Stochastic algorithms for meso-scale constant temperature simulations

Elena Akhmatskaya
Fujitsu Laboratories of Europe Limited

Sebastian Reich
Universität Potsdam, Institut für Mathematik

June 3, 2009

Contents

1 Introduction

Classical molecular dynamics (MD) simulations of all-atoms or coarse grained systems are among the most popular techniques for simulation of soft matter. Classical MD simulations are naturally performed under the conditions of constant energy E, constant volume V, and constant number of particles N. One refers to such simulations as microcanonical or NVE ensemble simulations (Allen and Tildesley, 1987; Frenkel and Smit, 2001). In addition to energy, volume and number of particles, microcanonical ensemble simulations also conserve linear and angular momentum, with the later not being a constant under periodic boundary conditions. Unfortunately, the microcanonical (NVE) ensemble does not correspond to the conditions under which most experiments are carried out. If one is interested in the behavior of the system at a specific temperature T, a constant NVT ensemble simulation using a thermostat is required. See, for example, Allen and Tildesley (1987); Frenkel and Smit (2001) for a description of popular thermostats. These thermostats can be classified into the following three categories:

1. scaling velocities (e.g., simple velocity scaling and the Berendsen thermostat),

2. stochastic thermostats (e.g., the Anderson thermostat, Langevin and Dissipative Particle dynamics (DPD))

3. extended system formulations (e.g., the Nosé-Hoover thermostat).

Each class of methods has advantages and disadvantages, depending on the application. In this paper, we will focus on stochastic thermostats and their numerical implementation since these methods appear to be robust and accurate in a wide range of applications. Furthermore, we restrict the discussion to local stochastic thermostats such as Langevin and dissipative particle dynamics (Frenkel and Smit, 2001). However, the techniques proposed in this paper essentially also apply to global stochastic thermostats such as the recently proposed Hoover-Langevin thermostat (Samoletov et al., 2007; Leimkuhler et al., 2009) and the global Langevin model of Bussi et al. (2007).

Note that local thermostats, such as Langevin dynamics, can be viewed as simplified reduced systems in the sense of Mori-Zwanzig reduction (see, for example, Allen and Tildesley (1987)). This aspect is particularly important for coarse grained models for which the fluctuation-dissipation contributions should not only keep the system at a desired target temperature but should also mimic the impact of non-resolved finer details of an all-atom model on the coarse grained length and time scales. Under those circumstances, the appropriate numerical treatment of fluctuation-dissipation terms also gains in importance.

The numerical methods proposed in this paper rely on the well-known splitting of stochastic thermostat equations into a conservative and a fluctuation-dissipation part (Allen and Tildesley, 1987). We propose a methodology to derive numerical approximation to the fluctuation-dissipation part that exactly samples from the underlying Boltzmann distribution. Our methodology applies to Langevin as well as dissipative particle dynamics (DPD) and, more generally, to arbitrary position

dependent fluctuation-dissipation terms. A composition method approach is used to derive a time-stepping method for the complete thermostat, where the conservative dynamics part is discretized by the standard Störmer-Verlet method (Allen and Tildesley, 1987; Leimkuhler and Reich, 2005). A Metropolis criterion is introduced to correct for inconsistency in the Störmer-Verlet time-stepping method with the underlying NVT ensemble and puts the resulting propagator into the context of Markov chain Monte Carlo (MCMC) methods (Liu, 2001). Our approach relies essentially on an appropriate adaptation of the generalized hybrid Monte Carlo method of Horowitz (1991); Kennedy and Pendleton (2001). Acceptance rates under the Metropolis criterion can be increased using shadow energies as demonstrated in the work of Akhmatskaya and Reich (2006, 2008). The proposed methods are applicable to meso-scale and coarse grained molecular dynamics simulations which require a highly accurate sampling from the given target NVT ensemble.

Following Akhmatskaya and Reich (2008), the techniques proposed in this paper can be combined with constant temperature T, pressure P, and number of particles N (NPT) ensemble simulation techniques such as proposed by Andersen (1980); Feller et al. (1995).

An outline of the paper is as follows. The stochastic equations of second-order Langevin and dissipative particle dynamics (DPD) are summarized in Section 2. In this section, we also state a unifying framework for general local stochastic thermostats which lead to position-dependent fluctuation-dissipation terms. We also provide a summary of state of the art numerical methods for stochastic thermostats. Section 3 gives a short summary on essential theoretical background concepts such as Markov chains, geometric ergodicity, Markov chain Monte Carlo (MCMC) methods and shadow energies for conservative time-stepping methods such as Störmer-Verlet. Numerical methods which exactly sample from the velocity Boltzmann distribution for force-free motion are proposed in Section 4, where we also discuss their composition with the Störmer-Verlet method for conservative dynamics. Since the composed approximation does not exactly sample from the NVT ensemble, Metropolis corrected implementations along the generalized hybrid Monte Carlo method is proposed in Section 5. In Section 5, we also discuss samples with respect to shadow energies which allows one to increase the acceptance rate under the Störmer-Verlet dynamics. Numerical results for some standard DPD test sets as well as from a membrane protein simulation are provided in Section 6. Concluding remarks can be found in Section 7.

2 Summary of local stochastic thermostat formulations

We consider an N-atom molecular system with atomistic positions $\mathbf{r}_i = (x_i, y_i, z_i)^T \in \mathbb{R}^3$, velocities $\mathbf{v}_i = (u_i, v_i, w_i)^T \in \mathbb{R}^3$, masses m_i, momenta $\mathbf{p}_i = m_i \mathbf{v}_i \in \mathbb{R}^3$, potential energy function V, and total energy function

$$E = \frac{1}{2} \sum_{i=1}^{N} \frac{\|\mathbf{p}_i\|^2}{m_i} + V(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N). \tag{1}$$

For easy of reference, we introduce the notations $\mathbf{r} = (\mathbf{r}_1^T, \ldots, \mathbf{r}_N^T)^T \in \mathbb{R}^{3N}$, $\mathbf{v} = (\mathbf{v}_1^T, \ldots, \mathbf{v}_N^T)^T \in \mathbb{R}^{3N}$, $\mathbf{p} = (\mathbf{p}_1^T, \ldots, \mathbf{p}_N^T)^T \in \mathbb{R}^{3N}$, and a diagonal mass matrix $M \in \mathbb{R}^{3N \times 3N}$ such that

$$\mathbf{p}^T M^{-1} \mathbf{p} = \sum_{i=1}^{N} m_i^{-1} \|\mathbf{p}_i\|^2. \tag{2}$$

The microcanonical equations of motion are

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = M^{-1}\mathbf{p}, \tag{3}$$

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\nabla_{\mathbf{r}} V(\mathbf{r}), \tag{4}$$

which conserve total energy (1).

