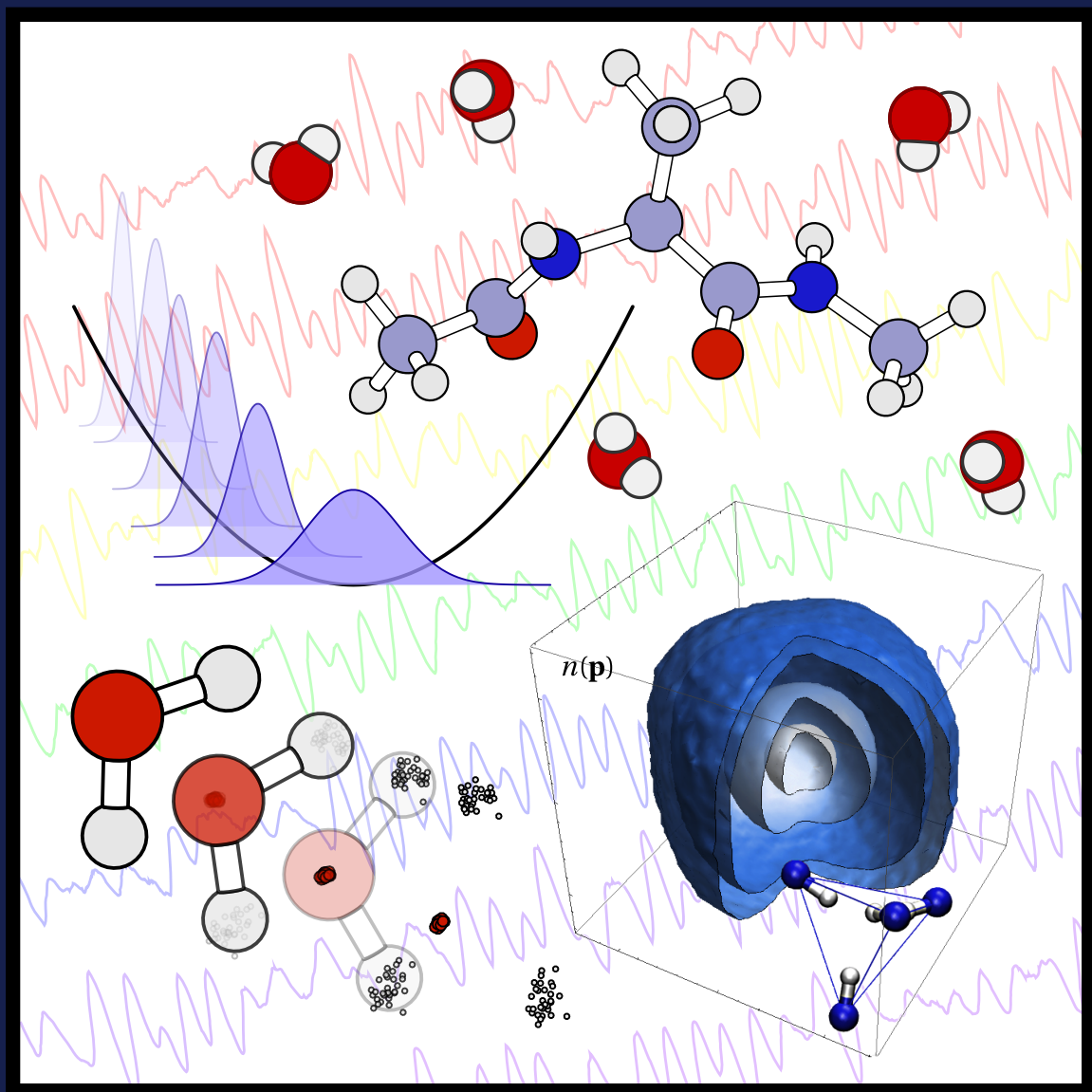


A novel framework for enhanced molecular dynamics based on the generalized Langevin equation

Michele Ceriotti



DISS. ETH. NO. 19038

**A novel framework for enhanced
molecular dynamics based on
the generalized Langevin equation**

A dissertation
submitted to
ETH Zürich
for the degree of
Doctor of Sciences

Presented by
Michele Ceriotti
M. Sc. Università di Milano - Bicocca
Born July 26th 1982

Citizen of
Italy

accepted on the recommendation of
Prof. Michele Parrinello
Prof. David E. Manolopoulos
Prof. Ulrich W. Suter
Prof. Bertram Batlogg

2010

Abstract

The modeling of Brownian motion by Einstein and Langevin, at the beginning of the past century, can be regarded as start point of the systematic use of stochastic methods in physics and chemistry. Besides the theory of Brownian motion, stochastic methods have been applied to countless applications, including the sampling of canonical, constant-temperature ensemble in molecular dynamics simulations.

In the vast majority of cases, the stochastic differential equations which are used as a physical model or as a computational tool are assumed to be Markovian, i.e. they are able to predict the stochastic evolution of the system based on a knowledge of its state at a single instant. On the other hand, many important developments can be derived by removing this assumption. In the Mori-Zwanzig theory, for instance, when one integrates out part of the dynamical variables from a Hamiltonian system, a non-Markovian stochastic equation for the remaining degrees of freedom arises, which contains finite-memory friction and noise.

By changing the properties of the memory in this generalized Langevin equation (GLE), one can greatly influence the static and dynamic properties. In this thesis we developed a robust and flexible framework for exploiting this class of stochastic differential equations in order to enhance and modify almost at will the properties of a molecular dynamics trajectory.

We begin by mapping the non-Markovian dynamics onto a Markovian one in an extended phase-space, as this is more convenient for both simulations and analytical derivations. We then show how a number of static and dynamic properties can be computed exactly in the case of a harmonic oscillator and how the predictions in this limit compare with the behavior of a real system. By performing a fitting procedure, we can then generate sets of parameters that impart to the stochastic dynamics the properties which are best suited to the sampling problem.

When a set of independent GLEs is applied to the Cartesian coordinates of an ensemble of coupled oscillators, the response is as if the stochastic terms had been applied in normal mode representation. This is a consequence of the linear nature of the generalized Langevin equation we employ. This feature allows us to tune the properties as a function of the vibrational frequency of the (quasi-)harmonic modes, without the need to know any frequency and displacement pattern explicitly.

This property has allowed us to develop and demonstrate many different applications. For instance, one can either enhance or degrade the efficiency of sampling of selected normal modes in constant-temperature simulations. One can then increase

the efficiency of sampling for all the frequencies relevant to a given problem. Therefore, an “optimal-sampling” thermostat is obtained which guarantees reliably low correlation times and hence fast convergence of averages. Conversely, by enforcing a long correlation time, one can ensure that the stochastic thermostat will not affect the fast degrees of freedom in simulation methods based on adiabatic separation, such as Car-Parrinello dynamics. This adiabatic decoupling would otherwise be disrupted by the conventional, uncorrelated-noise Langevin equation. One can also predict the extent to which the generalized Langevin equation affects the dynamical properties, and tune it to preserve them on a selected range of frequencies. We present benchmark calculations on a flexible model of liquid water, and discuss the relevance of optimal-sampling GLE in the challenging case of path-integral molecular dynamics.

A whole new set of possibilities opens up if one does not require the GLE to obey the fluctuation-dissipation theorem. In this case a non-equilibrium dynamics is obtained, which nonetheless reaches a stationary distribution with a frequency-dependent effective temperature. By tuning the form of this dependence, dynamics with unusual properties can be obtained. We demonstrate this idea by realizing a δ -thermostat; namely, a GLE which is tuned to set the normal modes within a narrow frequency to a finite temperature, and freeze all the others. We also prove that, by enforcing a frequency dependence of the fluctuations of position and momentum consistent with the finite-temperature density of a quantum harmonic oscillator, one obtains a good approximation for the nuclear quantum effects of light ions. Furthermore, this comes at a fraction of the computational effort that would be required by conventional techniques. After having presented a number of diverse one-dimensional examples and some benchmarks using simple classical forcefields, we discuss an application of this latter method to *ab-initio* molecular dynamics of lithium imide, a compound which is of interest as a medium for hydrogen storage.

The number and variety of methods we have presented within our framework demonstrates the adaptability of the generalized Langevin equation, and its relevance as a tool for improving significantly the efficiency and scope of molecular dynamics simulations. Many applications of the techniques we present will follow, and we also expect that other, completely different simulation schemes can be developed based on this work.

List of publications

Publications directly related to this thesis

1. “Efficient thermostating of path integral molecular dynamics”, M. Ceriotti, M. Parrinello, T. E. Markland and David E. Manolopoulos, *J. Chem. Phys.* **133**, 124104 (2010)
2. “The δ -thermostat: selective normal-modes excitation by colored-noise Langevin dynamics”, M. Ceriotti and M. Parrinello, *Proc. Computer Sci.* **1**, 1601 (2010)

3. "Colored-noise thermostats à la carte", M. Ceriotti, G. Bussi and M. Parrinello, *J. Chem. Theory Comput.* **6**, 1170 (2010)
4. "Nuclear quantum effects in solids using a colored-noise thermostat", M. Ceriotti, G. Bussi and M. Parrinello, *Phys. Rev. Lett.* **103**, 030603 (2009)
5. "Langevin equation with colored noise for constant-temperature molecular dynamics simulations", M. Ceriotti, G. Bussi and M. Parrinello, *Phys. Rev. Lett.* **102**, 020601 (2009)

Other publications submitted during the doctorate

1. "Static disorder and structural correlations in the low temperature phase of lithium imide", G. Miceli, M. Ceriotti, M. Bernasconi, and M. Parrinello, submitted
2. "Solid-liquid interface free energy through metadynamics simulations", S. Angioletti-Uberti, M. Ceriotti, P. D. Lee and M. W. Finnis, *Phys. Rev.* **B 81**, 125416 (2010)
3. "Ab initio study of the diffusion and decomposition pathways of SiH_x species on Si (100)", M. Ceriotti, S. Cereda, F. Montalenti, L. Miglio and M. Bernasconi, *Phys. Rev.* **B 79**, 165437 (2009)
4. "A Hybrid Approach to Fermi Operator Expansion", M. Ceriotti, T. D. Kühne and M. Parrinello, *AIP Proc.* 1148, 658 (2009)
5. "An efficient and accurate decomposition of the Fermi operator", M. Ceriotti, T. D. Kühne and M. Parrinello, *J. Chem. Phys.* **129**, 024707 (2008)
6. "First principles study of Ge/Si exchange mechanisms at the Si (100) surface", F. Zipoli, S. Cereda, M. Ceriotti, M. Bernasconi, L. Miglio and F. Montalenti, *Appl. Phys. Lett.* **92**, 191908 (2008)
7. "Diffusion and desorption of SiH_3 on hydrogenated H : Si (100) (2×1) from first principles", M. Ceriotti and M. Bernasconi, *Phys. Rev.* **B 76**, 245309 (2007)
8. "Conjugate gradient heat bath for ill-conditioned actions", M. Ceriotti, G. Bussi and M. Parrinello, *Phys. Rev.* **E 76**, 026707 (2007)
9. "Quantitative estimate of H abstraction by thermal SiH_3 on hydrogenated Si(001)(2×1)", S. Cereda, M. Ceriotti, F. Montalenti, M. Bernasconi and Leo Miglio, *Phys. Rev.* **B 75**, 235311 (2007)

Publications submitted before the doctorate

1. "Ab initio study of the vibrational properties of crystalline TeO_2 : the α , β , and γ phases", M. Ceriotti, F. Pietrucci and M. Bernasconi, *Phys. Rev.* **B 73**, 104304 (2006)
2. "Impact-driven effects in thin-film growth: steering and transient mobility at the Ag(110) surface", M. Ceriotti, R. Ferrando and F. Montalenti, *Nanotechnology* **17** 3556 (2006)

Riassunto

La modellizzazione del moto browniano messa in atto da Einstein e Langevin agli inizi del ventesimo secolo può essere considerata il momento centrale a partire dal quale i metodi stocastici sono stati impiegati in modo sistematico in fisica e chimica. Oltre che nella teoria del moto browniano, tecniche stocastiche sono state utilizzate in innumerevoli applicazioni, compreso il campionamento dell'*ensemble* canonico, per mezzo di simulazioni di dinamica molecolare a temperatura finita.

Nella maggior parte dei casi si considera che le equazioni differenziali stocastiche, utilizzate come modello fisico o come utile strumento computazionale, siano markoviane: in altre parole, che siano in grado di predire l'evoluzione stocastica del sistema a partire dalla conoscenza del suo stato ad un determinato istante. D'altra parte, eliminando questo assunto si possono ottenere molti risultati di grande interesse. Per esempio, quando nella teoria di Mori-Zwanzig si integrano le equazioni del moto per una parte delle variabili dinamiche di un sistema hamiltoniano, si ottiene un'equazione del moto non-markoviana per i gradi di libertà residui. Tale equazione è caratterizzata da una forza stocastica e da un termine di attrito, entrambi con una memoria finita.

Manipolando le proprietà di tali funzioni di memoria è possibile influenzare il comportamento statico e dinamico di questa equazione di Langevin generalizzata. In questa tesi abbiamo sviluppato un formalismo rigoroso e flessibile, che permette di sfruttare questa classe di equazioni differenziali stocastiche per modificare liberamente le proprietà di una traiettoria di dinamica molecolare, ad esempio rendendo più efficiente l'esplorazione di nuove configurazioni atomiche.

Il metodo che proponiamo prende spunto dall'isomorfismo esistente tra una dinamica non-markoviana ed una più semplice dinamica markoviana in uno spazio esteso. In questo modo diviene possibile calcolare analiticamente un gran numero di proprietà statiche e dinamiche della traiettoria, in particolare quando l'equazione di Langevin generalizzata (ELG) è applicata ad un oscillatore armonico. Per mezzo di una procedura di ottimizzazione è quindi possibile generare molteplici di insiemi di parametri in grado di modificare le proprietà della dinamica stocastica, rendendola adatta al problema in esame. Oltre a discutere questi risultati, osserveremo come il comportamento venga modificato nel caso di sistemi fortemente anarmonici.

La semplicità e l'efficacia del nostro approccio dipendono dal fatto che, in virtù della linearità delle equazioni stocastiche che usiamo, quando un insieme di ELG non correlate tra loro vengono applicate alle coordinate cartesiane di un insieme di oscillatori accoppiati la risposta è la stessa che si avrebbe qualora le equazioni fossero

state applicate direttamente nella base dei modi normali. Tale caratteristica permette di regolare le proprietà del metodo in funzione della frequenza dei modi normali presenti, senza conoscerne la frequenza esatta, né tanto meno gli spostamenti atomici a cui corrispondono.

Questa proprietà ci ha permesso di sviluppare e applicare numerose strategie diverse. Ad esempio, è possibile incrementare o diminuire l'efficienza dell'esplorazione configurazionale in funzione della frequenza. È quindi possibile migliorare l'efficienza del campionamento su un ampio intervallo di frequenze, ottenendo un termostato "ottimale" che assicura bassi tempi di correlazione per numerose osservabili, e di conseguenza una rapida convergenza delle corrispondenti medie. Al contrario, è anche possibile escludere dal campionamento un determinato intervallo di frequenze, imponendo un lungo tempo di correlazione. Questo secondo approccio è particolarmente utile nel caso di metodi basati sulla separazione adiabatica, quale la dinamica Car-Parrinello. Infatti, la dinamica di Langevin convenzionale, basata su un rumore bianco, provocherebbe la rottura del disaccoppiamento adiabatico ed il fallimento della tecnica di simulazione. Infine, è possibile valutare quantitativamente il grado di disturbo introdotto dall'ELG sulle proprietà dinamiche, e regolarlo in modo tale da preservare la dinamica di alcune frequenze selezionate. Discuteremo alcune simulazioni di prova, nel caso di un modello flessibile di acqua, discutendo anche i meriti dell'ELG nel difficile caso della dinamica *path integral*.

È possibile sviluppare una nuova categoria di metodi, oltre a quelli finora descritti, qualora non si richieda che l'ELG soddisfi il teorema fluttuazione-dissipazione. Ne risulta in tal caso una dinamica fuori equilibrio, che d'altra parte comporta una distribuzione stazionaria caratterizzata da una temperatura efficace dipendente dalla frequenza. Modificando tale dipendenza è quindi possibile realizzare traiettorie con proprietà inusuali. Come primo esempio, presentiamo un "termostato δ ", ovvero un'ELG costruita in modo tale da riscaldare esclusivamente i modi normali all'interno di un ristretto range di frequenze, congelando tutti gli altri. Con un approccio simile è anche possibile imporre una dipendenza delle fluttuazioni dalla frequenza compatibile con la distribuzione di equilibrio di un oscillatore armonico quantistico, ottenendo di conseguenza una buona approssimazione degli effetti quantistici legati ai nuclei leggeri. Il vantaggio principale di questa tecnica è il suo costo computazionale contenuto, ai fini pratici pari a quello di una dinamica convenzionale. Dopo aver dimostrato questo metodo in numerosi problemi monodimensionali, esso sarà applicato alla dinamica molecolare da principi primi dell'imide di litio, un composto di grande rilevanza per applicazioni di stoccaggio d'idrogeno.

Il numero e la varietà di metodi che abbiamo presentato nell'ambito del nostro formalismo dimostrano l'adattabilità dell'ELG, e la sua importanza come strumento per migliorare significativamente efficienza e campo di applicazione della dinamica molecolare. Non solo ci aspettiamo che queste tecniche trovino applicazione in numerosi casi concreti, ma riteniamo anche che sia possibile, basandosi sullo stesso approccio, sviluppare schemi di simulazione completamente nuovi.

Contents

Introduction	13
1 Theoretical background	17
1.1 Random processes	18
1.1.1 Markovian processes and the Fokker-Planck equation	19
1.1.2 The Wiener process	21
1.2 Stochastic differential equations	23
1.2.1 The Langevin equation	23
1.2.2 Stochastic integrals	25
1.2.3 SDEs and FP equation	27
1.3 The Ornstein-Uhlenbeck process	28
1.3.1 The integral of the OU process	30
1.3.2 The multivariate OU process	32
2 The GLE Thermostat	35
2.1 Markovian and non-Markovian formulations	35
2.1.1 Conditions for canonical sampling	37
2.1.2 Exponential correlations	39
2.1.3 δ -like memory kernels	40
2.2 The harmonic oscillator	41
2.3 Fitting the GLE parameters	43
2.3.1 Designing a figure of merit for the fit	43
2.3.2 Parametrization of GLE matrices	44
2.3.3 Minimization algorithm	45
2.4 Transformation rules	46
2.4.1 Matrix scaling	46
2.4.2 Combination rules	46
2.5 Implementation details	47

3	Canonical sampling	51
3.1	Optimal sampling	53
3.1.1	White noise on the harmonic oscillator	53
3.1.2	A uniform-efficiency thermostat	55
3.1.3	Diffusion and the free-particle limit	56
3.2	A thermostat for CPMD	58
3.2.1	A low-pass filtered noise	59
3.2.2	Enhancing the CPMD thermostat	61
3.2.3	<i>Ab initio</i> CPMD	61
3.3	Preserving dynamical properties	63
3.4	Liquid water: a case study	65
3.4.1	Local and global schemes	66
3.4.2	Optimal sampling	69
3.4.3	A thermostat for water	70
3.5	Thermostatting quasi-equilibrium dynamics	72
3.5.1	Metadynamics	72
3.5.2	Alanine dipeptide in water	73
4	A thermostat for path integral dynamics	77
4.1	Path-integral molecular dynamics	77
4.1.1	Normal-modes integrator for PIMD	79
4.2	Constant-temperature PIMD	80
4.2.1	White-noise Langevin thermostat	80
4.2.2	A path integral Langevin thermostat	81
4.2.3	A generalized Langevin thermostat for PIMD	82
4.3	Simulations of liquid water	83
4.4	Simulations of hydrogen in palladium	85
5	Frequency-dependent thermostats	89
5.1	The δ -thermostat	91
5.1.1	Selective phonon heating in ice <i>Ih</i>	93
5.1.2	δ -thermostat dynamics as an eigenvalue solver	94
5.2	The quantum thermostat	95
5.2.1	The quantum harmonic oscillator	97
5.2.2	One-dimensional anharmonic potentials	99
5.2.3	Zero-point energy leakage	104
5.2.4	Liquid water	107
5.2.5	A discussion of the accuracy	109
5.2.6	Prospective improvements	110

Contents	11
6 Nuclear quantum effects in Li_2NH	111
6.1 Crystal structure	111
6.2 Quantum-thermostatted simulations	113
6.2.1 Temperature of atomic species	115
6.2.2 Structural properties of NH bonds	115
6.3 Proton momentum distribution	117
6.3.1 Anisotropic momentum distribution	118
6.3.2 Spherically-averaged momentum distribution	119
Conclusions	121
Appendices	123
A Dimensional reduction	123
B Stability of the harmonic GLE	124
C Probability evolution in the harmonic limit	126
D Analytical results	127
E Comparison with Nosé-Hoover thermostat	129
F Momentum distribution in the harmonic limit	130
Bibliography	133
Acknowledgments	143

Nomenclature

CPMD	Car-Parrinello Molecular Dynamics
DOF	Degree(s) of freedom
DOS	Density of states
FP	Fokker-Planck
GLE	Generalized Langevin Equation
OU	Ornstein-Uhlenbeck
PILE-G	Path integral Langevin equation (global version)
PILE-L	Path integral Langevin equation (local version)
PIMD	Path integral molecular dynamics
PMD	Proton momentum distribution
SDE	Stochastic Differential Equation
SVR	Stochastic velocity rescaling
WNLE	White-noise Langevin equation
ZPE	Zero Point Energy

Introduction

Stochastic differential equations have been used to model the time evolution of processes characterized by random behavior, in fields as diverse as physics, chemistry and economics. In particular, the Langevin equation has been regularly applied in the study of Brownian motion in physical chemistry and has been used extensively in molecular dynamics computer simulations, as a convenient and efficient tool to obtain trajectories which sample the constant-temperature, canonical ensemble[1, 2].

In its original form, the Langevin equation is based on the assumption of instantaneous system-bath interactions, which corresponds to the values of the random force being uncorrelated at different times. A non-Markovian, generalized version of the Langevin equation arises in the context of Mori-Zwanzig theory[3–5]. When one integrates out part of the dynamical variables from a Hamiltonian system, a non-Markovian stochastic equation for the remaining degrees of freedom arises, containing finite-memory friction and noise. A simple example of these history-dependent stochastic equations, in the absence of an external potential, is

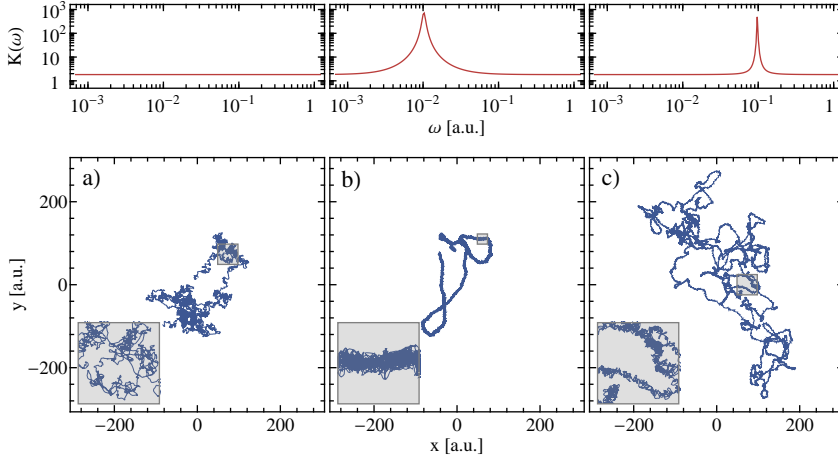
$$\ddot{x} + \int_{-\infty}^t K(t-s) \dot{x}(s) ds - \zeta = 0, \quad (1)$$

where $K(t)$ is a positive real function which describes the friction, and $\zeta(t)$ is a Gaussian process, with time correlation function $H(t) = \langle \zeta(t) \zeta(0) \rangle$. The fluctuation-dissipation theorem, which ensures that configurations visited along the trajectory are distributed according to canonical statistics, requires that $H(t) = k_B T K(t)$. The conventional Langevin equation is recovered as an approximation, in the limit of a clear separation between the characteristic time-scale of the system's dynamics and that of the system-bath interaction.

This class of non-Markovian stochastic differential equations has been extensively used to model the dynamics of open systems interacting with a physically-relevant bath (see e.g. Refs. [6–8]). In such cases, one wants to model how the dynamics of the system is affected by the interaction with its surroundings, without having to treat explicitly the dynamics of the full problem. While the effects of the system-bath coupling can be very significant, they can be often modeled with reasonable accuracy by reproducing their statistical properties by an appropriate colored-noise stochastic differential equation (see Figure 1).

If one wants to compute equilibrium properties, the trajectory of the system is

Figure 1: Trajectories of a particle in two dimensions, as described by the stochastic differential equation (1). Three different choices of memory (we report the Fourier transform of the friction kernel, $K(\omega)$) lead to very different dynamics, despite having the same diffusion coefficient. Panel (a) depicts the case of white-noise, where no time correlation is present. In panels (b) and (c), instead, two different history-dependent contribution are added on top of the same instantaneous friction used in (a). Each trajectory is $2.5 \cdot 10^5$ time units long.



irrelevant, as long as the configurations are sampled from the appropriate statistical ensemble. Monte Carlo methods can be regarded as a limiting case of this line of reasoning, where the procedure to generate new atomic configurations does not have to bear any resemblance to Hamilton's equations[9, 10]. However, it turns out that in many cases Hamilton's equations are a very natural and efficient way to generate ensemble-preserving collective moves and extensive research has been performed - and is still active - in the field of hybrid Monte Carlo methods, which try to combine the best features of molecular dynamics and stochastic approaches[11].

Hamiltonian trajectories alone sample the constant-energy, microcanonical ensemble, and they must be modified to allow energy fluctuations, if one wants to sample the canonical, constant-temperature ensemble when the number of simulated particles is far away from the thermodynamic limit. White-noise Langevin equation[1], or similar equations which employ uncorrelated noise[12, 13], have been used to this aim for a long time. One is then led to consider how non-Markovian Langevin dynamics can be used as a *tool* rather than as a *physical model*, i.e. whether an appropriate colored-noise stochastic dynamics can extend the range of physical observables accessible by molecular dynamics, and enhance the efficiency of conventional methods. As demonstrated in Figure 1, a linear, Gaussian stochastic process such as (1) can give rise to very different dynamical trajectories depending on the characteristics of its memory kernels, without affecting the sampled ensemble. In this thesis we will demonstrate that generalized Langevin equations can be used as a powerful and flexible instrument to aid molecular dynamics simulations, which can alter the sampling efficiency in a precise, targeted way, and even modify the

properties of the statistical ensemble which is sampled by the trajectory.

After having reviewed briefly the mathematical foundations of the theory of stochastic differential equations in chapter 1, we will introduce a comprehensive framework to predict the effect of a given generalized Langevin equation on the dynamical trajectory and on the stationary averages (chapter 2). This is based on the isomorphism between a non-Markovian equation on the physical variables and on a Markovian dynamics in an extended phase-space, and allows one to optimize the parameters so as to obtain the desired effect.

We will then demonstrate the flexibility of the approach, presenting some of the simulation techniques which can be developed within our GLE framework. We will first discuss in chapter 3 a series of strategies to enhance the sampling of the constant-temperature canonical ensemble, by making exploration of phase space more effective, or to prevent the SDE from affecting the equilibration or the dynamical properties of selected modes. In chapter 4 these ideas will be applied to the problem of sampling in path-integral molecular dynamics.

In chapter 5 we will instead survey the consequences of breaking the fluctuation-dissipation theorem and realize a non-equilibrium dynamics by which the effective temperature of distinct vibrational modes can be adjusted independently. Exploiting this concept we successfully obtain a stochastic method which can model to a good approximation the quantum effects connected with the presence of light ions, such as zero-point energy and non-constant specific heat. In chapter 6 this quantum thermostat will be applied to *ab-initio* molecular dynamics of lithium imide, a material which is being studied for hydrogen-storage applications. Finally, we will present our conclusions.

Theoretical background

Probabilistic concepts play a major role in modern physics, both in quantum mechanics and in statistical thermodynamics. Even Hamilton's equations, which are probably the prototype physical law for generating deterministic trajectories, can equally well be treated by Liouville's formalism, where the time evolution of a swarm of trajectories is described in terms of their probability density.

In a number of situations one cannot - or does not want to - tackle a physical problem based on a purely deterministic trajectory. Instead then from the beginning a certain degree of randomness is inserted into the model, and an attempt is made to understand the statistical properties of this stochastic contribution.

An archetypal example is Brownian motion, which is named after Robert Brown, a Scottish botanist who observed in 1827 that pollen grains suspended in a liquid droplet moved along irregular, unpredictable trajectories when observed under a microscope. It was soon ruled out that this chaotic motion was a manifestation of life, since it was demonstrated that inanimate particles also exhibit the very same behavior. No suitable model for this phenomenon was found until the beginning of 20th century, when first Einstein[14] and later Langevin[15] proposed satisfactory, quantitative explanations. The two approaches started from the same qualitative assumption; namely, that particles in suspension undergo apparently random motion due to fast, chaotic collisions with the molecules of the fluid. Therefore, a stochastic description of the problem was needed, if one wanted to avoid treating the liquid explicitly. The mathematical methods followed by the two authors were however very different. Langevin's argument introduced a stochastic equation for the trajectory, which only later was developed into a fully rigorous mathematical framework. Einstein's derivation meanwhile was based on estimating the diffusion of the particles in terms of their probability density. Both models gave results that agreed with the experiments, and it was quite clear that a stochastic description in terms of trajectories and one in terms of probability densities had a relationship much like that between Hamilton's and Liouville's formalisms in classical mechanics.

A satisfactory mathematical treatment of Langevin equation had to wait 40 years and the work of Kiyoshi Itô[16]. Even though we will not make heavy use of Itô calculus in this thesis, we will summarize its development briefly for the simple case

of the one-dimensional Langevin equation, without any pretense of being exhaustive. In the process we will encounter the fundamental ideas which are necessary for the description of random processes. This brief examination will start with a few selected concepts in probability theory, including the definition of the Fokker-Planck equation, and then move on to then sketch how stochastic integrals can be constructed. We will conclude by describing the Ornstein-Uhlenbeck stochastic process in some detail, as this is the basis of our generalized Langevin thermostat.

1.1 Random processes

Consider a system whose state is described by the value of a vector \mathbf{x} , which can evolve in time according to an unknown law, possibly characterized by a degree of random behavior. We now assume that we have collected several realizations of this process. We will refer to each trajectory as a *sample path* $\mathbf{x}(t)$. We let Ω be the set of all such paths, and label each path according to some index ω . One can then describe the random process in terms of the distribution of the points in phase space at a given time, and hence construct a probability density¹ (figure 1.1)

$$P(\mathbf{x}, t) \propto \int \delta(\mathbf{x}_\omega(t) - \mathbf{x}) d\omega.$$

This probability, however, does not characterize the random process completely, since one only has knowledge on the “snapshots” of the collection of sample paths at different times. No information regarding the identity of the paths in the different snapshots has been collected. One could compute the *joint* probability for a sample path to be at \mathbf{x}_1 at time t_1 , and at \mathbf{x}_2 at time t_2 ,²

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2) \propto \int \delta(\mathbf{x}_\omega(t_1) - \mathbf{x}_1) \delta(\mathbf{x}_\omega(t_2) - \mathbf{x}_2) d\omega. \quad (1.1)$$

It would then be necessary to define a three-times probability, and proceed to construct a hierarchy of joint probabilities and ultimately a *probability functional* $\mathcal{P}[\mathbf{x}(t)]$ [17]. Summation rules of the form

$$P(\mathbf{x}_1, t_1) = \int P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2) d\mathbf{x}_2 \quad (1.2)$$

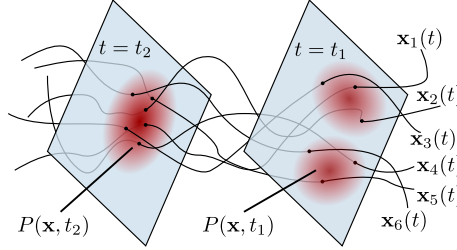
must hold, so that the information contained in the one-time probability density is completely embedded into the two-times P ; in general, knowledge of the n -times probability makes all the lower-order densities redundant.

This characterization of the stochastic process would be quite impractical, even though it is able to describe exhaustively its statistical properties. Fortunately, it is often justified to make a number of assumptions on the form of the joint probability (1.1) so as to bring it in a more treatable form. A first simplification requires

¹The integral here is just used to mean some “averaging” procedure to be performed over all the realizations of the random process.

²We will assume times to be ordered according to $t_1 \geq t_2 \geq \dots$

Figure 1.1: A collection of sample paths for a random process. Also shown is how the one-time probability density $P(\mathbf{x}, t)$ can be constructed as the distribution of the points of all the sample paths at a given time.



the process to be *stationary* in time, i.e. that the two-times joint probabilities only depend on the time difference,

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2) = P(\mathbf{x}_1, t_1 - t_2; \mathbf{x}_2, 0). \quad (1.3)$$

1.1.1 Markovian processes and the Fokker-Planck equation

Let us now introduce the *conditional probability* $P(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2)$, which is defined as the probability of the system being at \mathbf{x}_1 at a given time t_1 , given that it was as \mathbf{x}_2 at time t_2 . Its relation to the joint probability is

$$P(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2) = P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2) / P(\mathbf{x}_2, t_2). \quad (1.4)$$

We will say that a process is Markovian if its joint conditional probabilities depend only on the most recent time frame:

$$P(\mathbf{x}_1, t_1; \dots; \mathbf{x}_k, t_k | \mathbf{x}_{k+1}, t_{k+1}; \dots; \mathbf{x}_n, t_n) = P(\mathbf{x}_1, t_1; \dots; \mathbf{x}_k, t_k | \mathbf{x}_{k+1}, t_{k+1}). \quad (1.5)$$

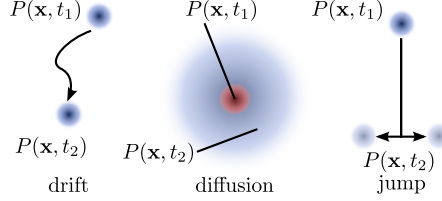
This ansatz means that at each time the model has no memory of the past history, and that further evolution is (probabilistically) determined uniquely by knowledge of the status at a given instant³. The description of the stochastic process is thus enormously simplified. Using the definition of conditional probability (1.4), one can write

$$\begin{aligned} P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2; \mathbf{x}_3, t_3) &= P(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2; \mathbf{x}_3, t_3) P(\mathbf{x}_2, t_2; \mathbf{x}_3, t_3) = \\ &= P(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2) P(\mathbf{x}_2, t_2 | \mathbf{x}_3, t_3) P(\mathbf{x}_3, t_3), \end{aligned}$$

i.e. any joint probability can be broken down to a product of the initial, single-time distribution and a series of conditional probabilities. If the process is also stationary, the conditional probability will depend only on the time difference, and hence its evolution is completely determined by the initial probability distribution and by the conditional probability $P(\mathbf{x}, t | \mathbf{x}_0, 0)$. Combining Eqs. (1.2), (1.4) and the Markov

³This might seem to be a very crude assumption, but it holds true at least approximately for a large number of physically-relevant problems.

Figure 1.2: The terms which govern the evolution of a probability density according to the differential Chapman-Kolmogorov equation (1.7) can be described as a deterministic drift part, which depends on a vector field $\mathbf{a}(\mathbf{x}, t)$, a diffusion term controlled by the matrix $\mathbf{B}(\mathbf{x}, t)$, and a discontinuous jump term which is described by the function $W(\mathbf{x}|\mathbf{x}', t)$.



property (1.5) one arrives at the Chapman-Kolmogorov equation,

$$P(\mathbf{x}_1, t_1 | \mathbf{x}_3, t_3) = \int P(\mathbf{x}_1, t_1 | \mathbf{x}_2, t_2) P(\mathbf{x}_2, t_2 | \mathbf{x}_3, t_3) d\mathbf{x}_2 \quad (1.6)$$

Eq. (1.6) is a complex functional equation for the conditional probability density. Under some mild conditions on the form of P , it is possible (see e.g. [16, p. 48ss]) to derive a corresponding differential equation, which describes in the most general form the evolution of a Markov process, that reads

$$\begin{aligned} \frac{\partial}{\partial t} P(\mathbf{x}, t | \mathbf{x}_0, t_0) = & - \sum_i \frac{\partial}{\partial x_i} [a_i(\mathbf{x}, t) P(\mathbf{x}, t | \mathbf{x}_0, t_0)] + & \leftarrow \text{drift} \\ & + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} [B_{ij}(\mathbf{x}, t) P(\mathbf{x}, t | \mathbf{x}_0, t_0)] + & \leftarrow \text{diffusion} \\ & + \int [W(\mathbf{x} | \mathbf{x}', t) P(\mathbf{x}', t | \mathbf{x}_0, t_0) - & \leftarrow \text{jump} \\ & W(\mathbf{x}' | \mathbf{x}, t) P(\mathbf{x}, t | \mathbf{x}_0, t_0) d\mathbf{x}'] . \end{aligned} \quad (1.7)$$

The differential Chapman-Kolmogorov equation can be split up in three terms, which describe different elementary processes contributing to the evolution of P during an infinitesimal time (figure 1.2).

Drift term The first term is dubbed the drift contribution as it corresponds, to deterministic trajectories treated in a probabilistic framework. One can see that it arises from taking the limit

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|\mathbf{y} - \mathbf{x}| < \epsilon} (\mathbf{y}_i - \mathbf{x}_i) P(\mathbf{y}, t + \Delta t | \mathbf{x}, t) d\mathbf{y} = a_i(\mathbf{x}, t) + \mathcal{O}(\epsilon). \quad (1.8)$$

A typical example of a FP equation containing the drift term only is the Liouville form of Hamilton's equations. In one dimension, with a potential $V(q, t)$ and given the boundary condition $P((q, p), t_0 | (q_0, p_0), t_0) = \delta[(q, p) - (q_0, p_0)]$, the solution in terms of a probability density is

$$P((q, p), t | (q_0, p_0), t_0) = \delta[(q, p) - (q(t), p(t))] \quad (1.9)$$

where $(q(t), p(t))$ is the phase-space trajectory stemming from the ordinary Hamilton's equations

$$\frac{\partial}{\partial t}(q, p) = \mathbf{a}((q, p), t) = \left(p, -\frac{\partial}{\partial q} V(q, t) \right)$$

with boundary condition $(q(t_0), p(t_0)) = (q_0, p_0)$, as one can see substituting (1.9) into

$$\frac{\partial}{\partial t} P = -\nabla \cdot [\mathbf{a}((q, p), t) P].$$

Diffusion term The second term corresponds to diffusion, i.e. to a spread of the probability density. This is closely connected with the random nature of the process, just like the drift term is related to its deterministic behavior. It can be shown to correspond to the second order analogue of Eq. (1.8):

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|\mathbf{y}-\mathbf{x}| < \epsilon} (y_i - x_i)(y_j - x_j) P(\mathbf{y}, t + \Delta t | \mathbf{x}, t) d\mathbf{y} = B_{ij}(\mathbf{x}, t) + \mathcal{O}(\epsilon). \quad (1.10)$$

Note that all higher-order terms are $\mathcal{O}(\epsilon)$ and can therefore be ignored. If only the diffusion term is present, the differential Chapman-Kolmogorov equation reduces to the same form obtained by Einstein for the diffusion of a Brownian particle. In one dimension, and setting the diffusion matrix to the constant D , one has

$$\frac{\partial}{\partial t} P(x, t | x_0, t_0) = \frac{1}{2} D \frac{\partial^2}{\partial x^2} P(x, t | x_0, t_0). \quad (1.11)$$

Its solution, subject to the boundary condition $P(x, 0 | 0, 0) = \delta(x)$, is $P(x, t | 0, 0) \propto e^{-x^2/2Dt}$. One finds that the mean square displacement as a function of time is $\langle x^2 \rangle(t) = Dt$, which correspond precisely to the behavior observed in Brownian motion.

Jump term The very last term is a jump term, as the solution for $a_i = B_{ij} = 0$ and the boundary condition $P(\mathbf{x}, t_0 | \mathbf{x}_0, t_0) = \delta(\mathbf{x} - \mathbf{x}_0)$ corresponds to discontinuous sample paths. It derives from the limit

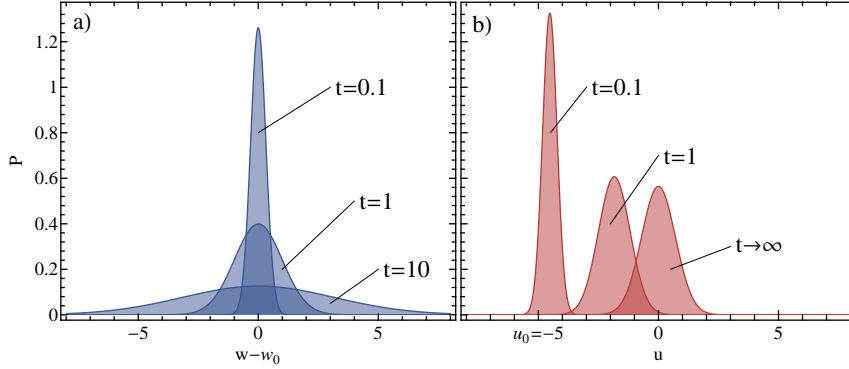
$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(\mathbf{y}, t + \Delta t | \mathbf{x}, t) = W(\mathbf{y} | \mathbf{x}, t) \quad [|\mathbf{y} - \mathbf{x}| \geq \epsilon].$$

In this thesis we will consider only processes with continuous sample paths, and will thus always set $W = 0$. In this limit the differential Chapman-Kolmogorov equation is often called a Fokker-Planck equation.

1.1.2 The Wiener process

The solution to the one-dimensional diffusion equation (1.11), with unitary diffusion coefficient and the initial condition $P(w, t_0 | w_0, t_0) = \delta(w - w_0)$ is known as a Wiener process, and it can be considered the prototype Markovian stochastic process. The corresponding conditional probability density can be readily found, and checked by

Figure 1.3: Representation of the probability density at different times for (a) the Wiener process (1.12), and (b) an Ornstein-Uhlenbeck process (1.24) with $a = b = 1$ and $u_0 = -5$. In the second case, a well-defined, stationary probability density exists for $t \rightarrow \infty$.



substitution. It reads

$$P(w, t | w_0, t_0) = \frac{1}{\sqrt{2\pi(t-t_0)}} e^{-(w-w_0)^2/2(t-t_0)}, \quad (1.12)$$

which is simply a Gaussian spreading in time, out of the initial Dirac delta (see figure 1.3(a)). This distribution can be as well defined in terms of its average w_0 and its variance which grows as $(t - t_0)$. Since P depends only on time differences, we can consider for all purposes $P(w, t - t_0 | w_0, 0)$.

In spite of its simplicity, the Wiener process has some rather disturbing properties. First of all, the variance of its sample path grows without bounds in time, which means that even if the average is equal to the position at time t_0 , the sample paths will end up being very far from it at large times, and that no stationary probability distribution exists.

The trajectory $w(t)$ is continuous, since

$$\begin{aligned} \frac{1}{\Delta t} \int_{|w-w_0| > \Delta w} P(w, \Delta t | w_0, 0) dw &= \frac{1}{\Delta t} \operatorname{erfc}\left(\frac{\Delta w}{\sqrt{2\Delta t}}\right) = \\ &= e^{-\frac{\Delta w^2}{2\Delta t}} \left(\sqrt{\frac{2\pi\Delta t}{\Delta w^2}} + \mathcal{O}(\Delta t^{3/2}) \right) \end{aligned}$$

tends to zero uniformly for small Δt . However, the derivative of the sample path is infinite almost everywhere. In fact, one can compute the probability of having a displacement per infinitesimal unit time Δt greater than a given value k . The resulting expression reads

$$\int_{|\Delta w| > k\Delta t} P(w_0 + \Delta w, \Delta t | w_0, 0) d\Delta w = \operatorname{erfc}\sqrt{\Delta t k^2/2}, \quad (1.13)$$

whose limit for $\Delta t \rightarrow 0$ is one, independent of k . Just as for Brownian motion, for which it is a model, the Wiener process has continuous but otherwise irregular and

chaotic sample paths.

Finally, let us derive a couple of results which will be useful in the development of Itô's stochastic integral. Consider the sample path at different time intervals, and compute the joint probability of the increments. Since $P(w_2, t_2 | w_1, t_1) = P(w_2 - w_1, t_2 - t_1 | 0, 0) = P(\Delta w, \Delta t)$, and since the process is Markovian one has⁴

$$P(\Delta w_2, \Delta t_2; \Delta w_1, \Delta t_1 | w_0, 0) = P(\Delta w_2, \Delta t_2) P(\Delta w_1, \Delta t_1),$$

which means that the increments Δw_i are all independent of each other.

The correlation function for the Wiener process starting off at (w_0, t_0) is

$$\begin{aligned} \langle w(t) w(s) | w_0, t_0 \rangle &= \int w_1 w_2 P(w_1, t; w_2, s | w_0, t_0) dw_1 dw_2 = \\ [t > s] &= \int \Delta w_1 w_2 P(\Delta w_1, t - s | 0, 0) P(w_2, s | w_0, t_0) d\Delta w_1 dw_2 + \\ &+ \int w_2^2 P(w_2, s | w_0, t_0) dw_2 = \\ &= 0 + s - t_0 + w_0^2 \end{aligned}$$

so that overall

$$\langle w(t) w(s) | w_0, t_0 \rangle = \min(t - t_0, s - t_0) + w_0^2. \quad (1.14)$$

1.2 Stochastic differential equations

1.2.1 The Langevin equation

So far, our discussion has dealt with the probability density associated with a given random process $x(t)$. This is reminiscent of the treatment of the Brownian motion developed by Einstein. As we mentioned at the beginning of this chapter, the alternative approach followed by Langevin was based on an equation of motion for the random process itself, which contained a “stochastic force” term. The Langevin equation, in the form which was used to model Brownian motion, reads

$$m\ddot{x} + m\gamma\dot{x} - \xi = 0, \quad (1.15)$$

where $\xi(t)$ is a Gaussian random force term which satisfies $\langle \xi \rangle = 0$, $\langle \xi(t) \xi(t') \rangle = \delta(t - t')$ and $\langle x\dot{x} \rangle = \langle \dot{x}\xi \rangle = 0$, x and \dot{x} being the position and the velocity of the Brownian particle respectively.

Langevin's derivation considered the product of Eq. (1.15) with x , which leads to

$$mx \frac{d^2}{dt^2} x + m\gamma x \frac{d}{dt} x - x\xi = \frac{m}{2} \frac{d^2}{dt^2} x^2 - m \left(\frac{d}{dt} x \right)^2 + \frac{m\gamma}{2} \frac{d}{dt} x^2 - x\xi.$$

⁴We are here considering $P(\Delta w_2, \Delta t_2; \Delta w_1, \Delta t_1 | w_0, 0) = P(w_2, t_2; w_1, t_1 | w_0, 0)$ with $w_1 = w_0 + \Delta w_1$, $w_2 = w_1 + \Delta w_2$, $t_1 = \Delta t_1$, $t_2 = t_1 + \Delta t_2$.

