SORBONNE UNIVERSITÉ
LJLL

Doctoral School **École Doctorale Sciences Mathématiques de Paris Centre**

University Department **Laboratoire Jacques-Louis Lions**

Thesis defended by **Pierre Le Bris**

Defended on **June 20, 2023**

In order to become Doctor from Sorbonne Université

Academic Field **Applied Mathematics**

Systems of particles in (singular) interaction : long-time behavior and propagation of chaos

Thesis supervised by Arnaud Guillin
Pierre Monmarché

Committee members

Referees	François Delarue	Professor at Université Côte d'Azur	
	Martin Hairer	Professor at EPFL and Imperial College London	
Examiners	Thierry Bodineau	Senior Researcher at IHES	
	Nicolas Fournier	Professor at Sorbonne Université	
	Sylvia Serfaty	Professor at New York University	
	Sylvie Méleard	Professor at Ecole Polytechnique	Committee President
Supervisors	Arnaud Guillin	Professor at Université Clermont-Auvergne	
	Pierre Monmarché	HDR Associate Professor at Sorbonne Université	

Colophon

Doctoral dissertation entitled "Systems of particles in (singular) interaction : long-time behavior and propagation of chaos", written by Pierre Le Bris, completed on August 1, 2023, typeset with the document preparation system LaTeX and the yathesis class dedicated to theses prepared in France.

Keywords: probability theory, stochastic calculus, coupling methods, propagation of chaos, logarithmic sobolev inequality.

Mots clés : probabilités, calcul stochastique, méthodes de couplage, propagation du chaos, inégalité de sobolev logarithmique.

This thesis has been prepared at

Laboratoire Jacques-Louis Lions
Sorbonne Université
Campus Pierre et Marie Curie
4 place Jussieu
75005 Paris
France

☎ +33 1 44 27 42 98
Web Site https://ljll.math.upmc.fr/

Alors, comment j'vais t'dire ?...

SYSTEMS OF PARTICLES IN (SINGULAR) INTERACTION : LONG-TIME BEHAVIOR AND PROPAGATION OF CHAOS

Abstract

This thesis is devoted to the study of certain systems of N particles in mean-field interaction, and of a specific phenomenon : in such systems, when $N \to \infty$, two given particles become "more and more" independent. This property is named *propagation of chaos*, and our aim is to prove it in several settings. We focus on three main cases: the kinetic one (i.e. with degenerate noise), the one with singular interactions, and the one with incomplete interactions. In each case, we seek to obtain quantitative and uniform in time results. We start by setting up a coupling method to prove the long time convergence of the *Vlasov-Fokker-Planck* equation describing the limit of particles in Lipschitz interactions and confined by a non-convex potential. The coupling method is then adapted to prove the propagation of chaos property for this system, as well as for the *FitzHugh-Nagumo* model describing neurons in the brain interacting in a mean field way. We then focus on a few particle systems in *Riesz* interactions. The first one is the *2D vortex* model, for which we prove uniform in time propagation of chaos using entropic methods. We then study a one-dimensional singular system, motivated by the *Dyson Brownian motion* derived from the study of random matrices, for which we prove this same phenomenon by a new coupling. Finally, we show the uniform in time mean field limit for a system of particles *interacting according to a graph*, random or not, before turning our attention to a method of numerical simulation of interacting particles. In particular, we study the *Random Batch Method*, and its effect on the phase transition that may exist for the nonlinear limit of the particle system. To this end, we look successively at the *Curie-Weiss* model and the double-well model for the overdamped Langevin equation.

Keywords: probability theory, stochastic calculus, coupling methods, propagation of chaos, logarithmic sobolev inequality.

Résumé

Cette thèse est consacrée à l'étude de certains systèmes de N particules en interaction en champ moyen, et d'un phénomène particulier : dans de tels systèmes, lorsque $N \to \infty$, deux particules données deviennent "de plus en plus" indépendantes. Cette propriété est appelée *propagation du chaos*, et notre objectif est de la prouver dans plusieurs contextes. Nous nous concentrons sur trois cas principaux : le cas cinétique (c'est-à-dire avec un bruit dégénéré), celui des interactions singulières et celui des interactions incomplètes. À chaque fois, nous cherchons à obtenir des résultats quantitatifs et uniformes en temps. Nous commençons ainsi par mettre en place une méthode de couplage afin de prouver la convergence en temps long de l'équation de *Vlasov-Fokker-Planck* décrivant la limite de particules en interaction lipschitzienne et confinées via un potentiel non convexe. Cette méthode est ensuite adaptée pour prouver la propriété de propagation du chaos pour ce système, ainsi que pour le modèle de *FitzHugh-Nagumo* décrivant des neurones dans le cerveau interagissant en champ moyen. Nous nous intéressons ensuite à quelques systèmes de particules en interaction de type *Riesz*. Le premier est le modèle de *vortex 2D*, pour lequel nous prouvons la propagation du chaos uniforme en temps en utilisant des méthodes entropiques. Nous étudions ensuite un système singulier en dimension 1, motivé par le *mouvement brownien de Dyson* provenant de l'étude de matrices aléatoires, pour lequel nous prouvons ce même phénomène par un nouveau couplage. Enfin, nous montrons la limite de champ moyen uniforme en temps pour un système de particules *interagissant selon un graphe*, aléatoire ou non, avant de nous intéresser à une méthode de simulation numérique de particules en interaction. En particulier, nous étudions la *Random Batch Method*, et son effet sur la transition de phase qui peut exister pour la limite non linéaire du système de particules. Pour cela, nous regardons successivement le modèle de *Curie-Weiss* et le modèle double-puits pour l'équation de Langevin sur-amortie.

Mots clés : probabilités, calcul stochastique, méthodes de couplage, propagation du chaos, inégalité de sobolev logarithmique.

Laboratoire Jacques-Louis Lions
Sorbonne Université – Campus Pierre et Marie Curie – 4 place Jussieu – 75005 Paris – France

Contents

List of Figures

Remerciements

> Un critique de film disais-je donc avant de m'ensabler dans les méandres sournois de mes aigreurs égarées entre deux virgules si éloignées du début de ma phrase que voilà-t-y pas que je sais plus de quoi que je cause.
>
> Pierre Desproges, *Chronique de la haine ordinaire : Criticon* (1986).

J'ai passé beaucoup de temps sur cette partie. Je ne garantis pas le résultat, mais j'y ai passé beaucoup de temps.

Il y a plusieurs raisons à cela.

La première, la plus pragmatique, provient de l'observation suivante : vous êtes probablement en train de lire ces mots pendant ma soutenance, et vous êtes vraiment en train de vous ennuyer. Un manuscrit traîne sur une table, ou quelqu'un vient de vous le passer, et vous l'avez ouvert à l'une des seules pages à peu près lisibles. On est en juin, il fait chaud, vous êtes en train de subir une présentation de probabilités en anglais par un rouquin qui galère au tableau, mais par politesse vous restez. Donc vous cherchez désespérément autre chose à faire. Je comprends. J'ai hésité à vous proposer un sudoku, mais j'ai finalement opté pour un petit texte à lire.

La deuxième raison, plus personnelle, est lié à mon incapacité à dire merci. J'ai tendance à en faire trop ou, au contraire, pas assez. J'aurais ici pu rester sobre[1] parce que, soyons honnêtes, ce n'est qu'une thèse parmi d'autres. Mais bon, quand même, c'est une thèse.

Alors, après réflexion, je tiens à faire le point, et à ce qu'il y ait une trace écrite, que ceux qui lisent ce manuscrit[2] sachent à qui ils le doivent. Au programme donc : la phrase la plus longue que j'ai jamais écrite, des fautes de grammaire, un brin de mégalomanie et surtout un festival de violons. Accrochez-vous, parce que je vais sûrement trop en faire, mais je vais le faire comme j'ai envie de le faire, et au pire je passe pour un con, mais ça, à force, je sais faire.

Bref, tout ça pour solennellement vous annoncer que je voudrais dire un petit mot :

À François Delarue et Martin Hairer, qui me font l'honneur de rapporter cette thèse, pour leurs retours et leurs remarques ; à Nicolas Fournier, Sylvie Méléard et Sylvia Serfaty, qui ont accepté de faire partie de mon jury ; et à Thierry Bodineau, qui a en plus bien voulu m'accueillir l'année prochaine ;

Aux professeurs qui m'ont accompagné avant et pendant la thèse ; à Nizar Touzi qui m'a donné l'envie de faire du calcul stochastique et qui m'a aiguillé dans mes choix de stage et de

[1] certains auraient même pu dire *rigide*.

[2] j'ai le droit de rêver...

doctorat ; à Djalil Chafaï, pour ses conseils et le temps qu'il m'a accordé au cours de ces trois années ; à Arnaud Guyader qui, en me recevant dans son bureau, a débloqué et mis en route ma recherche de thèse, et qui m'a ensuite régulièrement conseillé tout au long de celle-ci, que ce soit lors de nos rencontres informelles ou lors de mon comité de suivi, et à Irina Kourkova qui a aussi accepté de faire partie de ce dernier ;

À celles et ceux avec qui j'ai eu l'occasion de travailler, pour leur temps et leur aide ; à Laetitia Colombani avec qui ça a été un plaisir d'écrire le chapitre sur les neurones, à Patrick Cattiaux et Manon Costa pour leur encadrement dans cette entreprise, et à Samir Salem pour son aide ; enfin à Christophe Poquet pour être venu avec ce problème sur les graphes ;

À mes directeurs de thèse, Arnaud Guillin et Pierre Monmarché, pour leur encadrement, leurs conseils, leur bienveillance et leur patience, leurs réponses à mes mails de panique à 23h, la porte du bureau toujours (métaphoriquement) ouverte, pour m'avoir envoyé en conf, pour m'avoir donné l'occasion de parler, et pour les deux personnes incroyables qu'ils sont ; ce n'est pas la fin de nos travaux ensemble, en tout cas je l'espère, mais je tiens à vous dire que si j'ai passé trois superbes années, c'est en grande partie grâce à vous ;

Au LJLL qui m'a accueilli, à ses permanents, et aux secrétariats et services informatique de Jussieu et de Clermont-Ferrand qui m'ont donné tous les outils dont j'avais besoin pour cette thèse ;

Aux doctorants et doctorantes du LJLL qui ont contribué à l'ambiance et à la bonne humeur qui me donnaient envie de me lever le matin ; à Antoine L.[3], ses nombreuses activités au sein et en dehors du labo, et ses incroyables dîners ; à Charles, Guillaume et Lucas P. qui ont repris et amélioré le GTT, et surtout compris comment fonctionnait ce site web malgré mon aide inexistante ; et désormais à Ludovic et Zhe, à qui je souhaite beaucoup de courage ; au maté d'Agustín et au sel de Chourouk ; au spectacle d'Anatole et à la boîte à meuh de Thomas B. ; aux chocolats de Ruikang et au lièvre de Sylvain ; à Fabrice et sa quantique ; à Jules P., pour l'ensemble de son oeuvre, ses conseils, son engagement, ses débats ; à Juliette, qui a finalement réservé son train demain ; à Lucas E., sans qui je serais probablement encore de train de chercher un postdoc ; à la respo bureau qui travaille trop pendant les vacances, Liangying ; à María, camarade de GTT ; à Mingyue pour sa bonne humeur ; à Nga, sur qui je compte énormément pour me sauver de ces saletés de moustiques ; à Noemi, pour m'avoir appris la seule danse que je connais[4] ; à Rémi et ses montagnes ; à Robin, que j'ai probablement plus croisé en bas des tours qu'en haut, pour sa générosité[5] ; à Roxane, pour son engagement ; au thé de Rui, au futur thé d'Eleanor, et au TP d'Analyse numérique de Pauline ; à Jesús, même si l'Espagne a gagné contre la France ce triste 18 septembre ; à la team CROUS 1145T25, qui a su imposer son règne, et à son fidèle pilier Yipeng ; à Ramon donc, dont le départ a signé la fin de la team 1230 ; à Nicolas T. (¡Io no hablo español !) ; à Yvonne la plus française des canadiennes (une québécoise donc) ; et à toutes celles et ceux que j'ai pu croiser au labo ou en salle café, dans un ordre trombinoscopique, Gontran, Nicolas C., Allen, Willy, David, Alexandre R., Gong, Emma, Mathieu R., Alexiane, Emilio, Jules G., Antoine LC., Ioanna-Maria, Nathalie, Kala, Cristobal ;

Plus précisément ensuite au bureau 15-16-303 et ces illustres anciens, Alexandre P. pour son aide précieuse en début de thèse, et Liudi pour son fun ; à Giorgia, la première personne que j'ai rencontrée au labo, qui m'a assigné mon bureau ; à celui qui a commencé en même temps que moi mais un an avant et un an après, Lucas J., et dont je parlerai un peu plus bas ; puis à ceux qui sont arrivés après ; à Assane, à notre guerre du thermostat, mais surtout au petit éléphant que je garde précieusement ; à Charles qui a subi ma pression jusqu'à ce qu'il craque et accepte

[3] Antoine LB., mais en un seul mot, pas comme Pierre LB., et à ne pas confondre avec Antoine LC. donc, vous suivez ?

[4] et pour m'avoir remis cette chanson dans la tête rien qu'à l'écriture de ces mots. . .

[5] challenge accepted 16/12/22

le GTT, qui illumine le labo de son rire (trop) fort et franc, et qui a réussi l'exploit d'avoir un bureau plus bordélique que le mien ; à Toai, la libellule, la banquise, le foot, les échecs et les siestes ; à Victor, qui nous manque sur les photos :'(et au dernier entrant, Laurent, auprès de qui je m'excuse de vouloir laisser la porte du bureau ouverte[6] ; grâce à vous, j'ai passé trois super années, plus ou moins studieuses, et si un jour vous cherchez comment effacer ce tableau blanc de l'enfer, je n'ai qu'une chose à vous dire : *vous connaissez Google ?* ;

Aux thésards en interaction ; à la grande soeur d'abord, Lise, qui a traversé ce terrain miné qu'est la thèse pour nous prévenir des pièges ; à Lucas J., camarade du quotidien, compagnon de conf, qui a régulièrement corrigé mes absurdités probabilistes tout en me régalant d'un humour si... spécial ; et au petit dernier enfin, Nicolai, pour sa bonne humeur, son amour de la SF et sa présence[7] ;

À mes amis ; à Antoine, qui pendant deux ans de colocation a subi une autre forme de propagation du chaos ; à Cécile et Christian, heureusement indissociables, pour les bières et bien plus ; à François, pour son encadrement lors du PSC ; à Inès qui, même si ça me saoule de l'admettre, n'avait pas complètement tort ; à Matthieu[8], que j'aurais pu distinguer plus haut, mais dont la place est ici bas ; à Samy et ses plans (pas si) foireux ; à Sophie, pour sa gentillesse en toute circonstance ; au groupe de khôlle de l'espace, Moïse et Randy, à qui je souhaite bonne chance pour dans respectivement un an et un jour, car pour la première fois de l'histoire je finis (en) premier ; et enfin à Thomas, compagnon de baston et grand pratiquant de l'enchainement step-cross, pour l'ensemble des balayettes ;

À ma môman et mon pôpô pour leur présence, leurs conseils, et tout le temps qu'ils ont investi à enseigner les maths à un idiot comme moi[9] ; et à mon frère, l'idiot comme moi[10] ; je sais que je ne vous le dis pas souvent, parce qu'à chaque fois ça me semble évident, mais vous êtes le point (0,0,0,0) de mon repère espace-temps autour duquel je dois graviter ;

À mes grands parents ;

A te, per la cosa che avevo perso ;

À ceux qui sont partis, et surtout à ceux qui sont restés ;

Merci !

PS : Voici quand même un sudoku

	1				4		6	
	8		5				4	2
			3	8		1		
		2			5			4
	7			3			8	
	3							
	9		6	5			7	3
7			9					
		5		7	8			6

[6] Bon week-end d'ailleurs ;)

[7] même quand il n'est pas au labo "parce qu'aujourd'hui il travaille"...

[8] Docteur de Sorbonne Université

[9] $\sqrt{a}$ est le nombre positif qui mis au carré donne a, $\sqrt{a}$ est le nombre positif qui mis au carré donne a, $\sqrt{a}$ est le nombre positif qui mis au carré donne a,...

[10] mais pas que !

Résumé succinct en français

Comment ça merde alors ? But alors you are French ?

Louis de Funès, *La Grande Vadrouille* (1966), ~~written by~~ écrit par Georges et André Tabet.

This is the only chapter in French.

Nous étudions dans cette thèse des systèmes de particules en interaction. En particulier, nous cherchons à prouver dans plusieurs cas le phénomène suivant :

Dans un système de N particules en interaction, lorsque $N \to \infty$, deux particules données deviennent "de plus en plus" indépendantes.

Ce phénomène a été nommé *propagation du chaos* : le *chaos* fait bien entendu référence à la propriété d'indépendance, tandis que le terme *propagation* souligne le fait qu'il suffira de montrer l'indépendance à la limite $N \to \infty$ au temps $t = 0$ pour qu'elle soit également vraie à la limite à un temps ultérieur $t > 0$.

Le système typique de particules est décrit par un système d'équations différentielles stochastiques (EDS) de la forme suivante

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^{N} K(X_t^{i,N} - X_t^{j,N})dt + \sqrt{2\sigma_N} dB_t^i.$$

Chaque particule est représentée par une quantité $X \in \mathbb{R}^d$, généralement sa position, où $X_t^{i,N}$ désigne la position au temps t de la i-ème particule. K est une fonction (nous parlerons de noyau d'interaction), σ_N est un coefficient de diffusion qui peut dépendre ou non du nombre total de particules, et les $(B^i)_i$ sont des mouvements browniens indépendants en dimension d.

Comme mentionné, nous souhaitons comprendre la limite lorsque N tend vers l'infini de ce système. En notant $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ la mesure empirique, nous pouvons réécrire le système d'EDS comme suit

$$dX_t^{i,N} = K * \mu_t^N \left(X_t^{i,N}\right) dt + \sqrt{2\sigma_N} dB_t^i,$$

où $*$ désigne l'opération de convolution : $K*\mu(x) = \int K(x-y)\mu(dy)$. Si l'on s'attend effectivement à ce que les particules deviennent indépendantes à la limite $N \to \infty$, ainsi que identiquement distribuées, on devine que μ_t^N convergera vers une mesure $\bar\rho_t$, la loi d'une particule typique à la limite. De là, très formellement, un candidat naturel pour l'EDS limite

$$\begin{cases} d\bar X_t = K * \bar\rho_t(\bar X_t)dt + \sqrt{2\sigma}dB_t, \\ \bar\rho_t = \text{Law}(\bar X_t), \end{cases}$$

où $\sigma = \lim \sigma_N$. Cette EDS non linéaire est dite de type *McKean-Vlasov* en raison de la non linéarité induite par l'interaction avec sa propre loi.

Notons $\rho_t^{k,N} = \text{Loi}\left(X_t^{1,N}, ..., X_t^{k,N}\right)$ la loi jointe du sous-ensemble des k premières particules du système de N particules (avec la convention $\rho_t^N = \rho_t^{N,N}$) et $\bar\rho_t^{\otimes k} = \bar\rho_t \otimes .... \otimes \bar\rho_t$ la loi non linéaire limite $\bar\rho_t$ tensorisée k fois.

Notre but est ainsi de montrer un résultat de la forme

$$\textbf{Propagation du chaos :} \quad \lim_{N\to\infty} \rho_t^{k,N} = \bar\rho_t^{\otimes k}, \forall k \in \mathbb{N}, \forall t \geqslant 0, \text{ si cela est vrai pour } t = 0,$$

ou, de façon équivalente, (voir par exemple Proposition 2.2 de [162])

$$\textbf{Limite de champ moyen :} \quad \lim_{N\to\infty} \mu_t^N = \bar\rho_t, \forall t \geqslant 0, \text{ si cela est vrai pour } t = 0.$$

Nous ne précisons pas à ce stade le sens donné aux limites mentionnées ci-dessus.

Le document s'articule autour de trois grands axes, correspondant à des variations autour du système de particules donné ci-dessus.

- Le cas cinétique (ou bruit dégénéré) : le mouvement brownien n'agit pas sur toutes les coordonnées.
- Le cas singulier : le noyau K est singulier en 0.
- Avec des interactions incomplètes : les particules n'interagissent plus avec toutes les autres particules.

Partie I : Le cas cinétique (ou bruit dégénéré). Le premier travail [83] concerne la preuve de la propagation du chaos uniforme en temps pour un système de particules de la forme

$$\begin{cases} dX_t^{i,N} = V_t^{i,N}dt \\ dV_t^{i,N} = \sqrt{2\sigma}dB_t^i - V_t^{i,N}dt - \nabla U(X_t^{i,N})dt - \frac{1}{N}\sum_{j=1}^N \nabla W(X_t^{i,N} - X_t^j)dt, \end{cases}$$

où $X_t^{i,N}$ et V_t^i sont respectivement la position et la vitesse de la particule i dans $\mathbb{R}^d$, U est un potentiel de confinement non convexe tel que ∇U est lipschitzienne, et W est un potentiel d'interaction tel que ∇W est lipschitzienne avec un coefficient de Lipschitz suffisamment petit.

Nous considérons pour cela une méthode de couplage suggérée par A. Eberle [66], connue sous le nom de couplage *par réflexion*, conçue à l'origine pour traiter le comportement en temps long des processus de diffusion généraux, comme dans [66, 69], et étendue plus tard pour montrer la propagation du chaos uniforme en temps dans un système en champ moyen dans [64]. Cette méthode de couplage repose sur la construction minutieuse d'une semi-métrique prenant en compte les différents comportements des particules.

Cela correspond au Chapitre 2.

En utilisant les mêmes idées, il est également possible de prouver la propagation du chaos uniforme en temps pour un système de neurones dans le cerveau suivant le modèle de *FitzHugh-Nagumo*

$$\begin{cases} dX_t^{i,N} = (X_t^{i,N} - (X_t^{i,N})^3 - C_t^{i,N} - \alpha)dt + \frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N} - Z_t^{j,N}) + \sigma_X dB_t^{i,X} \\ dC_t^{i,N} = (\gamma X_t^{i,N} - C_t^{i,N} + \beta)dt + \frac{1}{N}\sum_{j=1}^N K_C(Z_t^{i,N} - Z_t^{j,N}) + \sigma_C dB_t^{i,C}, \end{cases}$$

où $X_t^{i,N}$ est la potentiel de membrane de la particule i et $C_t^{i,N}$ est une variable dite de récupération et où nous permettons à σ_X ou σ_C d'être égal à 0. Une fois encore, K_X et K_C sont des noyaux d'interaction supposés lipschitziens avec une constante de Lipschitz suffisamment petite. Au-delà du résultat concernant les réseaux de neurones, qui est en soi intéressant, ce travail témoigne surtout de la robustesse de la méthode de couplage mise en place dans le Chapitre 2.

Il s'agit d'un travail en collaboration avec Laetitia Colombani[11] [54], résumé dans le Chapitre 3.

Partie II : Le cas singulier. Le premier modèle singulier auquel nous nous intéressons est le modèle *vortex 2D*. Considérons le système de particules dans le tore en dimension 2 noté $\mathbb{T}^2$

$$dX_t^{i,N} = \frac{1}{N}\sum_{j\neq i} K\left(X_t^{i,N} - X_t^{j,N}\right)dt + \sqrt{2\sigma}dB_t^i,$$

pour le noyau de Biot-Savart $K(x) = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$. La limite non linéaire satisfait l'équation de vorticité, qui apparaît lorsqu'on considère le rotationnel de l'équation en 2D de Navier-Stokes incompressible.

Dans leur récent article [99], P.-E. Jabin et Z. Wang ont prouvé, quantitativement, la propagation du chaos pour ce système sur le tore. En s'appuyant sur leur travail, nous obtenons un résultat similaire dans [85], mais désormais uniforme en temps.

L'approche consiste à calculer l'évolution temporelle de l'entropie relative de ρ_t^N par rapport à $\bar\rho_t^{\otimes N}$, puis à utiliser une intégration par parties pour gérer la singularité de K grâce à la régularité de la densité de probabilité $\bar\rho_t$. Cette idée vient de l'observation que le noyau de Biot-Savart peut être explicitement écrit comme la divergence d'un champ matriciel borné. Afin d'améliorer cet argument pour obtenir une propagation du chaos uniforme en temps, notre principale contribution est la preuve de bornes pour la densité de $\bar\rho$, dont on déduit une inégalité de Sobolev logarithmique. De cette dernière, dans l'esprit des travaux de F. Malrieu [133] dans le cas lisse et convexe, l'information de Fisher apparaissant dans la dissipation d'entropie donne un contrôle sur l'entropie relative elle-même, induisant l'uniformité temporelle. Grâce à la décroissance rapide vers 0 des dérivées de la limite non linéaire, aucune hypothèse de petitesse sur l'interaction n'est requise. Nous prouvons ainsi l'uniformité en temps de la propagation du chaos sous un ensemble d'hypothèses satisfaites en particulier par le noyau de Biot-Savart.

La méthode et les résultats sont discutés dans le Chapitre 4.

Nous nous intéressons ensuite au cas du *mouvement brownien de Dyson (généralisé)* en dimension 1

$$dX_t^{i,N} = \sqrt{\frac{2\sigma}{N}}dB_t^i - \lambda X_t^{i,N}dt + \frac{1}{N}\sum_{j\neq i}\frac{1}{X_t^{i,N} - X_t^{j,N}}dt.$$

[11] University of Bern, Institute of Mathematical Statistics and Actuarial Science, CH-3012 Bern Switzerland. Email : laetitia.colombani[AT]stat.unibe.ch

La dynamique ci-dessus est satisfaite, pour $\lambda = 0$, par les valeurs propres d'un mouvement brownien dans l'espace des matrices symétriques $N \times N$, comme l'a observé F. J. Dyson en 1962 [65]. Pour $\lambda > 0$, elle correspond aux valeurs propres d'un processus d'Ornstein-Uhlenbeck à valeur dans l'espace des matrices symétriques $N \times N$ [47]. Ceci nous incite à considérer, dans un cadre légèrement plus général, le système 1D à N particules en champ moyen

$$dX_t^{i,N} = \sqrt{2\sigma_N} dB_t^i - \lambda X_t^{i,N} dt - \frac{1}{N} \sum_{j \neq i} V'(X_t^{i,N} - X_t^{j,N}) dt,$$

où $\sigma_N \xrightarrow[N\to\infty]{} 0$ et il existe $\alpha \in [1,2[$, tel que pour tout $x \in \mathbb{R}^*$, $V'(x) = -\frac{x}{|x|^{\alpha+1}}$.

Nous obtenons un résultat quantitatif uniforme en temps pour $\alpha \in [1,2[$ en couplant un système de N particules avec un système de M particules pour tous $N, M \geqslant 0$. Nous réussissons à prouver, en utilisant pleinement le fait qu'en dimension 1 les particules restent ordonnées et que le couplage optimal dans la distance de Wasserstein est explicite, que toute suite indépendante de mesures empiriques est une suite de Cauchy dans un certain sens. Ensuite, l'indépendance assure que la limite est une variable aléatoire presque sûrement constante dans l'espace des mesures de probabilité, qui peut alors être identifiée comme une solution faible de la limite non linéaire à laquelle nous nous attendons.

Cette méthode, qui, à notre connaissance, n'avait pas été utilisée auparavant, est décrite en détail dans [84], et semble jusqu'à présent limitée au cas de la dimension 1. Le Chapitre 5 comprend la description complète des résultats.

Partie III : Avec des interactions incomplètes. Enfin, nous considérons les cas où les particules n'interagissent pas avec toutes les autres particules, mais selon un graphe. Ceci rompt l'hypothèse d'échangeabilité du système, car *a priori* certaines peuvent interagir avec plus que d'autres et la structure du graphe sous-jacent dans de tels cas influence grandement le comportement du système.

Premièrement, nous remarquons que le couplage par réflexion permet de traiter des particules interagissant selon un graphe

$$dX_t^{i,N} = F\left(X_t^{i,N}, \omega_i\right) dt + \frac{\alpha_N}{N} \sum_{j=1}^{N} \xi_{i,j}^{(N)} \Gamma\left(X_t^{i,N}, \omega_i, X_t^{j,N}, \omega_j\right) dt + \sqrt{2\sigma} dB_t^i,$$

où $\xi^{(N)} = \left(\xi_{i,j}^{(N)}\right)_{1\leqslant i,j\leqslant N}$, $\xi_{i,j}^{(N)} \in \{0,1\}$ représente le graphe, $(\omega_i)_{1\leqslant i\leqslant N}$ est un désordre environnemental, $(\alpha_N)_{N\geqslant 1}$ est un paramètre d'échelle, $F : \mathbb{R}^d \times \mathbb{R}^{d'} \mapsto \mathbb{R}^d$ est une force extérieure lipschitzienne et $\Gamma : \left(\mathbb{R}^d \times \mathbb{R}^{d'}\right)^2 \mapsto \mathbb{R}^d$ est une interaction lipschitzienne. En supposant qu'il existe $p \in [0,1]$ tel que $\sup_{1\leqslant i\leqslant N} \left|\alpha_N \frac{\sum_{j=1}^N \xi_{i,j}^{(N)}}{N} - p\right| \xrightarrow[N\to\infty]{a.s} 0$, ce qui est vrai pour les graphes de Erdős-Rényi, la distribution empirique μ_t^N converge en distance de Wasserstein vers la limite non linéaire $\bar{\rho}_t$.

Il s'agit d'un travail conjoint avec Christophe Poquet[12] [115] disponible dans le Chapitre 6.

Ensuite, nous considérons le schéma numérique associé au système de particule en interaction

[12]Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, F-69622 Villeurbanne, France. Email : poquet[AT]math.univ-lyon1.fr

en champ moyen pour un pas de temps δ, qui s'écrit

$$\begin{cases} X_{k+1}^{i,N} = X_k^{i,N} - \delta \nabla U(X_k^{i,N}) - \frac{\delta}{N-1} \sum_{j \neq i} \nabla W(X_k^{i,N} - X_k^{j,N}) + \sqrt{2\sigma\delta} G_k^i, \\ G_k^i \quad \text{i.i.d} \quad \sim \mathcal{N}(0,1). \end{cases}$$

Remarquez que la complexité numérique de chaque pas de temps est de $O(N^2)$, correspondant au calcul de chaque interaction. Pour diminuer cette complexité, nous choisissons de considérer une méthode appelée *Random Batch Method*. L'idée est qu'au lieu de considérer que chaque particule interagit avec toutes les autres particules, nous divisons à chaque pas de temps le système en "lots" plus petits et nous calculons les interactions uniquement au sein de ces lots. Considérons, pour un pas de temps k, une partition $\mathcal{P}_k = \left(\mathcal{P}_k^1, ..., \mathcal{P}_k^{N/p}\right)$ de $\{1, ..., N\}$ en N/p sous-ensembles de taille p, et définissons

$$\mathcal{C}_k^i = \left\{ j \in \{1, .., N\} \text{ s.t. } \exists l \in \{1, ..., N/p\}, i, j \in \mathcal{P}_k^l \right\}.$$

En d'autres termes, $\mathcal{C}_k^i$ est l'ensemble des indices qui sont dans le même sous-ensemble que i au pas de temps k, avec la convention $i \in \mathcal{C}_k^i$. La partition est choisie aléatoirement et uniformément à chaque pas de temps, et donne le schéma numérique suivant

$$\begin{cases} Y_{k+1}^i = Y_k^i - \delta \nabla U(Y_k^i) - \frac{\delta}{p-1} \sum_{j \neq i \in \mathcal{C}_k^i} \nabla W(Y_k^i - Y_k^j) + \sqrt{2\sigma\delta} G_k^i, \\ G_k^i \quad \text{i.i.d} \quad \sim \mathcal{N}(0,1). \end{cases}$$

Le schéma numérique est maintenant de complexité $O(Np)$ pour chaque pas de temps, et la convergence de ce schéma, lorsque N tend vers l'infini et δ tend vers 0, vers la distribution invariante de l'EDS limite non linéaire peut par exemple être trouvée dans [177].

Il est intéressant de noter que cette méthode "ajoute" de l'aléatoire au système. Cela nous incite à analyser l'effet de la *Random Batch Method* sur les systèmes qui présentent une transition de phase puisque, si le système admet une température critique, cette température devrait intuitivement diminuer en raison de l'ajout d'aléa.

Pour mettre en évidence ce phénomène, nous étudions deux modèles-jouets : les limites non linéaires du modèle de *Curie-Weiss* et du système de particules dans un potentiel de confinement double-puits. Ces deux modèles présentent une transition de phase : pour une température (ou, de manière équivalente, un coefficient de diffusion) inférieure à un paramètre critique, il existe trois états stationnaires, et au-delà de ce paramètre critique, il n'en existe qu'un seul. Nous étudions l'évolution de cette température critique en fonction de la taille des "lots".

Il s'agit d'un travail en cours, dont vous pouvez trouver l'état d'avancement dans le Chapitre 7.

Chapter 1

Introduction

C'est curieux chez les marins ce besoin
de faire des phrases...

Francis Blanche, *Les Tontons flingueurs*
(1963), written by Michel Audiard.

Outline of the current chapter

1.1 Model(s) and motivation(s)

1.1.1 A matter of scale

When describing gases, several points of view are possible. The first and obvious one : a *macroscopic* point of view. There, one would be describing a gas through observable quantities such as its temperature or its pressure. But a gas is also a large system of interacting particles. One could therefore try and apply laws of motion to each individual particle, and thus take a *microscopic* point of view. This, however, quickly leads to several issues, a major one being that there are *a lot* of particles. Let us recall that the Avogadro number, which describes the number of molecules in thirty-two grams of oxygen, is around $\mathcal{N}_A \simeq 6 \times 10^{23}\text{mol}^{-1}$. Keeping track of each individual particle thus seems infeasible. Hence an "in-between" point of view : a *mesoscopic* one. The goal is no longer to follow each particle, but to give a statistical description of their behavior and study their probability distribution.

In 1872 Ludwig Boltzmann (1844-1906) thus derived, in a landmark article in the field of kinetic theory of gases [23], an equation that now bears his name and that describes the time evolution of the probability density of a typical particle within a gas of identical particles subject to elastic collisions. This equation allowed for the proof of the famed *H-theorem*, building upon the work of James Clerk Maxwell (1831-1879), which states that the entropy of a gas can only increase with time. Such a result however lead to much criticism[1], notably from Johann Josef Loschmidt (1821-1895), who argued that this theorem was in direct contradiction with Newton's law of motion, as the latter yields reversible processes (and therefore, reversing time, there should exist processes for which the entropy decreases).

The explanation for this seemingly intricate paradox lies in a crucial assumption by Boltzmann known as *molecular chaos* : the particles constituting the gas are assumed to be independent. However this trivially cannot hold true for a fixed number of particles, as after each collision between two particles, they are no longer uncorrelated. The rigorous justification of this independence became an important mathematical question in 1900 when David Hilbert (1862-1943) published a list of twenty-three (at the time) unsolved problems. Among these problems, the sixth stated that *"Boltzmann's work on the principles of mechanics suggests the problem of developing mathematically the limiting processes, there merely indicated, which lead from the atomistic view to the laws of motion of continua"* [82].

Among the various works sparked by this problem, let us focus on Mark Kac's (1914-1984) [106], whose goal was to derive Boltzmann's equation. He introduced the notion of *propagation of chaos*, which amounts to rigorously proving that, in a system of N particles interacting through collisions, as N goes to infinity, two particles become statistically independent, hence molecular chaos. Soon after Henry McKean (1930-) showed that this property actually holds for various diffusion models with other forms of interaction [135].

This property of independence at the limit nowadays finds application in fields other than statistical physics, and we will mention some of them throughout the document.

1.1.2 The two points of view

Let us start by very formally mentioning the various objects we will consider. The toy model for a particle system is described by a system of stochastic differential equations (SDE) of the form

$$dX_t^{i,N} = \frac{1}{N}\sum_{j=1}^{N} K(X_t^{i,N} - X_t^{j,N})dt + \sqrt{2\sigma_N}dB_t^i. \tag{1.1.1}$$

Each particle is represented through a quantity $X \in \mathbb{R}^d$, most likely its position, where $X_t^{i,N}$ denotes the position at time t of the i-th particle. K is at this stage an unspecified function (we will speak of an interaction kernel), σ_N is a diffusion coefficient that may or may not depend on the total number of particles, and $(B^i)_i$ are independent d-dimensional Brownian motions. We may sometimes refer to (1.1.1) as a type of *overdamped Langevin equation*, in contrast to the

[1] See [34] for an historical background on the H-theorem.

classical Langevin equation (see (1.3.1) later). Because of this scaling $\frac{1}{N}$ and the fact that each X^i interacts with every other X^j, the particles are said to be in mean-field interaction. A crucial assumption is the exchangeability of the particles.

Definition 1.1.1 (Exchangeable family). *A family $(X^i)_{i\in I}$ of random variables is said to be exchangeable when the law of $(X^i)_{i\in I}$ is invariant under every permutation of a finite number of indexes $i \in I$.*

One can easily show via Itô's calculus that ρ_t^N, the joint law of $(X_t^{1,N}, ..., X_t^{N,N})$, is a weak solution to the following partial differential equation (PDE)

$$\begin{aligned}\partial_t \rho_t^N(x_1, ..., x_N) = -\sum_{i=1}^N \nabla_i \cdot \left(\left(\frac{1}{N} \sum_{j=1}^N K(x_i - x_j) \right) \rho_t^N(x_1, ..., x_N) \right) \\ + \sigma_N \sum_{i=1}^N \Delta_i \rho_t^N(x_1, ..., x_N). \qquad (1.1.2)\end{aligned}$$

We will consider several variations of this toy model, either by adding an additional force on each particle (translating the effect of a confining potential), by considering a kinetic setting (in which case each particle is described through both its position and its velocity), or by considering different forms of interaction kernel K.

As mentioned, we wish to understand the limit as N goes to infinity of this system. Denoting $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ the empirical measure, we may rewrite (1.1.1) as

$$dX_t^{i,N} = K * \mu_t^N \left(X_t^{i,N} \right) dt + \sqrt{2\sigma_N} dB_t^i,$$

where $*$ denotes the operation of convolution : $K * \mu(x) = \int K(x-y)\mu(dy)$. If we indeed expect the particles to be independent at the limit $N \to \infty$ and identically distributed (because of the exchangeability), we guess μ_t^N will converge towards a measure $\bar{\rho}_t$, the law of one typical particle at the limit. Hence, very formally, a natural candidate for the limit SDE

$$\begin{cases} d\bar{X}_t = K * \bar{\rho}_t(\bar{X}_t) dt + \sqrt{2\sigma} dB_t, \\ \bar{\rho}_t = \text{Law}(\bar{X}_t), \end{cases} \qquad (1.1.3)$$

where $\sigma = \lim \sigma_N$. This nonlinear SDE is said to be of *McKean-Vlasov* type because of the nonlinearity induced by the interaction with its own law. We may also consider its related Fokker-Planck equation

$$\partial_t \bar{\rho}_t(x) = -\nabla \cdot ((K * \bar{\rho}_t(x))\bar{\rho}_t(x)) + \sigma \Delta \bar{\rho}_t(x). \qquad (1.1.4)$$

To come back to the initial motivation, (1.1.1) would thus be the microscopic point of view, while (1.1.4) would be the mesoscopic point of view.

Obviously, the solution $(\rho_t^N)_t$ of (1.1.2) takes its values in a space $\mathbb{R}^{dN}$, which tends to an infinite dimension space as N goes to infinity. The main convergence results will therefore not concern ρ_t^N *per se*, but rather its marginals.

Denote $\rho_t^{k,N} = \text{Law}\left(X_t^{1,N}, ..., X_t^{k,N} \right)$ the joint law of the subset of the first k particles within the N particle system (which means, by exchangeability, that $\rho_t^{k,N}$ is the law of *any* subset of k particles), and $\bar{\rho}_t^{\otimes k} = \bar{\rho}_t \otimes ... \otimes \bar{\rho}_t$ the measure $\bar{\rho}_t$ tensorized k times. The informal definition of *propagation of chaos* is the following.

Definition 1.1.2 (Propagation of chaos - Informal)**.** *We say that we have propagation of chaos for the particle system* (1.1.1) *if, provided for all* $k \in \mathbb{N}$ *we have* $\rho_0^{k,N} \xrightarrow[N\to\infty]{} \bar{\rho}_0^{\otimes k}$*, we have*

$$\forall t \geqslant 0, \quad \forall k \in \mathbb{N}, \quad \rho_t^{k,N} \xrightarrow[N\to\infty]{} \bar{\rho}_t^{\otimes k}.$$

We do not yet specify in what sense the measures converge.

Notice that this property yields "independence at the limit", as the joint law converges towards a tensorized law. The *chaos* aspect of this property obviously refers to the independence, while the *propagation* alludes to the fact that it will be sufficient to prove this convergence as N goes to infinity at time $t = 0$ for it to also hold at later times t.

Depending on the type of convergence or the distance between the measures we consider, we may define various types of propagation of chaos (weak or strong). The choice of metric will in fact motivate the method of proof.

Another form of convergence, as N goes to infinity, we might wish to consider concerns the empirical measure of the particle system $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$.

Definition 1.1.3 (Mean-field limit - Informal)**.** *Consider a sequence of deterministic empirical measures* $(\mu_0^N)_{N\geqslant 0}$ *such that* $\mu_0^N \xrightarrow[N\to\infty]{} \bar{\rho}_0$*. We say the mean-field limit holds if for all* $t \geqslant 0$*, denoting* μ_t^N *the empirical measure at time* $t \geqslant 0$ *associated to* μ_0^N *and* $\bar{\rho}_t$ *the solution of* (1.1.4) *with initial condition* $\bar{\rho}_0$*, we have* $\mu_t^N \xrightarrow[N\to\infty]{} \bar{\rho}_t$*.*

Likewise, we do not yet specify in what sense we consider the convergence, but keep in mind that μ_t^N *is a random variable in the space of probability measures.*

The convergences in Definitions 1.1.2 and 1.1.3 can be *quantitative*, in which case we will speak of the rate of convergence in N.

Furthermore, we consider the convergence in N without saying much of the time parameter t. But, a priori, for a distance between $\rho_t^{k,N}$ and $\bar{\rho}_t^{\otimes k}$ to be smaller than a given threshold $\epsilon > 0$, one has to choose an integer N that may depend on t. This leads us to a specific form of propagation of chaos or mean-field limit that we will mention on several occasions throughout this document, and that is a *uniform in time* version of these results, in which the convergence in N is independent of t.

In what follows, we denote $\mathcal{P}(\mathcal{X})$ (resp. $\mathcal{P}_p(\mathcal{X})$) the set of probability measures on a given space $\mathcal{X}$ (resp. probability measures with a p-th moment on $\mathcal{X}$).

1.1.3 Some preliminary remarks

Are the various objects well-defined ? The first question we should answer is whether or not the objects (1.1.1), (1.1.2), (1.1.3) and (1.1.4) exist and are uniquely defined. The answer obviously depends crucially on the kernel K and its properties, and depending on the model we consider, proving the well-posedness of the equations might require some work. We may however, at this point, mention the following classical result.

Lemma 1.1.1. *Consider, given two functions* $b, \sigma : \mathbb{R}^d \times \mathcal{P}\left(\mathbb{R}^d\right) \mapsto \mathbb{R}^d$*, the following d-dimensional McKean-Vlasov SDE*

$$\begin{cases} d\bar{X}_t = b(\bar{X}_t, \bar{\rho}_t)dt + \sigma(\bar{X}_t, \bar{\rho}_t)dB_t, \\ \bar{\rho}_t = Law(\bar{X}_t), \end{cases} \tag{1.1.5}$$

where $(B_t)_t$ is a d-dimensional Brownian motion, and the associated Fokker-Planck equation

$$\partial_t \bar{\rho}_t + \nabla_x \cdot (b(x, \bar{\rho}_t)\bar{\rho}_t) = \frac{1}{2} \sum_{i,j=1}^{d} \partial_{x_i} \partial_{x_j} \left(\sigma\sigma^T_{i,j}(x, \bar{\rho}_t)\bar{\rho}_t \right). \tag{1.1.6}$$

Assume b and σ are globally Lipschitz continuous, in the sense that

$$\begin{aligned} \exists C > 0, \quad \forall x, y \in \mathbb{R}^d, \quad \forall \mu, \nu \in \mathcal{P}_2\left(\mathbb{R}^d\right), \\ |b(x, \mu) - b(y, \nu)| + |\sigma(x, \mu) - \sigma(y, \nu)| \leqslant \quad C\left(|x - y| + \mathcal{W}_2(\mu, \nu)\right), \end{aligned}$$

where $\mathcal{W}_2$ is the L^2-Wasserstein distance (see Definition 1.2.1 below). Assume furthermore $\bar{\rho}_0 \in \mathcal{P}_2\left(\mathbb{R}^d\right)$. Then, for all $T > 0$, there exists a unique strong solution to SDE (1.1.5) *on* $[0, T]$ *and its law is the unique weak solution to the Fokker-Planck equation* (1.1.6)*.*

Remark 1.1.1. *The proof of Lemma 1.1.1 in the case σ constant and $b(x, \mu) = \int_{\mathbb{R}^d} b(x, y)\mu(dy)$, which will often be the case of interest to us, can be found as Theorem 1.1 in [162] and Theorem 2.2 in [137].*

Classically, the assumptions on b and σ given in Lemma 1.1.1 also ensure strong existence and uniqueness of the particle system

$$dX_t^{i,N} = b\left(X_t^{i,N}, \frac{1}{N}\sum_{j=1}^{N} \delta_{X_t^{j,N}}\right) dt + \sigma\left(X_t^{i,N}, \frac{1}{N}\sum_{j=1}^{N} \delta_{X_t^{j,N}}\right) dB_t^i.$$

We also refer to Theorems 2.4 and 2.5 of [112] for a weaker form of existence and uniqueness under more general assumptions on b and σ.

The proof of this lemma is based on a fixed point argument, sketched in Proposition 1 of [45].

On the equivalence of the limits for large number of particles. Let us now mention the fact that Definitions 1.1.2 and 1.1.3 describe, provided the system of particles is exchangeable, the same phenomenon, see for instance Proposition 2.2 of [162]. A quantitative link of this equivalence can be partially found in [91]. The following lemma and its proof, given for the sake of completeness in Appendix A.1, come from [80]. Recall that convergence in law towards a constant is equivalent to the convergence in probability.

Lemma 1.1.2. *Let $f \in \mathcal{P}(\mathbb{R}^d)$ and, for $N \geqslant 1$, let $F_N \in \mathcal{P}(\mathbb{R}^{Nd})$ be a probability measure symmetric in all its variables, i.e F_N is invariant under the transformation $(x_1, ..., x_N) \mapsto (x_{\sigma(1)}, ..., x_{\sigma(N)})$ for all $\sigma \in \mathfrak{S}_N$ the set of permutations of $\{1, ..., N\}$.*

Denoting $(X_1, ..., X_N)$ a random variable distributed according to F_N, the two following statements are equivalent :

1. *for all $\epsilon > 0$ and all bounded continuous function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$,*

$$\mathbb{P}^{F_N}\left(\left|\frac{1}{N}\sum_{i=1}^{N} \phi(X_i) - \int \phi(x)df(x)\right| \geqslant \epsilon\right) \xrightarrow[N\to\infty]{} 0.$$

2. *for all $k \geqslant 0$, the sequence $(F_N^k)_N$ of k-marginal distributions of F_N converges towards $f^{\otimes k}$ for the weak convergence of measures (we also say converges weakly-*), i.e for all bounded*

continuous function $\phi : \mathbb{R}^{kd} \mapsto \mathbb{R}$

$$\mathbb{E}^{F_N^k}\left(\phi(X_1, ..., X_k)\right) \xrightarrow[N\to\infty]{} \int \phi(x_1, .., x_k) df(x_1)...df(x_k).$$

Remark 1.1.2. *By Proposition 4.6 of Chapter 3 of [73], it is sufficient to obtain the convergence weakly-* of the marginals* F_N^k *to consider only tensorized test functions* $\phi = \phi_1 \otimes ... \otimes \phi_k$*, where for all* j *the function* $\phi_j : \mathbb{R}^d \mapsto \mathbb{R}$ *is bounded continuous.*

Furthermore, as can be seen in the proof, it is sufficient to prove the convergence of the k*-marginals for* $k = 1, 2$ *to obtain propagation of chaos.*

Remark 1.1.3. *Alternatively, the equivalence between Definitions 1.1.2 and 1.1.3 can be seen as a consequence of de Finetti and Hewitt-Savage theorem. See for instance Theorems 5.1 and 5.3 of [91].*

The long-time analysis of the processes. In this document, we will also focus on the long-time behavior of (1.1.1) and (1.1.3), i.e the possible convergence as $t \to \infty$. This is clearly an interesting problem in its own right, as it provides insight into the behavior of processes. However, in the context of this document, there may intuitively be an additional link between propagation of chaos and the long-time behavior of the system.

This will become clearer when we describe the methods of proof for propagation of chaos, since some of them are adapted from those for convergence towards a stationary distribution (and this will also be one of the main points of Chapter 2). Furthermore, in Chapter 5, we will also see how *uniform in time* propagation of chaos can be obtained from the long-time convergence of the particle system. Finally, the quantitative results of limits as $N \to \infty$ will often be interpretable as the propagation of the initial distance between μ_0^N and $\bar{\rho}_0$, corrected by a term coming from a form of law of large numbers (as there are interactions between particles) which will vanish as $N \to \infty$. See (1.2.4) later to get a better grasp on this particular remark.

In the meantime, it is easy to remark that propagation of chaos is linked to the stability of (1.1.4). Assume $\sigma_N = \sigma = 0$, meaning that we work in the purely deterministic case. Then for a smooth function f we have

$$\frac{d}{dt}\left(\frac{1}{N}\sum_{i=1}^N f\left(X_t^{i,N}\right)\right) = \frac{1}{N}\sum_{i=1}^N \frac{\nabla f\left(X_t^{i,N}\right)}{N} \cdot \sum_{j=1}^N K\left(X_t^{i,N} - X_t^{j,N}\right)$$

i.e

$$\frac{d}{dt}\int f d\mu_t^N = \int \nabla f \cdot K * \mu_t^N d\mu_t^N,$$

which in particular means that μ_t^N is a weak solution to the PDE (1.1.4) as

$$\partial_t \mu_t^N + \nabla \cdot \left((K * \mu_t^N)\mu_t^N\right) = 0.$$

Remark that, if $\sigma_N > 0$, there is an additional martingale term to the equation above (as well as a diffusion term), coming from the fact that μ_t^N is in this case a random variable because of the Brownian motions.

Assume now that we have a result of stability for (1.1.4) with $\sigma = 0$ of the following form. Let μ_t (resp. ν_t) be a solution of (1.1.4) with initial condition μ_0 (resp. ν_0) and let $d : \mathcal{P}(\mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d) \mapsto \mathbb{R}$ be a distance in $\mathcal{P}(\mathbb{R}^d)$. We suppose there is a constant C such that for all $t \geqslant 0$ we have

$$d(\mu_t, \nu_t) \leqslant e^{Ct} d(\mu_0, \nu_0).$$

Then we directly obtain from this stability result the property of propagation of chaos : μ^N and $\bar{\rho}$ being two solutions of (1.1.4), we have for all $t \geqslant 0$ the convergence $d(\mu_t^N, \bar{\rho}_t) \xrightarrow[N\to\infty]{} 0$ provided $d(\mu_0^N, \bar{\rho}_0) \xrightarrow[N\to\infty]{} 0$.

1.2 Proving propagation of chaos : some methods

In this document, we explore several methods to prove the propagation of chaos property. We thus start by giving a very basic idea of what some of these methods consist in, and what they require to be carried out. The goal is not to do an exhaustive review, but rather to give a flavor of how some of the recent and/or relevant proofs work. The choice of the methods we mention in what follows is motivated by the various comparisons we might wish to do, and many results from this very active field of research are unfortunately left out. We instead refer to the classical courses [162, 137], and to the more recent reviews [98] and [45, 46], as most of the proofs and calculations presented here are quite formal.

In addition, although we classify the various methods into fixed and specific categories, the boundaries between these categories are quite porous, and many works belong to more than one of them.

Let us just assume in the following, for the sake of simplicity, that we work in dimension one, that $\sigma_N = \sigma$ and that K is a Lipschitz continuous kernel, in the sense that

$$\exists L > 0,\ \forall x, y \in \mathbb{R},\ |K(x) - K(y)| \leqslant L|x - y|. \tag{1.2.1}$$

This yields strong existence and uniqueness for the solution of (1.1.1), (1.1.2), (1.1.3) and (1.1.4). We also assume that these solutions have finite p-moments, for p as large as required

$$\exists C_0 > 0,\ \forall t \geqslant 0,\ \forall N \geqslant 1,\ \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N} |X_t^{i,N}|^p\right) \leqslant C_0 \ \text{ and }\ \mathbb{E}\left(|\bar{X}_t|^p\right) \leqslant C_0. \tag{1.2.2}$$

1.2.1 Coupling approach

Historically, one of the first tool used to show propagation of chaos is a probabilistic tool, as used by H. P. McKean (see for instance [135]) and then popularised by A.-S. Sznitman [162], known as a coupling method. It comes from the idea that a natural distance between probability measures is the Wasserstein distance, strongly linked to the theory of optimal transport, and is based on R. L. Dobrushin's inequality in the deterministic case [61].

Definition 1.2.1 (Usual L^p-Wasserstein distances in $\mathbb{R}$)**.** *For μ and ν two probability measures on $\mathbb{R}$, denote by $\Pi(\mu, \nu)$ the set of couplings of μ and ν, i.e. the set of probability measures Γ on $\mathbb{R} \times \mathbb{R}$ with $\Gamma(A \times \mathbb{R}) = \mu(A)$ and $\Gamma(\mathbb{R} \times A) = \nu(A)$ for all Borel set A of $\mathbb{R}$. We define the L^p Wasserstein distance, with $p \geqslant 1$, as*

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{\Gamma \in \Pi(\mu,\nu)} \int |x - y|^p \Gamma(dxdy)\right)^{1/p},$$

or, equivalently,

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{(X,Y)\ s.t\ X\sim\mu, Y\sim\nu} \mathbb{E}\left(|X - Y|^p\right)\right)^{1/p}.$$

We would thus like to prove a property such as $\mathcal{W}_1(\rho_t^{k,N}, \bar{\rho}_t^{\otimes k}) \xrightarrow[N\to\infty]{} 0$. However, instead of considering the infimum over all possible couplings, we build a specific one : we construct simultaneously two solutions of (1.1.1) and (1.1.3) which tend to get closer together. The expectation of the distance between the two solutions for this specific coupling then yields an upper bound on the Wasserstein distance.

The method popularized by A.-S. Sznitman is the *synchronous coupling*, and consists in considering the same Brownian motions for (1.1.1) and (1.1.3), i.e

$$\begin{cases} dX_t^{i,N} = \frac{1}{N}\sum_{j=1}^N K(X_t^{i,N} - X_t^{j,N})dt + \sqrt{2\sigma}dB_t^i, \\ d\bar{X}_t^i = K * \bar{\rho}_t(\bar{X}_t^i)dt + \sqrt{2\sigma}dB_t^i, \\ \bar{\rho}_t = \text{Law}(\bar{X}_t). \end{cases}$$

Notice, by uniqueness of the solutions of (1.1.1) and (1.1.3), that $\rho_t^N = \text{Law}(X_t^{1,N}, ..., X_t^{N,N})$ and $\bar{\rho}_t^{\otimes N} = \text{Law}(\bar{X}_t^1, ..., \bar{X}_t^N)$. Then, for all $i \in \mathbb{N}$

$$d(X_t^{i,N} - \bar{X}_t^i) = \left(\frac{1}{N}\sum_{j=1}^N K(X_t^{i,N} - X_t^{j,N}) - K * \bar{\rho}_t(\bar{X}_t^i)\right) dt.$$

The Brownian motions cancel out, and we are left with an ordinary differential equation. Then

$$d|X_t^{i,N} - \bar{X}_t^i| = \text{sign}(X_t^{i,N} - \bar{X}_t^i)\left(\frac{1}{N}\sum_{j=1}^N K(X_t^{i,N} - X_t^{j,N}) - K * \bar{\rho}_t(\bar{X}_t^i)\right) dt,$$

where

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0. \end{cases}$$

We now have to consider the difference of the drifts. Notice

$$\begin{aligned} \left|\frac{1}{N}\sum_{j=1}^N K(X_t^{i,N} - X_t^{j,N}) - K * \bar{\rho}_t(\bar{X}_t^i)\right| \leqslant & \left|\frac{1}{N}\sum_{j=1}^N K(X_t^{i,N} - X_t^{j,N}) - \frac{1}{N}\sum_{j=1}^N K(\bar{X}_t^i - \bar{X}_t^j)\right| \\ & + \left|\frac{1}{N}\sum_{j=1}^N K(\bar{X}_t^i - \bar{X}_t^j) - K * \bar{\rho}_t(\bar{X}_t^i)\right|. \end{aligned}$$

Hence

$$\begin{aligned} |X_t^{i,N} - \bar{X}_t^i| \leqslant & |X_0^{i,N} - \bar{X}_0^i| + \int_0^t \left|\frac{1}{N}\sum_{j=1}^N K(X_s^{i,N} - X_s^{j,N}) - \frac{1}{N}\sum_{j=1}^N K(\bar{X}_s^i - \bar{X}_s^j)\right| ds \\ & + \int_0^t \left|\frac{1}{N}\sum_{j=1}^N K(\bar{X}_s^i - \bar{X}_s^j) - K * \bar{\rho}_s(\bar{X}_s^i)\right| ds. \end{aligned} \tag{1.2.3}$$

Thus, in order to prove some contraction or convergence, it is necessary (or at the very least it *seems* like it is necessary) to be able to control differences of the form $K(x) - K(y)$, to deal with the first integral of (1.2.3), while the second integral should be dealt with using some form of law of large number since the $(\bar{X}_t^i)_i$ are i.i.d. *A priori*, this coupling method thus only works with sufficiently *nice* interaction kernels K (think of Lipschitz continuous kernels...). Let us however note that such a method has been successfully applied to bounded discontinuous interactions [92].

We postpone the proof of the following lemma to Appendix A.2.

Lemma 1.2.1. *We have for all $t \geqslant 0$ and all $N \in \mathbb{N}$*

$$\mathbb{E}\left|\frac{1}{N}\sum_{j=1}^{N} K(X_t^{i,N} - X_t^{j,N}) - K * \bar{\rho}_t(\bar{X}_t^i)\right| \leqslant 2L\mathbb{E}|X_t^{i,N} - \bar{X}_t^i| + \frac{2LC_0^{1/2}}{N} + 2L\sqrt{\frac{C_0}{N-1}}.$$

There thus exists a constant $C > 0$, such that for all $i \in \mathbb{N}$, (1.2.3) yields

$$\mathbb{E}\left(|X_t^{i,N} - \bar{X}_t^i|\right) \leqslant \mathbb{E}\left(|X_0^{i,N} - \bar{X}_0^i|\right) + \int_0^t \left(2L\mathbb{E}\left(|X_s^{i,N} - \bar{X}_s^i|\right) + \frac{C}{\sqrt{N}}\right) ds,$$

and by Gronwall's lemma

$$\mathbb{E}\left(|X_t^{i,N} - \bar{X}_t^i|\right) \leqslant e^{2Lt}\left(\mathbb{E}\left(|X_0^{i,N} - \bar{X}_0^i|\right) + \frac{C}{2L\sqrt{N}}\right).$$

Finally, for all $k \in \{1, ..., N\}$, we have constructed a specific coupling, which controls the Wasserstein distance

$$\mathcal{W}_1\left(\rho_t^{k,N}, \bar{\rho}_t^{\otimes k}\right) \leqslant \mathbb{E}\left(\sum_{i=1}^{k} |X_t^{i,N} - \bar{X}_t^i|\right) \leqslant e^{2Lt}\left(\mathbb{E}\left(\sum_{i=1}^{k} |X_0^{i,N} - \bar{X}_0^i|\right) + \frac{Ck}{2L\sqrt{N}}\right).$$

Because this holds for any initial coupling of $\rho_0^{k,N}$ and $\bar{\rho}_0^{\otimes k}$, we can in particular choose initial conditions according to the optimal coupling for the Wasserstein distance, and obtain

$$\mathcal{W}_1\left(\rho_t^{k,N}, \bar{\rho}_t^{\otimes k}\right) \leqslant e^{2Lt}\left(\mathcal{W}_1\left(\rho_0^{k,N}, \bar{\rho}_0^{\otimes k}\right) + \frac{Ck}{2L\sqrt{N}}\right). \tag{1.2.4}$$

The control (1.2.4) above is the typical result of propagation of chaos one may obtain : provided $\mathcal{W}_1\left(\rho_0^{k,N}, \bar{\rho}_0^{\otimes k}\right) \xrightarrow[N\to\infty]{} 0$, we have $\mathcal{W}_1\left(\rho_t^{k,N}, \bar{\rho}_t^{\otimes k}\right) \xrightarrow[N\to\infty]{} 0$ for all $t \geqslant 0$. The bound consists in the propagation of the initial distance $\mathcal{W}_1\left(\rho_0^{k,N}, \bar{\rho}_0^{\otimes k}\right)$ and an additional error $\frac{Ck}{2L\sqrt{N}}$ which comes from the interactions (i.e the law of large number). Notice furthermore that this result is not uniform in time, because of the term e^{2Lt}.

Remark 1.2.1. *Some preliminary remarks on the method:*

- *This proof is, to some extent, the most "classical" proof given in the Lipschitz continuous case.*
- *These calculations yield an explicit rate of convergence.*
- *We have, from the start, cancelled out the Brownian motions. It may however be interesting to make use of this randomness, as explained later in Chapter 2, to obtain a uniform in time result.*

- *From a probabilistic point of view, this method of proof is quite satisfactory, as it provides a better understanding of the stochastic behavior of the particles. The back-and-forth between intuition and calculations will (hopefully) be more evident in the study of another coupling method later in this document.*

- *To the best of our knowledge, such methods still fail in dealing with singular kernels K. As mentioned, so far, they may only deal with Lipschitz continuous interactions, or possibly discontinuous but bounded interactions [92].*

1.2.2 Energy and entropy estimates

Instead of the Wasserstein distance, some physical quantities might be better suited to the use of tools from PDE analysis. The comparison between the law of the system of N interacting particles and the law of N independent particles satisfying the nonlinear equation (1.1.3) can for instance be done in terms of relative entropy.

Definition 1.2.2 (Relative entropy)**.** *Let μ and ν be two probability measures on $\mathbb{R}^d$. We consider the relative entropy*

$$\mathcal{H}(\nu,\mu) = \begin{cases} \mathbb{E}_\mu\left(\frac{d\nu}{d\mu}\log\frac{d\nu}{d\mu}\right) & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases} \tag{1.2.5}$$

It is an interesting quantity in part as it controls usual distances. With a bit more work (for instance proving a logarithmic Sobolev inequality, see [145]), a Talagrand's transportation inequality shows that the L^2-Wasserstein distance is bounded by the relative entropy. Classically, this is also the case of the total variation distance thanks to Pinsker's inequality.

Definition 1.2.3 (Total variation distance in $\mathbb{R}$)**.** *Let μ and ν be two probability measures on $\mathbb{R}$ and denote $\mathcal{B}(\mathbb{R})$ the Borel σ-algebra. The total variation distance is defined by*

$$d_{TV}(\mu,\nu) := 2 \sup_{A\in\mathcal{B}(\mathbb{R})} |\mu(A)-\nu(A)|.$$

Equivalently, the total variation distance can be written as a coupling problem

$$d_{TV}(\mu,\nu) := 2 \inf_{(X,Y)\ s.t.\ X\sim\mu, Y\sim\nu} \mathbb{P}(X\neq Y).$$

Again, the main idea of this method of proof is the following : compute the time derivative of the chosen "physical" quantity (here the relative entropy), use some form of law of large number or large deviation principle, and conclude using Gronwall's lemma, which allows us to bound the quantity at time t by its value at time 0 with an additional correction vanishing as N goes to infinity.

To compute the time derivative of the relative entropy, we have the following (formal) lemma

Lemma 1.2.2. *For $i = 1,2$ and $t > 0$, assume $P_t^i \in \mathcal{P}(\mathbb{R}^m)$ satisfies*

$$\partial_t P^i = -\nabla\cdot(b^i P^i) + \sigma\Delta P^i,$$

for some vector fields $b^i : [0,T]\times\mathbb{R}^m \mapsto \mathbb{R}^m$. For the sake of simplicity, assume $dP_t^i(x) = P_t^i(x)dx$

with $P_t^i \in \mathcal{C}^\infty$, $P_t^1 \ll P_t^2$ *and both* P_t^i *vanish to 0 at* $\pm\infty$. *Then*

$$\frac{d}{dt}\mathcal{H}\left(P_t^1, P_t^2\right) = -\sigma \int_{\mathbb{R}^m} P_t^1 \left|\nabla \log \frac{P_t^1}{P_t^2}\right|^2 - \int_{\mathbb{R}^m} P_t^1 \nabla \cdot \left(b^1 - b^2\right) - \int_{\mathbb{R}^m} P_t^1 \left(b^1 - b^2\right) \cdot \nabla \log P_t^2 \tag{1.2.6}$$

$$= \int_{\mathbb{R}^m} P_t^1 \left(\left(b^1 - b^2\right) \cdot \nabla \log \frac{P_t^1}{P_t^2} - \sigma \left|\nabla \log \frac{P_t^1}{P_t^2}\right|^2 \right). \tag{1.2.7}$$

We postpone the (formal) proof of the lemma to Appendix A.3.

Let us assume, in order to carry out the computations in a more comfortable way and use the lemma above, that $\bar{\rho}_t$ is a smooth strong solution of (1.1.4) in dimension one, which also vanishes to 0 at $\pm\infty$. Likewise, ρ_t^N is assumed to be a smooth strong solution of (1.1.2), which is such that $\rho_t^N \ll \bar{\rho}_t^{\otimes N}$. By definition,

$$\mathcal{H}(\rho_t^N, \bar{\rho}_t^{\otimes N}) = \int_{\mathbb{R}^N} \rho_t^N \log \frac{\rho_t^N}{\bar{\rho}_t^{\otimes N}}.$$

Applying (1.2.6) to $P_t^1 = \rho_t^N$ and $P_t^2 = \bar{\rho}_t^{\otimes N}$, we get for all $t \geqslant 0$ and all $N \in \mathbb{N}$

$$\begin{aligned}\frac{d}{dt}\mathcal{H}(\rho_t^N, \bar{\rho}_t^{\otimes N}) = &-\sigma \sum_i \int \rho_t^N \left|\partial_i \log \frac{\rho_t^N}{\bar{\rho}_t^{\otimes N}}\right|^2 \\ &- \frac{1}{N}\sum_i \int \rho_t^N \left(\frac{1}{N}\sum_j K(x_i - x_j) - K * \bar{\rho}_t(x_i)\right) \partial_i \log \bar{\rho}_t^{\otimes N} \\ &- \frac{1}{N}\sum_i \int \rho_t^N \left(\frac{1}{N}\sum_j \partial_i K(x_i - x_j) - \partial_i K * \bar{\rho}_t(x_i)\right).\end{aligned} \tag{1.2.8}$$

We often write

$$\mathcal{I}(\rho_t^N, \bar{\rho}_t^{\otimes N}) = \sum_i \int \rho_t^N \left|\partial_i \log \frac{\rho_t^N}{\bar{\rho}_t^N}\right|^2,$$

the Fisher information. As it is a non-negative quantity, we may discard it in order to give an upper bound on the derivative of the relative entropy.

We are left with bounding the last two terms in (1.2.8). These are, again, the differences of the drifts, which resemble the last term of (1.2.3) in expectation. We show later in Chapter 4 how such calculations are carried out, but in the meantime we accept that we obtain a result akin to

$$\frac{d}{dt}\mathcal{H}(\rho_t^N, \bar{\rho}_t^{\otimes N}) \leqslant C\left(\mathcal{H}(\rho_t^N, \bar{\rho}_t^{\otimes N}) + 1\right),$$

which, like previously, allows us to conclude using Gronwall's lemma.

We thus obtain a result that is $\int_{\mathbb{R}^N} \rho_t^N \log \frac{\rho_t^N}{\bar{\rho}_t^{\otimes N}} = O(1)$. To conclude on the convergence of the marginals, we use the property of sub-additivity of the entropy to get

$$\mathcal{H}(\rho_t^{k,N}, \bar{\rho}_t^{\otimes k}) = \int_{\mathbb{R}^k} \rho_t^{k,N} \log \frac{\rho_t^{k,N}}{\bar{\rho}_t^{\otimes k}} \leqslant C\frac{k}{N}, \tag{1.2.9}$$

for some constant C, independent of k and N, but not of time t.

Remark 1.2.2. *Some preliminary remarks on the method :*

- *Here we have discarded the Fisher information on the basis that it is a non-negative quantity. It can however be interesting to make use of it ; through a log-Sobolev inequality, one can bound the relative entropy thanks to the Fisher information. This is something we will do later in this document to obtain a uniform in time result.*
- *Instead of considering the entropy and its time evolution, it is possible to consider other quantities, such as a modulated energy or a modulated free energy, as suggested in [159, 152, 32].*
- *This methods allows us to deal with singular kernels, as it will be done later. The main idea is to use the regularity of $\bar{\rho}_t$ to compensate the singularity of the interaction kernel K. One thus ends up relying heavily on the properties of $\bar{\rho}_t$ that one may be able to prove (H^p estimates, regularity, etc).*
- *This proof only stays at the level of the PDE* (1.1.2)*, and does not concern itself with the SDE* (1.1.1)*, which is the "true" microscopic description of the system. This is not an issue provided we can can prove there is a one-to-one connection between the PDE and the SDE (as in : there exists a unique solution to* (1.1.2)*, which is the law of the particle system* (1.1.1)*). However, when considering singular kernels, this can be highly non trivial [114].*

1.2.3 BBGKY hierarchies

The method described above in Section 1.2.2 intuitively approaches the problem in a *global* way. Indeed, we consider the entropy (or energy, or else) of ρ_t^N with respect to $\bar{\rho}_t^{\otimes N}$, i.e of the joint law of the *entire* particle system.

But, in reality, what we wish to obtain is a result of convergence on the marginals (which we deduce from the global approach via sub-additivity).

Another idea would thus be to directly work at the level of the marginal distributions. From (1.1.2), by integrating over $(x_{k+1}, ..., x_N)$ and using the exchangeability of the particle system, we observe that $\rho_t^{k,N}$ satisfies the so-called *BBGKY hierarchy*

$$\begin{aligned}\partial_t \rho_t^{k,N} = &- \sum_{i=1}^k \partial_i \left(\left(\frac{1}{N} \sum_{j=1}^k K(x_i - x_j) \right) \rho_t^{k,N} \right) \\ &- \frac{N-k}{N} \sum_{i=1}^k \partial_i \left(\int_{\mathbb{R}} K(x_i - x_{k+1}) \rho_t^{k+1,N} dx_{k+1} \right) + \sigma \sum_{i=1}^k \partial_i^2 \rho_t^{k,N}. \end{aligned} \tag{1.2.10}$$

Equation (1.2.10) is not closed (except for $k = N$, for which it corresponds to (1.1.2)), as $\rho_t^{k,N}$ depends on $\rho_t^{k+1,N}$, thus the denomination of hierarchy.

Equivalently (see Lemma 4.1 of [111]), this corresponds to remarking that $(X^{1,N}, ..., X^{k,N})$ solves the SDE system

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^k K(X_t^{i,N} - X_t^{j,N})dt + \frac{N-k}{N} b_i^k(t, X^{1,N}, ..., X^{k,N})dt + \sqrt{2\sigma} dW_t^i, \tag{1.2.11}$$

for some independent Brownian motions $W^1, ..., W^k$ and a progressively measurable function b_i^k such that

$$b_i^k(t, X^{1,N}, ..., X^{k,N}) = \mathbb{E}\left(K(X_t^{i,N} - X_t^{k+1,N})|(X_s^{1,N}, ..., X_s^{k,N})_{s\leqslant t}\right) \qquad \text{a.s., a.e. } t,$$

i.e a progressively measurable version of the conditional expectation.

Starting from the hierarchies, we may now consider various approaches.

Propagating bounds

We may propagate L^p bounds along this hierarchy, as it was for instance done recently in [30]. From these L^p bounds, one is then able to extract a subsequence such that every $\rho^{k,N}$ converges weakly-* in some functional space.

For simplicity here, let us assume $K' \geqslant 0$. Let $p > 1$, and notice that, by multiplying (1.2.10) by $(\rho_t^{k,N})^{p-1}$ and integrating over $(x_1, ..., x_k)$, we (formally) obtain

$$\begin{aligned}
\frac{1}{p}\partial_t \int_{\mathbb{R}^k} (\rho_t^{k,N})^p = &- \sum_{i=1}^k \int_{\mathbb{R}^k} (\rho_t^{k,N})^{p-1}\partial_i \left(\left(\frac{1}{N}\sum_{j=1}^k K(x_i - x_j)\right)\rho_t^{k,N}\right) \\
&- \frac{N-k}{N}\sum_{i=1}^k \int_{\mathbb{R}^k} (\rho_t^{k,N})^{p-1}\partial_i \left(\int_{\mathbb{R}} K(x_i - x_{k+1})\rho_t^{k+1,N} dx_{k+1}\right) \\
&+ \sigma \sum_{i=1}^k \int_{\mathbb{R}^k} (\rho_t^{k,N})^{p-1}\partial_i^2 \rho_t^{k,N} \\
= &- \frac{p-1}{p}\sum_{i=1}^k \int_{\mathbb{R}^k} (\rho_t^{k,N})^p \left(\frac{1}{N}\sum_{j=1}^k K'(x_i - x_j)\right) \\
&- \frac{N-k}{N}\sum_{i=1}^k \int_{\mathbb{R}^k} (\rho_t^{k,N})^{p-1}\partial_i \left(\int_{\mathbb{R}} K(x_i - x_{k+1})\rho_t^{k+1,N} dx_{k+1}\right) \\
&- \sigma(p-1)\sum_{i=1}^k \int_{\mathbb{R}^k} (\rho_t^{k,N})^{p-2}\left(\partial_i \rho_t^{k,N}\right)^2. \qquad (1.2.12)
\end{aligned}$$

Recall Young's inequality

$$\forall a, b \geqslant 0, \quad \forall p, p^* > 1 \quad \text{s.t} \quad \frac{1}{p} + \frac{1}{p^*} = 1, \quad ab \leqslant \frac{a^p}{p} + \frac{b^{p^*}}{p^*},$$

which allows us to bound

$$\partial_t \int_{\mathbb{R}^k} (\rho_t^{k,N})^p dx_1...dx_k \leqslant C_{p,\sigma}\frac{N-k}{N}k \int_{\mathbb{R}^k} \left|\int_{\mathbb{R}} K(x_1 - x_{k+1})\rho_t^{k+1,N} dx_{k+1}\right|^p dx_1...dx_k,$$

where $C_{p,\sigma}$ is some universal constant depending only on p and σ. Notice that the derivative of $\rho_t^{k,N}$ appearing in the second to last term of (1.2.12), which is usually in these sort of calculations the main issue as the loss of a derivative at each step of the hierarchy constrains the study to a complicated analytical space, is here dealt with using the last term of (1.2.12) (via Young's inequality) i.e the diffusion part.

Another Hölder estimate then yields

$$\frac{d}{dt}||\rho_t^{k,N}||_{L^p} \leqslant C_{p,\sigma}\frac{k(N-k)}{N}||K||_{L^{p^*}}||\rho_t^{k+1,N}||_{L^p}. \tag{1.2.13}$$

Here again we obtain a hierarchy of ODE's. Provided at time $t=0$ we have $||\rho_0^{k,N}||_{L^p} \leqslant R^k$ for some $R>0$, which we can expect from an initially tensorized law, and if $p>1$ is sufficiently large to ensure $||K||_{L^{p^*}} < \infty$, one can show that there exists a time $T^* = T^*(d,p,\sigma,R)$ such that for $t \in [0,T^*]$ the estimate $||\rho_t^{k,N}||_{L^p} \leqslant (2R)^k$ remains true.

These uniform bounds let us extract a weakly-* converging subsequence in $L^\infty([0,T^*],L^p)$. Passing to the weak limit in each term of the hierarchy shows that the limit $\tilde{\rho}_k$ also satisfies an (infinite) Vlasov hierarchy

$$\partial_t\tilde{\rho}_k = -\sum_{i=1}^k \partial_i\left(\int_{\mathbb{R}} K(x_i - x_{k+1})\tilde{\rho}_{k+1}dx_{k+1}\right) + \sigma\sum_{i=1}^k \partial_i^2\tilde{\rho}_k.$$

It remains to prove uniqueness of the solutions to this Vlasov hierarchy, noticing that $\bar{\rho}^{\otimes k}$ is also a solution, to conclude on the convergence of the marginals.

Remark 1.2.3. *Some remarks on the method:*

- *These calculations are quite robust, and allow for singular kernels K : they only require $K \in L^p$ for some $p>1$. Furthermore, they can adapt to a kinetic setting (see below for an explanation of the kinetic SDE), as it was done in [30].*
- *It does not give an explicit rate of convergence in N, and so far only works on a finite time frame $[0,T^*]$ with other technical restrictions (the processes must be on the torus for instance).*

Local relative entropy

A recent approach, see for instance [111], consists in considering the relative entropy of the k-marginal $\rho_t^{k,N}$ with respect to $\bar{\rho}_t^{\otimes k}$.

Denote $H_t^k = \mathcal{H}\left(\rho_t^{k,N}, \bar{\rho}_t^{\otimes k}\right)$. Using (1.2.7) we have

$$\frac{d}{dt}H_t^k = \int_{\mathbb{R}^k}\left(\left(B^1(t,x_1,...,x_k) - B^2(t,x_1,...,x_k)\right)\cdot\nabla\log\frac{\rho_t^{k,N}}{\bar{\rho}_t^{\otimes k}} - \sigma\left|\nabla\log\frac{\rho_t^{k,N}}{\bar{\rho}_t^{\otimes k}}\right|^2\right)\rho_t^{k,N},$$

where B^1 and B^2 are the respective drifts of $\rho_t^{k,N}$ and $\bar{\rho}_t^{\otimes k}$, i.e

$$B^1(t,x_1,...,x_k) = \left(\frac{1}{N}\sum_{j=1}^k K(x_i - x_j) + \frac{N-k}{N}b_i^k(t,x_1,...,x_k)\right)_{1\leqslant i\leqslant k},$$

$$B^2(t,x_1,...,x_k) = \left(\int_{\mathbb{R}} K(x_i - x_{k+1})\bar{\rho}_t(dx_{k+1})\right)_{1\leqslant i\leqslant k}.$$

In particular, we may bound

$$\frac{d}{dt}H_t^k \leqslant \frac{1}{4\sigma}\int_{\mathbb{R}^k}\rho_t^{k,N}\left|B^1(t,x_1,...,x_k) - B^2(t,x_1,...,x_k)\right|^2$$

$$
\begin{aligned}
=&\frac{1}{4\sigma}\sum_{i=1}^{k}\mathbb{E}\left(\left|\frac{1}{N}\sum_{j=1}^{k}K(X_t^{i,N}-X_t^{j,N})+\frac{N-k}{N}b_i^k(t,X_t^{1,N},...,X_t^{k,N})\right.\right.\\
&\qquad\qquad\qquad\qquad\qquad\left.\left.-\int_{\mathbb{R}}K(X_t^{i,N}-y)\bar{\rho}_t(dy)\right|^2\right)\\
\leqslant&\frac{1}{2\sigma}\sum_{i=1}^{k}\mathbb{E}\left(\left|\frac{1}{N}\sum_{j=1}^{k}\left(K(X_t^{i,N}-X_t^{j,N})-\int_{\mathbb{R}}K(X_t^{i,N}-y)\bar{\rho}_t(dy)\right)\right|^2\right)\\
&+\frac{1}{2\sigma}\sum_{i=1}^{k}\mathbb{E}\left(\left|\frac{N-k}{N}\left(\mathbb{E}\left(K(X_t^{i,N}-X_t^{k+1,N})|(X_s^{1,N},...,X_s^{k,N})_{s\leqslant t}\right)\right.\right.\right.\\
&\qquad\qquad\qquad\qquad\qquad\left.\left.\left.-\int_{\mathbb{R}}K(X_t^{i,N}-y)\bar{\rho}_t(dy)\right)\right|^2\right)\\
=&\frac{k}{2\sigma N^2}\mathbb{E}\left(\left|\sum_{j=2}^{k}\left(K(X_t^{1,N}-X_t^{j,N})-K*\bar{\rho}_t(X_t^{1,N})\right)\right|^2\right)\\
&+\frac{k(N-k)^2}{2\sigma N^2}\mathbb{E}\left(\left|\mathbb{E}\left(K(X_t^{1,N}-X_t^{k+1,N})|(X_s^{1,N},...,X_s^{k,N})_{s\leqslant t}\right)-K*\bar{\rho}_t(X_t^{1,N})\right|^2\right),
\end{aligned}
$$

where we used the exchangeability for this last line. We first use the assumption on the bound of the second moments to obtain the crude bound

$$
\frac{k}{2\sigma N^2}\mathbb{E}\left(\left|\sum_{j=2}^{k}\left(K(X_t^{1,N}-X_t^{j,N})-K*\bar{\rho}_t(X_t^{1,N})\right)\right|^2\right)\leqslant C\frac{k(k-1)^2}{N^2},
$$

for some constant C independent of k and N.

For the second term, the key ingredient is a transport type inequality for $\bar{\rho}_t$

$$
\exists\gamma>0,\quad\forall\nu\in\mathcal{P}(\mathbb{R}),\quad\forall t,x,\qquad\left|\int_{\mathbb{R}}K(x-y)\nu(dy)-\int_{\mathbb{R}}K(x-y)\bar{\rho}_t(dy)\right|^2\leqslant\gamma\mathcal{H}\left(\nu,\bar{\rho}_t\right).\quad(1.2.14)
$$

Denoting $\rho^{(k+1|k)}_{t,X^{1,N},...,X^{k,N}}$ the conditional law of $X_t^{k+1,N}$ given $(X_s^{1,N},...,X_s^{k,N})_{s\leqslant t}$, we have

$$
\begin{aligned}
&\mathbb{E}\left(\left|\mathbb{E}\left(K(X_t^{1,N}-X_t^{k+1,N})|(X_s^{1,N},...,X_s^{k,N})_{s\leqslant t}\right)-K*\bar{\rho}_t(X_t^{1,N})\right|^2\right)\\
&=\mathbb{E}\left(\left|\int_{\mathbb{R}}K(X_t^{1,N}-y)\rho^{(k+1|k)}_{t,X^{1,N},...,X^{k,N}}(dy)-\int_{\mathbb{R}}K(X_t^{1,N}-y)\bar{\rho}_t(dy)\right|^2\right)\\
&\leqslant\gamma\mathbb{E}\left(\mathcal{H}\left(\rho^{(k+1|k)}_{t,X^{1,N},...,X^{k,N}},\bar{\rho}_t\right)\right).
\end{aligned}
$$

We then use a chain rule for the entropy to obtain

$$
\mathbb{E}\left(\mathcal{H}\left(\rho^{(k+1|k)}_{t,X^{1,N},...,X^{k,N}},\bar{\rho}_t\right)\right)=H_t^{k+1}-H_t^k,
$$

and thus obtain the key differential inequality

$$\frac{d}{dt}H_t^k \leqslant C\frac{k(k-1)^2}{N^2} + \gamma k(H_t^{k+1} - H_t^k). \tag{1.2.15}$$

By using Gronwall's lemma, and iterating (1.2.15) over $k \in [0, N]$ (closing the hierarchy with the crude bound $H_t^N = O(N)$), we may obtain $H_t^k = O\left(\frac{k^3}{N^2}\right)$. Plugging this estimate back into the previous calculations yields the final result

$$H_t^k \leqslant C\frac{k^2}{N^2}, \tag{1.2.16}$$

for some constant C, independent of k and N, but not of time t.

Remark 1.2.4. *Some remarks on the method:*

- *The main interest of this proof, and the reason we mention it, is the rate of convergence* (1.2.16) *which is a significant improvement over the rate* (1.2.9)*.*
- *The key ingredient* (1.2.14) *is satisfied by bounded kernels K (thanks to Pinsker's inequality), by Lipschitz continuous K (using the result of [60]), or by K with linear growth or Hölder continuous K. To the best of our knowledge,* (1.2.14) *is still the main issue when trying to apply the method to more general, and possibly singular, interaction kernels K.*
- *Once again, we may use the Fisher information (instead of discarding it) via a log-Sobolev inequality, as it was done in [113], to get a uniform in time estimate. Doing so, with some added technical assumptions, yields a constant C in* (1.2.16) *which is independent of the time t.*

1.2.4 Compactness

Consider $\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ the empirical measure. This object can have two meanings :

- either it is seen as a probability measure on the path space, i.e $\mu^N \in \mathcal{P}(\mathcal{C}([0,T],\mathbb{R}))$, in the sense that we randomly choose one of the particles that we follow along its path,
- or it is seen as a continuous function of the space of probability measures, i.e $\mu^N \in \mathcal{C}([0,T],\mathcal{P}(\mathbb{R}))$, in the sense that, at each time t, μ_t^N is a probability measure.

In both cases, it is a random variable (since the particles are random) and we consider ξ^N its law. Our goal is to show that ξ^N converges, for the weak topology, to $\delta_{\bar\rho}$ where $\bar\rho$ is the solution of (1.1.4). To do so, we rely on three steps:

1. **Tightness :** Show that the law of μ^N is tight, from which we deduce convergence up to extraction of a subsequence,
2. **Identification :** Show that the possible limits are made up of solutions of the nonlinear problem,
3. **Uniqueness :** Prove the uniqueness of the above limit object.

We show in the remaining of the section how one may obtain tightness in both cases.

A probability measure on the path space

Here we consider $\mu^N \in \mathcal{P}(\mathcal{C}([0,T],\mathbb{R}))$. Thanks to the following lemma, which can be found as Lemma 4.5 in [137], tightness boils down to the tightness of the trajectories.

Lemma 1.2.3. *The tightness of $\xi^N \in \mathcal{P}(\mathcal{P}(\mathcal{C}([0,T],\mathbb{R})))$ is equivalent to the tightness of the probability distribution of $X^{1,N} \in \mathcal{C}([0,T],\mathbb{R})$.*

We thus wish to prove that for all $\epsilon > 0$ and all $T > 0$, there exists a compact set $K_\epsilon \subset \mathcal{C}([0,T],\mathbb{R})$ such that

$$\sup_{N\geqslant 1} \mathbb{P}\left((X_t^{1,N})_{t\in[0,T]} \notin K_\epsilon\right) \leqslant \epsilon.$$

Let $\epsilon > 0$ and $T > 0$. To obtain the existence of such a compact set, we need some form of equicontinuity. The idea behind the following calculations comes from [76]. By definition, we have

$$X_t^{1,N} - X_s^{1,N} = \sqrt{2\sigma}\left(B_t^1 - B_s^1\right) + \int_s^t \frac{1}{N}\sum_{j=1}^N K\left(X_u^{1,N} - X_u^{j,N}\right) du.$$

Let $Z_T = \sqrt{2\sigma}\sup_{0\leqslant s<t\leqslant T}\frac{|B_t^1 - B_s^1|}{|t-s|^{1/3}}$. Since the Brownian motion is in particular $\frac{1}{3}$-Hölder continuous, Z_T is almost surely finite and $\mathbb{E}Z_T < \infty$ (see for instance Theorem 2.1 of [150]).

Furthermore

$$\begin{aligned}
\left|\int_s^t \frac{1}{N}\sum_{j=1}^N K\left(X_u^{1,N} - X_u^{j,N}\right) du\right| &\leqslant \frac{1}{N}\sum_{j=1}^N \int_s^t \left|K\left(X_u^{1,N} - X_u^{j,N}\right)\right| du \\
&\leqslant \frac{L}{N}\sum_{j=1}^N \int_0^T \mathbb{1}_{u\in[s,t]} \left|X_u^{1,N} - X_u^{j,N}\right| du \\
&\leqslant \frac{L}{N}\sum_{j=1}^N (t-s)^{1/3}\left(\int_0^T \left|X_u^{1,N} - X_u^{j,N}\right|^{3/2} du\right)^{2/3}, \quad (1.2.17)
\end{aligned}$$

where we used Hölder inequality for this last line. Thus

$$\left|X_t^{1,N} - X_s^{1,N}\right| \leqslant (Z_T + C_{N,T})(t-s)^{1/3}, \quad (1.2.18)$$

where $C_{N,T}$ is such that, using the assumption on the bounded second moment of the particle system, there exists a constant $C > 0$ such that $\sup_N \mathbb{E}C_{N,T} < C$.

In particular, there exists $R_\epsilon > 0$ such that for all $N \geqslant 1$ we have $\mathbb{P}(Z_T + C_{N,T} > R_\epsilon) \leqslant \frac{\epsilon}{2}$. Likewise, using the assumption on the bounded initial moments, there is $a_\epsilon > 0$ such that for all $N \geqslant 1$ we have $\mathbb{P}\left(|X_0^{1,N}| > a_\epsilon\right) \leqslant \frac{\epsilon}{2}$.

Let K_ϵ be the set of continuous functions $f : [0,T] \mapsto \mathbb{R}$ such that $|f(0)| \leqslant a_\epsilon$ and $|f(t) - f(s)| \leqslant R_\epsilon |t-s|^{1/3}$ for all $0 \leqslant s \leqslant t \leqslant T$. It is a compact subset of $\mathcal{C}([0,T],\mathbb{R})$ and for all $N \geqslant 1$

$$\mathbb{P}\left((X_t^{1,N})_{t\in[0,T]} \notin K_\epsilon\right) \leqslant \mathbb{P}\left(\left|X_0^{1,N}\right| > a_\epsilon\right) + \mathbb{P}(Z_T + C_{N,T} > R_\epsilon) \leqslant \epsilon.$$

We thus obtain the tightness of the probability distribution of $X^{1,N} \in \mathcal{C}([0,T],\mathbb{R})$, and thus of ξ^N. From Prokhorov theorem, we may extract a converging subsequence towards a limit ξ, which we may identify as the solution of a martingale problem for which we then need to obtain uniqueness.

A continuous function on the space of probability measures

To show tightness of the random variables $\mu^N \in \mathcal{C}([0,T],\mathcal{P}(\mathbb{R}))$ we use the following Kolmogorov criterion (see for instance Proposition 2.8 of [138] or Corollary 14.9 of [107]).

Proposition 1 (Kolmogorov criterion). *Let $(\mu^N)_N$ be a sequence of random variables in $\mathcal{C}([0,T],\mathcal{P}(\mathbb{R}))$. Assume*

- *$(\mu^N_t)_N$ is tight for all $t \in [0,T]$,*
- *there exists $C, p, \beta > 0$ such that*

$$\sup_N \mathbb{E}\left(\mathcal{W}_1(\mu^N_t,\mu^N_s)^p\right) \leqslant C(t-s)^{1+\beta}, \qquad \text{for all } s,t \in [0,T].$$

Then $(\mu^N)_N$ is tight.

The first point directly comes from the uniform bounds on the moments of the particle system. For the second point, we have

$$\mathcal{W}_1(\mu^N_t,\mu^N_s) \leqslant \frac{1}{N}\sum_{i=1}^N \left|X^{i,N}_t - X^{i,N}_s\right|.$$

Thus, similarly as (1.2.18), there are positive random variables Z^i_T and $C^i_{N,T}$ satisfying $\sup_N \sup_i \mathbb{E}(\left|Z^i_T + C^i_{N,T}\right|^4) < \infty$ (again, by Theorem 2.1 of [150] and the bounded moment assumption) such that

$$\mathcal{W}_1(\mu^N_t,\mu^N_s) \leqslant \left(\frac{1}{N}\sum_{i=1}^N Z^i_T + C^i_{N,T}\right)(t-s)^{1/3}.$$

We thus obtain

$$\sup_N \mathbb{E}\left(\mathcal{W}_1(\mu^N_t,\mu^N_s)^4\right) \leqslant C(t-s)^{1+\frac{1}{3}}.$$

The sequence $(\mu^N)_N$ is therefore tight. Up to extraction, μ^N converges in law to a random variable $\bar{\rho}$. It then remains to show that $\bar{\rho}$ is a solution of (1.1.4), and uniqueness of the latter yields the convergence.

Remark 1.2.5. *Some remarks :*

- *The main advantage of this approach is its wide range of applicability, as it can for instance be adapted to càdlàg processes, hence allowing jumps [137]. Other types of criteria for tightness are in this case necessary (see Aldous' criterion, given for instance in Theorem 16.10 of [18]).*

- *Obtaining tightness, regardless of the point of view, consists in "controlling" in expectation the interaction term (see for instance* (1.2.17)*). In particular, for kernels K with a singularity at 0, this amounts to controlling close encounters of particles [43, 76]. However, the uniqueness of the limit object may become difficult to obtain in those cases.*

- *This method does not yield any rate of convergence.*

- *The weak convergence Law* $\left(\mu^N\right) \in \mathcal{P}\left(\mathcal{P}\left(\mathcal{C}([0,T],\mathbb{R})\right)\right) \xrightarrow[N\to\infty]{} \delta_{\bar{\rho}}$ *is stronger than Law* $\left(\mu^N\right) \in \mathcal{P}\left(\mathcal{C}\left([0,T],\mathcal{P}\left(\mathbb{R}\right)\right)\right) \xrightarrow[N\to\infty]{} \delta_{\bar{\rho}}$*, as shown in Theorem 4.7 of [137].*

1.2.5 Weak derivatives

The goal is to provide an error estimation for quantities of the form

$$\left|\mathbb{E}\Phi\left(\mu_t^N\right) - \Phi\left(\bar{\rho}_t\right)\right|, \tag{1.2.19}$$

where $\Phi : \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$ is a test function chosen within a suitable class. The idea, to deal with these quantities, is to work with a semigroup that acts on the space of functions of measures.

Define $(P_t)_{t\geqslant 0}$ the semigroup generated by (1.1.4), whose action on a function $\Phi : \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$ reads

$$P_t\Phi : \bar{\rho}_0 \in \mathcal{P}(\mathbb{R}) \mapsto \Phi\left(\bar{\rho}_t\right),$$

where $(\bar{\rho}_t)_t$ is the solution of (1.1.4) associated to the initial condition $\bar{\rho}_0$.

Studying this semigroup, as done in [140] for instance, requires defining an appropriate form of differential calculus on the space of probability measures. The framework detailed in [48], introduced by P.-L. Lions in his lectures at *Collège de France* and for which we refer to [36], relies on the following notion of derivatives :

Definition 1.2.4 (Linear functional derivatives)**.** *A function $U : \mathcal{P}\left(\mathbb{R}^d\right) \mapsto \mathbb{R}$ is said to be continuously differentiable if there exists a continuous function $\frac{\delta U}{\delta m} : \mathcal{P}\left(\mathbb{R}^d\right) \times \mathbb{R}^d \mapsto \mathbb{R}$ such that for any $m, m' \in \mathcal{P}\left(\mathbb{R}^d\right)$,*

$$U(m') - U(m) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}\left((1-s)m + sm', y\right)(m'-m)(dy)ds,$$

We call $\frac{\delta U}{\delta m}$ the linear functional derivative fo U, with the convention $\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m,y)m(dy) = 0$ in order to ensure its unique definition.

By induction, one can introduce higher order derivatives

$$\begin{aligned}\frac{\delta^{p-1}U}{\delta m^{p-1}}(m', y_1, ..., y_{p-1}) - &\frac{\delta^{p-1}U}{\delta m^{p-1}}(m, y_1, ..., y_{p-1}) \\ &= \int_0^1 \int_{\mathbb{R}^d} \frac{\delta^p U}{\delta m^p}\left((1-s)m + sm', y_1, ..., y_{p-1}, y\right)(m'-m)(dy)ds,\end{aligned}$$

again with $\int_{\mathbb{R}^d} \frac{\delta^p U}{\delta m^p}\left(m, y_1, ..., y_{p-1}, y\right) m(dy) = 0$ to ensure uniqueness.

In particular, for $\Phi(\mu) := \int F(x)d\mu(x)$ with smooth function F, we can directly compute

$$\frac{\delta^p \Phi}{\delta m^p}\left(\mu, y_1, ..., y_p\right) = (-1)^p \left(\int F(x)d\mu(x) - F(y_p)\right). \tag{1.2.20}$$

The key fact concerning $(P_t\Phi)_{t\geqslant 0}$ is that, provided Φ is sufficiently regular, it is a solution of the so-called *master equation*. Denote, for a given function Φ,

$$\mathcal{U}(t,\mu) := P_t\Phi(\mu),$$

which thus satisfies (under suitable conditions on K),

$$\partial_t\mathcal{U}(t,\mu) = \int_{\mathbb{R}} \left(\partial_x \frac{\delta\mathcal{U}}{\delta m}(t,\mu)(x) K*\mu(x) + \sigma\partial_x^2 \frac{\delta\mathcal{U}}{\delta m}(t,\mu)(x)\right)\mu(dx),\ t\geqslant 0.$$

The starting point in order to obtain bounds on (1.2.19) is the following decomposition

$$\begin{aligned}\Phi\left(\mu_t^N\right) - \Phi\left(\bar\rho_t\right) =& \mathcal{U}\left(0,\mu_t^N\right) - \mathcal{U}\left(t,\bar\rho_0\right)\\ =& \left(\mathcal{U}\left(t,\mu_0^N\right) - \mathcal{U}\left(t,\bar\rho_0\right)\right) + \left(\mathcal{U}\left(0,\mu_t^N\right) - \mathcal{U}\left(t,\mu_0^N\right)\right).\end{aligned}$$

It is shown in [48, 57] that we have the two estimates

$$\mathbb{E}\left(\mathcal{U}\left(0,\mu_t^N\right) - \mathcal{U}\left(t,\mu_0^N\right)\right) = \frac{1}{N}\int_0^t \mathbb{E}\left(\int\left(\partial^2_{y_1y_2}\frac{\delta^2\mathcal{U}}{\delta m^2}(t-s,\mu_s^N)(z,z)\right)\mu_s^N(dz)\right)ds, \quad (1.2.21)$$

and

$$\mathbb{E}\left(\mathcal{U}\left(t,\mu_0^N\right)\right) - \mathcal{U}\left(t,\bar\rho_0\right) = \frac{1}{N}\int_0^1\int_0^1 \mathbb{E}\left(s\frac{\delta^2\mathcal{U}}{\delta m^2}(t,m^N_{ss_1})(\eta,\eta) - s\frac{\delta^2\mathcal{U}}{\delta m^2}(t,m^N_{ss_1})(\eta,X_0^{1,N})\right)ds_1ds, \quad (1.2.22)$$

where η is distributed according to $\bar\rho_0$ and

$$m^N_{ss_1} = \frac{ss_1}{N}(\delta_\eta - \delta_{X_0^{1,N}}) + \bar\rho_0 + s(\mu_0^N - \bar\rho_0).$$

Finally, it has been shown in [166] that under some regularity conditions on K and Φ,

$$\frac{\delta^2\mathcal{U}}{\delta m^2}(t,\mu)(z_1,z_2) = \frac{\delta^2\Phi}{\delta m^2}(\bar\rho_t)(m^{(1)}(t,\bar\rho_0,z_1), m^{(1)}(t,\bar\rho_0,z_2)) + \frac{\delta\Phi}{\delta m}(\bar\rho_t)(m^{(2)}(t,\bar\rho_0,z_1,z_2)), \quad (1.2.23)$$

where $m^{(1)}$ and $m^{(2)}$ are two solutions of some (explicit) Cauchy problem. Likewise, we can compute the second derivative of $\frac{\delta^2\mathcal{U}}{\delta m^2}$. The proofs of Equations (1.2.21), (1.2.22), and (1.2.23) are quite computationally involved, but can be done explicitly.

Notice how, for the case of observable $\Phi(\mu) := \int F(x)d\mu(x)$, that using (1.2.20) into (1.2.23), for a bounded function F, we can obtain via (1.2.21) and (1.2.22)

$$\left|\mathbb{E}\Phi\left(\mu_t^N\right) - \Phi\left(\bar\rho_t\right)\right| \leqslant \frac{C}{N}$$

The type of result obtained here is obviously weaker than the ones obtained in Wasserstein distance for instance, as the convergence only concerns an observable Φ, the regularity of which actually plays a huge role in the bounds one is able to obtain on the linear functional derivatives.

However, this general semigroup approach allows for more general types of processes, for instance for jump processes [140] or processes not driven by a Brownian motion but by an α-stable process [42]. See also [57] for a uniform in time result.

1.3 Contributions

Throughout this document, we will thus try and prove quantitative and uniform in time propagation of chaos in various settings. These could roughly be divided into three main groups, based on the type of difficulty they present.

- The kinetic (or degenerate noise) setting : the Brownian motion does not act on all coordinates.
- The singular setting : the kernel K in (1.1.1) is singular in 0.
- With incomplete interactions : particles no longer interact with every other particles.

1.3.1 Kinetic setting or degenerate noise

We consider here the second order dynamics (also referred to as the *(underdamped) Langevin dynamics*)

$$\begin{cases} dX_t^{i,N} = V_t^{i,N} dt \\ dV_t^{i,N} = \sqrt{2\sigma} dB_t^i + \frac{1}{N}\sum_{j=1}^N K\left(X_t^{i,N} - X_t^{j,N}\right) dt. \end{cases} \tag{1.3.1}$$

where $X_t^{i,N}$ and V_t^i denote respectively the position and velocity of particle i in $\mathbb{R}^d$. Notice that, for $\sigma = 0$, we obtain the classical deterministic Newton dynamics, which is fundamental in physics. In many applications, the function K can be polynomial (granular media), Newtonian (interacting stellar) or Coulombian (charged matter). See for instance [117] for an english translation of Paul Langevin's landmark paper on the physics behind the standard underdamped Langevin dynamics.

Observe that the Brownian motions only act on the velocities, hence the expression *degenerate noise*, and this fact will have several consequences. Mainly, proving propagation of chaos and/or long-time behavior combines the difficulties of dealing with nonlinear processes and hypoelliptic diffusions (see Chapter 2).

Remark that (1.1.1) can be seen as a small mass limit of (1.3.1), provided we add a friction term $-\gamma V_t^{i,N} dt$, with $\gamma > 0$ scaling properly, in the dynamics of $V_t^{i,N}$. See for instance Section 2.2.4 of [116].

Vlasov-Fokker-Planck equation. Our first work [83] concerns the proof of uniform in time propagation of chaos for a particle system of the form

$$\begin{cases} dX_t^{i,N} = V_t^{i,N} dt \\ dV_t^{i,N} = \sqrt{2\sigma} dB_t^i - V_t^{i,N} dt - \nabla U(X_t^{i,N}) dt - \frac{1}{N}\sum_{j=1}^N \nabla W(X_t^{i,N} - X_t^j) dt, \end{cases} \tag{1.3.2}$$

with $X_t^{i,N} \in \mathbb{R}^d$ and $V_t^i \in \mathbb{R}^d$, U is a non-convex confining potential such that ∇U is Lipschitz continuous, and W is an interaction potential such that ∇W is Lipschitz continuous with a sufficiently small Lipschitz coefficient (this is necessary for the uniformity in time as it ensures the uniqueness of the invariant measure).

We consider to this end a coupling method suggested by A. Eberle [66] known as *reflection* coupling, originally designed to deal with the long time behavior of general diffusion processes, as in [66, 69], and later extended to show uniform in time propagation of chaos in a mean-field system in [64]. This coupling method relies on the careful construction of a semimetric, taking into account the various behaviors of the particles. This back and forth between the calculations and the understanding of the processes is quite interesting as it yields a rather "robust" method.

You may find the article, containing the long-time convergence of the non linear limit and the uniform in time propagation of chaos of the particle system, as well as the precise description of the coupling method and its benefits, in Chapter 2.

The FitzHugh-Nagumo model. Using the same ideas, it is also possible to prove uniform in time propagation of chaos for a system of neurons in the brain satisfying the FitzHugh-Nagumo model

$$\left\{ \begin{array}{l} dX_t^{i,N} = (X_t^{i,N} - (X_t^{i,N})^3 - C_t^{i,N} - \alpha)dt + \frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N} - Z_t^{j,N}) + \sigma_X dB_t^{i,X} \\ dC_t^{i,N} = (\gamma X_t^{i,N} - C_t^{i,N} + \beta)dt + \frac{1}{N}\sum_{j=1}^N K_C(Z_t^{i,N} - Z_t^{j,N}) + \sigma_C dB_t^{i,C}, \end{array} \right. \tag{1.3.3}$$

where $X_t^{i,N}$ is the membrane potential of particle i and $C_t^{i,N}$ is a recovery variable, called the adaptation variable, and where we allow either σ_X or σ_C to be equal to 0. Once again, K_X and K_C are interaction kernels assumed to be Lipschitz continuous with a sufficiently small Lipschitz constant. Beyond the result of uniform in time propagation of chaos for the FitzHugh-Nagumo model, which is in itself an interesting result, the present work is also a testament to the robustness of the coupling method. This is joint work with Laetitia Colombani[2] [54].

Since the calculations are very similar to the ones of Chapter 2, you will find in Chapter 3 a summary of the model and the results.

1.3.2 Singular interactions

The case of singular kernel K holds importance because of the number of applications and links with other research areas. Consider the particle system in dimension d

$$dX_t^{i,N} = \frac{1}{N}\sum_{j\neq i} \mathbb{M}\nabla g\left(X_t^{i,N} - X_t^{j,N}\right) dt + \sqrt{2\sigma} dB_t^i, \tag{1.3.4}$$

where $\mathbb{M}$ is a constant $d\times d$ matrix either antisymmetric (in which case we speak of a conservative system) or $\mathbb{M} = \pm\mathbb{I}$ (dissipative, and resp. attractive or repulsive interactions), and assume there is $s \in [0, d[$ such that

$$g(x) = \left\{ \begin{array}{ll} -\log|x|, & \text{if } s = 0, \\ |x|^{-s}, & \text{if } s > 0. \end{array} \right.$$

These type of interactions are often referred to as *Riesz interactions*. The specific case of $s = d-2$ is known as the *Coulomb interaction*. We refer to [45, 98, 33], and more specifically to the recent review [120] and references therein, but to name a few cases of interest :

- $d \geqslant 2, s = d-2, \mathbb{M} = \mathbb{I}$ yields an approximation of the second-order Newtonian dynamics,
- $d = 1, s = 0, \mathbb{M} = -\mathbb{I}$ has links to random matrix theory (see (1.3.5) below and Chapter 5),
- $d = 2, s = 0, \mathbb{M} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ corresponds to the 2D vortex model (see below and Chapter 4),
- $d = 2, s = 0, \mathbb{M} = -\mathbb{I}$ for the Ginzburg-Landau vortices (see for instance [154, 158]),

[2]University of Bern, Institute of Mathematical Statistics and Actuarial Science, CH-3012 Bern Switzerland. Email : laetitia.colombani[AT]stat.unibe.ch

- $d = 2, s = 0, \mathbb{M} = \mathbb{I}$ corresponds to the Patlak-Keller-Segel model for chemotaxis in biology (see for instance [33]).

Among the most recent works, let us mention [30, 32, 143], which collectively seem to suggest that all results for mean-field limits have been obtained in the case $d \geqslant 1$ and $s < d$. However, our contributions yield both a *quantitative* and a *uniform in time* result, which complement said works and other such as [152].

The 2D vortex model. The first singular model which motivates us is the 2D vortex model. Consider, in Equation (1.1.1), the Biot-Savart kernel $K(x) = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$ in dimension 2 on the torus. The nonlinear limit satisfies the vorticity equation, which arises when considering the curl of the 2D incompressible Navier-Stokes system. In the deterministic case, a partial result may be found in [155]. With noise, for a convergence without rate, a first result appeared in [76], relying on proving that close encounters of particles are rare and that the possible limits of the particle system are made of solutions of the nonlinear SDE. In the recent work [99], P.-E. Jabin and Z. Wang have proven that propagation of chaos still holds on the torus with a quantitative rate. Building upon their work, we obtain a similar result in [85], but now uniform in time.