Recall now that the instantaneous value of the temperature T is related to the kinetic energy as follows:

$$\frac{k_B T}{2}(3N - N_c) = \frac{1}{2}\sum_{i=1}^{N} \frac{\|\mathbf{p}_i\|^2}{m_i}, \tag{5}$$

where k_B denotes Boltzmann's constant and N_c is the number of additionally conserved quantities. Since total linear and angular momentum are preserved for open boundary MD simulations we obtain $N_c = 6$, while only total linear momentum is preserved under periodic boundary conditions and $N_c = 3$.

We now introduce a unifying framework for locally coupled stochastic thermostats, which allow one to convert (3)-(4) to formulations suitable for constant temperature (NVT ensemble) simulations.

2.1 Langevin and dissipative particle dynamics (DPD)

Dissipative particle dynamics (DPD) has become a very popular method for meso-scale simulations of materials. In this section, we provide a short summary of the method and discuss a slightly more general framework that also includes traditional Langevin dynamics.

Following the notation of Español and Warren (1995), the standard DPD method of Hoogerbrugge and Koelman (1992) can be formulated as a stochastic differential equation (SDE):

$$\mathrm{d}\mathbf{r}_i = \frac{\mathbf{p}_i}{m_i}\,\mathrm{d}t, \tag{6}$$

$$\mathrm{d}\mathbf{p}_i = \left[\mathbf{F}_i - \gamma\sum_{j\neq i}\omega(r_{ij})(\mathbf{e}_{ij}\cdot\mathbf{v}_{ij})\mathbf{e}_{ij}\right]\mathrm{d}t + \sigma\sum_{j\neq i}\omega^{1/2}(r_{ij})\mathbf{e}_{ij}\,\mathrm{d}W_{ij}, \tag{7}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, $\mathbf{e}_{ij} = \mathbf{r}_{ij}/r_{ij}$, $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$, and $\mathbf{F}_i = -\nabla_{\mathbf{r}_i} V(\mathbf{r})$ is the conservative force acting on particle i. The dimensionless weight function $\omega(r)$ can be chosen in a rather arbitrary manner. However, to reproduce a constant temperature ensemble, the friction coefficient γ and the noise amplitude σ have to satisfy the fluctuation dissipation relation

$$\sigma = \sqrt{2k_B T\gamma}. \tag{8}$$

Finally, $W_{ij}(t) = W_{ji}(t)$ are independent Wiener processes (Oksendal, 2000). Recall that the finite-time increments $\Delta W_{ij}(\tau) = W_{ij}(t+\tau) - W_{ij}(t)$ of a Wiener process are Gaussian distributed with mean zero and variance $\sqrt{\tau}$, i.e., $\Delta W_{ij}(\tau) \sim \mathrm{N}(0,\tau)$. Here $\mathrm{N}(a,b)$ denotes the Gaussian distribution with mean a and variance b.

Following Cotter and Reich (2003), let us write the equations (6)-(7) in a more compact and general manner:

$$\mathrm{d}\mathbf{r} = M^{-1}\mathbf{p}\,\mathrm{d}t, \tag{9}$$

$$\mathrm{d}\mathbf{p} = -\nabla_{\mathbf{r}} V(\mathbf{r})\,\mathrm{d}t - \sum_{k=1}^{K}\nabla_{\mathbf{r}} h_k(\mathbf{r})\left[\gamma\dot{h}_k(\mathbf{r})\,\mathrm{d}t - \sigma\,\mathrm{d}W_k\right], \tag{10}$$

where

$$\dot{h}_k(\mathbf{r}) = \nabla_{\mathbf{r}} h_k(\mathbf{r})\cdot\mathbf{v} = \nabla_{\mathbf{r}} h_k(\mathbf{r})\cdot M^{-1}\mathbf{p}, \tag{11}$$

and the functions $h_k(\mathbf{r})$, $k = 1,\ldots,K$, can be chosen quite arbitrarily.

The choice

$$h_k(\mathbf{r}) = \phi(r_{ij}), \qquad \phi'(r) = \omega^{1/2}(r), \qquad k = 1,\ldots,K, \tag{12}$$

with $K = (N-1)N/2$ in (10) leads back to the standard DPD model. However, one can also set $K = 3N$ and

$$h_i(\mathbf{r}) = m_i^{1/2} x_i, \qquad h_{i+N}(\mathbf{r}) = m_i^{1/2} y_i, \qquad h_{i+2N}(\mathbf{r}) = m_i^{1/2} z_i, \tag{13}$$

$i = 1, \ldots, N$, in (10), which leads to the standard Langevin model

$$\begin{aligned} \mathrm{d}\mathbf{r} &= M^{-1}\mathbf{p}\,\mathrm{d}t, & (14)\\ \mathrm{d}\mathbf{p} &= -\nabla_{\mathbf{r}} V(\mathbf{r})\,\mathrm{d}t - \gamma\mathbf{p}\,\mathrm{d}t + \sigma M^{1/2}\,\mathrm{d}\mathbf{W}, & (15) \end{aligned}$$

where $\mathbf{W}(t) = (W_1(t), \ldots, W_N(t))^T$ is now a N-dimensional Wiener process.

An intriguing aspect of the DPD equations (6)-(7) is that they satisfy Newton's third law which implies conservation total linear momentum

$$\mathbf{P} = \sum_{i=1}^{N} \mathbf{p}_i \tag{16}$$

as well as total angular momentum

$$\mathbf{L} = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{p}_i. \tag{17}$$

The same is not true for the Langevin equations (14)-(15). Conservation of total linear and angular momentum has made DPD a popular method for meso-scale simulations (see, e.g., Español (2003)) as well as a stochastic thermostat for NVT simulations (Koopman and Lowe, 2006).

2.2 State of the art numerical treatment

Appropriate time-stepping methods for the conservative part (3)-(4) form the basis of any time-stepping methods for stochastic thermostats. It seems that there is widespread agreement that the conservative part should be solved numerically by the Störmer-Verlet method, which is written here in the velocity/momentum formulation:

$$\begin{aligned} \mathbf{p}^{n+1/2} &= \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n), & (18)\\ \mathbf{r}^{n+1} &= \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, & (19)\\ \mathbf{p}^{n+1} &= \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}), & (20) \end{aligned}$$

where Δt is the step size. The method was first used in the context of MD by Verlet (Verlet, 1967) and has been very popular with the MD community since (Allen and Tildesley, 1987). It is worthwhile noting that the method was essentially known to Newton already and was used in his *Principia* from 1687 to prove Kepler's second law (see Hairer et al. (2003)).