On averaging, the random force term disappears, since it is uncorrelated with the position, and one can set $m \langle \dot{x}^2 \rangle = k_B T$ because the system is in thermal equilibrium. Eventually, a differential equation for the mean square displacement $\langle x^2 \rangle$ is obtained,

$$\frac{m}{2} \frac{d^2}{dt^2} \langle x^2 \rangle + \frac{m\gamma}{2} \frac{d}{dt} \langle x^2 \rangle - k_B T = 0,$$

which can be solved to give

$$\frac{d}{dt} \langle x^2 \rangle = \frac{2k_B T}{\gamma m} + C e^{-\gamma t}.$$

Considering the long time limit, one recovers the expression for the diffusion coefficient $\frac{d}{dt} \langle x^2 \rangle = 2 \frac{k_B T}{\gamma m}$, which is the same result obtained by solving the diffusion equation (1.11) with the appropriate value for the diffusion coefficient D .

However, one might wonder about the mathematical consistency of Eq. (1.15), and with good reasons: consider the simple case of $\gamma = 0$ and $m = 1$, and solve for $w = \dot{x}$. We can tentatively write the corresponding integral

$$w(t) = \int_0^t \xi(s) ds, \quad (1.16)$$

and leave a more rigorous discussion of its definition till later. For now, we require only that $w(t)$ be continuous, and get as a consequence that $w(t)$ will be a Markovian process. In fact,

$$w(t_2) = \int_0^{t_2} \xi(s) ds = \int_0^{t_1} \xi(s) ds + \lim_{\epsilon \rightarrow 0} \int_{t_1+\epsilon}^{t_2} \xi(s) ds = w(t_1) + \lim_{\epsilon \rightarrow 0} \int_{t_1+\epsilon}^{t_2} \xi(s) ds,$$

i.e. the evolution of $w(t)$ from t_1 onwards depends only on $w(t_1)$, since $\xi(s)$ in the last integral term is uncorrelated with ξ at previous times. Being a continuous Markov process, one can therefore write a Fokker-Planck equation for $w(t)$, by computing the drift and diffusion terms according to Eqs. (1.8) and (1.10):

$$\begin{aligned} a(w_0, t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{w_0-\epsilon}^{w_0+\epsilon} (w - w_0) P(w, t + \Delta t | w_0, t) dw \sim \\ &\sim \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle w(t + \Delta t) - w_0 \rangle = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left\langle \int_t^{t+\Delta t} \xi(s) ds \right\rangle = 0 \end{aligned}$$

$$\begin{aligned} B(w_0, t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{w_0-\epsilon}^{w_0+\epsilon} (w - w_0)^2 P(w, t + \Delta t | w_0, t) dw \sim \\ &\sim \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (w(t + \Delta t) - w_0)^2 \rangle = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \int_t^{t+\Delta t} \langle \xi(s) \xi(s') \rangle ds ds' = \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \int_t^{t+\Delta t} \delta(s - s') ds ds' = 1 \end{aligned}$$

Note that these results imply that $w(t)$ obeys the same FP equation as the Wiener process, implying that it is a Wiener process! As a consequence, $\xi(t)$ is the derivative

of a Wiener process, which we have shown (see Eq. (1.13)) to be infinite almost everywhere.

To summarize, the Langevin equation (1.15) makes little sense as a differential equation. This is a consequence of the extremely irregular behavior of the stochastic force $\xi(t)$, which can have physical meaning only when thought of as the limiting case of a physical process with a short but finite correlation time. On the other hand, considering a noise with a finite correlation would break the Markovian assumption, and make the treatment even more complex, so one should develop the mathematical tools to give Eq. (1.15) a precise meaning.

1.2.2 Stochastic integrals

Generalizing Eq. (1.15), one can write a SDE of the form

$$\dot{x} = a(x, t) + b(x, t) \xi, \quad (1.17)$$

which is a shorthand notation for the integral equation

$$x(t) = x(0) + \int_0^t a(x(s), s) dt + \int_0^t b(x(s), s) \xi(s) ds.$$

To define $\int_0^t b(x(s), s) \xi(s) ds$, consider an arbitrary function of time $g(t)$. Since the integral of $\xi(s)$ alone is a Wiener process, we may write $\xi(s) ds = dw(s)$ and, in analogy with Riemann-Stieltjes integral⁵,

$$h(t) = \int_{t_0}^t g(s) dw(s) = \text{ms-}\lim_{n \rightarrow \infty} \sum_{i=1}^n g(\tau_i) [w(t_i) - w(t_{i-1})], \quad (1.18)$$

where the t_i 's correspond to a subdivision of the $[t_0, t]$ interval at $n-1$ points, ordered as $t_0 \leq t_1 \leq \dots \leq t_i \leq \dots \leq t_{n-1} \leq t$, and τ_i 's are points taken within each interval, $t_{i-1} \leq \tau_i \leq t_i$ (see figure 1.4). Given the random behavior of $w(t)$ and possibly $g(t)$, such a definition poses larger problems than in the Riemann-Stieltjes case; for example, the value of the limit can depend on the position of τ_i within each interval.

In fact, the precise meaning of the integral and ultimately of equation (1.17) depends on the relative positions of the points at which $g(t)$ and $w(t)$ are evaluated. Two schemes are commonly used to select these points. The first produces the Itô integral, which is defined as

$$\mathcal{I} \int_{t_0}^t g(s) dw(s) = \text{ms-}\lim_{n \rightarrow \infty} \sum_{i=1}^n g(t_{i-1}) [w(t_i) - w(t_{i-1})], \quad (1.19)$$

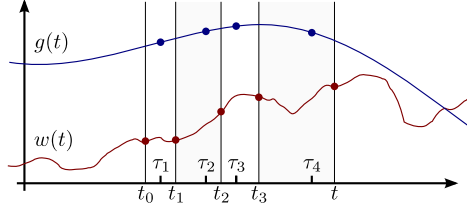
⁵Since we are dealing with random processes, what we really imply is a mean-square limit, which is defined as follows. Let f_n be a sequence of random processes; f_n is said to converge in a mean-square sense to f if

$$\lim_{n \rightarrow \infty} \langle (f_n - f)^2 \rangle = 0,$$

and one can in such case write

$$\text{ms-}\lim_{n \rightarrow \infty} f_n = f$$

Figure 1.4: A possible subdivision of the $[t_0, t]$ interval, to be used in the definition of the stochastic integral $\int g(t) dw(t)$. Since both $w(t)$ and $g(t)$ could be defined as a random process, and at least $w(t)$ is an extremely irregular one, the fact that w and g are evaluated at different points can lead to convergence problems for the partial sums in (1.18).



i.e. $g(t)$ is evaluated precisely at the initial point of each interval. This choice makes Itô's integral very convenient from a mathematical point of view, as the value of $g(t_{i-1})$ is independent⁶ of $[w(t_i) - w(t_{i-1})]$. For this reason, it is widely used in analytical derivations.

The Stratonovich approach can be introduced with reference to the stochastic differential equation (1.17) as

$$\mathcal{S} \int_{t_0}^t b(x(s), s) dw(s) = \text{ms-}\lim_{n \rightarrow \infty} \sum_{i=1}^n b\left(\frac{x(t_{i-1}) + x(t_i)}{2}, t_{i-1}\right) [w(t_i) - w(t_{i-1})],$$

where only the dependence on $x(t)$ is averaged in the mean-square limit procedure⁷. The Stratonovich integral arises quite naturally when considering ξ as the limit of a finite-memory noise, and has some nice properties. For instance, the same rules of ordinary calculus hold formally. Unfortunately, however, it also involves significant complications from a mathematical point of view, since the value of $x(t)$ is evaluated at time t_i , which might be correlated with $[w(t_i) - w(t_{i-1})]$. It is however possible to find the Itô equivalent of a Stratonovich SDE, so that one can convert one representations into the other as required[16].

A detailed discussion of the properties of the two flavors of stochastic integrals is beyond the scope of this introductory survey. In fact, we will use only an extremely simple form of Eq. (1.17), in which b is a constant. In this case, the two representations are equivalent, and most of the complications of stochastic integration are not a concern.

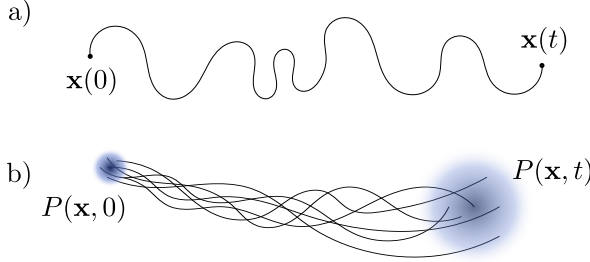
⁶Strictly speaking, this is true only if $g(t)$ is *non-anticipating*, a condition however which holds true in most of the physically relevant cases.

⁷The Stratonovich integral is usually defined just as

$$\mathcal{S} \int_{t_0}^t g(s) dw(s) = \text{ms-}\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{g(t_{i-1}) + g(t_i)}{2} [w(t_i) - w(t_{i-1})].$$

Only with the definition given in the text, however, it is possible to draw a simple connection between Stratonovich and Itô interpretations of Eq. (1.17). If $b(x, t)$ is a continuous function of t , the two definitions are equivalent.

Figure 1.5: Given a stochastic differential equation which generates trajectories in phase-space $\mathbf{x}(t)$ (a), the associated Fokker-Planck equation describes the evolution of a swarm of trajectories whose initial configuration is distributed according to the probability density $P(\mathbf{x}, 0)$ (b).



1.2.3 Stochastic differential equations and Fokker-Planck equation

Having examined the meaning of the stochastic differential equation (1.17), we can move on to discuss the connections between the resulting random process and a description in terms of a Fokker-Planck equation. In order to sketch a derivation, one should know that, if $x(t)$ is the solution of the Itô SDE and f is an arbitrary function of $x(t)$, f will obey

$$\frac{d}{dt} f[x(t), t] = \frac{\partial f}{\partial x} [x(t), t] (a(x, t) + b(x, t) \xi) + \frac{1}{2} b(x, t)^2 \frac{\partial^2 f}{\partial x^2} [x(t), t] + \frac{\partial f}{\partial t} [x(t), t]. \quad (1.20)$$

Consider now a function with no explicit dependence on time. The equation for the *average* of df/dt will have the dependence on the stochastic force removed, since ξ has zero average and is to be considered uncorrelated with any other part of the stochastic process:

$$\left\langle \frac{d}{dt} f(x) \right\rangle = \left\langle f'(x) a(x, t) + \frac{1}{2} b(x, t)^2 f''(x) \right\rangle.$$

However, the averages can also be computed as an integral over the conditional probability density, i.e. $\langle f(x(t)) \rangle = \int f(x) P(x, t|x_0, t_0) dx$, so that

$$\begin{aligned} \frac{d}{dt} \langle f(x) \rangle &= \int f(x) \frac{d}{dt} P(x, t|x_0, t_0) dx = \\ &= \int P(x, t|x_0, t_0) \left[f'(x) a(x, t) + \frac{1}{2} b(x, t)^2 f''(x) \right] dx. \end{aligned}$$

This second integral can be integrated by parts. Then, by considering that - in order to have a normalized probability density - $P(x, t|x_0, t_0)$ must vanish at infinity, the

boundary terms disappear⁸ and one is left with

$$\int f(x) \frac{d}{dt} P(x, t|x_0, t_0) dx = \int f(x) \left\{ -\frac{\partial}{\partial x} [a(x, t) P(x, t|x_0, t_0)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [b(x, t)^2 P(x, t|x_0, t_0)] \right\} dx.$$

Since $f(x)$ is arbitrary, we have proof that the Itô SDE (1.17) corresponds to a conditional probability density which satisfies the Fokker-Planck equation

$$\frac{d}{dt} P(x, t|x_0, t_0) = -\frac{\partial}{\partial x} [a(x, t) P(x, t|x_0, t_0)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [b(x, t)^2 P(x, t|x_0, t_0)]. \quad (1.21)$$

A similar derivation can also be performed in the case of the Stratonovich integral, leading to a different form of the FP equation. Again, if b does not depend on x , as will be the case for the SDEs treated in this thesis, there is no difference between the two approaches.

1.3 The Ornstein-Uhlenbeck process

In the previous sections we have only scratched the surface of the mathematical treatment of stochastic processes. We now have tools that are sufficient for the treatment of a very simple, one-dimensional process. This is probably the simplest SDE leading to a stationary probability distribution and the main ingredient behind our generalized Langevin equation thermostat. The Ornstein-Uhlenbeck process is characterized by the SDE⁹

$$\dot{u} = -au + b\xi, \quad (1.22)$$

i.e. there is a constant diffusion term b and a linear drift with a coefficient a . The corresponding Fokker-Planck equation is obtained by (1.21), and reads

$$\frac{\partial}{\partial t} P(u, t|u_0, 0) = \frac{\partial}{\partial u} (au P(u, t|u_0, 0)) + \frac{1}{2} d \frac{\partial^2}{\partial u^2} P(u, t|u_0, 0),$$

with $d = b^2$.

The corresponding stationary probability $P(u) = \lim_{t \rightarrow \infty} P(u, t|u_0, 0)$ is found by solving

$$\frac{\partial^2}{\partial u^2} P(u) = -\frac{2au}{d} \frac{\partial}{\partial u} P(u), \quad \Rightarrow \quad \frac{\partial}{\partial u} P(u) = -\frac{2au}{d} P(u)$$

⁸One obtains for the drift term

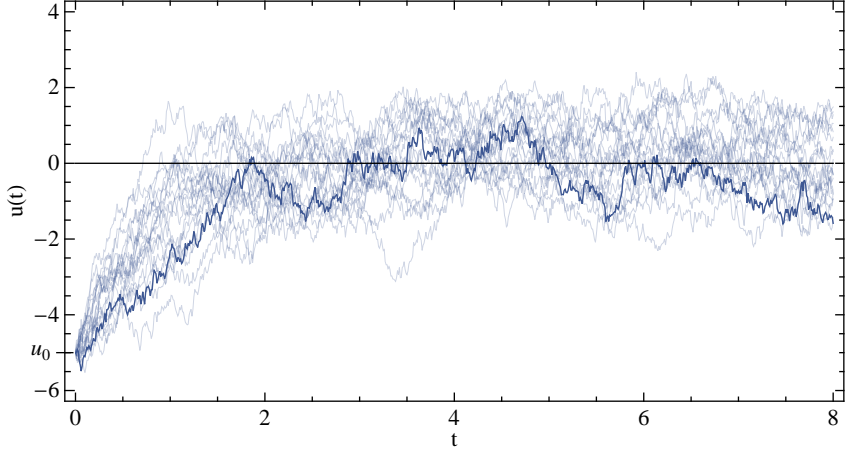
$$\int f'(x) a(x, t) P(x, t|x_0, t_0) dx = \left[f'(x) a(x, t) P(x, t|x_0, t_0) \right]_{-\infty}^{\infty} - \int f(x) \frac{\partial}{\partial x} [a(x, t) P(x, t|x_0, t_0)] dx$$

and for the diffusion part, integrating by parts twice,

$$\int P(x, t|x_0, t_0) b(x, t)^2 f''(x) dx = \int f(x) \frac{\partial^2}{\partial x^2} [b(x, t)^2 P(x, t|x_0, t_0)] dx.$$

⁹In this case, the Itô and Stratonovich interpretations are equivalent

Figure 1.6: Several realizations of the Ornstein-Uhlenbeck process with $a = b = 1$, starting at $u(0) = u_0$, are represented on the same plot. One of the trajectories is highlighted for clarity. One can see how the swarm of trajectories corresponds to the evolution of the probability density represented in figure 1.3(b).



which yields

$$P(u) = \frac{1}{\sqrt{2\pi c}} e^{-u^2/2c} \quad (1.23)$$

which is a Gaussian with variance $c = d/2a$. The solution for a generic time is still a Gaussian, although its breadth and center are time dependent (see figure 1.3(b)), which can be verified by substituting

$$P(u, t|u_0, 0) = \frac{1}{\sqrt{2\pi c(1 - e^{-2at})}} \exp \left[-\frac{(u - u_0 e^{-at})^2}{c(1 - e^{-2at})} \right], \quad (1.24)$$

into the FP equation.

The Ornstein-Uhlenbeck SDE (1.6) can also be solved using stochastic integration. Consider the stochastic process $y = ue^{at}$. According to the chain rule (1.20), the corresponding SDE will be¹⁰

$$\dot{y} = e^{at} \dot{u} + \frac{1}{2} b^2 \frac{d^2 y}{du^2} + aue^{at} = e^{at} (-au + b\xi) + aue^{at} = be^{at} \xi$$

¹⁰One must also consider the explicit dependence of y on time, hence the term aue^{at} .

Which can be solved to give the correlation properties¹¹

$$\langle y(t) \rangle = y(0), \quad \langle y(t) y(s) \rangle = c \left(e^{2a \min(t,s)} - 1 \right)$$

from which

$$\langle u(t) \rangle = e^{-at} u(0), \quad \langle (u(t) - \langle u(t) \rangle) (u(s) - \langle u(s) \rangle) \rangle = c \left(e^{-a|t-s|} - e^{-a(t+s)} \right) \quad (1.25)$$

At long times, the dependence on $x(0)$ is exponentially damped, and one is left with $\langle u(t)^2 \rangle \rightarrow c$, which is consistent with the stationary distribution (1.23). Note that Eqs. (1.25) provide a recipe for generating sample paths consistent with the OU stochastic differential equation (figure 1.6). If ξ_i are uncorrelated Gaussian numbers with zero mean and unit variance, the points

$$u(0) = u_0, \quad u(i\Delta t) = e^{-a\Delta t} u((i-1)\Delta t) + \sqrt{c(1 - e^{-2a\Delta t})} \xi_i$$

will be distributed in agreement with 1.24, and will hence correspond to points taken from an OU sample path with spacing Δt starting at u_0 .

1.3.1 The integral of the OU process

As a slightly more complicated example, let us consider the stochastic process described by the pair of coupled SDEs

$$\begin{aligned} \dot{x} &= u \\ \dot{u} &= -au + b\xi. \end{aligned} \quad (1.26)$$

We consider this case for two reasons. Firstly, it allows us to introduce the ideas which are necessary to extend the results of the previous section to arbitrary dimensions. Secondly, the integrated OU process x is the solution of the Langevin equation $\ddot{x} = -a\dot{x} + b\xi$, which allows us to explicitly compare Einstein's and Langevin's solutions for the Brownian motion problem.

One can proceed using the one-dimensional case as a guideline, considering a

¹¹To compute the Itô integral of a deterministic function $f(t)$, $h(t) = h(t_0) + \int_{t_0}^t f(s) dw(s)$ one must compute the partial sums, which read (with the shorthands $f(t_i) = f_i$ and $w(t_i) = w_i$)

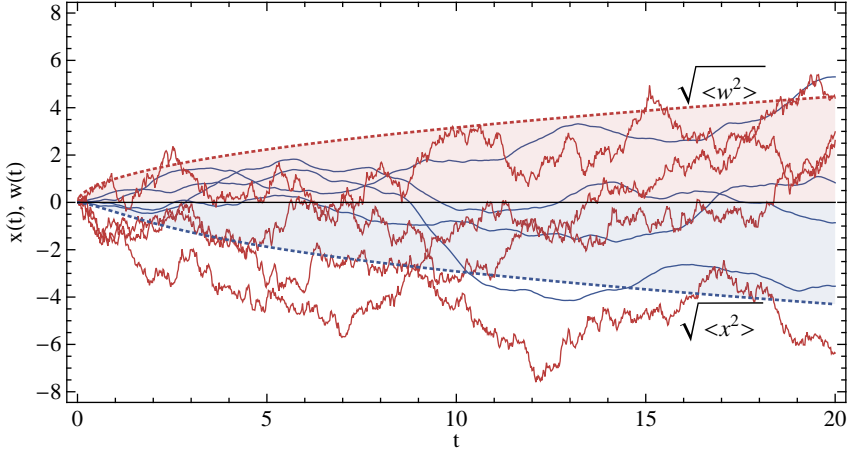
$$S_n = \sum_{i=1}^n f_{i-1} (w_i - w_{i-1});$$

each $\Delta w_i = w_i - w_{i-1}$ is a Gaussian-distributed random variable with zero mean and whose spread depends on the time interval $\Delta t_i = t_i - t_{i-1}$. To characterize the integral $h(t)$, we can compute its average, which is $h(t_0)$, since all the terms have zero mean, and its variance, which (exploiting the independence of intervals) reads

$$\sum_{i=1}^n \sum_{j=1}^n f_{i-1} f_{j-1} \langle \Delta w_i \Delta w_j \rangle = \sum_{i=1}^n (f_{i-1})^2 \langle \Delta w_i^2 \rangle = \sum_{i=1}^n (f_{i-1})^2 \Delta t_i$$

which converges to the Riemann integral $\int_{t_0}^t f(s)^2 ds$.

Figure 1.7: Four sample paths of a Wiener process (red) and of an integrated Ornstein-Uhlenbeck process with $a = b = 1$, starting at $x(0) = u(0) = 0$ (blue). They have the same long-time behavior, but the integrated OU process is much more regular on a short time scale. The time dependence of the root mean square deviation is also plotted, with thick, dashed lines.



state vector (x, u) and drift and diffusion *matrices* A and B , respectively:

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 0 & a \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix}$$

Eventually, one obtains the 2-dimensional analogue of Eqs (1.25), which read

$$\begin{aligned} \left\langle \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} \right\rangle &= e^{-\mathbf{A}t} \begin{pmatrix} x(0) \\ u(0) \end{pmatrix} \\ \left\langle \begin{pmatrix} x(t) - \langle x(t) \rangle \\ u(t) - \langle u(t) \rangle \end{pmatrix} \begin{pmatrix} x(t') - \langle x(t') \rangle \\ u(t') - \langle u(t') \rangle \end{pmatrix}^T \right\rangle &= \int_0^{\min(t, t')} e^{-(t-s)\mathbf{A}} \mathbf{B} \mathbf{B}^T e^{-(t'-s)\mathbf{A}^T} ds. \end{aligned}$$

for the average and the two-times correlation matrix.

In this simple two dimensional case an analytical solution can be found, allowing us to write down the correlation properties of x alone in a closed form:

$$\langle x(t) \rangle = x(0) + \frac{1 - e^{-at}}{a} u(0), \quad \langle (x(t) - \langle x(t) \rangle)^2 \rangle = \frac{c}{a^2} [2at + 4e^{-at} - 3 - e^{-2at}],$$

which completely determine the statistical properties of the integrated OU process.

It is instructive to compare this solution with the naive (and much simpler) derivation originally carried out by Langevin starting from (1.15). If x and u are interpreted as the position and velocity of a Brownian particle respectively, one finds that $c = \langle \dot{x}^2 \rangle = k_B T / m$, as $\lim_{t \rightarrow \infty} \langle u(t)^2 \rangle = c$. In addition, by comparing (1.15) and (1.26), one recognizes that $a = \gamma$. Combining these results, one can eventually

find that the long-time behavior of $\langle x(t)^2 \rangle$ is given by

$$\frac{d}{dt} \langle x(t)^2 \rangle = D = 2c/a = 2 \frac{k_B T}{m\gamma}.$$

We have therefore recovered the Langevin's original result with a more rigorous approach. We can now juxtapose the actual trajectory of $x(t)$ with those of a Wiener process with the same diffusion coefficient, and thus finally compare the trajectories resulting from Langevin's and Einstein's models for Brownian motion. For large t the two random processes have the same behavior, with a variance growing linearly with time. Their short-time behavior is however very different, with the integrated OU process having much smoother trajectories, since its derivative is continuous, and finite almost everywhere (see figure 1.7).

1.3.2 The multivariate OU process

We will conclude this chapter by briefly describing how the Ornstein-Uhlenbeck process can be extended to the cases with many variables. The calculation of the integrated OU process was already an example of this problem. However, in that case we had the additional complication of a drift matrix with a zero eigenvalue, leading to a covariance growing without bounds. Instead, we now focus on a positive-definite \mathbf{A} matrix, which guarantees that the OU process will have a well-defined stationary probability.

Let \mathbf{u} be a vector in \mathbb{R}^n which describes the state of the system, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ be the drift and diffusion matrices respectively and ξ be a vector of n uncorrelated Gaussian processes, such that $\langle \xi_i \rangle = 0$ and $\langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij} \delta(t - t')$. The SDE for the multivariate OU process then reads

$$\dot{\mathbf{u}} = -\mathbf{A}\mathbf{u} + \mathbf{B}\xi, \quad (1.27)$$

and the associated Fokker-Planck equation is

$$\frac{\partial}{\partial t} P(\mathbf{u}, t | \mathbf{u}_0, 0) = \sum_{ij} A_{ij} \frac{\partial}{\partial u_i} [u_j P(\mathbf{u}, t | \mathbf{u}_0, 0)] + \frac{1}{2} \sum_{ij} D_{ij} \frac{\partial^2}{\partial u_i \partial u_j} P(\mathbf{u}, t | \mathbf{u}_0, 0), \quad (1.28)$$

where $\mathbf{D} = \mathbf{B}\mathbf{B}^T$. Its solution can be found with conventional techniques, but the derivation is lengthy (see e.g. Ref.[18, p. 153ss]). A simpler, neat solution can be found by integration of (1.27) subject to the boundary condition $\mathbf{u}(0) = \mathbf{u}_0$ [16, p. 109ss]. Since $\mathbf{u}(t)$ is generated as a superposition of Gaussian increments, it will have a Gaussian distribution. Therefore, computing the average and covariance of $\mathbf{u}(t)$ is sufficient for its characterization. These can be obtained in the spirit described previously for the one-dimensional case:

$$\begin{aligned} \langle \mathbf{u}(t) \rangle &= e^{-\mathbf{A}t} \mathbf{u}_0 \\ \langle (\mathbf{u}(t) - \langle \mathbf{u}(t) \rangle)^T (\mathbf{u}(t) - \langle \mathbf{u}(t) \rangle) \rangle &= e^{-t\mathbf{A}} \left[\int_0^t e^{s\mathbf{A}} \mathbf{D} e^{s\mathbf{A}^T} ds \right] e^{-t\mathbf{A}^T}. \end{aligned} \quad (1.29)$$

Let us now define the static covariance matrix \mathbf{C} , as the symmetric solution of the linear equation

$$\mathbf{A}\mathbf{C} + \mathbf{C}\mathbf{A}^T = \mathbf{D} = \mathbf{B}\mathbf{B}^T. \quad (1.30)$$

One can then find an explicit solution for the integral in (1.29), and hence a closed form for the finite-time covariance:

$$\langle (\mathbf{u}(t) - \langle \mathbf{u}(t) \rangle)^T (\mathbf{u}(t) - \langle \mathbf{u}(t) \rangle) \rangle = \mathbf{C} - e^{-\mathbf{A}t} \mathbf{C} e^{-\mathbf{A}^T t}. \quad (1.31)$$

Now, if one starts out from a deterministic initial configuration \mathbf{u}_0 , which in probabilistic terms would correspond to the boundary condition $P(\mathbf{u}, 0 | \mathbf{u}_0, 0) = \delta(\mathbf{u} - \mathbf{u}_0)$, at time t a Gaussian probability would be obtained, whose mean and covariance are given in Eqs. (1.29) and (1.31). We can therefore write the solution to the Fokker-Planck equation (1.28), which reads

$$P(\mathbf{u}, t | \mathbf{u}_0, 0) = \frac{\exp \left[-\frac{1}{2} (\mathbf{u} - e^{-t\mathbf{A}} \mathbf{u}_0)^T (\mathbf{C} - e^{-t\mathbf{A}} \mathbf{C} e^{-t\mathbf{A}^T})^{-1} (\mathbf{u} - e^{-t\mathbf{A}} \mathbf{u}_0) \right]}{\sqrt{(2\pi)^n \det(\mathbf{C} - e^{-t\mathbf{A}} \mathbf{C} e^{-t\mathbf{A}^T})}}. \quad (1.32)$$

In the infinite-time limit, having taken \mathbf{A} to be positive definite, $P(\mathbf{u}, t | \mathbf{u}_0, 0)$ converges to the stationary probability density

$$P_s(\mathbf{u}) = \frac{e^{-\frac{1}{2} \mathbf{u}^T \mathbf{C}^{-1} \mathbf{u}}}{\sqrt{(2\pi)^n \det(\mathbf{C})}}.$$

A number of other analytical results can be obtained for the multidimensional OU process, including high-order time correlation functions (see appendix D). We will sketch the derivations as needed, when we will discuss their use in the development of the generalized Langevin equation thermostat, which is the subject of the next chapter.

The Generalized Langevin Equation Thermostat

As discussed in the introduction, the aim of this thesis is to present our approach for turning a generalized Langevin equation into a flexible and convenient tool for enhancing the performances or extending the capabilities of molecular dynamics. However, this far the only non-Markovian stochastic equation that has appeared is Eq. (1). Despite its apparent simplicity, dealing with an integral-differential equation such as (1) is a daunting task, both from the analytical and computational points of view.

The program of this chapter and the recipe for our method consist of three main points:

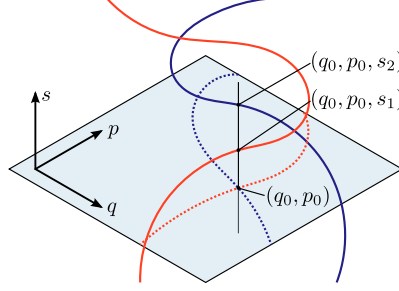
- expressing non-Markovian equations of motion in a Markovian form which is easy to treat analytically and to implement on a computer
- predicting the properties of the resulting dynamics in an inexpensive way, by using the exact, analytical results that can be obtained for a physically-motivated model system
- fitting the parameters which characterize the GLE so as to enhance the desired properties of the dynamics on the system of interest.

2.1 Markovian and non-Markovian formulations

A colored-noise Langevin dynamics for a particle with unit mass, position q and momentum p , subject to a potential $V(q)$ can be written in its non-Markovian form,

$$\begin{aligned}\dot{q} &= p \\ \dot{p} &= -V'(q) - \int_{-\infty}^t K(t-s)p(s)ds + \zeta,\end{aligned}\tag{2.1}$$

Figure 2.1: A cartoon representing how a Markovian dynamics can turn to a non-Markovian one when projected down to a lower-dimensionality phase space. The information contained in the s variable, which distinguishes between the red and the blue trajectories at the marked point, is embedded into the past history of the dynamics projected on the (q, p) subspace.



where $K(t)$ is a memory kernel describing the friction, and $\zeta(t)$ is a Gaussian random process whose time correlation function is $H(t) = \langle \zeta(t) \zeta(0) \rangle$. Algorithms exist that can be used to generate a sequence of random numbers with prescribed time correlation efficiently[19, 20]. However, in order to introduce a history-dependent friction kernel $K(t)$ in a straightforward manner, one would have to store the past trajectory of the momenta, leading to a procedure which would be inefficient from a computer time and memory requirements points of view.

On the other hand, it is known that integrating out some of the degrees of freedom in a Markovian dynamics gives rise to a non-Markovian dynamics in the remaining variables[3, 4] (see figure 2.1). Conversely, a Markovian dynamics in an extended phase-space has often been used as a practical method to simplify the treatment of non-Markovian problems[6, 21, 22]. In fact it has been proven that any non-Markovian dynamics whose memory kernels can be represented as a continued-fraction expansion can be expressed through a Markovian formalism[23].

Inspired by Mori-Zwanzig results, we supplement the dynamical variables (q, p) by a set of n additional momenta \mathbf{s} , which will be bilinearly coupled to the physical momentum p , so as to construct a Markovian Langevin equation:

$$\begin{aligned} \dot{q} &= p \\ \begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} &= \begin{pmatrix} -V'(q) \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_p^T \\ \bar{\mathbf{a}}_p & \mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} b_{pp} & \mathbf{b}_p^T \\ \bar{\mathbf{b}}_p & \mathbf{B} \end{pmatrix} \begin{pmatrix} \xi \\ \boldsymbol{\xi} \end{pmatrix}, \end{aligned} \quad (2.2)$$

where $\boldsymbol{\xi}$ is a vector of $n+1$, uncorrelated Gaussian numbers, i.e. $\langle \xi_i(t) \xi_j(0) \rangle = \delta_{ij} \delta(t)$. This looks very similar to the multivariate Ornstein-Uhlenbeck SDE (1.27), but for the possibly non-linear coupling between p and q . As such, we will use several of the results listed in section 1.3.2. In Eqs. (2.2) we introduced a compact notation to refer to the portions of the matrices which describe the coupling between different parts of the state vector. Given the extended phase-space vector $\mathbf{x} = (q, p, \mathbf{s})$, we will distinguish between a matrix acting on the full \mathbf{x} and one connecting subsets of its

components, according to the scheme

$$\begin{array}{c}
 \begin{array}{ccc}
 & q & p & s \\
 q & m_{qq} & m_{qp} & \mathbf{m}_q^T \\
 p & \tilde{m}_{qp} & m_{pp} & \mathbf{m}_p^T \\
 s & \tilde{\mathbf{m}}_q & \tilde{\mathbf{m}}_p & \mathbf{M}
 \end{array}
 \end{array}
 \left. \vphantom{\begin{array}{ccc} & q & p & s \\ q & m_{qq} & m_{qp} & \mathbf{m}_q^T \\ p & \tilde{m}_{qp} & m_{pp} & \mathbf{m}_p^T \\ s & \tilde{\mathbf{m}}_q & \tilde{\mathbf{m}}_p & \mathbf{M} \end{array}} \right\} \mathbf{M}_p \left. \vphantom{\begin{array}{ccc} & q & p & s \\ q & m_{qq} & m_{qp} & \mathbf{m}_q^T \\ p & \tilde{m}_{qp} & m_{pp} & \mathbf{m}_p^T \\ s & \tilde{\mathbf{m}}_q & \tilde{\mathbf{m}}_p & \mathbf{M} \end{array}} \right\} \mathbf{M}_{qp} \quad (2.3)$$

For example, in Eqs. (2.2) the assembly of a_{pp} , \mathbf{a}_p , $\tilde{\mathbf{a}}_p$ and \mathbf{A} corresponds to the deterministic drift matrix \mathbf{A}_p , while the (p, \mathbf{s}) diffusion matrix is designated \mathbf{B}_p .

The connection between Eqs. (2.2) and (2.1) can be obtained in the same manner as the classical application of the Mori-Zwanzig formalism to a harmonic bath[4]. The additional degrees of freedom \mathbf{s} can be integrated out easily, because the evolution of the (p, \mathbf{s}) variables in the free-particle limit of Eqs. (2.2) is described by a linear, Markovian stochastic differential equation. For the sake of completeness, we sketch the derivation in appendix A.

In the same appendix we also show that the memory kernel $K(t)$ and the noise correlation $H(t)$ are related to the drift and diffusion matrices by

$$\begin{aligned}
 K(t) &= 2a_{pp}\delta(t) - \mathbf{a}_p^T e^{-|t|\mathbf{A}} \tilde{\mathbf{a}}_p \\
 H(t) &= d_{pp}\delta(t) + \mathbf{a}_p^T e^{-|t|\mathbf{A}} [\mathbf{Z}\mathbf{a}_p - \mathbf{d}_p],
 \end{aligned} \quad (2.4)$$

where we defined¹ the matrices $\mathbf{D}_p = \mathbf{B}_p \mathbf{B}_p^T$ and $\mathbf{Z} = \int_0^\infty e^{-\mathbf{A}t} \mathbf{D} e^{-\mathbf{A}^T t} dt$. Therefore, the Markovian equations (2.2) can generate a trajectory corresponding to finite-memory kernels which are in principle an arbitrary combination of complex exponentials².

2.1.1 Conditions for canonical sampling

In order to define the conditions which are required for Eqs. (2.2) to sample the canonical, constant-temperature ensemble, it is useful to introduce also the static covariance matrix for the free-particle dynamics of (p, \mathbf{s}) , $\mathbf{C}_p = \langle (p, \mathbf{s})^T (p, \mathbf{s}) \rangle$, which is related to the drift and diffusion matrices by

$$\mathbf{A}_p \mathbf{C}_p + \mathbf{C}_p \mathbf{A}_p^T = \mathbf{B}_p \mathbf{B}_p^T. \quad (2.5)$$

¹The careful reader may have recognized the same integral appearing in (1.29), and wonder why we have not written

$$\int_0^\infty e^{-\mathbf{A}t} \mathbf{D} e^{-\mathbf{A}^T t} dt = \mathbf{C}, \quad \mathbf{A}\mathbf{C} + \mathbf{C}\mathbf{A}^T = \mathbf{D}.$$

While this would be perfectly legitimate, we want to avoid confusion between the matrix \mathbf{Z} (which is the covariance matrix of a hypothetical OU process based on the \mathbf{s} parts of the drift and diffusion matrices only) and \mathbf{C} which we will use below to refer to the \mathbf{s} block of a larger covariance matrix \mathbf{C}_p which corresponds to the OU process built out of \mathbf{A}_p and \mathbf{B}_p . We chose the notation (2.3) as we deem it to be a reasonable compromise that allows a concise labeling for the different bits of the large matrices we use. The overall size of the phase space to be considered can be understood from the context.

²There are a number of mathematical and physical constraints imposed on the kernels, which are discussed in section 2.3. Within those limits, any kernel can be represented approximately; clearly, it is not possible to obtain *exactly* an asymptotic decay $K(t) \sim t^{-n}$.

Eq. (2.5) is nothing but the equivalent of Eq. (1.30) for the OU process generated by \mathbf{A}_p and \mathbf{B}_p . Parts of \mathbf{C}_p enter the expression for the Fourier transform of the memory kernels (2.4):

$$\begin{aligned} K(\omega) &= 2a_{pp} - 2\mathbf{a}_p^T \frac{\mathbf{A}}{\mathbf{A}^2 + \omega^2} \bar{\mathbf{a}}_p \\ H(\omega) &= K(\omega) \left(c_{pp} - \mathbf{a}_p^T \frac{\mathbf{A}}{\mathbf{A}^2 + \omega^2} \mathbf{c}_p \right) + \\ &\quad + 2\omega^2 \left(\mathbf{a}_p^T \frac{1}{\mathbf{A}^2 + \omega^2} \mathbf{c}_p \right) \left(1 + \mathbf{a}_p^T \frac{1}{\mathbf{A}^2 + \omega^2} \bar{\mathbf{a}}_p \right). \end{aligned} \quad (2.6)$$

Eqs. (2.6) show clearly that, in order to fulfill the fluctuation-dissipation theorem[3, 24], which requires $H(\omega) = k_B T K(\omega)$, one must choose $c_{pp} = k_B T$ and $\mathbf{c}_p = 0$. Since the kernels - and therefore the dynamics of (q, p) - do not depend on the \mathbf{C} block in \mathbf{C}_p , one can simply choose $\mathbf{C}_p = k_B T$, by which (2.5) simplifies to³

$$\mathbf{D}_p = \mathbf{B}_p \mathbf{B}_p^T = k_B T (\mathbf{A}_p + \mathbf{A}_p^T). \quad (2.7)$$

Together with Eq. (2.5), the fluctuation-dissipation theorem fixes \mathbf{D}_p once \mathbf{A}_p is given. As we will see in chapter 5, a number of useful properties can be obtained if the fluctuation-dissipation theorem is violated, and we will therefore consider the general case throughout this chapter.

It is worth noting that one can also justify (2.7) and its connection to canonical sampling based on a Fokker-Planck formalism. One can write the Fokker-Planck equation associated with the Markovian GLE (2.2) as

$$\frac{\partial}{\partial t} P = [\mathcal{L}_{HAM} + \mathcal{L}_{GLE}] P,$$

where we define the operators

$$\begin{aligned} \mathcal{L}_{HAM} &= -p \frac{\partial}{\partial q} \cdot + V'(q) \frac{\partial}{\partial p} \cdot, \\ \mathcal{L}_{GLE} &= \frac{\partial}{\partial p} \left[a_{pp} p \cdot + \frac{1}{2} d_{pp} \frac{\partial}{\partial p} \cdot \right] + \sum_{ij} \frac{\partial}{\partial s_i} \left[(\mathbf{A})_{ij} s_j \cdot + \frac{1}{2} (\mathbf{D})_{ij} \frac{\partial}{\partial s_i} \cdot \right] + \\ &\quad + \frac{\partial}{\partial p} \sum_i \left[(\mathbf{a}_p)_i s_i \cdot + (\mathbf{d}_p)_i \frac{\partial}{\partial s_i} \cdot \right] + \sum_i \frac{\partial}{\partial s_i} \left[(\bar{\mathbf{a}}_p)_i p \cdot + (\mathbf{d}_p)_i \frac{\partial}{\partial p} \cdot \right]. \end{aligned} \quad (2.8)$$

Note that \mathcal{L}_{HAM} corresponds to the Liouville operator for a purely Hamiltonian trajectory, while \mathcal{L}_{GLE} would describe an infinitesimal evolution of the marginal probability of (p, \mathbf{s}) in the absence of any coupling between p and q . Consider now a

³Note that Eq. (2.7) is a *sufficient* condition to have canonical sampling on the physical (q, p) variables, but it is *necessary* to have a canonical-like distribution for \mathbf{s} as well. Since we will always use Eq. (2.7) to ensure that the fluctuation-dissipation theorem is satisfied, we will use “satisfies Eq. (2.7)” and “satisfies the FDT” as equivalents, even if strictly speaking they are not. This issue is discussed further in the introduction of chapter 5.

tentative stationary probability distribution

$$P(q, p, \mathbf{s}) \propto \exp \left[-\frac{1}{k_B T} \left(V(q) + \frac{1}{2} p^2 + \frac{1}{2} \mathbf{s}^2 \right) \right]. \quad (2.9)$$

It is simple albeit tedious to verify that if (2.7) holds, both \mathcal{L}_{HAM} and \mathcal{L}_{GLE} leave P unchanged⁴. A similar partitioning of the Liouville operator will be used in deriving a strategy for numerical integration of Eqs. (2.2) (see section 2.5 below).

The case in which (2.7) does not hold is much more complex, and finding an expression of the stationary probability distribution is exceedingly difficult. In fact, we were not even able to prove rigorously that a stationary probability distribution *exists* for an arbitrary potential. However, there are some clues that this is the case, which are supported by an argument in the harmonic limit (see appendix B) and by the empirical evidence of the many simulations we have performed (see chapter 5). Firstly, we know that if \mathbf{A}_p is positive definite the dynamics of (p, \mathbf{s}) alone will have a stationary distribution. An argument presented in Ref. [25] demonstrates that $K(\omega) > 0$ for all ω is a sufficient condition to guarantee that (2.1) with zero potential is consistent with the second law of thermodynamics. Since - roughly speaking - this means that any disturbance to the stable dynamics of the (p, \mathbf{s}) will be dissipated away, it seems that $K(\omega) > 0$ is a desirable condition to require in order to have a stable dynamics regardless of the details of the noise and of the potential.

Having discussed the relation between the non-Markovian Langevin equation (2.1) and its Markovian counterpart, let us now present a few simple examples, to get a flavor of the flexibility of this GLE framework. As we will see in the next chapters, some of these examples can also be used to demonstrate the concepts to be used in different applications of the GLE thermostat.

2.1.2 Exponential correlations

The case in which $K(t)$ and/or $H(t)$ decay as a simple exponential can be obtained easily, using a single auxiliary variable s . For example, to obtain a dynamics where both $K(t)$ and $H(t)$ are exponentials, one may set

$$\mathbf{A}_p = \begin{pmatrix} 0 & -\sqrt{\gamma/\tau} \\ \sqrt{\gamma/\tau} & 1/\tau \end{pmatrix}, \quad \mathbf{D}_p = \begin{pmatrix} 0 & 0 \\ 0 & 2k_B T/\tau \end{pmatrix}, \quad (2.10)$$

which corresponds to $K(t) = \gamma e^{-t/\tau}/\tau$ and $H(t) = k_B T K(t)$. The fluctuation-dissipation theorem is fulfilled, and the dynamics will sample the canonical ensemble. We will return to the parameters (2.10) in section 3.2, where we will exploit

⁴We have gathered together terms in the \mathcal{L}_{GLE} part of the Fokker-Plank equation to make the derivation more transparent. For example, the first term would generate

$$\left(a_{pp} - \frac{d_{pp}}{2k_B T} \right) p.$$

which is zero if $d_{pp} = 2k_B T a_{pp}$ as implied by (2.7).

the fact that the power spectrum of the corresponding noise is

$$H(\omega) \propto \frac{1}{1 + \tau^2 \omega^2}.$$

All the frequencies greater than $1/\tau$ are damped, as if a low-pass filter had been applied. Hence, we will use (2.10) as the prototype for a stochastic thermostat which does not disturb the high-frequency modes.

Dynamics which break the fluctuation-dissipation theorem are readily obtained. For instance, by taking

$$\mathbf{A}_p = \begin{pmatrix} \gamma & -1/\tau \\ 0 & 1/\tau \end{pmatrix}, \quad \mathbf{D}_p = \begin{pmatrix} 0 & 0 \\ 0 & 2\gamma k_B T \end{pmatrix} \quad (2.11)$$

one can obtain an exponentially correlated noise, $H(t) = k_B T \gamma e^{-t/\tau}/\tau$, and an instantaneous friction $K(t) = 2\gamma \delta(t)$. This dynamics has been used to compute approximate quantum corrections to specific heat[26].

Last of all the opposite case, with an exponential friction kernel $K(t) = \gamma e^{-t/\tau}/\tau$ and uncorrelated noise $H(t) = 2\gamma k_B T \delta(t)$, is obtained with the parameters

$$\mathbf{A}_p = \begin{pmatrix} 0 & -\sqrt{\gamma/\tau} \\ \sqrt{\gamma/\tau} & 1/\tau \end{pmatrix}, \quad \mathbf{D}_p = \begin{pmatrix} 2\gamma k_B T & 0 \\ 0 & 0 \end{pmatrix}. \quad (2.12)$$

2.1.3 δ -like memory kernels

Proceeding to a slightly more complex example, let us demonstrate how to construct a generalized Langevin equation whose memory kernels' Fourier transform has a Dirac- δ -like shape. Consider the stochastic differential equation which derives from the choice of parameters⁵

$$\mathbf{A}_p = \begin{pmatrix} 0 & \sqrt{\frac{\gamma}{2\pi}} & \sqrt{\frac{\gamma}{2\pi}} \\ -\sqrt{\frac{\gamma}{2\pi}} & \Delta\omega & \omega_0 \\ -\sqrt{\frac{\gamma}{2\pi}} & -\omega_0 & \Delta\omega \end{pmatrix}, \quad (2.13)$$

$$\mathbf{D}_p = k_B T (\mathbf{A}_p + \mathbf{A}_p^T). \quad (2.14)$$

⁵Also other choices are possible which lead to δ -like kernels[27], such as

$$\mathbf{A}_p = \begin{pmatrix} 0 & \sqrt{\gamma/2\pi} & 0 \\ -\sqrt{\gamma/2\pi} & \Delta\omega & \omega_0 \\ 0 & -\omega_0 & 0 \end{pmatrix},$$

which generates

$$K(\omega) = 2\gamma\omega^2 \frac{(\omega - \omega_0)^2 + \Delta\omega^2}{(\omega^2 - \omega_0^2)^2 + \Delta\omega^2\omega^2} \frac{1}{\pi} \frac{\Delta\omega}{(\omega - \omega_0)^2 + \Delta\omega^2}.$$

This kernel has the nice feature of being zero for $\omega = 0$ for finite $\Delta\omega$.

Plugging these parameters into Eqs. (2.6), one readily obtains the Fourier transform of the corresponding memory kernels:

$$K(\omega) = 2\gamma \frac{\Delta\omega^2 + \omega_0^2 + \omega^2}{(\omega + \omega_0)^2 + \Delta\omega^2} \frac{1}{\pi} \frac{\Delta\omega}{(\omega - \omega_0)^2 + \Delta\omega^2} \quad (2.15)$$

$$H(\omega) = k_B T K(\omega). \quad (2.16)$$

If $\Delta\omega$ is small $K(\omega)$ and $H(\omega)$ have a $\delta(\omega - \omega_0)$ -like shape. Eqs. (2.14) and (2.16) express the classical fluctuation-dissipation theorem. We note in passing that - together with the combination rules to be discussed below in section 2.4 - a combination of δ -like kernels centered on different frequencies can be used to construct a SDE which approximates arbitrary memory functions.