Their approach consists in computing the time evolution of the relative entropy of ρ_t^N with respect to $\bar{\rho}_t^{\otimes N}$, and then using an integration by parts to deal with the singularity of K thanks to the regularity of the probability density $\bar{\rho}_t$. This idea comes from the observation that the Biot-Savart kernel can be explicitly written as the divergence of a bounded matrix field. In order to improve this argument to get uniform in time propagation of chaos, our main contribution is the proof of time-uniform bounds for the density of $\bar{\rho}$, from which a time-uniform logarithmic Sobolev inequality is deduced. From the latter, in the spirit of the work of F. Malrieu [133] in the smooth and convex case, the Fisher information appearing in the entropy dissipation yields a control on the relative entropy itself, inducing the time uniformity. Thanks to the fast decay to 0 of the derivatives of the nonlinear limit, no smallness assumption on the interaction is required. We thus prove uniform in time propagation of chaos under a set of assumptions satisfied in particular by the Biot-Savart kernel.

The method and results are discussed in Chapter 4.

The 1D log and Riesz gases. We are then interested in the case of the (generalized) Dyson Brownian motion in dimension 1

$$dX_t^{i,N} = \sqrt{\frac{2\sigma}{N}} dB_t^i - \lambda X_t^{i,N} dt + \frac{1}{N}\sum_{j\neq i}\frac{1}{X_t^{i,N} - X_t^{j,N}} dt. \tag{1.3.5}$$

The dynamics above are satisfied, for $\lambda = 0$, by the eigenvalues of an $N \times N$ Hermitian matrix valued Brownian motion, as observed by F. J. Dyson in 1962 [65]. For $\lambda > 0$, it correspond to the eigenvalues of an $N \times N$ Hermitian matrix valued Ornstein-Uhlenbeck process [47]. This motivates us to consider, in a slightly more general setting, the 1D N-particle system in mean field interaction

$$dX_t^{i,N} = \sqrt{2\sigma_N} dB_t^i - \lambda X_t^{i,N} dt - \frac{1}{N}\sum_{j\neq i} V'(X_t^{i,N} - X_t^{j,N}) dt, \tag{1.3.6}$$

where $\sigma_N \xrightarrow[N\to\infty]{} 0$ and there exists $\alpha \in [1, 2[$, such that for all $x \in \mathbb{R}^*$, $V'(x) = -\frac{x}{|x|^{\alpha+1}}$.

The approaches discussed above fail on this system : a more PDE-oriented approach, which would rely on the regularity of the nonlinear limit $\bar{\rho}_t$ and on bounds on its derivatives, is hindered by the fact that there is no Laplacian term in the limit as $\sigma_N \to 0$. Likewise, since the noise vanishes and the interaction is singular, coupling methods also fail. A qualitative result of propagation of chaos for the (generalized) Dyson Brownian motion was proved in [151, 43, 125] using the tightness of the sequence of empirical measures and the uniqueness of the limit.

We provide a uniform in time quantitative result for $\alpha \in [1, 2[$ by coupling an N particle system with an M particle system for all $N, M \geqslant 0$. We succeed in proving, making full use of the fact that in dimension 1 the particles remain ordered and the optimal coupling in the Wasserstein distance is explicit, that any independent sequence of empirical measures is a Cauchy sequence in some sense. Then, independence ensures that the limit is an almost surely constant random variable in the space of probability measure, which can then be identified as a weak solution of the nonlinear limit we expect.

This method, which to the best of our knowledge had not been used before, is fully described in [84], and so far seems restricted to the dimension 1 case. You may find the complete description of the results in Chapter 5.

1.3.3 Incomplete interactions and non-exchangeability

Finally, we consider cases where the particles do not interact with every other particles, but according to a graph. This breaks the assumption of exchangeability of the system, as *a priori* some may interact with more than other, and the structure of the underlying graph in such settings influences greatly the behavior of the system. These systems of interacting particles find many applications in Physics, Biology, Economics or Social Sciences (see [97, 110, 173] and references therein), and model agents that are not identical.

Here, we focus on two problems concerning, first, the almost sure (with respect to the random graph) quantitative and uniform in time mean-field limit in graphs, and then a type of numerical scheme involving a random batch method.

Mean-field limit in graphs. We observe that the reflection coupling makes it possible to deal with particles interacting according to a graph

$$dX_t^{i,N} = F\left(X_t^{i,N}, \omega_i\right) dt + \frac{\alpha_N}{N}\sum_{j=1}^{N} \xi_{i,j}^{(N)} \Gamma\left(X_t^{i,N}, \omega_i, X_t^{j,N}, \omega_j\right) dt + \sqrt{2\sigma} dB_t^i, \tag{1.3.7}$$

where $\xi^{(N)} = \left(\xi_{i,j}^{(N)}\right)_{1\leqslant i,j\leqslant N}$, $\xi_{i,j}^{(N)} \in \{0, 1\}$ represents the graph, $\{\omega_i\}_{1\leqslant i\leqslant N}$ is an environmental disorder, $(\alpha_N)_{N\geqslant 1}$ is a scaling parameter, $F : \mathbb{R}^d \times \mathbb{R}^{d'} \mapsto \mathbb{R}^d$ is a Lipschitz continuous outside force and $\Gamma : \left(\mathbb{R}^d \times \mathbb{R}^{d'}\right)^2 \mapsto \mathbb{R}^d$ is a Lipschitz continuous interaction. Assuming there is $p \in [0, 1]$ such that $\sup_{1\leqslant i\leqslant N} \left|\alpha_N \frac{\sum_{j=1}^N \xi_{i,j}^{(N)}}{N} - p\right| \xrightarrow[N\to\infty]{a.s} 0$, which is true for Erdős-Rényi graphs, the empirical distribution μ_t^N converges in Wasserstein distance to the nonlinear limit $\bar{\rho}_t$. We prove furthermore that this mean-field limit for (1.3.7) is uniform in time.

This is joint work with Christophe Poquet[3] [115], available in Chapter 6.

[3]Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, F-69622 Villeurbanne, France. Email : poquet[AT]math.univ-lyon1.fr

The random batch method. The numerical scheme associated to the overdamped Langevin dynamics in mean field interaction for a timestep δ is the following

$$\begin{cases} X_{k+1}^{i,N} = X_k^{i,N} - \delta\nabla U(X_k^{i,N}) - \frac{\delta}{N-1}\sum_{j\neq i}\nabla W(X_k^{i,N} - X_k^{j,N}) + \sqrt{2\sigma\delta}G_k^i, \\ G_k^i \ \text{ i.i.d } \ \sim \mathcal{N}(0,1). \end{cases} \tag{1.3.8}$$

Notice, however, that the numerical complexity of each time step is $O(N^2)$, corresponding to the computation of each interaction. To decrease this complexity, we choose to consider a *Random Batch Method* (RBM). The idea is, instead of considering that each particle interacts with every other particles, we divide at each time step the system into smaller batches and compute the interactions only within these batches. Consider, for a timestep k, a partition $\mathcal{P}_k = \left(\mathcal{P}_k^1, ..., \mathcal{P}_k^{N/p}\right)$ of $\{1, ..., N\}$ into N/p subsets of size p, and define

$$\mathcal{C}_k^i = \left\{ j \in \{1, .., N\} \text{ s.t. } \exists l \in \{1, ..., N/p\}, i, j \in \mathcal{P}_k^l \right\}.$$

In other words, $\mathcal{C}_k^i$ is the set of indexes that are in the same subset as i at timestep k, with the convention $i \in \mathcal{C}_k^i$. The partition is chosen randomly and uniformly at each time step, and yields the following numerical scheme

$$\begin{cases} Y_{k+1}^i = Y_k^i - \delta\nabla U(Y_k^i) - \frac{\delta}{p-1}\sum_{j\neq i\in\mathcal{C}_k^i}\nabla W(Y_k^i - Y_k^j) + \sqrt{2\sigma\delta}G_k^i, \\ G_k^i \ \text{ i.i.d } \ \sim \mathcal{N}(0,1). \end{cases} \tag{1.3.9}$$

The numerical scheme is now of complexity $O(Np)$ at each time step, and the convergence of (1.3.9) as N goes to infinity and δ goes to 0 to the invariant distribution of the nonlinear limit SDE can for instance be found in [177].

It is interesting to notice that this RBM "adds" randomness to the system. This prompts us to analyze the effect of the method on systems that exhibit a phase transition since, if the system admits a critical temperature, this temperature should intuitively decrease due to the added randomness.

To get a better grasp on this phenomenon, we start by studying one of the simplest model admitting a phase transition : the *Curie-Weiss model*. For N spins, given by a configuration $\sigma = (\sigma_1, ..., \sigma_N)$, let $\Omega_N = \{-1, 1\}^N$ be the set of possible spin configurations. On this system we consider the following Hamiltonian

$$\forall \sigma \in \Omega_N, \quad H_N(\sigma) = -\frac{1}{2N}\sum_{i,j}\sigma_i\sigma_j.$$

Intuitively, each spin will tend to align with the others. The evolution for $(\sigma(n))_n$ in Ω_N is the following: at each discrete time step, a spin is chosen uniformly among the N possibilities. Let us denote i this spin, and $\sigma' = (\sigma'_1, ..., \sigma'_N)$ the configuration such that for all $j \neq i$, $\sigma'_j = \sigma(n)_j$, and $\sigma'_i = -\sigma(n)_i$. We accept σ' as the next step of $\sigma(n)$ with probability $\exp\left(-\beta(H_N(\sigma') - H_N(\sigma))_+\right)$. This parameter β will be called the inverse temperature.

In reality, this system can be entirely studied in terms of its *magnetization*, i.e the empirical mean of spins, $m_N(n) := \frac{1}{N}\sum_{i=1}^N \sigma(n)_i$.

The nonlinear limit as N goes to infinity of this system admits a phase transition : for $\beta > 1$, it admits three equilibrium points for the magnetization, and for $\beta \leqslant 1$ only one.

Now, instead of considering this well-known system, we choose a modification with random batches in which at each time step the chosen spin no longer evolves according to the entire system, but according to a subset of p spins containing the chosen spin. We can prove that the critical parameter is not 1, but $\beta_c > 1$, and that $\beta_c = 1 + \sqrt{\frac{2}{p\pi}} + o\left(\frac{1}{\sqrt{p}}\right)$.

The second model we consider is the overdamped Langevin diffusion in a double-well potential, for which it is known there is a phase transition, see for instance [168]. Similarly, we seek to understand how the critical parameter for this phase transition evolves with the size of the random batches.

This is ongoing work, the state of which you may find in Chapter 7.

All chapters are independent, with the exception of Chapter 3 which relies on Chapter 2.

1.4 Perspectives

Remaining singular problems. There currently remain a few cases of Riesz interactions for which quantitative and uniform in time propagation of chaos has not yet been obtained. However, some works are apparently in preparation and should get, using the modulated free energy approach, the desired result in some, if not most, of these remaining cases. This should mostly conclude the study of (1.3.4), although there may remain some technical assumptions to deal with. For instance, if the results hold on the torus, can we extend them to the entire space $\mathbb{R}^d$? A similar technical impediment can be found in our work on the 2D vortex model. This question of compactness becomes crucial when dealing with second order systems : we may bound the positions by working on the torus, but not the velocities. Obtaining quantitative and uniform in time results in the kinetic setting for singular interactions thus seems to be the next problem to tackle.

In both our work on the 2D vortex model and in [152], the uniform in time estimates are a consequence of the (sufficiently fast) decay of the L^∞ norms of the space derivatives of $\bar{\rho}_t$ (or a quantity involving said derivatives). See Chapter 4. Although it is not exactly how we prove these decays, we understand that it is linked to the fact that, on the torus for a divergence free kernel K, the invariant distribution is the Lebesgue measure, and likewise on the entire space with repulsive interaction the solution $\bar{\rho}_t$ goes to 0. Furthermore, we also work in a compact set in order to bound $\log \bar{\rho}_t$ and its derivatives which appear in the relative entropy.

To dispense with the need to be in a compact space, and in order to still be able to use the sufficiently fast decay of the nonlinear solution, we would like to try and work in weighted L^p spaces, changing the measure of reference to the stationary distribution. The solution $\bar{\rho}_t$ should stay bounded with respect to this invariant measure, and the bounds of the derivatives of $\frac{d\bar{\rho}_t}{d\bar{\rho}_\infty}$ should vanish. This may require the addition of a confining potential. Note however that working with this *a priori* unknown measure may prove to be difficult for lack of known properties...

Using probabilistic tools. The most exciting recent results concerning quantitative propagation of chaos for singular kernels use PDE-oriented methods (Modulated energy, BBGKY hierarchies...). As we hope to convince the reader, there is a real interest in obtaining probabilistic proofs (understand : proofs staying at the level of the SDE) as these may provide a better understanding of the processes, and this better understanding may then translate into robust and intuitive methods for other problems. See chapter 2 of [171], or how the proof of Chapter 2 adapts to Chapter 3. Thus, beyond the result, we would like to obtain coupling proofs for singular kernels to both deepen our understanding of the phenomenon and ensure a diversity of proofs.

For instance, the result in Chapter 5 which concerns propagation of chaos in the singular 1D case, indeed relies on the convexity of the Riesz interaction in dimension 1, which is arguably its biggest flaw, but also (and mainly) on the control of close encounters of particles, similarly as the proofs using tightness argument. An idea could thus be to combine the coupling method of Chapter 5, considering empirical measures, with the construction of a semimetrics of Chapter 2, that would thus involve Lyapunov functions. Close interactions would be controlled in the case of repulsive interactions in the kinetic case and regardless of the dimension.

Sharp rate of convergence. For a long time it had been assumed that the rate of convergence in N (in Wasserstein distance for instance) of $N^{-1/2}$ obtained by the usual coupling methods or via entropy dissipation was optimal, as it relates to the speed of convergence of the central limit theorem. However as we mentioned, in a recent work, D. Lacker [111] proved, using BBGKY hierarchies for the relative entropy, that the optimal rate should rather be N^{-1}.

So far, the difference in these rates of convergence has been justified by the fact that D. Lacker uses a local approach through this BBGKY hierarchies, while previous methods deduced the convergence of the law of a subset of k particles from the N particle system by subadditivity, thus considering a global approach to the problem.

Understanding why coupling methods fail in reaching that optimal rate and how we should adapt them is a direction we have in mind. Likewise, we are discussing with D. Lacker of the possibility of combining the BBGKY approach [111] with the calculations done in Chapter 4, in order to obtain this optimal rate of convergence for singular interaction kernels.

And other problems. The system (1.1.1) corresponds to a very specific (yet absolutely crucial) type of particle system in mean-field interaction corresponding to pairwise interactions. It would be interesting to go beyond and study

$$dX_t^{i,N} = b\left(X_t^{i,N}, \frac{1}{N}\sum_{j=1}^N X_t^{j,N}\right)dt + \sqrt{2\sigma}dB_t^i, \tag{1.4.1}$$

for a given $b : \mathbb{R}^d \times \mathcal{P}\left(\mathbb{R}^d\right) \mapsto \mathbb{R}^d$. The dynamics (1.1.1) correspond to the specific case $b(x,\mu) = \int K(x-y)d\mu(y)$. This type of system is currently considered for its applications, such as the training of neural networks, and has been studied using weak derivatives (see [96] or the recent [49] for a uniform in time result) or BBGKY hierarchies [111].

Furthermore, as the effective dynamics in Chapter 7 suggests, it may be interesting to take a look at systems in which the interaction (equivalently the nonlinearity) appears in the diffusion coefficient.

And finally, due to its wide range of applications, there is no doubt that the study of large systems of interacting agents will continue to provide microscopic problems that must be related to their mesoscopic limit.

Bonne lecture !

Part I

Degenerate noise and kinetic setting

Chapter 2

Convergence rates for the Vlasov-Fokker-Planck equation and uniform in time propagation of chaos in non convex cases

J'ai toujours eu du bol avec l'aléatoire, moi.

Lionnel Astier, *Kaamelott*, Livre II, *Le Sort Perdu* (2005), written by Alexandre Astier.

Outline of the current chapter

2.5 ∇U locally Lipschitz continuous 71

Joint work with Arnaud Guillin and Pierre Monmarché.

Published as: Arnaud Guillin, Pierre Le Bris, and Pierre Monmarché. "Convergence rates for the Vlasov-Fokker-Planck equation and uniform in time propagation of chaos in non convex cases." *Electronic Journal of Probability* 27 (2022): 1-44. [83]

Abstract: We prove the existence of a contraction rate for Vlasov-Fokker-Planck equation in Wasserstein distance, provided the interaction potential is Lipschitz continuous and the confining potential is both (locally) Lipschitz continuous and greater than a quadratic function, thus requiring no convexity conditions. Our strategy relies on coupling methods suggested by A. Eberle [66] adapted to the kinetic setting enabling also to obtain uniform in time propagation of chaos in a non convex setting.

Acknowledgements: This work has been funded by the project EFI ANR-17-CE40-0030.

2.1 Introduction

2.1.1 Framework

Let U and W be two functions in $\mathcal{C}^1\left(\mathbb{R}^d\right)$. We consider the Vlasov-Fokker-Planck equation:

$$\partial_t \nu_t (x,v) = -\nabla_x \cdot (v \nu_t (x,v)) + \nabla_v \cdot ((v + \nabla U (x) + \nabla W * \mu_t (x)) \nu_t (x,v) + \nabla_v \nu_t (x,v)) , \quad (2.1.1)$$

where $\nu_t(x,v)$ is a probability density in the space of positions $x \in \mathbb{R}^d$ and velocities $v \in \mathbb{R}^d$,

$$\mu_t (x) = \int_{\mathbb{R}^d} \nu_t (x, dv)$$

is the space marginal of ν_t and

$$\nabla W * \mu_t(x) = \int_{\mathbb{R}^d} \nabla W(x-y) \mu_t(dy).$$

It has the following probabilistic counterpart, the non linear stochastic differential equation of *McKean-Vlasov* type, i.e. ν_t is the density of the law at time t of the $\mathbb{R}^{2d}$-valued process $(X_t, V_t)_{t \geqslant 0}$ evolving as the mean field SDE (diffusive Newton's equations)

$$\begin{cases} dX_t = V_t dt \\ dV_t = \sqrt{2} dB_t - V_t dt - \nabla U (X_t) dt - \nabla W * \mu_t (X_t) dt \\ \mu_t = \text{Law} (X_t) . \end{cases} \quad (2.1.2)$$

Here, $(X_t, V_t) \in \mathbb{R}^d \times \mathbb{R}^d$, $(B_t)_{t \geqslant 0}$ is a Brownian motion in dimension d on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and μ_t is the law of the position X_t. The symbol ∇ refers to the gradient operator, and the symbol $*$ to the operation of convolution.

Both in the probability and in the partial differential equation community, existence and uniqueness of McKean-Vlasov processes have been well studied. See [135, 78, 162] for some historical milestones. In the specific case of (2.1.1) and (2.1.2), under the assumptions on U

and W introduced in the next section, existence and uniqueness follow from [137] for square integrable initial data.

A related process is the N particles system in $\mathbb{R}^d$ in mean field interaction

$$\forall i \in [\![1, N]\!], \qquad \begin{cases} dX_t^i &= V_t^i dt, \\ dV_t^i &= \sqrt{2} dB_t^i - V_t^i dt - \nabla U\left(X_t^i\right) dt - \dfrac{1}{N}\displaystyle\sum_{j=1}^N \nabla W\left(X_t^i - X_t^j\right) dt, \end{cases} \tag{2.1.3}$$

where X_t^i and V_t^i are respectively the position and the velocity of the i-th particle, and $\left(B_t^i, 1 \leqslant i \leqslant N\right)$ are independent Brownian motions in dimension d. One can see equation (2.1.3) as an approximation of equation (2.1.2), where the law μ_t is replaced by the empirical measure $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$.

It is well known, at least in a non kinetic setting [137, 162], that, under some weak conditions on U and W, μ_t^N converges in some sense toward the law μ_t of X_t solution of (2.1.2). This phenomenon has been stated under the name *propagation of chaos*, an idea motivated by M. Kac [106], and greatly developed by A.S. Sznitman [162]. See the recent reviews on propagation of chaos [45, 46] and references therein for an overview on the subject.

In statistical physics, (2.1.3) is a Langevin equation that describes the motion of N particles subject to damping, random collisions and a *confining potential* U and interacting with one another through an *interaction potential* W, which can be polynomial (granular media), Newtonian (interacting stellar) or Coulombian (charged matter). See for instance [117] for an english translation of P. Langevin's landmark paper on the physics behind the standard underdamped Langevin dynamics. Therefore, Equation (2.1.1) has the following natural interpretation: the solution ν_t is the density of the law at time t of the process $(X_t, V_t)_{t\geqslant 0}$ evolving according to (2.1.2), and thus describes the limit dynamic of a cloud of (charged) particles. In particular, it holds importance in plasma physics, see [172].

More recently, mean-field processes such as (2.1.3) have drawn much interest in the analysis of neuron networks in machine learning [51, 50]. In this context of stochastic algorithms, it is known that the underdamped Langevin dynamics (not necessarily with mean-field interactions) can converge faster than the overdamped (i.e non kinetic) Langevin dynamics [51, 88] toward its invariant measure. For example, the results on (2.1.2) could be applied to study the convergence of the Hamiltonian gradient descent algorithm for the overparametrized optimization as done in [108] for Generative Adversarial Network training.

The goal of the present work is twofold. We are interested, first, in the long-time convergence of the solution of (2.1.2) toward an equilibrium and, second, to a uniform in time convergence as $N \to +\infty$ of (2.1.3) toward (2.1.2). It is well known that such results cannot hold in full generality, as the non-linear equation (2.1.1) may have several equilibria. Here we will consider cases where the interaction is sufficiently small for the non-linear equilibrium to be unique and globally attractive, and for the propagation of chaos to be uniform in time.

There are various methods to study the long time behavior of kinetic type processes, such as Lyapunov conditions or hypocoercivity, and we will discuss these approaches and compare them with our results later on. We rely here on coupling methods following the guidelines of A. Eberle *et al.* in [68] where the convergence to equilibrium is established for (2.1.2) without interaction, and also extend the approach to handle only locally Lipschitz coefficient. In a second part, we also use reflection couplings (see [64]) for the propagation of chaos property.

Let us briefly describe the coupling method. 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 of (2.1.2) that have a trend to get closer with time. Have (X_t, V_t) be a solution of (2.1.2) driven by some Brownian motion $(B_t)_{t\geqslant 0}$ and let (X'_t, V'_t) solves

$$\begin{cases} dX'_t = V'_t dt \\ dV'_t = \sqrt{2} dB'_t - V'_t dt - \nabla U\left(X'_t\right) dt - \nabla W * \mu_t\left(X'_t\right) dt \\ \mu'_t = \text{Law}\left(X'_t\right). \end{cases}$$

with $(B'_t)_{t\geqslant 0}$ a d-dimensional Brownian motion. A coupling of (X, V) and (X', V') then follows from a coupling of the Brownian motions B and B'. Choosing $B = B'$ yields the so-called *synchronous* coupling, for which the Brownian noise cancels out in the infinitesimal evolution of the difference $(Z_t, W_t) = (X_t - X'_t, V_t - V'_t)$. In that case the contraction of a distance between the processes can only be induced by the deterministic drift, as in [21]. Such a deterministic contraction only holds under very restrictive conditions, in particular U should be strongly convex. Nevertheless, in more general cases, the calculation of the evolution of Z_t and W_t (see Section 2.3.1 below) shows that there is still some deterministic contraction when $Z_t + W_t = 0$. We can therefore use a synchronous coupling in the vicinity of this subspace.

Outside of $\{(z, w) \in \mathbb{R}^{2d}, z + w = 0\}$, 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 in B and B' being *antithetic* (i.e $B'_t = -B_t$) in the direction of space given by the difference of the processes, and synchronous in the orthogonal direction. In other words, writing

$$e_t = \begin{cases} \frac{Z_t + W_t}{|Z_t + W_t|} & \text{if } Z_t + W_t \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

we consider $dB'_t = \left(Id - 2 e_t e_t^T\right) dB_t$. Levy's characterization then ensures that it is indeed a Brownian motion.

Finally we construct a Lyapunov function H to take into account the trend of each process to come back to some compact set of $\mathbb{R}^{2d}$. We are then led to the study of a suitable distance between the two processes, which will be of the form $\rho_t := f(r_t)(1+\epsilon H(X_t, V_t)+\epsilon H(X'_t, V'_t))$, with $r_t = \alpha|Z_t| + |Z_t + W_t|$, where $\alpha, \epsilon > 0$ and the function f are some parameters to choose. More precisely, we have to choose these parameters carefully in order for $\mathbb{E}\rho_t$ to decay exponentially fast. This leads to several constraints on α, ϵ and on the parameters involved in the definition of f, and we have to prove that it is possible to meet all these conditions simultaneously. For the sake of clarity, in fact, we present the proof in a different order, namely we start by introducing very specific parameters and, throughout the proof, we check that our choice of parameters implies the needed constraints.

The study of the limit $N \to +\infty$ is based on a similar coupling, except that we couple a system of N interacting particles (2.1.3) with N independent non-linear processes (2.1.2).

The next subsections describe our main results and compare them to the few existing ones in the literature. Section 2.2 presents the precise construction of the aforementioned *ad hoc* Wasserstein distance. The proof of the long time behavior of the Vlasov-Fokker-Planck equation when confinement and interaction coefficient are Lipschitz continuous is done in Section 2.3, whereas the propagation of chaos property is proved in Section 2.4. An appendix gathers technical lemmas and the modifications of the main proofs when the confinement is only supposed locally Lipschitz continuous.

2.1.2 Main results

For μ and ν two probability measures on $\mathbb{R}^{2d}$, denote by $\Pi(\mu,\nu)$ the set of couplings of μ and ν, i.e. the set of probability measures Γ on $\mathbb{R}^{2d}\times\mathbb{R}^{2d}$ with $\Gamma(A\times\mathbb{R}^{2d})=\mu(A)$ and $\Gamma(\mathbb{R}^{2d}\times A)=\nu(A)$ for all Borel set A of $\mathbb{R}^{2d}$. We will define L^1 and L^2 Wasserstein distances as

$$\mathcal{W}_1(\mu,\nu)=\inf_{\Gamma\in\Pi(\mu,\nu)}\int\left(|x-\tilde{x}|+|v-\tilde{v}|\right)\Gamma\left(dxdvd\tilde{x}d\tilde{v}\right),$$

$$\mathcal{W}_2(\mu,\nu)=\left(\inf_{\Gamma\in\Pi(\mu,\nu)}\int\left(|x-\tilde{x}|^2+|v-\tilde{v}|^2\right)\Gamma\left(dxdvd\tilde{x}d\tilde{v}\right)\right)^{1/2}.$$

Our main results will be stated in terms of these distances, even if we work and get contraction in the Wasserstein distance defined with the aformentioned ρ. Let us detail the assumptions on the potentials U and W.

Assumption 2.1. *The potential U is non-negative and there exist $\lambda>0$ and $A\geqslant 0$ such that*

$$\forall x\in\mathbb{R}^d,\qquad \frac{1}{2}\nabla U(x)\cdot x\geqslant\lambda\left(U(x)+\frac{|x|^2}{4}\right)-A. \tag{2.1.4}$$

The condition (2.1.4) implies that the force $-\nabla U$ has a confining effect, bringing back particles toward some compact set. It implies the following:

Lemma 2.1.1. *If Assumption 2.1 holds, then there exists $\tilde{A}\geqslant 0$ such that for all $x\in\mathbb{R}^d$,*

$$U(x)\geqslant\frac{\lambda}{6}|x|^2-\tilde{A}. \tag{2.1.5}$$

The proof is postponed to Appendix B.1.1. In particular, it implies that U goes to infinity at infinity and is bounded below. Since only the gradient of U is involved in the dynamics, the condition $U\geqslant 0$ is thus not restrictive as it can be enforced without loss of generality by adding a sufficient large constant to U. This condition is added in order to simplify some calculations.

We will also assume that the potential U satisfies one of the two following conditions :

Assumption 2.2. *There is a constant $L_U>0$ such that*

$$\forall x,y\in\mathbb{R}^d\times\mathbb{R}^d,\qquad |\nabla U(x)-\nabla U(y)|\leqslant L_U|x-y|.$$

Assumption 2.3. *There exist $L_U>0$ and a function $\psi:\mathbb{R}^d\mapsto\mathbb{R}$ such that*

$$\forall x,y\in\mathbb{R}^d\times\mathbb{R}^d,\qquad |\nabla U(x)-\nabla U(y)|\leqslant\left(L_U+\psi(x)+\psi(y)\right)|x-y|,$$

and

$$\forall x\in\mathbb{R}^d,\qquad 0\leqslant\psi(x)\leqslant L_\psi\sqrt{\lambda|x|^2+24U(x)},$$

where $L_\psi>0$ is sufficiently small in the sense that

$$L_\psi\leqslant c_\psi(L_U,\lambda,\tilde{A},d,a),$$

where c_ψ is an explicit function given below in (2.5.9)*, L_U is given in Assumption 2.2, λ by Assumption 2.1, $\tilde{A}$ by Lemma 2.1.1, d is the dimension and a is a parameter such that* (2.5.1) *holds for some $\mathcal{C}^0$, namely is used to bound an initial moment.*

Obviously, Assumption 2.3 implies Assumption 2.2. We distinguish it as it yields simpler proofs. Actually, the proofs of our main results already rely on quite involved computations under Assumption 2.2, and thus for the convenience of the reader we present the proofs in this case with full details in a first step, and then in a second step we explain how the more general situation of Assumption 2.3 is tackled.

Remark 2.1.1. *In the literature, see for instance [146] or the recent [35], it is common to find the assumption U twice continuously differentiable with an hessian matrix satisfying*

$$||\nabla_x^2 U(x)|| \leqslant C(1 + |\nabla_x U(x)|), \tag{2.1.6}$$

where $||\nabla_x^2 U(x)||$ denotes the matrix norm of the hessian. Here, Assumption 2.3 together with Assumption 2.1 yield a stronger version of (2.1.6)*. Indeed, in dimension one for instance, we have*

$$\begin{aligned} |U''(x)| &= \lim_{y\to x} \frac{|U'(x) - U'(y)|}{|x-y|} \\ &\leqslant L_U + 2L_\psi \sqrt{\lambda |x|^2 + 24U(x)}. \end{aligned}$$

Using Assumption 2.1, we obtain the existence of a constant $\hat{A}$ such that

$$|U'(x)| \geqslant \frac{\lambda}{4}|x| - \hat{A},$$

which implies, once again using Assumption 2.1, that

$$|U'(x)| \left(\frac{4}{\lambda}|U'(x)| + \frac{4}{\lambda}\hat{A} \right) \geqslant |U'(x)||x| \geqslant U'(x)x \geqslant 2\lambda U(x) + \frac{\lambda}{2}|x|^2 - 2A.$$

In particular, there are constants c_1, c_2, c_3 and c_4 such that

$$c_1|U'(x)| + c_2 \geqslant \sqrt{c_3 U(x) + c_4|x|^2},$$

and therefore we obtain, for some constants C and η,

$$|U''(x)| \leqslant C(1 + \eta|U'(x)|),$$

where η has to be sufficiently small. This is no surprise as, in our work, we consider the "local Lipschitz condition" to be a perturbation of the global Lipschitz Assumption 2.2.

Example 2.1.1. *Assume d=1. The double-well potential given by*

$$U(x) = \begin{cases} \left(x^2 - 1\right)^2 & \text{if } |x| \leqslant 1, \\ \left(|x| - 1\right)^2 & \text{otherwise.} \end{cases}$$

satisfies Assumptions 2.1 and 2.2.

Example 2.1.2. *Likewise, we may consider $U(x) = \frac{1}{2}x^2 + \frac{3}{2}\cos(x)$ in dimension 1, which is neither strongly convex, nor strongly convex outside a ball, but satisfies Assumptions 2.1 and 2.2.*

Example 2.1.3. *Consider* $U(x) = \frac{1}{4}x^2 + \frac{b}{4}x^4$ *in dimension 1. We have*

$$\nabla U(x) \cdot x = \frac{x^2}{2} + bx^4 \geqslant \left(\frac{x^2}{4} + \frac{x^2}{4} + \frac{b}{4}x^4\right) = \left(U(x) + \frac{x^2}{4}\right),$$

hence U *satisfies Assumption 2.1.* U *is not Lipschitz continuous, however it satisfies*

$$\begin{aligned} |\nabla U(x) - \nabla U(y)| &= \frac{1}{2}|x-y| + b\left|x^3 - y^3\right| \\ &= \frac{1}{2}|x-y| + b|x-y|\left|x^2 + xy + y^2\right| \\ &\leqslant \frac{1}{2}|x-y| + \frac{3b}{2}|x-y|\left|x^2 + y^2\right| \\ &= \left(\frac{1}{2} + \psi(x) + \psi(y)\right)|x-y|, \end{aligned}$$

where,

$$\psi(x) = \frac{3b}{2}x^2 \leqslant \sqrt{b}\sqrt{24\frac{b}{4}x^4} \leqslant \sqrt{b}\sqrt{\lambda|x|^2 + 24U(x)}.$$

We then require b *to be sufficiently small for Assumption 2.3 to hold.*

Let us now give the assumption on the interaction potential.

Assumption 2.4. *The potential* W *is even, i.e.* $W(x) = W(-x)$ *for all* $x \in \mathbb{R}^d$*, in particular* $\nabla W(0) = 0$*. Moreover, there exists* $L_W < \lambda/8$ *(where* λ *is given in Assumption 2.1) such that*

$$\forall x, y \in \mathbb{R}^d \times \mathbb{R}^d, \qquad |\nabla W(x) - \nabla W(y)| \leqslant L_W|x-y|. \tag{2.1.7}$$

In particular $|\nabla W(x)| \leqslant L_W|x|$ *for all* $x \in \mathbb{R}^d$.

Here we consider an interaction force that is the gradient of a potential W, as we stick to the formalism of other related works (for instance [64]). Nevertheless, all the results and proofs still hold if ∇W is replaced by some $F : \mathbb{R}^d \mapsto \mathbb{R}^d$ satisfying the same conditions. The confinement potential may also be non gradient, however the fact that the confinement force ∇U is a gradient simplifies the construction of a Lyapunov function.

The condition $L_W \leqslant \lambda/8$ is related to the fact the interaction is considered as a perturbation of the non-interacting process studied in [68]. Therefore, ∇W has to be controlled by ∇U in some sense. Note that we immediately get the following bound on the non-linear drift:

Lemma 2.1.2. *Under Assumption 2.4, for all probability measures* μ *and* ν *on* $\mathbb{R}^d$ *and* $x, \tilde{x} \in \mathbb{R}^d$,

$$|\nabla W * \mu(x) - \nabla W * \nu(\tilde{x})| \leqslant L_W|x - \tilde{x}| + L_W\mathcal{W}_1(\mu, \nu).$$

See Appendix B.1.2 for the proof.

Example 2.1.4. *Assumption 2.4 is satisfied for an harmonic interaction* $W(x) = \pm L_W|x|^2/2$*, or a mollified Coulomb interaction for* $a, b > 0$ *and* $k \in \mathbb{N}^*$

$$W(x) = \pm\frac{a}{\left(|x|^k + b^k\right)^{\frac{1}{k}}}, \qquad i.e \qquad \nabla W(x) = \mp\frac{ax|x|^{k-2}}{\left(|x|^k + b^k\right)^{1+\frac{1}{k}}}.$$

The first of our main results concern the long-time convergence of the non-linear system (2.1.1).

Theorem 2.1.1. *Let U be continuously differentiable and satisfy Assumption 2.1 and Assumption 2.3. There is an explicit $c^W > 0$ such that, for all W continuously differentiable satisfying Assumption 2.4 with $L_W < c^W$, there is an explicit $\tau > 0$ such that for all probability measures ν_0^1 and ν_0^2 on $\mathbb{R}^{2d}$ with either a finite second moment (if Assumption 2.2 holds) or a finite Gaussian moment (if only Assumption 2.3 holds), there are explicit constants $C_1, C_2 > 0$ such that for all $t \geqslant 0$,*

$$\mathcal{W}_1\left(\nu_t^1, \nu_t^2\right) \leqslant e^{-\tau t} C_1\,, \qquad \mathcal{W}_2\left(\nu_t^1, \nu_t^2\right) \leqslant e^{-\tau t} C_2$$

where ν_t^1 and ν_t^2 are solutions of (2.1.1) with respective initial distributions ν_0^1 and ν_0^2.

In particular, we have existence and uniqueness of, as well as convergence towards, a stationary solution.

The second of our main results is a uniform in time convergence as $N \to +\infty$ of (2.1.3) toward (2.1.2).

Theorem 2.1.2. *Let $\tilde{\mathcal{C}}^0 > 0$ and $\tilde{a} > 0$. Let U be continuously differentiable and satisfy Assumptions 2.1 and 2.2. There is an explicit $c^W > 0$ such that, for all W continuously differentiable satisfying Assumption 2.4 with $L_W < c^W$, there exist explicit $B_1, B_2 > 0$, such that for all probability measures ν_0 on $\mathbb{R}^{2d}$ satisfying $\mathbb{E}_{\nu_0}\left(e^{\tilde{a}(|X|+|V|)}\right) \leqslant \tilde{\mathcal{C}}^0$,*

$$\mathcal{W}_1\left(\nu_t^{k,N}, \bar{\nu}_t^{\otimes k}\right) \leqslant \frac{kB_1}{\sqrt{N}}\,, \qquad \mathcal{W}_2^2\left(\nu_t^{k,N}, \bar{\nu}_t^{\otimes k}\right) \leqslant \frac{kB_2}{\sqrt{N}},$$

for all $k \in \mathbb{N}$, where $\nu_t^{k,N}$ is the marginal distribution at time t of the first k particles $\left((X_t^1, V_t^1), ..., (X_t^k, V_t^k)\right)$ of an N particle system (2.1.3) with initial distribution $(\nu_0)^{\otimes N}$, while $\bar{\nu}_t$ is a solution of (2.1.1) with initial distribution ν_0.

The organization of the chapter is as follows : in Section 2.2 we define the various tools involved in the construction of a good semimetrics. In Section 2.3 we study the long-time behavior of the Vlasov-Fokker-Planck equation (i.e Theorem 2.1.1) under the global Lipschitz Assumption 2.2 on U. Then, in Section 2.4, we prove propagation of chaos (i.e Theorem 2.1.2). Finally, in Section 2.5, we show how one may obtain the result of Theorem 2.1.1 under the local Lipschitz Assumption 2.3 on U.

We choose to present these proofs in this order, starting with the case in which the computations are the least cumbersome, in order to describe the method and motivate the construction of the semimetrics. Then, we add the tools to deal with the propagation of chaos. Finally, by combining the tools developed in Section 2.4 and the method of Section 2.3, we observe that it is possible to handle a small perturbation of the Lipschitz condition on U. In this way, we hope to gradually bring the difficulties and keep a form of clarity despite the sometimes involved calculations.

2.1.3 Comparison to existing works

Space homogeneous (i.e non kinetic) models of diffusive and interacting granular media, usually named McKean-Vlasov diffusions (see [14]), have attracted a lot of attention in the last twenty years. They have been treated by means of a stochastic interpretation and synchronous couplings as in [40] or in the recent [64] by reflection couplings 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 [93] for example). The granular media equations were interpreted as gradient flows in the space of probability measures in [38], leading to explicit exponential (or algebraic for non uniformly convex cases) rates of convergence to equilibrium of the non linear equation. Another approach relying on the dissipation of the Wasserstein distance and WJ inequalities was introduced in [20] handling small non convex cases. This approach was implemented in [153] to get propagation of chaos, under roughly the same type of assumptions. Mean-field limit using Γ-convergence tools has also been obtained in [37] for λ-convex potentials in this non kinetic setting.

Results on the long time behavior of the non-linear equation (2.1.2), i.e. space inhomogeneous, are few, as they combine the difficulty of getting explicit contraction rates for hypoelliptic diffusions as well as a non linear term. Recent works have tackled the question of contraction rate for the underdamped Langevin diffusion when there are no interaction (i.e $W = 0$). Results were obtained using hypocoercivity [62] and recently functional inequalities [1, 35], all in an L^2 setting that is not well adapted to the interacting particle system. For singular potential U, still without interaction, convergence rate in H^1 were obtained in [9]. Concerning the uniform in time propagation of chaos, there are no results except in the strictly convex case (with very small perturbation). We however refer to [170] for a result on the torus with W bounded with continuous derivative of all orders, see also [27]. Using functional inequalities (Poincaré or logarithmic Sobolev inequalities) for mean field models obtained in [87], other results were obtained provided the confining potential is a small perturbation of a quadratic function as in [142, 86, 89] which combines the hypocoercivity approach with independent of the number of particles constants appearing in the logarithmic Sobolev inequalities. The convergence of the Vlasov-Fokker-Planck equation to equilibrium for specific non-convex confining potentials and convex polynomial attractive interaction potentials using the free-energy approach has also been obtained in [63]. Our results generalize [89]. Indeed, we may consider non gradient interactions whereas it is crucial in their approach to know explicitly the invariant measure of the particles system, and also we may handle only locally Lipschitz confinement potential, whereas they impose at most quadratic growth of the potentials, and non strictly convex at infinity potential. It is however difficult to compare the smallness of the interaction potentials needed in both approaches. Note however that they obtain convergence to equilibrium in entropy whereas we get it in Wasserstein distance (controlled by entropy through a Talagrand inequality). Using a coupling strategy, and more precisely synchronous couplings, results under strict convexity assumption were obtained in [21] for contraction rates in Wasserstein distance, see also [108] but only for the nonlinear system.

As we mentioned, we adapt a proof from [68], which tackles (2.1.2) without interaction term. This chapter uses a Lyapunov condition that guarantees the recurrence of the process on a compact set. This idea is common when proving similar results through a probabilistic lens (see for instance [163] or [4]). Lyapunov conditions may also help to implement hypocoercivity techniques *à la Villani* to handle entropic convergence for non quadratic potentials, see [41]. Under the assumption U "greater than a quadratic function" at infinity and ∇W Lipschitz continuous, we too consider a Lyapunov function that allows us to construct a specific semimetric improving the convergence speed. But, and this is to our knowledge something new, when proving propagation of chaos we add a form of non linearity in the quantity we consider to tackle a part of the non linearity appearing in the dynamic (see Section 2.4 below). Let us also mention the very recent preprint by Schuh [156], posterior to our work, which also aims at proving long time behavior for the second-order Langevin dynamics and its non linear limit as well as uniform in time propagation of chaos, by constructing two separate metrics for small and large distances and showing contraction for both these quantities.

2.2 Modified semimetrics

As mentioned in the introduction, the proofs rely on the construction of suitable semimetrics on $\mathbb{R}^{2d}$ and $\mathbb{R}^{2dN}$. They are introduced in this section, together with some useful properties. In all this section, $\lambda, A, \tilde{A}, L_U$ and L_W are given by Assumptions 2.1, 2.2 and 2.4 and Lemma 2.1.1.

Before going into the details, let us highlight the main points of the construction of the semimetrics. It relies on the superposition of three ideas. The first idea is that, in order to deal with the kinetic process (2.1.2), the standard Euclidean norm $|x|^2 + |v|^2$ is not suitable and one should consider a linear change of variables, like $(x, v) \mapsto (x, x + \beta v)$ for some $\beta \in \mathbb{R}$. This is the case when using coupling methods as in [68, 21] but also when using hypocoercive modified entropies involving mixed derivatives as in [170, 163, 8, 41], the link being made in [141]. This motivates the definition of r below. The second idea is a modification of this distance r by some concave function f, which is related to the fact we are using, at least in some parts of the space, a reflection coupling. The concavity is well adapted to Itô's formula enabling the diffusion to provide a contraction effect (in a compact). This method has been considered for elliptic diffusions in [66], see also [69]. Intuitively, the contraction is produced by the fact that a random decrease in r has more effect on $f(r)$ than a random increase of the same amount. Finally, the third idea is the multiplication of a distance by a Lyapunov function G, which has first been used for Wasserstein distances in [90]. That way, on average, $f(r)G$ tends to decay because, when r is small, $f(r)$ tends to decay and, when r is large, G tends to decay.

2.2.1 A Lyapunov function

Let

$$\gamma = \frac{\lambda}{2(\lambda+1)}, \qquad B = 24\left(A + (\lambda - \gamma)\tilde{A} + d\right) \tag{2.2.1}$$

and, for $x, v \in \mathbb{R}^d$,

$$H(x,v) = 24U(x) + (6(1-\gamma) + \lambda)|x|^2 + 12x \cdot v + 12|v|^2 .$$

For μ a probability measure on $\mathbb{R}^d$ with finite first moment, ∇W being assumed Lipschitz continuous, denote by $\mathcal{L}_\mu$ the generator given by

$$\mathcal{L}_\mu \phi(x,v) = v \cdot \nabla_x \phi(x,v) - (v + \nabla U(x) + \nabla W * \mu(x)) \cdot \nabla_v \phi(x,v) + \Delta_v \phi(x,v) .$$

The main properties of H are the following.

Lemma 2.2.1. *Under Assumptions 2.1, 2.2 and 2.4, for all $x, v \in \mathbb{R}^d$ and μ,*

$$H(x,v) \geqslant 24U(x) + \lambda|x|^2 + 12\left|v + \frac{x}{2}\right|^2 , \tag{2.2.2}$$

$$\mathcal{L}_\mu H(x,v) \leqslant B + L_W(6+8\lambda)\left(\int |y| d\mu(y)\right)^2 - \left(\frac{3}{4}\lambda + \lambda^2\right)|x|^2 - \gamma H(x,v) , \tag{2.2.3}$$

$$\mathcal{L}_\mu H(x,v) \leqslant B + \left(\left(\int |y| d\mu(y)\right)^2 - |x|^2\right)\left(\frac{3}{4}\lambda + \lambda^2\right) - \gamma H(x,v) . \tag{2.2.4}$$

In particular H is non-negative and goes to $+\infty$ at infinity.

The proof follows from elementary computations and is detailed in Appendix B.1.3. Notice that the condition $L_W \leqslant \lambda/8$ is used here.

In the case of particular interest where $\mu = \mu_t$ is given by (2.1.2), taking the expectation in (2.2.4) and using Gronwall's lemma, we immediately get the following.

Lemma 2.2.2. *Under Assumptions 2.1, 2.2 and 2.4, let $(X_t, V_t)_{t\geqslant 0}$ be a solution of (2.1.2) with finite second moment at initial time. For all $t \geqslant 0$,*

$$\frac{d}{dt}\mathbb{E}H\left(X_t, V_t\right) \leqslant B - \gamma \mathbb{E}H\left(X_t, V_t\right), \tag{2.2.5}$$

$$\mathbb{E}H\left(X_t, V_t\right) - \frac{B}{\gamma} \leqslant \left(\mathbb{E}H\left(X_0, V_0\right) - \frac{B}{\gamma}\right) e^{-\gamma t}. \tag{2.2.6}$$

2.2.2 Change of variable and concave modification

We start by fixing the values of some parameters. The somewhat intricate expressions in this section are dictated by the computations arising in the proofs later on. Recall the definition of γ and B in (2.2.1). Set

$$\alpha = L_U + \frac{\lambda}{4}, \quad R_0 = \sqrt{\frac{24B}{5\gamma \min\left(3, \frac{\lambda}{3}\right)}}, \quad R_1 = \sqrt{\frac{24\left((1+\alpha)^2 + \alpha^2\right)}{5\gamma \min\left(3, \frac{\lambda}{3}\right)} B}.$$

For $x, \tilde{x}, v, \tilde{v} \in \mathbb{R}^d$, set

$$r(x, \tilde{x}, v, \tilde{v}) = \alpha|x - \tilde{x}| + |x - \tilde{x} + v - \tilde{v}|.$$

Lemma 2.2.3. *Under Assumptions 2.1, 2.2 and 2.4, for all $x, \tilde{x}, v, \tilde{v} \in \mathbb{R}^d$,*

$$r(x, \tilde{x}, v, \tilde{v})^2 \leqslant 2\frac{(1+\alpha)^2 + \alpha^2}{\min\left(\frac{1}{3}\lambda, 3\right)}\left(H\left(x, v\right) + H\left(\tilde{x}, \tilde{v}\right)\right), \tag{2.2.7}$$

so that, in particular,

$$r(x, \tilde{x}, v, \tilde{v}) \geqslant R_1 \qquad \Rightarrow \qquad \gamma H\left(x, v\right) + \gamma H\left(\tilde{x}, \tilde{v}\right) \geqslant \frac{12}{5}B.$$

We refer to Appendix B.1.4 for the proof. Let

$$c = \min\left\{\frac{\gamma}{36}, \frac{B}{3}, \frac{1}{7}\min\left(\frac{1}{2} - \frac{L_U + L_W}{2\alpha}, 2\sqrt{\frac{L_U + L_W}{2\pi\alpha}}\right)\right.$$
$$\left. \times \exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + 96\max\left(\frac{1}{2\alpha}, 1\right)\right) R_1^2\right)\right\}. \tag{2.2.8}$$

Set

$$\epsilon = \frac{3c}{B}, \qquad \mathbf{C} = c + 2\epsilon B$$

and, for $s \geqslant 0$,

$$\phi\left(s\right) = \exp\left(-\frac{1}{8}\left(\frac{1}{\alpha}\left(L_U + L_W\right) + \alpha + 96\epsilon\max\left(\frac{1}{2\alpha}, 1\right)\right) s^2\right), \qquad \Phi\left(s\right) = \int_0^s \phi\left(u\right) du$$

$$g(s) = 1 - \frac{\mathbf{C}}{4}\int_0^s \frac{\Phi(u)}{\phi(u)}du\,, \qquad f(s) = \int_0^{\min(s,R_1)} \phi(u)\, g(u)\, du.$$

Remark 2.2.1. *The parameters above are far from being optimal. They are somewhat roughly chosen as we only wish to convey the fact that every constant is explicit.*

The next lemma, proved in Appendix B.2, gathers the intermediary bounds that will be useful in the proofs of the main results.

Lemma 2.2.4. *Under Assumptions 2.1, 2.2 and 2.4,*

$$c \leqslant \frac{\gamma}{6}\left(1 - \frac{\frac{5\gamma}{6}}{2\epsilon B + \frac{5\gamma}{6}}\right), \tag{2.2.9}$$

$$L_U + L_W < \alpha, \tag{2.2.10}$$

$$c + 2\epsilon B \leqslant \frac{1}{2}\left(1 - \frac{L_U + L_W}{\alpha}\right) \inf_{r\in]0,R_1]} \frac{r\phi(r)}{\Phi(r)}, \tag{2.2.11}$$

$$c + 2\epsilon B \leqslant 2\left(\int_0^{R_1} \Phi(s)\,\phi(s)^{-1}\, ds\right)^{-1}, \tag{2.2.12}$$

$$\forall s \geqslant 0, \qquad 0 = 4\phi'(s) + \left(\frac{1}{\alpha}(L_U + L_W) + \alpha + 96\epsilon \max\left(\frac{1}{2\alpha}, 1\right)\right) s\phi(s). \tag{2.2.13}$$

The main properties of f are the following.

Lemma 2.2.5. *The function f is twice continuously differentiable on $(0, R_1)$ with $f'_+(0) = 1$ and $f'_-(R_1) > 0$, and constant on $[R_1, \infty)$. Moreover, it is non-negative, non-decreasing and concave, and for all $s \geqslant 0$,*

$$\min(s, R_1)\, f'_-(R_1) \leqslant f(s) \leqslant \min(s, f(R_1)) \leqslant \min(s, R_1).$$

Proof. First, notice that (2.2.12) ensures that $g(s) \geqslant \frac{1}{2}$ for all $s \geqslant 0$. Then, all the points immediately follow from the fact the functions ϕ and g are twice continuously differentiable, positive and decreasing, with $\phi(0) = g(0) = 1$. □

2.2.3 The modified semimetrics

For $x, \tilde{x}, v, \tilde{v} \in \mathbb{R}^d$, set

$$G(x, v, \tilde{x}, \tilde{v}) = 1 + \epsilon H(x, v) + \epsilon H(\tilde{x}, \tilde{v}),$$
$$\rho(x, v, \tilde{x}, \tilde{v}) = f(r(x, v, \tilde{x}, \tilde{v}))\, G(x, v, \tilde{x}, \tilde{v}).$$

An immediate corollary of Lemmas 2.2.3 and 2.2.5 is that ρ is a semimetric on $\mathbb{R}^{2d}$ which controls the usual L1 and L2 distances:

Lemma 2.2.6. *There are explicit constants $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_r, \mathcal{C}_z > 0$ such that for all $x, x', v, v' \in \mathbb{R}^d$,*

$$|x - x'| + |v - v'| \leqslant \mathcal{C}_1 \rho((x, v), (x', v'))$$
$$|x - x'|^2 + |v - v'|^2 \leqslant \mathcal{C}_2 \rho((x, v), (x', v'))$$
$$r(x, v, x', v') \leqslant \mathcal{C}_r \rho((x, v), (x', v'))$$
$$|x - x'| \leqslant \mathcal{C}_z f(r(x, v, x', v'))\left(1 + \epsilon\sqrt{H(x, v)} + \epsilon\sqrt{H(x', v')}\right).$$

We also mention a technical lemma, see Appendix B.1.6 for proof.

Lemma 2.2.7. *For all* $x, v, \tilde{x}, \tilde{v} \in \mathbb{R}^d$

$$|H(x,v) - H(\tilde{x},\tilde{v})| \leqslant C_{dH,1} r(x,\tilde{x},v,\tilde{v}) + C_{dH,2} r(x,\tilde{x},v,\tilde{v}) \left(\sqrt{H(x,v)} + \sqrt{H(\tilde{x},\tilde{v})}\right), \quad (2.2.14)$$

where

$$C_{dH,1} := \frac{24|\nabla U(0)|}{\alpha} \quad \text{and} \quad C_{dH,2} := \frac{24 L_U}{\alpha\sqrt{\lambda}} + \frac{6(1-\gamma)+\lambda-3}{\alpha\sqrt{\lambda}} + 2\sqrt{3}\max\left(1, \frac{1}{2\alpha}\right).$$

Finally, for μ and ν two probability measures on $\mathbb{R}^{2d}$ and a measurable function $h : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{R}$, we define

$$\mathcal{W}_h(\mu,\nu) = \inf_{\Gamma \in \Pi(\mu,\nu)} \int h(x,v,\tilde{x},\tilde{v})\, \Gamma\left(d(x,v)\, d(\tilde{x},\tilde{v})\right).$$

2.3 Proof of Theorem 2.1.1

In this section, for the sake of clarity, we only assume the potential U satisfies Assumption 2.1 and Assumption 2.2. We refer to Section 2.5 for the adjustment of the proof in the case ∇U locally Lipschitz continuous.

Our goal is to prove the following result

Theorem 2.3.1. *Let* $\mathcal{C}^0 > 0$. *Let* U *be continuously differentiable and satisfy Assumption 2.1 and Assumption 2.2. Let*

$$\tilde{\mathcal{C}}_K := \mathcal{C}_1 \left(1 + \frac{2\epsilon B}{\gamma} + 2\epsilon\mathcal{C}^0\right) + 2\epsilon\left(\frac{B}{\gamma} + \mathcal{C}^0\right)\frac{6+8\lambda}{\lambda}.$$

For all W *twice continuously differentiable satisfying Assumption 2.4 with* $L_W < c/\tilde{\mathcal{C}}_K$, *for all probability measures* ν_0^1 *and* ν_0^2 *on* $\mathbb{R}^{2d}$ *satisfying* $\mathbb{E}_{\nu_0^1} H \leqslant \mathcal{C}^0$ *and* $\mathbb{E}_{\nu_0^2} H \leqslant \mathcal{C}^0$

$$\forall t \geqslant 0, \qquad \mathcal{W}_\rho\left(\nu_t^1, \nu_t^2\right) \leqslant e^{-(c - L_W \tilde{\mathcal{C}}_K)t} \mathcal{W}_\rho\left(\nu_0^1, \nu_0^2\right),$$

where ν_t^1 *(resp.* ν_t^2*) is a solution of* (2.1.1) *with initial distribution* ν_0^1 *(resp.* ν_0^2*).*

2.3.1 Step one: Coupling and evolution of the coupling semimetric

Let $\xi > 0$, and let $rc,\ sc : \mathbb{R}^{2d} \mapsto [0,1]$ be two Lipschitz continuous functions such that :

$$\begin{aligned}
&rc^2 + sc^2 = 1, \\
&rc(z,w) = 0 \text{ if } |z+w| \leqslant \frac{\xi}{2} \text{ or } \alpha|z| + |z+w| \geqslant R_1 + \xi, \\
&rc(z,w) = 1 \text{ if } |z+w| \geqslant \xi \text{ and } \alpha|z| + |z+w| \leqslant R_1.
\end{aligned}$$

These two functions translate into mathematical terms the regions in which we use a reflection coupling (represented by $rc = 1$) and the ones where we use a synchronous coupling (represented by $sc = 1$). Finally, ξ is a parameter that will vanish to zero in the end. We therefore consider the following coupling :

$$\begin{cases} dX_t = V_t dt \\ dV_t = -V_t dt - \nabla U\left(X_t\right) dt - \nabla W * \mu_t\left(X_t\right) dt + \sqrt{2} rc\left(Z_t, W_t\right) dB_t^{rc} \\ \qquad +\sqrt{2} sc\left(Z_t, W_t\right) dB_t^{sc} \\ \mu_t = \text{Law}\left(X_t\right) \\ d\tilde{X}_t = \tilde{V}_t dt \\ d\tilde{V}_t = -\tilde{V}_t dt - \nabla U(\tilde{X}_t) dt - \nabla W * \tilde{\mu}_t(\tilde{X}_t) dt + \sqrt{2} rc\left(Z_t, W_t\right)\left(Id - 2 e_t e_t^T\right) dB_t^{rc} \\ \qquad +\sqrt{2} sc\left(Z_t, W_t\right) dB_t^{sc} \\ \tilde{\mu}_t = \text{Law}(\tilde{X}_t), \end{cases}$$

where B^{rc} and B^{sc} are independent Brownian motions, and

$$Z_t = X_t - \tilde{X}_t, \quad W_t = V_t - \tilde{V}_t, \quad Q_t = Z_t + W_t, \quad e_t = \begin{cases} \frac{Q_t}{|Q_t|} & \text{if } Q_t \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and e_t^T is the transpose of e_t. Then

$$\frac{dZ_t}{dt} = W_t = Q_t - Z_t. \tag{2.3.1}$$

So $\frac{d|Z_t|}{dt} = \frac{Z_t}{|Z_t|}\left(Q_t - Z_t\right)$ for every t such that $Z_t \neq 0$, and $\frac{d|Z_t|}{dt} \leqslant |Q_t|$ for every t such that $Z_t = 0$. In particular

$$\frac{d|Z_t|}{dt} \leqslant |Q_t| - |Z_t|.$$

We start by using Itô's formula to compute the evolution of $|Q_t|$. The following lemma is Lemma 7 of A. Durmus *et al.* [64] of which, for the sake of completeness, we give the proof.

Lemma 2.3.1. *Under Assumption 2.1, Assumption 2.2 and Assumption 2.4, we have almost surely for all $t \geqslant 0$.*

$$\begin{aligned} d|Q_t| = &- e_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - e_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\ &+ 2\sqrt{2} rc\left(Z_t, W_t\right) e_t \cdot dB_t^{rc} \end{aligned} \tag{2.3.2}$$

Proof. Let $t \geqslant 0$. We begin by considering the dynamics of Z_t, W_t and Q_t. We have

$$\begin{aligned} dZ_t &= W_t dt \\ dW_t &= -W_t dt - \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\ &\quad + 2\sqrt{2} rc\left(Z_t, W_t\right) e_t e_t \cdot dB_t^{rc} \\ dQ_t &= -\left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\ &\qquad + 2\sqrt{2} rc\left(Z_t, W_t\right) e_t e_t \cdot dB_t^{rc}. \end{aligned}$$

Therefore

$$\begin{aligned} d|Q_t|^2 = &-2 Q_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - 2 Q_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\ &+ 4\sqrt{2} rc\left(Z_t, W_t\right)\left(Q_t \cdot e_t\right) e_t \cdot dB_t^{rc} + 8 rc^2\left(Z_t, W_t\right) dt. \end{aligned}$$

We consider, for $\eta > 0$, the function $\psi_\eta(r) = (r+\eta)^{1/2}$ which is $\mathcal{C}^\infty$ on $]0,\infty[$ and satisfies

$$\forall r \geqslant 0, \quad \lim_{\eta\to 0} \psi_\eta(r) = r^{1/2}, \quad \lim_{\eta\to 0} 2\psi'_\eta(r) = r^{-1/2}, \quad \lim_{\eta\to 0} 4\psi''_\eta(r) = -r^{-3/2},$$
$$\text{and thus } \lim_{\eta\to 0} 2r\psi''_\eta(r) + \psi'_\eta(r) = 0.$$

Then

$$\begin{aligned} d\psi_\eta\left(|Q_t|^2\right) = &- 2\psi'_\eta\left(|Q_t|^2\right) Q_t \cdot \left(\nabla U(X_t) - \nabla U(\tilde{X}_t)\right) dt \\ &- 2\psi'_\eta\left(|Q_t|^2\right) Q_t \cdot \left(\nabla W * \mu_t(X_t) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\ &+ 4\psi'_\eta\left(|Q_t|^2\right)\sqrt{2} rc(Z_t, W_t)(Q_t \cdot e_t)\, e_t \cdot dB^{rc}_t + 8\psi'_\eta\left(|Q_t|^2\right) rc^2(Z_t, W_t)\, dt \\ &+ 16\psi''_\eta\left(|Q_t|^2\right) rc^2(Z_t, W_t)\, |Q_t|^2 dt. \end{aligned}$$

We make sure each individual term converges almost surely as $\eta \to 0$. First, we notice that

$$2|Q_t|\psi'_\eta\left(|Q_t|^2\right) = \frac{|Q_t|}{\left(|Q_t|^2+\eta\right)^{1/2}} \leqslant 1.$$

So

$$2\psi'_\eta\left(|Q_t|^2\right) Q_t \cdot \left(\nabla U(X_t) - \nabla U(\tilde{X}_t)\right) \leqslant |\nabla U(X_t) - \nabla U(\tilde{X}_t)| \leqslant L_U |Z_t|.$$

Then, by dominated convergence, for all $T \geqslant 0$ almost surely

$$\begin{aligned} \lim_{\eta\to 0}\int_0^T 2\psi'_\eta\left(|Q_t|^2\right) Q_t \cdot \left(\nabla U(X_t) - \nabla U(\tilde{X}_t)\right) dt &= \int_0^T \frac{Q_t}{|Q_t|} \cdot \left(\nabla U(X_t) - \nabla U(\tilde{X}_t)\right) dt \\ &= \int_0^T e_t \cdot \left(\nabla U(X_t) - \nabla U(\tilde{X}_t)\right) dt. \end{aligned}$$

Likewise for all $T \geqslant 0$

$$\begin{aligned} 2\psi'_\eta\left(|Q_t|^2\right) Q_t \cdot \left(\nabla W * \mu_t(X_t) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) &\leqslant \left|\nabla W * \mu_t(X_t) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right| \\ &\leqslant L_W |Z_t| + L_W \mathbb{E}|Z_t|, \end{aligned}$$

hence

$$\begin{aligned} &\lim_{\eta\to 0}\int_0^T 2\psi'_\eta\left(|Q_t|^2\right) Q_t \cdot \left(\nabla W * \mu_t(X_t) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\ &\quad = \int_0^T e_t \cdot \left(\nabla W * \mu_t(X_t) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt. \end{aligned}$$

Then, since $rc(Z_t, W_t) = 0$ for $|Q_t| \leqslant \frac{\xi}{2}$ and

$$8\psi'_\eta\left(|Q_t|^2\right) + 16\psi''_\eta\left(|Q_t|^2\right)|Q_t|^2 = 4\left(\frac{1}{\left(|Q_t|^2+\eta\right)^{1/2}} - \frac{|Q_t|^2}{\left(|Q_t|^2+\eta\right)^{3/2}}\right)$$

$$= 4\frac{\eta}{\left(|Q_t|^2+\eta\right)^{3/2}} \leqslant \frac{4\eta}{|Q_t|^3},$$

we have by dominated convergence

$$\lim_{\eta\to 0}\int_0^T \left(8d\psi_\eta'\left(|Q_t|^2\right) rc^2\left(Z_t,W_t\right) + 16\psi_\eta''\left(|Q_t|^2\right) rc^2\left(Z_t,W_t\right)|Q_t|^2\right) dt = 0.$$

Finally, by Theorem 2.12 chapter 4 of [150]

$$\lim_{\eta\to 0}\int_0^T 4\sqrt{2}\psi_\eta'\left(|Q_t|^2\right) rc\left(Z_t,W_t\right)\left(Q_t\cdot e_t\right) e_t\cdot dB_t^{rc} = \int_0^T 2\sqrt{2} rc\left(Z_t,W_t\right) e_t\cdot dB^{rc}.$$

For any t, we obtain the desired result almost surely. The continuity of $t\mapsto |Q_t|$ then allows us to conclude that (2.3.2) is almost surely true for all t. □

We denote

$$r_t := \alpha|X_t-\tilde{X}_t| + |X_t-\tilde{X}_t+V_t-\tilde{V}_t| = \alpha|Z_t|+|Q_t|, \tag{2.3.3}$$

$$\rho_t := f\left(r_t\right) G_t \text{ where } G_t = 1+\epsilon H\left(X_t,V_t\right)+\epsilon H(\tilde{X}_t,\tilde{V}_t). \tag{2.3.4}$$

Since $H\left(x,v\right)\geqslant 0$ we have $G_t\geqslant 1$. We now state the main lemma of this section.

Lemma 2.3.2. *Under Assumption 2.1, Assumption 2.2 and Assumption 2.4, let $c\in]0,\infty[$. Then almost surely for all $t\geqslant 0$*

$$\forall t\geqslant 0,\ e^{ct}\rho_t \leqslant \rho_0 + \int_0^t e^{cs}K_s ds + M_t, \tag{2.3.5}$$

where $\left(M_t\right)_t$ is a continuous local martingale and

$$\begin{aligned} K_t =\ & 4f''\left(r_t\right) rc\left(Z_t,W_t\right)^2 G_t + cf\left(r_t\right) G_t + 96\epsilon\max\left(1,\frac{1}{2\alpha}\right) r_t f'\left(r_t\right) rc\left(Z_t,W_t\right)^2 \\ & + \left(\alpha\frac{d|Z_t|}{dt} + \left(L_U+L_W\right)|Z_t| + L_W\mathbb{E}|Z_t|\right) f'\left(r_t\right) G_t \\ & + \epsilon\left(2B-\gamma H\left(X_t,V_t\right)-\gamma H(\tilde{X}_t,\tilde{V}_t)\right) f\left(r_t\right) \\ & + \epsilon L_W\left(6+8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2+\mathbb{E}(|\tilde{X}_t|)^2\right) f\left(r_t\right). \end{aligned}$$

Proof. Using (2.3.2)

$$\begin{aligned} |Q_t| &= |Q_0| + A_t^Q + M_t^Q \quad \text{with} \\ dA_t^Q &= -e_t\cdot\left(\nabla U\left(X_t\right)-\nabla U(\tilde{X}_t)\right) dt - e_t\cdot\left(\nabla W*\mu_t\left(X_t\right)-\nabla W*\tilde{\mu}_t(\tilde{X}_t)\right) dt \\ dM_t^Q &= 2\sqrt{2} rc\left(Z_t,W_t\right) e_t\cdot dB_t^{rc}. \end{aligned}$$

Therefore $r_t = |Q_0|+\alpha|Z_t|+A_t^Q+M_t^Q$. Let $c>0$. By Itô's formula

$$d\left(e^{ct}f\left(r_t\right)\right) = ce^{ct}f\left(r_t\right)dt + e^{ct}f'\left(r_t\right)dr_t + \frac{1}{2}e^{ct}f''\left(r_t\right)8rc^2\left(Z_t,W_t\right)dt.$$

Hence

$$\begin{aligned}
e^{ct}f(r_t) =& f(r_0)+\hat{A}_t+\hat{M}_t \text{ with}\\
d\hat{A}_t =& \Big(cf(r_t)+\alpha f'(r_t)\frac{d|Z_t|}{dt}-f'(r_t)\,e_t\cdot\Big(\nabla U(X_t)-\nabla U(\tilde{X}_t)\Big)\\
&-f'(r_t)\,e_t\cdot\Big(\nabla W*\mu_t(X_t)-\nabla W*\tilde{\mu}_t(\tilde{X}_t)\Big)+4f''(r_t)\,rc^2(Z_t,W_t)\Big)e^{ct}dt\\
d\hat{M}_t =& e^{ct}2\sqrt{2}f'(r_t)\,rc(Z_t,W_t)\,e_t\cdot dB_t^{rc}.
\end{aligned}$$

We now consider the evolution of $G_t=1+\epsilon H(X_t,V_t)+\epsilon H(\tilde{X}_t,\tilde{V}_t)$

$$\begin{aligned}
dG_t =& \epsilon\left(\mathcal{L}_{\mu_t}H(X_t,V_t)+\mathcal{L}_{\tilde{\mu}_t}H(\tilde{X}_t,\tilde{V}_t)\right)dt\\
&+\epsilon\sqrt{2}rc(Z_t,W_t)\left(\nabla_v H(X_t,V_t)-\nabla_v H(\tilde{X}_t,\tilde{V}_t)\right)\cdot e_te_t^T dB_t^{rc}\\
&+\epsilon\sqrt{2}rc(Z_t,W_t)\left(\nabla_v H(X_t,V_t)+\nabla_v H(\tilde{X}_t,\tilde{V}_t)\right)\cdot\left(Id-e_te_t^T\right)dB_t^{rc}\\
&+\epsilon\sqrt{2}sc(Z_t,W_t)\left(\nabla_v H(X_t,V_t)+\nabla_v H(\tilde{X}_t,\tilde{V}_t)\right)\cdot dB_t^{sc}.
\end{aligned}$$

Therefore $e^{ct}\rho_t=e^{ct}f(r_t)\,G_t=\rho_0+A_t+M_t$, where

$$\begin{aligned}
dA_t =& G_t d\hat{A}_t+\epsilon e^{ct}f(r_t)\left(\mathcal{L}_{\mu_t}H(X_t,V_t)+\mathcal{L}_{\tilde{\mu}_t}H(\tilde{X}_t,\tilde{V}_t)\right)dt\\
&+4\epsilon e^{ct}f'(r_t)\,rc^2(Z_t,W_t)\left(\nabla_v H(X_t,V_t)-\nabla_v H(\tilde{X}_t,\tilde{V}_t)\right)\cdot e_t dt,
\end{aligned}$$

and M_t is a continuous local martingale. This last equality uses the fact that B^{rc} and B^{sc} are independent Brownian motion and that $e_t\cdot\left(Id-e_te_t^T\right)=0$. Furthermore

$$\begin{aligned}
|\nabla_v H(X_t,V_t)-\nabla_v H(\tilde{X}_t,\tilde{V}_t)| &= 12|X_t+2V_t-\tilde{X}_t-2\tilde{V}_t| = 12|2Q_t-Z_t|\\
&\leqslant 24\left(\frac{1}{2}|Z_t|+|Q_t|\right)\\
&\leqslant 24\max\left(1,\frac{1}{2\alpha}\right)r_t,
\end{aligned}$$

so that $dA_t\leqslant e^{ct}\tilde{K}_t dt$, where

$$\begin{aligned}
\tilde{K}_t =& \Big(cf(r_t)+\alpha f'(r_t)\frac{d|Z_t|}{dt}-f'(r_t)\,e_t\cdot\Big(\nabla U(X_t)-\nabla U(\tilde{X}_t)\Big)\\
&-f'(r_t)\,e_t\cdot\Big(\nabla W*\mu_t(X_t)-\nabla W*\tilde{\mu}_t(\tilde{X}_t)\Big)+4f''(r_t)\,rc^2(Z_t,W_t)\Big)G_t\\
&+\epsilon\left(\mathcal{L}_t H(X_t,V_t)+\mathcal{L}_t H(\tilde{X}_t,\tilde{V}_t)\right)f(r_t)+96\epsilon\max\left(1,\frac{1}{2\alpha}\right)r_t f'(r_t)\,rc^2(Z_t,W_t).
\end{aligned}$$

And we conclude using Lemma 2.1.2, and Lemma 2.2.1. □

2.3.2 Step two : Contractivity in various regions of space

At this point, we have

$$\forall t\geqslant 0,\ e^{ct}\rho_t\leqslant\rho_0+\int_0^t e^{cs}K_s ds+M_t,$$

where M_t is a continuous local martingale and, by regrouping the terms according to how we will use them

$$K_t = \left(cf(r_t) + \left(\alpha\frac{d|Z_t|}{dt} + (L_U + L_W)|Z_t|\right) f'(r_t)\right) G_t \tag{2.3.6}$$

$$+ 4\left(f''(r_t) G_t + 24\epsilon \max\left(1, \frac{1}{2\alpha}\right) r_t f'(r_t)\right) rc(Z_t, W_t)^2 \tag{2.3.7}$$

$$+ \epsilon\left(2B - \gamma H(X_t, V_t) - \gamma H(\tilde{X}_t, \tilde{V}_t)\right) f(r_t) \tag{2.3.8}$$

$$+ L_W f'(r_t)\mathbb{E}(|Z_t|) G_t + \epsilon L_W (6 + 8\lambda)\left(\mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right) f(r_t). \tag{2.3.9}$$

Briefly,

- lines (2.3.6) and (2.3.7) will be non positive thanks to the construction of the function f when using the reflection coupling,
- when only using the synchronous coupling, i.e when the deterministic drift is contracting, line (2.3.6) alone will be sufficiently small,
- line (2.3.8) translates the effect the Lyapunov function has in bringing back processes that would have ventured at infinity,
- finally, line (2.3.9) contains the non linearity and will be tackled by taking L_W sufficiently small.

In this section, we thus prove the following lemma

Lemma 2.3.3. *Assume the confining potential U satisfies Assumption 2.1 and Assumption 2.2. Then there is a constant $c^W > 0$ such that for all interaction potential W satisfying Assumption 2.4 with $L_W < c^W$, the following holds for K_t defined in* (2.3.6)-(2.3.9)

$$\mathbb{E}K_t \leqslant (1+\alpha)\,\xi\mathbb{E}G_t + L_W\left(\mathcal{C}_K + \mathcal{C}_K^0 e^{-\gamma t}\right)\mathbb{E}\rho_t,$$

with

$$\mathcal{C}_K = \mathcal{C}_1\left(1 + \frac{2\epsilon B}{\gamma}\right) + \frac{2\epsilon B}{\gamma\lambda}(6 + 8\lambda), \tag{2.3.10}$$

$$\mathcal{C}_K^0 = \epsilon\left(\mathcal{C}_1 + \frac{6+8\lambda}{\lambda}\right)\left(\mathbb{E}H(X_0, V_0) + \mathbb{E}H(\tilde{X}_0, \tilde{V}_0)\right).$$

The constant c^W is explicit, as it will be shown in Appendix B.2.

To this end, we divide the space into three regions

$$\text{Reg}_1 = \left\{(X_t, V_t, \tilde{X}_t, \tilde{V}_t) \text{ s.t. } |Q_t| \geqslant \xi \text{ and } r_t \leqslant R_1\right\},$$
$$\text{Reg}_2 = \left\{(X_t, V_t, \tilde{X}_t, \tilde{V}_t) \text{ s.t. } |Q_t| < \xi \text{ and } r_t \leqslant R_1\right\},$$
$$\text{Reg}_3 = \left\{(X_t, V_t, \tilde{X}_t, \tilde{V}_t) \text{ s.t. } r_t > R_1\right\},$$

and consider

$$\mathbb{E}K_t = \mathbb{E}(K_t \mathbb{1}_{\text{Reg}_1}) + \mathbb{E}(K_t \mathbb{1}_{\text{Reg}_2}) + \mathbb{E}(K_t \mathbb{1}_{\text{Reg}_3}).$$

First region : $|Q_t| \geqslant \xi$ **and** $r_t \leqslant R_1$

In this region of space, we use the Brownian motion through the reflection coupling and the construction of the function f to bring the processes closer together. Here we have $rc\left(Z_t, W_t\right) = 1$. Recall $\alpha|Z_t| + |Q_t| = r_t$ and $G_t \geqslant 1$.

- We have

$$\begin{aligned}\alpha\frac{d|Z_t|}{dt} + \left(L_U + L_W\right)|Z_t| &\leqslant \alpha|Q_t| - \alpha|Z_t| + \left(L_U + L_W\right)|Z_t| \\ &= \alpha r_t - \alpha^2|Z_t| - \alpha|Z_t| + \left(L_U + L_W\right)|Z_t| \\ &\leqslant \left(\frac{1}{\alpha}\left(L_U + L_W\right) + \alpha\right) r_t.\end{aligned}$$

- Since $G_t = 1 + \epsilon H\left(X_t, V_t\right) + \epsilon H(\tilde{X}_t, \tilde{V}_t) \geqslant 1$,

$$cG_t + \epsilon\left(2B - \gamma H\left(X_t, V_t\right) - \gamma H(\tilde{X}_t, \tilde{V}_t)\right) \leqslant cG_t + 2\epsilon BG_t = \mathbf{C}G_t. \tag{2.3.11}$$

- We then have, by (2.2.13),

$$4\phi'\left(r_t\right) + \left(\frac{1}{\alpha}\left(L_U + L_W\right) + \alpha + 96\epsilon\max\left(\frac{1}{2\alpha}, 1\right)\right) r_t\phi\left(r_t\right) = 0.$$

  Hence

$$\begin{aligned}&4f''\left(r_t\right) + \left(\frac{1}{\alpha}\left(L_U + L_W\right) + \alpha + 96\epsilon\max\left(\frac{1}{2\alpha}, 1\right)\right) r_t f'\left(r_t\right) \\ &=4\phi'\left(r_t\right) g\left(r_t\right) + 4\phi\left(r_t\right) g'\left(r_t\right) \\ &\quad+ \left(\frac{1}{\alpha}\left(L_U + L_W\right) + \alpha + 96\epsilon\max\left(\frac{1}{2\alpha}, 1\right)\right) r_t\phi\left(r_t\right) g\left(r_t\right) \\ &=4\phi\left(r_t\right) g'\left(r_t\right),\end{aligned}$$

  and

$$4\phi\left(r_t\right) g'\left(r_t\right) + \mathbf{C}f\left(r_t\right) \leqslant -4\frac{\mathbf{C}}{4}\Phi\left(r_t\right) + \mathbf{C}\Phi\left(r_t\right) = 0.$$

- At this point, through this choice of function f, we are left with

$$K_t\mathbb{1}_{\text{Reg}_1} \leqslant L_W f'\left(r_t\right)\mathbb{E}\left(|Z_t|\right) G_t + \epsilon L_W\left(6 + 8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right) f\left(r_t\right).$$

  Using Lemma 2.2.6, $f'\left(r_t\right) \leqslant 1$ and $G_t \geqslant 1$,

$$\mathbb{E}\left(K_t\mathbb{1}_{\text{Reg}_1}\right) \leqslant L_W\mathcal{C}_1\mathbb{E}\left(\rho_t\right)\mathbb{E}\left(G_t\right) + \epsilon L_W\left(6 + 8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right)\mathbb{E}\left(\rho_t\right).$$

  Recall (2.2.6)

$$\begin{aligned}\mathbb{E}\left(G_t\right) =&1 + \epsilon\mathbb{E}H\left(X_t, V_t\right) + \epsilon\mathbb{E}H(\tilde{X}_t, \tilde{V}_t), \\ \leqslant&1 + \frac{2\epsilon B}{\gamma} + \epsilon\left(\mathbb{E}H\left(X_0, V_0\right) + \mathbb{E}H(\tilde{X}_0, \tilde{V}_0)\right) e^{-\gamma t},\end{aligned}$$

and, since $H(x,v) \geqslant \lambda|x|^2$,

$$\begin{aligned}\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \leqslant & \frac{1}{\lambda}\mathbb{E}H\left(X_t,V_t\right) + \frac{1}{\lambda}\mathbb{E}H(\tilde{X}_t,\tilde{V}_t),\\ \leqslant & \frac{2B}{\gamma\lambda} + \frac{1}{\lambda}\left(\mathbb{E}H\left(X_0,V_0\right) + \mathbb{E}H(\tilde{X}_0,\tilde{V}_0)\right)e^{-\gamma t}.\end{aligned}$$

Hence

$$\begin{aligned}\mathbb{E}\left(K_t\mathbb{1}_{\mathrm{Reg}_1}\right) \leqslant & L_W\left(\mathcal{C}_1\left(1+\frac{2\epsilon B}{\gamma}\right) + \frac{2\epsilon B}{\gamma\lambda}\left(6+8\lambda\right)\right)\mathbb{E}\left(\rho_t\right)\\ & + L_W\epsilon\left(\mathcal{C}_1 + \frac{6+8\lambda}{\lambda}\right)\left(\mathbb{E}H\left(X_0,V_0\right) + \mathbb{E}H(\tilde{X}_0,\tilde{V}_0)\right)\mathbb{E}\left(\rho_t\right)e^{-\gamma t}.\end{aligned}$$

We thus obtain $\mathbb{E}\left(K_t\mathbb{1}_{\mathrm{Reg}_1}\right) \leqslant L_W\left(\mathcal{C}_K + \mathcal{C}_K^0 e^{-\gamma t}\right)\mathbb{E}\rho_t$.

Second region : $|Q_t| < \xi$ **and** $r_t \leqslant R_1$

In this region of space, we use the naturally contracting deterministic drift thanks to a synchronous coupling. Here $R_1 \geqslant r_t \geqslant \alpha|Z_t| \geqslant r_t - \xi$ so that

$$\begin{aligned}K_t \leqslant & \mathbf{C}f\left(r_t\right)G_t + \left(\alpha\xi - r_t + \xi + \frac{1}{\alpha}\left(L_U+L_W\right)r_t\right)f'\left(r_t\right)G_t\\ & + \left(4f''\left(r_t\right)G_t + 96\epsilon\max\left(\frac{1}{2\alpha},1\right)r_tf'\left(r_t\right)\right)rc\left(Z_t,W_t\right)^2\\ & + L_Wf'\left(r_t\right)\mathbb{E}\left(|Z_t|\right)G_t + \epsilon L_W\left(6+8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right)f\left(r_t\right),\end{aligned}$$

where we use (2.3.11). First

$$4f''\left(r_t\right)G_t + 96\epsilon\max\left(\frac{1}{2\alpha},1\right)r_tf'\left(r_t\right) \leqslant 0.$$

We use (2.2.10) to obtain, since $f\left(r_t\right) \leqslant \Phi\left(r_t\right)$ and $\frac{1}{2}\phi\left(r_t\right) \leqslant f'\left(r_t\right) = \phi\left(r_t\right)g\left(r_t\right) \leqslant \phi\left(r_t\right)$ by (2.2.12),

$$\begin{aligned}K_t \leqslant & \xi\left(1+\alpha\right)\phi\left(r_t\right)g\left(r_t\right)G_t + G_t\left(\mathbf{C}\Phi\left(r_t\right) + \frac{1}{2}\left(\frac{1}{\alpha}\left(L_U+L_W\right)-1\right)r_t\phi\left(r_t\right)\right)\\ & + L_Wf'\left(r_t\right)\mathbb{E}\left(|Z_t|\right)G_t + \epsilon L_W\left(6+8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right)f\left(r_t\right).\end{aligned}$$

Then, thanks to (2.2.11), like in the first region of space

$$\begin{aligned}K_t\mathbb{1}_{\mathrm{Reg}_2} \leqslant & \xi\left(1+\alpha\right)\phi\left(r_t\right)g\left(r_t\right)G_t + L_Wf'\left(r_t\right)\mathbb{E}\left(|Z_t|\right)G_t\\ & + \epsilon L_W\left(6+8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right)f\left(r_t\right).\end{aligned}$$

Hence, since $\phi\left(r_t\right)g\left(r_t\right) \leqslant 1$

$$\mathbb{E}K_t\mathbb{1}_{\mathrm{Reg}_2} \leqslant \xi\left(1+\alpha\right)\mathbb{E}\left(G_t\right) + L_W\left(\mathcal{C}_K + \mathcal{C}_K^0e^{-\gamma t}\right)\mathbb{E}\rho_t.$$

Third region : $r_t > R_1$

In this region, we use the Lyapunov function. Here $f'(r_t) = f''(r_t) = 0$ so that

$$\begin{aligned} K_t \mathbb{1}_{\mathrm{Reg}_3} =& \left(cG_t + \epsilon \left(2B - \gamma H(X_t, V_t) - \gamma H(\tilde{X}_t, \tilde{V}_t) \right) \right) f(r_t) \mathbb{1}_{\mathrm{Reg}_3} \\ &+ \epsilon L_W (6 + 8\lambda) \left(\mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t) \mathbb{1}_{\mathrm{Reg}_3} \\ =& \Big[\epsilon (c - \gamma) \left(H(X_t, V_t) + H(\tilde{X}_t, \tilde{V}_t) \right) + 2\epsilon B + c \\ &+ \epsilon L_W (6 + 8\lambda) \left(\mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) \Big] f(r_t) \mathbb{1}_{\mathrm{Reg}_3} \end{aligned}$$

Since $c - \gamma < 0$ as a consequence of (2.2.9), and using Lemma 2.2.3

$$\begin{aligned} K_t &\leqslant \left((c - \gamma) \epsilon \frac{12}{5} \frac{B}{\gamma} + 2\epsilon B + c \right) f(r_t) + \epsilon L_W (6 + 8\lambda) \left(\mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t) \\ &\leqslant \left(c \left(\frac{12\epsilon}{5} \frac{B}{\gamma} + 1 \right) - \frac{2}{5} \epsilon B \right) f(r_t) + \epsilon L_W (6 + 8\lambda) \left(\mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t). \end{aligned}$$

Then, using (2.2.9), $\mathbb{E} K_t \mathbb{1}_{\mathrm{Reg}_3} \leqslant L_W \mathcal{C}_K \mathbb{E}\rho_t + L_W \mathcal{C}_K^0 \mathbb{E}\rho_t e^{-\gamma t}$.

2.3.3 Step three : Convergence

Let Γ be a coupling of ν_0^1 and ν_0^2 such that $\mathbb{E}_\Gamma \rho((x,v),(\tilde{x},\tilde{v})) < \infty$. We consider the coupling of (X_t, V_t) and $(\tilde{X}_t, \tilde{V}_t)$, with initial distribution $((X_0, V_0), (\tilde{X}_0, \tilde{V}_0)) \sim \Gamma$, previously introduced. Using Lemma 2.3.2 and Lemma 2.3.3, by taking the expectation in (2.3.5) at stopping times τ_n increasingly converging to t , we have by Fatou's lemma for $n \to \infty$, $\forall \xi > 0, \forall t \geqslant 0$,

$$e^{ct} \mathbb{E}\rho_t \leqslant \mathbb{E}\rho_0 + (1 + \alpha) \xi \int_0^t e^{cs} \mathbb{E}(G_s) ds + L_W \mathcal{C}_K \int_0^t e^{cs} \mathbb{E}\rho_s ds + L_W \mathcal{C}_K^0 \int_0^t e^{(c-\gamma)s} \mathbb{E}\rho_s ds. \tag{2.3.12}$$

Moreover, using Lemma 2.2.2 and the fact $\gamma > c$, for all $t \geqslant 0$,

$$\mathbb{E}(G_t) \leqslant \left(1 + \epsilon \mathcal{C}_H^0 + \epsilon \mathcal{C}_{\tilde{H}}^0\right), \qquad \int_0^t e^{(c-\gamma)s} \mathbb{E}\rho_s ds \leqslant \frac{f(R_1) \left(1 + \epsilon \mathcal{C}_H^0 + \epsilon \mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c},$$

$$\int_0^t e^{cs} ds = \frac{c}{c - L_W \mathcal{C}_K} \frac{e^{ct} - 1}{c} - \frac{L_W \mathcal{C}_K}{c - L_W \mathcal{C}_K} \int_0^t e^{cs} ds.$$

We get

$$\begin{aligned} & e^{ct} \left(\mathbb{E}\rho_t - \frac{(1+\alpha)\xi}{c - L_W \mathcal{C}_K} \left(1 + \epsilon \mathcal{C}_H^0 + \epsilon \mathcal{C}_{\tilde{H}}^0\right) \right) \\ \leqslant & \mathbb{E}\rho_0 - \frac{(1+\alpha)\xi}{c - L_W \mathcal{C}_K} \left(1 + \epsilon \mathcal{C}_H^0 + \epsilon \mathcal{C}_{\tilde{H}}^0\right) + L_W \mathcal{C}_K^0 \frac{f(R_1) \left(1 + \epsilon \mathcal{C}_H^0 + \epsilon \mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c} \\ & + L_W \mathcal{C}_K \int_0^t e^{cs} \left(\mathbb{E}\rho_s - \frac{(1+\alpha)\xi}{c - L_W \mathcal{C}_K} \left(1 + \epsilon \mathcal{C}_H^0 + \epsilon \mathcal{C}_{\tilde{H}}^0\right) \right) ds. \end{aligned}$$

Gronwall's lemma yields, for all $t \geqslant 0$

$$e^{ct}\left(\mathbb{E}\left(\rho_t\right) - \frac{(1+\alpha)\,\xi}{c - L_W\mathcal{C}_K}\left(1+\epsilon\mathcal{C}_H^0+\epsilon\mathcal{C}_{\tilde{H}}^0\right)\right) \\ \leqslant \left(\mathbb{E}\left(\rho_0\right) + L_W\mathcal{C}_K^0\frac{f(R_1)\left(1+\epsilon\mathcal{C}_H^0+\epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c}\right)e^{L_W\mathcal{C}_K t}.$$

Since $\mathcal{W}_\rho\left(\mu_t,\nu_t\right) \leqslant \mathbb{E}\left(\rho_t\right)$, we have thus obtained for all $t \geqslant 0$

$$\mathcal{W}_\rho\left(\nu_t^1,\nu_t^2\right) \leqslant \frac{(1+\alpha)\,\xi}{c - L_W\mathcal{C}_K}\left(1+\epsilon\mathcal{C}_H^0+\epsilon\mathcal{C}_{\tilde{H}}^0\right) \\ + \left(\mathbb{E}\left(\rho_0\right) + L_W\mathcal{C}_K^0\frac{f(R_1)\left(1+\epsilon\mathcal{C}_H^0+\epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c}\right)e^{(L_W\mathcal{C}_K - c)t}$$

Taking the infimum over all couplings Γ of the initial conditions and using the fact that the left hand side does not depend on ξ, so that we may take $\xi = 0$, we get finally that for all $t \geqslant 0$,

$$\mathcal{W}_\rho\left(\nu_t^1,\nu_t^2\right) \leqslant \left(\mathcal{W}_\rho\left(\nu_0^1,\nu_0^2\right) + L_W\mathcal{C}_K^0\frac{f(R_1)\left(1+\epsilon\mathcal{C}_H^0+\epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c}\right)e^{(L_W\mathcal{C}_K - c)t}, \tag{2.3.13}$$

and since, by Lemma 2.2.6, $\mathcal{C}_1\mathcal{W}_\rho\left(\nu_t^1,\nu_t^2\right) \geqslant \mathcal{W}_1\left(\nu_t^1,\nu_t^2\right)$ and $\mathcal{C}_2\mathcal{W}_\rho\left(\nu_t^1,\nu_t^2\right) \geqslant \mathcal{W}_2^2\left(\nu_t^1,\nu_t^2\right)$,

$$\mathcal{W}_1\left(\nu_t^1,\nu_t^2\right) \leqslant e^{-(c-L_W\mathcal{C}_K)t}C^1_{\nu_0^1,\nu_0^2},$$
$$\mathcal{W}_2^2\left(\nu_t^1,\nu_t^2\right) \leqslant e^{-(c-L_W\mathcal{C}_K)t}C^2_{\nu_0^1,\nu_0^2}.$$

Then, for all W such that $L_W < c/\mathcal{C}_K$, there will be contraction at rate $\tau := c - L_W\mathcal{C}_K > 0$. So, it only remains for L_W to satisfy

$$L_W \leqslant \frac{c}{\mathcal{C}_1\left(1+\frac{2\epsilon B}{\gamma}\right) + \frac{2\epsilon B}{\gamma\lambda}\left(6+8\lambda\right)}, \tag{2.3.14}$$

with

$$\mathcal{C}_1 = \max\left(\frac{2}{\alpha},1\right)\max\left(\frac{4\left((1+\alpha)^2+\alpha^2\right)}{\epsilon\min\left(\frac{2}{3}\lambda,6\right)f(1)},\frac{1}{\phi\left(R_1\right)g\left(R_1\right)}\right).$$

Remark 2.3.1. *We draw the reader's attention to the fact that Theorem 2.3.1 is then a consequence of everything we have done so far : if we have an upper bound on* $\mathbb{E}H\left(X_0,V_0\right) + \mathbb{E}H(\tilde{X}_0,\tilde{V}_0)$*, the constant* $\mathcal{C}_K^0$ *in Lemma 2.3.3 can be chosen equal to 0 provided we modify* $\mathcal{C}_K$*.*

Let us now show that there is existence and uniqueness of a stationary measure. Let $\mathcal{C}^0 > \frac{B}{\gamma}$ and μ_t a solution of (2.1.1) such that $\mathbb{E}_{\mu_0}H \leqslant \mathcal{C}^0$. Using (2.2.6), for all $t \geqslant 0$, $\mathbb{E}_{\mu_t}H \leqslant \mathcal{C}^0$. Thanks to Theorem 2.3.1, for L_W sufficiently small, there is $\tau > 0$ such that for all $t \geqslant s \geqslant 0$

$$\mathcal{W}_\rho\left(\mu_t,\mu_s\right) \leqslant e^{-\tau s}\mathcal{W}_\rho\left(\mu_{t-s},\mu_0\right) \leqslant f(R_1)\left(1+2\epsilon\mathcal{C}^0\right)e^{-\tau s},$$

and thus

$$\mathcal{W}_1\left(\mu_t,\mu_s\right)\leqslant \mathcal{C}_1 f(R_1)\left(1+2\epsilon\mathcal{C}^0\right)e^{-\tau s}.$$

The space of probability measure with first moments, equipped with the $\mathcal{W}_1$ distance, being a complete metric space (see for instance [19]), and μ_t being a Cauchy sequence, there exists μ_∞ such that

$$\mathcal{W}_1\left(\mu_t,\mu_\infty\right)\to 0 \text{ as } t\to\infty,$$

and μ_∞ stationary. Theorem 2.1.1 then ensures uniqueness and convergence towards this stationary measure.

2.4 Proof of Theorem 2.1.2

In this section, we show how we obtain similar results for the convergence of the particle system to the non-linear kinetic Langevin diffusion using the same tools. We start by introducing the coupling, the new Lyapunov function, we give a new definition for the various quantities we consider, and then prove contraction of the coupling semimetric.

2.4.1 Coupling

We consider the following coupling

$$\begin{cases} d\bar{X}_t^i=\bar{V}_t^i dt\\ d\bar{V}_t^i=-\bar{V}_t^i dt-\nabla U\left(\bar{X}_t^i\right)dt-\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)dt\\ \qquad\qquad +\sqrt{2}\left(rc\left(Z_t^i,W_t^i\right)dB_t^{rc,i}+sc\left(Z_t^i,W_t^i\right)dB_t^{sc,i}\right)\\ \bar{\mu}_t=\mathcal{L}\left(\bar{X}_t^i\right)\\ dX_t^{i,N}=V_t^{i,N}dt\\ dV_t^{i,N}=-V_t^{i,N}dt-\nabla U(X_t^{i,N})dt-\frac{1}{N}\sum_{j=1}^N\nabla W(X_t^{i,N}-X_t^{j,N})dt\\ \qquad\qquad +\sqrt{2}\left(rc\left(Z_t^i,W_t^i\right)\left(Id-2e_t^ie_t^{i,T}\right)dB_t^{rc,i}+sc\left(Z_t^i,W_t^i\right)dB_t^{sc,i}\right), \end{cases}$$

with, similarly as before,

$$Z_t^i=\bar{X}_t^i-X_t^{i,N},\quad W_t^i=\bar{V}_t^i-V_t^{i,N},\quad Q_t^i=Z_t^i+W_t^i,\ e_t^i=\begin{cases}\frac{Q_t^i}{|Q_t^i|}\text{ if } Q_t^i\neq 0,\\ 0 \text{ otherwise.}\end{cases}$$

Let $\mu_t^N:=\frac{1}{N}\sum_{i=1}^N\delta_{X_t^{i,N}}$ be the empirical distribution of the particle system, with i.i.d initial conditions $X_0^{i,N}\sim\nu_0$. We first notice that the particles are exchangeable. The generator of the process given by the particle system (2.1.3) is, for a function ϕ of $(x_1,...,x_N,v_1,...,v_N)$

$$\mathcal{L}^N\phi=\sum_{i=1}^N\mathcal{L}^{i,N}\phi,$$

with

$$\mathcal{L}^{i,N}\phi=v_i\cdot\nabla_{x_i}\phi-v_i\cdot\nabla_{v_i}\phi-\nabla U\left(x_i\right)\cdot\nabla_{v_i}\phi-\frac{1}{N}\sum_{j=1}^N\nabla W\left(x_i-x_j\right)\cdot\nabla_{v_i}\phi+\Delta_{v_i}\phi.$$

We define

$$r_t^i = \alpha|Z_t^i| + |Q_t^i|, \tag{2.4.1}$$

$$\tilde{H}(x,v) = \int_0^{H(x,v)} \exp\left(a\sqrt{u}\right) du, \tag{2.4.2}$$

$$G_t^i = 1 + \epsilon\tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \epsilon\tilde{H}(X_t^{i,N}, V_t^{i,N}) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N}), \tag{2.4.3}$$

$$\rho_t = \frac{1}{N}\sum_{i=1}^N f\left(r_t^i\right) G_t^i. \tag{2.4.4}$$

There are two ideas when constructing this new G_t^i compared to the previous section. First, we consider a modification of the Lyapunov function $\tilde{H}$, which we will describe in the next subsection. Second, we add these empirical means of the form $\frac{1}{N}\sum \tilde{H}$. This will allow us to deal with the non linearity appearing in the calculations. Recall the expectation in (2.3.9) : this term will become an empirical mean, see (2.4.23) and (2.4.24) below. When taking the expectation of what we will denote K_t^i (similar to K_t given in Lemma 2.3.2), we no longer have a product of expectations, which we were able to deal with using the uniform in time bounds, but an expectation of the product. We will therefore have to control a quantity on the particle i multiplied by a quantity on the particle j, and we do not have independence within the particle system. Hence the necessity, in the calculations, of adding this empirical mean of Lyapunov functions.

2.4.2 A modified Lyapunov function

Notice how in the expression of G^i above we did not consider the Lyapunov function H, but instead $\tilde{H}$. Let us assume there exist $\mathcal{C}_0, a > 0$ such that $\mathbb{E}_{\nu_0}\left(\tilde{H}(X,V)^2\right) \leqslant (\mathcal{C}_0)^2$ (which is equivalent to the existence of $\tilde{\mathcal{C}^0}, \tilde{a} > 0$ such that $\mathbb{E}_{\nu_0}\left(e^{\tilde{a}(|X|+|V|)}\right) \leqslant \tilde{\mathcal{C}^0}$, as it was stated in Theorem 2.1.2). First, notice

$$\tilde{H}(x,v) = \int_0^{H(x,v)} \exp\left(a\sqrt{u}\right) du = \frac{2}{a^2}\exp\left(a\sqrt{H(x,v)}\right)\left(a\sqrt{H(x,v)} - 1\right) + \frac{2}{a^2}.$$

The idea of considering the exponential of the Lyapunov function is common when trying to obtain a greater restoring force, see for instance [128].

Here, for technical reasons (more precisely when dealing with the last term of A_t^i given below in (2.4.18)) we have to ensure, when writing $\tilde{H} = f(H)$, that f' is of order $e^{\sqrt{x}}$ instead of e^x.

Direct calculations yield the following technical lemma.

Lemma 2.4.1. *We have, for all $x, v \in \mathbb{R}^d$*

$$H(x,v)\exp\left(a\sqrt{H(x,v)}\right) \geqslant \tilde{H}(x,v) \geqslant \exp\left(a\sqrt{H(x,v)}\right) - \frac{2}{a^2}\left(\exp\left(\frac{a^2}{2}\right) - 1\right), \tag{2.4.5}$$

$$\frac{2}{a}\sqrt{H(x,v)}\exp\left(a\sqrt{H(x,v)}\right) \geqslant \tilde{H}(x,v) \geqslant \frac{1}{a}\sqrt{H(x,v)}\exp\left(a\sqrt{H(x,v)}\right) - \frac{1}{a^2}(e-2), \tag{2.4.6}$$

$$\tilde{H}(x,v) \geqslant H(x,v) \tag{2.4.7}$$

We may calculate, using (2.2.2) and (2.2.3)

$$\begin{aligned}\mathcal{L}_\mu\left(\tilde{H}\right) &= \exp\left(a\sqrt{H}\right)\mathcal{L}_\mu H + \frac{a}{2\sqrt{H}}\exp\left(a\sqrt{H}\right)|\nabla_v H|^2\\ &= \exp\left(a\sqrt{H}\right)\mathcal{L}_\mu H + 24^2\frac{a}{2\sqrt{H}}\exp\left(a\sqrt{H}\right)\left|\frac{x}{2}+v\right|^2\\ &\leqslant \exp\left(a\sqrt{H}\right)\left(B + L_W\left(6+8\lambda\right)\mathbb{E}_\mu\left(|x|\right)^2 - \left(\frac{3}{4}\lambda+\lambda^2\right)|x|^2 - \gamma H\right) \qquad (2.4.8)\\ &\quad + 24a\sqrt{H}\exp\left(a\sqrt{H}\right)\\ &\leqslant \exp\left(a\sqrt{H}\right)\left(B + \frac{288a^2}{\gamma} + L_W\left(6+8\lambda\right)\mathbb{E}_\mu\left(|x|\right)^2 - \frac{\gamma}{2}H\right), \qquad (2.4.9)\end{aligned}$$

where for this last inequality we used Young's inequality $24a\sqrt{H} \leqslant \frac{\gamma}{2}H + 288\frac{a^2}{\gamma}$.

Notice that (2.4.9) ensures that this new Lyapunov function also tends to bring back particle which ventured at infinity, and at an even greater rate. This new rate $H\exp(\sqrt{H})$ however comes at a cost : the initial condition must have a finite exponential moment, and not just a finite second moment as in Section 2.3.

First, by (2.2.6) and (2.4.7),

$$\mathbb{E}(|\bar{X}_t^i|)^2 \leqslant \frac{1}{\lambda}\mathbb{E}\left(H\left(\bar{X}_t^i,\bar{V}_t^i\right)\right) \leqslant \frac{1}{\lambda}\left(\frac{B}{\gamma} + \mathbb{E}H\left(\bar{X}_0^i,\bar{V}_0^i\right)\right) \leqslant \frac{1}{\lambda}\left(\frac{B}{\gamma}+\mathcal{C}^0\right).$$

Furthermore, the function $h \mapsto \exp\left(a\sqrt{h}\right)\left(\tilde{B} - \frac{\gamma}{4}h\right)$ is bounded from above for $h \geqslant 0$ and $\tilde{B} \in \mathbb{R}$. We therefore obtain from (2.4.9) the existence of $\tilde{B}$ such that

$$\mathcal{L}_{\bar{\mu}_t^{\otimes N}}\tilde{H}\left(x_i,v_i\right) \leqslant \tilde{B} - \frac{\gamma}{4}\left(H\left(x_i,v_i\right)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)\right) \qquad (2.4.10)$$

$$\frac{d}{dt}\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right) \leqslant \tilde{B} - \frac{\gamma}{4}\mathbb{E}\left(H\left(\bar{X}_t^i,\bar{V}_t^i\right)\exp\left(a\sqrt{H\left(\bar{X}_t^i,\bar{V}_t^i\right)}\right)\right) \qquad (2.4.11)$$

$$\text{and}\quad \frac{d}{dt}\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right) \leqslant \tilde{B} - \frac{\gamma}{4}\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right), \qquad (2.4.12)$$

where for this last inequality, we used (2.4.5). While (2.4.10) and (2.4.11) will be useful in ensuring a sufficient restoring force, Equation (2.4.12) give us a uniform in time bound on $\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)$, provided we have an initial bound.

Now, for the system of particle, we have, using (2.4.9), $\forall i \in \{1,...,N\}$, $\forall x_i, v_i \in \mathbb{R}^d$,

$$\begin{aligned}&\mathcal{L}^{i,N}\tilde{H}\left(x_i,v_i\right)\\ &\quad\leqslant \exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)\left(B + \frac{288a^2}{\gamma} + L_W\left(6+8\lambda\right)\left(\frac{\sum_{j=1}^N|x_j|}{N}\right)^2 - \frac{\gamma}{2}H\left(x_i,v_i\right)\right).\end{aligned}$$

Summing over $i \in \{1,..,N\}$, we may calculate

$$L_W\left(6+8\lambda\right)\sum_{j=1}^N\left(\frac{\sum_{j=1}^N|x_j|}{N}\right)^2\sum_{i=1}^N\frac{\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)}{N}$$

$$
\begin{aligned}
&- \frac{\gamma}{8}\sum_{i=1}^{N} \frac{H\left(x_i,v_i\right)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)}{N} \\
&\leqslant \frac{\gamma}{8}\left(\sum_{i,j=1}^{N} \frac{H\left(x_i,v_i\right)}{N}\frac{\exp\left(a\sqrt{H\left(x_j,v_j\right)}\right)}{N} - \sum_{i=1}^{N} \frac{H\left(x_i,v_i\right)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)}{N}\right) \\
&\leqslant 0. \qquad (2.4.13)
\end{aligned}
$$

Here, we used (2.2.2), the fact that $\forall x,y \geqslant 0\ \ xe^{a\sqrt{y}}+ye^{a\sqrt{x}}-xe^{a\sqrt{x}}-ye^{a\sqrt{y}} = (e^{a\sqrt{x}}-e^{a\sqrt{y}})(y-x) \leqslant 0$ and assumed

$$
6\frac{L_W}{\lambda}\left(1+\frac{4}{3}\lambda\right) \leqslant \frac{\gamma}{8} \quad \text{i.e} \quad L_W \leqslant \frac{\gamma\lambda}{16(3+4\lambda)}.
$$

Likewise, there is a constant, which for the sake of clarity we will also denote $\tilde{B}$ (as we may take the maximum of the previous constants), such that we get

$$
\begin{aligned}
\mathcal{L}^{i,N}\tilde{H}(x_i,v_i) \leqslant & \tilde{B} + L_W\left(6+8\lambda\right)\left(\frac{\sum_{j=1}^N |x_i|}{N}\right)^2 \exp\left(a\sqrt{H\left(x_i,v_i\right)}\right) \\
& - \frac{\gamma}{4}H\left(x_i,v_i\right)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right), \qquad (2.4.14)\\
\mathcal{L}^{N}\left(\frac{1}{N}\sum_{i=1}^N \tilde{H}(x_i,v_i)\right) \leqslant & \tilde{B} - \frac{\gamma}{4}\left(\frac{1}{N}\sum_{i=1}^N H(x_i,v_i)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)\right), \qquad (2.4.15)
\end{aligned}
$$

and

$$
\mathcal{L}^{N}\left(\frac{1}{N}\sum_{i=1}^N \tilde{H}(x_i,v_i)\right) \leqslant \tilde{B} - \frac{\gamma}{4}\left(\frac{1}{N}\sum_{i=1}^N \tilde{H}(x_i,v_i)\right). \qquad (2.4.16)
$$

Once again, (2.4.14) and (2.4.15) will be ensure a sufficient restoring force, and (2.4.16) ensures a uniform in time bound on the expectation of $\tilde{H}(X_t^{i,N},V_t^{i,N})$, since $\mathbb{E}\left(\frac{1}{N}\sum_{j=1}^N \tilde{H}(X_t^{j,N},V_t^{j,N})\right) = \mathbb{E}\left(\tilde{H}(X_t^{i,N},V_t^{i,N})\right)$ by exchangeability of the particles.

