# Probability in computational physics and biology: some mathematical contributions

Mathias Rousset

**INRIA Rocquencourt — Ecole des Ponts ParisTech — Université Paris Est**

**HABILITATION A DIRIGER DES RECHERCHES**

Spécialité : **Mathématiques**

Présentée par

**Mathias Rousset**
**Equipe-projet MATHERIALS, INRIA Rocquencourt & CERMICS, Ecole des Ponts**

# Probability in computational physics and biology: some mathematical contributions

Soutenue publiquement le 27 novembre 2014 devant le jury composé de

| | | |
|---|---|---|
| Laurent DESVILLETTES | ENS Cachan | Examinateur |
| Erwan FAOU | INRIA & ENS | Examinateur |
| Josselin GARNIER | Université Paris 6 & 7 | Examinateur |
| Arnaud GUILLIN | Université Blaise Pascal | Rapporteur |
| Tony LELIEVRE | ENPC | Examinateur |
| Laurent MICLO | Université Paul Sabatier | Examinateur |
| Sylvie MELEARD | Ecole Polytechnique | Examinateur |

au vu du rapport écrit de

| | | |
|---|---|---|
| Luc REY-BELLET | University of Massachusetts | Rapporteur |

et de l'avis favorable de

| | | |
|---|---|---|
| Mireille Bossy | INRIA | Rapporteure |

# Remerciements

Mes premiers remerciements sont pour les membres du jury, qui ont accepté ma demande. Je remercie tout particulièrement mes rapporteurs, pour leur temps consacré à l'étude du manuscrit.

Je souhaiterais évoquer ensuite les rencontres scientifiques qui ont irrigué ma recherche années après années, par leurs conseils, leur culture, ou leur amitié: JM. Schlenker, P. Del Moral, L. Miclo, A. Doucet, G. Stoltz, T. Lelièvre, C. Le Bris, P. Plechac, T. Goudon, S. De Bièvre, G. Samaey, G. Dujardin, A. Gloria, JF. Coulombel, C. Besse, I. Violet, P. Lafitte, B. Jourdain, JF. Delmas, E. Cancès, S. Boyaval, V. Ehrlacher, L. Monasse, N. Fournier, C. Mouhot.

Finalement, rien ne serait possible sans le soutien sans faille de ma famille et de mes amis : mes parents, mon frère et son amie, tous les copains, et surtout, ma compagne L..

# Contents

# Part I

# Introduction

# General presentation of the dissertation

*Foreword*

The present Habilitation Thesis is structured into four chapters (or "contributions") that summarize the material contained in the articles [1–5, 7].

## Introduction

My scientific activity is mainly dedicated to the mathematical study of models coming from computational physics, and to a less extent, biology. These models are the following:

(i) Stochastically perturbed (thermostatted) Hamiltonian systems. Such systems are widely used in classical molecular simulation. They are the subject of the papers [4, 7], and of Chapter 1 of the present dissertation.

(ii) Fermionic Schrödinger operators, which describes non-relativistic systems of Fermionic (exchangeable) particles, and are central to computational chemistry (the particles are the electrons of molecules). Related probabilistic interpretations are studied in the paper [5], and are presented in Chapter 2.

(iii) Individual-based models of bacterial chemotaxis. Such models describe the random motion of each bacterium depending on the chemical environment. They are studied in the papers [2,3], and presented in Chapter 3.

(iv) Boltzmann's kinetic theory of rarefied gases; with focus on the space homogeneous simplification, as well as on the associated conservative $N$-particle system. It is the subject of the article [1], and of Chapter 4 of the present thesis.

Although the material is mainly written in mathematical style, the physics of the considered systems is a non-negligeable source of motivation and intuition. Some contributions are merely *theoretical*, with mathematical theorems analyzing some physically relevant features of the models. Other contributions ar more *applied*, with suggestions of numerical methods and realistic numerical tests.

Different standard mathematical tools are required in the basic analysis of the considered problems. For instance, some concepts of differential geometry for Hamiltonian systems with constraints; the spectral theorem for Schrödinger operators; and usual stochastic calculus associated with Markov processes in $\mathbb{R}^d$ in any case.

In the same way, the classical spectrum of probabilistic tools are used in the core of the presented contributions. For instance, the reader which is not an expert in probability theory shall not be completely unfamiliar with the following concepts: (i) changes of probability measures for stochastic processes, (ii) tightness and convergence in (probability distribution) law for processes;

(iii) basic Feynman-Kac representation of parabolic partial differential equations, probabilistic interpretation of boundary conditions; (iv) Coupling methods, which may be summarized for the unfamiliar reader with the aphorism: "the same random numbers are used to construct (or simulate) and compare two similar systems, or models".

## Summary of chapters

### Thermostatted molecular dynamics

Chapter 1 is devoted to classical Hamiltonian systems, with usual quadratic kinetic and potential energy decomposition. In the Euclidean space defined by the kinetic energy, the latter decomposition reads

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

with standard notation $(q,p)$ for position and momenta. For $N$ particles in Euclidean space, we have $(q,p) \in T^*\mathbb{R}^{3N} \simeq \mathbb{R}^{3N} \times \mathbb{R}^{3N}$, $T^*$ denoting the co-tangent bundle. The latter is coupled to a stochastic thermostat at a given temperature, resulting in a Markov diffusion process called Langevin processes, which is widely used in molecular simulation. This Markov process has a unique stationary probability distribution, called the *canonical Gibbs probability distribution*, defined up to a normalizing constant by

$$\mathrm{e}^{-\beta H(q,p)} dqdp,$$

where $dqdp$ is the phase-space (Liouville) measure which will be precised in Chapter 1. When distributed according this stationary probability distribution, the process is also *time reversible upon momenta reversal* $p \to -p$; so that the canonical Gibbs probability is also called *equilibrium stationary state* in statistical mechanics.

In practical molecular systems, the typical vibration period of a molecular bond is the femtosecond ($10^{-15}$ seconds); while the physically interesting behaviors, like conformation changes in proteins, are rare events happening on much larger time scales, at least of order $10^{-9}s$. Since the time step stability of a direct numerical time integration of the system is limited by the fastest degrees of freedom, millions of iterations are required to simulate and compute interesting physically phenomenons. This forms the **timescale problem**.

The main motivation of the contributions of Chapter 1 is to study some numerical methods that help to overcome this timescale problem, either by accelerating the slowest timescales, or on the contrary by slowing down the fastest ones.

*Slow variables and free energy*

Slow variables, also called "reaction coordinates", or "collective variables", are described mathematically by smooth functions

$$\xi_{\mathrm{slow}}(q) \in \mathbb{R}^m$$

of the system position. In Chapter 1, illustrative simulations will be presented where $\xi_{\mathrm{slow}}$ is the end-to-end length of a long atomic chain. We focus on methods relying on a *prescribed constraint* on these precise slow variables. The basic idea is to reduce by this mean the time required by the original dynamics to explore the different values of $\xi_{\mathrm{slow}}$.

In particular, the computation of **free energy**, which in this context is simply the *image probability distribution of* $\xi_{\mathrm{slow}}$ under the canonical distribution, is a central theme. When the constraint is *time-dependent*, such calculations can be formulated with the *Crooks-Jarzynski relation*. The Crooks-Jarzynski relation is the explicit computation of the relative probability distribution

between (i) a time-dependent stochastic dynamics with initial equilibrium distribution, and (ii) looking backward in time the stochastic dynamics obtained by time-reversing the time-dependence.

I have been initiated to related topics during my PhD by G. Stoltz and T. Lelivre, while studying the application of "Population (or Sequential) Monte-Carlo" methods to this field (see [12]). We also analyzed the case of "over-damped" diffusions (see [11]) where the momenta instantaneously equilibrated by the strong coupling with the thermostat, and state space of the process reduced to the the position only. Subsequently, we worked on *adaptive methods* (see [8, 9]), which broadly speaking adaptively compute an importance sampling biasing force (or drift). We then wrote a book (see [10]) summarizing our current understanding of the available techniques to compute free energy differences. We refer to the latter for a review of the associated computational physics literature.

The main contribution presented in this dissertation can be found in [4]. It contains the rigorous analysis of the case of Langevin processes subject to a prescribed, perhaps time-dependent, constraint on $\xi_{\text{slow}}$. This includes the rigorous mathematical justification of main identities (free energy calculations, Jarzynski-Crooks relation), as well as a suggestion of variants of numerical schemes, with some numerical analysis.

*Fast variables and mass-penalization*

Fast degrees of freedom are also described mathematically by smooths functions

$$\xi_{\text{fast}}(q) \in \mathbb{R}^m$$

of the system position; typically for molecular systems, they are related to the co-valent interactions between atoms. For instance in the case of a linear atomic chain: the so-called *bond lengths*, the *bond angles* (formed by three consecutive atoms of a chain), and perhaps the *torsion angles* (formed by the two consecutive planes spanned by four atoms in a chain). The main feature of such fast degrees of freedom is their *highly oscillatory* nature, which constrain the time step resolution of numerical schemes.

In [7] (initiated during my post-doc supervised by P. Plechac), I proposed an original method based on **implicit mass penalization**, based on an extended Lagrangian formulation. This idea generalizes *ad hoc* and explicit changes of the system mass matrix (see [11, 61]). Our main contribution is the proposition of the latter **new** method, together with a systematic mathematical analysis, as well ad numerical analysis and numerical demonstration of efficiency.

The main advantage of mass penalization is to remove the highly oscillatory nature of fast degrees of freedom. It does not modify the statistical behavior (here, the canonical distribution of positions) as do the usual direct constraints on these fast variables. It can also be used to construct consistant dynamical integrators with lower stiffness in a simple way, by using a time-step dependent mass penalization.

**Fermionic eigenstates**

Chapter 2 is devoted to a probabilistic analysis related to *skew-symmetric*[1] (or Fermionic) eigenfunctions of a standard Schrödinger operator, solutions of the eigenvalue problem

$$-\frac{\Delta}{2}\psi^* + V\psi^* = E^*\psi^*.$$

In the above the symmetry constraint refers to an underlying finite symmetry group of the operator, usually the particles permutations in the case of Fermions.

[1] Functions whose sign is changed under odd ($\det = -1$) symmetries

As is well-known, the eigenfunction with minimal energy $E^*$ (the *groundstate*) with a plain (as opposed to skew) symmetry constraint is a *signed* real valued function, and is amenable to probabilistic representations (through the *Feynman-Kac formula*), and thus, to associated Monte-Carlo numerical computations. This is the basic idea behind Quantum Monte-Carlo (QMC) calculations of groundstates, and allows a great numerical accuracy.

In the case of Fermions, the skew-symmetric groundstate has no longer a sign, which forbids any generic Monte-Carlo method. This is the so-called *sign problem* in computational physics. The standard method method used in QMC calculations to overcome this problem is to enforce the skew-symmetry with some *Dirichlet boundary conditions* given by an *i*nitially guessed skew-symmetric function; the boundary being given by the 0-level set of the latter called the *approximate nodal domain*. The fact that the latter approximate nodal domain is not the 0-level set of the Fermionic groundstate introduces an error () the difference between the Dirichlet energy and the true Fermionic energy) known as the Fixed Node Approximation (FNA).

I have been initiated to this topic after my PhD by E. Cancès, T. Lelièvre, M. Caffarel and R. Assaraf while analyzing the application of "Population (or Sequential) Monte-Carlo" methods to this field (also known as "Diffusion Monte-Carlo" by computational chemists, see [13]).

The main contribution presented in this chapter (see also [5]) has been to give an original probabilistic **characterization** of non-vanishing Fixed Node Approximation, in terms of a **lack of symmetry of the distribution a weighted random process killed** on the approximate **nodal boundary**. In fact, we will show more by relating this lack of symmetry with the **shape derivative** of a Dirichlet energy associated with the approximate nodes. This result can give insight into the attempts of numerical improvements to this Fixed Node Approximation, which however remains probably a very difficult task.

**Bacteria with internal state**

In Chapter 3, we consider the motion of flagellated bacteria. The latter consists of a sequence of run phases, during which a bacterium moves in a straight line at constant speed; then the bacterium changes direction, at random, in a tumble phase. To bias movement towards regions with high concentration of chemoattractant, the turning rate of the bacterium is adjusted by an internal variable, which acts as a memory. If the concentration used to be bigger, the bacterium's direction is more likely to change and the turning rate is higher. This leads to a **Markov process for the position, velocity and internal variables**. The velocity variable follows a piecewise constant jump process, whose rate of jump depend on the internal variables. The latter internal variables obey in turn a position dependent ordinary differential equation.

I have been initiated to this topic by G. Samaey, and we thought it was an original example to test *(asymptotic) variance reduction* for the simulation of stochastic particle systems (here, bacteria) using a coupling method. The basic idea is the following: (i) perform deterministic numerical computations of a simplified, lower dimensional model, (ii) use the obtained information to reduce the statistical variance of the full Monte-Carlo particle simulation, and thus the computational cost of the whole simulation. This led to the works [2, 3].

More precisely, we performed a **probabilistic coupling**, using the same random numbers defining the random jump times and the new velocities of bacteria, of **two different dynamics** of bacteria: (i) the dynamics with a **model of internal state** , of arbitrary high dimension; and (ii) the dynamics with direct **gradient sensing**, where the turning rate is not driven by an internal state, but directly by the gradient of the chemoattractant concentration. The state space of the gradient sensing process is then typically of dimension $2d-1$ ($d$ for position, $d-1$ for normalized velocity), and, say, for $d = 1, 2$ the probability density can be rather easily simulated with a deterministic grid method (*e.g.* finite difference or finite volume). The latter deterministic simulation can then be used as a *control variate* to reduce the variance of the simulation of dynamics with internal state. The whole variance reduction method can be sketched by the formula:

$$X_{\text{variance reduced}} := \underbrace{X_{\text{internal}} - X_{\text{gradient sensing}}}_{\text{coupled}} + \underbrace{x_{\text{gradient sensing}}}_{\text{deterministic control variate}}$$

The first main contribution of the presented works is a preliminary theoretical asymptotic analysis of processes. First, we consider an asymptotic regime defined by a small parameter $\varepsilon$: the ratio between the timescale of the turning rate (fast), and the chemoattractant concentration evolution felt by the bacterium (slow). We **prove the pathwise convergence** (in probability distribution, when $\varepsilon \to 0$) of the position evolution of both processes (both in the case with internal state, and in the case with gradient sensing) towards the **same drift-diffusion process** (with drift given by the chemoattractant gradient). This was missing in the literature.

The second main contribution consists of the analysis of the *coupling* which is shown to be *asymptotic* in the sense that the difference between the positions of the two coupled processes, on appropriate diffusive timescales, vanish with $\varepsilon$ with an estimated rate. Moreover, variance reduced numerical methods based on the latter coupling are proposed, and tested. In particular it is shown that **asymptotic variance reduction** can be achieved, in the sense that the statistical variance of the simulation vanishes with the small parameter $\varepsilon \to 0$. This implies that in the non-realistic situation where the deterministic computation is infinitely cheap as compared to the Monte-Carlo one, the theoretical gain of the method becomes infinite in the limit $\varepsilon \to 0$.

### Markov coupling of Boltzmann collisions

In Chapter 4, we focus on the kinetic theory of collisional gases, in the simplified *space homogeneous* case. The associated conservative stochastic Kac's $N$-particle system is considered with Maxwell collisions, and general scattering distribution. For the unfamiliar reader, the latter consists of $N$ (exchangeable) unit mass particles endowed with a velocity in Euclidean space. They perform at *a constant rate*[2] random two-body conservative collisions, where kinetic energy and momentum are preserved.

I have been initiated to this topic during a stay in IPAM (invited by T. Goudon) on kinetic theory. The main initial motivation was the following: in spite of a very large literature and tradition on that subject, the quantitative use of the probabilistic Markov coupling method on this model was lacking. Our initial motivation to study such couplings are the following: (i) the coupling point of view may give new insight into the *trend to equilibrium*, which have already been fully studied mainly by using entropy methods; (ii) the potential use of *variance reduction* in DSMC (Direct Simulation Monte-Carlo), which are used to simulate plasmas and rarefied gases with the stochastic, particle interpretation of kinetic equations (this aspect is still a research perspective, and will be discussed in the last chapter only).

Let us recall that if $t \mapsto V_t \in E$ is a time homogeneous Markov process in Euclidean state space $E$, then the time homogeneous Markov process $t \mapsto (U_t, V_t)$ is called an *(symmetric) Markov coupling* if the marginal dynamics of $t \mapsto V_t$ and $t \mapsto U_t$ are **both** distributed according to the **same Markov dynamics of interest**. We will be interested in the case where the latter is almost surely *weakly contractive* (a kind of "dissipative" behavior) in the sense that for some appropriate distance, here Euclidean, and any $0 \leqslant t \leqslant t + h$, one has

$$|U_{t+h} - V_{t+h}| \leqslant |U_t - V_t| \text{ a.s..} \tag{0.1}$$

Typical examples of such couplings are obtained by considering solutions with different initial conditions but single given Brownian motion $t \mapsto W_t$, of stochastic differential equations in $\mathbb{R}^d$ of the form $dU_t = -\nabla \mathcal{V}(U_t)dt + \sqrt{2}dW_t$, where $\mathcal{V}$ is a *convex* potential. If $\mathcal{V}$ is strongly convex with constant $c > 0$, then (0.1) becomes a strict contraction, with constant $\mathrm{e}^{-h/c} < 1$. A similar

[2] Independence with respect to positions defines *space homogeneity*, independence with respect to the relative speed of two collisional particles defines precisely Maxwell collisions.

phenomenon occurs for geometric Brownian motions on Riemannian manifolds with uniformly positive Ricci curvature tensor; the latter playing the same role as the Hessian of $\mathcal{V}$ (both cases are aggregated using the famous "Bakry-Emery" condition for general diffusions on manifolds).

To be more precise, we will focus on *coupling creation* functionals by computing, for instance in the quadratic, $L^2$ case,

$$\mathcal{C}_2(u,v) \stackrel{\text{def}}{=} -\frac{d}{dt}\Big|_{t=0} \mathbb{E}_{U_0=u,V_0=v} |U_t - V_t|^2 \geqslant 0.$$

It is then possible to define a quadratic uniform contraction inequality between coupling and coupling creation, of the form

$$\mathcal{C}_2(u,v) \geqslant 2\kappa |u-v|^2 \quad \forall u,v \in E, \tag{0.2}$$

for $\kappa > 0$. Then taking for the initial distribution of the process $t \mapsto U_t$ a stationary distribution $U_0 \sim \pi_\infty$ immediately yields exponential trend to equilibrium with respect to the quadratic Wasserstein distance in the sense that denoting,$\pi_t =^{\text{def}} \text{Law}(V_t)$,

$$d_{\mathcal{W}_2}(\pi_t, \pi_\infty) \leqslant \mathrm{e}^{-\kappa t} d_{\mathcal{W}_2}(\pi_0, \pi_\infty),$$

where the Wasserstein distance is defined by

$$d_{\mathcal{W}_2}(\pi_1,\pi_2) \stackrel{\text{def}}{=} \inf_{\pi_1=\mu_1,\pi_2=\mu_2} \left( \int |x-y|^2 \, \mu(dx \times dy) \right)^{1/2}, \tag{0.3}$$

where $\mu$ is spanning all probabilities of $E \times E$ with first (resp. second) marginal $\pi_1 = \mu_1$ (resp. $\pi_2 = \mu_2$).

The <u>first contribution</u> presented in Chapter 4 is the explicit construction of an almost surely increasing symmetric Markov coupling of the conservative $N$-particle system with state space $E = (\mathbb{R}^d)^N$ with Maxwell collisions. The coupling is based on: (i) *Simultaneous collisions*, where the same particles of the coupled systems perform each collision at the same time; (ii) *Spherical parallel coupling* of each collision, a **parallel coupling** of isotropic random walks on spheres using *geometric parallel transport* (a simple elementary rotation in the spherical case). In the present case, the sphere of interest is $\mathbb{S}^{d-1}$ and consists of the possible **directions of the relative velocity of two collisional particles**. Since in dimension greater than 3, spheres are uniformly positively curved, the whole coupling of the particle system is indeed weakly contractive. The associated coupling creation functional is then explicitly computed and given up to a multiplicative constant by the average (over all particles) of the following alignement functional

$$\mathcal{C}_2(u,v) = \frac{d-2}{2d-2} \left\langle |u-u_*| \, |v-v_*| - (u-u_*)\cdot(v-v_*) \right\rangle_N, \tag{0.4}$$

where $(u,v) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ is a coupled state of particle systems, but in the right hand side we denote (with a slight abuse) $(v, v_*) \in \mathbb{R}^d \times \mathbb{R}^d$ and $(u, u_*) \in \mathbb{R}^d \times \mathbb{R}^d$ two coupled *pair of particles* velocities, with $\langle \quad \rangle_N$ denoting averaging over the particles of the particle system.

The <u>second contribution</u> was to introduce a **weakened form** of the coupling - coupling creation uniform inequality (0.2) in the case of the conservative $N$-particle system with Maxwell collisions. The key ingredient is an original sharp inequality, which bounds from above the $L_2$ coupling distance of two normalized random variables with some sort of higher order relative alignement average. In its simplest form, it satisfies

$$\left\langle |u-v|^2 \right\rangle_N \leqslant 2\frac{d}{d-1} \left\langle |u-u_*|^2 \, |v-v_*|^2 - ((u-u_*)\cdot(v-v_*))^2 \right\rangle_N, \tag{0.5}$$

where $(u,v) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ are two centered and **energy normalized** states ($\left\langle |u|^2 \right\rangle_N = \left\langle |v|^2 \right\rangle_N = 1$) with the positive correlation $\langle u \cdot v \rangle_N \geqslant 0$, and the isotropy of co-variance condition $\langle (u \otimes u) \rangle_N = \frac{1}{d} \mathrm{Id}_d$.

Using Hölder inequality to compare (0.4) with the right hand side of (0.5), a weak coupling - coupling creation inequality is obtained. This yields our last, third contribution: for the Kac's $N$-particle system and with respect to an exchangeable version of the quadratic Wasserstein distance, we proved **uniformly in** $N$ an **"almost exponential" upper bound on the trend to equilibrium**, with a trade-off between power-law convergence, and higher moments ($> 2$) of velocity distributions in constants.

The latter result may be compared to some famous examples, counter-examples and conjectures in kinetic theory. Bobylev and Cercignani in [13], and Villani in [83] proved that moment dependence and power law behavior are typically necessary and sufficient, when quantifiying return to equilibrium with the **entropy dissipation** method. Carlen and Lu in [24], using the special Wild's expansion for Maxwell molecules, proved that some form of higher ($> 2$) moment dependence in the return to equilibrium is in fact inherent to the kinetic theory with Maxwell molecules. In particular, the latter results strongly support the idea that in the context of our result (Maxwell molecules and quantification of distance to equilibrium with a form of *euclidean* Wasserstein distance), strict contractivity **cannot be** satisfied, and some sort of weakened version should not come as a surprise.

In our study, we gave some counter-examples proving the sharpness of our result (using the equality case in (0.5)) for the *specific* considered coupling[3].

## Publications

I list here my publications by types, decomposing the publications in peer-reviewed journals into two categories, depending on whether the material has been produced or substantially initiated during my PhD, or afterwards. Publications from works done or substantially initiated during my PhD are referred to as [Pxx], while the subsequent contributions are listed as [Hxx]. Monographs are referred to as [Bxx]. The contributions summarized in the present thesis are distinguished with an asterisk in front ($*$).

### Submitted works

[H1] ($*$) M. Rousset, A $N$-uniform quantitative Tanaka's theorem for the conservative Kac's $N$-particle system with Maxwell molecules. *HAL preprint*, **01020012**, 2014.

### Publications from works initiated after the PhD defense

[H2] ($*$) M. Rousset and G. Samaey, *Simulating individual-based models of bacterial chemotaxis with asymptotic variance reduction*, *M3AS, Mathematical Models and Methods in Applied Sciences*, **23**, p. 2155-2191, 2013.

[H3] ($*$) M. Rousset and G. Samaey, *Individual-based models for bacterial chemotaxis in the diffusion asymptotics*, *M3AS, Mathematical Models and Methods in Applied Sciences*, **23**, p. 2005-2037, 2013.

[H4] ($*$) T. Lelièvre, M. Rousset and G. Stoltz, *Langevin dynamics with constraints and computation of free energy differences*, *Math. Comput.*, **81**, p. 2071-212, 2012.

