Theoretical and numerical analysis of non-reversible dynamics in computational statistical physics+

Julien Roussel

UNIVERSITÉ PARIS-EST

ÉCOLE DOCTORALE MSTIC
MATHÉMATIQUES, SCIENCES ET TECHNOLOGIES
DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE

pour obtenir le titre de

Docteur de l'Université Paris-Est
ès Mathématiques Appliquées

Présentée et soutenue par

Julien ROUSSEL

Analyse théorique et numérique de dynamiques non-réversibles en physique statistique computationnelle

Thèse dirigée par Gabriel Stoltz

Préparée au sein de l'équipe MODÉLISATION, ANALYSE ET SIMULATION
du laboratoire CERMICS de l'ÉCOLE DES PONTS PARISTECH

soutenue le 27 novembre 2018

JURY

Florent Malrieu	Professeur à l'Université de Tours	Président
Carsten Hartmann	Professeur à la Brandenburgische Technische Universität	Rapporteur
Stéphane Mischler	Professeur à l'Université Paris-Dauphine	Rapporteur
Djalil Chafaï	Professeur à l'Université Paris-Dauphine	Examinateur
Marielle Simon	Chargée de recherche à Inria Lille	Examinatrice
Gabriel Stoltz	Professeur à École des Ponts ParisTech	Directeur

Remerciements

J'aimerais avant tout adresser ma reconnaissance à Gabriel Stoltz, qui a été pour moi un directeur de thèse exceptionnel. Il a tenté pendant trois années de m'inculquer un peu de sa rigueur et de son sens de l'organisation. Il s'est toujours montré accessible et à l'écoute malgré ses nombreuses contraintes, et il a su trouver les mots justes pour me soutenir tout au long de ma thèse. Je m'excuse par avance d'encourager ainsi des générations de doctorants à le déconcentrer en lui posant des questions à l'intérêt douteux, comme je l'ai fait si souvent.

Merci à Greg Pavliotis qui m'a accueilli à l'Imperial College et à Nik Nüsken pour son enthousiasme et son amitié. Je pense aussi à Alain Durmus et Christophe Andrieu, avec qui j'ai eu beaucoup de plaisir à travailler. Les hasards de la science nous ont menés à travailler ensemble sur les PDMP, lorsque Pierre Monmarché nous a révélé que nous étudions le même sujet. La bienveillance de chacun a transformé cette situation délicate en une occasion de se rencontrer, de sympathiser et de produire une plus belle science.

Ma gratitude va bien sûr aux membres du jury: d'une part mes rapporteurs Carsten Hartmann et Stéphane Mischler qui ont lu le manuscrit avec attention, mais également à Djalil Chafaï, Marielle Simon et Florent Malrieu. Je leur suis sincèrement reconnaissant d'avoir pris le temps de se pencher sur mon travail de thèse.

Je remercie aussi Isabelle Simunic et Fatna Baouj, grâce à qui je n'ai raté qu'un seul avion et à peine un ou deux trains. Leur dévouement sans faille dans l'organisation des conférences et des divers aspects administratifs m'a été d'une grande aide. Merci à Isabelle qui a su rester calme malgré mes *quelques* étourderies, oublis et autres bourdes. Je suis aussi reconnaissant envers Tony Lelièvre, Eric Cancès et Frédéric Legoll. J'ai pris autant de plaisir à les avoir comme professeurs qu'à échanger avec eux pendant ma thèse. Et merci à Cédric Doucet pour m'avoir sauvé tant de fois de mes problèmes de compilation.

Aller au Cermics a toujours été un plaisir, dans une atmosphère souvent studieuse (selon les étages) et toujours empreinte de bonne humeur. À tous ceux avec qui j'aurais aimé plus échanger, si ma thèse ne me rappelait pas si vite à mon bureau ; à ceux avec qui nous avons essayé de résoudre des problèmes improbables dans la cafèt; à ceux qui ont essayé de me parler quand j'avais mon casque sur les oreilles ; à la cafetière moka du mec pénible derrière moi ; au gatoscope, mort trop tôt malgré quelques tentatives de réanimation ; à mes élèves de Dauphine et aux derniers thésards arrivés dont je n'aurai pas réussi à retenir le nom ; aux lumières de notre open space qui ont accepté de fonctionner le temps de ma thèse ; et à tous les autres êtres animés ou non je voudrais dire merci !

Vient le moment très risqué de donner les noms des nombreuses personnes qui ont donné à cette expérience de thèse toute sa saveur: alors mes pensées vont à Marc et nos parties d'échecs à Portland ; à Grégoire qui m'a prêté ce qu'il avait de plus précieux dans cette même ville ; au mec pénible assis derrière moi qui m'a prêté sa cafetière, et qui a tenté

inlassablement de m'emmener grimper ; à Laura (mention éspéciale à ses chats) pour ses fameuses histoires à son arrivée le "matin" ; à Antoine en espérant néanmoins qu'il arrêtera de me suivre partout ; à Boris même si je n'ai pas toujours compris de quoi il parlait ; à Henri pour ses histoires improbables, et pour avoir seulement joué au ballon au 2ème étage ; à Upanshu pour m'avoir toujours conseillé et remonté le moral quand il fallait ; à Adel et Alexandre pour leur influence démoniaque ; à Amina pour sa bonne humeur légendaire ; mais aussi à Zofia, Pierre, Sami, Raphaël, Adrien, Manon, Carol, Florent, Lingling, Pierre-Loïk, François ; et pour continuer en prenant moins de risque je reprendrai l'astuce du Maître en remerciant bien sûr G, E, W, M, A, S, T, R, D, B, C et jusqu'à Z en incluant les prénoms composés. Pour ceux qui ne trouveront pas leur nom dans cette liste je tiens à dire que c'est statistiquement improbable mais je vous remercie aussi.

Mes derniers remerciements vont à ma famille qui n'a pas trop cherché à me dissuader de faire une thèse, et qui m'a toujours soutenu. Pour son plus grand bonheur, je dédie toutes ces maths à Esther qui me supporte avec brio et qui m'apporte tant de joie.

Titre : Analyse théorique et numérique de dynamiques non-réversibles en physique statistique computationnelle

Résumé : Cette thèse traite de quatre sujets en rapport avec les dynamiques non-réversibles. Chacun fait l'objet d'un chapitre qui peut être lu indépendamment.

Le premier chapitre est une introduction générale présentant les problématiques et quelques résultats majeurs de physique statistique computationnelle.

Le second chapitre concerne la résolution numérique d'équations aux dérivées partielles hypoelliptiques, c'est-à-dire faisant intervenir un opérateur différentiel inversible mais non coercif. Nous prouvons la consistance de la méthode de Galerkin ainsi que des taux de convergence pour l'erreur. L'analyse est également conduite dans le cas d'une formulation point-selle, qui s'avère être la plus adaptée dans les cas qui nous intéressent. Nous démontrons que nos hypothèses sont satisfaites dans un cas simple et vérifions numériquement nos prédictions théoriques sur cet exemple.

Dans le troisième chapitre nous proposons une stratégie générale permettant de construire des variables de contrôle pour des dynamiques hors-équilibre. Cette méthode permet en particulier de réduire la variance des estimateurs de coefficient de transport par moyenne ergodique. Cette réduction de variance est quantifiée dans un régime perturbatif. La variable de contrôle repose sur la solution d'une équation aux dérivées partielles. Dans le cas de l'équation de Langevin cette équation est hypoelliptique, ce qui motive le chapitre précédent. La méthode proposée est testée numériquement sur trois exemples.

Le quatrième chapitre est connecté au troisième puisqu'il utilise la même idée de variable de contrôle. Il s'agit d'estimer la mobilité d'une particule dans le régime sous-amorti, où la dynamique est proche d'être Hamiltonienne. Ce travail a été effectué en collaboration avec G. Pavliotis durant un séjour à l'Imperial College London.

Le dernier chapitre traite des processus de Markov déterministes par morceaux, qui permettent l'échantillonnage de mesure en grande dimension. Nous prouvons la convergence exponentielle vers l'équilibre de plusieurs dynamiques de ce type sous un formalisme général incluant le processus de Zig-Zag (ZZP), l'échantillonneur à particule rebondissante (BPS) et la dynamique de Monte Carlo hybride randomisée (RHMC). La dépendances des bornes sur le taux de convergence que nous démontrons sont explicites par rapport aux paramètres du problème. Cela permet en particulier de contrôler la taille des intervalles de confiance pour des moyennes empiriques lorsque la dimension de l'espace des phases sous-jacent est grande. Ce travail a été fait en collaboration avec C. Andrieu, A. Durmus et N. Nüsken.

Mots-clés : Physique statistique, Hors-équilibre, Réduction de variance, Analyse numérique, Équations différentielles stochastiques, Processus de Markov déterministes par morceaux

Title: Theoretical and numerical analysis of non-reversible dynamics in computational statistical physics

Abstract: This thesis deals with four topics related to non-reversible dynamics. Each is the subject of a chapter which can be read independently.

The first chapter is a general introduction presenting the problematics and some major results of computational statistical physics.

The second chapter concerns the numerical resolution of hypoelliptic partial differential equations, i.e. involving an invertible but non-coercive differential operator. We prove the consistency of the Galerkin method as well as convergence rates for the error. The analysis is also carried out in the case of a saddle-point formulation, which is the most appropriate in the cases of interest to us. We demonstrate that our assumptions are met in a simple case and numerically check our theoretical predictions on this example.

In the third chapter we propose a general strategy for constructing control variates for nonequilibrium dynamics. In particular, this method reduces the variance of transport coefficient estimators by ergodic mean. This variance reduction is quantified in a perturbative regime. The control variate is based on the solution of a partial differential equation. In the case of Langevin's equation this equation is hypoelliptic, which motivates the previous chapter. The proposed method is tested numerically on three examples.

The fourth chapter is connected to the third since it uses the same idea of a control variate. The aim is to estimate the mobility of a particle in the underdamped regime, where the dynamics are close to being Hamiltonian. This work was done in collaboration with G. Pavliotis during a stay at Imperial College London.

The last chapter deals with Piecewise Deterministic Markov Processes, which allow measure sampling in high-dimension. We prove the exponential convergence towards the equilibrium of several dynamics of this type under a general formalism including the Zig-Zag process (ZZP), the Bouncy Particle Sampler (BPS) and the Randomized Hybrid Monte Carlo (RHMC). The dependencies of the bounds on the convergence rate that we demonstrate are explicit with respect to the parameters of the problem. This allows in particular to control the size of the confidence intervals for empirical averages when the size of the underlying phase space is large. This work was done in collaboration with C. Andrieu, A. Durmus and N. Nüsken.

Key words: Statistical physics, Nonequilibrium, Variance reduction, Numerical analysis, Stochastic differential equations, Piecewise Deterministic Markov Processes

List of publications

Scientific works performed during the PhD

[1] C. Andrieu, A. Durmus, N. Nüsken and J. Roussel, Hypocoercivity for Piecewise Deterministic Markov Processes, *arXiv:1808.08592*, 2018

[2] J. Roussel and G. Stoltz, A perturbative approach to control variates, *arXiv:1712.08022*, 2017

[3] J. Roussel and G. Stoltz, Spectral methods for Langevin dynamics and associated error estimates, *M2AN*, 52(3):1051–1083, 2018

Journal publications prior to PhD

[4] A.-A. Homman, J.-B. Maillet, J. Roussel, and G. Stoltz. New parallelizable schemes for integrating the Dissipative Particle Dynamics with Energy conservation, *J. Chem. Phys.*, 144(2):024112, 2016

[5] G. Faure, J.-B. Maillet, J. Roussel, and G. Stoltz, Size consistency in smoothed dissipative particle dynamics, *Physical Review E*, 94(4):043305, 2016

[6] C. Le Bris, P. Rouchon, and J. Roussel, Adaptive low-rank approximation and denoised Monte Carlo approach for high-dimensional Lindblad equations, *Physical Review A*, 92(6):062126, 2015

Contents

Résumé en français

Introduction

La physique statistique vise à faire le pont entre les caractéristiques microscopiques d'un système physique et son comportement macroscopique. Rappelons quelques ordres de grandeur pour appréhender le défi que cela représente. La distance typique entre deux atomes est de l'ordre de quelques Angstroms ($1\text{Å} = 10^{-10}\text{m}$), ce qui implique que le nombre d'atomes dans un échantillon macroscopique de matière est de l'ordre du nombre d'Avogadro $\mathcal{N}_\text{A} \sim 10^{23}$. Les échelles de temps sont également éloignées puisque l'unité pertinente pour l'évolution des systèmes macroscopiques est la seconde (éventuellement la minute, heure ou année selon l'application considérée), alors que les atomes des molécules vibrent à une fréquence comprise entre 10^{12}Hz et 10^{14}Hz.

Les simulations numériques en dynamique moléculaire se limitent généralement à des systèmes de moins d'un million d'atomes sur une durée de moins d'une milliseconde. Ils sont donc loin de pouvoir simuler des systèmes macroscopiques à l'échelle microscopique. Il y a cependant des domaines dans lesquels de tels calculs sont précieux, par exemple à des fins médicales telles que la conception de médicaments ou la compréhension du repliement des protéines, ou l'étude et la conception des matériaux, lorsque le comportement macroscopique peut être déduit de simulations à l'échelle microscopique. Il est par exemple possible de calculer l'équation d'état d'un système homogène, y compris des régimes de pression et de température extrêmes inaccessibles aux expériences. La dynamique moléculaire a conquis la reconnaissance de la communauté scientifique, avec notamment un prix Nobel partagé en 2013 entre Martin Karplus, Michael Levitt et Arieh Warshel "pour le développement de modèles multi-échelles pour des systèmes chimiques complexes". Pour une introduction plus complète à la dynamique moléculaire, nous renvoyons aux livres [2, 64, 101].

Les systèmes microscopiques typiques d'intérêt évoluent souvent à température constante, au contact d'un thermostat. Ces systèmes sont bien décrits à l'aide de la dynamique de Langevin, qui est une équation différentielle stochastique classique. De plus, ces systèmes peuvent être hors-équilibre. Cela signifie qu'un forçage externe induit un flux constant de masse ou d'énergie dans le système, de sorte que son évolution trahit la flèche du temps. La sensibilité du système à ces perturbations est quantifiée par les coefficients de transport.

Méthodes spectrales pour la dynamique de Langevin et estimées d'erreur associées

Chapitre 2 (publié dans [145]) traite de l'approximation numérique de la solution (unique) de l'équation aux dérivées partielles

$$-\mathcal{L}\Phi = R,$$

où $\mathcal{L}$ est un opérateur hypocoercif, inversible sur l'espace $\mathrm{L}^2_0(\mu)$ où μ est une mesure de probabilité, et $R \in \mathrm{L}^2_0(\mu)$. Ce type d'équation de Poisson apparaît notamment dans le contexte de dynamiques particulaires cinétiques telles que la dynamique de Langevin. Résoudre cette équation pour Φ fournit beaucoup d'information sur les propriétés dynamiques du processus. Il permet par exemple de calculer la variance asymptotique de l'observable R,

un coefficient de transport ou encore de construire une variable de contrôle comme expliqué au Chapitre 3.

Nous approchons numériquement la solution Φ en utilisant une méthode Galerkine. Dans le cas où le générateur $\mathcal{L}$ est coercif, le théorème de Lax-Milgram assure que la méthode est bien posée. De plus, les estimations d'erreur sont fournies par le Lemme de Céa dans ce cas. Nos résultats étendent ces garanties théoriques au cas où l'opérateur $\mathcal{L}$ n'est pas coercif mais seulement hypocoercif, sous des hypothèses adéquates. En particulier, nous dérivons des conditions sous lesquelles la restriction du générateur $\mathcal{L}$ à l'espace de Galerkine $V \subset \mathrm{L}^2_0(\mu)$ est encore hypocoercif. Une formulation de point-selle, impliquant un multiplicateur de Lagrange, est également proposée et analysée. Cette approche permet d'écrire la méthode Galerkine dans tout l'espace $\mathrm{L}^2(\mu)$, ce qui s'avère plus adapté.

Pour une particule piégée dans un potentiel sinusoïdal unidimensionnel, nous prédisons théoriquement que les erreurs de consistance et d'approximation décroîtront polyomialement avec un degré dépendant de la régularité de la fonction R mesurée dans des espaces de Sobolev. Nos résultats numériques dont la Figure 1 rend compte viennent confirmer nos résultats théoriques. Nous montrons également Figure 2 que pour ce système il est également possible de calculer le coefficient de mobilité de façon très rapide et très précise en utilisant la méthode de Galerkine.

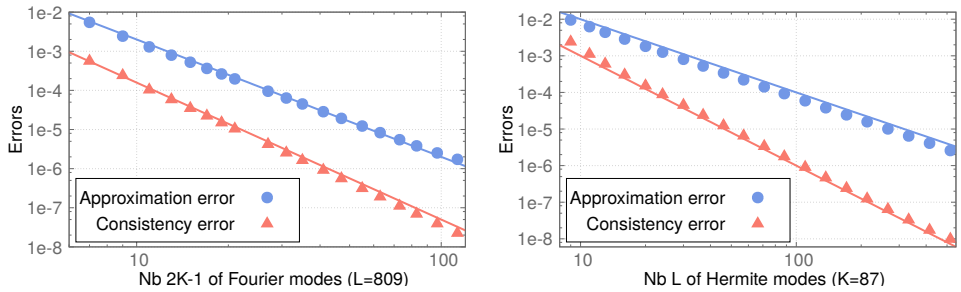


Figure 1: Erreurs d'approximation et de consistance en fonction du nombre de modes. Gauche: le nombre de modes de Fourier varie, pour un nombre fixe mais élevé de modes de Hermite ; l'erreur d'approximation décroît en K^{-3} tandis que l'erreur de consistance décroît en $K^{-7/2}$. Droite: le nombre de modes de Hermite varie, pour un nombre fixe mais élevé de modes de Fourier ; l'erreur d'approximation décroît en L^{-2} tandis que l'erreur de consistance décroît en L^{-3}.

Une approche perturbative des variables de contrôle en dynamique moléculaire

Le Chapitre 3 (voir la prépublication [144]) est motivée par l'estimation efficace des coefficients de transport. Nous présentons une stratégie générale de réduction de variance basée sur des variables de contrôle qui ne reposent pas sur la connaissance de la distribution de probabilité invariante. Cette dernière est en effet inconnue lorsque le système se trouve dans un état stationnaire hors-équilibre, de sorte que les techniques standard de

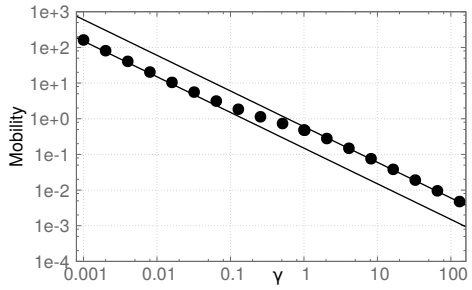


Figure 2: Mobilité en fonction de la friction γ. Elle décroît en γ^{-1} à la fois dans le régime des petits γ (avec un préfacteur 0.15) et dans le régime des grands γ (avec un préfacteur 0.6).

réduction de la variance (échantillonnage par importance, stratification, etc.) ne peuvent pas être utilisées. L'idée est que si la dynamique peut être approchée par une dynamique plus simple, pour laquelle des solutions aux équations de Poisson peuvent être calculées (par ex. une dynamique linéaire ou en faible dimension) alors nous pouvons construire une observable modifiée. Le nouvel estimateur est non biaisé par construction et sa variance asymptotique est grandement réduite si la dynamique simplifiée est proche de la dynamique initiale, comme nous le montrons dans des exemples précis.

Nous prouvons que la variance asymptotique de cet estimateur dépend quadratiquement de l'amplitude de la différence entre les deux dynamiques dans le régime perturbatif. Cette technique de réduction de variance est illustrée sur trois systèmes hors-équilibre : une particule dans un potentiel périodique unidimensionnel subissant un forçage non-gradient ; le calcul du flux thermique traversant une chaîne ; et l'estimation de la longueur moyenne d'un dimère dans un solvant en présence d'une force externe de cisaillement.

Dans le cas unidimensionnel, la construction de la variable de contrôle repose sur la résolution de l'équation de Poisson à l'équilibre avec un nombre M de fonctions de base. La Figure 3 montre que l'estimation de la mobilité n'est pas biaisée par la variable de contrôle, et que l'estimation est plus précise avec variable de contrôle, en particulier dans le cas qui nous intéresse où le forçage est faible. La Figure 4 vient quantifier ce phénomène, montrant la réduction de variance obtenue pour plusieurs valeurs de forçage et de taille de base M.

Le second système étudié est une chaîne d'atomes dont on cherche à estimer la conductivité thermique (Figure 5). Ici le générateur de l'évolution est approché par le générateur d'un système linéaire, correspondant à des interactions harmoniques. Nous montrons Figure 6 que lorsque le potentiel d'interaction est proche d'être harmonique la réduction de variance sur l'estimateur du flux thermique permettant de calculer la conductivité est importante. La taille de la chaîne est également étudiée.

Le dernier exemple numérique concerne l'estimation de l'élongation moyenne d'un dimère dans un solvant (Figure 7). Il s'agit d'estimer la dépendance de cette quantité avec le potentiel d'interaction caractérisant le solvant, ainsi qu'avec l'intensité du cisaillement généré par la force extérieure. On peut voir par exemple Figure 8 que le cisaillement a tendance à allonger le dimère. La réduction de variance permise par la variable de contrôle est tracée

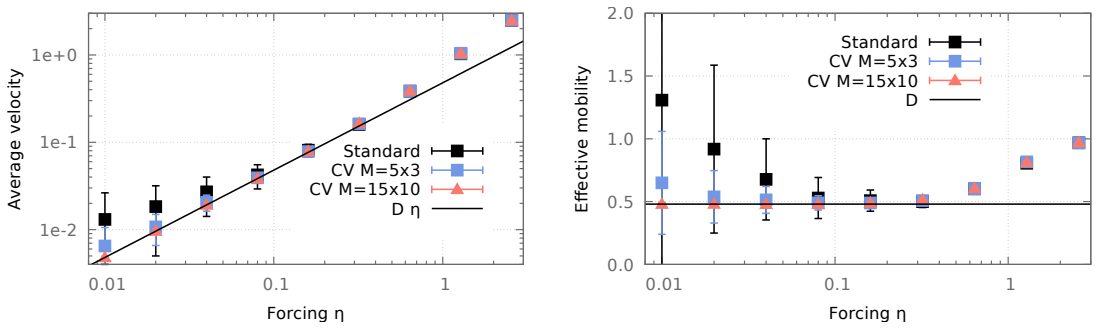


Figure 3: Réponse linéaire pour la simulation de Monte Carlo standard (carrés noirs) comparée à la version avec variable de contrôle (bleu, rouge) et à la réponse asymptotique $D\eta \approx 0.48\eta$ (ligne noire).

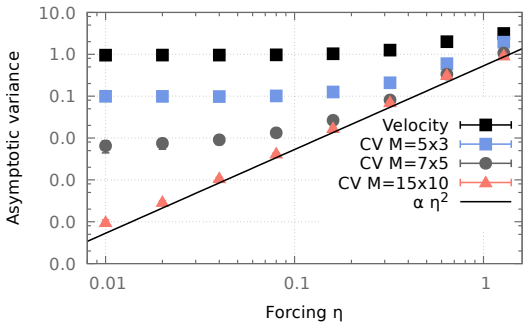


Figure 4: Variance asymptotique de la vitesse (carrés noirs) comparée à celle estimée lorsqu'une variable de contrôle est utilisée (bleu, gris, rouge) et à la variance réduite (ligne noire) prédite théoriquement ($\alpha \approx 0.53$ calculée avec la méthode de Galerkine).

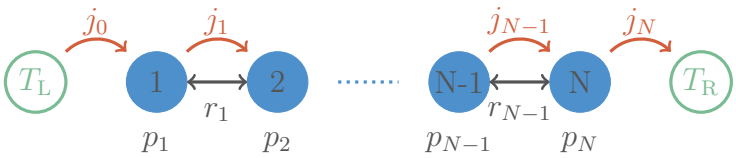


Figure 5: Transport thermique dans une chaîne d'atomes en une dimension.

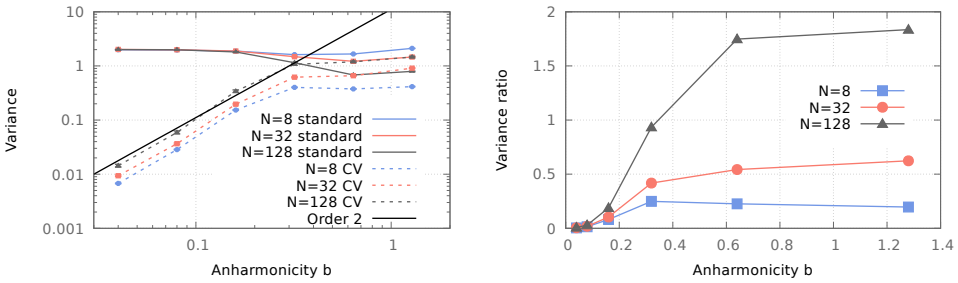


Figure 6: Gauche : Comparaison entre les variances du flux thermique standard et du flux modifié. Droite : Les ratios de variance estimés correspondent à la variance modifiée divisée par la variance standard.

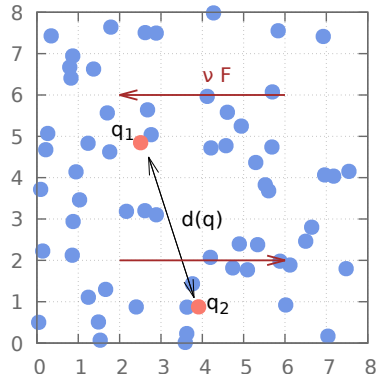


Figure 7: Dimère (rouge) dans un solvant (bleu) sous l'effet d'une force de cisaillement (brun).

sur la Figure 9. On peut voir que lorsque le potentiel régissant l'interaction avec le solvant est raide la réduction de variance est limitée, même en l'absence de cisaillement.

Estimation efficace de la mobilité dans le régime sous-amorti

Dans le chapitre 4, nous montrons comment la stratégie de réduction de la variance du chapitre 3 peut être adaptée au cas d'une dynamique Langevin de faible dimension dans le régime sous-amorti. Ce travail a été effectué à l'Imperial College de Londres avec G. Pavliotis dans le cadre d'un programme de mobilité de deux mois. Nous rappelons dans une première partie que dans cette limite la dynamique rescalée en temps converge vers un processus de diffusion sur un graphe. Nous construisons ensuite une variable de contrôle à l'aide de l'équation de Fokker-Planck correspondant à cette dynamique limite. Nous obtenons ainsi un nouvel estimateur de le mobilité. Nous illustrons par des simulations numériques qu'il se comporte bien dans le régime sous-amorti. La variable de contrôle ainsi construite nous permet d'étudier comment la mobilité dépend du coefficient de friction pour un système bidimensionnel non intégrable. Nous montrons numériquement que la mobilité

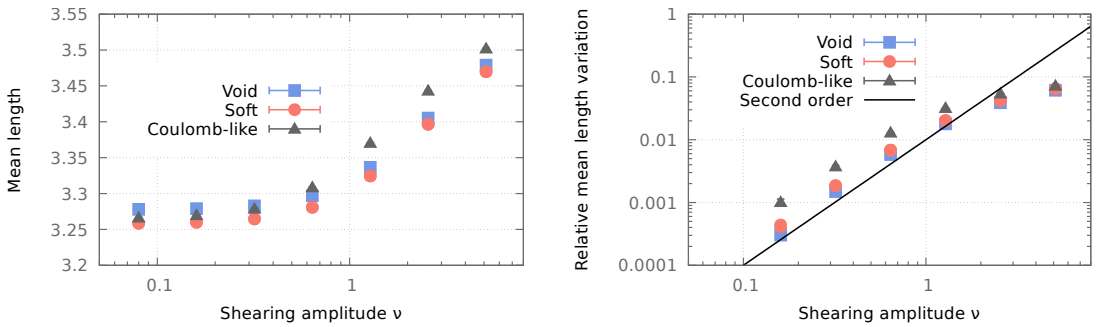


Figure 8: Gauche : Longueur moyenne du dimère, soit non-solvaté (dans le vide), soit dans un solvant avec un potentiel qui est soit "Soft", soit de type interaction de Coulomb. Droite : Variation relative de la longueur moyenne induite par le cisaillement. La ligne continue représente le scaling de référence ν^2.

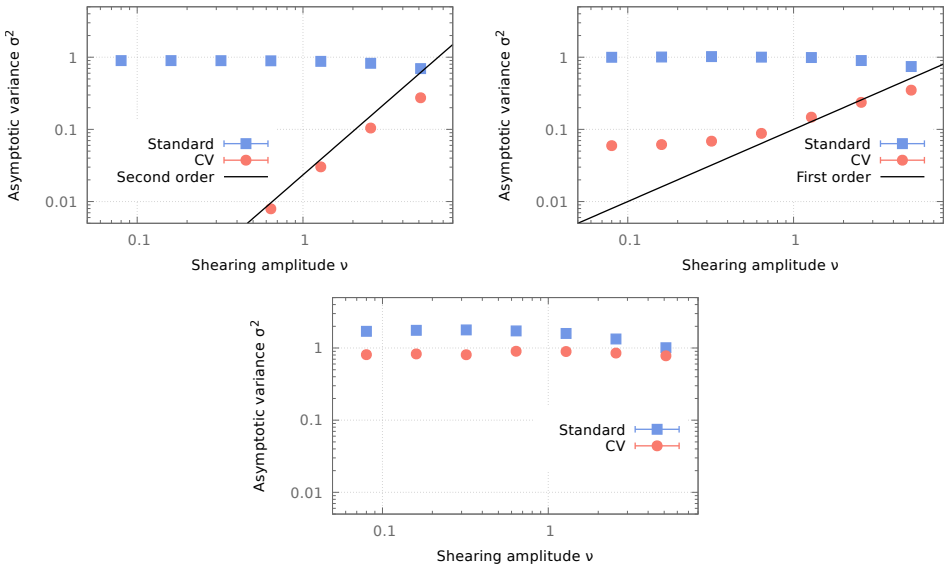


Figure 9: Variance asymptotique de la longueur du dimère, avec ou sans variable de contrôle. Gauche : Dimère non-solvaté. Droite : Solvant avec un potentiel "Soft". Bas : Solvant avec un potentiel de type Coulombien.

ne se comporte pas comme $1/\gamma$, contrairement au cas unidimensionnel. Cependant, ce résultat n'est pas bien compris du côté théorique.

Hypocoercivité des processus de Markov continus par morceaux

Dans le Chapitre 5 (voir la prépublication [4]) nous montrons qu'une grande classe de dynamiques cinétiques impliquant des sauts de la vitesse, appelées PDMPs, sont géométriquement ergodiques sous des hypothèses faibles sur le potentiel. Ce travail a été initié à l'Imperial College de Londres avec N. Nüsken et poursuivi en collaboration avec C. Andrieu et A. Durmus. Cette classe de PDMP comprend le processus Zig-Zag (ZZ), l'échantillonneur élastique de particules (BPS) ou le Monte Carlo Hamiltonien Randomisé (RHMC). Le ZZ et le BPS sont des outils récents en statistique et en physique statistique, où ils ont tous deux suscité un grand intérêt. Le RHMC, en revanche, est un échantillonneur classique.

Notre preuve, reposant sur des techniques d'hypocoercivité dans L^2 (voir Section 1.2.2.3), permet de prouver des estimations quantitatives de trou spectral. La dépendance de nos bornes par rapport au taux de rafraîchissement, le choix de l'espace des vitesses et surtout la dimension d, sont explicites. Nous prouvons en particulier que le trou spectral est uniformément séparé de zéro dans la limite des hautes dimensions pour le ZZ et le RHMC sous certaines hypothèses simples sur le potentiel. Ce résultat vaut également pour la dynamique de Langevin. Nous prouvons également sous les mêmes hypothèses que pour le BPS le trou spectral est supérieur à $d^{-1/2}$, ce qui semble être le bon taux au regard du résultat de [17].

Chapter 1

Introduction

Contents

Statistical physics aims at closing the gap between the microscopic features of a physical system and its macroscopic behavior. Let us recall some orders of magnitude to apprehend the challenge this represents. The typical distance between two atoms is of the order of a few Angstroms ($1\text{Å} = 10^{-10}\text{m}$), which implies that the number of atoms in a macroscopic sample of matter is of the order of the Avogadro number $\mathcal{N}_\text{A} \sim 10^{23}$. The time scales are also far apart since the relevant unit for the evolution of macroscopic systems is the second (possibly minutes, hours or years depending on the application under consideration), whereas atoms in molecules vibrate at a frequency in the range $10^{12} - -10^{14}\text{Hz}$.

Numerical simulations in molecular dynamics are typically restricted to systems of less than one million atoms over times of less than a millisecond. They are therefore far from being able to simulate macroscopic systems at a microscopic scale. There are however fields in which such computations are precious, *e.g.* for medical purposes such as drug design or the understanding of protein folding, or the study and design of materials, when the macroscopic behavior can be inferred from small-scale simulations. It is for example possible to compute the equation of state of a homogeneous system, including regimes of extreme pressure and temperature inaccessible to experiments. Molecular dynamics has earned the recognition of the scientific community, notably with a Nobel prize shared in 2013 between Martin Karplus, Michael Levitt and Arieh Warshel "for the development of multiscale models for complex chemical systems". For a more complete introduction to molecular dynamics we refer to the books [2, 64, 101].

Typical microscopic systems of interest often evolve at constant temperature, in contact with a thermostat. Such systems are well described using Langevin dynamics, which is a popular stochastic differential equation. Moreover these systems can be out of equilibrium, meaning that an external forcing induces a steady flux of mass or energy in the system, so that the arrow of time can be read off the evolution. The sensitivity of the system to these perturbation is quantified by the transport coefficients.

The introduction of this thesis is organized as follows: we describe the fundamentals of computational statistical physics in Section 1.1; we study Langevin dynamics focusing first on equilibrium systems in Section 1.2 and then on non-equilibrium systems in Section 1.3; we finally review variance reduction techniques in Section 1.4; and in Section 1.5 we highlight our contributions.

1.1 Fundamentals of statistical physics

Let us now describe molecular dynamics with more details. The general framework and notation are given in Section 1.1.1. Section 1.1.2 introduces the notion of macrostate, which is the fundamental tool allowing to make the connection between microscopic and macroscopic descriptions of matter. We end this part with Section 1.1.3, where microscopic dynamics are introduced as a tool to compute average thermodynamic quantities.

1.1.1 Microscopic description of matter

In statistical physics the configuration of a system at a given time can be described by the positions and velocities of every particle, which are typically atoms. These particles interact through a potential energy, in the framework of classical mechanics. Potential energies can be computed using *ab initio* simulations [30], which involve quantum physical models, though this is only computationally tractable for small systems. They can also be given by empirical formulas, the parameters of which are tuned in order to reproduce experimental measurements with numerical simulations.

Denoting by d the dimension of the space in which the particles live and by N the number of particles, the vector of all positions $q = (q_1, \cdots q_N)$ is an element of the domain $\mathcal{D} \subset \mathbb{R}^D$ with $D = dN$. Periodic boundary conditions are often considered, in which case the domain is a box of width $L > 0$ and $\mathcal{D} = (L\mathbb{T})^D$. Instead of the velocities of the particles we rather consider momenta, which are defined as the product of the mass m of a particle[1] by its velocity. The vector of all momenta is denoted by $p = (p_1, \cdots, p_N) \in \mathbb{R}^D$. In the following the configuration is often denoted by $x = (q, p)$.

The interaction potential associated to the position q is denoted by $V(q)$. Many types of potentials, depending on the applications, have been proposed in the literature. A very simple case is when the potential is pairwise and depends only on the distance between the particles:

$$V(q) = \sum_{1 \leqslant i < j \leqslant N} v(|q_i - q_j|),$$

where $|\cdot|$ is the Euclidean norm. When simulating atoms or molecules with non-bonded interactions, the pair potential v is repulsive at short range to account for exclusion of the electronic clouds and converges to 0 at long range to ensure local interactions. A popular potential, which has been proposed in the earliest days of molecular dynamics to simulate noble gases such as Argon, is the Lennard-Jones potential [152]:

$$v(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right].$$

It is the sum of a repulsive term scaling as r^{-12} which accounts for short range interaction and an attractive term scaling as $-r^{-6}$ corresponding to van der Waals contributions.

1.1.2 Macroscopic description of matter and thermodynamic properties

At the macroscopic time and space scales, only the averages of some observables such as the temperature, the pressure or the energy are measurable. These averages can be interpreted as the expectation of an observable with respect to a probability measure μ on the phase

[1] In the present work we will restrict to the case when all particles share the same mass for notational simplicity.

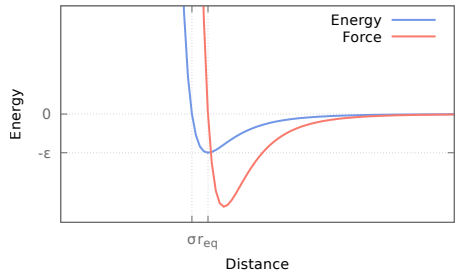


Figure 1.1: Lennard-Jones potential and the corresponding force.

space $\mathcal{E} = \mathcal{D} \times \mathbb{R}^D$:

$$\mathbb{E}_\mu[\varphi] = \int_{\mathcal{E}} \varphi(q,p)\mu(\mathrm{d}q, \mathrm{d}p).$$

The measure μ is called the macrostate of the system, by opposition to an element $(q,p) \in \mathcal{E}$ which is called a microstate. Macrostates are also named thermodynamic ensembles. They can be obtained by a maximization of the entropy under a given constraint [87]. The entropy of a macrostate measures the quantity of disorder. For any measure μ with a density with respect to the Lebesgue measure (this density being also denoted by μ with some abuse of notation), it writes

$$S(\mu) = -\int_{\mathcal{E}} \mu(q,p) \ln \mu(q,p) \mathrm{d}q\, \mathrm{d}p.$$

The thermodynamic ensemble corresponding to a given constraint, such as a fixed mean pressure or a fixed mean energy, corresponds to the maximizer of the entropy under this constraint [12, Chapter 3]. In other words, the thermodynamic ensemble is the most disordered macrostate compatible with the data. In particular, in the case when the state space is compact and no observation has been made, the maximizer of the entropy is the uniform measure. This represents the lack of information. In presence of measurements, the probability measure will depart from the uniform one as little as possible (in terms of S) in order to stay consistent with the data.

Closed systems are characterized by a constant number of particles N, volume V and energy E. The energy of a microstate is given by the Hamiltonian

$$H(q,p) = V(q) + \frac{1}{2m}|p|^2.$$

The associated macrostate is the microcanonical ensemble, denoted by NVE. It corresponds to the maximization of the entropy under the constraint that the probability measure is supported by the manifold of microstates with energy E. Note that this class of measures is not absolutely continuous with respect to the Lebesgue measure, so that the entropy has to be defined through a limiting process. The solution is the uniform probability measure on the manifold of codimension 1 of microstates with energy E.

In many physical situations, one considers open systems in contact with a thermostat at temperature T, so that the energy is fixed in average to a value E but it is not constant. The

corresponding thermodynamic ensemble is named the canonical ensemble and is denoted by NVT [12]. It is defined as the maximizer of the entropy over probability measures ρ on $\mathcal{E}$ such that $\int_{\mathcal{E}} H\rho = E$:

$$\sup_{\rho} \left\{ S(\rho) \;\middle|\; \rho \in \mathrm{L}^1(\mathcal{E}), \int_{\mathcal{E}} \rho = 1, \int_{\mathcal{E}} H\rho = E \right\}.$$

The solution will be denoted by μ in the following, and it writes

$$\mu(\mathrm{d}q\,\mathrm{d}p) = Z_\beta^{-1} \exp(-\beta H(q,p))\,\mathrm{d}q\,\mathrm{d}p =: \nu(\mathrm{d}q)\,\kappa(\mathrm{d}p),$$

where

$$\nu(\mathrm{d}q) = Z_{\nu,\beta}^{-1} \mathrm{e}^{-\beta V(q)}\,\mathrm{d}q, \qquad \kappa(\mathrm{d}p) = \left(\frac{\beta}{2\pi m}\right)^{D/2} \mathrm{e}^{-\beta|p|^2/(2m)}\,\mathrm{d}p,$$

and Z_β, $Z_{\nu,\beta}$ are normalization constants (the potential is such that $\mathrm{e}^{-\beta V} \in \mathrm{L}^1(\mathcal{E})$). The constant $\beta > 0$ is the opposite of the Lagrange multiplier associated with the energy constraint $\int_{\mathcal{E}} H\mu = E$, and it allows to define the temperature T through $\beta = \frac{1}{k_\mathrm{B} T}$ where k_B is Boltzmann's constant. We say that the NVT ensemble is at fixed temperature.

The measure μ is tensorized: the marginal measure ν in the position variable contains the physical information of the system, and it will be the center of attention in the following (see Section 1.1.3); whereas the marginal in the momentum variable is a normal distribution with covariance matrix $\frac{m}{\beta}\mathbf{I}$.

Other thermodynamic ensembles exist, for each of them three macroscopic invariants are usually considered, the value of which can be fixed a priori: number of particles N, volume V, temperature T, energy E, pressure P, chemical potential μ... We mention for example the isobaric-isothermal ensemble NPT (where the pressure and the temperature are fixed, but the volume can vary) [53] and the grand canonical ensemble μPT (where the chemical potential μ, the pressure and the temperature are fixed, but the number of particles and the volume can vary) [117, Chapter 3]. In this work we focus on the canonical ensemble NVT which is the most standard one.

1.1.3 Sampling

Observables of the system are functions of the microstate $x = (q,p) \in \mathcal{E}$, and they are typically given by a closed expression, such as the energy $H(q,p) = V(q) + |p|^2/(2m)$. Another example is the pressure of a three-dimensional fluid with periodic boundary conditions:

$$\varphi(q,p) = \frac{1}{3|\mathcal{D}|} \sum_{i=1}^{N} \left(\frac{|p_i|^2}{m} - q_i \cdot \nabla_{q_i} V(q) \right). \tag{1.1}$$

The contribution of the first term, which corresponds to the pressure of an ideal gas ($V = 0$), can be computed analytically under the canonical measure. Indeed we can define

$$\overline{\varphi}(q) := \int_{\mathbb{R}^D} \varphi(q,p)\kappa(\mathrm{d}p) = \frac{1}{3|\mathcal{D}|}\left(\frac{d}{\beta} - \sum_{i=1}^{N} q_i \cdot \nabla_{q_i} V(q)\right),$$

and we then have $\mathbb{E}_\mu[\varphi] = \mathbb{E}_\nu[\overline{\varphi}]$. Such a situation, where the momenta can often be integrated out, is quite common. The challenge indeed lies in the integration with respect to position variables, under the measure ν. The probability measure encoding the thermodynamic properties of the system is known explicitly, so that the mean of an observable φ writes as the integral over $\mathcal{E}$ of an integrand which has a closed expression. The challenge relies in the fact that this integral is over a very large dimensional space since D is typically large. Classical quadrature techniques are therefore not suited – a problem known as the curse of dimensionality.

Stochastic techniques are a generic way to tackle large dimensional sampling problems, and to avoid the computational complexity to explicitly scale exponentially with D. The most simple and efficient Monte Carlo method to compute an integral consists in averaging φ over a sample of independent microstates $(X_i)_{i\geqslant 1} = (q^i, p^i)_{i\geqslant 1}$ drawn under the probability measure μ. Owing to the Law of Large Numbers, the following holds almost surely

$$\frac{1}{n}\sum_{i=1}^{n} \varphi(X_i) \xrightarrow[n\to\infty]{} \mathbb{E}_\mu[\varphi] = \int_{\mathcal{E}} \varphi(q,p)\,\mu(\mathrm{d}q\,\mathrm{d}p). \tag{1.2}$$

Drawing such independent samples is however feasible only for particularly simple probability measures, such as Gaussian measures. Practical methods generate correlated samples, typically by realization of Markov chains. This is why we only assume a Markov property for the samples, instead of taking them independent. Conditions guaranteeing the almost sure convergence of ergodic averages (1.2) for the dynamics we consider are provided in Subsection 1.2.1. This property is known as pathwise ergodicity, see Proposition 1.1 for some sufficient conditions.

In this section we present several types of samplers. First we present a general manner to create a Markov chain sampling a given probability measure with the Metropolis-Hastings algorithm. Second we consider the Hamiltonian dynamics, which is deterministic. We then introduce stochastic differential equations using Brownian motions, and in particular the Langevin equation which will be the main focus of this thesis. We conclude with Piecewise Deterministic Markov Processes where Poisson processes are considered rather than Brownian motions.

1.1.3.1 Markov Chain Monte Carlo

The Metropolis-Hastings algorithm allows to sample from ν without bias. This method is extremely popular, especially in the Bayesian statistics community. It was first introduced in [112] and then refined in [77]. This method relies on a proposition kernel T whose role is to suggest a new configuration. A move is accepted with a probability depending on the

Metropolis-Hastings ratio r.

Algorithm 1.1 (Metropolis-Hastings)**.** *For a given initial configuration q^0, iterate on $n \geqslant 0$: 1) Propose a state $\widetilde{q}^{n+1}$ according to the proposition kernel $T(q^n, \cdot)$;*
2) Accept the proposition with probability

$$r(q, q') = \min\left(1, \frac{T(q', \mathrm{d}q)\nu(\mathrm{d}q')}{T(q, \mathrm{d}q')\nu(\mathrm{d}q)}\right),$$

and set in this case $q^{n+1} = \widetilde{q}^{n+1}$; otherwise set $q^{n+1} = q^n$.

Assuming that T is such that the Markov chain is irreducible (see Proposition 1.1 for a definition), the samples $(X^i)_{i \geqslant i}$ generated satisfy (1.2): the Markov chain is pathwise ergodic [113, Theorem 17.1.7]. Note that the acceptation probability is well defined assuming that the probability measures $T(q', \mathrm{d}q)\nu(\mathrm{d}q')$ and $T(q, \mathrm{d}q')\nu(\mathrm{d}q)$ are equivalent. Moreover this ratio does not depend on the normalization constant $Z_{\nu,\beta}$ of the measure ν, so that the algorithm can be performed even if this constant is not known.

In practice the efficiency of the method relies on the choice of the proposition kernel T. In order to illustrate this point we consider the Gaussian kernel $T(q, \mathrm{d}q') \propto \exp\left(-\frac{(q-q')^2}{2h^2}\right)\mathrm{d}q'$. A small spread h leads to correlated subsequent configurations, while with a large spread h the proposals are likely to be rejected since they often end up in unlikely regions. In this second case the Markov Chain can cross energetic barriers but the large rejection rate leads to a large correlation. Practitioners generally consider that a good trade-off is obtained when the rejection rate is of the order of a fraction of unity. This heuristic is comforted by theoretical studies for simplistic target measures [141, 142].

1.1.3.2 Hamiltonian dynamics

The evolution of an isolated system, for which the NVE ensemble is relevant, is given by Newton's equations of motion. This is why historically the microcanonical ensemble for a Hamiltonian H has been sampled using the associated Hamiltonian dynamics: the initial condition is given by $q(0) = q_0$, $p(0) = p_0$ and time averages are taken over trajectories

$$\begin{cases} \dot{q}(t) = \dfrac{1}{m}p(t), \\ \dot{p}(t) = -\nabla V(q(t)), \end{cases}$$

which also write in the more abstract form

$$\begin{cases} \dot{q}(t) = \partial_p H(q(t), p(t)), \\ \dot{p}(t) = -\partial_q H(q(t), p(t)). \end{cases}$$

The solutions are well defined for any time $t \geqslant 0$ if V is bounded from below and gradient Lipschitz continuous. The Hamiltonian is preserved along the trajectory since

$$\frac{\mathrm{d}}{\mathrm{d}t}H(q(t), p(t)) = \dot{q}(t)\partial_q H(q(t), p(t)) + \dot{p}(t)\partial_p H(q(t), p(t)) = 0.$$

The canonical distribution (NVE ensemble) can also be shown to be invariant by the dynamics. There is however no theoretical guarantee for the convergence (1.2) in the general case, since due to the fixed energy the dynamics may be unable to overcome some energetic barriers. Non-ergodicity is for example clear in the case of a multimodal potential with a sufficiently small energy. Pathwise ergodicity can however be shown rigorously for integrable systems and their perturbations [6]. The ergodicity issue can also be solved by considering dynamics involving randomness [56]. In any case, Hamiltonian dynamics are a useful building block for stochastic dynamics which are used to sample from the NVT ensemble.

1.1.3.3 Stochastic differential equations (SDE)

The first type of randomness which can be considered is based on the Brownian motion. Consider the general time-homogeneous stochastic differential equation on $\mathbb{R}^D$:

$$\mathrm{d}x_t = b(x_t)\ \mathrm{d}t + \sigma(x_t)\,\mathrm{d}W_t, \tag{1.3}$$

for a given initial condition $x_0 \in \mathbb{R}^D$ and standard Brownian motion $W_t \in \mathbb{R}^m$, and where $b : \mathbb{R}^D \to \mathbb{R}^D$ and $\sigma : \mathbb{R}^D \to \mathbb{R}^{D\times m}$ are assumed to be locally Lipschitz, so that there exist a unique (strong) solution local-in-time. We refer to [137, 76] for the existence of a global-in-time solution. In practice the solution is not analytical but it can be approximated numerically (see Section 1.2.4), leading to a bias in the sampling.

Given a $\mathcal{C}^\infty$ observable φ with compact support, we define the generator $\mathcal{L}$ of the SDE (1.3) by its action

$$\forall x \in \mathcal{E}, \quad \mathcal{L}\varphi(x) := \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\varphi(x_t)|x_0 = x]\bigg|_{t=0}. \tag{1.4}$$

The generator is a differential operator, obtained by Itô calculus:

$$\mathcal{L} = b \cdot \nabla + \frac{1}{2}\sigma\sigma^\top : \nabla^2,$$

where : denotes the Frobenius inner product:

$$\forall A, B \in \mathbb{R}^{D\times D}, \quad A : B := \mathrm{Tr}(A^\top B).$$

Langevin dynamics. The Langevin dynamics is a kinetic SDE, since the state can be decomposed as $x_t = (q_t, p_t)$ where q_t represents the configuration at time t and p_t is the momentum at time t. This dynamics is built from the Hamiltonian dynamics but it models the evolution of systems at constant temperature. It samples the canonical ensemble NVT and not the microcanonical ensemble NVE. Indeed the action of a thermostat is modeled using an additional Ornstein-Uhlenbeck process, which is composed of a friction term and

a stochastic fluctuation:

$$\begin{cases} \mathrm{d}q_t = \dfrac{1}{m} p_t \, \mathrm{d}t \,, \\ \mathrm{d}p_t = -\nabla V(q_t) \, \mathrm{d}t - \dfrac{\gamma}{m} p_t \, \mathrm{d}t + \sqrt{2\gamma\beta^{-1}} \mathrm{d}W_t, \end{cases} \tag{1.5}$$

where $\gamma > 0$ is the friction coefficient and $\beta^{-1} > 0$ is the temperature. The generator of the Langevin equation writes:

$$\mathcal{L} := \frac{1}{m} p \cdot \nabla_q - \nabla V \cdot \nabla_q - \frac{\gamma}{m} p \cdot \nabla_p + \gamma\beta^{-1} \Delta_p. \tag{1.6}$$

The Langevin dynamics is a particularly efficient sampler in practice [32], and understanding why it is so is an active field of research. It is also easy to implement in the existing production codes, especially starting from Hamiltonian dynamics. The main advantage compared to the latter is that pathwise ergodicity is proved under weak assumptions, so that the convergence of the empirical estimator (1.2) is granted.

The Langevin dynamics (1.5) will be the focus of a large part of the present work. Its mathematical properties, such as pathwise ergodicity and the existence of confidence intervals for the ergodic estimator, are discussed in Sections 1.2 and 1.3.

Overdamped Langevin dynamics In the limit of large frictions $\gamma \to \infty$, upon rescaling time as γt, the solution $q^\gamma_{\gamma t}$ to the Langevin equation (for the friction γ) converges in law [146] to the solution of the overdamped (or Smoluchowski) equation,

$$\mathrm{d}Q_t = -\nabla V(Q_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}W_t, \tag{1.7}$$

which admits for generator

$$\mathcal{L} = -\nabla V \cdot \nabla_Q + \beta^{-1} \Delta_Q.$$

This dynamics allows to sample the position marginal ν of the Gibbs measure μ since it is pathwise ergodic. For many observables $\varphi(q, p)$ the mean with the respect to the canonical distribution μ can be rewritten as a mean with respect to the probability measure ν, as it is the case for the pressure (1.1) or for any observable depending only on the position variable. Overdamped Langevin dynamics allows then to compute these means using an ergodic average.

1.1.3.4 Piecewise Deterministic Markov Processes (PDMP)

The invariant probability measure μ can also be sampled using stochastic dynamics involving Poisson processes instead of a Brownian motion [38]. For such dynamics the trajectory is deterministic, and sometimes straight, between two events. These events are momentum jumps, meaning that the new momentum vector is either drawn from a probability measure $Q(q, p, \mathrm{d}p')$, or modified according to a deterministic relation. The time between each event is given by an inhomogeneous Poisson law, so that the process is Markovian. This family

of processes encompasses the Bouncy Particle Sampler [25], Hybrid Monte Carlo [45], Run-and-Tumble [29], Event Chain Monte Carlo [114], and the Zig-Zag Process [16]. We present here the latter as an example, as it is both conceptually simple and numerically promising for certain applications. In the following the state space writes $\mathcal{E} = \mathcal{D} \times \mathcal{P}$: the momenta take value in the kinetic ensemble $\mathcal{P} \subset \mathbb{R}^D$.

The standard form of the Zig-Zag process involves a kinetic ensemble $\mathcal{P} = \{-1, +1\}^D$, so that for any $1 \leqslant i \leqslant D$, $p_i \in \{-1, +1\}$. Between two events the trajectory is a straight line:

$$\begin{cases} \mathrm{d}q_t = p_t \ \mathrm{d}t \,, \\ \mathrm{d}p_t = 0. \end{cases}$$

The events correspond to a superposition of D jump processes, one for each component. The i-th Poisson process has jump rate $(\partial_i V(q_t))_+$ at time t (where for any $x \in \mathbb{R}$, $x_+ := \max(0, x)$) and corresponds to a flip of the i-th momentum variable p_i. In practice the straightforward way to simulate such a dynamics is to draw for any $i \in [\![1, D]\!]$ a jump time $\tau_i \geqslant 0$ from the probability law with cumulative distribution:

$$\mathbb{P}[\tau_i > t] = \exp\Big(- \int_0^t (\partial_i V(q_0 + s p_0))_+ \mathrm{d}s \Big), \tag{1.8}$$

where q_0, p_0 is the state right after the last jump event. The next jump corresponds to the smallest time τ_i drawn, and it is a flip of the i-th coordinate of the momentum. Note that all jump times τ_i have to be redrawn after each jump. The jump time are drawn from (1.8) using a type of rejection method called Poisson thinning [107]. A jump time $\widetilde{\tau}_i$ is proposed using a law with cumulative distribution

$$\mathbb{P}[\widetilde{\tau}_i > t] = \exp\Big(- \int_0^t M_i(s) \mathrm{d}s \Big),$$

which can be sampled exactly for a certain class of functions M_i, and an accept/reject ratio is computed. This method can be optimized [16] by choosing properly the proposal function M_i, since its efficiency relies on the sharpness of the computational bound $(\partial_i V(q_0 + t p_0))_+ \leqslant M_i(t)$ on the instantaneous jump rate. The random times are drawn under the correct probability law without bias so that the full process can be simulated exactly, contrarily to what is generally the case for the SDEs. This is the major appeal of this method. The generator of this process writes

$$\mathcal{L}\varphi(q, p) = p \cdot \nabla_q \varphi(q, p) + \sum_{i=1}^{D} (\partial_i V(q))_+ \left(\varphi(q, p - 2 p_i e_i) - \varphi(q, p) \right),$$

where $(e_i)_{1 \leqslant i \leqslant D}$ is the canonical basis of $\mathbb{R}^D$.

The Zig-Zag process, the Bouncy Particle Sampler and Hybrid Monte Carlo admit an invariant probability measure with marginal distribution ν in the position variable, and it is used to sample the latter. The theoretical study of PDMPs is however hindered by technicalities due to the lack of regularization properties of the generator. We refer to [47] for a rigorous definition of PDMPs and a proof of pathwise ergodicity under adequate

assumptions. In the following we focus on SDEs, and more specifically on the Langevin dynamics, to avoid such issues.

1.2 Equilibrium Langevin dynamics

Many properties can be rigorously established for stochastic differential equations, in contrast to deterministic continuous dynamics or discrete Markov chains. We focus on the case of the Langevin dynamics which is of particular interest for sampling, and for which the analysis is particularly rich. The concepts introduced here are however relevant for general SDEs. We recall in Section 1.2.1 a practical criterion for ergodicity, then we define in Section 1.2.2 the evolution semi-group and review the most common techniques to prove its exponential decay in various functional frameworks. The latter decay estimates are a key element to establish that a Central Limit Theorem holds for ergodic means (see Section 1.2.3), justifying the use of confidence intervals in practice. Finally Section 1.2.4 recalls some concepts of numerical analysis for SDEs, and introduces two standard numerical schemes for the time integration of Langevin dynamics.

1.2.1 Ergodicity

The convergence of ergodic averages (1.2) is the fundamental basis for trajectorial averaging. It does not hold in general for deterministic dynamics, which is why we often prefer SDEs. Such stochastic processes can be proved to be ergodic using the following theorem, which a result from [89].

Proposition 1.1. *Considering a Markov process $(X_t)_{t\geqslant 0}$ on $\mathcal{E}$ such that*

- *the probability measure π is invariant for the stochastic process $(X_t)_t$: for any smooth observable φ,*
$$\forall t \geqslant 0, \quad \mathbb{E}[\varphi(X_t)|X_0 \sim \pi] = \mathbb{E}_\pi[\varphi];$$
- *the probability measure μ admits a positive density with respect to the Lebesgue measure, and the generator $\mathcal{L}$ is hypoelliptic: there exists $\varepsilon > 0$ such that, for any $s \in \geqslant 0$, $\mathcal{L}\varphi \in \mathrm{H}^s_{\mathrm{loc}}$ implies $\varphi \in \mathrm{H}^{s+\varepsilon}_{\mathrm{loc}}$.*

Then the process is pathwise ergodic: for any bounded measurable function φ and any given initial condition $X_0 = x_0$,

$$\frac{1}{T}\int_0^T \varphi(X_t)\,\mathrm{d}t \xrightarrow[t\to\infty]{} \mathbb{E}_\pi[\varphi] \quad \text{a.s..}$$

For a general SDE, it can be checked that the probability measure π is invariant by showing that

$$\mathcal{L}^\dagger \pi = 0,$$

where $\mathcal{L}^\dagger$ denotes the adjoint of $\mathcal{L}$ in $\mathrm{L}^2(\mathrm{d}q\,\mathrm{d}p)$. This condition is equivalent to: for any $\mathcal{C}^\infty$ function with compact support φ, $\int_{\mathcal{E}} \mathcal{L}\varphi\,\mathrm{d}\pi = 0$. For some dynamics, for example out of

equilibrium, it is not possible to exhibit an analytical invariant probability measure. In this case the existence of a unique invariant probability measure can be proven using Lyapunov techniques [74, 113, 138], see for example Chapter 3.

The second condition is satisfied for the overdamped Langevin dynamics, since its invariant probability measures ν admit a density with respect to the Lebesgue measure and its generators is elliptic, thus hypoelliptic. This condition is also satisfied for the Langevin dynamics. The hypoellipticity of the generator can be proven using Hörmander theorem [83].

The second assumption can be replaced by the assumption of aperiodic irreducibility: for any measurable set $A \subset \mathcal{E}$ such that $\pi(A) > 0$ and π-almost all initial condition $x_0 \in \mathcal{E}$,

$$\exists t_0 \geqslant 0,\ \forall t \geqslant t_0, \quad \mathbb{P}[X_t \in A | X_0 = x_0] > 0.$$

This property is often proved in two steps [137]. First the accessibility is proven for any open set A by constructing a control allowing to reach a point of A. Second we conclude using the regularity of the process.

1.2.2 Exponential decay of the semi-group

We just proved under some assumptions that ergodic averages converge to a mean value determined by the invariant probability measure μ. The rate of convergence of the process towards the equilibrium defined by μ can be estimated. This convergence can be quantified by looking at the decay of the evolution semi-group of the process, these concepts being explained in Subsection 1.2.2.1. We review some existing methods to prove that the decay is exponential for several metrics, for reversible dynamics in Subsection 1.2.2.2 and for non-reversible dynamics in Subsection 1.2.2.3.

1.2.2.1 Semi-group and Fokker-Planck equation

Consider a stochastic process with initial probability distribution $(q_0, p_0) \sim \rho_0 \in \mathcal{P}(\mathcal{E})$ and define the probability distribution at time t

$$\rho_t := \mathrm{Law}((q_t, p_t) \,|\, (q_0, p_0) \sim \rho_0).$$

Then ρ_t satisfies the Kolmogorov forward, or Fokker-Planck equation [139, 137]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t = \mathcal{L}^\dagger \rho_t.$$

Equivalently,

$$\rho_t = \mathrm{e}^{t\mathcal{L}^\dagger}\rho_0.$$

provided the semi-group generated by $\mathcal{L}^\dagger$ is well defined. The ergodicity property suggests that $\rho_t \xrightarrow[t\to\infty]{} \mu$ in some functional space. The algebra is in fact more simple from a dual point of view. We can indeed consider, for any observable φ_0,

$$\varphi_t(q,p) := \mathbb{E}[\varphi_0(q_t, p_t) \,|\, (q_0, p_0) = (q, p)],$$

which satisfies (see (1.4))

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t = \mathcal{L}\varphi_t.$$

Equivalently, $\varphi_t = \mathrm{e}^{t\mathcal{L}}\varphi_0$. We expect that $\varphi_t \xrightarrow[t\to\infty]{} \mathbb{E}_\mu[\varphi]$. We can prove that this convergence indeed happens at an exponential rate for several metrics such as relative entropies, total variation, $\mathrm{H}^1(\mu)$ or $\mathrm{L}^2(\mu)$ distances [104]. Note that since constants are left invariant by the semi-group, it holds for any test function φ,

$$\mathrm{e}^{t\mathcal{L}}\varphi - \mathbb{E}_\mu[\varphi] = \mathrm{e}^{t\mathcal{L}}(\varphi - \mathbb{E}_\mu[\varphi]).$$

Studying the convergence of $\varphi_t = \mathrm{e}^{t\mathcal{L}}\varphi$ towards $\mathbb{E}_\mu[\varphi]$ is therefore equivalent to studying the convergence to 0 of $\mathrm{e}^{t\mathcal{L}}\varphi$ for a function φ with mean zero. This is why we work in the Banach space

$$\mathrm{L}^1_0(\mu) = \left\{ \varphi \in \mathrm{L}^1(\mu) \;\middle|\; \int_{\mathcal{E}} \varphi \,\mathrm{d}\mu = 0 \right\}.$$

An exponential convergence result on the Banach space $\mathrm{X} \subset \mathrm{L}^1(\mu)$ has the following form: there exist $C, \kappa > 0$ such that, for any $\varphi \in \mathrm{X} \cap \mathrm{L}^1_0(\mu)$,

$$\|\mathrm{e}^{t\mathcal{L}}\varphi\|_{\mathrm{X}} \leqslant C\mathrm{e}^{-t\kappa}\|\varphi\|_{\mathrm{X}}. \tag{1.9}$$

This inequality holds for $C = 1$ when the generator is coercive, see the case of reversible dynamics. On the contrary, for non-reversible dynamics, the typical case is when the decay only holds for some $C > 1$. When X is a Hilbert space the generator is then said to be hypocoercive. The exponential decay of the semi-group implies the invertibility of the generator on $\mathrm{X} \cap \mathrm{L}^1_0(\mu)$:

$$\mathcal{L}^{-1} = -\int_0^\infty \mathrm{e}^{t\mathcal{L}} \,\mathrm{d}t\,, \quad \text{with} \quad \left\|\mathcal{L}^{-1}\right\|_{\mathcal{B}(X\cap \mathrm{L}^1_0(\mu))} \leqslant \frac{C}{\kappa},$$

where $\|A\|_{\mathcal{B}(X)} = \sup_{\|f\|_X=1} \|Af\|_X$ denotes the operator norm of the operator A defined on the Banach space X. The invertibility of the generator implies the existence of a spectral gap:

$$\tau := \inf\left\{\mathscr{R}(\lambda) \,|\, \lambda \in \mathrm{Sp}(-\mathcal{L})) \setminus \{0\}\right\} > 0,$$

where $\mathscr{R}(\lambda)$ denotes the real part of $\lambda \in \mathbb{C}$. Note that the converse implication is not true, and we refer to [51, Chapter 3] for further details. Using that the semi-group $\mathrm{e}^{t\mathcal{L}^\dagger}$ involved in the evolution of the probability measure ρ_t is the adjoint of the semi-group $\mathrm{e}^{t\mathcal{L}}$ involved in the evolution of the probability measure φ_t, it suffices to study the convergence of the latter to understand the long-time behavior of the dynamics.

1.2.2.2 Reversible case

We first consider the case of the overdamped Langevin dynamics, which is a reversible diffusion equation. We show how the Brownian motion acting on all variables can be used to prove exponential decay in this case. This serves as a starting point in the next section

to understand the Langevin dynamics, where the Brownian motion only acts on half the variables.

We show exponential decay in the space $\mathrm{L}^2(\nu)$, and more precisely on the hyperplane $\mathrm{L}^2_0(\nu)$ of $\mathrm{L}^2(\nu)$ composed of functions with mean zero with respect to ν.

We say that the measure ν satisfies a Poincaré inequality with constant R if for any $\varphi \in \mathrm{H}^1(\nu) \cap \mathrm{L}^2_0(\nu)$,

$$\|\varphi^2\| \leqslant \frac{1}{R}\|\nabla\varphi\|^2,$$

where the norm is taken in $\mathrm{L}^2(\nu)$. We refer to [9] for conditions implying such a property. In particular the Poincaré inequality holds when V is the sum of a uniformly convex [10] and a bounded function [82].

Proposition 1.2. *The measure ν satisfies a Poincaré inequality with constant R if and only if*

$$\forall \varphi \in \mathrm{L}^2_0(\nu), \quad \left\|\mathrm{e}^{t\mathcal{L}}\varphi\right\| \leqslant \mathrm{e}^{-R\beta^{-1}t}\|\varphi\|.$$

In this case the generator is invertible on $\mathrm{L}^2_0(\nu)$ and

$$\|\mathcal{L}^{-1}\|_{\mathcal{B}(\mathrm{L}^2_0(\nu))} \leqslant \frac{\beta}{R}.$$

Proof. Assume a Poincaré inequality and take $\varphi \in \mathcal{C}^\infty$ with compact support and mean zero. An integration by part provides

$$-\langle \mathcal{L}\varphi, \varphi\rangle_{\mathrm{L}^2(\nu)} = -\left\langle (-\nabla V \cdot \nabla_q + \beta^{-1}\Delta_q)\varphi, \varphi\right\rangle_{\mathrm{L}^2(\nu)} = \beta^{-1}\|\nabla_q\varphi\|^2 \geqslant R\beta^{-1}\|\varphi\|^2, \quad (1.10)$$

from which we deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\|\mathrm{e}^{t\mathcal{L}}\varphi\|^2\right) = 2\left\langle \mathrm{e}^{t\mathcal{L}}\varphi, \mathcal{L}\mathrm{e}^{t\mathcal{L}}\varphi\right\rangle_{\mathrm{L}^2(\nu)} \leqslant -2R\beta^{-1}\|\mathrm{e}^{t\mathcal{L}}\varphi\|^2.$$

We conclude to the exponential decay by the Gronwall lemma. This inequality can be extended to $\varphi \in \mathrm{L}^2_0(\nu)$ by density.

Conversely, assuming the exponential decay, the following inequality holds for any $\varphi \in \mathrm{L}^2_0(\nu)$:

$$\frac{\|\mathrm{e}^{t\mathcal{L}}\varphi\|^2 - \|\varphi\|^2}{t} \leqslant \|\varphi\|^2\frac{\mathrm{e}^{-2R\beta^{-1}t} - 1}{t}.$$

Taking the limit as $t \to 0$, the Poincaré inequality follows by using the two previous equations for $\varphi \in \mathcal{C}^\infty$ with compact support and mean zero. □

In this proof we used crucially the coercivity of the operator $\mathcal{L}$ in $\mathrm{L}^2_0(\nu)$ (1.10). However this property does not hold for the generators involved in the kinetic equations such as Langevin or the PDMPs.

1.2.2.3 Non-reversible case

It is numerically observed that kinetic dynamics such as the Langevin dynamics or Hamiltonian Monte-Carlo perform better than the Metropolis Algorithm or the overdamped Langevin dynamics. This is not yet completely understood from a theoretical perspective, but explicit spectral gap estimates is a step toward this direction. Proving the existence of a spectral gap is however more involved in the non-reversible case, and not all methods provide quantitative estimates of the decay rate.

For Langevin dynamics, the Dirichlet form associated to the generator writes

$$\forall \varphi \in \mathscr{C}, \quad -\langle \varphi, \mathcal{L}\varphi \rangle_{\mathrm{L}^2(\mu)} = \beta^{-1} \|\nabla_p \varphi\|^2.$$

In particular it vanishes for observables depending only on the position. The generator is not elliptic since second order derivatives in q are missing. The operator $-\mathcal{L}$ is therefore not coercive for the canonical scalar products in $\mathrm{L}^2(\mu)$ or subspaces of it. However we can prove the exponential decay of the semi-group and the invertibility of $\mathcal{L}$. Note that (1.9) cannot hold with $C = 1$, otherwise by a reasoning similar to the proof of Proposition 1.2, one could prove that $-\mathcal{L}$ is coercive on $\mathrm{X} \cap \mathrm{L}^2_0(\mu)$.

We briefly describe several types of proof, and compare their advantages. When the dynamics has a lot of structure, like the Langevin equation we consider here, we can rely on hypocoercivity techniques. In this case a scalar product equivalent to the canonical one introduces some mixed derivatives in q and p in order to retrieve some dissipation in q. On the other hand, approaches relying on a Lyapunov function apply to a wider set of dynamics but provide estimates for the decay rates whose dependence with the parameters of the problem is less explicit. These estimates can however be greatly improved using coupling techniques.

We are particularly interested in the scaling of the spectral gap with respect to two parameters. First we expect the spectral gap to scale with the friction γ like $\min(\gamma, \gamma^{-1})$, since it is the scaling which is derived in two analytical cases: in absence of potential ($V = 0$) on the torus ($\mathcal{D} = \mathbb{T}$) [95], and for a quadratic potential ($V(q) = |q|^2$) [111] in $\mathcal{D} = \mathbb{R}^D$. Good quantitative estimates should therefore have the same dependence with respect to γ. Qualitatively the spectral gap vanishes in the small γ limit because the energy varies on a slow time scale γ^{-1}, see Chapter 4 for further information. In the large γ limit, the convergence of the process rescaled in time by a factor γ to the overdamped dynamics (see Section 1.1.3.3) shows that the characteristic time is of order γ in this regime [121]. Second, when sampling in high dimension, the behavior of the convergence rate with respect to the dimension D is crucial. For a class of potentials satisfying some properties uniformly with D, the generator of the Langevin dynamics satisfies a spectral gap of size independent of D (see Chapter 5 for example).

Estimates in $\mathrm{H}^1(\mu)$. The first result of exponential decay was obtained in [153], introducing the idea of mixed derivatives, and generalized in [166] who proposed to replace the canonical scalar product by an equivalent one. This technique allows to show the following.

Proposition 1.3. *Assume that a Poincaré inequality holds for the measure ν and that there exists $\rho > 0$ such that*

$$\forall q \in \mathcal{D}, \qquad |\nabla^2 V(q)| \leqslant \rho(1 + |\nabla V(q)|).$$

Then there exist $C, \kappa > 0$ such that, for any $\varphi \in \mathrm{H}^1(\mu) \cap \mathrm{L}^2_0(\mu)$,

$$\forall t \geqslant 0, \quad \left\| \mathrm{e}^{t\mathcal{L}} \varphi \right\|_{\mathrm{H}^1(\mu)} \leqslant C \mathrm{e}^{-t\kappa} \|\varphi\|_{\mathrm{H}^1(\mu)}.$$

Moreover there exists $\bar{\kappa} > 0$ independent of γ such that $\kappa \geqslant \bar{\kappa} \min(\gamma, \gamma^{-1})$.

The proof of Proposition 1.3 relies on the fact that for any $a, b, c \in \mathbb{R}$ such that $a, c > 0$ and $ac - b^2 > 0$, the scalar products on $\mathrm{H}^1(\mu)$ defined by

$$\begin{aligned} \langle\langle \varphi, \psi \rangle\rangle = \langle \varphi, \psi \rangle_{\mathrm{L}^2(\mu)} + a \langle \nabla_p \varphi, \nabla_p \psi \rangle_{\mathrm{L}^2(\mu)} \\ - b \langle \nabla_p \varphi, \nabla_q \psi \rangle_{\mathrm{L}^2(\mu)} - b \langle \nabla_q \varphi, \nabla_p \psi \rangle_{\mathrm{L}^2(\mu)} + c \langle \nabla_q \varphi, \nabla_q \psi \rangle_{\mathrm{L}^2(\mu)}, \end{aligned}$$

are equivalent to the canonical scalar product (which corresponds to $a = c = 1$ and $b = 0$). Although the operator $-\mathcal{L}$ is not coercive on $\mathrm{H}^1(\mu)$ for the canonical scalar product, under the assumptions of Proposition 1.3 one can show that the coercivity property holds for some choice of a, b, c. This motivates the name hypocoercivity. The exponential decay for the norm induced by the equivalent scalar product follows by applying Gronwall's lemma, with a decay rate equal to the coercivity constant of $-\mathcal{L}$. One can conclude to the exponential decay in the canonical norm using the norm equivalence. The method provides an explicit bound on the decay rate: $\kappa \geqslant \bar{\kappa} \min(\gamma, \gamma^{-1})$, which is the expected scaling . The scaling with the dimension can also be estimated, similarly to what is done in Chapter 5.

Hypoelliptic regularization techniques, introduced in [79] and popularized in [75, 166], allow to deduce exponential decay in $\mathrm{L}^2(\mu)$ from the previous result. We present instead a more direct route to establish hypocoercivity results in $\mathrm{L}^2(\mu)$, as proposed in [44].

Estimates in $\mathrm{L}^2(\mu)$. Similarly to the previous approach, one constructs a scalar product on $\mathrm{L}^2(\mu)$ equivalent to the canonical one $\langle \cdot, \cdot \rangle_{\mathrm{L}^2(\mu)}$. Showing that the operator $-\mathcal{L}$ is coercive on $\mathrm{L}^2_0(\mu)$ for this equivalent scalar product, one can conclude to the exponential decay in $\mathrm{L}^2(\mu)$. We use this method of proof in Chapters 2 and 5, and we refer to the former for a detailed description of the proof in the case of Langevin dynamics.

This method, as well as the previous one, provides quantitative estimates on the spectral gap for the Langevin dynamics. The scaling of the estimates with the friction γ is correctly captured, and the scaling with the dimension D can be made explicit. We refer to Chapter 5 for the results. The spectral gap estimates can be extended to small perturbations of the Langevin equation [85], which is useful when studying the linear response to a nonequilibrium forcing. This is relevant in the context of the estimation of transport coefficient, which is the focus of Section 1.3. It can also be extended to the case of sub-exponential convergence for heavy tailed distributions [33].

Lyapunov approach. Lyapunov techniques allow to prove decay estimates in weighted L^∞ spaces. The early reference is [113], see also [137] for a presentation from a functional analysis viewpoint. We follow here the presentation of [74]. It is a versatile tool, as it applies to both Markov Chains and SDEs and it requires less structure than the previous approaches. In particular it allows to prove the existence of an invariant probability distribution as well as exponential decay of the semi-group for arbitrary large nonequilibrium perturbations of equilibrium dynamics, as done in Chapter 3.

The first assumption, called the Lyapunov condition (or Foster-Lyapunov drift), is that there exist a function $\mathcal{K} : \mathcal{E} \to [1, +\infty)$ and constants $a > 0, b \in \mathbb{R}$ such that

$$\mathcal{L}\mathcal{K} \leqslant -a\mathcal{K} + b. \tag{1.11}$$

The Lyapunov function $\mathcal{K}$ typically goes to infinity at infinity, so that the condition implies that the dynamics returns to regions of the phase space where $\mathcal{K}$ is not too large. We denote one of these regions by $\mathcal{C} = \{x \in \mathcal{E} \,|\, \mathcal{K}(x) \leqslant \mathcal{K}_{\max}(a, b)\}$, for some constant $\mathcal{K}_{\max}$ depending on the previous constants a and b.

The second assumption is a minorization condition (or Doeblin condition) on compact subsets: for any $t_0 > 0$, there exist a constant $a > 0$ and a probability measure λ such that

$$\inf_{x \in \mathcal{C}} \mathcal{P}_{t_0}(x, \mathrm{d}y) \geqslant \alpha\lambda(\mathrm{d}y), \tag{1.12}$$

where $\mathcal{P}_{t_0}$ is the evolution kernel over a time t_0. This kernel can be related to the semi-group using the following formula for $\varphi \in \mathcal{C}^\infty$ with compact support and $x \in \mathcal{E}$:

$$\left(\mathrm{e}^{t_0\mathcal{L}}\varphi\right)(x) = \int_{\mathcal{E}} \varphi(y)\mathcal{P}_{t_0}(x, \mathrm{d}y).$$

The existence of this kernel can be proved for SDEs using the irreducibility of the process (see Proposition 1.1) and the regularization properties (hypoellipticity) of the Fokker-Planck equation. The latter follows from Hörmander's theorem [83] which holds in particular for Langevin equation.

The two previous assumptions imply the existence of a unique invariant probability measure, as well as the exponential decay of the semi-group in the weighted L^∞ norm defined by

$$\|\varphi\|_{\mathrm{L}^\infty_{\mathcal{K}}} := \left\|\frac{\varphi}{\mathcal{K}}\right\|_{\mathrm{L}^\infty}.$$

More precisely, there exist $C, \kappa > 0$ such that for any $\varphi \in \mathrm{L}^\infty_{\mathcal{K}} \cap \mathrm{L}^2_0(\mu)$,

$$\forall t \geqslant 0, \quad \left\|\mathrm{e}^{t\mathcal{L}}\varphi\right\|_{\mathrm{L}^\infty_{\mathcal{K}}} \leqslant C\mathrm{e}^{-\kappa t}.$$

Note that, if $\mathcal{K} \in \mathrm{L}^2(\mu)$, then for any $\varphi \in \mathrm{L}^\infty_{\mathcal{K}}$,

$$\|\varphi\|_2 \leqslant \left\|\frac{\varphi}{\mathcal{K}}\right\|_{\mathrm{L}^\infty} \|\mathcal{K}\|_2,$$

so that in this case $\mathrm{L}^\infty_{\mathcal{K}} \subset \mathrm{L}^2(\mu)$.

For the Langevin dynamics, a Lyapunov function has been proposed in [153, 109]:

$$\mathcal{K}(q,p) = H(q,p) + \frac{\gamma}{2} p^\top q + \frac{\gamma^2}{4}|q|^2 + 1. \tag{1.13}$$

Under some assumptions on the growth of the potential V at infinity, we can show that the conditions (1.11) and (1.12) hold. Lyapunov functions with dominant term $\mathrm{e}^{\theta(H+p^\top q)}$ for some $\theta > 0$, plus some correction term, has been proposed in [81, 171]. They diverge faster to infinity than the Lyapunov function (1.13), allowing to prove the exponential decay of the semi-group on a larger space $\mathrm{L}^i_{\mathcal{K}} nfty$.