More precisely, we obtain from (2.4.12) and (2.4.16) the direct corollary

Lemma 2.4.2. *Provided the initial expectations $\mathbb{E}\left(G_0^1\right)$ and $\mathbb{E}\left(\left(G_0^1\right)^2\right)$ are finite, there are constants $\mathcal{C}_{G,1}$ and $\mathcal{C}_{G,2}$, depending on initial conditions, such that for all $t \geqslant 0$, for all $N \geqslant 0$, and all i*

$$
\mathbb{E}\left(G_t^i\right) \leqslant \mathcal{C}_{G,1} \quad \textit{and} \quad \mathbb{E}\left(\left(G_t^i\right)^2\right) \leqslant \mathcal{C}_{G,2}.
$$

Finally, since $\tilde{H}(x,v) \geqslant H(x,v)$, Lemma 2.2.6 still holds for our new semimetric.

2.4.3 New parameters

For the sake of completeness, and since this is similar to Section 2.2.2, we quickly give some explicit parameters that satisfy the various conditions arising from calculation. T hese parameters are far from optimal, and are just given to show that every constant is explicit. Let $\tilde{B}$ be given by (2.4.10)-(2.4.12), and (2.4.14)-(2.4.16). Define

$$\alpha = L_U + \frac{\lambda}{4}, \quad R_0 = \sqrt{\frac{160\tilde{B}}{\gamma \min\left(\frac{\lambda}{3}, 3\right)}} \quad \text{and} \quad R_1 = \sqrt{(1+\alpha)^2 + \alpha^2} R_0.$$

Recall the definition of $C_{dH,1}$ and $C_{dH,2}$ in (2.2.7). Denoting

$$C_{f,1} = 8\left(\left(\frac{96}{a^2}\max\left(1, \frac{1}{2\alpha}\right) + \frac{16\sqrt{3}}{a} C_{dH,1}\right)\left(\exp\left(\frac{a^2}{2}\right) - 1\right) + 16\sqrt{3}(e-2)C_{dH,2}\right)$$

$$C_{f,2} = 8\left(24\max\left(1, \frac{1}{2\alpha}\right) + 4\sqrt{3}C_{dH,1}a + 8\sqrt{3}C_{dH,2}a^2\right)$$

we set

$$c = \left\{\frac{2\tilde{B}}{5}, \frac{\gamma}{800}, \frac{1}{12}\min\left(2\sqrt{\frac{L_U + L_W}{2\pi\alpha R_1^2}}, \frac{1}{2}\left(1 - \frac{L_U + L_W}{\alpha}\right)\right)\right.$$
$$\left. \times \exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + C_{f,1} + C_{f,2}\right) R_1^2\right)\right\},$$

and $\epsilon = \frac{5c}{2\tilde{B}}$. For $s \geqslant 0$,

$$\phi(s) = \exp\left(-\frac{1}{8}\left(\frac{1}{\alpha}(L_U + L_W) + \alpha + \epsilon C_{f,1} + C_{f,2}\right) s^2\right), \qquad \Phi(s) = \int_0^s \phi(u)\, du$$

$$g(s) = 1 - \frac{c + 2\epsilon\tilde{B}}{2}\int_0^s \frac{\Phi(u)}{\phi(u)} du, \qquad f(s) = \int_0^{\min(s, R_1)} \phi(u)\, g(u)\, du.$$

This way we satisfy the following conditions

$$\begin{aligned}
c \leqslant & \frac{\gamma}{160}\left(1 - \frac{\gamma}{80\epsilon\tilde{B} + \gamma}\right) \\
\alpha > & L_U + L_W \\
\epsilon \leqslant & 1 \\
2c + 4\epsilon\tilde{B} \leqslant & 2\left(\int_0^{R_1} \frac{\Phi(u)}{\phi(u)} du\right)^{-1} \\
2c + 4\epsilon\tilde{B} \leqslant & \frac{1}{2}\left(1 - \frac{L_U + L_W}{\alpha}\right) \inf_{r \in]0, R_1]} \frac{r\phi(r)}{\Phi(r)} \\
\forall s \geqslant 0, 0 = & 4\phi'(s) + \left(\frac{1}{\alpha}(L_U + L_W) + \alpha + \epsilon C_{f,1} + C_{f,2}\right) s\phi(s)
\end{aligned}$$

2.4.4 Convergence

The goal of the section is to prove the following result

Theorem 2.4.1. *Let $U \in \mathcal{C}^1(\mathbb{R}^d)$ satisfy Assumption 2.1 and Assumption 2.2. For all $W \in$*

$\mathcal{C}^1\left(\mathbb{R}^d\right)$ *satisfying Assumption 2.4 with*

$$L_W \leqslant \min\left(\frac{\gamma\lambda}{16(3+4\lambda)}, \frac{c}{\mathcal{C}_1}, \frac{\gamma}{64\mathcal{C}_z}, \frac{\gamma a}{256\mathcal{C}_z\epsilon}\right), \tag{2.4.17}$$

and for all probability measures $\bar{\nu}_0$ *on* $\mathbb{R}^{2d}$ *such that* $\mathbb{E}_{\bar{\nu}_0}\tilde{H}^2(X,V) \leqslant (\mathcal{C}^0)^2$*, for all* N*,* $\xi > 0$*, and* $t \geqslant 0$*,*

$$e^{ct}\mathbb{E}\left(\rho_t\right) \leqslant \mathbb{E}\left(\rho_0\right) + \xi\left(1+\alpha\right)\mathcal{C}_{G,1}\int_0^t e^{cs}ds + L_W\frac{\mathcal{C}^0\mathcal{C}_{G,2}^{1/2}}{\lambda}\sqrt{\frac{8}{N}}\int_0^t e^{cs}ds.$$

Proof of Theorem 2.1.2 using Theorem 2.4.1

We first show how Theorem 2.1.2 is a consequence of Theorem 2.4.1. Let Γ be a coupling of $\nu_0^{\otimes N}$ and $\bar{\nu}_0^{\otimes N}$, such that $\mathbb{E}\rho_0 < \infty$. We consider the coupling previously introduced. For clarity, let us denote

$$A = L_W\frac{\mathcal{C}^0\mathcal{C}_{G,2}^{1/2}}{\lambda}\sqrt{8}, \quad B = (1+\alpha)\mathcal{C}_{G,1},$$

i.e

$$e^{ct}\mathbb{E}\left(\rho_t\right) \leqslant \mathbb{E}\left(\rho_0\right) + \xi B\int_0^t e^{cs}ds + \frac{A}{\sqrt{N}}\int_0^t e^{cs}ds.$$

Let us consider

$$u(t) = e^{ct}\left(\mathbb{E}\left(\rho_t\right) - \frac{A}{c}\frac{1}{\sqrt{N}} - \xi\frac{B}{c}\right)$$

Then $u(t) \leqslant u(0)$ i.e

$$\mathbb{E}\left(\rho_t\right) \leqslant \mathbb{E}\left(\rho_0\right)e^{-ct} + \frac{A}{c}\frac{1}{\sqrt{N}}\left(1-e^{-ct}\right) + \xi\frac{B}{c}\left(1-e^{-ct}\right).$$

We thus obtain the desired result, by taking the limit as $\xi \to 0$ uniformly in time, and by using the exchangeability of the particles to have $\mathbb{E}\left(\rho_t\right) = \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \rho_t^i\right) = \mathbb{E}\left(\frac{1}{k}\sum_{i=1}^k \rho_t^i\right)$ for all $k \in \mathbb{N}$.

Evolution of the coupling semimetric for the particle system

We thus need to start by considering the dynamic of ρ_t. Like in Lemma 2.3.1, we have almost surely for all $t \geqslant 0$

$$\begin{aligned} d|Q_t^i| = &-e_t^{i,T}\left(\nabla U\left(\bar{X}_t^i\right) - \nabla U(X_t^{i,N})\right)dt - e_t^{i,T}\left(\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right) - \nabla W * \bar{\mu}_t^N(X_t^{i,N})\right)dt \\ &+ 2\sqrt{2}rc\left(Z_t^i, W_t^i\right)e_t^{i,T}dB_t^{rc,i}. \end{aligned}$$

Hence $e^{ct} f\left(r_t^i\right) = f\left(r_0\right) + \hat{A}_t^i + \hat{M}_t^i$ with

$$\begin{aligned}
d\hat{A}_t^i = &\Bigg[cf\left(r_t^i\right) + \alpha f'\left(r_t^i\right) \frac{d|Z_t^i|}{dt} - f'\left(r_t^i\right) {e_t^i}^T \left(\nabla U\left(X_t^i\right) - \nabla U(X_t^{i,N})\right) \\
&- f'\left(r_t^i\right) {e_t^i}^T \left(\nabla W * \mu_t\left(X_t^i\right) - \frac{1}{N}\sum_{j=1}^N \nabla W(X_t^{i,N} - X_t^{j,N})\right) \\
&+ 4f''\left(r_t^i\right) rc^2\left(Z_t^i, W_t^i\right)\Bigg] e^{ct} dt, \\
d\hat{M}_t^i = & e^{ct} 2\sqrt{2} f'\left(r_t^i\right) rc\left(Z_t^i, W_t^i\right) {e_t^i}^T dB_t^{rc,i}.
\end{aligned}$$

We now consider the evolution of

$$G_t^i = 1 + \epsilon \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \epsilon \tilde{H}(X_t^{i,N}, V_t^{i,N}) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N}).$$

Notice how we have added new terms in G_t^i. Those additional quantities will help us in dealing with the non linearity, as will be shown later.

$$\begin{aligned}
dG_t^i = &\epsilon\left(\mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \mathcal{L}^N \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) dt \\
&+ \epsilon\sqrt{2} rc\left(Z_t^i, W_t^i\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot e_t^i {e_t^i}^T dB_t^{rc,i} \\
&+ \epsilon\sqrt{2} rc\left(Z_t^i, W_t^i\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot \left(Id - e_t^i {e_t^i}^T\right) dB_t^{rc,i} \\
&+ \epsilon\sqrt{2} sc\left(Z_t^i, W_t^i\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot dB_t^{sc,i} \\
&+ \frac{\epsilon}{N}\sum_{j=1}^N \left(\mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \mathcal{L}^N \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) dt \\
&+ \frac{\epsilon\sqrt{2}}{N}\sum_{j=1}^N rc\left(Z_t^j, W_t^j\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) - \nabla_v \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \cdot e_t^j {e_t^j}^T dB_t^{rc,j} \\
&+ \frac{\epsilon\sqrt{2}}{N}\sum_{j=1}^N rc\left(Z_t^j, W_t^j\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \nabla_v \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \cdot \left(Id - e_t^j {e_t^j}^T\right) dB_t^{rc,j} \\
&+ \frac{\epsilon\sqrt{2}}{N}\sum_{j=1}^N sc\left(Z_t^j, W_t^j\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \nabla_v \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \cdot dB_t^{sc,j}.
\end{aligned}$$

Therefore

$$e^{ct}\rho_t^i = e^{ct} f\left(r_t^i\right) G_t^i = \rho_0 + A_t^i + M_t^i, \tag{2.4.18}$$

with

$$\begin{aligned}
dA_t^i = & G_t^i d\hat{A}_t^i + \epsilon e^{ct} f\left(r_t^i\right)\Bigg(\mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \mathcal{L}^N \tilde{H}(X_t^{i,N}, V_t^{i,N}) \\
&+ \frac{1}{N}\sum_{j=1}^N \mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{1}{N}\mathcal{L}^N \sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N})\Bigg) dt
\end{aligned}$$

$$+ 4\epsilon\left(1+\frac{1}{N}\right)e^{ct}f'\left(r_t^i\right)rc^2\left(Z_t^i,W_t^i\right)\left(\nabla_v\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)-\nabla_v\tilde{H}(X_t^{i,N},V_t^{i,N})\right)\cdot e_t^i dt$$

and M_t^i is a continuous local martingale. Let us deal with this last line. For the sake of conciseness, from now on we denote for all i

$$\bar{H}_i := H\left(\bar{X}_t^i,\bar{V}_t^i\right), \quad \text{and} \quad H_i^N := H(X_t^{i,N},V_t^{i,N})$$

We have

$$\begin{aligned}
&|\nabla_v\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)-\nabla_v\tilde{H}(X_t^{i,N},V_t^{i,N})| \\
&=\left|\nabla_v\bar{H}_i\exp\left(a\sqrt{\bar{H}_i}\right)-\nabla_v H_i^N\exp\left(a\sqrt{H_i^N}\right)\right| \\
&\leqslant\left|12\bar{X}_t^i+24\bar{V}_t^i-12X_t^{i,N}-24V_t^{i,N}\right|\left(\exp\left(a\sqrt{\bar{H}_i}\right)+\exp\left(a\sqrt{H_i^N}\right)\right) \\
&\quad + a\left|12\bar{X}_t^i+24\bar{V}_t^i\right|\left|\sqrt{\bar{H}_i}-\sqrt{H_i^N}\right|\left(\exp\left(a\sqrt{\bar{H}_i}\right)+\exp\left(a\sqrt{H_i^N}\right)\right) \\
&\leqslant 24\max\left(1,\frac{1}{2\alpha}\right)r_t^i\left(\exp\left(a\sqrt{\bar{H}_i}\right)+\exp\left(a\sqrt{H_i^N}\right)\right) \\
&\quad + 4a\sqrt{3}\left|\bar{H}_i-H_i^N\right|\left(\exp\left(a\sqrt{\bar{H}_i}\right)+\exp\left(a\sqrt{H_i^N}\right)\right).
\end{aligned}$$

Now, using Lemma 2.2.7, we get

$$\begin{aligned}
&|\nabla_v\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)-\nabla_v\tilde{H}(X_t^{i,N},V_t^{i,N})| \\
&\leqslant\left(24\max\left(1,\frac{1}{2\alpha}\right)+4\sqrt{3}C_{dH,1}a\right)r_t^i\left(\exp\left(a\sqrt{\bar{H}_i}\right)+\exp\left(a\sqrt{H_i^N}\right)\right) \\
&\quad + 4\sqrt{3}C_{dH,2}ar_t^i\left(\sqrt{\bar{H}_i}+\sqrt{H_i^N}\right)\left(\exp\left(a\sqrt{\bar{H}_i}\right)+\exp\left(a\sqrt{H_i^N}\right)\right) \\
&\leqslant\left(24\max\left(1,\frac{1}{2\alpha}\right)+4\sqrt{3}C_{dH,1}a\right)r_t^i\left(\exp\left(a\sqrt{\bar{H}_i}\right)+\exp\left(a\sqrt{H_i^N}\right)\right) \\
&\quad + 8\sqrt{3}C_{dH,2}ar_t^i\left(\sqrt{\bar{H}_i}\exp\left(a\sqrt{\bar{H}_i}\right)+\sqrt{H_i^N}\exp\left(a\sqrt{H_i^N}\right)\right).
\end{aligned}$$

Hence why, using (2.4.5) and (2.4.6), we get

$$\begin{aligned}
&|\nabla_v\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)-\nabla_v\tilde{H}(X_t^{i,N},V_t^{i,N})| \\
&\leqslant\left(24\max\left(1,\frac{1}{2\alpha}\right)+4\sqrt{3}C_{dH,1}a\right)r_t^i\left(\frac{4}{a^2}\left(e^{\frac{a^2}{2}}-1\right)+\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)+\tilde{H}(X_t^{i,N},V_t^{i,N})\right) \\
&\quad + 8\sqrt{3}C_{dH,2}a^2r_t^i\left(\frac{2}{a^2}(e-2)+\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)+\tilde{H}(X_t^{i,N},V_t^{i,N})\right),
\end{aligned}$$

and thus

$$\begin{aligned}
&4\epsilon\left(1+\frac{1}{N}\right)e^{ct}f'\left(r_t^i\right)rc^2\left(Z_t^i,W_t^i\right)\left(\nabla_v\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)-\nabla_v\tilde{H}(X_t^{i,N},V_t^{i,N})\right)\cdot e_t^i dt \\
&\qquad\leqslant 8\epsilon r_t^i f'\left(r_t^i\right)e^{ct}rc^2\left(Z_t^i,W_t^i\right)
\end{aligned}$$

$$\times\left(\left(\frac{96}{a^2}\max\left(1,\frac{1}{2\alpha}\right)+\frac{16\sqrt{3}}{a}C_{dH,1}\right)\left(e^{\frac{a^2}{2}}-1\right)+16\sqrt{3}(e-2)C_{dH,2}\right)$$
$$+8r_t^if'\left(r_t^i\right)e^{ct}rc^2\left(Z_t^i,W_t^i\right)\left(24\max\left(1,\frac{1}{2\alpha}\right)+4\sqrt{3}C_{dH,1}a+8\sqrt{3}C_{dH,2}a^2\right)$$
$$\times\left(\epsilon\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)+\epsilon\tilde{H}(X_t^{i,N},V_t^{i,N})\right)$$
$$\leqslant\left(\epsilon C_{f,1}+C_{f,2}\right)r_t^if'\left(r_t^i\right)rc^2\left(Z_t^i,W_t^i\right)G_t^i.$$

Then we use

$$\begin{aligned}
&\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^N\nabla W(X_t^{i,N}-X_t^{j,N})\right|\\
&\quad\leqslant\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^N\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|\\
&\qquad+\left|\frac{1}{N}\sum_{j=1}^N\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)-\frac{1}{N}\sum_{j=1}^N\nabla W(X_t^{i,N}-X_t^{j,N})\right|,\\
&\quad\leqslant\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^N\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|\\
&\qquad+\frac{1}{N}\sum_{j=1}^N\left|\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)-\nabla W(X_t^{i,N}-X_t^{j,N})\right|,\\
&\quad\leqslant\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^N\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|+\frac{L_W}{N}\sum_{j=1}^N\left(\left|\bar{X}_t^i-X_t^{i,N}\right|+\left|\bar{X}_t^j-X_t^{j,N}\right|\right).
\end{aligned}$$

Thus

$$\begin{aligned}
&\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^N\nabla W(X_t^{i,N}-X_t^{j,N})\right|\\
&\qquad\leqslant\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^N\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|+L_W|Z_t^i|+L_W\frac{\sum_{j=1}^N|Z_t^j|}{N}.
\end{aligned}$$

And finally we use (2.4.10), (2.4.14) and (2.4.15) to have

$$\begin{aligned}
&\mathcal{L}_{\bar{\mu}_t^{\otimes N}}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)+\mathcal{L}^N\tilde{H}(X_t^{i,N},V_t^{i,N})\\
&\qquad+\frac{1}{N}\sum_{j=1}^N\mathcal{L}_{\bar{\mu}_t^{\otimes N}}\tilde{H}\left(\bar{X}_t^j,\bar{V}_t^j\right)+\frac{1}{N}\mathcal{L}^N\sum_{j=1}^N\tilde{H}(X_t^{j,N},V_t^{j,N})\\
&\quad\leqslant 4\tilde{B}+L_W\left(6+8\lambda\right)\left(\frac{\sum_{j=1}^N|X_t^{j,N}|}{N}\right)^2\exp\left(a\sqrt{H_i^N}\right)
\end{aligned}$$

$$- \frac{\gamma}{4} \bar{H}_i \exp\left(a\sqrt{\bar{H}_i}\right) - \frac{\gamma}{4} H_i^N \exp\left(a\sqrt{H_i^N}\right)$$
$$- \frac{\gamma}{4N} \sum_{j=1}^N \left(\bar{H}_j \exp\left(a\sqrt{\bar{H}_j}\right) + H_j^N \exp\left(a\sqrt{H_j^N}\right)\right).$$

We thus obtain

$$dA_t^i \leqslant e^{ct} K_t^i dt \tag{2.4.19}$$

with

$$K_t^i = f'\left(r_t^i\right) G_t^i \left(\alpha \frac{d|Z_t^i|}{dt} + (L_U + L_W)|Z_t^i| + (\epsilon C_{f,1} + C_{f,2}) r_t^i r c^2\left(Z_t^i, W_t^i\right)\right) + 2cf\left(r_t^i\right) G_t^i \tag{2.4.20}$$

$$+ 4f''\left(r_t^i\right) G_t^i r c^2\left(Z_t^i, W_t^i\right) + \left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right) - \frac{1}{N}\sum_{j=1}^N \nabla W\left(\bar{X}_t^i - \bar{X}_t^j\right)\right| f'\left(r_t^i\right) G_t^i \tag{2.4.21}$$

$$+ \epsilon f\left(r_t^i\right)\left(4\tilde{B} - \frac{\gamma}{16}\tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \frac{\gamma}{16}\tilde{H}(X_t^{i,N}, V_t^{i,N}) - \frac{\gamma}{16N}\sum_{j=1}^N \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right)\right.$$
$$\left. - \frac{\gamma}{16N}\sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \tag{2.4.22}$$

$$+ L_W \frac{\sum_{j=1}^N |Z_t^j|}{N} f'\left(r_t^i\right) G_t^i - cf\left(r_t^i\right) G_t^i$$
$$- \epsilon f\left(r_t^i\right)\left[\frac{\gamma}{16}\bar{H}_i \exp\left(a\sqrt{\bar{H}_i}\right) + \frac{\gamma}{16} H_i^N \exp\left(a\sqrt{H_i^N}\right)\right.$$
$$\left. + \frac{\gamma}{16N}\sum_{j=1}^N \bar{H}_j \exp\left(a\sqrt{\bar{H}_j}\right) + \frac{\gamma}{16N}\sum_{j=1}^N H_j^N \exp\left(a\sqrt{H_j^N}\right)\right] \tag{2.4.23}$$

$$+ \epsilon L_W (6 + 8\lambda) f\left(r_t^i\right)\left(\frac{\sum_{j=1}^N |X_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H_i^N}\right)$$
$$- \frac{\gamma\epsilon}{8} f\left(r_t^i\right)\left(H_i^N \exp\left(a\sqrt{H_i^N}\right) + \frac{1}{N}\sum_{j=1}^N H_j^N \exp\left(a\sqrt{H_j^N}\right)\right). \tag{2.4.24}$$

This formulation of K_t^i might seem cumbersome (and to some degree it is...) but we have actually grouped the various terms based on how we will have them compensate one another. Thus,

- lines (2.4.20) and (2.4.21) will be managed thanks to the construction of the function f like before, with a special care given to the last term of line (2.4.21), on which we will use a law of large number,
- line (2.4.22) will come into play when considering the "last region of space" introduced previously,
- line (2.4.23) will, under some conditions on L_W, be nonpositive when summing up all $\left(K_t^j\right)_j$,

- and finally, line (2.4.24) will be nonpositive thanks to Lemma 2.2.1, provided L_W is sufficiently small.

This highlights two important ideas in the construction of the function $\rho:$ we both added in G_t^i the empirical mean of $H(X_t^{i,N}, V_t^{i,N}) + H\left(\bar{X}_t^i, \bar{V}_t^i\right)$ and constructed a Lyapunov function with a greater restoring force. This is what allows us to tackle the non linearity appearing in (2.4.23) and (2.4.24) respectively in the terms $\frac{\sum_{j=1}^N |Z_t^j|}{N} G_t^i$ and $\left(\frac{\sum_{j=1}^N |X_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H_i^N}\right)$.

Some calculations

Like previously, we now have to show contraction in all three regions of space. Recall $f'\left(r_t^i\right) \leqslant 1$. The same calculations as before will be used, we only detail here the differences.

- First, since $\frac{L_W}{\lambda}(6+8\lambda) \leqslant \frac{\gamma}{8}$, by using Lemma 2.2.1 and since

$$H_j^N \exp\left(a\sqrt{H_i^N}\right) \leqslant H_i^N \exp\left(a\sqrt{H_i^N}\right) + H_j^N \exp\left(a\sqrt{H_j^N}\right)$$

  we obtain

$$\begin{aligned}\epsilon L_W (6+8\lambda) f\left(r_t^i\right) &\left(\frac{\sum_{j=1}^N |X_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H_i^N}\right) \\ &- \frac{\gamma\epsilon}{8N} f\left(r_t^i\right)\left(N H_i^N \exp\left(a\sqrt{H_i^N}\right) + \sum_{j=1}^N H_j^N \exp\left(a\sqrt{H_j^N}\right)\right) \leqslant 0.\end{aligned}$$

  This takes care of (2.4.24).

- We have, since $f'\left(r_t^i\right) \leqslant 1$: $\frac{1}{N}\sum_{i=1}^N \frac{\sum_{j=1}^N |Z_t^j|}{N} f'\left(r_t^i\right) G_t^i \leqslant \frac{\sum_{i,j=1}^N |Z_t^j| G_t^i}{N^2}$. Then, using Lemma 2.2.6

$$\begin{aligned}&\frac{1}{N^2}\sum_{i,j=1}^N |Z_t^i| G_t^j \\ =&\frac{1}{N}\sum_{i=1}^N |Z_t^i| + \frac{2\epsilon}{N^2}\sum_{i,j=1}^N |Z_t^i| \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{2\epsilon}{N^2}\sum_{i,j=1}^N |Z_t^i| \tilde{H}(X_t^{j,N}, V_t^{j,N}) \\ \leqslant&\frac{\mathcal{C}_1}{N}\sum_{i=1}^N \rho_t^i + \frac{2\mathcal{C}_z\epsilon}{N^2}\sum_{i,j=1}^N f(r_t^i)\left(\tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \\ &+ \frac{4\mathcal{C}_z\epsilon^2}{aN^2}\sum_{i,j=1}^N f(r_t^i)\left(\sqrt{\bar{H}_i} + \sqrt{H_i^N}\right)\left(\sqrt{\bar{H}_j}\exp\left(a\sqrt{\bar{H}_j}\right) + \sqrt{H_j^N}\exp\left(a\sqrt{H_j^N}\right)\right).\end{aligned}$$

  First, using (2.4.5)

$$\frac{2\mathcal{C}_z\epsilon}{N^2}\sum_{i,j=1}^N f(r_t^i)\left(\tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \tilde{H}(X_t^{j,N}, V_t^{j,N})\right)$$

$$\leqslant \frac{2\mathcal{C}_z\epsilon}{N^2}\sum_{i,j=1}^{N} f(r_t^i)\left(\bar{H}_j \exp\left(a\sqrt{\bar{H}_j}\right) + H_j^N \exp\left(a\sqrt{H_j^N}\right)\right).$$

Since

$$\begin{aligned}
&\left(\sqrt{\bar{H}_i}+\sqrt{H_i^N}\right)\left(\sqrt{\bar{H}_j}\exp\left(a\sqrt{\bar{H}_j}\right)+\sqrt{H_j^N}\exp\left(a\sqrt{H_j^N}\right)\right)\\
&\quad\leqslant 2\Big[\bar{H}_i\exp\left(a\sqrt{\bar{H}_i}\right)+\bar{H}_j\exp\left(a\sqrt{\bar{H}_j}\right)\\
&\qquad + H_i^N\exp\left(a\sqrt{H_i^N}\right)+H_j^N\exp\left(a\sqrt{H_j^N}\right)\Big],
\end{aligned}$$

we have

$$\begin{aligned}
&\frac{4\mathcal{C}_z\epsilon^2}{aN^2}\sum_{i,j=1}^{N} f(r_t^i)\left(\sqrt{\bar{H}_i}+\sqrt{H_i^N}\right)\left(\sqrt{\bar{H}_j}\exp\left(a\sqrt{\bar{H}_j}\right)+\sqrt{H_j^N}\exp\left(a\sqrt{H_j^N}\right)\right)\\
&\quad\leqslant \frac{8\mathcal{C}_z\epsilon^2}{aN^2}\sum_{i,j=1}^{N} f(r_t^i)\Big[\bar{H}_i\exp\left(a\sqrt{\bar{H}_i}\right)+\bar{H}_j\exp\left(a\sqrt{\bar{H}_j}\right)\\
&\qquad + H_i^N\exp\left(a\sqrt{H_i^N}\right)+H_j^N\exp\left(a\sqrt{H_j^N}\right)\Big].
\end{aligned}$$

This way, since $2\mathcal{C}_z L_W \leqslant \frac{\gamma}{32}$, $L_W\epsilon\frac{8\mathcal{C}_z}{a} \leqslant \frac{\gamma}{32}$, and $L_W\mathcal{C}_1 \leqslant c$, we get

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\Bigg(L_W\frac{\sum_{j=1}^N |Z_t^j|}{N}f'\left(r_t^i\right)G_t^i - cf\left(r_t^i\right)G_t^i\\
&\quad -\epsilon f\left(r_t^i\right)\left(\frac{\gamma}{16}\bar{H}_i\exp\left(a\sqrt{\bar{H}_i}\right)+\frac{\gamma}{16}H_i^N\exp\left(a\sqrt{H_i^N}\right)\right.\\
&\qquad \left.+\frac{\gamma}{16N}\sum_{j=1}^{N}\bar{H}_j\exp\left(a\sqrt{\bar{H}_j}\right)+\frac{\gamma}{16N}\sum_{j=1}^{N}H_j^N\exp\left(a\sqrt{H_j^N}\right)\right)\Bigg)\leqslant 0
\end{aligned}$$

- Using Cauchy-Schwarz inequality

$$\begin{aligned}
&\mathbb{E}\left(G_t^i\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|\right)\\
&\quad\leqslant \mathbb{E}\left({G_t^i}^2\right)^{1/2}\mathbb{E}\left(\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\right)^{1/2},\\
&\quad\leqslant \mathbb{E}\left({G_t^i}^2\right)^{1/2}\mathbb{E}\left(\mathbb{E}\left(\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\middle|\bar{X}_t^i\right)\right)^{1/2}.
\end{aligned}$$

Moreover, we notice that given $\bar{X}_t^i$, the random variables $\bar{X}_t^j$ for $j\neq i$ are i.i.d with law

$\bar{\mu}_t$. Hence

$$\mathbb{E}\left(\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right) - \frac{1}{N-1}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i - \bar{X}_t^j\right)\right|^2 \Big| \bar{X}_t^i\right)$$
$$= \frac{1}{N-1}\mathrm{Var}_{\bar{\mu}_t}\left(\nabla W\left(\bar{X}_t^i - \cdot\right)\right) \leqslant \frac{4L_W^2}{N-1}\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right),$$

so

$$\mathbb{E}\left(\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right) - \frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i - \bar{X}_t^j\right)\right|^2\right)$$
$$\leqslant \mathbb{E}\left(\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right) - \frac{1}{N-1}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i - \bar{X}_t^j\right)\right|^2\right)$$
$$+\mathbb{E}\left(\left|\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i - \bar{X}_t^j\right) - \frac{1}{N-1}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i - \bar{X}_t^j\right)\right|^2\right),$$
$$\leqslant \frac{4L_W^2}{N-1}\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right) + \left(\frac{1}{N-1} - \frac{1}{N}\right)^2 N\sum_{j=1}^{N} L_W^2\mathbb{E}\left(|\bar{X}_t^i - \bar{X}_t^j|^2\right),$$
$$\leqslant 4L_W^2\left(\frac{1}{N-1} + \frac{1}{(N-1)^2}\right)\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right).$$

We may then use $\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right) \leqslant \frac{\mathcal{C}^0}{\lambda}$.

Thus, by the same exact construction as before, we can obtain the existence of a function f and a constant $c > 0$ such that in all regions of space, for L_W sufficiently small,

$$\mathbb{E}\left(\frac{1}{N}\sum_i K_t^i\right) \leqslant \xi\left(1+\alpha\right)\mathcal{C}_{G,1} + L_W\frac{\mathcal{C}^0\mathcal{C}_{G,2}^{1/2}}{\lambda}\left(\frac{4}{N-1} + \frac{4}{(N-1)^2}\right)^{1/2}.$$

By taking the expectation in the dynamic of ρ_t given by (2.4.18) and(2.4.19) at stopping times τ_n increasingly converging to t , we prove Theorem 2.4.1 by using Fatou's lemma for $n \to \infty$.

2.5 ∇U locally Lipschitz continuous

As previously mentioned, the new Lyapunov function $\tilde{H}$ given in the previous section allows for a greater restoring force, recall (2.4.9). Let us now see how using this function allows for a perturbation of the global Lipschitz Assumption.

In this section we replace Assumption 2.2 with Assumption 2.3. We assume, for ν_0^1 and ν_0^2 the initial conditions,

$$\forall i \in \{1,2\},\ \mathbb{E}_{\nu_0^i}\left(\left(\int_0^{H(X,V)} e^{a\sqrt{u}}du\right)^2\right) \leqslant (\mathcal{C}^0)^2 \qquad (2.5.1)$$

We show how the proof can be modified to still obtain contraction. As explained in Assumption 2.3, the coefficient L_ψ will be considered sufficiently small with respect to the parameters of the problem. For now, let us simply assume L_ψ is smaller than an *a priori* bound, for instance $L_\psi \leqslant 1$. Some conditions on L_ψ will appear in the calculations below and we will deal with these later.

Like previously, we consider

$$G_t = 1 + \epsilon \tilde{H}\left(X_t, V_t\right) + \epsilon \tilde{H}(\tilde{X}_t, \tilde{V}_t).$$

Hence following the same method as previously we obtain

$$\begin{aligned} K_t \leqslant & G_t \Big(cf\left(r_t\right) + \alpha f'\left(r_t\right) \frac{d|Z_t|}{dt} + (L_U + L_W) f'\left(r_t\right) |Z_t| + L_W f'\left(r_t\right) \mathbb{E}\left(|Z_t|\right) && (2.5.2) \\ & \quad + 4 f''\left(r_t\right) rc^2\left(Z_t, W_t\right)\Big) + \frac{1}{2}\left(\epsilon \mathcal{C}_{f,1} + \mathcal{C}_{f,2}\right) r_t f'\left(r_t\right) rc\left(Z_t, W_t\right)^2 && (2.5.3) \\ & + \epsilon \left(2\tilde{B} - \frac{\gamma}{8}\left(\tilde{H}\left(X_t, V_t\right) + \tilde{H}(\tilde{X}_t, \tilde{V}_t)\right)\right) f\left(r_t\right) && (2.5.4) \\ & + \left(\psi\left(X_t\right) + \psi(\tilde{X}_t)\right) |Z_t| f'\left(r_t\right) G_t \\ & - \frac{\gamma\epsilon}{8}\left(H\left(X_t, V_t\right) \exp\left(a\sqrt{H\left(X_t, V_t\right)}\right) + H(\tilde{X}_t, \tilde{V}_t) \exp\left(a\sqrt{H(\tilde{X}_t, \tilde{V}_t)}\right)\right) && (2.5.5) \\ & + \epsilon \frac{L_W}{\lambda}\left(6 + 8\lambda\right) \\ & \times \left(\exp\left(a\sqrt{H\left(X_t, V_t\right)}\right) \mathbb{E} H\left(X_t, V_t\right) + \exp\left(a\sqrt{H(\tilde{X}_t, \tilde{V}_t)}\right) \mathbb{E} H(\tilde{X}_t, \tilde{V}_t)\right) f\left(r_t\right). && (2.5.6) \end{aligned}$$

We describe briefly how the terms will compensate each other before writing the calculations that are different.

- Like previously, lines (2.5.2) and (2.5.3) will be dealt with through the choice of function f, with the non linearity appearing at the end of (2.5.2) giving us a remaining expectation (cf bullet **1** below),
- line (2.5.4) will intervene like before in the last region of space (where we use that for all x in $\mathbb{R}$, $\tilde{H} \geqslant H$ to come back to calculations we've made in Section 2.3.2, cf bullet **2** below) and in the first two region of space to compensate line (2.5.5) (cf bullet **3** below),
- and line (2.5.6) will give us a remaining expectation (cf bullet **4** below).

Notice how we use the Lyapunov function to compensate ψ appearing when considering ∇U only locally Lipschitz continuous.

- **1**. We can find a constant $\mathcal{C}_{1,e}$ such that for all $x, v, \tilde{x}, \tilde{v} \in \mathbb{R}^d$,

$$|x - \tilde{x}| + |v - \tilde{v}| \leqslant \mathcal{C}_{1,e} \rho(x, v, \tilde{x}, \tilde{v}),$$

and thus

$$\mathbb{E}\left(\mathbb{E}\left(|Z_t|\right) G_t f'\left(r_t\right)\right) \leqslant \mathcal{C}_{1,e} \mathbb{E}\left(\rho_t\right) \mathbb{E}\left(G_t\right).$$

- **2**. In the last region of space, we use the fact that

$$K_t \mathbb{1}_{R_3} \leqslant \left(\left(c - \frac{\gamma}{8} \right) G_t + 2\epsilon \tilde{B} + \frac{\gamma}{8} \right) f(r_t) \tag{2.5.7}$$

We deal with (2.5.7) exactly like in Section 2.3.2.

- **3**. Let us deal with the only locally Lipschitz continuous aspect. In the first two regions of space we use $f'(r_t)|Z_t| \leqslant f'(r_t) r_t/\alpha \leqslant f(r_t)/\alpha$ and the upper bound in (2.4.6).

$$\begin{aligned}
& G_t \left(\psi(X_t) + \psi(\tilde{X}_t) \right) f'(r_t) |Z_t| \\
& \quad - \epsilon \frac{\gamma}{8} \left(H(X_t, V_t) \exp\left(a\sqrt{H(X_t, V_t)} \right) + H(\tilde{X}_t, \tilde{V}_t) \exp\left(a\sqrt{H(\tilde{X}_t, \tilde{V}_t)} \right) \right) f(r_t) \\
& \leqslant \left(\psi(X_t) + \psi(\tilde{X}_t) \right) f'(r_t) |Z_t| \\
& \quad + \frac{2\epsilon}{a\alpha} f(r_t) \left(\psi(X_t) + \psi(\tilde{X}_t) \right) \\
& \qquad \times \left(\sqrt{H(X_t, V_t)} \exp\left(a\sqrt{H(X_t, V_t)} \right) + \sqrt{H(\tilde{X}_t, \tilde{V}_t)} \exp\left(a\sqrt{H(\tilde{X}_t, \tilde{V}_t)} \right) \right) \\
& \quad - \epsilon \frac{\gamma}{8} \left(H(X_t, V_t) \exp\left(a\sqrt{H(X_t, V_t)} \right) + H(\tilde{X}_t, \tilde{V}_t) \exp\left(a\sqrt{H(\tilde{X}_t, \tilde{V}_t)} \right) \right) f(r_t),
\end{aligned}$$

On one hand, since $\psi(x) \leqslant L_\psi \sqrt{H(x,v)}$, we have

$$\psi(x) + \psi(\tilde{x}) \leqslant L_\psi \sqrt{H(x,v)} + L_\psi \sqrt{H(\tilde{x}, \tilde{v})} \leqslant L_\psi \left(\frac{H(x,v) + H(\tilde{x}, \tilde{v})}{2} + 1 \right),$$

and thus

$$\begin{aligned}
& (\psi(x) + \psi(\tilde{x})) f'(r_t) |Z_t| \\
& \quad \leqslant \frac{L_\psi}{2} \left(H(x,v) \exp\left(a\sqrt{H(x,v)} \right) + H(\tilde{x}, \tilde{v}) \exp\left(a\sqrt{H(\tilde{x}, \tilde{v})} \right) \right) \frac{f(r_t)}{\alpha} \\
& \qquad + L_\psi f'(r_t) |Z_t|
\end{aligned}$$

On the other hand

$$\begin{aligned}
& (\psi(x) + \psi(\tilde{x})) \left(\sqrt{H(x,v)} \exp\left(a\sqrt{H(x,v)} \right) + \sqrt{H(\tilde{x}, \tilde{v})} \exp\left(a\sqrt{H(\tilde{x}, \tilde{v})} \right) \right) \\
& \quad \leqslant L_\psi \left(\sqrt{H(x,v)} + \sqrt{H(\tilde{x}, \tilde{v})} \right) \\
& \qquad \times \left(\sqrt{H(x,v)} \exp\left(a\sqrt{H(x,v)} \right) + \sqrt{H(\tilde{x}, \tilde{v})} \exp\left(a\sqrt{H(\tilde{x}, \tilde{v})} \right) \right) \\
& \quad \leqslant 2L_\psi \left(H(x,v) \exp\left(a\sqrt{H(x,v)} \right) + H(\tilde{x}, \tilde{v}) \exp\left(a\sqrt{H(\tilde{x}, \tilde{v})} \right) \right).
\end{aligned}$$

This way, assuming in a first step that

$$\frac{L_\psi}{2\alpha} \leqslant \epsilon \frac{\gamma}{16} \quad \text{and} \quad \frac{4L_\psi \epsilon}{a\alpha} \leqslant \epsilon \frac{\gamma}{16}, \tag{2.5.8}$$

we get, since in the third region of space $f'(r_t) = 0$,

$$\begin{aligned} &G_t \left(\psi(X_t) + \psi(\tilde{X}_t)\right) f'(r_t)|Z_t| \\ &\quad - \epsilon\frac{\gamma}{8}\left(H(X_t,V_t)\exp\left(a\sqrt{H(X_t,V_t)}\right) + H(\tilde{X}_t,\tilde{V}_t)\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\right) f(r_t) \\ &\leqslant L_\psi |Z_t| f'(r_t) G_t. \end{aligned}$$

At this stage, lines (2.5.2), (2.5.3) and (2.5.5) (without the non linearity dealt with in bullet **1**), can be bounded by the quantity

$$\begin{aligned} \tilde{K}_t =& G_t\Big(cf(r_t) + \alpha f'(r_t)\frac{d|Z_t|}{dt} + (L_U + 1 + \frac{\lambda}{8}) f'(r_t)|Z_t| + 4f''(r_t) rc^2(Z_t,W_t)\Big) \\ &+ \frac{1}{2}(\epsilon\mathcal{C}_{f,1} + \mathcal{C}_{f,2}) r_t f'(r_t) rc(Z_t,W_t)^2, \end{aligned}$$

where we used the *a priori* bounds $0 \leqslant L_\psi \leqslant 1$ and $0 \leqslant L_W < \frac{\lambda}{8}$. The righthand side is then dealt with through the choice of the concave function f like previously.

- **4**. Likewise, we can bound

$$\begin{aligned} &\mathbb{E}\left(\epsilon\left(\exp\left(a\sqrt{H(X_t,V_t)}\right)\mathbb{E}H(X_t,V_t) + \exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\mathbb{E}H(\tilde{X}_t,\tilde{V}_t)\right) f(r_t)\right) \\ &\leqslant \mathcal{C}_{H,\tilde{H}}\mathbb{E}(\rho_t) + \mathcal{C}^0_{H,\tilde{H}}\mathbb{E}(\rho_t) e^{-\gamma t}, \end{aligned}$$

with $\mathcal{C}_{H,\tilde{H}}$ a constant independent of initial conditions and $\mathcal{C}^0_{H,\tilde{H}}$ another constant, possibly depending on initial conditions. Here, we used (2.2.6) and (2.4.5).

We can thus construct a function f and constants c and ϵ, through the same calculations as before, such that there are C and C^0 constants (resp. independent and dependent on initial conditions) such that

$$\begin{aligned} \forall t,\quad e^{ct}\mathbb{E}(\rho_t) \leqslant& \mathbb{E}(\rho_0) + \xi(1+\alpha)\mathbb{E}(G_t)e^{ct} + L_W C\int_0^t e^{cs}\mathbb{E}(\rho_s)\,ds \\ &+ L_W C^0\int_0^t e^{(c-\gamma\lambda)s}\mathbb{E}(\rho_s)\,ds. \end{aligned}$$

Since $\mathbb{E}G_t$ is bounded uniformly in time, we may now conclude using Gronwall's lemma.

We have used in the proof the assumption (2.5.8) on L_ψ. Let us explain why it can be enforced. Here, the parameter ϵ is independent of L_ψ (as above we have bounded it using the *a priori* bounds $0 \leqslant L_\psi \leqslant 1$) and is similar to the expression of ϵ given in Section 2.4.3. Using the fact that $\alpha > L_U$, we assume

$$L_\psi \leqslant c_\psi(L_U,\lambda,\tilde{A},d,a) := \min\left(\frac{L_U\gamma a}{64}, \frac{L_U\gamma\epsilon}{8}, 1\right) \tag{2.5.9}$$

with $\gamma = \frac{\lambda}{2(\lambda+1)}$.

Chapter 3

Propagation of chaos in mean field networks of FitzHugh-Nagumo neurons

Tata Yoyo qu'est-ce qu'y a sous ton grand chapeau ?

Annie Cordy, *Tata Yoyo* (1980) written by Jacques Mareuil.
(Celle-là vous l'aviez pas vue venir hein)

Outline of the current chapter

Joint work with Laetitia Colombani.

Based on : Laetitia Colombani, and Pierre Le Bris. "Propagation of chaos in mean field networks of FitzHugh-Nagumo neurons." *arXiv preprint arXiv:2206.13291* (2022) [54] **to appear in** Mathematical Neuroscience and Applications, Volume 3, (2023).

Abstract: In this chapter, we are interested in the behavior of a fully connected network of N neurons, where N tends to infinity. We assume that neurons follow the stochastic FitzHugh-Nagumo model, whose specificity is the non-linearity with a cubic term. We prove a result of uniform in time propagation of chaos of this model in a mean-field framework. We also exhibit explicit bounds.

Since the method is similar to that of the previous chapter, we only present the model and state the results.

Acknowledgements: L. C. is a PhD student under the supervision of Patrick Cattiaux and Manon Costa, and P. L B. is a PhD student under the supervision of Arnaud Guillin and Pierre Monmarché. The authors would like to thank them, as well as Samir Salem, for their help throughout the redaction of the article.

This work has been (partially) supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency.

3.1 Understanding the model

Understanding the brain activity is both a complex and important challenge in current research. Of course, interests are plentiful: characterizing brain functions, understanding structures and links between them and figuring out some phenomena - such as cyclic heartbeat. One way of modeling this activity consists in considering a very large number of individual neurons with interactions. Since the number of neurons in a human brain is around 10^{11}, and thus even "small" parts of the brain are constituted of a very large number of them, such a strategy can be considered coherent.

The main quantity we study is the membrane potential of the nerve cells: it can "easily" be observed and its modification characterizes a synapse (an interaction between neurons). Neurons regulate their electrical potential. In general, without interaction, the potential evolves with time but has quite small changes. Incoming potentials from other neurons are usually what make the neuron fire, i.e. send potential to other neurons. We will here focus on an homogeneous network of neurons and consider mean-field interactions. This way, each neuron will interact with every other one, as it can be the case in small regions of the brain. The parameters of the model will be considered the same for each neuron.

A classical model was introduced by Hodgkin and Huxley [95] using experimental data of the activity of the giant squid axon. It describes the ion exchanges K^+, Na^+ and Cl^- through the membrane and their effects on the potential. A simplification of this model is the FitzHugh-Nagumo model, which reduces the dimension: from a four-dimensional model (for one neuron) with Hodgkin-Huxley equations, we obtain a two-dimensional model. It's a compromise between the biological accuracy and the mathematical simplicity.

The deterministic FitzHugh-Nagumo model for one neuron (or one particle) is given by the following equations:

$$\begin{cases} dX_t = (X_t - (X_t)^3 - C_t - \alpha)dt \\ dC_t = (\gamma X_t - C_t + \beta)dt, \end{cases}$$

where X is the membrane potential and C is a recovery variable, called the adaptation variable. The parameters γ and β are positive constants that determine the duration of an excitation and the position of the equilibrium point of this system. Finally $\alpha \in \mathbb{R}$ is the magnitude of a stimulus current (an entrance current in the system). Note that the variable C isn't a physical quantity, and is used to allow X to mimic the behavior of the potential. This variable C has linear dynamics and provides a slower negative feedback.

This deterministic model has been widely studied. In Chapter 7 of [165], Thieullen describes the behavior of the solution of one deterministic FitzHugh-Nagumo system. The result is also extended to the case of a stochastic FitzHugh-Nagumo system, considering a noise on the dynamics of X.

In fact, noise can be introduced in both equations to model different types of randomness : when the noise appears in the first equation (dynamics of X) with a standard deviation $\sigma_X > 0$, it models a noisy presynaptic current. When it appears in the second equation (dynamics of C) with a standard deviation $\sigma_C > 0$, it describes a noisy conductance dynamic (a noise in the chemical behavior). In general, noise in this model is additive. Various mathematical questions can be studied. Some authors choose to focus on the properties of the natural macroscopic limit of the model as $N \to \infty$ when it is clearly defined (see system (3.2.2)), when others work on properties of the particles system for fixed N. These models can be quite complicated to study mathematically. The main objectives are to characterize the behavior of these models when the number of neurons N tends to $+\infty$ in a mean-field limit, and to prove whether or not there exists an equilibrium, a stationary behavior, when t tends to $+\infty$. The question of the synchronization of neurons can also be studied, since it is a phenomenon observed in different contexts, such as the generation of respiratory rhythm or complex neurological functionalities.

In [167], the authors work on the determination of firing time. They consider a stochastic FitzHugh-Nagumo model for one neuron, with Brownian noise on X, obtain approximation of firing times and compare them with numerical simulations.

Even if the majority of authors consider a noise only on one equation of the model, some study stochastic models with two noises. Berglund and Landon describe the behavior of the deterministic FitzHugh-Nagumo model for one neuron in [15], and consider the stochastic model, with noise on both equations, to work on the behavior of the interspike interval and the distribution of oscillations of the solution.

In [164], Tatchim Bemmo, Siewe Siewe and Tchawoua focus on a quite different stochastic model by considering additive noise η on the dynamics of X, and multiplicative noise ξ on the dynamics of C, both defined as sinusoidal functions of correlated Brownian motions (choosing to avoid Gaussian noises since it is an unbounded noise). They also consider a deterministic and periodic entrance signal in the first equation. They observe abrupt transitions of the membrane potential X when the intensity of the noise is gradually modified.

In general, as we mentionned, many authors focus on a noise on only one variable. In [118], León and Samson consider a FitzHugh-Nagumo model with a noise on C but not on X, i.e. $\sigma_X = 0$, and study the properties of the equations for one neuron. In particular, they focus on hypoellipticity of the model, the existence and uniqueness of an invariant probability and a mixing property by establishing a link between the model and the class of stochastic damping Hamiltonian systems. They also consider neuronal modeling questions and study the generation of spikes in function of parameters of the model. On the contrary, the article [169] focuses on stochastic FitzHugh-Nagumo model with noise in the dynamics of X, and $\sigma_C = 0$. They study one neuron in a periodically forced regime. This study relies on the theory of Markovian Random Dynamical System. The model is driven by a cosinus signal, and Uda studies the spike rate and compares it with the probability of two-points motion of membrane potential.

As we said, we consider mean-field interactions. These interactions are described by two functions K_X and K_C, applied on the difference between two states $((X_t^i, C_t^i) - (X_t^j, C_t^j))$. In particular, this type of interaction models electrical synapses.

In their article [6], Baladron, Fasoli, Faugeras and Touboul study FitzHugh-Nagumo and Hodgkin-Huxley models with mean-field interaction, only on X. They consider more general interactions, not only applied on the difference between two states, modeling chemical synapses and electrical synapses. For the FitzHugh-Nagumo model, they consider a noise on X, and prove propagation of chaos, i.e. the convergence of the law of k neurons towards the law of k independent solutions of the mean-field equations. This article is completed and clarified by the work of Bossy, Faugeras and Talay in [25].

Mischler, Quininao and Touboul consider a FitzHugh-Nagumo model in [139], with a linear interaction on X, and a noise only on X, i.e. $\sigma_C = 0$ and $K_X(z) = \lambda x$. The drift on X is not exactly the same as in the model above, but remains similar as it is a cubic function of X. They work on the properties of a solution of the McKean-Vlasov evolution PDE associated to this model and obtain the uniqueness of a global weak solution. Furthermore, they prove that there exists at least one stationary solution, and when the interaction is small, the stationary solution is unique and exponentially stable. They also exhibit numerical results with open problems, like attractive periodic solution in time. In a similar framework, Luçon and Poquet study the macroscopic limit of this mean-field model in [130], and in particular the periodicity of such a system. They analyze the influence of both noise and interaction on the emergence of periodic behavior, and prove the existence of periodic solution, exponentially attractive, when the parameters satisfy some assumptions and the drift is small enough with respect to interaction and noise. Their approach relies on a slow-fast analysis and Floquet theory.

This model can be complexified, by considering non mean-field interaction. In particular, Bayrak, Hövel and Vuksanović work on a stochastic FitzHugh-Nagumo model with a network interaction in [10]. Their type of interaction take into account a connectivity coefficient between two neurons, and a propagation velocity.

Other authors choose to complexify the model by considering stochastic FitzHugh-Nagumo with a spatial model. A second spatial derivative of X is added in the dynamics of X. Various authors study the behavior of such a model, and explore the notion of random attractors [131, 121, 126, 122] .

Finally, numerical schemes for the interacting particles system in the stochastic model can also be studied. In [149], the authors adapt Euler-Maruyama scheme to approximate the solution of the particles system.

3.2 Framework and results

Combining noise and interaction, we work on the following equations, for $1 \leqslant i \leqslant N$, where N is the number of neurons:

$$\begin{cases} dX_t^{i,N} = (X_t^{i,N} - (X_t^{i,N})^3 - C_t^{i,N} - \alpha)dt + \frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N} - Z_t^{j,N})dt + \sigma_X dB_t^{i,X} \\ dC_t^{i,N} = (\gamma X_t^{i,N} - C_t^{i,N} + \beta)dt + \frac{1}{N}\sum_{j=1}^N K_C(Z_t^{i,N} - Z_t^{j,N})dt + \sigma_C dB_t^{i,C}, \end{cases} \tag{3.2.1}$$

where we denote by $Z_t^{i,N}$ the pair $(X_t^{i,N}, C_t^{i,N})$ to simplify the notation.

We assume that $(B_t^{i,X})_i$ and $(B_t^{i,C})_i$ are independent Brownian motions. Here, we consider two Brownian noises B^X and B^C, one on each equation, and thus assume that each neuron has its own independent noises, and that there is no environmental (or shared) noise. We assume either σ_X or σ_C (or both) to be non zero.

We consider K_X and K_C to be Lipschitz continuous and respectively denote their Lipschitz constants by L_X and L_C.

The goal of this chapter is to describe the behavior of this network as the number N of neurons tends to infinity.

To describe its behavior, we consider the $\mathbb{R}^2$-valued process $(\bar{Z}_t)_{t\geqslant 0} = (\bar{X}_t, \bar{C}_t)_{t\geqslant 0}$ evolving according to the following non-linear stochastic differential equation of *McKean-Vlasov* type

$$\begin{cases} d\bar{X}_t = (\bar{X}_t - (\bar{X}_t)^3 - \bar{C}_t - \alpha)dt + K_X * \bar{\mu}_t(\bar{Z}_t)dt + \sigma_X d\bar{B}_t^X \\ d\bar{C}_t = (\gamma \bar{X}_t - \bar{C}_t + \beta)dt + K_C * \bar{\mu}_t(\bar{Z}_t)dt + \sigma_C d\bar{B}_t^C, \end{cases} \tag{3.2.2}$$

where $\bar{\mu}_t = \text{Law}(\bar{Z}_t)$ is the law at time t of the process $(\bar{X}_t, \bar{C}_t)$, and $*$ denotes the operation of convolution, i.e.

$$K_X * \bar{\mu}_t(u) = \int K_X(u - v)\bar{\mu}_t(dv).$$

To some extent, (3.2.1) can be seen as an approximation of (3.2.2) in which the operation of convolution is applied to the empirical measure $\mu_{t,\text{emp}} = \frac{1}{N}\sum_{i=1}^N \delta_{Z_t^{i,N}}$, and what we wish to prove is that, indeed, the law μ_t^N of the network (3.2.1) converges in some sense to $\bar{\mu}_t^{\otimes N}$ (i.e the law of the solution of (3.2.2) tensorized N times) as N tends to infinity. This phenomenon has been stated under the name *propagation of chaos* -an idea motivated by M. Kac [106]- as it amounts to saying that, as the number of particle increases in the system, two particles will become "more and more" independent, converging towards a tensorized law. The notion of "propagation" refers to the fact that proving such convergence at time 0 is sufficient to prove it at a later time t.

In order to prove the convergence of μ_t^N to $\bar{\mu}_t^{\otimes N}$, we follow the coupling method described in a recent work by one of the authors in [83], the result of which cannot be applied directly here. This method has been put forward by Eberle, following earlier works by Lindvall and Rogers [127].

Assumption 3.1. *K_X and K_C are Lipschitz continuous :*

$$\exists L_X \geqslant 0, \forall z, z' \in \mathbb{R}^2 \quad |K_X(z) - K_X(z')| \leqslant L_X(\|z - z'\|_1)$$

$$\exists L_C \geqslant 0, \forall z, z' \in \mathbb{R}^2 \quad |K_C(z) - K_C(z')| \leqslant L_C(\|z - z'\|_1)$$

$$K_X(0,0) = 0 \text{ and } K_C(0,0) = 0$$

Before any result on propagation of chaos, we prove that both systems (3.2.1) and (3.2.2) have well-defined solutions:

Proposition 2. *Let K_X and K_C satisfy Assumptions 3.1. We assume that for all $i \in \{1, ..., N\}$ the law of $(X_0^{i,N}, C_0^{i,N})$ and the law of $(\bar{X}_0, \bar{C}_0)$ have a moment of order 2. Then, there exists a unique strong solution for system* (3.2.1) *and a unique strong solution for system* (3.2.2).

We denote $\mathcal{W}_1$ and $\mathcal{W}_2$ the usual L^1 and L^2 Wasserstein distances.

Theorem 3.2.1. *[Non uniform in time propagation of chaos] Let K_X and K_C satisfy Assumptions 3.1. There exist explicit $C_1, C_2 > 0$, such that for all probability measures μ_0 on $\mathbb{R}^2$ with finite second moment,*

$$\mathcal{W}_1\left(\mu_t^{k,N}, \bar{\mu}_t^{\otimes k}\right) \leqslant C_1 e^{C_2 t}\frac{k}{\sqrt{N}},$$

for all $k \in \mathbb{N}$, where $\mu_t^{k,N}$ is the marginal distribution at time t of the first k neurons $\left((X_t^1, C_t^1), ..., (X_t^k, C_t^k)\right)$ of an N particles system (3.2.1) *with initial distribution $(\mu_0)^{\otimes N}$, while $\bar{\mu}_t$ is a solution of* (3.2.2) *with initial distribution μ_0.*

This first theorem is in accordance with the theorem from [109], and gives an explicit dependence in t.

Our main result consists in removing the time dependency in the previous upper bound. This uniform in time propagation of chaos however requires stronger assumptions on the interaction kernels.

Theorem 3.2.2. *[Uniform in time propagation of chaos] Let $L_{X,max}$ and $L_{C,max}$ be two (explicit) universal constants such that $L_X \leqslant L_{X,max}$ and $L_C \leqslant L_{C,max}$. Let $\mathcal{C}_{init,exp} > 0$ and $\tilde{a} > 0$. There is an explicit $c^K > 0$ such that, for all K_X and K_C satisfying Assumptions 3.1 with $L_X, L_C < c^K$, there exist explicit $B_1, B_2 > 0$, such that for all probability measures μ_0 on $\mathbb{R}^2$ satisfying $\mathbb{E}_{\mu_0}\left(e^{\tilde{a}(|X|+|C|)}\right) \leqslant \mathcal{C}_{init,exp}$,*

$$\mathcal{W}_1\left(\mu_t^{k,N}, \bar{\mu}_t^{\otimes k}\right) \leqslant B_1\frac{k}{\sqrt{N}}, \qquad \mathcal{W}_2^2\left(\mu_t^{k,N}, \bar{\mu}_t^{\otimes k}\right) \leqslant B_2\frac{k}{\sqrt{N}},$$

for all $k \in \mathbb{N}$, where $\mu_t^{k,N}$ is the marginal distribution at time t of the first k neurons $\big((X_t^1, C_t^1), ..., (X_t^k, C_t^k)\big)$ of an N particles system (3.2.1) *with initial distribution $(\mu_0)^{\otimes N}$, while $\bar{\mu}_t$ is a solution of* (3.2.2) *with initial distribution μ_0.*

At this stage we do not specify the constants $L_{X,max}$ and $L_{C,max}$. When we prove uniform in time propagation of chaos, $L_{X,max}$ and $L_{C,max}$ are *a priori* bounds : Theorem 3.2.2 above will be true for L_X and L_C sufficiently small. The condition $L_X \leqslant L_{X,max}$ and $L_C \leqslant L_{C,max}$ are therefore not restrictive conditions, and are useful in proving some parameters are independent of L_X and L_C.

The main interest of obtaining uniform in time estimates is that it allows the study and comparison of the long-time behavior of the particle system and its nonlinear limit. As previously mentioned, this work follows the method described in [83]. Beyond the result of uniform in time propagation of chaos for the FitzHugh-Nagumo model, which is in itself an interesting result, the present work is also a testament to the robustness of the coupling method.

3.3 Quick summary of how to adapt the calculations

In this section, we quickly explain how the calculations from Chapter 2 adapt to the case of (3.2.1). In what follows, we assume $\sigma_X > 0$. The case $\sigma_X = 0$ and $\sigma_C > 0$ has also been dealt with in the original article.

Modified Euclidean distance : Our goal is to find a naturaly contracting subspace. We start by considering a synchronous coupling between $(Z_t^{i,N})_i$ and $(\bar{Z}_t^i)_i$, i.e. $\tilde{B}_t^{i,X} = B_t^{i,X}$ and $\tilde{B}_t^{i,C} = B_t^{i,C}$, and we have

$$\begin{cases} dX_t^{i,N} = (X_t^{i,N} - (X_t^{i,N})^3 - C_t^{i,N} - \alpha)dt + \frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N} - Z_t^{j,N})dt + \sigma_X dB_t^{i,X} \\ dC_t^{i,N} = (\gamma X_t^{i,N} - C_t^{i,N} + \beta)dt + \frac{1}{N}\sum_{j=1}^N K_C(Z_t^{i,N} - Z_t^{j,N})dt + \sigma_C dB_t^{i,C} \end{cases}$$

and

$$\begin{cases} d\bar{X}_t^i = (\bar{X}_t^i - (\bar{X}_t^i)^3 - \bar{C}_t^i - \alpha)dt + K_X * \bar{\mu}_t(\bar{Z}_t^i)dt + \sigma_X dB_t^{i,X} \\ d\bar{C}_t^i = (\gamma \bar{X}_t^i - \bar{C}_t^i + \beta)dt + K_C * \bar{\mu}_t(\bar{Z}_t^i)dt + \sigma_C dB_t^{i,C}, \end{cases}$$

with $\bar{\mu}_t$ the law of $\bar{Z}_t^1$. Then,

$$\begin{aligned} d(X_t^{i,N} - \bar{X}_t^i) = \Bigg(&(X_t^{i,N} - \bar{X}_t^i) - \Big((X_t^{i,N})^3 - (\bar{X}_t^i)^3\Big) - (C_t^{i,N} - \bar{C}_t^i) \\ &+ \frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N} - Z_t^{j,N}) - K_X * \bar{\mu}_t(\bar{Z}_t^i)\Bigg)dt. \end{aligned}$$

We denote

$$\text{sign}(x) = \begin{cases} \frac{x}{|x|} \text{ if } x \neq 0, \\ 0 \text{ otherwise,} \end{cases}$$

and obtain, using Ito's formula,

$$\begin{aligned} &d|X_t^{i,N} - \bar{X}_t^i| \\ &= \Big(\text{sign}(X_t^{i,N} - \bar{X}_t^i)(X_t^{i,N} - \bar{X}_t^i) - \text{sign}(X_t^{i,N} - \bar{X}_t^i)\Big((X_t^{i,N})^3 - (\bar{X}_t^i)^3\Big) \end{aligned}$$

$$-\text{sign}(X_t^{i,N}-\bar{X}_t^i)(C_t^{i,N}-\bar{C}_t^i)+\text{sign}(X_t^{i,N}-\bar{X}_t^i)\frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N}-Z_t^{j,N})-K_X*\bar{\mu}_t(\bar{Z}_t^i)\Bigg)\,dt$$

$$\leqslant\left(|X_t^{i,N}-\bar{X}_t^i|-\left|(X_t^{i,N})^3-(\bar{X}_t^i)^3\right|+\left|C_t^{i,N}-\bar{C}_t^i\right|+\left|\frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N}-Z_t^{j,N})-K_X*\bar{\mu}_t(\bar{Z}_t^i)\right|\right)dt. \tag{3.3.1}$$

Similarly,

$$d(C_t^{i,N}-\bar{C}_t^i)=\left(\gamma(X_t^{i,N}-\bar{X}_t^i)-(C_t^{i,N}-\bar{C}_t^i)+\frac{1}{N}\sum_{j=1}^N K_C(Z_t^{i,N}-Z_t^{j,N})-K_C*\bar{\mu}_t(\bar{Z}_t^i)\right)dt,$$

and we obtain

$$d|C_t^{i,N}-\bar{C}_t^i|\leqslant\left(\gamma\left|X_t^{i,N}-\bar{X}_t^i\right|-|C_t^{i,N}-\bar{C}_t^i|+\left|\frac{1}{N}\sum_{j=1}^N K_C(Z_t^{i,N}-Z_t^{j,N})-K_C*\bar{\mu}_t(\bar{Z}_t^i)\right|\right)dt. \tag{3.3.2}$$

If $\sigma_X>0$, we may use the noise to get the processes $X_t^{i,N}$ and $\bar{X}_t^i$ closer, and then (3.3.2) shows that there is contraction for $|C_t^{i,N}-\bar{C}_t^i|$ when $\left|X_t^{i,N}-\bar{X}_t^i\right|$ is close to 0. We thus may consider the modified Euclidean distance : for $z=(x,c)\in\mathbb{R}^2$ and $z'=(x',c')\in\mathbb{R}^2$,

$$r(z,z')=r(x,c,x',c')=|x-x'|+\delta|c-c'|, \tag{3.3.3}$$

where $\delta>0$ is an explicit parameter.

Remark 3.3.1. *In the case $\sigma_X=0$, we then have to assume $\sigma_C>0$, and we do a change of variable, motivated by the following observation. We have, when $\sigma_X=0$*

$$\begin{aligned}d(X_t^i-\bar{X}_t^i)&=\big((X_t^i-\bar{X}_t^i)-((X_t^i)^3-(\bar{X}_t^i)^3)-(C_t^i-\bar{C}_t^i)\big)\,dt\\&\quad+\left(\frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N}-Z_t^{j,N})-K_X*\bar{\rho}(\bar{Z}_t^i)\right)dt,\\&=\big(2(X_t^i-\bar{X}_t^i)-(C_t^i-\bar{C}_t^i)-(X_t^i-\bar{X}_t^i)-((X_t^i)^3-(\bar{X}_t^i)^3)\big)\,dt\\&\quad+\left(\frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N}-Z_t^{j,N})-K_X*\bar{\rho}(\bar{Z}_t^i)\right)dt.\end{aligned}$$

Thus

$$\begin{aligned}d|X_t^i-\bar{X}_t^i|&=\big(sign(X_t^i-\bar{X}_t^i)\big(2(X_t^i-\bar{X}_t^i)-(C_t^i-\bar{C}_t^i)\big)-|(X_t^i)^3-(\bar{X}_t^i)^3|-|X_t^i-\bar{X}_t^i|\big)\,dt\\&\quad+sign(X_t^i-\bar{X}_t^i)\left(\frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N}-Z_t^{j,N})-K_X*\bar{\rho}(\bar{Z}_t^i)\right)dt.\end{aligned}$$

The quantity $|X_t^i-\bar{X}_t^i|$ is therefore naturally contracting when $|2(X_t^i-\bar{X}_t^i)-(C_t^i-\bar{C}_t^i)|$ is close to

0. *Thanks to the presence of a Brownian motion in the stochastic differential equations defining the potential C, we can now use a reflection coupling to have $|2(X_t^i - \bar{X}_t^i) - (C_t^i - \bar{C}_t^i)|$ go to 0. We thus consider the following modified distance $r(x,c,x',c') = \delta|x-x'| + |2(x-x') - (c-c')|$.*

(Exponential) Lyapunov function : One of the necessary tool is a Lyapunov function (in the same sense as in Chapter 2), for which we may consider $H : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$H(z) = H(x,c) = \frac{1}{2}\gamma x^2 + \beta x + \frac{1}{2}c^2 + \alpha c + H_0, \quad \text{with} \quad H_0 = \frac{\beta^2}{\gamma} + \alpha^2,$$

where γ, β and α are the parameters of the system (3.2.1). Like previously, we modify H into $\tilde{H}$ given for all $z \in \mathbb{R}^2$ by,

$$\tilde{H}(z) = \int_0^{H(z)} \exp\left(a\sqrt{u}\right) du = \frac{2}{a^2}\exp\left(a\sqrt{H(z)}\right)\left(a\sqrt{H(z)} - 1\right) + \frac{2}{a^2},$$

for some parameter $a > 0$.

Construction of a semimetric : Then, we consider a modified semimetric. Let $\rho((z_j, z_j')_{1\leqslant j\leqslant N})$ be defined as follows:

$$\rho\left((z_j, z_j')_{1\leqslant j\leqslant N}\right) = \frac{1}{N}\sum_{i=1}^N f\left(r\left(z_i, z_i'\right)\right) G^i\left((z_j, z_j')_j\right),$$

where f is a concave function and, for each $i \in \{1, ..., N\}$,

$$G^i\left((z_j, z_j')_j\right) = 1 + \epsilon\tilde{H}\left(z_i\right) + \epsilon\tilde{H}\left(z_i'\right) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}\left(z_j\right) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}\left(z_j'\right).$$

Coupling method : We now construct the coupling method. Let $\xi > 0$ be a parameter destined to vanish, and let $\varphi_{\text{sc}} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ and $\varphi_{\text{rc}} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ be two Lipschitz continuous functions such that

$$\begin{aligned} \forall x, \quad \varphi_{\text{sc}}^2(x) + \varphi_{\text{rc}}^2(x) &= 1 \\ \varphi_{\text{rc}}(x) &= 1 \text{ if } \xi \leqslant x \leqslant R \\ \varphi_{\text{rc}}(x) &= 0 \text{ if } x \leqslant \frac{\xi}{2} \text{ or } x \geqslant R + \xi. \end{aligned}$$

Intuitively, φ_{rc} represents the region of space in which we consider a reflection coupling, and φ_{sc} the one in which we consider a synchronous coupling. We thus simultaneously construct the following solutions

$$\begin{cases} dX_t^{i,N} = (X_t^{i,N} - (X_t^{i,N})^3 - C_t^{i,N} - \alpha)dt + \frac{1}{N}\sum_{j=1}^N K_X(Z_t^{i,N} - Z_t^{j,N})dt \\ \qquad\qquad + \sigma_X\varphi_{\text{sc}}\left(|X_t^{i,N} - \bar{X}_t^i|\right)dB_t^{i,sc,X} + \sigma_X\varphi_{\text{rc}}\left(|X_t^{i,N} - \bar{X}_t^i|\right)dB_t^{i,rc,X} \\ dC_t^{i,N} = (\gamma X_t^{i,N} - C_t^{i,N} + \beta)dt + \frac{1}{N}\sum_{j=1}^N K_C(Z_t^{i,N} - Z_t^{j,N})dt + \sigma_C dB_t^{i,C}, \end{cases}$$

and

$$\begin{cases} d\bar{X}_t^i = (\bar{X}_t^i - (\bar{X}_t^i)^3 - \bar{C}_t^i - \alpha)dt + K_X * \bar{\mu}_t(\bar{Z}_t^i)dt \\ \qquad\qquad +\sigma_X \varphi_{\mathrm{sc}}\left(|X_t^{i,N} - \bar{X}_t^i|\right) dB_t^{i,sc,X} - \sigma_X \varphi_{\mathrm{rc}}\left(|X_t^{i,N} - \bar{X}_t^i|\right) dB_t^{i,rc,X} \\ d\bar{C}_t^i = (\gamma \bar{X}_t^i - \bar{C}_t^i + \beta)dt + K_C * \bar{\mu}_t(\bar{Z}_t^i)dt + \sigma_C dB_t^{i,C}. \end{cases}$$

Notice that we consider a reflection coupling on the dynamics of X, to bring the processes closer, which will then ensure that the distance between $C^{i,N}$ and $\bar{C}^i$ will decrease via (3.3.2). We thus choose a symmetric coupling on the dynamics of C, since we do not use the Brownian motions (recall that we accept the possibility of having $\sigma_C = 0$).

Computation of the time evolution of the semimetric : With this coupling, it "only" remains to compute the time evolution of the semimetric. We denote $r_t^i = r(Z_t^{i,N}, \bar{Z}_t^i)$ and $G_t^i = G^i((Z_t^{j,N})_j, (\bar{Z}_t^j)_j)$. For all $c \in \mathbb{R}$, for each $i \in \{1, \dots, N\}$, we have by Itô's calculus

$$d(e^{ct} f(r_t^i) G_t^i) \leqslant e^{ct} K_t^i dt + dM_t^i, \tag{3.3.4}$$

where M_t^i is a continuous local martingale and K_t^i can be written as

$$K_t^i = \tilde{K}_t^i + I_t^{1,i} + I_t^{2,i} + I_t^{3,i}. \tag{3.3.5}$$

We define $\tilde{K}_t^i$, $I_t^{1,i}$, $I_t^{2,i}$ and $I_t^{3,i}$ as followed:

$$\begin{aligned}
\tilde{K}_t^i =& G_t^i \Big[2cf(r_t^i) + \frac{1}{2} f''(r_t^i) \left(2\sigma_X^2 \varphi_{\mathrm{rc}} \left(|X_t^{i,N} - \bar{X}_t^i| \right)^2 \right) \\
&+ f'(r_t^i) \Big((1 + \gamma\delta + L_X + \delta L_C)|X_t^{i,N} - \bar{X}_t^i| - |(X_t^{i,N})^3 - (\bar{X}_t^i)^3| \\
&+ (1 + L_X + \delta L_C - \delta)|C_t^{i,N} - \bar{C}_t^i| + (\epsilon \mathcal{C}_{f,1} + \mathcal{C}_{f,2})\, \sigma_X^2 \varphi_{\mathrm{rc}} \left(|X_t^{i,N} - \bar{X}_t^i| \right)^2 r_t^i \Big) \Big] \\
&+ \epsilon f(r_t^i) \left(4\tilde{B} - \frac{\lambda}{16} \tilde{H}(\bar{Z}_t^i) - \frac{\lambda}{16} \tilde{H}(Z_t^{i,N}) - \frac{\lambda}{16N} \sum_{j=1}^N \tilde{H}(\bar{Z}_t^j) - \frac{\lambda}{16N} \sum_{j=1}^N \tilde{H}(Z_t^{j,N}) \right), \\
I_t^{1,i} =& G_t^i f'(r_t^i) \left(\left| \frac{1}{N} \sum_{j=1}^N K_X(\bar{Z}_t^i - \bar{Z}_t^j) - K_X * \bar{\mu}_t(\bar{Z}_t^i) \right| \right) \\
&\qquad + \delta G_t^i f'(r_t^i) \left(\left| \frac{1}{N} \sum_{j=1}^N K_C(\bar{Z}_t^i - \bar{Z}_t^j) - K_C * \bar{\mu}_t(\bar{Z}_t^i) \right| \right), \\
I_t^{2,i} =& G_t^i f'(r_t^i) \left(\frac{L_X}{N} \left(\sum_{j=1}^N \| Z_t^{j,N} - \bar{Z}_t^j \|_1 \right) \right) + \delta G_t^i f'(r_t^i) \left(\frac{L_C}{N} \left(\sum_{j=1}^N \| Z_t^{j,N} - \bar{Z}_t^j \|_1 \right) \right) \\
&- cf(r_t^i) G_t^i - \epsilon f(r_t^i) \left[\frac{\lambda}{16} H(\bar{Z}_t^i) \exp\left(a\sqrt{H(\bar{Z}_t^i)} \right) + \frac{\lambda}{16} H(Z_t^{i,N}) \exp\left(a\sqrt{H(Z_t^{i,N})} \right) \right] \\
&- \epsilon f(r_t^i) \left[\frac{\lambda}{16N} \sum_{j=1}^N H(\bar{Z}_t^j) \exp\left(a\sqrt{H(\bar{Z}_t^j)} \right) + \frac{\lambda}{16N} \sum_{j=1}^N H(Z_t^{j,N}) \exp\left(a\sqrt{H(Z_t^{j,N})} \right) \right],
\end{aligned}$$

$$\text{and} \quad I_t^{3,i} = \epsilon f(r_t^i)\left((\alpha_X L_X + \beta_X L_C)\left(\frac{\sum_{j=1}^N |X_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H(Z_t^{i,N})}\right)\right.$$
$$+ (\alpha_C L_X + \beta_C L_C)\left(\frac{\sum_{j=1}^N |C_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H(Z_t^{i,N})}\right)$$
$$\left.-\frac{\lambda}{16}H(Z_t^{i,N})\exp\left(a\sqrt{H(Z_t^{i,N})}\right) - \frac{\lambda}{16N}\sum_{j=1}^N H(Z_t^{j,N})\exp\left(a\sqrt{H(Z_t^{j,N})}\right)\right).$$

We need a control on $\mathbb{E}(G_t^i)$, which is a consequence of the properties on $\tilde{H}$ and the assumption of the Theorem 3.2.2 on the initial condition.

Lemma 3.3.1. *There exists $\mathcal{C}_{G,1}$ and $\mathcal{C}_{G,2}$, such that for each $i \leqslant N$, for all $t > 0$, we have*

$$\mathbb{E}(G_t^i) \leqslant \mathcal{C}_{G,1} \quad \textit{and} \quad \mathbb{E}[(G_t^i)^2] \leqslant \mathcal{C}_{G,2}.$$

Each of the terms given in Equation (3.3.5) will be controlled in a different way. The following lemmas summarize it. The first term, $\tilde{K}_t^i$, contains the various behaviors we have previously identified : we deal with it either through a synchronous coupling (when the deterministic drift is contracting), or through a reflection coupling (notice the second derivative f" in $\tilde{K}_t^i$ which will provide contraction provided f is sufficiently concave). Finally, notice the effect of Lyapunov function $\tilde{H}$ which yields a restoring force in $\tilde{K}_t^i$, $I_t^{2,i}$ and $I_t^{3,i}$.

Lemma 3.3.2. *With the right choice of parameters and functions, for each $i \leqslant N$, for all $t > 0$,*

$$\mathbb{E}\tilde{K}_t^i \leqslant \xi\left(2 + \delta\gamma + L_X + \delta L_C - L_C - \frac{1 + L_X}{\delta}\right)\mathbb{E}G_t^i. \tag{3.3.6}$$

The interaction term $\frac{1}{N}K_X(Z_t^{j,N} - Z_t^{i,N}) - K_X * \bar{\mu}_t(\bar{Z}_t^i)$ can be decomposed into the following two parts : $\frac{1}{N}K_X(\bar{Z}_t^j - \bar{Z}_t^i) - K_X * \bar{\mu}_t(\bar{Z}_t^i)$ and $\frac{1}{N}\sum\left[K_X(Z_t^{j,N} - Z_t^{i,N}) - K_X(\bar{Z}_t^j - \bar{Z}_t^i)\right]$. The first part, which is in $I_t^{1,i}$, is dealt with using some form of law of large number.

Lemma 3.3.3. *With the right choice of parameters and functions, for each $i \leqslant N$, for all $t > 0$,*

$$\mathbb{E}(I_t^{1,i}) \leqslant 4\sqrt{\frac{\mathcal{C}_{init,2}\mathcal{C}_{G,2}}{N}}(L_X + L_C), \tag{3.3.7}$$

where $\mathcal{C}_{G,2}$ is defined in Lemma 3.3.1 and $\mathcal{C}_{init,2}$ is a bound on the second moment of the processes.

$I_t^{2,i}$ contains the leftovers of this decomposition and some of the additional terms of the Lyapunov function.

Lemma 3.3.4. *With the right choice of parameters and functions, for all $t > 0$,*

$$\frac{1}{N}\sum_{i=1}^N I_t^{2,i} \leqslant 0. \tag{3.3.8}$$

Finally, $I_t^{3,i}$ deals with the non linearity appearing in the dynamics of the Lyapunov function, and will be non positive for values of L_X and L_C sufficiently small.

Lemma 3.3.5. *With the right choice of parameters and functions, for each $i \leqslant N$, for all $t > 0$,*

$$I_t^{3,i} \leqslant 0. \tag{3.3.9}$$

With these four Lemmas, we can calculate

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}K_t^i =& \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\tilde{K}_t^i + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}I_t^{1,i} + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}I_t^{2,i} + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}I_t^{3,i}\\ \leqslant& \frac{1}{N}\sum_{i=1}^{N}\xi\left(2+\delta\gamma+L_X+\delta L_C - L_C - \frac{1+L_X}{\delta}\right)\mathbb{E}G_t^i + \frac{1}{N}\sum_{i=1}^{N}4\sqrt{\frac{\mathcal{C}_{init,2}\mathcal{C}_{G,2}}{N}}(L_X+L_C)\\ \leqslant& \xi\left(2+\delta\gamma+L_X+\delta L_C - L_C - \frac{1+L_X}{\delta}\right)\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}G_t^i + 4\sqrt{\frac{\mathcal{C}_{init,2}\mathcal{C}_{G,2}}{N}}(L_X+L_C)\end{aligned}$$

Since by Lemma 3.3.1, we have $\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}G_t^i \leqslant \mathcal{C}_{G,1}$, we obtain

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}K_t^i \leqslant \xi A + (L_X+L_C)\frac{B}{\sqrt{N}}$$

where A and B are constants.

For all initial couplings such that $\mathbb{E}\rho\left((Z_0^{j,N},\bar{Z}_0^j)_{1\leqslant j\leqslant N}\right) < \infty$, by taking the expectation of (3.3.4) along a sequence of increasing localizing stopping times, we have thanks to Fatou's lemma

$$\begin{aligned}e^{ct}\mathbb{E}\left(\rho\left((Z_t^{j,N},\bar{Z}_t^j)_{1\leqslant j\leqslant N}\right)\right) \leqslant& \mathbb{E}\left(\rho\left((Z_0^{j,N},\bar{Z}_0^j)_{1\leqslant j\leqslant N}\right)\right) + \xi A\int_0^t e^{cs}ds + (L_X+L_C)\frac{B}{\sqrt{N}}\int_0^t e^{cs}ds\\ \leqslant& \mathbb{E}\left(\rho\left((Z_0^{j,N},\bar{Z}_0^j)_{1\leqslant j\leqslant N}\right)\right) + \xi A\frac{e^{ct}-1}{c} + (L_X+L_C)\frac{B}{\sqrt{N}}\frac{e^{ct}-1}{c}.\end{aligned}$$

We obtain

$$\begin{aligned}\mathbb{E}\left(\rho\left((Z_t^{j,N},\bar{Z}_t^j)_{1\leqslant j\leqslant N}\right)\right) \leqslant \mathbb{E}\left(\rho\left((Z_0^{j,N},\bar{Z}_0^j)_{1\leqslant j\leqslant N}\right)\right)e^{-ct} + \frac{\xi A}{c}\left(1-e^{-ct}\right)\\ + \frac{(L_X+L_C)B}{c}\frac{1}{\sqrt{N}}\left(1-e^{-ct}\right).\end{aligned}$$

By using the exchangeability of the particles, we have $\mathbb{E}\left(\rho\left((Z_t^{j,N},\bar{Z}_t^j)_{1\leqslant j\leqslant N}\right)\right) = \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}f(r_t^i)G_t^i\right) = \mathbb{E}\left(\frac{1}{k}\sum_{i=1}^{k}f(r_t^i)G_t^i\right)$ for all $k \in \mathbb{N}$. Then

$$\mathbb{E}\left(\sum_{i=1}^{k}f(r_t^i)G_t^i\right) = k\mathbb{E}\left(\rho\left((Z_t^{j,N},\bar{Z}_t^j)_{1\leqslant j\leqslant N}\right)\right).$$

We may then conclude on uniform in time propagation of chaos.

Part II

Singular interactions

Chapter 4

Uniform in time propagation of chaos for the 2D vortex model and other singular stochastic systems

Et enfin, monsieur Bidochon qui, lui, est
"spécialiste sur le tas!!"

Binet, *Les Bidochon*, Tome 12
téléspectateurs (1991).

Outline of the current chapter

Joint work with Arnaud Guillin and Pierre Monmarché.

Submitted as: Arnaud Guillin, Pierre Le Bris, and Pierre Monmarché. "Uniform in time propagation of chaos for the 2D vortex model and other singular stochastic systems." *arXiv preprint*

arXiv:2108.08675 (2021). [85]

Abstract: In this chapter, we adapt the work of Jabin and Wang in [99] to show the first result of uniform in time propagation of chaos for a class of singular interaction kernels. In particular, our models contain the Biot-Savart kernel on the torus and thus the 2D vortex model.

Acknowledgements: This work has been (partially) supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency. The authors sincerely thank Didier Bresch, Laurent Chupin and Nicolas Fournier for many discussions around this topic.

4.1 Introduction

4.1.1 Framework

Our main subject is the convergence of the law of a stochastic particles system with mean field singular interactions towards its non linear limit. More precisely we will establish the first quantitative bounds in the number of particles uniformly in time. Let $K : \mathbb{T}^d \to \mathbb{R}^d$ be an *interaction kernel* on the d-dimensional ($d \geqslant 2$) 1-periodic torus $\mathbb{T}^d$ (represented as $[-\frac{1}{2}, \frac{1}{2}]^d$), on which we will specify some assumptions later. In this paper, we consider the non linear stochastic differential equation of *McKean-Vlasov type*

$$\begin{cases} dX_t = \sqrt{2}dB_t + K * \bar{\rho}_t(X_t)dt \\ \bar{\rho}_t = \text{Density of Law}(X_t), \end{cases} \tag{4.1.1}$$

where $X_t \in \mathbb{T}^d$, $(B_t)_{t\geqslant 0}$ is a d-dimensional Brownian motion and $f * g(x) = \int_{\mathbb{T}^d} f(x-y)g(y)dy$ stands for the convolution operation on the torus. The density $\bar{\rho}_t$ satisfies

$$\partial_t \bar{\rho}_t = -\nabla \cdot (\bar{\rho}_t (K * \bar{\rho}_t)) + \Delta \bar{\rho}_t. \tag{4.1.2}$$

In the other words, the non-linear Equation (4.1.2) has the following natural probabilistic interpretation : the solution $\bar{\rho}_t$ is the density of the law at time t of the $\mathbb{T}^d$ valued process $(X_t)_{t\geqslant 0}$ evolving according to (4.1.1). As we understand (4.1.1) to be the motion of a particle interacting with its own law, (4.1.2) thus describes the dynamic of a cloud of charged particles (where $(X_t)_{t\geqslant 0}$ would be one particle). In particular, it holds importance in plasma physics, see [172]. We also consider the associated system of particles, describing the motion of N particles interacting with one another through the interaction kernel K.

$$dX_t^i = \sqrt{2}dB_t^i + \frac{1}{N}\sum_{j=1}^N K(X_t^i - X_t^j)dt, \tag{4.1.3}$$

where $X_t^i \in \mathbb{T}^d$ is the position at time t of the i-th particle, and $(B_t^i, 1 \leqslant i \leqslant N)$ are independent Brownian motions in $\mathbb{T}^d$. We assume that $(X_0^i)_{i=1,..,N}$ are exchangeable, i.e. have a law which is invariant by permutation of the particles, so that this property is true for all times. We denote by ρ_N the density of the law of the system of particles, formally satisfying

$$\partial_t \rho_N = -\sum_{i=1}^N \nabla_{x_i} \cdot \left(\left(\frac{1}{N}\sum_{j=1}^N K(x_i - x_j) \right) \rho_N \right) + \sum_{i=1}^N \Delta_{x_i} \rho_N. \tag{4.1.4}$$

We define ρ_N^k the density of the law of the first k marginals of the N particles system

$$\rho_N^k(t, x_1, .., x_k) = \int_{\mathbb{T}^{(N-k)d}} \rho_N(t, x_1, .., x_N) dx_{k+1} ... dx_N,$$

which is also, thanks to the exchangeability of particles, the density of the law of any k marginals. More precisely, in this work, we focus on the equation (4.1.4) and we will not address the question of the well-posedness of the stochastic equation (4.1.3).

Here, although we will consider general assumptions on K, the main example motivating our work is the singular interaction kernel known as *the Biot-Savart* kernel, defined in $\mathbb{R}^2$ by

$$K(x) = \frac{1}{2\pi} \frac{x^\perp}{|x|^2} = \frac{1}{2\pi} \left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2} \right). \tag{4.1.5}$$

Consider the 2D incompressible Navier-Stokes system on $x \in \mathbb{R}^2$

$$\begin{aligned} \partial_t u =& - u \cdot \nabla u - \nabla p + \Delta u \\ \nabla \cdot u =& 0, \end{aligned}$$

where p is the local pressure. Taking the curl of the equation above, we get that $\omega(t,x) = \nabla \times u(t,x)$ satisfies (4.1.2) with K given by (4.1.5) (see for instance Chapter 1 of [134]).

One can see equation (4.1.3) as an approximation of equation (4.1.1), where the law $\bar{\rho}_t$ is replaced by the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$. It is well known, at least in a setting where the interaction kernel K is Lipschitz continuous ([137], [162]), that, under some mild conditions on K, for all fixed $k \in \mathbb{N}$ and all $t \geqslant 0$, $\rho_N^k(t,\cdot)$ converges toward $\bar{\rho}_k(t,\cdot) = \bar{\rho}_t^{\otimes k}$ as N tends to infinity, where $\bar{\rho}_t$ is the density of the law of X_t solution of (4.1.1). Thus, provided the particles start independent, they will stay (more or less) independent, as the law of any k-uplet of particles converges toward a tensorized law. The expression *propagation of chaos* to describe this behavior was coined by Kac [106]), and we refer to Sznitman [162] for a landmark study of the phenomenon. Of course there is a huge literature on propagation of chaos however limited for uniform in time results, and always when the interaction potential is regular, see Malrieu [133] for an example by a coupling approach under convexity conditions and the recent Durmus &*al* [64] via reflection coupling allowing non convexity but where the interaction is considered small and acts mainly as a perturbation. For more recent results we refer to [111] (and its uniform in time extension in [113]) for a nice new approach for propagation of chaos furnishing better speed but strong assumptions on the interactions (regularity, integrability), including a nice survey of the existing results, and [57] using Lions derivatives for uniform in time results on the torus but also under regularity assumptions on the interaction kernel.

Hence, both these classical and recent results do not apply to the Biot-Savart kernel, which is singular at 0. For a convergence without rate, and specific to the vortex 2D equation, a first striking result appeared in [76], relying on proving that close encounters of particles are rare and that the possible limits of the particles system are made of solutions of the nonlinear SDE. As a second step, in the recent work [99], Jabin and Wang have proven that propagation of chaos still holds in this case with a *quantitative* rate. The goal of the present paper is to extend their works and show a *quantitative propagation of chaos uniform in time.* We refer to [**BJW19b**, 76, 99, 32, 31] for detailed discussions on the literature concerning propagation of chaos with singular kernels, which is still at its beginning for quantitative rates. Shortly after this work was submitted, an alternative approach to global in time estimates was developed in [152], see also the very recent preprint [52].

Obtaining uniform in time estimates for propagation of chaos is an important challenge to tackle. One of its applications for instance concerns the use of particle system, which can easily be simulated numerically, to approximate the solution of a nonlinear physics motivated problem, such as here the vorticity equation arising from fluid mechanics. Likewise, it provides a framework for studying noisy gradient descent used in Machine Learning (see the recent [49]) and thus attracts some attention.

The approach of Jabin and Wang [99] is to compute the time evolution of the relative entropy of ρ_N with respect to $\bar{\rho}_N$ and then to use an integration by parts to deal with the singularity of K thanks to the regularity of the probability density $\bar{\rho}_t$. In order to improve this argument to get uniform in time propagation of chaos, our main contribution is the proof of time-uniform bounds for $\bar{\rho}_t$, in Lemma 4.2.1, from which a time-uniform logarithmic Sobolev inequality is deduced. From the latter, in the spirit of the work of Malrieu [133] in the smooth and convex case, the Fisher information appearing in the entropy dissipation yields a control on the relative entropy itself, inducing the time uniformity. However a major difficulty is that this quantities are expressed in terms of the solution of the nonlinear equation. We then have to prove a logarithmic Sobolev inequality, uniformly in time, for $\bar{\rho}_t$, and a sufficient decay of the derivatives of $\bar{\rho}_t$. To do so, it requires new estimates on regularity and a priori bounds of the solutions of non linear 2D vortex equation. Indeed, we prove that the bounds on the derivative of $\bar{\rho}_t$ decay sufficiently fast (see again Lemma 4.2.1) to ensure uniform in time convergence without smallness assumption on the interaction. Finally, the remaining error term in the entropy evolution due to the difference between (4.1.2) and (4.1.4) is tackled thanks to a law of large number already used in [99]. Compared to [76] we thus obtain a quantitative and uniform in time result.

The organization of the article is as follows. For the remaining of this section, we state the main theorem as well as the various assumptions on both the initial condition and the interaction kernel K. In Section 4.2 we gather various tools that will be useful later on: we state the regularity of the solutions, the existence of uniform in time bounds on the density and its derivatives, and we prove a logarithmic Sobolev inequality. Finally, in Section 4.3, we prove the uniform in time propagation of chaos following the method described in [99].

4.1.2 Main results

First, let us describe the assumptions made on the initial condition. Unless otherwise specified, L^p and H^p respectively refer to the spaces $L^p(\mathbb{T}^d)$ and $H^p(\mathbb{T}^d)$. Given $\lambda > 1$, we denote by $\mathcal{C}^\infty_\lambda(\mathcal{X})$ the set of functions f in $\mathcal{C}^\infty(\mathcal{X})$ such that $0 < \frac{1}{\lambda} \leqslant f \leqslant \lambda < \infty$, and $\mathcal{C}^\infty_{>0}(\mathcal{X}) = \cup_{\lambda>1}\mathcal{C}^\infty_\lambda(\mathcal{X})$, which is simply the set of positive smooth functions when $\mathcal{X}$ is compact.

Proposition 3. *We make the following assumptions on* $\bar{\rho}_0$ *:*

- *There is* $\lambda > 1$ *such that* $\bar{\rho}_0 \in \mathcal{C}^\infty_\lambda(\mathbb{T}^d)$
- *For all* $n \geqslant 1$, $C^0_n := \|\nabla^n \bar{\rho}_0\|_{L^\infty} < \infty$

Remark 4.1.1. *Let us discuss the smoothness assumption on the initial condition. Via Theorem 4.2.1 below, which follows from the result of [13], this will ensure the smoothness of* $\bar{\rho}_t$*. This fact (and the fact that we consider, as we will see later, a smooth solution* ρ_N *of* (4.1.4)*) allows us to justify all calculations in comfortable way. This could however be improved. First, as in [99], the calculations should hold for any entropy solution of* (4.1.4)*. Second, it is also shown in [13], in the case of the vorticity equation, that an initial condition in* L^1 *yields existence and uniqueness of a solution of* (4.1.2) *which is smooth for positive times. One could thus think of using the non-uniform in time result of [99] on a small time interval* $[0, \epsilon]$*, and then complete the proof on* $[\epsilon, \infty[$ *with our result. We would then require some bounds on* $\bar{\rho}_\epsilon$ *and its derivatives* *of a sufficient order (depending on the Sobolev embedding, see the proof of Lemma 4.2.1 below) that we could propagate in time.*

For the sake of clarity and conciseness however we choose not to insist in this direction.

Let us describe the assumptions on the interaction kernel K. Below, $\nabla\cdot$ stands for the divergence operator.

Proposition 4. *We make the following assumptions on K :*

- $\|K\|_{L^1} < \infty$.
- *In the sense of distributions,* $\nabla \cdot K = 0$,
- *There is a matrix field $V \in L^\infty$ such that $K = \nabla \cdot V$, i.e for $1 \leqslant \alpha \leqslant d$, $K_\alpha = \sum_{\beta=1}^d \partial_\beta V_{\alpha,\beta}$.*

The problem of finding a matrix field $V \in L^\infty(\mathbb{T}^d)$ such that $K = \nabla \cdot V$ for a given K is a complex mathematical question. We refer to [28] and [148] and the references therein for a more detailed discussion on the literature. As it was noted in Proposition 2 of [99], the existence of such a matrix V is true for any kernel $K \in L^d$ (using the results of [28]), or for any kernel K such that $\exists M > 0, \forall x \in \mathbb{T}^d, |K(x)| \leqslant M/|x|$ (using the results of [148]).

Remark 4.1.2. *If a function a satisfies $\nabla\cdot a = 0$, then for $\psi : \mathbb{T}^d \mapsto \mathbb{R}$ we have $\nabla\cdot(a\psi) = (a\cdot\nabla)\psi$*

Suppose $\tilde{K}$ is an interaction kernel in $\mathbb{R}^d$ (such as the Biot-Savart kernel). It is possible to periodize $\tilde{K}$ on the torus as follows. For f a function on the torus (identified as a 1-periodic function on $\mathbb{R}^d$), writing $f *_{\mathcal{X}} g(x) = \int_{\mathcal{X}} f(x-y)g(y)dy$ the convolution operator on a space $\mathcal{X}$,

$$\begin{aligned}\tilde{K} *_{\mathbb{R}^d} f(x) = \int_{\mathbb{R}^d} \tilde{K}(x-y)f(y)dy &= \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{T}^d} \tilde{K}(x-y+k)f(y-k)dy \\ &= \int_{\mathbb{T}^d} \left(\sum_{k\in\mathbb{Z}^d} \tilde{K}(x-y+k) \right) f(y)dy,\end{aligned}$$

and thus $\tilde{K} *_{\mathbb{R}^d} f(x) = K *_{\mathbb{T}^d} f(x)$, where $K(x) = \sum_{k\in\mathbb{Z}^d} \tilde{K}(x+k)$. In particular, the periodized Biot-Savart kernel obtained by taking $\tilde{K}$ from (4.1.5) reads

$$K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} + \frac{1}{2\pi} \sum_{k\in\mathbb{Z}^2, k\neq 0} \frac{(x-k)^\perp}{|x-k|^2} := \tilde{K}(x) + K_0(x). \tag{4.1.6}$$

It has been shown that the sum defining K_0 converges (in the sense that $K_0(x) = \lim_{N\to\infty} \sum_{|k|^2\leqslant N, k\neq 0} \frac{(x-k)^\perp}{|x-k|^2}$) in $\mathcal{C}^\infty$ (see for instance [155]). It is straightforward to check that K is periodic, bounded in L^1, and divergence free. Finally, Proposition 2 of [99] yields the existence of $V \in L^\infty$ such that $K = \nabla \cdot V$. As a consequence, Assumption 4 holds in the case of the periodized Biot-Savart kernel.

Remark 4.1.3. *Notice that, for the Biot-Savart kernel on the whole space $\mathbb{R}^2$*

$$\tilde{K}(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2},$$

the matrix field $\tilde{V}$ such that $\tilde{K} = \nabla \cdot \tilde{V}$ can be chosen explicitly

$$V(x) = \frac{1}{2\pi}\begin{pmatrix} -\arctan\left(\frac{x_1}{x_2}\right) & 0 \\ 0 & \arctan\left(\frac{x_2}{x_1}\right) \end{pmatrix}.$$

One could also consider collision-like interactions, as mentioned in [99]. Let $\phi \in L^1$ be a function on the torus, M be a smooth antisymmetric matrix field and consider the kernel $K = \nabla \cdot (M\mathbb{1}_{\phi(x)\leqslant 0})$. By construction, K is the divergence of a L^∞ matrix field, and since M is antisymmetric K is divergence free.

Example 4.1.1. *Consider in dimension 2 the function $\phi : x \mapsto |x|^2 - (2R)^2$ for a given radius $R > 0$ and the matrix*

$$M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which yield

$$K(x) = 2x^\perp \delta_{\phi(x)=0}.$$

This interaction kernel models particles, seen as balls of radius R, interacting via some form of collision.

The well-posedness of the equations (4.1.2) and (4.1.4) under Assumptions 3 and 4 will be discussed respectively in Sections 4.2.1 and 4.3.5. In particular we will see in Theorem 4.2.1 that $\bar{\rho}_t$ is in $\mathcal{C}^\infty_\lambda(\mathbb{R}^+ \times \mathbb{T}^d)$.

The comparison between the law of the system of N interacting particles and the law of N independent particles satisfying the non-linear equation (4.1.1) is stated in terms of relative entropy.

Definition 4.1.1. *Let μ and ν be two probability densities on $\mathbb{T}^{dN}$. We consider the rescaled relative entropy*

$$\mathcal{H}_N(\nu, \mu) = \begin{cases} \frac{1}{N}\mathbb{E}_\mu\left(\frac{\nu}{\mu}\log\frac{\nu}{\mu}\right) & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.1.7}$$

For the sake of conciseness, for all $k \in \mathbb{N}$ and $t \geqslant 0$, we denote $\rho_N(t) : \mathbf{x} \in \mathbb{T}^{dN} \mapsto \rho_N(t, \mathbf{x})$ and $\bar{\rho}_N(t) : \mathbf{x} \in \mathbb{T}^{dN} \mapsto \bar{\rho}_t^{\otimes N}(\mathbf{x})$. The main result is the following

Theorem 4.1.1. *Under Assumptions 3 and 4, there are constants C_1, C_2 and C_3 such that for all $N \in \mathbb{N}$ and all exchangeable density probability $\rho_N(0) \in \mathcal{C}^\infty_{>0}(\mathbb{T}^{dN})$ there exists a weak solution ρ_N of (4.1.4) such that for all $t \geqslant 0$*

$$\mathcal{H}_N(\rho_N(t), \bar{\rho}_N(t)) \leqslant C_1 e^{-C_2 t}\mathcal{H}_N(\rho_N(0), \bar{\rho}_N(0)) + \frac{C_3}{N} \tag{4.1.8}$$

In particular, if $\rho_N(0) = \bar{\rho}_N(0)$, the first term of the right-hand side vanishes, and this property has been called entropic propagation of chaos, see for example [91].

4.1.3 Strong propagation of chaos

We show that Theorem 4.1.1 yields strong propagation of chaos, uniform in time. For μ and ν two probability measures on $\mathbb{T}^{dk}$, denote by $\Pi(\mu, \nu)$ the set of couplings of μ and ν, i.e. the set of probability measures Γ on $\mathbb{T}^{dk} \times \mathbb{T}^{dk}$ with $\Gamma(A \times \mathbb{T}^{dk}) = \mu(A)$ and $\Gamma(\mathbb{T}^{dk} \times A) = \nu(A)$ for all

Borel set A of $\mathbb{T}^{dk}$. Let us define the usual L^2-Wasserstein distance by

$$\mathcal{W}_2(\mu,\nu) = \left(\inf_{\Gamma\in\Pi(\mu,\nu)} \int_{\mathbb{T}^{dk}} d_{\mathbb{T}^{dk}}(x,y)^2 \Gamma(dxdy) \right)^{1/2},$$

where $d_{\mathbb{T}^{dk}}$ is the usual distance on the torus. For $\mathbf{x} = (x_i)_{i\in[\![1,N]\!]} \in \mathbb{T}^{dN}$, we write $\pi(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ the associated empirical measure.

Corollary 4.1.1. *Under assumptions 3 and 4, assuming moreover that $\rho_N(0) = \bar\rho_N(0)$, there is a constant C such that for all $k \leqslant N \in \mathbb{N}$ and all $t \geqslant 0$,*

$$\|\rho_N^k(t) - \bar\rho_k(t)\|_{L^1} + \mathcal{W}_2\left(\rho_N^k(t), \bar\rho_k(t)\right) \leqslant C\left(\left\lfloor \frac{N}{k} \right\rfloor\right)^{-\frac{1}{2}}$$

and

$$\mathbb{E}_{\rho_N(t)}\left(\mathcal{W}_2(\pi(\boldsymbol{X}), \bar\rho_t)\right) \leqslant C\alpha(N)$$

where $\alpha(N) = N^{-1/2}\ln(1+N)$ if $d=2$ and $\alpha(N) = N^{-1/d}$ if $d>2$.

As shown in [22], the last result yields confidence interval in uniform norm when estimating $\bar\rho_t$ with $\pi(\mathbf{X}_t^N)$ convoluted to a smooth kernel.

We postpone the proof as it will rely on results shown later. It will however be a direct corollary of Theorem 4.1.1 and of the logarithmic Sobolev inequality proven in Corollary 4.2.1, which is a crucial ingredient in the proof of Theorem 4.1.1.

4.2 Preliminary work

4.2.1 First results on the non-linear PDE

We have the following result concerning the solution of (4.1.2).

Theorem 4.2.1. *Under Assumption 4, let $\mu_0 \in \mathcal{C}_\lambda^\infty(\mathbb{T}^d)$. Then the system*

$$\begin{cases} \partial_t\bar\rho_t = -\nabla\cdot\left((K*\bar\rho_t)\,\bar\rho_t\right) + \Delta\bar\rho_t, \text{ in } \mathbb{R}^+\times\mathbb{T}^d \\ \bar\rho_0 = \mu_0, \end{cases} \tag{4.2.1}$$

has, in the class of bounded solutions, a unique solution $\bar\rho(t,x) \in \mathcal{C}_\lambda^\infty(\mathbb{R}^+\times\mathbb{T}^d)$.

Proof. The first part of the theorem (existence, uniqueness and smoothness) can be proven by following closely the proof done by Ben-Artzi in [13]. For the sake of completeness, this is detailed in Appendix C.1. Note that a similar result has also been recently proven in [176], where the $\mathcal{C}^k$ regularity of $\bar\rho_t$ for any given k and any given t is shown. The proof relies heavily on the fact that the kernel K is divergence free, that the convolution operation tends to keep the regularity of the most regular term, and that the Fokker-Planck equation has a smoothing effect.

Let us now prove the second part of the result, namely the time uniform bounds on $\bar\rho_t$. Assume that $\mu_0 \in \mathcal{C}_\lambda^\infty(\mathbb{T}^d)$, which by definition implies $\frac{1}{\lambda} \leqslant \mu_0 \leqslant \lambda$, and consider $\bar\rho_t$ the unique solution of (4.2.1). We start by proving that $K*\bar\rho_t$ is in $\mathcal{C}^\infty$. By definition

$$K*\bar\rho_t(x) = \int_{\mathbb{T}^d} K(x-y)\bar\rho_t(y)dy = -\int_{\mathbb{T}^d} K(y)\bar\rho_t(x-y)dy.$$

Then

$$K * \bar{\rho}_t(x) = -\int_{\mathbb{T}^d} \nabla \cdot V(y)\bar{\rho}_t(x-y)dy = -\int_{\mathbb{T}^d} V(y)\nabla_y \bar{\rho}_t(x-y)dy.$$

Since $V \in L^\infty(\mathbb{T}^d)$ and $\bar{\rho} \in \mathcal{C}^\infty(\mathbb{R}^+ \times \mathbb{T}^d)$, we easily deduce that $K*\bar{\rho}$, as well as all its derivatives, are Lipschitz continuous on $[0,T] \times \mathbb{T}^d$ for all $T > 0$. Hence $K * \bar{\rho}$ is $\mathcal{C}^\infty$. Moreover, using that $\nabla \cdot K = 0$ (in the sense of distribution), we immediately get that $\nabla \cdot (K * \bar{\rho}_t) = 0$ for all $t \geqslant 0$.