Why is the Störmer-Verlet method so successful for constant energy MD simulations? Several reasons can be given. The method is easy to implement, it exactly conserves total linear and angular momentum, it is time-reversible, it conserves volume of phase space $\Omega = \mathbb{R}^{3N}$,[1] and the total energy (1) is very well conserved over long simulation times even for large and complex systems. See Hairer et al. (2002); Leimkuhler and Reich (2005) for a detailed theoretical exploration of the Störmer-Verlet and related methods. We will explore energy conservation under the Störmer-Verlet method further

[1] Conservation of volume in phase space is different from conservation of volume V in physical space. Conservation of volume in phase space means that the time-stepping map $\Psi_{\Delta t} : \Omega \to \Omega$, induced by the Störmer-Verlet method, has Jacobian with determinant equal to one.

in Section 3.4. In particular, we will summarize an explicit algorithm (Skeel and Hardy, 2001) for computing shadow (also called modified) energies along numerically computed trajectories. It can be shown that the Störmer-Verlet method allows for shadow energies of arbitrary high accuracy with respect to the step-size Δt (Hairer et al., 2002; Leimkuhler and Reich, 2005). Shadow energies can be used, for example, to assess the quality of NVT ensemble simulations (Engle et al., 2005) and to increase the acceptance rate in Markov chain Monte Carlo methods based on molecular dynamics proposal steps (Izaguirre and Hampton, 2004; Akhmatskaya and Reich, 2006, 2008).

2.2.1 Langevin dynamics

While there is common agreement that Störmer-Verlet provides the gold standard for constant energy MD simulations, the situation is less clear for Langevin dynamics. A popular method is the BBK method of Brünger et al. (1984). The BBK discretization for (14)-(15) is provided by

$$\mathbf{p}^{n+1/2} = \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n) - \frac{\gamma \Delta t}{2}\mathbf{p}^n + \sqrt{\frac{\gamma \Delta t}{2}} M^{1/2}\mathbf{R}^n, \tag{21}$$

$$\mathbf{r}^{n+1} = \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \tag{22}$$

$$\mathbf{p}^{n+1} = \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n) - \frac{\gamma \Delta t}{2}\mathbf{p}^{n+1} + \sqrt{\frac{\gamma \Delta t}{2}} M^{1/2}\mathbf{R}^{n+1} \tag{23}$$

where $\mathbf{R}^n$ is a $3N$-dimensional vector of independent Gaussian random variables with mean zero and variance equal to $k_B T$; i.e., $\mathbf{R}^n \sim [\mathrm{N}(0, k_B T)]^{3N}$. The BBK scheme reduces to the Störmer-Verlet method for $\gamma = 0$ and is second-order accurate in the deterministic part (i.e., for $\sigma = 0$ in (15)). The stochastic scheme is however only weakly first-order convergent. For a further discussion of the BBK scheme see Pastor et al. (1988).

The limited accuracy of the BBK can be overcome by more advanced integration techniques such as the schemes suggested by Gunsteren and Berendsen (1982) and Allen (1982). A further improved variant of these methods has been suggested by Skeel (1999); Skeel and Izaguirre (2002) and implemented by Izaguirre et al. (2001). The method suggested by Skeel (1999) relies on a splitting of the Langevin equations into exactly solvable sub-problems and an appropriate combination of the exact solution operators. More recently, related splitting techniques have been proposed by Fabritis et al. (2006); Bussi and Parrinello (2007).

A detailed comparison of various algorithms for Langevin dynamics has been provided by Wang and Skeel (2003). An elementary introduction to general numerical methods for SDEs can be found in Higham (2001).

We finally mention the generalized hybrid Monte Carlo (GHMC) method of Horowitz (1991); Kennedy and Pendleton (2001) as a Metropolis adjusted time-stepping method. The Metropolis criterion guarantees that the numerically computed temperature T^* is independent of the step-size Δt and converges to the target temperature T over sufficiently long simulation intervals. We will describe the GHMC method in more detail in Section 5.1. A somewhat related Monte Carlo method has been proposed by Scemama et al. (2006). The efficiency of the GHMC method has been improved through the use of shadow Hamiltonians by Akhmatskaya and Reich (2006, 2008); Akhmatskaya et al. (2009).

2.2.2 Dissipative particle dynamics

The optimal numerical treatment of the DPD equations (6)-(7) is still a subject of debate. See, for example, Pagonabarraga et al. (1998); Lowe (1999); Besold et al. (2000); Shardlow (2003); Vattulainen et al. (2002); Nikunen et al. (2003); Peters (2004); Hafskjold et al. (2004); Koopman and Lowe (2006);

Serrano et al. (2006). For further reference we state a method introduced by Besold et al. (2000) for the DPD equations (6)-(7) and which we extend to the generalized DPD equations (9)-(10):

$$\begin{aligned}
\mathbf{p}^{n+1/2} &= \mathbf{p}^n - \frac{\Delta t}{2}\left\{\nabla_{\mathbf{r}} V(\mathbf{r}^n) + \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^n) \times \right. \\
&\qquad \left. \left[\gamma \nabla_{\mathbf{r}} h_k(\mathbf{r}^n)\cdot \mathbf{M}^{-1}\mathbf{p}^n - \sqrt{\frac{2\gamma}{\Delta t}} R_k^n\right]\right\}, \qquad (24)\\
\mathbf{r}^{n+1} &= \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, \qquad (25)\\
\mathbf{p}^{n+1} &= \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\left\{\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}) + \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1}) \times \right. \\
&\qquad \left. \left[\gamma \nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1})\cdot \mathbf{M}^{-1}\mathbf{p}^{n+1} - \sqrt{\frac{2\gamma}{\Delta t}} R_k^{n+1}\right]\right\}, \qquad (26)
\end{aligned}$$

where R_k^n are independent standard Gaussian random variables with mean zero and variance equal to $k_B T$; i.e., $R_k^n \sim \mathrm{N}(0, k_B T)$.

This scheme is essentially a generalization of the BBK scheme (21)-(23) for Langevin dynamics and we will refer to (24)-(26) as the DPD-BBK scheme. Note that DPD-BBK is significantly more expensive to implement than the original BBK scheme since the linear system of equations in $\mathbf{p}^{n+1}$ arising from (26) is not diagonal.

For a comparison of several numerical algorithms for DPD see Vattulainen et al. (2002); Nikunen et al. (2003). In particular, it is found that the numerically observed temperature T^* depends on the step-size Δt and differs from the target temperature T. Methods are now available that lead to $T^* = T$ in the absence of conservative forces (Lowe, 1999; Peters, 2004; Koopman and Lowe, 2006; Serrano et al., 2006). However, none of the existing methods leads to $T^* = T$ under the full DPD dynamics.

