Mathematical modeling of collective behavior in socio-economic and life sciences*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston, Ltd., Boston, MA, 2010. ISBN 978-0-8176-4945-6. doi : 10.1007/978-0-8176-4946-3. URL https://doi.org/10.1007/978-0-8176-4946-3.

[194] F. Otto. The geometry of dissipative evolution equations : the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2) :101–174, 2001. ISSN 0360-5302. doi : 10.1081/PDE-100002243. URL https://doi.org/10.1081/PDE-100002243.

[195] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2) :361–400, 2000. ISSN 0022-1236. doi : 10.1006/jfan.1999.3557. URL https://doi.org/10.1006/jfan.1999.3557.

[196] F. Otto and C. Villani. Comment on : "Hypercontractivity of Hamilton-Jacobi equations" [J. Math. Pures Appl. (9) **80** (2001), no. 7, 669–696; MR1846020 (2003b :47073)] by S. G. Bobkov, I. Gentil and M. Ledoux. *J. Math. Pures Appl. (9)*, 80(7) :697–700, 2001. ISSN 0021-7824. doi : 10.1016/S0021-7824(01)01207-7. URL https://doi.org/10.1016/S0021-7824(01)01207-7.

[197] G. Pagès. *Numerical probability*. Universitext. Springer, Cham, 2018. ISBN 978-3-319-90274-6; 978-3-319-90276-0. doi : 10.1007/978-3-319-90276-0. URL https://doi.org/10.1007/978-3-319-90276-0. An introduction with applications to finance.

[198] F. Panloup. Recursive computation of the invariant measure of a stochastic differential equation driven by a Lévy process. *Ann. Appl. Probab.*, 18(2) :379–426, 2008. ISSN 1050-5164. doi : 10.1214/105051607000000285. URL http://dx.doi.org/10.1214/105051607000000285.

[199] F. Panloup and J. Reygner. Asymptotically unbiased approximation of the qsd of diffusion processes with a decreasing time step euler scheme, 2023.

[200] F. Panloup and J. Reygner. Asymptotically unbiased approximation of the qsd of diffusion processes with a decreasing time step euler scheme, 2023. URL https://arxiv.org/abs/2308.16196.

[201] R. Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 4 :1–79, 2007. ISSN 1549-5787. doi : 10.1214/07-PS094. URL http://dx.doi.org/10.1214/07-PS094.

[202] R. Pinnau, C. Totzeck, O. Tse, and S. Martin. A consensus-based model for global optimization and its mean-field limit. *Math. Models Methods Appl. Sci.*, 27(1) : 183–204, 2017. ISSN 0218-2025. doi : 10.1142/S0218202517400061. URL https://doi.org/10.1142/S0218202517400061.

[203] R. G. Pinsky. *Positive harmonic functions and diffusion*, volume 45 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. ISBN 0-521-47014-5. doi : 10.1017/CBO9780511526244. URL http://dx.doi.org/10.1017/CBO9780511526244.

[204] P. Pollett. Quasi-stationary distributions : A bibliography. http://www.maths.uq.edu.au/~pkp/papers/qsds/qsds.pdf.

[205] M. Pulvirenti. Kinetic limits for stochastic particle systems. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 96–126. Springer, Berlin, 1996. doi : 10.1007/BFb0093178. URL https://doi.org/10.1007/BFb0093178.

[206] M. Ramil. Quasi-stationary distribution for the Langevin process in cylindrical domains, part II : overdamped limit. *Electron. J. Probab.*, 27 :Paper No. 59, 18, 2022. doi : 10.1214/22-ejp789. URL https://doi.org/10.1214/22-ejp789.

[207] G. Royer. *Une initiation aux inégalités de Sobolev logarithmiques*, volume 5 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 1999. ISBN 2-85629-075-2.

[208] L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997. doi : 10.1007/BFb0092621. URL http://dx.doi.org/10.1007/BFb0092621.

[209] M. Scheutzow. Uniqueness and non-uniqueness of solutions of vlasov-mckean equations. *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics*, 43 :246 – 256, 1987. URL https://api.semanticscholar.org/CorpusID:121491716.

[210] T. Shiga and H. Tanaka. Central limit theorem for a system of Markovian particles with mean field interactions. *Z. Wahrsch. Verw. Gebiete*, 69(3) :439–459, 1985. ISSN 0044-3719. doi : 10.1007/BF00532743. URL https://doi.org/10.1007/BF00532743.

[211] J. Sirignano and K. Spiliopoulos. Mean field analysis of neural networks : a law of large numbers. *SIAM J. Appl. Math.*, 80(2) :725–752, 2020. ISSN 0036-1399. doi : 10.1137/18M1192184. URL https://doi.org/10.1137/18M1192184.