Coupling approach. In the previous Lyapunov approach the minimization condition is generally granted by qualitative arguments, leading to non-explicit spectral gap estimates. In particular the dependence of the decay rate with respect to the friction γ is not clear, contrarily to the results obtained with hypocoercive approaches. Coupling approaches aim at characterizing more explicitly the mixing properties of the dynamics, relying on a probabilistic viewpoint. The minorization condition is in fact equivalent [46] to: there exist $t > 0$, $\alpha > 0$ such that for all $x, \tilde{x} \in \mathcal{C}$,

$$\|\mathcal{P}_t(x,\cdot) - \mathcal{P}_t(\tilde{x},\cdot)\|_{\mathrm{TV}} \leqslant 2(1-\alpha). \tag{1.14}$$

Here $\|\cdot\|_{\mathrm{TV}}$ denotes the norm in total variation, defined for two probability measures $\lambda, \widetilde{\lambda}$ as

$$\|\lambda - \widetilde{\lambda}\|_{\mathrm{TV}} = \sup_{A \text{ measurable}} |\lambda(A) - \widetilde{\lambda}(A)|.$$

The condition (1.14) can be proved using probabilistic techniques, and quantitative values of α can be obtained. Given a solution (x_t) of the SDE for a realization (W_t) of a Brownian motion, the method relies on the construction of a solution $(\widetilde{x}_t)$ for the Brownian motion $(\widetilde{W}_t)$ related to W_t and x_t. The coupling procedure consists in defining this second Brownian motion in such a way that the two trajectories are coupled. The type of coupling depends on the equation. For example a synchronous coupling $\widetilde{W}_t = W_t$ works for the overdamped Langevin dynamics with a convex potential V but fails in most cases. A sticky coupling is proposed in [46] in the case of the Bouncy Particle Sampler, whereas in [48] a combination of synchronous and reflection couplings is considered for the Langevin equation. In this second case the method provides explicit scalings of the spectral gap estimates. In particular the bound on the decay rate satisfies $\kappa \geqslant \bar{\kappa} \min(\gamma, \gamma^{-1})$ for $\bar{\kappa} > 0$ independent of γ.

1.2.3 Central Limit Theorem

The convergence estimates for the semi-group obtained in Section 1.2.2 imply convergence rates for ergodic averages through a Central Limit Theorem. For a given observable φ let

us recall that the ergodic mean at time t, namely

$$\widehat{\varphi}_t := \frac{1}{t}\int_0^t \varphi(q_t, p_t)\ \mathrm{d}t\,,$$

converges to $\mathbb{E}_\mu[\varphi]$ in the large time limit under the assumptions of Proposition 1.1. The asymptotic variance of the observable $\varphi \in \mathrm{L}^2(\mu)$ is defined by

$$\sigma_\varphi^2 := \lim_{t\to\infty} t\mathrm{Var}[\widehat{\varphi}_t],$$

where the variance is taken over realizations of the process and over initial conditions $x_0 \sim \mu$. Assuming that the Poisson problem

$$-\mathcal{L}\Phi = \varphi - \mathbb{E}_\mu[\varphi],$$

admits a solution $\Phi \in \mathrm{L}^2(\mu)$, a Central Limit Theorem holds:

$$\sqrt{t}\,(\widehat{\varphi}_t - \mathbb{E}_\mu[\varphi]) \xrightarrow[t\to\infty]{\text{law}} \mathcal{N}(0, \sigma_\varphi^2),$$

where the asymptotic variance σ_φ^2 is well defined and given by

$$\sigma_\varphi^2 = 2\,\langle \Phi, \varphi\rangle_{\mathrm{L}^2(\mu)}\,. \tag{1.15}$$

This result can be extended to initial conditions which are not distributed according to the invariant probability measure [15]. This expression for the asymptotic variance comes from the following computation: for any φ such that $\mathbb{E}_\mu[\varphi] = 0$,

$$t\mathrm{Var}[\widehat{\varphi}_t] = t^{-1}\mathbb{E}\left[\left(\int_0^t \varphi(q_s, p_s)\mathrm{d}s\right)^2\right] = 2\int_0^t \left(1 - \frac{s}{t}\right)\left\langle \mathrm{e}^{s\mathcal{L}}\varphi\,\varphi\right\rangle_{\mathrm{L}^2(\mu)} \mathrm{d}s. \tag{1.16}$$

Formally, equation (1.15) is obtained by sending t to infinity and using that $\int_0^\infty \mathrm{e}^{t\mathcal{L}}\,\mathrm{d}t = -\mathcal{L}^{-1}$.

Assuming that the semi-group decays exponentially in $\mathrm{L}^2_0(\mu)$, then the generator $\mathcal{L}$ is invertible on this space and a Central Limit Theorem holds for any $\varphi \in \mathrm{L}^2(\mu)$. Moreover the asymptotic variance is bounded:

$$0 \leqslant \sigma_\varphi^2 = 2\left\langle -\mathcal{L}^{-1}(\varphi - \mathbb{E}_\mu[\varphi]), \varphi\right\rangle_{\mathrm{L}^2(\mu)} \leqslant 2\left\|\mathcal{L}^{-1}\right\|_{\mathcal{B}(\mathrm{L}^2_0(\mu))}\ \|\varphi\|^2.$$

Spectral gap estimates in $\mathrm{L}^2(\mu)$ are thus of prime interest since they imply confidence intervals for ergodic averages.

1.2.4 Numerical integration

In practice the solution to a stochastic differential equation (1.3) is not analytical in general, so that it needs to be approximated using a numerical integration scheme. The simplest

scheme for the general SDE (1.3) is the Euler-Maruyama discretization with time step $\Delta t > 0$, which writes

$$x^{n+1} = x^n + b(x^n)\,\Delta t + \sigma(x^n)\,\sqrt{\Delta t}\,G^n,$$

where $(G^n)_{n\geqslant 0}$ is a sequence of independent Gaussian variables with covariance matrix $\mathbf{I}_D$, and $x^0 \in \mathcal{E}$ is a given initial condition. This procedure defines a sequence of points $(x^n)_{n\geqslant 0}$ such that x^n approximates the exact solution $x_{n\Delta t}$.

The discretization error is either measured on finite time intervals, using weak or strong errors, or in terms of the difference between the ergodic invariant probability measures. We refer to [90], [132] and [116] for more general reviews on numerical analysis for SDEs. In practice the systems we consider are chaotic, so that it is impossible to simulate precisely the trajectory over long times. This is also useless since the initial conditions are arbitrary. This is why we put emphasis on the third type of error which does not measure the accuracy of the trajectories generated but rather their sampling properties. Note that weak error is relevant when studying dynamical properties of trajectories.

Weak error estimates. The integration scheme is of weak order $\alpha \in \mathbb{R}_+$ if, for any function $\varphi \in \mathrm{C}^\infty$ with compact support, time horizon T and initial condition x_{in}, there exist $C > 0$ and $\Delta t^* > 0$ such that, for all $0 < \Delta t \leqslant \Delta t^*$,

$$\sup_{0\leqslant n\leqslant T/\Delta t} \Big|\mathbb{E}[\varphi(x^n)|x^0 = x_{\mathrm{in}}] - \mathbb{E}[\varphi(x_{n\Delta t})|x_0 = x_{\mathrm{in}}]\Big| \leqslant C\Delta t^\alpha.$$

The weak order measures the error between the means.

Strong error estimates in L^p-norm. Set $p \geqslant 1$. The integration scheme is of strong order $\alpha \in \mathbb{R}_+$ if, for any time horizon T and initial condition x_0, there exist $C > 0$ and $\Delta t^* > 0$ such that, for any $0 < \Delta t \leqslant \Delta t^*$,

$$\sup_{0\leqslant n\leqslant T/\Delta t} \mathbb{E}\Big[|x^n - x_{n\Delta t}|^p\Big]^{1/p} \leqslant C\Delta t^\alpha.$$

Note that $x_{n\Delta t}$ and its discretized counterpart x^n correspond to the same realization of the Brownian motion in the sense that $G^n = \frac{W_{(n+1)\Delta t} - W_{n\Delta t}}{\sqrt{\Delta t}}$. The strong order measures the rate of convergence of the mean of the difference between the trajectories. It is indeed a stronger type of convergence since a scheme of strong order α is automatically of weak order α.

Long-time error. We are only interested in the invariant probability measure when estimating static properties. The long time error allows to quantify if this measure is well preserved by the numerical scheme. Before defining the systematic sampling bias associated to the numerical scheme, one first needs to show that the discrete trajectory $(x^n)_{n\geqslant 0}$ is ergodic. This requires to prove the irreducibility of the Markov chain and the existence of an invariant probability measure $\mu_{\Delta t}$. This is typically done by showing a Lyapunov and a minorization condition for a properly chosen Lyapunov function [102]. When the domain is non compact and the potential is not globally Lipschitz, explicit schemes such as Euler-

Maruyama may not be stable and the ergodicity may fail due to transient behaviors [109]. It may be necessary to resort to implicit schemes in this situation [94].

When $(x^n)_{n\geqslant 0}$ is ergodic with invariant probability measure $\mu_{\Delta t}$, the integration scheme is of order α if for any function $\varphi \in \mathrm{C}^\infty$ with compact support, there exist $C > 0$ and $\Delta t^* > 0$ such that, for any $0 < \Delta t \leqslant \Delta t^*$,

$$\left| \int_{\mathcal{E}} \varphi \, \mathrm{d}\mu_{\Delta t} - \int_{\mathcal{E}} \varphi \, \mathrm{d}\mu \right| \leqslant C \Delta t^\alpha. \tag{1.17}$$

This exponent α is larger than the order of the weak error, when the ergodicity assumption is satisfied. The error (1.17) can be expanded in powers of Δt, and an expression can be given for the dominant coefficient [154, 102], yet impossible to compute in practical situations. The total sampling error for finite integration times is therefore the sum of two errors, a finite time-step bias and a finite time statistical error: for $x^0 \sim \mu_{\Delta t}$,

$$\frac{1}{n}\sum_{i=1}^n \varphi(x^n) = \int_{\mathcal{E}} \varphi \, \mathrm{d}\mu + \underbrace{\int_{\mathcal{E}} \varphi \, \mathrm{d}\mu_{\Delta t} - \int_{\mathcal{E}} \varphi \, \mathrm{d}\mu}_{\text{bias } \mathrm{O}(\Delta \mathrm{t}^\alpha)} + \underbrace{\frac{1}{n}\sum_{i=1}^n \varphi(x^n) - \int_{\mathcal{E}} \varphi \, \mathrm{d}\mu_{\Delta t}}_{\text{statistical error } \mathrm{O}((\mathrm{n}\Delta \mathrm{t})^{-1/2})}.$$

In typical cases the statistical error dominates the bias, though for finely converged simulations there is a trade-off on Δt to minimize the total error for a given computational budget n.

Classical schemes. The Euler-Maruyama scheme is of strong order 1/2, weak order 1 and it corresponds to an approximate invariant probability measure which is correct at order 1. This scheme is very simple and works for a large variety of SDEs, but it generates rather large integration errors which require resorting to small time steps Δt.

Higher order schemes are usually designed by a splitting method. This consists in decomposing the dynamics into simpler dynamics which are successively integrated analytically. In the case of the Langevin a natural splitting is

$$\mathcal{L} = \mathcal{A} + \mathcal{B} + \mathcal{O},$$

where

$$\mathcal{A} = \frac{1}{m} p \cdot \nabla_q, \quad \mathcal{B} = -\nabla V \cdot \nabla_p, \quad \mathcal{O} = -\frac{\gamma}{m} p \cdot \nabla_p + \gamma \beta^{-1} \Delta_p.$$

The dynamics associated to each of the three generators can be integrated analytically. In other words the corresponding semi-group is analytic, so that the semi-group after a time Δt can be approximated by a product of explicit operators. For example the Trotter splitting writes

$$\mathrm{e}^{\Delta t \mathcal{L}} \approx \mathrm{e}^{\Delta t \mathcal{A}} \mathrm{e}^{\Delta t \mathcal{B}} \mathrm{e}^{\Delta t \mathcal{O}},$$

and it corresponds to a scheme where the parts of generator $\mathcal{O}$, $\mathcal{B}$ and $\mathcal{A}$ are successively integrated, in this order. This scheme can be shown to preserve the invariant probability measure with a first order accuracy.

More efficient schemes rely on higher order splitting such as the Strang splitting. We refer to [101, Section 7.3.1] for a precise introduction to splitting methods for the Langevin dynamics. A popular splitting scheme is the Geometric Langevin algorithm (GLA) [23], which corresponds to the splitting

$$\mathrm{e}^{\Delta t\mathcal{L}} \approx \mathrm{e}^{\Delta t\mathcal{O}}\mathrm{e}^{\Delta t/2\mathcal{A}}\mathrm{e}^{\Delta t\mathcal{B}}\mathrm{e}^{\Delta t/2\mathcal{A}},$$

of the semi-group. In more algorithmic terms, it writes as follows.

Algorithm 1.2 (Geometric Langevin algorithm)**.** *For a given initial configuration* (q^0, p^0)*, iterate on* $n \geqslant 0$*:*

$$\begin{aligned} p^{n+1/2} &= p^n - \nabla V(q^n)\frac{\Delta t}{2}, \\ q^{n+1} &= q^n + \frac{p^n}{m}\Delta t, \\ \tilde{p}^{n+1} &= p^{n+1/2} - \nabla V(q^{n+1})\frac{\Delta t}{2}, \\ p^{n+1} &= \alpha_{\Delta t}\tilde{p}^{n+1} + \sqrt{m\beta^{-1}(1-\alpha_{\Delta t}^2)}G^n, \end{aligned}$$

where $\alpha_{\Delta t} = \exp\Big(-\frac{\gamma}{m}\Delta t\Big)$.

The three first operations correspond to the integration of the Hamiltonian dynamics using the Velocity-Verlet scheme [164], while the last operation corresponds to the analytical integration of the Ornstein-Uhlenbeck process of generator $\mathcal{O}$.

This scheme requires one estimation of the forces ∇V per time step, as for the Euler-Maruyama scheme, so that they have similar computational costs. The GLA on the other hand preserves the invariant probability measure at second order in Δt, which allows to take significantly larger time steps.

1.3 Non-equilibrium Langevin dynamics

A nonequilibrium system can be modeled as an equilibrium system perturbed by an external field. The vast majority of systems which we encounter are not at thermodynamic equilibrium, since their macroscopic state evolve over time or can be triggered to do so. This is for example the case in life sciences, where processes under study are fundamentally non-reversible. All these systems are subject to fluxes of matter, electrical charge or energy, and sometimes chemical reactions. We refer for example to [52] for a review on non-equilibrium fluids.

We usually distinguish between weak and strong non-equilibrium perturbations, depending on whether the system is near equilibrium or not. The latter are not well understood, and few theoretical results exist. On the contrary linear response theory allows to study much easily small perturbations, and to characterize the regime of linear response. In this section we give some examples of non-equilibrium systems in Section 1.3.1, before defining transport coefficients in Section 1.3.2 and introducing linear response theory in Section 1.3.3 .

1.3.1 Non-equilibrium settings

Non-equilibrium systems are characterized by time irreversibility, which means that the arrow of time cannot be read off trajectories. A general stochastic differential equation

$$\mathrm{d}x_t = b(x_t)\,\mathrm{d}t + \sigma(x_t)\,\mathrm{d}W_t,$$

with unique invariant probability measure π is said to be reversible if for $x_0 \sim \pi$ the law of forward trajectories $(x_s)_{0\leqslant s\leqslant t}$ is the same as the law of backward trajectories $(x_{t-s})_{0\leqslant s\leqslant t}$ (note that $x_t \sim \pi$). This is equivalent to the self-adjointness of the generator $\mathcal{L}$ in $\mathrm{L}^2(\pi)$. This property is typically satisfied by diffusion equations such as the overdamped Langevin equation (1.7).

Kinetic equations such as the Langevin dynamics (1.5) or Piecewise Deterministic Markov Processes generate trajectories which are only reversible upon momentum reversal. Indeed, when $(q_0, p_0) \sim \mu$, the law of the forward paths $(q_s, p_s)_{0\leqslant s\leqslant t}$ is the same as the law of the backward paths $(q_{t-s}, -p_{t-s})_{0\leqslant s\leqslant t}$. Denoting by $\mathcal{R}$ the flip operator acting on smooth functions φ by $\mathcal{R}\varphi(q,p) = \varphi(q,-p)$ for any $(q,p) \in \mathcal{E}$, the adjoint of the generator in $\mathrm{L}^2(\pi)$ is self-adjoint up to a unitary transformation:

$$\mathcal{L}^* = \mathcal{R}\mathcal{L}\mathcal{R}.$$

In this case we still say that the system is at equilibrium, although it is not reversible.

There exist two classical ways to create a non-equilibrium system, which we illustrate with the following examples.

Non-gradient force. The first class of non-equilibrium systems involves an external force F which is not the gradient of a potential. This non-equilibrium forcing can be applied either to an overdamped Langevin dynamics:

$$\mathrm{d}q_t = (-\nabla V(q_t) + \eta F(q_t))\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}W_t,$$

where $\eta > 0$, or to a Langevin dynamics:

$$\begin{cases} \mathrm{d}q_t = \dfrac{1}{m} p_t\,\mathrm{d}t\,, \\ \mathrm{d}p_t = (-\nabla V(q_t) + \eta F(q_t))\,\mathrm{d}t - \dfrac{\gamma}{m} p_t\,\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\mathrm{d}W_t. \end{cases}$$

The external force typically generates a particle or mass flux in the direction given by the vector field F. The force field can be applied for example to a single tagged particle in an atomic fluid in a constant direction (see *e.g.* [143]). We refer to the discussion after (1.19) for more details on this dynamics. Another possibility is to push each half of the particle population in opposite directions [52, Section 6.2]. The external force can also mimic a shear, which allows to simulate a Couette flow [88, 159]. Such a shearing is considered in Section 3.5.

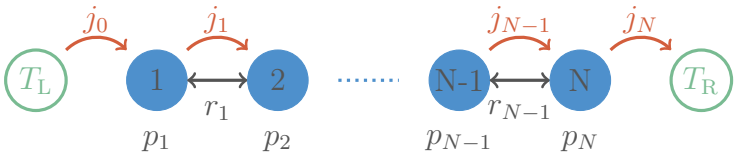


Figure 1.2: Heat transport in a one-dimensional chain.

Position dependent temperature. Langevin dynamics can also be perturbed by putting the system in contact with thermostats at temperature $T+\eta\Delta T(q)$ ($\beta^{-1} = k_{\mathrm{B}}T$ with $k_{\mathrm{B}} = 1$) depending on the position:

$$\begin{cases} \mathrm{d}q_t = \dfrac{1}{m} p_t \,\mathrm{d}t\,, \\ \mathrm{d}p_t = -\nabla V(q_t)\,\mathrm{d}t - \dfrac{\gamma}{m} p_t \,\mathrm{d}t + \sqrt{2\gamma(T+\eta\Delta T(q))}\mathrm{d}W_t, \end{cases}$$

where ΔT is a $\mathcal{C}^\infty$ function and $\eta > 0$ is such that $T + \eta\Delta T(q) \geqslant 0$ for any $q \in \mathcal{D}$. In this case the regions where $\Delta T(q) > 0$ are heated while the regions where $\Delta T(q) < 0$ are cooled. Physically, this triggers a heat flux from the hot source to the cold source. This type of simulation is used in practice to compute the thermal conductivity of a material.

An active field of research concerns heat transport in one-dimensional systems, see Section 3.4 for more details. This is modeled for instance as an atom chain, evolving under the following degenerate Langevin dynamics (see Figure 1.2):

$$\begin{cases} \mathrm{d}r_i = \dfrac{1}{m}(p_{i+1} - p_i)\,\mathrm{d}t, \\ \mathrm{d}p_1 = v'(r_1)\,\mathrm{d}t - \dfrac{\gamma}{m} p_1 \,\mathrm{d}t + \sqrt{2\gamma T_{\mathrm{L}}}\mathrm{d}W_t^L, \\ \mathrm{d}p_i = (v'(r_i) - v'(r_{i-1}))\,\mathrm{d}t, \\ \mathrm{d}p_N = -v'(r_{N-1})\,\mathrm{d}t - \dfrac{\gamma}{m} p_N \,\mathrm{d}t + \sqrt{2\gamma T_{\mathrm{R}}}\mathrm{d}W_t^R, \end{cases} \tag{1.18}$$

where $r_i = q_{i+1} - q_i$ represents the distance between two subsequent particles, v is a $\mathcal{C}^\infty$ pairwise potential and the potential energy writes $V(r) = \sum_{i=1}^{N-1} v(r_i)$. Both ends of the chain are in contact with a thermostats with temperature T_{L} and T_{R} such that $T_{\mathrm{L}} > T_{\mathrm{R}}$. The noise is therefore very degenerate since it only acts on the two variables p_1 and p_N. This dynamics corresponds to free boundary conditions at both ends, but other boundary conditions are sometimes preferred.

Numerical evidence show that thermal transport is sometimes anomalous, meaning that the asymptotic behavior in the macroscopic limit ($N \to \infty$) is not compatible with Fourier's law. The necessary microscopic ingredients, in terms of potential V or mass inhomogeneity along the chain, to observe (ab)normal thermal conductivity are still poorly understood. We refer to the review articles [20, 105] and [41] for more details.

1.3.2 Transport coefficients

The efficient numerical estimation of transport coefficients, such as mobility, thermal conductivity or shear viscosity is one of the main focus of this thesis. Transport coefficients relate a small external forcing (*e.g.* an electric field, a temperature gradient, a velocity field) to the average flux they induce (flux of electrical charges, energy, mass). The corresponding non-equilibrium steady state is a perturbation of the equilibrium Gibbs measure, but contrarily to the latter its expression is not given by a closed formula.

To concretely illustrate the above physical definitions, we focus on the computation of the mobility for Langevin dynamics on a domain with periodic boundary conditions $\mathcal{D} = \mathbb{T}^D$, for the sake of simplicity. In this case, an external constant force induces a flux of particles in the system, and the ratio between the average flux and the force in the small forcing limit defines the mobility. More precisely, the dynamics writes

$$\begin{cases} \mathrm{d}q_t = \dfrac{1}{m} p_t \,\mathrm{d}t\,, \\ \mathrm{d}p_t = (-\nabla V(q_t) + \eta F)\,\mathrm{d}t - \dfrac{\gamma}{m} p_t \,\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\mathrm{d}W_t, \end{cases} \tag{1.19}$$

where $\eta \geqslant 0$, $F \in \mathbb{R}^D$ is a constant unit vector ($|F| = 1$) and V is a $\mathcal{C}^\infty$ periodic potential. Note that $-\nabla V(q_t) + \eta F$ is not the gradient of a periodic function, so that the momentum distribution is not a centered Gaussian a priori. This allows the mean velocity to be non-zero and therefore a particle flux can appear. The generator of the non-equilibrium dynamics writes

$$\mathcal{L}_\eta = \mathcal{L}_0 + \eta \widetilde{\mathcal{L}},$$

where $\mathcal{L}_0$ is the generator at equilibrium (see (1.6)) and $\widetilde{\mathcal{L}} = F^\top \nabla_p$ is the generator of the perturbation.

Using Lyapunov techniques (see Section 1.2.2.3) and hypoellipticity one can show that the dynamics (1.19) admits a unique invariant probability measure μ_η for any $\eta \in \mathbb{R}$ [89], with a smooth density with respect to the Lebesgue measure. The particle flux is measured in the direction F of the external force, in the linear response regime which corresponds to a small forcing. Denoting by $\mathbb{E}_\eta$ the expectation with respect to μ_η, the mobility is defined as

$$\alpha := \lim_{\eta \to 0} \frac{\mathbb{E}_\eta \left[\frac{1}{m} F^\top p \right]}{\eta}. \tag{1.20}$$

1.3.3 Linear response

Linear response theory allows to reformulate the mobility as the autocorrelation of the particle flux [97].

Proposition 1.4 (Green–Kubo formula)**.** *The mobility coefficient for the dynamics* (1.19) *rewrites*

$$\alpha = \frac{\beta}{m^2} \left\langle -\mathcal{L}_0^{-1} F^\top p, F^\top p \right\rangle_{\mathrm{L}^2(\mu_0)} = \frac{\beta}{m^2} \int_0^\infty \mathbb{E}\left[F^\top p_0 F^\top p_t \right] \mathrm{d}t\,,$$

where the expectation is taken over initial conditions $(q_0, p_0) \sim \mu_0$ *and over all realizations of the dynamics at equilibrium.*

Proof. Transport coefficients can be reformulated using linear response theory by providing an expansion of the steady state μ_η in powers of η, for perturbations which are not too strong. For any $\varphi \in \mathrm{L}^2(\mu)$, we denote its orthogonal projection on $\mathrm{L}^2_0(\mu)$ by

$$\Pi_0 \varphi = \varphi - \mathbb{E}_\mu[\varphi].$$

We first note that the perturbation $\widetilde{\mathcal{L}}$ is $\mathcal{L}_0$-bounded. Indeed, for any $\varphi \in \mathrm{L}^2_0(\mu_0)$,

$$\begin{aligned}\|\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\varphi\|^2 \leqslant \|\nabla_p \mathcal{L}_0^{-1}\varphi\|^2 &= \left\langle \nabla_p^* \nabla_p \mathcal{L}_0^{-1}\varphi, \mathcal{L}_0^{-1}\varphi \right\rangle_{\mathrm{L}^2(\mu_0)} = \beta\gamma^{-1} \left\langle -\mathcal{L}_0\mathcal{L}_0^{-1}\varphi, \mathcal{L}_0^{-1}\varphi \right\rangle_{\mathrm{L}^2(\mu_0)} \\ &\leqslant \beta\gamma^{-1} \|\mathcal{L}_0^{-1}\|_{\mathcal{B}(\mathrm{L}^2_0(\mu_0))} \|\varphi\|^2,\end{aligned}$$

where we used that the symmetric part of $-\mathcal{L}_0$ is $\beta^{-1}\gamma\nabla_p^*\nabla_p$. We recall that the evolution semi-group of the Langevin dynamics on a compact domain decays exponentially, so that the generator $\mathcal{L}_0$ is invertible on $\mathrm{L}^2_0(\mu_0)$ (see Section 1.2.2.3). Therefore the operator $\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\Pi_0$ is bounded on $\mathrm{L}^2(\mu_0)$, so that there exist $\eta^* > 0$ such that, for any $0 \leqslant \eta \leqslant \eta^*$, the series $\sum_{k=0}^\infty \eta^k(-\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\Pi_0)^k$ converges in $\mathcal{B}(\mathrm{L}^2(\mu_0))$. Let us construct an invariant measure $\widetilde{\mu}_\eta$ of the form:

$$\widetilde{\mu}_\eta = \sum_{k=0}^\infty \eta^k f_k \mu_0,$$

for any $0 \leqslant \eta \leqslant \eta^*$, and prove that it is a probability measure. By uniqueness of the invariant probability measure μ_η, we will be able to conclude that $\mu_\eta = \widetilde{\mu}_\eta$ (we refer to the discussion before (1.20)).

Using the invariance of the measure, for any $\mathcal{C}^\infty$ function φ with compact support,

$$\int_{\mathcal{E}} (\mathcal{L}_0 + \eta\widetilde{\mathcal{L}})\varphi \, \mathrm{d}\widetilde{\mu}_\eta = 0,$$

so that by identifying each order in η we get $f_0 = \mathbf{1}$ and, for any $k \in \mathbb{N}$,

$$\mathcal{L}_0^* f_{k+1} + \widetilde{\mathcal{L}}^* f_k = 0.$$

The operator $(\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\Pi_0)^* = \Pi_0 \left(\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\Pi_0\right)^*$ is bounded on $\mathrm{L}^2(\mu_0)$, so that

$$\widetilde{\mu}_\eta = \left(\mathbf{1} + \sum_{k=1}^\infty \eta^k \left[(-\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\Pi_0)^*\right]^k \mathbf{1}\right) \mu_0.$$

is an invariant measure.

The measure $\widetilde{\mu}_\eta$ is normalized since for any $k \in \mathbb{N}$, $\left[(-\widetilde{\mathcal{L}}\mathcal{L}_0^{-1})^*\right]^k \mathbf{1} \in \mathrm{L}^2_0(\mu_0)$:

$$\int_{\mathcal{E}} \mathrm{d}\widetilde{\mu}_\eta = \int_{\mathcal{E}} \mathrm{d}\mu_0 = 1.$$

The positivity of $\widetilde{\mu}_\eta$ is proved using the ergodicity of the dynamics with respect to μ_η

in [104, Section 5.2.2]. This allows to conclude that $\widetilde{\mu}_\eta = \mu_\eta$ by uniqueness of the invariant probability measure.

In particular the first order term involved in (1.20) provides another expression for the mobility:

$$\alpha = \mathbb{E}_0\left[-\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\frac{1}{m}F^\top p\right] = \frac{\beta}{m^2}\left\langle -\mathcal{L}_0^{-1}F^\top p, F^\top p\right\rangle_{\mathrm{L}^2(\mu_0)},$$

where we used $\widetilde{\mathcal{L}}^*\mathbf{1} = \frac{\beta}{m}F^\top p$, and where $\langle\cdot,\cdot\rangle_{\mathrm{L}^2(\mu_0)}$ is the scalar product in $\mathrm{L}^2(\mu_0)$. The mobility coefficient can be reformulated as a velocity autocorrelation by noting that $-\mathcal{L}_0^{-1} = \int_0^\infty \mathrm{e}^{t\mathcal{L}_0}\,\mathrm{d}t$ in $\mathrm{L}^2_0(\mu_0)$, similarly (1.16):

$$\begin{aligned}\alpha &= \frac{\beta}{m^2}\left\langle -\int_0^\infty \mathrm{e}^{t\mathcal{L}_0}\,\mathrm{d}t\,F^\top p, F^\top p\right\rangle_{\mathrm{L}^2(\mu_0)} = \frac{\beta}{m^2}\int_0^\infty \left\langle \mathrm{e}^{t\mathcal{L}_0}F^\top p, F^\top p\right\rangle_{\mathrm{L}^2(\mu_0)}\mathrm{d}t \\ &= \frac{\beta}{m^2}\int_0^\infty \mathbb{E}\left[F^\top p_t\,F^\top p_0\right]\mathrm{d}t\,.\end{aligned}$$

This concludes the proof. □

1.3.4 Numerical estimation of transport coefficients

There exist several numerical methods allowing to estimate a transport coefficient. We discuss here the two most standard ways and also mention another one, reviewed in [52] and [161].

Equilibrium methods. They consist in estimating the transport coefficient as an integrated autocorrelation over equilibrium dynamics, using the Green–Kubo formula (Proposition 1.4):

$$\alpha = \frac{\beta}{m^2}\int_0^\infty \mathbb{E}\left[F^\top p_0 F^\top p_t\right]\,\mathrm{d}t\,.$$

The numerical integration of the dynamics induces a bias depending on the time step Δt in the estimation of the mobility [102]. The time integral must be truncated to be estimated numerically, which is a second source of numerical error. Choosing the correct truncation is tricky since the function $t \mapsto \mathbb{E}\left[F^\top p_0 F^\top p_t\right]$ is integrable but not necessarily monotonically decreasing for irreversible dynamics such as the Langevin dynamics. We refer to Section 3.9 for a short review of different estimators of this integral.

The mobility coefficient can also be written using the mean square displacement:

$$\alpha = \lim_{t\to\infty}\frac{\mathbb{E}\left[\left(F^\top(Q_t - Q_0)\right)^2\right]}{2t},$$

where the expectation is taken over realizations of the Brownian motion, and $Q_t - Q_0 = \frac{1}{m}\int_0^t p_t\,\mathrm{d}t$ is the unperiodized displacement. This provides another way to compute the mobility coefficient, relying on a collection of replicas of the system for which the square displacement is computed.

Non-equilibrium steady-state techniques. These techniques rely directly on the definition of the transport coefficient α, approximating the limit of the rate of increase by a finite difference. Its consists in simulating the non-equilibrium dynamics (e.g. (1.19)) for a perturbation amplitude η small enough, and to average the flux $\varphi(q,p) = \frac{1}{m}F^\top p$ over a simulation time T.

The estimator $\frac{1}{\eta T}\int_0^T \varphi\left(x_t^\eta\right) \mathrm{d}t$ has mean $\alpha + \mathrm{o}(1)$ and asymptotic standard deviation $\eta^{-1}\left(\sigma_{\varphi,0} + \mathrm{o}(1)\right)$. The relative statistical error committed when estimating the transport coefficient is therefore of order

$$\frac{T^{-1/2}\eta^{-1}\sigma_{\varphi,0}}{\alpha} = \frac{\sqrt{\beta^{-1}\alpha}}{\sqrt{T}\eta},$$

since the asymptotic variance of the observable φ at equilibrium satisfies $\sigma_{\varphi,0}^2 = \beta^{-1}\alpha$. For a given relative error $\varepsilon > 0$ one should therefore take a simulation time of the order $T \sim \eta^{-2}\varepsilon^{-2}$, which is very large. This leads in practice to very large simulation times.

Moreover the system cannot be initialized under the steady state distribution since it is not known explicitly. For certain boundary-driven dynamics such as the atom chain (1.18) this leads to transient regimes which can be extremely long. This part of the trajectory has to be discarded when averaging the flux response (burn-in), resulting in a loss of efficiency for the estimation. In the case of the estimation of the thermal conductivity for a one dimensional chain, the two previous problems add up and simulations typically take weeks or months for chains composed of tens of thousands of atoms.

Transient methods. The systems is initialized in a state which is not typical under the equilibrium probability measure, and by comparing the relaxation with the evolution given by a macroscopic model (such as the heat equation or Navier-Stokes equations) one is able to identify the transport coefficient (here a thermal conductivity or a viscosity) [84]. These methods are however less popular than the two previous types of techniques.

Each of these methods suffer from large variance issues, resulting in a poor accuracy. In this thesis we focus on the third strategy, which is also called non-equilibrium molecular dynamics (NEMD). Variance reduction methods designed for non-equilibrium systems are reviewed in Section 1.4.2. We propose in Chapter 3 a new variance reduction method relying on control variates to address this issue.

1.4 Variance reduction

The idea of variance reduction is to make the convergence of the ergodic average $\widehat{\varphi}_t$ faster towards its expectation $\mathbb{E}_\mu[\varphi]$, in order to reduce the statistical error for finite simulation times. In this section we first present variance reduction techniques for equilibrium systems. We then review the current approaches we know of for nonequilibrium systems.

1.4.1 Variance reduction at equilibrium

Large asymptotic variances are due to long correlation times of the states along the trajectory of the stochastic process. This happens when the trajectory stays trapped in a part of the state space during a large time before leaving it. This phenomenon is called metastability. A typical example is the following double well potential:

$$V(q) = q_1^2 + \frac{1}{\varepsilon}(q_2^2 - 1)^2,$$

for $\varepsilon > 0$. When the parameter ε is small compared to the temperature β^{-1} the energetic barrier at $q_2 = 0$ is hard to cross for a process such as the Langevin dynamics. This leads to a large asymptotic variance for observables depending on q_2 in a non-symmetric way. The most standard variance reduction techniques in molecular simulations rely on importance sampling or stratification. We refer to the review provided in [28], which also presents antithetic variables. These two methods are of general purpose as they tend to improve the sampling efficiency for a whole class of methods, by opposition to control variate techniques (see Chapter 3) which are target-oriented.

1.4.1.1 Importance sampling

Importance sampling consists in replacing the measure which is being sampled by another one which is easier to sample [96, 36, 108]. For any potential $\widetilde{V}$ one can generate a dynamics which is ergodic for the probability distribution $\mu_{\widetilde{V}} = \frac{Z_V}{Z_{\widetilde{V}}} \mathrm{e}^{\beta(V-\widetilde{V})}\mu = Z_{\widetilde{V}}^{-1}\mathrm{e}^{-\beta\widetilde{V}}\kappa$ associated with the potential $\widetilde{V}$, where Z_V and $Z_{\widetilde{V}}$ are normalization constants. We recall that κ is the marginal of μ for the momentum variable. The mean of the observable φ with respect to the Gibbs measure of interest μ is then given by

$$\mathbb{E}_\mu[\varphi] = \frac{\int_{\mathcal{E}} \varphi \, \mathrm{e}^{\beta(\widetilde{V}-V)} \mathrm{d}\mu_{\widetilde{V}}}{\int_{\mathcal{E}} \mathrm{e}^{\beta(\widetilde{V}-V)} \mathrm{d}\mu_{\widetilde{V}}}. \tag{1.21}$$

Note that the denominator is the proper normalization factor since $\mathbb{E}_\mu[\mathbf{1}] = 1$. The expectation (1.21) can be estimated by

$$\widehat{\varphi}_t^{\widetilde{V}} = \frac{\int_0^t \varphi(\widetilde{q}_s, \widetilde{p}_s) \mathrm{e}^{\beta(\widetilde{V}-V)(\widetilde{q}_s)} \, \mathrm{d}s}{\int_0^t \mathrm{e}^{\beta(\widetilde{V}-V)(\widetilde{q}_s)} \, \mathrm{d}s},$$

which is the ratio of two ergodic averages over the trajectory $(\widetilde{q}_s, \widetilde{p}_s)$ solution to the modified Langevin dynamics:

$$\begin{cases} \mathrm{d}\widetilde{q}_t = \dfrac{1}{m}\widetilde{p}_t \, \mathrm{d}t \,, \\ \mathrm{d}\widetilde{p}_t = -\nabla\widetilde{V}(\widetilde{q}_t) \, \mathrm{d}t - \dfrac{\gamma}{m}\widetilde{p}_t \, \mathrm{d}t + \sqrt{2\gamma\beta^{-1}} \mathrm{d}W_t. \end{cases}$$

Modifying the dynamics can allow to reduce the time correlation of the process, which reduces the variance associated to the two averages involved in the estimator. On the other hand, if the process $(\widetilde{q}_s, \widetilde{p}_s)$ stays too long in the region where V is larger than $\widetilde{V}$, then the numerator and the denominator of the estimator are small, leading to possibly large statistical errors.

In the example of the double well potential a natural (but maybe not optimal) choice for $\widetilde{V}$ would be the convex envelop of V, which amounts to taking the energetic barrier out. In this case the process goes freely from one well to the other one, which allows to overcome the metastability. In real applications, in particular in higher dimension, finding an appropriate $\widetilde{V}$ can however be challenging.

The metastable behavior is sometimes associated to a reaction coordinate, or collective variable, denoted by $q \mapsto \xi(q) \in \mathbb{R}$. In the case of the double well potential, the function $\xi(q) = q_2$ is a reaction coordinate since the metastability comes from the q_2 variable. When such a reaction coordinate is available, the modified potential $\widetilde{V}$ can be constructed using the free energy function F:

$$\widetilde{V}(q) = V(q) - F(\xi(q))$$

where the free energy $F : \mathbb{R} \to \mathbb{R}$ is such that, for any function $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\int_{\mathcal{D}} \varphi(\xi(q))\, \nu(\mathrm{d}q) = Z_F^{-1} \int_{\mathbb{R}} \varphi(z)\, \mathrm{e}^{-\beta F(z)}\, \mathrm{d}z,$$

for some normalization constant Z_F [103]. This construction allows to remove the metastability in the direction where ξ varies. Another way to take advantage from a reaction coordinate to reduce the variance is the stratification method.

1.4.1.2 Stratification

This second variance reduction technique relies on the decomposition of the potential V, which produces metastability, into a sum of potentials which are easier to sample from. When a reaction coordinate is known, the stratification consists in slicing the system into zones corresponding to given values of ξ. In the previous example of a doublewell with reaction coordinate $\xi(q) = q_2$, the technique relies on the sampling of the unimodal marginal distributions of density $\nu(\cdot, \nu_2)$ for every value of $\nu_2 \in \mathbb{R}$.

The first type of stratification technique consists in constraining the dynamics to staying on a manifold corresponding to a given value of the reaction coordinate: $\xi(q) = z$. In this case the state space is split into an infinite number of non-overlapping regions. The prescribed value z can vary in order to sample the full phase space. The sampling of each manifold defined by $\xi(q) = z$ is supposed to be easy, by construction of ξ, so that the overall sampling is efficient. This is known as thermodynamic integration and we refer to [103] for further details.

The second type is called umbrella sampling [160, 124, 42]. It relies on a set of overlapping regions, which are defined through a set a of K non-negative partition functions

χ_i:

$$\forall q \in \mathcal{D}, \quad \sum_{i=1}^{K} \chi_i(q) = 1.$$

These functions typically have compact support. For instance, they can depend only on the reaction coordinate:

$$\chi_i(q) = \widetilde{\chi}_i(\xi(q)), \quad \text{with} \quad \forall z \in \mathbb{R}, \ \sum_{i=1}^{K} \widetilde{\chi}_i(z) = 1.$$

The Gibbs measure is split as follows:

$$\mu(\mathrm{d}q, \mathrm{d}p) = \sum_{i=1}^{K} Z_i \, \mu_i(\mathrm{d}q, \mathrm{d}p), \qquad \mu_i(\mathrm{d}q, \mathrm{d}p) = Z_i^{-1} \, \chi_i(q) \, \mu(\mathrm{d}q, \mathrm{d}p),$$

where the $Z_i = \int_{\mathcal{E}} \chi_i(q) \, \mu(\mathrm{d}q, \mathrm{d}p)$ are normalization constants such that μ_i are probability distributions on $\mathcal{E}$. Then, for any measurable observable φ, its average with respect to the Gibbs measure μ can be expressed as a linear combination of averages with respect to the probability measures μ_i:

$$\mathbb{E}_{\mu}[\varphi] = \sum_{i=1}^{K} Z_i \, \mathbb{E}_{\mu_i}[\varphi]. \tag{1.22}$$

The local averages $\mathbb{E}_{\mu_i}[\varphi]$ are computed using a dynamics with an additional gradient force $\nabla \log \chi_i(q)$, and are aggregated using the weights $(Z_i)_{1 \leqslant i \leqslant K}$ which however need to be estimated. Anticipating on Section 1.4.2, note that we are here able to sample from $\mu_i = Z_i^{-1} \, \chi_i \, \mu$ because the system is at equilibrium. Out of equilibrium, the invariant probability measure μ is not explicit so that sampling from μ_i is not directly possible anymore.

Computing the weights Z_i is the main difficulty of the Umbrella Sampling approach. They can of course be estimated using averages with respect to μ, but it would bring us back to the original sampling problem. The different variants of Umbrella Sampling differ mostly through the way the weights Z_i are estimated. The most widely used technique in chemical physics application is the weighted histogram analysis method (WHAM). It is closely related to the multistate Bennett acceptance ratio (M-BAR) method, and both are derived from maximum-likelihood or minimum asymptotic variance principles [163, 98, 148]. Note that M-BAR is mathematically more satisfying than WHAM. We also mention the Eigenvector Method for Umbrella Sampling (EMUS), for which an error analysis is provided in [157].

1.4.2 Variance reduction out of equilibrium

Importance sampling and stratification methods require some knowledge on the Gibbs measure μ which is sampled. Indeed in both cases we use the property that adding a gradient force $-\nabla W$ to the dynamics amounts to multiplying the invariant probability measure by a factor proportional to $\mathrm{e}^{-\beta W}$. This property is not satisfied out of equilibrium in general, so that it is delicate to perturbe the dynamics. In particular the previous variance reduction

technique cannot be used in this context.

There is nonetheless a crying need for variance reduction when estimating transport coefficients through linear response, as motivated in Section 1.3.4. Some variance reduction techniques designed for non-equilibrium systems have been proposed in the literature. We present them here, and discuss their potential limitations.

1.4.2.1 Non-equilibrium Umbrella Sampling (NEUS)

A stratification technique derived from Umbrella Sampling has been proposed in [169], then simplified and generalized in [157, Section IV]. We refer to [42] for a numerical analysis of the method. The phase space $\mathcal{E}$ of the simulation is decomposed into boxes $(A_i)_{1\leqslant i\leqslant n}$ defining a lattice. With the notations of 1.4.1.2,

$$\chi_i = \mathbf{1}_{A_i}, \qquad Z_i = \int_{A_i} \mathrm{d}\mu,$$

and $Z = (Z_i)_{1\leqslant i\leqslant n}$ is the solution of the linear system $Z = GZ$ where $G_{i,j}$ is proportional to the probability flux from A_i to A_j. The stratification method consists in sampling each of these boxes independently, and aggregating the results using (1.22). The specificity of nonequilibrium dynamics is that detailed balance does not hold, so that there exist steady probability fluxes between neighboring boxes: $G_{i,j}$ and $G_{j,i}$ can differ. Moreover there is no straightforward method to sample the restriction μ_i of the invariant probability measure to a given box A_i to obtain $\mathbb{E}_{\mu_i}[\varphi]$.

The restricted distributions μ_i are sampled using fragments of trajectories with initial condition $\widetilde{\mu}_i$ and killing conditions on the boundary. The probability distribution $\widetilde{\mu}_i$ such that the fragments of trajectories preserve the restricted distribution μ_i are not known a priori, and they have to be estimated. The NEUS method estimates these distributions using a fixed-point iterations. This procedure also provides an estimator for the coefficients $G_{i,j}$ based on the empirical frequency corresponding to trajectories in the box A_i terminated at the boundary with the box A_j.

The stratification allows to reduce the variance in presence of metastability, as it has been described in Section 1.4.1.2. In high dimension the construction of the boxes A_i relies on the knowledge of a reaction coordinate ξ, which describes well the metastability. It is however not adapted as such to the computation of some transport coefficients such as the thermal conductivity of an atom chain or the mobility of a fluid. In both cases indeed the large statistical error comes from a large noise to signal ratio, and not from metastability. It may however be possible to adapt NEUS to this situation, possibly by considering the flux as the variable which respect to which stratification is performed. Moreover the method requires bookkeeping, which makes the implementation somewhat cumbersome.

1.4.2.2 Coupling control variates

A coupling control variance approach has been proposed in [68] for a steady flux of matter in a one-dimensional lattice gas and adapted to the context of continuous non-equilibrium dynamics for the numerical results reported in [102]. It consists in simulating both the

non-equilibrium dynamics $(X_t^\eta)_{t\geqslant 0}$ and the equilibrium dynamics $(X_t^0)_{t\geqslant 0}$ using coupled Brownian motions $(W_t^\eta)_{t\geqslant 0}$ and $(W_t^0)_{t\geqslant 0}$, and starting from the same initial condition. The quantity

$$\frac{1}{T}\int_0^T \Big(\mathbb{E}_0[\varphi] + \varphi(X_t^\eta) - \varphi(X_t^0)\Big)\,\mathrm{d}t\,, \tag{1.23}$$

is then an unbiased estimator of $\mathbb{E}_\eta[\varphi]$. Its variance is smaller than which of the standard empirical average $\frac{1}{T}\int_0^T \varphi(X_t^\eta)\,\mathrm{d}t$ when the processes $(X_t^\eta)_{t\geqslant 0}$ and $(X_t^0)_{t\geqslant 0}$ are strongly coupled. The difficulty of the method relies therefore on the construction of an adequate coupling of the Brownian motions, allowing to keep the two trajectories as close as possible even in the long time limit.

Consider for example the following non-equilibrium perturbation of the overdamped Langevin dynamics:

$$\mathrm{d}X_t^\eta = (-\nabla V(X_t^\eta) + \eta F(X_t^\eta))\,\mathrm{d}t + \sigma \mathrm{d}W_t, \tag{1.24}$$

for $\eta \geqslant 0$ small and F a bounded non-gradient function with $\|F\|_\infty = 1$. Consider $(X_t^0)_{t\geqslant 0}$ the equilibrium dynamics obtained for the same Brownian and initial condition, which corresponds to what is called a synchronous coupling, then

$$\mathrm{d}\Big(X_t^\eta - X_t^0\Big) = (-\nabla V(X_t^\eta) + \eta F(X_t^\eta))\,\mathrm{d}t + \nabla V(X_t^0)\;\mathrm{d}t\,,$$

so

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{|X_t^\eta - X_t^0|^2}{2}\right) = -(X_t^\eta - X_t^0)\cdot\Big(\nabla V(X_t^\eta) - \nabla V(X_t^0) + \eta F(X_t^\eta)\Big),$$

where we denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^D$. Assuming that the potential V is β-convex, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{|X_t^\eta - X_t^0|^2}{2}\right) \leqslant -\beta|X_t^\eta - X_t^0|^2 + \eta|X_t^\eta - X_t^0|,$$

so the distance between the two trajectories is bounded uniformly for any time $t \geqslant 0$:

$$|X_t^\eta - X_t^0| \leqslant \frac{\eta}{\beta}.$$

In practice two problems arise.

i) The previous estimation crucially relies on the convexity of the potential. When the potential V is multi-modal, the trajectory couple when they are in the same local minima. However they diverge and decouple, potentially for a long time, as soon as one of the two trajectories crosses an energetic barrier (which is locally concave).

ii) We are mostly interested in non-reversible dynamics such as Langevin dynamics. As pointed out in [48], the difference process is not contractive for the Langevin dynamics even for a strictly convex potential.

One could nevertheless try to construct a mix of synchronous and reflexion couplings, as proposed in [48] for theoretical purposes, to obtain a coupling control variate for the

Langevin dynamics with general potentials. Numerical results presented in [22] quantify the efficiency of some coupling methods for a kinetic dynamics in non convex potentials.

1.4.2.3 Tangent vector method

When $\mathbb{E}_0[\varphi] = 0$ the previous estimator (1.23) divided by η is a (biased) estimator of the transport coefficient α:

$$\frac{1}{T}\int_0^T \frac{\varphi(X_t^\eta) - \varphi(X_t^0)}{\eta}\,\mathrm{d}t \xrightarrow[T\to\infty]{} \alpha + \mathrm{O}(\eta),$$

for any $\eta > 0$. It is therefore tempting to send the perturbation amplitude η to 0 in order to remove the bias and to strengthen the coupling between X_t^η and X_t^0. We obtain formally the derivative of $\varphi(X_t^\eta)$ with respect to η taken at equilibrium:

$$\frac{1}{T}\int_0^T \frac{\mathrm{d}}{\mathrm{d}\eta}\varphi(X_t^\eta)\bigg|_{\eta=0}\,\mathrm{d}t \xrightarrow[T\to\infty]{} \alpha.$$

Denoting by $T_t := \frac{\mathrm{d}}{\mathrm{d}\eta}X_t^\eta\Big|_{\eta=0}$ the tangent dynamics, it holds:

$$\frac{\mathrm{d}}{\mathrm{d}\eta}\varphi(X_t^\eta)\bigg|_{\eta=0} = T_t \cdot \nabla\varphi(X_t^0).$$

It suffices therefore to compute T_t to obtain a practical method to compute the transport coefficient using the estimator:

$$\widehat{\alpha}_T := \frac{1}{T}\int_0^T T_t \cdot \nabla\varphi(X_t^0).$$

This method has been recently proposed in [8], where the quantity T_t is called the tangent vector. Note that T_t is a random variable since it depends on the realization of the Brownian motion, and on the initial condition X_0^0. In the context of the non-equilibrium overdamped Langevin equation (1.24), the state X_t^0 following the equilibrium trajectory and the tangent vector T_t satisfy the system:

$$\begin{cases} \mathrm{d}X_t^0 = -\nabla V(X_t^0)\,\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \\ \mathrm{d}T_t = \left[F(X_t^0) - \nabla^2 V(X_t^0)T_t\right]\,\mathrm{d}t\,, \end{cases}$$

where $\nabla^2 V$ denotes the Hessian of the potential.

In [8] the proofs are provided for the overdamped Langevin equation, for a general potential V. There is however no fundamental reason this method should not work for the Langevin equation, in a non-reversible setting.

1.4.2.4 Linearized Girsanov method

This last technique relies on Girsanov's theorem. Considering X_t^η solution to the general SDE (1.3) with constant diffusion coefficient $\sigma > 0$ and constant external force in the direction $F \in \mathbb{R}^D$:

$$\mathrm{d}X_t^\eta = b(X_t^\eta)\ \mathrm{d}t + \eta F\ \mathrm{d}t + \sigma \mathrm{d}W_t,$$

the theorem states that, for any finite time $t \geqslant 0$, forcing $\eta \geqslant 0$, observable φ anda any given initial condition x_0, the following equality holds:

$$\mathbb{E}_{x_0}\left[\frac{1}{t}\int_0^t \varphi(X_s^\eta)\mathrm{d}s\right] = \mathbb{E}_{x_0}\left[\frac{1}{t}\int_0^t \varphi(X_s^0)\mathrm{d}s\ \exp\left(\frac{\eta}{\sigma}\int_0^t F\cdot \mathrm{d}W_s - \frac{\eta^2}{2\sigma^2}\int_0^t |F|^2\,\mathrm{d}s\right)\right],$$

where the expectations are taken over all realizations of the Brownian motion. Denoting by $Z_t = \frac{1}{\sigma}\int_0^t F\cdot \mathrm{d}W_s$ and by $\widehat{\varphi}_t^\eta = \frac{1}{t}\int_0^t \varphi(X_s^\eta)\mathrm{d}s$, we obtain

$$\frac{1}{\eta}\left(\mathbb{E}_{x_0}[\widehat{\varphi}_t^\eta] - \mathbb{E}_{x_0}[\widehat{\varphi}_t^0]\right) = \mathbb{E}_{x_0}\left[\widehat{\varphi}_t^0\,\frac{1}{\eta}\left(\exp\left(\eta Z_t - \frac{\eta^2\beta}{2\gamma}\int_0^t |F|^2\,\mathrm{d}s\right) - 1\right)\right].$$

Noting that, by ergodicity

$$\lim_{\eta\to 0}\lim_{t\to\infty}\frac{1}{\eta}\left(\mathbb{E}_{x_0}[\widehat{\varphi}_t^\eta] - \mathbb{E}_{x_0}[\widehat{\varphi}_t^0]\right) = \lim_{\eta\to 0}\frac{1}{\eta}\left(\mathbb{E}_\eta[\varphi] - \mathbb{E}_0[\varphi]\right) = \alpha,$$

whereas

$$\lim_{t\to\infty}\lim_{\eta\to 0}\mathbb{E}_{x_0}\left[\widehat{\varphi}_t^0\,\frac{1}{\eta}\left(\exp\left(\eta Z_t - \frac{\eta^2\beta}{2\gamma}\int_0^t |F|^2\,\mathrm{d}s\right) - 1\right)\right] = \lim_{t\to\infty}\mathbb{E}_{x_0}\left[\widehat{\varphi}_t^0\,Z_t\right].$$

This suggests to consider

$$\widehat{\alpha}_t := \mathbb{E}_{x_0}\left[\frac{1}{t}\int_0^t \varphi(X_t^0)\,Z_t\,\mathrm{d}t\right],$$

as an estimator of the mobility coefficient. This estimator has been proposed in [167, 5] and analyzed in [67, 168] for jump Markov processes. The proof that $\widehat{\alpha}_t$ converges to α as t goes to infinity follows the same lines for SDEs.

1.5 Contributions of this work

I briefly present in this section the contributions of this PhD work.

1.5.1 Spectral methods for Langevin dynamics and associated error estimates

Chapter 2 (published in [145]) tackles the numerical approximation of the (unique) solution of the partial differential equation

$$-\mathcal{L}\Phi = R,$$

where $\mathcal{L}$ is a hypocoercive operator, invertible on the space $\mathrm{L}^2_0(\mu)$ where μ is a probability measure, and $R \in \mathrm{L}^2_0(\mu)$. This type of Poisson equation appears notably in the context of kinetic particle dynamics such as the Langevin dynamics. Solving this equation for Φ provides a lot of information on the dynamical properties of the process. It allows for example to compute the asymptotic variance of the observable R, a transport coefficient or to construct a control variate as made precise in Chapter 3.

We approximate numerically the solution Φ using a Galerkin method. In the case when the generator $\mathcal{L}$ is coercive, the Lax-Milgram theorem ensures the well-posedness of the method. Moreover error estimates are provided by Céa's Lemma in this case. Our results extend these theoretical guarantees to the case when the operator $\mathcal{L}$ is not coercive but only hypocoercive, under adequate assumptions. In particular we derive conditions under which the restriction of the generator $\mathcal{L}$ to the Galerkin space $V \subset \mathrm{L}^2_0(\mu)$ is still hypocoercive. A saddle point formulation, involving a Lagrange multiplier, is also proposed and analyzed. This approach allows to write the Galerkin method in the whole space $\mathrm{L}^2(\mu)$, which is in the end more convenient.

1.5.2 A perturbative approach to control variates in molecular dynamics

Chapter 3 (see the preprint in [144]) is motivated by the efficient estimation of transport coefficients. We present a general variance reduction strategy based on control variates which do not rely on the knowledge of the invariant probability distribution. The latter is indeed unknown when the system is in a nonequilibrium steady state, so that standard variance reduction techniques (importance sampling, stratification, ...) cannot be used. The idea is that if the dynamics can be simplified into a surrogate dynamics, for which solutions to the Poisson equations can be computed (e.g. low dimensional or linear dynamics), then we can construct a modified observable. The new estimator is unbiased by design, and its asymptotic variance is greatly reduced if the simplified dynamics is close to the original one, as we show on specific examples.

We prove that the asymptotic variance of this estimator scales quadratically with the amplitude of the difference between the two dynamics in the perturbative regime. This variance reduction technique is illustrated on three nonequilibrium systems: a single particle in a one-dimensional periodic potential under a non-gradient forcing; the computation of the thermal flux passing through a chain; and the estimation of the mean length of a dimer in a solvent under an external shearing stress.

1.5.3 Efficient mobility estimation in the underdamped regime

In Chapter 4 we show how the variance reduction strategy from Chapter 3 can be adapted to the case of a low dimensional Langevin dynamics in the underdamped regime. This work has been carried out at Imperial College London with G. Pavliotis during a two month PhD mobility program. We recall in a first part that in this limit the dynamics rescaled in time converges to a diffusion process on a graph. We then construct a control variate using the

Fokker-Planck equation corresponding to this limiting dynamics. We obtain this way a new estimator of the mobility. We illustrate with numerical simulations that it behaves well in the underdamped regime. The so-constructed control variate allows us to study the scaling of the mobility for a two-dimensional non integrable system with the friction coefficient. We show numerically that the mobility do not behave as $1/\gamma$, contrarily to the one-dimensional setup. However, this scaling is not well understood from a theoretical viewpoint.

1.5.4 Hypocoercivity of Piecewise Deterministic Markov Process Monte Carlo

In Chapter 5 (see the preprint in [4]) we show that a large class of kinetic dynamics involving velocity jumps, called PDMPs, are geometrically ergodic under weak assumptions on the potential. This work has been initiated at Imperial College London with N. Nüsken and continued in collaboration with C. Andrieu and A. Durmus. This class of PDMPs includes the Zig-Zag process (ZZ), the Bouncy Particle Sampler (BPS) or Randomized Hamiltonian Monte Carlo (RHMC). The ZZ and the BPS are recent tools in statistics and in statistical physics, where they have both attracted much interest. The RHMC on the other hand is a well established sampler, with has been widely used.

Our proof, reling on L^2 hypocoercivity techniques (see Section 1.2.2.3), allows to derive quantitative spectral gap estimates. The scaling of our bounds with the refreshment rate, the counterpart of the friction γ for Langevin dynamics, the choice of velocity space and most importantly the dimension d, is made explicit. We prove in particular that the spectral gap is uniformly bounded away from zero in the high dimension limit for the ZZ and the RHMC under some simple assumptions on the potential. This result also holds for the Langevin dynamics. We also prove under the same assumptions that for the BPS the spectral gap is larger than $d^{-1/2}$, which seems to be sharp considering the result from [17].

Chapter 2

Spectral methods for Langevin dynamics and associated error estimates

This chapter provides the content of [145] with some changes of notation and minor changes.

Contents

We prove the consistency of Galerkin methods to solve Poisson equations where the differential operator under consideration is hypocoercive. We show in particular how the hypocoercive nature of the generator associated with Langevin dynamics can be used at the discrete level to first prove the invertibility of the rigidity matrix, and next provide error bounds on the approximation of the solution of the Poisson equation. We present general convergence results in an abstract setting, as well as explicit convergence rates for a simple example discretized using a tensor basis. Our theoretical findings are illustrated by numerical simulations.

2.1 Introduction

Statistical physics gives a theoretical framework to bridge the gap between microscopic and macroscopic descriptions of matter [12]. This is done in practice with numerical methods known as molecular simulation [2, 64, 161, 101]. Despite its intrinsic limitations on spatial and timescales, molecular simulation has been used and developed over the past 50 years, and recently gained some recognition through the 2013 Chemistry Nobel Prize. One important aim of molecular dynamics is to quantitatively evaluate macroscopic properties of interest, obtained as averages of functions of the full microstate of the system (positions and velocities of all atoms in the system) with respect to some probability measure, called thermodynamic ensemble. Some properties of interest are static (a.k.a. thermodynamic properties): heat capacities; equations of state relating pressure, density and temperature; etc. Other properties of interest include some dynamical information. This is the case for transport coefficients (such as thermal conductivity, shear viscosity, etc) or time-dependent dynamic properties such as Arrhenius constants which parametrize chemical kinetics.

From a technical viewpoint, the computation of macroscopic properties requires in any case the sampling of high-dimensional measures. We consider in this work the computation of properties in the canonical ensemble, characterized by the Boltzmann–Gibbs measure, which models systems at constant temperature. One popular way to sample the canonical ensemble is provided by the Langevin dynamics. Denoting by D the dimension of the system, by $q \in \mathcal{D}$ the positions of the particles in the system and by $p \in \mathbb{R}^D$ their momenta, the Langevin dynamics reads

$$\begin{cases} \mathrm{d}q_t = \dfrac{p_t}{m}\,\mathrm{d}t, \\ \mathrm{d}p_t = \left(-\nabla V(q_t) - \gamma\dfrac{p_t}{m}\right)\,\mathrm{d}t + \sqrt{\dfrac{2\gamma}{\beta}}\,\mathrm{d}W_t, \end{cases} \tag{2.1}$$

where $\beta > 0$ is proportional to the inverse temperature, $m > 0$ is the mass of the particles[1], $\gamma > 0$ is the friction coefficient and W_t is a standard Brownian motion in dimension D. The potential energy $V : \mathcal{D} \to \mathbb{R}$ is supposed to be a smooth function. In practice, $\mathcal{D}$ is either a compact domain with periodic boundary conditions, as for example $\mathcal{D} = (a\mathbb{T})^D$

[1] Our results can be extended to the case of any symmetric positive definite mass matrix M but we focus on the case when M is proportional to the identity matrix for simplicity.

where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the unit torus and $a > 0$ denotes the size of the simulation cell; or the unbounded space $\mathcal{D} = \mathbb{R}^D$. When $\mathrm{e}^{-\beta V}$ is integrable, the Langevin dynamics admits as a unique invariant measure the canonical measure

$$\mu(\mathrm{d}q\,\mathrm{d}p) = Z_{\beta,\mu}^{-1}\mathrm{e}^{-\beta H(q,p)}\,\mathrm{d}q\,\mathrm{d}p, \qquad H(q,p) = V(q) + \frac{|p|^2}{2m}, \tag{2.2}$$

where the partition functions $Z_{\beta,\mu}$ is a normalization coefficient.

In several situations, one is interested in solutions of Poisson equations of the form

$$-\mathcal{L}\Phi = R - \mathbb{E}_\mu[R], \tag{2.3}$$

where $\mathcal{L}$ denotes the generator of the Langevin dynamics (2.1). For instance, asymptotic variances of ergodic averages or transport coefficients can be written as

$$\int_{\mathcal{E}} -\mathcal{L}^{-1}\left(R - \mathbb{E}_\mu[R]\right) S\,\mathrm{d}\mu \tag{2.4}$$

for some functions R and S. For the asymptotic variance related to the time average of an observable R, one has $S = 2R$. For transport coefficients, R would be the system response whereas S is the conjugate response (see for instance the presentation in [104, Section 5]). In practice, quantities such as (2.4) are evaluated by Monte Carlo strategies, where the quantity of interest is rewritten as the integral of a time-dependent correlation function (the famous Green–Kubo formula), which is approximated by independent realizations of the process. In some cases however, spectral methods are used to solve the Poisson equation (2.3), see for instance [139, 100, 135, 126].

The error analysis associated with spectral Galerkin methods faces several difficulties. The most important one probably is that the generator $\mathcal{L}$ of the Langevin dynamics is not an elliptic operator, and that it is not naturally associated with a quadratic form. Many approximation results exist for elliptic operators, see for instance [35]. In the context of molecular dynamics, elliptic operators correspond to overdamped Langevin dynamics, which are effective dynamics on the positions only. A Lax-Milgram theorem holds for the quadratic form associated with the generator of the overdamped Langevin dynamics, which makes it possible to quantify the error on the solution of Poisson equations, as recently done in [1]. In contrast, the generator $\mathcal{L}$ of the Langevin dynamics (2.1) is invertible but not coercive, so that a dedicated treatment is required to obtain error estimates. This is done here by a perturbation of the proof of invertibility obtained as a corollary of the decay estimates provided in [43, 44], which builds on the theory of hypocoercivity [166]. Note that this proof applies to a large class of hypocoercive operators. In this work we restrict ourselves to the Langevin dynamics, the proofs being directly transposable for operators satisfying the hypotheses presented in [44]. Let us also mention previous results on the numerical analysis of hypocoercive operators, relying on finite element or finite difference methods, and providing finite time estimates [59, 133].

This article is organized as follows. We first recall some fundamental properties of the Langevin dynamics in Section 2.2, where we describe in particular the approach developed

in [43, 44]. We next provide in Section 2.3 general a priori error estimates for the solutions of Poisson equations (2.3). One of the key point to state such error estimates is to prove the invertibility of the generator restricted to the Galerkin space, which can be shown by adapting the hypocoercive approach of [43, 44]. We finally turn in Section 2.4 to an application to a simple, one-dimensional setting, where explicit convergence rates can be obtained. Numerical simulations are also performed to test the relevance of the bounds we provide. Some technical results are gathered in the appendices.

2.2 Convergence of the Langevin dynamics

We recall in this section useful theoretical results on exponential convergence rates for the semigroup $\mathrm{e}^{t\mathcal{L}}$ associated with the generator of the Langevin dynamics, following the methodology introduced in [43, 44] and further made precise in [70] (note that the latter works rather considered the adjoint of the generator $\mathcal{L}$, the so-called Fokker–Planck operator, but this does not change the structure of the proof, see Remark 2.1 below); see also [85] for an application to Langevin dynamics. We formulate the result both for bounded and unbounded position spaces.

In the following we consider all operators as defined on the Hilbert space $L^2(\mu)$. The adjoint of a closed operator T on $\mathrm{L}^2(\mu)$ is denoted by T^*. The scalar product and norm on $\mathrm{L}^2(\mu)$ are respectively denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$. In fact, it is convenient in many cases to work in the subspace

$$\mathrm{L}^2_0(\mu) = \left\{ \varphi \in \mathrm{L}^2(\mu) \,\middle|\, \int_{\mathcal{E}} \varphi \, \mathrm{d}\mu = 0 \right\} \tag{2.5}$$

of $L^2(\mu)$. The orthogonal projector onto $\mathrm{L}^2_0(\mu)$ is defined by

$$\forall \varphi \in \mathrm{L}^2(\mu), \qquad \Pi_0 \varphi = \varphi - \mathbb{E}_\mu(\varphi). \tag{2.6}$$

Since

$$\left(\mathrm{e}^{t\mathcal{L}}\varphi\right)(q,p) = \mathbb{E}\left(\varphi(q_t,p_t) \,\middle|\, (q_0,p_0) = (q,p)\right)$$

where the expectation is over all the realizations of the Brownian motion in (2.1), it is expected that $\mathrm{e}^{t\mathcal{L}}\varphi$ converges to $\mathbb{E}_\mu(\varphi)$. Therefore, $\mathrm{e}^{t\mathcal{L}}\varphi$ converges to 0 for $\varphi \in \mathrm{L}^2_0(\mu)$. In order to state a precise convergence result, we need some conditions on the potential V, and on the marginal measure of μ in the position variable. The marginal measures in the position and momentum variables are respectively

$$\nu(\mathrm{d}q) = Z_{\beta,\nu}^{-1} \mathrm{e}^{-\beta V(q)} \, \mathrm{d}q, \qquad \kappa(\mathrm{d}p) = \left(\frac{\beta}{2\pi m}\right)^{D/2} \mathrm{e}^{-\beta \frac{|p|^2}{2m}} \, \mathrm{d}p. \tag{2.7}$$

We denote by $\mathrm{H}^s(\nu)$ the weighted Sobolev spaces of index $s \in \mathbb{N}$ composed of functions $\varphi(q)$ of the position variables for which $\partial_q^\alpha \varphi \in \mathrm{L}^2(\mu)$ for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_D) \in \mathbb{N}^D$ such that $|\alpha| = \alpha_1 + \ldots \alpha_D \leqslant s$ (where $\partial_q^\alpha = \partial_{q_1}^{\alpha_1} \ldots \partial_{q_D}^{\alpha_D}$). The spaces $\mathrm{H}^s(\kappa)$ and $\mathrm{H}^s(\mu)$ are defined in a similar way.

Assumption 2.1. *The potential V is smooth, and the marginal measure ν satisfies a Poincaré inequality with constant $C_\nu > 0$: for any function of the positions $\varphi \in \mathrm{H}^1(\nu)$,*

$$\left\| \varphi - \int_{\mathcal{D}} \varphi \, \mathrm{d}\nu \right\|^2_{\mathrm{L}^2(\nu)} \leqslant \frac{1}{C_\nu} \|\nabla_q \varphi\|^2_{\mathrm{L}^2(\nu)}. \tag{2.8}$$

Moreover, there exist $c_1 > 0$, $c_2 \in [0, 1)$ and $c_3 > 0$ such that V satisfies

$$\Delta V \leqslant c_1 + c_2 |\nabla V|^2, \quad |\nabla^2 V| \leqslant c_3 \left(1 + |\nabla V|\right). \tag{2.9}$$

$$\liminf_{|q| \to \infty} a\beta |\nabla V(q)|^2 - \Delta V(q) > 0. \tag{2.10}$$

The precise convergence result is then the following [43, 44] (the proof is recalled in Appendix 2.5).

Theorem 2.1 (Hypocoercivity in $\mathrm{L}^2(\mu)$)**.** *Suppose that Assumption 2.1 holds. Then there exist $C > 0$ and $\lambda_\gamma > 0$ (which are explicitly computable in terms of the parameters of the dynamics, C being independent of $\gamma > 0$) such that, for any initial datum $\varphi \in \mathrm{L}^2_0(\mu)$,*

$$\forall t \geqslant 0, \qquad \left\| \mathrm{e}^{t\mathcal{L}} \varphi \right\| \leqslant C \mathrm{e}^{-\lambda_\gamma t} \|\varphi\|. \tag{2.11}$$

Moreover, the convergence rate is of order $\min(\gamma, \gamma^{-1})$: there exists $\overline{\lambda} > 0$ such that

$$\lambda_\gamma \geqslant \overline{\lambda} \min(\gamma, \gamma^{-1}).$$

Remark 2.1. *Theorem 2.1 admits a dual version in terms of probability measures. Consider an initial condition $\psi_0 \in \mathrm{L}^2(\mu)$, which represents the density with respect to μ of a probability measure $f_0 = \psi_0 \mu$. In particular,*

$$\psi_0 \geqslant 0, \qquad \int_{\mathcal{E}} \psi_0 \, d\mu = 1.$$

Then the time-evolved probability measure $f_t = \psi_t \mu$ with $\psi_t = \mathrm{e}^{t\mathcal{L}^*} \psi_0$ converges exponentially fast to μ in the following sense:

$$\forall t \geqslant 0, \qquad \|\psi_t - \mathbf{1}\| \leqslant C \mathrm{e}^{-\lambda_\gamma t} \|\psi_0\|. \tag{2.12}$$

The convergence result (2.11) can be used to deduce that $\mathcal{L}$ is invertible on $\mathrm{L}^2_0(\mu)$. We denote by $\mathcal{B}(E)$ the Banach space of bounded operators on a given Banach space E, endowed with the norm

$$\|T\|_{\mathcal{B}(E)} = \sup_{\varphi \in E \setminus \{0\}} \frac{\|T\varphi\|_E}{\|\varphi\|_E}.$$

We simply denote by $\|T\|$ the operator norm on $\mathrm{L}^2(\mu)$.

Corollary 2.1. *The operator $\mathcal{L}$ is invertible on $\mathrm{L}^2_0(\mu)$, with*

$$\mathcal{L}^{-1} = -\int_0^\infty \mathrm{e}^{t\mathcal{L}}\,\mathrm{d}t \qquad \|\mathcal{L}^{-1}\|_{\mathcal{B}(\mathrm{L}^2_0(\mu))} \leqslant \frac{C}{\overline{\lambda}}\max(\gamma,\gamma^{-1}).$$

The upper bound on the resolvent is sharp in terms of the scaling with respect to γ, as shown in [75] for $\gamma \to 0$ and [102] for $\gamma \to +\infty$; see also [95] for the case $V = 0$.

In particular, the Poisson problem (2.3) admits a unique solution $\Phi \in \mathrm{L}^2_0(\mu)$ for any observable $R \in \mathrm{L}^2(\mu)$. In order to capture the solution Φ of (2.3) numerically, one possibility is to discretize the operator $\mathcal{L}$ on a Galerkin subspace of $\mathrm{L}^2_0(\mu)$. Section 2.3 proves the convergence of this method under appropriate assumptions.

Let us conclude this section by highlighting some elements of the proof of Theorem 2.1, which will be needed to establish a convergence result similar to (2.11) when a Galerkin discretization is considered. In order to formulate the result more rigorously, we introduce the core C composed of all $\mathcal{C}^\infty$ functions with compact support. The first key element in the proof is to use a modified norm equivalent to the standard $\mathrm{L}^2(\mu)$ norm. To define this norm, the generator $\mathcal{L}$ is decomposed into a symmetric part (corresponding to the fluctuation/dissipation) and an anti-symmetric part (corresponding to Hamiltonian transport):

$$\mathcal{L} = \mathcal{L}_{\mathrm{ham}} + \gamma\mathcal{L}_{\mathrm{FD}}, \qquad \text{with} \qquad \begin{cases} \mathcal{L}_{\mathrm{ham}} = \left(\dfrac{p}{m}\right)^\top \nabla_q - \nabla V^\top \nabla_p, \\ \mathcal{L}_{\mathrm{FD}} = -\left(\dfrac{p}{m}\right)^\top \nabla_p + \dfrac{1}{\beta}\Delta_p. \end{cases} \tag{2.13}$$

With this notation, $\mathcal{L}^*_{\mathrm{ham}} = -\mathcal{L}_{\mathrm{ham}}$ while $\mathcal{L}^*_{\mathrm{FD}} = \mathcal{L}_{\mathrm{FD}}$. In fact, since

$$\nabla_p^* = -\nabla_p^\top + \beta\frac{p^\top}{m}, \qquad \nabla_q^* = -\nabla_q^\top + \beta\nabla V^\top,$$

the two parts of the generator $\mathcal{L}$ can be reformulated as

$$\mathcal{L}_{\mathrm{FD}} = -\frac{1}{\beta}\nabla_p^*\nabla_p, \qquad \mathcal{L}_{\mathrm{ham}} = \frac{1}{\beta}\left(\nabla_p^*\nabla_q - \nabla_q^*\nabla_p\right). \tag{2.14}$$

We also need the orthogonal projector in $\mathrm{L}^2_0(\mu)$ on the subspace of functions depending only on positions:

$$\forall \varphi \in \mathrm{L}^2(\mu), \qquad (\Pi_p\varphi)(q) = \int_{\mathbb{R}^D} \varphi(q,p)\,\kappa(\mathrm{d}p). \tag{2.15}$$

Definition 2.1 (Modified squared $\mathrm{L}^2(\mu)$ norm)**.** *Fix $\varepsilon \in (-1,1)$. For any function $\varphi \in \mathsf{C}$,*

$$\mathcal{H}[\varphi] = \frac{1}{2}\|\varphi\|^2 - \varepsilon\langle A\varphi,\varphi\rangle, \qquad A = \Big(1 + (\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)\Big)^{-1}(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*. \tag{2.16}$$

A more explicit expression of the operator A is provided in (2.71). Since this operator is used in the sequel to state some conditions required for the error estimates, we gather some of its properties in the following lemma.

Lemma 2.1. *It holds* $A = \Pi_p A(1-\Pi_p)$. *Moreover, for any* $\varphi \in \mathrm{L}^2(\mu)$,

$$\|A\varphi\| \leqslant \frac{1}{2}\|(1-\Pi_p)\varphi\|, \qquad \|\mathcal{L}_{\mathrm{ham}} A\varphi\| \leqslant \|(1-\Pi_p)\varphi\|.$$

In particular, the operator A is in fact bounded in $\mathrm{L}^2(\mu)$ with operator norm smaller than 1, so that $\sqrt{\mathcal{H}}$ is a norm equivalent to the canonical norm of $\mathrm{L}^2(\mu)$ for $-1 < \varepsilon < 1$:

$$\frac{1-\varepsilon}{2}\|\varphi\|^2 \leqslant \mathcal{H}[\varphi] \leqslant \frac{1+\varepsilon}{2}\|\varphi\|^2. \tag{2.17}$$

The second key element is a coercivity property enjoyed by the time-derivative of the entropy functional. Denoting by $\langle\langle \cdot, \cdot \rangle\rangle$ the scalar product associated by polarization with $\mathcal{H}$, the following result can be proved.

Proposition 2.1. *There exists* $\overline{\varepsilon} \in (0,1)$ *and* $\overline{\lambda} > 0$, *such that, by considering* $\varepsilon = \overline{\varepsilon}\min(\gamma, \gamma^{-1})$ *in* (2.16),

$$\forall \varphi \in \Pi_0 \mathsf{C}, \qquad \mathscr{D}[\varphi] := \langle\langle -\mathcal{L}\varphi, \varphi \rangle\rangle \geqslant \widetilde{\lambda}_\gamma \|\varphi\|^2, \tag{2.18}$$

with $\widetilde{\lambda}_\gamma \geqslant \overline{\lambda}\min(\gamma, \gamma^{-1})$.

This coercivity property and a Gronwall inequality then allow to conclude to the exponential convergence to 0 of $\mathcal{H}[\mathrm{e}^{t\mathcal{L}}\varphi]$, from which (2.11) follows by the norm equivalence of $\sqrt{\mathcal{H}}$ and $\|\cdot\|$.

2.3 General a priori error estimates

In order to approximate the solution of the Poisson equation (2.3), we consider a Galerkin discretization characterized by a finite dimensional subspace $V_M \subset \mathrm{L}^2(\mu)$. We present the structure of the proof of error estimates in the conformal case (*i.e.* $V_M \subset \mathrm{L}^2_0(\mu)$) for the sake of clarity. Results in the non-conformal case are presented later on. Note that the results presented in this section for the Langevin generator can be generalized to other hypocoercive generators satisfying the assumptions required in [43, 44]. For conformal discretization spaces, the approximate solution Φ_M is defined by the variational formulation

$$\begin{cases} \text{Find } \Phi_M \in V_M \text{ such that} \\ \forall \psi \in V_M, \ -\langle \psi, \mathcal{L}\Phi_M \rangle = \langle \psi, R \rangle . \end{cases} \tag{2.19}$$

Note that $\Pi_0 R$ can be replaced by R on the right-hand side since functions $\psi \in V_M$ have average 0 with respect to μ. Denoting by Π_M the projector onto V_M, the variational formulation can be rewritten as

$$-\Pi_M \mathcal{L} \Pi_M \Phi_M = \Pi_M R.$$

We first prove in this section the existence and uniqueness of the solution Φ_M of (2.19) by studying the discretized operator $-\Pi_M \mathcal{L} \Pi_M$. A dedicated study is required since the

generator $\mathcal{L}$ is invertible but not coercive on $\mathrm{L}^2_0(\mu)$, so that the Lax-Milgram theorem cannot be applied. This is a major difference with overdamped Langevin dynamics for which the discretized problem is automatically well posed when a Poincaré inequality holds true [1]. Note that there are scalar products for which the quadratic form induced by $-\mathcal{L}$ is coercive, for instance the one induced by polarization from $\mathcal{H}$ or the scalar product on $\mathrm{H}^1(\mu)$ introduced in the hypocoercivity setting considered in [75, 166]. These scalar products however depend on parameters which are not explicitly known and on the friction γ, so that they cannot be considered for numerical simulations.