3 Theoretical background

3.1 Mathematical aspects of stochastic thermostats

From a mathematical point of view, all presented stochastic thermostats generate stochastic processes and, in fact, Markov processes (Meyn and Tweedie, 1993). Discrete time Markov processes are characterized by states $\Gamma \in \Omega$ and a transition probability density kernel $P(\Gamma'|\Gamma)$, which gives the probability density of going from a state Γ to a state Γ'.

In case of Langevin and dissipative particle dynamics, as considered in this paper, the possible states Γ are given by $\Gamma = (\mathbf{r}^T, \mathbf{p}^T)^T \in \mathbb{R}^{3N} = \Omega$. A discrete time Markov process is obtained by fixing an incremental time interval Δt and considering solutions $\Gamma(t)$ of any of the considered stochastic thermostats. We denote by $P_{\Delta t}(\Gamma'|\Gamma)$ the probability density of solutions $\Gamma(t)$, which start at $\Gamma = \Gamma(0)$ and terminate at $\Gamma' = \Gamma(\Delta t)$ at $t = \Delta t$.

A probability density π is called invariant under a general transition probability density kernel P if

$$\pi(\Gamma') = \int P(\Gamma'|\Gamma)\, \pi(\Gamma)\, \mathrm{d}\Gamma. \qquad (27)$$

It is relatively straightforward to show that the the canonical distribution

$$\pi_{\mathrm{can}}(\mathbf{r}, \mathbf{p}) \propto e^{-\beta E} \qquad (28)$$

is invariant under the stochastic thermostats, where $\beta = 1/k_B T$ denotes the inverse temperature and the total energy E is given by (1).

A useful sufficient criterion for the invariance of a probability density function π under a Markov chain P is given by detailed balance

$$P(\Gamma'|\Gamma)\,\pi(\Gamma) = P(\Gamma|\Gamma')\,\pi(\Gamma'), \tag{29}$$

which implies (27) since

$$\int P(\Gamma'|\Gamma)\,\pi(\Gamma)\,\mathrm{d}\Gamma = \int P(\Gamma|\Gamma')\,\pi(\Gamma')\,\mathrm{d}\Gamma = \pi(\Gamma') \int P(\Gamma|\Gamma')\,\mathrm{d}\Gamma = \pi(\Gamma'). \tag{30}$$

A more difficult question is that of uniqueness of the invariant distribution and a geometric rate of convergence towards the invariant distribution, which then implies ergodicity of the associated Markov process, i.e.,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\Gamma(t))\,\mathrm{d}t = \int f(\Gamma)\,\pi(\Gamma)\,\mathrm{d}\Gamma \tag{31}$$

for integrable functions $f : \Omega \to \mathbb{R}$ (Meyn and Tweedie, 1993). A useful concept is geometric ergodicity[2] (Roberts and Tweedie, 1995) which states that

$$\|P^n(\cdot|\Gamma) - \pi(\cdot)\|_W \le W(\Gamma) R\,\delta^n \tag{32}$$

for all states $\Gamma \in \Omega$, all $n \ge 1$, and for some positive constants $R < \infty$ and $\delta < 1$. Here $W(\Gamma) \ge 1$ is a given (Lyapunov) function, the norm $\|\cdot\|_W$ is defined by

$$\|\rho\|_W := \sup_{|f|\le W} \int f(\Gamma)\,\rho(\Gamma)\,\mathrm{d}\Gamma \tag{33}$$

for all integrable $\rho : \Omega \to \mathbb{R}$, and $P^n(\Gamma'|\Gamma)$ is recursively defined by

$$P^n(\Gamma'|\Gamma) = \int P(\Gamma'|\bar{\Gamma})\,P^{n-1}(\bar{\Gamma}|\Gamma)\,\mathrm{d}\bar{\Gamma}. \tag{34}$$

Geometric ergodicity has been discussed for the Langevin equations, for example, in Mattingly et al. (2002) and for dissipative particle dynamics in Shardlow and Yan (2006).

3.2 Numerical time-stepping aspects

All computer implementations of stochastic thermostats rely on a discretization of the underlying stochastic differential equations. On a more abstract level, this corresponds to replacing the exact time-Δt Markov process $P_{\Delta t}(\Gamma'|\Gamma)$ of the underlying stochastic thermostat by a numerically induced Markov process $\mathcal{P}_{\Delta t}(\Gamma'|\Gamma)$. It is desirable that $\mathcal{P}_{\Delta t}$ possesses the same qualitative properties as $P_{\Delta t}$; i.e., we expect a unique invariant measure $\pi_{\Delta t}$ and (geometric) ergodicity with respect to that measure. Ideally we would also like to obtain

$$\pi_{\mathrm{can}} = \pi_{\Delta t} \tag{35}$$

for all $\Delta t > 0$, π_{can} the canonical distribution (28). However, (35) does not necessarily hold for general time-stepping methods as we demonstrate next for a simple Brownian dynamics model system.

[2]There exists different definitions of geometric ergodicity in the literature. Here we use V-uniform ergodicity as defined by Meyn and Tweedie (1993); Roberts and Tweedie (1995).

Consider one-dimensional Brownian dynamics

$$\mathrm{d}x = -\nabla V(x)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W \tag{36}$$

and its (standard) Euler-Maruyama approximation

$$x^{n+1} = x^n - \Delta t \nabla V(x^n) + \sqrt{2\Delta t} R^n, \qquad R^n \sim \mathrm{N}(0,1). \tag{37}$$

We find that $\mathcal{P}_{\Delta t}(x^{n+1}|x^n)$ can be stated explicitly in this case; i.e.

$$\mathcal{P}_{\Delta t}(\cdot|x^n) = \mathrm{N}(x^{n+1} - x^n + \Delta t \nabla V(x^n), 2\Delta t). \tag{38}$$

It is also known (Allen and Tildesley, 1987) that (36) and, hence, the exact time-Δt transition probability kernel $P_{\Delta t}(x'|x)$ possess the invariant distribution

$$\pi_{\mathrm{bd}}(x) \sim e^{-V(x)}. \tag{39}$$

A fairly complete discussion of this problem can be found in Roberts and Tweedie (1995). Here we restrict the discussion even further by considering the two potentials $V(x) = x^2/2$ and $V(x) = x^4/4$. In both cases, it can be shown that Brownian dynamics is geometrically ergodic and that π_{bd} is the unique invariant measure (Roberts and Tweedie, 1995). The situation is more complex for the numerical approximation (37) (Roberts and Tweedie, 1995; Mattingly et al., 2002).