[3] Studying large time behavior with coupling, is, by essence, an upper bound method only.

[H5] (∗) M. Rousset, *On a probabilistic interpretation of shape derivatives of Dirichlet ground-states with application to Fermion nodes. ESAIM - Mathematical Modelling and Numerical Analysis*, **44**, 5, 2010.

[H6] T. Goudon and M. Rousset, *Stochastic Acceleration in an Inhomogeneous Time Random Force Field*, *Appl. Math. Res. Express*, **1**, p. 1-46, 2009.

[H7] (∗) P. Plechac and M. Rousset, *Implicit Mass-Matrix Penalization of Hamiltonian dynamics with application to exact sampling of stiff systems*, *SIAM MMS*, **8**, 2, 2009.

[H8] T. Lelièvre, M. Rousset and G. Stoltz, *Long time convergence of the Adaptive Biasing Force method*, *Nonlinearity*, **21**, p. 1155-1181, 2008.

[H9] T. Lelièvre, M. Rousset and G. Stoltz, *Computation of free energy profiles with parallel adaptive dynamics*, *J. Chem. Phys.*, **126**, 13, 2007.

**Book**

[B10] T. Lelièvre, M. Rousset and G. Stoltz, Free Energy Computations: A Mathematical Perspective (Imperial College Press, 2010)

**Publications from works completed or initiated before or during my PhD**

[P11] T. Lelièvre, M. Rousset and G. Stoltz, *Computation of free energy differences through non-equilibrium stochastic dynamics: the reaction coordinate case*, *J. Comp. Phys.*, **222** (2), p. 624-643, 2007.

[P12] M. Rousset and G. Stoltz, *Equilibrium sampling from non-equilibrium dynamics*, *J. Stat. Phys.*, **123** (6), 1251-1272, 2006.

[P13] M. Rousset, *On the control of an interacting particle approximation of Schrödinger ground-states*, *SIAM J. Math. Anal.*, **38** (3), 824-844, 2006.

[P14] M. Rousset, *Sur la rigidité de polyèdres hyperboliques en dimension 3: cas de volume fini, cas hyperidéal, cas fuchsien*, *Bull. SMF*, **132**, 233-261, 2004.

# Part II

## Presentation of some results

# 1

# Thermostatted (classical) molecular dynamics



## 1.1 The framework

*Foreword (On references)*

The literature in computational physics related to numerical methods in molecular dynamics is extremely vast. We refer to the monograph [57] (written by the author in collaboration with T. Lelièvre and G. Stoltz) as a reasonable summary of relevant references. We however recall mathematical references in the present dissertation.

### 1.1.1 Thermostatted classical systems

We consider a classical $N$-body (molecular) systems, with positions denoted $q \in \mathcal{M}$ lying in a smooth manifold $\mathcal{M}$. For simplicity, in the present exposition, the latter manifold is initially a periodic 3 dimensional box $\mathcal{M} = \mathbb{T}^{3N}$. Then mass-weighted coordinates are considered, that is to say we measure lengths associated with each atom $i = 1 \cdots N$ with a scale proportional to the square root of their respective mass $m_i^{1/2}$. As a consequence, with such a convention, the kinetic energy of the system is given by $\frac{1}{2}|p|^2$ where $p \in \mathbb{R}^{3N}$ is the momentum; the standard Euclidean norm and associated scalar product being respectively denoted as usual with $(\cdot, |\ |)$. The interaction energy potential is supposed to be a given smooth function $V : \mathbb{R}^{3N} \to \mathbb{R}$. Finally, the Hamiltonian of the system is thus of the form

$$H(q,p) \overset{\text{def}}{=} \frac{1}{2}|p|^2 + V(q),$$

with phase-space the co-tangent bundle $T^*\mathcal{M} \simeq \mathbb{T}^{3N} \times \mathbb{R}^{3N}$. The associated (deterministic) Hamilton's equations of motion then read

$$\begin{cases} dQ_t = P_t\,dt & = \partial_p H(Q_t, P_t)\,dt \\ dP_t = -\nabla V(Q_t)\,dt & = -\partial_q H(Q_t, P_t)\,dt \end{cases} \tag{1.1}$$

In practice, $V$ may be computed either via an *ab-initio* calculation of the electronic structure of the molecular system, or on the contrary, may be given as an *effective* potential (obtained from some *fitting, inductive* procedure). Note that electronic structure calculations requires the expensive numerical estimation of groundstates of Fermionic Schrdinger operators in $\mathbb{R}^{3M}$, where $M$ is the number of electrons.

Stochastic perturbations (of Gaussian type) are then considered. The latter are given, for each $q \in \mathbb{T}^{3N}$, in the form of a linear diffusion on momenta (an Orstein-Uhlenbeck process) whose stationary probability distribution is the kinetic Gaussian distribution $\propto \mathrm{e}^{-\beta\frac{1}{2}|p|^2}\,dp$ in $\mathbb{R}^{3N}$:

$$dP_t = -\gamma(q)P_t + \sigma(q)\,dW_t. \tag{1.2}$$

In the above, $\gamma(q)$ is assumed to be a strictly positive symmetric smooth tensor satisfying the fluctuation-dissipation relation:

$$\gamma = \frac{\beta}{2}\sigma\sigma^T. \tag{1.3}$$

When stationary ($\mathrm{Law}(P_t) = \mathrm{Gauss}(0, \mathrm{Id}), \quad \forall t$), the probability distribution of the path $(P_t)_{t\geqslant 0}$ of process (1.2) is both (i) reversible, that is to say invariant by time reversal, (ii) invariant by momenta reversal ($P_t \to -P_t$).

The process (obtained by combining from the Hamilton's equations (1.1) and the Gaussian fluctuation-dissipation (1.2)) is often called a Langevin process and reads

$$\begin{cases} dQ_t = P_t\,dt \\ dP_t = -\nabla V(Q_t)\,dt - \gamma(Q_t)P_t + \sigma(Q_t)\,dW_t. \end{cases} \tag{1.4}$$

It satisfies the following basic properties:

(i) It is ergodic[1], in the sense that for any initial condition $(Q_0, P_0) = (q_0, p_0)$ and bounded test function $\varphi$:

$$\lim_{t\to+\infty} \frac{1}{t}\int_0^t \varphi(Q_t, P_t)\,dt = \int_{T^*\mathcal{M}} \varphi(q,p)\mu(dqdp) \text{ a.s.},$$

with respect to the unique stationary probability distribution given by the *canonical* distribution:

$$\mu \overset{\text{def}}{=} \frac{1}{Z}\mathrm{e}^{-\beta H(q,p)}\,dqdp.$$

In the above, $Z$ is the normalization constant required to obtain a probability, and $dqdp$ is the *phase-space, or Liouville* reference positive measure induced by the symplectic structure of $T^*\mathcal{M}$ (the Lebesgue measure in the chosen canonical $(q,p)$ coordinates).

(ii) If stationary, the probability distribution of the whole process (in path space) is time reversible *up to momenta reversal* ($p \to -p$), which means that for any $T \geqslant 0$ the probability distribution of $(Q_t, P_t)_{t\in[0,T]}$ and $(Q_{T-t}, -P_{T-t})_{t\in[0,T]}$ are the same.

The proof of the latter assertion are a direct consequence of (i) Liouville's property of phase-space volume conservation, which follows from the symplecticity of Hamilton's equations, (ii) time reversibility up to momenta reversal of the Hamiltonian dynamics, which follows from the

[1] Technical assumptions of regularity and growth at infinity of $V$ and $\sigma$ are required.

invariance of the Hamiltonian $\frac{1}{2}|p|^2 + V(q)$ under momenta reversal, (iii) time reversibility up to momenta reversal of stationary Orstein-Uhlenbeck processes.

### 1.1.2 Numerical motivation: the timescale problem

As already mentioned, the main difficulty in many numerical computations of realiostic molecular systems is the presence of *separated time scales*. In this dissertation, we consider the case where the fastest or the slowest degrees of freedom of the system have been identified, in the form of smooth, independent (see next section for a definition) functions:

$$\xi_{\text{slow/fast}} : \mathcal{M} \to \mathbb{R}^m.$$

Such special degrees appear in practical situations in the following contexts:

(i) When $\xi_{\text{slow}}$ is a given *reaction coordinate* (also called *slow, or collective variables*) parameterizing a transition between "states" of chemical or physical interest. For instance one can take for $\xi_{\text{slow}}$ (with $m = 1$) the distance between two extremal atoms of a long linear molecule. In this context, an important information of interest is the computation of the $\xi_{\text{slow}}$-marginal probability density (the "*free energy*") defined as

$$F(z) \stackrel{\text{def}}{=} -\frac{1}{\beta} \ln \int_{\mathbb{R}^{3N}} \mathrm{e}^{-\beta V(q)} \delta_{\xi_{\text{slow}}(q)-z}(dq), \tag{1.5}$$

where $\delta_{\xi_{\text{slow}}(q)-z}(dq)$ is the *conditional distribution*[2] defined by the chain rule:

$$\delta_{\xi_{\text{slow}}(q)-z}(dq)\, dz = dq,$$

$dq$ being the (mass metric, here Euclidean) Lebesgue measure.

(ii) When $\xi_{\text{fast}}$ is a given molecular constraint, such as covalent bonds, bond angles (formed by three consecutive atoms of a chain), or even torsion angles (formed by the two consecutive planes spanned by four atoms in a chain). These degrees of freedom are of vibrational nature and generates a highly oscillatory dynamics. The direct numerical intergation of the latter is thus very stiff, with time step resctricted to the timescale of the fastest vibrations.
A typical simple numerical treatment is to enforce a constraint of the form $\xi_{\text{fast}}(q) = z_0$, where $z_0$ is the expected average value. The latter method may removes the stiffness by constraining the position manifold of the system into the slower manifold $\{q \in \mathcal{M} | \xi_{\text{fast}}(q) = z_0\}$, but by changing the phase-space introduces incontrolled sources of approximation.

### 1.1.3 Constrained dynamics

Hamiltonian systems can be rather easily *constrained* by using some independent position constraints defined by:

$$\xi(q) \stackrel{\text{def}}{=} (\xi_1(q), \ldots, \xi_m(q)) = z \in \mathbb{R}^m,$$

The resulting system is again an Hamiltonian system with Hamiltonian $H$, with phase-space the co-tangent bundle $T^*\Sigma_z$ where

$$T^*\Sigma_z \stackrel{\text{def}}{=} \Big\{ (q,p) \in \mathbb{T}^{3N} \times \mathbb{R}^{3N} \;\Big|\; \xi(q) = z, \quad p \cdot \nabla \xi(q) = 0 \Big\}, \tag{1.6}$$

$\Sigma_z$ being a sub-manifold of co-dimension $m$. The smoothness of the implicitly defined sub-manifold $\Sigma_z$ is ensured by the implicit funtcion theorem, as soon as the Gram tensor:

[2] NB: the latter is well-defined under the invertibility condition of the Gram tensor defined in Section 1.1.3.

$$G(q) \stackrel{\text{def}}{=} (\nabla_q \xi_a \cdot \nabla_q \xi_b)_{a,b=1\cdots m} \in \mathbb{R}^{m\times m}$$

is invertible for any configuration $q \in \Sigma_z$. Constrained Hamilton's equations of motion associated with the pair $(H, T^*\Sigma_z)$ can be derived (*e.g.* using the associated Lagrangian), and yields (we refer to [4, 56] for mathematical textbooks dealing with constrained Hamiltonian dynamics):

$$\begin{cases} dQ_t &= P_t\, dt, \\ dP_t &= -\nabla V(Q_t)\, dt + \nabla\xi(Q_t)\, d\Lambda_t, \\ \xi(Q_t) &= z. \qquad (C_q) \end{cases} \tag{1.7}$$

In the above $t \mapsto \Lambda_t \in \mathbb{R}^m$ is a uniquely defined Lagrange multiplier associated with the constraint $(C_q)$ and can be computed explicitly to be

$$d\Lambda_t = f_{\text{rgd}}(Q_t, P_t)\, dt$$

where the *constraining force* reads (matrix product notation is implicitly used in $\mathbb{R}^m$)

$$f_{\text{rgd}}(q,p) = G^{-1}(q)\nabla\xi(q)\cdot\nabla V(q) - G^{-1}(q)\text{Hess}_q(\xi)(p,p) \in \mathbb{R}^m. \tag{1.8}$$

Deriving the constraint through time $\frac{d}{dt}\xi(Q_t)$ yields the hidden momenta constraint

$$P_t \cdot \nabla\xi(Q_t) = 0. \qquad (C_p)$$

The system (1.7) is usually simulated using the so-called RATTLE scheme, an extension of the Störmer-Verlet scheme: with constraints:

$$\begin{cases} P^{n+1/2} = P^n - \dfrac{\Delta t}{2}\nabla V(Q^n) + \nabla\xi(Q^n)\,\Lambda^{n+1/4}, \\ Q^{n+1} = Q^n + \Delta t\, P^{n+1/2}, \\ P^{n+1} = P^{n+1/2} - \dfrac{\Delta t}{2}\nabla V(Q^{n+1}) + \nabla\xi(Q^{n+1})\,\Lambda^{n+3/4}, \\ \xi(Q^{n+1}) = z, \quad P^{n+1}\cdot\nabla\xi(Q^{n+1}) = 0, \qquad (C_q, C_p) \end{cases} \tag{1.9}$$

where typically the (non-linear) position constraint $(C_q)$ with associated Lagrange multiplier $\Lambda^{n+1/4}$ is enforced using a Newton iterating procedure. This typically adds a restriction on the time step size due to the non-linearity of $\xi$. The following facts are well established in the literature (see Chapter VII.1 in [47], or Chapter 7 in [56] for more precisions and historical references):

**Lemma 1.1.** *The RATTLE scheme* (1.9) *can be characterized by the following properties*

*(i) It is explicit with respect to force computations $\nabla_q V$,*
*(ii) It is invariant by combined reversal of time ($n \leftrightarrow n+1$) and momenta ($P \leftrightarrow -P$) (time symmetry),*
*(iii) Its flow defines a symplectic map on $T^*\Sigma_z$, which is variational with discrete Lagrangian*

$$\mathcal{L}_{\Delta t}(q^n, q^{n+1}) \stackrel{\text{def}}{=} \frac{\left|q^n - q^{n+1}\right|^2}{2\Delta t^2} - \frac{1}{2}V(q^n) - \frac{1}{2}V(q^{n+1}), \qquad \xi(q^n) = \xi(q^{n+1}) = z,$$

*and momentum map set by*

$$\begin{aligned} p^n &= \nabla_{q^n}\mathcal{L}_{\Delta t}(q^{n-1}, q^n) + \nabla\xi(q^n)\,\Lambda^{n-1/4} \\ &= -\nabla_{q^n}\mathcal{L}_{\Delta t}(q^n, q^{n+1}) - \nabla\xi(q^n)\,\Lambda^{n+1/4}. \end{aligned} \tag{1.10}$$

The latter properties explain the overwhelming practical sucess of this scheme. Let us recall that by "variational" it is meant that the positions evolution $(q^{n-1}, q^n) \to (q^n, q^{n+1})$ can be exactly written as the Euler-Lagrange equations extremizing the above discrete Lagrangian. Then the momenta solution of (1.7) defines a discrete version of the Legendre transformation from the Lagrangian viewpoint to the Haimltonian viewpoint. Because of this structure, such schemes are generically "symplectic", meaning that the numerical flow on the phase-space $T^*\Sigma_z$ preserves the natural symplectic form induced by the cotangent structure. This has the consequence that the Liouville phase-space measure[3] (induced by the symplectic differential form) on $T^*\Sigma_z$, and denoted

$$\sigma_{T^*\Sigma_z}(dqdp)$$

is also conserved by the latter numerical flow. Let us denote by $\Phi$ the numerical flow of the RATTLE scheme (1.9). The time-symmetry and volume conservation yields the so-called "detailed balance condition" up to momenta reversal:

$$\begin{aligned}\sigma_{T^*\Sigma_z}(dq^n dp^n)\delta_{\Phi(q^n,-p^n)}\left(dq^{n+1}dp^{n+1}\right)\\ = \sigma_{T^*\Sigma_z}\left(dq^{n+1}dp^{n+1}\right)\delta_{\Phi(q^{n+1},-p^{n+1})}\left(dq^n dp^n\right).\end{aligned} \tag{1.11}$$

To obtain stochastic thermostatting with mechanical constraints, it is useful to consider constrained stochastic Gaussian processes as follows:

$$dP_t = -\gamma_P(q)P_t\,dt + \sigma_P(q)\,dW_t, \tag{1.12}$$

where the tensors $(\gamma_P, \sigma_P)$ of the submanifols $\Sigma_z$: (i) still satisfy the fluctuation/dissipation relation (1.3); (i) have their images in the co-tangent space $\mathrm{Im}(\sigma_P(q)) \subset T_q^*\Sigma_z$ for any $q \in \Sigma_z$.

The resulting process (obtained by combining from the constrained Hamilton's equations (1.7) and the constrained Gaussian perturbation (1.12)) can be called a *constrained Langevin process* and reads

$$\begin{cases} dQ_t = P_t\,dt, \\ dP_t = -\nabla V(Q_t)\,dt + \nabla\xi(Q_t)\,d\Lambda_t - \gamma_P(Q_t)P_t\,dt + \sigma_P(Q_t)\,dW_t, \\ \xi(Q_t) = z. \qquad (C_q) \end{cases} \tag{1.13}$$

Again, $t \mapsto \Lambda_t \in \mathbb{R}^m$ is a Lagrange multiplier adapted with the filtration of the Brownian motion, and associated with $(C_q)$. The associated stationary probability distribution is now given by

$$\mu_z \overset{\text{def}}{=} \frac{1}{Z_z}\mathrm{e}^{-\beta H(q,p)}\,\sigma_{T^*\Sigma_z}(dqdp), \tag{1.14}$$

where $Z_z$ is the normalization constant. When the process (1.13) is stationary, the probability distribution of its trajectories satisfy the same invariance under combined time and momenta reversal.

Note that the constrained fluctuation dissipation part (1.12) can be easily discretized using a mid-point scheme, which can be shown to still be reversible (both in the plain sense and up to momenta reversal) with respect to the (kinetic part of the) canonical distribution (1.14). Typical numerical schemes for langevin processes with constraints can be thus constructed from the above discussion with splitting procedures between the Hamiltonian part (1.9), and the fluctuation/disspation part (1.12).

[3] In the present context, the phase-space measure identifies with the product of the Riemann volumes induced by the (mass) metric on positions, as well as momenta.

**Remark 1.1 (Metropolis acceptance-rejection).** *It is (at least in principle) interesting to keep in mind that the detailed balance condition* (1.11) *enables to construct Metropolis acceptation/rejectance rules out of Langevin processes constructed from the RATTLE scheme (we refer to [43, 50, 59] for historical references on Metropolis for Hamiltonian systems). The latter rule works as follows, using the notation of* (1.9)*:*

*(i) Accept* $(Q^{n+1}, P^{n+1})$ *from* (1.9) *with probability* $\min\left(1, \mathrm{e}^{-\beta\left(H(Q^{n+1},P^{n+1})-H(Q^n,P^n)\right)}\right)$,
*(ii) Otherwise reject and reverse momenta by setting* $(Q^{n+1}, P^{n+1}) = (Q^n, -P^n)$.

*The resulting random transition* $(Q^n, P^n) \to (Q^{n+1}, P^{n+1})$ *now satisifies "detailed balance" (up to momenta reversal) with respect to the* exact, canonical probability *distribution* (1.14)*, and lead to schemes without time step bias.*

*However, global Metropolization of large systems suffer from* two distinct *drawbacks: (i) the acceptance rate drastically vanishes with the number of degrees of freedom, due to the fluctuations in the Metropolis weight; (ii) the global momenta reversal in the rejection step removes the potentially positive effect of inertia on mixing times.*

## 1.2 Slow degrees of freedom and free energy calculation

We consider independent slow degrees of freedom $\xi_{\mathrm{slow}} : \mathcal{M} \to \mathbb{R}^m$, as described in the last section. The **main goal** in the present section is to analyze some numerical methods **calculating the free energy** (1.5) (that is to say, essentially, the equilibrium marginal distribution) associated with the slow degree of freedom $\xi_{\mathrm{slow}}$.

### 1.2.1 Free energy calculation and constrained canonical distributions

The free energy $F : \mathbb{R}^m \to \mathbb{R}$ defined in (1.5) by the $\xi$-marginal probability density can in fact be calculated from the normalization $Z_z$ of the constrained canonical distribution (1.14). It is given through the co-area(-like) formula:

$$F(z_1) - F(z_0) = -\frac{1}{\beta} \ln \frac{Z_{z_1}}{Z_{z_0}} - \frac{1}{\beta} \ln \frac{\displaystyle\int_{T^*\Sigma_z} (\det G)^{-1/2} d\mu_{z_0}}{\displaystyle\int_{T^*\Sigma_z} (\det G)^{-1/2} d\mu_{z_1}}. \tag{1.15}$$

As a consequence, we **will focus on the computation of ratios of the form** $\frac{Z_{z_1}}{Z_{z_0}}$ , keeping in mind that differences of free energy $F(z_1) - F(z_0)$ can be obtained by evaluating (1.15).

### 1.2.2 Jarzynski-Crooks identity and 'Thermodynamic Integration" (TI)

In the present section, we will lay the emphasis on the so-called *Jarzynski-Crooks identity*. The latter may be interpreted as a time-dependent generalization of the *time reversibility* for Langevin processes as (1.4). In a sense, the other topics (the time-independent "thermodynamic integration" method to calculate free energy, the physical concepts of work, free energy) can be deduced from it.

We first recall that the reversibility property of a random process can be written as follows for any time $T$:

$$\mathbb{E}_\mu\left(\Psi_{0\to T}(Q,P)\right) = \mathbb{E}_\mu\left(\Psi_{T\to 0}(Q,-P)\right),$$

where in the above:

1. $t \mapsto (Q_t, P_t)$ is a Langevin process as (1.4) or (1.13) (reversible up to momenta reversal) with stationary probability distribution $\mu$.
2. $\mathbb{E}_\mu$ implicitly define the initial probability distribution $(Q_0, P_0) \sim \mu$.
3. $\Psi_{0\to T}$ is an appropriate functional on *path space*, for instance in the finite dimensional case:

$$\Psi_{0\to T}(x) = \Psi(x_0, x_{t_1}, x_{t_2}, \cdots, x_{t_n}, x_T),$$

where $0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant T$ is some time ladder.

A more abstract way to state reversibility is to consider the relative probability of a stationary trajectory, and of the trajectory obtained by combined momenta and time reversal. It yields:

$$\ln \frac{\mathrm{Law}_\mu\left((Q_{T-t}, -P_{T-t})_{t\in[0,T]}\right)}{\mathrm{Law}_\mu\left((Q_t, P_t)_{t\in[0,T]}\right)} = 0$$

The most direct way to state Jarzynski-Crooks identity is to assume that the potential energy $V_z(q)$ depends on an external parameter $z \in \mathbb{R}^m$, and to consider the time-dependent Langevin process

$$\begin{cases} dQ_t = P_t\, dt, \\ dP_t = -\nabla V_{z(t)}(Q_t)\, dt - \gamma(Q_t)P_t + \sigma(Q_t)\, dW_t. \end{cases} \tag{1.16}$$

Until the end of this section only, we will abuse notation and denote, as in the constrained case,

$$\mu_z(dqdp) = \frac{1}{Z_z}\mathrm{e}^{-H_z(q,p)}dqdp$$

the canonical distribution with normalization $Z_z$, and associated with the Hamiltonian $H_z(q,p) = |p|^2/2 + V_z(q)$.

The Jarzynski-Crooks identity then reads, for any $\theta \in [0,1]$:

$$\begin{aligned} &\mathbb{E}_{\mu_{z(0)}}\left(\Psi_{0\to T}\left(Q,P\right)\,\mathrm{e}^{-\theta\beta\mathcal{W}_{0\to T}(Q,P)}\right) \\ &\qquad = \frac{Z_{z(T)}}{Z_{z(0)}}\mathbb{E}_{\mu_{z(T)}}\left(\Psi_{T\to 0}\left(Q^{\mathrm{b}}, -P^{\mathrm{b}}\right)\mathrm{e}^{-(1-\theta)\beta\mathcal{W}_{T\to 0}(Q^{\mathrm{b}},P^{\mathrm{b}})}\right). \end{aligned} \tag{1.17}$$

In the above:

(i) $\mathbb{E}_\mu$ indicates the initial distribution of the process.