We study instead the existence and the uniqueness of the solution Φ_M by a perturbation of the proof of Theorem 2.1, in two settings: the conformal case $V_M \subset \mathrm{L}^2_0(\mu)$ (see Subsection 2.3.1) and the non-conformal case $V_M \subset \mathrm{L}^2(\mu)$ but $V_M \not\subset \mathrm{L}^2_0(\mu)$ (the functions in the Galerkin basis are not of mean 0 with respect to μ, see Subsection 2.3.2).

In a second step, we prove a priori error estimates. To this end, we decompose the difference between Φ_M and the solution Φ of the equation (2.3) as the sum of two terms:

$$\Phi_M - \Phi = (\Phi_M - \Pi_M \Phi) - (1 - \Pi_M)\Phi. \tag{2.20}$$

The second term on the right-hand side is the approximation error $(1-\Pi_M)\Phi$, which depends only on the Galerkin space. We therefore postpone the study of this error to specific models (see Section 2.4.2). The first term is related to the consistency error $\eta_M = \Pi_M \mathcal{L} \Pi_M \Phi + \Pi_M R$ since $\Phi_M - \Pi_M \Phi = \left(-\Pi_M \mathcal{L} \Pi_M\right)^{-1} \eta_M$. We provide general error estimates on $\Phi_M - \Pi_M \Phi$ in Section 2.3.3. They can be made more precise in specific contexts, with explicit convergence rates; see Section 2.4.3.

We conclude the section with a practical reformulation of the variational problem (2.19) in a form more amenable to numerical computations (see Section 2.3.4).

2.3.1 Conformal case

In this section we suppose that $V_M \subset \mathrm{L}^2_0(\mu)$. The following theorem states that if the additional terms arising from the discretization in the expression of the entropy dissipation are sufficiently small, then hypocoercivity holds on the subspace V_M, and the exponential rate of convergence to 0 of the semigroup associated with $\Pi_M \mathcal{L} \Pi_M$ is uniform in M.

Theorem 2.2 (Discrete hypocoercivity)**.** *Fix $\gamma > 0$. Assume that the Galerkin space is composed of functions with mean 0 with respect to μ (i.e.$V_M \subset \mathrm{L}^2_0(\mu)$) and that*

$$\|(A + A^*)(1 - \Pi_M)\mathcal{L}\Pi_M\| \xrightarrow[M\to\infty]{} 0. \tag{2.21}$$

Then there exist $C \geqslant 1$ (independent of M, γ) and $M_0 \in \mathbb{N}$ such that, for any $M \geqslant M_0$, there is $\lambda_{\gamma,M} > 0$ for which

$$\forall \varphi \in V_M, \quad \forall t \geqslant 0, \qquad \left\| \mathrm{e}^{t\Pi_M \mathcal{L} \Pi_M} \varphi \right\| \leqslant C \mathrm{e}^{-\lambda_{\gamma,M} t} \|\varphi\|. \tag{2.22}$$

Moreover, $\lambda_{\gamma,M} \xrightarrow[M\to\infty]{} \lambda_\gamma$ where $\lambda_\gamma > 0$ is introduced in (2.11).

If in addition $\mathcal{L}_{\mathrm{FD}}$ stabilizes V_M (in the sense that $\Pi_M \mathcal{L}_{\mathrm{FD}} = \mathcal{L}_{\mathrm{FD}} \Pi_M$), then there exist $M_* \geqslant 1$ (independent of γ) such that, for any $M \geqslant M_*$, the following uniform bound holds:

$$\forall \gamma > 0, \qquad \lambda_{\gamma,M} \geqslant \overline{\lambda}_M \min(\gamma, \gamma^{-1}), \tag{2.23}$$

with $\overline{\lambda}_M \xrightarrow[M\to\infty]{} \overline{\lambda}$ where $\overline{\lambda} > 0$ is introduced in Proposition 2.1.

Let us emphasize that the condition (2.21) should be checked for the specific model under consideration; see Appendix 2.6 for an example. Note that the left hand side of (2.21) is constituted of a regularization operator $A + A^*$ applied to a residual off diagonal part of the operator $\mathcal{L}$. It is therefore expected that the norm of this operator goes to zero.

The stability of V_M by $\mathcal{L}_{\mathrm{FD}}$ is automatically ensured when the basis functions are tensor products of functions of the positions and eigenfunctions of $\mathcal{L}_{\mathrm{FD}}$ for the momentum part. The latter eigenfunctions turn out to be analytically known (they are in fact appropriately scaled Hermite functions, see Section 2.4.1), which makes it easy to conclude to (2.23).

Proof. Fix $\varphi_0 \in V_M$ and $\gamma > 0$, and consider $\varepsilon = \overline{\varepsilon} \min(\gamma, \gamma^{-1})$ as in Proposition 2.1. Introduce $\varphi_M(t) = \exp(t \Pi_M \mathcal{L} \Pi_M)\varphi_0$ and $\mathscr{H}_M(t) = \mathcal{H}[\varphi_M(t)]$. Note that the discretized generator $\Pi_M \mathcal{L} \Pi_M$ stabilizes the Galerkin space $V_M \subset \mathrm{L}^2_0(\mu)$. In particular, $\varphi_M(t) \in V_M \subset \mathrm{L}^2_0(\mu)$ for all $t \geqslant 0$ when $\varphi_0 \in V_M$. The time-derivative of the entropy functional is $\mathscr{H}_M'(t) = -\mathscr{D}_M[\varphi_M(t)]$, where $\mathscr{D}_M$ is similar to the entropy dissipation defined in (2.18) apart from two additional terms arising from the discretization. More precisely, for $\varphi \in V_M$,

$$\begin{aligned}
\mathscr{D}_M[\varphi] &= -\langle \varphi, \Pi_M \mathcal{L} \Pi_M \varphi \rangle - \varepsilon \langle A \Pi_M \mathcal{L} \Pi_M \varphi, \varphi \rangle - \varepsilon \langle A\varphi, \Pi_M \mathcal{L} \Pi_M \varphi \rangle \\
&= -\langle \varphi, \mathcal{L}\varphi \rangle - \varepsilon \langle A \Pi_M \mathcal{L}\varphi, \varphi \rangle - \varepsilon \langle \varphi, A^* \Pi_M \mathcal{L}\varphi \rangle \\
&= \mathscr{D}[\varphi] + \varepsilon \langle A(1-\Pi_M)\mathcal{L}\varphi, \varphi \rangle + \varepsilon \langle \varphi, A^*(1-\Pi_M)\mathcal{L}\varphi \rangle \\
&\geqslant \mathscr{D}[\varphi] - \varepsilon \|(A+A^*)(1-\Pi_M)\mathcal{L}\Pi_M\varphi\| \, \|\varphi\| \\
&\geqslant \Big(\widetilde{\lambda}_\gamma - \varepsilon \|(A+A^*)(1-\Pi_M)\mathcal{L}\Pi_M\| \Big) \|\varphi\|^2,
\end{aligned} \tag{2.24}$$

where the last inequality follows from Proposition 2.1. The conclusion then follows from the same reasoning as the one used at the end of Appendix 2.5 to prove Theorem 2.1, with an exponential convergence rate which is degraded uniformly in M:

$$\lambda_{\gamma,M} = \lambda_\gamma - \frac{\varepsilon}{1+\varepsilon} \|(A+A^*)(1-\Pi_M)\mathcal{L}\Pi_M\| > 0 \tag{2.25}$$

for M large enough.

Assume now that

$$\mathcal{L}_{\mathrm{FD}} \Pi_M = \Pi_M \mathcal{L}_{\mathrm{FD}} \tag{2.26}$$

so that $(1-\Pi_M)\mathcal{L}\Pi_M = (1-\Pi_M)\mathcal{L}_{\mathrm{ham}}\Pi_M$ does not depend on γ. The only γ-dependence on the right-hand side of (2.25) therefore arises from $\varepsilon = \overline{\varepsilon} \min(\gamma, \gamma^{-1})$. We then deduce the following lower bound from (2.1):

$$\lambda_{\gamma,M} \geqslant \Big(\overline{\lambda} - \overline{\varepsilon} \|(A+A^*)(1-\Pi_M)\mathcal{L}_{\mathrm{ham}}\Pi_M\| \Big) \min(\gamma, \gamma^{-1}),$$

which implies (2.23). □

Remark 2.2. *Another way to prove the hypocoercivity of the discretized generator on* $\mathrm{L}^2(\mu)$ *would be to first prove this property on* $\mathrm{H}^1(\mu)$ *(as in [166]), and then use hypoelliptic regularization [79]. This program is performed for Langevin dynamics in [75], with an emphasis on the Hamiltonian limit* $\gamma \to 0$ *(see also [104, Sections 2.3.3 and 2.3.4] for a careful analysis of the two limiting regimes* $\gamma \to 0$ *and* $\gamma \to +\infty$*). This approach introduces scalar products on* $\mathrm{H}^1(\mu)$ *depending on three coefficients* $a, b, c \in \mathbb{R}$*. The corresponding proofs are therefore more involved than the approach described here, and, more importantly, the conditions for* $\mathrm{H}^1(\mu)$ *hypocoercivity are incompatible with the conditions for* $\mathrm{L}^2(\mu)$ *regularization for the Galerkin space proposed in Section 2.4; see [144] for further precisions.*

An immediate consequence of the convergence result stated in Theorem 2.2 is the following corollary. It states that the discrete operator has a spectral gap, which does not vanish when the size of the Galerkin basis increases.

Corollary 2.2 (Discrete invertibility)**.** *For any* $M \geqslant M_0$*, the operator* $\Pi_M \mathcal{L} \Pi_M$ *is invertible on* V_M *and the following equality holds on* $\mathcal{B}(V_M)$*:*

$$(\Pi_M \mathcal{L} \Pi_M)^{-1} = -\int_0^\infty \mathrm{e}^{t \Pi_M \mathcal{L} \Pi_M} \, \mathrm{d}t \, .$$

Moreover,

$$\left\| (\Pi_M \mathcal{L} \Pi_M)^{-1} \right\|_{\mathcal{B}(V_M)} \leqslant \frac{C}{\lambda_{\gamma,M}}.$$

In particular, when $\mathcal{L}_{\mathrm{FD}}$ stabilizes V_M, the dependence on γ of the resolvent bound can be made explicit thanks to (2.23). Corollary 2.2 shows that the Galerkin problem (2.19) admits a unique solution, denoted by $\Phi_M = -\left(\Pi_M \mathcal{L} \Pi_M\right)^{-1} \Pi_M R$.

2.3.2 Non-conformal case

In practice the assumption $V_M \subset \mathrm{L}^2_0(\mu)$ is constraining since it may not be convenient to construct a basis of $\mathrm{L}^2_0(\mu)$ which is orthogonal for the associated scalar product. It seems easier in many situations to consider bases which are orthonormal on $\mathrm{L}^2(\mu)$ rather than $\mathrm{L}^2_0(\mu)$ (as we do here for the application treated in Section 2.4). Moreover, it may be preferable in practice to create bases adapted to the operators ∇_q, ∇_q^*, ∇_p and ∇_q^* in order to simplify the algebra involved in the computation of the elements of the rigidity matrix. For these two reasons basis functions are rarely of mean 0 with respect to μ in the literature, see for instance [100, 135, 1] for recent examples. We therefore need to extend the results of Section 2.3 to the non-conformal case $V_M \not\subset \mathrm{L}^2_0(\mu)$.

Now, the generator $\mathcal{L}$ is invertible on $\mathrm{L}^2_0(\mu)$ (by Corollary 2.1) but not on $\mathrm{L}^2(\mu)$ since $\mathcal{L}\mathbf{1} = 0$. The purpose of this subsection is to show how this degeneracy can be dealt with by introducing a Lagrangian formulation. We start by applying Theorem 2.2 to the Galerkin space $V_{M,0} = V_M \cap \mathrm{L}^2_0(\mu)$, whose associated orthogonal projector we denote by $\Pi_{M,0}$. The

issue is to control the solution in the direction associated with the function

$$u_M = \frac{\Pi_M \mathbf{1}}{\|\Pi_M \mathbf{1}\|} \in V_M, \tag{2.27}$$

which is not of zero mean. In this setting the approximate solution Φ_M is defined by the variational formulation

$$\begin{cases} \text{Find } \Phi_M \in V_{M,0} \text{ such that} \\ \forall \psi \in V_{M,0},\ -\langle \psi, \mathcal{L}\Phi_M \rangle = \langle \psi, R \rangle, \end{cases} \tag{2.28}$$

which can be rewritten as

$$-\Pi_{M,0}\mathcal{L}\Pi_{M,0}\Phi_M = \Pi_{M,0}R.$$

The precise result is the following.

Corollary 2.3 (Non-conformal Galerkin method)**.** *Assume that the Galerkin space V_M is such that* (2.21) *holds and additionally that*

$$\|\mathcal{L}^* u_M\| \xrightarrow[M\to\infty]{} 0. \tag{2.29}$$

Then there exist $C \geqslant 1$ (independent of M, γ) and $M_0 \geqslant 1$ such that, for any $M \geqslant M_0$, the operator $\Pi_{M,0}\mathcal{L}\Pi_{M,0}$ is invertible on V_M and there is $\widetilde{\lambda}_{\gamma,M} > 0$ for which

$$\left\| \left(\Pi_{M,0}\mathcal{L}\Pi_{M,0}\right)^{-1} \right\|_{\mathcal{B}(V_{M,0})} \leqslant \frac{C}{\widetilde{\lambda}_{\gamma,M}},$$

with $\widetilde{\lambda}_{\gamma,M} \xrightarrow[M\to\infty]{} \lambda_\gamma > 0$ where $\lambda_\gamma > 0$ is introduced in (2.11)*.*

If in addition $\mathcal{L}_{\mathrm{FD}}$ stabilizes V_M, then there exist $M_* \geqslant 1$ (independent of γ) such that, for any $M \geqslant M_*$, the following uniform bound holds:

$$\forall \gamma > 0, \qquad \widetilde{\lambda}_{\gamma,M} \geqslant \overline{\lambda}_M \min(\gamma, \gamma^{-1}),$$

with $\overline{\lambda}_M \xrightarrow[M\to\infty]{} \overline{\lambda}$ where $\overline{\lambda} > 0$ is introduced in Proposition 2.1.

Proof. Let us first decompose V_M as an orthogonal direct sum:

$$V_M = V_{M,0} \oplus \mathbb{R}u_M.$$

Denoting by Π_{u_M} the orthogonal projection onto $\mathbb{R}u_M$, it then holds $\Pi_M = \Pi_{M,0} + \Pi_{u_M}$. We can now show how the hypotheses on Π_M allow to apply Theorem 2.2 on the Galerkin space $V_{M,0}$. We follow the proof of Theorem 2.2 until (2.24), replacing Π_M with $\Pi_{M,0}$. It then suffices to prove that the following term is of order $\|\varphi\|^2$ for any $\varphi \in V_{M,0}$:

$$\langle (A + A^*)(1 - \Pi_{M,0})\mathcal{L}\varphi, \varphi \rangle = \langle (A + A^*)(1 - \Pi_M)\mathcal{L}\varphi, \varphi \rangle + \langle (A + A^*)\Pi_{u_M}\mathcal{L}\varphi, \varphi \rangle.$$

The first term on the right-hand side can be dealt with as in the proof of Theorem 2.2,

making use of (2.21). For the second one, we remark that

$$\langle (A+A^*)\Pi_{u_M}\mathcal{L}\varphi,\varphi\rangle = \langle \mathcal{L}\varphi,u_M\rangle \langle (A+A^*)u_M,\varphi\rangle ,$$

so that, using $\|A\| = \|A^*\| \leqslant 1/2$ (from Lemma 2.1):

$$|\langle (A+A^*)\Pi_{u_M}\mathcal{L}\varphi,\varphi\rangle| \leqslant \|\varphi\| \, \|\mathcal{L}^* u_M\| \, \|(A+A^*)u_M\| \, \|\varphi\| \leqslant \|\mathcal{L}^* u_M\| \, \|\varphi\|^2. \tag{2.30}$$

Plugging this additional term into the bound (2.24) obtained in the conformal case, it follows

$$\mathscr{D}_M[\varphi] \geqslant \left(\widetilde{\lambda}_\gamma - \varepsilon \, \|(A+A^*)(1-\Pi_M)\mathcal{L}\Pi_M\| - \varepsilon \|\mathcal{L}^* u_M\|\right) \|\varphi\|^2.$$

We can then conclude to the exponential convergence of the semi-group, with rate

$$\widetilde{\lambda}_{\gamma,M} = \lambda_\gamma - \frac{\varepsilon}{1+\varepsilon}\left(\|(A+A^*)(1-\Pi_M)\mathcal{L}\Pi_M\| + \|\mathcal{L}^* u_M\|\right) > 0, \tag{2.31}$$

when M is sufficiently large. The remainder of the proof follows the lines of the end of the proof of Theorem 2.2. □

Corollary 2.3 implies that the following saddle-point formulation is well-posed.

Proposition 2.2 (Saddle-point formulation)**.** *Assume that* (2.21) *and* (2.29) *hold. Then, for any $R \in \mathrm{L}^2(\mu)$, there exist a unique $\Phi_M \in V_M$ and a unique $\alpha_M \in \mathbb{R}$ such that*

$$\begin{cases} -\Pi_M \mathcal{L} \Pi_M \Phi_M + \alpha_M u_M = \Pi_M R, \\ \langle \Phi_M, u_M \rangle = 0. \end{cases} \tag{2.32}$$

Note that the unique solution Φ_M in fact belongs to $V_{M,0}$ since $\langle \Phi_M, u_M\rangle = 0$. Moreover, R does not need to be of mean 0 with respect to μ thanks to the term $\alpha_M u_M$ on the left-hand side of the first equality in (2.32). We show in the next subsection that Φ_M actually converges to the solution of the Poisson equation (2.3) with right-hand side $\Pi_0 R$.

Proof. Consider $R \in \mathrm{L}^2(\mu)$. In view of Corollary 2.3, there exists a unique $\Phi_M \in V_{M,0}$ such that

$$-\Pi_{M,0}\mathcal{L}\Phi_M = \Pi_{M,0} R.$$

Recalling that $\Pi_{M,0} = \Pi_M - \Pi_{u_M}$ it follows that

$$-\Pi_M \mathcal{L}\Pi_M \Phi_M + \Pi_{u_M}\left(\mathcal{L}\Pi_M\Phi_M + R\right) = \Pi_M R,$$

which leads to the saddle-point formulation (2.32) upon introducing the Lagrange multiplier $\alpha_M = \langle u_M, \mathcal{L}\Pi_M\Phi_M + R\rangle$ (which is uniquely defined). □

The system (2.32) can be reformulated as

$$\widetilde{\mathcal{L}}_M \begin{pmatrix} \Phi_M \\ \alpha_M \end{pmatrix} = \begin{pmatrix} \Pi_M R \\ 0 \end{pmatrix}, \tag{2.33}$$

where the Lagrangian operator $\widetilde{\mathcal{L}}_M$ on $V_M \times \mathbb{R}$ reads

$$\widetilde{\mathcal{L}}_M \begin{pmatrix} \varphi \\ \alpha \end{pmatrix} = \begin{pmatrix} -\Pi_M \mathcal{L} \Pi_M \varphi + \alpha u_M \\ \langle \varphi, u_M \rangle \end{pmatrix}. \tag{2.34}$$

Let us conclude this section by providing an estimate on the resolvent bound of $\widetilde{\mathcal{L}}_M$. This estimate is used in Section 2.3.4 to show that the matrix reformulation of (2.32) is well-posed, and in fact enjoys a good conditioning. Let us first prove that the Lagrangian operator $\widetilde{\mathcal{L}}_M$ is invertible on $V_M \times \mathbb{R}$ for $M \geqslant M_0$ (with M_0 the integer considered in Corollary 2.3). This is done by proving that the equation $\widetilde{\mathcal{L}}_M(\varphi, \alpha) = (\psi, s)$ admits a unique solution for an arbitrary element $(\psi, s) \in V_M \times \mathbb{R}$. Note that

$$\widetilde{\mathcal{L}}_M \begin{pmatrix} \varphi \\ \alpha \end{pmatrix} = \begin{pmatrix} \psi \\ s \end{pmatrix} \tag{2.35}$$

is equivalent to

$$\widetilde{\mathcal{L}}_M \begin{pmatrix} \varphi - s\, u_M \\ \alpha \end{pmatrix} = \begin{pmatrix} \psi - s \Pi_M \mathcal{L} u_M \\ 0 \end{pmatrix}.$$

For the latter equality to hold true, the function $\phi_{s,M} = \varphi - s u_M$ must satisfy the Poisson equation

$$-\Pi_M \mathcal{L} \Pi_M \phi_{s,M} = \psi - s \Pi_M \mathcal{L} u_M - \alpha u_M, \qquad \langle \phi_{s,M}, u_M \rangle = 0. \tag{2.36}$$

Then, $\phi_{s,M} \in V_{M,0}$ so that $\Pi_M \mathcal{L} \Pi_M \phi_{s,M} = \Pi_M \mathcal{L} \Pi_{M,0} \phi_{s,M} = \Pi_{M,0} \mathcal{L} \Pi_{M,0} \phi_{s,M} + \langle \mathcal{L} \phi_{s,M}, u_M \rangle\, u_M$. Therefore, (2.36) can be reformulated as

$$-\Pi_{M,0} \mathcal{L} \Pi_{M,0} \phi_{s,M} = \psi - s \Pi_M \mathcal{L} u_M + (\langle \mathcal{L} \phi_{s,M}, u_M \rangle\, u_M - \alpha) u_M, \qquad \langle \phi_{s,M}, u_M \rangle = 0.$$

Since $\Pi_{M,0} \mathcal{L} \Pi_{M,0}$ is invertible on $V_{M,0}$, the equation (2.36) admits a unique solution in $V_{M,0}$ if and only if the right-hand side of the above Poisson equation is in $V_{M,0}$, which is the case if and only if

$$\alpha = \langle u_M, \mathcal{L} \Pi_M (\varphi - s u_M) + (\psi - s \Pi_M \mathcal{L} u_M) \rangle\,. \tag{2.37}$$

This proves the existence and uniqueness of the solution to (2.35) since α and $\phi_{s,M}$ are completely identified through (2.37) and

$$\varphi = s u_M + \left(-\Pi_{M,0} \mathcal{L} \Pi_{M,0}\right)^{-1} \Pi_{M,0} \left(\psi - s \Pi_M \mathcal{L} u_M\right). \tag{2.38}$$

This allows to conclude that $\widetilde{\mathcal{L}}_M$ is invertible on $V_M \times \mathbb{R}$. Moreover, using Corollary 2.3,

$$\|\varphi\|^2 \leqslant s^2 + \left(\frac{C}{\widetilde{\lambda}_{\gamma,M}}\right)^2 \left(\|\psi\| + \|\mathcal{L} u_M\|\, |s|\right)^2,$$

and, in view of (2.37)-(2.38),

$$|\alpha| \leqslant \left(1 + \|\mathcal{L}^* u_M\| \frac{C}{\widetilde{\lambda}_{\gamma,M}}\right) \left(\|\psi\| + \|\mathcal{L} u_M\|\, |s|\right). \tag{2.39}$$

Therefore, endowing $V_M \times \mathbb{R}$ with the norm associated with the canonical scalar product, the following resolvent bound holds:

$$\left\|\widetilde{\mathcal{L}}_M^{-1}\right\|^2_{\mathcal{B}(V_M\times\mathbb{R})} \leqslant 1 + \left[\left(\frac{C}{\widetilde{\lambda}_{\gamma,M}}\right)^2 + \left(1 + \frac{C}{\widetilde{\lambda}_{\gamma,M}}\|\mathcal{L}^* u_M\|\right)^2\right](1 + \|\mathcal{L}u_M\|^2). \tag{2.40}$$

In fact, the operators $\widetilde{\mathcal{L}}_M^{-1}$ are bounded uniformly in $M \geqslant M_0$, since the upper bound on $\left\|\widetilde{\mathcal{L}}_M^{-1}\right\|_{\mathcal{B}(V_M\times\mathbb{R})}$ tends to $\sqrt{2 + (C/\lambda_\gamma)^2}$ as $M \to +\infty$.

2.3.3 Consistency error

We study in this section the error $\|\Phi_M - \Pi_{M,0}\Phi\|$ associated to the consistency error $\eta_{M,0} = \Pi_{M,0}\mathcal{L}\Phi_M + \Pi_{M,0}R$, sticking to the non-conformal case since this setting is the most appropriate for actual applications. With some abuse of terminology, we simply call $\|\Phi_M - \Pi_{M,0}\Phi\|$ the consistency error.

As in (2.20), the error can be decomposed as

$$\Phi_M - \Phi = (\Phi_M - \Pi_{M,0}\Phi) - (1 - \Pi_{M,0})\Phi. \tag{2.41}$$

Very similar results are obtained in the conformal case upon replacing $\Pi_{M,0}$ with Π_M. Moreover, we do not suppose in this section that R has mean 0 with respect to μ, but consider the Poisson problem (2.3) with R replaced by $\Pi_0 R$:

$$-\mathcal{L}\Phi = \Pi_0 R. \tag{2.42}$$

The solution Φ is approximated by the solution of the Poisson equation

$$-\Pi_{M,0}\mathcal{L}\Pi_{M,0}\Phi = \Pi_{M,0}R. \tag{2.43}$$

which is well-posed in view of Corollary 2.3.

Theorem 2.3. *Assume that* (2.21) *and* (2.29) *hold. Then the consistency error between the unique solution* $\Phi \in \mathrm{L}^2_0(\mu)$ *of* (2.42) *and the approximate solution* $\Phi_M \in V_{M,0}$ *of* (2.43) *can be bounded by*

$$\|\Phi_M - \Pi_{M,0}\Phi\| \leqslant \frac{C}{\widetilde{\lambda}_{\gamma,M}}\left(\|\Pi_M\mathcal{L}(1-\Pi_M)\Phi\| + \|\mathcal{L}u_M\|\|\Phi\|\right), \tag{2.44}$$

where $C, \widetilde{\lambda}_{\gamma,M}$ *are the constants introduced in Corollary 2.3.*

The extra term $\|\mathcal{L}u_M\|\|\Phi\|$ on the right-hand side of (2.44) arises from the fact that the Galerkin space is not conformal. It would not be present for conformal spaces.

Proof. Upon applying $\Pi_{M,0}$ to both sides of (2.42), it holds

$$-\Pi_{M,0}\mathcal{L}\Pi_{M,0}\Phi = \Pi_{M,0}R + \Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi.$$

After subtraction with (2.43), it follows

$$\Pi_{M,0}\mathcal{L}\Pi_{M,0}(\Phi_M - \Pi_{M,0}\Phi) = \Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi. \tag{2.45}$$

Therefore, using Corollary 2.3,

$$\begin{aligned}\|\Phi_M - \Pi_{M,0}\Phi\|_{\mathrm{L}^2(\mu)} &= \left\|(\Pi_{M,0}\mathcal{L}\Pi_{M,0})^{-1}\Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi\right\|_{\mathrm{L}^2(\mu)} \\ &\leqslant \left\|(\Pi_{M,0}\mathcal{L}\Pi_{M,0})^{-1}\right\|_{\mathcal{B}(V_{M,0})} \|\Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi\|_{\mathrm{L}^2(\mu)} \\ &\leqslant \frac{C}{\widetilde{\lambda}_{\gamma,M}} \|\Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi\|_{\mathrm{L}^2(\mu)}\,.\end{aligned} \tag{2.46}$$

Moreover

$$\begin{aligned}\|\Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi\|_{\mathrm{L}^2(\mu)} &\leqslant \|\Pi_M\mathcal{L}(1-\Pi_M+\Pi_{u_M})\Phi\|_{\mathrm{L}^2(\mu)} \\ &\leqslant \|\Pi_M\mathcal{L}(1-\Pi_M)\Phi\|_{\mathrm{L}^2(\mu)} + \|\Pi_M\mathcal{L}\Pi_{u_M}\Phi\|_{\mathrm{L}^2(\mu)} \\ &\leqslant \|\Pi_M\mathcal{L}(1-\Pi_M)\Phi\|_{\mathrm{L}^2(\mu)} + \|\Pi_M\mathcal{L}u_M\|_{\mathrm{L}^2(\mu)}\,|\langle\Phi,u_M\rangle|\,,\end{aligned} \tag{2.47}$$

which allows to conclude. □

There are several ways to bound the right-hand side of (2.44). It is difficult to state general results, and the strategy to be used depends on the model under consideration. One straightforward manner is to write

$$\|\Pi_M\mathcal{L}(1-\Pi_M)\Phi\|_{\mathrm{L}^2(\mu)} \leqslant \|\Pi_M\mathcal{L}(1-\Pi_M)\|_{\mathcal{B}(\mathrm{H}^2(\mu),\mathrm{L}^2(\mu))}\,\|(1-\Pi_M)\Phi\|_{\mathrm{H}^2(\mu)}\,,$$

and make use of the following (possible quite crude) bound which is independent of M:

$$\|\Pi_M\mathcal{L}(1-\Pi_M)\|_{\mathcal{B}(\mathrm{H}^2(\mu),\mathrm{L}^2(\mu))} \leqslant \|\mathcal{L}\|_{\mathcal{B}(\mathrm{H}^2(\mu),\mathrm{L}^2(\mu))}\,.$$

It remains then to show that the approximation error measured in the $\mathrm{H}^2(\mu)$ norm goes to zero. Possibly sharper estimates can be obtained by writing that

$$\|\Pi_M\mathcal{L}(1-\Pi_M)\Phi\|_{\mathrm{L}^2(\mu)} \leqslant \|\Pi_M\mathcal{L}(1-\Pi_M)\|_{\mathcal{B}(\mathrm{L}^2(\mu))}\,\|(1-\Pi_M)\Phi\|_{\mathrm{L}^2(\mu)}\,, \tag{2.48}$$

and showing that $\|\Pi_M\mathcal{L}(1-\Pi_M)\|_{\mathcal{B}(\mathrm{L}^2(\mu))}$ does not go too fast to infinity as M goes to infinity. We can then conclude in the case when the approximation error vanishes sufficiently fast in $\mathrm{L}^2(\mu)$. This is the path we follow in Section 2.6.

Remark 2.3. *We expect the operator $\Pi_{M,0}\mathcal{L}\Pi_{M,0}$ to be larger in a certain sense than $\Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})$ in $\mathrm{L}^2(\mu)$, so that (2.45) suggests that the consistency error is smaller than the approximation error $\|(1-\Pi_M)\Phi\|$. This is indeed what we observe in the numerical experiments we present in Figure 2.1. This shows that the way we bound the consistency error is probably not as sharp as it could be.*

2.3.4 Matrix conditioning and linear systems

We introduce in this section the linear system associated with the practical implementation of either the Galerkin formulation (2.19) in the conformal case $V_M \subset \mathrm{L}^2_0(\mu)$, or of (2.32) in the non-conformal case $V_M \subset \mathrm{L}^2(\mu)$ but $V_M \not\subset \mathrm{L}^2_0(\mu)$. In any case, we denote by $(e_j)_{1\leqslant j\leqslant M}$ an orthogonal basis of the Galerkin space V_M, assumed to be of dimension M.

Conformal case. The weak formulation (2.19) can be equivalently reformulated as the linear system

$$\mathbf{L}_M \mathbf{X}_M = \mathbf{Y}_M, \tag{2.49}$$

where

$$\forall 1 \leqslant i,j \leqslant M, \qquad (\mathbf{L}_M)_{i,j} = \langle e_i, -\mathcal{L} e_j \rangle\,, \quad (\mathbf{X}_M)_i = \langle \Phi_M, e_i \rangle\,, \quad (\mathbf{Y}_M)_i = \langle R, e_i \rangle\,.$$

When the assumptions of Theorem 2.2 hold, (2.49) admits a unique solution, so that $\mathbf{L}_M$ is invertible. Moreover $\|\mathbf{L}_M^{-1}\| \leqslant C/\lambda_{\gamma,M}$ is bounded uniformly in M for $M \geqslant M_0$. The linear system is therefore well-conditioned, and can be solved efficiently using any solver adapted to non-symmetric problems.

Non-conformal case. We suppose that the assumptions of Corollary 2.3 hold. Let us introduce the vector $\mathbf{U}_M \in \mathbb{R}^M$ corresponding to $u_M \in V_M$:

$$\forall 1 \leqslant i \leqslant M, \qquad (\mathbf{U}_M)_i = \langle u_M, e_i \rangle = \frac{\langle \mathbf{1}, e_i \rangle}{\|\Pi_M \mathbf{1}\|}.$$

Then the saddle-point problem (2.33) is equivalent to

$$\begin{cases} \mathbf{L}_M \mathbf{X}_M + \lambda \mathbf{U}_M = \mathbf{Y}_M, \\ \mathbf{U}_M^\top \mathbf{X}_M = 0, \end{cases}$$

with the same definition for $\mathbf{L}_M$ and $\mathbf{Y}_M$ as in the conformal case. With

$$\widehat{\mathbf{L}}_M = \left(\begin{array}{c|c} \mathbf{L}_M & \mathbf{U}_M \\ \hline \mathbf{U}_M^\top & 0 \end{array}\right), \qquad \widehat{\mathbf{X}}_M = \left(\begin{array}{c} \mathbf{X}_M \\ \hline \lambda \end{array}\right), \qquad \widehat{\mathbf{Y}}_M = \left(\begin{array}{c} \mathbf{Y}_M \\ \hline 0 \end{array}\right), \tag{2.50}$$

the saddle-point problem can finally be rewritten as

$$\widehat{\mathbf{L}}_M \widehat{\mathbf{X}}_M = \widehat{\mathbf{Y}}_M.$$

Proposition 2.2 and (2.40) imply that $\widehat{\mathbf{L}}_M$ is invertible, with $\left\|\widehat{\mathbf{L}}_M^{-1}\right\|$ uniformly bounded in M for $M \geqslant M_0$. This proves that the matrix $\widehat{\mathbf{L}}_M$ does not have vanishing eigenvalues, in contrast to $\mathbf{L}_M$ (since $\mathbf{L}_M \mathbf{U}_M \xrightarrow[M\to\infty]{} 0$). Therefore the linear system $\widehat{\mathbf{L}}_M \widehat{\mathbf{X}}_M = \widehat{\mathbf{Y}}_M$ can be

solved as efficiently as in the conformal case. In the following we choose to use a sparse LU factorization.

Remark 2.4. *Let us conclude this section with some criteria discriminating a good Galerkin space, and more generally a good function basis. Anticipating on the analysis of Section 2.4.1, a standard choice is to use tensorized bases. The difficult part is to find a basis to describe the position dependence of the function of consideration. This requires considering the following points:*

- *approximation errors and consistency errors should be small. It should be checked in particular that condition* (2.21) *holds and that the norm of the operator* $\Pi_M \mathcal{L}(1-\Pi_M)$ *does not grow too fast.*
- *the implementation is easier if the space is conformal, since it avoids the computation of* $\mathbf{U}_M$ *using integral quadratures.*
- *when the basis is non-orthogonal, the Gram matrix should be inverted. The latter can be ill conditioned, leading to numerically instability, specifically for unbounded position spaces.*

2.4 Application to a simple one-dimensional system

We present in this section an application of the theory developed in Section 2.3 to a specific example, described in Section 2.4.1 together with the Galerkin basis used to discretize the generator. This allows us to prove explicit convergence rates for the approximation error (Section 2.4.2) and the consistency error (Section 2.4.3). For the latter error, we have to further specify the potential in order to check the assumptions ensuring the hypocoercivity of the discretized generator. The final, global error estimate is summarized in (2.61). The technical proofs of some claims and bounds are postponed to Appendix 2.6. We finally present in Section 2.4.4 some numerical results illustrating the predicted error bounds.

2.4.1 Description of the system and the Galerkin space

We consider a single particle in a one-dimensional periodic potential: $D = 1$, $m = 1$ and $\mathcal{D} = 2\pi\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. The Galerkin space is constructed using the spectral tensor basis

$$e_{k,\ell}(q,p) = G_k(q)H_\ell(p),$$

where $0 \leqslant k < 2K-1$ and $0 \leqslant \ell < L$. Compared to the notation of Section 2.3, the basis size $M = (2K-1)L$ depends on two parameters K, L, which both have to go to infinity for the convergence results to hold. In this section we prefer the index KL instead of M, denoting thus V_{KL}, Π_{KL}, Φ_{KL},... In the remainder of this section we describe our choices for G_k and H_ℓ.

Note that the size of the matrix, namely the number of tensorized basis elements, increases exponentially with the dimension of the system. In larger dimension one could

consider resorting to tensor formats [73], as is done for the high-dimensional Schrödinger equation in [172], carefully making use of the symmetries and of the structure of the equation.

Weighted Fourier basis (G_k)**.** Fourier modes provide a natural basis to approximate periodic functions, such as functions of the positions here. Since the measure appearing in the scalar product is ν, we consider in fact the following $L^2(\nu)$-orthonormal modes:

$$\begin{aligned} G_0(q) &= \sqrt{\frac{Z_{\beta,\nu}}{2\pi}}\,\mathrm{e}^{\beta V(q)/2}, \\ G_{2k}(q) &= \sqrt{\frac{Z_{\beta,\nu}}{\pi}}\cos(kq)\,\mathrm{e}^{\beta V(q)/2}, \qquad k \geqslant 1, \\ G_{2k-1}(q) &= \sqrt{\frac{Z_{\beta,\nu}}{\pi}}\sin(kq)\,\mathrm{e}^{\beta V(q)/2}, \qquad k \geqslant 1. \end{aligned} \tag{2.51}$$

Note that the functions G_k for $k \geqslant 1$ do not have mean 0 with respect to ν (except for very specific potentials such as $V = 0$). The spanned discretization space is thus non-conformal: $V_{KL} \not\subset \mathrm{L}^2_0(\mu)$.

Hermite functions basis (H_ℓ)**.** Since the marginal measure κ in the momentum variables is Gaussian with variance β^{-1}, we consider the following orthonormal Hermite modes for $\ell \in \mathbb{N}$:

$$H_\ell(p) = \frac{1}{\sqrt{\ell!}}\widetilde{H}_\ell\left(\sqrt{\beta}p\right), \qquad \widetilde{H}_\ell(y) = (-1)^\ell \mathrm{e}^{\frac{y^2}{2}}\frac{d^\ell}{dy^\ell}\left(\mathrm{e}^{-\frac{y^2}{2}}\right).$$

They are well suited to our problem since they are the eigenfunctions of the symmetric part $\mathcal{L}_{\mathrm{FD}} = -\beta^{-1}\partial_p^*\partial_p$ of the generator. Indeed,

$$\forall \ell \in \mathbb{N}, \qquad \partial_p H_\ell = \sqrt{\beta\ell}H_{\ell-1} \qquad \text{and} \quad \partial_p^* H_\ell = \sqrt{\beta(\ell+1)}H_{\ell+1}, \tag{2.52}$$

so that

$$\forall \ell \in \mathbb{N}, \qquad \mathcal{L}_{\mathrm{FD}} H_\ell = -\ell H_\ell. \tag{2.53}$$

Remark 2.5. *The basis we consider is similar to the one used in [139] and [126], where the modes in position are the standard Fourier modes. The latter modes are orthogonal for the uniform measure on the compact position space $\mathcal{D}$ rather than on $\mathrm{L}^2(\nu)$. Therefore, the scalar product used in Subsection 2.3.4 should be replaced with the scalar product associated with the measure $\widetilde{\mu}(\mathrm{d}q\,\mathrm{d}p) = |\mathcal{D}|^{-1}\kappa(\mathrm{d}p)\,\mathrm{d}q$. The results of Section 2.3 could be adapted to this scalar product since the measures μ and $\widetilde{\mu}$ are equivalent. Note that the discretization based on the standard Fourier modes is a conformal one since one of the tensorized modes is proportional to* $\mathbf{1}$*, which simplifies the implementation. It is however not generalizable to unbounded position spaces because the uniform measure is not normalizable. An interesting question, not considered in this work, is to quantify the relative performances of the approaches based on orthonormal bases either on $\mathrm{L}^2(\mu)$ or $\mathrm{L}^2(\widetilde{\mu})$.*

Rigidity matrix. In order to give the expression of the rigidity matrix, we introduce, for a Fourier basis of $2K-1$ weighted Fourier modes, the matrix $\mathbf{Q}$ with entries

$$\mathbf{Q}_{k,k'} = \langle G_k, \partial_q G_{k'} \rangle_{\mathrm{L}^2(\nu)}, \tag{2.54}$$

and, for L Hermite modes, the matrix $\mathbf{P}$ with entries

$$\mathbf{P}_{\ell,\ell'} = \langle H_\ell, \partial_p H_{\ell'} \rangle_{\mathrm{L}^2(\kappa)} = \left\langle H_\ell, \sqrt{\beta \ell'} H_{\ell'-1} \right\rangle_{\mathrm{L}^2(\kappa)} = \sqrt{\beta \ell'} \delta_{\ell,\ell'-1}. \tag{2.55}$$

Note that $\mathbf{P}$ is sparse in view of (2.52). The matrix $\mathbf{Q}$ is, on the other hand, dense in general, except when V is a trigonometric polynomial. In the following, we choose $V(q) = 1-\cos(q)$ in order for $\mathbf{Q}$ to be tridiagonal. For a general, smooth potential V, $\mathbf{Q}$ would be dense but with coefficients which decay fast away from the diagonal.

The rigidity matrix which appears on the left-hand side of (2.49) has entries (for $0 \leqslant k \leqslant 2K-2$ and $0 \leqslant \ell \leqslant L-1$)

$$\begin{aligned}
\mathbf{L}_{k\ell,k'\ell'} &= \langle e_{k\ell}, -\mathcal{L} e_{k'\ell'} \rangle \\
&= -\beta^{-1} \left[\left\langle G_k H_\ell, \partial_q \partial_p^* G_{k'} H_{\ell'} \right\rangle - \left\langle G_k H_\ell, \partial_q^* \partial_p G_{k'} H_{\ell'} \right\rangle - \gamma \left\langle G_k H_\ell, \partial_p^* \partial_p G_{k'} H_{\ell'} \right\rangle \right] \\
&= -\beta^{-1} \mathbf{Q}_{k,k'} \mathbf{P}_{\ell',\ell} + \beta^{-1} \mathbf{Q}_{k',k} \mathbf{P}_{\ell,\ell'} + \gamma \mathbf{I}_{k,k'} \mathbf{N}_{\ell,\ell'},
\end{aligned}$$

where $\mathbf{I}_{k,k'} = \delta_{k,k'}$ and $\mathbf{N}_{\ell,\ell'} = \ell\, \delta_{\ell,\ell'}$. In practice we transform these tensors into matrices by a hashing function $\zeta : (k,\ell) \to \zeta(k,\ell) \in \mathbb{N}$. The matrix $\mathbf{L}$ is then of size $(2K-1)L$.

2.4.2 Approximation error for the tensor basis

We define the projectors Π_K^q and Π_L^p by

$$\Pi_K^q \varphi = \sum_{k=0}^{2K-2} \langle \varphi, G_k \rangle\, G_k, \qquad \Pi_L^p \varphi = \sum_{\ell=0}^{L-1} \langle \varphi, H_\ell \rangle\, H_\ell.$$

Their complements are $\Pi_K^{q\perp} = 1 - \Pi_K^q$ and $\Pi_L^{p\perp} = 1 - \Pi_L^p$. With this notation, the projector onto the Galerkin space is $\Pi_{KL} = \Pi_K^q \Pi_L^p$. The study of the approximation error $(1-\Pi_{KL})\Phi$ is performed by first estimating the error arising from the projection Π_K^q (see Lemma 2.2), and then the error arising from Π_L^p (see Lemma 2.3). The conclusion follows by remarking that

$$0 \leqslant 1 - \Pi_{KL} = 1 - \Pi_K^q + \Pi_K^q (1 - \Pi_L^p) \leqslant \Pi_K^{q\perp} + \Pi_L^{p\perp}, \tag{2.56}$$

see Proposition 2.3.

Lemma 2.2. *Assume that V is smooth. Then, for any $s \in \mathbb{N}$, there exists $M_s \in \mathbb{R}_+$ such that*

$$\forall \varphi \in \mathrm{H}^s(\nu), \quad \forall K \geqslant 1, \qquad \|\varphi - \Pi_K^q \varphi\|_{\mathrm{L}^2(\nu)} \leqslant \frac{M_s}{K^s} \|\varphi\|_{\mathrm{H}^s(\nu)}.$$

Proof. For $\varphi \in \mathrm{H}^s(\nu)$, we introduce $\widetilde{\varphi} = Z_{\beta,\nu}^{-1/2} \mathrm{e}^{-\beta V/2} \varphi \in \mathrm{L}^2(\mathrm{d}q)$, as well as the flat Fourier

basis $\widetilde{G}_k = Z_{\beta,\nu}^{-1/2} \mathrm{e}^{-\beta V/2} G_k$ which is orthonormal on $\mathrm{L}^2([0,2\pi])$. Since $\mathcal{D} = 2\pi\mathbb{T}$ is compact, $\mathrm{H}^s(\nu) = \mathrm{H}^s(\mathrm{d}q)$ for any $s \in \mathbb{N}$ and there exists $M_s \in \mathbb{R}_+$ such that

$$\|\partial_q^s \varphi\|_{\mathrm{L}^2(\mathrm{d}q)} \leqslant M_s \|\varphi\|_{\mathrm{H}^s(\nu)}. \tag{2.57}$$

By the Bessel-Parseval inequality,

$$\begin{aligned}
\|\varphi - \Pi_K^q \varphi\|_{\mathrm{L}^2(\nu)}^2 &= \sum_{k \geqslant 2K-1} \langle \varphi, G_k \rangle^2 = \sum_{k \geqslant 2K-1} \left(\int_0^{2\pi} \widetilde{\varphi}\, \widetilde{G}_k \, \mathrm{d}q \right)^2 \\
&= \frac{1}{\pi} \sum_{k \geqslant K} \left(\int_0^{2\pi} \widetilde{\varphi}(q)\, \cos(kq)\, \mathrm{d}q \right)^2 + \left(\int_0^{2\pi} \widetilde{\varphi}(q)\, \sin(kq)\, \mathrm{d}q \right)^2 \\
&\leqslant \frac{1}{\pi} \sum_{k \geqslant K} \left(\int_0^{2\pi} \widetilde{\varphi}(q)\, \frac{k^s}{K^s} \cos(kq)\, \mathrm{d}q \right)^2 + \left(\int_0^{2\pi} \widetilde{\varphi}(q)\, \frac{k^s}{K^s} \sin(kq)\, \mathrm{d}q \right)^2 \\
&= \frac{1}{\pi K^{2s}} \sum_{k \geqslant K} \left(\int_0^{2\pi} \widetilde{\varphi}(q)\, \partial_q^s \cos(kq)\, \mathrm{d}q \right)^2 + \left(\int_0^{2\pi} \widetilde{\varphi}(q)\, \partial_q^s \sin(kq)\, \mathrm{d}q \right)^2 \\
&\leqslant \frac{1}{K^{2s}} \|\partial_q^s \widetilde{\varphi}\|_{\mathrm{L}^2(\mathrm{d}q)}^2,
\end{aligned}$$

which allows to conclude with (2.57). □

Lemma 2.3. *For any $s \in \mathbb{N}$ and $\varphi \in \mathrm{H}^s(\kappa)$, it holds*

$$\forall L \geqslant s, \qquad \|\varphi - \Pi_L^p \varphi\|_{\mathrm{L}^2(\kappa)} \leqslant [\beta(L-s+1)]^{-s/2} \, \|\partial_p^s \varphi\|_{\mathrm{L}^2(\kappa)}.$$

Proof. Fix $L \geqslant s$. In view of (2.52), it holds

$$\left(\partial_p^*\right)^s H_{\ell-s} = \beta^{s/2} \sqrt{(\ell-s+1)\dots\ell}\, H_\ell,$$

with $\sqrt{(\ell-s+1)\dots\ell} \geqslant (L-s+1)^{s/2}$ when $\ell \geqslant L$. Therefore,

$$\begin{aligned}
\|\varphi - \Pi_L^p \varphi\|_{\mathrm{L}^2(\kappa)}^2 = \sum_{\ell \geqslant L} \langle \varphi, H_\ell \rangle^2 &\leqslant \sum_{\ell \geqslant L} \left\langle \varphi, \frac{\sqrt{(\ell-s+1)\dots\ell}}{(L-s+1)^{s/2}} H_\ell \right\rangle^2 \\
&= [\beta(L-s+1)]^{-s} \sum_{\ell \geqslant L} \left\langle \varphi, (\partial_p^*)^s H_{\ell-s} \right\rangle^2 \\
&\leqslant [\beta(L-s+1)]^{-s} \, \|\partial_p^s \varphi\|_{\mathrm{L}^2(\kappa)}^2,
\end{aligned}$$

from which the conclusion follows. □

The following approximation result is then directly deduced from the previous lemmas and (2.56).

Proposition 2.3. *Assume that V is smooth. Then, for any $s \in \mathbb{N}$, there exists $A_s \in \mathbb{R}_+$ such that*

$$\forall \varphi \in \mathrm{H}^s(\mu), \qquad \forall K \geqslant 1,\ L \geqslant s, \qquad \|\varphi - \Pi_{KL}\varphi\|_{\mathrm{L}^2(\mu)} \leqslant A_s \left(\frac{1}{K^s} + \frac{1}{L^{s/2}}\right) \|\varphi\|_{\mathrm{H}^s(\mu)}.$$

The approximation error $\|(1-\Pi_{KL})\Phi\|$ thus depends on the regularity of the solution Φ of the Poisson problem. Now, the operator $\mathcal{L}^{-1}$ is a bounded operator on $\mathrm{H}^s(\mu) \cap \mathrm{L}^2_0(\mu)$ for any $s \geqslant 0$ by the results of [153, Section 3.2] and [94] (see also [49, 80]). Therefore, when $R \in \mathrm{H}^s(\mu) \cap \mathrm{L}^2_0(\mu)$, the solution Φ belongs to $\mathrm{H}^s(\mu) \cap \mathrm{L}^2_0(\mu)$, and there is $\widetilde{A}_s \in \mathbb{R}_+$ such that

$$\|\Phi - \Pi_{KL}\Phi\|_{\mathrm{L}^2(\mu)} \leqslant \|\Phi - \Pi^q_K\Phi\|_{\mathrm{L}^2(\mu)} + \|\Phi - \Pi^p_L\Phi\|_{\mathrm{L}^2(\mu)} \leqslant \widetilde{A}_s \left(\frac{1}{K^s} + \frac{1}{L^{s/2}}\right) \|R\|_{\mathrm{H}^s(\mu)}. \quad (2.58)$$

Remark 2.6. *In fact, it can be expected that the operator $\mathcal{L}^{-1}$ further regularizes in the momentum variable; more precisely that $\partial_p\Phi \in \mathrm{H}^s(\mu)$ when $R \in \mathrm{H}^s(\mu)$. This is consistent with what we observe in the numerical simulations reported in Section 2.4.4. Note also that the estimates provided by [153, 49, 80, 94] are obtained for a fixed friction $\gamma > 0$. Some additional work is needed to carefully quantify their dependence upon γ, although we expect that the bounds on $\mathcal{L}^{-1}$ considered as an operator on $\mathrm{H}^s(\mu) \cap \mathrm{L}^2_0(\mu)$ should still scale as $\max(\gamma, \gamma^{-1})$.*

Let us conclude this section by an approximation result involving $\Pi_{KL,0}$ rather than Π_{KL} (see the decomposition (2.41), to be compared with (2.20)).

Corollary 2.4. *Assume that V is smooth. Then, for any $s \in \mathbb{N}$, there exists $A_s \in \mathbb{R}_+$ such that*

$$\forall \varphi \in \mathrm{H}^s(\mu) \cap \mathrm{L}^2_0(\mu), \qquad \forall K \geqslant 1,\ L \geqslant s, \qquad \|\varphi - \Pi_{KL,0}\varphi\|_{\mathrm{L}^2(\mu)} \leqslant A_s \left(\frac{1}{K^s} + \frac{1}{L^{s/2}}\right) \|\varphi\|_{\mathrm{H}^s(\mu)}.$$

Proof. Note first that $\langle H_\ell, \mathbf{1}\rangle = \delta_{\ell,0}$, so that $\Pi_{KL}\mathbf{1} = \Pi^q_K\mathbf{1}$ and $u_K = \Pi^q_K\mathbf{1}/\|\Pi^q_K\mathbf{1}\|$ depends only on the position variables for $L \geqslant 1$. Next, in view of the computations performed in the proof of Corollary 2.3,

$$\Pi_{KL,0}\varphi = \Pi_{KL}\varphi - \left\langle \frac{\Pi^q_K\mathbf{1}}{\|\Pi^q_K\mathbf{1}\|}, \varphi \right\rangle u_K,$$

where $\|u_K\| = 1$. Since $\varphi \in \mathrm{L}^2_0(\mu)$, it holds in fact

$$\left\langle \frac{\Pi^q_K\mathbf{1}}{\|\Pi^q_K\mathbf{1}\|}, \varphi \right\rangle = \left\langle \frac{(1-\Pi^q_K)\mathbf{1}}{\|\Pi^q_K\mathbf{1}\|}, \varphi \right\rangle,$$

which converges to 0 faster than any polynomial in K in view of Proposition 2.3. □

2.4.3 Consistency error

In order to simplify the computations (in particular to have some simple structure on the derivatives of the Fourier modes) we consider the following potential:

$$V(q) = 1 - \cos(q).$$

In this case, using the trigonometric identities

$$\begin{aligned} 2\cos(kq)\sin(q) &= \sin((k+1)q) - \sin((k-1)q) \\ 2\sin(kq)\sin(q) &= -\cos((k+1)q) + \cos((k-1)q), \end{aligned}$$

a straightforward computation shows that the derivatives of the basis functions satisfy

$$\begin{aligned} \partial_q G_0 &= \frac{\beta}{2\sqrt{2}} G_1, & \partial_q G_1 &= \frac{\beta}{2\sqrt{2}} G_0 + G_2 - \frac{\beta}{4} G_4, \\ \partial_q G_{2k} &= -\frac{\beta}{4} G_{2k-3} - k G_{2k-1} + \frac{\beta}{4} G_{2k+1}, & \partial_q G_{2k-1} &= \frac{\beta}{4} G_{2k-2} + k G_{2k} - \frac{\beta}{4} G_{2k+2}, \end{aligned} \tag{2.59}$$

where by convention $G_{-1} = 0$. The matrix $\mathbf{Q}$ defined in (2.54) is therefore a band matrix with width 4.

The well-posedness of the variational formulation associated with the Galerkin space is given by the following result.

Proposition 2.4. *The matrix $\widehat{\mathbf{L}}_{KL}$ defined in* (2.50) *is invertible for K, L sufficiently large. More precisely the resolvent bound satisfies*

$$\widehat{\lambda}_{\gamma,KL} \geqslant \lambda_\gamma - \frac{\varepsilon}{1+\varepsilon} \left[\frac{(1+\sqrt{2})\beta}{2K} + \frac{\beta^3}{16} \frac{\|(1-\Pi^q_{K-1})\mathbf{1}\|^2}{1 - \|(1-\Pi^q_K)\mathbf{1}\|^2} \right]. \tag{2.60}$$

In practice the term $\|(1-\Pi^q_{K-1})\mathbf{1}\|$ is very small (it decays faster than any polynomial in K by Lemma 2.2), so that the difference between the two estimates $\lambda_\gamma - \widehat{\lambda}_{\gamma,KL}$ scales as $1/K$ and in particular it does not depend on L. The proof presented in Appendix 2.6 consists in showing that the assumptions of Corollary 2.3 hold. Recall also that $\varepsilon, \lambda_\gamma \sim \min(\gamma, \gamma^{-1})$ by Proposition 2.1, so that the error term on the right-hand side of (2.60) is uniformly bounded with respect to λ_γ. This suggests that the relative error on the spectral gap is uniformly bounded with respect to $\gamma > 0$.

According to Theorem 2.3 and (2.79) the following rate of convergence can be deduced for the error $\Phi_M - \Pi^0_{KL}\Phi$ (which is related to the consistency error $\Pi^0_{KL}\mathcal{L}\Pi^0_{KL}\Phi + \Pi^0_{KL}R$).

Proposition 2.5. *The error $\|\Phi_{KL} - \Pi^0_{KL}\Phi\|$ is bounded by the approximation error as*

$$\|\Phi_{KL} - \Pi^0_{KL}\Phi\|_{\mathrm{L}^2(\mu)} \leqslant \frac{C}{\widehat{\lambda}_{\gamma,KL}} \left[\sqrt{\frac{L}{\beta}} (K - 1 + \beta) \|(1-\Pi_{KL})\Phi\|_{\mathrm{L}^2(\mu)} + \|\mathcal{L}u_K\| \|\Phi\|_{\mathrm{L}^2(\mu)} \right].$$

where $\|\mathcal{L}u_K\|$ decays faster than any polynomial (see (2.76) *for an explicit computation). Therefore, for any $s \geqslant 1$, there exists $A_{\gamma,s} \in \mathbb{R}_+$ such that, for all $R \in \mathrm{H}^s(\mu)$ and $\Phi =$*

$-\mathcal{L}^{-1}\Pi_0 R$,

$$\|\Phi_{KL} - \Pi^0_{KL}\Phi\|_{\mathrm{L}^2(\mu)} \leqslant A_{\gamma,s} K\sqrt{L}\left(\frac{1}{K^s} + \frac{1}{L^{s/2}}\right)\|R\|_{\mathrm{H}^s(\mu)}.$$

The second statement follows from the bounds on the approximation error $\|\Phi - \Pi^0_{KL}\Phi\|_{\mathrm{L}^2(\mu)}$ provided by Proposition 2.3, together with the fact that $\mathcal{L}^{-1}$ is a bounded operator on $\mathrm{H}^s(\mu) \cap \mathrm{L}^2_0(\mu)$ (see the discussion at the end of Section 2.4.2). The total error can thus be bounded as

$$\begin{aligned}\|\Phi_{KL} - \Phi\| &\leqslant \|\Phi_{KL} - \Pi^0_{KL}\Phi\| + \|\Phi - \Pi^0_{KL}\Phi\| \\ &\leqslant A_{\gamma,s} K\sqrt{L}\left(\frac{1}{K^s} + \frac{1}{L^{s/2}}\right)\|R\|_{\mathrm{H}^s(\mu)} + \widetilde{A}_s\left(\frac{1}{K^s} + \frac{1}{L^{s/2}}\right)\|R\|_{\mathrm{H}^s(\mu)} \\ &\leqslant \widehat{A}_{\gamma,s} K\sqrt{L}\left(\frac{1}{K^s} + \frac{1}{L^{s/2}}\right)\|R\|_{\mathrm{H}^s(\mu)}.\end{aligned} \tag{2.61}$$

2.4.4 Numerical results

In this section we call for simplicity consistency error the quantity $\|\Phi_{KL} - \Pi_{KL}\Phi\|$. In order to validate the results of Section 2.3 in the non-conformal case studied here, we compute the consistency error and the approximation error $\|\Phi - \Pi_{KL}\Phi\|$ as a function of the number K, L of modes and of the friction coefficient γ. We start by considering an observable which is not very regular; and then turn our attention to the case when $R(q,p) = p$ (which belongs to $\mathrm{H}^s(\mu)$ for any $s \in \mathbb{N}$). Solving the Poisson equation associated with this observable allows to predict the self-diffusion coefficient, which can be seen as the magnitude of the effective Brownian motion describing Langevin dynamics over diffusive timescales [127]. In all this section we set $\beta = 1$ and $m = 1$.

As a sanity check we also verified in the case $V = 0$ that the eigenvalues of the rigidity matrix $\mathbf{L}$ converge to their analytical expressions provided in [139].

Observable nearly in $\mathrm{H}^2(\mu)$. Fix $\gamma = 1$ and consider the observable

$$R = \sum_{k\in\mathbb{N},\ell\in\mathbb{N}} r_{k\ell} G_k H_\ell, \qquad r_{k\ell} = \max(1,k)^{-5/2}\max(1,\ell)^{-3/2}.$$

Note that

$$\|R\|^2 = \sum_{k\in\mathbb{N},\ell\in\mathbb{N}} |r_{k\ell}|^2 < +\infty.$$

Using (2.59) and (2.52) it can be shown that R is in $\mathrm{H}^1(\mu)$ but fails to be in $\mathrm{H}^2(\mu)$ (the exponents in $r_{k\ell}$ are critical). Note also that R does not have mean 0 with respect to μ, so that the solution of the saddle point problem (2.32) converges to the solution of the Poisson problem with $\Pi_0 R$ on the right-hand side. A very accurate approximation of the solution Φ, which serves as a reference value, is computed by setting $K = 100$ and $L = 1000$. The errors are plotted in Figure 2.1.

The polynomial power of the numerically observed decay of the approximation error is directly linked to the regularity of the solution Φ. Here the scalings K^{-3} and L^{-2} suggest

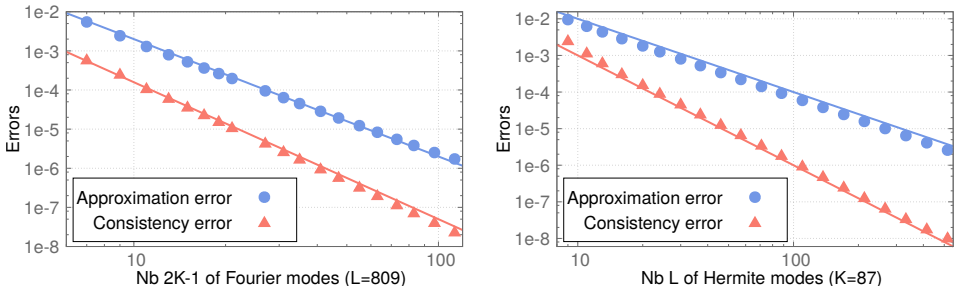


Figure 2.1: Approximation and consistency errors as a function of the number of modes. Left: varying number of Fourier modes for a large number of Hermite modes; the approximation error scales as K^{-3} while the consistency error scales as $K^{-7/2}$. Right: varying number of Hermite modes for a large number of Fourier modes; the approximation error scales as L^{-2} while the consistency error scales as L^{-3}.

that $\Phi, \partial_p \Phi \in \mathrm{H}^3(\mu)$, meaning that in this particular case $\mathcal{L}^{-1}$ regularizes one derivative of R in position and two in momenta, which is the most that could be expected. Note that the approximation error is therefore much smaller than predicted in (2.58), where we only stated that Φ is at least as regular as R. Moreover, we observe that the consistency error decays faster than the approximation error, as anticipated in Remark 2.3.

Velocity observable. The self-diffusion of a particle subjected to Langevin dynamics in dimension 1 is (see for instance [104, Section 5] for further background)

$$D = \int_0^\infty \mathbb{E}\left(p_t p_0\right) \mathrm{d}t = \left\langle -\mathcal{L}^{-1} p, p \right\rangle, \tag{2.62}$$

where the expectation is taken over all initial conditions $(q_0, p_0) \sim \mu$ and for all realizations of the Brownian motion in (2.1). This transport coefficient can be computed by approximating $\Phi = \mathcal{L}^{-1} p$ with the Galerkin method described in this article. The accurate reference is here computed by setting $K = 50$ and $L = 100$. We plot on Figure 2.2 the approximation error and the consistency error obtained for the observable $R(q, p) = p$. They decay faster than any polynomial since $p \in \mathrm{H}^s(\mu)$ for any $s \in \mathbb{N}$. They are in fact observed to decay exponentially fast with the number of modes. The error on the self-diffusion coefficient therefore also decays faster than any polynomial, in fact exponentially.

As an illustration of our approach, we plot the value of the self-diffusion as a function of γ in Figure 2.3, as already done in [127] using Monte-Carlo techniques and in [126] using a very similar spectral method. We indeed retrieve the scaling $D \sim \gamma^{-1}$ proved in [127]. This computation can be done in a matter of seconds as it involves a single inversion of a sparse matrix of size $KL = 5000$ for each value of the friction γ. It is thus much faster that a standard Monte-Carlo simulation. This approach however becomes intractable when the dimension increases.

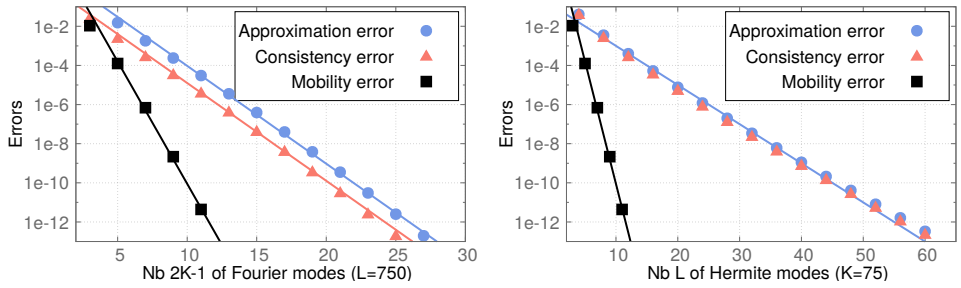


Figure 2.2: Approximation error, consistency error and error on the mobility as a function of the number of Fourier modes (Left) or Hermite modes (Right) for $\gamma = 1$. Logarithmic units are used on the ordinate axis. When the number of Hermite modes is large, the error on the mobility scales as $10^{-2.5K}$, while the approximation and consistency errors both scale as 10^{-K}. When the number of Fourier modes is large, the error on the mobility scales as $10^{-1.25L}$, while the approximation and consistency errors both scale as $10^{-0.2L}$.

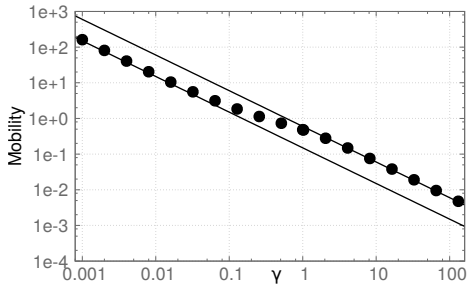


Figure 2.3: Self-diffusion as a function of the friction γ. It scales as γ^{-1} both for small γ (with prefactor 0.15) and large γ (with prefactor 0.6).

Estimates on the spectral gap. In order to illustrate the statements of Proposition 2.4, we compute the relative error between the spectral gap of $\mathcal{L}$ (approximated using a very large discretization basis) and the spectral gap of the matrix $\widehat{\mathbf{L}}$; see Figure 2.5. The spectral gap is close to the value $\min(\gamma, \gamma^{-1})$ obtained when $V = 0$ (see [95]), with deviations essentially around $\gamma = 1$. Note on Figure 2.4 that the relative error on the spectral gap decays exponentially with K and L. Let us also emphasize that, as suggested by (2.23), the relative error on the spectral gap is bounded uniformly with respect to γ for any K, L. We also observe that in the overdamped limit $\gamma \to \infty$ the relative error depends only on the discretization accuracy in the position variable. This is due to the fact that the resolvent $\mathcal{L}^{-1}$ converges in this regime to an operator acting only on the position variables [102].

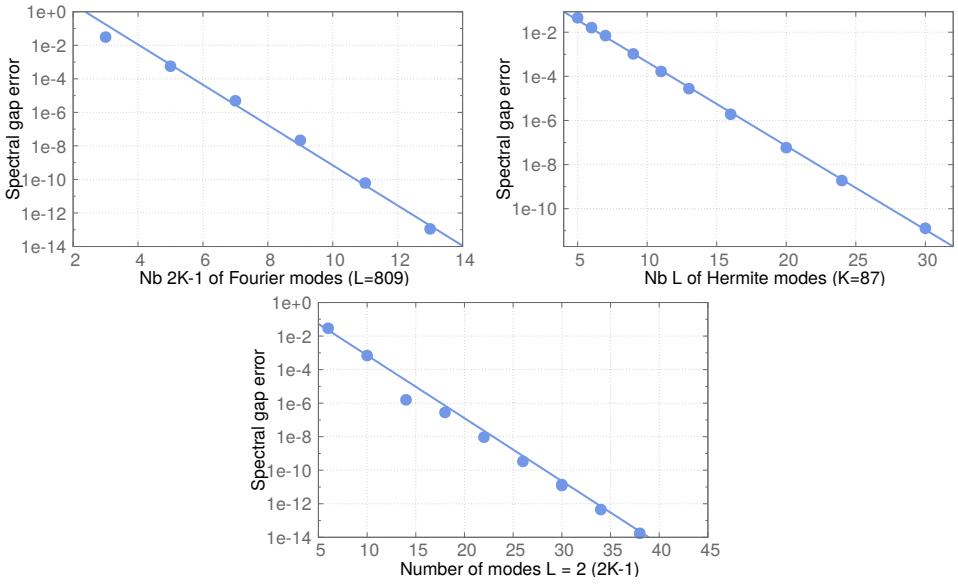


Figure 2.4: Error on the spectral gap as a function of the size of the basis in three cases for $\gamma = 1$. For a large number of Hermite modes the error scales approximatively as $10^{-1.2(2K-1)}$ (top left); for a large number of Fourier modes it scales approximatively as $10^{-0.32L}$ (top right); and for $L = 2(2K-1)$ it scales approximatively as $10^{-3.8L}$ (bottom).

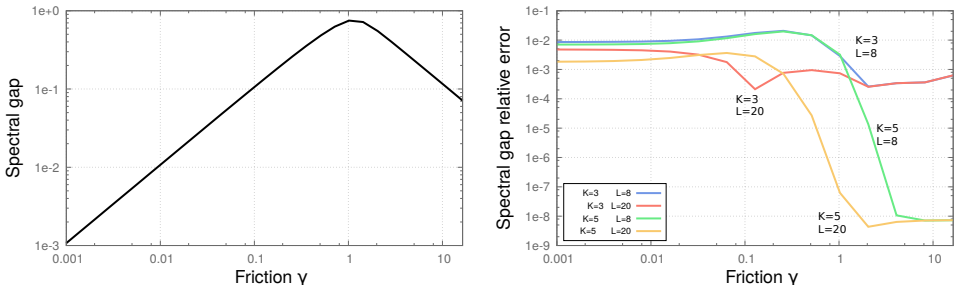


Figure 2.5: Left: Spectral gap as a function of the friction γ. Right: Relative error on the spectral gap as a function of γ for several couples K, L. Note that the curve corresponding to $K = 3, L = 8$ coincides with $K = 5, L = 8$ for γ small and with $K = 3, L = 20$ for γ large.

Acknowledgements

We thank Alexandre Ern, Tony Lelièvre and François Madiot (CERMICS), as well as Greg Pavliotis and Urbain Vaes (Imperial College) for helpful discussions. This work is supported by the Agence Nationale de la Recherche under grant ANR-14-CE23-0012 (COSMOS); as well as the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) – ERC Grant Agreement number 614492. We also benefited

from the scientific environment of the Laboratoire International Associé between the Centre National de la Recherche Scientifique and the University of Illinois at Urbana-Champaign. Finally, we acknowledge the support from the International Centre for Theoretical Sciences (ICTS) for the program *Non-equilibrium statistical physics* (ICTS/Prog-NESP/2015/10).

2.5 Proof of Theorem 2.1 ($\mathrm{L}^2(\mu)$ hypocoercivity)

We recall in this section the proof of Theorem 2.1, as presented in [43, 44]. We start with the proofs of the technical results presented at the end of Section 2.2.

Proof of Lemma 2.1. Consider $\varphi \in \mathsf{C}$. A simple computation shows that

$$\mathcal{L}_{\mathrm{ham}}\Pi_p = \frac{1}{\beta}\nabla_q \nabla_p^* \Pi_p = \left(\frac{p}{m}\right)^\top \nabla_q \Pi_p, \tag{2.63}$$

which immediately implies that $\mathcal{L}_{\mathrm{ham}}\Pi_p\varphi$ has average 0 with respect to $\kappa(\mathrm{d}p)$ for any $q \in \mathcal{D}$. Therefore, $\Pi_p \mathcal{L}_{\mathrm{ham}}\Pi_p = 0$, which implies $A = A(1-\Pi_p)$.

By definition of the operator A, it also holds

$$A\varphi + (\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)A\varphi = (\mathcal{L}_{\mathrm{ham}}\Pi_p)^*\varphi.$$

This identity immediately implies that $\Pi_p A = A$. Taking the scalar product with $A\varphi$, we obtain, using $\mathcal{L}_{\mathrm{ham}}A = \mathcal{L}_{\mathrm{ham}}\Pi_p A = (1-\Pi_p)\mathcal{L}_{\mathrm{ham}}A$:

$$\begin{aligned} \|A\varphi\|^2 + \|\mathcal{L}_{\mathrm{ham}}A\varphi\|^2 &= \langle \mathcal{L}_{\mathrm{ham}}A\varphi, \varphi\rangle = \langle \mathcal{L}_{\mathrm{ham}}A\varphi, (1-\Pi_p)\varphi\rangle \\ &\leqslant \|(1-\Pi_p)\varphi\|\,\|\mathcal{L}_{\mathrm{ham}}A\varphi\| \\ &\leqslant \frac{1}{4}\|(1-\Pi_p)\varphi\|^2 + \|\mathcal{L}_{\mathrm{ham}}A\varphi\|^2. \end{aligned} \tag{2.64}$$

The last inequality gives $\|A\varphi\| \leqslant \|(1-\Pi)\varphi\|/2$, while the second one implies that $\|\mathcal{L}_{\mathrm{ham}}A\varphi\| \leqslant \|(1-\Pi_p)\varphi\|$. The conclusion is finally obtained by density of C in $\mathrm{L}^2(\mu)$. □

The key element to prove Proposition 2.1 is the following coercivity estimates, respectively called "microscopic" and "macroscopic" coercivity in [43, 44].

Proposition 2.6 (Coercivity properties)**.** *The operators $\mathcal{L}_{\mathrm{FD}}$ and $\mathcal{L}_{\mathrm{ham}}\Pi_p$ satisfy the following coercivity properties:*

$$\forall \varphi \in \mathsf{C}, \qquad -\langle \mathcal{L}_{\mathrm{FD}}\varphi, \varphi\rangle \geqslant \frac{1}{m}\|(1-\Pi_p)\varphi\|^2, \tag{2.65}$$

$$\forall \varphi \in \mathsf{C} \cap \mathrm{L}^2_0(\mu), \qquad \|\mathcal{L}_{\mathrm{ham}}\Pi_p\varphi\|^2 \geqslant \frac{C_\nu}{\beta m}\|\Pi_p\varphi\|^2, \tag{2.66}$$

where C_ν is defined in (2.8)*. As a corollary, the following inequality holds in the sense of*

symmetric operators on $\mathrm{L}^2_0(\mu)$*:*

$$A\mathcal{L}_{\mathrm{ham}}\Pi_p \geqslant \lambda_{\mathrm{ham}}\Pi_p, \qquad \lambda_{\mathrm{ham}} = 1 - \left(1 + \frac{C_\nu}{\beta m}\right)^{-1} > 0. \tag{2.67}$$

Proof. The inequality (2.65) directly results from a Poincaré inequality for the Gaussian measure κ (see [14]), the position q being seen as a parameter. Indeed, for a given $\varphi \in \mathsf{C}$,

$$\forall q \in \mathcal{D}, \qquad \int_{\mathbb{R}^D} |\nabla_p \varphi(q,p)|^2 \, \kappa(\mathrm{d}p) \geqslant \frac{\beta}{m} \int_{\mathbb{R}^D} |(1-\Pi_p)\varphi(q,p)|^2 \, \kappa(\mathrm{d}p) \tag{2.68}$$

Integrating against ν and noting that $-\langle \mathcal{L}_{\mathrm{FD}}\varphi, \varphi\rangle = \beta^{-1}\|\nabla_p\varphi\|^2$ leads to the desired inequality.

To prove (2.66), we use (2.63), which leads to

$$\|\mathcal{L}_{\mathrm{ham}}\Pi_p\varphi\|^2_{\mathrm{L}^2(\mu)} = \|\frac{1}{m} p \nabla_q \Pi_p \varphi\|^2_{\mathrm{L}^2(\mu)} = \frac{1}{\beta m}\|\nabla_q \Pi_p \varphi\|^2_{\mathrm{L}^2(\nu)}. \tag{2.69}$$

The conclusion then follows from the Poincaré inequality (2.8), since, for $\varphi \in \mathsf{C} \cap \mathrm{L}^2_0(\mu)$, the function $\Pi_p\varphi$ has average 0 with respect to ν (namely, $\mathbb{E}_\nu[\Pi_p\varphi] = \mathbb{E}_\mu[\varphi] = 0$).

The macroscopic coercivity (2.66) allows to write $(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p) \geqslant \frac{C_\nu}{\beta m}\Pi_p$ in the sense of symmetric operators on $\mathrm{L}^2_0(\mu)$. Moreover,

$$A\mathcal{L}_{\mathrm{ham}}\Pi_p = \left[1 + (\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)\right]^{-1} (\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p).$$

Since $(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)$ is self-adjoint and the function $x \mapsto x/(1+x) = 1 - 1/(1+x)$ is increasing, the inequality (2.67) follows by spectral calculus. □

Another technical argument is the boundedness of certain operators, which appear in the proof of Proposition 2.1.

Lemma 2.4. *For any* $\ell \in \mathbb{N}^*$*,* $i \in \{1, 2, \ldots, D\}$ *and* $\varphi \in \mathrm{L}^2(\mu)$*,*

$$\|\Pi_p \partial_{p_i}^\ell \varphi\|_{\mathrm{L}^2(\nu)} \leqslant \sqrt{\left(\frac{\beta}{m}\right)^\ell \ell!}\, \|(1-\Pi_p)\varphi\|.$$

In particular, $\left\|\Pi_p \partial_{p_i}^\ell\right\| = \left\|\left(\partial_{p_i}^*\right)^\ell \Pi_p\right\| \leqslant \sqrt{\beta^\ell \ell!}$.