For $t \geqslant 0$ and $x \in \mathbb{T}^d$, consider Z_s the strong solution of the following stochastic differential equation for $s \in [0,t]$

$$dZ_s = \sqrt{2}dB_s - K * \bar{\rho}_{t-s}(Z_s)ds, \ Z_0 = x$$

which exists, is unique and non-explosive since $K * \bar{\rho}_{t-s}$ is smooth and bounded. Then

$$\bar{\rho}(t,x) = \mathbb{E}_x\left(\bar{\rho}_0(Z_t)\right).$$

The bounds on $\bar{\rho}_t$ follow. □

4.2.2 Higher order estimates

We have already established that $\bar{\rho}_t$ is bounded uniformly in time. In this section, we extend this result to all its derivatives.

Lemma 4.2.1. *For all $n \geqslant 1$ and $\alpha_1, ..., \alpha_n \in [\![1,d]\!]$, there exist $C_n^u, C_n^\infty > 0$ such that for all $t \geqslant 0$,*

$$\|\partial_{\alpha_1,\dots,\alpha_n}\bar{\rho}_t\|_{L^\infty} \leqslant C_n^u \quad \text{and} \quad \int_0^t \|\partial_{\alpha_1,\dots,\alpha_n}\bar{\rho}_s\|_{L^\infty}^2 ds \leqslant C_n^\infty$$

Proof. Thanks to Morrey's inequality and Sobolev embeddings, it is sufficient to prove such bounds in the Sobolev space H^m for all m, in other words it is sufficient to prove similar bounds for $\|\partial_{\alpha_1,\dots,\alpha_n}\bar{\rho}_s\|_{L^2}^2$ for all multi-indexes α. The proof is by induction on the order of the derivatives, we only detail the first iterations. We write $f = \nabla \cdot ((K * \bar{\rho}_t)\,\bar{\rho}_t) = (K * \bar{\rho}_t) \cdot \nabla \bar{\rho}_t$.

Integrated bound for $\|\nabla \bar{\rho}_t\|_{L^2}^2$. We have

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{T}^d} |\bar{\rho}_t|^2 = \int_{\mathbb{T}^d} \bar{\rho}_t \partial_t \bar{\rho}_t = \int_{\mathbb{T}^d} \bar{\rho}_t \Delta \bar{\rho}_t - \int_{\mathbb{T}^d} \bar{\rho}_t f.$$

On the one hand,

$$\int_{\mathbb{T}^d} \bar{\rho}_t \Delta \bar{\rho}_t = -\int_{\mathbb{T}^d} |\nabla \bar{\rho}_t|^2.$$

On the other hand,

$$\int_{\mathbb{T}^d} \bar{\rho}_t f = \int_{\mathbb{T}^d} \bar{\rho}_t \nabla \cdot ((K * \bar{\rho}_t)\,\bar{\rho}_t) = -\int_{\mathbb{T}^d} \nabla \bar{\rho}_t \cdot (K * \bar{\rho}_t)\,\bar{\rho}_t = -\int_{\mathbb{T}^d} \bar{\rho}_t f = 0.$$

Hence,

$$\frac{1}{2}\frac{d}{dt}\|\bar{\rho}_t\|_{L^2}^2 + \|\nabla \bar{\rho}_t\|_{L^2}^2 = 0.$$

By integrating the equality above, we get $\int_0^t \|\nabla \bar{\rho}_t\|_{L^2}^2 = \frac{\|\bar{\rho}_0\|_{L^2}^2 - \|\bar{\rho}_t\|_{L^2}^2}{2} \leqslant \frac{\lambda^2}{2} = C_1^\infty$.

Integrated bound for $\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t\|_{L^2}^2$ and uniform bound for $\|\nabla\bar{\rho}_t\|_{L^2}^2$. Similarly, we calculate

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{T}^d}|\partial_{\alpha_1}\bar{\rho}_t|^2 = \int_{\mathbb{T}^d}\partial_{\alpha_1}\bar{\rho}_t\partial_{\alpha_1}(\partial_t\bar{\rho}_t) = \int_{\mathbb{T}^d}\partial_{\alpha_1}\bar{\rho}_t\partial_{\alpha_1}\left(\Delta\bar{\rho}_t - f\right)$$
$$= -\sum_{\alpha_2}\int_{\mathbb{T}^d}|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t|^2 + \int_{\mathbb{T}^d}\partial_{\alpha_1,\alpha_1}\bar{\rho}_t f.$$

Bounding

$$\int_{\mathbb{T}^d}\partial_{\alpha_1,\alpha_1}\bar{\rho}_t f \leqslant \|\partial_{\alpha_1,\alpha_1}\bar{\rho}_t\|_{L^2}\|f\|_{L^2} \leqslant \frac{1}{2}\sum_{\alpha_2}\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t\|_{L^2}^2 + \frac{1}{2}\|f\|_{L^2}^2,$$

and

$$\|f\|_{L^2}^2 = \int_{\mathbb{T}^d}\left|\sum_{\gamma=1}^d\left(K_\gamma * \bar{\rho}_t\right)\partial_\gamma\bar{\rho}_t\right|^2 \leqslant \|K * \bar{\rho}_t\|_{L^\infty}^2\|\nabla\bar{\rho}_t\|_{L^2}^2 \leqslant \|K\|_{L^1}^2\|\bar{\rho}_t\|_{L^\infty}^2\|\nabla\bar{\rho}_t\|_{L^2}^2,$$

where we used Young's convolution inequality, we get

$$\frac{1}{2}\frac{d}{dt}\|\partial_{\alpha_1}\bar{\rho}_t\|_{L^2}^2 + \frac{1}{2}\sum_{\alpha_2}\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t\|_{L^2}^2 \leqslant \frac{1}{2}\|K\|_{L^1}^2\|\bar{\rho}_t\|_{L^\infty}^2\|\nabla\bar{\rho}_t\|_{L^2}^2.$$

By integrating the equality above and using Theorem 4.2.1, we get

$$\frac{\|\partial_{\alpha_1}\bar{\rho}_t\|_{L^2}^2 - \|\partial_{\alpha_1}\bar{\rho}_0\|_{L^2}^2}{2} + \frac{1}{2}\int_0^t\sum_{\alpha_2}\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_s\|_{L^2}^2 ds \leqslant \frac{1}{2}\|K\|_{L^1}^2\lambda^2\int_0^t\|\nabla\bar{\rho}_s\|_{L^2}^2 ds$$
$$\leqslant \frac{1}{2}\|K\|_{L^1}^2\lambda^2 C_1^\infty.$$

This provides both the existence of C_2^∞ such that for all $t \geqslant 0$, $\int_0^t\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_s\|_{L^2}^2 ds \leqslant C_2^\infty$, and the existence of C_1^u such that for all $t \geqslant 0$, $\|\partial_{\alpha_1}\bar{\rho}_t\|_{L^2}^2 \leqslant C_1^u$.

Integrated bound bound for $\|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_t\|_{L^2}^2$ and uniform bound for $\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t\|_{L^2}^2$. We have

$$\partial_\alpha f = \sum_\gamma(\partial_\alpha K_\gamma * \bar{\rho}_t)\partial_\gamma\bar{\rho}_t + \sum_\gamma(K_\gamma * \bar{\rho}_t)\partial_{\alpha,\gamma}\bar{\rho}_t,$$

and

$$\partial_\alpha K_\gamma * \bar{\rho}_t = \int_{\mathbb{T}^d}\partial_\alpha K_\gamma(x-y)\bar{\rho}_t(y)dy = -\int_{\mathbb{T}^d}\partial_\alpha K_\gamma(y)\bar{\rho}_t(x-y)dy = -\int_{\mathbb{T}^d}K_\gamma(y)\partial_\alpha\bar{\rho}_t(x-y)dy$$
$$= -\sum_\beta\int_{\mathbb{T}^d}V_{\gamma,\beta}(y)\partial_{\alpha,\beta}\bar{\rho}_t(x-y)dy = \sum_\beta V_{\gamma,\beta} * \partial_{\alpha,\beta}\bar{\rho}_t.$$

Hence

$$\sum_\gamma\left(\partial_\alpha K_\gamma * \bar{\rho}_t\right)\partial_\gamma\bar{\rho}_t = \sum_\gamma\left(\sum_\beta V_{\gamma,\beta} * \partial_{\alpha,\beta}\bar{\rho}_t\right)\partial_\gamma\bar{\rho}_t$$

$$= \left(V * \partial_\alpha \nabla \bar{\rho}_t\right) \nabla \bar{\rho}_t,$$

and thus

$$\| \sum_\gamma \left(\partial_\alpha K_\gamma * \bar{\rho}_t\right) \partial_\gamma \bar{\rho}_t \|_{L^2} \leqslant \|V * \partial_\alpha \nabla \bar{\rho}_t\|_{L^\infty} \|\nabla \bar{\rho}_t\|_{L^2} \leqslant \|V\|_{L^\infty} \|\partial_\alpha \nabla \bar{\rho}_t\|_{L^1} \|\nabla \bar{\rho}_t\|_{L^2}.$$

Therefore

$$\|\partial_\alpha f\|_{L^2}^2 \leqslant 2\|V\|_{L^\infty}^2 \|\partial_\alpha \nabla \bar{\rho}_t\|_{L^1}^2 \|\nabla \bar{\rho}_t\|_{L^2}^2 + 2\|K\|_{L^1}^2 \|\bar{\rho}_t\|_{L^\infty}^2 \|\partial_\alpha \nabla \bar{\rho}_t\|_{L^2}^2.$$

Similarly to previous computations,

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\int_{\mathbb{T}^d} |\partial_{\alpha_1,\alpha_2}\bar{\rho}_t|^2 &= \int_{\mathbb{T}^d} \partial_{\alpha_1,\alpha_2}\bar{\rho}_t \partial_{\alpha_1,\alpha_2}\left(\Delta \bar{\rho}_t - f\right) \\
&= -\sum_{\alpha_3} \int_{\mathbb{T}^d} |\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_t|^2 + \int_{\mathbb{T}^d} \partial_{\alpha_1,\alpha_2,\alpha_2}\bar{\rho}_t \partial_{\alpha_1} f \\
&\leqslant -\sum_{\alpha_3} \|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_t\|_{L^2}^2 + \|\partial_{\alpha_1,\alpha_2,\alpha_2}\bar{\rho}_t\|_{L^2} \|\partial_{\alpha_1} f\|_{L^2} \\
&\leqslant -\sum_{\alpha_3} \|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_t\|_{L^2}^2 + \frac{1}{2}\sum_{\alpha_3} \|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_t\|_{L^2}^2 \\
&\quad + \|V\|_{L^\infty}^2 \|\partial_{\alpha_1}\nabla \bar{\rho}_t\|_{L^1}^2 \|\nabla \bar{\rho}_t\|_{L^2}^2 + \|K\|_{L^1}^2 \|\bar{\rho}_t\|_{L^\infty}^2 \|\partial_{\alpha_1}\nabla \bar{\rho}_t\|_{L^2}^2 \\
&\leqslant -\frac{1}{2}\sum_{\alpha_3} \|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_t\|_{L^2}^2 + \|V\|_{L^\infty}^2 \|\nabla \bar{\rho}_t\|_{L^2}^2 \|\partial_{\alpha_1}\nabla \bar{\rho}_t\|_{L^2}^2 \\
&\qquad + \|K\|_{L^1}^2 \|\bar{\rho}_t\|_{L^\infty}^2 \|\partial_{\alpha_1}\nabla \bar{\rho}_t\|_{L^2}^2,
\end{aligned}$$

and thus

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t\|_{L^2}^2 + \frac{1}{2}\sum_{\alpha_3} \|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_t\|_{L^2}^2 \leqslant& \|V\|_{L^\infty}^2 \|\partial_{\alpha_1}\nabla \bar{\rho}_t\|_{L^2}^2 \|\nabla \bar{\rho}_t\|_{L^2}^2 \\
&+ \|K\|_{L^1}^2 \|\bar{\rho}_t\|_{L^\infty}^2 \|\partial_{\alpha_1}\nabla \bar{\rho}_t\|_{L^2}^2.
\end{aligned}$$

Integrating over time, and using Theorem 4.2.1

$$\begin{aligned}
\frac{\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t\|_{L^2}^2 - \|\partial_{\alpha_1,\alpha_2}\bar{\rho}_0\|_{L^2}^2}{2} + \frac{1}{2}\sum_{\alpha_3}\int_0^t \|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_s\|_{L^2}^2 ds& \\
\leqslant \|V\|_{L^\infty}^2 d C_1^u \int_0^t \|\partial_{\alpha_1}\nabla \bar{\rho}_s\|_{L^2}^2 ds + \|K\|_{L^1}^2 \lambda^2 \int_0^t \|\partial_{\alpha_1}\nabla \bar{\rho}_s\|_{L^2}^2 ds& \\
\leqslant d\left(d\|V\|_{L^\infty}^2 C_1^u + \|K\|_{L^1}^2 \lambda^2\right) C_2^\infty&
\end{aligned}$$

This provides both the existence of C_3^∞ such that for all $t \geqslant 0$, $\int_0^t \|\partial_{\alpha_1,\alpha_2,\alpha_3}\bar{\rho}_s\|_{L^2}^2 ds \leqslant C_3^\infty$, and the existence of C_2^u such that for all $t \geqslant 0$, $\|\partial_{\alpha_1,\alpha_2}\bar{\rho}_t\|_{L^2}^2 \leqslant C_2^u$.

The proof is then by induction on the order of derivative, iterating the same method. □

4.2.3 Logarithmic Sobolev inequality

We now establish a logarithmic Sobolev inequality (LSI) for $\bar{\rho}_t$ solution of (4.1.2). To this end, we use the fact that the uniform distribution u on $\mathbb{T}^d$ satisfies a LSI and that $\bar{\rho}_t$ is bounded (below and above) uniformly in time. Recall the following Holley-Stroock perturbation lemma, from [5, Prop. 5.1.6].

Lemma 4.2.2. *Assume that ν is a probability measure on $\mathbb{T}^d$ satisfying a logarithmic Sobolev inequality with constant C_ν^{LS}, i.e for all $f \in \mathcal{C}^\infty_{>0}(\mathbb{T}^d)$,*

$$\text{Ent}_\nu(f) := \int_{\mathbb{T}^d} f \log f d\nu - \int_{\mathbb{T}^d} f d\nu \log\left(\int_{\mathbb{T}^d} f d\nu\right) \leqslant C_\nu^{LS} \int_{\mathbb{T}^d} \frac{|\nabla f|^2}{f} d\nu.$$

Let μ be a probability measure with density h with respect to ν such that, for some constant $\lambda > 0$, $\frac{1}{\lambda} \leqslant h \leqslant \lambda$. Then μ satisfies a logarithmic Sobolev inequality with constant $C_\mu^{LS} = \lambda^2 C_\nu^{LS}$, i.e for all $f \in \mathcal{C}^\infty_{>0}(\mathbb{T}^d)$

$$\text{Ent}_\mu(f) \leqslant \lambda^2 C_\nu^{LS} \int_{\mathbb{T}^d} \frac{|\nabla f|^2}{f} d\mu.$$

We also know that the uniform distribution u (i.e the Lebesgue measure) on $\mathbb{T}^d$ satisfies a LSI. See for instance Proposition 5.7.5 of [5], or [79] for a proof in dimension 1, the results in higher dimension being a consequence of tensorization properties.

Lemma 4.2.3. *Let u be the uniform distribution on $\mathbb{T}^d$. Then u satisfies a logarithmic Sobolev inequality : for all $f \in \mathcal{C}^\infty_{>0}(\mathbb{T}^d)$*

$$\text{Ent}_u(f) \leqslant \frac{1}{8\pi^2} \int_{\mathbb{T}^d} \frac{|\nabla f|^2}{f} du \tag{4.2.2}$$

A direct consequence of Lemma 4.2.2, Lemma 4.2.3 and the bounds on $\bar{\rho}_t$ given in Theorem 4.2.1 is the following theorem, as well as its corollary. It establishes a uniform in time logarithmic Sobolev inequality for $\bar{\rho}_t$, crucial for the uniform control of the Fisher information appearing in the study of the dissipation of the entropy between the law of the particles system and the nonlinear ones.

Theorem 4.2.2. *Under Assumptions 3 and 4, for all $t \geqslant 0$ and all function $f \in \mathcal{C}^\infty_{>0}(\mathbb{T}^d)$,*

$$\text{Ent}_{\bar{\rho}_t}(f) \leqslant \frac{\lambda^2}{8\pi^2} \int_{\mathbb{T}^d} \frac{|\nabla f|^2}{f} d\bar{\rho}_t$$

Corollary 4.2.1. *Under Assumptions 3 and 4, for all $N \in \mathbb{N}$, $t \geqslant 0$ and all probability density $\mu_N \in \mathcal{C}^\infty_{>0}(\mathbb{T}^{dN})$,*

$$\mathcal{H}_N\left(\mu_N, \bar{\rho}_N(t)\right) \leqslant \frac{\lambda^2}{8\pi^2} \frac{1}{N} \sum_{i=1}^N \int_{\mathbb{T}^d} \mu_N \left|\nabla_{x_i} \log \frac{\mu_N}{\bar{\rho}_N(t)}\right|^2$$

Proof. By tensorization of the logarithmic Sobolev inequality (see for instance Proposition 5.2.7 of [5]), since $\bar{\rho}$ satisfies a LSI with constant $\frac{\lambda^2}{8\pi^2}$, so does $\bar{\rho}_N$. Using Theorem 4.2.2 for $f = \frac{\mu_N}{\bar{\rho}_N}$ we thus get

$$\mathcal{H}_N(\mu_N, \bar{\rho}_N(t)) = \frac{1}{N} \text{Ent}_{\bar{\rho}_N(t)}\left(\frac{\mu_N}{\bar{\rho}_N(t)}\right) \leqslant \frac{\lambda^2}{8\pi^2} \frac{1}{N} \mathbb{E}_{\bar{\rho}_N(t)}\left(\left|\nabla_x \frac{\mu_N}{\bar{\rho}_N(t)}\right|^2 \frac{\bar{\rho}_N(t)}{\mu_N}\right).$$

Hence the result. □

4.3 Proofs of the main results

From now on and up to Section 4.3.5 (excluded), in addition to Assumptions 3 and 4, we suppose that there exists a solution $\rho_N \in \mathcal{C}^\infty_{>0}(\mathbb{R}^+ \times \mathbb{T}^{dN})$ of (4.1.4). This justifies the validity of the various calculations conducted in this part of the proof. The question to lift this assumption (by taking a limit in a regularized problem) is addressed in Section 4.3.5

4.3.1 Time evolution of the relative entropy

We write

$$\mathcal{H}_N(t) = \mathcal{H}_N(\rho_N(t), \bar{\rho}_N(t)), \qquad \mathcal{I}_N(t) = \frac{1}{N}\sum_i \int_{\mathbb{T}^{dN}} \rho_N(t) \left|\nabla_{x_i} \log \frac{\rho_N(t)}{\bar{\rho}_N(t)}\right|^2 d\mathbf{x}.$$

as short hands for the relative entropy and relative Fisher information. We start by calculating the time evolution of the relative entropy.

Lemma 4.3.1. *For all $t \geqslant 0$,*

$$\frac{d}{dt}\mathcal{H}_N(t) \leqslant A_N(t) + \frac{1}{2}B_N(t) - \frac{1}{2}\mathcal{I}_N(t), \tag{4.3.1}$$

with

$$A_N(t) := \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_N \left(V(x_i - x_j) - V * \bar{\rho}_t(x_i)\right) : \frac{\nabla^2_{x_i}\bar{\rho}_N}{\bar{\rho}_N} d\boldsymbol{x}$$

$$B_N(t) := \frac{1}{N}\sum_i \int_{\mathbb{T}^{dN}} \rho_N \frac{|\nabla_{x_i}\bar{\rho}_N|^2}{\bar{\rho}_N^2} \left|\frac{1}{N}\sum_j V(x_i - x_j) - V * \bar{\rho}_t(x_i)\right|_f^2 d\boldsymbol{x}.$$

Here, $|\cdot|_f^2$ denotes the sum of the square of the coefficients of the matrix.

Proof. It has been shown, in [99], that

$$\frac{d}{dt}\mathcal{H}_N(t) \leqslant -\mathcal{I}_N(t) - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_N \left(K(x_i - x_j) - K * \bar{\rho}_t(x_i)\right) \cdot \nabla_{x_i} \log \bar{\rho}_N d\mathbf{x},$$

with

$$\begin{aligned}
-\frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} & \rho_N \left(K(x_i - x_j) - K * \bar{\rho}_t(x_i)\right) \cdot \nabla_{x_i} \log \bar{\rho}_N d\mathbf{x} \\
&= \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_N \left(V(x_i - x_j) - V * \bar{\rho}_t(x_i)\right) : \frac{\nabla^2_{x_i}\bar{\rho}_N}{\bar{\rho}_N} d\mathbf{x} \\
&\quad + \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \left(V(x_i - x_j) - V * \bar{\rho}_t(x_i)\right) : \nabla_{x_i}\bar{\rho}_N \otimes \nabla_{x_i} \frac{\rho_N}{\bar{\rho}_N} d\mathbf{x}.
\end{aligned}$$

Let us consider the latter term

$$\frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \left(V(x_i - x_j) - V * \bar{\rho}_t(x_i)\right) : \nabla_{x_i}\bar{\rho}_N \otimes \nabla_{x_i} \frac{\rho_N}{\bar{\rho}_N} d\mathbf{x}$$

$$= \frac{1}{N}\sum_i \sum_{\alpha,\beta} \int_{\mathbb{T}^{dN}} \left(\frac{1}{N}\sum_j V(x_i - x_j) - V * \bar{\rho}_t(x_i) \right)_{\alpha,\beta} (\nabla_{x_i}\bar{\rho}_N)_\alpha \left(\nabla_{x_i} \frac{\rho_N}{\bar{\rho_N}} \right)_\beta d\mathbf{x}.$$

Let

$$y^i_\beta := \left(\nabla_{x_i} \frac{\rho_N}{\bar{\rho_N}} \right)_\beta \frac{\bar{\rho}_N}{\sqrt{\rho_N}}, \quad z^i_\alpha := (\nabla_{x_i}\bar{\rho}_N)_\alpha \frac{\sqrt{\rho_N}}{\bar{\rho}_N}, \quad \text{and} \quad x^i_{\alpha,\beta} := \left(\frac{1}{N}\sum_j V(x_i - x_j) - V * \bar{\rho}(x_i) \right)_{\alpha,\beta},$$

then, using $xy \leqslant \frac{x^2}{2} + \frac{y^2}{2}$ for all $x, y \in \mathbb{R}$,

$$\sum_{\alpha,\beta} x^i_{\alpha,\beta} z^i_\alpha y^i_\beta = \sum_\beta y^i_\beta \left(\sum_\alpha x^i_{\alpha,\beta} z^i_\alpha \right) \leqslant \frac{1}{2}\sum_\beta (y^i_\beta)^2 + \frac{1}{2}\sum_\beta \left(\sum_\alpha x^i_{\alpha,\beta} z^i_\alpha \right)^2,$$

and thus, using the Cauchy-Schwarz inequality,

$$\begin{aligned} \sum_{\alpha,\beta} x^i_{\alpha,\beta} z^i_\alpha y^i_\beta \leqslant & \frac{1}{2}\sum_\beta (y^i_\beta)^2 + \frac{1}{2}\sum_\beta \left(\sum_\alpha (x^i_{\alpha,\beta})^2 \right) \left(\sum_\alpha (z^i_\alpha)^2 \right) \\ = & \frac{1}{2}\sum_\beta (y^i_\beta)^2 + \frac{1}{2} \left(\sum_\alpha (z^i_\alpha)^2 \right) \left(\sum_{\alpha,\beta} (x^i_{\alpha,\beta})^2 \right). \end{aligned}$$

Hence

$$\begin{aligned} & \frac{1}{N^2}\sum_{i,j} \int_{\mathbb{T}^{dN}} (V(x_i - x_j) - V * \bar{\rho}_t(x_i)) : \nabla_{x_i}\bar{\rho}_N \otimes \nabla_{x_i} \frac{\rho_N}{\bar{\rho_N}} d\mathbf{x} \\ \leqslant & \frac{1}{2N}\sum_i \int_{\mathbb{T}^{dN}} \frac{\bar{\rho}_N^2}{\rho_N} \left| \nabla_{x_i} \frac{\rho_N}{\bar{\rho_N}} \right|^2 d\mathbf{x} + \frac{1}{2N}\sum_i \int_{\mathbb{T}^{dN}} \rho_N \frac{|\nabla_{x_i}\bar{\rho}_N|^2}{\bar{\rho}_N^2} \left| \frac{1}{N}\sum_j V(x_i - x_j) - V * \bar{\rho}_t(x_i) \right|_f^2 d\mathbf{x} \\ = & \frac{1}{2}\mathcal{I}_N(t) + \frac{1}{2N}\sum_i \int_{\mathbb{T}^{dN}} \rho_N \frac{|\nabla_{x_i}\bar{\rho}_N|^2}{\bar{\rho}_N^2} \left| \frac{1}{N}\sum_j V(x_i - x_j) - V * \bar{\rho}_t(x_i) \right|_f^2 d\mathbf{x}. \end{aligned}$$

This yields the desired result. □

4.3.2 Change of reference measure and Law of Large Number

We now state three general results which will be useful in order to control the error terms A_N and B_N defined in Lemma 4.3.1. The first one will be used to perform a change of measure from ρ_N to $\bar{\rho}_N$.

Lemma 4.3.2. *Let $N \in \mathbb{N}$. For two probability densities μ and ν on $\mathbb{T}^{dN}$, and any $\Phi \in L^\infty(\mathbb{T}^{dN})$ and $\eta > 0$,*

$$\mathbb{E}^\mu \Phi \leqslant \eta \mathcal{H}_N(\mu, \nu) + \frac{\eta}{N} \log \mathbb{E}^\nu e^{N\Phi/\eta}.$$

Proof. Define

$$f = \frac{1}{\theta} e^{N\Phi/\eta} \nu, \qquad \theta = \int_{\mathbb{T}^{dN}} e^{N\Phi/\eta} \nu d\mathbf{x}.$$

Notice f is a probability density. By convexity of the entropy

$$\frac{1}{N}\int_{\mathbb{T}^{dN}} \mu \log f d\mathbf{x} \leqslant \frac{1}{N}\int_{\mathbb{T}^{dN}} \mu \log \mu d\mathbf{x}.$$

On the other hand

$$\frac{1}{N}\int_{\mathbb{T}^{dN}} \mu \log f d\mathbf{x} = \frac{1}{\eta}\int_{\mathbb{T}^{dN}} \mu \Phi d\mathbf{x} + \frac{1}{N}\int_{\mathbb{T}^{dN}} \mu \log \nu d\mathbf{x} - \frac{\log \theta}{N}.$$

□

The next two statements are crucial theorems of [99].

Theorem 4.3.1. *[Theorem 3 of [99]] Consider any probability measure μ on $\mathbb{T}^d$ and a scalar function $\psi \in L^\infty(\mathbb{T}^d \times \mathbb{T}^d)$ with $\|\psi\|_{L^\infty} < \frac{1}{2e}$ and such that for all $z \in \mathbb{T}^d$, $\int_{\mathbb{T}^d} \psi(z,x)\mu(dx) = 0$. Then*

$$\int_{\mathbb{T}^{dN}} \exp\Big(\frac{1}{N}\sum_{j_1,j_2=1}^{N} \psi(x_1,x_{j_1})\psi(x_1,x_{j_2})\Big)\mu^{\otimes N} d\boldsymbol{x} \leqslant C = 2\left(1 + \frac{10\alpha}{(1-\alpha)^3} + \frac{\beta}{1-\beta}\right), \quad (4.3.2)$$

where

$$\alpha = (e\|\psi\|_{L^\infty})^4 < 1\ ,\ \ \beta = \left(\sqrt{2e}\|\psi\|_{L^\infty}\right)^4 < 1.$$

The second one is a nice improvement of the usual level two large deviations bound for i.i.d. random variables.

Theorem 4.3.2. *[Theorem 4 of [99]] Consider any probability measure μ on $\mathbb{T}^d$ and $\phi \in L^\infty(\mathbb{T}^d \times \mathbb{T}^d)$ with*

$$\gamma := \left(1600^2 + 36e^4\right)\left(\sup_{p\geqslant 1} \frac{\|\sup_z |\phi(\cdot,z)|\|_{L^p(\mu))}}{p}\right)^2 < 1. \quad (4.3.3)$$

Assume that ϕ satisfies the following cancellations

$$\forall z \in \mathbb{T}^d, \quad \int_{\mathbb{T}^d} \phi(x,z)\mu(dx) = 0 = \int_{\mathbb{T}^d} \phi(z,x)\mu(dx)\,.$$

Then, for all $N \in \mathbb{N}$,

$$\int_{\mathbb{T}^{dN}} \exp\Big(\frac{1}{N}\sum_{i,j=1}^{N} \phi(x_i,x_j)\Big)\mu^{\otimes N} d\boldsymbol{x} \leqslant \frac{2}{1-\gamma} < \infty. \quad (4.3.4)$$

4.3.3 Bounding the error terms

Lemma 4.3.3. *The terms A_N and B_N introduced in Lemma 4.3.1 are such that*

$$A_N(t) + \frac{1}{2}B_N(t) \leqslant C\left(\mathcal{H}_N(t) + \frac{1}{N}\right)$$

with

$$C = \hat{C}_1 \lambda d \|\nabla^2 \bar{\rho}_t\|_{L^\infty}\|V\|_{L^\infty} + \hat{C}_2 \lambda^2 d^2 \|V\|_{L^\infty}^2 \|\nabla \bar{\rho}_t\|_{L^\infty}^2$$

where $\hat{C}_1, \hat{C}_2$ are universal constants.

Proof. Recall from Theorem 4.2.1 that $\bar{\rho}_t \in \mathcal{C}^\infty_\lambda(\mathbb{T}^d)$ for all $t \geqslant 0$. We first bound B_N. For $(X^i_t)_i$ given in (4.1.3), we have

$$\begin{aligned}
B_N =& \frac{1}{N}\sum_i \int_{\mathbb{T}^{dN}} \rho_N \frac{|\nabla\bar{\rho}_t|^2}{\bar{\rho}_t{}^2}(x_i)\left|\frac{1}{N}\sum_j V(x_i-x_j) - V*\bar{\rho}_t(x_i)\right|_f^2 d\mathbf{x}\\
=& \frac{1}{N}\sum_i \mathbb{E}\left(\left|\frac{\nabla\bar{\rho}_t}{\bar{\rho}_t}(X^i_t)\right|^2\left|\frac{1}{N}\sum_j V(X^i_t-X^j_t) - V*\bar{\rho}_t(X^i_t)\right|_f^2\right)\\
=& \frac{1}{N}\sum_i \sum_{\alpha,\beta=1}^d \mathbb{E}\left(\left|\frac{\nabla\bar{\rho}_t}{\bar{\rho}_t}(X^i_t)\right|^2\left(\frac{1}{N}\sum_j V_{\alpha,\beta}(X^i_t-X^j_t) - V_{\alpha,\beta}*\bar{\rho}_t(X^i_t)\right)^2\right)\\
\leqslant& \frac{\lambda^2\|\nabla\bar{\rho}_t\|^2_{L^\infty}}{N}\sum_i \sum_{\alpha,\beta=1}^d \mathbb{E}\left(\left(\frac{1}{N}\sum_j V_{\alpha,\beta}(X^i_t-X^j_t) - V_{\alpha,\beta}*\bar{\rho}_t(X^i_t)\right)^2\right).
\end{aligned}$$

We apply Lemma 4.3.2 to each

$$\Phi_{\alpha,\beta} = \left(\frac{1}{N}\sum_j V_{\alpha,\beta}(x_i-x_j) - V_{\alpha,\beta}*\bar{\rho}_t(x_i)\right)^2,$$

to get, for all $C_B > 0$,

$$\begin{aligned}
&\mathbb{E}\left(\left(\frac{1}{N}\sum_j V_{\alpha,\beta}(X^i_t-X^j_t) - V_{\alpha,\beta}*\bar{\rho}_t(X^i_t)\right)^2\right)\\
&\leqslant C_B\mathcal{H}_N(t) + \frac{C_B}{N}\log\mathbb{E}\left(\exp\left(\frac{1}{C_B}\left(\frac{1}{\sqrt{N}}\sum_j V_{\alpha,\beta}(\bar{X}^i_t-\bar{X}^j_t) - V_{\alpha,\beta}*\bar{\rho}_t(\bar{X}^i_t)\right)^2\right)\right).
\end{aligned}$$

This way,

$$\begin{aligned}
B_N \leqslant& \frac{C_B\lambda^2\|\nabla\bar{\rho}_t\|^2_{L^\infty}}{N^2}\sum_i\sum_{\alpha,\beta}\log\int_{\mathbb{T}^{dN}}\bar{\rho}_N\exp\left(\frac{1}{C_B}\left(\frac{1}{\sqrt{N}}\sum_j V_{\alpha,\beta}(x_i-x_j) - V_{\alpha,\beta}*\bar{\rho}_t(x_i)\right)^2\right)d\mathbf{x}\\
&+ C_B d^2\lambda^2\|\nabla\bar{\rho}_t\|^2_{L^\infty}\mathcal{H}_N(t).
\end{aligned}$$

In the following we choose $C_B = 64e^2\|V\|^2_{L^\infty}$. Applying Theorem 4.3.1 to

$$\psi(z,x) = \frac{1}{8e\|V\|_{L^\infty}}\left(V(z-x) - V*\bar{\rho}_t(z)\right),$$

which satisfies $\|\psi\|_{L^\infty} \leqslant \frac{1}{4e}$ and is such that

$$\int_{\mathbb{T}^d} \psi(z,x)\bar{\rho}_t(x)dx = \frac{1}{8e\|V\|_{L^\infty}} \int_{\mathbb{T}^d} V(z-x)\bar{\rho}_t(x)dx - \frac{1}{8e\|V\|_{L^\infty}} \int_{\mathbb{T}^d} V * \bar{\rho}_t(z)\bar{\rho}_t(x)dx = 0,$$

we get

$$B_N \leqslant \hat{C}_B \|V\|_{L^\infty}^2 \lambda^2 d^2 \|\nabla \bar{\rho}_t\|_{L^\infty}^2 \left(\mathcal{H}_N(t) + \frac{\tilde{C}_B}{N}\right), \tag{4.3.5}$$

where $\hat{C}_B$ and $\tilde{C}_B$ are universal constants.

We now proceed with the bound on A_N. Applying Lemma 4.3.2 to

$$\Phi = \frac{1}{N^2} \sum_{i,j} \left(V(x_i - x_j) - V * \bar{\rho}_t(x_i)\right) : \frac{\nabla^2_{x_i} \bar{\rho}_N}{\bar{\rho}_N},$$

we obtain, for all $C_A > 0$,

$$A_N \leqslant \frac{C_A}{N} \log \int_{\mathbb{T}^{dN}} \bar{\rho}_N \exp\left(\frac{1}{C_A N} \sum_{i,j} \left(V(x_i - x_j) - V * \bar{\rho}_t(x_i)\right) : \frac{\nabla^2_{x_i} \bar{\rho}_N}{\bar{\rho}_N}\right) d\mathbf{x} + C_A \mathcal{H}_N(t)$$

In the following we choose

$$C_A = 4\sqrt{1600^2 + 36e^4} \|\nabla^2 \bar{\rho}_t\|_{L^\infty} \|V\|_{L^\infty} \lambda d := \hat{C}_A \lambda d \|\nabla^2 \bar{\rho}_t\|_{L^\infty} \|V\|_{L^\infty}.$$

Then, we apply Theorem 4.3.2 to

$$\phi(z,x) = \frac{1}{C_A}\left(\left(V(z-x) - V * \bar{\rho}_t(z)\right) : \frac{\nabla^2 \bar{\rho}_t}{\bar{\rho}_t}(z)\right),$$

which satisfies, thanks to Assumption 4

$$\begin{aligned}\int_{\mathbb{T}^d} \phi(z,x)\bar{\rho}_t(z)dz =& \frac{1}{C_A} \int_{\mathbb{T}^d} \left(\left(V(z-x) - V * \bar{\rho}_t(z)\right) : \frac{\nabla^2 \bar{\rho}_t}{\bar{\rho}_t}(z)\right) \bar{\rho}_t(z)dz \\ =& \frac{1}{C_A} \int_{\mathbb{T}^d} \left(\mathrm{div}K(z-x) - \mathrm{div}K * \bar{\rho}_t(z)\right) \bar{\rho}_t(z)dz = 0,\end{aligned}$$

and, thanks to $\int_{\mathbb{T}^d} \left(V(z-x) - V * \bar{\rho}_t(z)\right) \bar{\rho}_t(x)dx = 0$,

$$\int_{\mathbb{T}^d} \phi(z,x)\bar{\rho}_t(x)dx = 0.$$

Through our choice of C_A, (4.3.3) is verified, as $\gamma \leqslant \left(1600^2 + 36e^4\right)\left(\frac{2d\|V\|_{L^\infty}\|\nabla^2\bar{\rho}_t\|_{L^\infty}\lambda}{C_A}\right)^2 = \frac{1}{4} < 1$. Hence

$$A_N \leqslant \hat{C}_A \|\nabla^2 \bar{\rho}_t\|_{L^\infty} \|V\|_{L^\infty} \lambda d \left(\mathcal{H}_N(t) + \frac{\tilde{C}_A}{N}\right), \tag{4.3.6}$$

where $\hat{C}_A$ and $\tilde{C}_A$ are universal constants. The conclusion easily follows.

□

4.3.4 Proof of Theorem 4.1.1 in the smooth case

It only remains to gather the previous results. Equations (4.3.1), (4.3.5) and (4.3.6) yield

$$\frac{d}{dt}\mathcal{H}_N(t) \leqslant \left(\hat{C}_A \lambda d \|\nabla^2 \bar{\rho}_t\|_{L^\infty} \|V\|_{L^\infty} + \frac{\hat{C}_B \|V\|_{L^\infty}^2 \lambda^2 \|\nabla \bar{\rho}_t\|_{L^\infty}^2 d^2}{2}\right) \mathcal{H}_N(t) + \frac{C_2}{N} - \frac{1}{2}\mathcal{I}_N(t),$$

and using Corollary 4.2.1 and $\hat{C}_A \|\nabla^2 \bar{\rho}_t\|_{L^\infty} \|V\|_{L^\infty} \lambda d \leqslant \frac{1}{2}\left(\frac{2\pi}{\lambda}\right)^2 + \frac{1}{2}\left(\frac{\lambda}{2\pi}\right)^2 \hat{C}_A^2 \|\nabla^2 \bar{\rho}_t\|_{L^\infty}^2 \|V\|_{L^\infty}^2 \lambda^2 d^2$

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_N(t) &\leqslant -\left(\left(\frac{2\pi}{\lambda}\right)^2 - \hat{C}_A \lambda d \|\nabla^2 \bar{\rho}_t\|_{L^\infty} \|V\|_{L^\infty} - \frac{\hat{C}_B \|V\|_{L^\infty}^2 \lambda^2 \|\nabla \bar{\rho}_t\|_{L^\infty}^2 d^2}{2}\right) \mathcal{H}_N(t) + \frac{C_2}{N} \\ &\leqslant -\frac{1}{2}\left(\left(\frac{2\pi}{\lambda}\right)^2 - \hat{C}_A^2 \frac{\lambda^4}{4\pi^2} d^2 \|\nabla^2 \bar{\rho}_t\|_{L^\infty}^2 \|V\|_{L^\infty}^2 - \hat{C}_B \|V\|_{L^\infty}^2 \|\nabla \bar{\rho}_t\|_{L^\infty}^2 \lambda^2 d^2\right) \mathcal{H}_N(t) + \frac{C_2}{N}.\end{aligned}$$

In a more concise way, using Lemma 4.2.1, it means there are constants $C_1, C_2^\infty, C_3 > 0$ and a function $t \mapsto C_2(t) > 0$ with $\int_0^t C_2(s)ds \leqslant C_2^\infty$ for all $t \geqslant 0$ such that for all $t \geqslant 0$

$$\frac{d}{dt}\mathcal{H}_N(t) \leqslant -(C_1 - C_2(t))\mathcal{H}_N(t) + \frac{C_3}{N}.$$

Multiplying by $\exp(C_1 t - \int_0^t C_2(s)ds)$ and integrating in time we get

$$\begin{aligned}\mathcal{H}_N(t) &\leqslant e^{-C_1 t + \int_0^t C_2(s)ds}\mathcal{H}_N(0) + \frac{C_3}{N}\int_0^t e^{C_1(s-t) + \int_s^t C_2(u)du} ds \\ &\leqslant e^{C_2^\infty - C_1 t}\mathcal{H}_N(0) + \frac{C_3}{C_1 N} e^{C_2^\infty},\end{aligned}$$

which concludes.

4.3.5 Dealing with the regularity of ρ_N

As mentioned at the beginning of Section 4.3, up to now we have proven the result under the additional assumption that there exists a smooth solution ρ_N to (4.1.4). Let us now remove this assumption. Consider $(\zeta_\epsilon)_{\epsilon \geqslant 0}$ a sequence of mollifiers such that $\|\zeta_\epsilon\|_{L^1} = 1$, whose compact support are assumed to be strictly contained within $\left[-\frac{1}{2}, \frac{1}{2}\right]^d$. Let us consider $K^\epsilon = K * \zeta_\epsilon$. We have $K^\epsilon \in \mathcal{C}^\infty(\mathbb{T}^d)$ and $\mathrm{div}(K^\epsilon) = 0$.

Let ρ_N^ϵ be the unique smooth solution (see Lemma 8 below) of the parabolic equation with smooth coefficients

$$\partial_t \rho_N^\epsilon + \frac{1}{N}\sum_{i,j=1}^N K^\epsilon(x_i - x_j) \cdot \nabla_{x_i}\rho_N^\epsilon = \sum_{i=1}^N \Delta_{x_i}\rho_N^\epsilon, \tag{4.3.7}$$

with initial condition $\rho_N^\epsilon(0,\cdot) = \rho_N(0,\cdot)$..

We have the following bounds

Lemma 4.3.4. *Let $\gamma > 1$ be such that $\rho_N(0) \in \mathcal{C}_\gamma^\infty(\mathbb{T}^{dN})$. Then, for all $t \geqslant 0$ and all $\epsilon > 0$, $\rho_N^\epsilon(t) \in \mathcal{C}_\gamma^\infty(\mathbb{T}^{dN})$.*

Proof. Let $\mathbf{x} \in \mathbb{T}^{dN}$ Consider the particle system $dX_i^\epsilon(t) = -\frac{1}{N}\sum_{j=1}^N K^\epsilon(X_i^\epsilon(t) - X_j^\epsilon(t))dt + \sqrt{2}dB_t^i$ with initial condition $\mathbf{X}_0^\epsilon = \mathbf{x}$, where we denote $\mathbf{X}_t^\epsilon = (X_1^\epsilon(t), \dots, X_N^\epsilon(t))$. We have strong existence and uniqueness for this SDE. Then

$$\rho_N^\epsilon(t, \mathbf{x}) = \mathbb{E}\left(\rho_N^\epsilon(0, \mathbf{X}_t^\epsilon)\right).$$

The bounds on ρ_N^ϵ follow. □

Using Lemma 4.3.4, we get $(\rho_N^\epsilon)_\epsilon$ is a sequence of smooth functions uniformly bounded in $L^\infty(\mathbb{R}^+ \times \mathbb{T}^{Nd})$. This yields two results.

First, we can extract a weakly-* converging subsequence in $L^\infty(\mathbb{R}^+ \times \mathbb{T}^{Nd})$, i.e there exists $\rho_N \in L^\infty(\mathbb{R}^+ \times \mathbb{T}^{Nd})$ such that for all $f \in L^1(\mathbb{R}^+ \times \mathbb{T}^{Nd})$ we have

$$\int_{\mathbb{T}^{Nd}} \rho_N^\epsilon f \underset{\epsilon \to 0^+}{\longrightarrow} \int_{\mathbb{T}^{Nd}} \rho_N f.$$

We finally check that ρ_N is indeed a weak solution of (4.1.4). For all $T \geqslant 0$ and for all f smooth test function on $[0,T] \times \mathbb{T}^{Nd}$

- We have, since $\partial_t f$ is smooth and therefore in $L^1([0,T] \times \mathbb{T}^{Nd})$

$$\int_{\mathbb{T}^{Nd}} \rho_N^\epsilon \partial_t f \to \int_{\mathbb{T}^{Nd}} \rho_N \partial_t f.$$

- Likewise, since $\Delta_{x_i} f$ is smooth and therefore in $L^1([0,T] \times \mathbb{T}^{Nd})$

$$\int_{\mathbb{T}^{Nd}} \rho_N^\epsilon \Delta_{x_i} f \to \int_{\mathbb{T}^{Nd}} \rho_N \Delta_{x_i} f.$$

- Finally

$$\begin{aligned}
&\int_{\mathbb{T}^{Nd}} \rho_N^\epsilon K^\epsilon(x_i - x_j) \cdot \nabla_{x_i} f - \int_{\mathbb{T}^{Nd}} \rho_N K(x_i - x_j) \cdot \nabla_{x_i} f \\
&\qquad = \int_{\mathbb{T}^{Nd}} \rho_N^\epsilon \left(K^\epsilon(x_i - x_j) - K(x_i - x_j)\right) \cdot \nabla_{x_i} f + \int_{\mathbb{T}^{Nd}} \left(\rho_N^\epsilon - \rho_N\right) K(x_i - x_j) \cdot \nabla_{x_i} f \\
&\qquad \leqslant \|\rho_N^\epsilon\|_{L^\infty} \|\nabla_{x_i} f\|_{L^\infty} \|K^\epsilon - K\|_{L^1} + \int_{\mathbb{T}^{Nd}} \left(\rho_N^\epsilon - \rho_N\right) K(x_i - x_j) \cdot \nabla_{x_i} f \\
&\qquad \to 0,
\end{aligned}$$

  as $\|K^\epsilon - K\|_{L^1} \to 0$ and $K(x_i - x_j) \cdot \nabla_{x_i} f \in L^1([0,T] \times \mathbb{T}^{Nd})$.

We have thus proven that ρ_N is a weak solution of (4.1.4).

Likewise, we may consider $(\bar{\rho}^\epsilon)_\epsilon$, which weakly-* converges to a solution which, by uniqueness, is $\bar{\rho}$.

Second, ρ_N^ϵ satisfies the assumption made at the beginning of Section 4.3, i.e $\rho_N^\epsilon \in \mathcal{C}^\infty_{>0}(\mathbb{R}^+ \times \mathbb{T}^d)$. Since by considering $V^\epsilon = V * \zeta_\epsilon$ we have $K^\epsilon = \operatorname{div}(V^\epsilon)$, we get that K^ϵ satisfies Assumption 4 and that the calculations done in Section 4.3 are valid for this specific kernel, i.e

$$\mathcal{H}_N(\rho_N^\epsilon(t), \bar{\rho}_N^\epsilon(t)) \leqslant \mathcal{H}_N(\rho_N(0), \bar{\rho}_N(0)) e^{-C_1^\epsilon t} e^{C^{\infty,\epsilon}} + \frac{C_3^\epsilon e^{C^{\infty,\epsilon}}}{C_1^\epsilon} \frac{1}{N}. \tag{4.3.8}$$

Notice how, in the proof of Lemma 4.2.1, the constants bounding the various derivatives of $\bar{\rho}$ only depend on the initial conditions, on $\|K\|_{L^1}$ and on $\|V\|_{L^\infty}$. Since $(\zeta_\epsilon)_{\epsilon\geqslant 0}$ is a sequence of mollifiers, we have $\|K^\epsilon\|_{L^1} \to \|K\|_{L^1}$ as ϵ tends to 0, and $\|V^\epsilon\|_{L^\infty} \leqslant \|V\|_{L^\infty}$. The righthand side of (4.3.8) can thus be chosen independent of ϵ.

We now use the fact that for $u \geqslant 0$ and $v \in \mathbb{R}$ we have $uv \leqslant u\log u - u + e^v$, to obtain the variational formulation of the entropy,

$$N\mathcal{H}_N(\rho_N^\epsilon(t), \bar{\rho}_N^\epsilon(t)) = \sup\left\{\mathbb{E}_{\rho_N^\epsilon(t)}(g) - \mathbb{E}_{\bar{\rho}_N^\epsilon(t)}\left(e^g\right) + 1, g \in L^\infty\right\}, \tag{4.3.9}$$

the equality being attained for $g = \log\left(\frac{\rho_N^\epsilon}{\bar{\rho}_N^\epsilon}\right)$. We thus consider, for $g \in L^\infty$,

$$\frac{1}{N}\left(\mathbb{E}_{\rho_N^\epsilon(t)}(g) - \mathbb{E}_{\bar{\rho}_N^\epsilon(t)}\left(e^g\right) + 1\right) \leqslant \mathcal{H}_N(\rho_N(0), \bar{\rho}_N(0))e^{-C_1 t}e^{C^\infty} + \frac{C_3 e^{C^\infty}}{1+C_1}\frac{1}{N}.$$

By definition of the weak-* convergence in L^∞ (since both g and e^g are thus in L^1), we have the following convergence $\mathbb{E}_{\rho_N^\epsilon(t)}(g) \longrightarrow \mathbb{E}_{\rho_N(t)}(g)$ and $\mathbb{E}_{\bar{\rho}_N^\epsilon(t)}\left(e^g\right) \longrightarrow \mathbb{E}_{\bar{\rho}_N(t)}\left(e^g\right)$ as ϵ tends to 0. Therefore, for all $g \in L^\infty$,

$$\frac{1}{N}\left(\mathbb{E}_{\rho_N(t)}(g) - \mathbb{E}_{\bar{\rho}_N(t)}\left(e^g\right) + 1\right) \leqslant \mathcal{H}_N(\rho_N(0), \bar{\rho}_N(0))e^{-C_1 t}e^{C^\infty} + \frac{C_3 e^{C^\infty}}{1+C_1}\frac{1}{N},$$

which yields Theorem 4.1.1, using (4.3.9) for $\mathcal{H}_N(\rho_N(t), \bar{\rho}_N(t))$.

4.3.6 Proof of Corollary 4.1.1

Let $k \in \mathbb{N}$, and $N \geqslant k$. The sub-additivity of the entropy (see for instance Theorem 10.2.3 of [3]) implies that the (rescaled) relative entropy of the marginals is bounded by the total (rescaled) relative entropy

$$k\left\lfloor\frac{N}{k}\right\rfloor \mathcal{H}_k(\rho_N^k(t), \bar{\rho}_k(t)) \leqslant N\mathcal{H}_N(\rho_N(t), \bar{\rho}_N(t)).$$

The logarithmic Sobolev inequality established in Corollary 4.2.1 implies a Talagrand's transportation inequality (see [145]), so that the L^2-Wasserstein distance is bounded by the relative entropy. Classically, this is also the case of the total variation thanks to Pinsker's inequality, and thus

$$\|\rho_N^k(t) - \bar{\rho}_k(t)\|_{L^1} + \mathcal{W}_2(\rho_N^k(t), \bar{\rho}_k(t)) \leqslant C\sqrt{k\mathcal{H}_k(\rho_N^k(t), \bar{\rho}_k(t))} \leqslant C\sqrt{\frac{N}{\left\lfloor\frac{N}{k}\right\rfloor}\mathcal{H}_N(\rho_N(t), \bar{\rho}_N(t))}.$$

With the additional assumption that $\mathcal{H}_N(\rho_N(0), \bar{\rho}_N(0)) = 0$, we thus get the result using Theorem 4.1.1. To obtain the result on the empirical measure, we recall for the sake of completeness the arguments of [105, Proposition 8]. Given $x, y \in \mathbb{T}^{dN}$, a coupling of $\pi(x)$ and $\pi(y)$ is obtained by considering (x_J, y_J) where J is uniformly distributed over $[\![1, N]\!]$. From this we get $\mathcal{W}_2(\pi(x), \pi(y)) \leqslant |x-y|/\sqrt{N}$. Considering $(\mathbf{X}, \mathbf{Y})$ an optimal coupling of $(\rho_N(t), \bar{\rho}_N(t))$, we bound

$$\begin{aligned}\mathbb{E}\left(\mathcal{W}_2(\pi(\mathbf{X}, \bar{\rho}_t)\right) &\leqslant \mathbb{E}\left(\mathcal{W}_2(\pi(\mathbf{X}), \pi(\mathbf{Y}))\right) + \mathbb{E}\left(\mathcal{W}_2(\pi(\mathbf{Y}), \bar{\rho}_t)\right) \\ &\leqslant \frac{1}{\sqrt{N}}\mathcal{W}_2(\rho_N(t), \bar{\rho}_N(t)) + \mathbb{E}\left(\mathcal{W}_2(\pi(\mathbf{Y}), \bar{\rho}_t)\right).\end{aligned}$$

The last term is tackled with the result for i.i.d. variables established in [75].

Chapter 5

On systems of particles in singular repulsive interaction in dimension one : log and Riesz gas

It is completely unimportant. That is why it is so interesting.

Agatha Christie, *The Murder of Roger Ackroyd* (1926).

Outline of the current chapter

Joint work with Arnaud Guillin and Pierre Monmarché.

Published as: Arnaud Guillin, Pierre Le Bris, and Pierre Monmarché. "On systems of particles in singular repulsive interaction in dimension one: log and Riesz gas." *Journal de l'École*

polytechnique — Mathématiques, Tome 10 (2023), pp. 867-916. [84]
with the addition of: Remark 5.3.2 and Section 5.6.

Abstract: In this chapter, we prove the first quantitative uniform in time propagation of chaos for a class of systems of particles in singular repulsive interaction in dimension one that contains the Dyson Brownian motion. We start by establishing existence and uniqueness for the Riesz gases, before proving propagation of chaos with an original approach to the problem, namely coupling with a Cauchy sequence type argument. We also give a general argument to turn a result of weak propagation of chaos into a strong and uniform in time result using the long time behavior and some bounds on moments, in particular enabling us to get a uniform in time version of the result of Cépa-Lépingle [43].

Acknowledgements: This work has been (partially) supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency.

The authors thank Prof. Martin Hairer for the remarks that led to the addition of Remark 5.3.2 and Section 5.6.

5.1 Introduction

We consider the one dimensional N-particle system in mean field interaction

$$dX_t^i = \sqrt{2\sigma_N} dB_t^i - U'\left(X_t^i\right) dt - \frac{1}{N}\sum_{j\neq i} V'(X_t^i - X_t^j)dt. \tag{5.1.1}$$

where for all $i \in \{1, ..., N\}$, X_t^i denotes the position in $\mathbb{R}$ of the i-th particle, $(B_t^i)_i$ are independent Brownian motions, and σ_N is a diffusion coefficient that may depend on N. We denote $\mathbf{X}_t^N = \left(X_t^1, ..., X_t^N\right)$. We will refer to U as the confining potential and V as the interaction potential, on which we will specify the assumptions later. Finally, we denote by ρ_t^N the law of $(X_t^1, ..., X_t^N)$

The goal of this chapter is to give various results concerning equation (5.1.1) in the case where V is a singular repulsive interaction potential. The main motivating example is the (generalized) Dyson Brownian motion

$$dX_t^i = \sqrt{\frac{2\sigma}{N}} dB_t^i - \lambda X_t^i dt + \frac{1}{N}\sum_{j\neq i} \frac{1}{X_t^i - X_t^j} dt. \tag{5.1.2}$$

Equation 5.1.2 is satisfied, for $\lambda = 0$, by the eigenvalues of an $N \times N$ Hermitian matrix valued Brownian motion, as observed by Dyson in 1962 [65]. For $\lambda > 0$, it corresponds to the eigenvalues of an $N \times N$ Hermitian matrix valued Ornstein-Uhlenbeck process (see for instance [47, 151]).

The work of Wigner [175] is often considered to be the starting point of Random Matrix Theory. The main observation is that, for a Wigner matrix (a symmetric $N \times N$ matrix whose entries above the main diagonal are independent centered variables), the empirical distribution of the eigenvalues converges weakly as $N \to \infty$ to the standard semi-circle distribution. We refer to [2] and references therein for a thorough introduction on Random Matrix Theory.

The main result of this chapter concerns the limit, as N goes to infinity, of (5.1.1), which can be considered as a dynamical version of the convergence of the eigenvalues of a Wigner matrix. What we wish to prove is that *in a system of N particles in mean-field interaction, as N goes to infinity, two particles become more and more statistically independent.* Kac [106] described this behavior as *propagation of chaos*, and we refer to Sznitman [162] for a landmark study of the phenomenon. The notion of *chaos* refers to the independence, while *propagation* alludes to the fact that having this property of independence at the limit at time 0 will be sufficient to ensure the same independence at later time t.

This limit for the Dyson Brownian motion was recently studied in [17] using a notion of *spectral dominance*, and obtained without convergence rate. Let us also mention the work [92] which proves propagation of chaos in 1D in a kinetic setting (i.e each particle is represented by a position and a velocity) for a discontinuous interaction corresponding to the sign of the difference of positions.

Throughout this chapter, we denote by $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$ the empirical measure at time t of the N particle system. As proven by Sznitman [162], the convergence of the empirical measure towards a constant random variable $\bar{\rho}_t$ is equivalent to the property of propagation of chaos. Very formally, this limit $\bar{\rho}_t$ is a weak solution to the non linear equation of *McKean-Vlasov* type

$$\partial_t \bar{\rho}_t = \partial_x \left(\left(U' + V' * \bar{\rho}_t \right) \bar{\rho}_t \right) + \sigma \partial_{xx}^2 \bar{\rho}_t, \tag{5.1.3}$$

where σ is the limit (possibly 0) of σ_N as $N \to \infty$ and $*$ is the (space) convolution operation. The stochastic differential equation associated to (5.1.3) is

$$\begin{cases} dX_t = \sqrt{2\sigma} dB_t - U'(X_t)dt - V' * \bar{\rho}_t(X_t)dt, \\ \bar{\rho}_t = \text{Law}(X_t), \end{cases} \tag{5.1.4}$$

and can also be seen as the formal limit of the stochastic differential equation (SDE) (5.1.1), noticing $\frac{1}{N}\sum_j V'(X_t^i - X_t^j) = V' * \mu_t^N(X_t^i)$. At this stage however, let us insist on the fact that the objects and solutions of (5.1.3) and (5.1.4) can be ill defined, especially when V' is singular.

As we aim at deriving *quantitative* propagation of chaos result, we need a *distance*. For μ and ν two probability measures on $\mathbb{R}$, denote by $\Pi(\mu, \nu)$ the set of couplings of μ and ν, i.e. the set of probability measures Γ on $\mathbb{R} \times \mathbb{R}$ with $\Gamma(A \times \mathbb{R}) = \mu(A)$ and $\Gamma(\mathbb{R} \times A) = \nu(A)$ for all Borel set A of $\mathbb{R}$. We define the L^p Wasserstein distance, with $p \geqslant 1$, as

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{\Gamma \in \Pi(\mu,\nu)} \int |x - y|^p \Gamma(dxdy) \right)^{1/p}.$$

It is important to notice (see for instance Remarks 3.28 and 3.30 of [147]) that in dimension 1 the optimal coupling (i.e the one realizing the infimum) for $\mathcal{W}_p$, $p \geqslant 1$, is known as it is the monotone map. In particular, for two sets of points $(x_i)_{i\in\{1,\dots,N\}}$ and $(y_j)_{j\in\{1,\dots,N\}}$, assuming without loss of generality that $x_1 \leqslant \dots \leqslant x_N$ and $y_1 \leqslant \dots \leqslant y_N$, and two measures $\mu = \frac{1}{N}\sum_i \delta_{x_i}$ and $\nu = \frac{1}{N}\sum_j \delta_{y_j}$, one has

$$\mathcal{W}_p(\mu, \nu)^p = \frac{1}{N}\sum_i |x_i - y_i|^p.$$

There exists many ways of proving propagation of chaos, let us mention some.

- The main probabilistic tool, as used by McKean (see for instance [136]) and then popularised by Sznitman [162], is the coupling method. It consists in coupling the solution of (5.1.1) with N independent copies $(\bar{X}_t^i)_i$ of the solution (5.1.4). The goal is to control the

Wasserstein distance, which by definition can be written as

$$\mathcal{W}_d\left(\rho_t^N, \bar{\rho}_t^{\otimes N}\right) = \inf_{\Gamma \in \Pi\left(\rho_t^N, \bar{\rho}_t^{\otimes N}\right)} \mathbb{E}^{\Gamma}\left(\sum_{i=1}^N d(X_t^i, \bar{X}_t^i)\right).$$

Here, the notation $\mathcal{W}_d$ refers to the fact that the Wasserstein distance depends on an underlying distance d. This is the only time we use this notation, not to be confused with the L^p Wasserstein distance $\mathcal{W}_p$ we use thereafter.

Instead of considering the minimum over all couplings, the key idea is to construct a specific one, which will therefore provide an upper bound on the Wasserstein distance. Well known coupling methods include the *synchronous* coupling [162, 40], or the more recent *reflection* coupling as suggested by Eberle [66, 69, 64]. The main benefit of this method of proof is that it allows for a better probabilistic understanding of the processes and gives quantitative speed of convergence in the case of Lipschitz continuous interactions. However, to the authors' knowledge, coupling methods have not yet given results in the case of singular interactions.

- Using tools from PDE analysis, and functional inequalities, in order to show convergence of ρ_t^N towards $\bar{\rho}_t^{\otimes N}$, recent progress have been made using a modulated energy [160, 159, 152], by considering the relative entropy of ρ_t^N with respect to $\bar{\rho}_t^{\otimes N}$ [99] or by combining these two quantities into a modulated free energy [32]. These quantities have proven useful in showing propagation of chaos for systems of particle in singular interaction by making full use of the regularity and bounds on the moments of the limit equation (5.1.3).
- Another method, that lies somewhere in between the fields of probability and PDE analysis, consists in proving the tightness or compactness of the set of empirical measure, showing that the limit of any convergent subsequence satisfies (5.1.3), and proving the uniqueness of the solution of (5.1.3). This has been for instance done for singular interaction kernels, in the specific case of (5.1.2) [151, 43, 125]. This method, however, does not provide quantitative convergence rates.

Notice that all the methods described above rely on the properties of the limit equation (5.1.3), because one needs to either give sense to the quantity $V' * \mu_t$ in (5.1.4) (and maybe show some properties in order to carry out computations) to use coupling methods, prove bounds and regularity on the solution in order to use PDE related methods, or at the very least prove the uniqueness of the solution of (5.1.3). This study of the limiting equation can be a quite challenging task.

In this chapter, we describe a method that relies only on the well posedness of the system of particles (5.1.1) and which provides a quantitative (and in some cases uniform in time) result of propagation of chaos. We make full use of the fact that in dimension one the particles will stay ordered, and that as a consequence the interaction we consider will be convex (See Remark 5.1.1 below). Using a coupling method, we prove that by taking any independent sequence of empirical measures, it is a Cauchy sequence. Then, independence ensures the fact that the limit measure is an almost surely constant random variable. To the authors' knowledge, such a method has not been used before to prove propagation of chaos.

Let us now introduce our main assumptions. The condition on the interaction potential is the following:

Assumption 5.1. *There exists $\alpha \geqslant 1$ such that*

$$\forall x \in \mathbb{R}^*,\ V'(x) = -\frac{x}{|x|^{\alpha+1}}\,, \tag{5.1.5}$$

and we thus consider

$$V(x) = \begin{cases} \frac{1}{\alpha-1}|x|^{-\alpha+1} & \text{if } \alpha > 1 \\ -\ln(|x|) & \text{if } \alpha = 1 \end{cases} \tag{5.1.6}$$

Notice that for all $x \in \mathbb{R}^*$, $V'(x) = -V'(-x)$, *and* $V''(x) = \frac{\alpha}{|x|^{\alpha+1}}$.

Let us consider the open set

$$\mathcal{O}_N := \left\{ \mathbf{X} = (x_1, ..., x_N) \in \mathbb{R}^N \quad \text{s.t.} \quad -\infty < x_1 < ... < x_N < \infty \right\}.$$

Remark 5.1.1. *We highlight the main geometrical property we will use :*
Denote $H_{int,\alpha} : \mathbb{R}^N \mapsto \mathbb{R}$ *the function given by*

$$\forall \boldsymbol{x} = (x_i)_{i \in \{1,...,N\}}, \quad H_{int,\alpha}(\boldsymbol{x}) = \frac{1}{2N} \sum_{i \neq j} V(x_i - x_j).$$

This way the particle system (5.1.1) *can be rewritten as the following Langevin diffusion*

$$d\boldsymbol{X}_t = \sqrt{2\sigma_N} dB_t - \mathcal{U}(\boldsymbol{X}_t)dt - \nabla H_{int,\alpha}(\boldsymbol{X}_t)dt,$$

where $\boldsymbol{X}_t = (X_t^1, ..., X_t^N) \in \mathbb{R}^N$, B *is a Brownian motion in* $\mathbb{R}^N$, *and* $\mathcal{U} : \mathbb{R}^N \mapsto \mathbb{R}^N$ *is the function given by* $\mathcal{U}(\boldsymbol{x}) = (U(x_i))_{1 \leqslant i \leqslant N}$.

Let $\boldsymbol{X} = (x_1, ..., x_N) \in \mathcal{O}_N$ *and* $\boldsymbol{Y} = (y_1, ..., y_N) \in \mathcal{O}_N$. *We have, since* $x \to V'(x)$ *is odd under Assumption 5.1*

$$\begin{aligned} &(\nabla H_{int,\alpha}(\boldsymbol{X}) - \nabla H_{int,\alpha}(\boldsymbol{Y})) \cdot (\boldsymbol{X} - \boldsymbol{Y}) \\ &= \frac{1}{N} \sum_{1 \leqslant j \neq i \leqslant N} (x_i - y_i) \left(V'(x_i - x_j) - V'(y_i - y_j) \right) \\ &= \frac{1}{N} \sum_{1 \leqslant j < i \leqslant N} \left((x_i - y_i) - (x_j - y_j) \right) \left(V'(x_i - x_j) - V'(y_i - y_j) \right) \\ &= \frac{1}{N} \sum_{1 \leqslant j < i \leqslant N} \left((x_i - x_j) - (y_j - y_j) \right) \left(V'(x_i - x_j) - V'(y_i - y_j) \right). \end{aligned}$$

Then, for $i > j$, *we have* $x_i - x_j > 0$ *and* $y_i - y_j > 0$. *Since the function* V' *given by Assumption 5.1 is an increasing function on* $\mathbb{R}^+$, *each term in the sum above is non-negative. We thus obtain*

$$(\nabla H_{int,\alpha}(\boldsymbol{X}) - \nabla H_{int,\alpha}(\boldsymbol{Y})) \cdot (\boldsymbol{X} - \boldsymbol{Y}) \geqslant 0.$$

The drift term appearing in the particle system seen as a Langevin diffusion in $\mathbb{R}^N$ *is therefore the gradient of a convex function on* $\mathcal{O}_N$. *This property will imply the long time convergence of the particle system (see the proof of Theorem 5.2.2 thereafter), and will be one of the main tools used to prove propagation of chaos (see the proof of Lemma 5.3.1).*

Consider $U \in \mathcal{C}^2(\mathbb{R})$, and make the following assumptions

Assumption 5.2. U' *is Lipschitz continuous, i.e there exists* L_U *such that for all* $x \in \mathbb{R}$ *we have* $|U''(x)| \leqslant L_U$. *This implies*

$$\forall x, y \in \mathbb{R}, \; |U'(x) - U'(y)| \leqslant L_U |x - y|,$$

and

$$\exists A>0,\ \forall x\in\mathbb{R},\ |U'(x)|\leqslant L_U|x|+A.$$

This first set of conditions will be used when establishing existence and uniqueness of solutions of (5.1.1) as well as non uniform in time propagation of chaos. For further results, for simplicity, the study will either be restricted to the convex case, namely:

Assumption 5.3. *U satisfies Assumption 5.2 and we have $U'(0)=0$. Furthermore, there is $\lambda>0$ such that*

$$\forall x,y\in\mathbb{R},\ (U'(x)-U'(y))(x-y)\geqslant\lambda(x-y)^2.$$

...or to the quadratic case :

Assumption 5.4. *There is $\lambda>0$ such that U is explicitly given by*

$$\forall x\in\mathbb{R},\ U(x)=\frac{\lambda}{2}x^2.$$

We choose to use the same notation λ in both Assumption 5.3 and Assumption 5.4 as it serves the same purpose.

Notice Assumption 5.4 is strictly stronger than Assumption 5.3, which itself is strictly stronger than Assumption 5.2.

We focus on the quadratic case as it is the main case of interest in the applications we presented. But, since the main results in the case $\alpha=1$ hold true in the convex case with little to no modification in the calculations, we distinguish this case. The assumption U convex should also be sufficient in the case $\alpha>1$, but requires more involved computations. We describe later in the proofs the key points that should be modified (see for instance Remark D.2.1). Likewise, assuming $U'(0)=0$ in Assumption 5.3 is purely technical.

We sum up the main results of the chapter in the following theorem.

Theorem 5.1.1. A] *Under Assumption 5.1 and 5.2, for $\alpha=1$ and $\sigma_N\leqslant\frac{1}{N}$ or for $\alpha>1$, there exists a unique strong solution to* (5.1.1).

B] *Under Assumptions 5.1 and 5.3, denoting by $\rho_t^{1,N}$ and $\rho_t^{2,N}$ the probability densities on $\mathcal{O}_N$ of the particle systems with respective initial conditions $\rho_0^{1,N}$ and $\rho_0^{2,N}$, we have*

$$\forall t\geqslant 0,\qquad \mathcal{W}_2\left(\rho_t^{1,N},\rho_t^{2,N}\right)\leqslant e^{-\lambda t}\mathcal{W}_2\left(\rho_0^{1,N},\rho_0^{2,N}\right).$$

C] *Under Assumptions 5.1 and 5.3 for $\alpha=1$ or under Assumptions 5.1 and 5.4 for $\alpha\in]1,2[$, let $\mu_t^N:=\frac{1}{N}\sum_{i=1}^N\delta_{X_t^i}$ be the empirical measure at time t of the solution of* (5.1.1). *Assume there exists $\bar\rho_0$ such that $\mathbb{E}\mathcal{W}_2^2(\mu_0^N,\bar\rho_0)\to 0$ as $N\to\infty$. With the additional assumption $\sigma_N\leqslant\frac{1}{N}$ for $\alpha=1$, there exist $(\rho_t)_{t\geqslant 0}\in\mathcal{C}(\mathbb{R}^+,\mathcal{P}_2(\mathbb{R}))$, as well as universal constants $C_1,C_2>0$ and a quantity $C_0^N>0$ that depends on the initial condition and such that $C_0^N\to 0$ as $N\to\infty$, such that for all $N\geqslant 1$ and all $t\geqslant 0$*

$$\mathbb{E}\left(\mathcal{W}_2(\mu_t^N,\bar\rho_t)^2\right)\leqslant e^{-2\lambda t}C_0^N+\frac{C_1}{N^{(2-\alpha)/\alpha}}+C_2\sigma_N,$$

where $(\bar\rho_t)_t$ satisfies, for all functions $f\in\mathcal{C}^2(\mathbb{R})$ with bounded derivatives such that f, f', $f'U'$, and f'' are Lipschitz continuous and that $f'U'$ is bounded, the following equation: $\forall t\geqslant 0$,

$$\int_{\mathbb{R}}f(x)\bar\rho_t(dx)=\int_{\mathbb{R}}f(x)\bar\rho_0(dx)-\int_0^t\int_{\mathbb{R}}f'(x)U'(x)\bar\rho_s(dx)ds$$

$$+\frac{1}{2}\int_0^t\int\int_{\{x\neq y\}}\frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}\bar{\rho}_s(dx)\bar{\rho}_s(dy)ds.$$

Remark that the statement B], as well as functional inequalities such as Poincaré or logarithmic Sobolev inequalities, has been obtained in [44] in the case $\alpha = 1$. Furthermore, statement C] extends the result of [16], in which similar systems are studied using the theory of gradient flows and (non uniform in time) propagation of chaos is obtained for $\alpha < 2$ without convergence rate.

We split Theorem 5.1.1 above into the more precise Theorems 5.2.1, 5.2.2, 5.3.1 and 5.4.1 below. The organization of the chapter is as follows :

- In Section 5.2 we prove various results concerning particle system (5.1.1). In Section 5.2.1, we show that, for $\alpha > 1$ with any diffusion coefficient σ_N or $\alpha = 1$ with $\sigma_N \leqslant \frac{1}{N}$, there exists a unique strong solution to (5.1.1) under Assumptions 5.1 and 5.2. Furthermore, the particles stay in the same order at all time. See Theorem 5.2.1. In Section 5.2.2, we show the long time convergence of the particle system under Assumptions 5.1 and 5.3. See Theorem 5.2.2. In Section 5.2.3, we prove bounds on the expectation of interaction that will be useful later.

- Section 5.3 contains the main proofs of the chapter concerning the propagation of chaos for (5.1.1) in the case $\sigma_N \to 0$. For clarity, we separate the case $\alpha = 1$ (in Section 5.3.1) and $\alpha \in]1, 2[$ (in Section 5.3.2), as the former allows for a proof that contains all the ideas with little technical difficulties, while the latter requires the more precise bounds obtained in Section 5.2.3. See Theorem 5.3.1.

- In Section 5.4, we identify the equation satisfied by the limit $\bar{\rho}_t$. See Theorem 5.4.1. In particular, we highlight an argument which intuitively suggests that $\alpha = 2$ should be the critical case for the well-posedness of the limit.

Finally, in Section 5.5, we show how one can turn a result of weak propagation of chaos, such as the one obtained in [151, 43, 125], into a strong and uniform in time result by using the long time convergence and some bounds on the moments of the particle system. See Theorem 5.5.1. This yields in particular strong uniform in time propagation of chaos in the case of $\alpha = 1$ and constant diffusion coefficient $\sigma_N = \sigma \neq 0$ that Theorem 5.3.1 cannot deal with, though without a quantitative rate of convergence, using the result of [43]. This Section 5.5 is independent of the previous ones. The result of this section is summarized in the next corollary.

Corollary 5.1.1. *Under Assumptions 5.1 and 5.3, for $\alpha = 1$, $\sigma_N = \sigma \in \mathbb{R}$, assume we have for all N an initial condition $(X_0^1, ..., X_0^N)$ with bounded fourth moments (i.e $\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_0^i|^4\right) < \infty$) such that $t \mapsto \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - X_0^i|^2\right)$ is continuous in $t = 0$ uniformly in N. Then we get strong uniform in time propagation of chaos, i.e*

$$\forall \epsilon > 0,\ \exists N \geqslant 0,\ \forall t \geqslant 0,\ \forall n \geqslant N,\ \mathbb{E}\left(\mathcal{W}_2\left(\mu_t^n, \bar{\rho}_t\right)\right) < \epsilon.$$

We give in Appendix D.3 some sufficient conditions for this assumption of continuity for the initial conditions.

Notation

We try to keep coherent notation throughout the chapter, but as the various objects and what they represent may become confusing, we list them here for reference :

- $\mathcal{P}(\mathcal{X})$ is the set of probability measures on the set $\mathcal{X}$, and $\mathcal{P}_p(\mathcal{X})$ is the set of probability measures on the set $\mathcal{X}$ with finite p-th moment,
- $(X_t^1, ..., X_t^N)$, or $(X_t^{1,N}, ..., X_t^{N,N})$ when we need to insist on the total number of particles, is the solution of the SDE defining our particle system. X_t^i denotes the position in $\mathbb{R}$ of the i-th particle,
- $\mathcal{O}_N := \left\{ \mathbf{X} = (x_1, ..., x_N) \in \mathbb{R}^N \text{ s.t. } -\infty < x_1 < ... < x_N < \infty \right\}$ is the set in which we prove the solutions are,
- $\mu_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i} \in \mathcal{P}(\mathbb{R})$ is the empirical measure at time t of the N particle system. Notice that it is a random variable on the set $\mathcal{P}(\mathbb{R})$,
- $\xi_t^N \in \mathcal{P}(\mathcal{P}(\mathbb{R}))$ is the law of μ_t^N,
- $\rho_t^N \in \mathcal{P}(\mathcal{O}_N)$ is the joint law of $(X_t^1, ..., X_t^N)$,
- $\bar{\rho}_t \in \mathcal{P}(\mathbb{R})$ is the limit towards which μ_t^N will converge,
- all the notation above used with $t = \infty$ refer to the stationary distribution (provided it exists),
- $\mathcal{C}(\mathbb{R}^+, \mathcal{P}_2(\mathbb{R}))$ is the space of continuous functions taking values in the space of probability measures $\mathcal{P}_2(\mathbb{R})$ endowed with the L^2 Wasserstein distance,
- $\mu^N = (\mu_t^N)_{t \geqslant 0} \in \mathcal{C}(\mathbb{R}^+, \mathcal{P}_2(\mathbb{R}))$ and $\bar{\rho} = (\bar{\rho}_t)_{t \geqslant 0} \in \mathcal{C}(\mathbb{R}^+, \mathcal{P}_2(\mathbb{R}))$,
- for a probability measure μ and a measurable function f, we may denote both $\mu(f) := \int f d\mu$ and $\mathbb{E}^\mu(f(X)) := \int f d\mu$.

5.2 Existence, uniqueness and long time behavior of the particles

We start by gathering some technical results on the particle system.

5.2.1 Existence, uniqueness and no collisions

The goal of this subsection is to prove the following result.

Theorem 5.2.1. *Consider $N \geqslant 2$, and $-\infty < x_1 < ... < x_N < \infty$. Under Assumptions 5.1 and 5.2 :*

- *If $\alpha > 1$, for any $\sigma_N \geqslant 0$, there exists a unique strong solution $X = (X^1, ..., X^N)$ to the stochastic differential equation* (5.1.1) *with initial condition $X_0^1 = x_1$, ..., $X_0^N = x_N$, which furthermore satisfies $X_t \in \mathcal{O}_N$ for all $t \geqslant 0$, $\mathbb{P}$-a.s.*
- *The same result holds for $\alpha = 1$ and $\sigma_N \leqslant \frac{1}{N}$.*

Remark 5.2.1. *In the case $\alpha = 1$ and $\sigma_N = \sigma > 0$, the existence of a unique strong solution has been written in Theorem 2.5 of [43], where the authors allow collisions between particles and show that the system still satisfies $X_t \in \overline{\mathcal{O}_N}$ i.e*

$$-\infty < X_t^1 \leqslant ... \leqslant X_t^N < \infty \quad \textit{for all } t \geqslant 0, \mathbb{P}\text{-a.s.}$$

In the case $\alpha = 1$ and $\sigma_N \leqslant \frac{1}{N}$, the proof of existence, uniqueness and absence of collision has been done in [151] or in the more recent [125]. For the sake of completeness, and because it uses similar calculations, we also write the proof in this case here.

Denote the infinitesimal generator

$$\mathcal{L}^{N,\alpha} f(x_1,...,x_N) := -\sum_{i=1}^{N} U'(x_i)\partial_i f(x_1,...,x_N) \\ - \frac{1}{N}\sum_{i\neq j} V'(x_i - x_j)\partial_i f(x_1,...,x_N) + \sigma_N \Delta f(x_1,...,x_N).$$

and consider, for $\mathbf{X} = (x_1,...,x_N) \in \mathbb{R}^N$

$$H_{int,\alpha}(\mathbf{X}) := \frac{1}{2N}\sum_{i\neq j} V(x_i - x_j), \qquad H_\alpha(\mathbf{X}) := H_{int,\alpha}(\mathbf{X}) + \sum_{i=1}^{N}\frac{x_i^2}{2},$$

where $H_{int,\alpha}$ denotes the interaction potential. We prove the following Lyapunov conditions for the particles system. Let us begin with the case $\alpha > 1$.

Lemma 5.2.1. *Let $N > 1$. Under Assumptions 5.1 and 5.2, for $\alpha > 1$, there exist $C^{N,\alpha} > 0$ and $D^{N,\alpha} > 0$ such that for all $\boldsymbol{X} \in \mathcal{O}_N$*

$$\mathcal{L}^{N,\alpha} H_\alpha(\boldsymbol{X}) \leqslant D^{N,\alpha} + C^{N,\alpha} H_\alpha(\boldsymbol{X}).$$

Under the additional Assumption 5.3, still for $\alpha > 1$, there exist $c^{N,\alpha} > 0$ and $D^{N,\alpha} > 0$ such that for all $\boldsymbol{X} \in \mathcal{O}_N$

$$\mathcal{L}^{N,\alpha} H_\alpha(\boldsymbol{X}) \leqslant D^{N,\alpha} - c^{N,\alpha} H_\alpha(\boldsymbol{X}).$$

This lemma shows that, for a force U' Lipschitz continuous, the energy does not explode in finite time, which will provide us the existence of the solution of (5.1.1), and the absence of collision between particles. For a potential U convex, it even yields a uniform in time bound on the second moment of the particles, and on the expectation of $\frac{1}{|X^i - X^j|^{\alpha-1}}$ (even though we do not use this result to bound these moments, as $D^{N,\alpha}$ and $C^{N,\alpha}$ depend rather badly on N).

Proof. We compute

$$\begin{aligned}\mathcal{L}^{N,\alpha} H_\alpha(\mathbf{X}) = &-\sum_{i=1}^{N}\left(U'(x_i) + \frac{1}{N}\sum_{j\neq i} V'(x_i - x_j)\right)\left(x_i + \frac{1}{N}\sum_{j\neq i} V'(x_i - x_j)\right) \\ &+ \sigma_N \sum_{i=1}^{N} 1 + \frac{\sigma_N}{N}\sum_{i=1}^{N}\sum_{j\neq i} V''(x_i - x_j) \\ = &-\sum_{i=1}^{N} U'(x_i)x_i - \frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i} U'(x_i)V'(x_i - x_j) - \frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i} x_i V'(x_i - x_j) + N\sigma_N \\ &- \sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j\neq i} V'(x_i - x_j)\right)^2 + \frac{\sigma_N}{N}\sum_{i=1}^{N}\sum_{j\neq i} V''(x_i - x_j).\end{aligned}$$

We have, under Assumptions 5.2 and 5.1

$$-\sum_{i=1}^{N} U'(x_i)x_i \leqslant \sum_{i=1}^{N} L_U|x_i|^2 + A|x_i| \leqslant \left(L_U + \frac{1}{2}\right)\sum_{i=1}^{N}|x_i|^2 + \frac{NA^2}{2} \tag{5.2.1}$$

$$\begin{aligned}-\frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i} U'(x_i)V'(x_i-x_j) =& -\frac{1}{N}\sum_{j<i}\left(U'(x_i)-U'(x_j)\right)\frac{x_i-x_j}{|x_i-x_j|^{\alpha+1}}\\ \leqslant& \frac{L_U}{N}\sum_{j<i}\frac{1}{|x_i-x_j|^{\alpha-1}},\end{aligned} \tag{5.2.2}$$

$$-\frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i} x_iV'(x_i-x_j) \leqslant \frac{1}{N}\sum_{j<i}\frac{1}{|x_i-x_j|^{\alpha-1}}. \tag{5.2.3}$$

Let us now consider $|\nabla H_{int,\alpha}(\mathbf{X})| = \left(\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j\neq i}V'(x_i-x_j)\right)^2\right)^{1/2}$. We follow the proof of Lemma 5.15 of [94]. Let $j<i$, which implies $x_j<x_i$, and denote

$$\eta_k(\mathbf{X}) = \begin{cases} 1 & \text{if } x_k < x_i \\ -1 & \text{otherwise.}\end{cases}$$

Then, considering $\eta(\mathbf{X}) = (\eta_1(\mathbf{X}), ..., \eta_N(\mathbf{X}))$, we have

$$\begin{aligned}\sqrt{N}\,|\nabla H_{int,\alpha}(\mathbf{X})| \geqslant& \xi(\mathbf{X})\cdot\nabla H_{int,\alpha}(\mathbf{X})\\ =& \frac{1}{N}\sum_{k}\eta_k(\mathbf{X})\sum_{l\neq k}V'(x_k-x_l)\\ =& -\frac{1}{N}\sum_{k<l}\left(\eta_k(\mathbf{X})-\eta_l(\mathbf{X})\right)\frac{x_k-x_l}{|x_k-x_l|^{\alpha+1}}.\end{aligned}$$

Notice that, for $k<l$, $\eta_k(\mathbf{X}) \neq \eta_l(\mathbf{X})$ if and only if $x_k < x_i \leqslant x_l$, in which case we have $\eta_k(\mathbf{X})-\eta_l(\mathbf{X}) = 2$ and $(x_k-x_l)<0$. Therefore, the sum above only contains nonpositive terms. In particular, choosing $k=j$ and $l=i$, we get

$$\sqrt{N}\,|\nabla H_{int,\alpha}(\mathbf{X})| \geqslant \frac{2}{N}\frac{1}{|x_i-x_j|^{\alpha}}.$$

This holds for any $j<i$, thus

$$\begin{aligned}\sqrt{N}\frac{N(N-1)}{2}\,|\nabla H_{int,\alpha}(\mathbf{X})| \geqslant& \frac{2}{N}\sum_{j<i}\frac{1}{|x_i-x_j|^{\alpha}}, \qquad \text{i.e}\\ |\nabla H_{int,\alpha}(\mathbf{X})| \geqslant& \frac{4}{N^2(N-1)\sqrt{N}}\sum_{j<i}\frac{1}{|x_i-x_j|^{\alpha}}.\end{aligned}$$

We therefore have

$$-\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j\neq i}V'(x_i-x_j)\right)^2 + \frac{\sigma_N}{N}\sum_{i=1}^{N}\sum_{j\neq i}V''(x_i-x_j)$$

$$\leqslant \sum_{j<i} \frac{\sigma_N}{N} \frac{2\alpha}{|x_i - x_j|^{\alpha+1}} - \left(\frac{4}{N^2(N-1)\sqrt{N}}\right)^2 \frac{1}{|x_i - x_j|^{2\alpha}}.$$

For $\alpha > 1$, thanks to Young's inequality, there is a constant C_N such that

$$\frac{2\sigma_N \alpha}{N} \frac{1}{|x_i - x_j|^{\alpha+1}} - \left(\frac{4}{N^2(N-1)\sqrt{N}}\right)^2 \frac{1}{|x_i - x_j|^{2\alpha}} < C_N.$$

Therefore, using this result along with (5.2.1) and (5.2.2), we prove the existence of two nonnegative constants C and D, possibly depending on N, such that

$$\mathcal{L}^{N,\alpha} H_\alpha(\mathbf{X}) \leqslant D + C H_\alpha(\mathbf{X}).$$

Let us now modify the various estimates under the additional Assumption 5.3. We may replace the control (5.2.1) by

$$-\sum_{i=1}^N U'(x_i) x_i \leqslant -2\lambda \sum_{i=1}^N \frac{x_i^2}{2} \tag{5.2.4}$$

Then, instead of (5.2.2) and (5.2.3), we use the fact that there are C_N and D_N such that

$$\frac{(L_U + 1)}{N} \frac{1}{|x_i - x_j|^{\alpha-1}} + \frac{2\sigma_N \alpha}{N} \frac{1}{|x_i - x_j|^{\alpha+1}} - \left(\frac{4}{N^2(N-1)\sqrt{N}}\right)^2 \frac{1}{|x_i - x_j|^{2\alpha}}$$
$$< C_N - \frac{D_N}{|x_i - x_j|^{\alpha-1}}.$$

Combining this inequality with (5.2.4), we prove the existence of two nonnegative constants C and D, possibly depending on N, such that

$$\mathcal{L}^{N,\alpha} H_\alpha(\mathbf{X}) \leqslant D - C H_\alpha(\mathbf{X}).$$

□

Let us now consider the Coulomb case.

Lemma 5.2.2. *Let $N > 1$. Under Assumptions 5.1 and 5.2, for $\alpha = 1$ and $\sigma_N \leqslant \frac{1}{N}$, there exists $C^{N,\alpha}, D^{N,\alpha} > 0$ such that for all $\boldsymbol{X} \in \mathcal{O}_N$*

$$\mathcal{L}^{N,\alpha} H_\alpha(\boldsymbol{X}) \leqslant D^{N,\alpha} + C^{N,\alpha} H_\alpha(\boldsymbol{X}).$$

Proof. We compute

$$\mathcal{L}^{N,\alpha} H_\alpha(\mathbf{X}) = -\sum_{i=1}^N U'(x_i) x_i - \frac{1}{N} \sum_{i=1}^N \sum_{j \neq i} U'(x_i) V'(x_i - x_j) - \frac{1}{N} \sum_{i=1}^N \sum_{j \neq i} x_i V'(x_i - x_j)$$
$$+ N\sigma_N - \sum_{i=1}^N \left(\frac{1}{N} \sum_{j \neq i} V'(x_i - x_j)\right)^2 + \frac{\sigma_N}{N} \sum_{i=1}^N \sum_{j \neq i} V''(x_i - x_j).$$

Let us consider

$$\begin{aligned}\sum_{i=1}^N\left(\frac{1}{N}\sum_{j\neq i}\frac{x_i-x_j}{|x_i-x_j|^2}\right)^2 =&\frac{1}{N^2}\sum_{i=1}^N\sum_{j,k\neq i}\frac{x_i-x_j}{|x_i-x_j|^2}\frac{x_i-x_k}{|x_i-x_k|^2}\\ =&\frac{1}{N^2}\sum_{i=1}^N\sum_{j\neq i}\frac{1}{|x_i-x_j|^2}\\ &+\frac{1}{N^2}\sum_{i,j,k\text{ distincts}}\frac{x_i-x_j}{|x_i-x_j|^2}\frac{x_i-x_k}{|x_i-x_k|^2},\end{aligned}$$

and using the fact that for $\mathbf{X}\in\mathcal{O}_N$ and $i<j$ we have $x_i<x_j$, we obtain

$$\begin{aligned}&\sum_{i,j,k\text{ distincts}}\frac{x_i-x_j}{|x_i-x_j|^2}\frac{x_i-x_k}{|x_i-x_k|^2}\\ &=2\sum_{i=1}^N\sum_{\substack{j<k\\ j,k\neq i}}\frac{x_i-x_j}{|x_i-x_j|^2}\frac{x_i-x_k}{|x_i-x_k|^2}\\ &=2\sum_{i<j<k}\left(\frac{x_i-x_j}{|x_i-x_j|^2}\frac{x_i-x_k}{|x_i-x_k|^2}+\frac{x_j-x_i}{|x_j-x_i|^2}\frac{x_j-x_k}{|x_j-x_k|^2}\right.\\ &\qquad\left.+\frac{x_k-x_j}{|x_k-x_j|^2}\frac{x_k-x_i}{|x_k-x_i|^2}\right)\\ &=2\sum_{i<j<k}\left(\frac{1}{x_j-x_i}\frac{1}{x_k-x_i}-\frac{1}{x_j-x_i}\frac{1}{x_k-x_j}+\frac{1}{x_k-x_i}\frac{1}{x_k-x_j}\right)\\ &=2\sum_{i<j<k}\frac{1}{x_j-x_i}\frac{1}{x_k-x_i}\frac{1}{x_k-x_j}\left(x_k-x_j-x_k+x_i+x_j-x_i\right)\\ &=0.\end{aligned}$$

Furthermore, the estimates (5.2.1), (5.2.2) and (5.2.3) still hold. We thus have

$$\begin{aligned}-\sum_{i=1}^N U'(x_i)x_i &\leqslant \left(L_U+\frac{1}{2}\right)\sum_{i=1}^N|x_i|^2+\frac{NA^2}{2},\\ -\frac{1}{N}\sum_{i=1}^N\sum_{j\neq i}(U'(x_i)+x_i)V'(x_i-x_j) &\leqslant \frac{1}{2}(L_U+1)(N-1).\end{aligned}$$

Therefore

$$\begin{aligned}\mathcal{L}^{N,\alpha}H_\alpha(\mathbf{X})\leqslant&\frac{(L_U+1)(N-1)}{2}+\frac{NA^2}{2}+N\sigma_N+\left(L_U+\frac{1}{2}\right)\sum_{i=1}^N|x_i|^2\\ &+2\sum_{i<j}\left(\frac{\sigma_N}{N}-\frac{1}{N^2}\right)\frac{1}{|x_i-x_j|^2}.\end{aligned}$$

Noticing that there exist constants C, D such that $\sum_{i=1}^N |x_i|^2 \leqslant CH_\alpha(\mathbf{X}) + D$, we obtain the result if $\frac{\sigma_N}{N} \leqslant \frac{1}{N^2}$. □

Proof of Theorem 5.2.1. For $R > 0$, define $\tau_R := \inf\{t \geqslant 0 \quad \text{s.t} \quad H_\alpha(\mathbf{X}_t) > R\}$, $\tau := \lim_{R\to\infty} \tau_R$, and $\tau_{\partial\mathcal{O}_N} := \inf\{t \geqslant 0 \quad \text{s.t} \quad \mathbf{X}_t \in \partial\mathcal{O}_N\}$. We have $\{\tau = \infty\} \subset \{\tau_{\partial\mathcal{O}_N} = \infty\}$. Equation (5.1.1) with initial condition $\mathbf{X}_0 = \mathbf{x} \in \mathcal{O}_N$ has a strong solution up to the stopping time τ. Let us show that $\mathbb{P}_\mathbf{x}(\tau = \infty) = 1$.

- $\alpha > 1$ **:** Itô's formula for the function $f(t, \mathbf{x}) = e^{-C^{N,\alpha}t} H_\alpha(\mathbf{x})$, using Lemma 5.2.1, yields for all $R > 0$ and $t \geqslant 0$

$$\mathbb{E}_\mathbf{x}\left(e^{-C^{N,\alpha}(t\wedge\tau_R)} H_\alpha(\mathbf{X}_{t\wedge\tau_R})\right) \leqslant H_\alpha(\mathbf{x}) + \frac{D^{N,\alpha}}{C^{N,\alpha}},$$

and thus, as $H_\alpha \geqslant 0$,

$$Re^{-C^{N,\alpha}t}\mathbb{P}_x(\tau_R \leqslant t) \leqslant H_\alpha(\mathbf{x}) + \frac{D^{N,\alpha}}{C^{N,\alpha}}.$$

We obtain, for all $t \geqslant 0$

$$\mathbb{P}_x(\tau \leqslant t) = \lim_{R\to\infty} \mathbb{P}_x(\tau_R \leqslant t) \leqslant \lim_{R\to\infty} \frac{H_\alpha(\mathbf{x}) + \frac{D^{N,\alpha}}{C^{N,\alpha}}}{R} e^{C^{N,\alpha}t} = 0.$$

- $\alpha = 1$ **:** There exists a constant $H_0 \in \mathbb{R}$, possibly depending on N, such that for all $\mathbf{x} \in \mathcal{O}_N$, $H_\alpha(\mathbf{x}) \geqslant H_0$. Considering Itô's formula for the function $f(t, \mathbf{x}) = e^{-C^{N,\alpha}t}\left(H_\alpha(\mathbf{x}) + H_0\right)$, using Lemma 5.2.2, yields for all $R > 0$ and $t \geqslant 0$

$$\mathbb{E}_\mathbf{x}\left(e^{-C^{N,\alpha}(t\wedge\tau_R)}\left(H_\alpha(\mathbf{X}_{t\wedge\tau_R}) + H_0\right)\right) \leqslant H_\alpha(\mathbf{x}) + H_0 + \frac{D^{N,\alpha}}{C^{N,\alpha}},$$

and thus, as $H_\alpha + H_0 \geqslant 0$,

$$e^{-C^{N,\alpha}t}(R + H_0)\mathbb{P}_x(\tau_R \leqslant t) \leqslant H_\alpha(\mathbf{x}) + H_0 + \frac{D^{N,\alpha}}{C^{N,\alpha}}.$$

We obtain, for all $t \geqslant 0$

$$\mathbb{P}_x(\tau \leqslant t) = \lim_{R\to\infty} \mathbb{P}_x(\tau_R \leqslant t) \leqslant \lim_{R\to\infty} \frac{H_\alpha(\mathbf{x}) + H_0 + \frac{D^{N,\alpha}}{C^{N,\alpha}}}{R + H_0} e^{C^{N,\alpha}t} = 0.$$

We thus have, in both cases, $\forall t \geqslant 0, \mathbb{P}_x(\tau > t) = 1$. This implies the particle system almost surely does not explode nor collide in finite time.

Uniqueness of the solution of (5.1.1) is a direct consequence of (5.2.5) in Theorem 5.2.2 below. □

5.2.2 Long time behavior

In this section we study the long time behavior of the particle system. Parts of the result will also allow us to conclude on the uniqueness of the particle system.

An important tool is the convexity of the interaction (see Remark 5.1.1).

Theorem 5.2.2. *Consider two solutions X and Y of* (5.1.1) *driven by the same Brownian motions.*

- *Under Assumptions 5.1 and 5.2, we have*

$$\sum_{i=1}^{N}\left(X_t^i - Y_t^i\right)^2 \leqslant e^{2L_U t}\sum_{i=1}^{N}\left(X_0^i - Y_0^i\right)^2. \tag{5.2.5}$$

This yields strong uniqueness of the solution of (5.1.1).

- *Under Assumptions 5.1 and 5.2, denoting by $\rho_t^{1,N}$ and $\rho_t^{2,N}$ the laws on $\mathcal{O}_N$ of the particle systems with respective initial conditions $\rho_0^{1,N}$ and $\rho_0^{2,N}$, we have*

$$\forall t \geqslant 0, \qquad \mathcal{W}_2\left(\rho_t^{1,N}, \rho_t^{2,N}\right) \leqslant e^{L_U t}\mathcal{W}_2\left(\rho_0^{1,N}, \rho_0^{2,N}\right). \tag{5.2.6}$$

- *Under Assumptions 5.1 and 5.3, we have*

$$\forall t \geqslant 0, \qquad \mathcal{W}_2\left(\rho_t^{1,N}, \rho_t^{2,N}\right) \leqslant e^{-\lambda t}\mathcal{W}_2\left(\rho_0^{1,N}, \rho_0^{2,N}\right). \tag{5.2.7}$$

Proof. Let $(X_t^i)_{1\leqslant i\leqslant N}$ and $(Y_t^i)_{1\leqslant i\leqslant N}$ be two solutions of (5.1.1) driven by the same set of Brownian motions (i.e coupled using a synchronous coupling), such that $X_t^1 < ... < X_t^N$ and $Y_t^1 < ... < Y_t^N$. Using Itô's formula,

$$\begin{aligned} d\left(\sum_{i=1}^{N}\left(X_t^i - Y_t^i\right)^2\right) =& -2\sum_{i=1}^{N}\left(U'(X_t^i) - U'(Y_t^i)\right)\left(X_t^i - Y_t^i\right)dt \\ & - \frac{1}{N}\sum_{i=1}^{N} 2\left(X_t^i - Y_t^i\right)\sum_{j\neq i}\left(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\right)dt, \end{aligned}$$

with, since $x \to V'(x)$ is odd and increasing for $x > 0$ under Assumption 5.1

$$\begin{aligned} &\frac{2}{N}\sum_{1\leqslant j\neq i\leqslant N}\left(X_t^i - Y_t^i\right)\left(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\right) \\ =&\frac{2}{N}\sum_{1\leqslant j< i\leqslant N}\left(\left(X_t^i - Y_t^i\right) - \left(X_t^j - Y_t^j\right)\right)\left(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\right) \\ =&\frac{2}{N}\sum_{1\leqslant j< i\leqslant N}\left(\left(X_t^i - X_t^j\right) - \left(Y_t^i - Y_t^j\right)\right)\left(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\right) \\ \geqslant& 0. \end{aligned}$$

Under Assumption 5.2, we obtain

$$\frac{d}{dt}\sum_{i=1}^{N}\left(X_t^i - Y_t^i\right)^2 \leqslant 2L_U\sum_{i=1}^{N}\left(X_t^i - Y_t^i\right)^2,$$

i.e

$$d\left(e^{-2L_U t}\sum_{i=1}^N \left(X_t^i - Y_t^i\right)^2\right) = K_t dt,$$

with $K_t \leqslant 0$. We thus obtain

$$\sum_{i=1}^N \left(X_t^i - Y_t^i\right)^2 \leqslant e^{2L_U t}\sum_{i=1}^N \left(X_0^i - Y_0^i\right)^2.$$

This yields the results (5.2.5) and (5.2.6).
Under Assumption 5.3, similar calculations yield

$$\frac{d}{dt}\sum_{i=1}^N \left(X_t^i - Y_t^i\right)^2 \leqslant -2\lambda\sum_{i=1}^N \left(X_t^i - Y_t^i\right)^2.$$

and thus (5.2.7). □

5.2.3 Some moment bounds

The aim of this section is to provide some explicit bounds on the second moment of the empirical measure, as well as on the expectation of the interaction potential. These bounds will be useful later when proving propagation of chaos. Let, for $\mathbf{x} \in \mathbb{R}^N$,

$$\mathcal{H}(\mathbf{x}) = \sum_{i=1}^N |x_i|^2 - \frac{1}{2N}\sum_{i\neq j}|x_i - x_j|. \tag{5.2.8}$$

The idea of considering this function comes from [128].

Lemma 5.2.3. *Consider Assumptions 5.1 and 5.4. The function $\mathcal{H}$ satisfies*

$$\forall \boldsymbol{x} \in \mathbb{R}^N, \quad \mathcal{H}(x) \geqslant \frac{1}{2}\sum_i |x_i|^2 - N. \tag{5.2.9}$$

Given $(\boldsymbol{X}_t)_t \geqslant 0$ a solution of (5.1.1), we have the uniform in time bound

$$\mathbb{E}\mathcal{H}(\boldsymbol{X}_t) \leqslant e^{-2\lambda t}\mathbb{E}\mathcal{H}(\boldsymbol{X}_0) + \frac{1}{\lambda}\left(N\sigma_N + \frac{C(\alpha,N)}{\alpha}\right), \tag{5.2.10}$$

as well as the following estimates

$$\mathbb{E}\left(\int_0^t \frac{e^{2\lambda s}}{N^2}\sum_{i>j}\frac{i-j}{\left|X_s^i - X_s^j\right|^\alpha}ds\right) \leqslant \frac{\alpha}{2}\left(\mathbb{E}\mathcal{H}(\boldsymbol{X}_0) + Ne^{2\lambda t} + 2N\sigma_N\frac{e^{2\lambda t}-1}{2\lambda}\right) + C(\alpha,N)\frac{e^{2\lambda t}-1}{2\lambda}, \tag{5.2.11}$$

$$\mathbb{E}\left(\int_0^t \frac{1}{N^2}\sum_{i>j}\frac{i-j}{\left|X_s^i - X_s^j\right|^\alpha}ds\right) \leqslant \frac{\alpha}{2}\left(\mathbb{E}\mathcal{H}(\boldsymbol{X}_0) + N + (2N\sigma_N + 2\lambda N)t\right) + C(\alpha,N)t, \tag{5.2.12}$$

where

$$C(\alpha,N)=\begin{cases}\frac{N-1}{2} & \textit{if } \alpha=1,\\ \frac{N}{2-\alpha} & \textit{if } \alpha\in]1,2[,\\ 2N\ln N & \textit{if } \alpha=2,\\ \left(1+\frac{1}{\alpha-2}\right)N^{\alpha-1} & \textit{if } \alpha>2.\end{cases}$$

The proof relies on the computation of the time evolution of $\mathcal{H}(\mathbf{X}_t)$ using Itô's formula. Using parts of the calculations of [128], as well as some technical results, we obtain the various bounds. We postpone the proof of this lemma to Appendix D.2 for the sake of clarity.

5.3 Limit for large number of particles with vanishing noise

Consider for a given $N \geqslant 1$ a solution $X_t = (X_t^1, ..., X_t^N)$ of (5.1.1). Our goal is to prove the following theorem

Theorem 5.3.1. *Consider a sequence of initial empirical measures $(\mu_0^N)_{N\geqslant 1}$ such that there exists $\bar\rho_0 \in \mathcal{P}_2(\mathbb{R})$ such that $\lim_{N\to 0}\mathbb{E}\left(\mathcal{W}_2(\mu_0^N,\bar\rho_0)^2\right)=0$. Under Assumptions 5.1 and 5.3 for $\alpha = 1$ or under Assumptions 5.1 and 5.4 for $\alpha \in]1,2[$ (with the additional assumption $\sigma_N \leqslant \frac{1}{N}$ for $\alpha = 1$), there exist a deterministic family of measures $(\bar\rho_t)_{t\geqslant 0} \in \mathcal{C}(\mathbb{R}^+,\mathcal{P}_2(\mathbb{R}))$, as well as universal constants $C_1, C_2 > 0$ and a quantity $C_0^N > 0$ that depends on the initial condition and such that $C_0^N \to 0$ as $N \to \infty$, such that for all $N \geqslant 1$ and all $t \geqslant 0$*

$$\mathbb{E}\left(\mathcal{W}_2(\mu_t^N,\bar\rho_t)^2\right) \leqslant e^{-2\lambda t}C_0^N + \frac{C_1}{N^{(2-\alpha)/\alpha}} + C_2\sigma_N.$$

In particular, notice that we require the diffusion coefficient σ_N to go to 0 in order to obtain the limit for the empirical measure.

Here, we do not identify the limit $\bar\rho_t$, we just prove its existence. The limit will be studied in Section 5.4 later.

The proof of Theorem 5.3.1 is divided in two parts. First, we prove a property on any sequence of independent empirical measures $(\mu^N)_N$ which is similar to a Cauchy property. Then, we use the independence of the random variables $(\mu^N)_N$ to conclude on the convergence towards a deterministic limit.

We will start by proving the Cauchy estimate in the case $\alpha = 1$, as this will allow us to describe the method in an easier case, before extending the result to $\alpha \in]1,2[$ using much more cumbersome computations.

5.3.1 The case $\alpha = 1$.

In this section, we prove the following lemma, which states a Cauchy property for the sequence of empirical measures :

Lemma 5.3.1. *Consider Assumption 5.1 and Assumption 5.3, with $\alpha = 1$ and $\sigma_N \leqslant \frac{1}{N}$. Let $(\mu^N)_{N\in\mathbb{N}}$ be any sequence of independent empirical measures, such that μ_t^N is the empirical measure of the N particle system at time t. We have for all $t \geqslant 0$ and all $N, M \geqslant 1$*

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N,\mu_t^M\right)^2\right) \leqslant e^{-2\lambda t}\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right) + \frac{1}{2\lambda}\left(\frac{1}{N}+\frac{1}{M}+2\left(\sigma_N+\sigma_M\right)\right), \tag{5.3.1}$$

There also are constants $C_1, C_2, C_3 > 0$ independent of N and M such that

$$\mathbb{E}\left(\sup_{s\in[0,t]} \mathcal{W}_2\left(\mu_s^N, \mu_s^M\right)^2\right) \leqslant e^{C_1 t}\left(\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2\right) + C_2(\sigma_M + \sigma_N) + C_3\left(\frac{1}{M} + \frac{1}{N}\right)\right). \tag{5.3.2}$$

Note that because of the expectation in (5.3.1), we cannot say that $(\mu_t^N)_{N\in\mathbb{N}}$ is a Cauchy sequence in the space of probability measures on $\mathbb{R}$. It is however a Cauchy sequence in the space of random probability measures, see Remark 5.3.2 below.

Proof. For $N, M \geqslant 1$, let $(\tilde{B}^i)_{i\in\{1,...,M\}}$ and $(\tilde{B}'^j)_{j\in\{1,...,N\}}$ be two independent families of Brownian motions, and consider $x_1 < ... < x_M$ and $y_1 < ... < y_N$ two sets of initial conditions. Denote by $(\tilde{X}^{i,M})_{i\in\{1,...,M\}}$ (resp. $(\tilde{Y}^{j,N})_{j\in\{1,...,N\}}$) the unique strong solution of (5.1.1) with initial conditions $x_1 < ... < x_M$ and Brownian motions $(\tilde{B}^i)_{i\in\{1,...,M\}}$ (resp. initial conditions $y_1 < ... < y_N$ and Brownian motions $(\tilde{B'}^j)_{j\in\{1,...,N\}}$).