(ii) $t \mapsto (Q_t^{\mathrm{b}}, P_t^{\mathrm{b}})$ is solution of (1.16), but with backward time dependence, $z^{\mathrm{b}}(t) = z(T-t)$, $t \in [0,T]$ replacing $z(t)$, $t \in [0,T]$.

(iii) The *work functional* $\mathcal{W}_{0\to T}$ is actually the *energy brought to the system* by the time-dependance in the Hamiltonian, and is defined by

$$\mathcal{W}_{t_0\to t_1}(q,p) \overset{\mathrm{def}}{=} \int_{t_0}^{t_1} \left(\partial_z H_z\right)_{z=z(t)}(q_t, p_t)\,\dot{z}(t)dt.$$

It satisfies the symmetry conditions

$$\begin{aligned} \mathcal{W}_{t_0\to t_1}(q,p) &= \mathcal{W}_{t_0\to t_1}(q,-p) \\ &= -\mathcal{W}_{t_1\to t_0}(q,p). \end{aligned}$$

It may be very enlightening to rewrite (1.17) (taking $\theta = 1$) as the relative probability of trajectories between, (i) the process (1.16) with backward time dependence $z^{\mathrm{b}}(t) = z(T-t)$, $t \in [0,T]$, and transformed by combined momenta and time reversal, (ii) the original time-dependent process (1.16). It yields

$$\boxed{\ln \frac{\mathrm{Law}_{\mu_{z(T)}}\left((Q^{\mathrm{b}}_{T-t}, -P^{\mathrm{b}}_{T-t})_{t\in[0,T]}\right)}{\mathrm{Law}_{\mu_{z(0)}}\left((Q_t,P_t)_{t\in[0,T]}\right)}(q,p) = -\ln \frac{Z_{z(T)}}{Z_{z(0)}} - \beta \mathcal{W}_{0\to T}(q,p).} \tag{1.18}$$

Note that ratios of normalizations of canonical distributions can then be computed $\frac{Z_{z(T)}}{Z_{z(0)}}$ using the identity (1.17) for $\Psi_{0\to T} \equiv 1$.

**Remark 1.2 (Second law of thermodynamics).** *By applying Jensen inequality in* (1.17)*, one obtains the "free energy version" of the second law of thermodynamics for any thermostatted Hamiltonian systems, namely:*

$$\mathbb{E}\left(\mathcal{W}_{0\to T}\left(Q,P\right)\right) \geqslant -\frac{1}{\beta}\ln\frac{Z_{z(T)}}{Z_{z(0)}}. \tag{1.19}$$

*The latter amounts to say that the expected dissipated work is always greater than the "free energy" variation defined by*

$$\Delta F \stackrel{\mathrm{def}}{=} -\frac{1}{\beta}\ln\frac{Z_{z(T)}}{Z_{z(0)}}.$$

*To relate the latter with the usual second law of thermodynamics, one can then assume that the time-dependence is stopped at $T$ and set $z(T+h) = z(T)$, $h > 0$. Then, we can define the random heat received by the system from the thermostat by using conservation of energy:*

$$\mathcal{Q}_{0\to T+h}(Q,P) \stackrel{\mathrm{def}}{=} \left(H_{z(T)}(Q_{T+h},P_{T+h}) - H_{z(0)}(Q_0,P_0)\right) - \mathcal{W}_{0\to T}\left(Q,P\right);$$

*and the random entropy received by the system from the thermostat is identified as in classical thermodynamics (up to the Boltzmann constant) with the exchanged heat:*

$$S_{\mathrm{exch},0\to T+h}(Q,P) \stackrel{\mathrm{def}}{=} \beta \mathcal{Q}_{0\to T+h}(Q,P).$$

*Then (i) averaging, (ii) taking the limit $h \to +\infty$ using the mixing property of Langevin processes, and (iii) using the inequality* (1.19) *yields*

$$\mathbb{E}\left(S_{\mathrm{exch},0\to+\infty}\right) \leqslant \beta\left(\Delta U - \Delta F\right) \stackrel{\mathrm{def}}{=} \Delta S,$$

*where the energy is defined as the average energy under canonical distribution $U(z) =^{\mathrm{def}} \mathbb{E}_{\mu_z}(H_z(Q,P))$. Thus Jarzynski relation can be interpreted as a quantitative and stochastic version (and also valid for small systems) of the usual second law of thermodynamics applied to equilibrium Gibbs states.*

Jarzynski's identity can also be seen as a time-dependent version of the *thermodynamic integration* method (TI), which amounts here to the direct differential identity

$$\frac{d}{dz}\left(-\frac{1}{\beta}\ln Z_z\right) = \int_{T^*\mathcal{M}} \left(\partial_z H_z\right)(q,p)\,\mu_z(dqdp), \tag{1.20}$$

which simply states that the derivative of free energy is the canonical average of the associated *virtual work* exerted on the system by a change of the considered parameters $dz$.

*Sketch of proof of* (1.17)

Since it can be explained very quickly, we detail the idea behind Jarzynski-Crooks formula. We assume for simplicity that the system is isolated (no thermostat ), and that time is discrete.

Denote by $\Phi_{z_{n+\frac{1}{2}}}$ the Hamiltonian flow of (1.16) between time $t_n$ and time $t_{n+1}$ with parameter $z_{n+1/2}$ (kept fixed). Consider the "detailed balance" (1.11) of Hamiltonian flows which can be rewritten here with energy conservation

$$\mathrm{e}^{-\beta H_{z_{n+1/2}}(q^n,p^n)} dq^n dp^n \delta_{\Phi_{z_{n+1/2}}(q^n,-p^n)}(dq^{n+1}dp^{n+1}) = \mathrm{e}^{-\beta H_{z_{n+1/2}}(q^{n+1},p^{n+1})} dq^{n+1}dp^{n+1}\delta_{\Phi_{z_{n+1/2}}(q^{n+1},-p^{n+1})}(dq^n dp^n).$$

then defining the physical work to be

$$\mathcal{W}_{n,n+1} = H_{z_{n+1/2}}(q^n,p^n) - H_{z_n}(q^n,p^n) + H_{z_{n+1}}(q^{n+1},p^{n+1}) - H_{z_{n+1/2}}(q^{n+1},p^{n+1}),$$

we exactly obtain a discrete, one step version of the Jarzynski-Crooks relation (1.17). The time continuous case is obtained formally by taking $\Delta t \to 0$. The case with thermostat coupling is obtained by using the detailed balance relation satisfied by the Orstein-Uhlenbeck process (1.2) on momenta.

### 1.2.3 Results

Assume now that the constrained Langevin process (1.13), is modified using a time dependent mechanical constraint of the form:

$$\begin{cases} dQ_t = P_t\, dt, \\ dP_t = -\nabla V(Q_t)\, dt + \nabla\xi_{\mathrm{slow}}(Q_t)\, d\Lambda_t - \gamma_P(Q_t)P_t\, dt + \sigma_P(Q_t)\, dW_t, \\ \xi_{\mathrm{slow}}(Q_t) = z(t), \qquad (C_q) \end{cases} \tag{1.21}$$

the Lagrange multiplier $t \mapsto \Lambda_t$ being similarly associated to the constraint $(C_q)$. In order to simplify the presentation of the results, we assume that

$$\dot{z}(0) = \dot{z}(T) = 0. \tag{1.22}$$

The first main contribution is the **rigorous proof of Jarzynski-Crooks relation** (1.17) for the process **with constraints** (1.21). The proof is based on the generalization of time symmetry by momenta reversal and symplecticity in the case of time-dependent position constraints. The time-independent case ("thermodynamic integration") is also considered.

**Result 1.1 (Jarzynski-Crooks and TI).**

*(i) Consider the time-dependent constrained Langevin process defined by* (1.21)*, as well as the constrained canonical distribution* (1.14)*. Then (assuming 1.22) the Jarzynski-Crooks relation* (1.17) *holds true with the following two (equivalent) definitions of the work*

$$\mathcal{W}_{z(0)\to z(T)}(Q,P) \overset{\mathrm{def}}{=} \int_0^T \dot{z}(t)\, d\Lambda_t, \tag{1.23}$$

$$= \int_0^T \frac{d}{dh} H\left(\Phi_{t,t+h}(Q_t,P_t)\right) dt, \tag{1.24}$$

*where $\Phi_{t,t+h}$ denotes the deterministic flow obtained by considering an isolated system (taking $(\gamma_P,\sigma_P) = (0,0)$ in* (1.21)*). Note that* (1.23) *can be interpreted as the variation of the reaction coordinate $\xi_{\mathrm{slow}}$ times the associated "exerted force" given by the Lagrange multiplier; while* (1.24) *can be interpreted as the virtual energy variation obtained by isolating the system from the thermostat.*

*(ii) For time-independent constraints ($z(t) = cte$ in* (1.21)*, or equivalently* (1.7)-(1.12)*), we have for any $T > 0$*

$$\frac{d}{dz}\left(-\frac{1}{\beta}\ln Z_z\right) = \mathbb{E}_{\mu_z}\left(\frac{1}{T}\int_0^T d\Lambda_t\right) = \lim_{T\to+\infty}\frac{1}{T}\int_0^T d\Lambda_t \quad a.s..$$

The second main contribution was to focus on the **associated numerical schemes** (based on the RATTLE scheme (1.7)).

**Result 1.2 (Numerical schemes).**

*(i) Consider the time-dependent constrained Langevin* (1.21)*, discretized using a splitting of the form:*

*Hamiltonian part with RATTLE* (1.7)
+
*Gaussian fluctuation-dissipation part* (1.12)
*(e.g. with mid-point)*

*Then, a discretization of the work based on the energy* (1.24) *(using the notation of* (1.7)*),*

$$\mathcal{W}_{n,n+1} \overset{\text{def}}{=} H\left(Q^{n+1}, P^{n+1}\right) - H\left(Q^n, P^n\right)$$

*yields an* exact *(no time-step discretization error) discrete Jarzynski-Crooks relation* (1.17)*. A discretization of the work based on force, or Lagrange multipliers* (1.23) *(using the notation of* (1.7)*),*

$$\mathcal{W}_{n,n+1} \overset{\text{def}}{=} (z(t_{n+1}) - z(t_n)) \left(\Lambda^{n+1/4} + \Lambda^{n+3/4}\right)$$

*has an order* 2 *time-step discretization error.*

*(ii) In the time-independent case ($z(t) = cte$), only the discretization based on force, or Lagrange multipliers* (1.23) *is stable. Up to the use of a Metropolis acceptance-rejection correction*[4]*, it yields an exact (at the cost of computing* $\mathrm{Hess}(\xi_{\mathrm{slow}})$ *in the formula* (1.8) *for* $f_{\mathrm{rgd}}$*), or a (time step) second order estimate:*

$$\begin{aligned} \frac{d}{dz}\left(-\frac{1}{\beta}\ln Z_z\right) &= \mathbb{E}_{\mu_z}\left(f_{\mathrm{rgd}}(Q,P)\right) \\ &= \mathbb{E}_{\mu_z}\left(\Lambda^{n+1/4} + \Lambda^{n+3/4}\right) + O(\Delta t^2). \end{aligned} \tag{1.25}$$

Finally, the last contribution was to remark that the above setting **contains as a special case the "overdamped" limit dynamics**. The so-called "overdamped" dynamics with constraints is solution to the stochastic differential equation:

$$\begin{cases} dQ_t = -\nabla V(Q_t)\, dt + \sqrt{\dfrac{2}{\beta}}\, dW_t + \nabla\xi(Q_t)\, d\Lambda_t^{\mathrm{od}}, \\ \xi(Q_t) = z(t), \end{cases} \tag{1.26}$$

where $\Lambda_t^{\mathrm{od}}$ is an adapted stochastic process (explicitly computable) such that $\xi(Q_t) = z(t)$.

**Result 1.3 (Overdamped limit).** *Consider the constrained Langevin process* (1.21) *(in the limit $\gamma_P \to +\infty$), discretized using a splitting of the form:*

*Hamiltonian part with RATTLE* (1.7)
+
*Indepedent sampling of momenta at equilibrium.*

*It yields the explicit Euler discretization of* (1.26) *with new time-step $\Delta s \overset{\text{def}}{=} \Delta t^2/2$, and with associated new numerical properties:*

*(i) Proof of the Jarzynski-Crooks relation without time-step discretization error (time-dependent case);*

[4] Time-step bias of discretization has to be added otherwise.

*(ii) Proof of the Jarzynski-Crooks relation with variance reduced work estimator based on the Lagrange multipliers* (1.25) *of the RATTLE scheme (time-dependent case);*
*(iii) TI relation with variance reduced work estimator* (1.25) *based on the Lagrange multipliers* (1.25) *of RATTLE (time-independent case);*
*(iv) Exact sampling of constrained canonical distributions* (1.14) *using the Metropolis acceptance-rejection rule defined in Remark 1.1 (time-independent case).*

## 1.3 Mass penalization of fast degrees of freedom

The main idea of mass penalization is to **remove the highly oscillatory nature of fast degrees of freedom**, leading to larger stable time-steps and faster numerical computation, but **without modifying the statistical behavior** (here, the canonical distribution of positions). Indeed, usual, direct constraints on these fast variables do change the phase-space and thus introduce an uncontrolled source of approximation. Mass-penalization can also be used to construct in a simple way consistent integrators of Hamiltonian dynamics, with lower stiffness, by using a time-step dependent mass penalization.

In [7], we have proposed a method, based on an artificial mass penalization, to simulate molecular systems with *known* fast degrees of freedom denoted:

$$\xi \equiv \xi_{\text{fast}} \in \mathbb{R}^m.$$

This method extends in a systematic way the idea of modifying the mass-tensor of a physical system in order to slow down its highly oscillatory components (see the previous works [11, 61]).

### 1.3.1 Presentation

We consider a classical (molecular) system in the framework of Section 1.1. The mass-metric of the system is then penalized with a tensor modification given by

$$M_\nu(q) \overset{\text{def}}{=} \text{Id}_{3\text{N}} + \nu^2 \sum_{a=1}^{m} \nabla_q \xi_{a,\text{fast}} \otimes \nabla_q \xi_{a,\text{fast}}, \tag{1.27}$$

where $\nu > 0$ denotes the penalty intensity. Intuitively, and by design, the latter modification does impact the part of momenta orthogonal to the fast degrees of freedom.

When analyzing the associated canonical distribution, the position dependence of the mass-penalization (hence of the kinetic energy) introduces a bias. This bias is corrected by introducing an *effective potential*

$$V_{\text{fix},\nu}(q) \overset{\text{def}}{=} \frac{1}{2\beta} \ln\left(\det(M_\nu(q))\right). \tag{1.28}$$

The latter can be interpreted as a perturbation near $\nu = +\infty$ (large penalty) of an effective potential originally introduced by Fixman to model thermostatted highly oscillatory systems around a slow manifold, and given by

$$V_{\text{fix}}(q) \overset{\text{def}}{=} \frac{1}{2\beta} \ln\left(G(q)\right). \tag{1.29}$$

It can be shown with a determinant identity that $V_{\text{fix},\nu} \to V_{\text{fix}}$ when $\nu \to +\infty$ up to a constant.

**Definition 1.1.** *The mass-penalized Hamiltonian dynamics in $T^*\mathcal{M}$ is equivalently defined by its Hamiltonian or its Lagrangian:*

$$\begin{cases} L_{\text{IMP}}(q,v) = \frac{1}{2} v^T M_\nu(q) v - V(q) - V_{\text{fix},\nu}(q) \\ H_{\text{IMP}}(q,p) = \frac{1}{2} p^T M_\nu^{-1}(q) p + V(q) + V_{\text{fix},\nu}(q) \end{cases} \tag{1.30}$$

*Moreover, we have the consistent position marginal probability distribution*

$$\int_{T_q^*\mathcal{M}} \mathrm{e}^{-\beta H_{\mathrm{IMP}}(q,p)} dqdp \propto \mathrm{e}^{-\beta V(q)} dq.$$

To avoid difficulties related to the (geometric) integration of *non-separable* Hamiltonians, the key point here is to use an implicit representation of the mass-penalty using the extended phase-space $T^*(\mathcal{M} \times \mathbb{R}^m)$ together with $m$ position constraints. The extended Lagrangian of the dynamics is thus a function of the extended tangent space, and reads for $(q, z, v, v_z) \in T(\mathcal{M} \times \mathbb{R}^m)$

$$\boxed{\begin{cases} L_{\mathrm{IMP}}(v, v_z, q, z) = \frac{1}{2}|v|^2 + \frac{1}{2}|v_z|^2 - V(q) - V_{\mathrm{fix},\nu}(q), \\ \xi_\nu(q, z) = 0, \qquad (C_\nu) \end{cases}} \tag{1.31}$$

where in the above we have defined

$$\xi_\nu(q, z) = \xi_{\mathrm{fast}}(q) - \frac{z}{\nu}.$$

The constraints $(C_\nu)$ are applied in order to identify the auxiliary variables and the fast degrees of freedom $\xi_{\mathrm{fast}}$ with a coupling intensity tuned by $\nu$. The position constraint $(C_\nu)$ also implies that the velocities satisfy $v \cdot \nabla\xi_{\mathrm{fast}}(q) = \frac{1}{\nu} v_z$, so that the Lagrangian (1.31) is equivalent to the Lagrangian (1.30) in original tangent space $T(\mathcal{M})$.

In the same way, the original mass-penalized Hamiltonian (1.30) can be interpreted as a constrained Hamilonian system in co-tangent space $T^*\{(q, z) \in \mathcal{M} \times \mathbb{R}^m | \xi_\nu(q, z) = 0\}$ which reads

$$\begin{cases} H_{\mathrm{IMP}}(p, p_z, q, z) = \frac{1}{2}|p|^2 + \frac{1}{2}|p_z|^2 + V(q) + V_{\mathrm{fix},\nu}(q), \\ \xi_\nu(q, z) = 0. \qquad (C_\nu) \end{cases} \tag{1.32}$$

The extended Hamilonian system with constraints obtained by this mean, can then be coupled to a thermostat, using a Langevin equation of the form (1.7)-(1.12), which yields a stochastically perturbed dynamics that samples the equilibrium canonical distribution (with marginal in position independent of the penalty $\nu$). It is defined by the following equations of motion

$$\begin{cases} dQ_t = P_t dt \\ dZ_t = P_{z,t} \\ dP_t = -\nabla V(Q_t)dt - \nabla V_{\mathrm{fix},\nu}(Q_t)dt - \gamma P_t dt + \sigma dW_t - \nabla\xi_{\mathrm{fast}}(Q_t)\, d\Lambda_t \\ dP_{z,t} = -\gamma_z P_{z,t} dt + \sigma_z dW_{z,t} + \frac{1}{\nu} d\Lambda_t \\ \xi_{\mathrm{fast}}(Q_t) = \frac{Z_t}{\nu} \qquad (C_\nu) \end{cases} \tag{1.33}$$

The process $t \mapsto (W_t, W_{z,t}) \in \mathbb{R}^{3N+m}$ is a standard multi-dimensional Brownian motion, $\gamma$ (resp. $\gamma_z$) a $3N \times 3N$ (resp. $m \times m$) non-negative symmetric dissipation matrix, assumed to be position independent for simplicity. The fluctuation- dissipation is supposed to hold: $\sigma\sigma^T = \frac{2}{\beta}\gamma$ (resp. $\sigma_z\sigma_z^T = \frac{2}{\beta}\gamma_z$). The processes $t \mapsto \Lambda_t \in \mathbb{R}^m$ are Lagrange multipliers associated with the constraints $(C_\nu)$ and adapted with the noise filtration. It can be checked that the stochastic process with constraints (1.33) is equivalent to a Langevin diffusion in $T^*\mathcal{M}$ with the mass-penalized Hamiltonian (1.30), and a dissipation tensor given by

$$\gamma_\nu(q) = \gamma + \nu^2 \nabla\xi_{\mathrm{fast}}(q)\, \gamma_z \nabla^T \xi_{\mathrm{fast}}(q).$$

The typical time scale of the fast degrees of freedom is thus enforced using the penalty $\nu$.

### 1.3.2 Results

The main contributions of the present section consist then in proving the following facts:

(i) In the limit of vanishing penalization ($\nu = 0$), the original dynamics enables the construction of dynamically consistent smoothed numerical schemes by taking a time-step dependent penalization, for instance satisfying $\nu(\Delta t) = O_{\Delta t}(1)$.
(ii) In the limit of infinite penalization, **the fast degrees of freedom are frozen**, and the dynamics is a standard effective constrained dynamics on the "slow" manifold $\Sigma_{\xi_{\text{fast}}(Q_0)} = \{q | \xi_{\text{fast}}(q) = \xi_{\text{fast}}(Q_0)\}$ prescribed by the initial condition.
(iii) For "stiff" potentials with stiffness parameter $1/\varepsilon$, a penalization of order $O(\frac{1}{\varepsilon})$ yields an effective dynamics on the associated slow manifold.
(iv) Numerical integrators can be obtained through a simple modification of the standard RATTLE integrator (1.9) for constrained Hamiltonian systems, with equivalent computational complexity. The latter is then **asymptotically stable**, in the sense that setting $\nu = 0$ or $\nu = +\infty$ yield consistent and stable discretization of the respective continuous limiting dynamics.
(v) Numerical tests on large alkane molecules (with penalized covalent bonds, bond angles, and torsion angles) illustrates the forementioned properties.

The first main result is the fact that the method yields an interpolated dynamics between the original, exact one, and the constrained one.

**Result 1.4 (Exact/Constrained Interpolation).**

*(i) Consider the constrained Langevin process* (1.33). *When $\nu \to 0$, the process $\{P_t, Q_t\}_{t \geqslant 0} \in T^*\mathcal{M}$ converges in the sense of probability distributions on continuous paths towards the original constrained Langevin process* (1.13).
*When $\nu \to +\infty$; if the initial condition satisfies*

$$\lim_{\nu \to +\infty} |\nu \left(\xi_{\text{fast}}(Q_{t=0}) - z_0\right)| < +\infty, \tag{1.34}$$

*and the Gram tensor $G$ is invertible in a neighborhood of $\Sigma_{z_0}$, then the process converges in the sense of probability distributions on continuous paths towards a decoupled constrained Langevin dynamics:*

$$\begin{cases} dQ_t = P_t dt \\ dZ_t = P_{z,t} \\ dP_t = -\nabla V(Q_t)dt - \nabla V_{\text{fix}}(Q_t)dt - \gamma P_t dt + \sigma dW_t - \nabla \xi_{\text{fast}}(Q_t)\, d\Lambda_t \\ dP_{z,t} = -\gamma_z P_{z,t} dt + \sigma_z dW_{z,t} \\ \xi_{\text{fast}}(Q_t) = z_0. \qquad (C_q) \end{cases} \tag{1.35}$$

*In particular, the process $\{P_t, Q_t\}_{t \geqslant 0} \in T^*\Sigma_{z_0}$ is a constrained Langevin process on co-tangent space $T^*\Sigma_{z_0}$ with effective Hamiltonian $\frac{1}{2}|p|^2 + V(q) + V_{\text{fix}}(q)$.*

*(ii) Numerical schemes constructed from a RATTLE discretization* (1.9) *of the Hamiltonian part and splitting with the thermostat part are "asymptotic preserving" in the two limits $\nu \to 0$ and $\nu \to +\infty$, in the sense that they approach the limiting continuous processes* **without stability restriction on $\nu$** .

