



HAL
open science

Theoretical and numerical study of a few stochastic models of statistical physics

Max Fathi

► **To cite this version:**

Max Fathi. Theoretical and numerical study of a few stochastic models of statistical physics. General Mathematics [math.GM]. Université Pierre et Marie Curie - Paris VI, 2014. English. NNT : 2014PA066349 . tel-01096886

HAL Id: tel-01096886

<https://theses.hal.science/tel-01096886>

Submitted on 18 Dec 2014

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

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



École Doctorale de Science Mathématiques de Paris Centre

THÈSE DE DOCTORAT

Discipline : Mathématiques

présentée par

Max FATHI

Étude théorique et numérique de quelques modèles stochastiques en physique statistique

dirigée par Cédric VILLANI

Soutenue le 3 décembre 2014 devant le jury composé de :

M. Thierry BODINEAU	École Polytechnique	examineur
M. Ivan GENTIL	Université Lyon 1	examineur
M. Michel LEDOUX	Université Toulouse 3	rapporteur
M. Tony LELIÈVRE	École des Ponts	examineur
M. Cédric VILLANI	Institut Henri Poincaré	directeur
M. Lorenzo ZAMBOTTI	Université Paris 6	examineur

Après avis de Michel LEDOUX (Université Toulouse 3) et S.R. Srinivasa VARADHAN (New York University).

Laboratoire de Probabilités et Modèles
aléatoires
4, place Jussieu
75 005 Paris

UPMC
Ecole Doctorale de Sciences
Mathématiques de Paris Centre
4 place Jussieu
75252 Paris Cedex 05
Boite courrier 290

*Quand on attaque l'empiriste,
l'empiriste contre-attaque.*
Pierre-Yves C., *Méditations à Cologne*

*The world of research has gone berserk
Too much paperwork*
Bob Dylan, Nettie Moore, *Modern Times*

Nan mais ça suffit là, avec les exergues débiles.
Catherine K., *Principia Impressoricae*

Remerciements

Tout d'abord, j'aimerais remercier Cédric Villani, pour m'avoir encadré pendant cette thèse. Il m'a proposé un sujet passionnant, et m'a beaucoup appris, tant en probabilités qu'en analyse et en géométrie. Par ses conseils avisés, il a été un guide fantastique pour ma découverte de la recherche.

I would like to thank S.R.S. Varadhan, for having supervised my stay at the Courant Institute in 2010-11, and for having accepted to referee this thesis. His profound understanding of hydrodynamic limits inspired several of the results included here. Un grand merci aussi à Michel Ledoux, pour avoir accepté de rapporter cette thèse.

Je souhaite ensuite remercier Thierry Bodineau, pour ses explications sur les modèles non-gradients qui ont beaucoup inspiré le chapitre 6 de cette thèse, et pour avoir accepté de participer au jury. Je remercie aussi Tony Lelièvre, pour l'organisation du CEMRACS 2013, pour sa participation au projet SMARTMC, à l'origine des chapitres 8 et 9 de cette thèse, et pour avoir accepté de participer au jury. Je remercie aussi bien entendu Ivan Gentil et Lorenzo Zambotti, pour m'avoir fait l'honneur de participer au jury.

J'aimerais rendre hommage à mes coauteurs, avec qui une grande partie de ce travail a été réalisé. Merci à Manh Hong Duong, Noufel Frikha, Ahmed-Amine Hommann, Georg Menz et Gabriel Stoltz. Ca a été un véritable plaisir de travailler avec vous, et j'espère qu'il y aura encore beaucoup d'autres occasions de travailler ensemble.

Mon travail a pu bénéficier de nombreuses discussions mathématiques. Merci en particulier à Emmanuel Boissard, Nathael Gozlan, Emanuel Indrei, Christian Leonard, Jan Maas, Bastien Mallein, Daniel Marahrens, Felix Otto, Mark Peletier et Marielle Simon pour leurs explications, leurs suggestions et leurs conseils.

Merci à tout le LPMA pour m'avoir accueilli pendant cette thèse. En particulier, j'aimerais remercier Julien Berestycki pour m'avoir aidé à naviguer les méandres administratifs de l'UPMC. Merci aussi à Florence Deschamps, Maria Pochot, Valérie Juvé, Jacques Portès et Philippe Macé pour leur disponibilité et leur efficacité.

J'ai eu le plaisir de cotoyer les thésards du LPMA, pour le GTT, les déjeuners ou les bières de début, milieu et fin d'année. Merci à mes co-bureaux : Yvain, Loic, Pierre-Antoine, Vincent, Cécile, Franck, Miraine. Les voisins : Alexandre G., Xan, Alexandre B., Florent, Guillaume C., Bastien, Nelo, Julien, Minmin, Olga. Les collègues de Paris 7 : Aser, Guillaume B., Vu Lan, Maud, Lorick, Marc-Antoine. Et aussi à tous les anciens : Clément, Eric, Christophe, Noufel, Oriane, Pablo, Antoine, Matthieu, Nikos, Stavros, Joachim et Xinxin.

Merci aux anciens de la République Populaire du C6¹ et aux affiliés, pour les nombreuses discussions à caractère (trop souvent) mathématique (ou pas). Silvain, pour son hospitalité, le service de déménagement et la référence indispensable [Rid]. Catherine, pour

1. A ma connaissance, aucun manuscrit de thèse n'a été mangé pendant la rédaction, conformément à l'article 4

les gâteaux², le whisky et le soutien moral. Bastien, pour les idées procrastinatrices. Yasmine, pour les débats sur la qualité des cookies. Vincent P., pour les vidéos sur Youtube. Marie A., pour les ragôts sur les célébrités du monde académique. David et Fathi, pour les parties de Magic. Julia, pour les blagues de latinistes. Pierre G., pour les parties de coinche et le pinot. Daphné, pour les leçons de diplomatie. Rémi V., pour ne pas avoir fait subir un sort abominable à Min. Irène et Gabriel, pour les excellents repas. Simon, pour les dégustations de bière. Camille, pour les dégustations de chocolat. Jean-François, pour les cours de physique. Aurélie et Pierre C., pour les conseils de lectures. Pierre-Yves, pour les best-of étudiantins. Tania, pour les discussions sur l'esthétique de la couleur rouge. Merci aussi à Jack, Martin, Julie, Julien, Delphine, Sham, Benoît, Lucie, Arnaud R. (× 2), Muy-Cheng, Nicolas, Titus, Marie D., Vincent B., Guillaume C., Stéphane, Manon, Poc, Olivier, Jérôme, Samy, Rafa, Kuba et Ruben.

J'aimerais aussi saluer mes partenaires de coinche : Marion, Pierre R., Guillaume M., Amael, William, Rémi S., Paul. Les anciens du debating : Cyrille, Claude, Noémie, Samuel, Arthur, Armand, Reda, Edouard, Charlotte, Dorothée, Claire. Les camarades de prépa : Arnaud S., Bruno, Amaury, Haojun. Les lyonnais : Alexis, Renaud, Mikael, Hannelore, Cécile.

Enfin, j'aimerais remercier mes parents, David, Lola et toute ma famille, qui m'ont soutenu non seulement ces trois dernières années, mais toutes les autres aussi.

2. En accord avec les règles de déontologie de l'UPMC, aucun crumble poire-banane-chocolat n'a été maltraité pendant la rédaction de cette thèse

Résumé

Résumé

Dans cette thèse, nous nous intéressons essentiellement à trois sujets : les inégalités fonctionnelles à contenu probabiliste, les limites hydrodynamiques pour les systèmes de spins continus en interaction et la discrétisation des équations différentielles stochastiques.

Ce document, outre l'introduction, comporte trois parties. La première s'intéresse aux inégalités fonctionnelles, et notamment aux inégalités de Sobolev logarithmiques, pour les mesures canoniques, ainsi qu'aux limites hydrodynamiques pour les systèmes des spins continus. La convergence vers la limite hydrodynamique pour plusieurs variantes du modèle de Ginzburg–Landau équipé de la dynamique de Kawasaki y est obtenue, avec notamment des bornes quantitatives en le nombre de spins. On y étudie également la convergence de l'entropie microscopique vers l'entropie hydrodynamique.

La deuxième partie étudie les liens entre flots gradients dans les espaces de mesures de probabilités et grandes déviations pour les suites de lois de solutions d'équations différentielles stochastiques. On y obtient l'équivalence entre le principe de grandes déviations et la Gamma-convergence d'une suite de fonctionnelles apparaissant dans la formulation en flots gradients du flot de marginales des lois des solutions des équations différentielles stochastiques. Comme application de ce principe, on obtient les grandes déviations par rapport à la limite hydrodynamique pour deux variantes du modèle de Ginzburg–Landau.

La troisième partie concerne la discrétisation des équations différentielles stochastiques. On y prouve une inégalité transport-entropie pour la loi du schéma d'Euler explicite. Cette inégalité implique des bornes sur les intervalles de confiance pour l'estimation de quantités de la forme $\mathbb{E}[f(X_T)]$. On y étudie également l'erreur de discrétisation pour l'évaluation des coefficients de transport avec l'algorithme MALA (qui est une combinaison du schéma d'Euler explicite et de l'algorithme de Metropolis–Hastings).

Mots-clefs

systèmes de spins, limites hydrodynamiques, flots gradients, inégalités de Sobolev logarithmiques, équations différentielles stochastiques

Theoretical and numerical study of a few stochastic models of statistical physics

Abstract

In this thesis, we are mainly interested in three topics : functional inequalities and their probabilistic aspects, hydrodynamic limits for interacting continuous spin systems and discretizations of stochastic differential equations.

This document, in addition to a general introduction (written in French), contains three parts. The first part deals with functional inequalities, especially logarithmic Sobolev inequalities, for canonical ensembles, and with hydrodynamic limits for continuous spin systems. We prove convergence to the hydrodynamic limit for several variants of the Ginzburg–Landau model endowed with Kawasaki dynamics, with quantitative bounds in the number of spins. We also study convergence of the microscopic entropy to its hydrodynamic counterpart.

In the second part, we study links between gradient flows in spaces of probability measures and large deviations for sequences of laws of solutions to stochastic differential equations. We show that the large deviations principle is equivalent to the Gamma-convergence of a sequence of functionals that appear in the gradient flow formulation of the flow of marginals of the laws of the diffusion processes. As an application of this principle, we obtain large deviations from the hydrodynamic limit for two variants of the Ginzburg–Landau model.

The third part deals with the discretization of stochastic differential equations. We prove a transport-entropy inequality for the law of the explicit Euler scheme. This inequality implies bounds on the confidence intervals for quantities of the form $\mathbb{E}[f(X_T)]$. We also study the discretization error for the evaluation of transport coefficients with the Metropolis-adjusted Langevin algorithm (which is a combination of the explicit Euler scheme and the Metropolis algorithm).

Keywords

Spin systems, hydrodynamic limits, gradient flows, logarithmic Sobolev inequalities, stochastic differential equations

Table des matières

1	Introduction générale	13
1.1	Transport optimal	15
1.1.1	Transport optimal et distances de Wasserstein	15
1.1.2	Transport optimal dynamique et formule de Benamou-Brenier	17
1.2	Concentration de la mesure et inégalités fonctionnelles	17
1.2.1	Inégalité de Sobolev logarithmique	19
1.2.2	Inégalités transport-entropie	23
1.3	Limites hydrodynamiques	24
1.3.1	Mesures canoniques et inégalité de Sobolev logarithmique	26
1.3.2	Un cadre abstrait pour l'étude des limites hydrodynamiques	30
1.3.3	Application au modèle de Ginzburg-Landau	31
1.4	Flots gradients dans les espaces de mesures	33
1.4.1	Définition	35
1.4.2	Stabilité des flots gradients	37
1.4.3	Flots gradients et grandes déviations	38
1.4.4	Grandes déviations pour deux variantes du modèle de Ginzburg-Landau	40
1.5	Discrétisation d'équations différentielles stochastiques	42
1.5.1	Erreur statistique pour le schéma d'Euler explicite	43
1.5.2	Erreur de discrétisation pour l'algorithme de Metropolis-Hastings	44
1.6	Liste des publications issues de la thèse	46
	Notation	47
I	Functional inequalities and hydrodynamic limits for interacting spin systems	49
2	Hydrodynamic limit for conservative spin systems with super-quadratic, partially inhomogeneous single-site potential.	51
2.1	Introduction	51
2.1.1	Logarithmic Sobolev inequalities	51
2.1.2	Hydrodynamic limits	52
2.2	Framework and Main Results	53
2.2.1	Microscopic and macroscopic scales	53
2.2.2	A new covariance estimate	55
2.2.3	Logarithmic Sobolev inequality	56
2.2.4	Hydrodynamic limit for Kawasaki dynamics	57
2.3	Proof of Proposition 2.6	60

2.4	Proof of Theorem 2.7	67
2.5	Proof of the hydrodynamic limit	69
2.5.1	Proof of Theorem 2.8	69
2.5.2	Proof of Theorem 2.9	74
2.5.3	Auxiliary result : Polynomial bounds on the energy	79
2.5.4	Auxiliary result : Convergence of the free energy	81
2.5.5	Proof of Proposition 2.21	85
2.6	Proof of the local Cramér theorem	90
2.7	Some classical criteria for the LSI	101
3	Local Gibbs behavior in the hydrodynamic limit	103
3.1	Introduction	103
3.2	Notation	104
3.3	Background and Main Results	105
3.3.1	Logarithmic Sobolev inequalities	105
3.3.2	Hydrodynamic limits	106
3.3.3	Kawasaki dynamics	110
3.4	Proof of Theorem 3.7	114
3.5	Application to Kawasaki Dynamics	121
3.5.1	Proof of Theorem 3.19	121
3.5.2	Proof of Theorem 3.20	124
3.5.3	Proof of Proposition 3.22	125
4	The two-scale approach for non-reversible dynamics	129
4.1	Introduction	129
4.2	Framework and main results	130
4.2.1	Abstract setting	130
4.2.2	Application to spin systems	134
4.3	Proof of the abstract results	136
4.3.1	Proof of Theorem 4.1	136
4.3.2	Sketch of proof of Theorem 4.5	145
4.4	Application to weakly asymmetric Kawasaki dynamics	146
5	Modified logarithmic Sobolev inequalities for canonical ensembles	155
5.1	Introduction	155
5.2	Background and Main Results	156
5.3	The iterated two-scale approach for modified logarithmic Sobolev inequalities	158
5.4	An application to Kawasaki dynamics	167
5.5	Appendix : Standard criteria for modified LSI	167
II	Gradient flows and large deviations	171
6	A gradient flow approach to large deviations	173
6.1	Introduction	173
6.2	Framework and Method	175
6.2.1	Gradient flows in $P_2(\mathbb{R}^n)$	175
6.2.2	Relative entropy and large deviations	177
6.2.3	Relative entropy for the law of processes	179
6.2.4	Some questions	180

6.3	Proof of Theorem 6.17	181
6.4	Large deviations for the Ginzburg-Landau model	184
6.4.1	The model	184
6.4.2	Some technical estimates	185
6.4.3	Large deviations for the GL model in a random environment	191
6.4.4	Large deviations for the non-gradient Ginzburg-Landau model	197
III Discretization of stochastic differential equations		205
7	Transport entropy inequalities for stochastic approximation schemes	207
7.1	Introduction	207
7.1.1	Euler-like Scheme of a Diffusion Process	208
7.1.2	Transport-Entropy inequalities	210
7.2	Main Results	212
7.3	Euler Scheme : Proof of the Main Results	215
7.3.1	Proof of Theorem 7.6	215
7.3.2	Proof of Theorem 7.11	220
7.4	Proof of Proposition 7.5	223
8	Error analysis of the transport properties of Metropolized schemes	225
8.1	Introduction	225
8.2	Description of the model	226
8.2.1	Brownian dynamics	226
8.2.2	Self-diffusion	227
8.2.3	Numerical estimation of the self-diffusion	228
8.3	A priori error estimates on the self-diffusion	229
8.3.1	Error estimates for the Green-Kubo formula	230
8.3.2	Error bounds on the Einstein formula	231
8.4	Numerical illustration	232
8.4.1	A simple one-dimensional case	232
8.4.2	The more realistic case of solvated ions	232
8.5	Possible work tracks to reduce the error on the estimation of the self-diffusion	237
8.6	Proof of the results	237
8.6.1	Proof of Theorem 8.1	237
8.6.2	Proof of Lemma 8.2	239
8.6.3	Proof of Lemma 8.3	241
8.6.4	Proof of Theorem 8.4	242
8.6.5	Proof of Theorem 8.5	243
9	Improving trajectorial accuracy for the Metropolis-adjusted Langevin algorithm	249
	Bibliographie	257

Chapitre 1

Introduction générale

Au cours de cette thèse, je me suis essentiellement intéressé à trois sujets : les inégalités fonctionnelles à contenu probabiliste, les limites hydrodynamiques pour les systèmes de spins continus en interaction et la discrétisation des équations différentielles stochastiques. Le premier chapitre donne une présentation générale de ces sujets, et des résultats que j'ai obtenus. Le reste de cette thèse contient les énoncés précis, les développements et les preuves de ces différents résultats.

Dans cette introduction, je vais commencer par présenter le problème du transport optimal, afin de définir les distances de Wasserstein sur les espaces de mesures de probabilité. Ces distances ont été un outil essentiel dans l'étude de la plupart des problèmes auxquels je me suis intéressé. Je vais ensuite présenter les inégalités de concentration, et les différentes inégalités fonctionnelles que j'ai étudié (inégalité de Sobolev logarithmique et inégalités de transport-entropie). J'expliquerai comment elles sont reliées entre elles, et quelques unes de leurs conséquences et applications.

Dans la troisième section, je vais présenter les limites hydrodynamiques et la méthode à deux échelles, introduite dans [GOVW09] pour l'étude de ces limites. En particulier, je parlerai des mesures canoniques, qui ont été l'un des objets que j'ai étudié pendant ma thèse, et de la limite hydrodynamique du modèle de Ginzburg-Landau sur lequel on fait agir la dynamique de Kawasaki. Ce modèle donne l'évolution en temps d'un système de spins continus sur un réseau, avec une interaction entre spins voisins.

Dans la quatrième section, je vais parler des flots gradients, qui sont une manière de définir des équations différentielles dans les espaces métriques. Lorsqu'on utilise cette construction sur l'espace des mesures de probabilité équipé de la distance de Wasserstein W_2 , on peut obtenir notamment l'équation de Fokker-Planck, qui donne l'évolution en temps de la loi de la solution d'une équation différentielle stochastique réversible. On verra comment cette formulation est liée à l'étude des grandes déviations, et une méthode de preuve de principes de grandes déviations pour les trajectoires d'équations différentielles stochastiques. Cette méthode sera ensuite appliquée à deux variantes du modèle de Ginzburg-Landau, pour étudier les grandes déviations par rapport à leur limite hydrodynamique.

Enfin, dans la dernière section, je présenterais mes contributions à l'étude de l'erreur pour des discrétisation d'équations différentielles stochastiques. On y verra comment on peut utiliser les inégalités de transport-entropie pour obtenir des bornes sur l'erreur statistique du schéma d'Euler-Maruyama (ou Euler explicite). Je parlerai également de l'étude de l'erreur de discrétisation de l'algorithme MALA, utilisé pour simuler les trajectoires d'équations différentielles stochastiques réversibles.

Les autres chapitres sont les différents articles de recherche que j'ai réalisés. Une grande

partie de ces résultats ont été obtenus en collaboration avec d'autres chercheurs.

- Le chapitre 2 généralise la méthode à deux échelles de Grunewald, Otto, Villani et Westdickenberg (introduite dans [GOVW09]) au cas où le potentiel de la mesure invariante (dite mesure canonique) est donné par une fonction super-quadratique, et contient un terme linéaire inhomogène supplémentaire, qui modélise l'effet d'un potentiel chimique. On y obtient une inégalité de Sobolev logarithmique indépendante de la dimension pour les mesures canoniques, ainsi que la limite hydrodynamique du modèle de Ginzburg Landau dans le cas où on ajoute un potentiel chimique donné par des variables aléatoires indépendantes, identiquement distribuées et bornées. Ces résultats ont été obtenus en collaboration avec Georg Menz, et sont tirés de l'article [FM14]
- Le chapitre 3 étudie le comportement de l'entropie microscopique dans le cadre de [GOVW09]. J'y ai montré qu'elle converge vers l'entropie associée à l'équation hydrodynamique. J'ai également montré que, à tout temps $t > 0$, l'état du système est proche au sens de l'entropie relative d'une mesure de Gibbs locale. Ceci est une notion de convergence plus forte que la convergence en distance de Wasserstein, qui est l'hypothèse qu'on suppose être satisfaite par les conditions initiales. Ce travail a été publié à *ALEA* en 2013, dans l'article [Fat13b].
- Le chapitre 4 généralise les résultats de [GOVW09] et du chapitre 2 au cas de dynamiques faiblement asymétriques, obtenues en rajoutant un terme supplémentaire à la dynamique étudiée dans les chapitres 2 et 3. Ce terme supplémentaire ne modifie pas la mesure invariante, mais rend le processus de Markov non-réversible. Ces résultats ont été obtenus en collaboration avec Manh Hong Duong et ont été soumis pour publication dans l'article [DF14].
- Le chapitre 5 s'intéresse à une variante de l'inégalité de Sobolev logarithmique, appelée inégalité de Sobolev logarithmique modifiée, qui a été introduite par Bobkov et Ledoux dans [BL00]. On y donne une condition suffisante pour que les mesure canoniques utilisées dans les chapitres précédents vérifient cette inégalité, avec une constante indépendante de la dimension. Ce chapitre correspond à l'article [Fat13a]
- Le chapitre 6 s'intéresse aux liens entre flots gradients dans les espaces de mesure et grandes déviations pour les solutions d'équations différentielles stochastiques. On y montre que la suite des lois des trajectoires de solutions d'équations différentielles stochastiques réversibles vérifie un principe de grandes déviations si et seulement si une certaine suite de fonctionnelles qui apparait dans la formulation en flot gradient du flots de marginales des processus Γ -converge vers la fonction de taux du principe de grandes déviations. Comme application, on obtient les grandes déviations par rapport à la limite hydrodynamique pour deux variantes du modèle de Ginzburg-Landau. Ces résultats sont tirés de l'article [Fat14]
- Le chapitre 7 concerne les inégalités transport-entropie pour des schémas d'approximation numérique. Le résultat principal est une inégalité transport-entropie pour la loi à un instant donné du schéma d'Euler-Maruyama utilisé pour simuler la solution d'une équation différentielle stochastique, dans le cas où le coefficient de diffusion n'est pas borné. Cette inégalité implique ensuite des bornes de concentration sur l'erreur statistique pour de telles approximations, avec un ordre de grandeur optimal sous nos hypothèses. Ces résultats ont été obtenus en collaboration avec Noufel Frikha, et ont été publiés à *Electronic Journal of Probability* en 2013, dans l'article [FF13].
- Le chapitre 8 s'intéresse à l'erreur de discrétisation pour l'estimation des coefficients de transport pour l'équation de Langevin, lorsqu'on utilise l'algorithme de Metropolis

pour stabiliser le schéma d'Euler-Maruyama, connu sous le nom d'algorithme MALA. On y montre que cette erreur est d'ordre 1 en le pas de discrétisation. Ces résultats ont été obtenus en collaboration avec Ahmed-Amine Hommann et Gabriel Stoltz, et ont été accepté pour publication à *ESAIM : Proceedings*, dans l'article [FHS].

- Le chapitre 9 présente une modification de l'algorithme MALA utilisé au chapitre 8. Le schéma numérique modifié obtenu a l'avantage d'avoir une meilleure erreur trajectorielle forte à l'équilibre que l'algorithme MALA standard. Ce résultat a été obtenu en collaboration avec Ahmed-Amine Hommann et Gabriel Stoltz.

1.1 Transport optimal

1.1.1 Transport optimal et distances de Wasserstein

Dans cette section, je vais faire une brève introduction au transport optimal, afin de pouvoir introduire les distances de Wasserstein sur un espace de mesures de probabilité. Les preuves des résultats énoncés ici pourront être trouvés dans [Vil03] et [Vil09] par exemple.

Au XVIIIème siècle, Gaspard Monge a introduit dans [Mon84] le problème du transport d'une masse d'un emplacement à l'autre. En termes modernes, on représente une répartition de masse par une mesure sur un espace E , et en supposant que la masse totale est égale à 1, on prend une mesure de probabilité μ . On considère ensuite une fonction $c : E \times E \rightarrow \mathbb{R}_+$, où $c(x, y)$ représente le coût du déplacement d'une unité de masse du point x vers le point y . Il nous faut encore choisir comment représenter une manière de déplacer la masse de la configuration représentée par μ vers celle représentée par une autre mesure de probabilité ν . Ceci nous mène à la définition suivante :

Définition 1.1. Soit (E, μ) et (F, ν) deux espaces de probabilité. Un couplage de μ et de ν est une mesure π sur $E \times F$ dont les marginales sont μ et ν , *i.e.* pour toute partie mesurable A de E (resp. B de F) on a $\pi(A \times F) = \mu(A)$ (resp. $\pi(E \times B) = \nu(B)$)

Remarque 1.2. Un couplage π est la loi d'un couple de variables aléatoires (X, Y) , où X a pour loi μ et Y a pour loi ν

Exemple 1.3. La mesure produit est un couplage, qui correspond au cas où U et V sont indépendantes.

Un couplage entre deux mesures de probabilité μ et ν va alors représenter une manière de transporter la masse de μ vers ν . En effet, on peut voir $\pi(A, B)$ comme la quantité de masse présente dans une région A de l'espace que l'on va transporter dans la région B . Dans certains cas, le support du couplage sera le graphe d'une fonction $T : E \rightarrow F$. Ceci correspond au cas où toute la masse présente en un point x est transportée en un unique point $T(x)$. Intuitivement, cela veut dire qu'on ne sépare pas la masse. On parle alors de couplage déterministe (le point d'arrivée ne dépend que du point de départ).

On peut ensuite s'intéresser à la meilleure manière de transporter cette masse, pour la fonction de coût c . Comme le coût du transport pour un couplage π est donné par $\int c(x, y)\pi(dx, dy)$, on va s'intéresser au couplage qui minimise cette quantité. En général, ce couplage n'existe pas toujours, mais dans des cadres plus restreints on peut affirmer son existence. Mais ici, on ne s'intéressera qu'au cas particulier des distances de Wasserstein, définies de la manière suivante :

Définition 1.4. Soit (E, d) un espace métrique séparable complet, et p un réel supérieur ou égal à 1. Pour tout couple de mesures (μ, ν) sur E , on définit la distance de Wasserstein

d'ordre p entre μ et ν par

$$W_{p,d}(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left(\int d(x, y)^p \pi(dx, dy) \right)^{\frac{1}{p}}.$$

Dans ce cadre (i.e. E espace polonais, et c est une puissance d'une distance), le couplage optimal existe :

Proposition 1.5. *Dans la définition précédente, il existe un couplage tel que l'infimum soit atteint pour ce couplage.*

Remarque 1.6. On peut écrire cette définition en termes de variables aléatoires, et elle devient alors

$$W_{p,d}(\mu, \nu) = \inf_{L(X, Y) \in \Pi(\mu, \nu)} \mathbb{E}[d(X, Y)^p]^{\frac{1}{p}},$$

où la notation $L(X, Y)$ désigne la loi jointe du couple de variables aléatoires (X, Y) .

Une des raisons pour lesquelles ce cas particulier de transport optimal est intéressant est donnée par le théorème suivant :

Théorème 1.1. W_p définit une distance sur $P_p(E) = \{\mu \in P(E); \int d(x_0, x)^p \mu(dx) < \infty\}$

Remarque 1.7. L'inégalité $d(x, y)^p \leq 2^{p-1}(d(x, x_0)^p + d(x_0, y)^p)$ montre que la définition de $P_p(E)$ ne dépend pas du choix de x_0 , et que W_p est finie sur $P_p(E) \times P_p(E)$.

On a le résultat suivant sur la topologie de l'espace métrique obtenu :

Théorème 1.2. *Une suite (μ_n) de mesures de $P_p(E)$ converge vers $\mu \in P_p(E)$ pour la distance de Wasserstein ssi elle converge étroitement et les moments d'ordre p convergent. $(P_p(E), W_p)$ est séparable complet.*

Les résultats de cette thèse utilisent les distances de Wasserstein dans trois cadres : pour leurs liens avec le phénomène de concentration de la mesure (qui sera introduit dans la section suivante), pour quantifier la vitesse de convergence de suites de lois de probabilités vers leur limite et pour définir la notion de flots gradients dans les espaces de mesures de probabilités.

Voici un résultat important sur la structure du transport optimal : le théorème de Brenier [Bre91]

Théorème 1.3. *On se place dans le cas où E est l'espace Euclidien \mathbb{R}^d . Supposons que la mesure μ est absolument continue par rapport à la mesure de Lebesgue. Alors le couplage optimal pour la distance de Wasserstein W_2 existe, et est déterministe (ie de la forme $\pi(dx, dy) = \delta_{y=T(x)}\mu(dx)$). De plus, la fonction T est le gradient d'une fonction convexe.*

On interprète le caractère déterministe de ce couplage optimal comme signifiant qu'on ne sépare pas la masse : toute la masse présente à un point x est envoyé en un même point $T(x)$. Le fait que ce couplage est le gradient d'une fonction convexe vient du fait que, si ce n'était pas le cas, on pourrait faire des permutations pour diminuer le coût du transport.

Ceci n'est pas la forme la plus générale possible d'un tel résultat. Il est valide pour des coûts de transport plus généraux par exemple. Il a également été étendu au cas où on se place sur une variété Riemannienne par McCann dans [McC01].

À noter qu'en dimension 1, ce couplage optimal peut être rendu explicite, à l'aide des fonctions de répartition de μ et ν , car il n'y a qu'une seule fonction croissante qui envoie μ sur ν . Ce couplage optimal est d'ailleurs le même pour toutes les distances W_p .

1.1.2 Transport optimal dynamique et formule de Benamou-Brenier

Une fois que l'on dispose du transport optimal entre μ et ν , on peut naturellement se demander de quelle manière il faut déplacer la masse, et pas seulement quel est le point d'arrivée. Ceci nous mène à construire un chemin de mesures de probabilités (μ_t) tel que $\mu_0 = \mu$ et $\mu_1 = \nu$ tel que la variation du coût de transport le long du chemin (ie la dérivée du coût de transport entre μ et μ_t) soit constante.

Dans le cas des distances de Wasserstein, ceci revient à construire une géodésique dans l'espace des mesures de probabilité. Pour W_2 , il s'avère que ceci est possible si et seulement si l'espace métrique sous-jacent est lui-même géodésique. Si de plus les géodésiques sont uniques (ceci est le cas dans \mathbb{R}^d par exemple), μ_t est alors l'image d'un transport optimal par l'application $e_t : (x, y) \rightarrow \gamma_{x,y}(t)$, où $\gamma_{x,y}$ est la géodésique allant de x à y . Ce concept a été introduit par McCann dans [McC97], sous le nom de « displacement interpolation », ou interpolation par déplacement.

Cette géodésique est alors le chemin reliant μ_0 à μ_1 qui minimise une certaine énergie. On peut alors réécrire le problème de transport optimal pour le coût quadratique comme un problème de minimisation de l'énergie le long des chemins. Ceci mène au résultat suivant, appelé formule de Benamou-Brenier.

Théorème 1.4. *Soit μ_0 et μ_1 deux mesures de probabilité sur \mathbb{R}^d . Alors*

$$W_2(\mu_0, \mu_1) = \inf \left\{ \int_0^1 \int |v(t, x)|^2 d\mu_t(x) dt; \dot{\mu}_t + \nabla \cdot (v\mu) = 0 \right\}.$$

L'inf porte sur tous les couples de chemins μ_t et de champs de vecteurs $v(t, x)$ qui satisfont la relation $\dot{\mu}_t + \nabla \cdot (v\mu) = 0$. Cette condition s'interprète comme la conservation de la masse, puisque l'on ne considère que des chemins de mesures de probabilités.

A noter qu'on peut ne considérer que des champs de vecteurs de la forme $v = \nabla\psi$ sans modifier la conclusion.

Ce résultat a été formellement justifié par Benamou et Brenier [BB99]. Des preuves rigoureuses peuvent être trouvées dans [AGS08] et [Vil03]. Il est encore valable, sous certaines conditions, sur les variétés riemanniennes (par exemple, sur toute variété riemannienne compacte).

Ce résultat permet notamment de munir formellement l'espace $(P_2(\mathbb{R}^d), W_2)$ d'une structure de variété Riemannienne de dimension infinie. Ce formalisme a été proposé dans [Ott01], et permet de mieux comprendre la notion de flots gradients que l'on verra plus loin, ainsi que de réinterpréter certaines inégalités fonctionnelles comme des propriétés de l'espace métrique mesuré, cf. [OV00].

Les géodésiques en distance W_2 se sont avérées un outil très utile dans l'étude de la géométrie des espaces métriques mesurés. La notion de borne inférieure sur la courbure de Ricci s'est avérée équivalente à une propriété de convexité de l'entropie le long des géodésiques pour W_2 . En courbure positive, on peut ensuite déduire certaines inégalités fonctionnelles (inégalité de Sobolev logarithmique, inégalité HWI,...) que nous allons voir dans la section suivante.

1.2 Concentration de la mesure et inégalités fonctionnelles

L'étude du phénomène de concentration de la mesure consiste à regarder la manière dont se répartit la masse dans un espace métrique muni d'une mesure, que l'on supposera ici toujours être une mesure de probabilités. Typiquement, on considère des observables F à

valeurs réelles, et on cherche des bornes supérieures sur les quantités $\mu(F(x) \geq \mu(F) + r)$, où r est un réel positif.

Définition 1.8. Soit $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction croissante, avec $\Phi(0) = 0$. Une mesure de probabilité μ sur un espace métrique (E, d) satisfait une inégalité de concentration avec profil Φ si, pour toute fonction $f : X \rightarrow \mathbb{R}$ 1-lipschitzienne, on a pour tout $r \geq 0$

$$\mu \left\{ x \in X; f(x) - \int f d\mu \geq r \right\} \leq \exp(-\Phi(r)).$$

En prenant l'inégalité analogue avec $-f$, on a alors également

$$\mu \left\{ x \in X; \left| f(x) - \int f d\mu \right| \geq r \right\} \leq 2 \exp(-\Phi(r)).$$

Une inégalité de concentration implique que, pour tout borélien A avec $\mu(A) \geq 1/2$, si on définit

$$A_r := \{x \in E; d(x, A) \geq r\}$$

on ait

$$\mu(A_r^c) \leq \exp(-\Phi(r/2)).$$

L'intérêt d'une telle inégalité est le suivant : imaginons qu'on cherche à deviner le résultat d'une expérience dans laquelle intervient un aléa. Une hypothèse raisonnable est la valeur moyenne du résultat. Une inégalité de concentration nous donne une borne sur la probabilité que le vrai résultat soit très éloigné de cette moyenne. Ceci nous permet donc d'avoir une idée a priori de si cette hypothèse est raisonnable ou pas, par exemple à travers des intervalles de confiance.

Les deux cas particuliers typiquement étudiés sont la concentration gaussienne et la concentration exponentielle, qui correspondent respectivement à $\Phi(r) = cr^2$ et $\Phi(r) = cr$. Comme l'indiquent ces noms, les exemples emblématiques de mesures satisfaisant ces cas de concentration sont respectivement la mesure gaussienne et la mesure exponentielle sur \mathbb{R}^d .

Une inégalité de concentration donne une borne sur la probabilité d'être « loin » d'une zone de grande masse. Un tel résultat est très lié à la géométrie de l'espace métrique mesuré considéré. L'étude de ces inégalités remonte à la preuve par V. Milman du théorème de Dvoretzky sur la structure locale des espaces de Banach [Mil71], qui est basée sur l'inégalité isopérimétrique sur les sphères. Elles ont de nombreuses applications :

- En géométrie. Un exemple célèbre est le théorème de Lévy-Gromov, qui nous dit que le volume sur la sphère de dimension n et de rayon 1 se concentre sur une bande autour de l'équateur, de largeur d'ordre $1/\sqrt{n}$ (voir [Gro80], et le chapitre 3 $\frac{1}{2}$ de [Gro07] pour le cadre Riemannien). Plus généralement, on a de nombreux liens entre la concentration de la mesure et la courbure des variétés Riemanniennes. Pour plus de détails, on pourra consulter le chapitre 22 de [Vil09] ou le chapitre 3 de [Led01]
- En probabilités. On en verra plusieurs dans cette thèse, mais on peut aussi mentionner les applications aux matrices aléatoires, en combinatoire, en percolation... Un certain nombre de ces applications sont décrites dans le chapitre 8 de [Led01]
- En statistiques, pour obtenir des bornes sur les intervalles de confiance pour des estimateurs. On peut consulter pour plus de détails [Mas07]

Dans la suite de cette section, nous allons voir quelques inégalités fonctionnelles liées au phénomène de concentration de la mesure. Pour plus de détails, on pourra consulter [Led01] et le chapitre 22 de [Vil09].

1.2.1 Inégalité de Sobolev logarithmique

L'inégalité de Sobolev logarithmique (qu'on abrègera souvent en LSI dans la suite, pour « logarithmic Sobolev inequality ») a été introduite par Gross dans [Gro75] pour étudier certaines propriétés du semigroupe généré par l'opérateur $L = \Delta - \nabla V \cdot \nabla$. Elle s'est avérée très utile pour étudier le comportement en temps long des équations d'évolution, et le phénomène de concentration de la mesure. Elles sont particulièrement utiles pour les problèmes en grande dimension. Dans cette section, je vais définir cette inégalité, présenter certaines de ses propriétés et quelques conditions suffisantes pour sa validité. Pour une étude plus complète, on pourra se référer à [ABC⁺00] et à [Led01].

Définition 1.9. L'entropie d'une fonction positive $f : E \rightarrow \mathbb{R}$ par rapport à une mesure μ sur E est définie comme

$$\text{Ent}_\mu(f) := \int f \log f d\mu - \left(\int f d\mu \right) \log \left(\int f d\mu \right).$$

Si ν est une autre mesure sur E , on définit l'entropie relative de ν par rapport à μ , notée $\text{Ent}_\mu(\nu)$, comme $\text{Ent}_\mu(f)$ si ν admet une densité f par rapport à μ , et $+\infty$ sinon.

D'après l'inégalité de Csiszár-Kullback-Pinsker, étant donné deux mesures de probabilité μ et ν sur un espace polonais, on a l'inégalité

$$\|\mu - \nu\|_{TV} \leq \sqrt{2 \text{Ent}_\mu(\nu)}$$

où $\|\cdot\|_{TV}$ est la norme en variation totale. On pourra se référer à [BV05] pour une présentation de cette inégalité, ainsi que plusieurs généralisations. L'entropie relative est donc une manière de mesurer à quel point deux mesures de probabilité sont distantes, à priori plus forte que la convergence étroite.

Pour définir les inégalités de Sobolev logarithmiques, on va commencer par étendre la définition de la longueur du gradient aux fonctions localement lipschitziennes sur un espace métrique.

Définition 1.10. Soit E un espace euclidien et $f : E \rightarrow \mathbb{R}$ une fonction localement lipschitzienne. On pose $|\nabla f|(x) = \limsup_{y \rightarrow x} \frac{|f(x) - f(y)|}{d(x,y)}$. Cette définition coïncide bien avec la définition usuelle si f est une fonction différentiable sur un espace euclidien.

On définit aussi l'information de Fisher :

Définition 1.11. Soit (E, d) un espace métrique, et μ et ν deux mesures boréliennes sur E . On définit l'information de Fisher de μ par rapport à ν par

$$I_\nu(\mu) = \int \frac{|\nabla f|^2}{f} d\nu$$

si μ admet une densité f par rapport à ν qui soit localement lipschitzienne, et par $+\infty$ sinon.

Proposition 1.12. On a aussi les définitions alternatives (équivalentes) de l'information de Fisher suivantes :

$$\begin{aligned} I_\nu(\mu) &= \int |\nabla(\log f)|^2 f d\nu \\ &= \int 4|\nabla(\sqrt{f})|^2 d\nu. \end{aligned}$$

Ceci va nous servir pour la définition des inégalités de Sobolev logarithmiques.

Définition 1.13. Soit E un espace métrique. On dit qu'une mesure de probabilité μ sur E satisfait une inégalité log-Sobolev de constante ρ (et on le notera $LSI(\rho)$) si, pour toute fonction f positive, localement lipschitzienne et L^1 , on a

$$Ent_\mu(f) \leq \frac{1}{\rho} \int \frac{|\nabla f|^2}{2f} d\mu \quad (1.1)$$

Remarque 1.14. Si f est une densité de probabilité, on reconnaît dans le terme de droite $\frac{I_\mu(f\mu)}{2\rho}$.

Remarque 1.15. On peut se contenter de montrer que (1.1) est vérifiée pour les fonctions positives, localement lipschitziennes et bornées, car si f est positive, localement lipschitzienne et L^1 , alors $f \wedge n$ l'est aussi, et la longueur de son gradient est en tout point inférieure à celle de f , et ne diffère qu'en les points où $f = n$. Si, pour tout n , (1.1) est vérifiée pour $f \wedge n$, alors pour tout n on a

$$Ent_\mu(f \wedge n) \leq \frac{1}{\rho} \int \frac{|\nabla f|^2}{2f} d\mu$$

et le terme de gauche converge vers $Ent_\mu(f)$ lorsque n tend vers l'infini, par convergence monotone.

On appelle ce type d'inégalités des inégalités de Sobolev logarithmique car elles donnent un critère du même type que les injections de Sobolev, mais plutôt que de dire qu'une fonction de H^1 appartient à un espace L^p , avec $p > 2$, elles disent que de telles fonctions sont dans un espace $L^2 \log L$, appelé espace d'Orlicz. Elle a des applications intéressantes pour l'étude de la convergence vers l'équilibre de processus de Markov, et en concentration de la mesure, que nous verrons plus loin.

Un exemple important d'espace de probabilité admettant une inégalité de Sobolev logarithmique est l'espace \mathbb{R}^n muni de la mesure gaussienne :

Proposition 1.16. *L'espace \mathbb{R}^n muni de la mesure gaussienne standard*

$$\gamma_n(dx) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\sum_{i=1}^n x_i^2/2\right) dx$$

vérifie $LSI(1)$.

La première preuve de ce résultat a été donnée par Gross dans [Gro75]. La constante 1 est optimale.

Une des applications importantes des inégalités de Sobolev logarithmiques est qu'elles impliquent un phénomène de concentration gaussienne :

Proposition 1.17. *Si une mesure de probabilités vérifie $LSI(\rho)$, alors elle vérifie une inégalité de concentration avec profil $\Phi(r) = \rho r^2/2$.*

L'implication inverse est en général fausse, mais si l'espace métrique mesuré (X, d, μ) vérifie une condition de courbure dimension $CD(-\kappa, +\infty)$, une concentration gaussienne suffisamment forte implique une inégalité de Sobolev logarithmique. Ceci a été prouvé dans [Mil10] dans le cas des variétés Riemanniennes, et dans [GRS11a] pour le cas des espaces métriques. On réfère au chapitre 16 de [Vil09] pour une présentation des conditions de courbure-dimension.

Une seconde propriété importantes des inégalités de Sobolev logarithmiques est qu'elles impliquent une inégalité de Poincaré.

Définition 1.18. Une mesure μ vérifie une inégalité de Poincaré de paramètre $\rho > 0$ si, pour toute fonction localement lipschitzienne $f : E \rightarrow \mathbb{R}$, on a

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\rho} \int |\nabla f|^2 d\mu.$$

On notera cette condition $\mathrm{SG}(\rho)$ (pour « spectral gap »).

En linéarisant l'inégalité de Sobolev logarithmique, on peut montrer que $\mathrm{LSI}(\rho)$ implique une inégalité de Poincaré de paramètre ρ . Cette inégalité est reliée à des propriétés de concentration exponentielle, et à l'existence d'une borne inférieure strictement positive sur les valeurs propres non nulles de l'opérateur linéaire $\Delta - \nabla V \cdot \nabla$ sur $H^1(\mu)$ (qui est l'espace des fonctions dont le gradients est de carré intégrable par rapport à μ), où V est donnée par la relation $\mu = \exp(-V)dx$. Elle est souvent utilisée pour étudier la vitesse de convergence vers l'équilibre de processus de Markov, cf. [MT09]. A noter que l'inégalité de Poincaré est valable dans un cadre beaucoup plus général que l'inégalité de Sobolev logarithmique. Par exemple, elle est valable pour les mesures de probabilités qui sont simplement log-concaves (voir [BBCG08]). Ainsi, la mesure exponentielle sur \mathbb{R}^d vérifie une inégalité de Poincaré, mais ne peut pas satisfaire une inégalité de Sobolev logarithmique.

Une inégalité de Sobolev logarithmique permet donc d'obtenir des bornes sur les variances. Plus généralement, on peut aussi obtenir des inégalités sur les covariances :

Proposition 1.19. *Supposons que μ vérifie la condition $\mathrm{LSI}(\rho)$. Alors, pour toutes fonctions f et g à valeurs réelles, avec f positive, on a*

$$\left| \int f g d\mu - \left(\int f d\mu \right) \left(\int g d\mu \right) \right| \leq \frac{\|\nabla g\|_\infty}{\rho} \sqrt{\left(\int f d\mu \right) \left(\int \frac{|\nabla f|^2}{f} d\mu \right)}.$$

Une autre application des inégalités de Sobolev logarithmique est qu'elles impliquent une convergence vers l'équilibre pour les équations aux dérivées partielles de la forme

$$\partial_t \rho = \Delta \rho - \nabla V \cdot \nabla \rho,$$

avec V une fonction à valeurs réelles sur \mathbb{R}^d . Cette équation aux dérivées partielles donne l'évolution en temps de la loi des solutions de l'équation différentielle stochastique

$$dX_t = -\nabla V(X_t) + \sqrt{2}dB_t$$

où B est un mouvement Brownien standard sur \mathbb{R}^d . On peut voir que la mesure de probabilité $\mu(dx) = Z^{-1} \exp(-V)dx$ est une mesure invariante pour cette dynamique (si elle existe). Si elle vérifie une inégalité de Sobolev logarithmique, on a alors, en écrivant $\rho_t = f_t \mu$ une solution de l'équation aux dérivées partielles,

$$\frac{d}{dt} \mathrm{Ent}_\mu(\rho_t) = -I(\rho_t) \leq -C \mathrm{Ent}_\mu(\rho_t)$$

et donc

$$\mathrm{Ent}_\mu(\rho_t) \leq \exp(-Ct) \mathrm{Ent}_\mu(\rho_0).$$

Il y a donc alors convergence exponentielle vers l'équilibre, en entropie relative. On remarque au passage qu'on peut interpréter l'inégalité de Sobolev logarithmique comme une inégalité entropie/production d'entropie pour l'équation d'évolution associée à l'opérateur $L = \Delta - \nabla V \cdot \nabla$, où V est l'opposé du logarithme de la densité de μ .

Une dernière application est que cette inégalité est équivalente à l'hypercontractivité du semigroupe P_t associé au générateur $L = \Delta - \nabla V \cdot \nabla$, c'est à dire que l'espace L^2 est envoyé dans l'espace L^p après un temps $t(p)$, pour tout $p \geq 2$. L'étude du phénomène d'hypercontractivité était d'ailleurs la motivation originale de l'introduction de l'inégalité dans [Gro75].

On dispose de trois critères classiques pour montrer qu'une mesure satisfait une inégalité de Sobolev logarithmique : le principe de tensorisation, le lemme de Holley-Stroock, et le théorème de Bakry-Emery :

Proposition 1.20. (i) Si deux mesures μ et ν satisfont une inégalité de Sobolev logarithmique avec constantes ρ_1 et ρ_2 , alors la mesure $\mu \otimes \nu$ sur l'espace produit vérifie $LSI(\min(\rho_1, \rho_2))$;

(ii) Soit f une fonction bornée. Si une mesure μ vérifie $LSI(\rho)$, alors la mesure de probabilité $\nu(dx) := Z^{-1} \exp(f(x))\mu(dx)$ vérifie $LSI(\rho \exp(-2 \operatorname{osc}(f)))$, où $\operatorname{osc}(f) := \sup f - \inf f$;

(iii) Si V vérifie la condition $\operatorname{Hess} V \geq \lambda Id$ avec $\lambda > 0$, alors $\mu(dx) = Z^{-1} \exp(-V(x))dx$ vérifie $LSI(\lambda)$.

Grâce à la propriété de tensorisation, les inégalités de Sobolev logarithmiques sont particulièrement pratiques dans les problèmes en grande dimension.

Pour le théorème de Bakry-Émery, la propriété $\operatorname{Hess} f \geq \lambda Id$ avec $\lambda > 0$ est équivalente à une propriété d'uniforme convexité de l'entropie relative Ent_μ le long des géodésiques de $\mathcal{P}_2(\mathbb{R}^d)$ pour la distance W_2 . L'information de Fisher apparait alors comme le carré de la norme du « gradient » de l'entropie, et l'inégalité de Sobolev logarithmique peut en être ensuite déduite. On reparlera de cette interprétation dans la section 1.4. On peut trouver une présentation complète de cet argument dans [OV00].

On s'intéressera aussi à la généralisation suivante de l'inégalité de Sobolev logarithmique, appelée inégalité de Sobolev logarithmique modifiée, qui a été introduite par Bobkov et Ledoux dans [BL00] :

Définition 1.21. Une mesure μ sur \mathbb{R}^d vérifie une p-inégalité de Sobolev logarithmique modifiée de paramètre ρ si, pour toute fonction localement lipschitzienne positive f , on a

$$\operatorname{Ent}_\mu(f) \leq \frac{1}{\rho} \int \frac{\|\nabla f\|_q^q}{f^{q-1}} d\mu.$$

Il existe un analogue du théorème de Bakry-Émery pour ces inégalités. Si V est uniformément p-convexe, c'est à dire si il existe $c > 0$ tel que

$$V(tx + (1-t)y) \leq tV(x) + (1-t)V(y) - \frac{c}{p}t(1-t)\|x - y\|_p^p \quad (1.2)$$

pour tout x et y , alors la mesure $\mu = \exp(-V(x))dx$ vérifie une p-inégalité de Sobolev logarithmique modifiée. Ce résultat a été prouvé dans [BL00].

Les inégalités de Sobolev logarithmique modifiées impliquent également une propriété de concentration, pour la norme ℓ^p .

Proposition 1.22. Supposons que la mesure μ sur \mathbb{R}^d vérifie une p-inégalité de Sobolev logarithmique modifiée de paramètre ρ . Alors, pour toute fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ 1-Lipschitzienne pour la norme $\|\cdot\|_p$ et pour tout $r \geq 0$, on a

$$\mu\left(f \geq \int f d\mu + r\right) \leq \exp\left(-\frac{\rho r^p}{p(p-1)^{p-1}}\right).$$

1.2.2 Inégalités transport-entropie

L'inégalité de Talagrand (ou T_2) classique a été introduite par Talagrand [Tal96].

Définition 1.23. Une mesure de probabilité μ vérifie $T_2(\rho)$ si, pour toute mesure de probabilité ν , on a

$$W_2(\mu, \nu) \leq \sqrt{\frac{2 \text{Ent}_\mu(\nu)}{\rho}}.$$

L'une des raisons pour laquelle on s'est intéressé à cette inégalité est qu'elle implique la concentration gaussienne, qui a été démontrée dans [Tal96]. Comme pour les inégalités de Sobolev logarithmiques, l'implication inverse est fautive en générale (on peut trouver un contre-exemple dans [CG06]), mais devient vraie sous une hypothèse de borne inférieure sur la courbure. Ceci a été prouvé par E. Milman dans [Mil10]. Il a été démontré dans [Goz09] que l'inégalité de Talagrand est équivalente à une forme renforcée de concentration gaussienne, indépendante de la dimension.

Comme la topologie associée à la distance W_2 est plus forte que celle associée à la distance en variation totale, si μ vérifie une inégalité T_2 , la convergence en entropie relative vers μ est strictement plus forte que la convergence en variation totale, ce qui renforce l'inégalité de Csiszàr–Kullback–Pinsker.

Comme pour l'inégalité de Sobolev logarithmique, l'inégalité de Talagrand est stable par tensorisation (cf. [Tal96]) et par perturbation bornée (cf [GRS11b])

Comme l'inégalité de Talagrand et l'inégalité de Sobolev logarithmique impliquent toutes les deux la concentration gaussienne, une question naturelle est de savoir si ces deux inégalités sont comparables. La réponse suivante a été apportée par Otto et Villani dans [OV00]

Théorème 1.5 ([OV00]). *Si une mesure μ vérifie $LSI(\rho)$, alors elle vérifie $T_2(\rho)$.*

Plusieurs preuves de ce résultat ont été données ([OV00], [BGL01], [Goz09], [GL13]). La preuve de [Goz09] est la plus générale, valable dans n'importe quel espace métrique. À noter que l'implication inverse n'est pas vraie en général (un contre-exemple a été donné dans [CG06]), mais le devient si on suppose en plus le critère de courbure dimension $CD(-K, \infty)$ valide, avec une constante positive K suffisamment petite.

Pour résumer, on a la hiérarchie suivante des inégalités fonctionnelles :

$$LSI(\rho) \Rightarrow T_2(\rho) \Rightarrow SG(\rho).$$

L'inégalité de Sobolev logarithmique n'est pas l'inégalité la plus forte à laquelle on peut s'intéresser. En effet, une inégalité isopérimétrique implique une inégalité de Sobolev logarithmique. [Led] est une bonne introduction au sujet.

Dans cette thèse, je me suis intéressé à plusieurs cas de la généralisation suivante de l'inégalité de Talagrand :

Définition 1.24. Soit $p \geq 1$ et $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction convexe, croissante vérifiant $\Phi(0) = 0$. Une mesure μ satisfait l'inégalité de transport-entropie $T_{p,\Phi}$ si, pour toute mesure ν , on a

$$\Phi(W_p(\mu, \nu)) \leq \text{Ent}_\mu(\nu).$$

Lorsque $p = 2$ et $\Phi(r) = \rho r^2/2$, on reconnaît $T_2(\rho)$. À fonction Φ donnée, plus p est grand, plus cette inégalité est contraignante.

Pour le cas $p = 1$, on a la caractérisation suivante, obtenue par Bobkov et Gotze dans [BG99] :

Proposition 1.25. *Une mesure de probabilité μ vérifie $T_{1,\Phi}$ ssi, pour toute fonction $f : X \rightarrow \mathbb{R}$ 1-lipschitzienne, on a pour tout $\lambda \geq 0$*

$$\int e^{\lambda f} d\mu \leq \exp\left(\lambda \int f d\mu + \Phi^*(\lambda)\right)$$

où Φ^* est la transformée de Legendre de Φ .

Comme pour l'inégalité de Talagrand, on a un lien entre inégalités transport-entropie et concentration de la mesure. Par exemple, si une mesure de probabilité vérifie $T_{1,\Phi}$, alors elle vérifie une inégalité de concentration avec profil Φ . Les liens dans un cadre plus général ont été étudiés notamment par Marton dans [Mar96]

Dans le cas des mesures log-concaves, on dispose également de l'inégalité HWI suivante, qui interpole entre l'inégalité de Sobolev logarithmique et l'inégalité de Talagrand :

Théorème 1.6. *Soit $\mu = \exp(-V)dx$ une mesure de probabilité sur \mathbb{R}^d . Si $\text{Hess } V \geq K \text{Id}$, avec $K \in \mathbb{R}$, alors pour toute mesure de probabilité ν on a l'inégalité*

$$\text{Ent}_\mu(\nu) \leq W_2(\mu, \nu) \sqrt{I_\mu(\nu)} - \frac{K}{2} W_2(\mu, \nu)^2.$$

Cette inégalité est encore valide sur les variété Riemanniennes, sous une condition de borne inférieure sur la courbure de Ricci.

On peut remarquer que l'inégalité HWI implique le théorème de Bakry-Emery, puisque d'après l'inégalité de Young on a

$$W_2(\mu, \nu) \sqrt{I_\mu(\nu)} \leq \frac{K}{2} W_2(\mu, \nu)^2 + \frac{1}{2K} I_\mu(\nu).$$

Le théorème d'Otto-Villani peut être généralisé au cas des inégalités de Sobolev logarithmiques modifiées, qui impliquent alors une inégalité de transport-entropie T_{p,cr^p} . Ce résultat a été démontré dans [GRS14].

Théorème 1.7. *Si une mesure de probabilité μ sur \mathbb{R}^d vérifie une p -inégalité de Sobolev logarithmique modifiée de paramètre ρ , alors elle vérifie l'inégalité*

$$\inf_{\pi \in \Pi(\mu, \nu)} \int \|x - y\|_p^p \pi(dx, dy) \leq \frac{p}{\tilde{\rho}} \text{Ent}_\mu(\nu)$$

avec $\tilde{\rho} = ((p-1)\rho)^{p-1}$.

On reconnaît dans le terme de gauche une distance de Wasserstein W_p^p , associée à la norme ℓ^p sur \mathbb{R}^d . Cette inégalité est à la p -inégalité de Sobolev logarithmique modifiée ce que l'inégalité de Talagrand T_2 est à l'inégalité de Sobolev logarithmique usuelle.

1.3 Limites hydrodynamiques

Le but de l'étude des limites hydrodynamiques est de justifier rigoureusement les équations aux dérivées partielles utilisées par les physiciens pour décrire l'évolution des quantités thermodynamiques d'un fluide.

Lorsqu'on s'intéresse à l'évolution d'un gaz de particules, comme le nombre de particules est très grand (typiquement, de l'ordre du nombre d'Avogadro $N_a = 6.10^{23}$), il serait irréaliste d'essayer de décrire son état en donnant la liste des positions et des vitesses de toutes les particules. À la place, on s'intéresse à l'évolution d'un petit nombre de variables

macroscopiques qui caractérisent les états d'équilibre (densité, pression, température par exemple). Pour un système hors équilibre, il est naturel de considérer, dans une petite région de l'espace, dite mésoscopique (de volume faible devant le volume total, mais suffisamment grand pour contenir un grand nombre de particules), que le système est dans un état d'équilibre local, caractérisé par les valeurs (dans ce volume mésoscopique) des grandeurs thermodynamiques.

L'état du système est alors donné par un paramètre $p(t, x)$, qui donne la valeur des grandeurs thermodynamiques au voisinage de x à l'instant t . On s'attend à ce que ce paramètre évolue selon une équation aux dérivées partielles, qu'on appelle équation hydrodynamique. On peut formellement obtenir ces équations en appliquant les lois de Newton, non pas aux particules, mais à des éléments infinitésimaux de volume. Pour un fluide visqueux newtonien, on obtient les équations de Navier-Stokes, et les équations d'Euler pour un fluide non visqueux. Pour ces équations, on distingue deux régimes, dits compressible et incompressible, selon si on peut négliger les variations de masse volumique du fluide (ce qui est le cas en général pour l'eau liquide, mais pas pour les gaz à faible nombre de Mach).

La question de savoir si on peut rigoureusement obtenir les équations de la mécanique des fluides comme limites de systèmes de particules avec une évolution microscopique donnée par les lois de Newton fait partie du 6ème problème de Hilbert [Hil02]. Très peu de résultats dans cette direction ont été obtenus. Par exemple, le problème d'obtenir la loi de Fourier à partir d'une chaîne d'oscillateurs anharmoniques demeure ouvert à ce jour ([BLRB00]). À ma connaissance, la seule avancée récente dans cette direction est l'article [BGSR], qui obtient l'équation de la chaleur comme limite d'un gaz de sphères dures.

L'une des difficultés principales pour les systèmes évoluant selon les lois de Newton est l'absence de bonnes propriétés d'ergodicité pour la dynamique. En effet, pour montrer que la description microscopique est bien approximée par une description macroscopique, il faut que la distribution soit dans un état d'équilibre local. Il faut donc réussir à montrer que cette propriété d'équilibre local reste vérifiée au cours du temps malgré l'effet de la dynamique. Pour la plupart des dynamiques physiquement réalistes, comme les systèmes hamiltoniens, on ne sait pas prouver un tel résultat. La seule exception semble être le cas des sphères dures avec collisions élastiques sur un tore. Pour ce modèle, les cas de deux ([Sin70]) ou trois sphères ([KSS91]) ont été résolus, et une résolution du cas général a été récemment annoncée dans [Sim13].

Pour contourner cette difficulté, deux types de simplification ont été considérés. La première est le cas de gaz à faible densité, pour limiter les corrélations entre les différentes particules. La deuxième est l'ajout d'un bruit stochastique, ou le cas de dynamiques stochastiques. Les modèles auxquels je me suis intéressé dans cette thèse utilisent cette deuxième simplification.

Pour le cas des gaz à faible densité, on peut mentionner les travaux sur la dérivation des équations de la mécanique des fluides (Navier-Stokes incompressible, Euler incompressible) comme limites hydrodynamiques de l'équation de Boltzman (qui donne une description cinétique d'un gaz raréfié de N particules, en interaction à travers des collisions réversibles et élastiques). Le sujet a été présenté dans [Vil02], et [SR09] fait une description complète des avancées récentes.

L'obtention des équations de la mécanique des fluides à partir de modèles microscopiques stochastiques est un problème qui a été proposé par Morrey dans [Mor55]. Dans ce cadre, on considère des modèles dans lesquels deux mécanismes agissent :

- une force qui tend à faire diminuer une certaine énergie ;
- un bruit stochastique, qui empêche le système de rester bloqué en un minimum local

de l'énergie.

La combinaison de ces deux effets fera que la loi de l'état du système microscopique stochastique converge vers une mesure invariante. Dans cette thèse, on ne considèrera que des modèles où cette mesure invariante est unique, et explicitement connue.

Il y a deux méthodes générales qui ont été développées pour étudier les limites hydrodynamiques de modèles stochastiques :

- La méthode entropique, développée pour la première fois dans [GPV88], consiste à utiliser des bornes a priori sur l'entropie par rapport à la mesure invariante du système, ainsi que sur la production d'entropie, pour obtenir la convergence vers un objet limite. Cet objet limite est ensuite correctement identifiée à l'aide d'arguments de martingales. Une présentation a été faite dans [Com91].
- La méthode de l'entropie relative, due à Yau [Yau91], qui consiste à prouver une inégalité de type Gronwall pour l'entropie relative du système par rapport à une mesure dite de Gibbs locale, qui varie en temps et qui converge vers la solution de la limite hydrodynamique. Ceci implique alors que si la loi initiale est proche au sens de l'entropie relative d'une mesure de Gibbs locale, alors c'est encore le cas à tout temps positif. Cette méthode demande toutefois que l'équation limite admette des solutions régulières.

On pourra se référer à [KL99] pour une présentation de ces deux méthodes et de leur application à divers systèmes de particules. Parmi les résultats obtenus à ce jour sur les limites hydrodynamiques de modèles stochastiques, il faut mentionner :

- L'obtention des solutions faibles de Leray des équations de Navier-Stokes incompressibles à partir d'un modèle de gaz sur réseau par Quastel et Yau dans [QY98]
- La dérivation des équations d'Euler à partir d'un système de particules en interaction par Olla, Varadhan et Yau dans [OVY93], tant qu'il existe des solutions régulières.
- Les travaux sur la limite hydrodynamique de chaînes d'oscillateurs bruités, comme [BO05] et [LO12].

Le point de départ de cette thèse, l'article [GOVW09], propose une autre méthode, appelée méthode à deux échelles. Elle permet d'obtenir des estimées quantitatives en distance de Wasserstein entre l'état du système microscopique d'une taille donnée et la mesure de Dirac en un état déterministe obtenu en discrétisant l'EDP limite. Elle est également reliée à une méthode proposée dans [GOVW09] pour prouver des inégalités de Sobolev logarithmique. Ces deux aspects seront présentés dans les sections 1.3.2 et 1.3.1 respectivement.

On peut aussi mentionner que récemment, une nouvelle méthode a été proposée par Mischler et Mouhot dans [MM13]. En comparant les générateurs du processus discret et de l'équation limite, ils parviennent à obtenir des estimées quantitatives sur la convergence du modèle de Kac vers l'équation de Boltzman homogène en espace. Un intérêt particulier de leur méthode est qu'ils obtiennent des estimées uniformes en temps. Une présentation de ces résultats a été faite dans [Des]. Leur méthode a également été adaptée par Marahrens et Mouhot pour l'étude de la limite hydrodynamique du processus zero-range ([Mar12]).

1.3.1 Mesures canoniques et inégalité de Sobolev logarithmique

La loi jointe de N variables indépendantes, identiquement distribuées sur \mathbb{R} , chacune de loi $\mu(dx) = \exp(-\psi(x))dx$ est donnée par la mesure produit de densité

$$\mu_N(dx) = \exp\left(-\sum_{i=1}^N \psi(x_i)\right) dx.$$

En physique, on s'intéresse souvent à des phénomènes qui possèdent une loi de conservation : une certaine quantité $\Phi(x_1, \dots, x_N) \in \mathbb{R}^d$ est préservée par la dynamique. Cette quantité peut être une énergie ou une quantité de mouvement par exemple. Si la condition initiale (qui peut être donnée par une mesure de probabilités) vérifie presque sûrement $\Phi(x_1, \dots, x_N) = m$, toute loi limite (pour la convergence en temps long, à nombre de particules fixé) doit aussi avoir cette propriété. Ceci motive l'introduction des mesures canoniques, pour lesquelles on fixe la valeur de $\Phi(x_1, \dots, x_N)$:

Définition 1.26. La mesure canonique associée au potentiel ψ et à la loi de conservation $\Phi(x_1, \dots, x_N) = m$ est

$$\mu_{N,m}(dx) := \frac{1}{Z} \mathbb{1}_{\Phi(x)=m} \exp\left(-\sum_{i=1}^N \psi(x_i)\right) \mathcal{H}(dx)$$

où \mathcal{H} est la mesure de Hausdorff sur la sous-variété de \mathbb{R}^N $\{x \in \mathbb{R}^N; \Phi(x) = m\}$.

Les mesures canoniques représentent donc la loi jointe de particules indépendantes auxquelles on a imposé une contrainte.

Remarque 1.27. Pour assurer que $\{x \in \mathbb{R}^N; \Phi(x) = m\}$ est bien une sous-variété, il suffit de supposer que Ψ est lisse et que la matrice de Gram $\nabla\Phi\nabla\Phi^t$ est toujours inversible.

Dans cette thèse, on ne s'intéressera qu'au cas de la loi de conservation $\Phi(x_1, \dots, x_N) = \frac{1}{N} \sum x_i$. Ceci correspond par exemple à la conservation du moment pour des systèmes de particules de même masse en dimension 1.

Dans l'article [Var93], Varadhan a posé la question de savoir à quelle condition sur ψ les mesures canoniques $\mu_{N,m}$ vérifient une inégalité de Poincaré avec une constante indépendante de la dimension N et de la valeur m . Une généralisation naturelle de cette question est de demander à quelle condition on a l'inégalité de Sobolev logarithmique, qui est plus forte. Cette question est motivée par un problème de limite hydrodynamique dont on reparlera plus loin.

Ce problème a motivé un certain nombre de travaux. Le cas où ψ est à support compact a été résolu dans [LY93], avec une méthode basée sur une décomposition en martingale de la configuration. Puis a été résolu le cas où ψ est une perturbation bornée d'un potentiel quadratique, dans [LPY02] et [Cha03], avec la même méthode. Le cas de l'inégalité de Poincaré a finalement été résolu dans [Cap03]. Une autre méthode a été récemment proposée dans [BM13], qui utilise le fait que, sous des hypothèses de minoration de la courbure, une concentration gaussienne suffisamment forte implique une inégalité de Sobolev logarithmique. Cette méthode permet de retrouver le résultat de [Cap03] pour l'inégalité de Poincaré, mais ne s'applique qu'aux potentiels quadratiques dans le cas de l'inégalité de Sobolev logarithmique.

Les travaux de cette thèse sont basés sur une autre stratégie, proposée dans [GOVW09], qui a permis une preuve alternative dans le cas où ψ est une perturbation bornée d'un potentiel quadratique. Elle est basée sur une décomposition de la configuration dans l'espace des phases : une composante macroscopique, qui est l'image de la configuration par une projection dans un espace de dimension plus petite, et une composante de fluctuations par rapport à cette configuration macroscopique. La méthode développée dans [GOVW09] permet de déduire l'inégalité de Sobolev logarithmique pour une mesure de probabilités à partir d'une inégalité pour la loi de la projection macroscopique et pour la loi des fluctuations. C'est cette méthode qui a ensuite été adaptée par Menz et Otto dans [MO13] pour obtenir la réponse suivante à la question de Varadhan :

Théorème 1.8 ([MO13]). *Supposons que ψ soit de la forme $\psi = \psi_c + \delta\psi$, où ψ_c est une fonction uniformément convexe, et $\delta\psi$ est une fonction bornée, de dérivée bornée. Alors la mesure canonique $\mu_{N,m}$ vérifie une inégalité de Sobolev logarithmique avec une constante indépendante de N et m .*

En particulier, contrairement aux résultats antérieurs, ce théorème s'applique dans le cas où la dérivée seconde du potentiel ψ n'est pas bornée, et notamment au cas du puits double quartique $\psi(x) = (x^2 - 1)^2$.

Je vais maintenant donner une présentation de la méthode à deux échelles telle qu'elle est utilisée dans [GOVW09].

Dans un cadre abstrait, l'opérateur de projection est un opérateur linéaire

$$P : X \longrightarrow Y$$

où X est l'espace des configurations microscopiques, et Y l'espace des configurations macroscopiques, tous deux munis d'une structure euclidienne. On suppose que cet opérateur vérifie la relation

$$PNP^t = \text{id}_Y$$

pour un certain réel N qui mesure la taille du système, et qui sera le nombre de particules dans les applications que l'on considèrera. Typiquement, on aura $N = \dim X$. Sous cette condition, les opérateurs P^tNP et $\text{id}_X - P^tNP$ sont des projecteurs orthogonaux.

Une mesure de probabilités μ sur X admet alors la décomposition

$$\mu(dx) = \mu(dx|y)\bar{\mu}(dy)$$

où $\bar{\mu} = \mu \circ P^{-1}$ est la loi de $y = Px$ lorsque x est distribué selon μ , et $\mu(dx|y)$ est la loi de x conditionnellement à $Px = y$. On écrit $\bar{\mu}$ sous la forme $\bar{\mu}(dy) = \bar{Z}^{-1} \exp(-N\bar{H}(y))dy$. Cette décomposition a été utilisée dans [GOVW09] pour montrer que, lorsque $\text{Hess } H$ est bornée, si le potentiel macroscopique \bar{H} est uniformément convexe, et si les mesures $\mu(dx|y)$ vérifient une inégalité de Sobolev logarithmique avec une constante uniforme en y , alors la mesure μ vérifie aussi une inégalité de Sobolev logarithmique, avec une constante explicite. La raison pour laquelle on a besoin de supposer que $\text{Hess } H$ est bornée est qu'on utilise la Proposition 1.19 avec $g = \nabla H$.

En particulier, cette méthode a été utilisé dans [GOVW09] pour montrer que les mesures canoniques vérifient une inégalité de Sobolev logarithmique avec une constante uniforme, dans les où ψ est une perturbation bornée d'un potentiel quadratique. Ces mesures sont les

$$\mu_{N,m}(dx) := \frac{1}{Z} \mathbb{1}_{\Phi(x)=m} \exp\left(-\sum_{i=1}^N \psi(x_i)\right) \mathcal{H}^{N-1}(dx)$$

dont le support est l'hyperplan

$$X_{N,m} := \left\{ x \in \mathbb{R}^N; \frac{1}{N} \sum_{i=1}^N x_i = m \right\},$$

que l'on munit du produit scalaire ℓ^2 habituel. Pour cette application, on construit l'opérateur P en séparant les N spins en M boîtes de K spins chacune, avec $N = KM$. Le profil macroscopique $y = Px$ est un élément de \mathbb{R}^M , dont les coordonnées sont les moyennes des spins sur chacune de ces boîtes. Le potentiel macroscopique \bar{H} associé à la mesure macroscopique $\bar{\mu}$ est alors de la forme

$$\bar{H}(y) := \frac{1}{M} \sum_{i=1}^M \psi_K(y_i).$$

Dès que K est suffisamment grand, ψ_K est uniformément convexe, et donc \bar{H} l'est aussi.

Cette méthode a aussi été utilisée dans [Men11] pour le cas où le hamiltonien est de la forme

$$H(x) = \sum_i \psi(x_i) + \epsilon \sum_i m_{i,j} (x_i - x_j)^2$$

avec ϵ suffisamment faible. Elle a ensuite été adaptée dans [Lel09] pour des notions d'état macroscopique plus générales, lorsque l'opérateur P n'est pas linéaire, et dans [MS12] pour étudier l'asymptotique à basse température de la constante de l'inégalité de Sobolev logarithmique pour des potentiels avec plusieurs minimums, en dimension fixe. Toutefois, elle ne peut pas être utilisée telle quelle pour obtenir le théorème 1.8 à cause de la restriction $\text{Hess } H < \infty$. Pour pouvoir contourner cette difficulté, on a besoin de bornes de covariances plus générales que celle de la Proposition 1.19.

Pour contourner cette dernière difficulté, Menz et Otto ont utilisé dans [MO13] une version itérative de la méthode à deux échelles. En utilisant comme opérateur de projection macroscopique la fonction qui a une configuration de N spins associe la configuration de $N/2$ spins obtenue en faisant la moyenne des paires de spins voisins, ils ont obtenu une version restreinte du critère abstrait de [GOVW09] sans avoir à supposer que $\text{Hess } H$ est bornée. L'outil technique principal est une version asymétrique de l'inégalité de Brascamp-Lieb, utilisée pour borner la covariance entre ∇H et la densité f de la loi de la configuration de spins. Comme cette inégalité n'est valable que pour des densités de probabilité sur \mathbb{R} , l'argument de [MO13] ne marche pas pour l'opérateur de projection macroscopique de [GOVW09] dès que la taille des boîtes dépasse deux. Mais pour l'inégalité de Sobolev logarithmique, en appliquant cet argument un nombre fini de fois, ils parviennent à obtenir le théorème 1.8.

Une méthode reliée à cette méthode à deux échelles a été proposée dans [OR07] et généralisée dans [Men13] et [MN13]. Elle utilise également une inégalité de Sobolev logarithmique pour les mesures conditionnelles et pour les marginales associées pour obtenir une inégalité de Sobolev logarithmique globale. Cette méthode a été utilisée pour le cas des mesures sans loi de conservation, mais avec des interactions, reprenant et généralisant des résultats de [BH99], [BH00], [Yos99] et [Zeg96].

Dans le chapitre 2, qui est un travail réalisé en collaboration avec Georg Menz, nous avons utilisé l'inégalité de Brascamp-Lieb asymétrique pour obtenir une nouvelle inégalité de covariance. Cette inégalité nous a ensuite permis de prouver l'inégalité de Sobolev logarithmique pour les potentiels superquadratiques en utilisant la méthode de [GOVW09], sans avoir à introduire plus de deux échelles. Nous avons également généralisé ce résultat au cas où le Hamiltonien est de la forme

$$H(x) = \sum \psi(x_i) + a_i x_i$$

avec les a_i des réels bornés par une constante L . Ces a_i jouent le rôle d'un potentiel chimique additionnel, introduisant une inhomogénéité spatiale dans le système à l'équilibre. Nous avons obtenu une inégalité de Sobolev logarithmique, avec une constante indépendante de la dimension et du spin moyen, mais qui dépend de la borne L . Ceci généralise un résultat de [LNN05], où l'inégalité est démontrée pour un potentiel quadratique et des a_i donnés par des variables aléatoires indépendantes et identiquement distribuées.

Dans le chapitre 5, j'ai démontré que si le potentiel ψ est une perturbation bornée d'un potentiel uniformément p -convexe (notion définie en (1.2)), alors les mesures canoniques vérifient une inégalité de Sobolev logarithmique p -modifiée avec une constante indépendante du nombre de spins et du spin moyen. La preuve consiste en une adaptation de

la méthode itérative de [MO13]. Ce résultat implique également une inégalité transport-entropie.

1.3.2 Un cadre abstrait pour l'étude des limites hydrodynamiques

L'article [GOVW09] utilise aussi cette notion de profil macroscopique pour développer un cadre abstrait pour l'étude des limites hydrodynamiques. Ils ont considéré des dynamiques données par l'équation de Fokker-Planck

$$\frac{\partial}{\partial t}(f\mu) = \nabla \cdot (A\nabla f\mu) \quad (1.3)$$

avec A une matrice définie-positive sur X . $f = f(t, x)$ est une densité de probabilité par rapport à la mesure invariante μ . Le principe de la méthode de [GOVW09] est de montrer que la mesure $f(t)\mu$ est proche, pour une distance de Wasserstein, d'un état déterministe $NP^t\eta(t)$, où η est un profil macroscopique donné par une équation différentielle ordinaire sur Y , de la forme

$$\frac{d}{dt}\eta = -\bar{A}\nabla\bar{H}(\eta). \quad (1.4)$$

Dans le cas où \bar{H} est uniformément convexe et $\text{Hess } H$ est borné, et si A vérifie l'inégalité

$$|(\text{id}_X - NP^tP)x|^2 \leq \frac{C}{M^2}\langle Ax, x \rangle \quad \forall x \in X, \quad (1.5)$$

avec M la dimension de Y , il est prouvé dans [GOVW09] qu'on peut obtenir une inégalité de la forme

$$\frac{d}{dt} \left(\frac{1}{N} W_{2,A^{-1}}(f(t)\mu, \delta_{NP^t\eta(t)})^2 \right) \leq C \left(\frac{M}{N} + \frac{1}{M} \right). \quad (1.6)$$

En intégrant ceci, on obtient

$$\frac{1}{N} W_{2,A^{-1}}(f(t)\mu, \delta_{NP^t\eta(t)})^2 \leq \frac{1}{N} W_{2,A^{-1}}(f(0)\mu, \delta_{NP^t\eta(0)})^2 + CT \left(\frac{1}{M} + \frac{M}{N} \right).$$

On peut ensuite utiliser une telle inégalité pour obtenir des résultats de convergence vers la limite hydrodynamique, comme on le verra dans la section suivante.

Sous une contrainte supplémentaire sur A , il est possible d'améliorer la borne (1.6) en obtenant un terme M^{-2} à la place de M^{-1} . Ce résultat a été obtenu dans [Diz07].

Dans le chapitre 2, qui est un travail réalisé en collaboration avec Georg Menz, nous avons développé une nouvelle inégalité de covariance, qui nous a permis d'éliminer l'hypothèse de borne supérieure sur $\text{Hess } H$, dans le cas de Hamiltoniens de la forme $H(x) = \sum \psi(x_i)$. Ceci nous a permis d'étudier les limites hydrodynamiques dans le cas où ψ se comporte en $|x|^p$ avec $p \geq 2$.

Dans le chapitre 3, je me suis intéressé au comportement de l'entropie $\text{Ent}_\mu(f\mu)$. J'ai démontré une borne de la forme

$$\left| \int_0^T \frac{1}{N} \text{Ent}_\mu(f(t)\mu) - \bar{H}(\eta(t)) dt \right| \leq C \sqrt{\frac{M}{N} + \frac{1}{M} + \frac{1}{N} W_{2,A^{-1}}(f(0)\mu, \delta_{NP^t\eta(0)})^2},$$

ainsi que des bornes sur l'entropie relative de $f(t)\mu$ par rapport à un état de Gibbs local, c'est à dire une mesure de la forme $Z^{-1} \exp(\lambda(t) \cdot x)\mu(dx)$, avec $\lambda(t)$ choisi pour que cette mesure soit une bonne approximation de $\delta_{\eta(t)}$. La proximité en entropie relative à une mesure de Gibbs locale est la notion de convergence utilisée dans la méthode de

l'entropie relative développée dans [Yau91], et est plus forte que la convergence en distance de Wasserstein de [GOVW09].

Dans le chapitre 4, qui est un travail réalisé en collaboration avec Manh Hong Duong, nous avons étendu ce résultat (avec une dépendance exponentielle en T plutôt que linéaire) pour les équations de la forme

$$\frac{\partial}{\partial t}(f\mu) = \nabla \cdot ((A + J)(\nabla f)\mu) \quad (1.7)$$

avec J une matrice antisymétrique, telle que $-J^2 \leq cA$, avec c une constante positive indépendante de N . L'ajout du terme supplémentaire fait que l'opérateur différentiel n'est plus symétrique dans $L^2(\mu)$, mais ne modifie pas la mesure invariante. Ceci étend donc le résultat de [GOVW09] au cas des dynamiques faiblement asymétriques. Le résultat du chapitre 3 sur la convergence de l'entropie reste aussi valable dans ce contexte.

1.3.3 Application au modèle de Ginzburg-Landau

Nous allons maintenant voir comment on peut appliquer les méthodes et résultats décrits dans un cadre abstrait dans la section précédente à un modèle concret.

Les systèmes de spins continus en interaction sont un type de modèle classique de la physique statistique. C'est une généralisation possible des systèmes de spins discrets, comme le modèle d'Ising ou le modèle de Potts, qui a l'avantage de pouvoir être étudiée avec des techniques d'analyse. Ils peuvent être utilisés comme modèles en physique, par exemple pour les matériaux magnétiques ([KPDKT05]) ou comme discrétisation de surfaces aléatoires ([Fun05], [Fun12]). Ils sont également utilisés en physique statistique ([LP76]) et en mécanique quantique ([GRS75]).

On considère un système de N spins continus X^i , dont l'évolution est donné par les équations différentielles stochastiques

$$dX_t^i = -N^2(2\psi(X_t^i) - \psi(X_t^{i+1}) - \psi(X_t^{i-1}))dt + \sqrt{2}N(dB_t^{i+1} - dB_t^i)$$

où les B^i sont des mouvements browniens indépendants, et ψ est une fonction à valeurs réelles, qui croît de manière au moins linéaire à l'infini. Le facteur N correspond à un scaling en temps de la dynamique. Un exemple typique est le potentiel $\psi(x) = (x^2 - 1)^2$, et on a alors un puits double dont les minimums correspondent aux deux valeurs possibles du modèle d'Ising discret.

Comme cette dynamique conserve le spin moyen $m = \frac{1}{N} \sum_{i=1}^N X^i$, pour obtenir de bonnes propriétés d'ergodicité, on se restreint à un hyperplan

$$X_{N,m} := \left\{ x \in \mathbb{R}^N; \frac{1}{N} \sum_{i=1}^N x_i = m \right\},$$

qu'on munit du produit scalaire ℓ^2 habituel.

La dynamique admet alors une unique mesure de probabilité invariante, qui est la mesure canonique

$$\mu_{N,m}(dx) := \frac{1}{Z} \mathbb{1}_{\Phi(x)=m} \exp\left(-\sum_{i=1}^N \psi(x_i)\right) \mathcal{H}^{N-1}(dx)$$

où \mathcal{H}^{N-1} est la mesure de Hausdorff de dimension $N - 1$ sur l'hyperplan $X_{N,m}$.

L'évolution de la loi du système est de la forme (1.3), avec la matrice A donnée par le Laplacien discret, c'est à dire

$$A = N^2 \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & -1 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ -1 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

Comme on l'a vu précédemment, dans le cas où ψ est une perturbation bornée d'un potentiel uniformément convexe, cette mesure vérifie une inégalité de Sobolev logarithmique, avec une constante indépendante de la taille du système N et du spin moyen m .

En combinant l'inégalité de Sobolev logarithmique et l'inégalité de Poincaré discrète, si on note $f(t)\mu$ la loi de X_t au temps t , on obtient

$$\frac{d}{dt} \text{Ent}_{\mu_N}(f(t)\mu) = - \int \frac{\langle A\nabla f, \nabla f \rangle}{f} d\mu \leq -C \text{Ent}_{\mu}(f(t)\mu).$$

En conséquence, le système converge exponentiellement vite vers l'équilibre, uniformément en la taille du système et en le spin moyen.

Pour étudier la limite d'échelle de cette dynamique lorsque $N \rightarrow +\infty$, il nous faut plonger les espaces $X_{N,m}$ dans un même espace. Ici, on peut les plonger dans l'espace $L^2(\mathbb{T})$. A toute configuration de spins $x \in \mathbb{R}^N$, on associe une fonction en escalier sur le tore \bar{x} , définie par

$$\bar{x}(\theta) = x_i \quad \forall \theta \in \left[\frac{i-1}{N}, \frac{i}{N} \right). \quad (1.8)$$

On munit l'espace obtenu de la norme H^{-1} . L'adhérence de la réunion des espaces $X_{N,m}$ pour cette norme est l'espace de Sobolev H^{-1} .

Si on applique les résultats abstraits présentés à la section précédente, on peut montrer que si la suite des conditions initiales $f_{0,N}\mu_N$ vérifie

$$\int \|\bar{x} - \zeta_0\|_{H^{-1}}^2 f_{0,N}(x) \mu_N(dx) \rightarrow 0$$

alors, pour tout $T > 0$,

$$\int \|\bar{x} - \zeta(t)\|_{H^{-1}}^2 f_N(t, x) \mu_N(dx) \rightarrow 0$$

où ζ est la solution d'une équation de la chaleur non-linéaire, de la forme

$$\frac{\partial}{\partial t} \zeta = \frac{\partial^2}{\partial \theta^2} \varphi'(\rho)$$

avec

$$\varphi(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \log \int \exp(\sigma x - \psi(x)) dx \right)$$

Ce résultat avait déjà été prouvé dans [GPV88]. La méthode a deux échelles permet toutefois d'obtenir un résultat quantitatif, avec une vitesse de convergence en $N^{-1/2}$. Le résultat de [GOVW09] est toutefois restreint au cas où ψ est de la forme $|x|^2 + \delta(x)$, avec δ une perturbation bornée. La raison pour laquelle on peut passer des bornes en distance $W_{2,A^{-1}}$ à une convergence en norme H^{-1} est que, lorsque N est grand,

$$\frac{1}{N} \langle A^{-1}x, x \rangle \approx \|\bar{x}\|_{H^{-1}}^2.$$

Dans le chapitre 2, qui présente des résultats obtenus en collaboration avec Georg Menz, on étend le résultat de [GOVW09] au cas où ψ est de la forme $|x|^p + \delta(x)$, avec $p \geq 2$. On étudie aussi l'effet d'un potentiel chimique aléatoire, en étudiant le cas où le système est donné par les équations différentielles stochastiques

$$dX_t^i = -N^2 \left[2\psi(X_t^i) + 2a_i - \psi(X_t^{i+1}) - a_{i+1} - \psi(X_t^{i-1}) - a_{i-1} \right] dt + \sqrt{2}N(dB_t^{i+1} - dB_t^i)$$

où les a_i sont des variables aléatoires i.i.d. bornées, qui jouent le rôle de potentiels chimiques. On montre alors que pour presque toutes les réalisations du champ de potentiels chimiques aléatoires, il y a convergence vers la limite hydrodynamique, qui est donnée par l'équation aux dérivées partielles

$$\frac{\partial}{\partial t} \zeta = \frac{\partial^2}{\partial \theta^2} \tilde{\varphi}'(\rho).$$

avec

$$\tilde{\varphi}(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \mathbb{E} \left[\log \int \exp(\sigma x - a_1 x - \psi(x)) dx \right] \right).$$

Dans le chapitre 3, on applique le théorème abstrait sur la convergence en entropie relative au modèle de Ginzburg Landau. On y obtient la convergence

$$\int_0^T \frac{1}{N} \text{Ent}_{\mu_N}(f(t)\mu_N) dt \longrightarrow \int_0^T \int \varphi(\zeta(t, \theta)) d\theta - \varphi \left(\int \zeta(t, \theta) d\theta \right) dt,$$

c'est à dire la convergence de l'énergie libre microscopique vers l'énergie libre hydrodynamique. J'ai également montré que cette convergence a lieu pour tout $t > 0$, uniformément sur les intervalles de la forme $[\epsilon, T]$. Ce résultat avait déjà été prouvé dans [Kos01a] avec une autre méthode, qui ne permet pas d'obtenir des bornes quantitatives, mais qui a l'avantage de s'appliquer aussi à des systèmes discrets comme les processus d'exclusion [Kos01b].

Dans le chapitre 4, travail réalisé en collaboration avec Manh Hong Duong, on applique les résultats sur les dynamiques de la forme (1.7) pour obtenir la limite hydrodynamique lorsque

$$J = \frac{N}{2} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & -1 \\ -1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & -1 & 0 \end{pmatrix}.$$

Ce choix correspond à une perturbation faiblement asymétrique de la dynamique de Kawasaki étudiée précédemment. La limite hydrodynamique est alors donnée par l'équation aux dérivées partielles

$$\frac{\partial}{\partial t} \zeta = \frac{\partial^2}{\partial \theta^2} \varphi'(\rho) + \frac{\partial}{\partial \theta} \varphi'(\rho).$$

On y étudie également la convergence de l'entropie.

1.4 Flots gradients dans les espaces de mesures

Un flot gradient dans l'espace \mathbb{R}^d est une équation différentielle ordinaire de la forme

$$\dot{x}_t = -\nabla F(x_t). \tag{1.9}$$

Cette équation est celle qui, étant donné un point de départ x_0 et une énergie $F : \mathbb{R}^d \rightarrow \mathbb{R}$, nous donne le chemin $(x_t)_{t \in \mathbb{R}_+}$ qui fait décroître « le plus vite possible » $F(x_t)$. Ceci n'est pas exactement vrai (il suffit de considérer l'équation $\dot{x}_t = -\lambda \nabla F(x_t)$ avec $\lambda > 1$), mais vient de l'interprétation suivante. Si on discrétise en temps cette équation avec un schéma d'Euler implicite avec pas de temps $\Delta t > 0$, on obtient le schéma

$$x_{n+1} = x_n - \Delta t \nabla F(x_{n+1}).$$

Si la dérivée seconde de F est minorée, et si Δt est suffisamment petit, alors x_{n+1} est l'unique minimiseur de la fonctionnelle

$$x \rightarrow F(x) + \frac{1}{2\Delta t} d(x, x_n)^2. \quad (1.10)$$

x_{n+1} est donc le point qui minimise l'énergie F , sous la contrainte de ne pas trop s'éloigner de x_n .

De nombreux systèmes peuvent être modélisés par des équations de la forme (1.9). On peut par exemple penser à l'évolution de la vitesse d'une particule amortie, ou celle d'un fluide visqueux le long d'une pente.

Le second principe de la thermodynamique nous dit que l'entropie (physique) S d'un système isolé est une fonction croissante du temps. Par conséquent, on s'attend à ce que l'énergie libre $F = U - TS$ (où U est l'énergie interne et T est la température) soit une fonction décroissante du temps, au moins dans le cas d'un système fermé subissant une transformation isotherme.

Ici, je m'intéresse à l'entropie relative par rapport à une mesure $\mu = \exp(-H)dx$. Pour une mesure $\nu = f dx$, l'entropie relative s'écrit alors

$$\text{Ent}_\mu(\nu) = \int f \log f dx + \int H(x) \nu(dx).$$

La quantité $\int f \log f dx$ est l'opposé de l'entropie physique, et la quantité $\int H(x) \nu(dx)$ peut être vue comme une énergie interne. L'entropie relative peut donc être vue comme une énergie libre du système, au moins pour des situations où la température ne joue aucun rôle (ce qui sera le cas des systèmes considérés dans cette thèse). Si on avait voulu tenir compte de la température, on aurait remplacé H par βH , où β est l'inverse de la température, et l'énergie libre aurait alors été $\beta^{-1} \text{Ent}_\mu(\nu)$.

Par abus de langage, et pour suivre la convention habituelle en mathématiques, ce que j'appellerai entropie dans la suite de ce texte est cette énergie libre. Cette entropie est donc une fonction *décroissante* du temps.

Une conséquence du second principe de la thermodynamique est que les états d'équilibre possible d'un système isolé sont ceux dont l'entropie physique est maximale. Ce principe est très utile en physique statistique à l'équilibre, et conduit à chercher les mesures invariantes possibles d'une dynamique en minimisant une fonctionnelle thermodynamique, comme l'énergie libre.

Il s'avère en fait qu'on peut aussi comprendre certains phénomènes de la physique statistique *hors équilibre* à partir de principes de minimisation d'une fonctionnelle. Cette idée est apparue en physique dans les travaux de Prigogine. Puisque le système tend vers l'état de moindre entropie, on peut chercher la *trajectoire* que suit le système comme celle qui va le plus vite possible vers cet état. Ceci nous mène à chercher à définir une équation de la forme

$$\dot{\nu} = -\nabla \text{Ent}_\mu(\nu).$$

Mais si une telle équation a un sens dans \mathbb{R}^d , on ne peut pas vraiment définir un gradient dans $\mathcal{P}(\mathbb{R}^d)$. On va donc chercher à donner un sens à cette équation dans un cadre plus général.

Une première manière de faire est de remarquer que la formulation (1.10) demande seulement l'existence d'une distance, et la semi-convexité de la fonctionnelle F . On va donc pouvoir utiliser cette formulation pour construire des suites d'éléments de manière itérative et avec un pas de temps $\Delta t > 0$, et les voir comme une discrétisation en temps d'une équation continue. Si, dans la limite $\Delta t \rightarrow 0$, il existe une courbe limite, alors on peut définir les flots gradients dans des espaces métriques comme les limites de tels schémas. Cette idée est attribuée à De Giorgi, et a été introduite dans [DGMT80]. Elle a été utilisée dans le cadre des espaces de mesures équipés de la distance W_2 dans [JKO98] et [AGS08].

Une deuxième manière de faire, qui est en quelque sorte la version intégrée en temps de la formulation précédente, est de remarquer que, pour toute courbe de $\mathcal{C}^1([0, T], \mathbb{R}^d)$, on a

$$\begin{aligned} F(x_T) - F(x_0) &= \int_0^T \dot{x}_t \cdot \nabla F(x_t) dt \\ &\geq -\frac{1}{2} \int_0^T |\dot{x}_t|^2 dt - \frac{1}{2} \int_0^T |\nabla F(x_t)|^2 dt \end{aligned} \quad (1.11)$$

avec égalité si et seulement si $t \rightarrow x_t$ est une solution de (1.9). On peut donc définir les solutions de l'équation différentielle ordinaire comme les seules courbes annulant la fonctionnelle positive

$$J((x_t)_{0 \leq t \leq T}) := F(x_T) - F(x_0) + \frac{1}{2} \int_0^T |\dot{x}_t|^2 dt + \frac{1}{2} \int_0^T |\nabla F(x_t)|^2 dt. \quad (1.12)$$

Utiliser cette formulation ne demande pas de définir le gradient ∇F , mais seulement sa longueur $|\nabla F|$, ce qu'on parviendra à faire dans le cadre métrique. C'est ce deuxième point de vue qu'on va plutôt utiliser ici.

Dans le cas où F est semi-convexe, on peut donner une formulation infinitésimale un peu plus simple de cette notion de flots gradients, qui a par exemple été utilisée dans [Vil09] pour les flots gradients de l'entropie dans les espaces de mesures.

1.4.1 Définition

Je vais maintenant donner la définition des flots gradients. Pour cette présentation, je ne vais pas être complètement rigoureux, et passer sous silence un certain nombre d'aspects techniques. En particulier, je ne donnerai pas les conditions nécessaires pour que la construction que je vais présenter fonctionne rigoureusement. Pour une description rigoureuse et complète, on pourra se référer à [AGS08].

Comme vu précédemment, pour définir les flots gradients, il nous faut donner un sens aux quantités $|\dot{x}_t|$ et $|\nabla F(x)|$. Ceci n'est pas faisable pour n'importe quelle courbe, il faut qu'elle vérifie une propriété de continuité. La bonne notion s'avère être celle de courbe absolument continue :

Définition 1.28. Une courbe $t \rightarrow x_t$ sur un intervalle $[0, T]$, à valeurs dans un espace métrique (E, d) , est dite absolument continue si il existe une fonction $g \in L^1([0, T], \mathbb{R}_+)$ telle que, pour tout $0 \leq s \leq t \leq T$, on a

$$d(x_s, x_t) \leq \int_s^t g(r) dr.$$

On notera $AC([0, T], E)$ l'ensemble des courbes absolument continues.

Pour de telles courbes, on peut définir $|\dot{x}_t|$:

Définition 1.29. Soit (E, d) un espace métrique, et $t \rightarrow x_t$ une courbe de $AC([0, T], E)$.

$$|\dot{x}_t| := \limsup_{h \rightarrow 0^+} \frac{d(x_t, x_{t+h})}{h}.$$

Il nous reste encore à définir l'analogue de $|\nabla F(x)|$. Dans le cas de \mathbb{R}^d , on aurait pu utiliser

$$|\nabla F(x)| = \limsup_{y \rightarrow x} \frac{|F(y) - F(x)|}{|y - x|}.$$

Dans le cadre métrique, cette définition devient :

Définition 1.30. On définit la longueur du gradient $|\nabla F|$ de la fonctionnelle $F : E \rightarrow \mathbb{R}$ comme

$$|\nabla F|(x) := \limsup_{y \rightarrow x} \frac{(F(y) - F(x))_+}{d(x, y)}.$$

Dans les bonnes conditions, on peut alors définir les flots gradients comme les courbes absolument continues qui annullent la fonctionnelle

$$J((x_t)_t) := F(x_T) - F(x_0) + \frac{1}{2} \int_0^T |\dot{x}_t|^2 dt + \frac{1}{2} \int_0^T |\nabla F|(x_t)^2 dt. \quad (1.13)$$

On peut alors utiliser cette définition dans $\mathcal{P}_2(\mathbb{R}^d)$ muni de la distance W_2 et pour la fonctionnelle Ent_μ , avec μ une mesure absolument continue par rapport à la mesure de Lebesgue, et avec une densité $\mu(dx) = \exp(-H(x))dx$. On obtient des courbes de mesures de probabilité ν_t , qui s'avèrent être les solutions faibles de l'équation

$$\frac{\partial}{\partial t} \nu_t = \Delta \nu_t - \nabla H \cdot \nabla \nu_t.$$

Plus généralement, on peut travailler sur une variété riemannienne, avec une mesure absolument continue par rapport au volume, et on obtient le même résultat (en prenant la définition appropriée du laplacien et du gradient).

Dans ce cadre, la longueur du gradient $|\nabla F|$ est l'information de Fisher relative à μ , et $|\dot{\nu}|$ est la norme H^{-1} sur l'espace $L^2(\nu)$. La fonctionnelle J est donnée par

$$J((\nu_t)) = \text{Ent}_\mu(\nu_T) - \text{Ent}_\mu(\nu_0) + \frac{1}{2} \int_0^T I_\mu(\nu_t) dt + \frac{1}{2} \int_0^T \|\dot{\nu}_t\|_{H^{-1}(\nu_t)}^2 dt. \quad (1.14)$$

Dans le cadre de l'espace des mesures de probabilités sur \mathbb{R}^d (ou sur une variété Riemannienne), on dispose d'un autre point de vue (équivalent au précédent) sur les flots gradients, dû à Otto. Il consiste à utiliser la formule de Benamou-Brenier pour munir formellement cet espace de mesures d'une structure de variété riemannienne de dimension infinie. La valeur de $W_2(\nu_0, \nu_1)$ donnée par la formule de Benamou-Brenier peut être vue comme la minimisation de l'action le long des courbes reliant ν_0 à ν_1 , ce qui correspond bien à la formulation des distances riemanniennes habituelles. On pourra consulter [Ott01] et [OV00] pour une présentation de ce point de vue.

1.4.2 Stabilité des flots gradients

Un des intérêts de cette notion de flots gradients est qu'elle est stable : lorsqu'on a une suite de flots gradients (pour des métriques et fonctionnelles d'énergie données) qui converge vers un objet limite, on s'attend à ce que cet objet limite soit lui-même un flot gradient. De plus, on s'attend à ce que la métrique et la fonctionnelle d'énergie du flot gradient limite apparaissent elles-mêmes comme limites des suites de métriques et de fonctionnelles d'énergie.

Un exemple important d'application de ce principe est la stabilité des équations de la chaleur sur des suites d'espaces métriques convergeant au sens de Gromov-Hausdorff, prouvée dans [AGS14].

Ce principe n'est pas vrai en général. Un contre-exemple est donné par l'homogénéisation des équations d'évolution périodiques sur \mathbb{R}^2 . On considère une fonction $A : \mathbb{R}^2 \rightarrow \mathcal{S}^{+,2}(\mathbb{R})$, qu'on suppose \mathbb{Z}^2 -périodique. Alors les solutions de l'équation aux dérivées partielles

$$\frac{\partial f}{\partial t}(x, t) = \operatorname{div}(A(x/\epsilon)\nabla f(x, t)) \quad (1.15)$$

convergent lorsque ϵ tend vers 0 vers une solution de l'équation

$$\frac{\partial f}{\partial t}(x, t) = \operatorname{div}(\bar{A}\nabla f(x, t)) \quad (1.16)$$

pour une certaine matrice constante \bar{A} . L'équation (1.15) est le flot gradient de l'entropie (par rapport à la mesure de Lebesgue) pour la distance W_2 associée à la distance riemannienne donnée par le tenseur métrique $x \rightarrow A^{-1}(x/\epsilon)$. Cette distance riemannienne sur \mathbb{R}^2 converge vers une certaine distance limite, qui en général n'est pas riemannienne, mais finslerienne (ce résultat peut être trouvé dans [BD98]). Toutefois, l'équation (1.16) *n'est pas* en général le flot gradient de l'entropie pour cette distance limite. En effet, l'équation aux dérivées partielles homogénéisée est linéaire, alors que le flot gradient de l'entropie pour une norme non-euclidienne sur \mathbb{R}^2 ne l'est pas. On ne peut donc pas toujours se contenter d'étudier le comportement asymptotique des métriques et des fonctionnelles d'énergie pour prouver une convergence.

Toutefois, on peut donner des cadres ou des critères pour que ce principe fonctionne. Par exemple, dans [ASZ09], la stabilité des équations de Fokker-Planck est démontrée, dans le cas où les drifts sont toujours le gradient d'une fonction convexe. Cette hypothèse entraîne la convexité de l'entropie, et empêche le phénomène qui se produit pour l'homogénéisation périodique.

Dans l'article [Ser11], Serfaty a proposé un théorème abstrait permettant d'identifier la limite d'une suite de flots gradients. Ce théorème généralise des idées développées dans [SS04]. Informellement, il dit que, étant donné une suite de distances d_n et une suite de fonctionnelles F_n , si il existe une fonctionnelle F et une distance d telles que pour toute suite convergente $(x^n)_n$, de limite x , on a

$$\liminf F(x_n) \geq F(x)$$

et si pour toute suite de courbes absolument continues $x^n(t)$ convergeant vers une courbe $x(t)$ (pour une topologie sur l'espace des courbes continues donnée) on a

$$\liminf \int_0^T |\dot{x}^n(t)|_n^2 dt \geq \int_0^T |\dot{x}(t)|^2 dt$$

et

$$\liminf \int_0^T |\nabla F_n|_n^2(x^n(t)) dt \geq \int_0^T |\nabla F|^2(x(t)) dt$$

alors pour toute suite convergente de flots gradients dont les données initiales sont bien préparées (c'est à dire que $F_n(x^n(0))$ converge vers $F(x(0))$), la limite est un flot gradient pour la fonctionnelle F et la distance d .

Comme la convergence vers un objet limite fait partie des hypothèses, il doit être utilisé conjointement à un résultat de compacité pour la suite de flots gradients.

On peut aussi mentionner les travaux [AMP⁺12] et [PSV10], qui utilisent des idées similaires pour étudier un modèle de réactions chimiques.

Le cadre abstrait de [Ser11] peut être utilisé pour prouver la convergence vers la limite hydrodynamique du modèle de Ginzburg-Landau, retrouvant ainsi le résultat de [GPV88], sous l'hypothèse additionnelle que les conditions initiales convergent en entropie relative (mais comme on l'a vu, cette condition supplémentaire devient immédiatement vérifiée si on attend un temps $\epsilon > 0$).

1.4.3 Flots gradients et grandes déviations

Dans cette section, je vais discuter des liens entre les flots gradients dans les espaces de mesures et les grandes déviations. Pour commencer, je vais donner une brève introduction aux grandes déviations. Pour plus de détails, on pourra se référer à [DZ10], [FK06] ou [Var08].

Le but d'étudier les grandes déviations est d'étudier les probabilités d'événements très rares. Typiquement, on cherche à prouver des estimées de la forme

$$\mathbb{P}[X_n \in A] \approx \exp(-c_n I(A))$$

où (X_n) est une suite de variables aléatoires, (c_n) une suite de réels positifs allant à l'infini, A est un borélien et I une fonction positive. La probabilité $\mathbb{P}[X_n \in A]$ est donc exponentiellement petite, et on cherche à quantifier à quel point.

Définition 1.31. Soit $(a_n)_n$ une suite croissante de réels strictement positifs, qui tend vers $+\infty$, et $I : X \rightarrow \mathbb{R}_+$ une fonction semi-continue inférieurement, telle que les ensembles de niveau $\{x \in X, I(x) \leq c\}$ soient compacts pour tout $c \in \mathbb{R}$.

Une suite de mesures de probabilité $(\nu_n)_n$ vérifie un principe de grandes déviations si, pour tout ensemble fermé F de X , on a

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \nu_n(F) \leq - \inf_{x \in F} I(x)$$

et si, pour tout ensemble ouvert O , on a

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log \nu_n(O) \geq - \inf_{x \in O} I(x).$$

Une conséquence d'un principe de grandes déviations pour une suite (X_n) est que, si la fonction de taux admet un unique minimiseur x_0 , alors la suite (X_n) converge en probabilité vers x_0 .