In order to compare the two sets $(\tilde{X}^{i,M})_{i\in\{1,...,M\}}$ and $(\tilde{Y}^{j,N})_{j\in\{1,...,M\}}$ despite the difference in the number of particles, we consider N exact copies of the system $(\tilde{X}^{i,M})_{i\in\{1,...,M\}}$, and M exact copies of $(\tilde{Y}^{j,N})_{j\in\{1,...,N\}}$. We denote $(X^i)_{i\in\{1,...,NM\}}$ and $(Y^i)_{i\in\{1,...,NM\}}$ the resulting processes, numbered such that for all $t \geqslant 0$

$$\begin{aligned}
-\infty <& X_t^1 = ... = X_t^N < ... < X_t^{N(M-1)+1} = ... = X_t^{NM} < \infty \\
-\infty <& Y_t^1 = ... = Y_t^M < ... < Y_t^{M(N-1)+1} = ... = Y_t^{NM} < \infty.
\end{aligned}$$

Thus

$$\mu_t^M = \frac{1}{M}\sum_{i=1}^{M} \delta_{\tilde{X}_t^{i,M}} = \frac{1}{NM}\sum_{i=1}^{NM} \delta_{X_t^i}, \qquad \mu_t^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\tilde{Y}_t^{i,N}} = \frac{1}{NM}\sum_{i=1}^{NM} \delta_{Y_t^i},$$

and

$$\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 = \frac{1}{NM}\sum_{i=1}^{NM} \left|X_t^i - Y_t^i\right|^2.$$

By convention, and for the sake of clarity, consider $V'(0) = 0$. Then, for all $i \in \{1, ..., NM\}$, we have the following dynamics

$$\begin{aligned}
dX_t^i =& -U'(X_t^i)dt - \frac{1}{NM}\sum_j V'(X_t^i - X_t^j)dt + \sqrt{2\sigma_M}dB_t^i \\
dY_t^i =& -U'(Y_t^i)dt - \frac{1}{NM}\sum_j V'(Y_t^i - Y_t^j)dt + \sqrt{2\sigma_N}dB_t'^i,
\end{aligned}$$

where the Brownian motions $(B_t^i)_i$ and (resp. $(B_t'^i)_i$) are such that for all $k \in \{1, ..., M\}$, we have $B^{N(k-1)+1} = ... = B^{Nk} = \tilde{B}^k$, (resp. for all $l \in \{1, ..., N\}$, $B'^{M(l-1)+1} = ... = B'^{Ml} = \tilde{B}'^l$). Thus

$$d\left(X_t^i - Y_t^i\right)^2 = -2\left(U'(X_t^i) - U'(Y_t^i)\right)\left(X_t^i - Y_t^i\right)dt + 2\sigma_M dt + 2\sigma_N dt$$

$$
\begin{aligned}
&- 2\left(X_t^i - Y_t^i\right)\frac{1}{NM}\sum_j \Big(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\Big)\,dt \\
&+ 2\sqrt{2\sigma_M}\left(X_t^i - Y_t^i\right)dB_t^i - 2\sqrt{2\sigma_N}\left(X_t^i - Y_t^i\right)dB_t^{'i},
\end{aligned}
$$

and

$$
\begin{aligned}
d\left(\frac{1}{NM}\sum_i \left(X_t^i - Y_t^i\right)^2\right) = &-\frac{2}{NM}\sum_i \left(U'(X_t^i) - U'(Y_t^i)\right)\left(X_t^i - Y_t^i\right)dt + 2(\sigma_M + \sigma_N)dt \\
&+ \frac{2\sqrt{2\sigma_M}}{NM}\sum_i \left(X_t^i - Y_t^i\right)dB_t^i - \frac{2\sqrt{2\sigma_N}}{NM}\sum_i \left(X_t^i - Y_t^i\right)dB_t^{'i} \\
&- \frac{2}{(NM)^2}\sum_i \left(X_t^i - Y_t^i\right)\sum_j \Big(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\Big)\,dt.
\end{aligned}
$$

We first compute

$$
\begin{aligned}
&\sum_i \left(X_t^i - Y_t^i\right)\sum_j \Big(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\Big) \\
&\qquad = \sum_{i>j} \Big(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\Big)\Big(\left(X_t^i - Y_t^i\right) - \left(X_t^j - Y_t^j\right)\Big) \\
&\qquad = \sum_{i>j} \Big(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\Big)\Big(\left(X_t^i - X_t^j\right) - \left(Y_t^i - Y_t^j\right)\Big).
\end{aligned}
$$

Remember that the function $x \to V'(x)$ is increasing for $x > 0$. Thus, all choices of indexes $i > j$ such that $X_t^i \neq X_t^j$ (which therefore imply, by the choice of numbering, that $X_t^i > X_t^j$) and $Y_t^i \neq Y_t^j$ yield nonnegative terms in the sum above. If $X_t^i = X_t^j$, by convention, we have $V'(X_t^i - X_t^j) = 0$.

$$
\begin{aligned}
\sum_i \left(X_t^i - Y_t^i\right)\sum_j \Big(V'(X_t^i - X_t^j) - V'(Y_t^i - Y_t^j)\Big) &\geqslant \sum_{i>j \text{ s.t } Y_t^i = Y_t^j} V'(X_t^i - X_t^j)\left(X_t^i - X_t^j\right) \\
&\quad + \sum_{i>j \text{ s.t } X_t^i = X_t^j} V'(Y_t^i - Y_t^j)\left(Y_t^i - Y_t^j\right) \qquad (5.3.3) \\
&\geqslant \sum_{i>j \text{ s.t } Y_t^i = Y_t^j} -1 + \sum_{i>j \text{ s.t } X_t^i = X_t^j} -1 \\
&= -\frac{M(M-1)}{2}N - \frac{N(N-1)}{2}M.
\end{aligned}
$$

Since $-\left(U'(X_t^i) - U'(Y_t^i)\right)\left(X_t^i - Y_t^i\right) \leqslant \lambda\left(X_t^i - Y_t^i\right)^2$, we thus obtain, for all $t \geqslant 0$

$$
\begin{aligned}
\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 \leqslant &\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2 - 2\lambda\int_0^t \mathcal{W}_2\left(\mu_s^N, \mu_s^M\right)^2 ds \\
&+ \int_0^t \left(\frac{2}{(NM)^2}\left(\frac{N(N-1)}{2}M + \frac{M(M-1)}{2}N\right) + 2(\sigma_M + \sigma_N)\right)ds
\end{aligned}
$$

$$+\frac{2\sqrt{2\sigma_M}}{NM}\sum_i\int_0^t\left(X_s^i-Y_s^i\right)dB_s^i-\frac{2\sqrt{2\sigma_N}}{NM}\sum_i\int_0^t\left(X_s^i-Y_s^i\right)dB_s^{'i}. \quad (5.3.4)$$

Considering the expectation of the inequality above, and using Gronwall's lemma yields (5.3.1). Let us now take the supremum

$$\begin{aligned}\mathbb{E}\left(\sup_{s\in[0,t]}\mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right)\leqslant&\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right)+\left(\frac{1}{M}+\frac{1}{N}\right)t+2(\sigma_M+\sigma_N)t\\&+\mathbb{E}\left(\frac{2\sqrt{2\sigma_M}}{NM}\sup_{s\in[0,t]}\sum_i\int_0^s\left(X_u^i-Y_u^i\right)dB_u^i\right)\\&+\mathbb{E}\left(\frac{2\sqrt{2\sigma_N}}{NM}\sup_{s\in[0,t]}\sum_i-\int_0^s\left(X_u^i-Y_u^i\right)dB_u^{'i}\right).\end{aligned}$$

We use Burkholder-Davis-Gundy inequality to show that there exists a constant C_{BDG} such that

$$\begin{aligned}\mathbb{E}&\left(\frac{2\sqrt{2\sigma_M}}{NM}\sup_{s\in[0,t]}\sum_i\int_0^s\left(X_u^i-Y_u^i\right)dB_u^i\right)\\&\leqslant\frac{2\sqrt{2\sigma_M}}{NM}\sum_i\mathbb{E}\left(\sup_{s\in[0,t]}\int_0^s\left(X_u^i-Y_u^i\right)dB_u^i\right)\\&\leqslant C_{BDG}\frac{2\sqrt{2\sigma_M}}{NM}\sum_i\mathbb{E}\left(\left(\int_0^t\left(X_s^i-Y_s^i\right)^2ds\right)^{1/2}\right)\\&\leqslant C_{BDG}\frac{2\sqrt{2\sigma_M}}{NM}\sum_i\left(\frac{1}{2\sqrt{2\sigma_M}}\mathbb{E}\left(\int_0^t\left(X_s^i-Y_s^i\right)^2ds\right)+\frac{\sqrt{2\sigma_M}}{2}\right)\\&=C_{BDG}\mathbb{E}\left(\int_0^t\frac{1}{NM}\sum_i\left(X_s^i-Y_s^i\right)^2ds\right)+2\sigma_MC_{BDG}\\&=C_{BDG}\mathbb{E}\left(\int_0^t\mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2ds\right)+2\sigma_MC_{BDG}.\end{aligned}$$

Using the same control on the second local martingale, we get

$$\begin{aligned}\mathbb{E}\left(\sup_{s\in[0,t]}\mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right)\leqslant\ &\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right)+\left(\frac{1}{N}+\frac{1}{M}\right)t+2(\sigma_M+\sigma_N)t\\&+2C_{BDG}(\sigma_M+\sigma_N)+2C_{BDG}\int_0^t\mathbb{E}\left(\mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right)ds,\end{aligned}$$

and thus, denoting

$$C_{N,M}=\frac{1}{2C_{BDG}}\left(\frac{1}{M}+\frac{1}{N}+2(\sigma_M+\sigma_N)\right),$$

we get

$$\mathbb{E}\left(\sup_{s\in[0,t]}\mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right)+C_{N,M}\leqslant\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right)+2C_{BDG}(\sigma_M+\sigma_N)+C_{N,M}$$

$$+ 2C_{BDG} \int_0^t \left(\mathbb{E}\left(\mathcal{W}_2\left(\mu_s^N, \mu_s^M\right)^2 \right) + C_{N,M} \right) ds.$$

Gronwall's lemma yields (5.3.2). □

Remark 5.3.1. *For $\lambda = 0$, the proof above still yields a quantitative result of propagation of chaos, though no longer uniform in time : considering* (5.3.4)*, we get for all $t \geqslant 0$*

$$\mathbb{E}\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 \leqslant \mathbb{E}\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2 + \left(\frac{1}{N} + \frac{1}{M} + 2(\sigma_N + \sigma_M) \right) t.$$

Likewise, under Assumption 5.2 instead of Assumption 5.3, we get a similar (non uniform in time) result

$$\mathbb{E}\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 \leqslant e^{2L_U t} \left(\mathbb{E}\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2 + \frac{1}{2L_U} \left(\frac{1}{N} + \frac{1}{M} + 2(\sigma_N + \sigma_M) \right) \right).$$

5.3.2 The case $\alpha \in]1, 2[$.

Let us now show the proof of Lemma 5.3.1 can be extended to other values of α. Notice that we use the assumption $\alpha = 1$ to deal with (5.3.3). To account for this quantity for $\alpha > 1$, we now use the bound (5.2.11) and obtain, using the definition of $\mathcal{H}$ given in (5.2.8), the following lemma.

Lemma 5.3.2. *Consider Assumptions 5.1 and 5.4, with $\alpha \in]1, 2[$. Let $(\mu^N)_{N \in \mathbb{N}}$ be any sequence of independent empirical measures, such that μ_t^N is the empirical measure of the N particle system at time t. We have for all $t \geqslant 0$ and all $N, M \geqslant 1$*

$$\begin{aligned} &\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 \right) \qquad (5.3.5) \\ \leqslant & e^{-2\lambda t} \left(\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2 \right) + \frac{3(\alpha-1)}{N^{(2-\alpha)/\alpha}} \mathbb{E}^{\mu_0^M}\left(|X|^2\right) + \frac{3(\alpha-1)}{M^{(2-\alpha)/\alpha}} \mathbb{E}^{\mu_0^N}\left(|Y|^2\right) \right) \\ &+ \frac{1}{\lambda}(\sigma_M + \sigma_N) + \frac{1}{\lambda} \left(\frac{3\alpha - 2}{\alpha(2-\alpha)} + 3\lambda(\alpha - 1) \right) \left(\frac{1}{N^{(2-\alpha)/\alpha}} + \frac{1}{M^{(2-\alpha)/\alpha}} \right) \\ &+ \frac{3(\alpha-1)}{2\lambda} \left(\frac{\sigma_N}{M^{(2-\alpha)/\alpha}} + \frac{\sigma_M}{N^{(2-\alpha)/\alpha}} \right) \end{aligned}$$

Proof. Consider a similar set up as the proof of Lemma 5.3.1, and define $(\tilde{X}_t^{i,M})_i$,$(\tilde{Y}_t^{j,N})_j$, $(X_t^i)_i$, $(Y_t^j)_j$ in the same manner. We compute, like previously

$$\begin{aligned} d\left(e^{2\lambda t} \mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 \right) =& 2\lambda e^{2\lambda t} \mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 dt + e^{2\lambda t} d\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2 \\ =& e^{2\lambda t} A_t dt + e^{2\lambda t} dM_t \end{aligned}$$

where M_t is a local martingale, and

$$\begin{aligned} A_t \leqslant & -\frac{2}{(NM)^2} \sum_{i>j \text{ s.t } Y_t^i = Y_t^j} V'(X_t^i - X_t^j)\left(X_t^i - X_t^j \right) \\ & -\frac{2}{(NM)^2} \sum_{i>j \text{ s.t } X_t^i = X_t^j} V'(Y_t^i - Y_t^j)\left(Y_t^i - Y_t^j \right) + 2(\sigma_M + \sigma_N). \end{aligned}$$

Using Young's inequality, we have, for all $\gamma > 0$ and $i > j$

$$\frac{1}{|x|^{\alpha-1}} \leqslant \gamma^{\frac{\alpha}{\alpha-1}} \frac{\alpha-1}{\alpha} \frac{\lfloor \frac{i-j}{N} \rfloor + 1}{|x|^{\alpha}} + \frac{1}{\alpha\gamma^{\alpha}\left(\lfloor \frac{i-j}{N} \rfloor + 1\right)^{\alpha-1}}.$$

Hence

$$\begin{aligned}\frac{1}{(NM)^2} \sum_{i>j \text{ s.t } \substack{Y_t^i = Y_t^j \\ X_t^i \neq X_t^j}} \frac{1}{|X_t^i - X_t^j|^{\alpha-1}} \leqslant& \frac{1}{(NM)^2} \gamma^{\frac{\alpha}{\alpha-1}} \frac{\alpha-1}{\alpha} \sum_{i>j \text{ s.t } \substack{Y_t^i = Y_t^j \\ X_t^i \neq X_t^j}} \frac{\lfloor \frac{i-j}{N} \rfloor + 1}{|X_t^i - X_t^j|^{\alpha}} \\ &+ \frac{1}{(NM)^2} \frac{1}{\alpha\gamma^{\alpha}} \sum_{i>j \text{ s.t } Y_t^i = Y_t^j} \frac{1}{\left(\lfloor \frac{i-j}{N} \rfloor + 1\right)^{\alpha-1}}.\end{aligned}$$

We calculate, since $\lfloor \frac{i-j}{N} \rfloor + 1 \geqslant \frac{i-j}{N}$

$$\sum_{i>j \text{ s.t } Y_t^i = Y_t^j} \frac{1}{\left(\lfloor \frac{i-j}{N} \rfloor + 1\right)^{\alpha-1}} \leqslant \sum_{i>j \text{ s.t } Y_t^i = Y_t^j} \frac{N^{\alpha-1}}{(i-j)^{\alpha-1}} \quad \text{and}$$

$$\sum_{i>j \text{ s.t } Y_t^i = Y_t^j} \frac{1}{(i-j)^{\alpha-1}} = \sum_{i=1}^{NM} \sum_{j=\lfloor \frac{i-1}{M} \rfloor M+1}^{i-1} \frac{1}{(i-j)^{\alpha-1}} = \sum_{i=1}^{NM} \sum_{j=1}^{i-1-\lfloor \frac{i-1}{M} \rfloor M} \frac{1}{j^{\alpha-1}}$$

which implies using Lemma D.1.1

$$\sum_{i>j \text{ s.t } Y_t^i = Y_t^j} \frac{1}{\left(\lfloor \frac{i-j}{N} \rfloor + 1\right)^{\alpha-1}} \leqslant N^{\alpha-1} \sum_{i=1}^{NM} \frac{1}{2-\alpha} \left(i-1-\left\lfloor \frac{i-1}{M} \right\rfloor M\right)^{2-\alpha} \leqslant \frac{N^{\alpha} M M^{2-\alpha}}{2-\alpha}.$$

Hence

$$\frac{1}{(NM)^2} \frac{1}{\alpha\gamma^{\alpha}} \sum_{i>j \text{ s.t } Y_t^i = Y_t^j} \frac{1}{\left(\lfloor \frac{i-j}{N} \rfloor + 1\right)^{\alpha-1}} \leqslant \frac{1}{N^{2-\alpha}} \frac{1}{M^{\alpha-1}} \frac{1}{\alpha(2-\alpha)\gamma^{\alpha}}.$$

We consider $\gamma^{\frac{\alpha}{\alpha-1}} = \frac{1}{M^{1+\delta}}$ for some yet unspecified $\delta > 0$. Thus

$$\frac{1}{(NM)^2} \frac{1}{\alpha\gamma^{\alpha}} \sum_{i>j \text{ s.t } Y_t^i = Y_t^j} \frac{1}{\left(\lfloor \frac{i-j}{N} \rfloor + 1\right)^{\alpha-1}} \leqslant \frac{1}{\alpha(2-\alpha)N^{2-\alpha}} \frac{M^{(1+\delta)(\alpha-1)}}{M^{\alpha-1}} = \frac{M^{\delta(\alpha-1)}}{\alpha(2-\alpha)N^{2-\alpha}}.$$

Furthermore

$$\begin{aligned}\frac{1}{(NM)^2} \sum_{i>j \text{ s.t } \substack{Y_t^i = Y_t^j \\ X_t^i \neq X_t^j}} \frac{\lfloor \frac{i-j}{N} \rfloor + 1}{|X_t^i - X_t^j|^{\alpha}} \leqslant& \frac{1}{(NM)^2} \sum_{i=0}^{M-1} \sum_{j=0}^{i-1} \sum_{k=1}^{N} \sum_{l=1}^{N} \frac{\lfloor i-j+\frac{k-l}{N} \rfloor + 1}{|X_t^{iN+k} - X_t^{jN+l}|^{\alpha}} \\ \leqslant& \frac{1}{(NM)^2} \sum_{i=1}^{M} \sum_{j=1}^{i-1} N^2 \frac{i-j+2}{|\tilde{X}_t^i - \tilde{X}_t^j|^{\alpha}},\end{aligned}$$

and thus

$$\gamma^{\frac{\alpha}{\alpha-1}}\frac{\alpha-1}{\alpha}\frac{1}{(NM)^2}\sum_{\substack{i>j \text{ s.t } \begin{subarray}{c} Y_t^i = Y_t^j \\ X_t^i \neq X_t^j \end{subarray}}}\frac{\lfloor\frac{i-j}{N}\rfloor+1}{|X_t^i - X_t^j|^\alpha} \leqslant 3\frac{\alpha-1}{\alpha}\frac{1}{M^{1+\delta}}\frac{1}{M^2}\sum_{i>j}\frac{i-j}{|\tilde{X}_t^i - \tilde{X}_t^j|^\alpha}.$$

Using the same calculations to deal with

$$\frac{1}{(NM)^2}\sum_{\substack{i>j \text{ s.t } \begin{subarray}{c} X_t^i = X_t^j \\ Y_t^i \neq Y_t^j \end{subarray}}}\frac{1}{|Y_t^i - Y_t^j|^{\alpha-1}},$$

we obtain by taking the expectation in Itô's formula that for all $t \geqslant 0$

$$\begin{aligned}
&e^{2\lambda t}\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2\right) \\
\leqslant&\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2\right) + 2(\sigma_M + \sigma_N)\int_0^t e^{2\lambda s} ds \\
&+ 2\int_0^t \mathbb{E}\left(\frac{e^{2\lambda s}}{(NM)^2}\sum_{i>j \text{ s.t } Y_s^i = Y_s^j}\frac{1}{|X_s^i - X_s^j|^{\alpha-1}}\right) ds \\
&+ 2\int_0^t \mathbb{E}\left(\frac{e^{2\lambda s}}{(NM)^2}\sum_{i>j \text{ s.t } X_s^i = X_s^j}\frac{1}{|Y_s^i - Y_s^j|^{\alpha-1}}\right) ds \\
\leqslant&\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2\right) + 2(\sigma_M + \sigma_N)\int_0^t e^{2\lambda s} ds + \int_0^t \frac{2M^{\delta(\alpha-1)}}{\alpha(2-\alpha)N^{2-\alpha}}e^{2\lambda s} ds \\
&+ \int_0^t \frac{2N^{\tilde{\delta}(\alpha-1)}}{\alpha(2-\alpha)M^{2-\alpha}}e^{2\lambda s} ds \\
&+ 6\frac{\alpha-1}{\alpha}\frac{1}{M^{1+\delta}}\int_0^t \mathbb{E}\left(\frac{e^{2\lambda s}}{M^2}\sum_{i>j}\frac{i-j}{|\tilde{X}_s^{i,M} - \tilde{X}_s^{j,M}|^\alpha}\right) ds \\
&+ 6\frac{\alpha-1}{\alpha}\frac{1}{N^{1+\tilde{\delta}}}\int_0^t \mathbb{E}\left(\frac{e^{2\lambda s}}{N^2}\sum_{i>j}\frac{i-j}{|\tilde{Y}_s^{i,N} - \tilde{Y}_s^{j,N}|^\alpha}\right) ds,
\end{aligned}$$

and we use (5.2.11) to get

$$\begin{aligned}
\mathbb{E}\left(e^{2\lambda t}\mathcal{W}_2\left(\mu_t^N, \mu_t^M\right)^2\right) \leqslant&\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N, \mu_0^M\right)^2\right) + 2(\sigma_M + \sigma_N)\int_0^t e^{2\lambda s} ds \\
&+ \int_0^t \frac{2M^{\delta(\alpha-1)}}{\alpha(2-\alpha)N^{2-\alpha}}e^{2\lambda s} ds + \int_0^t \frac{2N^{\tilde{\delta}(\alpha-1)}}{\alpha(2-\alpha)M^{2-\alpha}}e^{2\lambda s} ds \\
&+ \frac{3(\alpha-1)}{M^{1+\delta}}\left(\mathbb{E}\left(\mathcal{H}((\tilde{X}_0^{i,M})_i)\right) + Me^{2\lambda t} + 2M\sigma_M\frac{e^{2\lambda t-1}}{2\lambda}\right) \\
&+ \frac{3(\alpha-1)}{N^{1+\tilde{\delta}}}\left(\mathbb{E}\left(\mathcal{H}((\tilde{Y}_0^{j,N})_j)\right) + Ne^{2\lambda t} + 2N\sigma_N\frac{e^{2\lambda t-1}}{2\lambda}\right)
\end{aligned}$$

$$+\frac{6(\alpha-1)}{\alpha(2-\alpha)}\frac{e^{2\lambda t-1}}{2\lambda}\left(\frac{1}{M^{\delta}}+\frac{1}{N^{\tilde{\delta}}}\right).$$

We now choose the coefficients δ and $\tilde{\delta}$. Consider

$$\delta=\frac{2-\alpha}{\alpha}\frac{\ln N}{\ln M}\quad\text{ and }\quad\tilde{\delta}=\frac{2-\alpha}{\alpha}\frac{\ln M}{\ln N}.$$

This way, we have both

$$\frac{M^{\delta(\alpha-1)}}{N^{2-\alpha}}=\frac{e^{\delta(\alpha-1)\ln M}}{N^{2-\alpha}}=\frac{e^{\frac{(2-\alpha)(\alpha-1)}{\alpha}\ln N}}{N^{2-\alpha}}=N^{-\frac{2-\alpha}{\alpha}}\text{ and, }\frac{N^{\tilde{\delta}(\alpha-1)}}{M^{2-\alpha}}=M^{-\frac{2-\alpha}{\alpha}},$$

and

$$M^{-\delta}=M^{-\frac{2-\alpha}{\alpha}\frac{\ln N}{\ln M}}=e^{-\frac{2-\alpha}{\alpha}\ln N}=N^{-\frac{2-\alpha}{\alpha}}\text{ and, likewise, }N^{-\tilde{\delta}}=M^{-\frac{2-\alpha}{\alpha}}.$$

And thus

$$\begin{aligned}\mathbb{E}\left(e^{2\lambda t}\mathcal{W}_2\left(\mu_t^N,\mu_t^M\right)^2\right)\leqslant&\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right)\\&+3(\alpha-1)\left(\frac{1}{N^{\frac{2-\alpha}{\alpha}}}\mathbb{E}\left(\frac{1}{M}\mathcal{H}((\tilde{X}_0^{i,M})_i)\right)+\frac{1}{M^{\frac{2-\alpha}{\alpha}}}\mathbb{E}\left(\frac{1}{N}\mathcal{H}((\tilde{Y}_0^{j,N})_j)\right)\right)\\&+\frac{e^{2\lambda t}-1}{2\lambda}\left(2(\sigma_M+\sigma_N)+\frac{2}{\alpha(2-\alpha)}\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right)\right.\\&\qquad+6(\alpha-1)\left(\frac{\sigma_N}{M^{(2-\alpha)/\alpha}}+\frac{\sigma_M}{N^{(2-\alpha)/\alpha}}\right)\\&\qquad\left.+\frac{6(\alpha-1)}{\alpha(2-\alpha)}\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right)\right)\\&+3(\alpha-1)e^{2\lambda t}\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right).\end{aligned}$$

This yields the result. □

Lemma 5.3.3. *Consider Assumptions 5.1 and 5.4, with $\alpha\in]1,2[$. Let $(\mu^N)_{N\in\mathbb{N}}$ be any sequence of independent empirical measures, such that μ_t^N is the empirical measure of the N particle system at time t. There exist positive constants C_1,C_2 and C_3 such that for all $t\geqslant 0$ and all $N,M\geqslant 1$*

$$\begin{aligned}\mathbb{E}\left(\sup_{s\in[0,t]}\mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right)\leqslant&e^{C_1t}\left(\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right)+3(\alpha-1)\left(\frac{\mathbb{E}^{\mu_0^M}\left(|X|^2\right)}{N^{(2-\alpha)/\alpha}}+\frac{\mathbb{E}^{\mu_0^N}\left(|X|^2\right)}{M^{(2-\alpha)/\alpha}}\right)\right.\\&\left.+C_2(\sigma_N+\sigma_M)+C_3\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right)\right).\end{aligned}\tag{5.3.6}$$

Proof. Consider a similar set up as the proof of Lemma 5.3.1, and define $(\tilde{X}_t^{i,M})_i$, $(\tilde{Y}_t^{j,N})_j$, $(X_t^i)_i$, $(Y_t^j)_j$ in the same manner. With similar calculations, Itô's formula yields

$$\mathcal{W}_2\left(\mu_t^N,\mu_t^M\right)^2\leqslant\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2+\left(2(\sigma_M+\sigma_N)+\frac{2}{\alpha(2-\alpha)}\left(\frac{1}{N^{\frac{2-\alpha}{\alpha}}}+\frac{1}{M^{\frac{2-\alpha}{\alpha}}}\right)\right)t$$

$$
\begin{aligned}
&+\frac{6(\alpha-1)}{\alpha M N^{\frac{2-\alpha}{\alpha}}} \int_0^t \frac{1}{M^2} \sum_{i>j} \frac{i-j}{|\tilde{X}_s^{i,M}-\tilde{X}_s^{j,M}|^\alpha} ds \\
&+\frac{6(\alpha-1)}{\alpha N M^{\frac{2-\alpha}{\alpha}}} \int_0^t \frac{1}{N^2} \sum_{i>j} \frac{i-j}{|\tilde{Y}_s^{i,N}-\tilde{Y}_s^{j,N}|^\alpha} ds \\
&+\frac{2\sqrt{2\sigma_M}}{NM} \sum_i \int_0^t (X_s^i-Y_s^i) dB_s^i - \frac{2\sqrt{2\sigma_N}}{NM} \sum_i \int_0^t (X_s^i-Y_s^i) dB_s'^i.
\end{aligned}
$$

Similarly as Lemma 5.3.1, we use Burkholder-Davis-Gundy inequality to show that there exists a constant C_{BDG} such that

$$
\mathbb{E}\left(\frac{2\sqrt{2\sigma_M}}{NM} \sup_{s\in[0,t]} \sum_i \int_0^s \left(X_u^i-Y_u^i\right) dB_u^i\right) \leqslant 2C_{BDG}\sigma_M + C_{BDG}\mathbb{E}\left(\int_0^t \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2 ds\right),
$$

and

$$
\mathbb{E}\left(\frac{2\sqrt{2\sigma_N}}{NM} \sup_{s\in[0,t]} \sum_i -\int_0^s \left(X_u^i-Y_u^i\right) dB_u'^i\right) \leqslant 2C_{BDG}\sigma_N + C_{BDG}\mathbb{E}\left(\int_0^t \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2 ds\right).
$$

We now use (5.2.12) to obtain

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{s\in[0,t]} \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right) \\
\leqslant &\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right) + \left(2(\sigma_M+\sigma_N) + \frac{2}{\alpha(2-\alpha)}\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right)\right)t \\
&+\frac{6(\alpha-1)}{\alpha M N^{(2-\alpha)/\alpha}}\left(\frac{\alpha}{2}\left(\mathbb{E}\mathcal{H}(\mathbf{X}_0)+M+(2M\sigma_M+M)t\right)+C(\alpha,M)t\right) \\
&+\frac{6(\alpha-1)}{\alpha N M^{(2-\alpha)/\alpha}}\left(\frac{\alpha}{2}\left(\mathbb{E}\mathcal{H}(\mathbf{Y}_0)+N+(2N\sigma_N+N)t\right)+C(\alpha,N)t\right) \\
&+2C_{BDG}(\sigma_N+\sigma_M)+2C_{BDG}\mathbb{E}\left(\int_0^t \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2 ds\right),
\end{aligned}
$$

and thus

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{s\in[0,t]} \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right) \\
\leqslant &\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right) + \frac{3(\alpha-1)}{N^{(2-\alpha)/\alpha}}\mathbb{E}\left(\frac{1}{M}\mathcal{H}(\mathbf{X}_0)\right) + \frac{3(\alpha-1)}{M^{(2-\alpha)/\alpha}}\mathbb{E}\left(\frac{1}{N}\mathcal{H}(\mathbf{Y}_0)\right) \\
&+2C_{BDG}(\sigma_N+\sigma_M)+3(\alpha-1)\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right) \\
&+\Bigg(2(\sigma_M+\sigma_N)+6(\alpha-1)\left(\frac{\sigma_M}{N^{(2-\alpha)/\alpha}}+\frac{\sigma_N}{M^{(2-\alpha)/\alpha}}\right) \\
&\quad+\frac{6\alpha-4}{\alpha(2-\alpha)}\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right) \\
&\quad+3\lambda(\alpha-1)\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right)\Bigg)t
\end{aligned}
$$

$$+ 2C_{BDG}\mathbb{E}\left(\int_0^t \sup_{u\in[0,s]} \mathcal{W}_2\left(\mu_u^N,\mu_u^M\right)^2 ds\right).$$

Denote

$$\begin{aligned}
C_{\text{prop}}(N,M,\alpha) :=& \left(\frac{6\alpha-4}{\alpha(2-\alpha)}+3\lambda(\alpha-1)\right)\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right) \\
&+ 6(\alpha-1)\left(\frac{\sigma_M}{N^{(2-\alpha)/\alpha}}+\frac{\sigma_N}{M^{(2-\alpha)/\alpha}}\right)+2(\sigma_M+\sigma_N), \\
D_{\text{prop}}(N,M,\alpha) :=& \mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N,\mu_0^M\right)^2\right)+3(\alpha-1)\left(\frac{\mathbb{E}\left(\frac{1}{M}\mathcal{H}(\mathbf{X}_0)\right)}{N^{(2-\alpha)/\alpha}}+\frac{\mathbb{E}\left(\frac{1}{N}\mathcal{H}(\mathbf{Y}_0)\right)}{M^{(2-\alpha)/\alpha}}\right) \\
&+ 2C_{BDG}(\sigma_N+\sigma_M)+3(\alpha-1)\left(\frac{1}{N^{(2-\alpha)/\alpha}}+\frac{1}{M^{(2-\alpha)/\alpha}}\right),
\end{aligned}$$

such that, for the sake of conciseness, we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{s\in[0,t]} \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right) \\
&\quad\leqslant D_{\text{prop}}(N,M,\alpha)+2C_{BDG}\mathbb{E}\left(\int_0^t\left(\sup_{u\in[0,s]} \mathcal{W}_2\left(\mu_u^N,\mu_u^M\right)^2+\frac{C_{\text{prop}}(N,M,\alpha)}{2C_{BDG}}\right)ds\right).
\end{aligned}$$

Using Gronwall lemma on $t\mapsto \mathbb{E}\left(\sup_{s\in[0,t]} \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right)+\frac{C_{\text{prop}}(N,M,\alpha)}{2C_{BDG}}$, we get for all $t\geqslant 0$

$$\mathbb{E}\left(\sup_{s\in[0,t]} \mathcal{W}_2\left(\mu_s^N,\mu_s^M\right)^2\right)+\frac{C_{\text{prop}}(N,M,\alpha)}{2C_{BDG}} \leqslant e^{2C_{BDG}t}\left(D_{\text{prop}}(N,M,\alpha)+\frac{C_{\text{prop}}(N,M,\alpha)}{2C_{BDG}}\right).$$

□

5.3.3 Conclusion

We now wish to prove that the Cauchy-like estimates (5.3.1) and (5.3.5) are sufficient to conclude on the convergence, at any given $t>0$, of the empirical measures.

Lemma 5.3.4. *For any sequence $(\mu^n)_{n\in\mathbb{N}}$ of independent random measures in $\mathcal{P}_2(\mathbb{R})$, if*

$$\forall\epsilon>0,\ \exists N\geqslant 0,\ \forall n,m\geqslant N,\ \mathbb{E}\mathcal{W}_2\left(\mu^n,\mu^m\right)\leqslant\epsilon, \tag{5.3.7}$$

then there exists a deterministic measure $\rho\in\mathcal{P}_2(\mathbb{R})$ such that

$$\mathbb{E}\mathcal{W}_2\left(\mu^n,\rho\right)\to 0 \quad \text{as} \quad n\to\infty.$$

Proof. Let us start by mentioning the result of [19], which states that if (X,d) is a complete metric space, then so is $(\mathcal{P}(X),\mathcal{W}_d)$, where $\mathcal{W}_d$ is the Wasserstein distance associated to d. Denote, for ξ and ζ two probability measures on the space $\mathcal{P}_2(\mathbb{R})$ and Γ the set of couplings of ξ and ζ, the Wasserstein distance

$$\mathbb{W}(\xi,\zeta)=\inf_{(\mu,\nu)\sim\Gamma}\mathbb{E}\mathcal{W}_2(\mu,\nu). \tag{5.3.8}$$

The metric space $(\mathcal{P}_1(\mathcal{P}_2(\mathbb{R})), \mathbb{W})$ is complete.

Let ξ^n be the law of μ^n. The assumption (5.3.7) implies, since $\mathbb{W}(\xi^n, \xi^m) \leqslant \mathbb{E}\mathcal{W}_2(\mu^n, \mu^m)$, that there exists a measure $\zeta \in \mathcal{P}(\mathcal{P}_2(\mathbb{R}))$ such that

$$\mathbb{W}(\xi^n, \zeta) \to 0 \quad \text{as} \quad n \to \infty.$$

Denote by π^n the optimal coupling between ξ^n and ζ for the Wasserstein distance above. Considering $\pi^1 \otimes \pi^2 \otimes ...$, there exists a sequence $(\rho^n)_n$, of independent measures identically distributed according to ζ, such that

$$\mathbb{E}\mathcal{W}_2(\mu^n, \rho^n) \to 0 \quad \text{as} \quad n \to \infty.$$

We now wish to prove that all ρ^n are almost surely equal. To do so, we will make use of the assumption of independence of the sequence μ^n. We have

$$\begin{aligned}
&\forall \epsilon > 0,\ \exists N \geqslant 0,\ \forall n \geqslant N,\ \forall p > 0,\ \mathbb{E}\mathcal{W}_2\left(\mu^n, \mu^{n+p}\right) \leqslant \epsilon,\\
&\forall \epsilon > 0,\ \exists N \geqslant 0,\ \forall n \geqslant N,\ \mathbb{E}\mathcal{W}_2(\mu^n, \rho^n) \leqslant \epsilon.
\end{aligned}$$

Direct triangle inequalities using the two assertions above yield

$$\begin{aligned}
&\forall \epsilon > 0,\ \exists N \geqslant 0,\ \forall n \geqslant N,\ \forall p > 0,\ \mathbb{E}\mathcal{W}_2\left(\mu^n, \rho^{n+p}\right) \leqslant \epsilon,\\
&\forall \epsilon > 0,\ \exists N \geqslant 0,\ \forall n \geqslant N,\ \forall p > 0,\ \mathbb{E}\mathcal{W}_2\left(\rho^n, \rho^{n+p}\right) \leqslant \epsilon.
\end{aligned}$$

The fact that $\mathbb{E}\mathcal{W}_2(\mu^n, \rho^n) \to 0$ implies $\mathbb{E}\mathcal{W}_1(\mu^n, \rho^n) \to 0$. The dual formulation of the L^1 Wasserstein distance yields

$$\mathbb{E} \sup_{||\psi||_{Lip} \leqslant 1} |\mu^n(\psi) - \rho^n(\psi)| \to 0.$$

Let f be a bounded Lipschitz continuous function. We have

$$\mathbb{E}\,|\mu^n(f) - \rho^n(f)| \to 0. \tag{5.3.9}$$

In particular we get $\mathbb{E}\mu^n(f) \to \mathbb{E}\rho(f)$, with $\rho \sim \zeta$. Likewise

$$\mathbb{E}\left|\rho^n(f) - \rho^{n+1}(f)\right| \to 0. \tag{5.3.10}$$

On the one hand, using the independence of the sequence,

$$\mathbb{E}\left(\mu^n(f)\mu^{n+1}(f)\right) = \mathbb{E}\left(\mu^n(f)\right)\mathbb{E}\left(\mu^{n+1}(f)\right) \to \mathbb{E}\left(\rho(f)\right)^2.$$

On the other hand

$$\begin{aligned}
\mathbb{E}\left(\mu^n(f)\mu^{n+1}(f)\right) = \mathbb{E}\Big((\rho^n(f))^2 &+ \rho^n(f)\left(\mu^n(f) - \rho^n(f)\right) + \mu^n(f)\left(\mu^{n+1}(f) - \rho^{n+1}(f)\right)\\
&+ \mu^n(f)\left(\rho^{n+1}(f) - \rho^n(f)\right)\Big).
\end{aligned}$$

Let us consider each term individually

$$\begin{aligned}
\mathbb{E}\left((\rho^n(f))^2\right) &= \mathbb{E}\left((\rho(f))^2\right),\\
\mathbb{E}\left(\rho^n(f)\left(\mu^n(f) - \rho^n(f)\right)\right) &\leqslant ||f||_\infty \mathbb{E}\,|\mu^n(f) - \rho^n(f)| \to 0 \text{ using (5.3.9)},\\
\mathbb{E}\left(\mu^n(f)\left(\mu^{n+1}(f) - \rho^{n+1}(f)\right)\right) &\leqslant ||f||_\infty \mathbb{E}\left|\mu^{n+1}(f) - \rho^{n+1}(f)\right| \to 0 \text{ using (5.3.9)},
\end{aligned}$$

$$\mathbb{E}\left(\mu^n(f)\left(\rho^{n+1}(f)-\rho^n(f)\right)\right) \leqslant ||f||_\infty \mathbb{E}\left|\rho^{n+1}(f)-\rho^n(f)\right| \to 0 \text{ using (5.3.10).}$$

Thus

$$\mathbb{E}\left(\mu^n(f)\mu^{n+1}(f)\right) \to \mathbb{E}\left((\rho(f))^2\right).$$

We have obtained

$$\mathbb{E}\left(\rho(f)\right)^2 = \mathbb{E}\left((\rho(f))^2\right),$$

which implies that for any bounded and Lipschitz continuous function f, $\rho(f)$ is almost surely constant. Let ρ_1 and ρ_2 be two random variables with law ζ, considering two random variables $X \sim \rho_1$ and $Y \sim \rho_2$, we get for all Lipschitz continuous bounded function h that $\mathbb{E}h(X) = \mathbb{E}h(Y)$. Let $a < b$ be two real numbers. Consider

$$g_m(x) = \begin{cases} 1 & \text{if } x \in [a+\frac{1}{m}, b-\frac{1}{m}] \\ 0 & \text{if } x \leqslant a \text{ or } x \geqslant b \\ m(x-a) & \text{if } a < x < a+\frac{1}{m} \\ m(b-x) & \text{if } b-\frac{1}{m} < x < b \end{cases}$$

By construction, $(g_m)_{m\in\mathbb{N}}$ is an increasing sequence of bounded Lipschitz continuous functions such that for all m, $g_m \leqslant \mathbb{1}_{]a,b[}$ and for all $x \in \mathbb{R}$, $g_m(x) \to \mathbb{1}_{]a,b[}(x)$ as $m \to \infty$. We thus have for all $m \in \mathbb{N}$ the equality $\mathbb{E}g_m(X) = \mathbb{E}g_m(Y)$ and, by the monotone convergence theorem, $\mathbb{E}\mathbb{1}_{]a,b[}(X) = \mathbb{E}\mathbb{1}_{]a,b[}(Y)$. Then, again by the monotone convergence theorem and by considering an increasing sequence of simple functions, the equality $\mathbb{E}h(X) = \mathbb{E}h(Y)$ holds true for all bounded measurable function h. The variables X and Y are thus equal in law, and ρ is therefore a deterministic probability measure.

□

Lemma 5.3.4 allows us to conclude on the convergence, at any given $t \geqslant 0$, of the sequence of empirical measures $(\mu_t^N)_N$ towards $\bar{\rho}_t$ where, at least formally, $\bar{\rho}$ is a solution of the non linear limit equation (we refer to the next Section 5.4 for a more rigorous identification of the equation satisfied by the limit $\bar{\rho}$). However, *a priori*, if the limit equation admits several solutions, nothing guarantees that the sequence converges towards the *same* solution at two different times t_1 and t_2. To show this, we now use the estimates (5.3.2) and (5.3.6) which, even though on their own do not ensure uniform in time convergence because of the exponential term, show that on any time interval $[0,T]$ there is uniform convergence towards a unique solution of the limit equation. This result, combined with the uniform in time pointwise convergence given by Lemma 5.3.4, will yield the desired result. Denote $\mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}))$ the space of continuous functions taking values in the space of probability measures $\mathcal{P}_2(\mathbb{R})$ endowed with the L^2 Wasserstein distance.

Lemma 5.3.5. *Let $T \geqslant 0$. For any sequence $(\mu^n)_{n\in\mathbb{N}}$ of independent random variables in $\mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}))$, if*

$$\forall\epsilon > 0,\ \exists N \geqslant 0,\ \forall n,m \geqslant N,\ \mathbb{E}\sup_{t\in[0,T]} \mathcal{W}_2\left(\mu_t^n,\mu_t^m\right) \leqslant \epsilon, \tag{5.3.11}$$

then there exists a deterministic measure $(\rho_t)_{t\in[0,T]} \in \mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}))$ such that

$$\mathbb{E}\sup_{t\in[0,T]} \mathcal{W}_2\left(\mu_t^n,\rho_t\right) \to 0 \quad \textit{as} \quad n\to\infty.$$

Proof. Let, for ξ and ζ two probability measures on the space $\mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}))$ and Γ the set of couplings of ξ and ζ, the Wasserstein distance be defined by

$$\mathbb{W}_s(\xi,\zeta) = \inf_{(\mu,\nu)\sim\Gamma} \mathbb{E}\left(\sup_{t\in[0,T]} \mathcal{W}_2(\mu_t,\nu_t)\right). \tag{5.3.12}$$

Let ξ^n be the law of μ^n. By completeness, Assumption (5.3.11) implies that there exists a probability measure ζ on $\mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}))$ such that

$$\mathbb{W}_s(\xi^n,\zeta) \to 0 \quad \text{as} \quad n\to\infty.$$

Thus, there exists a sequence $(\rho^n)_n$, identically distributed according to ζ, such that

$$\mathbb{E} \sup_{t\in[0,T]} \mathcal{W}_2\left(\mu_t^n,\rho_t^n\right) \to 0 \quad \text{as} \quad n\to\infty.$$

By the same proof as Lemma 5.3.4, we get that for all $t \geqslant 0$, all ρ_t^n are almost surely equal, hence the result. □

Remark 5.3.2. ***Addendum :*** *In this section we thus have shown that the sequence of the laws $(\xi_t^N)_N$ of the empirical measures $(\mu_t^N)_N$ is a Cauchy sequence in the space of probability measures on the space of probability measures endowed with its Wasserstein distance. This is the way the result has been proved and published, and the way the authors find the most simple.*

However, as it was suggested to us, given the space $(\Omega,\mathbb{P})$ on which the random variable $(\mu_t^N)_N$ are realised, μ_t^N can be seen as an element in $L^2(\Omega,\mathcal{P}_2(\mathbb{R}))$, a space which is complete when endowed by the distance $d(\mu,\nu) = \mathbb{E}\left(\mathcal{W}_2^2(\mu,\nu)\right)^{1/2}$ (which is the L^2 distance associated to the Wasserstein distance). Therefore there exists a random measure ρ_t defined on the same probability space such that $\mathbb{E}\mathcal{W}_2(\mu_t^N,\rho_t)\to 0$, and since ρ_t is independent of each of the μ_t^N, it is deterministic.

5.4 Identification of the limit

The goal of this subsection is to identify, in a more rigorous way than the formal calculations of the introduction, the limit $\bar\rho_t$, and more precisely the PDE it satisfies. We prove the following theorem

Theorem 5.4.1. *For $\alpha \in]1,2[$ under Assumptions 5.1 and 5.4, or for $\alpha = 1$ under Assumptions 5.1 and 5.2, both with $\sigma_N \to 0$, the limit $(\bar\rho_t)_{t\geqslant 0}$, obtained in Theorem 5.3.1, of the sequence of empirical measures $((\mu_t^N)_{t\geqslant 0})_{N\geqslant 2}$ satisfies, for all functions $f \in \mathcal{C}^2(\mathbb{R})$ with bounded derivatives such that f, f', $f'U'$, and f'' are Lipschitz continuous and that $f'U'$ is bounded, the following equation, for all $t \geqslant 0$,*

$$\begin{aligned}\int_{\mathbb{R}} f(x)\bar\rho_t(dx) = \int_{\mathbb{R}} f(x)\bar\rho_0(dx) - \int_0^t \int_{\mathbb{R}} f'(x)U'(x)\bar\rho_s(dx)ds \\ + \frac{1}{2}\int_0^t \int\int_{\{x\neq y\}} \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}} \bar\rho_s(dx)\bar\rho_s(dy)ds.\end{aligned}$$

The proof of the theorem above consists in rigorously applying the dominated convergence theorem in the PDE satisfied by the empirical measure.

To do so, let us first mention the following lemma, which is a consequence of previous calculations.

Lemma 5.4.1. *For $\alpha \in]1,2[$ and under Assumptions 5.1 and 5.4, for all $t \geqslant 0$, there exists a constant C_{int} such that for all $N \geqslant 2$, we have the following estimates*

$$\mathbb{E}\left(\int_0^t \int \int_{\{x\neq y\}} \frac{1}{|x-y|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}} \mu_s^N(dx)\mu_s^N(dy)ds\right) \leqslant C_{int}. \tag{5.4.1}$$

Proof. Let $(X_t^1, ..., X_t^N)_t$ be the unique strong solution of (5.1.1), and μ_t^N the associated empirical measure. By definition

$$\mathbb{E}\left(\int_0^t \int \int_{\{x\neq y\}} \frac{1}{|x-y|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}} \mu_s^N(dx)\mu_s^N(dy)ds\right) = 2\mathbb{E}\left(\int_0^t \frac{1}{N^2}\sum_{i>j} \frac{1}{|X_s^i - X_s^j|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}} ds\right).$$

Young's inequality yields, for $i > j$, for $\beta > 0$, $\gamma > 0$, and $p > 1$ and $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$

$$\frac{1}{|X_t^i - X_t^j|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}} \leqslant \frac{\gamma^p}{p}\left(\frac{(i-j)^\beta}{|X_t^i - X_t^j|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}}\right)^p + \frac{1}{q\gamma^q}\frac{1}{(i-j)^{\beta q}}.$$

We choose

$$\beta = \frac{\alpha-1}{\alpha}, \qquad p = \frac{\alpha}{\alpha-1}, \qquad q = \alpha, \qquad \gamma = N^{-\frac{\alpha-1}{\alpha}}.$$

which yields

$$\frac{1}{|X_t^i - X_t^j|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}} \leqslant \frac{\alpha-1}{\alpha}\frac{1}{N}\frac{i-j}{|X_t^i - X_t^j|^{\frac{\alpha+2}{2}}} + \frac{N^{\alpha-1}}{\alpha}\frac{1}{(i-j)^{\alpha-1}}.$$

From Lemma D.1.1, we have

$$\sum_{i>j}\frac{1}{(i-j)^{\alpha-1}} \leqslant \frac{N}{2-\alpha}N^{2-\alpha}.$$

Thus

$$\begin{aligned}\int_0^t \frac{1}{N^2}\sum_{i>j}\frac{1}{|X_s^i - X_s^j|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}}ds \leqslant & \frac{\alpha-1}{\alpha}\frac{1}{N}\left(\int_0^t \frac{1}{N^2}\sum_{i>j}\frac{i-j}{|X_s^i - X_s^j|^{\frac{\alpha+2}{2}}}ds\right) \\ & + \frac{1}{N^2}\frac{N^{\alpha-1}}{\alpha}\frac{N}{2-\alpha}N^{2-\alpha}t.\end{aligned}$$

This yields the result using (5.2.12), as $\frac{\alpha+2}{2} \in]1,2[$, and noticing that $\frac{1}{N}\mathbb{E}\mathcal{H}(X_0)$ is bounded from above by the initial second moment, and is thus bounded uniformly in N. ☐

Proof of Theorem 5.4.1. As $\mathbb{E}\mathcal{W}_1(\mu_t^N, \bar{\rho}_t) \to 0$, we get by the dual formulation of the Wasserstein distance that for all function g Lipshitz continuous

$$\mathbb{E}\int_{\mathbb{R}} g(x)\mu_t^N(dx) \to \int_{\mathbb{R}} g(x)\bar{\rho}_t(dx).$$

Likewise, since

$$
\begin{aligned}
\mathcal{W}_1(\mu_t^N \otimes \mu_t^N, \bar\rho_t \otimes \bar\rho_t) \leqslant & \mathcal{W}_1(\mu_t^N \otimes \mu_t^N, \bar\rho_t \otimes \mu_t^N) + \mathcal{W}_1(\bar\rho_t \otimes \mu_t^N, \bar\rho_t \otimes \bar\rho_t) \\
= & \inf_{\substack{(X^1,X^2)\sim \mu_t^N \otimes \mu_t^N \\ (Y^1,Y^2)\sim \bar\rho_t \otimes \mu_t^N}} \mathbb{E}\left(|X^1 - Y^1| + |X^2 - Y^2|\right) \\
& + \inf_{\substack{(X^1,X^2)\sim \bar\rho_t \otimes \mu_t^N \\ (Y^1,Y^2)\sim \bar\rho_t \otimes \bar\rho_t}} \mathbb{E}\left(|X^1 - Y^1| + |X^2 - Y^2|\right) \\
\leqslant & \inf_{X^1\sim\mu_t^N, Y^1\sim\bar\rho_t} \mathbb{E}\left(|X^1 - Y^1|\right) + \inf_{X^2\sim\mu_t^N, Y^2\sim\bar\rho_t} \mathbb{E}\left(|X^2 - Y^2|\right) \\
= & 2\mathcal{W}_1(\mu_t^N, \bar\rho_t),
\end{aligned}
$$

we get that for all function g Lipshitz continuous

$$
\mathbb{E}\int_{\mathbb{R}}\int_{\mathbb{R}} g(x,y)\mu_t^N(dx)\mu_t^N(dy) \to \int_{\mathbb{R}}\int_{\mathbb{R}} g(x,y)\bar\rho_t(dx)\bar\rho_t(dy).
$$

Let us now consider a function $f \in \mathcal{C}^2(\mathbb{R})$ with bounded derivatives such that f, f', $f'U'$ and f'' are Lipschitz continuous and $f'U'$ is bounded. By Itô's formula, we have

$$
\begin{aligned}
\int_{\mathbb{R}} f(x)\mu_t^N(dx) = & \int_{\mathbb{R}} f(x)\mu_0^N(dx) - \int_0^t\int_{\mathbb{R}} f'(x)U'(x)\mu_s^N(dx)ds \\
& + \int_0^t\int_{\mathbb{R}} \sigma_N f''(x)\mu_s^N(dx)ds + \int_0^t \frac{\sqrt{2\sigma_N}}{N}\sum_{i=1}^N f'(X_s^i)dB_s^i \\
& + \frac{1}{2}\int_0^t\int\int_{\{x\neq y\}} \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}\mu_s^N(dx)\mu_s^N(dy)ds \\
:= & I_0(N) - I_1(N) + I_2(N) + I_3(N) + I_4(N).
\end{aligned}
$$

Let us deal with each terms.

- **$I_0(N)$:** Since we assume f to be Lipschitz continuous

$$
\mathbb{E}I_0(N) = \mathbb{E}\int_{\mathbb{R}} f(x)\mu_0^N(dx) \longrightarrow \int_{\mathbb{R}} f(x)\bar\rho_0(dx).
$$

- **$I_1(N)$:** $f'U'$ being Lipschitz continuous, we have

$$
\mathbb{E}\int_{\mathbb{R}} f'(x)U'(x)\mu_s^N(dx) \to \int_{\mathbb{R}} f'(x)U'(x)\bar\rho_s(dx).
$$

  Furthermore, since $f'U'$ is bounded, $\left|\int_{\mathbb{R}} f'(x)U'(x)\mu_s^N(dx)\right| \leqslant ||f'U'||_\infty$ and we have by dominated convergence

$$
\begin{aligned}
\mathbb{E}I_1(N) = \mathbb{E}\int_0^t\int_{\mathbb{R}} f'(x)U'(x)\mu_s^N(dx)ds = & \int_0^t \mathbb{E}\int_{\mathbb{R}} f'(x)U'(x)\mu_s^N(dx)ds \\
& \longrightarrow \int_0^t\int_{\mathbb{R}} f'(x)U'(x)\bar\rho_s(dx)ds
\end{aligned}
$$

- $I_2(N)$ **:** Since we assume f'' to be Lipschitz continous

$$\mathbb{E}I_2(N) = \mathbb{E}\int_0^t \int_{\mathbb{R}} \sigma_N f''(x)\mu_s^N(dx)ds = \int_0^t \sigma_N \mathbb{E}\left(\int_{\mathbb{R}} f''(x)\mu_s^N(dx)\right) ds$$
$$\longrightarrow 0 \quad \text{(by dominated convergence).}$$

- $I_3(N)$ **:** As f' is bounded, $I_3(N)$ is a true martingale, and thus $\mathbb{E}I_3(N) = 0$.

- $I_4(N)$ **:** Let, for $R > 0$, ϕ_R be a Lipshitz continous function such that

$$\phi_R(x) = \begin{cases} 1 & \text{if } x \leqslant R \\ \frac{2R-x}{R} & \text{if } R \leqslant x \leqslant 2R \\ 0 & \text{if } x \geqslant 2R. \end{cases}$$

We have

$$\begin{aligned} &\int_0^t \int\int_{\{x\neq y\}} \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}} \mu_s^N(dx)\mu_s^N(dy)ds \\ &= \int_0^t \int\int_{\{x\neq y\}} \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}} \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds \\ &+ \int_0^t \int\int \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}} (1-\phi_R(|x-y|))\mu_s^N(dx)\mu_s^N(dy)ds. \end{aligned} \tag{5.4.2}$$

Let us now find the limit as R goes to 0 of the limit as N goes to infinity of the expectation of the first term of (5.4.2). By Hölder's inequality

$$\begin{aligned} &\mathbb{E}\int_0^t \int\int_{\{x\neq y\}} \left| \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}} \phi_R(|x-y|)\right| \mu_s^N(dx)\mu_s^N(dy)ds \\ &\leqslant ||f''||_\infty \mathbb{E}\left(\int_0^t \int\int_{\{x\neq y\}} \frac{1}{|x-y|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}} \mu_s^N(dx)\mu_s^N(dy)ds\right)^{\frac{2\alpha}{\alpha+2}} \\ &\quad \times \mathbb{E}\left(\int_0^t \int\int \phi_R(|x-y|)^{\frac{\alpha+2}{2-\alpha}} \mu_s^N(dx)\mu_s^N(dy)ds\right)^{\frac{2-\alpha}{\alpha+2}}, \end{aligned}$$

and since $0 \leqslant \phi_R \leqslant 1$,

$$\begin{aligned} &\mathbb{E}\left(\int_0^t \int\int \phi_R(|x-y|)^{\frac{\alpha+2}{2-\alpha}} \mu_s^N(dx)\mu_s^N(dy)ds\right)^{\frac{2-\alpha}{\alpha+2}} \\ &\qquad \leqslant \mathbb{E}\left(\int_0^t \int\int \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds\right)^{\frac{2-\alpha}{\alpha+2}}. \end{aligned}$$

We now use (5.4.1) to get

$$\mathbb{E}\int_0^t \int\int_{\{x\neq y\}} \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}} \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds$$

$$\leqslant ||f''||_\infty C_{int}^{\frac{2\alpha}{\alpha+2}} \mathbb{E}\left(\int_0^t \int\int \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds\right)^{\frac{2-\alpha}{\alpha+2}}. \tag{5.4.3}$$

We then use

$$\begin{aligned}\int_0^t \int\int \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds \leqslant &\int_0^t \int\int_{\{x=y\}} \mu_s^N(dx)\mu_s^N(dy)ds \\ &+ \int_0^t \int\int_{\{x\neq y\}} \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds.\end{aligned}$$

First

$$\int_0^t \int\int_{\{x=y\}} \mu_s^N(dx)\mu_s^N(dy)ds = \frac{t}{N}.$$

Then $\phi_R(|x|) \leqslant \mathbb{1}_{|x|\leqslant 2R} \leqslant \left(\frac{2R}{|x|}\right)^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}$, which implies

$$\begin{aligned}&\mathbb{E}\left(\int_0^t \int\int_{\{x\neq y\}} \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds\right) \\ &\leqslant (2R)^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}} \mathbb{E}\left(\int_0^t \int\int_{\{x\neq y\}} \frac{1}{|x-y|^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}}} \mu_s^N(dx)\mu_s^N(dy)ds\right) \\ &\leqslant (2R)^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}} C_{int}.\end{aligned}$$

Thus,

$$\mathbb{E}\int_0^t \int\int \phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds \leqslant \frac{t}{N} + (2R)^{\frac{(\alpha-1)(\alpha+2)}{2\alpha}} C_{int}. \tag{5.4.4}$$

Thus, for the first term of (5.4.2), using (5.4.3) and (5.4.4), taking the limit as $N\to\infty$ and then as $R\to 0$ yields

$$\lim_{R\to 0}\lim_{N\to\infty} \mathbb{E}\int_0^t \int\int_{\{x\neq y\}} \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}\phi_R(|x-y|)\mu_s^N(dx)\mu_s^N(dy)ds = 0. \tag{5.4.5}$$

Let us find the limit as R goes to 0 of the limit as N goes to infinity of the expectation of the second term of (5.4.2). Since $\frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}(1-\phi_R(|x-y|))$ is bounded and Lipschitz continous, we have

$$\begin{aligned}&\mathbb{E}\int_0^t \int\int \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}(1-\phi_R(|x-y|))\mu_s^N(dx)\mu_s^N(dy)ds \\ &\longrightarrow \int_0^t \int\int \frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}(1-\phi_R(|x-y|))\bar\rho_s(dx)\bar\rho_s(dy)ds.\end{aligned}$$

We now want to use dominated convergence to consider the limit as R goes to 0. We have

$$\left|\frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}(1-\phi_R(|x-y|))\right| \leqslant ||f''||_\infty \frac{\mathbb{1}_{x\neq y}}{|x-y|^{\alpha-1}}.$$

Let us show that $\int_0^t\int\int \frac{\mathbb{1}_{x\neq y}}{|x-y|^{\alpha-1}}\bar{\rho}_s(dx)\bar{\rho}_s(dy)ds<\infty$. Using (5.4.1), and Young's inequality as $\alpha-1\leqslant\frac{(\alpha-1)(\alpha+2)}{2\alpha}$, we get

$$\begin{aligned}\mathbb{E}\int_0^t\int\int\frac{1-\phi_R(|x-y|)}{|x-y|^{\alpha-1}}\mu_s^N(dx)\mu_s^N(dy)ds &\leqslant \mathbb{E}\int_0^t\int\int\frac{\mathbb{1}_{x\neq y}}{|x-y|^{\alpha-1}}\mu_s^N(dx)\mu_s^N(dy)ds\\ &\leqslant \tilde{C}_{int},\end{aligned}$$

where $\tilde{C}_{int}$ is a constant independent of N (depending on C_{int}). The righthand side being independent of N and R, and since $\frac{1-\phi_R(|x-y|)}{|x-y|^{\alpha-1}}$ is bounded and Lipschtiz continous, we have taking the limit as $N\to\infty$

$$\mathbb{E}\int_0^t\int\int\frac{1-\phi_R(|x-y|)}{|x-y|^{\alpha-1}}\bar{\rho}_s(dx)\bar{\rho}_s(dy)ds\leqslant\tilde{C}_{int},$$

and by monotone convergence theorem

$$\mathbb{E}\int_0^t\int\int\frac{\mathbb{1}_{x\neq y}}{|x-y|^{\alpha-1}}\bar{\rho}_s(dx)\bar{\rho}_s(dy)ds\leqslant\tilde{C}_{int}.$$

This implies

$$\begin{aligned}\lim_{R\to 0}\lim_{N\to\infty}\mathbb{E}\int_0^t\int\int&\frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}(1-\phi_R(|x-y|))\mu_s^N(dx)\mu_s^N(dy)ds\\ &=\int_0^t\int\int_{\{x\neq y\}}\frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}\bar{\rho}_s(dx)\bar{\rho}_s(dy)ds. \end{aligned}\tag{5.4.6}$$

From (5.4.5) and (5.4.6), we obtain

$$\mathbb{E}I_4(N)\longrightarrow\frac{1}{2}\int_0^t\int\int_{\{x\neq y\}}\frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}\bar{\rho}_s(dx)\bar{\rho}_s(dy)ds.$$

Hence the result.

□

Remark 5.4.1. *Notice how above we rely on the fact that* $\frac{(f'(x)-f'(y))(x-y)}{|x-y|^{\alpha+1}}\mathbb{1}_{x\neq y}$ *is integrable with respect to* $\bar{\rho}_t\otimes\bar{\rho}_t$ *for a Lipschitz continuous function* f'*. This amounts to being able to prove*

$$\int\int_{\{x\neq y\}}\frac{1}{|x-y|^{\alpha-1}}\bar{\rho}_t(dx)\bar{\rho}_t(dy)<\infty.$$

For the sake of the argument, let us assume that $\bar{\rho}_t=\mathbb{1}_{[0,1]}$ *is the uniform distribution on* $[0,1]$*. Then*

$$\int\int_{[0,1]\times[0,1]}\frac{1}{|x-y|^{\alpha-1}}dxdy<\infty,$$

if and only if $\alpha < 2$. Although this is no proof, this small estimate seems to indicate that $\alpha = 2$ is indeed a critical value.

5.5 From weak propagation of chaos to strong uniform in time propagation of chaos

In this section, which is somewhat independent of the previous ones, we wish to show how one could improve a result of weak propagation of chaos, as for instance obtained in [151], [43] or [125], into a result of strong and uniform in time propagation of chaos. We consider (5.1.1) for any potentials U and V and any diffusion σ_N, and assume there is a strong solution $(X_t^i)_t$ of (5.1.1). In this general framework, we assume one has been able to prove the following assertions :

Assumption 5.5. *[Weak prop. of chaos] For an initial distribution μ_0 converging in L^2 Wasserstein distance to a measure $\bar{\rho}_0$, and for all $t \geqslant 0$, the empirical measure $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$ converges weakly to a probability density $\bar{\rho}_t$.*

Assumption 5.6. *[Bounded moments] Assume there is $C_0 \geqslant 0$ such that for all N and all $t \geqslant 0$*

$$\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i|^4\right) \leqslant C_0.$$

Assumption 5.7. *[Long time convergence] Denoting by $\rho_t^{1,N}$ and $\rho_t^{2,N}$ the probability densities of the N particle systems in $\mathcal{O}_N$ with respective initial conditions $\rho_0^{1,N}$ and $\rho_0^{2,N}$, there exists $\lambda > 0$ such that we have*

$$\forall t \geqslant 0, \mathcal{W}_2\left(\rho_t^{1,N}, \rho_t^{2,N}\right) \leqslant e^{-\lambda t}\mathcal{W}_2\left(\rho_0^{1,N}, \rho_0^{2,N}\right)$$

.

Assumption 5.8. *[Continuity in 0] The function $t \mapsto \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - X_0^i|^2\right)$ is continuous in $t = 0$, uniformly in N, in the sense that*

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall 0 \leqslant t < \delta,\ \forall N \geqslant 0,\ \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - X_0^i|^2\right) \leqslant \epsilon$$

Remark 5.5.1. *These assumptions are satisfied in the case $\alpha = 1$. We have shown in Theorem 5.2.2 the long time convergence of the particle system. In Appendix D.3 we prove continuity in 0 for a well chosen initial condition. To prove the bounded 4-th moments, considering $\phi : (x_1, .., x_N) \mapsto \frac{1}{N}\sum_{i=1}^N |x_i|^4$, we have*

$$\begin{aligned}\mathcal{L}^{N,\alpha}\phi &= -\sum_{i=1}^N U'(x_i)\left(\frac{4}{N}x_i^3\right) + \sum_{i=1}^N\left(\frac{1}{N}\sum_{j\neq i}^N \frac{1}{x_i - x_j}\right)\left(\frac{4x_i^3}{N}\right) + \sigma_N\sum_{i=1}^N \frac{12}{N}x_i^2 \\ &\leqslant 12\sigma_N\left(\frac{1}{N}\sum_{i=1}^N x_i^2\right) - \frac{4\lambda}{N}\sum_{i=1}^N |x_i|^4 + \frac{4}{N^2}\sum_{i\neq j}\frac{x_i^3}{x_i - x_j}.\end{aligned}$$

We get

$$\frac{4}{N^2}\sum_{i\neq j}\frac{x_i^3}{x_i-x_j}=\frac{4}{N^2}\sum_{j<i}\frac{x_i^3-x_j^3}{x_i-x_j}=\frac{4}{N^2}\sum_{j<i}x_i^2+x_ix_j+x_j^2$$
$$\leqslant \frac{6}{N^2}\sum_{j<i}x_i^2+x_j^2\leqslant\frac{6}{N}\sum_{i=1}^N x_i^2.$$

This way

$$\begin{aligned}\mathcal{L}^{N,\alpha}\phi \leqslant & 6\left(2\sigma_N+1\right)\left(\frac{1}{N}\sum_{i=1}^N x_i^2\right)-\frac{4\lambda}{N}\sum_{i=1}^N|x_i|^4\\ \leqslant & \frac{9\left(2\sigma_N+1\right)^2}{2\lambda}-\frac{2\lambda}{N}\sum_{i=1}^N|x_i|^4 \qquad \textit{since} \qquad x_i^2\leqslant\frac{\lambda x_i^4}{3(1+2\sigma_N)}+\frac{3(1+2\sigma_N)}{4\lambda}\\ = & \frac{9\left(2\sigma_N+1\right)^2}{2\lambda}-2\lambda\phi.\end{aligned}$$

We thus obtain the uniform in time bound on the 4-th moment provided we have an initial bound.

Our goal is to show

Theorem 5.5.1. *Under Assumptions 5.5, 5.6, 5.7 and 5.8, we get strong uniform in time propagation of chaos, i.e*

$$\forall\epsilon>0,\ \exists N\geqslant 0,\ \forall t\geqslant 0,\ \forall n\geqslant N,\ \mathbb{E}\left(\mathcal{W}_2\left(\mu_t^n,\bar\rho_t\right)\right)<\epsilon.$$

The outline of the proof is the following

- Using the weak propagation of chaos and the bounded moments, we get a strong convergence in Wasserstein distance.
- Using the long time convergence of the particle system and the strong propagation of chaos, we get the long time convergence for the limiting process, as well as strong propagation of chaos for the stationnary measures.
- Thanks to the long time convergence of both the particle system and the limiting process, and using the continuity in 0 of the particle system for the Wasserstein distance, we get uniform continuity in time for the Wasserstein distance between the empirical measure and the limiting process, this continuity being uniform in N.
- Finally, thanks to all the previous results, we get uniform in time propagation of chaos.

The following result is the characterization of the $\mathcal{W}_2$-convergence, as given in Theorem 6.9 of [171].

Lemma 5.5.1. *[Strong propagation of chaos] Under Assumptions 5.5 and 5.6, we also have the following convergence*

$$\forall t\geqslant 0,\ \lim_{N\to\infty}\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N,\bar\rho_t\right)^2\right)=0 \tag{5.5.1}$$

Remark 5.5.2. *We use here the assumption on the bounded 4-th moment of the empirical measure, to have by Cauchy-Schwarz inequality for* $R > 0$

$$\frac{1}{N}\sum_{i=1}^{N}\left|X^i\right|^2 \mathbb{1}_{|X^i|\geqslant \frac{R}{2}} \leqslant \left(\frac{1}{N}\sum_{i=1}^{N}\left|X^i\right|^4\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{|X^i|\geqslant \frac{R}{2}}\right)^{1/2}$$

$$\begin{aligned}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\left|X^i\right|^2 \mathbb{1}_{|X^i|\geqslant \frac{R}{2}}\right) &\leqslant \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\left|X^i\right|^4\right)^{1/2}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{|X^i|\geqslant \frac{R}{2}}\right)^{1/2}\\ &\leqslant C_0^{1/2}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{|X^i|\geqslant \frac{R}{2}}\right)^{1/2}.\end{aligned}$$

We have, by weak convergence since $x \mapsto \mathbb{1}_{|x|\geqslant \frac{R}{2}}$ *is a bounded upper semi continuous function*

$$\limsup_{N\to\infty}\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\mu_t^N \leqslant \int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\bar{\rho}_t.$$

Then, since $\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\mu_t^N$ *is a sequence of positive functions such that* $\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\mu_t^N \leqslant 1$, *we have by Fatou's lemma*

$$\limsup_{N\to\infty}\mathbb{E}\left(\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\mu_t^N\right) \leqslant \mathbb{E}\left(\limsup_{N\to\infty}\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\mu_t^N\right)$$

and by dominated convergence

$$\lim_{R\to\infty}\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\bar{\rho}_t = 0.$$

Therefore

$$\begin{aligned}\lim_{R\to\infty}\limsup_{N\to\infty}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{|X^i|\geqslant \frac{R}{2}}\right) &\leqslant \lim_{R\to\infty}\mathbb{E}\left(\limsup_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{|X^i|\geqslant \frac{R}{2}}\right)\\ &\leqslant \lim_{R\to\infty}\mathbb{E}\left(\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\bar{\rho}_t\right)\\ &= \lim_{R\to\infty}\int \mathbb{1}_{|x|\geqslant \frac{R}{2}} d\bar{\rho}_t\\ &=0.\end{aligned}$$

This yields the necessary property to use Theorem 6.9 of [171]. In reality, any assumption on a bounded p-th moment with $p > 2$ *would have been sufficient, using Hölder's inequality instead of Cauchy-Schwarz's.*

Lemma 5.5.2 (Long time behavior of the limiting equation)**.** *Under Assumptions 5.5, 5.6 and 5.7, consider* μ_t^N *(resp.* ν_t^N*) the empirical distribution of the solution* $(X_t^i)_t$ *(with* $(X_t^i)_i \in \mathcal{O}_N$*) weakly converging as* N *goes to infinity to* $\bar{\mu}_t$ *(resp.* $\bar{\nu}_t$*). We have*

$$\mathcal{W}_2\left(\bar{\mu}_t, \bar{\nu}_t\right) \leqslant e^{-\lambda t}\mathcal{W}_2\left(\bar{\mu}_0, \bar{\nu}_0\right).$$

Proof. Denoting $\rho_t^{1,N}$ (resp. $\rho_t^{2,N}$) the law in $\mathcal{O}_N$ the law of the N particle system which yields μ_t^N (resp. ν_t^N). We have, under π_t^N the optimal coupling between $\rho_t^{1,N}$ and $\rho_t^{2,N}$ for the L^2

Wasserstein distance.

$$\mathcal{W}_2\left(\bar{\mu}_t, \bar{\nu}_t\right) \leqslant \mathbb{E}^{\pi_t^N}\left(\mathcal{W}_2\left(\bar{\mu}_t, \mu_t^N\right) + \mathcal{W}_2\left(\mu_t^N, \nu_t^N\right) + \mathcal{W}_2\left(\nu_t^N, \bar{\nu}_t\right)\right).$$

Since

$$\begin{aligned}\mathbb{E}^{\pi_t^N}\left(\mathcal{W}_2\left(\mu_t^N, \nu_t^N\right)\right) \leqslant \mathbb{E}^{\pi_t^N}\left(\mathcal{W}_2\left(\mu_t^N, \nu_t^N\right)^2\right)^{1/2} &= \mathbb{E}^{\pi_t^N}\left(\frac{1}{N}\sum_{i=1}^N (X_t^i - Y_t^i)^2\right)^{1/2} \\ &= \frac{1}{\sqrt{N}}\mathcal{W}_2(\rho_t^{1,N}, \rho_t^{2,N}),\end{aligned}$$

and by Assumption 5.7 (recall that we assume $\rho_t^{1,N}$ is the law in $\mathcal{O}_N$ of the particle system)

$$\mathcal{W}_2(\rho_t^{1,N}, \rho_t^{2,N}) \leqslant e^{-\lambda t}\mathcal{W}_2(\rho_0^{1,N}, \rho_0^{2,N}) \leqslant e^{-\lambda t}\mathbb{E}\left(\sum_{i=1}^N (X_0^i - Y_0^i)^2\right)^{1/2} = \sqrt{N}e^{-\lambda t}\mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N, \nu_0^N\right)\right),$$

(where this last expectation is taken for any coupling of $\rho_0^{1,N}$ and $\rho_0^{2,N}$) we get, for all $N \geqslant 0$

$$\mathcal{W}_2\left(\bar{\mu}_t, \bar{\nu}_t\right) \leqslant \mathbb{E}\left(\mathcal{W}_2\left(\bar{\mu}_t, \mu_t^N\right) + e^{-\lambda t}\left(\mathcal{W}_2\left(\mu_0^N, \bar{\mu}_0\right) + \mathcal{W}_2\left(\bar{\mu}_0, \bar{\nu}_0\right) + \mathcal{W}_2\left(\bar{\nu}_0, \nu_0^N\right)\right) + \mathcal{W}_2\left(\nu_t^N, \bar{\nu}_t\right)\right).$$

Recall from Lemma 5.5.1 $\mathbb{E}\left(\mathcal{W}_2\left(\bar{\mu}_t, \mu_t^N\right)\right) \to 0$ as N tends to infinity, thus using Assumptions 5.5 and 5.6. By taking the limit as N tends to infinity in the righthand side of the inequality above , we obtain

$$\mathcal{W}_2\left(\bar{\mu}_t, \bar{\nu}_t\right) \leqslant e^{-\lambda t}\mathcal{W}_2\left(\bar{\mu}_0, \bar{\nu}_0\right).$$

□

Recall from [19] that the space of probability measures endowed with the Wasserstein distance is a complete metric space. Thus, thanks to the Banach fixed point theorem, the contraction of the Wasserstein distance for the non linear limit yields the existence of a stationary distribution.

Lemma 5.5.3 (Propagation of chaos for the stationary distribution)**.** *Under Assumptions 5.5, 5.6, and 5.7, denote by $\bar{\rho}_\infty$ (resp. ρ_∞^N) the stationary measure for the non linear process (resp. for the particle system), and let μ_∞^N be an empirical measure associated to ρ_∞^N. We have*

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)^2\right) \to 0 \qquad \textit{as} \qquad N \to \infty.$$

Proof. We have

$$\begin{aligned}\mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_\infty\right) &\leqslant e^{-\lambda t}\mathcal{W}_2\left(\bar{\rho}_0, \bar{\rho}_\infty\right), \\ \mathcal{W}_2\left(\rho_t^N, \rho_\infty^N\right) &\leqslant e^{-\lambda t}\mathcal{W}_2\left(\rho_0^N, \rho_\infty^N\right).\end{aligned}$$

Let μ_∞^N be an empirical measure associated to ρ_∞^N. We have, for all $t \geqslant 0$, under π_t the optimal coupling between ρ_t^N and ρ_∞^N

$$\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)^2\right) \leqslant 3\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_\infty^N, \mu_t^N\right)^2\right) + 3\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)^2\right) + 3\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_\infty\right)^2\right).$$

We consider an initial condition $\bar{\rho}_0 = \bar{\rho}_\infty$ (and thus for all $t \geqslant 0$, $\bar{\rho}_t = \bar{\rho}_\infty$) and $X_0^1, ..., X_0^N$ i.i.d initial condition (reordered) distributed according to $\bar{\rho}_\infty$ (this way $\mathbb{E}(\mathcal{W}_2(\mu_0^N, \bar{\rho}_\infty)) \to 0$). We get

$$\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)^2\right) \leqslant 3\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_\infty^N, \mu_t^N\right)^2\right) + 3\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)^2\right).$$

On one hand, since the optimal transport map for the $\mathcal{W}_2$ distance between two sets of points in dimension one is the map that transports the first point to the first point, the second to the the second, etc, when the two sets are ordered,

$$\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_\infty^N, \mu_t^N\right)^2\right) = \mathbb{E}^{\pi_t}\left(\frac{1}{N}\sum_{i=1}^N (X^i - Y^i)^2\right) = \frac{1}{N}\mathcal{W}_2\left(\rho_t^N, \rho_\infty^N\right)^2 \leqslant \frac{e^{-2\lambda t}}{N}\mathcal{W}_2\left(\rho_0^N, \rho_\infty^N\right)^2. \tag{5.5.2}$$

Then, there exists a constant C_0, depending on the uniform bounds on the second moments of the non linear process and the empirical measure of the particle system such that

$$\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_\infty^N, \mu_t^N\right)^2\right) \leqslant C_0 e^{-2\lambda t}$$

On the other hand, as $\bar{\rho}_t$ is a deterministic measure, we have

$$\mathbb{E}^{\pi_t}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)^2\right) = \mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)^2\right) \to 0 \quad \text{as} \quad N \to \infty.$$

This yields

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)^2\right) \leqslant C_0 e^{-2\lambda t} + 3\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)^2\right).$$

Consider $\epsilon > 0$. There is t_ϵ such that for all $t \geqslant t_\epsilon$ we have $C_0 e^{-2\lambda t} \leqslant \frac{\epsilon}{2}$ and, given t_ϵ, there is a N_ϵ such that for all $N \geqslant N_\epsilon$ we have $3\mathbb{E}\left(\mathcal{W}_2\left(\mu_{t_\epsilon}^N, \bar{\rho}_{t_\epsilon}\right)^2\right) \leqslant \frac{\epsilon}{2}$. This way

$$\forall \epsilon > 0, \quad \exists N_\epsilon \geqslant 0, \quad \forall N \geqslant N_\epsilon, \quad \mathbb{E}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)^2\right) \leqslant \epsilon,$$

i.e

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)^2\right) \to 0 \quad \text{as} \quad N \to \infty.$$

□

Lemma 5.5.4 (Uniform continuity in t, uniformly in N)**.** *Under Assumptions 5.5, 5.6, 5.7 and 5.8, the function* $t \to \mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)\right)$ *is uniformly continuous in* t*, uniformly in* N*.*

Proof. Let us begin by showing the function $t \to \mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_0\right)$ is continuous in $t = 0$. We have, for all $N \geqslant 0$

$$\mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_0\right) \leqslant \mathbb{E}\left(\mathcal{W}_2\left(\bar{\rho}_t, \mu_t^N\right) + \mathcal{W}_2\left(\mu_t^N, \mu_0^N\right) + \mathcal{W}_2\left(\mu_0^N, \bar{\rho}_0\right)\right).$$

Let $\epsilon > 0$. First, we have

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \mu_0^N\right)^2\right) = \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - X_0^i|^2\right),$$

and thus, by Assumption 5.8, there exists $\delta > 0$ such that for all $t \leqslant \delta$ and $N \geqslant 0$

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \mu_0^N\right)\right) \leqslant \frac{\epsilon}{3}.$$

Then, let $t \leqslant \delta$. Using the strong propagation of chaos, there exists $N_t \geqslant 0$ and $N_0 \geqslant 0$ such that for $N = \max(N_t, N_0)$

$$\mathbb{E}\left(\mathcal{W}_2\left(\bar{\rho}_t, \mu_t^N\right)\right) \leqslant \frac{\epsilon}{3} \quad \text{and} \quad \mathbb{E}\left(\mathcal{W}_2\left(\mu_0^N, \bar{\rho}_0\right)\right) \leqslant \frac{\epsilon}{3}$$

Hence

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall t < \delta,\ \mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_0\right) < \epsilon,$$

and the continuity of the function $t \to \mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_0\right)$ in $t = 0$.

Now, let $t \geqslant 0$ and $(t_n)_{n\in\mathbb{N}}$ a sequence converging to t. We have

$$\begin{aligned}
&\left|\mathbb{E}^{\pi^N_{t,t_n}}\left(\mathcal{W}_2\left(\mu_{t_n}^N, \bar{\rho}_{t_n}\right)\right) - \mathbb{E}^{\pi^N_{t,t_n}}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)\right)\right| \\
&\quad \leqslant \left|\mathbb{E}^{\pi^N_{t,t_n}}\left(\mathcal{W}_2\left(\mu_{t_n}^N, \mu_t^N\right)\right) + \mathbb{E}^{\pi^N_{t,t_n}}\left(\mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_{t_n}\right)\right)\right| \\
&\quad \leqslant e^{-\lambda(t\wedge t_n)}\left(\frac{1}{\sqrt{N}}\mathcal{W}_2\left(\rho^N_{|t-t_n|}, \rho_0^N\right) + \mathcal{W}_2\left(\bar{\rho}_{|t-t_n|}, \bar{\rho}_0\right)\right),
\end{aligned}$$

where the expectation is taken under π^N_{t,t_n} the optimal coupling between $\rho^N_{t_n}$ and ρ^N_t and the last inequality comes from the fact that

$$\begin{aligned}
\mathbb{E}^{\pi^N_{t,t_n}}\left(\mathcal{W}_2\left(\mu_{t_n}^N, \mu_t^N\right)\right) &\leqslant \mathbb{E}^{\pi^N_{t,t_n}}\left(\mathcal{W}_2\left(\mu_{t_n}^N, \mu_t^N\right)^2\right)^{1/2} = \mathbb{E}^{\pi^N_{t,t_n}}\left(\frac{1}{N}\sum_{i=1}^N (X^i - Y^i)^2\right)^{1/2} \\
&= \frac{1}{\sqrt{N}}\mathcal{W}_2\left(\rho_t^N, \rho_{t_n}^N\right).
\end{aligned}$$

We have

$$\frac{1}{N}\mathcal{W}_2\left(\rho^N_{|t-t_n|}, \rho_0^N\right)^2 \leqslant \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X^i_{|t-t_n|} - X_0^i|^2\right).$$

The continuity in 0 of $t \to \mathcal{W}_2\left(\bar{\rho}_t, \bar{\rho}_0\right)$, and the continuity in 0 (uniform in N) of $t \to \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - X_0^i|^2\right)$ are therefore sufficient to yield the result. □

Lemma 5.5.5. *Under Assumptions 5.5, 5.6, 5.7 and 5.8, there exists a non-decreasing sequence* $(t_N)_{N\geqslant 0}$ *that goes to infinity such that for all* $N \geqslant 0$

$$\sup_{s\in[0,t_N]} \mathbb{E}\left(\mathcal{W}_2\left(\mu_s^N, \bar{\rho}_s\right)\right) \to 0 \qquad \textit{as} \qquad N \to \infty. \tag{5.5.3}$$

Proof. By strong propagation of chaos,

$$\forall \epsilon > 0, \quad \forall t \geqslant 0, \quad \exists N \geqslant 0, \quad \forall n \geqslant N, \quad \mathbb{E}\left(\mathcal{W}_2\left(\mu_t^n, \bar{\rho}_t\right)\right) \leqslant \epsilon. \tag{5.5.4}$$

Denote $g(t, N) = \mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)\right)$. By Lemma 5.5.4, g is uniformly continuous in t, uniformly in N. Let $\epsilon > 0$ and $t > 0$. There exists $N_1 \geqslant 0$ such that for all $n \in \mathbb{N}$ and all $x, y \in [0, t]$

$$|x - y| \leqslant \frac{t}{N_1} \implies |g(x, n) - g(y, n)| \leqslant \frac{\epsilon}{2}.$$

We also have

$$\exists N \geqslant 0, \quad \forall n \geqslant N, \quad \forall i \in \{0, .., N_1\}, \quad g\left(t\frac{i}{N_1}, n\right) \leqslant \frac{\epsilon}{2}.$$

This way

$$\exists N \geqslant 0, \quad \forall n \geqslant N, \quad \forall s \in [0, t], \quad g(s, n) \leqslant \epsilon.$$

Denoting $f(t, N) = \sup_{s\in[0,t]} \mathbb{E}\left(\mathcal{W}_2\left(\mu_s^N, \bar{\rho}_s\right)\right)$, we thus obtain

$$\forall \epsilon > 0, \quad \forall t \geqslant 0, \quad \exists N \geqslant 0, \quad \forall n \geqslant N, \quad f(t, n) \leqslant \epsilon. \tag{5.5.5}$$

There exists a non-decreasing function $\phi : \mathbb{R} \mapsto \mathbb{N}$ such that for all $t \geqslant 0$ and all $n \geqslant \phi(t)$ we have $f(t, n) \leqslant \frac{1}{t}$ and $\lim_{t\to\infty} \phi(t) = +\infty$. By convention $\phi(0) = 0$.

Consider $t_0 = 0$ and

$$t_N = \sup\{t \geqslant t_{N-1} \text{ s.t. } t \in \phi^{-1}(\{0, 1, .., N\})\}.$$

The sequence $(t_N)_{N\geqslant 0}$ thus defined is non-decreasing by construction. Because $\lim_{t\to\infty} \phi(t) = +\infty$, the set $\phi^{-1}(\{0, 1, .., N\})$ is non-empty and its supremum goes to infinity as N goes to infinity. Therefore $\lim_{N\to\infty} t_N = +\infty$ and $t_N \neq 0$ eventually.

We have $N \geqslant \phi(t_{N-1})$, and therefore by definition of ϕ, we eventually get for N sufficiently large

$$f(t_{N-1}, N) \leqslant \frac{1}{t_{N-1}}.$$

This concludes the proof. □

We may now conclude.

Proof of Theorem 5.5.1. We have, for μ_∞^N an empirical measure associated to ρ_∞^N, and π_t^N the optimal coupling between ρ_t^N and ρ_∞^N

$$\begin{aligned}\mathbb{E}^{\pi_t^N}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)\right) \leqslant & \mathbb{E}^{\pi_t^N}\left(\mathcal{W}_2\left(\mu_t^N, \mu_\infty^N\right)\right) + \mathbb{E}^{\pi_t^N}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)\right) + \mathbb{E}^{\pi_t^N}\left(\mathcal{W}_2\left(\bar{\rho}_\infty, \bar{\rho}_t\right)\right) \\ \leqslant & e^{-\lambda t}\left(\frac{1}{\sqrt{N}}\mathcal{W}_2\left(\rho_0^N, \rho_\infty^N\right) + \mathcal{W}_2\left(\bar{\rho}_\infty, \bar{\rho}_0\right)\right) + \mathbb{E}\left(\mathcal{W}_2\left(\mu_\infty^N, \bar{\rho}_\infty\right)\right).\end{aligned}$$

Since

$$\frac{1}{\sqrt{N}}\mathcal{W}_2\left(\rho_0^N, \rho_\infty^N\right) \leqslant \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_0^i - X_\infty^i|^2\right)^{1/2} \leqslant 2C_0^{1/2},$$

we obtain

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)\right) \leqslant \tilde{C}(t) + \tilde{f}(N), \tag{5.5.6}$$

where $\tilde{C}$ is decreasing and goes to 0, and $\tilde{f}$ tends to 0.

Let $t \geqslant 0$. If $t \leqslant t_N$ where t_N is given in Lemma 5.5.5, we have using (5.5.3)

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)\right) \leqslant \sup_{s\in[0,t_N]} \mathbb{E}\left(\mathcal{W}_2\left(\mu_s^N, \bar{\rho}_s\right)\right) \to 0 \quad \text{as} \quad N \to \infty,$$

and if $t > t_N$, using (5.5.6)

$$\mathbb{E}\left(\mathcal{W}_2\left(\mu_t^N, \bar{\rho}_t\right)\right) \leqslant \tilde{C}(t) + \tilde{f}(N) \leqslant \tilde{C}(t_N) + \tilde{f}(N) \to 0 \quad \text{as} \quad N \to \infty.$$

Those two bounds being independent of t, we obtain uniform in time propagation of chaos. □

5.6 Addendum : More general version of Section 5.5

As suggested to us by Prof. Martin Hairer, Section 5.5 could be written in a more general way, which extends the result beyond dimension one and its linear order. We detail here the new proof.

Let us start by explaining why the Assumptions that we make on the joint law (now seen as a symmetric joint law on the whole space) can be extended to the law of the empirical measure.

Let $(\mathcal{X}, d)$ be a Polish space, and consider the map associating to N points the corresponding empirical distribution

$$\Pi_N : (X_1, ..., X_N) \in \mathcal{X}^N \mapsto \frac{1}{N}\sum_{i=1}^N \delta_{X_i} \in \mathcal{P}(\mathcal{X}).$$

Denote $\mathcal{X}_q^N$ the space $\mathcal{X}^N$ endowed with the distance

$$d_q^q(X, Y) = \frac{1}{N}\sum_{i=1}^N d^q(X_i, Y_i),$$

and write $\mathcal{P}_q(\mathcal{X})$ the space of probability measures on $\mathcal{X}$ endowed with the q-Wasserstein distance.

Lemma 5.6.1. *The map $\Pi_N : \mathcal{X}_q^N \mapsto \mathcal{P}_q(\mathcal{X})$ is a contraction.*

This lemma is a direct consequence of the definitions, as

$$\mathcal{W}_q^q\left(\Pi_N(X), \Pi_N(Y)\right) \leqslant d_q^q(X, Y).$$

Lemma 5.6.2. *The pushforward map $\Pi_N^* : \mathcal{P}_p(\mathcal{X}_q^N) \mapsto \mathcal{P}_p(\mathcal{P}_q(\mathcal{X}))$, which associates the law of the empirical distribution to the joint law of the points in $\mathcal{X}_q^N$, is a contraction.*

Proof. Denote μ^N and ν^N two probability measure in $\mathcal{P}_p(\mathcal{X}_q^N)$ and let $X \sim \mu^N$ and $Y \sim \nu^N$. In particular, $\Pi_N(X) \sim \Pi_N^*(\mu^N)$ and $\Pi_N(Y) \sim \Pi_N^*(\nu^N)$. Thus

$$\begin{aligned}\mathcal{W}_p^p\left(\Pi_N^*(\mu^N), \Pi_N^*(\nu^N)\right) \leqslant& \mathbb{E}\left(\mathcal{W}_q^p(\Pi_N(X), \Pi_N(Y))\right) \\ \leqslant& \mathbb{E}\left(d_q^q(X, Y)\right).\end{aligned}$$

This being true for all coupling of (X, Y) where $X \sim \mu^N$ and $Y \sim \nu^N$, we may consider the optimal coupling between μ^N and ν^N, and obtain

$$\mathcal{W}_p^p\left(\Pi_N^*(\mu^N), \Pi_N^*(\nu^N)\right) \leqslant \inf_{X\sim\mu^N, Y\sim\nu^N} \mathbb{E}\left(d_q^q(X, Y)\right) = \mathcal{W}^p(\mu^N, \nu^N)$$

□

Remark 5.6.1. *Staying coherent with our notations so far, we thus have* $\Pi_N(\boldsymbol{X}_t) = \mu_t^N$ *and* $\Pi_N^*(\rho_t^N) = \xi_t^N$*, with* $\boldsymbol{X}_t$ *the solution of the N-particle system,* μ_t^N *the associated empirical law,* ρ_t^N *the joint law of the particle system, and* ξ_t^N *the law of* μ_t^N*.*

Remark 5.6.2. *In fact, the pushforward map* Π_N^* *is an isometry from the space of symmetric (i.e. exchangeable) probability measures on* $\mathcal{X}^N$ *to* $\mathcal{P}(\mathcal{P}(\mathcal{X}))$ *(see Proposition 2.14 of [91] for case in Wasserstein-1 and Lemma 11 of [37] for the proof in the case of Wasserstein-2).*

We consider Assumption 5.5-5.8 written in the more general case of processes in a Polish space $\mathcal{X}$. In particular, Assumption 5.7 now concerns the long time convergence of the symmetric probability measures ρ_t^N on $\mathcal{X}^N$ or, equivalently by Remark 5.6.2 above, of the probability measure ξ_t^N on the space $\mathcal{P}_2(\mathcal{X})$.

Proposition 5. *Let* $(\mathcal{X}, d)$ *be a complete metric space, and let* Y_t^N *be a collection of elements in* $\mathcal{Y}$ *with the following properties:*

1. *For every fixed* $t \geqslant 0$*, there exists* $Y_t \in \mathcal{Y}$ *such that* $\lim_{N\to\infty} Y_t^N = Y_t$*,*
2. *The map* $t \mapsto Y_t^N$ *is uniformly continuous, uniformly in* N*,*
3. *There exist constants* C *and* $\lambda > 0$ *such that* $d(Y_s^N, Y_t^N) \leqslant Ce^{-\lambda(s\wedge t)}$*.*

Then one has $\lim_{N\to\infty} \sup_{t\geqslant 0} d(Y_t^N, Y_t) = 0$*.*

Proof of Theorem 5.5.1. We use Proposition 5 with $\mathcal{Y} = \mathcal{P}_2(\mathcal{P}_2(\mathbb{R}))$ endowed with the Wasserstein distance, $Y_t^N = \xi_t^N$ and $Y_t = \delta_{\bar{\rho}_t}$.

In the spirit of Lemma 5.5.1, the first point of Proposition 5 is a consequence of Assumptions 5.5 and 5.6. Likewise, similarly as Lemma 5.5.4, the second point is a consequence of 5.8.

The third point of Proposition 5 is obtain from Assumption 5.7 from

$$d(Y_s^N, Y_t^N) \leqslant e^{-\lambda(s\wedge t)} d(Y_0^N, Y_{|t-s|}^N),$$

and $d(Y_0^N, Y_{|t-s|}^N) \leqslant d(Y_0^N, \delta_{\delta_0}) + d(\delta_{\delta_0}, Y_{|t-s|}^N) \leqslant C$ thanks to Assumption 5.6. □

Proof of Proposition 5. Let us start by showing that for all $0 \leqslant s \leqslant t$ we have $d(Y_s, Y_t) \leqslant Ce^{-\lambda s}$. For all $N \geqslant 0$, we have

$$d(Y_s, Y_t) \leqslant d(Y_s, Y_s^N) + d(Y_t^N, Y_s^N) + d(Y_t^N, Y_t) \leqslant d(Y_s, Y_s^N) + Ce^{-\lambda s} + d(Y_t^N, Y_t),$$

where C and λ are independent of N, s and t. Taking the limit $N \to \infty$ above yields $d(Y_s, Y_t) \leqslant Ce^{-\lambda s}$. From the first and third points, we then have for all $0 \leqslant s \leqslant t$

$$d(Y_t^N, Y_t) \leqslant d(Y_t^N, Y_s^N) + d(Y_s^N, Y_s) + d(Y_s, Y_t) \leqslant d(Y_s^N, Y_s) + 2Ce^{-\lambda s}.$$

Let $\epsilon > 0$. There exists $s_* \geqslant 0$ such that $2Ce^{-\lambda s_*} \leqslant \epsilon$. Furthermore, by the second point, let $\delta > 0$ be such that

$$\forall N \geqslant 0,\ \forall u, v \geqslant 0 \ \text{ s.t. } |u - v| \leqslant \delta,\ d(Y_u^N, Y_v^N) \leqslant \epsilon.$$

Finally, let N be large enough so that $d(Y_u^N, Y_u) \leqslant \epsilon$ for all $u < s_*$ such that $u = k\delta$ with $k \in \mathbb{N}$ (note that there are thus only a finite number of possible values for u). We have

$$\begin{aligned}\sup_{t \geqslant s_*} d(Y_t^N, Y_t) \leqslant & d(Y_s^N, Y_{s_*}) + 2Ce^{-\lambda s_*} \\ \leqslant & d(Y_{s_*}^N, Y^N_{\lfloor \frac{s_*}{\delta} \rfloor \delta}) + d(Y^N_{\lfloor \frac{s_*}{\delta} \rfloor \delta}, Y_{\lfloor \frac{s_*}{\delta} \rfloor \delta}) + 2Ce^{-\lambda s_*} \\ \leqslant & 3\epsilon,\end{aligned}$$

and

$$\begin{aligned}\sup_{t < s_*} d(Y_t^N, Y_t) \leqslant & d(Y_t^N, Y^N_{\lfloor \frac{t}{\delta} \rfloor \delta}) + d(Y^N_{\lfloor \frac{t}{\delta} \rfloor \delta}, Y_{\lfloor \frac{t}{\delta} \rfloor \delta}) \\ \leqslant & 2\epsilon.\end{aligned}$$

Hence we obtain the result. □

In particular, notice that there is no need to use the linear order of the 1D case and the result is fairly general.

We however choose to keep the previous proof of Theorem 5.5.1, since we proved the long time convergence of the limit and the property of propagation of chaos for the stationary distribution (whereas this new proof uses a given reference time s) which are in themselves interesting results.

Part III

Incomplete interactions

Chapter 6

A note on uniform in time mean-field limit in graphs

Attention, car les ressemblances sont grandes. Ce sont des ressemblances identiques. Ecoutez, suivez bien...

Eugène Ionesco, *La leçon* (1950).

Outline of the current chapter

Joint work with Christophe Poquet.

Submitted as: Pierre Le Bris, and Christophe Poquet. "A note on uniform in time mean-field limit in graphs." *arXiv preprint arXiv:2211.11519* (2022). [115]

Abstract: In this chapter we show, in a concise manner, a result of uniform in time propagation of chaos for non exchangeable systems of particles interacting according to a random graph. Provided the interaction is Lipschitz continuous, the restoring force satisfies a general one-sided Lipschitz condition (thus allowing for non-convex confining potential) and the graph is dense enough, we use a coupling method suggested by Eberle [67] known as *reflection* coupling to obtain uniform in time mean-field limit with bounds that depend explicitly on the graph structure.

Acknowledgements: P.LB. and C.P. acknowledge the support of ANR-17-CE40-0030 (EFI). C.P. also acknowledges the support of ANR-19-CE40-0023 (PERISTOCH).

6.1 Introduction

6.1.1 Model and motivation

Let $N \in \mathbb{N}$ and consider an adjacency matrix $\xi^{(N)} = \left(\xi^{(N)}_{i,j}\right)_{i,j\in\{1,...,N\}}$ with coefficients $\xi^{(N)}_{i,j} \in \{0,1\}$. Denote by $G^{(N)} = (V^{(N)}, E^{(N)})$ the graph associated to this adjacency matrix, in the sense $V^{(N)} := \{1,...,N\}$ and $E^{(N)} := \left\{(i,j) \in V^{(N)} \times V^{(N)} \text{ s.t. } \xi^{(N)}_{i,j} = 1\right\}$. We assume by convention that $\xi^{(N)}_{i,i} = 0$ for all $i \in \{1,...,N\}$.

We will consider in this note a system of particles interacting according to this graph, more precisely the system of N SDEs in $\mathbb{R}^d$

$$dX^i_t = F\left(X^i_t, \omega_i\right) dt + \frac{\alpha_N}{N}\sum_{j=1}^N \xi^{(N)}_{i,j}\Gamma\left(X^i_t, \omega_i, X^j_t, \omega_j\right) dt + \sqrt{2}\sigma dB^i_t, \quad i \in \{1,...,N\}, \tag{IPS}$$

where $\left(B^i_\cdot\right)_{i=1,...,N}$ is a sequence of independent standard Brownian motions, $\{\omega_i\}_{i\in\{1,...,N\}}$ is a sequence of elements in $\mathbb{R}^{d'}$ (with the convention $d' = 0$ if Γ does not depend on ω) which represents some environmental disorder, $(\alpha_N)_{N\geqslant 1}$ is a positive scaling, $F : \mathbb{R}^d \times \mathbb{R}^{d'} \mapsto \mathbb{R}^d$ is an outside force, $\Gamma : \left(\mathbb{R}^d \times \mathbb{R}^{d'}\right)^2 \mapsto \mathbb{R}^d$ is an interaction kernel and σ is a positive diffusion coefficient. We will assume that $(\omega_i)_{i=1,...,N}$ is a sequence of IID random variables, and that the Brownian motions are independent from the initial condition $(X^i_0, \omega_i)_{i=1...,N}$. We will denote by $\mathbb{E}$ the expectation with respect to the Brownian motions, the initial condition and the disorder.

One of the main difficulties arising in the study of this model comes from the fact that the particles are not exchangeables as, *a priori*, some may interact with more particles than others. This motivates us to consider the empirical distribution, defined for $(X^1_t, ..., X^N_t)$ a solution of (IPS) with disorder $(\omega_1, \dots, \omega_N)$, by

$$\mu^N_t := \frac{1}{N}\sum_{i=1}^N \delta_{(X^i_t,\omega_i)}.$$

Notice that μ^N_t is a random variable.

We are interested in the limit, as the number N of particles goes to infinity, of (IPS). Intuitively one expects the empirical measure to converge towards a measure $\bar{\rho}$ which would represent the law of one typical particle and its disorder within a cloud of interacting disordered particles. Assuming that $\frac{\alpha_N}{N}\sum_{j=1}^N \xi^{(N)}_{i,j}$ converges in some sense (given later) to a parameter p, this typical particle $\bar{X}^\omega$ with disorder ω would then in the limit evolve according to the non-linear diffusion

$$\begin{cases} d\bar{X}^\omega_t = F\left(\bar{X}^\omega_t, \omega\right) dt + p\int_{\mathbb{R}^d\times\mathbb{R}^{d'}} \Gamma\left(\bar{X}^\omega_t, \omega, y, \tilde{\omega}\right) \bar{\rho}_t(dy, d\tilde{\omega})dt + \sqrt{2}\sigma dB_t, \\ \bar{\rho}_t = \mathrm{Law}(\bar{X}^\omega_t, \omega) \end{cases}, \tag{NL}$$

where B is a standard Brownian motion. This limit was proven rigorously on finite time horizon $[0,T]$, where T does not depend on N, under some hypotheses on the graph structure, which are in particular satisfied by sufficiently dense Erdős-Rényi graphs [58, 55]. Our aim in the present paper is to obtain uniform in time estimates of the distance between the empirical measure μ^N_t and the limit distribution $\bar{\rho}_t$, with estimates that depend explicitly on the graph.

Note that proofs of convergence of particle systems interacting via random graphs possessing a spatial structure (for example converging to a graphon) were recently obtained [144, 129, 11, 12]. In particular uniform in time estimates in the context of graphons were obtained in [12], where the empirical measure is shown to be close to the limit distribution with high probability with respect to the distribution of the random graph. In this note we aim at obtaining *quenched* results, i.e. obtaining estimates that hold for almost every realization of the graph.

These recent results generalize the classical case of complete graph of interaction ($\alpha_N = 1$ and $\xi \equiv 1$) and without any dependence on the environment ω, for which it is well known that under some weak conditions on F and Γ the empirical measure μ_t^N converges towards the non-linear limit $\bar{\rho}_t$ [137, 162]. This phenomenon has been named *propagation of chaos*, an idea motivated by Kac [106] : it is equivalent, in the case of exchangeable particles, to the convergence of all k marginals of the law of $(X_t^1, ..., X_t^N)$ to $\bar{\rho}_t^{\otimes k}$ (the non linear limit tensorized k times). Thus, as N goes to infinity, two particles become "more and more" independent, converging to a tensorized law, hence *chaos*. The term *propagation* emphasizes the fact that it is sufficient to show independence at the limit at time 0 for it to also hold true at the limit at later time $t > 0$. We refer to the recent [45, 46], and references therein, for a thorough reviews on propagation of chaos.

To quantify the convergence of the empirical measure towards the non-linear limit, we will use the L^1-Wasserstein distance defined as follows.

Definition 6.1.1. *For μ and ν two probability measures on $\mathbb{R}^d$, denote by $\Pi(\mu,\nu)$ the set of couplings of μ and ν, i.e. the set of probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with $\pi(A \times \mathbb{R}^d) = \mu(A)$ and $\pi(\mathbb{R}^d \times A) = \nu(A)$ for all Borel set A of $\mathbb{R}^d$. The L^1-Wasserstein distance is given by*

$$\mathcal{W}_1(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int |x - \tilde{x}| \pi(dx d\tilde{x}) . \tag{6.1.1}$$

Equivalently, we may write in probabilistic terms

$$\mathcal{W}_1(\mu,\nu) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}(|X - Y|) ,$$

where $X \sim \mu$ is a random variable distributed according to μ. This distance is a usual distance in optimal transport and in the study of measures in general, as the space of probability measures on $\mathbb{R}^d$, equipped with the L^1-Wasserstein distance, is a complete and separable metric space (see for instance [19]). To prove the convergence in Wasserstein distance, we use a *coupling method*. The idea is, instead of considering the minimum over all couplings of the law of the particle system and the non-linear limit as should be done according to (6.1.1), we construct simultaneously two solutions of (IPS) and (NL) such that the expectation of the L^1 distance between these solutions tends to decrease. We would thus construct a specific coupling, that controls the L^1-Wasserstein distance, providing a quantitative bound. To construct this coupling, we may act on the Brownian motions and on the random variables ω.

The approach we consider was motivated by the work of Eberle [67]. Let us describe the idea of the coupling method. Assume, for the sake of the explanation, that $F = -\nabla U$ where U is therefore a confinement potential. Constructing a solution of (IPS) and N independent solutions of (NL) simultaneously by choosing the same Brownian motions yields the so-called *synchronous* coupling, for which the Brownian noise cancels out in the infinitesimal evolution of the difference $Z_t^i = X_t^i - \bar{X}_t^{\omega,i}$. In that case the contraction of a distance between the processes can only be induced by the deterministic drift. Such a deterministic contraction only holds under very restrictive conditions, in particular U should be strongly convex. In the case of a non-convex confinement potential U, it is necessary to make use of the noise to obtain contraction. Constructing the solutions choosing the two Brownian motions to be antithetic (or opposite) in the direction of space given by the difference of the processes maximises the variance of the noise in the desired direction. However, *a priori*, nothing ensures the noise will bring the processes closer rather than further. We thus modify the Euclidean distance by some concave function f, in order for a random decrease of the difference to have more effect than a random increase of the same amount.