The harmonic potential $V(x) = x^2/2$ leads to

$$\mathcal{P}_{\Delta t}(\cdot|x^n) = \mathrm{N}(x^{n+1} - x^n + \Delta t x^n, 2\Delta t) \tag{40}$$

and one finds that

$$\int \mathcal{P}(x'|x)\,\pi_{\mathrm{bd}}(x)\,\mathrm{d}x \sim \mathrm{N}\left(0, \frac{1+2\Delta t}{(1-\Delta t)^2}\right) \tag{41}$$

which is not equal to $\pi_{\mathrm{bd}}(x) \sim \mathrm{N}(0,1)$. However, there is an invariant density $\pi_{\Delta t}$ and $\mathcal{P}_{\Delta t}$ is geometrically ergodic provided $\Delta t \leq 2$. It also holds that $\pi_{\Delta t} \to \pi_{\mathrm{bd}}$ for $\Delta t \to 0$. See Section 3.5 for a numerical demonstration.

On the other hand, the numerically induced Markov chain $\mathcal{P}_{\Delta t}$ is not geometrically ergodic for $V(x) = x^4/4$ and, even worse, for any Δt there is a non-vanishing probability that the numerical results will increase without bounds (numerical instability) (Mattingly et al., 2002). Numerically one finds that there is a critical step-size Δt_c such that numerically instabilities are practically not observed for all $\Delta t \leq \Delta t_c$ (even though, theoretically, the probability of blow-up remains non-zero). In the following section, we discuss a Monte Carlo approach which can control such instabilities by re-enforcing the invariant distribution of the analytic problem during the simulation via an appropriate Metropolis-Hastings acceptance criterion.

3.3 Markov chain Monte Carlo methods

Markov chain Monte Carlo (MCMC) methods (Liu, 2001) are an extremely flexible tool to generate Markov chains with a $P(\Gamma'|\Gamma)$ transition probability density kernel which satisfy detailed balance with respect to a given distribution π. The principle idea is to take an available transition probability density kernel $A(\Gamma'|\Gamma)$ and to combine it with a Metropolis-Hastings acceptance criterion. More specifically, given a current state Γ a proposal state Γ' is generated according to $A(\Gamma'|\Gamma)$. The proposal state Γ' is then accepted according to the Metropolis-Hastings criterion $r(\Gamma', \Gamma) \geq \xi$, where $\xi \in [0,1]$ is a uniformly distributed random number, i.e., $\xi \sim \mathrm{U}([0,1])$,

$$r(\Gamma', \Gamma) = \frac{\delta(\Gamma', \Gamma)}{\pi(\Gamma)\,A(\Gamma'|\Gamma)}, \tag{42}$$

and $\delta(\Gamma', \Gamma)$ is any function with $\delta(\Gamma', \Gamma) = \delta(\Gamma, \Gamma')$ that makes $r(\Gamma', \Gamma) \leq 1$. A common choice for $\delta(\Gamma', \Gamma)$ is

$$\delta(\Gamma', \Gamma) = \min\{\pi(\Gamma)\, A(\Gamma', \Gamma), \pi(\Gamma')\, A(\Gamma, \Gamma')\} \tag{43}$$

which leads to

$$r(\Gamma', \Gamma) = \min\left(1, \frac{\pi(\Gamma')\, A(\Gamma, \Gamma')}{\pi(\Gamma)\, A(\Gamma', \Gamma)}\right). \tag{44}$$

In case the proposal step is rejected, one continues with Γ. One can show, that the complete Markov chain (proposal step plus Metropolis-Hastings acceptance criterion) satisfies detailed balance (29) (Liu, 2001).

If $A(\Gamma', \Gamma)$ comes from a numerical time-stepping method, then one can think of the Metropolis-Hastings criterion as a correction step that re-enforces the invariant distribution π of the analytic problem. In this case, the rejection rate should be low for sufficiently small time-step sizes Δt.

To make this point more transparent, let us return to the numerical time-stepping method (37) and its induced Markov chain (40) for $V(x) = x^2/2$. We already found that

$$\mathcal{P}_{\Delta t}(x'|x) = \mathrm{N}(x' - x + \Delta t x, 2\Delta t). \tag{45}$$

We also find that

$$\mathcal{P}_{\Delta t}(x|x') = \mathrm{N}(x - x' + \Delta t x', 2\Delta t) \tag{46}$$

and, hence,

$$r(x', x) = \min\left(1, \frac{e^{-(x)'^2/2}\mathcal{P}_{\Delta t}(x|x')}{e^{-x^2/2}\mathcal{P}_{\Delta t}(x'|x)}\right), \tag{47}$$

which can be explicitly evaluated. This is essentially the Metropolis adjusted Langevin algorithm (MALA) of Roberts and Tweedie (1995).

An equivalent but more elegant approach is the following. We write down the Newtonian equations of motion for a particle with position x, mass $m = 1$, and potential energy $V(x)$, i.e.,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = p, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\nabla V(x). \tag{48}$$

The equations are approximated numerically by a single time-step with the Störmer-Verlet method (18)-(20) with time-step size $h = \sqrt{2\Delta t}$ and initial momentum $p = R \sim \mathrm{N}(0, 1)$, which we write as

$$\bar{p} = p - \frac{h}{2}\nabla V(x), \tag{49}$$

$$x' = x + h\bar{p}, \tag{50}$$

$$p' = \bar{p} - \frac{h}{2}\nabla V(x'). \tag{51}$$

We find that the propagation in x is given by

$$x' = x + hp - \frac{h^2}{2}\nabla V(x) = x - \Delta t \nabla V(x) + \sqrt{2\Delta t} R, \tag{52}$$

which, as desired, coincides with the Euler-Maruyama approximation (37). However, since the proposal step (49)-(51) is volume conserving and time-reversible (Liu, 2001), the Metropolis-Hastings acceptance criterion simplifies to

$$r = \min\left(1, \exp(E - E')\right) \geq \xi, \qquad \xi \sim \mathrm{U}([0, 1]), \tag{53}$$

with new and old energy values

$$E = p^2/2 + V(x), \qquad E' = (p')^2/2 + V(x'), \tag{54}$$

respectively. What we have just described is a special instance of the hybrid Monte Carlo (HMC) method of Duane et al. (1987); Mehlig et al. (1992); Forrest and Suter (1994).

As shown in Roberts and Tweedie (1995), MALA/HMC is geometrically ergodic for $V(x) = x^2/2$ but loose this property for $V(x) = x^4/4$. Numerical experiments reveal that MALA/HMC never produce unstable sample trajectories but might have long periods of rejections if Δt is chosen too large. Furthermore, the performance remains satisfactory for much larger values of Δt compared to the unadjusted algorithm (37). See Section 3.5 for a numerical demonstration. In the following section, we introduce the concept of shadow energies for symplectic time-stepping methods such as Störmer-Verlet. Shadow energies can be used to increase the acceptance rate in hybrid Monte Carlo methods (Izaguirre and Hampton, 2004; Akhmatskaya and Reich, 2006, 2008). We will come back to this issue in Section 5.2.