2.2 The harmonic oscillator

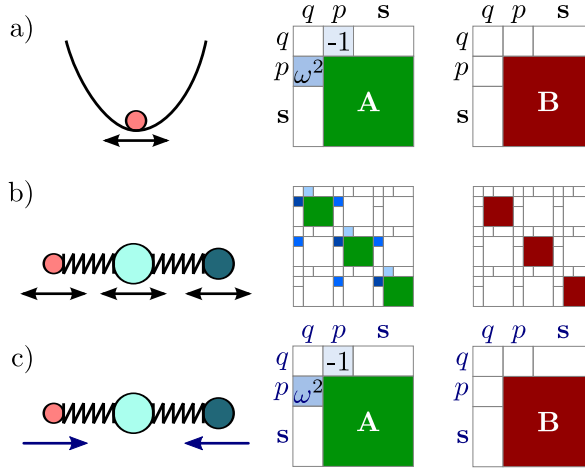
In the previous section we have discussed the connections between the Markovian and non-Markovian formulations of the generalized Langevin equation thermostat, Eqs. (2.2) and (2.1). This relation, which is expressed by the connection between the drift and diffusion matrices and the non-Markovian memory, holds regardless of the potential $V(q)$. However, the actual properties of the trajectory will obviously depend on the details of $V(q)$. Hence, if one wants to optimize a selected property for a given problem, the behavior of the colored-noise thermostat in those specific circumstances must be studied.

Typically, the thermostats used in molecular simulations have a few parameters, that are chosen by trial and error. A thermostat based on Eqs. (2.2) depends on a much larger number of parameters, and hence a purely empirical approach is impractical. It is therefore important to find ways to compute *a priori* analytical estimates which can be used to guide the tuning of the thermostat. To this end, we examine the harmonic oscillator, which is commonly used to model physical and chemical systems. By choosing $V(q) = \frac{1}{2}\omega^2 q^2$ the force term in (2.2) is linear, and the dynamics of $\mathbf{x} = (q, p, \mathbf{s})^T$ becomes the OU process $\dot{\mathbf{x}} = -\mathbf{A}_{qp}\mathbf{x} + \mathbf{B}_{qp}\boldsymbol{\xi}$. In Eqs. (2.2) the \mathbf{s} degrees of freedom are coupled to the momentum only. Therefore, most of the additional entries in \mathbf{A}_{qp} and \mathbf{B}_{qp} are zero, and the equations for \mathbf{x} read

$$\begin{pmatrix} \dot{q} \\ \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} = - \begin{pmatrix} 0 & -1 & \mathbf{0} \\ \omega^2 & a_{pp} & \mathbf{a}_p^T \\ \mathbf{0} & \bar{\mathbf{a}}_p & \mathbf{A} \end{pmatrix} \begin{pmatrix} q \\ p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} 0 & 0 & \mathbf{0} \\ 0 & \mathbf{B}_p \\ \mathbf{0} & \mathbf{B}_p \end{pmatrix} \begin{pmatrix} \xi \end{pmatrix} \quad (2.17)$$

The exact finite-time propagator for Eqs. (2.17) can be computed, and it is therefore possible to obtain any ensemble average or time-correlation function analytically. Of course, one is most interested in the expectation values which depend on the physical variables q and p . In particular, one can obtain the fluctuations $\langle q^2 \rangle$ and $\langle p^2 \rangle$ and correlation functions of the form $\langle q^2(t) q^2(0) \rangle$, which can be used to measure the coupling between the thermostat and the system. We will discuss these and other quantities in the following chapters, where we report the different applications

Figure 2.2: The colored-noise dynamics of a 1D harmonic oscillator (a) can be described by a linear stochastic differential equation (2.17). A system composed of several, coupled oscillators, where each degree of freedom is thermostatted with an independent Langevin equation with identical parameters (b) can be thought again as a linear, OU process, with a large, sparse drift matrix. This matrix can be transformed in block-diagonal form by writing the dynamics in normal-modes representation. Since the stochastic part is built out of uncorrelated Gaussian processes, one can prove that each individual normal mode (c) automatically behaves in the same way as if the generalized Langevin equation had been applied in normal-modes representation.



we have developed so far. We refer the reader to appendix D for a complete list.

The estimates computed for an oscillator of frequency ω can be used to evaluate and hence optimize the response of a normal mode of a similar frequency in the system being studied. By performing such optimization for several frequencies at the same time, one can achieve good performance over all the relevant phonon frequencies. Furthermore, thanks to the properties of Eq. (2.17), one does not need to perform a normal-modes analysis to turn this idea into a practical method. Consider a perfect harmonic crystal, and apply an independent instance of the GLE thermostat to the three Cartesian coordinates of each atom. It is easy to see that, since Eqs. (2.17) are linear and contain Gaussian noise, the thermostatted equations of motion are invariant under any orthogonal transformation of the coordinates. Therefore, the resulting dynamics can be described on the basis of the normal modes just as in ordinary Hamiltonian lattice dynamics (see figure 2.2). As a consequence, each phonon will respond independently as a 1-D oscillator with its own characteristic frequency.

The possibility of enforcing the desired, frequency-dependent behavior without the need to know the exact vibrational frequency or the phonons' displacement patterns is the basis for the flexibility of our approach, and suggests countless applications which we have only now just begun to investigate. The amount of information needed to tune a GLE thermostat on a complex system is minimal. In practice, one

only needs to know a rough estimate of the range of frequencies spanned by its vibrational spectrum. The analytical results in the one-dimensional case, can then be evaluated for the relevant values of ω , and the parameters can be optimized to obtain the desired behavior.

Fokker-Planck equation In passing, one can find the solution for the evolution of the probability density (1.28) in the case of a harmonic potential of frequency ω , project it on the $\mathbf{x} = (q, p)$ subspace only, and average the initial configuration in the \mathbf{s} subspace over the stationary probability. In a sense, this would give the solution for the time-dependent probability density under the equivalent, non-Markovian equation (2.1)⁶.

The derivation is cumbersome, and is performed in appendix C. The resulting probability evolves as a Gaussian

$$P(\mathbf{x}, t | \mathbf{x}_0, 0) \propto \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{U}_{xx}(t) \mathbf{x}_0)^T \mathbf{W}_{xx}^{-1}(t) (\mathbf{x} - \mathbf{U}_{xx}(t) \mathbf{x}_0) \right].$$

The time-dependent matrices \mathbf{U}_{xx} and \mathbf{W}_{xx} are given in Eq. (C.3).

2.3 Fitting the GLE parameters

The last point in our strategy requires the fitting of the parameters in the generalized Langevin thermostat to a set of desired properties. Just as for most optimization schemes, three main ingredients enter the procedure. First, the definition of a merit function which assesses quantitatively how close a set of parameters is to satisfying the requirements of the user. The details of such a function depend on what property one wants to optimize, and we will discuss them in chapters 3 and 5. Then, one needs a parametrization of the drift and diffusion matrix, which possibly regularizes the search space, and automatically enforces any existing constraints. Finally, an optimization algorithm is needed that finds the extremal values of the merit function, without getting trapped in local minima.

2.3.1 Designing a figure of merit for the fit

Since the number and type of properties to be optimized depends on the particular application, we will discuss here the general principles we have used to design a function χ whose minimum corresponds to the optimal parameters. We begin identifying those properties of a GLE dynamics that depend only on \mathbf{A}_p and \mathbf{B}_p , and those which also have a dependence on a frequency ω . Examples of the first kind of properties (which we shall label as $h(\mathbf{A}_p, \mathbf{B}_p)$) are the white-noise term a_{pp} or the free-particle diffusion coefficient (3.4), while the second category (which will be labeled $h(\mathbf{A}_p, \mathbf{B}_p, \omega)$) includes properties such as the Fourier transforms of the

⁶Since the status of the \mathbf{s} variables encode the “memory” of the (q, p) degrees of freedom (see figure 2.1), what we are really computing is the evolution of the probability density according to (2.1), given the boundary condition $P((q, p), 0 | (q_0, p_0), 0) = \delta((q, p) - (q_0, p_0))$ and averaged over all the possible previous histories compatible with the equilibrium probability distribution.

memory kernels, $K(\omega)$ and $H(\omega)$, and the properties of the dynamics in the harmonic limit. We build the overall merit function as a weighted sum of terms stemming from the different properties,

$$\chi = \sum_i w_i \chi_i. \quad (2.18)$$

Each term evaluates the distance of the actual values from the target. For frequency-independent properties, it will simply read

$$\chi_i = d(h_i(\mathbf{A}_p, \mathbf{B}_p), \tilde{h}_i), \quad (2.19)$$

while for frequency-dependent ones we take some representative frequencies ω_{ij} over the range relevant for the system being studied, and compute

$$\chi_i = \left[\sum_j w_{ij} d(h_i(\mathbf{A}_p, \mathbf{B}_p, \omega_{ij}), \tilde{h}_i(\omega_{ij}))^{m_i} \right]^{1/m_i}. \quad (2.20)$$

The form of (2.20) is chosen to mimic a p -norm. By tuning the exponent m_i , one can interpolate between a 1-norm behavior, where the discrepancies from individual frequencies are simply summed, to an ∞ -norm which singles out the term with the maximum discrepancy. The latter is useful when one would like to have $h_i(\mathbf{A}_p, \mathbf{B}_p, \omega)$ converge uniformly to the target $\tilde{h}_i(\omega)$.

Finally, depending on the property being considered, we have chosen the distance function to be simply $d(x_1, x_2) = |x_1 - x_2|$, or a logarithmic distance $d(x_1, x_2) = |\log(x_1/x_2)|$. In cases where an inequality target was desired (i.e. $h_i \geq \tilde{h}_i$) we have taken functions of the form $d_{>}(x_1, x_2) = \max(x_2 - x_1, 0)$ or $d_{<}(x_1, x_2) = \max(\log(x_1/x_2), 0)$.

2.3.2 Parametrization of GLE matrices

As already discussed, a number of constraints must be enforced on the drift and diffusion matrices in order to guarantee that the resulting stochastic differential equation is well-behaved. It is therefore important to find a representation of the matrices that ensures these conditions are automatically enforced during the fit, and that the parameters space is efficiently explored. A first condition, required to yield a memory kernel with exponential decay, is that all the eigenvalues of \mathbf{A}_p have a positive real part. A second requirement is that the kernel $K(\omega)$ is positive for all real ω . This ensures that the stochastic process will be consistent with the second law of thermodynamics[25].

Finding the general conditions for \mathbf{A}_p to satisfy this second constraint is not simple. However, we can state that a sufficient condition for $K(\omega) > 0$ is that $\mathbf{A}_p + \mathbf{A}_p^T$ is positive definite (see appendix B). For simplicity we shall assume such a positivity condition to hold, since we found empirically that this modest loss of generality does not significantly affect the accuracy or the flexibility of the fit. Moreover, in the case of canonical sampling, $\mathbf{A}_p + \mathbf{A}_p^T > 0$ is also required in order to obtain a real diffusion matrix, since $\mathbf{B}_p \mathbf{B}_p^T = k_B T (\mathbf{A}_p + \mathbf{A}_p^T)$ according to Eq. (2.5).

One would like to find a convenient parametrization, which automatically enforces these constraints. This is best done by writing \mathbf{A}_p as the sum of a symmetric and antisymmetric part, $\mathbf{A}_p^{(S)} + \mathbf{A}_p^{(A)}$. Since any orthogonal transform of the \mathbf{s} degrees of freedom will not change the dynamics (see appendix A), one can assume without loss of generality that the $\mathbf{A}^{(S)}$ block in $\mathbf{A}_p^{(S)}$ is diagonal (see Eq. (2.3) for the naming convention). Since in the general case the two components of \mathbf{A}_p do not commute, we will assume $\mathbf{A}_p^{(A)}$ to be full, and write $\mathbf{A}_p^{(S)}$ in the form

$$\mathbf{A}_p^{(S)} = \begin{pmatrix} a & a_1 & a_2 & \cdots & a_n \\ a_1 & \alpha_1 & 0 & \cdots & 0 \\ a_2 & 0 & \alpha_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_n & 0 & 0 & \cdots & \alpha_n \end{pmatrix}.$$

In order to enforce the positive-definiteness, one use an analytical Cholesky decomposition $\mathbf{A}_p^{(S)} = \mathbf{Q}_p \mathbf{Q}_p^T$, with

$$\mathbf{Q}_p = \begin{pmatrix} q & q_1 & q_2 & \cdots & q_n \\ 0 & d_1 & 0 & \cdots & 0 \\ 0 & 0 & d_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}, \quad (2.21)$$

$\alpha_i = d_i^2$, $a_i = d_i q_i$, and $a = q^2 + \sum_i q_i^2$. Such a parametrization guarantees that \mathbf{A}_p will generate a dynamics with a stationary probability distribution, and requires $2n + 1$ parameters for the symmetric part (the elements of \mathbf{Q}_p , Eq. (2.21)), and $n(n + 1)/2$ for the antisymmetric part $\mathbf{A}_p^{(A)}$. If we want the equilibrium distribution to be canonical we must enforce the fluctuation-dissipation theorem. The matrix $\mathbf{B}_p \mathbf{B}_p^T$ is then uniquely determined by (2.7).

If instead we aim at a generalized formulation, which allows for frequency-dependent thermalization, there are no constraints on the choice of \mathbf{B}_p other than the fact that both $\mathbf{B}_p \mathbf{B}_p^T$ and the covariance \mathbf{C}_p must be positive-definite. Clearly, a real, lower-triangular \mathbf{B}_p is the most general parametrization of a positive-definite $\mathbf{B}_p \mathbf{B}_p^T$, and amounts at introducing $(n + 1)(n + 2)/2$ extra parameters. Together with the assumption that $\mathbf{A}_p^{(S)} > 0$, the condition $\mathbf{B}_p \mathbf{B}_p^T > 0$ is sufficient to ensure that the unique symmetric \mathbf{C}_p which satisfies (2.5) is also positive-definite.

2.3.3 Minimization algorithm

We have developed a robust parametrization, which however does not solve all the minimization problems associated with the highly nonlinear search space. The major problems we encounter are the presence of local minima and, more specifically, the existence of “degenerate” states when relatively large matrices are used ($n \gtrsim 4$). In these cases, the optimization does not exploit the full parameter space, which

means that increasing n does not necessarily improve the accuracy. This issue can be reduced by adding penalty terms which depend on the eigenvalue spectrum of \mathbf{A} , and enforce an even distribution of the eigenvalues over the range of interest⁷.

Since computing analytical estimates of GLE properties is inexpensive, we used very simple, robust algorithms; namely, a combination of simulated annealing[28], which we employed for an initial phase-space search, and downhill simplex[29] minimization to refine the fit. As is often the case in complex non-linear optimizations, having a good initial guess helps tremendously in improving the final outcome.

We used several heuristic strategies to obtain such initial configurations. A very successful prescription is to gather several small matrices with known properties, either by using an analytical form such as (2.13) or by performing a preliminary fit, and then combining them using the rules of section 2.4, to generate the initial parameters for further optimization at higher n .

2.4 Transformation rules for GLE dynamics

Given that fitting generalized Langevin thermostat parameters is not a straightforward task, it is useful to have some simple recipes to transform and combine existing matrices, so as to adapt them to systems different from the one they had initially been fitted for. A collection of optimized parameters is can also be obtained from an online repository.⁸

2.4.1 Matrix scaling

The colored-noise dynamics (2.2) has some invariance properties, which make it possible to predict how the dynamics change when the \mathbf{A}_p and \mathbf{C}_p matrix are multiplied by a positive scalar. In fact, one can see that if the drift and covariance matrices ($\mathbf{A}_p, \mathbf{C}_p$) correspond to the memory kernels $K(\omega)$ and $H(\omega)$, the scaled matrices ($\alpha\mathbf{A}_p, \beta\mathbf{C}_p$) yield the kernels $\alpha K(\alpha^{-1}\omega)$ and $\alpha\beta H(\alpha^{-1}\omega)$. Furthermore, as we will discuss later, most of the properties of the dynamics are then transformed by a simple scaling (as a rule of thumb, quantities with the units of a frequency transform as $h(\omega) \leftarrow \alpha h(\alpha^{-1}\omega)$, and quantities with the units of an energy as $h(\omega) \leftarrow \beta h(\alpha^{-1}\omega)$).

2.4.2 Combination rules

We now discuss a set of combination rules which can be used to blend two generalized Langevin equations into a new one, whose memory kernel is a weighted sum of those of the original SDEs. At variance with matrix rescaling, which generates a dynamics whose statistical properties can be obtained from the original ones, there is no simple relationship between the static and dynamic properties of the mixture GLE and those of its components.

⁷One can compute for instance the average and the spread of the spectrum, and pin them to the desired values by a term of the form (2.19), whose weight in Eq. (2.18) is then reduced as the optimization procedure approaches convergence.

⁸See <http://gle4md.berlios.de>, where sample routines in different programming languages are also available.

Consider the Markovian stochastic equations

$$\begin{aligned} \begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}}_1 \end{pmatrix} &= - \begin{pmatrix} {}^1a_{pp} & {}^1\mathbf{a}_p^T \\ {}^1\bar{\mathbf{a}}_p & {}^1\mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s}_1 \end{pmatrix} + \begin{pmatrix} {}^1b_{pp} & {}^1\mathbf{b}_p^T \\ {}^1\bar{\mathbf{b}}_p & {}^1\mathbf{B} \end{pmatrix} \begin{pmatrix} \xi_1 \end{pmatrix} \\ \begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}}_2 \end{pmatrix} &= - \begin{pmatrix} {}^2a_{pp} & {}^2\mathbf{a}_p^T \\ {}^2\bar{\mathbf{a}}_p & {}^2\mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} {}^2b_{pp} & {}^2\mathbf{b}_p^T \\ {}^2\bar{\mathbf{b}}_p & {}^2\mathbf{B} \end{pmatrix} \begin{pmatrix} \xi_2 \end{pmatrix}, \end{aligned} \quad (2.22)$$

which have the corresponding non-Markovian form

$$\begin{aligned} \dot{p} &= - \int_{-\infty}^t K_1(t-s)p(s)ds + \zeta_1(t) \\ \dot{p} &= - \int_{-\infty}^t K_2(t-s)p(s)ds + \zeta_2(t). \end{aligned} \quad (2.23)$$

For simplicity we have written down only the free-particle limit of Eqs. (2.2) and (2.1). The position-momentum coupling can be introduced without changing the reasoning nor the final result. It is to be understood that ${}^1\mathbf{A}_p$ and ${}^1\mathbf{B}_p$ generate the memory kernels $K_1(t)$ and $H_1(t)$ and so on.

It is simple to prove, by direct substitution into Eqs. (2.4), that the non-Markovian SDE with memory kernels $K(t) = w_1 K_1(t) + w_2 K_2(t)$ and $H(t) = w_1 H_1(t) + w_2 H_2(t)$ can be obtained from a Markovian process whose matrices \mathbf{A}_p and $\mathbf{D}_p = \mathbf{B}_p \mathbf{B}_p^T$ are built from the following combination of the parameters of Eqs. (2.22):

$$\begin{aligned} \mathbf{A}_p &= \begin{pmatrix} {}^1a_{pp}w_1 + {}^2a_{pp}w_2 & {}^1\mathbf{a}_p^T\sqrt{w_1} & {}^2\mathbf{a}_p^T\sqrt{w_2} \\ {}^1\bar{\mathbf{a}}_p\sqrt{w_1} & {}^1\mathbf{A} & \mathbf{0} \\ {}^2\bar{\mathbf{a}}_p\sqrt{w_2} & \mathbf{0} & {}^2\mathbf{A} \end{pmatrix} \\ \mathbf{D}_p &= \begin{pmatrix} {}^1d_{pp}w_1 + {}^2d_{pp}w_2 & {}^1\mathbf{d}_p^T\sqrt{w_1} & {}^2\mathbf{d}_p^T\sqrt{w_2} \\ {}^1\mathbf{d}_p\sqrt{w_1} & {}^1\mathbf{D} & \mathbf{0} \\ {}^2\mathbf{d}_p\sqrt{w_2} & \mathbf{0} & {}^2\mathbf{D} \end{pmatrix} \end{aligned} \quad (2.24)$$

We note in passing that by combining this result with the possibility of having δ -like kernels (2.15), one obtains a recipe to construct kernels of arbitrary shape. Let us stress that this idea does not substitute the fitting procedure described in section 2.3, since there is not a straightforward connection between the memory kernels and all the properties of the dynamical trajectory. Moreover, an explicit fitting procedure allows to obtain a better accuracy with the same number of additional degrees of freedom, by exploiting the full dimensionality of the search space. However, as we have discussed, mixing a few δ -like kernels can provide a suitable starting configuration for the nonlinear optimization.

2.5 Implementation details

Having discussed how to derive, characterize and fit a set of parameters for a general Langevin equation thermostat, we must face the problem of implementing an integrator for Eqs. (2.2) in a molecular-dynamics code. Thankfully, this is a straight-

forward task. Here we consider the case of a velocity-Verlet integrator, which updates positions and momenta by a time step Δt , according to the scheme:

$$\begin{aligned} p &\leftarrow p - V'(q)\Delta t/2 \\ q &\leftarrow q + p\Delta t \\ p &\leftarrow p - V'(q)\Delta t/2. \end{aligned} \quad (2.25)$$

Eqs. (2.25) can be obtained using Trotter splitting in a Liouville operator formalism[30]. In the same spirit one can introduce our GLE thermostat by performing two free-particle steps by $\Delta t/2$ on the (p, \mathbf{s}) variables[31] (with reference to Eq. (2.8), the Verlet step corresponds to an approximation of the propagator $e^{\Delta t \mathcal{L}_{HAM}}$, and the (p, \mathbf{s}) to $e^{\Delta t \mathcal{L}_{GLE}/2}$):

$$\begin{aligned} (p, \mathbf{s}) &\leftarrow \mathcal{P}[(p, \mathbf{s}), \Delta t/2] \\ p &\leftarrow p - V'(q)\Delta t/2 \\ q &\leftarrow q + p\Delta t \\ p &\leftarrow p - V'(q)\Delta t/2. \\ (p, \mathbf{s}) &\leftarrow \mathcal{P}[(p, \mathbf{s}), \Delta t/2] \end{aligned} \quad (2.26)$$

At variance with thermostats based on second-order equations of motion such as Nosé-Hoover, where a multiple time-step approach is required to obtain accurate trajectories[32, 33], this free-particle step can be performed without introducing additional sampling errors. The exact finite-time propagator for (p, \mathbf{s}) reads⁹:

$$\mathcal{P}[(p, \mathbf{s}), \Delta t]^T = \mathbf{T}(\Delta t)(p, \mathbf{s})^T + \mathbf{S}(\Delta t)\boldsymbol{\xi}^T \quad (2.27)$$

where $\boldsymbol{\xi}$ is a vector of $n + 1$ uncorrelated Gaussian numbers, and the matrices \mathbf{T} and \mathbf{S} can be computed once, at the beginning of the simulation and for all degrees of freedom[16, 34, 35]. The relations between \mathbf{T} , \mathbf{S} , \mathbf{A}_p , \mathbf{C}_p and Δt read¹⁰

$$\mathbf{T} = e^{-\Delta t \mathbf{A}_p}, \quad \mathbf{S}\mathbf{S}^T = \mathbf{C}_p - e^{-\Delta t \mathbf{A}_p} \mathbf{C}_p e^{-\Delta t \mathbf{A}_p^T}.$$

It is worth pointing out that when fluctuation-dissipation theorem holds, the canonical distribution is invariant under the action of (2.27), whatever the size of the time-step. A useful consequence of this property is that, in the rare cases where applying (2.27) introduces a significant overhead over the force calculation, the thermostat can be applied every m steps of dynamics, using a stride of $m \Delta t$. This will change the trajectory, but does not affect the accuracy of sampling.

The velocity-Verlet algorithm (2.25) introduces finite- Δt errors, whose effect

⁹Compare with Eq. (1.32). The finite-time propagator generates the new position in such a way to fulfill the Fokker-Planck equation associated with the free-particle OU process, subject to the boundary condition of being initially at point (p, \mathbf{s}) .

¹⁰So far we have simplified expressions by choosing a unit mass. In a real system, one has two choices: either using mass-scaled momentum ($\tilde{p}_i = \sqrt{m_i} \dot{q}_i$) and the expressions derived for unit mass, or applying Eq. (2.27) with \mathbf{S} multiplied by $\sqrt{m_i}$ (if the integrator stores momenta) or by $1/\sqrt{m_i}$ (if the integrator stores velocities).

needs to be monitored. In microcanonical simulations, this is routinely done by checking conservation of the total energy H . Following the work of Bussi *et al.*[13] we introduce a conserved quantity \tilde{H} , which can be used to the same purpose:

$$\tilde{H} = H - \sum_i \Delta K_i \quad (2.28)$$

where ΔK_i is the change in kinetic energy due to the action of the thermostat at the i -th time-step, and the sum is extended over the past trajectory. In cases where the fluctuation-dissipation theorem holds, the drift of the effective energy quantitatively measures the violation of detailed balance induced by the velocity-Verlet step, similarly to Refs. [13, 31]. In the cases where the FDT does not hold, the conservation of this quantity measures only the accuracy of the integration, similarly to Refs. [36, 37].

One of the major advantages of our approach is that, once Eqs. (2.26) are implemented in a MD code, all the different applications that will be described in the following chapters can be realized within the very same framework, by simply changing the parameters of the GLE.

Canonical sampling

After having described in detail how to define and tune a generalized Langevin thermostat in the abstract, we now turn our attention to a first set of applications; namely, various methods for enhancing and modifying the sampling of the constant-temperature, canonical ensemble. The use of molecular dynamics as a means for computing equilibrium averages is well established. This approach is based on the assumption of ergodicity[10, 38], i.e. that one can compute an ensemble average by performing a dynamical trajectory which visits points with a probability distribution consistent with the ensemble,

$$\frac{1}{\mathcal{Z}} \int d\mathbf{p} d\mathbf{q} \mathcal{A}(\mathbf{q}, \mathbf{p}) e^{-\mathcal{H}(\mathbf{q}, \mathbf{p})/k_B T} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \mathcal{A}(\mathbf{q}(t), \mathbf{p}(t)) dt.$$

Serious problems arise when different regions of phase space are separated by high energetic barriers, which are seldom crossed during dynamical trajectories. These activated, rare events are one of the major problems in computer simulation of many important systems, and several techniques have been and are being developed to deal with them¹.

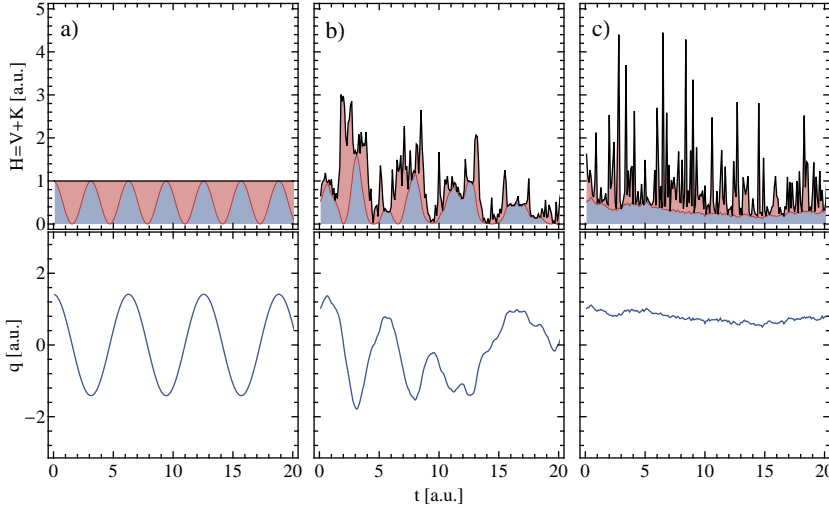
Even when these activated processes are not a concern, one would like to speed up the convergence of the averaging procedure, which basically amounts to generating uncorrelated configurations in as few steps as possible. To this aim, many different schemes, both stochastic and deterministic, have been developed which modify Hamilton's equations to allow for accurate and efficient constant-temperature sampling. During this process, the real dynamical properties of the system are lost, to an extent which depends to the details of the system and of the thermostating scheme (see figure 3.1 for a simple example).

So-called *global* thermostating schemes[13, 39] act only on the total kinetic energy K of the system, which is changed so as to sample the appropriate distribution

$$P(K) dK \propto K^{(N_f/2-1)} e^{-K/k_B T} dK.$$

¹We will briefly sketch one of these methods in section 3.5.

Figure 3.1: Total energy (divided in its potential (blue) and kinetic (red) energy contributions) and position q of a one-dimensional harmonic oscillator of unitary ω and mass, during microcanonical dynamics (a) and subject to a white-noise dynamics with friction $\gamma = 1$ (b) and $\gamma = 10^3$ (c).



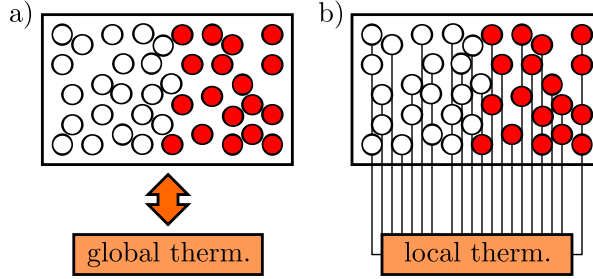
The effect on the trajectory is small, since the perturbation on individual components of the momentum becomes negligible as the number of degrees of freedom N_f is increased. Hence, in general, global thermostats will affect the dynamical properties only slightly. However, there is a drawback with these approaches which can be understood by considering the following: if half of the particles of the system are initialized with zero temperature, and half at twice the target temperature, the system would be in a massively non-equilibrium state, yet the overall kinetic temperature would be correct (figure 3.2). Consequentially, a global thermostat would not help the system to reach equilibrium, and instead one would be forced to rely on the spontaneous heat flow which occurs because of anharmonic coupling between the components of the system.

A *local* thermostat, such as white-noise Langevin or the massive version of Nosé-Hoover thermostat[1, 40], couples to the individual atomic components of the momentum. As a result, it is able to detect and actively counterbalance uneven distributions of energy between the different parts of the system. This comes however at the expense of a greater influence on the real-time dynamics.

The extreme example given in Figure 3.2 serves as a guide for understanding the properties of the two classes of thermostats in more realistic situations. Global schemes will perform well in anharmonic and homogeneous systems, where the coupling between different particles or normal modes is strong. Meanwhile, local schemes are better suited to situations where one wants to compute a local property, and cannot rely on internal dissipation to ensure efficient loss of memory. This would be the case for harmonic or inhomogeneous systems[41].

Although it would be interesting to explore the effect of correlated noise on the stochastic velocity rescaling[13] global thermostat, the generalized Langevin

Figure 3.2: Comparison of the behavior of a global (a) and local (b) thermostats in the case where part of the system is much hotter than the target temperature, while the remainder is much cooler. The global thermostat only “feels” the total kinetic energy, which has a reasonable value, while in the local scheme each degree of freedom is taken care of independently, and hence the thermostat would actively counter the temperature imbalance. Equilibrium will be reached eventually in the global case also, but only because of internal coupling between components.



thermostat we present in this thesis is a local scheme. In this chapter we will discuss how one can fine-tune the effect of the thermostat as a function of the characteristic time of molecular motion, and thereby enhance or reduce the efficiency of sampling in the constant-temperature ensemble, while maintaining a degree of control over the degree of disturbance introduced to the dynamics.

3.1 Optimal sampling

As a first example application for the colored-noise framework we have introduced in chapter 2, we will discuss how it is possible to construct a thermostat which aims to sample the canonical ensemble as efficiently as possible. A precise measure of efficiency would necessitate an in-depth knowledge of the details of the system and the observables one wishes to compute. Instead, we endeavor to provide a “fool-proof” scheme, which ensures a close-to-optimal efficiency for a broad spectrum of cases, while allowing enough flexibility to be further tuned to more specific requirements.

3.1.1 White noise on the harmonic oscillator

Let us consider again the simple example in figure 3.1, where white-noise Langevin dynamics is performed on a one-dimensional harmonic oscillator. Clearly, in the microcanonical, zero-friction limit the position oscillates back and forth periodically and total energy remains constant. In this case the canonical ensemble is not sampled properly, and a high degree of correlation is present between the configurations adopted at different times.

Introducing a friction and noisy force terms in the form of a white-noise Langevin equation (1.15) causes fluctuations in the total energy, and guarantees that the canonical ensemble will be sampled. However, the efficiency of sampling is not a monotonic function of the intensity of the noise. Compare the panels in figure 3.1. In panel (b)

a friction $\gamma = \omega$ is used: total energy fluctuates wildly, and both the potential and kinetic energy components vary significantly over the time scale of few vibrational periods. In panel (c), where a much higher friction $\gamma = 10^3 \omega$ was chosen, we observe that the kinetic energy fluctuates much faster than in panel (b). However, exploration of configuration space is hampered significantly.

Sampling momentum space is easy, as the equilibrium probability distribution is just $e^{-p^2/2mk_B T}$, and hence one can pick a random, independent momentum at any time. By contrast, configuration sampling is considerably more difficult, and is thus our major concern ².

To turn these observations into a quantitative assessment of the efficiency of sampling, we define the correlation times of the potential, kinetic and total energies, $V(t)$, $K(t)$ and $H(t) = V(t) + K(t)$:

$$\begin{aligned}\tau_V &= \frac{1}{\langle V^2 \rangle - \langle V \rangle^2} \int_0^\infty \langle (V(t) - \langle V \rangle) (V(0) - \langle V \rangle) \rangle dt, \\ \tau_K &= \frac{1}{\langle K^2 \rangle - \langle K \rangle^2} \int_0^\infty \langle (K(t) - \langle K \rangle) (K(0) - \langle K \rangle) \rangle dt, \\ \tau_H &= \frac{1}{\langle H^2 \rangle - \langle H \rangle^2} \int_0^\infty \langle (H(t) - \langle H \rangle) (H(0) - \langle H \rangle) \rangle dt.\end{aligned}\tag{3.1}$$

Our strategy is then to compute these correlation times for a harmonic oscillator of frequency ω , and subsequently minimize their value as a function of the friction coefficient γ .

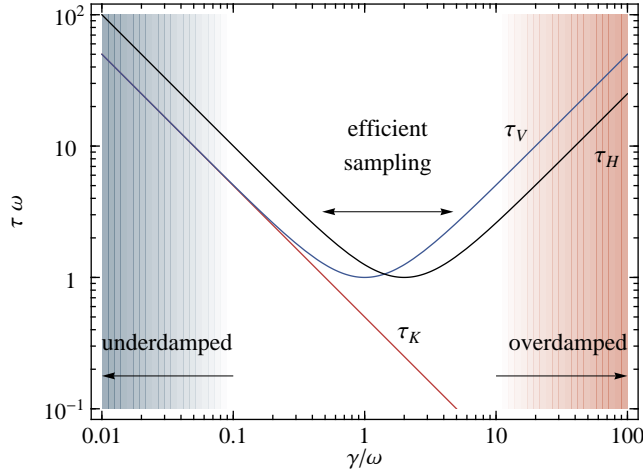
In the case of white-noise, it is possible to obtain simple, closed expressions, which read

$$\begin{aligned}\tau_V(\omega) &= \frac{1}{2\gamma} + \frac{\gamma}{2\omega^2}, \\ \tau_K(\omega) &= \frac{1}{2\gamma}, \\ \tau_H(\omega) &= \frac{1}{\gamma} + \frac{\gamma}{4\omega^2}.\end{aligned}\tag{3.2}$$

While τ_K can be decreased without bound, an optimum value of the friction exists for both potential and total energy which minimizes the corresponding correlation time (see figure 3.3). These optimal values are $\gamma_V = \omega$ and $\gamma_H = 2\omega$ respectively. In a real system, many different frequencies are present. Therefore, by selecting a friction one chooses a particular frequency $\omega_V = \gamma$ ($\omega_H = \gamma/2$) which will be optimally coupled with the thermostat. The remaining normal modes will be sampled sub-optimally, and we can define an efficiency parameter $\kappa_{V(H)} = [\tau_{V(H)}(\omega)\omega]^{-1}$, which is equal to one for $\omega_{V(H)}$, and decreases for smaller and higher ω 's.

²A complete re-sampling of momenta at given intervals is the main idea behind the Andersen thermostat[42]. In this simple harmonic model, because the exact Hessian is known, the position can be randomized just as easily.

Figure 3.3: Correlation times for the kinetic, potential and total energies (τ_K , τ_V and τ_H respectively, see Eqs. (3.2)) of a harmonic oscillator of frequency ω , subject to a white-noise Langevin equation with friction γ . As in figure 3.1, we can distinguish a low-friction, underdamped regime (panel 3.1(a)) and a high-friction limit where the oscillator is overdamped and diffusion within configuration space is hampered (panel 3.1(c)). Between these two extremes, there is an intermediate range of values, for which the noisy term disturbs the periodic dynamics without completely masking the inertial oscillations, which help the sampling.



3.1.2 A uniform-efficiency thermostat

The results in the previous section serve as a useful introduction to our strategy. They also support the common wisdom, which is to choose a value of the friction which is commensurate to the reciprocal of the characteristic time scale of the process being monitored. Armed with the generalized-Langevin equation machinery we have introduced in chapter 2 we can however attempt something more sophisticated; namely, we can enhance the value of $\kappa_{V(H)}(\omega)$ over a broad frequency range that encompasses all the relevant vibrational modes of the system being studied[43, 44]. This is possible because the generalized Langevin equation is linear, as discussed in section 2.2. Consequentially, if an independent thermostat is applied to every Cartesian component of a harmonic crystal, the resulting dynamics will behave as if the thermostats had been applied on its normal modes. Deterministic thermostats such as Nosé-Hoover[39, 45, 46] do not possess this rotational invariance, because they are based on second-order equations of motion. In appendix E we present a simple example in which this lack of invariance leads to significant deterioration of the sampling properties.

There are several reasons why this kind of a “flat efficiency” thermostat is desirable. Firstly, a particular observable could depend, in a complex manner, on a number of different vibrational modes. This may well make a white-noise thermostat, tuned on a specific frequency, a non-optimal choice. Secondly, in order to optimize a conventional thermostat in an empirical way one has to perform very long

preliminary runs and compute the relevant correlation times as a function of the parameters; we refer the reader to chapter 4, where we compare different sampling schemes in the challenging case of path-integral molecular dynamics, for a more thorough discussion. Finally, in a number of situations (ab-initio dynamics to name just one) one cannot afford to perform simulations just to optimize the sampling parameters. It would be desirable therefore to have a “black-box” strategy in which the sampling is guaranteed to be close to optimal regardless of the physical property of interest.

The fitting procedure described in section 2.3 is readily adapted to this case. Since we want to perform canonical sampling, the FDT has to hold, so that $\mathbf{C}_p = k_B T$. This in turns fixes \mathbf{B}_p by means of Eq. (2.5). Consequentially, the entries of \mathbf{A}_p are the only independent parameters. One then chooses a set of m frequencies $\omega_i = \omega_{\min}(\omega_{\max}/\omega_{\min})^{i/(m-1)}$, that are equally spaced on a logarithmic scale over a broad range $(\omega_{\min}, \omega_{\max})$. Finally, the function to be optimized is built using the model given in Eq. (2.20) and a logarithmic distance between the actual values $\kappa_{V(H)}(\omega_i)$ and the optimal value of 1:

$$\chi = \left[\sum_i |\log \kappa_{V(H)}(\omega_i)|^m \right]^{1/m}. \quad (3.3)$$

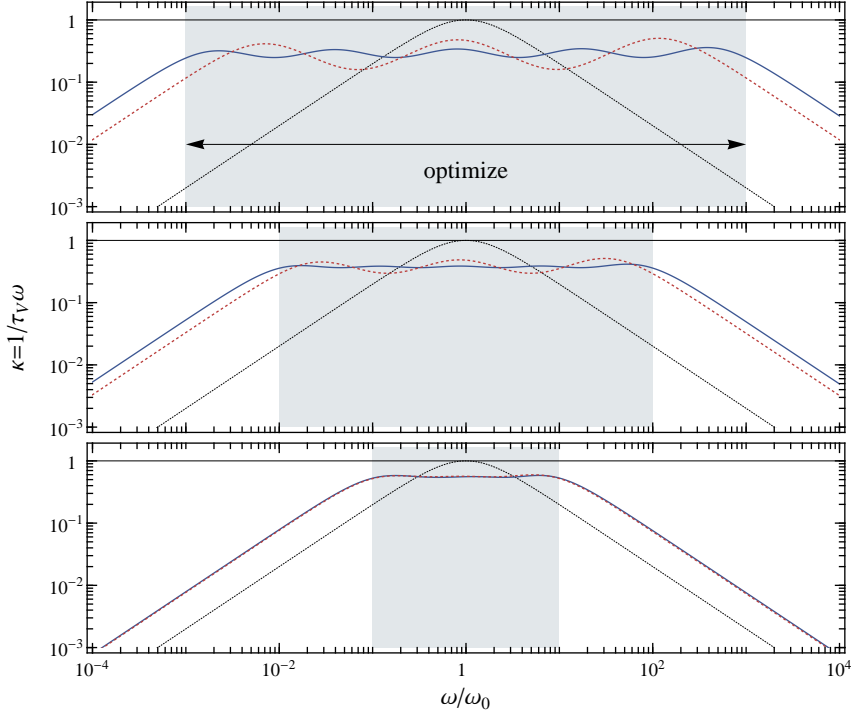
A large value for the exponent m enforces a flat efficiency curve over the whole frequency range. However as the exponent is increased the optimization procedure gets stiffer. Therefore, starting off with a large value will almost certainly lead to unsatisfactory local minima. To remedy this one can start the optimization with a low value for m , and gradually increase it to higher values as the optimization progresses.

In Figure 3.4 we compare the optimized $\kappa(\omega)$ for different frequency ranges and numbers of additional degrees of freedom. Although we could not obtain a rigorous proof, we find empirically that $\kappa(\omega) = 1$ is the best result which can be attained, and that a nearly-optimal efficiency can be reached over a very broad range of frequencies. However, this constant efficiency decreases slightly as the fitted range is extended. We were not able to contrast this effect, neither by increasing the number n of s_i employed. We found however that by using a larger n it is possible to make the response flatter.

3.1.3 Diffusion and the free-particle limit

The scheme we have described so far works optimally in harmonic or quasi-harmonic systems. Anharmonicity will introduce deviations from the predicted behavior so, in the extreme case of diffusive systems such as liquids, one has to ask how much will diffusion be affected by the thermostat? Not least because, for an overdamped Langevin equation, the diffusive modes are slowed down considerably (see e.g. Ref. [41]). To estimate the impact of the thermostat on the diffusion, we define the free-particle diffusion coefficient D^* as the diffusion constant in the absence of physical forces,

Figure 3.4: Sampling efficiency as estimated from Eq. (3.1) for a harmonic oscillator, plotted as a function of the frequency ω . The $\kappa(\omega)$ curve for a white-noise Langevin thermostat optimized for $\omega = \omega_0$ (black, dotted lines, Eq. (3.2)) is contrasted with those for a set of optimized GLE thermostats. The panels, from bottom to top, contain the results fitted over frequency ranges spanning two, four and six orders of magnitude around $\omega = 1$ respectively. Blue, continuous lines correspond to matrices with $n = 4$, while the red, dashed lines are for $n = 2$.



which, for a GLE thermostat, is given by:

$$\begin{aligned} \frac{mD^*}{k_B T} &= \frac{1}{\langle p^2 \rangle} \int_0^\infty \langle p(t)p(0) \rangle dt = \\ &= \left[\mathbf{A}_p^{-1} \right]_{pp} = \left(a_{pp} - \mathbf{a}_p^T \mathbf{A}^{-1} \bar{\mathbf{a}}_p \right)^{-1}, \end{aligned} \quad (3.4)$$

if the FDT holds³. In practical cases, if an estimate of the unthermostated (intrinsic) diffusion coefficient D is available, one should choose the matrix \mathbf{A}_p such that $D^* \gg D$, as then the thermostat will not become an additional bottleneck for diffusion. Equation (3.4) has the interesting consequence that D^* can be enhanced either by reducing the overall strength of the noise, as in white-noise LE, or by carefully balancing the terms in the denominator.

³In the general case, Eq. (3.4) would read

$$\frac{mD^*}{k_B T} = \left[\mathbf{A}_p^{-1} \mathbf{C}_p \right]_{pp} / [\mathbf{C}_p]_{pp}.$$

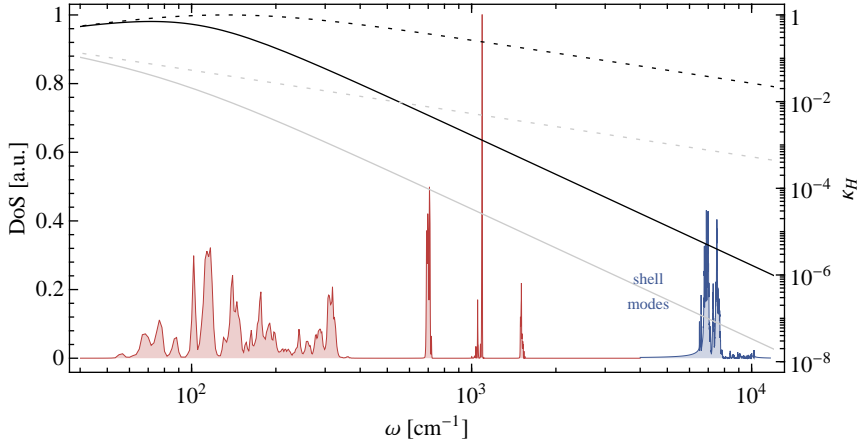
We have found empirically that for an \mathbf{A}_p matrix fitted to a set of harmonic modes in the frequency range $(\omega_{min}, \omega_{max})$, the diffusion coefficient computed using (3.4) is about $D^* \approx k_B T / m \omega_{min}$. This expression therefore gives a useful recipe for choosing the minimum frequency to be considered in the fit, when the diffusion coefficient is known experimentally or can be estimated. We will return to the problem of constant-temperature simulation of diffusive systems in section 3.4, where we will apply these ideas to a practical example. Now instead we turn our attention to achieving *poor* sampling for selected frequency ranges, which is precisely what is required for simulation methods based on adiabatic separation.

3.2 A thermostat for Car-Parrinello dynamics

Car-Parrinello (CP)-like, extended Lagrangian schemes [47, 48] would greatly benefit from the development of an improved, tunable thermostat. The idea behind this approach is very general, and applies to any system where the forces are the result of an expensive optimization procedure. Essentially, the minimization is circumvented by extending the dynamical degrees of freedom so as to include the parameters which would normally be optimized. An artificial dynamics is then introduced, which allows these extra variables to be maintained close to the ground state, by adiabatic decoupling from the remaining degrees of freedom. In the prototypical example of CP molecular dynamics (CPMD) a fictitious mass is assigned to the electronic degrees of freedom so that they can be evolved at the same time as the ions. If the fictitious mass is small, the dynamics of the electrons are adiabatically separated from those of the ions, so the electrons are kept close to the ground state while the nuclei are evolved at the correct temperature. An extended Lagrangian scheme can also be used in classical simulations that use polarizable force fields [49, 50]. In these cases, the adiabatically-separated degrees of freedom describe the charge polarization. Similar approaches have also been suggested in the field of rare-events sampling, to separate the oscillations of the microscopic degrees of freedom from those of a few selected slow reaction coordinates [51–53].