Proof. Fix $\varphi \in \mathsf{C}$. For $q \in \mathcal{D}$,

$$\left(\Pi_p \partial_{p_i}^n \varphi\right)(q) = \int_{\mathbb{R}^D} \left(\partial_{p_i}^n (1-\Pi_p)\varphi\right)(q,p)\, \kappa(\mathrm{d}p) = \int_{\mathbb{R}^D} (1-\Pi_p)\varphi(q,p)\, (\partial_{p_i}^*)^n \mathbf{1}\, \kappa(\mathrm{d}p).$$

Denoting by $H_\ell(p_i) = (m/\beta)^{\ell/2}\ell!^{-1/2}(\partial_{p_i}^*)^\ell \mathbf{1}$ the Hermite polynomials in the variable p_i

(which, we recall, are such that $\|H_\ell\|_{\mathrm{L}^2(\kappa)} = 1$), a Cauchy–Schwarz inequality shows that

$$\begin{aligned}\|\Pi_p \partial_{p_i}^\ell \varphi\|_{\mathrm{L}^2(\nu)}^2 &\leqslant \int_{\mathcal{D}} \left(\int_{\mathbb{R}^D} |(1-\Pi_p)\varphi(q,p)| \left| \sqrt{\left(\frac{\beta}{m}\right)^\ell \ell!} H_\ell(p_i) \right| \kappa(\mathrm{d}p) \right)^2 \nu(\mathrm{d}q) \\ &\leqslant \left(\frac{\beta}{m}\right)^\ell \ell! \int_{\mathcal{D}} \|(1-\Pi_p)\varphi(q,\cdot)\|_{\mathrm{L}^2(\kappa)}^2 \|H_\ell\|_{\mathrm{L}^2(\kappa)}^2 \nu(\mathrm{d}q) = \left(\frac{\beta}{m}\right)^\ell \ell! \|(1-\Pi_p)\varphi\|^2,\end{aligned}$$

which gives the claimed result. □

Proposition 2.7 (Boundedness of auxiliary operators)**.** *There exist $R_{\mathrm{ham}} > 0$ such that*

$$\forall \varphi \in \mathsf{C}, \qquad \begin{cases} \|A\mathcal{L}_{\mathrm{ham}}(1-\Pi_p)\varphi\| \leqslant R_{\mathrm{ham}} \|(1-\Pi_p)\varphi\|, \\ \|A\mathcal{L}_{\mathrm{FD}}\varphi\| \leqslant \dfrac{1}{2m} \|(1-\Pi_p)\varphi\|. \end{cases} \tag{2.70}$$

Proof. The first task is to give a more explicit expression of the operator A. In the following we use frequently the fact that operators acting only on the variables q (such as ∇_q and ∇_q^*) commute with operators acting only on variables p (such as ∇_p, ∇_p^* and Π_p). Moreover the relations $\partial_{p_i}\Pi_p = 0$, $\Pi_p \partial_{p_i}^* = 0$ and $\Pi_p \partial_{p_i}\partial_{p_j}^* = \partial_{p_i}\partial_{p_j}^* \Pi_p = \frac{\beta}{m}\Pi_p \delta_{ij}$ allow to simplify the action of $(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)$ as follows:

$$\begin{aligned}(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p) &= -\frac{1}{\beta^2}\Pi_p(\nabla_p^*\nabla_q - \nabla_q^*\nabla_p)(\nabla_p^*\nabla_q - \nabla_q^*\nabla_p)\Pi_p \\ &= \frac{1}{\beta^2}\Pi_p(\nabla_q^*\nabla_p)(\nabla_p^*\nabla_q)\Pi_p = \frac{1}{\beta m}\nabla_q^*\nabla_q\Pi_p.\end{aligned}$$

The operator A can therefore be reformulated as

$$A = \frac{1}{\beta}\left(1 + \frac{1}{\beta m}\nabla_q^*\nabla_q\right)^{-1} \nabla_q^*\Pi_p\nabla_p. \tag{2.71}$$

To obtain bounds on the operator $A\mathcal{L}_{\mathrm{ham}}(1-\Pi_p)$, we next consider its adjoint:

$$\begin{aligned}-(1-\Pi_p)\mathcal{L}_{\mathrm{ham}}A^* &= -\frac{1}{\beta^2}(1-\Pi_p)\left(\nabla_p^*\nabla_q - \nabla_q^*\nabla_p\right)\nabla_p^*\nabla_q\Pi_p\left(1 + \frac{1}{\beta m}\nabla_q^*\nabla_q\right)^{-1} \\ &= -\frac{1}{\beta^2}(1-\Pi_p)\left(\nabla_p^*\nabla_q\nabla_p^*\nabla_q - \frac{\beta}{m}\nabla_q^*\nabla_q\right)\Pi_p\left(1 + \frac{1}{\beta m}\nabla_q^*\nabla_q\right)^{-1} \\ &= -\frac{1}{\beta^2}(1-\Pi_p)\nabla_p^*\nabla_q\nabla_p^*\nabla_q\Pi_p\left(1 + \frac{1}{\beta m}\nabla_q^*\nabla_q\right)^{-1},\end{aligned}$$

where we used $(1-\Pi_p)\nabla_q^*\nabla_q\Pi_p = 0$ in the last line. Moreover, the operator

$$\nabla_p^*\nabla_q\nabla_p^*\nabla_q\Pi_p = \sum_{i,j=1}^{D} \partial_{p_i}^*\partial_{p_j}^*\Pi_p\partial_{q_i}\partial_{q_j}$$

is bounded from $\mathrm{H}^2(\nu)$ to $\mathrm{L}^2(\mu)$ according to Lemma 2.4. Moreover, as proved in [44], Assumption 2.1 ensures that the operator $\Pi_p\left(1+\frac{1}{\beta m}\nabla_q^*\nabla_q\right)^{-1}$ is bounded from $\mathrm{L}^2(\mu)$ to $\mathrm{H}^2(\nu)$. In conclusion, $-(1-\Pi_p)\mathcal{L}_{\mathrm{ham}}A^*$ is bounded on $\mathrm{L}^2(\mu)$.

The boundedness of the operator $A\mathcal{L}_{\mathrm{FD}}$ comes from the fact that

$$\begin{aligned}\Pi_p\mathcal{L}_{\mathrm{ham}}\mathcal{L}_{\mathrm{FD}} &= -\frac{1}{\beta^2}\Pi_p\left(\nabla_p^*\nabla_q - \nabla_q^*\nabla_p\right)\nabla_p^*\nabla_p = \frac{1}{\beta^2}\Pi_p\nabla_q^*\nabla_p\nabla_p^*\nabla_p \\ &= \frac{1}{\beta m}\Pi_p\nabla_q^*\nabla_p = -\frac{1}{m}\Pi_p\mathcal{L}_{\mathrm{ham}}.\end{aligned}$$

In conclusion, $A\mathcal{L}_{\mathrm{FD}} = -A/m$, which gives the claimed result with Lemma 2.1. □

We can now proceed with the proof of Proposition 2.1.

Proof of Proposition 2.1. Note first that, for a given $\varphi \in \mathsf{C}$, the entropy dissipation $\mathscr{D}[\varphi]$ can be explicitly written as

$$\begin{aligned}\mathscr{D}[\varphi] = \langle -\gamma\mathcal{L}_{\mathrm{FD}}\varphi, \varphi\rangle + \varepsilon\langle A\mathcal{L}_{\mathrm{ham}}\Pi_p\varphi, \varphi\rangle + \varepsilon\langle A\mathcal{L}_{\mathrm{ham}}(1-\Pi_p)\varphi, \varphi\rangle \\ - \varepsilon\langle \mathcal{L}_{\mathrm{ham}}A\varphi, \varphi\rangle + \varepsilon\gamma\langle A\mathcal{L}_{\mathrm{FD}}\varphi, \varphi\rangle,\end{aligned} \tag{2.72}$$

since $\mathcal{L}_{\mathrm{FD}}A = \mathcal{L}_{\mathrm{FD}}\Pi_p A = 0$. Using respectively the properties (2.65), (2.67), (2.70) and Lemma 2.1, it follows

$$\begin{aligned}\mathscr{D}[\varphi] \geqslant \frac{\gamma}{m}\|(1-\Pi_p)\varphi\|^2 + \varepsilon\lambda_{\mathrm{ham}}\|\Pi_p\varphi\|^2 - \varepsilon\left(R_{\mathrm{ham}} + \frac{\gamma}{2m}\right)\|(1-\Pi_p)\varphi\|\,\|\Pi_p\varphi\| \\ - \varepsilon\langle \mathcal{L}_{\mathrm{ham}}A\varphi, \varphi\rangle.\end{aligned} \tag{2.73}$$

Since, by Lemma 2.1,

$$\langle \mathcal{L}_{\mathrm{ham}}A\varphi, \varphi\rangle = \langle (1-\Pi_p)\mathcal{L}_{\mathrm{ham}}\Pi_p A(1-\Pi_p)\varphi, \varphi\rangle \leqslant \|(1-\Pi_p)\varphi\|^2,$$

it holds $\mathscr{D}[\varphi] \geqslant X^\top\mathbf{S}X$, where

$$X = \begin{pmatrix}\|\Pi_p\varphi\| \\ \|(1-\Pi_p)\varphi\|\end{pmatrix}, \qquad \mathbf{S} = \begin{pmatrix}S_{--} & S_{-+}/2 \\ S_{-+}/2 & S_{++}\end{pmatrix},$$

with

$$S_{--} = \varepsilon\lambda_{\mathrm{ham}}, \qquad S_{-+} = -\varepsilon\left(R_{\mathrm{ham}} + \frac{\gamma}{2m}\right), \qquad S_{++} = \frac{\gamma}{m} - \varepsilon.$$

The smallest eigenvalue of $\mathbf{S}$ is

$$\Lambda(\gamma,\varepsilon) = \frac{S_{--}+S_{++}}{2} - \frac{1}{2}\sqrt{(S_{--}-S_{++})^2+(S_{-+})^2}.$$

In the limit $\gamma \to 0$, the parameter ε should be chosen of order γ in order for $\Lambda(\gamma,\varepsilon)$ to be positive (in particular for S_{++} to remain positive). When $\gamma \to +\infty$, the parameter ε should be chosen of order $1/\gamma$ in order for the determinant of $\mathbf{S}$ to remain positive. We therefore consider the choice

$$\varepsilon = \overline{\varepsilon}\min(\gamma,\gamma^{-1}). \tag{2.74}$$

It is then easy to check that there exists $\overline{\varepsilon} > 0$ sufficiently small such that $\Lambda(\gamma,\overline{\varepsilon}\min(\gamma,\gamma^{-1})) > 0$ for all $\gamma > 0$. Moreover, it can be proved that $\Lambda(\gamma,\overline{\varepsilon}\min(\gamma,\gamma^{-1}))/\gamma$ converges to a positive value as $\gamma \to 0$, while $\gamma\Lambda(\gamma,\overline{\varepsilon}\min(\gamma,\gamma^{-1}))$ converges to a positive value as $\gamma \to +\infty$. This gives the claimed result with $\widetilde{\lambda}_\gamma = \Lambda(\gamma,\overline{\varepsilon}\min(\gamma,\gamma^{-1}))$. □

The proof of Theorem 2.1 is now easy to obtain. Consider $\varphi_0 \in \mathrm{Dom}(\mathcal{L}) \cap \mathrm{L}^2_0(\mu)$ (which contains $\mathrm{H}^2(\mu) \cap \mathrm{L}^2_0(\mu)$) and introduce $\mathscr{H}(t) = \mathcal{H}[\varphi(t)]$, where $\varphi(t) = \mathrm{e}^{t\mathcal{L}}\varphi_0 \in \mathrm{Dom}(\mathcal{L})$ for any $t \geqslant 0$. Then,

$$\mathscr{H}'(t) = -\mathscr{D}[\varphi(t)] \leqslant -\widetilde{\lambda}_\gamma \|\varphi(t)\|^2.$$

Using the norm equivalence (2.17) and the choice (2.74) for $\overline{\varepsilon} < 1$, it follows that

$$\mathscr{H}'(t) \leqslant -\frac{2\widetilde{\lambda}_\gamma}{1+\overline{\varepsilon}\min(\gamma,\gamma^{-1})}\mathscr{H}(t),$$

so that, by a Gronwall estimate,

$$\mathscr{H}(t) \leqslant \mathscr{H}(0)\exp\left(-\frac{2\widetilde{\lambda}_\gamma}{1+\overline{\varepsilon}\min(\gamma,\gamma^{-1})}t\right).$$

Using again the norm equivalence (2.17), it follows that

$$\|\varphi(t)\|^2 \leqslant \frac{1+\overline{\varepsilon}}{1-\overline{\varepsilon}}\,\mathrm{e}^{-2\lambda_\gamma t}\|\varphi(0)\|^2,$$

with the decay rate

$$\lambda_\gamma = \frac{\widetilde{\lambda}_\gamma}{1+\overline{\varepsilon}\min(\gamma,\gamma^{-1})}.$$

The desired estimate finally follows by density of $\mathrm{Dom}(\mathcal{L})$ in $\mathrm{L}^2(\mu)$.

2.6 Proof of technical estimates for the system considered in Section 2.4

We prove in this section that the conditions (2.29) and (2.21) allowing to apply the results of Section 2.3 hold for the system considered in Section 2.4. Recall that the condition

$M \to +\infty$ should be understood as $K, L \to +\infty$. Let us also emphasize that, although we perform the computations for the simple potential $V(q) = 1 - \cos(q)$, the extension to a general trigonometric polynomial V is straightforward.

Condition (2.29) **and bound on** $\|\mathcal{L}u_K\|$**.** Since u_M depends only on the positions, it is denoted u_K and

$$\|\mathcal{L}u_K\|^2 = \|\mathcal{L}^* u_K\|^2 = \frac{1}{\|\Pi_K^q \mathbf{1}\|^2} \|p\partial_q \Pi_K^q \mathbf{1}\|^2 = \beta \frac{\|\partial_q \Pi_K^q \mathbf{1}\|^2}{\|\Pi_K^q \mathbf{1}\|^2}.$$

In order to estimate $\|\partial_q \Pi_K^q \mathbf{1}\|$, we decompose $\Pi_K^q \mathbf{1}$ in the basis under consideration as follows:

$$\Pi_K^q \mathbf{1} = \sum_{j=0}^{2K-2} g_j G_j, \qquad g_j = \langle \Pi_K^q \mathbf{1}, G_j \rangle = \int_{\mathcal{D}} G_j \, d\nu.$$

Then, using $\partial_q \Pi_K^q \mathbf{1} = -\partial_q (1 - \Pi_K^q)\mathbf{1}$ and (with (2.59))

$$\forall k \geqslant 1, \qquad \partial_q^* G_{2k} = -\frac{\beta}{4} G_{2k-3} + k G_{2k-1} + \frac{\beta}{4} G_{2k+1}, \qquad \partial_q^* G_{2k-1} = \frac{\beta}{4} G_{2k-2} - k G_{2k} - \frac{\beta}{4} G_{2k+2}, \tag{2.75}$$

it follows that, for $K \geqslant 1$,

$$\begin{aligned}
\|\partial_q \Pi_K^q \mathbf{1}\|^2 &= \sum_{j \in \mathbb{N}} \langle \partial_q \Pi_K^q \mathbf{1}, G_j \rangle^2 \\
&= \sum_{j=0}^{2K-2} \langle -\partial_q (1 - \Pi_K^q)\mathbf{1}, G_j \rangle^2 + \sum_{j=2K-1}^{+\infty} \langle \partial_q \Pi_K^q \mathbf{1}, G_j \rangle^2 \\
&= \sum_{j=0}^{2K-2} \mathbb{E}_\nu \left[(1 - \Pi_K^q) \partial_q^* G_j \right]^2 + \sum_{j=2K-1}^{+\infty} \mathbb{E}_\nu \left[\Pi_K^q \partial_q^* G_j \right]^2 \\
&= \mathbb{E}_\nu [(1 - \Pi_K^q)\partial_q^* G_{2K-3}]^2 + \mathbb{E}_\nu [(1 - \Pi_K^q)\partial_q^* G_{2K-2}]^2 + \mathbb{E}_\nu [\Pi_K^q \partial_q^* G_{2K-1}]^2 + \mathbb{E}_\nu [\Pi_K^q \partial_q^* G_{2K}]^2 \\
&= \frac{\beta^2}{16} \left(g_{2K}^2 + g_{2K-1}^2 + g_{2K-2}^2 + g_{2K-3}^2 \right) \leqslant \frac{\beta^2}{16} \left\| (1 - \Pi_{K-1}^q)\mathbf{1} \right\|^2.
\end{aligned}$$

Since $\mathbf{1} \in \mathrm{H}^s(\nu)$ for any $s \in \mathbb{N}$, it follows that $\|(1 - \Pi_{K-1}^q)\mathbf{1}\|$ vanishes faster than any polynomial in K in view of Lemma 2.2. This implies that $\|\partial_q \Pi_K^q \mathbf{1}\|$, and hence $\|\mathcal{L}u_K\|$ and $\|\mathcal{L}^* u_K\|$, vanish faster than any polynomial in K. More precisely,

$$\|\mathcal{L}^* u_K\|^2 = \|\mathcal{L}u_K\|^2 \leqslant \frac{\beta^3}{16} \frac{\left\| (1 - \Pi_{K-1}^q)\mathbf{1} \right\|^2}{\|\Pi_K^q \mathbf{1}\|^2} \leqslant \frac{\beta^3}{16} \frac{\left\| (1 - \Pi_{K-1}^q)\mathbf{1} \right\|^2}{1 - \|(1 - \Pi_K^q)\mathbf{1}\|^2}. \tag{2.76}$$

Condition (2.21)**.** Let us now prove that $\|(A + A^*)(1 - \Pi_{KL})\mathcal{L}\Pi_{KL}\| \xrightarrow[K,L \to \infty]{} 0$ for the model under consideration. Introducing $\mathcal{L}_{KL}^{+-} = (1 - \Pi_{KL})\mathcal{L}\Pi_{KL}$, we prove in fact that $A\mathcal{L}_{KL}^{+-}$ and $A^*\mathcal{L}_{KL}^{+-}$ are bounded operators whose norms converge to 0 as $K, L \to +\infty$. In all this proof, we consider $K \geqslant 1$ and $L \geqslant 2$.

The first task is to provide a more explicit expression of $\mathcal{L}_{KL}^{+-}$. We introduce to this end the operator $D_K^{+-} = \Pi_K^{q\perp}\partial_q\Pi_K^q$. In view of (2.59),

$$D_K^{+-}\varphi = \sum_{j'=2K-1}^{+\infty}\sum_{j=0}^{2K-2}\langle\varphi, G_j\rangle\,\langle\partial_q G_j, G_{j'}\rangle\,G_{j'} = \frac{\beta}{4}\Big(\langle\varphi, G_{2K-2}\rangle\,G_{2K-1} - \langle\varphi, G_{2K-3}\rangle\,G_{2K}\Big).$$

This shows that the operator D_K^{+-} is bounded on $\mathrm{L}^2(\mu)$, and in fact

$$\|D_K^{+-}\varphi\| \leqslant \frac{\beta}{4}\|\Pi_{K+1}^{q\perp}\Pi_K^q\varphi\|. \tag{2.77}$$

Comparing (2.75) and (2.59), we also see that $D_K^{+-} = \Pi_K^{q\perp}\partial_q\Pi_K^q = \Pi_K^{q\perp}\partial_q^*\Pi_K^q$. We can now compute more explicitly the action of $\mathcal{L}_{KL}^{+-}$ by noting that

$$\beta\mathcal{L}_{KL}^{+-} = (1-\Pi_{KL})\partial_q\partial_p^*\Pi_{KL} - (1-\Pi_{KL})\partial_q^*\partial_p\Pi_{KL} - \gamma(1-\Pi_{KL})\partial_p^*\partial_p\Pi_{KL},$$

where $(1-\Pi_{KL})\partial_p^*\partial_p\Pi_{KL} = 0$ by (2.53), while (using (2.52) to write $\Pi_{L-1}^p\partial_p = \partial_p\Pi_L^p$ and $\Pi_{L+1}^p\partial_p^* = \partial_p^*\Pi_L^p$)

$$\begin{aligned}
(1-\Pi_{KL})\partial_q\partial_p^*\Pi_{KL} &= (1-\Pi_K^q\Pi_L^p)\partial_q\Pi_{L+1}^p\partial_p^*\Pi_K^q\\
&= (1-\Pi_K^q\Pi_L^p)\partial_q(\Pi_L^p + \Pi_L^{p\perp})\Pi_{L+1}^p\partial_p^*\Pi_K^q\\
&= (\Pi_L^p + \Pi_L^{p\perp}\Pi_{L+1}^p - \Pi_K^q\Pi_L^p)\partial_q\partial_p^*\Pi_K^q\\
&= \Pi_L^p(1-\Pi_K^q)\partial_q\partial_p^*\Pi_K^q + \partial_q\Pi_{L+1}^p\Pi_L^{p\perp}\partial_p^*\Pi_K^q\\
&= \Pi_L^p\partial_p^*D_K^{+-} + \partial_q\partial_p^*\Pi_L^p\Pi_{L-1}^{p\perp}\Pi_K^q\\
&= \partial_p^*\Pi_{L-1}^p D_K^{+-} + \partial_q\partial_p^*\Pi_{L-1}^{p\perp}\Pi_{KL},
\end{aligned}$$

and

$$\begin{aligned}
(1-\Pi_{KL})\partial_q^*\partial_p\Pi_{KL} &= \partial_p(1-\Pi_K^q\Pi_{L+1}^p)\partial_q^*\Pi_L^p\Pi_K^q = \partial_p\Pi_L^p(1-\Pi_K^q)\partial_q^*\Pi_K^q\\
&= D_K^{+-}\partial_p\Pi_L^p.
\end{aligned}$$

Therefore,

$$\beta\mathcal{L}_{KL}^{+-} = \partial_p^*\Pi_{L-1}^p D_K^{+-} + \partial_q\partial_p^*\Pi_{L-1}^{p\perp}\Pi_{KL} - D_K^{+-}\partial_p\Pi_L^p. \tag{2.78}$$

Moreover $\|\partial_p^*\Pi_{L-1}^p\| \leqslant \sqrt{\beta(L-1)}$, $\|\partial_p\Pi_L^p\| \leqslant \sqrt{\beta(L-1)}$ and using the Gerschgorin theorem (see [134] for example) $\|\partial_q\Pi_K^q\| \leqslant K-1+\beta/2$, so the operator $\mathcal{L}_{KL}^{+-}$ is bounded, with

$$\begin{aligned}
\left\|\mathcal{L}_{KL}^{+-}\right\| &\leqslant \beta^{-1}\sqrt{\beta(L-1)}\frac{\beta}{4} + \beta^{-1}\sqrt{\beta L}\left(K-1+\frac{\beta}{2}\right) + \beta^{-1}\sqrt{\beta(L-1)}\frac{\beta}{4}\\
&\leqslant \sqrt{\frac{L}{\beta}}\,(K-1+\beta)\,.
\end{aligned} \tag{2.79}$$

We are now in position to provide a more explicit expression of $A\mathcal{L}_{KL}^{+-}$ and $A^*\mathcal{L}_{KL}^{+-}$ based on (2.78). Recalling the definition (2.15) of $\Pi_p = \Pi_1^p$, it holds $\Pi_p\,\Pi_{L-1}^{p\perp} = 0$ and $\Pi_p\,\Pi_{L-1}^p = \Pi_p$

for $L \geqslant 2$. Using also the relation $\Pi_p \partial_p \partial_p^* = \beta$, we obtain

$$\begin{aligned}(\mathcal{L}_{\rm ham}\Pi_p)^* \mathcal{L}_{KL}^{+-} &= \beta^{-1}\Pi_p \partial_q^* \partial_p \mathcal{L}_{KL}^{+-} \\ &= \beta^{-1}\Pi_p \partial_q^* \Pi_{L-1}^p D_K^{+-} + \beta^{-1}\Pi_p \partial_q^* \partial_q \Pi_{L-1}^{p\perp} \Pi_{KL} - \beta^{-2}\Pi_p \partial_q^* \partial_p^2 D_K^{+-} \Pi_L^p \\ &= \beta^{-1}\Pi_p \partial_q^* D_K^{+-} - \beta^{-2}\Pi_p \partial_q^* \partial_p^2 D_K^{+-} \\ &= \beta^{-1}\Pi_p \left(1 - \beta^{-2}\partial_p^2\right) \partial_q^* D_K^{+-}\end{aligned}$$

since $L \geqslant 2$. Introducing the generator of the overdamped Langevin dynamics (for $m = 1$ here)

$$\mathcal{L}_{\rm ovd} = -\beta^{-1}\partial_q^* \partial_q,$$

it is possible to rewrite (2.71) as $A = (1 - \mathcal{L}_{\rm ovd})^{-1} \Pi_p \partial_p \partial_q^*$, so that

$$A\mathcal{L}_{KL}^{+-} = \left(\beta^{-1}\Pi_p - \beta^{-2}\Pi_p \partial_p^2\right)(1 - \mathcal{L}_{\rm ovd})^{-1} \partial_q^* D_K^{+-}. \tag{2.80}$$

Similar computations show that (using $\Pi_p \partial_p^* = 0$)

$$\begin{aligned}A^* \mathcal{L}_{KL}^{+-} &= -\beta^{-2}\partial_p^* \partial_q (1 - \mathcal{L}_{\rm ovd})^{-1} \Pi_p \partial_p D_K^{+-} \Pi_L^p \\ &= -\beta^{-2}\partial_p^* \Pi_p \partial_p \partial_q (1 - \mathcal{L}_{\rm ovd})^{-1} D_K^{+-}.\end{aligned} \tag{2.81}$$

The momentum operators Π_p, $\Pi_p \partial_p^2$ and $\partial_p^* \Pi_p \partial_p$ are bounded according to Lemma 2.4:

$$\left\|\beta^{-2}\Pi_p \partial_p^2 - \beta^{-1}\Pi_p\right\|_{\mathcal{B}(\mathrm{L}^2(\kappa))} \leqslant \frac{\sqrt{2}+1}{\beta}, \qquad \left\|\partial_p^* \Pi_p \partial_p\right\|_{\mathcal{B}(\mathrm{L}^2(\kappa))} \leqslant \beta,$$

so that

$$\begin{aligned}\left\|A\mathcal{L}_{KL}^{+-}\right\|_{\mathcal{B}(\mathrm{L}^2(\mu))} &\leqslant \frac{\sqrt{2}+1}{\beta} \left\|(1 - \mathcal{L}_{\rm ovd})^{-1} \partial_q^* D_K^{+-}\right\|_{\mathcal{B}(\mathrm{L}^2(\nu))}, \\ \left\|A^* \mathcal{L}_{KL}^{+-}\right\|_{\mathcal{B}(\mathrm{L}^2(\mu))} &\leqslant \frac{1}{\beta} \left\|\partial_q (1 - \mathcal{L}_{\rm ovd})^{-1} D_K^{+-}\right\|_{\mathcal{B}(\mathrm{L}^2(\nu))}.\end{aligned} \tag{2.82}$$

At this stage, it remains to prove that the operators on $\mathrm{L}^2(\nu)$ in the right-hand sides of the previous inequalities are bounded, with vanishing norms as $K \to +\infty$. We use to this end the following decompositions:

$$(1 - \mathcal{L}_{\rm ovd})^{-1} \partial_q^* D_K^{+-} = T_1 S_{1,K} D_K^{+-}, \qquad \partial_q (1 - \mathcal{L}_{\rm ovd})^{-1} D_K^{+-} = T_2 S_{2,K} D_K^{+-},$$

with (using $D_K^{+-} = \Pi_{K-1}^{q\perp} D_K^{+-}$)

$$\begin{aligned}T_1 &= (1 - \mathcal{L}_{\rm ovd})^{-1} \partial_q^* (1 - \widetilde{\mathcal{L}}_{\rm ovd})^{1/2}, & S_{1,K} &= (1 - \widetilde{\mathcal{L}}_{\rm ovd})^{-1/2} \Pi_{K-1}^{q\perp}, \\ T_2 &= \partial_q (1 - \mathcal{L}_{\rm ovd})^{-1/2}, & S_{2,K} &= (1 - \mathcal{L}_{\rm ovd})^{-1/2} \Pi_{K-1}^{q\perp},\end{aligned} \tag{2.83}$$

where we introduced the symmetric negative operator $\widetilde{\mathcal{L}}_{\rm ovd} = -\beta^{-1}\partial_q \partial_q^*$. Let us show that T_1 and T_2 are bounded and $S_{1,K}$ and $S_{2,K}$ can be made small for K sufficiently large. Note

first that

$$T_1T_1^* = (1-\mathcal{L}_{\rm ovd})^{-1}\partial_q^*\left(1-\widetilde{\mathcal{L}}_{\rm ovd}\right)\partial_q(1-\mathcal{L}_{\rm ovd})^{-1}$$
$$= (1-\mathcal{L}_{\rm ovd})^{-1}\left(\partial_q^*\partial_q+\beta^{-1}\partial_q^*\partial_q\partial_q^*\partial_q\right)(1-\mathcal{L}_{\rm ovd})^{-1} = -\beta(1-\mathcal{L}_{\rm ovd})^{-1}\mathcal{L}_{\rm ovd},$$

so that, by spectral calculus, $0 \leqslant T_1T_1^* \leqslant \beta$. This shows that T_1^* and T_1 are bounded operators on $\mathrm{L}^2(\nu)$, with $\|T_1^*\| = \|T_1\| \leqslant \sqrt{\beta}$. Similarly,

$$T_2^*T_2 = -\beta(1-\mathcal{L}_{\rm ovd})^{-1/2}\mathcal{L}_{\rm ovd}(1-\mathcal{L}_{\rm ovd})^{-1/2},$$

from which we deduce $\|T_2^*\| = \|T_2\| \leqslant \sqrt{\beta}$. We next prove that the operators $S_{1,K}$ and $S_{2,K}$ can be made as small as wanted by increasing K. We start by proving the following lemma.

Lemma 2.5. *For $K \geqslant 2$, the following inequalities hold in the sense of symmetric operators:*

$$1-\mathcal{L}_{\rm ovd} \geqslant \beta^{-1}(K-1)^2\Pi_{K-1}^{q\perp}, \qquad 1-\widetilde{\mathcal{L}}_{\rm ovd} \geqslant \beta^{-1}(K-1)^2\Pi_{K-1}^{q\perp}.$$

Proof. The operator $1-\mathcal{L}_{\rm ovd}$ can be expressed in the $\mathrm{L}^2(\mu)$-orthonormal basis G_k as

$$\begin{cases} (1-\mathcal{L}_{\rm ovd})G_{2k-1} = -\dfrac{\beta}{16}(G_{2k-5}+G_{2k+3}) - \dfrac{1}{4}(G_{2k-3}+G_{2k+1}) + \left(1+\dfrac{\beta}{8}+\dfrac{k^2}{\beta}\right)G_{2k-1}, \\ (1-\mathcal{L}_{\rm ovd})G_{2k} = -\dfrac{\beta}{16}(G_{2k-4}+G_{2k+4}) - \dfrac{1}{4}(G_{2k-2}+G_{2k+2}) + \left(1+\dfrac{\beta}{8}+\dfrac{k^2}{\beta}\right)G_{2k}. \end{cases} \tag{2.84}$$

Similar formulas hold for $1-\widetilde{\mathcal{L}}_{\rm ovd}$, upon changing the factors $-1/4$ into $1/4$ in the above expressions. Therefore, the symmetric operators $1-\mathcal{L}_{\rm ovd}-\left(\beta^{-1}(K-1)^2+\frac{1}{2}\right)\Pi_{K-1}^{q\perp}$ and $1-\widetilde{\mathcal{L}}_{\rm ovd}-\left(\beta^{-1}(K-1)^2+\frac{3}{2}\right)\Pi_{K-1}^{q\perp}$ can be represented by diagonally dominant matrices in the basis (G_k), which shows that these operators are positive. □

Lemma 2.6. *There exists $K_0 \in \mathbb{N}$ such that, for any $K \geqslant K_0$, the following inequalities hold in the sense of symmetric operators:*

$$0 \leqslant \Pi_{K-1}^{q\perp}(1-\mathcal{L}_{\rm ovd})^{-1}\Pi_{K-1}^{q\perp} \leqslant \frac{2\beta}{K^2}, \qquad 0 \leqslant \Pi_{K-1}^{q\perp}\left(1-\widetilde{\mathcal{L}}_{\rm ovd}\right)^{-1}\Pi_{K-1}^{q\perp} \leqslant \frac{2\beta}{K^2}.$$

Proof. We write the proof for the operator $\mathcal{A} = 1-\mathcal{L}_{\rm ovd}$, the result for $1-\widetilde{\mathcal{L}}_{\rm ovd}$ being obtained by similar manipulations. Consider the following block decomposition with respect to $\Pi_{K-1}^{q\perp}$ for K fixed:

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}^{--} & \mathcal{A}^{-+} \\ \mathcal{A}^{+-} & \mathcal{A}^{++} \end{pmatrix}.$$

More precisely, $\mathcal{A}^{--} = \Pi_{K-1}^{q}\mathcal{A}\Pi_{K-1}^{q}$, $\mathcal{A}^{-+} = \Pi_{K-1}^{q}\mathcal{A}\Pi_{K-1}^{q\perp}$, $\mathcal{A}^{+-} = \Pi_{K-1}^{q\perp}\mathcal{A}\Pi_{K-1}^{q}$ and $\mathcal{A}^{++} = \Pi_{K-1}^{q\perp}\mathcal{A}\Pi_{K-1}^{q\perp}$. A similar decomposition holds for $\mathcal{A}^{-1}$. With this notation, the goal is to

estimate $(\mathcal{A}^{-1})^{++} = \Pi_{K-1}^{q\perp}(1-\mathcal{L}_{\rm ovd})^{-1}\Pi_{K-1}^{q\perp}$. By the Schur complement formula,

$$\left(\mathcal{A}^{-1}\right)^{++} = \left[\mathcal{A}^{++} - \mathcal{A}^{+-}\left(\mathcal{A}^{--}\right)^{-1}\mathcal{A}^{-+}\right]^{-1},$$

provided the operators under consideration are all invertible. By Lemma 2.5,

$$\mathcal{A}^{++} - \mathcal{A}^{+-}\left(\mathcal{A}^{--}\right)^{-1}\mathcal{A}^{-+} \geqslant \left(\frac{(K-1)^2}{\beta} - \|\mathcal{A}^{+-}\|^2\left\|\left(\mathcal{A}^{--}\right)^{-1}\right\|\right)\Pi_{K-1}^{q\perp}.$$

Since $(\mathcal{A}^{--})^{-1} \leqslant 1$ (because $\mathcal{A}^{--} \geqslant 1$) and, in view of (2.84),

$$\|\mathcal{A}^{+-}\|^2 \leqslant \frac{1}{8} + \frac{\beta^2}{64},$$

the Schur complement is invertible for K sufficiently large, and its inverse is a symmetric operator satisfying

$$0 \leqslant \left(\mathcal{A}^{-1}\right)^{++} \leqslant \left[\frac{(K-1)^2}{\beta} - \left(\frac{1}{8} + \frac{\beta^2}{64}\right)\right]^{-1}\Pi_{K-1}^{q\perp}.$$

The right-hand side is, in turn, smaller than $2\beta/K^2$ for $K \geqslant K_0$ with K_0 sufficiently large. □

Since $S_{2,K}^* S_{2,K} = \Pi_{K-1}^{q\perp}(1-\mathcal{L}_{\rm ovd})^{-1}\Pi_{K-1}^{q\perp}$ and $S_{1,K}^* S_{1,K} = \Pi_{K-1}^{q\perp}(1-\widetilde{\mathcal{L}}_{\rm ovd})^{-1}\Pi_{K-1}^{q\perp}$, Lemma 2.6 immediately implies that

$$\forall K \geqslant K_0, \qquad \|S_{1,K}\|_{\mathrm{L}^2(\nu)} \leqslant \frac{\sqrt{2\beta}}{K}, \qquad \|S_{2,K}\|_{\mathrm{L}^2(\nu)} \leqslant \frac{\sqrt{2\beta}}{K}. \tag{2.85}$$

The conclusion now follows from (2.77) (which implies that $\left\|D_K^{+-}\right\|_{\mathrm{L}^2(\nu)} \leqslant \beta/4$) and (2.80)-(2.81), which lead to

$$\left\|T_1 S_{1,K} D_K^{+-}\right\|_{\mathrm{L}^2(\nu)} \leqslant \frac{\sqrt{2}\beta^2}{4K}, \qquad \left\|T_2 S_{2,K} D_K^{+-}\right\|_{\mathrm{L}^2(\nu)} \leqslant \frac{\sqrt{2}\beta^2}{4K}.$$

Using (2.82), we finally obtain

$$\left\|(A+A^*)\mathcal{L}_{KL}^{+-}\right\|_{\mathcal{B}(\mathrm{L}^2(\mu))} \leqslant \frac{(1+\sqrt{2})\beta}{2K}.$$

Final explicit estimates. Using the bounds provided in this appendix, it is easily seen that the constant $\widehat{\lambda}_{\gamma,KL}$ introduced in Corollary 2.3 satisfies (2.60). It is then possible to make explicit the resolvent bound (2.40).

Chapter 3

A perturbative approach to control variates in molecular dynamics

This chapter provides the content of [144] with some changes of notation and minor changes.

Contents

We propose a general variance reduction strategy to compute averages with diffusion processes. Our approach does not require the knowledge of the measure which is sampled, which may indeed be unknown as for nonequilibrium dynamics in statistical physics. We show by a perturbative argument that a control variate computed for a simplified version of the model can provide an efficient control variate for the actual problem at hand. We illustrate our method with numerical experiments and show how the control variate is built in three practical cases: the computation of the mobility of a particle in a periodic potential; the thermal flux in atom chains, relying on a harmonic approximation; and the mean length of a dimer in a solvent under shear, using a non-solvated dimer as the approximation.

3.1 Introduction

Diffusion processes have won an increasing interest in the past years in the statistical physics community, to model physical phenomena and to sample the underlying probability measure characterizing the state of the system [12]. The average value of a thermodynamic function R (as the energy, the pressure, a length, a flux, ...) under this probability distribution is given by an integral over the very high-dimensional configurational space. An important motivation for this work is the averaging of mean properties for systems subject to an external driving induced by non-reversible dynamics which change the invariant probability measure in a non trivial way. In this case the invariant probability measure is often not know. The goal can be to compute a transport coefficient, a free energy or more generally the response to a non-equilibrium forcing. From a practical point of view, the unknown probability measure is sampled by integrating a stochastic dynamics [2, 64, 161, 102]

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t. \tag{3.1}$$

Two prototypical dynamics in molecular simulation are the Langevin and overdamped Langevin dynamics [2, 101]. At equilibrium, the Langevin dynamics evolves positions q and momenta p as

$$\begin{cases} \mathrm{d}q_t = \dfrac{p_t}{m}\,\mathrm{d}t, \\ \mathrm{d}p_t = -\nabla V(q_t)\,\mathrm{d}t - \dfrac{\gamma}{m}p_t\,\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\,\mathrm{d}W_t, \end{cases} \tag{3.2}$$

where $\gamma > 0$ is the friction coefficient, $m > 0$ is the mass of a particle and $\beta > 0$ is proportional to the inverse temperature. The potential energy function is denoted by V and W_t is a multi-dimensional standard Brownian motion. In the limit of large frictions γ, this equation becomes after proper rescaling the overdamped Langevin dynamics [63]:

$$\mathrm{d}q_t = -\nabla V(q_t)\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}W_t. \tag{3.3}$$

Nonequilibrium versions of the above dynamics are obtained for instance by considering non-gradient forces rather than $-\nabla V$.

For any ergodic dynamics (3.1), macroscopic properties are computed via averages over

a trajectory as

$$\mathbb{E}[R] := \lim_{T\to\infty} \widehat{\varphi}_T \quad \text{a.s.}, \qquad \widehat{\varphi}_T = \frac{1}{T}\int_0^T R(X_t)\,\mathrm{d}t.$$

The statistical error for these estimators is characterized by the asymptotic variance:

$$\sigma_R^2 = \lim_{T\to\infty} T\,\mathrm{Var}\left[\widehat{\varphi}_T\right]. \tag{3.4}$$

In many cases of interest ergodic means converge very slowly, requiring the use of variance reduction techniques to speed up the computation. Two types of phenomena can lead to large statistical errors: first, the metastability arising from multimodal potentials, which can greatly increase the correlation of the trajectory in time and lead to large variances; second, a high signal-to-noise ratio, which is typical when averaging small linear responses as for the computation of transport coefficients [54, 161].

When the system is at equilibrium, by which we mean that detailed balance holds, the invariant probability measure is often known and it is possible to use standard variance reduction techniques [147, 58, 28, 108, 99] such as importance sampling [36, 108] or stratification [66, 110, 91, 155, 148]. This allows to address both metastability issues and high noise-to-signal ratios.

For non-equilibrium systems, and more generally when the invariant probability measure is not known, reducing the variance is challenging since standard variance reduction methods cannot be used. Note that reducing the metastability would require to modify the dynamics while keeping the invariant probability measure unchanged, or at least knowing how it changes (see for instance [104, Section 3.4]). When the latter is not known, this task is hard to perform. On the other hand reducing the noise-to-signal ratio is feasible even for non-equilibrium dynamics. This is the goal of the present work, where we rely on control variates. Control variates laying on the concept of "zero-variance" principle have been already used in molecular simulation [7]. This approach was also studied in Bayesian inference simulations [78, 118, 40, 119, 122], where configurations are sampled with Markov chains rather than diffusion processes. This type of techniques has however been restricted to cases where the invariant probability measure is known, except for specific settings such as [68], where a coupling strategy is described.

In the present work, which relies on ideas announced in [104, Section 3.4.2], control variates are constructed without any knowledge of the expression of the invariant probability measure. We build an unbiased modified observable $R+\xi = R+\mathcal{L}\Phi$, where $\mathcal{L}$ is the generator of the dynamics, which is of smaller variance (at least in some asymptotic regime):

$$\mathbb{E}[R+\xi] = \mathbb{E}[R] \quad \text{and} \quad \sigma_{R+\xi}^2 < \sigma_R^2.$$

The optimal choice for the control variate ξ is $\xi = \mathcal{L}\Phi$ where Φ is the solution of the following Poisson equation:

$$-\mathcal{L}\Phi = R - \mathbb{E}[R].$$

The general strategy we consider consists in approximating this partial differential equation (PDE) by a simplified one, with an operator $\mathcal{L}_0$, for which the solution Φ_0 can be analytically

computed or numerically approximated with a good precision. Theorems 3.1 and 3.2 provide an analysis of the asymptotic variance $\sigma^2_{R+\xi}$.

We present numerical results illustrating the general method in three practical cases. In particular we provide in each case a simplified process, associated to a simplified Poisson problem. In these applications we are interested in averaging the linear response of an observable with respect to a non-equilibrium perturbation. This is a challenging class of problems since the average quantity is small and thus the relative statistical error is large. We present the problems we consider by increasing complexity of the setup. We start with the computation of the mobility of a particle in a periodic two-dimensional potential. The control variate can be approached with a very high precision by a numerical method based on a spectral basis, allowing to illustrate Theorem 3.2. We next estimate the conductivity of an atom chain [20, 106, 41]. The number of state variables is much larger (up to several hundreds of degrees of freedom in our simulations) but the geometrical setting is one-dimensional. The control variate can be computed analytically when taking a harmonic model as a reference, and thus it does not require any additional numerical procedure. The third application is a dimer in a solvent, whose mean length is estimated under an external shearing force. In the latter case the difficulty comes from the fact that the system is high-dimensional and not as structured as the atom chain.

This article is organized as follows. We present in Section 3.2 the general strategy for building control variates and state a result making precise how control variates behave in a perturbative framework. We then turn to the case of a single particle in a one-dimensional periodic potential under a non-gradient forcing in Section 3.3; the computation of the thermal flux passing through a chain in Section 3.4; and the estimation of the mean length of a dimer in a solvent under an external shearing stress in Section 3.5. Some technical results are gathered in the appendices.

3.2 General strategy

The definition of the asymptotic variance of time averages along a trajectory requires to introduce more precisely the generator of the process and some associated functional spaces, which is done in Section 3.2.1. The concept of control variate is then explained Section 3.2.2, as well as the so-called "zero-variance principle". We give in Section 3.2.3 our perturbative construction of control variate in an abstract setting and state the main theorem quantifying the variance reduction in a limit regime. Finally Section 3.2.4 provides a generalized version of this theorem in the case when an approximate solver is used.

3.2.1 Asymptotic variance

The state space $\mathcal{X}$ is typically the full space $\mathbb{R}^d$ or a bounded domain with periodic boundary conditions $\mathbb{T}^d$. For some dynamics such as Langevin dynamics, auxiliary variables with values in $\mathbb{R}^d$ are added, so that in this case $\mathcal{X} = \mathbb{R}^d \times \mathbb{R}^d$ or $\mathcal{X} = \mathbb{T}^d \times \mathbb{R}^d$. As suggested in the introduction, we decompose the generator of the process (3.1) as a sum $\mathcal{L} = \mathcal{L}_0 + \widetilde{\mathcal{L}}$ of a reference generator $\mathcal{L}_0$ and a perturbation $\widetilde{\mathcal{L}}$. In order to study the asymptotic regime

corresponding to small perturbations, we use a parameter $\eta \in \mathbb{R}$ to interpolate smoothly between $\mathcal{L}_0$ and $\mathcal{L}$, and define

$$\mathcal{L}_\eta = \mathcal{L}_0 + \eta\widetilde{\mathcal{L}}.$$

We suppose that these operators $\mathcal{L}_\eta$ write:

$$\mathcal{L}_\eta = b_\eta \cdot \nabla + \frac{1}{2}\sigma_\eta\sigma_\eta^\top : \nabla^2 = \sum_{i=1}^{d} (b_\eta)_i \, \partial_{x_i} + \frac{1}{2}\sum_{i,j=1}^{d} \left(\sigma_\eta\sigma_\eta^\top\right)_{i,j} \partial_{x_i x_j},$$

with b_η and σ_η are continuous. They are then the generators of the following stochastic processes on $\mathcal{X}$, indexed by η:

$$\mathrm{d}X_t^\eta = b_\eta(X_t^\eta)\,\mathrm{d}t + \sigma_\eta(X_t^\eta)\,\mathrm{d}W_t, \tag{3.5}$$

where W_t is a d-dimensional standard Brownian motion. Let us assume that b_η and σ_η are such that the following holds.

Assumption 3.1. *The dynamics* (3.5) *admits a unique invariant probability measure* π_η *for any* $\eta \in \mathbb{R}$. *Moreover, trajectorial ergodicity holds: for any observable* $R \in \mathrm{L}^1(\pi_\eta)$,

$$\mathbb{E}_\eta[R] := \int_{\mathcal{X}} R(x)\,\mathrm{d}\pi_\eta(x) = \lim_{T\to\infty} \frac{1}{T}\int_0^T R(X_t^\eta)\,\mathrm{d}t \quad \text{a.s.}$$

Sufficient conditions for this to hold are discussed after Assumption 3.2. Let us now make the functional spaces precise. We denote by $(\mathcal{K}_n)_{n\in\mathbb{N}}$ a family of so-called Lyapunov functions with values in $[1, +\infty)$. The associated weighted L^∞ spaces are:

$$\forall n \in \mathbb{N}, \quad \mathrm{L}^\infty_n = \left\{\varphi \text{ measurable } \middle|\ \|\varphi\|_{\mathrm{L}^\infty_n} < \infty\right\}, \qquad \|\varphi\|_{\mathrm{L}^\infty_n} = \left\|\frac{\varphi}{\mathcal{K}_n}\right\|_{\mathrm{L}^\infty}.$$

We make the following assumption on the Lyapunov functions.

Assumption 3.2. *For any* $\eta \in \mathbb{R}$, *the function* $\mathcal{K}_n$ *belongs to* $\mathrm{L}^2(\pi_\eta)$. *In particular,*

$$\forall n \in \mathbb{N}, \forall \eta \in \mathbb{R}, \quad \mathrm{L}^\infty_n \subset \mathrm{L}^2(\pi_\eta).$$

We also assume that for any $n, n' \in \mathbb{N}$, *there exists* $m \in \mathbb{N}$ *such that* $\mathcal{K}_n\mathcal{K}_{n'} \in \mathrm{L}^\infty_m$.

The first part of Assumption 3.2 is typically obtained by using a family of Lyapunov functions satisfying conditions of the form

$$\mathcal{L}_\eta\mathcal{K}_n \leqslant -\alpha_{n,\eta}\mathcal{K}_n + b_{n,\eta}, \tag{3.6}$$

for some $\alpha_{n,\eta} > 0$ and $b_{n,\eta} \in \mathbb{R}$. Indeed, after integration against π_η,

$$0 = \int_{\mathcal{X}} \mathcal{L}_\eta\mathcal{K}_n \mathrm{d}\pi_\eta \leqslant -\alpha_{n,\eta}\int_{\mathcal{X}} \mathcal{K}_n \mathrm{d}\pi_\eta + b_{n,\eta},$$

so that

$$1 \leqslant \int_{\mathcal{X}} \mathcal{K}_n \mathrm{d}\pi_\eta \leqslant \frac{b_{n,\eta}}{\alpha_{n,\eta}}.$$

We can then conclude with the second part of Assumption 3.2 since, for any $n \geqslant 1$, there exist $C_n > 0$ and $m \geqslant 1$ such that $1 \leqslant \mathcal{K}_n^2 \leqslant C_n \mathcal{K}_m$. The condition (3.6) also implies the existence of an invariant probability measure for any $\eta \in \mathbb{R}$ when a minorization condition holds [74]. A typical choice for the Lyapunov functions are the polynomials $\mathcal{K}_n(x) = 1 + |x|^n$. This choice satisfies the second part of Assumption 3.2 when the invariant probability measure has moments of any order. Unless otherwise mentioned, we always consider this choice in the sequel (which is standard for Langevin and overdamped Langevin dynamics, see [153, 109, 93, 94]). As for Assumption 3.1, trajectorial ergodicity holds when the generator $\mathcal{L}_\eta$ is elliptic or hypoelliptic, and there exists an invariant probability measure with positive density with respect to the Lebesgue measure [89]. The latter condition follows if the measure which appears in the minorization condition has a positive density with respect to the Lebesgue measure.

We denote for any function $\varphi \in \mathrm{L}^1(\pi_\eta)$ the projection on the space of mean zero functions by:

$$\Pi_\eta \varphi = \varphi - \mathbb{E}_\eta[\varphi].$$

For any operator $A \in \mathcal{B}(E)$ (bounded on the Banach space E), the operator norm is defined as

$$\|A\|_{\mathcal{B}(E)} = \sup_{\|\varphi\|_E = 1} \|A\varphi\|_E.$$

Let us now make the following assumption.

Assumption 3.3. *For any $n \in \mathbb{N}$, the $\mathrm{L}^2(\pi_\eta)$ norms of the Lyapunov functions are uniformly bounded on compact sets of η: for any $\eta_* > 0$ there exists a constant C_{n,η_*} such that*

$$\forall |\eta| \leqslant \eta_*, \quad \|\mathcal{K}_n\|_{\mathrm{L}^2(\pi_\eta)} \leqslant C_{n,\eta_*}. \tag{3.7}$$

Moreover $\mathcal{L}_\eta$ is invertible on $\Pi_\eta \mathrm{L}^\infty_n$. Finally the inverse generator is bounded uniformly on compact sets of η:

$$\forall |\eta| \leqslant \eta_*, \quad \left\| -\mathcal{L}_\eta^{-1} \right\|_{\mathcal{B}(\Pi_\eta \mathrm{L}^\infty_n)} \leqslant C_{n,\eta_*}. \tag{3.8}$$

The invertibility of $\mathcal{L}_\eta$ on $\Pi_\eta \mathrm{L}^\infty_n$ is a standard result which follows typically from the Lyapunov conditions (3.6) and a so-called minorization condition [74]. It has been proved for a large variety of problems [50, 153, 109, 93, 94]. Conditions (3.7) and (3.8) are needed to prove Theorems 3.1 and 3.2 to come. Condition 3.7 can be obtained by showing uniform bounds on the coefficients which appear in the Lyapunov conditions, while condition (3.8) additionally requires some uniformity on the minorization condition.

When Assumptions 1 to 3 hold, the asymptotic variance introduced in (3.4) is finite for any $\varphi \in \mathrm{L}^\infty_n$ and the following formula holds [104]:

$$\sigma^2_{\varphi,\eta} = 2 \left\langle \varphi, -\mathcal{L}_\eta^{-1} \Pi_\eta \varphi \right\rangle_\eta, \tag{3.9}$$

where $\langle \cdot, \cdot \rangle_\eta$ denotes the canonical scalar product on $\mathrm{L}^2(\pi_\eta)$. We refer to Appendix 3.9 for

more details on the numerical estimation of the asymptotic variance.

Remark 3.1. *We choose to work directly with weighted* L^∞ *spaces as this is the relevant setting for Theorems 3.1 and 3.2. Note that the asymptotic variance of an observable* $\varphi \in \mathrm{L}^2(\pi_\eta)$ *can also be defined using perturbative arguments relatively to an equilibrium reference dynamics [44, 85]. Contrarily to the* L^∞_n *framework one would however be restricted in this case to small non-equilibrium perturbations.*

3.2.2 Ideal control variate

We recall in this section what a control variate is in our context and show how the construction of an optimal control variate can be reformulated as solving a Poisson problem. The functional framework is made precise in a second step using Assumption 3.3. We say that a function ξ is a control variate of the observable R for the process X_t^η with generator $\mathcal{L}_\eta$ if

$$\mathbb{E}_\eta[R+\xi] = \mathbb{E}_\eta[R] \quad \text{and} \quad \sigma^2_{R+\xi,\eta} < \sigma^2_{R,\eta}.$$

The principle of our method, already explained in [104], is based on the equation which characterizes the invariance of the measure π_η: for any function Φ,

$$\mathbb{E}_\eta[\mathcal{L}_\eta \Phi] = 0.$$

This shows that control variates ξ of the form $\xi = \mathcal{L}_\eta\Phi$ automatically ensure that $\mathbb{E}_\eta[R+\xi] = \mathbb{E}_\eta[R]$, whatever the choice of Φ. In order for ξ to be a good control variate, the modified observable $R + \xi = R + \mathcal{L}_\eta\Phi$ should however be of small asymptotic variance. The optimal choice, denoted by Φ_η and referred to as the "zero-variance principle" [7, 118], is to make the modified observable constant. This constant is then necessarily equal to $\mathbb{E}_\eta[R]$, and Φ_η is the solution of the Poisson problem:

$$-\mathcal{L}_\eta\Phi_\eta = R - \mathbb{E}_\eta[R]. \tag{3.10}$$

Assuming that $R \in \mathrm{L}^\infty_n$ for some $n \in \mathbb{N}$, the problem (3.10) admits a unique solution $\Phi_\eta \in \mathrm{L}^\infty_n$ when Assumption 3.3 holds.

In practice two problems arise when trying to solve (3.10). First, the equation (3.10) is a very high-dimensional PDE for most purposes and the complexity of such problems scales exponentially with the dimension, contrarily to stochastic sampling. Second, $\mathbb{E}_\eta[R]$ is not known since it is precisely the quantity we are trying to compute. We discuss in the next section how to approximate the solution of (3.10), at least for small η.

3.2.3 Perturbative control variate

The key assumption in our approach is to assume that we can compute the solution Φ_0 of the reference Poisson problem corresponding to $\eta = 0$:

$$-\mathcal{L}_0\Phi_0 = R - \mathbb{E}_0[R]. \tag{3.11}$$

In practice $\mathcal{L}_0$ is the generator of a simplified dynamics (depending on the problem), and $\mathcal{L}_\eta$ is the generator of the problem at hand (say for $\eta = 1$). Let us emphasize that the dynamics associted with $\mathcal{L}_0$ need not be an equilibrium dynamics (see the example discussed in Section 3.4). The small parameter η is used to quantify the discrepancy between the optimal function Φ_η and its approximation Φ_0 in a perturbative framework. We refer to Section 3.2.4 for a discussion on the numerical resolution of (3.11).

We define the so-called core space $\mathcal{S}$ as the set of all $\mathcal{C}^\infty$ functions which grow at infinity at most like K_n for some n, and whose derivatives also grow at most like K_n for some n. Such a space was considered in [153] for instance. More precisely,

$$\mathcal{S} = \Big\{ \varphi \in \mathcal{C}^\infty(\mathcal{X}) \Big| \ \forall k \in \mathbb{N}, \exists n \in \mathbb{N}, \quad \partial_k \varphi \in \mathrm{L}^\infty_n \Big\}. \tag{3.12}$$

The space $\mathcal{S}$ is dense in $\mathrm{L}^2(\pi_\eta)$ under Assumption 3.2, since the $\mathcal{C}^\infty$ functions with compact support are included in $\mathcal{S}$. We need an additional assumption in our analysis to ensure that $\Phi_0 \in \mathcal{S}$.

Assumption 3.4. *The space $\mathcal{S}$ is stable by the generator $\mathcal{L}_0$ and $\mathcal{L}_0$ is invertible on the space $\Pi_0\mathcal{S}$ composed of functions with average 0 with respect to the invariant probability measure π_0. This means that, for any $\varphi \in \Pi_0\mathcal{S}$, there exists a unique solution $\psi \in \Pi_0\mathcal{S}$ to the Poisson equation*

$$-\mathcal{L}_0 \psi = \varphi.$$

Assumption 3.4 can be proved to hold for Langevin and overdamped Langevin dynamics at equilibrium under certain assumptions on the potential V, see [153, 93, 94]. The generator of the perturbation should satisfy the following.

Assumption 3.5. *The generator $\widetilde{\mathcal{L}}$ of the perturbation is such that $\mathcal{S}$ is stable by $\widetilde{\mathcal{L}}$ and $\widetilde{\mathcal{L}}^* \mathbf{1} \in \mathrm{L}^2(\pi_0)$.*

Here and in the following we denote by B^* the adjoint of a closed operator B on the functional space $\mathrm{L}^2(\pi_0)$. Assumption 3.5 is easy to check, since $\widetilde{\mathcal{L}}$ is typically a differential operator with coefficients in $\mathcal{S}$.

Let us define the following modified observable involving $\Phi_0 \in \Pi_0\mathcal{S}$, defined in (3.11):

$$\phi_\eta := R + \mathcal{L}_\eta \Phi_0.$$

The following theorem makes precise the main properties of this modified observable.

Theorem 3.1. *Fix $R \in \mathcal{S}$. Under Assumptions 1 to 5, $\phi_\eta \in \mathcal{S}$ is well defined for any $\eta \in \mathbb{R}$, and $\mathbb{E}_\eta[\phi_\eta] = \mathbb{E}_\eta[R]$. Moreover, for any $\eta_* > 0$, there exists $C_{R,\eta_*} > 0$ such that, for any $|\eta| \leqslant \eta_*$, the asymptotic variance satisfies*

$$\sigma^2_{\phi_\eta,\eta} = 2\eta^2 \left\langle AR, -\mathcal{L}_0^{-1} AR \right\rangle_0 + \eta^3 E_{R,\eta}, \tag{3.13}$$

with $A = -\widetilde{\mathcal{L}}\mathcal{L}_0^{-1}\Pi_0$ and $|E_{R,\eta}| \leqslant C_{R,\eta_*}$.

The scalar products involved in the previous theorem are well defined since $\mathcal{S}$ is stable by A, and $\mathcal{S} \subset \mathrm{L}^2(\pi_\eta)$ for any $\eta \in \mathbb{R}$. Equation (3.13) shows that the standard error $\sqrt{\frac{\sigma^2_{\phi_\eta,\eta}}{T}}$ committed on the empirical estimator $\widehat{\varphi}_T$ of $\mathbb{E}_\eta[R]$ after a time T is of leading order η.

The scaling η^2 of the asymptotic variance formally comes from the fact that the modified observable writes $\phi_\eta = \mathbb{E}_\eta[R] + \mathrm{O}(\eta)$. Indeed,

$$\mathcal{L}_\eta \mathcal{L}_0^{-1} \Pi_0 = \Pi_\eta (\mathcal{L}_0 + \eta \widetilde{\mathcal{L}}) \mathcal{L}_0^{-1} \Pi_0 = \Pi_\eta + \eta \Pi_\eta \widetilde{\mathcal{L}} \mathcal{L}_0^{-1} \Pi_0 = \Pi_\eta (1 - \eta A), \tag{3.14}$$

so that the modified observable can be rewritten as:

$$\phi_\eta = R + \mathcal{L}_\eta \Phi_0 = R - \mathcal{L}_\eta \mathcal{L}_0^{-1} \Pi_0 R = \mathbb{E}_\eta[R] + \eta \Pi_\eta A R. \tag{3.15}$$

In particular,

$$\Pi_\eta \phi_\eta = \eta \Pi_\eta A R. \tag{3.16}$$

The remainder of the proof consists in carefully estimating remainders in some truncated series expansion of $-\mathcal{L}_\eta^{-1} \Pi_\eta$; see Appendix 3.6.

Remark 3.2. *The formula* (3.13) *can in fact be replaced by an expansion in powers of* η *with a truncation at an arbitrarily high order and a remainder controlled uniformly in* $|\eta| \leqslant \eta_*$. *This can be proved by an immediate generalization of the proof we provide in Appendix 3.6.*

In the following applicative sections we cannot always prove that Assumptions 3.3 and 3.4 hold true, but the scaling of the variance predicted by Theorem 3.1 is nevertheless numerically observed to hold.

3.2.4 Numerical resolution of the reference Poisson problem

We discuss in this section a strategy to compute the solution to (3.11) when this equation cannot be analytically solved. We rely for this on a Galerkin strategy, and look for an approximation of the solution Φ_0 to the Poisson problem (3.11) in a subspace $V_M \subset \mathrm{L}^2(\pi_0)$ of finite dimension M. For simplicity we suppose that $V_M \subset \Pi_0 \mathrm{L}^2(\pi_0)$ (which corresponds to a conformal approximation). This implies in particular that $\Phi_M \in \Pi_0 \mathrm{L}^2(\pi_0)$ has mean zero with respect to π_0. We also assume that $V_M \subset \mathrm{H}^2(\pi_0)$ to avoid regularity issues. The optimal choice for the approximation $\Phi_{0,M}$ is to minimize the variance of the modified observable $R + \mathcal{L}_0 \Phi_{0,M}$ for the reference dynamics:

$$\min_{\varphi \in V_M} \sigma^2_{R+\mathcal{L}_0\varphi,0} = \sigma^2_{R,0} + \min_{\varphi \in V_M} \frac{1}{2} \left\langle (\mathcal{L}_0 + \mathcal{L}_0^*)\varphi, 2\mathcal{L}_0^{-1}\Pi_0 R + \varphi \right\rangle_0 . \tag{3.17}$$

The latter equality follows from Lemma 3.2 in Appendix 3.6, in the particular case when $\eta = 0$. In the case when $\mathcal{L}_0$ is not the generator of a stochastic differential equation the quantity $\sigma^2_{R+\mathcal{L}_0\varphi,0}$ cannot be interpreted as a variance but the analysis we provide here remains valid. The necessary optimality condition for a minimizer $\Phi_{0,M}$ of (3.17) is given

by the following Euler-Lagrange equation:

$$\forall \psi \in V_M, \quad \left\langle (\mathcal{L}_0 + \mathcal{L}_0^*)(\Phi_{0,M} + \mathcal{L}_0^{-1}\Pi_0 R), \psi \right\rangle_0 = 0.$$

Introducing the orthogonal projector Π_M on V_M (with respect to the scalar product on $\mathrm{L}^2(\pi_0)$), the latter equation can be rewritten as

$$\Pi_M(\mathcal{L}_0 + \mathcal{L}_0^*)(\Phi_{0,M} + \mathcal{L}_0^{-1}\Pi_0 R) = 0. \tag{3.18}$$

In practice we distinguish two cases.

(i) For reversible dynamics such as the overdamped Langevin dynamics (3.3) $\mathcal{L}_0 = \mathcal{L}_0^*$, so the equation reduces to

$$-\Pi_M \mathcal{L}_0 \Phi_{0,M} = \Pi_M R. \tag{3.19}$$

(ii) For Langevin dynamics at equilibrium (see (3.2)), we consider a tensorized basis involving Hermite elements as in Section 3.3 or in [145]. The symmetric part of the generator:

$$\frac{1}{2}(\mathcal{L}_0 + \mathcal{L}_0^*) = -\gamma\beta^{-1}\nabla_p^*\nabla_p,$$

diagonalizes the Hermite polynomials so we have the commutation rule $\Pi_M(\mathcal{L}_0+\mathcal{L}_0^*) = (\mathcal{L}_0+\mathcal{L}_0^*)\Pi_M$. Moreover the kernel of $\mathcal{L}_0+\mathcal{L}_0^*$ is composed of functions depending only on the position variables. The condition (3.18) then implies that there exists $g = g(q)$ in $\mathrm{L}^2(\pi_0)$ such that

$$\Phi_{0,M} = -\Pi_M \mathcal{L}_0^{-1}\Pi_0 R + g. \tag{3.20}$$

The solution to (3.19) coincides with the result provided by the Galerkin method on the approximation space V_M. For (3.20) the optimal solution, for $g = 0$, is given by the Galerkin method apart from a consistency error (see [145] for a detailed analysis). This justifies the use of the Galerkin methods to determine a good approximation $\Phi_{0,M}$ of Φ_0 in the general case.

For the Langevin equation the operator $\mathcal{L}$ is not coercive on $\mathrm{L}^2_0(\pi)$ so the associated rigidity matrix is not automatically invertible. The existence of a unique solution $\Phi_{0,M} \in V_M$ converging to $\Phi_0 = -\mathcal{L}^{-1}\Pi_0 R$ when $M \to +\infty$, as well as error estimates and a discussion on non-conformal approximations, can be found in [145].

When a Galerkin method (or any other approximation method) is used, the error committed on Φ_0 induces an error on the modified observable ϕ_η. The modified asymptotic variance is then the sum of terms coming from (3.13) (depending on η) and terms coming from the approximation error (of order ε) committed on Φ_0, as made precise in the following result.

Theorem 3.2. *Fix $R \in \mathcal{S}$ and assume that Φ_0 is approximated by $\Phi_{0,\varepsilon} = \Phi_0 + \varepsilon f$ with $f \in \mathcal{S}$ and $\varepsilon \geqslant 0$. Denote by $\phi_{\eta,\varepsilon} = R + \mathcal{L}_\eta \Phi_{0,\varepsilon}$ the modified observable. Under Assumptions 1*

to 5, for any $\eta_*, \varepsilon_* > 0$, there exists $E_{R,\eta_*,\varepsilon_*} > 0$ such that, for any $|\eta| \leqslant \eta_*$ and $|\varepsilon| \leqslant \varepsilon_*$,

$$\begin{aligned}\sigma^2_{\phi_{\eta,\varepsilon},\eta} = 2\varepsilon^2 \left\langle -\mathcal{L}_0 f, f \right\rangle_0 - 2\varepsilon\eta \left\langle (\mathcal{L}_0 + \mathcal{L}_0^*) f, \mathcal{L}_0^{-1}\Pi_0 AR \right\rangle_0 \\ + 2\eta^2 \left\langle AR, -\mathcal{L}_0^{-1}\Pi_0 AR \right\rangle_0 + (\eta^3 + \varepsilon^3) C_{R,\eta,\varepsilon},\end{aligned} \tag{3.21}$$

with $|C_{R,\eta,\varepsilon}| \leqslant E_{R,\eta_*,\varepsilon_*}$.

The proof of this result can be read in Appendix 3.6. It shows that the variance is globally of order 2 with respect to both η and ε. This suggests to take η and ε of the same order.

Remark 3.3. *The dependence of $\widetilde{C}_{R,\eta_*,\varepsilon_*}$ with respect to R can be made more explicit (see for instance the discussion in [102]).*

The error committed on Φ_0 can also arise from additional approximations on the right hand side of the Poisson problem, in situations when the observable $R_\eta = R_0 + \eta\widetilde{R}$ depends on η. A result similar to Theorem 3.2 can be obtained upon assuming that $\widetilde{R} \in \mathcal{S}$, where Φ_0 is the solution to (3.11) with R replaced by R_0.

Remark 3.4. *Note that in the expression* (3.21) *the error term f only appears through $(\mathcal{L}_0 + \mathcal{L}_0^*)f$. This term may vanish even if f is not identically zero. For example, for a Langevin process at equilibrium, $(\mathcal{L}_0 + \mathcal{L}_0^*)f$ vanishes when f is a function depending only on the positions.*

3.3 One-dimensional Langevin dynamics

We construct in this section a control variate for a one-dimensional system by solving a simplified Poisson equation using a spectral Galerkin method. The simplification consists in neglecting a non-equilibrium perturbation, the small parameter η being the amplitude of this perturbation. We first present in Section 3.3.1 the model and define the quantity of interest, namely the mobility. We next construct in Section 3.3.2 the approximate control variate and conclude in Section 3.3.3 with some numerical results.

3.3.1 Full dynamics

We consider the following Langevin process on the state space $\mathcal{X} = 2\pi\mathbb{T} \times \mathbb{R}$:

$$\begin{cases} \mathrm{d}q_t = \dfrac{p_t}{m}\,\mathrm{d}t, \\ \mathrm{d}p_t = (-v'(q_t) + \eta)\,\mathrm{d}t - \dfrac{\gamma}{m}p_t\,\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\,\mathrm{d}W_t, \end{cases} \tag{3.22}$$

where $\gamma, m, \beta > 0$ and v is a smooth 2π-periodic potential. The particle experiences a constant external driving of amplitude $\eta \in \mathbb{R}$. This force is not the gradient of a periodic function, so the system is out of equilibrium and the invariant measure is not known. We

are interested in the average velocity $R(q,p) = \frac{p}{m}$ induced by the non-gradient force η, which can also be seen as a mass flux. The linear response of the average velocity with respect to the external driving is characterized by the mobility of the particle [143]:

$$D = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta[R]}{\eta}.$$

The generator of (3.22) is the sum of the generator associated with the Langevin dynamics at equilibrium and of a non-equilibrium perturbation:

$$\mathcal{L}_\eta = \mathcal{L}_0 + \eta \widetilde{\mathcal{L}},$$

where

$$\mathcal{L}_0 = -v'(q)\partial_p + \frac{p}{m}\partial_q - \frac{\gamma}{m} p \partial_p + \gamma\beta^{-1}\partial_p^2, \qquad \widetilde{\mathcal{L}} = \partial_p.$$

In this setting the Lyapunov functions are defined for all $n \in \mathbb{N}$ as:

$$\forall (q,p) \in \mathcal{X}, \quad \mathcal{K}_n(q,p) = 1 + |p|^n,$$

and Assumptions 3.1, 3.2, 3.3 and 3.4 correspond to standard results for Langevin dynamics [153, 137, 94, 102, 104]. Assumption 3.5 trivially holds: the core space $\mathcal{S}$ is stable by $\widetilde{\mathcal{L}} = \partial_p$ by definition, while $\widetilde{\mathcal{L}}^* \mathbf{1} = (-\partial_p + \frac{\beta}{m} p)\mathbf{1} = \frac{\beta}{m} p \in \mathrm{L}^2(\pi_0)$.

3.3.2 Simplified dynamics and control variate

Solving the Poisson problem $-\mathcal{L}_\eta \Phi_\eta = R - \Pi_\eta[R]$ associated with the nonequilibrium dynamics is not practical because $\mathbb{E}_\eta[R]$ is not know. For this simple one-dimensional example it would still be technically doable. Since our purpose is however to illustrate both Theorems 3.1 and 3.2, we do not follow this path. We therefore consider the control variate associated to a reference Poisson problem, namely

$$-\mathcal{L}_0 \Phi_0 = R. \tag{3.23}$$

Note that the average drift vanishes at equilibrium: $\mathbb{E}_0[R] = 0$. Equation (3.23) cannot be solved analytically, but it is possible to approach its solution by a Galerkin method as explained in Section 3.2.4. The modified observable is

$$\phi_{\eta,M} = R + \mathcal{L}_\eta \Phi_{0,M},$$

with the notation of Theorem 3.2 (the error committed when estimating Φ_0 is indexed by M instead of ε). In practice we construct a basis $(e_m)_m$ of V_M and write $\Phi_{0,M} = \sum_{m=1}^M a_m e_m$ where the coefficients $(a_m)_m \in \mathbb{R}^M$ are the solution of a linear system obtained from (3.18) (see [145]). The modified observable is then $\phi_{\eta,M} = R + \sum_{m=1}^M a_m \mathcal{L}_\eta e_m$ where the functions $\mathcal{L}_\eta e_m$ are explicit for appropriate choice of basis functions $(e_m)_m$; see Section 3.3.3.

The mobility is estimated with an ergodic average of R along a trajectory during a time T, for a forcing η, by $\widehat{D}_{\eta,T} = \frac{1}{\eta}\widehat{R}_T$. This estimator has an expectation of order 1 and a

large variance when η is small and T is large [103, 104]:

$$\mathbb{E}[\widehat{D}_{\eta,T}] = D + \mathrm{O}\left(\eta + \frac{1}{T}\right), \qquad \mathrm{Var}[\widehat{D}_{\eta,T}] \sim \frac{\sigma_R^2}{\eta^2 T},$$

so that the relative statistical error scales as

$$\frac{\sqrt{\mathrm{Var}[\widehat{D}_{\eta,T}]}}{\mathbb{E}[\widehat{D}_{\eta,T}]} \sim \frac{\sigma_R}{D\eta\sqrt{T}}.$$

In order for this quantity to be small, the simulation time should be taken of order $T \sim \frac{1}{\eta^2}$, which is very large since η is small.

When the Poisson problem is exactly solved, the modified observable is

$$\phi_\eta = R + \mathcal{L}_\eta \Phi_0 = \eta \widetilde{\mathcal{L}} \Phi_0,$$

which is proportional to η, so that the associated asymptotic variance scales as $\sigma^2_{\phi_\eta} \sim \eta^2$. The relative statistical error for the mobility estimator $\widetilde{D}_{\eta,T} = \frac{1}{\eta}\widehat{\phi}_{\eta,T}$ is then bounded with respect to η:

$$\frac{\mathrm{Var}[\widetilde{D}_{\eta,T}]}{\mathbb{E}[\widetilde{D}_{\eta,T}]} \sim \frac{1}{D\sqrt{T}},$$

and the simulation time can be fixed independently of the value of η. Now if an error of order ε_M is committed on Φ_0 the asymptotic variance of $\phi_{\eta,M}$ scales like $\eta^2 + \varepsilon_M^2$ so the relative statistical error on $\widetilde{D}_{\eta,T}$ is of order

$$\frac{|\eta| + \varepsilon_M}{\sqrt{T}|\eta|} = \frac{1}{\sqrt{T}}\left(1 + \frac{\varepsilon_M}{|\eta|}\right).$$

This implies that the simulation time T should be taken of order $1 + \left(\frac{\varepsilon_M}{\eta}\right)^2$ instead of η^{-2}.

3.3.3 Numerical results

In order to simplify the numerical resolution of the Galerkin problem (see Section 3.2.4) we consider the simple potential:

$$\forall q \in 2\pi\mathbb{T}, \quad v(q) = 1 - \cos(q).$$

We construct V_M using a tensorized basis made of weighted Fourier modes in position and Hermite modes in momenta. The particular weights of the Fourier modes are chosen so that the basis is orthogonal for the $\mathrm{L}^2(\pi_0)$ scalar product. Obtaining error estimates on $\Phi_0 - \Phi_{0,M}$ requires some work, see [145, Section 4] for a detailed analysis and a precise expression of the basis. In the following we take either $M = 15 \times 10$ basis elements (15 Fourier modes and 10 Hermite modes), $M = 7 \times 5$ or 5×3 basis elements. Estimating Φ_0

allows to construct a control variate, and also to compute directly the mobility since the Green–Kubo formula [97] states that

$$D = \beta \left\langle R, -\mathcal{L}_0^{-1} R \right\rangle_0 = \beta \left\langle R, \Phi_0 \right\rangle_0 . \tag{3.24}$$

In order to compute the mobility using Monte-Carlo simulations, we fix $\eta > 0$ small and rely on the estimators $\widehat{D}_{\eta,T}$ or $\widetilde{D}_{\eta,T}$ defined in Section 3.3.2. We are interested in the reduction of the asymptotic variance provided by our control variate, *i.e.* comparing $\mathrm{Var}[\widetilde{D}_{\eta,T}]$ and $\mathrm{Var}[\widehat{D}_{\eta,T}]$. The Langevin dynamics is integrated over a time $T = 2 \times 10^4$ with time steps $\Delta t = 0.02$ for an external forcing η ranging from 0.01 to 2.56. The numerical integration is done with a Geometric Langevin Algorithm [23]. This scheme ensures that the invariant probability measure is correct up to terms of order $\mathrm{O}(\Delta t^2)$ at equilibrium. Moreover the transport coefficients estimated by linear response are also correct up to terms of order Δt^2 (see [102]). The scheme writes:

$$\begin{cases} p^{k+1/2} = p^k - v'(q^k)\dfrac{\Delta t}{2}, \\ q^{k+1} = q^k + \dfrac{p^{k+1/2}}{m}\Delta t, \\ \widetilde{p}^{k+1} = p^{k+1/2} - v'(q^{k+1})\dfrac{\Delta t}{2}, \\ p^{k+1} = \alpha_{\Delta t}\widetilde{p}^{k+1} + \sqrt{m\beta^{-1}(1-\alpha_{\Delta t}^2)}\; G^k, \end{cases} \tag{3.25}$$

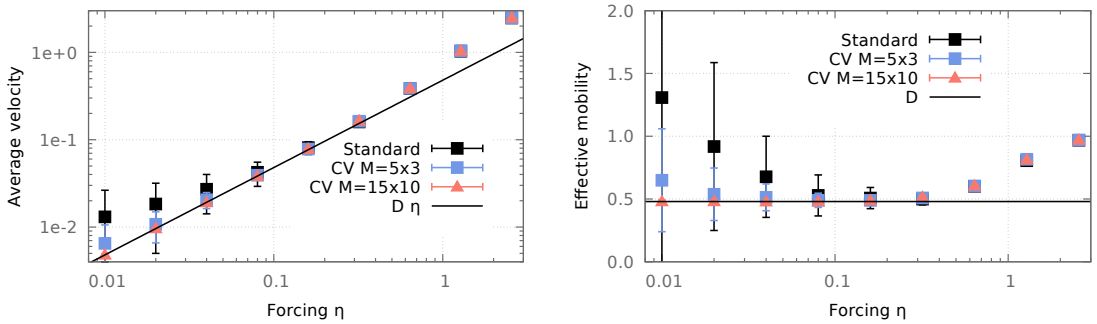
where the superscript k is the iteration index, $\alpha_{\Delta t} = \exp\Big(-\frac{\gamma}{m}\Delta t\Big)$ and the $(G^k)_{k\in\mathbb{N}}$ are independent and identically distributed (i.i.d.) standard one-dimensional Gaussian random variables. The results which are reported are obtained for $m = \gamma = \beta = 1$.

Linear response. The results presented in Figure 3.1 (Left) show that the average velocity scales linearly with respect to the forcing for η small, as predicted by linear response theory. The slope, which is the mobility D, matches the one computed using (3.24) for M large. An effective mobility is obtained by dividing the average velocity by the forcing, see Figure 3.1 (Right). We are interested in its limiting value for a small forcing. On the one hand the result is biased if η is too large, but on the other hand the statistical error scales like $\frac{1}{\eta}$. For the standard observable we see that the optimal trade-off value of η is around 0.2 for the chosen simulation time T. When using a control variate the variance is much smaller in the small forcing regime, so η can be taken very small to reduce the bias while keeping the statistical error under control. We discuss next the estimation of the error bars plotted on Figure 3.1.

Correlation profiles. The asymptotic variance of an observable $\varphi \in \mathcal{S}$ writes, using the Green–Kubo formula [97],

$$\sigma_\varphi^2 = 2\int_0^\infty C_\varphi(t)\,\mathrm{d}t, \qquad C_\varphi(t) = \mathbb{E}[(\varphi(X_0) - \mathbb{E}[\varphi])(\varphi(X_t) - \mathbb{E}[\varphi])],$$

Figure 3.1: Linear response for the standard MC simulation (black squares) compared to the version with control variate (blue, red) and to the asymptotic response $D\eta \approx 0.48\eta$ (black line).



where the autocorrelation function C_φ involves an expectation over all initial condition $X_0 = (q_0, p_0) \in \mathcal{X}$ distributed according to the invariant probability measure and all realizations of the Brownian motion. The integrability of $C_\varphi(t)$ can be guaranteed when the semi-group $\mathrm{e}^{t\mathcal{L}}$ decays sufficiently fast in L^∞_n [104]. The function C_φ is characterized by three major features explaining the value of the asymptotic variance; see Figure 3.2 for an illustration.

(i) The first one is the amplitude of the signal $\|\varphi - \mathbb{E}[\varphi]\|^2_{\mathrm{L}^2(\pi)} = C_\varphi(0)$ corresponding to the value of the autocorrelation at $t = 0$.

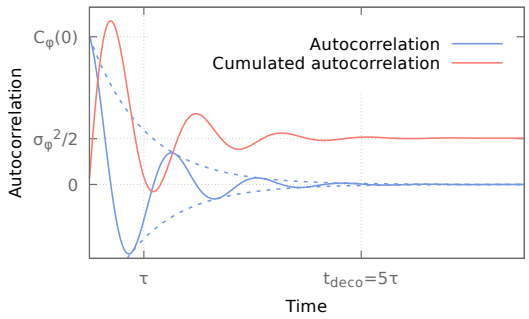
(ii) The second one is the characteristic decay time τ of the autocorrelation, which can be related to the decay of its exponential envelope.

(iii) The last one is the presence of anticorrelations, which arise only for non-reversible dynamics such as Langevin dynamics.

A proper estimation of the asymptotic variance requires to compute autocorrelation profiles on a sufficiently long time interval $[0, t_{\mathrm{deco}}]$ (here $t_{\mathrm{deco}} = 6$). One can check a posteriori that this time is sufficient by looking at the convergence of the cumulated autocorrelation $t \mapsto \int_0^t C_\varphi$ toward its limit $\frac{\sigma^2_\varphi}{2} = \int_0^\infty C_\varphi$ (see Appendix 3.9 for more details on the variance estimators, and the computation of error bars for these quantities).

Figure 3.3 compares the autocorrelation profile of the velocity with the ones for the modified observables, for two different Galerkin basis sizes M and two different forcing amplitudes η. For a small forcing $\eta = 0.08$ the two modified observables have an amplitude and a decorrelation time which are both much smaller than for the standard velocity. Note that the two modified observables do not exhibit any anti-correlation, contrarily to the velocity observable. The cumulated plots show that the control variates drastically reduce the asymptotic variance in this case, especially for the one based on a more accurate Galerkin approximation. For a larger value $\eta = 1.28$ the modified observables have a significantly larger amplitude (*i.e.* $C_\varphi(0)$ is larger), especially in the case of a low accuracy M. However

Figure 3.2: Illustrative autocorrelation profile. The dashed line is the exponential envelope of the correlation function.



the decorrelation times are small and there is anti-correlation, resulting in a reasonable variance reduction in both cases.