We have then extend the study in the case where the potential energy $V$ has the highly oscillatory form

$$V_\varepsilon(q) = U(q, \frac{\xi_{\rm fast}(q) - z_0}{\varepsilon}), \tag{1.36}$$

with a confining assumption for instance of power-law form ($c_1, c_2, \alpha > 0$)

$$\inf_{q\in\mathbb{R}^d} U(q,y) \geqslant c_1 + c_2\,|y|^\alpha \xrightarrow{y\to+\infty} +\infty.$$

The infinite stiffness limit ($\varepsilon \to 0$) of highly oscillatory dynamics has been studied in a series of papers [16, 68, 69, 71, 75, 75, 81]. The limiting dynamics can be fully characterized in special cases, for instance through adiabatic effective potentials. When the system is thermostatted, one can postulate an "ad hoc" effective dynamics (see [69]) that we have recovered in the result below. We obtain:

**Result 1.5 (Highly oscillatory limit).**

*(i) Consider a constrained Langevin process* (1.35)*, with highly oscillatory potential* (1.36) *and initial condition of the form* (1.34)*. Setting a $\varepsilon$-dependent penalization*

$$\nu = \nu(\varepsilon) \underset{\varepsilon\to 0}{\sim} \frac{1}{\varepsilon},$$

*the latter process converges to an effective Langevin process on $T^*(\Sigma_{z_0}\times\mathbb{R}^m)$ with effective Hamiltonian*

$$\left\{\begin{array}{l} H_{\rm eff}(p,p_z,q,z) = \frac{1}{2}\,|p|^2 + \frac{1}{2}\,|p_z|^2 + U(q,z) + V_{\rm fix}(q) \\ \xi_{\rm fast}(q) = z_0. \quad (C) \end{array}\right. \tag{1.37}$$

*Explicitly:*

$$\left\{\begin{array}{l} dQ_t = P_t dt \\ dZ_t = P_{z,t} \\ dP_t = -\nabla_1 U(Q_t,Z_t)dt - \nabla V_{\rm fix}(Q_t)dt - \gamma P_t dt + \sigma dW_t - \nabla\xi_{\rm fast}(Q_t)\,d\Lambda_t \\ dP_{z,t} = -\nabla_2 U(Q_t,Z_t)dt - \gamma_z P_{z,t}dt + \sigma_z dW_{z,t} \\ \xi_{\rm fast}(Q_t) = z_0. \qquad (C_q) \end{array}\right. \tag{1.38}$$

*(ii) The long time stationary distribution of the process $t \mapsto Q_t$ is the $\varepsilon \to 0$ limit of the canonical distribution $\propto \mathrm{e}^{-\beta V_\varepsilon(q)}dq$ associated with the highly oscillatory potential* (1.36) *(with support on the slow manifold).*

*(iii) Consider a numerical scheme constructed from a RATTLE discretization* (1.9) *of the Hamiltonian part. Then the latter is asymptotic preserving (***no stability restriction***) in the limit $\varepsilon \to 0$ with $\nu \sim_{\varepsilon\to 0} \frac{1}{\varepsilon}$.*

### 1.3.3 Example of numerical simulations

The method is simulated for the $N$-alkane model in dimension 3 (a linear chain of $N$-atoms with effective short range 2-body and 3-body potentials). The penalized fast degrees of freedom $\xi_{\rm fast}$ are the *interatomic distances*, the *bond angles*, and the *torsion angles*.

*Exact sampling for butane ($N = 4$)*

Exact sampling of the equilibrium distribution on very large times, whatever the value of the mass penalization, is shown in fig. 1.1. In the latter figure, the distribution of the butane length for

constrained bond angles is clearly distorted. Then mixing time to equilibrium is studied with a time-step defined by a 95% acceptance rate in a Metropolis acception-rejection rule. The autocorrelation function of the length evolution in terms of iteration steps is given in fig.1.1, and the faster convergence of the IMMP method is demonstrated. In the latter figure, the decorrelation time is enhanced by a factor $\sim 2$ using the mass penalization method.

*Dynamics and mixing for a large molecule*

The dynamics of the end-to-end length for a larger molecule is shown in fig. 1.2, and the associated frquency is not perturbed by the IMMP method. One can observe a small group of fast oscillations in the middle of the Verlet dynamics plot in fig. 1.2 which is not present in the IMMP case. This translates in the top of spectral plots in fig. 1.3 where a cut-off of the fastest oscillatory scales for the IMMP case occurs.

Finally, the precise ratio of the $l^2$ decorrelation time between the mass-penalized integrator and the Verlet one is given in fig. 1.4 for different system sizes, with the use of the Metropolis acception/rejection step (with a reference acceptance rate $\sim 95\%$). It increases again more than linearly in $N$, in fact exponentially here; this shows that in the present case, the **mass penalization method heals the decrease of the Metropolis rejection rate** of for large systems (see also [53]).

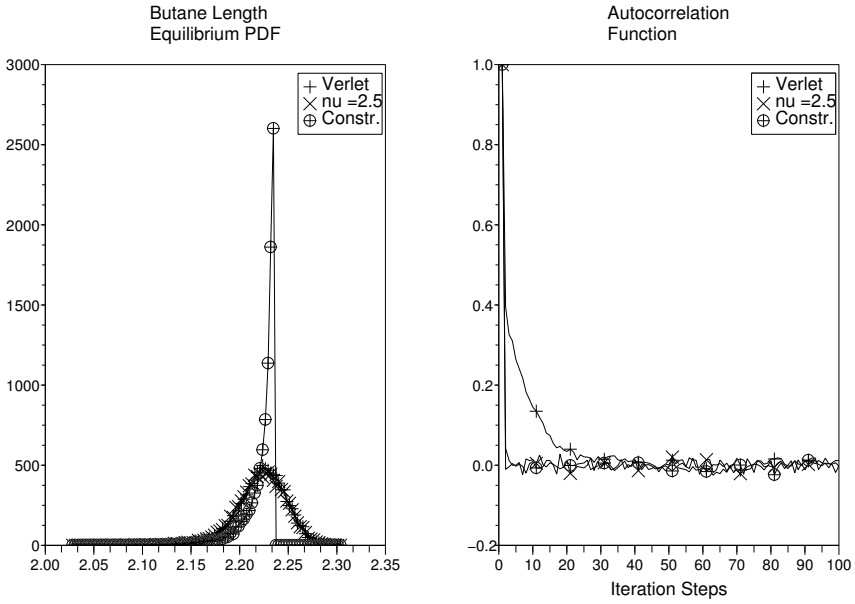


**Fig. 1.1.** Left: Equilibrium probability density of the end-to-end length of the butane molecule with a Metropolis step, using Verlet and IMMP integrator (penalty $\nu$) on the one end, and direct constraints on bonds/angle bonds (with RATTLE integrator) on the other hand. Note that the constrained integrator do not sample the appropriate measure. Right: The autocorrelation function in terms of iteration steps comparing the IMMP and the Verlet integrator. The gain in $l^2$ decorrelation time is 1.8.

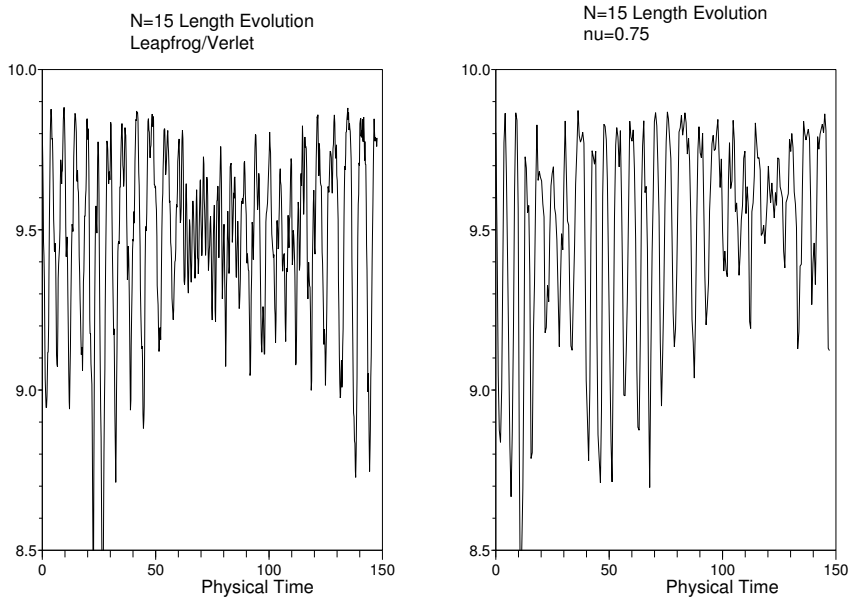


**Fig. 1.2.** The trajectory of the $N = 20$-alkane dynamics for the Verlet scheme and the IMMP scheme. Note that the IMMP penalty do not modify substantially the slow frequency/varying components.

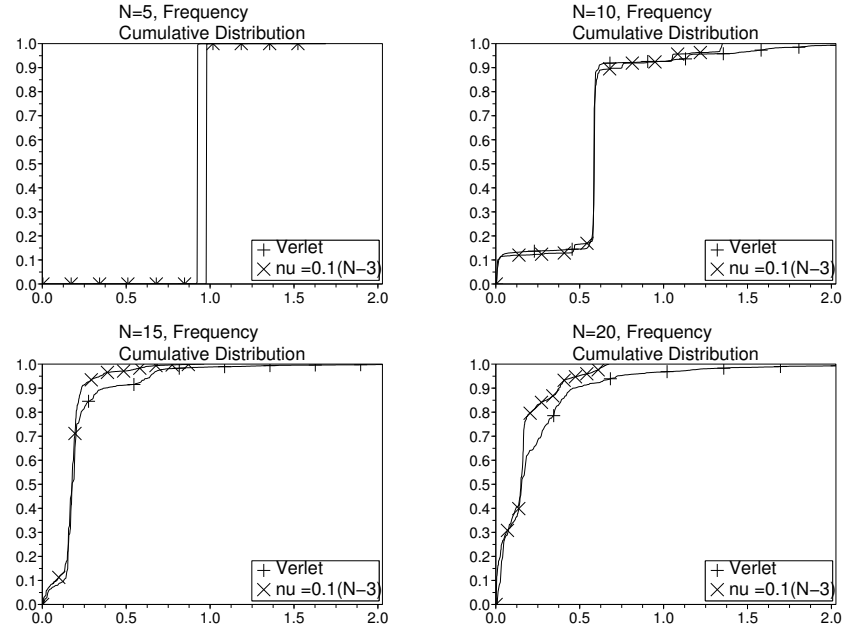


**Fig. 1.3.** Spectral densities of the end-to-en length trajectories of the alkane for $N = 5, 10, 15, 20$. Note that the IMMP penalty do not modify substantially the slow frequencies/varying components.

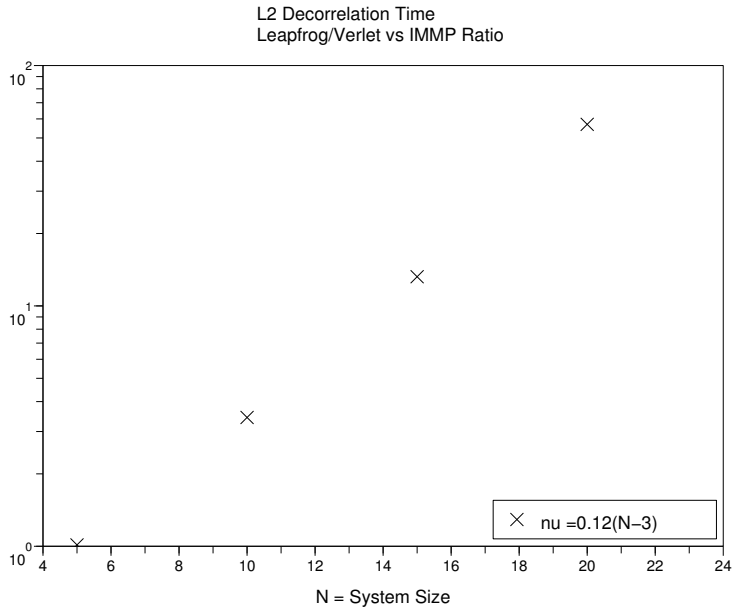


**Fig. 1.4.** Decorrelation time of the end-to-end alkane length in terms of Monte-Carlo iteration steps. The ratio between the Verlet integration and the IMMP integration is given. Note that the $y$-axis is in logarithmic scale, and an exponential gain occurs.

# 2

# Fermionic eigenstates



## 2.1 Presentation

### 2.1.1 Fermionic groundstates

In this chapter, we consider a self-adjoint[1] Schrödinger operator in the Hilbert space $L^2(\mathbb{R}^{3N})$ associated with $N$ identical particles of the form:

$$H = -\frac{\Delta}{2} + V, \tag{2.1}$$

where $V : \mathbb{R}^{3N} \to \mathbb{R}$ is a potential function assumed to be *invariant under the permutation group* $\mathcal{S}_N$ . The operator (2.1) can be defined on the sub-Hilbert space of skew-symmetric functions:

$$L^2_{\text{skew}}(\mathbb{R}^{3N}) \stackrel{\text{def}}{=} \left\{\psi \in L^2(\mathbb{R}^{3N}),\, \psi \circ S = \det(S)\psi \quad \forall S \in \mathcal{S}_N\right\},$$

where $\det(S) \in \{1, -1\}$ denotes the parity of permutation $S$ (odd for transpositions). This precisely defines the Hamiltonian of $N$ quantum particles in the usual three dimensional Euclidean space, obeying the so-called Fermi statistics, and thus called "Fermions".

We are interested in Fermionic groundstates $(\psi^*_{\text{F}}, E^*_{\text{F}})$, solutions to the minimization problem:

$$E^*_{\text{F}} =^{\text{def}} \inf \left( \frac{\displaystyle\int_{\mathbb{R}^{3N}} \psi H\left(\psi\right)}{\displaystyle\int_{\mathbb{R}^{3N}} \psi^2}, \quad \psi \in L^2_{\text{skew}}(\mathbb{R}^{3N}) \subset L^2(\mathbb{R}^{3N}) \right)$$
$$= \frac{\displaystyle\int_{\mathcal{N}_\theta} \psi^*_{\text{F}} H\left(\psi^*_{\text{F}}\right)}{\displaystyle\int_{\mathcal{N}_\theta} \left(\psi^*_{\text{F}}\right)^2},$$

[1] as standardly defined by the spectral theory of Hilbertian unbounded operators.

and defining the quantum state of minimal energy. This state is by far the most common state of the $N$-electronic structure found in everyday molecular systems, and can be used to describe the vast majority of associated chemical processes (quantum chemistry).

Introducing a standard Wiener process (Brownian motion)

$$t \mapsto W_t \in \mathbb{R}^{3N},$$

with a given initial distribution $\mathrm{Law}(W_0)$ that may vary through the section. We recall that the semi-group $\mathrm{e}^{-tH}$ associated with the operator (2.1) on the Hilbert space $L^2(\mathbb{R}^{3N})$

$$P_t^V \stackrel{\mathrm{def}}{=} \mathrm{e}^{-tH}$$

has a Feynman-Kac representation formula given by ($\varphi$ is a measurable test function, say with compact support):

$$P_t^V(\varphi)(x) \stackrel{\mathrm{def}}{=} \mathbb{E}_{W_0=x}\left(\varphi(W_t)\,\mathrm{e}^{-\int_0^t V(W_s)\,ds}\right). \tag{2.2}$$

The well-posedness of the latter Feynman-Kac formula, as well as the self-adjointness of $H$, are the basic minimal assumptions necessary to work in such a context, and fortunately are valid for a large class of potentials which includes the physical Coulombian ones. For instance, if $V = V^+ - V^-$, both properties can be justified for instance for $V^+ \in L^2_{\mathrm{loc}}(\mathbb{R}^{3N})$, and $V^- \in L^2(\mathbb{R}^{3N}) + L^\infty(\mathbb{R}^{3N})$; see *e.g.* [74].

**Remark 2.1 (The sign problem).** *Computing directly $(\psi_F^*, E_F^*)$ using Monte-Carlo methods is one instance of untractable problems of computational physics known as the* sign problems. *The latter can be summarized as follows. Assuming discreteness of the spectrum of $H$, and indexing the eigenvalues with the countable index set $I$, symmetry implies that the associated eigenfunctions of $H$ in $L^2(\mathbb{R}^{3N})$, denoted $(E^{(n)}, \psi^{(n)})_{n\in I}$, are either symmetric or skew-symmetric functions. Computing* (2.2) *by a probabilistic method is in principle possible as follows. Consider*

*(i) a* **non-symmetric** *positive initial distribution $W_0 \sim \psi_{\mathrm{init}}(x)dx > 0$ in $L^2(\mathbb{R}^{3N})$ with $\int_{\mathbb{R}^{3N}} \psi_{\mathrm{init}} = 1$,*

*(ii) a* **skew-symmetric** *(say, continuous) test functions denoted by $\psi_{\mathrm{skew}} \in L^2_{\mathrm{skew}}(\mathbb{R}^{3N}) \cap C(\mathbb{R}^{3N})$,*

*and compute with some Monte-Carlo method the average*

$$\mathbb{E}\left(\psi_{\mathrm{skew}}(W_T)\mathrm{e}^{-\int_0^T V(W_s)ds}\right) = \sum_{n\in I} \mathrm{e}^{-E^{(n)}T} \int_{\mathbb{R}^{3N}} \psi^{(n)}\psi_{\mathrm{init}} \int_{\mathbb{R}^{3N}} \psi^{(n)}\psi_{\mathrm{skew}}. \tag{2.3}$$

*Since $\psi_{\mathrm{skew}}$ is skew-symmetric, $\int_{\mathbb{R}^{3N}} \psi^{(n)}\psi_{\mathrm{skew}} = 0$ for* **symmetric** *eigenstates $\psi^{(n)}$. As a consequence, for large times, the dominant term in* (2.3) *is exactly given by the Fermionic ground-state(s) as defined by* (2.2). *Comparing the result obtained for different test function $\psi_{\mathrm{skew}}$ enables in principle to extract all the information on the Fermionic groundstate.*

*Unfortunately, it is obvious that the average typical error in the latter calculation, given by $\mathbb{E}\left(|\psi_{\mathrm{skew}}(W_T)|\,\mathrm{e}^{-\int_0^T V(W_s)ds}\right)$, is now dominated by the* **symmetric or Bosonic** *groundstate, which identifies with the* **full** *groundstate with energy $E^*_{\mathrm{B}} < E^*_{\mathrm{F}}$, since $|\psi_{\mathrm{skew}}|$ is a symmetric function.*

*Thus one is compelled when computing* (2.3) *with Monte-Carlo estimators, to generate some spurious variance associated with the symmetric eigenfunctions of the bottom of the spectrum. These terms have vanishing bias but an exponentially dominating variance when $T$ is large.*

*Worse than that, a simple classical computation shows that the ratio $\frac{E^*_{\mathrm{B}}}{E_{\mathrm{F}}}$ vanishes with the number $N$ of particles, for instance linearly in the case of the harmonic oscillator. As a conse-*

*quence, the exponential explosion of the relative variance has a stronger rate for larger systems and quickly gets out of range of any reasonable computer simulation. This forms the sign problem.*

### 2.1.2 Fixed nodal domains

In practice, $(\psi_{\mathrm{F}}^*, E_{\mathrm{F}}^*)$ is computed using an **analytical parametrization of skew-symmetric functions**, and a numerical optimization procedure associated to the minimization problem (2.2). For instance, the parametrization is given as a finite sum of physically meaningful determinants[2] multiplied by a strictly positive symmetric function (called the *Jastrow factor*). Such optimization problems are central to computational Quantum Chemistry, which forms a huge scientific field. We refer to [20] for a mathematical introduction with a consequent bibliography. See also the following two typical papers [79, 80] involving wave function optimization using a Monte-Carlo method. Monte Carlo methods in computational Quantum Chemistry are usually referred to as Quantum Monte Carlo (QMC) methods.

We thus now assume given a family of skew-symmetric functions

$$\left\{\psi_\theta^{\mathrm{I}}\right\}_{\theta\in\mathbb{R}^p},$$

with an explicit analytical expression, which includes $\psi_{\theta_0}^{\mathrm{I}}$ (the *trial wave function*) obtained with a preliminary optimization scheme.

The latter are used to define the *nodal domain*

$$\mathcal{N}_\theta \stackrel{\mathrm{def}}{=} \mathcal{N}_\theta^+ \cup \mathcal{N}_\theta^-,$$

where:

$$\begin{cases} \mathcal{N}_\theta^+ \stackrel{\mathrm{def}}{=} \left\{x\in\mathbb{R}^{3N} \,|\, \psi_\theta^{\mathrm{I}}(x) > 0\right\} \\ \mathcal{N}_\theta^- \stackrel{\mathrm{def}}{=} \left\{x\in\mathbb{R}^{3N} \,|\, \psi_\theta^{\mathrm{I}}(x) < 0\right\} \\ \partial\mathcal{N}_\theta = \left\{x\in\mathbb{R}^{3N} \,|\, \psi_\theta^{\mathrm{I}}(x) = 0\right\}. \end{cases} \tag{2.4}$$

We will assume that $\mathcal{N}_\theta$ satisfy the *tiling property* for $\theta$ in a neighborhood of $\theta_0$, which means that the set $\mathcal{N}_\theta/\mathcal{S}_N$ obtained by identifying points using the symmetric group $\mathcal{S}_N$ is a *connected open set*. Fortunately, it can be shown that the nodal domain of Fermionic groundstates indeed do satisfy the tiling property under mild assumptions [19].

The *Fixed Node Approximation* (FNA) then consists in computing with a Monte-Carlo method the Dirichlet groundstate $(\psi_\theta^*, E_\theta^*)$ of the variational problem associated with the Hamiltonian $H$, but with *Dirichlet boundary conditions* on $\partial\mathcal{N}_\theta$. The latter operator will be denoted

$$-\frac{1}{2}\Delta_{\mathcal{N}_\theta} + V,$$

and assumed to be self-adjoint on the Hilbert space $L^2_{\mathrm{skew}}(\mathbb{R}^{3N})$. The associated energy minimization problem then reads

[2] called Slater determinants, especially when constructed from explicit eigenfunctions.

$$
\begin{aligned}
E_\theta^* &\stackrel{\text{def}}{=} \inf \left( \frac{\displaystyle\int_{\mathbb{R}^{3N}} \psi H(\psi)}{\displaystyle\int_{\mathbb{R}^{3N}} \psi^2}, \quad \psi|_{\partial\mathcal{N}_\theta} = 0, \quad \psi \in L^2_{\text{skew}}(\mathbb{R}^{3N}) \right) \\
&= \frac{\displaystyle\int_{\mathbb{R}^{3N}} \psi_\theta H(\psi_\theta^*)}{\displaystyle\int_{\mathbb{R}^{3N}} (\psi_\theta^*)^2} \\
&\geqslant E_F^*,
\end{aligned}
\tag{2.5}
$$

where the last inequality follows from the fact that $E_F^*$ is solution to the variational problem without Dirichlet boundary condition.

### 2.1.3 Stopped Feynman-Kac formula and long time behavior

We will now recall some classical extension of the Feynman-Kac formula. We first denote the first exit time of the domain $\mathcal{N}_\theta$ by

$$
\tau \stackrel{\text{def}}{=} \inf (t \geqslant 0 | W_t \in \partial\mathcal{N}_\theta) . \tag{2.6}
$$

**Lemma 2.1.** *Assume that: (i) $\mathcal{N}_\theta$ define a regular[3] domain, (ii) $V$ is smooth and going to infinity at infinity, $-\frac{1}{2}\Delta_{\mathcal{N}_\theta} + V$ is self-adjoint in $L^2_{\text{skew}}(\mathbb{R}^{3N})$. Let $\varphi$ be a smooth function with compact support, and $\lambda \in \mathbb{R}$. The classical solution $(t,x) \mapsto h_t(x)$ of the parabolic problem with inhomogeneous Dirichlet conditions*

$$
\begin{cases}
\partial_t h_t(\varphi) = \left( \dfrac{1}{2}\Delta - V + \lambda \right) (h_t(\varphi)) = 0 \\
h_t(\varphi)|_{\partial\mathcal{N}_\theta} = \varphi|_{\partial\mathcal{N}_\theta}.
\end{cases}
\tag{2.7}
$$

*has the probabilistic representation*

$$
h_t(\varphi)(x) = \mathbb{E}_x \left( \varphi(W_{t\wedge\tau}) \mathrm{e}^{-\int_0^{t\wedge\tau} (V(W_s)-\lambda)\mathrm{d}s} \right). \tag{2.8}
$$

*If $\lambda < E_\theta^*$ then the following stationary representation holds*

$$
h_\infty(\varphi)(x) = \mathbb{E}_x \left( \varphi(W_\tau) \mathrm{e}^{-\int_0^{\tau} (V(W_s)-\lambda)\mathrm{d}s} \right). \tag{2.9}
$$

The above assumptions are not optimal, but the proof in [5] is classical, and does not require advanced potential theory. See [54] for the classical treatment, and *e.g.* [74] for possible generalizations.