Controlling the temperature in these CP-like techniques requires separate treatment of the ionic degrees of freedom, which must sample the correct canonical ensemble, and the variational parameters, which must always remain at low temperature so as to minimize the error in the forces [54]. One might think that only applying the thermostat to the ionic degrees of freedom would be sufficient, since the ions' dynamics are adiabatically separated from those of the electrons. However, if one applies a white-noise Langevin thermostat, the motion of the ions will be disturbed, and high-frequency components will be introduced which will disrupt adiabatic decoupling causing a fast increase in the temperature of the electrons (see figure 3.8). For this reason, so far deterministic thermostats of the Nosé-Hoover (NH) type [39, 45, 46] have been adopted, generally in the form of Nosé-Hoover chains which circumvents the well-known ergodicity problems of the original version of the algorithm [46]. However, this scheme introduces a large number of parameters, whose effect on the ions' dynamics is not easy to predict and control, and suffers problems connected with the lack of rotational invariance in the Nosé-Hoover equa-

Figure 3.5: The autocorrelation time of the total energy for harmonic oscillators of frequency ω (Eq (3.2)) is plotted for different values of the thermostat parameters (2.10). Namely, dark curves correspond to high friction ($\gamma^{-1} = 20$ fs) whereas light ones correspond to a more gentle thermostat ($\gamma^{-1} = 1$ ps). Dotted lines correspond to white noise ($\tau \rightarrow 0$) and full ones to colored noise with $\tau = 2$ fs. The curves are superimposed on the vibrational spectrum, obtained from the Fourier transform of the velocity-velocity autocorrelation function, for a simulation of calcium carbonate using a polarizable force-field. The shell vibrational modes we report for reference are obtained from a run where we artificially heated the shells to 300 K.



tions of motion (see appendix E). The GLE thermostat approach, however, allows one to design a Langevin equation where the noisy force does not contain the high-frequency components that induce the high-frequency motion of the ions that are so damaging to the adiabatic decoupling. We will show that this is a viable strategy for obtaining a stochastic thermostat suitable for CPMD dynamics.

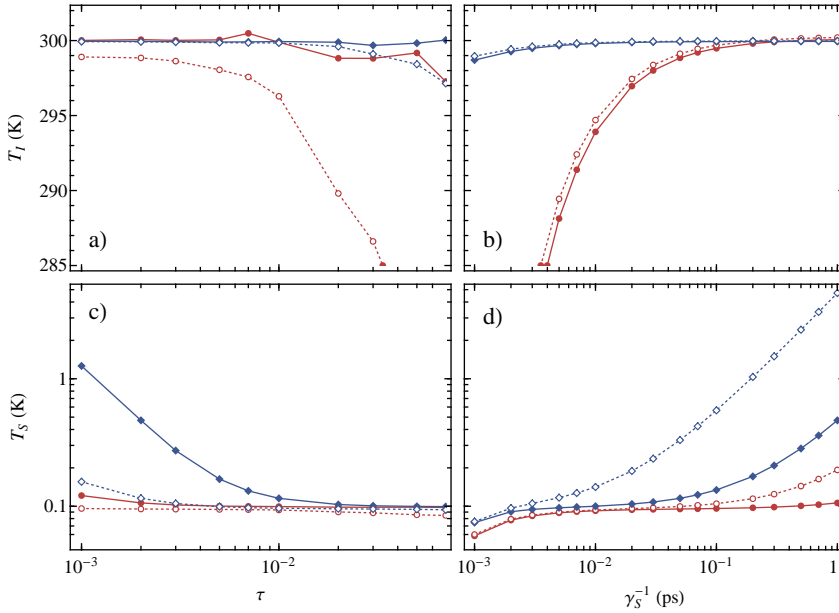
3.2.1 A low-pass filtered noise

As a first example, and to test our ideas, we consider the simple exponentially-correlated noise generated by the parameters given by Eq. (2.10). The memory kernel and its Fourier transform read

$$K(t) = \frac{\gamma}{\tau} e^{-|t|/\tau}, \quad K(\omega) = \frac{\gamma}{\pi} \frac{1}{1 + \tau^2 \omega^2},$$

respectively. This corresponds to a low-pass filter which damps the components of the power spectrum of the noise for ω greater than a cutoff frequency $\omega_F = 1/\tau$. By tuning τ one can reduce the frequencies which would otherwise couple with the fast electronic modes (see figure 3.5). Initially, to perform extensive tests, we did not apply this thermostat to ab-initio CPMD, but instead focused on classical molecular dynamics, using a polarizable force field. In this approach the electronic DOF are represented by charged shells, bound by a harmonic potential to their corresponding atomic core. In particular, we consider the simulation of crystalline calcite[36], using a force field in which the Ca^{2+} cores are treated as non-polarizable ions, while the

Figure 3.6: Shells (T_S) and ions (T_I) temperature for calcite, as a function of the thermostat parameters. Simulations have been performed for several values of thermostat parameters, which are joined by continuous lines, for the sake of clarity. In all panels we distinguish the strength of the ion thermostat by the line color. Blue and red curves correspond to strong ($\gamma^{-1} = 20$ fs) or mild ($\gamma^{-1} = 1$ ps) frictions respectively. In panel (a) and (c) we vary τ , and choose two extreme values of the shell friction, $\gamma_S^{-1} = 1$ ps and $\gamma_S^{-1} = 50$ fs, which are represented with full and dashed lines respectively. In panel (b) and (d), we vary γ_S instead. Here full and dashed lines correspond to a physically meaningful filter ($\tau = 2$ fs) and to white noise ($\tau = 0$), while the red and blue lines correspond to the same values of γ used in panels (a) and (c).



polarizability of the CO_3^{2-} anions is described by a charged shell attached to each oxygen. Thermostats are applied to the non-polarizable ions and, in the case of the oxygens, to the center of mass of the system formed by the ion plus its shell.

A small drift of the electronic temperature is unavoidable even in a microcanonical simulation due to imperfect decoupling. Since we aim to perform long simulations, it is necessary to apply a zero temperature, memory-less thermostat of friction γ_S to the electrons so as to balance this leakage of kinetic energy from the ionic degrees of freedom. This simulation scheme amounts to a non-equilibrium dynamics, in which heat is injected into the ionic DOF and systematically subtracted from the electronic ones. It would therefore be desirable to use a very small γ_S , since a dynamics too far from equilibrium would raise serious concerns about the accuracy of sampling.

We simulated a box containing 96 CaCO_3 units, with a timestep of 1 fs. NVT runs with target temperature $T = 300$ K were performed, and we systematically tested what effects varying τ , γ and γ_S has (Fig. 3.6). Averages are computed from 1 ns-long runs, where we discard the first 100 ps for equilibration.

For a large range of parameters, the procedure performs as expected; namely, the temperature of the shells remains below a few K, and the ions equilibrate to the desired temperature. When τ is larger than zero the heat transferred to the electronic DOF is reduced. Some care must be taken in choosing the friction γ_S , as the shell thermostat induces a small drag on the ions. If, at the same time, a weak friction γ is used for the ions their temperature sets to a lower value than the target. Moreover, since τ_H decays rather slowly to zero for $\omega > \tau^{-1}$, to prevent the thermostat from heating up the shells one must choose a cutoff frequency much lower than the highest ionic frequency. As a consequence, the relaxation time for high-frequency phonons is increased, which makes the effects of shell-induced drag more pronounced. However, the thermostat can be systematically improved by using more s_i 's, so as to obtain a filter which is more sharply defined.

3.2.2 Enhancing the CPMD thermostat

In order to improve the performance of the simple low-pass filter GLE (2.10) we can introduce a fitting procedure, which is necessary if we wish move on to a larger parameter space. We can optimize the sampling efficiency by using (3.3) over the ionic frequencies range, up to a cutoff frequency ω_{CP} . We then introduce an additional term so as to force a steep decrease of the effectiveness of the thermostat for higher frequency modes,

$$\chi_{CP} = \left[\sum_{\omega_i > \omega_{CP}} \max \left[\log \kappa(\omega_i) - k \log \frac{\omega_{CP}}{\omega_i}, 0 \right]^m \right]^{1/m}. \quad (3.5)$$

The term above ensures a decrease of $\kappa(\omega)$ above ω_{CP} , with a slope k on a logarithmic scale, and values of k as large as 9 can be used. Such a steep slope guarantees an abrupt drop in thermalization efficiency, hence removing the high-frequency components from the noise (see Figure 3.7).

3.2.3 *Ab initio* CPMD

Thermostatting an *ab-initio* CPMD is more challenging. Since wavefunctions are not atom-centered, the coupling of the dynamics of the electronic degrees of freedom to the ions is stronger than in the shell-model case, so the presence of high-frequency components in the noise quickly heats up the electrons. In addition the considerable computational effort involved in *ab-initio* CPMD makes speeding up equilibration and sampling mandatory.

As a test example, we ran simulations of a single heavy water molecule in vacuum, using a standard setup from the literature (see figure 3.9 and Ref. [56]). We ran several independent trajectories for a total of 90 ps, starting from ionic configurations equilibrated at 300 K and from wavefunctions quenched to the Born-Oppenheimer surface [57]. We then take as a reference a massive Nosé-Hoover-chains simulation [46, 58], and compare it with the GLE thermostat results. We used 5 additional momenta and fitted the parameters using the prescriptions described above (the resulting $\kappa_H(\omega)$ curve is reported in the inset of figure 3.9(b)).

Figure 3.7: Thermostatting efficiency, as estimated from Eq. (3.1), for a colored-noise thermostat optimized for Car-Parrinello dynamics. Sampling efficiency is optimized for $\omega \in (10^{-3}, 1)$, and an abrupt drop in efficiency is enforced for $\omega \in (1, 10)$, using the penalty function (3.5) in the fitting. The thick curve corresponds to $k = 9$, the thin curve to $k = 6$ and the dotted one to $k = 3$. The $\kappa(\omega)$ curve for a white-noise thermostat centered on the optimized range is also reported for reference (dotted black curve).

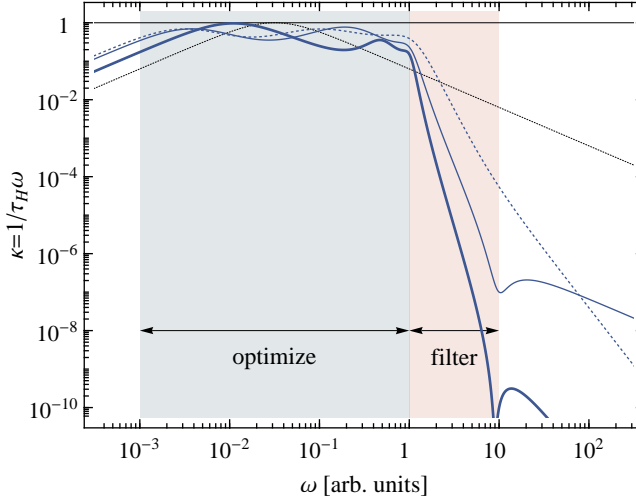


Figure 3.8: Temperature of the electronic degrees of freedom during the initial phases of a Car-Parrinello dynamics of a heavy water molecule in vacuum. The huge drift for a white-noise Langevin thermostat (with a rather strong friction $\gamma^{-1} = 15$ fs) (a) is in direct contrast with the performance of a GLE thermostat (b) (see text and figure 3.9 for details about the parameters) and a Nosé-Hoover chains thermostat of length four, and with coupling frequency $\omega_Q = 2000$ cm $^{-1}$ (c).

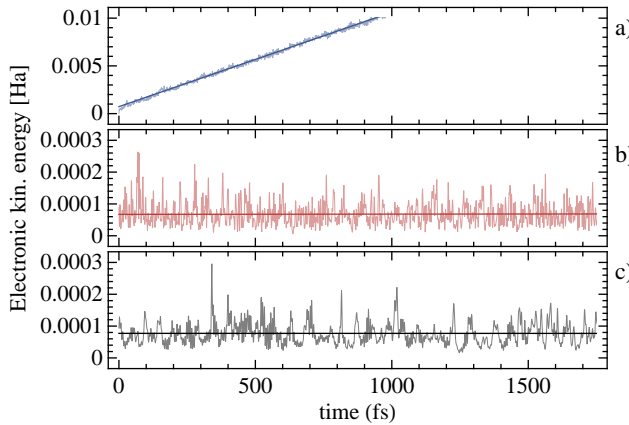
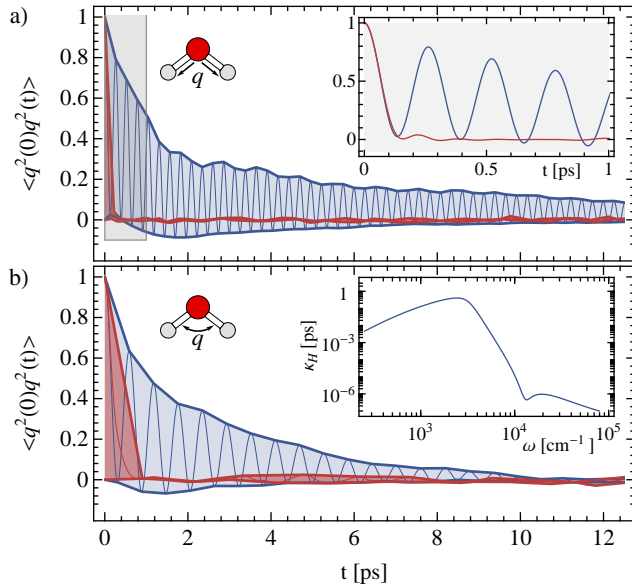


Figure 3.9: Autocorrelation functions for the squared displacements along the symmetric stretching (a) and the bending modes (b) of a heavy water molecule in vacuum, performed in the NVT ensemble at $T = 300$ K. We used a fictitious mass $\mu = 200$ a.u., and a timestep of 4 a.u., in order to minimize the errors on the forces [55]. The Nosé-Hoover thermostat with chain length 4 has been used, with its mass chosen so as to maximize the coupling to the stretching mode. The NH correlation functions (blue) are highly oscillating and decay very slowly (the shading highlights the curve's envelope). In contrast using the new thermostat (red) we find a much sharper decay, which in the case of the stretching requires display on an enlarged scale to be appreciated (inset of panel (a)). In the inset of panel (b) we show the relation between κ_H and ω for our thermostat. The parameters have been optimized so as to obtain a sharp decay of the response for frequencies above the stretching mode.



With the present, very conservative choice of parameters the drift in electronic energy is negligible for both thermostats (see figure 3.8). The strength of the thermostats is such that the underlying dynamics of the ions is severely altered. In Figure 3.9 we plot the autocorrelation function of the squares of the normal modes. The integral of these functions measures the time required to lose memory of the initial configuration. It is evident that the use of an optimized colored-Langevin thermostat dramatically reduces this time. Therefore, a thermostat tuned in this way helps to reduce statistical uncertainty in *ab initio* molecular dynamics simulations, which is particularly relevant given their high computational cost.

3.3 Preserving dynamical properties

Introducing a stochastic term on top of Hamilton's equations will inevitably alter the time-dependent properties of the resulting trajectory. When one is sampling static averages, this is not a concern. Ideally, to compute dynamical properties

one should perform several microcanonical simulations starting from independent configurations at the desired temperature.

In this section we will explore the possibility of obtaining an *a-priori* estimate for the extent of disturbance introduced to the dynamics by a given GLE. Then, extending this concept, one could select which ranges of frequencies should be affected, and which should be undisturbed.

To this end, we consider once more our harmonic oscillator model. The power spectrum of a harmonic dynamics, as computed from the velocity-velocity correlation function, is a Dirac δ -function, centered at the oscillator's characteristic frequency ω_0 . Any sort of dissipation will cause a broadening of this δ function and leave it with a finite peak width, and/or a shift in the peak position. We can therefore compute the power spectrum of the actual GLE dynamics and integrate it in a narrow window of width $\Delta\omega$ around ω_0 , and compare it with the total integral. Then, the larger the intensity leaking out of the window, the larger the disturbance introduced to the dynamics.

Let $\mathbf{x}(t) = (q(t), p(t), \mathbf{s}(t))$ be the vector describing the position in the extended phase space and \mathbf{A}_{qp} and \mathbf{C}_{qp} be the drift and static covariance matrix for a harmonic oscillator of frequency ω_0 (see section 2.2). The Fourier transform of the normalized velocity-velocity correlation function reads (see Eq. (D.2))

$$\mathcal{C}_{pp}(\omega) = \frac{1}{[\mathbf{C}_{qp}]_{pp}} \left[\frac{\mathbf{A}_{qp}}{\mathbf{A}_{qp}^2 + \omega^2} \mathbf{C}_{qp} \right]_{pp}.$$

The integral of $\mathcal{C}_{pp}(\omega)$ between 0 and ∞ is $\pi/2$, therefore the fraction of the power spectrum between two frequencies is

$$W(\omega_1, \omega_2) = \frac{2}{\pi} \int_{\omega_1}^{\omega_2} \mathcal{C}_{pp}(\omega) d\omega = \frac{2}{\pi} \frac{1}{[\mathbf{C}_{qp}]_{pp}} \left\{ \left[-\tan^{-1} \left(\frac{1}{\omega} \mathbf{A}_{qp} \right) \mathbf{C}_{qp} \right]_{\omega_1}^{\omega_2} \right\}_{pp}.$$

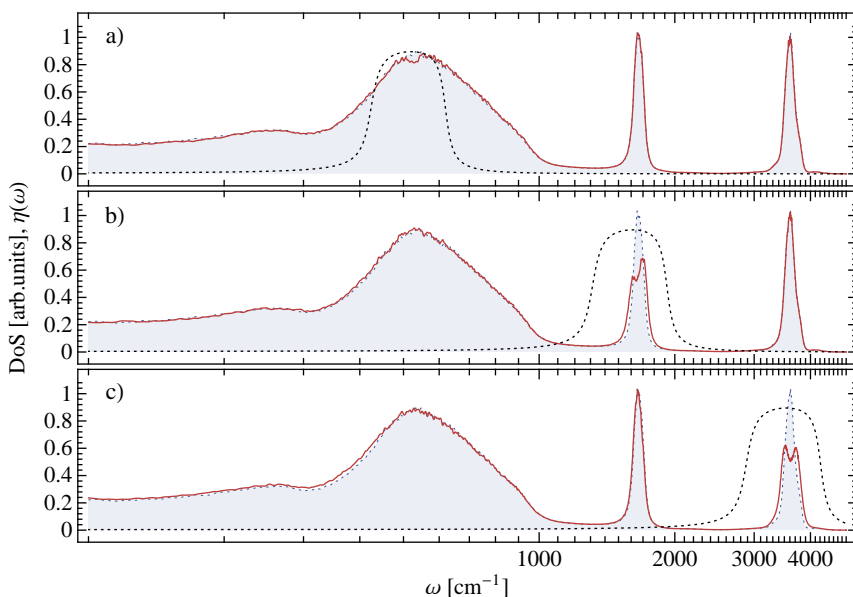
We can thus define a simple measure of the disturbance introduced by the thermostat

$$\eta(\omega_0, \Delta) = 1 - W(\omega_0(1 - \Delta/2), \omega_0(1 + \Delta/2)), \quad (3.6)$$

where Δ is the width of the integration window relative to ω_0 . As usual, we construct a target function modeled using Eq. (2.20) to define regions where we want to preserve the dynamic properties, and then carry out a fitting procedure to obtain a set of parameters for a thermostat that will preserve the dynamics in the specified region.

In the following we will use this idea to design a thermostat which seeks to efficiently couple in the neighborhood of a selected frequency, while disturbing as little as possible the dynamics of the other normal modes. This setup could be used to guarantee efficient, local thermalization of high-frequency modes, without disturbing the slow, diffusive dynamics, which is similar to the use of a dissipative particle dynamics thermostat in MD[12]. While we will not compare the two approaches here, it is interesting to remark that our method has some computational advantage, as one needs only couple individual particles, rather than particle pairs, to a stochastic dynamics.

Figure 3.10: The vibrational spectrum for a flexible water model[59] as computed from the Fourier transform of the velocity-velocity correlation function of a microcanonical dynamics (shaded) is compared with the spectrum measured from a thermostatted simulation (red). We employed a GLE thermostat built so that the disturbance to the dynamics is minimal outside a narrow frequency window. The parameters are then adapted (see section 2.4) to center this disturbance on 530, 1650, 3600 cm^{-1} in panels (a), (b), (c) respectively. The “disturbance” parameter (3.6) with $\Delta = 0.01$ is shown as a dashed line.



In figure 3.10 we demonstrate the behavior of this thermostatting scheme on the dynamics of a flexible-water model at 298 K. The dynamical properties, as estimated from the vibrational density of states, are not perturbed unless $\eta \gg 0$, in accordance with the predictions performed in the harmonic limit. The disturbance is less apparent but still present when the region with $\eta \neq 0$ overlaps with anharmonic, broad features of the spectrum. Also the diffusion coefficient, which corresponds to the $\omega \rightarrow 0$ limit, is the same within statistical uncertainty in all the three cases and in the microcanonical dynamics⁴.

3.4 Liquid water: a case study

We started this chapter comparing global and local thermostatting schemes. We discussed how the former disturb the dynamics less, while the latter are more suitable for dealing with the imbalance in the energies of different parts of the system, or

⁴As a side note, we remark that in order to compute the vibrational spectrum or the diffusion coefficient from NVE dynamics, one must perform several independent simulations starting from configurations equilibrated at the desired temperature as, if this is not done, one risks finding spurious effects due to the infinitely long correlation time of total energy.

more generally for enhancing the sampling when the natural lifetime of phonons is very long.

When one is simulating a strongly anharmonic system such as a liquid, the diffusive modes are the most challenging sampling problem. Molecular dynamics is particularly successful in this respect, and for this reason global thermostats, which have a negligible effect on the dynamics, are a very good for systems where diffusion is important[13, 41]. We discussed in section 3.1.3 how the optimal sampling GLE thermostat can address this problem in a local framework, since it automatically reduces the disturbance on long-time dynamics, when fitted to a very broad frequency range.

We will conclude the present chapter by applying our GLE framework to a flexible model of water[59]. This test is particularly challenging, since a broad range of vibrational frequencies are present, ranging from localized and relatively harmonic stretchings, to anharmonic hydrogen-bond network modes and eventually diffusive motion. Note that the model was originally fitted for path-integral simulations. Therefore, we do not expect to achieve quantitative agreement with experiments by running purely classical dynamics. Our reason for choosing this force field was purely pragmatic - it will be discussed also in a proper path-integral framework (chapter 4) and in an application of the quantum thermostat (section 5.2). We will carry out a similar comparison in chapter 4, where we further demonstrate the advantages of the optimal-sampling GLE, which behaves as a “one-size-fits-all” thermostat, that is able to provide satisfactory, robust results with minimal knowledge of the properties of the system being studied.

3.4.1 Local and global schemes

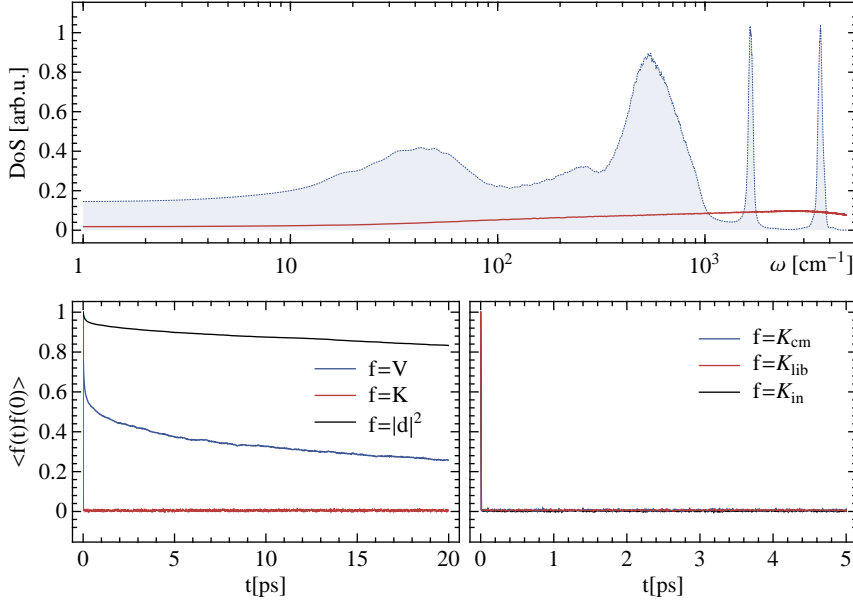
We first perform a couple of benchmark calculations using conventional, white-noise Langevin dynamics and stochastic velocity rescaling[13], which we take as examples of a local a global thermostat respectively. This will allow us to set the standards for comparison with GLE thermostating, but also to highlight the strengths and deficiencies of these two approaches.

For each choice of parameters⁵ we monitor the vibrational spectrum as computed from the velocity-velocity correlation function. The deviation from the constant-energy density of states then serves as a measure of the magnitude of the disturbance introduced in the dynamics. We also compute the correlation properties of the total potential and kinetic energies, and of the square modulus of the dipole moment of the supercell. Finally, we also compute correlation properties for the projections of the kinetic energy on the internal, rotational and translational motions of the water molecules⁶. The first three quantities serve as examples of global properties. The

⁵The simulations have been performed with a modified version of the DL-POLY[60] MD code. A cubic supercell containing 216 water molecules at the experimental room-temperature density have been used, with a time-step of 0.5 fs. Velocity-velocity correlation functions were computed out of four independent 12.5 ps-long runs; the correlation function has been multiplied by a triangle windowing function before being Fourier transformed. Other autocorrelation functions have been computed from 4 independent runs, each 1 ns long.

⁶These three components are defined as follows. m_O , m_H and $m = 2m_H + m_O$ are the masses of oxygen, hydrogen and of a water molecule respectively, and $\mathbf{v}_{H1(2)}$ and \mathbf{v}_O are the velocities of the

Figure 3.11: Dynamical and correlation properties for the dynamics of a flexible water force field[59], thermostatted at 298 K with a white-noise Langevin equation, with the friction set to $\gamma^{-1} = 1$ fs. The NVE density of states (shaded) is compared with the actual vibrational spectrum. Lower panels show the computed autocorrelation functions for a number of observables. On the left, the total potential and kinetic energies (V and K respectively), and the squared modulus of the total dipole moment are shown, while, on the right, the kinetic temperature is decomposed into the contributions from center-of-mass velocity of individual water molecules, K_{cm} , from hindered-rotational motion, K_{lib} and from internal vibrations (stretchings and bending) K_{in} .

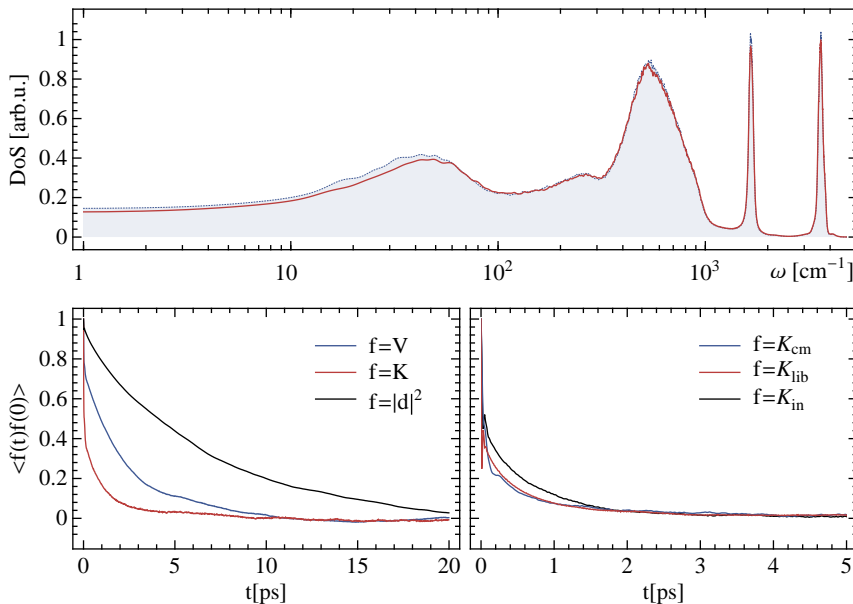


fluctuations of the dipole moment (which can be related to the dielectric constant of the liquid), in particular, are known to converge slowly. The decomposition of the different components of the kinetic energy, instead, allows us to discuss the efficiency in thermostating “semi-local” properties, i.e. the ability of the thermostat to contrast the unbalance between different portions of the system.

White-noise Langevin dynamics The first example calculation was done with a white-noise, local Langevin thermostat with a very strong friction, $\gamma^{-1} = 1$ fs. The results are shown in figure 3.11. Just as in the overdamped case of figure 3.1, the dynamical properties are completely lost. All the components of the kinetic energy are randomized very quickly, however the sampling of configuration space is slowed

two hydrogen atoms and of the oxygen. First, we compute the velocity of the center of mass as $\mathbf{v}_{cm} = m^{-1} (m_O \mathbf{v}_O + m_H \mathbf{v}_{H1} + m_H \mathbf{v}_{H2})$, from which it is trivial to compute the kinetic energy of the center of mass, $K_{cm} = m \mathbf{v}_{cm}^2 / 2$. The residual velocities, after subtraction of \mathbf{v}_{cm} , must be further processed. To compute a (hindered-)rotational contribution K_{lib} we compute the inertia tensor and total angular momentum relative to the center of mass, \mathbf{I} and \mathbf{L} , and get $K_{lib} = \mathbf{L}^T \mathbf{I}^{-1} \mathbf{L} / 2$. The remainder is the internal (stretching and bending) contribution, $K_{in} = K - K_{cm} - K_{lib}$. These values are then summed over all individual molecules to find the overall terms, whose autocorrelation functions are plotted in the figures of this section.

Figure 3.12: Dynamical and correlation properties for the dynamics of a flexible water force field[59], thermostatted at 298 K with a white-noise Langevin equation. The friction here is set to $\gamma^{-1} = 1$ ps. The caption of figure 3.11 provides a description of the quantities shown.

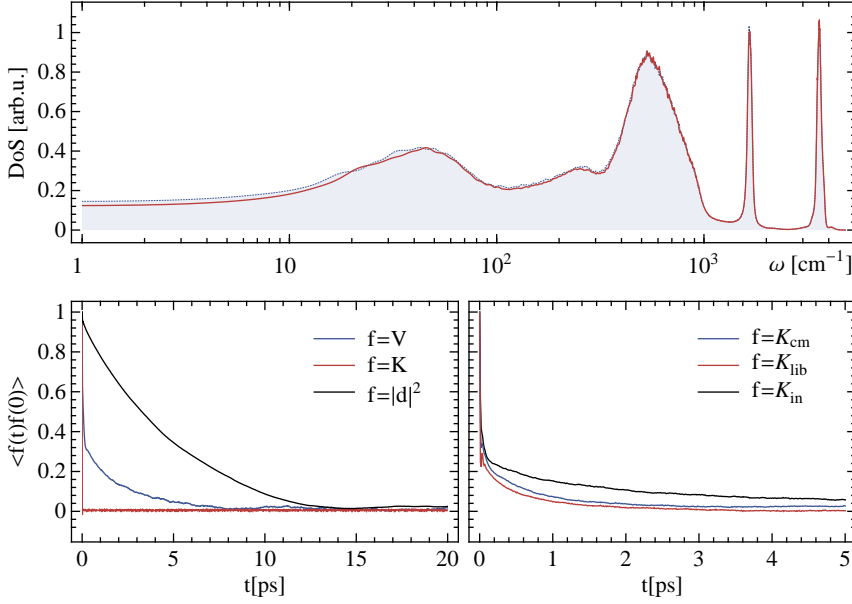


down considerably. In particular, relaxation time of the squared dipole moment of the cell is extremely long, which would make converging the statistical error problematic.

The correlation properties are much better with a friction tuned to $\gamma^{-1} = 1$ ps (see figure 3.12). The disturbance to dynamical properties is smaller, and the configuration sampling is more efficient. Note however that the sampling of fast modes will be much less efficient than that obtained in the overdamped case. This is not a major concern, if one is interested only in equilibrium sampling of slow, diffusive modes. On the other hand, during equilibration or when performing *ab initio* simulations one may wish to apply more aggressive thermostating. In summary, finding the best γ for a white-noise LE requires a lot of trial and error, not least because the efficiency of sampling depends strongly on the observable of interest.

Stochastic velocity rescaling Let us examine the effect of a *global* thermostat; namely, stochastic velocity rescaling[13] with a very strong friction, $\gamma^{-1} = 1$ fs (figure 3.13). This method has only a marginal effect on the dynamics, as the sole *total* kinetic energy is modified. Therefore, the effect on individual particles becomes negligible for systems of moderate size. Hence, a strong coupling can be used without affecting the sampling of the diffusive modes[13, 41]. The sampling of a difficult quantity such as the squared dipole moment is very efficient and, at the same time, the total kinetic energy is thermalized almost instantaneously. Note however that the projection of the kinetic energy on the various types of molecular motion has a relatively long correlation time, which reflects a similar, slow relaxation of the corre-

Figure 3.13: Dynamical and correlation properties for the dynamics of a flexible water force field[59], thermostatted at 298 K with a stochastic velocity rescale method[13]. The friction is set to $\gamma^{-1} = 1$ fs. The caption of figure 3.11 provides a description of the quantities shown.

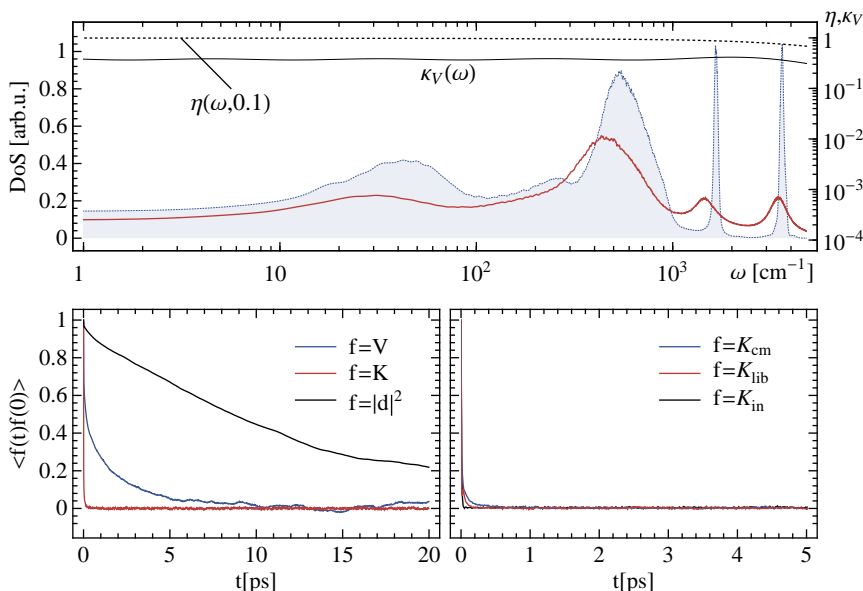


sponding vibrational modes. This observation highlights the main drawback of global thermostating discussed in figure 3.2; namely because only the total kinetic energy is taken into account, the relaxation time of different portions of the system is dictated by their intrinsic dynamics. During equilibration, or when quasi-equilibrium dynamics are performed (see section 3.5) this effect can lead to sampling of a significantly distorted statistical ensemble. For this reason, one must be particularly careful and monitor both global and local properties so as to ensure that equilibrium has been reached.

3.4.2 Optimal sampling

We now undertake a practical test of the ideas discussed in section 3.1. We use a GLE which satisfies the FDT, and whose parameters have been fitted to provide constant efficiency over a broad frequency range that encompasses most of the vibrational spectrum of liquid water. We measure the efficiency by the value of κ_V predicted in a purely harmonic case (Eq. (3.1)). In figure 3.14 we report the relaxation properties of the resulting dynamics. As predicted from the computed η (Eq. (3.6)) the dynamics are heavily modified. However, the modifications are not nearly as severe as they are in the overdamped case (figure 3.11). As the thermostat is tuned so as to be close to critical damping conditions over the whole vibrational spectrum, we observed fast relaxation of all the observables we compute. We underscore that the parameters for the GLE thermostat have been chosen based on general arguments, and only

Figure 3.14: Dynamical and correlation properties for the dynamics of a flexible water force field[59], thermostatted at 298 K with a GLE thermostat fitted so as to provide optimal sampling for harmonic modes spanning four order of magnitude in vibrational frequencies, between 0.5 and 5000 cm^{-1} . The theoretical predictions for κ_V (full line) and η (dashed line) are also reported on top of the vibrational spectrum. In the lower panels, we report the correlation functions for a number of observables (see the caption of figure 3.11 for a description of the quantities shown).



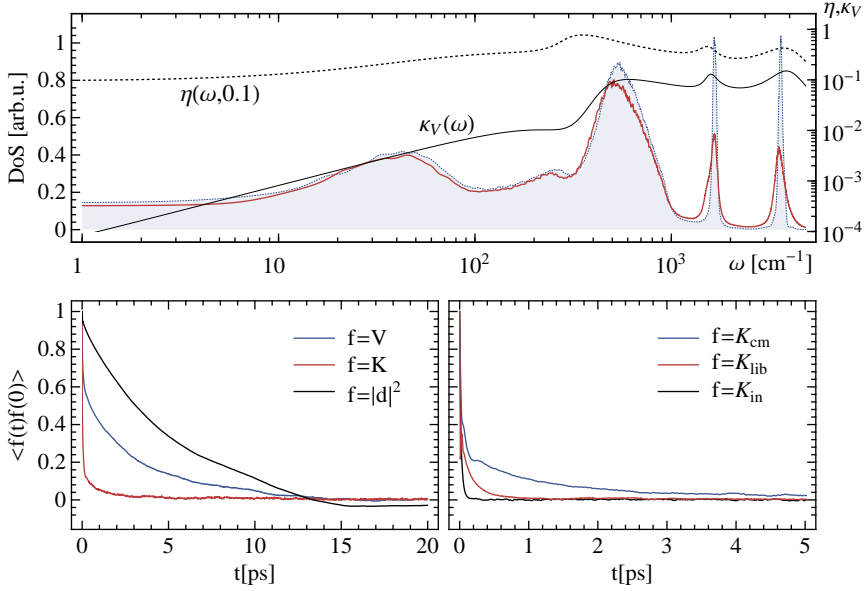
an order-of-magnitude estimate of the extent of relevant vibrational frequencies. Thus, our fitting strategy turns a method with tens of parameters into a black-box procedure, which enforces efficient sampling on very different properties without any need for empirical testing. Moreover, as will be discussed in chapter 4, such broadband optimal-sampling noise is resilient to huge shifts in the optimized frequency range, as opposed to the dramatic degradation of efficiency which is witnessed in the case of uncorrelated noise.

One flaw is that the observed correlation time of the squared dipole moment is significantly larger (even if not dramatically so) than in the case of stochastic rescaling (figure 3.13). In fact, fluctuations in such observable depend on concerted rearrangements of the hydrogen-bond network, which are completely anharmonic and are disturbed by any aggressive thermostating. However, the flexibility of GLE framework even provides tools for tackling this problem, as described below.

3.4.3 A thermostat for water

The thorough testing and discussion performed above leads to a list of desiderata for the thermostating scheme. Ideally, we would like to achieve an efficient coupling to the internal degrees of freedom and the to the stiffer librations, while leaving the anharmonic, diffusive portion of the vibrational spectrum as undisturbed as possible.

Figure 3.15: We designed a GLE thermostat to be used in simulations containing liquid water. The efficiency of sampling for the potential energy κ_V is optimized over the frequency range containing stretchings, bendings and hydrogen-bond network modes. For frequencies below 300 cm^{-1} , we require a reduced disturbance on the dynamics, as measured by η (see Eq. (3.6)). These predictions in the harmonic limit are compared with the actual vibrational spectrum computed from a thermostatted dynamics of a flexible water model[59] at 298 K, and with the NVE density of states (shaded), which is drawn as a reference. In the lower panels, we report the correlation functions for a number of observables (see figure 3.11 for a description of the quantities shown).



We can enforce the first requirement during fitting by using κ_V to measure sampling efficiency of stiff modes. The disturbance on the low-frequency region meanwhile can be gauged by monitoring η .

This is a difficult fitting problem, since we want to achieve somewhat contradictory results; namely, we would like the colored noise to greatly disturb the dynamics of the fast modes, while at the same time it should not harass the slower vibrations, which will be particularly sensitive to the presence of the noise. To resolve this, it is important to start with good initial parameter set. To this end, we use a superposition of three δ -like contributions to the memory kernel (see section 2.1.3) centered on the stretching, bending and libration peaks of the DOS. We then let the fitting algorithm optimize the GLE properties.

The resulting $\kappa_V(\omega)$ and $\eta(\omega, 0.1)$ are plotted in the upper panel of figure 3.15, together with the vibrational spectrum computed from the thermostatted dynamics. One can see that the high-frequency, localized modes are efficiently thermalized (despite the coupling being far from the optimal $\kappa_V = 1$), while the diffusive modes experience only a minor disturbance.

The relaxation properties computed for the flexible-water model conform with expectations. Fast, internal modes are thermostatted efficiently, as are the hindered

rotation motions corresponding to frequencies between 500 and 1000 cm^{-1} . Slower, diffusive behavior meanwhile is minimally affected, and we obtain a short correlation time for the squared dipole moment, as desired. The price to pay is that the thermostating of K_{cm} is less effective than in the optimal-sampling case, and closer to the behavior of a global thermostat. However, as discussed below, there are great advantages in using a local thermostat whenever fast equilibration is required.

3.5 Thermostating quasi-equilibrium dynamics

As a final example of the relevance of an efficient thermostat in the context of molecular-dynamics simulations we turn to the problem of non-equilibrium dynamics. Many techniques for estimating free-energy differences must be performed at quasi-equilibrium conditions, either by slowly dragging the system along a pre-defined reaction coordinate (e.g. steered MD[61]) or by introducing a history-dependent bias potential (this is the case of metadynamics[62, 63]).

These methods share the common feature that energy is injected in the system to artificially accelerate what would otherwise be rare, activated events. In order to guarantee that quantitative estimates for the free-energy differences are computed, one must assume that such energy supply will be dissipated quickly and that the system will remain close to equilibrium under the imposed bias. This condition holds true only in the limit of infinitely slow steering or infinitely slow addition of bias. This is obviously in conflict with the user's desire to reduce the computational effort. However, an efficient, well-designed thermostat can help speed up equilibration, and provide more accurate results in a given simulation time.

3.5.1 Metadynamics

While a critical discussion of the relative merits of different free-energy techniques is beyond the scope of this section, we need to introduce the concepts and the important parameters for one of those methods, as we will use it in the examples that follow. Metadynamics aim is to reconstruct the free energy with respect to a set of collective variables $\{s_i(\mathbf{x})\}$

$$F(\mathbf{s}_0) = -k_B T \ln \int e^{-\beta V(\mathbf{x})} \delta(\mathbf{s}(\mathbf{x}) - \mathbf{s}_0) d\mathbf{x}$$

by performing a dynamical trajectory in which the physical potential is supplemented by an history-dependent repulsive bias, constructed as a superposition of Gaussians centered at points previously visited by the trajectory

$$\tilde{V}(\mathbf{x}, t) = w \int_0^t e^{-(\mathbf{s}(\mathbf{x}) - \mathbf{s}(\mathbf{x}(t')))^2 / 2\sigma^2} dt'. \quad (3.7)$$

The relevant parameters here are the rate of deposition w and the width of the Gaussian functions σ , which define the resolution at which the free-energy surface can be reconstructed[62–66]. As the bias builds up, the system is discouraged from

remaining in the same region of collective-variable space. It has been shown[67] in a high-friction Langevin model that the negative of the bias, once the relevant portion of the phase space has been explored, provides an unbiased estimator of the underlying free energy.

However, we have observed that, when an overdamped Langevin equation is used, the sampling of diffusive modes is hampered considerably. This could eventually result in slower convergence of the bias to the correct free-energy profile. In the previous section we have seen that a global thermostat, such as stochastic velocity rescaling, works well to enforce a canonical distribution of the total kinetic energy, while also ensuring minimal disturbance of the dynamical properties and of the diffusive, collective modes of a liquid in particular. On the other hand, we have also noticed that local properties are equilibrated slowly, as this equilibration is dependent on internal anharmonic coupling between the different components of the system. Intuitively, this drawback is particularly worrisome in a non-equilibrium simulation such as the one stemming from the history-dependent bias (3.7). Energy is injected into the system, often into a few degrees of freedom, as most collective variables depend on the coordinates of few atoms. If quantitative results are to be obtained, this additional energy must be dissipated quickly, and it is therefore questionable whether one can rely on internal couplings only.

3.5.2 Alanine dipeptide in water

In order to discuss this issue in more detail, we present a realistic example, i.e. the calculation of the free-energy profile for alanine dipeptide in explicit water, a classic model system which, in spite of its simplicity, captures some of the features of more complex - and computationally expensive - problems⁷.

The conformation of the backbone of the dipeptide is described by two dihedral angles, and the energetic landscape is relatively complex, containing valleys and mounds spanning a range of several tens of kJ/mol (see figure 3.16). The history-dependent bias introduces forces on few atoms, and we can monitor the kinetic temperature of the peptide atoms only. We note in passing that monitoring the temperature of a small group of particles is not simple, since fluctuations are very large and one needs to average several independent trajectories to observe a trend which is statistically significant.

We ran several simulations with different values for the bias deposition rate, in which the filling rate (the *volume* of bias deposited per unit time) was kept constant by adjusting the Gaussian width σ . We performed two sets of simulations. In the first case, we used a strong stochastic velocity rescaling thermostat, with friction

⁷We performed simulations using the GROMACS4[68] code, patched with the PLUMED[69] plug-in for metadynamics and with custom routines to implement the GLE thermostat. We have used the TIP3P[70] water potential and OPLS[71] force field for proteins. We prepared a simulation box containing 205 water molecules together with the alanine dipeptide, and equilibrated it at constant temperature and ambient pressure for 10 ns. No bonds or angle constraints have been used. For each set of parameters we then performed 64 independent metadynamics runs at constant volume, each of which was 1 ns long. A timestep of 0.5 fs has been chosen to integrate the equations of motion. As is customary in metadynamics simulations, free-energy profiles have been constructed from the average of snapshots of the bias taken from the point of first complete filling (about 0.5 ns) until the end of each trajectory.

Figure 3.16: Free-energy surface for an alanine dipeptide molecule in water. Contour lines are separated by 5 kJ/mol. The dihedral angles used as collective variables are also illustrated.