Imaginons maintenant qu'on dispose d'une suite de processus stochastiques $(X_n(t))$, qui vérifie un principe de grandes déviations sur un espace fonctionnel (typiquement, $\mathcal{C}([0, T], \mathbb{R}^d)$). Alors, dans la limite, les trajectoires vont se concentrer avec une forte probabilité autour des trajectoires $t \rightarrow x_t$ qui minimisent la fonction de taux I . Les "bonnes" trajectoires apparaissent donc comme les minimiseurs d'une fonctionnelle. Les liens entre grandes déviations et minimisation de fonctionnelles thermodynamiques apparaissent dans les travaux de Jona-Lasinio et ses collaborateurs, pour l'étude des fluctuations et des

grandes déviations du courant dans les systèmes de particules dont la dynamique n'est pas réversible ([BDSG⁺01], [BDSG⁺02], [BLM09]).

Ce point de vue présente des similarités avec la formulation (1.12) des flots gradients. Dans les deux cadres, les bonnes trajectoires sont les minimiseurs d'une certaine fonctionnelle. On peut donc se demander si il y a des liens entre ces deux notions.

Les travaux [ADPZ11], [ADPZ13] et [DLR13] ont cherché à lier la formulation flot gradient du flot de marginales de processus stochastiques, et le principe de grandes déviations pour la moyenne empirique d'un grand nombre de copies indépendantes de ce processus. Il s'avère que, dans de nombreux cas, la fonction de taux et la fonctionnelle J apparaissant dans la formulation (1.12) sont reliées.

Sur les liens entre transport optimal et grandes déviations, il faut aussi mentionner [GL07], qui utilise des arguments de grandes déviations comme le théorème de Sanov pour démontrer des inégalités de transport-entropie. On retrouve aussi des arguments de transport optimal et de flots gradients dans le chapitre 9 de [FK06], pour l'étude des diffusions avec interaction faible en champ moyen.

Dans le chapitre 6, j'ai cherché à voir si on pouvait utiliser la notion de flots gradients pour *prouver* des principes de grandes déviations pour des suites de processus stochastiques. Il s'est avéré qu'il y a un lien fort entre grandes déviations et flots gradients, au moins dans le cas des processus de diffusion réversibles, c'est à dire les solutions d'équations différentielles stochastiques de la forme

$$dX_t = -A(X_t)\nabla H(X_t)dt + \operatorname{div}(A)(X_t)dt + \sqrt{2A(X_t)}dB_t$$

où A est une fonction à valeurs dans l'espace des matrices symétriques, définies positives, et B est un mouvement brownien standard sur \mathbb{R}^d . Ce processus laisse invariante la mesure $\mu(dx) = \exp(-H(x))dx$. Le flot des marginales d'un tel processus est un flot gradient, pour la fonctionnelle $\nu \rightarrow \operatorname{Ent}_\mu(\nu)$, et la distance de Wasserstein $W_{2,A^{-1}}$ associée à la distance riemannienne donnée par le tenseur $x \rightarrow A^{-1}(x)$.

On considère une suite (X_t^n) de diffusions de cette forme, avec des conditions initiales X_0^n de loi $\mathbb{P}_{0,n}$, des coefficients de diffusion A_n et des drifts donnés par des fonctions H_n . On autorise aussi la dimension à changer, et on la notera d_n . On supposera qu'on s'est donné une manière d'injecter tous les espaces \mathbb{R}^{d_n} dans un même espace métrique, séparable et complet X , et on munit $\mathcal{C}([0, T], X)$ d'une topologie qui le rend métrique, séparable et complet.

Le flot des marginales de la diffusion X^n est un flot gradient, auquel on peut donc associer une fonctionnelle de la forme (1.12).

On supposera que la suite des lois de nos processus est exponentiellement tendue, avec une vitesse (a_n) . La définition de cette notion est la suivante :

Définition 1.32. Une suite de mesures de probabilité (ν_n) sur un espace topologique X est dite exponentiellement tendue, avec vitesse (a_n) , si pour tout $\alpha > 0$ il existe un compact K_α tel que

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \nu_n(K_\alpha^c) \leq -\alpha.$$

Cette notion revient à demander qu'asymptotiquement la masse à l'infini est exponentiellement petite.

Le théorème que j'ai obtenu est le suivant :

Théorème 1.9. *La suite des lois des processus (X_t^n) sur $[0, T]$ vérifie un principe de grandes déviations de vitesse (a_n) et avec une fonction de taux I ssi elle est exponentiellement tendue avec vitesse (a_n) et si*

(i) Pour toute suite de lois sur $\mathcal{C}([0, T], X)$ convergeant faiblement vers la mesure de Dirac en une trajectoire $\rho \in \mathcal{C}([0, T], X)$, si on note $\nu_{n,t}$ le flot de marginale de la n -ième loi, on a

$$\liminf \frac{1}{a_n} \text{Ent}_{\mathbb{P}_{0,n}}(\nu_{n,0}) + \frac{1}{2a_n} J_n((\nu_{n,t})) \geq I(\rho);$$

(ii) Pour toute trajectoire $\rho \in \mathcal{C}([0, T], X)$, il existe une suite de lois sur $\mathcal{C}([0, T], X)$ convergeant faiblement vers la mesure de Dirac en ρ telle que

$$\lim \frac{1}{a_n} \text{Ent}_{\mathbb{P}_{0,n}}(\nu_{n,0}) + \frac{1}{2a_n} J_n((\nu_{n,t})) = I(\rho).$$

La condition (i) correspond à la borne supérieure dans le principe de grandes déviations, et la condition (ii) à la borne inférieure. L'ajout du terme $\text{Ent}_{\mathbb{P}_{0,n}}(\nu_{n,0})$ intervient pour retranscrire les grandes déviations pour la suite de conditions initiale (X_0^n) . Dans le cadre d'une application concrète de ce principe, on sait en général trouver des conditions initiales pour lesquelles un tel principe de grandes déviations est vérifié. Par exemple, pour l'application aux grandes déviations par rapport à la limite hydrodynamique pour le modèle de Ginzburg–Landau, un bon choix est donné par des mesures de Gibbs locales bien choisies.

Ce théorème permet donc de ramener le principe de grandes déviations à l'étude des fonctionnelles J_n intervenant dans la formulation en flot gradient des flots de marginales des équations différentielles stochastiques.

Ce résultat est une généralisation et reformulation d'une méthode utilisée dans [DG87] pour étudier les grandes déviations de la mesure empirique de mouvements browniens indépendants. On trouvera également une présentation de ce résultat dans la deuxième partie de [Föl88]

1.4.4 Grandes déviations pour deux variantes du modèle de Ginzburg–Landau

Pour des modèles dont on a prouvé la limite hydrodynamique, une question naturelle est de voir ensuite à quel point une réalisation typique du système à N particules est proche de la limite. Les bornes quantitatives de [GOVW09] apportent une réponse possible à cette question. Une autre manière de voir le problème serait d'étudier les fluctuations autour de la limite. Un tel résultat est l'analogue du Théorème Central Limite dans ce contexte, et a été établi dans [CY92] pour le modèle de Ginzburg–Landau.

Une troisième manière de faire est d'étudier les grandes déviations par rapport à la limite hydrodynamique. Typiquement, de tels résultats montrent que la probabilité d'observer une déviation significative est exponentiellement petite en la taille du système. Un principe de grandes déviations a été établi pour le modèle de Ginzburg–Landau dans [DV89]. On peut également mentionner les résultats sur les grandes déviations pour le processus d'exclusion obtenu dans [QRV99], et les grandes déviations pour les gaz sur réseau de [QY98]. Pour une revue du sujet, on pourra consulter le chapitre 10 de [KL99].

Dans le chapitre 6, j'ai étudié les grandes déviations par rapport à la limite hydrodynamique pour deux variantes du modèle de Ginzburg–Landau : une version avec des conductances données par des variables aléatoires i.i.d., et une version introduite par Varadhan dans [Var93] dont la conductance entre deux spins voisins est fonction de la valeur de ces deux spins. Ce dernier modèle est dit non-gradient, car les courants entre les sites ne peuvent pas s'écrire comme le gradient discret d'une fonction de la configuration (contrairement au modèle standard présenté dans la section 1.3.3). Les grandes déviations pour ce

deuxième modèle avaient déjà été obtenues dans [Qua95] avec la méthode proposée dans [DV89]. La méthode que j'ai employée a été d'utiliser le critère abstrait présenté dans la section 1.4.3.

La dynamique du modèle avec conductances aléatoires est donnée par le système d'équations différentielles stochastiques

$$dX_t^i = N^2 a_{i+1} (\psi'(X_t^{i+1}) - \psi'(X_t^i)) dt + N^2 a_i (\psi'(X_t^{i-1}) - \psi'(X_t^i)) dt + N a_{i+1} dB_t^{i+1} - N a_i dB_t^i$$

où les a_i sont des variables iid bornées. J'ai aussi supposé que la condition initiale suit la loi de densité

$$\exp\left(\sum \varphi'(\zeta_0(i/N)) x_i\right) \mu_{N,m}(dx)$$

pour une fonction lisse ζ_0 de moyenne m .

Pour ce modèle, j'ai montré qu'on obtient pour presque toutes les réalisations des variables a_i un principe de grandes déviations de vitesse N et avec la fonction de taux

$$I(\rho) := \int \varphi(\rho_0) - \varphi(\zeta_0) - \varphi'(\zeta_0)(\rho_0 - \zeta_0) d\theta \\ + \frac{1}{4\bar{a}} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \bar{a} \frac{\partial^2}{\partial \theta^2} \varphi'(\rho) \right\|_{H^{-1}}^2 dt$$

où \bar{a} est la moyenne harmonique de a_1 , c'est à dire $\bar{a}^{-1} := \mathbb{E}[a_1^{-1}]$.

Le modèle non-gradient est donné par le système d'EDS

$$dX_t^i = N^2 a(X_i, X_{i+1}) (\psi'(X_t^{i+1}) - \psi'(X_t^i)) dt + N^2 a(X_{i-1}, X_i) (\psi'(X_t^{i-1}) - \psi'(X_t^i)) dt \\ + N^2 \partial_1 a(X_i, X_{i+1}) dt - N^2 \partial_2 a(X_{i-1}, X_i) dt + N a(X_i, X_{i+1}) dB_t^{i+1} - N a(X_{i-1}, X_i) dB_t^i$$

où a est une fonction de deux variables de classe C^1 , bornée par deux constantes strictement positives, et $\partial_1 a$ et $\partial_2 a$ sont ses dérivées partielles.

Por ce modèle, avec les même conditions initiales que précédemment, on retrouve le principe de grandes déviations de [Qua95], de vitesse N et de bonne fonction de taux

$$I(\rho) := \int \varphi(\rho_0) - \varphi(\zeta_0) - \varphi'(\zeta_0)(\rho_0 - \zeta_0) d\theta \\ + \frac{1}{4} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \frac{\partial}{\partial \theta} \left(\hat{a}(\rho) \frac{\partial}{\partial \theta} \varphi'(\rho) \right) \right\|_{H^{-1}(\hat{a}(\rho))}^2 dt$$

où

$$\|g\|_{H^{-1}(\hat{a}(\rho))}^2 := \sup_h 2 \int g(\theta) h(\theta) d\theta - \int \hat{a}(\rho(\theta)) \left(\frac{\partial}{\partial \theta} h \right)^2 d\theta$$

avec le sup sur l'ensemble des fonctions h de classe C^1 et de moyenne nulle sur le tore. Dans ces définitions, \hat{a} est une certaine fonction d'une seule variable, bornée, donnée par une formulation variationnelle, et qui dépend à la fois de a et de ψ .

L'intérêt de ce deuxième modèle est qu'il constitue une adaptation des arguments à une situation riemannienne : c'est le flot gradient de l'entropie, mais pour la distance riemannienne sur \mathbb{R}^N associée au tenseur $x \rightarrow A^{-1}(x)$. La preuve que j'ai donné de ce résultat a pour intérêt de faire apparaître le coefficient \hat{a} dans le problème de minimisation de l'information de Fisher. Ceci nous permet de comprendre l'apparition du coefficient de diffusion comme étant celui qui *minimise la dissipation d'entropie*. Ceci coïncide bien avec l'interprétation en flot gradient de la dynamique : pour faire décroître l'entropie (mathématique) le moins vite possible, on veut suivre des trajectoires pour lesquelles la dissipation d'entropie est la plus petite possible. Ceci correspond bien à l'intuition de Prigogine, qui est d'identifier la bonne trajectoire comme celle minimisant la production d'entropie physique (et donc la dissipation d'entropie mathématique).

1.5 Discrétisation d'équations différentielles stochastiques

Au cours de cette thèse, je me suis également intéressé au problème de la simulation numérique d'équations différentielles stochastiques (EDS)

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \quad (1.17)$$

où B est un mouvement brownien (sur \mathbb{R}^d ou \mathbb{T}^d). Le but est l'évaluation numérique de quantités de la forme

$$\mathbb{E}(f(X_t))$$

ou

$$\mathbb{E} \left[\int_0^T f(X_t)dt \right],$$

où f est une fonction à valeurs réelles. Pour évaluer de telles quantités il nous faut tout d'abord choisir une manière de discrétiser en temps cette EDS, avec un pas de temps Δt petit. Ceci nous permet de définir des variables aléatoires $X_n^{\Delta t}$ qui approximent X_t lorsque $\Delta t \rightarrow 0$ et $n\Delta t \rightarrow t$. Ensuite, pour évaluer $\mathbb{E}(f(X_t))$, il nous faut générer un grand nombre N de copies indépendantes $X_n^{\Delta t, i}$ de cette variable aléatoire. D'après la loi des grands nombres, si le nombre de copies tend vers l'infini, on a presque sûrement

$$\frac{1}{N} \sum_{i=1}^N f(X_n^{\Delta t, i}) \rightarrow \mathbb{E}(f(X_n^{\Delta t})).$$

Lorsqu'on utilise cette méthode, on a deux sources d'erreur numérique :

- L'erreur de discrétisation $\mathbb{E}(f(X_t)) - \mathbb{E}(f(X_n^{\Delta t}))$, due au fait que le schéma numérique discrétisé qu'on utilise ne simule pas exactement X_t , et qui dépend du pas de temps Δt et du schéma numérique employés. Si le schéma est bien adapté, cette erreur tend vers 0 lorsque Δt tend vers 0.
- L'erreur statistique $\mathbb{E}(f(X_n^{\Delta t})) - \frac{1}{N} \sum_{i=1}^N f(X_n^{\Delta t, i})$, qui est une variable aléatoire dépendant du nombre de copies N qu'on utilise. D'après le théorème central limite, cette erreur est typiquement d'ordre $N^{-1/2}$.

Pour réduire l'erreur a priori, il nous faut donc choisir un pas de temps Δt tel que l'erreur de discrétisation soit suffisamment faible, et déterminer le nombre de copies N nécessaire pour que l'erreur statistique ait une forte probabilité d'être suffisamment faible.

On distingue deux types d'erreur de discrétisation :

- L'erreur forte, définie comme

$$\mathbb{E} \left[\sup_{k \leq T/\Delta t} |X_k^{\Delta t} - X_{k\Delta t}|^2 \right]^{1/2}.$$

Cette définition demande qu'on se soit donné un couplage entre la solution de l'EDS et son approximation.

- L'erreur faible, associée à une fonction f à valeurs réelles et suffisamment régulière donnée, définie comme

$$\left| \mathbb{E}[f(X_T)] - \mathbb{E}[f(X_n^{\Delta t})] \right|,$$

avec $T = n\Delta t$.

L'erreur forte donne un contrôle trajectorielle, tandis que l'erreur faible est une évaluation à un temps donné. De plus, pour obtenir une petite erreur forte, il faut un couplage tel que la réalisation du schéma soit proche de la réalisation de la solution de l'EDS, tandis que l'erreur faible est une espérance, et ne dépend donc des lois des deux processus. Lorsque l'erreur se comporte en Δt^α , on parle d'erreur d'ordre α . Pour une fonction f lipschitzienne, l'ordre de l'erreur forte est toujours plus petit que l'ordre de l'erreur faible.

Dans cette thèse, je me suis essentiellement intéressé à deux algorithmes : le schéma d'Euler explicite, et l'algorithme de Metropolis-Hastings.

1.5.1 Erreur statistique pour le schéma d'Euler explicite

Le schéma d'Euler explicite associé à l'équation (1.17) est la chaîne de Markov $(X_n^{\Delta t})$ construite itérativement de la manière suivante : si la variable aléatoire $X_n^{\Delta t}$ a pris la valeur $x \in \mathbb{R}^d$, on génère une variable aléatoire U_{n+1} , qui suit une loi normale centrée réduite, indépendante de tout ce qui a été fait précédemment, et on pose

$$X_{n+1}^{\Delta t} := X_n^{\Delta t} - \Delta t b(n\Delta t, X_n^{\Delta t}) + \sqrt{\Delta t} \sigma(n\Delta t, X_n^{\Delta t}) U_{n+1}.$$

Cet algorithme est connu sous le nom de schéma d'Euler explicite, ou encore schéma d'Euler-Maruyama. C'est la généralisation naturelle du schéma d'Euler classique utilisé pour discrétiser en temps les équations différentielles ordinaires.

Lorsque les coefficients b et σ sont lipschitziens, si l'on fait tendre le pas de discrétisation Δt vers 0, la loi de ce schéma converge bien vers celle de l'EDS (1.17). Dans [TT90], il a été prouvé que l'erreur faible de discrétisation de ce schéma est d'ordre 1.

Pour l'erreur de discrétisation, dans le cas où σ est borné, il a été prouvé dans [FM12] que la loi de $X_n^{\Delta t}$ vérifie une inégalité de transport-entropie dont la constante ne dépend de n et Δt qu'à travers $t = n\Delta t$ (ce qui signifie qu'elle est stable par passage à la limite $\Delta t \rightarrow 0$). Cette inégalité implique un phénomène de concentration gaussienne, avec une constante explicite, qui nous permet de borner a priori l'erreur statistique, et donc de donner des intervalles de confiance pour la valeur de $\mathbb{E}[f(X_t)]$ obtenue en prenant la moyenne sur des copies indépendantes de X_t , lorsque f est une fonction lipschitzienne.

Les résultats du chapitre 7 constituent une extension du résultat de [FM12] au cas où σ n'est plus borné. On ne peut plus s'attendre à obtenir de la concentration gaussienne, mais on peut obtenir des inégalités de transport-entropie plus faibles que T_1 , et qui permettent encore d'obtenir des intervalles de confiance pour $\mathbb{E}[f(X_t)]$. Le résultat principal est le suivant :

Théorème 1.10. *Supposons que les coefficients b et σ de l'équation différentielle stochastique soient des fonctions Lipschitziennes en espace, uniformément en temps. Supposons de plus qu'il existe une fonction $V : \mathbb{R}^d \rightarrow \mathbb{R}_+^*$ de classe C^2 , de hessienne bornée et telle qu'il existe une constante $\alpha \in [1/2, 1]$ et deux constantes positives C_b et C_σ telles que pour tout $x \in \mathbb{R}^d$ et $t \in [0, T]$ on ait*

$$|b(t, x)|^2 \leq C_b V(x) \quad \text{Tr}(\sigma\sigma^*(t, x)) \leq C_\sigma V^{1-\alpha}(x).$$

Alors, pour Δt suffisamment petit, la loi du schéma d'Euler au temps $T = n\Delta t$ vérifie une inégalité transport-entropie T_{1, Φ_α^} , avec Φ_α^* une fonction qui dépend explicitement des paramètres et de T , mais pas de n et Δt . La fonction $\Phi^*(r)$ se comporte de manière quadratique pour r petit, et comme $Cr^{2\alpha}$ pour r grand.*

Dans ce résultat, on peut voir α comme un paramètre mesurant la croissance du coefficient de diffusion σ à l'infini. C'est ce paramètre qui donne le comportement à l'infini de la fonction Φ_α^* impliquée dans l'inégalité transport-entropie.

Une exemple de diffusions auxquelles ce résultat s'applique est celui des processus de la forme

$$dX_t = \lambda X_t dt + \left(1 + |X_t|^2\right)^{(1-\alpha)/2} dB_t.$$

Ce résultat implique les bornes de concentration suivantes :

Corollaire 1.33. *Sous les hypothèses précédentes, si les $X_T^{\Delta t, k}$ sont des variables aléatoires iid obtenues en utilisant le schéma d'Euler explicite, on a les bornes*

$$\mathbb{P}_x \left[\left| \frac{1}{M} \sum_{k=1}^M f(X_T^{\Delta t, k}) - \mathbb{E}[f(X_T^{\Delta t})] \right| \geq r \right] \leq 2 \exp(-M\Phi_\alpha^*(r));$$

où f est une fonction 1-lipschitzienne, et

$$\mathbb{P}_x \left[\sup_{f \text{ 1-lip}} \left(\frac{1}{M} \sum_{k=1}^M f(X_T^{\Delta t, k}) - \mathbb{E}[f(X_T^{\Delta t})] \right) \geq \frac{C(T, \Delta t)}{M^{1/(d+2)}} + r \right] \leq \exp(-M\Phi_\alpha^*(r)).$$

Des résultats similaires sont également obtenus pour une famille d'algorithmes d'approximation stochastiques, utilisés pour estimer numériquement les zéros d'une fonction continue sur \mathbb{R}^d , dont les valeurs sont coûteuses à obtenir (en terme de nombre de calculs à réaliser). Ils n'ont pas été repris dans cette thèse, mais peuvent être trouvés dans l'article [FF13].

1.5.2 Erreur de discrétisation pour l'algorithme de Metropolis-Hastings

On considère maintenant le cadre plus restreint des équations différentielles stochastiques de la forme

$$dq_t = -\nabla V(q_t)dt + \sqrt{2}dB_t. \quad (1.18)$$

Comme on l'a déjà vu, les diffusions de cette forme sont des processus de Markov réversibles, et convergent vers une unique mesure de probabilité invariante

$$\mu(dx) = \frac{1}{Z} \exp(-V(x))dx,$$

au moins lorsque V croît suffisamment vite vers $+\infty$ à l'infini.

Dans le cas où V n'est pas globalement lipschitzienne, il s'avère que le schéma d'Euler explicite fonctionne mal. Au lieu d'être stable en temps long, ses moments partent à l'infini. L'une des sources de ce problème est que, contrairement au processus lui-même, cette approximation n'est pas une chaîne de Markov réversible. Pour pouvoir éviter ce phénomène d'explosion, il est donc naturel de chercher un moyen de modifier le schéma d'Euler explicite pour le rendre réversible. Une manière possible de faire ceci est d'utiliser l'algorithme de Metropolis-Hastings pour stabiliser le schéma.

L'algorithme de Metropolis-Hastings est une manière de construire une chaîne de Markov réversible ayant une mesure de probabilité invariante spécifiée. Elle a été introduite pour la première fois dans [MRR⁺53]. Elle est souvent utilisée pour échantillonner la mesure π , par exemple dans le cadre d'une méthode de Monte Carlo. On pourra consulter [DSC98] et les références qui y sont incluses pour plus d'information sur l'algorithme de Metropolis et ses applications.

La construction de la chaîne (X_n) se fait de la manière suivante : on suppose qu'on dispose d'une mesure de probabilité π sur un espace X , que l'on va pour l'instant supposer discret. On suppose aussi qu'on dispose d'un ensemble de mesures de probabilités

$T(x, \cdot)$ sur X , de préférence faciles à simuler. À partir de là, on définit les probabilités d'acceptation

$$R(x, y) := \frac{T(y, x)\pi(y)}{T(x, y)\pi(x)} \quad (1.19)$$

Alors, conditionnellement à la valeur de X_n , on génère une proposition Y_{n+1} , de loi $T(X_n, \cdot)$, et une variable aléatoire U_{n+1} , de loi uniforme sur $[0, 1]$. On considère alors deux possibilités :

- Si $U_{n+1} \leq R(X_n, Y_{n+1})$, on pose $X_{n+1} := Y_{n+1}$;
- Sinon, on pose $X_{n+1} := X_n$.

Dans le cas où la chaîne de Markov est à valeurs dans un espace continu, on remplace juste T et π par leurs densités par rapport à une mesure de référence (typiquement, la mesure de Lebesgue si on se place sur \mathbb{R}^d).

On remarque que, pour construire cette chaîne de Markov, on peut se contenter de connaître π à une constante multiplicative près. Ceci est très pratique lorsqu'on utilise cet algorithme pour simuler une mesure de la forme $\mu(dx) = \frac{1}{Z} \exp(-V(x))dx$ lorsqu'on connaît le potentiel V , pour laquelle il peut parfois être très difficile de calculer efficacement la fonction de partition Z . C'est pour cette raison que cet algorithme est souvent utilisé en physique statistique.

On peut maintenant appliquer cette méthode au schéma d'Euler explicite. Le taux d'acceptation est donné par

$$R_{\Delta t}(x, y) := \frac{e^{-V(y)}T_{\Delta t}(y, x)}{e^{-V(x)}T_{\Delta t}(x, y)},$$

où $T(x, y)$ est la probabilité de transition de x à y du schéma d'Euler explicite, donnée par

$$T_{\Delta t}(x, y) := \left(\frac{1}{4\pi}\right)^{d/2} \exp\left(-\frac{1}{4\Delta t}|y - x + \Delta t \nabla V(x)|^2\right).$$

Ce schéma est connu dans la littérature sous le nom de SmartMC, ou Metropolis-adjusted Langevin algorithm (MALA). Comme la chaîne de Markov obtenue est réversible, la dynamique est stable, et converge vers la mesure invariante μ . Toutefois, comme cette modification est sans relation avec la dynamique, elle affecte en mal l'erreur de discrétisation. Par exemple, il a été montré dans [BRVE09] que l'erreur de discrétisation trajectorielle forte à l'équilibre, donnée par

$$\mathbb{E}^\mu \left[\mathbb{E}^x \left[|X_{[t/\Delta t]}^{\Delta t} - X_t|^2 \right] \right]^{1/2},$$

se comporte en $C(t)\Delta t^{3/4}$. Ceci est à comparer à l'erreur en Δt pour le schéma d'Euler explicite avec drift lipschitzien.

Dans le chapitre 8, je me suis intéressé à l'erreur de discrétisation dans le calcul du coefficient de transport, défini par

$$\mathcal{D} := \lim_{t \rightarrow \infty} \mathbb{E} \left[\frac{X_t - X_0}{\sqrt{t}} \otimes \frac{X_t - X_0}{\sqrt{t}} \right].$$

Ce coefficient est également donné par la formule

$$\mathcal{D} = 2 \left(\text{Id} - \int_0^{+\infty} \mathbb{E}[\nabla V(q_t) \otimes \nabla V(q_0)] dt \right).$$

Les résultats du chapitre 8, obtenus en collaboration avec Ahmed-Amine Hommann et Gabriel Stoltz montrent que lorsqu'on utilise l'algorithme MALA pour simuler une dynamique de la forme (1.18) sur le tore, l'erreur sur l'évaluation du coefficient de transport faite en discrétisant l'une des deux formules précédentes est en $\Delta t \mathcal{D}^* + O(\Delta t^{3/2})$.

Dans le chapitre 9, qui présente un résultat obtenu en collaboration avec Ahmed-Amine Hommann et Gabriel Stoltz, on a cherché à modifier l'algorithme pour obtenir un schéma avec une meilleure erreur de discrétisation trajectorielle forte à l'équilibre. On a montré qu'en ajoutant des perturbations faibles au schéma d'Euler explicite, puis en utilisant l'algorithme de Metropolis-Hastings, on obtient un schéma avec une erreur forte en Δt (plutôt que $(\Delta t)^{3/4}$). A noter qu'il n'est pas possible de faire un schéma avec une meilleure erreur forte (voir par exemple l'article de revue [MGR06]). Ce schéma est obtenu en appliquant l'algorithme de Metropolis à la chaîne de Markov donnée par

$$X_{n+1}^{\Delta t} = X_n^{\Delta t} - \Delta t \nabla V(X_n^{\Delta t}) + \frac{\Delta t^2}{6} \left(-\text{Hess } V(X_n^{\Delta t}) \nabla V(X_n^{\Delta t}) + \nabla \Delta V(X_n^{\Delta t}) \right) \\ + \sqrt{2\Delta t} \left(\text{Id} + \frac{\Delta t}{3} \text{Hess } V(X_n^{\Delta t}) \right)^{-1/2} U_{n+1}.$$

Pour tous ces résultats, l'élément important est l'étude du comportement du taux d'acceptation $R_{\Delta t}$ lorsque Δt tend vers 0. Pour l'algorithme MALA standard, il se comporte en $1 - c\Delta t^{3/2}$. Pour pouvoir obtenir des meilleures bornes d'erreur, il nous faut réussir à augmenter le taux d'acceptation, pour mieux coller à la dynamique continue. Le schéma du chapitre 9 a un taux d'acceptation en $1 - c\Delta t^{5/2}$, et c'est ce qui permet d'obtenir une erreur forte trajectorielle d'ordre 1.

1.6 Liste des publications issues de la thèse

Voici la liste des articles (publiés, acceptés pour publication ou soumis) repris dans les différents chapitres de cette thèse

- [Fat13b] : A two-scale approach to the hydrodynamic limit part II : local Gibbs behavior, *ALEA Lat. Am. J. Probab. Math. Stat.* 10 (2013), no. 2, 625–651.
- [FF13] : Transport-Entropy inequalities and deviation estimates for stochastic approximation schemes, avec N. Frikha, *Electronic Journal of Probability*, Volume 18 (2013), no. 67, 1-36.
- [FHS] : Error analysis of the transport properties of Metropolized schemes, avec A.–A. Hommann et G. Stoltz, accepté pour publication à ESAIM : Proceedings (2014).
- [Fat13a] : Modified logarithmic Sobolev inequalities for canonical ensembles, soumis (2013).
- [DF14] : The two-scale approach for non-reversible dynamics, avec M. H. Duong, soumis (2014).
- [FM14] : Hydrodynamic limit for conservative spin systems with super-quadratic single-site potential, avec G. Menz, soumis (2014).
- [Fat14] : A gradient flow approach to large deviations for diffusion processes, soumis (2014).

Notation

- \mathcal{L}^k is the k -dimensional Lebesgue measure
- ∇ is the gradient, Hess stands for Hessian, $|\cdot|$ is the norm and $\langle \cdot, \cdot \rangle$ is an inner product. If necessary, a subscript will indicate the space on which these are taken.
- A^t is the transpose of the operator A .
- $\text{Ran}(A)$ is the range of the operator A .
- $\text{Ent}_\mu(f) = \int f(\log f)\mu - (\int f\mu) \log(\int f\mu)$ is the entropy of the positive function f with respect to the measure μ .
- $\text{osc } \Psi = \sup_X \Psi - \inf_X \Psi$ is the oscillation of the function Ψ .
- Z is a constant enforcing unit mass for a probability measure.
- C is a positive constant, which may vary from line to line, or even within a line.¹
- id_X is the identity map $X \rightarrow X$.
- LSI is an abbreviation for logarithmic Sobolev inequality.
- $\|f\|_{\infty, E} = \sup_{x \in E} |f(x)|$ is the supremum norm of f over the set E .
- $a \lesssim b$ means that there is a constant C such that $a \leq Cb$.
- $\mathcal{C}^{k, \ell}(A \times B)$ is the space of real-valued functions on $A \times B$ which are C^k with respect to the first variable and C^ℓ with respect to the second variable.
- $\Pi(\mu, \nu)$ is the set of all couplings of μ and ν , that is the probability measures on the product space whose first marginal μ and whose second marginal is ν .
- $AC([0, T], E)$ is the set of absolutely continuous curves $x : t \in [0, T] \rightarrow x_t \in E$, with E some metric space.

1. Only in this text. Elsewhere, C is for cookie.

Première partie

**Functional inequalities and
hydrodynamic limits for
interacting spin systems**

Chapitre 2

Hydrodynamic limit for conservative spin systems with super-quadratic, partially inhomogeneous single-site potential.

Ce chapitre a été écrit en collaboration avec Georg Menz, et est tiré de l'article [FM14].

Abstract We consider an interacting unbounded spin system with conservation of the mean spin. We derive quantitative rates of convergence to the hydrodynamic limit provided the single-site potential is a bounded perturbation of a strictly convex function with polynomial growth, and with an additional random inhomogeneous linear term. The argument adapts the two-scale approach of Grunewald, Otto, Villani and Westdickenberg from the quadratic to the general case. The main ingredient is the derivation of a covariance estimate that is uniform in the system size. We also show that this covariance estimate can be used to change the iterative argument of [MO13] for deducing the optimal scaling LSI for the canonical ensemble into a two-scale argument in the sense of [GOVW09]. We also prove the LSI for canonical ensembles with an inhomogeneous linear term.

2.1 Introduction

In this work, we are interested in generalizing the results of [GOVW09] to a larger setting. There are two topics we are interested in : logarithmic Sobolev inequalities and quantitative rates of convergence to the hydrodynamic limit for continuous spin systems on a lattice. We use the two-scale approach introduced in [GOVW09], which is based on coarse-graining arguments.

2.1.1 Logarithmic Sobolev inequalities

Definition 2.1. A probability measure μ on a Riemannian manifold X is said to satisfy a logarithmic Sobolev inequality with parameter $\rho > 0$ if, for all smooth, non-negative

functions $f : X \rightarrow \mathbb{R}_+$, we have

$$\text{Ent}_\mu(f) \leq \frac{1}{\rho} \int \frac{|\nabla f|^2}{f} \mu.$$

This inequality can be very useful in the study of diffusion processes. It was originally introduced by Gross in [Gro75] to study hypercontractivity phenomenon for diffusion processes. For example, it is well known that such an inequality implies exponentially fast convergence to equilibrium for the diffusion equation naturally associated with μ (namely $\Delta - \nabla V$, when $\mu = \exp(-V)$). Due to tensorization properties (see Criterion I in Appendix B), it also behaves well in large dimensions. For example, it has been proven in [Goz09] that it is equivalent to dimension-free Gaussian concentration. A well-written introduction to this topic can be found in [Led01], and we recall the three most common criterion that are used to derive such inequalities in Appendix B.

In [GOVW09], a new criteria, based on a decomposition of the system in two scales, is devised. It is then applied to prove the LSI for the canonical ensemble with a bounded perturbation of a quadratic Ginzburg-Landau type potential. Due to a technical restriction on a covariance estimate they used, their criteria could not be directly applied to the case of a superquadratic Ginzburg-Landau type potential. To overcome this problem, the two-scale strategy was modified into an iterative approach in [MO13]. In this paper, we will devise a new covariance estimate (Proposition 2.4), which will allow us to apply the original two-scale approach. We also obtain the LSI when we add an inhomogeneous linear term to the Hamiltonian, which generalizes to the superquadratic case results of [Men11] and [LNN05].

2.1.2 Hydrodynamic limits

Our second theme of interest is the derivation of scaling limits for stochastic interacting spin systems. Typically, such results consist in proving that, under a suitable time-space scaling, random systems with a large number of particles behave like a deterministic object, given as the solution of a partial differential equation.

There are two main general methods in the literature to prove hydrodynamic limit results for stochastic systems. The first, often called the entropy method, was introduced in [GPV88], and is based on martingale methods and entropy estimates. The second, called the relative entropy method, was devised in [Yau91], and is based on proving a Gronwall-type inequality for the relative entropy with respect to a well-chosen time-varying reference state. It yields stronger results than the entropy method, and is in general simpler, but relies on stronger assumptions, namely convergence of the initial data in the sense of relative entropy rather than of the macroscopic observables. Both methods are reviewed in [KL].

In [GOVW09], a new method is presented, based on estimates in Wasserstein distances. Its assumptions as well as its results, are at an intermediate level between the two classical methods, and uses a logarithmic Sobolev inequality. It has the additional interest of giving explicit quantitative rates of convergence. This method was then applied to Kawasaki dynamics with an equilibrium measure that is a bounded perturbation of a quadratic potential. Here we will generalize their results to the case where the single-site potential is a bounded perturbation of a uniformly convex function with polynomial growth, and with an additional inhomogeneous linear term, with coefficients given by bounded iid random variables. The additional linear term corresponds to the effect of a random chemical potential. It seems that convergence to the hydrodynamic limit when the Hamiltonian includes such a random chemical potential is entirely new.

The plan of the paper is as follows : In Section 2, we present the framework and our main results. Section 3 is devoted to the new covariance estimate. Section 4 presents the proof of the logarithmic Sobolev inequality, and Section 5 contains the proofs of our results on hydrodynamic limits.

2.2 Framework and Main Results

2.2.1 Microscopic and macroscopic scales

We consider a (periodic) lattice spin system of N continuous variables governed by a Ginzburg-Landau type potential $\psi : \mathbb{R} \rightarrow \mathbb{R}$ and an additional inhomogeneous linear term given by a family of real numbers $a = (a_i)_{1 \leq i \leq N}$. The grand canonical measure on \mathbb{R}^N has density

$$\frac{d\mu_{N,a}}{d\mathcal{L}^N}(x) = \frac{1}{Z} \exp \left(- \sum_{i=1}^N \psi(x_i) + a_i x_i \right). \quad (2.1)$$

We shall assume that the potential ψ is of class C^2 and is a bounded perturbation of a strictly convex potential, that is

$$\psi(x) = \psi_c(x) + \delta\psi(x); \quad \psi_c'' \geq \lambda > 0; \quad \|\delta\psi\|_{C^2} < +\infty. \quad (2.2)$$

We shall also assume that

$$\sup |a_i| \leq L < +\infty. \quad (2.3)$$

To simplify notations, we define the Hamiltonian

$$H(x) := \sum_{i=1}^N \psi(x_i) + a_i x_i - \log Z, \quad (2.4)$$

so that $\mu(dx) = \exp(-H(x))dx$.

We will add to the situation a constraint of fixed mean spin. Our phase state space will be

$$X_{N,m} := \left\{ x \in \mathbb{R}^N, \frac{1}{N} \sum_{i=1}^N x_i = m \right\}.$$

where m is an arbitrary real number. This space is a hyperplane of \mathbb{R}^N with a fixed mean constraint. We endow this space with the ℓ^2 inner product

$$\langle x, \tilde{x} \rangle_X = \sum_{i=1}^N x_i \tilde{x}_i. \quad (2.5)$$

For a given $m \in \mathbb{R}$, we consider the restriction $\mu_{N,a,m}$ of the grand canonical measure to $X_{N,m}$, that is

$$\frac{d\mu_{N,a,m}}{d\mathcal{L}^{N-1}}(x) = \frac{1}{Z} 1_{(1/N) \sum x_i = m} \exp \left(- \sum_{i=1}^N \psi(x_i) + a_i x_i \right). \quad (2.6)$$

This measure is called the canonical ensemble. It gives the distribution of the random variables x_i conditioned on the event that their mean value is given by m .

We now consider two integers M and K such that $N = KM$. To each microscopic profile $x \in X$, we associate a macroscopic profile $y \in \mathbb{R}^M$, given by

$$y_i := \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} x_j. \quad (2.7)$$

We denote by $P_{N,K}$ the operator that associates the macroscopic profile y to the microscopic profile x . The macroscopic profiles live in the space

$$Y_{M,m} := \left\{ y \in \mathbb{R}^M, \frac{1}{M} \sum_{i=1}^M y_i = m \right\}, \quad (2.8)$$

which we endow with the L^2 inner product

$$\langle y, \tilde{y} \rangle_Y := \frac{1}{M} \sum y_i \tilde{y}_i. \quad (2.9)$$

Remark 2.2. Note that we are *not* using the same scalar product for both the microscopic and the macroscopic spaces.

If we denote by $X_{N,K,y}$ the space of microscopic profiles such that $Px = y$, we have the factorization

$$X_{N,K,y} = \bigotimes_{i=1}^M X_{M,y_i}, \quad (2.10)$$

and $P_{N,K}$ factorizes on the fibers $X_{N,K,y}$.

An important feature of the coarse-graining operator $P_{N,K}$ is the relation

$$PNP^t = id_Y, \quad (2.11)$$

so that the symmetric operator P^tNP is the orthogonal projection of X to $(\text{Ker } P)^\perp$. This induces a decomposition of X into macroscopic variables in $Y \approx (\text{Ker } P)^\perp$ and microscopic fluctuations in $\text{Ker } P$, as $x = P^tNPx + (\text{id} - P^tNP)x$. This decomposition of variables induces a decomposition of measures and a decomposition of gradients :

(i) a decomposition of measures : We denote by $\bar{\mu}$ the image of the measure μ by P . This measure is the distribution of the macroscopic variable when the microscopic variable is distributed according to μ . Let $\mu(dx|y)$ be the distribution of the microscopic variable conditioned on the event $\{Px = y\}$. For each y , $\mu(dx|y)$ is a probability measure, and we have

$$\mu(dx) = \mu(dx|y)\bar{\mu}(dy). \quad (2.12)$$

We can define a macroscopic Hamiltonian associated to $\bar{\mu}$ by

$$\bar{H}(y) := -\frac{1}{N} \log \left(\frac{d\bar{\mu}}{dy} \right), \quad (2.13)$$

so that $\bar{\mu}(dy) = \exp(-N\bar{H}(y))dy$. We can explicitly compute \bar{H} as

$$\begin{aligned} \bar{H}(y) &= -\frac{1}{N} \left(\log \int_{\{Px=y\}} \exp(-H(x)) dx - \log \bar{Z} \right) \\ &= -\frac{1}{N} \sum_{i=1}^M \log \int_{X_{K,y_i}} \exp \left(- \sum_{j=(i-1)K+1}^{iK} \psi(x_j) + a_j x_j \right) dx - \frac{1}{N} \log \bar{Z}. \end{aligned}$$

If we define

$$\psi_{K,a_1..a_K}(m) := -\frac{1}{K} \log \int_{X_{K,m}} \exp(-\sum a_i x_i + \psi(x_i)) dx, \quad (2.14)$$

we get

$$\bar{H}(y) = \frac{1}{M} \sum_{i=1}^M \psi_{K,a_{(i-1)K+1}..a_{iK}}(y_i) - \frac{1}{N} \log \bar{Z}. \quad (2.15)$$

Also, the fluctuations measure $\mu(dx|y)$ on $X_{N,K,y}$ can be decomposed as a product measure along the factorization (2.10)

$$\mu_{N,a,m}(dx|y) = \bigotimes_{i=1}^M \mu_{K,(a_j)_{(i-1)K+1 \leq j \leq iK}, y_i} \quad (2.16)$$

(ii) a decomposition of gradients : If f is a smooth function on X , we can decompose its gradient into a macroscopic gradient and a fluctuation gradient :

$$\nabla^{\text{macro}} f(x) = NP^t P \nabla f(x), \quad \nabla^{\text{fluct}} f(x) = (id_X - NP^t P) \nabla f(x). \quad (2.17)$$

This decomposition makes sense, since $(id_X - NP^t P)$ is the orthogonal projection onto $\ker P$, which is the tangent space to the fiber $\{Px = y\}$.

An important use of the introduction of a macroscopic scale will be that, for large enough K , \bar{H} will be uniformly convex :

Proposition 2.3. *Assume that (2.2) and (2.3) hold. There exists an integer K_0 and a positive real number λ such that, for any $K \geq K_0$ and $M \in \mathbb{N}$, we have for any $y, \tilde{y} \in Y$,*

$$\langle \text{Hess } \bar{H}(y) \tilde{y}, \tilde{y} \rangle_Y \geq \lambda |\tilde{y}|_Y^2.$$

This result is a consequence of a local Cramèr theorem, which we prove in Appendix A.

2.2.2 A new covariance estimate

To explain our results, let us first define the covariance of two functions.

Definition 2.4. The covariance of a non-negative function f and a vector-valued function g with respect to a measure μ is given by

$$\text{cov}_\mu(f, g) = \int gf \, d\mu - \left(\int f \, d\mu \right) \left(\int g \, d\mu \right).$$

In [GOVW09], the following covariance estimate to bound some crucial terms which appeared because of the interaction between the microscopic and macroscopic scales.

Proposition 2.5. *Let μ be a probability measure on X satisfying LSI(ρ) for some $\rho > 0$. Then for any two Lipschitz functions $f : X \rightarrow \mathbb{R}_+$ and $g : X \rightarrow \mathbb{R}$, we have*

$$|\text{cov}_\mu(f, g)| \leq \|\nabla g\|_{L^\infty} \sqrt{\frac{2}{\rho} \int f \, d\mu \text{Ent}_\mu(f)}.$$

To be usefully applied, this proposition requires g to have a bounded gradient. Since this result was used in [GOVW09] with $g = \nabla H$, this restricted the method of [GOVW09] to the case where the function ψ they used as a single-site potential has bounded Hessian, and therefore it did not apply to the superquadratic case. For the logarithmic Sobolev inequality, this problem was circumvented in [MO13] through the use of an iterated coarse-graining, but this strategy cannot be transported very well for hydrodynamic limits, where the use of a finite number of scales is crucial. Therefore we devised a new covariance estimate, which applies well to our problem :

Proposition 2.6. *Assume that the single-site potential ψ is perturbed strictly convex in the sense that there is a splitting $\psi = \psi_c + \delta\psi$ such that*

$$\psi_c'' \gtrsim 1 \quad \text{and} \quad |\delta\psi| + |\delta\psi'| \lesssim 1. \quad (2.18)$$

Then the canonical ensemble $\mu_{K,m}$ satisfies for any nonnegative function $f \geq 0$ with $\int f d\mu_{K,m} = 1$ the following covariance estimate :

$$\left| \text{cov}_{\mu_{K,a,m}} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \right|^2 \leq C_0 K \left(\int \frac{\sum_{i=1}^{K-1} \left| \frac{d}{dx_i} f \right|^2}{f} d\mu_{K,a,m} \right)$$

for some constant C_0 which only depends on ψ and L .

Note that, in the case of a quadratic potential, this covariance bound is suboptimal, since it has an extra K factor. However, when used in the proof of the hydrodynamic limit, this extra factor doesn't change the optimal order of magnitude of the quantitative bounds (in $N^{-1/2}$) which is obtained when taking the best possible choices for K and M .

2.2.3 Logarithmic Sobolev inequality

One of our main results is to prove that, when the potential ψ is a bounded perturbation of a strictly convex function, the (generalized) canonical ensemble with inhomogeneous linear part, as defined by (5.7), satisfies a LSI with a parameter ρ that doesn't depend on the dimension N , the mean m or the a_i .

Theorem 2.7. *Assume that ψ is a bounded perturbation of a strictly convex potential, as in (2.2). Let a_1, \dots, a_N be a collection of N real numbers, such that*

$$\sup_i |a_i| \leq L.$$

The measure $\mu_{N,a,m}$ satisfies $LSI(\rho)$ for some constant $\rho > 0$ that is independent of the dimension N , the mean m and the sequence (a_i) . However, it does depend on the bound L .

This result constitutes an extension of the main result of [MO13] (which generalized a result of [LPY02] with a different method), where they did not add an inhomogeneous linear part. This result doesn't rule out a dependence of ρ on L , although it does not in the case of grand canonical measures without a conservation law. The case where ψ is a bounded perturbation of a quadratic potential was already addressed in [Men11] by using the two-scale approach. It should be noted that the assumption of boundedness of the a_i is not used in [Men11], and the constant obtained still does not depend at all on the a_i . Therefore it is possible that our assumption of boundedness on the a_i is not

optimal. However, the local Cramér theorem we use to obtain convexity of the coarse-grained Hamiltonian does not hold for unbounded a_i , so it seems our method cannot be extended to cover that case.

Another approach to this type of result was also proposed in [BM13], but does not obtain the result of [MO13] in full generality. It consists in using curvature lower bounds and Gaussian concentration estimates to recover the LSI.

The logarithmic Sobolev inequality for Kawasaki dynamics (which is weaker than the LSI we prove here, see [Cha03]) when the a_i are bounded iid random variables and for a quadratic single-site potential was also addressed in [LNN05].

2.2.4 Hydrodynamic limit for Kawasaki dynamics

We shall now consider the (reversible) stochastic dynamics on $X_{N,m}$ described by the time-evolution

$$\frac{\partial}{\partial t}(f\mu) = \nabla \cdot (A\nabla f\mu_{N,a,m}), \quad (2.19)$$

which is to be understood in a weak sense : for any smooth test function ξ , we have

$$\frac{d}{dt} \int \xi(x) f(t, x) \mu_{N,a,m}(dx) = - \int \nabla \xi(x) \cdot A\nabla f(t, x) \mu_{N,a,m}(dx).$$

In this equation, the operator A is an $N \times N$ matrix defined by

$$A_{i,j} := N^2(-\delta_{i,j-1} + 2\delta_{i,j} - \delta_{i,j+1}). \quad (2.20)$$

In this definition, δ is the Kronecker symbol and by convention $N + 1 = 1$. A is a symmetric, positive operator when restricted to $X_{N,0}$. We start with an initial condition f_0 such that $f_0\mu$ is a probability measure. Then $f(t, \cdot)\mu$ represents the law at time t of the solution of the stochastic differential equation

$$dX_t = -A\nabla H(X_t)dt + \sqrt{2A}dB_t$$

with initial value distributed according to $f_0\mu$, and where B_t is a standard Brownian motion on \mathbb{R}^N .

Our first result is a quantitative bound on how close a typical realization of the diffusion is to a deterministic vector, whose time-evolution will be given by the following equation :

$$\frac{d\eta}{dt} = -\bar{A}\nabla \bar{H}_{N,K,a}(\eta(t)) \quad (2.21)$$

where the (positive, symmetric) operator \bar{A} on Y is defined through

$$\bar{A}^{-1} = PA^{-1}NP^t. \quad (2.22)$$

Note that η depends on the a_i .

Our main abstract result on the hydrodynamic limit is the following.

Theorem 2.8. *Let us consider a solution of (2.19) in dimension $N = KM$, and a collection a_1, \dots, a_N of real numbers in $[-L, L]$. Let η be a solution of (2.21) with initial condition η_0 , and*

$$\Theta(t) := \frac{1}{2N} \int \langle A^{-1}(x - NP^t\eta(t)), (x - NP^t\eta(t)) \rangle f(t, x) \mu_{N,a,m}(dx), \quad (2.23)$$

Assume that

- (i) $\text{Ent}_{\mu_{N,a,m}}(f_0) \leq C_1 N$;
- (ii) $\bar{H}(\eta_0) \leq C_2$ for some $C_2 > 0$;
- (iii) $\inf \bar{H} \geq -\beta$ for some $\beta > 0$;
- (iv) $\int |x|^2 \mu_{N,a,m}(dx) \leq \alpha N$ for some $\alpha > 0$;
- (v) The coarse-grained Hamiltonian $\bar{H}_{N,K,a}$ is λ -uniformly convex.

Then we have, for any $T > 0$,

$$\begin{aligned} & \max \left\{ \sup_{0 \leq t \leq T} \Theta(t), \frac{\lambda}{2} \int_0^T |y - \eta(t)|_Y^2 \bar{f}(t, y) \bar{\mu}(dy) \right\} \\ & \leq \Theta(0) + T \frac{M}{N} + \frac{\gamma C_0 C_1 K}{2\lambda M^2} + \frac{\gamma^{1/2}}{M} \left(2\alpha + \frac{2C_1}{\rho} \right)^{1/2} (C_1 + C_2 + \beta)^{1/2}, \end{aligned} \quad (2.24)$$

where C_0 is the constant in the covariance estimate of Proposition 2.6.

The statement of Theorem 2.8 can be understood as analog of Theorem 8 in [GOVW09]. Comparing the estimate (2.24) to Theorem 8 in [GOVW09], one sees that instead of the term with scaling $\frac{1}{M^2}$ one gets a term with scaling $\frac{K}{M^2}$, which is a relic of the fact that our covariance estimate has the factor K in contrast to the covariance estimate used in [GOVW09].

Our aim will then be to prove a hydrodynamic limit theorem by applying Theorem 2.8. To do this, we will need to embed the spaces X_N into a single functional space. To simplify notations, we arbitrarily take the mean m to be zero. Everything works fine for any $m \in \mathbb{R}$. With this aim in mind, we identify $X_{N,0}$ with the space \bar{X}_N of piecewise constant functions on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$

$$\bar{X}_N := \left\{ \bar{x} : \mathbb{T} \rightarrow \mathbb{R}; \bar{x} \text{ is constant on } \left(\frac{j-1}{N}; \frac{j}{N} \right], j = 1, \dots, N \right\}.$$

From now on, we will call "the step function associated to $x \in X_N$ " the step function $\bar{x} \in \bar{X}_N$ with

$$\bar{x}(\theta) = x_j, \quad \theta \in \left(\frac{j-1}{N}; \frac{j}{N} \right]. \quad (2.25)$$

Similarly, the "vector associated to $\bar{x} \in \bar{X}_N$ " will denote the vector $x \in X_N$ with

$$x_j = \bar{x}(j/N).$$

It turns out the L^2 norm is not well-adapted to our problem, since it is too sensitive to local fluctuations. Instead, we endow the spaces \bar{X}_N with the H^{-1} norm, which we define as follows : if $f : \mathbb{T} \rightarrow \mathbb{R}$ is a locally integrable function with zero mean, define

$$\|f\|_{H^{-1}}^2 := \int_{\mathbb{T}} w(\theta)^2 d\theta, \quad w' = f, \quad \int_{\mathbb{T}} w(\theta) d\theta = 0. \quad (2.26)$$

As an application of Theorem 2.8, we consider a sequence of bounded iid random variables $(a_q)_{q \in \mathbb{Q}}$, and we define $a_{i,N} = a_{i/N}$ for $i \in \{1, \dots, N\}$. Given a realization of this sequence, we define $\mu_{N,a,0}(dx) = \exp(-H_N(x)) dx$ a probability measure on $X_{N,m}$, where

$$H_N(x) := \sum_{i=1}^N \psi(x_i) + a_{i,N} x_i.$$

We also consider a sequence (F_N) of probability densities on $X_{N,0}$ such that for every N , F_N is absolutely continuous with respect to the Lebesgue measure on $X_{N,0}$. We denote by $f_{0,N,a}$ the density of F_N with respect to $\mu_{N,a,0}$.

We denote by μ_N^0 the generalized canonical ensemble on X_N where the linear part is equal to 0, that is

$$\mu_N^0(dx) = \frac{1}{Z} 1_{(1/N) \sum x_i = m} \exp\left(-\sum \psi(x_i)\right). \quad (2.27)$$

We make the following extra assumption of polynomial growth on ψ_c , which we will need to obtain some a priori uniform L^p estimates on η .

$$c_1(1 + |x|^{p-2}) \leq \psi_c(x) \leq c_2(1 + |x|^{p-2}), \quad p \geq 2. \quad (2.28)$$

This assumption is probably not the most general possible, but covers the physically relevant situations. It should be possible to extend it to an assumption of subexponential growth, but this would require manipulating some Orlicz norm associated to ψ_c rather than the L^p norm we shall use in Section 5 in the energy estimates.

Theorem 2.9. *Assume ψ satisfies (2.2) and (2.28). Given a realization of the random variables (a_q) , let $f = f(t, x)$ be a time-dependent probability density on $(X_N, \mu_{N,a})$, solving*

$$\frac{\partial}{\partial t}(f \mu_{N,a}) = \nabla \cdot (A \nabla f \mu_{N,a})$$

where $f(0, \cdot) = f_{0,a}$ is the density of a probability measure F_N with respect to $\mu_{N,a}$. We assume F_N to be absolutely continuous with respect to the Lebesgue measure, and to satisfy the bound

$$\int_{X_N} \left(\frac{dF_N}{d\mu_N^0} \right) \log \left(\frac{dF_N}{d\mu_N^0} \right) \mu_N^0 \leq CN \quad (2.29)$$

for some constant $C > 0$, where μ_N^0 was defined in (2.27). Assume that

$$\lim_{N \uparrow \infty} \int_{X_N} \|\bar{x} - \zeta_0\|_{H^{-1}}^2 F_N(dx) = 0 \quad (2.30)$$

for some $\zeta_0 \in L^p(\mathbb{T})$ (initial macroscopic profile) with $\int \zeta_0 d\theta = 0$. Then for any $T > 0$ we have, for almost every realization of the field of random variables (a_q) ,

$$\lim_{N \uparrow \infty} \sup_{0 \leq t \leq T} \int_{X_N} \|\bar{x} - \zeta(t, \cdot)\|_{H^{-1}}^2 f(t, x) \mu_{N,a}(dx) = 0 \quad (2.31)$$

where ζ is the unique weak solution of the nonlinear parabolic equation

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} \tilde{\varphi}'(\zeta) \quad (2.32)$$

with initial condition $\zeta(0, \cdot) = \zeta_0$, and where $\tilde{\varphi}$ is defined by

$$\tilde{\varphi}(m) = \sup_{\sigma \in \mathbb{R}} \left\{ \sigma m - \mathbb{E}^a \left[\log \int_{\mathbb{R}} \exp((\sigma - a)x - \psi(x)) dx \right] \right\}. \quad (2.33)$$

In this theorem, the notion of weak solution we have used is the following :

Definition 2.10. We will call $\zeta(t, \theta)$ a weak solution of (2.32) on $[0, T] \times \mathbb{T}$ if

$$\zeta \in L_t^\infty(L_\theta^p), \quad \frac{\partial \zeta}{\partial t} \in L_t^2(H_\theta^{-1}), \quad \tilde{\varphi}'(\zeta) \in L_t^\infty(L_\theta^q); \quad (2.34)$$

and

$$\left\langle \xi, \frac{\partial \zeta}{\partial t} \right\rangle_{H^{-1}} = - \int_{\mathbb{T}} \xi \tilde{\varphi}'(\zeta) d\theta \quad \text{for all } \xi \in L^2, \quad \text{for almost every } t \in [0, T]. \quad (2.35)$$

Remark 2.11. Our method also works (and the proof is much easier) in the case where the a_i are of the form $a_i = a(i/N)$ for some continuous function $a : \mathbb{T} \rightarrow \mathbb{R}$. In such a case, the hydrodynamic equation is

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} (\varphi'(\zeta) - a).$$

2.3 Proof of Proposition 2.6

Our proof of the covariance estimate relies on the following result, which already was an important component of the proof of the LSI in [MO13].

Proposition 2.12 (One-dimensional asymmetric Brascamp-Lieb inequality). *Let $\nu(dx) = Z^{-1} \exp(-\psi(x)) dx$ a probability measure on \mathbb{R} , where $\psi = \psi_c + \delta\psi$ is a bounded perturbation of a strictly convex potential, as in (2.2). Then for any functions f and g , we have*

$$|\text{cov}_\nu(f, g)| \leq \exp(-3 \text{osc } \delta\psi) \sup_x \left| \frac{g'(x)}{\psi_c''(x)} \right| \int |f'| d\nu.$$

We refer to [MO13] for a proof of this result. This covariance estimate was generalized in [CCEL13] to cover other situations, but here this estimate will suffice. Using this inequality, we shall prove the following two lemmas, which will imply Proposition 2.6.

Lemma 2.13. *Let $\mu_{N,a,m}(dx) = Z^{-1} \exp(-\sum \psi(x_i) + a_i x_i) dx$ be a probability measure on $X_{N,m}$, where the single-site potential ψ is strictly convex and the a_i are some real numbers. Then we have, for any non-negative function f such that $\int f d\mu_{N,a,m} = 1$, the following covariance estimate :*

$$\left| \text{cov}_{\mu_{N,a,m}} \left(f, \frac{1}{N} \sum_{i=1}^N \psi'(x_i) \right) \right|^2 \lesssim K \int \frac{\sum \left| \frac{d}{dx_i} f \right|^2}{f} \mu_{N,a,m} \quad (2.36)$$

Lemma 2.14. *Assume that the single-site potential ψ is perturbed strictly convex in the sense of (2.18). Let $\tilde{K} \in \mathbb{N}$ be an arbitrary integer. Then the canonical ensemble $\mu_{\tilde{K},m}$ satisfies for any nonnegative function $f \geq 0$ with $\int f d\mu_{\tilde{K},a,m} = 1$ the following covariance estimate :*

$$\left| \text{cov}_{\mu_{\tilde{K},a,m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \psi'(x_i) \right) \right|^2 \lesssim C(\tilde{K}) \int \frac{\sum_{i=1}^{\tilde{K}} \left| \frac{d}{dx_i} f \right|^2}{f} d\mu_{\tilde{K},a,m}.$$

Using these two lemmas and a two-scale decomposition of the variance, we can prove Proposition 2.6

Proof of Proposition 2.6. We decompose the system of K spins into \tilde{M} blocks, each containing \tilde{K} many spins i.e. $K = \tilde{K} \tilde{M}$. We disintegrate $\mu_{K,a,m}$ according to the coarse-graining operator

$$Px = \left(\frac{1}{\tilde{K}} \sum_{i \in B(1)} x_i, \dots, \frac{1}{\tilde{K}} \sum_{i \in B(\tilde{M})} x_i \right).$$

Here, the index set $B(l)$ of the l -th block is given by $B(l) = \{(l-1)\tilde{K} + 1, l\tilde{K}\}$, $l \in \{1, \dots, \tilde{M}\}$. Then the canonical ensemble $\mu_{K,m}$ can be written as

$$\mu_{K,m} = \mu(dx | Px = y) \bar{\mu}(dy),$$

where $\mu(dx|Px = y)$ are the conditional measures and $\bar{\mu}(dy)$ is the marginal. This disintegration implies the following decomposition of the covariance

$$\begin{aligned} \text{cov}_{\mu_{K,m}} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) &= \int \text{cov}_{\mu(dx|Px=y)} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \bar{\mu}(dy) \\ &\quad + \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px = y), \frac{1}{K} \int \sum_{i=1}^K \psi'(x_i) \mu(dx|Px = y) \right), \end{aligned} \quad (2.37)$$

and therefore

$$\begin{aligned} & \left| \text{cov}_{\mu_{K,m}} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \right|^2 \\ & \leq 2 \left(\int \text{cov}_{\mu(dx|Px=y)} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \bar{\mu}(dy) \right)^2 \\ & \quad + \left| \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px = y), \frac{1}{K} \int \sum_{i=1}^K \psi'(x_i) \mu(dx|Px = y) \right) \right|^2 \\ & \leq 2 \int \left(\text{cov}_{\mu(dx|Px=y)} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \right)^2 \bar{\mu}(dy) \\ & \quad + 2 \left| \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px = y), \frac{1}{K} \int \sum_{i=1}^K \psi'(x_i) \mu(dx|Px = y) \right) \right|^2. \end{aligned} \quad (2.38)$$

Let us now estimate the first term on the right hand side of (2.38) :
Note that the conditional measures $\mu(dx|Px = y)$ have a product structure i.e. one can write

$$\mu(dx|Px = y) = \bigotimes_{j=1}^{\tilde{M}} \mu_{\tilde{K},y_j}(dx^j),$$

where the probability measures $\mu_{\tilde{K},y_j}$ are defined as in 5.7. Additionally, we used the notation $x^j = (x_i)_{i \in B(j)}$. The latter yields the following representation of the microscopic covariance

$$\begin{aligned} \text{cov}_{\bar{\mu}(dx|Px=y)} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \\ = \frac{1}{\tilde{M}} \sum_{j=1}^{\tilde{M}} \text{cov}_{\mu_{\tilde{K},y_j}} \left(\int f \otimes_{i \neq j} \mu_{\tilde{K},y_i}(dx_i), \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \psi'(x_i) \right). \end{aligned}$$

On the right hand side we apply Lemma 2.14. This leads to the estimate

$$\begin{aligned} & \text{cov}_{\bar{\mu}(dx|Px=y)} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right)^2 \\ & \leq C(\tilde{K}) \frac{1}{\tilde{M}} \sum_{j=1}^{\tilde{M}} \int \frac{\sum_{i \in B(j)} \left| \frac{d}{dx_i} f \right|^2}{f} \mu(dx|Px = y) \\ & = C(\tilde{K}) \frac{1}{\tilde{M}} \int \frac{\sum_{i=1}^K \left| \frac{d}{dx_i} f \right|^2}{f} \mu(dx|Px = y). \end{aligned} \quad (2.39)$$

Let us now estimate the second term on the right hand side of (2.38) :
We take a closer look at the structure of the marginal $\bar{\mu}$. It turns out that the Hamiltonian \bar{H} of

$$\bar{\mu}(dy) = \frac{1}{Z} \exp\left(-\bar{H}(y)\right) \mathcal{H}^{\tilde{M}-1} \left\{ \frac{1}{\tilde{M}} \sum_{j=1}^{\tilde{M}} y_j = m \right\} (dx)$$

is given by a sum of single-site potentials $\psi_{\tilde{K},j}$ which depend on the site j because of the inhomogeneous linear part.

$$\bar{H}(y) = \sum_{j=1}^{\tilde{M}} \tilde{K} \psi_{\tilde{K},j}(y_j),$$

where the single-site potential $\psi_{\tilde{K},j}$ is given by

$$\psi_{\tilde{K}}(y_j) = -\frac{1}{\tilde{K}} \log \int \exp\left(-\sum_{i=1}^{\tilde{K}} \psi(x_i) + a_{(j-1)\tilde{K}+i}\right) \mathcal{H}(dx)^{\tilde{K}-1} \left\{ \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} x_i = y_j \right\}.$$

Without loss of generality we may assume that $\tilde{K} \gg 1$. Hence by the generalized local Cramér theorem (see Appendix A), the function $\psi_{\tilde{K}}$ is uniformly strictly convex. A direct calculation yields that the first derivative of $\psi_{\tilde{K},j}$ is given by

$$\psi'_{\tilde{K}}(y_j) = \frac{1}{\tilde{K}} \int \sum_{i=1}^{\tilde{K}} \psi'(x_i) + a_{(j-1)\tilde{K}+i} \mu_{\tilde{K},y_j}(dx). \quad (2.40)$$

Let us now estimate the macroscopic covariance term. Since the constants a_i disappear in the covariance, rewriting the macroscopic covariance using (2.40) yields

$$\begin{aligned} \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px = y), \frac{1}{\tilde{K}} \int \sum_{i=1}^{\tilde{K}} \psi'(x_i) \mu(dx|Px = y) \right) \\ = \frac{1}{\tilde{K}\tilde{M}} \sum_{j=1}^{\tilde{M}} \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px = y), \tilde{K} \psi'_{\tilde{K}}(y_j) \right) \\ = \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px = y), \frac{1}{\tilde{K}\tilde{M}} \sum_{j=1}^{\tilde{M}} \tilde{K} \psi'_{\tilde{K}}(y_j) \right). \end{aligned}$$

Therefore an application of Lemma 2.13 yields the estimate

$$\begin{aligned} \left| \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px = y), \frac{1}{\tilde{K}} \int \sum_{i=1}^{\tilde{K}} \psi'(x_i) \mu(dx|Px = y) \right) \right|^2 \\ \leq C\tilde{M} \int \frac{\sum_{j=1}^{\tilde{M}} \left| \frac{d}{dy_j} \int f \mu(dx|Px = y) \right|^2}{\int f(x) \mu(dx|y)} \bar{\mu}(dy). \quad (2.41) \end{aligned}$$

Because

$$\begin{aligned} \frac{d}{dy_j} \int f \mu(dx|Px = y) \\ = \int \sum_{i \in B(j)} \frac{d}{dx_i} f \mu(dx|Px = y) - \text{cov}_{\mu(dx|Px=y)} \left(f, \sum_{i \in B(j)} \psi'(x_i) \right), \end{aligned}$$

we get the estimate

$$\begin{aligned} & \left| \frac{d}{dy_j} \int f \mu(dx|Px=y) \right| \\ & \lesssim \tilde{K} \int \sum_{i \in B(j)} \left| \frac{d}{dx_i} f \right| \mu(dx|Px=y) \\ & + \left| \text{cov}_{\mu(dx|Px=y)} \left(f, \sum_{i \in B(j)} \psi'(x_i) \right) \right|. \end{aligned}$$

The covariance terms can be estimated by an application of Lemma 2.14 yielding the overall inequality

$$\begin{aligned} & \left| \frac{d}{dy_j} \int f \mu(dx|Px=y) \right|^2 \\ & \lesssim C(\tilde{K}) \left(\int f(x) \mu(dx|y) \right) \int \frac{\sum_{i \in B(j)} \left| \frac{d}{dx_i} f \right|^2}{f} \mu(dx|Px=y). \quad (2.42) \end{aligned}$$

Now combining the estimates (2.41) and (2.42) implies

$$\begin{aligned} & \left| \text{cov}_{\bar{\mu}} \left(\int f \mu(dx|Px=y), \frac{1}{K} \int \sum_{i=1}^K \psi'_c(x_i) \mu(dx|Px=y) \right) \right|^2 \\ & \lesssim C(\tilde{K}) \tilde{M} \int \frac{\sum_{i=1}^K \left| \frac{d}{dx_i} f \right|^2}{f} \mu_{K,m}(dx). \quad (2.43) \end{aligned}$$

Now, it is only left to insert the estimate (2.39) for the conditional covariance and the estimate (2.43) for the macroscopic covariance in the decomposition (2.38). Using $K = \tilde{K} \tilde{M}$ this yields

$$\left| \text{cov}_{\mu_{K,m}} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \right|^2 \lesssim C(\tilde{K}) K \int \frac{\sum_{i=1}^K \left| \frac{d}{dx_i} f \right|^2}{f} \mu_{K,m}(dx) \quad (2.44)$$

□

Proof of Lemma 2.13. We will apply a recursive approach. In the first step, we will estimate the term

$$\text{cov}_{\mu_{K,m}} \left(f, \frac{1}{K} \psi'(x_n) \right) \quad (2.45)$$

for a fixed site $n \in \{1, \dots, K\}$. Let us choose another arbitrary site $\ell \in \{1, \dots, K\}$ and disintegrate the canonical ensemble $\mu_{K,m}$ by fixing the spin x_ℓ i.e.

$$\mu_{K,a,m}(dx) = \mu(d\bar{x}_\ell|x_\ell) \bar{\mu}(dx_\ell), \quad (2.46)$$

where we used the notation

$$\bar{x}_\ell = (x_1, \dots, x_{\ell-1}, x_{\ell+1}, \dots, x_K).$$

Then, the marginal $\bar{\mu}(dx_\ell)$ is given by

$$\bar{\mu}(dx_\ell) = \frac{1}{Z} \exp \left(-\psi(x_\ell) - a_\ell - (K-1) \psi_{K-1, \bar{a}_\ell} \left(\frac{K}{K-1} m - \frac{1}{K-1} x_\ell \right) \right),$$

where the function ψ_{K-1, \bar{a}_ℓ} is given by

$$\psi_{K-1}(z) = -\frac{1}{K-1} \log \int \exp \left(-\sum_{i=1}^{K-1} \psi(x_i) - \langle \bar{a}_\ell, x \rangle \right) \mathcal{H}_{\left\{ \frac{1}{K-1} \sum_{i=1}^{K-1} x_i = z \right\}}^{K-2}(dx).$$

Note that the single-site potential ψ_{K-1, \bar{a}_ℓ} inherits the uniform strict convexity from the single-site potential ψ by a standard argument using the (symmetric) Brascamp-Lieb inequality. The disintegration (2.46) of $\mu_{K, a, m}$ yields the following representation of the covariance

$$\text{cov}_{\mu_{K, m}} \left(f, \frac{1}{K} \psi'(x_n) \right) = \text{cov}_{\bar{\mu}(dx_\ell)} \left(\int f \mu(d\bar{x}_\ell | x_\ell), \frac{1}{K} \psi'(x_n) \right).$$

Because $\bar{\mu}(dx_\ell)$ is an one-dimensional measure with strictly convex Hamiltonian, an application of the one-dimensional asymmetric Brascamp-Lieb inequality yields

$$\begin{aligned} & \text{cov}_{\mu_{K, m}} \left(f, \frac{1}{K} \psi'(x_n) \right) \\ & \leq \frac{1}{K} \underbrace{\left| \frac{\psi''(x_\ell)}{\psi''(x_\ell) + \frac{1}{K-1} \psi_K'' \left(\frac{K}{K-1} m - \frac{1}{K-1} x_\ell \right)} \right|}_{\leq 1} \int \left| \frac{d}{dx_\ell} \int f \mu(d\bar{x}_\ell | x_\ell) \right| \bar{\mu}(dx_\ell). \end{aligned} \quad (2.47)$$

Direct calculation reveals that

$$\begin{aligned} & \frac{d}{dx_\ell} \int f \mu(d\bar{x}_\ell | x_\ell) \\ & = \int \left(\frac{d}{dx_\ell} f - \frac{1}{K-1} \sum_{i=1}^{K-1} \frac{d}{dx_i} f \right) \mu(d\bar{x}_\ell | x_\ell) - \frac{1}{K-1} \text{cov}_{\mu(d\bar{x}_\ell | x_\ell)} \left(f, \sum_{i \neq \ell, i=1}^K \psi'(x_i) \right) \\ & = \frac{1}{K-1} \int \left(\sum_{i=1}^K \frac{d}{dx_\ell} f - \frac{d}{dx_i} f \right) \mu(d\bar{x}_\ell | x_\ell) - \frac{1}{K-1} \text{cov}_{\mu(d\bar{x}_\ell | x_\ell)} \left(f, \sum_{i \neq \ell, i=1}^K \psi'(x_i) \right) \end{aligned}$$

Using the last identity one can directly deduce from (2.47) that

$$\begin{aligned} \text{cov}_{\mu_{K, m}} \left(f, \frac{1}{K} \psi'(x_n) \right) & \leq \frac{1}{K-1} \int \sum_{i=1}^K \left| \frac{d}{dx_\ell} f - \frac{d}{dx_i} f \right| \mu_{K, m}(dx) \\ & \quad + \int \left| \frac{1}{K-1} \text{cov}_{\mu(d\bar{x}_\ell | x_\ell)} \left(f, \sum_{i \neq \ell, i=1}^K \psi'(x_i) \right) \right| \bar{\mu}(dx_\ell). \end{aligned} \quad (2.48)$$

One sees that on the right hand side, one has to estimate a covariance term

$$\frac{1}{K-1} \text{cov}_{\mu(d\bar{x}_\ell | x_\ell)} \left(f, \sum_{i \neq \ell, i=1}^K \psi'(x_i) \right)$$

that has the same structure as the covariance term we started with (cf. (2.45)). The reason is that the conditional measure $\mu(d\bar{x}_\ell | x_\ell)$ has the structure of a canonical ensemble i.e.

$$\mu(d\bar{x}_\ell | x_\ell) = \mu_{K-1, \frac{K}{K-1} m - \frac{1}{K-1} x_\ell}(d\bar{x}_\ell).$$

Therefore one can apply the estimate (2.48) recursively, until there is no covariance term left. On the right hand side of the covariance estimate, there only occur terms of the form

$$\int \left| \frac{d}{dx_i} f - \frac{d}{dx_j} f \right| \mu_{K,m}$$

for some indexes $1 \leq i < j \leq K$. Therefore, one only has to determine the prefactors in front of these terms. By symmetry of the system this prefactor is independent of the particular choice i and j and can be determined by combinatorics. Indeed, we will need the following observation that follows from the recursive formula (2.48) :

The term $\int \left| \frac{d}{dx_i} f - \frac{d}{dx_j} f \right| \mu_{K,m}$ only occurs, if one conditions on the spin value x_i and the spin value x_j is free or if one conditions on the spin value x_j and the spin value x_i is free.

Let us start to determine the prefactors. In the first step, there are K free spins. Let us condition on an arbitrary spin $x_l, l \in \{1, \dots, K\}$. Then the probability that one conditions on the spin x_i or x_j is given by $\frac{2}{K}$. In this case, one gets by (2.48) the prefactor

$$\frac{2}{K} \frac{1}{K-1}.$$

Note that in the remaining conditional steps, the term $\int \left| \frac{d}{dx_i} f - \frac{d}{dx_j} f \right| \mu_{K,m}$ does not occur anymore.

So, let us consider the remaining cases. We assume that in the first step one conditioned on a spin x_l with $i \neq l \neq j$ and in the second step one conditions on the spin x_i or x_j . It follows from (2.48) that the resulting prefactor is given by

$$\frac{K-2}{K} \frac{2}{K-1} \frac{1}{K-2}.$$

With this argument it becomes clear that the overall prefactor c is given by

$$\begin{aligned} c &= \frac{2}{K} \frac{1}{K-1} + \frac{K-2}{K} \frac{2}{K-1} \frac{1}{K-2} + \frac{K-2}{K} \frac{K-3}{K-1} \frac{2}{K-2} \frac{1}{K-3} + \dots \\ &= \frac{2}{K} \frac{1}{K-1} \sum_{i=1}^{K-1} 1 = \frac{2}{K}. \end{aligned}$$

Overall, this deduces the covariance estimate

$$\text{cov}_{\mu_{K,m}} \left(f, \frac{1}{K} \psi'(x_i) \right) \lesssim \frac{1}{K} \int \sum_{i,j=1}^K \left| \frac{d}{dx_i} f - \frac{d}{dx_j} f \right| \mu_{K,m} \quad (2.49)$$

and therefore

$$\begin{aligned} & \left| \text{cov}_{\mu_{K,m}} \left(f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \right|^2 \lesssim \left(\int \sum_{i,j=1}^K \left| \frac{d}{dx_i} f - \frac{d}{dx_j} f \right| \mu_{K,m} \right)^2 \\ & \lesssim \left(\int \frac{\sum_{i,j=1}^K \left| \frac{d}{dx_i} f - \frac{d}{dx_j} f \right|^2}{f} \mu_{K,m} \right) \left(\int f \mu_{K,m} \right) \\ & \lesssim \int \frac{\sum_{i,j=1}^K 2 \left| \frac{d}{dx_i} f \right|^2 + 2 \left| \frac{d}{dx_j} f \right|^2}{f} \mu_{K,m} \\ & \lesssim K \int \frac{\sum_{i=1}^K \left| \frac{d}{dx_i} f \right|^2}{f} \mu_{K,m} \end{aligned} \quad (2.50)$$

□

Proof of Lemma 2.14. By the splitting $\psi = \psi_c + \delta\psi$, we can decompose the covariance term according to

$$\begin{aligned} \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \psi'(x_i) \right) \\ = \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \psi'_c(x_i) \right) + \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \delta\psi'(x_i) \right). \end{aligned} \quad (2.51)$$

Because $|\delta\psi'| \lesssim 1$ by assumption, we can estimate the second term on the right hand side of the last equality in the same way as in [MO13] by an application of the classical Csiszár-Kullback-Pinsker inequality

$$\|f\mu - \mu\|_{TV} \leq \sqrt{2 \text{Ent}_\mu(f)}$$

and the LSI for $\mu_{\tilde{K},m}$, which holds a priori with a bad constant that depends on \tilde{K} (see Lemma 2.15 in the next section). Indeed,

$$\begin{aligned} \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \delta\psi'(x_i) \right) &\leq \|\delta\psi'\|_\infty \int |f(x) - 1| \mu_{\tilde{K},m}(dx) \\ &= \|f\mu_{\tilde{K},m} - \mu_{\tilde{K},m}\|_{TV} \\ &\leq \sqrt{2 \text{Ent}_{\mu_{\tilde{K},m}}(f)} \\ &\leq C(\tilde{K}) \left(\int \frac{\sum \left| \frac{df}{dx_i} \right|^2}{f} \mu_{\tilde{K},m} \right)^{1/2}. \end{aligned} \quad (2.52)$$

A proof of the Csiszár-Kullback-Pinsker inequality can be found for example in [Led01]. Now, let us have a look at the first term on the right hand side of (2.51). This term can be estimated with the same strategy as for Lemma 2.13. The only difference is that instead of the asymmetric Brascamp-Lieb inequality for strictly convex single-site potentials one has to apply the version for perturbed strictly convex single-site potentials. However, this does not matter, because we do not care about constants depending on \tilde{K} . There is still a little bit of work, because ψ_K is not convex for small K . This is not a problem because ψ_K can be represented as a sum of a strictly convex potential and a perturbation (see [MO13]). What we get is

$$\begin{aligned} \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \psi'_c(x_i) \right) &\leq C(\tilde{K}) \int \sum_{i=1}^{\tilde{K}} \left| \frac{d}{dx_i} f \right| \mu_{\tilde{K},m}, \\ &\leq C(\tilde{K}) \left(\int \frac{\sum_{i=1}^{\tilde{K}} \left| \frac{d}{dx_i} f \right|^2}{f} \mu_{\tilde{K},m} \right)^{1/2} \left(\int f \mu_{\tilde{K},m} \right)^{1/2}. \end{aligned} \quad (2.53)$$

Inserting (2.52) and (2.53) into (2.51), we get

$$\begin{aligned}
& \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \psi'(x_i) \right)^2 \\
& \leq 2 \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \psi'_c(x_i) \right)^2 + 2 \text{cov}_{\mu_{\tilde{K},m}} \left(f, \frac{1}{\tilde{K}} \sum_{i=1}^{\tilde{K}} \delta \psi'(x_i) \right)^2 \\
& \leq C(\tilde{K}) \left(\int \frac{\sum_{i=1}^{\tilde{K}} \left| \frac{d}{dx_i} f \right|^2}{f} \mu_{\tilde{K},m} \right)^{1/2} \left(\int f \mu_{\tilde{K},m} \right). \tag{2.54}
\end{aligned}$$

□

2.4 Proof of Theorem 2.7

As announced, to prove this result we will use the two-scale approach that was introduced in [GOVW09] and the covariance estimate of Proposition 2.6. First, we will prove that the logarithmic Sobolev inequalities are satisfied at macroscopic scale and at the scale of fluctuations around a fixed macroscopic profile, before showing how we can combine the two LSIs to obtain a full LSI at microscopic scale.

Lemma 2.15. *The measure μ satisfies $LSI(\rho e^{-cN})$, with constants c and ρ independent of N , m and the sequence (a_i) . As a consequence, $\mu(dx|y)$ satisfies $LSI(\rho_0)$ for some constant ρ_0 that depends on K , but not on y .*

Démonstration. The first part is a consequence of a combination of the Bakry-Emery theorem and the Holley-Stroock perturbation Lemma (see Appendix B).

Recall (2.16) :

$$\mu_{N,a,m}(dx|y) = \bigotimes_{i=1}^M \mu_{K,(a_j)_{(i-1)K+1 \leq j \leq iK}, y_i}$$

so by the tensorization principle (Criterion I of Appendix B), $\mu(dx|y)$ satisfies $LSI(\rho_0)$ for some constant ρ_0 that depends on K , but not on y , since $\mu_{K,(a_j)_{(i-1)K+1 \leq j \leq iK}, y_i}$ satisfies $LSI(\rho e^{-cK})$. □

Note that, since $\mu_{K,(a_j)_{(i-1)K+1 \leq j \leq iK}, y_i}$ has the same form as μ , Theorem 2.7 will imply that a posteriori ρ_0 does not actually depend on K .

Lemma 2.16. *There exists a positive constant λ and an integer K_0 such that for all $K \geq K_0$, the measure $\bar{\mu}$ satisfies $LSI(\lambda N)$, independently of the mean m and the sequence (a_i) .*

Démonstration. By the generalized local Cramèr theorem of Appendix A, for K large enough, $\bar{H}(y) = \frac{1}{M} \sum \psi_{K,i,a}(y_i)$ is λ -uniformly convex, with constant λ independent of K , m and the sequence (a_i) . An application of the Bakry-Emery Theorem then yields the desired result. □

To get the full logarithmic Sobolev inequality at microscopic scale, we rely on the following decomposition of the entropy :

$$\text{Ent}_\mu(f) = \text{Ent}_{\bar{\mu}}(\bar{f}) + \int_Y \text{Ent}_{\mu(dx|y)}(f) \bar{\mu}(dy). \quad (2.55)$$

This identity separates the entropy into the sum of a macroscopic entropy and an entropy at the scale of fluctuations. We can then use the fact that we have a LSI at both scales to bound these two terms separately. Indeed, by using Lemma 2.15, we have, for a decomposition with large but fixed K ,

$$\text{Ent}_{\mu(dx|y)}(f) \leq \frac{1}{\rho_0} \int_X \frac{|(id_X - P^t N P) \nabla f|^2}{f} \mu(dx|y). \quad (2.56)$$

In the same way, at macroscopic scale, we have

$$\text{Ent}_{\bar{\mu}}(\bar{f}) \leq \frac{1}{\lambda N} \int_Y \frac{|\nabla \bar{f}|^2}{\bar{f}} \bar{\mu}. \quad (2.57)$$

To use this second estimate, we need to relate $\nabla \bar{f}$ and ∇f . This is the point of the following Lemma :

Lemma 2.17. *For any $f \in \text{Lip}(X)$ and any $y \in Y$,*

$$\int P \nabla f \mu(dx|y) = \frac{1}{N} \nabla_Y \bar{f}(y) + P \text{cov}_{\mu(dx|y)}(f, \nabla H). \quad (2.58)$$

The proof of this Lemma will be deferred to the end of this section. Using this result and the convexity of the function $(x, b) \rightarrow |x|^2/b$ with Jensen's inequality, we get

$$\begin{aligned} \frac{|\nabla \bar{f}|^2}{\bar{f}} &= \frac{\left| N \int P \nabla f \mu(dx|y) + N P \text{cov}_{\mu(dx|y)}(f, \nabla H) \right|^2}{\int f(x) \mu(dx|y)} \\ &\leq 2N \frac{1}{\bar{f}(y)} \left| \int P^t N P \nabla f \mu(dx|y) \right|^2 + 2N \frac{1}{\bar{f}(y)} \left| N P^t P \text{cov}(f, \nabla H) \right|^2 \\ &\leq 2N \int \frac{|P^t N P \nabla f|^2}{f} \mu(dx|y) + 2N \frac{1}{\bar{f}(y)} \left| N P^t P \text{cov}(f, \nabla H) \right|^2 \end{aligned} \quad (2.59)$$

We can now use the covariance estimate of Proposition 2.6 to bound the second term on the right-hand side of this equation.

$$\begin{aligned} \left| N P^t P \text{cov}(f, \nabla H) \right|^2 &= \sum_{i=1}^M \left| \text{cov}_{\mu_{K,a,y_i}} \left(\int f \otimes_{k \neq i} \mu_{K,a,y_k}, \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \psi'(x_j) \right) \right|^2 \\ &\leq CK \bar{f}(y) \int \frac{|(id_X - P^t N P) \nabla f|^2}{f} \mu_{N,a,m}(dx|y). \end{aligned} \quad (2.60)$$

Plugging (2.56), (2.57), (2.59) and (2.60) into (2.55), we get

$$\begin{aligned} \text{Ent}_{\mu_{N,a,m}}(f) &\leq \frac{2}{\lambda} \int \frac{|P^t N P \nabla f|^2}{f} \mu_{N,a,m}(dx) \\ &\quad + \left(\frac{\exp(cK)}{\rho} + \frac{2CK}{\lambda} \right) \int \frac{|(id_X - P^t N P) \nabla f|^2}{f} \mu_{N,a,m}(dx) \\ &\lesssim (1 + K) \exp(cK) \int \frac{|\nabla f|^2}{f} \mu_{N,a,m}(dx). \end{aligned} \quad (2.61)$$

Since as long as K is large enough for the coarse-grained Hamiltonian to be strictly convex, this bound holds, we just have to take K large but fixed for Theorem 2.7 to be proved.

We finish this section with the proof of Lemma 2.17, which is unchanged from [GOVW09] Lemma 21.

Proof of Lemma 2.17. By definition, we have

$$\begin{aligned}\bar{f}(y) &= \int f(x) \mu_{N,a,m}(dx|y) \\ &= \frac{1}{\int_{\{P_x=0\}} \exp(-H(NP^t y + z)) dz} \int_{\{P_x=0\}} f(NP^t y + z) \exp(-H(NP^t y + z)) dz.\end{aligned}$$

Thus, for any $\tilde{y} \in Y$, we have

$$\begin{aligned}\nabla_Y \bar{f}(y) \cdot \tilde{y} &= N \int \nabla f(x) \cdot P^t \tilde{y} \mu(dx|y) - N \int f(x) \nabla H(x) \cdot P^t \tilde{y} \mu(dx|y) \\ &\quad - N \left(\int f(x) \mu(dx|y) \right) \left(\int (-H(x) \cdot P^t \tilde{y}) \mu(dx|y) \right) \\ &= N \left[\int P \nabla f(x) \mu(dx|y) - \int f(x) P \nabla H(x) \mu(dx|y) \right. \\ &\quad \left. + \left(\int f(x) \mu(dx|y) \right) \left(\int P \nabla H(x) \mu(dx|y) \right) \right] \cdot \tilde{y},\end{aligned}$$

which is what we were aiming for. \square

2.5 Proof of the hydrodynamic limit

This section is devoted to the proof of the hydrodynamic limit, namely the abstract quantitative estimate of Theorem 2.8 and the concrete application to the Kawasaki dynamics stated in Theorem 2.9. This section is organized in the following way :

- In Section 2.5.1 we give the complete proof of Theorem 2.8.
- In Section 2.5.2 we give the complete proof of Theorem 2.9 up to two missing ingredients, namely Proposition 2.21 and Proposition 2.25. Proposition 2.21 contains the convergence of the deterministic macroscopic ODE given by (2.21) to the nonlinear heat equation given by 2.32 on the continuum. Proposition 2.25 contains polynomial bounds on the free energy.
- In Section 2.5.3, we state and prove Proposition 2.25.
- In Section 2.5.4, we state and prove auxiliary results needed for Proposition 2.21.
- And finally in Section 2.5.5, we prove Proposition 2.21.

2.5.1 Proof of Theorem 2.8

The first element of the proof is the following lemma, which will allow to control fluctuations.

Lemma 2.18. *There exists a constant $\gamma > 0$, which is independent of N and M , such that for any $x \in X_{N,0}$, we have*

$$|(id_X - P^t NP)x|_X^2 \leq \frac{\gamma}{M^2} \langle Ax, x \rangle_X; \quad (2.62)$$

$$\langle A^{-1}(id_X - P^t NP)x, (id_X - P^t NP)x \rangle \leq \frac{\gamma}{M^2} |x|_X^2. \quad (2.63)$$

Of course, these statements are not new and we copy the argument of (54) in [GOVW09].

Démonstration. Bound (2.62) is a consequence of the discrete Poincaré inequality applied to the blocks : there exists a universal constant γ such that for any $N \in \mathbb{N}$ and real numbers x_1, \dots, x_N with $\sum x_i = 0$, we have

$$\sum_{i=1}^N x_i^2 \leq \gamma N^2 \sum_{i=1}^N (x_{i+1} - x_i)^2.$$

Estimate (2.63) follows from (2.62) and the Cauchy-Schwartz inequality :

$$\begin{aligned} \langle A^{-1}(id_X - P^t NP)x, (id_X - P^t NP)x \rangle &= \langle x, (id_X - P^t NP)A^{-1}(id_X - P^t NP)x \rangle \\ &\leq |x| |(id_X - P^t NP)A^{-1}(id_X - P^t NP)x| \\ &\leq |x| \left(\frac{\gamma}{M^2} \langle AA^{-1}(id_X - P^t NP)x, A^{-1}(id_X - P^t NP)x \rangle \right)^{1/2} \\ &= |x| \left(\frac{\gamma}{M^2} \langle (id_X - P^t NP)x, A^{-1}(id_X - P^t NP)x \rangle \right)^{1/2}. \end{aligned}$$

□

We will also need the following energy and moment estimates, which are those of Proposition 24 in [GOVW09]. For the sake of completeness we give the proof of Proposition 2.19 at the end of the section.

Proposition 2.19. *If $f(t, x)$ and $\eta(t)$ satisfy the assumptions of Theorem 2.8, then for any $T < +\infty$ we have*

$$\text{Ent}_\mu(f(T, \cdot)) + \int_0^T \left(\int_X \frac{\langle A \nabla f, \nabla f \rangle}{f} d\mu \right) dt = \text{Ent}_\mu(f_0); \quad (2.64)$$

$$\bar{H}(\eta(T)) + \int_0^T \left\langle \frac{d\eta}{dt}, \bar{A}^{-1} \frac{d\eta}{dt} \right\rangle dt = \bar{H}(\eta(0)); \quad (2.65)$$

$$\left(\int |x|^2 f(t, x) \mu(dx) \right)^{1/2} \leq \left(\frac{2}{\rho} \text{Ent}_\mu(f_0) \right)^{1/2} + \left(\int |x|^2 \mu(dx) \right)^{1/2} \quad (2.66)$$

We will now prove estimate (2.24) in three steps : first we will compute the time derivative of Θ , then we will bound it, and finally integrate it in time to obtain our result. The calculations are exactly the same as in [GOVW09], except for the use of the new covariance estimate of Proposition 2.6

Step 1 : Our aim is to obtain the exact formula

$$\begin{aligned} \frac{d}{dt} \Theta(t) &= - \int (\nabla_Y \bar{H}(y) - \nabla_Y \bar{H}(\eta)) \cdot (y - \eta) \bar{f}(y) \bar{\mu}(dy) + \frac{M-1}{N} \\ &\quad - \int (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu}(dy) \\ &\quad - \int \frac{1}{N} (id - P^t NP)x \cdot \nabla f \mu(dx) - \int \frac{d\eta}{dt} \cdot P A^{-1}(id - P^t NP)x f \mu(dx). \end{aligned} \quad (2.67)$$

Using the definition of the stochastic dynamic, of the coarse-grained evolution (2.21) and the splitting $x = P^t NPx + (id_X - P^t NP)x$, we have

$$\begin{aligned}
\frac{d}{dt}\Theta(t) &= \frac{d}{dt} \frac{1}{2N} \int \langle A^{-1}(x - NP^t\eta(t)), (x - NP^t\eta(t)) \rangle f(t, x) \mu_{N,a,m}(dx) \\
&\stackrel{(2.19)}{=} - \int \frac{1}{N} A^{-1}(x - NP^t\eta(t)) \cdot A \nabla f(t, x) \mu(dx) - \int P^t \frac{d\eta}{dt} \cdot A^{-1}(x - NP^t\eta) f \mu(dx) \\
&= - \int P^t (Px - \eta) \cdot \nabla f \mu(dx) - \int PA^{-1} NP^t \frac{d\eta}{dt} \cdot (Px - \eta) f \mu(dx) \\
&\quad - \int \frac{1}{N} (id - P^t NP) x \cdot \nabla f \mu(dx) - \int \frac{d\eta}{dt} \cdot PA^{-1} (id - P^t NP) x f \mu(dx) \\
&= - \int (Px - \eta) \cdot P \nabla f \mu(dx) - \int \bar{A}^{-1} \frac{d\eta}{dt} \cdot (Px - \eta) f \mu(dx) \\
&\quad - \int \frac{1}{N} (id - P^t NP) x \cdot \nabla f \mu(dx) - \int \frac{d\eta}{dt} \cdot PA^{-1} (id - P^t NP) x f \mu(dx) \\
&= - \int (y - \eta) \cdot P \int \nabla f \mu(dx|y) \bar{\mu}(dy) + \int \nabla_Y \bar{H}(\eta) \cdot (y - \eta) \bar{f}(y) \bar{\mu}(dy) \\
&\quad - \int \frac{1}{N} (id - P^t NP) x \cdot \nabla f \mu(dx) - \int \frac{d\eta}{dt} \cdot PA^{-1} (id - P^t NP) x f \mu(dx). \tag{2.68}
\end{aligned}$$

We keep the last three terms unchanged, and transform the first one according to Lemma 2.17

$$\begin{aligned}
&- \int (y - \eta) \cdot P \int \nabla f \mu(dx|y) \bar{\mu}(dy) \\
&= - \frac{1}{N} \int (y - \eta) \cdot \nabla_Y \bar{f}(y) \bar{\mu}(dy) - \int (y - \eta) \cdot P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu}(dy). \tag{2.69}
\end{aligned}$$

Using integration by parts, the first term on the right-hand side of this last equation becomes

$$\begin{aligned}
&- \frac{1}{N} \int (y - \eta) \cdot \nabla_Y \bar{f}(y) \bar{\mu}(dy) \\
&= \frac{1}{N} \int (\nabla_Y \cdot y) \bar{f}(y) \bar{\mu}(dy) - \int (y - \eta) \cdot \nabla_Y \bar{H}(y) \bar{\mu}(dy) \\
&= \frac{M-1}{N} - \int (y - \eta) \cdot \nabla_Y \bar{H}(y) \bar{\mu}(dy). \tag{2.70}
\end{aligned}$$

This final estimate allows us to conclude step 1.

Step 2 : *Upper bound on the time-derivative* Our aim here is to get the upper bound

$$\begin{aligned}
\frac{d}{dt}\Theta(t) &+ \frac{\lambda}{2} \int |y - \eta|_Y^2 \bar{f} \bar{\mu}(dy) \\
&\leq \frac{M-1}{N} + \frac{\gamma C_0 K}{2\lambda N M^2} \int \frac{\langle A \nabla f, \nabla f \rangle_X}{f} \mu(dx) \\
&\quad + \left(\frac{\gamma}{M^2} \int \frac{1}{N f} \langle A \nabla f, \nabla f \rangle \mu(dx) \cdot \int \frac{1}{N} |x|_X^2 f(x) \mu(dx) \right)^{1/2} \\
&\quad + \left(\frac{d\eta}{dt} \cdot \bar{A}^{-1} \frac{d\eta}{dt} \right)^{1/2} \left(\int \frac{\gamma}{N M^2} |x|_X^2 f(x) \mu(dx) \right)^{1/2} \tag{2.71}
\end{aligned}$$

We will now individually bound each term in (2.67). The first term can be bounded by using the assumption that K is large enough such that \bar{H} is λ - uniformly convex, which yields

$$- \int (\nabla_Y \bar{H}(y) - \nabla_Y \bar{H}(\eta)) \cdot (y - \eta) \bar{f}(y) \bar{\mu}(dy) \leq -\lambda \int |y - \eta|_Y^2 \bar{f} \bar{\mu}(dy). \tag{2.72}$$

The third term can be treated in the same way as was done for estimate (2.60) in the proof of the LSI, and by using Lemma 2.18 :

$$\begin{aligned}
& \int (y - \eta) \cdot P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu}(dy) \\
& \leq \left(\int |y - \eta|_Y^2 \bar{f}(y) \bar{\mu}(dy) \right)^{1/2} \left(\int \frac{|P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|^2}{\bar{f}(y)} \bar{\mu}(dy) \right)^{1/2} \\
& \leq \frac{\lambda}{2} \int |y - \eta|_Y^2 \bar{f}(y) \bar{\mu}(dy) + \frac{1}{2\lambda N} \int \frac{|P^t N P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|^2}{\bar{f}(y)} \bar{\mu}(dy) \\
& \leq \frac{\lambda}{2} \int |y - \eta|_Y^2 \bar{f}(y) \bar{\mu}(dy) + \frac{C_0 K}{2\lambda N} \int \frac{|(id_X - P^t N P) \nabla f|^2}{f} \mu(dx) \\
& \leq \frac{\lambda}{2} \int |y - \eta|_Y^2 \bar{f}(y) \bar{\mu}(dy) + \frac{\gamma C_0 K}{2\lambda N M^2} \int \frac{\langle A \nabla f, \nabla f \rangle_X}{f} \mu(dx). \tag{2.73}
\end{aligned}$$

The fourth term can be controlled using (2.63) and the Cauchy-Schwartz inequality :

$$\begin{aligned}
& \left| \int \frac{1}{N} (id - P^t N P) x \cdot \nabla f \mu(dx) \right| \\
& \leq \left(\frac{1}{N} \int \langle A^{-1} (id - P^t N P) x, (id - P^t N P) x \rangle f(x) \mu(dx) \cdot \int \frac{1}{N f} \langle A \nabla f, \nabla f \rangle \mu(dx) \right)^{1/2} \\
& \leq \left(\frac{\gamma}{M^2} \int \frac{1}{N f} \langle A \nabla f, \nabla f \rangle \mu(dx) \cdot \int \frac{1}{N} |x|_X^2 f(x) \mu(dx) \right)^{1/2} \tag{2.74}
\end{aligned}$$

Similarly, we have for the fifth term

$$\begin{aligned}
& \left| \int \frac{d\eta}{dt} \cdot P A^{-1} (id - P^t N P) x f \mu(dx) \right| \\
& \leq \left(\int P^t \frac{d\eta}{dt} \cdot A^{-1} N P^t \frac{d\eta}{dt} f x \mu(dx) \right)^{1/2} \\
& \quad \cdot \left(\int \frac{1}{N} \langle A^{-1} (id - P^t N P) x, (id - P^t N P) x \rangle f(x) \mu(dx) \right)^{1/2} \\
& \leq \left(\frac{d\eta}{dt} \cdot \bar{A}^{-1} \frac{d\eta}{dt} \right)^{1/2} \left(\int \frac{\gamma}{N M^2} |x|_X^2 f(x) \mu(dx) \right)^{1/2} \tag{2.75}
\end{aligned}$$

Combining these estimates gives the desired upper bound.

Step 3 : *Time integration and conclusion* Integrating (2.71) with respect to time yields

$$\begin{aligned}
& \max \left\{ \sup_{0 \leq t \leq T} \Theta(t), \frac{\lambda}{2} \int_0^T \int_Y |y - \eta(t)|_Y^2 \bar{f} \bar{\mu}(dy) dt \right\} \\
& \leq \Theta(0) + T \frac{M-1}{N} + \frac{\gamma C_0 K}{2\lambda N M^2} \int_0^T \int \frac{\langle A \nabla f, \nabla f \rangle_X}{f} \mu(dx) dt \\
& \quad + \frac{\gamma^{1/2}}{M} \int_0^T \left(\int \frac{1}{N} |x|_X^2 f(x) \mu(dx) \right)^{1/2} \\
& \quad \times \left(\left(\int \frac{1}{N f} \langle A \nabla f, \nabla f \rangle \mu(dx) \right)^{1/2} + \left(\frac{d\eta}{dt} \cdot \bar{A}^{-1} \frac{d\eta}{dt} \right)^{1/2} \right) dt. \tag{2.76}
\end{aligned}$$

According to Proposition 2.19, since the entropy is positive, and the initial entropy satisfies the bound of assumption (i), we have

$$\int_0^T \int \frac{\langle A\nabla f, \nabla f \rangle_X}{f} \mu(dx) dt \leq C_1 N. \quad (2.77)$$

The bounds of Proposition 2.19 also yield the bound

$$\begin{aligned} & \int_0^T \left(\int \frac{1}{N} |x|_X^2 f(x) \mu(dx) \right)^{1/2} \times \left(\left(\int \frac{1}{Nf} \langle A\nabla f, \nabla f \rangle \mu(dx) \right)^{1/2} + \left(\frac{d\eta}{dt} \cdot \bar{A}^{-1} \frac{d\eta}{dt} \right)^{1/2} \right) dt \\ & \leq \left(\int_0^T \int \frac{1}{N} |x|_X^2 f(t, x) \mu(dx) dt \right)^{1/2} \\ & \quad \times \left(\left(\int_0^T \int \frac{1}{Nf} \langle A\nabla f, \nabla f \rangle \mu(dx) dt \right)^{1/2} + \left(\int_0^T \frac{d\eta}{dt} \cdot \bar{A}^{-1} \frac{d\eta}{dt} dt \right)^{1/2} \right) \\ & \leq \left(2 \int_0^T \left(\int \frac{1}{N} |x|_X^2 \mu(dx) + \frac{2}{N\rho} \text{Ent}_\mu(f_0) \right) dt \right)^{1/2} \\ & \quad \times \left(\left(\frac{1}{N} \text{Ent}_\mu(f_0) \right)^{1/2} + \left(\bar{H}(\eta_0) - \bar{H}(\eta(T)) \right)^{1/2} \right) \\ & \leq \left(2\alpha + \frac{2C_1}{\rho} \right)^{1/2} (C_1 + C_2 + \beta)^{1/2} \end{aligned} \quad (2.78)$$

Plugging (2.77) and (2.78) into (2.76) ends the proof.

We finish this section with the proof of Proposition 2.19. For the proof we will use the following lemma. It was proven in [GOVW09] using a semigroup argument. Here we give a proof based on Talagrand's inequality.

Lemma 2.20. *Assume μ is a probability measure on an Euclidean space X that satisfies $\text{LSI}(\rho)$ for some $\rho > 0$. Then for any probability density f with respect to μ , we have*

$$\left(\int |x|^2 f(x) \mu(dx) \right)^{1/2} \leq \left(\int |x|^2 \mu(dx) \right)^{1/2} + \left(\frac{2}{\rho} \text{Ent}_\mu(f) \right)^{1/2}.$$

Démonstration. Since μ satisfies $\text{LSI}(\rho)$, it also satisfies Talagrand's transport-entropy inequality

$$W_2^2(\mu, \nu) \leq \frac{2}{\rho} \text{Ent}_\mu(\nu)$$

for any probability measure ν , where W_2 is the Wasserstein distance

$$W_2^2(\mu, \nu) = \inf_{\pi \in \Pi} \int |x - y|^2 \pi(dx, dy),$$

where Π is the set of couplings of μ and ν , that is the set of all probability measure π on X^2 such that its first marginal is μ and its second marginal is ν . Proofs of the fact that the LSI implies this transport-entropy inequality (a result which is known as the Otto-Villani theorem) can be found in [OV00] and [Go].

Let π be a coupling of μ and $f\mu$. We have, by the Cauchy-Schwartz inequality

$$\int \langle x, y \rangle \pi(dx, dy) \geq - \left(\int |x|^2 f(x) \mu(dx) \right)^{1/2} \left(\int |x|^2 \mu(dx) \right)^{1/2}$$

so that

$$\begin{aligned} \int |x - y|^2 \pi(dx, dy) &= \int |x|^2 f(x) \mu(dx) + \int |x|^2 \mu(dx) - 2 \int \langle x, y \rangle \pi(dx, dy) \\ &\geq \int |x|^2 f(x) \mu(dx) + \int |x|^2 \mu(dx) - 2 \left(\int |x|^2 f(x) \mu(dx) \right)^{1/2} \left(\int |x|^2 \mu(dx) \right)^{1/2} \\ &= \left(\left(\int |x|^2 f(x) \mu(dx) \right)^{1/2} - \left(\int |x|^2 \mu(dx) \right)^{1/2} \right)^2. \end{aligned}$$

Taking the infimum over $\pi \in \Pi$ and applying the transport-entropy inequality above, we get

$$\left(\int |x|^2 f(x) \mu(dx) \right)^{1/2} - \left(\int |x|^2 \mu(dx) \right)^{1/2} \leq \left(\frac{2}{\rho} \text{Ent}_\mu(f) \right)^{1/2}$$

which, after rearranging the terms, is the desired inequality. \square

Proof of Proposition 2.19. To get (2.64), we use the definition of the microscopic dynamics to get

$$\frac{d}{dt} \int f(t, x) \log f(t, x) \mu(dx) = - \int \frac{\langle A \nabla f, \nabla f \rangle}{f} \mu$$

and integrating yields

$$\int_0^T \int \frac{\langle A \nabla f, \nabla f \rangle}{f} \mu dt = \text{Ent}_\mu(f_0) - \text{Ent}_\mu(f(T, \cdot))$$

which is what we want.