Asymptotic variances. Let us now compute the asymptotic variance for a whole range of Galerkin accuracies M and forcing amplitudes η. The results presented in Figure 3.4 confirm that for a very accurate Galerkin resolution the variance of the modified observable scales as η^2 with the prefactor $\alpha = \left\langle \Pi_0 AR, -\mathcal{L}_0^{-1}\Pi_0 AR \right\rangle_0$ predicted theoretically in Theorem 3.1. This prefactor has been computed independently by solving (3.23) using a Galerkin method and plugging this approximation in (3.24). When the Galerkin discretization is not sufficiently accurate, the variance reaches a plateau in the region of small forcings as predicted by Theorem 3.2.

3.4 Thermal transport in atom chains

Thermal transport in one-dimensional systems has been the topic of many investigations, both from theoretical and numerical points of view [20, 106, 41]. Determining which microscopic ingredients influence the scaling of the conductivity with respect to the length of the chain is still an active line of research. Studying numerically this scaling requires to simulate chains of thousands of particles. In these systems the temperature gradient and the thermal flux are both very small, which induces large statistical errors when estimating the conductivity. Introducing variance reduction techniques not requiring the knowledge of the invariant probability measure could alleviate (at least partly) these difficulties.

3.4.1 Full dynamics

3.4.1.1 Equations of motion

We consider a chain composed of N particles interacting through a nearest-neighbor potential v (see Figure 3.5). The evolution is dictated by a Hamiltonian dynamics and a

Figure 3.3: Left: Autocorrelation profile of the velocity compared to the one of the modified observables for two different accuracies, either for a small forcing $\eta = 0.08$ (top) or a larger one $\eta = 1.28$ (bottom). Right: Corresponding cumulated autocorrelations. The limit value is half the asymptotic variance of the observable.

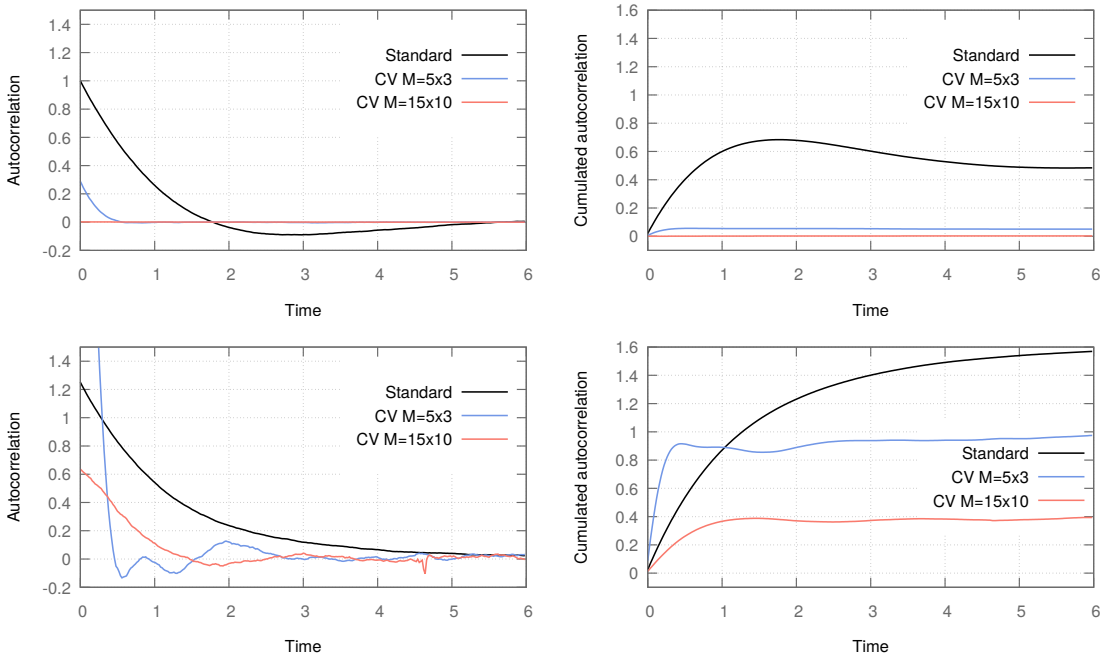
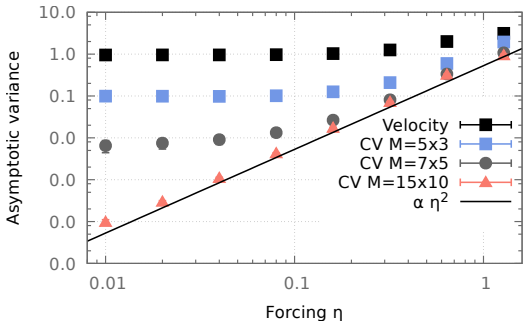


Figure 3.4: Asymptotic variance of the velocity (black squares) compared to its counterpart when using a control variate (blue, grey, red) and to the reduced variance (black line) predicted theoretically ($\alpha \approx 0.53$ computed with a Galerkin discretization).



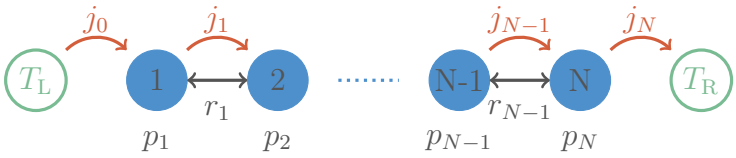


Figure 3.5: Heat transport in a one-dimensional chain.

thermalization mechanism at the boundaries, where the first and the last particles are submitted to Ornstein-Uhlenbeck processes, at temperatures T_{L} and T_{R} respectively. The unknowns are the momenta $p = (p_n)_{1 \leqslant n \leqslant N}$ of the particles and the interparticle distances $r = (r_n)_{1 \leqslant n \leqslant N-1}$. In these variables, the dynamics reads:

$$\begin{cases} \mathrm{d}r_n = \dfrac{1}{m}(p_{n+1} - p_n)\,\mathrm{d}t, \\ \mathrm{d}p_1 = v'(r_1)\,\mathrm{d}t - \dfrac{\gamma}{m}p_1\,\mathrm{d}t + \sqrt{2\gamma T_{\mathrm{L}}}\mathrm{d}W_t^L, \\ \mathrm{d}p_n = (v'(r_n) - v'(r_{n-1}))\,\mathrm{d}t, \\ \mathrm{d}p_N = -v'(r_{N-1})\,\mathrm{d}t - \dfrac{\gamma}{m}p_N\,\mathrm{d}t + \sqrt{2\gamma T_{\mathrm{R}}}\mathrm{d}W_t^R, \end{cases} \tag{3.26}$$

where $m > 0$ is the mass of a particle, $\gamma > 0$ is the friction coefficient, $T_{\mathrm{L}} \geqslant T_{\mathrm{R}}$ and W_t^L, W_t^R are two independent standard one-dimensional Brownian motions. Notice that the ends of the chain are free. The Hamiltonian of the system is the sum of the potential and kinetic energies:

$$H(r,p) = V(r) + \sum_{n=1}^{N} \frac{p_n^2}{2m}, \quad V(r) = \sum_{n=1}^{N-1} v(r_n).$$

The infinitesimal generator of the dynamics (3.26) reads

$$\begin{aligned} \mathcal{L} = \frac{1}{m}\sum_{n=1}^{N-1}(p_{n+1} - p_n)\partial_{r_n} + \sum_{n=1}^{N}(v'(r_n) - v'(r_{n-1}))\partial_{p_n} \\ - \frac{\gamma}{m}p_1\partial_{p_1} + \gamma T_L\partial_{p_1}^2 - \frac{\gamma}{m}p_N\partial_{p_N} + \gamma T_R\partial_{p_N}^2, \end{aligned}$$

using the convention $v'(r_0) = v'(r_N) = 0$.

3.4.1.2 Properties of the dynamics

Let us recall some properties of the dynamics (3.26) which hold under the following assumption.

Assumption 3.6. *The interaction potential v is $\mathcal{C}^\infty$ and there exist $k \geqslant 2$ and $a > 0$ such that*

$$\forall r_1 \in \mathbb{R}, \quad \lim_{\tau \to +\infty} \tau^{-k} v(\tau r_1) = a|r_1|^k, \quad \lim_{\tau \to +\infty} \tau^{1-k} v'(\tau r_1) = ka|r_1|^{k-1}\mathrm{sign}(r_1).$$

Moreover the interaction potential is not degenerate: for any $r_1 \in \mathbb{R}$ *there exists* $m = m(r_1) \geqslant 2$ *such that* $\partial_m v(r_1) \neq 0$.

These conditions hold for the potentials we use in the numerical simulations reported in Section 3.4.3. When Assumption 3.6 holds, the dynamics admits a unique invariant probability measure π (see [34]). This invariant probability measure is explicit when the chain is at equilibrium ($T_{\mathrm{L}} = T_{\mathrm{R}} = \beta^{-1}$), in which case it has the tensorized form

$$\pi_{\mathrm{eq}}(\mathrm{d}r\,\mathrm{d}p) = Z_\beta^{-1} \exp\left(-\beta\left(\frac{|p|^2}{2m} + V(r)\right)\right) \mathrm{d}r\,\mathrm{d}p, \tag{3.27}$$

where Z_β^{-1} is a normalization constant. Let us emphasize that the reference system considered later on in Section 3.4.2 is not at equilibrium. Following the framework considered in [34] (which is itself based on [138]), we consider in this section the Lyapunov functions $\mathcal{K}_\theta = \mathrm{e}^{\theta H}$. There exist $\theta_* > 0$ such that $\mathcal{K}_\theta \subset \mathrm{L}^2(\pi)$ for any $\theta \in [0, \theta_*)$. The functional spaces we use are also indexed by the continuous parameter $\theta \in [0, \theta_*)$:

$$\mathrm{L}^\infty_\theta = \left\{\varphi \text{ measurable} \;\middle|\; \|\varphi\,\mathrm{e}^{-\theta H}\|_{\mathrm{L}^\infty} < +\infty\right\},$$

and the space $\mathcal{S}$ is defined similarly to (3.12). For $\theta \in [0, \theta_*)$, we also define the vector space $\mathrm{L}^\infty_{\theta,0}$ of functions of L^∞_θ with mean zero with respect to π. One can prove the exponential decay of the semi-group on the associated functional space $\mathrm{L}^\infty_{\theta,0}$ (see [34]): for any $\theta \in [0, \theta_*)$, there exist $C, \lambda > 0$ such that, for any $\varphi \in \mathrm{L}^\infty_{\theta,0}$,

$$\forall t \geqslant 0, \quad \|\mathrm{e}^{t\mathcal{L}}\varphi\|_{\mathrm{L}^\infty_\theta} \leqslant C\mathrm{e}^{-\lambda t}\|\varphi\|_{\mathrm{L}^\infty_\theta}.$$

This implies that $\mathcal{L}$ is invertible on $\mathrm{L}^\infty_{\theta,0}$, and that its inverse is bounded.

Validity of Assumptions 1 and 2. Assumption 3.1 holds true since there exist a unique invariant probability measure with positive density with respect to the Lebesgue measure, and the generator of the dynamics is hypoelliptic [34, 89, 138]. The first part of Assumption 3.2 is also satisfied for $\theta \in [0, \theta_*)$. Note that at equilibrium ($T_{\mathrm{L}} = T_{\mathrm{R}} = \beta^{-1}$) the invariant probability measure π is explicit and $\theta^* = \beta/2$. The product of two Lyapunov functions $\mathcal{K}_\theta$ and $\mathcal{K}_{\theta'}$ is in a Lyapunov space only if $\theta + \theta' < \theta_*$, so the second part of Assumption 3.2 is not satisfied.

3.4.1.3 Heat flux and conductivity

When studying heat transport in atom chains the typical quantity of interest is the thermal flux through the chain:

$$\forall n \in [1, N-1], \quad j_n(r,p) = -\frac{p_n + p_{n+1}}{2} v'(r_n), \tag{3.28}$$

see for example the review [105] on thermal transport in low-dimensional lattices for further background material. We also make use of the two boundary elementary fluxes:

$$j_0(r,p) = \frac{\gamma}{m}\left(T_{\rm L} - \frac{p_1^2}{m}\right), \qquad j_N(r,p) = \frac{\gamma}{m}\left(\frac{p_N^2}{m} - T_{\rm R}\right). \tag{3.29}$$

The definition of the elementary fluxes j_n is motivated by the local energy balance, centered on particle n:

$$\forall n \in [1,N], \quad \mathcal{L}\varepsilon_n = j_{n-1} - j_n, \qquad \varepsilon_n(r,p) = \frac{v(r_{n-1})}{2} + \frac{p_n^2}{2m} + \frac{v(r_n)}{2}. \tag{3.30}$$

The quantity $\mathcal{L}\varepsilon_n$ is of mean zero since it is in the image of the generator. Therefore the elementary fluxes j_n all have the same stationary values:

$$\mathbb{E}_\pi[j_0] = \mathbb{E}_\pi[j_1] = \cdots = \mathbb{E}_\pi[j_N]. \tag{3.31}$$

Any linear combinations of such fluxes, namely

$$J_\lambda = \sum_{n=0}^{N} \lambda_n j_n, \qquad \sum_{n=0}^{N} \lambda_n = 1, \tag{3.32}$$

has the same stationary value. The most common choice is the spatial mean

$$\widetilde{R} = \frac{1}{N-1}\sum_{n=1}^{N-1} j_n. \tag{3.33}$$

We call the latter observable "standard heat flux" in the sequel. Notice that it does not depend on the boundary fluxes j_0 and j_N. The linear response of $\widetilde{R}$ (or of any flux J_λ) with respect to the temperature gradient $\frac{T_{\rm L}-T_{\rm R}}{N-1}$ defines the effective conductivity:

$$\kappa = \frac{N-1}{T_{\rm L}-T_{\rm R}}\mathbb{E}_\pi[\widetilde{R}], \tag{3.34}$$

which is here the transport coefficient of interest. There exist infinitely many observables having the same expectation as $\widetilde{R}$, see (3.32) for example. Let us first discuss the choice of observable for the heat flux, before trying to reduce its variance by constructing a control variate.

Asymptotic variance of the heat fluxes at equilibrium. The chain is supposed to be at equilibrium in all this paragraph ($T_{\rm L} = T_{\rm R} = \beta^{-1}$). The conclusions remain unchanged for nonequilibrium systems when $T_{\rm L} - T_{\rm R}$ is small since the results are only perturbed to the first order with respect to this quantity. In the remainder of Section 3.4, an index 'eq' refers to the equilibrium dynamics and to the equilibrium invariant probability measure $\pi_{\rm eq}$. In this setting the asymptotic variance $\sigma^2_{\widetilde{R},\rm eq}$ for the standard heat flux $\widetilde{R}$ is not smaller than the one associated with any elementary flux $(j_n)_{1\leqslant n\leqslant N-1}$. These two variances are in fact

equal, as made precise in the following proposition (similar in spirit to Remark 3.4).

Proposition 3.1. *Consider an observable $\varphi \in \mathcal{S}$ and a function $U \in \mathcal{S}$ which does not depend on p_1 and p_N. Then adding $\mathcal{L}U$ to the observable does not modify the variance:*

$$\sigma^2_{\varphi+\mathcal{L}U,\mathrm{eq}} = \sigma^2_{\varphi,\mathrm{eq}}. \tag{3.35}$$

Proof. At equilibrium, the invariant probability measure π_{eq} is explicit. The symmetric part of the generator can then be computed and it corresponds to the fluctuation-dissipation part of the process:

$$\frac{1}{2}(\mathcal{L} + \mathcal{L}^*) = \mathcal{L}_{\mathrm{FD}} := -\frac{\gamma}{\beta}\left(\partial^*_{p_1}\partial_{p_1} + \partial^*_{p_N}\partial_{p_N}\right), \tag{3.36}$$

where adjoints are considered on $\mathrm{L}^2(\pi_{\mathrm{eq}})$. When U does not depend on p_1 nor p_N, it holds $\mathcal{L}_{\mathrm{FD}}U = 0$. The claimed result then follows from Lemma 3.2. □

The equality (3.35) is perturbed by terms of order $T_{\mathrm{L}} - T_{\mathrm{R}}$ for out of equilibrium dynamics according to linear response theory. Upon taking for U a linear combination of the energies $(\varepsilon_n)_{2\leqslant n\leqslant N-1}$, we directly obtain, thanks to (3.30), that all the fluxes of the form (3.32) which do not depend on the boundary fluxes ($\lambda_0 = \lambda_N = 0$) share the same asymptotic variance at equilibrium; in particular

$$\forall 1 \leqslant n \leqslant N-1, \quad \sigma^2_{j_n,\mathrm{eq}} = \sigma^2_{\widetilde{R},\mathrm{eq}}. \tag{3.37}$$

Remark 3.5. *Linear response theory indicates that the previous asymptotic variances are related to the conductivity through the Green–Kubo formula [97]:*

$$\sigma^2_{\widetilde{R},\mathrm{eq}} = \frac{2\kappa}{\beta^2(N-1)}. \tag{3.38}$$

We show in Appendix 3.7.2 that, in this equilibrium situation, the variance of the two boundary fluxes j_0 and j_N is also related to the conductivity as:

$$\sigma^2_{j_0,\mathrm{eq}} = \frac{\gamma}{m\beta^2} - \frac{2\kappa}{\beta^2(N-1)}, \qquad \sigma^2_{j_N,\mathrm{eq}} = \frac{\gamma}{m\beta^2} - \frac{2\kappa}{\beta^2(N-1)}. \tag{3.39}$$

Apart from special cases such as integrable systems (as the harmonic system considered in Appendix 3.7.4), we generically observe numerically that $\frac{\kappa(N)}{N-1} \xrightarrow[N\to\infty]{} 0$. The boundary fluxes therefore have (asymptotically in N) a larger variance than bulk fluxes. Since the variances are perturbed to first order in $T_{\mathrm{L}} - T_{\mathrm{R}}$ in nonequilibrium situations, the same conclusion holds for temperature differences which are not too large.

In the next section we construct a modified observable by adding a control variate to the reference observable $R = \frac{1}{2}(j_0 + j_N)$. This particular heat flux does not depend on the potential energy function v, which simplifies the computation of the control variate (see Appendix 3.7.4). In Appendix 3.7.1 we prove using Proposition 3.1 that, at equilibrium, the asymptotic variance of the resulting modified observable does not depend on the choice of the reference observable R.

3.4.2 Simplified dynamics and control variate

We split the interaction potential into a harmonic part with parameters $\hat{\omega} > 0$ and $\hat{r} \in \mathbb{R}$, and an anharmonic part w:

$$v(r_1) = v_0(r_1) + w(r_1), \qquad v_0(r_1) = \frac{1}{2} m\hat{\omega}^2 (r_1 - \hat{r})^2. \tag{3.40}$$

The potential energy is then decomposed as

$$V(r) = V_0(r) + W(r), \qquad V_0(r) = \sum_{n=1}^{N-1} v_0(r_n), \qquad W(r) = \sum_{n=1}^{N-1} w(r_n).$$

Following the general strategy outlined in Section 3.2 we decompose the generator as

$$\mathcal{L} = \mathcal{L}_0 + \widetilde{\mathcal{L}},$$

where $\mathcal{L}_0$ is the generator of the harmonic chain corresponding to the harmonic interaction potential v_0 and $\widetilde{\mathcal{L}}$ is the generator of the anharmonic perturbation:

$$\widetilde{\mathcal{L}} = \sum_{n=1}^{N} \left(w'(r_n) - w'(r_{n-1}) \right) \partial_{p_n},$$

with the same convention $w'(r_0) = w'(r_N) = 0$ as for the potential v. We simplify the Poisson problem for the optimal control variate

$$-\mathcal{L}\Phi = R - \mathbb{E}[R],$$

into the harmonic Poisson problem

$$-\mathcal{L}_0 \Phi_0 = R - \mathbb{E}_0[R]. \tag{3.41}$$

Note that the observable R does not depend on the potential, contrarily to other heat fluxes such has $\widetilde{R}$ in (3.33), so that the right hand side of the Poisson equation needs not be changed when looking for an approximate control variate. The equation (3.41) can be solved analytically for Φ_0. In Appendix 3.7.4 we show that

$$\begin{aligned} \mathbb{E}_0[R] &= \frac{\nu^2}{1+\nu^2} \frac{\gamma(T_{\mathrm{L}} - T_{\mathrm{R}})}{2m}, \\ \Phi_0(r,p) &= \frac{m}{2\gamma\left(1+\nu^2\right)} \left[-\hat{\omega}^2 \sum_{n=1}^{N-1} (r_n - \hat{r})(p_n + p_{n+1}) + \frac{\gamma}{2m^2} \left(p_N^2 - p_1^2 \right) \right] + C, \end{aligned} \tag{3.42}$$

where $\nu = \frac{m\hat{\omega}}{\gamma}$ and $C \in \mathbb{R}$ is such that $\mathbb{E}_0[\Phi_0] = 0$. The modified observable is therefore

$$\begin{aligned}
&(R + \mathcal{L}\Phi_0)(r,p) = \mathbb{E}_0[R] + \widetilde{\mathcal{L}}\Phi_0(r,p) \\
&= \frac{1}{2(1+\nu^2)}\left[\nu\hat{\omega}(T_{\rm L} - T_{\rm R}) - \nu\hat{\omega}\sum_{n=1}^{N-1}(r_n - \hat{r})\left(w'(r_{n+1}) - w'(r_{n-1})\right) - \left(\frac{p_1}{m}w'(r_1) + \frac{p_N}{m}w'(r_{N-1})\right)\right] \\
&= \frac{1}{2(1+\nu^2)}\left[\nu\hat{\omega}(T_{\rm L} - T_{\rm R}) - \sum_{n=1}^{N-1}\left(\widetilde{v}_{n+1}(r,p) - \widetilde{v}_{n-1}(r,p)\right)w'(r_n)\right],
\end{aligned} \tag{3.43}$$

where

$$\widetilde{v}_n(r,p) = \begin{cases} -\frac{p_1}{m} & \text{if} \quad n = 0, \\ -\nu\hat{\omega}(r_n - \hat{r}) & \text{if} \quad 1 \leqslant n \leqslant N-1, \\ \frac{p_N}{m} & \text{if} \quad n = N. \end{cases}$$

Notice that, by construction, this observable is constant when the chain is harmonic (*i.e.* $w = 0$).

Harmonic fitting. For a given pair potential $v = v(r)$, there is some freedom in the decomposition (3.40), namely the choice of the parameters $\hat{\omega}$ and $\hat{r}$. The optimal choice would be such that the variance of the modified observable (3.43) is minimal, but this condition is not practical. A possible (and simpler) heuristic is to choose these coefficients in order to minimize the $\mathrm{L}^2(\pi_{\rm eq})$ norm of the anharmonic force $-\nabla W$ at equilibrium, namely when $T_{\rm L} = T_{\rm R} = \beta^{-1}$. In view of the tensorized form (3.27) of the invariant probability measure at equilibrium,

$$\|\nabla W(r)\|_{\rm eq}^2 = (N-1)z_\beta^{-1}\int_{\mathbb{R}}\left(v'(r_1) - m\hat{\omega}^2(r_1 - \hat{r})\right)^2 \mathrm{e}^{-\beta v(r_1)}\mathrm{d}r_1,$$

where $z_\beta = \int_{\mathbb{R}} \mathrm{e}^{-\beta v(r_1)}\mathrm{d}r_1$. Therefore the minimization problem defining $\hat{r}$ and $\hat{\Omega} = m\hat{\omega}^2$ writes

$$\operatorname*{argmin}_{\hat{\omega},\hat{r}} \int_{\mathbb{R}}\left(v'(r_1) - \hat{\Omega}(r_1 - \hat{r})\right)^2 \mathrm{e}^{-\beta v(r_1)}\mathrm{d}r_1. \tag{3.44}$$

There exists a minimizer $(\hat{r}, \hat{\Omega})$ since the function to be minimized is continuous and coercive; uniqueness is proved in Appendix 3.7.3. Define the moments of the marginal measure for inter-particle distances as:

$$\mathcal{M}_k = \int_{\mathbb{R}} r_1^k \mathrm{e}^{-\beta v(r_1)}\mathrm{d}r_1.$$

The Euler-Lagrange equation associated with (3.44) provides the expression of the minimizer (see Appendix 3.7.3):

$$\hat{r} = \frac{\mathcal{M}_1}{\mathcal{M}_0}, \qquad \hat{\Omega} = m\hat{\omega}^2 = \beta^{-1}\frac{\mathcal{M}_0^2}{\mathcal{M}_0\mathcal{M}_2 - \mathcal{M}_1^2}, \tag{3.45}$$

where, by the Cauchy-Schwarz inequality, $\mathcal{M}_0\mathcal{M}_2 - \mathcal{M}_1^2 > 0$ for any continuous potential v which is not constant.

Validity of Assumptions 3 to 5. The standard way to verify Assumption 3.3 would be to show that the coefficients in the Lyapunov condition exhibited in [34] depend continuously on perturbations of the potential v. This is not straightforward, especially if the exponent k introduced in Assumption 3.6 is discontinuous with the perturbation amplitude at $\eta = 0$ (which corresponds to the harmonic chain). For example, for the FPU potential (3.46), $k = 2$ for the harmonic chain ($\eta = 0$) and $k = 4$ for $\eta > 0$. We show that Assumption 3.4 holds under Assumption 3.6 in Appendix 3.7.5. Moreover it is clear that $\mathcal{S}$ is stable by $\widetilde{\mathcal{L}}$. A simple computation shows that π_0 is a Gaussian probability measure, which implies that $\widetilde{\mathcal{L}}^*\mathbf{1} \in \mathrm{L}^2(\pi_0)$. Therefore Assumption 3.5 holds as well.

3.4.3 Numerical results

We consider a Fermi-Pasta-Ulam (FPU) potential

$$v(r_1) = \frac{a}{2}r_1^2 + \frac{b}{3}r_1^3 + \frac{c}{4}r_1^4, \tag{3.46}$$

where $c = \frac{b^2}{3a}$ is such that $v''(r_1) = a + 2br_1 + 3cr_1^2 = \frac{1}{a}(a + br_1)^2$ is positive except at a single point where it vanishes. This choice makes the potential both asymmetric and convex. Symmetric potentials indeed exhibit special behaviors [151], whereas we want to be as general as possible. On the other hand, non-convex potentials are not typical in the literature on the computation of thermal conduction in one-dimensional chains. In the following we fix $a = 1$ and vary b only. The parameters $\hat{r}$ and $\hat{\omega}$ are given by (3.45), where the moments $\mathcal{M}_k$ are computed using one-dimensional numerical quadratures.

The system is simulated for a time $T = 10^8$ with time steps $\Delta t = 10^{-2}$, and with $m = \gamma = 1$. The atom chain is simulated either at equilibrium with $T_\mathrm{L} = T_\mathrm{R} = 2$, or for $T_\mathrm{L} = 3$ and $T_\mathrm{R} = 1$. The dynamics is discretized using a Geometric Langevin Algorithm scheme as in (3.25). The estimator of the asymptotic variance is made precise in Appendix 3.9. The decorrelation time is set to $t_\mathrm{deco} = 3N$ for the standard flux R and to $t_\mathrm{deco} = 32$ for the modified observable.

We plot in Figure 3.6 the autocorrelation profiles of the heat flux for a chain of size $N = 128$ at equilibrium, for two different anharmonicities. Let us comment this picture in greater detail. The results first show that the chosen decorrelation time t_deco is sufficiently large. Similar plots were used to check this is also the case for all the range of anharmonicities b and numbers of particles N we consider. They can also be used to understand the eventual variance reduction granted by the control variate (3.42). For a small anharmonicity $b = 0.08$ we see that the asymptotic variance, which is twice the limit of the cumulated autocorrelation (right plot) is greatly reduced for two reasons. First, the signal amplitude, which is related to the autocorrelation value at $t = 0$ (left plot), is slightly smaller, and (right plot) the contribution of the times $0 \leqslant t \leqslant 2$ is twice smaller for the modified flux. Second, and that is the actual reason for the variance reduction, there is anticorrelation for $2 \leqslant t \leqslant 5$. For a larger anharmonicity it appears that the amplitude of the modified observable is much larger, but this is compensated by a long-time anticorrelation. The resulting asymptotic variance of the modified flux is slightly smaller than the one of the

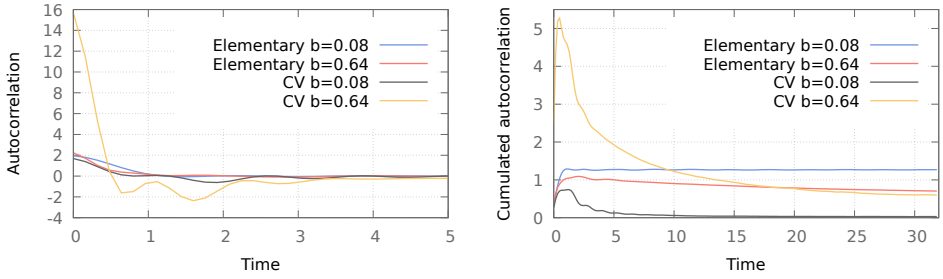


Figure 3.6: Comparison of results obtained either with the elementary flux j_n (3.28) or with the modified observable (3.43), for two different anharmonicities $b = 0.08, 0.64$, and for a chain of $N = 128$ particles at equilibrium. Left: Autocorrelation profile. Right: Cumulated autocorrelation $t \mapsto \int_0^t C$ at longer times.

standard flux. The plots are essentially the same out of equilibrium, when $T_{\mathrm{L}} = 3$ and $T_{\mathrm{R}} = 1$ (numerical results not presented here).

The asymptotic variances, with associated error bars (see Appendix 3.9), are plotted on Figure 3.7 for a whole range of anharmonicities b and numbers of particles N. The two left plots are at equilibrium ($T_{\mathrm{L}} = T_{\mathrm{R}} = 2$) while the two right plots are out of equilibrium ($T_{\mathrm{L}} = 3$ and $T_{\mathrm{R}} = 1$). We check that the variances are extremely similar in both cases, which is expected by linear response theory. We observe that the asymptotic variance of the modified flux scales as b^2 for $b \ll 1$, as expected from Theorem 3.1, providing an excellent variance reduction in this case. Note that, in the limit $b \to 0$, the variance of the standard flux tends to $\frac{\gamma\nu^2}{m\beta^2(1+\nu^2)} = 2$ (since $\nu = \frac{m\omega}{\gamma} = 1$ here), which is the theoretical value predicted at equilibrium in view of the expression of the mean flux for a harmonic chain (see Equations (3.42), (3.38), and (3.34)). The modified flux can sometimes have a larger variance than the standard one, for example in the regime $b = \mathrm{O}(1)$ and N large. Note that for the particular choice $\hat{\omega} = 0$ and $\hat{r} = 0$ the modified flux is $\frac{1}{2}(j_1 + j_{N-1})$, which has the same asymptotic variance as the standard flux $\widetilde{R}$ according to (3.37). Therefore, for any set of parameters, there exist an optimal choice of the coefficients $\hat{\omega}, \hat{r}$ providing a modified flux whose asymptotic variance is smaller or equal to its counterpart without control variate. In the present application these two coefficients are instead chosen according to the heuristic (3.45), leading to a degradation of the asymptotic variance in certain cases.

3.5 Solvated dimer under shear

A solvated dimer is a pair of bonded particles in a bath constituted of many other particles. It serves as a prototypical model of a molecule in solvent (e.g. peptide in water). This model has been used in the context of free energy computations [103]. We apply to this system an external shearing force as in [88], also coined sinusoidal transverse field in the physics literature (see [159, Section 9.1] and [54, 158]), so that the invariant measure of the system is not known. A typical question is the influence of the shear force on the average

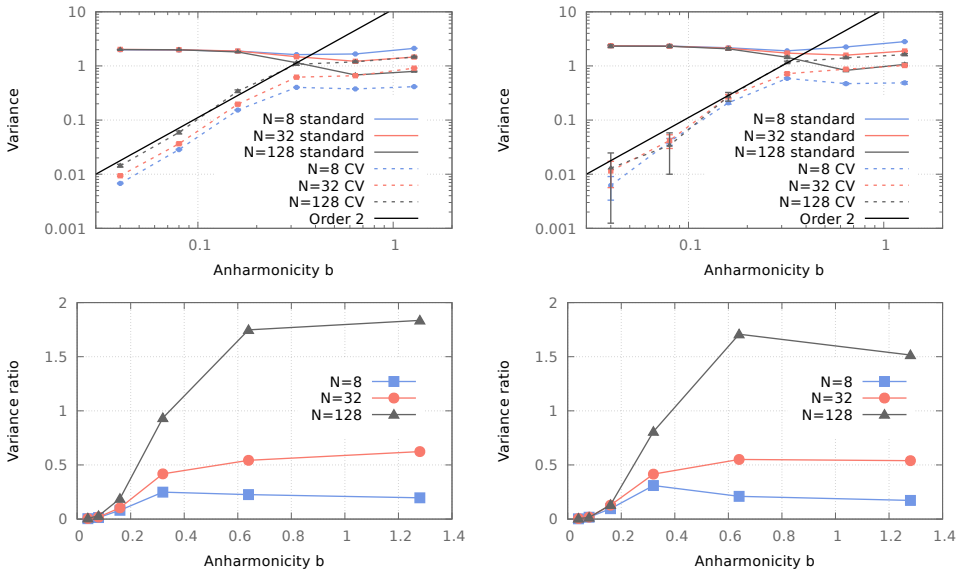


Figure 3.7: Comparison of variances for the standard and modified fluxes. The reported variance ratio corresponds to the modified variance divided by the standard one. Left: Chain at equilibrium ($T_{\mathrm{L}} = T_{\mathrm{R}} = 2$). Right: Chain out of equilibrium ($T_{\mathrm{L}} = 3$ and $T_{\mathrm{R}} = 1$).

bond length of the dimer.

3.5.1 Full dynamics

We consider N particles in a two-dimensional box of length L with periodic boundary conditions, with positions $q = (q_1, \cdots, q_N) \in \mathcal{D} = (L\mathbb{T})^{2N}$. Two of these particles, with positions q_1 and q_2, form a dimer whereas the other $N-2$ particles, with positions $q_3, \cdots, q_N$, constitute the solvent. The potential energy of the system is composed of three parts:

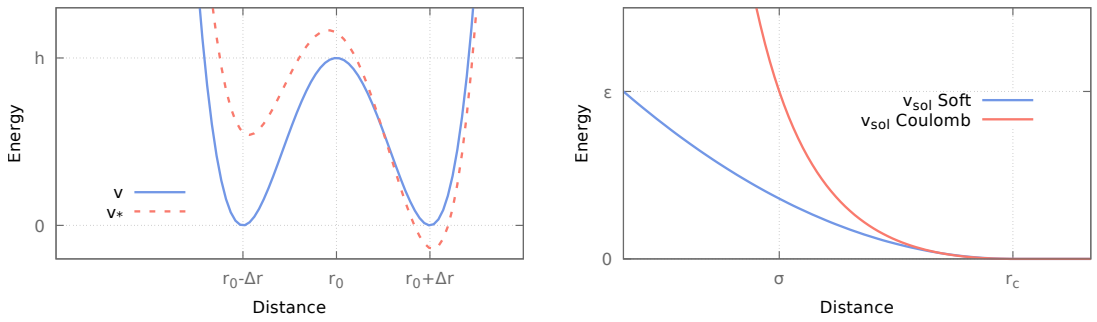
$$\begin{aligned} V(q) &= v(|q_1 - q_2|) + \sum_{i\in\{1,2\}} \sum_{j=3}^{N} v_{\mathrm{sol}}(|q_i - q_j|) + \sum_{3\leqslant i<j\leqslant N} v_{\mathrm{sol}}(|q_i - q_j|) \\ &=: V_{\mathrm{dim}}(q) + V_{\mathrm{inter}}(q) + V_{\mathrm{sol}}(q), \end{aligned}$$

where V_{dim} is the potential energy of the dimer, V_{inter} is the interaction energy between the dimer and the solvent and V_{sol} is the potential energy of the solvent. The two particles forming the dimer interact via a double-well potential: denoting by $r = |q_1 - q_2|$ the bond length,

$$v(r) = h\left[1 - \left(\frac{r - r_0}{\Delta r}\right)^2\right]^2, \tag{3.47}$$

where $r_0, \Delta r > 0$ (see Figure 3.8, Left). The potential v presents two minima: one associated with a compact state of length $r = r_0 - \Delta r$ and one associated with a stretched state of length $r = r_0 + \Delta r$. These minima are separated by a potential barrier of height h. The

Figure 3.8: Pairwise potentials for the solvated dimer model. Left: Potential for the dimer and associated free energy in vacuum (see (3.52)). Right: Potentials for the solvent interaction.



particles of the solvent interact both with the other particles of the solvent and the particles of the dimer through a purely repulsive potential. In the following we consider two types of potentials with compact support (see Figure 3.8, Right): a soft repulsion potential (used in [69] for example)

$$\forall r > 0, \quad v_{\mathrm{sol}}(r) = \varepsilon \left(1 - \frac{r}{r_{\mathrm{cut}}}\right)^2 \mathbb{1}_{r \leqslant r_{\mathrm{cut}}}, \tag{3.48}$$

where $\varepsilon, r_{\mathrm{cut}} > 0$; and a singular Coulomb-like potential:

$$\forall r > 0, \quad v_{\mathrm{sol}}(r) = \varepsilon \left(\frac{\frac{1}{\sqrt{r}} - \frac{1}{\sqrt{r_{\mathrm{cut}}}}}{\frac{1}{\sqrt{\sigma}} - \frac{1}{\sqrt{r_{\mathrm{cut}}}}}\right)^2 \mathbb{1}_{\mathrm{r<r_{cut}}} = \varepsilon \frac{\sigma}{r} \left(\frac{1 - \sqrt{\frac{r}{r_{\mathrm{cut}}}}}{1 - \sqrt{\frac{\sigma}{r_{\mathrm{cut}}}}}\right)^2 \mathbb{1}_{r \leqslant r_{\mathrm{cut}}}. \tag{3.49}$$

This potential behaves like $\frac{1}{r}$ for $r \to 0$, reaches the value ε at $r = \sigma$ and vanishes at $r = r_{\mathrm{cut}}$,where its derivative also vanishes. Note that we recover the Coulomb potential $\varepsilon \frac{\sigma}{r}$ in the limit $r_{\mathrm{cut}} \to +\infty$.

The system is driven out of equilibrium by a shearing force of amplitude ν. More precisely, a particle located at a position $(q_{i,x}, q_{i,y})$ experiences the force [88, 159]:

$$(0, f(q_{i,x})) = \left(0, \nu \sin\left(2\pi \frac{q_{i,x}}{L}\right)\right).$$

This force is in the y direction and depends only on x. It therefore induces a non-equilibrium forcing since it is not of gradient type. We are interested in computing the mean length of the dimer $R(q,p) = |q_1 - q_2|$ as a function of this external forcing. The corresponding average is denoted by $\mathbb{E}[|q_1 - q_2|]$.

For simplicity we study the overdamped dynamics associated with V, but everything can be adapted to the Langevin case. Since the space $\mathcal{D}$ is compact and the noise in the dynamics is non degenerate, there exist a unique invariant probability measure π by the Doeblin condition when the potentials under consideration are smooth. This invariant

measure depends on ν, and is not explicit. Proving a similar result for singular potentials such as the Coulomb-like potential (3.49) would require more work.

The generator can be decomposed as:

$$\mathcal{L} = -\nabla V(q)\cdot\nabla + \beta^{-1}\Delta + \nu\sum_{i=1}^{N} f(q_{i,x})\partial_{q_{i,y}} = \mathcal{L}_{\rm dim} + \mathcal{L}_{\rm inter} + \mathcal{L}_{\rm sol} + \nu\mathcal{L}_{\rm pert},$$

where

$$\mathcal{L}_{\rm dim} = \sum_{i=1,2}\left(-\nabla_{q_i}V_{\rm dim}(q)\cdot\nabla_{q_i} + \beta^{-1}\Delta_{q_i}\right), \qquad \mathcal{L}_{\rm inter} = -\nabla V_{\rm inter}(q)\cdot\nabla,$$
$$\mathcal{L}_{\rm sol} = \sum_{i=3}^{N}\left(-\nabla_{q_i}V_{\rm sol}(q)\cdot\nabla_{q_i} + \beta^{-1}\Delta_{q_i}\right), \qquad \mathcal{L}_{\rm pert} = \sum_{i=1}^{N} f(q_{i,x})\partial_{q_{i,y}}.$$

Note that $\mathcal{L}_{\rm dim}$ is the generator of the dynamics of the dimer at equilibrium in vacuum and $\mathcal{L}_{\rm sol}$ is the generator of the dynamics of the solvent at equilibrium and without dimer.

3.5.2 Simplified dynamics and control variate

We consider the following reference Poisson equation where the system is at equilibrium and the interaction between the dimer and the solvent has been switched off:

$$-\mathcal{L}_0\Phi_0 = R - \mathbb{E}_0[R], \tag{3.50}$$

where

$$\mathcal{L}_0 = \mathcal{L}_{\rm dim} + \mathcal{L}_{\rm sol}.$$

Let us show that this equation admits a solution Φ_0 depending only on the length $|q_1-q_2|$ of the dimer. In order to highlight the dependence on the dimension of the underlying space, let us denote by $d=2$ this dimension. Assume that Φ_0 is defined for any $q\in(L\mathbb{T})^{dN}$ by $\Phi_0(q) = \frac{1}{2}\psi(|q_1-q_2|)$ for some smooth function ψ. The Laplacian of Φ_0 can be rewritten using spherical coordinates as:

$$\Delta\Phi_0(q) = \psi''(|q_1-q_2|) + \frac{d-1}{|q_1-q_2|}\psi'(|q_1-q_2|), \tag{3.51}$$

where $d=2$ is the dimension of the underlying physical space. We obtain by substituting Φ_0 into (3.50) that ψ satisfies the following one-dimensional differential equation:

$$\forall r>0, \quad v_*'(r)\psi'(r) - \beta^{-1}\psi''(r) = r - r^*, \tag{3.52}$$

where $v_*(r) = v(r) - \frac{d-1}{\beta}\ln(r)$ and $r^* = \mathbb{E}_*[r]$ is the expectation of the length r with respect to the probability measure $\pi_*(\mathrm{d}r) = Z_*^{-1}\mathrm{e}^{-\beta v_*(r)}\,\mathrm{d}r$. Note the additional term $-\frac{d-1}{\beta}\ln(r)$ in the expression of v_* coming from (3.51), which can be interpreted as an entropic contribution.

Let us first discuss the well-posedness of (3.52). The double-well potential (3.47) consid-

ered here is such that v_* is a bounded perturbation of a convex function. Therefore $\pi_*(\mathrm{d}r)$ satisfies a log-Sobolev inequality and thus a Poincaré inequality by the Holley-Stroock theorem [82] and the Bakry-Emery criterion [10]. This implies that the one-dimensional Poisson problem (3.52) then admits a unique solution in

$$\mathrm{H}^1(\pi_*) \cap \mathrm{L}^2_0(\pi_*) = \left\{ \varphi \in H^1(\pi_*),\ \int_0^{+\infty} \varphi \, \mathrm{d}\pi_* = 0 \right\}$$

by the Lax-Milgram theorem for the variational formulation:

$$\forall u \in \mathrm{H}^1(\pi_*) \cap \mathrm{L}^2_0(\pi_*), \quad \beta^{-1} \int_0^\infty \psi'(r) u'(r) \, \pi_*(\mathrm{d}r) = \int_0^\infty (r - r_*) u(r) \, \pi_*(\mathrm{d}r).$$

We discuss precisely in Appendix 3.8 how we numerically solve (3.52). Knowing the solution ψ, the corresponding modified observable then writes

$$\begin{aligned}(R + \mathcal{L}\Phi_0)(q) &= |r_{12}| + \beta^{-1}\psi''(|r_{12}|) \\ &\quad + \left[\frac{1}{2}\Big(\nabla_{q_1} V(q) - \nabla_{q_2} V(q) - \nu(f(q_{1,x}) - f(q_{2,x}))e_y \Big) \cdot \frac{r_{12}}{|r_{12}|} + \frac{d-1}{\beta |r_{12}|} \right] \psi'(|r_{12}|),\end{aligned}$$

where $r_{12} = q_2 - q_1$ and $e_y = (0, 1)$. Note that $\nabla_{q_1} V$ and $\nabla_{q_2} V$ are the forces that apply on particles 1 and 2 respectively, which depend also on the solvent variables.

3.5.3 Numerical results

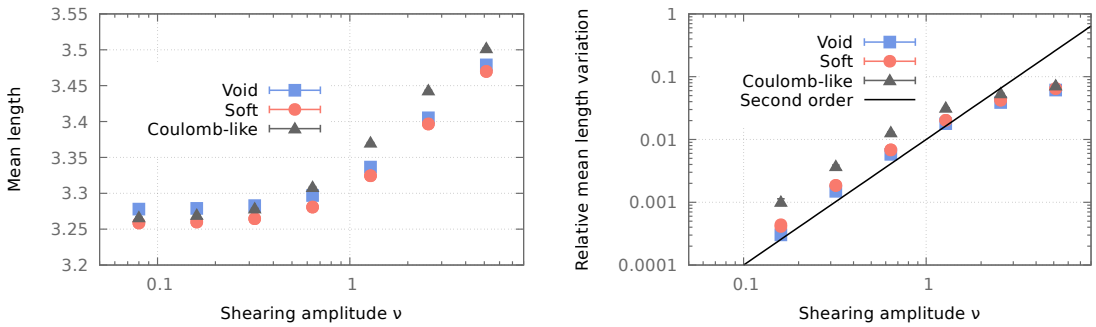
We simulate a system of $N = 64$ particles in $d = 2$ dimensions, using periodic boundary conditions. We fix $L = 8$ (so that the particle density is 1), and $\beta = 1$. The parameters of the potentials are set to $r_{\mathrm{cut}} = 2.5$, $\varepsilon = 1$, $h = 1$, $r_0 = 3$ and $\Delta r = 1$ (see Figure 3.8). For the finite difference method used to solve the Poisson equation (3.52) we use a mesh size $\Delta r = 10^{-3}$ on an interval $[0, r_{\max}]$ with $r_{\max} = 10$ (see Appendix 3.8).

The influence of the shearing on the average dimer length is plotted Figure 3.9. We see that a shear force of amplitude $\nu = 1$ increases the mean length by roughly 1%, and that the response of the mean length to the nonequilibrium forcing is of order 2. The response is small thus difficult to estimate accurately, hence the need for control variates to alleviate this issue.

In the case of an unsolvated dimer, Figure 3.10 (Left) shows that the variance of the modified observable scales like ν^2, as predicted by Theorem 3.1. Note that in the limit $\nu \to 0$ the modified observable is the constant $\mathbb{E}_0[R] = r_*$, which is computed by a numerical quadrature, so that the variance converges to zero.

When the solvent interacts with the dimer the variance of the modified observable plateaus at a certain value when $\nu \to 0$, as expected from Theorem 3.2. For the soft potential (3.48), the variance scales like ν for a forcing amplitude of order 1, which is expected from Theorem 3.2 (see Figure 3.10, Right). The variance stabilizes at a value which is ten times smaller than the initial one. For the Coulomb-like potential the influence of the solvent on the dimer is stronger and the control variate does not perform as well, as seen

Figure 3.9: Left: Mean length of a dimer, either unsolvated (in vacuum) or in a solvent with soft or Coulomb-like potential. Right: Relative variation of this mean length induced by the shearing. The solid line represents the reference scaling ν^2.



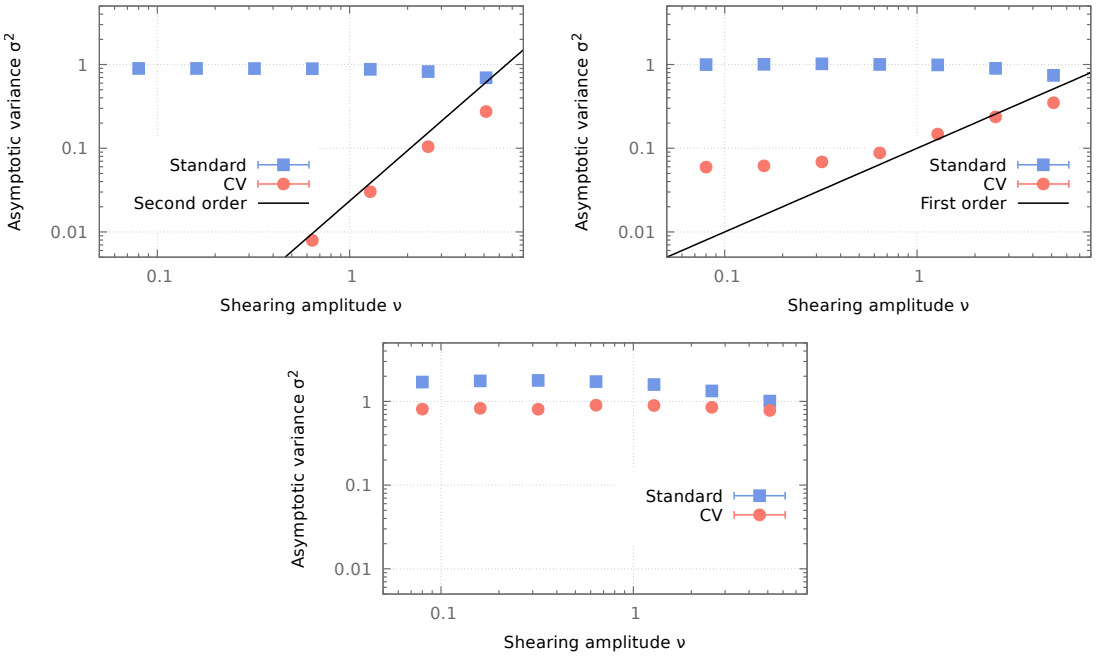
on Figure 3.10 (Bottom). For a small shearing the variance is however reduced by a factor 4.

Generalization. The variance reduction strategy discussed here can be easily adapted to similar systems. For example Langevin dynamics can be treated by replacing (3.52) by a two-dimensional PDE where the variables are the dimer length and the radial part of the momentum associated to this length. The Poisson equation should then be solved using a Galerkin approximation similar to what is done Section 3.3. One could also consider a solvated molecule more complex than a dimer. In this case (3.52) would be posed in several dimensions and thus becomes rapidly impossible to solve in practice. In general one has to reduce the system to a few relevant variables corresponding to a simplified Poisson equation. This is quite connected to the issue of coarse-graining or identifying an appropriate molecular backbone. Another route, which does not require a priori physical knowledge, would be to use greddy methods [120, 156, 31, 57]. Additionally if the system possesses a specific symmetry or structure, one can make profit of dedicated tensor formats [73] as done for the Schrödinger equation in [172].

Acknowledgements

The idea of using control variates came out of discussions with Antonietta Mira (USI) while Gabriel Stoltz was participating to the workshop "Free-energy calculations: a mathematical perspective" at Oaxaca. We thank Antoine Levitt (ENPC), Greg Pavliotis (Imperial College), Stefano Lepri (ISC) and Jonathan Goodman (NYU) for helpful discussions. This work is supported by the Agence Nationale de la Recherche under grant ANR-14-CE23-0012 (COSMOS); as well as the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) – ERC Grant Agreement number 614492. We also benefited from the scientific environment of the Laboratoire International Associé

Figure 3.10: Asymptotic variance of the length of the dimer, with or without control variate. Left: Unsolvated dimer. Right: Solvent with the soft potential (3.48). Bottom: Solvent with the Coulomb-like potential (3.49).



between the Centre National de la Recherche Scientifique and the University of Illinois at Urbana-Champaign. Part of this work was done during the authors' stay at the Institut Henri Poincaré - Centre Emile Borel during the trimester "Stochastic Dynamics Out of Equilibrium" (April-July 2017). The authors warmly thank this institution for its hospitality.

3.6 Proofs of Theorems 3.1 and 3.2

Let us first prove Theorem 3.1, and deduce Theorem 3.2 in a second step. We suppose in all this section that Assumptions 1 to 5 hold true. The norm and scalar product indexed by η correspond to the canonical ones on $\mathrm{L}^2(\pi_\eta)$. We start by giving a useful technical result.

Lemma 3.1. *For any $\eta_* > 0$ and $n \in \mathbb{N}$, there exists $C_{n,\eta_*} \in \mathbb{R}_+$ such that, for any $|\eta| \leqslant \eta_*$,*

$$\forall \varphi \in \mathrm{L}^\infty_n, \quad |\mathbb{E}_\eta[\varphi] - \mathbb{E}_0[\varphi]| \leqslant C_{n,\eta_*}\, \eta \|\varphi\|_{\mathrm{L}^\infty_n}.$$

Proof. For any $\psi \in \mathcal{S}$,

$$\mathbb{E}_0[\mathcal{L}_\eta \psi] = \eta \mathbb{E}_0[\widetilde{\mathcal{L}}\psi],$$

so that, for a given $\varphi \in \mathcal{S}$, the previous equality applied to $\psi = \mathcal{L}_\eta^{-1}\Pi_\eta \varphi$ leads to

$$\mathbb{E}_0[\Pi_\eta \varphi] = \eta \mathbb{E}_0[\widetilde{\mathcal{L}}\mathcal{L}_\eta^{-1}\Pi_\eta \varphi] = \eta \left\langle \widetilde{\mathcal{L}}\mathcal{L}_\eta^{-1}\Pi_\eta \varphi, \mathbf{1} \right\rangle_0 = \eta \left\langle \mathcal{L}_\eta^{-1}\Pi_\eta \varphi, \widetilde{\mathcal{L}}^* \mathbf{1} \right\rangle_0.$$

Since $\mathbb{E}_0[\Pi_\eta \varphi] = \mathbb{E}_0[\varphi] - \mathbb{E}_\eta[\varphi]$ and $|\varphi| \leqslant \|\varphi\|_{\mathrm{L}^\infty_n} \mathcal{K}_n$, we obtain

$$|\mathbb{E}_\eta[\varphi] - \mathbb{E}_0[\varphi]| \leqslant \eta \left\| \mathcal{L}_\eta^{-1} \right\|_{\mathcal{B}(\Pi_\eta \mathrm{L}^\infty_n)} \|\varphi\|_{\mathrm{L}^\infty_n} \left\| \mathcal{K}_n \right\|_0 \left\| \widetilde{\mathcal{L}}^* \mathbf{1} \right\|_0 \leqslant C_{\eta_*,n} \eta \|\varphi\|_{\mathrm{L}^\infty_n},$$

since $\widetilde{\mathcal{L}}^* \mathbf{1} \in \mathrm{L}^2(\pi_0)$ by Assumption 3.4 and $\left\| \mathcal{K}_n \right\|_0 < +\infty$ by Assumption 3.2. The proof is concluded by the density of $\mathcal{S}$ in L^∞_n. □

Corollary 3.1. *For any $\eta_* > 0$ and $n, n' \in \mathbb{N}$, there exists $C_{n,n',\eta_*} \in \mathbb{R}_+$ such that, for any $|\eta| \leqslant \eta_*$,*

$$\forall \varphi \in \mathrm{L}^\infty_n,\ \forall \psi \in \mathrm{L}^\infty_{n'}, \quad |\langle \varphi, \psi \rangle_\eta - \langle \varphi, \psi \rangle_0| \leqslant C_{n,n',\eta_*} \eta \|\varphi\|_{\mathrm{L}^\infty_n} \|\psi\|_{\mathrm{L}^\infty_{n'}}; \tag{3.53}$$

and, for a given $\psi \in \mathcal{S}$, there exists $C_{\psi,n,\eta_*} \mathbb{R}_+$ such that

$$\forall \varphi \in \mathrm{L}^\infty_n, \quad |\langle \varphi, \mathcal{L}_\eta \psi \rangle_\eta - \langle \varphi, \mathcal{L}_0 \psi \rangle_0| \leqslant C_{\psi,n,\eta_*} \eta \|\varphi\|_{\mathrm{L}^\infty_n}. \tag{3.54}$$

Proof. In view of Assumption 3.2 there exist $m \in \mathbb{N}$ depending only on n and n' such that $\|\mathcal{K}_n \mathcal{K}_{n'}\|_{\mathrm{L}^\infty_m} < +\infty$. Therefore, writing

$$\varphi\psi = \frac{\varphi}{\mathcal{K}_n} \frac{\psi}{\mathcal{K}_{n'}} \mathcal{K}_n \mathcal{K}_{n'}.$$

we obtain

$$\|\varphi\psi\|_{\mathrm{L}^\infty_m} \leqslant \|\varphi\|_{\mathrm{L}^\infty_n} \|\psi\|_{\mathrm{L}^\infty_{n'}} \|\mathcal{K}_n \mathcal{K}_{n'}\|_{\mathrm{L}^\infty_m}.$$

The estimate (3.53) then follows from Lemma 3.1 since $\langle \varphi, \psi \rangle_\eta = \mathbb{E}_\eta[\varphi\psi]$. Fix now $\psi \in \mathcal{S}$. There exist $n', n'' \in \mathbb{N}$ such that $\mathcal{L}_0\psi \in \mathrm{L}^\infty_{n'}$ and $\widetilde{\mathcal{L}}\psi \in \mathrm{L}^\infty_{n''}$. Therefore, using Lemma 3.1 twice,

$$\begin{aligned}
|\langle \varphi, \mathcal{L}_\eta \psi \rangle_\eta - \langle \varphi, \mathcal{L}_0 \psi \rangle_0| &\leqslant \left| \langle \varphi, \mathcal{L}_0 \psi \rangle_\eta - \langle \varphi, \mathcal{L}_0 \psi \rangle_0 \right| + |\eta| \left| \left\langle \varphi, \widetilde{\mathcal{L}} \psi \right\rangle_\eta \right| \\
&\leqslant C_{n,n',\eta_*} \eta \|\varphi\|_{\mathrm{L}^\infty_n} \|\mathcal{L}_0 \psi\|_{\mathrm{L}^\infty_{n'}} + |\eta| \left| \left\langle \varphi, \widetilde{\mathcal{L}} \psi \right\rangle_0 \right| + \eta^2 C_{n,n'',\eta_*} \|\varphi\|_{\mathrm{L}^\infty_n} \|\widetilde{\mathcal{L}}\psi\|_{\mathrm{L}^\infty_{n''}}.
\end{aligned}$$

This implies (3.54) since n' and n'' depend only on ψ. □

We can now provide the proof of Theorem 3.1.

Proof of Theorem 3.1. When $\Pi_0 A$ is not bounded one needs to define an approximation of $-\mathcal{L}_\eta^{-1}\Pi_\eta\phi_\eta$ at order K in η, as done in [135] for instance:

$$Q^K := -\Pi_\eta \mathcal{L}_0^{-1} \Pi_0 \sum_{k=1}^K \eta^k A^k R \in \mathcal{S}.$$

Let us show that this is indeed a good approximation. Using successively (3.14) and (3.16) the corresponding truncation error reads:

$$\begin{aligned}
\Pi_\eta \phi_\eta + \mathcal{L}_\eta Q^K &= \Pi_\eta \phi_\eta - \mathcal{L}_\eta \mathcal{L}_0^{-1} \Pi_0 \sum_{k=1}^K \eta^k A^k R \\
&= \eta \Pi_\eta A R - \Pi_\eta (1 - \eta A) \sum_{k=1}^K \eta^k A^k R \\
&= \eta^{K+1} \Pi_\eta A^{K+1} R,
\end{aligned}$$

which implies:

$$Q^K + \mathcal{L}_\eta^{-1} \Pi_\eta \phi_\eta = \eta^{K+1} \mathcal{L}_\eta^{-1} \Pi_\eta A^{K+1} R. \tag{3.55}$$

Let us first show that the corresponding approximated asymptotic variance $\sigma^2_{\phi_{\eta,K}} := 2 \left\langle \Pi_\eta \phi_\eta, Q^K \right\rangle_\eta$ is close to $\sigma^2_{\phi_\eta,\eta}$ (defined in (3.9)). Indeed,

$$\sigma^2_{\phi_\eta,\eta} - \sigma^2_{\phi_{\eta,K}} = 2 \left\langle \Pi_\eta \phi_\eta, -\mathcal{L}_\eta^{-1} \Pi_\eta \phi_\eta - Q^K \right\rangle_\eta = 2\eta^{K+1} \left\langle \Pi_\eta \phi_\eta, -\mathcal{L}_\eta^{-1} \Pi_\eta A^{K+1} R \right\rangle_\eta.$$

Note that $\Pi_\eta A^{K+1} R \in \mathcal{S}$ because $\mathcal{S}$ is stable by $\mathcal{L}_0^{-1}\Pi_0$ and $\widetilde{\mathcal{L}}$ in view of Assumptions 3.4 and 3.5. Since $\Pi_\eta \phi_\eta \in \mathcal{S}$ as well, there exist $n \in \mathbb{N}$ (depending on R and K) and $m \in \mathbb{N}$ (depending on R) such that $\Pi_\eta A^{K+1} \Pi_0 R \in \mathrm{L}^\infty_n$ and $\Pi_\eta \phi_\eta \in \mathrm{L}^\infty_m$. Note that m does not depend on η in view of the expression (3.15) of ϕ_η. Using Assumption 3.4 we obtain, for

any $\eta_* > 0$ and $|\eta| \leqslant \eta_*$,

$$\begin{aligned} |\sigma^2_{\phi_\eta,\eta} - \sigma^2_{\phi_{\eta,K}}| &\leqslant 2|\eta|^{K+1} \|\Pi_\eta \phi_\eta\|_{\mathrm{L}^\infty_m} \left\| \mathcal{L}_\eta^{-1} \Pi_\eta A^{K+1} R \right\|_{\mathrm{L}^\infty_n} \langle \mathcal{K}_m, \mathcal{K}_n \rangle_\eta \\ &\leqslant 2|\eta|^{K+2} \|\Pi_\eta A R\|_{\mathrm{L}^\infty_m} \left\| \mathcal{L}_\eta^{-1} \right\|_{\mathcal{B}(\Pi_\eta \mathrm{L}^\infty_n)} \left\| A^{K+1} R \right\|_{\mathrm{L}^\infty_n} \|\mathcal{K}_m\|_\eta \, \|\mathcal{K}_n\|_\eta, \end{aligned} \tag{3.56}$$

where the four terms on the right hand side are uniformly bounded for $|\eta| \leqslant \eta_*$ in view of Assumption 3.3 and Lemma 3.1. This shows that there exists $C_{R,\eta_*,K} \in \mathbb{R}_+$ such that, for any $|\eta| \leqslant \eta_*$,

$$\sigma^2_{\phi_\eta,\eta} - \sigma^2_{\phi_{\eta,K}} = \eta^{K+2} E_{R,\eta,K}, \tag{3.57}$$

where $|E_{R,\eta,K}| \leqslant C_{R,\eta_*,K}$.

At this stage it is sufficient to prove the expansion (3.13) for $\sigma^2_{\phi_{\eta,K}}$. The approximate variance $\sigma^2_{\phi_{\eta,K}}$ can be expanded in powers of η as follows:

$$\sigma^2_{\phi_{\eta,K}} = 2 \left\langle \Pi_\eta \phi_\eta, Q^K \right\rangle_\eta = 2 \left\langle \eta \Pi_\eta A R, Q^K \right\rangle_\eta = -2\eta \sum_{k=1}^{K} \eta^k \left\langle \Pi_\eta A R, \mathcal{L}_0^{-1} \Pi_0 A^k R \right\rangle_\eta .$$

In fact it suffices to consider $K = 1$. We use Lemma 3.1 and Corollary 3.1 to replace integrals with respect to π_η by integrals with respect to π_0: there exists $C_{R,\eta_*} \in \mathbb{R}_+$ such that, for any $|\eta| \leqslant \eta_*$,

$$\sigma^2_{\phi_{\eta,1}} = -2\eta^2 \left\langle \Pi_\eta A R, \mathcal{L}_0^{-1} \Pi_0 A R \right\rangle_\eta = -2\eta^2 \left\langle A R, \mathcal{L}_0^{-1} \Pi_0 A R \right\rangle_0 + \eta^3 \widetilde{E}_{R,\eta},$$

with $|\widetilde{E}_{R,\eta}| \leqslant C_{R,\eta_*}$. The claimed result then follows by (3.57). □

The following lemma is useful for the proof of Theorem 3.2 and is also used in Section 3.4. Denote by $\mathcal{L}_\eta^{\mathrm{S}}$ the symmetric part of $\mathcal{L}_\eta$ on $\mathrm{L}^2(\pi_\eta)$, defined as

$$\forall \varphi, \psi \in \mathcal{S}, \qquad \left\langle \mathcal{L}_\eta^{\mathrm{S}} \varphi, \psi \right\rangle_\eta = \frac{1}{2} \Big(\langle \mathcal{L}_\eta \varphi, \psi \rangle_\eta + \langle \varphi, \mathcal{L}_\eta \psi \rangle_\eta \Big).$$

Note that the action of this operator is not explicit when π_η is not known.

Lemma 3.2. *For any $\varphi, U \in \mathcal{S}$,*

$$\sigma^2_{\varphi + \mathcal{L}_\eta U,\eta} = \sigma^2_{\varphi,\eta} + \left\langle -\mathcal{L}_\eta^{\mathrm{S}} U, 2\mathcal{L}_\eta^{-1} \Pi_\eta \varphi + \Pi_\eta U \right\rangle_\eta .$$

Proof. By definition of the asymptotic variance,

$$\begin{aligned} \sigma^2_{\varphi+\mathcal{L}_\eta U,\eta} &= \left\langle \varphi + \mathcal{L}_\eta U, -\mathcal{L}_\eta^{-1} \Pi_\eta \left(\varphi + \mathcal{L}_\eta U \right) \right\rangle_\eta \\ &= \sigma^2_{\varphi,\eta} - \langle \varphi, \Pi_\eta U \rangle_\eta - \left\langle \mathcal{L}_\eta U, \mathcal{L}_\eta^{-1} \Pi_\eta \varphi \right\rangle_\eta - \langle \mathcal{L}_\eta U, \Pi_\eta U \rangle_\eta \\ &= \sigma^2_{\varphi,\eta} - \left\langle \Pi_\eta U, \mathcal{L}_\eta \mathcal{L}_\eta^{-1} \Pi_\eta \varphi \right\rangle_\eta + \left\langle \mathcal{L}_\eta U, -\mathcal{L}_\eta^{-1} \Pi_\eta \varphi \right\rangle_\eta + \left\langle -\mathcal{L}_\eta^{\mathrm{S}} U, \Pi_\eta U \right\rangle_\eta \\ &= \sigma^2_{\varphi,\eta} + \left\langle -\mathcal{L}_\eta^{\mathrm{S}} U, 2\mathcal{L}_\eta^{-1} \Pi_\eta \varphi + \Pi_\eta U \right\rangle_\eta , \end{aligned}$$

which is the desired result. □

We now deduce Theorem 3.2 from Theorem 3.1 using Lemma 3.2.

Proof of Theorem 3.2. We use Lemma 3.2 with $U = \varepsilon f$ to compute the asymptotic variance of

$$\phi_{\eta,\varepsilon} = \phi_\eta + \varepsilon \mathcal{L}_\eta f,$$

with ϕ_η given by (3.15). Noting that (from (3.55) with $K = 1$)

$$\mathcal{L}_\eta^{-1} \Pi_\eta \phi_\eta = \eta^2 \mathcal{L}_\eta^{-1} \Pi_\eta A^2 R + \eta \Pi_\eta \mathcal{L}_0^{-1} \Pi_0 A R,$$

it comes

$$\begin{aligned}
\sigma^2_{\phi_{\eta,\varepsilon},\eta} &= \sigma^2_{\phi_\eta,\eta} + \varepsilon \left\langle -\mathcal{L}_\eta^{\mathrm{S}} f, 2\mathcal{L}_\eta^{-1} \Pi_\eta \phi_\eta + \varepsilon \Pi_\eta f \right\rangle_\eta \\
&= \sigma^2_{\phi_\eta,\eta} + 2\varepsilon\eta \left\langle -\mathcal{L}_\eta^{\mathrm{S}} f, \Pi_\eta \mathcal{L}_0^{-1} \Pi_0 A R \right\rangle_\eta + \varepsilon^2 \left\langle -\mathcal{L}_\eta^{\mathrm{S}} f, \Pi_\eta f \right\rangle_\eta \\
&\quad + 2\varepsilon\eta^2 \left\langle -\mathcal{L}_\eta^{\mathrm{S}} f, \mathcal{L}_\eta^{-1} \Pi_\eta A^2 R \right\rangle_\eta \\
&= \sigma^2_{\phi_\eta,\eta} + \varepsilon\eta \left\langle \mathcal{L}_\eta f, -\mathcal{L}_0^{-1} \Pi_0 A R \right\rangle_\eta - \varepsilon\eta \left\langle f, \mathcal{L}_\eta \mathcal{L}_0^{-1} \Pi_0 A R \right\rangle_\eta \\
&\quad + \varepsilon^2 \left\langle -\mathcal{L}_\eta f, f \right\rangle_\eta + \varepsilon\eta^2 \left\langle \mathcal{L}_\eta f, -\mathcal{L}_\eta^{-1} \Pi_\eta A^2 R \right\rangle_\eta - \varepsilon\eta^2 \left\langle f, \Pi_\eta A^2 R \right\rangle_\eta .
\end{aligned} \tag{3.58}$$

In order to retain only the leading order terms in the expansion in η and ε we first bound the two last terms in the last equation of (3.58) in a fashion similar to (3.56). Then we change the scalar products in $\mathrm{L}^2(\pi_\eta)$ by their equivalents in $\mathrm{L}^2(\pi_0)$ and replace $\mathcal{L}_\eta$ by $\mathcal{L}_0$ (controlling the error with Corollary 3.1). All higher order terms are gathered in the remainder, using the inequalities $|\varepsilon|\eta^2 \leqslant |\varepsilon|^3 + |\eta|^3$ and $\varepsilon^2|\eta| \leqslant |\varepsilon|^3 + |\eta|^3$. Finally, there exist $\varepsilon_* > 0$ and $C_{R,\eta_*,\varepsilon_*,f} \in \mathbb{R}_+$ such that, for any $|\eta| \leqslant \eta_*$ and any $|\varepsilon| \leqslant \varepsilon_*$,

$$\sigma^2_{\phi_{\eta,\varepsilon},\eta} = \sigma^2_{\phi_\eta,\eta} + \varepsilon\eta \left\langle (\mathcal{L}_0 + \mathcal{L}_0^*) f, -\mathcal{L}_0^{-1} \Pi_0 A R \right\rangle_0 + \varepsilon^2 \left\langle -\mathcal{L}_0 f, \Pi_0 f \right\rangle_0 + (\varepsilon^3 + \eta^3) E_{R,\eta,\varepsilon,f},$$

where $|E_{R,\eta,\varepsilon,f}| \leqslant C_{R,\eta_*,\varepsilon_*,f}$. Formula (3.21) then follows in view of Theorem 3.1. □

3.7 Technical results used in Section 3.4

3.7.1 Equivalence of modified flux observables

There exist infinitely many observables whose average is the average heat flux in the chain. In particular (see (3.31)) any linear combination of the elementary fluxes with weights summing to 1 (*i.e.* of the form (3.32)) has the same average. The procedure described in Section 3.2 allows to construct a modified observable ϕ starting from any observable R. A legitimate question is which choice of R provides the modified observable with the smallest asymptotic variance. We show here that, starting from any linear combination of the form (3.32), the resulting modified observable has the same asymptotic variance in the

equilibrium setting. Note that the linear combination can involve the fluxes at the ends of the chain j_0 and j_N.

Consider two fluxes R^1 and $R^2 = R^1 + \mathcal{L}U$, where R^1, R^2 and U are linear combinations of the $(j_n)_{0\leqslant n\leqslant N}$ and the $(\varepsilon_n)_{1\leqslant n\leqslant N}$, respectively. The function U can indeed be assumed to be a combination of the energies $(\varepsilon_n)_{1\leqslant n\leqslant N}$ since $j_{n+1} = j_n - \mathcal{L}\varepsilon_n$ in view of (3.30). The functions R^1, R^2 and U have their counterparts in the simplified (harmonic) setting: $R_0^2 = R_0^1 + \mathcal{L}_0 U_0$. The two associated simplified Poisson equations read

$$\begin{cases} -\mathcal{L}_0\Phi_0^1 = R_0^1 - \mathbb{E}_0[R_0^1], \\ -\mathcal{L}_0\Phi_0^2 = R_0^2 - \mathbb{E}_0[R_0^2]. \end{cases}$$

The right hand side of these two equations is modified as well since the definition of the fluxes j_n depends on the potential v. The average $\mathbb{E}_0[R_0^1] = \mathbb{E}_0[R_0^2]$ is the heat flux for the harmonic chain. The solutions of these Poisson equations satisfy $\Phi_0^2 = \Phi_0^1 - U_0$ (up to elements of the kernel of $\mathcal{L}_0$, which are constants [34]), so the two corresponding modified observables are such that

$$\phi_2 = R^2 + \mathcal{L}\Phi_0^2 = R^1 + \mathcal{L}U + \mathcal{L}(\Phi_0^1 - U_0) = \phi_1 + \mathcal{L}(U - U_0).$$

Assume now that the chain is at equilibrium ($T_{\mathrm{L}} = T_{\mathrm{R}}$). In view of Proposition 3.1, the two modified observables thus have the same asymptotic variance (*i.e.* $\sigma_{\phi_1}^2 = \sigma_{\phi_2}^2$) as soon as $U - U_0$ does not depend on p_1 nor on p_N. This is indeed the case for the elementary energies $\varepsilon_n - \varepsilon_{n,0} = \frac{1}{2}(w(r_{n-1}) + w(r_n))$ for $0 \leqslant n \leqslant N$, where $\varepsilon_{n,0}$ is defined by (3.30) with v replaced by v_0. This is in particular true at the ends of the chain ($n = 0$ and $n = N$), so the boundary flux R defined in (3.33) and the standard (bulk) flux $\widetilde{R}$ provide two modified observables with the same asymptotic variance.

When the temperature difference $T_{\mathrm{L}} - T_{\mathrm{R}}$ is not too large, the asymptotic variances of the two modified observables are approximately equal. This shows that the choice of the linear combination of the form (3.32), from which the modified observable ϕ is constructed, does not significantly change the asymptotic variance in this regime.

3.7.2 Computation of the asymptotic variances of j_0 and j_N

Since we are in the setting of Remark 3.5, we assume in this section that the system is at equilibrium ($T_{\mathrm{L}} = T_{\mathrm{R}} = \beta^{-1}$). Recall that $\mathcal{L}\varepsilon_1 = j_0 - j_1$, and more precisely $\mathcal{L}_{\mathrm{FD}}\varepsilon_1 = j_0$ where $\mathcal{L}_{\mathrm{FD}}$ is the symmetric part of the generator at equilibrium, which is known explicitly

(see (3.36)). Therefore, using Lemma 3.2 with $\varphi = j_0$ and $U = -\varepsilon_1$ (so that $\varphi + \mathcal{L}U = j_1$),

$$\begin{aligned}\sigma_{j_1}^2 &= \sigma_{j_0}^2 + \left\langle \mathcal{L}_{\mathrm{FD}}\varepsilon_1, 2\mathcal{L}^{-1}j_0 - \varepsilon_1 \right\rangle_{\mathrm{eq}} \\ &= \sigma_{j_0}^2 + 2\left\langle j_0, \mathcal{L}^{-1}j_0 \right\rangle_{\mathrm{eq}} + \left\langle \gamma\beta^{-1}\partial_{p_1}^*\partial_{p_1}\varepsilon_1, \varepsilon_1 \right\rangle_{\mathrm{eq}} \\ &= -\sigma_{j_0}^2 + \gamma\beta^{-1}\left\| \partial_{p_1}\left(\frac{p_1^2}{2m}\right)\right\|_{\mathrm{eq}}^2 \\ &= -\sigma_{j_0}^2 + \frac{\gamma}{m}\beta^{-2}.\end{aligned}$$

Therefore,

$$\sigma_{j_0}^2 = \frac{\gamma}{m}\beta^{-2} - \sigma_{j_1}^2,$$

from which (3.39) follows in view of (3.38). Similar computations give the result for j_N.

3.7.3 Euler-Lagrange equation for (3.45)

Denoting by $\hat{\Omega} = m\hat{\omega}^2$, the minimization problem (3.44) can be recast as minimizing the following function for $(\hat{r}, \hat{\Omega}) \in \mathbb{R} \times (0, +\infty)$:

$$\begin{aligned}f(\hat{r}, \hat{\Omega}) &= \int_{\mathbb{R}} \left[v'(r_1) - \hat{\Omega}(r_1 - \hat{r})\right]^2 \mathrm{e}^{-\beta v(r_1)}\,\mathrm{d}r_1 \\ &= \int_{\mathbb{R}} \left[v'(r_1)^2 - 2\hat{\Omega}v'(r_1)(r_1 - \hat{r}) + \hat{\Omega}^2(r_1 - \hat{r})^2\right] \mathrm{e}^{-\beta v(r_1)}\,\mathrm{d}r_1 \\ &= \int_{\mathbb{R}} \left[v'(r_1)^2 - 2\beta^{-1}\hat{\Omega} + \hat{\Omega}^2(r_1 - \hat{r})^2\right] \mathrm{e}^{-\beta v(r_1)}\,\mathrm{d}r_1 \\ &= C - 2\beta^{-1}\hat{\Omega}\mathcal{M}_0 + \hat{\Omega}^2(\mathcal{M}_2 - 2\mathcal{M}_1\hat{r} + \mathcal{M}_0\hat{r}^2),\end{aligned}$$

with $C = \int_{\mathbb{R}} v'(r_1)^2 \mathrm{e}^{-\beta v(r_1)}\,\mathrm{d}r_1$ and where the third line is obtained with an integration by parts. The gradient of f vanishes if and only if:

$$\begin{cases} 0 = \hat{\Omega}^2(-2\mathcal{M}_1 + 2\mathcal{M}_0\hat{r}), \\ 0 = -2\beta^{-1}\mathcal{M}_0 + 2\hat{\Omega}(\mathcal{M}_2 - 2\mathcal{M}_1\hat{r} + \mathcal{M}_0\hat{r}^2). \end{cases}$$

The only solution of this system is indeed given by (3.45).