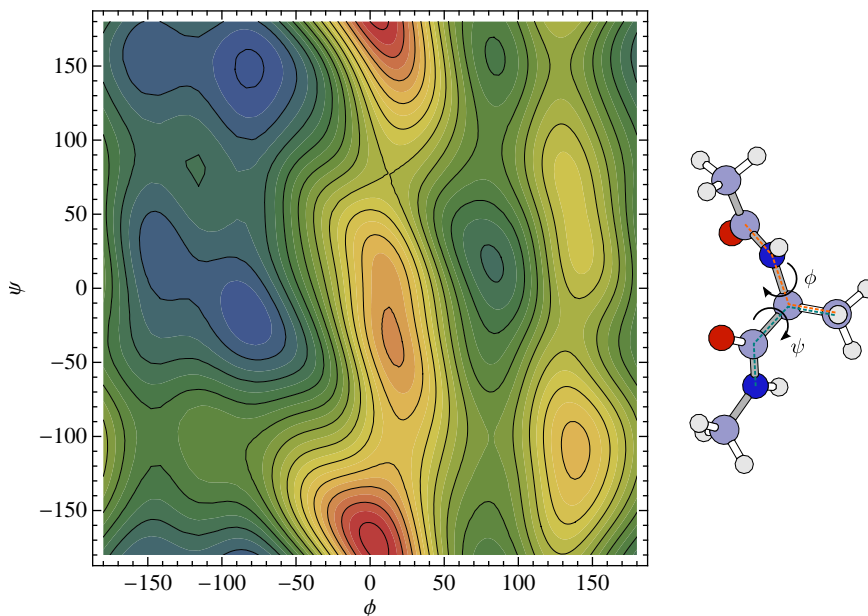
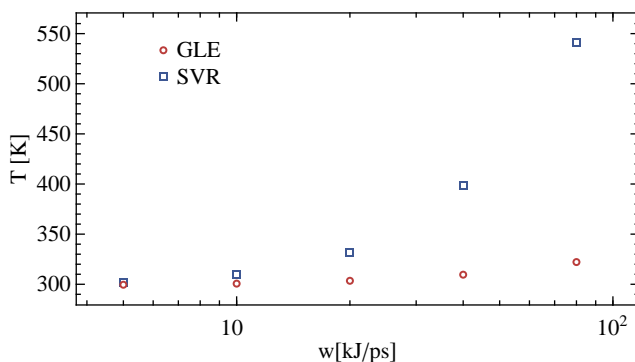


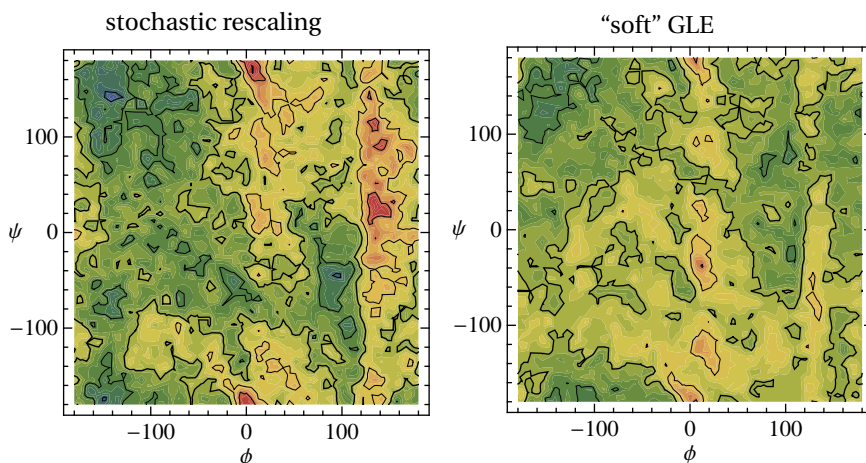
Figure 3.17: Average kinetic temperature of the alanine dipeptide atoms only, during a metadynamics run with bias deposition rate w . In all cases, the average temperature of the *whole* simulation box is within 1 K of the target of 300 K. Results from an optimized GLE thermostat and from stochastic velocity rescaling[13] are compared.



$\gamma^{-1} = 10$ fs, while, in the second case, we resorted to a local GLE scheme, with the parameters described in figure 3.15. As we discussed above, this choice of parameters achieves concurrently the goals of a strong local coupling with minimal disturbance of diffusive, collective modes.

We start with a relatively high deposition rate $w = 5$ kJ mol⁻¹ps⁻¹, and with broad hills with $\sigma = 0.40$ radians. With these parameters, both thermostats succeed in maintaining the kinetic temperature of the whole system and of the protein at

Figure 3.18: Differences between the free-energy surface computed with the highest deposition rate ($80 \text{ kJ mol}^{-1} \text{ps}^{-1}$) and the converged result at low bias deposition rate. Left panel shows the result with a global thermostat, while the rightmost panel shows the GLE case, where we used the same parameters employed in figure 3.15. The thick line corresponds to the zero, and contour lines are separated by 1 kJ/mol . Note that the statistical uncertainty is about 1 kJ/mol , so the GLE result can be considered converged within statistical accuracy.



300 K, within the statistical error bar ($< 1 \text{ K}$). The resulting free-energy surface is identical to the one reported in figure 3.16 within the statistical error, which is around 1 kJ/mol . If the deposition rate is increased, and the width of the hills decreased (so as to maintain a constant “volume”) the average kinetic temperature of the atoms of the peptide deviates from the target value (figure 3.17). The local GLE thermostat - despite using parameters far from optimal coupling (see section 3.4.3) - effectively counterbalances this local deviation from equilibrium, even with a deposition rate as high as $80 \text{ kJ mol}^{-1} \text{ps}^{-1}$ where the deviation from the target temperature is less than 10 %. The global thermostat is considerably less effective, and the dialanine reaches a stationary temperature as high as 550 K.

One would expect to observe major systematic errors in the reconstructed FES, when using such high values of w , especially given the large deviations in the local temperature from the correct ensemble. Interestingly enough, this is not the case. Even for the global thermostat, with $w = 80 \text{ kJ mol}^{-1} \text{ps}^{-1}$, the reconstructed free energy is within 3 kJ/mol of the exact profile (figure 3.18). At lower deposition rate, and even for $w = 80 \text{ kJ mol}^{-1} \text{ps}^{-1}$ when the GLE thermostat is used, the deviation is not statistically significant.

This observation is reassuring, and confirms the reliability of metadynamics as a free-energy technique. We must however underline the fact that alanine dipeptide is a textbook example, where very effective collective variables are known, and that the presence of an anharmonic liquid bath around the peptide mitigates the deficiencies of the global thermostat. In other examples, with a more harmonic environment, or in which there are transverse degrees of freedom that are not described within

the space of CVs, the effects on the free-energy profile might be much more severe. Whenever a high deposition rate is necessary - as it is generally the case in *ab-initio* simulations, for instance - one should monitor the kinetic temperature of the degrees of freedom directly affected by CVs, and opt for a local thermostat whenever possible.

An optimal-sampling GLE, as described in section 3.1, provides very strong coupling over a broad range of frequencies, even if a moderate slow-down of diffusive dynamics is introduced as a side effect. If one wishes to avoid such disturbances, a different colored-noise setup can be obtained with little effort. By balancing the requirements of an intense coupling to quasi-harmonic modes and minimal hindrance of slow collective motion, we obtained the thermostat described in figure 3.15. We have demonstrated in this section that this setup is capable of rapidly balancing local deviations from the desired ensemble, despite being less intensely coupled than an optimal-sampling GLE.

A thermostat for path integral dynamics

Having discussed how the GLE method can be used to enhance the exploration of phase space in constant-temperature molecular dynamics, we now turn to an application of our ideas to a more specific sampling problem. Path-integral methods comprise a set of techniques which are useful for sampling quantum-mechanical, finite-temperature probability distributions, without solving the Schrödinger equation explicitly. As we will discuss below, they require the evaluation of a modified Hamiltonian, which contains some very stiff harmonic modes together with the vibrational modes characteristic of the system. The broad vibrational spectrum makes path-integral methods a challenging sampling problem[32, 40]. At the same time, it is possible to predict the approximate frequencies and displacement patterns for many of the normal modes present, which makes this problem an ideal benchmark for our method.

In this chapter we will examine the sampling of the path-integral probability distribution in conjunction with white and colored-noise Langevin dynamics. We will also introduce as a benchmark a new stochastic approach, which is based on knowledge of the approximate vibrational frequencies of the ring polymer. We will present correlation times for a number of different observables in two very different condensed-phase problems; namely, liquid water at room temperature and the diffusion of atomic hydrogen in noble metals.

4.1 Path-integral molecular dynamics

Before describing the sampling problems connected with path-integral molecular dynamics (PIMD), we must briefly survey the foundations of this technique. In particular, we must introduce a viable strategy to integrate the equations of motion in the absence of any thermostat. To this end, consider a system of N distinguishable

atoms described by a Hamiltonian of the form

$$\mathcal{H} = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i} + V(q_1, \dots, q_{3N}),$$

in which the potential energy $V(q_1, \dots, q_{3N})$ is such that the canonical quantum mechanical partition function $Z = \text{Tr} e^{-\mathcal{H}/k_B T}$ is well defined. After a standard Trotter product factorization of the trace, and inserting a number of complete sets of states, the partition function can be written as the configurational integral

$$Z = \int d\mathbf{q}_1 \dots d\mathbf{q}_P \langle \mathbf{q}_1 | e^{-\mathcal{H}/Pk_B T} | \mathbf{q}_2 \rangle \dots \langle \mathbf{q}_P | e^{-\mathcal{H}/Pk_B T} | \mathbf{q}_1 \rangle.$$

For large values of P , one can make a high-temperature expansion of the density matrix $\langle \mathbf{q}_1 | e^{-\mathcal{H}/Pk_B T} | \mathbf{q}_2 \rangle$, and obtain an expression which is equivalent to the classical partition function for a fictitious ring polymer consisting of P copies of the system connected by harmonic springs (figure 4.1). The error in this expansion is $\mathcal{O}(1/P^2)$ and so vanishes in the limit $P \rightarrow \infty$. In view of the analogy between this classical partition function and that of a polymer with harmonic bonds, we will refer to the ensemble of replicas as a ring polymer (RP) or necklace, and to individual instances of the system as a bead. For each particle in the system we will refer to the center of mass of the corresponding ring polymer as the centroid for that atom.

The sampling of this configurational partition function can be performed with Monte Carlo methods[72, 73]. However, as pointed out by Parrinello and Rahman[74], it is also possible, and often advantageous, to introduce momenta and use molecular dynamics to facilitate the exploration of configuration space. The partition function containing both positions and momenta can be written as

$$Z \sim \frac{1}{(2\pi\hbar)^{3NP}} \int d^{3NP} \mathbf{p} \int d^{3NP} \mathbf{q} e^{-\mathcal{H}_P(\mathbf{p}, \mathbf{q})/Pk_B T}. \quad (4.1)$$

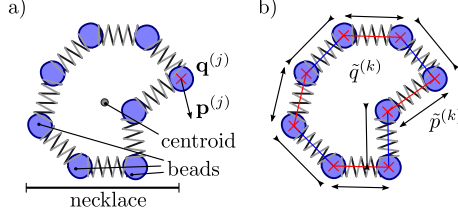
Here $\mathcal{H}_P(\mathbf{p}, \mathbf{q})$ is the classical ring polymer Hamiltonian which is given by

$$\begin{aligned} \mathcal{H}_P(\mathbf{p}, \mathbf{q}) &= \mathcal{H}_P^0(\mathbf{p}, \mathbf{q}) + V_P(\mathbf{q}) \\ \mathcal{H}_P^0(\mathbf{p}, \mathbf{q}) &= \sum_{i=1}^{3N} \sum_{j=1}^P \left(\frac{[p_i^{(j)}]^2}{2m_i} + \frac{1}{2} m_i \omega_P^2 [q_i^{(j)} - q_i^{(j-1)}]^2 \right) \\ V_P(\mathbf{q}) &= \sum_{j=1}^P V(q_1^{(j)}, \dots, q_{3N}^{(j)}), \end{aligned} \quad (4.2)$$

with $\omega_P = Pk_B T/\hbar$ and closed-path boundaries $q_i^{(0)} \equiv q_i^{(P)}$ (see figure 4.1)¹. Superscripts label the different replicas. Path integral methods use this classical isomorphism as a computational tool for calculating quantum mechanical equilibrium

¹As we have written it, the free ring polymer Hamiltonian $\mathcal{H}_P^0(\mathbf{p}, \mathbf{q})$ corresponds to a particular choice of the Parrinello-Rahman mass matrix, in which the physical particle mass is assigned to each ring polymer bead. This choice is used in the ring polymer molecular dynamics approximation to the real-time quantum correlation functions[75, 76].

Figure 4.1: In path integral methods each particle is represented by a ring polymer, which consists of several replicas connected by harmonic springs. The configuration of the necklace can be described equally well in terms of the Cartesian coordinates of individual beads (a), or on the basis of the normal modes of the free ring polymer (b).



properties of the form $\langle \mathcal{A} \rangle = Z^{-1} \text{Tr} [e^{-\mathcal{H}/k_B T} \mathcal{A}]$, in which the operator \mathcal{A} can be expressed in terms of the positions \mathbf{q} only. In these cases, the quantum mechanical expectation value can be recovered by an ensemble average performed on a set of configurations consistent with the PI classical partition function (4.1)

$$\langle A \rangle \sim \frac{Z^{-1}}{(2\pi\hbar)^{3NP}} \int d^{3NP} \mathbf{p} \int d^{3NP} \mathbf{q} e^{-\mathcal{H}_P(\mathbf{p}, \mathbf{q})/k_B T} \mathcal{A}_P(\mathbf{q}) \quad (4.3)$$

where $\mathcal{A}_P(\mathbf{q}) = \sum_{j=1}^P \mathcal{A}(q_1^{(j)}, \dots, q_{3N}^{(j)})$.

4.1.1 Normal-modes integrator for PIMD

Any method for exploring the canonical distribution by molecular dynamics can be used in this context. However, the robustness and the efficiency of sampling, which can be measured by the correlation time of the observables of interest, as discussed in chapter 3, should be considered when making this choice. Before going on to describe the various strategies available for enforcing efficient canonical sampling, let us first discuss an algorithm that can be used to integrate the set of Hamilton's equations that result from $\mathcal{H}_P(\mathbf{p}, \mathbf{q})$,

$$\dot{\mathbf{q}} = \frac{\partial}{\partial \mathbf{p}} \mathcal{H}_P(\mathbf{p}, \mathbf{q}), \quad \dot{\mathbf{p}} = -\frac{\partial}{\partial \mathbf{q}} \mathcal{H}_P(\mathbf{p}, \mathbf{q}).$$

This is in itself a difficult problem, as the free ring polymer dynamics contains fast modes with frequencies

$$\omega_k = 2\omega_P \sin \frac{k\pi}{P} \quad (4.4)$$

which arise from the harmonic springs between neighboring beads. Consequentially, a very small time step is required to integrate these stiff vibrations properly. A convenient way to circumvent this problem is based on splitting of the Hamiltonian (4.2) into a sum of a free ring polymer part $\mathcal{H}_P^0(\mathbf{p}, \mathbf{q})$ and a potential energy part $V_P(\mathbf{q})$. The evolution generated by these two parts can be combined using a symplectic integration scheme in which the phase space density evolves under the influence of

the symmetric split operator propagator[30]

$$e^{-\Delta t \mathcal{L}} \sim e^{-\frac{\Delta t}{2} \mathcal{L}_V} e^{-\Delta t \mathcal{L}_0} e^{-\frac{\Delta t}{2} \mathcal{L}_V}, \quad (4.5)$$

where $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_V$ is the Liouvillian associated with $\mathcal{H}_P(\mathbf{p}, \mathbf{q})$, and \mathcal{L}_0 and \mathcal{L}_V are those associated with $\mathcal{H}_P^0(\mathbf{p}, \mathbf{q})$ and $V_P(\mathbf{q})$. The exact evolution generated by $\mathcal{H}_P^0(\mathbf{p}, \mathbf{q})$ can be greatly simplified by transforming the ring polymer from the bead representation into the normal mode representation (figure 4.1(b)),

$$\{\tilde{\mathbf{p}}^{(k)}\} \leftarrow \{\mathbf{p}^{(j)}\}, \quad \{\tilde{\mathbf{q}}^{(k)}\} \leftarrow \{\mathbf{q}^{(j)}\},$$

which can be done very efficiently using fast Fourier transform routines.

The first step in the algorithm is thus to evolve the momenta through a time interval $\Delta t/2$ under the influence of $V_P(\mathbf{q})$. Then, one transforms to the normal mode representation, and performs the exact evolution through a time interval Δt under the influence of the free ring polymer Hamiltonian $\mathcal{H}_P^0(\mathbf{p}, \mathbf{q})$. Coordinates and momenta are then transformed back to the bead representation, forces are computed and momenta are evolved again under the influence of the potential.

4.2 Constant-temperature PIMD

Constant-temperature sampling can be enforced in path integral dynamics based on the same methods which are applied to conventional molecular dynamics, because PIMD is simply classical MD in an extended phase space. In particular, the integrator described in section 4.1.1 is a direct generalization of the velocity Verlet algorithm, so it is therefore straightforward to adapt any stochastic thermostat to the present context.

4.2.1 White-noise Langevin thermostat

As discussed in length in previous chapters, Langevin's equation is a very convenient method for including fluctuations of the total energy, consistent with the constant-temperature canonical ensemble, in the context of molecular dynamics[1]. Once the target temperature has been fixed, a white-noise Langevin equation (WNLE) thermostat has a single parameter, the friction γ (see e.g. Eq. (1.15)). In what follows we express the friction as the inverse of the characteristic relaxation time for the Langevin equation, $\tau_0 = 1/\gamma$, so as to set a common scale for comparing different strategies.

A convenient integrator can be obtained using a symmetric Trotter factorization[31], where (4.5) is modified to

$$e^{-\Delta t \mathcal{L}} \sim e^{-\frac{\Delta t}{2} \mathcal{L}_\gamma} e^{-\frac{\Delta t}{2} \mathcal{L}_V} e^{-\Delta t \mathcal{L}_0} e^{-\frac{\Delta t}{2} \mathcal{L}_V} e^{-\frac{\Delta t}{2} \mathcal{L}_\gamma},$$

where \mathcal{L}_γ is the part of the Liouvillian in the FP equation for the Langevin phase-space density that introduces the friction and the thermal noise. From an algorithmic

point of view, this splitting can be implemented by introducing the Langevin propagator

$$p_i^{(j)} \leftarrow e^{-\gamma\Delta t/2} p_i^{(j)} + \sqrt{Pm_i k_B T} \sqrt{1 - e^{-\gamma\Delta t}} \xi_i^{(j)} \quad (4.6)$$

before and after the steps described in section 4.1.1. Here ξ 's are uncorrelated Gaussian processes, $\langle \xi_i^{(j)}(t) \xi_{i'}^{(j')}(t') \rangle = \delta_{ii'} \delta_{jj'} \delta(t - t')$.

4.2.2 A path integral Langevin thermostat

The stochastic part of the propagator (4.6) can be applied just as well in the normal-modes representation

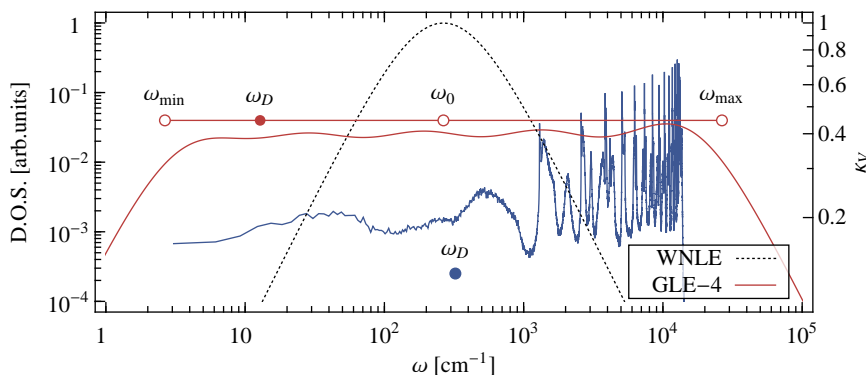
$$\tilde{p}_i^{(k)} \leftarrow e^{-\gamma_k \Delta t/2} \tilde{p}_i^{(k)} + \sqrt{Pm_i k_B T} \sqrt{1 - e^{-\gamma_k \Delta t}} \xi_i^{(k)}.$$

However, the advantage of working in the normal mode representation is that one can use different friction coefficients γ_k for each of the vibrational modes. This choice can be performed so as to give an optimal sampling of the canonical distribution for the free ring polymer. To this end, consider a ring polymer whose beads interact with an external harmonic potential corresponding to the frequency ω_0 . In this simple model the k -th normal mode in the necklace dynamics has a frequency given by $\omega'_k = \sqrt{\omega_0^2 + \omega_k^2}$, where the ω_k 's are the free ring-polymer normal mode frequencies, given by Eq. (4.4). One can then assume $\omega'_k \approx \omega_k$, as this is a very good approximation for all but the softest necklace modes. When this approximation is valid, the choice for the frictions $\gamma_k = 2\omega_k$ gives the most efficient convergence of the Boltzmann-weighted average total energy (see section 3.1, and equation (3.2) in particular).

The price paid for this simple and transferable choice is that the lowest-lying eigenfrequencies may be shifted by interatomic forces to well above the free ring polymer limit. In particular, this prescription cannot be used to thermostat the centroid mode, as its spectrum corresponds roughly to the classical vibrations of the system. Therefore consider the friction $\gamma = 1/\tau_0$ applied on the centroid as a free parameter of this method, which we dub the local version of a path integral Langevin equation thermostat (PILE-L).

One could also thermostat the centroid using a global scheme (PILE-G), and so we used a stochastic velocity rescaling approach[13]. As discussed in Ref. [41], there is a direct connection between the white-noise Langevin equation and this global thermostating strategy, which is made apparent if the relaxation time (see e.g. Eq. (4) in Ref. [13]) is set to $\tau = 1/2\gamma = \tau_0/2$. As was seen in the results discussed in sections 3.4 and 3.5, this choice turns out to be very effective when measuring global properties of a homogeneous, strongly anharmonic system. However care must be taken as this approach fails when effective, local thermalization is needed, as is the case for harmonic systems or for quasi-equilibrium simulation schemes.

Figure 4.2: The vibrational spectrum for a path integral dynamics of a flexible water model. In order to also include the modes of the necklace, the density of states has been computed from the velocity autocorrelation of the beads. This vibrational spectrum is compared to the harmonic-limit sampling efficiency κ_V for a white-noise Langevin equation and an optimal-sampling GLE thermostat. The latter has been optimized so as to give constant sampling efficiency over a frequency range four orders of magnitude wide, geometrically centered on $\omega_0 = 1/2\tau_0$. The diffusion coefficient for the water model, and that computed for the GLE dynamics in the free-particle limit, Eq. (3.4), are also shown by means of the quantity $\omega_D = k_B T / mD$.



4.2.3 A generalized Langevin thermostat for PIMD

In chapter 3 we discussed how a generalized Langevin equation can be used to enhance the sampling of the canonical ensemble by molecular dynamics. The same concepts can be applied in the case of PIMD. In this context, we would like to tune the colored noise so as to efficiently thermostat a very wide range of frequencies, including both the physical vibrations of the system and the high frequency ring polymer internal modes (see figure 4.2).

This approach has a number of advantages over the method described in section 4.2.2. Not least, one can avoid going into the normal mode representation². Moreover, deviations of the lowest-lying necklace vibrations from the free ring polymer limit are dealt with automatically. Finally, the problems associated with thermostating the centroid are also solved, as one can expect the same level of efficiency obtained for the conventional MD simulations presented in section 3.4.

An additional advantage is that, for an optimal-sampling GLE, the sampling efficiency κ_V can be optimized over the relevant frequency range, which can in turn be chosen based on minimal amount of information, as discussed in section 3.1. As a result, this approach can be considered parameterless. However, for the sake of comparison, we will perform several simulations in which we vary the width of the fitted range and its geometrical center $\omega_0 = 1/2\tau_0$. We will demonstrate that the

²In practice, going into normal mode representation is necessary to integrate the equations of motion, and does not involve a dramatic computational overhead. A scenario in which one would like to reduce the number of Fourier transforms to the minimum is when PI simulation is parallelized over several processors, as each transformation requires all the nodes to share information about the coordinates and velocities of all the beads.

breadth of the optimally-sampled region means that the sampling efficiency is not greatly affected when a wrong estimate is used for ω_0 .

4.3 Simulations of liquid water

Having outlined a number of different strategies, we now discuss two realistic applications, and demonstrate how, by using a good thermostat, one can increase the speed at which ensemble averages in path integral molecular dynamics converge. As a first benchmark we have chosen the simulation of a flexible classical force field model of water[59]. This is an important example, in which path integral methods should be used, as the nuclear quantum effects are large. Moreover, we used the very same potential as a benchmark in section 3.4, and will use it in chapter 5 to demonstrate some applications of non-equilibrium GLEs.

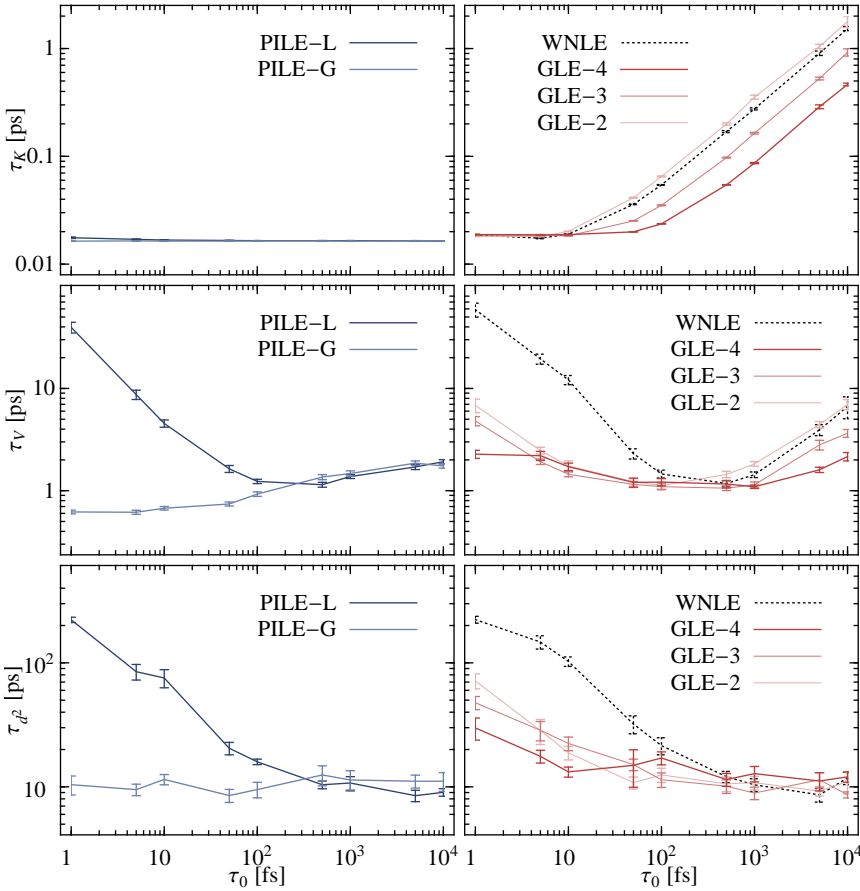
Among the many physical properties which exhibit a non-trivial dependence on the friction γ we will discuss the virial estimator for the kinetic energy, which depends strongly on the internal modes of the necklace, the potential energy which depends mostly on the lowest frequency modes, and the squared dipole moment of the supercell, which depends almost exclusively on the centroid and which, as shown in section 3.4, converges slowly and is very sensitive to the details of the stochastic dynamics. As we discussed above, we performed calculations using various strategies: a simple white-noise local Langevin thermostat (WNLE), a white-noise Langevin adapted to the free ring polymer normal modes, with both a local (PILE-L) and global (PILE-G) thermostat applied to the centroid, and a colored noise, generalized Langevin equation thermostat (GLE- n , where n designates the width of the frequency range on which the sampling has been optimized, expressed in orders of magnitude on a logarithmic scale). We performed an extensive set of simulations at the target temperature $T = 298$ K, with the thermostat time constant τ_0 ranging from 1 fs to 10 ps. For each choice of thermostat and parameters we have performed 8 statistically independent runs, each of which was 1.5 ns long. A time step of 0.75 fs was employed and the simulation box contained 216 water molecules, at the experimental room-temperature density. The ring polymer contained 32 beads, which is necessary to obtain converged nuclear quantum effects for water at ambient conditions³.

In figure 4.3 we report the correlation times for the quantities introduced above. For an observable which depends strongly on internal modes of the necklace such as the kinetic energy, PILE thermostats do an excellent job. The GLE thermostat performs equally well, provided that the upper bound of the fitted range encompasses the whole vibrational spectrum⁴. This is a clear demonstration of the power of the optimal sampling GLE, especially considering that no explicit information on the vibrations of the ring polymer was used.

³The simulations we performed would correspond to a total of $32 \times 6 \times 9 \times 8 \times 1.5 \sim 20 \mu\text{s}$ of dynamics, if performed with conventional PIMD methods. We have however used the ring-polymer contraction scheme for the calculation of electrostatic interactions[77], with the Ewald summation evaluated on a single, contracted mode and a radius of 5 Å for the calculation of short-range electrostatics. This reduces the overhead with respect to a classical simulation to a number much smaller than 32 (typically 2-4) and made simulations dramatically less time-consuming.

⁴Compare with figure 4.2, considering that $\omega_0 = 300 \text{ cm}^{-1}$ corresponds to a τ_0 value of about 10 fs.

Figure 4.3: The correlation time has been computed for the virial estimator to the kinetic energy (τ_K), the total potential energy (τ_V) and the squared dipole moment of the cell (τ_{d^2}) for a path integral MD simulation of liquid water. In the six panels we compare the performance of the path integral Langevin thermostat - both using a (L)ocal and (G)lobal thermostat on the centroid, and the optimal sampling GLE, fitted over a frequency range which is 2, 3 and 4 orders of magnitude wide. A white-noise Langevin reference is also shown.



For the total potential energy, which depends strongly on the motion of the centroid, the thermostating of the internal modes is largely irrelevant. Consequentially, much like in section 3.4, one must choose a small centroid friction in order to avoid hampering diffusion and sampling. The local version of PILE shows an improvement with respect to conventional white noise Langevin thermostat, but only for a very large τ_0 . In contrast, PILE-G, which uses a global thermostat on the centroid, yields small values for τ_V , particularly when the relaxation time is very small. These observations are equivalent to the results we obtained with conventional MD. They can be explained by the fact that the global thermostat minimally hampers the collective modes which are involved in the complex rearrangements of the hydrogen-bonding network that characterize the dynamics of liquid water. Just as in the classical case, the GLE yields a significant improvement over the white-noise Langevin for most choices of τ_0 . It is interesting to note that, particularly for the broadest fitted range, the results are insensitive to the choice of the central frequency, as in all cases significant overlap between the optimally sampled region and the important part of the vibrational spectrum of the centroid is guaranteed. However, the efficiency is not as good as for the global thermostating scheme. Most likely, further improvement of the performance of the GLE can be obtained without sacrificing locality, using a strategy analogous to that described in section 3.4.3. Similar arguments also apply to τ_{d^2} , which depends almost exclusively on the sampling of the centroid motion and for which the advantage of a global thermostating scheme is even clearer.

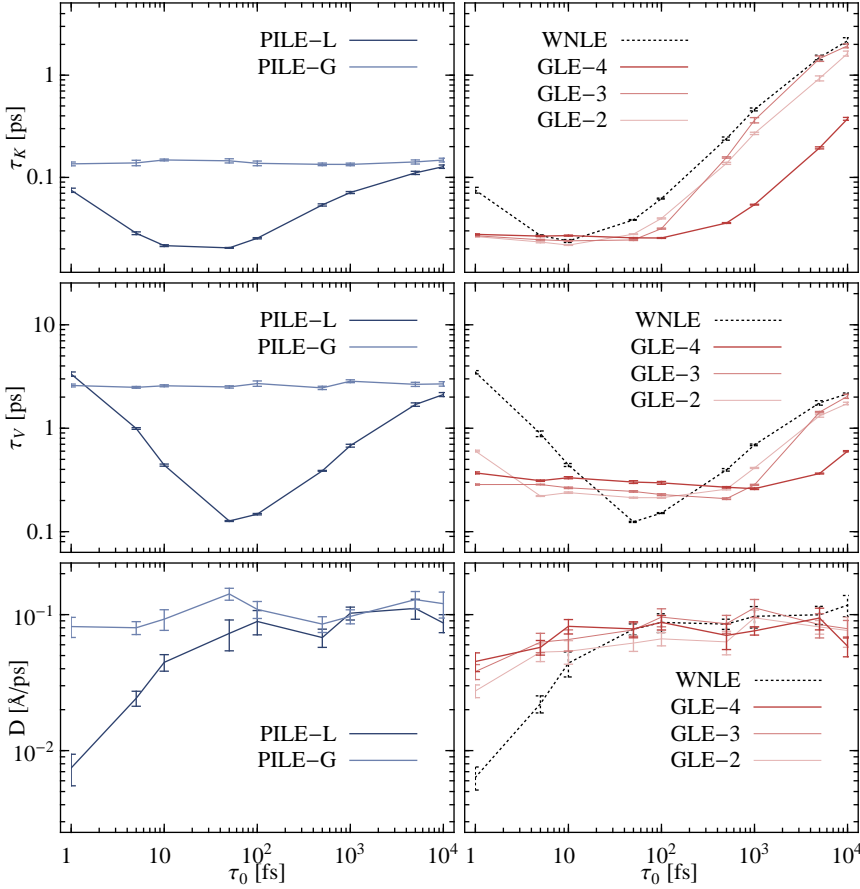
As we observed for classical MD, whenever an observable which converges very slowly is to be computed, efficient sampling of other, faster quantities becomes largely irrelevant. For example, by the time the squared dipole moment has been properly sampled, the estimate for $\langle K \rangle$ would have converged to a small statistical uncertainty even if the least efficient WNLE with $\tau_0 = 10$ ps had been used. However, in a different context one might want to focus solely on $\langle V \rangle$ or $\langle K \rangle$, and one can hardly afford to perform such a detailed analysis for every single application. It is therefore extremely valuable to have thermostats available which provide consistently efficient sampling for a broad spectrum of observables and parameters.

4.4 Simulations of hydrogen in palladium

A radically different solid-state problem was chosen as a second benchmark; namely, the diffusion of a hydrogen atom within the lattice of a palladium crystal, using an embedded-atom potential[78] to compute interatomic forces. Contrary to liquid water, this is a fairly harmonic system, and we decided to monitor local properties in order to test the various methods in a different context. We computed the contributions to the virial kinetic energy estimator and to the potential energy coming from the single hydrogen atom present⁵. We also computed the diffusion coefficient of the

⁵In this example, strong anticorrelations are present when the underdamped limit of the thermostats are considered. Because of this, correlation times turn out to be quite small, even if the actual envelope of the correlation function decays very slowly. To avoid this artifact, when speaking of correlation times of a quantity A in the present chapter, what we really mean is the integral of the *absolute value* of the

Figure 4.4: The correlation time has been computed for the virial estimator to the kinetic energy (τ_K) and the potential energy (τ_V) for a path integral MD simulation of an individual hydrogen atom in a palladium crystal. Only the contributions for the hydrogen atoms have been considered, in order to measure a local property. We also show the diffusion constant D of H within the Pd lattice. In the six panels we compare the performance of the path integral Langevin thermostat - both using a (L)ocal and (G)lobal thermostat on the centroid, and the optimal sampling GLE, fitted over frequency ranges which are 2, 3 and 4 orders of magnitude wide respectively. A white-noise Langevin reference is also shown.



H atom, which can be used to gauge the disturbance introduced by the thermostat on the most anharmonic motion present in this system.

We have used a supercell containing 256 palladium atoms and a single hydrogen. Given their large mass, it is safe to treat Pd atoms as classical particles, whereas a 10 bead polymer was used for the hydrogen atom. The thermostating of the classical

normalized correlation function,

$$\tau_A = \frac{1}{\langle A^2 \rangle - \langle A \rangle^2} \int_0^\infty |\langle A(t) - \langle A \rangle \rangle \langle A(0) - \langle A \rangle \rangle| dt$$

and PI dynamics has been performed separately and the same scheme has been used for both palladium and the centroid of the necklace.

With this setup, the velocity rescaling on the centroid corresponds to only 3 degrees of freedom. Nevertheless, we will see that there are significant differences from the case of a fully local method. For each choice of thermostat parameters, eight independent 1 ns-long trajectories have been performed, with a time step of 0.5 fs. The target temperature was set to 350 K, to ensure a significant diffusion during the simulated time span.

Results are reported in figure 4.4, and tell a story which is quite different from the case of liquid water. First of all, it is evident that global thermostats fail when the aim is to compute local properties in a harmonic lattice. In this particular case the rescaling was applied to the velocity of a single atom, but this problem can only be exacerbated if more hydrogen atoms are included and thermostatted together. An advantage of PILE-G is that, just as for water, it does not greatly disturb the diffusive modes. In contrast, conventional Langevin equation in the overdamped limit slows down diffusion by an order of magnitude.

As far as the local schemes are concerned, we first observe that, excepting the virial kinetic energy which depends strongly on the fast modes of the necklace, there is no clear advantage in using the local version of PILE over a plain white-noise Langevin equation. It is also interesting to look at the plot of τ_V for the optimal-sampling GLE thermostat. In this system very few normal modes contribute to the potential energy of the hydrogen atom. As such, a carefully-tuned white noise Langevin (or a narrow-range GLE) is advantageous because, as we discuss in figure 3.4, the price to pay for a constant sampling efficiency over a broad range of frequencies is a small degradation of the coupling with respect to a thermostat targeted on a single mode.

Concluding this chapter, and before discussing a class of non-equilibrium GLEs in which ensembles different from the canonical one are sampled, let us summarize our findings. The success of PILE when sampling the virial estimator of the kinetic temperature demonstrates that the conventional LE can be very effective, as long as it is applied directly to normal modes, and tuned to individual frequencies. This suggests that when atomic nuclei with very different masses are present, different frictions should be used to enhance the efficiency of sampling. However, this normal-modes Langevin strategy fails when an inexpensive prediction of vibrational properties is impossible, as is the case for the motion of the centroid in PIMD. We also confirmed that - as discussed in section 3.4 - Hamiltonian dynamics is very efficient in sampling diffusive modes, and global thermostating schemes which do not disturb this property are very effective in liquids and strongly anharmonic problems in general.

In the end, for the vast majority of cases, one would simply like to avoid the hassle and the computational effort associated with fine-tuning the parameters of the thermostat. From this point of view, the performance of the optimal-sampling generalized Langevin equation is impressive. The same parameters we have used in figure 3.14 to thermalize a classical MD simulation of liquid water have been used successfully for a PI dynamics of the same model, and even for the radically different example of a hydrogen atom diffusing in a Pd crystal. This approach has delivered close-to-optimal sampling efficiency in all cases and for all observables

tested, without any transformation to a normal mode representation.

If it is necessary - and worthwhile - to push the sampling efficiency further, our framework provides the flexibility required to fine-tune the GLE, using the different analytical estimates we introduced in chapter 3. One could imagine computing the vibrational spectrum for the different species using the velocity-velocity correlation function, and then fitting a narrow-band optimal sampling GLE to specifically target those regions with a high density of states. If a region of strongly anharmonic modes is present, the disturbance introduced by the thermostat can be minimized based on the estimate of Eq. (3.6). However, it is our opinion that, for most cases, the consistent performance on different observables, the transferability and the simplicity of the optimal sampling GLE largely offset the slight reduction in sampling efficiency.

Frequency-dependent thermostats

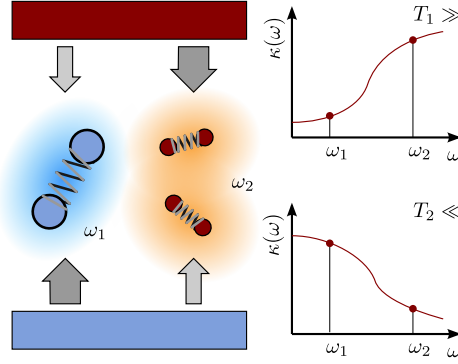
In previous chapters we have demonstrated how a generalized Langevin equation can be used to enhance the sampling in molecular dynamics. These enhancements were achieved by fine-tuning the coupling with different vibrational modes. To ensure that the configurations visited along the trajectories are consistent with a canonical statistics, we enforced the fluctuation-dissipation theorem, by means of the condition (2.7). In this chapter we will explore the possibilities which open up when this constraint is relaxed. In particular we will see how sampling the resulting, unusual statistical ensembles, provides tools for exciting selected normal modes and for modeling nuclear quantum effects with negligible computational overhead.

The prototypical simulation which breaks fluctuation-dissipation theorem is represented in figure 5.1. Here, two thermostats with different target temperatures, $T_1 \gg T_2$ and different coupling curves $\kappa(\omega)$ are applied to an ensemble of harmonic oscillators with different frequencies, ω_1 and ω_2 . This is clearly a non-equilibrium system, as energy will be constantly injected by the high-temperature bath, and removed by the low-temperature one. However, a steady state will be reached, with the different normal modes characterized by effective temperatures $T^*(\omega_1)$ and $T^*(\omega_2)$, which will depend on the relative strength of the coupling of the system with the two thermostats. To demonstrate that our framework can be adapted to this circumstance, let the two GLEs in the setup represented in figure 5.1 be characterized by the drift matrices ${}^1\mathbf{A}_p$ and ${}^2\mathbf{A}_p$ and the diffusion matrices ${}^1\mathbf{B}_p$ and ${}^2\mathbf{B}_p$. Then, the parameters of the overall GLE can be obtained using the combination rule given in Eq. (2.24), with unit weights $w_1 = w_2 = 1$. If in addition the two initial thermostats satisfy the fluctuation-dissipation theorem at temperatures T_1 and T_2 one can write

$$\mathbf{D}_p = k_B \begin{pmatrix} 2T_1 {}^1a_{pp} + 2T_2 {}^2a_{pp} & T_1 ({}^1\mathbf{a}_p^T + {}^1\bar{\mathbf{a}}_p) & T_2 ({}^2\mathbf{a}_p^T + {}^2\bar{\mathbf{a}}_p) \\ T_1 ({}^1\mathbf{a}_p^T + {}^1\bar{\mathbf{a}}_p) & T_1 ({}^1\mathbf{A} + {}^1\mathbf{A}^T) & \mathbf{0} \\ {}^2T_2 ({}^2\mathbf{a}_p^T + {}^2\bar{\mathbf{a}}_p) & \mathbf{0} & T_2 ({}^2\mathbf{A} + {}^2\mathbf{A}^T) \end{pmatrix}.$$

If $T_1 \neq T_2$ and the symmetric parts of ${}^1\mathbf{A}_p$ and ${}^2\mathbf{A}_p$ are non-zero, one cannot find

Figure 5.1: A cartoon representing a two-thermostat setup, which we take as the simplest example of a stochastic process violating the fluctuation-dissipation theorem. If the relaxation time versus frequency curves for the two thermostats are different, a steady-state will be reached in which the normal modes corresponding to different frequencies will equilibrate at different effective temperatures.



a temperature T such that $\mathbf{D}_p = k_B T (\mathbf{A}_p + \mathbf{A}_p^T)$. As a consequence, the stationary distribution cannot be expressed in the simple form given by Eq. (2.9)¹.

We will not discuss this two-thermostats example further, and instead consider the general case where (2.7) is not satisfied, and an arbitrary, positive-definite \mathbf{D}_p is selected. In appendix B we demonstrate that the dynamics of a GLE coupled with a harmonic potential with frequency ω has a stationary probability distribution provided that the symmetric part of \mathbf{A}_p is positive-definite. The resulting distribution is Gaussian, and is fully determined by an ω -dependent covariance matrix $\mathbf{C}_{qp}(\omega)$, which can be obtained by inserting the $\mathbf{A}_{qp}(\omega)$ and \mathbf{B}_{qp} matrices, which describe the Ornstein-Uhlenbeck process in the full (q, p, \mathbf{s}) space, into Eq. (D.1) (see section 2.2). The elements of the stationary covariance are rational functions of ω and of the elements of \mathbf{A}_p and \mathbf{D}_p . Not much can be said in general about their properties, besides the fact that $c_{qp}(\omega) = 0$, i.e. that p and q remain uncorrelated in the long-time limit. It is not even granted that an equipartition property holds, i.e. $\omega^2 c_{qq}(\omega) \neq c_{pp}(\omega)$. When speaking of an “effective kinetic temperature”, therefore, we refer to $T^*(\omega) = c_{pp}(\omega) / k_B$.

By tuning the parameters of the generalized Langevin equation it is possible to change at will the ω -dependence of $c_{qq}(\omega)$ and $c_{pp}(\omega)$, and to fit them to the desired target function by means of terms of the form (2.20). Before going on to describe the useful properties which can be imbued into a MD simulation using this strategy, a word of caution is necessary. The inevitable deviations from the harmonic-limit predictions affect the *efficiency* of sampling and the properties of the dynamics. When (2.7) holds the *static* properties at least are well-defined, since the

¹It is also possible to consider this problem in the non-Markovian representation. In that case, one can verify that the overall memory kernels will be $K(t) = K_1(t) + K_2(t)$ and $H(t) = H_1(t) + H_2(t) = k_B (T_1 K_1(t) + T_2 K_2(t))$. If $K_1 = K_2$, the resulting sampling for (q, p) will be canonical at temperature $(T_1 + T_2)/2$, even if Eq. (2.7) does not hold. The \mathbf{s} variables meanwhile will be distributed in a more complicated manner than they are in Eq. (2.9).

stationary distribution is (2.9) regardless of the potential. However, in the applications discussed in the present chapter, we were not able to find an analytical expression for the stationary distribution in the presence of an anharmonic potential. Therefore, the static properties can deviate significantly from the predictions, which means that one must be more careful and resort to empirical testing for assessing the reliability of the results. We will demonstrate that, despite these caveats, non-equilibrium GLEs can provide very useful applications, ranging from selective normal-modes excitation to inexpensive simulation of nuclear quantum effects.

5.1 The δ -thermostat

In section 2.1.3 we discussed a strategy that can be used to obtain Dirac- δ -like memory kernels, while in figure 3.10 we demonstrated how the disturbance on the dynamics induced by a colored-noise thermostat can be restricted to a narrow range of frequencies. The effectiveness of the thermostat in affecting few normal modes in a highly selective way demonstrates the flexibility of our scheme. In addition, it almost invariably implies that the frequency dependencies of the desired properties can be made arbitrarily complex. Hence, as a first example of a non-equilibrium GLE thermostat, we present an application in which the thermostat induces large fluctuations for the modes within a narrow frequency range, while ensuring that all the others remain cold.

In order to realize this δ -thermostat idea, we return to the δ -like memory kernels described in section 2.1.3. The power spectrum of the noise will be dominated by frequencies close to ω_0 . Intuitively, the thermostat would be expected to couple more strongly to the normal modes with a frequency close to ω_0 . If one imposes a constant friction on top of the matrices in Eq. 2.13 (i.e. by changing the a_{pp} element from zero to a finite value γ' , without changing \mathbf{D}_p)

$$\mathbf{A}_p = \begin{pmatrix} \gamma' & \sqrt{\gamma/2\pi} & 0 \\ -\sqrt{\gamma/2\pi} & \Delta\omega & \omega_0 \\ 0 & -\omega_0 & 0 \end{pmatrix}, \quad \mathbf{D}_p = k_B T \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2\Delta\omega & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (5.1)$$

one would expect modes far away from the peak frequency to be affected by the uniform friction, and to thus freeze almost completely. Modes close to ω_0 , instead, would equilibrate at a much larger average temperature. Note that changing \mathbf{A}_p without performing the changes to \mathbf{D}_p prescribed by Eq. (2.14) amounts to breaking the classical FDT.

For a given peak frequency ω_0 one is left with four parameters which can be optimized ($\Delta\omega$, γ , γ' , T), so as to obtain the most convenient set of calculated response parameters as a function of the oscillator frequency ω (see figure 5.2). Even with this relatively simple model, satisfactory results can be achieved, with fluctuations in positions and momenta which correspond to a desired effective temperature in a neighborhood of ω_0 , and drop quickly to a very low value elsewhere. However, the dependency of the temperature profile $T^*(\omega)$ on the memory kernels is not trivial. For example, reducing $\Delta\omega$ is not sufficient to obtain a sharper peak in the normal-modes

Figure 5.2: Effective temperature $T^*(\omega)$ as a function of ω and $\Delta\omega$ for three different choices of the friction parameters γ and γ' (see Eq. (5.1)). Note that as the coupling parameters are decreased, the sampling efficiency degrades. Therefore, it is wise to perform a fitting procedure in which the profiles of $c_{pp}(\omega)$ and $\kappa_V(\omega)$ are both monitored.