In the same way, to get (2.65), we use (2.21) to get

$$\frac{d}{dt} \bar{H}(\eta(t)) = - \langle \bar{A} \nabla \bar{H}(\eta(t)), \nabla \bar{H}(\eta(t)) \rangle_Y$$

and integrating this yields the desired result.

Finally, we get (2.66) simply by applying Lemma 2.20 and the fact that the relative entropy is decreasing, which means that by (2.64) it holds

$$\text{Ent}_\mu(f(t)) \leq \text{Ent}_\mu(f_0).$$

.

\square

2.5.2 Proof of Theorem 2.9

We shall prove Theorem 2.9 in two steps : first we shall use Theorem 2.8 to show that our data is asymptotically close, in the sense of the H^{-1} norm, to a sequence of deterministic macroscopic vectors given by (2.21), and then we will show that the step functions associated to these vectors converge to the solution of the hydrodynamic equation (2.32).

To apply Theorem 2.8, we consider a sequence $\{M_\ell, N_\ell\}_{\ell=1}^\infty$ of integers with

$$M_\ell \uparrow +\infty, \quad N_\ell \uparrow +\infty, \quad K_\ell := \frac{N_\ell}{M_\ell} \uparrow +\infty. \quad (2.79)$$

To simplify notations, we shall often not explicitly write the dependence of our various objects on ℓ , and just write N , M and K for N_ℓ , M_ℓ and K_ℓ .

Let $\bar{\eta}_0^\ell \in \bar{Y}_M$ be a step function approximation of ζ_0 with

$$\lim_{\ell \uparrow \infty} \|\bar{\eta}_0^\ell - \zeta_0\|_{L^2} = 0. \quad (2.80)$$

Since by assumption ζ_0 lies in $L^p(\mathbb{T})$, we can take a sequence such that

$$\sup \|\bar{\eta}_0^\ell\|_{L^p} \leq C. \quad (2.81)$$

Let η_0^ℓ be the vector associated to $\bar{\eta}_0^\ell$, and consider the solutions η^ℓ of

$$\frac{d\eta^\ell}{dt} = -A\nabla_Y \bar{H}_{N,K,a}(\eta^\ell), \quad \eta^\ell(0) = \eta_0^\ell. \quad (2.82)$$

Then we have the following result, which will be the key to pass from Theorem 2.8 to Theorem 2.9.

Proposition 2.21. *Under the notations above, and for almost every realization of the family of random variables (a_q) , the step functions $\bar{\eta}^\ell$ converge strongly in $L^\infty(H^{-1})$ to the unique weak solution of*

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} \tilde{\varphi}(\zeta), \quad \zeta(0, \cdot) = \zeta_0.$$

We shall defer the proof of this result to Section 5.5, and use it to prove our theorem. It will use several auxiliary results on the behavior of the coarse-grained Hamiltonian \bar{H} , that will be proven in Sections 5.3 and 5.4.

First, let us show that the assumptions of Theorem 2.8 hold for every N and M large enough, with uniform constants.

Assumption (iv) is given by the following lemma :

Lemma 2.22. *Assume that the real numbers a_i satisfy $\sup |a_i| \leq L$. Then there exists a constant α , which only depends on ψ , L and m such that*

$$\int |x|^2 \mu_{N,a,m}(dx) \leq \alpha N.$$

Démonstration. Our proof of this result will be based on Lemma 2.20. We apply its result with reference measure

$$\mu_{N,m}^0 = \frac{1}{Z} \exp\left(-\sum \psi(x_i)\right) dx. \quad (2.83)$$

Note that the measure $\mu_{N,m}^0$ satisfies LSI(ρ) and $f\mu_{N,m}^0 = \mu_{N,a,m}$, so that

$$\begin{aligned} & \left(\int |x|^2 \mu_{N,a,m}(dx)\right)^{1/2} \leq \left(\int |x|^2 \mu_{N,m}^0(dx)\right)^{1/2} \\ & + \left(\frac{2}{\rho} \int \frac{1}{Z} \left(-\sum a_i x_i - \log Z\right) \exp\left(-\sum a_i x_i\right) \mu_{N,m}^0(dx)\right)^{1/2}. \end{aligned}$$

Using this and

$$\begin{aligned} 0 & \leq \text{Ent}_{\mu_{N,m}^0} \left(\frac{1}{Z} \exp\left(-\sum a_i x_i\right)\right) \\ & = \left(\frac{2}{\rho} \int \frac{1}{Z} \left(-\sum a_i x_i - \log Z\right) \exp\left(-\sum a_i x_i\right) \mu_{N,m}^0(dx)\right)^{1/2} \end{aligned}$$

we get

$$\begin{aligned}
\int |x|^2 \mu_{N,a,m}(dx) &\leq 2 \int |x|^2 \mu_{N,m}^0(dx) - \frac{4}{\rho} \int \left(- \sum a_i x_i \right) \mu_{N,a,m}(dx) \\
&\quad - \frac{4}{\rho} \log \left(\int \exp \left(- \sum a_i x_i \right) \mu_{N,m}^0(dx) \right) \\
&\leq 2 \int |x|^2 \mu_{N,m}^0(dx) + \frac{4}{\rho} \left(\sum a_i^2 \right)^{1/2} \left(\int |x|^2 \mu_{N,a,m}(dx) \right)^{1/2} \\
&\quad - \frac{4}{\rho} \int \left(\sum a_i x_i \right) \mu_{N,m}^0(dx) \\
&\leq 2 \int |x|^2 \mu_{N,m}^0(dx) + \frac{4}{\rho} \sqrt{NL} \left(\int |x|^2 \mu_{N,a,m}(dx) \right)^{1/2} \\
&\quad + \frac{4}{\rho} \sqrt{NL} \left(\int |x|^2 \mu_{N,m}^0(dx) \right)^{1/2}. \tag{2.84}
\end{aligned}$$

Applying the Poincaré inequality to $\mu_{N,m}^0(dx)$, we have

$$\int x_i^2 \mu_{N,m}^0(dx) \leq \frac{1}{\rho} \int |\nabla x_i|^2 \mu_{N,m}^0(dx) + \left(\int x_i \mu_{N,m}^0(dx) \right)^2 \leq \frac{1}{\rho} + m^2$$

since, for $\mu_{N,m}^0(dx)$, the coordinates are exchangeable and therefore all have the same mean m . This implies that

$$\int |x|^2 \mu_{N,m}^0(dx) \leq \left(\frac{1}{\rho} + m^2 \right) N$$

and therefore

$$\int |x|^2 \mu_{N,a,m}(dx) \leq CN + C\sqrt{N} \sqrt{\int |x|^2 \mu_{N,a,m}(dx)}.$$

Using Young's inequality on the last term on the right hand side, it is then easy to show that

$$\int |x|^2 \mu_{N,a,m}(dx) \leq \alpha N$$

for some constant α which doesn't depend on N or the a_i . □

Let us now turn to Assumption (i) : Recalling our choice of $f_{0,a}$ (see Theorem 2.9) and the definition (2.83) of the measure μ_N^0 , we have

$$\begin{aligned}
\text{Ent}_{\mu_{N,a}}(f_{0,a}) &= \int \log \left(\frac{dF_N}{d\mu_{N,a}} \right) dF_N \\
&= \int \log \left(\frac{dF_N}{d\mu_N^0} \right) dF_N + \int \log \left(\frac{d\mu_N^0}{d\mu_{N,a}} \right) dF_N \\
&\stackrel{(2.29)}{\leq} CN + \int \left(\sum a_i x_i \right) F_N(dx) - \log \int \exp \left(\sum a_i x_i \right) \mu_{N,a}(dx) \\
&\leq CN + \sqrt{\sum a_i^2} \sqrt{\int |x|^2 F_N(dx)} - \int \left(\sum a_i x_i \right) \mu_{N,a}(dx) \\
&\leq CN + \sqrt{LN} \sqrt{\int |x|^2 F_N(dx)} + \sqrt{LN} \sqrt{\int |x|^2 \mu_{N,a}(dx)}. \tag{2.85}
\end{aligned}$$

We already know by Lemma 2.22 that $\int |x|^2 \mu_{N,a}(dx) \leq CN$. Moreover, we can use Lemma 2.20, the fact that $\text{Ent}_{\mu_N^0}(F_N) \leq CN$ (see (2.29)) and Lemma 2.22 to show that $\int |x|^2 F_N(dx) \leq CN$, and we then obtain assumption (i) of Theorem 2.8 with uniform constant C_1 .

Let us turn to the Assumption (ii) and (iii), which follow from Proposition 2.25. Indeed, recall that

$$\bar{H}_{N,K,a}(y) = \frac{1}{M} \sum_{i=1}^M \psi_{N,K,a}(y_i) + \frac{1}{N} \log Z.$$

Since

$$\frac{1}{N} \log Z = \frac{1}{N} \log \int \exp\left(-\sum \psi(x_i) + a_i x_i\right) dx,$$

it is not hard to deduce from (2.28) and the boundedness of the a_i that this quantity is bounded both from above and from below, so that, in order to prove (ii) and (iii), we can assume without loss of generality that \bar{H} is given by

$$\bar{H}_{N,K,a}(y) = \frac{1}{M} \sum_{i=1}^M \psi_{N,K,a}(y_i).$$

Assumption (ii) then follows from (2.81) and part (iv) of Proposition 2.25 from below. Assumption (iii) also follows from part (iv) of Proposition 2.25.

Finally, Assumption (v) follows directly from Proposition 2.3.

Before turning to the proof of Theorem 2.9, we need two more ingredients : they are estimates comparing the H^{-1} norm to the A^{-1} norm on X_N .

Lemma 2.23. *There exists a constant $C < \infty$ such that for any $x \in X$, if \bar{x} is the associated step function, then*

(i)

$$\frac{1}{C} \langle \bar{x}, \bar{x} \rangle_{H^{-1}} \leq \frac{1}{N} \langle A^{-1}x, x \rangle_X \leq C \langle \bar{x}, \bar{x} \rangle_{H^{-1}};$$

(ii) *If x is bounded in L^2 , then*

$$\left| \langle \bar{x}, \bar{x} \rangle_{H^{-1}} - \frac{1}{N} \langle A^{-1}x, x \rangle_X \right| \leq \frac{C}{N}.$$

These estimates were already included in [GOVW09], but for the sake of completeness we reproduce the proof.

Démonstration. First, we can express the discrete norm as

$$\frac{1}{N} \langle A^{-1}x, x \rangle_X = \frac{1}{N} \sum_{i=1}^N F_i^2$$

$$\text{where } x_i = N(F_{i+1} - F_i) \quad \text{and} \quad \sum_{i=1}^N F_i = 0.$$

To estimate $\langle \bar{x}, \bar{x} \rangle_{H^{-1}}$ through the F_i , recall that by definition

$$\langle \bar{x}, \bar{x} \rangle_{H^{-1}} = \int_{\mathbb{T}} w(\theta)^2 d\theta, \quad \text{where } w' = f \quad \text{and} \quad \int_{\mathbb{T}} w(\theta) d\theta = 0.$$

It is easy to check that we can define such a w with

$$w(\theta) = F_i + N(F_{i+1} - F_i) \left(\theta - \frac{i-1}{N} \right) \quad \text{for } \theta \in \left[\frac{i-1}{N}, \frac{i}{N} \right).$$

Consequently, we have

$$\begin{aligned} \langle \bar{x}, \bar{x} \rangle_{H^{-1}} &= \sum_{i=1}^N \int_0^{N^{-1}} (F_i + N(F_{i+1} - F_i)\theta)^2 d\theta \\ &= \frac{1}{N} \sum_{i=1}^N \left(F_i^2 + (F_{i+1} - F_i)F_i + \frac{1}{3}(F_{i+1} - F_i)^2 \right) \\ &= \frac{1}{N} \langle A^{-1}x, x \rangle_X + \frac{1}{N} \sum_{i=1}^N \left((F_{i+1} - F_i)F_i + \frac{1}{3}(F_{i+1} - F_i)^2 \right). \end{aligned} \quad (2.86)$$

We have the bounds on the second term on the right-hand side

$$-\frac{2}{3N} \sum_{i=1}^N F_i^2 \leq \frac{1}{N} \sum_{i=1}^N \left((F_{i+1} - F_i)F_i + \frac{1}{3}(F_{i+1} - F_i)^2 \right) \leq 0$$

so that

$$\langle \bar{x}, \bar{x} \rangle_{H^{-1}} \leq \frac{1}{N} \langle A^{-1}x, x \rangle_X \leq 3 \langle \bar{x}, \bar{x} \rangle_{H^{-1}}.$$

Moreover, if \bar{x} is bounded in L^2 , then

$$\frac{1}{N} \sum_{i=1}^N x_i^2 = \sum_{i=1}^N N(F_{i+1} - F_i)^2 \leq C \quad \text{and} \quad \frac{1}{N} \sum_{i=1}^N F_i^2 \leq C,$$

so we get from (2.86)

$$\begin{aligned} &\left| \langle \bar{x}, \bar{x} \rangle_{H^{-1}} - \frac{1}{N} \langle A^{-1}x, x \rangle_X \right| \\ &= \frac{1}{N} \left| \sum_{i=1}^N \left((F_{i+1} - F_i)F_i + \frac{1}{3}(F_{i+1} - F_i)^2 \right) \right| \\ &\leq \frac{1}{N} \left(\sum_{i=1}^N (F_{i+1} - F_i)^2 \right)^{1/2} \left(\sum_{i=1}^N F_i^2 \right)^{1/2} + \frac{1}{3N} \sum_{i=1}^N (F_{i+1} - F_i)^2 \\ &\leq \frac{C}{N} \end{aligned}$$

□

We are now ready to give the proof of our main result.

Proof of Theorem 2.9. First, we have

$$\begin{aligned} &\lim_{\ell \uparrow \infty} \sup_{t \in [0, T]} \int_{X_N} \|\bar{x} - \zeta(t, \cdot)\|_{H^{-1}}^2 f(t, x) \mu_{N, a}(dx) \\ &\leq 2 \lim_{\ell \uparrow \infty} \sup_{t \in [0, T]} \int_{X_N} \|\bar{x} - \bar{\eta}_\ell(t)\|_{H^{-1}}^2 f(t, x) \mu_{N, a}(dx) + 2 \lim_{\ell \uparrow \infty} \sup_{t \in [0, T]} \|\bar{\eta}_\ell(t) - \zeta(t, \cdot)\|_{H^{-1}}^2. \end{aligned}$$

Applying Proposition 2.21, we immediately see that the second term on the right-hand side of this equation goes to zero almost surely. For the first term, using bound (i) of Lemma 2.23, we see that, for any realization of the random field,

$$\begin{aligned} & \int \frac{1}{N} \langle (x - NP^t \eta_{0,a}), A^{-1}(x - NP^t \eta_{0,a}) \rangle F_N(dx) \\ & \leq C \int_{X_N} \|\bar{x} - \bar{\eta}_{\ell,0}\|_{H^{-1}}^2 f_{0,a} \mu_{N,a}(dx) \longrightarrow 0 \end{aligned}$$

so that an application of Theorem 2.8 yields that

$$\lim_{\ell \uparrow \infty} \sup_{t \in [0, T]} \int \frac{1}{N} \langle (x - NP^t \eta_{0,a}), A^{-1}(x - NP^t \eta_{\ell,a}(t)) \rangle f(t, x) \mu_{N,a}(dx) = 0$$

for any realization of the random field. Another application of bound (i) in Lemma 2.23 then yields

$$\lim_{\ell \uparrow \infty} \sup_{t \in [0, T]} \int_{X_N} \|\bar{x} - \bar{\eta}_\ell(t)\|_{H^{-1}}^2 f(t, x) \mu_{N,a}(dx) = 0$$

also for any realization of the random field, which concludes the proof.

The remainder of this section is devoted to the proof of Proposition 2.21

□

2.5.3 Auxiliary result : Polynomial bounds on the energy

In this section, our aim will be to obtain polynomial bounds on the macroscopic and hydrodynamic free energies $\psi_{N,K,i,a}$, $\varphi_{N,K,i,a}$ and $\tilde{\varphi}$ (for the definitions see (2.87), (2.88), and (2.33) respectively). The main tool will be the following Lemma, due to Caputo (see Lemma 2.4 in [Cap03]).

Lemma 2.24. *Let*

$$s(\sigma)^2 := \int (x - m)^2 \mu_\sigma(dx)$$

where $\mu_\sigma(dx) = Z^{-1} \exp(\sigma x - \psi(x)) dx$ and $m = \int x \mu_\sigma(dx)$. Then

$$\frac{1}{C \psi_c''(m)} \leq s(\sigma)^2 \leq \frac{C}{\psi_c''(m)}.$$

For a proof of this statement we refer to [Cap03].

We define

$$\psi_{N,K,i,a}(m) := -\frac{1}{K} \log \int_{X_{K,m}} \exp \left(- \sum_{j=(i-1)K+1}^{iK} a_{j/N} x_j + \psi(x_j) \right) dx \quad (2.87)$$

and

$$\varphi_{N,K,i,a}(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \log \int_{\mathbb{R}} \exp((\sigma - a_{j/N})x - \psi(x)) dx \right) \quad (2.88)$$

Theorem 2.34 from the appendix tells us that these two functions behave in the same manner. We will deduce bounds on $\varphi_{N,K,i,a}$ from Lemma 2.24, which will then carry on to $\psi_{N,K,i,a}$. Since these bounds will only depend on K and $\sup |a_i|$, we simplify notations by writing ψ_K instead of $\psi_{N,K,i,a}$ and φ_K instead of $\varphi_{N,K,i,a}$ in this section.

Proposition 2.25. *There exists K_0 and $C > 0$ which only depend on ψ and $\sup a_i$ such that, for any $m \in \mathbb{R}$ and $K \geq K_0$*

- (i) $\frac{1}{C}(1 + |m|^{p-2}) \leq \varphi_K''(m) \leq C(1 + |m|^{p-2})$;
- (ii) $\frac{1}{C}(|m|^{p-1} - 1) \leq |\varphi_K'(m)| \leq C(1 + |m|^{p-1})$;
- (iii) $\frac{1}{C}(|m|^p - 1) \leq \varphi_K(m) \leq C(1 + |m|^p)$;
- (iv) $\frac{1}{C}(|m|^p - 1) \leq \psi_K(m) \leq C(|m|^p - 1)$.

Moreover, bounds (i), (ii) and (iii) are also valid for $\tilde{\varphi}$.

Démonstration. We start with proving these bounds for φ_K . We only need to prove (i), since (ii) and (iii) directly follow by integrating (i) and using some uniform bounds on the value at $m = 0$, which directly follow from the boundedness of the random field.

We start with showing that

$$\varphi_K''(m) = \left(\frac{1}{K} s(\sigma - a_{j/N})^2 \right)^{-1}, \quad (2.89)$$

where we used $\sigma = \varphi_K'(m)$ and

$$\frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \int x \mu_{\sigma - a_{j/N}}(dx) = m.$$

Indeed, because φ_K is the Legendre transform of

$$\frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \log \int_{\mathbb{R}} \exp((\sigma - a_{j/N})x - \psi(x)) dx$$

we get that by fundamental properties of the Legendre transform

$$\varphi_K''(m) = \left(\frac{d^2}{ds^2} \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \log \int_{\mathbb{R}} \exp((\sigma - a_{j/N})x - \psi(x)) dx \right)^{-1}.$$

Straightforward calculation yields

$$\frac{d^2}{ds^2} \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \log \int_{\mathbb{R}} \exp((\sigma - a_{j/N})x - \psi(x)) dx = \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} s(\sigma - a_{j/N})^2,$$

which satisfies the desired identity (2.89).

It follows from the boundedness of the random field that

$$\frac{1}{C} s(\sigma_0) \leq s(\sigma - a_{j/N}) \leq C s(\sigma_0)$$

where σ_0 is such that $\sum \int x \mu_{\sigma_0}(dx) = m$ and C only depends on ψ and $L = \sup_j |a_{j/N}|$ (see the proof of (2.123) in the appendix). Therefore

$$\frac{1}{Cs(\sigma_0)^2} \leq \varphi_K''(m) \leq \frac{C}{s(\sigma_0)^2},$$

which yields statement (i) by Lemma 2.24 and Assumption (2.28).

Let us now turn to the verification of (i), (ii), and (iii) for the function $\tilde{\varphi}$. We recall the definition (2.33) of $\tilde{\varphi}$, namely

$$\tilde{\varphi}(m) = \sup_{\sigma \in \mathbb{R}} \left\{ \sigma m - \mathbb{E}^a \left[\log \int_{\mathbb{R}} \exp((\sigma - a)x - \psi(x)) dx \right] \right\}.$$

Because $\tilde{\varphi}$ is again a Legendre transform one can use the same argument as for φ_K to obtain similar bounds.

Let us now turn to the verification of (iv). By the local Cramér Theorem (cf. Theorem 2.34 in the appendix) one can directly transfer the polynomial bound of φ_K to the function ψ_K .

□

2.5.4 Auxiliary result : Convergence of the free energy

In the previous section we derived polynomial bounds on the macroscopic and hydrodynamic free energies $\psi_{N,K,i,a}$, $\varphi_{N,K,i,a}$ and $\tilde{\varphi}$ (for the definitions see (2.87), (2.88) and (2.33) respectively). In this section we will show that the macroscopic free energy $\varphi_{N,K,i,a}$ converges to the hydrodynamic free energy $\tilde{\varphi}$.

We recall that, following (2.15), we have

$$\bar{H}_{N,K,a}(y) = \frac{1}{M} \sum_{i=1}^M \psi_{N,K,a}(y_i) + \frac{1}{N} \log \bar{Z} \quad (2.90)$$

By Theorem 2.34, we know that, for any compact subset E of \mathbb{R} , we have

$$\sup \{ |\psi_{N,K,i,a}(m) - \varphi_{N,K,i,a}(m)| ; m \in E, \sup |a_q| \leq L \} \longrightarrow 0 \quad \text{as } N, K \rightarrow \infty. \quad (2.91)$$

This implies that, in order to study the asymptotic behavior of $\psi_{N,K,i,a}$, we can study the behavior of $\varphi_{N,K,i,a}$.

Proposition 2.26. *For almost every realization of the random variables a_q , we have for every compact subset E of \mathbb{R} , and for any sequence (i_ℓ) with $i_\ell \in \{1, \dots, M_\ell\}$,*

$$\varphi_{N,K,i,a}(m) \longrightarrow \tilde{\varphi}(m) \quad (2.92)$$

uniformly in $m \in E$. Moreover,

$$\mathbb{E}[\varphi_{N,K,i,a}(m)] \xrightarrow{K \uparrow \infty} \tilde{\varphi}(m) \quad (2.93)$$

and

$$\tilde{\varphi}(m) = \inf_K \mathbb{E}[\varphi_{N,K,i,a}(m)]. \quad (2.94)$$

Démonstration. Due to the polynomial bounds on φ_K and $\tilde{\varphi}$, we can show that the optimum in

$$\varphi_{N,K,i,a}(m) = \sup_{\sigma} \left(\sigma m - \varphi_{N,K,i,a}^*(\sigma) \right)$$

is reached for $\sigma = \varphi'_{N,K,i,a}(m)$ which satisfies the bound (see statement (iii) of Proposition 2.25)

$$|\sigma| \leq C(1 + |m|^{p-1})$$

with a constant C uniform in K , which yields

$$|\sigma|^q \leq C(1 + |m|^p),$$

after taking the $q = \frac{p}{p-1}$ -th power. Therefore we can localize the variational formulation as

$$\varphi_{N,K,i,a}(m) = \sup_{|\sigma|^q \leq C(1+|m|^p)} \left(\sigma m - \varphi_{N,K,i,a}^*(\sigma) \right).$$

We have the same localization property for $\tilde{\varphi}$, and therefore

$$|\varphi_{N,K,i,a}(m) - \tilde{\varphi}(m)| \tag{2.95}$$

$$= \left| \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \varphi_{N,K,i,a}^*(\sigma)) - \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \tilde{\varphi}^*(\sigma)) \right| \tag{2.96}$$

. We will now show that

$$\begin{aligned} & \left| \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \varphi_{N,K,i,a}^*(\sigma)) - \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \tilde{\varphi}^*(\sigma)) \right| \\ & \leq \sup_{|\sigma|^q \leq C(1+|m|^p)} |\varphi_{N,K,i,a}^*(\sigma) - \tilde{\varphi}^*(\sigma)|. \end{aligned}$$

Indeed, let us assume that (the other case works with the same argument)

$$\begin{aligned} & \left| \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \varphi_{N,K,i,a}^*(\sigma)) - \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \tilde{\varphi}^*(\sigma)) \right| \\ & = \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \varphi_{N,K,i,a}^*(\sigma)) - \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \tilde{\varphi}^*(\sigma)). \end{aligned}$$

By elementary properties of the Legendre transform, there is $\tilde{\sigma}$ such that

$$\sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \varphi_{N,K,i,a}^*(\sigma)) = (\tilde{\sigma} m - \varphi_{N,K,i,a}^*(\tilde{\sigma})).$$

Then we get

$$\begin{aligned} & \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \varphi_{N,K,i,a}^*(\sigma)) - \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \tilde{\varphi}^*(\sigma)) \\ & \leq (\tilde{\sigma} m - \varphi_{N,K,i,a}^*(\tilde{\sigma})) - (\tilde{\sigma} m - \tilde{\varphi}^*(\tilde{\sigma})) \\ & = \tilde{\varphi}^*(\tilde{\sigma}) - \varphi_{N,K,i,a}^*(\tilde{\sigma}) \\ & \leq \sup_{|\sigma|^q \leq C(1+|m|^p)} |\varphi_{N,K,i,a}^*(\sigma) - \tilde{\varphi}^*(\sigma)|. \end{aligned}$$

Hence we get overall that

$$\begin{aligned} & |\varphi_{N,K,i,a}(m) - \tilde{\varphi}(m)| \\ & = \left| \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \varphi_{N,K,i,a}^*(\sigma)) - \sup_{|\sigma|^q \leq C(1+|m|^p)} (\sigma m - \tilde{\varphi}^*(\sigma)) \right| \\ & \leq \sup_{|\sigma|^q \leq C(1+|m|^p)} |\varphi_{N,K,i,a}^*(\sigma) - \tilde{\varphi}^*(\sigma)| \\ & = \sup_{|\sigma|^q \leq C(1+|m|^p)} \left| \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \varphi^*(\sigma - a_{j/N}) - \mathbb{E}[\varphi^*(\sigma - a)] \right| \end{aligned}$$

Assume now that $|m| \leq m_0$ for some positive number m_0 . Then we have that

$$\begin{aligned} & \sup_{|m| \leq m_0} |\varphi_{N,K,i,a}(m) - \tilde{\varphi}(m)| \\ & \leq \sup_{|m| \leq m_0} \sup_{|\sigma|^q \leq C(1+|m|^p)} \left| \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \varphi^*(\sigma - a_{j/N}) - \mathbb{E}[\varphi^*(\sigma - a)] \right| \\ & = \sup_{|\sigma|^q \leq C(1+m_0^p)} \left| \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \varphi^*(\sigma - a_{j/N}) - \mathbb{E}[\varphi^*(\sigma - a)] \right|. \end{aligned}$$

If we denote by ν is the distribution of the random variable a , and $\nu^{N,K,i,a}$ the (random) probability measure given by

$$\nu^{N,K,i,a}(dx) := \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \delta_{a_j},$$

we have

$$\begin{aligned} & \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \varphi^*(\sigma - a_{j/N}) - \mathbb{E}[\varphi^*(\sigma - a)] \\ &= \int \varphi^*(\sigma - x) \nu^{N,K,i,a}(dx) - \int \varphi^*(\sigma - x) \nu(dx) \\ &= \int (\varphi^*(\sigma - x) - \varphi^*(\sigma - y)) \pi(dx, dy), \end{aligned}$$

where $\nu^{N,K,i,a}$ is the empirical measure associated to the a_i and $\pi * (da, d\tilde{a})$ is an arbitrary coupling of the measures $\nu^{N,K,i,a}(dx)$ and $\nu(d\tilde{y})$.

Our main argument will be that this random measure converges almost surely to ν (this is a variant of the strong law of large numbers).

The function $a \rightarrow \varphi^*(\sigma - a)$ is $C(L)(1 + |\sigma|^{q-1})$ -Lipschitz on $[-L, L]$. This follows from the fact that $(\varphi^*)'(\sigma) \leq C(1 + |\sigma|^{q-1})$, which we deduce now. Indeed, by duality of the Legendre transform we have $(\varphi^*)'(\sigma) = m$. By part (iii) of Proposition 2.25, we have that

$$|\sigma| \geq \frac{1}{C} (|m|^{p-1} - 1),$$

which yields

$$|(\varphi^*)'(\sigma)| = |m| \leq C(|\sigma| + 1)^{\frac{1}{p-1}} \leq C(|\sigma|^{\frac{1}{p-1}} + 1),$$

which yields the bound $(\varphi^*)'(\sigma) \leq C(1 + |\sigma|^{q-1})$ by observing that $q-1 = \frac{p}{p-1} - \frac{p-1}{p-1} = \frac{1}{p-1}$. By this fact we can use the Kantorovitch-Rubinstein duality formula to obtain

$$\begin{aligned} & \sup_{|\sigma|^q \leq C(1+|m|^p)} \left| \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \varphi^*(\sigma - a_{j/N}) - \mathbb{E}[\varphi^*(\sigma - a)] \right| \\ & \leq C(1 + |m|^p)^{(q-1)/q} W_1(\nu^{N,K,i,a}, \nu) \\ & = C(1 + |m|) W_1(\nu^{N,K,i,a}, \nu), \end{aligned} \tag{2.97}$$

where $W_1(\nu^{N,K,i,a}, \nu)$ denotes the L^1 -Wasserstein distance between the measures $\nu^{N,K,i,a}$ and ν . By Varadarajan's Theorem (see section 11.4 in [Dud02]), almost surely $W_1(\nu^K, \nu) \rightarrow 0$, and therefore almost surely, $\varphi_{N,K,i,a}(m)$ converges to $\tilde{\varphi}(m)$ uniformly on bounded sets. This proves (2.92).

Since the a_q are bounded random variables, $\varphi_{N,K,i,a}(m)$ is also bounded, and an application of the Dominated Convergence Theorem yields (2.93).

Finally, we prove (2.94) : since the expectation of $\varphi_{N,K,i,a}(m)$ only depends on K , we drop the subscript and consider

$$\varphi_K(m) = \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \frac{1}{K} \sum_{j=1}^K \varphi^*(\sigma - a_j) \right),$$

with the a_i a sequence of iid random variables. We then have, for any $K_1, K_2 \in \mathbb{N}$

$$\begin{aligned} (K_1 + K_2)\varphi_{K_1+K_2}(m) &= (K_1 + K_2)\sup_{\sigma \in \mathbb{R}} \left(\sigma m - \frac{1}{K_1 + K_2} \sum_{j=1}^{K_1+K_2} \varphi^*(\sigma - a_j) \right) \\ &= \sup_{\sigma \in \mathbb{R}} \left((K_1 + K_2)\sigma m - \sum_{j=1}^{K_1+K_2} \varphi^*(\sigma - a_j) \right) \\ &\leq \sup_{\sigma \in \mathbb{R}} \left(K_1 \sigma m - \sum_{j=1}^{K_1} \varphi^*(\sigma - a_j) \right) \\ &\quad + \sup_{\sigma \in \mathbb{R}} \left(K_2 \sigma m - \sum_{j=K_1+1}^{K_1+K_2} \varphi^*(\sigma - a_j) \right) \end{aligned}$$

So that

$$(K_1 + K_2)\mathbb{E}[\varphi_{K_1+K_2}(m)] \leq K_1\mathbb{E}[\varphi_{K_1}(m)] + K_2\mathbb{E}[\varphi_{K_2}(m)].$$

Therefore $K \mathbb{E}[\varphi_K(m)]$ is a sub-additive sequence, and an application of the sub-additivity theorem yields

$$\mathbb{E}[\varphi_K(m)] \longrightarrow \inf_K \mathbb{E}[\varphi_K(m)].$$

Since $\mathbb{E}[\varphi_K(m)]$ also converges to $\tilde{\varphi}(m)$, this implies (2.94). □

We will also need convergence of the free energy when taking the average along some L^p function :

Lemma 2.27. *Let $c(K) := \mathbb{E}[W_1(\nu^{N,K,i,a}, \nu)^q]$, which only depends on K , and assume that*

$$\sum_{\ell} c(K_{\ell}) < \infty \tag{2.98}$$

for some $\epsilon > 0$. Then, almost surely, for any sequence of adapted step functions $\bar{\zeta}_{\ell}$ that is bounded in L^p , we have

$$\lim_{\ell} \frac{1}{M} \sum_{i=1}^M |\varphi_{N,K,i,a}(\zeta_{\ell,i}) - \tilde{\varphi}(\zeta_{\ell,i})| = 0.$$

(By adapted step function, we mean that $\bar{\zeta}_{\ell}$ is the step function associated to an element of Y_{ℓ} , ie that its mesh size is $1/M_{\ell}$.)

Remark 2.28. Note that, since $W_1(\nu^{N,K,i,a}, \nu)$ almost surely converges to zero and is a bounded random variable (since the a_q are bounded), by the Dominated Convergence Theorem, $c(K) \longrightarrow 0$ as K goes to infinity.

Démonstration. With the same arguments as in the proof of Proposition 2.26, we have

$$\frac{1}{M} \sum_{i=1}^M |\varphi_{N,K,i,a}(\zeta_{\ell,i}) - \tilde{\varphi}(\zeta_{\ell,i})| \leq \frac{1}{M} \sum_{i=1}^M C(1 + |\zeta_{\ell,i}|)W_1(\nu^{N,K,i,a}, \nu). \tag{2.99}$$

Using Holder's inequality, this yields

$$\begin{aligned}
& \frac{1}{M} \sum_{i=1}^M |\varphi_{N,K,i,a}(\zeta_{\ell,i}) - \tilde{\varphi}(\zeta_{\ell,i})| \\
& \leq \left(\frac{1}{M} \sum_{i=1}^M C(1 + |\zeta_{\ell,i}|)^p \right)^{1/p} \left(\frac{1}{M} \sum_{i=1}^M W_1(\nu^{N,K,i,a}, \nu)^q \right)^{1/q} \\
& \leq C(1 + \|\bar{\zeta}_{\ell}\|_p^p) \left(\frac{1}{M} \sum_{i=1}^M W_1(\nu^{N,K,i,a}, \nu)^q \right)^{1/q}. \tag{2.100}
\end{aligned}$$

Therefore, all we need to show is that $\frac{1}{M} \sum_{i=1}^M W_1(\nu^{N,K,i,a}, \nu)^q$ converges almost surely to zero. By Markov's inequality, for any $\epsilon > 0$, we have

$$\begin{aligned}
\mathbb{P} \left[\frac{1}{M} \sum_{i=1}^M W_1(\nu^{N,K,i,a}, \nu)^q > \epsilon \right] & \leq \frac{1}{\epsilon} \mathbb{E} \left[\frac{1}{M} \sum_{i=1}^M W_1(\nu^{N,K,i,a}, \nu)^q \right] \\
& = \frac{c(K)}{\epsilon}
\end{aligned}$$

Applying the Borell-Cantelli Lemma with assumption (2.98) then yields the result. \square

2.5.5 Proof of Proposition 2.21

Our proof will follow the same structure as the one in [GOVW09]. The main differences are the use of Proposition 2.26, which is needed because of the additional random linear term in the Hamiltonian, and the use of L^p bounds on the data (rather than L^2). They mostly appear in the proof of Lemma 5.14. The proof will rely on the following five lemmas.

Lemma 2.29. *Consider the sequence (η^ℓ) of solutions of (2.82), subject to (2.81). There exists $C < \infty$ (independent of ℓ) such that*

$$\sup_{0 \leq t \leq T} \|\eta^\ell(t)\|_{L^p} \leq C \tag{2.101}$$

$$\int_0^T \left\langle \frac{d\eta^\ell}{dt}(t), \bar{A}^{-1} \frac{d\eta^\ell}{dt}(t) \right\rangle_Y dt \leq C \tag{2.102}$$

In particular, (2.101) implies that the η^ℓ are uniformly bounded in $L^\infty(L^p)$, and therefore that, for the sequence of associated step functions $\bar{\eta}^\ell$, there is a subsequence such that

$$\bar{\eta}^\ell \longrightarrow \eta_* \text{ weak-}^* \text{ in } L^\infty(L^p) = (L^1(L^q))^*$$

for some limit η_* . We will now state several lemmas, which we will use to show that η_* must be the unique weak solution of (2.32).

Lemma 2.30. *Any limit η_* of η^ℓ for the weak- * topology on $L^\infty(L^p)$, we have*

$$\eta_* \in L^\infty(L^p), \quad \frac{\partial \eta_*}{\partial t} \in L^2(H^{-1}), \quad \tilde{\varphi}'(\eta_*) \in L^\infty(L^q). \tag{2.103}$$

Lemma 2.31 (Inequality formulation for convex potential). *Assume $\bar{H}_{N,K,a}$ is convex. Then η is a solution of (2.82) iff for all $\xi \in Y_M$ and smooth $\beta : [0, T] \rightarrow \mathbb{R}_+$ with $\beta(0) = \beta(T) = 0$, we have*

$$\int_0^T \bar{H}_{N,K,a}(\eta)\beta(t)dt \leq \int_0^T \bar{H}_{N,K,a}(\eta + \xi)\beta(t)dt - \int_0^T \langle \xi, (\bar{A})^{-1}\eta \rangle_Y \dot{\beta}(t)dt. \quad (2.104)$$

Similarly, ζ satisfies (2.32) if and only if

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}} \tilde{\varphi}(\zeta(t, \theta))\beta(t)d\theta dt \\ & \leq \int_0^T \int_{\mathbb{T}} \tilde{\varphi}(\zeta(t, \theta) + \xi(\theta))\beta(t)d\theta dt - \int_0^T \langle \xi(\cdot), \zeta(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t)dt \end{aligned} \quad (2.105)$$

for all $\xi \in L^p(\mathbb{T})$ and smooth $\beta : [0, T] \rightarrow [0, \infty)$ that takes value 0 at 0 and T .

Lemma 2.32. *Suppose that the sequence $\{\eta_\ell\}$ satisfies (2.32), (2.101) and (2.102) and consider a subsequence such that the associated step functions weak- $*$ converge in $(L^1(L^q))^*$ to a limit η_* . Moreover, assume that (2.98) holds.*

Let $\xi_\ell := \pi_\ell(\xi + \eta_*) - \eta_\ell$, where ξ is an arbitrary L^p function and π_ℓ is the L^p projection onto elements of Y_ℓ . Then we have for almost every realization of the random field a

(i)

$$\liminf_{\ell \uparrow \infty} \int_0^T (\bar{H}_{N,K,a}(\eta_\ell(t)) - \frac{1}{N} \log \bar{Z}_{N,K,a})\beta(t)dt \geq \int_0^T \beta(t) \int_{\mathbb{T}} \tilde{\varphi}(\eta_*(t, \theta))d\theta dt;$$

(ii)
$$\limsup_{\ell \uparrow \infty} \int_0^T (\bar{H}_{N,K,a}(\eta_\ell(t) + \xi_\ell(t)) - \frac{1}{N} \log \bar{Z}_{N,K,a})\beta(t)dt \leq \int_0^T \beta(t) \int_{\mathbb{T}} \tilde{\varphi}(\eta_*(t, \theta) + \xi(\theta))d\theta dt;$$

(iii)

$$\lim_{\ell \uparrow \infty} \int_0^T \langle \xi_\ell(t), (\bar{A})^{-1}\eta_\ell(t) \rangle_Y \dot{\beta}(t)dt = \int_0^T \langle \xi(\cdot), \eta_*(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t)dt. \quad (2.106)$$

Lemma 2.33. *Equation (2.32) has at most one weak solution with initial condition ζ_0 .*

Using these, we can prove Proposition 2.21. The main difference with [GOVW09] will be the proof of Lemma 2.32.

Proof of Proposition 2.21. As stated earlier, by applying Lemma 2.29, we can extract a subsequence such that

$$\bar{\eta}^\ell \longrightarrow \eta_* \text{ weak-}^* \text{ in } L^\infty(L^p) = (L^1(L^q))^*.$$

We want to use Lemma 2.32, so we need to check (2.98). But we already know that $c(K_\ell) \rightarrow 0$ (see Remark 2.28), so we can further extract a subsequence such that (2.98) is satisfied, without changing the limit. Since all we care about here is to identify the limit η_* , there is no loss of generality here.

According to Lemma 2.31, η^ℓ satisfies (2.104). By applying Lemma 2.32 to this inequality, we get that η_* satisfies (2.105), so that another application of Lemma 2.31 and of Lemma 2.30 yields that η_* is a solution of the hydrodynamic equation. The unicity result of Lemma 2.33 then guarantees that the full sequence converges to the unique weak solution of (2.32). \square

We will now prove the previous Lemmas.

Proof of Lemma 2.29. The proof is exactly the same as in Lemma 34 of [GOVW09].

Argument for (2.101) : By (2.65) it holds

$$\bar{H}_K(\eta^\ell(t)) \leq \bar{H}_K(\eta^\ell(0)) \stackrel{(2.15)}{=} \frac{1}{M} \sum_{i=1}^M \psi_{K, a_{(i-1)K+1}, \dots, a_{iK}}(y_i) - \frac{1}{N} \log \bar{Z}. \quad (2.107)$$

Now, the desired bound (2.101) follows from a combination of the polynomial bounds on $\psi_{K, a_{(i-1)K+1}, \dots, a_{iK}}$ of Proposition 2.25 and (2.81).

Argument for (2.102) : Using By (2.65) yields

$$\int_0^T \left\langle \frac{d\eta^\ell}{dt}(t), \bar{A}^{-1} \frac{d\eta^\ell}{dt}(t) \right\rangle_Y dt = \bar{H}_K(\eta^\ell(0)) - \bar{H}_K(\eta^\ell(t)).$$

The first term on the right hand side is bounded as in (2.107). To bound the second term, observe that

$$\inf_y \bar{H}_K(y) \geq -C,$$

uniformly, which is a consequence of the uniform convergence of ψ_K to φ_K , which is a strictly convex function. \square

Proof of Lemma 2.30. By weak lower semicontinuity, and recalling (2.101) we have for all $t \in [0, T]$ that

$$\|\eta_*(t, \cdot)\|_{L^p} \leq \liminf \|\eta^\ell(t)\|_{L^p} \leq C.$$

The first estimate is therefore proved. The third one immediately follows, since $\tilde{\varphi}'$ satisfies the polynomial bound $\tilde{\varphi}'(m) \leq C(1 + |m|^{p-1})$ and $q = p/(p-1)$. For the second estimate, we use Lemma 2.23, and, recalling that $NP^t \eta^\ell = \bar{\eta}^\ell$, we have

$$\begin{aligned} \int_0^T \langle \dot{\bar{\eta}}^\ell, \dot{\bar{\eta}}^\ell \rangle_{H^{-1}} dt &\leq \frac{C}{N} \int_0^T \langle NP^t \dot{\eta}^\ell, A^{-1} NP^t \dot{\eta}^\ell \rangle_X dt \\ &= C \int_0^T \langle \dot{\eta}^\ell, \bar{A}^{-1} \dot{\eta}^\ell \rangle_Y dt \leq C \end{aligned}$$

where the last inequality used (2.102). Once more, by weak lower semicontinuity, we then have

$$\int_0^T \langle \dot{\eta}_*, \dot{\eta}_* \rangle_{H^{-1}} dt \leq C. \quad \square$$

Proof of Lemma 2.33. Once more, this is exactly the same as in Lemma 38 of [GOVW09]. Consider two solutions ζ_1 and ζ_2 with same initial condition. We have the weak formulation

$$\langle \dot{\zeta}_i, \xi \rangle_{H^{-1}} = - \int_{\mathbb{T}} \tilde{\varphi}'(\zeta_i) \xi d\theta \text{ for all } \xi \in L^2, \text{ for a.e. } t \in [0, T],$$

so that

$$\langle (\dot{\zeta}_1 - \dot{\zeta}_2), \xi \rangle_{H^{-1}} = - \int_{\mathbb{T}} (\tilde{\varphi}'(\zeta_1) - \tilde{\varphi}'(\zeta_2)) \xi d\theta.$$

Since both ζ_1 and ζ_2 lie in $L^\infty(L^p)$, we can take $\zeta_1 - \zeta_2$ as a test function, so that for a.e. $t \in [0, T]$ we have

$$\frac{d}{dt} \langle (\zeta_1 - \zeta_2), (\zeta_1 - \zeta_2) \rangle_{H^{-1}} = -2 \int_{\mathbb{T}} (\tilde{\varphi}'(\zeta_1) - \tilde{\varphi}'(\zeta_2)) (\zeta_1 - \zeta_2) d\theta \leq 0$$

by convexity of $\tilde{\varphi}$. This implies that $\zeta_1 = \zeta_2$. \square

Proof of Lemma 2.31. Both parts of this lemma are proved in the same way, so we shall only prove (2.105). Let ζ be a bounded solution of (2.32). First let us rewrite the equation as

$$\int_0^T \langle \xi(\cdot), \zeta(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t) dt = \int_0^T \int_{\mathbb{T}} \varphi'(\zeta(t, \theta)) \xi(\theta) d\theta \beta(t) dt \quad (2.108)$$

for all $\xi \in L^2$ and smooth $\beta : [0, T] \rightarrow \mathbb{R}_+$. By density of L^p in L^2 , it is enough if (2.108) holds for all $\xi \in L^p$.

Let us show that (2.108) implies (2.105). By convexity of φ , we have

$$\xi(\theta) \varphi'(\eta_*(\theta)) \leq -\varphi(\eta_*(\theta)) + \varphi(\eta_*(\theta) + \xi(\theta)).$$

Inserting this inequality into (2.108) gives

$$\int_0^T \langle \xi(\cdot), \zeta(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t) dt \leq \int_0^T \int_{\mathbb{T}} \varphi(\zeta(t, \theta) + \xi(\theta)) d\theta \beta(t) dt - \int_0^T \int_{\mathbb{T}} \varphi(\zeta(t, \theta)) d\theta \beta(t) dt$$

which, after rearranging terms, is (2.105).

Let us now show that (2.105) implies (2.108). We substitute ξ for $\epsilon \xi$ in (2.105), for some $\epsilon > 0$. Dividing both sides by ϵ and rearranging then yields

$$\int_0^T \langle \xi(\cdot), \zeta(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t) dt \leq \int_0^T \int_{\mathbb{T}} \frac{\varphi(\zeta + \epsilon \xi) - \varphi(\zeta)}{\epsilon} \beta(t) d\theta dt.$$

Taking the limit $\epsilon \rightarrow 0$ then gives us

$$\int_0^T \langle \xi(\cdot), \zeta(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t) dt \leq \int_0^T \int_{\mathbb{T}} \varphi'(\zeta(t, \theta)) \xi(\theta) d\theta \beta(t) dt.$$

Repeating this process with $-\epsilon \xi$ instead of $\epsilon \xi$ returns

$$\int_0^T \langle \xi(\cdot), \zeta(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t) dt \geq \int_0^T \int_{\mathbb{T}} \varphi'(\zeta(t, \theta)) \xi(\theta) d\theta \beta(t) dt,$$

so that (2.108) holds. □

Proof of Lemma 2.32. Since the proof of (iii) is exactly the same as the proof of (iii) in Lemma 37 of [GOVW09], we will skip its proof, and concentrate on (i) and (ii).

In order to prove (i), it is enough to show that

$$\liminf \frac{1}{M} \sum_{i=1}^M \varphi_{N,K,a,i}(\eta_i(t)) \geq \int_{\mathbb{T}} \tilde{\varphi}(\eta_*(t, \theta)) d\theta, \quad (2.109)$$

because it already follows from Theorem 2.34 that, for a sequence that is bounded in L^p , we can replace $\psi_{N,K,a,i}$ with $\varphi_{N,K,a,i}$, and since an application of Fatou's Lemma would then immediately yield (i). Since we have

$$\begin{aligned} \frac{1}{M} \sum \varphi_{N,K,i,a}(\eta_i(t)) &= \frac{1}{M} \sum_{i=1}^M \sup_{\sigma \in \mathbb{R}} \left(\sigma \eta_i - \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} \varphi^*(\sigma - a_{i/N}) \right) \\ &\geq \sup_{\sigma \in \mathbb{R}} \left(\sigma \frac{1}{M} \sum_{i=1}^M \eta_i - \frac{1}{N} \sum_{j=1}^N \varphi^*(\sigma - a_{i/N}) \right) \\ &= \varphi_{N,a} \left(\frac{1}{M} \sum_{i=1}^M \eta_i \right) \longrightarrow \tilde{\varphi} \left(\int_{\mathbb{T}} \eta_*(\theta) d\theta \right) \end{aligned} \quad (2.110)$$

almost surely, we have

$$\liminf \frac{1}{M} \sum_{i=1}^M \varphi_{N,K,i,a}(\eta_i(t)) \geq \tilde{\varphi} \left(\int_{\mathbb{T}} \eta_*(t, \theta) d\theta \right).$$

But, in exactly the same way, we almost surely have

$$\liminf \frac{1}{M/2} \sum_{i=1}^{M/2} \varphi_{N,K,i,a}(\eta_i(t)) \geq \tilde{\varphi} \left(\int_0^{1/2} \eta_*(t, \theta) d\theta \right)$$

and

$$\liminf \frac{1}{M/2} \sum_{i=M/2}^M \varphi_{N,K,i,a}(\eta_i(t)) \geq \tilde{\varphi} \left(\int_{1/2}^1 \eta_*(t, \theta) d\theta \right),$$

so that

$$\liminf \frac{1}{M} \sum_{i=1}^M \varphi_{N,K,i,a}(\eta_i(t)) \geq \frac{1}{2} \tilde{\varphi} \left(\int_0^{1/2} \eta_*(t, \theta) d\theta \right) + \frac{1}{2} \tilde{\varphi} \left(\int_{1/2}^1 \eta_*(t, \theta) d\theta \right).$$

Iterating this argument gives us, for any $L \in \mathbb{N}$,

$$\liminf \frac{1}{M} \sum_{i=1}^M \varphi_{N,K,i,a}(\eta_i(t)) \geq \frac{1}{2^L} \sum_{j=1}^{2^L} \tilde{\varphi} \left(\int_{(j-1)/2^L}^{j/2^L} \eta_*(\theta) d\theta \right) \quad (2.111)$$

almost surely. Letting L go to infinity, since η_* belongs to L^p , we get (2.109), and therefore (i).

Let us now prove (ii). We fix $\xi \in L^p$. We have

$$\begin{aligned} & \left| \int_0^T (\bar{H}_{N,K,a}(\eta_\ell(t) + \xi_\ell(t)) - \frac{1}{N} \log \bar{Z}_{N,K,a} \beta(t)) dt - \int_0^T \beta(t) \int_{\mathbb{T}} \tilde{\varphi}(\eta_*(t, \theta) + \xi(\theta)) d\theta dt \right| \\ &= \left| \int_0^T \left(\sum_{i=1}^{M_\ell} \psi_{N,K,i,a}(\eta_i^\ell + \xi_i^\ell) \right) \beta(t) dt - \int_0^T \beta(t) \int_{\mathbb{T}} \tilde{\varphi}(\eta_*(t, \theta) + \xi(\theta)) d\theta dt \right| \\ &\leq \int_0^T \beta(t) \left(\sum_{i=1}^{M_\ell} |\psi_{N,K,i,a}(\eta_i^\ell + \xi_i^\ell) - \varphi_{N,K,i,a}(\eta_i^\ell + \xi_i^\ell)| \right) dt \\ &\quad + \int_0^T \beta(t) \left| \left(\sum_{i=1}^{M_\ell} \varphi_{N,K,i,a}(\eta_i^\ell + \xi_i^\ell) - \tilde{\varphi}(\eta_i^\ell + \xi_i^\ell) \right) \right| dt \\ &\quad + \left| \int_0^T \int_{\mathbb{T}} \tilde{\varphi}(\bar{\xi}^\ell + \bar{\eta}^\ell) \beta(t) d\theta dt - \int_0^T \int_{\mathbb{T}} \tilde{\varphi}(\eta_* + \xi) \beta(t) d\theta dt \right|. \end{aligned} \quad (2.112)$$

The first term on the right-hand side is controlled by Theorem 2.34, since the sequence $\eta_i^\ell + \xi_i^\ell$ is bounded in L^p . The second term goes to 0 as a consequence of Lemma 2.27, which can be used since we assumed (2.98) holds.

Finally, since we have $\eta^\ell + \xi^\ell = \pi_\ell(\xi + \eta_*)$, $\bar{\xi}^\ell + \bar{\eta}^\ell$ converges to $\xi + \eta_*$ strongly in L^p . From the polynomial upper bound on $\tilde{\varphi}$, $\rho \rightarrow \int \tilde{\varphi}(\rho) d\theta$ is continuous with respect to strong L^p convergence, and therefore

$$\int_{\mathbb{T}} \tilde{\varphi}(\bar{\xi}^\ell + \bar{\eta}^\ell) d\theta \rightarrow \int_{\mathbb{T}} \tilde{\varphi}(\eta_* + \xi) d\theta. \quad (2.113)$$

By the Dominated Convergence Theorem, it follows that

$$\int_0^T \int_{\mathbb{T}} \tilde{\varphi}(\bar{\xi}^\ell + \bar{\eta}^\ell) \beta(t) d\theta dt \longrightarrow \int_0^T \int_{\mathbb{T}} \tilde{\varphi}(\eta_* + \xi) \beta(t) d\theta dt, \quad (2.114)$$

so the third term on the right-hand side of (2.112) goes to zero. □

2.6 Proof of the local Cramér theorem

In this section, we give the proof of the local Cramér theorem we used to obtain convexity of the coarse-grained Hamiltonian (cf. Proposition 2.3). It is a combination of the proofs of the local Cramér theorem of [GOVW09] and [MO13], modified to take into account the additional inhomogeneous linear term in the Hamiltonian.

Theorem 2.34. *Let*

$$\psi_K(m) := -\frac{1}{K} \log \int_{X_{K,m}} \exp\left(-\sum a_i x_i + \psi(x_i)\right) dx$$

and

$$\varphi_K(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \frac{1}{K} \sum_{i=1}^K \log \int_{\mathbb{R}} \exp((\sigma - a_i)x - \psi(x)) dx \right).$$

There exists $K_0 \in \mathbb{N}$, $C > 0$ (which only depends on ψ , K_0 and $\sup |a_i|$) such that for any m

$$|\psi_K(m) - \varphi_K(m)| \leq \frac{C}{K} |\varphi_K''(m)|^{\frac{1}{2}}.$$

In particular, for any $L > 0$ and any compact subset E of \mathbb{R} ,

$$\lim_{K \rightarrow \infty} \sup_{a_1 \dots a_K \in [-L, L]} \|\psi_K - \varphi_K\|_{\infty, E} = 0.$$

Moreover, there exists $\lambda > 0$ and $K_0 \in \mathbb{N}$ such that, for any $K \geq K_0$ and any collection of real numbers $a_1 \dots a_K \in [-L, L]$, we have

$$\psi_K'' \geq \lambda. \quad (2.115)$$

The starting point of the proof is the explicit representation

$$g_{K,m}(0) = \exp(K\varphi_K(m) - K\psi_K(m)), \quad (2.116)$$

where $g_{K,m}$ is the Lebesgue density of the random variable

$$\frac{1}{\sqrt{K}} \sum_{i=1}^K X_i - m_i,$$

and the X_i are independent random variables, distributed as

$$\mu_{\sigma,i}(dx) := \exp(-\varphi^*(\sigma - a_i) + (\sigma - a_i)x - \psi(x)) dx,$$

with m_i the mean of X_i .

Since φ_K is the Legendre transform of the strictly convex function

$$\varphi_K^*(\sigma) := \frac{1}{K} \sum_{i=1}^K \log \int_{\mathbb{R}} \exp((\sigma - a_i)x - \psi(x)) dx,$$

there exists for any $m \in \mathbb{R}$ a unique $\sigma = \sigma(m)$ such that

$$\varphi_K(m) = \sigma m - \varphi_K^*(\sigma).$$

It is a basic property of the Legendre transform that σ is determined by

$$\frac{d}{d\sigma} \varphi_K^*(\sigma) = \frac{1}{K} \sum_{i=1}^K \frac{\int x \mu_{\sigma,i}(dx)}{\int \mu_{\sigma,i}(dx)} = \frac{1}{K} \sum_{i=1}^K m_i = m. \quad (2.117)$$

Since the derivative of the Legendre transform of a convex function is the inverse of the derivative of the function, we also have

$$\sigma(m) = \varphi'_K(m). \quad (2.118)$$

Proposition 2.35. *Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded perturbation of a uniformly convex function. Let us define*

$$\mu_{\sigma}(dx) := \exp(-\varphi^*(\sigma) + \sigma x - \psi(x)) dx; \quad (2.119)$$

$$m(\sigma) := \int_{\mathbb{R}} x \mu_{\sigma}(dx); \quad s(\sigma)^2 := \int_{\mathbb{R}} (x - m)^2 \mu_{\sigma}(dx). \quad (2.120)$$

We assume that the following bounds hold uniformly in σ and j :

$$\int_{\mathbb{R}} |x - m(\sigma)|^k \mu_{\sigma}(dx) \leq C s^k \text{ for } k = 1..5; \quad (2.121)$$

$$\left| \int_{\mathbb{R}} \exp(i(x - m)\xi) \mu_{\sigma}(dx) \right| \leq \frac{C}{s|\xi|} \text{ for all } \xi \in \mathbb{R}. \quad (2.122)$$

Moreover, we assume that

$$\sup_{\sigma \in \mathbb{R}_a} \sup_{b \in [-L, L]} \frac{s(\sigma - a)}{s(\sigma - b)} \lesssim 1. \quad (2.123)$$

Then we have

$$\left| g_{K,m}(0) - \left(\frac{2\pi}{K} \sum s(\sigma - \lambda_i)^2 \right)^{-1/2} \right| \lesssim \frac{1}{\sqrt{K}} \left(\frac{1}{K} \sum s(\sigma - \lambda_i)^2 \right)^{-1/2}; \quad (2.124)$$

$$\left| \frac{d}{d\sigma} g_{K,m}(0) \right| \lesssim 1 \quad (2.125)$$

$$\left| \frac{d^2}{d\sigma^2} g_{K,m}(0) \right| \lesssim \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{1/2}. \quad (2.126)$$

Using this result, we can prove our local Cramer theorem.

Proof of Theorem 2.34. Let us first check that the Assumptions of Proposition 2.35 are verified. It has been shown in [MO, Lemma 3.2] that (2.121) and (2.122) hold when ψ is a bounded perturbation of a uniformly convex potential. We refer to [MO13] for a proof, which is based on simple general bounds on log-concave probability densities.

To show that (2.123) is satisfied, we will show that $\sigma \rightarrow \log s(\sigma)$ is Lipschitz continuous, by showing that its derivative is bounded. We have

$$\frac{d}{d\sigma} \log s(\sigma) = \frac{1}{s} \frac{d}{d\sigma} s = \frac{1}{2s^2} \frac{d}{d\sigma} s^2,$$

and we will show later (see (2.144)) that

$$\frac{d}{d\sigma} s(\sigma)^2 = \int (x - m)^3 \mu_\sigma(dx),$$

so that, by (2.121), we have

$$\left| \frac{d}{d\sigma} \log s(\sigma) \right| \leq C s(\sigma).$$

Since $\sigma \rightarrow s(\sigma)$ is bounded above (see Lemma 2.38), this quantity is bounded independently of σ . Therefore $\log s$ is Lipschitz-continuous, and (2.123) immediately follows.

We will now apply Proposition 2.35 to prove our theorem. First, we have

$$\varphi_K(m) - \psi_K(m) = \frac{1}{K} \log g_{K,m}(0).$$

The first part of the theorem then immediately follows from Proposition 2.35 and the fact that (cf. (2.89))

$$\varphi_K''(m) = \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1}.$$

Since the variance s is a continuous, positive function, the bound (2.124) implies that $\log g_{K,m}(0)$ is uniformly bounded when σ varies in a compact subset of \mathbb{R} and the real numbers are bounded by some constant L . As an immediate consequence, we get the second part of Theorem 2.34.

For the last part, we have

$$\begin{aligned} \varphi_K''(m) - \psi_K''(m) &= \frac{1}{K} \frac{d^2}{dm^2} \log g_{K,m}(0) \\ &= \frac{1}{K} \frac{1}{g_{K,m}(0)} \frac{d^2}{dm^2} g_{K,m}(0) - \frac{1}{K} \left(\frac{1}{g_{K,m}(0)} \frac{d}{dm} g_{K,m}(0) \right)^2 \\ &= \frac{1}{K} \frac{1}{g_{K,m}(0)} \frac{d^2}{dm^2} \sigma \frac{d}{d\sigma} g_{K,m}(0) + \frac{1}{K} \frac{1}{g_{K,m}(0)} \left(\frac{d}{dm} \sigma \right)^2 \frac{d^2}{d\sigma^2} g_{K,m}(0) \\ &\quad - \frac{1}{K} \left(\frac{1}{g_{K,m}(0)} \frac{d\sigma}{dm} \frac{d}{d\sigma} g_{K,m}(0) \right)^2 \\ &\leq \frac{1}{K} \frac{1}{g_{K,m}(0)} \frac{d^2}{dm^2} \sigma \frac{d}{d\sigma} g_{K,m}(0) + \frac{1}{K} \frac{1}{g_{K,m}(0)} \left(\frac{d}{dm} \sigma \right)^2 \frac{d^2}{d\sigma^2} g_{K,m}(0) \end{aligned} \tag{2.127}$$

We can explicitly compute the derivatives of σ with respect to m . Direct calculation (see 2.143 below) yields

$$\frac{d}{d\sigma} m = \frac{1}{K} \sum_{i=1}^K \frac{d}{d\sigma} m_i = \frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2,$$

so that

$$\frac{d^2}{dm^2}\varphi_K = \frac{d}{dm}\sigma = \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1}. \quad (2.128)$$

Another direct calculation yields

$$\begin{aligned} \frac{d^2}{dm^2}\sigma &= \frac{d}{dm} \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1} \\ &= - \left(\frac{1}{K} \sum_{i=1}^K \frac{d}{dm} s(\sigma - a_i)^2 \right) \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-2} \\ &= - \left(\frac{1}{K} \sum_{i=1}^K \int (x - m_i)^3 \mu_{\sigma,i}(dx) \right) \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-3}. \end{aligned}$$

This equality, (2.121) and (2.123) immediately imply

$$\begin{aligned} \left| \frac{d^2}{dm^2}\sigma \right| &\lesssim \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-3/2} \frac{\sup_i s(\sigma - a_i)^3}{\inf_i s(\sigma - a_i)^3} \\ &\lesssim \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-3/2}. \end{aligned} \quad (2.129)$$

Plugging the bounds (2.124), (2.126), (2.125), (2.128) and (2.129) into (2.127) yields

$$\varphi_K''(m) - \psi_K''(m) \leq \frac{C}{K} \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1}.$$

But since we have (2.128), which states that

$$\varphi_K''(m) = \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1},$$

we get

$$\begin{aligned} \psi_K''(m) &\geq \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1} - \frac{C}{K} \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1} \\ &\geq \frac{1}{2} \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{-1} \end{aligned}$$

for K large enough. Since $\sigma \rightarrow s(\sigma)$ is bounded above (see Lemma 2.38), we immediately obtain (2.115) □

Now all that is left is to prove Proposition 2.35.

Notation 2.36. We denote by $h(\sigma, \xi)$ the Fourier transform of the law of $X - m$, with X distributed according to μ_σ , and m its mean, that is

$$h(\sigma, \xi) := \int_{\mathbb{R}} \exp(i(x - m)\xi) \mu_\sigma(dx).$$

By using the inverse Fourier transform, we have the explicit formula for $g_{K,m}(0)$:

$$g_{K,m}(0) = \frac{1}{2\pi} \int_{\mathbb{R}} \prod_{i=1}^K h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) d\xi. \quad (2.130)$$

To control this quantity, we will split this integral in two : one integral over large values of ξ , and one over small values of ξ .

Lemma 2.37. *Let $A(C_0)$ be the set of probability measures ν on \mathbb{R} such that*

$$\int_{\mathbb{R}} |x| \nu(dx) \leq C_0$$

and

$$\left| \int_{\mathbb{R}} \exp(ix\xi) \nu(dx) \right| \leq \frac{C_0}{|\xi|} \text{ for all } \xi \in \mathbb{R}$$

for some $C_0 < \infty$. Then, for any $\delta > 0$, there exists $\lambda < 1$ such that

$$\left| \int_{\mathbb{R}} \exp(ix\xi) \nu(dx) \right| \leq \lambda \text{ for all } \xi \geq \delta$$

for all ν in $A(C_0)$.

Proof of Lemma 2.37. We can assume without loss of generality that $\delta < 1$. Because of assumption (2.122), we have for any $|\xi| \geq 2C_0$ that $|\int_{\mathbb{R}} \exp(ix\xi) \nu(dx)| \leq \frac{1}{2}$, so it is enough to show that

$$\left| \int_{\mathbb{R}} \exp(ix\xi) \nu(dx) \right| \leq \lambda \text{ for all } \xi \in [\delta, 2C_0].$$

Lets assume that this result doesn't hold. Then there exists a sequence (ν_n) of probability measures in $A(C_0)$ and a sequence $(\xi_n) \in [\delta, 2C_0]^{\mathbb{N}}$ such that

$$\liminf_{n \uparrow \infty} \left| \int_{\mathbb{R}} \exp(ix\xi_n) \nu_n(dx) \right| \geq 1. \quad (2.131)$$

Since for all n we have $\int_{\mathbb{R}} |x| \nu_n(dx) \leq C_0$, we have compactness of the sequence (ν_n) for weak convergence, and we can extract a subsequence such that

$$\int f(x) \nu_n(dx) \longrightarrow \int f(x) \nu_{\infty}(dx)$$

for all bounded continuous function f , and

$$\xi_n \longrightarrow \xi_{\infty}.$$

Therefore, since $|\exp(ix\xi_n) - \exp(ix\xi_{\infty})| \leq |x| |\xi_n - \xi_{\infty}|$, we also have

$$\int_{\mathbb{R}} \exp(ix\xi_n) \nu_n(dx) \longrightarrow \int_{\mathbb{R}} \exp(ix\xi_{\infty}) \nu_{\infty}(dx),$$

so that (2.131) saturates to

$$\int_{\mathbb{R}} \exp(ix\xi_{\infty}) \nu_{\infty}(dx) = 1.$$

The classical result on the equality cases in the triangle inequality then tells us that $x \rightarrow \exp(ix\xi_{\infty})$ is ν_{∞} -almost surely a constant $\zeta \in S^1$, and therefore

$$\int_{\mathbb{R}} \exp(ikx\xi_{\infty}) \nu_{\infty}(dx) = \zeta^k \text{ for all } k \in \mathbb{N}.$$

But since $\xi_{\infty} \neq 0$, $|k\xi_{\infty}|$ goes to infinity when k goes to infinity, and since $\nu_{\infty} \in A(C_0)$, $\int_{\mathbb{R}} \exp(ikx\xi_{\infty}) \nu_{\infty}(dx)$ should go to zero. We therefore have a contradiction, and (2.131) cannot hold.

□

Lemma 2.38. *There exists a constant $C > 0$ such that for all $\sigma \in \mathbb{R}$ we have*

$$s(\sigma) = \int_{\mathbb{R}} (x - m(\sigma))^2 \mu_{\sigma}(dx) \leq C.$$

Since, by this lemma, the variances $s(\sigma)$ are bounded above independently of σ , we can apply Lemma 2.37, so for any $\delta > 0$, there exists $\lambda < 1$ such that if $|\xi| > \delta$, then for any σ we have $|h(\sigma, \xi)| \leq \lambda$.

Proof of Lemma 2.38. It is easy to show (with a combination of the Bakry-Emery and Holley-Stroock criterions) that the measures μ_{σ} satisfy a spectral gap inequality with a uniform constant, yielding the desired uniform upper bound on the variances. \square

Lemma 2.39. *Under the assumptions of Proposition 2.35, there exists a family of complex-valued functions v_{σ} such that*

$$\int_{\mathbb{R}} \exp(i(x - m(\sigma))\xi) \mu_{\sigma}(dx) = \exp(-v_{\sigma}(\xi)) \quad (2.132)$$

and

$$|v_{\sigma}(\xi) - \frac{s(\sigma)^2}{2} \xi^2| \leq Cs(\sigma)^3 |\xi|^3 \quad (2.133)$$

for all ξ small enough, independently of σ .

Démonstration. We can use (2.132) as a definition of the functions v_{σ} , so we just have to prove (2.133). We have

$$\frac{d^k}{d\xi^k} \int_{\mathbb{R}} \exp(i(x - m(\sigma))\xi) \mu_{\sigma}(dx) = i^k \int_{\mathbb{R}} (x - m)^k \exp(i(x - m)\xi) \mu_{\sigma}(dx),$$

and, since we assumed 2.121, we get

$$\left| \frac{d^k}{d\xi^k} \int_{\mathbb{R}} \exp(i(x - m(\sigma))\xi) \mu_{\sigma}(dx) \right| \leq Cs^k,$$

and a Taylor expansion of the function v_{σ} then gives us the desired result. \square

Let us now fix $\delta > 0$ small enough, so that the bound of Lemma 2.39 holds for all $|\xi| \leq \delta$. We split (2.130) in two integrals, depending on the value of $|\xi|$:

$$\begin{aligned} g_{K,m}(0) &= \frac{1}{2\pi} \int_{\mathbb{R}} \prod_{i=1}^K h(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}) d\xi \\ &= \frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| \leq \delta} \prod_{i=1}^K h(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}) d\xi \\ &\quad + \frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| > \delta} \prod_{i=1}^K h(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}) d\xi \end{aligned} \quad (2.134)$$

We will show that the first term is of leading order, and converges to 1. Let us first take care of the second term. Since we have a uniform bound on the variances of the probability measures μ_{σ} , we can apply Lemma 2.38, and obtain the bound

$$h(\sigma, \xi) \leq \lambda$$

for some $\lambda < 1$, for all $|\xi| \leq \delta$, independently of σ . We use this estimate on $K - 2$ of the K elements of the product, and estimate (2.122) on the last two, to obtain

$$\begin{aligned} \left| \prod_{i=1}^K h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) \right| &\leq C\lambda^{K-2} \left(\frac{1}{1 + |\xi|/\sqrt{K}} \right)^2 \\ &\leq C\lambda^{K-2} \frac{K}{K + \xi^2} \\ &\leq C\lambda^{K-2} \frac{K}{1 + \xi^2} \end{aligned} \tag{2.135}$$

We therefore have

$$\begin{aligned} \left| \frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| > \delta} \prod_{i=1}^K h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) d\xi \right| &\leq CK\lambda^{K-2} \int_{\mathbb{R}} \frac{1}{1 + \xi^2} d\xi \\ &\leq \frac{C}{K} \end{aligned} \tag{2.136}$$

where the last estimate is because we have $\lambda < 1$, so this term is exponentially small in K , and therefore negligible when compared to $1/K$. All that is now left is to take care of the integral over small values of ξ . Recalling definition (2.132), we have

$$\frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| \leq \delta} \prod_{i=1}^K h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) d\xi = \frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| \leq \delta} \exp\left(-\sum_{i=1}^K v_{\sigma - \lambda_i}(\xi/\sqrt{K})\right) d\xi$$

Since we assumed δ small enough, according to Lemma 2.39, we have, independently of σ and for all ξ such that $|\xi|/\sqrt{K} \leq \delta$

$$|v_{\sigma}(\xi) - \frac{s(\sigma)^2}{2}\xi^2| \leq Cs(\sigma)^3|\xi|^3.$$

This bound implies that

$$\begin{aligned} \operatorname{Re}(v_{\sigma}(\xi)) &\geq \frac{s(\sigma)^2}{2}\xi^2 - Cs(\sigma)^3|\xi|^3 \\ &\geq \frac{s(\sigma)^2}{4}\xi^2 \end{aligned}$$

independently of σ , as long as we have chosen δ small enough (since $s(\sigma)$ is bounded above). But since $y \rightarrow \exp(y)$ is Lipschitz continuous on $\{\operatorname{Re}(y) \leq -c\xi^2\}$, with constant $\exp(-c\xi^2)$, we have

$$\begin{aligned} &\left| \exp\left(-\sum_{i=1}^K v_{\sigma - \lambda_i}(\xi/\sqrt{K})\right) - \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{2K}\xi^2\right) \right| \\ &\leq C \left(\sum s(\sigma - \lambda_i)^3 \right) \frac{|\xi|^3}{K^{3/2}} \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{4K}\xi^2\right) \\ &\leq C \left(\frac{1}{K} \sum s(\sigma - \lambda_i)^3 \right) \frac{|\xi|^3}{\sqrt{K}} \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{4K}\xi^2\right) \end{aligned} \tag{2.137}$$

Consequently, we have

$$\begin{aligned}
& \left| \frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| \leq \delta} \prod_{i=1}^K h(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}) d\xi - \frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| \leq \delta} \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{2K} \xi^2\right) d\xi \right| \\
& \leq C \left(\frac{1}{K} \sum s(\sigma - \lambda_i)^3 \right) \frac{1}{\sqrt{K}} \int_{\frac{1}{\sqrt{K}}|\xi| \leq \delta} |\xi|^3 \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{4K} \xi^2\right) d\xi \\
& \leq C \left(\frac{1}{K} \sum s(\sigma - \lambda_i)^3 \right) \frac{1}{\sqrt{K}} \int_{\mathbb{R}} |\xi|^3 \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{4K} \xi^2\right) d\xi \\
& \leq \frac{C}{\sqrt{K}} \frac{\left(\frac{1}{K} \sum s(\sigma - \lambda_i)^3\right)}{\left(\frac{1}{K} \sum s(\sigma - \lambda_i)^2\right)^2} \\
& \leq \frac{C}{\sqrt{K}} \left(\frac{1}{K} \sum s(\sigma - a_i)^2 \right)^{-1/2} \frac{\sup s(\sigma - a_i)^3}{\inf s(\sigma - a_i)^3} \\
& \leq \frac{C}{\sqrt{K}} \left(\frac{1}{K} \sum s(\sigma - a_i)^2 \right)^{-1/2}. \tag{2.138}
\end{aligned}$$

Since $\int_{\frac{1}{\sqrt{K}}|\xi| > \delta} \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{2K} \xi^2\right) d\xi$ is exponentially small in K , we finally obtain

$$\begin{aligned}
& \left| \frac{1}{2\pi} \int_{\frac{1}{\sqrt{K}}|\xi| \leq \delta} \prod_{i=1}^K h(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}) d\xi - \frac{1}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{\sum s(\sigma - \lambda_i)^2}{2K} \xi^2\right) d\xi \right| \\
& \leq \frac{C}{\sqrt{K}} \left(\frac{1}{K} \sum s(\sigma - \lambda_i)^2 \right)^{-1/2}.
\end{aligned}$$

In the end, what we obtain is

$$\left| g_{K,m}(0) - \left(\frac{2\pi}{K} \sum s(\sigma - \lambda_i)^2 \right)^{-1/2} \right| \leq \frac{C}{\sqrt{K}} \left(\frac{1}{K} \sum s(\sigma - \lambda_i)^2 \right)^{-1/2}. \tag{2.139}$$

To bound the derivatives of $g_{K,m}(0)$ with respect to σ , we will need the following estimates :

Lemma 2.40. *Under the assumptions of Proposition 2.35 we have, uniformly in σ*

$$\left| \frac{d}{d\sigma} h(\sigma, \xi) \right| \leq C s^3 |\xi|^2; \tag{2.140}$$

$$\left| \frac{d^2}{d\sigma^2} h(\sigma, \xi) \right| \leq C s^4 (1 + s^2 |\xi|^2) |\xi|^2. \tag{2.141}$$

Démonstration. To prove this lemma, we will rely on the following identity :

$$\frac{d}{d\sigma} \int f(x) \mu_\sigma(dx) = \int (x - m) f(x) \mu_\sigma(dx). \tag{2.142}$$

which can be directly computed easily. Using this identity, we can show that

$$\frac{d}{d\sigma} m = \frac{d}{d\sigma} \int x \mu_\sigma(dx) = \int x(x - m) \mu_\sigma(dx) = s^2, \tag{2.143}$$

$$\frac{d}{d\sigma} s^2 = \frac{d}{d\sigma} \int (x - m)^2 \mu_\sigma(dx) = \int (x - m)^3 \mu_\sigma(dx). \tag{2.144}$$

To bound the derivatives in σ of the Fourier transform, we will use a Taylor expansion in 0. We have

$$\frac{d^k}{d\xi^k} \frac{d}{d\sigma} h(\sigma, \xi) = \frac{d}{d\sigma} \frac{d^k}{d\xi^k} h(\sigma, \xi) = i^k \frac{d}{d\sigma} \int (x - m)^k \exp(i(x - m)\xi) \mu_\sigma(dx). \quad (2.145)$$

In particular, $\frac{d}{d\sigma} h(\sigma, \xi)$ vanishes to the first order in 0 in ξ , and, by assumption (2.121),

$$\left| \frac{d^2}{d\xi^2} \frac{d}{d\sigma} h(\sigma, \xi) \right| = \left| \frac{d}{d\sigma} s^2 \right| \leq C s^3,$$

so that

$$\left| \frac{d}{d\sigma} h(\sigma, \xi) \right| \leq C s^3 |\xi|^2. \quad (2.146)$$

In the same way,

$$\frac{d^k}{d\xi^k} \frac{d^2}{d\sigma^2} h(\sigma, \xi) = \frac{d^2}{d\sigma^2} \frac{d^k}{d\xi^k} h(\sigma, \xi) = i^k \frac{d^2}{d\sigma^2} \int (x - m)^k \exp(i(x - m)\xi) \mu_\sigma(dx), \quad (2.147)$$

so that $\frac{d^2}{d\sigma^2} h(\sigma, \xi)$ also vanishes to the first order in 0 in ξ , and

$$\begin{aligned} \frac{d^2}{d\xi^2} \frac{d^2}{d\sigma^2} h(\sigma, \xi) &= -\frac{d^2}{d\sigma^2} \int (x - m)^2 \exp(i(x - m)\xi) \mu_\sigma(dx) \\ &= \frac{d}{d\sigma} \left[-\int (x - m)^3 \exp(i(x - m)\xi) \mu_\sigma(dx) \right. \end{aligned} \quad (2.148)$$

$$\begin{aligned} &\quad \left. + s^2 \int (2(x - m) + i\xi(x - m)^2) \exp(i(x - m)\xi) \mu_\sigma(dx) \right] \\ &= -\int (x - m)^4 \exp(i(x - m)\xi) \mu_\sigma(dx) \end{aligned} \quad (2.149)$$

$$\begin{aligned} &+ s^2 \int (3(x - m)^2 + i\xi(x - m)^3) \exp(i(x - m)\xi) \mu_\sigma(dx) \\ &+ \left(\int (x - m)^3 \mu_\sigma(dx) \right) \left(\int (2(x - m) + i\xi(x - m)^2) \exp(i(x - m)\xi) \mu_\sigma(dx) \right) \\ &- 2s^4 h(\sigma, \xi) - 4i\xi s^4 \int (x - m) \exp(i(x - m)\xi) \mu_\sigma(dx) \\ &+ s^4 \xi^2 \int (x - m)^2 \exp(i(x - m)\xi) \mu_\sigma(dx), \end{aligned}$$

so that, in the end, using assumption 2.121, we get

$$\left| \frac{d^2}{d\sigma^2} h(\sigma, \xi) \right| \leq C s^4 (1 + s^2 |\xi|^2) |\xi|^2. \quad (2.150)$$

□

Let us now compute the derivatives with respect to σ of $g_{K,\sigma}$

$$\begin{aligned} \frac{d}{d\sigma} g_{K,\sigma}(0) &= \frac{d}{d\sigma} \frac{1}{2\pi} \int_{\mathbb{R}} \prod_{i=1}^K h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) d\xi \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \sum_{i=1}^K \frac{d}{d\sigma} h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) \prod_{j \neq i} h\left(\sigma - \lambda_j, \frac{\xi}{\sqrt{K}}\right) d\xi; \end{aligned}$$

$$\begin{aligned}
\frac{d^2}{d\sigma^2}g_{K,\sigma}(0) &= \frac{1}{2\pi} \int_{\mathbb{R}} \sum_{i=1}^K \frac{d^2}{d\sigma^2}h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) \prod_{j \neq i} h\left(\sigma - \lambda_j, \frac{\xi}{\sqrt{K}}\right) d\xi \\
&+ \frac{1}{2\pi} \int_{\mathbb{R}} \sum_{i,j=1, i \neq j}^K \frac{d}{d\sigma}h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) \frac{d}{d\sigma}h\left(\sigma - \lambda_j, \frac{\xi}{\sqrt{K}}\right) \prod_{k \neq i,j} h\left(\sigma - \lambda_k, \frac{\xi}{\sqrt{K}}\right) d\xi. \quad (2.151)
\end{aligned}$$

We are seeking to bound $\frac{d^2}{d\sigma^2}g_{K,\sigma}(0)$. Once again, we separate the integrals into inner and outer integrals. Once more, the leading order term will be the inner integrals. For the outer integrals, we have, by using Lemma 2.37 on $K - 6$ of the $K - 1$ factors, estimate (2.122) on the remaining 6 and estimate (2.141) on the second derivative of h to obtain (still remembering that $s(\sigma)$ is bounded above independently of σ)

$$\begin{aligned}
&\frac{1}{2\pi} \left| \int_{|\xi|/\sqrt{K} > \delta} \sum_{i=1}^K \frac{d^2}{d\sigma^2}h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) \prod_{j \neq i} h\left(\sigma - \lambda_j, \frac{\xi}{\sqrt{K}}\right) d\xi \right| \\
&\leq C \sum_{i=1}^K \left| \int_{|\xi|/\sqrt{K} > \delta} s(\sigma - \lambda_i)^4 (1 + s(\sigma - \lambda_i)^2 \frac{|\xi|^2}{K}) \frac{|\xi|^2}{K} \left(\frac{K}{K + |\xi|^2}\right)^6 \lambda^{K-6} d\xi \right| \\
&\leq \lambda^{K-6} \frac{C}{K} \sum_{i=1}^K s(\sigma - \lambda_i)^4 \int_{\mathbb{R}} \frac{1}{1 + \xi^2} d\xi \\
&\leq \frac{C}{K}. \quad (2.152)
\end{aligned}$$

The same strategy (using estimate (2.140)) allows us to show that

$$\begin{aligned}
&\frac{1}{2\pi} \left| \int_{|\xi|/\sqrt{K} > \delta} \sum_{i,j=1, i \neq j}^K \frac{d}{d\sigma}h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) \frac{d}{d\sigma}h\left(\sigma - \lambda_j, \frac{\xi}{\sqrt{K}}\right) \prod_{k \neq i,j} h\left(\sigma - \lambda_k, \frac{\xi}{\sqrt{K}}\right) d\xi \right| \\
&\leq \frac{C}{K}.
\end{aligned}$$

We will now show that the inner integrals are bounded.

$$\begin{aligned} & \frac{1}{2\pi} \left| \int_{|\xi|/\sqrt{K} \leq \delta} \sum_{i=1}^K \frac{d^2}{d\sigma^2} h(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}) \prod_{j \neq i} h(\sigma - \lambda_j, \frac{\xi}{\sqrt{K}}) d\xi \right| \\ & \leq C \sum_{i=1}^K \left| \int_{|\xi|/\sqrt{K} \leq \delta} s(\sigma - \lambda_i)^4 \left(1 + s(\sigma - a_i)^2 \frac{|\xi|^2}{K} \right) \frac{|\xi|^2}{K} \right. \\ & \quad \left. \times \exp \left(- \sum_{j \neq i} v_{\sigma - a_j}(\xi/\sqrt{K}) \right) d\xi \right| \end{aligned} \quad (2.153)$$

$$\begin{aligned} & \leq C \sum_{i=1}^K \int_{|\xi|/\sqrt{K} \leq \delta} s(\sigma - a_i)^4 (1 + s(\sigma - a_i)^2 \delta^2) \frac{|\xi|^2}{K} \\ & \quad \times \exp \left(- \sum_{j \neq i} \frac{s(\sigma - a_j)^2}{2K} |\xi|^2 \right) d\xi \end{aligned} \quad (2.154)$$

$$\begin{aligned} & + C \sum_{i=1}^K \int_{|\xi|/\sqrt{K} \leq \delta} s(\sigma - a_i)^4 (1 + s(\sigma - a_i)^2 \delta^2) \frac{|\xi|^5}{K^{5/2}} \left(\sum_{j \neq i} s(\sigma - a_j)^3 \right) \\ & \quad \times \exp \left(- \sum_{j \neq i} \frac{s(\sigma - a_j)^2}{4K} |\xi|^2 \right) d\xi \end{aligned} \quad (2.155)$$

$$\begin{aligned} & \leq \frac{C}{K} \sum_{i=1}^K s(\sigma - a_i)^4 \int_{\mathbb{R}} |\xi|^2 \exp \left(- \sum_{j \neq i} \frac{s(\sigma - \lambda_j)^2}{2K} |\xi|^2 \right) d\xi \\ & + \frac{C}{K^{3/2}} \sum_{i=1}^K s(\sigma - a_i)^4 \left(\frac{1}{K} \sum_{j \neq i} s(\sigma - a_j)^3 \right) \int_{\mathbb{R}} |\xi|^5 \exp \left(- \sum_{j \neq i} \frac{s(\sigma - a_j)^2}{4K} |\xi|^2 \right) d\xi \\ & \leq \frac{C}{K} \sum_{i=1}^K \frac{s(\sigma - a_i)^4}{\left(\sum_{j \neq i} \frac{s(\sigma - a_j)^2}{2K} \right)^{3/2}} \int_{\mathbb{R}} u^2 \exp(-u^2) du \\ & + \frac{C}{K^{3/2}} \sum_{i=1}^K s(\sigma - a_i)^4 \left(\frac{1}{K} \sum_{j \neq i} s(\sigma - a_j)^3 \right) \left(\sum_{j \neq i} \frac{s(\sigma - a_j)^2}{4K} \right)^{-3} \\ & \quad \times \int_{\mathbb{R}} |u|^5 \exp(-u^2) du \end{aligned} \quad (2.156)$$

$$\begin{aligned} & \leq \frac{C}{K} \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{1/2} \sum_{i=1}^K \frac{\left(\sum_{j \neq i} s(\sigma - a_j)^2 \right)^{1/2} \sup_k s(\sigma - a_k)^4}{\left(\sum_{k=1}^K s(\sigma - a_k)^2 \right)^{1/2} \inf_k s(\sigma - a_k)^4} \\ & + \frac{C}{K^{3/2}} \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{1/2} \sum_{i=1}^K \frac{\left(\sum_{j \neq i} s(\sigma - a_j)^2 \right)^{1/2} \sup_k s(\sigma - a_k)^7}{\left(\sum_{k=1}^K s(\sigma - a_k)^2 \right)^{1/2} \inf_k s(\sigma - a_k)^7} \\ & \leq C \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{1/2}. \end{aligned} \quad (2.157)$$

We can show in the same way that

$$\frac{1}{2\pi} \left| \int_{|\xi|/\sqrt{K} \leq \delta} \sum_{i,j=1, i \neq j}^K \frac{d}{d\sigma} h\left(\sigma - \lambda_i, \frac{\xi}{\sqrt{K}}\right) \frac{d}{d\sigma} h\left(\sigma - \lambda_j, \frac{\xi}{\sqrt{K}}\right) \right. \quad (2.158)$$

$$\begin{aligned} & \times \left. \prod_{k \neq i,j} h\left(\sigma - \lambda_k, \frac{\xi}{\sqrt{K}}\right) d\xi \right| \\ & \leq C \sum_{i,j=1, i \neq j}^K \int_{|\xi|/\sqrt{K} \leq \delta} s(\sigma - \lambda_i)^3 s(\sigma - \lambda_j)^3 \frac{|\xi|^4}{K^2} \exp\left(-\sum_{k \neq i,j} v_{\sigma - \lambda_k}(\xi/\sqrt{K})\right) d\xi \\ & \leq C \left(\frac{1}{K} \sum_{i=1}^K s(\sigma - a_i)^2 \right)^{1/2}. \quad (2.159) \end{aligned}$$

The same technique applied to (2.151) also yields (2.125).

Combining the bounds on the inner and outer integrals yields (2.126).

2.7 Some classical criteria for the LSI

Here we recall the three main criteria that are commonly used to obtain logarithmic Sobolev inequalities.

Criterion I (Tensorization principle)

If $\mu_1 \in P(X_1)$ and $\mu_2 \in P(X_2)$ satisfy $\text{LSI}(\rho_1)$ and $\text{LSI}(\rho_2)$ respectively, then $\mu_1 \otimes \mu_2$ satisfies $\text{LSI}(\min\{\rho_1, \rho_2\})$.

Criterion II (Holley-Stroock perturbation lemma)

Let $\mu \in P(X)$ satisfy $\text{LSI}(\rho)$ and let $\delta\psi : X \rightarrow \mathbb{R}$ be a bounded function. Let $\tilde{\mu} \in P(X)$ be defined through

$$\frac{d\tilde{\mu}}{d\mu}(x) = \frac{1}{Z} \exp(-\delta\psi(x)).$$

Then $\tilde{\mu}$ satisfies $\text{LSI}(\tilde{\rho})$, where

$$\tilde{\rho} = \rho \exp(-\text{osc}(\delta\psi)).$$

Criterion III (Bakry-Emery theorem)

Let X be a K -dimensional Riemannian manifold, let $H \in C^2(X)$ and let $\mu \in P(X)$ be defined by

$$\frac{d\mu}{dH^K}(x) = \frac{1}{Z} \exp(-H(x)).$$

If there is $\rho > 0$ such that $\text{Hess } H(x) \geq \rho$ for all $x \in X$, then μ satisfies $\text{LSI}(\rho)$.

Acknowledgement The first author would like to thank S.R.S. Varadhan, who suggested the problem to him during a stay at the Courant Institute in 2010-2011, and Cédric Villani, for discussions on the two-scale approach. The second author wants to thank Maria Westdickenberg (née Reznikoff) and Felix Otto for the inspiring discussions on this topic. He also wants to thank *Max-Planck Institute for Mathematics in the Sciences* in Leipzig for funding during the years 2010 to 2012 and Stanford University for funding during the years 2013 to 2014, where most of the content of this article originated.

Chapitre 3

Local Gibbs behavior in the hydrodynamic limit

Les résultats de ce chapitre ont été publiés dans [Fat13b]

Abstract : This work is a follow-up on [GOVW09]. In that previous work a two-scale approach was used to prove the logarithmic Sobolev inequality for a system of spins with fixed mean whose potential is a bounded perturbation of a Gaussian, and to derive an abstract theorem for the convergence to the hydrodynamic limit. This strategy was then successfully applied to Kawasaki dynamics. Here we shall use again this two-scale approach to show that the microscopic variable in such a model behaves according to a local Gibbs state. As a consequence, we shall prove the convergence of the microscopic entropy to the hydrodynamic entropy.

3.1 Introduction

A local Gibbs measure is a vague term used to designate a measure whose density (with respect to the plain Gibbs measure) takes the form $G(x) = \exp(\sum \lambda_i x_i)$, where the coefficients λ_i vary "at macroscopic scale". They have been used in [GPV88] for the Ginzburg-Landau model, and also play a crucial role in the relative entropy method devised by [Yau91]. They represent in some sense a "typical" microscopic distribution having the correct hydrodynamic profile. The main result in [Yau91] can be informally summarized as follows : if the initial datum is in local Gibbs state, then at later times the microscopic variable is very close (in the sense of Kullback information) to be in local Gibbs state too. The local Gibbs state Yau used is defined in terms of the hydrodynamic equation, and chosen so that it is close to being a solution of the microscopic equation.

In a more recent contribution, [Kos01a] proved that the solution of the Ginzburg-Landau model behaves like a local Gibbs state for all positive times even if it does not at initial time. That is, there is a time-dependent family of vectors $\lambda(t)$ such that, if $f(t, \cdot)$ is the density at time t with respect to the equilibrium measure μ of a system of N continuous spins x_i interacting according to Kawasaki dynamics, then the relative entropy of $f\mu$ with respect to the measure $\nu(dx) = \frac{1}{Z} \exp(\lambda \cdot x)\mu(dx)$, given by

$$Ent_\nu(f\mu) = \int \rho \log \rho ,$$

with ρ being the density of $f\mu$ with respect to ν , goes to 0 when N goes to infinity for any time $t > 0$. The equilibrium measure $\mu(dx) := \exp(-H(x))dx$ is assumed to have a

Ginzburg–Landau type potential, that is

$$H(x) = \sum \psi(x_i),$$

where ψ is the single-site potential. Kosygina’s proof relied on the logarithmic Sobolev inequality, and she used an assumption of uniform convexity of the Ginzburg–Landau potential ψ . In the present work we shall generalize these results to cover a certain class of nonconvex potentials. At the same time we shall point out the role of another information-theoretical inequality, the so-called HWI inequality introduced by [OV00]. This inequality will allow us to pass from a convergence in a Wasserstein distance sense to a convergence in relative entropy. To be used efficiently in this setting, the HWI inequality needs a log-concave reference measure, which is not the case for the microscopic equilibrium measure when ψ is not convex. This is why we shall use, like in [GOVW09], the convexification induced by the macroscopic block decomposition. Since we will then need our local Gibbs state to be compatible with the passage to macroscopic scale, we will use a local Gibbs state slightly different from the one used in [Yau91] (the value of $\lambda \cdot x$ must only depend on the macroscopic profile associated with x), but such that when N goes to infinity, the relative entropy with respect to either measure behaves in the same way.

As in [Kos01a], we shall also prove the (physically relevant) convergence of the microscopic entropy to the macroscopic (hydrodynamic) entropy, that is

$$\frac{1}{N} \int f(t, x) \log f(t, x) \mu(dx) \rightarrow \int_{\mathbb{T}} \varphi(\zeta(t, \theta)) d\theta - \varphi \left(\int_{\mathbb{T}} \zeta(t, \theta) d\theta \right),$$

where ζ is the hydrodynamic limit, and φ is the Cramér transform of the potential. However, we shall deduce it from the local Gibbs behavior, while Kosygina does it the other way round.

Our two-scale approach will only yield convergences in $L^1(dt)$. However, by using a method of [Kos01a] in conjunction with these results, we will be able to immediately prove that the convergence is uniform in time, as long as we stay away from time $t = 0$.

It should be noted that, while Kosygina used an assumption of convexity, it was mainly required to ensure that the equilibrium measures satisfies a logarithmic Sobolev inequality. In light of the recent work [MO13], it seems that her method can be adapted to the class of nonconvex potentials covered here. On the other hand, our method does not cover potentials with superquadratic growth. This restriction is inherited from the results in [GOVW09]. If the two scale approach could be extended to cover superquadratic potentials, our method could also be extended, with minor technical modifications. However, even with the logarithmic Sobolev inequality obtained in [MO13], extending the section of [GOVW09] that concerns the hydrodynamic limit to superquadratic potentials is nontrivial, and remains to be done.

The plan of this paper is as follows : in Section 3, we will recall the framework and results of [GOVW09] which will be used in this article and present our main results, in both the abstract framework and their application for Kawasaki dynamics. Sections 4 and 5 will then give the details of the proofs.

3.2 Notation

- ∇ stands for the gradient, Hess for Hessian, $|\cdot|$ for norm and $\langle \cdot, \cdot \rangle$ for inner product. Whenever necessary, the space to which these are associated will be indicated with a subscript.

- A^t is the transpose of the operator A .
- $\text{Ran}(A)$ is the range of the operator A .
- $\Phi(x) = x \log x$.
- $\text{Ent}_\mu(f) := \int \Phi(f) d\mu - \Phi(\int f d\mu)$ is the (negative of the) entropy of the positive function f with respect to the probability measure μ .
- C is a positive constant, which may change from line to line, or even within a line.
- Z is a positive constant enforcing unit mass of a given probability measure.
- id_E is the identity map $E \rightarrow E$
- LSI is an abbreviation for Logarithmic Sobolev Inequality.
- $\Gamma(Y, |\cdot|_Y) := \int \exp(-|y|_Y^2/2) dy$ is the Gaussian integral on the space Y with respect to the norm $|\cdot|_Y$.
- $C^{1,2}(A \times B)$ is the space of real-valued functions on $A \times B$ which are C^1 with respect to the first variable and C^2 with respect to the second variable.
- $W_2(\mu, \nu)$ is the Wasserstein distance between two probability measures μ and ν with finite second moment. It is defined as $W_2(\mu, \nu)^2 := \inf_{\pi \in \Pi} \int |x - y|^2 \pi(dx, dy)$, where Π is the set of all coupling of μ and ν ;
- $I_\mu(\nu)$ is the Fisher information of the probability measure ν with respect to μ , given by $\int \frac{|\nabla f|^2}{f} d\mu$ if $\nu = f\mu$ for some density f , and $+\infty$ if not.

3.3 Background and Main Results

The aim of this section is to recall the setting and the main results of [GOVW09], as well as to present the new results brought by the present paper.

3.3.1 Logarithmic Sobolev inequalities

Throughout this work, X and Y are two Euclidean spaces. It is convenient to think of X as the space of microscopic variables, and Y as the space of macroscopic variables. We consider a linear operator $P : X \rightarrow Y$, that associates to the microscopic profile x the corresponding macroscopic profile $y = Px$. We shall assume that there is an integer $N \in \mathbb{N}$, which measures the size of the microscopic system, such that

$$PNP^t = id_Y. \quad (3.1)$$

We shall keep the same framework as in [GOVW09], by considering a measure $\mu(dx) = \exp(-H(x)) dx$ on X , and its decomposition, as induced by the operator P . The measure $\bar{\mu} = P\#\mu$ is the distribution of the macroscopic profile, and for all $y \in Y$, $\mu(dx|y)$ is the conditional distribution of x given $Px = y$. This decomposition induces a natural coarse-graining of the microscopic Hamiltonian H , defined by $\bar{H}(y) := -\frac{1}{N} \log \left(\frac{d\bar{\mu}}{dy} \right)$, so that

$$\bar{\mu}(dy) = \exp(-N\bar{H}(y)) dy.$$

One of the tools frequently used to study particle systems is a logarithmic Sobolev inequality. Let us first recall the definition :

Definition 3.1. Let X be a Riemannian manifold. A probability measure μ on X is said to satisfy a LSI with constant $\rho > 0$ if, for any locally Lipschitz, nonnegative function $f \in L^1(\mu)$,

$$\int f \log(f) d\mu - \left(\int f d\mu \right) \log \left(\int f d\mu \right) \leq \frac{1}{\rho} \int \frac{|\nabla f|^2}{2f} d\mu.$$

There are many criteria and applications for LSI in the literature. [Led01] contains a nice introduction to the topic. One of the main results of [GOVW09] is the following sufficient condition for LSI, based on the two-scale decomposition of μ .

Theorem 3.2 (Two-scale LSI). *Let $\mu(dx) = \exp(-H(x))dx$ be a probability measure on X , and let $P : X \rightarrow Y$ satisfy (3.1). Assume that*

(i) $\kappa :=$

$$\max_{x \in X} \left\{ \langle \text{Hess } H(x) \cdot u, v \rangle, u \in \text{Ran}(NP^t P), v \in \text{Ran}(id_X - NP^t P), |u| = |v| = 1 \right\} \quad (3.2)$$

is finite;

(ii) There is $\rho > 0$ such that $\mu(dx|y)$ satisfies LSI(ρ) for all y ;

(iii) There is $\lambda > 0$ such that $\bar{\mu}$ satisfies LSI(λN).

Then μ satisfies LSI($\hat{\rho}$), with

$$\hat{\rho} := \frac{1}{2} \left(\rho + \lambda + \frac{\kappa^2}{\rho} - \sqrt{(\rho + \lambda + \frac{\kappa^2}{\rho})^2 - 4\rho\lambda} \right) > 0. \quad (3.3)$$

3.3.2 Hydrodynamic limits

Let us now recall the setting of the abstract criterion for hydrodynamic limits proved in [GOVW09]. We endow the space X with a Gibbs probability measure μ , and we consider a positive definite symmetric linear operator $A : X \rightarrow X$. The stochastic dynamics on X that is studied is described by the equation

$$\frac{\partial}{\partial t}(f\mu) = \nabla \cdot (A\nabla f\mu). \quad (3.4)$$

This equation is to be understood in a weak sense. That is, for all smooth test function ξ , we have $\frac{d}{dt} \int \xi(x)f(t,x)\mu(dx) = - \int \nabla \xi(x) \cdot A\nabla f(t,x)\mu(dx)$. Given an initial condition $f(0, \cdot)$ such that $f(0, x)\mu(dx)$ is a probability measure, the solution $f(t, x)$ is at all times the microscopic density of a probability measure with respect to μ .

The aforementioned abstract criterion states that, under certain conditions, and in a precise sense, the macroscopic profile $y = Px$, with law given by $\bar{f}(t, y) = \int_{\{Px=y\}} f(t, x)\mu(dx)$ is close to the solution of the following differential equation :

$$\frac{d\eta}{dt} = -\bar{A}\nabla\bar{H}(\eta(t)) \quad (3.5)$$

where \bar{A} is the symmetric, positive definite operator on Y defined by

$$\bar{A}^{-1} = PA^{-1}NP^t. \quad (3.6)$$

We can now recall the abstract theorem proved in [GOVW09] :

Theorem 3.3. *Let $\mu(dx) = \exp(-H(x))dx$ be a probability measure on X , and let $P : X \rightarrow Y$ satisfy (3.1). We define $M := \dim Y + 1$. Let $A : X \rightarrow X$ be a symmetric, definite positive operator, and $f(t, x)$ and $\eta(t)$ be the solutions of (3.4) and (3.5), with initial data $f(0, \cdot)$ and η_0 respectively. Assume that :*

(i) κ as defined by (3.2) is finite;

(ii) There is $\rho > 0$ such that $\mu(dx|y)$ satisfies LSI(ρ) for all y ;

(iii) There is $\lambda > 0$ such that $\langle \text{Hess } \bar{H}(y)\tilde{y}, \tilde{y} \rangle_Y \geq \lambda \langle \tilde{y}, \tilde{y} \rangle_Y$ for all $y, \tilde{y} \in Y$;

- (iv) There is $\alpha > 0$ such that $\int_X |x|^2 \mu(dx) \leq \alpha N$;
(v) There is $\beta > 0$ such that $\inf_{y \in Y} \bar{H}(y) \geq -\beta$;
(vi) There is $\gamma > 0$ such that for all $x \in X$,

$$|(id_X - NP^t P)x|^2 \leq \gamma M^{-2} \langle x, Ax \rangle_X;$$

- (vii) There are constants C_1 and C_2 such that the initial datum satisfy

$$\int f(0, x) \log f(0, x) \mu(dx) \leq C_1 N \quad \text{and} \quad \bar{H}(\eta_0) \leq C_2.$$

Define

$$\Theta(t) := \frac{1}{2N} \int \left\langle (x - NP^t \eta(t)), A^{-1}(x - NP^t \eta(t)) \right\rangle f(t, x) \mu(dx).$$

Then for any $T > 0$, we have, with $\hat{\rho}$ given by (3.3),

$$\begin{aligned} & \max \left\{ \sup_{0 \leq t \leq T} \Theta(t), \frac{\lambda}{2} \int_0^T \left(\int_Y |y - \eta(t)|_Y^2 \bar{f}(t, y) \bar{\mu}(dy) \right) dt \right\} \\ & \leq \Theta(0) + T \left(\frac{M}{N} \right) + \frac{1}{M^2} \left(\frac{C_1 \gamma \kappa^2}{2\lambda \hat{\rho}^2} \right) + \frac{1}{M} \left[\sqrt{2\gamma T} \left(\alpha + \frac{2C_1}{\hat{\rho}} \right)^{1/2} (\sqrt{C_1} + \sqrt{C_2 + \beta}) \right] \\ & =: \Xi(T, M, N). \end{aligned}$$

This theorem means that, if we consider a sequence of data

$\{X_\ell, Y_\ell, N_\ell, P_\ell, A_\ell, \mu_\ell, f_{0,\ell}, \eta_{0,\ell}\}_{\ell \in \mathbb{N}}$ that satisfies the previous assumptions with uniform constants, and if we assume that

$$M_\ell \uparrow \infty; \quad N_\ell \uparrow \infty; \quad \frac{N_\ell}{M_\ell} \uparrow \infty \quad (3.7)$$

and that the initial data $\Theta_\ell(0)$ goes to 0, then for all $T > 0$ we have

$$\lim_{\ell \uparrow \infty} \sup_{0 \leq t \leq T} \frac{1}{N_\ell} \int (x - N_\ell P_\ell^t \eta_\ell(t)) \cdot A_\ell^{-1}(x - N_\ell P_\ell^t \eta_\ell(t)) f_\ell(t, x) \mu_\ell(dx) = 0 \quad (3.8)$$

and

$$\lim_{\ell \uparrow \infty} \int_0^T \int_{Y_\ell} |y - \eta_\ell(t)|_{Y_\ell}^2 \bar{f}_\ell(t, y) \bar{\mu}(dy) dt = 0.$$

Remark 3.4. As noted in [GOVW09], hypothesis (iii) of this theorem implies hypothesis (iii) of Theorem 3.2 by the Bakry–Émery theorem, a proof of which can be found in [Led01].