3.7.4 Harmonic chain

We establish in this section the formulas (3.42) using the linear structure of the harmonic chain, see [105, Appendix B] for similar computations. The interaction potential writes

$v_0(r) = \frac{1}{2}m\omega^2(r-\hat{r})^2$, so (3.26) reduces to

$$\begin{cases} \mathrm{d}r_n = \dfrac{1}{m}(p_{n+1}-p_n)\,\mathrm{d}t, \\ \mathrm{d}p_1 = m\omega^2(r_1-\hat{r})\,\mathrm{d}t - \dfrac{\gamma}{m}p_1\,\mathrm{d}t + \sqrt{2\gamma T_{\mathrm{L}}}\mathrm{d}W_t^L, \\ \mathrm{d}p_n = m\omega^2(r_n-r_{n-1})\,\mathrm{d}t, \\ \mathrm{d}p_N = -m\omega^2(r_{N-1}-\hat{r})\,\mathrm{d}t - \dfrac{\gamma}{m}p_N\,\mathrm{d}t + \sqrt{2\gamma T_{\mathrm{R}}}\mathrm{d}W_t^R. \end{cases} \tag{3.59}$$

In order to simplify the algebra we make the change of variables

$$x = (p_1, m\omega(r_1-\hat{r}), p_2, \cdots, p_{N-1}, m\omega(r_{N-1}-\hat{r}), p_N) \in \mathbb{R}^{2N-1},$$

and denote by $\nu = \frac{m\omega}{\gamma} > 0$ the dimensionless ratio between the respective time scales of the harmonic potential and of the fluctuation-dissipation process. The process (3.59) is in fact a generalized Ornstein-Uhlenbeck process:

$$\mathrm{d}x = \frac{\gamma}{m}\mathbf{A}x\,\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\left(\mathbf{S} + \frac{1}{2}\beta(T_{\mathrm{L}}-T_{\mathrm{R}})\,\mathbf{R}\right)^{1/2}\mathrm{d}W_t, \tag{3.60}$$

where $\beta^{-1} = (T_{\mathrm{L}}+T_{\mathrm{R}})/2$ and

$$\mathbf{A} = \nu\left(\mathbf{J} - \mathbf{J}^\top\right) - \mathbf{S} \in \mathbb{R}^{2N-1\times 2N-1},$$

with

$$\mathbf{J} = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{pmatrix},\ \mathbf{S} = \begin{pmatrix} 1 & & (0) \\ & (0) & \\ (0) & & 1 \end{pmatrix},\ \mathbf{R} = \begin{pmatrix} 1 & & (0) \\ & (0) & \\ (0) & & -1 \end{pmatrix}.$$

The generator of this process writes, for any smooth function φ:

$$\mathcal{L}_0\varphi(x) = \frac{\gamma}{m}x^\top\mathbf{A}^\top\nabla\varphi(x) + \gamma\beta^{-1}\left(\mathbf{S} + \frac{1}{2}\beta(T_{\mathrm{L}}-T_{\mathrm{R}})\,\mathbf{R}\right) : \nabla^2\varphi(x).$$

Recall that the observable we consider is the heat flux $R = \frac{1}{2}(j_0 + j_N)$ at the ends of the chain, with j_0 and j_N given by (3.29). This corresponds to the following quadratic form:

$$R(x) = -\frac{\gamma}{2m^2}x^\top\mathbf{R}x + \frac{\gamma(T_{\mathrm{L}}-T_{\mathrm{R}})}{2m}.$$

We look for the solution Φ_0 to the Poisson equation

$$-\mathcal{L}_0\Phi_0 = R - \mathbb{E}_0[R]. \tag{3.61}$$

The observable R is the sum of a quadratic part and a constant. Since $\mathcal{L}_0$ stabilizes the space of functions $x \mapsto a + x^\top \mathbf{M} x$ with $a \in \mathbb{R}$ and $\mathbf{M}$ a symmetric matrix, we consider the ansatz

$$\Phi_0(x) = \frac{1}{2m} x^\top \mathbf{K} x + C,$$

where $\mathbf{K} \in \mathbb{R}^{(2N-1)\times(2N-1)}$ is symmetric and $C \in \mathbb{R}$ is chosen such that $\mathbb{E}_0[\Phi_0] = 0$. The Poisson equation (3.61) then writes: for all $x \in \mathbb{R}^{2N-1}$,

$$-\frac{\gamma}{m^2} x^\top \mathbf{A}^\top \mathbf{K} x - \gamma\beta^{-1}(\mathbf{S} + \beta(T_\mathrm{L} - T_\mathrm{R})\ \mathbf{R}) : \frac{1}{m}\mathbf{K} = -\frac{\gamma}{2m^2} x^\top \mathbf{R} x + \frac{\gamma(T_\mathrm{L} - T_\mathrm{R})}{2m} - \mathbb{E}_0[R],$$

which is equivalent to

$$\begin{cases} \mathbf{A}^\top \mathbf{K} + \mathbf{K}\mathbf{A} = \mathbf{R}, \\ \qquad \mathbb{E}_0[R] = \dfrac{\gamma(T_\mathrm{L} - T_\mathrm{R})}{2m} + \dfrac{\gamma\beta^{-1}}{m}\left(\mathbf{S} + \beta\dfrac{T_\mathrm{L} - T_\mathrm{R}}{2}\ \mathbf{R}\right) : \mathbf{K}, \end{cases} \tag{3.62}$$

by separating the constant and the quadratic term. The solution is in fact fully explicit since there is an analytical formula for $\mathbf{K}$.

Proposition 3.2. *The solution to* (3.62) *is the following symmetric matrix*

$$\mathbf{K} = -\frac{1}{2(1+\nu^2)}\Big[\nu(\mathbf{J} + \mathbf{J}^\top) + \mathbf{R}\Big]. \tag{3.63}$$

In particular,

$$\mathbb{E}_0[R] = \frac{\nu^2}{1+\nu^2}\frac{\gamma(T_\mathrm{L} - T_\mathrm{R})}{2m}.$$

Proof. Denoting by $\mathbf{M} = \mathbf{J} - \mathbf{J}^\top$ and $\mathbf{N} = \mathbf{J} + \mathbf{J}^\top$, the following relations hold true

$$\mathbf{M}\mathbf{N} - \mathbf{N}\mathbf{M} = 2\mathbf{R}, \quad \mathbf{M}\mathbf{R} = -\mathbf{N}\mathbf{S}, \quad \mathbf{R}\mathbf{M} = \mathbf{S}\mathbf{N}, \quad \mathbf{R}\mathbf{S} = \mathbf{S}\mathbf{R} = \mathbf{R},$$
$$\mathbf{R} : \mathbf{R} = 2, \quad \mathbf{R} : \mathbf{S} = 0, \quad \mathbf{S} : \mathbf{N} = 0, \quad \mathbf{R} : \mathbf{N} = 0.$$

This allows to develop $\mathbf{A}\mathbf{K} + \mathbf{K}\mathbf{A}^\top$ with $\mathbf{K}$ defined in (3.63) and obtain $\mathbf{R}$. By injecting the expression of $\mathbf{K}$ into (3.62) we obtain the expression of $\mathbb{E}_0[R]$. □

Remark 3.6. *There exists in fact a unique solution to the Lyapunov equation* (3.62) *for any right hand side, since* $\mathbf{A}$ *is Hurwitz [13]. This latter assertion is equivalent to the exponential decay of the semigroup* $\mathrm{e}^{t\mathcal{L}}$*, proved in [34] for example for more general interaction potentials. To prove that* $\mathbf{A}$ *is Hurwitz, take a non-zero eigenvector* x *associated to an eigenvalue* $\lambda \in \mathbb{C}$*. Suppose that* $\mathscr{R}(\lambda) \geqslant 0$*. Then,*

$$-|x_1|^2 - |x_{2N-1}|^2 = \bar{x}^\top \mathbf{A} x = \mathscr{R}(\lambda)|x|^2 \geqslant 0,$$

so $\mathscr{R}(\lambda) = 0$ *and* $x_1 = x_{2N-1} = 0$*. Using* $\mathbf{A}x = \lambda x$ *we iteratively obtain* $x_2 = 0$*, then* $x_3 = 0$*, and so on until* $x = 0$*. The contradiction proves that any eigenvalue of* $\mathbf{A}$ *has a negative real part.*

The optimal harmonic control variate Φ_0 is thus

$$\begin{aligned}\Phi_0(x) &= -\frac{1}{2m(1+\nu^2)}\left[\nu\sum_{k=1}^{2N-2}x_kx_{k+1}+\frac{1}{2}x_1^2-\frac{1}{2}x_{2N-1}^2\right]+C\\ &= \frac{m}{2\gamma(1+\nu^2)}\left[-\omega^2\sum_{n=1}^{N-1}(r_n-\hat{r})(p_n+p_{n+1})+\frac{\gamma}{2m^2}\left(p_N^2-p_1^2\right)\right]+C\\ &= \frac{m}{2\gamma(1+\nu^2)}\sum_{n=0}^{N}(j_{n,0}-\mathbb{E}_0[R]),\end{aligned}$$

where $j_{n,0}$ is the n-th elementary flux (3.28) with v replaced by v_0. This function indeed has the dimensions of an energy since it is the product of some characteristic time by a heat flux.

3.7.5 Proof of Assumption 3.4 for the harmonic chain

The space $\mathcal{S}$ is easily seen to be stable by $\mathcal{L}$. We prove next that $\mathcal{L}^{-1}\varphi$ is in $\mathcal{S}$ when $\varphi\in\mathcal{S}$. Note first that it is possible to analytically integrate the dynamics (3.60) as

$$x_t = \mathrm{e}^{\gamma t\mathbf{A}/m}x_0+\sqrt{\frac{2\gamma}{\beta}}\int_0^t\mathrm{e}^{\gamma(t-s)\mathbf{A}/m}\left(\mathbf{S}+\frac{1}{2}\beta(T_\mathrm{L}-T_\mathrm{R})\mathbf{R}\right)^{1/2}\mathrm{d}W_t. \tag{3.64}$$

The matrix $\mathbf{A}$ is Hurwitz so there exist $\lambda>0$ and $C_\mathbf{A}\geqslant 1$ such that the Frobenius norm of the associated semi-group decays exponentially with rate λ:

$$\left\|\mathrm{e}^{\gamma t\mathbf{A}/m}\right\|\leqslant C_\mathbf{A}\mathrm{e}^{-\lambda t}\leqslant C_\mathbf{A}.$$

Take $\varphi\in\mathcal{S}$ with mean zero with respect to π. There exist $\theta_0,\theta_1\in[0,\theta_*/2)$ such that $\varphi\in\mathrm{L}^\infty_{\theta_0}$ and, for any $n\in[1,2N-1]$, $\partial_{x_n}\varphi\in\mathrm{L}^\infty_{\theta_1}$. By the results of [34] (recalled in Section 3.4.1.2) we know already that $\mathcal{L}^{-1}\varphi\in\mathrm{L}^\infty_{\theta_0}$. Denoting by $|\cdot|$ the Euclidean norm in $\mathbb{R}^{2N-1}$, and using (3.64),

$$\begin{aligned}\left|\nabla_{x_0}\left(\mathrm{e}^{t\mathcal{L}}\varphi\right)(x_0)\right| &= \left|\nabla_{x_0}\mathbb{E}_{x_0}[\varphi(x_t)]\right| = \left|\mathbb{E}_{x_0}[\mathrm{e}^{\gamma t\mathbf{A}/m}\nabla\varphi(x_t)]\right|\\ &\leqslant \left\|\mathrm{e}^{\gamma t\mathbf{A}/m}\right\|\left|\mathbb{E}_{x_0}[\nabla\varphi(x_t)]\right| \leqslant C_\mathbf{A}\mathrm{e}^{-\lambda t}\|\nabla\varphi\|_{\mathrm{L}^\infty_{\theta_1}}\mathbb{E}_{x_0}[\mathcal{K}_{\theta_1}(x_t)].\end{aligned} \tag{3.65}$$

By the exponential decay of the semi-group $\mathrm{e}^{t\mathcal{L}}$ on the functional space $\mathrm{L}^\infty_{\theta_1}$ (see [34]), there exist C_{θ_1},λ' such that

$$\left|\mathrm{e}^{t\mathcal{L}}\mathcal{K}_{\theta_1}(x_0)-\mathbb{E}[\mathcal{K}_{\theta_1}]\right|\leqslant C_{\theta_1}\mathrm{e}^{-\lambda' t}\mathcal{K}_{\theta_1}(x_0),$$

so that

$$\mathbb{E}_{x_0}[\mathcal{K}_{\theta_1}(x_t)]\leqslant\mathbb{E}[\mathcal{K}_{\theta_1}]+C_{\theta_1}\mathrm{e}^{-\lambda' t}\mathcal{K}_{\theta_1}(x_0)\leqslant C'_{\theta_1}\mathcal{K}_{\theta_1}(x_0),$$

with $C'_{\theta_1} = \max\left(\mathbb{E}[\mathcal{K}_{\theta_1}], C_{\theta_1}\right)$. Using this result and integrating (3.65) from $t = 0$ to ∞,

$$\left|\nabla_{x_0}\mathcal{L}^{-1}\varphi(x_0)\right| \leqslant \int_0^\infty \left|\nabla_{x_0}\mathrm{e}^{t\mathcal{L}}\varphi(x_0)\right| \mathrm{d}t \leqslant \int_0^\infty C_{\mathbf{A}}\mathrm{e}^{-\lambda t}\|\nabla\varphi\|_{\mathrm{L}^\infty_{\theta_1}} C'_{\theta_1}\mathcal{K}_{\theta_1}(x_0)\,\mathrm{d}t,$$

so that

$$\left\|\nabla_{x_0}\mathcal{L}^{-1}\varphi\right\|_{\mathrm{L}^\infty_{\theta_1}} \leqslant \frac{C_{\mathbf{A}}C'_{\theta_1}}{\lambda}\|\nabla\varphi\|_{\mathrm{L}^\infty_{\theta_1}}.$$

This implies that $\nabla\mathcal{L}^{-1}\varphi \in \mathrm{L}^\infty_{\theta_1}$. Similar formulas hold for higher order derivatives. This allows to show that $\mathcal{L}^{-1}\varphi \in \mathcal{S}$, proving that the core $\Pi_0\mathcal{S}$ is stable by $\mathcal{L}^{-1}$.

3.8 Resolution of the differential equation (3.52)

The Poisson equation (3.52) can be easily solved using finite differences. In order to provide a stable numerical resolution of this equation, let us first determine its boundary conditions. Denoting by $\varphi = \psi'$, (3.52) can be reformulated as

$$\beta^{-1}\varphi'(r) = r_* - r + v'_*(r)\varphi(r). \tag{3.66}$$

Note that it is sufficient to determine φ in order to evaluate $\mathcal{L}\Phi_0$.

Proposition 3.3. *Assume that $v \in \mathcal{C}^1((0,+\infty),\mathbb{R})$ is such that*

$$\limsup_{r\to 0} v'(r) < +\infty \quad \textit{and} \quad \frac{v'(r)}{r} \xrightarrow[r\to+\infty]{} +\infty. \tag{3.67}$$

Then (3.66) admits a unique solution $\varphi \in \mathrm{L}^2(\pi_*)$ whose primitives are in $\mathrm{L}^2(\pi_*)$. Moreover this solution in continuous on $[0,+\infty)$, $\varphi(0) = 0$ and φ converges to 0 at $+\infty$.

The conditions (3.67) are satisfied for the double-well potential (3.47) and for many potentials used in practice. They imply in particular that $v_*(r) \xrightarrow[r\to 0]{} +\infty$ and that π_* vanishes at 0 and $+\infty$.

Proof. Let us introduce the function

$$f(r) = \int_0^r \beta\,(r_* - s)\,\mathrm{e}^{-\beta v_*(s)}\mathrm{d}s.$$

We prove that $\varphi(r) = f(r)\mathrm{e}^{\beta v_*(r)}$ is the only bounded solution to (3.66), and that it vanishes at the boundary of the domain. We first obtain bounds on f to this end. The function f satisfies $f(0) = 0$ and $f'(r) = \beta\,(r_* - r)\,\mathrm{e}^{-\beta v_*(r)}$. Using the short-hand notation $z(r) = \int_0^r \mathrm{e}^{-\beta v_*}$ (with limiting value z_∞ as $r \to +\infty$) and $e(r) = \int_0^r s\mathrm{e}^{-\beta v_*(s)}\mathrm{d}s$ (with limiting value e_∞ as $r \to +\infty$), f can be rewritten as

$$\begin{aligned} f(r) &= \beta r_* z(r) - \beta e(r) \\ &= \beta e_\infty\left(\frac{z(r)}{z_\infty} - \frac{e(r)}{e_\infty}\right) = \beta e_\infty\left(\frac{e_\infty - e(r)}{e_\infty} - \frac{z_\infty - z(r)}{z_\infty}\right), \end{aligned} \tag{3.68}$$

since $r_* = e_\infty/z_\infty$, which shows that $f(r) \xrightarrow[r\to\infty]{} 0$. Note that f is increasing on $[0, r_*]$, decreasing on $[r_*, +\infty]$ and vanishes at 0 and infinity. Therefore, $f \geqslant 0$. Let us now bound the behavior of this function near 0 and $+\infty$, in order to prove that φ vanishes at 0 and at $+\infty$. In view of (3.67), there exist $0 < \varepsilon < M < +\infty$ such that $v_*'(r) = v_*'(r) - \frac{d-1}{\beta r}$ is negative on $(0, \varepsilon]$ and positive on $[M, +\infty)$. Define

$$\overline{v_*'}(r) = \sup_{0<s\leqslant r} v_*'(s), \qquad \underline{v_*'}(r) = r \inf_{s\geqslant r} \frac{v_*'(s)}{s}.$$

The functions $\overline{v_*'}$ and $\underline{v_*'}$ are increasing on $(0, +\infty)$, $\overline{v_*'}$ converges to $-\infty$ as $r \to 0$ while $\underline{v_*'}$ converges to $+\infty$ as $r \to +\infty$. Moreover, by definition,

$$\forall 0 < s \leqslant r \leqslant \varepsilon, \quad 1 \leqslant \frac{v_*'(s)}{\overline{v_*'}(r)}, \qquad \text{and} \quad \forall M \leqslant r \leqslant s, \quad 1 \leqslant \frac{v_*'(s)/s}{\underline{v_*'}(r)/r}.$$

Therefore,

$$\begin{aligned} \forall r \leqslant \varepsilon, \quad z(r) &= \int_0^r \mathrm{e}^{-\beta v_*(s)} \mathrm{d}s \leqslant \frac{1}{\beta \overline{v_*'}(r)} \int_0^r \beta v_*'(s) \mathrm{e}^{-\beta v_*(s)} \mathrm{d}s = -\frac{1}{\beta \overline{v_*'}(r)} \mathrm{e}^{-\beta v_*(r)}, \\ \forall r \geqslant M, \quad e_\infty - e(r) &= \int_r^\infty s \mathrm{e}^{-\beta v_*(s)} \mathrm{d}s \leqslant \frac{r}{\beta \underline{v_*'}(r)} \int_r^\infty \beta v_*'(s) \mathrm{e}^{-\beta v_*(s)} \mathrm{d}s = \frac{r}{\beta \underline{v_*'}(r)} \mathrm{e}^{-\beta v_*(r)}. \end{aligned} \tag{3.69}$$

From (3.68) and (3.69) we deduce that the solution $\varphi(r) = f(r)\mathrm{e}^{\beta v_*(r)}$ of (3.66) is non negative on $\mathbb{R}_+^*$ and satisfies

$$\begin{aligned} \forall r \leqslant \varepsilon, \quad 0 \leqslant \varphi(r) &\leqslant \beta r_* z(r) \mathrm{e}^{\beta v_*(r)} \leqslant r_* \frac{1}{\left|\overline{v_*'}(r)\right|}, \\ \forall r \geqslant M, \quad 0 \leqslant \varphi(r) &\leqslant \beta(e_\infty - e(r)) \mathrm{e}^{\beta v_*(r)} \leqslant \frac{r}{\underline{v_*'}(r)}. \end{aligned}$$

This shows that φ vanishes at 0 and $+\infty$. Moreover, any primitive ψ of φ is in $\mathrm{L}^2(\pi_*)$ (because $\psi' = \varphi$ is bounded and π_* integrates functions which increase linearly). The other solutions of (3.66) differ from this one by a factor proportional to $\mathrm{e}^{\beta v_*(r)}$ (which is the solution of the homogeneous equation associated with (3.66)) so that their primitives ψ are not in $\mathrm{L}^2(\pi_*)$. □

Proposition 3.3 shows that the solution φ of (3.66) we are interested in corresponds to the boundary condition $\varphi(0) = 0$. This solution is estimated numerically using a finite difference method. The expectation $r_* = \mathbb{E}_*[r]$ is computed with a one-dimensional numerical quadrature. The so-obtained solution is then interpolated by a function $\widehat{\varphi}$ which is affine on each mesh, so that $\mathcal{L}\widehat{\psi}$ can be evaluated exactly at any point. This ensures that the modified observable is not biased since the control variate indeed belongs to the image of $\mathcal{L}$.

3.9 Asymptotic variance estimator

In the three applications we consider, we provide estimators of the asymptotic variances associated with some function φ together with error bars on this quantity. We make precise in this section this estimator of the variance and the way the error bars on these variance estimates are computed. Under Assumptions 3.1 to 5, the stochastic process admits a unique invariant probability measure π, and the asymptotic variance is well defined for an observable $\varphi = \Pi\varphi + \mathbb{E}[\varphi] \in \mathcal{S}$. The empirical mean of φ is

$$\widehat{\varphi}_t = \frac{1}{t}\int_0^t \varphi(x_t)\,\mathrm{d}t.$$

The associated asymptotic variance (3.4) can be computed using the Green–Kubo formula [97]

$$\begin{aligned}\sigma_\varphi^2 &= 2\int_{\mathcal{X}} \varphi\left(-\mathcal{L}^{-1}\Pi\varphi\right)\,\mathrm{d}\pi = 2\int_0^\infty \mathbb{E}_{x_0}\left[\Pi\varphi(x_s)\Pi\varphi(x_0)\right]\,\mathrm{d}s \\ &= 2\int_0^\infty \left(\mathbb{E}_{x_0}\left[\varphi(x_s)\varphi(x_0)\right] - \mathbb{E}[\varphi]^2\right)\,\mathrm{d}s,\end{aligned}$$

where $\mathbb{E}$ denotes the expectation with respect to initial conditions x_0 distributed according to the invariant probability measure π and for all realizations of the Brownian motion. All these expressions are well defined if we assume that a sufficiently fast decay of the associated semi-group (see [104, Section 3.1.2]). In order to approximate σ_φ^2 we first truncate the time integral as

$$\sigma_\varphi^2 \approx 2\int_0^{t_{\mathrm{deco}}} \mathbb{E}_{x_0}\left[\varphi(x_s)\varphi(x_0)\right]\,\mathrm{d}s - 2t_{\mathrm{deco}}\mathbb{E}[\varphi]^2,$$

where the integrand $\mathbb{E}\left[\varphi(x_s)\varphi(x_0)\right]$ is neglected for $s > t_{\mathrm{deco}}$. The expectations in the integrand are estimated using an empirical average over all the continuous trajectory $(x_t)_{t\in[0,T]}$ (see [3]):

$$\widehat{\sigma_\varphi}^2 = \frac{1}{T}\int_0^T \int_{-t_{\mathrm{deco}}}^{t_{\mathrm{deco}}} \varphi(x_t)\varphi(x_{t+s})\,\mathrm{d}t\mathrm{d}s - 2t_{\mathrm{deco}}\widehat{\varphi}_T^2, \tag{3.70}$$

which is a biased estimator of σ_φ^2:

$$\mathbb{E}\left[\widehat{\sigma_\varphi}^2\right] = 2\int_0^{t_{\mathrm{deco}}} \mathbb{E}_{x_0}\left[\varphi(x_s)\varphi(x_0)\right]\,\mathrm{d}s - 2t_{\mathrm{deco}}\mathbb{E}[\varphi]^2. \tag{3.71}$$

Of course, in practice, the formula for $\widehat{\sigma_\varphi}^2$ is slightly changed not to involve x_t for $t < 0$ or $t > T$. The double integral is approximated using a Riemann sum or a trapezoidal rule for instance. Consider a discretization $(x^n)_{1\leqslant n\leqslant N_{\mathrm{iter}}}$ of the trajectory $(x_t)_{t\in[0,T]}$ with a timestep Δt, of length $T = N_{\mathrm{iter}}\Delta t$. Introducing $N_{\mathrm{deco}} = t_{\mathrm{deco}}/\Delta t$, the discretized version of the estimator (3.70) is

$$\widehat{\widehat{\sigma_\varphi}}^2 = \frac{\Delta t}{N_{\mathrm{iter}}}\sum_{i=1}^{N_{\mathrm{iter}}}\sum_{j=-N_{\mathrm{deco}}}^{N_{\mathrm{deco}}} \varphi(x^i)\varphi(x^{i+j}) - 2t_{\mathrm{deco}}\left(\frac{1}{N_{\mathrm{iter}}}\sum_{i=0}^{N_{\mathrm{iter}}}\varphi(x^n)\right)^2. \tag{3.72}$$

This is the estimator we use throughout this work to provide error bars on average properties. The leading term of the variance of the estimator $\widehat{\widehat{\sigma_\varphi}}^2$ in the regime $\Delta t \ll 1$ and $1 \ll N_{\mathrm{deco}} \ll N_{\mathrm{iter}}$ is

$$\mathrm{Var}\left[\widehat{\widehat{\sigma_\varphi}}^2\right] \approx \frac{2(2N_{\mathrm{deco}}+1)}{N_{\mathrm{iter}}}\sigma_\varphi^4 \approx \frac{4t_{\mathrm{deco}}}{T}\sigma_\varphi^4.$$

Here we made the assumption that Isserlis' theorem [86] holds, as if $(x_t)_t$ was a Gaussian process. It is thus straightforward to provide error bars for the estimator $\widehat{\widehat{\sigma_\varphi}}^2$, and even to choose the simulation time T a priori. Indeed the relative standard statistical error on the variance is very explicit:

$$\frac{\sqrt{\mathrm{Var}\left[\widehat{\widehat{\sigma_\varphi}}^2\right]}}{\sigma_\varphi^2} \approx 2\sqrt{\frac{t_{\mathrm{deco}}}{T}}.$$

For example to estimate the variance with an uncertainty of 1% one should run the simulation for a time $T = 10^4 \times t_{\mathrm{deco}}$. There is a trade-off concerning the choice of t_{deco}: if it is too small the estimators of the integrals are biased in view of (3.71), but if it is too large the variance of the estimator increases. In practice one picks a large value of t_{deco} and uses the cumulated empirical autocorrelation profile to check a posteriori that this value is indeed sufficiently large.

Block averaging. Let us relate the previous estimator of the variance to the common variance estimator $\widetilde{\sigma_\varphi}^2$ considered in the method of block averaging (or batch means); see [131] as well as the references in [103, Section 2.3.1.3]. This method consists in cutting the trajectory into several blocks, computing the empirical average of φ on each block, and estimating the variance of these random variables (considered as independent and identically distributed). If the size of the blocks is $2t_{\mathrm{deco}}$ this estimator has the same variance as $\widehat{\sigma_\varphi}^2$ but the bias is different since

$$\mathbb{E}\left[\widetilde{\sigma_\varphi}^2\right] = 2\int_0^{2t_{\mathrm{deco}}} \left(1 - \frac{s}{2t_{\mathrm{deco}}}\right) \mathbb{E}_{x_0}\left[\varphi(x_s)\varphi(x_0)\right] \,\mathrm{d}s - 2t_{\mathrm{deco}}\mathbb{E}[\varphi]^2.$$

Implementation. It is crucial to compute on-the-fly the first term of the estimator (3.72), without resorting to a double sum which is computationally prohibitive. In practice the sum $S_i = \sum_{j=0}^{N_{\mathrm{deco}}} \varphi(x^{i-j})$ is not recomputed from scratch at every time step but updated using $S_{i+1} = S_i + \varphi(x^{i+1}) - \varphi(x^{i-N_{\mathrm{deco}}})$. The complexity of this algorithm is thus independent of the choice of t_{deco}.

Chapter 4

Mobility estimation in the underdamped regime using control variates

This chapter accounts for a work carried on during a two month stay at Imperial College London, in collaboration with Grigorios Pavliotis.

Contents

4.1 Introduction

The mobility is a transport coefficient (see Section 1.3.2) which can be defined for the Langevin dynamics. We recall that the nonequilibrium Langevin equation (1.19) writes:

$$\begin{cases} \mathrm{d}q_t = \dfrac{1}{m} p_t \, \mathrm{d}t \,, \\ \mathrm{d}p_t = (-\nabla V(q_t) + \eta F) \, \mathrm{d}t - \dfrac{\gamma}{m} p_t \, \mathrm{d}t + \sqrt{2\gamma\beta^{-1}} \mathrm{d}W_t. \end{cases}$$

The system undergoes an external forcing of amplitude $\eta > 0$ in the direction given by the vector F of norm 1. The parameter $\beta > 0$ is proportional to the inverse temperature, $m > 0$ is the mass and $\gamma > 0$ is the friction parameter, which quantifies the intensity of the coupling of the system with the thermostat. The generator writes:

$$\mathcal{L}_{\gamma,\eta} = \mathcal{L}_{\mathrm{ham}} + \gamma \mathcal{L}_{\mathrm{OU}} + \eta \mathcal{L}_{\mathrm{pert}},$$

where

$$\mathcal{L}_{\mathrm{ham}} = \frac{1}{m} p^\top \nabla_p - \nabla V(q)^\top \nabla_q, \quad \mathcal{L}_{\mathrm{OU}} = -\frac{1}{m} p^\top \nabla_p + \beta^{-1} \Delta_p, \quad \mathcal{L}_{\mathrm{pert}} = F^\top \nabla_p,$$

are respectively the generators of the Hamiltonian part, the fluctuation-dissipation part and the nonequilibrium perturbation.

For any values of the parameters γ and η, there exists a unique invariant probability measure (see Section 1.3.2) denoted by $\mu_{\gamma,\eta}$. We denote by $\mathbb{E}_{\gamma,\eta}$ the expectation with respect to this probability measure. For dynamics at equilibrium ($\eta = 0$), the invariant probability measure admits the same explicit expression for any value of γ:

$$\mu_{\gamma,0}(\mathrm{d}q \, \mathrm{d}p) = Z_\beta^{-1} \mathrm{e}^{-\beta H(q,p)} \, \mathrm{d}q \, \mathrm{d}p,$$

where Z_β is a normalization factor. This equilibrium measure is denoted by μ_0 in the following. On the other hand, the invariant probability measure is modified when an external force is applied to the system ($\eta > 0$)Moreover the corresponding steady state $\mu_{\gamma,\eta}$ depends on γ. As a consequence, the mobility

$$\alpha_\gamma := \lim_{\eta \to 0} \frac{\mathbb{E}_{\gamma,\eta}\left[\frac{1}{m} F^\top p\right]}{\eta} = \frac{\beta}{m^2} \left\langle -\mathcal{L}_\gamma^{-1} F^\top p, F^\top p \right\rangle,$$

depends on the friction γ. In this equation, and in the remainder of this chapter, the norms and the scalar products are considered with respect to the equilibrium measure μ_0. This transport coefficient characterizes the behavior of the system in the diffusive limit, so that understanding the dependence of α_γ with γ sheds some light on the influence of γ on the dynamical properties of the Langevin equation.

Two limiting regimes are of particular interest: the overdamped limit $\gamma \to \infty$ and the underdamped limit $\gamma \to 0$. The first case is well understood in all dimensions [121, 60, 75,

102], and in particular

$$\alpha_\gamma \underset{\gamma\to\infty}{\propto} \gamma^{-1}.$$

In this chapter the symbol $\propto$ means that the ratio between the two terms converges to a constant. The underdamped limit is not as well understood, except in dimension one for which the same scaling holds [61, 62, 75]:

$$\alpha_\gamma \underset{\gamma\to 0}{\propto} \gamma^{-1} \quad \text{in 1D.}$$

Consider now the case of the underdamped limit for two-dimensional systems. Note that when the potential V is additively separable:

$$\forall q \in \mathcal{D}, \quad V(q) = V_1(q_1) + V_2(q_2),$$

the variable q_1 satisfies a one-dimensional Langevin equation with potential V_1. The mobility in the direction given by $F = e_1 := (1, 0)$ scales therefore as γ^{-1}, as in dimension one. More generally, one can go back to the one-dimensional case when the Hamiltonian is separable. The analysis of the underdamped limit when the dimension is larger than one is much more involved in the non-separable case [75]. It has been conjectured in [27] that the diffusion should generically converge to a constant in the latter case, though there is yet no numerical evidence. Formally the generator $\mathcal{L}_\gamma$ converges to the Hamiltonian $\mathcal{L}_{\rm ham}$, which is not elliptic and not invertible. To help obtaining some intuition on the actual behavior of underdamped systems, accurate numerical studies of the scaling of the mobility are precious. Computing the mobility with a good accuracy is challenging because of large relative statistical errors. Let us give some scalings in the one-dimensional case to make this point clear, similarly to what is done in Chapter 3.

Averages with respect to the invariant probability measure $\mu_{\gamma,\eta}$ can be written as the following power expansion for any smooth function φ with compact support:

$$\mathbb{E}_{\gamma,\eta}[\varphi] = \mathbb{E}_0[\varphi] + \sum_{k=1}^{\infty} \eta^k \mathbb{E}_0\left[(-\widetilde{\mathcal{L}}\mathcal{L}_\gamma^{-1}\Pi_0)^k \varphi\right],$$

where we denote by $\mathcal{L}_\gamma = \mathcal{L}_{\gamma,0} = \mathcal{L}_{\rm ham} + \gamma\mathcal{L}_{\rm OU}$ the generator of the equilibrium dynamics. Recall that $\Pi_0\varphi = \varphi - \mathbb{E}_0[\varphi]$. Note that this expansion is rigorous since $\mathcal{L}_{\rm pert}$ is $\mathcal{L}_\gamma$-bounded [104, Section 5.2], so that the operator $\sum_{k=1}^{\infty} \eta^k(-\widetilde{\mathcal{L}}\mathcal{L}_\gamma^{-1}\Pi_0)^k$ is well defined in $\mathcal{B}(\mathrm{L}^2(\mu_0))$ for η sufficiently small. In particular, using that

$$\left\|\mathcal{L}_\gamma^{-1}\Pi_0\right\|_{\mathcal{B}(\mathrm{L}^2(\mu_0))} \underset{\gamma\to 0}{\propto} \gamma^{-1}, \quad \left\|\mathcal{L}_{\rm pert}\mathcal{L}_\gamma^{-1}\Pi_0\right\|_{\mathcal{B}(\mathrm{L}^2(\mu_0))} = \mathrm{O}(\gamma^{-1}),$$

for the operator norm defined in (2.2), the mean drift can be developed to the second order in η:

$$\mathbb{E}_{\gamma,\eta}[p] = \eta\alpha_\gamma + \eta^2\mathbb{E}_0\left[\left(-\mathcal{L}_{\rm pert}\mathcal{L}_\gamma^{-1}\Pi_0\right)^2 p\right] + \mathrm{O}\left(\frac{\eta^3}{\gamma^3}\right).$$

We expect that

$$\mathbb{E}_0\left[\left(-\mathcal{L}_{\rm pert}\mathcal{L}_\gamma^{-1}\Pi_0\right)^2 p\right] \underset{\gamma\to 0}{\propto} \gamma^{-2},$$

so that the relative error on the mobility due to the finiteness of the perturbation η scales as

$$\frac{\eta^{-1}\mathbb{E}_{\gamma,\eta}[p]-\alpha_\gamma}{\alpha_\gamma} = \alpha_\gamma^{-1}\left(\eta\mathbb{E}_0\left[\left(-\mathcal{L}_{\rm pert}\mathcal{L}_\gamma^{-1}\Pi_0\right)^2 p\right] + \mathrm{O}(\frac{\eta^2}{\gamma^2})\right) \underset{\substack{\gamma\to 0\\ \frac{\eta}{\gamma}\to 0}}{\propto} \frac{\eta}{\gamma}.$$

Therefore the linear regime corresponds in the underdamped limit to the scaling $\eta \ll \gamma \ll 1$, so that the forcing η should be taken very small in the underdamped regime. On the other hand the relative statistical error committed on the statistical estimator of the mobility for a simulation time T scales

$$\frac{T^{-1/2}\eta^{-1}\sigma_{p,\gamma}}{\alpha_\gamma} \propto \sqrt{\frac{\gamma}{T\eta^2}},$$

where $\sigma_{p,\gamma}^2 = \beta^{-1}\alpha_\gamma \propto \gamma^{-1}$ is the asymptotic variance of the momentum for a friction γ. Hence the computational cost of a simulation achieving a relative error of order ε is of order $\frac{\gamma}{\eta^2\varepsilon^2} = \frac{1}{\gamma\varepsilon^2}\left(\frac{\gamma}{\eta}\right)^2 g1$. The large variance issue is therefore even more acute than in standard linear response estimation, since there is an additional factor $\frac{1}{\gamma}$.

Contrarily to the situation described in Chapter 3, the Poisson problem at equilibrium:

$$-\gamma^{-1}\mathcal{L}_\gamma\Phi_\gamma = p, \tag{4.1}$$

is ill-conditionned for γ small since the spectral gap of $\mathcal{L}_\gamma$ scales like γ in the underdamped regime. This is why we propose here another control variate through a homogenization procedure, and show that it allows to construct a relevant modified observable. Numerical simulations suggest that the asymptotic variance of this estimator scales like $\gamma^{-\alpha}$ with $0 < \alpha < 1$, instead of γ^{-1} in the underdamped regime. The exponent can be made arbitrarily close to 0, at the price of a diverging prefactor.

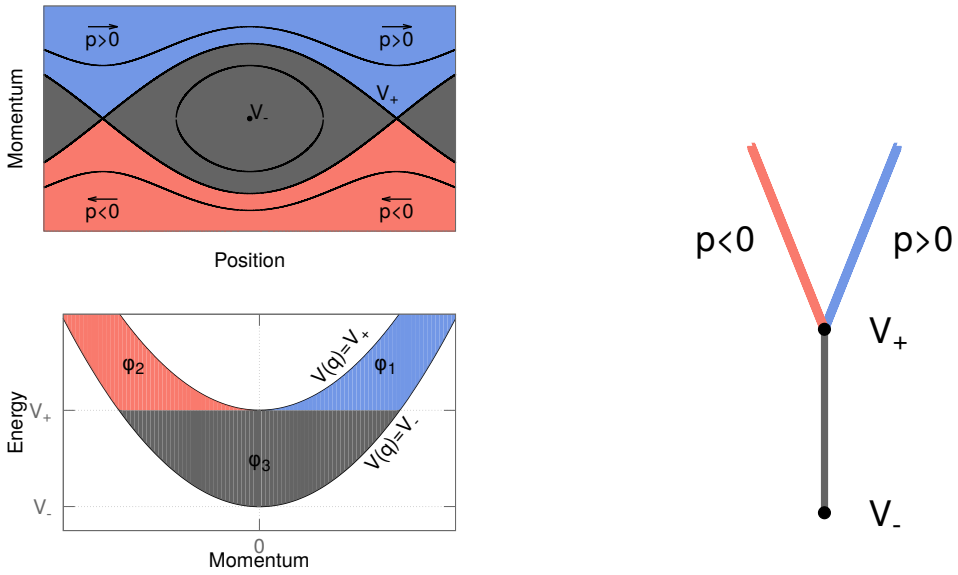
4.2 Underdamped limit

Let us first study the system at equilibrium ($\eta = 0$) on the one-dimensional unit torus ($\mathcal{D} = \mathbb{T}$). We first recall that the solution Φ_γ to (4.1) converges in $\mathrm{L}^2(\mu_0)$ towards a limit Φ_0 when $\gamma \to 0$, and that this limit is known analytically. This summary is based on the work of Hairer and Pavliotis [75]. We remark in a second step that the function Φ_0 provides an approximation of Φ_γ in the underdamped limit, so that $\mathcal{L}_\gamma\Phi_0$ is good candidate for a control variate.

4.2.1 Homogenized equation

In the underdamped limit the energy varies on a timescale γ^{-1} whereas the Hamiltonian orbits are visited at a rate independent of γ. The dynamics can therefore be intuitively split into these two processes, one slow and one fast. When the potential V is smooth, the set of

Figure 4.1: Top left: Phase portrait of the Hamiltonian dynamics. Bottom left: Accessible states for a given energy and momentum. Right: Graph Γ giving the structure of the Hamiltonian invariants.



the orbits of the Hamiltonian dynamics can be described by a graph Γ. In other words, each element of the graph corresponds to a connected level set of H. We denote this mapping by $\widetilde{H} : \mathcal{E} \mapsto \Gamma$ where $\mathcal{E}$ is the phase space. In the sequel, we denote by $V_- := \min V$ and $V_+ := \max V$.

Unimodal potential. Consider a quasi-convex potential V. We plot in Figure 4.1 the phase portrait of the Hamiltonian dynamics in the (q, p) variables. We also represent the states using (p, h) variables where $h = V(q) + \frac{1}{2}p^2$ is the energy. In this case the horizontal axis is the fast variable, the vertical axis is the slow variable and a connected orbit of the Hamiltonian dynamics corresponds to a horizontal segment. It appears that an energy greater than V_+ corresponds to two orbits, one for particles with $p > 0$ and one for particles with $p < 0$. In this case the graph of the invariants is composed of an edge which represents the bounded trajectories, connected with two edges which represent unbounded trajectories (when unfolded on $\mathbb{R}^2$) with momentum either positive or negative. Each element z of the graph is identified by an energy $h = H(z)$ and the index of its edge.

Limiting process. The limiting process of Langevin dynamics in the underdamped regime is a diffusion process on Γ. We define the following isometric embedding operator from observables of the graph Γ to observables of the state space $\mathcal{E}$ depending only on the Hamil-

tonian invariants:

$$\tau : \mathrm{L}^2(\widetilde{H}\sharp\mu) \to \mathrm{Ker}(\mathcal{L}_{\mathrm{ham}}) \subset \mathrm{L}^2(\mu_0)$$
$$f \mapsto \varphi = f \circ \widetilde{H}.$$

We denote by $\widetilde{H}\sharp\mu$ the probability measure on Γ which is the pushforward measure of μ by $\widetilde{H}$. It satisfies, for any test function f,

$$\int_{\mathcal{E}} (\tau f)(q,p)\,\mu(\mathrm{d}q\,\mathrm{d}p) = \int_{\mathcal{E}} f\left(\widetilde{H}(q,p)\right)\,\mu(\mathrm{d}q\,\mathrm{d}p) = \int_{\Gamma} f(z)\,\widetilde{H}\sharp\mu(\mathrm{d}z).$$

Denoting by Π_H the orthogonal projector on the kernel of $\mathcal{L}_{\mathrm{ham}}$ in $\mathrm{L}^2(\mu_0)$, composed of functions depending only on the Hamiltonian invariants. Let us define a diffusion process on $\mathrm{Ker}(\mathcal{L}_{\mathrm{ham}}) = \Pi_H \mathrm{L}^2(\mu_0)$ by its generator

$$\mathcal{S} := \Pi_H \mathcal{L}_{\mathrm{OU}} \Pi_H = \tau \mathcal{S}_\Gamma \tau^{-1} \Pi_H,$$

where

$$\mathcal{S}_\Gamma = \tau^{-1} \Pi_H \mathcal{L}_{\mathrm{OU}} \tau,$$

is the diffusion operator on the graph. In order to give a more constructive expression of the latter we introduce the period $T(z)$ of the orbit corresponding to $z \in \Gamma$ [126], and an antiderivative S of T:

$$\begin{aligned} T(z) &:= \int_{\widetilde{H}^{-1}(z)} |p|^{-1} \mathrm{d}\mu_0 = \int_{\substack{V(q)<H(z) \\ q\in\mathrm{edge(z)}}} \frac{\mathrm{d}q}{\sqrt{2}\sqrt{H(z)-V(q)}}, \\ S(z) &:= \int_{\widetilde{H}^{-1}(z)} |p|\,\mathrm{d}\mu_0 = \int_{\substack{V(q)<H(z) \\ q\in\mathrm{edge(z)}}} \sqrt{2}\sqrt{H(z)-V(q)}\,\mathrm{d}q. \end{aligned} \tag{4.2}$$

With this notation,

$$\mathcal{S}_\Gamma f(z) = \frac{1}{\beta T(z)} \partial_z (S(z) \partial_z f(z)) - \frac{S(z)}{T(z)} \partial_z f(z),$$

for z in the interior of the edges of the graph Γ [75]. The domain of $\mathcal{S}_\Gamma$ is composed of functions f such that the above expression is square integrable on Γ, and which satisfy the following gluing condition for every interior vertex z_0 of Γ:

$$\sum_{k\sim z_0} \sigma(z_0,k) \lim_{z \xrightarrow{k} z_0} S(z) \partial_z f(z) = 0,$$

where the sum is over the edges connected to z_0, the symbol $z \xrightarrow{k} z_0$ means that the point z converges to z_0 on the edge k and $\sigma(z_0,k)$ is 1 if $H(z) > H(z_0)$ and -1 otherwise.

Proposition 4.1. *The operator $-\Pi_H \mathcal{L}_{\mathrm{OU}} \Pi_H$ is coercive on $\mathrm{Ker}(\mathcal{L}_{\mathrm{ham}})$ in $\mathrm{L}^2(\mu_0)$, and in particular the Poisson equation*

$$-\Pi_H \mathcal{L}_{\mathrm{OU}} \Pi_H \Phi_0 = \Pi_H p, \tag{4.3}$$

admits a unique solution. Moreover

$$\Phi_\gamma \xrightarrow[\gamma\to 0]{\mathrm{L}^2(\mu_0)} \Phi_0.$$

A proof of the convergence is provided in [75] using probabilistic tools. We propose here an alternative proof for the convergence relying on analysis arguments.

Proof. Consider a sequence $(\gamma_k)_{k\in\mathbb{N}}$ with $\gamma_k > 0$ converging to 0. The sequence (Φ_{γ_k}) is bounded in $\mathrm{H}^1(\mu_0)$, due to spectral gap estimates for $\mathcal{L}_\gamma$ in $\mathrm{H}^1(\mu_0)$ (see [75, Proposition 1.7]). By the Rellich compactness theorem (extended to weighted spaces) [136, Th. XIII.65], $\mathrm{H}^1(\mu)$ is compactly embedded in $\mathrm{L}^2(\mu_0)$. This relies on a unitary transformation from $\mathrm{L}^2(\mu_0)$ to $\mathrm{L}^2(\mathrm{d}q\,\mathrm{d}p)$, which do not require any further assumption on V since the domain is compact. Therefore there exists a subsequence $\gamma_{k'}$ such that $\Phi_{\gamma_{k'}}$ converges in $\mathrm{L}^2(\mu_0)$ to some limit $\widetilde{\Phi} \in \mathrm{L}^2(\mu_0)$.

We denote by γ the general term of the subsequence $\gamma_{k'}$, and fix a smooth function $\psi \in \mathscr{C}$. By definition of Φ_γ,

$$\langle \mathcal{L}_{\mathrm{ham}}\Phi_\gamma, \psi\rangle = \gamma \left\langle p - \mathcal{L}_{\mathrm{OU}}\Phi_\gamma, \psi\right\rangle,$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product in $\mathrm{L}^2(\mu_0)$, so that

$$|\left\langle \Phi_\gamma, \mathcal{L}_{\mathrm{ham}}\psi\right\rangle| \leqslant \gamma |\left\langle p, \psi\right\rangle| + \gamma \|\Phi_\gamma\|\,\|\mathcal{L}_{\mathrm{OU}}\psi\|,$$

where we used that $\mathcal{L}_{\mathrm{ham}}$ and $\mathcal{L}_{\mathrm{OU}}$ are respectively skew-symmetric and symmetric on $\mathrm{L}^2(\mu_0)$. By passing to the limit $\gamma \to 0$ we get using the skew-symmetry of $\mathcal{L}_{\mathrm{ham}}$ that $\left\langle -\mathcal{L}_{\mathrm{ham}}\widetilde{\Phi}, \psi\right\rangle = 0$ for any $\psi \in \mathscr{C}$. Therefore $\mathcal{L}_{\mathrm{ham}}\widetilde{\Phi} = 0$ in the sense of distributions. Moreover,

$$\begin{aligned}\langle \Pi_H\mathcal{L}_{\mathrm{OU}}\Phi_\gamma + \Pi_H p, \psi\rangle &= \left\langle \Pi_H\mathcal{L}_{\mathrm{OU}}\Phi_\gamma - \gamma^{-1}\Pi_H\left(\mathcal{L}_{\mathrm{ham}} + \gamma\mathcal{L}_{\mathrm{OU}}\right)\Phi_\gamma, \psi\right\rangle \\ &= -\gamma^{-1}\left\langle \Phi_\gamma, \mathcal{L}_{\mathrm{ham}}\Pi_H\psi\right\rangle = 0,\end{aligned}$$

since $\mathcal{L}_{\mathrm{ham}}\Pi_H = 0$. Therefore $\langle\Phi_\gamma, \mathcal{L}_{\mathrm{OU}}\Pi_H\psi\rangle = \langle\Pi_H p, \psi\rangle$ and by passing to the limit in γ we obtain

$$-\Pi_H\mathcal{L}_{\mathrm{OU}}\Pi_H\widetilde{\Phi} = \Pi_H p, \tag{4.4}$$

in the sense of distributions. Therefore, by uniqueness of the solution of (4.4) in $\mathrm{L}^2(\mu_0)$, $\widetilde{\Phi} = \Phi_0$ is the only possible accumulation point of the sequence (Φ_{γ_k}), which implies the claimed convergence. □

Remark 4.1. *In order to formally derive equations for the underdamped limit of Φ_γ it is tempting to write the following expansion*

$$\Phi_\gamma = \Phi^0 + \gamma\Phi^1 + \gamma^2\Phi^2 + \cdots,$$

and consider the following hierarchy obtained by substituting in (4.1)*:*

$$\begin{aligned} \mathcal{L}_{\mathrm{ham}}\Phi^0 &= 0, \\ \mathcal{L}_{\mathrm{OU}}\Phi^0 + \mathcal{L}_{\mathrm{ham}}\Phi^1 &= -p, \\ \mathcal{L}_{\mathrm{OU}}\Phi^1 + \mathcal{L}_{\mathrm{ham}}\Phi^2 &= 0, \\ &\vdots \end{aligned}$$

In fact the lack of regularization properties of the Hamiltonian generator invalidates this expansion. In particular it has been suggested that a term of order $\gamma^{1/2}$ should be taken into account [75, 149, 150].

The function Φ_0 is obtained by solving an ODE on each edge of the graph. The full computation is provided in [126].

Corollary 4.1. *The solution Φ_0 to* (4.3) *writes*

$$\Phi_0(q,p) = \begin{cases} 0 & \text{if} \quad z \leqslant V_+ \\ \mathrm{sgn}(p) \displaystyle\int_{V_+}^{z} \frac{1}{S} & \text{if} \quad z > V_+ \end{cases}, \quad \text{where} \quad z = H(q,p) = \frac{p^2}{2m} + V(q),$$

and we recall that S is defined in (4.2)*.*

4.2.2 Control variate

Let us now construct a control variate using the function Φ_0 previously introduced. We recall that in our context a control variate is a function $\xi_{\gamma,\eta}$ on $\mathcal{E}$ such that $\mathbb{E}_{\gamma,\eta}[\xi_{\gamma,\eta}] = 0$ and the asymptotic variance of $p + \xi_{\gamma,\eta}$ is smaller than the variance of p. Using that, for any function Φ, it holds $\mathbb{E}_{\gamma,\eta}[\mathcal{L}_{\gamma,\eta}\Phi] = 0$, we consider $\xi_{\gamma,\eta} = \gamma^{-1}\mathcal{L}_{\gamma,\eta}\Phi$ for some function Φ. As discussed in Chapter 3, the optimal choice for Φ is the solution to the Poisson problem

$$-\gamma^{-1}\mathcal{L}_{\gamma,\eta}\Phi_{\gamma,\eta} = p - \mathbb{E}_{\gamma,\eta}[p],$$

since the modified observable $p + \xi_{\gamma,\eta}$ would be constant. However, this Poisson equation is impossible to solve in practice since $\mathbb{E}_{\gamma,\eta}[p]$ is not known. The idea here is to replace $\Phi_{\gamma,\eta}$ by its limit Φ_0 in the regime $\gamma, \frac{\eta}{\gamma} \ll 1$ under consideration. The modified observable then writes

$$\zeta_{\gamma,\eta}(q,p) = p + \gamma^{-1}\mathcal{L}_{\gamma,\eta}\Phi_0 = p + \mathcal{L}_{\mathrm{OU}}\Phi_0 + \frac{\eta}{\gamma}\partial_p\Phi_0,$$

since $\mathcal{L}_{\mathrm{ham}}\Phi_0 = 0$ by definition of Φ_0 (see (4.3)).

Integrability of the modified observable. Let us now study the integrability of $\zeta_{\gamma,\eta}$ in $\mathrm{L}^1(\mu_0)$ and $\mathrm{L}^2(\mu_0)$. We note first that Φ_0 vanishes for $H(q,p) \leqslant V_+$ and that it is $\mathcal{C}^\infty$ on $\{(q,p) \in \mathcal{E} \,|\, H(q,p) > V_+\}$. The factor $\mathrm{sgn}(p)$ do not produce any discontinuity since Φ_0 vanishes for $p = 0$, so that we can compute $\partial_p\Phi_0(q,p) = \mathrm{sgn}(p)\mathbf{1}_{H(q,p)>V_+}\frac{p}{mS(H(q,p))}$. An

additional straightforward computation leads to

$$\zeta_{\gamma,\eta}(q,p) = p + \operatorname{sgn}(p)\mathbf{1}_{H(q,p)>V_+}\left[\frac{\beta^{-1}+\frac{\eta}{\gamma}p-\frac{p^2}{m}}{mS(H(q,p))} - \frac{p^2T(H(q,p))}{\beta m^2 S(H(q,p))^2}\right].$$

It should be clear that what matters is the integrability of $\mathbf{1}_{H(q,p)>V_+}\left[\frac{\beta^{-1}+\frac{\eta}{\gamma}p-p^2}{S(H(q,p))} - \frac{p^2T(H(q,p))}{\beta S(H(q,p))^2}\right]$. The domain $\mathcal{D}$ is compact so the two regimes to be investigated correspond to p large (say positive) and $H(q,p) = V_+ + \varepsilon$ for $\varepsilon \to 0^+$, which is the neighborhood of the separatrix $\{(q,p) \in \mathcal{E} \,|\, H(q,p) = V_+\}$. Since $p \mapsto S(H(q,p))$ is increasing p and $p \mapsto T(H(q,p))$ is decreasing (see Equation (4.2)), it is clear that $\zeta_{\gamma,\eta}(q,p)$ grows at most like p^2 in the limit $p \to +\infty$. Moreover the invariant probability measure integrates any power law. Indeed,

$$\begin{aligned}
\mathcal{L}_{\gamma,\eta}\mathrm{e}^{\beta p^2/4} &= \left(\beta\frac{p}{2m}V'(q) + \frac{\gamma}{2m} + \frac{p^2\beta\gamma}{4m^2} - \frac{\beta p^2\gamma}{2m^2} + \frac{\beta\eta p}{2m}\right)\mathrm{e}^{\beta p^2/4} \\
&\leqslant \left(\beta\frac{|p|}{2m}\|V'\|_\infty + \frac{\gamma}{2m} - \frac{p^2\beta\gamma}{4m^2} + \frac{\beta\eta p}{2m} + 1\right)\mathrm{e}^{\beta p^2/4} - \mathrm{e}^{\beta p^2/4} \\
&\leqslant a - \mathrm{e}^{\beta p^2/4},
\end{aligned}$$

where $a := \max_p \left(\beta\frac{|p|}{2m}\|V'\|_\infty + \frac{\gamma}{2m} - \frac{p^2\beta\gamma}{4m^2} + \frac{\beta\eta p}{2m} + 1\right)\mathrm{e}^{\beta p^2/4} < +\infty$ since the function of p in the parentheses becomes negative outside a compact. By integrating both sides against $\mu_{\gamma,\eta}$ we obtain

$$\int_{\mathcal{E}} \mathrm{e}^{\beta p^2/4}\,\mathrm{d}\mu_{\gamma,\eta} \leqslant a < +\infty,$$

so that in particular $\mu_{\gamma,\eta}$ integrates any polynomial in p.

The sole potential integrability issue is therefore at the separatrix. For any z such that $H(z) > V_+$, S depends only on the energy $H(z)$ associated to the orbit z. Hence we abuse notation and identify z and $H(z)$: using that $\{q \in \mathcal{D} \,|\, V(q) \leqslant V_+\} = \mathbb{T}$, (4.2) rewrites

$$S(z) = \int_{\mathbb{T}} \sqrt{2}\sqrt{z - V(q)}\mathrm{d}q, \quad T(z) = S'(z) = \int_{\mathbb{T}} \frac{\mathrm{d}q}{\sqrt{2}\sqrt{z-V(q)}},$$

so S is continuous on $[V_+, +\infty)$ with

$$S(V_+) = \int_{\mathbb{T}} \sqrt{2}\sqrt{V_+ - V(q)}\,\mathrm{d}q.$$

On the other hand

$$T(V_+ + \varepsilon) = \int_{\mathbb{T}} \frac{\mathrm{d}q}{\sqrt{2}\sqrt{\varepsilon + V_+ - V(q)}} = \frac{1}{\sqrt{2\varepsilon}}\int_{\mathbb{T}} \left(1 + \frac{V_+ - V(q)}{\varepsilon}\right)^{-1/2}\mathrm{d}q \leqslant \frac{1}{\sqrt{2\varepsilon}},$$

using that $\left(1 + \frac{V_+ - V(q)}{\varepsilon}\right)^{-1/2} \leqslant 1$. When V is constant, $V(q) = V_+$, then $T(V_+ + \varepsilon) = \frac{2\pi}{\sqrt{2\varepsilon}}$.

Let us now study the case when V_+ is reached at a unique isolated point (we assume that

this point is in 0) and V is locally quadratic around this minimizer: there exist $0 < a < b$ and $0 < \delta < 1/2$ such that for any $q \in [-\delta, \delta]$,

$$aq^2 \leqslant V_+ - V(q) \leqslant bq^2,$$

and there exist $0 < a'$ such that for any $q \in \mathbb{T} \setminus [-\delta, \delta]$, $a' \leqslant V_+ - V(q)$. Then

$$\int_{[-\delta,\delta]} \left(1 + \frac{bq^2}{\varepsilon}\right)^{-1/2} \mathrm{d}q \leqslant \sqrt{2\varepsilon} T(V_+ + \varepsilon) \leqslant \int_{[-\delta,\delta]} \left(1 + \frac{aq^2}{\varepsilon}\right)^{-1/2} \mathrm{d}q + \int_{[-\delta,\delta]^c} \left(1 + \frac{a'}{\varepsilon}\right)^{-1/2} \mathrm{d}q,$$

where

$$\int_{[-\delta,\delta]} \left(1 + \frac{bq^2}{\varepsilon}\right)^{-1/2} \mathrm{d}q = \sqrt{\varepsilon/b} \left[\log\left(x + \sqrt{x^2+1}\right)\right]_{-\delta\sqrt{b/\varepsilon}}^{\delta\sqrt{b/\varepsilon}} = 2\sqrt{\varepsilon/b} \log\left(2\delta\sqrt{b/\varepsilon}\right) \mathrm{d}q + \mathrm{o}(\sqrt{\varepsilon}),$$

so that

$$\sqrt{2/b} \log\left(2\delta\sqrt{b/\varepsilon}\right) + \mathrm{o}(1) \leqslant T(V_+ + \varepsilon) \leqslant \sqrt{2/a} \log\left(2\delta\sqrt{a/\varepsilon}\right) + \mathrm{o}(1) + 2\delta,$$

and in conclusion

$$-\log(\varepsilon)/\sqrt{2b} + \mathrm{O}(1) \leqslant T(V_+ + \varepsilon) \leqslant -\log(\varepsilon)/\sqrt{2a} + \mathrm{O}(1). \tag{4.5}$$

In this case $\zeta_{\gamma,\eta}(q,p)$ diverges like $-\log(\varepsilon)$ when $H(q,p) = V_+ + \varepsilon$, so that it is in $\mathrm{L}^p_{\mathrm{loc}}(\mathrm{d}q\,\mathrm{d}p)$ for any $p \geqslant 1$. Using the fact that the measures $\mu_{\gamma,\eta}$ integrate any polynomial in p, and that they are locally bounded, the function $\zeta_{\gamma,\eta}(q,p)$ is in $\mathrm{L}^p(\mu_{\gamma,\eta})$ for any $p \geqslant 1$.

4.3 Numerical results

4.3.1 One-dimensional oscillator

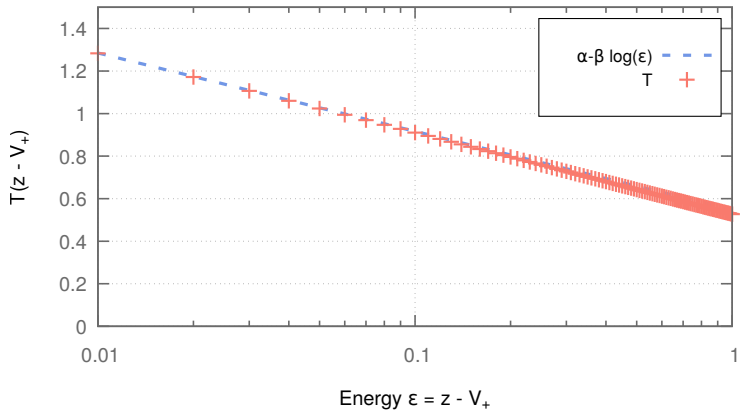
We set the temperature and the mass to 1, and consider the potential

$$V(q) = 1 - \cos(2\pi q).$$

This potential admits a unique maximizer and is locally quadratic around minimizer. We check on Figure 4.2 that the function T diverges logarithmically at the separatrix, with a prefactor $1/(2\pi)$ as expected from (4.5) (a and b can be taken arbitrarily close to $V''(0)/2 = 2\pi^2$). The dynamics is numerically integrated using a Geometric Langevin Algorithm (GLA) [23], which ensures that the invariant probability measure is correct up to terms of order $\mathrm{O}(\Delta t^2)$ at equilibrium. We choose a time step of $\Delta t = 0.02$.

In order to evaluate the modified observable $\zeta_{\gamma,\eta}$ at each time step we would need to evaluate S and T for the energy of the current step. The functions S and T are however not explicit so that we can only provide an approximation of their values using numerical integral quadratures. Therefore the corresponding approximation of $\zeta_{\gamma,\eta}$ is not exactly of

Figure 4.2: Function $\varepsilon \mapsto T(V_+ + \varepsilon)$ for a cosine potential. We observe that the divergence at $\varepsilon = 0$ is logarithmic as expected.



mean $\mathbb{E}_{\gamma,\eta}[p]$ with respect to the invariant measure.

For this reason we construct a piecewise linear approximation $\chi^{\Delta z}$ of the discontinuous function $\chi : z \mapsto \frac{\mathbf{1}_{z>V_+}}{S(z)}$, where the space step $\Delta z > 0$ is a discretization parameter. The value of the function χ is computed for $z = \left(k + \frac{1}{2}\right)\Delta z$ with $k \in \mathbb{N}$ using integral quadratures. In the end the modified observable we consider writes

$$\zeta^{\Delta z}_{\gamma,\eta}(q,p) = p + \mathcal{L}_{\gamma,\eta}\Phi^{\Delta z}_0(q,p), \qquad \Phi^{\Delta z}_0(q,p) = p + \mathrm{sgn}(p)\int_0^z \chi^{\Delta z},$$

where $z \mapsto \int_0^z \chi^{\Delta z}$ is twice differentiable and piecewise quadratic.

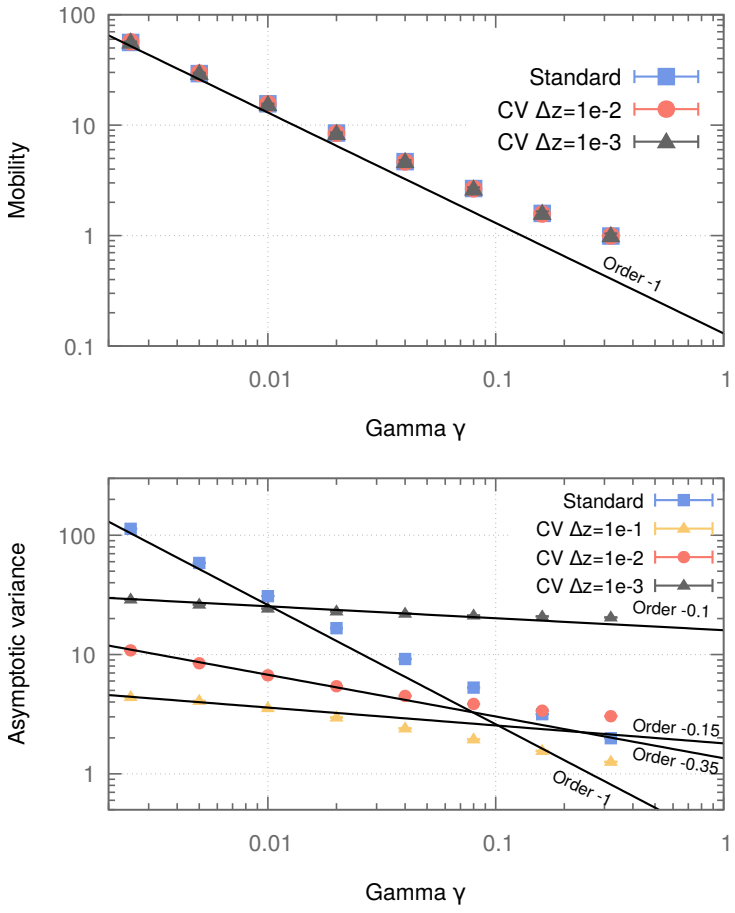
We check on Figure 4.3 (Left) that the estimator of the mobility relying on the control variate is unbiased, up to time discretization error depending on Δt. The mobility scales as γ^{-1} in the underdamped limit, as expected.

Figure 4.3 (Right) shows that the variance of the modified observable $\zeta^{\Delta z}_{\gamma,\eta}$ behaves better than the variance of the momentum observable p in the small γ regime. In particular for $\gamma = 0.0025$ and $\Delta z = 0.1$, we reach a speed-up ratio of order 20. Surprisingly we observe that the efficiency of the method decays when $\Delta z \to 0$, which is when the modified observable $\zeta^{\Delta z}_{\gamma,\eta}$ converges to $\zeta_{\gamma,\eta}$. In practice a coarse approximation with $\Delta z = 0.1$ gives very good results. In the present regime where $\eta \ll \gamma$, the asymptotic variances which we consider are only slightly different from the ones obtained for $\eta = 0$. This justifies a posteriori the resolution of the underdamped limiting equation at equilibrium.

4.3.2 Two-dimensional oscillator

Now that the variance reduction approach has been validated on a simple one-dimensional problem, we can use it to investigate the scaling of the mobility in dimension two when the

Figure 4.3: Left: Estimated mobility with and without control variate. Right: Asymptotic variance as a function of γ at equilibrium ($\eta = 0$), either for the standard observable p (Direct) or for the modified one (CV). The latter observable is computed using an energy step $\Delta z = 0.01$, $\Delta z = 0.001$ or $\Delta z = 0.0001$.



system is non-integrable. We consider the family of potentials

$$V_\delta(q) = 2 - \cos(q_1) - \cos(q_2) + \delta \exp(\sin(q_1 + q_2)),$$

which are separable if and only if $\delta = 0$. The exponential function introduces a non-linearity aiming at breaking the separable structure of the Hamiltonian. We estimates the mobility in the direction $F = e_1 = (1, 0)$, so that the Poisson problem writes

$$-\mathcal{L}_{\eta,\gamma,\delta}\Phi_{\eta,\gamma,\delta} = p_1 - \mathbb{E}_{\eta,\gamma,\delta}[p_1].$$

We approximate $\Phi_{\eta,\gamma,\delta}$ by the function $\Phi_0(q, p) = \Phi_0(q_1, p_1)$ previously computed, so that the approach is perturbative in η, γ and δ. Indeed for $\delta = 0$, the solution $\Phi_{\eta,\gamma,0}(q, p) = \Phi_{\eta,\gamma}(q_1, p_1)$ depends only on q_1 and p_1, so we expect Φ_0 to be a good approximation is the small δ regime. The modified observable is defined by

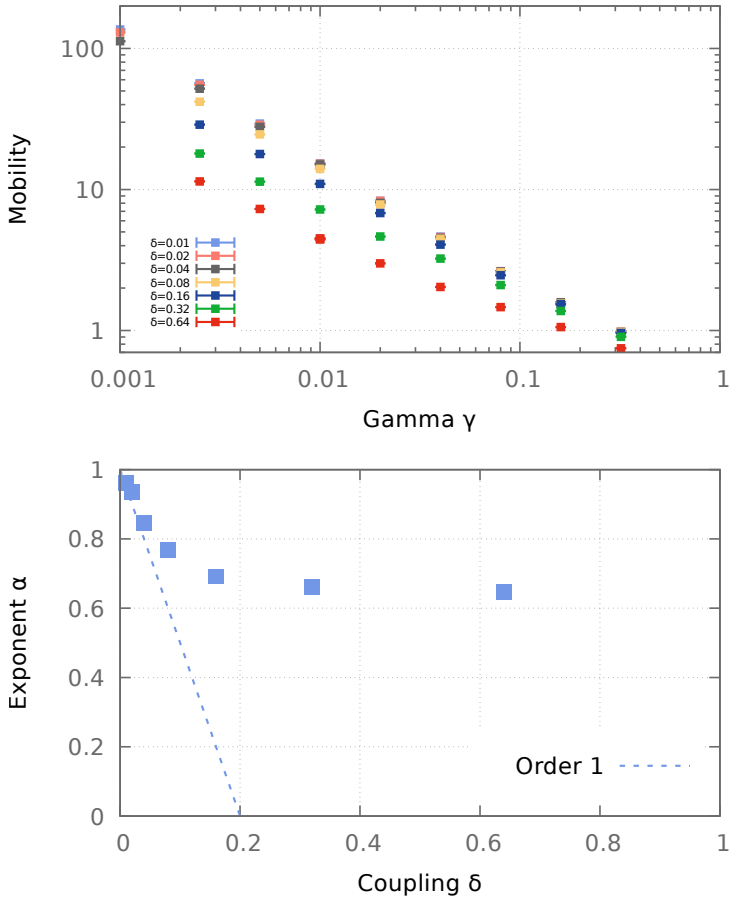
$$\zeta_{\gamma,\eta,\delta}(q, p) = p + \gamma^{-1}\mathcal{L}_{\gamma,\eta,\delta}\Phi_0(q, p),$$

which is by construction an unbiased estimator of the mean drift $\mathbb{E}_{\gamma,\eta,\delta}[p]$.

We run the two-dimensional Langevin dynamics with mass 1 and temperature 1, using the Geometric Langevin Algorithm (see Section 1.2.4) with time step $\Delta t = 0.02$. The energy step is set to $\Delta z = 0.01$ for the computation of the function Φ_0. which has already been computed previously. We vary γ and δ in a range of values, keeping a forcing amplitude $\eta = 0.08\gamma$ as this seems to be a good trade-off between bias and statistical error in the estimation of the mobility.

We plot in Figure 4.4 (Left) the mobility as a function of the friction γ for several values of the coupling δ. The asymptotic behavior in the underdamped regime is compatible with a power law scaling $\gamma^{-\alpha}$ with an exponent α depending on δ (Right). As expected the exponent α converges to 1 as $\delta \to 0$ since the potential is separable in this limit. We observe that $1-\alpha$ depends linearly on δ in the regime of small couplings δ. In particular the mobility does not seem to be bounded in the underdamped regime for these non-integrable systems, contrarily to the conjecture made in [27].

Figure 4.4: Top: Estimated mobility for different values δ of the coupling as a function of the friction γ. Bottom: Exponent α involved in the dependence of the mobility with respect to γ, as a function of the coupling δ.



Chapter 5

Hypercoercivity of Piecewise Deterministic Markov Process-Monte Carlo

This chapter provides the content of [4] with some changes of notation and minor changes.

Contents

In this paper we derive spectral gap estimates for several Piecewise Deterministic Markov Processes, namely the Randomized Hamiltonian Monte Carlo, the Zig-Zag process and the Bouncy Particle Sampler. The hypocoercivity technique we use, presented in [44], produces estimates with explicit dependence on the parameters of the dynamics. Moreover the general framework we consider allows to compare quantitatively the bounds found for the different methods.

5.1 Introduction

Consider a probability distribution π defined on the Borel σ-field $\mathcal{X}$ of some domain $\mathsf{X} = \mathbb{R}^d$ or $\mathsf{X} = \mathbb{T}^d$ where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Assume that π has a density with respect to the Lebesgue measure also denoted $\pi(x)$ and of the form

$$\pi \propto \mathrm{e}^{-U}, \tag{5.1}$$

where $U \colon \mathsf{X} \to \mathbb{R}$ is referred to as the potential. Sampling from such distributions is of interest in computational statistical mechanics and in Bayesian statistics and allows one, for example, to compute efficiently expectations of functions f with respect to π by invoking empirical process limit theorems, e.g. the law of large numbers. In practical set-ups, sampling exactly from π directly is either impossible or computationally prohibitive. A standard and versatile approach to sampling from such distributions with arbitrary precision consists of using Markov Chain Monte Carlo (MCMC) techniques [65, 108, 140], where the ability of simulating realisations of ergodic Markov chains leaving π invariant is exploited. Markov Process Monte Carlo (MPMC) methods are the continuous time counterparts of MCMC but their exact implementation is most often impossible on computers and requires additional approximation, such as time discretisation of the process in the case of the Langevin diffusion. A notable exception, which has recently attracted significant attention, is the class of MPMC relying on Piecewise Deterministic Markov Processes (PDMP) [38], which in addition to being simpler to simulate than earlier MPMC are nonreversible, offering the promise of better performance. We now briefly introduce a class of processes covering existing algorithms.

Known PDMP Monte Carlo methods rely on the use of the auxiliary variable trick, that is the introduction of an instrumental variable and probability distribution μ defined on an extended domain, of which π is a marginal distribution, which may facilitate simulation. In the present set-up one introduces the velocity variable $v \in \mathsf{V} \subset \mathbb{R}^d$ associated with a probability distribution ν defined on the σ-field $\mathcal{V}$ of V, where the subset V is assumed closed. Standard choices for ν include the centered normal distribution of covariance $m_2 \, \mathrm{I}_d$, where I_d is d-dimensional identity matrix, the uniform distribution on the unit sphere $\mathbb{S}^{d-1}$, or the uniform distribution on $\mathsf{V} = \{-1, 1\}^d$. Let $\mathsf{E} = \mathsf{X} \times \mathsf{V}$ and define the probability measure $\mu = \pi \otimes \nu$. The aim is now to sample from the probability distribution μ.

We denote by $\mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$ the set of bounded functions of $\mathrm{C}^2(\mathsf{E})$. The PDMP Monte Carlo algorithms we are aware of fall in a class of processes associated with generators of the form,

for $f \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$ and $(x, v) \in \mathsf{E}$,

$$\mathcal{L}_1 f(x,v) = v^\top \nabla_x f(x,v) + \sum_{k=1}^{K} (v^\top F_k(x))_+ \left(\mathcal{B}_k - \mathrm{Id}\right) f(x,v) + m_2^{1/2} \lambda_{\mathrm{ref}}(x) \mathcal{R}_v f(x,v) \;, \quad (5.2)$$

where $K \in \mathbb{N}$, $F_k \colon \mathsf{X} \to \mathsf{X}$ for $k \in \{1, \ldots, K\}$, $(\mathcal{R}_v, \mathrm{D}(\mathcal{R}_v))$ and $(\mathcal{B}_k, \mathrm{D}(\mathcal{B}_k))$, for $k \in \{1, \ldots, K\}$, are operators we specify below, and

$$m_2 = \|v_1\|_2^2 = \int_{\mathsf{V}} v_1^2 \, \mathrm{d}\nu(v) \;, \quad (5.3)$$

which is assumed to be finite. The choice of jump rates $(x, v) \mapsto (v^\top F_k(x))_+$, which together with other conditions ensures that μ is an invariant distribution of the associated semi-group. In the case where $\mathsf{V} = \mathbb{R}^d$ and ν is the zero-mean Gaussian distribution on $\mathbb{R}^d$ with covariance matrix $m_2 \, \mathrm{I}_d$, we also consider generators of the form, for any $f \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$ and $(x, v) \in \mathsf{E}$,

$$\mathcal{L}_2 f(x,v) = \mathcal{L}_1 f(x,v) - m_2 F_0(x)^\top \nabla_v f(x,v) \;, \quad (5.4)$$

where $F_0 : \mathsf{X} \to \mathsf{X}$. In all the paper we assume the following condition for either $\mathcal{L}_1$ or $\mathcal{L}_2$ which is satisfied for most examples we consider in this document. Denote by $\mathrm{L}^2(\mu)$ the set of measurable function $g : \mathsf{E} \to \mathbb{R}$ such that $\int_{\mathsf{E}} g^2 \, \mathrm{d}\mu < +\infty$. We let $\| \cdot \|_2$ be the norm induced by the scalar product

$$\text{for all } f, g \in \mathrm{L}^2(\mu) \;, \quad \langle f, g \rangle_2 = \int_{\mathsf{E}} f \, g \, \mathrm{d}\mu \;, \quad (5.5)$$

making $\mathrm{L}^2(\mu)$ a Hilbert space.

Assumption 5.1. *The operator $\mathcal{L}$ is closed in $\mathrm{L}^2(\mu)$, generates a strong contraction semi-group $(P_t)_{t \geqslant 0}$ on $\mathrm{L}^2(\mu)$ for which μ is a stationary measure.*

As we shall see, the $K + 1$ vector fields F_k are tied to the potential U by the relation $\nabla U(x) = \sum_{k=0}^{K} F_k(x)$, required to ensure that μ is left invariant by the associated semi-group. Informally, assuming for the moment that $\lambda_{\mathrm{ref}} = 0$, the corresponding process follows the solution of Hamilton's equations $(\dot{x}_t, \dot{y}_t) = (y_t, 0)$ for a random time of distribution governed by an inhomogeneous Poisson process of intensity $\sum_{k=1}^{K} (v^\top F_k(x))^+$. When an event occurs one chooses between the K possible updates of the state available, with probability proportional to $(v^\top F_1(x))_+, \ldots, (v^\top F_K(x))_+$, with the particularity here that the position x is left unchanged.

We will refer to $\mathcal{R}_v$ as the refresh operator, a standard example of which is $\mathcal{R}_v = \Pi_v - \mathrm{Id}$ where Π_v the following orthogonal projector in $\mathrm{L}^2(\mu)$

$$\Pi_v f(x,v) = \int_{\mathsf{V}} f(x,w) \, \mathrm{d}\nu(w) \;, \quad (5.6)$$

in which case the velocity is drawn afresh from the marginal invariant distribution, while the position is left unchanged. In this scenario the informal description of the process given above carries on with $\lambda_{\mathrm{ref}} \neq 0$ added to the Poisson intensity and Π_v now one of $K + 1$

possible updates, chosen with probability proportional to λ_{ref}. Another possible choice is the generator of an Ornstein-Uhlenbeck operator leaving ν invariant. For any $k \in \{1, \dots, K\}$, the jump operators $\mathcal{B}_k$ we consider are of the form

$$\mathcal{B}_k f(x,v) = f\left(x, v - 2(v^\top \mathrm{n}_k(x))\,\mathrm{n}_k(x)\right) \;, \quad \mathrm{n}_k(x) = \begin{cases} F_k(x)/|F_k(x)| & \text{if } F_k(x) \neq 0 \;, \\ 0 & \text{otherwise} \;. \end{cases} \tag{5.7}$$

They correspond to bounces on the hyperplanes orthogonal to $F_k(x)$ at the event position x, *i.e.* a flip of the component of the velocity in the direction given by F_k. We now describe how various choices of K and F_k lead to known algorithms. For simplicity of exposition we assume for the moment that $\mathsf{V} = \mathbb{R}^d$, ν is the zero-mean Gaussian distribution with covariance matrix $m_2\,\mathrm{I}_d$ and $\mathcal{R}_v = \Pi_v - \mathrm{Id}$, but as we shall see later our results cover more general scenarios.

- The particular choice $K = 0$ and $F_0 = \nabla U$ corresponds to the procedure described in [45] as a motivation for the popular hybrid Monte Carlo method. This process is also known as the Linear Boltzman/kinetic equation in the statistical physics literature or randomized Hamiltonian Monte Carlo and was recently studied theoretically in [44, 24, 22]. In this scenario the process follows the isocontours of μ for random times distributed according to an inhomogeneous Poisson law of parameter λ_{ref}, triggering events where the velocity is sampled afresh from ν.

- The scenario where $K = d$, $F_0 = 0$ and for $k \in \{1, \dots, d\}$, $x \in \mathsf{X}$, $F_k(x) = \partial_k U(x)\mathbf{e}_k$ where $(\mathbf{e}_k)_{k\in\{1,\dots,k\}}$ is the canonical basis, corresponds to the Zig-Zag (ZZ) process [16], where the x component of the process follows straight lines in direction v which remains constant between events. In this scenario, the choice of $\mathcal{B}_k$ to update velocity consists of negating the k-th component of the velocity; see also [55] for related ideas motivated by other applications.

- The standard Bouncy Particle Sampler (BPS) of [130], extended by [25], corresponds to the choice $K = 1$, $F_0 = 0$ and $F_1 = \nabla U$.

- More elaborate versions of the ZZ and BPS processes, motivated by computational considerations, take advantage of the possibility to decompose the energy as $U = \sum_{k=0}^{K} U_k$ and corresponds to the choice $F_k = \nabla U_k$ [114, 25], where in the former the sign flip operation is replaced with a component swap.

- It should be clear that one can consider more general deterministic dynamics with $F_0 \neq 0$, effectively covering the Hamiltonian Bouncy Particle Sampler, suggested in [162].