On the other hand, the Feynman-Kac formula for the operator $-\frac{1}{2}\Delta_{\partial\mathcal{N}_\theta} + V$ can be easily derived by interpreting the Dirichlet boundary conditions as an infinite potential $V = +\infty$ outside the domain $\partial\mathcal{N}_\theta$. This leads to a Feynman-Kac representation similar to (2.2):

$$
P_t^{\theta,V}(\varphi)(x) = \mathbb{E}_x \left( \varphi(W_t) \, \mathrm{e}^{-\int_0^t V(W_s)\, ds} \mathbb{1}_{t\leqslant\tau} \right),
$$

if

$$
P_t^{\theta,V} = \mathrm{e}^{-t\left(-\frac{1}{2}\Delta_{\partial\mathcal{N}_\theta} + V\right)},
$$

and as soon as $\mathcal{N}_\theta$ is sufficiently regular. Here again see [74] for some possible generalizations.

It is then possible to study the long-time behavior of the latter semi-group, as a direct consequence of the spectral theorem for self-adjoint operators:

[3] with some usual cone condition (no cusp) enabling to define traces of functions, integration by parts, and so on.

**Lemma 2.2.** *Assume that $\mathcal{N}_\theta$ satisfy the tiling property, $-\frac{1}{2}\Delta_{\partial\mathcal{N}_\theta}+V$ has a spectral gap, and that the signed groundstate $\psi^*_\theta \in L^2(\mathbb{R}^{3N})$ is integrable ($\psi^*_\theta \in L^1(\mathbb{R}^{3N})$). Then, for any initial distribution* Law$(W_0)$ *with support in $\mathcal{N}^+_\theta$, the long time probability distribution of the Feynman-Kac semi-group conditioned to remain in the domain $\mathcal{N}^+_\theta$ (the so-called* Quasi-Stationary Distribution *(QSD) ), has a probability density function exactly given by $\psi^*_\theta$:*

$$\lim_{T\to+\infty} \frac{\mathbb{E}\left(\varphi(W_T)\mathbf{l}_{T\leqslant\tau}\mathrm{e}^{-\int_0^T V(W_s)ds}\right)}{\mathbb{E}\left(\mathbf{l}_{T\leqslant\tau}\mathrm{e}^{-\int_0^T V(W_s)ds}\right)} = \frac{\displaystyle\int_{\mathcal{N}^+_\theta}\varphi\,\psi^*_\theta \mathrm{d}x}{\displaystyle\int_{\mathcal{N}^+_\theta}\psi^*_\theta \mathrm{d}x}. \tag{2.10}$$

*The exponential rate of the evolution of the weighted extinction probability yields the groundstate energy:*

$$\lim_{T\to+\infty} -\frac{1}{T}\ln\mathbb{E}\left(\mathbf{l}_{T\leqslant\tau}\mathrm{e}^{-\int_0^T V(W_s)ds}\right) = E^*_\theta. \tag{2.11}$$

**Remark 2.2 (Variants with drift, Population Monte-Carlo).** *The probabilistic interpretations* (2.10)-(2.11) *may be modified using a change of probability (sometimes called Doob's transformations). In the latter case, a drift is added to the Wiener process, and the range of the potential in the Feynman-Kac weight $\mathrm{e}^{-\int_0^T V(W_s)ds}$ may be reduced. These transformations may be interpreted within spectral theory of self-adjoint operators as unitary transformations obtained by products with functions: for a strictly positive $\mathcal{S}_N$-*symmetric *function $\psi_{\mathrm{B}} \in L^2_{\mathrm{sym}}(\mathbb{R}^{3N})$, $\psi_{\mathrm{B}} > 0$, we can consider the unitary mapping:*

$$\left(H = \frac{\Delta}{2}+V, L^2\left(\mathbb{R}^{3N},dx\right)\right) \simeq \left(\psi_{\mathrm{B}}^{-1}H(\psi_{\mathrm{B}}\,.\,), L^2\left(\mathbb{R}^{3N},\psi^2_{\mathrm{B}}(x)dx\right)\right), \tag{2.12}$$

*and all the discussion of this chapter still holds:*

*(i) by replacing the Brownian motion $t\mapsto W_t\in\mathbb{R}^{3N}$ (Markov process with generator $\frac{\Delta}{2}$); by a drifted diffusion $t\mapsto X_t$ with generator of the form $L=^{\mathrm{def}}\frac{\Delta}{2}+\nabla\ln\psi_{\mathrm{B}}\cdot\nabla$;*
*(ii) replacing the potential $V$ by the new potential $U=^{\mathrm{def}}V-\psi_{\mathrm{B}}^{-1}\frac{\Delta}{2}(\psi_{\mathrm{B}})$.*

*A different case appears when the importance function is chosen with* skew-symmetry *(the latter is now denoted $\psi^{\mathrm{I}}_\theta$), and vanishes on the fixed nodal domain $\partial\mathcal{N}_\theta := (\psi^{\mathrm{I}}_\theta)^{-1}(\{0\})$. It generates a singular drift on the nodal boundary $\partial\mathcal{N}_\theta$, and under some technical conditions on the (nicely bounded) behavior of new potential $U$ (see [19]), the mapping similar to* (2.12) *is unitary but with* **added Dirichlet conditions**

$$\left(H_{\mathcal{N}_\theta} = \frac{\Delta_{\mathcal{N}_\theta}}{2}+V, L^2\left(\mathbb{R}^{3N},dx\right)\right) \simeq \left((\psi^{\mathrm{I}}_\theta)^{-1}H(\psi^{\mathrm{I}}_\theta\,.\,), L^2\left(\mathbb{R}^{3N},(\psi^{\mathrm{I}}_\theta)^2(x)dx\right)\right), \tag{2.13}$$

*In any case, the latter methods may leads to Monte-Carlo methods with some* importance sampling *variance reduction defined by the addition of a drift which can efficiently compute the Bosonic groundstate $(\psi^*_{\mathrm{B}},E^*_{\mathrm{B}})$ with* (2.12)*, or the* **Fixed Node** *groundstate $(\psi^*_\theta,E^*_\theta)$ with* (2.13)*. This method has been widely used and studied in many fields. In chemistry, this is the essence of for instance of the so-called Diffusion Monte-Carlo (DMC) method under the Fixed Node Approximation (see [5, 7, 19, 26, 28, 48]).*

*In such methods, a set of "replicas" (or "clones", or "particles") of the considered process are simulated, and a* selection, birth-death *step has to be carried out at regular time intervals, according to the Feynman-Kac weights associated with each replica. The latter step enables to avoid the degeneracy of the weights. We refer the reader to [41, 42] for applications in Bayesian statistics (referred to as "Sequential Monte-Carlo" methods), and to [19, 32–34, 70] for the associated mathematical analysis.*

## 2.2 Nodal shape derivation of the Fixed Node energy

### 2.2.1 Context in computational chemistry

The issue of **optimizing the nodes** $\partial\mathcal{N}_\theta$ of the trial wave function $\psi_\theta^{\mathrm{I}}$ with respect to the parameters $\theta \in \mathbb{R}^p$ in the fixed node approximation in order to **minimize the Dirichlet energy** $E_\theta^*$ in (2.5) was pointed out in [27], where a long discussion on the structure of Fermion nodes and appropriate (from this perspective) trial wave functions is provided. This problem remains a partly unsolved problem and motivates the material presented in this dissertation.

Note however that methods to approximately compute the gradient $\nabla_\theta E_\theta^*$ were already suggested in the literature in the more general context of the calculation of physical properties (or "forces"), (see *e.g.* [6, 8, 9, 25, 78]).

### 2.2.2 Shape derivation

The computation of the shape derivative of the fixed node groundstate

$$\nabla_\theta E_\theta^*. \tag{2.14}$$

is given by standard calculus of variations enable to compute the shape derivative of the groundstate energy through the formula:

$$\nabla_\theta E_\theta^* = -\frac{1}{2\int_{\mathbb{R}^{3N}}(\psi_\theta^*)^2}\int_{\partial\mathcal{N}_\theta}\left(\left|\nabla^+\psi_\theta^*\right|^2 - \left|\nabla^-\psi_\theta^*\right|^2\right) r_\theta^+ \, d\sigma, \tag{2.15}$$

where in the above $\nabla^+$ (resp. $\nabla^-$) denotes the trace on $\partial\mathcal{N}_\theta$ of the gradient in $\mathcal{N}_\theta^+$ (resp. $\mathcal{N}_\theta^-$), $\sigma$ is the usual surface measure induced by the Euclidean structure $\mathbb{R}^{3N}$, and $r_\theta^+$ is the *shape derivative* as seen from $\mathcal{N}_\theta^+$. Shape derivatives are given by a smooth, compactly supported field $r_\theta^+ : \partial\mathcal{N}_\theta \to \mathbb{R}^p$ such that formally the boundary variation writes down:

$$\partial\mathcal{N}_{\theta+d\theta} = \left\{x + n^+(x) r_\theta^+(x)\cdot d\theta \,|\, x \in \partial\mathcal{N}_\theta\right\},$$

where $n^+(x)$ is the exterior normal vector at $x \in \partial\mathcal{N}_\theta$, pointing in $\mathcal{N}_\theta^-$.

## 2.3 Results

An exact **probabilistic representation** of the shape derivative (2.15) leading to a purely Monte-Carlo estimation is probably impossible. However, and this is main contribution of the present chapter, it possible to give an *approximate* probabilistic representation, with nonetheless the **correct sign of the shape derivative of the Fixed Node energy**. This will lead to the following characterization: the approximate nodal domain $\mathcal{N}_{\theta_0}$ is in fact the *exact nodal domain* of a Fermionic eigenstate on the full space $\mathbb{R}^{3N}$ if and only if the probably distribution of a certain weighted process killed on the boundary is *fully* $\mathcal{S}_N$*-symmetric*. In the latter case, $\nabla_\theta E_\theta^*|_{\theta=\theta_0} = 0$ for any reasonable parametrization.

The key point consists in considering the elliptic differential equation satisfied by the Feynman-Kac formula (2.9), and then integrate by parts with a careful handling of the symmetry. By this mean we relate: (i) the weighted distribution of the Wiener process at the hitting time $\tau$, when the process is **initially distributed according to the Dirichlet groundstate** , (ii) the trace of gradient of the Dirichlet groundstate $\nabla^+\psi_\theta^*$ on the nodal boundary $\partial\mathcal{N}_\theta$.

**Result 2.1.** *Under the assumptions of Section 2.1.3, the following identity holds true, for any* $\lambda < E_\theta^*$*, and any* **odd** *permutation* $p_{\mathrm{odd}} \in \mathcal{S}_N$*:*

$$\boxed{\begin{aligned}&\mathbb{E}\left((\varphi(W_\tau)-\varphi\circ p_{\mathrm{odd}}(W_\tau))\,\mathrm{e}^{-\int_0^\tau (V(W_s)-\lambda)\mathrm{d}s}\,|\,\mathrm{Law}(W_0)\propto\psi_\theta^*\mathbf{1}_{\mathcal{N}_\theta^+}\right)\\&\quad=\frac{1}{2(E_\theta^*-\lambda)\displaystyle\int_{\mathcal{N}_\theta^+}\psi_\theta^*}\int_{\partial\mathcal{N}_\theta}\varphi\left(\left|\nabla^+\psi_\theta^*\right|-\left|\nabla^-\psi_\theta^*\right|\right)\mathrm{d}\sigma.\end{aligned}} \tag{2.16}$$

*In particular the* sign of the real valued measure *defined by the hitting distribution difference* (2.16) *is the same as the Dirichlet energy shape derivative* (2.15).

Approximations[4] of $\nabla_\theta E_\theta^*$ with Monte-Carlo methods can then be, in principle, carried out by relating the formulations (2.15) and (2.16) using a (deterministic) approximation of the symmetric part of the Dirichlet groundstate gradient on the nodal boundary $\mathcal{N}_\theta$.

The claimed characterization of exact Fermionic nodal domains is then the following.

**Result 2.2.** *Assume the assumptions of Section 2.1.3 hold true. Then we have the following* equivalent *assertions:*

*(i) The hitting distribution difference* (2.16) *on the boundary $\partial\mathcal{N}_\theta$ vanishes (i.e. the weighted hitting distribution is invariant by the action of the* full *permutation $\mathcal{S}_N$, and not the alternate sub-group only).*

*(ii) The trace of the gradient of the ("fixed node") Dirichlet groundstate $\nabla^+\psi_\theta^*$ on $\partial\mathcal{N}_\theta^+$ and $\nabla^-\psi_\theta^*$ on $\partial\mathcal{N}_\theta^-$ are identical (continuity): $\nabla^+\psi_\theta^*=\nabla^-\psi_\theta^*$.*

*(iii) The ("fixed node") Dirichlet goundstate $\psi_\theta^*$ is a skew-symmetric eigenfunction of $H$ on the whole space $\mathbb{R}^{3N}$.*

*(iv) The nodal domain $\mathcal{N}_\theta$ is critical for the "fixed node" Dirichlet energy: $\nabla_\theta E_\theta^*=0$ for any (smooth) shape perturbation of the domain.*

[4] There is no known direct probabilistic expressions for $\nabla_\theta E_\theta^*$. This is a general fact, that holds for any "force" (energy derivative with respect to some parameter).

# 3

# Bacteria with internal state variables



## 3.1 Context

We consider a hierarchy of three models describing the motion of a bacterium influenced by a chemical environment (motion called "chemotaxis"). The concentration of the different chemical species (usually called "chemoattractants" in the case of attractive chemicals) influences the orientation of the bacterium. The main objective is to model and simulate the motion of bacteria.

*The internal state model*

The motion of flagellated bacteria typically consists of a sequence of running phases, during which a bacterium moves in a straight line at constant speed. The bacterium changes direction in a tumbling phase, which takes much less time than the run phase and acts as a reorientation. To bias the movement towards regions with high concentration of chemoattractants, the bacterium adjusts its turning rate to increase, resp. decrease, the probability of tumbling when moving in an unfavorable, resp. favorable, direction [2, 73]. Since many species are unable to sense chemoattractant gradients reliably due to their small size, this adjustment is often done via an intracellular mechanism that allows the bacterium to retain information on the history of the chemoattractant concentrations along its path [18]. The resulting model can be formulated as a velocity-jump process, combined with an ordinary differential equation (ODE) that describes the evolution of an internal state that incorporates this memory effect [44, 45]. Some recent studies have assessed the biological relevance of such a model through experimental validation of travelling pulses [72]. This model will be called the model with *internal state*.

In [44,45,86], such models have been shown to formally converge (under appropriate timescale separation, see below) to a drift-diffusion equation (satisfied by the density of bacteria in space), the parameters of internal dynamics appearing in the expression for the chemotactic sensitivity. Existence/long time behavior results when the model is coupled to a mean-field production of chemo-attractants are also available [17, 46].

*The gradient sensing model*

Several works have also considered the motion of a bacterium to be governed by a velocity-jump process with a gradient sensing rate [1,65,67]. These models have a drift-diffusion limit similar to the internal state model, see e.g. [31, 49, 66].

*The drift-diffusion model*

In chemotaxis, the most standard description is obtained by neglecting velocity and internal variables, and considering the bacterial position density on large space and time scales. One then postulates an advection-diffusion equation for this bacterial position density, in which a chemotactic sensitivity coefficient incorporates the effect of chemoattractant concentrations gradients on the density fluxes. When coupled to a model for chemoattractant production by the bacteria, this assumption leads to the classical (non-linear) Keller–Segel drift-diffusion equations (see [55], and [51, 52] for numerous historical references).

## 3.2 Models

The precise mathematical description of the three models mentioned above are given in the present section.

### 3.2.1 Dimensional analysis

We consider the following two dimensional parameters of the problem:

(i) A typical length $l_s > 0$ of the chemoattractant concentration variations (which we assume similar for all the different species of chemoattractants);
(ii) A typical time $t_\lambda > 0$ between two changes of the bacterium velocity direction (tumbling).

The speed $v_0 > 0$ of a bacterium being assumed to be constant, it is possible to consider the adimensional parameter

$$\varepsilon := \frac{t_\lambda v_0}{l_s}. \tag{3.1}$$

We will focus on the asymptotic regime where the latter is a small parameter ($\varepsilon \ll 1$); this amounts to assume that the typical time between two velocity changes (reorientation) is much smaller than the typical time on which we can observe the macroscopic motion of the bacteria (diffusive regime).

The models below will be presented directly in dimensionless form. The position and (rescaled to unity) velocity of bacteria are denoted with adimensional variables

$$(x, v) \in \mathbb{R}^d \times \mathbb{S}^{d-1},$$

where $d \geqslant 1$ is the space dimension, and $\mathbb{S}^{d-1}$ is the associated unit sphere.

The probability distribution of new jumps after reorientation is denoted $\mathcal{M}(dv)$, it is the same for the internal state and the gradient sensing model, and it is assumed to be centered

$$\int_{\mathbb{S}^{d-1}} v\mathcal{M}(dv) = 0.$$

The necessary information on the concentrations of the different chemoattractants at a space point $x \in \mathbb{R}^d$ is described by fields. For the internal state model, it is given by a smooth field:

$$S : \mathbb{R}^d \to \mathbb{R}^n$$

for some $n \geqslant 1$; and for the gradient sensing and drift-diffusion models by another smooth field

$$A : \mathbb{R}^d \to \mathbb{R}^d.$$

The consistency between the two descriptions in the regime $\varepsilon \to 0$ will be given below in (3.10).

### 3.2.2 Internal state kinetic model

The evolution of each bacterium is given by a *pure jump Markov process* denoted

$$t \mapsto (X_t, V_t, Y_t) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \times \mathbb{R}^n \text{ (position, velocity, internal state)}.$$

The velocity of each bacterium is modified (tumbling) at random jump times $(T_n)_{n\geqslant 1}$ that are generated via a Poisson process with a time dependent rate; the resulting stochastic evolution is then described by the pure jump Markov process:

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \varepsilon V_t \\ \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = \tau_\varepsilon^{-1}\left(S(X_t) - Y_t\right) \\ \displaystyle\int_{T_n}^{T_{n+1}} \lambda(S(X_t) - Y_t)\mathrm{d}t = \Theta_{n+1} \ (\sim \mathrm{Exp}(1)) \\ V_t = \mathcal{V}_n \ (\sim \mathcal{M}(dv)) \qquad \text{for } t \in [T_n, T_{n+1}[\ . \end{cases} \tag{3.2}$$

with initial condition $X_0, V_0 \in \mathbb{R}^d$, $Y_0 \in \mathbb{R}^n$ and $T_0 = 0$. In the above:

(i) The internal state is linearly[1] attracted by the function $S : \mathbb{R}^d \to \mathbb{R}^n$; the (exponential) rates of convergence being given by $\tau_\varepsilon \in \mathbb{R}^{n\times n}$, a symmetric positive matrix that satisfies (in the sense of symmetric matrices), for some $C, \delta > 0$:

$$\tau_\varepsilon \leqslant C\varepsilon^{\delta-1}.$$

(ii) The tumbling rate $z \mapsto \lambda(z) > 0$ is a smooth positive function bounded above and below, with Taylor expansion

$$\lambda(z) = \lambda_0 - b \cdot z + \mathcal{O}\left(|z|^k\right), \tag{3.3}$$

with $k \geqslant 2$ and $b \in \mathbb{R}^n$.

(iii) $(\Theta_n)_{n\geqslant 1}$ are i.i.d. with normalized exponential distribution Exp(1).

(iv) The new velocities $(\mathcal{V}_n)_{n\geqslant 1}$ are i.i.d. with probability distribution $\mathcal{M}(dv)$ of $\mathbb{S}^{d-1}$.

**Remark 3.1.** *We thus consider* three *different time scales:*

- *A fast time scale of order $\mathcal{O}(1)$ given by the rate of change of the velocity direction;*
- *Some at least intermediate time scales of order $\mathcal{O}(\varepsilon^{\delta-1})$ with any $\delta - 1 > -1$ given by the internal state evolution;*
- *A slow time scale $\mathcal{O}(\varepsilon^{-1})$ given by the evolution of the chemoattractant concentration as seen from the bacteria.*

[1] In [3], a technical non-linear generalization with a non-symmetric linear part $\tau_\varepsilon$ is provided.

The probability density of the velocity-jump process then evolves according to a kinetic equation, in which, besides position and velocity, the internal state appears as additional variables. The associated operator is given by the adjoint of the following Markov generator of (3.2), defined for test functions $\varphi \in C_c^\infty\left(\mathbb{R}^d \times \mathbb{S}^{d-1} \times \mathbb{R}^n\right)$ by:

$$\mathcal{L}_\varepsilon \varphi \stackrel{\text{def}}{=} \varepsilon v \cdot \nabla_x \varphi + \left(\tau_\varepsilon^{-1}\left(S(x)\right) - y\right) \cdot \nabla_y \varphi + \lambda\left(S(x) - y\right)\left(\int_{\mathbb{S}^{d-1}} \varphi \, d\mathcal{M}(dv) - \varphi\right). \tag{3.4}$$

**Remark 3.2.** *Except for small $n$ and $d = 1, 2$ a direct deterministic simulation of the density distribution of all the variables of the model is therefore prohibitively expensive. Hence, it is of interest to study the relation of this model with simplified, coarse-level descriptions of the bacteria dynamics.*

### 3.2.3 Gradient sensing kinetic model

We now turn to a simplified model, in which the internal state process (3.2), and the corresponding state variables, are eliminated. Instead, the turning rate depends directly on the chemoattractant gradient. The process with direct gradient sensing is a Markov process in position-velocity variables

$$t \mapsto (X_t^c, V_t^c) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \text{ (position,velocity)};$$

the velocity of each bacterium being here again switched at random jump times $(T_n^c)_{n \geqslant 1}$ that are generated via a Poisson process with a time dependent rate. The latter satisfies for any $n \geqslant 1$:

$$\begin{cases} \dfrac{\mathrm{d}X_t^c}{\mathrm{d}t} = \varepsilon V_t^c, \\ \displaystyle\int_{T_n^c}^{T_{n+1}^c} \lambda_\varepsilon^c(X_t^c, V_t^c)\mathrm{d}t = \Theta_{n+1} \left(\sim \mathrm{Exp}(1)\right), \\ V_t^c = \mathcal{V}_n \left(\sim \mathcal{M}(dv)\right) \quad \forall t \in [T_n^c, T_{n+1}^c[, \end{cases} \tag{3.5}$$

with initial condition $X_0, \mathcal{V}_0 \in \mathbb{R}^d$. In the above:

(i) The rate $z \mapsto \lambda_\varepsilon^c(x, v)$ is a smooth positive function bounded from above and below, satisfying

$$\lambda_\varepsilon^c(x, v) := \lambda_0 - \varepsilon\, A(x) \cdot v + \mathcal{O}\left(\varepsilon^2\right), \tag{3.6}$$

for some "gradient sensing" vector field $A : \mathbb{R}^d \to \mathbb{R}^d$ that may depend on $\nabla S$, usually as a linear combination of the columns of $\nabla S(x) \in \mathbb{R}^{d \times n}$.

(ii) $(\Theta_n)_{n \geqslant 1}$ are i.i.d. with normalized exponential distribution Exp(1),

(iii) The new velocities $(\mathcal{V}_n)_{n \geqslant 1}$ are i.i.d. with centered probability distribution $\mathcal{M}(dv)$ of $\mathbb{S}^{d-1}$.