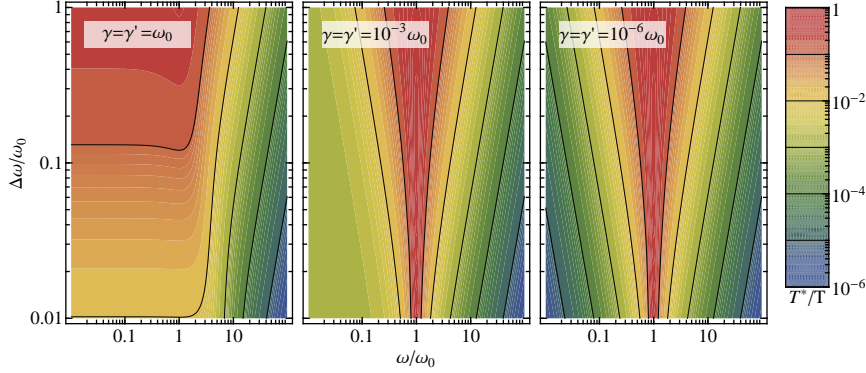
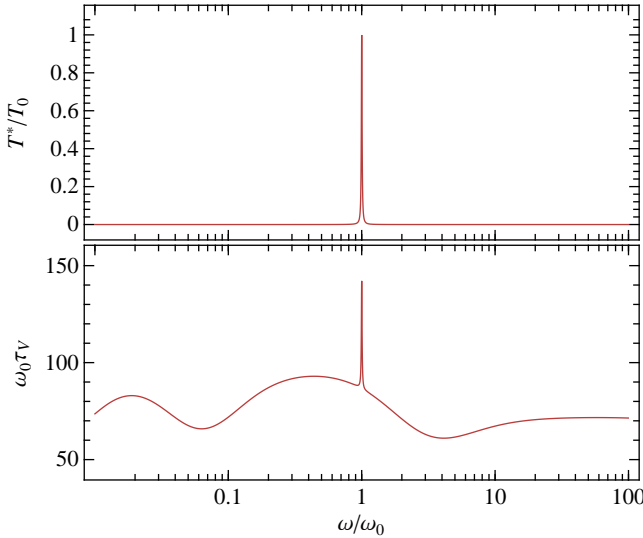


Figure 5.3: Effective temperature $T^*(\omega)$ and correlation time of potential energy $\tau_V(\omega)$, for the dynamics of a harmonic oscillator of frequency ω coupled to a δ -thermostat with peak frequency ω_0 and peak temperature T_0 .

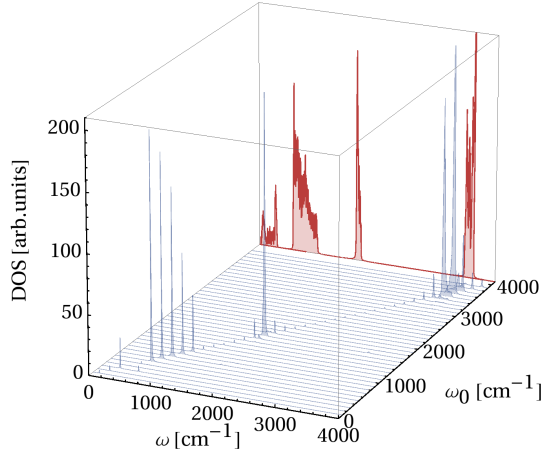


response.

One would like to find the best possible compromise between the sharpness of the response $T^*(\omega)$ and the efficiency of sampling in the neighborhood of ω_0 . For this reason we resorted to a fitting procedure, in which we enforce $T^*(\omega_0) = T_0$ and that $T^*(\omega)$ is small elsewhere, while we aiming for a low $\tau_V(\omega)$ in a neighborhood of ω_0 . We have selected a satisfactory set of parameters, which has the response properties described in figure 5.3, to be used for the examples discussed in this section.

By simply scaling the drift and diffusion matrix, the parameters depicted in Figure 5.2 can be adapted to any peak frequency ω_0 and any effective temperature

Figure 5.4: The vibrational density of states for ice *Ih* at 10 K (thick line, fully-shaded in red), as computed from the Fourier transform of the velocity-velocity correlation function obtained from a molecular-dynamics simulation of a flexible-water model[59]. This is compared to a stacked plot of a series of δ -thermostat runs, with different peak excitation frequencies ω_0 .



T_0 (see section 2.4). In fact, for a set of matrices \mathbf{A}_p and \mathbf{D}_p , which yield an effective temperature versus frequency curve $T^*(\omega)$ peaked around $\omega_0 = 1$, one can prove that any GLE which uses the scaled drift matrix $\omega_0 \mathbf{A}_p$ and an unscaled \mathbf{D}_p matrix will produce the response curve

$$T^*(\omega, \omega_0) = T^*(\omega/\omega_0)/\omega_0 \quad (5.2)$$

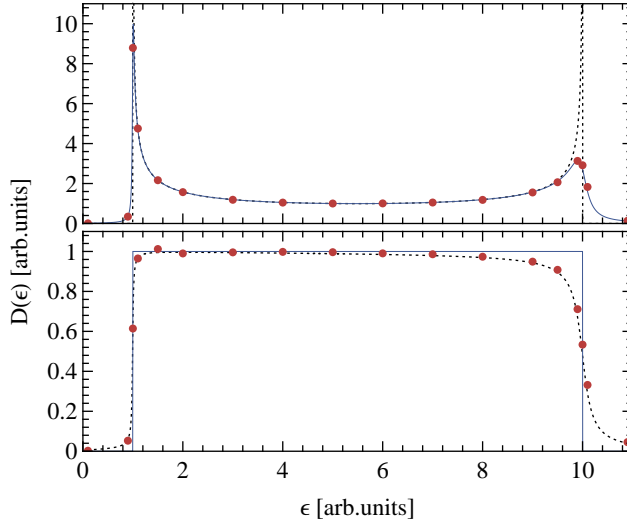
which is peaked around ω_0 and has the same peak integral for all values of ω_0 .

5.1.1 Selective phonon heating in ice *Ih*

As a first example of this δ -thermostat we describe its effect on the molecular dynamics of a solid. To this end, we performed simulations using a flexible water force field[59], and ran several trajectories scaling the parameters of the thermostat to different peak frequencies. Throughout, we have scaled the matrices so as to obtain a constant peak temperature $T_0 = 10$ K, which is sufficiently low enough to limit the effects due to anharmonicities.

In Figure 5.4 we report the vibrational density of states computed from the Fourier transform of the velocity-velocity correlation function for the different trajectories, and compare it with the vibrational spectrum computed from constant-energy dynamics for a sample equilibrated at 10 K. The δ -thermostat succeeds in selectively exciting a narrow band of normal modes. Interestingly, even at this low temperature, anharmonic effects cause slight deviations from the expected behavior. In particular, an excitation is also observed when the peak frequency falls within a vibrational gap of the intrinsic phonon density of states for the ice model. However, this excitation is

Figure 5.5: The eigenvalues density $D(\epsilon)$ as computed from a δ -thermostat dynamics, for a matrix with uniformly-distributed eigenvalues between 1 and 10 (lower panel) and for a matrix corresponding to a linear chain of masses connected by springs, with periodic boundaries (upper panel). In both cases, a 1000×1000 matrix has been used for \mathbf{H} , and $4 \cdot 10^6$ timesteps of 0.05 time units each have been performed for each run. The exact DOS (dashed line) is compared with an analytical predictions for $\tilde{T}(\omega_0)$ (continuous line) and with the average kinetic temperatures computed from actual runs (dots). For the dots, the error bar is smaller than the size of the points.



a couple of orders of magnitude smaller than the ones observed when proper normal modes are present.

5.1.2 δ -thermostat dynamics as an eigenvalue solver

The possibility to excite normal modes within a narrow frequency band, without any explicit knowledge of the vibrational spectrum of the system, suggests an application of the δ -thermostat as a probe of the spectrum of a positive-definite matrix \mathbf{H} of size N . This would work treating \mathbf{H} as the Hessian of a multi-dimensional harmonic potential, on which a dynamics is to be performed. This method would be particularly useful whenever matrix-vector multiplications are considerably less expensive than explicit diagonalization. This is the case for sparse matrices, or when the Hamiltonian has a favorable structure[47].

To demonstrate this method, we used \mathbf{H} as the Hessian of a multidimensional, harmonic potential, $V(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x}$, and implemented an artificial stochastic dynamics based on the GLE parameters described in figure 5.3. Contrary to the ice case, no anharmonicities are present here, so the equilibrium distribution will be the one predicted analytically, except for negligible finite-timestep errors.

There are many possible applications for this technique. For instance, one could search for eigenvalues within a preassigned frequency range[79]. Also, by performing several runs with a different choice of ω_0 it is possible to obtain an estimate of the

whole density of states (DOS). With this in mind, we will assume that the eigenvalues ϵ_i of \mathbf{H} are known, and show how the overall average kinetic temperature, obtained by a δ -thermostat trajectory, is related to the DOS.

The frequencies of the normal modes of the artificial dynamics will be $\omega_i = \sqrt{\epsilon_i}$, while the average temperature is just

$$\bar{T}(\omega_0) = \frac{1}{N} \sum_i T^*(\omega_i, \omega_0) \quad (5.3)$$

where $T^*(\omega_i, \omega_0)$ is the scaled temperature profile (5.2). By construction, $T^*(\omega_i, \omega_0) \sim \delta(\omega_0 - \omega_i)$, which means that the connection between (5.3) and the density of states (DOS) of the eigenvalue spectrum of \mathbf{H} is readily found:

$$D(\epsilon) = \sum_i \delta(\epsilon - \epsilon_i) \sim \bar{T}(\sqrt{\epsilon}) / \sqrt{\epsilon}. \quad (5.4)$$

We will now consider, as a simple illustration, how the spectrum of two simple matrices can be analyzed. We considered a matrix with evenly-spaced eigenvalues between 1 and 10, and one corresponding to the dynamical matrix of a linear chain of masses connected by springs, with periodic boundary conditions and an on-site term. In this second case, the eigenvalues are $\epsilon_i = \alpha - 2\beta \cos(2\pi i / N)$, where N is the dimensionality of the Hessian and α and β are chosen so that the ϵ_i 's lie between 1 and 10.

The results for these two problems are reported in Figure 5.5. The actual outcome for the dynamics is identical to the predictions one can make using Eq. (5.3), which demonstrates that the errors due to the finite time-step in the integration are negligible. The distribution of the eigenvalues is quantitatively reproduced, except for the points where the actual DOS has features that are sharper than the resolution of the δ -thermostat. It is worth noting that the scaling procedure we used to shift the fitted parameters to different peak frequencies guarantees constant *relative* resolution, i.e. it ensures the breadth of the peak is a constant fraction of ω_0 . This explains the asymmetry between the two peaks of the DOS in the linear chain panel of Figure 5.5.

The resolution can be further improved, but this comes at the expense of a lower sampling efficiency. The performance of this approach compared to traditional methods depends on the size of the matrix, its condition number and the desired resolution and accuracy. As is the case for any stochastic method, the statistical error decreases slowly, with the square root of the sampling time. We can therefore expect our approach to be more useful as a qualitative probe of the eigenvalue spectrum than as a tool for obtaining very precise estimates of the density of states.

5.2 The quantum thermostat

The δ -thermostat we described in the previous section can be used in many applications, but in this thesis it has been presented mainly as a proof of concept that demonstrates how a non-equilibrium GLE is capable to enforce a frequency-dependent distribution in a quasi-harmonic system. In this section we will discuss

another application of the GLE thermostat. This is probably the most exciting application we have developed thus far, as it allows one to include an approximate treatment of the quantum nature of ionic cores in conventional, classical molecular dynamics with practically no additional computational overhead.

Nuclear quantum effects are extremely important in many condensed-phase systems. For instance, zero-point fluctuations affect static correlations, while energy quantization causes deviations from the classical value of the specific heat at low temperatures. A quantum treatment of the ionic degrees of freedom is mandatory for capturing these effects, which are particularly important when light atoms or stiff vibrational modes are present. Conventional techniques for including quantum effects are computationally demanding and for this reason the nuclei in molecular simulations are often treated classically, even when the electronic degrees of freedom are accounted for quantum-mechanically [48].

When exchange-symmetry effects are not relevant, the method of choice for studying equilibrium expectation values is path-integrals molecular dynamics (PIMD) [80, 81], which was described in chapter 4. This approach involves a high computational cost, as many replicas of the system must be simulated in parallel. Approximate but less expensive methods such as Feynman-Hibbs effective potentials [72, 82] and semiclassical approaches to treat zero-point energy (ZPE) have also been used [83–85]. However, their range of validity is limited to weak quantum behavior and to cases where the Hessian of the potential is available and cheap to compute. As a result, the interest in methods to introduce quantum effects in classical trajectories is still very high, see e.g. Ref. [7, 86].

A recently-developed, promising strategy to reduce the overhead of PI calculations [77] works by computing the long-range contribution to the electrostatic interactions on an appropriate contraction of the ring polymer, rather than computing this term for each individual replica. For the majority of classical force fields the evaluation of these terms is by far the most expensive part of the calculation, and so this method leads to substantial computational savings. However, these techniques are only helpful whenever if it is possible to separate an expensive, long-range interaction from a simpler, short-ranged one. In density functional simulations, for instance, this is not possible. Therefore, *ab initio* molecular dynamics is beyond the scope of these optimized PI methods.

In this section we will present a radically different approach. We will show how, by constructing an appropriate GLE thermostat, one can obtain an accurate approximation for nuclear quantum effects in a variety of systems, without any limitation on the kind of potential energy surfaces which can be treated. The idea of using a non-Markovian Langevin equation to model quantum systems can be traced back to the semiclassical approximation to the quantum Langevin equation [25, 87, 88]. This equation has been used to describe quantum systems in contact with a quantum harmonic bath. A similar idea has been recently used by Buyukdagli *et. al.* [26], in order to compute quantum specific heat in harmonic systems. However, even for a harmonic oscillator, Buyukdagli's scheme is only qualitatively correct and neglects ZPE completely. Our approach is more general, allowing us to obtain high accuracy in reproducing quantum specific heats, and also to tackle the more challenging task

of introducing zero-point motion effects[89]. We will also demonstrate how very anharmonic condensed-phase problems can be treated with good accuracy, by carefully choosing the GLE parameters.

5.2.1 The quantum harmonic oscillator

To develop a robust approach for introducing nuclear quantum effects, by means of an appropriate generalized Langevin equation, we first consider the harmonic oscillator as an exact treatment is available in both the quantum-mechanical and SDE frameworks. The finite-temperature density matrix for a quantum harmonic oscillator of frequency ω at temperature T reads

$$\begin{aligned} \rho(q, q') &\propto \sum_n e^{-\epsilon_n/k_B T} \psi_n^*(q) \psi_n(q') \\ &\propto \exp \left[-\omega^2 \left(\frac{(q+q')^2/4\hbar\omega}{\coth \hbar\omega/2k_B T} + \frac{(q-q')^2/4\hbar\omega}{\tanh \hbar\omega/2k_B T} \right) \right], \end{aligned}$$

where ϵ_n and ψ_n are the eigenvalues and eigenvectors of the harmonic Hamiltonian. For any temperature, the density is a Gaussian, whose width depends on ω and T

$$\rho(q) = \rho(q, q) \propto \exp \left[-\frac{1}{2} \omega^2 q^2 \left(\frac{\hbar\omega}{2} \coth \frac{\hbar\omega}{2k_B T} \right)^{-1} \right].$$

For a classical oscillator the probability distribution of the displacement q is also a Gaussian, $P(q) \propto \exp(-\omega^2 q^2/2k_B T)$. It is therefore possible to reproduce the stationary distribution of q , in a purely classical framework, by performing canonical sampling using an ω -dependent target temperature

$$T^* = \frac{\hbar\omega}{2k_B} \coth \frac{\hbar\omega}{2k_B T}. \quad (5.5)$$

The distribution of momenta, which can be obtained by writing the density matrix in the momentum representation, is also Gaussian, and has a width consistent with the classical distribution at the effective temperature T^* .

In passing, it is interesting to note that this isomorphism has consequences well beyond the marginal probabilities of the phase-space distribution. Consider any observable which depends locally on positions only. Its finite-temperature expectation value can be written in the position representation in terms of the one-particle density

$$\langle \mathcal{A} \rangle = \frac{\sum_n e^{-\epsilon_n/k_B T} \int \mathcal{A}(q) |\psi_n(q)|^2 dq}{\sum_n e^{-\epsilon_n/k_B T}} = \int \mathcal{A}(q) \rho(q) dq.$$

The same expectation value can be computed *exactly* as the ensemble average of $\mathcal{A}(q)$ over a classical trajectory at the effective temperature (5.5).

Computing free-energy differences It is straightforward to show that quantum free-energy differences can be computed with conventional methods, by simply

performing classical trajectories with a frequency-dependent effective temperature. For example, to compute these differences using thermodynamic integration[10], one must introduce an order parameter λ , which when varied smoothly describes the transformation between the initial and final states. The free-energy difference is then obtained by integrating the thermodynamic force

$$\frac{\partial F(\lambda)}{\partial \lambda} = -k_B T \frac{\partial \log Q}{\partial \lambda} = \frac{\text{Tr}(e^{-\beta \mathcal{H}} \partial \mathcal{H} / \partial \lambda)}{\text{Tr} e^{-\beta \mathcal{H}}}$$

along λ . When \mathcal{H} depends on λ by through the potential only, one can write the thermodynamic force in terms of a Boltzmann-weighted combination of expectation values computed on the eigenstates of the Hamiltonian

$$\frac{\partial F(\lambda)}{\partial \lambda} = \frac{\sum_i e^{-\epsilon_i / k_B T} \langle \psi_i | \partial \mathcal{H} / \partial \lambda | \psi_i \rangle}{\sum_i e^{-\epsilon_i / k_B T}} = \int \frac{\partial V(x, \lambda)}{\partial \lambda} \rho(x) dx.$$

This equation implies that $\partial F(\lambda) / \partial \lambda$, and hence the correct free-energy difference, can be computed without explicit knowledge of the eigenstates. It is instead sufficient to reproduce the finite-temperature density only. Similar derivations can be carried out for other free-energy methods.

Fitting a quantum thermostat To turn the isomorphism described above in a viable method for computing nuclear quantum effects, one must be able to enforce the ω -dependent effective temperature (5.5) simultaneously on different normal modes, without any explicit knowledge of phonon spectrum and displacement patterns. We know already that this goal can be achieved using a generalized Langevin equation, provided that the fluctuations computed for an individual harmonic oscillator, as a function of its frequency, are consistent with Eq. (5.5).

To enforce this behavior we compute the stationary distribution in the harmonic limit, and fit the fluctuations $c_{pp}(\omega)$ and $c_{qq}(\omega)$ to the target functions $\tilde{c}_{pp}(\omega) = \omega^2 \tilde{c}_{qq}(\omega) = \frac{\hbar \omega}{2} \coth \frac{\hbar \omega}{2 k_B T}$. These conditions must be imposed separately, since there is no guarantee that a non-equilibrium GLE will satisfy equipartition. Overall, the two conditions result in a term

$$\chi_{qt} = \left[\sum_i \left| \log \frac{c_{qq}(\omega_i)}{\tilde{c}_{qq}(\omega_i)} \right|^m + \left| \log \frac{c_{pp}(\omega_i)}{\tilde{c}_{pp}(\omega_i)} \right|^m \right]^{1/m}, \quad (5.6)$$

that must be included in the merit function for the fit. The $\omega \rightarrow 0$, classical limit ($c_{pp}(\omega) = \omega^2 c_{qq}(\omega) = k_B T$) can be proven to correspond to two conditions on the elements of the free-particle covariance matrix \mathbf{C}_p ; namely, $c_{pp} = k_B T$ and $\mathbf{a}_p^T \mathbf{A}^{-1} \mathbf{c}_p = 0$. One could enforce these constraints exactly, by treating the entries of \mathbf{C}_p as independent fitting parameters and obtaining the diffusion matrix from Eq. (2.5). We found however that this choice makes it difficult to obtain a positive-definite $\mathbf{B}_p \mathbf{B}_p^T$, and that the fitting becomes more complex and inefficient.

As an alternative, we decided to enforce the low-frequency limit with an appro-

priate penalty function included in the fit

$$\chi_{cl} = (c_{pp}/k_B T - 1)^2 + \left(\mathbf{a}_p^T \mathbf{A}^{-1} \mathbf{c}_p / k_B T \right)^2. \quad (5.7)$$

Transferability of fitted parameters The terms discussed in the previous paragraph must be included in the fit, along with others (which will be discussed in section 5.2.3) that control the strength of coupling and the efficiency of sampling. All these terms, must then be fitted over a range of frequencies which encompasses all the relevant vibrational modes. This range should start at the onset of quantum effects and finish with the largest frequency present, ω_{max} .

In practice one does not need to perform a separate fit for each system and temperature. Instead, it is sufficient to prepare a library of matrices fitted using $\omega_{max} = 1$, $k_B T = 1$ and a number of different values for the dimensionless parameter $\varphi = \hbar \omega_{max} / k_B T$. The resulting GLE matrices \mathbf{A}_p and \mathbf{C}_p can then be adapted to a simulation at temperature T by means of the transformation rules described in section 2.4. In this case, the appropriate scalings read $\mathbf{A}_p \leftarrow \mathbf{A}_p k_B T \varphi / \hbar$ and $\mathbf{C}_p \leftarrow \mathbf{C}_p k_B T$. The transformed SDE will enforce quantum-corrected fluctuations up to the frequency $\omega_{max} = k_B T \varphi / \hbar$. Hence, a library of few matrices, optimized for different values of φ , can be adapted to any ensemble of harmonic frequencies and any temperature².

5.2.2 One-dimensional anharmonic potentials

Besides negligible, finite-time step errors, the distribution of q and p , as well as the average values for physically-relevant observables such as the potential and kinetic energy will correspond to the analytical predictions for a given set of GLE parameters. These parameters can in turn be fitted to quantum-mechanical results within few percent. Many physical problems are well-described in terms of small perturbations on top of a quadratic potential. Therefore we should first demonstrate that our method can provide a reasonable approximation of quantum expectation values for anharmonic potentials. Consequentially, our first benchmarks are some one-dimensional examples - the stronger the anharmonicity, the more challenging the test.

Quasi-harmonic potential The first case we consider is a one-dimensional, anharmonic potential of the form

$$V(x) = \frac{\omega^2}{2} x^2 \frac{1 - e^{-kx}}{kx} \quad (5.8)$$

For fixed values of k and the temperature T , this potential allows one to explore a range of behaviors ranging from the highly anharmonic, classical limit at $\omega \rightarrow 0$ to

²As we will discuss later, the larger the value of φ , the more difficult the fitting procedure, and the larger the deviations from the expected behavior in anharmonic cases. In practice, values as high as $\varphi = 50$ and even 100 can be obtained, with an error between a few percents and 10%, depending on the details of the fit.

Figure 5.6: In the left panel, the one-dimensional anharmonic potential (5.8) is plotted for $k = 1 \text{ \AA}^{-1}$ and different values of the frequency ω . To give physically-relevant conditions, we considered the mass to be that of a hydrogen atom. For a target temperature $T = 100 \text{ K}$ the system will, depending on the frequency, sample either a nearly classical but highly anharmonic regime, a strongly quantum but nearly harmonic region for stiff vibrations, or an intermediate situation where both anharmonic and quantum effects are relevant. In the rightmost panel, we plot the characteristics of the GLE parameters used in these one-dimensional examples. The dashed black line corresponds to the correlation time of the kinetic energy (see section 5.2.3 for a discussion of its relevance), the blue and red lines to the predicted fluctuations of position and momentum respectively.

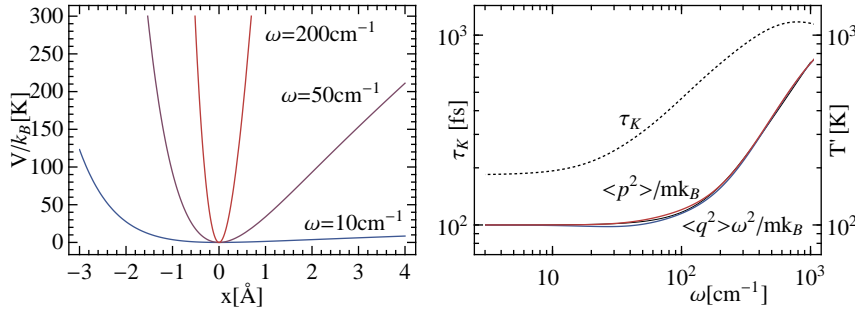
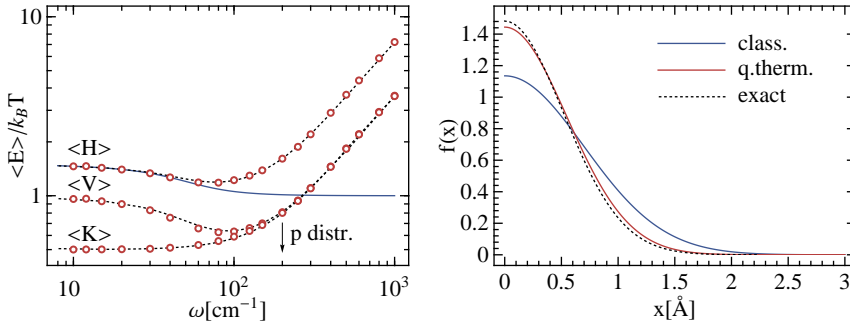


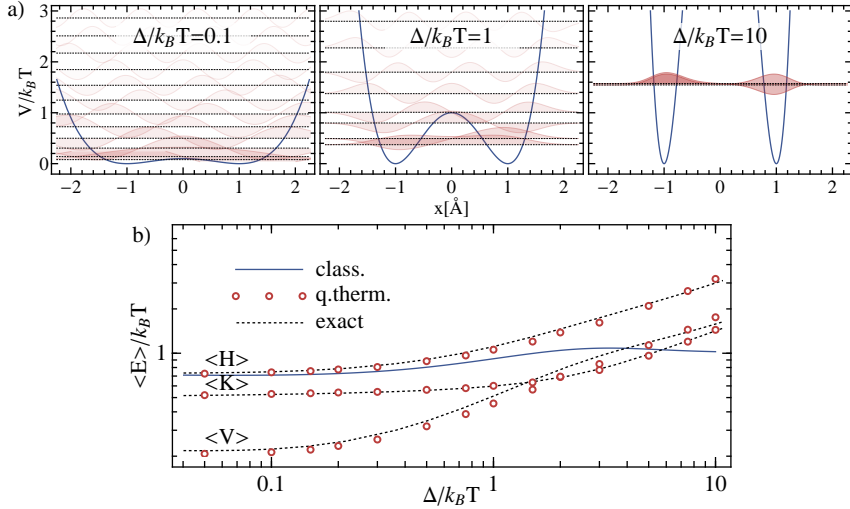
Figure 5.7: In the left panel, the average total ($\langle H \rangle$), potential ($\langle V \rangle$) and kinetic ($\langle K \rangle$) energies for a proton in the external potential of Eq. (5.8) are plotted as a function of ω , for $k = 1 \text{ \AA}^{-1}$ and $T = 100 \text{ K}$. Dashed black lines correspond to the exact quantum solution, the red series are results from quantum-thermostatted dynamics, using a set of parameters fitted over a frequency range from 2 to 2000 cm^{-1} . Blue lines correspond to the classical expectation values. In the right panel the Fourier transform of the momentum distribution is reported for the fully quantum mechanical, classical and colored-noise thermostat simulations at $\omega = 200 \text{ cm}^{-1}$.



a quantum, harmonic regime at high ω (see figure 5.6). In Figure 5.7 we compare the exact, quantum solution with the averages obtained using our colored-noise thermostat.

There is a remarkably good quantitative agreement not only in the asymptotic $\omega \rightarrow 0$ and $\omega \rightarrow \infty$ limits, but also in the intermediate region, where both quantum-mechanical and anharmonic effects are significant, which suggests that both effects can be captured, albeit not fully. The momentum distribution is also in good agreement with its quantum-mechanical counterpart. This is particularly appealing, since

Figure 5.8: (a) Schematic representation of the energy levels for the double-well potential (5.9), using three different values of the barrier Δ . The distance between minima was set to $d = 1 \text{ \AA}^{-1}$ and a particle with the mass of a proton was used. Wavefunctions are shaded according to their thermal occupation. (b) The average total ($\langle H \rangle$), potential ($\langle V \rangle$) and kinetic ($\langle K \rangle$) energies for a proton in the double-well potential as a function of Δ , for $d = 1 \text{ \AA}^{-1}$. Dashed black lines correspond to the exact quantum solution, while the red series are the results from quantum-thermostatted dynamics, using a set of parameters fitted over a frequency range from 2 to 2000 cm^{-1} . Blue lines correspond to the classical expectation values.



conventional PIMD can only sample positions, and special procedures must be introduced if momenta are to be sampled as well [73, 90, 91].

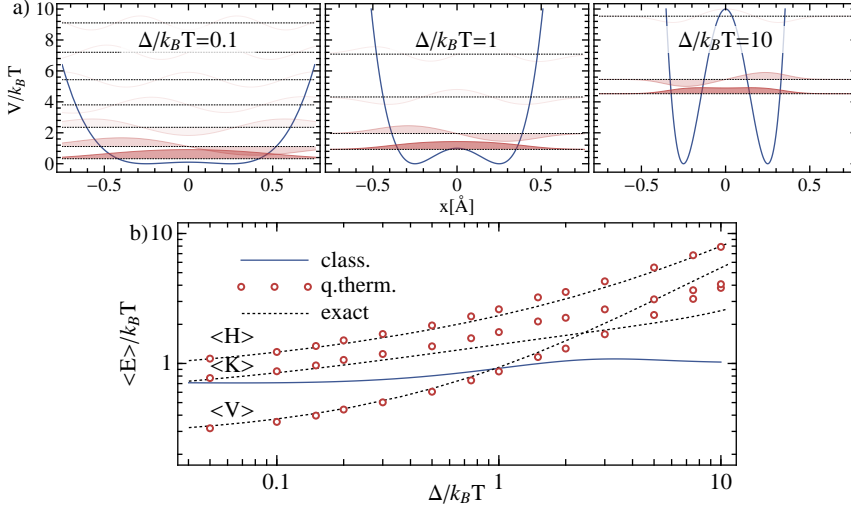
Double-well potential Moving on to a more challenging example, we considered a quartic double well potential,

$$V(x) = \Delta \left(1 + \frac{x^2}{d^2} \left(\frac{x^2}{d^2} - 2 \right) \right) \quad (5.9)$$

with a spacing of $2d$ between the two minima, and a barrier of height Δ . This potential is strongly anharmonic both for small Δ 's, when it reduces to a flat-bottom quartic potential, and for intermediate values, where there is a significant coupling between the bound states in the two wells (see figure 5.8(a)). For large Δ , the anharmonicity depends on the distance between the minima as well as on the curvature of the potential. Each of the two wells is locally quasi-harmonic, and hence the main anharmonic effect is the splitting of the energy levels due to tunnelling. In spite of this complex behavior, the results in figure 5.8(b) are in good agreement with exact, quantum-mechanical expectation values, and capture their qualitative features, including the crossover between $\langle V \rangle$ and $\langle K \rangle$.

Wavefunction-phase effects are completely absent in our model. Nonetheless our model suffers no catastrophic failures even in cases for which tunneling through

Figure 5.9: (a) Schematic representation of the energy levels for the double-well potential (5.9), using three different values of the barrier Δ , $d = 0.25 \text{ \AA}^{-1}$ and a particle with the mass of a proton. (b) The average total ($\langle H \rangle$), potential ($\langle V \rangle$) and kinetic ($\langle K \rangle$) energies for a proton in the double-well potential as a function of Δ , for $d = 0.25 \text{ \AA}^{-1}$. See figure 5.8 for an example with larger separation between the minima.



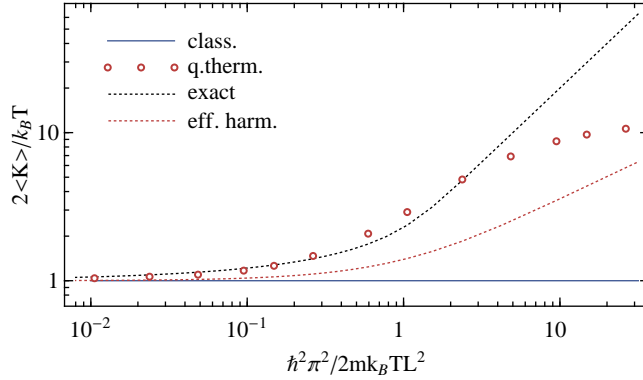
the barrier would have a sizable effect on energetics. Transitions between the two wells are more frequent than in purely classical dynamics at the same temperature due to zero-point energy³. However, one cannot compute dynamical properties of any sort, as they are heavily disturbed by the stochastic term.

In figure 5.8 we report analogous results for a much smaller distance between the two minima. This results in a more pronounced coupling between the two wells, and hence much stronger quantum effects. Here, the agreement between the results of our method and the exact quantum-mechanical expectation values is less satisfactory, particularly for the largest values of Δ . It is interesting to observe that there is a cancellation of errors between the overestimation of $\langle K \rangle$ and the underestimation of $\langle V \rangle$, which leads to a good agreement for the total energy $\langle H \rangle$. This is a consequence of our method's poor description of tunneling. Quantum tunneling arises because the system reduces its kinetic energy by smoothening the curvature of the wavefunction. This in turn compensates for the penetration of probability density into regions where the potential is large. In contrast, in the quantum thermostat simulation, the particle is more confined. Therefore, it has a larger kinetic energy and, because it does not penetrate the barrier, a smaller average for $\langle V \rangle$.

So far, we were unable to develop a satisfactory formal treatment for the behavior of non-equilibrium GLEs with anharmonic potentials. It is therefore difficult to quantitatively explain the successes and failures of our approach. Based on the

³We recall, in passing, that in recent years several attempts have been made to exploit the properties of quantum dynamics to enhance sampling in the context of simulated annealing minimization. This has led to the development of so-called *quantum annealing* methods[28, 92–94]. In these cases it is not necessary to reproduce quantum properties exactly, therefore it would be interesting to test whether or not our quantum thermostat can be used in this context rather than more expensive path integral methods.

Figure 5.10: Average kinetic energy computed for an infinite square-well potential, as a function of the dimensionless parameter $\hbar^2 / (8mL^2 k_B T)$ (the ratio between the zero-point and thermal energy). Dashed black lines correspond to the exact quantum solution, the red series are results from quantum-thermostatted dynamics, using a set of parameters fitted over a frequency range from 2 to 2000 cm^{-1} . Blue lines correspond to the classical expectation values. The dashed red line corresponds to the result which would be expected for an “effective” harmonic oscillator (see text).



qualitative arguments discussed above, however, one can expect a good agreement whenever the splitting between the levels in the two wells is smaller than either $k_B T$ or the zero-point energy, as in these cases tunneling is weak, and hence its effect on energetics is small.

Particle in a box As a final example, we tried to reproduce the motion of a particle in an infinite square-well potential. This is an extreme anharmonic case, as locally the particle experiences no force and hence the only “information” that can influence the response of the SDE is the reversal of velocity which occurs when one of the walls is reached. Despite these unfavorable conditions, the average kinetic energy is reproduced to a decent accuracy up to strongly quantum conditions (figure 5.10)⁴. It should be noted, however, that the histogram of positions is very different from the exact finite-temperature density. On the plus side it is not constant, and the probability of finding the particle is slightly higher at the center of the well. On the down side, though, it does not get close to zero at the boundaries. We could not expect much more from a method based on a harmonic reference, and note that even path integral methods require special care in the presence of hard walls[95].

The fact that we observe some effect is remarkable in itself, if one considers that other approximate techniques such as the Feynman-Hibbs method require the local curvature of the potential energy surface explicitly, and would still fail to realize that the particle is confined. One could speculate that the periodic reversal of velocities leads the GLE to respond as if it were applied to an “effective” harmonic oscillator, with a frequency equal to the average frequency of the reversal of velocity⁵.

⁴To give some physical reference, the parameter used on the x -axis in figure 5.10 takes on a value of one for a proton in a 1 Å-wide box potential at $T \approx 240$ K. We can therefore expect to be able to simulate reliably a light molecule confined in a pore or a nanotube.

⁵One can compute this effective frequency $\bar{\omega}$ from the mean thermal velocity and the width L of the

In actuality, the quantum thermostat performs better than it would be expected if this argument were the case. This provides further indication that it can provide a qualitative description of nuclear quantum effects even for strongly anharmonic problems.

5.2.3 Zero-point energy leakage

We have demonstrated that a quantum thermostat can provide a reasonable estimate of important physical observables also for very anharmonic, one-dimensional cases. One would thus expect - given the rotational invariance properties described in section 2.2 - that applications to quasi-harmonic multidimensional problems pose only a little problem. While this is certainly true for perfectly harmonic systems, one must be aware of the so-called zero-point energy (ZPE) leakage which plagues semi-classical approaches to the computation of nuclear quantum effects[83, 96]. In these techniques, particle momenta are initialized based on the quantum-mechanical distribution consistent with the local curvature of the potential energy surface. The system is then evolved by microcanonical dynamics. Unfortunately this means that, from a classical viewpoint, large differences in the kinetic energies of soft and stiff modes are introduced. Therefore, even a minute phonon-phonon coupling leads to energy exchange and to equilibration of the whole system at an intermediate temperature.

This effect is likely to affect our quantum thermostat, and any frequency-dependent thermostat when it is applied to a multidimensional, quasi-harmonic system. When this effect is included, the different modes will reach a steady-state configuration with an effective temperature which is still frequency-dependent, but which deviates significantly from the values predicted in the limit of perfect decoupling. In order to examine the magnitude of this error, we tested the method on a prototype quasi-harmonic problem.

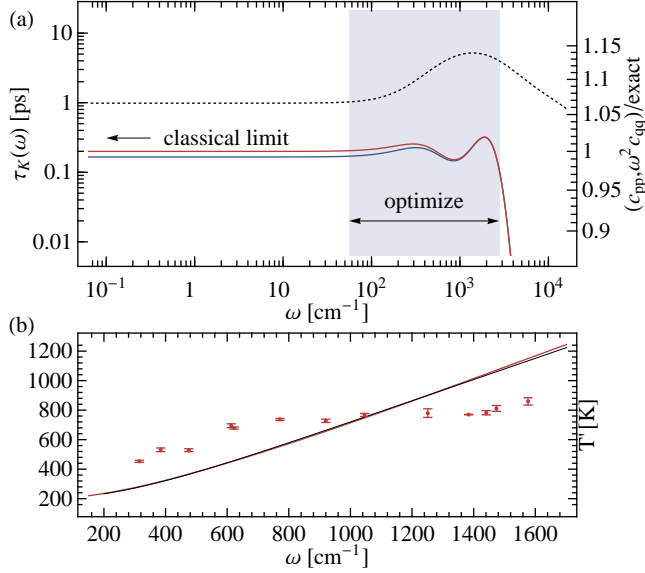
Diamond: a test case The ultimate test for assessing the accuracy of the quantum thermostat is a comparison with a path-integral calculation on a similar but computationally cheaper model, such as a smaller-size box or a simpler force field. However, it would be desirable to obtain some qualitative measure of the quality of the fit, and to gauge the transferability of a given set of parameters without having to perform a time-consuming comparison with path-integral methods.

To this end, we first tested the quantum thermostat on a Tersoff model[97] of diamond at a temperature $T = 200$ K. At this low temperature, slightly below 10% of its Debye temperature Θ_D , quantum effects are very strong, and we therefore expect problems in maintaining the large difference in temperature between the stiff and soft phonons. Using a harmonic system such as diamond is particularly useful, as one can monitor directly the efficiency of the thermostat by projecting the atomic velocities on a selection of normal modes. Hence, a projected kinetic temperature

well,

$$\bar{\omega} = 2\pi \frac{\langle v \rangle}{L} = \sqrt{\frac{8\pi k_B T}{mL^2}}$$

Figure 5.11: (a): ω -dependence of the kinetic energy correlation time $\tau_k(\omega)$ (light, dotted line), the ratio of the fitted fluctuations $c_{pp}(\omega)$ (red line) and of $\omega^2 c_{qq}(\omega)$ (blue line) with the exact, quantum-mechanical target function. (b): normal-mode-projected kinetic temperature for a few, selected phonons. The dashed line is the value expected from the fitting $c_{pp}(\omega)$, while the full line is the exact, quantum-mechanical expectation value for a harmonic oscillator of the same frequency.



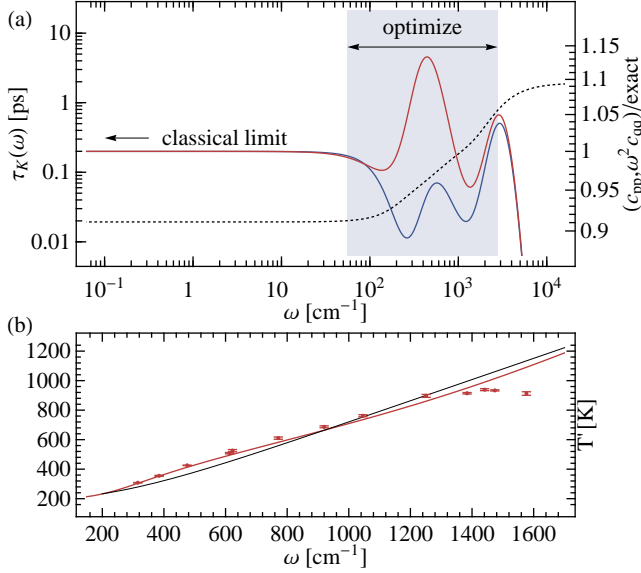
$T'(\omega)$ can be computed and its value can be checked against the predictions in the harmonic limit $T^*(\omega) = c_{pp}(\omega) / k_B$.

We first performed a fit which takes into account only the terms (5.6) and (5.7). It would be reasonable to also include a term that optimizes the efficiency of sampling, but at this stage we only wanted to ensure that the *predicted* discrepancy from the exact quantum result is as low as possible.

In Figure 5.11 we report the results with a matrix that was fitted taking into account only Eqs. (5.6) and (5.7). Even in a harmonic system such as diamond there are major errors due to ZPE leakage from the high-frequency modes to the low-frequency modes, which are compensated for only partially by the thermostat. Clearly, it makes little sense to obtain a very high accuracy on the fitting of c_{pp} and c_{qq} , if anharmonicities then lead to discrepancies of more than 100 % between the predicted and actual expectation values.

In principle one could fit a thermostat that explicitly takes into account the anharmonic cross-coupling between different normal modes, so as to compensate exactly for energy leakage. However this solution is not very practical, as it results in a thermostat that is specific to each particular system and because of the difficulties associated with computing all the relevant coupling terms. A simpler strategy is instead to enhance the system-bath coupling, so as to actively counterbalance the transfer of kinetic energy between different normal modes. This approach is transferable, and therefore requires only information on the largest vibrational frequency.

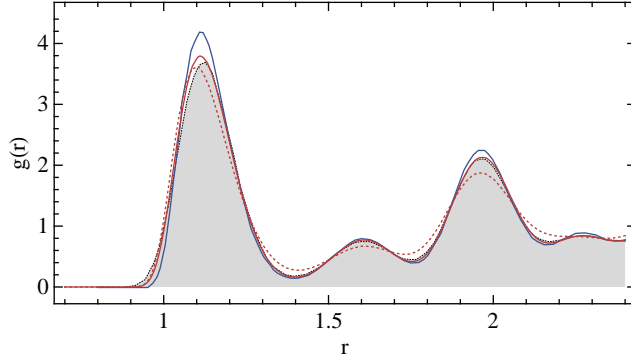
Figure 5.12: (a): ω -dependence of the kinetic energy correlation time $\tau_K(\omega)$ (light, dotted line), the ratio of the fitted fluctuations $c_{pp}(\omega)$ (red line) and of $\omega^2 c_{qq}(\omega)$ (blue line) with the exact, quantum-mechanical target function. (b): normal-mode-projected kinetic temperature for a few, selected phonons. The dashed line is the value expected from the fitting $c_{pp}(\omega)$, while the full line is the exact, quantum-mechanical expectation value for a harmonic oscillator.



To test this idea, we performed another fit, where we tried to reduce the correlation time of kinetic energy τ_K , which we took as an estimate of the intensity of the system-bath coupling. In doing so, we had to sacrifice slightly the accuracy of fitting for c_{pp} and c_{qq} . Fit data and results for Tersoff diamond are reported in figure 5.12. The projected kinetic temperature now agrees almost perfectly with the analytical predictions for the fluctuations $c_{pp}(\omega)$ of most of the modes. The only ones displaying significant deviations are the fast ones, for which the value of $\tau_K(\omega)$ is slightly larger. The $c_{pp}(\omega)$ curve deviates by nearly 10% from the exact, quantum-mechanical expectation value. However, thanks to the more efficient coupling, the errors due to anharmonicities are better compensated and so, in actuality, the overall error is much smaller than for the parameters presented in figure 5.11.

To test whether these prescriptions work for less harmonic problems, we now turn to a completely different system; namely, the structural properties of solid neon at 20 K. At variance with diamond, quantum-ion effects are less pronounced, but the system is close to its melting temperature and is thus significantly anharmonic. As shown in figure 5.13, the agreement between our results and those of accurate path-integral calculations[98] is almost perfect if the parameters of figure 5.12 are used, scaled to the appropriate temperature. As expected, large errors are present if those of figure 5.11 are chosen.

Figure 5.13: Radial distribution function computed from fully-converged path-integral calculations[98] (shaded curve), and from a quantum-thermostat MD trajectory for a Lennard-Jones model of solid neon at $T = 20$ K. Distances are in reduced units. The full line corresponds to the parameters of figure 5.12 and the dashed line to the set used in figure 5.11. The outcome of a purely classical simulation is also reported, in blue.



5.2.4 Liquid water

The results for solid neon at its melting point are very promising, as they agree nicely with PI data in spite of the large anharmonicities present. We thus decided to attempt the simulation of a flexible water model at liquid conditions, to push our method to its limits. We used the TIP4P-like potential[59] we employed in section 3.4 and in chapter 4, with the same set of GLE parameters described in figure 5.12, for which we sacrificed the accuracy of the fitting of the the fluctuations of p and q to the quantum expectation values in order to obtain a stronger system-bath coupling. As discussed above, this feature is of paramount importance if we are to attempt the modeling of an extremely anharmonic system such as a liquid.

We ran 2 ns of dynamics at constant volume, using a time step of 0.5 fs to guarantee an accurate integration of the equations of motion. We simulated a box containing 216 water molecules, and computed the average kinetic temperature projected on the various degrees of freedom. The results were encouraging, as the thermostat was able to sustain a huge temperature gradient between modes. In the final stationary distribution the internal modes equilibrated at $T'_{in} = 1290$ K, the hindered-rotations equilibrated at $T'_{ib} = 575$ K and the center-of-mass motion had a mean temperature of $T'_{cm} = 317$ K, close to the classical limit. Recent PIMD results[96] yield a value of $T'_{cm} = 320$ K for the kinetic temperature of the molecular center of mass, in good agreement with our result. To appreciate the strength of anharmonic coupling in this system, consider that in the absence of the thermostat, and starting from velocities obtained from the quantum distribution corresponding to the local Hessian, the centers of mass would reach a temperature of 700 K within 1 ps, due to ZPE leakage from internal modes[96].