3.4 Shadow energies

It is well-known that the Störmer-Verlet method (18)-(20) conserves the total energy (1) over very long simulation intervals. This observation is a consequence of the symplectic nature of the scheme (Hairer et al., 2002; Leimkuhler and Reich, 2005). Furthermore, it can be shown that the Störmer-Verlet method conserves, so called, shadow or modified energies to even higher accuracy in Δt. For example, a fourth-order shadow energy is given by the expression

$$E^{[4]}_{\Delta t}(\mathbf{r},\mathbf{p}) = E(\mathbf{r},\mathbf{p}) + \frac{\Delta t^2}{24}\left[2(\nabla_{\mathbf{r}} V(\mathbf{r}))^T \mathbf{M}^{-1} \nabla_{\mathbf{r}} V(\mathbf{r}) - \mathbf{p}^T M^{-1} V_{\mathbf{rr}}(\mathbf{r}) M^{-1}\mathbf{p}\right], \tag{55}$$

where E is given by (1) and $V_{\mathbf{rr}}(\mathbf{r})$ denotes the Hessian matrix of the potential energy V at $\mathbf{r}$ (Hairer et al., 2002; Leimkuhler and Reich, 2005).

We now summarize the elegant construction of Skeel and Hardy (2001)) for monitoring shadow energies along numerically computed trajectories which does not require the computation of the Hessian or higher-order derivatives of V. An alternative construction can be found in Akhmatskaya and Reich (2006, 2008).

Following Skeel and Hardy (2001), we introduce an additional degree of freedom $(a, b) \in \mathbb{R}^2$ and the extended formulation

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = M^{-1}\mathbf{p}, \quad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\nabla_{\mathbf{r}} V(\mathbf{r}), \quad \frac{\mathrm{d}b}{\mathrm{d}t} = \mathbf{r}^T \nabla_{\mathbf{r}} V(\mathbf{r}) - 2V(\mathbf{r}), \quad \frac{\mathrm{d}a}{\mathrm{d}t} = 0. \tag{56}$$

Let $\mathbf{Y}(t) = (\mathbf{r}(t)^T, 1, \mathbf{p}(t)^T, b(t)) \in \mathbb{R}^{3N+2}$ denote a solution of (56). A key result of Skeel and Hardy (2001) states that the energy (1) satisfies

$$E = \frac{1}{2}\left\{\dot{\mathbf{r}}(t)^T \mathbf{p}(t) - \dot{\mathbf{p}}(t)^T \mathbf{r}(t) - \dot{b}(t)\right\} = \frac{1}{2}\dot{\mathbf{Y}}(t)^T J \mathbf{Y}(t), \tag{57}$$

which, using (56), is easy to verify. Here $\dot{\mathbf{z}}$ denotes the time derivative of a variable $\mathbf{z}$, J is the skew-symmetric matrix

$$J = \begin{bmatrix} 0_{3N+1} & I_{3N+1} \\ -I_{3N+1} & 0_{3N+1} \end{bmatrix} \in \mathbb{R}^{(6N+2)\times(6N+2)} \tag{58}$$

and I_{3N+1} denotes the identity matrix of dimension $3N + 1$, N the number of atoms, and $\mathbf{0}_{3N+1}$ denotes the zero matrix of the same dimension.

We now introduce an appropriate extension of the Störmer-Verlet method (18)-(20) to the extended formulation (56); i.e.

$$\begin{aligned}
\mathbf{p}^{n+1/2} &= \mathbf{p}^n - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^n), && (59)\\
b^{n+1/2} &= b^n + \frac{\Delta t}{2}\left[(\mathbf{r}^n)^T \nabla_{\mathbf{r}} V(\mathbf{r}^n) - 2V(\mathbf{r}^n)\right], && (60)\\
\mathbf{r}^{n+1} &= \mathbf{r}^n + \Delta t \mathbf{M}^{-1}\mathbf{p}^{n+1/2}, && (61)\\
b^{n+1} &= b^{n+1/2} + \frac{\Delta t}{2}\left[(\mathbf{r}^{n+1})^T \nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}) - 2V(\mathbf{r}^{n+1})\right], && (62)\\
\mathbf{p}^{n+1} &= \mathbf{p}^{n+1/2} - \frac{\Delta t}{2}\nabla_{\mathbf{r}} V(\mathbf{r}^{n+1}). && (63)
\end{aligned}$$

For further reference, we denote the induced time-stepping method in the extended state variable $\mathbf{Y}$ by

$$\mathbf{Y}^{n+1} = \widehat{\Psi}_{\Delta t}(\mathbf{Y}^n). \tag{64}$$

Once an appropriate extension of a symplectic time-stepping method has been formulated, Skeel and Hardy (2001) provide a straightforward approach for evaluating higher-order shadow Hamiltonians. Here we only summarize the final formulas for fourth and eighth order shadow Hamiltonians.

Assume that we wish to determine the value of the shadow Hamiltonian about $\mathbf{Y}^n = ((\mathbf{r}^n)^T, 1, (\mathbf{p}^n)^T, b^n)^T$. Then $k = 1$ (fourth-order) or $k = 2$ (eighth-order) time-steps forward and backward in time are performed with (64) (unless those steps have already been performed as part of the simulation). Hence we have $2k + 1$ discrete approximations $\mathbf{Y}^i$, $i = n - k, \ldots, n, \ldots, n + k$, centered about $\mathbf{Y}^n$ available and, for $k = 2$, we define

$$\begin{aligned}
\mathbf{A}_0 &= \mathbf{Y}^i, && (65)\\
\mathbf{A}_1 &= \frac{1}{2}\left(\mathbf{Y}^{i+1} - \mathbf{Y}^{i-1}\right), && (66)\\
\mathbf{A}_2 &= \mathbf{Y}^{i+1} - 2\mathbf{Y}^i + \mathbf{Y}^{i-1}, && (67)\\
\mathbf{A}_3 &= \frac{1}{2}\left[\mathbf{Y}^{i+2} - 2\mathbf{Y}^{i+1} + 2\mathbf{Y}^{i-1} - \mathbf{Y}^{i-2}\right], && (68)\\
\mathbf{A}_4 &= \mathbf{Y}^{i+2} - 4\mathbf{Y}^{i+1} + 6\mathbf{Y}^i - 4\mathbf{Y}^{i-1} + \mathbf{Y}^{i-2}. && (69)
\end{aligned}$$

as well as

$$A_{lm} = \frac{1}{2\Delta t}\mathbf{A}_l^T J \mathbf{A}_m, \quad l, m = 0, \ldots, 4, \tag{70}$$

with the skew-symmetric matrix J defined by (58). In case of $k = 1$, we only need to compute $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2$ and A_{lm} for $m, l = 0, \ldots, 2$.