The model (3.6) describes a large bacterium that is able to directly sense chemoattractant gradients. The associated Markov generator acts on test functions $\varphi \in C_c^\infty\left(\mathbb{R}^d \times \mathbb{S}^d\right)$:

$$\mathcal{L}_\varepsilon^c(\varphi) \stackrel{\text{def}}{=} \varepsilon v \cdot \nabla_x \varphi + \lambda_\varepsilon^c(x, v)\left(\int_{\mathbb{S}^{d-1}} \varphi \, d\mathcal{M}(dv) - \varphi\right). \tag{3.7}$$

### 3.2.4 Advection-diffusion model

The time parameter at diffusive time scale is denoted

$$\bar{t} := t\varepsilon^2.$$

The drift-diffusion equation of chemotaxis is represented with a stochastic differential equation (SDE), in the following form:

$$\mathrm{d}X^0_{\bar{t}} = \left( \frac{DA(X^0_{\bar{t}})}{\lambda_0} \mathrm{d}\bar{t} + \left( \frac{2D}{\lambda_0} \right)^{1/2} \mathrm{d}W_{\bar{t}} \right), \tag{3.8}$$

where in the above $\bar{t} \mapsto W_{\bar{t}}$ is a standard Brownian motion, $D \in \mathbb{R}^{d\times d}$ is the positive symmetric matrix defined by:

$$D \stackrel{\text{def}}{=} \int_{\mathbb{S}^{d-1}} v \otimes v \, \mathcal{M}(dv) \in \mathbb{R}^{d\times d}.$$

and $A : \mathbb{R}^d \to R^d$ is the smooth vector field defined in (3.6). The associated Markov generator is given by ($\varphi \in C^\infty_c(\mathbb{R}^d)$):

$$\mathcal{L}_0\varphi \stackrel{\text{def}}{=} (A(x) + \nabla_x) \cdot \frac{D}{\lambda_0} \nabla_x \varphi. \tag{3.9}$$

## 3.3 Results

### 3.3.1 Diffusion approximation

The first contribution of the present chapter, is to rigorously prove, using explicit probabilistic arguments, the convergence with respect to *pathwise* probability distribution (a.k.a. convergence in distribution for stochastic processes) of the position variable of two velocity-jump processes described above ((3.5)-(3.2)) towards the stochastic differential equation (3.8).

The two proofs are based on an *asymptotic expansion of the jump times* with respect to $\varepsilon > 0$ (the time between two tumble phases), and a *comparison with a simpler random walk* (a standard diffusion approximation).

**Result 3.1.** *Assume that the parameters of the internal state dynamics* (3.2) *satisfy*

*(i)* $\frac{1}{k} > \delta$ *(tamed non-linearity of the rate).*

*(ii) Uniformly on compact sets (consistency between the two chemoattractants field description* $A(x) \in \mathbb{R}^d$ *and* $S(x) \in \mathbb{R}^n$*):*

$$b \cdot \lim_{\varepsilon\to 0} \frac{\tau_\varepsilon}{\lambda_0\tau_\varepsilon + 1} \partial_{x_i} S = A^i(x) \quad i = 1 \dots d, \tag{3.10}$$

*(iii) The initial condition satisfies:*

$$|S(X_0) - Y_0| = \mathcal{O}(\varepsilon^\delta),$$

*Then, the position process* $\bar{t} \mapsto X^\varepsilon_{\bar{t}} = X_{t/\varepsilon^2}$ *at the diffusive timescale converges in distribution (for the uniform convergence topology) towards* $\bar{t} \mapsto X^0_{\bar{t}}$ *solution to the SDE* (3.8)*. The same results hold for the gradient sensing dynamics* $\bar{t} \mapsto X^{c,\varepsilon}_{\bar{t}} = X^c_{t/\varepsilon^2}$*.*

### 3.3.2 Asymptotically stable coupling

We then have estimated the coupling distance between the two velocity-jump processes, when the *same random variables* are used: (i) the same exponentially distributed seeds $(\Theta_n)_{n\geqslant 1}$ defining jump times, (ii) the same random reorientations (in $\mathbb{S}^{d-1}$) $(\mathcal{V}_n)_{n\geqslant 1}$ of velocities.

**Result 3.2.** *Under the assumptions of Result 3.1, the difference, at the diffusive timescale, between the process with internal state* (3.2) *and the coupled gradient sensing process* (3.5) *defined with the same random numbers* $(\Theta_n, \mathcal{V}_n)_{n\geqslant 1}$ *satisfies* ($\forall p \geq 1, \bar{t} \geqslant 0$)

$$\boxed{\mathbb{E}\left(\left(X_{\bar{t}/\varepsilon^2} - X^c_{\bar{t}/\varepsilon^2}\right)^p\right)^{1/p} = \mathcal{O}(\varepsilon + \varepsilon^\delta + \varepsilon^{k\delta-1}).} \tag{3.11}$$

The coupling is thus "asymptotic" in the sense that the $L^p$-distance between the two processes vanishes with $\varepsilon \to 0$ on diffusive time scales. It requires that $0 > \frac{1}{k} > \delta$ which implies that the time scale of the ordinary differential equation of the internal state in (3.2) is sufficiently fast, with a constraint coming from the non-linearity of the tumbling rate $\lambda(S(x) - y)$.

**Remark 3.3.** *The latter result may interpreted together with Result 3.1. Indeed, assuming additional convergence of p-moments in Result 3.1, we obtain the respective convergence in p-Wasserstein distance (denoted $d_{W_p}$) when $\varepsilon \to 0$ of the two probability distributions*

$$d_{W_p}\left(\mathrm{Law}(X_{\bar{t}/\varepsilon^2}), \mathrm{Law}(X^0_{\bar{t}})\right) \xrightarrow[\varepsilon\to 0]{} 0, \quad d_{W_p}\left(\mathrm{Law}(X_{\bar{t}/\varepsilon^2}), \mathrm{Law}(X^0_{\bar{t}})\right) \xrightarrow[\varepsilon\to 0]{} 0.$$

*towards the same solution of the drift-diffusion equation. Result 3.2 yields by explicit coupling a quantitative upper bound on the speed:*

$$d_{W_p}\left(\mathrm{Law}(X_{\bar{t}/\varepsilon^2}), \mathrm{Law}(X^c_{\bar{t}/\varepsilon^2})\right) = \mathcal{O}(\varepsilon + \varepsilon^\delta + \varepsilon^{k\delta-1}).$$

*The question wether a* lower bound *with a similar scale ($\varepsilon + \varepsilon^\delta + \varepsilon^{k\delta-1}$) holds remains an open problem.*

### 3.3.3 Application: asymptotic variance reduction of simulations

Due to the possibly high number of dimensions of the kinetic model with internal state, the evolution of the bacterial density away from the diffusive limit may rather be simulated using a stochastic particle method. However, a direct stochastic particle-based simulation suffers from a large statistical variance, raising the important issue of variance reduced simulation.

In this section, we will present a *numerical method based on the coupling* presented in Result 3.2, between the fine-scale model for bacteria with internal dynamics and the simpler, coarse model for bacteria with direct gradient sensing. We then show that Result 3.2 implies that the *variance reduction is asymptotic*, in the sense that the statistical variance of the method vanishes asymptotically, with upper bound on the speed.

Let us first assume that we are able to *numerically compute* the exact solution of the kinetic equation for the control, gradient sensing process (3.7), with infinite precision in space and time. The associated semi-group evolution on probability measures will be denoted:

$$\mu_0 \mapsto \left(\mathrm{e}^{\bar{t}/\varepsilon^2 \mathcal{L}^c_\varepsilon}\right)^* \mu_0 \in \mathcal{P}(\mathbb{R}^d \times \mathbb{S}^{d-1}), \tag{3.12}$$

where $\mu_0(dxdv)$ is the initial distribution. The latter will be used to as a control variate.

The algorithm of asymptotic variance reduction is then based on an ensemble of replicas (or "particles") $\left\{X^i_t, V^i_t, Y^i_t\right\}_{1\leqslant i\leqslant N, t\geqslant 0}$ evolving according to the process with internal state (3.2), with empirical distribution of positions and velocities denoted:

$$\mu^N_{\bar{t}}(dxdv) = \frac{1}{N}\sum_{i=1}^N \delta_{X^i_{\bar{t}/\varepsilon^2}, V^i_{\bar{t}/\varepsilon^2}}(dxdv).$$

In the same way, an ensemble of replicas of the control process (3.5) is considered, with empirical distribution of positions and velocities denoted:

$$\mu_{\bar{t}}^{c,N}(dxdv) = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_{\bar{t}/\varepsilon^2}^{c,i}, V_{\bar{t}/\varepsilon^2}^{c,i}}\,(dxdv)\,.$$

The coupling between the two ensembles is obtained by ensuring that both simulations use *the same random numbers* $(\Theta_n)_{n\geq 1}$ and $(\mathcal{V}_n)_{n\geq 0}$.

We also denote by $\overline{\mu}_{\bar{t}}^N$ the variance reduced measure, which will be defined by the algorithm below. Since, with increasing diffusive time, the variance of the algorithm increases due to a loss of coupling between the particles with internal state and the control particles, the variance reduced algorithm will also make use of a reinitialization time step $\overline{\delta t}$. The corresponding time instances are denoted as $\bar{t}_n = n\overline{\delta t}$.

We then use the following algorithm to advance from $\bar{t}_n$ to $\bar{t}_{n+1}$, (see also Figure 3.1) :

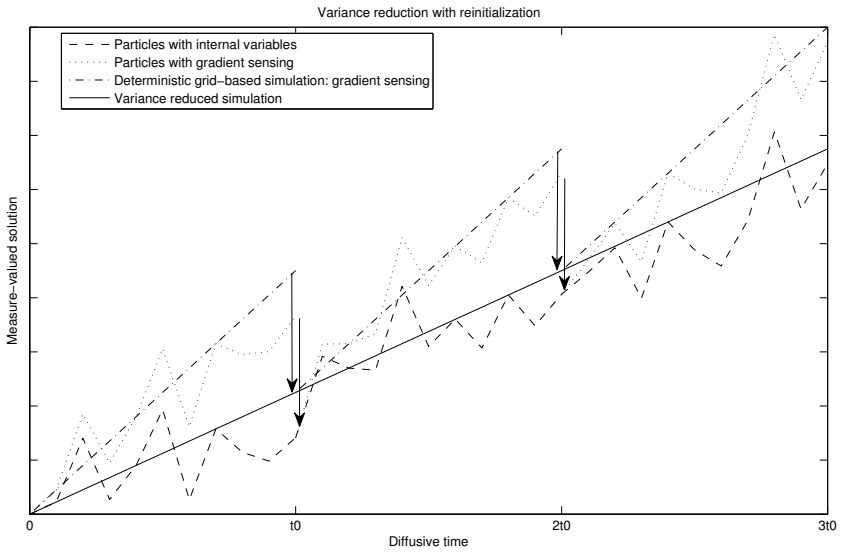


**Fig. 3.1.** A schematic description of Algorithm 3.1. The dashed line represent the evolution of $N$ bacteria with internal state. The dotted line represent the coupled evolution of $N$ bacteria with gradient sensing, subject to regular reinitializations. The dashed-dotted line is computed according to a deterministic method simulating the density of the model with gradient sensing, and subject to reinitializations at regular time intervals. The solid line is the variance reduced simulation of the internal state dynamics, and is computed by comparison.

**Algorithm 3.1** *To advance from time $\bar{t}_n$ to $\bar{t}_{n+1}$, we perform the following steps :*

*(i) Evolve the particles $\left\{X_t^i, V_t^i, Y_t^i\right\}_{i=1}^N$ from $t_n$ to $t_{n+1}$, according to* (3.2)*,*

*(ii) Evolve the particles $\left\{X_t^{i,c}, V_t^{i,c}\right\}_{i=1}^N$ from $t_n$ to $t_{n+1}$, according to* (3.5)*, using the same random numbers as for the process with internal state,*

*(iii) Compute the variance reduced evolution according to*

$$\overline{\mu}_{\bar{t}_{n+1}}^N := \mu_{\bar{t}_{n+1}}^N - \mu_{\bar{t}_{n+1}}^{c,N} + \left(\mathrm{e}^{\overline{\delta t}/\varepsilon^2 \mathcal{L}_\varepsilon^c}\right)^* \overline{\mu}_{\bar{t}_n}^N \tag{3.13}$$

*(iv) Reinitialize the* control particles *by setting*

$$X_{t_{n+1}}^{i,c} = X_{t_{n+1}}^i, \qquad V_{t_{n+1}}^{i,c} = V_{t_{n+1}}^i, \qquad i = 1, \ldots, N,$$

*i.e., we set the state of the control particles to be identical to the state of the particles with internal state.*

By construction, the variance of the algorithm is controlled by the coupling difference between the two processes so that

$$\mathrm{var}(\overline{\mu}^N_{\bar{t}_n}(\varphi)) \leqslant \sum_{k=1}^{n} \frac{\|\nabla\varphi\|^2_\infty}{N} \mathbb{E}\left(\left|X_{\bar{t}_k/\varepsilon^2} - X^c_{\bar{t}_k/\varepsilon^2}\right|^2\right),$$

and thus, from Result 3.2:

**Result 3.3.** *Algorithm 3.1 satisfies* asymptotic variance reduction*, in the sense that the variance on diffusive timescales vanishes with $\varepsilon$ for a* fixed *number of particles $N$:*

$$\mathrm{var}^{1/2}(\overline{\mu}^N_{\bar{t}_n}(\varphi)) \leqslant C\,\|\nabla\varphi\|_\infty \sqrt{n}\frac{\varepsilon + \varepsilon^\delta + \varepsilon^{k\delta-1}}{\sqrt{N}}, \tag{3.14}$$

*where in the last line, $C$ is independent of $n$, $\varepsilon$, and $N$.*

## 3.4 Example of simulation

Simulations are performed in dimension $d = 1$, with a single chemoattractant concentration, with bi-modal distribution. The simulation time is sufficiently large so that the computed distribution may be considered as a (meta-stable) stationary distribution.

### 3.4.1 Simulation without variance reduction.

First, we simulate both stochastic processes, and estimate the density of each of these processes, without variance reduction. The density is obtained via binning in a histogram, in which the grid points of the deterministic simulation are the centers of the bins. Figure 3.2 (left) shows the results

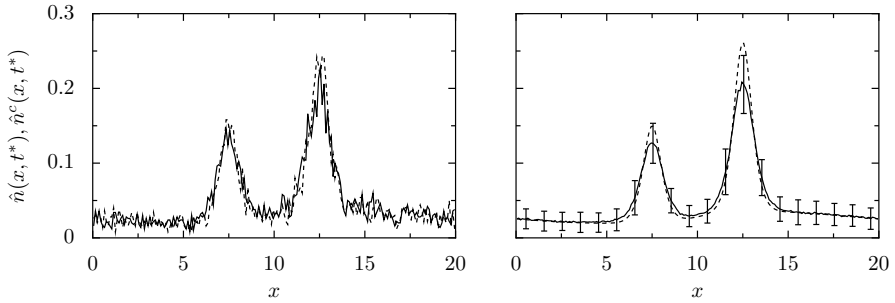


**Fig. 3.2.** Bacterial density as a function of space at without variance reduction. Left: one realization. Right: mean over 100 realizations and 95% confidence interval. The solid line is the estimated density from a particle simulation using the process with internal state; the dashed line is estimated from a particle simulation using the control process. Both used $N = 5000$ particles. The dotted line is the solution of the deterministic evolution (3.12).

for a single realization. We see that, given the fluctuations on the obtained density, it is impossible to conclude on differences between the two models. The mean densities are shown in figure 3.2 (right), which also reveals that the mean density of the control process is (almost) within the 95% confidence interval of the process with internal state. Both figures also show the density that is computed using the continuum description, which coincides with the mean of the density of the control particles.

### 3.4.2 Simulation with variance reduction.

Next, we compare the variance reduced estimation (3.13) with the density of the control PDE. We reinitialize the control particles after each coarse-scale step, i.e., each $k$ steps of the particle scheme, where $k\delta t = \delta t_{pde}$, (here $k = 1$). The results are shown in figure 3.3. We see that, using

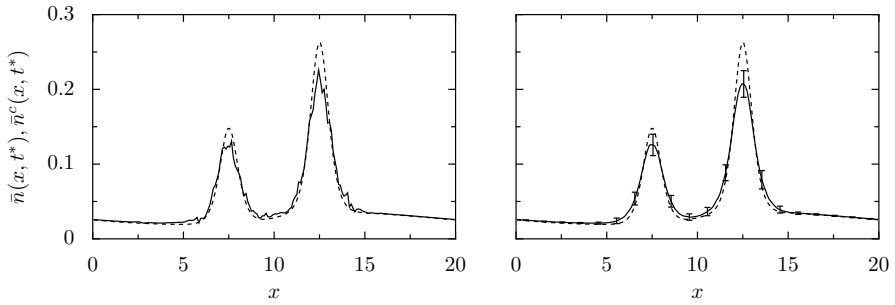


**Fig. 3.3.** Bacterial density as a function of space with variance reduction and reinitialization. Left: variance reduced density estimation of one realization with $N = 5000$ particles (solid) and deterministic solution for the control process (3.12) (dashed). Right: mean over 100 realization and 95% confidence interval (solid) and the deterministic solution for the control process (3.12) (dashed).

this reinitialization, the difference between the behaviour of the two processes is visually clear from one realization (left figure). Also, the resulting variance is such that the density of the gradient-sensing evolution is no longer within the 95% confidence interval of the variance reduced density estimation (right figure).

**Remark 3.4 (Modeling interpretation).** *We see that there is a significant difference between both models: the density corresponding to the control process is more peaked, indicating that bacteria that follow the control, gradient sensing process are* **more sensitive to sudden changes in chemoattractant** *gradient. This difference can be interpreted from the fact that the bacteria with internal state do not adjust themselves instantaneously to their environment, but instead with a delay.*

# 4

# Coupling of Boltzmann collisions (and trend to equilibrium)

## 4.1 Presentation

### 4.1.1 Random collisions

A classical (non-relativistic) elastic collision of two particles can be parametrized as follows. The pre-collisional velocities of a pair of particles are usually denoted $(v, v_*) \in \mathbb{R}^d \times \mathbb{R}^d$, and the post-collisional velocities $(v', v'_*) \in \mathbb{R}^d \times \mathbb{R}^d$. The latter are related through a (one-to-one) *conservative collision mapping*, which **conserves kinetic energy and total momentum**, and thus exactly amount to a **change of direction of two particles velocity difference**, and is denoted

$$\begin{cases} v' = \frac{1}{2}(v + v_*) + \frac{1}{2}\,|v - v_*|\,n'_v, \\ v'_* = \frac{1}{2}(v + v_*) - \frac{1}{2}\,|v - v_*|\,n'_v. \end{cases} \tag{4.1}$$

In the above, $(n_v, n'_v) = \left(\frac{v-v_*}{|v-v_*|}, \frac{v'-v'_*}{|v'-v'_*|}\right) \in \mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ denote the *pre-collisional and post-collisional directions* of velocity differences. The *scattering or deviation angle* $\theta \in [0, \pi]$ of the collision is then uniquely defined as the half-line angle between the collisional and the post-collisional direction:

$$\cos\theta \overset{\text{def}}{=} n'_v \cdot n_v \left(= \frac{v' - v'_*}{|v' - v'_*|} \cdot \frac{v - v_*}{|v - v_*|}\right).$$

Physical Galilean invariance implies that any two-body random collision is necessarily an **isotropic** random step of angle $\theta$ on the Euclidean sphere of possible collisional directions. This yields to the definition of an *isotropic probability transition on the sphere* $\mathbb{S}^{d-1}$ with scattering angle $\theta \in [0, \pi]$ as the unique probability transition

$$\mathrm{unif}_\theta(n_v, \mathrm{d}n'_v) \overset{\text{def}}{=} \mathrm{Unif}_{\{n'_v \in \mathbb{S}^{d-1} \,|\, n_v \cdot n'_v = \cos\theta\}}\,(\mathrm{d}n'_v)\,. \tag{4.2}$$

Then, the *angular collision kernel* $\beta(\mathrm{d}\theta)$ is defined as a positive Levy measure on $[0, \pi]$ (generating the random steps of the scattering angle), with normalization

$$\int_{[0,\pi]} \sin^2\theta\,\beta(\mathrm{d}\theta) = 1. \tag{4.3}$$

The latter objects enable to construct Galilean invariant collision Levy processes for two particles in $\mathbb{R}^d \times \mathbb{R}^d$, which conserve total energy and momentum. These are Markov processes with the following generator on $\mathbb{R}^d \times \mathbb{R}^d$:

$$L(\varphi)(v, v_*) \overset{\text{def}}{=} \int_{\mathbb{S}^{d-1}\times[0,\pi]} (\varphi(v', v'_*) - \varphi(v, v_*))\,\mathrm{unif}_\theta(n_v, \mathrm{d}n'_v)\,\beta(\mathrm{d}\theta). \tag{4.4}$$

Note that the generator of the isotropic *diffusion* on the sphere (called "Landau generator"), and obtained as a limit from (4.4), can also be considered (it is not done for notational simplicity). The reader may consider here only the simpler case of *angular cut-off* (*i.e.* bounded jump generator $L$):

$$b_0 := \int_0^\pi \beta(\mathrm{d}\theta) < +\infty. \tag{4.5}$$

In kinetic theory, the general full kernel is usually denoted

$$b(v - v_*, \mathrm{d}n'_v) \equiv \mathrm{unif}_\theta(n_v, \mathrm{d}n'_v)\, \beta(|v - v_*|, \mathrm{d}\theta),$$

and is called the *Boltzmann collision kernel*.

**Remark 4.1 (Maxwell collisions).** *Throughout the present chapter, Galilean invariance allows that $\beta$ may depend on the system state $(v, v_*)$ through the absolute collision speed $|v - v_*|$ (a conserved quantity). We will only consider here the case where* collision rate $\beta$ is in fact independent of $|v - v_*|$ and thus constant*, which is exactly what is called* Maxwell collisions *in kinetic theory.*

### 4.1.2 Space homogenous kinetic theory and particle systems

The *space homogenous* kinetic theory is a simplification obtained by looking at random collisions between physical particles *without considering their evolution in space*. It is then possible to focus on the associated Kac's conservative stochastic $N$-particle system. The latter is a Markov process denoted

$$t \mapsto V_t = (V_{t,(1)}, \dots, V_{t,(N)}) \in \left(\mathbb{R}^d\right)^N, \tag{4.6}$$

with distribution

$$\pi_t \stackrel{\mathrm{def}}{=} \mathrm{Law}(V_t),$$

and satisfying the conservation laws (momenta and energy), for any $t \geqslant 0$:

$$\langle V_t \rangle_N = 0 \quad \text{a.s.}, \qquad \left\langle |V_t|^2 \right\rangle_N = 1 \quad \text{a.s.}. \tag{4.7}$$

In the above, the bracket denotes the averaging over particles ($\langle\, . \,\rangle_N \equiv \frac{1}{N}\sum_{n=1}^N$). The Markov dynamics of (4.6) is specified by using , for each pair particle $(n, m)$, a two-body collision generator $L^{(n,m)}$ of (Levy jump) type (4.4); the full particle system generator then reads

$$\mathcal{L}_N \stackrel{\mathrm{def}}{=} \frac{1}{2N} \sum_{1 \leqslant n \neq m \leqslant N} L_{(n,m)}. \tag{4.8}$$

and we have (the following can be proven by a simple tensorization argument in the product space $(\mathbb{R}^d)^N$):

**Lemma 4.1.** *The particle system with generator* (4.8) *is reversible with respect to the invariant uniform probability distribution*

$$\pi_\infty = \mathrm{unif}_{\mathbb{S}^{dN-d-1}} \stackrel{\mathrm{def}}{=} \mathrm{unif}\left\{ v \in \left(\mathbb{R}^d\right)^N \;\middle|\; (4.7)\ \textit{holds} \right\}.$$

Finally, the particle system can be explicitly constructed using Grad's angular cut-off (4.5) as follows:

(i) Each particle perform a collision with a fixed rate $b_0$, and with a uniformly randomly chosen other particle.
(ii) The scattering angle of each two-body collision is independently distributed according to the probability $\dfrac{\beta(\mathrm{d}\theta)}{b_0}$.

(iii) The random post-collisional directions $n'_v$ (with scattering angle prescribed by (ii)) are sampled using the isotropic probability transition on sphere $\mathrm{unif}_\theta$.

The general case of Levy (grazing) collisions can then be considered as limits of the latter.