- We remark that the well-known Langevin algorithm corresponds to $K = 0$, $F_0 = \nabla U$ and the situation where $\mathcal{R}_v$ is the Ornstein-Uhlenbeck process.

More general bounces involving randomisation (see [162, 170, 115]) can also be considered in our framework, at the cost of additional complexity and reduced tightness of our bounds.

The main aim of the present paper is the study of hypercoercivity [165, 166] for the class of processes described above. More precisely, consider $(P_t)_{t\geqslant 0}$ the semigroup associated to the PDMP with generator $\mathcal{L}$ as above, we aim to find simple and verifiable conditions on $U, F_k, \mathcal{R}_v$ and λ_{ref} ensuring the existence of $C \geqslant 1$ and $\alpha > 0$, and their explicit computation in terms of characteristics of the data of the problem, such that for any $f \in \mathrm{L}^2_0(\mu) = \{g \in \mathrm{L}^2(\mu) : \int_{\mathsf{E}} g \,\mathrm{d}\mu = 0\}$,

$$\|P_t f\|_2 \leqslant C\mathrm{e}^{-\alpha t} \|f\|_2 \ , \tag{5.8}$$

We will use the same notation for vector and matrix fields $\Phi, \Gamma \in \left(\mathbb{R}^d\right)^{\mathsf{E}}$ or $\left(\mathbb{R}^{d\times d}\right)^{\mathsf{E}}$, i.e. $\langle \Phi, \Gamma \rangle_2 = \int_{\mathsf{E}} \mathrm{Tr}\left(\Phi^\top \Gamma\right) \mathrm{d}\mu$ and no confusion should be possible. For $\Phi, \Gamma \in \mathbb{R}^d$ or $\mathbb{R}^{d\times d}$, associated to the usual Frobenius inner product $\mathrm{Tr}(\Phi^\top \Gamma)$ is the norm $|\Phi|$.

Establishing such a result is of interest for multiple reasons. Explicit bounds may provide insights into expected performance properties of the algorithm in various situations or regimes. For example the above leads to an upper bound of the asymptotic variance, for any $f \in \mathrm{L}^2(\mu)$,

$$\lim_{T\to\infty} T \,\mathrm{Var}\left(T^{-1} \int_0^T f(X_t, V_t) \,\mathrm{d}t\right) \leqslant (2C/\alpha) \left\| f - \int_{\mathsf{E}} f \,\mathrm{d}\mu \right\|_2^2 \ , \tag{5.9}$$

where $(X_t, V_t)_{t\geqslant 0}$ is a PDMP process corresponding to $\mathcal{L}_i$ for $i = 1, 2$, which is a performance measure of the Monte Carlo estimator of $\int_{\mathsf{E}} f \,\mathrm{d}\mu$. For a class of problems of, say, increasing dimension $d \to \infty$, weak dependence of C and α on d indicates scalability of the method. It is worth pointing out that the result above is equivalent to the existence of $C \geqslant 1$ and $\alpha > 0$ such that for any $\rho_0 \ll \mu$ such that $\|\mathrm{d}\rho_0/\mathrm{d}\mu\|_2 < \infty$

$$\|\rho_0 P_t - \mu\|_{\mathrm{TV}} = \int_{\mathsf{E}} |\mathrm{d}(\rho_0 P_t)/\mathrm{d}\mu - 1| \mathrm{d}\mu \leqslant \|\mathrm{d}(\rho_0 P_t)/\mathrm{d}\mu - 1\|_2 \leqslant C\mathrm{e}^{-\alpha t} \|\mathrm{d}\rho_0/\mathrm{d}\mu - 1\|_2 \ ,$$

where the leftmost inequality is standard and a consequence of the Cauchy-Schwarz inequality. Our hypocoercivity result therefore also allows characterisation of convergence to equilibrium of PDMP in various scenarios and regimes and, for example, that the method is scalable.

Notation

The canonical basis of $\mathbb{R}^d$ is denoted by $(\mathbf{e}_i)_{i\in\{1,\ldots,d\}}$ and the d-dimensional identity matrix I_d. The Euclidean norm on $\mathbb{R}^d$ is denoted by $|\cdot|$. For $f : \mathsf{X} \to \mathbb{R}$ and $i \in \{1, \ldots, d\}$, $x \mapsto \partial_{x_i} f(x)$ stands for the partial derivative of f with respect to the i^{th}-coordinate, if it exists. Similarly, for $f : \mathsf{X} \to \mathbb{R}$, $i, j \in \{1, \ldots, d\}$, denote by $\partial_{x_i,x_j} f = \partial_{x_i}\partial_{x_j} f$ when $\partial_{x_i}\partial_{x_j} f$ exists. For any $k \in \mathbb{N}$, denote by $\mathrm{C}^k(\mathsf{X}, \mathbb{R}^m)$ the set of k-times differentiable functions from X to $\mathbb{R}^m$, $\mathrm{C}^k_{\mathrm{b}}(\mathsf{X}, \mathbb{R}^m)$ stands for the subset of bounded functions in $\mathrm{C}^k(\mathsf{X}, \mathbb{R}^m)$. $\mathrm{C}^k(\mathsf{X})$ and $\mathrm{C}^k_{\mathrm{b}}(\mathsf{X})$ stand for $\mathrm{C}^k(\mathsf{X}, \mathbb{R})$ and $\mathrm{C}^k_{\mathrm{b}}(\mathsf{X}, \mathbb{R})$ respectively. For $f = (f_1, \ldots, f_m) \in \mathrm{C}^1(\mathsf{X}, \mathbb{R}^m)$, $\nabla_x f$ stands for the gradient of f defined for any $x \in \mathsf{X}$ by $\nabla_x f(x) = (\partial_{x_j} f_i(x))_{i\in\{1,\ldots,m\}, j\in\{1,\ldots,d\}} \in \mathbb{R}^{m\times d}$.

For any $f \in \mathrm{C}^k(\mathsf{X})$, $k \in \mathbb{N}$ and $p \geqslant 0$, define

$$\|f\|_{k,p} = \sup_{x \in \mathsf{X}} \sup_{(i_1,\ldots,i_k) \in \{1,\ldots,d\}^k} \left\{ |\partial_{x_{i_1},\ldots,x_{i_k}} f(x)| / (1 + \|x\|^p) \right\} .$$

We set for $k \geqslant 0$,

$$\mathrm{C}^k_{\mathrm{poly}}(\mathsf{X}) = \left\{ f \in \mathrm{C}^k(\mathsf{X}) \, : \, \inf_{p \geqslant 0} \|f\|_{k,p} < +\infty \right\} .$$

Id stands for the identity operator. For two self-adjoint operators $(\mathcal{A}, \mathrm{D}(\mathcal{A}))$ and $(\mathcal{B}, \mathrm{D}(\mathcal{B}))$ on a Hilbert space H equipped with the scalar product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, denote by $\mathcal{A} \succeq \mathcal{B}$ if $\langle f, \mathcal{A} f \rangle \geqslant \langle f, \mathcal{B} f \rangle$ for all $f \in \mathrm{D}(\mathcal{A}) \cap \mathrm{D}(\mathcal{B})$. Then, define $(\mathcal{A}\mathcal{B}, \mathrm{D}(\mathcal{A}\mathcal{B}))$ with domain $\mathrm{D}(\mathcal{A}\mathcal{B}) = \mathrm{D}(\mathcal{B}) \cap \{\mathcal{B}^{-1}\mathrm{D}(\mathcal{A})\}$. For a bounded operator $\mathcal{A}$ on H, we let $|||\mathcal{A}||| = \sup_{f \in \mathsf{H}} \|\mathcal{A} f\| / \|f\|$. Denote by 1_{F} the constant function equals to 1 from a set F to $\mathbb{R}$. For any unbounded operator $(\mathcal{A}, \mathrm{D}(\mathcal{A}))$, we denote by $\mathrm{Ran}(\mathcal{A}) = \{\mathcal{A} f \, : \, f \in \mathrm{D}(\mathcal{A})\}$ and $\mathrm{Ker}(\mathcal{A}) = \{f \in \mathrm{D}(\mathcal{A}) \, : \, \mathcal{A} f = 0\}$. For any probability measure m on a measurable space $(\mathsf{M}, \mathcal{F})$, we denote by $\mathrm{L}^2(\mathrm{m})$ the Hilbert space of measurable functions f satisfying $\int_{\mathsf{M}} f^2 \mathrm{dm} < +\infty$ and $\mathrm{L}^2_0(\mathrm{m}) = \{f \in \mathrm{L}^2(\mathrm{m}) \, : \, \int_{\mathsf{M}} f \mathrm{dm} = 0\}$. For any $x \in \mathsf{M}$ denote by δ_x the Dirac distribution at x. We define the total variation distance between two probability measures $\mathrm{m}_1, \mathrm{m}_2$ on $(\mathsf{M}, \mathcal{F})$ by $\|\mathrm{m}_1 - \mathrm{m}_2\|_{\mathrm{TV}} = \sup_{\mathsf{A} \in \mathcal{F}} |\mathrm{m}_1(\mathsf{A}) - \mathrm{m}_2(\mathsf{A})|$. For a square matrix A we let $\mathrm{diag}(A)$ be its main diagonal and for a vector a we let $\mathrm{diag}(a)$ be the square matrix of diagonal a and zeros elsewhere. For $a, b \in \mathbb{R}$ we let $a \wedge b$ denote their minimum. For $a, b \in \mathbb{R}^d$ ($A, B \in \mathbb{R}^{d \times d}$), we denote by $a \odot b \in \mathbb{R}^d$ ($A \odot B \in \mathbb{R}^{d \times d}$) the Hadamard product between a and b defined for any $i \in \{1, \ldots, d\}$ ($i, j \in \{1, \ldots, d\}$) by $(a \odot b)_i = a_i b_i$ ($(A \odot B)_{i,j} = A_{i,j} B_{i,j}$). For any $i, j \in \mathbb{N}$, $\delta_{i,j}$ denotes the Kronecker symbol which is 1 if $i = j$ and 0 otherwise.

5.2 Main results and organisation of the paper

We now state our main results. In the following, for any closable operator $(\mathcal{C}, \mathrm{D}(\mathcal{C}))$ we let $(\mathcal{C}^*, \mathrm{D}(\mathcal{C}^*))$ denote its $\mathrm{L}^2(\mu)$-adjoint. First we specify conditions satisfied by the potential U.

Assumption 5.2. *The potential $U \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$ and satisfies*

(a) there exists $c_1 \geqslant 0$ such that, for any $x \in \mathsf{X}$,

$$\nabla^2 U(x) \succeq -c_1 \, \mathrm{I}_d ; \tag{5.10}$$

(b)

$$\liminf_{|x| \to \infty} \left\{ |\nabla U(x)|^2 / 2 - \Delta U(x) \right\} > 0 . \tag{5.11}$$

From [129, 9], A5.2-(b) is equivalent to assuming that π satisfies a Poincaré inequality on X, that is the existence of $C_{\mathrm{P}} > 0$ such that, for any $f \in \mathrm{L}^2_0(\pi)$,

$$\|\nabla_x f\|_2 \geqslant C_{\mathrm{P}} \|f\|_2 . \tag{5.12}$$

Further, A5.2-(b) also implies the existence of $c_2 > 0$ such that for any $x \in \mathsf{X}$,

$$\Delta U(x) \leqslant dc_2 + |\nabla U(x)|^2/2 \; . \tag{5.13}$$

A5.2-(b) indeed implies that the quantity considered is bounded from below, the scaling in d in front of c_2 will appear natural in the following. We have opted for this formulation of the assumption required of the potential to favour intuition and link it to the necessary and sufficient condition for geometric convergence of Langevin diffusions, but our quantitative bounds below will be given in terms of the Poincaré constant C_{P} for simplicity (see [11, Section 4.2] for quantitative estimates of C_{P} depending on potentially further conditions on U). A5.2-(a) is realistic in most applications, can be checked in practice and has the advantage of leading to simplified developments. It is possible to replace this assumption with $\sup_{x\in\mathsf{X}}\{|\nabla_x^2 U(x)|/(1+|\nabla_x U(x)||)\} < \infty$ and rephrase our results in terms of any finite upper bound of this quantity (see [44, Sections 2 and 3]). Finally the Poincaré inequality (5.12) implies by [11, Proposition 4.4.2] that there exists $s > 0$ such that

$$\int_{\mathbb{R}^d} \mathrm{e}^{s|x|} \, \mathrm{d}\pi(x) < +\infty \; . \tag{5.14}$$

Assumption 5.3. *The family of vector fields $\{F_k : \mathsf{X} \to \mathbb{R}^d \, ; \, k \in \{0, \dots, K\}\}$ satisfies*

(a) for $k \in \{0, \dots, K\}$, $F_k \in \mathrm{C}^2(\mathsf{X}, \mathbb{R}^d)$;

(b) for all $x \in \mathsf{X}$,

$$\nabla U(x) = \sum_{k=0}^{K} F_k(x) \; ; \tag{5.15}$$

(c) for all $k \in \{0, \dots, K\}$ there exists $a_k \geqslant 0$ such that for all $x \in \mathsf{X}$,

$$|F_k|(x) \leqslant a_k \left(1 + |\nabla U|(x)\right) \; . \tag{5.16}$$

This assumption is in particular trivially true for the Zig-Zag and the Bouncy Particle Samplers.

Assumption 5.4. *Assume that V and ν satisfy the following conditions.*

(a) V is stable under bounces, i.e. for all $(x, v) \in \mathsf{E}$ and $k \in \{1, \dots, d\}$, $v - 2(v^\top \mathrm{n}_k(x)) \, \mathrm{n}_k(x) \in \mathsf{V}$;

(b) for any $\mathsf{A} \in \mathcal{V}$, $x \in \mathsf{X}$, we have $\nu\left(\left\{\mathrm{Id} - 2\mathrm{n}_k(x)\mathrm{n}_k(x)^\top\right\} \mathsf{A}\right) = \nu(\mathsf{A})$;

(c) for any bounded and measurable function $g : \mathbb{R}^2 \to \mathbb{R}$, $i, j \in \{1, \dots, d\}$ such that $i \neq j$, $\int_{\mathsf{V}} g(v_i, v_j) \, \mathrm{d}\nu(v) = \int_{\mathsf{V}} g(v_1, v_2) \, \mathrm{d}\nu(v)$;

(d) ν has finite fourth order marginal moment

$$m_4 = (1/3) \left\| v_1^2 \right\|_2^2 = (1/3) \int_{\mathsf{V}} v_1^4 \, \mathrm{d}\nu(v) < +\infty \; , \tag{5.17}$$

and for any $i,j,k,l \in \{1,\dots,d\}$ *such that* $\mathrm{card}(\{i,j,k,l\}) > 2$, $\int_{\mathsf{V}} v_i v_j v_k v_l \,\mathrm{d}\nu(v) = 0$.

Note that in the case where V and ν are rotation invariant, *i.e.* for any rotation O on $\mathbb{R}^d$, $O\mathsf{V} = \mathsf{V}$ and for any $\mathsf{A} \in \mathcal{V}$, $\nu(O\mathsf{A}) = \nu(\mathsf{A})$, then A5.4-(a)-(b)-(c) are automatically satisfied. Under A5.4-(d), from the Cauchy-Schwarz inequality, we obtain that

$$m_{2,2} = \|v_1 v_2\|_2^2 = \int_{\mathsf{V}} v_1^2 v_2^2 \,\mathrm{d}\nu(v) < \infty \;, \tag{5.18}$$

and note that in the Gaussian case we have the relation $m_4 = m_{2,2} = m_2^2$.

In this paper we consider operators $(\mathcal{R}_v, \mathrm{D}(\mathcal{R}_v))$ on $\mathrm{L}^2(\mu)$ satisfying the following conditions.

Assumption 5.5. *The refresh operator* $\mathcal{R}_v$ *satisfies the following conditions.*

(a) Functions depending only on the position belong to the kernel of $\mathcal{R}_v$*:* $\mathrm{L}^2(\pi) \subset \mathrm{D}(\mathcal{R}_v)$ *and for any* $f \in \mathrm{L}^2(\pi)$, $\mathcal{R}_v f = 0$.

(b) $\mathcal{R}_v$ *satisfies the detailed balance condition:* $\mathcal{R}_v = \mathcal{R}_v^*$ *and* $\mathrm{C}_{\mathrm{b}}^2(\mathsf{E}) \subset \mathrm{D}(\mathcal{R}_v)$*;*

(c) $\mathcal{R}_v$ *admits a spectral gap of size* 1 *on* $\mathrm{L}_0^2(\nu)$*: for any* $g \in \mathrm{L}_0^2(\nu) \cap \mathrm{D}(\mathcal{R}_v)$, $\langle -\mathcal{R}_v g, g\rangle_2 \geqslant \|g\|_2$*; in addition, for any* $f \in \mathrm{L}^2(\pi)$*, it holds* $v_1 f \in \mathrm{D}(\mathcal{R}_v)$ *and* $-\mathcal{R}_v(v_1 f) = v_1 f$.

Typically, $\mathcal{R}_v$ is of the form $\mathrm{Id} \otimes \tilde{\mathcal{R}}_v$ where $(\tilde{\mathcal{R}}_v, \mathrm{D}(\tilde{\mathcal{R}}_v))$ is a self-adjoint operator on $\mathrm{L}^2(\nu)$ with spectral gap equals 1. Then, condition A5.5-(a) is equivalent to $\tilde{\mathcal{R}}_v(1_{\mathsf{V}}) = 0$, which implies that for any $g \in \mathrm{D}(\tilde{\mathcal{R}}_v)$, we have

$$\int_{\mathsf{V}} \mathcal{R}_v g \,\mathrm{d}\nu = \langle 1_{\mathsf{E}}, \mathcal{R}_v g\rangle_2 = \langle \mathcal{R}_v^*(1_{\mathsf{V}}), g\rangle_2 = \langle \mathcal{R}_v(1_{\mathsf{V}}), g\rangle_2 = 0 \;, \tag{5.19}$$

so that the process associated with $\tilde{\mathcal{R}}_v$ preserves the probability measure ν.

Note that A5.5-(a) implies that $\mathcal{R}_v \Pi_v = 0$, whereas A5.5-(c) implies that $-\mathcal{R}_v(v_1 \Pi_v) = v_1 \Pi_v$, where Π_v is defined by (5.6). Assumption A5.5 is satisfied when $\mathcal{R}_v = \Pi_v$, or $\mathcal{R}_v = \mathrm{Id} \otimes \tilde{\mathcal{R}}_v$ with $\tilde{\mathcal{R}}_v$ the generator of the Ornstein-Uhlenbeck process defined for any $g \in \mathrm{C}_{\mathrm{b}}^2(\mathbb{R}^d)$ by

$$\tilde{\mathcal{R}}_v g = -\nabla g^\top v + \Delta g \;.$$

Assumption 5.6. *The refreshment rate* $\lambda_{\mathrm{ref}} : \mathsf{X} \to \mathbb{R}_+$ *is bounded from below and from above as follows: there exists* $\underline{\lambda} > 0$ *and* $c_\lambda \geqslant 0$ *such that for all* $x \in \mathsf{X}$,

$$0 < \underline{\lambda} \leqslant \lambda_{\mathrm{ref}}(x) \leqslant \underline{\lambda}(1 + c_\lambda |\nabla U(x)|) \;. \tag{5.20}$$

Under the previous assumptions we can prove exponential convergence of the semigroup. The proof of the theorem and its corollaries can be found in Subsection 5.3.3.

Theorem 5.1. *Assume that A5.1, A5.2, A5.3, A5.4, A5.5 and A5.6 hold. In addition, assume that* $\mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$ *is a core for* $\mathcal{L}$*. Then there exist* $C > 0$ *and* $\alpha > 0$ *such that, for any* $f \in \mathrm{L}_0^2(\mu)$*, and* $t \in \mathbb{R}_+$,

$$\|P_t f\|_2 \leqslant C \mathrm{e}^{-\alpha t} \|f\|_2 \;. \tag{5.21}$$

The constants C and α are explicitly given in Theorem 5.2, with

$$\lambda_v = \underline{\lambda}\,, \qquad \lambda_x = C_{\mathrm{P}}/(1+C_{\mathrm{P}})\,, \tag{5.22}$$

and

$$R_0 = \frac{\sqrt{2m_{2,2}+3(m_4-m_{2,2})_+}}{m_2}\left\{\frac{2^{3/2}\kappa_1}{\kappa_2}\sum_{k=1}^K a_k + \kappa_1\right\} + \underline{\lambda}\left\{2^{-1/2} + \frac{2^{1/2}c_\lambda\kappa_1}{\kappa_2}\right\}\,, \tag{5.23}$$

where $\kappa_1 = (1+c_1/2)^{1/2}$ and $\kappa_2^{-1} = C_{\mathrm{P}}^{-1}(1+4dc_2+16C_{\mathrm{P}}^2)^{1/2}$.

By Remark 5.1, in the case where C_{P} is fixed, there exist $C_1(C_{\mathrm{P}}), C_2(C_{\mathrm{P}}) > 0$ such that for $R_0 \geqslant 1 \vee (\underline{\lambda}/2)^{1/2}$,

$$\begin{aligned} C_1(C_{\mathrm{P}})\sqrt{\frac{1+\underline{\lambda}R_0^{-2}}{1-\underline{\lambda}R_0^{-2}}} \leqslant C \leqslant C_2(C_{\mathrm{P}})\sqrt{\frac{1+\underline{\lambda}R_0^{-2}}{1-\underline{\lambda}R_0^{-2}}} \\ \underline{\lambda}m_2^{1/2}C_1(C_{\mathrm{P}})R_0^{-2}/(1+R_0^{-2}\underline{\lambda}) \leqslant \alpha \leqslant \underline{\lambda}m_2^{1/2}C_2(C_{\mathrm{P}})R_0^{-2}/(1+R_0^{-2}\underline{\lambda})\,. \end{aligned} \tag{5.24}$$

In the following, we assume that C_{P} does not depend on the dimension and is fixed. By [11, Proposition 5.1.3, Corollary 5.7.2], this condition is satisfied for strongly convex potential U: *i.e.* there exists $m > 0$ such that $\nabla^2 U(x) \succeq m\,\mathrm{I}_d$ for any $x \in \mathbb{R}^d$ which implies that $C_{\mathrm{P}} = m$. We note that recent progress in the precise quantitative estimation of spectral gaps of certain probability measures [19, 21] allow for the strong convexity property to be relaxed to convexity and beyond, leading to quantitative estimates for C_{P}.

First since C_{P} does not depend on the dimension d, then C is always of order $\mathcal{O}(1)$ as $d \to +\infty$ in the case where $R_0 \geqslant 1 \vee (\underline{\lambda}/2)^{1/2}$, which can always be assumed without any loss of generality.

To discuss the dependence of α, given by Theorem 5.1, on the dimension, we need to specify $m_2, m_{2,2}, m_4$ since they depend in some cases on d, while we assume for this discussion that C_{P}, c_1, c_2 and $(a_k)_{k\in\{1,\dots,K\}}$ are fixed. In particular, we impose that $m_2^{-1}\sqrt{2m_{2,2}+3(m_4-m_{2,2})_+}$ can be upper bounded by a constant m_b independent of the dimension d. This condition is satisfies in many cases, for example in the case where ν is the uniform distribution on $\sqrt{d}\mathbb{S}^{d-1}$ or $\{-1,1\}^d$, or ν is a zero-mean Gaussian distribution with covariance matrix $m\,\mathrm{I}_d$, with m independent of d. In fact, by Lemma 5.12, it is satisfied if ν is a spherically symmetric distribution on $\mathbb{R}^d$ corresponding to random variables $V = B^{1/2}W$ for W uniformly distributed on the hypersphere $\sqrt{d}\mathbb{S}^{d-1}$ and B a non-negative random variable independent of W and of first and second order moments γ_1 and γ_2 respectively, and $\gamma_2^{1/2}/\gamma_1$ upper bounded by a constant independent of the dimension.

By (5.23), if $c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b$ are fixed, there exist $C_1^R(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b) > 0$ and $C_2^R(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b) > 0$, independent of $d, \underline{\lambda}$ and c_λ, such that

$$C_1^R(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b)\bar{R}_0 \leqslant R_0 \leqslant C_2^R(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b)\bar{R}_0\,, \tag{5.25}$$

where $\bar{R}_0 = d^{1/2}K + \underline{\lambda}(1 + c_\lambda d^{1/2})$. Therefore by (5.24), there exists a constant $C^\alpha(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b) > 0$, independent of $d, \underline{\lambda}$ and c_λ, such that for d large enough,

$$\alpha > C^\alpha(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b)\, \underline{\lambda}\, m_2^{1/2} [1 \vee (d^{1/2}K) \vee (\underline{\lambda}(1 + c_\lambda d^{1/2}))]^{-2} \,. \tag{5.26}$$

Thus, if $\underline{\lambda}$ and c_λ are fixed, we get that α^{-1} is of order $\mathcal{O}(m_2^{-1/2}(1 + dK^2))$. Note that the spectral gap is expected to be proportional to $m_2^{1/2}$, since if $(X_t, V_t)_{t\geqslant 0}$ is a PDMP with generator of the form (5.2) or (5.4), and with m_2, then $(X_{m^{1/2}t}, m^{1/2}V_{m^{1/2}t})_{t\geqslant 0}$ is a PDMP with generator of the same form with $m_2 = m$. We consider then in the following that $m_2 = 1$ for any d, this property being satisfied for the uniform distribution on the sphere $\sqrt{d}\,\mathbb{S}^{d-1}$ or the d-dimensional zero-mean Gaussian distribution with covariance matrix I_d. We also assume that the refresh rate is bounded, so that $c_\lambda = 0$. We can now specify the conclusion of (5.26).

- For the BPS process, we get

$$\alpha > C^\alpha(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b) \left[(\underline{\lambda} d^{-1}) \wedge \underline{\lambda}^{-1})\right] \,.$$

  This bound scales optimally with d when there exist $C_1^{\underline{\lambda}}, C_2^{\underline{\lambda}} > 0$ such that $C_1^{\underline{\lambda}}\, d^{1/2} \leqslant \underline{\lambda} \leqslant C_2^{\underline{\lambda}}\, d^{1/2}$, and in this case $\alpha > C^\alpha_{\mathrm{BPS}}(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b)\, d^{-1/2}$.

- For the randomized Hamiltonian Monte Carlo, we get

$$\alpha > C^\alpha(C_{\mathrm{P}}, c_1, c_2, \sup_{k\in\{1,\dots,K\}} a_k, m_b) \left[\underline{\lambda} \wedge \underline{\lambda}^{-1}\right] \,,$$

  which suggests that $\underline{\lambda}$ should be fixed independent of d, and in this case we obtain that α^{-1} is bounded by a constant independent of d.

- For the ZZ process, the dependence on d is worse than for the BPS process, which is suboptimal. We refine our results for this particular process in Section 5.4, for which we obtain convergence rates independent of the dimension in some particular cases.

While nonreversibily of the processes considered here may be practically beneficial, it is only recently that the tools allowing our work have been developed [165, 166]. Our method of proof relies on the framework proposed recently in [43, 44, 26] (see also [70, 71, 72]) to study the solutions of the forward Kolmogorov equation associated with the linear kinetic process, but we study the dual backward Kolmogorov equation for a broader class of processes. This, combined with the flexibility of the framework of [44, 26] explains the differing inner product used throughout, which we have found to lead to simpler computations while yielding identical conclusions. The estimate (5.8) (with constant $C = 1$) would follow straightforwardly from a Gronwall argument if the generator $\mathcal{L}$ of the semigroup was coercive on $\mathrm{L}^2(\mu)$ equipped with $\langle\cdot,\cdot\rangle_2$. Unfortunately, the symmetric part of the generator corresponding to a PDMP is degenerate in general, in the sense that it has a nontrivial null space. Hence, the aforementioned coercivity clearly fails to hold. However, it is possible to

equip $\mathrm{L}^2(\mu)$ with an equivalent scalar product derived from $\langle \cdot, \cdot \rangle_2$ with respect to which $\mathcal{L}$ is coercive. The constant α is then given by the coercivity bound, while the constant C can be obtained from estimates relating the two equivalent scalar products.

The paper is organised as follows. In Section 5.3 we outline the main result of [44], providing optimized constants and indicating which assumptions must be checked in every scenario, and prove Theorem 5.1. In Section 5.4 we specialise our results to the case of the Zig-Zag process for which better estimates are possible, leading to attractive scaling properties with dimension d.

5.3 The DMS framework for hypocoercivity

As explained earlier our results rely on the approach proposed by [44], which we summarize and adapt slightly to our choices of Hilbert space and notation–we further provide explicit and precise estimates of the constants involved. We first present abstract results which form the core of all of our proofs and then establish more specific results common to all the processes considered in this paper, implying some of the abstract conditions. More specific results relating to the Zig-Zag process is treated in Section 5.4.

5.3.1 Abstract DMS results

We let $\mathcal{S}$ and $\mathcal{T}$ be the $\mathrm{L}^2(\mu)$-symmetric and $\mathrm{L}^2(\mu)$-skew-symmetric parts of a generator $\mathcal{L}$ satisfying A5.1, that is

$$\mathcal{S} = \frac{\mathcal{L} + \mathcal{L}^*}{2} \quad \text{and} \quad \mathcal{T} = \frac{\mathcal{L} - \mathcal{L}^*}{2} , \tag{5.27}$$

and define the operator $\mathcal{A}$ as follows,

$$\mathcal{A} = \left(m_2 \operatorname{Id} + (\mathcal{T}\Pi_v)^*(\mathcal{T}\Pi_v)\right)^{-1} (-\mathcal{T}\Pi_v)^* , \tag{5.28}$$

where Π_v is given by (5.6) and m_2 by (5.3).

Lemma 5.1. *Assume that $(\mathcal{T}\Pi_v, \mathrm{D}(\mathcal{T}\Pi_v))$ is a densely defined closable operator, then the operators $(m_2 + (\mathcal{T}\Pi_v)^*(\mathcal{T}\Pi_v))^{-1}$ and $\mathcal{A}$ are bounded on $\mathrm{L}^2(\mu)$. In addition $\mathcal{A}$ satisfies*

$$|||\mathcal{A}|||_2 \leqslant 1/(2m_2^{1/2}) , \qquad |||\mathcal{T}\mathcal{A}|||_2 \leqslant 1 . \tag{5.29}$$

Proof. The proof is a direct consequence of Proposition 5.5 and Remark 5.4. □

The main result of [44] can be formulated under the following abstract assumption and the proof of our main theorem relies on sharp estimates of the constants involved.

Assumption 5.7 (DMS abstract conditions)**.** *Assume that there exists a core $\mathsf{C} \subset \mathrm{D}(\mathcal{L})$ for $\mathcal{L}$ such that*

(a) there exists $\lambda_v > 0$ satisfying for any $f \in \mathsf{C}$

$$- \langle \mathcal{S}f, f \rangle_2 \geqslant \lambda_v m_2^{1/2} \left\| (\operatorname{Id} - \Pi_v) f \right\|_2^2 ; \tag{5.30}$$

(b) there exists $\lambda_x \in (0,1)$ *satisfying for any* $f \in \mathsf{C}$

$$-\langle \mathcal{A}\mathcal{T}\Pi_v f, f\rangle_2 \geqslant \lambda_x \left\|\Pi_v f\right\|_2^2 \ ; \tag{5.31}$$

(c) there exist $R_0 \geqslant 0$ *satisfying for any* $f \in \mathsf{C}$

$$|\langle \mathcal{A}\mathcal{T}(I-\Pi_v)f, f\rangle_2 + \langle \mathcal{A}\mathcal{S}f, f\rangle_2| \leqslant R_0 \left\|(\mathrm{Id}-\Pi_v)f\right\|_2 \left\|\Pi_v f\right\|_2 \ ;$$

(d) $\Pi_v \mathcal{T} \Pi_v = 0$ *and* $\mathrm{Ran}(\Pi_v) \subset \mathrm{Ker}(\mathcal{S})$.

Theorem 5.2. *Assume that* $\mathcal{L}$ *satisfies A5.1 and A5.7.*

(a) Then, for any $f \in \mathrm{L}_0^2(\mu)$, $t \in \mathbb{R}_+$ *and* $\epsilon \in (0, \lambda_v^{-1} \wedge \{4\lambda_x/(4\lambda_x + R_0^2)\})$

$$\left\|P_t f\right\|_2 \leqslant C(\epsilon) \mathrm{e}^{-\alpha(\epsilon)t} \left\|f\right\|_2 \ , \tag{5.32}$$

with

$$\alpha(\epsilon) = \lambda_v m_2^{1/2} \frac{\Lambda(\epsilon)}{1+\lambda_v \epsilon} > 0 \quad \textit{and} \quad C(\epsilon) = \sqrt{\frac{1+\lambda_v\epsilon}{1-\lambda_v\epsilon}} \ , \tag{5.33}$$

where

$$\Lambda(\epsilon) = \frac{1-\epsilon(1-\lambda_x) - \sqrt{[1-\epsilon(1-\lambda_x)]^2 - 4\epsilon\lambda_x(1-\epsilon) + \epsilon^2 R_0^2}}{2} \ . \tag{5.34}$$

(b) Further, if $R_0 \geqslant (\lambda_v/2)^{1/2}$ *then* $\alpha \colon (0, 4\lambda_x/(4\lambda_x + R_0^2)) \to \mathbb{R}_+$ *has a unique maximum at* ϵ^* *such that*

$$\alpha(\epsilon_0) < \alpha(\epsilon^*) < 3\alpha(\epsilon_0),$$

and $C(\epsilon_0) < +\infty$ *is well defined, with*

$$\epsilon_0 = \frac{1+\lambda_x - (1-\lambda_x)\sqrt{\frac{{R_0}^2}{{R_0}^2+4\lambda_x}}}{(1+\lambda_x)^2 + {R_0}^2}. \tag{5.35}$$

Remark 5.1. *Note that if* $\lambda_x \in (0,1)$ *is fixed, then there exist* $C_1^{\epsilon_0}(\lambda_x), C_2^{\epsilon_0}(\lambda_x) \in (0,2)$ *such that for* $R_0 \geqslant 1$,

$$C_1^{\epsilon_0}(\lambda_x) R_0^{-2} \leqslant \epsilon_0 \leqslant C_2^{\epsilon_0}(\lambda_x) R_0^{-2} \ . \tag{5.36}$$

Therefore, we get by (5.33)*, that there exist* $C_1^C(\lambda_x), C_2^C(\lambda_x) > 0$ *such that for* $R_0 \geqslant 1 \vee (\lambda_v/2)^{1/2}$,

$$C_1^C(\lambda_x)\sqrt{\frac{1+\lambda_v R_0^{-2}}{1-\lambda_v R_0^{-2}}} \leqslant C(\epsilon_0) \leqslant C_2^C(\lambda_x)\sqrt{\frac{1+\lambda_v R_0^{-2}}{1-\lambda_v R_0^{-2}}} \ .$$

Using that $[1-\epsilon(1-\lambda_x)]^2 - 4\epsilon\lambda_x(1-\epsilon) \geqslant 0$ *and for any* $a \in [0,1]$, $a/2 \leqslant 1-(1-a)^{1/2} \leqslant a$, *there exist* $C_1^\Lambda(\lambda_x), C_2^\Lambda(\lambda_x) > 0$ *such that* $C_1^\Lambda(\lambda_x)\epsilon_0 \leqslant \Lambda(\epsilon_0) \leqslant C_2^\Lambda(\lambda_x)\epsilon_0$ *if* $\epsilon_0 \leqslant \lambda_v^{-1} \wedge \{4\lambda_x/(4\lambda_x + R_0^2)\}$. *As a result, using* (5.36) *again and since* $\epsilon_0 \leqslant \lambda_v^{-1} \wedge \{4\lambda_x/(4\lambda_x +$

$R_0^2)\} \leqslant \{4\lambda_x/(4\lambda_x + R_0^2)\}$, *if* $R_0 \geqslant (\lambda_v/2)^{1/2}$, *there exist* $C_1^\alpha(\lambda_x), C_2^\alpha(\lambda_x) > 0$ *such that for* $R_0 \geqslant 1 \vee (\lambda_v/2)^{1/2}$,

$$\lambda_v m_2^{1/2} C_1^\alpha(\lambda_x) R_0^{-2}/(1 + R_0^{-2}\lambda_v) \leqslant \alpha(\epsilon_0) \leqslant \lambda_v m_2^{1/2} C_2^\alpha(\lambda_x) R_0^{-2}/(1 + R_0^{-2}\lambda_v) \, .$$

The main idea of [44] behind the proof of Theorem 5.2 is the introduction of an equivalent norm for $\varepsilon \in \mathbb{R}_+$ (instead of the $\mathrm{L}^2(\mu)$ norm, which corresponds to $\varepsilon = 0$)

$$\mathscr{H}_\varepsilon(f) = (1/2) \|f\|_2^2 + \varepsilon \langle f, \mathcal{A}f \rangle_2 \, , \tag{5.37}$$

for which $(P_t)_{t\geqslant 0}$ is exponentially contracting. More precisely, [44, Theorem 2] shows that for any $\varepsilon \in (-m_2^{1/2}, m_2^{1/2})$, there exists $\alpha_\varepsilon > 0$ such that for any $f \in \mathrm{L}_0^2(\mu)$, $\mathscr{H}_\varepsilon(P_t f) \leqslant \mathrm{e}^{-\alpha_\varepsilon t} \mathscr{H}_\varepsilon(f)$. Then, the convergence in $\mathrm{L}_0^2(\mu)$ follows by Lemma 5.1 which implies that $\mathscr{H}_\varepsilon$ defines a norm which is equivalent to $\|\cdot\|_2$: for $\varepsilon \in (-m_2^{1/2}, m_2^{1/2})$ and for any $f \in \mathrm{L}^2(\mu)$, it holds

$$(1 - m_2^{-1/2}\varepsilon) \|f\|_2^2 \leqslant 2\mathscr{H}_\varepsilon(f) \leqslant (1 + m_2^{-1/2}\varepsilon) \|f\|_2^2 \, . \tag{5.38}$$

Therefore exponential decay of $t \mapsto \mathcal{H}[f_t]$ is equivalent to that of $t \mapsto \|f_t\|_2^2$, a property exploited in the following proof.

Proof of Theorem 5.2. Let $f \in \mathsf{C}$ and $\int_{\mathsf{E}} f \mathrm{d}\mu = 0$. For ease of notation, set for any $t \geqslant 0$, $f_t = P_t f \in \mathsf{C}$. Using that for any $t \mapsto f_t$ is continuously differentiable on $\mathbb{R}_+$and $\mathrm{d}f_t/\mathrm{d}t = \mathcal{L}f_t = (\mathcal{S} + \mathcal{T})f_t$ for any $t \in \mathbb{R}_+$, we obtain

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}_\varepsilon(f_t) &= -\langle f_t, (\mathcal{S}+\mathcal{T})f_t\rangle_2 - \varepsilon\left[\langle \mathcal{A}(\mathcal{S}+\mathcal{T})f_t, f_t\rangle_2 + \langle \mathcal{A}f_t, (\mathcal{S}+\mathcal{T})f_t\rangle_2\right] \\
&= -\langle f_t, \mathcal{S}f_t\rangle_2 - \varepsilon\left[\langle \mathcal{A}\mathcal{T}(\Pi_v + \mathrm{Id} - \Pi_v)f_t, f_t\rangle_2 + \langle \mathcal{A}\mathcal{S}f_t, f_t\rangle_2 + \langle \mathcal{S}\mathcal{A}f_t, f_t\rangle_2 - \langle \mathcal{T}\mathcal{A}f_t, f_t\rangle_2\right] \\
&= -\langle f_t, \mathcal{S}f_t\rangle_2 - \varepsilon\left[\langle \mathcal{A}\mathcal{T}\Pi_v f_t, f_t\rangle_2 - \langle \mathcal{T}\mathcal{A}f_t, f_t\rangle_2 + \langle \mathcal{A}\mathcal{T}(\mathrm{Id} - \Pi_v)f_t, f_t\rangle_2 + \langle \mathcal{A}\mathcal{S}f_t, f_t\rangle_2\right] \, ,
\end{aligned} \tag{5.39}$$

where we have used that $\mathcal{S}$ and $\mathcal{T}$ are self adjoint and anti-self adjoint respectively, by Lemma 5.1 and $\langle \mathcal{S}\mathcal{A}f_t, f_t\rangle_2 = \langle \mathcal{S}\Pi_v\mathcal{A}f_t, f_t\rangle_2 = 0$ from Lemma 5.9. This, together with A5.7 and Lemma 5.1 imply since $f_t \in \mathsf{C}$ for any $t \in \mathbb{R}_+$,

$$\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}_\varepsilon(f_t) \\
&\quad \geqslant \lambda_v m_2^{1/2} \|(\mathrm{Id} - \Pi_v)f_t\|_2^2 + \varepsilon\lambda_x \|\Pi_v f_t\|_2^2 - \varepsilon \|(\mathrm{Id} - \Pi_v)f_t\|_2^2 - \varepsilon R_0 \|(\mathrm{Id} - \Pi_v)f_t\|_2 \|\Pi_v f_t\|_2 \\
&\quad = \begin{pmatrix} \|\Pi_v f_t\|_2 \\ \|(\mathrm{Id} - \Pi_v)f_t\|_2 \end{pmatrix}^\top \begin{pmatrix} \varepsilon\lambda_x & -\varepsilon R_0/2 \\ -\varepsilon R_0/2 & \lambda_v m_2^{1/2} - \varepsilon \end{pmatrix} \begin{pmatrix} \|\Pi_v f_t\|_2 \\ \|(\mathrm{Id} - \Pi_v)f_t\|_2 \end{pmatrix} \geqslant \Lambda(\varepsilon) \|f_t\|_2^2 \, ,
\end{aligned}$$

where

$$\Lambda(\varepsilon) = \frac{\lambda_v m_2^{1/2} - \varepsilon(1-\lambda_x) - \sqrt{(\lambda_v m_2^{1/2} - \varepsilon(1-\lambda_x))^2 - [4\varepsilon\lambda_x(\lambda_v m_2^{1/2} - \varepsilon) - \varepsilon^2 R_0^2]}}{2} \, .$$

is the smallest eigenvalue of the symmetric matrix, positive for $\varepsilon \leqslant 4\lambda_x\lambda_v m_2^{1/2}/(4\lambda_x + R_0^2)$ from Lemma 5.5 (as $\lambda_x \leqslant 1$ by A5.7-(b)). Using (5.38), we get

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}_\varepsilon(f_t) \geqslant \frac{2\Lambda(\varepsilon)}{1 + m_2^{-1/2}\varepsilon}\mathscr{H}_\varepsilon(f_t) \; .$$

From Grönwall lemma and Lemma 5.1, we obtain for $\varepsilon \leqslant m_2^{1/2} \wedge \{4\lambda_x\lambda_v m_2^{1/2}/(4\lambda_x + R_0^2)\}$,

$$\|f_t\|_2 \leqslant C(\varepsilon)\mathrm{e}^{\alpha(\varepsilon)t}\|f\|_2 \;, \text{ where } \alpha(\varepsilon) = \frac{\Lambda(\varepsilon)}{1 + m_2^{-1/2}\varepsilon} \quad \text{and} \quad C(\varepsilon) = \sqrt{\frac{1 + m_2^{-1/2}\varepsilon}{1 - m_2^{-1/2}\varepsilon}} \; . \tag{5.40}$$

For notational simplicity we let $\epsilon = \varepsilon/(\lambda_v m_2^{1/2})$ and note that for $\epsilon < 4\lambda_x/(4\lambda_x + R_0^2)$, $\alpha(\epsilon) > 0$, where $\epsilon \to \alpha(\epsilon)$ is defined by (5.33) which concludes the proof of (a).

From Proposition 5.4 and associated notation in Section 5.6, $\epsilon \mapsto \alpha(\epsilon)$ has a unique, but intractable, maximum, $\epsilon^* \in (0, 4\lambda_x/(4\lambda_x + R_0^2))$. However from Lemma 5.6 and Proposition 5.4 the unique maximum $\epsilon_0 \in (\epsilon^*, 4\lambda_x/(4\lambda_x + R_0^2))$ of $\epsilon \mapsto \Lambda(\epsilon)$, defined by (5.66), provides us with a tractable proxy such that $\alpha(\epsilon_0) < \alpha(\epsilon^*) < 3\alpha(\epsilon_0)$. In addition, since $\lambda_x \leqslant 1$, for $R_0 \geqslant (\lambda_v/2)^{1/2}$, we get

$$\epsilon_0 < \frac{(1 + \lambda_x)}{(1 + \lambda_x)^2 + \lambda_v/2} \leqslant \lambda_v^{-1} \; ,$$

which implies that $C(\epsilon_0)$ is well defined.

□

5.3.2 DMS for PDMP: generic results

The following provides expressions for $\mathrm{L}^2(\mu)$-symmetric and $\mathrm{L}^2(\mu)$-skew-symmetric parts of $\mathcal{L}$ for all the PDMP processes considered in this paper. We define the directional derivative operator for any

$$f \in \mathrm{D}(\mathcal{D}) = \mathrm{C}_{\mathrm{b}}^{1,\mathsf{X}}(\mathsf{E}) \; , \; \mathcal{D}f(x,v) = v^\top \nabla_x f(x,v) \; , \tag{5.41}$$

where

$$\mathrm{C}_{\mathrm{b}}^{1,\mathsf{X}}(\mathsf{E}) = \{f \in \mathrm{L}^2(\mu) \; : \; \text{for any } v \in \mathsf{V},\, x \mapsto f(x,v) \in \mathrm{C}_{\mathrm{b}}^1(\mathsf{X})\} \; . \tag{5.42}$$

Note that $\mathcal{D}$ is densely defined on $\mathrm{L}^2(\mu)$ and closable.

Proposition 5.1. *Assume that A5.1, A5.2, A5.4, A5.5 and A5.6 hold. Let $\mathcal{L}_i$ for $i \in \{1,2\}$ be defined in* (5.2) *or* (5.4) *with $\mathcal{B}_k$ as in* (5.7)*, its symmetric part $\mathcal{S}_i$ defined by* (5.27)*, and the operator $\mathcal{A}_i$ defined by* (5.28) *relatively to $\mathcal{T}_i$. Then for any $f \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$,*

(a)

$$\mathcal{T}_i f = v^\top \nabla_x f - \delta_{i,2} m_2 F_0^\top \nabla_v f + \frac{1}{2}\sum_{k=1}^K (v^\top F_k)\,(\mathcal{B}_k - \mathrm{Id}) f \;,$$

$$\mathcal{S}_i f = \frac{1}{2}\sum_{k=1}^K |v^\top F_k|\,(\mathcal{B}_k - \mathrm{Id}) f + m_2^{1/2}\lambda_{\mathrm{ref}}\mathcal{R}_v f \;.$$

(b) $\mathcal{T}_i \Pi_v f = \mathcal{D}\Pi_v f$ *and* $\mathcal{T}_i, \mathcal{S}_i$ *satisfy A5.7-(d).*

Note in particular that $\mathcal{S}_1 = \mathcal{S}_2$ and $\mathcal{T}_1\Pi_v = \mathcal{T}_2\Pi_v$.

Proof. (a) follows from Proposition 5.7 and the definitions of $\mathcal{S}$ and $\mathcal{T}$. The first statement of (b) follows from the fact that for any $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$ and $x \in \mathsf{X}$, $v \mapsto \Pi_v f(x,v)$ is a constant function, therefore $\nabla_v \Pi_v f = 0$ and the definition of $(\mathcal{B}_k)_{k\in\{1,\dots,K\}}$ (5.7) which implies that $(\mathcal{B}_k - \mathrm{Id})\Pi_v f = 0$. The second statement of (b) then follows from the first statement and A5.5-(a). □

Establishing A5.7-(a) (referred to as microscopic coercivity in [44]) for the processes considered is fairly straightforward in the present framework.

Proposition 5.2. *Consider the generator* $\mathcal{L}$ *given by* (5.2) *and its symmetric part* $\mathcal{S}$ *given by Proposition 5.1-(a). Assume A5.5 and A5.6 hold. Then A5.7-(a) holds with* $\lambda_v = \underline{\lambda}$.

Proof. From A5.5-(c) and A5.6, it holds

$$-\lambda_{\mathrm{ref}} m_2^{1/2}\mathcal{R}_v \succeq \underline{\lambda} m_2^{1/2}(\mathrm{Id} - \Pi_v) \;. \tag{5.43}$$

In addition, note that for any $f \in \mathrm{L}^2(\mu)$ satisfying $\max_{k\in\{1,\dots,K\}} \left\|v^\top F_k f\right\|_2 < +\infty$, for any $k \in \{1,\dots,K\}$, by the Cauchy-Schwarz inequality and definition of $\mathcal{B}_k$ (5.7), we obtain $\left\langle \left|v^\top F_k\right| \mathcal{B}_k f, f\right\rangle_2 \leqslant \| \left|v^\top F_k\right|^{1/2} f\|_2$, therefore, we get $|v^\top F_k|\,(\mathrm{Id} - \mathcal{B}_k) \succeq 0$ and using Proposition 5.1-(a) it follows that in the sense of symmetric operators

$$-\mathcal{S} = \frac{1}{2}\sum_{k=1}^K |v^\top F_k|(\mathrm{Id} - \mathcal{B}_k) - m_2^{1/2}\lambda_{\mathrm{ref}}\mathcal{R}_v \succeq \underline{\lambda} m_2^{1/2}(\mathrm{Id} - \Pi_v). \tag{5.44}$$

Therefore the microscopic coercivity holds with $\lambda_v \geqslant \underline{\lambda}$. □

The following lemma establishes equivalence between A5.7-(b) and the Poincaré inequality A5.2 , which allows one to refer to the expansive body of literature on the topic and implies dependence on the properties of the potential U only.

Lemma 5.2. *Consider a closed and densely defined generator* $\mathcal{L}$ *and its anti-symmetric part* $\mathcal{T}$ *given by* (5.27). *Assume A5.2-(b) and* $\mathcal{T}\Pi_v = \mathcal{D}\Pi_v$ *on a common core which is dense in* $\mathrm{L}^2(\mu)$. *Then,* $\mathcal{A} = m_2^{-1}(\mathrm{Id} + \nabla_x^* \nabla_x \Pi_v)^{-1}(-\mathcal{D}\Pi_v)^*$ *and for any* $f \in \mathrm{D}(\mathcal{T}\Pi_v)$, *A5.7-(b) i.e.* (5.31) *holds with*

$$\lambda_x = C_{\mathrm{P}}/(1 + C_{\mathrm{P}}) \;. \tag{5.45}$$

Proof. The first statement is a direct consequence of the definition of $\mathcal{D}$ and the assumptions. As regards the second point, note that from the identity: for any $f \in \mathrm{D}(\mathcal{D}\Pi_v)$

$$\int \left(\mathcal{D}\Pi_v f\right)^2 \mathrm{d}\mu = m_2 \left\|\nabla_x \Pi_v f\right\|_2^2 ,$$

and the assumed Poincaré inequality (5.12) we have for any $f \in \mathrm{D}((\mathcal{D}\Pi_v)^* \mathcal{D}\Pi_v)$,

$$\left\langle f, m_2^{-1} (\mathcal{D}\Pi_v)^* \mathcal{D}\Pi_v f \right\rangle_2 = \left\| m_2^{-1/2} \mathcal{D}\Pi_v f \right\|_2^2 = \left\|\nabla_x \Pi_v f\right\|_2^2 \geqslant C_{\mathrm{P}}^2 \left\|\Pi_v f\right\|_2^2 . \tag{5.46}$$

Since $\mathcal{T}\Pi_v = \mathcal{D}\Pi_v$ on a dense subset of $\mathrm{L}^2(\mu)$ and $\mathcal{D}\Pi_v$ is densely defined, we get $(\mathcal{D}\Pi_v)^* = (\mathcal{T}\Pi_v)^*$. This result and (5.46) implies that $\mathrm{Spec}(m_2^{-1}(\mathcal{T}\Pi_v)^*\mathcal{T}\Pi_v) \subseteq [C_{\mathrm{P}}, \infty)$ by [128, Theorem 5.1.9] and [37, Theorem 4.3.1].

Second, note that for any $f \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$,

$$\mathcal{A}\mathcal{T}\Pi_v f = \left(m_2 \,\mathrm{Id} + (\mathcal{D}\Pi_v)^* \mathcal{D}\Pi_v\right)^{-1} (\mathcal{D}\Pi_v)^* \mathcal{D}\Pi_v f = \Phi\left(m_2^{-1} (\mathcal{D}\Pi_v)^* \mathcal{D}\Pi_v\right) f ,$$

where $\Phi(z) = z/(1+z)$. From the spectral mapping theorem [37, Theorem 2.5.1] and the fact that $\Phi \colon [0,\infty) \to [0,1]$ is non-decreasing $\mathrm{Spec}\left(\mathcal{A}\mathcal{T}\Pi_v\right) \subseteq [\Phi(C_{\mathrm{P}}), 1)$. In addition, from the fact that Π_v is a projector, we deduce $\mathcal{A}\mathcal{T}\Pi_v f = \Pi_v \Phi\left(m_2^{-1}(\mathcal{D}\Pi_v)^*\mathcal{D}\Pi_v\right)\Pi_v f$ and therefore, we get for any $f \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$,

$$\langle \Pi_v f, \mathcal{A}\mathcal{T}\Pi_v f \rangle_2 = \left\langle \Pi_v f, \Phi\left(m_2^{-1}(\mathcal{D}\Pi_v)^*\mathcal{D}\Pi_v\right)\Pi_v f \right\rangle_2 \geqslant \frac{C_{\mathrm{P}}}{1+C_{\mathrm{P}}} \left\|\Pi_v f\right\|_2^2 = \lambda_x \left\|\Pi_v f\right\|_2^2 ,$$

which concludes the proof. □

Corollary 5.1. *Consider the generator $\mathcal{L}$ given by* (5.2) *and its anti-symmetric part $\mathcal{T}$ given by Proposition 5.1-(a). Then A5.2 implies A5.7-(b).*

A5.7-(c) is usually a more involved condition to check. For $f \in \mathrm{L}^2(\mu)$ denote by

$$u = m_2^{-1}(\mathrm{Id} + \nabla_x^* \nabla_x)^{-1} g \in \mathrm{H}^2(\pi) , \text{ where } g = \Pi_v f , \tag{5.47}$$

omitting dependence on f for ease of notation. In the scenarios considered here, condition A5.7-(c) relies on estimates of $\|u\|_2$, $\|\nabla_x u\|_2$ and $\|\nabla_x^2 u\|_2$ which are obtained by noticing that by definition u is solution of the following partial differential equation

$$m_2(\mathrm{Id} + \nabla_x^* \nabla_x) u = \Pi_v f . \tag{5.48}$$

In Lemma 5.10 and Lemma 5.11 we show how general, but potentially rough, estimates can be obtained, while in Section 5.4 we show how tighter bounds can be obtained in specific scenarios where we can take advantage of the structure at hand, in particular when interested in the scaling properties of the algorithm with d.

5.3.3 Proof of Theorem 5.1

In this section we prove that A5.7 holds for the dynamics described in Section 5.2 in order to obtain Theorem 5.1 as a consequence of the abstract Theorem 5.2.

Under the assumptions of the theorem, we can set C to be $\mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$. Then, A5.7-(a) follows from Proposition 5.2 with $\lambda_v = \underline{\lambda}$. A5.7-(b) follows from Lemma 5.2 and its corollary, with $\lambda_x = C_{\mathrm{P}}/(1 + C_{\mathrm{P}})$. A5.7-(d) follows from Proposition 5.1. We are left with checking A5.7-(c). By Lemma 5.10-(b), Remark 5.6, Lemma 5.11-(b), we get

$$
\begin{aligned}
&\|\mathcal{A}\mathcal{T}(\mathrm{Id} - \Pi_v) f\|_2 + \|\mathcal{A}\mathcal{S} f\|_2 \\
&\qquad \leqslant \sqrt{2m_{2,2} + 3(m_4 - m_{2,2})_+} \left\{ \|\nabla_x^2 u\|_2 + 2 \sum_{k=1}^K \|F_k^\top \nabla_x u\|_2 \right\} + m_2^{1/2} \|\lambda_{\mathrm{ref}} \nabla_x u\|_2 \\
&\qquad \leqslant \left[\frac{\sqrt{2m_{2,2} + 3(m_4 - m_{2,2})_+}}{m_2} \left\{ \frac{2^{3/2}\kappa_1}{\kappa_2} \sum_{k=1}^K a_k + \kappa_1 \right\} + \frac{\underline{\lambda}}{(2m_2)^{1/2}} \left\{ 1 + \frac{2c_\lambda \kappa_1}{\kappa_2} \right\} \right] \|\Pi_v f\|_2 \ ,
\end{aligned}
$$

where we have used that $\|\nabla_x^2 u\|_2 \leqslant m_2^{-1} \kappa_1 \|\Pi_v f\|_2$ by Proposition 5.6 and Remark 5.8, with κ_1 and κ_2 given in (5.77) and (5.80) respectively. The proof of A5.7-(c) is then completed using Lemma 5.10-(a) and Lemma 5.11-(a).

5.4 The Zig-Zag sampler

In this section, we specify our results in the case of the Zig-Zag sampler for which better estimates can be obtained, leading to better scaling properties with d. The Zig-Zag process corresponds to the instantiation of (5.2) for which $F_0 = 0$, $K = d$, $F_i(x) = \partial_{x_i} U(x) \mathbf{e}_i$ and $\mathrm{n}_i(x) = \mathbf{e}_i$ for $i \in \{1, \ldots, d\}$ and $x \in \mathsf{X}$, $\lambda_{\mathrm{ref}}(x) = \underline{\lambda} > 0$ for any $x \in \mathsf{X}$ (which corresponds to $c_\lambda = 0$ in A5.6) and $\mathcal{R}_v = \Pi_v - \mathrm{Id}$. The corresponding generator takes the simplified form, for $f \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$ and any $(x, v) \in \mathsf{E}$

$$
\mathcal{L} f(x, v) = v^\top \nabla_x f(x) + \sum_{i=1}^d (v_i \partial_{x_i} U(x))_+ \left[f\left(x, (\mathrm{Id} - 2\mathbf{e}_i \mathbf{e}_i^\top) v\right) - f(x, v) \right] + \lambda_{\mathrm{ref}}(x) m_2^{1/2} \mathcal{R}_v f(x, v) \ . \tag{5.49}
$$

In the next two subsections we first consider general velocity distributions and then show how our results can be specialised to the scenario where $\mathsf{V} = \{-1, +1\}^d$.

5.4.1 General velocity distribution

Theorem 5.3. *Consider the Zig-Zag process with generator defined by* (5.49) *with* $\lambda_{\mathrm{ref}} = \underline{\lambda}$ *and* $\mathcal{R}_v = \Pi_v - \mathrm{Id}$. *Assume A5.1, A5.2, A5.3, A5.4, A5.5, A5.6 hold and that there exists* $c_3 \geqslant 0$ *such that for any* $g \in \mathrm{L}^2(\pi)^d$

$$
\left\langle g, \left[\nabla_x^2 U - \mathrm{diag}(\nabla_x^2 U) \right] g \right\rangle_2 \geqslant -c_3 \|g\|_2^2 \ . \tag{5.50}
$$

In addition, assume that $\mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$ *is a core for* $\mathcal{L}$*. Then, Theorem 5.2 holds with* λ_x *as in* (5.45)*,* $\lambda_v = \underline{\lambda}$ *and*

$$R_0 = \frac{3(6m_4)^{1/2}}{m_2}\left((1+c_1/2)^{1/2} + 1 + (c_3/2)^{1/2}\right) + \underline{\lambda}/2^{1/2} \; . \tag{5.51}$$

Remark 5.2. *From A5.2 we have for any* $g \in \mathrm{L}^2(\pi)^d$

$$\left\langle g, \nabla_x^2 U g \right\rangle_2 \geqslant -c_1 \left\| g \right\|_2^2$$

and therefore (5.50) *holds if there exist* $\bar{c}_1 > 0$ *such that for any* $g \in \mathrm{L}^2(\pi)^d$*,*

$$\left\langle g, \mathrm{diag}(\nabla_x^2 U) g \right\rangle_2 \leqslant \bar{c}_1 \left\| g \right\|_2^2 \; ,$$

which is itself implied by $\bar{c}_1 \, \mathrm{Id} \succeq \mathrm{diag}(\nabla_x^2 U(x))$ *for all* $x \in \mathsf{X}$*, since* $\mathrm{diag}(\nabla_x^2 U(x))$ *is symmetric. Note that this is the case when* $|\mathrm{diag}(\nabla_x^2 U(x))| \leqslant \bar{c}_1$ *or* $|\nabla_x^2 U(x)| \leqslant \bar{c}_1$ *for all* $x \in \mathsf{X}$*, for example.*

Proof. We again apply Theorem 5.2. Checking A5.7-(a)-(b) and (d) is identical to the work in the proof of Theorem 5.1 with the constants $\lambda_v = \underline{\lambda}$ and λ_x given by (5.45). We are left with checking A5.7-(c). Let $f \in \mathrm{C}^2_{\mathrm{b}}(\mu)$ and u be defined by (5.85) where $\mathcal{T}$ is defined by (5.27) with respect to $\mathcal{L}$ given by (5.49). By [125, Theorem 2], $u \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$. From Corollary 5.2 $\left\| \nabla_x u \right\|_2 = m^{-1/2} \left\| \mathcal{D} \Pi_v u \right\|_2 \leqslant \left\| \Pi_v f \right\|_2 / (2^{1/2} m_2)$ and $\left\| \nabla_x^* \nabla_x u \right\|_2 \leqslant \left\| \Pi_v f \right\|_2 / m_2$, and from Proposition 5.6 $\left\| \nabla_x^2 u \right\|_2 \leqslant m_2^{-1} [1 + c_1/2]^{1/2} \left\| \Pi_v f \right\|_2$. Therefore, by Lemma 5.10-(a) and Lemma 5.11-(a) and the improved bounds from Lemma 5.3 and Lemma 5.4, we deduce that for any $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$,

$$\begin{aligned}
\left\| \mathcal{A}\mathcal{T}(\mathrm{Id} - \Pi_v) f \right\|_2 + \left\| \mathcal{A}\mathcal{S} f \right\|_2 &\leqslant 3(6m_4)^{1/2} \left(\left\| \nabla_x^2 u \right\|_2 + \left\| \nabla_x^* \nabla_x u \right\|_2 + c_3^{1/2} \left\| \nabla_x u \right\|_2 \right) + \underline{\lambda} m_2 \left\| \nabla_x u \right\|_2 \\
&\leqslant \left\{ \frac{3(6m_4)^{1/2}}{m_2} \left((1 + c_1/2)^{1/2} + 1 + (c_3/2)^{1/2} \right) + \underline{\lambda}/2^{1/2} \right\} \left\| \Pi_v f \right\|_2 \; ,
\end{aligned}$$

and we conclude. □

We discuss in the following the dependence on the dimension of the convergence rate $\alpha(\epsilon_0)$ and the constant $C(\epsilon_0)$ given by Theorem 5.2 based on the constant provided by Theorem 5.3. Similarly to the general case, we need to impose some conditions on m_2, m_4. Here, we assume that $m_4^{1/2}/m_2$ does not depend on d, which holds in the case where ν is the uniform distribution on $\mathsf{V} = \{-1, 1\}^d$ or the d-dimensional zero-mean Gaussian distribution with covariance matrix I_d.

In the case where π is the i.i.d. product of one-dimensional distributions π_i on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ associated with potentials $U_i : \mathbb{R} \to \mathbb{R}$ satisfying A5.2, *i.e.* for any $x \in \mathsf{X}$, $U(x) = \sum_{i=1}^d U_i(x_i)$, then $\nabla^2 U(x) = \mathrm{diag}(\nabla^2 U(x))$ for any $x \in \mathsf{X}$ and therefore (5.50) holds with $c_3 = 0$. Then, the convergence rate $\alpha(\varepsilon_0)$ and the constant $C(\varepsilon_0)$ in Theorem 5.2 does not depend on the dimension but only on the constants c_1, c_2, $\underline{\lambda}$, c_λ and C_{P} associated to each U_i.

Consider now the case where the potential U is strongly convex and gradient Lipschitz, *i.e.* there exist $m, L > 0$ such that $m\,\mathrm{I}_d \preceq \nabla^2 U(x) \preceq L\,\mathrm{I}_d$ for any $x \in \mathsf{X}$. Then, since for any $i \in \{1, \dots, d\}$ and $x \in \mathsf{X}$, $\partial_{x_i, x_i} U(x) = \mathbf{e}_i^\top \nabla^2 U(x) \mathbf{e}_i \leqslant L$ by assumption, Remark 5.2 implies that (5.50) holds for $c_3 = L - m$. In addition, A5.2 holds with $c_1 = 0$ and $c_2 = L$ and by [11, Proposition 5.1.3, Corollary 5.7.2], U satisfies (5.12) with $C_{\mathrm{P}} = m$. Then, the convergence rate $\alpha(\varepsilon_0)$ and the constant $C(\varepsilon_0)$ in Theorem 5.2 do not depend on the dimension but only on L, m, $\underline{\lambda}$ and $\overline{\lambda}$. In addition, we observe that the larger $L - m$ is, the larger R_0 given in (5.51) is, which in turn make the convergence rate $\alpha(\varepsilon_0)$ worse since it is a $\mathcal{O}(1/R_0^2)$ as $R_0 \to +\infty$ by Remark 5.1. This result is expected in the Gaussian case $U(x) = x^\top \Sigma x$ for any $x \in \mathsf{X}$, since $L - m$ is the diameter of the set of eigenvalues of Σ which is a characterization of the conditioning of the problem.

Lemma 5.3. *Consider the Zig-Zag process with generator defined by* (5.49) *with* $\lambda_{\mathrm{ref}} = \underline{\lambda}$ *and* $\mathcal{R}_v = \Pi_v - \mathrm{Id}$. *Assume A5.1, A5.2, A5.3, A5.4, A5.5, A5.6 and* (5.50) *hold. Then for any* $f \in \mathrm{L}^2(\mu)$

$$\left\|\{\mathcal{A}\mathcal{T}(\mathrm{Id} - \Pi_v)\}^* f\right\|_2 \leqslant [6(4m_4 - m_{2,2})]^{1/2} \left(\left\|\nabla_x^2 u\right\|_2 + \left\|\nabla_x^* \nabla_x u\right\|_2 + c_3^{1/2} \left\|\nabla_x u\right\|_2 \right) ,$$

where $\mathcal{A}$ *is defined in* (5.28), u *is defined by* (5.85) *with respect to* $\mathcal{T}$ *and* $\mathcal{L}$ *given in* (5.49) *and* (5.27) *respectively.*

Proof. Note that it is sufficient to show this result for $f \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$. Let $f \in \mathrm{C}_{\mathrm{b}}^2(\mu)$ and u be defined by (5.85) where $\mathcal{T}$ is defined by (5.27) with respect to $\mathcal{L}$ given by (5.49). By [125, Theorem 2], $u \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$. We use Lemma 5.10 and its notations, where $K = d$, for $k \in \{1, \dots, d\}$, $F_k = \partial_{x_k} U \mathbf{e}_k$ and $\mathrm{n}_k = \mathrm{sgn}(\partial_{x_k} U)\mathbf{e}_k$. In this setting and by (5.87), it follows that

$$\mathbf{M}(x) = \nabla_x^2 u(x) + \mathrm{diag}\Big(\nabla_x u \odot \nabla_x U\Big),$$

Since $\|\mathbf{M}\|_2^2 = \|\mathrm{diag}(\mathbf{M})\|_2^2 + \|\mathbf{M} - \mathrm{diag}(\mathbf{M})\|_2^2$, we obtain

$$\begin{aligned} 2m_{2,2} \|\mathbf{M}\|_2^2 + 3(m_4 - m_{2,2}) \|\mathrm{diag}(\mathbf{M})\|_2^2 &= 2m_{2,2} \|\mathbf{M} - \mathrm{diag}(\mathbf{M})\|_2^2 + (3m_4 - m_{2,2}) \|\mathrm{diag}(\mathbf{M})\|_2^2 \\ &\leqslant 2m_{2,2} \left\|\nabla_x^2 u\right\|_2^2 + (3m_4 - m_{2,2}) \|\mathrm{diag}(\mathbf{M})\|_2^2 . \end{aligned} \tag{5.52}$$

We now bound $\|\mathrm{diag}(\mathbf{M})\|_2^2$. First, we apply the triangle inequality and use Lemma 5.7-(a), to deduce that

$$\begin{aligned} \|\mathrm{diag}(\mathbf{M})\|_2^2 = \sum_{k=1}^d \left\|2\partial_{x_k}^2 u - \partial_{x_k}^2 u + \partial_{x_k} U \partial_{x_k} u\right\|_2^2 &\leqslant \sum_{k=1}^d \left(2 \left\|\partial_{x_k}^2 u\right\|_2 + \left\|-\partial_{x_k}^2 u + \partial_{x_k} U \partial_{x_k} u\right\|_2\right)^2 \\ &\leqslant \sum_{k=1}^d \left(8 \left\|\partial_{x_k}^2 u\right\|_2^2 + 2 \left\|\partial_{x_k}^* \partial_{x_k} u\right\|_2^2\right) , \end{aligned} \tag{5.53}$$

where we have used for the last inequality that $(a+b)^2 \leqslant 2a^2 + 2b^2$ for any $a, b \in \mathbb{R}$. By Lemma 5.7-(a), (5.75), (5.14) and the fact that $U \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$ using A5.2, using that same

reasoning as to establish (5.78), it holds for any $k \in \{1, \dots, d\}$,

$$\left\|\partial^*_{x_k}\partial_{x_k}u\right\|_2^2 = \left\|\partial^2_{x_k}u\right\|_2^2 + \langle \partial_{x_k}u, \partial_{x_k,x_k}U\,\partial_{x_k}u\rangle_2 \ , \text{ and } \quad \|\nabla_x^*\nabla_x u\|_2^2 = \left\|\nabla_x^2 u\right\|_2^2 + \left\langle \nabla_x u, \nabla_x^2 U \nabla_x u\right\rangle_2 \ .$$

These identities and the condition (5.50) imply

$$\begin{aligned}\sum_{i=1}^d \left\|\partial^*_{x_i}\partial_{x_i}u\right\|_2^2 &= \left\|\mathrm{diag}\left(\nabla_x^2 u\right)\right\|_2^2 + \left\langle \nabla_x u, \mathrm{diag}\left(\nabla_x^2 U\right)\nabla_x u\right\rangle_2 \leqslant \left\|\nabla_x^2 u\right\|_2^2 + \left\langle \nabla_x u, \mathrm{diag}\left(\nabla_x^2 U\right)\nabla_x u\right\rangle_2 \\ &\leqslant \|\nabla_x^*\nabla_x u\|_2^2 - \left\langle \nabla_x u, \left(\nabla_x^2 U - \mathrm{diag}(\nabla_x^2 U)\right)\nabla_x u\right\rangle_2 \leqslant \|\nabla_x^*\nabla_x u\|_2^2 + c_3\|\nabla_x u\|_2^2 \ . \end{aligned} \tag{5.54}$$

Combining (5.53) and (5.54), we obtain

$$\|\mathrm{diag}(\mathbf{M})\|_2^2 \leqslant 8\sum_{k=1}^d \left\|\partial^2_{x_k}u\right\|_2^2 + 2(\|\nabla_x^*\nabla_x u\|_2^2 + c_3\|\nabla_x u\|_2^2) \ .$$

From this inequality, (5.52) and Lemma 5.10, we deduce

$$\begin{aligned}\|\{\mathcal{A}\mathcal{T}(\mathrm{Id} - \Pi_v)\}^* f\|_2^2 &\leqslant 6(4m_4 - m_{2,2})\left\|\nabla_x^2 u\right\|_2^2 + 2(3m_4 - m_{2,2})\left(\|\nabla_x^*\nabla_x u\|_2^2 + c_3\|\nabla_x u\|_2^2\right) \\ &\leqslant 6(4m_4 - m_{2,2})\left(\left\|\nabla_x^2 u\right\|_2 + \|\nabla_x^*\nabla_x u\|_2 + c_3^{1/2}\|\nabla_x u\|_2\right)^2 \ , \end{aligned}$$

since for $a, b, c \geqslant 0$, $a^2 + b^2 + c^2 \leqslant (a + b + c)^2$. □

Lemma 5.4. *Consider the Zig-Zag process with generator defined by* (5.49) *with* $\lambda_{\mathrm{ref}} = \underline{\lambda}$ *and* $\mathcal{R}_v = \Pi_v - \mathrm{Id}$. *Assume A5.1, A5.2, A5.3, A5.4, A5.5, A5.6 and* (5.50) *hold. Then for any* $f \in \mathrm{L}^2(\mu)$

$$\|\left\{\mathcal{A}\mathcal{S}(\mathrm{Id} - \Pi_v)\right\}^* f\|_2 \leqslant (6m_4)^{1/2}\left(\left\|\nabla_x^2 u\right\|_2 + \|\nabla_x^*\nabla_x u\|_2 + c_3^{1/2}\|\nabla_x u\|_2\right) + \underline{\lambda} m_2\|\nabla_x u\|_2 \ . \tag{5.55}$$

where $\mathcal{S}, \mathcal{A}$ *are defined in* (5.27)- (5.28), u *is defined by* (5.85) *with respect to* $\mathcal{T}$ *and* $\mathcal{L}$ *given in* (5.49) *and* (5.27) *respectively.*

Proof. Note that it is sufficient to show this result for $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$. Let $f \in \mathrm{C}^2_{\mathrm{b}}(\mu)$ and u be defined by (5.85) where $\mathcal{T}$ is defined by (5.27) with respect to $\mathcal{L}$ given by (5.49). By [125, Theorem 2], $u \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$. We use Lemma 5.11 and its notation, where $K = d$, for $k \in \{1, \dots, d\}$, $F_k = \partial_{x_k}U$ and $\mathrm{n}_k = \mathrm{sgn}(\partial_{x_k}U)\mathbf{e}_k$. In this setting and by (5.92), it follows that for any $(x, v) \in \mathsf{E}$,

$$\mathbf{G}(x, v) = \sum_{k=1}^d \mathrm{sgn}(v_k)v_k^2\partial_{x_k}U(x)\mathbf{e}_k + \underline{\lambda}m_2^{1/2}v \ .$$

From the triangle inequality and since for $i, j \in \{1, \dots, d\}$, $\int_{\mathsf{V}} \mathrm{sgn}(v_i v_j)v_i^2 v_j^2 \mathrm{d}\nu(v) = m_4\delta_{i,j}$,

$$\left\|\mathbf{G}^\top\nabla_x u\right\|_2 \leqslant \sqrt{3m_4\textstyle\sum_{k=1}^d \|\partial_{x_k}U\partial_{x_k}u\|_2^2} + \underline{\lambda}m_2\|\nabla_x u\|_2 \ . \tag{5.56}$$

To bound the sum we note that for $k \in \{1, \dots, d\}$ $\partial_{x_k} U \partial_{x_k} u = \partial^2_{x_k} u + \partial^*_{x_k} \partial_{x_k} u$ by Lemma 5.7-(a), which together with the fact $(a+b)^2 \leqslant 2(a^2+b^2)$ leads to

$$\|\partial_{x_i} U \partial_{x_i} u\|_2^2 \leqslant 2\Big(\left\|\partial^2_{x_i} u\right\|_2^2 + \left\|\partial^*_{x_i} \partial_{x_i} u\right\|_2^2 \Big) .$$

Then, using that for $a, b \geqslant 0$ $\sqrt{a+b} \leqslant \sqrt{a} + \sqrt{b}$ twice and (5.54), we deduce

$$\begin{aligned}
\left(\sum_{k=1}^d \|\partial_{x_k} U \partial_{x_k} u\|_2^2 \right)^{1/2} &\leqslant 2^{1/2} \left\{ \sum_{k=1}^d \left(\left\|\partial^2_{x_k} u\right\|_2^2 + \left\|\partial^*_{x_k} \partial_{x_k} u\right\|_2^2 \right) \right\}^{1/2} \\
&\leqslant 2^{1/2} \left\{ \left(\sum_{k=1}^d \left\|\partial^2_{x_k} u\right\|_2^2 \right)^{1/2} + \left(\sum_{k=1}^d \left\|\partial^*_{x_k} \partial_{x_k} u\right\|_2^2 \right)^{1/2} \right\} \\
&\leqslant 2^{1/2} \left(\left\|\nabla_x^2 u\right\|_2 + \|\nabla_x^* \nabla_x u\|_2 + c_3^{1/2} \|\nabla_x u\|_2 \right) . \qquad (5.57)
\end{aligned}$$

Then combining (5.56) and (5.57) completes the proof by Lemma 5.11.

□

5.4.2 d-dimensional Radmacher distribution

We now consider the case $\mathsf{V} = \{-1, +1\}^d$ and ν is the uniform distribution on V which corresponds to the original setting of the Zig-Zag process. This process has been proved to be ergodic [18] even in the absence of refreshment, that is $\lambda_{\mathrm{ref}} = 0$. We note that in this scenario $m_2 = 1$, $m_4 = \frac{1}{3}$ and $m_{2,2} = 1$ which leads to simplified expressions for the bounds in Lemma 5.3 and Lemma 5.11 upon revisiting their proofs. However this has no qualitative impact. In this section we show that hypocoercivity holds with our techniques for $\lambda_{\mathrm{ref}}(x) = 0$ for "most of X" for a particular type of partial refreshment update.

Consider the scenario where $\mathcal{R}_v$ is a mixture of the bounces $\{\mathcal{B}_k, k = 1, \dots, d\}$, for any $f \in \mathrm{L}^2(\mu)$, $(x, v) \in \mathsf{E}$,

$$\lambda_{\mathrm{ref}} \mathcal{R}_v f(x, v) = \sum_{k=1}^d \lambda_{\mathrm{ref},k}(x) \Big[f\Big(x, v - 2v_i \mathbf{e}_i\Big) - f(x, v) \Big] , \qquad (5.58)$$

with $\lambda_{\mathrm{ref},k} \colon \mathsf{X} \to \mathbb{R}_+$ for $k \in \{1, \dots, d\}$ satisfying A5.6, and $\lambda_{\mathrm{ref}} = \sum_{k=1}^d \lambda_{\mathrm{ref},k}$, that is when the process refreshes, $k \in \{1, \dots, d\}$ is chosen at random with probability proportional to $(\lambda_{\mathrm{ref},1}, \dots, \lambda_{\mathrm{ref},d})$ and the component v_k of v is updated to $-v_k$.

Proposition 5.3. *Assume A5.2. Consider the Zig-Zag process with generator and refreshment operator as in* (5.49) *and* (5.58) *respectively, with* $\lambda_{\mathrm{ref},k} \colon \mathsf{X} \to \mathbb{R}_+$ *for* $k \in \{1, \dots, d\}$ *satisfying A5.6. Then,*

(a) the symmetric part of the generator is given for any $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$, $(x, v) \in \mathsf{E}$ *by*

$$\mathcal{S} f(x, v) = \sum_{k=1}^d \left(\frac{1}{2} |\partial_{x_k} U|(x) + \lambda_{\mathrm{ref},k}(x) \right) \Big[f\Big(x, v - 2v_i \mathbf{e}_i\Big) - f(x, v) \Big] ; \qquad (5.59)$$

(b) the microscopic coercivity condition A5.7-(a) is satisfied, i.e. for any $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$, $(x,v) \in \mathsf{E}$

$$-\langle \mathcal{S} f, f\rangle_2 \geqslant \lambda_v \left\|(\mathrm{Id} - \Pi_v) f\right\|_2^2 \quad \textit{with} \quad \lambda_v = \min_{k\in\{1,\dots,d\}, x\in\mathsf{X}} \frac{|\partial_{x_k} U(x)|}{2} + \lambda_{\mathrm{ref},k}(x) \ . \tag{5.60}$$

Remark 5.3. *In other words A5.7-(a) holds if for any* $\varepsilon > 0$, *for all* $k \in \{1,\dots,d\}$, $\lambda_{\mathrm{ref},k}$ *vanishes everywhere, except on* $\{x \in \mathsf{X} : \exists k \in \{1,\dots,d\} \mid |\partial_{x_k} U|(x) < \varepsilon\}$. *We also note that a similar result holds for the case where* $\mathcal{R}_v = \Pi_v - \mathrm{Id}$, *that is A5.7-(a) holds whenever* λ_{ref} *vanishes everywhere, except on* $\{x \in \mathsf{X} : \exists k \in \{1,\dots,d\}, \mid |\partial_{x_k} U|(x) < \varepsilon\}$ *for some* $\varepsilon > 0$.