A fourth-order shadow Hamiltonian at $\mathbf{Y}^n$ is now defined by

$$\mathcal{E}^{[4]}_{\Delta t}(\mathbf{r}^n, \mathbf{p}^n) = A_{10} - \frac{1}{6}A_{12} \tag{71}$$

and a eighth-order shadow Hamiltonian by

$$\mathcal{E}^{[8]}_{\Delta t}(\mathbf{r}^n, \mathbf{p}^n) = A_{10} - \frac{2}{7}A_{12} + \frac{5}{42}A_{30} + \frac{13}{105}A_{32} - \frac{19}{210}A_{14} - \frac{1}{140}A_{34}, \tag{72}$$

respectively. Note that (55) and (71) agree up to terms of $\mathcal{O}(\Delta t^4)$. However, (71) is easier to implement than (55) which requires the explicit evaluation of the Hessian matrix $V_{\mathbf{rr}}(\mathbf{r})$ of the potential energy V.

We will use shadow energies in Section 5.2 to improve the acceptance rate of a generalized hybrid Monte Carlo method. See also Akhmatskaya and Reich (2006, 2008).

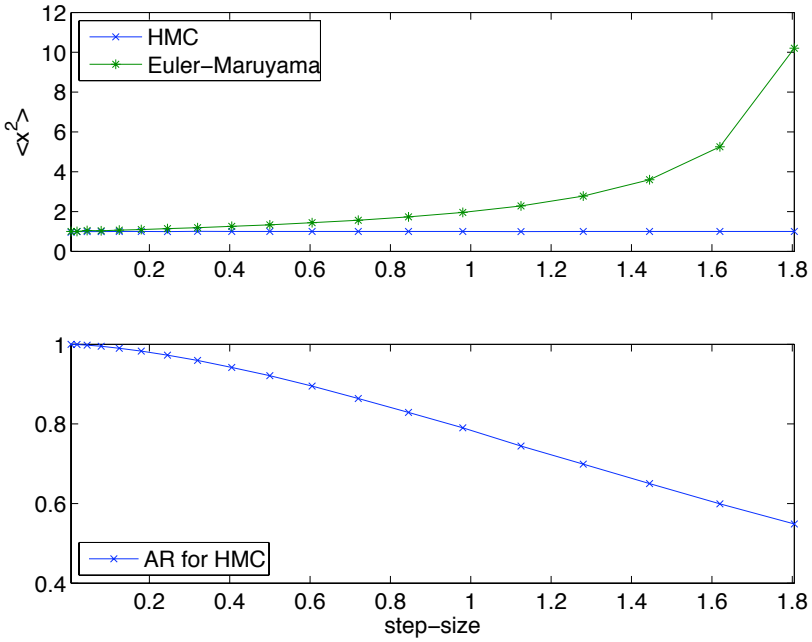

Figure 1: Displayed are the computed expectation value $\langle x^2 \rangle$ for the Euler-Maruyama scheme and the MALA/HMC method. The exact value is $\langle x^2 \rangle = 1$. The acceptance rates (AR) is provided for the MALA/HMC method. Here AR = 1 corresponds to a 100% acceptance rate. Results are displayed for step-sizes $\Delta t \in [0.05, 1.8]$. While the Euler-Maruyama scheme leads to large sampling errors, the MALA/HMC method leads to correct sampling results independent of the step-size Δt.

3.5 Numerical demonstration

We compare the behavior of the Euler-Maruyama approximation (37) for Brownian dynamics (36) with quadratic potential $V(x) = x^2/2$ with the behavior of the Metropolis adjusted (MALA)/hybrid Monte Carlo (HMC) scheme as described in Section 3.3. We monitor the computed expectation values of x^2 for values of the step-size in the range $\Delta t \in [0.05, \ldots, 1.8]$. Note that the stability limit of the Euler-Maruyama scheme is given by $\Delta t < 2$. It can be seen that the Euler-Maruyama scheme leads to large sampling errors while the MALA/HMC method leads to correct sampling results independent of the step-size Δt. In Figure 1, we also provide the acceptance rate (AR) for the MALA/HMC method. Numerically computed distribution functions in the position variable x are displayed in Figure 2 for $\Delta t = 1.8$. While the Euler-Maruyama scheme leads to a severely distorted distribution, the MALA/HMC scheme leads to perfect agreement (up to finite sample size effects) with the exact distribution.

In a second experiment we study conservation of energy under the Störmer-Verlet method applied to the microcanonical equations (3)-(4). As already discussed in Section 3.3, conservation of energy plays a crucial role in the acceptance rate of hybrid Monte Carlo methods. For simplicity, we consider a one-dimensional period chain of $N = 10$ particles with positions q_i and mass $m = 0.5$ interacting through a pair-wise Lennard-Jones potential

$$V(r) = \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}, \qquad \sigma = (1/2)^{1/6}, \qquad r = q_i - q_{i-1}.$$

The length of the domain is $l = 10$. Relative changes in the numerical values of the energy (1) as well as fourth and eighth-order shadow energies (71) and (72), respectively, can be found in Figure 3. The excellent conservation of shadow energies has motivated the development of hybrid Monte

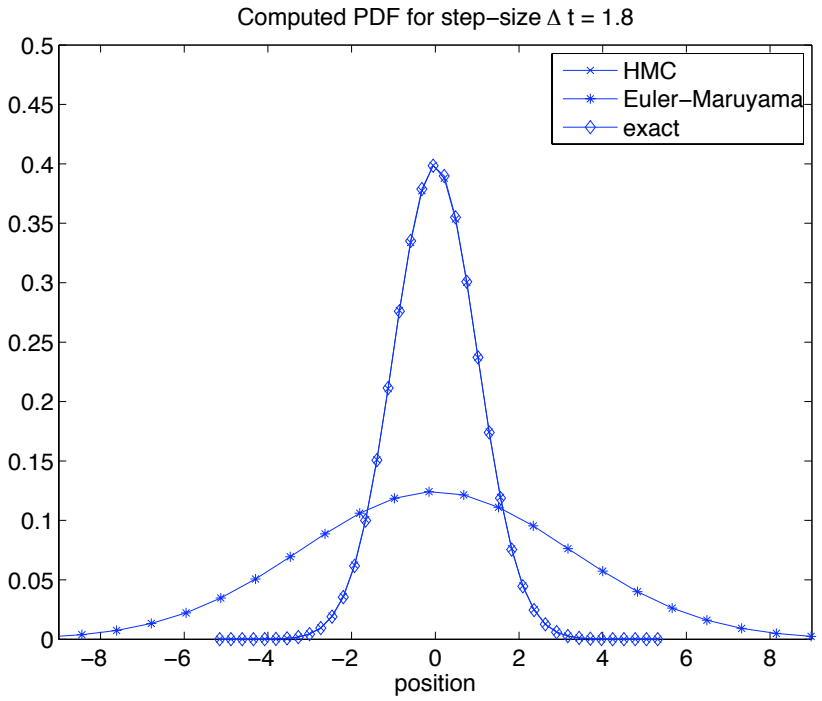

Figure 2: Displayed are the computed probability density functions for step-size $\Delta t = 1.8$ and for the Euler-Maruyama scheme and the MALA/HMC scheme. The agreement between the exact and the numerical distribution from the MALA/HMC scheme is almost perfect, while the distribution from the Euler-Maruyama scheme is severely distorted.