This method was originally designed to deal with the long time behavior of general diffusion processes, as in [67, 70], and later extended to show uniform in time propagation of chaos in a mean-field system in [64]. The main difference of this work when compared to [64] comes from the non-exchangeability of the particles, as we thus need careful estimates with respect to the graph. For instance, since the particles do not share a common law, we cannot restrict our analysis to the study of $\mathbb{E}[Z_t^1]$ and then conclude using the fact that all Z_t^i have the same expectation ; the proof requires a more global approach to the system, by considering the empirical measure, and thus other tools.

The framework of this article was inspired by [58], and we improve their result, obtaining a uniform in time estimate, while removing some of the boundedness assumptions on the various functions.

Uniform in time propagation of chaos has recently attracted a lot of attention. The ideas behind this coupling method were used to prove such estimates in a kinetic setting (i.e a particle is represented by both its position and velocity, and the Brownian motion only acts on the latter) in [83, 156]. In [133, 132], uniform in time propagation of chaos was proved using synchronous coupling assuming a convexity condition on the interaction. Likewise, a similar result was obtained in [40], using functional inequalities, under some assumptions of convexity at infinity. Also using functional inequalities for mean field models developed in [87], uniform in time propagation of chaos was proved in a kinetic setting in [86, 89, 142] combining the hypocoercivity approach with uniform in the number of particles logarithmic Sobolev inequalities. Let us also mention the optimal coupling approach of [153] using a WJ inequality, which is also used in [56], which enables to recover the results in [64]. Thanks to an analysis of the relative entropy through the BBGKY hierarchy, building upon the work [111], a result of uniform in time propagation of chaos was obtained, with a sharp rate in N, in [113]. Finally, in the recent work [57], uniform in time weak propagation of chaos (i.e observable by observable) was shown on the torus via Lions derivative. Notably, this result may extend to the case the McKean-Vlasov limit has several invariant measures, as in the Kuramoto model for instance.

All the works mentioned above assume the interaction to be "sufficiently nice" (either gradient of a convex potential, smooth, bounded, etc), and we will also consider an Lipschitz continuous interaction, but not according to a random graph. Let us also quickly mention the case of singular interactions which is, because of the various applications in biology, physics and others, also of great interest. Though some recent works have obtained quantitative mean-field convergence for some singular potential (for instance using entropy dissipation in [99], modulated energy in [159], a mix of both in [32], or BBGKY hierarchies in [30]), few still have obtained uniform in time estimates. We mention the results dealing with singular repulsive interactions of the type $-\log|x|$ or $|x|^{-s}$, $0 < s < d-2$, in [152] using the modulated energy, dealing with the specific case of the 2D vortex model in [85] (building upon the work [99]), or dealing with repulsive singular interactions in dimension one in [84] using another type of coupling method.

There again, the particles are not interacting according to a graph.

6.1.2 Assumptions and main result

Denote by $d_i^{(N)} := \sum_{j=1}^N \xi_{i,j}^{(N)}$ and $\tilde{d}_i^{(N)} := \sum_{j=1}^N \xi_{j,i}^{(N)}$ the degrees of vertex i. The family $\xi^{(N)}$ may be deterministic or random, in this second case we assume that it is independent from the Brownian motions and from $(X_0^i, \omega_i)_{i=1,\dots,N}$ and that the following assumption is verified almost surely. These assumptions are similar to the ones made in [58].

Assumption 6.1 (On the graph)**.** *The adjacency matrix $\xi^{(N)}$ satisfies the following assertions for all $N \geqslant 1$.*

6.1-1 *There exists a positive constant C_g such that*

$$\limsup_{N\to\infty} D_{N,g} \leqslant C_g,$$

where

$$D_{N,g} := \sup_{i\in\{1,\dots,N\}} \alpha_N \left(\frac{d_i^{(N)}}{N} + \frac{\tilde{d}_i^{(N)}}{N} \right).$$

6.1-2 *There exists $p \in [0,1]$ such that*

$$I_{N,g} := \sup_{i\in\{1,\dots,N\}} \left| \alpha_N \frac{d_i^{(N)}}{N} - p \right| \xrightarrow[N\to\infty]{a.s} 0.$$

Example 6.1.1.

- *Regular graphs: if $\xi^{(N)}$ defines a regular graph of degree d_N with $\frac{d_N}{N} \xrightarrow[N\to\infty]{} p$, then $\xi^{(N)}$ satisfies Assumption 6.1 with $\alpha_N = 1$.*
- *Erdős-Rényi graphs: Let $\xi_{i,j}^{(N)}$ be a sequence of IID Bernouilli variables of parameter q_N. There are two possible scalings. First, the positive edge-density, in which $q_N \xrightarrow[N\to\infty]{} p > 0$ and $\alpha_N = 1$. Second, the vanishing edge density, in which $q_N \xrightarrow[N\to\infty]{} 0$. In this case, to obtain a non trivial limit, we assume $\frac{1}{q_N} = o\left(\frac{N}{\log N}\right)$ and consider $\alpha_n = \frac{1}{q_N}$ (Note that by rescaling α_N we can assume $p = 1$ here). Then in both cases $\xi^{(N)}$ satisfies Assumption 6.1 (see Proposition 1.3 of [58] for a proof).*
- *Community models: more generally, suppose that the whole population is divided in r sub-populations of size m (so that $N = rm$), the graph structure being then defined by independent random variables $\xi_{i,j}^{(N,k,k')}$ for $k, k' \in \{1,\dots,r\}$ and $i,j \in \{1,\dots,m\}$. Suppose moreover that the intra-community interaction variables $\xi_{i,j}^{(N,k,k)}$ are of Bernoulli distribution with parameter q_N satisfying $\frac{1}{q_N} = o\left(\frac{N}{\log N}\right)$, while the inter-community interaction variables $\xi_{i,j}^{(N,k,k')}$ are of Bernoulli distribution with parameter $q_N^{k,k'}$ satisfying $q_N^{k,k'} = o(q_N)$. Then, for r fixed and $m \to \infty$, $\xi^{(N)}$ satisfies Assumption 6.1 with $\alpha_N = \frac{1}{q_N}$ and $p = \frac{1}{r}$. For more details and the proof of this result see Appendix E.1.*

Assumption 6.2 (On the restoring force)**.** *There exists a continuous function $\kappa : \mathbb{R}^+ \mapsto \mathbb{R}$ satisfying*
$\liminf_{r\to\infty} \kappa(r) > 0$ *such that*

$$\forall x, y \in \mathbb{R}^d, \quad \forall \omega \in \mathbb{R}^{d'}, \quad (F(x,\omega) - F(y,\omega)) \cdot (x-y) \leqslant -\kappa(|x-y|)|x-y|^2.$$

In particular, this implies that there exist $M_F \geqslant 0$ and $m_F > 0$ such that

$$\forall x, y \in \mathbb{R}^d, \quad \forall \omega \in \mathbb{R}^{d'}, \quad (F(x,\omega) - F(y,\omega)) \cdot (x-y) \leqslant M_F - m_F|x-y|^2.$$

The added one-sided assumption on F when compared to [58] is both classical (see [64]) and necessary to ensure that the particles tend to come back to a compact set.

Example 6.1.2. *Let us give some examples of functions F satisfying Assumption 6.2. Let $F(x,\omega) = -V'(x)$ in dimension 1 with :*

- $V(x) = \frac{x^2}{2}$ *: then F satisfies Assumption 6.2 with $\kappa \equiv 1$.*
- $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$ *: then*

$$\begin{aligned}(F(x) - F(y))(x-y) &= -(x^3 - y^3)(x-y) + (x-y)^2 \\ &= -(x-y)^2\left(x^2 + xy + y^2 - 1\right) \\ &\leqslant -(x-y)^2\left(\frac{1}{4}(x-y)^2 - 1\right).\end{aligned}$$

Hence, F satisfies Assumption 6.2 with $\kappa(x) = \frac{x^2}{4} - 1$.

Likewise, we may consider disordered restoring forces such as $F(x,\omega) = -x^3 + \omega x$, provided ω belongs to a bounded subset of $\mathbb{R}$, or $F(x,\omega) = -\omega x^3$ provided ω is positive bounded from below.

Assumption 6.3 (On the interaction)**.** *Γ satisfies 6.3-1 below, and either 6.3-2 or 6.3-2-bis.*

6.3-1 *$\Gamma : (x,\omega,y,\omega') \to \Gamma(x,\omega,y,\omega')$ is Lipschitz-continuous in (x,y) uniformly in ω and ω':*

$$\begin{aligned}&\exists L_\Gamma \geqslant 0,\ \forall x,y,t,s \in \mathbb{R}^d,\ \forall \omega,\omega' \in \mathbb{R}^{d'},\\ &\qquad |\Gamma(x,\omega,t,\omega') - \Gamma(y,\omega,s,\omega')| \leqslant L_\Gamma\left(f(|x-y|) + f(|t-s|)\right),\end{aligned}$$

where f is a function given below in (6.1.3) *such that $x \mapsto f(|x|)$ is equivalent to the usual L^1 distance in $\mathbb{R}$.*

Furthermore, for simplicity, we have $\Gamma(0,0,0,0) = 0$.

6.3-2 *Γ is Lipschitz-continuous in ω and ω' at $(x,y) = (0,0)$:*

$$\begin{aligned}&\exists L_\Gamma \geqslant 0,\ \forall \omega_1,\omega_1',\omega_2,\omega_2' \in \mathbb{R}^{d'},\\ &|\Gamma(0,\omega_1,0,\omega_1') - \Gamma(0,\omega_2,0,\omega_2')| \leqslant L_\Gamma\left(|\omega_1 - \omega_2| + |\omega_1' - \omega_2'|\right).\end{aligned}$$

6.3-2-bis *Γ is bounded*

$$\exists L_\infty \geqslant 0,\ \forall x,y \in \mathbb{R}^d,\ \forall \omega,\omega' \in \mathbb{R}^{d'},\ |\Gamma(x,\omega,y,\omega')| \leqslant L_\infty.$$

These are usual assumptions when proving mean-field limits using coupling methods. In particular, Assumptions 6.2 and 6.3 imply strong existence and uniqueness for the solutions of both (IPS) and (NL).

Assumption 6.4 (On the initial distributions)**.**

6.4-1 *The sequence of disorder $(w_i)_{i=1,\dots,N}$ is IID of distribution ν, satisfying*

$$\int_{\mathbb{R}^{d'}} \left(|\omega|^2 + |F(0,\omega)|^2\right)\nu(d\omega) \leqslant C_{\text{dis}}.$$

6.4-2 *The random variables* $(X_0^i)_{i=1,\dots,N}$ *are exchangeable, independent from the disorder* $(w_i)_{i=1,\dots,N}$ *and satisfy*

$$\mathbb{E}\left(|X_0^1|\right) < \infty.$$

6.4-3 *The initial distribution* $\bar{\rho}_0$ *is a product measure with second marginal equal to* ν*, i.e.* $\bar{\rho}_0(dx, d\omega) = \bar{\rho}_0^1(dx)\nu(d\omega)$*. Moreover there exists a positive constant* $\bar{C}$ *such that*

$$\int_{\mathbb{R}^d} |x|^2 \bar{\rho}_0^1(dx) \leqslant \bar{C}.$$

We may now state the main theorem.

Theorem 6.1.1. *Consider Assumptions 6.1,6.2, 6.3 and 6.4. There exist explicit positive constants* $c_\Gamma, \tilde{C}, \tilde{c}$ *that do not depend on* N *and the graph such that for all* $t \geqslant 0$ *and all* $N \geqslant 1$*, provided* $L_\Gamma \leqslant c_\Gamma / D_{N,g}$ *(recall that* $\limsup_{N\to\infty} D_{N,g} < C_g$*),*

$$\mathbb{E}\mathcal{W}_1\left(\mu_t^N, \bar{\rho}_t\right) \leqslant \tilde{C}\left(e^{-\tilde{c}t}\mathbb{E}\mathcal{W}_1\left(\mu_0^N, \bar{\rho}_0\right) + L_\Gamma\sqrt{\frac{\alpha_N D_{N,g}}{N}} + L_\Gamma I_{N,g} + h(N)\right), \tag{6.1.2}$$

where $h : \mathbb{N} \mapsto \mathbb{R}^+$ *is an explicit decreasing function such that* $h(N) \xrightarrow[N\to\infty]{} 0$ *that only depends on the dimensions* d *and* d' *and the second moment of* ρ *and* $\bar{\rho}_0$*.*

Remark 6.1.1. *The smallness assumption on the Lipschitz coefficient of* Γ *is natural to obtain uniform in time propagation of chaos, as for large interactions the non linear limit may have several stationary measures (see [93] for instance). Non uniqueness of the stationary measures of* (NL) *prevents time-uniform estimate for the mean field limit, since on the other hand there is uniqueness of the stationary distribution of* (IPS)*.*

Remark 6.1.2. *We may write the order of magnitude of the rate function* h *depending on the dimension*

$$h(N) \lesssim N^{-\frac{1}{3}}\mathbb{1}_{d+d'\leqslant 2} + N^{-\frac{1}{d+d'}}\mathbb{1}_{d+d'\geqslant 3}.$$

In reality, this term is a consequence of the approximation of the measure $\bar{\rho}_t$ *by the empirical measure given by* N *independent random variables distributed according to* $\bar{\rho}_t$*, as it is given by [75]. We notice that it could be improved, however at a cost, as there is a tradeoff between the speed of convergence and the moments we impose on the initial condition. If we assumed that* $\bar{\rho}_0$ *admits a* q*-th moment with* $q > 2$*, we could show the following bound:*

$$h(N) \lesssim N^{-\frac{1}{2}}\mathbb{1}_{d+d'=1} + N^{-\frac{1}{2}}\log(1+N)\mathbb{1}_{d+d'=2} + N^{-\frac{1}{d+d'}}\mathbb{1}_{d+d'\geqslant 3}.$$

Remark 6.1.3. *Let us give an example of the rate of convergence in the case of Erdős-Rényi graphs as in Example 6.1.1. Assume that for some* $\alpha \in]0,1[$ *we have* $q_N = N^{-\alpha}$ *and* $\alpha_N = 1/q_N$*. In this case the rate* $\sqrt{\frac{\alpha_N D_{N,g}}{N}} + I_{N,g}$ *can be bounded by* $\sqrt{\frac{\log N}{N q_N}} = \sqrt{\frac{\log N}{N^{1-\alpha}}}$ *(the bound on* $I_{N,g}$ *is given by the proof of Proposition 1.3 of [58]). The rate* $\mathbb{E}\mathcal{W}_1\left(\mu_0^N, \bar{\rho}_0\right) + h(N)$ *depends on the moments of* $\bar{\rho}_0$ *and the dimensions as explained in Remark 6.1.2 above*

Remark 6.1.4. *If Assumption 6.3-2-bis holds instead of Assumption 6.3-2, the coefficients* L_Γ *within the parentheses in* (6.1.2) *are to be replaced by* L_∞*.*

Remark 6.1.5. *Although we do not write the calculations for the sake of conciseness, a similar theorem can be proved if* $p = 0$ *for weaker assumptions on* Γ. *In this case, the limit is a linear Ornstein-Uhlenbeck process, and we do not rely on a form of Law of Large Number to get the limit as* N *goes to infinity, we only have to show that the interaction term vanishes sufficiently fast. Thus we need the expectation of* Γ *to be bounded uniformly in time, which can be done under weaker assumptions and in particular doesn't require a Lipschitz assumption, as the convergence to 0 of* $\frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)}$ *then yields the result of propagation of chaos.*

6.1.3 Semimetric and preliminary results

As mentioned previously, we use a concave function to modify the Euclidean distance in order to use the reflection coupling. Define

$$\begin{aligned}
R_0 &:= \inf\left\{s \geqslant 0 \ : \ \forall r \geqslant s,\ \kappa(r) \geqslant 0\right\},\\
R_1 &:= \inf\left\{s \geqslant R_0 \ : \ \forall r \geqslant s,\ s(s-R_0)\kappa(r) \geqslant 8\sigma^2\right\},
\end{aligned}$$

and the functions

$$\begin{aligned}
\phi(r) &:= \exp\left(-\frac{1}{4\sigma^2}\int_0^r s\kappa_-(s)ds\right),\\
\Phi(r) &:= \int_0^r \phi(s)ds,\\
g(r) &:= 1 - \frac{c}{2}\int_0^{r\wedge R_1} \Phi(s)\phi(s)^{-1}ds,
\end{aligned}$$

where $\kappa_- = \max(0,-\kappa)$ and $c = \left(\int_0^{R_1}\Phi(s)\phi(s)^{-1}ds\right)^{-1}$. Finally, define

$$f(x) = \int_0^x g(t)\phi(t)dt. \tag{6.1.3}$$

Note that ϕ and g are positive non-increasing on $\mathbb{R}^+$ and that $\phi(r) = \phi(R_0) \leqslant 1$ for $r \geqslant R_0$, and $g(r) = \frac{1}{2}$ for $r \geqslant R_1$. In particular, for $r \geqslant R_1$ we simply have $f(r) = f(R_1) + \frac{\phi(R_0)(r-R_1)}{2}$. The function f satisfies moreover some useful properties gathered in the following Lemma, the proof of which can be found in [64].

Lemma 6.1.1 (Some properties of the semimetric)**.** *The function* f *satisfies the following properties :*

- $f : \mathbb{R}^+ \to \mathbb{R}^+$ *is non-negative and increasing. Furthermore* $0 < f'(x) \leqslant 1$ *for all* $x \geqslant 0$.
- *There exist* $c_f, C_f > 0$ *such that for all* $x \in \mathbb{R}$, *we have* $c_f|x| \leqslant f(|x|) \leqslant C_f|x|$.
- *We have*
$$\forall r \in \mathbb{R}^+ \setminus \{R_1\}, \quad f''(r) - \frac{1}{4\sigma^2}r\kappa(r)f'(r) \leqslant -\frac{c}{2}f(r). \tag{6.1.4}$$

We now give a uniform in time moment bound for the non linear process (NL), relying in particular on Assumption 6.2.

Lemma 6.1.2 (Uniform in time bound on the second moment)**.** *Consider Assumption 6.2, 6.3 and 6.4, and let* $(\bar{X}_t^\omega)_t$ *be the unique strong solution of* (NL)*. Assuming* $2pC_fL_\Gamma < m_F$*, there exists a constant* $\bar{\mathcal{C}}_2 > 0$ *such that for all* $t \geqslant 0$

$$\mathbb{E}\left(|\bar{X}_t^\omega|^2\right) \leqslant \bar{\mathcal{C}}_2.$$

Proof. Using Itô's formula on the function $H(x) = \frac{x^2}{2}$, we obtain

$$dH\left(\bar{X}_t^\omega\right) = A_t dt + dM_t, \tag{6.1.5}$$

where $(M_t)_t$ is a continuous local martingale and

$$A_t = \bar{X}_t^\omega \cdot F\left(\bar{X}_t^\omega, \omega\right) + p\int_{\mathbb{R}^d\times\ \mathbb{R}^{d'}} \bar{X}_t^\omega \cdot \Gamma\left(\bar{X}_t^\omega, \omega, y, \bar{\omega}\right)\bar{\rho}_t(dy, d\bar{\omega}) + \sigma^2 d.$$

First, using Assumption 6.2,

$$\bar{X}_t^\omega \cdot F\left(\bar{X}_t^\omega, \omega\right) \leqslant M_F - m_F\left|\bar{X}_t^\omega\right|^2 + \bar{X}_t^\omega \cdot F\left(0, \omega\right).$$

Then, using Assumption 6.3

$$\begin{aligned}\bar{X}_t^\omega\Gamma\left(\bar{X}_t^\omega, \omega, y, \bar{\omega}\right) =& \bar{X}_t^\omega \cdot \left(\Gamma\left(\bar{X}_t^\omega, \omega, y, \bar{\omega}\right) - \Gamma\left(0, \omega, 0, \bar{\omega}\right)\right) + \bar{X}_t^\omega \cdot \left(\Gamma\left(0, \omega, 0, \bar{\omega}\right) - \Gamma\left(0, 0, 0, 0\right)\right)\\ \leqslant& C_fL_\Gamma\left|\bar{X}_t^\omega\right|\left(\left|\bar{X}_t^\omega\right| + |y|\right) + L_\Gamma\left|\bar{X}_t^\omega\right|\left(|\omega| + |\bar{\omega}|\right).\end{aligned}$$

Note that if Assumption 6.3-2-bis holds rather than Assumption 6.3-2, the term above can directly be bounded by $L_\infty|\bar{X}_t^\omega|$. Then

$$\begin{aligned}p\int_{\mathbb{R}^d\times\mathbb{R}^{d'}} &\bar{X}_t^\omega \cdot \Gamma\left(\bar{X}_t^\omega, \omega, y, \bar{\omega}\right)\bar{\rho}_t(dy, d\bar{\omega})\\ &\leqslant pC_fL_\Gamma\left|\bar{X}_t^\omega\right|^2 + pC_fL_\Gamma\left|\bar{X}_t^\omega\right|\mathbb{E}\left|\bar{X}_t^\omega\right| + pL_\Gamma\left|\bar{X}_t^\omega\right||\omega| + pL_\Gamma C_{dis}^{1/2}\left|\bar{X}_t^\omega\right|,\end{aligned}$$

where for this last term we used Assumption 6.4. Finally

$$\mathbb{E}A_t \leqslant M_F + \sigma^2 d - \left(m_F - 2pC_fL_\Gamma\right)\mathbb{E}\left(\left|\bar{X}_t^\omega\right|^2\right) + \mathbb{E}\left(\left|\bar{X}_t^\omega\right|\left(\left|F\left(0, \omega\right)\right| + pL_\Gamma|\omega| + pL_\Gamma C_{dis}^{1/2}\right)\right).$$

Assuming $2pC_fL_\Gamma < m_F$, using the inequality $\forall x, y \in \mathbb{R},\ \forall \alpha > 0,\ xy \leqslant \frac{\alpha x^2}{2} + \frac{y^2}{2\alpha}$, we can ensure there exist two non negative constant B_1 and B_2 such that

$$\mathbb{E}A_t \leqslant B_1 - B_2\mathbb{E}H(\bar{X}_t^\omega).$$

Taking the expectation in (6.1.5), remarking that M_t is a martingale, and using Gronwall's lemma yields the desired result. □

6.2 Mean-field limit

Let $\phi_s, \phi_r : \mathbb{R}^+ \to \mathbb{R}$ be two Lipschitz continuous functions satisfying, for some parameter $\delta > 0$, the following conditions

$$\forall x \in \mathbb{R}^+, \quad \phi_s(x)^2 + \phi_r(x)^2 = 1, \quad \phi_r(x) = \begin{cases} 1 & \text{if } x \geqslant \delta, \\ 0 & \text{if } x \leqslant \delta/2 \end{cases} \tag{6.2.1}$$

These functions describe the regions of space in which we either use a synchronous coupling ($\phi_s \equiv 1$ and $\phi_r \equiv 0$) and a reflection coupling ($\phi_s \equiv 0$ and $\phi_r \equiv 1$). Ideally, we would like to use $\phi_r(x) = \mathbb{1}_{x>0}$, but the indicator function is not continuous, hence the reason we use a Lipschitz approximation.

Consider the initial conditions $(X_0^i, \omega_i)_{i\in\{1,...,N\}}$ for (IPS), and consider N independent random variables $(\bar{X}_0^i)_{i=1,...,N}$ identically distributed according to $\bar{\rho}_0^1$ (recall Assumption 6.4). We know (see for instance Proposition 2.1 of [147]) that there exists a least one permutation $\tau : \{1, ..., N\} \to \{1, ..., N\}$ such that

$$\mathcal{W}_1\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_0^i}, \frac{1}{N}\sum_{i=1}^N \delta_{\bar{X}_0^i}\right) = \frac{1}{N}\sum_{i=1}^N \left|X_0^i - \bar{X}_0^{\tau(i)}\right|. \tag{6.2.2}$$

If there exists more that one such permutation we choose one of them uniformly. Up to renumbering, we assume $\tau(i) = i$ for all $i \in \{1, ..., N\}$. The random variables $(\bar{X}_0^i, \omega_i)$ are then IID with distribution $\bar{\rho}_0$. Using these initial conditions, we now consider the following coupling

$$\begin{cases} dX_t^i &= F\left(X_t^i, \omega_i\right)dt + \frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)}\Gamma\left(X_t^i, \omega_i, X_t^j, \omega_j\right)dt + \sqrt{2}\sigma\phi_s(|X_t^i - \bar{X}_t^i|)d\tilde{B}_t^i \\ & \quad + \sqrt{2}\sigma\phi_r(|X_t^i - \bar{X}_t^i|)dB_t^i, \\ d\bar{X}_t^i &= F\left(\bar{X}_t^i, \omega_i\right)dt + p\int \Gamma(\bar{X}_t^i, \omega_i, x, \omega)\bar{\rho}_t(dx, d\omega)dt + \sqrt{2}\sigma\phi_s(|X_t^i - \bar{X}_t^i|)d\tilde{B}_t^i \\ & \quad + \sqrt{2}\sigma(Id - 2e_t^i(e_t^i)^T)\phi_r(|X_t^i - \bar{X}_t^i|)dB_t^i, \end{cases} \tag{6.2.3}$$

where

$$e_t^i := \begin{cases} \frac{X_t^i - \bar{X}_t^i}{|X_t^i - \bar{X}_t^i|} & \text{if } X_t^i - \bar{X}_t^i \neq 0, \\ 0 & \text{if } X_t^i - \bar{X}_t^i = 0. \end{cases},$$

and $\left(B_\cdot^i\right)_{i=1,...,N}$ and $\left(\tilde{B}_\cdot^i\right)_{i=1,...,N}$ are sequences of independent Brownian motions, and $\bar{\rho}_t$ is the distribution of the non linear diffusion (NL). In particular, Levy's characterization of Brownian motion ensures that $\left(\bar{X}_\cdot^i, \omega_i\right)_i$ are N independent copies of the same diffusion process and thus $\bar{\rho}_t = \text{Law}(\bar{X}_t^1, \omega_1) = ... = \text{Law}(\bar{X}_t^N, \omega_N)$.

Let us denote $Z_t^i = X_t^i - \bar{X}_t^i$. The following lemma concerning the dynamics of $|Z_t^i|$, which can be found in [64], relies on dominated convergence and the fact that $\phi_r(x)$ is zero around $x = 0$.

Lemma 6.2.1 (Lemma 7 of [64]). *For all $t \geqslant 0$ and all $i \in \{1, ..., N\}$,*

$$d\left|Z_t^i\right| = \left(F\left(X_t^i, \omega_i\right) - F\left(\bar{X}_t^i, \omega_i\right)\right)\cdot e_t^i dt + A_t^i dt + 2\sqrt{2}\sigma\phi_r(\left|Z_t^i\right|)e_t^i \cdot dB_t^i,$$

where $(A_t^i)_t$ is an adapted stochastic process such that

$$A_t^i \leqslant \left|\frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)}\Gamma\left(X_t^i, \omega_i, X_t^j, \omega_j\right) - p\int \Gamma(\bar{X}_t^i, \omega_i, x, \omega)\bar{\rho}_t(dx, d\omega)\right|.$$

Applying Itô-Tanaka's formula, as the function f is $\mathcal{C}^1$ and piecewise $\mathcal{C}^2$ and concave, and relying on Lemma 6.2.1 we obtain

$$df\left(\left|Z_t^i\right|\right) = f'\left(\left|Z_t^i\right|\right)\left(\left(F\left(X_t^i, \omega_i\right) - F\left(\bar{X}_t^i, \omega_i\right)\right)\cdot e_t^i + A_t^i\right)dt + 4f''\left(\left|Z_t^i\right|\right)\sigma^2\phi_r^2(\left|Z_t^i\right|)dt + dM_t^i, \tag{6.2.4}$$

where, with a slight abuse of notation, f' denotes the left derivative of f and f'' its almost everywhere defined second derivative, and $(M_t^i)_t$ is a continuous martingale (recall f' is bounded). Let us define $\omega : \mathbb{R}^+ \mapsto \mathbb{R}^+$ by

$$\omega(r) := \sup_{s\in[0,r]} s\kappa_-(s).$$

Relying on Assumption 6.2, (6.1.4) and (6.2.1) we then get the following inequality:

$$\begin{aligned}
\left(F\left(X_t^i,\omega_i\right) - F\left(\bar{X}_t^i,\omega_i\right)\right)\cdot e_t^i f'\left(\left|Z_t^i\right|\right) + 4f''\left(\left|Z_t^i\right|\right)\sigma^2\phi_r^2(\left|Z_t^i\right|)&\\
\leqslant -\left|Z_t^i\right|\kappa\left(\left|Z_t^i\right|\right)f'\left(\left|Z_t^i\right|\right) + 4f''\left(\left|Z_t^i\right|\right)\sigma^2\phi_r^2(\left|Z_t^i\right|)&\\
\leqslant -2c\sigma^2 f\left(\left|Z_t^i\right|\right)\phi_r^2(\left|Z_t^i\right|) - \left|Z_t^i\right|\kappa\left(\left|Z_t^i\right|\right)f'\left(\left|Z_t^i\right|\right)\phi_s^2(\left|Z_t^i\right|)&\\
\leqslant -2c\sigma^2 f\left(\left|Z_t^i\right|\right)\phi_r^2(\left|Z_t^i\right|) + \omega(\delta)&\\
\leqslant -2c\sigma^2 f\left(\left|Z_t^i\right|\right) + \omega(\delta) + 2c\sigma^2 f(\delta).&
\end{aligned}$$

We deduce that there exists an adapted process K^i satisfying

$$K_t^i \leqslant \omega(\delta) + 2\sigma^2 c f(\delta),$$

and such that for all $\tilde{\kappa} \in [0, 2\sigma^2 c]$

$$\begin{aligned}
d(e^{(2c\sigma^2-\tilde{\kappa})t} f\left(\left|Z_t^i\right|\right) =& e^{(2c\sigma^2-\tilde{\kappa})t} df\left(\left|Z_t^i\right|\right) + (2c\sigma^2 - \tilde{\kappa}e^{(2c\sigma^2-\tilde{\kappa})t} f\left(\left|Z_t^i\right|\right)dt\\
=& e^{(2c\sigma^2-\tilde{\kappa})t}(-\tilde{\kappa}f\left(\left|Z_t^i\right|\right) + K_t^i + A_t^i)dt + e^{(2c\sigma^2-\tilde{\kappa})t} dM_t^i,
\end{aligned} \tag{6.2.5}$$

with A_t^i given in Lemma 6.2.1. The next step is to deal with A_t^i. We have

$$\begin{aligned}
&\left|\frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)}\Gamma\left(X_t^i,\omega_i,X_t^j,\omega_j\right) - p\int\Gamma(\bar{X}_t^i,\omega_i,x,\omega)\bar{\rho}_t(dx,d\omega)\right|\\
&\quad\leqslant \left|\frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)}\left(\Gamma\left(X_t^i,\omega_i,X_t^j,\omega_j\right) - \Gamma\left(\bar{X}_t^i,\omega_i,\bar{X}_t^j,\omega_j\right)\right)\right|\\
&\quad\quad + \left|\frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)}\left(\Gamma\left(\bar{X}_t^i,\omega_i,\bar{X}_t^j,\omega_j\right) - \int\Gamma(\bar{X}_t^i,\omega_i,x,\omega)\bar{\rho}_t(dx,d\omega)\right)\right|\\
&\quad\quad + \left|\left(\frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)} - p\right)\int\Gamma(\bar{X}_t^i,\omega_i,x,\omega)\bar{\rho}_t(dx,d\omega)\right|\\
&\quad = I_{1,i} + I_{2,i} + I_{3,i}.
\end{aligned}$$

We deal with each of these three terms individually.

Dealing with $I_{1,i}$: Lipschitz continuity of Γ. Using Assumption 6.3,

$$I_{1,i} = \left|\frac{\alpha_N}{N}\sum_{j=1}^N \xi_{i,j}^{(N)}\left(\Gamma\left(X_t^i,\omega_i,X_t^j,\omega_j\right) - \Gamma\left(\bar{X}_t^i,\omega_i,\bar{X}_t^j,\omega_j\right)\right)\right|$$

$$\begin{aligned}
&\leqslant \frac{L_\Gamma \alpha_N}{N} \sum_{j=1}^N \xi_{i,j}^{(N)} \left(f(|X_t^i - \bar{X}_t^i|) + f(|X_t^j - \bar{X}_t^j|) \right) \\
&= L_\Gamma f(|Z_t^i|) \frac{\alpha_N}{N} \sum_{j=1}^N \xi_{i,j}^{(N)} + \frac{L_\Gamma \alpha_N}{N} \sum_{j=1}^N \xi_{i,j}^{(N)} f(|Z_t^j|).
\end{aligned}$$

We then deduce, relying on Assumption 6.1,

$$\frac{1}{N} \sum_{i=1}^N I_{1,i} \leqslant \frac{L_\Gamma}{N} \sum_{i=1}^N f(|Z_t^i|) \frac{\alpha_N}{N} d_i^{(N)} + \frac{L_\Gamma}{N} \sum_{j=1}^N f(|Z_t^j|) \frac{\alpha_N}{N} \tilde{d}_j^{(N)} \leqslant \frac{L_\Gamma D_{N,g}}{N} \sum_{i=1}^N f(|Z_t^i|).$$

Dealing with $I_{2,i}$: some law of large numbers. Let us denote

$$\bar{\Gamma}(x,\omega,y,\omega') = \Gamma(x,\omega,y,\omega') - \int \Gamma(x,\omega,z,\tilde{\omega}) \bar{\rho}_t(dz,d\tilde{\omega}).$$

After expansion, we obtain (recall that we have made the hypothesis $\xi_{i,i}^{(N)} = 0$)

$$\begin{aligned}
I_{2,i}^2 =& \left| \frac{\alpha_N}{N} \sum_{j=1}^N \xi_{i,j}^{(N)} \bar{\Gamma}\left(\bar{X}_t^i, \omega_i, \bar{X}_t^j, \omega_j\right) \right|^2 \\
=& \frac{\alpha_N^2}{N^2} \sum_{j=1, j\neq i}^N \xi_{i,j}^{(N)} \bar{\Gamma}\left(\bar{X}_t^i, \omega_i, \bar{X}_t^j, \omega_j\right)^2 \\
&+ \frac{\alpha_N^2}{N^2} \sum_{j,k=1, j,k\neq i, j\neq k}^N \xi_{i,j}^{(N)} \xi_{i,k}^{(N)} \bar{\Gamma}\left(\bar{X}_t^i, \omega_i, \bar{X}_t^j, \omega_j\right) \bar{\Gamma}\left(\bar{X}_t^i, \omega_i, \bar{X}_t^k, \omega_k\right).
\end{aligned}$$

The expectation of the last term conditioned to $(\bar{X}_t^i, \omega^i)$ is equal to 0, and thus, relying in particular on Assumption 6.3-2 and Lemma 6.1.1,

$$\begin{aligned}
&\mathbb{E}\left(I_{2,i} \middle| \bar{X}_t^i, \omega_i \right) \\
&\quad\leqslant \mathbb{E}\left(I_{2,i}^2 \middle| \bar{X}_t^i, \omega_i \right)^{1/2} \\
&\quad= \mathbb{E}\left(\frac{\alpha_N^2}{N^2} \sum_{j=1, j\neq i}^N \xi_{i,j}^{(N)} \bar{\Gamma}\left(\bar{X}_t^i, \omega_i, \bar{X}_t^j, \omega_j\right)^2 \middle| \bar{X}_t^i, \omega_i \right)^{1/2} \\
&\quad\leqslant \mathbb{E}\left(\frac{3\alpha_N^2}{N^2} \sum_{j=1, j\neq i}^N \xi_{i,j}^{(N)} \left| \Gamma\left(\bar{X}_t^i, \omega_i, \bar{X}_t^j, \omega_j\right) - \Gamma(0,\omega_i,0,\omega_j) \right|^2 \middle| \bar{X}_t^i, \omega_i \right)^{1/2} \\
&\qquad+ \mathbb{E}\left(\frac{3\alpha_N^2}{N^2} \sum_{j=1, j\neq i}^N \xi_{i,j}^{(N)} \left| \int \left(\Gamma(0,\omega_i,0,\omega_j) - \Gamma(0,\omega_i,0,\omega)\right) \bar{\rho}_t(dx,d\omega) \right|^2 \middle| \bar{X}_t^i, \omega_i \right)^{1/2} \\
&\qquad+ \mathbb{E}\left(\frac{3\alpha_N^2}{N^2} \sum_{j=1, j\neq i}^N \xi_{i,j}^{(N)} \left| \int \left(\Gamma(0,\omega_i,0,\omega) - \Gamma(\bar{X}_t^i,\omega_i,x,\omega)\right) \bar{\rho}_t(dx,d\omega) \right|^2 \middle| \bar{X}_t^i, \omega_i \right)^{1/2}
\end{aligned}$$

$$\leqslant 2\left[\frac{6L_\Gamma^2 C_f^2\alpha_N^2}{N^2}d_i^{(N)}\left(|\bar{X}_t^i|^2+\int|x|^2\bar{\rho}_t(dx,d\omega)\right)\right]^{1/2}+\left[\frac{12L_\Gamma^2\alpha_N^2}{N^2}d_i^{(N)}\int|\omega|^2\bar{\rho}_t(dx,d\omega)\right]^{1/2}.$$

We deduce, recalling Lemma 6.1.2,

$$\mathbb{E}\left(I_{2,i}\right)\leqslant 2\sqrt{3}L_\Gamma(2C_f\bar{\mathcal{C}}_2^{1/2}+C_{dis}^{1/2})\sqrt{\frac{\alpha_N D_{N,g}}{N}}.$$

Remark that if Γ satisfies Assumption 6.3-2-bis, then we simply have

$$\mathbb{E}\left(I_{2,i}\right)\leqslant 2L_\infty\sqrt{\frac{\alpha_N D_{N,g}}{N}}.$$

Dealing with $I_{3,i}$: convergence of the graph. We immediately get

$$\mathbb{E}\left(I_{3,i}\right)\leqslant I_{N,g}\mathbb{E}\left(\left|\int\Gamma(\bar{X}_t^i,\omega_i,x,\omega)\bar{\rho}_t(dx,d\omega)\right|\right),$$

and thus, if Assumption 6.3-2-bis holds, this directly implies $\mathbb{E}\left(I_{3,i}\right)\leqslant L_\infty I_{N,g}$. Otherwise, if Assumption 6.3-2 holds, we obtain

$$\begin{aligned}\mathbb{E}\left(\left|\int\Gamma(\bar{X}_t^i,\omega_i,x,\omega)\bar{\rho}_t(dx,d\omega)\right|\right)\leqslant&\mathbb{E}\left(\int\left|\Gamma(\bar{X}_t^i,\omega_i,x,\omega)-\Gamma(0,\omega_i,0,\omega)\right|\bar{\rho}_t(dx,d\omega)\right)\\&+\mathbb{E}\left(\int\left|\Gamma(0,\omega_i,0,\omega)-\Gamma(0,0,0,0)\right|\bar{\rho}_t(dx,d\omega)\right)\\\leqslant&2L_\Gamma\left(C_f\mathbb{E}\left|\bar{X}_t^i\right|+\mathbb{E}\left|\omega_i\right|\right).\end{aligned}$$

So, using Lemma 6.1.2, we get

$$\mathbb{E}\left(I_{3,i}\right)\leqslant 2L_\Gamma\left(C_f\bar{\mathcal{C}}_2^{1/2}+\mathcal{C}_{2,\omega}^{1/2}\right)I_{N,g}.$$

Conclusion Recalling (6.2.5) and choosing $\tilde{\kappa}=L_\Gamma D_{N,g}$ we obtain

$$d\left(e^{(2\sigma^2c-L_\Gamma D_{N,g})t}f(|Z_t^i|)\right)=e^{(2\sigma^2c-L_\Gamma D_{N,g})t}\tilde{K}_t^i dt+e^{(2\sigma^2c-L_\Gamma D_{N,g})t}dM_t^i,$$

where there exists a constant C_0, depending on the parameters as well as possibly on $\bar{\rho}_0$, but that do not depend on N and on the graph, such that

$$\frac{1}{N}\sum_{i=1}^N\mathbb{E}\tilde{K}_t^i\leqslant C_0L_\Gamma\left(\sqrt{\frac{\alpha_N D_{N,g}}{N}}+I_{N,g}\right)+\omega(\delta)+2\sigma^2cf(\delta).$$

Then

$$\begin{aligned}\mathbb{E}&\left(\frac{e^{(2\sigma^2c-L_\Gamma D_{N,g})t}}{N}\sum_{i=1}^N f\left(\left|Z_t^i\right|\right)\right)-\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N f\left(\left|Z_0^i\right|\right)\right)\\&\leqslant\frac{e^{(2\sigma^2c-L_\Gamma D_{N,g})t}-1}{2\sigma^2c-L_\Gamma D_{N,g}}\left(C_0L_\Gamma\left(\sqrt{\frac{\alpha_N D_{N,g}}{N}}+I_{N,g}\right)+\omega(\delta)+2\sigma^2cf(\delta)\right),\end{aligned}$$

i.e.

$$\begin{aligned}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N\left|X_t^i-\bar{X}_t^i\right|\right)\leqslant&\frac{C_fe^{-(2\sigma^2c-L_\Gamma D_{N,g})t}}{c_f}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N\left|X_0^i-\bar{X}_0^i\right|\right)\\&+\frac{1}{c_f(2\sigma^2c-L_\Gamma D_{N,g})}\left(C_0L_\Gamma\left(\sqrt{\frac{\alpha_ND_{N,g}}{N}}+I_{N,g}\right)+\omega(\delta)+2\sigma^2cf(\delta)\right)\end{aligned}$$

Thus, denoting $\bar{\mu}_t^N$ the empirical measure associated with the system of independent non-linear particles $\left((\bar{X}_t^1,\omega_1),...,(\bar{X}_t^N,\omega_N)\right)$, we obtain, for $c_\Gamma=\sigma^2c$ and $L_\Gamma\leqslant c_\Gamma/D_{N,g}$,

$$\begin{aligned}\mathbb{E}\mathcal{W}_1\left(\mu_t^N,\bar{\mu}_t^N\right)\leqslant&\frac{C_fe^{-\sigma^2ct}}{c_f}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N\left|X_0^i-\bar{X}_0^i\right|\right)\\&+\frac{1}{c_f\sigma^2c}\left(C_0L_\Gamma\left(\sqrt{\frac{\alpha_ND_{N,g}}{N}}+I_{N,g}\right)+\omega(\delta)+2\sigma^2cf(\delta)\right).\end{aligned}$$

Notice that

$$\begin{aligned}\mathcal{W}_1\left(\mu_0^N,\bar{\mu}_0^N\right)=&\mathcal{W}_1\left(\frac{1}{N}\sum_{i=1}^N\delta_{(X_0^i,\omega_i)},\frac{1}{N}\sum_{i=1}^N\delta_{(\bar{X}_0^i,\omega_i)}\right)\\=&\min_{\tau\text{ permutation}}\left\{\frac{1}{N}\sum_{i=1}^N|X_0^i-\bar{X}_0^{\tau(i)}|+|\omega_i-\omega_{\tau(i)}|\right\}\\=&\frac{1}{N}\sum_{i=1}^N|X_0^i-\bar{X}_0^i|,\end{aligned}$$

as both terms to minimize are minimal for τ the identity. By having $\delta\to0$, we thus have

$$\begin{aligned}\mathbb{E}\mathcal{W}_1\left(\mu_t^N,\bar{\mu}_t^N\right)\leqslant&\frac{C_fe^{-\sigma^2ct}}{c_f}\mathbb{E}\mathcal{W}_1\left(\mu_0^N,\bar{\mu}_0^N\right)\\&+\frac{1}{c_f\sigma^2c}\left(C_0L_\Gamma\left(\sqrt{\frac{\alpha_N}{N}}+I_{N,g}\right)\right).\end{aligned}$$

Since $\left((\bar{X}_t^1,\omega_1),...,(\bar{X}_t^N,\omega_N)\right)$ are N independent random variables with law $\bar{\rho}_t$ by construction, and since $\bar{\rho}_t$ admits a second moment, Theorem 1 of [75] yields the existence of a constant C, depending only on the dimensions d and d', such that

$$\mathbb{E}\mathcal{W}_1\left(\bar{\mu}_t^N,\bar{\rho}_t\right)\leqslant C\left(C_{\text{dis}}+\bar{\mathcal{C}}_2\right)^{1/2}\begin{cases}N^{-\frac{1}{2}}+N^{-\frac{1}{3}}&\text{if }d+d'=1,\\N^{-\frac{1}{2}}\log(1+N)+N^{-\frac{1}{3}}&\text{if }d+d'=2,\\N^{-\frac{1}{2}}+N^{-\frac{1}{d+d'}}&\text{if }d+d'\geqslant3.\end{cases}$$

The convergence rates could be improved (with respective rates $N^{-\frac{1}{2}}$, $N^{-\frac{1}{2}}\log(1+N)$ and $N^{-\frac{1}{d+d'}}$) provided we can prove uniform in time bounds on a moment of order $q>2$ for $\bar{\rho}_t$. This can be done, but requires a similar great moment assumption on the initial distribution $\bar{\rho}_0$.

Chapter 7

Some remarks on the effect of the Random Batch Method on phase transition

> And now for something completely different.
>
> Monty Python, *Monty Python's Flying Circus* (1969-1974), written by the Monty Python.

Outline of the current chapter

Joint work with Arnaud Guillin and Pierre Monmarché.

Ongoing work.

Abstract: In this chapter, we focus on two toy models : the *Curie-Weiss* model and the system of N particles in linear interactions in a *double well confining potential*. Both models, which have been extensively studied, describe a large system of particles with a mean-field limit that admits a phase transition. We are concerned with the numerical simulation of these particle systems. To deal with the quadratic complexity of the numerical scheme, corresponding to the computation of the $O(N^2)$ interactions per time step, the *Random Batch Method* (RBM) has been suggested. It consists in randomly (and uniformly) dividing the particles into batches of size $p > 1$, and computing the interactions only within each batch, thus reducing the numerical complexity to $O(Np)$ per time step. The convergence of this numerical method has been proved in other works.

This work is motivated by the observation that the RBM, via the random constructions of batches, artificially adds noise to the particle system. The goal of this chapter is to study the effect of this added noise on the phase transition of the nonlinear limit, and more precisely we study the *effective dynamics* of the models to show how the critical temperature decreases with the RBM.

Acknowledgements: This work has been (partially) supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency.

7.1 Introduction

7.1.1 Motivation

Consider a system of N particles $(X^i)_{i\in\{1,...,N\}}$ in interaction

$$dX_t^i = -\nabla U(X_t^i)dt - \frac{1}{N-1}\sum_{j\neq i}\nabla W(X_t^i - X_t^j)dt + \sqrt{2\sigma}dB_t^i, \tag{IPS}$$

where for all $i \in \{1, ..., N\}$ and $t \geqslant 0$ we have $X_t^i \in \mathbb{R}^d$. U and W are two twice continuously differentiable functions, respectively called *confining potential* and *interaction potential*, $\sigma > 0$ is a *diffusion coefficient* or *temperature*, and $(B^i)_i$ are independent d-dimensional Brownian motions. The name (IPS) refers to *Interacting Particle System*.

It is well known (see [45, 46] and references therein) that, under suitable assumptions on U and W, the particle system (IPS) converges as $N \to \infty$ towards its nonlinear mean-field limit, a stochastic differential equation (SDE) of *McKean-Vlasov* type

$$\begin{cases} d\bar{X}_t = -\nabla U(\bar{X}_t)dt - \nabla W * \bar{\rho}_t(\bar{X}_t)dt + \sqrt{2\sigma}dB_t, \\ \bar{\rho}_t = \text{Law}(\bar{X}_t). \end{cases} \tag{NL}$$

Here, the name (NL) refers to *Nonlinear Limit*, and this equation arises in the modelling of granular media [40].

The quantitative link between of (IPS) and (NL) can be exploited in various ways. On one hand, as it was historically motivated, the study of (way too) large systems of particles cannot be feasible, and boiling it down to the study of the nonlinear limit yields exploitable results. On the other hand, one can see (IPS) as an approximation of (NL), and in particular an approximation that can be numerically simulated. Consider the *Euler-Maruyama* scheme associated to (IPS) with a timestep $\delta > 0$

$$\begin{cases} X_{t+1}^{i,\delta} = X_t^{i,\delta} - \delta \nabla U(X_t^{i,\delta}) - \frac{\delta}{N-1} \sum_{j \neq i} \nabla W(X_t^{i,\delta} - X_t^{j,\delta}) + \sqrt{2\sigma\delta} G_t^i, \\ G_t^i \ \text{ i.i.d } \ \sim \mathcal{N}(0,1), \quad t \in \mathbb{N}. \end{cases} \tag{D-IPS}$$

Its name (D-IPS) comes from *Discrete - Interacting Particle System.* The convergence of (D-IPS) towards (NL) has been extensively studied : with bounded Lipschitz coefficients [26], with Hölder continuous coefficients [7], non-Lipschitz coefficients [59]. The quantitative convergence of the implicit Euler-Maruyama scheme can also be found in [132].

Notice that this numerical scheme requires $O(N^2)$ operations per time step, corresponding to the total number of interactions of pairs $(i,j)_{i,j \in \{1,...,N\}}$. To cope with this possibly limiting complexity, several works have suggested using the *Random Batch Method* (RBM) (see for instance [102]), motivated by the Stochastic Gradient Langevin Dynamics [174].

Consider, for a time step $t \in \mathbb{N}$, a partition $\mathcal{P}_t = \left(\mathcal{P}_t^1, ..., \mathcal{P}_t^{N/p}\right)$ of $\{1, ..., N\}$ into N/p subsets of size $p > 1$, assuming for the sake of simplicity that N is a multiple of p, and define

$$\mathcal{C}_t^i = \left\{ j \in \{1, .., N\} \text{ s.t. } \exists l \in \{1, ..., N/p\}, i, j \in \mathcal{P}_t^l \right\}. \tag{7.1.1}$$

In other words, $\mathcal{C}_t^i$ is the set of indexes that are in the same subset as i at time step t, with the convention $i \in \mathcal{C}_t^i$. We now consider the following numerical scheme

$$\begin{cases} Y_{t+1}^{i,\delta,p} = Y_t^{i,\delta,p} - \delta \nabla U(Y_t^{i,\delta,p}) - \frac{\delta}{p-1} \sum_{j \in \mathcal{C}_t^i \setminus \{i\}} \nabla W(Y_t^{i,\delta,p} - Y_t^{j,\delta,p}) + \sqrt{2\sigma\delta} G_t^i, \\ G_t^i \ \text{ i.i.d } \ \sim \mathcal{N}(0,1), \quad i \in \{1, ..., N\}, \quad t \in \mathbb{N}, \end{cases} \tag{D-RB-IPS}$$

where for each time step t the partition $\mathcal{P}_t$ is random and each partition has the same probability of occurring. The name (D-RB-IPS) refers to *Discrete - Random Batch - Interacting Particle System.* The convergence of (D-RB-IPS) towards (NL) can be found in [101, 104, 177].

The idea of using random batches has been shown to be efficient for computing the evolution of large interacting system of quantum particles [81], of particles with Coulomb interactions in molecular dynamics [103], but also for Markov Chain Monte Carlo [124], or for solving PDEs [39, 123]. See also references therein.

The starting point of this work is the following observation : the RBM, via the random construction of a partition of $\{1, ..., N\}$, artificially adds noise (or temperature) to a system. We thus ask the following question :

Does the critical temperature of (the mean-field limit of) a system of interacting particles admitting a phase transition decreases when considering a version with random batches ? If so, can we quantify it ?

To partially answer this question, we focus on two specific types of particle systems for which the mean-field limit admits a phase transition : the first one is the *Curie-Weiss* model and the second one is the system (IPS) with attractive and quadratic interaction potential W and the *double well confining potential* U.

The nonlinear mean-field limits of both models admit, as we will discuss, a phase transition occurring at a certain critical parameter. We consider a version with random batches of size p of each system, consider the limit as $N \to \infty$ (with fixed p) towards a nonlinear model, and then study the phase transition of said limit.

7.1.2 The Curie-Weiss model

The classical system. The Curie-Weiss model is, and it is the reason we start by studying it, one of the most simple system admitting a phase transition. Consider N spins, given by a configuration $\sigma = (\sigma_1, ..., \sigma_N)$, and $\Omega_N = \{-1, 1\}^N$ the set of possible configurations for the system. On this system we consider the following Hamiltonian

$$\forall \sigma \in \Omega_N, \quad H_N(\sigma) = -\frac{1}{2N} \sum_{i,j} \sigma_i \sigma_j. \tag{7.1.2}$$

Intuitively, each spin will tend to align with the others. It is a mean field model as H_N only depends in reality on the mean magnetization $m_N(\sigma) := \frac{1}{N} \sum_{i=1}^N \sigma_i$, by

$$H_N(\sigma) = -\frac{N}{2} m_N(\sigma)^2.$$

The evolution for $(\sigma(n))_{n\geqslant 0}$ in Ω_N is the following : at each discrete time step, a spin is chosen uniformly among the N possible spins. Let us denote i this spin, and $\sigma' = (\sigma'_1, ..., \sigma'_N)$ the configuration such that for all $j \neq i$, $\sigma'_j = \sigma(n)_j$, and $\sigma'_i = -\sigma(n)_i$. We accept σ' as the next step of $\sigma(n)$ with probability $\exp\left(-\beta(H_N(\sigma') - H_N(\sigma))_+\right)$ (i.e if the Hamiltonian decreases then with probability 1, otherwise with a positive probability depending on a parameter β), otherwise the system remains at $\sigma(n)$. This parameter β is known as the *inverse temperature*. This yields the following transition probabilities for the Markov chain $(\sigma(n))_{n\geqslant 0}$:

$$p(\sigma, \sigma') = \begin{cases} \frac{1}{N} \exp\left(-\beta(H_N(\sigma') - H_N(\sigma))_+\right) & \text{if } ||\sigma - \sigma'||_1 = 2 \\ 0 & \text{if } ||\sigma - \sigma'||_1 > 2 \\ 1 - \sum_{\eta \neq \sigma} p(\sigma, \eta) & \text{if } \sigma' = \sigma \end{cases}$$

This dynamics $(\sigma(n))_{n\geqslant 0}$, which is an irreducible and aperiodic Markov chain on a finite state space Ω_N, is reversible with respect to the Gibbs measure

$$\mu_{\beta,N}(\sigma) = \frac{1}{Z_{\beta,N}} \exp(-\beta H_N(\sigma)), \tag{7.1.3}$$

where $Z_{\beta,N}$ is a normalizing constant. Instead of studying the dynamics of σ, we look at the mean magnetization $m_N(n) = m_N(\sigma(n))$, which is still a Markov chain. This quantity, at each time step, can only increase or decrease by $\frac{2}{N}$, and the transition probabilities are given by

$$r(m, m') = \begin{cases} \frac{1-m}{2} \exp\left(-\frac{\beta N}{2}(m^2 - m'^2)_+\right) & \text{if } m' = m + \frac{2}{N} \\ \frac{1+m}{2} \exp\left(-\frac{\beta N}{2}(m^2 - m'^2)_+\right) & \text{if } m' = m - \frac{2}{N} \\ 1 - r\left(m, m + \frac{2}{N}\right) - r\left(m, m - \frac{2}{N}\right) & \text{if } m' = m \\ 0 & \text{otherwise.} \end{cases} \tag{7.1.4}$$

Likewise, this dynamics is reversible with respect to the Gibbs measure

$$\nu_{\beta,N}(m) = \frac{1}{Z_{\beta,N}} \binom{N}{\frac{1+m}{2}N} \exp\left(\frac{\beta N m^2}{2}\right).$$

Many works (see for instance [53, 72, 119], the classical reference that is Chapter 4 of [71] or more recently Chapter 2 of [77]) have studied Large Deviation Principles for this system, and have shown that there exists a critical inverse temperature $\beta_c = 1$. For the sake of completeness, and because the method will be similar in the case with random batches, we give a proof in Section 7.2.1 of the phase transition happening in the following sense : the process $M_t^{(N)} = m_N(\lfloor Nt \rfloor)$ weakly converges to the solution of an ordinary differential equation (ODE). For $\beta > 1$, the limit ODE admits three equilibrium states, and for $\beta \leqslant 1$ only one. In both cases, 0 is an equilibrium state, and is stable in the case $\beta \leqslant 1$ and unstable in the case $\beta > 1$.

The Curie-Weiss model with random batches. We then consider the same system, but using the Random Batch Method. At each time step, the chosen spin no longer evolves according to the entire system, but according to a subset of p spins containing the chosen spin.

We thus consider a new evolution for $(\sigma^p(n))_{n\geqslant 0}$ in Ω_N, where σ^p denotes the new sequence of spin configurations. At each discrete time step, a spin is chosen uniformly among the N possible spins. Let us denote it i, and $\sigma' = (\sigma'_1, ..., \sigma'_N)$ the configuration such that for all $j \neq i$, $\sigma'_j = \sigma^p(n)_j$, and $\sigma'_i = -\sigma^p(n)_i$. We then sample a subset of $\{1, ..., N\}$ of size p containing i, denoted $\mathcal{C}^{i,p}$, uniformly over such subsets, and accept σ' as the next step of $\sigma^p(n)$ with probability $\exp\left(-\beta(H_{N,p}(\sigma', \mathcal{C}^{i,p}) - H_{N,p}(\sigma^p(n), \mathcal{C}^{i,p}))_+\right)$, where

$$H_{N,p}(\sigma, \mathcal{C}^{i,p}) = -\frac{1}{2p} \sum_{j,k\in\mathcal{C}^{i,p}} \sigma_j \sigma_k. \tag{7.1.5}$$

Likewise, we may study this system in terms of its magnetization, denoted $(m_{N,p}(n))_n$, for which we can explicitly write the transition probabilities (see Lemma 7.2.1).

This system resembles to some extent the *dilute Curie-Weiss model* [29], in which the spins interact according to an Erdős-Rényi random graph with edge probability $\tilde{p} = \frac{p}{N} \in]0, 1[$, the main difference being that the "graph", in our case, is modified at each time step and there are exactly $p - 1$ spins interacting with a given one.

Studying the Curie-Weiss model with random batches, which is done in Section 7.2.2, yields the following results.

Theorem 7.1.1. *Let $p \in \mathbb{N} \setminus \{0, 1\}$ and $\beta > 0$.*

- *Define*

$$\begin{aligned}
S_1^{p,\beta}(m) &= \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^k \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} \\
S_2^{p,\beta}(m) &= \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^k \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+}, \\
f_p(\beta, m) &= \left(S_1^{p,\beta}(m) - S_2^{p,\beta}(m)\right) - m\left(S_1^{p,\beta}(m) + S_2^{p,\beta}(m)\right).
\end{aligned}$$

  The process $M_t^{(N,p)} = m_{N,p}(\lfloor Nt \rfloor)$, i.e the magnetization rescaled in time, weakly converges as $N \to \infty$ to the solution of the ODE

$$\frac{d}{dt} m(t) = f_p(\beta, m(t)). \tag{7.1.6}$$

  For all $\beta > 0$, 0 is an equilibrium state for the solution of (7.1.6).

- *For $p \in \{2, 3\}$, 0 is the unique equilibrium state, and it is stable.*

- *For $p \geqslant 4$, there exists $\beta_{c,p}$ such that for all $\beta > \beta_{c,p}$, the equilibrium state 0 is unstable, and for all $\beta \leqslant \beta_{c,p}$ it is stable. Furthermore, we have the estimate*

$$\beta_{c,p} = 1 + \sqrt{\frac{2}{p\pi}} + o\left(\frac{1}{\sqrt{p}}\right). \tag{7.1.7}$$

This theorem thus gives a first answer to the main question of the chapter : the RBM does increase the critical inverse temperature of the system (i.e decreases the critical temperature).

7.1.3 Numerical scheme and double-well potential

We then go back to the initial motivation concerning numerical scheme for interacting particle systems.

The effective dynamics. Just like we may consider the nonlinear limit of (IPS), we may also consider the limit as $N \to \infty$ of (1.3.9). Define

$$\begin{cases} \bar{Y}^{\delta,p}_{t+1} = \bar{Y}^{\delta,p}_t - \delta\nabla U(\bar{Y}^{\delta,p}_t) - \frac{\delta}{p-1}\sum_{j=1}^{p-1}\nabla W(\bar{Y}^{\delta,p}_t - Y^j) + \sqrt{2\sigma\delta}G_t, \\ G_t \ \text{ i.i.d } \ \sim \mathcal{N}(0,1), \quad (Y^j)_j \ \text{ i.i.d } \ \sim \text{Law}(\bar{Y}^{\delta,p}_t). \end{cases} \tag{D-RB-NL}$$

The name (D-RB-NL) stands for *Discrete - Random Batch - Nonlinear Limit.* The convergence of (1.3.9) towards (D-RB-NL) can be found in [100]. The proof relies on a coupling method, noticing that, as $N \to \infty$, the probability of constructing batches of fixed size p in (1.3.9) with independent and identically distributed particles goes to 1, thus giving a convergence in total variation distance.

We then, in the spirit of [157], construct a continuous process, parameterized by the timestep and the batch size, which is closer to the numerical scheme (1.3.9) than the target (NL). In the dynamics of (D-RB-NL), notice that

$$\mathbb{E}\left(\frac{\delta}{p-1}\sum_{j=1}^{p-1}\nabla W\left(\bar{Y}^{\delta,p}_t - Y^j\right)\Big|\bar{Y}^{\delta,p}_t\right) = \delta\nabla W * \bar{\rho}^{\delta,p}_t(\bar{Y}^{\delta,p}_t),$$

and

$$\begin{aligned}\text{Var}\left(\frac{\delta}{p-1}\sum_{j=1}^{p-1}\nabla W\left(\bar{Y}^{\delta,p}_t - Y^j\right)\Big|\bar{Y}^{\delta,p}_t\right) =& \frac{\delta^2}{p-1}\text{Var}_{\bar{\rho}_k}\left(\nabla W(\bar{Y}^{\delta,p}_t - \cdot)\Big|\bar{Y}^{\delta,p}_t\right) \\ =& \frac{\delta^2}{p-1}\left((\nabla W)^2 * \bar{\rho}^{\delta,p}_t(\bar{Y}^{\delta,p}_t) - (\nabla W * \bar{\rho}^{\delta,p}_t(\bar{Y}^{\delta,p}_t))^2\right).\end{aligned}$$

Denoting $\Sigma(x,\rho) = (\nabla W)^2 * \rho(x) - (\nabla W * \rho(x))^2$, we thus have

$$\frac{\delta}{p-1}\sum_{j=1}^{p-1}\nabla W\left(\bar{Y}^{\delta,p}_t - Y^j\right) = \delta\nabla W * \bar{\rho}^{\delta,p}_t(\bar{Y}^{\delta,p}_t) + \delta\sqrt{\frac{\Sigma(\bar{Y}^{\delta,p}_t, \bar{\rho}^{\delta,p}_t)}{p-1}}G_p, \tag{7.1.8}$$

where G_p converges in law as $p \to \infty$, via the Central Limit Theorem, to a random variable $G \sim \mathcal{N}(0,1)$. This added random variable G_p, close to a normal distribution, suggests we should consider for given $\delta > 0$ and $p \in \mathbb{N}$, the following non-linear SDE, that we call the *effective* *dynamics*:

$$\begin{cases} d\bar{X}_t^{e,\delta,p} = -\nabla U(\bar{X}_t^{e,\delta,p})dt - \nabla W * \bar{\rho}_t^{e,\delta,p}(\bar{X}_t^{e,\delta,p})dt + \left(2\sigma + \frac{\delta}{p-1}\Sigma(\bar{X}_t^{e,\delta,p}, \bar{\rho}_t^{e,\delta,p})\right)^{1/2} dB_t, \\ \bar{\rho}_t^{e,\delta,p} = \mathrm{Law}(\bar{X}_t^{e,\delta,p}). \end{cases} \tag{Eff}$$

Remark 7.1.1. *Let us quickly insist on the fact that the formal justification for* (Eff) *given above using the Central Limit Theorem is far from correct ! In this work we indeed consider fixed values of p, and we do not assume that the error G_p is Gaussian. In reality, the Gaussian random variable appears when summing the interactions over several time steps, and as a consequence* (Eff) *should not be understood as an approximation when $p \to \infty$ but when $\delta \to 0$.*

Such dynamics are also known as modified equations in various works considering the backward error analysis of SDEs [161, 178], improving upon a technique that had already provided a better understanding of the numerical methods for ODEs.

Our goal now is to study this dynamics. A better justification of this effective dynamics is one of the remaining questions of this work, as explained in Section 7.4.

The double well confining potential. We now choose in (NL) the dimension to be $d = 1$ and the potentials

$$U(x) = \frac{x^4}{4} - \frac{x^2}{2}, \quad W(x) = L_W \frac{x^2}{2} \quad \text{with } L_W > 0. \tag{7.1.9}$$

Recall the following result adapted from [168].

Theorem 7.1.2 (Theorem 2.1 of [168])**.** *For U and W given by* (7.1.9)*, there exists $\sigma_c > 0$ such that*

- *For all $\sigma \geqslant \sigma_c$, there exists a unique stationary distribution $\mu_{\sigma,0}$ for* (NL)*. Furthermore, $\mu_{\sigma,0}$ is symmetric.*
- *For all $\sigma < \sigma_c$, there exist three stationary distributions for* (NL)*. One is symmetric, also denoted $\mu_{\sigma,0}$, and the other two, denoted $\mu_{\sigma,+}$ and $\mu_{\sigma,-}$, satisfy $\pm \int x d\mu_{\sigma,\pm}(dx) > 0$.*

By convention, in the case $\sigma \geqslant \sigma_c$, we may denote $\mu_\sigma = \mu_{\sigma,\pm} = \mu_{\sigma,0}$.

Our goal is now to study the stationary distribution(s) for the effective dynamics (Eff) in the specific case of the double-well potential (7.1.9). We wish to understand if, similarly as Theorem 7.1.2, there exists a phase transition, and if so compare the critical parameters. We thus prove in Section 7.3 the following theorem.

Theorem 7.1.3. *Let $\sigma_0 \in]0, \sigma_c[$ where σ_c is defined in Theorem 7.1.2. For U and W given by* (7.1.9)*, there exists $c_0 > 0$ such that for all (δ, p) satisfying $\frac{\delta}{p-1} \leqslant c_0$, denoting*

$$\sigma_c^{eff} = \sigma_c \left(1 - \frac{\delta L_W}{2(p-1)}\right), \tag{7.1.10}$$

we have the following phase transition for the dynamics (Eff)

- *For all $\sigma \geqslant \sigma_c^{eff}$, there exists a unique stationary distribution $\mu_{\sigma,0}^{\delta,p}$ for* (Eff)*. Furthermore, $\mu_{\sigma,0}^{\delta,p}$ is symmetric.*

- *For all* $\sigma \in [\sigma_0, \sigma_c^{eff}[$, *there exists exactly three stationary distributions for* (Eff). *One is symmetric, also denoted* $\mu_{\sigma,0}^{\delta,p}$, *and the other two, denoted* $\mu_{\sigma,+}^{\delta,p}$ *and* $\mu_{\sigma,-}^{\delta,p}$, *have nonzero means* $\pm \int x d\mu_{\sigma,\pm}^{\delta,p}(x) > 0$.

Remark 7.1.2. *Let us quickly discuss the form of* (7.1.10). *In the specific case of* (7.1.9), *as discussed in Section 7.3, one has* $\Sigma(\bar{X}_t^{e,\delta,p}, \bar{\rho}_t^{e,\delta,p}) = L_W^2 \, Var(\bar{\rho}_t^{e,\delta,p})$. *To insist on the dependence on* σ *rather than* (δ, p), *let us denote, only in this remark,* $\Sigma_\sigma := \Sigma(\bar{X}_t^{e,\delta,p}, \bar{\rho}_t^{e,\delta,p})$.

We will show, but this can be intuitively understood at this stage, that any stationary distribution for (NL) *is also a stationary distribution for* (Eff), *but for a smaller value of* σ. *We thus have to study the stationary distribution at the critical value* σ_c.

As proved in Lemma 7.3.2, the variance of the stationary distribution for (NL) *at the critical value is* $Var(\mu_{\sigma_c,0}) = \frac{\sigma_c}{L_W}$. *By considering the diffusion term in* (Eff), *and considering the added noise* $\frac{\delta}{p-1}\Sigma_{\sigma_c}$, *we intuitively obtain* $2\sigma_c = 2\sigma_c^{eff} + \frac{\delta}{p-1}\Sigma_{\sigma_c}$ *and thus* (7.1.10).

Let us sum up the organization of the chapter.

- The Curie-Weiss model is studied in Section 7.2. We start by showing the phase transition of the classical Curie-Weiss model in Section 7.2.1 since the same ideas will be used in what follows. The study of the Curie-Weiss model with random batches and the proof of Theorem 7.1.1 are then done in Section 7.2.2,
- In Section 7.3 we study the Random Batch Method for interacting particle systems. More specifically we prove Theorem 7.1.3 in the specific case of the double-well potential,
- Finally, we end by stating the remaining questions of this work, mainly concerning the effective dynamics, in Section 7.4.

Notation

For the Curie-Weiss model, with and without random batches:

- $\Omega_N = \{-1, ..., 1\}^N$: the set of possible configurations,
- $\sigma(n) = (\sigma_1(n), ..., \sigma_N(n))$: the spin configuration at time step n,
- β : the inverse temperature,
- β_c : the critical inverse temperature,
- H_N: the Hamiltonian of the Curie-Weiss model given in (7.1.2),
- $m_N(n) = \frac{1}{N}\sum_{i=1}^N \sigma_i(n)$: the magnetization at time step n,
- $r(\sigma, \sigma')$: transition probability for the Markov chain $(m_N(n))_n$, given in (7.1.4).
- $\sigma^p(n) = (\sigma_1^p(n), ..., \sigma_N^p(n))$: the spin configuration of the system with random batches of size p at time step n,
- $H_{N,p}$: the Hamiltonian for the system with random batches of size p, given in (7.1.5),
- $m_{N,p}(n) = \frac{1}{N}\sum_{i=1}^N \sigma_i^p(n)$: the magnetization at time step n for the system with random batches of size p,
- $r_p(m, m')$: transition probability for the Markov chain $(m_{N,p}(n))_n$, given in Lemma 7.2.1.
- $\beta_{c,p}$: the critical inverse temperature for the system with random batches of size p.

For the Random Batch Method for interacting particle system :

- U, W : two twice continuously differentiable functions, respectively the confining potential and the interacting potential (see (IPS)),
- $\sigma > 0$: a diffusion coefficient (see (IPS)),
- $(X_t^i)_{i\in\{1,...,N\}}$: the solution at time $t \in \mathbb{R}^+$ of the interacting particle system (IPS),
- $\bar{X}_t$, $\bar{\rho}_t$: the solution at time $t \in \mathbb{R}^+$ of the nonlinear limit (NL) and its law,
- $\delta > 0$: a timestep used in the various numerical schemes,
- $(X_t^{i,\delta})_{i\in\{1,...,N\}}$: the solution at time step $t \in \mathbb{N}$ of the Euler-Maruyama numerical scheme (D-IPS),
- $p \in \mathbb{N} \setminus \{0, 1\}$: the batch size,
- $\mathcal{P}_t$: the partition of $\{1, ..., N\}$ at time step t into subsets of size p,
- $\mathcal{C}_t^i$: the cluster containing index i at time step t (see (7.1.1)),
- $(Y_t^{i,\delta,p})_{i\in\{1,...,N\}}$: the solution at time step $t \in \mathbb{N}$ of the numerical scheme with random batches (D-RB-IPS),
- $\bar{Y}_t^{\delta,p}$: the solution at time step $t \in \mathbb{N}$ of (D-RB-NL), the nonlinear limit of (D-RB-IPS) as $N \to \infty$,
- $\bar{X}_t^{e,\delta,p}$, $\bar{\rho}_t^{e,\delta,p}$: the effective dynamics (Eff) at time $t \in \mathbb{R}^+$ and its law,
- $\mu_{\sigma,*}$ for $* \in \{0, \pm\}$, σ_c : stationary distributions and critical parameter of (NL) given in Theorem 7.1.2,
- $\mu_{\sigma,*}^{\delta,p}$ for $* \in \{0, \pm\}$, σ_c^{eff} : stationary distributions and critical parameter of (Eff) given in Theorem 7.1.3.

7.2 Understanding the problem on the Curie-Weiss model

In order to get a better grasp on the phenomenon we focus on, we begin by studying one of the simplest model admitting a phase transition : the Curie-Weiss model. In Section 7.2.1, we show how we obtain the value of the critical parameter in the classical case. Then, in Section 7.2.2, we follow the same steps to compute the new critical inverse temperature in the case with random batches.

7.2.1 ...without the Random Batch Method

In order to study this critical inverse temperature, we choose to look at the limit of the dynamics with time step $\frac{1}{N}$ as N goes to infinity. $(m_N(n))_n$ is a discrete-time Markov chain with transition operator $U^{(N)}$ given by $U^{(N)} = \left(U_{i,j}^{(N)}\right)_{0\leqslant i,j\leqslant N}$ where $U_{i,j}^{(N)} = r\left(-1 + \frac{2i}{N}, -1 + \frac{2j}{N}\right)$. We denote $A_N = N\left(U^{(N)} - I\right)$. We have, for all continuously differentiable functions f,

$$A_N f(m) = N\frac{1-m}{2}e^{-\beta N\left(\frac{m^2}{2} - \frac{\left(m+\frac{2}{N}\right)^2}{2}\right)_+}\left(f\left(m + \frac{2}{N}\right) - f(m)\right)$$

$$+ N\frac{1+m}{2}e^{-\beta N\left(\frac{m^2}{2}-\frac{\left(m-\frac{2}{N}\right)^2}{2}\right)_+}\left(f\left(m-\frac{2}{N}\right)-f(m)\right).$$

We thus get

$$\begin{aligned}A_N f(m) =& N\frac{1-m}{2}e^{-2\beta\left(-m-\frac{1}{N}\right)_+}\left(f\left(m+\frac{2}{N}\right)-f(m)\right)\\&+N\frac{1+m}{2}e^{-2\beta\left(m+\frac{1}{N}\right)_+}\left(f\left(m-\frac{2}{N}\right)-f(m)\right)\\=&N\frac{1-m}{2}e^{-2\beta\left(-m-\frac{1}{N}\right)_+}\left(\frac{2}{N}f'(m)+O\left(\frac{1}{N^2}\right)\right)\\&+N\frac{1+m}{2}e^{-2\beta\left(m+\frac{1}{N}\right)_+}\left(-\frac{2}{N}f'(m)+O\left(\frac{1}{N^2}\right)\right)\\=&(1-m)e^{-2\beta\left(-m-\frac{1}{N}\right)_+}f'(m)-(1+m)e^{-2\beta\left(m+\frac{1}{N}\right)_+}f'(m)+O\left(\frac{1}{N}\right)\\&\xrightarrow[N\to\infty]{} f'(m)\left((1-m)e^{-2\beta(-m)_+}-(1+m)e^{-2\beta m_+}\right),\end{aligned}$$

which finally yields

$$A_N f(m) \xrightarrow[N\to\infty]{} 2f'(m)e^{-\beta|m|}\left(\sinh(m\beta)-m\cosh(m\beta)\right).$$

By Theorem 17.28 of [107], the process $M_t^{(N)} = m_N(\lfloor Nt\rfloor)$ weakly converges to the solution of

$$\frac{d}{dt}m(t) = 2e^{-\beta|m(t)|}\left(\sinh(\beta m(t)) - m(t)\cosh(\beta m(t))\right).$$

Denote $f(\beta, m) = 2e^{-\beta|m|}(\sinh(\beta m) - m\cosh(\beta m))$. We have

$$f(\beta, m) = 0 \quad \iff \quad \tanh(\beta m) = m.$$

For $\beta > 1$, the equation $f(\beta, m) = 0$ admits three solutions, and for $\beta \leqslant 1$ only one. Notice that for all $\beta > 0$, $f(\beta, 0) = 0$: 0 is thus always an equilibrium state for the magnetization. Furthermore

$$\begin{aligned}\forall \beta > 0, \forall m \neq 0, \partial_m f(\beta, m) =& -2\beta \mathrm{sign}(m)e^{-\beta|m|}\left(\sinh(\beta m) - m\cosh(\beta m)\right)\\&+2e^{-\beta|m|}\left((\beta-1)\cosh(\beta m) - \beta m\sinh(\beta m)\right),\end{aligned}$$

and, extending by continuity, we have $\partial_m f(\beta, 0) = 2(\beta - 1)$. Therefore, for $\beta > 1$, 0 is unstable as $\partial_m f(\beta, 0) > 0$, and for $\beta \leqslant 1$ it is stable.

Hence a critical inverse temperature $\beta_c = 1$, above which there are two stable equilibrium states, and under which there is only one.

7.2.2 ...with the Random Batch Method

To follow the same steps in the case with random batches, we need to compute the transition operator before finding its limit.

Transition probabilities

Let us start by giving explicit values for the transitions probabilities for the magnetization using the Random Batch Method. The proof, which relies on combinatorics arguments, is double-checked via numerical simulations in Figure 7.1.

Lemma 7.2.1. *In a system of size N, the transition probabilities for the magnetization with random batches of size p are given by*

$$r_p(m,m') = \begin{cases} \frac{1-m}{2}\binom{N-1}{p-1}^{-1}\sum_{k=0}^{p-1}\binom{(\frac{1-m}{2})N-1}{k}\binom{(\frac{1+m}{2})N}{p-1-k}e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} \\ \qquad \text{if } m' = m+\frac{2}{N} \\ \frac{1+m}{2}\binom{N-1}{p-1}^{-1}\sum_{k=0}^{p-1}\binom{(\frac{1-m}{2})N}{k}\binom{(\frac{1+m}{2})N-1}{p-1-k}e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+} \\ \qquad \text{if } m' = m-\frac{2}{N} \\ 1 - r_p\left(m, m+\frac{2}{N}\right) - r_p\left(m,m-\frac{2}{N}\right) \\ \qquad \text{if } m' = m \\ 0 \qquad \text{otherwise.} \end{cases} \tag{7.2.1}$$

Proof. Notice that, for a given m, the number of positive spins is given by $\frac{1+m}{2}N$ and the number of negative spins by $\frac{1-m}{2}N$.

Going right. Let us calculate the probability of going from m to $m+\frac{2}{N}$. To do so, the chosen spin, denoted i, must be of value -1, and this will happen with probability $\frac{1-m}{2}$. Then, depending on the cluster $\mathcal{C}$ to which spin i belongs, switching the spin from -1 to $+1$ happens with probability

$$\mathbb{P}(\sigma_i^p(n+1)=1|\sigma_i^p(n)=-1,\mathcal{C}) = \exp\left(-\beta\left(-\frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma'_j\sigma'_l + \frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma_j^p(n)\sigma_l^p(n)\right)_+\right),$$

where σ' denotes the configuration such that for all $j\neq i$, $\sigma'_j = \sigma_j^p(n)$, and $\sigma'_i = -\sigma_i^p(n)$. We have

$$\begin{aligned} -\frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma'_j\sigma'_l + \frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma_j^p(n)\sigma_l^p(n) \\ = -\frac{1}{2p}\Bigg(\sum_{j,l\in\mathcal{C},j\neq i,l\neq i}\sigma'_j\sigma'_l - \sum_{j,l\in\mathcal{C},j\neq i,l\neq i}\sigma_j^p(n)\sigma_l^p(n) + 2\sum_{j\in\mathcal{C},j\neq i}\sigma'_j\sigma'_i \\ -2\sum_{j\in\mathcal{C},j\neq i}\sigma_j^p(n)\sigma_i^p(n) + (\sigma'_i)^2 - (\sigma_i^p(n))^2\Bigg) \\ = -\frac{1}{2p}\left(-2\sigma_i^p(n)\sum_{j\in\mathcal{C},j\neq i}\sigma_j^p(n) - 2\sigma_i^p(n)\sum_{j\in\mathcal{C},j\neq i}\sigma_j^p(n)\right) \\ = \frac{2}{p}\sigma_i^p(n)\sum_{j\in\mathcal{C},j\neq i}\sigma_j^p(n). \end{aligned}$$

We classify the possible clusters containing i based on the number of negative spins. The number of clusters containing i and k other negative spins is $\binom{\frac{1-m}{2}N-1}{k}\binom{\frac{1+m}{2}N}{p-1-k}$ (choosing k spins among the $\frac{1-m}{2}N-1$ negative spins that are not i, then the $p-1-k$ spins that remain to construct cluster $\mathcal{C}$ among the positive spins). For k negative spins in cluster $\mathcal{C}$ (without counting i), we have

$$\sum_{j\in\mathcal{C},j\neq i}\sigma_j^p(n)=\sum_{j\in\mathcal{C},j\neq i,\sigma_j^p(n)=1}1-\sum_{j\in\mathcal{C},j\neq i,\sigma_j^p(n)=-1}1=p-1-k-k,$$

and thus, since $\sigma_i^p(n)=-1$

$$-\frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma_j'\sigma_l'+\frac{1}{2p}\sum_{j,l\in\mathcal{C}}\sigma_j^p(n)\sigma_l^p(n)=2\frac{2k+1-p}{p}$$

The total number of possible choices for $\mathcal{C}$ is $\binom{N-1}{p-1}$ (choosing the $(p-1)$ spins that are not i). Hence

$$r_p\left(m,m+\frac{2}{N}\right)=\frac{1-m}{2}\frac{1}{\binom{N-1}{p-1}}\sum_{k=0}^{p-1}\binom{\left(\frac{1-m}{2}\right)N-1}{k}\binom{\left(\frac{1+m}{2}\right)N}{p-1-k}e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}$$

Going left. Similar calculations yield the probability of going left : the probability of choosing a spin of value $+1$ is $\frac{1+m}{2}$, then we classify the possible clusters containing this spin based on the number of negative spins. □

Remark 7.2.1. *The values given in* (7.2.1) *are coherent in the case* $p=N$. *Observe for instance that the only nonzero term in the sum defining* $r_N\left(m,m+\frac{2}{N}\right)$ *is obtained for* $k=\frac{1-m}{2}N-1$. *Thus*

$$r_N\left(m,m+\frac{2}{N}\right)=\frac{1-m}{2}e^{-2\beta\left(-m-\frac{1}{N}\right)_+}=r\left(m,m+\frac{2}{N}\right),$$

where the value of r *is given in* (7.1.4).

Remark 7.2.2. *We observe how the transition probabilities evolve with the parameter* p *in Figure 7.2. Furthermore, the values given in* (7.2.1) *allow us to define, on the state space* $\{-1,1+\frac{2}{N},...,1-\frac{2}{N},1\}$, *a transition matrix for the magnetization. The latter is an irreducible and aperiodic Markov chain on a finite state space, and thus admits a unique invariant measure. We can numerically obtain it by iterating the transition matrix (see Figure 7.3)*

Study of the critical parameter

We now wish to show how adding random batches artificially increases the temperature of the system, thus decreasing the critical temperature (or, equivalently, increasing the critical inverse temperature).

Limit ODE. Let us, like previously, find the limit as N goes to infinity of the dynamics of $(m_{N,p}(n))_n$ with time step $\frac{1}{N}$. This discrete-time Markov chain admits a transition operator $U^{(N,p)}$ given by

$$U^{(N,p)}=\left(U_{i,j}^{(N,p)}\right)_{0\leqslant i,j\leqslant N}\quad\text{where}\quad U_{i,j}^{(N,p)}=r_p\left(-1+\frac{2i}{N},-1+\frac{2j}{N}\right).$$

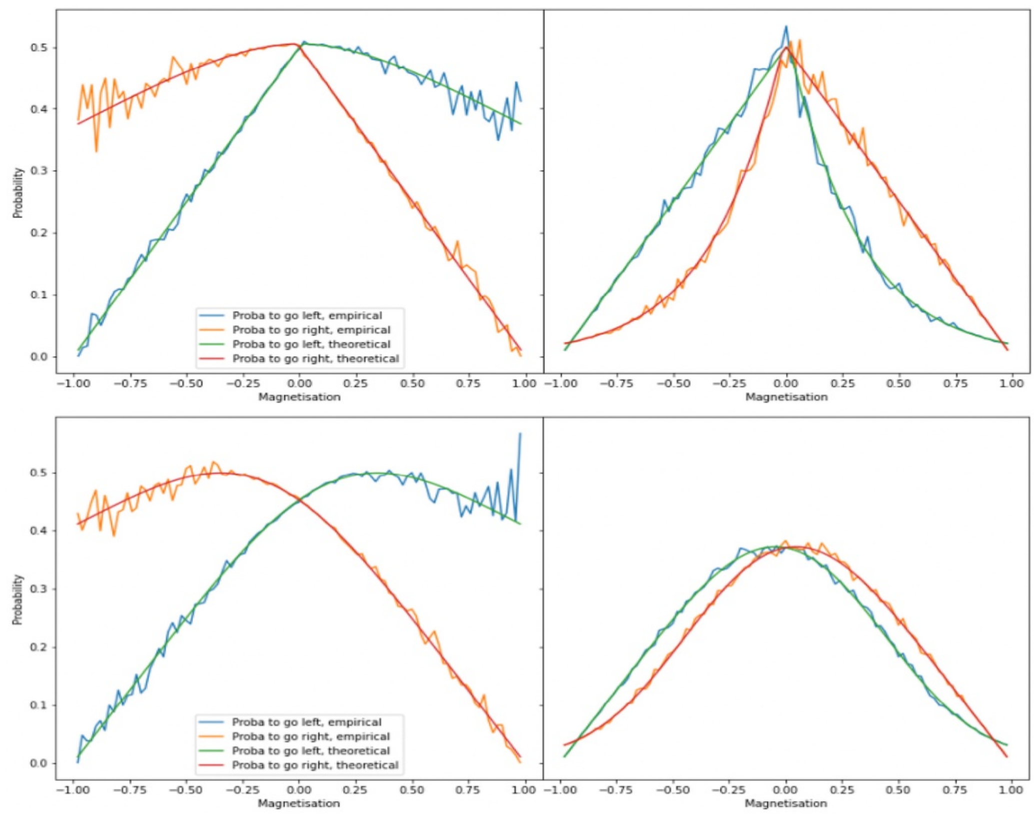

Figure 7.1: Comparison of theoretical and empirical transition probabilities, for $N = 100$. The theoretical values are those given in Lemma 7.2.1. To numerically compute the empirical transition probabilities, for each initial magnetization in $\{-1, -1 + \frac{2}{N}, ..., 1 - \frac{2}{N}, 1\}$, 10 processes are simulated during 1000 timesteps, and we consider the proportion of times the processes go left or right. **Top :** without random batches. **Bottom :** with random batches of size $p = 10$. **Left :** for $\beta = 0.5$. **Right :** for $\beta = 2$.

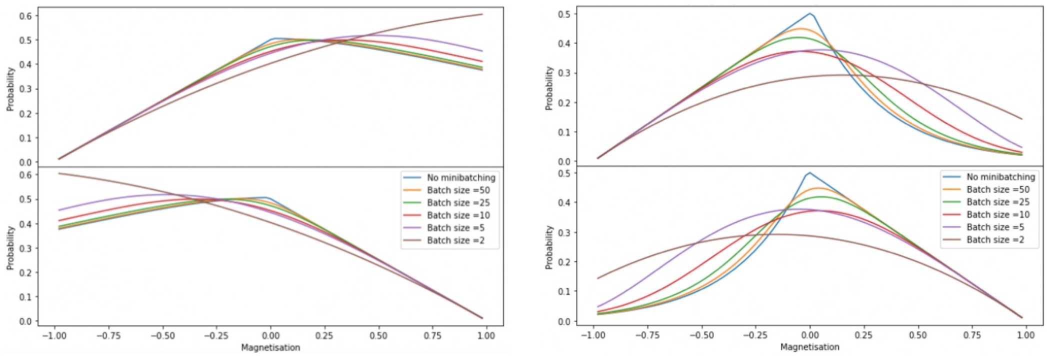

Figure 7.2: Comparison of transition probabilities depending on batch size, for $N = 100$. The values given are from Lemma 7.2.1. **Top :** probability of going left. **Bottom :** probability of going right. **Left :** for $\beta = 0.5$. **Right :** for $\beta = 2$.

We denote $A_N^{(p)} = N(U^{(N,p)} - I)$ and have, for all continuously differentiable functions f,

$$\begin{aligned} A_N^{(p)} f(m) =& N r_p\left(m, m+\frac{2}{N}\right)\left(f(m+\frac{2}{N}) - f(m)\right) \\ &+ N r_p\left(m, m-\frac{2}{N}\right)\left(f(m-\frac{2}{N}) - f(m)\right) \\ =& r_p\left(m, m+\frac{2}{N}\right)\left(2f'(m) + O\left(\frac{1}{N}\right)\right) - r_p\left(m, m-\frac{2}{N}\right)\left(2f'(m) + O\left(\frac{1}{N}\right)\right). \end{aligned}$$

We have, by standard computations

$$\begin{aligned} r_p\left(m, m+\frac{2}{N}\right) &= \frac{1-m}{2}\binom{N-1}{p-1}^{-1}\sum_{k=0}^{p-1}\binom{\left(\frac{1-m}{2}\right)N-1}{k}\binom{\left(\frac{1+m}{2}\right)N}{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} \\ &\xrightarrow[N\to\infty]{} \frac{1-m}{2}\sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^k\left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}, \end{aligned}$$

and likewise

$$r_p\left(m, m-\frac{2}{N}\right) \xrightarrow[N\to\infty]{} \frac{1+m}{2}\sum_{k=0}^{p-1}\binom{p-1}{k}\left(\frac{1-m}{2}\right)^k\left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+}.$$

Hence

$$A_N^{(p)} f(m) \xrightarrow[N\to\infty]{} A^{(p)} f(m),$$

where

$$A^{(p)} f(m) = f'(m)\left(S_1^{p,\beta}(m) - S_2^{p,\beta}(m)\right) - m f'(m)\left(S_1^{p,\beta}(m) + S_2^{p,\beta}(m)\right),$$

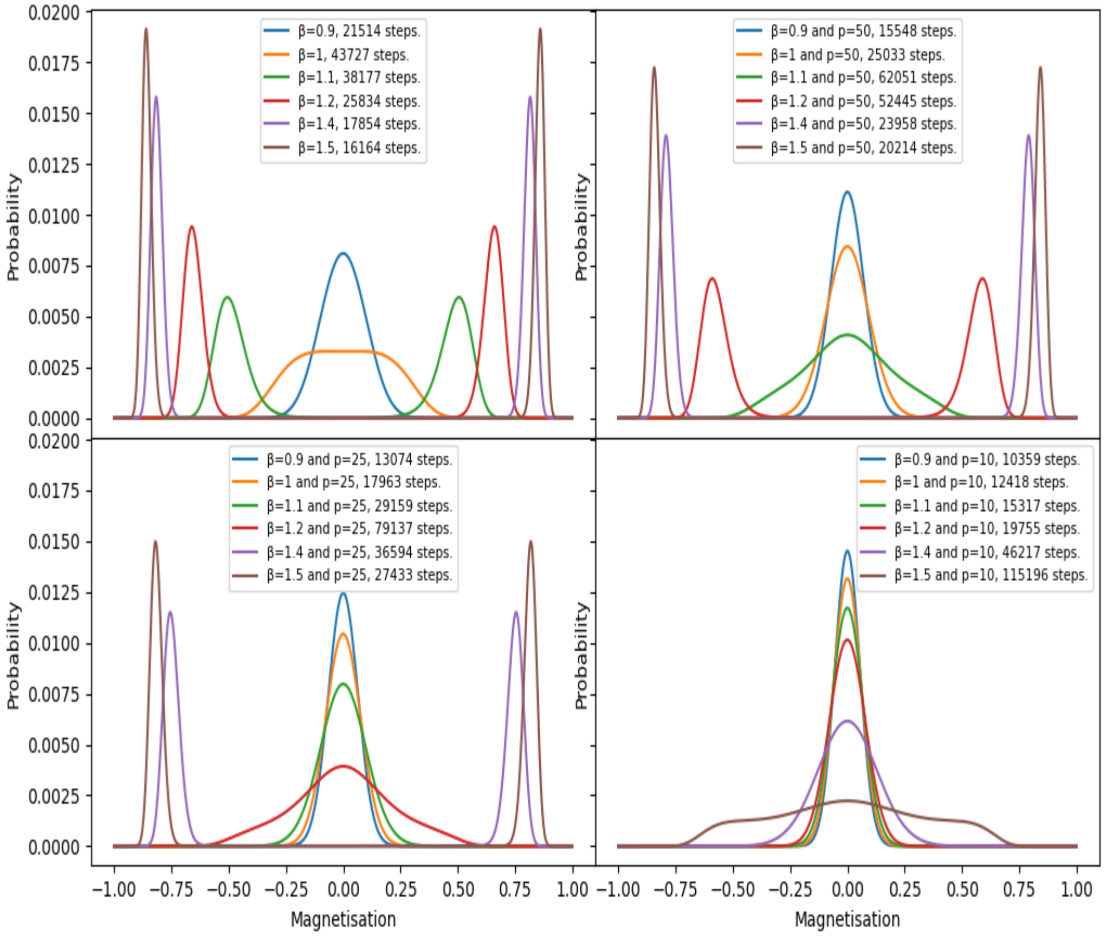

Figure 7.3: Numerical observation of the invariant distribution for the Curie-Weiss model with N spins. Starting from the uniform distribution for the magnetization, we iterate the transition matrix (given in Lemma 7.2.1) until the L^1 distance between two consecutive iterations is less than a threshold $N\epsilon$, with $N = 1000$, $\epsilon = 10^{-9}$ and various values for β. We indicate the number of iterations (or steps) needed before convergence. **Top left :** with no random batches. **Top right :** with $p = 50$. **Bottom left :** with $p = 25$. **Bottom right :** with $p = 10$.

$$S_1^{p,\beta}(m) = \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^k \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}$$
$$S_2^{p,\beta}(m) = \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^k \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+}.$$

Remark 7.2.3. *Notice that*

$$S_1^{p,\beta}(m) = \mathbb{E}\left(e^{-2\beta\left(\frac{2X_{m,p}+1-p}{p}\right)_+}\right), \quad S_2^{p,\beta}(m) = \mathbb{E}\left(e^{-2\beta\left(\frac{p-1-2X_{m,p}}{p}\right)_+}\right),$$

where $X_{m,p}$ is a random variable following a binomial distribution of parameters $p-1$ and $\frac{1-m}{2}$. Intuitively, for an infinite number of spins, the dynamics of the system relies on the construction of a cluster of size p (containing the chosen spin that may change), which is done by independently taking the remaining $p-1$ spins from an infinite pool containing a proportion of $\frac{1-m}{2}$ negative spins.

Denoting $f_p(\beta, m) = \left(S_1^{p,\beta}(m) - S_2^{p,\beta}(m)\right) - m\left(S_1^{p,\beta}(m) + S_2^{p,\beta}(m)\right)$, by Theorem 17.28 of [107], the process $M_t^{(N,p)} = m_{N,p}(\lfloor Nt \rfloor)$ weakly converges to the solution of

$$\frac{d}{dt}m(t) = f_p(\beta, m(t)).$$

The cases $p = 2$ and $p = 3$. We may directly compute

$$S_1^{2,\beta}(m) = \frac{1+m}{2} + \frac{1-m}{2}e^{-\beta}, \qquad S_2^{2,\beta}(m) = \frac{1+m}{2}e^{-\beta} + \frac{1-m}{2},$$
$$S_1^{3,\beta}(m) = \left(\frac{1+m}{2}\right)^2 + 2\left(\frac{1+m}{2}\right)\left(\frac{1-m}{2}\right) + \left(\frac{1-m}{2}\right)^2 e^{-\frac{4\beta}{3}},$$
$$S_2^{3,\beta}(m) = \left(\frac{1+m}{2}\right)^2 e^{-\frac{4\beta}{3}} + 2\left(\frac{1+m}{2}\right)\left(\frac{1-m}{2}\right) + \left(\frac{1-m}{2}\right)^2,$$

which yield

$$\begin{aligned} f_2(\beta, m) =& m(1-e^{-\beta}) - m(1+e^{-\beta}) = -2me^{-\beta}, \\ f_3(\beta, m) =& \left(\left(\frac{1+m}{2}\right)^2 - \left(\frac{1-m}{2}\right)^2\right)\left(1 - e^{-\frac{4\beta}{3}}\right) \\ & - m\left(\left(\left(\frac{1+m}{2}\right)^2 + \left(\frac{1-m}{2}\right)^2\right)\left(1 + e^{-\frac{4\beta}{3}}\right) + (1+m)(1-m)\right) \\ =& m\left(1 - e^{-\frac{4\beta}{3}}\right) - m\left(\frac{(1+m^2)}{2}\left(1 + e^{-\frac{4\beta}{3}}\right) + 1 - m^2\right) \\ =& -\frac{m}{2}\left(1 + 3e^{-\frac{4\beta}{3}}\right) + \frac{m^3}{2}\left(1 - e^{-\frac{4\beta}{3}}\right). \end{aligned}$$

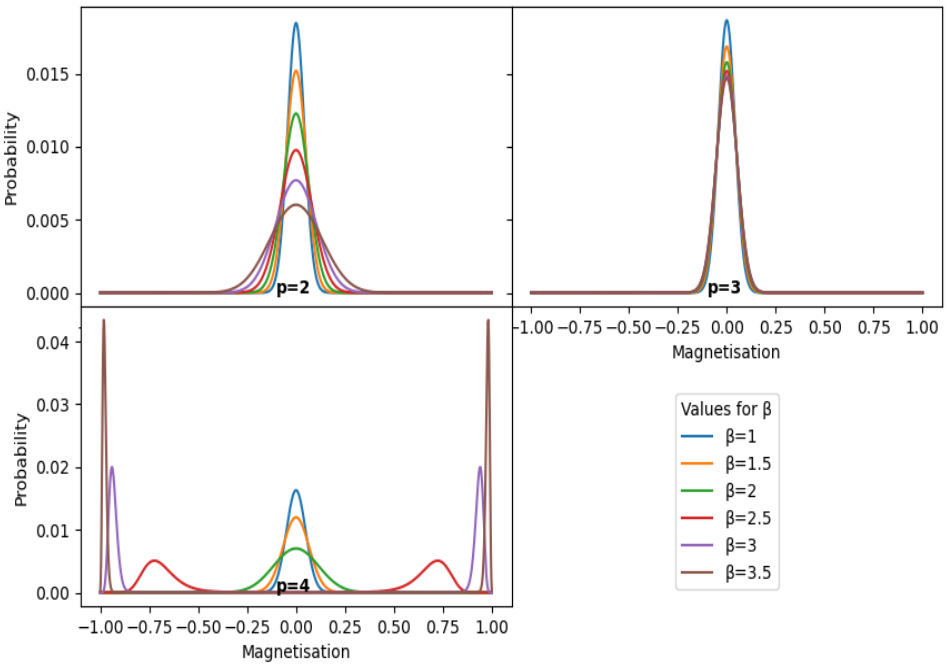

Figure 7.4: Numerical observation of the invariant distribution for the Curie-Weiss model with N spins in the cases $p = 2$ (Top right), $p = 3$ (Top left) and $p = 4$ (Bottom left). Starting from the uniform distribution for the magnetization, we iterate the transition matrix (given in Lemma 7.2.1) until the L^1 distance between two consecutive iterations is less than a threshold $N\epsilon$, with $N = 1000$, $\epsilon = 10^{-9}$ and various values for β.

For $p = 2$ we thus have, $f_2(\beta, m) = 0 \iff m = 0$, and furthermore notice that $\partial_m f_2(\beta, 0) < 0$, which means that 0 is the unique equilibrium state, and it is stable. For $p = 3$,

$$f_3(\beta, m) = 0 \quad \iff \quad m = 0 \quad \text{or} \quad m = \pm\sqrt{\frac{1 + 3e^{-\frac{4\beta}{3}}}{1 - e^{-\frac{4\beta}{3}}}}.$$

However, for all $\beta > 0$ we have $\sqrt{\frac{1+3e^{-\frac{4\beta}{3}}}{1-e^{-\frac{4\beta}{3}}}} > 1$, and furthermore $\partial_m f_3(\beta, 0) = -\frac{1+3e^{-\frac{4\beta}{3}}}{2} < 0$. The point 0 is thus the unique equilibrium state, and it is stable.

We may observe this phenomenon in Figure 7.4, in which we compare the case $p = 2$ and $p = 3$ with $p = 4$.

Existence of a phase transition for $p \geqslant 4$. First notice that

$$\begin{aligned} f_p(\beta, 0) =& \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1}{2}\right)^{p-1} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} - \sum_{k=0}^{p-1} \binom{p-1}{k} \left(\frac{1}{2}\right)^{p-1} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+} \\ =& 0 \quad \text{by change of variables} \quad k' = p - 1 - k. \end{aligned}$$

Thus $m = 0$ is for all $\beta > 0$ an equilibrium state. The remaining questions, in order to prove Theorem 7.1.1, are

- is there $\beta_{c,p} > 0$ such that for all $\beta < \beta_{c,p}$ we have $\partial_m f(\beta, 0) < 0$ (in which case $m = 0$ is stable) and such that for all $\beta > \beta_{c,p}$ we have $\partial_m f(\beta, 0) > 0$ (in which case $m = 0$ is unstable) ?
- do we have $\beta_{c,p} > 1$ (in which case the critical temperature has indeed decreased when compared to the case without random batches) ?
- can we give an estimate of $\beta_{c,p}$?

To answer the first question, we may calculate

$$S_1'^{p,\beta}(m) = -\sum_{k=0}^{p-1} \frac{k}{2} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^{k-1} \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} + \sum_{k=0}^{p-1} \frac{p-1-k}{2} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^{k} \left(\frac{1+m}{2}\right)^{p-2-k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}$$

and

$$S_2'^{p,\beta}(m) = -\sum_{k=0}^{p-1} \frac{k}{2} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^{k-1} \left(\frac{1+m}{2}\right)^{p-1-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+} + \sum_{k=0}^{p-1} \frac{p-1-k}{2} \binom{p-1}{k} \left(\frac{1-m}{2}\right)^{k} \left(\frac{1+m}{2}\right)^{p-2-k} e^{-2\beta\left(\frac{p-1-2k}{p}\right)_+},$$

which yields

$$\partial_m f_p(\beta, m) = \left(S_1'^{p,\beta}(m) - S_2'^{p,\beta}(m)\right) - \left(S_1^{p,\beta}(m) + S_2^{p,\beta}(m)\right) - m\left(S_1'^{p,\beta}(m) + S_2'^{p,\beta}(m)\right).$$

We thus have

$$\partial_m f_p(\beta, 0) = 2S_1'^{p,\beta}(0) - 2S_1^{p,\beta}(0) = 2\left(\frac{1}{2}\right)^{p-1} \sum_{k=0}^{p-1} (p-2-2k) \binom{p-1}{k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+}.$$

First, notice

$$\partial_\beta(\partial_m f_p(\beta, 0)) = -4\left(\frac{1}{2}\right)^{p-1} \sum_{k=0}^{p-1} (p-2-2k) \left(\frac{2k+1-p}{p}\right)_+ \binom{p-1}{k} e^{-2\beta\left(\frac{2k+1-p}{p}\right)_+} > 0.$$

The function $\beta \mapsto \partial_m f_p(\beta, 0)$ is therefore an increasing function, which furthermore satisfies $\partial_m f_p(0, 0) < 0$ and $\lim_{\beta\to\infty} \partial_m f_p(\beta, 0) > 0$, hence a unique critical parameter $\beta_{c,p} > 0$.

Remark 7.2.4. *We use the assumption $p \geqslant 4$ in order to prove $\lim_{\beta\to\infty} \partial_m f_p(\beta, 0) > 0$. Indeed*

$$\lim_{\beta\to\infty} \partial_m f_p(\beta, 0) = 2\left(\frac{1}{2}\right)^{p-1} \sum_{k=0}^{p-1} (p-2-2k) \binom{p-1}{k} \mathbb{1}_{k \leqslant \frac{p-1}{2}}.$$

If p is even, all the terms in the sum are nonnegative, and if $p \geqslant 4$, at least one term is positive. If p is odd, one term is negative, and if $p \geqslant 5$ it can easily be shown that it is compensated by the positive terms.

Estimation of the critical parameter. Denoting X_p a random variable following a binomial distribution of parameters $p-1$ and $\frac{1}{2}$, we have

$$\partial_m f_p(\beta, 0) = 2\mathbb{E}\left((p-2-2X_p) e^{-2\beta\left(\frac{2X_p+1-p}{p}\right)_+} \right) := g_p(\beta).$$

We are thus looking for the unique $\beta_{c,p} > 0$ such that $g_p(\beta_{c,p}) = 0$.

Let $Y_p = 2\frac{X_p}{p} - \frac{p-1}{p}$. We have

$$g_p(\beta) = \mathbb{E}\left(2(-pY_p - 1)e^{-2\beta(Y_p)_+}\right). \tag{7.2.2}$$

Since X_p and $p-1-X_p$ have the same law, Y_p has the same law as $2\frac{p-1-X_p}{p} - \frac{p-1}{p} = \frac{p-1}{p} - 2\frac{X_p}{p} = -Y_p$. Thus

$$\begin{aligned}
g_p(\beta) =& -\mathbb{E}\left(pY_p e^{-2\beta(Y_p)_+}\right) - \mathbb{E}\left(p(-Y_p)e^{-2\beta(-Y_p)_+}\right) - \mathbb{E}\left(e^{-2\beta(Y_p)_+}\right) - \mathbb{E}\left(e^{-2\beta(-Y_p)_+}\right) \\
=& -\mathbb{E}\left(pY_p\left(e^{-2\beta(Y_p)_+} - e^{-2\beta(-Y_p)_+}\right)\right) - \mathbb{E}\left(e^{-2\beta(Y_p)_+} + e^{-2\beta(-Y_p)_+}\right) \\
=& \mathbb{E}\left(2pY_p e^{-\beta|Y_p|}\sinh(\beta Y_p)\right) - \mathbb{E}\left(2e^{-\beta|Y_p|}\cosh(\beta Y_p)\right) \\
=& 2\mathbb{E}\left(\cosh(\beta Y_p)e^{-\beta|Y_p|}\left(pY_p\tanh(\beta Y_p) - 1\right)\right).
\end{aligned}$$

As this is an increasing function in β, in order to prove that $\beta_{c,p} > 1$, it is sufficient to prove that $g_p(1) < 0$. The Law of Large Number and the Central Limit Theorem yield

$$Y_p \xrightarrow[p\to\infty]{a.s} 0 \quad \text{and} \quad \frac{p}{\sqrt{p-1}}Y_p \xrightarrow[p\to\infty]{law} \mathcal{N}(0,1).$$

We have

$$\begin{aligned}
g_p(\beta) =& 2\mathbb{E}\left(\left(1 + \frac{\beta^2 Y_p^2}{2} + o(Y_p^2)\right)\left(1 - \beta|Y_p| + \frac{\beta^2 Y_p^2}{2} + o(Y_p^2)\right)\right. \\
& \left.\times\left(pY_p\left(\beta Y_p - \frac{\beta^3 Y_p^3}{3} + o(Y_p^3)\right) - 1\right)\right) \\
=& 2\mathbb{E}\left(\left(\frac{p-1}{p}\beta\frac{p^2}{p-1}Y_p^2 - 1\right) - \beta|Y_p|\left(\frac{p-1}{p}\beta\frac{p^2}{p-1}Y_p^2 - 1\right) + O(Y_p^2) + O(pY_p^4)\right) \\
=& 2\left(\beta\frac{p-1}{p}\mathbb{E}\left(\frac{p^2}{p-1}Y_p^2\right) - 1 - \frac{(p-1)^{3/2}}{p^2}\left(\beta^2\mathbb{E}\left(\left|\frac{p}{\sqrt{p-1}}Y_p\right|^3\right) - \beta\mathbb{E}\left(\left|\frac{p}{\sqrt{p-1}}Y_p\right|\right)\right)\right. \\
& \left.+\frac{1}{p}\mathbb{E}\left(O(pY_p^2) + O(p^2Y_p^4)\right)\right) \\
=& 2\left(\beta\frac{p-1}{p}\mathbb{E}\left(Z^2\right) - 1 - \frac{(p-1)^{3/2}}{p^2}\left(\beta^2\left(\mathbb{E}\left|Z\right|^3 + o(1)\right) - \beta\left(\mathbb{E}\left(|Z|\right) + o(1)\right)\right) + O\left(\frac{1}{p}\right)\right)
\end{aligned}$$

$$=2(\beta-1)-\frac{2}{\sqrt{p}}\sqrt{\frac{2}{\pi}}\left(2\beta^2-\beta\right)+o\left(\frac{1}{\sqrt{p}}\right),$$

where for this last equality, we use Lemma F.1.1 and the fact that, for $Z\sim\mathcal{N}(0,1)$, $\mathbb{E}|Z|=\sqrt{\frac{2}{\pi}}$ and $\mathbb{E}(|Z|^3)=2\sqrt{\frac{2}{\pi}}$.