[212] T. H. R. Skyrme. A non-linear field theory. *Proc. Roy. Soc. London Ser. A*, 260 :127–138, 1961. ISSN 0962-8444,2053-9169. doi : 10.1098/rspa.1961.0018. URL https://doi.org/10.1098/rspa.1961.0018.

[213] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. ISBN 978-3-540-28998-2 ; 3-540-28998-4. Reprint of the 1997 edition.

[214] D. W. Stroock and B. a. Zegarliński. The logarithmic Sobolev inequality for continuous spin systems on a lattice. *J. Funct. Anal.*, 104(2) :299–326, 1992. ISSN 0022-1236. doi : 10.1016/0022-1236(92)90003-2. URL https://doi.org/10.1016/0022-1236(92)90003-2.

[215] D. W. Stroock and B. a. Zegarliński. The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition. *Comm. Math. Phys.*, 144 (2) :303–323, 1992. ISSN 0010-3616. URL http://projecteuclid.org/euclid.cmp/1104249318.

[216] D. W. Stroock and B. a. Zegarliński. The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Comm. Math. Phys.*, 149(1) :175–193, 1992. ISSN 0010-3616. URL http://projecteuclid.org/euclid.cmp/1104251144.

[217] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991. doi : 10.1007/BFb0085169. URL https://doi.org/10.1007/BFb0085169.

[218] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991. doi : 10.1007/BFb0085169. URL http://dx.doi.org/10.1007/BFb0085169.

[219] P. Tarrès. Pièges répulsifs. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(2) :125–130, 2000. ISSN 0764-4442. doi : 10.1016/S0764-4442(00)00139-7. URL http://dx.doi.org/10.1016/S0764-4442(00)00139-7.

[220] A. Thalmaier. On the differentiation of heat semigroups and Poisson integrals. *Stochastics Stochastics Rep.*, 61(3-4) :297–321, 1997. ISSN 1045-1129. doi : 10.1080/17442509708834123. URL https://doi.org/10.1080/17442509708834123.

[221] A. Thalmaier and F.-Y. Wang. A stochastic approach to a priori estimates and Liouville theorems for harmonic maps. *Bull. Sci. Math.*, 135(6-7) :816–843, 2011. ISSN 0007-4497,1952-4773. doi : 10.1016/j.bulsci.2011.07.014. URL https://doi.org/10.1016/j.bulsci.2011.07.014.

[222] O. Tough and J. Nolen. The Fleming-Viot process with McKean-Vlasov dynamics. *Electron. J. Probab.*, 27 :Paper No. 101, 72, 2022. doi : 10.1214/22-ejp820. URL https://doi.org/10.1214/22-ejp820.

[223] O. Tough and J. Nolen. The Fleming-Viot process with McKean-Vlasov dynamics. *Electron. J. Probab.*, 27 :Paper No. 101, 72, 2022. ISSN 1083-6489. doi : 10.1214/22-ejp820. URL https://doi.org/10.1214/22-ejp820.

[224] S. H. J. Tugaut. Non uniqueness of stationary measures for self-stabilizing processes, 2009. URL https://arxiv.org/abs/0903.2460.

[225] E. A. van Doorn and P. K. Pollett. Quasi-stationary distributions for discrete-state models. *European J. Oper. Res.*, 230(1) :1–14, 2013. ISSN 0377-2217. doi : 10.1016/j.ejor.2013.01.032. URL http://dx.doi.org/10.1016/j.ejor.2013.01.032.

[226] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. ISBN 978-3-540-71049-3. doi : 10.1007/978-3-540-71050-9. URL https://doi.org/10.1007/978-3-540-71050-9. Old and new.

[227] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950) :iv+141, 2009. ISSN 0065-9266. doi : 10.1090/S0065-9266-09-00567-5. URL https://doi.org/10.1090/S0065-9266-09-00567-5.

[228] C. Villani. Fisher information in kinetic theory, 2025. URL https://arxiv.org/abs/2501.00925.

[229] D. Villemonais. Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift. *Electron. J. Probab.*, 16 :no. 61, 1663–1692, 2011. ISSN 1083-6489. doi : 10.1214/EJP.v16-925. URL http://dx.doi.org/10.1214/EJP.v16-925.

[230] D. Villemonais. *Distributions quasi-stationnaires et méthodes particulaires pour l'approximation de processus conditionné.* PhD thesis, École Polytechnique, 2011.

[231] D. Villemonais. General approximation method for the distribution of Markov processes conditioned not to be killed. *ESAIM Probab. Stat.*, 18 :441–467, 2014. ISSN 1292-8100. doi : 10.1051/ps/2013045. URL http://dx.doi.org/10.1051/ps/2013045.

[232] M.-K. von Renesse and K.-T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7) :923–940, 2005. ISSN 0010-3640. doi : 10.1002/cpa.20060. URL http://dx.doi.org/10.1002/cpa.20060.