### 4.1.3 Kinetic equation

Under *propagation of chaos*[1], the limit of the *one body* distribution $\pi_t \in \mathcal{P}(\mathbb{R}^d)$ of the particle system satisfies formally an evolution equation in closed form (here, with a quadratic non-linearity) given by ($\varphi$ is a test function of $\mathbb{R}^d$):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\varphi\,\mathrm{d}\pi_t = \int_{\mathbb{R}^d\times\mathbb{R}^d} L\left(\varphi\otimes\mathbb{1}\right)\mathrm{d}\pi_t\otimes\pi_t, \tag{4.9}$$

When $L$ is the collision operator (4.4), then the non-linear equation (4.9) is exactly the **Boltzmann equation**[2] in $\mathbb{R}^d$ with Maxwell collision operator. The usual expression on the particle velocity density, denoted $\pi_t(dv) = f_t(v)\mathrm{d}v$, is then:

$$\frac{\mathrm{d}}{\mathrm{d}t}f_t(v) = \int_{\mathbb{R}^d\times\mathbb{S}^{d-1}\times[0,\pi]}\left(f_t(v')f_t(v'_*) - f_t(v)f_t(v_*)\right)\mathrm{d}v_*\,\mathrm{unif}_\theta(n_v,\mathrm{d}n'_v)\,\beta(\mathrm{d}\theta). \tag{4.10}$$

In the above, the collision mapping (4.1) is used implicitly, and detailed balance has been used to remove test functions.

## 4.2 Context

The main objective is then to **quantify the speed of the large time convergence of $\pi_t$ towards its equilibrium limit $\pi_\infty$.**

### 4.2.1 Convergence to equilibrium

First, let us recall standard strategies to study convergence to equilibrium for probability flows solutions of evolution (linear or non-linear) equations defined by a reversible Markovian mechanism.

(i) The *(relative) entropy dissipation* method. Denote $t\mapsto\pi_t$ a probability flow with state space $E$, expected to converge to $\pi_\infty$. The entropy method computes the variation of the relative entropy

$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\int_E \frac{\mathrm{d}\pi_t}{\mathrm{d}\pi_\infty}\ln\frac{\mathrm{d}\pi_t}{\mathrm{d}\pi_\infty}\mathrm{d}\pi_\infty}_{E(\pi_t)} = -D(\pi_t)\leqslant 0,$$

and try to obtain exponential convergence to equilibrium by obtaining a so-called *modified log-Sobolev inequality* of the form

$$E(\pi)\leqslant\frac{1}{2c_{\mathrm{ls}}}D(\pi)\qquad\forall\pi\in\mathcal{P}(E), \tag{4.11}$$

for some constant $c_{\mathrm{ls}}>0$. When $t\mapsto\pi_t$ is the distribution flow of a reversible diffusion on a Riemannian manifold, the famous curvature condition $CD(c_{\mathrm{cd}},\infty)$ of Bakry and Emery (a

[1] For a permutation symmetric $N$-particle system, we say that *propagation of chaos* holds if the marginal distribution of $k$-particles ($k$ being fixed) is converging (in law) to a product measure when $N\to+\infty$

[2] the general Boltzmann equation can then be derived formally by adding position dependence $f_t(x,v)$, transport $v\cdot\nabla_x$ from velocity, and relative velocity dependence in the collision kernel $b$.

mixture of strong convexity of the diffusion drift's potential, and uniform positive curvature of the metric, with lower bound $c_{\mathrm{cd}} > 0$, see [10] and references therein) yields such an exponential convergence by proving the inequality $0 < c_{\mathrm{cd}} \leqslant c_{\mathrm{ls}}$ by deriving entropy *two times with respect to time*. This topic has received considerable interest recently, due to Otto's gradient's flow interpretation: the probability flow of a reversible diffusion on a manifold is in fact the gradient flow of the relative entropy $E(\pi)$ with respect to the probability metric given by the quadratic Wasserstein distance $W_2$ (see the monographs [3, 84]). This has led to the interpretation of $c_{\mathrm{cd}}$ has a uniform displacement convexity constant, and yielded a conceptual explanation for the inequality $0 < c_{\mathrm{cd}} \leqslant c_{\mathrm{ls}}$.

(ii) The weaker ($0 < c_{\mathrm{ls}} \leqslant c_{\mathrm{sg}}$) *spectral gap* method which computes

$$\frac{\mathrm{d}}{\mathrm{d}t} \underbrace{\int_E \left(1 - \frac{\mathrm{d}\pi_t}{\mathrm{d}\pi_\infty}\right)^2 \mathrm{d}\pi_\infty}_{E_2(\pi_t)} = -D_2(\pi_t) \leqslant 0,$$

and try to obtain exponential convergence to equilibrium by obtaining a so-called *spectral gap inequality* of the form

$$E_2(\pi) \leqslant \frac{1}{c_{\mathrm{sg}}} D_2(\pi) \qquad \forall \pi \in \mathcal{P}(E).$$

When $t \mapsto \pi_t$ is the flow of a reversible Markov process, the latter is indeed the spectral gap of $D_2$ (the so-called Dirichlet form) seen as a self-adjoint operator in $\mathcal{L}_2(E, \pi_\infty)$.

(iii) The *Markov coupling* method, which amount to construct an explicit Markov coupling, a probabilistic coupling of two copies of the Markov process of interest which is itself again Markov:

$$t \mapsto (U_t, V_t) \in E \times E.$$

If the latter coupling contracts with respect to some distance in an average $L_p$ sense:

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}\left(d(U_t, V_t)^p\right)^{1/p} \leqslant -c_p \, \mathbb{E}\left(d(U_t, V_t)^p\right)^{1/p}$$

for any initial condition, then the method yields an upper bound on the contractivity (with constant $c_{w_p} \geqslant c_p > 0$) with respect to the related probability Wasserstein distance $W_p$. Exponential trend to equilibrium follows by taking $t \mapsto U_t$ in a stationary state: $\mathrm{Law}(U_t) = \pi_\infty$. Here again, for reversible diffusion on a manifold, the curvature condition $CD(c_{\mathrm{cd}}, \infty)$ is typically required to obtain a **contractive coupling**, obtained using the parallel transport along geodesics defined by the underlying metric. Using again the gradient's flow interpretation, it can be shown that the $CD(c_{\mathrm{cd}}, \infty)$ condition is essentially equivalent to contractivity in quadratic Wasserstein distance $c_{\mathrm{cd}} = c_{w_2}$, (see for instance [3, 85]).

**Remark 4.2.** *In practice, some more or less weakened versions of the above inequalities, especially of the modified log-Sobolev ("entropy / entropy dissipation") inequality can be obtained. They are of the form:*

$$E(\pi)^{1+1/\delta} \leqslant \frac{1}{2c_{\mathrm{ls},\delta}(\pi)} D(\pi) \qquad \forall \pi \in \mathcal{P}(E), \tag{4.12}$$

*and yields* **algebraic or power law trends** *of order $t^{-\delta}$, for $\delta \in ]0, +\infty]$ ($\delta = +\infty$ formally stands for the exponential case). $c_{\mathrm{ls},\delta}(\pi)$ is typically* **dependent on moments and regularity** *of $\pi$. In particular, in kinetic theory, "Cercignani's conjecture" refers to the exponential case $c_{\mathrm{ls},\delta=+\infty}(\pi) > 0$, where the type of dependence with respect to $\pi$ (moments, regularity) is known to be propagated by the probability flow. Usually, probabilists speak of modified log-Sobolev inequalities when $c_{\mathrm{ls}}$ is unconditionally bounded below (independent of $\pi$).*

### 4.2.2 Literature

The mathematical literature studying the convergence to equilibrium of the space homogenous Boltzmann kinetic equation, and its related Kac's conservative $N$-particle system is extremely vast, and we refer to the classical reviews [29, 82]. In the same way, the use of explicit coupling methods to study the trend to equilibrium of Markov processes (or Markov chains) is now a classical topic on its own, especially for discrete models (see *e.g.* [58]). It is also a well-established topic for continuous models, as well as for non-linear partial differential equations that have an interpretation in terms of Markovian particles. Let us mention some classical papers more closely related to the present study, with a sample of more recent references.

In kinetic theory, several types of collision rates and jumps give rise to different large time behavior. From physical scattering theory, the case (of interest here) of *constant* collision rate is called "Maxwell molecules", the case where collisions with higher relative speed are more likely is often called "hard potential case", while the case where collisions with higher relative speed are less likely is called "soft potential case". The latter case is the most badly behaved as far as large time convergence is concerned. For the type of collisional jumps the two extreme cases are the case of bounded jump kernels which is called "angular cut-off", while the purely diffusive case is called the "Landau case".

*Entropy method, $N < +\infty$.* First, the most studied method for trend to equilibrium in kinetic theory is by far the entropy method, in the case of the kinetic ($N = +\infty$) equation. Some famous counterexamples (see [12, 13, 83]) have shown that a weak entropy-entropy dissipation inequality of the form (4.12) (called "Cercignani's conjecture" in kinetic theory) cannot not hold for $\delta = +\infty$ (*i.e.* $c_{\mathrm{ls},\delta=+\infty}(\pi) = 0$), even when restricting to reasonable conditions on $\pi$ (moments, regularity,...). This counter-example contains several physically realistic collisions which includes Maxwell collisions in the angular cut-off case, of interest here.

On the other hand, a modified log-Sobolev inequality has been shown to hold for several other models, for instance for the diffusive case with Maxwell collisions in [36], or the particle system with a quadratically enhanced (by energetic collisions) collisions rate in [83]. It has been conjectured in [35] from rigorous proofs in meaningful particular examples, that a necessary and sufficient criteria for a modified log-Sobolev inequality to hold is given by a joint contribution of (i) a high probability of high energy collisions (the more likely, the faster), and (ii) the singularity of the angular kernels $\beta$ (the closer to diffusion, the faster). **For Maxwell molecules, modified log-Sobolev inequality hold for the diffusive case only**.

Meanwhile, many studies have been developed in the cases where exponential entropy convergence is known to fail, say $c_{\mathrm{ls}} = 0$. Some weakened versions of the "entropy / entropy dissipation" analysis of the form (4.12) (here is a sample: [13, 21, 30, 77, 83]) in order to obtain algebraic or power law trends with some a priori estimates on $\pi$ that has to be obtained separately.

*Spectral gap and Wild's expansion, $N = +\infty$.* For the case of interest in the present paper (Maxwell molecules), an expansion method, known as Wild's expansion, enables to give precise estimates using some refined form of the central limit theorem. It has been shown in [24], that arbitrary high moments of a velocity distribution necessarily lead to arbitrary slow decay to equilibrium (in $L^1$). In [24,38–40] a full theory of convergence to equilibrium for Maxwell molecules is then developed using Wild's method, showing that the convergence is essentially exponential with rate given by the spectral gap, but requires some moment and regularity condition on the initial condition, and a constant which is sub-optimal for short time. In [63], the case of hard potentials is treated with a spectral method that essentially prove exponential convergence with rate given by the spectral gap of the linearized near equilibrium equation, and rely on moment creation in the case of hard potentials.

*Spectral gap and entropy method $N < +\infty$.* Direct studies of the trend of equilibrium of the Kac's $N$-particle system have been undertaken [22, 23, 37, 64]. The main striking feature of the latter list is the **difficulty to achieve the so-called "Kac's program" for large time**

**behavior** (see [62] ): obtaining a **scalable ($N$-uniform) analysis** of the trend to equilibrium of the $N$-particle system. A famous result (see [22, 23]) exactly computes the *spectral gap* for Maxwell molecules, and proves that the latter is $N$-uniform ($\lim_{N\to+\infty} c_{N,\mathrm{sg}} > 0$). However, the $L_2((\mathbb{R}^d)^N, \pi_\infty)$-norm used in the spectral gap case, is usually thought to be an unsatisfactory $N$-scalable measure of trend to equilibrium (see [62], or the last section of [83] for longer discussions). By extensivity of entropy, the modified log-Sobolev constant $c_{N,\mathrm{ls}}$ is believed to be a more reliable quantity. According to [83], it is conjectured (and proven in the case of Kac's caricature) that the modified log-Sobolev constant of Kac's $N$-particle system with Maxwell molecules is of order $c_{N,\mathrm{ls}} \sim N^{-1}$, and thus not $N$-scalable.

*Coupling method, $N \leqslant +\infty$.* The use of explicit coupling methods to study the trend to equilibrium of Markov processes (or Markov chains) is now a classical topic on its own, especially for discrete models (see *e.g.* the classical textbook [58]). It is also a well-established topic for continuous models, as well as for non-linear partial differential equations that have an interpretation in terms of a Markovian mechanism. For the granular media equation (diffusive particles interacting through a smooth pairwise potential), and its related $N$-particle system, Markov coupling can give exponential trend to equilibrium, by using a "strong coupling/coupling creation inequality" (see for instance [14, 15, 60], using $CD(c_{N,\mathrm{cd}}, \infty)$-type convexity assumptions on potentials, with $\lim_{N\to+\infty} c_{N,\mathrm{cd}} > 0$). For the Kac's $N$-particle system of kinetic theory, the only paper known to us quantitatively using a Markov coupling is in [64]. In the latter, the (almost optimal, and not $N$-uniform) **estimate ($c_{N,\mathrm{w2}} \sim 1/(N \ln N)$) is obtained for Kac's caricature**, in accordance with the result cited in [83]: $c_{N,\mathrm{ls}} \sim 1/N$.

### 4.2.3 Motivation of the presented results

The main contribution described in the present chapter have been to study the contractivity of the latter coupling, uniformly in the number of particles $N$. We have developped on the Kac's particle system with Maxwell molecules a **"weak approach" of the (quadratic) coupling method, uniformly in the number of particles** $N$. The latter results extend in spirit the classical paper by Tanaka [76], where the quadratic Wasserstein distance between the solution of the kinetic equation with Maxwell collisions and the equilibrium Gaussian distribution (the Maxwellian) is shown to be decreasing through time, with a similar coupling argument, but without quantitative analysis. In a sense, the analysis in the present paper makes Tanaka's argument quantitative (with respect to time), and available for the Kac's $N$-particle system.

More precisely, we will obtain power law trends to equilibrium with respect to a permutation invariant version of the quadratic Wasserstein distance, and upon estimates on higher moments of the velocity distribution. **Up to our knowledge, this is the first time this type of estimate is obtained directly on the Kac's particle system**. Moreover, the counterexamples of Cercignani's conjecture for the entropy method in the angular's cut-off case, **motivates the moment dependence and power law behavior which are obtained in the present manuscript**. The lower bounds obtained using the Wild's expansion method support the idea that moment dependence is necessary, but that power law behavior is sub-optimal for long times.

## 4.3 Results

### 4.3.1 The Markov coupling

We have introduced an explicit symmetric Markov coupling (*i.e.* a probabilistic coupling of two copies of a Markov process which is itself again Markov) denoted

$$t \mapsto (U_t, V_t) \equiv (U_{t,(1)}, V_{t,(1)}, \dots, U_{t,(N)}, V_{t,(N)}) \in \left(\mathbb{R}^d \times \mathbb{R}^d\right)^N, \tag{4.13}$$

such that both processes $(U_t)_{t\geqslant 0}$ and $(V_t)_{t\geqslant 0}$ evolve according to the generator of the Kac's system (4.4)-(4.8), with collisions coupled using the following set of rules:

**Definition 4.1 (Simultaneous parallel coupling).** *The Simultaneous Parallel Coupling between $t \mapsto U_t$ and $t \mapsto V_t$ is obtained by the following set of rules:*

*(i) Collision times and collisional particles are the same (simultaneous collisions).*
*(ii) For each collision, the scattering angles $\theta \in [0, \pi]$ of are the same.*
*(iii) For each coupled collision, the post-collisional directions $n'_u \in \mathbb{S}^{d-1}$ and $n'_v \in \mathbb{S}^{d-1}$ are parallelly coupled: they can be obtained from each other using the elementary rotation along the great circle (the geodesic) of $\mathbb{S}^{d-1}$ joining $n_u$ and $n_v$. The resulting coupled probability is denoted*

$$\mathrm{unif}_{c,\theta}\left(n_u, n_v; \mathrm{d}n'_u \mathrm{d}n'_v\right).$$

The **sphere being a strictly positively curved manifold in dimension** $d \geqslant 3$, the latter coupling is bound to be almost surely decreasing.

We give a more explicit expression (in spherical coordinates) of the *spherical parallel coupling* of collisional directions used in (iii).

**Lemma 4.2.** *Let $(n_u, n_v) \in \mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ be given. A pair $(n'_u, n'_v) \in \mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ of post-collisional directions is* spherically coupled *if and only if, using spherical coordinates,*

$$\begin{cases} n'_u = \cos\theta\, n_u + \sin\theta \cos\varphi\, m_u + \sin\theta \sin\varphi\, l, \\ n'_v = \cos\theta\, n_v + \sin\theta \cos\varphi\, m_v + \sin\theta \sin\varphi\, l, \end{cases} \tag{4.14}$$

*where in the above $(n_u, m_u, l)$ and $(n_v, m_v, l)$ are identically oriented orthonormal sets of vectors such that $(n_u, m_u)$ and $(n_v, m_v)$ are* spanning the same plane. *Then, the image of the probability distribution*

$$\sin^{d-3}\varphi \, \frac{\mathrm{d}\varphi}{w_{d-3}} \mathrm{unif}_{(n_v, m_v)^\perp \cap \mathbb{S}^{d-1}}(\mathrm{d}l), \tag{4.15}$$

*($w_{d-3}$ denotes the Wallis integral normalization) by* (4.14) *yields the coupled probability* $\mathrm{unif}_{c,\theta}$ *introduced in Definition 4.1.*

### 4.3.2 The coupling creation functional

We summarize below the (elementary) computation of the "coupling creation" of (4.13). In the above, and in the rest of the chapter, the following notation is used for $(u, v) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$:

$$\langle o\,(u, v, u_*, v_*)\rangle_N \overset{\text{def}}{=} \frac{1}{N^2} \sum_{n_1, n_2 = 1}^{N} o(u_{(n_1)}, u_{(n_1)}, u_{(n_2)}, v_{(n_2)}),$$

in order to account for averages over particles of a two-body observable $o : \left(\mathbb{R}^d \times \mathbb{R}^d\right)^2 \to \mathbb{R}$.

**Result 4.1.** *Consider the coupled collisions process* (4.13) *as defined in Section 4.3.1. For any initial condition and $0 \leqslant t \leqslant t + h$, the $L^2$-coupling distance is almost surely decreasing*

$$\left\langle |U_{t+h} - V_{t+h}|^2 \right\rangle_N \leqslant \left\langle |U_t - V_t|^2 \right\rangle_N \qquad \text{a.s..} \tag{4.16}$$

*Moreover the average coupling creation*

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E} \left\langle |U_t - V_t|^2 \right\rangle_N = -\mathbb{E}\mathcal{C}_2\left(U_t, V_t\right) \leqslant 0, \tag{4.17}$$

*is given by the following functional (the average of alignements between the velocity difference of pairs $v - v_* \in \mathbb{R}^d$, and their coupled counterpart $u - u_* \in \mathbb{R}^d$:)*

$$\mathcal{C}_2(u,v) = \frac{d-2}{2d-2} \left\langle |u-u_*|\,|v-v_*| - (u-u_*)\cdot(v-v_*) \right\rangle_N \geqslant 0. \tag{4.18}$$

In order to relate the coupling and the coupling creation, we will introduce in the present paper an original general sharp inequality holding for any couple of centered and normalized random variables in $\mathbb{R}^d$.

**Result 4.2.** *Let $(U,V) \in \mathbb{R}^d \times \mathbb{R}^d$ a couple of centered and normalized ($\mathbb{E}\,|U|^2 = \mathbb{E}\,|V|^2 = 1$) random variables in Euclidean space. Let $(U_*, V_*) \in \mathbb{R}^d \times \mathbb{R}^d$ be an i.i.d. copy. Assume the positive correlation condition $\mathbb{E}U \cdot V \geqslant 0$. Then we have:*

$$\begin{aligned} \frac{1}{2}\mathbb{E}\,|U-V|^2 \leqslant & \min\left(\kappa_{\mathbb{E}(U\otimes U)}, \kappa_{\mathbb{E}(V\otimes V)}\right) \\ & \times \mathbb{E}\left(|U-U_*|^2\,|V-V_*|^2 - ((U-U_*)\cdot(V-V_*))^2\right), \end{aligned} \tag{4.19}$$

*where in the above*

$$\kappa_S = (1-\lambda_{\max}(S))^{-1} \in [d/(d-1), +\infty]$$

*is a condition number for a symmetric positive matrix $S$ of trace 1 and maximal eigenvalue $\lambda_{\max}(S)$ (it is finite if and only if $S$ is of rank at least 2). Moreover, a sufficient condition for the equality case in (4.19) is given by the following isotropy and co-linear coupling conditions*

*(i)* $\frac{U}{|U|} = \frac{V}{|V|}$ a.s..
*(ii) Either* $\mathbb{E}(U\otimes V) = \mathbb{E}(U\otimes U) = \frac{1}{d}1$ *or* $\mathbb{E}(U\otimes V) = \mathbb{E}(V\otimes V) = \frac{1}{d}1$.

In what follows, the inequality (4.19) will be used with respect to particle averaging, that is to say in the form

$$\begin{aligned} \frac{1}{2}\left\langle |u-v|^2 \right\rangle_N \leqslant & \min\left(\kappa_{\langle u\otimes u\rangle_N}, \kappa_{\langle v\otimes v\rangle_N}\right) \\ & \times \left\langle |u-u_*|^2\,|v-v_*|^2 - ((u-u_*)\cdot(v-v_*))^2 \right\rangle_N, \end{aligned} \tag{4.20}$$

for any vectors $u \in (\mathbb{R}^d)^N$ and $v \in (\mathbb{R}^d)^N$ both satisfying the conservation laws (4.7) and such that $\langle u \cdot v\rangle_N \geqslant 0$.

It is then of interest to compare that the alignement functional in the right hand side of (4.19) (which is a sharp upper bound of the square coupling distance), and the coupling creation functional (4.18). They differ by a weight of the form $|u-u_*|\,|v-v_*|$ which forbids any strong "coupling/coupling creation" inequality of the form

$$\frac{\mathcal{C}_2(u,v)}{\left\langle |u-v|^2 \right\rangle_N} \geqslant 2\kappa > 0$$

for some universal constant $\kappa > 0$ independent of $N$ and of the the pair $(u,v) \in (\mathbb{R}^d \times \mathbb{R}^d)^N$ both satisfying the conservation laws (4.7).

Using Hölder inequality, and taking $\mathrm{Law}(U_t) = \pi_\infty = \mathrm{unif}_{\mathbb{S}^{Nd-N-1}}$ (equilibrium), we can however obtain some weaker power law versions for any $\delta \in ]0, +\infty[$. For this purpose, we define

**Definition 4.2.** *The "two-step" or "symmetric" quadratic Wasserstein distance on exchangeable (permutation symmetric) probabilities, denoted $d_{W_2,\mathrm{sym}}$, is defined as the usual quadratic Wasserstein distance (0.3) on the quotient space $(\mathbb{R}^d)^N/\mathrm{Sym}_N$ ($\mathrm{Sym}_N$ is the permutation group) endowed with the quotient distance associated with $d(u,v) = \left\langle |u-v|^2 \right\rangle_N$. The latter quotient distance is*

*also the restriction of the Euclidean quadratic Wasserstein distance on* $\mathbb{R}^d$ *on* empirical distributions *formed by the particle system*[3].

And obtain:

**Result 4.3.** *Let* $t \mapsto V_t \in \left(\mathbb{R}^d\right)^N$ *a Kac's conservative particle system with Maxwell molecules and normalization conditions* (4.7)-(4.3)*. Denote* $\pi_t =^{\text{def}} \text{Law}(V_t)$*. For any* $\delta > 0, q > 1$*, the following trend to equilibrium holds:*

$$\frac{\mathrm{d}^+}{\mathrm{d}t} d_{\mathcal{W}_2,\mathrm{sym}}(\pi_t, \pi_\infty) \leqslant -c_{\delta,q,N}(\pi_t)\, d_{\mathcal{W}_2,\mathrm{sym}}(\pi_t, \pi_\infty)^{1+1/\delta},$$

*where in the above*

$$c_{\delta,q,N}(\pi_t) = k_{\delta,q} \mathbb{E}\left(\left\langle |V_t|^{2q(1+\delta)} \right\rangle_N\right)^{-1/2q\delta} > 0.$$

*with* $k_{\delta,q}$ *a numerical constant (independent of the initial condition and of the angular kernel).*

The moment can be explicitly estimated, uniformly in $N$, in the case of order 4 moments.