Using this result, we will deduce bounds on the relative entropy with respect to a well-chosen local Gibbs state. Let us first give a precise definition of what we mean by a local Gibbs state.

Definition 3.5. Let $\eta \in Y$. The local Gibbs state associated with η is the probability measure on X whose density is given by

$$G^\eta(x) \mu(dx) = Z^{-1} \exp(\vec{\lambda} \cdot x) \mu(dx), \quad \vec{\lambda} = NP^t \nabla \bar{H}(\eta). \quad (3.9)$$

Remark 3.6. Notice that, in this definition, $G^\eta(x)$ only depends on the macroscopic profile Px . This differs from the local Gibbs measure used in [Yau91], which (slowly) varied at the microscopic scale. But here, we force the maximum of the macroscopic probability density to be reached at η , which makes this definition convenient.

We can now formulate our results in this abstract setting :

Theorem 3.7. *Let $G(t, \cdot) = G^{\eta(t)}$ denote the local Gibbs state associated with $\eta(t)$, where $\eta(t)$ solves the macroscopic equation (3.5). Suppose assumptions (i) to (vii) from Theorem 3.3 hold. Further assume that*

(viii) *There is $\tau > 0$ such that $A \geq \tau \text{Id}_X$*

(ix) *The Hessian of \bar{H} is bounded above, i.e. there exists $\Lambda > 0$ such that for all $y \in Y$ we have $\text{Hess } \bar{H}(y) \leq \Lambda \text{Id}$;*

Then

(a) *The relative entropy with respect to the local Gibbs state is controlled as follows :*

$$\int_0^T \frac{1}{N} \int \Phi \left(\frac{f(t, x)}{G(t, x)} \right) G(t, x) \mu(dx) dt = O \left(\sqrt{\Theta(0) + \frac{M}{N} + \frac{1}{M}} \right) \quad (3.10)$$

where the actual constants in the bound depend on $T, \lambda, \alpha, \gamma, \rho, \kappa, \tau, C_1$ and C_2 , but not on M and N ;

(b) *The difference between the microscopic free energy and the free energy associated with the macroscopic profile η is bounded as follows :*

$$\begin{aligned} & \int_0^T \left| \frac{1}{N} \int \Phi(f(t, x)) \mu(dx) - \bar{H}(\eta(t)) \right| dt \\ &= O \left(\sqrt{\Theta(0) + \frac{M}{N} + \frac{1}{M}} \right) \\ &+ O \left(\frac{M}{N} \right) \times \max \left(\left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right) \right|, \left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\lambda N} \right) \right| \right) \end{aligned} \quad (3.11)$$

Remark 3.8 (On the assumptions). Assumption (viii) is always true, since we assumed A to be a positive symmetric operator on X , but I write it down this way because, in the next Corollary, I will require this lower bound to be uniform in N , and setting it this way makes this requirement clearer. When the Hessian of H is bounded above (which will be the case in the next section for the application to Kawasaki dynamics), both assumptions (i) and (ix) will be satisfied. As for $\Gamma(Y, |\cdot|_Y)$, it will have a nice behavior when $|\cdot|_Y$ is comparable to the L^2 norm, as we shall see in the proof of Theorem 3.19.

With this theorem, we can obtain quantitative controls in the hydrodynamic limit :

Corollary 3.9. *Consider a sequence of data $\{X_\ell, Y_\ell, N_\ell, P_\ell, A_\ell, \mu_\ell, f_{0,\ell}, \eta_{0,\ell}\}$ satisfying the previous assumptions, with uniform constants $\alpha, \lambda, \beta, C_1, C_2, \tau$ and Λ . Assume moreover that*

$$N_\ell \rightarrow \infty; \quad M_\ell \rightarrow \infty; \quad \frac{M_\ell}{N_\ell} \rightarrow 0; \quad (3.12)$$

$$\frac{M_\ell}{N_\ell} \log \left(\frac{\Gamma(Y_\ell, |\cdot|_Y)^{1/(M-1)}}{N_\ell} \right) \rightarrow 0 \quad (3.13)$$

and that the sequence of initial data satisfies

$$\Theta_\ell(0) \rightarrow 0.$$

Then we have, for all $T > 0$,

(a')

$$\int_0^T \frac{1}{N} \int \Phi \left(\frac{f_\ell(t, x)}{G_\ell(t, x)} \right) G_\ell(t, x) \mu_\ell(dx) dt \rightarrow 0;$$

(b')

$$\int_0^T \left| \frac{1}{N} \int \Phi(f_\ell(t, x)) \mu_\ell(dx) - \bar{H}_\ell(\eta_\ell(t)) \right| dt \longrightarrow 0.$$

Let us summarize these results in the language of statistical physics :

-The microscopic variables are approximately distributed according to a local Gibbs state, in the sense of relative Kullback information, in a time-integrated sense on $[0, T]$.

-The microscopic free energy converges to the hydrodynamic free energy, in $L^1([0, T])$.

In the next section, in the case of a concrete example, we will reinforce this into a convergence uniformly in time as long as we stay away from zero.

Remark 3.10. Using the Otto-Villani theorem, which states that the Wasserstein distance $W_2(\nu, \mu)^2$ is controlled by the entropy $\text{Ent}_\mu(\nu)$ when μ satisfies a logarithmic Sobolev inequality (see [OV00]), it is possible to show that (a') implies

$$\int_0^T \frac{1}{N} W_2(f(t)\mu, G(t)\mu)^2 dt \longrightarrow 0,$$

with the Wasserstein distance associated to the L^2 structure, rather than the penalized A^{-1} scalar product that appears in [GOVW09]. Since the L^2 norm is strictly stronger than the A^{-1} norm, this shows that our convergence in entropy result is strictly stronger than the convergence (3.8), as long as we integrate in time. We will later see that this convergence also holds pointwise, for strictly positive times, even if it only holds in the weaker A^{-1} sense at time zero.

Our results also imply that, at macroscopic scale, we have

$$\int |y - \eta(t)|_Y^2 \bar{f}(t, y) \bar{\mu}(dy) \longrightarrow 0$$

for any time $t > 0$, while this convergence was only proven in a time-integrated sense in [GOVW09]. This statement follows from the convergence to 0 of $W_2(\bar{f}\bar{\mu}, \bar{G}\bar{\mu})$ and $W_2(\bar{G}\bar{\mu}, \delta_\eta)$, and the triangle inequality for Wasserstein distances.

One of the main tools we shall use is the following interpolation inequality, due to Otto and Villani ([OV00], Theorem 5) :

Theorem 3.11. *Let $\mu(dx) = e^{-H(x)} dx$ be a probability measure on \mathbb{R}^n with a finite moment of order 2 such that $H \in C^2(\mathbb{R}^n)$ and $\text{Hess } H \geq \lambda I_n$, $\lambda \in \mathbb{R}$. Then for any probability measure ν on \mathbb{R}^n that is absolutely continuous with respect to μ , we have*

$$\text{Ent}_\mu(\nu) \leq W_2(\mu, \nu) \sqrt{I_\mu(\nu)} - \frac{\lambda}{2} W_2(\mu, \nu)^2.$$

In particular, if H is convex, then

$$\text{Ent}_\mu(\nu) \leq W_2(\mu, \nu) \sqrt{I_\mu(\nu)}. \quad (3.14)$$

We refer to the original article [OV00] for a proof of this theorem. This will allow us to transform a convergence in Wasserstein distance and a bound on the Fisher information into a convergence of the relative entropy. However, if we apply this result immediately in the microscopic scale, if we use the usual Euclidean structure, the Wasserstein distance between $f\mu$ and the local Gibbs state does not go to zero. And if we use the penalized Euclidean structure $\langle A^{-1}\cdot, \cdot \rangle$, the lower bound on the Hessian will grow too fast, and the additional term $(\inf \text{Hess } H) W_{A^{-1}, 2}(f\mu, G\mu)$ will go to infinity. So, in order to get rid of the additional term, we will go to macroscopic scale, where the Hessian of \bar{H} is convex, and use inequality (3.14).

Remark 3.12. In this context, Kosygina's method would suggest to decompose the macroscopic relative entropy

$$\begin{aligned} \frac{1}{N} \text{Ent}_{\bar{G}\bar{\mu}}(\bar{f}/\bar{G}) &= \frac{1}{N} \int_Y \tilde{f} \log \tilde{f} dy + \int \bar{H}(y) \bar{f}(y) \bar{\mu}(dy) \\ &+ \frac{1}{N} \log \bar{Z} - \int \nabla \bar{H}(\eta) \cdot y \bar{f}(y) \bar{\mu}(dy) \end{aligned}$$

where $\tilde{f} = e^{-N\bar{H}} f$ is the density of the coarse-grained state with respect to the Lebesgue measure. We can get a bound on the time-integral of the sum of the last three terms of the same type as those in Theorem 3.7. So the problem would be to bound $\frac{1}{N} \int_Y \tilde{f} \log \tilde{f} dy$.

For the application to Kawasaki dynamics, Kosygina proved a vanishing upper bound on $\frac{1}{N} \int_{t'}^t \int_Y f(s) \log f(s) dy ds$ for times $t > t' > 0$, which has the same order of magnitude in the system size $N = KM$ as ours. Her proof consists in showing that we can replace the law of our process Kawasaki dynamics with another process for which this problem is easier. This method uses Girsanov's theorem and specific information on the operator A , and I do not know how to replicate it in the abstract setting considered here. Moreover, unlike our method, it does not work when $t' = 0$. However, for discrete spins, the quantity that would play the role of $\int_Y \tilde{f} \log \tilde{f} dy$ is non-positive, which makes her method very convenient when applied to particle systems such as exclusion processes. It is not clear whether the two-scale approach can be successfully applied to the study of discrete systems.

3.3.3 Kawasaki dynamics

We shall now present the application of the two previous theorems to Kawasaki dynamics. We consider a one dimensional N -periodic lattice system with continuous spin variables. The law of each variable is given by a Ginzburg-Landau potential $\psi : \mathbb{R} \rightarrow \mathbb{R}$, which we shall assume to be of the form

$$\psi(x) = \frac{1}{2}x^2 + \delta\psi(x), \quad \|\delta\psi\|_{C^2(\mathbb{R})} < \infty. \quad (3.15)$$

We shall also force the mean spin to take a given value $m \in \mathbb{R}$. That is, the random vector $x = (x_1, \dots, x_N)$ will take its values in the $(N - 1)$ -dimensional hyperplane with mean $m \in \mathbb{R}$:

$$X_{N,m} := \left\{ (x_1, \dots, x_N) \in \mathbb{R}^N; \frac{1}{N} \sum x_i = m \right\}$$

equipped with the ℓ^2 inner product,

$$\langle x, \tilde{x} \rangle_{X_{N,m}} := \sum x_i \tilde{x}_i.$$

We shall consider the canonical ensemble $\mu_{N,m}$, which is the distribution of the random variables x_1, \dots, x_N conditioned on the event that their mean value is given by $m \in \mathbb{R}$. Its density with respect to the Lebesgue measure on $X_{N,m}$ is given by

$$\mu_{N,m}(dx) = \frac{1}{Z} 1_{\sum x_i = Nm} \exp \left(- \sum_{i=1}^N \psi(x_i) \right). \quad (3.16)$$

The logarithmic density H is evidently given by $H(x) = \sum_{i=1}^N \psi(x_i) + \log Z$.

We shall now introduce the macroscopic state necessary to apply the abstract results of the previous section. We first divide the N spins into M blocks. To fix ideas, we shall assume

that all these blocks have the same size K , such that $N = KM$. This assumption is not necessary (all that will be needed is that the sizes of all the blocks are of same order) but it will make things a lot clearer. See Remark 30 of [GOVW09] for a full explanation about this. We will now define the macroscopic variables as the mean of each block. Therefore they form a set of M real numbers that still have mean m . The associated macroscopic space is thus

$$Y_{M,m} := \left\{ (y_1, \dots, y_M) \in \mathbb{R}^M; \frac{1}{M} \sum y_i = m \right\}$$

which we endow with the L^2 inner product

$$\langle y, \tilde{y} \rangle_Y := \frac{1}{M} \sum y_j \tilde{y}_j.$$

Then the projection operator $P_{N,K} : X_{N,m} \rightarrow Y_{M,m}$ that associates to a given microscopic profile its macroscopic profile is given by

$$P_{N,K}(x_1, \dots, x_N) = (y_1, \dots, y_M); \quad y_j = \frac{1}{K} \sum_{i=(j-1)K+1}^{jK} x_i,$$

and it is easy to check that $PNP^t = id_Y$. We can explicitly compute the coarse-grained Hamiltonian \bar{H} :

$$\bar{H}(y) = \frac{1}{M} \sum_{j=1}^M \psi_K(y_j) + \frac{1}{N} \log \bar{Z}$$

where

$$\psi_K(m) = -\frac{1}{K} \log \left(\int_{X_{K,m}} \exp\left(-\sum_{i=1}^K \psi(x_i)\right) dx \right) \quad (3.17)$$

and \bar{Z} is the normalization constant. The gradient and Hessian of \bar{H} are then given by

$$(\nabla_Y \bar{H}(y))_Y = \psi'_K(y_i); \quad (\text{Hess}_Y \bar{H})_{ij} = \psi''_K(y_i) \delta_{ij}. \quad (3.18)$$

As a consequence of the principle of equivalence of ensembles (quantified through a local version of the Cramér theorem), the following proposition explains the behavior of ψ_K when K is large. It was proven in the Appendix of [GOVW09].

Proposition 3.13. *If ψ satisfies (3.15) and ψ_K is defined by (3.17), then*

$$\psi_K \xrightarrow{K \uparrow \infty} \varphi \quad \text{in the uniform } C^2 \text{ topology,}$$

where φ is the Cramér transform of ψ , defined by

$$\varphi(m) = \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \log \int \exp(\sigma x - \psi(x)) dx \right). \quad (3.19)$$

Using this proposition, the strict convexity of φ and the expression of the Hessian (3.18), the following lemma is easily deduced :

Lemma 3.14 (Convexity of the coarse-grained Hamiltonian). *There exists $K_0 < \infty$ and $\lambda > 0$ depending only on ψ such that, for any $K \geq K_0$,*

$$\langle \tilde{y}, \text{Hess } \bar{H}(y) \tilde{y} \rangle_Y \geq \lambda \langle \tilde{y}, \tilde{y} \rangle_Y.$$

This lemma, among others, allowed to apply the abstract criterion for logarithmic Sobolev inequalities to the present setting, and obtain

Theorem 3.15. *Let ψ satisfy (3.15) and let $\mu_{N,m}$ be defined by (3.16). Then there exists $\rho > 0$ such that for any $N \in \mathbb{N}$ and $m \in \mathbb{R}$, $\mu_{N,m}$ satisfies LSI(ρ).*

This result was recently extended in [MO13] to the case where ψ is a bounded perturbation of a uniformly convex function (rather than strictly quadratic), using a technique of iterated coarse-graining.

We shall now present the Kawasaki dynamics for such a system of spins. We (arbitrarily) set the mean m in the setting just explained to be 0, and we consider a dynamics of the form described by (3.4), with the matrix $A = (A_{ij})$ defined by

$$A_{ij} = N^2(-\delta_{i,j-1} + 2\delta_{i,j} - \delta_{i,j+1}) \quad (3.20)$$

We also identify the space $X_{N,0}$ with the space \bar{X} of piecewise constant functions on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$:

$$\bar{X} = \left\{ \bar{x} : \mathbb{T} \rightarrow \mathbb{R}; \quad \bar{x} \text{ is constant on } \left(\frac{j-1}{N}, \frac{j}{N} \right], \quad j = 1, \dots, N \right\}$$

by associating to the vector $x \in X_{N,0}$ the function $\bar{x} \in \bar{X}$ such that

$$\bar{x}(\theta) = x_j, \quad \theta \in \left(\frac{j-1}{N}, \frac{j}{N} \right].$$

To obtain the final hydrodynamic limit, we must embed all of these spaces \bar{X}_N in a common functional space. We consider the space of functions $f : \mathbb{T} \rightarrow \mathbb{R}$ of locally integrable functions of mean zero, which we equip of the following norm :

$$\|f\|_{H^{-1}}^2 = \int_{\mathbb{T}} w^2(\theta) d\theta; \quad w' = f, \quad \int_{\mathbb{T}} w(\theta) d\theta = 0. \quad (3.21)$$

Then the closure of all the spaces \bar{X}_N for this norm is the usual Sobolev space $H^{-1}(\mathbb{T})$. We can now formulate the following theorem on the hydrodynamic limit of the Kawasaki dynamics.

Theorem 3.16. *Assume that ψ satisfies (3.15). Let $f_N = f_N(t, x)$ be a time-dependent probability density on $(X_{N,0}, \mu_{N,0})$ solving*

$$\frac{\partial}{\partial t}(f \mu_{N,0}) = \nabla \cdot (A \nabla f \mu_{N,0})$$

where $f_N(0, \cdot) = f_{0,N}(\cdot)$ satisfies

$$\int f_{0,N}(x) \log f_{0,N}(x) \mu_{N,0}(dx) \leq CN \quad (3.22)$$

for some constant $C > 0$. Assume that

$$\lim_{N \uparrow \infty} \int \|\bar{x} - \zeta_0\|_{H^{-1}}^2 f_{0,N}(x) \mu_{N,0}(dx) = 0 \quad (3.23)$$

for some $\zeta_0 \in L^2(\mathbb{T})$ which has mean zero. Then for any $T > 0$ we have

$$\lim_{N \uparrow \infty} \sup_{0 \leq t \leq T} \int \|\bar{x} - \zeta(t, \cdot)\|_{H^{-1}}^2 f_N(t, x) \mu_{N,0}(dx) = 0, \quad (3.24)$$

where ζ is the unique weak solution of the nonlinear parabolic equation

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) \quad (3.25)$$

with initial condition $\zeta(0, \cdot) = \zeta_0(\cdot)$, where φ is defined as in (3.19).

In this theorem, a weak solution of (3.25) is defined in the following way :

Definition 3.17. We will call $\zeta = \zeta(t, \theta)$ a weak solution of (3.25) on $[0, T] \times \mathbb{T}$ if

$$\zeta \in L_t^\infty(L_\theta^2), \quad \frac{\partial \zeta}{\partial t} \in L_t^2(H_\theta^{-1}), \quad \varphi'(\zeta) \in L_t^2(L_\theta^2), \quad (3.26)$$

and

$$\left\langle \xi, \frac{\partial \zeta}{\partial t} \right\rangle_{H^{-1}} = - \int_{\mathbb{T}^1} \xi \varphi'(\zeta) d\theta \quad \text{for all } \xi \in L^2, \text{ for a.e. } t \in [0, T] \quad (3.27)$$

One of the main steps of the proof, which will also be used in this paper, is the convergence of $\bar{\eta}$ to ζ :

Proposition 3.18. Let $\bar{\eta}_0^\ell \in \bar{Y}_\ell$ be a step function approximation of ζ_0 , η_0^ℓ the vector of Y associated with it, and η_ℓ the solution of (3.5) with initial condition η_0^ℓ . Then the step functions $\bar{\eta}_\ell$ converge strongly in $L_t^\infty(H_\theta^{-1})$ to the unique weak solution of (3.25) with initial condition ζ_0 .

We now state the result obtained when applying the previous abstract theorem to this setting.

Theorem 3.19 (Convergence of the entropy for Kawasaki dynamics). *Under the same assumptions as Theorem 3.16, the relative entropy with respect to the local Gibbs state goes to zero, in a time-integrated sense :*

$$\int_0^T \int_{X_N} \Phi \left(\frac{f_N(t, x)}{G_N(t, x)} \right) G_N(t, x) \mu_N(dx) dt, \quad (3.28)$$

where $G_N(t, \cdot)$ is the local Gibbs state given by $\eta_N(t)$. As a consequence, we have convergence of the microscopic entropy to the hydrodynamic entropy, in a time-integrated sense :

$$\int_0^T \left| \frac{1}{N} \int \Phi(f_N(t, x)) \mu_N(dx) - \left(\int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta - \varphi \left(\int_{\mathbb{T}} \zeta(t, \theta) d\theta \right) \right) \right| dt \xrightarrow{N \rightarrow \infty} 0. \quad (3.29)$$

Moreover, in this setting, we will be able to get a pointwise convergence of the entropy, as long as we stay away from the origin. It will follow from the time-integrated convergence and the fact that the entropy is decreasing in time.

Theorem 3.20 (Pointwise convergence of the relative entropy). *Assume that ζ is continuous in both variables. Let $0 < \epsilon < T$. Then*

$$\frac{1}{N} \int \Phi(f_N(t, x)) \mu_N(dx) \xrightarrow{N \rightarrow \infty} \int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta - \varphi \left(\int_{\mathbb{T}} \zeta(t, \theta) d\theta \right)$$

uniformly on $[\epsilon, T]$.

Remark 3.21. This convergence will in general not hold true at initial time, since no relation is assumed between the initial microscopic entropy and the initial hydrodynamic entropy. However, if it does hold true initially, then it will hold true uniformly on $[0, T]$ for any $T > 0$. (This is the main outcome of Yau's relative entropy method.)

Since we do not necessarily assume our initial data to be smooth, ζ is not in general smooth at $t = 0$. However, as long as ζ_0 lies in $L^2(\mathbb{T})$, ζ will satisfy the smoothness assumptions of Theorem 3.20 :

Proposition 3.22. *Assume φ is a C^3 function, with $\varphi'' \geq \lambda > 0$, $\|\varphi''\|_\infty < \infty$ and $\|\varphi^{(3)}/\varphi''\|_\infty < \infty$. Let ζ the weak solution of*

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\zeta)$$

with initial data $\zeta_0 \in L^2(\mathbb{T})$. Then, for any $\epsilon > 0$, ζ lies in $C^{1,2}([\epsilon, T] \times \mathbb{T})$, and $\frac{\partial}{\partial t} \varphi'(\zeta)$ and $\frac{\partial^2}{\partial \theta^2} \varphi'(\zeta)$ are uniformly continuous.

This result is well-known in the PDE community, we give a proof for the sake of completeness.

3.4 Proof of Theorem 3.7

Let us first state some properties of the Local Gibbs state, which we shall use for the proof.

Proposition 3.23 (Study of the local Gibbs state). *(i) At the macroscopic scale the density $\bar{G}\bar{\mu}$ is given by*

$$\bar{G}(y)\bar{\mu}(dy) = \frac{1}{Z} \exp\left(N(\nabla \bar{H}(\eta) \cdot y - \bar{H}(y))\right) dy.$$

(ii) At the macroscopic scale, we have the following bound of the Wasserstein distance between the local Gibbs state given by $\bar{G}(\cdot)\bar{\mu}$ and δ_η :

$$\int |y - \eta|_Y^2 \bar{G}(y)\bar{\mu}(dy) \leq \frac{M}{\lambda N}.$$

(iii) The free energy associated with G^η is close to the energy associated with η , with the explicit bound

$$\begin{aligned} & \left| \frac{1}{N} \int \Phi(G^\eta) d\mu - \bar{H}(\eta) \right| \\ & \leq \frac{(M-1)}{2N} \max \left(\left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right) \right|, \left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\lambda N} \right) \right| \right) \\ & + \sqrt{\frac{M}{\lambda N}} |\nabla \bar{H}(\eta)|. \end{aligned}$$

Remark 3.24. We can use the same techniques as in [GOVW09] to pass from macroscopic to microscopic scale, and deduce from (ii) a bound on the penalized Wasserstein distance $\frac{1}{N} W_{A^{-1}}(G d\mu, \delta_{NP^t \eta(t)}) = \frac{1}{N} \int \langle A^{-1}(x - NP^t \eta), x - NP^t \eta \rangle G(x) \mu(dx)$.

Part (iii) will be how we deduce the convergence of the entropy from the local Gibbs behavior. The proof of this proposition will hinge on the following lemma, which tells us that among all the probability measures on \mathbb{R}^M of the form $\exp(-f(x))dx$, with $\text{Hess} f \geq \lambda Id$ and where f reaches its minimum in 0, the one with the highest second moment is the centered Gaussian of covariance matrix λid .

Lemma 3.25. *If $f : \mathbb{R}^M \rightarrow \mathbb{R}$ is C^2 and uniformly convex, with $\text{Hess } f \geq \lambda \text{Id}$, $\lambda > 0$, and $\min f = f(0)$, then*

$$\int |x|^2 e^{-f(x)} dx \leq \frac{M}{\lambda} \int e^{-f(x)} dx.$$

The following proof of this lemma was pointed out to us by S.R.S. Varadhan.

Démonstration. Fix $x \in \mathbb{R}^M$. The function $g(t) = f(tx) - \frac{\lambda}{2} \|tx\|_2^2$ is convex, and reaches its minimum for $t = 0$. Therefore, we have $g'(1) \geq g'(0) = 0$. Since $g'(1) = \langle x, \nabla f(x) \rangle - \lambda |x|^2$, we obtain

$$\lambda \int |x|^2 e^{-f(x)} dx \leq \int \langle x, \nabla f(x) \rangle e^{-f(x)} dx$$

and, by integration by parts, the term on the right-hand side is equal to $M \int e^{-f(x)} dx$, which concludes the proof. \square

This lemma will allow us to bound the Wasserstein distance in (ii).

Proof of Proposition 3.23. The proof of (i) is trivial, since we have constructed our local Gibbs state such that $G(x)$ actually only depends on Px . For all t , since $y \rightarrow \psi_t(y) = \bar{H}(y) - \nabla \bar{H}(\eta(t)) \cdot y$ is uniformly convex, with its Hessian bound below by λId , and reaches its minimum for $y = \eta(t)$, applying lemma 3.25 with $f = N\psi_t$, after translating by $\eta(t)$, we obtain

$$\begin{aligned} \int |y - \eta|_Y^2 \bar{G}(y) \bar{\mu}(dy) &= \frac{1}{\bar{Z}} \int |y - \eta|_Y^2 e^{-N\psi_t(y)} dy \\ &\leq \frac{M \bar{Z}}{\lambda N \bar{Z}} \end{aligned} \quad (3.30)$$

which yields (ii).

For (iii), we have

$$\begin{aligned} \left| \frac{1}{N} \int \Phi(G^\eta) d\mu - \bar{H}(\eta) \right| &= \left| \frac{1}{N} \int (N \nabla \bar{H}(\eta) \cdot y - \log \bar{Z}) \bar{G}^\eta(y) \bar{\mu}(dy) - \bar{H}(\eta) \right| \\ &\leq \left| -\frac{1}{N} \log \bar{Z} - \bar{H}(\eta) + \nabla \bar{H}(\eta) \cdot \eta \right| + \left| \frac{1}{N} \int \nabla \bar{H}(\eta) \cdot (y - \eta) \bar{G}(y) \bar{\mu}(dy) \right| \end{aligned} \quad (3.31)$$

Since $\lambda \text{Id} \leq \text{Hess } \bar{H} \leq \Lambda \text{Id}$, we have the bounds

$$\begin{aligned} \frac{\lambda}{2} |y - \eta|^2 &\leq -\nabla \bar{H}(\eta) \cdot (y - \eta) - \bar{H}(\eta) + \bar{H}(y) \\ &\leq \frac{\Lambda}{2} |y - \eta|^2 \end{aligned}$$

for all $y \in Y$. We now multiply by N and integrate. We obtain the upper bound

$$\begin{aligned}
& -\frac{1}{N} \log \bar{Z} - \bar{H}(\eta) + \nabla \bar{H}(\eta) \cdot \eta \\
&= -\frac{1}{N} \log \int \exp(N(\nabla \bar{H}(\eta) \cdot y - \bar{H}(y))) dy - \bar{H}(\eta) + \nabla \bar{H}(\eta) \cdot \eta \\
&= -\frac{1}{N} \log \int \exp(N(\nabla \bar{H}(\eta) \cdot (y - \eta) - (\bar{H}(y) - \bar{H}(\eta)))) dy \\
&\leq -\frac{1}{N} \log \int \exp\left(-\frac{\Lambda N}{2} |y - \eta|_Y^2\right) dy \\
&= -\frac{1}{N} \log \left(\left(\frac{1}{\sqrt{\Lambda N}} \right)^{M-1} \int \exp(-|y|_Y^2/2) dy \right) \\
&= -\frac{M-1}{2N} \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right)
\end{aligned}$$

In the same way, we obtain a lower bound, so that when we take the absolute value we get the bound

$$\begin{aligned}
& \left| \frac{1}{N} \log \bar{Z} - \bar{H}(\eta) + \nabla \bar{H}(\eta) \cdot \eta \right| \\
&\leq \frac{M-1}{2N} \max \left(\left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\lambda N} \right) \right|, \left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right) \right| \right). \quad (3.32)
\end{aligned}$$

Combined with (3.31), this means the final element we need is a bound on the quantity $\int \nabla \bar{H}(\eta) \cdot (y - \eta) \bar{G}^\eta(y) \bar{\mu}(dy)$, and simply using the Cauchy-Schwartz inequality and (ii) gives us the desired result. \square

We shall now use these properties of the local Gibbs State to prove Theorem 3.7.

Proof of Theorem 3.7. We shall divide this proof into three steps : first we shall reduce the problem to the study of the time-integrated relative entropy between the macroscopic state $\bar{f}(t, \cdot) \bar{\mu}$ and the local Gibbs state $G^{\eta(t)} \bar{\mu}$, then we shall use the HWI interpolation inequality to show that this relative entropy goes to 0. Finally, we shall reintroduce the microscopic terms to get the full bound (a). Then (b) shall follow, since we have already proved in Proposition 3.25 that the free energy associated with G^η is asymptotically close to $\bar{H}(\eta(t))$.

Step 1 : Let us consider

$$\begin{aligned}
\frac{1}{N} H_N(t) &= \int \Phi \left(\frac{f(t, x)}{G(t, x)} \right) G(t, x) \mu_N(dx) \\
&= \frac{1}{N} \text{Ent}_\mu(f) - \frac{1}{N} \int f \log(G) \mu(dx) \\
&= \frac{1}{N} \int \text{Ent}_{\mu(dx|y)}(f) \bar{\mu}(dy) + \frac{1}{N} \text{Ent}_{\bar{\mu}}(\bar{f}) - \frac{1}{N} \int f \log(G) \mu(dx) \\
&= \frac{1}{N} \int \text{Ent}_{\mu(dx|y)}(f) \bar{\mu}(dy) + \frac{1}{N} \text{Ent}_{\bar{G}\bar{\mu}}\left(\frac{\bar{f}}{\bar{G}}\right). \quad (3.33)
\end{aligned}$$

where the last equality is obtained because $G(x)$ only depends on the macroscopic state Px . Therefore, to reduce the problem to the study of the macroscopic entropy, we just have to produce an appropriate bound on

$\frac{1}{N} \int \text{Ent}_{\mu(dx|y)}(f) \bar{\mu}(dy)$. Since $\mu(dx|y)$ satisfies the condition LSI(ρ) by assumption (ii), we have

$$\begin{aligned} & \frac{1}{N} \int_Y \text{Ent}_{\mu(dx|y)}(f) \bar{\mu}(dy) \\ & \leq \frac{1}{N\rho} \int_X \frac{|(id_X - P^t NP) \nabla f(x)|^2}{2f(x)} \mu(dx) \\ & \leq \frac{\gamma}{2NM^2\rho} \int \frac{\langle \nabla f(x), A \nabla f(x) \rangle}{f(x)} \mu(dx), \end{aligned} \quad (3.34)$$

where the last inequality is due to hypothesis (vi). By integrating by parts (3.4), as was done in Proposition 24 of [GOVW09], we deduce that

$$\begin{aligned} & \int f(T, x) \log f(T, x) \mu(dx) + \int_0^T \left(\int \frac{\nabla f \cdot A \nabla f}{f}(t, x) \mu(dx) \right) dt \\ & = \int f(0, x) \log f(0, x) \mu(dx). \end{aligned} \quad (3.35)$$

Since, by assumption (vii), $\int \Phi(f(0, x)) \mu(dx) \leq C_1 N$ and the (mathematical) entropy is non-negative, this tells us that

$$\int_0^T \frac{1}{N} \int \text{Ent}_{\mu(dx|y)}(f(t, \cdot)) \bar{\mu}(dy) dt \leq \frac{\gamma C_1}{2M^2\rho}, \quad (3.36)$$

which concludes this first step of the proof.

Step 2 : We shall now study the macroscopic relative entropy. This is where we shall use the HWI inequality (Theorem 3.11), applied at the macroscopic scale with reference measure $\bar{\mu}$. Indeed, since $\bar{G}\bar{\mu}$ is log-concave, we have for all $T > 0$

$$\begin{aligned} & \int_0^T \frac{1}{N} \text{Ent}_{\bar{G}(t, \cdot) \bar{\mu}} \left(\frac{\bar{f}(t, \cdot)}{\bar{G}(t, \cdot)} \right) dt \leq \int_0^T \frac{1}{N} W_2(\bar{f}(t, \cdot) \bar{\mu}, \bar{G}(t, \cdot) \bar{\mu}) \sqrt{I_{\bar{G}(t, \cdot) \bar{\mu}}(\bar{f}(t, \cdot) \bar{\mu})} dt \\ & \leq \sqrt{\int_0^T W_2(\bar{f}(t, \cdot) \bar{\mu}, \bar{G}(t, \cdot) \bar{\mu})^2 dt} \sqrt{\frac{1}{N^2} \int_0^T I_{\bar{G}(t, \cdot) \bar{\mu}}(\bar{f}(t, \cdot) \bar{\mu}) dt} \end{aligned} \quad (3.37)$$

We already know that both $\bar{f}\bar{\mu}$ and $\bar{G}\bar{\mu}$ are asymptotically close to δ_η . By the triangle inequality for the Wasserstein distance,

$$W_2(\bar{f}(t, \cdot) \bar{\mu}, \bar{G}(t, \cdot) \bar{\mu})^2 \leq 2 W_2(\bar{f}(t, \cdot) \bar{\mu}, \delta_{\eta(t)})^2 + 2 W_2(\delta_{\eta(t)}, \bar{G}(t, \cdot) \bar{\mu})^2 \quad (3.38)$$

Theorem 3.3 states that

$$\begin{aligned} & \int_0^T W_2(\bar{f}(t, \cdot) \bar{\mu}, \delta_{\eta(t)})^2 dt = \int_0^T \int |y - \eta(t)|_Y^2 \bar{f}(t, y) \bar{\mu}(dy) dt \\ & \leq \frac{2}{\lambda} \Xi(T, M, N) \end{aligned} \quad (3.39)$$

where Ξ was defined in Theorem 3.3. Moreover, part (ii) of Proposition 3.23 tells us that

$$\int_0^T \int |y - \eta(t)|_Y^2 \bar{G}(t, y) \bar{\mu}(dy) dt \leq T \frac{M}{\lambda N} \quad (3.40)$$

so we have a bound on the time-integral of the Wasserstein distance that, under suitable assumptions, will go to 0.

We must now produce a bound on the macroscopic Fisher information. We have

$$\begin{aligned}
\frac{1}{N^2} I_{\bar{G}\bar{\mu}} \left(\frac{\bar{f}}{\bar{G}} \right) &= \frac{1}{N^2} \int \frac{|\nabla(\bar{f}/\bar{G})|^2}{\bar{f}/\bar{G}} \bar{G} d\bar{\mu} \\
&= \frac{1}{N^2} \int \frac{|(\nabla\bar{f})/\bar{G} - \bar{f}\nabla\bar{G}/\bar{G}^2|^2}{\bar{f}} \bar{G}^2 d\bar{\mu} \\
&\leq \frac{2}{N^2} \int \frac{|\nabla\bar{f}|^2}{\bar{f}} d\bar{\mu} + \frac{2}{N^2} \int \frac{|\nabla\bar{G}|^2}{\bar{G}^2} \bar{f} d\bar{\mu} \\
&= \frac{2}{N^2} \int \frac{|\nabla\bar{f}|^2}{\bar{f}} d\bar{\mu} + 2|\nabla\bar{H}(\eta)|^2
\end{aligned} \tag{3.41}$$

Now, since we have a lower bound τ on the spectral values of A , $1/\tau$ is an upper bound on the spectral values of A^{-1} . Since for any $y \in Y$

$$\begin{aligned}
\langle \bar{A}^{-1}y, y \rangle_Y &= \langle PA^{-1}NP^t y, y \rangle_Y \\
&= \frac{1}{N} \langle A^{-1}NP^t y, NP^t y \rangle_X \\
&\leq \frac{1}{N\tau} \langle NP^t y, NP^t y \rangle_X \\
&= \frac{1}{\tau} |y|_Y^2,
\end{aligned}$$

so $1/\tau$ is an upper bound on the spectral values of \bar{A}^{-1} , and thus τ is also a lower bound on the spectral values of \bar{A} . Therefore

$$\begin{aligned}
\int_0^T |\nabla\bar{H}(\eta(t))|^2 dt &\leq \frac{1}{\tau} \int_0^T \langle \bar{A}\nabla\bar{H}(\eta(t)), \nabla\bar{H}(\eta(t)) \rangle dt \\
&= -\frac{1}{\tau} \int_0^T \left\langle \frac{d\eta}{dt}(t), \nabla\bar{H}(\eta(t)) \right\rangle dt \\
&= \frac{1}{\tau} (\bar{H}(\eta(0)) - \bar{H}(\eta(T))) \\
&\leq \frac{C_2 + \beta}{\tau}.
\end{aligned} \tag{3.42}$$

To obtain a bound on $\int_0^T \int \frac{|\nabla\bar{f}|^2}{\bar{f}} d\bar{\mu} dt$, we shall use the following proposition, that was proved in Proposition 20 of [GOVW09].

Proposition 3.26. *Assume that κ as given by (3.2) is finite and that for all $y \in Y$, $\mu(dx|y)$ satisfies LSI(ρ). Then, for any positive, C^1 function on X one has, for any $y \in Y$ and $s \in (0, 1)$*

$$\begin{aligned}
\frac{1}{N} \frac{|\nabla_Y \bar{f}(y)|_Y^2}{\bar{f}(y)} &\leq \frac{1}{1-s} \left(\frac{\kappa^2}{\rho^2} \right) \int \frac{|(id_X - PNP^t)\nabla f(x)|^2}{f(x)} \mu(dx|y) \\
&\quad + \frac{1}{s} \int \frac{|PNP^t \nabla f(x)|^2}{f(x)} \mu(dx|y)
\end{aligned} \tag{3.43}$$

Applying this bound to the density f with $s = \frac{\rho^2}{\kappa^2 + \rho^2}$ gives us the bound

$$\begin{aligned}
\frac{1}{N^2} \int_0^T \int \frac{|\nabla \bar{f}|^2}{\bar{f}} d\bar{\mu} dt &\leq \frac{\kappa^2 + \rho^2}{\rho^2} \frac{1}{N} \int_0^T \int \frac{|\nabla f(x)|^2}{f(x)} \mu(dx) dt \\
&\leq \frac{\kappa^2 + \rho^2}{\rho^2} \frac{1}{N} \frac{1}{\tau} \int_0^T \int \frac{\nabla f \cdot A \nabla f}{f} d\mu dt \\
&\leq \frac{C_1(\kappa^2 + \rho^2)}{\tau \rho^2}
\end{aligned} \tag{3.44}$$

where the final inequality was already proved in step 1.

Combining (3.37), (3.38), (3.39), (3.40), (3.41), (3.42) and (3.44) gives us the bound

$$\begin{aligned}
&\int_0^T \frac{1}{N} Ent_{\bar{G}(t, \cdot) \bar{\mu}} \left(\frac{\bar{f}(t, \cdot)}{\bar{G}(t, \cdot)} \right) dt \\
&\leq \sqrt{\frac{2TM}{N} + \frac{4}{\lambda} \Xi(T, M, N)} \\
&\quad \times \sqrt{2 \frac{C_2 + \beta}{N^2 \tau} + 2 \frac{C_1(\kappa^2 + \rho^2)}{\tau \rho^2}}.
\end{aligned} \tag{3.45}$$

This concludes Step 2.

Step 3 : Recombining (3.36) and (3.45) gives us the full bound

$$\begin{aligned}
&\int_0^T \frac{1}{N} \int \Phi \left(\frac{f(t, x)}{G(t, x)} \right) G(t, x) \mu_N(dx) dt \\
&\leq \sqrt{\frac{2TM}{N} + \frac{4}{\lambda} \Xi(T, M, N)} \times \sqrt{2 \frac{C_2 + \beta}{N^2 \tau} + 2 \frac{C_1(\kappa^2 + \rho^2)}{\tau \rho^2}} + \frac{\gamma C_1}{2M^2 \rho}.
\end{aligned} \tag{3.46}$$

To prove (b), we have

$$\begin{aligned}
&\frac{1}{N} \int \Phi \left(\frac{f(t, x)}{G(t, x)} \right) G(t, x) \mu_N(dx) \\
&= \left| \frac{1}{N} \int f(t, x) \log f(t, x) \mu(dx) - \frac{1}{N} \int f(t, x) \log G(t, x) \mu(dx) \right| \\
&= \left| \frac{1}{N} \int f(t, x) \log f(t, x) \mu(dx) - \frac{1}{N} \int \bar{f}(t, y) \log \bar{G}(t, y) \bar{\mu}(dy) \right|
\end{aligned}$$

and thus

$$\begin{aligned}
& \int_0^T \left| \frac{1}{N} \int f(t, x) \log f(t, x) \mu(dx) - \bar{H}(\eta(t)) \right| dt \\
& \leq \int_0^T \left| \frac{1}{N} \int f(t, x) \log f(t, x) \mu(dx) - \frac{1}{N} \int f(t, x) \log G(t, x) \mu(dx) \right| dt \\
& \quad + \int_0^T \left| \frac{1}{N} \int f(t, x) \log G(t, x) \mu(dx) - \bar{H}(\eta(t)) \right| dt \\
& \leq \int_0^T \frac{1}{N} \int \Phi \left(\frac{f(t, x)}{G(t, x)} \right) G(t, x) \mu_N(dx) dt \\
& \quad + \int_0^T \left| \frac{1}{N} \int \bar{f}(t, y) \log \bar{G}(t, y) \bar{\mu}(dy) - \bar{H}(\eta(t)) \right| dt \\
& = \int_0^T \frac{1}{N} \int \Phi \left(\frac{f(t, x)}{G(t, x)} \right) G(t, x) \mu_N(dx) dt \\
& \quad + \int_0^T \left| \frac{1}{N} \int (\bar{f}(t, y) - \bar{G}(t, y)) \log \bar{G}(t, y) \bar{\mu}(dy) \right| dt \\
& \quad + \int_0^T \left| \frac{1}{N} \int \Phi(G(t, \cdot)) d\mu_N - \bar{H}(\eta(t)) \right| dt. \tag{3.47}
\end{aligned}$$

This leaves us with three quantities to bound. The first one is exactly the quantity that is bounded by (a). The third quantity can be bounded using part (iii) of Proposition **3.23** :

$$\begin{aligned}
& \int_0^T \left| \frac{1}{N} \int \Phi(G^\eta) d\mu_N - \bar{H}(\eta) \right| dt \\
& \leq \frac{T(M-1)}{2N} \max \left(\left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right) \right| \left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\lambda N} \right) \right| \right) \\
& \quad + \int_0^T \sqrt{\frac{M}{\lambda N}} |\nabla \bar{H}(\eta)| dt \\
& \leq \frac{T(M-1)}{2N} \max \left(\left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right) \right| \left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\lambda N} \right) \right| \right) \\
& \quad + \sqrt{\frac{TM C_2 + \beta}{\lambda N} \frac{1}{\tau}} \tag{3.48}
\end{aligned}$$

where the final inequality is due to (3.42). To conclude, we just have to bound the second quantity in the right-hand side of (3.47). This will be possible because, as $\log \bar{G}$ is slowly varying (it is an affine function), since $\bar{f}\mu$ and $\bar{G}\mu$ are close for the second Wasserstein distance, when integrating against $\log \bar{G}$ they act the same way.

$$\begin{aligned}
& \int_0^T \left| \frac{1}{N} \int (\bar{f}(t, y) - \bar{G}(t, y)) \log \bar{G}(t, y) \bar{\mu}(dy) \right| dt \\
&= \int_0^T \left| \frac{1}{N} \int (\bar{f}(t, y) - \bar{G}(t, y)) (N \nabla \bar{H}(\eta(t)) \cdot y) \bar{\mu}(dy) \right| dt \\
&\leq \int_0^T \left| \int \langle (y - \eta(t)), \nabla \bar{H}(\eta(t)) \rangle \bar{f}(t, y) \bar{\mu}(dy) \right| dt \\
&\quad + \int_0^T \left| \int \langle (y - \eta(t)), \nabla \bar{H}(\eta(t)) \rangle \bar{G}(t, y) \bar{\mu}(dy) \right| dt \\
&\leq \sqrt{\int_0^T \int |y - \eta(t)|_Y^2 \bar{f}(t, y) \bar{\mu}(dy) dt} \sqrt{\int_0^T |\nabla \bar{H}(\eta(t))|_Y^2 dt} \\
&\quad + \sqrt{\int_0^T \int |y - \eta(t)|_Y^2 \bar{G}(t, y) \bar{\mu}(dy) dt} \sqrt{\int_0^T |\nabla \bar{H}(\eta(t))|_Y^2 dt} \\
&\leq \sqrt{\frac{2(C_2 + \beta) \Xi(T, M, N)}{\lambda \tau}} + \sqrt{\frac{(C_2 + \beta) M}{N \lambda \tau}} \tag{3.49}
\end{aligned}$$

which was the last element needed to get the full bound (b). □

3.5 Application to Kawasaki Dynamics

3.5.1 Proof of Theorem 3.19

We shall now prove Theorem 3.19 as a consequence of Corollary 3.9. We consider, in the same way as in [GOVW09], a sequence of step functions $\bar{\eta}_{0,\ell} \in \bar{Y}_\ell$ such that

$$\|\bar{\eta}_{0,\ell} - \zeta_0\|_{L^2} \longrightarrow 0,$$

canonically associate to each of them a vector $\eta_{0,\ell} \in Y_\ell$ and consider the solution η_ℓ of

$$\frac{d\eta_\ell}{dt} = -\bar{A} \nabla \bar{H}(\eta_\ell), \quad \eta_\ell(0) = \eta_{0,\ell}.$$

We also assume (3.12), that is

$$N_\ell \rightarrow \infty; \quad M_\ell \rightarrow \infty; \quad \frac{M_\ell}{N_\ell} \rightarrow 0; \tag{3.50}$$

which, in this setting, will imply (3.13).

The following proposition was proven in [GOVW09] :

Proposition 3.27. *With the above notations, the step functions $\bar{\eta}_\ell$ converge strongly in $L^\infty(H^{-1})$ to the unique weak solution of*

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\zeta), \quad \zeta(0, \cdot) = \zeta_0.$$

We first have to check that the assumptions of Theorem 3.7 hold with uniform constants. It has already been checked in [GOVW09] that this is the case for assumptions (i) to (vii), so we just have to check assumptions (viii) and (ix).

- It is easy to compute the spectral values of A , since it is a circulant matrix. We have

$$\text{Sp}(A) = \left\{ 2N^2 \left(1 - \cos\left(\frac{2k\pi}{N}\right) \right), \quad k = 0, \dots, N-1 \right\}$$

as an operator on \mathbb{R}^N , and the spectral value 0 corresponds to the action of A on $\mathbb{R}(1, \dots, 1)$, which we don't take into account, since we only consider the action of A on the hyperplane of mean 0. The lowest spectral value of A is then

$$\inf \text{Sp}(A) = 2N^2 \left(1 - \cos\left(\frac{2\pi}{N}\right) \right) \xrightarrow{N \uparrow \infty} 4\pi^2 > 0.$$

Since the sequence of lowest spectral values converges to a strictly positive limit, we have a strictly positive lower bound on the whole sequence, which proves assumption (viii) with a uniform constant τ .

- Since, by Proposition **3.15**, ψ''_K converges uniformly to φ'' and $(\text{Hess}\bar{H}(y))_{ij} = \psi''_K(y_i)\delta_{ij}$, to prove assumption (ix) with a uniform constant, we just have to prove that φ'' is bounded above. This was actually already proved in ([GOVW09], Lemma 41), where is proved both a lower and an upper bound on the second derivative of $\varphi^*(\sigma) = \log \int \exp(\sigma x - \psi(x)) dx$. It is a property of the Legendre transform that, if f is a strictly convex function, its Legendre transform f^* satisfies $(f^*)^* = f$ and $(f^*)' = (f')^{-1}$, so the strictly positive upper and lower bounds on the second derivative of φ^* translate into strictly positive upper and lower bounds on φ'' .

We will also check that, in this case, (3.12) implies (3.13). This is easy to check : since we have $|y|_Y^2 = \frac{1}{M} \sum y_i^2$, we can explicitly compute

$$\Gamma(Y, |\cdot|_Y) = (\sqrt{2\pi M})^{M-1}, \quad (3.51)$$

and (3.13) follows.

We can therefore apply Corollary **3.9**. (3.28) is then a direct application of (a'), so we will concentrate on the proof of (3.29). Part (b') of Corollary **3.9** states that

$$\int_0^T \left| \frac{1}{N} \int \Phi(f_\ell(t, x)) \mu_\ell(dx) - \bar{H}_\ell(\eta_\ell(t)) \right| dt \longrightarrow 0,$$

so we now just have to prove that

$$\int_0^T \left| \bar{H}_\ell(\eta_\ell(t)) - \int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta + \varphi \left(\int_{\mathbb{T}} \zeta(t, \theta) d\theta \right) \right| dt \xrightarrow{t \uparrow \infty} 0. \quad (3.52)$$

We have the expression

$$\begin{aligned} \bar{H}(y) &= \frac{1}{M} \sum \psi_K(y_i) + \frac{1}{N} \log \bar{Z} \\ &= \int_{\mathbb{T}} \psi_K(\bar{y}) d\theta + \frac{1}{N} \log \bar{Z} \end{aligned} \quad (3.53)$$

As a consequence of Proposition **3.27**, we shall prove that $\int_{\mathbb{T}} \psi_K(\bar{\eta}) d\theta$ converges to $\int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta$ in a time-integrated sense, and then we shall prove that $\frac{1}{N} \log \bar{Z}$ converges to $-\varphi(\int_{\mathbb{T}} \zeta(t, \theta) d\theta)$, which will yield (3.52). By the triangle inequality

$$\begin{aligned} &\left| \int_{\mathbb{T}} \psi_K(\bar{\eta}(\theta, t)) d\theta - \int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta \right| \\ &\leq \int_{\mathbb{T}} |\psi_K(\bar{\eta}_\ell(t, \theta)) - \varphi(\bar{\eta}_\ell(t, \theta))| d\theta + \int_{\mathbb{T}} |\varphi(\bar{\eta}_\ell(t, \theta)) - \varphi(\zeta(t, \theta))| d\theta. \end{aligned} \quad (3.54)$$

But

$$\int_{\mathbb{T}} |\psi_K(\bar{\eta}_\ell(t, \theta)) - \varphi(\bar{\eta}_\ell(t, \theta))| d\theta \leq \|\psi_K - \varphi\|_\infty \xrightarrow{K \rightarrow \infty} 0$$

and by convexity, and since $\varphi'' \leq \Lambda$,

$$\begin{aligned} \varphi'(\zeta(t, \theta)) (\bar{\eta}_\ell(t, \theta) - \zeta(t, \theta)) &\leq \varphi(\bar{\eta}_\ell(t, \theta)) - \varphi(\zeta(t, \theta)) \\ &\leq \varphi'(\zeta(t, \theta)) (\bar{\eta}_\ell(t, \theta) - \zeta(t, \theta)) + \frac{\Lambda}{2} |\bar{\eta}_\ell(t, \theta) - \zeta(t, \theta)|^2 \end{aligned} \quad (3.55)$$

We know that $\bar{\eta}$ converges to ζ in $L^\infty(H^{-1})$. Since

$$\int_0^T \int_{\mathbb{T}} |\varphi'(\zeta(t, \theta)) (\bar{\eta}_\ell(t, \theta) - \zeta(t, \theta))| d\theta dt \leq \sqrt{\int_0^T \|\varphi'(\rho(t))\|_{H^1}^2 dt} \sqrt{\int_0^T \|\bar{\eta}_\ell(t) - \zeta(t)\|_{H^{-1}}^2 dt}$$

and $\varphi'(\zeta) \in L^2(H^1)$ we deduce that $\int_{\mathbb{T}} |\varphi(\bar{\eta}_\ell(t, \theta)) - \varphi(\zeta(t, \theta))| d\theta$ converges to 0 in a time-integrated sense, and thus

$$\int_0^T \left| \int_{\mathbb{T}} \psi_K(\bar{\eta}_\ell(\theta, t)) d\theta - \int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta \right| dt \longrightarrow 0.$$

Note that, if we have a time-uniform bound on $\|\varphi'(\rho(t))\|_{H^1}$, this convergence actually holds uniformly in time. In the proof of Proposition 3.22, we show that such a bound holds on time intervals $[\epsilon, +\infty[$, for any $\epsilon > 0$.

To conclude the proof of (3.52), it is enough to prove that

$$\frac{1}{N} \log \bar{Z} \longrightarrow \varphi \left(\int_{\mathbb{T}} \zeta(t, \theta) d\theta \right) \quad (3.56)$$

uniformly in time.

First of all, recall that at any time t we have $\int_{\mathbb{T}} \zeta(t, \theta) d\theta = m$. Also recall that

$$\bar{Z} = \int \exp \left(-\frac{N}{M} \sum_{i=1}^M \psi_K(y_i) \right) dx.$$

Since $y \mapsto \frac{1}{M} \sum_{i=1}^M \psi_K(y_i)$ is strictly convex, it has a unique minimum on Y , and since the variables are exchangeable this minimum can only be reached for $y_1 = \dots = y_M$, and by definition of Y this can only be the case if all the y_i are equal to m . Since $\lambda \leq \psi_K'' \leq \Lambda$ and $\|y\|_Y = \frac{1}{M} \sum_{i=1}^M |y_i|^2$, by convexity, for all $y \in Y$ we have

$$\psi_K(m) + \frac{\lambda}{2M} \|y - m\|_2^2 \leq \frac{1}{M} \sum_{i=1}^M \psi_K(y_i) \leq \psi_K(m) + \frac{\Lambda}{2M} \|y - m\|_2^2$$

where $\|\cdot\|_2$ is the usual Euclidean norm, and we identify the mean m and the vector of Y where all coordinates are equal to m . We take the exponential of this inequality multiplied by $-N$ and integrate, which gives us, since for any $y \in Y$, $y - m$ is of mean 0,

$$\begin{aligned} -\inf \psi_K + \frac{1}{N} \log \int_{\mathbb{R}^{M-1}} \exp \left(-\frac{\Lambda N}{2M} \|y\|_2^2 \right) dy &\leq \frac{1}{N} \log \bar{Z} \\ &\leq -\inf \psi_K + \frac{1}{N} \log \int_{\mathbb{R}^{M-1}} \exp \left(-\frac{\lambda N}{2M} \|y\|_2^2 \right) dy. \end{aligned}$$

Since

$$\frac{1}{N} \log \int_{\mathbb{R}^{M-1}} \exp\left(-\frac{\Lambda N}{2M} \|y\|_2^2\right) dy = \frac{M-1}{2N} \log\left(\frac{\Lambda N}{2\pi}\right) \rightarrow 0$$

and the same goes for $\frac{1}{N} \log \int \exp(-\frac{\Lambda N}{2M} \|y\|_2^2) dy$, we deduce that $|\frac{1}{N} \log \bar{Z} + \inf \psi_K|$ goes to 0 uniformly in time. Finally, since ψ_K converges uniformly to φ , $\psi_K(m)$ converges to $\varphi(m)$, which implies the desired result.

3.5.2 Proof of Theorem 3.20

We shall now use the time-integrated convergence of the entropy we just proved to show that this convergence actually holds pointwise. Our proof closely follows an idea of [Kos01a]. This method was pointed out to us by the (anonymous) referee. It is also possible to deduce the pointwise convergence from the time integrated convergence by using the relative entropy method devised in [Yau91], but this yields a much longer proof.

In a first step, we will show pointwise convergence of the entropy, by showing that

$$\liminf \frac{1}{N} \text{Ent}_{\mu_N}(f_N(t)) \geq \int \varphi(\zeta(t, \theta)) d\theta - \varphi\left(\int \zeta(t, \theta) d\theta\right) \quad (3.57)$$

and

$$\limsup \frac{1}{N} \text{Ent}_{\mu_N}(f_N(t)) \leq \int \varphi(\zeta(t, \theta)) d\theta - \varphi\left(\int \zeta(t, \theta) d\theta\right). \quad (3.58)$$

In a second step, we will show that this pointwise convergence actually holds uniformly in time, as long as we stay away from time $t = 0$.

Let us start with the upper bound. We know that

$$\frac{d}{dt} \int f(t, x) \log f(t, x) \mu(dx) = - \int \frac{\langle A \nabla f, \nabla f \rangle}{f} d\mu \leq 0,$$

so that, for any N , the entropy $\frac{1}{N} \text{Ent}_{\mu_N}(f_N)$ is decreasing in time. This is just the H-theorem expressed in the context of our model.

Therefore, for any N , any $t > 0$ and ϵ small enough, we have

$$\frac{1}{N} \text{Ent}_{\mu_N}(f_N(t)) \leq \frac{1}{\epsilon} \int_{t-\epsilon}^t \frac{1}{N} \text{Ent}_{\mu_N}(f_N(s)) ds \quad (3.59)$$

we know from Theorem 3.19 that $\int_{t-\epsilon}^t \frac{1}{N} \text{Ent}_{\mu_N}(f_N(s)) ds$ converges to $\int_{t-\epsilon}^t \int_{\mathbb{T}} \varphi(\zeta(s, \theta)) d\theta - \varphi\left(\int \zeta(s, \theta) d\theta\right) ds$. Therefore, for any $t > 0$ and any ϵ small enough, we have

$$\limsup \frac{1}{N} \text{Ent}_{\mu_N}(f_N(t)) \leq \frac{1}{\epsilon} \int_{t-\epsilon}^t \int_{\mathbb{T}} \varphi(\zeta(s, \theta)) d\theta - \varphi\left(\int \zeta(s, \theta) d\theta\right) ds.$$

Since ρ is smooth, by Proposition 3.22, letting ϵ go to zero yields (3.58). (3.57) can be obtained in the same way, by using the inequality

$$\frac{1}{N} \text{Ent}_{\mu_N}(f_N(t)) \geq \frac{1}{\epsilon} \int_t^{t+\epsilon} \frac{1}{N} \text{Ent}_{\mu_N}(f_N(s)) ds.$$

Since the functions $t \rightarrow \frac{1}{N} \text{Ent}_{\mu_N}(f_N(t))$ are continuous and decreasing, and the function $t \rightarrow \int \varphi(\zeta(t, \theta)) d\theta - \varphi\left(\int \zeta(t, \theta) d\theta\right)$ is continuous, Dini's second theorem implies that this pointwise convergence is actually uniform on the compact sets $[\epsilon, T]$, for any $T > \epsilon > 0$.

3.5.3 Proof of Proposition 3.22

To prove the regularity of the solution of the hydrodynamic equation, we shall use the following interpolation inequality, which is a particular case of a family of inequalities that can be found in the second chapter of [LSU68].

Lemma 3.28. *For any $u \in H^1(\mathbb{T})$ with $\int_{\mathbb{T}} u d\theta = 0$ we have*

$$\|u\|_{L^4} \leq 2^{1/4} \|u\|_{L^2}^{3/4} \|u'\|_{L^2}^{1/4}.$$

Démonstration. Let us take such a function u . We have

$$|u(\theta)|^4 = |u(\theta)|^2 |u(\theta)|^2 \leq |u(\theta)|^2 \left(\int_{\mathbb{T}} 2|u(s)| |u'(s)| ds \right)$$

Using Hölder's inequality, we have

$$\int_{\mathbb{T}} |u(s)| |u'(s)| ds \leq \|u\|_{L^2} \|u'\|_{L^2},$$

so that

$$\int_{\mathbb{T}} |u(\theta)|^4 d\theta \leq 2 \|u\|_{L^2}^3 \|u'\|_{L^2}$$

and the result immediately follows. \square

To prove the regularity of our function, we shall prove bounds on the L^2 norms of the derivatives of $\varphi'(\zeta)$, using differential inequalities, and then Sobolev injections. We first have

$$\begin{aligned} \frac{d}{dt} \int \zeta(t, \theta)^2 d\theta &= 2 \int \zeta \frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) d\theta \\ &= -2 \int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right) \frac{\partial \zeta}{\partial \theta} d\theta \\ &= -2 \int \varphi''(\zeta) \left(\frac{\partial \zeta}{\partial \theta} \right)^2 d\theta \\ &\leq -2(\inf \varphi'') \int \left(\frac{\partial \zeta}{\partial \theta} \right)^2 d\theta \end{aligned} \tag{3.60}$$

Integrating this inequality yields

$$\begin{aligned} \int_0^T \int \left(\frac{\partial \zeta}{\partial \theta} \right)^2 d\theta dt &\leq \frac{1}{2(\inf \varphi'')} \left(\|\zeta(0, \cdot)\|_{L^2}^2 - \|\zeta(T, \cdot)\|_{L^2}^2 \right) \\ &\leq \frac{1}{2(\inf \varphi'')} \|\zeta(0, \cdot)\|_{L^2}^2. \end{aligned} \tag{3.61}$$

Since $\|\zeta(0, \cdot)\|_{L^2}^2$ is finite, we obtain

$$\int_0^\infty \int \left(\frac{\partial \zeta}{\partial \theta} \right)^2 d\theta dt < \infty.$$

Moreover, since φ'' is bounded, we also get the bound on $\partial \varphi'(\zeta) / \partial \theta = \varphi''(\zeta) \partial \zeta / \partial \theta$:

$$\int_0^\infty \int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta dt < \infty. \tag{3.62}$$

We then have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta &= \int \frac{\partial \varphi'(\zeta)}{\partial \theta} \frac{\partial}{\partial \theta} \left(\varphi''(\zeta) \frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) \right) d\theta \\ &= - \int \varphi''(\zeta) \left(\frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) \right)^2 d\theta \\ &\leq -\lambda \int \left(\frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) \right)^2 d\theta \end{aligned} \quad (3.63)$$

$$\leq -\lambda \pi^2 \int \left(\frac{\partial}{\partial \theta} \varphi'(\zeta) \right)^2 d\theta \quad (3.64)$$

where the last inequality is a consequence of the Poincaré inequality $\|u\|_{L^2} \leq \pi \|u'\|_{L^2}$ for all functions in $H^1(\mathbb{T})$ with mean zero.

Combining (3.62) and (3.64), we get for any $t_1 > t_2 > 0$

$$\int_{\mathbb{T}} \left(\frac{\partial \varphi'(\zeta(t_2, \theta))}{\partial \theta} \right)^2 d\theta \leq \frac{C}{t_1} \exp(-2\lambda\pi^2(t_2 - t_1)). \quad (3.65)$$

Moreover, using (3.63), we get

$$\int_{\epsilon}^T \int \left(\frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) \right)^2 d\theta \leq \frac{CT}{\epsilon} \quad (3.66)$$

for any $0 < \epsilon < T$.

In the same way, we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta &= - \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \frac{\partial}{\partial \theta} \left(\varphi''(\zeta) \frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \\ &= - \int \varphi''(\zeta) \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta - \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi''(\zeta)}{\partial \theta} \right) d\theta \\ &\leq -\lambda \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta - \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi''(\zeta)}{\partial \theta} \right) d\theta. \end{aligned} \quad (3.67)$$

A simple calculation yields

$$\frac{\partial \varphi''(\zeta)}{\partial \theta} = \frac{\varphi^{(3)}(\zeta)}{\varphi''(\zeta)} \frac{\partial \varphi'(\zeta)}{\partial \theta}$$

and our assumption of boundedness on $\varphi^{(3)}/\varphi''$ then yields the bound

$$\begin{aligned} &\left| \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi''(\zeta)}{\partial \theta} \right) d\theta \right| \\ &\leq C \int \left| \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right) \right| d\theta. \end{aligned}$$

Using Hölder's inequality, we then have

$$\begin{aligned} &\left| \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi''(\zeta)}{\partial \theta} \right) d\theta \right| \\ &\leq \left(\int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta \right)^{1/2} \left(\int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^4 d\theta \right)^{1/4} \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^4 d\theta \right)^{1/4} \end{aligned} \quad (3.68)$$

By an application of Lemma 3.28, we have

$$\left(\int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^4 d\theta \right)^{1/4} \leq C \left(\int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta \right)^{3/8} \left(\int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta \right)^{1/8} \quad (3.69)$$

and

$$\left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^4 d\theta \right)^{1/4} \leq C \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right) d\theta \right)^{3/8} \left(\int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta \right)^{1/8}. \quad (3.70)$$

Plugging (3.69) and (3.70) into (3.68), we get

$$\begin{aligned} & \left| \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right) d\theta \right| \\ & \leq \left(\int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta \right)^{5/8} \left(\int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta \right)^{1/2} \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta \right)^{3/8}. \end{aligned} \quad (3.71)$$

Using the classical interpolation inequality $\|u'\|_{L^2}^2 \leq \|u\|_{L^2} \|u''\|_{L^2}$, we get

$$\left(\int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta \right)^{1/2} \leq \left(\int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta \right)^{1/4} \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta \right)^{1/4}$$

and therefore

$$\begin{aligned} & \left| \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right) d\theta \right| \\ & \leq \left(\int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta \right)^{7/8} \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta \right)^{5/8}. \end{aligned} \quad (3.72)$$

Finally, using Young's inequality $ab \leq 7a^{8/7}/8 + b^8/8$, we get for any $\delta > 0$

$$\begin{aligned} & \left| \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right) \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right) \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right) d\theta \right| \\ & \leq C\delta^{8/7} \left(\int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta \right) + \frac{C}{\delta^8} \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta \right)^5. \end{aligned} \quad (3.73)$$

Taking δ small enough and inserting this inequality into (3.67), we get

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta \\ & \leq -\frac{\lambda}{2} \int \left(\frac{\partial^3 \varphi'(\zeta)}{\partial \theta^3} \right)^2 d\theta + C \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta \right)^5 \\ & \leq -\frac{\lambda}{2\pi^2} \int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta + C \left(\int \left(\frac{\partial \varphi'(\zeta)}{\partial \theta} \right)^2 d\theta \right)^5. \end{aligned} \quad (3.74)$$

Combining (3.65), (3.66) and (3.74), it is easy to see that $\int \left(\frac{\partial^2 \varphi'(\zeta)}{\partial \theta^2} \right)^2 d\theta$ is uniformly bounded for t in $[\epsilon, T]$, for all $T > \epsilon > 0$. Since we can inject $H^2(\mathbb{T})$ into $C^{1+\alpha}(\mathbb{T})$ for some $\alpha > 0$, $\varphi'(\zeta(t, \cdot))$ lies in $C^{1+\alpha}(\mathbb{T})$ for all t in $[\epsilon, T]$. Since φ' is invertible and φ'' is positive, this implies that $\zeta(t, \cdot)$ also lies in $C^{1+\alpha}(\mathbb{T})$ for all t in $[\epsilon, T]$. Using this fact, we can rewrite the PDE as

$$\frac{\partial \zeta}{\partial t} = \varphi''(\zeta) \frac{\partial^2 \zeta}{\partial \theta^2} + \varphi^{(3)}(\zeta) \left(\frac{\partial \zeta}{\partial \theta} \right)^2.$$

Taking $a(t, \theta) = \varphi''(\zeta(t, \theta))$ and $b(t, \theta) = \varphi^{(3)}(\zeta) \left(\frac{\partial \zeta}{\partial \theta} \right)$, we get that ζ is a solution of the linear parabolic PDE

$$\frac{\partial \zeta}{\partial t} = a(t, \theta) \frac{\partial^2 \zeta}{\partial \theta^2} + b(t, \theta) \frac{\partial \zeta}{\partial \theta}$$

with coefficients a and b that belong to C^α . We can then use the theory for regularity of the solutions of linear parabolic equations (see for example [LSU68]) to show that $\zeta(t, \cdot)$ lies in $C^{2+\alpha}(\mathbb{T})$ for all t in $[\epsilon, T]$. The fact that $\frac{\partial \zeta}{\partial t}$ lies in C^α for all t in $[\epsilon, T]$ immediately follows from the PDE.

Chapitre 4

The two-scale approach for non-reversible dynamics

Les résultats de ce chapitre ont été obtenus en collaboration avec Manh Hong Duong, et ont été soumis pour publication dans l'article [DF14].

Abstract : In [GOVW09], a new method to study hydrodynamic limits was developed for reversible dynamics. In this work, we generalize this method to a family of non-reversible dynamics. As an application, we obtain quantitative rates of convergence to the hydrodynamic limit for a weakly asymmetric version of the Ginzburg-Landau model endowed with Kawasaki dynamics. These results also imply local Gibbs behavior, following the method developed in Chapter 2.

4.1 Introduction

In this work, we are interested in generalizing the results of [GOVW09] on hydrodynamic limits to the case of weakly asymmetric interacting spin systems. We obtain quantitative rates of convergence to the hydrodynamic limit for such dynamics. Our main contribution is a method of controlling the effects of the antisymmetric component of the dynamic.

A typical result of convergence to the hydrodynamic limit consists in proving that, under a suitable time-space scaling and for nice initial conditions, a random systems with a large number of particles behaves like a deterministic object, given as the solution of a partial differential equation.

In [GOVW09], a new method to study such problems was developed. It consists in establishing estimates in Wasserstein distance between the distribution of the system and a well-chosen macroscopic state, given as the solution of a differential equation. The main elements are a coarse-graining argument and a logarithmic Sobolev inequality. It was applied to dynamics of the form

$$dX_t = -A\nabla H(X_t)dt + \sqrt{2A}dW_t$$

on some Euclidean space, where A is a positive definite matrix, H is the Hamiltonian and W is a Wiener process. In the case where A and H correspond to the Ginzburg-Landau model endowed with Kawasaki dynamics, they obtained scaling limits of the form

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\rho).$$

In this work, we add an extra term to the previous dynamic, and study

$$dX_t = -A\nabla H(X_t)dt - J\nabla H(X_t)dt + \sqrt{2A}dW_t$$

where J is an antisymmetric matrix. This extra term makes the dynamic non-reversible, but does not modify the invariant measure. For a particular choice of J , we obtain a scaling limit of the form

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\rho) + \frac{\partial}{\partial \theta} \varphi'(\rho).$$

Our method is restricted to the case where the square of the antisymmetric part $-J^2$ is controlled by A (in the sense of symmetric matrices). This is because if the antisymmetric component becomes dominant in the scaling limit, we would expect the limiting PDE to be hyperbolic (rather than parabolic), and estimates in Wasserstein distances would not be adapted.

These estimates in Wasserstein distance also allow us to study local Gibbs behavior (which is stronger form of convergence) by using an interpolation inequality, following a method developed in [Fat13b]. We also obtain quantitative rates of convergence for the microscopic free energy to its scaling limit.

The plan of the paper is as follows : in Section 2, we present the framework and our main results. Section 3 contains the proofs of our results in the abstract setting. In section 4, we give the proofs of convergence to the hydrodynamic limit for the Ginzburg-Landau model endowed with a weakly asymmetric version of Kawasaki dynamics.

4.2 Framework and main results

4.2.1 Abstract setting

Let X, Y be two Euclidean spaces with $X \subset \mathbb{R}^N, Y \subset \mathbb{R}^M$. We think of X as the microscopic space and Y as the macroscopic space. N and M can then be thought of as the size of the microscopic and macroscopic data respectively. Let A and J be respectively a positive definite symmetric and an anti-symmetric linear operators on X . Let $H: X \rightarrow \mathbb{R}$ be a given function. We consider the stochastic dynamic on X that is given by the following stochastic differential equation

$$dX_t = -A\nabla H(X_t) dt - J\nabla H(X_t) dt + \sqrt{2A} dW_t, \quad (4.1)$$

where W_t is a Wiener process, and \sqrt{A} is the square root of the matrix A . When $J \neq 0$, this is a non-reversible process, and the Fokker-Planck equation associated to this SDE is

$$\partial_t(f\mu) = \div[\mu(A + J)\nabla f], \quad (4.2)$$

where μ is the invariant measure of the dynamic, which is

$$\mu(dx) := \frac{1}{Z} \exp(-H(x))dx.$$

In the application we have in mind, which we shall present in the next section, A will be the discrete Laplacian, and J the discrete derivation.

We now introduce an abstract framework for the notion of coarse-graining operator. Let $P: X \rightarrow Y$ be a linear operator such that

$$NPP^t = \text{id}_Y, \quad (4.3)$$

where P^t is the adjoint operator of P . We think of $y = Px$ as the macroscopic state associated to the microscopic state x . This operator induces a decomposition of the invariant measure into a macroscopic component and a fluctuation component. Let $\bar{\mu}(dy) = P_{\#}\mu$ be the push-forward of μ under the operator P and $\mu(dx|y)$ be the conditional measure of μ given $Px = y$, i.e., for each y , $\mu(dx|y)$ is a probability measure on X and satisfies that for any test function φ

$$\int_X \varphi(x) d\mu(x) = \int_Y \left(\int_{Px=y} \varphi(x) \mu(dx|y) \right) \bar{\mu}(dy). \quad (4.4)$$

Applying the technique in [GOVW09], we show that under certain conditions, the macroscopic profile $y = Px$, with law given by $\bar{f}(t, y) = \int_{Px=y} f(t, x) \mu(dx)$, is close to the solution of the following differential equation

$$\frac{d\eta}{dt} = -(\bar{A} + \bar{J})\nabla\bar{H}(\eta(t)). \quad (4.5)$$

In this equation, \bar{A} is a symmetric, positive definite operator and \bar{J} is another operator on Y defined by

$$\bar{A}^{-1} = PA^{-1}NP^t, \quad \bar{J} = \bar{A}PA^{-1}NJP^t, \quad (4.6)$$

and $\bar{H}: Y \rightarrow \mathbb{R}$ is the macroscopic Hamiltonian that satisfies

$$\bar{\mu}(dy) = \exp(-N\bar{H}(y))dy. \quad (4.7)$$

In order to state the assumptions, we need to recall the definition of the Logarithmic Sobolev inequality. A probability measure $\nu \in \mathcal{P}(X)$ is said to satisfy an LSI with constant $\rho > 0$ (abbreviation $\text{LSI}(\rho)$) if, for any locally Lipschitz, nonnegative function $f \in L^1(\nu)$,

$$\int \Phi(f) d\nu - \Phi\left(\int f d\nu\right) \leq \frac{1}{2\rho} \int \frac{|\nabla f|^2}{f} d\nu.$$

Assumptions : Throughout the paper, we assume that

- (i) $\kappa := \max_{x \in X} \{\langle \text{Hess } H(x) \cdot u, v \rangle, u \in \text{Ran}(NP^tP), v \in \text{Ran}(\text{id}_X - NP^tP), |u| = |v| = 1\} < \infty$;
- (ii) There is $\rho > 0$ such that $\mu(dx|y)$ satisfies $\text{LSI}(\rho)$ for all y ;
- (iii) There exist $\lambda, \Lambda > 0$ such that $\lambda \text{Id} \leq \text{Hess} \bar{H} \leq \Lambda \text{Id}$;
- (iv) There is $\alpha > 0$ such that $\int_X |x|^2 f \mu(dx) \leq \alpha N$;
- (v) There is $\beta > 0$ such that $\inf_{y \in Y} \bar{H}(y) \geq -\beta$;
- (vi) There is $\gamma > 0$ such that for all $x \in X$,

$$|(\text{id}_X - NP^tP)x|^2 \leq \gamma M^{-2} \langle x, Ax \rangle_X;$$

- (vii) There are constants C_1 and C_2 such that the initial datum satisfy

$$\int \Phi(f(0, x)) \mu(dx) \leq C_1 N \quad \text{and} \quad \bar{H}(\eta_0) \leq C_2;$$

- (viii) There is a $\tau > 0$ such that $A \geq \tau \text{Id}$;
- (ix) $-J^2 \leq cA$;
- (x) J and A commute.

Under these assumptions, we have the following bound on the Wasserstein distances between $f\mu$ and δ_η .

Theorem 4.1. *Let $\mu(dx) = \exp(-H(x)) dx$ be a probability measure on X , and let $P: X \rightarrow Y$ satisfy (4.3). Let $A: X \rightarrow X$ be a symmetric, definite positive operator, and $f(t, x)$ and $\eta(t)$ be the solutions of (4.2) and (4.5), with initial data $f(t, \cdot)$ and η_0 respectively. Suppose that the assumptions above hold. Define*

$$\Theta(t) := \frac{1}{2N} \int_X (x - NP^t \eta(t)) \cdot A^{-1}(x - NP^t \eta(t)) f(t, x) \mu(dx).$$

Then for any $T > 0$, we have

$$\max \left\{ \sup_{0 \leq t \leq T} \Theta(t), \frac{\lambda}{8} \int_0^T \left(\int_Y |y - \eta(t)|_Y^2 \bar{f}(t, y) \bar{\mu}(dy) \right) dt \right\} \leq e^{\frac{8c\Lambda^2}{\lambda} T} \left[\Theta(0) + E(T, M, N) \right],$$

where $E(T, M, N) \rightarrow 0$ as $N \uparrow \infty, M \uparrow \infty, \frac{N}{M} \uparrow \infty$. More precisely,

$$\begin{aligned} E(T, M, N) = & T \left(\frac{M}{N} \right) + \frac{4c\gamma\Lambda^2 T}{\lambda} \left(\alpha + \frac{2C_1}{\hat{\rho}} \right) \frac{1}{M} + C_1 \left(\frac{\gamma\kappa^2}{2\lambda\rho^2} + \frac{2c\gamma\kappa^2}{\tau\lambda\rho^2} + \frac{4\gamma c}{\lambda\tau} \right) \frac{1}{M^2} \\ & + \sqrt{2T\gamma} \left(\alpha + \frac{2C_1}{\hat{\rho}} \right)^{\frac{1}{2}} \left\{ \left(1 + \sqrt{\frac{c}{\tau}} + \frac{\sqrt{2c\gamma}}{M} \right) \sqrt{C_1} \right. \\ & \left. + \sqrt{2} \left(1 + \sqrt{\frac{c}{\tau}} \right) (H(\eta_0) - H(\eta_T)) + CT(1 + e^{CT} \bar{H}(\eta_0))^{\frac{1}{2}} \right\} \frac{1}{M}, \end{aligned}$$

where

$$\hat{\rho} := \frac{1}{2} \left(\rho + \lambda + \frac{\kappa^2}{\rho} - \sqrt{\left(\rho + \lambda + \frac{\kappa^2}{\rho} \right)^2 - 4\rho\lambda} \right).$$

Remark 4.2 (Remarks on the assumptions). Assumptions (i) to (viii) are collected from [GOVW09] and [Fat13b]. Assumption (ix) means that the asymmetric effect is controlled by the symmetric one. Its main use is to rule out situations where the scaling limit is a hyperbolic equation (this would be the case for a continuous analog of the asymmetric exclusion process), which the two-scale approach doesn't seem to handle. Assumption (x) is natural if we think of J and A are finite approximations of first and second derivatives operators, which is the application we have in mind. It could be replaced by an appropriate bound on the symmetric part of $PA^{-1}JNP^t$ (which is the macroscopic component of the commutator between A^{-1} and J), and an additional bound of the form $|\text{Tr}(PJA^{-1}NP^t)| \leq CM$. But since our proof is already fairly technical, and we do not have an application in mind that would warrant the greater generality, we decided to just assume that A and J commute, and simplify the proof. All these assumptions will be used in Lemma 3.4 to estimate the time derivative of $\Theta(t)$. In particular, (ii) and (vi) are used to handle the covariance and fluctuations terms respectively.

The hydrodynamic limit is obtained as a consequence.

Corollary 4.3. *Consider a sequence $\{X_\ell, Y_\ell, P_\ell, A_\ell, J_\ell, \mu_\ell, f_{0,\ell}, \eta_{0,\ell}\}_\ell$ satisfying the assumptions (i) to (x) with uniform constants $\kappa, \rho, \lambda, \Lambda, \alpha, \beta, \gamma, C_1, C_2$ and c . Suppose that*

$$N_\ell \xrightarrow{\ell \uparrow \infty} \infty; \quad M_\ell \xrightarrow{\ell \uparrow \infty} \infty; \quad \frac{M_\ell}{N_\ell} \xrightarrow{\ell \uparrow \infty} 0.$$

Further assume that

$$\lim_{\ell \uparrow \infty} \frac{1}{N_\ell} \int (x - N_\ell P^t \eta_{0,\ell}) \cdot A_\ell^{-1} (x - N_\ell P^t \eta_{0,\ell}) f_{0,\ell}(x) \mu_\ell(dx) = 0.$$

Then, for any $T > 0$:

(a) The microscopic variables are close to the solution of (4.5) in the penalized norm induced by A_ℓ^{-1} , uniformly in $t \in [0, T]$:

$$\lim_{\ell \uparrow \infty} \sup_{0 \leq t \leq T} \frac{1}{N_\ell} \int (x - N_\ell P^t \eta_\ell) \cdot A_\ell^{-1} (x - N_\ell P^t \eta_\ell) f_\ell(t, x) \mu_\ell(dx) = 0;$$

(b) The macroscopic variables are close to the solution of (4.5) in the strong $L^2(Y)$ norm, in a time-integrated sense :

$$\lim_{\ell \uparrow \infty} \int_0^T \int |y - \eta_\ell|_Y^2 \bar{f}(t, y) \bar{\mu}(dy) dt = 0.$$

Another topic of interest is whether the data behaves like a local Gibbs state.

Definition 4.4. The local Gibbs state with macroscopic profile $\eta \in Y$ is the probability measure on X whose density with respect to μ is given by

$$G(x) \mu(dx) := \frac{1}{Z} \exp(N P^t \nabla \bar{H}(\eta) \cdot x) \mu(dx).$$

Such a probability measure is close (in Wasserstein distance) to the associated macroscopic profile η .

In [Yau91], it is shown that, if the initial data is close (in the sense of relative entropy) to a local Gibbs state, then this also holds at any positive time, for a time-dependent local Gibbs state. Since closeness in relative entropy is stronger (in the current setting) than closeness in Wasserstein distance, the kind of results obtained with Yau's method are stronger than those of the previous Corollary, but require a stronger assumption on the initial data.

In [Kos01a], it was shown that convergence in relative entropy actually holds at positive times, even if the initial data converges only in a weaker sense. In [Fat13b], the second author obtained a new proof of this fact in the reversible setting, using the two-scale approach. This method also yields quantitative rates of convergence in relative entropy. Now that we have generalized the two-scale approach to the non-reversible setting, the extension of the results of [Fat13b] follows.

Theorem 4.5. Let $G(t, x)$ be the time-dependent local Gibbs state associated to the solution η of (4.5). Under our assumptions, the following holds :

(a) The relative entropy with respect to the local Gibbs state is controlled as follows :

$$\int_0^T \frac{1}{N} \int \Phi \left(\frac{f(t, x)}{G(t, x)} \right) G(t, x) \mu(dx) dt = O \left(\sqrt{\Theta(0) + \frac{M}{N} + \frac{1}{M}} \right) \quad (4.8)$$

where the actual constants in the bound (which can be made explicit) depend on $T, \lambda, \Lambda, \alpha, \gamma, \rho, \kappa, \tau, c, C_1$ and C_2 , but not on M and N ;

(b) The difference between the microscopic free energy and the free energy associated with the macroscopic profile η is bounded as follows :

$$\begin{aligned} & \int_0^T \left| \frac{1}{N} \int \Phi(f(t, x)) \mu(dx) - \bar{H}(\eta(t)) \right| dt \\ &= O \left(\sqrt{\Theta(0) + \frac{M}{N} + \frac{1}{M}} \right) \\ &+ O \left(\frac{M}{N} \right) \times \max \left(\left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right) \right|, \left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\lambda N} \right) \right| \right), \end{aligned} \quad (4.9)$$

where $\Gamma(Y, |\cdot|_Y)$ is the Gaussian integral on the space Y with respect to the norm $|\cdot|_Y$.

4.2.2 Application to spin systems

We now give an application of Theorem 4.1 to a system of interacting continuous spins. The application we have in mind is when the matrices A and J are given by

$$A = N^2 \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & -1 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ -1 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix} \quad (4.10)$$

and

$$J = \frac{N}{2} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & -1 \\ -1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & -1 & 0 \end{pmatrix} \quad (4.11)$$

As in [GOVW09], let

$$H(x) := \sum_{i=1}^N \psi(x_i) \quad (4.12)$$

where $\psi : \mathbb{R} \rightarrow \mathbb{R}$ satisfies the following assumptions :

$$\psi(x) = \frac{1}{2}x^2 + \delta\psi(x); \quad \|\delta\psi\|_{C^2} < \infty \quad (4.13)$$

We consider the dynamic where A and J are given by (4.10) and (4.11) respectively. This corresponds to the system of stochastic differential equations

$$dX_i(t) = -N^2(2\psi(X_i) - \psi(X_{i+1}) - \psi(X_{i-1}))dt - \frac{N}{2}(\psi(X_{i+1}) - \psi(X_{i-1}))dt + N\sqrt{2}(dB_t^{i+1} - dB_t^i).$$

This is the dynamic studied in [GPV88] and [GOVW09], to which we have added a weak asymmetric perturbation. This model is to the symmetric dynamic what the weakly asymmetric exclusion process is to the simple symmetric exclusion process, i.e., we have added an extra asymmetric term which has a scaling of lower order in N .

Since this dynamic conserves the mean spin $m = N^{-1} \sum X_i$, the natural space on which to work is

$$X_{N,m} := \left\{ x \in \mathbb{R}^N; \quad \frac{1}{N} \sum_{i=1}^N x_i = m \right\},$$

which we endow with the usual ℓ^2 scalar product. Following [GOVW09], the macroscopic space is

$$Y_{M,m} := \left\{ y \in \mathbb{R}^M; \quad \frac{1}{N} \sum_{i=1}^M y_i = m \right\},$$

which we endow with the L^2 scalar product

$$\langle y, \tilde{y} \rangle_Y := \frac{1}{M} \sum_{i=1}^M y_i \tilde{y}_i.$$

The coarse graining operator P is defined as

$$(Px)_i := \frac{1}{K} \sum_{j=(i-1)K+1}^{iK} x_j$$

where K is an integer such that $N = KM$. We can think of this coarse-graining operator as taking local averages of the microscopic profile over boxes of size K . This operator does satisfy the relation $PNP^t = \text{id}_Y$.

When K is large enough, it has been shown that the coarse-grained Hamiltonian \bar{H} is uniformly convex, so we will be able to apply the previous abstract Theorem.

Without loss of generality, we shall assume in the sequel that $m = 0$, since it does not play a role in our estimates.

To study the scaling limit, we need to embed our spaces $X_{N,m}$ into a single functional space. To a macroscopic profile $x \in X_{N,0}$, we associate the step function on the torus \bar{x} , defined by

$$\bar{x}(\theta) := x_i \quad \forall \theta \in \left[\frac{i-1}{N}, \frac{i}{N} \right).$$

We endow the space $L^2(\mathbb{T})$ with the H^{-1} norm, defined by

$$\|w\|_{H^{-1}}^2 = \int g^2 d\theta, \quad g' = w, \quad \int g d\theta = 0.$$

The closure of the spaces $X_{N,0}$ for this norm is the usual H^{-1} space of functions of average 0, which is the dual of the Sobolev space H^1 for the L^2 norm.

We can now state the hydrodynamic limit result we obtain for this model :

Theorem 4.6. *Let A_ℓ and J_ℓ be given by (4.10) and (4.11) respectively. Assume that ψ satisfies (4.13). Let $f(t, x)$ be a time-dependent probability density on $(X_{N,0}, \mu_{N,0})$ solving (4.2), with $f(0, \cdot) = f_0$ such that*

$$\int f_0 \log f_0 d\mu_{N,0} \leq CN$$

for some $C > 0$ and

$$\lim_{N \uparrow \infty} \int \|\bar{x} - \zeta_0\|_{H^{-1}}^2 f_0(x) \mu_{N,0}(dx) = 0$$

for some initial macroscopic profile $\zeta_0 \in L^2(\mathbb{T})$. Then, for any $T > 0$, we have

$$\lim_{N \uparrow \infty} \sup_{0 \leq t \leq T} \int \|\bar{x} - \zeta(t, \cdot)\|_{H^{-1}}^2 f(t, x) \mu_{N,0}(dx) = 0$$

where ζ is the unique solution of

$$\begin{cases} \frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) + \frac{\partial}{\partial \theta} \varphi'(\zeta), \\ \zeta(0, \cdot) = \zeta_0, \end{cases} \quad (4.14)$$

where φ is the Cramér transform of ψ , i.e.

$$\varphi(m) = \sup_{\sigma \in \mathbb{R}} \left\{ \sigma m - \log \int_{\mathbb{R}} \exp(\sigma x - \psi(x)) dx \right\}. \quad (4.15)$$

We can also use [Fat13b] to study local Gibbs behavior, and convergence of the relative entropy.

Theorem 4.7. *Under the same assumptions as in Theorem 4.6, the following holds :*

$$\int_0^T \int_{X_N} \Phi \left(\frac{f_N(t, x)}{G_N(t, x)} \right) G_N(t, x) \mu_N(dx) dt \longrightarrow 0, \quad (4.16)$$

where $G_N(t, \cdot)$ is the local Gibbs state given by $\eta_N(t)$. As a consequence, we have convergence of the microscopic entropy to the hydrodynamic entropy, in a time-integrated sense :

$$\int_0^T \left| \frac{1}{N} \int \Phi(f_N(t, x)) \mu_N(dx) - \left(\int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta - \varphi \left(\int_{\mathbb{T}} \zeta(t, \theta) d\theta \right) \right) \right| dt \xrightarrow{N \rightarrow \infty} 0. \quad (4.17)$$

Moreover, convergence of $\frac{1}{N} \int \Phi(f_N(t, x)) \mu_N(dx)$ to $\int_{\mathbb{T}} \varphi(\zeta(\theta, t)) d\theta - \varphi \left(\int_{\mathbb{T}} \zeta(t, \theta) d\theta \right)$ holds uniformly on any time-interval $[\epsilon, T]$, for any $0 < \epsilon < T$.

Since deducing this result from 4.5 is nearly the same as in [Fat13b], we omit the proof. The only significant difference is proving that the solution of the hydrodynamic equation ζ is smooth on $[\epsilon, T]$, which is a known result, that can be proven by a straightforward adaptation of the proof of Proposition 3.22 in [Fat13b].

4.3 Proof of the abstract results

In this section, we prove Theorem 4.1 and provide a sketch of proof of Theorem 4.5.

4.3.1 Proof of Theorem 4.1

Following the approach of [GOVW09], we prove Theorem 4.5 in three steps : first we differentiate with respect to time the Wasserstein distance between $f(t)\mu$ and the macroscopic profile $\eta(t)$, then we derive an upper bound for the quantity we obtain, before integrating in time and applying Gronwall's Lemma to obtain the result.

Lemma 4.8. *The time-derivative of $\Theta(t)$ is given by the following formula*

$$\begin{aligned}
& \frac{d}{dt} \frac{1}{2N} \int_X (x - NP^t \eta(t)) \cdot A^{-1}(x - NP^t \eta(t)) f(t, x) \mu(dx) \\
&= \frac{M}{N} - \int_Y (y - \eta) \cdot (\nabla_Y \bar{H}(y) - \nabla_Y \bar{H}(\eta)) \bar{f}(t, y) \bar{\mu} dy \\
&\quad - \int_Y PJA^{-1} NP^t (y - \eta) \cdot (\nabla_Y \bar{H}(y) - \nabla_Y \bar{H}(\eta)) \bar{f}(t, y) \bar{\mu} dy \\
&\quad - \int_Y (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu} dy \\
&\quad - \frac{1}{N} \int_X (\text{id}_X - NP^t P)x \cdot \nabla f(t, x) \mu(dx) \\
&\quad + \int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1} (\text{id}_X - NP^t P)x f \mu(dx) \\
&\quad + \int_Y PJA^{-1} NP^t (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{d}y \\
&\quad + \int_Y PJA^{-1} (\text{id}_X - NP^t P)x \cdot P \nabla f(t, x) \mu(dx) \\
&\quad + \frac{1}{N} \int_X (\text{id}_X - NP^t P)JA^{-1}(x - NP^t \eta) \cdot \nabla f(t, x) \mu(dx) \\
&\quad + \int_X PA^{-1} (\text{id}_X - NP^t P)x \cdot \bar{J} \nabla_Y \bar{H}(\eta) f(t, x) \mu(dx). \tag{4.18}
\end{aligned}$$

Démonstration. We have

$$\begin{aligned}
& \frac{d}{dt} \frac{1}{2N} \int_X (x - NP^t \eta(t)) \cdot A^{-1}(x - NP^t \eta(t)) f(t, x) \mu(dx) \\
&\stackrel{(4.2)}{=} -\frac{1}{N} \int_X A^{-1}(x - NP^t \eta) \cdot (A + J) \nabla f \mu(dx) - \int P^t \frac{d\eta}{dt} \cdot A^{-1}(x - NP^t \eta) f \mu(dx) \\
&\stackrel{(4.5)}{=} -\frac{1}{N} \int_X A^{-1}(x - NP^t \eta) \cdot A \nabla f \mu(dx) + \int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1}(x - NP^t \eta) f \mu(dx) \\
&\quad - \frac{1}{N} \int A^{-1}(x - NP^t \eta) \cdot J \nabla f \mu(dx) + \int A^{-1}(x - NP^t \eta) \cdot P^t \bar{J} \nabla \bar{H}(\eta) f \mu(dx) \\
&= (I) + (II) + (III) + (IV). \tag{4.19}
\end{aligned}$$

We now use the decomposition $x = NP^t Px + (\text{id}_X - NP^t P)x$ to transform each term on the right hand side of (4.19). We need the following definition of the μ -covariance of two functions $f, g \in L^2(\mu)$

$$\text{cov}_{\mu}(f, g) = \int fg d\mu - \left(\int f d\mu \right) \left(\int g d\mu \right). \tag{4.20}$$

The first two terms, (I) and (II), are already done in [GOVW09]. We repeat here for the sake of completeness.

$$\begin{aligned}
(I) &= -\frac{1}{N} \int_X (x - NP^t \eta) \cdot \nabla f \mu(dx) \\
&= -\int_X P^t (Px - \eta) \cdot \nabla f \mu(dx) - \frac{1}{N} \int (\text{id}_X - NP^t P)x \cdot \nabla f \mu(dx). \tag{4.21}
\end{aligned}$$

We now transform the first term in (4.21) using (4.4) and Lemma 21 in [GOVW09].

$$\begin{aligned}
& - \int_X P^t(Px - \eta) \cdot \nabla f \mu(dx) = - \int (Px - \eta) \cdot P \nabla f \mu(dx) \\
& \stackrel{(4.4)}{=} - \int_Y (y - \eta) \cdot P \int_{Px=y} \nabla f \mu(dx|y) \bar{\mu} dy \\
& \stackrel{[\text{GOVW09},(36)]}{=} - \frac{1}{N} \int (y - \eta) \cdot \nabla_Y \bar{f} \bar{\mu} dy - \int (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu} dy \\
& \stackrel{(4.7)}{=} \frac{1}{N} \int \nabla_Y \cdot y \bar{f} \bar{\mu} dy - \int (y - \eta) \cdot \nabla_Y \bar{H}(y) \bar{f} \bar{\mu} dy - \int (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu} dy \\
& = \frac{\dim Y}{N} - \int (y - \eta) \cdot \nabla_Y \bar{H}(y) \bar{f} \bar{\mu} dy - \int (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu} dy.
\end{aligned}$$

We obtain

$$(I) = \frac{\dim Y}{N} - \int (y - \eta) \cdot \nabla_Y \bar{H}(y) \bar{f} \bar{\mu} dy - \int (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu} dy - \frac{1}{N} \int (\text{id}_X - NP^t P)x \cdot \nabla f \mu(dx). \quad (4.22)$$

Now we proceed with (II).

$$\begin{aligned}
(II) &= \int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1} NP^t (Px - \eta) f \mu(dx) + \int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1} (\text{id}_X - NP^t P)x f \mu(dx) \\
&\stackrel{(4.6)}{=} \int \nabla_Y \bar{H}(\eta) \cdot (Px - \eta) f \mu(dx) + \int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1} (\text{id}_X - NP^t P)x f \mu(dx) \\
&\stackrel{(4.4)}{=} \int_Y (y - \eta) \cdot \nabla_Y \bar{H}(\eta) \bar{f} \bar{\mu}(dy) + \int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1} (\text{id}_X - NP^t P)x f \mu(dx). \quad (4.23)
\end{aligned}$$

Next, we continue with (III).

$$\begin{aligned}
(III) &= \frac{1}{N} \int JA^{-1}(x - NP^t \eta) \cdot \nabla f \mu(dx) \\
&= \frac{1}{N} \int PJA^{-1}(x - NP^t \eta) \cdot NP \nabla f \mu(dx) + \frac{1}{N} \int (\text{id}_X - NP^t P)JA^{-1}(x - NP^t \eta) \cdot \nabla f \mu(dx) \\
&= \frac{1}{N} \int PJA^{-1} NP^t (Px - \eta) \cdot NP \nabla f \mu(dx) + \frac{1}{N} \int PJA^{-1} (\text{id}_X - NP^t P)x \cdot NP \nabla f \mu(dx) \\
&\quad + \frac{1}{N} \int (\text{id}_X - NP^t P)JA^{-1}(x - NP^t \eta) \cdot \nabla f \mu(dx).
\end{aligned}$$

The first term on the right hand side of the expression above can be transformed further using Lemma 21 in [GOVW09] as done for (I).

$$\begin{aligned}
\frac{1}{N} \int PJA^{-1}(x - NP^t \eta) \cdot NP \nabla f \mu(dx) &= \frac{1}{N} \int_Y \left(\int_{Px=y} PJA^{-1} NP^t (y - \eta) \cdot NP \nabla f \mu(dx|y) \right) \bar{\mu} dy \\
&= \frac{1}{N} \int_Y PJA^{-1} NP^t (y - \eta) \cdot \left[\nabla_Y \bar{f}(y) + NP \text{cov}_{\mu(dx|y)}(f, \nabla H) \right] \bar{\mu}(dy) \\
&= \frac{1}{N} \int_Y PJA^{-1} NP^t (y - \eta) \cdot \nabla_Y \bar{f}(y) \bar{\mu}(dy) + \int_Y PJA^{-1} NP^t (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu}(dy) \\
&= - \frac{\text{Tr}(PJA^{-1} NP^t)}{N} + \int_Y PJA^{-1} NP^t (y - \eta) \cdot \nabla_Y \bar{H}(y) \bar{f} \bar{\mu}(dy) \\
&\quad + \int_Y PJA^{-1} NP^t (y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu}(dy).
\end{aligned}$$

Since $PJA^{-1}NP^t$ is anti-symmetric, $\text{Tr}(PJA^{-1}NP^t)=0$, and we obtain

$$\begin{aligned}
(III) &= \int_Y PJA^{-1}NP^t(y - \eta) \cdot \nabla_Y \bar{H}(y) \bar{f} \bar{\mu}(dy) \\
&\quad + \int_Y PJA^{-1}NP^t(y - \eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu}(dy) \\
&\quad + \frac{1}{N} \int PJA^{-1}(\text{id}_X - NP^t P)x \cdot NP \nabla f \mu(dx) \\
&\quad + \frac{1}{N} \int (\text{id}_X - NP^t P)JA^{-1}(x - NP^t \eta) \cdot \nabla f \mu(dx). \tag{4.24}
\end{aligned}$$

Finally, we now transform (IV).

$$\begin{aligned}
(IV) &= \int PA^{-1}NP^t(Px - \eta) \cdot \bar{J} \nabla_Y \bar{H}(\eta) f \mu(dx) + \int PA^{-1}(\text{id}_X - NP^t P) \cdot \bar{J} \nabla_Y \bar{H}(\eta) f \mu(dx) \\
&\stackrel{(4.4)}{=} \int_Y PA^{-1}NP^t(y - \eta) \cdot \bar{J} \nabla_Y \bar{H}(\eta) \bar{f} \bar{\mu}(dy) + \int PA^{-1}(\text{id}_X - NP^t P) \cdot \bar{J} \nabla_Y \bar{H}(\eta) f \mu(dx) \\
&\stackrel{(4.6)}{=} - \int PJA^{-1}NP^t(y - \eta) \cdot \nabla_Y \bar{H}(\eta) \bar{f} \bar{\mu}(dy) + \int PA^{-1}(\text{id}_X - NP^t P) \cdot \bar{J} \nabla_Y \bar{H}(\eta) f \mu(dx). \tag{4.25}
\end{aligned}$$

Substituting (4.22)-(4.25) into (4.19), we obtain (4.18) and the lemma is proven. \square

The following auxiliary lemma will be helpful in the sequel. The second and the third parts are respectively (54) and (52) in [GOVW09]; we put them here for readers' convenience.

Lemma 4.9. *We have the following estimate*

1. *For every $y \in Y$*

$$|PJA^{-1}NP^t y|^2 \leq c \langle \bar{A}^{-1} y, y \rangle \leq \frac{c}{\tau} |y|_Y^2, \tag{4.26}$$

$$\langle \bar{A} PJA^{-1}NP^t y, PJA^{-1}NP^t y \rangle \leq c |y|^2. \tag{4.27}$$

2. *For every $x \in X$*

$$(\text{id}_X - NP^t P)x \cdot A^{-1}(\text{id}_X - NP^t P)x \leq \frac{\gamma}{M^2} |x|^2. \tag{4.28}$$

3. *It holds that*

$$|NP^t P \text{cov}_{\mu(dx|y)}(f, \nabla H)|^2 \leq \gamma \frac{\kappa^2}{\rho^2} \frac{1}{M^2} \bar{f} \int \frac{1}{f} \nabla f \cdot A \nabla f \mu(dx|y). \tag{4.29}$$

Démonstration. We only need to prove the first part.