Given these promising observations, we proceeded to evaluate the structural properties of the liquid, and compared the different radial distribution functions with those computed using purely classical and full path-integral calculations (figures 5.14 and 5.15). The agreement between PIMD and the quantum thermostat is very good.

Figure 5.14: Oxygen-oxygen (left) and hydrogen-hydrogen (right) radial distribution functions for a molecular dynamics simulation of a water model, at 298 K and experimental density. The shaded curve corresponds to a reference PIMD simulation, the blue curve to a purely classical dynamics, and the red curve has been obtained with a GLE quantum thermostat, with the parameters described in figure 5.12 scaled to the appropriate target temperature.

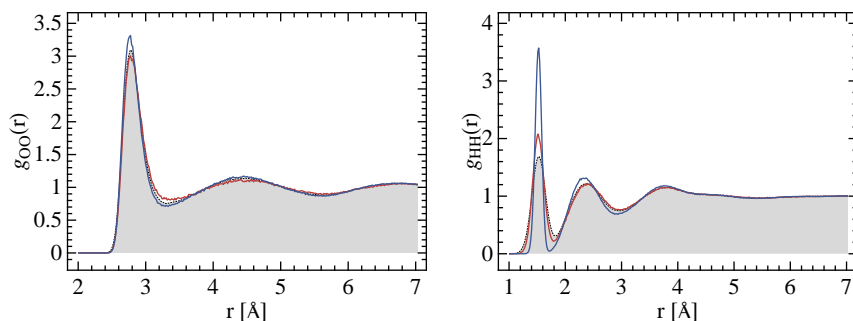
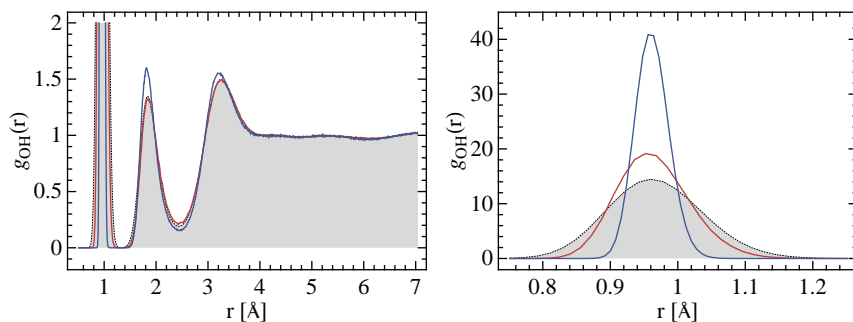


Figure 5.15: Oxygen-hydrogen radial distribution function for a molecular dynamics simulation of a water model, at 298 K and experimental density. The shaded curve corresponds to a reference PIMD simulation, the blue curve to a purely classical dynamics, and the red curve has been obtained with a GLE quantum thermostat, with the parameters described in figure 5.12 scaled to the appropriate target temperature. The rightmost panel shows the intramolecular peak in detail.



The intermolecular part of the radial distribution functions is described almost perfectly, while for the intra-molecular peaks, where the quantum nature of the proton is evident, the improvement over the classical result is impressive.

This satisfactory agreement in a strongly anharmonic system, which at the same time exhibits large nuclear quantum effects, suggests that our method can be applied to almost any condensed-phase problem. However, as is apparent from the significant deviation of the intra-molecular peak in $g_{OH}(r)$ from PIMD results, a residual error is present. We will discuss how these inaccuracies can be estimated in the following section.

5.2.5 A discussion of the accuracy

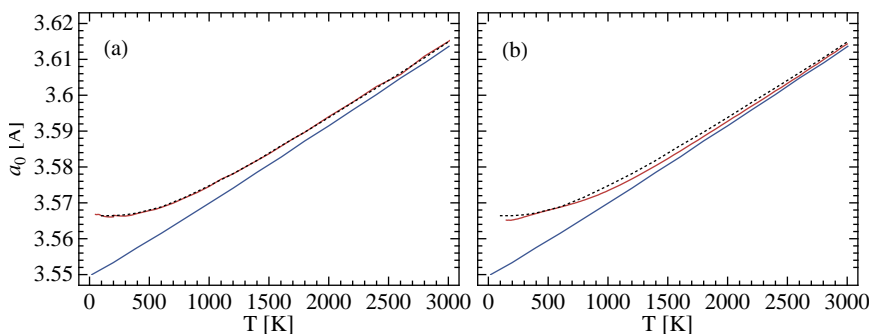
The agreement between the structural properties of liquid water computed using PIMD and quantum thermostat GLE dynamics is impressive, as discussed in the previous section. The fact that similar, satisfactory agreement is obtained for diamond and neon (figures 5.12 and 5.13), with the same set of parameters scaled to the appropriate target temperature, is a convincing proof that our simulation strategy can capture a significant portion of the nuclear quantum effects in many interesting systems. However, one must be aware that this is an approximate scheme, and that significant discrepancies might arise, which will depend on the property being studied.

A possible strategy for obtaining an order-of-magnitude estimate of the uncertainty in the results of a quantum thermostat calculation, that does not require one to do an expensive path-integral reference calculations, is to perform two simulations with two different sets of parameters, that are fitted to slightly different frequency ranges, or with different coupling strengths. The difference between the two results provides a lower bound on the inaccuracy in the treatment of quantum effects for the observables of interest.

To test this strategy, let us return to the example of Tersoff diamond we discussed in section 5.2.3. In Ref. [89] we have compared a number of properties calculated using our quantum thermostat with those obtained using path-integral methods, and obtained an excellent agreement down to a temperature of $0.1 \Theta_D$. To achieve these results, we fitted five different matrices, which we then used for different temperature ranges between 50 K and 3000 K, and we carefully balanced the accuracy of the fitting of c_{pp} and c_{qq} (Eq. (5.6)) and the coupling strength, which does not need to be *very* high in a harmonic system such as diamond. It is interesting to compare these refined results with those which could be obtained by simply rescaling the parameters used to create figure 5.12. Given that the fluctuations were fitted with an error as large as 10%, we cannot expect a better accuracy for the kinetic temperature and for corresponding properties.

Compare the temperature-dependence of the lattice parameter as obtained from the two sets of GLE matrices (figure 5.16). The purpose-fit parameters give a nearly-perfect agreement with the path-integral results [99] down to 10% of the Debye temperature but are poorly transferable, since we sacrificed in part the coupling strength. The general-purpose parameters meanwhile capture the main features of

Figure 5.16: Lattice parameter as a function of temperature for an NPT simulations of a Tersoff model of diamond. The black line corresponds to a set of path-integral results[99], the blue line to a set of purely classical simulations and the red line to trajectories using the quantum thermostat. In panel (a), parameters were fitted for five temperature windows, and an accurate fit of fluctuations was required, together with mild requirements on the coupling time. Panel (b) uses the general-purpose parameters used in figure 5.12 over the whole range of temperatures.



the thermal expansion, and improve enormously with respect to a purely classical simulation. However, they give less accurate results than could be obtained with a more careful fit.

5.2.6 Prospective improvements

The level of accuracy of figure 5.16(b) is sufficient for most applications, and can be improved by performing custom fits when the system is very harmonic and the requirements of extremely strong coupling can be relaxed. We have also shown that in more controlled, one-dimensional cases, the qualitative features stemming from nuclear quantum effects can be captured even in extremely anharmonic cases, such as the particle in a box.

Currently one cannot systematically increase the accuracy of our approach, not even through higher computational expenditure. This possibility would make it much easier to test the accuracy of results, and would increase its reliability. A viable strategy for reaching this goal, which we are now beginning to explore, would be to combine the quantum thermostat with a path-integral formulation. In this approach a quantum-thermostat simulation could be seen as a modified one-bead dynamics, in which a set of GLE parameters is used which enforces a strongly out-of-equilibrium phase-space distribution. At the other end of the scale, a fully-converged PIMD trajectory requires $\mathcal{O}(\hbar\omega_{max}/k_B T)$ replicas, on each of which a conventional thermostat is applied to maintain the system at the appropriate temperature (see chapter 4). One could then envisage that a PI simulation, with a small number of beads, could be brought closer to convergence by enforcing a stationary phase-space distribution intermediate between the canonical distribution and the one generated by the bare quantum thermostat. Finding the form of this distribution, and verifying the rate of convergence of the error will be the subject of further investigation.

Nuclear quantum effects in lithium imide

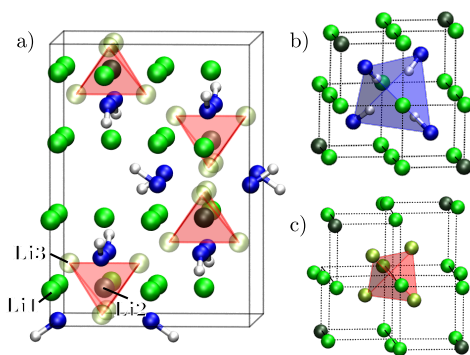
Lithium amide (LiNH_2) and imide (Li_2NH) have been extensively studied in recent years as they are promising materials for hydrogen storage[100–103]. Hydrogen release occurs in the mixture $\text{LiNH}_2 + \text{LiH}$ via a reversible solid-state decomposition reaction into lithium imide and molecular hydrogen, $\text{LiNH}_2 + \text{LiH} \longrightarrow \text{Li}_2\text{NH} + \text{H}_2$. The typical operating temperature for this system is around 280 °C. This is probably too high for on-board applications. Nevertheless, the amide/imide system is under deep scrutiny since it represents a prototypical, relatively simple system, which could shed light on the mechanisms for reversible H-release in the more complex, and technologically promising, reactive hydrides. In spite of a substantial amount of experimental work, several issues on the basic properties of this system are still open, including its structural properties.

Atomistic modeling has been proved to be very useful for obtaining insight into the properties of Li_2NH . However, to ensure reliable results, it is vitally important to use an accurate, density-functional description of the interactions. In addition, as this compound contains light atoms, it is also desirable to include a treatment of nuclear quantum effects. In this chapter we will present the results we have obtained by combining *ab-initio* molecular dynamics with the quantum thermostat. We will demonstrate how a number of observables relevant for the interpretation of experiments can be obtained with good accuracy and affordable computational effort.

6.1 Crystal structure

Differential thermal analysis and NMR measurements[104, 105] in the late 60's revealed a reversible phase transition at 356 K between an unknown low temperature (LT) structure and a high temperature antiferroite phase of Li_2NH . Further structure refinement from x-ray and neutron diffraction measurements of deuterated imide (Li_2ND)[106] has been published recently. The high temperature phase has a diffraction pattern consistent with an anti-fluorite structure, in which hydrogen atoms

Figure 6.1: (a) Tetragonal unit cell consistent with the experimental diffraction pattern. Atoms are arranged according to the $Fd\bar{3}m$ space group. Li2 and Li1 sites are fully occupied, whereas Li3 positions have partial occupation. (b) A detail of the tetrahedral arrangement of imide groups which surrounds constitutional Li^+ vacancies. (c) A Li2 vacancy - Li3 interstitials cluster. We observed formation of this structure in *ab initio* molecular dynamics trajectories starting from a realization of the structure in panel (a), in which the Li interstitials were distributed according to Ref. [108]. These clusters are the building blocks of the structure we propose.



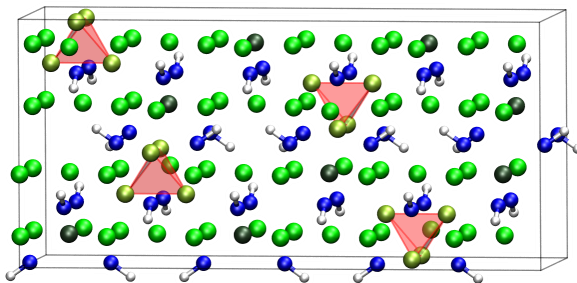
occupy the $192l$ positions of the $Fm\bar{3}m$ space group. At low temperature (100-300 K) the diffraction data were best fitted by a cubic crystal with $Fd\bar{3}m$ space group (figure 6.1(a)). This LT crystal can be seen as a superstructure of the antifluorite phase in which one out of every 8 Li atoms is displaced to an interstitial site. This gives rise to an ordered arrangement of Li vacancies, which are tetrahedrally coordinated to four NH groups (figure 6.1(b)). This structure is stabilized by an electrostatic interaction between the formally negatively-charged Li^+ vacancy and the H^+ atoms which points towards it. Similar tetrahedral arrangements of NH groups are present in other imides, such as $\text{Li}_2\text{Mg}(\text{NH})_2$ [107].

In the LT phase the H and N atoms occupy the $32e$ sites. The Li atoms meanwhile are distributed over three different sites: $48f$ (Li1), $8a$ (Li2) and $32e$ (Li3). Structural refinement yields a fractional occupation for the last of these positions of about $1/3$ at low temperatures. This is well above the value required by stoichiometry ($1/4$), which is only observed in experiments at room temperature.

The presence of sites which are partially occupied is a sign of static disorder. It is necessary to resort to theoretical modeling to resolve the local structure and the correlations between occupations. A structure obtained by populating a single Li3 site around each Li2 position has been proposed [108]. However, this arrangement is unstable and upon local relaxation lithium interstitials move to occupy octahedral sites, in contradiction with the diffraction pattern [106, 108, 109]. Moreover, the calculated enthalpy of formation for this model is higher than that obtained in experiments. In recent years, other structures have been suggested on the basis of *ab-initio* calculations, which have lower formation energies than the one in Ref. [108]. However, these structures do not reproduce the experimental diffraction data [110, 111].

Based on a combination of *ab-initio* molecular dynamics and static optimization,

Figure 6.2: Supercell corresponding to $3 \times 1 \times 1$ tetragonal cells (figure 6.1(a)), where we have distributed four lithium interstitials-vacancy clusters, so as to obtain the appropriate stoichiometry. Other arrangements of tetrahedra exist with nearly degenerate energy. However, the one in figure is the most stable, and so we used it as the starting point for further simulations.



we have recently found a different pattern for the occupation of the $Fd\bar{3}m$ sites[112], which is both more stable than the one in Ref. [108], and in excellent agreement with all the experimental evidence. During a short *ab-initio* molecular dynamics trajectory¹, started from the configuration given in Ref. [108], we have observed fast diffusion of the lithium interstitials, and the eventual formation of tetrahedral clusters in which a vacancy in Li2 position is surrounded by four occupied Li3 sites (see figure 6.1(c)). A superstructure in which all the interstitials are arranged in this fashion has a considerably lower enthalpy than the relaxed structure in Ref. [108]. Considering that various arrangements of the clusters have nearly degenerate energies, a long-range disorder is predicted. This disordered structure has the experimental space group with fractional occupations of the Li3 sites (1/3) and Li2 sites (2/3).

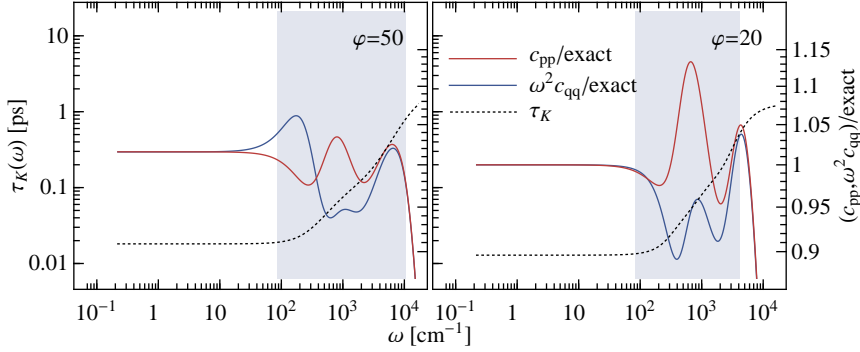
The symmetry-adapted average structure which can be inferred from our most stable arrangement (figure 6.2) is in excellent agreement with experimental data. The only discrepancy is the presence of fractional occupation for the Li2 sites, which probably was not considered in the Rietveld refinement. In the simulations described in the remainder of this chapter we used a representative supercell, based on the most stable arrangement of lithium interstitial-vacancy clusters we found. We employed the quantum thermostat to perform trajectories at 300 K, as will be discussed in the following sections.

6.2 Quantum-thermostatted simulations

Li_2NH is a challenging test for the quantum thermostat for several reasons. Being

¹A $\sqrt{2} \times 1 \times \sqrt{2}$, 128-atoms supercell corresponding to two tetragonal unit cells (cf. figure 6.1(a)) was used. We performed Born-Oppenheimer molecular dynamics simulations within Density Functional Theory with gradient corrected exchange and correlation functional[113] as implemented in the CPMD[57] package. Ultrasoft[114] and Goedecker-type[115] pseudopotentials were used for N and H atoms and for Li with three valence electrons, respectively. Kohn-Sham orbitals were expanded in plane waves up to a kinetic-energy cutoff of 50 Ry. Brillouin Zone (BZ) integration was restricted to the Γ point only. A time step of 0.6 fs was used and a constant temperature of 300 K was enforced by an optimal-sampling generalized Langevin equation thermostat.

Figure 6.3: Comparison of the fitted properties for the two set of parameters we used in our simulations. In the two panels, the ratio between the fitted and exact fluctuations (cfr. Eq. (5.6)) are reported, together with the correlation time of kinetic energy τ_K , for $\varphi = \hbar\omega_{\max}/k_B T = 50$ (left) and $\varphi = 20$ (right). The shaded region corresponds to the range on which Eq. (5.6) has been enforced.

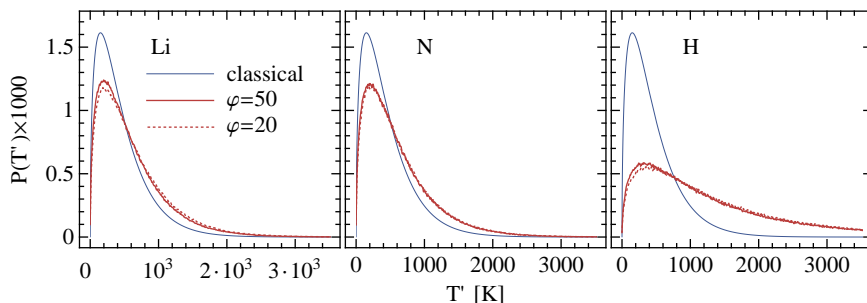


composed exclusively of light nuclei, we expect its properties to be significantly affected by nuclear quantum effects. Moreover, it is a relatively soft material, in which anharmonicities are important. Last of all, we want to perform *ab-initio* trajectories. This makes it hard to obtain a reliable reference to gauge the accuracy of our method, as an *ab-initio* PIMD trajectory would be prohibitively expensive and, to our knowledge, no good-quality classical force field is available.

We therefore applied the empirical strategy we described in section 5.2.5 to test our approach. We prepared two sets of parameters using different fitting strategies, and quantified inaccuracy by calculating the discrepancy between the results obtained with the two choices. The highest frequencies present, which set the upper bound of the range to be included in the fit, can be ascribed to the NH bond stretching. Consequentially, the same GLE parameters we used for water (fitted for $\varphi = \hbar\omega_{\max}/k_B T = 20$, see section 5.2.4) are sufficient to cover the relevant range of frequencies at 300 K, and can be used as a first reference. As a second option we prepared a pair of $\{\mathbf{A}_p, \mathbf{C}_p\}$ matrices with target $\varphi = 50$, as we know that $\varphi > 20$ is required in the low-temperature simulations that we intend to perform in the future.

The properties of the two sets are represented in figure 6.3. For $\varphi = 50$ the coupling is slightly less intense, since non-equilibrium, frequency-dependent thermalization has been enforced over a larger range. On the other hand, we obtained a better fit of $c_{pp}(\omega)$ and $c_{qq}(\omega)$. While there are no compelling reasons to prefer one choice over the other, there are arguments in favor of $\varphi = 50$ which has tighter fit to the quantum fluctuations. If we consider that Li_2NH , albeit anharmonic, is a crystalline solid, we can safely assume that phonon-phonon coupling terms will be less dramatic than in liquid water. Hence we can expect the results to be little affected by the weaker coupling intensity. We therefore consider the $\varphi = 50$ parameters as our best estimate, and use $\varphi = 20$ only to gauge the inaccuracies in the physical observables. As we will demonstrate, the two choices yield similar results, which makes our preference somewhat irrelevant.

Figure 6.4: Comparison of the distribution of kinetic temperature, as computed separately for individual atomic species, for a purely classical (blue line) and two quantum-thermostatted simulations of Li_2NH . The continuous red line corresponds to $\varphi = 50$, while the dotted red line to $\varphi = 20$.



6.2.1 Temperature of atomic species

A normal-modes analysis of the kind performed for diamond in section 5.2.3 would not be very significant here, as it would be hard to tell apart the discrepancies due to phonon-phonon coupling and those due to the fact that the vibrations are not strictly harmonic. We can however compute the kinetic temperature of different elements separately, and compare the outcome resulting from the different choices of GLE parameters.

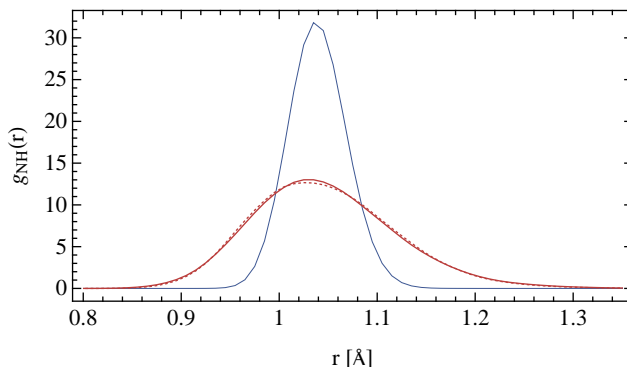
The results of this analysis are reported in figure 6.4, which shows that nuclear quantum effects are very significant for hydrogen atoms. They are also sizable for lithium and nitrogen, which both have an average effective temperature around 400 K. In addition, it is worth noting that the discrepancy between the results at $\varphi = 50$ and 20 is small if not negligible. This gives us some confidence in the reliability of our method when applied to Li_2NH . We will return in more detail to the distribution of the proton momentum in section 6.3, and will discuss the relevance of this result with regards to experiment.

6.2.2 Structural properties of NH bonds

We now comment further on the accuracy of our results by considering the structural properties of NH groups. As discussed in section 6.1, the distinctive trait of the low-temperature phase of Li_2NH is the presence of Li^+ vacancies in the lithium antiferroite lattice which are coordinated by four protons, in a tetrahedral arrangement (see figure 6.1(b)). The stability of this structure has been disputed, and other arrangements of the bonds have been suggested[110, 116, 117]. For instance, if the Li^+ ions are pinned so that they fully occupy the antiferroite sublattice, the imide groups can rotate freely even at room temperature. However, the resulting structure is unstable, as Li^+ Frenkel pairs form spontaneously if the lithium atoms are not constrained.

It is therefore both useful and instructive to analyze in some detail the structural properties of the imide groups, and the influence of quantum effects on their bond

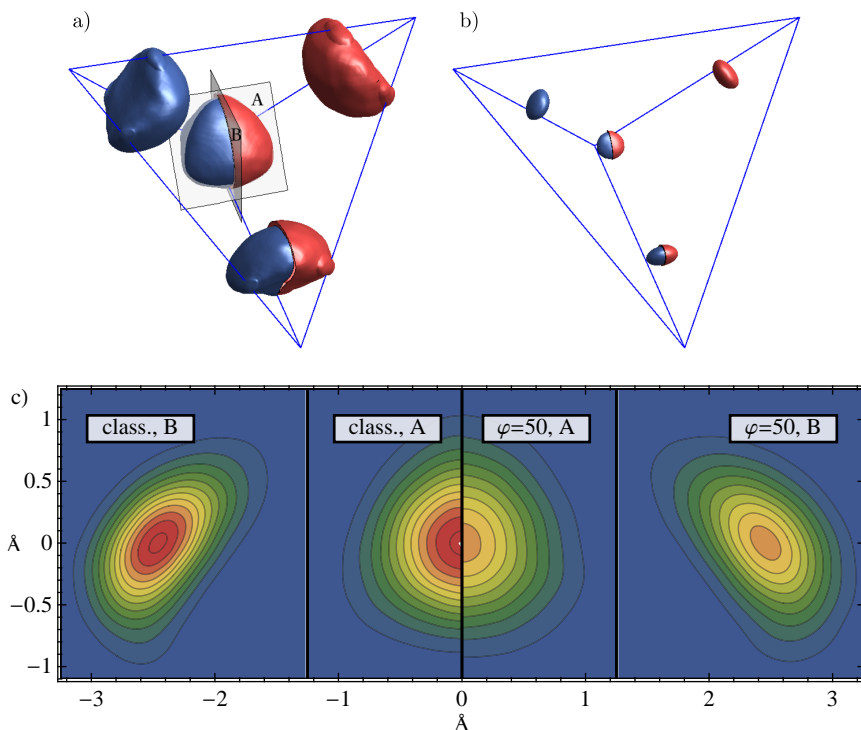
Figure 6.5: Short-range radial distribution function for the hydrogen-nitrogen pairs. The result from a purely classical trajectory (blue line) is compared to the results from quantum-thermostatted simulations (red lines). The continuous line corresponds to $\varphi = 50$, and the dotted line to $\varphi = 20$.



lengths and orientations. With this in mind, we first computed the NH radial distribution function, so as to characterize the distribution of bond lengths. The resulting curve is reported in figure 6.5, which demonstrates the importance of nuclear quantum effects. Figure 6.5 also demonstrates the reliability of our method, as it shows that the discrepancy between $\varphi = 50$ and $\varphi = 20$ is small, and minute when compared with the difference from the classical estimate. The broadened peak is asymmetric, which indicates a significant anharmonicity of the underlying potential. Indeed, the average bond length predicted by the classical simulation is 1.041 Å. This increases to 1.057 Å when the GLE thermostat is used.

As far as orientational order is concerned, we observed that the tetrahedral structures of figure 6.1(a) are very stable, but that imide groups undergo large vibrations about the (111) and equivalent directions. In figure 6.6 we plot the three-dimensional probability density for hydrogen atoms, symmetry-adapted to the $Fd\bar{3}m$ space group, so as to provide a quantitative picture of the typical orientation of the bonds. It is clear that in both classical and quantum-thermostatted dynamics the imide bonds do not undergo free rotations, but can oscillate considerably around their equilibrium position. The distribution in figure 6.6 is smeared by the vibrations of the entire imide groups and by static disorder, as it is averaged over all the hydrogen atoms in the supercell. This demonstrates that nuclear quantum effects are important, but their effect on this averaged orientations is not as evident as their effect on NH bond length.

Figure 6.6: Three-dimensional probability distribution of the proton. In panels (a) and (b) we report the isoproability contours enclosing 90% and 10% of the density, respectively. Blue contours correspond to classical simulations, and red ones to the GLE thermostat with $\varphi = 50$. The blue lines represent the tetrahedron formed by the symmetry-averaged positions of the nitrogen atoms. In panel (c) we also report the plot of the density along the A and B planes (see panel (a)). Contours are drawn every 0.1 \AA^{-3} .

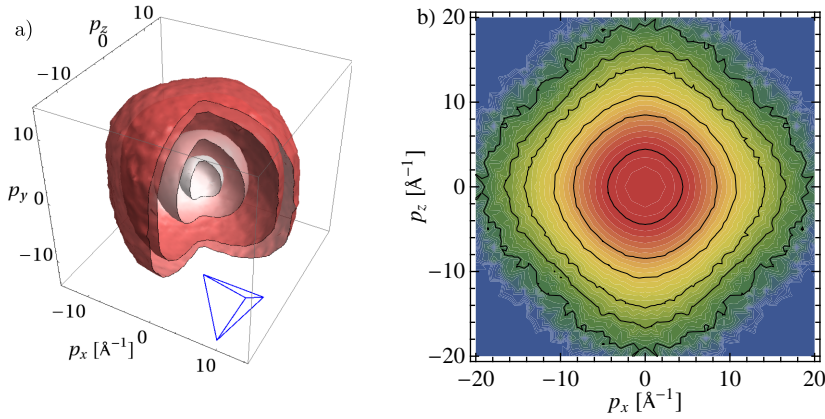


6.3 Proton momentum distribution

In a purely classical framework the constant-temperature probability distribution of the momentum of any particle is rather uninteresting, as it depends only on the temperature and the mass of the particle. In contrast, when significant quantum-mechanical effects are present, the momentum distribution depends strongly on the underlying potential. Consequentially, this quantity can be used as a powerful probe of the local environment in materials which are disordered, or difficult to analyze using conventional techniques.

In recent years the development of pulsed neutron sources has made it possible to use inelastic neutron scattering to probe the momentum distribution of individual atoms[118–120]. Hydrogen is an ideal target for these studies, because of its large incoherent cross section and also because the scattered neutrons have a distribution of energies that is well separated from those of neutrons scattered by other species. Moreover, because of its small mass, the proton has a strong quantum behavior, which makes its momentum distribution very sensitive to the local potential energy

Figure 6.7: (a) Direction-resolved proton momentum distribution computed from a molecular dynamics trajectory where we employed the GLE parameters $\varphi = 50$. The orientation of a NH tetrahedron is also drawn, to give the orientation of the axes relative to the crystal structure. Isosurfaces enclose (from the most external to the innermost surface) 95%, 90%, 50% and 10% of the density. (b) A contour plot of the 2D PMD computed on the xz plane. Contour lines are traced for $n(p_x, p_z) = 5 \times 10^{-3}, 10^{-3}, 5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5} \text{ \AA}^2$.



surface.

However, measuring and interpreting the proton momentum distribution (PMD) is not a simple task and is a significant challenge also from the point of view of simulation. For instance, conventional path-integral methods can only generate statistically-representative configurations of *positions* - one must adopt more refined *open path* methods to access the distribution of momenta[73, 80]. These techniques are complex, and lead to slowly-converging averages. As a result, only ever in exceptional cases and with great computational effort[91], have they been applied together with a first-principles computation of interatomic forces.

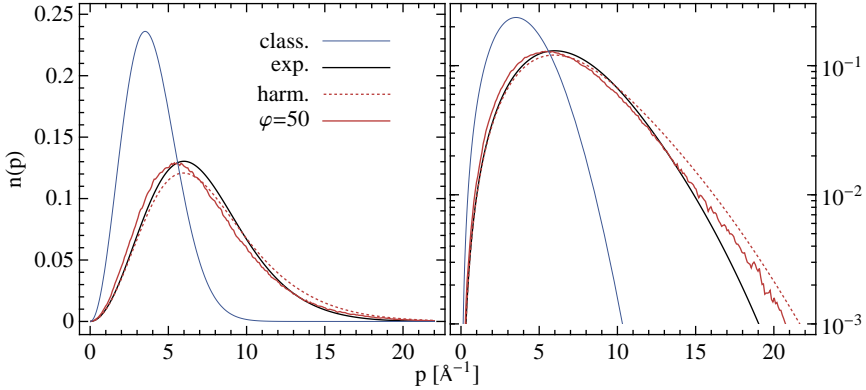
In contrast, for the quantum thermostat the distribution of velocities is an approximation of the quantum-mechanical finite temperature distribution, by construction. Hence, we can obtain the PMD easily, and compare it with experimental data which have been gathered very recently at ISIS (Oxford) by A. Pietropaolo.

6.3.1 Anisotropic momentum distribution

Most of the experimental assessments of PMD to date have been collected on disordered materials or on powder samples. In these cases, a spherically-averaged distribution is recovered, which we will discuss below. Whenever it is possible to perform measurements on single-crystal samples, one can infer information on the preferential orientation hydrogen atoms by observing the directional dependence of the PMD $n(\mathbf{p})$ [121, 122].

The anisotropy of the PMD is a purely quantum phenomenon, and is clearly visible in the 3D and 2D distributions we have computed from our quantum-

Figure 6.8: Spherical average of the proton momentum distribution, $n(p)$ for Li_2NH at 300 K. The classical Maxwell-Boltzmann distribution is shown for reference, together with experimental data. We have computed the PMD by ab-initio calculations, both from quantum-thermostatted dynamics and in the harmonic limit, using the phonons obtained at the Γ point by diagonalization of the dynamical matrix. In the rightmost panel, PMDs are shown on a logarithmic scale, to magnify the difference in the tail region.



thermostatted trajectory (see figure 6.7). This anisotropy arises because the proton momentum has larger fluctuations in the direction parallel to the stiffer bonds. As discussed above, protons in Li_2NH are oriented towards the centers of the tetrahedral arrangements of four imide groups, and so the three-dimensional $n(\mathbf{p})$ has cubic symmetry. This is particularly evident in the tails of the distribution.

6.3.2 Spherically-averaged momentum distribution

When measurements are performed on powder samples, the orientational dependence of $n(\mathbf{p})$ is lost, and only the spherically-averaged $n(p)$ can be recovered. However, a remnant of the original anisotropy is left, because, for a purely harmonic but non-spherical potential, the shape of $n(p)$ will not be a Gaussian (see appendix F).

We have been given some preliminary, unpublished experimental data recorded on a powder sample of Li_2NH at ISIS (Oxford), and we can thus perform a comparison with our simulations (see figure 6.8).

As a further theoretical comparison, we also computed the proton momentum distribution predicted by a purely harmonic model, based on normal modes computed at the Γ point using finite displacements of the atoms. To compare this result with quantum-thermostatted simulations and with experiments it would be necessary to take into account anharmonic corrections to the potential. However, the improvement over a classical simulation is impressive, and there is qualitative agreement between the experimental and both theoretical curves. There are on the other hand discrepancies which are significant compared to the error bar of the experiments, and further analysis is being performed so we can obtain a quantitative interpretation of the results.

In conclusion, we have demonstrated that in this challenging practical application a non-equilibrium generalized Langevin equation, fitted to reproduce nuclear

quantum effects in the harmonic limit, is capable of qualitatively capture the features of the properties of anharmonic solids. An error of the order of 10% on the magnitude of quantum-ions corrections is to be expected, given the discrepancy between the results obtained with different sets of parameters. We would like to stress that the computational effort required in our study was at least one and a half orders of magnitude smaller than what would have been required in conventional PI techniques. This makes our method particularly useful for performing preliminary studies, to quantify the importance of quantum nuclei in the modeling of novel, cutting edge applications. It also provides quantitative estimates which offer a significant improvement over purely classical results, even if more accurate - and expensive - methods are required if high-precision values are needed.

Conclusions

In this thesis the development and several applications of a generalized Langevin framework for molecular dynamics simulations have been presented. Building on an extended phase-space Markovian realization of a non-Markovian Langevin equation, we have discussed the prediction of static and dynamic response properties of a stochastic dynamics applied to a harmonic potential. We have then used these predictions to tune the parameters of the thermostat. This allowed us to develop new methods that extend the range of applications which can be examined using molecular dynamics. The efficacy of these methods has then been demonstrated for a number of applications that despite their apparent simplicity represent complex sampling problems.

For instance, we have shown how the efficiency of sampling in constant-temperature molecular dynamics can be improved, and how an optimal-sampling GLE can be constructed. We have tested the reliability of our method for problems as diverse as liquid water - both in classical and path integral dynamics - and for examining the diffusion of hydrogen atoms in noble metals, where the H atoms are treated using PIMD. In addition, we have shown how similar concepts can be used to achieve the opposite effect; namely, GLE can be used to prevent the thermalization of a range of frequencies. This is useful for Car-Parrinello MD and similar methods based on adiabatic separation, where high-frequency components of the vibrational spectrum should be undisturbed.

In passing we have also compared more traditional, uncorrelated-noise stochastic thermostats, and have rationalized their positive and negative features using a set of quantitative parameters we have developed for characterizing the behavior of GLEs. During the course of this analysis we observed that, to efficiently sample configurations in liquids, it is essential that the diffusive, concerted modes are undisturbed. To fulfill this aim we have devised an analytical estimate of such disturbance, that can be used during the fit of the parameters of a GLE.

A radically different set of applications can be obtained if a class of non-equilibrium generalized Langevin equations is employed. When used in this way the effective temperature reached by different normal modes can be tuned as a function of their frequency. We have shown how this frequency-dependent response can be exploited to selectively excite a very narrow range of frequencies. We have then applied this method to both molecular dynamics and to the more general problem of spectral analysis of large matrices.

Using a similar philosophy, we have demonstrated how a phase-space distribution for positions and momenta, which is consistent with the quantum-mechanical finite-temperature density matrix, can be enforced in harmonic and quasi-harmonic problems. We have shown that this approach yields satisfactory results for strongly anharmonic, one-dimensional potentials, and have also discussed its application to a series of condensed-phase problems, ranging from diamond to liquid water. In all cases, we obtained good agreement with the exact, path integral approach, but at a fraction of the computational cost.

The final system we examined was lithium imide, a compound which is relevant for hydrogen storage. To model this correctly it is essential to include both nuclear quantum effects and an *ab initio* treatment of the interatomic forces. A study of this system with path integral methods would be prohibitively expensive. However, using our quantum thermostat, we obtained results in qualitative agreement with experiments, with a relatively small computational effort.

Our approach has been proved useful for performing both constant-temperature sampling and for including nuclear quantum effects in expensive, *ab initio* simulations. These ideas can be routinely used to gain insight in a number of important simulation problems[123]. They are being implemented in CP2K[124], CPMD[57], DLPOLY[60] and GROMACS[68]. Furthermore, we believe that we have only scratched the surface of possible applications of stochastic concepts in molecular dynamics. Many of the techniques we have employed lend themselves quite naturally to generalizations and to usage in different contexts, and we are beginning to explore these new possibilities. Moreover, our framework is based on a fairly simple form of a stochastic differential equation. Many other, more complex classes of dynamics have been explored by the mathematical community and could provide novel, useful tools in the realm of molecular dynamics and atomistic simulations.

Appendices

A Dimensional reduction

Elimination of dynamical variables The connection between the Markovian (2.2) and non-Markovian (2.1) formulations of the colored-noise Langevin equation can be understood using techniques similar to those adopted in Mori-Zwanzig theory[4, 22]. Let us first consider a very general, multidimensional OU process, where we single out some degrees of freedom (\mathbf{y}) that we wish to integrate out, leaving only the variables marked as \mathbf{x} .

$$\begin{pmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \end{pmatrix} = - \begin{pmatrix} \mathbf{A}_{xx} & \mathbf{A}_{xy} \\ \mathbf{A}_{yx} & \mathbf{A}_{yy} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + \begin{pmatrix} \mathbf{B}_{x\xi} \\ \mathbf{B}_{y\xi} \end{pmatrix} \begin{pmatrix} \xi \end{pmatrix} \quad (\text{A.1})$$

Assuming that the dynamics has finite memory, one can safely take $\mathbf{y}(-\infty) = 0$, and the ansatz²

$$\mathbf{y}(t) = \int_{-\infty}^t e^{-(t-t')\mathbf{A}_{yy}} [-\mathbf{A}_{yx}\mathbf{x}(t') + \mathbf{B}_{y\xi}\xi(t')] dt'. \quad (\text{A.2})$$

Substituting into (A.1), one sees that \mathbf{y} can be eliminated from the dynamics of \mathbf{x} , and arrives at

$$\begin{aligned} \dot{\mathbf{x}}(t) &= - \int_{-\infty}^t \mathbf{K}(t-t')\mathbf{x}(t') dt' + \zeta(t) \\ \mathbf{K}(t) &= 2\mathbf{A}_{xx}\delta(t) - \mathbf{A}_{xy}e^{-t\mathbf{A}_{yy}}\mathbf{A}_{yx} \quad (t \geq 0) \\ \zeta(t) &= \mathbf{B}_{x\xi}\xi(t) - \int_{-\infty}^t \mathbf{A}_{xy}e^{-(t-t')\mathbf{A}_{yy}}\mathbf{B}_{y\xi}\xi(t') dt'. \end{aligned} \quad (\text{A.3})$$

Incidentally, one can then verify that (A.3) are invariant under any orthogonal transformation of the \mathbf{y} dynamical variables, meaning that such a transformation leaves the dynamics of the \mathbf{x} 's unchanged.

The colored noise is better described in terms of its time-correlation function, $\mathbf{H}(t) = \langle \zeta(t)\zeta(0)^T \rangle$. Let us first introduce the symmetric matrix $\mathbf{D} = \mathbf{B}\mathbf{B}^T$, whose parts we shall label using the same scheme used for \mathbf{A} in Eq. (A.1). We shall also need

²For brevity, we will write $\int f(s)\xi_i(s)ds$ meaning in fact an Itô integral $\int f(s)dw_i(s)$.

$\mathbf{Z}_{yy} = \int_0^\infty e^{-\mathbf{A}_{yy}t} \mathbf{D}_{yy} e^{-\mathbf{A}_{yy}^T t} dt$. With these definitions in mind, one finds

$$\mathbf{H}(t) = \delta(t) \mathbf{D}_{xx} + \mathbf{A}_{xy} e^{-t\mathbf{A}_{yy}} \left[\mathbf{Z}_{yy} \mathbf{A}_{xy}^T - \mathbf{D}_{yx} \right] \quad (t \geq 0). \quad (\text{A.4})$$

Note that the value of $\mathbf{H}(t)$ for $t < 0$ is determined by the constraint $\mathbf{H}(-t) = \mathbf{H}(t)^T$; the value of $\mathbf{K}(t)$ instead, is irrelevant for negative times: we will assume $\mathbf{K}(-t) = \mathbf{K}(t)^T$ to hold, since this will simplify some algebra below.

Let's now switch to the case of the free-particle counterpart of Eqs. (2.2), which is relevant to the memory functions entering Eqs. (2.1). Here, we want to integrate away all the \mathbf{s} degrees of freedom, retaining only the momentum p . Hence, we can transform Eqs. (A.3) and (A.4) to the less cumbersome form

$$\begin{aligned} K(t) &= 2a_{pp}\delta(t) - \mathbf{a}_p^T e^{-|t|\mathbf{A}} \bar{\mathbf{a}}_p \\ H(t) &= d_{pp}\delta(t) - \mathbf{a}_p^T e^{-|t|\mathbf{A}} [\mathbf{Z}\mathbf{a}_p - \mathbf{d}_p] \end{aligned} \quad (\text{A.5})$$

This compact notation hides certain relevant property of the memory kernels, which are more apparent when the kernels are written in their Fourier representation. If $\mathbf{D}_p = \mathbf{B}_p \mathbf{B}_p^T$ is transformed according to Eq. (2.5). $K(\omega)$ and $H(\omega)$ read

$$\begin{aligned} K(\omega) &= 2a_{pp} - 2\mathbf{a}_p^T \frac{\mathbf{A}}{\mathbf{A}^2 + \omega^2} \bar{\mathbf{a}}_p \\ H(\omega) &= K(\omega) \left(c_{pp} - \mathbf{a}_p^T \frac{\mathbf{A}}{\mathbf{A}^2 + \omega^2} \mathbf{c}_p \right) + \\ &\quad + 2\omega^2 \left(\mathbf{a}_p^T \frac{1}{\mathbf{A}^2 + \omega^2} \mathbf{c}_p \right) \left(1 + \mathbf{a}_p^T \frac{1}{\mathbf{A}^2 + \omega^2} \bar{\mathbf{a}}_p \right). \end{aligned} \quad (\text{A.6})$$

It is seen that the memory functions (hence the dynamical trajectory) are independent of the value of \mathbf{C} , the covariance of the fictitious degrees of freedom. Moreover, a sufficient condition for the FDT to hold is readily found. By setting $c_{pp} = k_B T$ and $\mathbf{c}_p = 0$, one obtains $H(\omega) = k_B T K(\omega)$, which is precisely the FDT for a non-Markovian Langevin equation. Since the value of \mathbf{C} is irrelevant we can take $\mathbf{C}_p = k_B T$, which simplifies the algebra and leads to numerically-stable trajectories.

Note that the same argument (with a slightly more cumbersome notation) can be carried on when a potential is present in Eqs. (2.2), since the additional degrees of freedom \mathbf{s} are not coupled with the position q , and can be therefore integrated out in the same way.

B Stability of the harmonic GLE

In chapter 2 we have discussed how condition (2.7) guarantees that the dynamics generated by Eqs. (2.2) will sample the canonical ensemble. We have seen that it is very difficult to treat the general case, where fluctuation-dissipation theorem does not hold. However, we can prove that for an harmonic potential a positive-definite $\mathbf{A}_p + \mathbf{A}_p^T$ guarantees a well-defined stationary distribution. Firstly, let us remark by

transforming $q \leftarrow \omega q$ in (2.17) the rescaled \mathbf{A}_{qp} and \mathbf{D}_{qp} will be obtained:

$$\mathbf{A}'_{qp} = \begin{pmatrix} 0 & (-\omega, \mathbf{0}) \\ (\omega, \mathbf{0})^T & \mathbf{A}_p \end{pmatrix}, \quad \mathbf{D}'_{qp} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{D}_p \end{pmatrix}.$$

We now want to prove that for $\mathbf{A}_p + \mathbf{A}_p^T > 0$ and any positive-definite \mathbf{D}_p , the symmetric solution to

$$\mathbf{A}'_{qp} \mathbf{C}'_{qp} + \mathbf{C}'_{qp} (\mathbf{A}'_{qp})^T = \mathbf{D}'_{qp} \quad (\text{B.1})$$

will be positive-definite: in that case, a stationary solution to the Fokker-Planck equation exists, which is a Gaussian with \mathbf{C}'_{qp} as covariance.

Let \mathbf{x} be an arbitrary vector of complex numbers, and \cdot^\dagger represent Hermitian conjugation. Then³

$$0 \leq \mathbf{x}^\dagger \mathbf{D}'_{qp} \mathbf{x} = \mathbf{x}^\dagger \mathbf{A}'_{qp} \mathbf{C}'_{qp} \mathbf{x} + \mathbf{x}^\dagger \mathbf{C}'_{qp} (\mathbf{A}'_{qp})^T \mathbf{x} = 2 \operatorname{Re} \mathbf{x}^\dagger \mathbf{A}'_{qp} \mathbf{C}'_{qp} \mathbf{x}. \quad (\text{B.2})$$

Now, let us assume that \mathbf{C}'_{qp} has a negative eigenvalue v , and \mathbf{u} is the corresponding eigenvector. Then one would have

$$\operatorname{Re} \mathbf{u}^\dagger \mathbf{A}'_{qp} \mathbf{C}'_{qp} \mathbf{u} = v \operatorname{Re} \mathbf{u}^\dagger \mathbf{A}'_{qp} \mathbf{u} \leq 0$$

since $\mathbf{A}'_{qp} + (\mathbf{A}'_{qp})^T$ is positive semidefinite. This would be in contradiction with (B.2), except for the case in which $\mathbf{u} = \tilde{\mathbf{u}} = (1, 0, \mathbf{0})$, which is the null eigenvector of \mathbf{D}'_{qp} and $\mathbf{A}'_{qp} + (\mathbf{A}'_{qp})^T$.

If we consider this case separately, we see that the only possibility to have a negative v is that $\tilde{\mathbf{u}}$ is also the associated eigenvector of \mathbf{C}'_{qp} . On the other hand, this would mean⁴ that $v = \tilde{\mathbf{u}}^\dagger \mathbf{C}'_{qp} \tilde{\mathbf{u}} = c_{qq}$ and $c_{qp} = \mathbf{c}_q = 0$, which in turns implies⁵ that the \mathbf{C}_p part of \mathbf{C}'_{qp} will be just the solution of $\mathbf{A}_p \mathbf{C}_p + \mathbf{C}_p \mathbf{A}_p^T = \mathbf{D}_p$, positive-definite just by the argument above, without the nuisances of a zero eigenvalue. One also finds that in this case the relation $v = c_{pp}$ would hold, which means that indeed v must be positive since c_{pp} is the diagonal element of a positive-definite matrix.