[233] A. Q. Wang, G. O. Roberts, and D. Steinsaltz. An approximation scheme for quasi-stationary distributions of killed diffusions. *Stochastic Process. Appl.*, 130(5) :3193–3219, 2020. ISSN 0304-4149,1879-209X. doi : 10.1016/j.spa.2019.09.010. URL https://doi.org/10.1016/j.spa.2019.09.010.

[234] S. Wang. Uniform log-sobolev inequalities for mean field particles with flat-convex energy, 2024. URL https://arxiv.org/abs/2408.03283.

[235] L. Wu. Gradient estimates of Poisson equations on Riemannian manifolds and applications. *J. Funct. Anal.*, 257(12) :4015–4033, 2009. ISSN 0022-1236. doi : 10.1016/j.jfa.2009.07.013. URL https://doi.org/10.1016/j.jfa.2009.07.013.

[236] N. Yoshida. Application of log-Sobolov inequality to the stochastic dynamics of unbounded spin systems on the lattice. *J. Funct. Anal.*, 173(1) :74–102, 2000. ISSN 0022-1236. doi : 10.1006/jfan.1999.3558. URL https://doi.org/10.1006/jfan.1999.3558.

[237] N. Yoshida. The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(2) :223–243, 2001. ISSN 0246-0203. doi : 10.1016/S0246-0203(00)01066-9. URL https://doi.org/10.1016/S0246-0203(00)01066-9.

[238] B. Zegarlinski. The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice. *Comm. Math. Phys.*, 175(2) :401–432, 1996. ISSN 0010-3616. URL http://projecteuclid.org/euclid.cmp/1104275930.

[239] B. a. Zegarliński. Dobrushin uniqueness theorem and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 105(1) :77–111, 1992. ISSN 0022-1236. doi : 10.1016/0022-1236(92)90073-R. URL https://doi.org/10.1016/0022-1236(92)90073-R.

COLLEGE ED MATHEMATIQUES
DOCTORAL STIC
PAYS DE LA LOIRE

université angers

Titre : Ergodicité de dynamiques de McKean-Vlasov

Mots clés : mesure (quasi-)stationnaire ; champ moyen ; inégalité fonctionnelle ; semi-flot

Résumé : Guillin et al. ont établi des inégalités fonctionnelles non linéaires pour l'équation des milieux granulaires. Le chapitre 1 généralise ces résultats à l'équation de Langevin de champ moyen. Bolley et al. ont montré par couplage synchrone une contraction exponentielle en métrique de Wasserstein pour l'équation de Vlasov-Fokker-Planck. Ils en ont déduit une propagation du chaos uniforme en temps. Le chapitre 2 étend ces résultats aux équations de Langevin de champ moyen. Benaim et al. ont construit et étudié un algorithme basé sur l'(auto-)interaction d'une dynamique avec sa mesure d'occupation pour approximer des distributions quasi-stationnaires (QSDs) de chaînes de Markov générales conditionnées à rester dans un ensemble compact. Le chapitre 3 propose d'aborder le cas des dynamiques de McKean-Vlasov absorbées. Nous présentons des conditions garantissant que les limites faibles de ces suites de mesures aléatoires sont des QSDs de la dynamique donnée. Nous prouvons également des résultats de tension dans le cas non compact. Ces conditions générales sont ensuite appliquées aux schémas d'Euler des EDS de McKean-Vlasov et le comportement de ces QSDs lorsque le pas h tend vers 0 est étudié dans le cas compact. Nos résultats permettent également d'envisager quelques exemples dans le cas non compact et de nouveaux critères de tension sont fournis dans ce cadre. Enfin, nous illustrons nos résultats théoriques par plusieurs simulations.

Title: Ergodicity of McKean-Vlasov dynamics

Keywords: (quasi-)stationary measure; mean field; functional inequality; semi-flow

Abstract: Guillin et al. established nonlinear functional inequalities for the equation of granular media. The chapter 1 generalizes these results to the mean-field Langevin equation. Bolley et al. have shown by synchronous coupling an exponential contraction in Wasserstein metric for the Vlasov-Fokker-Planck equation. They deduced a uniform propagation of chaos in time. The chapter 2 extends these results to the mean-field Langevin equations. Benaim et al. built and studied an algorithm based on the (self-)interaction of a dynamics with its occupation measure to approximate Quasi-Stationary Distributions (QSD) of general Markov chains conditioned to stay in a compact set. The chapter 3 propose to tackle the case of absorbed McKean-Vlasov dynamics. We are able to exhibit some conditions which guarantee that weak limits of these sequences of random measures are QSDs of the given dynamics. We also prove tightness results in the non-compact case. These general conditions are then applied to Euler schemes of McKean-Vlasov SDEs and the behavior of these QSDs when the step h goes to 0 is investigated in the compact case. Our results also allow to consider some examples in the non-compact case and some new tightness criterions are also provided in this setting. Finally, we illustrate our theoretical results with several simulations.