We start with (4.26). The first inequality is obtained using the assumption (2) and the fact that $NP^t P$ is an orthogonal projection as follows.

$$\begin{aligned}
\langle PJA^{-1}NP^t y, PJA^{-1}NP^t y \rangle &= \frac{1}{N} \langle NP^t PJA^{-1}NP^t y, JA^{-1}NP^t y \rangle \\
&\leq \frac{1}{N} \langle JA^{-1}NP^t y, JA^{-1}NP^t y \rangle \\
&= -\frac{1}{N} \langle J^2 A^{-1}NP^t y, A^{-1}NP^t y \rangle \\
&\leq \frac{c}{N} \langle A^{-1}NP^t y, NP^t y \rangle \quad (\text{used assumption (ix) here}) \\
&= c \langle NPA^{-1}P^t y, y \rangle \\
&= c \langle \bar{A}^{-1} y, y \rangle.
\end{aligned}$$

Now we prove the second one. Since τ is a lower bound on the spectral value of A , $\frac{1}{\tau}$ is an upper bound on that of A^{-1} . Hence

$$\langle \bar{A}^{-1}y, y \rangle = \langle PA^{-1}NP^t y, y \rangle = \frac{1}{N} \langle A^{-1}NP^t y, NP^t y \rangle \leq \frac{1}{N\tau} \langle NP^t y, NP^t y \rangle = \frac{1}{\tau} |y|_Y^2.$$

Next, we prove (4.27). By duality, we have

$$\begin{aligned} \langle \bar{A}PJA^{-1}NP^t y, PJA^{-1}NP^t y \rangle &= \sup_z \{2\langle PJA^{-1}NP^t y, z \rangle - \langle \bar{A}^{-1}z, z \rangle\} \\ &\stackrel{(4.26)}{\leq} \sup_z \{2\langle y, PJA^{-1}NP^t z \rangle - c^{-1}|PJA^{-1}NP^t z|^2\} \\ &\leq \sup_z \{2\langle y, z \rangle - c^{-1}|z|^2\} \\ &\leq c|y|^2. \end{aligned}$$

□

Lemma 4.10. *If $f(t, x)$ and $\eta(t)$ satisfy the assumptions of theorem 4.1, then for any $T < \infty$ we have*

$$\int_0^T \int \frac{1}{f} \nabla f \cdot A \nabla f(t, x) \mu(dx) dt = \int \Phi(f(0, x)) \mu(dx) - \int \Phi(f(T, x)) \mu(dx); \quad (4.30)$$

$$\int_0^T \langle \bar{A} \nabla_Y \bar{H}(\eta), \nabla_Y \bar{H}(\eta) \rangle dt \leq 2(H(\eta_0) - H(\eta_T)) + CT(1 + e^{CT} \bar{H}(\eta_0)), \quad (4.31)$$

where $C > 0$ is a constant;

$$\left(\int |x|^2 f(t, x) \mu(dx) \right)^{\frac{1}{2}} \leq \left(\int |x|^2 \mu(dx) \right)^{\frac{1}{2}} + \left(\frac{2}{\hat{\rho}} \int \Phi(f(0, x)) \mu(dx) \right)^{\frac{1}{2}}. \quad (4.32)$$

Démonstration. The proof of this lemma is similar to that of proposition 24 in [GOVW09]. We prove (4.30) first. We have

$$\begin{aligned} \frac{d}{dt} \int \Phi(f(t, x)) \mu(dx) &= \int (\log f + 1) \partial_t (f \mu) \\ &= \int (\log f + 1) \operatorname{div}(\mu(A + J) \nabla f) \\ &= - \int (A + J) \nabla f \cdot \frac{\nabla f}{f} \mu(dx) \\ &= - \int \frac{1}{f} A \nabla f \cdot \nabla f \mu(dx) \quad (\text{since } J \text{ is anti-symmetric}). \end{aligned} \quad (4.33)$$

Thus (4.30) follows. Next we prove (4.31). We have

$$\begin{aligned} \frac{d}{dt} \bar{H}(\eta(t)) &= \langle \dot{\eta}(t), \nabla_Y \bar{H}(\eta) \rangle \\ &\stackrel{(4.5)}{=} - \langle \bar{A} \nabla_Y \bar{H}(\eta), \nabla_Y \bar{H}(\eta) \rangle - \langle \bar{J} \nabla_Y \bar{H}(\eta), \nabla_Y \bar{H}(\eta) \rangle \\ &= - \langle \bar{A} \nabla \bar{H}(\eta), \nabla \bar{H}(\eta) \rangle - \langle \bar{A} P J A^{-1} N P^t \nabla \bar{H}(\eta), \nabla \bar{H}(\eta) \rangle \\ &\leq \frac{1}{2} \langle \bar{A} P J A^{-1} N P^t \nabla \bar{H}(\eta), P J A^{-1} N P^t \nabla \bar{H}(\eta) \rangle - \frac{1}{2} \langle \bar{A} \nabla_Y \bar{H}(\eta), \nabla_Y \bar{H}(\eta) \rangle \\ &\stackrel{(4.27)}{\leq} \frac{c}{2} |\nabla \bar{H}(\eta)|^2 - \frac{1}{2} \langle \bar{A} \nabla_Y \bar{H}(\eta), \nabla_Y \bar{H}(\eta) \rangle \end{aligned}$$

and therefore

$$\begin{aligned} \frac{d}{dt}\bar{H}(\eta(t)) + \frac{1}{2}\langle \bar{A}\nabla_Y\bar{H}(\eta), \nabla_Y\bar{H}(\eta) \rangle \\ \leq \frac{c}{2}|\nabla\bar{H}(\eta)|^2 \\ \leq C(|\eta|^2 + 1) \\ \leq C(\bar{H}(\eta) + 1). \end{aligned}$$

In the above estimate, $C > 0$ is a general constant. Note that we have used the assumption (iii). The above Gronwall-type inequality implies that for every $t \geq 0$, we have $\bar{H}(\eta(t)) \leq e^{C(T+1)}\bar{H}(\eta_0)$, and

$$\int_0^T \langle \bar{A}\nabla_Y\bar{H}(\eta), \nabla_Y\bar{H}(\eta) \rangle dt \leq 2(H(\eta_0) - H(\eta_T)) + CT(1 + e^{CT}\bar{H}(\eta_0)).$$

By (4.33), $\int \Phi(f(t, x))\mu(dx)$ is non-increasing in t ; hence the proof of (4.32) is the same as that of (46) in [GOVW09]. \square

Lemma 4.11. *We have the following estimate*

$$\begin{aligned} \frac{d}{dt}\Theta(t) - \frac{8c\Lambda^2}{\lambda}\Theta(t) + \frac{\lambda}{8}\int |y - \eta|^2 \bar{f}\bar{\mu}(dy) \\ \leq \frac{M}{N} + \frac{4c\gamma\Lambda^2}{2\lambda NM^2}\int |x|^2 f\mu(dx) \\ + \left(\frac{\gamma\kappa^2}{2\lambda\rho^2 M^2} + \frac{2c\gamma\kappa^2}{\tau\lambda\rho^2 M^2} + \frac{4\gamma c}{\lambda\tau M^2} \right) \int \frac{1}{Nf}\nabla f \cdot A\nabla f\mu(dx) \\ + \frac{\sqrt{\gamma}}{M} \left(\int \frac{1}{N}|x|^2 f\mu(dx) \right)^{\frac{1}{2}} \left[\left(1 + \sqrt{\frac{c}{\tau}} + \frac{\sqrt{2c\gamma}}{M} \right) \left(\int \frac{1}{Nf}\nabla f \cdot A\nabla f\mu(dx) \right)^{\frac{1}{2}} \right. \end{aligned} \quad (4.34)$$

$$\left. + \left(1 + \sqrt{\frac{c}{\tau}} \right) \left(\bar{A}\nabla_Y\bar{H}(\eta) \cdot \nabla_Y\bar{H}(\eta) \right)^{\frac{1}{2}} \right]. \quad (4.35)$$

Démonstration. We estimate each term in (4.18). The 2nd, 4th and 5th terms are already done in [GOVW09]. We get

$$-\int_Y (y - \eta) \cdot (\nabla_Y\bar{H}(y) - \nabla_Y\bar{H}(\eta))\bar{f}\bar{\mu}dy \leq -\lambda\int |y - \eta|_Y^2 \bar{f}\bar{\mu}dy, \quad (4.36)$$

$$\left| \int (y - \eta) \cdot P\text{cov}_{\mu(dx|y)}(f, \nabla H)\bar{\mu}dy \right| \leq \frac{\gamma\kappa^2}{2\lambda\rho^2 M^2} \int \frac{1}{Nf}\nabla f \cdot A\nabla f\mu(dx) + \frac{\lambda}{2}\int |y - \eta|_Y^2 \bar{f}\bar{\mu}(dy), \quad (4.37)$$

$$\left| \frac{1}{N} \int (\text{id}_X - NP^tP)x \cdot \nabla f\mu(dx) \right| \leq \left(\frac{\gamma}{M^2} \int \frac{1}{Nf}\nabla f \cdot A\nabla f\mu(dx) \cdot \int \frac{1}{N}|x|^2 f\mu(dx) \right)^{\frac{1}{2}}. \quad (4.38)$$

We estimate the 3rd term. Since

$$\begin{aligned} |PJA^{-1}NP^t(y - \eta)| \cdot |\nabla_Y\bar{H}(y) - \nabla_Y\bar{H}(\eta)| &\leq \Lambda|y - \eta| \cdot |PJA^{-1}NP^t(y - \eta)| \\ &\stackrel{(4.26)}{\leq} \Lambda|y - \eta|\sqrt{c\langle \bar{A}^{-1}(y - \eta), y - \eta \rangle} \\ &\leq \frac{\lambda}{8}|y - \eta|^2 + \frac{2c\Lambda^2}{\lambda}\langle \bar{A}^{-1}(y - \eta), y - \eta \rangle, \end{aligned}$$

we have

$$\begin{aligned}
& \left| \int_Y PJA^{-1}NP^t(y-\eta) \cdot (\nabla_Y \bar{H}(y) - \nabla_Y \bar{H}(\eta)) \bar{f} \bar{\mu}(dy) \right| \\
& \leq \int_Y |PJA^{-1}NP^t(y-\eta)| |\nabla_Y \bar{H}(y) - \nabla_Y \bar{H}(\eta)| \bar{f} \bar{\mu}(dy) \\
& \leq \frac{\lambda}{8} \int_Y |y-\eta|^2 \bar{f} \bar{\mu}(dy) + \frac{2c\Lambda^2}{\lambda} \int_Y \langle \bar{A}^{-1}(y-\eta), y-\eta \rangle \bar{f} \bar{\mu}(dy) \\
& = \frac{\lambda}{8} \int_Y |y-\eta|^2 \bar{f} \bar{\mu}(dy) + \frac{2c\Lambda^2}{\lambda} \frac{1}{N} \int_X \langle A^{-1}NP^t(Px-\eta), NP^t(Px-\eta) \rangle f \mu(dx) \\
& \leq \frac{\lambda}{8} \int_Y |y-\eta|^2 \bar{f} \bar{\mu}(dy) + \frac{2c\Lambda^2}{\lambda} \frac{2}{N} \int_X \langle A^{-1}(x-NP^t\eta), (x-NP^t\eta) \rangle f \mu(dx) \\
& \quad + \frac{2c\Lambda^2}{\lambda} \frac{2}{N} \int_X \langle A^{-1}(\text{id}_X - NP^tP)x, (\text{id}_X - NP^tP)x \rangle f \mu(dx) \\
& \leq \frac{\lambda}{8} \int_Y |y-\eta|^2 \bar{f} \bar{\mu}(dy) + \frac{8c\Lambda^2}{\lambda} \Theta(t) + \frac{4c\gamma\Lambda^2}{\lambda NM^2} \int |x|^2 f \mu(dx). \tag{4.39}
\end{aligned}$$

Next we estimate the 6th term.

$$\begin{aligned}
& \int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1}(\text{id}_X - NP^tP)x f \mu(dx) \\
& = \int P^t \bar{A} \nabla_Y \bar{H}(\eta) \cdot A^{-1}(\text{id}_X - NP^tP)x f \mu(dx) \\
& \leq \left(\int P^t \bar{A} \nabla_Y \bar{H}(\eta) \cdot A^{-1}NP^t \bar{A} \nabla_Y \bar{H}(\eta) f \mu(dx) \right)^{\frac{1}{2}} \left(\frac{1}{N} \int (\text{id}_X - NP^tP)x \cdot A^{-1}(\text{id}_X - NP^tP)x f \mu(dx) \right)^{\frac{1}{2}}
\end{aligned}$$

Since

$$P^t \bar{A} \nabla_Y \bar{H}(\eta) \cdot A^{-1}NP^t \bar{A} \nabla_Y \bar{H}(\eta) = \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1}NP^t \bar{A} \nabla_Y \bar{H}(\eta) = \bar{A} \nabla_Y \bar{H}(\eta) \cdot \nabla_Y \bar{H},$$

and from (4.28), we have

$$\int \bar{A} \nabla_Y \bar{H}(\eta) \cdot PA^{-1}(\text{id}_X - NP^tP)x f \mu(dx) \leq \left(\bar{A} \nabla_Y \bar{H}(\eta) \cdot \nabla_Y \bar{H}(\eta) \right)^{\frac{1}{2}} \left(\frac{\gamma}{NM^2} \int |x|^2 f \mu(dx) \right)^{\frac{1}{2}}. \tag{4.40}$$

Next, we estimate the 7th term.

$$\begin{aligned}
& \left| \int PJA^{-1}NP^t(y-\eta) \cdot P \text{cov}_{\mu(dx|y)}(f, \nabla H) \bar{\mu}(dy) \right| \\
& \leq \left(\int |PJA^{-1}NP^t(y-\eta)|^2 \bar{f} \bar{\mu}(dy) \cdot \int \frac{1}{\bar{f}} |P \text{cov}_{\mu(dx|y)}|_Y^2 \bar{\mu}(dy) \right)^{\frac{1}{2}} \\
& \stackrel{(4.26), (4.29)}{\leq} \left(\frac{2c}{\tau} \gamma \frac{\kappa^2}{\rho^2} \frac{1}{M^2} \int |y-\eta|^2 \bar{f} \bar{\mu}(dy) \int \frac{1}{Nf} \nabla f \cdot A \nabla f \mu(dx) \right)^{\frac{1}{2}} \\
& \leq \frac{2c\gamma\kappa^2}{\tau\lambda\rho^2 M^2} \int \frac{1}{Nf} \nabla f \cdot A \nabla f \mu(dx) + \frac{\lambda}{8} \int |y-\eta|^2 \bar{f} \bar{\mu}(dy). \tag{4.41}
\end{aligned}$$

For the 8th term, we have

$$\begin{aligned}
& \int PJA^{-1}(\text{id}_X - NP^tP)x \cdot P \nabla f \mu(dx) = \frac{1}{N} \int NP^t PJA^{-1}(\text{id}_X - NP^tP)x \nabla f \mu(dx) \\
& \leq \left(\int \frac{1}{N} NP^t PJA^{-1}(\text{id}_X - NP^tP)x \cdot A^{-1}NP^t PJA^{-1}(\text{id}_X - NP^tP)x f \mu(dx) \right)^{\frac{1}{2}} \left(\int \frac{1}{Nf} \nabla f \cdot A \nabla f \mu(dx) \right)^{\frac{1}{2}}
\end{aligned}$$

Since

$$\begin{aligned}
& \langle NP^t PJA^{-1}(\text{id}_X - NP^t P)x, A^{-1}NP^t PJA^{-1}(\text{id}_X - NP^t P)x \rangle \\
& \leq \frac{1}{\tau} \langle NP^t PJA^{-1}(\text{id}_X - NP^t P)x, NP^t PJA^{-1}(\text{id}_X - NP^t P)x \rangle \\
& \leq \frac{1}{\tau} \langle JA^{-1}(\text{id}_X - NP^t P)x, JA^{-1}(\text{id}_X - NP^t P)x \rangle \\
& = \frac{1}{\tau} \langle -J^2 A^{-1}(\text{id}_X - NP^t P)x, A^{-1}(\text{id}_X - NP^t P)x \rangle \\
& \leq \frac{c}{\tau} \langle (\text{id}_X - NP^t P)x, A^{-1}(\text{id}_X - NP^t P)x \rangle \\
& \stackrel{(4.28)}{\leq} \frac{c\gamma}{\tau M^2} |x|^2,
\end{aligned}$$

we obtain

$$\int PJA^{-1}(\text{id}_X - NP^t P)x \cdot P\nabla f \mu(dx) \leq \left(\frac{c\gamma}{\tau M^2} \int \frac{1}{N} |x|^2 f \mu(dx) \cdot \int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) \right)^{\frac{1}{2}}. \quad (4.42)$$

Next we estimate the 9th term. Set $z = JA^{-1}(x - NP^t \eta)$. We have

$$\begin{aligned}
& \frac{1}{N} \int (\text{id}_X - NP^t P)JA^{-1}(x - NP^t \eta) \cdot \nabla f \mu(dx) = \frac{1}{N} \int (\text{id}_X - NP^t P)z \cdot \nabla f \mu(dx) \\
& \leq \left(\int \frac{1}{N} (\text{id}_X - NP^t P)z \cdot A^{-1}(\text{id}_X - NP^t P)z f \mu(dx) \int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) \right)^{\frac{1}{2}} \\
& \stackrel{(4.28)}{\leq} \left(\frac{\gamma}{M^2} \int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) \int \frac{1}{N} |z|^2 f \mu(dx) \right)^{\frac{1}{2}}.
\end{aligned}$$

We estimate the second integral inside the parentheses. It holds that

$$\begin{aligned}
|z|^2 & = \langle JA^{-1}(x - NP^t \eta), JA^{-1}(x - NP^t \eta) \rangle \\
& = \langle -J^2 A^{-1}(x - NP^t \eta), A^{-1}(x - NP^t \eta) \rangle \\
& \stackrel{(ix)}{\leq} c \langle A^{-1}(x - NP^t \eta), x - NP^t \eta \rangle \\
& \leq 2c \left(\langle A^{-1}NP^t(Px - \eta), NP^t(Px - \eta) \rangle + \langle A^{-1}(\text{id}_X - NP^t P)x, (\text{id}_X - NP^t P)x \rangle \right) \\
& \stackrel{(4.28)}{\leq} 2c \left(\frac{1}{\tau} |NP^t(Px - \eta)|^2 + \frac{\gamma}{M^2} |x|^2 \right) \\
& = 2c \left(\frac{N}{\tau} |Px - \eta|^2 + \frac{\gamma}{M^2} |x|^2 \right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
& \frac{1}{N} \int (\text{id}_X - NP^t P)JA^{-1}(x - NP^t \eta) \cdot \nabla f \mu(dx) \\
& \leq \left(\frac{\gamma}{M^2} \int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) \right)^{\frac{1}{2}} \left(\frac{2c}{\tau} \int |y - \eta|^2 \bar{f} \bar{\mu}(dy) + \frac{2c\gamma}{M^2 N} \int |x|^2 f \mu(dx) \right)^{\frac{1}{2}} \\
& \leq \frac{4\gamma c}{M^2 \lambda \tau} \int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) + \frac{\lambda}{8} \int |y - \eta|^2 \bar{f} \bar{\mu}(dy) \\
& \quad + \left(\frac{\gamma}{M^2} \int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) \right)^{\frac{1}{2}} \left(\frac{2c\gamma}{M^2 N} \int |x|^2 f \mu(dx) \right)^{\frac{1}{2}}. \quad (4.43)
\end{aligned}$$

Finally, we estimate the 10th term. Since

$$\begin{aligned} & \langle PA^{-1}NP^t\bar{J}\nabla\bar{H}(\eta), \bar{J}\nabla\bar{H}(\eta) \rangle \\ &= \langle PJA^{-1}NP^t\nabla\bar{H}(\eta), \bar{A}PJA^{-1}NP^t\nabla\bar{H}(\eta) \rangle \\ &\leq c|\nabla\bar{H}(\eta)|^2 \\ &\leq \frac{c}{\tau}\langle \bar{A}\nabla_Y\bar{H}(\eta), \nabla_Y\bar{H}(\eta) \rangle, \end{aligned}$$

we have

$$\begin{aligned} & \left| \int A^{-1}(\text{id}_X - NP^tP)x \cdot P^t\bar{J}\nabla_Y\bar{H}(\eta) f\mu(dx) \right| \\ &\leq \left(\int P^t\bar{J}\nabla_Y\bar{H}(\eta) \cdot A^{-1}NP^t\bar{J}\nabla_Y\bar{H}(\eta) f\mu(dx) \right)^{\frac{1}{2}} \left(\int \frac{1}{N}(\text{id}_X - NP^tP)x \cdot A^{-1}(\text{id}_X - NP^tP)x f\mu(dx) \right)^{\frac{1}{2}} \\ &\leq \left(\frac{c}{\tau}\bar{A}\nabla_Y\bar{H}(\eta) \cdot \nabla_Y\bar{H}(\eta) \right)^{\frac{1}{2}} \left(\frac{\gamma}{NM^2} \int |x|^2 f\mu(dx) \right)^{\frac{1}{2}}. \end{aligned} \quad (4.44)$$

Summing up from (4.36) to (4.44), we obtain

$$\begin{aligned} & \frac{d}{dt}\Theta(t) - \frac{8c\Lambda^2}{\lambda}\Theta(t) + \frac{\lambda}{8} \int |y - \eta|^2 \bar{f}\bar{\mu}(dy) \\ &\leq \frac{M}{N} + \frac{4c\gamma\Lambda^2}{2\lambda NM^2} \int |x|^2 f\mu(dx) \\ &\quad + \left(\frac{\gamma\kappa^2}{2\lambda\rho^2 M^2} + \frac{2c\gamma\kappa^2}{\tau\lambda\rho^2 M^2} + \frac{4\gamma c}{\lambda\tau M^2} \right) \int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) \\ &\quad + \frac{\sqrt{\gamma}}{M} \left(\int \frac{1}{N} |x|^2 f\mu(dx) \right)^{\frac{1}{2}} \left[\left(1 + \sqrt{\frac{c}{\tau}} + \frac{\sqrt{2c\gamma}}{M} \right) \left(\int \frac{1}{Nf} \nabla f \cdot A\nabla f \mu(dx) \right)^{\frac{1}{2}} \right. \\ &\quad \left. \left(1 + \sqrt{\frac{c}{\tau}} \right) \left(\bar{A}\nabla_Y\bar{H}(\eta) \cdot \nabla_Y\bar{H}(\eta) \right)^{\frac{1}{2}} \right]. \end{aligned}$$

□

Proof of Theorem 4.1. Denote by $R(t)$ the right hand side of (4.35). Set $D = \frac{8c\Lambda^2}{\lambda}$. For any $0 < t \leq T$, we have

$$\begin{aligned} \frac{d}{dt} \left(e^{-Dt}\Theta(t) \right) + e^{-Dt} \frac{\lambda}{8} \int |y - \eta|^2 \bar{f}\bar{\mu}(dy) &\leq \frac{d}{dt} \left(e^{-Dt}\Theta(t) \right) + e^{-Dt} \frac{\lambda}{8} \int |y - \eta|^2 \bar{f}\bar{\mu}(dy) \\ &\leq e^{-Dt}S(t) \leq S(t). \end{aligned} \quad (4.45)$$

Integrating (4.35) with respect to time, for any $0 < t \leq T$, we have

$$\begin{aligned} e^{-DT}\Theta(t) + \frac{\lambda}{8} e^{-DT} \int_0^T \int_Y |y - \eta|^2 \bar{f}\bar{\mu}(dy) &\leq e^{-Dt}\Theta(t) + \frac{\lambda}{8} e^{-DT} \int_0^T \int_Y |y - \eta|^2 \bar{f}\bar{\mu}(dy) \\ &\leq \Theta(0) + \int_0^T S(t)dt. \end{aligned} \quad (4.46)$$

It follows that for any $T > 0$

$$\max \left\{ \sup_{t \in (0, T)} \Theta(t), \frac{\lambda}{8} \int_0^T \int_Y |y - \eta|^2 \bar{f}\bar{\mu}(dy) \right\} \leq e^{DT} \left(\Theta(0) + \int_0^T S(t)dt \right). \quad (4.47)$$

□

It remains to take care of each term in the right hand side of (4.47). Let $a, b > 0$ be two constants.

$$\begin{aligned}
& \int_0^T \int \frac{1}{N} |x|^2 f(t, x) \mu(dx) dt \stackrel{(4.32)}{\leq} 2 \left(\alpha + \frac{2C_1}{\hat{\rho}} \right) T; \\
& \int_0^T \int \frac{1}{Nf} \nabla \cdot A \nabla f \mu(dx) dt \stackrel{(4.30)}{\leq} C_1; \\
& \int_0^T \left(\int \frac{1}{N} |x|^2 f \mu(dx) \right)^{\frac{1}{2}} \left(a \left(\int \frac{1}{Nf} \nabla f \cdot A \nabla f \mu(dx) \right)^{\frac{1}{2}} + b \left(\bar{A} \nabla_Y \bar{H}(\eta) \cdot \nabla_Y \bar{H}(\eta) \right)^{\frac{1}{2}} \right) dt \\
& \leq \left(\int_0^T \int \frac{1}{N} |x|^2 f \mu(dx) dt \right)^{\frac{1}{2}} \left(a \left(\int_0^T \int \frac{1}{Nf} \nabla f \cdot A \nabla f \mu(dx) dt \right)^{\frac{1}{2}} + b \left(\int_0^T \bar{A} \nabla_Y \bar{H}(\eta) \cdot \nabla_Y \bar{H}(\eta) dt \right)^{\frac{1}{2}} \right) \\
& \leq \sqrt{2T} \left(\alpha + \frac{2C_1}{\hat{\rho}} \right)^{\frac{1}{2}} \left(a \sqrt{C_1} + \sqrt{2} b (H(\eta_0) - H(\eta_T)) + CT(1 + e^{CT} \bar{H}(\eta_0))^{\frac{1}{2}} \right).
\end{aligned}$$

Substituting these estimate to (4.47) concludes the proof of Theorem 4.1.

4.3.2 Sketch of proof of Theorem 4.5

In this section, we give the main arguments of the proof of Theorem 4.5, which exactly follows the method of [Fat13b].

- First, we decompose the relative entropy with respect to the local Gibbs state into a macroscopic component and a fluctuations component. Since $G(x)$ only depends on the macroscopic profile $y = Px$, we have

$$\text{Ent}_{G\mu}(f\mu) = \text{Ent}_{\bar{G}\bar{\mu}}(\bar{f}\bar{\mu}) + \int_Y \text{Ent}_{\mu(dx|y)}(f\mu) \bar{G}(y) \bar{\mu}(dy).$$

- The fluctuations component $\int_0^T \int_Y \text{Ent}_{\mu(dx|y)}(f\mu) \bar{G}(y) \bar{\mu}(dy) dt$ can be bounded using the logarithmic Sobolev inequality for $\mu(dx|y)$, assumption (vi) and the bound on the microscopic entropy production of Lemma 4.10.
- For the macroscopic component, since $\bar{G}\bar{\mu}$ is log-concave, we can use the HWI inequality of [OV00], which states that

$$\text{Ent}_{\bar{G}\bar{\mu}}(\bar{f}\bar{\mu}) \leq W_2(\bar{f}\bar{\mu}, \bar{G}\bar{\mu}) \sqrt{I_{\bar{G}\bar{\mu}}(\bar{f}\bar{\mu})},$$

where the Wasserstein distance W_2 is taken with respect to the norm $|\cdot|_Y$, and I is the Fisher information

$$I_{\bar{G}\bar{\mu}}(\bar{f}\bar{\mu}) := \int \frac{|\nabla(\bar{f}/\bar{G})|^2}{\bar{f}/\bar{G}} \bar{G} d\bar{\mu}.$$

As a consequence, to obtain convergence in relative entropy, we only require convergence in Wasserstein distance and a bound on the Fisher information.

- We already have a bound on $\int_0^T W_2(\bar{f}\bar{\mu}, \delta_{\eta(t)})^2 dt$ from Theorem 4.1. Moreover,

$$W_2(\bar{G}\bar{\mu}, \delta_\eta)^2 \leq \frac{M}{\lambda N}$$

by Proposition 4.1 of [Fat13b]. A bound on $\int_0^T W_2(\bar{f}\bar{\mu}, \bar{G}\bar{\mu}) dt$ immediately follows from the triangle inequality.

- Finally, the time-integral of the Fisher information can be bounded using the bounds on the entropy production of Lemma 4.10. This concludes the proof of (a).
- (b) can be deduced from (a) using elementary inequalities and the bound

$$\begin{aligned} & \left| \frac{1}{N} \int \Phi(G^n) d\mu - \bar{H}(\eta) \right| \\ & \leq \frac{(M-1)}{2N} \max \left(\left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\Lambda N} \right) \right|, \left| \log \left(\frac{\Gamma(Y, |\cdot|_Y)^{2/(M-1)}}{\lambda N} \right) \right| \right) \\ & \quad + \sqrt{\frac{M}{\lambda N}} |\nabla \bar{H}(\eta)|_Y, \end{aligned}$$

which was proven in Proposition 4.1 of [Fat13b].

4.4 Application to weakly asymmetric Kawasaki dynamics

In this section, we prove Theorem 4.6. First, we give a precise definition of the notion of weak solution to the limiting equation (4.14).

Definition 4.12. $\zeta = \zeta(t, \theta)$ is called a weak solution of (4.14) on $[0, T] \times \mathbb{T}^1$ if

$$\zeta \in L^\infty(L^2), \quad \frac{\partial \zeta}{\partial t} \in L^2(H^{-1}), \quad \varphi'(\zeta) \in L^2(L^2), \quad (4.48)$$

and

$$\left\langle g, \frac{\partial \zeta}{\partial t} \right\rangle_{H^{-1}} = - \int_{\mathbb{T}^1} g \varphi'(\zeta) d\theta + \int_{\mathbb{T}^1} G \varphi'(\zeta) d\theta, \quad \text{for all } g \in L^2(\mathbb{T}^1), \quad \text{for almost every } t \in [0, T], \quad (4.49)$$

where G is the (unique up to a set of Lebesgue measure 0) function on the torus such that $\int_{\mathbb{T}^1} G d\theta = 0$ and $G' = g$.

As in Corollary 4.3, consider a sequence $\{M_\ell, N_\ell\}_{\ell=1}^\infty$ such that

$$M_\ell \uparrow \infty; \quad N_\ell \uparrow \infty; \quad \frac{N_\ell}{M_\ell} \uparrow \infty.$$

Let $\bar{\eta}_0^\ell$ be a step-function approximation of ζ_0 , such that

$$\|\bar{\eta}_0^\ell - \zeta_0\|_{L^2} \xrightarrow{\ell \uparrow \infty} 0. \quad (4.50)$$

Consider η^ℓ the solutions to

$$\frac{d\eta^\ell}{dt} = -(\bar{A} + \bar{J}) \nabla_Y \bar{H}(\eta^\ell), \quad \eta^\ell(0) = \bar{\eta}_0^\ell.$$

To obtain Theorem 4.6 from Theorem 4.3, we shall need to study the convergence of the sequence η^ℓ . It is given by the following result.

Proposition 4.13. *With the notations above, the sequence of step functions $\bar{\eta}^\ell$ converge strongly in $L^\infty(H^{-1})$ to the unique weak solution of (4.14) with initial condition ζ_0 .*

The key estimate which will allow us to pass to the limit is the fact that, when N goes to infinity, the Euclidean product associated to A^{-1} behaves like the H^{-1} norm. This is the content of the following lemma :

Lemma 4.14. *There exists $C < +\infty$ such that, for any $x \in X$, if \bar{x} is the associated step function, then*

$$\frac{1}{C} \|\bar{x}\|_{H^{-1}}^2 \leq \frac{1}{N} \langle A^{-1}x, x \rangle \leq C \|\bar{x}\|_{H^{-1}}^2.$$

Moreover, if \bar{x} is bounded in L^2 , then

$$\left| \|\bar{x}\|_{H^{-1}}^2 - \frac{1}{N} \langle A^{-1}x, x \rangle \right| \leq \frac{C}{N}.$$

These estimates have been proven in section 6.3 of [GOVW09].

We delay the proof of Proposition 4.13, and first prove Theorem 4.6

Proof of Theorem 4.6. Our aim is to apply Corollary 4.3. To do this, we need to check that assumptions (i) to (x) hold with uniform constants. Assumptions (i) to (vii) have been checked in [GOVW09], and assumption (viii) in [Fat13b]. Assumption (x) can be immediately check by the direct computation of JA and AJ . Finally, it is easy to see that for any $x \in X$, we have

$$\begin{aligned} \langle -J^2x, x \rangle &= |Jx|^2 \\ &= \frac{N^2}{4} \sum_{i=1}^N (x_{i+1} - x_{i-1})^2 \\ &\leq \frac{N^2}{4} \sum_{i=1}^N 2(x_{i+1} - x_i)^2 + 2(x_i - x_{i-1})^2 \\ &= N^2 \sum_{i=1}^N (x_{i+1} - x_i)^2 \\ &= \langle Ax, x \rangle \end{aligned} \tag{4.51}$$

and therefore assumption (ix) holds with $c = 1$.

Applying Corollary 4.3, we get

$$\lim_{\ell \uparrow \infty} \sup_{0 \leq t \leq T} \int \langle (x - NP^t \eta^\ell(t)), A^{-1}(x - NP^t \eta^\ell(t)) \rangle f(t, x) \mu(dx) = 0.$$

By Lemma 4.14, this implies

$$\lim_{\ell \uparrow \infty} \sup_{0 \leq t \leq T} \int \|\bar{x} - \bar{\eta}^\ell(t)\|_{H^{-1}}^2 f(t, x) \mu(dx) = 0.$$

Applying Proposition 4.13 and using the triangle inequality then concludes the proof. \square

We now turn to the proof of Proposition 4.13. It is based on the following six lemmas, and closely follows the method of [GOVW09], with additional arguments to take into account the extra first-order term.

Lemma 4.15. *Assume \bar{H} is convex. Then η satisfies (4.5) with initial condition $\eta(0) = \eta_0$ if and only if*

$$2 \int_0^T \bar{H}(\eta) \beta(t) dt \leq \int_0^T \left[\bar{H}(\eta + g) + \bar{H}(\eta - PA^{-1}NJP^t g) \right] \beta(t) dt - \int_0^T \langle g, (\bar{A})^{-1} \eta \rangle_Y \dot{\beta}(t) dt, \tag{4.52}$$

for all $g \in Y$ and smooth $\beta: [0, T] \rightarrow [0, \infty)$.

Similarly, assume that φ is convex. Then ζ satisfies (4.49) if and only if

$$\begin{aligned}
& 2 \int_0^T \int_{\mathbb{T}^1} \varphi(\zeta(t, \theta)) \beta(t) d\theta dt \\
& \leq \int_0^T \int_{\mathbb{T}^1} [\varphi(\zeta(t, \theta) + g(\theta)) + \varphi(\zeta(t, \theta) - G(\theta))] \beta(t) d\theta dt - \int_0^T \langle g(\cdot), \zeta(t, \cdot) \rangle_{H^{-1}} \dot{\beta}(t) dt,
\end{aligned} \tag{4.53}$$

for all $g \in L^2\mathbb{T}^1$ and smooth $\beta: [0, T] \rightarrow [0, \infty)$, where G is the (unique up to a set of Lebesgue measure 0) function on the torus such that $\int_{\mathbb{T}^1} G d\theta = 0$ and $G' = g$.

Démonstration. The proof of this Lemma is modified from that of Lemma 36 in [GOVW09]. We show that (4.5) is equivalent to (4.52). The equivalence of (4.49) and (4.53) follows analogously.

The weak form of (4.5) is given by

$$\int_0^T \langle g, (\bar{A})^{-1} \eta \rangle_Y \dot{\beta}(t) dt = \int_0^T \left[\langle g, \nabla_Y \bar{H}(\eta) \rangle_Y - \langle PA^{-1}NJP^t g, \nabla_Y \bar{H}(\eta) \rangle_Y \right] \beta(t) dt, \tag{4.54}$$

for all $g \in Y$ and smooth $\beta: [0, T] \rightarrow [0, \infty)$. We now show that (4.54) implies (4.52). Since \bar{H} is convex, we have

$$\begin{aligned}
\langle g - PA^{-1}NJP^t g, \nabla_Y \bar{H}(\eta) \rangle_Y & \leq (\bar{H}(\eta + g) - \bar{H}(\eta)) + (\bar{H}(\eta - PA^{-1}NJP^t g) - \bar{H}(\eta)) \\
& = -2\bar{H}(\eta) + \bar{H}(\eta + g) + \bar{H}(\eta - PA^{-1}NJP^t g).
\end{aligned} \tag{4.55}$$

Substituting (4.55) into (4.54), we obtain (4.52)

$$\begin{aligned}
& \int_0^T \langle g, (\bar{A})^{-1} \eta \rangle_Y \dot{\beta}(t) dt \\
& \leq -2 \int_0^T \bar{H}(\eta) \beta(t) dt + \int_0^T \left[\bar{H}(\eta + g) + \bar{H}(\eta - PA^{-1}NJP^t g) \right] \beta(t) dt.
\end{aligned} \tag{4.56}$$

Next we show (4.52) implies (4.54). Take $\tilde{g} = \varepsilon g$ in (4.54), for some $\varepsilon > 0$ and $g \in Y$, we get

$$\int_0^T \langle g, (\bar{A})^{-1} \eta \rangle_Y \dot{\beta}(t) dt \leq \int_0^T \left[\frac{\bar{H}(\eta + \varepsilon g) - \bar{H}(\eta)}{\varepsilon} + \frac{\bar{H}(\eta - \varepsilon PA^{-1}NJP^t g) - \bar{H}(\eta)}{\varepsilon} \right] \beta(t) dt.$$

By passing to the limit $\varepsilon \rightarrow 0$, we get

$$\int_0^T \langle g, (\bar{A})^{-1} \eta \rangle_Y \dot{\beta}(t) dt \leq \int_0^T \langle g - PA^{-1}NJP^t g, \nabla_Y \bar{H}(\eta) \rangle_Y \beta(t) dt.$$

Similarly now by taking $\tilde{g} = -\varepsilon g$, we obtain the opposite inequality

$$\int_0^T \langle g, (\bar{A})^{-1} \eta \rangle_Y \dot{\beta}(t) dt \geq \int_0^T \langle g - PA^{-1}NJP^t g, \nabla_Y \bar{H}(\eta) \rangle_Y \beta(t) dt.$$

Thus (4.54) is proven. \square

Lemma 4.16. *Let $\{\eta^\ell\}_{\ell=1}^\infty$ be a sequence of solutions of (4.5) with initial data η_0^ℓ satisfying $\|\bar{\eta}_0^\ell\|_{L^2} \leq C$. There exists a constant C independent of ℓ such that*

$$\int_0^T \left\langle \frac{d\eta^\ell}{dt}(t), (\bar{A})^{-1} \frac{d\eta^\ell}{dt}(t) \right\rangle dt \leq C, \tag{4.57}$$

$$\sup_{t \in [0, T]} \langle \eta^\ell(t), \eta^\ell(t) \rangle_Y \leq C. \tag{4.58}$$

As a consequence, there is a subsequence of the sequence of the associated step functions $\bar{\eta}^\ell$ and a function η_* such that

$$\bar{\eta}^\ell \rightharpoonup \eta_* \quad \text{weak-}^* \text{ in } L^\infty(L^2) = (L^1(L^2))^*.$$

Démonstration. According to proof of (4.31), we have

$$\bar{H}(\eta^\ell(t)) \leq e^{C(T+1)} \bar{H}(\eta_0^\ell) \quad \text{for all } t \in [0, T]. \quad (4.59)$$

Since \bar{H} is strictly convex, we obtain

$$\langle \eta^\ell(t), \eta^\ell(t) \rangle_Y \leq C(\bar{H}(\eta^\ell(t)) + 1) \leq C e^{C(T+1)} \bar{H}(\eta_0^\ell) \leq C,$$

which is (4.58). Now we establish (4.57). From (4.5), we have

$$\begin{aligned} \langle \dot{\eta}^\ell(t), (\bar{A})^{-1} \dot{\eta}^\ell(t) \rangle &= \langle \bar{A}(I + PJA^{-1}NP^t) \nabla_Y \bar{H}(\eta^\ell(t)), (I + PJA^{-1}NP^t) \nabla_Y \bar{H}(\eta^\ell(t)) \rangle \\ &\leq 2(\langle \bar{A} \nabla_Y \bar{H}(\eta^\ell(t)), \nabla_Y \bar{H}(\eta^\ell(t)) \rangle + \langle \bar{A} PJA^{-1}NP^t \nabla_Y \bar{H}(\eta^\ell(t)), PJA^{-1}NP^t \nabla_Y \bar{H}(\eta^\ell(t)) \rangle) \\ &\stackrel{(4.27)}{\leq} 2(\langle \bar{A} \nabla_Y \bar{H}(\eta^\ell(t)), \nabla_Y \bar{H}(\eta^\ell(t)) \rangle + c |\nabla_Y \bar{H}(\eta^\ell(t))|^2) \\ &\stackrel{(iii)}{\leq} 2(\langle \bar{A} \nabla_Y \bar{H}(\eta^\ell(t)), \nabla_Y \bar{H}(\eta^\ell(t)) \rangle + C(\bar{H}(\eta^\ell(t)) + 1)) \end{aligned}$$

Therefore,

$$\begin{aligned} \int_0^T \langle \dot{\eta}^\ell(t), (\bar{A})^{-1} \dot{\eta}^\ell(t) \rangle dt &\leq 2 \int_0^T (\langle \bar{A} \nabla_Y \bar{H}(\eta^\ell(t)), \nabla_Y \bar{H}(\eta^\ell(t)) \rangle + C(\bar{H}(\eta^\ell(t)) + 1)) dt \\ &\stackrel{(4.31), (4.59)}{\leq} C, \end{aligned}$$

which is (4.57). \square

Lemma 4.17. *Let $\{\eta^\ell\}_1^\infty$ be a sequence of solutions of (4.5) satisfying Lemma 4.16. We take any subsequence that the associated step functions weak- * convergence in $(L^1(L^2))^*$ to a limit η_* . Then on any bounded time interval, we have*

$$\eta_* \in L^\infty(L^2), \quad \frac{\partial \eta_*}{\partial t} \in L^2(H^{-1}), \quad \varphi'(\eta_*) \in L^2(L^2). \quad (4.60)$$

Démonstration. Having the estimate in Lemma 4.16, the proof of this Lemma is the same as that of Lemma 35 in [GOVW09]; hence we omit it here. \square

Lemma 4.18. *If $\bar{g}^\ell \rightarrow \bar{g}$ strongly in $H^{-1}(\mathbb{T})$, then $-\overline{PA^{-1}JNP^t g^\ell} \rightarrow G$ strongly in $L^2(\mathbb{T})$ where G is the primitive of \bar{g} .*

Démonstration. Set

$$D = N \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ -1 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}, \quad (4.61)$$

then we can write

$$A = DD^T, \quad J = \frac{1}{2}(D^T - D).$$

Hence

$$JA^{-1} = \frac{1}{2}(D^T - D)(D^T D)^{-1} = \frac{1}{2}(D^T - D)D^{-1}(D^T)^{-1} = \frac{1}{2}(D^{-1} - (D^T)^{-1}). \quad (4.62)$$

The inverse of D and D^T can be computed explicitly

$$D^{-1} = \frac{1}{2N} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & -1 & & & 1 \end{pmatrix}, \quad (D^T)^{-1} = (D^{-1})^T.$$

So we obtain

$$D^{-1} - (D^T)^{-1} = \frac{1}{N} \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & \\ -1 & & & & 0 \end{pmatrix}, \quad (4.63)$$

Let $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_M \end{pmatrix} \in Y = \mathbb{R}^M$ be given. We now compute $\overline{PA^{-1}JNP^t\xi}$ explicitly in three steps.

First, by definition of P^t , we have

$$NP^t\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_1 \\ \xi_2 \\ \vdots \\ \xi_2 \\ \vdots \\ \xi_M \\ \vdots \\ \xi_M \end{pmatrix} \in \mathbb{R}^N = \mathbb{R}^{KM}. \quad (4.64)$$

Second, from (4.62), (4.63) and (4.64), we have

$$A^{-1}JNP^t\xi = \frac{1}{2N} \begin{pmatrix} K(\xi_1 + \cdots + \xi_M) \\ (K-1)\xi_1 + K(\xi_2 + \cdots + \xi_M) \\ \vdots \\ K(\xi_2 + \cdots + \xi_M) \\ (K-1)\xi_2 + K(\xi_3 + \cdots + \xi_M) \\ \vdots \\ K(\xi_3 + \cdots + \xi_M) \\ \vdots \\ \xi_M \end{pmatrix} - \frac{1}{2N} \begin{pmatrix} \xi_1 \\ 2\xi_1 \\ \vdots \\ K\xi_1 \\ K\xi_1 + \xi_2 \\ \vdots \\ K(\xi_1 + \cdots + \xi_M) \end{pmatrix}.$$

Therefore, by definition of P ,

$$\begin{aligned} PA^{-1}JNP^t\xi &= \frac{1}{2M} \begin{pmatrix} \xi_1 + \xi_2 + \cdots + \xi_M \\ \xi_2 + \cdots + \xi_M \\ \vdots \\ \xi_M \end{pmatrix} - \frac{1}{2M} \begin{pmatrix} \xi_1 \\ \xi_1 + \xi_2 \\ \vdots \\ \xi_1 + \xi_2 + \cdots + \xi_M \end{pmatrix} \\ &= -\frac{1}{2M} \begin{pmatrix} \xi_1 \\ 2\xi_1 + \xi_2 \\ \vdots \\ 2(\xi_1 + \cdots + \xi_{M-1}) + \xi_M \end{pmatrix}. \end{aligned}$$

This implies that $-\overline{PA^{-1}JNP^t\xi} = \Upsilon_\xi$, where Υ_ξ is the primitive of $\bar{\xi}$. The assertion then follows since

$$(\bar{g}^\ell \rightarrow \bar{g} \text{ strongly in } H^{-1}(\mathbb{T})) \iff (\Upsilon_{g^\ell} \rightarrow \Upsilon_g \equiv G \text{ strongly in } L^2(\mathbb{T})).$$

□

Lemma 4.19. *Suppose that the sequence η^ℓ satisfies (4.57), (4.58) and (4.52), and consider a subsequence such that*

$$\bar{\eta}^\ell \rightharpoonup \eta_* \quad \text{weak-}^* \text{ in } L^\infty(L^2) = (L^1(L^2))^*.$$

holds. Let $\xi^\ell = \pi_\ell(\xi + \eta^\ell) - \eta^\ell$, where ξ is an arbitrary L^2 function and π_ℓ is the L^2 -projection onto elements of Y . Let Ξ be the primitive with average 0 of ξ . Then we have

(i)

$$\liminf_\ell \int_0^T \bar{H}(\eta^\ell(t))\beta(t)dt \geq \int_0^T \int_{\mathbb{T}} \varphi(\eta_*(t, \theta))\beta(t)d\theta dt;$$

(ii)

$$\lim_\ell \int_0^T \bar{H}(\eta^\ell(t) + \xi^\ell(t))\beta(t)dt = \int_0^T \int_{\mathbb{T}} \varphi(\eta_*(t, \theta) + \xi(\theta))\beta(t)d\theta dt;$$

(iii)

$$\lim_\ell \int_0^T \bar{H}(\eta^\ell(t) - PA^{-1}JNP^t\xi^\ell(t))\beta(t)dt = \int_0^T \int_{\mathbb{T}} \varphi(\eta_*(t, \theta) - \Xi(\theta))\beta(t)d\theta dt;$$

(iv)

$$\lim_\ell \int_0^T \langle \xi^\ell(t), \bar{A}^{-1}\eta^\ell(t) \rangle_Y \dot{\beta}(t)dt = \int_0^T \langle \xi(\theta), \eta_*(t, \theta) \rangle_{H^{-1}} \dot{\beta}(t)dt.$$

Démonstration. (i), (ii) and (iv) have already been proven in Lemma 37 of [GOVW09], so we only have to prove (iii).

Since η^ℓ converges to η_* and $PA^{-1}JNP^t\xi^\ell(t)$ converges to Ξ , by weak lower-semicontinuity and the uniform convergence of ψ_K to φ we immediately get

$$\liminf_\ell \int_0^T \bar{H}(\eta^\ell(t) - PA^{-1}JNP^t\xi^\ell(t))\beta(t)dt \geq \int_0^T \int_{\mathbb{T}} \varphi(\eta_*(t, \theta) - \Xi(\theta))\beta(t)d\theta dt$$

so we only need to prove the associated upper bound. Let $g^\ell(t)$ be a sequence of elements of Y such that \bar{g}^ℓ strongly converges in $L^\infty(L^2)$ to $\eta_* - \Xi$. Since we then have

$$\int_0^T \bar{H}(g^\ell(t))\beta(t)dt \longrightarrow \int_0^T \int_{\mathbb{T}} \varphi(\eta_*(t, \theta) - \Xi(\theta))\beta(t)d\theta dt$$

we only need to show that

$$\limsup_{\ell} \int_0^T \bar{H}(\eta^\ell(t) - PA^{-1}JNP^t\xi^\ell(t))\beta(t)dt - \int_0^T \bar{H}(g^\ell(t))\beta(t)dt \leq 0.$$

Let A_M be the discrete Laplacian with scaling factor M^2 on Y . Since ψ_K is convex, we have

$$\begin{aligned} \bar{H}(\eta^\ell(t) - PA^{-1}JNP^t\xi^\ell(t)) - \bar{H}(g^\ell(t)) &= \frac{1}{M} \sum_{i=1}^M \psi_K(\eta_i^\ell - (PA^{-1}JNP^t\xi^\ell)_i) - \psi_K(g_i^\ell) \\ &\leq \frac{1}{M} \sum_{i=1}^M \psi'_K(\eta_i^\ell - (PA^{-1}JNP^t\xi^\ell)_i)(\eta_i^\ell - (PA^{-1}JNP^t\xi^\ell)_i - g_i^\ell) \\ &= \langle \nabla \bar{H}(\eta^\ell - PA^{-1}JNP^t\xi^\ell), (\eta^\ell - PA^{-1}JNP^t\xi^\ell - g^\ell) \rangle_Y \\ &\leq \langle A_M \nabla \bar{H}(\eta^\ell - PA^{-1}NP^t\xi^\ell), \nabla \bar{H}(\eta^\ell - PA^{-1}NP^t\xi^\ell) \rangle_Y^{1/2} \\ &\quad \times \langle A_M^{-1}(\eta^\ell - PA^{-1}NP^t\xi^\ell - g^\ell), (\eta^\ell - PA^{-1}NP^t\xi^\ell - g^\ell) \rangle_Y^{1/2} \end{aligned}$$

Since $\langle A_M^{-1}\cdot, \cdot \rangle_Y$ behaves like the H^{-1} norm, the fact that $\eta^\ell - PA^{-1}NP^t\xi^\ell$ and g^ℓ converge to the same limit in $L^\infty(H^{-1})$ implies that

$$\langle A_M^{-1}(\eta^\ell - PA^{-1}NP^t\xi^\ell - g^\ell), (\eta^\ell - PA^{-1}NP^t\xi^\ell - g^\ell) \rangle_Y \longrightarrow 0,$$

and therefore it will be enough to show that

$$\int_0^T \langle A_M \nabla \bar{H}(\eta^\ell - PA^{-1}NP^t\xi^\ell), \nabla \bar{H}(\eta^\ell - PA^{-1}NP^t\xi^\ell) \rangle_Y dt < C.$$

Since under our assumptions ψ'_K is bi-Lipschitz, we have

$$\begin{aligned} &\langle A_M \nabla \bar{H}(\eta^\ell - PA^{-1}NP^t\xi^\ell), \nabla \bar{H}(\eta^\ell - PA^{-1}NP^t\xi^\ell) \rangle_Y \\ &= M \sum_{i=1}^M (\psi'_K(\eta_{i+1}^\ell - (PA^{-1}NP^t\xi^\ell)_{i+1}) - \psi'_K(\eta_i^\ell - (PA^{-1}NP^t\xi^\ell)_i))^2 \\ &\leq CM \sum_{i=1}^M (\eta_{i+1}^\ell - (PA^{-1}NP^t\xi^\ell)_{i+1} - (\eta_i^\ell - (PA^{-1}NP^t\xi^\ell)_i))^2 \\ &\leq CM \sum_{i=1}^M (\eta_{i+1}^\ell - \eta_i^\ell)^2 + ((PA^{-1}NP^t\xi^\ell)_{i+1} - (PA^{-1}NP^t\xi^\ell)_i)^2 \\ &\leq CM \sum_{i=1}^M (\psi'_K(\eta_{i+1}^\ell) - \psi'_K(\eta_i^\ell))^2 + \frac{C}{M} \sum_{i=1}^M (\xi_{i+1}^\ell - \xi_i^\ell)^2 \\ &\leq C \langle A_M \nabla \bar{H}(\eta^\ell), \bar{H}(\eta^\ell) \rangle + C \|\bar{\xi}^\ell\|_{L^2}^2. \end{aligned}$$

Since $\bar{\xi}^\ell$ converges in L^2 , $\|\bar{\xi}^\ell\|_{L^2}^2$ is bounded. To conclude, we then only require (4.57) and the fact that

$$\langle A_M y, y \rangle \leq C \langle \bar{A} y, y \rangle \quad \forall y \in Y. \quad (4.65)$$

This statement is equivalent to bounding from below A_M^{-1} by \bar{A}^{-1} . This does hold, since

we have

$$\begin{aligned} \langle \bar{A}^{-1}y, y \rangle_Y &= \frac{1}{N} \langle A^{-1}NP^t y, NP^t y \rangle_X \\ &\leq C \|N\bar{P}^t y\|_{H^{-1}}^2 \\ &\leq C \|\bar{y}\|_{H^{-1}}^2 \\ &\leq C \langle A_M^{-1}y, y \rangle_Y \end{aligned}$$

which concludes the proof. \square

Finally, we need to prove uniqueness of solutions to the limiting PDE :

Lemma 4.20. *Given an initial condition ζ_0 , there is at most one solution to (4.14).*

Démonstration. Let ζ_1 and ζ_2 be two solutions of (4.14) with same initial condition. Let $F(t) := 2^{-1} \|\zeta_1(t, \cdot) - \zeta_2(t, \cdot)\|_{H^{-1}}^2$, and let g_1 and g_2 be mean-zero primitives (in space) of ζ_1 and ζ_2 . Then, for any $\lambda > 0$,

$$\begin{aligned} F'(t) &= - \int_{\mathbb{T}} (\varphi'(\zeta_1) - \varphi'(\zeta_2))(\zeta_1 - \zeta_2) d\theta + \int_{\mathbb{T}} (\varphi'(\zeta_1) - \varphi'(\zeta_2))(g_1 - g_2) d\theta \\ &\leq - \frac{\inf \varphi''}{2} \int_{\mathbb{T}} (\zeta_1 - \zeta_2)^2 d\theta + \frac{\lambda}{2} \int_{\mathbb{T}} (\varphi'(\zeta_1) - \varphi'(\zeta_2))^2 d\theta \\ &\quad + \frac{1}{2\lambda} \int_{\mathbb{T}} (g_1 - g_2)^2 d\theta \\ &\leq - \frac{\inf \varphi''}{2} \int_{\mathbb{T}} (\zeta_1 - \zeta_2)^2 d\theta + \frac{\lambda \sup \varphi''}{2} \int_{\mathbb{T}} (\zeta_1 - \zeta_2)^2 d\theta + \frac{1}{\lambda} F(t) \end{aligned}$$

Taking $\lambda = \frac{\inf \varphi''}{\sup \varphi''}$, we obtain a differential inequality which, by Gronwall's lemma, implies that $\zeta_1 = \zeta_2$. \square

We can now prove Proposition 4.13 :

Proof of Proposition 4.13. According to Lemma 4.16, we can consider a subsequence such that

$$\bar{\eta}^\ell \rightharpoonup \eta_* \quad \text{weak-}^* \text{ in } L^\infty(L^2) = (L^1(L^2))^*$$

and strongly in $L^\infty(H^{-1})$. By Lemma 4.17, η_* satisfies (4.48). According to Lemma 4.15, η^ℓ satisfies (4.52). Passing to the limit using Lemma 4.19, we see that η_* satisfies (4.53), and therefore is a weak solution of (4.14).

Since Lemma 4.20 guarantees uniqueness of the weak solution, the full sequence $(\eta^\ell)_\ell$ converges to the unique weak solution of (4.14). \square

Chapitre 5

Modified logarithmic Sobolev inequalities for canonical ensembles

Ce chapitre reprend les résultats de l'article [Fat13a], qui a été soumis pour publication.

Abstract : In this chapter, we prove modified logarithmic Sobolev inequalities for canonical ensembles with superquadratic single-site potential. These inequalities were introduced by Bobkov and Ledoux, and are closely related to concentration of measure and transport-entropy inequalities. Our method is an adaptation of the iterated two-scale approach that was developed by Menz and Otto to prove the usual logarithmic Sobolev inequality in this context. As a consequence, we obtain convergence in Wasserstein distance W_p for Kawasaki dynamics on the Ginzburg-Landau model.

5.1 Introduction

The logarithmic Sobolev inequality is an inequality allowing to embed the Sobolev space $H^1(\mu)$ in the Orlicz space $L^2 \log L(\mu)$, just like the usual Sobolev inequalities embed H^1 in L^p for some $p > 2$. It was introduced by Gross in [Gro75], and has been shown to be very useful in some problems of statistical physics, such as long-time convergence to equilibrium, and hydrodynamic limits (see for example [GOVW09]).

One case of measures where such an inequality has been useful is for canonical ensembles, which are probability measures $\mu(dx) = \exp(-\sum \psi(x_i))$ on the hyperplane $\{\sum x_i = Nm\}$ of \mathbb{R}^N . In the recent contribution [MO13], Menz and Otto proved that, if the function ψ is a bounded perturbation of a uniformly convex function, then the canonical ensemble satisfies a logarithmic Sobolev inequality, with a constant independent of the mean m and the dimension N .

The result of [MO13] covers potentials which behave like $|x|^p$ for some $p \geq 2$. A natural question is whether we can improve the LSI when p is strictly larger than 2. For this purpose, we investigate whether a variant of the LSI called the modified logarithmic Sobolev inequality, which was introduced by Bobkov and Ledoux in [BL00], is satisfied by canonical ensembles. Our method is a generalization of the iterated two-scale approach that was used in [MO13] to obtain the usual LSI.

5.2 Background and Main Results

In this paper, we are interested in the following family of inequalities, which generalizes the logarithmic Sobolev inequality.

Definition 5.1. A probability measure μ satisfies a p -modified logarithmic Sobolev inequality with parameter ρ if, for all positive compactly supported C^1 function f , we have

$$\text{Ent}_\mu(f) \leq \frac{1}{\rho} \int \frac{\|\nabla f\|_q^q}{f^{q-1}} d\mu, \quad (5.1)$$

where q is the dual exponent of p , that is

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Equivalently, μ satisfies this inequality if for any such function f , we have

$$\text{Ent}_\mu(f^q) \leq \frac{q^q}{\rho} \int \|\nabla f\|_q^q d\mu. \quad (5.2)$$

In the case $p = 2$, this is the usual logarithmic Sobolev inequality. Many results on these inequalities can be found in [BZ05], and we recall some of them in the sequel. It is well known that the usual LSI implies Gaussian concentration properties. In the same way, modified logarithmic Sobolev inequalities are linked to the following form of concentration of measure :

Definition 5.2. A probability measure μ on a metric space (X, d) has the p -exponential concentration property with parameter c if, for any 1-Lipschitz function $f : X \rightarrow \mathbb{R}$ and every $r \geq 0$, we have

$$\mu\left(f \geq \int f d\mu + r\right) \leq \exp\left(-\frac{cr^p}{p(p-1)^{p-1}}\right).$$

Theorem 5.3. *If μ satisfies p -LSI(ρ), then μ satisfies p -exponential concentration for the ℓ^p distance*

We refer to Theorem 1.3 in [BZ05] for a proof of this result.

We consider a (periodic) lattice spin system of N continuous variables governed by a Ginzburg-Landau type potential $\psi : \mathbb{R} \rightarrow \mathbb{R}$. The grand canonical measure on \mathbb{R}^N has density

$$\frac{d\mu_N}{d\mathcal{L}^N}(x) = \frac{1}{Z} \exp\left(-\sum_{i=1}^N \psi(x_i)\right). \quad (5.3)$$

We shall assume that the potential ψ is of class C^1 and is of the form

$$\psi(x) = \psi_c(x) + \delta\psi(x); \quad \psi_c''(x) \geq c(1 + |x|^{p-2}); \quad \|\delta\psi\|_\infty + \|\delta\psi'\|_\infty < +\infty. \quad (5.4)$$

Under these assumptions, ψ_c is a uniformly p -convex and uniformly convex function. A typical example would be the quartic double-well potential $\psi(x) = (x^2 - 1)^2$. For a definition of p -convexity, see Theorem 5.20

Remark 5.4. Our results are still valid if we only ask ψ_c to satisfy $\psi_c''(x) \geq c(1 + |x - x_0|^{p-2})$ for some x_0 . The proof is exactly the same, but the extra assumption makes the calculations easier to read.

To simplify notations, we define the Hamiltonian

$$H(x) := \sum_{i=1}^N \psi(x_i) + \log Z, \quad (5.5)$$

so that $\mu(dx) = \exp(-H(x))dx$.

We will add to the situation a constraint of fixed mean spin. The phase state space is

$$X_{N,m} := \left\{ x \in \mathbb{R}^N, \frac{1}{N} \sum_{i=1}^N x_i = m \right\},$$

where m is an arbitrary real number. This space is a hyperplane of \mathbb{R}^N with a fixed mean constraint. We equip this space with the ℓ^2 inner product

$$\langle x, \tilde{x} \rangle_X = \sum_{i=1}^N x_i \tilde{x}_i. \quad (5.6)$$

For a given $m \in \mathbb{R}$, we consider the restriction $\mu_{N,m}$ of the grand canonical measure to $X_{N,m}$, that is

$$\frac{d\mu_{N,m}}{d\mathcal{L}^{N-1}}(x) = \frac{1}{Z} \mathbb{1}_{(1/N) \sum x_i = m} \exp \left(- \sum_{i=1}^N \psi(x_i) \right). \quad (5.7)$$

This measure is called the canonical ensemble. It gives the distribution of the random variables x_i conditioned on the event that their mean value is given by m .

It was shown in [MO13] that when the single site potential satisfies assumption (5.4) with $p = 2$, then the canonical ensemble satisfies the classical logarithmic Sobolev inequality for some constant $\rho > 0$ that is independent of both m and N . Our aim in this paper is to generalize this result for the modified LSI, and we obtain the following :

Theorem 5.5. *Under the assumption (5.4), the canonical ensemble $\mu_{N,m}$ satisfies p -LSI(ρ) for some constant $\rho > 0$ that is independent of both N and m .*

The proof in [MO13] uses a method called the iterated two-scale approach, which generalizes a method that was developed in [GOVW09]. The idea is to use a decomposition of the system into a macroscopic component and a fluctuations component. There are then two main ideas : the first is to prove that if the laws of both the macroscopic and fluctuations part satisfy the desired functional inequality, then the law of the full system also satisfies the inequality. The second idea is to show that, if we iterate this decomposition often enough for the successive macroscopic component, then we obtain additional convexity properties, which allow us to prove that the macroscopic component satisfies the inequality we are looking for.

Our proof here follows the iterated two-scale approach, but uses several new ingredients :

- To deduce the modified LSI for the full measure from the inequality for the macroscopic measure, we use the L^1 Poincaré inequality to bound a crucial covariance term ;
- In addition to uniform convexity, we must prove uniform p -convexity for the macroscopic Hamiltonian, as soon as we have coarse-grained the system often enough ;
- We use the Prekopa-Leindler inequality to show that, if the single-site potential satisfies assumption (5.4), then the coarse-grained potential also does.

It was shown in [OV00] (and then in [BGL01] and [Goz09] with alternative proofs) that the classical logarithmic Sobolev inequality implies that the square root of the entropy controls the Wasserstein distance of order two (up to a multiplicative constant). Such an inequality is known as Talagrand’s inequality. Similarly, we can define a class of inequalities which generalizes the Talagrand inequality to Wasserstein distances of order p , which is linked to the modified logarithmic Sobolev inequality we just defined.

Definition 5.6. A probability measure μ satisfies a Talagrand inequality with parameter p and constant ρ if, for any probability measure ν , we have

$$W_p^p(\mu, \nu) \leq \frac{p}{\rho} \text{Ent}_\mu(\nu).$$

We will denote this inequality by $T_p(\rho)$.

Remark 5.7. Some people define $T_p(\rho)$ as $W_p(\mu, \nu) \leq \sqrt{\frac{2}{\rho} \text{Ent}_\mu(\nu)}$. These two definitions are **not** equivalent.

It was shown by Marton in [Mar96] that transport-entropy inequalities such as Talagrand inequalities imply concentration properties. These inequalities are also linked to modified logarithmic Sobolev inequalities through the following result, which was proven in [GRS13] :

Proposition 5.8. *If μ satisfies p -LSI(ρ), then it satisfies $T_p(\tilde{\rho})$, with constant $\tilde{\rho} = ((p - 1)\rho)^{p-1}$ and the ℓ^p distance.*

Combining this Proposition and Theorem 5.5, we obtain

Theorem 5.9. *Under the assumption (5.4), the canonical ensemble $\mu_{N,m}$ satisfies $T_p(\tilde{\rho})$ for some constant $\tilde{\rho} > 0$ that is independent of N and m .*

In section 3, an application of these modified LSI is presented, to obtain rates of convergence in the Wasserstein distance W_p for the Kawasaki dynamic on the Ginzburg-Landau model.

5.3 The iterated two-scale approach for modified logarithmic Sobolev inequalities

In this section, we shall prove Theorem 5.5.

Démonstration. Assume $N = 2^K$ for some large $K \in \mathbb{N}$. We define $P : X_{N,m} \rightarrow X_{N/2,m}$ by

$$P(x_1, x_2, \dots, x_N) := \left(\frac{x_1 + x_2}{2}, \frac{x_3 + x_4}{2}, \dots, \frac{x_{N-1} + x_N}{2} \right). \quad (5.8)$$

Using this operator, we can decompose $\mu_{N,m}$ as

$$\mu_{N,m}(dx) = \mu(dx|y)\bar{\mu}(dy)$$

where $\bar{\mu}$ is the push forward of μ under P and $\mu(dx|y)$ is the conditional measure of x given $Px = y$.

Proposition 5.10. *If $\bar{\mu}$ satisfies p -LSI(ρ) with ρ independent of N and m , then $\mu_{N,m}$ satisfies p -LSI($\tilde{\rho}$) with $\tilde{\rho}$ also independent of N and m .*

Démonstration. First we use the decomposition

$$\text{Ent}_\mu(f) = \text{Ent}_{\bar{\mu}}(\bar{f}) + \int \text{Ent}_{\mu(\cdot|y)}(f)\bar{\mu}(dy), \quad (5.9)$$

which can easily be verified through conditioning. We will then bound the two terms on the right-hand side of (5.9) by using modified LSI for the measures $\mu(dx|y)$ and $\bar{\mu}$.

Lemma 5.11. *There exists $\lambda > 0$ such that $\mu(dx|y)$ satisfies p-LSI(λ) for all $y \in Y$.*

Démonstration. Since $\mu(dx|y) = \otimes \mu_{2,y_i}(dx_{2i-1}, dx_{2i})$, by the tensorization property (see Proposition 5.24), we just have to show that $\mu_{2,m}$ satisfies p-LSI(λ) for some $\lambda > 0$ which does not depend on the real number m .

We have

$$\begin{aligned} \mu_{2,m}(dx_1, dx_2) &= \frac{1}{Z} \mathbb{1}_{x_1+x_2=2m} \exp(-\psi(x_1) - \psi(x_2)) dx \\ &= \frac{1}{Z} \mathbb{1}_{x_1+x_2=2m} \exp(-\psi_c(x_1) - \psi_c(x_2) - \delta\psi(x_1) - \delta\psi(x_2)) dx \end{aligned}$$

We immediately see that $(x_1, x_2) \rightarrow \psi_c(x_1) + \psi_c(x_2)$ is uniformly p-convex, so an application of Theorem 5.20 yields that the measure $\tilde{\mu}(dx) = Z^{-1} \mathbb{1}_{x_1+x_2=2m} \exp(-\psi_c(x_1) - \psi_c(x_2)) dx$ satisfies p-LSI($\tilde{\lambda}$) for some $\tilde{\lambda} > 0$ which doesn't depend on m . Since $\delta\psi$ is bounded, $\mu_{2,m}$ is a bounded perturbation of $\tilde{\mu}$, and we immediately deduce from Proposition 5.25 that it satisfies p-LSI(λ) for some $\lambda > 0$ which does not depend on m . \square

We therefore have

$$\begin{aligned} \int \text{Ent}_{\mu(\cdot|y)}(f)\bar{\mu}(dy) &\leq \int_Y \frac{1}{\lambda} \int_{\{Px=y\}} \frac{|(id_X - 2P^t P)\nabla f|_q^q}{f^{q-1}} \mu(dx|y)\bar{\mu}(dy) \\ &= \frac{1}{\lambda} \int_X \frac{|(id_X - 2P^t P)\nabla f|_q^q}{f^{q-1}} \mu(dx). \end{aligned} \quad (5.10)$$

By assumption, $\bar{\mu}$ satisfies p-LSI(ρ), so that

$$\text{Ent}_{\bar{\mu}}(\bar{f}) \leq \frac{1}{\rho} \int_Y \frac{|\nabla_Y \bar{f}|_q^q}{\bar{f}^{q-1}} \bar{\mu}(dy) \quad (5.11)$$

To deduce from this inequality a bound on the macroscopic entropy by a function of the microscopic gradient, we need to relate $\nabla_Y \bar{f}$ and ∇f . This is the point of the following lemma :

Lemma 5.12.

$$\nabla_Y \bar{f}(y) = 2P \int \nabla f(x)\mu(dx|y) + 2P \text{cov}_{\mu(dx|y)}(f, \nabla H).$$

This lemma was already used for the same reasons in [GOVW09] and [MO13]. For now, we defer its proof. Using this result, the convexity of $(x, b) \rightarrow \|x\|_q^q/b^{q-1}$ and the inequality $|a + b|^q \leq C(q)(|a|^q + |b|^q)$, we get

$$\begin{aligned} \text{Ent}_{\bar{\mu}}(\bar{f}) &\leq \frac{1}{\rho} \int \frac{|\nabla_Y \bar{f}|_q^q}{\bar{f}^{q-1}} \bar{\mu}(dy) \\ &= \frac{1}{\rho} \int \frac{\left| 2P \int \nabla f(x)\mu(dx|y) + 2P \text{cov}_{\mu(dx|y)}(f, \nabla H) \right|_q^q}{\left(\int f(x)\mu(dx|y) \right)^{q-1}} \bar{\mu}(dy) \\ &\leq \frac{C}{\rho} \int_X \frac{|2P \nabla f(x)|_q^q}{f^{q-1}} \mu(dx) + \frac{C}{\rho} \int \frac{|2P \text{cov}_{\mu(dx|y)}(f, \nabla H)|_q^q}{\bar{f}} \bar{\mu}(dy) \end{aligned} \quad (5.12)$$

We have

$$\begin{aligned}
|2P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|_q^q &= \sum_{i=1}^{N/2} \int |\operatorname{cov}_{\mu_{2,y_i}}(f, (2P\nabla H)_i)|^q \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \\
&= \sum_{i=1}^{N/2} \int |\operatorname{cov}_{\mu_{2,y_i}}(f, \psi'(x_{2i-1}) + \psi'(x_{2i}))|^q \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \\
&\leq C(q) \sum_{i=1}^{N/2} \int |\operatorname{cov}_{\mu_{2,y_i}}(f, \psi'_c(x_{2i-1}) + \psi'_c(x_{2i}))|^q \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \\
&\quad + C(q) \sum_{i=1}^{N/2} \int |\operatorname{cov}_{\mu_{2,y_i}}(f, \delta\psi'(x_{2i-1}) + \delta\psi'(x_{2i}))|^q \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \quad (5.13)
\end{aligned}$$

To bound the first part term, we use the following inequality, due to [MO13] :

Lemma 5.13 (Asymmetric Brascamp-Lieb inequality). *Let $\nu(dx) = \frac{1}{Z} \exp(-\psi(x))dx$ a probability measure on \mathbb{R} , where $\psi = \psi_c + \delta\psi$ is a bounded perturbation of a strictly convex potential. Then for any functions f and g , we have*

$$|\operatorname{cov}_{\nu}(f, g)| \leq \exp(-3 \operatorname{osc} \delta\psi) \sup_x \left| \frac{g'(x)}{\psi_c''(x)} \right| \int |f'| d\nu.$$

Using this lemma, we get

$$\begin{aligned}
&\int |\operatorname{cov}_{\mu_{2,y_i}}(f, \psi'_c(x_{2i-1}) + \psi'_c(x_{2i}))|^q \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \\
&\leq C \int \left(\int \left| \frac{df}{dx_{2i-1}} \right| + \left| \frac{df}{dx_{2i}} \right| \mu_{2,y_i}(dx_{2i-1}, dx_{2i}) \right)^q \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \\
&\leq \left(\int \left(\int f(x) \mu_{2,y_i}(dx_{2i-1}, dx_{2i}) \right)^{q/p} \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \right) \\
&\quad \times \left(\int \left(\int \frac{|\frac{df}{dx_{2i-1}}|^q + |\frac{df}{dx_{2i}}|^q}{f^{q-1}} \mu_{2,y_i}(dx_{2i-1}, dx_{2i}) \right) \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \right) \\
&\leq C \bar{f}(y)^{q-1} \left(\int \frac{|\frac{df}{dx_{2i-1}}|^q + |\frac{df}{dx_{2i}}|^q}{f^{q-1}} \mu(dx|y) \right), \quad (5.14)
\end{aligned}$$

where the last inequality uses the fact that $q/p = q - 1 \leq 1$, and therefore $a \rightarrow a^{q-1}$ is concave. Summing up, we obtain

$$\begin{aligned}
&\sum_{i=1}^{N/2} \int |\operatorname{cov}_{\mu_{2,y_i}}(f, \psi'_c(x_{2i-1}) + \psi'_c(x_{2i}))|^q \bigotimes_{j \neq i} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \\
&\leq C \bar{f}(y)^{q-1} \int \frac{|\nabla f|_q^q}{f^{q-1}} \mu(dx|y). \quad (5.15)
\end{aligned}$$

For the second part of (5.13), we use the following L^1 Poincaré inequality, which is Proposition 1.8 in [Led99] :

Theorem 5.14. *Consider a measure $\mu = \exp(-H)dx$ on \mathbb{R}^d , and assume that H is a bounded perturbation of a uniformly convex potential. Then there exists a constant $\alpha > 0$ such that, for any smooth function f , we have*

$$\int \left| f(x) - \int f(y)\mu(dy) \right| \mu(dx) \leq \alpha \int |\nabla f(x)|\mu(dx).$$

Since $\delta\psi'$ is bounded, we have

$$\begin{aligned} & |\text{cov}_{\mu_{2,y_i}}(f, \delta\psi'(x_{2i-1}) + \delta\psi'(x_{2i}))|^q \\ & \leq (2\|\delta\psi'\|_\infty)^q \left(\int \left| f(x) - \int f d\mu_{2,y_i} \right| \mu_{2,y_i}(dx_{2i-1}, dx_{2i}) \right)^q \\ & \leq C \left(\int \left| \frac{df}{dx_{2i-1}} \right| + \left| \frac{df}{dx_{2i}} \right| \mu_{2,y_i}(dx_{2i-1}, dx_{2i}) \right)^q \\ & \leq C \left(\int f(x)\mu_{2,y_i}(dx_{2i-1}, dx_{2i}) \right)^{q-1} \int \frac{\left| \frac{df}{dx_{2i-1}} \right|^q + \left| \frac{df}{dx_{2i}} \right|^q}{f(x)^{q-1}} \mu_{2,y_i}(dx_{2i-1}, dx_{2i}), \end{aligned} \quad (5.16)$$

where we have used Theorem 5.14 and the convexity of the function $(a, b) \rightarrow a^q/b^{q-1}$.

With the previous two bounds, we get

$$\int \frac{|2P \text{cov}_{\mu(dx|y)}(f, \nabla H)|_q^q}{\bar{f}} \bar{\mu}(dy) \leq C \int \frac{|\nabla f|_q^q}{f^{q-1}} \mu(dx). \quad (5.17)$$

We then state the elementary inequalities

$$|2Px|_q^q = \sum_i |x_{2i-1} + x_{2i}|^q \leq C(q) \sum_j |x_j|^q = C(q)|x|_q^q$$

and

$$|(id - 2P^tP)x|_q^q = \sum_i \left| \frac{x_{2i-1} - x_{2i}}{2} \right|^q + \left| \frac{x_{2i} - x_{2i-1}}{2} \right|^q \leq \frac{C(q)}{2^q} |x|_q^q.$$

Using these bounds, (5.10), (5.12) and (5.17), we get the result. \square

Before using this result to prove our main Theorem, here is a short proof of Lemma 5.12, which is taken from [GOVW09].

Proof of Lemma 5.12. Recall that

$$\begin{aligned} \bar{f}(y) &= \int_{\{Px=y\}} f(x)\mu(dx|y) \\ &= \frac{1}{\int_{\{Px=0\}} \exp(-H(2P^ty + z))dz} \int_{\{Px=0\}} f(2P^ty + z) \exp(-H(2P^ty + z))dz, \end{aligned}$$

and therefore, for any $\tilde{y} \in Y$, we have

$$\begin{aligned} \nabla_Y \bar{f}(y) \cdot \tilde{y} &= 2 \int \nabla f(x) \cdot P^t \tilde{y} \mu(dx|y) - 2 \int f(x) \nabla H(x) \cdot P^t \tilde{y} \mu(dx|y) \\ &\quad - 2 \left(\int f(x)\mu(dx|y) \right) \left(\int -H(x) \cdot P^t \tilde{y} \mu(dx|y) \right) \\ &= 2 \int P \nabla f(x) \cdot \tilde{y} \mu(dx|y) - 2 \int f(x) P \nabla H(x) \cdot \tilde{y} \mu(dx|y) \\ &\quad + 2 \left(\int f(x)\mu(dx|y) \right) \left(\int PH(x) \cdot \tilde{y} \mu(dx|y) \right), \end{aligned}$$

which is what we wanted to prove. \square

To deduce Theorem 5.5 from Proposition 5.10, we still need to show that $\bar{\mu}$ still has the structure $\exp(-\sum \tilde{\psi}(y_i))$ with $\tilde{\psi}$ a bounded perturbation of a p-convex and uniformly convex function, and that after enough iterations, $\bar{\mu}$ satisfies p-LSI with constant independent of the dimension. To do this, let's look at the structure of $\bar{\mu}$. We have

$$\bar{\mu}(dy) = \frac{1}{Z} \exp \left(-2 \sum_{i=1}^{N/2} R\psi(y_i) \right) dy$$

where

$$R\psi(y) := -\frac{1}{2} \log \left(\int_{\mathbb{R}} \exp(-\psi(x+y) - \psi(-x+y)) dy \right) \quad (5.18)$$

is the renormalized single-site potential. We denote by $R^M\psi$ the M-times renormalized single-site potential.

Lemma 5.15. *If $\psi = \psi_c + \delta\psi$ is a bounded perturbation of a p-convex, uniformly convex potential, then $R\psi$ also is.*

Démonstration. We define

$$\bar{\psi}_c(m) := -\frac{1}{2} \log \int \exp(-\psi_c(m+x) - \psi_c(m-x)) dx$$

and

$$\begin{aligned} \delta\bar{\psi}(m) &:= -\frac{1}{2} \log \int_{\mathbb{R}} \exp(-\psi(m+x) - \psi(m-x)) dx \\ &\quad + \frac{1}{2} \log \int \exp(-\psi_c(m+x) - \psi_c(m-x)) dx. \end{aligned} \quad (5.19)$$

Our aim is to show that $\delta\bar{\psi}$ is bounded in the C^1 topology, and that $\bar{\psi}_c$ is uniformly convex and p-convex. Since $R\psi = \bar{\psi}_c + \delta\bar{\psi}$, this will show that $\bar{\mu}$ has the desired structure.

The fact that $\bar{\psi}_c$ is uniformly convex has been proven in [MO13], using the (symmetric) Brascamp-Lieb inequality. Here we also need to prove that $\bar{\psi}_c$ is uniformly p-convex. To do this, we shall use the Prekopa-Leindler inequality, and the same method will also show that $\bar{\psi}_c$ is uniformly convex (which is not surprising, since the Prékopa-Leindler inequality is stronger than the Brascamp-Lieb inequality, as was shown in [BL00]).

Theorem 5.16. *Let $t \in (0, 1)$ and f, g, h be non-negative measurable functions defined on \mathbb{R} . Suppose that these functions satisfy*

$$h(tx + (1-t)y) \geq f(x)^t g(y)^{1-t}$$

for all x and y in \mathbb{R} . Then

$$\int h(x) dx \geq \left(\int f(x) dx \right)^t \left(\int g(x) dx \right)^{1-t}.$$

Let $h(x, m) = \exp(-\psi_c(x+m) - \psi_c(-x+m))$. We have for any $t \in (0, 1)$

$$\begin{aligned}
 & h(tx + (1-t)y, tm + (1-t)m') \\
 &= \exp(-\psi_c(tx + (1-t)y + tm + (1-t)m') - \psi_c(-tx - (1-t)y + tm + (1-t)m')) \\
 &\geq \exp(-t\psi_c(x+m) - t\psi_c(-x+m) - (1-t)\psi_c(y+m') - (1-t)\psi_c(-y+m')) \\
 &\quad \times \exp(ct(1-t)|m-m'+x-y|^p + ct(1-t)|m-m'+y-x|^p) \\
 &\geq \exp(-t(\psi_c(x+m) + \psi_c(-x+m)) - (1-t)(\psi_c(y+m') + \psi_c(-y+m')) + 2ct(1-t)|m-m'|^p) \\
 &= \exp(-\psi_c(x+m) - \psi_c(-x+m) + c(1-t)|m-m'|^p)^t \\
 &\quad \times \exp(-\psi_c(y+m') - \psi_c(-y+m') + ct|m-m'|^p)^{1-t} \tag{5.20}
 \end{aligned}$$

Applying the Prékopa-Leindler inequality with

$$h(x) = h(x, tm + (1-t)m'),$$

$$f(x) = \exp(-\psi_c(x+m) - \psi_c(-x+m) + c(1-t)|m-m'|^p)$$

and

$$g(x) = \exp(-\psi_c(y+m') - \psi_c(-y+m') + ct|m-m'|^p)$$

then yields

$$\begin{aligned}
 & \int_{\mathbb{R}} \exp(-\psi_c(x + tm + (1-t)m') - \psi_c(-x + tm + (1-t)m')) dx \\
 & \geq \left(\int_{\mathbb{R}} \exp(-\psi_c(x+m) - \psi_c(-x+m) + c(1-t)|m-m'|^p) dx \right)^t \\
 & \quad \times \left(\int_{\mathbb{R}} \exp(-\psi_c(y+m') - \psi_c(-y+m') + ct|m-m'|^p) dx \right)^{1-t} \tag{5.21}
 \end{aligned}$$

so that

$$\bar{\psi}_c(tm + (1-t)m') \leq t\bar{\psi}_c(m) + (1-t)\bar{\psi}_c(m') - ct(1-t)|m-m'|^p,$$

which is the inequality we were aiming for.

The same arguments, applied with $p = 2$ also show that $\bar{\psi}_c$ inherits uniform convexity from ψ_c .

We still need to prove bounds on $\delta\bar{\psi}$ and its first derivative. These were already proven in [MO13], we reproduce their argument here.

It will be convenient to introduce the probability measures

$$\nu(dx) = \frac{1}{Z} \exp(-\psi(-x+m) - \psi(x+m))dx$$

and

$$\nu_c(dx) = \frac{1}{Z} \exp(-\psi_c(-x+m) - \psi_c(x+m))dx,$$

so that we have

$$\delta\bar{\psi} = -\frac{1}{2} \log \int \exp(-\delta\psi(-x+m) - \delta\psi(x+m))\nu_c(dx)$$

and the bound $\|\delta\bar{\psi}\|_{\infty} < \infty$ immediately follows from $\|\delta\psi\|_{\infty} < \infty$.

A direct calculation yields

$$2\delta\bar{\psi}'(m) = \int (\psi'(-x+m) + \psi'(x+m))\nu(dx) - \int (\psi'_c(-x+m) + \psi'_c(x+m))\nu_c(dx).$$

We introduce the family of measures $(\nu^s)_{s \in [0,1]}$, defined by

$$\nu^s(dx) := \frac{1}{Z} \exp(-\psi_c(-x+m) - \psi_c(x+m) - s\delta\psi(-x+m) - s\delta\psi(x+m))dx.$$

This family interpolates between $\nu^0 = \nu_c$ and $\nu^1 = \nu$. By the mean-value theorem, there exists $s \in [0, 1]$ such that

$$\begin{aligned} 2\bar{\delta}\psi'(m) &= \frac{d}{ds} \int (\psi'_c(-x+m) + \psi'_c(x+m) + s\delta\psi'(-x+m) + s\delta\psi'(x+m))\nu^s(dx) \\ &= \int (\delta\psi(-x+m) + \delta\psi(x+m))\nu^s(dx) \\ &\quad - \text{cov}_{\nu^s}(\psi'_c(-x+m) + \psi'_c(x+m), \delta\psi(-x+m) + \delta\psi(x+m)) \\ &\quad - \text{cov}_{\nu^s}(s\delta\psi'(-x+m) + s\delta\psi'(x+m), \delta\psi(-x+m) + \delta\psi(x+m)) \end{aligned}$$

The first and third term on the right-hand side of this equation can be bounded uniformly in m by using the assumption that $\delta\psi$ and $\delta\psi'$ are bounded. For the second term, we also use these bounds, as well as the asymmetric Brascamp-Lieb inequality of Lemma 5.13 to show that

$$\begin{aligned} &\text{cov}_{\nu^s}(\psi'_c(-x+m) + \psi'_c(x+m), \delta\psi(-x+m) + \delta\psi(x+m)) \\ &\leq C \sup_x \left| \frac{\psi''_c(-x+m) - \psi''_c(x+m)}{\psi''_c(-x+m) + \psi''_c(x+m)} \right| \int |-\delta\psi'(-x+m) + \delta\psi'(x+m)|\nu^s(dx) \\ &\leq C, \end{aligned}$$

which finishes the proof of $\|\bar{\delta}\psi'\|_\infty < \infty$. □

The final step will be to show that, if we coarse-grain enough our system, the single-site potential becomes uniformly p -convex. Our proof will rely on a local Cramer theorem, which was already used in [MO13]. This convexification phenomenon is well-known in statistical physics, as a consequence of the equivalence of ensembles principle.

Lemma 5.17. *Let ψ be a bounded perturbation of a p -convex, uniformly convex potential. Then there is an integer M_0 such that for all $M \geq M_0$ the M -times renormalized single-site potential $R^M\psi$ is uniformly p -convex with constant ρ independent of the system size N , M and of the mean m .*

Démonstration. We define

$$\varphi(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \log \int_{\mathbb{R}} \exp(\sigma x - \psi(x))dx \right). \tag{5.22}$$

It is the Legendre transform of the function

$$\varphi^*(\sigma) := \log \int \exp(\sigma x - \psi(x))dx. \tag{5.23}$$

$$\mu^\sigma(dx) = \exp(\sigma x - \psi(x) - \varphi^*(\sigma))dx \tag{5.24}$$

Theorem 5.18 (Local Cramér theorem, Menz-Otto 2011). *Let*

$$\psi_K(m) := -\frac{1}{K} \log \left(\int_{X_{K,m}} \exp(-\sum \psi(x)) dx \right).$$

If ψ is a bounded perturbation of a uniformly convex potential, we have

$$|\psi_K''(m) - \varphi''(m)| \leq \frac{C}{K} \varphi''(m)$$

uniformly in $m \in \mathbb{R}$.

Since the proof of this result is quite long, we will not reproduce it here, and refer the interested reader to [MO13].

We apply this theorem, and since $R^M \psi = \psi_{2^M}$, for M large enough we have, uniformly in m ,

$$R^M \psi''(m) \geq \frac{1}{2} \varphi''(m).$$

Direct calculation on expression (5.22)

$$\varphi''(m) = \frac{1}{s(\sigma_m)^2},$$

where

$$s(\sigma)^2 = \int (x - m)^2 \mu^\sigma(dx),$$

$$\mu^\sigma(dx) := \frac{1}{Z} \exp(\sigma x - \psi(x)) dx,$$

and σ is the unique real number such that $\int x \mu^\sigma(dx) = m$.

The measures μ^σ satisfy a Poincaré inequality with constant independent of *sigma*, therefore we can show that $s(\sigma)^2$ is bounded above independently of σ :

$$s(\sigma)^2 \leq \frac{1}{\rho} \int |\nabla x|^2 \mu^\sigma(dx) = \frac{1}{\rho},$$

and the uniform convexity of $R^M \psi''$ follows.

To show that $R^M \psi$ is p-convex, it is therefore enough to show that

$$\varphi''(m) \geq C |m - m_0|^{p-2} \tag{5.25}$$

for some $C > 0$ and $m_0 \in \mathbb{R}$. Let

$$m_0 := \int x \mu^0(dx).$$

Since, by the usual properties of the Legendre transform, the real number σ_m such that $\varphi(m) = m\sigma_m - \varphi^*(\sigma_m)$ is given by $\varphi'(m) = \sigma_m$, we have $\varphi'(m_0) = 0$, and the unique minimum of φ is reached at m_0 . Since μ^0 satisfies p-LSI(ρ) for some $\rho > 0$ (to show this, use the p-convexity of ψ_c and the Holley-Stroock lemma), applying Proposition 5.26, we have

$$\frac{1}{\int \exp(-\psi(x)) dx} \int \exp(\sigma x - \psi(x)) dx \leq \exp \left(\sigma \int x \mu^0(dx) + \frac{|\sigma|^q}{\rho(q-1)} \right)$$

and therefore

$$\varphi^*(\sigma) \leq \varphi^*(0) + \sigma m_0 + \frac{|\sigma|^q}{\rho(q-1)}.$$

We then have

$$\begin{aligned}
 \varphi(m) &= \sup_{\sigma \in \mathbb{R}} (\sigma m - \varphi^*(\sigma)) \\
 &\geq \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \varphi^*(0) - \sigma m_0 - \frac{|\sigma|^q}{\rho(q-1)} \right) \\
 &= \varphi(m_0) + \sup_{\sigma \in \mathbb{R}} \left(\sigma(m - m_0) - \frac{|\sigma|^q}{\rho(q-1)} \right) \\
 &= \varphi(m_0) + c|m - m_0|^p
 \end{aligned} \tag{5.26}$$

where c is a positive constant which only depends on ρ and p . We then consider $f(m) = (m - m_0)\varphi'(m) - \varphi(m)$. Since φ'' is positive, f reaches its minimum at m_0 , so that for all $m \in \mathbb{R}$ we have $(m - m_0)\varphi'(m) - \varphi(m) \geq -\varphi(m_0)$, and therefore, using (5.26) and the fact that φ' is increasing, we get

$$|\varphi'(m)| \geq c|m - m_0|^{p-1}. \tag{5.27}$$

To study the behavior of φ'' , we shall now look at $\varphi^{(3)}$. An explicit calculation shows that

$$\begin{aligned}
 \varphi^{(3)}(m) &= \frac{d}{dm} \left(\int (x - m)^2 \mu^\sigma(dx) \right)^{-1} \\
 &= \frac{d\sigma}{dm} \times \frac{d}{d\sigma} \left(\int (x - m)^2 \mu^\sigma(dx) \right)^{-1} \\
 &= - \left(\int (x - m)^3 \mu^\sigma(dx) \right) \left(\int (x - m)^2 \mu^\sigma(dx) \right)^{-3}
 \end{aligned} \tag{5.28}$$

so that $\varphi^{(3)}(m) = 0$ iff $\int (x - m)^3 \mu^\sigma(dx) = 0$. But we have

$$\frac{d}{d\sigma} \int (x - m)^3 \mu^\sigma(dx) = \int (x - m)^4 \mu^\sigma(dx) > 0$$

so that $\int (x - m)^3 \mu^\sigma(dx)$ is a strictly increasing function, and cancels for at most one value of m . Therefore there exists some $m_1 \in \mathbb{R}$ such that $\varphi^{(3)}$ has constant sign on $(m_1, +\infty)$. Without loss of generality, we can assume $m_1 > m_0$. We consider two cases :

If $\varphi^{(3)}$ is non-negative on $(m_1, +\infty)$, then for any $\alpha \in [0, 1]$ the function $(m - m_0)\varphi''(m) - \alpha\varphi'(m)$ is increasing on $(m_1, +\infty)$. Moreover, since $m_1 > m_0$, $\varphi'(m_1) > 0$, and if we take $\alpha = \min(1, \frac{(m_1 - m_0)\varphi''(m_1)}{\varphi'(m_1)})$, this function is nonnegative at $m = m_1$. Therefore, for any $m \in (m_1, +\infty)$, we have

$$\begin{aligned}
 \varphi''(m) &\geq \alpha \frac{\varphi'(m)}{m - m_0} \\
 &\geq c|m - m_0|^{p-2}.
 \end{aligned} \tag{5.29}$$

If $\varphi^{(3)}$ is negative on $(m_1, +\infty)$, then φ'' is decreasing, and since it is bounded below by a positive constant, it converges to some positive constant $\lambda > 0$ in $+\infty$. We then have

$$\varphi'(m) = \int_{m_0}^m \varphi''(s) dx \underset{m \rightarrow +\infty}{\sim} \lambda m.$$

But since we know that $\varphi'(m) \geq c|m - m_0|^{p-1}$ with $p > 2$, this is a contradiction, so $\varphi^{(3)}$ must be non-negative on $(m_1, +\infty)$. Therefore we have

$$\varphi''(m) \geq c|m - m_0|^{p-2}$$

for all $m > m_1$. With the same reasoning, we can show that $\varphi''(m) \geq c|m - m_0|^{p-2}$ for all $m < m_2$ for some $m_2 < m_0$. But since φ'' is bounded below by a strictly positive constant, if we take C small enough, we also have $\varphi''(m) \geq c|m - m_0|^{p-2}$ for all $m \in [m_2, m_1]$, and therefore (5.25) holds. \square

Using the previous two lemmas, we can apply Proposition 5.10 a large but finite number of times to get our result. \square

5.4 An application to Kawasaki dynamics

There are many results on convergence to equilibrium in relative entropy for various dynamics in the literature. Theorem 5.9 says that, when we have such a convergence and if the invariant measure is the canonical ensemble $\mu_{N,m}$, then we also have convergence in the Wasserstein distance W_p . An example of such a dynamic with conservation law is given by the Kawasaki dynamic on \mathbb{R}^N :

$$dX_t = -A\nabla H(X_t)dt + \sqrt{2A}dB_t$$

where B_t is a Brownian motion on R^N and A is the discrete Laplacian on R^N , that is

$$A_{i,j} := 2\delta_{i,j} - \delta_{i,j+1} - \delta_{i,j-1}.$$

If we assume that the law of the initial value X_0 is absolutely continuous with respect to $\mu = \exp(-H)dx$, then the law f_t of X_t satisfies (in a weak sense) the PDE

$$\frac{\partial f_t}{\partial t} = \nabla \cdot (A\nabla f_t \mu).$$

Since this dynamic conserves the average $\sum x_i$, we restrict it to the hyperplane $\{\sum x_i = Nm\}$. It is a consequence of the LSI proved in [MO13] that, when $H(x) = \sum \psi(x_i)$ with ψ a bounded perturbation of a uniformly convex potential, the entropy satisfies the bound

$$\text{Ent}_\mu(f_t) \leq \exp(-\rho t/N^2) \text{Ent}_\mu(f_0),$$

and the order of magnitude t/N^2 is optimal. The following result is then an immediate consequence of this bound and our results :

Proposition 5.19. *Assume that f_t is the law of a solution of the Kawasaki dynamics with initial condition $f_0\mu$. Assume that the single-site potential satisfies (5.4). Then we have convergence to equilibrium for W_p , in the following sense :*

$$W_p^p(f_t\mu, \mu) \leq C \exp(-\rho t/N^2) \text{Ent}_\mu(f_0),$$

with constants C and ρ independent of the dimension N and the mean spin m , and the ℓ^p distance.

5.5 Appendix : Standard criteria for modified LSI

In this section, we state some standard criterion for a measure to satisfy a modified LSI. These criterion are the natural equivalents of the Bakry-Emery theorem, the tensorization principle and the Holley-Stroock lemma for the classical LSI.

Theorem 5.20. *Let V be a uniformly p -convex function with constant ρ on \mathbb{R}^N , that is for any $x, y \in \mathbb{R}^N$ and $t \in [0, 1]$, we have*

$$V(tx + (1-t)y) \leq tV(x) + (1-t)V(y) - \rho \frac{t(1-t)}{p} \|x - y\|_p^p.$$

Then the probability measure $\mu(dx) = \frac{1}{Z} \exp(-V(x))dx$ satisfies p -LSI($(\rho/q)^{q-1}$).

For a proof of this result, we refer to [BL00].

Example 5.21. $\mu(dx) = \exp(-\|x\|_p^p)dx$ satisfies p -LSI(c) for some $c > 0$.

Proposition 5.22. *If $V : \mathbb{R} \rightarrow \mathbb{R}$ satisfies $V''(x) \geq c(p-1)|x|^{p-2}$, then V is p -convex with constant c .*

Remark 5.23. This is not a necessary condition. $x \rightarrow (x-1)^4$ is 4-convex with constant 4, yet we do not have $12(x-1)^2 \geq 12x^2$.

Proposition 5.24. *If μ (resp. ν) is a probability measure on X_1 (resp. X_2) satisfying p -LSI(ρ_1) (resp. p -LSI(ρ_2)), then $\mu \otimes \nu$ satisfies p -LSI($\min(\rho_1, \rho_2)$).*

Démonstration. It is proven in the same way as for the usual LSI, by using the inequality

$$\text{Ent}_{\mu \otimes \nu}(f^q) \leq \int_{X_2} \text{Ent}_{\mu}(f(\cdot, x_2)^q) \nu(dx_2) + \int_{X_1} \text{Ent}_{\nu}(f(x_1, \cdot)^q) \mu(dx_1)$$

and applying the p -LSI for each measure. See for example [Led01], Proposition 5.6 for a proof of this inequality. \square

Proposition 5.25. *If μ satisfies p -LSI(ρ) and ψ is a bounded function, then the probability measure $\nu = \frac{1}{Z} \exp(\psi)d\mu$ satisfies p -LSI($e^{2 \text{osc}(\psi)}\rho$), where $\text{osc}(\psi) = \sup \psi - \inf \psi$.*

Démonstration. This is the analogue of the Holley-Stroock lemma for the usual LSI, and we can prove it in the same way. The identity (valid for any probability measure μ)

$$\text{Ent}_{\mu}(f) = \inf_{t \geq 0} \int_X f \log f - t \log t + (t-f)(1 + \log t) d\mu$$

implies that

$$\text{Ent}_{\nu}(f^q) \geq \exp(\text{osc}(\psi)) \text{Ent}_{\mu}(f^q).$$

It is also easy to show that

$$\int \|\nabla f\|_q^q d\mu \leq \exp(\text{osc}(\psi)) \int \|\nabla f\|_q^q d\nu,$$

so that, if μ satisfies p -LSI(ρ), then ν satisfies p -LSI($e^{2 \text{osc}(\psi)}\rho$). \square

Proposition 5.26. *If a probability measure μ on \mathbb{R}^n (equipped with the L^p norm) satisfies p -LSI(ρ), then for any 1-Lipschitz function f such that $\int f d\mu = 0$, we have $\int e^{\lambda f} d\mu \leq \exp\left(\frac{\lambda^q}{\rho(q-1)}\right)$ for all $\lambda \geq 0$.*

Démonstration. Let f be a smooth 1-Lipschitz function on X for the $\|\cdot\|_p$ norm, with mean 0, and

$$H(\lambda) := \int \exp(\lambda f - c\lambda^q \|f\|_{lip}^q) d\mu.$$

Then

$$\begin{aligned} \frac{d}{d\lambda} H(\lambda) &= \int (f - qc\lambda^{q-1} \|f\|_{lip}^q) \exp(\lambda f - c\lambda^q \|f\|_{lip}^q) d\mu \\ &= \frac{1}{\lambda} \int (\lambda f - cq\lambda^q \|f\|_{lip}^q) \exp(\lambda f - c\lambda^q \|f\|_{lip}^q) d\mu \\ &= \frac{1}{\lambda} \int (\lambda f - c\lambda^q \|f\|_{lip}^q) \exp(\lambda f - c\lambda^q \|f\|_{lip}^q) d\mu + \frac{c(1-q)^{q-1}}{\lambda} \|f\|_{lip}^q H(\lambda) \\ &= \frac{1}{\lambda} \text{Ent}_\mu(\exp(f - c\lambda^q \|f\|_{lip}^q)) + \frac{c(1-q)^{q-1}}{\lambda} \|f\|_{lip}^q H(\lambda) \end{aligned}$$

We can use the assumption that μ satisfies the p-LSI with parameter ρ under the form (5.2) to bound the entropy term, and we obtain

$$\frac{d}{d\lambda} H(\lambda) \leq \frac{1}{\lambda\rho} \int \lambda^q \|\nabla f\|_q^q \exp(f - c\lambda^q \|f\|_{lip}^q) d\mu + \frac{c(1-q)^{q-1}}{\lambda} \|f\|_{lip}^q H(\lambda). \quad (5.30)$$

Since we assumed f to be 1-Lipschitz for the L^p norm, $\|\nabla f\|_q \leq \|f\|_{lip}$ almost everywhere, and therefore

$$\begin{aligned} \frac{d}{d\lambda} H(\lambda) &\leq \frac{1}{\lambda\rho} \int \lambda^q \|f\|_{lip}^q \exp(f - c\lambda^q \|f\|_{lip}^q) d\mu + \frac{c(1-q)^{q-1}}{\lambda} \|f\|_{lip}^q H(\lambda) \\ &= \left(\frac{1}{\rho} + c(1-q) \right) \lambda^{q-1} \|f\|_{lip}^q H(\lambda). \end{aligned}$$

Taking $c = 1/\rho(q-1)$, we get $\frac{d}{d\lambda} H(\lambda) \leq 0$, therefore $H(\lambda) \leq H(0) = 1$ for all $\lambda \geq 0$, so that

$$\int \exp(\lambda f) d\mu \leq \exp\left(\frac{\lambda^q \|f\|_{lip}^q}{\rho(q-1)}\right)$$

for all $\lambda \geq 0$, which implies the desired result. \square

Acknowledgments : I would like to thank Emmanuel Boissard, Nathael Gozlan, Georg Menz and Cédric Villani for discussions on this topic. I would also like to thank the (anonymous) referee for pointing out a mistake in a previous version of the proof of Lemma 5.17.

Deuxième partie

Gradient flows and large
deviations

Chapitre 6

A gradient flow approach to large deviations

Les résultats de ce chapitre correspondent à l'article [Fat14].

Abstract : In this work, we investigate links between the formulation of the flow of marginals of reversible diffusion processes as gradient flows in the space of probability measures and path wise large deviation principles for sequences of such processes. An equivalence between the LDP principle and Gamma-convergence for a sequence of functionals appearing in the gradient flow formulation is proved. As an application, we study large deviations from the hydrodynamic limit for two variants of the Ginzburg-Landau model endowed with Kawasaki dynamics.

6.1 Introduction

In this work, we are interested in the links between the gradient flow formulation of the flow of marginals of stochastic differential equations, and path wise large deviations for sequences of such processes.

Interest in gradient flows on the space of probability measures goes back to [JKO98], where it was observed that the heat equation can be viewed as the gradient flow of the entropy

$$\text{Ent}(\rho) = \int \rho \log \rho dx$$

for the Wasserstein distance W_2 . Note that what we will call here entropy is the negative of the physical entropy. This was later developed into a notion of formal Riemannian structure on $\mathcal{P}(\mathbb{R}^n)$ by Otto in [Ott01]. While a powerful tool to predict the behavior of certain partial differential equations, the point of view of Otto is formal, and we must rely on other tools for proofs.

Another point of view was developed by Ambrosio, Gigli and Savaré in [AGS08], which uses the notion of "minimizing movement" schemes, developed by De Giorgi and which first appeared in [DGMT80], to provide a rigorous framework to define gradient flows on spaces of probability measures. It is based on the idea that gradient flows on \mathbb{R}^n of the form

$$\dot{x}(t) = -\nabla F(x(t))$$

are the only solutions of

$$F(x(T)) - F(x(0)) + \frac{1}{2} \int_0^T |\nabla F(x(t))|^2 dt + \frac{1}{2} \int_0^T |\dot{x}(t)|^2 dt = 0. \quad (6.1)$$

While the usual gradient flow equation only makes sense in a Riemannian setting (at least in the classical sense), this alternative formulation can be given a meaning in a purely metric setting, as long as we can define a "length of the gradient" functional $|\nabla F|$. Section 1.1 concerns this formulation in the setting of the space of probability measures on \mathbb{R}^n endowed with a Wasserstein distance W_2 , when the functional F is the relative entropy with respect to a nonnegative measure μ , that is

$$\text{Ent}_\mu(\nu) := \int f \log f d\mu$$

if $\nu = f\mu$, and $+\infty$ if ν is not absolutely continuous with respect to μ . This is the framework developed in the first sections of [AGS08].

Several recent papers have been interested in using abstract gradient flow formulations to study convergence of sequences of solutions to partial differential equations. One method, tailored for the case of diffusion processes and based on the discrete approximation of gradient flows, has been devised in [ASZ09]. Another, more general, method has been presented in [Ser11]. It consists in studying the asymptotic behavior of sequence of functionals of the form (6.1) for given functions F_n . Informally, it consists in showing that, if the sequence converges in a certain sense to a limiting functional of the same form, with a function F_∞ , we can directly identify limits of solutions of (6.1) as gradients flows for the limiting function F_∞ .

In the context of statistical physics, the method developed in [Ser11] can be used to prove convergence to the hydrodynamic limit for some models of interacting diffusion processes, such as the Ginzburg-Landau model (see [GPV88] or [GOVW09] for a presentation of the model, and its hydrodynamic limit). Such results consist in a convergence in probability of the dynamics to a deterministic limiting object, given in general as the solution to some partial differential equation.

Our aim here is to use the notion of gradient flows to study large deviations from the hydrodynamic limit for interacting spin systems. Such a result consists in proving that the probabilities of a significant deviation from the hydrodynamic limit decays exponentially fast in the system size. A standard textbook on the topic of large deviations is [DZ10], and [KL] contains a review of the literature in the context of large deviations from the hydrodynamic limit for particle systems.

In the recent series of contribution [ADPZ1], [ADPZ13] and [DLR13], links between gradient flows in spaces of probability measures and large deviations for many examples of processes arising in statistical physics have been investigated. The main contribution is to show that the gradient flow formulation for partial differential equations such as the heat equation can be deduced from the large deviation principle for N independent stochastic processes given by the stochastic differential equation whose flow of marginals is the solution to the PDE.

In this paper, we prove that process-level large deviations for sequences of diffusion processes are equivalent to the Gamma-convergence of a sequence of functionals that naturally appear in the gradient-flow formulation of these processes. This result generalizes a method used in [DG87] and [Föl88] to obtain process-level large deviations for the empirical measure of independent Brownian motions. Although these previous works do not discuss gradient flows or optimal transport, there are a lot of similarities between the formalism we use here and their framework, and the proof is based on a similar method.

As an application of this equivalence, we investigate the large deviations for two variants of the Ginzburg-Landau model endowed with Kawasaki dynamics, giving an alternative approach to obtaining the results of [DV89] and [Qua95]. The first model is a random conductance model, and the second one is the non-gradient Ginzburg-Landau model of [Var93] and [Qua95]. As far as the author knows, the large deviation principle for the random conductance model obtained here is new.

6.2 Framework and Method

6.2.1 Gradient flows in $P_2(\mathbb{R}^n)$

In this section, we endow \mathbb{R}^n with a Riemannian structure, given by a metric tensor $G(x)$, and a positive measure μ that is absolutely continuous with respect to the Lebesgue measure. We also consider the functional on the space of probability measures

$$\text{Ent}_\mu(\nu) := \int f \log f d\mu \quad (6.2)$$

when $\nu = f\mu$, and that takes value $+\infty$ for probability measures that are not absolutely continuous with respect to μ . Note that although we call this functional the entropy, it is the *negative* of the physical entropy. When $\mu = \exp(-H)$, this functional can be written as

$$\text{Ent}_\mu(\nu) = \text{Ent}_{dx}(\nu) + \int H(x)\nu(dx)$$

where $\text{Ent}_{dx}(\nu)$ is the relative entropy with respect to the Lebesgue measure. In the sequel, H will always denote a function such that $\mu = \exp(-H)$. We shall always assume H is at least C^1 .

We can endow the space of probability measures with finite second moments with the Wasserstein distance associated to the Riemannian metric structure $W_{2,G}$.