In passing, we can also prove that a positive-definite $\mathbf{A}_p + \mathbf{A}_p^T$ is a sufficient condition to have a positive Fourier transform of the memory kernel $K(\omega)$ for all ω . To this aim, we consider that it is possible to write

$$\frac{1}{2} K(\omega)^{-1} = [\boldsymbol{\Omega}^{-1}]_{pp} = \left[\left(\mathbf{A}_p + \omega^2 \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{-1} \end{pmatrix} \right)^{-1} \right]_{pp}$$

³Let us recall that if \mathbf{M} is an arbitrary matrix in $\mathbb{C}^{n \times n}$,

$$\left[\forall \mathbf{x} \in \mathbb{C}^n, \operatorname{Re}(\mathbf{x}^\dagger \mathbf{M} \mathbf{x}) > 0 \right] \Leftrightarrow \mathbf{M} + \mathbf{M}^\dagger \text{ is positive definite.}$$

And that if \mathbf{M} is Hermitian, $\mathbf{x}^\dagger \mathbf{M} \mathbf{x} \in \mathbb{R}$. Also, by taking $\mathbf{x} = \hat{\mathbf{e}}_i = (0, \dots, 1, \dots, 0)$, one can see that if $\operatorname{Re}(\mathbf{x}^\dagger \mathbf{M} \mathbf{x}) > 0$ all the diagonal elements of \mathbf{M} must have a positive real part.

⁴For instance, $c_{qp} = (0, 1, \mathbf{0})^T \mathbf{C}'_{qp} \tilde{\mathbf{u}} = v (0, 1, \mathbf{0})^T (1, 0, \mathbf{0}) = 0$.

⁵Just write out all the separate parts of Eq. (B.1) in the case where $c_{qp} = \mathbf{c}_q = 0$.

and verify that $\mathbf{\Omega}$ is such that $\text{Re}(\mathbf{x}^\dagger \mathbf{\Omega} \mathbf{x}) > 0$ for any \mathbf{x} and ω . Then, one checks that the same property holds for $\mathbf{\Omega}^{-1}$, and hence its diagonal elements must be positive.

C Probability evolution in the harmonic limit

In the case of a harmonic potential it is possible to find an explicit solution for the spread of the probability density out of a point in (q, p) space, integrating away the dependence on the additional degrees of freedom. To this end, let us first review a few results for multidimensional Gaussian probability densities.

Product of Gaussians Consider two normalized Gaussian distributions with dimensionality d , means $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ and covariances \mathbf{C}_1 and \mathbf{C}_2 respectively. Their product will be a non-normalized Gaussian, with covariance $\mathbf{C} = (\mathbf{C}_1^{-1} + \mathbf{C}_2^{-1})^{-1}$ and mean $\bar{\mathbf{x}} = \mathbf{C}(\mathbf{C}_1^{-1}\bar{\mathbf{x}}_1 + \mathbf{C}_2^{-1}\bar{\mathbf{x}}_2)$. The integral of such non-normalized gaussian is

$$Z = \frac{1}{\sqrt{(2\pi)^d \det(\mathbf{C}_1 + \mathbf{C}_2)}} \exp -\frac{1}{2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T (\mathbf{C}_1 + \mathbf{C}_2)^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

Marginal probability It is possible to show that the marginal probability distribution for a subset of the variables in a larger set, knowing that, as a whole, they have a Gaussian distribution. Let again (\mathbf{x}, \mathbf{y}) be the full state vector, and \mathbf{y} the variables one wants to eliminate. The overall covariance can be splitted into the parts corresponding to the different sets of variables

$$\mathbf{C} = \left(\begin{array}{c|c} \mathbf{C}_{xx} & \mathbf{C}_{xy} \\ \hline \mathbf{C}_{xy}^T & \mathbf{C}_{yy} \end{array} \right)$$

and the overall probability distribution is $P(\mathbf{x}, \mathbf{y}) \propto \exp \left[-(\mathbf{x}, \mathbf{y})^T \mathbf{C}^{-1} (\mathbf{x}, \mathbf{y}) / 2 \right]$. One wants then to perform an integration over the \mathbf{y} 's to find⁶

$$P(\mathbf{x}) = \int P(\mathbf{x}, \mathbf{y}) d\mathbf{y} \propto \exp \left[-\mathbf{x}^T (\mathbf{C}_{xx})^{-1} \mathbf{x} / 2 \right]. \quad (\text{C.1})$$

Conditional probability As a final result, we recall that the conditional probability which can be obtained fixing the values of some of the variables in a multivariate Gaussian is still a Gaussian distribution. Using the same notation as above, and considering also that \mathbf{x} and \mathbf{y} components have averages $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ respectively, one

⁶It is useful here to recall that given a set of data points $\{\mathbf{x}_i\}$ which are normally distributed with mean $\bar{\mathbf{x}}$ and covariance \mathbf{C} , the data points $\{\mathbf{S}\mathbf{x}_i\}$ will be distributed as a Gaussian with mean $\mathbf{S}\bar{\mathbf{x}}$ and covariance $\mathbf{S}\mathbf{C}\mathbf{S}^T$. Eq. (C.1) follows by taking

$$\mathbf{S} = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \end{array} \right)$$

which is a matrix with as many columns as the total number of variables, and as many rows as the \mathbf{x} 's.

finds that $P(\mathbf{x}|\mathbf{y})$ has average $\bar{\mathbf{x}}' = \bar{\mathbf{x}} + \mathbf{C}_{xy}\mathbf{C}_{yy}^{-1}(\mathbf{y} - \bar{\mathbf{y}})$ and covariance $\mathbf{C}'_{xx} = \mathbf{C}_{xx} - \mathbf{C}_{xy}\mathbf{C}_{yy}^{-1}\mathbf{C}_{xy}^T = [(\mathbf{C}^{-1})_{xx}]^{-1}$.

One can then start from the solution of the Fokker-Planck equation Eq. (1.32). Let us define the shorthands $\mathbf{T} = e^{-t\mathbf{A}}$ and $\mathbf{R} = \mathbf{C} - e^{-t\mathbf{A}}\mathbf{C}e^{-t\mathbf{A}^T}$. We also drop normalizations, and write

$$P((\mathbf{x}, \mathbf{y}), t | (\mathbf{x}_0, \mathbf{y}_0), 0) \propto \exp \left[-\frac{1}{2} ((\mathbf{x}, \mathbf{y}) - \mathbf{T}(\mathbf{x}_0, \mathbf{y}_0))^T \mathbf{R}^{-1} ((\mathbf{x}, \mathbf{y}) - \mathbf{T}(\mathbf{x}_0, \mathbf{y}_0)) \right].$$

A first step is to integrate out the dependence on \mathbf{y} , to find the marginal probability

$$P(\mathbf{x}, t | (\mathbf{x}_0, \mathbf{y}_0), 0) \propto \exp \left[-\frac{1}{2} (\mathbf{x} - (\mathbf{T}_{xx}\mathbf{x}_0 + \mathbf{T}_{xy}\mathbf{y}_0))^T \mathbf{R}_{xx}^{-1} (\mathbf{x} - (\mathbf{T}_{xx}\mathbf{x}_0 + \mathbf{T}_{xy}\mathbf{y}_0)) \right]. \quad (\text{C.2})$$

Since we aim at an expression which does not contain any reference to additional momenta, we decide to consider them to be picked at random, in a way which is consistent with the stationary distribution and the choice of \mathbf{x}_0 :

$$P(\mathbf{x}, t | \mathbf{x}_0, 0) = \int d\mathbf{y}_0 P(\mathbf{x}, t | (\mathbf{x}_0, \mathbf{y}_0), 0) P(\mathbf{y}_0 | \mathbf{x}_0).$$

Based on the discussion above for the conditional probability of a multivariate Gaussian, we find

$$P(\mathbf{y}_0 | \mathbf{x}_0) \propto \exp \left[-\frac{1}{2} (\mathbf{y}_0 - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{x}_0)^T (\mathbf{C}^{-1})_{yy} (\mathbf{y}_0 - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{x}_0) \right].$$

After straightforward albeit tedious calculations, one see that $P(\mathbf{x}, t | \mathbf{x}_0, 0)$ can be written as

$$P(\mathbf{x}, t | \mathbf{x}_0, 0) \propto \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{U}_{xx}(t)\mathbf{x}_0)^T \mathbf{W}_{xx}^{-1}(t) (\mathbf{x} - \mathbf{U}_{xx}(t)\mathbf{x}_0) \right],$$

where the two matrices can be written as

$$\begin{aligned} \mathbf{W}_{xx}^{-1} &= \mathbf{R}_{xx}^{-1} - \mathbf{R}_{xx}^{-1}\mathbf{T}_{xy} \left((\mathbf{C}^{-1})_{yy} + \mathbf{T}_{xy}^T \mathbf{R}_{xx}^{-1} \mathbf{T}_{xy} \right)^{-1} (\mathbf{R}_{xx}^{-1}\mathbf{T}_{xy})^T \\ \mathbf{U}_{xx} &= \mathbf{T}_{xx} + \mathbf{W}_{xx} \mathbf{R}_{xx}^{-1} \mathbf{T}_{xy} \left((\mathbf{C}^{-1})_{yy} + \mathbf{T}_{xy}^T \mathbf{R}_{xx}^{-1} \mathbf{T}_{xy} \right)^{-1} (\mathbf{C}^{-1})_{yy} \mathbf{C}_{yx} \mathbf{C}_{xx}^{-1}. \end{aligned} \quad (\text{C.3})$$

One also sees that, in the limit as $t \rightarrow \infty$, $\mathbf{W}_{xx} \rightarrow \mathbf{C}_{xx}$ and $\mathbf{U}_{xx} \rightarrow 0$, so that the stationary distribution is recovered.

D Analytical results

We collect in this appendix a number of simple but cumbersome analytical results, concerning the static and dynamical properties of a multivariate Ornstein-Uhlenbeck process. Substituting \mathbf{A} and \mathbf{C} matrices (the drift term and the static covariance for a generic OU process) with \mathbf{A}_p and \mathbf{C}_p or \mathbf{A}_{qp} and \mathbf{C}_{qp} for the various matrices, one

can obtain the relevant quantities for the free-particle and the harmonic oscillator cases of the GLE thermostat.

The diffusion matrix \mathbf{B} can be obtained by an expression analogous to Eq. (2.5). The same relation can be used to obtain the elements of \mathbf{C} given the drift and diffusion matrices, by solving the linear system. However, the covariance matrix can be computed more efficiently by finding the eigendecomposition of $\mathbf{A} = \mathbf{O} \text{diag}(\alpha_i) \mathbf{O}^{-1}$, and computing

$$C_{ij} = \sum_{kl} \frac{O_{ik} \left[\mathbf{O}^{-1} \mathbf{B} \mathbf{B}^T \mathbf{O}^{-1T} \right]_{kl} O_{jl}}{\alpha_k + \alpha_l}. \quad (\text{D.1})$$

Since the equilibrium distribution for an OU process is a Gaussian, the covariance \mathbf{C} describes exhaustively the static properties. We then move on to compute the dynamical properties, by means of the correlation functions.

Let $\mathbf{x}(t)$ be the vector describing the trajectory of the OU process. The first non-normalized correlation function is readily computed to be

$$\mathcal{C}_{ij}(t) = \langle x_i(t) x_j(0) \rangle = [e^{-|t|\mathbf{A}} \mathbf{C}]_{ij} \quad (\text{D.2})$$

and its integral to time t is

$$\int_0^t \mathcal{C}_{ij}(s) ds = [\mathbf{A}^{-1} (1 - e^{-t\mathbf{A}}) \mathbf{C}]_{ij}.$$

In order to compute τ_H or τ_V (Eq. (3.1)) for the harmonic oscillator, one needs time correlation functions of the second order combinations of coordinates, of the form $x_i(t) x_j(t)$. It reads

$$\begin{aligned} \mathcal{C}_{ijkl}(t) &= \langle x_i(t) x_j(t) x_k(0) x_l(0) \rangle - \langle x_i x_j \rangle \langle x_k x_l \rangle = \\ &= [e^{-|t|\mathbf{A}} \mathbf{C}]_{ik} [e^{-|t|\mathbf{A}} \mathbf{C}]_{jl} + [e^{-|t|\mathbf{A}} \mathbf{C}]_{il} [e^{-|t|\mathbf{A}} \mathbf{C}]_{jk} \end{aligned}$$

The corresponding, non-normalized integrals

$$\tau_{ijkl} = \int_0^\infty \mathcal{C}_{ijkl}(s) ds$$

can be computed in terms of the tensorial quantity

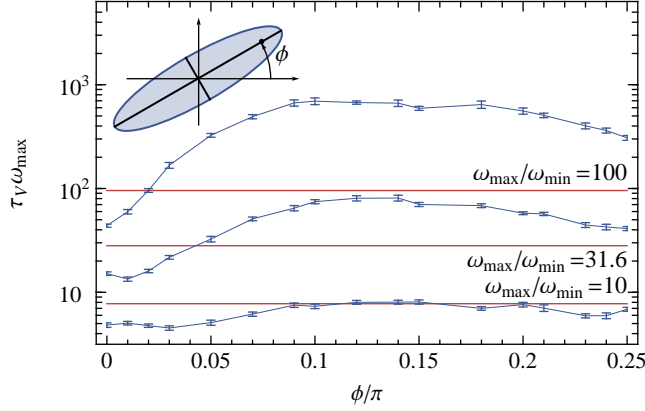
$$X_{ijkl} = \sum_{mn} \frac{O_{im} [\mathbf{O}^{-1} \mathbf{C}]_{ml} [\mathbf{O}^{-1} \mathbf{C}]_{nk} O_{jn}}{\alpha_m + \alpha_n} \quad (\text{D.3})$$

as $\tau_{ijkl} = \frac{1}{4} (X_{ijkl} + X_{ijlk} + X_{klij} + X_{lkij})$. For example - if we consider the full (q, p, \mathbf{s}) OU process in the harmonic case - one computes

$$\tau_H = \frac{\omega^4 \tau_{qqqq} + 2\omega^2 \tau_{qqpp} + \tau_{pppp}}{\omega^4 c_{qq}^2 + 2\omega^2 c_{qp}^2 + c_{pp}^2}, \quad \tau_V = \frac{\tau_{qqqq}}{c_{qq}^2} \quad (\text{D.4})$$

where we use an obvious notation for the indexes in τ_{ijkl} .

Figure 9: Correlation time for the potential energy of a 2-D harmonic oscillator, as a function of the angle between the eigenmodes and the Cartesian axes. τ_V is computed for different values of the condition number $\omega_{\max}/\omega_{\min}$, from bottom to top 10, 31.6 and 100. Thin lines serve as an aid for the eye, connecting the results obtained in the three cases using a massive NH chains thermostat, with four additional degrees of freedom and $Q = k_B T / \omega_{\max}^2$. Error bars are also shown for individual data points. Thick lines correspond to the (constant) result predicted for a GLE thermostat, using respectively the thermostat parameters fitted to give a flat, optimal τ_V over a frequency range between ω_{\min} and ω_{\max} . The values obtained in actual GLE simulations agree with the predictions within the statistical error bar, and are not reported.



E Comparison with Nosé-Hoover thermostat

The most widespread techniques for canonical sampling in MD are probably white-noise Langevin and Nosé-Hoover chains (NHC). White-noise Langevin can be considered as a limit case of the thermostating method we describe in this work, but NHC is based on a radically different philosophy. It is therefore worth performing a brief comparison between the latter and the GLE thermostat.

In the “massive” version of the NH thermostat[39, 45], each component of the physical momentum is coupled to an additional degree of freedom with a fictitious mass Q , by means of a second-order equation of motion. The resulting dynamics ensures that the physically-relevant degrees of freedom will sample the correct, constant-temperature ensemble, with the advantage of having deterministic equations of motion, and a well-defined conserved quantity. However, in the harmonic case, trajectories are poorly ergodic. This problem can be addressed by coupling the fictitious momentum to a second bath variable with a similar equation of motion. By repeating this process further a “Nosé-Hoover chain” can be formed, which ensures that the dynamics is sufficiently chaotic to achieve efficient sampling[40, 46]

$$\begin{aligned} \dot{q} &= p/m & \dot{p} &= -V'(q) - p s_1/Q & \dot{s}_1 &= [p^2/m - k_B T] - s_1 s_2/Q \\ & & \dot{s}_i &= [s_{i-1}^2/Q - k_B T] - s_i s_{i+1}/Q & \dot{s}_n &= [s_{n-1}^2/Q - k_B T] \end{aligned}$$

The drawback of this approach is that the thermostat equations are second-order in momenta. It is therefore difficult to obtain analytical predictions for the properties

of the dynamics, and the integration of the additional degrees of freedom must be performed with a multiple time-step approach, which makes the thermostat more expensive.

To examine the performances of NHC and GLE, one could envisage comparing the sampling efficiency as defined by the correlation times (3.1). Obtaining such estimates is not straightforward, not only because the harmonic case cannot be treated analytically, but also because in the multidimensional case the properties of the trajectory will not be invariant under an orthogonal transformation of coordinates. The simplest model we can conceive for comparing NHC and GLE is therefore a two-dimensional harmonic oscillator, with different vibrational frequencies on the two normal modes and adjustable relative orientations of the eigenvectors with respect to the thermostatted coordinates.

The resulting τ_V is reported in Figure 9: in the highly anisotropic cases, the efficiency of the NH chains depends dramatically on the orientation of the axes, while for well-conditioned problems is almost constant. The linear stochastic thermostat, on the other hand, has a predictable response, which is completely independent of orthogonal transforms of the coordinates. In the one-dimensional case, or when eigenvectors are perfectly aligned with axes, NH chains are very efficient for all modes with frequency $\omega < \omega_Q = \sqrt{\frac{k_B T}{Q}}$. One should however consider that, in the absence of an exact propagator, choosing a small Q implies that integration of the trajectory for the chains will become more expensive.

Obviously, such a simple toy model does not give quantitative information on the behavior in real-life cases, where modes of different frequencies coexist with anharmonicity and diffusive behavior. However, it demonstrates that the colored-noise Langevin thermostat performs almost as well as the axis-aligned NH chains. Furthermore, unlike the NHC, there are no unpredictable failures for anisotropic potentials.

F Momentum distribution in the harmonic limit

To recover an expression for the spherical average of the PMD in a harmonic solid, let us first consider a three-dimensional harmonic oscillator, with different modes oriented parallel to the Cartesian axes. For a start, we do not write the dependence of the Gaussian spread along the axes on the frequencies, but consider the widths as parameters, and write

$$n(\mathbf{p}) \propto e^{-\frac{1}{2} \left(\frac{p_x^2}{\bar{p}_x^2} + \frac{p_y^2}{\bar{p}_y^2} + \frac{p_z^2}{\bar{p}_z^2} \right)}.$$

With no loss of generality, we assume $\bar{p}_x \leq \bar{p}_y \leq \bar{p}_z$, write $p' = p/\bar{p}_z$, $\bar{p}_x = \bar{p}_y/\sqrt{1+\alpha^2}$, $\bar{p}_y = \bar{p}_z/\sqrt{1+\beta^2}$ and integrate away the angular dependence in polar coordinates, so as to get

$$n(p) dp = p'^2 \int d\phi d\theta e^{-\frac{1}{2} p'^2 [\cos^2 \theta + \sin^2 \theta (1+\beta^2)(1+\alpha^2 \cos^2 \phi)]} \sin \theta \bar{p}_z dp' \quad (\text{E1})$$

For the cylindrical symmetric limit $\alpha = 0$ we found a closed solution, and recovered a working expression for small xy -plane anisotropy by writing the first terms of a series expansion in terms of α^2 ,

$$n(p) \approx \frac{1}{p_z} e^{-p'^2/2} p' (I_0(p') + \alpha^2 I_1(p') + \alpha^4 I_2(p')) + \mathcal{O}(\alpha^6)$$

where

$$\begin{aligned} I_0(x) &= 2 \frac{1+\beta^2}{\sqrt{\pi}\beta} D(\beta x/\sqrt{2}) \\ I_1(x) &= \frac{1+\beta^2}{2\sqrt{\pi}\beta^3} \left[(1+\beta^2) \beta x/\sqrt{2} - D(\beta x/\sqrt{2}) (1-\beta^2 + \beta^2(1+\beta^2)x^2) \right] \\ I_2(x) &= \left[D(\beta x/\sqrt{2}) (9+10\beta^2-7\beta^4+2\beta^2(1+\beta^2)(3-\beta^2)p^2+3\beta^4(1+\beta^2)^2p^4) \right. \\ &\quad \left. - (1+\beta^2)(9+\beta^2+3\beta^2(1+\beta^2)x^2) \beta x/\sqrt{2} \right] \frac{1+\beta^2}{32\sqrt{\pi}\beta^5} \end{aligned}$$

and $D(x)$ is the Dawson integral $e^{-x^2} \int_0^x e^{y^2} dy$.

Quantum harmonic crystal We now move on to the treatment of the harmonic crystal. At first, we observe that the quantum momentum distribution can be written as a Gaussian, with frequency-dependent widths along the same normal modes which can be obtained by classical lattice dynamics. Then, let the indexes i, k, α label atoms, phonons and Cartesian coordinates respectively, and ω_k and $e_{i\alpha}^{(k)}$ be the frequency and the components of the k -th normalized eigenvector of the dynamical matrix. The covariance matrix which describes the multivariate-Gaussian distribution of momenta at temperature T reads

$$\langle p_{i\alpha} p_{j\beta} \rangle = \hbar \sqrt{m_i m_j} \sum_k e_{i\alpha}^{(k)\star} e_{j\beta}^{(k)} \frac{\omega_k}{2} \coth \frac{\hbar \omega_k}{2k_B T}.$$

In order to compute the spherically-averaged proton momentum distribution, one should then consider each hydrogen atom in the structure individually, and diagonalize the 3×3 matrix $\langle p_{i\alpha} p_{i\beta} \rangle$. The eigenvalues must then be plugged into Eq. (E1). The overall PMD is finally recovered by averaging over the probability distribution computed for all the protons.

Curriculum vitae

Michele Ceriotti was born in Busto Arsizio, Italy, on the 26th of July 1982. After having received his diploma at Liceo Scientifico E. Tosi (2001), he enrolled in Materials Science at the University of Milano - Bicocca. Here he obtained his B. Sc. in 2004, with a thesis on “*Ab-initio* calculation of the structural and vibrational properties of the crystalline phases of TeO_2 ”, and his M. Sc. in 2006, with a thesis concerning “*Ab-initio* simulations of diffusion and decomposition processes of SiH_x species on Si (100) surface”, both under the supervision of Prof. Marco Bernasconi. Since January 2007 he enrolled in a doctorate at the department of Physics at ETH Zürich, supervised by Prof. Michele Parrinello. His present research deals primarily with the development of methods to improve atomistic computer simulations, in particular by the use of stochastic differential equations on top of Hamiltonian dynamics. See p. 4 for a list of publications.

Bibliography

- [1] T. Schneider and E. Stoll, “Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions”, *Phys. Rev. B*, **17**, 1302 (1978).
- [2] S. A. Adelman and C. L. Brooks, “Generalized Langevin models and condensed-phase chemical reaction dynamics”, *J. Phys. Chem.*, **86**, 1511 (1982).
- [3] R. Zwanzig, “Memory Effects in Irreversible Thermodynamics”, *Phys. Rev.*, **124**, 983 (1961).
- [4] R. Zwanzig, *Nonequilibrium statistical mechanics* (Oxford University Press, New York, 2001).
- [5] H. Mori, “Transport, Collective Motion, and Brownian Motion”, *Prog. Th. Phys.*, **33**, 423 (1965).
- [6] C. C. Martens, “Qualitative dynamics of generalized Langevin equations and the theory of chemical reaction rates”, *J. Chem. Phys.*, **116**, 2516 (2002).
- [7] J. Wang, “Quantum Thermal Transport from Classical Molecular Dynamics”, *Phys. Rev. Lett.*, **99**, 160601 (2007).
- [8] L. Kantorovich, “Generalized Langevin equation for solids. I. Rigorous derivation and main properties”, *Phys. Rev. B*, **78**, 094304 (2008).
- [9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of State Calculations by Fast Computing Machines”, *J. Chem. Phys.*, **21**, 1087 (1953).
- [10] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Second ed. (Academic Press, London, 2002).
- [11] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, “Hybrid Monte Carlo”, *Phys. Lett. B*, **195**, 216 (1987).
- [12] T. Soddemann, B. Dünweg, and K. Kremer, “Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations”, *Phys. Rev. E*, **68**, 046702 (2003).

- [13] G. Bussi, D. Donadio, and M. Parrinello, “Canonical sampling through velocity rescaling”, *J. Chem. Phys.*, **126**, 014101 (2007).
- [14] A. Einstein, “Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen”, *Ann. Phys.*, **322**, 549 (1905).
- [15] P. Langevin, “The theory of Brownian movement”, *CR Acad. Sci*, **146**, 530 (1908).
- [16] C. W. Gardiner, *Handbook of Stochastic Methods*, Third ed. (Springer, Berlin, 2003).
- [17] R. L. Stratonovich, *Topics in the Theory of Random Noise*, Vol. 1 (Gordon and Breach, New York, 1963).
- [18] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1996).
- [19] M. J. Levin, “Generation of a Sampled Gaussian Time Series Having a Specified Correlation Function”, *IRE Trans. Inf. Th.*, **6**, 545 (1960).
- [20] A. A. Maradudin, T. Michel, A. R. McGurn, and E. R. Méndez, “Enhanced backscattering of light from a random grating”, *Ann. Phys.*, **203**, 255 (1990).
- [21] F. Marchesoni and P. Grigolini, “On the extension of the Kramers theory of chemical relaxation to the case of nonwhite noise”, *J. Chem. Phys.*, **78**, 6287 (1983).
- [22] J. Łuczka, “Non-Markovian stochastic processes: Colored noise”, *Chaos*, **15**, 026107 (2005).
- [23] H. Mori, “A Continued-Fraction Representation of the Time-Correlation Functions”, *Prog. Th. Phys.*, **34**, 399 (1965).
- [24] R. Zwanzig, “Nonlinear generalized Langevin equations”, *J. Stat. Phys.*, **9**, 215 (1973).
- [25] G. W. Ford, J. T. Lewis, and R. F. O’Connell, “Quantum Langevin equation”, *Phys. Rev. A*, **37**, 4419 (1988).
- [26] S. Buyukdagli, A. V. Savin, and B. Hu, “Computation of the temperature dependence of the heat capacity of complex molecular systems using random color noise”, *Phys. Rev. E*, **78**, 066702 (2008).
- [27] J.-D. Bao, Y. Zhou, and K. Lü, “Anomalous diffusion in periodic potentials under self-similar colored noise”, *Phys. Rev. E (Statistical, Nonlinear, and Soft Matter Physics)*, **74**, 041125 (2006).
- [28] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by Simulated Annealing”, *Science, New Series*, **220**, 671 (1983).
- [29] J. A. Nelder and R. Mead, “A Simplex Method for Function Minimization”, *Comp. J.*, **7**, 308 (1965).

- [30] M. Tuckerman, B. J. Berne, and G. J. Martyna, "Reversible multiple time scale molecular dynamics", *J. Chem. Phys.*, **97**, 1990 (1992).
- [31] G. Bussi and M. Parrinello, "Accurate sampling using Langevin dynamics", *Phys. Rev. E*, **75**, 056707 (2007).
- [32] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, "Efficient and general algorithms for path integral Car-Parrinello molecular dynamics", *J. Chem. Phys.*, **104**, 5579 (1996).
- [33] S. Jang and G. A. Voth, "Simple reversible molecular dynamics algorithms for Nosé-Hoover chain dynamics", *J. Chem. Phys.*, **107**, 9514 (1997).
- [34] R. F. Fox, I. R. Gatland, R. Roy, and G. Vemuri, "Fast, accurate algorithm for numerical simulation of exponentially correlated colored noise", *Phys. Rev. A*, **38**, 5938 (1988).
- [35] D. T. Gillespie, "Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral", *Phys. Rev. E*, **54**, 2084 (1996).
- [36] F. Bruneval, D. Donadio, and M. Parrinello, "Molecular Dynamics Study of the Solvation of Calcium Carbonate in Water", *J. Phys. Chem. B*, **111**, 12219 (2007).
- [37] B. Ensing, S. O. Nielsen, P. B. Moore, M. L. Klein, and M. Parrinello, "Energy Conservation in Adaptive Hybrid Atomistic/Coarse-Grain Molecular Dynamics", *J. Chem. Theory Comput.*, **3**, 1100 (2007).
- [38] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford University Press, USA, 1990).
- [39] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods", *J. Chem. Phys.*, **81**, 511 (1984).
- [40] M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, "Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals", *J. Chem. Phys.*, **99**, 2796 (1993).
- [41] G. Bussi and M. Parrinello, "Stochastic thermostats: comparison of local and global schemes", *Comp. Phys. Comm.*, **179**, 26 (2008).
- [42] H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature", *J. Chem. Phys.*, **72**, 2384 (1980).
- [43] M. Ceriotti, G. Bussi, and M. Parrinello, "Colored noise Thermostats à la carte", *Journal of Chemical Theory and Computation*, **6**, 1170 (2010).
- [44] M. Ceriotti, G. Bussi, and M. Parrinello, "Langevin Equation with Colored Noise for Constant-Temperature Molecular Dynamics Simulations", *Phys. Rev. Lett.*, **102**, 020601 (2009).
- [45] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions", *Phys. Rev. A*, **31**, 1695 (1985).

- [46] G. J. Martyna, M. E. Tuckerman, and M. L. Klein, "Nosé-Hoover chains: The canonical ensemble via continuous dynamics", *J. Chem. Phys.*, **97**, 2635 (1992).
- [47] R. Car and M. Parrinello, "Unified Approach for Molecular Dynamics and Density-Functional Theory", *Phys. Rev. Lett.*, **55**, 2471 (1985).
- [48] D. Marx and J. Hutter, "Ab initio molecular dynamics: theory and implementation", in *Modern Methods and Algorithms of Quantum Chemistry*, Vol. 1, edited by J. Grotendorst (NIC Series, Jülich, Germany, 2000) pp. 301–449.
- [49] M. Sprik, "Computer simulation of the dynamics of induced polarization fluctuations in water", *J. Phys. Chem.*, **95**, 2283 (1991).
- [50] S. W. Rick, S. J. Stuart, and B. J. Berne, "Dynamical fluctuating charge force fields: Application to liquid water", *J. Chem. Phys.*, **101**, 6141 (1994).
- [51] L. Rosso, P. Mináry, Z. Zhu, and M. E. Tuckerman, "On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles", *J. Chem. Phys.*, **116**, 4389 (2002).
- [52] J. VandeVondele and U. Rothlisberger, "Canonical Adiabatic Free Energy Sampling (CAFES): A Novel Method for the Exploration of Free Energy Surfaces", *J. Phys. Chem. B*, **106**, 203 (2002).
- [53] L. Maragliano and E. Vanden-Eijnden, "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations", *Chem. Phys. Lett.*, **426**, 168 (2006).
- [54] P. E. Blöchl and M. Parrinello, "Adiabaticity in first-principles molecular dynamics", *Phys. Rev. B*, **45**, 9413 (1992).
- [55] P. Tangney and S. Scandolo, "How well do Car–Parrinello simulations reproduce the Born–Oppenheimer surface?: Theory and examples", *J. Chem. Phys.*, **116**, 14 (2002).
- [56] P. H. L. Sit and N. Marzari, "Static and dynamical properties of heavy water at ambient conditions from first-principles molecular dynamics", *J. Chem. Phys.*, **122**, 204510 (2005).
- [57] <http://www.cpmc.org/>, Copyright IBM Corp 1990-2006, Copyright MPI für Festkörperforschung Stuttgart 1997-2001.
- [58] D. J. Tobias, G. J. Martyna, and M. L. Klein, "Molecular dynamics simulations of a protein in the canonical ensemble", *J. Phys. Chem.*, **97**, 12959 (1993).
- [59] S. Habershon, T. E. Markland, and D. E. Manolopoulos, "Competing quantum effects in the dynamics of a flexible water model.", *J. Chem. Phys.*, **131**, 024501 (2009).
- [60] W. Smith, T. R. Forester, I. T. Todorov, and M. Leslie, "DL-POLY", CCLRC Daresbury Laboratory, Daresbury, Warrington, UK (2006).

- [61] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten, "Molecular dynamics study of unbinding of the avidin-biotin complex", *Biophys. J.*, **72**, 1568 (1997).
- [62] A. Laio and M. Parrinello, "Escaping free-energy minima", *Proc. Natl. Acad. Sci. USA*, **99**, 12562 (2002).
- [63] M. Iannuzzi, A. Laio, and M. Parrinello, "Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics", *Phys. Rev. Lett.*, **90**, 238302 (2003).
- [64] A. Laio, A. Rodriguez-Forte, F. L. Gervasio, M. Ceccarelli, and M. Parrinello, "Assessing the accuracy of metadynamics", *J. Phys. Chem. B*, **109**, 6714 (2005).
- [65] A. Barducci, G. Bussi, and M. Parrinello, "Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method", *Phys. Rev. Lett.*, **100**, 020603 (2008).
- [66] A. Laio and F. L. Gervasio, "Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science", *Rep. Prog. Phys.*, **71**, 126601 (2008).
- [67] G. Bussi, A. Laio, and M. Parrinello, "Equilibrium free energies from nonequilibrium metadynamics", *Physical review letters*, **96**, 90601 (2006).
- [68] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation", *J. Chem. Theory Comput.*, **4**, 435 (2008).
- [69] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, "PLUMED: A portable plugin for free-energy calculations with molecular dynamics", *Comp. Phys. Comm.*, **180**, 1961 (2009).
- [70] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water", *J. Chem. Phys.*, **79**, 926 (1983).
- [71] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids", *J. Am. Chem. Soc.*, **118**, 11225 (1996).
- [72] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1964).
- [73] D. M. Ceperley and E. L. Pollock, "Path-integral computation of the low-temperature properties of liquid/sup 4/He", *Phys. Rev. Lett.*, **56**, 351 (1986).
- [74] M. Parrinello and A. Rahman, "Study of an F center in molten KCl", *J. Chem. Phys.*, **80**, 860 (1984).

- [75] I. R. Craig and D. E. Manolopoulos, "Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics", *J. Chem. Phys.*, **121**, 3368 (2004).
- [76] J. O. Richardson and S. C. Althorpe, "Ring-polymer molecular dynamics rate-theory in the deep-tunneling regime: Connection with semiclassical instanton theory", *J. Chem. Phys.*, **131**, 214106 (2009).
- [77] T. E. Markland and D. E. Manolopoulos, "A refined ring polymer contraction scheme for systems with electrostatic interactions", *Chem. Phys. Lett.*, **464**, 256 (2008).
- [78] X. W. Zhou, J. A. Zimmerman, B. M. Wong, and J. J. Hoyt, "An embedded-atom method interatomic potential for Pd-H alloys", *J. Mat. Res.*, **23**, 704 (2008).
- [79] L. Wang and A. Zunger, "Solving Schrödinger's equation around a desired energy: Application to silicon quantum dots", *J. Chem. Phys.*, **100**, 2394 (1994).
- [80] D. M. Ceperley, "Path integrals in the theory of condensed helium", *Rev. Mod. Phys.*, **67**, 279 (1995).
- [81] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, "The nature of the hydrated excess proton in water", *Nature*, **397**, 601 (1999).
- [82] G. A. Voth, "Calculation of equilibrium averages with Feynman-Hibbs effective classical potentials and similar variational approximations", *Phys. Rev. A*, **44**, 5302 (1991).
- [83] R. Alimi, A. García-Vela, and R. B. Gerber, "A remedy for zero-point energy problems in classical trajectories: A combined semiclassical/classical molecular dynamics algorithm", *J. Chem. Phys.*, **96**, 2034 (1992).
- [84] Y. Guo, D. L. Thompson, and T. D. Sewell, "Analysis of the zero-point energy problem in classical trajectory simulations", *J. Chem. Phys.*, **104**, 576 (1996).
- [85] J. S. Bader and B. J. Berne, "Quantum and classical relaxation rates from classical simulations", *J. Chem. Phys.*, **100**, 8359 (1994).
- [86] G. Stock, "Classical Simulation of Quantum Energy Flow in Biomolecules", *Phys. Rev. Lett.*, **102**, 118301 (2009).
- [87] A. Schmid, "On a quasiclassical Langevin equation", *J. Low Temp. Phys.*, **49**, 609 (1982).
- [88] C. W. Gardiner, "Quantum noise and quantum Langevin equations", *IBM J. Res. Dev.*, **32**, 127 (1988).
- [89] M. Ceriotti, G. Bussi, and M. Parrinello, "Nuclear Quantum Effects in Solids Using a Colored-Noise Thermostat", *Phys. Rev. Lett.*, **103**, 030603 (2009).

- [90] J. A. Morrone, V. Srinivasan, D. Sebastiani, and R. Car, "Proton momentum distribution in water: an open path integral molecular dynamics study", *J. Chem. Phys.*, **126**, 234504 (2007).
- [91] J. A. Morrone and R. Car, "Nuclear Quantum Effects in Water", *Phys. Rev. Lett.*, **101**, 017801 (2008).
- [92] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, "Theory of Quantum Annealing of an Ising Spin Glass", *Science, New Series*, **295**, 2427 (2002).
- [93] Y.-H. Lee and B. J. Berne, "Global Optimization: Quantum Thermal Annealing with Path Integral Monte Carlo", *J. Phys. Chem. A*, **104**, 86 (2000).
- [94] G. E. Santoro and E. Tosatti, "Optimization using quantum mechanics: quantum annealing through adiabatic evolution", *J. Phys. A*, **39**, R393 (2006).
- [95] G. Jacucci and E. Omerti, "Monte Carlo calculation of the radial distribution function of quantum hard spheres at finite temperatures using path integrals with boundary conditions", *J. Chem. Phys.*, **79**, 3051 (1983).
- [96] S. Habershon and D. E. Manolopoulos, "Zero point energy leakage in condensed phase dynamics: An assessment of quantum simulation methods for liquid water", *J. Chem. Phys.*, **131**, 244518 (2009).
- [97] J. Tersoff, "Modeling solid-state chemistry: Interatomic potentials for multi-component systems", *Phys. Rev. B*, **39**, 5566 (1989).
- [98] K. Singer and W. Smith, "Path integral simulations of condensed phase Lennard-Jones systems", *Mol. Phys.*, **64**, 1215 (1988).
- [99] C. P. Herrero and R. Ramírez, "Structural and thermodynamic properties of diamond: A path-integral Monte Carlo study", *Phys. Rev. B*, **63**, 024103 (2000).
- [100] P. Chen, Z. Xiong, J. Luo, J. Lin, and K. L. Tan, "Interaction of hydrogen with metal nitrides and imides", *Nature*, **420**, 302 (2002).
- [101] P. Chen and M. Zhu, "Recent progress in hydrogen storage", *Materials Today*, **11**, 36 (2008).
- [102] D. H. Gregory, "Lithium nitrides, imides and amides as lightweight, reversible hydrogen stores", *J. Mat. Chem.*, **18**, 2321 (2008).
- [103] S. Orimo, Y. Nakamori, J. R. Eliseo, A. Zuttel, and C. M. Jensen, "Complex hydrides for hydrogen storage", *Chem. Rev.*, **107**, 4111 (2007).
- [104] P. J. Haigh, R. A. Forman, and R. C. Frisch, "Nuclear Magnetic Resonance of Li and H in Solid Lithium Imide, Lithium Amide, and Lithium Nitride", *J. Chem. Phys.*, **45**, 812 (1966).
- [105] R. A. Forman, "Order-Disorder Transition in Lithium Imide, LiNH", *J. Chem. Phys.*, **55**, 1987 (1971).

- [106] M. P. Balogh, C. Y. Jones, J. F. Herbst, L. G. Hector, *et al.*, “Crystal structures and phase transformation of deuterated lithium imide, Li_2ND ”, *J. Alloys Compd.*, **420**, 326 (2006).
- [107] J. Rijssenbeek, Y. Gao, J. Hanson, Q. Huang, C. Jones, and B. Toby, “Crystal structure determination and reaction pathway of amide-hydride mixtures”, *J. Alloys Compd.*, **454**, 233 (2008).
- [108] J. F. Herbst and L. G. Hector Jr, “Energetics of the Li amide/Li imide hydrogen storage reaction”, *Phys. Rev. B*, **72**, 125120 (2005).
- [109] L. G. Hector Jr and J. F. Herbst, “Density functional theory for hydrogen storage materials: successes and opportunities”, *J. Phys. Cond. Mat.*, **20**, 064229 (2008).
- [110] T. Mueller and G. Ceder, “Effective interactions between the NH bond orientations in lithium imide and a proposed ground-state structure”, *Phys. Rev. B*, **74**, 134104 (2006).
- [111] B. Magyari-Köpe, V. Ozoliņš, and C. Wolverton, “Theoretical prediction of low-energy crystal structures and hydrogen storage energetics in Li_2NH ”, *Phys. Rev. B*, **73**, 220101 (2006).
- [112] G. Miceli, M. Ceriotti, M. Bernasconi, and M. Parrinello, “Static disorder and structural correlations in the low temperature phase of lithium imide”, submitted (2010).
- [113] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized Gradient Approximation made simple”, *Phys. Rev. Lett.*, **77**, 3865 (1996).
- [114] D. Vanderbilt, “Soft self-consistent pseudopotentials in a generalized eigenvalue formalism”, *Phys. Rev. B*, **41**, 7892 (1990).
- [115] S. Goedecker, M. Teter, and J. Hutter, “Separable dual-space Gaussian pseudopotentials”, *Phys. Rev. B*, **54**, 1703 (1996).
- [116] C. M. Araújo, A. Blomqvist, R. H. Scheicher, P. Chen, and R. Ahuja, “Superionicity in the hydrogen storage material Li_2NH : Molecular dynamics simulations”, *Phys. Rev. B*, **79**, 172101 (2009).
- [117] C. Zhang, M. Dyer, and A. Alavi, “Quantum delocalization of hydrogen in the Li_2NH crystal”, *J. Phys. Chem. B*, **109**, 22089 (2005).
- [118] G. Reiter, J. C. Li, J. Mayers, T. Abdul-Redah, and P. Platzman, “The proton momentum distribution in water and ice”, *Braz. J. Phys.*, **34**, 142 (2004).
- [119] C. Andreani, D. Colognesi, J. Mayers, G. F. Reiter, and R. Senesi, “Measurement of momentum distribution of light atoms and molecules in condensed matter systems using inelastic neutron scattering”, *Adv. Phys.*, **54**, 377 (2005).
- [120] C. Andreani, E. Degiorgi, R. Senesi, F. Cilloco, D. Colognesi, J. Mayers, M. Nardone, and E. Pace, “Single particle dynamics in fluid and solid hydrogen sulphide: An inelastic neutron scattering study”, *J. Chem. Phys.*, **114**, 387 (2001).

- [121] J. Mayers, G. F. Reiter, and P. Platzman, "Measurement of proton momentum distributions by eV neutron scattering", *Journal of Molecular Structure*, **615**, 275 (2002).
- [122] G. F. Reiter, J. Mayers, and J. Noreland, "Momentum-distribution spectroscopy using deep inelastic neutron scattering", *Phys. Rev. B*, **65**, 104305 (2002).
- [123] R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, "Graphite-diamond phase coexistence study employing a neural-network mapping of the ab initio potential energy surface", *Phys. Rev. B*, **81**, 100103 (2010).
- [124] "CP2K", <http://cp2k.berlios.de>.

Acknowledgments

This thesis collects a good deal of the things I've learned and of the work I've been doing during the past three years of my life. Much more has happened, and much more I've learned, but it is impossible to mention it all. I can only try my best to express my gratitude to all the people who guided me through this journey, and were even more helpful just by staying by my side.

Much I owe to Michele Parrinello, who did way more than he was supposed to as a supervisor, by any standard. He gave me as much guidance as freedom, being supportive to my research and sharing with me bits of his massive experience about both science and research. Every single time I entered his office I was sure to end up with a - sometimes heated - discussion about what I was doing, a reference to an obscure paper of the eighties which solved exactly my problem, or just some motivation to get back and try a bit harder. Then I need to thank Marco Bernasconi, again: this is the third time I mention him in a thesis! I enjoyed the possibility of working with him for six years in a row, and at every interaction he managed to transfer me some momentum. As if two outstanding supervisors in a row were not enough, I am glad to know that the next one coming will continue my lucky series. I'm in debt to David Manolopoulos for the help and the encouragement he gave me since we met last May, and for his exquisite kindness. Working with him on a couple of projects this year was as instructing as pleasant and I am looking forward to begin more learning and work, next year in Oxford.

The subject of this dissertation would be quite different if I had not met Giovanni Bussi. It all started with "a couple of calculations with pink noise", then the colors ended up covering a much larger gamut. I owe him a big slice of what I learned during this doctorate, and I think that anyone who had the chance of discussing science with him agrees that his clear explanations are second only to his fancy movies (and to his gnocco fritto). A big thank you is also due to Thomas Kühne: I would have been scared away by all the nasty algebra which plagues these pages if I hadn't sweated for two years on sparse matrices and linear scaling. I would also like to express my gratitude to Tom Markland, who worked with me on thermostating PIMD, Michela Ottobre and Grigorios Pavliotis with whom I discussed a lot about GLEs and who pointed me at useful references.

Sir Gareth Aneurin Tribello deserves a whole paragraph and at least a barrel of ale for having tried really hard to teach me how to write as a real British. Were not for him, this thesis would look like Italian processed by Google translator. I haven't learned it all, but at least I know that. I. Should. Not. Spare. Full. Stops. While I am at

it, I also want to thank him for convincing me that English food is not THAT bad, for the beers at “Tra” and endless political discussions.

One of the main reasons I try to make some science is that I like the many, diverse, interesting people one meets in the process. Three and a half years in Lugano allowed me to interact with so many exceptional people that I fear that I’ll leave out someone really important. Nevertheless, I want to take the risk and list a few, who contributed in some way to this thesis, by sharing with me their impressions and insight. Giacomo, Paolo, Max, Ivan, Stefano, Hagai, Rustam, Oliviero, Davide, Tilde, Alessandro, Carlo, Monica, thank you! I would also like express my gratitude to all the people who were in the lab during these years, making it a pleasant and stimulating environment. Many thanks go to Daniela, for having taken care of most of the paperwork, leaving me much more time and fewer worries.

Sure I cannot leave out my friends, those in Milan and those in Busto! Many are leaving, many are staying, but you are all making this world a better place to live. You are one of the few things I’ll miss about Italy.

Finally, I owe everything to my parents and my not-so-little sister. They are always there to help, and will never get from me all the credit they deserve. Thank you!

To Lu, who stayed by my side, pushing me to do my best and keeping her temper even when I was at my worst: I don’t want to say thanks, I just want to say yes.

Molecular dynamics simulations provide insight into the microscopic properties of a variety of systems, in areas ranging from biochemistry to materials science, by reproducing on a computer the motion of their atoms. Stochastic differential equations have been used extensively in this context, both to model physical effects and as a computational tool to make the exploration of configurational space more effective.

Here a comprehensive framework is introduced, based on generalized Langevin equations, which allows one to enhance molecular dynamics simulations in a controlled, predictable fashion. Based on this approach, a number of applications is described, along with demonstrations on realistic problems. By tuning the correlation of the noise in the generalized Langevin equation, it is possible to enhance or otherwise modulate the sampling properties of constant-temperature molecular dynamics. One can also independently control the effective temperature of individual normal modes, ultimately obtaining an inexpensive method to model zero-point energy and related nuclear quantum effects.