Proof. The first statement is a direct application of Proposition 5.1-(a). For the second statement, using that ν is the uniform distribution on $\mathsf{V} = \{-1,1\}^d$, from the polarization identity, we get for any $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$,

$$\begin{aligned} -\langle \mathcal{S} f, f\rangle_2 &= \frac{1}{4} \int_{\mathsf{E}} \sum_{k=1}^d \left(|\partial_{x_k} U|(x) + 2\lambda_{\mathrm{ref},k}(x,v)\right) \left[f(x,v) - f\Big(x, (\mathrm{Id} - 2\mathbf{e}_k \mathbf{e}_k^\top) v\Big)\right]^2 \mathrm{d}\mu(x,v) \\ &\geqslant (\lambda_v/2) \int_{\mathsf{E}} \sum_{k=1}^d \left[f(x,v) - f\Big(x, (\mathrm{Id} - 2\mathbf{e}_k \mathbf{e}_k^\top) v\Big)\right]^2 \mathrm{d}\mu(x,v) \ , \end{aligned} \tag{5.61}$$

where λ_v is defined in (5.60). Now by the Poincaré inequality for any $g \in \mathrm{L}^2_0(\nu)$, see e.g. [123, p. 52], it holds that

$$(1/2) \int_{\mathsf{V}} \sum_{k=1}^d \left[g(v) - g\Big((\mathrm{Id} - 2\mathbf{e}_i \mathbf{e}_i^\top) v\Big)\right]^2 \mathrm{d}\nu(v) \geqslant \int_{\mathsf{V}} \sum_{k=1}^d g^2(v) \, \mathrm{d}\nu(v) \ . \tag{5.62}$$

Now since for any $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$, $\langle \mathcal{S} f, f\rangle_2 = \langle \mathcal{S}(\mathrm{Id} - \Pi_v) f, (\mathrm{Id} - \Pi_v) f\rangle_2$ and for any $x \in \mathsf{X}$, $v \mapsto (\mathrm{Id} - \Pi_v) f(x,v) \in \mathrm{L}^2_0(\nu)$, then combining (5.61) and (5.62) and using Fubini's theorem concludes the proof of (5.60). □

5.5 Discussion and link to earlier work

As pointed out earlier the scenario $K = 0$ where $F_0 = \nabla_x U$ is considered in [44] where the authors establish hypercoercivity but also in [24] where the authors establish V-geometric convergence, that is the existence of constants $C \geqslant 0$, $\alpha > 0$ and a Lyapunov function $V \colon \mathsf{E} \to \mathbb{R}_+$ [24, Theorem 3.9] such that for any $f \colon \mathsf{E} \to \mathbb{R}$ satisfying $\int_{\mathsf{E}} f \, \mathrm{d}\mu = 0$ and $|f|_\infty := \sup_{(x,v)\in\mathsf{E}} |f(x,v)| < +\infty$ then for any $(x,v) \in \mathsf{E}$ and $t \geqslant 0$

$$|P_t f(x,v)| \leqslant C V(x,v) \mathrm{e}^{-\alpha t} |f|_V \ . \tag{5.63}$$

Similar results have been obtained in [39] and [46] for the Bouncy Particle Sampler and in [18] for the Zig-Zag process. All these methods rely on guessing such a suitable Lyapounov

function V and establishing a so-called drift condition for this function, in conjunction with a minorization condition [113].

The existence of an $\mathrm{L}^2(\mu)$ spectral gap (which corresponds to $C = 1$ in (5.21)) always implies V-geometric ergodicity but the latter does not, in general, imply the former, except for reversible processes [92]. To our knowledge it is unclear when hypocoercivity and V-geometric convergence are equivalent, if at all. We note that our results do not allow for the initial probability distribution ρ_0 to be a delta Dirac mass. However an advantage of our approach is that it provides explicit and relatively simple bounds in terms of interpretable quantities which, we show, are informative, which is in contrast with those on minorization and drift conditions in most scenarios. One exception is the study of BPS on the torus carried out in [46] for $U = 0$, using an appropriate coupling argument, which leads to a rate of convergence for the total variation distance with a favourable $\mathcal{O}(d^{1/2})$ scaling. Further we note that if a suitable Lyapunov function can be identified and an associated drift condition found then our results automatically imply V-geometric ergodicity, but with our bounds on the spectral gap . Although we have shown that for the Zig-Zag sampler with Rademacher distribution λ_{ref} is not required to be bounded away from zero on X, the results of [18] hold with $\lambda_{\mathrm{ref}} = 0$. It would be interesting to further investigate whether our results can be specialized to consider the scenario $\lambda_{\mathrm{ref}} = 0$. The hypocoercivity approach introduced in [44] which we follow here has been extended to the case of heavy tailed distribution satisfying a weak Poincaré inequality [33]. Results on the algebraic decay of the semi-group can certainly be extended to PDMP dynamics.

Although we have shown that the theory developed in this paper covers numerous scenarios in a unified set-up, various possible extensions are possible. For example we have restricted this first investigation to deterministic bounces of the type given in (5.7), but there does not seem to be any obstacle to the extension of our results to the more general set-ups such as considered in [162, 170, 115]. In the same vein, great parts of our calculations could be used to consider distributions of the velocity ν that are neither Gaussian, nor the uniform distribution on the hypersphere. For ν of density proportional to $\exp(-K(v))$ with $K\colon \to \mathbb{R}$ the Liouville operator involved in the definition of (5.4) would take the form $\nabla_v K(v)^\top \nabla_x f(x,v) - m_2 F_0^\top \nabla_v f(x,v)$, leading to a different expression for $\mathcal{T}$. Such modified kinetic energies have been proposed to speed up the computation, introducing the Modified Langevin Dynamics for which convergence to equilibrium has been studied in [135].

5.6 Optimization of the rate of convergence $\alpha(\epsilon)$

We let

$$R(\epsilon) = [1-\epsilon(1-\lambda_x)]^2 - 4\epsilon\lambda_x(1-\epsilon) + \epsilon^2 R_0^2 = R_1^2\left(\epsilon - \frac{1+\lambda_x}{R_1^2}\right)^2 + 1 - \frac{(1+\lambda_x)^2}{R_1^2} > 0\ , \quad (5.64)$$

with

$$R_1^2 = (1+\lambda_x)^2 + R_0^2\ , \qquad \tilde{\alpha}(\epsilon) = \frac{\Lambda(\epsilon)}{1+\lambda_v\epsilon}\ , \quad (5.65)$$

where Λ is defined by (5.34). We show that optimizing $\epsilon \mapsto \Lambda(\epsilon)$ is a good enough proxy for optimizing $\epsilon \mapsto \tilde{\alpha}(\epsilon)$, whose maximum is unique, but intractable. Since $\epsilon \mapsto \alpha(\epsilon)$ defined by (5.33) is proportional to $\epsilon \mapsto \tilde{\alpha}(\epsilon)$, the same conclusion holds for this function.

Lemma 5.5. *Let $\Lambda\colon \mathbb{R} \to \mathbb{R}$ be defined by (5.34). Then with $\lambda_x \in (0,1)$,*

(a) $\Lambda(\epsilon) \geqslant 0$ for $0 \leqslant \epsilon \leqslant 4\lambda_x/(4\lambda_x + R_0^2)$ and $\Lambda(0) = 0$.

(b) Λ has first order derivative

$$\Lambda'(\epsilon) = -\frac{1}{2}\Big[(1-\lambda_x)R(\epsilon)^{1/2} + \epsilon R_1^2 - (1+\lambda_x)\Big]R(\epsilon)^{-1/2}\,,$$

and $\Lambda_0'(0) = \lambda_x > 0$.

(c) $\Lambda\colon \mathbb{R}_+ \to \mathbb{R}$ has a unique critical point ($\Lambda'(\epsilon_0) = 0$)

$$\epsilon_0 = \frac{(1+\lambda_x) - (1-\lambda_x)\sqrt{\frac{R_0^2}{R_0^2+4\lambda_x}}}{(1+\lambda_x)^2 + R_0^2} > 0\,, \tag{5.66}$$

such that $\Lambda(\epsilon_0) > 0$.

Proof. From (5.34) we see that $\Lambda(\epsilon) \geqslant 0$ requires

$$\epsilon \leqslant \frac{1}{1-\lambda_x} \wedge \frac{4\lambda_x}{4\lambda_x + R_0^2} = \frac{4\lambda_x}{4\lambda_x + R_0^2}\,,$$

where the equality follows from $\lambda_x > 0$, which completes the proof of (a). The proof of (b) is a simple calculation and is omitted. We now show (c). If we set $\Lambda'(\epsilon) = 0$, it implies that $\epsilon > 0$ satisfies

$$(1+\lambda_x) - \epsilon R_1^2 = R(\epsilon)^{1/2}(1-\lambda_x),$$

and imposes the condition

$$0 \leqslant \epsilon \leqslant \frac{1+\lambda_x}{(1+\lambda_x)^2 + R_0^2}\ . \tag{5.67}$$

Squaring both sides of the equality above implies the following sequence of equalities

$$(1-\lambda_x)^2 R(\epsilon) = \Big[\epsilon R_1^2 - (1+\lambda_x)\Big]^2\,,$$
$$(1-\lambda_x)^2\Big[R_1^2\epsilon^2 - 2(1+\lambda_x)\epsilon + 1\Big] = R_1^4\epsilon^2 - 2R_1^2(1+\lambda_x)\epsilon + (1+\lambda_x)^2,$$

that is

$$\begin{aligned}\epsilon^2 R_1^2\left[(1-\lambda_x)^2 - R_1^2\right] - 2\epsilon(1+\lambda_x)\left[(1-\lambda_x)^2 - R_1^2\right] - 4\lambda_x &= 0\\ \epsilon^2 R_1^2\left[-4\lambda_x - R_0^2\right] - 2\epsilon(1+\lambda_x)\left[-4\lambda_x - R_0^2\right] - 4\lambda_x &= 0\\ R_1^2\epsilon^2 - 2(1+\lambda_x)\epsilon + \frac{4\lambda_x}{R_0^2+4\lambda_x} &= 0.\end{aligned}$$

The two strictly positive roots are

$$\epsilon_\pm = \frac{(1+\lambda_x) \pm \sqrt{(1+\lambda_x)^2 - 4\lambda_x \frac{(1+\lambda_x)^2+R_0^2}{R_0^2+4\lambda_x}}}{(1+\lambda_x)^2 + R_0^2} > 0,$$

where the inequality follows from $\lambda_x > 0$. Further

$$(1+\lambda_x)^2\Big(R_0^2 + 4\lambda_x\Big) - 4\lambda_x\Big[(1+\lambda_x)^2 + R_0^2\Big] = R_0^2\Big[(1+\lambda_x)^2 - 4\lambda_x\Big]\ ,$$

and since $\lambda_x \leqslant 1$, this yields the simplified expression for the two roots

$$\epsilon_\pm = \frac{(1+\lambda_x) \pm (1-\lambda_x)\sqrt{\frac{R_0^2}{R_0^2+4\lambda_x}}}{(1+\lambda_x)^2 + R_0^2}\ .$$

From the conditions on ϵ given by (a) and (5.67), and the fact that $\lambda_x \leqslant 1$, we retain $\epsilon_0 = \epsilon_-$ only. The last statement follows from the second statement and the fact that Λ' is continuous. □

The following lemma establishes in particular that ϵ_0 is a global maximum.

Lemma 5.6. $\Lambda\colon \mathbb{R}_+ \to \mathbb{R}$

(a) is such that $\epsilon \mapsto \Lambda''(\epsilon) \leqslant 0$ *(implying concavity),*

(b) is maximized at ϵ_0 *defined by* (5.66) *and*

$$\epsilon_0 \leqslant \frac{4\lambda_x}{4\lambda_x + R_0^2}\ .$$

Proof. We differentiate $\epsilon \mapsto -2\Lambda(\epsilon) = -[1-\epsilon(1-\lambda_x)] + \sqrt{R(\epsilon)}$ twice, yielding the first order derivative

$$(1-\lambda_x) + \frac{1}{2}R'(\epsilon)R(\epsilon)^{-1/2}$$

and the second order derivative follows

$$\frac{1}{2}\left(R''(\epsilon)R(\epsilon)^{-1/2} - \frac{1}{2}R'(\epsilon)^2 R(\epsilon)^{-3/2}\right) = \frac{1}{4}R(\epsilon)^{-3/2}\left(2R''(\epsilon)R(\epsilon) - R'(\epsilon)^2\right)\ .$$

Now from (5.64), $R(\epsilon) = aR_0(\epsilon)$ with $R_0(\epsilon) = (\epsilon - b)^2 + c$ with all constants non-negative. Further $R_0'(\epsilon) = 2(\epsilon - b)$ and $R_0''(\epsilon) = 2$ and therefore

$$2R_0''(\epsilon)R_0(\epsilon) - R_0'(\epsilon)^2 = 4[(\epsilon - b)^2 + c - (\epsilon - b)^2] = 4c \geqslant 0.$$

From the concavity we deduce that ϵ_0 is a maximum, and the inequality on ϵ_0 follows from the fact that this is required for $\Lambda(\epsilon_0) \geqslant 0$. □

Proposition 5.4. *The function $\tilde{\alpha}\colon \mathbb{R}_+ \to \mathbb{R}_+$, defined by* (5.65)*, has a unique maximizer $0 < \epsilon^* < \epsilon_0$ and if $R_0 \geqslant \lambda_v/2$ then*

$$\tilde{\alpha}(\epsilon_0) \leqslant \tilde{\alpha}(\epsilon^*) \leqslant 3\tilde{\alpha}(\epsilon_0) \ .$$

Proof. First note that

$$\tilde{\alpha}'(\epsilon) = \frac{\Psi(\epsilon)}{(1 + \lambda_v\epsilon)^2} \ ,$$

with

$$\Psi(\epsilon) = \Lambda'(\epsilon)(1 + \lambda_v\epsilon) - \lambda_v\Lambda(\epsilon) \ .$$

Then from Lemma 5.6

$$\Psi'(\epsilon) = (1 + \lambda_v\epsilon)\Lambda''(\epsilon) < 0 \ .$$

Together with $\Psi(0) = \Lambda'(0) = \lambda_x > 0$, $\Psi(\epsilon_0) = -\lambda_v\Lambda(\epsilon_0) < 0$ and the fact that $\epsilon \to \Psi'(\epsilon)$ is continuous, we deduce the existence and uniqueness of $\epsilon^* \in (0, \epsilon_0)$ satisfying $\tilde{\alpha}'(\epsilon^*) = 0$, and maximizing $\tilde{\alpha}$ on $\mathbb{R}_+$. Further since $\tilde{\alpha}'(\epsilon^*) = 0$ and $\epsilon \mapsto \Psi(\epsilon)$ is non-increasing we deduce

$$\sup_{\epsilon \in [\epsilon^*, \epsilon_0]} |\tilde{\alpha}'(\epsilon)| \leqslant \frac{|\Psi(\epsilon_0)|}{(1 + \lambda_v\epsilon^*)^2} = \lambda_v \frac{1 + \lambda_v\epsilon_0}{(1 + \lambda_v\epsilon^*)^2}\tilde{\alpha}(\epsilon_0) \ ,$$

and from classical calculus

$$\tilde{\alpha}(\epsilon^*) - \tilde{\alpha}(\epsilon_0) \leqslant (\epsilon_0 - \epsilon^*)\lambda_v \frac{1 + \lambda_v\epsilon_0}{(1 + \lambda_v\epsilon^*)^2}\tilde{\alpha}(\epsilon_0) \ ,$$

from which we conclude that

$$\tilde{\alpha}(\epsilon_0) \leqslant \tilde{\alpha}(\epsilon^*) \leqslant \left[1 + (\epsilon_0 - \epsilon^*)\lambda_v \frac{1 + \lambda_v\epsilon_0}{(1 + \lambda_v\epsilon^*)^2}\right] \tilde{\alpha}(\epsilon_0) \ .$$

Now if we use $R_0^2 \geqslant \lambda_v^2/4$ we have by (5.66) that

$$\lambda_v\epsilon_0 < \frac{(1 + \lambda_x)\lambda_v}{(1 + \lambda_x)^2 + \lambda_v^2/4} \leqslant 1 \ ,$$

implying

$$(\epsilon_0 - \epsilon^*)\lambda_v \frac{1 + \lambda_v\epsilon_0}{(1 + \lambda_v\epsilon^*)^2} \leqslant \lambda_v\epsilon_0(1 + \lambda_v\epsilon_0) \leqslant 2 \ ,$$

so we have

$$\tilde{\alpha}(\epsilon_0) \leqslant \tilde{\alpha}(\epsilon^*) \leqslant 3\tilde{\alpha}(\epsilon_0) \; .$$

□

5.7 Elliptic regularity estimates

5.7.1 Proof of Lemma 5.1 and more

In this section we gather classical results concerning densely defined closed operators on a Hilbert space to which we repeatedly refer throughout the manuscript.

Proposition 5.5. *Let $\mathcal{B}$ be a closed and densely defined operator on a Hilbert space H of inner product $\langle \cdot, \cdot \rangle$, induced norm $\|\cdot\|$ and operator norm $|||\cdot|||$.*

(a) $\mathrm{Id} + \mathcal{B}^*\mathcal{B}$ *is a positive self-adjoint operator on* H *bijective from* $\mathrm{D}(\mathcal{B}^*\mathcal{B})$ *to* H. *In addition,* $(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}$ *is a positive self-adjoint bounded operator on* H *and* $\mathcal{B}(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}$ *is a bounded operator.*

(b) For any $h \in \mathsf{H}$,

$$\|(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1} h\|^2 + 2\,\|\mathcal{B}(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1} h\|^2 \leqslant \|h\|^2 \; .$$

(c) $\mathcal{B}^*\mathcal{B}(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}$ *is a bounded operator on* $\mathcal{H}$ *which satisfies* $|||\mathcal{B}^*\mathcal{B}(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}||| \leqslant 1$.

Remark 5.4. *Note that under the condition of Proposition 5.5, we get that* $(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}\mathcal{B}^*$ *can be extended to a bounded operator and*

$$|||(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}||| \leqslant 1 \; , \quad |||\mathcal{B}(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}||| = |||(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}\mathcal{B}^*||| \leqslant 1/2^{1/2} \; . \tag{5.68}$$

Proof. (a) and (b) follow from [128, Theorem 5.1.9] and inspection of the proof. It remains to show (c). First note that $(\mathrm{Id} + \mathcal{B}^*\mathcal{B} - \mathrm{Id})(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1} = \mathrm{Id} - (\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}$, from which we deduce that it is a self-adjoint and bounded operator by the triangle inequality with norm less or equal than 2. To prove the tighter upper bound we use [128, Proposition 3.2.27 p. 99] (twice), the identity for any $h \in \mathsf{H}$

$$\left| \left\langle \mathcal{B}^*\mathcal{B}(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1} h, h \right\rangle \right| = \max \left\{ \|h\|^2 - \left\langle (\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1} h, h \right\rangle , \left\langle (\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1} h, h \right\rangle - \|h\|^2 \right\} \; ,$$

that $(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}$ is positive and $|||(\mathrm{Id} + \mathcal{B}^*\mathcal{B})^{-1}||| \leqslant 1$ from the first statement. □

The operator ∇_x on $\mathrm{L}^2(\mu)$ can be extended as an operator on $\mathrm{L}^2(\mu)^d$ as follows: for any $(f_1, \ldots, f_d) \in \mathrm{L}^2(\mu)^d$, $f_1 \in \mathrm{D}(\nabla_x)$, $\nabla_x f = \nabla_x f_1$. Therefore a direct consequence of Proposition 5.5 applied to the operator $m^{-1/2}\nabla_x$ for $m > 0$, on $\mathrm{L}^2(\mu)^d$ is the following.

Corollary 5.2. *Let $m > 0$. The operators $\nabla_x(m\,\mathrm{Id} + \nabla_x^*\nabla_x)^{-1}$ and $\nabla_x^*\nabla_x(m\,\mathrm{Id} + \nabla_x^*\nabla_x)^{-1}$ are bounded on $\mathrm{L}^2(\mu)^d$ with*

$$\left|\left|\left| \nabla_x(m\,\mathrm{Id} + \nabla_x^*\nabla_x)^{-1} \right|\right|\right|_2 \leqslant 1/(2m)^{1/2} \; , \; \left|\left|\left| \nabla_x^*\nabla_x(m\,\mathrm{Id} + \nabla_x^*\nabla_x)^{-1} \right|\right|\right|_2 \leqslant 1 \; . \tag{5.69}$$

In addition, for any $f \in \mathrm{L}^2(\mu)$,

$$\left\|(m\,\mathrm{Id} + \nabla_x^* \nabla_x)^{-1} f\right\|_2^2 + (2/m) \left\|\nabla_x (m\,\mathrm{Id} + \nabla_x^* \nabla_x)^{-1} f\right\|_2^2 \leqslant \left\{\|f\|_2 / m\right\}^2 , \tag{5.70}$$

and

$$\left\|\nabla_x^* \nabla_x (m\,\mathrm{Id} + \nabla_x^* \nabla_x)^{-1} f\right\|_2 \leqslant \|f\|_2 \ . \tag{5.71}$$

5.7.2 Improved Poincaré inequalities

We preface this section with some complements on the adjoint of ∇_x seen as an operator on $\mathrm{L}^2(\mu)^p$ for $p \in \mathbb{N}^*$.

Lemma 5.7. *Consider the operator* ∇_x *from the Hilbert space* $\mathrm{L}^2(\mu)$ *to* $\mathrm{L}^2(\mu)^d$ *endowed with the inner product defined by* (5.5). *Then it holds*

(a) for any $i \in \{1, \ldots, d\}$, *the* $\mathrm{L}^2(\mu)$*-adjoint of* ∂_{x_i} *is given for any* $f \in \mathrm{C}_{\mathrm{b}}^1(\mathsf{E})$ *by*

$$\partial_{x_i}^* f = -\partial_{x_i} f + f \partial_{x_i} U \ ; \tag{5.72}$$

(b) the $\mathrm{L}^2(\mu)$*-adjoint of* ∇_x *is given for any* $G \in \mathrm{C}_{\mathrm{b}}^1(\mathsf{E}, \mathbb{R}^d)$ *by*

$$\nabla_x^* G = -\operatorname{div}_x G + \nabla_x U^\top G \ ; \tag{5.73}$$

(c) if ν *is the zero-mean Gaussian distribution on* $\mathbb{R}^d$ *with covariance matrix* $m_2\,\mathrm{I}_d$, *the* $\mathrm{L}^2(\mu)$*-adjoint of* ∇_v *is*

$$\nabla_v^* G = -\operatorname{div}_v G + m_2^{-1} v^\top G \ . \tag{5.74}$$

Remark 5.5. *Note that Lemma 5.7 implies that for any* $g \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E})$ *and* $G \in \mathrm{C}_{\mathrm{b}}^2(\mathsf{E}, \mathbb{R}^d)$, *we have*

$$\nabla_x^* \nabla_x g = -\Delta_x g + \nabla_x U^\top \nabla_x g \text{ and } \nabla_x \nabla_x^* G = \nabla_x^* \nabla_x G + \nabla_x^2 U G \ , \tag{5.75}$$

where we have defined $\nabla_x^* \nabla_x G \in \mathrm{C}_{\mathrm{b}}(\mathsf{E}, \mathbb{R}^d)$ *for any* $(x, v) \in \mathsf{E}$ *and* $i \in \{1, \ldots, d\}$ *by*

$$\{\nabla_x^* \nabla_x G(x, v)\}_i = \nabla_x^* \partial_{x_i} G(x, v) = \sum_{j=1}^d -\partial_{x_j, x_i} G_j(x, v) + \partial_{x_j} U(x) \partial_{x_i} G(x, v) \ . \tag{5.76}$$

Proof. (a) and (b) follow from integration by parts whereas (c) is a consequence of the first point. □

Proposition 5.6. *Let* $m > 0$ *and assume A5.2. Then for any* $f \in \mathrm{L}^2(\mu)$,

$$\|\nabla_x^2 (m\,\mathrm{Id} + \nabla_x^* \nabla_x)^{-1} \Pi_v f\|_2 \leqslant \kappa_1 \|\Pi_v f\|_2 \quad \text{where} \quad \kappa_1 = (1 + c_1/(2m))^{1/2} \ . \tag{5.77}$$

Proof. Setting $g = \Pi_v f \in \mathrm{L}^2(\pi)$, it is enough to show that for any $g \in \mathrm{L}^2(\pi)$, we have

$$\|\nabla_x^2 (m\,\mathrm{Id} + \nabla_x^* \nabla_x)^{-1} g\|_2 \leqslant \kappa_1 \|g\|_2 \ .$$

In addition, by density, we only need to deal with $g \in \mathrm{C}^\infty_{\mathrm{c}}(\mathsf{X})$. Let $g \in \mathrm{C}^\infty_{\mathrm{c}}(\mathsf{X})$ and consider $u = (m\,\mathrm{Id} + \nabla_x^* \nabla_x)^{-1} g$. By [125, Theorem 2], $u \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$. Therefore we obtain by (5.75), (5.14) and the fact that $U \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$ using A5.2,

$$\begin{aligned}\|\nabla_x^2 u\|_2^2 = \langle \nabla_x^2 u, \nabla_x^2 u\rangle_2 = \langle \nabla_x u, (\nabla_x^* \nabla_x)[\nabla_x u]\rangle_2 = \langle \nabla_x u, (\nabla_x \nabla_x^*)[\nabla_x u] - \nabla_x^2 U \nabla_x u\rangle_2 \\ = \|\nabla_x^* \nabla_x u\|_2^2 - \langle \nabla_x u, \nabla_x^2 U \nabla_x u\rangle_2 \,. \quad (5.78)\end{aligned}$$

From the definition of u, using Corollary 5.2 and A5.2-(a) we conclude that

$$\|\nabla_x^2 u\|_2^2 \leqslant \|g\|_2^2 + c_1 \|\nabla_x u\|_2^2 \leqslant \|g\|_2^2 + c_1 \|g\|_2^2/(2m) \,.$$

□

In order to bound terms of the form $\|F_k^\top \nabla_x u\|$ in Section 5.8 we need the following Lemma which is a quantitative version of [44, Lemma 6]. Consider the function $W : \mathbb{R}^d \to \mathbb{R}_+$ defined for any $x \in \mathbb{R}^d$ by

$$W(x) = \left\{1 + |\nabla_x U(x)|^2\right\}^{1/2} \,. \quad (5.79)$$

Lemma 5.8 ([44, Lemma 6])**.** *Assume A5.2. Then for any $\varphi \in \mathrm{H}^1(\pi)$,*

$$\|\nabla_x \varphi\|_2 \geqslant \left[4\left(1 + dc_2/(4C_{\mathrm{P}}^2)\right)^{1/2}\right]^{-1} \|\varphi \nabla_x U\|_2 \,,$$

where c_2 and C_{P} are defined in (5.13) *and* (5.12) *respectively. As a corollary, it holds for any $\varphi \in \mathrm{H}^1(\pi)$,*

$$\|\nabla_x \varphi\|_2 \geqslant \kappa_2 \|\varphi W\|_2 \,, \text{where}\, \kappa_2^{-1} = \left(C_{\mathrm{P}}^{-2} + 16(1 + dc_2/(4C_{\mathrm{P}}^2))\right)^{1/2} = C_{\mathrm{P}}^{-1}\left(1 + 4dc_2 + 16C_{\mathrm{P}}^2\right)^{1/2} \geqslant C_{\mathrm{P}}^{-1} \,. \quad (5.80)$$

Proof. Note that we only need to consider $\varphi \in \mathrm{C}^\infty_{\mathrm{c}}(\mathsf{X})$. First since $\nabla_x U \in \mathrm{L}^2(\pi)$, for any $\varepsilon > 0$, we get

$$2\langle \varphi \nabla_x U, \nabla_x \varphi\rangle_2 \leqslant \varepsilon^{-1} \|\nabla_x \varphi\|_2^2 + \varepsilon \|\varphi \nabla_x U\|_2^2 \,. \quad (5.81)$$

We then bound from below the left-hand side. Using the *carré du champ* identity, *i.e.* for any $f, g \in \mathrm{C}^2_{\mathrm{poly}}(\mathsf{X})$, $\langle \nabla_x f, \nabla_x g\rangle_2 = \left\langle \nabla_x U^\top \nabla_x f - \Delta f, g\right\rangle_2$, we get using that $\nabla_x[\varphi^2] = 2\varphi \nabla_x \varphi$,

$$2\langle \varphi \nabla_x U, \nabla_x \varphi\rangle_2 = \left\langle \nabla_x[\varphi^2], \nabla_x U\right\rangle_2 = \|\varphi \nabla_x U\|_2^2 - \left\langle \varphi^2, \Delta U\right\rangle_2 \,.$$

By (5.13) and (5.12), we obtain

$$2\langle \varphi \nabla_x U, \nabla_x \varphi\rangle_2 \geqslant \|\varphi \nabla_x U\|_2^2/2 - dc_2 \|\varphi\|_2^2 \geqslant \|\varphi \nabla_x U\|_2^2/2 - (dc_2/C_{\mathrm{P}}^2) \|\nabla_x \varphi\|_2^2 \,.$$

From this result and (5.81), it follows that

$$\|\varphi \nabla_x U\|_2^2/2 - (dc_2/C_{\mathrm{P}}^2) \|\nabla_x \varphi\|_2^2 \leqslant \varepsilon^{-1} \|\nabla_x \varphi\|_2^2 + \varepsilon \|\varphi \nabla_x U\|_2^2 \,.$$

Rearranging terms and setting $\varepsilon = 1/4$ completes the proof. The last statement is a direct consequence of the first one using the definition of W in (5.79). □

Putting this with Proposition 5.6, this implies the following.

Corollary 5.3. *Let $m > 0$ and assume A5.2 and A5.3. For any $f \in \mathrm{L}^2(\mu)$ and $k \in \{1, \dots, K\}$, we have*

$$\left\| F_k^\top \{\nabla_x (m \operatorname{Id} + \nabla_x^* \nabla_x)^{-1} \Pi_v f\} \right\|_2 \leqslant 2^{1/2} a_k \left\| W \{\nabla_x (m \operatorname{Id} + \nabla_x^* \nabla_x)^{-1} \Pi_v f\} \right\|_2 \leqslant \frac{2^{1/2} a_k \kappa_1}{\kappa_2} \left\| \Pi_v f \right\|_2 , \tag{5.82}$$

where a_k, W, κ_1 and κ_2 are defined by (5.16), (5.79), (5.77) and (5.80) respectively.

Proof. Note first that it is sufficient to show this result for $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$. Let $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$ and $u = (m + \nabla_x^* \nabla_x)^{-1} \Pi_v f$. By [125, Theorem 2], $u \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$ and therefore $u \in \mathrm{H}^2(\pi)$. Second since for any $t, s \geqslant 0$, $s + t \leqslant 2^{1/2} \sqrt{s^2 + t^2}$, A5.3-(c) implies for any $x \in \mathsf{X}$,

$$|F_k|(x) \leqslant a_k (1 + |\nabla_x U|(x)) \leqslant 2^{1/2} a_k W(x) .$$

Therefore using Lemma 5.8 and Proposition 5.6 successively, we obtain

$$\begin{aligned} \left\| F_k^\top \nabla_x u \right\|_2 &\leqslant \| \, |F_k| \, \nabla_x u \, \|_2 \leqslant 2^{1/2} a_k \left\| W \nabla_x u \right\|_2 = 2^{1/2} a_k \left(\sum_{i=1}^d \left\| W \partial_{x_i} u \right\|_2^2 \right)^{1/2} \\ &\leqslant (2^{1/2} a_k / \kappa_2) \left(\sum_{i=1}^d \left\| \nabla_x \left[\partial_{x_i} u \right] \right\|_2^2 \right)^{1/2} = (2^{1/2} a_k / \kappa_2) \left\| \nabla_x^2 u \right\|_2 \leqslant (2^{1/2} a_k \kappa_1 / \kappa_2) \left\| \Pi_v f \right\|_2 . \end{aligned}$$

□

5.8 Computation of R_0

We first establish general results used throughout the paper.

Proposition 5.7. *Assume that A5.1, A5.2, A5.3, A5.4, A5.5 and A5.6 hold. Then the $\mathrm{L}^2(\mu)$-adjoint of $\mathcal{L}_i$ for $i \in \{1, 2\}$ defined by (5.2) or (5.4) (with $\mathcal{B}_k$ as in (5.7)) is given for any $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$ by*

$$\mathcal{L}_i^* f = -v^\top \nabla_x f + \delta_{i,2} m_2 F_0^\top \nabla_v f + \sum_{k=1}^K (-v^\top F_k)_+ [(\mathcal{B}_k - \operatorname{Id}) f] + m_2^{1/2} \lambda_{\mathrm{ref}} \mathcal{R}_v f .$$

Proof. We only consider the case $i = 2$ since the proof for $i = 1$ follows the same lines. In addition, since $\mathcal{R}_v$ is self-adjoint by assumption, we can consider the case $\lambda_{\mathrm{ref}}(x) = 0$ for any $x \in \mathsf{X}$. It can be easily checked that for any $k \in \{1, \dots, K\}$, $\mathcal{B}_k$ is $\mathrm{L}^2(\mu)$-self-adjoint and further satisfies $\mathcal{B}_k (v^\top F_k)_+ = \left(-v^\top F_k \right)_+$. Based on (5.2)-(5.4), this result, $(a)_+ = (-a)_- + a$

for any $a \in \mathbb{R}$ and Lemma 5.7, for any $f, g \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{E})$, we obtain

$$\begin{aligned}\langle g, \mathcal{L}f\rangle_2 &= \left\langle -v^\top \nabla_x g + (v^\top \nabla U) g + m_2 F_0^\top \nabla_v g - (v^\top F_0) g + \sum_{k=1}^K (\mathcal{B}_k - \mathrm{Id})[(v^\top F_k)_+ g], f \right\rangle_2 \\ &= \left\langle -v^\top \nabla_x g + [v^\top (\nabla U - F_0)] g + m_2 F_0^\top \nabla_v g + \sum_{k=1}^K \{(-v^\top F_k)_+ \mathcal{B}_k g - (v^\top F_k)_+ g\}, f \right\rangle_2 \\ &= \left\langle -v^\top \nabla_x g + [v^\top (\nabla U - F_0)] g + m_2 F_0^\top \nabla_v g + \sum_{k=1}^K \{(-v^\top F_k)_+ (\mathcal{B}_k - \mathrm{Id}) g - (v^\top F_k) g\}, f \right\rangle_2 .\end{aligned}$$

Using that $\sum_{k=0}^K F_k = \nabla U$ by A5.3-(b) concludes the proof. □

Corollary 5.4. *Note that $\mathcal{L}^* \mathbf{1} = 0$, which implies that μ is an invariant probability measure.*

Lemma 5.9. *Assume that $\mathcal{L}$ satisfies A5.1and A5.7-(d). Then,*

$$\mathcal{T}\Pi_v = (\mathrm{Id} - \Pi_v)\mathcal{T}\Pi_v , \qquad \mathcal{S} = (\mathrm{Id} - \Pi_v)\mathcal{S}(\mathrm{Id} - \Pi_v) , \tag{5.83}$$

$$\Pi_v \mathcal{A} = \mathcal{A} \quad \text{and} \qquad \mathcal{A} = \Pi_v \mathcal{A}(\mathrm{Id} - \Pi_v) , \tag{5.84}$$

where $\mathcal{S}, \mathcal{T}, \mathcal{A}$ and Π_v are defined by (5.27), (5.28) *and* (5.6) *respectively.*

Proof. The first equality is straightforward by A5.7-(d). For the second statement it suffices to show that $\mathcal{S}\Pi_v = 0$ and $\Pi_v \mathcal{S} = 0$. The first equality follows from A5.7-(d) and it remains to show the second equality. Using that $\mathcal{S}$ is self-adjoint, we have $\mathrm{Ker}(\mathcal{S}) = \mathrm{Ran}(\mathcal{S})^\perp$ and by A5.7-(d) $\mathrm{Ran}(\Pi_v) \subset \mathrm{Ran}(\mathcal{S})^\perp$. Therefore, using that Π_v is an orthogonal projection in $\mathrm{L}^2(\mu)$, for any $f \in \mathrm{D}(\mathcal{S})$, we have $\|\Pi_v \mathcal{S} f\|_2 = \langle \Pi_v \mathcal{S} f, \mathcal{S} f\rangle_2 = 0$. The third follows from $(m_2 \mathrm{Id} + (\mathcal{T}\Pi_v)^*(\mathcal{T}\Pi_v))\, \Pi_v \mathcal{A} = (m_2 \mathrm{Id} + (\mathcal{T}\Pi_v)^*(\mathcal{T}\Pi_v))\, \mathcal{A}$ and Lemma 5.1. The last statement follows from the third and first and the definition of $\mathcal{A}$. □

For any $f \in \mathrm{L}^2(\mu)$, consider u_f defined by

$$u_f = (m_2 \mathrm{Id} + (\mathcal{T}_i \Pi_v)^*(\mathcal{T}_i \Pi_v))^{-1} \Pi_v f = m_2^{-1} \left\{\mathrm{Id} + \nabla_x^* \nabla_x\right\}^{-1} \Pi_v f , \tag{5.85}$$

where $\mathcal{T}_i$ is defined by (5.27) relatively to the generator $\mathcal{L}_i$, for $i \in \{1, 2\}$ defined by (5.2) or (5.4), but $\mathcal{T}_i \Pi_v$ does not depend on $i = 1, 2$. Note that we used Lemma 5.2 and Proposition 5.1. To alleviate notation and whenever confusion is not possible, we may use u instead of u_f.

Lemma 5.10. *Assume A5.1, A5.2, A5.3, A5.4, A5.5 and A5.6 hold. Consider $\mathcal{L}_i$ for $i \in \{1, 2\}$ defined by* (5.2) *or* (5.4)*, its anti-symmetric part $\mathcal{T}_i$ defined by* (5.27)*, and the operator $\mathcal{A}_i$ defined by* (5.28) *relatively to $\mathcal{T}_i$.*

(a) For any $f \in \mathrm{D}(\mathcal{T}(\mathrm{Id} - \Pi_v))$, we get

$$|\langle \mathcal{A}_i \mathcal{T}_i (\mathrm{Id} - \Pi_v) f, f\rangle_2| \leqslant \|(\mathrm{Id} - \Pi_v) f\|_2 \left\| \left\{\mathcal{A}_i \mathcal{T}_i (\mathrm{Id} - \Pi_v)\right\}^* f \right\|_2 .$$

(b) For any $f \in \mathrm{L}^2(\mu)$

$$\|\Big\{\mathcal{A}_i\mathcal{T}_i(\mathrm{Id}-\Pi_v)\Big\}^* f\|_2^2 = 2m_{2,2}\|\mathbf{M}\|_2^2 + 3(m_4 - m_{2,2})\|\mathrm{diag}(\mathbf{M})\|_2^2 \,, \tag{5.86}$$

with

$$\mathbf{M} = \nabla_x^2 u + \sum_{k=1}^K (F_k^\top \nabla_x u)\mathrm{n}_k\mathrm{n}_k^\top \,, \tag{5.87}$$

and u *defined by* (5.85).

Remark 5.6. *A general, but potentially rough, bound on the right hand side of* (5.86) *can be obtained as follows. From the fact that* $\|\mathrm{diag}(\mathbf{M})\|_2 \leqslant \|\mathbf{M}\|_2$*, it holds that*

$$\|\Big\{\mathcal{A}\mathcal{T}(\mathrm{Id}-\Pi_v)\Big\}^* f\|_2 \leqslant \sqrt{2m_{2,2} + 3(m_4 - m_{2,2})_+}\|\mathbf{M}\|_2$$

where from the triangle inequality and the property $|\mathrm{n}_k(x)\mathrm{n}_k(x)^\top| = 1$

$$\|\mathbf{M}\|_2 \leqslant \|\nabla_x^2 u\|_2 + \sum_{k=1}^K \|F_k^\top \nabla_x u\|_2 \,.$$

Remark 5.7. *Specific scenarios lead to simplifications of these bounds and the bounds in Lemma 5.4:*

(a) from Lemma 5.12 for radial distributions $m_4 = m_{2,2}$ *leading to a simplification of this bound,*

(b) further if ν *is the centred normal distribution of covariance* $m_2\,\mathrm{I}_d$*, then* $m_{2,2} = m_2^2$*, leading to further simplifications,*

(c) if $K = 0$*, and hence* $F_0 = \nabla_x U$*, the scenario considered by [44], then one finds that the bound depends on* $\|\nabla_x^2 u\|_2$ *only.*

Proof. We only consider the case $i = 2$ since the case $i = 1$ is similar by taking $F_0 = 0$. (a) is a direct application of the Cauchy-Schwarz inequality. As for the proof of (b), note first that using that $\mathcal{T}_2$ is anti-symmetric, $\mathcal{A}_2$ is a bounded operator by Lemma 5.1, we have

$$\{\mathcal{A}_2\mathcal{T}_2(\mathrm{Id}-\Pi_v)\}^* = -(\mathrm{Id}-\Pi_v)\mathcal{T}_2\mathcal{A}_2^* \,. \tag{5.88}$$

Now, consider $f \in \mathrm{C}_\mathrm{b}^2(\mathsf{M})$ and u defined by (5.85). By [125, Theorem 2], $u \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$. Therefore, we obtain, using Lemma 5.2, and the fact that $\Pi_v\mathcal{A}_2 = \mathcal{A}_2$ and $\mathcal{T}_2\Pi_v = \mathcal{D}\Pi_v$ on $\mathrm{C}_\mathrm{b}^2(\mathsf{X})$,

$$\mathcal{A}_2^* f = m_2^{-1}(\mathcal{D}\Pi_v)\Big[\mathrm{Id} + \nabla_x^*\nabla_x\Big]^{-1}\Pi_v f = \mathcal{D}\Pi_v u = \mathcal{D}u \,, \tag{5.89}$$

and therefore by (5.88)

$$\{\mathcal{A}_2\mathcal{T}_2(\mathrm{Id}-\Pi_v)\}^* f = -(\mathrm{Id}-\Pi_v)\mathcal{T}_2\mathcal{D}u \,. \tag{5.90}$$

From Proposition 5.1-(a) and (5.7), using that for any $x \in \mathsf{X}$ the mapping $v \mapsto \mathcal{D}u(x,v)$ is linear and $F_k = \mathrm{n}_k |F_k|$ by definition, we deduce that for any $(x,v) \in \mathsf{E}$,

$$\begin{aligned}\mathcal{T}_2\mathcal{D}u(x,v) &= v^\top \nabla_x^2 u(x) v - m_2 F_0^\top(x) \nabla_x u(x) - \sum_{k=1}^K (v^\top F_k) \Big(\mathrm{n}_k(x)\mathrm{n}_k(x)^\top v\Big)^\top \nabla_x u(x) \\ &= v^\top \mathbf{M}(x) v - m_2 F_0^\top(x) \nabla_x u(x) \;,\end{aligned}$$

and as a result

$$(\mathrm{Id} - \Pi_v)\mathcal{T}_2\mathcal{D}u(x,v) = v^\top \mathbf{M}(x) v - m_2 \,\mathrm{Tr}(\mathbf{M}(x)) \;.$$

Combining this result and (5.90), and using Lemma 5.13, we deduce

$$\begin{aligned}\|\{\mathcal{A}_2\mathcal{T}_2(\mathrm{Id} - \Pi_v)\}^* f\|_2^2 &= 2m_{2,2}\|\mathbf{M}\|_2^2 + 3(m_4 - m_{2,2})\|\mathrm{diag}(\mathbf{M})\|_2^2 \\ &\leqslant \Big[2m_{2,2} + 3(m_4 - m_{2,2})_+\Big]\|\mathbf{M}\|_2^2 \;,\end{aligned}$$

which completes the proof. □

Lemma 5.11. *Assume A5.1, A5.2, A5.3, A5.4, A5.5 and A5.6 hold. Consider $\mathcal{L}_i$ for $i \in \{1,2\}$ defined by* (5.2) *or* (5.4)*, its symmetric part $\mathcal{S}_i$ defined by* (5.27)*, and the operator $\mathcal{A}_i$ defined by* (5.28) *relatively to $\mathcal{T}_i$.*

(a) For any $f \in \mathrm{D}(\mathcal{A}_i\mathcal{S}_i(\mathrm{Id} - \Pi_v))$

$$|\,\langle \mathcal{A}_i\mathcal{S}_i(\mathrm{Id} - \Pi_v) f, f\rangle\,| \leqslant \|(\mathrm{Id} - \Pi_v) f\|_2 \Big\|\Big\{\mathcal{A}_i\mathcal{S}_i(\mathrm{Id} - \Pi_v)\Big\}^* f\Big\|_2 \;.$$

(b) For any $f \in \mathrm{L}^2(\mu)$,

$$\Big\|\Big\{\mathcal{A}_i\mathcal{S}_i(\mathrm{Id} - \Pi_v)\Big\}^* f\Big\|_2 = \|\mathbf{G}^\top \nabla_x u\|_2 \;, \tag{5.91}$$

with $\mathbf{G}$ given for any $(x,v) \in \mathsf{E}$ by

$$\mathbf{G}(x,v) = \sum_{k=1}^K \mathrm{sgn}(\mathrm{n}_k^\top(x) v)\Big(\mathrm{n}_k^\top(x) v\Big)^2 F_k + m_2^{1/2} \lambda_{\mathrm{ref}}(x) v \;, \tag{5.92}$$

and u defined by (5.85)*. In addition*

$$\|\mathbf{G}^\top \nabla_x u\|_2 \leqslant m_2 \|\lambda_{\mathrm{ref}} \nabla_x u\|_2 + \sqrt{2m_{2,2} + 3(m_4 - m_{2,2})_+} \sum_{k=1}^K \|F_k^\top \nabla_x u\|_2 \;. \tag{5.93}$$

Proof. We proceed as in the proof of Lemma 5.10 and use Lemma 5.9. We only consider the case $i = 2$ since the case $i = 1$ is obtained by taking $F_0 = 0$. (a) is a direct application of the Cauchy-Schwarz inequality. As for (b), first note that since $\mathcal{S}_2$ is symmetric and $\mathcal{A}_2$ is a bounded operator by Lemma 5.1, we have

$$\Big\{\mathcal{A}_i\mathcal{S}_i(\mathrm{Id} - \Pi_v)\Big\}^* = (\mathrm{Id} - \Pi_v)\mathcal{S}_i\mathcal{A}_i^* \;. \tag{5.94}$$

Now, consider $f \in \mathrm{C}^2_{\mathrm{b}}(\mathsf{M})$ and u defined by (5.85). By [125, Theorem 2], $u \in \mathrm{C}^3_{\mathrm{poly}}(\mathsf{X})$. By (5.89), Proposition 5.1-(a) and A5.4- A5.5 it holds

$$\begin{aligned}\mathcal{S}_2\mathcal{A}_2^* f &= \left(\frac{1}{2}\sum_{k=1}^K |v^\top F_k|(\mathcal{B}_k - \mathrm{Id}) - m_2^{1/2}\lambda_{\mathrm{ref}}\mathcal{R}_v\right) v^\top \nabla_x u \\ &= -\sum_{k=1}^K |v^\top F_k|\,(v^\top \mathrm{n}_k)(\mathrm{n}_k^\top \nabla_x u) - m_2^{1/2}\lambda_{\mathrm{ref}} v^\top \nabla_x u = -\mathbf{G}^\top \nabla_x u\ , \end{aligned} \tag{5.95}$$

where we used that by definition for any $k \in \{1,\dots,K\}$, $F_k = \mathrm{n}_k |F_k|$. Therefore, $\Pi_v \mathcal{S}_2\mathcal{A}_2^* f = 0$ and combining (5.94) and (5.95) completes the proof of (5.91). Finally (5.93) is a direct consequence of the triangle inequality and Lemma 5.13.

□

Remark 5.8. *Combining Corollary 5.2 and Corollary 5.3, by definition of u in* (5.85) *and using A5.6, we obtain that*

$$\sum_{k=1}^K \|F_k^\top \nabla_x u\|_2 \leqslant \frac{2^{1/2}\kappa_1}{m_2\kappa_2}\sum_{k=1}^K a_k \|\Pi_v f\|_2\ , \qquad m_2 \|\lambda_{\mathrm{ref}}\nabla_x u\|_2 \leqslant \underline{\lambda}\left\{2^{-1/2} + \frac{2^{1/2} c_\lambda \kappa_1}{\kappa_2}\right\}\|\Pi_v f\|_2\ . \tag{5.96}$$

5.9 Radial distributions

The following gathers standard results on spherically symmetric distributions on $\mathbb{R}^d$ for which we could not find a single reference. In particular we establish that A5.4-(a) and conditions required in Lemma 5.13 are satisfied in this scenario.

Lemma 5.12. *Let $d \geqslant 2$.*

(a) Assume ν is the uniform distribution on the unit hypersphere $\mathbb{S}^{d-1}$, then

(i) for $i,j,k,l \in \{1,\dots,d\}$ such that $\mathrm{card}(\{i,j,k,l\}) > 2$, we have $\int_{\mathbb{S}^{d-1}} v_i v_j v_k v_l \,\mathrm{d}\nu(v) = 0$,

(ii) otherwise,

$$m_2 = \frac{1}{d}\ , \qquad m_{2,2} = \int_{\mathbb{S}^{d-1}} v_1^2 v_2^2 \,\mathrm{d}\nu(v) = \frac{1}{d(d+2)} \text{ and } \qquad m_4 = \frac{1}{3}\int_{\mathbb{S}^{d-1}} v_1^4 \,\mathrm{d}\nu(v) = \frac{1}{d(d+2)}\ .$$

(b) For any spherically symmetric distribution ν i.e. corresponding to random variables $V = B^{1/2}W$ for W uniformly distributed on the unit hypersphere $\mathbb{S}^{d-1}$ and B a non-negative random variable independent of w and of first and second order moments γ_1 and γ_2 respectively,

(i) for $i,j,k,l \in \{1,\dots,d\}$ such that $\mathrm{card}(\{i,j,k,l\}) > 2$, we have

$$\int_{\mathbb{R}^d} v_i v_j v_k v_l \,\mathrm{d}\nu(v) = 0,$$

(ii) otherwise,

$$m_2 = \frac{\gamma_1}{d}\ , \qquad m_{2,2} = \frac{\gamma_2}{d(d+2)} \text{ and } \qquad m_4 = \frac{\gamma_2}{d(d+2)}\ .$$

Remark 5.9. *Naturally the zero-mean d-dimensional Gaussian distribution on $\mathbb{R}^d$ with covariance matrix I_d. corresponds to B distributed according to $\chi^2(d)$, in which case $m_4 = m_{2,2} = m_2^2$.*

Proof. We use the polar parametrisation of the multivariate normal distribution. Let

$$v(\phi) = \Big(\cos\phi_1, \sin\phi_1\cos\phi_2, \ldots, \cos(\phi_k)\prod_{i=1}^{k-1}\sin(\phi_i), \ldots, \prod_{i=1}^{d-1}\sin(\phi_i)\Big),$$

$\phi \in [0,\pi]^{d-2}\times[0,2\pi]$. The probability distribution for ϕ ensuring uniformity of $v(\phi)$ on the surface of the $d-$sphere has density

$$\varpi(\phi) \propto \prod_{i=1}^{d-2}\sin^{d-i-1}(\phi_i)\mathbb{1}_{[0,\pi]^{d-2}\times[0,2\pi]}(\phi)\ ,$$

with respect to the Lebesgue measure on $\mathbb{R}^{d-1}$. Let Φ be random variable with distribution ϖ. Further let $B \sim \chi^2(d)$ be independent of Φ then it is standard knowledge that $W = B^{1/2}v(\Phi)$ follows the zero-mean d-dimensional Gaussian distribution on $\mathbb{R}^d$ with covariance matrix I_d. Therefore, by construction,

$$\begin{aligned}\mathbb{E}\Big[W_iW_jW_kW_l\Big] = \mathbb{E}\Big[B^2v_i(\Phi)v_j(\Phi)v_k(\Phi)v_l(\Phi)\Big] &= \mathbb{E}\Big[B^2\Big]\mathbb{E}\Big[v_i(\Phi)v_j(\Phi)v_k(\Phi)v_l(\Phi)\Big]\\ &= d(d+2)\mathbb{E}\Big[v_i(\Phi)v_j(\Phi)v_k(\Phi)v_l(\Phi)\Big]\ ,\end{aligned}$$

and the latter term vanishes when the leftmost term does. We also deduce that

$$\mathbb{E}\Big[W_1^2\Big]\mathbb{E}\Big[W_2^2\Big] = \mathbb{E}\Big[W_1^2W_2^2\Big] = d(d+2)\mathbb{E}\Big[v_1^2(\Phi)v_2^2(\Phi)\Big]\ ,$$

from which we obtain $\mathbb{E}\Big[v_1^2(\Phi)v_2^2(\Phi)\Big]$. Similarly using properties of the moments of the normal distribution,

$$3\mathbb{E}\Big[W_1^2\Big]^2 = \mathbb{E}\Big[W_1^4\Big] = d(d+2)\mathbb{E}\Big[v_1^4(\Phi)\Big]\ ,$$

leading to the expression for $\mathbb{E}\Big[v_1^4(\Phi)\Big]$. The last statement is straightforward.

□

5.10 Expectation of quadratic forms of the velocity

This section provides expressions for second order moments of quadratic forms of v for a large class of distributions for which we could not find adequate references.

Lemma 5.13. *Let $M \in \mathbb{R}^{d\times d}$ be a symmetric matrix, $c \in \mathbb{R}$ and assume the distribution ν of v is such that*

(a) for any bounded and measurable function $f : \mathbb{R}^2 \to \mathbb{R}$, $i, j \in \{1, \dots, d\}$ such that $i \neq j$, $\int f(v_i, v_j)\,\mathrm{d}\nu(v) = \int f(v_1, v_2)\,\mathrm{d}\nu(v)$,

(b) for $i, j, k, l \in \{1, \dots, d\}$, we have $\int v_i v_j v_k v_l\,\mathrm{d}\nu(v) = 0$ whenever $\mathrm{card}(\{i, j, k, l\}) > 2$.

Then

$$\left\|v^\top M v - c\right\|_\nu^2 = 3(m_4 - m_{2,2})\mathrm{Tr}(M \odot M) + (m_2\mathrm{Tr}(M) - c)^2 + 2m_{2,2}\mathrm{Tr}(M^2), \tag{5.97}$$

where $\odot$ denotes the Hadamard product.

Proof. Using that M is symmetric, and the expectation symbol for expectations with respect to ν,

$$\mathbb{E}\left[\left(\sum_{i,j=1}^d M_{ij}v_i v_j - c\right)^2\right] = \sum_{i,j,k,\ell=1}^d M_{ij}M_{k\ell}\mathbb{E}[v_i v_j v_k v_\ell] - 2c\sum_{i,j=1}^d M_{ij}\mathbb{E}[v_i v_j] + c^2 \tag{5.98}$$

where

$$\begin{aligned}\sum_{i,j,k,\ell=1}^d M_{ij}M_{k\ell}\mathbb{E}[v_i v_j v_k v_\ell] &= 3m_4\sum_{i=1}^d M_{ii}^2 + m_{2,2}\sum_{i\neq j} M_{ii}M_{jj} + 2m_{2,2}\sum_{i\neq j} M_{ij}^2 \\ &= (3m_4 - 3m_{2,2})\sum_{i=1}^d M_{ii}^2 + m_{2,2}\sum_{i,j=1}^d \left(M_{ii}M_{jj} + 2M_{ij}^2\right) \\ &= (3m_4 - 3m_{2,2})\mathrm{Tr}(M \odot M) + m_{2,2}\left(\mathrm{Tr}(M)^2 + 2\mathrm{Tr}(M^2)\right).\end{aligned} \tag{5.99}$$

Therefore

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{i,j=1}^d M_{ij}v_i v_j - c\right)^2\right] &= (3m_4 - 3m_{2,2})\mathrm{Tr}(M \odot M) + m_{2,2}\mathrm{Tr}(M)^2 + 2m_{2,2}\mathrm{Tr}(M^2) \\ &\quad - 2cm_2\mathrm{Tr}(M) + c^2,\end{aligned} \tag{5.100}$$

which implies the desired result. □

Corollary 5.5. *Given a symmetric matrix $M \in \mathbb{R}^{d\times d}$ and a constant $c \in \mathbb{R}$,*

$$\left\|v^\top M v - m_2\mathrm{Tr}(M)\right\|_\nu \leqslant \sqrt{2m_{2,2} + 3(m_4 - m_{2,2})_+}|M|. \tag{5.101}$$

Acknowledgment

JR would like to thank Pierre Monmarché for showing him how ZZ and BPS fall under a general framework. CA acknowledges support from EPSRC "Intractable Likelihood: New

Challenges from Modern Applications (ILike)" (EP/K014463/1). All the authors acknowledge the support of the Institute for Statistical Science in Bristol. AD acknowledges support from Chaire BayeScale "P. Laffitte".

Bibliography

[1] A. Abdulle, G. A. Pavliotis, and U. Vaes. Spectral methods for multiscale stochastic differential equations. *SIAM/ASA J. Uncertain. Quantif.*, 5(1):720–761, 2017.

[2] M. Allen and D. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publications, 1987.

[3] T. W. Anderson. *The Statistical Analysis of Time Series*, volume 19. John Wiley & Sons, 1971.

[4] C. Andrieu, A. Durmus, N. Nüsken, and J. Roussel. Hypercoercivity of Piecewise Deterministic Markov Process-Monte Carlo. *arXiv:1808.08592*, 2018.

[5] G. Arampatzis, M. A. Katsoulakis, and L. Rey-Bellet. Efficient estimators for likelihood ratio sensitivity indices of complex stochastic dynamics. *J. Chem. Phys.*, 144(10):104107, 2016.

[6] V. I. Arnol'd. *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.

[7] R. Assaraf and M. Caffarel. Zero-variance principle for Monte Carlo algorithms. *Phys. Rev. Lett.*, 83:4682–4685, 1999.

[8] R. Assaraf, B. Jourdain, T. Lelièvre, and R. Roux. Computation of sensitivities for the invariant measure of a parameter dependent diffusion. *Stoch. Partial Differ. Equ. Anal. Comput.*, 6(2):125–183, 2018.

[9] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Elect. Comm. in Probab.*, 13:60–66, 2008.

[10] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[11] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[12] R. Balian. *From Microphysics to Macrophysics. Methods and Applications of Statistical Physics*, volume I - II. Springer, 2007.

[13] R. H. Bartels and G. W. Stewart. Solution of the matrix equation AX + XB = C. *Commun. ACM*, 15(9):820–826, 1972.

[14] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *P. Am. Math. Soc.*, 105(2):397–400, 1989.

[15] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Probab. Theory Relat. Fields*, 60(2):185–201, 1982.

[16] J. Bierkens, P. Fearnhead, and G. Roberts. The zig-zag process and super-efficient sampling for Bayesian analysis of big data. *arXiv:1607.03188*, 2016.

[17] J. Bierkens, K. Kamatani, and G. O. Roberts. High-dimensional scaling limits of piecewise deterministic sampling algorithms. *arXiv:1807.11358*, 2018.

[18] J. Bierkens, G. Roberts, and P.-A. Zitt. Ergodicity of the zigzag process. *arXiv:1712.09875*, 2018.

[19] S. G. Bobkov. *Spectral Gap and Concentration for Some Spherically Symmetric Probability Measures*, pages 37–43. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[20] F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet. Fourier's law: a challenge to theorists. In *Mathematical physics 2000*, pages 128–150. Imperial College Press, London, 2000.

[21] M. Bonnefont, A. Joulin, and Y. Ma. Spectral gap for spherically symmetric log-concave probability measures, and beyond. *Journal of Functional Analysis*, 270(7):2456 – 2482, 2016.

[22] N. Bou-Rabee, A. Eberle, and R. Zimmer. Coupling and convergence for Hamiltonian Monte Carlo. *arXiv:1805.00452*, 2018.

[23] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, 48(1):278–297, 2010.

[24] N. Bou-Rabee and J. M. a. Sanz-Serna. Randomized Hamiltonian Monte Carlo. *Ann. Appl. Probab.*, 27(4):2159–2194, 2017.

[25] A. Bouchard-Côté, S. J. Vollmer, and A. Doucet. The Bouncy Particle Sampler: A nonreversible rejection-free Markov chain Monte Carlo method. *J. Am. Stat. Assoc.*, pages 1–13, 2018.

[26] E. Bouin, J. Dolbeault, S. Mischler, C. Mouhot, and C. Schmeiser. Hypocoercivity without confinement. *arXiv:1708.06180*, 2017.

[27] O. M. Braun and R. Ferrando. Role of long jumps in surface diffusion. *Phys. Rev. E*, 65:061107, 2002.

[28] R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7:1–49, 1998.

[29] V. Calvez, G. Raoul, and C. Schmeiser. Confinement by biased velocity jumps: aggregation of *escherichia coli*. *Kinet. Relat. Models*, 8(4):651–666, 2015.

[30] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris, and Y. Maday. Computational quantum chemistry: a primer. In *Handbook of numerical analysis, Vol. X*, Handb. Numer. Anal., X, pages 3–270. North-Holland, Amsterdam, 2003.

[31] E. Cancès, V. Ehrlacher, and T. Lelièvre. Convergence of a greedy algorithm for high-dimensional convex nonlinear problems. *Math. Models Methods Appl. Sci.*, 21(12):2433–2467, 2011.

[32] E. Cancès, F. Legoll, and G. Stoltz. Theoretical and numerical comparison of some sampling methods for molecular dynamics. *M2AN*, 41(2):351–389, 2007.

[33] C. Cao. The kinetic Fokker-Planck equation with weak confinement force. *arXiv:1801.10354*, 2018.

[34] P. Carmona. Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths. *Stoch. Proc. Appl.*, 117(8):1076–1092, 2007.

[35] F. Chatelin. *Spectral Approximation of Linear Operators*, volume 65 of *Classics in Applied Mathematics*. SIAM, 2011.

[36] M.-H. Chen and Q.-M. Shao. On Monte Carlo methods for estimating ratios of normalizing constants. *Ann. Statist.*, 25(4):1563–1594, 1997.

[37] E. B. Davies. *Spectral Theory and Differential Operators*, volume 42 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[38] M. H. A. Davis. Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, 46(3):353–388, 1984.

[39] G. Deligiannidis, A. Bouchard-Côté, and A. Doucet. Exponential Ergodicity of the Bouncy Particle Sampler. *arXiv:1705.04579*, 2017.

[40] P. Dellaportas and I. Kontoyiannis. Control variates for estimation based on reversible Markov chain Monte Carlo samplers. *J. R. Stat. Soc. Series B Stat. Methodol.*, 74(1):133–161, 2012.

[41] A. Dhar. Heat transport in low-dimensional systems. *Adv. Phys.*, 57(5):457–537, 2008.

[42] A. R. Dinner, J. C. Mattingly, J. O. B. Tempkin, B. Van Koten, and J. Weare. Trajectory stratification of stochastic dynamics. *arXiv:1610.09426*, 2016.

[43] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for kinetic equations with linear relaxation terms. *C. R. Math. Acad. Sci. Paris*, 347(9-10):511–516, 2009.

[44] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Trans. AMS*, 367:3807–3828, 2015.

[45] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Physics Letters B*, 195(2):216 – 222, 1987.

[46] A. Durmus, A. Guillin, and P. Monmarché. Geometric ergodicity of the bouncy particle sampler. *arXiv:1807.05401*, 2018.

[47] A. Durmus, A. Guillin, and P. Monmarché. Piecewise Deterministic Markov Processes and their invariant measure. *arXiv:1807.05421*, 2018.

[48] A. Eberle, A. Guillin, and R. Zimmer. Couplings and quantitative contraction rates for Langevin dynamics. *arXiv:1703.01617*, 2017.

[49] J.-P. Eckmann and M. Hairer. Spectral properties of hypoelliptic operators. *Commun. Math. Phys.*, 235(2):233–253, 2003.

[50] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Commun. Math. Phys.*, 201(3):657–697, 1999.

[51] K.-J. Engel and R. Nagel. *A Short Course On Operator Semigroups*. Universitext. Springer, New York, 2006.

[52] D. J. Evans and G. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. ANU Press, 1990.

[53] D. J. Evans and G. P. Morriss. The isothermal/isobaric molecular dynamics ensemble. *Phys. Rev. A*, 98(8-9):433–436, 1983.

[54] D. J. Evans and G. P. Morriss. *Computer Simulation Algorithms*. ANU Press, 2013.

[55] A. Faggionato, D. Gabrielli, and M. Ribezzi Crivellari. Non-equilibrium thermodynamics of piecewise deterministic markov processes. *Journal of Statistical Physics*, 137(2):259, 2009.

[56] E. Faou and T. Lelièvre. Conservative stochastic differential equations: mathematical and numerical analysis. *Math. Comp.*, 78(268):2047–2074, 2009.

[57] L. E. Figueroa and E. Süli. Greedy approximation of high-dimensional Ornstein-Uhlenbeck operators. *Found. Comput. Math.*, 12(5):573–623, 2012.

[58] G. S. Fishman. *Monte Carlo: Concepts, Algorithms and Applications*. Springer, 1996.

[59] E. L. Foster, J. Lohéac, and M.-B. Tran. A structure preserving scheme for the Kolmogorov–Fokker–Planck equation. *J.Comput. Phys.*, 330:319 – 339, 2017.

[60] M. Freidlin. Some remarks on the Smoluchowski-Kramers approximation. *J. Statist. Phys.*, 117(3-4):617–634, 2004.

[61] M. Freidlin and M. Weber. Random perturbations of nonlinear oscillators. *Ann. Probab.*, 26(3):925–967, 1998.

[62] M. Freidlin and M. Weber. A remark on random perturbations of the nonlinear pendulum. *Ann. Appl. Probab.*, 9(3):611–628, 1999.

[63] M. I. Freidlin and A. D. Wentzell. Diffusion processes on an open book and the averaging principle. *Stochastic Process. Appl.*, 113(1):101–126, 2004.

[64] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, 2002.

[65] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*. Texts in Statistical Science Series. CRC Press, Boca Raton, FL, third edition, 2014.

[66] C. J. Geyer. Estimating normalizing constants and reweighting mixtures. Technical report 565. 1994.

[67] P. W. Glynn and M. Olvera-Cravioto. Likelihood ratio gradient estimation for steady-state parameters. *arXiv:1707.02659*, 2017.

[68] J. B. Goodman and K. K. Lin. Coupling control variates for Markov chain Monte Carlo. *J. Comput. Phys.*, 228(19):7127–7136, 2009.

[69] R. D. Groot and P. B. Warren. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.*, 107(11):4423–4435, 1997.

[70] M. Grothaus and P. Stilgenbauer. Hypocoercivity for Kolmogorov backward evolution equations and applications. *J. Funct. Anal.*, 267:3515–3556, 2014.

[71] M. Grothaus and P. Stilgenbauer. A hypocoercivity related ergodicity method for singularly distorted non-symmetric diffusions. *Integral Equations Operator Theory*, 83(3):331–379, 2015.

[72] M. Grothaus and P. Stilgenbauer. Hilbert space hypocoercivity for the Langevin dynamics revisited. *Methods Funct. Anal. Topology*, 22(2):152–168, 2016.

[73] W. Hackbusch. *Tensor Spaces and Numerical Tensor Calculus*, volume 42. Springer Science & Business Media, 2012.

[74] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.

[75] M. Hairer and G. A. Pavliotis. From ballistic to diffusive behavior in periodic potentials. *J. Stat. Phys.*, 131(1):175–202, 2008.

[76] R. Z. Has'minskiĭ. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn—Germantown, Md., 1980.

[77] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[78] S. G. Henderson. *Variance Reduction via an Approximating Markov Process*. PhD thesis, Stanford University, 1997.

[79] F. Hérau. Short and long time behavior of the Fokker-Planck equation in a confining potential and applications. *J. Funct. Anal.*, 244(1):95–118, 2007.

[80] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, 171:151–218, 2004.

[81] D. P. Herzog and J. C. Mattingly. Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials. *arXiv:1711.02250*, 2017.

[82] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.*, 46(5-6):1159–1194, 1987.

[83] L. Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119(1):147–171, 1967.

[84] R. Hulse, R. Rowley, and W. Wilding. Transient nonequilibrium molecular dynamic simulations of thermal conductivity: 1. Simple fluids. *Int. J. Thermophys.*, 26(1):1–12, 2005.

[85] A. Iacobucci, S. Olla, and G. Stoltz. Convergence rates for nonequilibrium Langevin dynamics. *Ann. Math. Quebec*, 2017.

[86] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1-2):134–139, 1918.

[87] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.

[88] R. Joubaud and G. Stoltz. Nonequilibrium shear viscosity computations with Langevin dynamics. *Multiscale Model. Simul.*, 10(1):191–216, 2012.

[89] W. Kliemann. Recurrence and invariant measures for degenerate diffusions. *Ann. Probab.*, 15(2):690–707, 1987.

[90] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 1991.

[91] A. Kong, P. McCullagh, X.-L. Meng, D. Nicolae, and Z. Tan. A theory of statistical models for Monte Carlo integration. *J. R. Stat. Soc. Series B Stat. Methodol.*, 65(3):585–604, 2003.

[92] I. Kontoyiannis and S. P. Meyn. Geometric ergodicity and the spectral gap of nonreversible markov chains. *Probability Theory and Related Fields*, 154(1-2):327–339, 2012.

[93] M. Kopec. Weak backward error analysis for overdamped Langevin processes. *IMA J. Numer. Anal.*, 35(2):583–614, 2014.

[94] M. Kopec. Weak backward error analysis for Langevin process. *BIT*, 55(4):1057–1103, 2015.

[95] S. M. Kozlov. Effective diffusion in the Fokker-Planck equation. *Math. Notes*, 45(5):360–368, 1989.

[96] D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo Methods*, volume 706. John Wiley & Sons, 2013.

[97] R. Kubo, M. Toda, and N. Hashitsume. *Statistical Physics. II*, volume 31 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, second edition, 1991. Nonequilibrium Statistical Mechanics.

[98] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.*, 13(8):1011–1021, 1992.

[99] B. Lapeyre, E. Pardoux, and R. Sentis. *Introduction to Monte-Carlo Methods for Transport and Diffusion Equations*, volume 6 of *Oxford Texts in Applied and Engineering Mathematics*. Oxford University Press, Oxford, 2003.

[100] J. C. Latorre, G. A. Pavliotis, and P. R. Kramer. Corrections to Einstein's relation for Brownian motion in a tilted periodic potential. *J. Stat. Phys.*, 150(4):776–803, 2013.

[101] B. Leimkuhler and C. Matthews. *Molecular Dynamics*. Springer, 2016.

[102] B. Leimkuhler, C. Matthews, and G. Stoltz. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.*, 36(1):13–79, 2016.

[103] T. Lelièvre, M. Rousset, and G. Stoltz. *Free Energy Computations: A Mathematical Perspective*. Imperial College Press, London, 2010.

[104] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numerica*, 25:681–880, 2016.

[105] S. Lepri, R. Livi, and A. Politi. Thermal conduction in classical low-dimensional lattices. *Phys. Rep.*, 377(1):1–80, 2003.

[106] S. Lepri, R. Livi, and A. Politi. Heat transport in low dimensions: introduction and phenomenology. In *Thermal Transport in Low Dimensions*, volume 921 of *Lecture Notes in Phys.*, pages 1–37. Springer, 2016.

[107] P. A. W. Lewis and G. S. Shedler. Simulation of nonhomogeneous Poisson processes by thinning. *Naval Res. Logist. Quart.*, 26(3):403–413, 1979.

[108] J. S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer Series in Statistics. Springer-Verlag, New York, 2001.

[109] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, 101(2):185–232, 2002.

[110] X.-L. Meng and W. H. Wong. Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Statist. Sinica*, 6(4):831–860, 1996.

[111] G. Metafune, D. Pallara, and E. Priola. Spectrum of Ornstein-Uhlenbeck operators in L^p spaces with respect to invariant measures. *J. Funct. Anal.*, 196(1):40 – 60, 2002.

[112] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

[113] S. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, Cambridge, second edition, 2009.

[114] M. Michel, S. C. Kapfer, and W. Krauth. Generalized event-chain Monte Carlo: Constructing rejection-free global-balance algorithms from infinitesimal steps. *J. Chem. Phys.*, 140(5):054116, 2014.

[115] M. Michel and S. Sénécal. Forward Event-Chain Monte Carlo: a general rejection-free and irreversible Markov chain simulation method. *arXiv:1702.08397*, 2017.

[116] G. N. Milstein and M. V. Tretyakov. *Stochastic Numerics For Mathematical Physics*. Scientific Computation. Springer-Verlag, Berlin, 2004.

[117] R. A. Minlos. *Introduction to Mathematical Statistical Physics*, volume 19 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.

[118] A. Mira and D. J. Sargent. A new strategy for speeding Markov chain Monte Carlo algorithms. *Stat. Methods Appt.*, 12(1):49–60, 2003.

[119] A. Mira, R. Solgi, and D. Imparato. Zero variance Markov chain Monte Carlo for Bayesian estimators. *Stat. Comput.*, 23(5):653–662, 2013.

[120] B. Mokdad, E. Pruliere, A. Ammar, and F. Chinesta. On the simulation of kinetic theory models of complex fluids using the Fokker-Planck approach. *Appl. Rheol.*, 17(2):26494, 2007.

[121] E. Nelson. *Dynamical Theories of Brownian motion*. Princeton University Press, Princeton, N.J., 1967.

[122] C. J. Oates, M. Girolami, and N. Chopin. Control functionals for Monte Carlo integration. *J. R. Stat. Soc. Series B Stat. Methodol.*, 79(3):695–718, 2017.

[123] R. O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014.

[124] C. Pangali, M. Rao, and B. Berne. A Monte Carlo simulation of the hydrophobic interaction. *J. Chem. Phys.*, 71(7):2975–2981, 1979.

[125] E. Pardoux and Y. Veretennikov. On the Poisson equation and diffusion approximation. i. *Ann. Probab.*, 29(3):1061–1085, 2001.

[126] G. Pavliotis and A. Vogiannou. Diffusive transport in periodic potentials: underdamped dynamics. *FNL*, 8(02):L155–L173, 2008.

[127] G. A. Pavliotis and A. M. Stuart. *Multiscale Methods: Averaging and Homogenization*. Springer Science & Business Media, 2008.

[128] G. K. Pedersen. *Analysis Now*, volume 118. Springer Science & Business Media, 1995.

[129] A. Persson. Bounds for the discrete part of the spectrum of a semi-bounded Schrödinger operator. *Mathematica Scandinavica*, 8(1):143–153, 1960.

[130] E. A. Peters et al. Rejection-free monte carlo sampling for general potentials. *Physical Review E*, 85(2):026703, 2012.

[131] H. Petersen and H. Flyvbjerg. Error estimates in molecular dynamics simulations. *J. Chem. Phys*, 91:461–467, 1989.

[132] E. Platen. An introduction to numerical methods for stochastic differential equations. *Acta Numerica*, 8:197–246, 1999.

[133] A. Porretta and E. Zuazua. Numerical hypocoercivity for the Kolmogorov equation. *Math. Comp.*, 86(303):97–119, 2017.

[134] L. Qi. Some simple estimates for singular values of a matrix. *Linear Algebra Appl.*, 56:105 – 119, 1984.

[135] S. Redon, G. Stoltz, and Z. Trstanova. Error analysis of modified Langevin dynamics. *J. Stat. Phys.*, 164(4):735–771, 2016.

[136] M. Reed and B. Simon. *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.

[137] L. Rey-Bellet. Ergodic properties of Markov processes. In *Open Quantum Systems. II*, volume 1881 of *Lecture Notes in Math.*, pages 1–39. Springer, Berlin, 2006.

[138] L. Rey-Bellet. Open classical systems. In *Open Quantum Systems. II*, volume 1881 of *Lecture Notes in Math.*, pages 41–78. Springer, Berlin, 2006.

[139] H. Risken and T. Frank. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer Berlin Heidelberg, 1996.

[140] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Science & Business Media, 2013.

[141] G. O. Roberts, A. Gelman, and W. R. Gilks. Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.*, 7(1):110–120, 1997.

[142] G. O. Roberts and J. S. Rosenthal. Optimal scaling of discrete approximations to Langevin diffusions. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 60(1):255–268, 1998.

[143] H. Rodenhausen. Einstein's relation between diffusion constant and mobility for a diffusion model. *J. Stat. Phys.*, 55(5-6):1065–1088, 1989.

[144] J. Roussel and G. Stoltz. A perturbative approach to control variates. *arXiv:1712.08022*, 2017.

[145] J. Roussel and G. Stoltz. Spectral methods for Langevin dynamics and associated error estimates. *M2AN*, 52(3):1051–1083, 2018.

[146] M. Rousset, Y. Xu, and P.-A. Zitt. A Weak Overdamped Limit Theorem for Langevin Processes. *arXiv:1709.09866*, 2017.

[147] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, 2017.

[148] M. R. Shirts and J. D. Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.*, 129(12):124105, 2008.

[149] R. B. Sowers. Stochastic averaging near homoclinic orbits via singular perturbations. In *IUTAM Symposium on Nonlinear Stochastic Dynamics*, volume 110 of *Solid Mech. Appl.*, pages 83–94. Kluwer Acad. Publ., Dordrecht, 2003.

[150] R. B. Sowers. A boundary layer theory for diffusively perturbed transport around a heteroclinic cycle. *Comm. Pure Appl. Math.*, 58(1):30–84, 2005.

[151] H. Spohn. Nonlinear fluctuating hydrodynamics for anharmonic chains. *J. Stat. Phys.*, 154(5):1191–1227, 2014.

[152] J. E. Straub, M. Borkovec, and B. J. Berne. Molecular dynamics study of an isomerizing diatomic in a Lennard-Jones fluid. *J. Chem. Phys.*, 89(8):4833–4847, 1988.

[153] D. Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Proc. Rel. Fields*, 8(2):163–198, 2002.

[154] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509, 1990.

[155] Z. Tan. On a likelihood approach for Monte Carlo integration. *J. Am. Stat. Assoc.*, 99(468):1027–1036, 2004.

[156] V. N. Temlyakov. Greedy approximation. *Acta Numer.*, 17:235–409, 2008.

[157] E. H. Thiede, B. Van Koten, J. Weare, and A. R. Dinner. Eigenvector method for umbrella sampling enables error analysis. *J. Chem. Phys.*, 145(8):084115, 2016.

[158] B. D. Todd and P. J. Daivis. Homogeneous non-equilibrium molecular dynamics simulations of viscous flow: techniques and applications. *Mol. Simul.*, 33(3):189–229, 2007.

[159] B. D. Todd and P. J. Daivis. *Nonequilibrium Molecular Dynamics: Theory, Algorithms and Applications*. Cambridge University Press, 2017.

[160] G. M. Torrie and J. P. Valleau. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.*, 23(2):187–199, 1977.

[161] M. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press, 2010.

[162] P. Vanetti, A. Bouchard-Côté, G. Deligiannidis, and A. Doucet. Piecewise Deterministic Markov Chain Monte Carlo. *arXiv:1707.05296*, 2017.

[163] Y. Vardi. Empirical distributions in selection bias models. *Ann. Statist.*, 13(1):178–203, 1985.

[164] L. Verlet. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159(1):98, 1967.

[165] C. Villani. Hypocoercive diffusion operators. In *International Congress of Mathematicians*, volume 3, pages 473–498, 2006.

[166] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950), 2009.

[167] H. Wang, C. Hartmann, and C. Schütte. Linear response theory and optimal control for a molecular system under non-equilibrium conditions. *Mol. Phys.*, 111(22-23):3555–3564, 2013.

[168] T. Wang and P. Plechac. Steady state sensitivity analysis of continuous time Markov Chains. *arXiv:1804.00585*, 2018.

[169] A. Warmflash, P. Bhimalapuram, and A. R. Dinner. Umbrella sampling for nonequilibrium processes. *J. Chem. Phys.*, 127(15):114109, 2007.

[170] C. Wu and C. P. Robert. Generalized bouncy particle sampler. *arXiv:1706.04781*, 2017.

[171] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.*, 91(2):205 – 238, 2001.

[172] H. Yserentant. *Regularity and Approximability of Electronic Wave Functions*, volume 2000 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2010.