In the sequel, we will consider curves $(\nu_t)_{t \in [0, T]}$ in $\mathcal{P}_2(\mathbb{R}^d)$ (which is the space of probability measures with finite second moment) which are absolutely continuous. That is there exists a nonnegative function $\ell \in L^1([0, T])$ (which depends on $(\nu_t)_{t \in [0, T]}$) such that for any $s \leq t$ we have

$$W_{2,G}(\nu_s, \nu_t) \leq \int_s^t \ell(r) dr. \quad (6.3)$$

We will require the following technical assumptions on the metric tensor G :

$$\frac{1}{c}|\xi|^2 \leq \langle G(x)\xi, \xi \rangle \leq c|\xi|^2, \forall x \in \mathbb{R}^n, \forall \xi \in \mathbb{R}^n \quad (6.4)$$

for some constant c , and

$$x \longrightarrow \langle G(x)\xi, \xi \rangle \quad \text{is lower semicontinuous } \forall \xi \in \mathbb{R}^n. \quad (6.5)$$

Since we now have a metric structure on $P_2(\mathbb{R}^n)$, we can define the metric derivative of an absolutely continuous curve (ν_t) as

$$|\dot{\nu}|(t) := \limsup_{h \rightarrow 0^+} \frac{1}{h} W_{2,G}(\nu_t, \nu_{t+h}). \quad (6.6)$$

Definition 6.1. Let $A = G^{-1}$. We denote by $\|h\|_\nu$ the H^1 of the norm of the smooth function h , defined by

$$\|h\|_\nu^2 := \int \langle A\nabla h, \nabla h \rangle d\nu \quad (6.7)$$

and $\|\rho\|_{\nu,*}$ its dual norm, given by

$$\|\rho\|_{\nu,*}^2 := \sup_h 2 \int h \rho d\nu - \|h\|_\nu^2, \quad (6.8)$$

where the supremum runs over all smooth function $h : \mathbb{R}^n \rightarrow \mathbb{R}$.

Let g be defined by

$$g(\nu) := \left(\int \left\langle \frac{A\nabla\nu}{\nu} + A\nabla H, \frac{\nabla\nu}{\nu} + \nabla H \right\rangle d\nu \right)^{1/2} \quad (6.9)$$

if ν is absolutely continuous with respect to the Lebesgue measure, and $+\infty$ else wise. $\nabla\nu$ denote the gradient of the density of ν with respect to the Lebesgue measure.

The following result explains how the functional g can be used to control the variation in relative entropy for absolutely continuous curves in the space of probability measures. Its proof in this context can be found in [Lis09] (which generalizes previous results of [AGS08]).

Proposition 6.2. g is an upper gradient for Ent_μ , i.e. for every absolutely continuous curve $(\nu_t)_{0 \leq t \leq T}$ we have

$$|\text{Ent}_\mu(\nu_t) - \text{Ent}_\mu(\nu_s)| \leq \int_s^t g(\nu_r) |\dot{\nu}|(r) dr$$

for every $0 \leq s \leq t \leq T$.

Definition 6.3. Let $(\nu_t)_{t \in [0, T]}$ be a time-dependent family of measures that is absolutely continuous. We say it is a gradient flow of the functional Ent_μ if

$$\frac{d}{dt} \text{Ent}_\mu(\nu_t) = -\frac{1}{2} g(\nu_t)^2 - \frac{1}{2} |\dot{\nu}|(t)^2 \quad (6.10)$$

for almost every $t \in [0, T]$.

Gradient flows for the Wasserstein structure on $\mathcal{P}(\mathbb{R}^d)$ endowed with an Euclidean structure have been studied in [AGS08], and it turns out they are related to the heat equation. Their results were then generalized to the case of a Riemannian structure in [Lis09] :

Proposition 6.4. $(\nu_t)_{t \in [0, T]}$ satisfies (6.10) iff the densities $f(t, \cdot) = \frac{d\nu_t}{d\mu}$ form a weak solution of the parabolic PDE

$$\frac{\partial(f\mu)}{\partial t} = \nabla \cdot (A(\nabla f)\mu), \quad (6.11)$$

where $A = G^{-1}$.

As a consequence of this result and Ito's formula, we also have a representation of gradient flows as the flow of laws of the solution to a SDE :

Proposition 6.5. *If $(\nu_t)_t$ is a gradient flow of Ent_μ with $\mu = \exp(-H)dx$, then it is the flow of marginals of a solution of the SDE*

$$dX_t = -A(X_t)\nabla H(X_t)dt + \text{div}(A)(X_t)dt + \sqrt{2A} dB_t \quad (6.12)$$

with initial condition X_0 that has law ν_0 . Here, B is a standard Brownian motion on \mathbb{R}^d .

Remark 6.6. Diffusion processes that can be written in the form (6.12) are necessarily reversible.

We shall now give the definition of a key functional, which allows us to characterize gradient flows :

Proposition 6.7. *Let*

$$J((\nu_t)_t) := \text{Ent}_\mu(\nu_T) - \text{Ent}_\mu(\nu_0) + \frac{1}{2} \int_0^T g(\nu_t)^2 + |\dot{\nu}|(t)^2 dt. \quad (6.13)$$

Then $(\nu_t)_{t \in [0, T]}$ is a gradient flow of the functional H_μ iff $J((\nu_t)_t) = 0$.

Remark 6.8. In this setting, we have the following alternate formulation for the functional J :

$$J((\nu_t)_t) = \frac{1}{2} \int_0^T \|\dot{\nu} - \nabla \cdot (A(\nabla f)\mu)\|_{\nu_t, *}^2 dt, \quad (6.14)$$

at least for smooth functions. This formulation may seem more convenient, but in this context, (6.13) will be easier to manipulate.

Remark 6.9. We can make an interpretation of the notion of gradient flows in a statistical physics framework. It is a well-known principle in equilibrium statistical physics that steady states can be identified as minimizers of a thermodynamic functional, such as free energy, as a consequence of the second principle of thermodynamics. Here it is the relative entropy Ent_μ which plays the role of free energy, and indeed its minimizer is the equilibrium state μ . The gradient flow formulation identifies the correct trajectory as the minimizer of some action functional. This can be seen as an extension of the minimization principle to non-equilibrium statistical physics, with correct trajectories being those that decrease the free energy as fast as possible.

6.2.2 Relative entropy and large deviations

In this section, we introduce the notions of relative entropy and large deviations, and the links between the two.

Definition 6.10 (Relative entropy). Given two probability measures P and Q on a Polish space X , the relative entropy of P with respect to Q is given by

$$H(P; Q) := \sup_{f \in C_b(X)} \mathbb{E}_P(f) - \log \mathbb{E}_Q(e^f).$$

The following result is well-known, and is obtained by a computation of the Legendre transform (see Lemma 6.2.13 in [DZ10]).

Proposition 6.11. *We have*

$$H(P; Q) = \mathbb{E}_P \left[\ln \left(\frac{dP}{dQ} \right) \right]$$

if P is absolutely continuous with respect to Q , and $H(P; Q) = +\infty$ if not.

We now define large deviations :

Definition 6.12. Let I be a lower semicontinuous, nonnegative function on a Polish space X and $(a_n)_{n \in \mathbb{N}}$ a sequence of increasing, positive real numbers that goes to infinity. A sequence of probability measures \mathbb{P}_n on X is said to satisfy a large deviation principle with speed $(a_n)_n$ and good rate function I iff

- (i) For any closed set F , $\limsup a_n^{-1} \log \mathbb{P}_n(F) \leq -\inf_{x \in F} I(x)$;
- (ii) For any open set O , $\liminf a_n^{-1} \log \mathbb{P}_n(O) \geq -\inf_{x \in O} I(x)$.

Informally, this definition means that $\mathbb{P}_n(X_n \approx x) \approx \exp(-a_n I(x))$. We refer to the textbook [DZ10] for an introduction to large deviations.

In the recent contribution [Mar], Mariani proved the equivalence between large deviation principles for sequences of probability measures and the Gamma convergence (which we define below) of the associated relative entropy functionals.

Definition 6.13 (Gamma convergence). Let X be a space endowed with a notion of convergence. A sequence (I_n) of functionals on X is said to Γ -converge to a functional I at point $x \in X$ if the two following conditions are met :

- (i) For any sequence (x_n) that converges to x , we have $\liminf_{n \rightarrow \infty} I_n(x_n) \geq I(x)$;
- (ii) There exists a sequence (x_n) that converges to x such that $\lim_{n \rightarrow \infty} I_n(x_n) = I(x)$.

The sequence (I_n) is said to Γ -converge to I if it Γ -converges to I at every point.

Before we give a statement of Mariani's result, we define the notion of exponential tightness :

Definition 6.14 (Exponential tightness). A sequence of probability measures (μ_n) on a topological space X is said to be exponentially tight with speed (a_n) if, for any $\alpha > 0$, there exists a compact set K_α such that

$$\limsup_n \frac{1}{a_n} \log \mu_n(K_\alpha^c) \leq -\alpha.$$

Mariani's result can be stated as follows :

Theorem 6.15 ([Mar], 2012). *Let (μ_n) be a sequence of probability measures on a Polish space X , (a_n) a sequence of positive real numbers such that $\lim_n a_n = +\infty$ and $I : X \rightarrow [0, +\infty]$ a measurable, lower semicontinuous functional. We endow the space of probability measures with the topology of weak convergence.*

(i) *The sequence (μ_n) satisfies a large deviations upper bound with speed (a_n) and rate function I iff it is exponentially tight with speed $(a_n)_n$ and if for any sequence (ν_n) of probability measures on X that weakly converges to a Dirac measure δ_x , we have*

$$\liminf \frac{1}{a_n} H(\nu_n, \mu_n) \geq I(x);$$

(ii) *The sequence (μ_n) satisfies a large deviations upper bound with speed (a_n) and rate function I iff for any point x , there exists a sequence (ν_n) of probability measures on X that weakly converges to δ_x such that*

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} H(\nu_n, \mu_n) \leq I(x).$$

A heuristic explanation of Theorem 6.15 can be made in terms of the Bryc-Varadhan theorem (see sections 4.3 and 4.4 in [DZ10]). One can relate the relative entropy functional and exponential moments of functions by the relation

$$\int \exp(f) d\mu = \sup_{\nu \in \mathcal{P}(X)} \int f d\nu - \text{Ent}_\mu(\nu).$$

Since the Bryc-Varadhan lemma states that we can understand the large deviations for sequences of measures by looking at $\frac{1}{a_n} \log \int \exp(a_n f) d\mu_n$ for bounded continuous functions, the above relation translates the problem to investigating the behavior of the sequence of relative entropy functionals.

Remark 6.16. For the lower bound, it is enough to check the existence of a recovery sequence for every point y in a subset Y of X , such that given $x \in X$, there exists a sequence (y_k) of elements of Y that converges to x , and such that $I(y_k)$ converges to $I(x)$.

6.2.3 Relative entropy for the law of processes

Our first result is a relation between relative entropy with respect to the law of a solution of (6.12) and the functional J associated to the gradient flow formulation of the flow of marginals. It is a generalization of Theorem (1.31) in part II of [Föl88] (which dealt with the case of independent Brownian motions).

Theorem 6.17. *Let Q be the law of a solution to a SDE of type (6.12) on a space \mathbb{R}^d , and P the law of a process with finite relative entropy with respect to Q , with flow of marginals (ν_t) . Then :*

(i) *We have the lower bound*

$$H(P; Q) \geq H(P_0, Q_0) + \frac{1}{2} J((\nu_t))$$

where P_0 and Q_0 are the laws of the initial conditions.

(ii) *There exists a process with law \tilde{P} that has the same flow of marginals as P , such that*

$$H(\tilde{P}, Q) = H(P_0, Q_0) + \frac{1}{2} J((\nu_t)).$$

As a direct consequence of this relation and Theorem 6.15, we can use the functionals J to study large deviations.

We consider a sequence of diffusion processes of the form (6.12). The parameters we allow to vary are the drift ∇H , the diffusion coefficient A , and the dimension of the underlying space d_n .

To be able to state a large deviation principle for the laws of these diffusion processes, we need to embed their trajectories into a single space. We therefore implicitly assume that all the spaces \mathbb{R}^{d_n} have been embedded into a single metric space X . We then endow the space $\mathcal{C}([0, T], X)$ with some topology that makes it a metric, separable space. A typical choice would be the supremum norm.

Corollary 6.18. *Let Q_n be the law of a stochastic differential equation of the form 6.12, with $Q_{0,n}$ the law of the initial condition, and let J_n be the functional involved in the gradient flow formulation of the flow of marginals of Proposition 6.13. Then for any continuous trajectory $x \rightarrow x_t$, the sequence of normalized relative entropy functionals*

$a_n^{-1}H(\cdot, Q_n)$ Γ -converge at point δ_x iff the functionals $\frac{1}{a_n}(H(\cdot, Q_{0,n}) + \frac{1}{2}J_n(\cdot))$ also do, with the same Γ -limit.

As a consequence, solutions of gradients flows satisfy a large deviations principle with speed (a_n) and rate function I iff the functionals $\frac{1}{a_n}(H(\cdot, Q_{0,n}) + \frac{1}{2}J_n(\cdot))$ Γ -converge to I at the Dirac measure $\delta_{(x_t)}$, for every continuous trajectory $t \rightarrow x_t$, and if the sequence of laws is exponentially tight.

Of course, studying relative entropy to understand large deviations is a known technique (see [DG87] and [Föl88]). Our contribution is to show that instead of studying relative entropy (which depends on the law of the whole trajectory), we can study the functional J , which only depends on the flow of marginals, and is easier to manipulate, at least in some cases of interest, due to its connexion with optimal transport. It should be noted that some of the ideas we use here (relative entropy, variational formulations for rate functions) are reminiscent of those used in [DG87] to study large deviations for weakly-interacting mean-field models.

Remark 6.19. We expect the rate function I to be of the form $I(x_t) = I_0(x_0) + J(x_t)$, with J the function involved in the formulation (6.1) of a gradient flow in a certain metric space. This comes from the fact that often the rate function I will have a unique minimizer, which will be the deterministic limit of our sequence of processes. So we can reformulate the "correct" limit as the unique minimizer of a function. But gradient flow formulations also characterize some "correct" path as the unique minimizer of a functional. The similarity between these two point of views make us expect that they will be related, and this turns out to often be the case. See [ADPZ11] and [ADPZ13] for a study of this link in the case of sequences of independent processes.

Remark 6.20. An important element of the study of the functional J is the study of the Fischer information, or entropy-production functional. The importance of this functional can be understood in terms of statistical physics. It is a well-known principle in equilibrium statistical physics that equilibrium states can be obtained by optimizing some thermodynamic quantity, such as the free energy. This principle sometimes carries through to *non-equilibrium* statistical physics. Since the system seeks to increase the physical entropy (and therefore decrease the mathematical entropy), we can look at the entropy production functional, which we seek to optimize. Gamma-convergence corresponds to convergence of minimizers, so we can expect the "correct" trajectories to be those that, in the limit, make the entropy production functional as small as possible.

6.2.4 Some questions

- Is there a similar phenomenon for the large deviations of discrete systems, such as interacting particle systems? In the recent paper [Maa11], Maas showed that any reversible Markov chain on a finite space can be written as a gradient flow of a relative entropy for a well-chosen Riemannian structure on the space of probability measures. Can we exploit this structure to get the large deviations of systems such as a zero-range process, or exclusion processes?
- Many partial differential equations can be written as gradient flows for energy functionals which are not the relative entropy, such as porous medium equations (see [Ott01]). The energy production functional J still characterizes such gradient flows. Are there any nice properties implied by Gamma-convergence of the functional J for such systems?

- We apply in Section 3 this principle to get the large deviations for a system of diffusions with nearest neighbor interaction. It would also be interesting to look at mean-field models, where each diffusion interacts with all the others. In the case of smooth mean-field interactions, the question has been solved in [DG87], with a method that is very similar to the one we use here. A natural question is whether this extends to singular interactions. A case of interest is that of Coulomb interactions.
- Another natural question is whether we can use this principle in a context of modelization. Say we wish to approximate a phenomenon characterized as the unique solution to a partial differential equation of the form $\partial_t \rho = H(\rho, \nabla \rho, \dots)$ with a system of N interacting diffusion, with N large. If we can find a sequence of diffusion processes on \mathbb{R}^N and a sequence of positive numbers a_N such that $\frac{1}{a_N} J_N$ Gamma-converges to a lower semicontinuous functional that has the solution to the PDE as sole minimizer, then these diffusion processes form a good approximation. Can this idea be exploited in this context? This would be particularly interesting if we can extend our results to sequences of interacting particle systems.
- Our method works for reversible diffusion processes. Is there a similar method that works for non reversible processes, such as interacting diffusion processes with a boundary condition, or second-order diffusion processes? A method has been recently developed in [DPZ13] in what is called the GENERIC framework.

6.3 Proof of Theorem 6.17

The proof is a generalization of ideas coming from Part II of [Föl88]. It will consist in three steps : first we shall use Girsanov's theorem to give a representation of the law of processes that are absolutely continuous with respect to the law of the gradient flow. In a second step, we shall give a representation of the relative entropy of such a process, and finally we shall use this representation to obtain the lower bound of our theorem.

The following result is a direct application of Girsanov's theorem (see for example Theorem 2.4 in [Léo12]) :

Proposition 6.21. *Let Q be the law of the solution of (6.12) on $[0, T]$, and P the law of a process that is absolutely continuous with respect to Q . Then there exists an adapted process b_t valued in \mathbb{R}^d such that*

$$\frac{dP}{dQ}((X_t)_{0 \leq t \leq T}) = 1_{\frac{dP}{dQ} > 0} \frac{dP_0}{dQ_0}(X_0) \exp \left(\int_0^T b_t \cdot \sqrt{2A(X_t)} dB_t^P + \int_0^T \langle A(X_t) b_t, b_t \rangle dt \right). \quad (6.15)$$

In this equation, B^P is a P -Brownian motion, that is a local martingale under P which P -almost surely has quadratic variation equal to t . Moreover, P can be viewed as the law of a solution to the SDE

$$dX_t = A(X_t)(2b_t - \nabla H(X_t))dt + \operatorname{div}(A)(X_t)dt + \sqrt{2A(X_t)}dB_t^P \quad (6.16)$$

As a consequence, the relative entropy is given by

$$H(P; Q) = H(P_0, Q_0) + \mathbb{E}_P \left[\int_0^T \langle A(X_t) b_t, b_t \rangle dt \right]. \quad (6.17)$$

Note that the relative entropy doesn't only depend on the flow of marginals P_t , but on the law of the whole trajectory, unlike the functional J .

Lemma 6.22 (Markov version of the process). *Let $b_t = b(t, (X_s)_{0 \leq s \leq t})$ be the adapted process associated to a law P . Define*

$$\tilde{b}_t(x) := \mathbb{E}_P[b(t, (X_s)_{0 \leq s \leq t}) | X_t = x]. \quad (6.18)$$

Then the process defined by

$$dX_t = A(X_t)(2\tilde{b}_t(X_t) - \nabla H(X_t))dt + \operatorname{div}(A)(X_t)dt + \sqrt{2A(X_t)}dB_t \quad (6.19)$$

with initial condition $X_0 \sim P_0$ is a Markov process, and its law has the same flow of marginals as P .

Démonstration. The fact that this process is a Markov process is a classic result on SDEs, so we shall concentrate on proving that both processes have same marginals. Let g be a smooth function, X_t be a solution of (6.16) and \tilde{X}_t a solution of (6.19). We have

$$\begin{aligned} \mathbb{E}[g(\tilde{X}_t)] &= \int_0^t \mathbb{E} \left[A(\tilde{X}_s) \nabla g(\tilde{X}_s) \cdot (2\tilde{b}_s(X_s) - \nabla H(X_s)) \right] ds \\ &\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(\tilde{X}_s) \nabla g(\tilde{X}_s)) \right] ds \\ &= \int_0^t \mathbb{E} \left[A(\tilde{X}_s) \nabla g(\tilde{X}_s) \cdot (2\mathbb{E}_P[b(s, (X_r)_{0 \leq r \leq s}) | X_s] - \nabla H(X_s)) \right] ds \\ &\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(\tilde{X}_s) \nabla g(\tilde{X}_s)) \right] ds \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}[g(X_t)] &= \int_0^t \mathbb{E} \left[A(X_s) \nabla g(X_s) \cdot (2b(s, (X_r)_{0 \leq r \leq s}) - \nabla H(X_s)) \right] ds \\ &\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(X_s) \nabla g(X_s)) \right] ds \\ &= \int_0^t \mathbb{E} \left[A(X_s) \nabla g(X_s) \cdot (2\mathbb{E}_P[b(s, (X_r)_{0 \leq r \leq s}) | X_s] - \nabla H(X_s)) \right] ds \\ &\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(X_s) \nabla g(X_s)) \right] ds \\ &= \int_0^t \mathbb{E} \left[A(X_s) \nabla g(X_s) \cdot (2\tilde{b}_s(X_s) - \nabla H(X_s)) \right] ds \\ &\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(X_s) \nabla g(X_s)) \right] ds. \end{aligned}$$

This shows that the marginals of the laws of X and \tilde{X} satisfy the same parabolic PDE

$$\frac{\partial f}{\partial t} = \operatorname{div}(A \nabla f) + \operatorname{div}(A(2\tilde{b} - \nabla H)f).$$

Since they have the same initial condition, and since solutions to such PDEs are unique, they are the same. \square

Lemma 6.23. *Let \tilde{P} be the law of the solution of (6.18).*

We have

$$H(P; Q) \geq H(\tilde{P}; Q) \quad (6.20)$$

Démonstration. We already know that

$$H(\tilde{P}; Q) = H(P_0; Q_0) + \mathbb{E}_{\tilde{P}} \left[\int_0^T \langle A(X_t) \tilde{b}_t, \tilde{b}_t \rangle dt \right]. \quad (6.21)$$

An application of Jensen's inequality and the definition of \tilde{b} yields

$$\begin{aligned} H(P; Q) - H(P_0, Q_0) &= \mathbb{E}_P \left[\int_0^T \langle A(X_t) b_t, b_t \rangle dt \right] \\ &= \int_0^T \mathbb{E}_P [\langle A(X_t) b_t, b_t \rangle] dt \\ &= \int_0^T \mathbb{E}_{P_t} [\mathbb{E}_P [\langle A(X_t) b_t, b_t \rangle | X_t]] dt \\ &\geq \int_0^T \mathbb{E}_{P_t} [\langle A(X_t) \tilde{b}(X_t), \tilde{b}(X_t) \rangle] dt \\ &= H(\tilde{P}, Q) - H(P_0, Q_0), \end{aligned} \quad (6.22)$$

which is the desired lower bound. \square

Lemma 6.24 (Entropy of the Markov process). *The entropy of the Markov version of the process satisfies*

$$H(\tilde{P}; Q) = H(\nu_0, Q_0) + \frac{1}{2} J((\nu_t)_t)$$

where ν_t is the flow of marginals of the process P .

Démonstration. Let g be a smooth, compactly supported function. Itô's formula applied to the SDE (6.19) yields

$$\begin{aligned} \mathbb{E}[g(X_t)] &= \mathbb{E}[g(X_0)] + \int_0^t \mathbb{E}[A \nabla g(X_s) \cdot (2\tilde{b}_s(X_s) - \nabla H(X_s))] ds \\ &\quad + \int_0^t \mathbb{E}[(\nabla \cdot A(X_s) \nabla g(X_s))] ds \end{aligned} \quad (6.23)$$

It is easy to deduce from the Ito formulation (6.23) that the flow of marginals ν_t solves (in a weak sense) the PDE

$$\dot{\nu}_t = -\operatorname{div} \left(2A\tilde{b}_t \nu_t - \nu_t A \nabla H - A \nabla \nu_t \right). \quad (6.24)$$

Therefore the variation of the entropy of the marginals is given by

$$\begin{aligned} \operatorname{Ent}_\mu(\nu_T) - \operatorname{Ent}_\mu(\nu_0) &= \int_0^T \int A \nabla \nu_t(x) \cdot (2\tilde{b}_t(x) - \nabla H(x)) dx dt \\ &\quad - \int_0^T \int \frac{\langle A(x) \nabla \nu_t(x), \nabla \nu_t(x) \rangle}{\nu_t(x)} dx dt \\ &\quad + \int_0^T \int A(2\tilde{b}_t - \nabla H) \cdot \nabla H d\nu_t dt - \int_0^T \int A \nabla H \cdot \nabla \nu_t dx dt \\ &= \int_0^T \int 2A\tilde{b}_t \cdot \nabla \nu_t + \nu_t \nabla H dx dt - \int \frac{A(\nabla \nu_t + \nu_t \nabla H) \cdot (\nabla \nu_t + \nu_t \nabla H)}{\nu_t} dx \end{aligned} \quad (6.25)$$

From the Benamou-Brenier formula for W_2 (see for example [Vil03]) :

$$W_{2,G}^2(\nu_0, \nu_1) = \inf \left\{ \int_0^1 \int \langle Gv, v \rangle d\nu_t dt; \quad \dot{\nu} + \operatorname{div}(v\nu_t) = 0 \right\}$$

we deduce

$$\begin{aligned} & \frac{1}{2} \int_0^T |\dot{\nu}_t|^2 dt \\ &= \frac{1}{2} \int_0^T \left\langle A \left(2\tilde{b}_t - \nabla H - \frac{\nabla \nu_t}{\nu_t} \right), \left(2\tilde{b}_t - \nabla H + \frac{\nabla \nu_t}{\nu_t} \right) \right\rangle d\nu_t dt \\ &= 2 \int_0^T \int \langle A\tilde{b}_t, \tilde{b}_t \rangle d\nu_t dt + \frac{1}{2} \int_0^T \int \left\langle A \left(\nabla H + \frac{\nabla \nu_t}{\nu_t} \right), \left(\nabla H + \frac{\nabla \nu_t}{\nu_t} \right) \right\rangle d\nu_t dt \\ &\quad - \int_0^T \int \left\langle 2A\tilde{b}_t, \left(\nabla H + \frac{\nabla \nu_t}{\nu_t} \right) \right\rangle d\nu_t dt. \end{aligned} \tag{6.26}$$

By the definition (6.9) of the upper gradient g , we have

$$\frac{1}{2} \int_0^T g(\nu_t)^2 dt = \frac{1}{2} \int_0^T \int \left\langle \frac{A\nabla \nu_t}{\nu_t} + A\nabla H, \frac{\nabla \nu_t}{\nu_t} + \nabla H \right\rangle d\nu_t dt. \tag{6.27}$$

Combining (6.25), (6.26) and (6.27), we get

$$J((\nu_t)_t) = 2 \int_0^T \int \langle A\tilde{b}_t, \tilde{b}_t \rangle d\nu_t dt, \tag{6.28}$$

and then the lemma immediately follows from (6.21). \square

To deduce Corollary 6.18 from Theorem 6.17, the only thing we still have to prove is that, if (ν_t) is a flow of marginals such that $J((\nu_t))$ is finite, there exists a process whose law is absolutely continuous with respect to Q , and with flow of marginals (ν_t) .

Let (ν_t) be an absolutely continuous flow of marginals such that $J((\nu_t))$ is finite. From [L, Theorem 2.4], we know that there exists a vector field (v_t) such that the continuity equation

$$\dot{\nu}_t = \nabla \cdot (Av_t\nu_t) \tag{6.29}$$

is satisfied. On the other hand, we know that (ν_t) is the flow of marginals of the solution to an SDE of type

$$dX_t = 2A(X_t)b_t(X_t)dt + \sqrt{2A(X_t)}dB_t,$$

whose law would then be absolutely continuous with respect to Q , if the flow solves in a weak sense the PDE

$$\dot{\nu}_t = \operatorname{div}(A(\nabla \nu_t + 2b_t\nu_t)).$$

Since $J((\nu_t))$ is finite, the upper gradient $g(\nu_t)$ is finite for almost every t , and $\nabla \nu_t$ exists. We therefore only have to take $2b_t(x) = v_t(x) - \frac{\nabla \nu_t}{\nu_t}(x)$ to see that the flow solves the above PDE.

6.4 Large deviations for the Ginzburg-Landau model

6.4.1 The model

The (classical) Ginzburg-Landau model equipped with Kawasaki dynamics is a sequence of N diffusions, interacting according to the SDE

$$dX_t^i = N^2(\psi(X_t^{i+1}) + \psi(X_t^{i-1}) - 2\psi(X_t^i))dt + \sqrt{2}N(dB_t^{i+1} - dB_t^i).$$

This diffusion is not ergodic on the whole space \mathbb{R}^N , since it preserves the quantity $\sum X_t^i$, but it is ergodic when restricted to a hyperplane

$$X_{N,m} := \left\{ x \in \mathbb{R}^N; \quad \frac{1}{N} \sum_{i=1}^N x_i = m \right\}. \quad (6.30)$$

It then has an invariant measure

$$\mu(dx) := \frac{1}{Z} \exp\left(-\frac{1}{N} \sum_{i=1}^N \psi(x_i)\right) \mathbb{1}_{x \in X_{N,m}} \mathcal{L}^{N-1}(dx) \quad (6.31)$$

where \mathcal{L} is the Lebesgue measure on $X_{N,m}$. Equivalently, its law is given by the solution of the PDE

$$\frac{\partial f \mu}{\partial t} = \operatorname{div}(A_0 \nabla f \mu), \quad (6.32)$$

where f is the density with respect to μ , and A_0 is the discrete Laplacian scaled by N^2 , that is

$$(A_0)_{i,j} := N^2(\delta_{i,j+1} + \delta_{i,j-1} - 2\delta_{i,j}).$$

When N goes to infinity, if the initial condition behaves deterministically in the limit, the (properly rescaled) solutions concentrate around a deterministic profile, called the hydrodynamic limit, which has been studied in [GPV88]. Large deviations from this hydrodynamic limit have been studied in [DV89].

We shall investigate the large deviations as N goes to infinity for two versions of this model, with conductances. The first case will involve random conductances, and the second will involve conductances depending on the configuration of spins. Our method will rely on Corollary 6.18, and reduce the problem to the study of the behavior of the functional J_N associated with the gradient flow formulation of these dynamics.

For technical reasons, we shall assume that the initial data follows a local Gibbs state, that is

$$f_0(x) = \frac{1}{Z} \exp\left(\sum x_i \varphi'(\rho_0(i/N))\right) \quad (6.33)$$

for some continuous function ρ_0 . We will later see that this initial data concentrates around the deterministic profile ρ_0 . It can be shown that, for initial data that behaves deterministically in the limit, solutions at any positive time are close (in the sense of relative entropy) to such a local Gibbs state. See [Kos01a] or [Fat13b] for a proof.

We will also assume that the single-site potential ψ is of the form

$$\psi(x) = \frac{1}{p} x^p + \delta\psi(x) \quad (6.34)$$

for some $p \geq 0$ and a perturbation $\delta\psi$ that is C^2 , bounded and with bounded first and second derivative.

It is likely that the results hold for more general functions ψ , but such a result would require more general technical tools than those developed in the next section. For example, if ψ doesn't grow at least as $|x|^2$ as x goes to infinity, then the logarithmic Sobolev inequality doesn't hold. In [DV89], the LDP is proved for the case where ψ is only superquadratic, and $\psi' = o(\psi)$.

It turns out that the proof of the upper bound in the convergence of J_N is the same as the proof of [Qua95] of the lower bound for the LDP. We shall therefore only sketch the proofs of the upper bounds, and concentrate on the lower bounds in the Γ -convergence.

6.4.2 Some technical estimates

In this section, we give a few technical results, collected from various sources, which we shall use in the proofs of the large deviation principles. Most of them are classical results in the study of hydrodynamic limits, and we will often only give a brief sketch of the proofs, or simply refer to the original source.

We will use logarithmic Sobolev inequalities, which we now define :

Definition 6.25. Let X be a Riemannian manifold. A probability measure μ on X is said to satisfy a LSI with constant $\rho > 0$ if, for any locally Lipschitz, nonnegative function $f \in L^1(\mu)$,

$$\int f \log(f) d\mu - \left(\int f d\mu \right) \log \left(\int f d\mu \right) \leq \frac{1}{\rho} \int \frac{|\nabla f|^2}{2f} d\mu.$$

The following result was proven in [MO13] :

Theorem 6.26. Under the assumption (6.34), the measures μ_N satisfy the logarithmic Sobolev inequality

$$\text{Ent}_{\mu_N}(g) \leq C \int \frac{|\nabla g|^2}{g} d\mu_N$$

for any nonnegative, locally Lipschitz function g , with constant C independent of the dimension N and the mean spin m . Combined with the discrete Poincaré inequality, this implies

$$\text{Ent}_{\mu_N}(g) \leq C \int \frac{\langle A_0 \nabla g, \nabla g \rangle}{g} d\mu_N$$

for some constant C that is independent of the dimension and the mean spin.

As a consequence of this result and of [GOVW, Lemma 26], we have the following result :

Lemma 6.27. Let f_N be a sequence of probability densities with respect to μ_N such that

$$\sup_N \frac{1}{N} \int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_N < +\infty.$$

Then we also have

$$\sup_N \frac{1}{N} \int \sum |x_i|^2 f(x) \mu_N(dx) < +\infty.$$

This result still holds if we replace μ_N by another sequence of measures with bounded second moment and which satisfy a LSI with uniform constant.

This can be generalized to the following result :

Lemma 6.28. *Assume that ψ is of the form $\frac{1}{p}|x|^p + \delta\psi(x)$. Let f_N be a sequence of probability densities with respect to μ_N such that*

$$\sup_N \frac{1}{N} \int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_N < +\infty.$$

Then we also have

$$\sup_N \frac{1}{N} \int \sum |x_i|^p f(x) \mu_N(dx) < +\infty.$$

Démonstration. It has been shown in [Fat13a] that, under our assumptions on ψ , μ_N satisfies the following transport-entropy inequality : for any probability measure ν_N ,

$$W_p^p(\nu_N, \mu_N) \leq C \text{Ent}_{\mu_N}(\nu_N)$$

for some constant $C > 0$ that does not depend on N , and W_p is the L^p Wasserstein distance

$$W_p^p(\nu, \mu) := \inf_{\pi} \int \sum |x_i - y_i|^p \pi(dx, dy).$$

From the W_p -Lipschitz continuity of p-moments (see [Vi2, Proposition 7.29]), we know that

$$\left(\int \sum |x_i|^p \nu_N(dx) \right)^{1/p} - \left(\int \sum |x_i|^p \mu_N(dx) \right)^{1/p} \leq W_p(\mu_N, \nu_N)$$

so that

$$\int \sum |x_i|^p \nu_N(dx) \leq C \text{Ent}_{\mu_N}(\nu_N) + C \int \sum |x_i|^p \mu_N(dx)$$

Since μ_N satisfies a logarithmic Sobolev inequality, $\text{Ent}_{\mu_N}(\nu_N) \leq CN$, and it is also easy to see that

$$\int \sum |x_i|^p \mu_N(dx) \leq CN,$$

which concludes the proof. \square

We now give a version the version of the local Cramér theorem we shall use :

Theorem 6.29. *Let (a_i) be some sequence of real numbers. We define*

$$\psi_K(m) := -\frac{1}{K} \log \int_{X_{K,m}} \exp(\sum a_i x_i + \psi(x_i)) dx$$

and

$$\varphi_K(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \frac{1}{K} \sum_{i=1}^K \log \int_{\mathbb{R}} \exp((\sigma + a_i)x - \psi(x)) dx \right).$$

We then have , for any $L > 0$ and any compact subset E of \mathbb{R} ,

$$\lim_{K \rightarrow \infty} \sup_{a_1 \dots a_K \in [-L, L]} \|\psi_K - \varphi_K\|_{\infty, E} = 0.$$

In particular, if $a_i = \lambda(i/K)$ for some smooth function λ , then ψ_K converges to

$$\varphi_\lambda(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \int_0^1 \log \int_{\mathbb{R}} \exp((\sigma + \lambda(\theta))x - \psi(x)) dx d\theta \right).$$

uniformly on compact sets.

A proof of this result can be found in [FM14] or [K, Appendix A]. Roughly speaking, it says that local averages of a large number K of spins behave like random variables satisfy a large deviation principle of speed K and rate function φ .

The following proposition is a consequence of [GPV, Theorem 4.1].

Proposition 6.30. *Let f_N be a sequence of probability densities with respect to μ_N which weakly converges to a deterministic profile ρ . Assume that*

$$\frac{1}{N} \int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_N \leq C.$$

Then, for any smooth function $J : \mathbb{T} \rightarrow \mathbb{R}$ and bounded continuous function $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$, we have

$$\frac{1}{N} \int \sum J(i/N) F(x_{i-k}, \dots, x_{i+k}) f_N(x) \mu_N(dx) \longrightarrow \int_{\mathbb{T}} J(\theta) \tilde{F}(\rho(\theta)) d\theta$$

where

$$\tilde{F}(y) := \int F(x_1, \dots, x_{2k+1}) \mu^{\lambda, \otimes 2k+1}(dx), \quad (6.35)$$

with $\mu^\lambda(dx) = \frac{1}{Z} \exp(\lambda x - \psi(x)) dx$ and $\lambda = \varphi'(y)$.

Note that, in [GPV88], it was also required that there exists a superlinear function ω such that $\int \sum \omega(x_i) f(x) \mu_N(dx) \leq CN$. However, under our assumptions on ψ , the bound on $\int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_N$ implies that $\int \sum |x_i|^2 f(x) \mu_N(dx) \leq CN$, as we have seen in Lemma 6.27.

Proposition 6.31. *Let ρ be a smooth function on the torus, and define the probability density with respect to μ*

$$G_N(x) = \frac{1}{Z} \exp \left(\sum_{i=1}^N \varphi'(\rho(i/N)) x_i \right).$$

Then

- (i) *The measures $G_N \mu_N$ weakly converge to the deterministic profile ρ ;*
- (ii) *They satisfy a logarithmic Sobolev inequality, with a constant that only depends on ρ and ψ , but which is uniform in N ;*
- (iii) *For any sequence of probability measures ν_N on \mathbb{R}^N , if*

$$\frac{1}{N} \text{Ent}_{G_N \mu_N}(\nu_N) \longrightarrow 0,$$

then the sequence weakly converges to the deterministic profile ρ . Moreover, we then have

$$\frac{1}{N} \text{Ent}_{\mu_N}(\nu_N) \longrightarrow \int \varphi(\rho) d\theta - \varphi \left(\int \rho d\theta \right).$$

Démonstration. (i) is a classic large deviation result. See for example [Yau91]. (ii) was proven in [FM14]. (iii) is a consequence of these two results, and we can prove it as follows.

Since the measures $G_N \mu_N$ satisfy a logarithmic Sobolev inequality, they also satisfy a transport entropy inequality, that is

$$W_{2, A_0^{-1}}(\nu_N, G_N \mu_N)^2 \leq C \text{Ent}_{G_N \mu_N}(\nu_N).$$

The fact that we can use the inner product given by A_0 rather than the usual inner product follows from the discrete Poincaré inequality. Therefore, we have

$$\frac{1}{N} W_{2, A_0^{-1}}(\nu_N, G_N \mu_N)^2 \longrightarrow 0.$$

The result then follows from the fact that $(G_N \mu_N)$ weakly converges to ρ , and that $\frac{1}{N} \langle A_0 x, x \rangle \leq C \|\bar{x}\|_{H^{-1}}^2$.

The second part is a consequence of the identity

$$\frac{1}{N} \text{Ent}_{\mu_N}(\nu_N) = \frac{1}{N} \text{Ent}_{G_N \mu_N}(\nu_N) + \frac{1}{N} \int \log G_N d\nu_N$$

and the convergence

$$\begin{aligned} \frac{1}{N} \int \log G_N d\nu_N &= \frac{1}{N} \int \sum_{i=1}^N \varphi'(\rho_i) x_i \nu_N(dx) \\ &\quad - \frac{1}{N} \log \int \exp \left(\sum_{i=1}^N \varphi'(\rho(i/N)) x_i - \psi(x_i) \right) dx + \frac{1}{N} \log \int \exp \left(\sum_{i=1}^N -\psi(x_i) \right) dx \\ &\longrightarrow \int \varphi(\rho(\theta)) d\theta - \varphi \left(\int \rho(\theta) d\theta \right). \end{aligned} \quad (6.36)$$

This last convergence follows from the convergence of (ν_N) to the deterministic profile ρ and Theorem 6.29. A complete proof is given in Lemma 7.1 of [K]. \square

Proposition 6.32. *Let f_N be a sequence of probability densities with respect to μ_N which weakly converges to a deterministic profile ρ . Assume that*

$$\frac{1}{N} \int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_N \leq C$$

and

$$\frac{1}{N} \int \sum \omega(x_i) f_N(x) \mu_N(dx) \leq C.$$

Then, for any sequence (J^N) of step functions on the torus that are constant on the intervals $[(i-1)/N, i/N)$ and which converges in H^1 to a function J , we have

$$\frac{1}{N} \int \sum J^N(i/N) x_i f_N \mu_N(dx) \longrightarrow \int_{\mathbb{T}} J(\theta) \rho(\theta) d\theta \quad (6.37)$$

and

$$\frac{1}{N} \int \sum J^N(i/N) \psi'(x_i) f_N \mu_N(dx) \longrightarrow \int_{\mathbb{T}} J(\theta) \varphi'(\rho(\theta)) d\theta. \quad (6.38)$$

Démonstration. For the first part, notice that

$$\begin{aligned} &\left| \frac{1}{N} \int \sum J^N(i/N) x_i f_N \mu_N(dx) - \int \int_{\mathbb{T}} J(\theta) \bar{x}(\theta) d\theta f_N(x) \mu_N(dx) \right| \\ &\leq \|J^N - J\|_{H^1} \left(\int \|\bar{x}\|_{H^{-1}}^2 f_N(x) \mu_N(dx) \right)^{1/2} \longrightarrow 0 \end{aligned}$$

and

$$\begin{aligned} & \left| \int \int_{\mathbb{T}} J(\theta) \bar{x}(\theta) d\theta f_N(x) \mu_N(dx) - \int_{\mathbb{T}} J(\theta) \rho(\theta) d\theta \right| \\ & \leq \|J\|_{H^1} \left(\int \|\bar{x} - \rho\|_{H^{-1}} f_N(x) \mu_N(dx) \right) \\ & \quad \|J\|_{H^1} \left(\int \|\bar{x} - \rho\|_{L^2} f_N(x) \mu_N(dx) \right) \end{aligned}$$

so we just have to show that

$$\int \|\bar{x} - \rho\|_{L^2} f_N(x) \mu_N(dx) \longrightarrow 0.$$

This quantity is the Wasserstein distance W_1 between $f_N \mu_N$ and δ_ρ for the L^2 distance. Since we already have weak convergence, to show that there is convergence for W_1 , according to [Vi1, Theorem 7.12], we just have to prove the following tightness estimate

$$\lim_{R \rightarrow \infty} \limsup_{N \rightarrow \infty} \int_{\sum |x_i|^2 \geq NR^2} \sqrt{\frac{1}{N} \sum |x_i|^2} f_N(x) \mu_N(dx) = 0.$$

This estimate automatically follows from the bound

$$\sup_N \frac{1}{N} \int \sum |x_i|^2 f_N(x) \mu_N(dx) < +\infty$$

that was given by Lemma 6.27.

For the second part, we give a very brief sketch of the method of proof that was used in [GPV88]. Let ψ_ℓ be a cutoff of ψ' at level $\ell > 0$, that is

$$\psi_\ell(x) = \psi'(x) \text{ if } |\psi'(x)| \leq \ell, \psi_\ell(x) = \pm \ell \text{ if not.}$$

From the bound of Lemma 6.28, we can deduce

$$\frac{1}{N} \int \sum J^N(i/N) \psi'_\ell(x_i) f_N \mu_N(dx) \underset{\ell \rightarrow \infty}{\lim} \frac{1}{N} \int \sum J^N(i/N) \psi'(x_i) f_N \mu_N(dx) \quad (6.39)$$

uniformly in N , since ψ goes to infinity faster than $|\psi'|$. Moreover, from Proposition 6.30, we obtain

$$\frac{1}{N} \int \sum J^N(i/N) \psi'_\ell(x_i) f_N \mu_N(dx) \underset{N \rightarrow \infty}{\lim} \int_{\mathbb{T}} J(\theta) \tilde{\psi}_\ell(\rho(\theta)) d\theta, \quad (6.40)$$

so all we need to do is show that $\tilde{\psi}_\ell$ converges to φ' , which was done in [GPV, Lemma 6.4]. \square

Remark 6.33. Similarly, under the assumption that $\int_0^T \int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} f \mu_N dt \leq CN$ uniformly in N , then (6.37) and (6.38) hold in a time-integrated sense.

We also give a priori estimates on weak limits of sequences of probability measures, obtained as direct consequences of [GPV88], Lemmas 6.3 and 6.6 :

Lemma 6.34. *Under our assumptions on ψ , for any sequence of probability f_N with respect to μ_N that weakly converges to a deterministic trajectory ρ , such that*

$$\sup_N \frac{1}{N} \int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_N \leq C$$

we have

$$\int_{\mathbb{T}} \varphi(\rho(\theta)) d\theta \leq C$$

and

$$\int_{\mathbb{T}} (\partial_{\theta} \varphi(\rho)(\theta))^2 d\theta \leq C.$$

Finally, we shall need the following lower bound on the slope of absolutely continuous curves.

Lemma 6.35. *Let (ν_t) be an absolutely continuous curve of probability measures on \mathbb{R}^n , which is equipped with a Riemannian tensor $(A^{-1}(x))$ satisfying the assumptions (6.4) and (6.5). Then we have, for any smooth function $V : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$,*

$$\begin{aligned} \int |\dot{\nu}_t|^2 dt &\geq 2 \int V(T, x) \nu_T(dx) - 2 \int V(0, x) \nu_0(dx) \\ &\quad - 2 \int_0^T \int \frac{\partial V}{\partial t}(t, x) \nu_t(dx) dt - \int_0^T \int \langle A(x) \nabla V, \nabla V \rangle \nu_t(dx) dt \end{aligned} \quad (6.41)$$

Démonstration. From [L, Theorem 2.4], we know that there exists a vector field v_t such that

$$\dot{\nu}_t + \operatorname{div}(v_t \nu_t) = 0 \quad (6.42)$$

and

$$\int_0^T |\dot{\nu}_t|^2 dt = \int_0^T \int \langle A^{-1}(x) v_t(x), v_t(x) \rangle \nu_t(dx) dt. \quad (6.43)$$

Since we have

$$\langle A^{-1}(x) v_t(x), v_t(x) \rangle \geq 2 \langle v_t(x), \nabla V(t, x) \rangle - \langle A(x) \nabla V(t, x), \nabla V(t, x) \rangle$$

for any t and x , we get

$$\begin{aligned} \int_0^T |\dot{\nu}_t|^2 dt &\geq 2 \int_0^T \int \langle v_t(x), \nabla V(t, x) \rangle \nu_t(dx) dt \\ &\quad - \int_0^T \int \langle A(x) \nabla V(t, x), \nabla V(t, x) \rangle \nu_t(dx) dt \end{aligned}$$

Using (6.42) to do an integration by parts on the first term, the result immediately follows. \square

6.4.3 Large deviations for the GL model in a random environment

In this section, we shall be interested in the large deviations for a version of the process (6.32) in a random environment, where the operator A is replaced by a realization of the symmetric random matrix

$$A_{i,j}(\omega) := N^2 a_{i+1}(\omega) (\delta_{i,j-1} - \delta_{i,j}) - N^2 a_i(\omega) (\delta_{i,j} - \delta_{i,j+1}) \quad (6.44)$$

where the a_i are iid random variable defined on a probability space Ω , and we assume there exists a constant $c > 0$ such that we almost surely have

$$c \geq a_i \geq 1/c. \quad (6.45)$$

This assumption corresponds to an ellipticity assumption on the operator A that is uniform in the realization of the random field. Therefore, for any x and any realization of the random field, we have

$$\frac{1}{c}\langle A_0x, x \rangle \leq \langle Ax, x \rangle \leq c\langle A_0x, x \rangle. \quad (6.46)$$

Under these assumptions, the quantity

$$\bar{a} := \mathbb{E} \left[\frac{1}{a_1} \right] \quad (6.47)$$

is well defined and finite.

The associated SDE is

$$dX_t^i = N^2 (a_i(\psi(X_{i+1}) - \psi(X_i)) - a_{i-1}(\psi(X_i) - \psi(X_{i-1}))) + \sqrt{2a_i}dB_t^i - \sqrt{2a_{i-1}}dB_t^{i-1}. \quad (6.48)$$

Given a realization of the random environment, we denote by $L_{a,N}$ the generator of this diffusion.

The following hydrodynamic limit result for the random environment model has been proven in [Fri89] :

Theorem 6.36. *Assume that the sequence of initial data $f_{0,N}\mu_N$ weakly converges to a deterministic profile $\rho_0 \in H^1(\mathbb{T})$. Then, for any time $t > 0$ and any smooth function $J : \mathbb{T} \rightarrow \mathbb{R}$, the random variable $\frac{1}{N} \sum J(i/N)X_t^i$ converges in probability to $\int_{\mathbb{T}} J(\theta)\rho(t,\theta)d\theta$, where $\rho(t,\theta)$ is the unique solution to the PDE*

$$\frac{\partial \rho}{\partial t} = \bar{a}\Delta\varphi'(\rho)$$

with initial condition ρ_0 . This convergence holds for almost every realization of the random field.

We are interested in the following quenched large deviation principle for this model, using the gradient flow approach we developed in section 1.

Theorem 6.37. *Assume that the sequence of initial data is of the form (6.33) for some smooth initial profile ρ_0 . Then, for almost every realization of the random field, the sequence of random functions satisfies a LDP in $L^\infty(H^{-1})$ with speed N and rate function*

$$I(\rho) := \int \varphi(\rho(0,\theta)) - \varphi(m_0(\theta)) - \varphi'(m_0(\theta))(\rho(\theta) - m_0(\theta))d\theta + \frac{1}{4\bar{a}} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \bar{a} \frac{\partial^2}{\partial \theta^2} \varphi'(\rho) \right\|_{H^{-1}}^2 dt.$$

This generalizes the large deviations principle of [DV89] to the case of random environment.

In terms of gradient flows, this result follows from two facts :

- The relative entropy with respect to the invariant measure, divided by N , Γ -converges to $\rho \rightarrow \int \varphi(\rho) - \varphi(f\rho)$. This corresponds to a large deviations principle for the sequence of invariant measures μ_N ;
- The sequence of metrics given by $A^{-1}(w)$ almost surely converge to the H^{-1} norm, divided by a factor \bar{a} .

As a technical tool, we shall need the following convergence result, which will be used to formalize the convergence of the discrete norms.

Lemma 6.38. *Let $(a_q)_{q \in \mathbb{Q}}$ be a sequence of positive, bounded, iid random variables, and let $a_i^N := a_{i/N}$. With probability one, for any sequence (h^N) of step functions on \mathbb{T} that converges to a function h in L^1 , such that h^N is constant on $(\frac{i-1}{N}, \frac{i}{N}]$, and denoting by h_i^N the value of h^N on such an interval, we have*

$$\lim \frac{1}{N} \sum_{i=1}^N a_i^N h_i^N = \mathbb{E}(a) \int h(\theta) d\theta.$$

Démonstration. Let M be an integer. The strong law of large numbers implies that, with probability 1, for any step function h that is constant on the intervals $(\frac{i-1}{M}, \frac{i}{M}]$, we have

$$\lim \frac{1}{N} \sum_{i=1}^N a_i^N h(i/N) = \mathbb{E}(a) \int h(\theta) d\theta.$$

This then remains true simultaneously for every integer M , still with probability 1. An approximation argument in $L^1(\mathbb{T})$ then yields the desired result. \square

Proof of Theorem 6.37. Given a realization of our random environment, the functional J_N is given by

$$J_N(\nu_t) = \text{Ent}_\mu(\nu_T) - \text{Ent}_\mu(\nu_0) + \frac{1}{2} \int_0^T \|\partial_t \nu_t\|_{H^{-1}(A^{-1})} dt + \frac{1}{2} \int_0^T \int \frac{\langle A \nabla g_t, \nabla g_t \rangle}{g_t} d\mu dt \quad (6.49)$$

where g_t is the density of ν_t with respect to μ .

Lemma 6.39. *The functional $\frac{1}{N} \text{Ent}_{\mu_N} \Gamma$ -converges at every Dirac mass to $\rho \rightarrow \int \varphi(\rho) d\theta - \varphi(\int \rho d\theta)$.*

Démonstration. Let ν_N be a sequence that weakly converges to a deterministic profile $\rho : \mathbb{T} \rightarrow \mathbb{R}$, and let

$$G_N(x) := \frac{1}{Z_N} \exp \left(\sum_{i=1}^N \varphi'(\rho_{i,N}) x_i \right) \quad (6.50)$$

be a local Gibbs profile with respect to μ , where $\rho_{i,N} = \int_{(i-1)/N}^{i/N} \rho(\theta) d\theta$ and Z_N is the normalization constant such that $G_N \mu$ is a probability measure.

We then have the decomposition

$$\text{Ent}_{G_N \mu}(\nu_N) = \text{Ent}_\mu(\nu_N) - \int \log G_N d\nu_N. \quad (6.51)$$

By definition of G_N , since ν_N weakly converges to the deterministic profile ρ , as we have seen in the proof of Proposition 6.31, $\frac{1}{N} \int \log G_N d\nu_N$ converges to $\int \varphi(\rho) d\theta - \varphi(\int \rho d\theta)$. Since the relative entropy is nonnegative, we can deduce from (6.51) the inequality

$$\liminf \frac{1}{N} \text{Ent}_\mu(\nu_N) \geq \int \varphi(\rho) d\theta - \varphi \left(\int \rho d\theta \right).$$

Moreover, the measures $G_N \mu$ weakly converge to the deterministic profile ρ , so that they provide the recovery sequence for this Gamma-convergence result. \square

Lemma 6.40. *Let (ν_N) be a sequence of probability measures that converges to a profile ρ , such that, for any N , ν_N is absolutely continuous with respect to μ_N .*

Then

$$\frac{1}{N} \text{Ent}_{f_{0,N}\mu_N}(\nu_N) - \frac{1}{N} \text{Ent}_{\mu_N}(\nu_N) \longrightarrow \int_{\mathbb{T}} \varphi'(m_0)(m_0 - \rho) - \varphi(m_0) d\theta.$$

Démonstration. Since ν_N is absolutely continuous with respect to μ_N (and therefore to $f_{0,N}\mu_N$), we have

$$\begin{aligned} & \frac{1}{N} \text{Ent}_{f_{0,N}\mu_N}(\nu_N) - \frac{1}{N} \text{Ent}_{\mu_N}(\nu_N) \\ &= -\frac{1}{N} \int \log f_{0,N} d\nu_N \\ &= -\frac{1}{N} \int \sum \varphi'(m_0(i/N))x_i \nu_N(dx) + \frac{1}{N} \log \int \exp\left(\sum \varphi'(m_0(i/N))x_i\right) \mu_N(dx) \\ &= -\frac{1}{N} \int \sum \varphi'(m_0(i/N))x_i \nu_N(dx) + \frac{1}{N} \log \int \exp\left(\sum \varphi'(m_0(i/N))x_i - \psi(x_i)\right) dx \\ &\quad - \frac{1}{N} \log \int \exp\left(\sum -\psi(x_i)\right) dx \\ &\longrightarrow -\int \varphi'(m_0(\theta))\rho(\theta) d\theta + \int \varphi'(m_0(\theta))m_0(\theta) d\theta \\ &\quad - \int \varphi(m_0(\theta)) d\theta \end{aligned}$$

since (ν_N) has asymptotic profile ρ , and applying Theorem 6.29. □

We will now investigate the behavior of the slope :

Lemma 6.41 (Lower bound for the time-derivative). *For any time t and subsequence such that*

$$\sup_N \frac{1}{N} J_N(f_N) < +\infty,$$

we have

$$\liminf \frac{1}{N} \int_0^T |\dot{\nu}_{N,t}|_{\nu_{N,t}}^2 dt \geq \frac{1}{a} \int_0^T \|\partial\rho/\partial t\|_{H^{-1}}^2 dt.$$

Démonstration. Let $J : [0, T] \times \mathbb{T} \longrightarrow \mathbb{R}$ be a smooth function. Applying Lemma 6.35 with $J_N(t, x) := \frac{1}{N} \sum_i J(t, i/N)x_i + \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(t, j/N)b_j x_i$ and $b_i = \frac{\bar{a}}{a_i} - 1$, we have

$$\begin{aligned} & \frac{1}{N} \int_0^T |\dot{\nu}_{N,t}|_{\nu_{N,t}}^2 dt \geq \frac{2}{N} \int \sum J(T, i/N)x_i \nu_{N,T}(dx) + \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(T, j/N)b_j x_i \nu_{N,T}(dx) \\ & \quad - \frac{2}{N} \int \sum J(0, i/N)x_i \nu_{N,T}(dx) - \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(0, j/N)b_j x_i \nu_{N,0}(dx) \\ & \quad - \frac{2}{N} \int_0^T \int \sum \frac{\partial J}{\partial t}(t, i/N)x_i \nu_{N,t}(dx) dt - \frac{2}{N^2} \int_0^T \int \sum_i \sum_{j=1}^{i-1} \frac{\partial^2 J}{\partial t \partial \theta}(t, j/N)b_j x_i \nu_{N,t}(dx) dt \\ & \quad - \frac{1}{N} \int_0^T \sum a_i \left(N J\left(t, \frac{i+1}{N}\right) - N J\left(t, \frac{i}{N}\right) + b_i \frac{\partial J}{\partial \theta}\left(t, \frac{i}{N}\right) \right)^2 dt. \end{aligned}$$

Taking the limit $N \rightarrow +\infty$ and using the second-moment bounds of Lemma 6.27, we get

$$\begin{aligned} & \frac{2}{N} \int \sum J(T, i/N) x_i \nu_{N,T}(dx) + \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(T, j/N) b_j x_i \nu_{N,T}(dx) \\ &= \frac{2}{N} \int \sum J(T, i/N) x_i \nu_{N,T}(dx) + O\left(\frac{1}{N}\right) \\ &\rightarrow \int_{\mathbb{T}} J(T, \theta) \rho(T, \theta) d\theta. \end{aligned}$$

In the same way,

$$\begin{aligned} & \frac{2}{N} \int \sum J(0, i/N) x_i \nu_{N,T}(dx) + \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(0, j/N) b_j x_i \nu_{N,0}(dx) \\ &\rightarrow \int_{\mathbb{T}} J(0, \theta) \rho(0, \theta) d\theta \end{aligned}$$

and

$$\begin{aligned} & \frac{2}{N} \int_0^T \int \sum \frac{\partial J}{\partial t}(t, i/N) x_i \nu_{N,t}(dx) dt + \frac{2}{N^2} \int_0^T \int \sum_i \sum_{j=1}^{i-1} \frac{\partial^2 J}{\partial t \partial \theta}(t, j/N) b_j x_i \nu_{N,t}(dx) dt \\ &\rightarrow \int_0^T \int_{\mathbb{T}} \frac{\partial J}{\partial t}(t, \theta) \rho(t, \theta) d\theta dt. \end{aligned}$$

Finally, using Lemma 6.38, we get

$$\begin{aligned} & \frac{1}{N} \int_0^T \sum a_i \left(N J\left(t, \frac{i+1}{N}\right) - N J\left(t, \frac{i}{N}\right) + b_i \frac{\partial J}{\partial \theta}\left(t, \frac{i}{N}\right) \right)^2 dt \\ &\rightarrow \int_0^T \int_{\mathbb{T}} \bar{a} \left(\frac{\partial J}{\partial \theta}(t, \theta) \right)^2 d\theta dt. \end{aligned}$$

Combining these lower bounds, we get

$$\begin{aligned} \liminf \frac{1}{N} \int_0^T |\dot{\nu}_{N,t}|_{\nu_{N,t}}^2 dt &\geq 2 \int_{\mathbb{T}} J(T, \theta) \rho(T, \theta) d\theta - 2 \int_{\mathbb{T}} J(0, \theta) \rho(0, \theta) d\theta \\ &\quad - 2 \int_0^T \int_{\mathbb{T}} \frac{\partial J}{\partial t}(t, \theta) \rho(t, \theta) d\theta dt - \int_0^T \int_{\mathbb{T}} \bar{a} \left(\frac{\partial J}{\partial \theta}(t, \theta) \right)^2 d\theta dt. \end{aligned}$$

Taking the supremum over all smooth functions J yields the result. \square

Lemma 6.42 (Lower bound for the upper gradient). *For any time t and subsequence such that*

$$\sup_N \frac{1}{N} \int \frac{\langle A_0(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} < +\infty,$$

we almost surely have

$$\liminf_N \frac{1}{N} \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} \geq \bar{a} \int_{\mathbb{T}} (\partial_\theta \varphi'(\rho(t, \theta)))^2 d\theta.$$

Démonstration. From Lemma 6.34, we know that, under these assumptions, $\varphi'(\rho)$ lies in $H^1(\mathbb{T})$.

Let $J : \mathbb{T} \rightarrow \mathbb{R}$ be a smooth function, and define

$$G_N(x) := \exp \left(\sum_{i=1}^N \left(\sum_{j=1}^i \frac{\bar{a}}{a_j} J(i/N) \right) x_i \right). \quad (6.52)$$

Since the upper gradient takes value $+\infty$ when ν_N is not absolutely continuous with respect to μ_N , and we are looking for a lower bound, we can assume without loss of generality that ν_N is absolutely continuous with respect to μ_N , and therefore to $G_N\mu_N$. Let $g_{N,t}$ be the density of $\nu_{N,t}$ with respect to μ_N .

We consider the quantity $\int \frac{\langle A\nabla \left(\frac{g_N}{G_N} \right), \nabla \left(\frac{g_N}{G_N} \right) \rangle}{g_{N,t}/G_N} G_N d\mu_N$, which is nonnegative. We have

$$\begin{aligned} & \int \frac{\langle A\nabla \left(\frac{g_N}{G_N} \right), \nabla \left(\frac{g_N}{G_N} \right) \rangle}{g_{N,t}/G_N} G_N d\mu_N \\ &= \int \frac{\langle A\nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_N - 2 \int \frac{\langle A\nabla g_{N,t}, \nabla G_N \rangle}{G_N} d\mu_N \\ & \quad + \int \frac{\langle A\nabla G_N, \nabla G_N \rangle}{G_N^2} d\mu_N \\ &= \int \frac{\langle A\nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_N - 2 \int \frac{\langle A\nabla H, \nabla G_N \rangle}{G_N} g_{N,t} d\mu_N \\ & \quad + \int \frac{\langle A\nabla G_N, \nabla G_N \rangle}{G_N^2} d\mu_N. \end{aligned}$$

Therefore, we have

$$\int \frac{\langle A\nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_N \geq 2 \int \frac{\langle A\nabla H, \nabla G_N \rangle}{G_N} g_{N,t} d\mu_N - \int \frac{\langle A\nabla G_N, \nabla G_N \rangle}{G_N^2} d\mu_N \quad (6.53)$$

for any realization of the random field, any N and any t .

Applying Lemma 6.38, we have

$$\begin{aligned} \frac{1}{N} \int \frac{\langle A\nabla G_{N,t}, \nabla G_{N,t} \rangle}{G_{N,t}^2} d\mu_N &= \frac{1}{N} \sum_{i=1}^N \frac{\bar{a}^2}{a_i} J(i/N)^2 \\ &\rightarrow \bar{a} \int_{\mathbb{T}} J(\theta)^2 d\theta. \end{aligned} \quad (6.54)$$

We also have

$$\begin{aligned} \frac{1}{N} \int \frac{\langle A\nabla H, \nabla G_N \rangle}{G_N} g_{N,t} d\mu_N &= \frac{\bar{a}}{N} \int \sum (\psi'(x_{i+1}) - \psi'(x_i)) J(i/N) g_N(x) \mu_N(dx) \\ &= \frac{\bar{a}}{N} \int \sum \psi'(x_i) (J((i-1)/N) - J(i/N)) g_N(x) \mu_N(dx) \\ &= \frac{\bar{a}}{N} \int \sum \psi'(x_i) J'(i/N) g_N(x) \mu_N(dx) \\ & \quad + O \left(\frac{1}{N^2} \int \sum |\psi'(x_i)| g_N(x) \mu_N(dx) \right) \\ &\rightarrow \bar{a} \int_{\mathbb{T}} \varphi'(\rho(\theta)) J'(\theta) d\theta \end{aligned} \quad (6.55)$$

Combining these two lower bounds and taking the supremum over smooth functions J , we get the lower bound of our Lemma. \square

From the previous Lemma and Fatou's Lemma, we can then deduce that, for a sequence that converges to a Dirac mass, and such that $J_N(f_N) \leq CN$, we have

$$\liminf \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \geq \int_0^T \int_{\mathbb{T}} \bar{a} \int_{\mathbb{T}} (\partial_\theta \varphi'(\rho(t, \theta)))^2 d\theta dt$$

which was the last element we needed for the lower bound of the Γ -convergence.

We now turn to the recovery sequence. Given a profile $\rho(t, \theta)$ that is weakly continuous in time, and such that $\int_0^T \|\bar{a} \partial_\theta^2 \varphi'(\rho) - \partial_t \rho\|_{H^{-1}}^2 dt$ is finite, there exists a sequence of smooth profiles ρ_k that converge to ρ , and such that $\int_0^T \|\bar{a} \partial_\theta^2 \varphi'(\rho_k) - \partial_t \rho_k\|_{H^{-1}}^2 dt$ converges to $\int_0^T \|\bar{a} \partial_\theta^2 \varphi'(\rho) - \partial_t \rho\|_{H^{-1}}^2 dt$. Therefore, in view of Remark 6.16, we only have to prove the existence of a recovery sequence for profiles ρ that are smooth.

Given such a smooth profile ρ , there exists a continuous function $h : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ such that

$$\frac{\partial \rho}{\partial t} = \bar{a} \frac{\partial}{\partial \theta} \left(h(t, \theta) + \frac{\partial}{\partial \theta} \varphi'(\rho) \right). \quad (6.56)$$

We now consider a dynamic with law given by the time-dependent generator

$$\tilde{L}_{a,N} = L_{a,N} + N \sum_{i=1}^N \frac{\bar{a}}{a_i} h(t, i/N) \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_i} \right) \quad (6.57)$$

and initial condition given by the local Gibbs state associated to $\rho(0, \cdot)$.

We need to prove two things : that the solutions to such dynamics converge to the deterministic profile ρ , and that $N^{-1} J_N(f_N)$ has the correct limit (where f_N are the marginals of the law of the solution). The first part can be done in the same way as in [Q, Section 3], so we concentrate on the second part. We have

$$\begin{aligned} \frac{1}{N} J_N(f_N) &= \frac{1}{2} \int_0^T \left\| \frac{\partial f_N}{\partial t} - L_{a,N} f_N \right\|_{H^{-1}(A)}^2 dt \\ &= \frac{1}{2} \int_0^T \left\| \tilde{L}_{a,N} f_N - L_{a,N} f_N \right\|_{H^{-1}(A)}^2 dt \\ &= \frac{1}{2} \int_0^T \sum_{i=1}^N \frac{\bar{a}^2}{a_i} h(t, i/N)^2 dt \\ &\rightarrow \frac{1}{2} \int_0^T \int_{\mathbb{T}} \bar{a} h(t, \theta)^2 d\theta dt \end{aligned} \quad (6.58)$$

and, using (6.56), it is easy to see that this is equal to $\frac{1}{2\bar{a}} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \bar{a} \frac{\partial^2}{\partial \theta^2} \varphi'(\rho) \right\|_{H^{-1}}^2 dt$, which was what we needed to prove.

We still have to prove exponential tightness for the laws of solutions to (6.48). It is given by the following two results :

Lemma 6.43. *Let \mathbb{P}_N be the law of a solution to the SDE (6.48) with initial condition X_0 having a distribution $f_0 \mu_N$ that satisfies $\text{Ent}_{\mu_N}(f_0) \leq CN$. Then*

$$\lim_{\ell \rightarrow +\infty} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq t \leq T} \frac{1}{N} \sum |X_t^i| \geq \ell \right) = -\infty.$$

Lemma 6.44. *Under the same assumptions as the previous lemma, for any $\epsilon > 0$ and any smooth function on the torus J , we have*

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq s \leq t \leq T, |s-t| \leq \delta} \left| \frac{1}{N} \sum J(i/N)(X_t^i - X_s^i) \right| \geq \epsilon \right) = -\infty.$$

Proof of Lemma 6.43. This proof is exactly the same as in [DV89], we give a brief sketch to show that the random field (a_i) does not make any difference..

Let $\tilde{P}^{eq,N}$ be the law of a solution to the SDE starting from the equilibrium measure μ_N . From Lemma 1.12 in [KV86], we know that, for any symmetric function g on \mathbb{R}^N , we have

$$\tilde{P}^{eq,N} \left(\sup_{0 \leq t \leq T} g(X_t) \geq \ell \right) \leq \frac{3}{\ell} \sqrt{a + Tb}$$

with $a = \int g^2 d\mu_N$ and $b = \int \frac{A \nabla g, \nabla g}{g} d\mu_N$. When $g(x) = \exp(\sum |x_i|)$, there exists $C > 0$ such that $a \leq C^N$ and $b \leq N^2 C^N$. Using the Tchebychev inequality, we obtain

$$\tilde{P}^{eq,N} \left(\sup_{0 \leq t \leq T} \frac{1}{N} \sum |X_t^i| \geq \ell \right) \leq \sqrt{C^N(1 + N^2)} e^{-N\ell} \leq C' e^{-C'' N\ell}.$$

We can then use the basic entropy inequality

$$\mathbb{P}_N(A) \leq \frac{\log(2) + H_N}{\log(1 + 1/\tilde{P}^{eq,N}(A))},$$

where H_N is the relative entropy of the non equilibrium process with respect to the equilibrium process, which satisfies the bound $H_N \leq CN$, to get the result. \square

In the same way, we refer the reader interested in the proof of Lemma 6.44 to [DV, Lemma 2.8]. The proof is exactly the same, since the random variables a_i are bounded. \square

6.4.4 Large deviations for the non-gradient Ginzburg-Landau model

We consider the SDE given by

$$dX_t^i = N^2(W_{i,i+1} - W_{i-1,i})dt + N\sqrt{2a(X_t^i, X_t^{i+1})}dB_t^{i+1} - N\sqrt{2a(X_t^{i-1}, X_t^i)}dB_t^i \quad (6.59)$$

where

$$W_{i,i+1} = a(X_t^i, X_t^{i+1})(\psi'(X_t^i) - \psi'(X_t^{i+1})) - \frac{\partial a}{\partial x}(X_t^i, X_t^{i+1}) + \frac{\partial a}{\partial y}(X_t^i, X_t^{i+1}). \quad (6.60)$$

The marginals of the law of a solution to this SDE solve the PDE

$$\frac{\partial f \mu_N}{\partial t} = \nabla \cdot (A(x) \nabla f \mu_N) \quad (6.61)$$

where the matrix $A(x)$ is given by

$$A(x)_{i,j} := a(x_{i-1}, x_i)(\delta_{i,j} - \delta_{i,j+1}) + a(x_i, x_{i+1})(\delta_{i,j} - \delta_{i,j-1}).$$

Once more, we will assume that the initial data f_0 is of the form (6.33).

The generator of this dynamic is given by

$$Lf = N^2 \sum e^H \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_{i+1}} \right) e^{-H} a(x_i, x_{i+1}) \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_{i+1}} \right) f \quad (6.62)$$

It has been shown in [Va] that trajectories of such a dynamic concentrate around the solution to the PDE

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \theta} \left(\hat{a}(\rho) \frac{\partial}{\partial \theta} \varphi'(\rho) \right) \quad (6.63)$$

where φ is the same function as in the previous section, and \hat{a} is a bounded continuous function, which we shall now define.

Let $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$ be a bounded, smooth function of a finite number of variables. The function $\xi : x \in \mathbb{R}^\infty \rightarrow \sum_{i=-\infty}^{+\infty} F(x_{i-k}, \dots, x_{i+k})$ is not well-defined, but its partial derivatives are. We can therefore define

$$a_F(y) := \int a(x_0, x_1) \left(1 - \frac{\partial \xi}{\partial x_1} + \frac{\partial \xi}{\partial x_0} \right)^2 \mu^{\infty, y}(dx) \quad (6.64)$$

where $\mu^{\infty, y}$ is the product measure on \mathbb{R}^∞ with every one-dimensional marginal having density $Z^{-1} \exp(\lambda x - \psi(x))$, with λ the unique real number such that this measure has expectation y .

The function \hat{a} is then given by

$$\hat{a}(y) := \inf_F a_F(y) \quad (6.65)$$

with the infimum running over the set of all smooth, bounded functions of a finite number of variables.

In the next proposition, we summarize a few properties of the function \hat{a} :

Proposition 6.45. (i) \hat{a} is a bounded, continuous function.

(ii) For any $\epsilon > 0$ and $C < +\infty$, there exists a smooth real-valued function $g(x_k, \dots, x_k, y)$ on \mathbb{R}^{2k+2} with bounded first derivatives such that

$$\sup_{|y| \leq C} \left(a_{g(\cdot, y)}(y) - \hat{a}(y) \right) < \epsilon$$

and

$$\sup_{y \in \mathbb{R}} \left(a_{g(\cdot, y)}(y) - \hat{a}(y) \right) \leq \|a\|_\infty.$$

Part (i) of this Proposition comes from [Va], and part (ii) from [Qua95].

Our aim is to prove the following large deviations result :

Theorem 6.46. Assume that the sequence of initial data is of the form (6.33) for some smooth initial profile ρ_0 . The sequence of random functions satisfies a LDP in $L^\infty(H^{-1})$ with speed N and rate function

$$\begin{aligned} I(\rho) := & \int \varphi(\rho(0, \theta)) - \varphi(m_0(\theta)) - \varphi'(m_0(\theta))(\rho(\theta) - m_0(\theta)) d\theta \\ & + \frac{1}{4} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \frac{\partial}{\partial \theta} \left(\hat{a}(\rho) \frac{\partial}{\partial \theta} \varphi'(\rho) \right) \right\|_{H^{-1}(\hat{a}(\rho(t, \cdot)))}^2 dt. \end{aligned}$$

In the rate function, the norm is defined as

$$\|u\|_{H^{-1}(\hat{a}(\rho(t, \cdot)))}^2 := \sup_{v \in H^1(\mathbb{T})} 2 \int_{\mathbb{T}} u(\theta) v(\theta) d\theta - \int \hat{a}(\rho(t, \theta)) \left(\frac{\partial v}{\partial \theta} \right)^2 d\theta. \quad (6.66)$$

This result was already proved in [Qua95], under the assumption that the single site potential ψ is uniformly convex, and that its second derivative is bounded above. Our assumptions are a priori more general, since they allow for superquadratic potentials, but it seems likely that, using the logarithmic Sobolev inequality proved in [MO13], the method of [Qua95] could be extended for such functions.

The following result is the key technical estimate to prove large deviations for nongradient models. It has been proven in [Var93].

Let R_N be the law of the stationary solution to the SDE (6.59), with the initial condition X_0 having law μ_N . Let $\rho(t, \theta)$ be a deterministic profile, which we assume to be in $L^\infty(H^1)$.

Denote by $\rho_\theta^c(t) := \int_{\theta-c}^{\theta+c} \rho(t, s) ds$. For a given smooth function $J : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ and a function g as in part (ii) of Proposition 6.45, we define

$$\begin{aligned} V(t) := & \sum_{i=1}^N J(t, i/N) \left[W_{i, i+1} - \frac{1}{N^2} Lg \left(X_t^{i-k}, \dots, X_t^{i+k}, \rho_{i/N}^{\ell/N}(t) \right) \right] \\ & + \frac{1}{N} \sum_{i=2}^N J(t, i/N) \hat{a}(\rho_{i/N}^{\epsilon_1}(t)) \left(\frac{\varphi'(\rho_{i/N+\epsilon_2}^{\epsilon_1}(t)) - \varphi'(\rho_{i/N-\epsilon_2}^{\epsilon_1}(t))}{2\epsilon_2} \right) \\ & - \frac{\alpha}{N} \sum_{i=1}^N J(t, i/N)^2 (a_{g(\cdot, \rho_{i/N}^{\epsilon_1}(t))}(\rho_{i/N}^{\epsilon_1}(t)) - \hat{a}(\rho_{i/N}^{\epsilon_1}(t))) \end{aligned} \quad (6.67)$$

where α , ϵ_1 and ϵ_2 are positive numbers, and ℓ is a positive integer.

Under these notations, we have the following exponential estimate, which is due to Varadhan [Var93] :

Theorem 6.47. *For any profile ρ , any $\alpha > 0$, any J and g , we have*

$$\lim_{\epsilon_2 \rightarrow 0} \limsup_{\epsilon_1 \rightarrow 0} \limsup_{\ell \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}^{R_N} \left[\exp \left(\alpha N \int_0^T V(t) dt \right) \right] \leq 0.$$

Using the result, we obtain what is known in the hydrodynamic literature as the gradient replacement estimate.

Corollary 6.48. *Let $(\nu_{N,t})$ be a sequence of flows of time-marginals of a sequence of processes that weakly converge to a deterministic flow $\rho(t, \theta)$. Assume moreover that $J_N(\nu_{N,t}) \leq CN$ for some $C > 0$.*

Then for any smooth, bounded functions $J : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ and $g : \mathbb{R}^{2k+2} \rightarrow \mathbb{R}$, we have

$$\begin{aligned} \lim_{\ell \rightarrow \infty} \lim_N \int_0^T \int \sum J(t, i/N) (W_{i, i+1}(x) - \frac{1}{N^2} (Lg)(x_{i-k}, \dots, x_{i+k}, \rho_{i/N}^{\ell/N})) \nu_{N,t}(dx) dt \\ = \int_0^T \int_{\mathbb{T}} J(t, \theta) \hat{a}(\rho(t, \theta)) \partial_\theta \varphi'(\rho(t, \theta)) d\theta dt. \end{aligned}$$

Démonstration. First, let us show that we can build a diffusion process of law P with time-marginals $(\nu_{N,t})$ such that $H(P_N, R_N) = O(N)$. Since the equilibrium process is a solution of (6.59) with initial condition μ_N , we know that there exists a process with marginals $(\nu_{N,t})$ such that

$$\begin{aligned} H(P_N, R_N) &= \frac{1}{2} \text{Ent}_{\mu_N}(\nu_{0,N}) + \frac{1}{2} \text{Ent}_{\mu_N}(\nu_{N,T}) + \frac{1}{4} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \\ &+ \frac{1}{4} \int_0^T |\dot{\nu}_t|^2 dt \\ &= J_N(\nu_{N,t}) + \text{Ent}_{\mu_N}(\nu_{N,0}) - \text{Ent}_{f_{0,N} \mu_N}(\nu_{N,0}) \end{aligned}$$

so we only have to get an upper bound on $\text{Ent}_{\mu_N}(\nu_{N,0}) - \text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0})$. Moreover, the bound on $J_N(\nu_{N,t})$ implies that $\text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0}) \leq CN$.

Denoting $\nu_{N,0} = \rho_N \mu_N$, we have

$$\begin{aligned} \text{Ent}_{\mu_N}(\nu_{N,0}) - \text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0}) &= \int \rho \log f_{0,N} d\mu_N \\ &= \int \sum \varphi'(m_0(i/N)) x_i \nu_{N,0}(dx) - \log Z_N \\ &\leq CN \times \sqrt{\frac{1}{N} \int \sum |x_i|^2 \nu_{N,0}(dx)} + CN \\ &\leq CN \end{aligned} \tag{6.68}$$

where the last bound follows from $\text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0}) \leq CN$, Lemma 6.27 and the fact that the measures $f_{0,N}\mu_N$ satisfy a LSI with uniform constant, and have uniformly bounded second moments.

Consequently, using the entropy inequality, we have

$$\mathbb{E}^{P_N} \left[\int_0^T V(t) dt \right] \leq \frac{C}{\alpha} + \frac{1}{\alpha N} \log \mathbb{E}^{R_N} \left[\exp \left(\alpha N \int_0^T V(t) dt \right) \right] \tag{6.69}$$

and the result immediately follows from Theorem 6.47 and using the fact that the inequality is valid for both J and $-J$. \square

We also recall the exponential tightness estimates that have been proven in [Var93], and which we need to apply Corollary 6.18 :

Lemma 6.49. *Let \mathbb{P}_N be the law of a solution to the SDE 6.59 with initial condition X_0 having a distribution $f_0 \mu_N$ that satisfies $\text{Ent}_{\mu_N}(f_0) \leq CN$. Then*

$$\lim_{\ell \rightarrow +\infty} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq t \leq T} \frac{1}{N} \sum |X_t^i| \geq \ell \right) = -\infty.$$

and for any $\epsilon > 0$ and any smooth function on the torus J , we have

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq s \leq t \leq T, |s-t| \leq \delta} \left| \frac{1}{N} \sum J(i/N) (X_t^i - X_s^i) \right| \geq \epsilon \right) = -\infty.$$

The proof of these estimates is exactly the same as for the gradient case studied in [DV89]. Once more, the use of a variable function $a(x_i, x_{i+1})$ does not make a difference as long as it is bounded.

Lemmas 6.39 and 6.40 remain valid, so that, to prove Theorem 6.46, we only have to study the behavior of the slopes.

Lemma 6.50.

$$\liminf_N \frac{1}{N} \int_0^T |\dot{\nu}_t|^2 dt \geq \int_0^T \|\partial_t \rho\|_{H^{-1}(\hat{a})}^2 dt$$

Démonstration. Let $J : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ be a smooth function, and $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$ be a smooth, bounded function. Applying Lemma 6.35 with $V(t, x) = \sum J(t, i/N) x_i + \frac{1}{N} J'(t, i/N) F(x_{i-k}, \dots, x_{i+k})$, we get

$$\begin{aligned} \frac{1}{N} \int_0^T |\dot{\nu}_t|^2 dt &\geq 2 \frac{1}{N} \int V(T, x) \nu_T(dx) - 2 \frac{1}{N} \int V(0, x) \nu_0(dx) \\ &\quad - 2 \frac{1}{N} \int_0^T \int \frac{\partial V}{\partial t}(x) \nu_t(dx) dt \\ &\quad - \frac{1}{N} \int_0^T \int \langle A(x) \nabla V(t, x), \nabla V(t, x) \rangle \nu_t(dx) dt. \end{aligned} \tag{6.70}$$

We have

$$\begin{aligned} \frac{1}{N} \int V(T, x) \nu_T(dx) &= \frac{1}{N} \int J(T, i/N) x_i \nu_T(dx) + \frac{1}{N^2} \int J'(T, i/N) F(x_{i-k}, \dots, x_{i+k}) \nu_T(dx) \\ &= \frac{1}{N} \int J(T, i/N) x_i \nu_T(dx) + O\left(\frac{1}{N}\right) \\ &\longrightarrow \int J(T, \theta) \rho(T, \theta) d\theta. \end{aligned} \quad (6.71)$$

In the same way,

$$\frac{1}{N} \int V(0, x) \nu_0(dx) \longrightarrow \int J(0, \theta) \rho(0, \theta) d\theta \quad (6.72)$$

and

$$\frac{1}{N} \int_0^T \int \frac{\partial V}{\partial t}(x) \nu_t(dx) dt \longrightarrow \int_0^T \int_{\mathbb{T}} \frac{\partial J}{\partial t}(t, \theta) \rho(t, \theta) d\theta dt. \quad (6.73)$$

For the last term, we have

$$\begin{aligned} &\frac{1}{N} \int_0^T \int \langle A(x) \nabla V(t, x), \nabla V(t, x) \rangle \nu_t(dx) dt \\ &= \frac{1}{N} \int_0^T \int \sum_i N^2 a(x_i, x_{i+1}) \left(J(t, (i+1)/N) - J(t, i/N) \right. \\ &\quad \left. + \frac{1}{N} \sum_{j=i-k}^{i+k} \frac{\partial}{\partial x_{i+1}} F(x_{j-k}, \dots, x_{j+k}) - \frac{\partial}{\partial x_i} F(x_{j-k}, \dots, x_{j+k}) \right)^2 \nu_t(dx) dt \\ &= \frac{1}{N} \int_0^T \int \sum_i a(x_i, x_{i+1}) J' \left(t, \frac{k}{N} \right)^2 \left(1 - \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_i} \right) \sum_{j=i-k}^{i+k} F(x_{j-k}, \dots, x_{j+k}) + O\left(\frac{k}{N}\right) \right)^2 \nu_t(dx) dt \\ &\longrightarrow \int_0^T \int_{\mathbb{T}} a_F(\rho(t, \theta)) \left(\frac{\partial J}{\partial \theta} \right)^2 d\theta dt. \end{aligned} \quad (6.74)$$

We combine these convergence estimates, and then optimize in F and J to get the desired result. \square

Lemma 6.51. *For a flow of marginals $\nu_{N,t}$ that weakly converges to ρ , and such that $\frac{1}{N} J_N((\nu_{N,t})_t)$ is bounded, we have*

$$\liminf_N \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \geq \int_0^T \int_{\mathbb{T}} \hat{a}(\rho(t, \theta)) (\partial_\theta \varphi'(\rho))^2 d\theta dt.$$

Démonstration. Let $J(t, \theta)$ be a smooth function, $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$ a smooth function and let $\xi(x) = \sum F(x_{i-k}, \dots, x_{i+k})$. We define $\vec{J}_N(t, x)$ the element of \mathbb{R}^N given by $\vec{J}_N(t, x)_i := \sum_{j=1}^{i-1}$

$$\begin{aligned}
& J'(t, j/N) + \frac{\partial}{\partial x_i} \left[\frac{1}{N} \sum_{j=1}^N J'(t, j/N) F(x_{j-k}, \dots, x_{j+k}) \right]. \text{ We have} \\
& \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \\
& \geq \frac{2}{N} \int_0^T \int \langle A(\nabla \nu_t + \nu_t \nabla H), \bar{J} \rangle dt - \frac{1}{N} \int_0^T \int \langle A \bar{J}, \bar{J} \rangle \nu_t(dx) dt \\
& = 2 \int_0^T \int \sum J'(t, i/N) W_{i, i+1}(x) \nu_t(dx) dt \\
& \quad - 2 \int_0^T \int \sum \frac{J'(t, i/N)}{N^2} (LF)(x_{i-k}, \dots, x_{i+k}) \nu_t(dx) dt \\
& \quad - \int_0^T \int \sum a(x_i, x_{i+1}) J'(t, i/N)^2 \left(1 - \frac{\partial \xi}{\partial x_{i+1}} + \frac{\partial \xi}{\partial x_i} \right)^2 \nu_t(dx) dt + o(1) \quad (6.75)
\end{aligned}$$

We then just have to use Corollary 6.48 and optimize in F to obtain

$$\begin{aligned}
& \liminf \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \\
& \geq 2 \int_0^T \int_{\mathbb{T}} \varphi'(\rho(t, \theta)) J'(t, \theta) d\theta dt - \int_0^T \int_{\mathbb{T}} \hat{a}(\rho) J(t, \theta)^2 d\theta dt.
\end{aligned}$$

Taking the supremum over all smooth functions J then yields our Lemma.

For the Gamma-convergence upper bound, the method we use is pretty much the same as the proof of the LDP lower bound in [Qua95], so we only give a rough sketch. We fix a smooth profile ρ , for which there exists a continuous function $h : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ such that

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \theta} \left(\hat{a}(\rho) \left(\frac{\partial}{\partial \theta} \varphi'(\rho) + h \right) \right). \quad (6.76)$$

We consider an evolution given by the generator

$$\tilde{L}_N f := L_N f + N \sum_{i=1}^N h(t, i/N) a(x_i, x_{i+1}) \left(1 + \frac{\partial}{\partial x_{i+1}} \xi - \frac{\partial}{\partial x_i} \xi \right) \left(\frac{\partial}{\partial x_{i+1}} f - \frac{\partial}{\partial x_i} f \right).$$

and initial condition the local Gibbs state associated to $\rho(0, \cdot)$. It is shown in [Q, Section 3] that the solutions converge to the deterministic profile ρ . We write $f_{N,F}$ the law of the solution. We have

$$\begin{aligned}
\frac{1}{N} J_N(f_N) &= \frac{1}{2N} \int_0^T \int a(x_i, x_{i+1}) h(t, i/N)^2 \left(1 + \frac{\partial}{\partial x_{i+1}} \xi - \frac{\partial}{\partial x_i} \xi \right)^2 f_N(dx) dt \\
&\rightarrow \frac{1}{2} \int_0^T \int_{\mathbb{T}} a_F(\rho) \left(\frac{\partial h}{\partial \theta} \right)^2 d\theta dt \quad (6.77)
\end{aligned}$$

If, instead of using a fixed function F , we use a sequence (F_N) such that $a_{F_N}(\rho)$ uniformly converges to $\hat{a}(\rho)$ on compact sets (which is possible, see Proposition 6.45), we obtain the upper bound

$$\limsup \frac{1}{N} J_N(f_{N, F_N}) \leq \frac{1}{2} \int_0^T \int_{\mathbb{T}} \hat{a}(\rho) \left(\frac{\partial h}{\partial \theta} \right)^2 d\theta dt$$

which, by representation (6.76), is the one we needed to prove the upper bound in the Gamma convergence. \square

Acknowledgments : I would like to thank Christian Léonard for having pointed out to me the use of Girsanov's theorem to understand relative entropy for diffusion processes. I would also like to thank Thierry Bodineau, Georg Menz, Felix Otto, S.R.S. Varadhan and Cédric Villani for discussions about gradient flows and hydrodynamic limits.

Troisième partie

**Discretization of stochastic
differential equations**

Chapitre 7

Transport entropy inequalities for stochastic approximation schemes

Les résultats de cette section ont été obtenus en collaboration avec Noufel Frikha, et sont issus de l'article [FF13].

Abstract : We obtain new transport-entropy inequalities and, as a by-product, new deviation estimates for the law of an Euler like discretization scheme of a diffusion process at a fixed deterministic date. Our results notably improve and complete those obtained in [FM12]. The key point is to properly quantify the contribution of the diffusion term to the concentration regime. We also derive a general non-asymptotic deviation bound for the difference between a function of the trajectory of a continuous Euler scheme associated to a diffusion process and its mean.

7.1 Introduction

In this work, we derive transport-entropy inequalities and, as a consequence, non-asymptotic deviation estimates for the laws at a given time step of d -dimensional Euler-Maruyama schemes. More generally, the method we use here applies to stochastic schemes of the form

$$X_{n+1} = X_n + \gamma_{n+1}H(n, X_n, U_{n+1}), \quad n \geq 0, X_0 = x \in \mathbb{R}^d, \quad (7.1)$$

where $(\gamma_n)_{n \geq 1}$ is a deterministic positive sequence of time steps, the $(U_i)_{i \in \mathbb{N}^*}$ are i.i.d. \mathbb{R}^q -valued random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with law μ and the function $H : \mathbb{N} \times \mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}^d$ is a measurable function satisfying for all $x \in \mathbb{R}^d$, for all $n \in \mathbb{N}$, $H(n, x, \cdot) \in \mathcal{L}^1(\mu)$, and $\mu(du)$ -a.s., $H(n, \cdot, u)$ is continuous. Here and below, we will also assume that μ satisfies a *Gaussian concentration property*, that is there exists $\beta > 0$ such that for every real-valued 1-Lipschitz function f defined on \mathbb{R}^q and for all $\lambda \geq 0$:

$$\mathbb{E}[\exp(\lambda f(U_1))] \leq \exp(\lambda \mathbb{E}[f(U_1)] + \frac{\beta \lambda^2}{4}). \quad (GC(\beta))$$

It is well known that $(GC(\beta))$ implies the following deviation bound

$$\mathbb{P}[f(U_1) - \mathbb{E}[f(U_1)] \geq r] \leq \exp(-\frac{r^2}{\beta}) \quad \forall r \geq 0,$$

Examples of random variables satisfying this property include Gaussians, as well as bounded random variables. A characterization of $(GC(\beta))$ due to Djellout, Guillin and

Wu [DGW04] is given by Gaussian tail of U_1 , that is there exists $\varepsilon > 0$ such that $\mathbb{E}[\exp(\varepsilon|U_1|^2)] < +\infty$, see also Bolley and Villani [BV05] for another proof with a simple link between the involved constants. The two claims are actually equivalent.

We are interested in furthering the discussion, initiated in [FM12], about giving non asymptotic deviation bounds for the deviation between a function of an Euler like discretization scheme of a diffusion process at a fixed deterministic date and its mean. Under some mild assumptions, in particular the assumption that the function $u \mapsto H(n, x, u)$ is lipschitz uniformly in space and time, it is proved in [FM12] that recursive schemes share the Gaussian concentration property of the innovation.

In the present work, we point out the contribution of the diffusion term to the concentration rate which to our knowledge is new. This covers many situations and gives rise to different regimes ranging from exponential to Gaussian. We also derive a general non-asymptotic deviation bound for the difference between a function of the trajectory of a *continuous Euler scheme* associated to a diffusion process and its mean. It turns out that, under mild assumptions, the concentration regime is log-normal. In the article [FF13], we also apply the method to stochastic approximation schemes, and obtain similar results.

7.1.1 Euler-like Scheme of a Diffusion Process

We consider a Brownian diffusion process $(X_t)_{t \geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, satisfying the usual conditions, and solution to the following stochastic differential equation (SDE)

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad (SDE_{b,\sigma})$$

where $(W_t)_{t \geq 0}$ is a q -dimensional $(\mathcal{F}_t)_{t \geq 0}$ Brownian motion and the coefficients b, σ are assumed to be uniformly Lipschitz continuous in space and measurable in time.

A basic problem in Numerical Probability is to compute quantities like $\mathbb{E}_x[f(X_T)]$ for a given Lipschitz continuous function f and a fixed deterministic time horizon T using Monte Carlo simulation. For instance, it appears in mathematical finance and represents the price of a European option with maturity T when the dynamics of the underlying asset is given by $(SDE_{b,\sigma})$. To this end, we first introduce some discretization schemes of $(SDE_{b,\sigma})$ that can be easily simulated. For a fixed time step $\Delta = T/N$, $N \in \mathbb{N}^*$, we set $t_i := i\Delta$, for all $i \in \mathbb{N}$ and define an Euler like scheme by

$$X_0^\Delta = x, \quad \forall i \in \llbracket 0, N-1 \rrbracket, \quad X_{t_{i+1}}^\Delta = X_{t_i}^\Delta + b(t_i, X_{t_i}^\Delta)\Delta + \sigma(t_i, X_{t_i}^\Delta)\Delta^{1/2}U_{i+1}, \quad (7.2)$$

where $(U_i)_{i \in \mathbb{N}^*}$ is a sequence of \mathbb{R}^q -valued i.i.d. random variables with law μ satisfying : $\mathbb{E}[U_1] = 0_q$, $\mathbb{E}[U_1 U_1^*] = I_q$, where U_1^* denotes the transpose of the column vector U_1 and $0_q, I_q$ respectively denote the zero vector of \mathbb{R}^q and the identity matrix of $\mathbb{R}^q \otimes \mathbb{R}^q$. We also assume that μ satisfies $(GC(\beta))$ for some $\beta > 0$. The main advantage of such a situation is that it includes the case of the standard Euler scheme where $U_1 \stackrel{d}{=} \mathcal{N}(0, I_q)$ and the case of the Bernoulli law where $U_1 \stackrel{d}{=} (B_1, \dots, B_q)$, $(B_k)_{k \in \llbracket 1, q \rrbracket}$ are i.i.d random variables with law $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$, both satisfying $(GC(\beta))$ with $\beta = 2$.

The weak error $\mathcal{E}_D(f, \Delta, T, b, \sigma) = \mathbb{E}_x[f(X_T)] - \mathbb{E}_x[f(X_T^\Delta)]$ corresponds to the discretization error when replacing the diffusion X by its Euler scheme X^Δ for the computation of $\mathbb{E}_x[f(X_T)]$. Since the seminal work of [TT90], it is known that, under smoothness assumption on the coefficients b, σ , the *standard Euler scheme* produces a weak error of order Δ . In a hypoelliptic setting for the coefficients b and σ and for a bounded measurable function f , Bally and Talay [BT96] obtained the expected order using Malliavin calculus.

Let us also mention the recent work [AJKH14] where the authors study the *weak trajectorial error* using coupling techniques. More precisely, they prove that the Wasserstein distance between the law of a uniformly elliptic and one-dimensional diffusion process and the law of its *continuous Euler scheme* $X^{c,\Delta}$ with time step $\Delta := T/N$ is smaller than $\mathcal{O}(N^{-2/3+\epsilon})$, $\forall \epsilon > 0$.

The expansion of \mathcal{E}_D also allows to improve the convergence rate to 0 of the discretization error using Richardson-Romberg extrapolation techniques, see e.g. [TT90].

In order to have a global control of the numerical procedure for the computation of $\mathbb{E}_x[f(X_T)]$, it remains to approximate the expectation $\mathbb{E}_x[f(X_T^\Delta)]$ using a Monte Carlo estimator $M^{-1} \times \sum_{k=1}^M f((X_T^\Delta)^k)$ where the $((X_T^\Delta)^k)_{k \in [1, M]}$ are M independent copies of the scheme (7.2) starting at the initial value x at time 0. This gives rise to an *empirical error* defined by $\mathcal{E}_{Emp}(M, f, \Delta, T, b, \sigma) = \mathbb{E}_x[f(X_T^\Delta)] - M^{-1} \times \sum_{j=1}^M f((X_T^\Delta)^j)$. Consequently, the global error associated to the computation of $\mathbb{E}_x[f(X_T)]$ writes as

$$\begin{aligned} \mathcal{E}_{Glob}(M, \Delta) &= \mathbb{E}_x[f(X_T)] - \mathbb{E}_x[f(X_T^\Delta)] + \mathbb{E}_x[f(X_T^\Delta)] - \frac{1}{M} \times \sum_{j=1}^M f((X_T^\Delta)^j) \\ &:= \mathcal{E}_D(f, \Delta, T, b, \sigma) + \mathcal{E}_{Emp}(M, f, \Delta, T, b, \sigma). \end{aligned}$$

It is well-known that if $f(X_T^\Delta)$ belongs to $L^2(\mathbb{P})$ the central limit theorem provides an *asymptotic* rate of convergence of order $M^{1/2}$. Moreover, if $f(X_T^\Delta) \in L^3(\mathbb{P})$, a non-asymptotic result is given by the Berry-Essen theorem. However, in practical implementation, one is interested in obtaining deviation bounds in probability for a fixed M and a given threshold $r > 0$, that is explicitly controlling $\mathbb{P}(|\mathcal{E}_{Emp}(M, f, \Delta, T, b, \sigma)| \geq r)$.

In this context, Malrieu and Talay [MT06] obtained Gaussian deviation bounds in an ergodic framework and for a constant diffusion coefficient. Using optimal transportation techniques, Blower and Bolley [BB06] obtained Gaussian concentration inequalities and transportation inequalities for the joint law of the first n positions of a stochastic processes with state space some Polish space. Concerning the *standard* Euler scheme, Menozzi and Lemaire [LM10] obtained two-sided Gaussian bounds up to a systematic bias under the assumptions that the diffusion coefficient is uniformly elliptic, $\sigma\sigma^*$ is Hölder-continuous, bounded and that b is bounded. Frikha and Menozzi [FM12], getting rid of the non-degeneracy assumption on σ , recently obtained Gaussian deviation bound under the mild smoothness condition that b, σ are uniformly Lipschitz-continuous in space (uniformly in time) and that σ is bounded. It should be noted that it is the boundedness of σ that gives rise to the Gaussian concentration regime for the deviation of the *empirical error*.

In the current work, we get rid of the boundedness of σ and we only need the Gaussian concentration property of the innovation. We suppose that the coefficients satisfy the following smoothness and domination assumptions

- (HS)** The coefficients b, σ are uniformly Lipschitz continuous in space uniformly in time.
(HD $_\alpha$) There exists a $\mathcal{C}^2(\mathbb{R}^d, \mathbb{R}_+^*)$ function V satisfying $\exists C_V > 0, |\nabla V|^2 \leq C_V V$, $\eta := \frac{1}{2} \sup_{x \in \mathbb{R}^d} \|\nabla^2 V(x)\| < +\infty$ and $\exists \alpha \in (0, 1]$, such that for all $x \in \mathbb{R}^d$,

$$\exists C_b > 0, \quad \sup_{t \in [0, T]} |b(t, x)|^2 \leq C_b V(x), \quad , \quad \exists C_\sigma > 0, \quad \sup_{t \in [0, T]} \text{Tr}(a(t, x)) \leq C_\sigma V^{1-\alpha}(x).$$

where $a = \sigma\sigma^*$.

The idea behind assumption **(HD $_\alpha$)** is to parameterize the growth of the diffusion coefficient in order to quantify its contribution to the concentration regime. Indeed, under

(HS) and **(HD $_{\alpha}$)**, with $\alpha \in [1/2, 1]$, and if the innovations satisfy $(GC(\beta))$, for some positive β , we derive non-asymptotic deviation bounds for the empirical error $\mathcal{E}_{Emp}(M, f, \Delta, T, b, \sigma)$ ranging from exponential (if $\alpha = 1/2$) to Gaussian (if $\alpha = 1$) regimes. Therefore, we greatly improve the results obtained in [FM12].

Our approach here is different from [FM12]. Indeed, in [FM12], the key tool consists in writing the deviation using the same kind of decompositions that are exploited in [TT90] for the analysis of the discretization error. In the current work, we will use the fact that the Euler-like scheme (7.2) defines an inhomogenous Markov chain having Feller transitions P_k , $k = 0, \dots, N - 1$, defined for non negative or bounded Borel function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ by

$$P_k(f)(x) = \mathbb{E} \left[f(X_{t_{k+1}}^{\Delta}) \mid X_{t_k}^{\Delta} = x \right] = \mathbb{E} \left[f \left(x + b(t_k, x)\Delta + \sigma(t_k, x)\Delta^{1/2}U \right) \right].$$

For every $k, p \in \{0, \dots, N - 1\}$, $k \leq p$, we also define the iterative kernels $P_{k,p}$ by

$$P_{k,p}(f)(x) = P_k \circ \dots \circ P_{p-1}(f)(x) = \mathbb{E} \left[f(X_{t_p}^{\Delta}) \mid X_{t_k}^{\Delta} = x \right].$$

Now using that the law μ of the innovation satisfies $(GC(\beta))$ for some positive β , for every 1-Lipschitz function f and for all $\lambda \geq 0$, we obtain

$$\begin{aligned} P_{N-1}(\exp(\lambda f))(x) &= \mathbb{E} \left[\exp \left(\lambda f \left(x + b(t_{N-1}, x)\Delta + \sigma(t_{N-1}, x)\Delta^{1/2}U \right) \right) \right] \\ &\leq \exp \left(\lambda P_{N-1}(f)(x) + \beta \frac{\lambda^2}{4} \Delta |\sigma(t_{N-1}, x)|^2 \right) \end{aligned}$$

If σ is bounded, the Gaussian concentration property will readily follow provided the iterated kernel functions $P_{k,p}(f)$ are uniformly Lipschitz. Under the mild smoothness assumption **(HS)**, this can be easily derived, see Proposition 7.17. Otherwise, using **(HD $_{\alpha}$)**, we obtain

$$P_{N-1}(\exp(\lambda f))(x) \leq \exp \left(\lambda P_{N-1}(f)(x) + \frac{C_{\sigma}\beta\Delta}{4} \lambda^2 V^{1-\alpha}(x) \right). \quad (7.3)$$

The last inequality is the first step of our analysis. To investigate the empirical error, the key idea is to exploit recursively from (7.3) that the increments of the scheme (7.2) satisfy $(GC(\beta))$ and to adequately quantify the contribution of the diffusion term $V^{1-\alpha}(x)$ to the concentration rate. Under **(HS)** and **(HD $_{\alpha}$)**, the latter is addressed using flow techniques and integrability results on the law of the scheme (7.2), see Propositions 7.13 and 7.18.

7.1.2 Transport-Entropy inequalities

As a by-product of our analysis, we derive transport-entropy inequalities for the law of both stochastic approximation schemes. We recall here basic definitions and properties. For a complete overview and recent developments in the theory of transport inequalities, the reader may refer to the recent survey [GL10]. We will denote by $\mathcal{P}(\mathbb{R}^d)$ the set of probability measures on \mathbb{R}^d .

For $p \geq 1$, we consider the set $\mathcal{P}_p(\mathbb{R}^d)$ of probability measures with finite moment of order p . The Wasserstein metric $W_p(\mu, \nu)$ of order p between two probability measures $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ is defined by

$$W_p^p(\mu, \nu) = \inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \pi(dx, dy) : \pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d), \pi_0 = \mu, \pi_1 = \nu \right\}$$

where π_0 and π_1 are two probability measures standing for the first and second marginals of $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$. For $\mu \in \mathcal{P}(\mathbb{R}^d)$, we define the relative entropy w.r.t $\nu \in \mathcal{P}(\mathbb{R}^d)$ as

$$H(\mu, \nu) = \int_{\mathbb{R}^d} \log \left(\frac{d\mu}{d\nu} \right) d\mu$$

if $\mu \ll \nu$ and $H(\mu, \nu) = +\infty$ otherwise. We are now in position to define the notion of transport-entropy inequality. Here as below, $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a convex, increasing function with $\Phi(0) = 0$.

Definition 7.1. A probability measure μ on \mathbb{R}^d satisfies a transport-entropy inequality with function Φ if for all $\nu \in \mathcal{P}(\mathbb{R}^d)$, one has

$$\Phi(W_1(\nu, \mu)) \leq H(\nu, \mu)$$

For the sake of simplicity, we will write that μ satisfies T_Φ .

The following proposition comes from Corollary 3.4. of [GL10].

Proposition 7.2. *The following propositions are equivalent :*

- *The probability measure μ satisfies T_Φ .*
- *For all 1-Lipschitz function f , one has*

$$\forall \lambda \geq 0, \quad \int \exp(\lambda f) d\mu \leq \exp \left(\lambda \int f d\mu + \Phi^*(\lambda) \right),$$

where Φ^* is the monotone conjugate of Φ defined on \mathbb{R}_+ as $\Phi^*(\lambda) = \sup_{\rho \geq 0} \{\lambda\rho - \Phi(\rho)\}$.