In the end, we obtain, again, the fact that $g_p(1)\xrightarrow[p\to\infty]{}0$ (hence the right critical parameter at the limit) and the fact that, at least for p sufficiently large, $g_p(1)<0$. For smaller values of p, we rely on numerical simulations to verify $g_p(1)<0$ (See Figure 7.5). Let us find an approximation of $\beta_{c,p}$ by using the fact that $g_p(\beta_{c,p})=0$. We have

$$2\sqrt{\frac{2}{p\pi}}\beta_{c,p}^2-\left(1+\sqrt{\frac{2}{p\pi}}\right)\beta_{c,p}+1+o\left(\frac{1}{\sqrt{p}}\right)=0,$$

i.e

$$\begin{aligned}\beta_{c,p,\pm}=&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(\left(1+\sqrt{\frac{2}{p\pi}}\right)\pm\left(\left(1+\sqrt{\frac{2}{p\pi}}\right)^2-8\sqrt{\frac{2}{p\pi}}\left(1+o\left(\frac{1}{\sqrt{p}}\right)\right)\right)^{1/2}\right)\\=&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(1+\sqrt{\frac{2}{p\pi}}\pm\left(1+\frac{2}{p\pi}-6\sqrt{\frac{2}{p\pi}}+o\left(\frac{1}{p}\right)\right)^{1/2}\right)\\=&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(1+\sqrt{\frac{2}{p\pi}}\pm\left(1+\frac{1}{p\pi}-3\sqrt{\frac{2}{p\pi}}-\frac{9}{p\pi}+o\left(\frac{1}{p}\right)\right)\right),\end{aligned}$$

thus

$$\begin{aligned}\beta_{c,p}=&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(1+\sqrt{\frac{2}{p\pi}}-\left(1-3\sqrt{\frac{2}{p\pi}}-\frac{8}{p\pi}+o\left(\frac{1}{p}\right)\right)\right)\\=&\frac{1}{4}\sqrt{\frac{p\pi}{2}}\left(4\sqrt{\frac{2}{p\pi}}+\frac{8}{p\pi}+o\left(\frac{1}{p}\right)\right)\\=&1+\sqrt{\frac{2}{p\pi}}+o\left(\frac{1}{\sqrt{p}}\right).\end{aligned}$$

We have thus proved Theorem 7.1.1.

7.3 Random Batch Method for interacting particle systems and stationary distribution(s)

We now turn our attention to the study of (1.3.9) for a given batch size $p\in\mathbb{N}\setminus\{0,1\}$. In the specific case of U and W given in (7.1.9), we study the phase transition for (Eff) and prove Theorem 7.1.3. The stationary distributions of (NL), provided there exists one, are defined by the solutions of

$$\mu_\sigma(dx)=\frac{\exp\left(-\frac{1}{\sigma}\left(U(x)+W*\mu_\sigma(x)\right)\right)}{\int\exp\left(-\frac{1}{\sigma}\left(U(y)+W*\mu_\sigma(y)\right)\right)dy}dx. \tag{7.3.1}$$

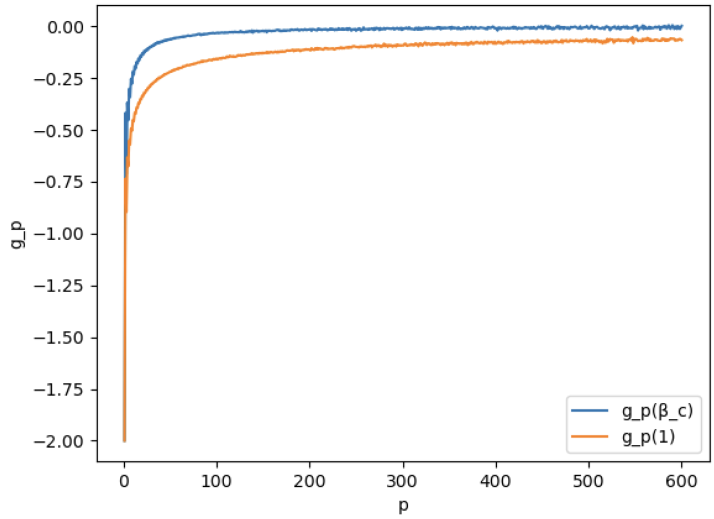

Figure 7.5: Numerical values for $g_p(1)$ and $g_p\left(\tilde{\beta}_c\right)$, with $\tilde{\beta}_c = 1 + \sqrt{\frac{2}{p\pi}}$. The value of g_p given in (7.2.2) is computed via Monte-Carlo approximation using $M = 10^8$ samples for Y_p.

We consider the case of linear interactions in a double well potential in dimension one, i.e U and W given in (7.1.9), which in particular implies

$$\Sigma(x,\rho) = (\nabla W)^2 * \rho(x) - (\nabla W * \rho(x))^2 = L_W^2 \left(\int y^2 \rho(dy) - \left(\int y \rho(dy) \right)^2 \right) = L_W^2 \mathrm{Var}(\rho).$$

Denote, for a measure μ,

$$\kappa_1(\mu) = \int_{\mathbb{R}} x \mu(dx) \quad \text{and} \quad \kappa_2(\mu) = \int_{\mathbb{R}} (x - \kappa_1(\mu))^2 \mu(dx).$$

The stationary distributions of (Eff), provided there exist one, are thus similarly defined by the solutions of

$$\mu_\sigma^{\delta,p}(dx) = \frac{\exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2\left(\mu_\sigma^{\delta,p}\right)} \left(U(x) + \frac{L_W}{2} \left| x - \kappa_1\left(\mu_\sigma^{\delta,p}\right) \right|^2 \right) \right)}{\int \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2\left(\mu_\sigma^{\delta,p}\right)} \left(U(y) + \frac{L_W}{2} \left| y - \kappa_1\left(\mu_\sigma^{\delta,p}\right) \right|^2 \right) \right) dy} dx. \tag{7.3.2}$$

The pair $(\kappa_1(\mu_\sigma^{\delta,p}), \kappa_2(\mu_\sigma^{\delta,p}))$ is therefore a solution of

$$\kappa_1 = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2} \left(U(x) + \frac{L_W}{2} |x - \kappa_1|^2 \right) \right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma + \frac{\delta L_W^2}{p-1}\kappa_2} \left(U(x) + \frac{L_W}{2} |x - \kappa_1|^2 \right) \right) dx} \tag{7.3.3}$$

$$\kappa_2 = \frac{\int_{\mathbb{R}} (x-\kappa_1)^2 \exp\left(-\frac{2}{2\sigma+\frac{\delta L_W^2}{p-1}\kappa_2}\left(U(x)+\frac{L_W}{2}|x-\kappa_1|^2\right)\right)dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma+\frac{\delta L_W^2}{p-1}\kappa_2}\left(U(x)+\frac{L_W}{2}|x-\kappa_1|^2\right)\right)dx}. \tag{7.3.4}$$

Thanks to (7.3.2), solving for (κ_1,κ_2) the system of equations (7.3.3)-(7.3.4) is equivalent to finding a stationary distribution of (Eff). Define

$$g(x,\sigma,\kappa) = \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa|^2\right)\right), \tag{7.3.5}$$

$$f_1(\sigma,\kappa) = \frac{\int_{\mathbb{R}} x g(x,\sigma,\kappa)dx}{\int_{\mathbb{R}} g(x,\sigma,\kappa)dx}, \tag{7.3.6}$$

$$f_2(\sigma,\kappa) = \frac{\int_{\mathbb{R}} (x-\kappa)^2 g(x,\sigma,\kappa)dx}{\int_{\mathbb{R}} g(x,\sigma,\kappa)dx}, \tag{7.3.7}$$

such that, for the symbol $* \in \{0,\pm\}$ and $\mu_{\sigma,*}$ defined in Theorem 7.1.2, $\kappa_1(\mu_{\sigma,*})$ is a solution of $\kappa_1(\mu_{\sigma,*}) = f_1(\sigma,\kappa_1(\mu_{\sigma,*}))$ and $\kappa_2(\mu_{\sigma,*}) = f_2(\sigma,\kappa_1(\mu_{\sigma,*}))$ is the corresponding variance.

Our goal is to compare the stationary distribution(s) for (Eff) to the stationary distribution(s) for (NL), in particular in regards to this critical parameter σ_c. To do so, we begin by showing that any stationary distribution of (NL) is a stationary distribution of (Eff), and conversely.

Lemma 7.3.1. *Let μ be a probability measure on $\mathbb{R}$.*

- *If μ is a solution of (7.3.2) for a diffusion coefficient σ', then μ is a solution of (7.3.1) for a diffusion coefficient $\sigma = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$,*
- *If μ is a solution of (7.3.1) for a diffusion coefficient σ and $\frac{\delta}{p-1} < \frac{2\sigma}{L_W^2\kappa_2(\mu)}$, then μ is a solution of (7.3.2) for a diffusion coefficient $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$.*

Proof. Let us prove the two points.

A stationary distribution of (Eff) **is a stationary distribution of** (NL)**.** Assume μ satisfies (7.3.2) for a given (δ,p,σ'), which in particular is equivalent to the pair $(\kappa_1(\mu),\kappa_2(\mu))$ satisfying

$$\kappa_1(\mu) = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right)dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right)dx}$$

$$\kappa_2(\mu) = \frac{\int_{\mathbb{R}} (x-\kappa_1(\mu))^2 \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right)dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right)dx}.$$

Denoting $\sigma = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$, we thus have

$$\kappa_1(\mu) = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right)dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right)dx}$$

$$\kappa_2(\mu) = \frac{\int_{\mathbb{R}} (x-\kappa_1(\mu))^2 \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}.$$

Thus, if μ is a stationary distribution for (Eff) with diffusion coefficient σ' and parameters δ and p, it is also a stationary distribution for (NL) with diffusion coefficient $\sigma = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$.

A stationary distribution of (NL) **is a stationary distribution of** (Eff)**.** Assume μ satisfies (7.3.2) for a given σ, which in particular is equivalent to the pair $(\kappa_1(\mu), \kappa_2(\mu))$ satisfying

$$\kappa_1(\mu) = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}$$

$$\kappa_2(\mu) = \frac{\int_{\mathbb{R}} (x-\kappa_1(\mu))^2 \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}.$$

Consider parameters δ and p, and denote $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$. Notice $\kappa_2(\mu)$ is independent of δ and p thus, provided $\frac{\delta}{(p-1)}$ is small enough, we may ensure $\sigma' > 0$ and have

$$\kappa_1(\mu) = \frac{\int_{\mathbb{R}} x \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}$$

$$\kappa_2(\mu) = \frac{\int_{\mathbb{R}} (x-\kappa_1(\mu))^2 \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}{\int_{\mathbb{R}} \exp\left(-\frac{2}{2\sigma'+\frac{\delta L_W^2}{p-1}\kappa_2(\mu)}\left(U(x)+\frac{L_W}{2}|x-\kappa_1(\mu)|^2\right)\right) dx}.$$

Thus, given two parameters δ and p, if μ is a stationary distribution for (NL) with diffusion coefficient σ, it is also a stationary distribution for (Eff) with diffusion coefficient $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$. □

Unfortunately, this lemma does not directly imply the existence of a phase transition for (Eff). Several issues arise :

- the existence of a symmetric stationary distribution $\mu_{\sigma,0}$ for (NL) with diffusion coefficient $\sigma > 0$ only yields the existence of a symmetric stationary distribution for (Eff) for a specific diffusion coefficient $\sigma' = \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu)$. We need to show that, for any $\sigma > 0$, there exists a symmetric stationary distribution for (Eff).
- likewise, the existence of non symmetric stationary distribution for (Eff) is only ensured for specific diffusion coefficients.
- we cannot infer the uniqueness of the symmetric stationary distribution for (Eff) from the uniqueness of the symmetric stationary distribution for (NL). Given $\sigma' > 0$, there may *a priori* be two stationary distributions for (Eff), denoted μ_1 and μ_2, with different variance $\kappa_2(\mu_1) \neq \kappa_2(\mu_2)$, which thus correspond to two different symmetric stationary distributions for (NL) with diffusion coefficients $\sigma_1 = \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_1) \neq \sigma' + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_2) = \sigma_2$.

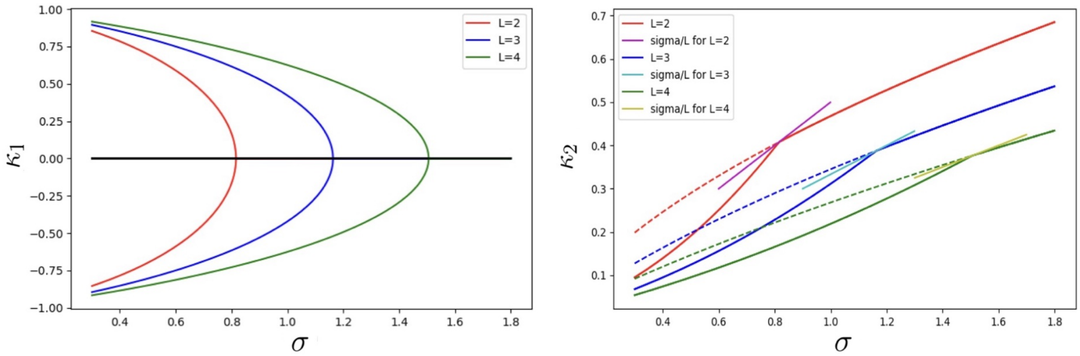

Figure 7.6: **Left** : The means of $\mu_{\sigma,\pm}$ as a function of σ for different values of L_W, as given in Theorem 7.1.2. **Right** : The variances of $\mu_{\sigma,\pm}$ as a function of σ for different values of L_W. In dotted line the variance of $\mu_{\sigma,0}$, and in solid line the variance of $\mu_{\sigma,\pm}$.

We therefore dedicate the remainder of this document to the proof of Theorem 7.1.3.

7.3.1 Some results on the stationary distribution(s) of (NL)

The study of the critical parameter of (Eff) relies, by Lemma 7.3.1, on the study of the one of (NL). We gather here some results concerning the latter. They are numerically illustrated in Figure 7.6, and the proofs are postponed to Appendix F.2.

Lemma 7.3.2. *We have the following results concerning the stationary distribution(s) of* (NL).

- ***Symmetry.*** *We have $\kappa_1(\mu_{\sigma,+}) = -\kappa_1(\mu_{\sigma,-})$ and $\kappa_2(\mu_{\sigma,+}) = \kappa_2(\mu_{\sigma,-})$.*

- ***Moment bound.*** *Let the symbol $* \in \{0, \pm\}$. Consider $\mu_{\sigma,*}$ given in Theorem 7.1.2, and $\kappa_1(\mu_{\sigma,*})$ (resp. $\kappa_2(\mu_{\sigma,*})$) the corresponding mean (resp. variance). There exists $C_{\kappa_1}, C_{\kappa_2} > 0$ such that for $\sigma \in [0, \sigma_c]$ we have*

$$|\kappa_1(\mu_{\sigma,*})| \leqslant C_{\kappa_1}, \qquad \text{and} \qquad |\kappa_2(\mu_{\sigma,*})| \leqslant C_{\kappa_2}. \tag{7.3.8}$$

- ***Critical variance.*** *We have the equality*

$$\kappa_2\left(\mu_{\sigma_c}\right) = \frac{\sigma_c}{L_W}. \tag{7.3.9}$$

  Furthermore, for $\sigma < \sigma_c$ we have $\kappa_2\left(\mu_{\sigma,\pm}\right) < \frac{\sigma}{L_W}$ and $\kappa_2\left(\mu_{\sigma,0}\right) > \frac{\sigma}{L_W}$, and for $\sigma > \sigma_c$ we have $\kappa_2\left(\mu_{\sigma,0}\right) < \frac{\sigma}{L_W}$.

- ***Continuity.*** *The function $\sigma \mapsto \kappa_1(\mu_{\sigma,+})$, with the convention $\mu_{\sigma,+} = \mu_{\sigma,0}$ for $\sigma \geqslant \sigma_c$, is continuous on $]0, \infty[$. In particular, this also yields the continuity of $\sigma \mapsto \kappa_2(\mu_{\sigma,+}) = f_2(\sigma, \kappa_1(\mu_{\sigma,+}))$.*

- ***Lipschitz continuity.*** *Let $\sigma_0 > 0$. The functions $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ and $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$ are Lipschitz continuous, respectively on $[\sigma_0, \infty[$ and on $[\sigma_0, \sigma_c]$. More precisely, there exists $C > 0$ such that for respectively $\sigma > \sigma_0$ and $\sigma \in]\sigma_0, \sigma_c[$ we have $\left|\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})\right| \leqslant C$.*

Remark 7.3.1. *The bound* (7.3.8)*, combined with the knowledge of the fact that for* $\sigma \geqslant \sigma_c$ *there only exists a symmetric stationary distribution for* (NL) *as well as Lemma 7.3.1, shows that we can restrict our study of the stationary distribution for both* (NL) *and* (Eff) *to a compact set of means* $m \in [-C_{\kappa_1}, C_{\kappa_1}]$.

Remark 7.3.2. *The main technical difficulty lies in the proof of the Lipschitz continuity of* $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ *and* $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$*, since it turns out that the mean* $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$ *is not Lipschitz continuous near the critical parameter* σ_c *(See Figure 7.6). It therefore requires a careful estimation of the mean and variance around* σ_c*, and the proof is a section of its own (See Appendix F.2.2).*

7.3.2 Phase transition for the effective dynamics

Let $\sigma_0 > 0$ and define, for $* \in \{0, \pm\}$, the function $g_{eff,*} : \sigma \mapsto \sigma - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma,*})$.

From Lemma 7.3.1, if $\mu_{\sigma,*}$ is a stationary distribution for (NL), then it is a stationary distribution for (Eff) with diffusion coefficient $\sigma' = g_{eff,*}(\sigma)$.

By Lemma 7.3.2, $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ is a Lipschitz continuous function on $[\sigma_0, \infty[$ and $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$ is a Lipschitz continuous function on $[\sigma_0, \sigma_c]$, and, more precisely, in both cases we obtain that $\left|\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})\right|$ is bounded by some constant $C > 0$.

In this case, the function $\sigma \mapsto g_{eff,*}(\sigma)$ is such that $g'_{eff,*}(\sigma) = 1 - \frac{\delta L_W^2}{2(p-1)}\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})$ and thus $g'_{eff,*}(\sigma) \in \left[1 - \frac{\delta L_W^2}{2(p-1)}C, 1 + \frac{\delta L_W^2}{2(p-1)}C\right]$. In particular, for $\frac{\delta}{p-1}$ sufficiently small, $g_{eff,*}(\sigma)$ is both an increasing continuous function and positive.

Thus, $g_{eff,0}$ and $g_{eff,\pm}$ are two injective functions. In particular, $g_{eff,\pm}$ is a bijection from $[\sigma_0, \sigma_c]$ to $[g_{eff,\pm}(\sigma_0), g_{eff,\pm}(\sigma_c)]$.

Finally, notice that $g_{eff,\pm}(\sigma_c) = g_{eff,0}(\sigma_c) = \sigma_c\left(1 - \frac{\delta L_W}{2(p-1)}\right)$, that $g_{eff,*}(\sigma_0) \leqslant \sigma_0$, and that, up to the additional assumption $\frac{2(\sigma_c - \sigma_0)}{\sigma_c L_W} > \frac{\delta}{p-1}$, we may assume $g_{eff,*}(\sigma_c) > \sigma_0$.

We may now state the following facts concerning the stationary distribution(s) for (Eff).

- **There exists at least one symmetric stationary distribution.** Since $g'_{eff,0} \geqslant 1 - \frac{\delta L_W^2}{2(p-1)}C$, $g_{eff,0}$ is an increasing function such that $g_{eff,0}(x) \xrightarrow[x\to\infty]{} \infty$. Thus, $g_{eff,0}$ is a bijection from $[\sigma_0, \infty[$ to $[g_{eff,0}(\sigma_0), \infty[$. Therefore, for all $\sigma \in [g_{eff,0}(\sigma_0), \infty[$, there exists $\tilde{\sigma}$ such that $g_{eff,0}(\tilde{\sigma}) = \sigma$. In other words, $\mu_{\tilde{\sigma},0}$ is also a symmetric stationary distribution for (Eff) with diffusion coefficient σ.

- **There exists at most one symmetric stationary distribution.** Let $\sigma \geqslant \sigma_0$ and $\kappa_1 = 0$, and assume there are two symmetric stationary distributions of (Eff) with diffusion coefficient σ. This yields two coefficients $\sigma', \sigma'' \geqslant \sigma > 0$ such that

$$\sigma' = \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right)$$
$$\sigma'' = \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right),$$

  where $\mu_{\sigma,1}^{\delta,p}$ and $\mu_{\sigma,2}^{\delta,p}$ denote the two stationary distributions. We consider $\mu_{\sigma'}$ $(= \mu_{\sigma,1}^{\delta,p})$ and $\mu_{\sigma''}$ $(= \mu_{\sigma,2}^{\delta,p})$ the corresponding (unique) symmetric stationary distributions of (NL).

Because σ' and σ'' are greater than σ_0, there exists a constant K, possibly depending on σ_0 and L_W, such that by Lemma 7.3.2

$$\begin{aligned} &|\kappa_2(\mu_{\sigma''}) - \kappa_2(\mu_{\sigma'})| = |f_2(\sigma'',0) - f_2(\sigma',0)| \leqslant K|\sigma'' - \sigma'| \\ \text{i.e.} \quad &|\kappa_2(\mu_{\sigma''}) - \kappa_2(\mu_{\sigma'})| \leqslant \frac{K\delta L_W^2}{2(p-1)} \left| \kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right) - \kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right) \right|. \end{aligned} \tag{7.3.10}$$

Because the stationary distributions of (Eff) are uniquely defined by their mean and variance,

$$\kappa_2(\mu_{\sigma''}) = \kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right) \quad \text{and} \quad \kappa_2(\mu_{\sigma'}) = \kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right),$$

and we obtain from (7.3.10) that, for $\frac{\delta}{p-1}$ sufficiently small, $\kappa_2\left(\mu_{\sigma,1}^{\delta,p}\right) = \kappa_2\left(\mu_{\sigma,2}^{\delta,p}\right)$ and thus that $\mu_{\sigma,1}^{\delta,p} = \mu_{\sigma,2}^{\delta,p}$.

- **For $\sigma \in [g_{eff,\pm}(\sigma_0), g_{eff,\pm}(\sigma_c)[$, there exists at least two stationary distributions with nonzero mean.** Because $g_{eff,\pm}$ is a bijection, consider $g_{eff,\pm}^{-1}(\sigma) \in [\sigma_0, \sigma_c[$. There are three stationary distribution for (NL) with diffusion coefficient $g_{eff,\pm}^{-1}(\sigma)$. By Lemma 7.3.1, $\mu_{g_{eff,\pm}^{-1}(\sigma),+}$ and $\mu_{g_{eff,\pm}^{-1}(\sigma),-}$ are also stationary distributions for (Eff) with diffusion coefficient σ, and they have nonzero mean.

- **For $\sigma \in [g_{eff,\pm}(\sigma_0), g_{eff,\pm}(\sigma_c)[$, there exists at most two stationary distributions with nonzero mean.** By symmetry, it is sufficient to prove that there is at most one stationary distribution with positive mean. Assume there are two such solutions $\mu_{\sigma,+,1}^{\delta,p}$ and $\mu_{\sigma,+,2}^{\delta,p}$.

  Let, for $i \in \{1,2\}$, $\sigma_i = \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma,+,i}^{\delta,p})$. Then, $\mu_{\sigma,+,i}^{\delta,p}$ is a stationary distribution with a positive mean for (NL) with diffusion coefficient σ_i, i.e $\mu_{\sigma,+,i}^{\delta,p} = \mu_{\sigma_i,+}$.

  Thus $\sigma = \sigma_i - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma,+,i}^{\delta,p}) = \sigma_i - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma_i,+}) = g_{eff,+}(\sigma_i)$. Since $g_{eff,+}$ is an injective function, we obtain that $\sigma_1 = \sigma_2$. In particular, $\mu_{\sigma,+,1}^{\delta,p}$ and $\mu_{\sigma,+,2}^{\delta,p}$ are two stationary distribution with a positive mean for (NL) with diffusion coefficient $\sigma_1 = \sigma_2$, thus by uniqueness $\mu_{\sigma,+,1}^{\delta,p} = \mu_{\sigma,+,2}^{\delta,p}$.

- **For $\sigma \geqslant g_{eff,+}(\sigma_c)$, there does not exists stationary distribution with nonzero mean.** The result is direct if $\sigma \geqslant \sigma_c$, because if μ is a stationary measure for (Eff) with diffusion coefficient σ, it is is a stationary measure for (NL) with diffusion coefficient $\sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu) \geqslant \sigma > \sigma_c$, hence it cannot have a nonzero mean.

  Assume $\sigma_c > \sigma \geqslant g_{eff,+}(\sigma_c) > \sigma_0$ and that there exists such a solution $\mu_{\sigma,+}^{\delta,p}$. Consider

  $$\sigma' = \sigma + \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma,+}^{\delta,p}), \tag{7.3.11}$$

  such that $\mu_{\sigma,+}^{\delta,p} = \mu_{\sigma',+}$ is a stationary distribution with positive mean for (NL). We thus necessarily have $\sigma' < \sigma_c$. Let

  $$\tilde{\sigma} = g_{eff,+}(\sigma') = \sigma' - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma',+}). \tag{7.3.12}$$

We obtain $\mu^{\delta,p}_{\sigma,+} = \mu_{\sigma',+} = \mu^{\delta,p}_{\tilde{\sigma},+}$. Since $g_{eff,+}$ is increasing we have $g_{eff,+}(\sigma_c) > g_{eff,+}(\sigma') = \tilde{\sigma}$, we obtain from (7.3.12)

$$\begin{aligned} g_{eff,+}(\sigma_c) >& \sigma' - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu_{\sigma',+}) = \sigma' - \frac{\delta L_W^2}{2(p-1)}\kappa_2(\mu^{\delta,p}_{\sigma,+}), \\ \text{i.e} \quad \kappa_2(\mu^{\delta,p}_{\sigma,+}) >& \frac{2(p-1)}{\delta L_W^2}\left(\sigma' - g_{eff,+}(\sigma_c)\right). \end{aligned}$$

Plugging that back into (7.3.11), we obtain

$$\sigma' > \sigma + \sigma' - g_{eff,+}(\sigma_c), \quad \text{i.e} \quad \sigma < g_{eff,+}(\sigma_c),$$

which contradicts the initial assumption.

This concludes the proof of Theorem 7.1.3.

7.4 Some remaining questions before submission

We list here the main questions we would like to tackle before submitting the work presented in this chapter for publication.

A better link to the effective dynamics : The main goal of this work is the study of the nonlinear effective dynamics (Eff) linked to the numerical scheme (D-RB-IPS), and this is justified by the fact that this numerical scheme is closer to the effective dynamics than it is to the McKean-Vlasov SDE (NL). However this last affirmation is not quantified in the present work. This is due to several issues.

- The usual way of quantifying the distance between a numerical scheme and its effective dynamics, or at least the one we are aware of, is done in terms of weak error, i.e show that there is $\beta > 0$ such that for any given smooth function f we have $\left|\mathbb{E}f\left(\bar{X}^{e,\delta,p}_t\right) - \mathbb{E}f\left(Y^{1,\delta,p}_t\right)\right| = O(\delta^\beta)$ (see for instance [157]), and often in linear models. On the other hand, works quantifying the propagation of chaos or the convergence of the numerical scheme to the nonlinear limit, for instance [100, 177], focus on a stronger error in Wasserstein distance via coupling methods. We are thus trying to show the convergence of (D-RB-NL) towards (Eff) by using a coupling method and comparing the random variable G_p in (7.1.8) with a Gaussian variable that can then be compared to the Brownian motion in (Eff). The construction of the closest Brownian motion to G_p can be done thanks to a Central Limit Theorem in L^2-Wasserstein distance (see [24]). But we also keep in mind that weak errors and strong errors might be of different orders.

- We could try and go the other way : first consider the effective dynamics of the linear numerical scheme (D-RB-IPS) and then, via usual propagation of chaos results that allow for the diffusion coefficient to depend on the empirical measure, show that (Eff) is the limit as $N \to \infty$ of said (linear) effective dynamics. However, doing so, we lose the independence of the particles that facilitates the construction of the Gaussian random variable appearing when $\delta \to 0$.

- The convergence as $N \to \infty$ of (D-RB-IPS), either towards (NL) or (D-RB-NL), is obtained under (one-sided) Lipschitz assumptions [100, 177], and uniformly in time in models where such mean field limits are known for (IPS). Here, we cannot expect to obtain a uniform in time result, as this is precisely prevented by the non uniqueness of the stationary distribution for (NL). Some results of uniform in time propagation of chaos for models without a unique invariant measure have been obtained [57], but to our knowledge in specific cases. The lack of uniformity in time of the convergence in N prevents us from quantifying the distance between the invariant measures that we study.

- Giving a quantitative link between the various processes (IPS), (D-IPS), (D-RB-IPS), (D-RB-NL), (Eff), and (NL) would thus require an entire separate analysis, even though some results are already known, and we believe that it would dilute the main message of this work concerning the phase transition of the effective dynamics.

Improving some proofs : The proof of Lemma F.1.2 below relies on a numerical simulation, and we would like to rid of this. Similarly, the proof that the variance is Lipschitz continuous in Lemma 7.3.2 feels a bit too technical, and we seek to simplify it.

Part IV

Appendices

Appendix A

Appendix of the introduction

A.1 Proof of Lemma 1.1.2

1. implies 2. : Let $k \geqslant 1$ and $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be bounded continuous. Since convergence in probability implies convergence in distribution, and since ϕ is assumed to be bounded, we deduce from 1. that

$$\mathbb{E}^{F_N}\left(\left|\frac{1}{N}\sum_{i=1}^{N}\phi(X_i) - \int \phi(x)df(x)\right|^k\right) \xrightarrow[N\to\infty]{} 0.$$

For $k = 1$, this yields

$$\mathbb{E}^{F_N}\left(\frac{1}{N}\sum_{i=1}^{N}\phi(X_i)\right) \xrightarrow[N\to\infty]{} \int \phi(x)df(x),$$

and by symmetry of F_N we have $\mathbb{E}^{F_N}\left(\frac{1}{N}\sum_{i=1}^{N}\phi(X_i)\right) = \mathbb{E}^{F_N^1}\phi(X)$, and thus F_N^1 converges in distribution towards f.

For $k = 2$, let $\phi, \psi : \mathbb{R}^d \mapsto \mathbb{R}$ be two bounded continuous functions. By Cauchy-Schwarz,

$$\begin{aligned}
&\mathbb{E}^{F_N}\left(\left(\frac{1}{N}\sum_{i=1}^{N}\phi(X_i) - \int \phi(x)df(x)\right)\left(\frac{1}{N}\sum_{i=1}^{N}\psi(X_i) - \int \psi(x)df(x)\right)\right)\\
&\quad\leqslant \mathbb{E}^{F_N}\left(\left|\frac{1}{N}\sum_{i=1}^{N}\phi(X_i) - \int \phi(x)df(x)\right|^2\right)^{1/2}\mathbb{E}^{F_N}\left(\left|\frac{1}{N}\sum_{i=1}^{N}\psi(X_i) - \int \psi(x)df(x)\right|^2\right)^{1/2}\\
&\quad\xrightarrow[N\to\infty]{} 0.
\end{aligned}$$

We have

$$\mathbb{E}^{F_N}\left(\left(\frac{1}{N}\sum_{i=1}^{N}\phi(X_i) - \int \phi(x)df(x)\right)\left(\frac{1}{N}\sum_{i=1}^{N}\psi(X_i) - \int \psi(x)df(x)\right)\right)$$

$$=\mathbb{E}^{F_N}\left(\left(\frac{1}{N}\sum_{i=1}^N \phi(X_i)\right)\left(\frac{1}{N}\sum_{i=1}^N \psi(X_i)\right)\right)+\left(\int \phi(x)df(x)\right)\left(\int \psi(x)df(x)\right)$$
$$-\int \phi(x)df(x)\mathbb{E}^{F_N}\left(\frac{1}{N}\sum_{i=1}^N \psi(X_i)\right)-\int \psi(x)df(x)\mathbb{E}^{F_N}\left(\frac{1}{N}\sum_{i=1}^N \phi(X_i)\right).$$

Since

$$\mathbb{E}^{F_N}\left(\left(\frac{1}{N}\sum_{i=1}^N \phi(X_i)\right)\left(\frac{1}{N}\sum_{i=1}^N \psi(X_i)\right)\right)$$
$$=\frac{1}{N^2}\mathbb{E}^{F_N}\left(\sum_{i=1}^N \phi(X_i)\psi(X_i)\right)+\frac{1}{N^2}\mathbb{E}^{F_N}\left(\sum_{i\neq j}\phi(X_i)\psi(X_j)\right)$$
$$=\frac{1}{N}\mathbb{E}^{F_N^1}\phi\psi+\frac{N-1}{N}\mathbb{E}^{F_N^2}\phi\otimes\psi,$$

where we used the symmetry of F_N for this last equality, we have

$$\frac{1}{N}\mathbb{E}^{F_N^1}\phi\psi+\frac{N-1}{N}\mathbb{E}^{F_N^2}\phi\otimes\psi+\left(\int \phi(x)df(x)\right)\left(\int \psi(x)df(x)\right)$$
$$-\int \phi(x)df(x)\mathbb{E}^{F_N^1}\psi-\int \psi(x)df(x)\mathbb{E}^{F_N^1}\phi \xrightarrow[N\to\infty]{} 0,$$

i.e

$$\frac{1}{N}\mathbb{E}^{F_N^1}\phi\psi+\left(\frac{N-1}{N}\mathbb{E}^{F_N^2}\phi\otimes\psi-\left(\int \phi(x)df(x)\right)\left(\int \psi(x)df(x)\right)\right)$$
$$+\int \phi(x)df(x)\left(\int \psi(x)df(x)-\mathbb{E}^{F_N^1}\psi\right)+\int \psi(x)df(x)\left(\int \phi(x)df(x)-\mathbb{E}^{F_N^1}\phi\right)$$
$$\xrightarrow[N\to\infty]{} 0.$$

This implies $\mathbb{E}^{F_N^2}\phi\otimes\psi \xrightarrow[N\to\infty]{} \left(\int \phi(x)df(x)\right)\left(\int \psi(x)df(x)\right)$ and thus, by Remark 1.1.2, F_N^2 converges towards $f^{\otimes 2}$ for the weak convergence of measures.

The extension to values of k larger than 2 is done in the same manner.

2. implies 1. : Let $\epsilon > 0$. By Bienaymé-Chebyshev,

$$\mathbb{P}^{F_N}\left(\left|\frac{1}{N}\sum_{i=1}^N \phi(X_i)-\int \phi(x)df(x)\right|\geqslant\epsilon\right)\leqslant\frac{1}{\epsilon^2}\mathbb{E}^{F_N}\left(\left|\frac{1}{N}\sum_{i=1}^N \phi(X_i)-\int \phi(x)df(x)\right|^2\right).$$

Furthermore,

$$\mathbb{E}^{F_N}\left(\left(\frac{1}{N}\sum_{i=1}^N \phi(X_i)-\int \phi(x)df(x)\right)^2\right)=\frac{1}{N}\mathbb{E}^{F_N^1}\phi^2+\frac{N-1}{N}\mathbb{E}^{F_N^2}\phi^{\otimes 2}+\left(\int \phi(x)df(x)\right)^2$$
$$-2\int \phi(x)df(x)\mathbb{E}^{F_N^1}\phi,$$

so that, since by using the convergence of the k-marginals for $k=1,2$ we have $\mathbb{E}^{F_N^1}\phi \xrightarrow[N\to\infty]{} \int \phi(x)df(x)$ and $\mathbb{E}^{F_N^2}\phi^{\otimes 2} \xrightarrow[N\to\infty]{} \left(\int \phi(x)df(x)\right)^2$, we obtain the desired result.

A.2 Proof of Lemma 1.2.1

In this section, we shall detail the calculations leading up to the lemma starting from (1.2.3). The first term is controlled using the regularity of K.

$$\begin{aligned}\left|\frac{1}{N}\sum_{j=1}^N K(X_s^i - X_t^j) - \frac{1}{N}\sum_{j=1}^N K(\bar{X}_s^i - \bar{X}_t^j)\right| \leqslant & \frac{1}{N}\sum_{j=1}^N |K(X_t^i - X_t^j) - K(\bar{X}_t^i - \bar{X}_t^j)| \\ \leqslant & L|X_t^i - \bar{X}_t^i| + \frac{L}{N}\sum_{j=1}^N |X_t^j - \bar{X}_t^j|.\end{aligned}$$

Then, the second term is controlled by

$$\begin{aligned}\left|\frac{1}{N}\sum_{j=1}^N K(\bar{X}_t^i - \bar{X}_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right| \leqslant & \left|\frac{1}{N}\sum_{j=1}^N K(\bar{X}_t^i - \bar{X}_t^j) - \frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i - \bar{X}_t^j)\right| \\ & + \left|\frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i - \bar{X}_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right| \\ \leqslant & L\left(\frac{1}{N-1} - \frac{1}{N}\right)\sum_{j\neq i} |\bar{X}_t^i - \bar{X}_t^j| \\ & + \left|\frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i - \bar{X}_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right| \\ \leqslant & L\left(\frac{1}{N-1} - \frac{1}{N}\right)\left((N-1)|\bar{X}_t^i| + \sum_{j\neq i} |\bar{X}_t^j|\right) \\ & + \left|\frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i - \bar{X}_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right|\end{aligned}$$

Let us focus on the last term. Notice that the random variables $\bar{X}_t^i$ are independent identically distributed, with law $\bar{\rho}_t$, such that for all $j \neq i$, we have $K*\bar{\rho}_t(\bar{X}_t^i) = \mathbb{E}\left(K(\bar{X}_t^i - \bar{X}_t^j)\middle|\bar{X}_t^i\right)$. We thus deal with this last term through some form of law of large number.

$$\begin{aligned}&\mathbb{E}\left|\frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i - \bar{X}_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right| \\ &\qquad \leqslant \mathbb{E}\left(\left|\frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i - \bar{X}_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right|^2\right)^{1/2}\end{aligned}$$

$$
\begin{aligned}
&\leqslant \mathbb{E}\left(\mathbb{E}\left(\left|\frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i-\bar{X}_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right|^2 \middle| \bar{X}_t^i\right)\right)^{1/2}\\
&= \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{N-1}\sum_{j\neq i} K(\bar{X}_t^i-\bar{X}_t^j)\middle|\bar{X}_t^i\right)\right)^{1/2}\\
&= \mathbb{E}\left(\frac{1}{N-1}\mathrm{Var}\left(K(\bar{X}_t^i-\bar{X}_t^j)\middle|\bar{X}_t^i\right)\right)^{1/2}\\
&\leqslant \frac{L}{\sqrt{N-1}}\mathbb{E}\left(\mathrm{Var}\left(|\bar{X}_t^i-\bar{X}_t^j|\middle|\bar{X}_t^i\right)\right)^{1/2}\\
&\leqslant \frac{2L}{\sqrt{N-1}}\mathbb{E}\left(|\bar{X}_t^i|^2\right)^{1/2} \leqslant 2L\sqrt{\frac{C_0}{N-1}},
\end{aligned}
$$

where we used Assumption (1.2.2). Thus

$$
\begin{aligned}
&\mathbb{E}\left|\frac{1}{N}\sum_{j=1}^N K(X_t^i-X_t^j) - K*\bar{\rho}_t(\bar{X}_t^i)\right|\\
&\qquad\leqslant L\mathbb{E}|X_t^i-\bar{X}_t^i| + \frac{L}{N}\sum_{j=1}^N \mathbb{E}|X_t^j-\bar{X}_t^j| + \frac{2LC_0^{1/2}}{N} + 2L\sqrt{\frac{C_0}{N-1}}.
\end{aligned}
$$

By exchangeability $\forall j, \mathbb{E}|X_t^j-\bar{X}_t^j| = \mathbb{E}|X_t^i-\bar{X}_t^i|$, which concludes the proof.

A.3 Proof of Lemma 1.2.2

We present here the formal proof of Lemma 1.2.2. We have

$$
\begin{aligned}
\partial_t \log P_t^2 = \frac{\partial_t P_t^2}{P_t^2} &= -\nabla\cdot b^2 - b^2\cdot\frac{\nabla P_t^2}{P_t^2} + \sigma\frac{\Delta P_t^2}{P_t^2}\\
&= -\nabla\cdot b^2 - b^1\cdot\frac{\nabla P_t^2}{P_t^2} + \sigma\frac{\Delta P_t^2}{P_t^2} - (b^2-b^1)\cdot\frac{\nabla P_t^2}{P_t^2}
\end{aligned}
$$

Hence

$$
\begin{aligned}
\partial_t\left(P_t^1\log P_t^2\right) =& P_t^1\partial_t\log P_t^2 + (\partial_t P_t^1)\log P_t^2\\
=& -\nabla\cdot b^2 P_t^1 - b^1\cdot\nabla P_t^2\frac{P_t^1}{P_t^2} + \sigma\Delta P_t^2\frac{P_t^1}{P_t^2} - (b^2-b^1)\cdot\nabla P_t^2\frac{P_t^1}{P_t^2}\\
&-\nabla\cdot b^1 P_t^1\log P_t^2 - b^1\cdot\nabla P_t^1\log P_t^2 + \sigma\Delta P_t^1\log P_t^2.
\end{aligned}
$$

We may now calculate

$$
\begin{aligned}
\frac{d}{dt}\int P_t^1\log P_t^2 =& \int P_t^1\left(b^1-b^2\right)\nabla\log P_t^2 + \sigma\int\left(\Delta P_t^2\frac{P_t^1}{P_t^2} + \Delta P_t^1\log P_t^2\right) - \int\nabla b^2 P_t^1\\
&-\int b^1\cdot\nabla\log P_t^2 P_t^1 - \int\nabla\cdot b^1 P_t^1\log P_t^2 - \int b^1\cdot\nabla P_t^1\log P_t^2.
\end{aligned}
$$

By integration by part, the last line yields a term equal to 0. Furthermore

$$\partial_t(P_t^1 \log P_t^1) = \partial_t P_t^1(\log P_t^1 + 1),$$

hence

$$\frac{d}{dt}\int P_t^1 \log P_t^1 = \int \partial_t P_t^1 \log P_t^1 + \frac{d}{dt}\int P_t^1.$$

Since we are dealing with probability measures, of total mass one, the last term is 0.

$$\frac{d}{dt}\int P_t^1 \log P_t^1 = -\int \nabla\cdot b^1 P_t^1 \log P_t^1 - \int b^1\cdot\nabla P_t^1 \log P_t^1 + \sigma\int \Delta P_t^1 \log P_t^1.$$

We may now calculate

$$\begin{aligned}
-\frac{d}{dt}\mathcal{H}(P_t^1, P_t^2) =& \frac{d}{dt}\int P_t^1 \log P_t^2 - \frac{d}{dt}\int P_t^1 \log P_t^1\\
=& \int P_t^1\left(b^1 - b^2\right)\nabla \log P_t^2 + \sigma\int\left(\Delta P_t^2\frac{P_t^1}{P_t^2} + \Delta P_t^1 \log P_t^2 - \Delta P_t^1 \log P_t^1\right)\\
&-\int \nabla\cdot b^2 P_t^1 + \int \nabla\cdot b^1 P_t^1 \log P_t^1 + \int b^1\cdot\nabla P_t^1 \log P_t^1.
\end{aligned}$$

Integrating by parts, we get

$$\int \nabla\cdot b^1 P_t^1 \log P_t^1 = -\int b^1\cdot\nabla P_t^1 \log P_t^1 - \int b^1\cdot\nabla P_t^1,$$

thus

$$\int \nabla\cdot b^1 P_t^1 \log P_t^1 + \int b^1\cdot\nabla P_t^1 \log P_t^1 = \int \nabla\cdot b^1 P_t^1.$$

Finally

$$\begin{aligned}
-\frac{d}{dt}\mathcal{H}(P_t^1, P_t^2) = &\int P_t^1\left(b^1 - b^2\right)\cdot\nabla\log P_t^2 - \int P_t^1\left(\nabla\cdot b^2 - \nabla\cdot b^1\right)\\
&+\sigma\int\left(\Delta P_t^2\frac{P_t^1}{P_t^2} + \Delta P_t^1 \log P_t^2 - \Delta P_t^1 \log P_t^1\right).
\end{aligned}$$

Direct calculations, involving various integration by parts, yield

$$\int\left(\Delta P_t^2\frac{P_t^1}{P_t^2} + \Delta P_t^1 \log P_t^2 - \Delta P_t^1 \log P_t^1\right) = \int P_t^1\left|\nabla\log\frac{P_t^1}{P_t^2}\right|^2.$$

Hence (1.2.6). To obtain (1.2.7), it is sufficient to observe

$$-\int P_t^1\left(\nabla\cdot b^1 - \nabla\cdot b^2\right) = \int P_t^1\nabla\log P_t^1\cdot\left(b^1 - b^2\right).$$

Appendix B

Appendix of Chapter 2

B.1 Various results

B.1.1 Proof of lemma 2.1.1

The property only depends on the distance to the the origin, not the direction. We therefore only need to prove it in dimension 1, making sure the constant $\tilde{A}$ is independent of the direction. There is $x_0 > 0$ such that $\frac{\lambda}{2}x_0^2 = 2A$. Therefore, for $x \geqslant 0$, using (2.1.4):

$$U'(x_0+x)(x_0+x) \geqslant 2\lambda U(x_0+x) + \frac{\lambda}{2}(x_0+x)^2 - 2A = 2\lambda U(x_0+x) + \frac{\lambda}{2}x^2 + \lambda x x_0.$$

Then, for $x \geqslant 0$:

$$\begin{aligned} U(x_0+x) - U(x_0) &= \int_0^1 U'(x_0+tx)\,x\,dt = \int_0^1 U'(x_0+tx)(x_0+tx)\frac{x}{x_0+tx}dt \\ &\geqslant \frac{x}{x_0+x}\int_0^1 2\lambda U(x_0+tx) + \frac{\lambda}{2}t^2x^2 + \lambda t x x_0 dt \\ &\geqslant \frac{x}{x_0+x}\left(\frac{\lambda}{6}x^2 + \frac{\lambda}{2}x x_0\right) \qquad \text{since } U \geqslant 0 \\ &= \frac{\lambda}{6}\frac{x^3}{x_0+x} + \frac{\lambda}{2}\frac{x^2x_0}{x_0+x}. \end{aligned}$$

We thus have for all $x \geqslant x_0$:

$$\begin{aligned} U(x) - U(x_0) &\geqslant \frac{\lambda}{6}\frac{(x-x_0)^3}{x} + \frac{\lambda}{2}\frac{(x-x_0)^2x_0}{x} \\ &= \frac{\lambda}{6}x^2 - \frac{\lambda}{2}x x_0 + \frac{\lambda}{2}x_0^2 - \frac{\lambda}{6}\frac{x_0^3}{x} + \frac{\lambda}{2}x x_0 - \lambda x_0^2 + \frac{\lambda}{2}\frac{x_0^3}{x} \\ &= \frac{\lambda}{6}x^2 - \frac{\lambda}{2}x_0^2 + \frac{\lambda}{3}\frac{x_0^3}{x}. \end{aligned}$$

However, $-\frac{\lambda}{2}x_0^2 + \frac{\lambda}{3}\frac{x_0^3}{x} \geqslant -\frac{\lambda}{2}x_0^2 = -2A$ for $x \geqslant x_0$. We therefore have the desired result for $x \geqslant x_0$. The same reasoning gives us the result for $x \leqslant -x_0$.

Hence, if $|x| \geqslant |x_0| = \sqrt{\frac{4A}{\lambda}}$, $U(x) - U(x_0) \geqslant \frac{\lambda}{6}x^2 - 2A$. We then use the fact that $U(x)$ is continuous on the sphere of center 0 and radius $\sqrt{\frac{4A}{\lambda}}$, hence bounded on this set, to give a lower bound on $U(x_0)$ independent of the direction. Finally, for $|x| \in [-x_0, x_0]$, the function $x \mapsto U(x) - \frac{\lambda}{6}x^2$ is continuous, therefore bounded.

B.1.2 Proof of lemma 2.1.2

We have

$$\nabla W * \mu(x) - \nabla W * \nu(\tilde{x}) = \nabla W * \mu(x) - \nabla W * \mu(\tilde{x}) + \nabla W * \mu(\tilde{x}) - \nabla W * \nu(\tilde{x})$$

Let $(X, \tilde{X})$ be a coupling of μ and ν. Then

$$\begin{aligned}|\nabla W * \mu_t(x) - \nabla W * \tilde{\mu}_t(\tilde{x})| =& \left|\mathbb{E}\left(\nabla W(x - X) - \nabla W(\tilde{x} - \tilde{X})\right)\right| \\ \leqslant& L_W \mathbb{E}\left(|x - X - \tilde{x} + \tilde{X}|\right) \\ \leqslant& L_W \mathbb{E}\left(|x - \tilde{x}| + |X - \tilde{X}|\right)\end{aligned}$$

This being true for all coupling, we obtain the desired result.

B.1.3 Proof of Lemma 2.2.1

Remark B.1.1. *With γ given by* (2.2.1)*, we have $\gamma \leqslant \frac{1}{2}$.*

We have

$$\begin{aligned}\mathcal{L}_\mu H(x,v) =& v \cdot \nabla_x H(x,v) - v \cdot \nabla_v H(x,v) - \nabla U(x) \cdot \nabla_v H(x,v) \\ & - \nabla W * \mu(x) \cdot \nabla_v H(x,v) + \Delta_v H(x,v) \\ =& v \cdot (24\nabla U(x) + 12(1-\gamma)x + 2\lambda x + 12v) - v \cdot (12x + 24v) \\ & - \nabla U(x) \cdot (12x + 24v) - \nabla W * \mu(x) \cdot (12x + 24v) + 24d \\ =& 24d - 12\nabla U(x) \cdot x + x \cdot v(12(1-\gamma) + 2\lambda - 12) \\ & - \nabla W * \mu(x) \cdot (12x + 24v) - 12|v|^2,\end{aligned}$$

with

$$\begin{aligned}-\gamma H(x,v) =& -24\gamma U(x) - 6\gamma(1-\gamma)|x|^2 - \gamma\lambda|x|^2 - 12\gamma x \cdot v - 12\gamma|v|^2 \\ -12\nabla U(x) \cdot x \leqslant& -24\lambda U(x) - 6\lambda|x|^2 + 24A,\end{aligned}$$

and

$$\begin{aligned}&-\nabla W * \mu(x) \cdot (12x + 24v) \\ &\quad \leqslant (L_W|x| + L_W \mathbb{E}_\mu(|\cdot|))(12|x| + 24|v|) \\ &\quad \leqslant 12L_W|x|^2 + 24L_W|x||v| \\ &\qquad + L_W \mathbb{E}_\mu(|\cdot|)\left(6\frac{|x|^2}{a_x \mathbb{E}_\mu(|\cdot|)} + 6a_x \mathbb{E}_\mu(|\cdot|) + 12\frac{|v|^2}{a_v \mathbb{E}_\mu(|\cdot|)} + 12a_v \mathbb{E}_\mu(|\cdot|)\right),\end{aligned}$$

where this last inequality holds for any $a_x, a_v > 0$. Therefore

$$\begin{aligned}\mathcal{L}_\mu H(x,v) \leqslant & 24A + 24d + 6L_W \mathbb{E}_\mu(|\cdot|)^2(a_x + 2a_v) - \gamma H(x,v) + 24\gamma U(x) + 6\gamma(1-\gamma)|x|^2 \\ & + \gamma\lambda|x|^2 + 12\gamma x\cdot v + 12\gamma|v|^2 - 24\lambda U(x) - 6\lambda|x|^2 + 12L_W|x|^2 + 24L_W|x||v| \\ & + \frac{6L_W}{a_x}|x|^2 + \frac{12L_W}{a_v}|v|^2 + x\cdot v\,(12(1-\gamma) + 2\lambda - 12) - 12|v|^2,\end{aligned}$$

and then

$$\begin{aligned}\mathcal{L}_\mu H(x,v) \leqslant & 24A + 24d + 6L_W \mathbb{E}_\mu(|\cdot|)^2(a_x + 2a_v) - \gamma H(x,v) + 24U(x)(\gamma - \lambda) \\ & + |x||v|\left(|12\gamma + 12(1-\gamma) + 2\lambda - 12| + 24L_W\right) \\ & + |x|^2\left(6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + 12L_W + \frac{6L_W}{a_x}\right) \\ & + |v|^2\left(12\gamma - 12 + \frac{12L_W}{a_v}\right).\end{aligned}$$

We now use $|x||v| \leqslant \frac{\lambda}{3}|x|^2 + \frac{3}{4\lambda}|v|^2$, and $|12\gamma\lambda + 12(1-\gamma\lambda) + 2\lambda - 12| = 2\lambda$.

We have $(\gamma - \lambda) < 0$. Hence $24U(x)(\gamma-\lambda) \leqslant 4\lambda(\gamma-\lambda)|x|^2 - 24(\gamma-\lambda)\tilde{A}$ using Lemma 2.1.1. Then

$$\begin{aligned}\mathcal{L}_\mu H(x,v) \leqslant & \, 24A - 24(\gamma-\lambda)\tilde{A} + 24d + 6L_W\mathbb{E}_\mu(|\cdot|)^2(a_x + 2a_v) - \gamma H(x,v) \\ & + |x|^2\left(4\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{6L_W}{a_x} + 12L_W + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3}\right) \\ & + |v|^2\left(12\gamma - 12 + \frac{12L_W}{a_v} + \frac{3}{2} + \frac{3}{4\lambda}24L_W\right).\end{aligned}$$

We now consider each term individually.

Coefficient of $|x|^2$. We have, using $0 < \gamma < 1$ and $L_W \leqslant \frac{\lambda}{8}$

$$\begin{aligned}4\lambda(\gamma-\lambda) + & 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W \\ & \leqslant \gamma(5\lambda+6) - \left(4\lambda^2 + 6\lambda - \frac{2\lambda^2}{3} - \lambda^2 - \frac{3\lambda}{2}\right).\end{aligned}$$

Therefore, it is sufficient that

$$\gamma \leqslant \lambda\frac{\frac{7}{3}\lambda + \frac{9}{2}}{5\lambda+6}.$$

We check this holds for $\gamma = \frac{\lambda}{2\lambda+2}$. Then

$$\begin{aligned}4\lambda(\gamma-\lambda) + & 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W + \frac{6L_W}{a_x} \\ & \leqslant (5\lambda+6)\left(\gamma - \frac{\frac{7}{3}\lambda^2 + \frac{9}{2}\lambda}{5\lambda+6}\right) + \frac{3\lambda}{4a_x}.\end{aligned}$$

We therefore choose

$$\frac{3\lambda}{4a_x} \leqslant -(5\lambda+6)\left(\gamma - \frac{\frac{7}{3}\lambda + \frac{9}{2}\lambda}{5\lambda+6}\right) = \frac{7}{3}\lambda^2 + \frac{9}{2}\lambda - \frac{5\lambda^2+6\lambda}{2\lambda+2}.$$

It is, for that, sufficient to take

$$\frac{3\lambda}{4a_x} = \frac{3}{4}\lambda, \quad \text{i.e} \quad a_x = 1.$$

Furthermore

$$\begin{aligned} &6\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W + \frac{6L_W}{a_x} \\ &\leqslant \frac{5\lambda^2+6\lambda}{2\lambda+2} - \frac{7}{3}\lambda^2 - \frac{9}{2}\lambda + \frac{3}{4}\lambda = \frac{5\lambda^2+6\lambda}{2\lambda+2} - \frac{7}{3}\lambda^2 - \frac{15}{4}\lambda. \end{aligned}$$

We then observe

$$6\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{6L_W}{a_x} + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W \leqslant -\lambda^2 - \frac{3}{4}\lambda.$$

And finally, for all $\lambda > 0$ and for all x

$$\begin{aligned} &|x|^2\left(6\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{6L_W}{a_x} + \frac{2\lambda^2}{3} + \frac{48L_W\lambda}{3} + 12L_W\right) \\ &\leqslant -\lambda^2|x|^2 - \frac{3}{4}\lambda|x|^2 \end{aligned}$$

Coefficient of $|v|^2$. We have, using $0 < \gamma \leqslant \frac{1}{2}$ and $L_W \leqslant \lambda/8$

$$12\gamma - 12 + \frac{3}{2} + \frac{3}{4\lambda}24L_W \leqslant -6 + \frac{3}{2} + \frac{18}{\lambda}\cdot\frac{\lambda}{8} = -6 + \frac{3}{2} + \frac{9}{4} = -\frac{9}{4}.$$

We then choose

$$\frac{12\lambda}{8a_v} = \frac{9}{4}, \quad \text{i.e} \quad a_v = \frac{2}{3}\lambda.$$

Therefore

$$\forall\lambda > 0,\ \forall v, \quad |v|^2\left(12\gamma - 12 + \frac{3}{2} + \frac{3}{4\lambda}24L_W + \frac{12L_W}{a_v}\right) \leqslant 0.$$

We thus obtain

$$\begin{aligned} &\mathcal{L}_\mu H(x,v) \\ &\leqslant 24\left(A - (\gamma-\lambda)\tilde{A} + d\right) + 6L_W\mathbb{E}_\mu(|\cdot|)^2\left(1 + \frac{4}{3}\lambda\right) - \lambda^2|x|^2 - \frac{3}{4}\lambda|x|^2 - \gamma H(x,v), \end{aligned}$$

i.e

$$\mathcal{L}_\mu H(x,v) \leqslant 24\left(A-(\gamma-\lambda)\tilde{A}+d\right)+\mathbb{E}_\mu(|\cdot|)^2\left(\frac{3}{4}\lambda+\lambda^2\right)-\lambda^2|x|^2-\frac{3}{4}\lambda|x|^2-\gamma H(x,v). \tag{B.1.1}$$

B.1.4 Proof of Lemma 2.2.3

Using $1-\gamma \geqslant \frac{1}{2}$, we get

$$H(x,v) \geqslant 24U(x)+(3+\lambda)|x|^2+12\left|v+\frac{x}{2}\right|^2-3|x|^2,$$

which is (2.2.2). We then have

$$H(x,v) \geqslant \min\left(\frac{2}{3}\lambda,6\right)\left(|v|^2+|x+v|^2\right).$$

Thus

$$\begin{aligned}
r(x,\tilde{x},v,\tilde{v})^2 &\leqslant \left((1+\alpha)|x-\tilde{x}+v-\tilde{v}|+\alpha|v-\tilde{v}|\right)^2\\
&\leqslant 2(1+\alpha)^2|x-\tilde{x}+v-\tilde{v}|^2+2\alpha^2|v-\tilde{v}|^2\\
&\leqslant 4\left((1+\alpha)^2+\alpha^2\right)\left(|x+v|^2+|v|^2+|\tilde{x}+\tilde{v}|^2+|\tilde{v}|^2\right).
\end{aligned}$$

Therefore we obtain the final point.

B.1.5 Proof of control of L1 and L2 Wasserstein distances

We prove Lemma 2.2.6. Using the definition of R_1 and (2.2.1), and since $B \geqslant d \geqslant 1$ and $\gamma \leqslant \frac{1}{2}$, we have $R_1 \geqslant 1$.

- First for the L1-Wasserstein distance

  $$|x-x'|+|v-v'| \leqslant |v-v'+x-x'|+2|x-x'| \leqslant \max\left(\frac{2}{\alpha},1\right)r((x,v),(x',v')).$$

  If $r((x,v),(x',v')) \leqslant 1 \leqslant R_1$

  $$r((x,v),(x',v')) \leqslant \frac{f(r)}{f'_-(R_1)} \leqslant \frac{\rho((x,v),(x',v'))}{\phi(R_1)g(R_1)}.$$

  If $r((x,v),(x',v')) \geqslant 1$, we have shown (2.2.7)

  $$r((x,v),(x',v')) \leqslant r^2((x,v),(x',v')) \leqslant 4\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(H(x,v)+H(x',v')\right).$$

  Thus

  $$r((x,v),(x',v')) \leqslant \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(\epsilon H(x,v)+\epsilon H(x',v')\right)$$

$$\leqslant \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left(\left(x,v\right),\left(x',v'\right)\right)}{f\left(r\right)}$$
$$\leqslant \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left(\left(x,v\right),\left(x',v'\right)\right)}{f\left(1\right)}.$$

Therefore

$$|x-x'|+|v-v'| \leqslant \max\left(\frac{2}{\alpha},1\right)\max\left(\frac{4\left((1+\alpha)^2+\alpha^2\right)}{\epsilon\min\left(\frac{2}{3}\lambda,6\right)f\left(1\right)},\frac{1}{\phi\left(R_1\right)g\left(R_1\right)}\right)\rho\left(\left(x,v\right),\left(x',v'\right)\right).$$

- Then for the L2-Wasserstein distance

$$|v-v'|^2=|v-v'+x-x'-\left(x-x'\right)|^2\leqslant 2|v-v'+x-x'|^2+2|x-x'|^2.$$

Hence

$$|x-x'|^2+|v-v'|^2\leqslant 3\left(|v-v'+x-x'|^2+|x-x'|^2\right).$$

But

$$\begin{aligned}
r^2\left(\left(x,v\right),\left(x',v'\right)\right) &= \left(\alpha|x-x'|+|x-x'+v-v'|\right)^2\\
&\geqslant \alpha^2|x-x'|^2+|x-x'+v-v'|^2\\
&\geqslant \left(1+\alpha^2\right)\left(|x-x'|^2+|x-x'+v-v'|^2\right)\\
&\geqslant \frac{1+\alpha^2}{3}\left(|x-x'|^2+|v-v'|^2\right).
\end{aligned}$$

If $r\left(\left(x,v\right),\left(x',v'\right)\right)\leqslant 1\leqslant R_1$

$$r^2\left(\left(x,v\right),\left(x',v'\right)\right)\leqslant r\left(\left(x,v\right),\left(x',v'\right)\right)\leqslant\frac{f\left(r\right)}{f'_-\left(R_1\right)}\leqslant\frac{\rho\left(\left(x,v\right),\left(x',v'\right)\right)}{\phi\left(R_1\right)g\left(R_1\right)}.$$

If $r\left(\left(x,v\right),\left(x',v'\right)\right)\geqslant 1$, we have shown (2.2.7)

$$r^2\left(\left(x,v\right),\left(x',v'\right)\right)\leqslant 4\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(H\left(x,v\right)+H\left(x',v'\right)\right).$$

Thus

$$\begin{aligned}
r\left(\left(x,v\right),\left(x',v'\right)\right) &\leqslant \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(\epsilon H\left(x,v\right)+\epsilon H\left(x',v'\right)\right)\\
&\leqslant \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left(\left(x,v\right),\left(x',v'\right)\right)}{f\left(r\right)}\\
&\leqslant \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left(\left(x,v\right),\left(x',v'\right)\right)}{f\left(1\right)}.
\end{aligned}$$

Therefore

$$|x-x'|^2+|v-v'|^2 \leqslant \frac{3}{1+\alpha^2}\max\left(\frac{4\left((1+\alpha)^2+\alpha^2\right)}{\epsilon\min\left(\frac{2}{3}\lambda,6\right)f(1)},\frac{1}{\phi(R_1)\,g(R_1)}\right)\rho\left((x,v),(x',v')\right).$$

B.1.6 Proof of Lemma 2.2.7

We have

$$\begin{aligned}H(x,v)-H(\tilde{x},\tilde{v}) &=24\left(U(x)-U(\tilde{x})\right)+\left(6(1-\gamma)+\lambda\right)\left(|x|^2-|\tilde{x}|^2\right)+12\left(x\cdot v-\tilde{x}\cdot\tilde{v}\right)+12\left(|v|^2-|\tilde{v}|^2\right)\\ &=24\left(U(x)-U(\tilde{x})\right)+\left(6(1-\gamma)+\lambda-3\right)\left(|x|^2-|\tilde{x}|^2\right)+12\left(\left|v+\frac{x}{2}\right|^2-\left|\tilde{v}+\frac{\tilde{x}}{2}\right|^2\right).\end{aligned}$$

We first have

$$\left||x|^2-|\tilde{x}|^2\right|\leqslant|x-\tilde{x}|\left(|x|+|\tilde{x}|\right)\leqslant\frac{r(x,v,\tilde{x},\tilde{v})}{\alpha\sqrt{\lambda}}\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right).$$

Then

$$\begin{aligned}\left|\left|v+\frac{x}{2}\right|^2-\left|\tilde{v}+\frac{\tilde{x}}{2}\right|^2\right| &\leqslant\left|v+\frac{x}{2}-\tilde{v}-\frac{\tilde{x}}{2}\right|\left(\left|v+\frac{x}{2}\right|+\left|\tilde{v}+\frac{\tilde{x}}{2}\right|\right)\\ &\leqslant\frac{1}{\sqrt{12}}\left|v-\tilde{v}+\frac{1}{2}(x-\tilde{x})\right|\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right)\\ &\leqslant\frac{1}{2\sqrt{3}}\max\left(1,\frac{1}{2\alpha}\right)r(x,v,\tilde{x},\tilde{v})\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right).\end{aligned}$$

And finally

$$\begin{aligned}|U(x)-U(\tilde{x})| &=\left|\int_0^1\nabla U\left(\tilde{x}+t(x-\tilde{x})\right)\cdot(x-\tilde{x})dt\right|\\ &\leqslant\sup_{t\in[0,1]}\left|\nabla U\left(\tilde{x}+t(x-\tilde{x})\right)\right||x-\tilde{x}|\\ &\leqslant\left(\nabla U(0)+L_U(|x|+|\tilde{x}|)\right)|x-\tilde{x}|\\ &\leqslant\left(\nabla U(0)+\frac{L_U}{\sqrt{\lambda}}\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right)\right)\frac{r(x,v,\tilde{x},\tilde{v})}{\alpha}.\end{aligned}$$

These three inequalities yield the desired result.

B.2 Proof of Lemma 2.2.4

We first rewrite the various conditions on the parameters.

- Since for all $u\geqslant 0$, $0<\phi(u)\leqslant 1$, we have $0<\Phi(s)=\int_0^s\phi(u)\,du\leqslant s$, i.e $s/\Phi(s)\geqslant 1$.

Therefore

$$\inf_{r\in]0,R_1]}\frac{r\phi\left(r\right)}{\Phi\left(r\right)}\geqslant\inf_{r\in]0,R_1]}\phi\left(r\right)=\phi\left(R_1\right).$$

It is thus sufficient for (2.2.11) that

$$c+2\epsilon B\leqslant\frac{1}{2}\left(1-\frac{1}{\alpha}\left(L_U+L_W\right)\right)\phi\left(R_1\right).$$

- We have

$$\phi\left(r\right)\leqslant\exp\left(-\frac{L_U+L_W}{8\alpha}r^2\right).$$

So

$$\Phi\left(r\right)\leqslant\int_0^\infty\exp\left(-\frac{L_U+L_W}{8\alpha}s^2\right)ds=\sqrt{\frac{2\pi\alpha}{L_U+L_W}}.$$

Then

$$\int_0^{R_1}\frac{\Phi\left(r\right)}{\phi\left(r\right)}dr\leqslant\sqrt{\frac{2\pi\alpha}{L_U+L_W}}R_1\frac{1}{\phi\left(R_1\right)}.$$

It is thus sufficient for (2.2.12) that

$$c+2\epsilon B\leqslant2\sqrt{\frac{L_U+L_W}{2\pi\alpha}}\frac{\phi\left(R_1\right)}{R_1}.$$

At this point, we have now proven that under Assumption 2.1, Assumption 2.2 and Assumption 2.4, for the parameters to satisfy Lemma 2.2.4 it is sufficient for them to satisfy

$$\alpha>L_U+L_W, \tag{B.2.1}$$

$$c\leqslant\frac{\gamma}{6}\left(1-\frac{\frac{5}{6}\gamma}{2\epsilon B+\frac{5}{6}\gamma}\right), \tag{B.2.2}$$

$$c+2\epsilon B\leqslant\frac{1}{2}\left(1-\frac{1}{\alpha}\left(L_U+L_W\right)\right)\phi\left(R_1\right), \tag{B.2.3}$$

$$c+2\epsilon B\leqslant2\sqrt{\frac{L_U+L_W}{2\pi\alpha}}\frac{\phi\left(R_1\right)}{R_1}, \tag{B.2.4}$$

with, again

$$B=24\left(A+\left(\lambda-\gamma\right)\tilde{A}+d\right),\quad R_1=\sqrt{\left(1+\alpha\right)^2+\alpha^2}\sqrt{\frac{24}{5\gamma\min\left(3,\frac{1}{3}\lambda\right)}B}.$$

Let us show that there are positive parameters ϵ, α, L_W and c satisfying those conditions.

For inequality (B.2.1) it is sufficient, as $L_W < \frac{\lambda}{8}$, to consider

$$\alpha = L_U + \frac{\lambda}{4},$$

while inequality (B.2.2) first invites us to consider $2\epsilon B$ of a comparable order to c

$$2\epsilon B = \delta c.$$

We have thus switched parameter ϵ for δ. First we translate (B.2.2) into our new parameter:

$$\begin{aligned} c \leqslant \frac{\gamma}{6}\left(1 - \frac{\frac{5}{6}\gamma}{2\epsilon B + \frac{5}{6}\gamma}\right) &\iff c \leqslant \frac{\gamma}{6}\frac{\delta c}{\delta c + \frac{5}{6}\gamma} \\ &\iff 1 \leqslant \frac{\gamma}{6}\frac{\delta}{\delta c + \frac{5}{6}\gamma} \text{ (since } c \geqslant 0) \\ &\iff c \leqslant \frac{\gamma}{6}\frac{\delta - 5}{\delta}. \end{aligned}$$

The appearance of $\phi(R_1)$ in (B.2.3) and (B.2.4) suggests we should try to minimize it. Let us assume, for simplicity, that $\epsilon \leqslant 1$, which is equivalent to having $c \leqslant \frac{2B}{\delta}$. We then have

$$\begin{aligned} \phi(r) &= \exp\left(-\frac{1}{8}\left(\frac{1}{\alpha}(L_U + L_W) + \alpha + 96\epsilon \max\left(\frac{1}{2\alpha}, 1\right)\right) r^2\right) \\ &\geqslant \exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + 96 \max\left(\frac{1}{2\alpha}, 1\right)\right) r^2\right) \text{ on } [0, R_1]. \end{aligned} \tag{B.2.5}$$

Now, using (B.2.5), we have for (B.2.3) and (B.2.4) that it is sufficient that

$$c \leqslant \frac{1}{2(\delta+1)}\left(1 - \frac{1}{\alpha}(L_U + L_W)\right)\exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + 96 \max\left(\frac{1}{2\alpha}, 1\right)\right) R_1^2\right),$$

and

$$c \leqslant \frac{2}{\delta+1}\sqrt{\frac{L_U + L_W}{2\pi\alpha}} \exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + 96 \max\left(\frac{1}{2\alpha}, 1\right)\right) R_1^2\right).$$

We could now optimize parameter δ, but for the sake of conciseness, we choose $\delta = 6$.

Recall $0 \leqslant L_W < \frac{\lambda}{8}$. This way, both in (2.2.8) and (2.3.14), c and $\mathcal{C}_1$ can be bounded independently of L_W. Hence why L_W, in (2.3.14) and (2.4.17), can be chosen last, and every other quantities can be chosen independently of L_W.

Appendix C

Appendix of Chapter 4

C.1 Proof of Theorem 4.2.1

The proof is based on an iterative procedure, and relies heavily on the work of Ben-Artzi [13]. Let $\bar{\rho}^{(-1)} := 0$, and then for $k \in \mathbb{N}$ solve

$$\partial_t \bar{\rho}^{(k)} = -\left(u^{(k-1)} \cdot \nabla\right) \bar{\rho}^{(k)} + \Delta \bar{\rho}^{(k)}, \text{ in } \mathbb{R}^+ \times \mathbb{T}^d \tag{C.1.1}$$

$$u^{(k)} = K * \bar{\rho}^{(k)} \tag{C.1.2}$$

$$\bar{\rho}^{(k)}(0, \cdot) = \mu_0. \tag{C.1.3}$$

Let us recall the following lemma concerning the regularity of the second order parabolic equation. We refer to Chapter 7 of [74] for a proof on a bounded domain that can be extended to the torus.

Lemma C.1.1. *Let $a(t, x)$ be a $\mathcal{C}^\infty$ function on $\mathbb{R}^+ \times \mathbb{T}^d$ and $\psi_0 \in \mathcal{C}^\infty(\mathbb{T}^d)$. Then the problem*

$$\begin{aligned} &\partial_t \psi = -a \cdot \nabla \psi + \Delta \psi, \text{ in } \mathbb{R}^+ \times \mathbb{T}^d \\ &\psi(0, \cdot) = \psi_0, \end{aligned}$$

has a unique solution, which is $\mathcal{C}^\infty$.

Lemma C.1.2. *Suppose $\mu_0 \in \mathcal{C}^\infty(\mathbb{T}^d)$. Then the system* (C.1.1)-(C.1.3) *defines successively a sequence of $\mathcal{C}^\infty$ solutions $\{\bar{\rho}^{(k)}, u^{(k)}\}_{k \in \mathbb{N}}$.*
Furthermore, for all $t \geqslant 0$ and all $k \in \mathbb{N}$, $\|\bar{\rho}^{(k)}(t, \cdot)\|_{L^\infty} \leqslant \|\mu_0\|_{L^\infty}$ and $\|u^{(k)}(t, \cdot)\|_{L^\infty} \leqslant \|K\|_{L^1} \|\mu_0\|_{L^\infty}$.
Finally, given a final time $T \geqslant 0$, $\bar{\rho}^{(k)}$ (resp. $u^{(k)}$) and all its derivatives, both in time and in space, are bounded on $[0, T] \times \mathbb{T}^d$ uniformly in k

Proof. We use induction on k. The assertion is clear for $\bar{\rho}^{(0)}$ as the explicit solution to the heat equation. Suppose $\{\bar{\rho}^{(j)}, u^{(j-1)}\}_{j=0,\ldots,k}$ have be shown to be $\mathcal{C}^\infty$ solutions bounded uniformly in time.

Regularity. By definition

$$u^{(k)}(t, x) = K * \bar{\rho}^{(k)}(t, x) = \int_{\mathbb{T}^d} K(x - y) \bar{\rho}^{(k)}(t, y) dy = -\int_{\mathbb{T}^d} K(y) \bar{\rho}^{(k)}(t, x - y) dy.$$

Then

$$u^{(k)}(t,x) = -\int_{\mathbb{T}^d} \mathrm{div} V(y)\bar{\rho}^{(k)}(t,x-y)dy = -\int_{\mathbb{T}^d} V(y)\nabla_y \bar{\rho}^{(k)}(t,x-y)dy.$$

Since we are in the compact set $\mathbb{T}^d$, that $V \in L^\infty(\mathbb{T}^d)$, and that $\bar{\rho}^{(k)} \in \mathcal{C}^\infty(\mathbb{R}^+ \times \mathbb{T}^d)$ by induction hypothesis, we can easily show that $u^{(k)}$, as well as all its derivatives, are Lipschitz continuous. Hence $u^{(k)}$ is $\mathcal{C}^\infty$. Using Lemma C.1.1 in (C.1.1) with k replaced by $k+1$ yields the desired result for $\bar{\rho}^{(k+1)}$.

Boundedness of $\bar{\rho}^{(k+1)}$ and $u^{(k)}$. Let us show that for all $T \geqslant 0$, $\bar{\rho}^{(k+1)}$ and $u^{(k)}$ are both bounded on $[0,T] \times \mathbb{T}^d$, with a bound independent of T. We have, using Young's convolution inequality and the induction hypothesis

$$\|u^{(k)}(t,\cdot)\|_{L^\infty} \leqslant \|K\|_{L^1}\|\bar{\rho}^{(k)}(t,\cdot)\|_{L^\infty} \leqslant \|K\|_{L^1}\|\mu_0\|_{L^\infty}.$$

Now $\bar{\rho}^{(k+1)}$ is the unique solution of

$$\begin{aligned}\partial_t \bar{\rho}^{(k+1)} =& -\left(u^{(k)} \cdot \nabla\right)\bar{\rho}^{(k+1)} + \Delta \bar{\rho}^{(k+1)} \\ \bar{\rho}^{(k+1)}(0,x) =& \mu_0(x).\end{aligned}$$

For $t \geqslant 0$, consider $Z_s^{(k+1)}$ the strong solution of the following stochastic differential equation for $s \in [0,t]$

$$dZ_s^{(k+1)} = \sqrt{2}dB_s - u^{(k)}(t-s, Z_s)ds,$$

which exists, is unique and non-explosive since $u^{(k)}$ is smooth, bounded and Lipschitz continuous. Then

$$\bar{\rho}^{(k+1)}(t,x) = \mathbb{E}_x\left(\mu_0(Z_t^{(k+1)})\right).$$

We thus get

$$\|\bar{\rho}_t^{(k+1)}\|_{L^\infty} \leqslant \|\mu_0\|_{L^\infty}.$$

Notice that this is simply a probabilistic way of presenting the use of the maximum principle.

Boundedness of the derivatives of $\bar{\rho}^{(k+1)}$ and $u^{(k)}$. The boundedness of the derivatives of $u^{(k)}$ is a direct consequence of the boundedness of the derivatives of $\bar{\rho}^{(k)}$ thanks to Young's convolution inequality. Then, the proof for $\bar{\rho}^{(k+1)}$ similar to the proof of Lemma 4.2.1, using the boundedness of the derivatives of $u^{(k)}$. To show that the bounds are in fact independent of k, we follow the proof of Lemma 4.2.1, i.e by induction on the order of the derivative, and within each induction step we prove that both the integrated and uniform bounds are independent of k. This comes from the fact that the proof initially only relies on the bounds on $\|\bar{\rho}_t^{(k+1)}\|_{L^\infty}$ and $\|u_t^{(k)}\|_{L^\infty}$ -which, as we have shown, only depend on $\|\mu_0\|_{L^\infty}$- and then for each induction step on the initial condition and on the bounds constructed at the previous step (therefore independent of k). The bounds concerning the derivatives involving time are then obtained thanks to the bounds on the space derivatives using (C.1.1).

□

Proof of Theorem 4.2.1. It is sufficient to prove existence and uniqueness of the solution in $[0,T] \times \mathbb{T}^d$ for all $T \geqslant 0$, since then the solutions on $[0,T_1] \times \mathbb{T}^d$ and $[0,T_2] \times \mathbb{T}^d$, with $T_1 < T_2$, must coincide in $[0, T_1] \times \mathbb{T}^d$, leading to the existence and uniqueness of the global solution in $\mathbb{R}^+ \times \mathbb{T}^d$. Let us consider $T \geqslant 0$

Existence in $[0,T] \times \mathbb{T}^d$ for T small enough: Let us show the existence of the limit solution. We consider here T to be small enough (an explicit bound will be given later) Let $G(t,x) = \sum_{k\in\mathbb{Z}^d} \frac{1}{(4\pi t)^{\frac{d}{2}}} \exp(-\frac{|x+k|^2}{4t})$ be the heat kernel on the d dimensional torus. We have

$$\bar{\rho}^{(k)}(t,x) = G(t,\cdot) * \mu_0(x) - \int_0^t \int_{\mathbb{T}^d} G(t-s,x-y) u^{(k-1)}(s,y) \cdot \nabla_y \bar{\rho}^{(k)}(s,y) dy ds.$$

Let us denote $N_k(t) = \sup_{0\leqslant s\leqslant t} \|\bar{\rho}^{(k+1)}(s,\cdot) - \bar{\rho}^{(k)}(s,\cdot)\|_{L^\infty}$. We have, using $\nabla_y \cdot u^{(k)} = 0$

$$\begin{aligned}\bar{\rho}^{(k+1)}(t,x) - \bar{\rho}^{(k)}(t,x) = &- \int_0^t \int_{\mathbb{T}^d} \nabla_y G(t-s,x-y) \left(\bar{\rho}^{(k+1)}(s,y) - \bar{\rho}^{(k)}(s,y)\right) u^{(k)}(s,y) dy ds \\ &- \int_0^t \int_{\mathbb{T}^d} \nabla_y G(t-s,x-y) \bar{\rho}^{(k)}(s,y) \left(u^{(k)}(s,y) - u^{(k-1)}(s,y)\right) dy ds.\end{aligned}$$

Remark (using the first moment of the chi distribution) that, for some constant $\beta > 0$

$$\int_{\mathbb{T}^d} |\nabla_x G(t,x)| dx \leqslant \beta t^{-\frac{1}{2}}.$$

We thus get

$$\begin{aligned}\|\bar{\rho}^{(k+1)}(t,\cdot) - \bar{\rho}^{(k)}(t,\cdot)\|_{L^\infty} \leqslant &\beta \|K\|_{L^1} \|\mu_0\|_{L^\infty} \int_0^t (t-s)^{-\frac{1}{2}} \|\bar{\rho}^{(k+1)}(s,\cdot) - \bar{\rho}^{(k)}(s,\cdot)\|_{L^\infty} ds \\ &+ \beta \|\mu_0\|_{L^\infty} \int_0^t (t-s)^{-\frac{1}{2}} \|u^{(k)}(s,\cdot) - u^{(k-1)}(s,\cdot)\|_{L^\infty} ds,\end{aligned}$$

and

$$\|u^{(k)}(s,\cdot) - u^{(k-1)}(s,\cdot)\|_{L^\infty} \leqslant \|K\|_{L^1} \|\bar{\rho}^{(k)}(s,\cdot) - \bar{\rho}^{(k-1)}(s,\cdot)\|_{L^\infty}.$$

Therefore

$$N_k(t) \leqslant \beta \|K\|_{L^1} \|\mu_0\|_{L^\infty} \int_0^t (t-s)^{-\frac{1}{2}} N_k(s) ds + \beta \|K\|_{L^1} \|\mu_0\|_{L^\infty} \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds.$$

Denoting $C = \beta \|K\|_{L^1} \|\mu_0\|_{L^\infty}$ we get

$$N_k(t) \leqslant C \int_0^t (t-s)^{-\frac{1}{2}} \left(N_k(s) + N_{k-1}(s)\right) ds. \tag{C.1.4}$$

Since N_k is continuous, there exists $R > 0$ such that for all $t \in [0,T]$ we have $N_k(t) \leqslant R$. We thus have, using this bound in (C.1.4) and assuming $2C\sqrt{T} \leqslant \frac{1}{2}$

$$\begin{aligned}N_k(t) \leqslant &RC \int_0^t (t-s)^{-\frac{1}{2}} ds + C \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds \\ \leqslant &\frac{R}{2} + C \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds.\end{aligned}$$

We use this bound in (C.1.4)

$$\begin{aligned} N_k(t) \leqslant & \frac{R}{2} C \int_0^t (t-s)^{-\frac{1}{2}} ds + C \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds \\ & + C^2 \int_0^t \int_0^s (t-s)^{-\frac{1}{2}} (s-u)^{-\frac{1}{2}} N_{k-1}(u) du ds. \end{aligned}$$

We deal with this last term

$$\begin{aligned} C^2 \int_0^t \int_0^s & (t-s)^{-1/2} (s-u)^{-\frac{1}{2}} N_{k-1}(u) du ds \\ & = C^2 \int_0^t N_{k-1}(u) \int_u^t (t-s)^{-\frac{1}{2}} (s-u)^{-\frac{1}{2}} ds du \\ & = C^2 \pi \int_0^t N_{k-1}(u) du. \end{aligned}$$

Let $\alpha = \sqrt{T}\pi C$ and choose T such that $\alpha \leqslant \frac{1}{2}$ (which in turns also yields the previous condition $2C\sqrt{T} \leqslant \frac{1}{2}$). We have

$$\alpha C \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds - C^2 \pi \int_0^t N_{k-1}(s) du = C \int_0^t N_{k-1}(s) \left(\alpha (t-s)^{-\frac{1}{2}} - \pi C \right) ds,$$

and since $\alpha = \sqrt{T}\pi C \geqslant \sqrt{t-s}\pi C$ for $0 \leqslant s \leqslant t \leqslant T$, we get

$$\alpha C \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds \geqslant C^2 \pi \int_0^t N_{k-1}(s) du,$$

and thus

$$N_k(t) \leqslant \frac{R}{4} + C(1+\alpha) \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds.$$

Iterating this method, we obtain for all $n \in \mathbb{N}$

$$N_k(t) \leqslant 2^{-n} R + C(1 + \alpha + \cdots + \alpha^{n-1}) \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds,$$

and thus

$$N_k(t) \leqslant 2C \int_0^t (t-s)^{-\frac{1}{2}} N_{k-1}(s) ds.$$

We now show that this implies that

$$N_k(t) \leqslant N_0(T) \left(2C\Gamma\left(\frac{1}{2}\right) \right)^k t^{k/2} \Gamma\left(\frac{k+2}{2}\right)^{-1}, \tag{C.1.5}$$

where we denote $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$. We have that for $k = 0$, (C.1.5) is satisfied and, by induction, we have

$$\int_0^t (t-s)^{-\frac{1}{2}} s^{\frac{k}{2}} ds = t^{\frac{k+1}{2}} \int_0^1 (1-u)^{-\frac{1}{2}} u^{\frac{k}{2}} du = t^{\frac{k+1}{2}} \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{k+2}{2}\right)}{\Gamma\left(\frac{k+3}{2}\right)}$$

Using the fact that $\Gamma(k+1) = k!$ and $\Gamma(k+\frac{3}{2}) = k!\Gamma(\frac{1}{2})$, we get that $\sum_{k=0}^\infty N_k(t)$ converges uniformly for $t \in [0,T]$ and the limits

$$\bar{\rho}(t,x) = \lim_{k\to\infty} \bar{\rho}^{(k)}(t,x) \quad \text{and} \quad u(t,x) = \lim_{k\to\infty} u^{(k)}(t,x)$$

exist in $\mathcal{C}([0,T] \times \mathbb{T}^d)$. Now, since for all $l, n \in \mathbb{N}$ and all $\alpha_1, ..., \alpha_n$, $\|\partial_t^l \partial_{\alpha_1,...,\alpha_n} \bar{\rho}^{(k)}\|_{L^\infty}$ and $\|\partial_t^l \partial_{\alpha_1,...,\alpha_n} u^{(k)}\|_{L^\infty}$ are bounded uniformly in k, using Arzela-Ascoli theorem, we have uniform convergence, up to an extraction, of the derivatives. Hence the validity of the limits in $\mathcal{C}^\infty([0,T]\times\mathbb{T}^d)$, i.e there is convergence of the functions along with their derivatives of all order in $[0,T]\times\mathbb{T}^d$. This gives us the fact that the limit $\bar{\rho}$ satisfies (4.2.1).

Uniqueness in $[0,T]\times\mathbb{T}^d$. Suppose $\bar{\rho}^1$ and $\bar{\rho}^2$ are two bounded solutions of (4.2.1) on $[0,T]\times\mathbb{T}^d$. Then

$$\partial_t\left(\bar{\rho}^1 - \bar{\rho}^2\right) - \Delta\left(\bar{\rho}^1 - \bar{\rho}^2\right) = -\left(K * \bar{\rho}^1\right)\cdot\nabla\left(\bar{\rho}^1 - \bar{\rho}^2\right) - \nabla\cdot\left(\left(K*\bar{\rho}^1 - K*\bar{\rho}^2\right)\bar{\rho}^2\right),$$

so that

$$\begin{aligned}\bar{\rho}^1(t,x) - \bar{\rho}^2(t,x) = &-\int_0^t\int_{\mathbb{T}^d} \nabla_y G(t-s,x-y)\cdot\left(K *_y \bar{\rho}^1(s,y)\right)\left(\bar{\rho}^1(s,y)-\bar{\rho}^2(s,y)\right)dyds\\ &-\int_0^t\int_{\mathbb{T}^d} \nabla_y G(x-y,t-s)\cdot\left(K *_y \bar{\rho}^1(s,y) - K *_y \bar{\rho}^2(s,y)\right)\bar{\rho}^2(s,y)dyds.\end{aligned}$$

Let $N(t) := \sup_{0\leqslant s\leqslant t}\|\bar{\rho}^1(s,\cdot) - \bar{\rho}^2(s,\cdot)\|_{L^\infty}$. Recall

$$\|K*\bar{\rho}^1(s,\cdot) - K*\bar{\rho}^2(s,\cdot)\|_{L^\infty} \leqslant \|K\|_{L^1}\|\bar{\rho}^1(s,\cdot)-\bar{\rho}^2(s,\cdot)\|_{L^\infty},$$

which implies, like previously, the existence of a constant C such that

$$N(t) \leqslant C\int_0^t (t-s)^{-1/2}N(s)ds.$$

We choose $L>0$ such that $C\int_0^T s^{-\frac{1}{2}}e^{-Ls}ds \leqslant \frac{1}{2}$, and let $Q(t) = e^{-Lt}N(t)$, which satisfies for all t

$$Q(t) \leqslant C\int_0^t (t-s)^{-1/2}Q(s)e^{-L(t-s)}ds.$$

Let $R>0$ be such that $Q(t)\leqslant R$.

Then

$$Q(t) \leqslant RC\int_0^t (t-s)^{-1/2}e^{-L(t-s)}ds \leqslant \frac{R}{2}.$$

By induction, we get $N(t) = 0$ for $t \in [0, T]$. This concludes the proof of uniqueness.

Existence in $\mathbb{R}^+ \times \mathbb{T}^d$. For T small enough, there exists a solution in $[0, T] \times \mathbb{T}^d$. Notice that T only depends on constants independent of time (it depends on the L^∞ bound of the initial condition, which we have shown propagates). It is therefore possible to construct the (unique) smooth solution on all intervals $[t_0, T + t_0] \times \mathbb{T}^d$. Uniqueness allows us to iteratively construct the (unique) smooth solution on $\mathbb{R}^+ \times \mathbb{T}^d$. This concludes the proof.

□

Appendix D

Appendix of Chapter 5

D.1 Technical results

Lemma D.1.1. *We have the following inequality*

$$\forall N \geqslant 1, \quad \sum_{i=1}^{N} \frac{1}{i^{\alpha-1}} \leqslant \begin{cases} \frac{1}{2-\alpha} N^{2-\alpha} & \text{if } \alpha \in [1,2[, \\ 2\ln N & \text{if } \alpha = 2 \text{ (and } N \geqslant 2\text{)}, \\ 1 + \frac{1}{\alpha-2} & \text{if } \alpha > 2. \end{cases}$$

Proof. Let $f_\alpha : x \to \frac{1}{x^{\alpha-1}}$. For $\alpha \geqslant 1$, f_α is a non increasing function on $]0, +\infty[$, and for all $x \in [i, i+1]$, $f_\alpha(i+1) \leqslant f_\alpha(x) \leqslant f_\alpha(i)$. This implies

$$f_\alpha(i+1) \leqslant \int_i^{i+1} f_\alpha(x)dx \leqslant f_\alpha(i),$$

and thus

$$\sum_{i=1}^{N} f_\alpha(i) \leqslant \begin{cases} \int_0^N \frac{1}{x^{\alpha-1}} dx & \text{if } \alpha \in [1,2[, \\ f_\alpha(1) + \int_1^N \frac{1}{x^{\alpha-1}} dx & \text{if } \alpha \geqslant 2. \end{cases} = \begin{cases} \left[\frac{x^{2-\alpha}}{2-\alpha}\right]_0^N & \text{if } \alpha \in [1,2[, \\ 1 + [\ln x]_1^N & \text{if } \alpha = 2, \\ 1 + \left[\frac{x^{2-\alpha}}{2-\alpha}\right]_1^N & \text{if } \alpha > 2. \end{cases}$$

Hence

$$\sum_{i=1}^{N} \frac{1}{i^{\alpha-1}} \leqslant \begin{cases} \frac{N^{2-\alpha}}{2-\alpha} & \text{if } \alpha \in [1,2[, \\ 1 + \ln N & \text{if } \alpha = 2, \\ 1 + \frac{1}{\alpha-2} - \frac{1}{(\alpha-2)N^{\alpha-2}} & \text{if } \alpha > 2. \end{cases}$$

□

Lemma D.1.2. *Let, for* $\boldsymbol{x} = (x_i)_{i \in \{1,\dots,N\}}$

$$A(\boldsymbol{x}) = \left(\sum_{j \neq i} \frac{1}{x_i - x_j} \right)_{1 \leqslant i \leqslant N}.$$

There is a constant C such that for all $N \geqslant 0$ and for the set of points $\boldsymbol{x} = (x_i)_{i \in \{1,\dots,N\}}$ with $x_i = \frac{i}{N}$, we have $|A(\boldsymbol{x})| \leqslant CN^{3/2}$.

Proof. Throughout this proof, we denote by C the various universal constants appearing for the sake of conciseness. We have, for all i

$$\sum_{j \neq i} \frac{1}{x_i - x_j} = N \sum_{j \neq i} \frac{1}{i-j} = N \sum_{j=1}^{i-1} \frac{1}{|i-j|} - N \sum_{j=i+1}^{N} \frac{1}{|i-j|} = N \left(\sum_{j=1}^{i-1} \frac{1}{j} - \sum_{j=1}^{N-i} \frac{1}{j} \right),$$

and thus

$$\sum_{j \neq i} \frac{1}{x_i - x_j} = \begin{cases} -N \sum_{j=i}^{N-i} \frac{1}{j} & \text{if } i \leqslant \lfloor \frac{N+1}{2} \rfloor \\ N \sum_{j=N-i+1}^{i-1} \frac{1}{j} & \text{if } i \geqslant \lfloor \frac{N+1}{2} \rfloor. \end{cases}$$

We obtain

$$|A(\mathbf{x})| = \left(N^2 \sum_{i=1}^{\lfloor \frac{N+1}{2} \rfloor} \left(\sum_{j=i}^{N-i} \frac{1}{j} \right)^2 + N^2 \sum_{i=1+\lfloor \frac{N+1}{2} \rfloor}^{N} \left(\sum_{j=N-i+1}^{i-1} \frac{1}{j} \right)^2 \right)^{1/2}.$$

The change of variable $\tilde{i} = N + 1 - i$ in this last sum yields

$$|A(\mathbf{x})| = \left(2N^2 \sum_{i=1}^{\lfloor \frac{N+1}{2} \rfloor} \left(\sum_{j=i}^{N-i} \frac{1}{j} \right)^2 \right)^{1/2}.$$

There exists a universal constant C such that for all $n \geqslant 0$

$$\ln(n) - C \leqslant \sum_{i=1}^{n} \frac{1}{i} \leqslant \ln(n) + C.$$

This yields, for $i \geqslant 2$

$$\left(\sum_{j=i}^{N-i} \frac{1}{j} \right)^2 \leqslant \left(\ln(N-i) - \ln(i-1) + 2C \right)^2,$$

and for $i = 1$

$$\left(\sum_{j=i}^{N-i} \frac{1}{j} \right)^2 \leqslant \left(\ln(N-i) + C \right)^2$$

This way

$$|A(\mathbf{x})| \leqslant \left(2N^2 \left(\ln(N-i) + C \right)^2 + 2N^2 \sum_{i=2}^{\lfloor \frac{N+1}{2} \rfloor} \left(2(\ln(N-i) - \ln(i-1))^2 + 8C^2 \right) \right)^{1/2}.$$

Then

$$(\ln(N-i)-\ln(i-1))^2 = \ln\left(\frac{1-\frac{i}{N}}{\frac{i-1}{N}}\right)^2 \leqslant 2\ln\left(1-\frac{i}{N}\right)^2 + 2\ln\left(\frac{i-1}{N}\right)^2,$$

and there is a universal constant, which we also denote by C, such that

$$\frac{1}{N}\sum_{i=2}^{\lfloor\frac{N+1}{2}\rfloor}\ln\left(1-\frac{i}{N}\right)^2 \leqslant C + \int_0^{1/2}\ln(1-x)^2 dx \leqslant C,$$

and

$$\frac{1}{N}\sum_{i=2}^{\lfloor\frac{N+1}{2}\rfloor}\ln\left(\frac{i-1}{N}\right)^2 \leqslant C + \int_0^{1/2}\ln(x)^2 dx \leqslant C.$$

And thus

$$|A(\mathbf{x})| \leqslant \left(CN^3\right)^{1/2},$$

hence the result. □

D.2 Proof of Lemma 5.2.3

Recall we work under Assumptions 5.1 and 5.4.

Proving (5.2.9)**.** For $\mathbf{x} \in \mathbb{R}^N$, we have

$$\mathcal{H}(\mathbf{x}) = \sum_i |x_i|^2 - \frac{1}{2N}\sum_{i\neq j}|x_i - x_j| \geqslant \sum_i |x_i|^2 - \frac{1}{2N}\sum_{i\neq j}\left(2 + \frac{1}{8}|x_i - x_j|^2\right),$$

and thus

$$\begin{aligned}\mathcal{H}(\mathbf{x}) \geqslant& \sum_i |x_i|^2 - \frac{N(N-1)}{N} - \frac{1}{8N}\sum_{i\neq j}|x_i|^2 + |x_j|^2 \\ \geqslant& \sum_i |x_i|^2\left(1 - \frac{(N-1)}{4N}\right) - \frac{N(N-1)}{N} \\ \geqslant& \frac{1}{2}\sum_i |x_i|^2 - N.\end{aligned}$$

Hence the result.

Time evolution of $\mathcal{H}(\mathbf{X}_t)$. We consider $(\mathbf{X}_t)_t \geqslant 0$ a solution of (5.1.1) such that for all $t \geqslant 0$ we have $X_t^1 < ... < X_t^N$. We apply Itô's formula to get, as almost surely $\forall t \geqslant 0,\ \mathbf{X}_t \in \mathcal{O}_N$.

$$d\mathcal{H}(\mathbf{X}_t) = -\sum_i 2U'\left(X_t^i\right)X_t^i dt - 2\sum_i \frac{X_i}{N}\sum_{j\neq i}V'(X_t^i - X_t^j)dt + 2N\sigma_N dt$$

$$+ 2\sqrt{2\sigma_N}\sum_i X_i dB_t^i + \sum_i \frac{U'\left(X_t^i\right)}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|}dt$$
$$+ \sum_i \left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|}\right)\left(\frac{1}{N}\sum_{j\neq i}V'(X_t^i - X_t^j)\right)dt$$
$$- \sqrt{2\sigma_N}\sum_i \left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|}\right)dB_t^i.$$

We have

$$\sum_i X_i \sum_{j\neq i} V'(X_t^i - X_t^j) = \sum_{i>j} V'(X_t^i - X_t^j)(X_t^i - X_t^j) = -\sum_{i>j}\frac{1}{\left|X_t^i - X_t^j\right|^{\alpha-1}}$$

$$\begin{aligned}\sum_i \frac{U'\left(X_t^i\right)}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|} =& \frac{1}{N}\sum_{i>j}\frac{\left(U'\left(X_t^i\right) - U'(X_t^j)\right)\left(X_t^i - X_t^j\right)}{\left|X_t^i - X_t^j\right|}\\ =& \frac{\lambda}{N}\sum_{i>j}\left|X_t^i - X_t^j\right| = \frac{\lambda}{2N}\sum_{i\neq j}\left|X_t^i - X_t^j\right|. \end{aligned} \tag{D.2.1}$$

Hence

$$\begin{aligned} d\mathcal{H}(\mathbf{X}_t) =& -2\lambda\mathcal{H}(\mathbf{X}_t)dt - \frac{\lambda}{2N}\sum_{i\neq j}\left|X_t^i - X_t^j\right| + \frac{2}{N}\sum_{i>j}\frac{1}{\left|X_t^i - X_t^j\right|^{\alpha-1}}dt\\ &+ 2N\sigma_N dt + \sum_i \left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|}\right)\left(\frac{1}{N}\sum_{j\neq i}V'(X_t^i - X_t^j)\right)\\ &+ 2\sqrt{2\sigma_N}\sum_i X_t^i dB_t^i - \sqrt{2\sigma_N}\sum_i \left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|}\right)dB_t^i. \end{aligned}$$

We now use the calculations of Lemma 4.2 of [128] and write

$$\begin{aligned} &\sum_i \left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{|X_t^i - X_t^j|}\right)\left(\frac{1}{N}\sum_{j\neq i}V'(X_t^i - X_t^j)\right)\\ &= -\sum_i \left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{|X_t^i - X_t^j|}\right)\left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{|X_t^i - X_t^j|^{\alpha+1}}\right)\\ &= -\frac{1}{N^2}\sum_i\sum_{j,l\neq i}\frac{X_t^i - X_t^j}{|X_t^i - X_t^j|}\frac{X_t^i - X_t^l}{\left|X_t^i - X_t^l\right|^{\alpha+1}}\\ &= -\frac{1}{N^2}\sum_i\sum_{j\neq i}\frac{1}{|X_t^i - X_t^j|^{\alpha}} - \frac{1}{N^2}\sum_i\sum_{\substack{j,\,l\neq i\\ j\neq l}}\frac{X_t^i - X_t^j}{|X_t^i - X_t^j|}\frac{X_t^i - X_t^l}{\left|X_t^i - X_t^l\right|^{\alpha+1}}, \end{aligned}$$

and

$$\begin{aligned}
\sum_i \sum_{j,l\neq i, j\neq l} \frac{X_t^i - X_t^j}{|X_t^i - X_t^j|} \frac{X_t^i - X_t^l}{|X_t^i - X_t^l|^{\alpha+1}} \\
=& \sum_i \sum_{\substack{j,l\neq i \\ j<l}} \frac{X_t^i - X_t^j}{|X_t^i - X_t^j|} \frac{X_t^i - X_t^l}{|X_t^i - X_t^l|^{\alpha+1}} + \frac{X_t^i - X_t^l}{|X_t^i - X_t^l|} \frac{X_t^i - X_t^j}{|X_t^i - X_t^j|^{\alpha+1}} \\
=& \sum_i \sum_{\substack{j,l\neq i \\ j<l}} \frac{(X_t^i - X_t^j)(X_t^i - X_t^l)}{|X_t^i - X_t^j||X_t^i - X_t^l|} \left(\frac{1}{|X_t^i - X_t^j|^{\alpha}} + \frac{1}{|X_t^i - X_t^l|^{\alpha}} \right) \\
=& \sum_{i<j<l} \frac{(X_t^i - X_t^j)(X_t^i - X_t^l)}{|X_t^i - X_t^j||X_t^i - X_t^l|} \left(\frac{1}{|X_t^i - X_t^j|^{\alpha}} + \frac{1}{|X_t^i - X_t^l|^{\alpha}} \right) \\
&+ \frac{(X_t^j - X_t^i)(X_t^j - X_t^l)}{|X_t^j - X_t^i||X_t^i - X_t^l|} \left(\frac{1}{|X_t^j - X_t^i|^{\alpha}} + \frac{1}{|X_t^j - X_t^l|^{\alpha}} \right) \\
&+ \frac{(X_t^l - X_t^j)(X_t^l - X_t^i)}{|X_t^l - X_t^j||X_t^l - X_t^i|} \left(\frac{1}{|X_t^l - X_t^j|^{\alpha}} + \frac{1}{|X_t^l - X_t^i|^{\alpha}} \right) \\
=& \sum_{i<j<l} \frac{1}{|X_t^i - X_t^j|^{\alpha}} + \frac{1}{|X_t^i - X_t^l|^{\alpha}} - \frac{1}{|X_t^j - X_t^i|^{\alpha}} - \frac{1}{|X_t^j - X_t^l|^{\alpha}} \\
&+ \frac{1}{|X_t^l - X_t^j|^{\alpha}} + \frac{1}{|X_t^l - X_t^i|^{\alpha}} \\
=& 2 \sum_{i<j<l} \frac{1}{|X_t^i - X_t^l|^{\alpha}} \\
=& 2 \sum_{i<j} \frac{j-i-1}{|X_t^i - X_t^j|^{\alpha}}.
\end{aligned}$$

We therefore have

$$\sum_i \left(\frac{1}{N} \sum_{j\neq i} \frac{X_t^i - X_t^j}{|X_t^i - X_t^j|} \right) \left(\frac{1}{N} \sum_{j\neq i} V'(X_t^i - X_t^j) \right) = -\frac{2}{N^2} \sum_{i>j} \frac{i-j}{|X_t^i - X_t^j|^{\alpha}}.$$

We now compute

$$\begin{aligned}
d\left(e^{2\lambda t}\mathcal{H}(\mathbf{X}_t)\right) =& 2\lambda e^{2\lambda t}\mathcal{H}(\mathbf{X}_t)dt + e^{2\lambda t} d\mathcal{H}(\mathbf{X}_t) \\
=& e^{2\lambda t} \left(\frac{2}{N} \sum_{i>j} \frac{1}{|X_t^i - X_t^j|^{\alpha-1}} - \frac{2}{N^2} \sum_{i>j} \frac{i-j}{|X_t^i - X_t^j|^{\alpha}} \right. \\
& \left. - \frac{\lambda}{2N} \sum_{i\neq j} \left| X_t^i - X_t^j \right| + 2N\sigma_N \right) dt
\end{aligned}$$

$$+\sqrt{2\sigma_N}e^{2\lambda t}\sum_i\left(2X_t^i-\frac{1}{N}\sum_{j\neq i}\frac{X_t^i-X_t^j}{|X_t^i-X_t^j|}\right)dB_t^i. \tag{D.2.2}$$

Proving (5.2.10) **and** (5.2.11) **for** $\alpha=1$**.** Let $\alpha=1$. We get from (D.2.2)

$$e^{2\lambda t}\mathbb{E}\mathcal{H}(\mathbf{X}_t)\leqslant\mathbb{E}\mathcal{H}(\mathbf{X}_0)+(N-1+2N\sigma_N)\frac{e^{2\lambda t}-1}{2\lambda}-\mathbb{E}\left(\int_0^t\frac{2e^{2\lambda s}}{N^2}\sum_{j<i}\frac{i-j}{|X_s^i-X_s^j|}ds\right).$$

This first yields (5.2.10), and then (5.2.11) using (5.2.9), for $\alpha=1$.