**Result 4.4.** *Consider the case* $0 < \delta < 1$*,* $2q(1+\delta) = 4$*, in Result 4.3. We have the upper bound estimate:*

$$d_{W_2,\mathrm{sym}}(\pi_t, \pi_\infty) \leqslant \left(d_{W_2,\mathrm{sym}}(\pi_0, \pi_\infty)^{-1/\delta} + c_\delta (t - t_*)^+\right)^{-\delta}.$$

*where the cut-off time depends logarithmically on the initial radial order-4 moment and is defined by:*

$$t_* = 2\left(\ln\left(\frac{d}{d+2}\mathbb{E}\left\langle |V_0|^4 \right\rangle_N - 1\right)\right)^+.$$

*and* $c_\delta > 0$ *is a numerical constant (independent of* $N$*, of the initial condition and of the angular kernel). For instance, denoting* $c_{\delta,N}$ *the constant for a given particle system size* $N$*, we found*

$$\lim_{\delta\to 1} \lim_{d\to+\infty} \lim_{N\to+\infty} c_{\delta,N} > 10^{-3},$$

*which although sub-optimal is* **physically meaningful**.

Finally let us mention that we have suggested an analysis of the **sharpness** of the obtained estimates, in the form of counter-examples. They provide information on the limitation of the specific choice of the coupling (the simultaneous parallel coupling), but not directly on the trend to equilibrium of the model. Here are the counterexamples:

(i) Velocity distributions with sufficiently heavy tails can make the coupling creation vanish. This first counterexample shows that the obtained "coupling/coupling creation inequality" *must* involve some higher order (say, $> 2$) velocity distribution moments.

(ii) There exists a continuous perturbation of the identity coupling at equilibrium for which however the coupling creation is sub-linearly smaller than the coupling itself. This second type of counterexample shows that even with moment restrictions, a *sub-exponential trend is unavoidable.*

[3] hence the appellation "two-steps"

# Part III

# Perspectives

# 5

# Future directions



I present in this chapter a small sample of some of the research tracks I intend to explore (or propose) in the near future, and that can be thought as sequels of the different chapters of the present dissertation. Here is the list: (i) analysis and numerical simulation of Hamiltonian systems with *path-integral* potential (Section 5.1), (ii) analysis of Fermion Monte-Carlo particle methods (Section 5.2), (iii) asymptotic variance reduction for numerical particle methods (Section 5.3), (iv) coupling methods for convergence to equilibrium of conservative particle systems (Section 5.4).

## 5.1 Classical molecular simulation with *ab initio* potentials

### 5.1.1 Path-integral potentials

Path integral molecular dynamics (PIMD) is a class of methods incorporating quantum mechanics into the classical molecular dynamics simulations of *nuclei* using the Feynman path-integral formulation. In principle, this is motivated for usual temperatures and molecular systems by the fact that the vibrational shortest timescales of covalent bonds compare with the semi-classical parameter.

The resulting model is nonetheless a **classical** Hamiltonian describing the classical dynamics of the nuclei of atoms. The latter Hamiltonian consists in the sum of the usual classical kinetic energy and of an effective potential energy, constructed as follows:

(i) An *ab-initio* electronic structure calculation (e.g. the standard Hartree-Fock method) yields a "raw" potential energy $V_{\rm el}$.
(ii) The nuclei of atoms are considered as non-exchangeable quantum particles.
(iii) The effective, path-integral, potential $V_{{\rm pi},\beta}$ is expressed using an average of $V_{\rm el}$ over small[1] random loops (with probability distribution given by "Brownian bridges"). The latter effective potential is precisely a "free energy" in the sense that it is defined as the probability

[1] with respect to a semi-classical parameter, *i.e.* an adimensional Planck constant

distribution of the *nuclei positions observables* when the system is assumed to be in a *quantum Gibbs state*[2]. The path integral potential can then be discretized by using a *ring of M replicas* of the system where neighbors in the "ring" are coupled by stiff harmonic potentials, the stiffness depending on the semi-classical parameter.

(iv) The system dimension is now $\mathbb{R}^{3NM}$ if $N$ is the number of nuclei, and 3 the space dimension.

Note that at this level, the *only* source of errors due to modeling on the canonical distribution of the system positions is the Born-Oppenheimer approximation (the decoupling assumption between electrons and nuclei).

The molecular dynamics methods based on such principles are thus potentially very accurate methods, particularly useful for studying nuclear quantum effects in light atoms and molecules. They can also very efficiently be used in short time simulations in order to accurately fit effective classical potentials for larger molecular systems.

### 5.1.2 Analysis and numerical issues

The overall objective is then to extend the mass-penalization method described in Section 1.3 to path-integral potentials. One of the main difficulty is the increasing number of force computations when $M$ becomes large. However, we would like to investigate several topics:

(i) The formulation of general method penalizing *at the same time* the fast degrees of freedom of the molecular vibrations and the stiff harmonic interactions introduced by the *ring formulation* of the path-integral potential. This will require the use of a preliminary guess of the fast degrees of freedom (the covalent structure), but in principle, the accuracy of the method is robust on errors on the latter.

(ii) The formal asymptotic analysis when a stiffness parameter is introduced, with a "slow manifold" given by some molecular structure. The semi-classical parameter is taken simultaneously to 0, and has an influence on the description of the effective system dynamics on the slow manifold.

(iii) When $M$ is finite, and not too large, a bias is introduced in the path-integral effective potential. However, it may be possible to use the asymptotic analysis of (ii) above to decrease the bias, or at least to understand the associated error, without increasing $M$, or computing higher derivatives of $V_{\text{el}}$.

## 5.2 Fermion Monte-Carlo methods

The Schrödinger operator for $N$ Fermions can be computed in a *complete basis of Slater determinants*[3]; the operator does no longer possess a sign structure, and the groundstate calculation problem amounts to compute the bottom eigenelements of a very large symmetric matrix without any special structure.

Then, a stochastic representation is obtained but using a population of random processes that evolve in the discrete state space consisting of the labels of each vector of the latter Hilbert basis. Since the matrix lacks a sign structure, walkers may also hold *weights with opposite signs* and may *annihilate* when they are in the same state. This is the idea behind Fermion Monte-Carlo (see [BooTho,CleBoo] where numerical results are exhibited).

However, there is currently no rigorous mathematical understanding of the details of the Fermion Monte-Carlo (FMC) method. In particular, the stabilization of the population size, and the annihilation step need to be understood. The results in [BooTho] indicate that the FMC algorithm exhibits very original behaviors, such as the stabilization of the population size due to

[2] with given temperature and fixed number of particles

[3] the alternate tensor basis obtained from an explicit 1-body Hilbert basis

the annihilation step in a stationary regime, or the phase separation between processes holding different signs. Such behaviors currently lack an appropriate mathematical understanding.

More down-to-earth topics to be tackled in the first place are probably: (i) prove the consistency of the method; (ii) understand the possibility (or impossibility) of effective variance reduction methods for the method, especially when an approximate deterministic solution is available and can be used as an initial guess. The existence of such an asymptotic variance reduction technique in this context (in the sense that the Monte-Carlo error scales with the quality of the ) might be considered as an important breakthrough.

## 5.3 Coupling and variance reduction for particle simulations

### 5.3.1 Moment equations

Consider a particle model, say in phase-space (position and velocity), and subjected to a collisional mechanism. When the collisional mechanisms are important, the velocities are in "local equilibrium", and the first moments of the velocity distribution satisfy a diffusive (if the collisions are non conservative) and/or hyperbolic (if the collisions are conservative) partial differential evolution equation in closed form. Within this point of view, the latter is usually called the *moment equation*. Two classic examples are the following:

(i) The transport of independent particles whose velocities are subjected to damping and random fluctuations (non conservation law). This is the standard Langevin model of Chapter 1. When the damping and fluctuation become large, the probability density of the position of the particles satisfies an ("overdamped") diffusion equation (order 0 moment equation).

(ii) In the case of kinetic equations of Boltzmann type, the particles model is the stochastic particle description of the Boltzmann kinetic equation, where particles are transported according to their velocity, and are subjected to local (in space) *conservative* (momentum and energy) collisions. The moment equation (order $0, 1, 2$: particle density, local mean velocity, local kinetic temperature) is the so-called hydrodynamical limit, the Euler system of compressible gas dynamics.

### 5.3.2 Variance reduction

The overall objective is to study, in a systematic way, *coupling methods* associated with such particle models that uses the moment equation as a control variate. Or, in other (less technical) words, we want to use as much as possible the deterministic information of this moment equations to decrease the statistical error of the simulated particle model evolution.

Chapter 3 is related to this program, the associated moment equation being the described advection-diffusion equation (the chemotaxis equation). However, the control variate is not the moment equation, but another (simplified, kinetic) model.

Anyway, two cases will have to be considered:

(i) The first case is the simplest: it occurs when it is possible to couple directly the collisional process with its equilibrium distribution. This happens for instance in the case (i) of the last section: Langevin processes. The collision mechanism is indeed a Orstein-Uhlenbeck Gaussian process that have an explicit expression as a linear combination of the initial condition and a normalized normal distribution. One can then use the latter expression as the common representation in the coupling.

(ii) The second case happens when such a direct coupling is not explicitly possible, as for Kac's conservative particle system studied in Chapter 4. One then can resort on an additional collisional mechanism on the control variable to *enforce equilibrium*, which requires some additional parameters to tune, and probably a bias versus variance trade-off.

## 5.4 Trend to equilibrium and coupling of conservative collisions

The Markov coupling described in Chapter 4 is based on *simultaneous coupling* of collisions, which yield a *symmetric and Markov* coupling (the two coupled copies have the same marginal distribution, and are jointly Markovian). This deeply rely on the fact that the rate of collisions is constant (Maxwell collisions).

We would like to investigate more general couplings to more general conservative collision processes. In particular, some results in kinetic theory [35, 36, 83] suggests that for certain type of collisions (like diffusive collision with Maxwell constant rate, or bounded jump collisions with super-quadratic rates) a modified log-Sobolev does hold, and exponential entropic convergence holds. This suggests that a more intrinsic *geometric coupling* of the Kac's particle system (without imposing simultaneous coupling) may lead to Wasserstein contractivity for these special cases. A possible first route is to consider the diffusive Kac's system ("Landau case") generator as a *Riemannian metric* on the sphere defined by collision invariants $\mathbb{S}^{dN-N-1}$, and try to compare it to the usual uniform metric. This should lead to inequality analysis very close to the special inequality in Result 4.2.

# Part IV

# References

## References


[1] W Alt, Biased random-walk models for chemotaxis and related diffusion approximations, *J Math Biol* **9**(2) (1980) 147–177.

[2] W Alt, Orientation of cells migrating in a chemotactic gradient, *Adv Appl Probab* **12**(3) (1980) 566–566.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows: in metric spaces and in the space of probability measures* (Springer, 2006).

[4] V. I. Arnol'd, *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics* (Springer-Verlag, 1989).

[5] R. Assaraf and M. Caffarel, *A pedagogical introduction to Quantum Monte Carlo*, volume 74 (Springer, 2000).

[6] Roland Assaraf and Michel Caffarel, Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces, *The Journal of Chemical Physics* **119**(20) (2003) 10536–10552.

[7] R. Assaraf, M. Caffarel, and A. Khelif, Diffusion Monte Carlo with a fixed number of walkers, *Phys. Rev. E* **61**(4) (2000) 4566–4575.

[8] A. Badinski, P. D. Haynes, and R. J. Needs, Nodal Pulay terms for accurate diffusion quantum Monte Carlo forces, *Phys. Rev. B* **77**(8) (2008) 085111.

[9] A. Badinski and R. J. Needs, Total forces in the diffusion Monte Carlo method with nonlocal pseudopotentials, *Phys. Rev. B* **78**(3) (2008) 035134.

[10] Dominique Bakry, Ivan Gentil, Michel Ledoux, et al., *Analysis and geometry of Markov diffusion operators* (Springer, 2014).

[11] C.H. Bennett, Mass tensor molecular dynamics, *J. Comp. Phys.* **19** (1975) 267–279.

[12] AV Bobylev, The theory of the nonlinear boltzmann equation for maxwell molecules, *Mathematical physics reviews* **7** (1988) 111.

[13] AlexanderV. Bobylev and Carlo Cercignani, On the rate of entropy production for the boltzmann equation, *Journal of Statistical Physics* **94**(3-4) (1999) 603–618.

[14] François Bolley, Ivan Gentil, and Arnaud Guillin, Convergence to equilibrium in wasserstein distance for fokker–planck equations, *Journal of Functional Analysis* (2012).

[15] François Bolley, Ivan Gentil, and Arnaud Guillin, Uniform convergence to equilibrium for granular media, *Archive for Rational Mechanics and Analysis* (2012) 1–17.

[16] F. Bornemann and C. Schütte, Homogenization of Hamiltonian system with a strong constraining potential, *Physica D* **102** (1992) 57–77.

[17] Nikolaos Bournaveas and Vincent Calvez, Global existence for the kinetic chemotaxis model without pointwise memory effects, and including internal variables, *Kinet. Relat. Models* **1**(1) (2008) 29–48.

[18] A Bren and M Eisenbach, How signals are heard during bacterial chemotaxis: Protein-protein interactions in sensory signal propagation, *Journal of bacteriology* **182**(24) (2000) 6865–6873.

[19] E. Cancès, B. Jourdain, and T. Lelièvre, Quantum monte-carlo simulations of fermions. a mathematical analysis of the fixed-node approximation, *Math. Mod. and Meth. in App. Sci.* **16** (2006) 1403–1440.

[20] E. Cancès, C. Le Bris, and Y. Maday, *Méthodes mathématiques en chimie quantique: Une introduction* (Springer-Verlag, 2006).

[21] EA Carlen and MC Carvalho, Strict entropy production bounds and stability of the rate of convergence to equilibrium for the boltzmann equation, *Journal of statistical physics* **67**(3-4) (1992) 575–608.

[22] Eric A Carlen, Maria C Carvalho, and Michael Loss, Determination of the spectral gap for kacś master equation and related stochastic evolution, *Acta mathematica* **191**(1) (2003) 1–54.

[23] Eric A Carlen, Jeffrey S Geronimo, and Michael Loss, Determination of the spectral gap in the kac's model for physical momentum and energy-conserving collisions, *SIAM Journal on Mathematical Analysis* **40**(1) (2008) 327–364.

[24] Eric A Carlen and Xuguang Lu, Fast and slow convergence to equilibrium for maxwellian molecules via wild sums, *Journal of statistical physics* **112**(1-2) (2003) 59–134.

[25] Mosé Casalegno, Massimo Mella, and Andrew M. Rappe, Computing accurate forces in quantum Monte Carlo using Pulay's corrections and energy minimization, *The Journal of Chemical Physics* **118**(16) (2003) 7193–7201.

[26] D. Ceperley, G. V. Chester, and M. H. Kalos, Monte-Carlo simulation of a many-fermion study, *Phys. Rev. B* **16** (1977) 3081–3099.

[27] D. M. Ceperley, Fermion nodes, *Journal of Statistical Physics* **63** (1991) 1237–1267.

[28] D. M. Ceperley and B. J. Alder, Ground state of the electron gas by a stochastic method, *Phys. Rev. Lett.* **45**(7) (1980) 566–569.

[29] Carlo Cercignani, *Mathematical methods in kinetic theory* (Plenum Press New York, 1969).

[30] C Cercignani, H-theorem and trend to equilibrium in the kinetic theory of gases, *Archiv of Mechanics, Archiwum Mechaniki Stosowanej* **34** (1982) 231–241.

[31] FACC Chalub, PA Markowich, B Perthame, and Christian Schmeiser, Kinetic models for chemotaxis and their drift-diffusion limits, *Monatsh Math* **142**(1-2) (2004) 123–141.

[32] P. Del Moral, *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications*, Springer Series Probability and its Applications (Springer, 2004).

[33] P. Del Moral and L. Miclo, Branching and Interacting Particle Systems approximations of Feynman-Kac formulae with applications to nonlinear filtering, *Lecture notes in Mathematics* **1729** (2000) 1–145.

[34] P. Del Moral and L. Miclo, Particle approximations of Lyapounov exponents connected to Schrdinger operators and Feynman-Kac semigroups, *ESAIM Proba. Stat.* **7** (2003) 171–208.

[35] Laurent Desvillettes, Clément Mouhot, Cédric Villani, et al., Celebrating cercignani's conjecture for the boltzmann equation, *Kinetic and related models* **4**(1) (2011) 277–294.

[36] Laurent Desvillettes and Cédric Villani, On the spatially homogeneous landau equation for hard potentials part ii: h-theorem and applications: H-theorem and applications, *Communications in Partial Differential Equations* **25**(1-2) (2000) 261–298.

[37] Persi Diaconis and Laurent Saloff-Coste, Bounds for kacś master equation, *Communications in Mathematical Physics* **209**(3) (2000) 729–755.

[38] Emanuele Dolera, Ester Gabetta, and Eugenio Regazzini, Reaching the best possible rate of convergence to equilibrium for solutions of kac equation via central limit theorem, *The Annals of Applied Probability* **19**(1) (2009) 186–209.

[39] Emanuele Dolera and Eugenio Regazzini, The role of the central limit theorem in discovering sharp rates of convergence to equilibrium for the solution of the kac equation, *The Annals of Applied Probability* **20**(2) (2010) 430–461.

[40] Emanuele Dolera and Eugenio Regazzini, Proof of a mckean conjecture on the rate of convergence of boltzmann-equation solutions, *Probability Theory and Related Fields* (2012) 1–75.

[41] A. Doucet, N. de Freitas, and N.J. Gordon, *Sequential Monte-Carlo Methods in Practice*, Series Statistics for Engineering and Information Science (Springer, 2001).
[42] A. Doucet, P. Del Moral, and A. Jasra, Sequential monte carlo samplers, *J. Roy. Stat.. Soc. B* **68**(3) (2006) 411–436.
[43] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Hybrid Monte-Carlo, *Phys. Lett. B* **195**(2) (1987) 216–222.
[44] R Erban and HG Othmer, From individual to collective behavior in bacterial chemotaxis, *SIAM Journal on Applied Mathematics* **65**(2) (2004) 361–391.
[45] R Erban and HG Othmer, From signal transduction to spatial pattern formation in e-coli: A paradigm for multiscale modeling in biology, *Multiscale Model Sim* **3**(2) (2005) 362–394.
[46] J. Fontbona, H. Guérin, and F. Malrieu, Quantitative estimates for the long time behavior of a PDMP describing the movement of bacteria, *ArXiv e-prints* (2010).
[47] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, volume 31 of *Springer Series in Computational Mathematics* (Springer-Verlag, 2006).
[48] B.L. Hammond, W.A. Lester, and P.J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, 1994).
[49] T Hillen and HG Othmer, The diffusion limit of transport equations derived from velocity-jump processes, *SIAM Journal on Applied Mathematics* **61**(3) (2000) 751–775.
[50] A. M. Horowitz, A generalized guided Monte Carlo algorithm, *Phys. Lett. B* **268** (1991) 247–252.
[51] D Horstman, From 1970 until present: the Keller–Segel model in chemotaxis and its consequences I, *Jahresber. Deutsh. Math.-Verein* **105**(3) (2003) 103–165.
[52] D Horstman, From 1970 until present: the Keller–Segel model in chemotaxis and its consequences II, *Jahresber. Deutsh. Math.-Verein* **106**(2) (2004) 51–69.
[53] J. A. Izaguirre and S. S. Hampton, Shadow hybrid Monte Carlo: an efficient propagator in phase space of macromolecules, *J. Comput. Phys.* **200**(2) (2004) 581–604.
[54] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus* (Springer, 1988).
[55] E Keller and L Segel, Initiation of slime mold aggregation viewed as an instability, *Journal on Theoretical Biology* **26** (1970) 399–415.
[56] B. J. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics* (Cambridge University Press, 2005).
[57] T. Lelièvre, M. Rousset, and G. Stoltz, *Free Energy Computations. A Mathematical Perspective* (Imperial College Press, 2010).
[58] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer, *Markov chains and mixing times* (AMS Bookstore, 2009).
[59] P. B. Mackenzie, An improved hybrid Monte Carlo method, *Phys. Lett. B* **226**(3-4) (1989) 369–371.
[60] Florient Malrieu, Logarithmic sobolev inequalities for some nonlinear pde's, *Stochastic processes and their applications* **95**(1) (2001) 109–132.
[61] B. Mao and A.R. Friedman, Molecular dynamics simulation by atomic mass weighting, *Biophysical Journal* **58** (1990) 803–805.
[62] Stéphane Mischler and Clément Mouhot, About kac's program in kinetic theory, *Comptes Rendus Mathematique* **349**(23) (2011) 1245–1250.
[63] Clément Mouhot, Rate of convergence to equilibrium for the spatially homogeneous boltzmann equation with hard potentials, *Communications in mathematical physics* **261**(3) (2006) 629–672.
[64] Roberto Imbuzeiro Oliveira, On the convergence to equilibrium of kacś random walk on matrices, *The Annals of Applied Probability* **19**(3) (2009) 1200–1231.
[65] HG Othmer, SR Dunbar, and W Alt, Models of dispersal in biological-systems, *J Math Biol* **26**(3) (1988) 263–298.
[66] HG Othmer and T Hillen, The diffusion limit of transport equations ii: Chemotaxis equations, *SIAM Journal on Applied Mathematics* **62**(4) (2002) 1222–1250.

[67] C Patlak, Random walk with persistence and external bias, *B Math Biol* **15** (1953) 311–338.
[68] S. Reich, Smoothed dynamics of highly oscillatory Hamiltonian systems, *Physica D* **89** (1995) 28–42.
[69] S. Reich, Smoothed Langevin dynamics of highly oscillatory systems, *Physica D* **138** (2000) 210–224.
[70] M. Rousset, On the control of an interacting particle approximation of Schrödinger ground-states, *SIAM J. Math. Anal.* **38**(3) (2006) 824–844.
[71] H. Rubin and P. Ungar, Motion under a strong constraining force, *Commun. Pure Appl. Math.* **10** (1957) 65–87.
[72] Jonathan Saragosti, Vincent Calvez, Nikolaos Bournaveas, Axel Buguin, Pascal Silberzan, and Benoit Perthame, Mathematical description of bacterial traveling pulses, *PLoS Comput Biol* **6**(8) (2010) e1000890.
[73] AM Stock, A nonlinear stimulus-response relation in bacterial chemotaxis, *P Natl Acad Sci Usa* **96**(20) (1999) 10945–10947.
[74] Alain-Sol Sznitman, *Brownian motion, obstacles and random media* (Springer, 1998).
[75] F. Takens. Motion under the influence of a strong constraining force. In *Global theory of dynamical systems (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1979)* (1980), volume 819 of *Lect. Notes Math.*, Springer, pp. 425–445.
[76] Hiroshi Tanaka, Probabilistic treatment of the boltzmann equation of maxwellian molecules, *Probability Theory and Related Fields* **46**(1) (1978) 67–105.
[77] G Toscani and C Villani, Sharp entropy dissipation bounds and explicit rate of trend to equilibrium for the spatially homogeneous boltzmann equation, *Communications in mathematical physics* **203**(3) (1999) 667–706.
[78] J. Toulouse, R. Assaraf, and C. J. Umrigar, Zero-variance zero-bias quantum Monte Carlo estimators of the spherically and system-averaged pair density, *The Journal of Chemical Physics* **126**(24) (2007) 244112.
[79] J. Toulouse and C. J. Umrigar, Optimization of quantum Monte Carlo wave functions by energy minimization, *The Journal of Chemical Physics* **126**(8) (2007) 084102.
[80] C. J. Umrigar and C. Filippi, Energy and Variance Optimization of Many-Body Wave Functions, *Physical Review Letters* **94**(15) (2005) 150201–+.
[81] N. G. van Kampen, Elimination of fast variables, *Phys. Rep.* **124**(2) (1985) 9–160.
[82] Cédric Villani, A review of mathematical topics in collisional kinetic theory, *Handbook of mathematical fluid dynamics* **1** (2002) 71–74.
[83] Cédric Villani, Cercignani's conjecture is sometimes true and always almost true, *Communications in mathematical physics* **234**(3) (2003) 455–490.
[84] C. Villani, *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics* (American Mathematical Society, 2003).
[85] Max-K von Renesse and Karl-Theodor Sturm, Transport inequalities, gradient estimates, entropy and ricci curvature, *Communications on pure and applied mathematics* **58**(7) (2005) 923–940.
[86] Chuan Xue and Hans G. Othmer, Multiscale models of taxis-driven patterning in bacterial populations, *SIAM J. Appl. Math.* **70**(1) (2009) 133–167.