Such transport-entropy inequalities are very attractive especially from a numerical point of view since they are related to the concentration of measure phenomenon which allows to establish non-asymptotic deviation estimates. The three next results put an emphasis on this point. Suppose that $(X_n)_{n \geq 1}$ is a sequence of i.i.d. \mathbb{R}^d -valued random variables with common law μ .

Corollary 7.3. *If μ satisfies T_Φ then for all 1-Lipschitz function f and for all $r \geq 0$, for all $M \geq 1$, one has*

$$\mathbb{P} \left(\left| \frac{1}{M} \sum_{k=1}^M f(X_k) - \mathbb{E}[f(X_1)] \right| \geq r \right) \leq 2 \exp(-M\Phi(r))$$

Deriving non-asymptotic deviation bounds for $W_1(\mu_M, \mu)$ is of interest for many applications in the fields of numerical probability and statistic. In its present form, next result is due to Gozlan and Leonard [GL07], Theorem 12.

Proposition 7.4. *If μ satisfies T_Φ then the empirical measure μ_M defined as $\mu_M = \frac{1}{M} \sum_{k=1}^M \delta_{X_k}$ satisfies the following concentration bound*

$$\mathbb{P}(W_1(\mu_M, \mu) \geq \mathbb{E}[W_1(\mu_M, \mu)] + r) \leq \exp(-M\Phi(r)).$$

where for $x \in \mathbb{R}^d$, δ_x stands for the Dirac mass at point x .

The quantity $\mathbb{E}[W_1(\mu_M, \mu)]$ will go to zero as M goes to infinity, by convergence of empirical measures, but we still need quantitative bounds. The next result is an adaptation of Theorem 10.2.1 in [RR98] on similar bounds but for the distance W_2 . For sake of completeness, we provide a proof in Appendix 7.4.

Proposition 7.5. *Assume that μ has a finite moment of order $d + 3$. Then, one has*

$$\mathbb{E}[W_1(\mu_M, \mu)] \leq C(d, \mu)M^{-1/(d+2)}$$

where

$$C(d, \mu) := 4\sqrt{d} + 2\sqrt{\int_{\mathbb{R}^d} (1 + |x|^{d+1})^{-1} dx \sqrt{2^{-2d} + 2^{3-d} \int |y|^{d+3} \mu(dy) + 2^{3-d} d(d+3)!}}.$$

This bound is not optimal in general, but has the advantage of having very explicit constants. In the case of a distribution with compact support, it has been shown in [BB13], Section 7, that $\mathbb{E}[W_1(\mu_M, \mu)]$ is of order $O(M^{-1/d})$, and that this is the optimal exponent in d when $d \geq 3$.

In view of Kantorovich-Rubinstein duality formula, namely

$$W_1(\mu, \nu) = \sup \left\{ \int f d\mu - \int f d\nu : [f]_1 \leq 1 \right\}$$

where $[f]_1$ denotes the Lipschitz-modulus of f , the latter result provides the following concentration bounds $\forall r \geq 0, \forall M \geq 1$

$$\mathbb{P} \left(\sup_{f: [f]_1 \leq 1} \left(\frac{1}{M} \sum_{k=1}^M f(X_k) - \mathbb{E}[f(X_1)] \right) \geq C(d, \mu)M^{-1/(d+2)} + r \right) \leq \exp(-M\Phi(r)).$$

Similar results were first obtained for different concentration regimes by Bolley, Guillin, Villani [BGV07] relying on a non-asymptotic version of Sanov's Theorem. Some of these results have also been derived by Boissard [Boi11] using concentration inequalities, and were also extended to ergodic Markov chains up to some contractivity assumptions in the Wasserstein metric on the transition kernel.

Some applications are proposed in [BGV07]. Such results can indeed provide non-asymptotic deviation bounds for the estimation of the density of the invariant measure of a Markov chain. Let us note that the (possibly large) constant $C(d, \mu)$ appears as a trade-off to obtain uniform deviations over all Lipschitz functions.

As a consequence of the transport-entropy inequalities obtained for the laws at a given time step of Euler like schemes and stochastic approximation algorithm, we will derive non-asymptotic deviation bounds in the Wasserstein metric.

7.2 Main Results

Theorem 7.6 (Transport-Entropy inequalities for Euler like schemes). *Denote by X_T^Δ the value at time T of the scheme (7.2) associated to the diffusion $(SDE_{b,\sigma})$ starting from x at time 0. Denote the Lipschitz modulus of b and σ appearing in the diffusion process $(SDE_{b,\sigma})$ by $[b]_1$ and $[\sigma]_1$, respectively and by μ_T^Δ the law of X_T^Δ . Assume that the innovations $(U_i)_{i \geq 1}$ in (7.2) satisfy $(GC(\beta))$ for some $\beta > 0$ and that the coefficients b, σ satisfy **(HS)** and **(HD $_\alpha$)** for $\alpha \in [\frac{1}{2}, 1]$.*

Then, μ_T^Δ satisfies $T_{\Phi_\alpha^}$ with $\Phi_\alpha^*(\lambda) = \sup_{\rho \geq 0} \{\lambda\rho - \Phi_\alpha(\rho)\}$ and one has :*

- If $\alpha \in (\frac{1}{2}, 1]$, for all $\rho \geq 0$

$$\Phi_\alpha(\rho) = \Psi_\alpha(T, \Delta, b, \sigma, x)(\rho^2 \vee \rho^{\frac{2\alpha}{2\alpha-1}}),$$

– If $\alpha = \frac{1}{2}$, for all $\rho \in [0, \varphi(T, b, \sigma, \Delta)^{-1/2} \lambda_{3.2})$

$$\Phi_{1/2}(\rho) = K_{3.2} \frac{(\rho \varphi(T, b, \sigma, \Delta)^{1/2} / \lambda_{3.2})^2}{1 - (\rho \varphi(T, b, \sigma, \Delta)^{1/2} / \lambda_{3.2})}.$$

Moreover, we have $\Psi_\alpha(T, \Delta, b, \sigma, x) = K_{3.1}(\varphi(T, b, \sigma, \Delta)^2 \vee \varphi(T, b, \sigma, \Delta)^{\frac{\alpha}{2\alpha-1}})$, $\varphi(T, b, \sigma, \Delta) = C_\sigma \beta^{\frac{(1+C(\Delta)\Delta)}{4C(\Delta)}} e^{3C(\Delta)T}$, $C(\Delta) := 2[b]_1 + [\sigma]_1^2 + \Delta[b]_1^2$, the constants $K_{3.1}$, $\lambda_{3.2}$ and $K_{3.2}$ being defined in Corollaries 7.14 and 7.16 respectively.

Note that in the above theorem, we do not need any non-degeneracy condition on the diffusion coefficient.

In the case $\alpha \in (\frac{1}{2}, 1]$, one easily gets the following explicit formula :

- If $\lambda \in [0, 2\Psi]$, then $\Phi_\alpha^*(\lambda) = \frac{1}{4\Psi} \lambda^2$;
- If $\lambda \in [\frac{2\alpha}{2\alpha-1} \Psi, +\infty)$, then $\Phi_\alpha^*(\lambda) = \frac{1}{2\alpha} \left(\frac{2\alpha-1}{2\alpha\Psi} \right)^{2\alpha-1} \lambda^{2\alpha}$;
- If $\lambda \in (2\Psi, \frac{2\alpha}{2\alpha-1} \Psi)$, then $\Phi_\alpha^*(\lambda) = \lambda - \Psi$.

Let us note that the linear behavior of Φ_α^* on a small interval is due to the fact that Φ_α is not \mathcal{C}^1 . One may want to replace $\rho^2 \vee \rho^{\frac{2\alpha}{2\alpha-1}}$ by $\rho^2 + \rho^{\frac{2\alpha}{2\alpha-1}}$ (up to a factor 2) in the expression of Φ_α . However, in this case, an explicit expression for Φ_α^* does not exist (except for the case $\alpha = 1$) and only its asymptotic behavior can be derived so that one is led to compute it numerically in practical situations.

In the case $\alpha = 1/2$, tedious but simple computations show that

$$\Phi_{1/2}^*(\lambda) = \left(\left(1 + \frac{\lambda_{3.2}}{K_{3.2} \varphi(T, b, \sigma, \Delta)^{1/2}} \lambda \right)^{\frac{1}{2}} - 1 \right)^2.$$

This behavior corresponds to a concentration profile that is Gaussian at short distance, and exponential at large distance.

Remark 7.7. The order of magnitude of our bounds is actually optimal in α under our general assumptions. For example, if we consider the diffusion process $dX_t = (1 + X_t^2)^{(1-\alpha)/2} dB_t$, then the process $Y_t = V(X_t)$, with $V(x) := \int_0^x (1 + s^2)^{(\alpha-1)/2} ds$, satisfies the SDE $dY_t = dB_t + b(Y_t)dt$, where b is a bounded drift. This process therefore has the same concentration properties as a Brownian motion, which are known to be Gaussian. From this, we deduce

$$\mathbb{P}_x(X_t \geq r) = \mathbb{P}_x(Y_t \geq V(r)) \leq \exp(-cV(r)^2).$$

This is indeed the order of magnitude of the concentration bounds given by Theorem 7.6.

Corollary 7.8. (Non-asymptotic deviation bounds) Under the same assumptions as Theorem 7.6, one has :

- for all real-valued 1-Lipschitz function f defined on \mathbb{R}^d , for all $\alpha \in [1/2, 1]$ for all $M \geq 1$ and all $r \geq 0$,

$$\mathbb{P}_x \left(\left| \frac{1}{M} \sum_{k=1}^M f((X_T^\Delta)^k) - \mathbb{E}_x[f(X_T^\Delta)] \right| \geq r \right) \leq 2 \exp(-M\Phi_\alpha^*(r)),$$

- for all $\alpha \in [1/2, 1]$, for all $M \geq 1$ and all $r \geq 0$,

$$\mathbb{P}_x \left(\sup_{f:|f|_1 \leq 1} \left(\frac{1}{M} \sum_{k=1}^M f((X_T^\Delta)^k) - \mathbb{E}_x[f(X_T^\Delta)] \right) \geq \frac{C(d, \mu_T^\Delta)}{M^{1/(d+2)}} + r \right) \leq \exp(-M\Phi_\alpha^*(r)),$$

where the $((X_T^\Delta)^k)_{1 \leq k \leq M}$ are M independent copies of the scheme (7.2).

The constant $C(d, \mu_T^\Delta)$ depends on the moment of order $d+3$ of μ_T^Δ . Hence, an explicit control in terms of x, b, σ, Δ can be easily obtained under our general assumptions. We leave the computational details to the reader.

Remark 7.9 (Extension to smooth functions of a finite number of time step). The previous transport-inequalities and non-asymptotic bounds could be extended to smooth functions of a finite number of time step such as the maximum of a scalar Euler like scheme. In that case, it suffices to introduce the additional state variable $(M_{t_i}^\Delta)_{i \geq 1} := (\max_{k \in \llbracket 0, i \rrbracket} X_{t_k}^\Delta)_{i \geq 1}$. Now, the couple $(X_{t_i}^\Delta, M_{t_i}^\Delta)_{1 \leq i \leq N}$ is Markovian and similar arguments could be easily extended to the couple for Lipschitz functions of both variables.

Remark 7.10 (Transport-Entropy inequalities for the law of a diffusion process). The previous transport-inequalities and non-asymptotic bounds could be extended to the law at time T of the diffusion process solution to $(SDE_{b,\sigma})$ by passing to the limit $\Delta \rightarrow 0$. Indeed, it is well-known that under **(HS)**, one has $X_T^\Delta \xrightarrow{a.s.} X_T$, as $\Delta \rightarrow 0$ and by Lebesgue theorem, one deduces from the first result of Corollary 7.8 that the empirical error (empirical mean) of X_T itself satisfies a non-asymptotic deviation bound with a similar deviation function (just pass to the limit $\Delta \rightarrow 0$ in all constants). Then, using Corollary 5.1 in [GL10] (equivalence between deviation of the empirical mean and transport-entropy inequalities), one easily derives that the law of X_T satisfies a similar transport-entropy inequalities when $\alpha \in (1/2, 1]$.

We want to point out that it is the growth of σ that gives the concentration regime ranging from Gaussian concentration bound if $\alpha = 1$ to exponential when $\alpha = \frac{1}{2}$. However, in many popular models in finance, the diffusion coefficient is linear, for instance practitioners often have to deal with Black-Scholes like dynamics of the form

$$X_t = x_0 + \int_0^t b(X_s)X_s ds + \int_0^t \sigma(X_s)X_s dW_s$$

for smooth, bounded coefficients b, σ . This corresponds to assumption **(HD $_\alpha$)** where $\alpha = 0$ and $V(x) = 1 + |x|^2$, $x \in \mathbb{R}^d$. For the estimation of $\mathbb{E}_x[f(X_T^\Delta)]$ for a Lipschitz function $f : \mathbb{R}^d \rightarrow \mathbb{R}$, or even in more general situations, the estimation of $\mathbb{E}_x[f(X^\Delta)]$ for a Lipschitz function $f : \mathcal{C} \rightarrow \mathbb{R}$, where $\mathcal{C} := \mathcal{C}([0, T], \mathbb{R}^d)$ stands for the space of \mathbb{R}^d -valued continuous functions on $[0, T]$, equipped with the uniform norm $\|f\|_\infty := \sup_{0 \leq t \leq T} |f(t)|$, the expected concentration is the log-normal one. To deal with the latter case, we consider the continuous Euler scheme $X^{c,\Delta}$ associated to $(SDE_{b,\sigma})$ and writing

$$\forall t \in [0, T], X_t^{c,\Delta} = x + \int_0^t b(\phi(s), X_{\phi(s)}^{c,\Delta}) ds + \int_0^t \sigma(\phi(s), X_{\phi(s)}^{c,\Delta}) dW_s, \quad x \in \mathbb{R}^d. \quad (7.4)$$

where we set $\phi(t) := t_i$ for $t_i \leq t < t_{i+1}$, $i \in \mathbb{N}$. The next result provides a general non-asymptotic deviation bound for the empirical error under very mild assumptions.

Theorem 7.11 (General non-asymptotic deviation bounds). *Denote by $X^{c,\Delta} := (X_t^{c,\Delta})_{0 \leq t \leq T}$ the path of the scheme (7.4) with step Δ starting from point x at time 0. Assume that $\forall t \in [0, T]$, the coefficients $b(t, \cdot)$ and $\sigma(t, \cdot)$ are continuous functions in x and that they satisfy the linear growth assumption :*

$$\forall x \in \mathbb{R}^d, \quad \sup_{t \in [0, T]} |b(t, x)| \leq C_b(1 + |x|), \quad \sup_{t \in [0, T]} Tr(a(t, x)) \leq C_\sigma(1 + |x|^2).$$

Then, for all 1-Lipschitz function $f : \mathcal{C} \rightarrow \mathbb{R}$, for all $M \in \mathbb{N}^*$, for all $r \geq 0$, one has

$$\mathbb{P}_x \left(\left| \frac{1}{M} \sum_{k=1}^M f((X^{c,\Delta})^k) - \mathbb{E}_x[f(X^{c,\Delta})] \right| \geq r \right) \leq \begin{cases} 2e^{-\frac{r^2 M}{(2(1+|x|))^2 \exp(2\kappa(b,\sigma,T))}}, & \text{if } \frac{r\sqrt{M}}{2(1+|x|)} \leq e^{\kappa(b,\sigma,T)} \\ 2e^{-\frac{1}{4\kappa(b,\sigma,T)} \log\left(\frac{r^2 M}{(2(1+|x|))^2}\right)^2}, & \text{otherwise} \end{cases}$$

where $\kappa(b, \sigma, T) := 28(1 + (C_\sigma \vee C_b)T)$ and $((X^{c,\Delta})^k)_{1 \leq k \leq M}$ are M independent copies of the scheme (7.4). The result remains valid when one considers the path of the diffusion X solution to $(SDE_{b,\sigma})$ instead of the continuous Euler scheme.

Remark 7.12. We want to point out that though the constants appearing in the above non-asymptotic deviation bound are all-purpose and rough estimates, the decay in r is optimal. Indeed, if we select $b(t, x) = 0$, $\sigma(t, x) = \sigma x$, $\sigma > 0$, so that $X_t = x_0 \exp(\sigma W_t - \sigma^2 t/2)$, $M = 1$ and $f = \Pi_T$, where Π_T denotes the projection at time T , sharp bounds can be easily derived and it is plain to see that in this simple example the concentration regime for large values of r is the log-normal one and gaussian for small values of r .

7.3 Euler Scheme : Proof of the Main Results

In this section we will assume that **(HS)** and **(HD $_\alpha$)** are in force.

7.3.1 Proof of Theorem 7.6

The proof of Theorem 7.6 is divided into several propositions. Next proposition stresses the key role played by the Gaussian concentration property of the innovations law, that is a weaker concentration regime will lead to a lower integrability rate with respect to α .

Proposition 7.13. Denote by $X^\Delta := (X_{t_k}^\Delta)_{0 \leq k \leq N}$ the scheme (7.2) with time step $\Delta = T/N$, $N \in \mathbb{N}^*$ associated to the diffusion $(SDE_{b,\sigma})$ starting from x at time 0. Assume that the innovations $(U_i)_{i \geq 1}$ of (7.2) satisfy $(GC(\beta))$ for some $\beta > 0$. Then, there exists $\varepsilon_\beta > 0$ which only depends on the law μ such that for all $\lambda < \min(1, \varepsilon_\beta(2\eta\alpha C_\sigma T \exp(CT))^{-1})$, one has

$$\sup_{0 \leq n \leq N} \log \left(\mathbb{E}_x \left[\exp(\lambda V^\alpha(X_{t_n}^\Delta)) \right] \right) \leq \lambda \exp(CT) V^\alpha(x) + \frac{1}{2} \log \left(\mathbb{E} \left[\exp \left(\lambda 2\eta\alpha C_\sigma T \exp(CT) |U_1|^2 \right) \right] \right).$$

with $C := C(b, \sigma, V, \alpha, \Delta) = \alpha(C_V C_b)^{\frac{1}{2}} + \beta C_\sigma \alpha^2 (1 + 2\eta\Delta)^2 (C_V + C_b) + \alpha\eta C_b \Delta$.

Démonstration. Using the concavity of $x \mapsto x^\alpha$, $\alpha \in (0, 1]$, we have for all $k \geq 0$

$$V^\alpha(X_{t_{k+1}}^\Delta) - V^\alpha(X_{t_k}^\Delta) \leq \alpha V^{\alpha-1}(X_{t_k}^\Delta) (V(X_{t_{k+1}}^\Delta) - V(X_{t_k}^\Delta)).$$

A Taylor expansion of order 2 of the function V , recalling that $2\eta = \sup_{x \in \mathbb{R}^d} \|\nabla^2 V(x)\| < +\infty$, yields

$$V(X_{t_{k+1}}^\Delta) - V(X_{t_k}^\Delta) \leq \nabla V(X_{t_k}^\Delta) \cdot (X_{t_{k+1}}^\Delta - X_{t_k}^\Delta) + \eta |X_{t_{k+1}}^\Delta - X_{t_k}^\Delta|^2,$$

which together with the previous inequality leads to

$$\begin{aligned} V^\alpha(X_{t_{k+1}}^\Delta) - V^\alpha(X_{t_k}^\Delta) &\leq \alpha \Delta \frac{\nabla V(X_{t_k}^\Delta) \cdot b(t_k, X_{t_k}^\Delta)}{V^{1-\alpha}(X_{t_k}^\Delta)} + \alpha \Delta^{\frac{1}{2}} \frac{\nabla V(X_{t_k}^\Delta) \cdot \sigma(t_k, X_{t_k}^\Delta) U_{k+1}}{V^{1-\alpha}(X_{t_k}^\Delta)} \\ &\quad + \alpha \eta \Delta^2 \frac{|b(t_k, X_{t_k}^\Delta)|^2}{V^{1-\alpha}(X_{t_k}^\Delta)} + 2\alpha \eta \Delta^{\frac{3}{2}} \frac{b(t_k, X_{t_k}^\Delta) \cdot \sigma(t_k, X_{t_k}^\Delta) U_{k+1}}{V^{1-\alpha}(X_{t_k}^\Delta)} \\ &\quad + \alpha \eta \Delta \frac{|\sigma(t_k, X_{t_k}^\Delta) U_{k+1}|^2}{V^{1-\alpha}(X_{t_k}^\Delta)}. \end{aligned}$$

From **(HD_α)**, for all $(x, u) \in \mathbb{R}^d \times \mathbb{R}^q$, we clearly have $\sup_{t \in [0, T]} |\nabla V(x) \cdot b(t, x)| \leq (C_V C_b)^{\frac{1}{2}} V(x)$ and $\sup_{t \in [0, T]} |\sigma(t, x)u|^2 \leq C_\sigma V^{1-\alpha}(x)|u|^2$ which yields

$$\begin{aligned} V^\alpha(X_{t_{k+1}}^\Delta) &\leq V^\alpha(X_{t_k}^\Delta)(1 + \alpha(C_V C_b)^{\frac{1}{2}}\Delta + \alpha\eta C_b \Delta^2) \\ &\quad + \alpha\Delta^{\frac{1}{2}}(1 + 2\eta\Delta) \frac{(\nabla V(X_{t_k}^\Delta) + b(X_{t_k}^\Delta)) \cdot \sigma(X_{t_k}^\Delta)U_{k+1}}{V^{1-\alpha}(X_{t_k}^\Delta)} + C_\sigma \alpha\eta\Delta |U_{k+1}|^2. \end{aligned}$$

Using **(HD_α)**, $\forall x \in \mathbb{R}^d$ the functions $g(x, \cdot) : u \mapsto \frac{(\nabla V(x) + b(x)) \cdot \sigma(x)u}{V^{1-\alpha}(x)}$ are Lipschitz, and more precisely satisfy

$$\forall x \in \mathbb{R}^d, \quad \sup_{(u, u') \in (\mathbb{R}^q)^2} \frac{|g(x, u) - g(x, u')|}{|u - u'|} \leq (C_V^{1/2} + C_b^{1/2})C_\sigma^{1/2}V^{\frac{\alpha}{2}}(x).$$

Hence, from the Cauchy Schwarz inequality and since the law of the innovations satisfy **(GC(β))** for some $\beta > 0$, there exists $\epsilon_\beta > 0$ such that for $\lambda < \min(1, \epsilon_\beta(2\eta\alpha C_\sigma \Delta)^{-1})$, one has

$$\begin{aligned} \mathbb{E} \left[\exp(\lambda V^\alpha(X_{t_{k+1}}^\Delta)) \middle| \mathcal{F}_{t_k} \right] &\leq \exp(\lambda V^\alpha(X_{t_k}^\Delta)(1 + \alpha(C_V C_b)^{\frac{1}{2}}\Delta + \alpha\eta C_b \Delta^2)) \\ &\quad \times \mathbb{E} \left[\exp(2\lambda\alpha\Delta^{\frac{1}{2}}(1 + 2\eta\Delta)g(X_{t_k}^\Delta, U_{k+1})) \middle| \mathcal{F}_{t_k} \right]^{\frac{1}{2}} \\ &\quad \times \mathbb{E} \left[\exp(2\lambda\eta\alpha C_\sigma \Delta |U_{k+1}|^2) \middle| \mathcal{F}_{t_k} \right]^{\frac{1}{2}} \\ &\leq \exp(\lambda V^\alpha(X_{t_k}^\Delta)(1 + \alpha(C_V C_b)^{\frac{1}{2}}\Delta + \alpha\eta C_b \Delta^2)) \\ &\quad \times \exp(\lambda^2 \beta \alpha^2 \Delta (1 + 2\eta\Delta)^2 (C_V + C_b) C_\sigma V^\alpha(X_{t_k}^\Delta)) \\ &\quad \times \mathbb{E} \left[\exp(2\lambda\eta\alpha C_\sigma \Delta |U_1|^2) \right]^{\frac{1}{2}} \\ &\leq \exp(\lambda C(\Delta) V^\alpha(X_{t_k}^\Delta)) \mathbb{E} \left[\exp(2\lambda\eta\alpha C_\sigma \Delta |U_1|^2) \right]^{\frac{1}{2}}, \end{aligned}$$

where $C(\Delta) := 1 + \Delta \left(\alpha(C_V C_b)^{\frac{1}{2}} + \beta C_\sigma \alpha^2 (1 + 2\eta\Delta)^2 (C_V + C_b) + \alpha\eta C_b \Delta \right)$. Now define $V_k = \frac{V^\alpha(X_{t_k}^\Delta)}{C(\Delta)^k}$, for $k \in \{0, \dots, N\}$. Taking expectation in both sides of the previous inequality clearly implies

$$\mathbb{E} [\exp(\lambda V_{k+1})] \leq \mathbb{E} [\exp(\lambda V_k)] \mathbb{E} \left[\exp \left(\lambda \frac{2\eta\alpha C_\sigma \Delta}{C(\Delta)^{k+1}} |U_1|^2 \right) \right]^{\frac{1}{2}}$$

and by a straightforward induction, for $n \in \{0, \dots, N\}$ we have

$$\mathbb{E} [\exp(\lambda V_n)] \leq \exp(\lambda V_0) \prod_{k=0}^{n-1} \mathbb{E} \left[\exp \left(\lambda \frac{2\eta\alpha C_\sigma \Delta}{C(\Delta)^{k+1}} |U_1|^2 \right) \right]^{\frac{1}{2}},$$

which finally yields, for $\lambda < \min(1, \epsilon_\beta(2\eta\alpha C_\sigma \Delta C(\Delta)^n)^{-1})$,

$$\mathbb{E} \left[\exp(\lambda V^\alpha(X_{t_n}^\Delta)) \right] \leq \exp(\lambda C(\Delta)^n V^\alpha(x)) \prod_{k=0}^{n-1} \mathbb{E} \left[\exp \left(\lambda 2\eta\alpha C_\sigma \Delta C(\Delta)^{k+1} |U_1|^2 \right) \right]^{\frac{1}{2}}.$$

Observe now that $C(\Delta)^N \leq \exp(CT)$ with $C := C(b, \sigma V, \alpha, \Delta) = \alpha(C_V C_b)^{\frac{1}{2}} + \beta C_\sigma \alpha^2 (1 + 2\eta\Delta)^2 (C_V + C_b) + \alpha\eta C_b \Delta$. Using Jensen's inequality, the latter bound clearly provides the following control of the quantity of interest for $\lambda < \min(1, \epsilon_\beta(2\eta\alpha C_\sigma T \exp(CT))^{-1})$

$$\sup_{0 \leq n \leq N} \log \left(\mathbb{E} \left[\exp(\lambda V^\alpha(X_{t_n}^\Delta)) \right] \right) \leq \lambda \exp(CT) V^\alpha(x) + \frac{1}{2} \log \left(\mathbb{E} \left[\exp \left(\lambda 2\eta\alpha C_\sigma T \exp(CT) |U_1|^2 \right) \right] \right).$$

□

Corollary 7.14. *Assume that the assumptions of Proposition 7.13 are satisfied. Then, there exists a constant $K_{3.1}$ such that for all $\alpha \in (\frac{1}{2}, 1]$, one has*

$$\forall \lambda \geq 0, \quad \sup_{0 \leq n \leq N} \log \left(\mathbb{E}_x \left[\exp(\lambda V^{1-\alpha}(X_{t_n}^\Delta)) \right] \right) \leq K_{3.1}(\lambda \vee \lambda^{\frac{\alpha}{2\alpha-1}}).$$

Remark 7.15. The constant $K_{3.1}$ can be explicitly computed. Indeed, one has $K_{3.1} := \max(\Psi_1(T, \Delta, x, b, \sigma), \Psi_2(T, \Delta, x, b, \sigma))$ with

$$\begin{aligned} \Psi_1(T, \Delta, x, b, \sigma) &:= e^{\frac{2\alpha-1}{\alpha} \rho^{-\frac{1-\alpha}{2\alpha-1}}} \exp \left(\rho \frac{1-\alpha}{\alpha} e^{CT} V^\alpha(x) + \frac{1}{2} \log \mathbb{E}[e^{\frac{\varepsilon_\beta(1-\alpha)}{2\alpha} |U|^2}] \right) \\ &\quad + \left(V^{1-\alpha}(x) + \left(\frac{C_\sigma \mathbb{E}[|U|^2]}{K} \right)^{\frac{1-\alpha}{\alpha}} \right) e^{(1-\alpha)KT}, \\ \Psi_2(T, \Delta, x, b, \sigma) &:= \rho^{-\frac{1-\alpha}{2\alpha-1}} \frac{2\alpha-1}{\alpha} + \rho \frac{1-\alpha}{\alpha} e^{CT} V^\alpha(x) + \frac{1}{2} \log \mathbb{E} \left[\exp \left(\frac{\varepsilon_\beta(1-\alpha)}{2\alpha} |U|^2 \right) \right], \\ \rho &:= \frac{1}{2} \min(1, \varepsilon_\beta(2\eta\alpha C_\sigma T \exp(CT))^{-1}), \\ C &:= C(b, \sigma V, \alpha, \Delta) = \alpha(C_V C_b)^{\frac{1}{2}} + \beta C_\sigma \alpha^2 (1 + 2\eta\Delta)^2 (C_V + C_b) + \alpha\eta C_b \Delta \\ K &:= K(V, b, \Delta) = (C_V C_b)^{\frac{1}{2}} + \eta C_b \Delta \end{aligned}$$

Démonstration. For $\lambda \in [0, 1]$, one has

$$\begin{aligned} \mathbb{E}_x[\exp(\lambda V^{1-\alpha}(X_{t_n}^\Delta))] &= 1 + \lambda \mathbb{E}_x[V^{1-\alpha}(X_{t_n}^\Delta)] + \sum_{k \geq 2} \frac{\lambda^k}{k!} \mathbb{E}_x[V^{(1-\alpha)k}(X_{t_n}^\Delta)] \\ &\leq 1 + \lambda \mathbb{E}_x[V^{1-\alpha}(X_{t_n}^\Delta)] + \lambda \sum_{k \geq 0} \frac{1}{k!} \mathbb{E}_x[V^{(1-\alpha)k}(X_{t_n}^\Delta)] \\ &\leq \exp \left(\lambda (\mathbb{E}_x[V^{1-\alpha}(X_{t_n}^\Delta)] + \mathbb{E}_x[\exp(V^{1-\alpha}(X_{t_n}^\Delta))]) \right), \end{aligned}$$

Tedious but simple computations, in the spirit of Proposition 7.13, show that

$$\mathbb{E}_x[V^{1-\alpha}(X_{t_n}^\Delta)] \leq \mathbb{E}_x[V^\alpha(X_{t_n}^\Delta)]^{\frac{1-\alpha}{\alpha}} \leq \left(V^{1-\alpha}(x) + \left(\frac{C_\sigma \mathbb{E}[|U|^2]}{K} \right)^{\frac{1-\alpha}{\alpha}} \right) e^{(1-\alpha)KT}.$$

with $K := K(V, b, \Delta) = (C_V C_b)^{\frac{1}{2}} + \eta C_b \Delta$.

Thanks to the following Young inequality, for all $(\rho, x) \in \mathbb{R}_+ \times \mathbb{R}^d$, $V^{1-\alpha}(x) \leq \frac{1-\alpha}{\alpha} \rho V^\alpha(x) + \frac{2\alpha-1}{\alpha} \rho^{-\frac{1-\alpha}{2\alpha-1}}$, which is valid if $\alpha \in (\frac{1}{2}, 1]$, one has for $\rho = \underline{\rho} := \frac{1}{2} \min(1, \varepsilon_\beta(2\eta\alpha C_\sigma T \exp(CT))^{-1})$

$$\begin{aligned} \sup_{0 \leq n \leq N} \mathbb{E}_x[\exp(V^{1-\alpha}(X_{t_n}^\Delta))] &\leq e^{\frac{2\alpha-1}{\alpha} \rho^{-\frac{1-\alpha}{2\alpha-1}}} \sup_{0 \leq n \leq N} \mathbb{E}_x \left[\exp \left(\frac{1-\alpha}{\alpha} \rho V^\alpha(X_{t_n}^\Delta) \right) \right] \\ &\leq e^{\frac{2\alpha-1}{\alpha} \rho^{-\frac{1-\alpha}{2\alpha-1}}} e^{\rho \frac{1-\alpha}{\alpha} e^{CT} V^\alpha(x) + \frac{1}{2} \log \mathbb{E}[\exp(\frac{\varepsilon_\beta(1-\alpha)}{2\alpha} |U|^2)]} \end{aligned}$$

where we used Proposition 7.13 for the last inequality.

Now, for all $\lambda > 1$, using the Young inequality $\lambda V^{1-\alpha}(X_{t_n}^\Delta) \leq (\frac{2\alpha-1}{\alpha}) \rho^{-\frac{1-\alpha}{2\alpha-1}} \lambda^{\frac{\alpha}{2\alpha-1}} + (\frac{1-\alpha}{\alpha}) \rho V^\alpha(X_{t_n}^\Delta)$, valid for all $\rho > 0$ (to be chosen later on) and for all $\alpha \in (\frac{1}{2}, 1]$, one derives

$$\begin{aligned} \mathbb{E}_x[\exp(\lambda V^{1-\alpha}(X_{t_n}^\Delta))] &\leq \exp \left(\left(\frac{2\alpha-1}{\alpha} \right) \rho^{-\frac{1-\alpha}{2\alpha-1}} \lambda^{\frac{\alpha}{2\alpha-1}} \right) \mathbb{E}_x \left[\exp \left(\left(\frac{1-\alpha}{\alpha} \right) \rho V^\alpha(X_{t_n}^\Delta) \right) \right] \\ &\leq \exp \left(K \lambda^{\frac{\alpha}{2\alpha-1}} \right) \end{aligned}$$

with $K(\rho) := \frac{2\alpha-1}{\alpha}\rho^{-\frac{1-\alpha}{2\alpha-1}} + \log(\mathbb{E}_x [e^{(\frac{1-\alpha}{\alpha})\rho V^\alpha(X_{t_n}^\Delta)})]$ and $\frac{1-\alpha}{\alpha}\rho < \min(1, \varepsilon_\beta(2\eta\alpha C_\sigma T e^{CT})^{-1})$. We select $\rho = \underline{\rho}$ in the last inequality to complete the proof and use Proposition 7.13 to bound the quantity $K(\underline{\rho})$. \square

Corollary 7.16. *Under the same assumptions as Proposition 7.13, one has*

$$\forall \lambda \in [0, \lambda_{3.2}), \quad \sup_{0 \leq n \leq N} \log \left(\mathbb{E}_x \left[\exp(\lambda^2 V^{1/2}(X_{t_n}^\Delta)) \right] \right) \leq K_{3.2} \frac{(\lambda/\lambda_{3.2})^2}{1 - (\lambda/\lambda_{3.2})}$$

with $K_{3.2} := \lambda_{3.2}^2 e^{CT} (2V^{\frac{1}{2}}(x) + 2\eta\alpha C_\sigma \mathbb{E}[|U_1|^2]T)$ and $\lambda_{3.2}$ satisfies $\mathbb{E}[e^{\lambda_{3.2}^2 2\eta\alpha C_\sigma T \exp(CT)|U_1|^2}] \leq 2$.

Démonstration. By definition of $\lambda_{3.2}$, we have $\forall k \geq 1, \lambda_{3.2}^{2k} (2\eta\alpha C_\sigma T \exp(CT))^k \mathbb{E}[|U_1|^{2k}] \leq 2k!$. Consequently, setting temporarily $C_1 := \exp(CT)V^{1/2}(x), C_2 := 2\eta\alpha C_\sigma T \exp(CT)$ for sake of simplicity, simple computations show that for $\lambda < \lambda_{3.2}$

$$\begin{aligned} \log \mathbb{E} \left[\exp \left(\lambda^2 C_2 |U_1|^2 \right) \right] - \lambda^2 C_2 \mathbb{E}[|U_1|^2] &= \log \left(1 + \sum_{k \geq 1} \frac{\lambda^{2k} C_2^k \mathbb{E}[|U_1|^{2k}]}{k!} \right) - \lambda^2 C_2 \mathbb{E}[|U_1|^2] \\ &\leq \sum_{k \geq 2} \frac{\lambda^{2k} C_2^k \mathbb{E}[|U_1|^{2k}]}{k!} \\ &\leq 2 \sum_{k \geq 2} \left(\frac{\lambda}{\lambda_{3.2}} \right)^{2k} \leq 2 \frac{(\lambda/\lambda_{3.2})^2}{1 - (\lambda/\lambda_{3.2})} \end{aligned}$$

hence, using Proposition 7.13 for $\alpha = \frac{1}{2}$, we clearly get

$$\begin{aligned} \sup_{0 \leq n \leq N} \log \left(\mathbb{E}_x \left[\exp(\lambda^2 V^{1/2}(X_{t_n}^\Delta)) \right] \right) &\leq \lambda_{3.2}^2 \left(C_1 + \frac{C_2 \mathbb{E}[|U_1|^2]}{2} \right) (\lambda/\lambda_{3.2})^2 + \frac{(\lambda/\lambda_{3.2})^2}{1 - (\lambda/\lambda_{3.2})} \\ &\leq 2\lambda_{3.2}^2 \left(C_1 + \frac{C_2 \mathbb{E}[|U_1|^2]}{2} \right) \frac{(\lambda/\lambda_{3.2})^2}{1 - (\lambda/\lambda_{3.2})}. \end{aligned}$$

This completes the proof. \square

Proposition 7.17. *(Control of the Lipschitz modulus of iterative kernels) Denote the Lipschitz modulus of b and σ appearing in the diffusion process $(SDE_{b,\sigma})$ by $[b]_1$ and $[\sigma]_1$, respectively. Denote by P_k and $P_{k,p} = P_k \circ \dots \circ P_{p-1}$, $k, p \in \{0, \dots, N-1\}$, $k \leq p$ the (Feller) transition kernel and the iterative kernels of the Markov chain defined by the scheme (7.2), respectively. Then, for all real-valued Lipschitz function f and for all $k, p \in \{0, \dots, N-1\}$, $k \leq p$ the functions $P_k(f)$ are Lipschitz-continuous and one has*

$$[P_{k,p}(f)]_1 := \sup_{(x,x') \in (\mathbb{R}^d)^2} \frac{|P_{k,p}(f)(x) - P_{k,p}(f)(x')|}{|x - x'|} \leq [f]_1 (1 + C(b, \sigma, \Delta) \Delta)^{\frac{p-k}{2}}$$

where $[f]_1$ stands for the Lipschitz modulus of the function f and $C(b, \sigma, \Delta) = 2[b]_1 + [\sigma]_1^2 + \Delta[b]_1^2$.

Démonstration. Using the Cauchy Schwarz inequality and **(HS)**, for all $(x, y) \in (\mathbb{R}^d)^2$ and for all $k \in \{0, \dots, N-1\}$, one has

$$\begin{aligned} |P_k(f)(x) - P_k(f)(y)| &\leq \mathbb{E} \left[\left| f(x + b(t_k, x)\Delta + \sigma(t_k, x)U_1) - f(y + b(t_k, y)\Delta + \Delta^{\frac{1}{2}}\sigma(t_k, y)U_1) \right| \right] \\ &\leq [f]_1 \mathbb{E} \left[\left| x - y + (b(t_k, x) - b(t_k, y))\Delta + \Delta^{\frac{1}{2}}(\sigma(t_k, x) - \sigma(t_k, y))U_1 \right|^2 \right]^{\frac{1}{2}} \\ &\leq [f]_1 (1 + C(b, \sigma, \Delta) \Delta)^{\frac{1}{2}} |x - y|. \end{aligned}$$

A straightforward induction argument completes the proof. \square

Proposition 7.18. (*Control of the Laplace transform*) Denote by X_T^Δ the value at time T of the scheme (7.2) associated to the diffusion (SDE $_{b,\sigma}$). Assume that the innovations $(U_n)_{n \geq 1}$ in (7.2) satisfy (GC(β)) for some $\beta > 0$. Let f be a real-valued 1-Lipschitz-continuous function defined on \mathbb{R}^d . For all $\lambda \geq 0$ and for all $\alpha \in (\frac{1}{2}, 1]$, one has

$$\mathbb{E}_x \left[\exp(\lambda f(X_T^\Delta)) \right] \leq \exp(\lambda \mathbb{E}_x [f(X_T^\Delta)]) \exp \left(K_{3.1}(\varphi(T, b, \sigma, \Delta) \vee \varphi(T, b, \sigma, \Delta)^{\frac{\alpha}{2\alpha-1}})(\lambda^2 \vee \lambda^{\frac{2\alpha}{2\alpha-1}}) \right),$$

with $\varphi(T, b, \sigma, \Delta) := C_\sigma \beta \frac{(1+C(\Delta)\Delta)}{4C(\Delta)} e^{3C(\Delta)T}$ and $C(\Delta) := 2[b]_1 + [\sigma]_1^2 + \Delta[b]_1^2$.

If $\alpha = \frac{1}{2}$, for all $\lambda \in [0, \varphi(T, b, \sigma, \Delta)^{-1/2} \lambda_{3.2})$, one has

$$\mathbb{E}_x \left[\exp(\lambda f(X_T^\Delta)) \right] \leq \exp(\lambda \mathbb{E}_x [f(X_T^\Delta)]) \exp \left(K_{3.2} \frac{(\lambda \varphi(T, b, \sigma, \Delta)^{1/2} / \lambda_{3.2})^2}{1 - (\lambda \varphi(T, b, \sigma, \Delta)^{1/2} / \lambda_{3.2})} \right).$$

Démonstration. As mentioned earlier on in the introduction, we begin our proof using that the law μ of the innovation satisfies (GC(β)) and (HD $_\alpha$). Hence, for $\lambda \geq 0$ and $k \in \{0, \dots, N-1\}$, one has

$$\begin{aligned} P_k(\exp(\lambda f))(x) &= \mathbb{E} \left[\exp \left(\lambda f \left(x + b(t_k, x)\Delta + \sigma(t_k, x)\Delta^{1/2}U_{k+1} \right) \right) \right] \\ &\leq \exp \left(\lambda P_k(f)(x) + \beta \frac{\lambda^2}{4} [f]_1^2 \Delta |\sigma(t_k, x)|^2 \right) \\ &\leq \exp \left(\lambda P_k(f)(x) + C_\sigma \beta \frac{\lambda^2}{4} [f]_1^2 \Delta V^{1-\alpha}(x) \right). \end{aligned} \quad (7.5)$$

Taking expectation from both sides of the last inequality and using the Hölder inequality with conjugate exponents (p, q) (to be specified later on) leads to

$$\mathbb{E}_x \left[\exp(\lambda f(X_{t_{k+1}}^\Delta)) \right] \leq \mathbb{E}_x \left[\exp(\lambda p P_k(f)(X_{t_k}^\Delta)) \right]^{\frac{1}{p}} \mathbb{E}_x \left[\exp \left(\frac{q C_\sigma \beta}{4} \Delta \lambda^2 [f]_1^2 V^{1-\alpha}(X_{t_k}^\Delta) \right) \right]^{\frac{1}{q}}. \quad (7.6)$$

Now, we apply the last inequality for $f := P_{k+1,N}(f)$ and obtain

$$\begin{aligned} \mathbb{E}_x \left[\exp(\lambda P_{k+1,N}(f)(X_{t_{k+1}}^\Delta)) \right] &\leq \mathbb{E}_x \left[\exp(\lambda p P_{k,N}(f)(X_{t_k}^\Delta)) \right]^{\frac{1}{p}} \\ &\quad \times \mathbb{E}_x \left[\exp \left(\frac{q C_\sigma \beta}{4} \Delta \lambda^2 [P_{k+1,N}(f)]_1^2 V^{1-\alpha}(X_{t_k}^\Delta) \right) \right]^{\frac{1}{q}}. \end{aligned}$$

Consequently, an elementary induction yields

$$\begin{aligned} \mathbb{E}_x \left[\exp(\lambda f(X_T^\Delta)) \right] &= \mathbb{E}_x \left[\exp(\lambda P_{N,N}(f)(X_{t_N}^\Delta)) \right] \\ &\leq \mathbb{E}_x \left[\exp(\lambda p^N P_{0,N}(f)(x)) \right]^{\frac{1}{p^N}} \\ &\quad \times \prod_{k=0}^{N-1} \left(\mathbb{E}_x \left[\exp \left(\frac{C_\sigma \beta}{4} \lambda^2 q p^{2k} \Delta [P_{N-k,N}(f)]_1^2 V^{1-\alpha}(X_{t_{N-k-1}}^\Delta) \right) \right]^{\frac{1}{q}} \right)^{\frac{1}{p^k}} \\ &\leq \exp(\lambda \mathbb{E}_x [f(X_T^\Delta)]) \\ &\quad \times \exp \left(\sum_{k=0}^{N-1} \frac{1}{p^k} \frac{1}{q} \sup_{0 \leq n \leq N} \log \left(\mathbb{E}_x \left[e^{\frac{C_\sigma \beta}{4} \lambda^2 \Delta q p^{2N} (1+C(\Delta)\Delta)^N V^{1-\alpha}(X_{t_n}^\Delta)} \right] \right) \right) \end{aligned}$$

where we used Proposition 7.17 for the last inequality. Observe now that since (p, q) are conjugate exponents, we have $\frac{1}{q} \sum_{k=0}^{N-1} \frac{1}{p^k} = \frac{1}{q} (1 - \frac{1}{p^N}) \frac{1}{1 - \frac{1}{p}} \leq \frac{1}{q} \frac{p}{p-1} = 1$, so that

$$\mathbb{E}_x \left[\exp(\lambda f(X_T^\Delta)) \right] \leq \exp(\lambda \mathbb{E}_x [f(X_T^\Delta)]) e^{\sup_{0 \leq n \leq N} \log \left(\mathbb{E}_x \left[e^{\frac{C_\sigma \beta}{4} \lambda^2 \Delta q p^{2N} (1+C(\Delta)\Delta)^N V^{1-\alpha}(X_{t_n}^\Delta)} \right]} \right).$$

Setting $p := 1 + C(\Delta)\Delta$, $q = \frac{p}{p-1} = \frac{1+C(\Delta)\Delta}{C(\Delta)\Delta}$ and using the straightforward inequality $(1 + C(\Delta)\Delta)^{3N} \leq \exp(3C(\Delta)T)$, we derive

$$\mathbb{E}_x \left[\exp(\lambda f(X_T^\Delta)) \right] \leq \exp(\lambda \mathbb{E}_x [f(X_T^\Delta)]) e^{\sup_{0 \leq n \leq N} \log \left(\mathbb{E}_x \left[e^{\frac{C_\sigma \beta (1+C(\Delta)\Delta)}{4C(\Delta)} e^{3C(\Delta)T} \lambda^2 V^{1-\alpha}(X_{t_n}^\Delta)} \right]} \right).$$

We set $\varphi(T, b, \sigma, \Delta) := C_\sigma \beta \frac{(1+C(\Delta)\Delta)}{4C(\Delta)} e^{3C(\Delta)T}$. For $\alpha \in (\frac{1}{2}, 1]$, Corollary 7.14 clearly implies

$$\mathbb{E}_x \left[\exp(\lambda f(X_T^\Delta)) \right] \leq \exp(\lambda \mathbb{E}_x [f(X_T^\Delta)]) \exp \left(K_{3.1} (\varphi(T, b, \sigma, \Delta) \vee \varphi(T, b, \sigma, \Delta)^{\frac{\alpha}{2\alpha-1}}) (\lambda^2 \vee \lambda^{\frac{2\alpha}{2\alpha-1}}) \right)$$

and for $\alpha = \frac{1}{2}$, according to Proposition 7.16, for $\lambda < \varphi(T, b, \sigma, \Delta)^{-1/2} \lambda_{3.2}$, one has

$$\mathbb{E}_x \left[\exp(\lambda f(X_T^\Delta)) \right] \leq \exp(\lambda \mathbb{E}_x [f(X_T^\Delta)]) \exp \left(K_{3.2} \frac{(\lambda \varphi(T, b, \sigma, \Delta)^{1/2} / \lambda_{3.2})^2}{1 - (\lambda \varphi(T, b, \sigma, \Delta)^{1/2} / \lambda_{3.2})} \right).$$

□

7.3.2 Proof of Theorem 7.11

We will prove the result for the process X solution of $(SDE_{b,\sigma})$. The proof for the continuous Euler scheme is similar. The following lemma is standard. As will become clear in the discussion below, the only interest it holds is to know the explicit behavior with respect to p . This behavior is optimal as it can be readily checked by taking $X_t = x_0 \exp(\sigma W_t - \sigma^2 t/2)$, $t \in [0, T]$, see also remark 7.12.

Lemma 7.19. *Under the assumptions of Theorem 7.11, for all $p \geq 1$, one has*

$$\mathbb{E}_x \left[\sup_{0 \leq t \leq T} |X_t|^{2p} \right] \leq (1 + |x|)^{2p} \exp(26p^2(1 + (C_b \vee C_\sigma)T)).$$

Démonstration. Let $g : x \mapsto \sqrt{1 + |x|^2}$ satisfying for all $x \in \mathbb{R}^d$, $\nabla g(x) = g^{-1}(x)x$, $\nabla^2 g(x) = g^{-1}(x)I_d - g^{-3}(x)xx^*$ and $V : x \mapsto g^{2p}(x)$. We apply Itô's formula to the process $V(X_t)$ with $\nabla V(x) = 2pg(x)^{2p-1} \nabla g(x)$, $\nabla^2 V(x) = 2pg(x)^{2p-1} \nabla^2 g(x) + 2p(2p-1)g(x)^{2p-2} \nabla g(x) \nabla g(x)^*$ noticing that for all $t \in [0, T]$

$$\begin{aligned} \nabla V(x) \cdot b(t, x) + \frac{1}{2} \text{Tr}(\sigma^* \nabla^2 V \sigma)(t, x) &\leq 2pC_b g(x)^{2p-1} (1 + |x|) + \frac{1}{2} C_\sigma (1 + |x|^2) \|\nabla^2 V(x)\| \\ &\leq 4pC_b g(x)^{2p} \\ &\quad + \frac{1}{2} C_\sigma (1 + |x|^2) (4pg(x)^{2p-2} + 2p(2p-1)g(x)^{2p-2}) \\ &\leq 4p(C_b \vee C_\sigma) g(x)^{2p} + 2p(C_b \vee C_\sigma) g(x)^{2p} \\ &\quad + p(2p-1)(C_b \vee C_\sigma) g(x)^{2p} \\ &\leq 8p^2 (C_b \vee C_\sigma) V(x) \end{aligned}$$

we clearly obtain,

$$V(X_t^{\tau_m}) \leq V(x) + 8p^2(C_b \vee C_\sigma) \int_0^t V(X_s^{\tau_m}) ds + \int_0^{t \wedge \tau_m} (\nabla V^* \sigma)(X_s^{\tau_m}) dW_s, \quad (7.7)$$

where we classically introduced the stopping time $\tau_m := \inf \{t \geq 0 : |X_t - x| \geq m\}$ for $m \in \mathbb{N}^*$ and $X^{\tau_m} := (X_{t \wedge \tau_m})_{t \geq 0}$. The stochastic integral $M_t^m := \int_0^{t \wedge \tau_m} (\nabla V^* \sigma)(X_s^{\tau_m}) dW_s$ defines a continuous martingale so that taking expectation in the previous inequality clearly yields

$$\mathbb{E}_x[V(X_t^{\tau_m})] \leq V(x) + 8p^2(C_b \vee C_\sigma) \int_0^t \mathbb{E}_x[V(X_s^{\tau_m})] ds.$$

Now, using Gronwall's lemma we derive

$$\forall m \in \mathbb{N}^*, \quad \sup_{t \in [0, T]} \mathbb{E}_x[V(X_t^{\tau_m})] \leq (1 + |x|)^{2p} \exp(8p^2(C_b \vee C_\sigma)T)$$

As $\tau_m \rightarrow +\infty$ a.s., as $m \rightarrow +\infty$ (since $\sup_{s \in [0, t]} |X_s| < +\infty$) using Fatou's lemma, we finally obtain for all $p \geq 1$

$$\sup_{0 \leq t \leq T} \mathbb{E}_x[V(X_t)] = \sup_{0 \leq t \leq T} \mathbb{E}_x[g(X_t)^{2p}] \leq (1 + |x|)^{2p} \exp(8p^2(C_b \vee C_\sigma)T). \quad (7.8)$$

We then observe that Itô's formula also implies

$$\mathbb{E}_x[\sup_{0 \leq s \leq t} V(X_s^{\tau_m})] \leq V(x) + 8p^2(C_b \vee C_\sigma) \int_0^t \mathbb{E}_x[\sup_{0 \leq u \leq s} V(X_u^{\tau_m})] ds + \mathbb{E}_x[(M_t^m)^*] \quad (7.9)$$

where $(M_t^m)^* := \sup_{0 \leq s \leq t} M_s^m$. Combining Jensen's and Doob's inequalities, one clearly gets

$$\begin{aligned} \mathbb{E}_x[(M_t^m)^*]^2 &\leq \mathbb{E}_x[((M_t^m)^*)^2] \leq 4\mathbb{E}_x[(M_t^m)^2] \leq 16p^2 C_\sigma \int_0^t \mathbb{E}_x[g(X_s^{\tau_m})^{4p}] ds \\ &\leq 16p^2 C_\sigma T (1 + |x|)^{4p} \exp(32p^2(C_b \vee C_\sigma)T) \end{aligned}$$

where we used $\forall x \in \mathbb{R}^d$, $(\nabla V^* \sigma)^2(x) \leq 4p^2 C_\sigma g(x)^{4p-2} (1 + |x|^2) = 4p^2 C_\sigma g(x)^{4p}$ and (7.8) for the last inequality. Consequently, plugging the latter estimate into (7.9), one has for all $t \in [0, T]$

$$\begin{aligned} \mathbb{E}_x[\sup_{0 \leq s \leq t} V(X_s^{\tau_m})] &\leq V(x) + 4p(C_\sigma T)^{\frac{1}{2}} (1 + |x|)^{2p} \exp(16p^2(C_b \vee C_\sigma)T) \\ &\quad + 8p^2(C_b \vee C_\sigma) \int_0^t \mathbb{E}_x[\sup_{0 \leq u \leq s} V(X_u^{\tau_m})] ds \\ &\leq (1 + |x|)^{2p} (1 + 4p(C_\sigma T)^{\frac{1}{2}} \exp(16p^2(C_b \vee C_\sigma)T)) \\ &\quad + 8p^2(C_b \vee C_\sigma) \int_0^t \mathbb{E}_x[\sup_{0 \leq u \leq s} V(X_u^{\tau_m})] ds \end{aligned}$$

so that using Gronwall's lemma yields and passing to the limit $m \rightarrow +\infty$, for all $p \geq 1$

$$\mathbb{E}_x[\sup_{0 \leq t \leq T} |X_t|^{2p}] \leq \mathbb{E}_x[\sup_{0 \leq s \leq T} V(X_s)] \leq 2(1 + |x|)^{2p} \exp(26p^2(1 + (C_b \vee C_\sigma)T)).$$

□

For all real-valued and 1-Lipschitz function f defined on \mathcal{C} and for all $p \geq 1$, one has

$$\begin{aligned} \mathbb{E}_x[|f(X) - \mathbb{E}_x[f(X)]|^{2p}] &= \mathbb{E}_x[|f(X) - f(0) + f(0) - \mathbb{E}_x[f(X)]|^{2p}] \leq 2^{2p} \mathbb{E}_x[|X|_\infty^{2p}] \\ &\leq 2^{2p+1} (1 + |x|)^{2p} \exp(26p^2(1 + (C_b \vee C_\sigma)T)) \end{aligned} \quad (7.10)$$

where we used Lemma 7.19 for the last inequality. Now, combining the Chebyshev and Rosenthal inequalities for independent zero-mean random variables (see e.g. [JSZ85]), for all $p \geq 1$, there exists $C_{2p} > 0$ such that

$$\begin{aligned} \mathbb{P}_x \left(\frac{1}{M} \left| \sum_{k=1}^M f(X^k) - \mathbb{E}_x[f(X)] \right| \geq r \right) &\leq \frac{\mathbb{E}_x[(\sum_{k=1}^M f(X^k) - \mathbb{E}_x[f(X)])^{2p}]}{r^{2p} M^{2p}} \\ &\leq C_{2p} \frac{\mathbb{E}_x[|f(X) - \mathbb{E}_x[f(X)]|^{2p}]}{r^{2p} M^p} \\ &\leq 2 \frac{(2(1 + |x|))^{2p} \exp(28p^2(1 + (C_b \vee C_\sigma)T))}{r^{2p} M^p} \\ &:= 2 \exp(-\varphi(p)) \end{aligned}$$

with $\varphi(p) := -\kappa(b, \sigma, T)p^2 + p \log(\frac{r^2 M}{(2(1+|x|))^2})$ and where we used for all $p \geq 1$, $C_{2p} \leq (2p)^{2p} \leq \exp(2p^2)$, see e.g. p.235-236 in [JSZ85], and (7.10) for the last inequality. Optimizing the latter inequality with respect to p with $p \geq 1$, i.e. selecting $p = \frac{1}{2\kappa(b, \sigma, T)} \log(\frac{r^2 M}{(2(1+|x|))^2})$, we obtain

$$\mathbb{P}_x \left(\frac{1}{M} \left| \sum_{k=1}^M f(X^k) - \mathbb{E}_x[f(X)] \right| \geq r \right) \leq 2 \exp \left(-\frac{1}{4\kappa(b, \sigma, T)} \log \left(\frac{r^2 M}{(2(1 + |x|))^2} \right)^2 \right)$$

for $r^2 M \geq (2(1 + |x|))^2 \exp(2\kappa(b, \sigma, T))$. Otherwise, using the Jensen and Rosenthal inequalities, one has for all $p \in [0, 1]$

$$\begin{aligned} \mathbb{E}_x[(\sum_{k=1}^M f(X^k) - \mathbb{E}_x[f(X)])^{2p}] &\leq \mathbb{E}_x[(\sum_{k=1}^M f(X^k) - \mathbb{E}_x[f(X)])^2]^p \leq (MC_2 \mathbb{E}_x[|f(X) - \mathbb{E}_x[f(X)]|^2])^p \\ &\leq M^p \left(4(2(1 + |x|))^2 \exp(\kappa(b, \sigma, T)) \right)^p \end{aligned}$$

where we used (7.10) for the last inequality. Now, noticing that we have $4e \leq \exp(\kappa(b, \sigma, T))$, Chebyshev's inequality yields

$$\mathbb{P}_x \left(\frac{1}{M} \left| \sum_{k=1}^M f(X^k) - \mathbb{E}_x[f(X)] \right| \geq r \right) \leq \frac{C^p}{r^{2p} M^p} \leq 2 \frac{(Cp)^p}{r^{2p} M^p} \leq 2 \exp(-\varphi(p))$$

with $\varphi(p) := -p \log(p) + p \log(\frac{r^2 M}{C})$, $C := (2(1 + |x|))^2 \exp(2\kappa(b, \sigma, T) - 1)$ and where we used that for all $p \geq 0$, $C^p \leq 2(Cp)^p$ since the function $p \mapsto 2p^p$ is minimized for $p = \exp(-1)$ and $2 \exp(-1/e) > 1$. Consequently, optimizing over p such that $p \leq 1$, i.e. selecting $p = \frac{r^2 M}{Ce}$, one has

$$\mathbb{P}_x \left(\frac{1}{M} \left| \sum_{k=1}^M f(X^k) - \mathbb{E}_x[f(X)] \right| \geq r \right) \leq 2 \exp \left(-\frac{r^2 M}{(2(1 + |x|))^2 \exp(2\kappa(b, \sigma, T))} \right)$$

for $r^2 M \leq Ce = (2(1 + |x|))^2 \exp(2\kappa(b, \sigma, T))$. This completes the proof.

7.4 Proof of Proposition 7.5

Let $e_\sigma := \frac{1}{2\sigma} \exp(-|x|/\sigma)$ be the density of the exponential distribution with variance $2\sigma^2$ on \mathbb{R} . If μ is a probability measure on \mathbb{R}^d , we define μ^σ as the convolution of μ with $e_\sigma^{\otimes d}$, that is

$$\mu^\sigma(dx) := \int \prod_{i=1}^d \frac{1}{2\sigma} \exp(-|x_i - y_i|/\sigma) \mu(dy).$$

Lemma 7.20. *If μ is a probability measure on \mathbb{R}^d with finite first moment, then $W_1(\mu, \mu^\sigma) \leq \sqrt{2d}\sigma$.*

Démonstration. Let X and Y be independent random vectors with laws μ and $e_\sigma^{\otimes d}$ respectively. Then $(X, X + Y)$ is a coupling of μ and μ^σ , and

$$W_1(\mu, \mu^\sigma) \leq \mathbb{E}[|Y|] \leq \mathbb{E}[|Y|^2]^{1/2} \leq \sqrt{2d}\sigma.$$

□

We therefore have the bound

$$W_1(\mu_M, \mu) \leq W_1(\mu_M, \mu_M^\sigma) + W_1(\mu_M^\sigma, \mu^\sigma) + W_1(\mu^\sigma, \mu) \leq W_1(\mu_M^\sigma, \mu^\sigma) + \sqrt{8d}\sigma, \quad (7.11)$$

so what is left is to bound $\mathbb{E}[W_1(\mu_M^\sigma, \mu^\sigma)]$ and to optimize with respect to σ .

The density of μ_M^σ with respect to the Lebesgue measure is given by $g_{1,\sigma,M}(x) := \frac{1}{M} \sum_{k=1}^M e_\sigma^{\otimes d}(x - X_k)$, and the density of μ^σ is $g_{2,\sigma}(x) := \mathbb{E}_\mu(e_\sigma^{\otimes d}(x - X))$.

By the Kantorovitch-Rubinstein duality formula, we have

$$W_1(\mu_M^\sigma, \mu^\sigma) = \sup_{f:|f|_1 \leq 1} \int f(x) g_{1,\sigma,M}(x) dx - \int f(x) g_{2,\sigma}(x) dx \leq \int |x| |g_{1,\sigma,M}(x) - g_{2,\sigma}(x)| dx$$

To bound this quantity, we use the Cauchy-Schwarz inequality, namely for any nonnegative measurable function f on \mathbb{R}^d , we have

$$\int f(x) dx \leq C_d \sqrt{\int (1 + |x|^{d+1}) f(x)^2 dx}, \quad C_d := \sqrt{\int_{\mathbb{R}^d} \frac{1}{1 + |x|^{d+1}} dx}.$$

Using this inequality, we get the bound

$$\begin{aligned} W_1(\mu_M^\sigma, \mu^\sigma) &\leq C_d \sqrt{\int (1 + |x|^{d+1}) |x|^2 |g_{1,\sigma,M}(x) - g_{2,\sigma}(x)|^2 dx} \\ &\leq C_d \sqrt{\int (1 + 2|x|^{d+3}) |g_{1,\sigma,M}(x) - g_{2,\sigma}(x)|^2 dx}. \end{aligned} \quad (7.12)$$

Therefore,

$$\begin{aligned} \mathbb{E}[W_1(\mu_M^\sigma, \mu^\sigma)] &\leq C_d \mathbb{E} \left[\sqrt{\int (1 + 2|x|^{d+3}) \left| \frac{1}{M} \sum_{k=1}^M e_\sigma^{\otimes d}(x - X_k) - \mathbb{E}_\mu(e_\sigma^{\otimes d}(x - X)) \right|^2 dx} \right] \\ &\leq \frac{C_d}{\sqrt{M}} \sqrt{\int (1 + 2|x|^{d+3}) \text{Var}_\mu(e_\sigma^{\otimes d}(x - X)) dx} \\ &\leq \frac{C_d}{\sqrt{M}} \sqrt{\int (1 + 2|x|^{d+3}) \mathbb{E}[e_\sigma^{\otimes d}(x - X)^2] dx}. \end{aligned}$$

Note that $e_{\sigma}^{\otimes d}(x)^2 = 2^{-2d}\sigma^{-d}e_{\sigma/2}^{\otimes d}(x)$, so that we get

$$\begin{aligned}
 \mathbb{E}[W_1(\mu_M^{\sigma}, \mu^{\sigma})] &\leq \frac{C_d}{2^d\sigma^{d/2}\sqrt{M}} \sqrt{\int (1 + 2|x|^{d+3}) \int e_{\sigma/2}^{\otimes d}(x-y)\mu(dy)dx} \\
 &\leq \frac{C_d}{2^d\sigma^{d/2}\sqrt{M}} \sqrt{\int \int (1 + 2|u+y|^{d+3})e_{\sigma/2}^{\otimes d}(u)du\mu(dy)} \\
 &\leq \frac{C_d}{2^d\sigma^{d/2}\sqrt{M}} \sqrt{\int \int (1 + 2^{d+3}(|u|^{d+3} + |y|^{d+3}))e_{\sigma/2}^{\otimes d}(u)du\mu(dy)} \\
 &\leq \frac{C_d}{2^d\sigma^{d/2}\sqrt{M}} \sqrt{1 + 2^{d+3} \int |y|^{d+3}\mu(dy) + 2^{d+3} \int |u|^{d+3}e_{\sigma/2}^{\otimes d}(u)du} \\
 &\leq \frac{C_d}{2^d\sigma^{d/2}\sqrt{M}} \sqrt{1 + 2^{d+3} \int |y|^{d+3}\mu(dy) + \sigma^{d+3} \int |u|^{d+3}e_1^{\otimes d}(u)du} \\
 &\leq \frac{C_d}{2^d\sigma^{d/2}\sqrt{M}} \sqrt{1 + 2^{d+3} \int |y|^{d+3}\mu(dy) + 2^{d+3}\sigma^{d+3}d(d+3)!}
 \end{aligned}$$

In the end, assuming $\sigma \leq 1$, we obtain

$$\begin{aligned}
 \mathbb{E}[W_1(\mu, \mu^{\sigma})] &\leq \sqrt{8d}\sigma + \frac{C_d}{2^d\sigma^{d/2}\sqrt{M}} \sqrt{1 + 2^{d+3} \int |y|^{d+3}\mu(dy) + 2^{d+3}\sigma^{d+3}d(d+3)!} \\
 &\leq C(d, \mu)\left(\sigma + \frac{\sigma^{-d/2}}{\sqrt{M}}\right)
 \end{aligned}$$

Taking $\sigma = M^{-1/(d+2)}$, we get the upper bound we were aiming for.

Chapitre 8

Error analysis of the transport properties of Metropolized schemes

Ce chapitre a été écrit en collaboration avec Ahmed-Amine Hommann et Gabriel Stoltz, et a été soumis pour publication. Les simulations numériques ont été réalisées par Ahmed-Amine Hommann et Gabriel Stoltz.

Ce travail a été réalisé lors du CEMRACS 2013, qui a eu lieu au Centre International de Recherches Mathématiques de Marseille. Le projet a été proposé par Marie Jardat et Vincent Dahirel, et a financé par NEEDS « milieux poreux ».

Abstract : In this work, we are interested in the numerical computation of transport coefficients of Brownian dynamics. We investigate the discretization error arising when we simulate the dynamics using the Smart MC algorithm (also known as Metropolis-adjusted Langevin algorithm). We prove that the error is of order one as the time step goes to zero, when using either the Green-Kubo or the Einstein formula to estimate the transport coefficients. We illustrate our results with numerical simulations.

8.1 Introduction

Molecular simulation is nowadays a very common tool to quantitatively predict macroscopic properties of matter starting from a microscopic description. These macroscopic properties can be either static properties (such as the average pressure or energy in a system at fixed temperature and density), or transport properties (such as thermal conductivity or shear viscosity). Molecular simulation can be seen as the computational version of statistical physics, and is therefore often used by practitioners of the field as a black box to extract the desired macroscopic properties from some model of interparticle interactions. Most of the work in the physics and chemistry fields therefore focuses on improving the microscopic description, most notably developing force fields of increasing complexity. In comparison, not so much attention has been paid to the estimation of errors in the quantities actually computed by numerical simulation. Usually, due to the very high dimensionality of the systems under consideration, macroscopic properties are computed as ergodic averages over a very long trajectory of the system, evolved under some appropriate dynamics. There are two main types of errors in this approach : (i) statistical errors ari-

sing from incomplete sampling, and (ii) systematic errors (bias) arising from the fact that continuous dynamics are numerically integrated using a finite time-step $\Delta t > 0$.

The aim of this work is to understand the bias arising from the use of finite time steps in the computation of transport coefficients. We consider the case of the self-diffusion, for a certain type of dynamics called Brownian dynamics in the chemistry literature, discretized using the so-called ‘‘Smart MC’’ algorithm [RDF78, JBTK99] (this algorithm was also rediscovered later on in the computational statistics literature [RT96]).

This proceedings is organized as follows. We start by describing in Section 8.2 the Brownian dynamics and its discretization, and define the self-diffusion. We then provide in Section 8.3 a priori error estimates for the numerical estimation of the self-diffusion, through two different routes. Numerical simulations illustrate our error bounds in Section 8.4. We conclude in Section 8.5 with some tracks to reduce the numerical error by appropriately modifying the numerical scheme. The proofs of our results are gathered in Section 8.6.

8.2 Description of the model

8.2.1 Brownian dynamics

Consider N particles with positions $q = (q_1, \dots, q_N) \in \mathcal{M} = (L\mathbb{T})^{dN}$, $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ being the standard one dimensional torus and d being the physical dimension (usually $d = 3$). The positions of the particles evolve according to the following dynamics :

$$dq_t = -\beta \nabla V(q_t) dt + \sqrt{2} dW_t, \quad (8.1)$$

where $\beta = 1/(k_B T)$ is the inverse temperature and W_t is a standard dN -dimensional Brownian motion. The function $V : \mathcal{M} \rightarrow \mathbb{R}$ is the potential energy, assumed to be smooth for the mathematical analysis. However, the numerical results presented in Section 8.4.2 correspond to a potential energy function with singularities.

Standard results show that (8.1) admits the Boltzmann-Gibbs measure

$$\mu(dq) = Z^{-1} e^{-\beta V(q)} dq \quad (8.2)$$

as unique invariant probability measure. In fact, (8.1) is ergodic with respect to this measure, where ergodicity is understood both as (i) the longtime (almost-sure) convergence of averages along trajectories

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t A(q_s) ds = \int_{\mathcal{M}} A(q) \mu(dq) \quad \text{a.s.}$$

for any initial condition $q_0 \in \mathcal{M}$ and all observables $A \in L^1(\mu)$; or as (ii) the convergence of the law $\psi(t, q) dq$ of the process (8.1), happening here at an exponential rate, for instance in total variation : There exist $C, \lambda > 0$ such that

$$\|\psi(t) - \mu\|_{\text{TV}} \leq C e^{-\lambda t},$$

where the total variation distance between two measures ν_1, ν_2 is defined as (we use the normalization from [BRH13])

$$\|\nu_1 - \nu_2\|_{\text{TV}} = 2 \sup_{S \in \mathcal{B}(\mathcal{M})} |\nu_1(S) - \nu_2(S)| = \sup_{|\varphi| \leq 1} \left| \int_{\mathcal{M}} \varphi d\nu_1 - \int_{\mathcal{M}} \varphi d\nu_2 \right|,$$

the suprema being taken over all measurable sets of \mathcal{M} for the first one, and over all bounded, measurable functions for the second one.

For further purposes, we introduce the generator of (8.1), namely the operator

$$\mathcal{L} = -\beta \nabla V \cdot \nabla + \Delta. \quad (8.3)$$

This operator (defined with domain $D(\mathcal{L}) = H^2(\mu)$) is self-adjoint on the Hilbert space $L^2(\mu)$, whose associated scalar product is

$$\langle \varphi, \psi \rangle_{L^2(\mu)} = \int_{\mathcal{M}} \varphi \psi d\mu.$$

The operator \mathcal{L} moreover has a positive spectral gap. Indeed, a simple computation shows that

$$\langle \mathcal{L}\varphi, \varphi \rangle_{L^2(\mu)} = \frac{1}{\beta} \|\nabla \varphi\|_{L^2(\mu)}^2. \quad (8.4)$$

The Poincaré inequality $\|\varphi\|_{L^2(\mu)} \leq C_{\mathcal{M},V} \|\nabla \varphi\|_{L^2(\mu)}$, valid for any function belonging to

$$\tilde{L}^2(\mu) = \left\{ \varphi \in L^2(\mu) \mid \int_{\mathcal{M}} \varphi d\mu = 0 \right\},$$

allows to conclude that

$$\forall \varphi \in \tilde{L}^2(\mu), \quad \langle \mathcal{L}\varphi, \varphi \rangle_{L^2(\mu)} \geq C_{\mathcal{M},V}^{-1} \|\varphi\|_{L^2(\mu)}^2, \quad (8.5)$$

which shows that the spectral gap is larger or equal to $C_{\mathcal{M},V}^{-1}$. In particular, the resolvent \mathcal{L}^{-1} is a well-defined operator on $\tilde{L}^2(\mu)$, and the following estimate holds :

$$\left\| \mathcal{L}^{-1} \right\|_{\mathcal{B}(\tilde{L}^2(\mu))} \leq C_{\mathcal{M},V}. \quad (8.6)$$

8.2.2 Self-diffusion

The positions q_t are restricted to the periodic domain \mathcal{M} and are therefore uniformly bounded in time. To obtain a diffusive behavior from the evolution of q_t , we consider the following additive functional defined on the whole space \mathbb{R}^d : starting from $Q_0 = q_0$,

$$Q_t - Q_0 = -\beta \int_0^t \nabla V(q_s) ds + \sqrt{2} W_t. \quad (8.7)$$

The difference with q_t is that Q_t is not reprojected in \mathcal{M} by the periodization procedure. The diffusion tensor is then given by the following limit (provided it exists) :

$$\mathcal{D} = \lim_{t \rightarrow +\infty} \mathbb{E} \left(\frac{Q_t - Q_0}{\sqrt{t}} \otimes \frac{Q_t - Q_0}{\sqrt{t}} \right), \quad (8.8)$$

where the expectation is over all realizations of the continuous dynamics (8.1), starting from initial conditions distributed according to the Boltzmann-Gibbs measure (8.2). The following result shows that the diffusion tensor (8.8) is well defined, and naturally arises in a diffusive time-rescaling of the dynamics (8.1).

Theorem 8.1. *Consider for $\varepsilon > 0$ the diffusively rescaled process $Q_t^\varepsilon = \varepsilon Q_{t/\varepsilon^2}$. Then, as $\varepsilon \rightarrow 0$, the process Q_t^ε starting from a given initial condition Q_0 weakly converges on*

finite time intervals to an effective Brownian motion starting from Q_0 and with covariance matrix \mathcal{D} given by (8.8). Moreover, \mathcal{D} is a real, positive definite $dN \times dN$ matrix, satisfying

$$0 \leq \mathcal{D} \leq 2\text{Id}$$

in the sense of symmetric matrices, and which can alternatively be expressed as

$$\mathcal{D} = 2 \left(\text{Id} - \beta^2 \int_0^{+\infty} \mathbb{E} [\nabla V(q_t) \otimes \nabla V(q_0)] dt \right), \quad (8.9)$$

where the expectation is over all realizations of the continuous dynamics (8.1), starting from initial conditions distributed according to the Boltzmann-Gibbs measure (8.2).

The proof of this statement is standard, and follows from arguments presented in [BLP11, Chapter 3] for instance. We nonetheless provide a short proof in Section 8.6.1 since the proofs of the discrete counterparts of Theorem 8.1 rely on an appropriate extension of the argument used in the continuous case (see Section 8.6.5).

A straightforward consequence of Theorem 8.1 is that the self-diffusion constant \mathcal{D} , defined as the average mean-square displacement of the individual particles, is well defined, and has two equivalent expressions :

$$\mathcal{D} = \frac{1}{2dN} \text{Tr}(\mathcal{D}) = \lim_{t \rightarrow +\infty} \mathbb{E} \left(\frac{1}{2dNt} \sum_{i=1}^N (Q_{i,t} - Q_{i,0})^2 \right) \quad (8.10)$$

$$= 1 - \frac{\beta^2}{dN} \int_0^{+\infty} \mathbb{E} [\nabla V(q_t)^T \nabla V(q_0)] dt. \quad (8.11)$$

The expression (8.10) is called the *Einstein formula*, and the second one involves an integrated autocorrelation function. In accordance with the standard physics and chemistry nomenclature, we call (8.11) the *Green-Kubo formula* of the self-diffusion in the sequel.

8.2.3 Numerical estimation of the self-diffusion

In order to compute approximations of formulas such as (8.10) or (8.11), the first task is to numerically integrate realizations of the continuous dynamics (8.1). The most straightforward way would be to resort to a Euler-Maruyama scheme : given a time-step $\Delta t > 0$ and denoting by q^n an approximation of $q_{n\Delta t}$, this scheme reads

$$q^{n+1} = q^n - \beta \Delta t \nabla V(q^n) + \sqrt{2\Delta t} G^n, \quad (8.12)$$

where $(G^n)_{n \geq 0}$ is a sequence of dN dimensional standard Gaussian random variables. However, this simple scheme has been shown to fail to be ergodic when the dynamics is considered on unbounded spaces and the potential energy function is not globally Lipschitz [MSH02]. In simulations of Brownian dynamics for ionic solutions, potential energy functions with Coulomb-type singularities are used and it has been observed that the energy blows up along trajectories of (8.12) (see Section 8.4.2).

A way to stabilize the Euler-Maruyama scheme is to consider the configuration (8.12) as a proposal move in a Metropolis-Hastings algorithm [MRR⁺53, Has70], which is precisely the Smart MC algorithm proposed in [RDF78] and later called Metropolis adjusted Langevin algorithm in the computational statistics literature [RT96]. More precisely, starting from a configuration $q^n \in \mathcal{M}$ (in fact seen as an element of \mathbb{R}^{dN}), a new configuration $\tilde{q}^{n+1} \in \mathbb{R}^{dN}$ is proposed according to (8.12), and then accepted with probability

$$R_{\Delta t} \left(q^n, \tilde{q}^{n+1} \right) = \frac{e^{-\beta V(\tilde{q}^{n+1})} T_{\Delta t}(\tilde{q}^{n+1}, q^n)}{e^{-\beta V(q^n)} T_{\Delta t}(q^n, \tilde{q}^{n+1})},$$

where

$$T_{\Delta t}(q, q') = \left(\frac{1}{4\pi\Delta t} \right)^{dN/2} \exp \left(-\frac{|q' - q + \beta\Delta t \nabla V(q)|^2}{4\Delta t} \right)$$

is the probability transition of the Markov chain (8.12). When the proposition is accepted, we project \tilde{q}^{n+1} into the periodic simulation cell \mathcal{M} . If the proposal is rejected, the previous configuration is counted twice : $q^{n+1} = q^n$ (It is very important to count rejected configuration as many times as needed to ensure that the Boltzmann-Gibbs measure μ is invariant). In conclusion,

$$q^{n+1} = q^n + \mathbf{1}_{U^n \leq R_{\Delta t}(q^n, \tilde{q}^{n+1})} \left(-\beta\Delta t \nabla V(q^n) + \sqrt{2\Delta t} G^n \right), \quad (8.13)$$

where, in this formula, q^{n+1} is considered as an element of the periodic box \mathcal{M} , while the proposed configuration \tilde{q}^{n+1} is not reprojected into the simulation cell \mathcal{M} using the periodic boundary conditions and is therefore considered as an element of \mathbb{R}^{dN} .

In order to avoid confusions, we call this scheme “metropolized Euler-Maruyama” in the sequel, and denote by $P_{\Delta t}$ its evolution operator :

$$P_{\Delta t}f(q) = \mathbb{E} \left(f \left(q^{n+1} \right) \mid q^n = q \right).$$

By construction, the measure (8.2) is an invariant probability measure for this scheme, which is a reversible Markov chain. We refer to [BRH13] for a study of the ergodic properties of the dynamics (in the more complicated case of dynamics on the full configuration space \mathbb{R}^{dN} , subjected to a confining potential).

Of course, the fact that some configurations are rejected destroys the trajectorial accuracy of the dynamics, see [BRVE09] for precise statements. It is unclear how to account for these rejections in the computation of the diffusion coefficient. The next section quantifies the errors in the approximation of (8.10) and (8.11) when the metropolized Euler-Maruyama scheme is used. Although the formulas (8.10) and (8.11) are equivalent for continuous dynamics, they lead to different numerical methods.

8.3 A priori error estimates on the self-diffusion

As discussed in [LMS13], error bounds on transport properties in fact depend on approximation properties of the evolution operator (similar to the one used to prove weak error estimates), rather than strong error estimates. A key building block in this framework is the following expansion of the evolution operator.

Lemma 8.2. *There exist an operator A and $\Delta t^* > 0$ such that, for any $0 < \Delta t \leq \Delta t^*$ and any smooth function ψ ,*

$$P_{\Delta t}\psi = \psi + \Delta t \mathcal{L}\psi + \Delta t^2 A\psi + \Delta t^{5/2} r_{\psi, \Delta t}, \quad (8.14)$$

with a remainder $r_{\psi, \Delta t}$ uniformly bounded for $0 < \Delta t \leq \Delta t^*$. Moreover,

$$\int_{\mathcal{M}} A\psi d\mu = 0. \quad (8.15)$$

Finally, the average rejection rate scales as $\Delta t^{3/2}$: There exists a constant $C \geq 0$ and a bounded function $\bar{\xi}$ such that, for any $0 < \Delta t \leq \Delta t^*$,

$$\left| \mathbb{E}_G \left[1 - R_{\Delta t} \left(q, q - \Delta t \nabla V(q) + \sqrt{2\Delta t} G \right) \right] - \Delta t^{3/2} \bar{\xi}(q^n) \right| \leq C\Delta t^2, \quad (8.16)$$

where the expectation is over all possible realizations of the standard, dN dimensional Gaussian random variable G .

The precise expression of the operator A is unimportant. It is however given in the proof of this result, see Section 8.6.2. Note that the method can be proved to be weakly first order accurate since

$$P_{\Delta t}\psi - e^{\Delta t\mathcal{L}}\psi = \Delta t^2 \left(A - \frac{1}{2}\mathcal{L}^2 \right) \psi + \Delta t^{5/2}r_{\psi,\Delta t}.$$

Another important result which we will repeatedly use in the analysis below is the following uniform geometric ergodicity of the metropolized Euler-Maruyama scheme, easily obtained from the results of [BRH13] (we nonetheless provide for completeness elements of proof in Section 8.6.3). To state the result, we introduce the following functional space

$$\tilde{L}^\infty(\mathcal{M}) = \left\{ f \in L^\infty(\mathcal{M}) \mid \int_{\mathcal{M}} f d\mu = 0 \right\}.$$

Lemma 8.3. *There exists $\Delta t^* > 0$ and $C, \lambda > 0$ such that, for any $0 < \Delta t < \Delta t^*$, for all $n \in \mathbb{N}$ and any $f \in \tilde{L}^\infty(\mathcal{M})$,*

$$\|P_{\Delta t}^n f\|_{L^\infty} \leq C e^{-\lambda n \Delta t} \|f\|_{L^\infty}. \quad (8.17)$$

As a consequence, there exists $K > 0$ such that

$$\left\| \left(\frac{\text{Id} - P_{\Delta t}}{\Delta t} \right)^{-1} \right\|_{\mathcal{B}(\tilde{L}^\infty)} \leq K,$$

where $\mathcal{B}(\tilde{L}^\infty)$ denotes the Banach space of bounded operators on $\tilde{L}^\infty(\mathcal{M})$.

8.3.1 Error estimates for the Green-Kubo formula

We first give error estimates on (8.11) by appropriately adapting the results from [LMS13]. The result is stated for two general observables $\psi, \varphi \in \tilde{L}^\infty(\mathcal{M})$. Error estimates for (8.11) are obtained by setting $\psi = \varphi = -\partial_{q_{i,\alpha}} V$, with $1 \leq i \leq N$ and $1 \leq \alpha \leq d$. Note indeed that a simple integration by parts shows that $\partial_{q_{i,\alpha}} V$ has average 0 with respect to μ , so $\partial_{q_{i,\alpha}} V \in \tilde{L}^\infty(\mathcal{M})$.

Theorem 8.4. *Consider two observables $\psi, \varphi \in \tilde{L}^\infty(\mathcal{M})$, and define the modified observable*

$$\tilde{\psi}_{\Delta t} = (\text{Id} + \Delta t A \mathcal{L}^{-1})\psi,$$

where the operator A is defined in (8.14). Then, there exists $\Delta t^* > 0$ such that, for any $0 < \Delta t \leq \Delta t^*$,

$$\int_0^{+\infty} \mathbb{E}[\psi(q_t) \varphi(q_0)] dt = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} [\tilde{\psi}_{\Delta t}(q^n) \varphi(q^0)] + \Delta t^{3/2} r_{\psi,\varphi,\Delta t},$$

with $r_{\psi,\varphi,\Delta t}$ uniformly bounded, and where the expectation on the left hand side is with respect to initial conditions $q_0 \sim \mu$ and for all realizations of the dynamics (8.1), while the expectation on the right hand side is with respect to initial conditions $q^0 \sim \mu$ and for all realizations of the metropolized Euler-Maruyama scheme.

As a corollary, we obtain first order error bounds on the computation of the self-diffusion through (8.11) :

$$\mathcal{D} = \mathcal{D}_{\Delta t}^{\text{GK}} + \Delta t \tilde{\mathcal{D}}^{\text{GK},1} + \Delta t^{3/2} \tilde{\mathcal{D}}_{\Delta t}^{\text{GK}}, \quad (8.18)$$

where $\tilde{\mathcal{D}}_{\Delta t}$ is uniformly bounded for Δt sufficiently small, and where the numerically computed self-diffusion reads

$$\mathcal{D}_{\Delta t}^{\text{GK}} = \text{Id} - \frac{\beta^2}{dN} \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} \left[\nabla V(q^n)^T \nabla V(q^0) \right]. \quad (8.19)$$

The appearance of subleading fractional correction term (here, of order $\Delta t^{3/2}$) is typical of Metropolis algorithms, and is usually not encountered for standard, un-metropolized discretizations of SDEs.

8.3.2 Error bounds on the Einstein formula

In this section, we investigate the discretization error made when using the metropolized Euler-Maruyama scheme to approximate the self-diffusion using (8.10). In accordance with the definition (8.7), we introduce a discrete additive functional allowing to keep track of the diffuse behavior of the Markov chain : Starting from $Q^0 = q^0$,

$$Q^n = \sum_{k=0}^{n-1} \delta_{\Delta t} \left(q^k, G^k, U^k \right),$$

with

$$\delta_{\Delta t} \left(q^k, G^k, U^k \right) = \mathbf{1}_{U^k \leq R_{\Delta t}(q^k, q^k - \beta \Delta t \nabla V(q^k) + \sqrt{2\Delta t} G^k)} \left(-\beta \Delta t \nabla V(q^k) + \sqrt{2\Delta t} G^k \right). \quad (8.20)$$

While the Markov chain $(q^n)_{n \geq 0}$ defined by (8.13) remains in \mathcal{M} , the additive functional $(Q^n)_{n \geq 0}$ has values in \mathbb{R}^{dN} . With this notation, the diffusion tensor actually computed by the numerical scheme is

$$\mathcal{D}_{\Delta t}^{\text{Einstein}} = \lim_{n \rightarrow +\infty} \mathbb{E}_{\Delta t} \left[\frac{Q^n - Q^0}{\sqrt{n\Delta t}} \otimes \frac{Q^n - Q^0}{\sqrt{n\Delta t}} \right], \quad (8.21)$$

where, as in Theorem 8.4, the expectation on the right hand side is with respect to initial conditions $Q^0 = q^0 \sim \mu$ and for all realizations of the metropolized Euler-Maruyama scheme.

Theorem 8.5. *There exists $\Delta t^* > 0$ and a symmetric matrix $\tilde{\mathcal{D}}^1 \in \mathbb{R}^{dN \times dN}$ such that, for any $0 < \Delta t \leq \Delta t^*$,*

$$\mathcal{D} = \mathcal{D}_{\Delta t}^{\text{Einstein}} + \Delta t \tilde{\mathcal{D}}_{\Delta t},$$

where the coefficients of the matrix $\tilde{\mathcal{D}}_{\Delta t} \in \mathbb{R}^{dN \times dN}$ are uniformly bounded.

The proof of this result can be read in Section 8.6.5. In fact, a slight extension of our technique of proof would allow to show that the diffusively rescaled process generated by the metropolized Euler-Maruyama scheme, namely $\varepsilon Q^{\lfloor t/(\Delta t \varepsilon^2) \rfloor}$ (where $\lfloor x \rfloor$ denote the unique integer such that $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$), weakly converges on finite time intervals to a Brownian motion with covariance matrix $\mathcal{D}_{\Delta t}$.

An immediate corollary of Theorem 8.5 is the following a priori error estimate on the self-diffusion :

$$\mathcal{D} = \mathcal{D}_{\Delta t}^{\text{Einstein}} + \Delta t \tilde{\mathcal{D}}_{\Delta t}^{\text{Einstein}}, \quad \mathcal{D}_{\Delta t}^{\text{Einstein}} = \frac{1}{2dN} \text{Tr} \left(\mathcal{D}_{\Delta t}^{\text{Einstein}} \right), \quad (8.22)$$

where $\tilde{\mathcal{D}}_{\Delta t}^{\text{Einstein}}$ is uniformly bounded for Δt sufficiently small. Some more work would allow to prove that the subleading correction term is of order $\Delta t^{3/2}$, as in the Green-Kubo case (see Remark 8.9).

8.4 Numerical illustration

The aim of this section is to illustrate the errors bounds (8.18) and (8.22). We perform long computations in order for the statistical errors to be negligible.

8.4.1 A simple one-dimensional case

We start by considering a simple one dimensional example ($N = d = 1$) : a single particle in the unit torus $\mathcal{M} = \mathbb{T}$, with the periodic potential $V(q) = \cos(2\pi q)$, at $\beta = 1$. Computations are performed by approximating expectations by realizations over M replicas evolving independently, denoted by $q^{m,n}$ with $1 \leq m \leq M$ and where n still is the step index. Initial conditions are prepared incrementally over the replicas. More precisely, starting from $q^{1,0} = 0$, we obtain the initial condition $q^{m+1,0}$ for the replica number $m + 1$ by evolving the initial condition $q^{m,0}$ over 10 steps of the metropolized Euler-Maruyama scheme with step size $\Delta t_{\text{thm}} = 0.01$. We have checked that the equilibrium distribution is very well reproduced by the empirical measure produced by $\{q^{m,0}\}_{1 \leq m \leq M}$ provided M is reasonably large (say, $M \geq 10^3$).

The self-diffusion coefficient $\mathcal{D}_{\Delta t}^{\text{Einstein}}$ for the Einstein approach is approximated by fitting the unnormalized self-diffusion

$$D_n^M = \frac{1}{M} \sum_{m=1}^M (Q^{m,n} - Q^{m,0})^2 \quad (8.23)$$

by a linear function $\mathcal{D}_{\Delta t}^{\text{Einstein},M} n \Delta t$, the slope being the estimation of the self-diffusion for the time step under consideration. This is indeed confirmed by Figure 8.1 (Left), which presents the evolution of D_n^M as a function of the physical time $n \Delta t$. The results produced in Figures 8.1 and (8.3) have been obtained with $M = 10^7$ replicas and $n_{\text{Einstein}} = 3 \times 10^5$ steps. Let us also note that, in accordance with (8.16), the rejection rate scales as $\Delta t^{3/2}$.

A numerical approximation of (8.19) requires both a discretization using finitely many replicas, but also a truncation of the integration in time with an upper bound τ . We consider the following numerical estimation of the self-diffusion coefficient obtained with the Green-Kubo formula :

$$\mathcal{D}_{\Delta t}^{\text{GK},M,\tau} = \text{Id} - \beta^2 \Delta t \sum_{n=0}^{\lfloor \tau/\Delta t \rfloor} \frac{1}{M} \left[\nabla V(q^{m,n})^T \nabla V(q^{m,0}) \right]. \quad (8.24)$$

The correlation functions shown in Figure 8.2 suggest that the autocorrelation of ∇V is exponentially decreasing. It can be considered as negligible for times larger than 0.2. The numerical results reported in Figure 8.2 and (8.3) have been obtained with $M = 2 \times 10^8$ replicas and a time cut-off $\tau = 0.3$.

The results presented in Figure 8.3 indeed confirm that, for small time steps Δt , the error in the self-diffusion is of order Δt for both methods. For larger time steps, nonlinearities appear. Note also that estimates obtained with the Green-Kubo formula are more reliable in this simple case. Note also that, in accordance with the statements of Theorem 8.1, the self-diffusion is between 0 and 1 when $\Delta t \rightarrow 0$.

8.4.2 The more realistic case of solvated ions

We consider in this section a more physical system : a large, fixed ion interacting with smaller particles, typically smaller ions. The N smaller particles evolve in a three-dimensional simulation box of length L with periodic boundary conditions. The smaller

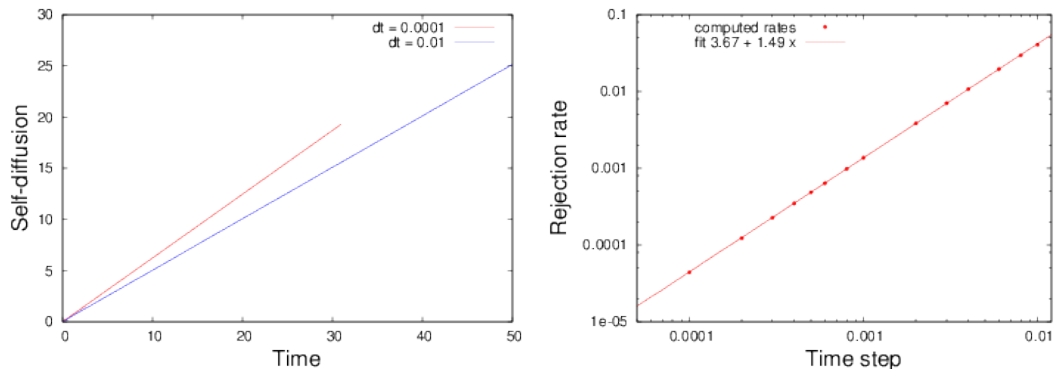


FIGURE 8.1 – Left : Self-diffusion D_n^M as a function of the physical time $n\Delta t$ for two values of the time step Δt . Right : average rejection as a function of the time step Δt , in a log-log scale.

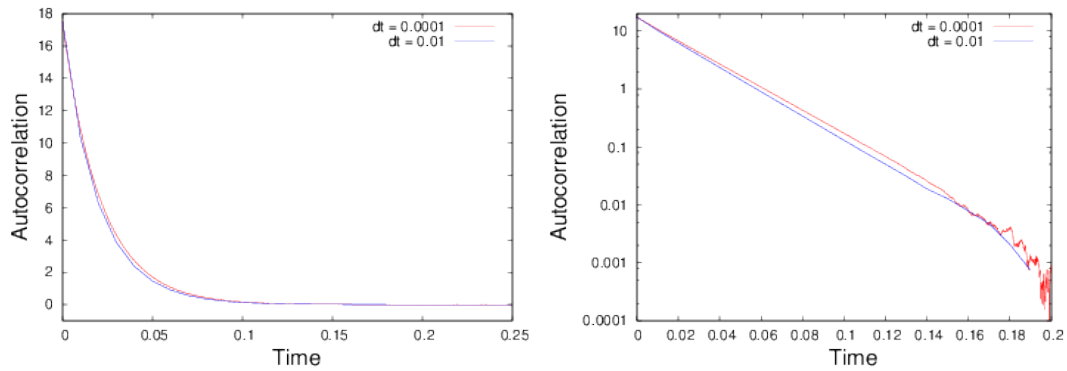


FIGURE 8.2 – Plot of the approximated correlation functions $\mathbb{E}(\nabla V(q_t)^T \nabla V(q_0))$. Left : standard view. Right : logarithmic scale on the ordinates.

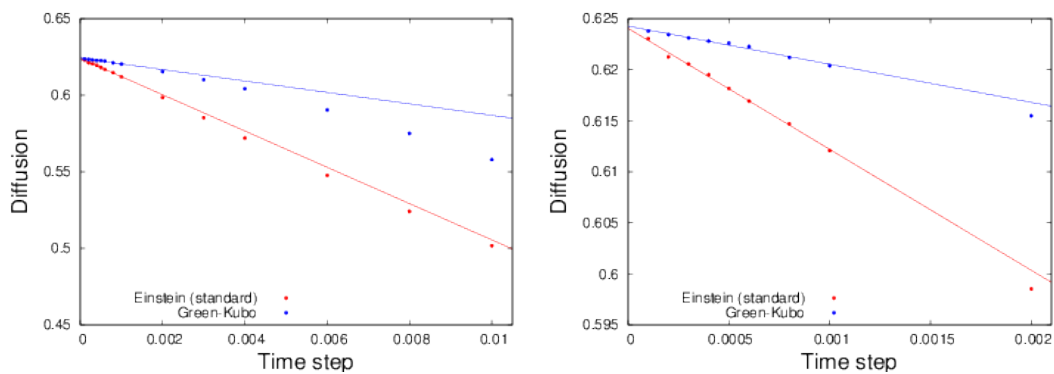


FIGURE 8.3 – Diffusion constant as a function of the time step Δt for the one-dimensional potential $V(q) = \cos(2\pi q)$ at $\beta = 1$, with a zoom on the smaller time steps on the right picture.

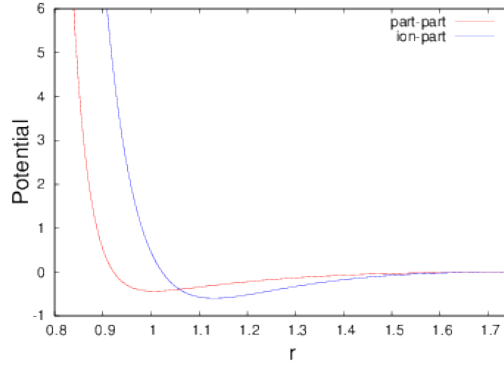


FIGURE 8.4 – Plot of the particle-particle interaction v (red), and of the ion-particle interaction v_{ion} (blue).

particles have the same mass m , and their positions are denoted by $q_i \in (LT)^3$. The potential energy functions are inspired from standard choices in the modelling of ionic solutions [JBTK99]. The interaction between the small particles is governed by an appropriately truncated Lennard-Jones potential :

$$v(r) = 4\varepsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right) - \varepsilon_{\text{shift}} - f_{\text{spline}}(r - r_{\text{cut}}), \quad \text{when } r \leq r_{\text{cut}}, \quad (8.25)$$

and $v(r) = 0$ for $r \geq r_{\text{cut}}$. The parameter $\varepsilon > 0$ is some reference energy, while $\sigma > 0$ is some reference distance. The parameters $\varepsilon_{\text{shift}}$ and f_{spline} ensure that v is a C^1 function. When $r_{\text{cut}} \rightarrow +\infty$, the minimal energy of v converges to $-\varepsilon$, a value obtained at a distance $r_{\text{min}} = 2^{1/6}\sigma$.

We additionally consider a large ion, modelled as fixed particle at position q_{ion} (the center of the simulation box), whose interaction with the solvent particles is described by an attractive Yukawa potential (screened Coulomb interaction) plus some repulsive potential preventing the small particles to come too close to the ion. More precisely, for a solvent particle at position q , the interaction reads $v_{\text{ion}}(|q - q_{\text{ion}}|)$, with

$$v_{\text{ion}}(r) = E_{\text{min}} \left(1 - \frac{1 + \kappa\sigma}{24} \right)^{-1} \left(\frac{1 + \kappa\sigma}{24} \left(\frac{\sigma}{r} \right)^{24} - \frac{\sigma}{r} e^{-\kappa(r-\sigma)} \right) - \varepsilon_{\text{shift}}^{\text{ion}} - f_{\text{spline}}^{\text{ion}}(r - r_{\text{cut}}^{\text{ion}}), \quad r \leq r_{\text{cut}}^{\text{ion}}, \quad (8.26)$$

and $v_{\text{ion}}(r) = 0$ for $r \geq r_{\text{cut}}^{\text{ion}}$. The parameters $\varepsilon_{\text{shift}}^{\text{ion}}$ and $f_{\text{spline}}^{\text{ion}}$ ensure as above that v_{ion} is C^1 . The energy $-E_{\text{min}} \leq 0$ is the minimal value of the potential, obtained when $r = \sigma$ (in the limit when $r_{\text{cut}}^{\text{ion}} \rightarrow +\infty$ and $f_{\text{shift}}^{\text{ion}} = 0$) while κ is some inverse length. The total potential energy of the N small particles finally reads

$$V(q_1, \dots, q_N) = \sum_{1 \leq i < j \leq N} v(|q_i - q_j|) + \sum_{i=1}^N v_{\text{ion}}(|q_i - q_{\text{ion}}|).$$

The potentials v and v_{ion} are plotted in Figure 8.4.

The results of this section are expressed in the reduced units obtained from the Lennard-Jones energy ε , the Lennard-Jones distance σ and the mass m . In particular, the reference time is $t^* = \sigma\sqrt{m/\varepsilon}$. Simulations were performed using the following parameters : $N = 20$, solvent density $\rho = N/L^3 = 0.4$, $E_{\text{min}} = 0.8347$, $\kappa = 1.7025$, temperature $k_{\text{B}}T = 1$, and $r_{\text{cut}} = r_{\text{cut}}^{\text{ion}} = 1.76$. For this choice of parameters, we have observed that

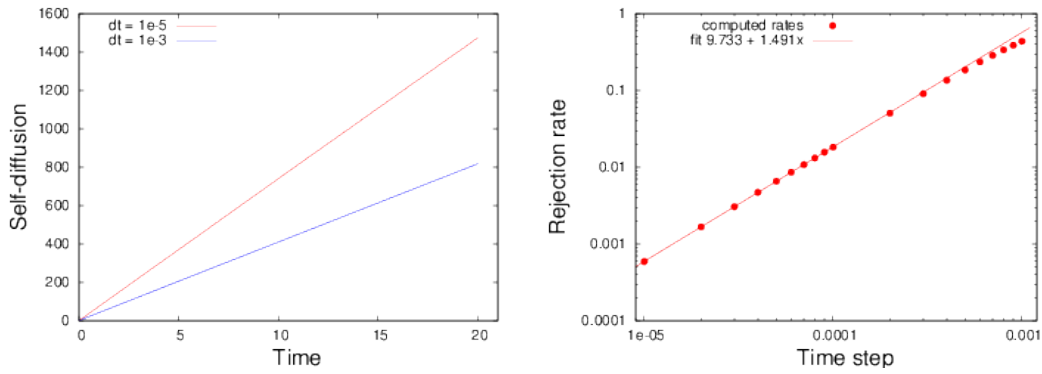


FIGURE 8.5 – Left : Self-diffusion D_n^M as a function of the physical time $n\Delta t$ for two values of the time step Δt . Right : average rejection as a function of the time step Δt , in a log-log scale.

the simulations blow up for time-steps Δt of the order of 4×10^{-4} when using the Euler-Maruyama (unmetropolized) scheme (8.12) ; whereas the metropolized Euler scheme (8.13) allows for much larger time steps.

Expectations are approximated using M trajectories of the system. We integrate trajectories one after the other, using the metropolized Euler scheme (8.13), with initial conditions for the $(m + 1)$ th trajectory obtained by taking the last configuration of the m th trajectory. The self-diffusion coefficient calculated with the Green-Kubo formula is obtained as in the previous section, using the estimator (8.24). The values obtained by the Einstein equation are computed by dividing the unnormalized diffusion (8.23) by the final time of the simulation : for a simulation time τ ,

$$D_{\Delta t}^{\text{Einstein}, M, \tau} = \frac{1}{2dN\tau} D_{[\tau/\Delta t]}^M. \quad (8.27)$$

The results presented in Figure 8.5 show that the unnormalized mean squared displacement indeed grows linearly in time, as expected. Note also that, in accordance with (8.16), the rejection rate scales as $\Delta t^{3/2}$. The results of Figure 8.5 and (8.7) have been obtained with $M = 10^5$ trajectories, with integrations performed up to $\tau = 20$.

The results presented in Figure 8.6 suggest that the force autocorrelation decay cannot be represented by a single exponential function. The force autocorrelation can be considered to be small in relative value for times of the order of 0.1. The numerical results reported in Figure 8.6 and 8.7 were obtained by averaging $M = 5 \times 10^5$ trajectories with an integration time $\tau = 0.3$.

Error estimates for the diffusion coefficients are gathered in Figure 8.7. For small time steps Δt , the error in the self-diffusion is linear in Δt , while nonlinearities appear for larger time steps. The continuous lines are linear fits obtained over the values corresponding to the 10 smallest time steps. As in the simple example discussed in the previous section, estimations obtained with Green-Kubo's formula seem more reliable than those obtained with Einstein's formula. Note also that, in accordance with Theorem 8.1, the self-diffusion is between 0 and 1 in all cases.

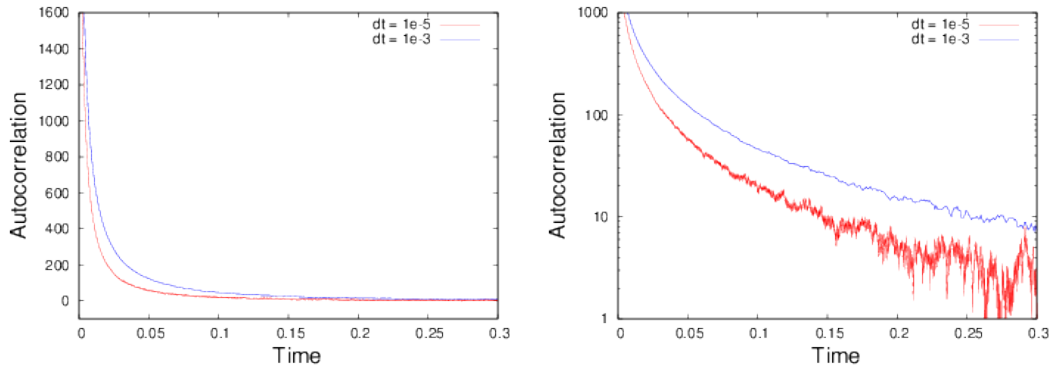


FIGURE 8.6 – Plot of the approximated correlation functions $\mathbb{E}(\nabla V(q_t)^T \nabla V(q_0))$. Left : standard view. Right : logarithmic scale on the ordinates.