Proving (5.2.11) **for** $\alpha>1$**.** Let $\alpha>1$. Using Young's inequality, we have, for all $\gamma>0$ and $i>j$

$$\frac{1}{|x|^{\alpha-1}}\leqslant\gamma^{\frac{\alpha}{\alpha-1}}\frac{\alpha-1}{\alpha}\frac{i-j}{|x|^\alpha}+\frac{1}{\alpha\gamma^\alpha(i-j)^{\alpha-1}}.$$

Hence

$$\frac{1}{N}\sum_{i>j}\frac{1}{|X_t^i-X_t^j|^{\alpha-1}}\leqslant\gamma^{\frac{\alpha}{\alpha-1}}\frac{\alpha-1}{\alpha}\frac{1}{N}\sum_{i>j}\frac{i-j}{|X_t^i-X_t^j|^\alpha}+\frac{1}{\alpha\gamma^\alpha}\frac{1}{N}\sum_{i>j}\frac{1}{(i-j)^{\alpha-1}}.$$

We consider $\gamma^{\frac{\alpha}{\alpha-1}}=\frac{1}{N}$, i.e $\gamma^\alpha=\frac{1}{N^{\alpha-1}}$.

$$\frac{1}{N}\sum_{i>j}\frac{1}{|X_t^i-X_t^j|^{\alpha-1}}\leqslant\frac{\alpha-1}{\alpha}\frac{1}{N^2}\sum_{i>j}\frac{i-j}{|X_t^i-X_t^j|^\alpha}+\frac{N^{\alpha-2}}{\alpha}\sum_{i>j}\frac{1}{(i-j)^{\alpha-1}}. \tag{D.2.3}$$

Let us now assume $\alpha\in]1,2[$, using Lemma D.1.1

$$\begin{aligned}\sum_{i>j}\frac{1}{(i-j)^{\alpha-1}}&=\sum_{i=1}^N\sum_{j=1}^{i-1}\frac{1}{(i-j)^{\alpha-1}}=\sum_{i=1}^N\sum_{j=1}^{i-1}\frac{1}{j^{\alpha-1}}=\sum_{j=1}^N\frac{N-j}{j^{\alpha-1}}\\&\leqslant N\sum_{j=1}^N\frac{1}{j^{\alpha-1}}\leqslant\frac{NN^{2-\alpha}}{2-\alpha}.\end{aligned}$$

Hence

$$\frac{1}{N}\sum_{i>j}\frac{1}{|X_t^i-X_t^j|^{\alpha-1}}\leqslant\frac{\alpha-1}{\alpha}\frac{1}{N^2}\sum_{i>j}\frac{i-j}{|X_t^i-X_t^j|^\alpha}+\frac{N}{\alpha(2-\alpha)},$$

and thus

$$\frac{2}{N}\sum_{i>j}\frac{1}{|X_t^i-X_t^j|^{\alpha-1}}-\frac{2}{N^2}\sum_{i>j}\frac{i-j}{|X_t^i-X_t^j|^\alpha}\leqslant-\frac{2}{\alpha}\frac{1}{N^2}\sum_{i>j}\frac{i-j}{|X_t^i-X_t^j|^\alpha}+\frac{2N}{\alpha(2-\alpha)}.$$

Using (D.2.2), we get

$$e^{2\lambda t}\mathbb{E}\mathcal{H}(\mathbf{X}_t) \leqslant \mathbb{E}\mathcal{H}(\mathbf{X}_0) - \frac{2}{\alpha}\mathbb{E}\left(\int_0^t \frac{e^{2\lambda s}}{N^2}\sum_{i>j}\frac{i-j}{|X_s^i - X_s^j|^\alpha}ds\right) + \int_0^t e^{2\lambda s}\left(2N\sigma_N + \frac{2N}{\alpha(2-\alpha)}\right)ds. \tag{D.2.4}$$

This yields

$$\begin{aligned}\frac{2}{\alpha}\mathbb{E}\left(\int_0^t \frac{e^{2\lambda s}}{N^2}\sum_{i>j}\frac{i-j}{|X_s^i - X_s^j|^\alpha}ds\right) &\leqslant \mathbb{E}\mathcal{H}(\mathbf{X}_0) - e^{2\lambda t}\mathbb{E}\mathcal{H}(\mathbf{X}_t) + \int_0^t e^{2\lambda s}\left(2N\sigma_N + \frac{2N}{\alpha(2-\alpha)}\right)ds \\ &\leqslant \mathbb{E}\mathcal{H}(\mathbf{X}_0) + Ne^{2\lambda t} + \left(2N\sigma_N + \frac{2N}{\alpha(2-\alpha)}\right)\left(\frac{e^{2\lambda t}-1}{2\lambda}\right),\end{aligned}$$

where we used (5.2.9) for this last inequality. This yields the desired result for $\alpha \in]1, 2[$.

Let $\alpha = 2$. Instead of the control (D.2.3), we have, by Lemma D.1.1

$$\sum_{i>j}\frac{1}{(i-j)^{\alpha-1}} \leqslant 2N\ln N,$$

which then yields

$$2\mathbb{E}\left(\int_0^t \frac{e^{2\lambda s}}{N^2}\sum_{i>j}\frac{i-j}{\left|X_s^i - X_s^j\right|^\alpha}ds\right) \leqslant \alpha\left(\mathbb{E}\mathcal{H}(\mathbf{X}_0) + Ne^{2\lambda t} + 2N\sigma_N\frac{e^{2\lambda t}-1}{2\lambda}\right) + 4N\ln N\frac{e^{2\lambda t}-1}{2\lambda}.$$

Finally, let $\alpha > 2$. By Lemma D.1.1

$$\sum_{i>j}\frac{1}{(i-j)^{\alpha-1}} \leqslant \left(1 + \frac{1}{\alpha-2}\right)N,$$

which then yields

$$\begin{aligned}2\mathbb{E}\left(\int_0^t \frac{e^{2\lambda s}}{N^2}\sum_{i>j}\frac{i-j}{\left|X_s^i - X_s^j\right|^\alpha}ds\right) \leqslant &\alpha\left(\mathbb{E}\mathcal{H}(\mathbf{X}_0) + Ne^{2\lambda t} + 2N\sigma_N\frac{e^{2\lambda t}-1}{2\lambda}\right) \\ &+ 2\left(1 + \frac{1}{\alpha-2}\right)N^{\alpha-1}\frac{e^{2\lambda t}-1}{2\lambda}.\end{aligned}$$

Proving (5.2.10) **for** $\alpha > 1$**.** Using (D.2.4) for $\alpha > 1$

$$e^{2\lambda t}\mathbb{E}\mathcal{H}(\mathbf{X}_t) \leqslant \mathbb{E}\mathcal{H}(\mathbf{X}_0) + \left(2N\sigma_N + \frac{2C(\alpha, N)}{\alpha}\right)\frac{e^{2\lambda t}-1}{2\lambda},$$

i.e

$$\mathbb{E}\mathcal{H}(\mathbf{X}_t) \leqslant e^{-2\lambda t}\mathbb{E}\mathcal{H}(\mathbf{X}_0) + \frac{1}{\lambda}\left(N\sigma_N + \frac{C(\alpha,N)}{\alpha}\right).$$

We thus obtain the uniform in time bound.

Proving (5.2.12)**.** Using the previous calculations, we have

$$\begin{aligned} d\mathcal{H}(\mathbf{X}_t) \leqslant & -2\lambda\mathcal{H}(\mathbf{X}_t)dt + 2\sum_{i>j}\frac{1}{\left|X_t^i - X_t^j\right|^{\alpha-1}}dt + 2N\sigma_N dt - \frac{2}{N^2}\sum_{i<j}\frac{j-i}{\left|X_t^i - X_t^j\right|^{\alpha}}dt \\ & + 2\sqrt{2\sigma_N}\sum_i X_i dB_t^i - \sqrt{2\sigma_N}\sum_i\left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|}\right)dB_t^i, \end{aligned}$$

as well as

$$\begin{aligned} d\mathcal{H}(\mathbf{X}_t) \leqslant & -2\lambda\mathcal{H}(\mathbf{X}_t)dt + 2N\sigma_N dt + \frac{2C(\alpha,N)}{\alpha}dt - \frac{2}{\alpha N^2}\sum_{i>j}\frac{i-j}{|X_t^i - X_t^j|^{\alpha}}dt \\ & + 2\sqrt{2\sigma_N}\sum_i X_i dB_t^i - \sqrt{2\sigma_N}\sum_i\left(\frac{1}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|}\right)dB_t^i. \end{aligned} \tag{D.2.5}$$

Hence, from (D.2.5), we get

$$\begin{aligned} \frac{2}{\alpha}\mathbb{E}\left(\frac{1}{N^2}\int_0^t\sum_{i>j}\frac{i-j}{|X_s^i - X_s^j|^{\alpha}}ds\right) \leqslant & \mathbb{E}\mathcal{H}(\mathbf{X}_0) - \mathbb{E}\mathcal{H}(\mathbf{X}_t) - \int_0^t 2\lambda\mathbb{E}\mathcal{H}(\mathbf{X}_s)ds \\ & + \left(2N\sigma_N + \frac{2C(\alpha,N)}{\alpha}\right)t \\ \leqslant & \mathbb{E}\mathcal{H}(\mathbf{X}_0) + N \\ & + \left(2\lambda N + 2N\sigma_N + \frac{2C(\alpha,N)}{\alpha}\right)t, \end{aligned}$$

where we used (5.2.9) for this last inequality.

This concludes the proof.

Remark D.2.1. *One of the reasons we choose to focus on a quadratic potential U can be found in* (D.2.1) *: convexity is not sufficient to deal with this term and one should instead use the Lipschitz condition to bound*

$$\sum_i \frac{U'\left(X_t^i\right)}{N}\sum_{j\neq i}\frac{X_t^i - X_t^j}{\left|X_t^i - X_t^j\right|} \leqslant \frac{L_U}{2N}\sum_{i\neq j}\left|X_t^i - X_t^j\right|.$$

This is not surprising as a convex potential U would tend to bring the particles closer together, thus increasing the interactions we are trying to bound.

If one wishes to work in the general convex case, the calculations should then be adapted. First notice that assuming $L_U \leqslant 2\lambda$, where λ would be the convexity coefficient, would be sufficient as we would keep a nonpositive coefficient in front of the sum $\sum_{i\neq j}\left|X_t^i - X_t^j\right|$ in (D.2.2)*. If $L_U > 2\lambda$, the function $\mathcal{H}$ should be however modified to take into account this constant L_U. We do not address this question.*

D.3 Establishing the continuity in time

In this section, we show the continuity in 0, uniform in N, of $t \mapsto \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - X_0^i|^2\right)$ (where we denote $\mathbf{X}_t = (X_t^1, ..., X_t^N)$ the solution of (5.1.1) with initial condition $\mathbf{X}_0 = (X_0^1, ..., X_0^N) \in \mathcal{O}_N$), under some assumptions on $\mathbf{X}_0$.

Lemma D.3.1. *Under Assumptions 5.1 and 5.3, let $\boldsymbol{X}_t = (X_t^1, ..., X_t^N)$ be the solution of* (5.1.1) *with (deterministic) initial condition $\boldsymbol{x}_0 = (x_0^1, ..., x_0^N) \in \mathcal{O}_N$. For $\alpha \in [1, 2[$, there exists a constant C_{cont} (depending only on α and λ) such that for all $t \geqslant 0$ and $N \in \mathbb{N}$*

$$\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - x_0^i|^2\right) \leqslant C_{cont}\left(\frac{|A(\boldsymbol{x}_0)|\mathcal{H}(\boldsymbol{x}_0)}{N^{5/2}} + \frac{(1+\sigma_N)|A(\boldsymbol{x}_0)|}{N^{3/2}} + \frac{\mathcal{H}(\boldsymbol{x}_0)}{N} + 1 + \sigma_N\right) t,$$

where

$$A(\boldsymbol{x}) = \left(-\sum_{j\neq i} V'(x^i - x^j)\right)_{1\leqslant i\leqslant N}. \tag{D.3.1}$$

Proof. Itô's formula yields

$$\sum_i |X_t^i - x_0^i|^2 = -2\int_0^t \sum_i U'(X_s^i)(X_s^i - x_0^i)ds - \frac{2}{N}\int_0^t \sum_i \left(\sum_{j\neq i} V'(X_s^i - X_s^j)\right)(X_s^i - x_0^i)ds$$
$$+ 2N\sigma_N t + 2\sqrt{2\sigma_N}\sum_i \int_0^t (X_s^i - x_0^i)dB_s^i. \tag{D.3.2}$$

We have, using the convexity of A,

$$-\sum_i \left(\sum_{j\neq i} V'(X_s^i - X_s^j)\right)(X_s^i - x_0^i) = A(\mathbf{X}_s)\cdot(\mathbf{X}_s - \mathbf{x}_0) \leqslant |A(\mathbf{x}_0)|\,|\mathbf{X}_s - \mathbf{x}_0|\,,$$

and thus

$$\mathbb{E}\left(-\frac{2}{N}\int_0^t \sum_i \left(\sum_{j\neq i} V'(X_s^i - X_s^j)\right)(X_s^i - x_0^i)ds\right) \leqslant \frac{2|A(\mathbf{x}_0)|}{N}\mathbb{E}\left(\int_0^t \left(\sum_i (X_s^i - x_0^i)^2\right)^{1/2} ds\right).$$

Then

$$\left(\sum_i (X_s^i - x_0^i)^2\right)^{1/2} \leqslant \frac{\sqrt{N}}{2} + \frac{1}{2\sqrt{N}}\sum_i (X_s^i - x_0^i)^2$$

$$\leqslant \frac{\sqrt{N}}{2} + \frac{1}{\sqrt{N}} \left(\sum_i (X_s^i)^2 + \sum_i (x_0^i)^2 \right),$$

and thus, using (5.2.9)

$$\left(\sum_i (X_s^i - x_0^i)^2 \right)^{1/2} \leqslant \frac{\sqrt{N}}{2} + 4\sqrt{N} + \frac{2}{\sqrt{N}} \left(\mathcal{H}(\mathbf{X}_s) + \mathcal{H}(\mathbf{x}_0) \right).$$

This way

$$\mathbb{E}\left(\int_0^t \left(\sum_i (X_s^i - x_0^i)^2 \right)^{1/2} ds \right) \leqslant \frac{9}{2}\sqrt{N}t + \frac{2}{\sqrt{N}}\mathcal{H}(\mathbf{x}_0)t + \frac{2}{\sqrt{N}} \int_0^t \mathbb{E}\left(\mathcal{H}(\mathbf{X}_s) \right) ds.$$

We now use (5.2.10) to get that there exists a universal constant C, depending only on α, such that

$$\mathbb{E}\left(\int_0^t \left(\sum_i (X_s^i - x_0^i)^2 \right)^{1/2} ds \right) \leqslant \left(\frac{9}{2}\sqrt{N} + \frac{4}{\sqrt{N}}\mathcal{H}(\mathbf{x}_0) + \frac{2\sqrt{N}\sigma_N}{\lambda} + \frac{2C\sqrt{N}}{\lambda} \right) t,$$

which finally yields

$$\mathbb{E}\left(-\frac{2}{N} \int_0^t \sum_i \left(\sum_{j\neq i} V'(X_s^i - X_s^j) \right) (X_s^i - x_0^i) ds \right)$$
$$\leqslant \left(\frac{9|A(\mathbf{x}_0)|}{\sqrt{N}} + \frac{8|A(\mathbf{x}_0)|\mathcal{H}(\mathbf{x}_0)}{N^{3/2}} + \frac{4|A(\mathbf{x}_0)|\sigma_N}{\lambda\sqrt{N}} + \frac{2C|A(\mathbf{x}_0)|}{\lambda\sqrt{N}} \right) t.$$

We then have, using again (5.2.9)

$$\begin{aligned}
-2\int_0^t \sum_i U'(X_s^i)(X_s^i - x_0^i) ds &= -2\int_0^t \sum_i U'(X_s^i) X_s^i ds + 2\int_0^t \sum_i U'(X_s^i) x_0^i ds \\
&\leqslant -2\lambda \int_0^t \sum_i (X_s^i)^2 ds + 2L_U \int_0^t \sum_i |X_s^i||x_0^i| ds \\
&\leqslant -\lambda \int_0^t \sum_i (X_s^i)^2 ds + \frac{L_U^2}{\lambda} \int_0^t \sum_i |x_0^i|^2 ds \\
&\leqslant \frac{2L_U^2}{\lambda} (\mathcal{H}(\mathbf{x}_0) + N) t.
\end{aligned}$$

and thus

$$\mathbb{E}\left(-2\int_0^t \sum_i U'(X_s^i)(X_s^i - x_0^i) ds \right) \leqslant \frac{2L_U^2}{\lambda} (\mathcal{H}(\mathbf{x}_0) + N) t$$

Going back to (D.3.2), we obtain

$$\mathbb{E}\left(\frac{1}{N}\sum_i |X_t^i - x_0^i|^2\right) \leqslant \frac{2L_U^2}{\lambda}\left(\frac{\mathcal{H}(\mathbf{x}_0)}{N}+1\right)t + 2\sigma_N t$$
$$+\left(\frac{9|A(\mathbf{x}_0)|}{N^{3/2}} + \frac{8|A(\mathbf{x}_0)|\mathcal{H}(\mathbf{x}_0)}{N^{5/2}} + \frac{4|A(\mathbf{x}_0)|\sigma_N}{\lambda N^{3/2}} + \frac{2C|A(\mathbf{x}_0)|}{\lambda N^{3/2}}\right)t,$$

hence

$$\mathbb{E}\left(\frac{1}{N}\sum_i |X_t^i - x_0^i|^2\right)$$
$$\leqslant \left(\frac{|A(\mathbf{x}_0)|}{N^{3/2}}\left(9 + \frac{4\sigma_N}{\lambda} + \frac{2C}{\lambda}\right) + \frac{2L_U^2}{\lambda}\frac{\mathcal{H}(\mathbf{x}_0)}{N} + \frac{8|A(\mathbf{x}_0)|\mathcal{H}(\mathbf{x}_0)}{N^{5/2}} + \frac{2L_U^2}{\lambda} + 2\sigma_N\right)t.$$

This yields the result. □

Remark D.3.1. *We thus have to assume the initial condition* $\boldsymbol{X}_0 = \left(X_0^i\right)_i$ *is such that* $|A(X_0)| \lesssim N^{3/2}$ *and* $\mathcal{H}(\boldsymbol{X}_0) \lesssim N$*, and still satisfies* $\mathcal{W}_2(\mu_0^N, \bar{\rho}_0) \to 0$ *as* $N \to \infty$*. As shown in Lemma D.1.2, for* $\alpha = 1$ *and* $\rho_0 = \mathbb{1}_{[0,1]}$*, such a choice is possible.*

Appendix E

Appendix of Chapter 6

E.1 Graph estimates

Lemma E.1.1. *Let us fix an integer r, consider an integer m, define the total size of the population $N = mr$, and define independent random variables $\xi_{i,j}^{(N,k,k')}$ for $k, k' \in \{1, \dots, r\}$ and $i, j \in \{1, \dots, m\}$ such that $\xi_{i,j}^{(N,k,k)}$ are of Bernoulli distribution with parameter $q_N^{k,k} = q_N$ satisfying $\frac{1}{q_N} = o\left(\frac{N}{\log N}\right)$, while for $k \neq k'$ $\xi_{i,j}^{(N,k,k')}$ are of Bernoulli distribution with parameter $q_N^{k,k'}$ satisfying $q_N^{k,k'} = o(q_N)$. Then, defining $d_i^{(N,k)} = \sum_{k'=1}^r \sum_{j=1}^m \xi_{i,j}^{(N,k,k')}$ and $\tilde{d}_i^{(N,k)} = \sum_{k'=1}^r \sum_{j=1}^m \xi_{j,i}^{(N,k,k')}$, there exists a constant C such that*

$$\limsup_{m\to\infty} \sup_{k\in\{1,\dots,r\}} \sup_{i\in\{1,\dots m\}} \frac{1}{q_N} \left(\frac{d_i^{(N,k)}}{N} + \frac{\tilde{d}_i^{(N,k)}}{N} \right) \leqslant C,$$

and moreover

$$\sup_{k\in\{1,\dots,r\}} \sup_{i\in\{1,\dots,m\}} \left| \frac{d_i^{(N,k)}}{Nq_N} - \frac{1}{r} \right| \xrightarrow[N\to\infty]{a.s} 0.$$

Proof. We only prove the second estimate, the first one being a consequence of the second claim and the fact that $d_i^{(N,k)}$ and $\tilde{d}_i^{(N,k)}$ have the same distribution. Remarking that the independent random variables $Z_{i,j}^{(N,k,k')} := \frac{1}{Nq_N}\left(\xi_{i,j}^{(N,k,k')} - q_N^{k,k'}\right)$ satisfy $\left|Z_{i,j}^{(N,k,k')}\right| \leqslant \frac{1}{Nq_N}$ and $\mathbb{E}\left[\left|Z_{i,j}^{(N,k,k')}\right|^2\right] \leqslant \frac{q_N^{k,k'}}{N^2 q_N^2}$, Benrstein inequality leads to

$$\mathbb{P}\left(\left|\sum_{k'=1}^r \sum_{j=1}^m Z_{i,j}^{(N,k,k')}\right| > t\right) \leqslant 2\exp\left(-\frac{1}{2}\frac{Nq_N t^2}{\sum_{k'=1}^r \frac{mq_N^{k,k'}}{Nq_N} + \frac{t}{3}}\right).$$

Taking $t = \sqrt{\frac{c\log N}{Nq_N}}$ for some positive constant c and remarking that $q_N^{k,k'} \leqslant q_N$ and $\frac{1}{3}\sqrt{\frac{c\log N}{Nq_N}} \leqslant$

1 for N large enough and we get

$$\mathbb{P}\left(\left|\sum_{k'=1}^{r}\sum_{j=1}^{m} Z_{i,j}^{(N,k,k')}\right| > \sqrt{\frac{c\log N}{Nq_N}}\right) \leqslant 2\exp\left(-\frac{1}{2}\frac{c\log N}{\sum_{k'=1}^{r}\frac{mq_N^{k,k'}}{Nq_N} + \frac{1}{3}\sqrt{\frac{c\log N}{Nq_N}}}\right) \leqslant 2N^{-\frac{c}{4}}.$$

So

$$\mathbb{P}\left(\sup_{k\in\{1,\dots,r\}}\sup_{i\in\{1,\dots,m\}}\left|\frac{d_i^{(N,k)}}{Nq_N} - \sum_{k'=1}^{r}\frac{mq_N^{k,k'}}{Nq_N}\right| > \sqrt{\frac{c\log N}{Nq_N}}\right) \leqslant 2N^{1-\frac{c}{4}},$$

and we conclude by applying Borel-Cantelli Lemma, taking c large enough and noting that $\sum_{k'=1}^{r}\frac{mq_N^{k,k'}}{Nq_N}$ converges to $\frac{1}{r}$ as N goes to infinity (recall that $q_N^{k,k} = q_N$). □

Appendix F

Appendix of Chapter 7

F.1 Technical lemmas

We start with this slight extension of the central limit theorem:

Lemma F.1.1. *Let $(X_i)_{i\geqslant 1}$ be a sequence of i.i.d random variables in $\mathbb{R}$ such that $\mathbb{E}X_1 = 0$ and $\mathbb{E}(|X_1|^2) = 1$. Assume also that $\mathbb{E}(|X_1|^4) < \infty$. Denote $Z_p = \frac{1}{\sqrt{p}}\sum_{i=1}^{p} X_i$.*

Then, we have for $Z \sim \mathcal{N}(0,1)$

$$\mathbb{E}\left(|Z_p|^2\right) = \mathbb{E}\left(|Z|^2\right), \qquad \text{and} \qquad \mathbb{E}\left(|Z_p|^4\right) = \mathbb{E}\left(|Z|^4\right) + O\left(\frac{1}{p}\right).$$

This in particular also yields $\mathbb{E}\left(|Z_p|\right) \xrightarrow[p\to\infty]{} \mathbb{E}\left(|Z|\right)$ and $\mathbb{E}\left(|Z_p|^3\right) \xrightarrow[p\to\infty]{} \mathbb{E}\left(|Z|^3\right)$

Proof. Direct computations give $\mathbb{E}\left(|Z_p|^2\right) = \mathbb{E}\left(|Z|^2\right)$. Likewise, we may explicitly compute $\mathbb{E}|Z_p|^4$. Keeping only the terms with nonzero expectation, we have

$$\begin{aligned}\mathbb{E}|Z_p|^4 =& \frac{1}{p^2}\sum_{i=1}^{p}\mathbb{E}|X_i|^4 + \frac{6}{p^2}\sum_{i>j}\mathbb{E}|X_i|^2\mathbb{E}|X_j|^2 \\ =& \frac{\mathbb{E}|X_1|^4}{p} + \frac{6}{p^2}\frac{p(p-1)}{2}.\end{aligned}$$

Noticing that $\mathbb{E}|Z|^4 = 3$ yields the convergence $\mathbb{E}\left(|Z_p|^4\right) = \mathbb{E}\left(|Z|^4\right) + O\left(\frac{1}{p}\right)$. Thus, we have both

- Z_p converges in law to $Z \sim \mathcal{N}(0,1)$,
- and the convergence of the fourth moment $\mathbb{E}\,|Z_p|^4 \xrightarrow[p\to\infty]{} \mathbb{E}\,|Z|^4$.

By Theorem 6.9 of [171], we have the convergence in L^4 Wasserstein distance of the law of Z_p to a law $\mathcal{N}(0,1)$. This implies the convergence in both L^1 Wasserstein distance and L^3 Wasserstein distance, thus the convergence of the first and third moments of Z_p. □

Lemma F.1.2. *The function $\sigma \in]0,\infty[\mapsto \partial_\kappa^3 f_1(\sigma,0)$ (with f_1 given by (7.3.6)) is continuous and satisfies*

$$\forall \sigma > 0, \quad \partial_\kappa^3 f_1(\sigma,0) < 0.$$

Proof. We have

$$\partial_\kappa^3 f_1(\sigma,0) = \left(\frac{L_W}{\sigma}\right)^3 \left(\frac{\int_{\mathbb{R}} x^4 g}{\int_{\mathbb{R}} g} - 3\left(\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g}\right)^2\right),$$

with g given by (7.3.5). We wish to prove

$$A(\sigma, L_W) := \frac{\int_{\mathbb{R}} x^4 g \int_{\mathbb{R}} g}{\left(\int_{\mathbb{R}} x^2 g\right)^2} < 3.$$

Remark that $A(\sigma, L_W)$ is by definition the kurtosis of a random variable with probability density $\frac{g}{\int g}$. Let us rewrite, for $\alpha > 0$

$$\begin{aligned} U(x) + \frac{L_W}{2}x^2 =& \frac{x^4}{4} + \frac{L_W - 1}{2}x^2 \\ =& \frac{x^4}{4} + \frac{L_W - 1 - \alpha}{2}x^2 + \left(\frac{L_W - 1 - \alpha}{2}\right)^2 + \frac{\alpha}{2}x^2 - \left(\frac{L_W - 1 - \alpha}{2}\right)^2 \\ =& \frac{1}{4}\left(x^2 + L_W - 1 - \alpha\right)^2 + \frac{\alpha}{2}x^2 - \left(\frac{L_W - 1 - \alpha}{2}\right)^2, \end{aligned}$$

such that

$$g(x,\sigma,0) = \exp\left(-\frac{1}{4\sigma}\left(x^2 + L_W - 1 - \alpha\right)^2\right)\exp\left(\frac{1}{\sigma}\left(\frac{L_W - 1 - \alpha}{2}\right)^2\right)e^{-\frac{\alpha}{2\sigma}x^2}.$$

This way we can write

$$\begin{aligned} A(\sigma, L_W) =& \frac{\mathbb{E}\left(\exp\left(-\frac{1}{4\sigma}\left(Y^2 + L_W - 1 - \alpha\right)^2\right)\right)\mathbb{E}\left(Y^4\exp\left(-\frac{1}{4\sigma}\left(Y^2 + L_W - 1 - \alpha\right)^2\right)\right)}{\mathbb{E}\left(Y^2\exp\left(-\frac{1}{4\sigma}\left(Y^2 + L_W - 1 - \alpha\right)^2\right)\right)^2}, \\ & \text{with } \; Y \sim \mathcal{N}\left(0, \frac{\sigma}{\alpha}\right), \\ =& \frac{\mathbb{E}\left(\exp\left(-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2 + L_W - 1 - \alpha\right)^2\right)\right)\mathbb{E}\left(X^4\exp\left(-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2 + L_W - 1 - \alpha\right)^2\right)\right)}{\mathbb{E}\left(X^2\exp\left(-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2 + L_W - 1 - \alpha\right)^2\right)\right)^2}, \\ & \text{with } \; X \sim \mathcal{N}(0,1). \end{aligned}$$

We have

$$-\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}X^2 + L_W - 1 - \alpha\right)^2 = -\frac{1}{4\sigma}\left(\frac{\sigma}{\alpha}\right)^2\left(X^2 + \frac{\alpha\left(L_W - 1 - \alpha\right)}{\sigma}\right)^2.$$

We thus choose $\alpha = \frac{1}{2}\left(L_W - 1 + \sqrt{(L_W - 1)^2 + 4\sigma}\right) > 0$ in order to ensure $\frac{\alpha(L_W - 1 - \alpha)}{\sigma} = -1$.

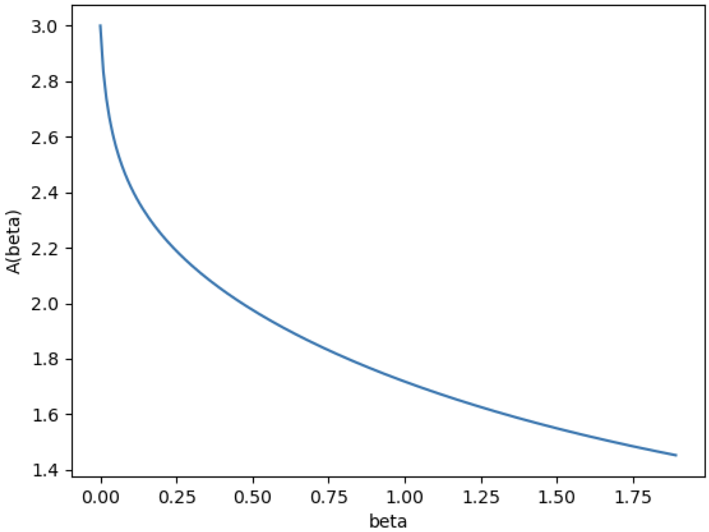

Figure F.1: Numerical simulation of the quantity given in (F.1.1).

Finally, denoting

$$\beta(\sigma, L_W) = \left(\frac{L_W - 1}{\sqrt{\sigma}} + \sqrt{\left(\frac{L_W - 1}{\sqrt{\sigma}}\right)^2 + 4} \right)^{-2} > 0,$$

we have

$$\begin{aligned} A(\sigma, L_W) =& \frac{\mathbb{E}\left(\exp\left(-\beta(\sigma, L_W)\left(X^2-1\right)^2\right)\right)\mathbb{E}\left(X^4\exp\left(-\beta(\sigma, L_W)\left(X^2-1\right)^2\right)\right)}{\mathbb{E}\left(X^2\exp\left(-\beta(\sigma, L_W)\left(X^2-1\right)^2\right)\right)^2}, \quad X \sim \mathcal{N}(0,1) \\ :=& A(\beta(\sigma, L_W)). \end{aligned} \tag{F.1.1}$$

The quantity A given above can be expressed as a function of $\beta(\sigma, L_W) \in]0, \infty[$, that we denote $A(\beta)$. We may then numerically check that $A(\beta) < 3$ for all $\beta > 0$. (see Figure F.1).

Notice that the function $\beta(\sigma, L_W)$, which is in reality a function of the quantity $\frac{L_W-1}{\sqrt{\sigma}}$, is a bijection from $\frac{L_W-1}{\sqrt{\sigma}} \in \mathbb{R}$ to $]0, \infty[$, which satisfies $\beta(\sigma, L_W) \xrightarrow[\frac{L_W-1}{\sqrt{\sigma}} \to \infty]{} 0$. And, for $\beta(\sigma, L_W) = 0$, direct calculations knowing the moments of the Gaussian law yield $A(0) = 3$ and $A'(0) = -24 < 0$ □

Lemma F.1.3. *Consider the function $F_1 : \sigma \mapsto \partial_\kappa f_1(\sigma, 0)$. It is continuously differentiable and, for $\sigma > 0$, satisfies $F_1'(\sigma) < 0$.*

Proof. We have

$$F_1(\sigma) = \frac{L_W}{\sigma} \frac{\int x^2 \exp\left(-\frac{x^4}{4\sigma} + \frac{1-L_W}{2\sigma} x^2\right) dx}{\int \exp\left(-\frac{x^4}{4\sigma} + \frac{1-L_W}{2\sigma} x^2\right) dx}.$$

Consider the change of variable $y = \frac{x}{\sqrt{\sigma}}$. We have

$$\begin{aligned} F_1(\sigma) =& \frac{L_W}{\sigma} \frac{\int \sigma y^2 \exp\left(-\sigma^2 \frac{y^4}{4\sigma} + \frac{1-L_W}{2\sigma}\sigma y^2\right) dy}{\int \exp\left(-\sigma^2 \frac{y^4}{4\sigma} + \frac{1-L_W}{2\sigma}\sigma y^2\right) dy} \\ =& L_W \frac{\int y^2 \exp\left(-\frac{\sigma}{4}y^4 + \frac{1-L_W}{2}y^2\right) dy}{\int \exp\left(-\frac{\sigma}{4}y^4 + \frac{1-L_W}{2}y^2\right) dy}, \end{aligned}$$

which then yields

$$F_1'(\sigma) = \frac{L_W}{4}\left(-\mathbb{E}\left(Y^6\right) + \mathbb{E}\left(Y^2\right)\mathbb{E}\left(Y^4\right)\right),$$

where Y is a random variable with probability density (up to renormalization) $\exp\left(-\frac{\sigma}{4}y^4 + \frac{1-L_W}{2}y^2\right) dy$. By Jensen inequality (in a strictly convex case with a non almost surely constant random variable), we have $F_1'(\sigma) < 0$. □

F.2 Proofs of Lemma 7.3.2

In this section we prove the various results of Lemma 7.3.2.

F.2.1 Moment bounds, critical variance and continuity

To prove the first properties stated in Lemma 7.3.2, we recall the following result, extracted from Theorem 2.1 of [168] and its proof.

Lemma F.2.1. *The equation (with unknown σ)*

$$\int_{\mathbb{R}^+} \left(x^2 - \frac{1}{2L_W}\right) \exp\left((1-L_W)x^2 - \sigma x^4\right) dx = 0, \tag{F.2.1}$$

admits a unique solution, that is the critical value σ_c.

Finally, consider the function

$$\xi(\sigma,\kappa) = \int_{\mathbb{R}} (x-\kappa) \exp\left(-\frac{1}{\sigma}\left(U(x) + \frac{L_W}{2}x^2 - L_W x m\right)\right) dx. \tag{F.2.2}$$

We have the following properties on ξ :

- *The function $\sigma \mapsto \partial_\kappa \xi(\sigma,0)$ is decreasing : for $\sigma < \sigma_c$ we have $\partial_\kappa \xi(\sigma,0) > 0$, for $\sigma > \sigma_c$ we have $\partial_\kappa \xi(\sigma,0) < 0$, and finally $\partial_\kappa \xi(\sigma_c,0) = 0$.*

- *For $\sigma \geqslant \sigma_c$ and $\kappa \geqslant 0$, the function $\kappa \mapsto \xi(\sigma,\kappa)$ is decreasing (which, in fact, ensures uniqueness of the stationary solution for* (NL)*).*

- *For $\sigma < \sigma_c$ and $\kappa \geqslant 0$, the function $\kappa \mapsto \xi(\sigma,\kappa)$ is increasing and then decreasing (which,in fact, ensures the thirdness of the stationary solution for* (NL)*).*

Moment bound : Consider $(\bar{X}_t)_t$ the solution of (NL). Itô's formula yields

$$dU(\bar{X}_t) = A_t dt + dM_t,$$

where M_t is a continuous local martingale and

$$A_t = -U'\left(\bar{X}_t\right)^2 - L_W\left(\bar{X}_t - \mathbb{E}\left(\bar{X}_t\right)\right)U'\left(\bar{X}_t\right) + \sigma U''\left(\bar{X}_t\right).$$

There exists $\lambda > 0$ and $C > 0$, both independent of $\sigma \in [0, \sigma_c]$, such that for all $x \in \mathbb{R}$

$$\sigma U''(x) + 2\lambda U(x) \leqslant \frac{U'(x)^2}{2} + C, \qquad \text{and} \qquad 2L_W^2 x^2 \leqslant \lambda U(x) + C.$$

Consider for instance $\lambda = 1$ and $C = \max\left(2\sigma + (2\sigma)^{3/2} - \frac{1}{2}, \frac{(1+4L_W^2)^2}{4}\right)$ for $U(x) = \frac{x^4}{4} - \frac{x^2}{2}$. Thus

$$A_t \leqslant C - \lambda U\left(\bar{X}_t\right) + \left(L_W^2\left(\bar{X}_t^2 + \mathbb{E}\left(\bar{X}_t\right)^2\right) - \lambda U\left(\bar{X}_t\right)\right),$$

and, using Fatou's lemma to deal with the local martingale, finally we obtain thanks to Gronwall's lemma

$$\mathbb{E}U\left(\bar{X}_t\right) \leqslant e^{-\lambda t}\mathbb{E}U\left(\bar{X}_0\right) + \frac{2C}{\lambda}.$$

Since $\mathbb{E}\left(\bar{X}_t\right)^2 \leqslant \mathbb{E}\left(\bar{X}_t^2\right) \leqslant \frac{1}{2L_W^2}\left(\lambda\mathbb{E}U\left(\bar{X}_t\right) + C\right)$, and considering $\bar{X}_0$ distributed according to a stationary distribution, we may conclude.

Value of $\kappa_2(\mu_{\sigma_c})$: We rewrite Equation (F.2.1) defining σ_c, first by using the symmetry in x to obtain

$$\int_{\mathbb{R}}\left(x^2 - \frac{1}{2L_W}\right)\exp\left((1 - L_W)x^2 - \sigma x^4\right)dx = 0,$$

and then, by a change of variable $x = \frac{y}{\sqrt{2\sigma}}$, this is equivalent to

$$\int_{\mathbb{R}}\left(y^2 - \frac{\sigma}{L_W}\right)\exp\left(-\frac{1}{\sigma}\left(\frac{y^4}{4} - \frac{1 - L_W}{2}y^2\right)\right)dy = 0.$$

Finally, this amounts to having

$$\frac{\sigma}{L_W} = \frac{\int_{\mathbb{R}} y^2 \exp\left(-\frac{1}{\sigma}\left(\frac{y^4}{4} - \frac{1-L_W}{2}y^2\right)\right)dy}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\sigma}\left(\frac{y^4}{4} - \frac{1-L_W}{2}y^2\right)\right)dy},$$

which, since $\kappa_1(\mu_{\sigma_c}) = 0$, yields the value of $\kappa_2(\mu_{\sigma_c})$.

Consider then the function ξ given (F.2.2). We have $\xi(\sigma, \kappa) = (f_1(\sigma, \kappa) - \kappa)\int g(x, \sigma, \kappa)dx$, and thus

$$\partial_\kappa\xi(\sigma, \kappa) = (\partial_\kappa f_1(\sigma, \kappa) - 1)\int g(x, \sigma, \kappa)dx + (f_1(\sigma, \kappa) - \kappa)\int \partial_\kappa g(x, \sigma, \kappa)dx.$$

Considering the equation above for $\kappa = \kappa_1(\mu_{\sigma,*})$, we obtain

$$\partial_\kappa \xi(\sigma, \kappa_1(\mu_{\sigma,*})) = (\partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,*})) - 1) \int g(x, \sigma, \kappa_1(\mu_{\sigma,*}))dx.$$

We may compute the derivatives of f_1 (see (F.2.5) below), and obtain

$$\partial_\kappa \xi(\sigma, \kappa_1(\mu_{\sigma,*})) = (\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*}) - 1) \int g(x, \sigma, \kappa_1(\mu_{\sigma,*}))dx.$$

The values of $\partial_\kappa \xi(\sigma, \kappa)$ for $\kappa = 0$ and $\kappa = \kappa_1(\mu_{\sigma,+})$ depending on σ, as given in Lemma F.2.1, yields the result.

Continuity of the moments : Notice that f_1 given in (7.3.6) is continuous on $(\sigma, \kappa) \in \mathbb{R}^{+,*} \times \mathbb{R}^+$. We start by proving the continuity of $\sigma \mapsto \kappa_1(\mu_{\sigma,+})$, with the convention $\mu_{\sigma,+} = \mu_{\sigma,0}$ for $\sigma > \sigma_c$. In this latter case, the function $\sigma \mapsto \kappa_1(\mu_{\sigma,+})$ is trivially continuous as $\kappa_1(\mu_{\sigma,+}) = 0$.

Let us show the continuity at the point σ_c. Let $(\sigma_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers such that $\sigma_n \xrightarrow[n\to\infty]{} \sigma_c$, and consider the (bounded) sequence $(\kappa_1(\mu_{\sigma_n,+}))_n$. Up to extraction, we can assume $\kappa_1(\mu_{\sigma_n,+}) \xrightarrow[n\to\infty]{} \kappa_1 \geqslant 0$. We have, by definition, $\kappa_1(\mu_{\sigma_n,+}) = f_1(\sigma_n, \kappa_1(\mu_{\sigma_n,+}))$ and, by considering the limit $n \to \infty$, thanks to the continuity of f_1, we obtain $\kappa_1 = f_1(\sigma_c, \kappa_1)$. Uniqueness of the fixed point for σ_c then ensures $\kappa_1 = 0 = \kappa_1(\mu_{\sigma_c,+})$. Hence we obtain the desired continuity.

We now consider $\sigma < \sigma_c$. Assume there exists $\epsilon > 0$ and a sequence $(\sigma_n)_{n\in\mathbb{N}}$ such that $\sigma_n \xrightarrow[n\to\infty]{} \sigma$ and $|\kappa_1(\mu_{\sigma_n,+}) - \kappa_1(\mu_{\sigma,+})| > \epsilon$. Again, up to extraction, we have $\kappa_1(\mu_{\sigma_n,+}) \xrightarrow[n\to\infty]{} \kappa_1 \geqslant 0$ and, since κ_1 is a fixed point that cannot be $\kappa_1(\mu_{\sigma,+})$, we have $\kappa_1 = 0$. Consider the (at least) twice continuously differentiable function ξ given in (F.2.2). On one hand, we have by continuity $\partial_\kappa \xi(\sigma_n, \kappa_1(\mu_{\sigma_n,+})) \xrightarrow[n\to\infty]{} \partial_\kappa \xi(\sigma, 0) > 0$. On the other hand, by the properties of ξ given in Theorem 7.1.2 , we have $\partial_\kappa \xi(\sigma_n, \kappa_1(\mu_{\sigma_n,+})) < 0$. Hence a contradiction, and $\kappa_1(\mu_{\sigma_n,+}) \xrightarrow[n\to\infty]{} \kappa_1(\mu_{\sigma,+})$ for any sequence $\sigma_n \xrightarrow[n\to\infty]{} \sigma$. We thus obtain the continuity.

F.2.2 On the variance of the stationary distribution(s) of (NL)

Let $\sigma_0 > 0$, and let us show that the functions $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ and $\sigma \mapsto \kappa_2(\mu_{\sigma,\pm})$ are Lipschitz continuous. This is useful, as can be seen in Section 7.3.2, in proving that there exists a phase transition for the effective dynamics (Eff).

Throughout this section, the constant C holds no importance and may change from one line to the next.

We start by showing that $\sigma \mapsto \kappa_2(\mu_{\sigma,0})$ is Lipschitz continuous.

Lemma F.2.2. *Let $\sigma_0 > 0$ and $\kappa_1 \in [-C_{\kappa_1}, C_{\kappa_1}]$ (where C_{κ_1} is given in (7.3.8)). The function $\sigma \mapsto f_2(\sigma, \kappa_1)$ is Lipschitz continuous on $[\sigma_0, \infty[$ uniformly in $\kappa_1 \in [-C_{\kappa_1}, C_{\kappa_1}]$.*

Applying this lemma for $\kappa_1 = 0$ yields the desired Lipschitz continuity for $\kappa_2(\mu_{\sigma,0})$.

Proof of Lemma F.2.2. Recall g defined in (7.3.5). and consider $C = \frac{1+2L_W\kappa_1^2}{4}$, a constant such that, for U given by (7.1.9), we ensure $U(x) + \frac{L_W}{2}x^2 - L_W x\kappa_1 + C \geqslant 0$. We have

$$f_2(\sigma, \kappa_1) = \frac{\int_{\mathbb{R}} (x - \kappa_1)^2 g(x, \sigma, \kappa_1) e^{-\frac{C}{\sigma}} dx}{\int_{\mathbb{R}} g(x, \sigma, \kappa_1) e^{-\frac{C}{\sigma}} dx},$$

and thus

$$\begin{aligned}
|\partial_\sigma f_2(\sigma,\kappa_1)| =&\frac{1}{\sigma^2}\left|\frac{\int_{\mathbb{R}}(x-\kappa_1)^2\left(U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C\right)g(x,\sigma,\kappa_1)dx\int_{\mathbb{R}}g(x,\sigma,\kappa_1)dx}{\left(\int_{\mathbb{R}}g(x,\sigma,\kappa_1)dx\right)^2}\right.\\
&\left.-\frac{\int_{\mathbb{R}}(x-\kappa_1)^2g(x,\sigma,\kappa_1)dx\int_{\mathbb{R}}\left(U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C\right)g(x,\sigma,\kappa_1)dx}{\left(\int_{\mathbb{R}}g(x,\sigma,\kappa_1)dx\right)^2}\right|\\
\leqslant&\frac{1}{\sigma^2}\frac{\int_{\mathbb{R}}(x-\kappa_1)^2\left(U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C\right)g(x,\sigma,\kappa_1)dx}{\int_{\mathbb{R}}g(x,\sigma,\kappa_1)dx} \qquad \text{(F.2.3)}\\
&+\frac{1}{\sigma^2}\frac{\int_{\mathbb{R}}(x-\kappa_1)^2g(x,\sigma,\kappa_1)dx}{\int_{\mathbb{R}}g(x,\sigma,\kappa_1)dx}\frac{\int_{\mathbb{R}}\left(U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C\right)g(x,\sigma,\kappa_1)dx}{\int_{\mathbb{R}}g(x,\sigma,\kappa_1)dx}. \qquad \text{(F.2.4)}
\end{aligned}$$

First

$$\left(\int_{\mathbb{R}}g(x,\sigma,\kappa_1)dx\right)^{-1}\leqslant\left(\int_{\mathbb{R}}\exp\left(-\frac{1}{\sigma_0}\left(U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C\right)\right)dx\right)^{-1}.$$

Then, for all $x\in\mathbb{R}$ and all $\alpha\geqslant 0$, we have

$$U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C=\frac{x^4}{4}-\frac{x^2}{2}+\frac{1}{4}+\frac{L_W}{2}|x-\kappa_1|^2\geqslant\alpha x^2-\beta_\alpha,$$

with

$$\beta_\alpha=\frac{(2\alpha+1)^2}{4}-\frac{1}{4}=\alpha^2+\alpha.$$

Thus, for all integers $k\geqslant 0$, we have

$$\begin{aligned}
\int_{\mathbb{R}}x^{2k}\exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C\right)\right)dx\leqslant&e^{\frac{\alpha^2+\alpha}{\sigma}}\int_{\mathbb{R}}x^{2k}\exp\left(-\frac{\alpha x^2}{\sigma}\right)dx\\
=&e^{\frac{\alpha^2+\alpha}{\sigma}}\sqrt{2\pi}\frac{(2k)!}{2^kk!}\left(\frac{\sigma}{2\alpha}\right)^{k+\frac{1}{2}}.
\end{aligned}$$

Choosing $\alpha=\frac{\sqrt{\sigma}}{2}$, we obtain

$$\int_{\mathbb{R}}x^{2k}\exp\left(-\frac{1}{\sigma}\left(U(x)+\frac{L_W}{2}x^2-L_Wx\kappa_1+C\right)\right)dx\leqslant e^{\frac{1}{4}+\frac{1}{2\sqrt{\sigma}}}\sqrt{2\pi}\frac{(2k)!}{2^kk!}\sigma^{\frac{k}{2}+\frac{1}{4}}.$$

Hence there exists C, independent of σ, such that for $k\leqslant\frac{7}{2}$, $\frac{1}{\sigma^2}\int_{\mathbb{R}}x^{2k}g(\sigma,x)dx\leqslant C$ (which allows us to deal with (F.2.3)) and for $k\leqslant 2$, $\frac{1}{\sigma^{5/4}}\int_{\mathbb{R}}x^{2k}g(\sigma,x)dx\leqslant C$ and for $k\leqslant 1$, $\frac{1}{\sigma^{3/4}}\int_{\mathbb{R}}x^{2k}g(\sigma,x)dx\leqslant C$ (both allow us to deal with (F.2.4)). Thus, for $\sigma\geqslant\sigma_0$, $|\partial_\sigma f_2(\sigma,\kappa_1)|$ is bounded uniformly in κ_1, which yields the result. □

We now show that $\sigma\mapsto\kappa_2(\mu_{\sigma,+})$ is Lipschitz continuous. Let $\sigma_c>\sigma_0>0$. We have already proved, in Lemma F.2.2, that $\sigma\mapsto f_2(\sigma,\kappa_1)$ is Lipschitz continuous uniformly in $\kappa_1\in[-C_{\kappa_1},C_{\kappa_1}]$. However, the difficulty lies in the fact that, for $\kappa_2(\mu_{\sigma,+})$ given by $\kappa_2(\mu_{\sigma,+})=f_2(\sigma,\kappa_1(\mu_{\sigma,+}))$, the mean $\sigma\mapsto\kappa_1(\mu_{\sigma,+})$ is not Lipschitz continuous around σ_c. We will work our way around this fact (and, doing so, also prove it) in the rest of the subsection, but in the meantime this can be numerically observed in Figure 7.6.

Let us compute the various derivatives of f_1 and f_2 given in (7.3.6) and (7.3.7).

$$\begin{aligned}
\partial_\sigma g(x,\sigma,\kappa) =& \frac{\left(U(x) + \frac{L_W}{2}|x-\kappa|^2\right)}{\sigma^2} g(x,\sigma,\kappa),\\
\partial_\kappa g(x,\sigma,\kappa) =& \frac{L_W}{\sigma}(x-\kappa) g(x,\sigma,\kappa),\\
\partial_\sigma f_1(\sigma,\kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\int_{\mathbb{R}} g \int_{\mathbb{R}} x\partial_\sigma g - \int_{\mathbb{R}} \partial_\sigma g \int_{\mathbb{R}} xg\right)\\
=& \frac{1}{\sigma^2}\left(\frac{\int_{\mathbb{R}} x\left(U(x)+\frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}} \left(U(x)+\frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g}\right),\\
\partial_\kappa f_1(\sigma,\kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\int_{\mathbb{R}} g \int_{\mathbb{R}} x\partial_\kappa g - \int_{\mathbb{R}} \partial_\kappa g \int_{\mathbb{R}} xg\right)\\
=& \frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}} x(x-\kappa) g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}} (x-\kappa) g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g}\right)\\
=& \frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}} x^2 g}{\int_{\mathbb{R}} g} - \left(\frac{\int_{\mathbb{R}} xg}{\int_{\mathbb{R}} g}\right)^2\right),\\
\partial_\sigma f_2(\sigma,\kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\int_{\mathbb{R}} g \int_{\mathbb{R}} (x-\kappa)^2\partial_\sigma g - \int_{\mathbb{R}} \partial_\sigma g \int_{\mathbb{R}} (x-\kappa)^2 g\right)\\
=& \frac{1}{\sigma^2}\left(\frac{\int_{\mathbb{R}} (x-\kappa)^2\left(U(x)+\frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}} \left(U(x)+\frac{L_W}{2}|x-\kappa|^2\right) g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} (x-\kappa)^2 g}{\int_{\mathbb{R}} g}\right),\\
\partial_\kappa f_2(\sigma,\kappa) =& \frac{1}{\left(\int_{\mathbb{R}} g\right)^2}\left(\left(2\int_{\mathbb{R}} (\kappa - x) g + \int_{\mathbb{R}} (x-\kappa)^2\partial_\kappa g\right)\int_{\mathbb{R}} g - \int_{\mathbb{R}} \partial_\kappa g \int_{\mathbb{R}} (x-\kappa)^2 g\right)\\
=& \frac{2\int_{\mathbb{R}} (\kappa-x) g}{\int_{\mathbb{R}} g} + \frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}} (x-\kappa)^3 g}{\int_{\mathbb{R}} g} - \frac{\int_{\mathbb{R}} (x-\kappa)^2 g}{\int_{\mathbb{R}} g}\frac{\int_{\mathbb{R}} (x-\kappa) g}{\int_{\mathbb{R}} g}\right).
\end{aligned}$$

In particular, notice that

$$\partial_\kappa f_1(\sigma,\kappa_1(\mu_{\sigma,*})) = \frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*}) \quad \text{and} \quad \partial_\kappa f_2(\sigma,\kappa_1(\mu_{\sigma,*})) = \frac{L_W}{\sigma}\frac{\int_{\mathbb{R}} (x-\kappa_1(\mu_{\sigma,*}))^3 g}{\int_{\mathbb{R}} g}. \tag{F.2.5}$$

We have

$$\begin{aligned}
\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}) =& \partial_\sigma f_1(\sigma,\kappa_1(\mu_{\sigma,*})) + \partial_\kappa f_1(\sigma,\kappa_1(\mu_{\sigma,*}))\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}),\\
\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*}) =& \partial_\sigma f_2(\sigma,\kappa_1(\mu_{\sigma,*})) + \partial_\kappa f_2(\sigma,\kappa_1(\mu_{\sigma,*}))\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}).
\end{aligned}$$

Thus

$$\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*}) = \frac{\partial_\sigma f_1(\sigma,\kappa_1(\mu_{\sigma,*}))}{1-\partial_\kappa f_1(\sigma,\kappa_1(\mu_{\sigma,*}))} = \frac{\partial_\sigma f_1(\sigma,\kappa_1(\mu_{\sigma,*}))}{1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})}. \tag{F.2.6}$$

By the results on the critical variance in Lemma 7.3.2, $\kappa_1(\mu_{\sigma,*})$ is continuously differentiable on

$]\sigma_0, \sigma_c[$. Likewise

$$\begin{aligned}\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,*})\left(1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})\right)\\ =\left(1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})\right)\partial_\sigma f_2(\sigma,\kappa_1(\mu_{\sigma,*}))+\partial_\kappa f_2(\sigma,\kappa_1(\mu_{\sigma,*}))\partial_\sigma f_1(\sigma,\kappa_1(\mu_{\sigma,*})).\end{aligned} \tag{F.2.7}$$

The fact that $1-\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,*})$ goes to 0 as $\sigma \to \sigma_c$ is what prevents us from giving an upper bound on $\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,*})$. By the result on the critical variance in Lemma 7.3.2 and by continuity, there may be a problem in the limit $\sigma \to \sigma_c^-$, but we wan already say that $\sigma \mapsto \kappa_2(\mu_{\sigma,+})$ is Lipschitz continuous on any interval of the form $[\sigma_1, \sigma_2]$ with $0 < \sigma_1 < \sigma_2 < \sigma_c$. The following lemma gives a more precise speed of convergence of the mean to 0 around the critical parameter.

Lemma F.2.3. *There exists $C > 0$ such that*

$$\frac{\kappa_1(\mu_{\sigma,+})}{\sqrt{\sigma_c-\sigma}} \xrightarrow[\sigma\to\sigma_c^-]{} C.$$

Proof. We restrict the study to $\sigma \in [\sigma_0, \sigma_c[$ for some arbitrary $\sigma_0 > 0$. We compute

$$\begin{aligned}\partial_\kappa f_1(\sigma,\kappa) =&\frac{L_W}{\sigma}\left(\frac{\int_{\mathbb{R}}x^2g}{\int_{\mathbb{R}}g}-\left(\frac{\int_{\mathbb{R}}xg}{\int_{\mathbb{R}}g}\right)^2\right),\\ \partial_\kappa^2 f_1(\sigma,\kappa) =&\left(\frac{L_W}{\sigma}\right)^2\left(\frac{\int_{\mathbb{R}}x^3g}{\int_{\mathbb{R}}g}-3\frac{\int_{\mathbb{R}}x^2g}{\int_{\mathbb{R}}g}\frac{\int_{\mathbb{R}}xg}{\int_{\mathbb{R}}g}+2\left(\frac{\int_{\mathbb{R}}xg}{\int_{\mathbb{R}}g}\right)^3\right),\\ \partial_\kappa^3 f_1(\sigma,\kappa) =&\left(\frac{L_W}{\sigma}\right)^3\left(\frac{\int_{\mathbb{R}}x^4g}{\int_{\mathbb{R}}g}-4\frac{\int_{\mathbb{R}}x^3g}{\int_{\mathbb{R}}g}\frac{\int_{\mathbb{R}}xg}{\int_{\mathbb{R}}g}+12\frac{\int_{\mathbb{R}}x^2g}{\int_{\mathbb{R}}g}\left(\frac{\int_{\mathbb{R}}xg}{\int_{\mathbb{R}}g}\right)^2-6\left(\frac{\int_{\mathbb{R}}xg}{\int_{\mathbb{R}}g}\right)^4-3\left(\frac{\int_{\mathbb{R}}x^2g}{\int_{\mathbb{R}}g}\right)^2\right).\end{aligned}$$

In particular

$$\begin{aligned}f_1(\sigma,0) =&0,\\ \partial_\kappa f_1(\sigma,0) =&\frac{L_W}{\sigma}\frac{\int_{\mathbb{R}}x^2g}{\int_{\mathbb{R}}g} > 0,\\ \partial_\kappa^2 f_1(\sigma,0) =&0,\end{aligned}$$

and, by Lemma F.1.2

$$\partial_\kappa^3 f_1(\sigma,0) = \left(\frac{L_W}{\sigma}\right)^3\left(\frac{\int_{\mathbb{R}}x^4g}{\int_{\mathbb{R}}g}-3\left(\frac{\int_{\mathbb{R}}x^2g}{\int_{\mathbb{R}}g}\right)^2\right) \xrightarrow{\sigma\to\sigma_c^-} \partial_\kappa^3 f_1(\sigma_c,0) < 0. \tag{F.2.8}$$

Let us start by proving that there exists $C > 0$ such that, in the limit $\sigma \to \sigma_c$ (or equivalently $\kappa_1(\mu_{\sigma,+}) \to 0$), we have

$$\frac{\sigma_c-\sigma}{\kappa_1(\mu_{\sigma,+})^2} < C + o(1). \tag{F.2.9}$$

We compute

$$\partial_\kappa f_1(\sigma,\kappa) = \partial_\kappa f_1(\sigma,0) + \kappa\partial_\kappa^2 f_1(\sigma,0) + \frac{\kappa^2}{2}\partial_\kappa^3 f_1(\sigma,0) + o(\kappa^2).$$

By Lemma F.1.3, there exists $C > 0$ such that

$$\partial_\kappa f_1(\sigma,0) \geqslant \partial_\kappa f_1(\sigma_c,0) + C(\sigma_c - \sigma) = 1 + C(\sigma_c - \sigma).$$

Since $\partial_\kappa f_1(\sigma_c,0) = \partial_\kappa f_1(\sigma_c,\kappa_1(\mu_{\sigma_c})) = \frac{L_W}{\sigma_c}\kappa_2(\mu_{\sigma_c}) = 1$ by (7.3.9)

$$\partial_\kappa f_1(\sigma,\kappa) \geqslant 1 + C(\sigma_c - \sigma) + \frac{\kappa^2}{2}\left(\partial_\kappa^3 f_1(\sigma_c,0) + o_{\sigma\to\sigma_c}(1)\right) + o(\kappa^2),$$

where the $o(\kappa^2)$ is uniform in $\sigma \in [\sigma_0,\sigma_c]$. Considering $\kappa = \kappa_1(\mu_{\sigma,+})$, which goes to 0 as $\sigma \to \sigma_c$, in the equation above yields

$$\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,+}) \geqslant 1 + C(\sigma_c - \sigma) + \frac{\kappa_1(\mu_{\sigma,+})^2}{2}\left(\partial_\kappa^3 f_1(\sigma_c,0) + o_{\sigma\to\sigma_c}(1)\right) + o(\kappa_1(\mu_{\sigma,+})^2).$$

By the results of Lemma 7.3.2 concerning the critical variance, $\frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,+}) \leqslant 1$, which gives

$$0 \geqslant \frac{C(\sigma_c - \sigma)}{\kappa_1(\mu_{\sigma,+})^2} + \frac{\partial_\kappa^3 f_1(\sigma_c,0)}{2} + o_{\sigma\to\sigma_c}(1).$$

This gives (F.2.9) and this in turns allows us to state that $\sigma_c - \sigma = O(\kappa_1(\mu_{\sigma,+})^2)$.

We then have

$$\begin{aligned} f_1(\sigma,\kappa) =& f_1(\sigma,0) + \kappa\partial_\kappa f_1(\sigma,0) + \frac{\kappa^2}{2}\partial_\kappa^2 f_1(\sigma,0) + \frac{\kappa^3}{6}\partial_\kappa^3 f_1(\sigma,0) + o(\kappa^3) \\ =& \kappa\partial_\kappa f_1(\sigma,0) + \frac{\kappa^3}{6}\partial_\kappa^3 f_1(\sigma,0) + o(\kappa^3). \end{aligned}$$

Furthermore, defining F_1 as in Lemma F.1.3, we obtain

$$\begin{aligned} \partial_\kappa f_1(\sigma,0) = F_1(\sigma) =& F_1(\sigma_c) - (\sigma_c - \sigma)F_1'(\sigma_c) + o(\sigma_c - \sigma) \\ =& 1 - (\sigma_c - \sigma)F_1'(\sigma_c) + o(\sigma_c - \sigma), \end{aligned}$$

which then yields

$$f_1(\sigma,\kappa) = \kappa\left(1 - (\sigma_c - \sigma)F_1'(\sigma_c) + o(\sigma_c - \sigma)\right) + \frac{\kappa^3}{6}\left(\partial_\kappa^3 f_1(\sigma_c,0) + o(1)\right) + o(\kappa^3).$$

Since $\kappa_1(\mu_{\sigma,+}) = f_1(\sigma,\kappa_1(\mu_{\sigma,+}))$, we thus get in the limit $\sigma \to \sigma_c^-$

$$\begin{aligned} 0 =& -\kappa_1(\mu_{\sigma,+})(\sigma_c - \sigma)F_1'(\sigma_c) + \frac{\kappa_1(\mu_{\sigma,*})^3}{6}\partial_\kappa^3 f_1(\sigma_c,0) \qquad \text{(F.2.10)} \\ & + o(\kappa_1(\mu_{\sigma,+})^3) + \kappa_1(\mu_{\sigma,+})o(\sigma_c - \sigma) \\ 0 =& -(\sigma_c - \sigma)F_1'(\sigma_c) + \frac{\kappa_1(\mu_{\sigma,+})^2}{6}\partial_\kappa^3 f_1(\sigma_c,0) + o(\kappa_1(\mu_{\sigma,+})^2). \qquad \text{(F.2.11)} \end{aligned}$$

Thus, thanks to (F.2.8) and Lemma F.1.3, there exists $C > 0$ such that

$$\frac{\sigma_c - \sigma}{\kappa_1(\mu_{\sigma,+})^2} \xrightarrow{\sigma \to \sigma_c^-} C,$$

which yields the final result. □

Lemma F.2.4. *Let $\sigma_0 \in]0, \sigma_c[$. Then $\sigma \mapsto \kappa_2(\mu_{\sigma,+})$ is Lispchitz continuous on $[\sigma_0, \sigma_c]$.*

Proof. Let us write for all $(\sigma, \kappa) \in [\sigma_0, \sigma_c] \times [-C_{\kappa_1}, C_{\kappa_1}]$

$$\begin{aligned}\partial_\kappa f_1(\sigma,\kappa) =& \partial_\kappa f_1(\sigma,0) + \kappa \partial_\kappa^2 f_1(\sigma,0) + \frac{\kappa^2}{2}\partial_\kappa^3 f_1(\sigma,0) + o(\kappa^3)\\ =& \left(F_1(\sigma_c) - (\sigma_c-\sigma)F_1'(\sigma_c) + o(\sigma_c-\sigma)\right) + \frac{\kappa^2}{2}\left(\partial_\kappa^3 f_1(\sigma_c,0) + O(\sigma_c-\sigma)\right) + o(\kappa^3),\end{aligned}$$

where we used the notation F_1 from Lemma F.1.3, the fact that $\partial_\kappa^2 f_1(\sigma, 0) = 0$, and where all notation $o(\cdot)$ and $O(\cdot)$ are uniform in $(\sigma, \kappa) \in [\sigma_0, \sigma_c] \times [-C_{\kappa_1}, C_{\kappa_1}]$ by continuity. Since $F_1(\sigma_c) = 1$, we obtain

$$1 - \partial_\kappa f_1(\sigma,\kappa) = (\sigma_c-\sigma)F_1'(\sigma_c) - \frac{\kappa^2}{2}\partial_\kappa^3 f_1(\sigma_c,0) + o(\sigma_c-\sigma) + o(\kappa^3) + \kappa^2 O(\sigma_c-\sigma).$$

Applying this for $\kappa = \kappa_1(\mu_{\sigma,+})$, and using (F.2.11) and Lemma F.2.3, we obtain

$$\begin{aligned}1 - \partial_\kappa f_1(\sigma,\kappa_1(\mu_{\sigma,+})) =& \frac{\kappa_1(\mu_{\sigma,+})^2}{6}\partial_\kappa^3 f_1(\sigma_c,0) - \frac{\kappa_1(\mu_{\sigma,+})^2}{2}\partial_\kappa^3 f_1(\sigma_c,0) + o(\kappa_1(\mu_{\sigma,+})^2)\\ =& -\frac{\kappa_1(\mu_{\sigma,+})^2}{3}\partial_\kappa^3 f_1(\sigma_c,0) + o(\kappa_1(\mu_{\sigma,+})^2).\end{aligned}$$

Hence, using (F.2.6), for all $\sigma \in [\sigma_0, \sigma_c[$

$$\left|\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,+})\right| = \frac{|\partial_\sigma f_1(\sigma,\kappa_1(\mu_{\sigma,+}))|}{|1-\partial_\kappa f_1(\sigma,\kappa_1(\mu_{\sigma,+}))|} \lesssim \frac{C\kappa_1(\mu_{\sigma,+})}{\kappa_1(\mu_{\sigma,+})^2 + o(\kappa_1(\mu_{\sigma,+})^2)} \lesssim \frac{1}{\kappa_1(\mu_{\sigma,+})} + o(\kappa_1(\mu_{\sigma,+})),$$

where we used that $\partial_\sigma f_1(\sigma, 0) = 0$ and that $\partial_\kappa \partial_\sigma f_1$ is bounded over $[\sigma_0, \sigma_c] \times [-C_{\kappa_1}, C_{\kappa_1}]$ to bound the numerator, and Lemma F.1.2 for the denominator.

Besides, since $\partial_\kappa f_2(\sigma, 0) = 0$ (by (F.2.5) and symmetry),

$$|\partial_\kappa f_2(\sigma,\kappa_1(\mu_{\sigma,+})| = |\partial_\kappa f_2(\sigma,\kappa_1(\mu_{\sigma,+})) - \partial_\kappa f_2(\sigma,0)| \lesssim |\kappa_1(\mu_{\sigma,+})| + o(\kappa_1(\mu_{\sigma,+}))$$

and thus we bound for all $\sigma \in [\sigma_0, \sigma_c[$

$$\begin{aligned}\left|\frac{d}{d\sigma}\kappa_2(\mu_{\sigma,+})\right| \leqslant& |\partial_\sigma f_2(\sigma,\kappa_1(\mu_{\sigma,+}))| + \left|\frac{d}{d\sigma}\kappa_1(\mu_{\sigma,+})\right| |\partial_\kappa f_2(\sigma,\kappa_1(\mu_{\sigma,+}))| + o(\kappa_1(\mu_{\sigma,+}))\\ \lesssim& 1 + \frac{|\partial_\kappa f_2(\sigma,\kappa_1(\mu_{\sigma,+})|}{\kappa_1(\mu_{\sigma,+})} + o(1)\\ \lesssim& 1 + o(1).\end{aligned}$$

By continuity (recall that $1 - \partial_\kappa f_1(\sigma, \kappa_1(\mu_{\sigma,+})) = 1 - \frac{L_W}{\sigma}\kappa_2(\mu_{\sigma,+}) > 0$ for $\sigma < \sigma_c$) , this proves that $\sigma \mapsto \kappa_2(\mu_{\sigma,+})$ is Lipschitz on $[\sigma_0, \sigma_c]$. □

Bibliography

[1] D. Albritton, S. Armstrong, J. C. Mourrat, and M. Novack. “Variational methods for the kinetic Fokker-Planck equation”. In: *arXiv e-prints*, arXiv:1902.04037 (Feb. 2019), arXiv:1902.04037. arXiv: `1902.04037 [math.AP]`.

[2] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*. Vol. 118. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2010, pp. xiv+492.

[3] C. Ané et al. *Sur les inégalités de Sobolev logarithmiques*. Vol. 10. Panoramas et Synthèses [Panoramas and Syntheses]. With a preface by Dominique Bakry and Michel Ledoux. Société Mathématique de France, Paris, 2000, pp. xvi+217.

[4] D. Bakry, P. Cattiaux, and A. Guillin. “Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré”. In: *J. Funct. Anal.* 254.3 (2008), pp. 727–759.

[5] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*. Vol. 348. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2014, pp. xx+552.

[6] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul. “Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons”. In: *J. Math. Neurosci.* 2 (2012), Art. 10, 50.

[7] J. Bao and X. Huang. “Approximations of Mckean-Vlasov SDEs with Irregular Coefficients”. In: *arXiv e-prints*, arXiv:1905.08522 (May 2019), arXiv:1905.08522. arXiv: `1905.08522 [math.PR]`.

[8] F. Baudoin. “Bakry-Émery meet Villani”. In: *J. Funct. Anal.* 273.7 (2017), pp. 2275–2291.

[9] F. Baudoin, M. Gordina, and D. P. Herzog. “Gamma calculus beyond Villani and explicit convergence estimates for Langevin dynamics with singular potentials”. In: *Arch. Ration. Mech. Anal.* 241.2 (2021), pp. 765–804.

[10] S. Bayrak, P. Hövel, and V. Vuksanović. “Modeling functional connectivity on empirical and randomized structural brain networks”. In: *Differential Equations and Dynamical Systems* 29.4 (2021), 789–805.

[11] E. Bayraktar, S. Chakraborty, and R. Wu. “Graphon mean field systems”. In: *arXiv e-prints*, arXiv:2003.13180 (Mar. 2020), arXiv:2003.13180. arXiv: `2003.13180 [math.PR]`.

[12] E. Bayraktar and R. Wu. “Graphon particle system: Uniform-in-time concentration bounds”. In: *Stochastic Process. Appl.* 156 (2023), pp. 196–225.

[13] M. Ben-Artzi. “Global solutions of two-dimensional Navier-Stokes and Euler equations”. In: *Arch. Rational Mech. Anal.* 128.4 (1994), pp. 329–358.

[14] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti. "A non-Maxwellian steady distribution for one-dimensional granular media". In: *J. Statist. Phys.* 91.5-6 (1998), pp. 979–990.

[15] N. Berglund and D. Landon. "Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh-Nagumo model". In: *Nonlinearity* 25.8 (2012), pp. 2303–2335.

[16] R. J. Berman and M. Önnheim. "Propagation of chaos for a class of first order models with singular mean field interactions". In: *SIAM J. Math. Anal.* 51.1 (2019), pp. 159–196.

[17] C. Bertucci, M. Debbah, J.-M. Lasry, and P.-L. Lions. "A spectral dominance approach to large random matrices". In: *J. Math. Pures Appl. (9)* 164 (2022), pp. 27–56.

[18] P. Billingsley. *Convergence of probability measures.* Second. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. x+277.

[19] F. Bolley. "Separability and completeness for the Wasserstein distance". In: *Séminaire de probabilités XLI.* Vol. 1934. Lecture Notes in Math. Springer, Berlin, 2008, pp. 371–377.

[20] F. Bolley, I. Gentil, and A. Guillin. "Uniform convergence to equilibrium for granular media". In: *Arch. Ration. Mech. Anal.* 208.2 (2013), pp. 429–445.

[21] F. Bolley, A. Guillin, and F. Malrieu. "Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation". In: *M2AN Math. Model. Numer. Anal.* 44.5 (2010), pp. 867–884.

[22] F. Bolley, A. Guillin, and C. Villani. "Quantitative concentration inequalities for empirical measures on non-compact spaces". In: *Probab. Theory Related Fields* 137.3-4 (2007), pp. 541–593.

[23] L. Boltzmann. Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*, Sitzungsberichte der Akademie der Wissenschaften,* **66** (1872), 275–370, Translation: Further studies on the thermal equilibrium of gas molecules, in *Kinetic Theory 2*, 88–174, Ed. S.G. Brush, Pergamon, Oxford (1966).

[24] T. Bonis. "Stein's method for normal approximation in Wasserstein distances with application to the multivariate central limit theorem". In: *Probab. Theory Related Fields* 178.3-4 (2020), pp. 827–860.

[25] M. Bossy, O. Faugeras, and D. Talay. "Clarification and complement to "Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons"". In: *J. Math. Neurosci.* 5 (2015), Art. 19, 23.

[26] M. Bossy and D. Talay. "A stochastic particle method for the McKean-Vlasov and the Burgers equation". In: *Math. Comp.* 66.217 (1997), pp. 157–192.

[27] F. Bouchut and J. Dolbeault. "On long time asymptotics of the Vlasov-Fokker-Planck equation and of the Vlasov-Poisson-Fokker-Planck system with Coulombic and Newtonian potentials". In: *Differential Integral Equations* 8.3 (1995), pp. 487–514.

[28] J. Bourgain and H. Brezis. "On the equation $\operatorname{div} Y = f$ and application to control of phases". In: *J. Amer. Math. Soc.* 16.2 (2003), pp. 393–426.

[29] A. Bovier, S. Marello, and E. Pulvirenti. "Metastability for the dilute Curie-Weiss model with Glauber dynamics". In: *Electron. J. Probab.* 26 (2021), Paper No. 47, 38.

[30] D. Bresch, P.-E. Jabin, and J. Soler. "A new approach to the mean-field limit of Vlasov-Fokker-Planck equations". In: *arXiv e-prints*, arXiv:2203.15747 (Mar. 2022), arXiv:2203.15747. arXiv: 2203.15747 [math.AP].

[31] D. Bresch, P.-E. Jabin, and Z. Wang. "Mean-field limit and quantitative estimates with singular attractive kernels". In: *arXiv e-prints*, arXiv:2011.08022 (Nov. 2020), arXiv:2011.08022. arXiv: `2011.08022 [math.AP]`.

[32] D. Bresch, P.-E. Jabin, and Z. Wang. "Modulated free energy and mean field limit". In: *Séminaire Laurent Schwartz—EDP et applications* (2019), pp. 1–22.

[33] D. Bresch, P.-E. Jabin, and Z. Wang. "On mean-field limits and quantitative estimates with a large class of singular kernels: application to the Patlak-Keller-Segel model". In: *C. R. Math. Acad. Sci. Paris* 357.9 (2019), pp. 708–720.

[34] H. R. Brown, W. Myrvold, and J. Uffink. "Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics". In: *Stud. Hist. Philos. Sci. B Stud. Hist. Philos. Modern Phys.* 40.2 (2009), pp. 174–191.

[35] Y. Cao, J. Lu, and L. Wang. "On explicit L^2-convergence rate estimate for underdamped Langevin dynamics". In: *arXiv e-prints*, arXiv:1908.04746 (Aug. 2019), arXiv:1908.04746. arXiv: `1908.04746 [math.AP]`.

[36] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The master equation and the convergence problem in mean field games.* Vol. 201. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 2019, pp. x+212.

[37] J. A. Carrillo, M. G. Delgadino, and G. A. Pavliotis. "A λ-convexity based proof for the propagation of chaos for weakly interacting stochastic particles". In: *J. Funct. Anal.* 279.10 (2020), pp. 108734, 30.

[38] J. A. Carrillo, R. J. McCann, and C. Villani. "Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates". In: *Rev. Mat. Iberoamericana* 19.3 (2003), pp. 971–1018.

[39] J. A. Carrillo, S. Jin, and Y. Tang. "Random batch particle methods for the homogeneous Landau equation". In: *Commun. Comput. Phys.* 31.4 (2022), pp. 997–1019.

[40] P. Cattiaux, A. Guillin, and F. Malrieu. "Probabilistic approach for granular media equations in the non-uniformly convex case". In: *Probab. Theory Related Fields* 140.1-2 (2008), pp. 19–40.

[41] P. Cattiaux, A. Guillin, P. Monmarché, and C. Zhang. "Entropic multipliers method for Langevin diffusion and weighted log Sobolev inequalities". In: *J. Funct. Anal.* 277.11 (2019), pp. 108288, 24.

[42] T. Cavallazzi. "Quantitative weak propagation of chaos for stable-driven McKean-Vlasov SDEs". In: *arXiv e-prints*, arXiv:2212.01079 (Dec. 2022), arXiv:2212.01079. arXiv: `2212.01079 [math.AP]`.

[43] E. Cépa and D. Lépingle. "Diffusing particles with electrostatic repulsion". In: *Probab. Theory Related Fields* 107.4 (1997), pp. 429–449.

[44] D. Chafaï and J. Lehec. "On Poincaré and logarithmic Sobolev inequalities for a class of singular Gibbs measures". In: *Geometric aspects of functional analysis. Vol. I.* Vol. 2256. Lecture Notes in Math. Springer, Cham, 2020, pp. 219–246.

[45] L.-P. Chaintron and A. Diez. "Propagation of chaos: a review of models, methods and applications. I. Models and methods". In: *Kinet. Relat. Models* 15.6 (2022), pp. 895–1015.

[46] L.-P. Chaintron and A. Diez. "Propagation of chaos: a review of models, methods and applications. II. Applications". In: *Kinet. Relat. Models* 15.6 (2022), pp. 1017–1173.

[47] T. Chan. “The Wigner semi-circle law and eigenvalues of matrix-valued diffusions”. In: *Probab. Theory Related Fields* 93.2 (1992), pp. 249–272.

[48] J.-F. Chassagneux, L. Szpruch, and A. Tse. “Weak quantitative propagation of chaos via differential calculus on the space of measures”. In: *Ann. Appl. Probab.* 32.3 (2022), pp. 1929–1969.

[49] F. Chen, Z. Ren, and S. Wang. “Uniform-in-Time Propagation of Chaos for Mean Field Langevin Dynamics”. In: *arXiv e-prints*, arXiv:2212.03050 (Dec. 2022), arXiv:2212.03050. arXiv: `2212.03050 [math.PR]`.

[50] X. Cheng, N. S. Chatterji, Y. Abbasi-Yadkori, P. L. Bartlett, and M. I. Jordan. *Sharp convergence rates for Langevin dynamics in the nonconvex setting.* May 2018. arXiv: `1805.01648 [stat.ML]`.

[51] X. Cheng, N. S. Chatterji, P. L. Bartlett, and M. I. Jordan. “Underdamped Langevin MCMC: A non-asymptotic analysis”. In: *Proceedings of the 31st Conference On Learning Theory.* Ed. by S. Bubeck, V. Perchet, and P. Rigollet. Vol. 75. Proceedings of Machine Learning Research. PMLR, 2018, pp. 300–323.

[52] A. Chodron de Courcel, M. Rosenzweig, and S. Serfaty. “Sharp uniform-in-time mean-field convergence for singular periodic Riesz flows”. In: *arXiv e-prints*, arXiv:2304.05315 (Apr. 2023), arXiv:2304.05315. arXiv: `2304.05315 [math.AP]`.

[53] F. Collet and R. C. Kraaij. “Dynamical moderate deviations for the Curie-Weiss model”. In: *Stochastic Process. Appl.* 127.9 (2017), pp. 2900–2925.

[54] L. Colombani and P. Le Bris. “Chaos propagation in mean field networks of FitzHugh-Nagumo neurons”. In: *arXiv e-prints*, arXiv:2206.13291 (June 2022), arXiv:2206.13291. arXiv: `2206.13291 [math.PR]`.

[55] F. Coppini, H. Dietert, and G. Giacomin. “A law of large numbers and large deviations for interacting diffusions on Erdős-Rényi graphs”. In: *Stoch. Dyn.* 20.2 (2020), pp. 2050010, 19.

[56] P. Del Moral and J. Tugaut. “Uniform propagation of chaos and creation of chaos for a class of nonlinear diffusions”. In: *Stoch. Anal. Appl.* 37.6 (2019), pp. 909–935.

[57] F. Delarue and A. Tse. “Uniform in time weak propagation of chaos on the torus”. In: *arXiv e-prints*, arXiv:2104.14973 (Apr. 2021), arXiv:2104.14973. arXiv: `2104.14973 [math.AP]`.

[58] S. Delattre, G. Giacomin, and E. Luçon. “A note on dynamical models on random graphs and Fokker-Planck equations”. In: *J. Stat. Phys.* 165.4 (2016), pp. 785–798.

[59] X. Ding and H. Qiao. “Euler-Maruyama approximations for stochastic McKean-Vlasov equations with non-Lipschitz coefficients”. In: *J. Theoret. Probab.* 34.3 (2021), pp. 1408–1425.

[60] H. Djellout, A. Guillin, and L. Wu. “Transportation cost-information inequalities and applications to random dynamical systems and diffusions”. In: *Ann. Probab.* 32.3B (2004), pp. 2702–2732.

[61] R. L. Dobrušin. “Vlasov equations”. In: *Funktsional. Anal. i Prilozhen.* 13.2 (1979), pp. 48–58, 96.

[62] J. Dolbeault, C. Mouhot, and C. Schmeiser. “Hypocoercivity for linear kinetic equations conserving mass”. In: *Trans. Amer. Math. Soc.* 367.6 (2015), pp. 3807–3828.

[63] M. H. Duong and J. Tugaut. “The Vlasov-Fokker-Planck equation in non-convex landscapes: convergence to equilibrium”. In: *Electron. Commun. Probab.* 23 (2018), Paper No. 19, 10.

[64] A. Durmus, A. Eberle, A. Guillin, and R. Zimmer. "An elementary approach to uniform in time propagation of chaos". In: *Proc. Amer. Math. Soc.* 148.12 (2020), pp. 5387–5398.

[65] F. J. Dyson. "A Brownian-motion model for the eigenvalues of a random matrix". In: *J. Mathematical Phys.* 3 (1962), pp. 1191–1198.

[66] A. Eberle. "Reflection couplings and contraction rates for diffusions". In: *Probab. Theory Related Fields* 166.3-4 (2016), pp. 851–886.

[67] A. Eberle. "Reflection couplings and contraction rates for diffusions". In: *Probab. Theory Related Fields* 166.3-4 (2016), pp. 851–886.

[68] A. Eberle, A. Guillin, and R. Zimmer. "Couplings and quantitative contraction rates for Langevin dynamics". In: *Ann. Probab.* 47.4 (2019), pp. 1982–2010.

[69] A. Eberle, A. Guillin, and R. Zimmer. "Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes". In: *Trans. Amer. Math. Soc.* 371.10 (2019), pp. 7135–7173.

[70] A. Eberle, A. Guillin, and R. Zimmer. "Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes". In: *Trans. Amer. Math. Soc.* 371.10 (2019), pp. 7135–7173.

[71] R. S. Ellis. *Entropy, large deviations, and statistical mechanics.* Vol. 271. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, New York, 1985, pp. xiv+364.

[72] R. S. Ellis and C. M. Newman. "The statistics of Curie-Weiss models". In: *J. Statist. Phys.* 19.2 (1978), pp. 149–161.

[73] S. N. Ethier and T. G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Characterization and convergence. John Wiley & Sons, Inc., New York, 1986, pp. x+534.

[74] L. C. Evans. *Partial differential equations.* Second. Vol. 19. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2010, pp. xxii+749.

[75] N. Fournier and A. Guillin. "On the rate of convergence in Wasserstein distance of the empirical measure". In: *Probab. Theory Related Fields* 162.3-4 (2015), pp. 707–738.

[76] N. Fournier, M. Hauray, and S. Mischler. "Propagation of chaos for the 2D viscous vortex model". In: *J. Eur. Math. Soc. (JEMS)* 16.7 (2014), pp. 1423–1466.

[77] S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction.* Cambridge University Press, 2017.

[78] T. Funaki. "A certain class of diffusion processes associated with nonlinear parabolic equations". In: *Z. Wahrsch. Verw. Gebiete* 67.3 (1984), pp. 331–348.

[79] I. Gentil. "Inégalités de Sobolev logarithmique et de Poincaré pour la loi uniforme". In: *Unpublished note, see http://math.univ-lyon1.fr/homes-www/gentil/logsob-segment.pdf* (2004).

[80] F. Golse. "The mean-field limit for the dynamics of large particle systems". In: *Journées "Équations aux Dérivées Partielles".* Univ. Nantes, Nantes, 2003, Exp. No. IX, 47.

[81] F. Golse, S. Jin, and T. Paul. "The random batch method for N-body quantum dynamics". In: *J. Comput. Math.* 39.6 (2021), pp. 897–922.

[82] A. N. Gorban. "Hilbert's sixth problem: the endless road to rigour". In: *Philos. Trans. Roy. Soc. A* 376.2118 (2018), pp. 20170238, 10.

[83] A. Guillin, P. Le Bris, and P. Monmarché. "Convergence rates for the Vlasov-Fokker-Planck equation and uniform in time propagation of chaos in non convex cases". In: *Electron. J. Probab.* 27 (2022), Paper No. 124, 44.

[84] A. Guillin, P. Le Bris, and P. Monmarché. "On systems of particles in singular repulsive interaction in dimension one: log and Riesz gas". en. In: *Journal de l'École polytechnique — Mathématiques* 10 (2023), pp. 867–916.

[85] A. Guillin, P. Le Bris, and P. Monmarché. "Uniform in time propagation of chaos for the 2D vortex model and other singular stochastic systems". In: *arXiv e-prints*, arXiv:2108.08675 (Aug. 2021), arXiv:2108.08675. arXiv: `2108.08675 [math.PR]`.

[86] A. Guillin, W. Liu, L. Wu, and C. Zhang. "The kinetic Fokker-Planck equation with mean field interaction". In: *J. Math. Pures Appl. (9)* 150 (2021), pp. 1–23.

[87] A. Guillin, W. Liu, L. Wu, and C. Zhang. "Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems". In: *The Annals of Applied Probability* 32.3 (2022), pp. 1590 –1614.

[88] A. Guillin and P. Monmarché. "Optimal linear drift for the speed of convergence of an hypoelliptic diffusion". In: *Electron. Commun. Probab.* 21 (2016), Paper No. 74, 14.

[89] A. Guillin and P. Monmarché. "Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes". In: *J. Stat. Phys.* 185.2 (2021), Paper No. 15, 20.

[90] M. Hairer and J. C. Mattingly. "Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations". In: *Ann. Probab.* 36.6 (2008), pp. 2050–2091.

[91] M. Hauray and S. Mischler. "On Kac's chaos and related problems". In: *J. Funct. Anal.* 266.10 (2014), pp. 6055–6157.

[92] M. Hauray and S. Salem. "Propagation of chaos for the Vlasov-Poisson-Fokker-Planck system in 1D". In: *Kinet. Relat. Models* 12.2 (2019), pp. 269–302.

[93] S. Herrmann and J. Tugaut. "Non-uniqueness of stationary measures for self-stabilizing processes". In: *Stochastic Process. Appl.* 120.7 (2010), pp. 1215–1246.

[94] D. P. Herzog and J. C. Mattingly. "Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials". In: *Comm. Pure Appl. Math.* 72.10 (2019), pp. 2231–2255.

[95] A. L. Hodgkin and A. F. Huxley. "A quantitative description of membrane current and its application to conduction and excitation in nerve". In: *The Journal of Physiology* 117.4 (1952), 500–544.

[96] K. Hu, Z. Ren, D. Šiška, and L. u. Szpruch. "Mean-field Langevin dynamics and energy landscape of neural networks". In: *Ann. Inst. Henri Poincaré Probab. Stat.* 57.4 (2021), pp. 2043–2065.

[97] P.-E. Jabin, D. Poyato, and J. Soler. "Mean-field limit of non-exchangeable systems". In: *arXiv e-prints*, arXiv:2112.15406 (Dec. 2021), arXiv:2112.15406. arXiv: `2112.15406 [math.PR]`.

[98] 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. Birkhäuser/Springer, Cham, 2017, pp. 379–402.

[99] P.-E. Jabin and Z. Wang. "Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels". In: *Invent. Math.* 214.1 (2018), pp. 523–591.

[100] S. Jin and L. Li. "On the mean field limit of the random batch method for interacting particle systems". In: *Sci. China Math.* 65.1 (2022), pp. 169–202.

[101] S. Jin, L. Li, and J.-G. Liu. "Convergence of the random batch method for interacting particles with disparate species and weights". In: *SIAM J. Numer. Anal.* 59.2 (2021), pp. 746–768.

[102] S. Jin, L. Li, and J.-G. Liu. "Random batch methods (RBM) for interacting particle systems". In: *J. Comput. Phys.* 400 (2020), pp. 108877, 30.

[103] S. Jin, L. Li, Z. Xu, and Y. Zhao. "A random batch Ewald method for particle systems with Coulomb interactions". In: *SIAM J. Sci. Comput.* 43.4 (2021), B937–B960.

[104] S. Jin, L. Li, X. Ye, and Z. Zhou. "Ergodicity and long-time behavior of the Random Batch Method for interacting particle systems". In: *arXiv e-prints*, arXiv:2202.04952 (Feb. 2022), arXiv:2202.04952. arXiv: `2202.04952 [math.PR]`.

[105] L. Journel and P. Monmarché. "Convergence of a particle approximation for the quasi-stationary distribution of a diffusion process: uniform estimates in a compact soft case". In: *ESAIM Probab. Stat.* 26 (2022), pp. 1–25.

[106] M. Kac. "Foundations of kinetic theory". In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III.* University of California Press, Berkeley-Los Angeles, Calif., 1956, pp. 171–197.

[107] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, 1997, pp. xii+523.

[108] A. Kazeykina, Z. Ren, X. Tan, and J. Yang. "Ergodicity of the underdamped mean-field Langevin dynamics". In: *arXiv e-prints*, arXiv:2007.14660 (July 2020), arXiv:2007.14660. arXiv: `2007.14660 [math.PR]`.

[109] C. Kumar, Neelima, C. Reisinger, and W. Stockinger. "Well-posedness and tamed schemes for McKean-Vlasov equations with common noise". In: *Ann. Appl. Probab.* 32.5 (2022), pp. 3283–3330.

[110] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence.* Springer Berlin Heidelberg, 1984.

[111] D. Lacker. "Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions". In: *arXiv e-prints*, arXiv:2105.02983 (May 2021), arXiv:2105.02983. arXiv: `2105.02983 [math.PR]`.

[112] D. Lacker. "On a strong form of propagation of chaos for McKean-Vlasov equations". In: *Electron. Commun. Probab.* 23 (2018), Paper No. 45, 11.

[113] D. Lacker and L. Le Flem. "Sharp uniform-in-time propagation of chaos". In: *arXiv e-prints*, arXiv:2205.12047 (May 2022), arXiv:2205.12047. arXiv: `2205.12047 [math.PR]`.

[114] C. Le Bris and P.-L. Lions. "Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications". In: *Ann. Mat. Pura Appl. (4)* 183.1 (2004), pp. 97–130.

[115] P. Le Bris and C. Poquet. "A note on uniform in time mean-field limit in graphs". In: *arXiv e-prints*, arXiv:2211.11519 (Nov. 2022), arXiv:2211.11519. arXiv: `2211.11519 [math.PR]`.

[116] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations.* A mathematical perspective. Imperial College Press, London, 2010, pp. xiv+458.

[117] D. S. Lemons and A. Gythiel. "Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) 146, 530-533 (1908)]". In: *American Journal of Physics* 65.11 (Nov. 1997), pp. 1079–1081.

[118] J. R. León and A. Samson. "Hypoelliptic stochastic FitzHugh-Nagumo neuronal model: mixing, up-crossing and estimation of the spike rate". In: *Ann. Appl. Probab.* 28.4 (2018), pp. 2243–2274.

[119] D. A. Levin, M. J. Luczak, and Y. Peres. "Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability". In: *Probab. Theory Related Fields* 146.1-2 (2010), pp. 223–265.

[120] M. Lewin. "Coulomb and Riesz gases: the known and the unknown". In: *J. Math. Phys.* 63.6 (2022), Paper No. 061101, 77.

[121] F. Li. "Dynamics for stochastic Fitzhugh-Nagumo systems with general multiplicative noise on thin domains". In: *Math. Methods Appl. Sci.* 44.6 (2021), pp. 5050–5078.

[122] F. Li and D. Xu. "Regular dynamics for stochastic Fitzhugh-Nagumo systems with additive noise on thin domains". In: *Discrete Contin. Dyn. Syst. Ser. B* 26.7 (2021), pp. 3517–3542.

[123] L. Li, J.-G. Liu, and Y. Tang. "Some random batch particle methods for the Poisson-Nernst-Planck and Poisson-Boltzmann equations". In: *Commun. Comput. Phys.* 32.1 (2022), pp. 41–82.

[124] L. Li, Z. Xu, and Y. Zhao. "A random-batch Monte Carlo method for many-body systems with singular kernels". In: *SIAM J. Sci. Comput.* 42.3 (2020), A1486–A1509.

[125] S. Li, X.-D. Li, and Y.-X. Xie. "On the law of large numbers for the empirical measure process of generalized Dyson Brownian motion". In: *J. Stat. Phys.* 181.4 (2020), pp. 1277–1305.

[126] Y. Li and F. Li. "Limiting dynamics for stochastic FitzHugh-Nagumo equations on large domains". In: *Stoch. Dyn.* 19.5 (2019), pp. 1950037, 25.

[127] T. Lindvall and L. C. G. Rogers. "Coupling of multidimensional diffusions by reflection". In: *Ann. Probab.* 14.3 (1986), pp. 860–872.

[128] Y. Lu and J. C. Mattingly. "Geometric ergodicity of Langevin dynamics with Coulomb interactions". In: *Nonlinearity* 33.2 (2020), pp. 675–699.

[129] E. Luçon. "Quenched asymptotics for interacting diffusions on inhomogeneous random graphs". In: *Stochastic Process. Appl.* 130.11 (2020), pp. 6783–6842.

[130] E. Luçon and C. Poquet. "Periodicity induced by noise and interaction in the kinetic mean-field FitzHugh-Nagumo model". In: *Ann. Appl. Probab.* 31.2 (2021), pp. 561–593.

[131] Y. Lv and W. Wang. "Limit dynamics for the stochastic FitzHugh-Nagumo system". In: *Nonlinear Anal. Real World Appl.* 11.4 (2010), pp. 3091–3105.

[132] F. Malrieu. "Convergence to equilibrium for granular media equations and their Euler schemes". In: *Ann. Appl. Probab.* 13.2 (2003), pp. 540–560.

[133] F. Malrieu. "Logarithmic Sobolev inequalities for some nonlinear PDE's". In: *Stochastic Process. Appl.* 95.1 (2001), pp. 109–132.

[134] C. Marchioro and M. Pulvirenti. *Mathematical theory of incompressible nonviscous fluids.* Vol. 96. Applied Mathematical Sciences. Springer-Verlag, New York, 1994, pp. xii+283.

[135] H. P. McKean Jr. "A class of Markov processes associated with nonlinear parabolic equations". In: *Proc. Nat. Acad. Sci. U.S.A.* 56 (1966), pp. 1907–1911.

[136] 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)*. Air Force Office Sci. Res., Arlington, Va., 1967, pp. 41–57.

[137] 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)*. Vol. 1627. Lecture Notes in Math. Springer, Berlin, 1996, pp. 42–95.

[138] G. Miermont. *Théorèmes limites et processus de Poisson.* Notes de cours, `http://perso.ens-lyon.fr/gregory.miermont/thlim.pdf`.

[139] S. Mischler, C. Quiñinao, and J. Touboul. "On a kinetic Fitzhugh-Nagumo model of neuronal network". In: *Comm. Math. Phys.* 342.3 (2016), pp. 1001–1042.

[140] S. Mischler, C. Mouhot, and B. Wennberg. "A new approach to quantitative propagation of chaos for drift, diffusion and jump processes". In: *Probab. Theory Related Fields* 161.1-2 (2015), pp. 1–59.

[141] P. Monmarché. "Generalized Γ calculus and application to interacting particles on a graph". In: *Potential Anal.* 50.3 (2019), pp. 439–466.

[142] P. Monmarché. "Long-time behaviour and propagation of chaos for mean field kinetic particles". In: *Stochastic Process. Appl.* 127.6 (2017), pp. 1721–1737.

[143] Q.-H. Nguyen, M. Rosenzweig, and S. Serfaty. "Mean-field limits of Riesz-type singular flows". In: *Ars Inven. Anal.* (2022), Paper No. 4, 45.

[144] R. I. Oliveira and G. H. Reis. "Interacting diffusions on random graphs with diverging average degrees: hydrodynamics and large deviations". In: *J. Stat. Phys.* 176.5 (2019), pp. 1057–1087.

[145] F. Otto and C. Villani. "Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality". In: *J. Funct. Anal.* 173.2 (2000), pp. 361–400.

[146] G. A. Pavliotis. *Stochastic processes and applications.* Vol. 60. Texts in Applied Mathematics. Diffusion processes, the Fokker-Planck and Langevin equations. Springer, New York, 2014, pp. xiv+339.

[147] G. Peyré, M. Cuturi, et al. "Computational optimal transport: With applications to data science". In: *Foundations and Trends® in Machine Learning* 11.5-6 (2019), pp. 355–607.

[148] N. C. Phuc and M. Torres. "Characterizations of the existence and removable singularities of divergence-measure vector fields". In: *Indiana Univ. Math. J.* 57.4 (2008), pp. 1573–1597.

[149] C. Reisinger and W. Stockinger. "An adaptive Euler-Maruyama scheme for McKean-Vlasov SDEs with super-linear growth and application to the mean-field FitzHugh-Nagumo model". In: *J. Comput. Appl. Math.* 400 (2022), Paper No. 113725, 23.

[150] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Third. Vol. 293. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xiv+602.

[151] L. C. G. Rogers and Z. Shi. "Interacting Brownian particles and the Wigner law". In: *Probab. Theory Related Fields* 95.4 (1993), pp. 555–570.

[152] M. Rosenzweig and S. Serfaty. "Global-in-time mean-field convergence for singular Riesz-type diffusive flows". In: *Ann. Appl. Probab.* 33.2 (2023), pp. 754–798.

[153] S. Salem. "A gradient flow approach of propagation of chaos". In: *Discrete Contin. Dyn. Syst.* 40.10 (2020), pp. 5729–5754.

[154] E. Sandier and S. Serfaty. "From the Ginzburg-Landau model to vortex lattice problems". In: *Comm. Math. Phys.* 313.3 (2012), pp. 635–743.

[155] S. Schochet. "The point-vortex method for periodic weak solutions of the 2-D Euler equations". In: *Comm. Pure Appl. Math.* 49.9 (1996), pp. 911–965.

[156] K. Schuh. "Global contractivity for Langevin dynamics with distribution-dependent forces and uniform in time propagation of chaos". In: *arXiv e-prints*, arXiv:2206.03082 (June 2022), arXiv:2206.03082. arXiv: `2206.03082 [math.PR]`.

[157] I. Sekkat and G. Stoltz. "Removing the mini-batching error in Bayesian inference using Adaptive Langevin dynamics". In: *arXiv e-prints*, arXiv:2105.10347 (May 2021), arXiv:2105.10347. arXiv: `2105.10347 [stat.ML]`.

[158] S. Serfaty. *Coulomb gases and Ginzburg-Landau vortices*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2015, pp. viii+157.

[159] S. Serfaty. "Mean field limit for Coulomb-type flows". In: *Duke Math. J.* 169.15 (2020). With an appendix by Mitia Duerinckx and Serfaty, pp. 2887–2935.

[160] S. Serfaty. "Systems of points with Coulomb interactions". In: *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures.* World Sci. Publ., Hackensack, NJ, 2018, pp. 935–977.

[161] T. Shardlow. "Modified equations for stochastic differential equations". In: *BIT* 46.1 (2006), pp. 111–125.

[162] A.-S. Sznitman. "Topics in propagation of chaos". In: *École d'Été de Probabilités de Saint-Flour XIX—1989.* Vol. 1464. Lecture Notes in Math. Springer, Berlin, 1991, pp. 165–251.

[163] D. Talay. "Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme". In: vol. 8. 2. Inhomogeneous random systems (Cergy-Pontoise, 2001). 2002, pp. 163–198.

[164] D. Tatchim Bemmo, M. Siewe Siewe, and C. Tchawoua. "Combined effects of correlated bounded noises and weak periodic signal input in the modified FitzHugh-Nagumo neural model". In: *Commun. Nonlinear Sci. Numer. Simul.* 18.5 (2013), pp. 1275–1287.

[165] M. Thieullen. "Deterministic and stochastic FitzHugh-Nagumo systems". In: *Stochastic biomathematical models.* Vol. 2058. Lecture Notes in Math. Springer, Heidelberg, 2013, pp. 175–186.

[166] A. Tse. "Higher order regularity of nonlinear Fokker-Planck PDEs with respect to the measure component". In: *J. Math. Pures Appl. (9)* 150 (2021), pp. 134–180.

[167] H. C. Tuckwell, R. Rodriguez, and F. Y. M. Wan. "Determination of firing times for the stochastic FitzHugh-Nagumo neuronal model". In: *Neural Computation* 15.1 (2003), 143–159.

[168] J. Tugaut. "Phase transitions of McKean-Vlasov processes in double-wells landscape". In: *Stochastics* 86.2 (2014), pp. 257–284.

[169] K. Uda. "Ergodicity and spike rate for stochastic FitzHugh-Nagumo neural model with periodic forcing". In: *Chaos Solitons Fractals* 123 (2019), pp. 383–399.

[170] C. Villani. "Hypocoercivity". In: *Mem. Amer. Math. Soc.* 202.950 (2009), pp. iv+141.

[171] C. Villani. *Optimal transport*. Vol. 338. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Old and new. Springer-Verlag, Berlin, 2009, pp. xxii+973.

[172] A. A. Vlasov. "The vibrational properties of an electron gas". In: *Soviet Physics Uspekhi* 10.6 (1968), pp. 721–733.

[173] D. J. Watts and S. H. Strogatz. "Collective dynamics of 'small-world' networks". In: *Nature* 393.6684 (1998), pp. 440–442.

[174] M. Welling and Y. W. Teh. "Bayesian Learning via Stochastic Gradient Langevin Dynamics". In: *Proceedings of the 28th International Conference on International Conference on Machine Learning*. ICML'11. Bellevue, Washington, USA: Omnipress, 2011, 681–688.

[175] E. P. Wigner. "Characteristic vectors of bordered matrices with infinite dimensions". In: *Ann. of Math. (2)* 62 (1955), pp. 548–564.

[176] D. Wynter. "Quantitative propagation of chaos for the mixed-sign viscous vortex model on the torus". In: *Kinet. Relat. Models* 16.2 (2023), pp. 294–310.

[177] X. Ye and Z. Zhou. "Error Analysis of Time-Discrete Random Batch Method for Interacting Particle Systems and Associated Mean-Field Limits". In: *arXiv e-prints*, arXiv:2206.02166 (June 2022), arXiv:2206.02166. arXiv: `2206.02166 [math.PR]`.

[178] K. C. Zygalakis. "On the existence and the applications of modified equations for stochastic differential equations". In: *SIAM J. Sci. Comput.* 33.1 (2011), pp. 102–130.

-Hey, Napoleon, that sounds like the end.
-Wait a minute. *I'm* the leader. *I'll* say when it's the end.
...
It's the end.

The Aristocats (1970) directed by Wolfgang Reitherman.

Systems of particles in (singular) interaction : long-time behavior and propagation of chaos

Abstract

This thesis is devoted to the study of certain systems of N particles in mean-field interaction, and of a specific phenomenon : in such systems, when $N \to \infty$, two given particles become "more and more" independent. This property is named *propagation of chaos*, and our aim is to prove it in several settings. We focus on three main cases: the kinetic one (i.e. with degenerate noise), the one with singular interactions, and the one with incomplete interactions. In each case, we seek to obtain quantitative and uniform in time results. We start by setting up a coupling method to prove the long time convergence of the *Vlasov-Fokker-Planck* equation describing the limit of particles in Lipschitz interactions and confined by a non-convex potential. The coupling method is then adapted to prove the propagation of chaos property for this system, as well as for the *FitzHugh-Nagumo* model describing neurons in the brain interacting in a mean field way. We then focus on a few particle systems in *Riesz* interactions. The first one is the *2D vortex* model, for which we prove uniform in time propagation of chaos using entropic methods. We then study a one-dimensional singular system, motivated by the *Dyson Brownian motion* derived from the study of random matrices, for which we prove this same phenomenon by a new coupling. Finally, we show the uniform in time mean field limit for a system of particles *interacting according to a graph*, random or not, before turning our attention to a method of numerical simulation of interacting particles. In particular, we study the *Random Batch Method*, and its effect on the phase transition that may exist for the nonlinear limit of the particle system. To this end, we look successively at the *Curie-Weiss* model and the double-well model for the overdamped Langevin equation.

Keywords: probability theory, stochastic calculus, coupling methods, propagation of chaos, logarithmic sobolev inequality.

Résumé

Cette thèse est consacrée à l'étude de certains systèmes de N particules en interaction en champ moyen, et d'un phénomène particulier : dans de tels systèmes, lorsque $N \to \infty$, deux particules données deviennent "de plus en plus" indépendantes. Cette propriété est appelée *propagation du chaos*, et notre objectif est de la prouver dans plusieurs contextes. Nous nous concentrons sur trois cas principaux : le cas cinétique (c'est-à-dire avec un bruit dégénéré), celui des interactions singulières et celui des interactions incomplètes. À chaque fois, nous cherchons à obtenir des résultats quantitatifs et uniformes en temps. Nous commençons ainsi par mettre en place une méthode de couplage afin de prouver la convergence en temps long de l'équation de *Vlasov-Fokker-Planck* décrivant la limite de particules en interaction lipschitzienne et confinées via un potentiel non convexe. Cette méthode est ensuite adaptée pour prouver la propriété de propagation du chaos pour ce système, ainsi que pour le modèle de *FitzHugh-Nagumo* décrivant des neurones dans le cerveau interagissant en champ moyen. Nous nous intéressons ensuite à quelques systèmes de particules en interaction de type *Riesz*. Le premier est le modèle de *vortex 2D*, pour lequel nous prouvons la propagation du chaos uniforme en temps en utilisant des méthodes entropiques. Nous étudions ensuite un système singulier en dimension 1, motivé par le *mouvement brownien de Dyson* provenant de l'étude de matrices aléatoires, pour lequel nous prouvons ce même phénomène par un nouveau couplage. Enfin, nous montrons la limite de champ moyen uniforme en temps pour un système de particules *interagissant selon un graphe*, aléatoire ou non, avant de nous intéresser à une méthode de simulation numérique de particules en interaction. En particulier, nous étudions la *Random Batch Method*, et son effet sur la transition de phase qui peut exister pour la limite non linéaire du système de particules. Pour cela, nous regardons successivement le modèle de *Curie-Weiss* et le modèle double-puits pour l'équation de Langevin sur-amortie.

Mots clés : probabilités, calcul stochastique, méthodes de couplage, propagation du chaos, inégalité de sobolev logarithmique.

Laboratoire Jacques-Louis Lions

Sorbonne Université – Campus Pierre et Marie Curie – 4 place Jussieu – 75005 Paris – France