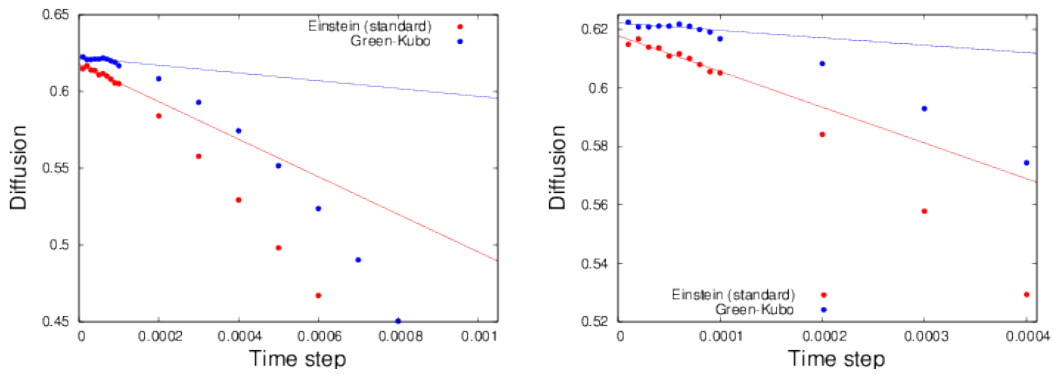


FIGURE 8.7 – Diffusion constant as a function of the time step Δt for the solvated ion system at $\beta = 1$, with a zoom on the smaller time steps on the right picture.

8.5 Possible work tracks to reduce the error on the estimation of the self-diffusion

Both the Green-Kubo and the Einstein approaches lead to discretization errors of order Δt , as proved theoretically and verified numerically. A natural question is how to reduce this error. In the chemistry literature, it was proposed to renormalize the time in Einstein's method by replacing the simulation time $n\Delta t$ appearing in the denominator of (8.21) by some effective time $\theta_{\Delta t}n\Delta t$, where $\theta_{\Delta t}$ is the average acceptance rate of the Metropolis algorithm [HB98] for a given time step Δt . However, since the average acceptance rate is of order $1 - C\Delta t^3$ for small time steps (see (8.16)), such a correction cannot possibly cancel out the Δt -error in the diffusion coefficient.

A more promising work track, which we started working on at the end of our stay at CIRM, is to modify the proposed move (instead of simply considering the Euler-Maruyama scheme (8.12)), and possibly the invariant measure as well, in order to increase the weak and strong orders of the associated Metropolized scheme. In fact, the proofs of Theorem 8.4 and 8.5 show that, in order to gain accuracy on the computation of transport coefficients, it is sufficient to find a numerical scheme such that

$$P_{\Delta t}\psi = \left(\text{Id} + \Delta t \mathcal{L} + \Delta t^2 \mathcal{L}^2\right)\psi + \Delta t^{5/2}r_{\psi,\Delta t}.$$

A key element to obtain such equalities is to decrease the rejection rate for small time steps.

8.6 Proof of the results

In all the proofs, the constants $C > 0$ and the critical time steps Δt^* may change from line to line. Upon changing V into βV , we may also assume that $\beta = 1$.

8.6.1 Proof of Theorem 8.1

The idea is to rewrite the part of the additive functional involving $-\nabla V(q_t)$ as an approximate martingale. To this end, we introduce the solutions $\Phi_0 = (\Phi_{0,1}, \dots, \Phi_{0,dN})$ of the following Poisson equations

$$\mathcal{L}\Phi_{0,j} = -\nabla_{q_j} V. \tag{8.28}$$

In view of the resolvent estimate (8.6), the functions $\Phi_{0,j}$ are well defined since

$$\int_{\mathcal{M}} \nabla V d\mu = -\nabla \left(\int_{\mathcal{M}} d\mu \right) = 0.$$

In addition, by elliptic regularity, the functions $\Phi_{0,j}$ are smooth. By Ito's lemma, we therefore obtain

$$d\Phi_{0,j}(q_t) = \mathcal{L}\Phi_{0,j}(q_t)dt + \sqrt{2}\nabla\Phi_{0,j}(q_t) \cdot dW_t,$$

so that the integrated displacement with respect to the origin can be rewritten as

$$Q_t - Q_0 = -\int_0^t \nabla V(q_s) ds + \sqrt{2}W_t = \Phi_0(q_t) - \Phi_0(q_0) + \sqrt{2} \int_0^t \left(\text{Id} - \nabla\Phi_0(q_s)\right) dW_s. \tag{8.29}$$

Since Φ_0 is bounded, $(\Phi_0(X_t) - \Phi_0(X_0))/\sqrt{t}$ uniformly vanishes as t goes to infinity. The long time behavior of the process $F^T(Q_t^\varepsilon - Q_0^\varepsilon)/\sqrt{t}$ (for a given direction $F \in \mathbb{R}^{dN}$) is therefore determined by the martingale

$$\mathcal{M}_t^{F,\varepsilon} = \varepsilon \sqrt{\frac{2}{t}} \int_0^{t/\varepsilon^2} \left(F - \nabla \left(F^T \Phi_0(q_s)\right)\right) \cdot dW_s,$$

whose quadratic variation is

$$\langle \mathcal{M}^{F,\varepsilon} \rangle_t = \frac{2\varepsilon^2}{t} \int_0^{t/\varepsilon^2} |F|^2 - 2\nabla \left(F^T \Phi_0(q_s) \right) \cdot F + \left| \nabla \left(F^T \Phi_0(q_s) \right) \right|^2 ds.$$

The ergodic properties of the diffusion process q_t allow to prove that

$$\begin{aligned} F^T \mathcal{D}F &= \lim_{t \rightarrow +\infty} \langle \mathcal{M}^{F,\varepsilon} \rangle_t = 2 \int_{\mathcal{M}} |F|^2 - 2\nabla \left(F^T \Phi_0 \right) \cdot F + \left| \nabla \left(F^T \Phi_0 \right) \right|^2 d\mu \\ &= 2 \int_{\mathcal{M}} |F|^2 - 2 \left(F^T \Phi_0 \right) \left(F^T \nabla V \right) + \mathcal{L} \left[\left(F^T \Phi_0 \right) \right] \left(F^T \Phi_0 \right) d\mu \\ &= 2 \int_{\mathcal{M}} |F|^2 - \left(F^T \Phi_0 \right) \left(F^T \nabla V \right) d\mu \\ &= 2F^T \left(\text{Id} - \int_{\mathcal{M}} \mathcal{L}^{-1} \left(\nabla V \right) \otimes \nabla V d\mu \right) F \end{aligned} \quad (8.30)$$

where, to pass from the first to the second line, we have used an integration by parts to transform the second term in the integral and (8.4) for the third one, while the two last equalities follows from the definition of Φ_0 . At this stage, we note that $F^T \mathcal{D}F = 0$ implies that the integrand of the first equality vanishes almost everywhere, which, by a Cauchy-Schwarz inequality, in turn implies that $\nabla \left(F^T \Phi_0 \right)$ is proportional to F . This is however impossible, and we therefore conclude that \mathcal{D} is positive definite. In addition, the last expression shows that $\mathcal{D} \leq 2\text{Id}$ since \mathcal{L}^{-1} is a positive operator (replace φ by $\mathcal{L}^{-1}\phi$ in (8.5)).

Since \mathcal{L} has a positive spectral gap on $L^2(\mu)$ (see (8.5)), we can write the following operator equality on $\tilde{L}^2(\mu)$:

$$\mathcal{L}^{-1} = - \int_0^{+\infty} e^{t\mathcal{L}} dt.$$

Therefore, for general functions ψ, φ with vanishing average with respect to μ ,

$$\int_0^{+\infty} \mathbb{E} \left[\psi(q_t) \varphi(q_0) \right] dt = \int_{\mathcal{M}} \left(-\mathcal{L}^{-1} \psi \right) \varphi d\mu. \quad (8.31)$$

Combining this result with (8.30) leads to the expression (8.9) of the diffusion matrix \mathcal{D} . To prove the convergence of the processes, two arguments should be made precise (see [Oll94] for an elementary account) :

- the convergence of the finite dimensional laws, which can be obtained very simply here by considering the exponential martingales

$$\exp \left[i\theta \left(\mathcal{M}_t^{F,\varepsilon} - \mathcal{M}_s^{F,\varepsilon} \right) + \frac{\theta^2}{2} \left(\langle \mathcal{M}^{F,\varepsilon} \rangle_t - \langle \mathcal{M}^{F,\varepsilon} \rangle_s \right) \right],$$

which are such that the conditional expectations converge to those of a Brownian motion as $\varepsilon \rightarrow 0$:

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left(\exp \left[i\theta \left(\mathcal{M}_t^{F,\varepsilon} - \mathcal{M}_s^{F,\varepsilon} \right) + \frac{\theta^2}{2} \left(\langle \mathcal{M}^{F,\varepsilon} \rangle_t - \langle \mathcal{M}^{F,\varepsilon} \rangle_s \right) \right] \middle| \mathcal{F}_{s/\varepsilon^2} \right) = \exp \left(-\frac{\theta^2}{2} (t-s) F^T \mathcal{D}F \right),$$

$\mathcal{F}_{s/\varepsilon^2}$ denoting the filtration of events until the time s/ε^2 . Finite-dimensional laws are then obtained by a simple induction, as made precise in [BLP11, Oll94] for instance.

– the tightness of the process, proved using Prohorov’s criterion (see for instance [Bil99])

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \mathbb{P} \left(\sup_{\substack{|t-s| < \delta \\ 0 \leq s < t \leq \tau}} |F^T(Q_t^\varepsilon - Q_s^\varepsilon)| \geq \alpha \right) = 0.$$

This criterion is satisfied in view of tightness of the martingale $\mathcal{M}_t^{F,\varepsilon}$, itself easily obtained using Doob’s inequality.

8.6.2 Proof of Lemma 8.2

We first determine the magnitude of the acceptance probability in the Metropolis algorithm, which reads

$$R_{\Delta t}(q^n, \tilde{q}^{n+1}) = \min \left(1, e^{-\alpha_{\Delta t}(q^n, \tilde{q}^{n+1})} \right),$$

with

$$\begin{aligned} \alpha_{\Delta t}(q, q') &= V(q') - V(q) + \frac{1}{4\Delta t} \left[(q - q' + \Delta t \nabla V(q'))^2 - (q' - q + \Delta t \nabla V(q))^2 \right] \\ &= V(q') - V(q) - \frac{1}{2} \langle q' - q, \nabla V(q') + \nabla V(q) \rangle + \frac{\Delta t}{4} \left(|\nabla V(q')|^2 - |\nabla V(q)|^2 \right). \end{aligned}$$

Using the following expansions (with integral remainders)

$$\begin{aligned} V(q') - V(q) &= \langle \nabla V(q), q' - q \rangle + \frac{1}{2} (q' - q)^T \left[\nabla^2 V(q) \right] (q' - q) + \frac{1}{6} D^3 V(q) \left((q' - q)^{\otimes 3} \right) \\ &\quad + \frac{1}{6} \int_0^1 (1-t)^3 D^4 V((1-t)q + tq') \left((q' - q)^{\otimes 4} \right) dt, \quad (8.32) \\ \nabla V(q') &= \nabla V(q) + \nabla^2 V(q) (q' - q) + \frac{1}{2} D^3 V(q) \left((q' - q)^{\otimes 2} \right) \\ &\quad + \frac{1}{2} \int_0^1 (1-t)^2 D^3 V((1-t)q + tq') \left((q' - q)^{\otimes 3} \right) dt, \end{aligned}$$

a simple computation shows that

$$\alpha_{\Delta t}(q^n, \tilde{q}^{n+1}) = \Delta t^{3/2} \xi(q^n, G^n) + \Delta t^2 \tilde{\xi}_{\Delta t}(q^n, G^n), \quad (8.33)$$

where

$$\xi(q, G) = -\frac{\sqrt{2}}{6} D^3 V(q) \left(G^{\otimes 3} \right) + \frac{\sqrt{2}}{2} \nabla V(q)^T \nabla^2 V(q) G, \quad (8.34)$$

while there exists a constant $C > 0$ such that $|\tilde{\xi}_{\Delta t}(q, G)| \leq C(1 + |G|^6)$ for any $0 \leq \Delta t \leq 1$. Therefore, in view of the inequalities $1 - \theta \leq e^{-\theta} \leq 1 - \theta + \theta^2/2$ when $\theta \geq 0$, we have the following bounds : a lower bound

$$R_{\Delta t}(q^n, \tilde{q}^{n+1}) \geq \min \left(1, 1 - \Delta t^{3/2} \xi(q^n, G^n) + \Delta t^2 \tilde{\xi}_{\Delta t}(q^n, G^n) \right) \geq 1 - \Delta t^{3/2} \xi_+(q^n, G^n) + \Delta t^2 \hat{\xi}_{\Delta t}^{\min}(q^n, G^n)$$

with $\xi_+(q, G) = \max(0, \xi(q, G))$ and $|\hat{\xi}_{\Delta t}^{\min}(q, G)| \leq C(1 + |G|^6)$; as well as an upper bound

$$\begin{aligned} R_{\Delta t}(q^n, \tilde{q}^{n+1}) &\leq \min \left(1, 1 - \Delta t^{3/2} \xi(q^n, G^n) + \Delta t^2 \tilde{\xi}_{\Delta t}(q^n, G^n) + \frac{1}{2} \left(\Delta t^{3/2} \xi(q^n, G^n) - \Delta t^2 \tilde{\xi}_{\Delta t}(q^n, G^n) \right)^2 \right) \\ &\leq 1 - \Delta t^{3/2} \xi_+(q^n, G^n) + \Delta t^2 \hat{\xi}_{\Delta t}^{\max}(q^n, G^n) \end{aligned}$$

with $\left| \widehat{\xi}_{\Delta t}^{\max}(q, G) \right| \leq C(1 + |G|^{12})$. This shows that

$$R_{\Delta t}(q^n, \tilde{q}^{n+1}) = 1 - \Delta t^{3/2} \xi_+(q^n, G^n) + \Delta t^2 \widehat{\xi}_{\Delta t}(q^n, G^n), \quad (8.35)$$

where $\left| \widehat{\xi}_{\Delta t}(q, G) \right| \leq C(1 + |G|^{12})$. The estimate on the average acceptance rate (8.16) is obtained by taking the expectation over all possible realizations of G^n .

Now, using the expression of the Metropolis transition operator (see for instance [LRS10, Section 2.1.2])

$$P_{\Delta t} \psi(q) = \int_{\mathcal{M}} R_{\Delta t}(q, q') T_{\Delta t}(q, q') \psi(q') \mu(dq) + \left(1 - \int_{\mathcal{M}} R_{\Delta t}(q, q') T_{\Delta t}(q, q') \mu(dq) \right) \psi(q),$$

it holds

$$P_{\Delta t} \psi(q) - \psi(q) = \int_{\mathcal{M}} R_{\Delta t}(q, q') T_{\Delta t}(q, q') (\psi(q') - \psi(q)) dq'.$$

We now write $q' = q - \Delta t \nabla V(q) + \sqrt{2\Delta t} g$, so that, using (8.35) to estimate the rejection rate,

$$\begin{aligned} & P_{\Delta t} \psi(q) - \psi(q) \\ &= \int_{\mathbb{R}^{dN}} R_{\Delta t}(q, q - \Delta t \nabla V(q) + \sqrt{2\Delta t} g) \left(\psi(q - \Delta t \nabla V(q) + \sqrt{2\Delta t} g) - \psi(q) \right) \frac{e^{-g^2/2}}{(2\pi)^{dN/2}} dg \\ &= \int_{\mathbb{R}^{dN}} \left(\psi(q - \Delta t \nabla V(q) + \sqrt{2\Delta t} g) - \psi(q) \right) \frac{e^{-g^2/2}}{(2\pi)^{dN/2}} dg \\ &+ \int_{\mathbb{R}^{dN}} \left[R_{\Delta t}(q, q - \Delta t \nabla V(q) + \sqrt{2\Delta t} g) - 1 \right] \left(\psi(q - \Delta t \nabla V(q) + \sqrt{2\Delta t} g) - \psi(q) \right) \frac{e^{-g^2/2}}{(2\pi)^{dN/2}} dg \\ &= \Delta t (\mathcal{L}\psi)(q) + \frac{\Delta t^2}{2} \left([\mathcal{L}^2 + D_1 + D_2] \psi \right)(q) + \Delta t^{5/2} r_{\psi, \Delta t}, \end{aligned}$$

where we have used for the first integral a Taylor expansion at fourth order similar to (8.32) to obtain (see the computations in [LMS13, Section 3.9])

$$D_1 \psi = 2\nabla^2 V : \nabla^2 \psi + \nabla(\Delta V) \cdot \nabla \psi - \nabla V^T (\nabla^2 V) \nabla \psi;$$

and a Taylor expansion at first order for the term involving the rejection rate to obtain

$$D_2 \psi = -\sqrt{2} \left(\int_{\mathbb{R}^{dN}} \xi_+(q, g) g^T \frac{e^{-g^2/2}}{(2\pi)^{dN/2}} dg \right) \cdot \nabla \psi.$$

The remainder $r_{\psi, \Delta t}$ is uniformly bounded in $L^\infty(\mathcal{M})$ for Δt sufficiently small. The conclusion follows by setting $A = (\mathcal{L}^2 + D_1 + D_2)/2$.

Finally, the invariance of μ by $P_{\Delta t}$ implies

$$\forall \Delta t > 0, \quad \int_{\mathcal{M}} P_{\Delta t} \psi d\mu = \int_{\mathcal{M}} \psi d\mu.$$

This equality, together with the expansion (8.14) proves (8.15).

8.6.3 Proof of Lemma 8.3

The proof below is a simplification of the argument presented in [BRH13], made possible since we work on a compact state space. The idea of the proof is to compare the Metropolis dynamics to the continuous dynamics, using the standard, un-metropolized Euler-Maruyama scheme as an intermediate. Alternatively, it would be possible to directly compare the metropolized and un-metropolized schemes, by proving as in [LMS13, Section 3.2] that the standard, un-metropolized Euler-Maruyama scheme is geometrically ergodic since it satisfies a minorization condition.

To prove the first part of the Lemma, it is enough to show that there exists $\rho < 1$ such that, for any $0 < \Delta t \leq \Delta t^*$ and any $q \in \mathbb{T}^{dN}$,

$$\left\| P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - \mu \right\|_{\text{TV}} \leq C\rho^n.$$

Since we work on a compact state space, it is enough to show by Harris' theorem (see the presentation in [HM11, BRH13]) that there exist $\alpha \in (0, 1)$ such that

$$\left\| P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q', \cdot) \right\|_{\text{TV}} \leq 2(1 - \alpha), \quad (8.36)$$

uniformly in $0 < \Delta t \leq \Delta t^*$ and $(q, q') \in \mathcal{M}^2$. We now introduce the transition kernel Q_t of the continuous dynamics (8.1), defined as

$$Q_t \varphi(q) = \mathbb{E}(\varphi(q_t) | q_0 = q) = \left(e^{t\mathcal{L}} \varphi \right) (q) = \int_{\mathcal{M}} Q_t(q, q') \varphi(q') dq',$$

and consider it at time $t = 1$. The transition kernel is well defined and regular since the generator is elliptic. By the triangle inequality,

$$\begin{aligned} \sup_{(q, q') \in \mathcal{M}^2} \left\| P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q', \cdot) \right\|_{\text{TV}} &\leq \sup_{(q, q') \in \mathcal{M}^2} \left\| Q_1(q, \cdot) - Q_1(q', \cdot) \right\|_{\text{TV}} \\ &\quad + 2 \sup_{q \in \mathcal{M}} \left\| P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - Q_1(q, \cdot) \right\|_{\text{TV}}. \end{aligned} \quad (8.37)$$

From [BRH13, Lemma 2.7], since we work on a compact space, we know that there exists $\varepsilon > 0$ such that

$$\sup_{(q, q') \in \mathcal{M}^2} \left\| Q_1(q, \cdot) - Q_1(q', \cdot) \right\|_{\text{TV}} \leq 2(1 - \varepsilon).$$

To control the second term in (8.37), we introduce the transition kernel of the standard, un-metropolized Euler-Maruyama scheme (8.12), denoted by $\tilde{P}_{\Delta t}$, and write

$$\begin{aligned} \sup_{q \in \mathcal{M}} \left\| P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - Q_1(q, \cdot) \right\|_{\text{TV}} &\leq \sup_{q \in \mathcal{M}} \left\| \tilde{P}_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - Q_1(q, \cdot) \right\|_{\text{TV}} \\ &\quad + \sup_{q \in \mathcal{M}} \left\| P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - \tilde{P}_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) \right\|_{\text{TV}} \end{aligned}$$

By [BRH13, Lemma 4.2], the transition kernel of the standard, un-metropolized dynamics is uniformly close to the transition kernel of the continuous dynamics when the state space is compact : There exists $C_{\text{EM}} > 0$ and $\Delta t^* > 0$ such that, for any $0 < \Delta t \leq \Delta t^*$,

$$\sup_{q \in \mathcal{M}} \left\| \tilde{P}_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - Q_1(q, \cdot) \right\|_{\text{TV}} \leq C_{\text{EM}} \sqrt{\Delta t}.$$

It therefore remains to control the distance between the transition rate of the metropolized and un-metropolized dynamics. It is at this stage that the argument of [BRH13,

Lemma 4.6] can be simplified. As in the proof of this lemma, we use a coupling argument, and consider two chains q^i and \tilde{q}^i , corresponding respectively to the metropolized dynamics and the standard un-metropolized one, starting from the same initial condition $q^0 \in \mathcal{M}$. The probability that $q^n \neq \tilde{q}^n$ is bounded from above by the probability that $q^i \neq \tilde{q}^i$ for some $1 \leq i \leq n$, *i.e.* at least one rejection occured along the discrete trajectory. Since the probability to reject the move from q^i to q^{i+1} is $1 - R_{\Delta t}(q^i, q^{i+1})$, it holds

$$\|P_{\Delta t}^n(q, \cdot) - \tilde{P}_{\Delta t}^n(q, \cdot)\|_{\text{TV}} \leq 2 \mathbb{P} [q^n \neq \tilde{q}^n \mid q^0 = q] \leq 2 \sum_{i=1}^n \mathbb{E} \left[\left(1 - R_{\Delta t}(q^i, q^{i+1})\right) \mid q^0 = q \right].$$

In view of (8.35), there exists C_{reject} such that $\mathbb{E} \left[\left(1 - R_{\Delta t}(q^i, q^{i+1})\right) \mid q^0 = q \right] \leq C_{\text{reject}} \Delta t^{3/2}$ for Δt sufficiently small. The sum from $i = 1$ to $n = \lfloor 1/\Delta t \rfloor$ may then be estimated as

$$\|P_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot) - \tilde{P}_{\Delta t}^{\lfloor 1/\Delta t \rfloor}(q, \cdot)\|_{\text{TV}} \leq 2 C_{\text{reject}} \sqrt{\Delta t}.$$

The combination of all previous estimates finally gives (8.36).

For the second part of the lemma, we note that the bounds on the powers of $P_{\Delta t}$ imply that the sum $\sum_{n=0}^{+\infty} P_{\Delta t}^n$ is absolutely convergent in $\tilde{L}^\infty(\mathcal{M})$, and it is then easily checked that

$$(\text{Id} - P_{\Delta t})^{-1} = \sum_{n=0}^{+\infty} P_{\Delta t}^n. \tag{8.38}$$

In fact,

$$\|(\text{Id} - P_h)^{-1} f\|_{L^\infty} = \left\| \sum_{n=0}^{+\infty} P_{\Delta t}^n f \right\|_{L^\infty} \leq \sum_{n=0}^{+\infty} \|P_{\Delta t}^n f\|_{L^\infty} \leq \sum_{n=0}^{+\infty} e^{-\lambda n \Delta t} \|f\|_{L^\infty} \leq \frac{1}{1 - e^{-\lambda \Delta t}} \|f\|_{L^\infty},$$

from which the result immediately follows.

8.6.4 Proof of Theorem 8.4

We follow the strategy of [LMS13, Section 3.8]. The proof starts by noting that the integrated correlation function can be written using \mathcal{L}^{-1} , as made precise in (8.31). The strategy of the proof is to write an approximation of \mathcal{L}^{-1} using the discrete evolution operators $P_{\Delta t}$.

In view of (8.38) and (8.14), it holds

$$\begin{aligned} (-\mathcal{L})^{-1} \psi &= \left(\Delta t \sum_{n=0}^{+\infty} P_{\Delta t}^n \right) \left(\frac{\text{Id} - P_{\Delta t}}{\Delta t} \right) (-\mathcal{L}^{-1}) \psi \\ &= \left(\Delta t \sum_{n=0}^{+\infty} P_{\Delta t}^n \right) \left((\text{Id} + \Delta t A \mathcal{L}^{-1}) \psi + \Delta t^{3/2} r_{\mathcal{L}^{-1}\psi, \Delta t} \right). \end{aligned}$$

Since $\mathcal{L}^{-1}\psi$ still is a smooth function (by elliptic regularity), the remainder $r_{\mathcal{L}^{-1}\psi, \Delta t}$ is uniformly bounded by Lemma 8.14. Note also that since $(\text{Id} - P_{\Delta t})\mathcal{L}^{-1}\psi$ and $A\mathcal{L}^{-1}\psi$ have vanishing averages with respect to μ , the remainder $r_{\mathcal{L}^{-1}\psi, \Delta t}$ has a vanishing average with respect to μ . The above equality shows that

$$\int_{\mathcal{M}} (-\mathcal{L}^{-1}\psi) \varphi d\mu = \Delta t \sum_{n=0}^{+\infty} \int_{\mathcal{M}} \left[P_{\Delta t}^n \left(\tilde{\psi}_{\Delta t} + \Delta t^{3/2} r_{\mathcal{L}^{-1}\psi, \Delta t} \right) \right] \varphi d\mu,$$

where the sum is convergent in view of (8.17). In conclusion,

$$\int_{\mathcal{M}} \left(-\mathcal{L}^{-1}\psi\right) \varphi d\mu = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} \left(\tilde{\psi}_{\Delta t}(q^n) \varphi(q^0)\right) + \Delta t^{3/2} \int_{\mathcal{M}} \left[\left(\frac{\text{Id} - P_{\Delta t}}{\Delta t}\right)^{-1} r_{\mathcal{L}^{-1}\psi, \Delta t}\right] \varphi d\mu,$$

which gives the result, in view of the boundedness of the operator $\left(\frac{\text{Id} - P_{\Delta t}}{\Delta t}\right)^{-1}$ (given by Lemma 8.3).

8.6.5 Proof of Theorem 8.5

We start by highlighting the martingale part of the increments $\delta_{\Delta t}(q^n, U^n, G^n) = Q^{n+1} - Q^n$, similarly to the continuous case (compare (8.29)) :

$$\delta_{\Delta t}(q^n, U^n, G^n) = \left(\delta_{\Delta t}(q^n, U^n, G^n) - \mathbb{E}_{G,U}(\delta_{\Delta t}(q^n, U, G))\right) + \mathbb{E}_{G,U}(\delta_{\Delta t}(q^n, U, G)), \quad (8.39)$$

where the expectation on the right-hand side is over the Gaussian random variable G and the uniform variable U (the configuration q^n being fixed). It will be useful to decompose the increment as

$$\delta_{\Delta t}(q, G, U) = \sqrt{2\Delta t}G - \Delta t \nabla V(q) - \delta_{\Delta t}^{\text{reject}}(q, G, U),$$

with

$$\delta_{\Delta t}^{\text{reject}}(q, G, U) = \mathbf{1}_{U > R_{\Delta t}(q, q - \Delta t \nabla V(q) + \sqrt{2\Delta t}G)} \left(\sqrt{2\Delta t}G - \Delta t \nabla V(q)\right).$$

In view of Lemma 8.6 (which shows in particular that $\delta_{\Delta t}^{\text{reject}}(q, G, U)$ can be thought of as being of order Δt^2), the first term on the right-hand side of (8.39) is equal to $\sqrt{2\Delta t}G^n$ at dominant order in Δt . This term therefore corresponds to the term $\sqrt{2}W_t$ in the decomposition (8.29). The second term in the right-hand side of (8.39) is handled by introducing an appropriate Poisson equation, which is the discrete analogue of (8.28).

Lemma 8.6. *The average increment has the following expansion in powers of Δt :*

$$\mathbb{E}_{G,U}[\delta_{\Delta t}(q, G, U)] = -\Delta t \nabla V(q) - \sqrt{2} \Delta t^2 \mathbb{E}_G[\xi_+(q, G)G] + \Delta t^{5/2} r_{\delta, \Delta t}, \quad (8.40)$$

where $r_{\delta, \Delta t}$ is uniformly bounded for Δt sufficiently small. In addition, there exists a constant $K > 0$ such that $|\delta_{\Delta t}^{\text{reject}}(q, G, U)| \leq K(1 + \sqrt{\Delta t}|G|)$ and, for any $p > 0$,

$$\mathbb{E}_{G,U} \left| \delta_{\Delta t}^{\text{reject}}(q, G, U) \right|^p \leq C \Delta t^{(p+3)/2}. \quad (8.41)$$

Lemma 8.7. *There exists, for any $\Delta t > 0$, a unique function $\Phi_{\Delta t} = (\Phi_{\Delta t, 1}, \dots, \Phi_{\Delta t, dN})$ such that*

$$(P_{\Delta t} - \text{Id}) \Phi_{\Delta t}(q) = \mathbb{E}_{G,U}[\delta_{\Delta t}(q, G, U)]. \quad (8.42)$$

Moreover, recalling the definition (8.28) of Φ_0 ,

$$\Phi_{\Delta t} = \Phi_0 + \Delta t \tilde{\Phi}^1 + \Delta t^{3/2} \Psi_{\Delta t},$$

where $\Psi_{\Delta t}$ is uniformly bounded, and $\tilde{\Phi}^1$ is the unique solution of the Poisson equation

$$\mathcal{L}\tilde{\Phi}^1 = A\Phi_0 - \sqrt{2} \mathbb{E}_G[\xi_+(q, G)G], \quad \int_{\mathcal{M}} \tilde{\Phi}^1 d\mu = 0. \quad (8.43)$$

In view of these results,

$$\begin{aligned} \delta_{\Delta t}(q^n, U^n, G^n) &= \left(\delta_{\Delta t}(q^n, U^n, G^n) - \mathbb{E}_{G,U}(\delta_{\Delta t}(q^n, U, G)) \right) + P_{\Delta t} \Phi_{\Delta t}(q^n) - \Phi_{\Delta t}(q^n) \\ &= M_{\Delta t}^n + \Phi_{\Delta t}(q^{n+1}) - \Phi_{\Delta t}(q^n), \end{aligned}$$

with, upon rewriting $P_{\Delta t} \Phi_{\Delta t}(q^n)$ as $\mathbb{E}_{G,U} [\Phi_{\Delta t}(q^n + \delta_{\Delta t}(q^n, U, G))]$,

$$\begin{aligned} M_{\Delta t}^n &= \left(\delta_{\Delta t}(q^n, U^n, G^n) - \mathbb{E}_{G,U}(\delta_{\Delta t}(q^n, U, G)) \right) - \left(\Phi_{\Delta t}(q^{n+1}) - P_{\Delta t} \Phi_{\Delta t}(q^n) \right) \\ &= \left(\delta_{\Delta t}(q^n, U^n, G^n) - \mathbb{E}_{G,U}(\delta_{\Delta t}(q^n, U, G)) \right) \\ &\quad - \left(\Phi_{\Delta t}(q^n + \delta_{\Delta t}(q^n, U^n, G^n)) - \mathbb{E}_{G,U} [\Phi_{\Delta t}(q^n + \delta_{\Delta t}(q^n, U, G))] \right). \end{aligned} \tag{8.44}$$

The interest of this rewriting is to highlight the fact that $M_{\Delta t}^0$ can be fully understood in terms of the increments $\delta_{\Delta t}(q^n, U^n, G^n)$ (in order to use Lemma 8.8 below). Note that $(M_{\Delta t}^n)_{n \geq 0}$ are stationary, independent martingale increments when $q^0 \sim \mu$ (since in this case $q^n \sim \mu$ for all $n \geq 0$). This shows that

$$Q^n - Q^0 = \Phi_{\Delta t}(q^n) - \Phi_{\Delta t}(q^0) + \sum_{m=0}^{n-1} M_{\Delta t}^m.$$

Since $\Phi_{\Delta t}$ is uniformly bounded as $\Delta t \rightarrow 0$, we obtain

$$\mathcal{D}_{\Delta t}^{\text{Einstein}} = \mathbb{E} \left(\frac{M_{\Delta t}^0}{\sqrt{\Delta t}} \otimes \frac{M_{\Delta t}^0}{\sqrt{\Delta t}} \right).$$

We now expand $M_{\Delta t}^0$ in powers of Δt . By Lemma 8.7, it is possible to replace the function $\Phi_{\Delta t}$ in the second term on the right-hand side of (8.44) by $\Phi_0 + \Delta t \tilde{\Phi}^1$, up to a remainder of order $\Delta t^{3/2}$. We next use the following lemma to compute the cross correlation between the various functions of Δt appearing in $M_{\Delta t}^0 \otimes M_{\Delta t}^0$.

Lemma 8.8. *For any smooth functions f, g , growing at most polynomially,*

$$\frac{1}{\Delta t} \mathbb{E}_{G,U} \left[\left(f(\delta_{\Delta t}(q, G, U)) - \bar{f}(q) \right) \left(g(\delta_{\Delta t}(q, G, U)) - \bar{g}(q) \right) \right] = 2 \nabla f(0)^T \nabla g(0) + \Delta t r_{f,g,\Delta t},$$

with $\bar{f}(q) = \mathbb{E}_{G,U} [f(\delta_{\Delta t}(q, G, U))]$, and where the remainder $r_{f,g,\Delta t}$ is uniformly bounded for Δt sufficiently small.

The conclusion then follows by applying this result with the functions f, g replaced by $x \mapsto F^T x$, $x \mapsto F^T \Phi_0(q + x)$ and $x \mapsto F^T \tilde{\Phi}^1(q + x)$ for a given test direction F . Indeed,

$$F^T \mathcal{D}_{\Delta t}^{\text{Einstein}} F = 2 \int_{\mathcal{M}} |F|^2 - 2 \nabla \left(F^T \Phi_0 \right) \cdot F + \left| \nabla \left(F^T \Phi_0 \right) \right|^2 d\mu + \Delta t r_{F,\Delta t},$$

with $|r_{F,\Delta t}|/|F|^2$ uniformly bounded as $\Delta t \rightarrow 0$. Manipulations similar to the ones leading to (8.30) finally give the claimed result.

Remark 8.9. In order to characterize the leading order term in the error and prove that the subleading order term indeed is of order $\Delta t^{3/2}$, it would be necessary to compute correlation terms involving components of the remainder term $\Psi_{\Delta t}$. This is not possible as such because the regularity of $\Psi_{\Delta t}$ is not established, and obtaining regularity result from (8.42) is difficult. An expansion of $\Phi_{\Delta t}$ up to Δt^2 terms (as $\Phi_0 + \Delta t \tilde{\Phi}^1 + \Delta t^{3/2} \tilde{\Phi}^{3/2} + \Delta t^2 \tilde{\Psi}_{\Delta t}$) is therefore needed in order not have to treat correlations involving the remainder $\Psi_{\Delta t}$. This, in turn, would require an expansion of $P_{\Delta t}$ up to remainders of order Δt^3 , instead of $\Delta t^{5/2}$ as in (8.14). Although this does not pose any problem in principle, we chose not to follow this path in order to keep the arguments as simple as possible.

We conclude this section with the proofs of the technical results quoted above.

Proof of Lemma 8.6. We first write an expansion of $\mathbb{E}_{G,U}(\delta_{\Delta t}(q, G, U))$ in fractional powers of Δt :

$$\begin{aligned} \mathbb{E}_{G,U}[\delta_{\Delta t}(q, G, U)] + \Delta t \nabla V(q) &= \mathbb{E}_{G,U} \left[\delta_{\Delta t}(q, G, U) + \Delta t \nabla V(q) - \sqrt{2\Delta t} G \right] \\ &= \mathbb{E}_{G,U} \left[\left(\mathbf{1}_{U \leq R_{\Delta t}(q, q - \Delta t \nabla V(q) + \sqrt{2\Delta t} G)} - 1 \right) \left(-\Delta t \nabla V(q) + \sqrt{2\Delta t} G \right) \right] \\ &= \mathbb{E}_G \left[\left(R_{\Delta t}(q, q - \Delta t \nabla V(q) + \sqrt{2\Delta t} G) - 1 \right) \left(-\Delta t \nabla V(q) + \sqrt{2\Delta t} G \right) \right]. \end{aligned}$$

In view of (8.35), it holds $\mathbb{E}_{G,U}[\delta_{\Delta t}(q, G, U)] + \Delta t \nabla V(q) = -\sqrt{2}\Delta t^2 \mathbb{E}_G[\xi_+(q, G)G] + \Delta t^{5/2} r_{\delta, \Delta t}$, which gives (8.40).

The bound (8.41) is a straightforward consequence of the equality

$$\mathbb{E}_{G,U} \left| \delta_{\Delta t}^{\text{reject}}(q, G, U) \right|^p = \mathbb{E}_G \left[\left(1 - R_{\Delta t}(q, q - \Delta t \nabla V(q) + \sqrt{2\Delta t} G) \right) \left| \sqrt{2\Delta t} G - \Delta t \nabla V(q) \right|^p \right],$$

while $\left| \delta_{\Delta t}^{\text{reject}}(q, G, U) \right| \leq \left| \sqrt{2\Delta t} G - \Delta t \nabla V(q) \right|$ immediately shows that $\left| \delta_{\Delta t}^{\text{reject}}(q, G, U) \right| \leq K \left(1 + \sqrt{\Delta t} |G| \right)$. \square

Proof of Lemma 8.7. We introduce the normalized average increment, defined as the following periodic function :

$$\bar{\delta}_{\Delta t}(q) = \frac{\mathbb{E}_{G,U}[\delta_{\Delta t}(q, G, U)]}{\Delta t}.$$

Lemma 8.3 shows that $\Phi_{\Delta t}$ is well defined provided the periodic function $\bar{\delta}_{\Delta t}$ has a vanishing average with respect to μ . To prove this statement, we start from

$$\mathbb{E}_{G,U}[\delta_{\Delta t}(q, G, U)] = \int_{\mathbb{R}^{dN}} R_{\Delta t}(q, q') T_{\Delta t}(q, q') (q' - q) dq'.$$

It is easily seen that

$$\begin{aligned} \int_{\mathcal{M}} \mathbb{E}_{G,U}[\delta_{\Delta t}(q, G, U)] \mu(dq) &= \int_{\mathcal{M}} \int_{\mathbb{R}^{dN}} R_{\Delta t}(q, q') T_{\Delta t}(q, q') (q' - q) dq' \mu(dq) \\ &= \sum_{n \in \mathbb{Z}^{dN}} \int_{\mathcal{M}} \int_{\mathcal{M}} R_{\Delta t}(q, q' + nL) T_{\Delta t}(q, q' + nL) (q' - q + nL) dq' \mu(dq) = \sum_{n \in \mathbb{Z}^{dN}} I_n, \end{aligned}$$

with

$$I_n = \int_{\mathcal{M}} \int_{\mathcal{M}} \min \left(\mu(dq') T_{\Delta t}(q' + nL, q), \mu(dq) T_{\Delta t}(q, q' + nL) \right) (q' - q + nL) dq' dq.$$

Since q and q' play symmetric roles, I_0 vanishes. For $n \geq 0$, we obtain, by first exchanging the names of the dummy variables q and q' and then using $T_{\Delta t}(q + nL, q') = T_{\Delta t}(q, q' - nL)$ as well as the invariance of μ by translations of the periodic cell,

$$\begin{aligned} I_n &= \int_{\mathcal{M}} \int_{\mathcal{M}} \min \left(\mu(dq') T_{\Delta t}(q' + nL, q), \mu(dq) T_{\Delta t}(q, q' + nL) \right) (q' - q + nL) dq' dq \\ &= \int_{\mathcal{M}} \int_{\mathcal{M}} \min \left(\mu(dq) T_{\Delta t}(q + nL, q'), \mu(dq') T_{\Delta t}(q', q + nL) \right) (q - q' + nL) dq dq' \\ &= \int_{\mathcal{M}} \int_{\mathcal{M}} \min \left(\mu(dq) T_{\Delta t}(q, q' - nL), \mu(dq') T_{\Delta t}(q' - nL, q) \right) (q - (q' - nL)) dq dq' \\ &= - \int_{\mathcal{M}} \int_{\mathcal{M}} \min \left(\mu(dq') T_{\Delta t}(q' - nL, q), \mu(dq) T_{\Delta t}(q, q' - nL) \right) (q' - q - nL) dq dq'. \end{aligned}$$

Therefore, $I_n = -I_{-n}$, which allows to conclude that

$$\forall \Delta t > 0, \quad \int_{\mathcal{M}} \bar{\delta}_{\Delta t}(q) \mu(dq) = 0. \quad (8.45)$$

To obtain an expansion of $\Phi_{\Delta t}$ in terms of fractional powers of Δt , we rely on (8.40), which implies that

$$\bar{\delta}_{\Delta t}(q) = -\nabla V(q) - \sqrt{2} \Delta t \mathbb{E}_G [\xi_+(q, G)G] + \Delta t^{3/2} r_{\delta, \Delta t}. \quad (8.46)$$

The function $\tilde{\Phi}^1$ introduced in (8.43) is indeed well defined since Af has a vanishing average with respect to μ for any function f (see (8.15)), while the condition (8.45), together with the expansion (8.46), shows that the average of $\mathbb{E}_G [\xi_+(q, G)G]$ with respect to μ also vanishes. Now, consider the following difference, relying on the expansion (8.14) and the definition of Φ_0 :

$$\begin{aligned} - \left(\frac{\text{Id} - P_{\Delta t}}{\Delta t} \right) (\Phi_{\Delta t} - \Phi_0 - \Delta t \tilde{\Phi}^1) &= \bar{\delta}_{\Delta t}(q) - (\mathcal{L} + \Delta t A) (\Phi_0 + \Delta t \tilde{\Phi}^1) + \Delta t^{3/2} r_{\Delta t} \\ &= \Delta t (A\Phi_0 - \sqrt{2} \mathbb{E}_G [\xi_+(q, G)G] - \mathcal{L}\tilde{\Phi}^1) + \Delta t^{3/2} \tilde{r}_{\Delta t}. \end{aligned}$$

The first term on the right-hand side of the last equality vanishes by definition of $\tilde{\Phi}^1$. The remainder $\tilde{r}_{\Delta t}$ has a vanishing average with respect to μ since it belongs to $\text{Ran}(I - P_{\Delta t})$. Lemma 8.3 then shows that there exists a constant $C > 0$ and $\Delta t^* > 0$ such that, for any $0 < \Delta t \leq \Delta t^*$,

$$\left\| \Phi_{\Delta t} - \Phi_0 - \Delta t \tilde{\Phi}^1 \right\|_{L^\infty(\mathcal{M})} \leq C \Delta t^{3/2}.$$

This gives the result upon defining $\Psi_{\Delta t} = \Delta t^{-3/2} (\Phi_{\Delta t} - \Phi_0 - \Delta t \tilde{\Phi}^1)$. □

Proof of Lemma (8.8). A Taylor expansion with integral remainder gives

$$\begin{aligned} f(\delta_{\Delta t}(q, G, U)) &= f(0) + \nabla f(0)^T \delta_{\Delta t}(q, G, U) + \frac{1}{2} \delta_{\Delta t}(q, G, U)^T \nabla^2 f(0) \delta_{\Delta t}(q, G, U) \\ &\quad + \frac{1}{2} \int_0^1 (1 - \theta)^2 D^3 f(\theta \delta_{\Delta t}(q, G, U)) (\delta_{\Delta t}(q, G, U)^{\otimes 3}) d\theta. \end{aligned}$$

A simple computation using (8.40) and (8.41) shows that

$$\bar{f}(q) = f(0) - \Delta t (\nabla f(0)^T \nabla V(q) + \Delta f(0)) + \Delta t^2 r_{f, \Delta t},$$

where $r_{f, \Delta t}$ is uniformly bounded for Δt sufficiently small. Therefore,

$$\begin{aligned} &f(\delta_{\Delta t}(q, G, U)) - \bar{f}(q) \\ &= \nabla f(0)^T (\delta_{\Delta t}(q, G, U) + \Delta t \nabla V(q)) + \frac{1}{2} \nabla^2 f(0) : (\delta_{\Delta t}(q, G, U) \otimes \delta_{\Delta t}(q, G, U) - 2\Delta t \text{Id}) \\ &\quad + \frac{1}{2} \int_0^1 (1 - \theta)^2 D^3 f(\theta \delta_{\Delta t}(q, G, U)) (\delta_{\Delta t}(q, G, U)^{\otimes 3}) d\theta - \Delta t^2 r_{f, \Delta t} \\ &= \sqrt{2\Delta t} \nabla f(0)^T G + \Delta t \nabla^2 f(0) : (G \otimes G - \text{Id}) - \frac{\Delta t^{3/2}}{\sqrt{2}} \nabla^2 f(0) : (\nabla V(q) \otimes G + G \otimes \nabla V(q)) \\ &\quad + \nabla f(0)^T \delta_{\Delta t}^{\text{reject}}(q, G, U) \\ &\quad + \frac{1}{2} \nabla^2 f(0) : (\delta_{\Delta t}^{\text{reject}}(q, G, U) \otimes \delta_{\Delta t}(q, G, U) + \delta_{\Delta t}(q, G, U) \otimes \delta_{\Delta t}^{\text{reject}}(q, G, U) + \delta_{\Delta t}^{\text{reject}}(q, G, U) \otimes \delta_{\Delta t}^{\text{reject}}(q, G, U)) \\ &\quad + \frac{1}{2} \int_0^1 (1 - \theta)^2 D^3 f(\theta \delta_{\Delta t}(q, G, U)) (\delta_{\Delta t}(q, G, U)^{\otimes 3}) d\theta - \Delta t^2 r_{f, \Delta t}. \end{aligned}$$

A simple computation finally shows that

$$\begin{aligned} & \frac{1}{\Delta t} \mathbb{E}_{G,U} \left[\left(f(\delta_{\Delta t}(q, G, U)) - \bar{f}(q) \right) \left(g(\delta_{\Delta t}(q, G, U)) - \bar{g}(q) \right) \right] \\ & = 2 \mathbb{E}_G \left[\left(\nabla f(0)^T G \right) \left(\nabla g(0)^T G \right) \right] + \Delta t r_{f,g,\Delta t}, \end{aligned}$$

where the remainder $r_{f,g,\Delta t}$ is uniformly bounded. □

Chapitre 9

Improving trajectorial accuracy for the Metropolis-adjusted Langevin algorithm

Les résultats de cette section ont été obtenus en collaboration avec Ahmed-Amine Hommann et Gabriel Stoltz.

As in the previous chapter, we consider the Brownian dynamics

$$dq_t = -\nabla V(q_t) dt + \sqrt{2} dW_t \quad (9.1)$$

with some smooth potential V , but this time the space is \mathbb{R}^d , rather than the torus. We denote by

$$\mu(dq) := \frac{1}{Z} \exp(-V(q)) dq$$

its invariant measure. We also denote by L the generator of the dynamic, which acts on smooth test functions f as

$$Lf(q) = -\nabla V(q) \cdot \nabla f(q) + \Delta f(q) \quad (9.2)$$

Since we work on a non-compact space, we shall need a few growth assumptions on the potential V :

Assumption 9.1. We will make the following growth assumptions on the potential V (which are taken from [BRVE09], except for (vi)) :

- (i) There exists a real constant $K > 0$ such that $V(q) \geq K|q|$;
- (ii) For every integer $\ell \geq 1$, there exists constants $\delta_\ell > 0$ and $M_\ell > 0$ such that

$$L\{V(q)^\ell\} \leq -\delta_\ell V(q) + M_\ell;$$

- (iii) There exists a constant $K > 0$ such that

$$|\nabla V(q) - \nabla V(q')| \leq K(V(q) + V(q'))|q - q'| \quad \forall q, q' \in \mathbb{R}^d;$$

- (iv) There exists a real constant $K > 0$ such that

$$\langle -\nabla V(q) + \nabla V(q'), q - q' \rangle \leq K|q - q'|^2 \quad \forall q, q' \in \mathbb{R}^d;$$

- (v) There exists a real constant $K > 0$ such that

$$\|D^5 V(q)\| \vee \|D^4 V(q)\| \vee \|D^3 V(q)\| \vee \|D^2 V(q)\| \vee |\nabla V(q)| \leq K(1 + V(q));$$

- (vi) The second derivative of the potential Hess V is bounded from below (in the sense of symmetric matrices) by some $-\kappa Id$.

As previously, we can consider the Metropolis-adjusted Langevin algorithm. Proposals are generated according to the forward Euler scheme

$$q^{n+1} = q^n - \beta \Delta t \nabla V(q^n) + \sqrt{2\Delta t} G^n, \quad (9.3)$$

and accepted with probability

$$R_{\Delta t}(q^n, \tilde{q}^{n+1}) = \frac{e^{-\beta V(\tilde{q}^{n+1})} T_{\Delta t}(\tilde{q}^{n+1}, q^n)}{e^{-\beta V(q^n)} T_{\Delta t}(q^n, \tilde{q}^{n+1})},$$

where

$$T_{\Delta t}(q, q') = \left(\frac{1}{4\pi\Delta t} \right)^{dN/2} \exp \left(-\frac{|q' - q + \beta \Delta t \nabla V(q)|^2}{4\Delta t} \right).$$

If the proposal is rejected, we remain at q^n .

This time, we are interested in the strong trajectorial accuracy at equilibrium of the scheme, defined as

$$\sup_{t \in [0, T]} \left(\mathbb{E}^\mu \left[\mathbb{E}^x [|q_{[t/\Delta t]} - q_t|^2] \right] \right)^{1/2}. \quad (9.4)$$

In [BRVE09], Bou-Rabee and Vanden-Eijnden proved the following sharp bound for the strong trajectorial accuracy at equilibrium of MALA

Theorem 9.1. *Under the assumptions 9.1, for any $T > 0$ there exists $\Delta t^* > 0$ and $C(T) > 0$ such that, for any $\Delta t < \Delta t^*$ and for all $t \in [0, T]$,*

$$\left(\mathbb{E}^\mu \left[\mathbb{E}^x [|q_{[t/\Delta t]} - q_t|^2] \right] \right)^{1/2} \leq C(T) (\Delta t)^{3/4}.$$

The reason why it is of order 3/4 is because the small rejection probability of MALA, while it allows to avoid blowup, corrupts the dynamic and reduces its accuracy.

In this chapter, we introduce a modified MALA scheme that has better strong trajectorial accuracy at equilibrium. We do this by adding perturbations to the explicit Euler proposal, small enough to disappear in the scaling limit, but large enough to reduce the rejection rate of the Metropolis-Hastings procedure. Rather than being in $\Delta t^{3/2}$, the rejection rate will behave like Δt^2 , which will allow us to obtain better estimates on the strong accuracy of the scheme.

Once we will have identified the right perturbation, the scheme of proof will be the same as in [BRVE09], except for the extra terms and increased accuracy that come from the perturbation.

We consider the modified explicit Euler proposal

$$q^{n+1} = q^n - \Delta t \nabla V(q^n) + \Delta t^2 F(q^n) + \sqrt{2\Delta t} (Id + \Delta t \sigma(q^n))^{-1/2} G^n, \quad (9.5)$$

where $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \rightarrow \mathcal{S}_d(\mathbb{R})$ will be specified later. We assume that, for Δt small enough, $(Id + \Delta t \sigma(q))$ is a positive-definite matrix for all $q \in \mathbb{R}^d$ (which will be the case for the choice we shall later make).

The Metropolis-Hastings procedure then uses the acceptance rate

$$R_{\Delta t}(q^n, \tilde{q}^{n+1}) = \frac{e^{-\beta V(\tilde{q}^{n+1})} T_{\Delta t}(\tilde{q}^{n+1}, q^n)}{e^{-\beta V(q^n)} T_{\Delta t}(q^n, \tilde{q}^{n+1})},$$

where

$$T_{\Delta t}(q, q') = \left(\frac{1}{4\pi\Delta t} \right)^{d/2} \det(Id + h\sigma(q))^{1/2}$$

$$\times \exp \left(- \frac{\langle (Id + \Delta t \sigma(q))(q' - q + \Delta t \nabla V(q) - \Delta t^2 F(q)), (q' - q + \Delta t \nabla V(q) - \Delta t^2 F(q)) \rangle}{4\Delta t} \right).$$

We therefore have

$$\begin{aligned} -\log R_{\Delta t}(q, q') &= V(q') - V(q) - \frac{1}{2} \langle q' - q, \nabla V(q') + \nabla V(q) \rangle \\ &+ \frac{\Delta t}{4} (|\nabla V(q')|^2 - |\nabla V(q)|^2) - \frac{1}{2} \log \det(Id + \Delta t \sigma(q')) + \frac{1}{2} \log \det(Id + \Delta t \sigma(q)) \\ &+ \frac{\Delta t}{2} \langle q' - q, F(q') + F(q) \rangle - \frac{\Delta t^2}{2} \langle \nabla V(q'), F(q') \rangle + \frac{\Delta t^2}{2} \langle \nabla V(q), F(q) \rangle \\ &+ \frac{\Delta t^3}{4} (|F(q')|^2 - |F(q)|^2) + \frac{1}{4} \langle \sigma(q')(q' - q), (q' - q) \rangle - \frac{1}{4} \langle \sigma(q)(q' - q), (q' - q) \rangle \\ &- \frac{\Delta t}{2} \langle \sigma(q')(q' - q), \nabla V(q') \rangle - \frac{\Delta t}{2} \langle \sigma(q)(q' - q), \nabla V(q) \rangle \\ &- \frac{\Delta t^2}{2} \langle \sigma(q')(q' - q), F(q') \rangle - \frac{\Delta t^2}{2} \langle \sigma(q)(q' - q), F(q) \rangle \\ &+ \frac{\Delta t^2}{4} \langle \sigma(q') \nabla V(q'), \nabla V(q') \rangle - \frac{\Delta t^2}{4} \langle \sigma(q) \nabla V(q), \nabla V(q) \rangle \\ &- \frac{\Delta t^3}{2} \langle \sigma(q') \nabla V(q'), F(q') \rangle + \frac{\Delta t^3}{2} \langle \sigma(q) \nabla V(q), F(q) \rangle \\ &+ \frac{\Delta t^4}{4} \langle \sigma(q') F(q'), F(q') \rangle - \frac{\Delta t^4}{4} \langle \sigma(q) F(q), F(q) \rangle. \end{aligned} \tag{9.6}$$

If we take $q' = q - \Delta t \nabla V(q) + \Delta t^2 F(q) + \sqrt{2h}(\text{Id} + \sigma(q))^{-1/2} G$ and do a Taylor expansion in Δt , the first non-zero term is in $\Delta t^{3/2}$, and is

$$\begin{aligned} &-\frac{2\sqrt{2}}{12} D^3 V(q)(G^{\otimes 3}) + \frac{\sqrt{2}}{2} \langle \text{Hess } V(q) \nabla V(q), G \rangle - \frac{\sqrt{2}}{2} \langle \nabla \text{Tr}(\sigma)(q), G \rangle \\ &+ \sqrt{2} \langle F(q), G \rangle + \frac{\sqrt{2}}{2} D\sigma(q)(G^{\otimes 3}) - \sqrt{2} \langle \sigma(q) G, \nabla V(q) \rangle. \end{aligned} \tag{9.7}$$

Since we want this to be equal to zero, it leads to the conditions

$$D\sigma(q) - \frac{1}{3} D^3 V(q) = 0, \quad \text{Hess } V(q) \nabla V(q) - \nabla \text{Tr}(\sigma)(q) + 2F(q) - 2\sigma(q) \nabla V(q) = 0 \tag{9.8}$$

for all $q \in \mathbb{R}^d$. This automatically leads to the choices

$$\sigma(q) = \frac{1}{3} \text{Hess } V(q), \quad F(q) = -\frac{1}{6} \text{Hess } V(q) \nabla V(q) + \frac{1}{6} \nabla \Delta V(q). \tag{9.9}$$

Under assumption (vi), for Δt small enough $(\text{Id} + \frac{\Delta t}{3} \text{Hess } V(q))^{-1/2}$ makes sense for every q , and the scheme is well defined. Moreover, since we use a Metropolis-Hastings procedure, it is reversible, and has invariant measure μ .

With this modified proposal, we have the following strong trajectorial accuracy estimate :

Theorem 9.2. *The modified MALA scheme with proposal given by (9.5), with σ and F chosen as in (9.9), has strong trajectorial accuracy at equilibrium of order 1. That is for any $T > 0$ there exists $\Delta t^* > 0$ and $C(T) > 0$ such that, for any $\Delta t < \Delta t^*$ and for all $t \in [0, T]$,*

$$\left(\mathbb{E}^\mu \left[\mathbb{E}^x [|q^{[t/\Delta t]} - q_t|^2] \right] \right)^{1/2} \leq C(T) \Delta t.$$

Remark 9.3. As for the usual MALA, the main reason we cannot extend this result to the case where we start from a measure that is not the equilibrium measure is the lack of geometric ergodicity for MALA in the case of a non-compact space.

Remark 9.4. For practical purposes, to reduce the number of computations we have to make when generating the proposal, it seems that it would be better to replace the term $(\text{Id} + \frac{\Delta t}{3} \text{Hess } V(q))^{-1/2} G$ by $(\text{Id} - \frac{\Delta t}{6} \text{Hess } V(q))G$, or even $G - \frac{1}{6}(\nabla V(q + \Delta t G) - \nabla V(q))$. In this case, the strong trajectorial accuracy at equilibrium would still be of order 1.

To prove our theorem, we continue to investigate the behavior of the acceptance rate of the scheme. We now look at the term in Δt^2 , and we obtain

$$\begin{aligned}
 & -\frac{1}{6}D^4V(q)(G^{\otimes 4}) + \frac{1}{2}D^3V(q)(\nabla V(q), G^{\otimes 2}) - \frac{1}{2}\langle \text{Hess } V(q)\nabla V(q), \nabla V(q) \rangle \\
 & + \frac{1}{2}D^3V(q)(G^{\otimes 2}), \nabla V(q) \rangle + \frac{1}{2}|\text{Hess } V(q)G|^2 + \frac{1}{2}\langle \nabla \text{Tr}(\sigma)(q), \nabla V(q) \rangle \\
 & - \frac{1}{2}D^2\text{Tr}(\sigma)(q)(G^{\otimes 2}) - \langle \nabla V(q), F(q) \rangle + \langle DF(q)G, G \rangle \\
 & - \frac{3}{2}D\sigma(q)(\nabla V(q), G^{\otimes 2}) + \frac{1}{2}D^2\sigma(q)(G^{\otimes 4}) + \langle \sigma(q)\nabla V(q), \nabla V(q) \rangle \\
 & - D\sigma(q)(\nabla V(q), G^{\otimes 2}) - \langle \sigma(q)G, \text{Hess } V(q)G \rangle
 \end{aligned} \tag{9.10}$$

When we plug in the expressions of (9.9), we get zero (!). This surprising fact is confirmed by numerical simulations, which show that the acceptance rate behaves like $1 - C\Delta t^{5/2}$.

So we need to look at the term in $\Delta t^{5/2}$ the development, which is

$$\begin{aligned}
& -\frac{\sqrt{2}}{20}D^5V(q)(G^{\otimes 5}) + \frac{\sqrt{2}}{4}D^3V(q)(G^{\otimes 2}, \sigma(q)G) - \frac{\sqrt{2}}{4}D^3V(q)(G, \nabla V(q), \nabla V(q)) \\
& + \frac{\sqrt{2}}{3}D^4V(q)(G^{\otimes 3}, \nabla V(q)) - \frac{\sqrt{2}}{4}\langle \nabla V(q), \text{Hess } V(q)\sigma(q)G \rangle \\
& - \frac{\sqrt{2}}{2}D^3V(q)(G, \nabla V(q), \nabla V(q)) + \frac{\sqrt{2}}{2}D^3V(q)(G^{\otimes 2}, \text{Hess } V(q)G) \\
& - \frac{\sqrt{2}}{2}\langle \text{Hess } V(q)G, \text{Hess } V(q)\nabla V(q) \rangle + \frac{\sqrt{2}}{6}D^4V(q)(G^{\otimes 3}, \nabla V(q)) \\
& + \frac{\sqrt{2}}{4}\langle \nabla \text{Tr}(\sigma)(q), \sigma(q)G \rangle + \frac{\sqrt{2}}{2}\text{Tr}(\sigma(q))\langle \nabla \text{Tr}(\sigma)(q)G \rangle - \frac{\sqrt{2}}{2}D_{\text{Id}}^2(\det)(D\sigma(q)G, \sigma(q)) \\
& - \frac{\sqrt{2}}{2}\langle \sigma(q)G, F(q) \rangle - \frac{\sqrt{2}}{2}\langle \nabla V(q), DF(q)G \rangle - \frac{\sqrt{2}}{2}\langle G, DF(q)\nabla V(q) \rangle \\
& + \frac{\sqrt{2}}{2}\langle G, D^2F(q)(G^{\otimes 2}) \rangle - \frac{\sqrt{2}}{2}\langle \text{Hess } V(q)G, F(q) \rangle - \frac{\sqrt{2}}{2}\langle \nabla V(q), DF(q)G \rangle \\
& + \frac{\sqrt{2}}{6}\langle D^3\sigma(q)(G^{\otimes 3})G, G \rangle - \sqrt{2}D^2\sigma(q)(G^{\otimes 3}, \nabla V(q)) + \frac{3\sqrt{2}}{4}D\sigma(q)(G, \nabla V(q), \nabla V(q)) \\
& - \frac{3\sqrt{2}}{4}D\sigma(q)(G^{\otimes 2}, \sigma(q)G) + \frac{\sqrt{2}}{2}\langle \sigma(q)^2G, \nabla V(q) \rangle - \frac{\sqrt{2}}{2}\langle D^2\sigma(q)(G^{\otimes 2})G, \nabla V(q) \rangle \\
& + \sqrt{2}\langle D\sigma(q)(G)\nabla V(q), \nabla V(q) \rangle - \frac{\sqrt{2}}{2}\langle \sigma(q)G, D^3V(q)(G^{\otimes 2}) \rangle \\
& - \sqrt{2}\langle D\sigma(q)(G)G, \text{Hess } V(q)G \rangle + \frac{\sqrt{2}}{2}\langle \sigma(q)\nabla V(q), \text{Hess } V(q)G \rangle \\
& + \frac{\sqrt{2}}{2}\langle \sigma(q)G, \text{Hess } V(q)\nabla V(q) \rangle - \sqrt{2}\langle \sigma(q)G, F(q) \rangle \\
& + \frac{\sqrt{2}}{4}\langle D\sigma(q)(G)\nabla V(q), \nabla V(q) \rangle + \frac{\sqrt{2}}{2}\langle \sigma(q)\nabla V(q), \text{Hess } V(q)G \rangle \tag{9.11}
\end{aligned}$$

When we plug in the expressions of σ , we get

$$\begin{aligned}
& \frac{\sqrt{2}}{180}D^5V(q)(G^{\otimes 5}) - \frac{\sqrt{2}}{12}D^3V(q)(G, \nabla V(q), \nabla V(q)) \\
& - \frac{\sqrt{2}}{36}\langle \text{Hess } V(q)G, \text{Hess } V(q)\nabla V(q) \rangle - \frac{\sqrt{2}}{6}D^3V(q)(G^{\otimes 2}, \text{Hess } V(q)G) \\
& - \frac{\sqrt{2}}{18}D_{\text{Id}}^2(\det)(D^3V(q)G, \text{Hess } V(q)) + \frac{\sqrt{2}}{36}\langle \nabla \Delta V(q), \text{Hess } V(q)G \rangle \\
& + \frac{\sqrt{2}}{18}\Delta V(q)\langle \nabla \Delta V(q), G \rangle + \frac{\sqrt{2}}{2}\langle G, D^2F(q)(G^{\otimes 2}) \rangle \\
& - \frac{3\sqrt{2}}{2}\langle \nabla V(q), DF(q)G \rangle - \sqrt{2}\langle \text{Hess } V(q)G, F(q) \rangle. \tag{9.12}
\end{aligned}$$

And finally, when we plug in the expression of F , we get

$$\begin{aligned}
& \frac{\sqrt{2}}{180} D^5 V(q)(G^{\otimes 5}) + \frac{\sqrt{2}}{6} D^3 V(q)(G, \nabla V(q), \nabla V(q)) \\
& + \frac{7\sqrt{2}}{18} \langle \text{Hess } V(q)G, \text{Hess } V(q)\nabla V(q) \rangle - \frac{5\sqrt{2}}{12} D^3 V(q)(G^{\otimes 2}, \text{Hess } V(q)G) \\
& - \frac{\sqrt{2}}{18} D_{\text{Id}}^2(\det)(D^3 V(q)G, \text{Hess } V(q)) - \frac{5\sqrt{2}}{36} \langle \nabla \Delta V(q), \text{Hess } V(q)G \rangle \\
& + \frac{\sqrt{2}}{18} \Delta V(q) \langle \nabla \Delta V(q), G \rangle - \frac{\sqrt{2}}{4} \langle \text{Hess}(\Delta V)(q)G, \nabla V(q) \rangle \\
& + \frac{\sqrt{2}}{12} \langle D^3(\Delta V)(q)(G^{\otimes 3}) - \frac{\sqrt{2}}{12} D^4 V(q)(G^{\otimes 3}, \nabla V(q)) \rangle \tag{9.13}
\end{aligned}$$

This is of course a very unwieldy expression. Thankfully, we only need it in order to obtain the following Lemma :

Lemma 9.5. *Under our growth assumptions on V , there exists a constant $C_\ell > 0$ such that $\mathbb{E}^q[(1 - R_{\Delta t}(q, q'))^{2\ell}] \leq C_\ell(\Delta t)^{5\ell}(1 + V(q)^{8\ell})$ for any $q \in \mathbb{R}^d$ and any $\ell \geq 1$, where q' is the proposal generated by our modified Euler scheme.*

To prove this Lemma, we just have to plug in the Taylor expansion we just obtained in the expression of $R_{\Delta t}$, and use Laplace's method to rigorously justify the bound. The exponent 8ℓ comes from the fact that the largest term in the previous expression (when q is large) is $D_{\text{Id}}^2(\det)(D^3 V(q)G, \text{Hess } V(q))$, which under our growth assumptions on V is bounded by $C(1 + V(q)^4)$.

For the proof of our main result, we shall need some estimates on the solutions to the SDE, which are collected from [BRVE09] (with the notations modified to fit with those we use) :

Proposition 9.6 (Properties of the SDE). *For $s \leq t$, let $Q_{t,s}(x)$ denote the evolution operator of (9.1). (i) For any integer $\ell \geq 1$, there exists $K_\ell > 0$ such that, for any $t \geq 0$ and any q ,*

$$\mathbb{E}[|Q_{t,0}(x) - x|^{2\ell}] \leq K_\ell(1 + t^\ell V(q)^{2\ell})t^\ell;$$

(ii) *For all $K > 0$, there exists $\Delta t^* > 0$ such that, for any $q, q' \in \mathbb{R}^d$, $s \geq 0$ and $\Delta t < \Delta t^*$, we have*

$$\mathbb{E}[|Q_{s+h,s}(q) - Q_{s+h,s}(q')|^2] \leq |q - q'|^2(1 + Kh);$$

and, with $\Delta := Q_{s+h,s}(q) - Q_{s+h,s}(q') - (q - q')$,

$$\mathbb{E}[|\Delta|^2] \leq Kh^2(1 + V(q)^2 + V(q')^2)|q - q'|.$$

Lemma 9.7 (Single step accuracy of the modified EM scheme). *There exists $\Delta t^* > 0$ such that, for any $\Delta t < \Delta t^*$,*

$$\mathbb{E}^x[|\tilde{q} - q_{\Delta t}|^2] \leq C(1 + V(q)^4)\Delta t^3$$

where

$$\tilde{q} = q - \Delta t \nabla V(q) + \Delta t^2 \left(\frac{1}{6} \nabla \Delta V(q) - \frac{1}{6} \text{Hess } V(q) \nabla V(q) \right) + \sqrt{2\Delta t} \left(\text{Id} + \frac{\Delta t}{3} \text{Hess } V(x) \right)^{-1/2} B_{\Delta t}$$

and $q_{\Delta t}$ is the value at time h of the solution to (9.1) with initial condition q . Moreover, with the same notations, we have

$$|\mathbb{E}^q[\tilde{q} - q_{\Delta t}]| \leq C(1 + V(q)^2)\Delta t^2.$$

Démonstration. Let $\hat{q} := q - \Delta t \nabla V(q) + \sqrt{2\Delta t} B_{\Delta t}$. We have

$$\begin{aligned} \mathbb{E}^q[|\tilde{q} - q_{\Delta t}|^2] &\leq 2\mathbb{E}^x[|\hat{q} - q_{\Delta t}|^2] + 8\Delta t \mathbb{E}[|(Id + \frac{h}{3} \text{Hess } V(x))^{-1/2} - Id| B_{\Delta t}|^2] \\ &\quad + 4\Delta t^4 \mathbb{E}[|\frac{1}{6} \nabla \Delta V(q) - \frac{1}{6} \text{Hess } V(q) \nabla V(q)|^2] \end{aligned}$$

It has been proven in Lemma 4.3 of [BRVE09] that

$$\mathbb{E}^q[|\hat{q} - q_{\Delta t}|^2] \leq C(1 + V(q)^4)\Delta t^3.$$

We also have

$$\begin{aligned} \mathbb{E}[|(Id + \frac{h}{3} \text{Hess } V(q))^{-1/2} - Id| B_{\Delta t}|^2] &= \text{Tr}(((Id + \frac{h}{3} \text{Hess } V(q))^{-1/2} - Id)^2) \\ &= \text{Tr}((Id + \frac{\Delta t}{3} \text{Hess } V(q))^{-1} + \text{Tr}(Id) - 2 \text{Tr}((Id + \frac{\Delta t}{3} \text{Hess } V(q))^{-1/2})) \\ &\leq C\Delta t^2(1 + V(q)^2) \end{aligned}$$

where we have used the assumed bounds on the derivatives of V . Finally,

$$\Delta t^4 \mathbb{E}[|\frac{1}{6} \nabla \Delta V(q) - \frac{1}{6} \text{Hess } V(q) \nabla V(q)|^2] \leq C\Delta t^4(1 + V(q)^4)$$

This concludes the proof of the first part.

For the second part, we just have to notice that

$$\mathbb{E}^q[\tilde{q} - q_{\Delta t}] = \mathbb{E}^q[\bar{q} - q_{\Delta t}] + \Delta t^2 (\frac{1}{6} \nabla \Delta V(q) - \frac{1}{6} \text{Hess } V(q) \nabla V(q)),$$

and use the accuracy estimate for the local mean deviation of the regular EM scheme in [BV, Lemma 4.3], as well as the bound on the derivatives of V . \square

Lemma 9.8 (Local accuracy of MALA). *Let \bar{q} be the random variable obtained from a single step of the modified MALA scheme starting from q . Then, for $\Delta t < \Delta t^*$, we have*

$$\begin{aligned} (i) \quad &\mathbb{E}[|\bar{q} - q_{\Delta t}|^2] \leq C\Delta t^3(1 + V(q)^6); \\ (ii) \quad &|\mathbb{E}[\bar{q} - q_{\Delta t}]| \leq C(1 + V(q)^6)\Delta t^2 \end{aligned}$$

Démonstration. By definition of the scheme, we have

$$\mathbb{E}^q[|\bar{q} - q_{\Delta t}|^2] = \mathbb{E}^q[|\tilde{q} - q_{\Delta t}|^2(\alpha(q, \tilde{q}) \wedge 1)] + \mathbb{E}^q[|q - Y_{\Delta t}|^2(1 - (\alpha(q, \tilde{q}) \wedge 1))]$$

For the first term on the right-hand side, we just have to apply Lemma 9.7. For the second term, applying the Cauchy-Schwartz inequality, we get

$$\mathbb{E}^q[|q - q_{\Delta t}|^2(1 - (R_{\Delta t}(q, \tilde{q}) \wedge 1))] \leq \mathbb{E}^q[|q - q_{\Delta t}|^4]^{1/2} \mathbb{E}^q[(1 - R_{\Delta t}(q, \tilde{q}))^2]^{1/2} \quad (9.14)$$

From part (i) of Proposition 9.6, we know that $\mathbb{E}^q[|q - q_{\Delta t}|^4] \leq C(1 + \Delta t^2 V(q)^4)\Delta t^2$. Applying Lemma 9.5 to bound $\mathbb{E}^q[(1 - R_{\Delta t}(q, \tilde{q}))^2]$ then concludes the proof.

The proof of (ii) is based on the same arguments. We have

$$\begin{aligned} |\mathbb{E}[\bar{q} - q_{\Delta t}]| &\leq |\mathbb{E}[(\tilde{q} - q_{\Delta t})(\alpha(q, \tilde{q}) \wedge 1)]| + |\mathbb{E}^q[(q - q_{\Delta t})(1 - (\alpha(q, \tilde{q}) \wedge 1))]| \\ &\leq |\mathbb{E}[(\tilde{q} - q_{\Delta t})]| \\ &\quad + \mathbb{E}^q[|(q - q_{\Delta t})|^2]^{1/2} \mathbb{E}^q[(1 - R_{\Delta t}(q, \tilde{q}))^2]^{1/2} \\ &\leq C(1 + V(q)^6)\Delta t^2 \end{aligned} \quad (9.15)$$

where we have used Proposition 9.6 and Lemma 9.7. \square

Proof of Theorem 9.2. For $k = 0, \dots, \lfloor T/h \rfloor$ let $t_k := k\Delta t$ and

$$\epsilon_k := \mathbb{E}^{\bar{\mu}} \mathbb{E}^q [|q^k - q_{t_k}|^2].$$

We shall prove the inequality

$$\epsilon_{k+1} \leq (1 + A\Delta t)\epsilon_k + K\Delta t^3 \quad (9.16)$$

which implies Theorem 9.2 by a simple iteration.

Let \mathcal{F}_k be the sigma-algebra of events up to the k -th iteration. We have

$$\begin{aligned} |q^{k+1} - q_{t_{k+1}}|^2 &= |q^{k+1} - Q_{t_{k+1}, t_k}(q^k) + Q_{t_{k+1}, t_k}(q^k) - Q_{t_{k+1}, t_k}(Q_{t_k, 0}(q))|^2 \\ &= |q^{k+1} - Q_{t_{k+1}, t_k}(q^k)|^2 + |Q_{t_{k+1}, t_k}(q^k) - Q_{t_{k+1}, t_k}(Q_{t_k, 0}(q))|^2 \\ &\quad + 2\langle q^{k+1} - Q_{t_{k+1}, t_k}(q^k), Q_{t_{k+1}, t_k}(q^k) - Q_{t_{k+1}, t_k}(Q_{t_k, 0}(q)) \rangle. \end{aligned} \quad (9.17)$$

From Lemma 9.8, we know that

$$\mathbb{E}[|q^{k+1} - Q_{t_{k+1}, t_k}(q^k)|^2 | \mathcal{F}_k] \leq C\Delta t^3(1 + V(q^k)^4)$$

and therefore, since the scheme leaves the measure μ invariant,

$$\mathbb{E}^{\mu} \mathbb{E}^q [|q^{k+1} - Q_{t_{k+1}, t_k}(q^k)|^2] \leq C\Delta t^3. \quad (9.18)$$

Similarly, using part (ii) of Proposition 9.6, we get

$$\mathbb{E}^{\mu} \mathbb{E}^q [|Q_{t_{k+1}, t_k}(q^k) - Y_{t_{k+1}, t_k}(Y_{t_k, 0}(q))|^2] \leq (1 + A\Delta t)\epsilon_k. \quad (9.19)$$

All that is left is to bound the third term on the right-hand side in (9.17). Setting

$$\Delta := Q_{t_{k+1}, t_k}(q^k) - Q_{t_{k+1}, t_k}(Q_{t_k, 0}(q)) - (q^k - Q_{t_k, 0}(q))$$

we can rewrite it as

$$\langle q^{k+1} - Q_{t_{k+1}, t_k}(q^k), q^k - Q_{t_k, 0}(q) \rangle + \langle q^{k+1} - Q_{t_{k+1}, t_k}(q^k), \Delta \rangle.$$

Using the Cauchy-Schwartz inequality and Lemma 9.8, we have

$$\begin{aligned} &\mathbb{E}^{\mu} \mathbb{E}^q \left[\langle \mathbb{E}[q^{k+1} - Q_{t_{k+1}, t_k}(q^k) | \mathcal{F}_k], q^k - Q_{t_k, 0}(q) \rangle \right] \\ &\leq \mathbb{E}^{\mu} \mathbb{E}^q \left[|\mathbb{E}[q^{k+1} - Q_{t_{k+1}, t_k}(q^k) | \mathcal{F}_k]|^2 \right]^{1/2} \epsilon_k^{1/2} \\ &\leq C\Delta t^2 \epsilon_k^{1/2} \\ &\leq C(\Delta t \epsilon_k + \Delta t^3). \end{aligned} \quad (9.20)$$

In the same way,

$$\begin{aligned} &\mathbb{E}^{\bar{\mu}} \mathbb{E}^q [\langle q^{k+1} - Q_{t_{k+1}, t_k}(q^k), \Delta \rangle] \\ &\leq \mathbb{E}^{\mu} \mathbb{E}^q \left[\mathbb{E}[|q^{k+1} - Q_{t_{k+1}, t_k}(q^k)|^2 | \mathcal{F}_k] \right]^{1/2} \mathbb{E}^{\mu} \mathbb{E}^q \left[\mathbb{E}[|\Delta|^2 | \mathcal{F}_k] \right]^{1/2} \\ &\leq C\Delta t^{3/2} \mathbb{E}^q \left[\mathbb{E}[|\Delta|^2 | \mathcal{F}_k] \right]^{1/2} \end{aligned}$$

According to Proposition 9.6,

$$\mathbb{E}^q \left[\mathbb{E}[|\Delta|^2 | \mathcal{F}_k] \right]^{1/2} \leq C\Delta t \epsilon_k^{1/4},$$

and therefore

$$\begin{aligned} \mathbb{E}^{\bar{\mu}} \mathbb{E}^q [\langle q^{k+1} - Q_{t_{k+1}, t_k}(q^k), \Delta \rangle] &\leq C\Delta t^{5/2} \epsilon_k^{1/4} \\ &\leq C(\Delta t \epsilon_k + \Delta t^3). \end{aligned}$$

This concludes the proof of (9.16). □

Bibliographie

- [ABC⁺00] Cécile ANÉ, Sébastien BLACHÈRE, Djalil CHAFAÏ, Pierre FOUGÈRES, Ivan GENTIL, Florent MALRIEU, Cyril ROBERTO et Grégory SCHEFFER : *Sur les inégalités de Sobolev logarithmiques*, volume 10 de *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.
- [ADPZ11] Stefan ADAMS, Nicolas DIRR, Mark A. PELETIER et Johannes ZIMMER : From a large-deviations principle to the Wasserstein gradient flow : a new micro-macro passage. *Comm. Math. Phys.*, 307(3):791–815, 2011.
- [ADPZ13] Stefan ADAMS, Nicolas DIRR, Mark PELETIER et Johannes ZIMMER : Large deviations and gradient flows. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 371(2005):20120341, 17, 2013.
- [AGS08] Luigi AMBROSIO, Nicola GIGLI et Giuseppe SAVARÉ : *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second édition, 2008.
- [AGS14] Luigi AMBROSIO, Nicola GIGLI et Giuseppe SAVARÉ : Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.
- [AJKH14] A. ALFONSI, B. JOURDAIN et A. KOHATSU-HIGA : Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme. *Ann. Appl. Probab.*, 24(3):1049–1080, 2014.
- [AMP⁺12] Steffen ARNRICH, Alexander MIELKE, Mark A. PELETIER, Giuseppe SAVARÉ et Marco VENERONI : Passing to the limit in a Wasserstein gradient flow : from diffusion to reaction. *Calc. Var. Partial Differential Equations*, 44(3-4):419–454, 2012.
- [ASZ09] Luigi AMBROSIO, Giuseppe SAVARÉ et Lorenzo ZAMBOTTI : Existence and stability for Fokker-Planck equations with log-concave reference measure. *Probab. Theory Related Fields*, 145(3-4):517–564, 2009.
- [BB99] Jean-David BENAMOU et Yann BRENIER : A numerical method for the optimal time-continuous mass transport problem and related problems. In *Monge Ampère equation : applications to geometry and optimization (Deerfield Beach, FL, 1997)*, volume 226 de *Contemp. Math.*, pages 1–11. Amer. Math. Soc., Providence, RI, 1999.
- [BB06] G. BLOWER et F. BOLLEY : Concentration inequalities on product spaces with applications to Markov processes. *Studia Mathematica*, 175-1:47–72, 2006.
- [BB13] Franck BARTHE et Charles BORDENAVE : Combinatorial optimization over two random point sets. In *Séminaire de Probabilités XLV*, volume 2078 de *Lecture Notes in Math.*, pages 483–535. Springer, Cham, 2013.

- [BBCG08] Dominique BAKRY, Franck BARTHE, Patrick CATTIAUX et Arnaud GUILLIN : A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13:60–66, 2008.
- [BD98] Andrea BRAIDES et Anneliese DEFRANCESCHI : *Homogenization of multiple integrals*, volume 12 de *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1998.
- [BDSG⁺01] L. BERTINI, A. DE SOLE, D. GABRIELLI, G. JONA-LASINIO et C. LANDIM : Fluctuations in stationary non equilibrium states of irreversible processes. *Phys. Rev. Lett.*, 87, 2001.
- [BDSG⁺02] L. BERTINI, A. DE SOLE, D. GABRIELLI, G. JONA-LASINIO et C. LANDIM : Macroscopic fluctuation theory for stationary non equilibrium states. *J. Stat. Phys.*, 107, 2002.
- [BG99] S. G. BOBKOV et F. GÖTZE : Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.
- [BGL01] Sergey G. BOBKOV, Ivan GENTIL et Michel LEDOUX : Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 80(7):669–696, 2001.
- [BGSR] Thierry BODINEAU, Isabelle GALLAGHER et Laure SAINT-RAYMOND : The brownian motion as the limit of a deterministic system of hard-spheres.
- [BGV07] François BOLLEY, Arnaud GUILLIN et Cédric VILLANI : Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields*, 137(3-4):541–593, 2007.
- [BH99] T. BODINEAU et B. HELFFER : The log-Sobolev inequality for unbounded spin systems. *J. Funct. Anal.*, 166(1):168–178, 1999.
- [BH00] T. BODINEAU et B. HELFFER : Correlations, spectral gap and log-Sobolev inequalities for unbounded spins systems. In *Differential equations and mathematical physics (Birmingham, AL, 1999)*, volume 16 de *AMS/IP Stud. Adv. Math.*, pages 51–66. Amer. Math. Soc., Providence, RI, 2000.
- [Bil99] Patrick BILLINGSLEY : *Convergence of probability measures*. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley & Sons, Inc., New York, second édition, 1999. A Wiley-Interscience Publication.
- [BL00] S. G. BOBKOV et M. LEDOUX : From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.*, 10(5):1028–1052, 2000.
- [BLM09] Lorenzo BERTINI, Claudio LANDIM et Mustapha MOURRAGUI : Dynamical large deviations for the boundary driven weakly asymmetric exclusion process. *Ann. Probab.*, 37(6):2357–2403, 2009.
- [BLP11] A. BENSOUSSAN, J.-L. LIONS et G. PAPANICOLAOU : *Asymptotic analysis for periodic structures*. AMS Chelsea Publishing, 2011.
- [BLRB00] F. BONETTO, J. L. LEBOWITZ et L. REY-BELLET : Fourier’s law : a challenge to theorists. In *Mathematical physics 2000*, pages 128–150. Imp. Coll. Press, London, 2000.
- [BM13] Franck BARTHE et Emanuel MILMAN : Transference principles for log-Sobolev and spectral-gap with applications to conservative spin systems. *Comm. Math. Phys.*, 323(2):575–625, 2013.

- [BO05] Cédric BERNARDIN et Stefano OLLA : Fourier's law for a microscopic model of heat conduction. *J. Stat. Phys.*, 121(3-4):271–289, 2005.
- [Boi11] E. BOISSARD : Simple bounds for the convergence of empirical and occupation measures in 1-Wasserstein distance. *Electronic Journal of Probability*, 16, 2011.
- [Bre91] Yann BRENIER : Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.
- [BRH13] N. BOU-RABEE et M. HAIRER : Nonasymptotic mixing of the MALA algorithm. *IMA J. Numer. Anal.*, 33:80–110, 2013.
- [BRVE09] N. BOU-RABEE et E. VANDEN-EIJNDEN : Pathwise accuracy and ergodicity of metropolized integrators for SDEs. *Commun. Pure Appl. Math.*, 63(5):655–696, 2009.
- [BT96] V. BALLY et D. TALAY : The law of the Euler scheme for stochastic differential equations, II. Convergence rate of the density. *Monte-Carlo methods and Appl.*, 2:93–128, 1996.
- [BV05] F. BOLLEY et C. VILLANI : Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. *Annales de la Faculté des Sciences de Toulouse (6)*, 14(3):331–352, 2005.
- [BZ05] S. G. BOBKOV et B. ZEGARLINSKI : Entropy bounds and isoperimetry. *Mem. Amer. Math. Soc.*, 176(829):x+69, 2005.
- [Cap03] Pietro CAPUTO : Uniform Poincaré inequalities for unbounded conservative spin systems : the non-interacting case. *Stochastic Process. Appl.*, 106(2):223–244, 2003.
- [CCEL13] E. CARLEN, D. CORDERO-ERAUSQUIN et E. LIEB : Asymmetric covariance estimates of brascamp-lieb type and related inequalities for log-concave measures. *Ann. Inst. H. Poincaré Probab. Statist.*, 49:1–12, 2013.
- [CG06] Patrick CATTIAUX et Arnaud GUILLIN : On quadratic transportation cost inequalities. *J. Math. Pures Appl. (9)*, 86(4):341–361, 2006.
- [Cha03] D. CHAFAÏ : Glauber versus Kawasaki for spectral gap and logarithmic Sobolev inequalities of some unbounded conservative spin systems. *Markov Process. Related Fields*, 9(3):341–362, 2003.
- [Com91] Francis COMETS : Limites hydrodynamiques. *Astérisque*, (201-203):Exp. No. 735, 167–192 (1992), 1991. Séminaire Bourbaki, Vol. 1990/91.
- [CY92] Chih Chung CHANG et Horng-Tzer YAU : Fluctuations of one-dimensional Ginzburg-Landau models in nonequilibrium. *Comm. Math. Phys.*, 145(2):209–234, 1992.
- [Des] Laurent DESVILLETES : Progrès récents concernant le programme de kac en théorie cinétique (d'après Stéphane Mischler et Clément Mouhot). *Astérisque*, page Exp. No. 1076. Séminaire Bourbaki, Vol. 2013/2014.
- [DF14] Manh Hong DUONG et Max FATHI : The two-scale approach for non-reversible dynamics. 2014.
- [DG87] Donald A. DAWSON et Jürgen GÄRTNER : Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

- [DGMT80] Ennio DE GIORGI, Antonio MARINO et Mario TOSQUES : Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3):180–187, 1980.
- [DGW04] H. DJELLOUT, A. GUILLIN et L. WU : Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Prob.*, 32:no. 3B, 2702–2732, 2004.
- [Diz07] Deniz DIZDAR : Towards an optimal rate of convergence in the hydrodynamic limit for kawasaki dynamics. Mémoire de D.E.A., Universitat Bonn, 2007.
- [DLR13] Manh Hong DUONG, Vaïos LASCHOS et Michiel RENGER : Wasserstein gradient flows from large deviations of many-particle limits. *ESAIM Control Optim. Calc. Var.*, 19(4):1166–1188, 2013.
- [DPZ13] Manh Hong DUONG, Mark A. PELETIER et Johannes ZIMMER : GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to large-deviation principles. *Nonlinearity*, 26(11):2951–2971, 2013.
- [DSC98] P. DIACONIS et L. SALOFF-COSTE : What do we know about the Metropolis algorithm? *J. Comput. System Sci.*, 57(1):20–36, 1998. 27th Annual ACM Symposium on the Theory of Computing (STOC'95) (Las Vegas, NV).
- [Dud02] R. M. DUDLEY : *Real analysis and probability*, volume 74 de *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.
- [DV89] M. D. DONSKER et S. R. S. VARADHAN : Large deviations from a hydrodynamic scaling limit. *Comm. Pure Appl. Math.*, 42(3):243–270, 1989.
- [DZ10] Amir DEMBO et Ofer ZEITOUNI : *Large deviations techniques and applications*, volume 38 de *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [Fat13a] Max FATHI : Modified logarithmic sobolev inequalities for canonical ensembles. *preprint*, 2013.
- [Fat13b] Max FATHI : A two-scale approach to the hydrodynamic limit part II : local Gibbs behavior. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(2):625–651, 2013.
- [Fat14] Max FATHI : A gradient flow approach to large deviations for diffusion processes. *preprint*, 2014.
- [FF13] Max FATHI et Noufel FRIKHA : Transport-entropy inequalities and deviation estimates for stochastic approximations schemes. *Electron. J. Probab.*, 18:no. 67, 36, 2013.
- [FHS] Max FATHI, Ahmed-Amine HOMMANN et Gabriel STOLTZ : Error analysis of the transport properties of metropolized schemes. *preprint*.
- [FK06] Jin FENG et Thomas G. KURTZ : *Large deviations for stochastic processes*, volume 131 de *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.
- [FM12] N. FRIKHA et S. MENOZZI : Concentration bounds for stochastic approximations. *Electron. Commun. Probab.*, 17:no. 47, 1–15, 2012.
- [FM14] Max FATHI et Georg MENZ : Hydrodynamic limit for conservatic spin systems with super-quadratic single-site potential. *preprint*, 2014.
- [Föl88] Hans FÖLLMER : Random fields and diffusion processes. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 de *Lecture Notes in Math.*, pages 101–203. Springer, Berlin, 1988.

- [Fri89] J. FRITZ : Hydrodynamics in a symmetric random medium. *Comm. Math. Phys.*, 125(1):13–25, 1989.
- [Fun05] Tadahisa FUNAKI : Stochastic interface models. In *École d'Été de Probabilités de Saint-Flour 2003*, volume 1869 de *Lecture Notes in Math.*, pages 103–274. Springer, Berlin, 2005.
- [Fun12] Tadahisa FUNAKI : Hydrodynamic limit for the $\nabla\varphi$ interface model via two-scale approach. *Probability in Complex Physical Systems : In Honour of Erwin Bolthausen and Jürgen Gärtner*, pages 463–490, 2012.
- [GL07] N. GOZLAN et C. LÉONARD : A large deviation approach to some transportation cost inequalities. *Prob. Th. Rel. Fields*, 139(1-2):235–283, 2007.
- [GL10] N. GOZLAN et C. LÉONARD : Transport inequalities. A survey. *Markov Process. Related Fields*, 16(4):635–736, 2010.
- [GL13] Nicola GIGLI et Michel LEDOUX : From log Sobolev to Talagrand : a quick proof. *Discrete Contin. Dyn. Syst.*, 33(5):1927–1935, 2013.
- [GOVW09] Natalie GRUNEWALD, Felix OTTO, Cédric VILLANI et Maria G. WESTDIKCKENBERG : A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(2):302–351, 2009.
- [Goz09] Nathael GOZLAN : A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.*, 37(6):2480–2498, 2009.
- [GPV88] M. Z. GUO, G. C. PAPANICOLAOU et S. R. S. VARADHAN : Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.*, 118(1):31–59, 1988.
- [Gro75] Leonard GROSS : Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.
- [Gro80] M. GROMOV : Paul Lévy's isoperimetric inequality. *Prépublication I.H.E.S.*, 1980.
- [Gro07] Misha GROMOV : *Metric structures for Riemannian and non-Riemannian spaces*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, english édition, 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.
- [GRS75] F. GUERRA, L. ROSEN et B. SIMON : The $P(\phi)_2$ Euclidean quantum field theory as classical statistical mechanics. I, II. *Ann. of Math.*, 1975.
- [GRS11a] Nathael GOZLAN, Cyril ROBERTO et Paul-Marie SAMSON : From concentration to logarithmic Sobolev and Poincaré inequalities. *J. Funct. Anal.*, 260(5):1491–1522, 2011.
- [GRS11b] Nathael GOZLAN, Cyril ROBERTO et Paul-Marie SAMSON : A new characterization of Talagrand's transport-entropy inequalities and applications. *Ann. Probab.*, 39(3):857–880, 2011.
- [GRS13] N. GOZLAN, C. ROBERTO et P.-M. SAMSON : Characterization of Talagrand's transport-entropy inequalities in metric spaces. *Ann. Probab.*, 41(5):3112–3139, 2013.
- [GRS14] Nathael GOZLAN, Cyril ROBERTO et Paul-Marie SAMSON : Hamilton Jacobi equations on metric spaces and transport entropy inequalities. *Rev. Mat. Iberoam.*, 30(1):133–163, 2014.

- [Has70] W. K. HASTINGS : Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.
- [HB98] D. M. HEYES et A. C. BRAŃKA : Monte Carlo as Brownian dynamics. *Mol. Phys.*, 94(3):447–454, 1998.
- [Hil02] David HILBERT : Mathematical problems. *Bull. Amer. Math. Soc.*, 8(10):437–479, 1902.
- [HM11] M. HAIRER et J. C. MATTINGLY : Yet another look at Harris’ ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 de *Progr. Probab.*, pages 109–117. Birkhäuser/Springer, 2011.
- [JBTK99] M. JARDAT, O. BERNARD, P. TURQ et G. R. KNELLER : Transport coefficients of electrolyte solutions from Smart Brownian dynamics simulations. *J. Chem. Phys.*, 110(16):7993–7999, 1999.
- [JKO98] Richard JORDAN, David KINDERLEHRER et Felix OTTO : The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [JSZ85] W. B. JOHNSON, G. SCHECHTMAN et J. ZINN : Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probab.*, 13(1):234–253, 1985.
- [KL99] Claude KIPNIS et Claudio LANDIM : *Scaling limits of interacting particle systems*, volume 320 de *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.
- [Kos01a] E. KOSYGINA : The behaviour of the specific entropy in the hydrodynamic scaling limit for Ginzburg-Landau model. *Markov Process. Related Fields*, 7(3):383–417, 2001. MR1856489.
- [Kos01b] Elena KOSYGINA : The behavior of the specific entropy in the hydrodynamic scaling limit. *Ann. Probab.*, 29(3):1086–1110, 2001.
- [KPKDT05] M. A. KATSOUKAKIS, P. PLECHÁČ et D. K. D. K. TSAGKAROGIANNIS : Mesoscopic modeling for continuous spin lattice systems : model problems and micromagnetics applications. *J. Stat. Phys.*, (119):347–389, 2005.
- [KSS91] A. KRÁMLI, N. SIMANYI et D. SZÁSZ : The K-property of three billiard balls. *Annals of Mathematics*, 133:32–72, 1991.
- [KV86] C. KIPNIS et S. R. S. VARADHAN : Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.
- [Led] Michel LEDOUX : Inégalités isopérimétriques en analyse et probabilités. *Astérisque*, page Exp. No. 773. Séminaire Bourbaki, Vol. 1992/1993.
- [Led99] Michel LEDOUX : Logarithmic Sobolev inequalities for spin systems revisited. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.36.4917>, 1999.
- [Led01] Michel LEDOUX : *The concentration of measure phenomenon*, volume 89 de *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.
- [Lel09] Tony LELIÈVRE : A general two-scale criteria for logarithmic Sobolev inequalities. *J. Funct. Anal.*, 256(7):2211–2221, 2009.

- [Léo12] Christian LÉONARD : Girsanov theory under a finite entropy condition. *In Séminaire de Probabilités XLIV*, volume 2046 de *Lecture Notes in Math.*, pages 429–465. Springer, Heidelberg, 2012.
- [Lis09] Stefano LISINI : Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, 15(3):712–740, 2009.
- [LM10] V. LEMAIRE et S. MENOZZI : On some non asymptotic bounds for the Euler scheme. *Electronic Journal of Probability*, 15:1645–1681, 2010.
- [LMS13] B. LEIMKUHNER, Ch. MATTHEWS et G. STOLTZ : The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *arXiv preprint*, 1308.5814, 2013.
- [LNN05] C. LANDIM et J. NORONHA NETO : Poincaré and logarithmic Sobolev inequality for Ginzburg-Landau processes in random environment. *Probab. Theory Related Fields*, 131(2):229–260, 2005.
- [LO12] Carlangelo LIVERANI et Stefano OLLA : Toward the Fourier law for a weakly interacting anharmonic crystal. *J. Amer. Math. Soc.*, 25(2):555–583, 2012.
- [LP76] J. L. LEBOWITZ et E. PRESUTTI : Statistical mechanics of systems of unbounded spins. *Comm. Math. Phys.*, 50:195–218, 1976.
- [LPY02] C. LANDIM, G. PANIZO et H. T. YAU : Spectral gap and logarithmic Sobolev inequality for unbounded conservative spin systems. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(5):739–777, 2002.
- [LRS10] T. LELIÈVRE, M. ROUSSET et G. STOLTZ : *Free-energy computations : a mathematical perspective*. Imperial College Press, 2010.
- [LSU68] O. A. LADYŽENSKAJA, V. A. SOLONNIKOV et N. N. URALCEVA : *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968. MR0241822.
- [LY93] Sheng Lin LU et Horng-Tzer YAU : Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics. *Comm. Math. Phys.*, 156(2):399–433, 1993.
- [Maa11] Jan MAAS : Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.
- [Mar] Mauro MARIANI : A gamma-convergence approach to large deviations. <http://arxiv.org/abs/1204.0640>.
- [Mar96] K. MARTON : A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3):556–571, 1996.
- [Mar12] D. MARAHRENS : *On some nonlinear partial differential equations for classical and quantum many body problems*. Thèse de doctorat, University of Cambridge, 2012.
- [Mas07] Pascal MASSART : *Concentration inequalities and model selection*. Lecture Notes in Math. 1896. Springer, 2007.
- [McC97] Robert J. MCCANN : A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
- [McC01] Robert J. MCCANN : Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

- [Men11] Georg MENZ : LSI for Kawasaki dynamics with weak interaction. *Comm. Math. Phys.*, 307(3):817–860, 2011.
- [Men13] Georg MENZ : The approach of Otto-Reznikoff revisited. *preprint*, 2013.
- [MGR06] T. MULLER-GRONBACH et K. RITTER : Minimal errors for strong and weak approximation of stochastic differential equations. In Berlin SPRINGER-VERLAG, éditeur : *Monte Carlo and Quasi-Monte Carlo Methods (A. Keller, S. Heinrich, H. Niederreiter, eds.)*, pages 53–82. 2006.
- [Mil71] V. D. MILMAN : New proof of the theorem of Dvoretzky on sections of convex bodies. *Funct. Anal. Appl.*, (5):28–37, 1971.
- [Mil10] Emanuel MILMAN : Isoperimetric and concentration inequalities : equivalence under curvature lower bound. *Duke Math. J.*, 154(2):207–239, 2010.
- [MM13] Stéphane MISCHLER et Clément MOUHOT : Kac’s program in kinetic theory. *Invent. Math.*, 193(1):1–147, 2013.
- [MN13] Georg MENZ et Robin NITTKA : One-phase region of one-dimensional lattice systems of continuous spins and long-range interaction. *preprint*, 2013.
- [MO13] G. MENZ et F. OTTO : Uniform logarithmic sobolev inequalities for conservative spin systems with super-quadratic single-site potential. *Ann. Probab.*, 41(3B):2182–2224, 2013. DOI: 10.1214/11-AOP715.
- [Mon84] G. MONGE : *Mémoire sur la théorie des déblais et des remblais*. 1784.
- [Mor55] Charles B. MORREY, Jr. : On the derivation of the equations of hydrodynamics from statistical mechanics. *Comm. Pure Appl. Math.*, 8:279–326, 1955.
- [MRR⁺53] N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER et E. TELLER : Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1091, 1953.
- [MS12] G. MENZ et A. SCHLICHTING : Poincaré $\frac{1}{2}$ and logarithmic Sobolev inequalities by decomposition of the energy landscape. *to appear in Ann. Probab.*, 2012.
- [MSH02] J. C. MATTINGLY, A. M. STUART et D. J. HIGHAM : Ergodicity for SDEs and approximations : locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, 101(2):185–232, 2002.
- [MT06] F. MALRIEU et D. TALAY : Concentration inequalities for Euler schemes. In *Monte Carlo and quasi-Monte Carlo methods 2004*, pages 355–371. Springer, Berlin, 2006.
- [MT09] S. P. MEYN et R. L. TWEEDIE : *Markov chains and stochastic stability (2nd edition)*. Cambridge University Press, 2009.
- [Oll94] S. OLLA : Homogenization of diffusion processes in random fields, 1994. Lecture notes from Ecole polytechnique, www.ceremade.dauphine.fr/~olla/lho.ps.
- [OR07] Felix OTTO et Maria G. REZNIKOFF : A new criterion for the logarithmic Sobolev inequality and two applications. *J. Funct. Anal.*, 243(1):121–157, 2007.
- [Ott01] Felix OTTO : The geometry of dissipative evolution equations : the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

- [OV00] F. OTTO et C. VILLANI : Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.
- [OVY93] S. OLLA, S. R. S. VARADHAN et H.-T. YAU : Hydrodynamical limit for a Hamiltonian system with weak noise. *Comm. Math. Phys.*, 155(3):523–560, 1993.
- [PSV10] Mark A. PELETIER, Giuseppe SAVARÉ et Marco VENERONI : From diffusion to reaction via Γ -convergence. *SIAM J. Math. Anal.*, 42(4):1805–1825, 2010.
- [QRV99] J. QUASTEL, F. REZAKHANLOU et S. R. S. VARADHAN : Large deviations for the symmetric simple exclusion process in dimensions $d \geq 3$. *Probab. Theory Related Fields*, 113(1):1–84, 1999.
- [Qua95] Jeremy QUASTEL : Large deviations from a hydrodynamic scaling limit for a nongradient system. *Ann. Probab.*, 23(2):724–742, 1995.
- [QY98] J. QUASTEL et H.-T. YAU : Lattice gases, large deviations, and the incompressible Navier-Stokes equations. *Ann. of Math. (2)*, 148(1):51–108, 1998.
- [RDF78] P. J. ROSSKY, J. D. DOLL et H. L. FRIEDMAN : Brownian dynamics as smart Monte Carlo simulation. *J. Chem. Phys.*, 69(10):4628–4633, 1978.
- [Rid] S. RIDEAU : De la preuve par contraposition et ses applications. *personal communication*.
- [RR98] S. T. RACHEV et L. RÜSCHENDORF : *Mass transportation problems. Vol. II. Probability and its Applications* (New York). Springer-Verlag, New York, 1998. Applications.
- [RT96] G. O. ROBERTS et R. L. TWEEDIE : Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996.
- [Ser11] Sylvia SERFATY : Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.*, 31(4):1427–1451, 2011.
- [Sim13] N. SIMANYI : Singularities and nonhyperbolic manifolds do not coincide. *Nonlinearity*, 26:1703–1717, 2013.
- [Sin70] Y. G. SINAI : Dynamical systems with elastic reflections. *Russ. Math. Survey*, 25:137–189, 1970.
- [SR09] Laure SAINT-RAYMOND : *Hydrodynamic limits of the Boltzmann equation*, volume 1971 de *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.
- [SS04] Etienne SANDIER et Sylvia SERFATY : Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.
- [Tal96] M. TALAGRAND : Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.
- [TT90] Denis TALAY et Luciano TUBARO : Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.
- [Var93] S. R. S. VARADHAN : Nonlinear diffusion limit for a system with nearest neighbor interactions. II. In *Asymptotic problems in probability theory : stochastic models and diffusions on fractals (Sanda/Kyoto, 1990)*, volume 283

- de *Pitman Res. Notes Math. Ser.*, pages 75–128. Longman Sci. Tech., Harlow, 1993.
- [Var08] S. R. S. VARADHAN : Large deviations. *Ann. Probab.*, 36(2):397–419, 2008.
- [Vil02] Cédric VILLANI : Limites hydrodynamiques de l'équation de Boltzmann (d'après C. Bardos, F. Golse, C. D. Levermore, P.-L. Lions, N. Masmoudi, L. Saint-Raymond). *Astérisque*, (282):Exp. No. 893, ix, 365–405, 2002. Séminaire Bourbaki, Vol. 2000/2001.
- [Vil03] Cédric VILLANI : *Topics in optimal transportation*, volume 58 de *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [Vil09] Cédric VILLANI : *Optimal transport*, volume 338 de *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [Yau91] Horng-Tzer YAU : Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.*, 22(1):63–80, 1991.
- [Yos99] Nobuo YOSHIDA : The log-Sobolev inequality for weakly coupled lattice fields. *Probab. Theory Related Fields*, 115(1):1–40, 1999.
- [Zeg96] Boguslaw ZEGARLINSKI : The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice. *Comm. Math. Phys.*, 175(2):401–432, 1996.