Carlo methods which sample with respect to the canonical distribution associated with a shadow energy. See Izaguirre and Hampton (2004) for the shadow hybrid Monte Carlo (SHMC) method, Akhmatskaya and Reich (2006, 2008) for generalized shadow hybrid Monte Carlo (GSHMC) methods, and Section 5.2 for a generalization of the GSHMC methods to DPD-type momentum dynamics.

4 A splitting approach for stochastic thermostats

In this section, we develop a general framework for numerical implementations of stochastic thermostats. The approach is based on a splitting of the stochastic differential equation into an energy conserving part and fluctuation-dissipation part. This splitting is used since each part separately conserves the canonical distribution (28) under its analytically generated Markov processes. None of the two parts is, however, ergodic with respect to (28).

While splitting methods have been discussed in the context of stochastic differential equations before (see, for example, Skeel (1999); Izaguirre et al. (2001); Skeel and Izaguirre (2002); Shardlow (2003); Fabritis et al. (2006); Serrano et al. (2006); Bussi and Parrinello (2007)), we propose to use simplified numerical methods for the fluctuation-dissipation contributions which, however, exactly conserve the momentum Boltzmann distribution (Allen and Tildesley, 1987). We next describe these methods in the context of Langevin and dissipative particle dynamics before we formulate the complete time-stepping approach at the end of this section.

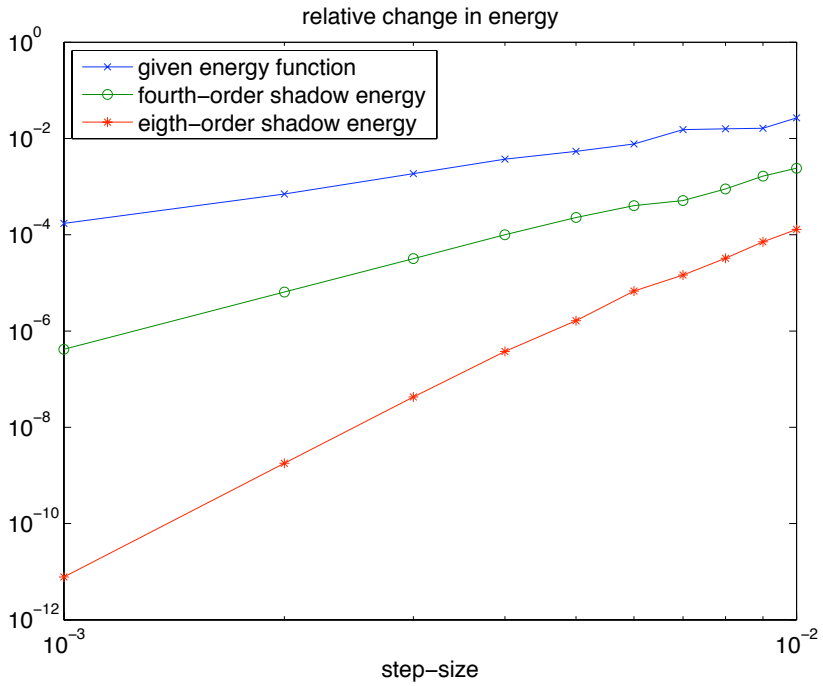

Figure 3: Displayed are relative changes in the total energy and fourth and eighth-order shadow energies as a function of the step-size Δt. Simulations are performed with the Störmer-Verlet method for a one-dimensional chain of Lennard-Jones particles. The slope of the graphs correctly reflects the order of the method and its shadow energies for sufficiently small step-sizes.

4.1 Time-stepping methods for fluctuation-dissipation terms

We wish to investigate numerical time-stepping methods for stochastic thermostats and their ability to reproduce the canonical distribution as well as geometric ergodicity properties. To begin with, we consider the Langevin and dissipative particle dynamics for zero potential energy; i.e., $V(\mathbf{r}) = 0$ (force free motion), and $\mathbf{r} = const.$

4.1.1 Langevin dynamics

We consider the equation

$$\mathrm{d}\mathbf{p} = -\gamma \mathbf{p} dt + \sigma M^{1/2}\, \mathrm{d}\mathbf{W} \tag{73}$$

where, as before, $\mathbf{W}(t)$ is a N-dimensional Wiener process and $\sigma = \sqrt{2\gamma k_B T}$. We assume that the mass matrix M is diagonal. Obviously, the equation (73) is linear and can be solved exactly using stochastic quadrature. See, for example, Skeel (1999); Skeel and Izaguirre (2002), for a numerical implementation of such an approach. However, exact solutions require multiple sets of random numbers and cannot easily be generalized to more complex fluctuation-dissipation terms.

On the other hand, a straightforward and easy to implement numerical discretization is

$$\mathbf{p}^{n+1} = \mathbf{p}^n - \gamma \Delta t \mathbf{p}^n + \sqrt{2\gamma \Delta t} M^{1/2} \mathbf{R}^n, \qquad \mathbf{R}^n \sim [\mathrm{N}(0, k_B T)]^{3N}. \tag{74}$$

As already discussed for the related method (37) with $V(x) = x^2/2$, the discretization (74) is geometrically ergodic but it converges to a stationary distribution which is a step-size dependent perturbation of the desired momentum Boltzmann distribution

$$\pi_{\mathrm{b}}(\mathbf{p}) \propto e^{-\beta \|\mathbf{p}\|^2/2}, \tag{75}$$

$\beta = 1/k_B T$. Following Kennedy and Pendleton (2001), an alternative time-stepping method is

$$\mathbf{p}^{n+1} = \cos(\phi)\,\mathbf{p}^n + \sin(\phi)\,M^{1/2}\mathbf{R}^n \tag{76}$$

with $\mathbf{R}^n \sim [\mathrm{N}(0, k_B T)]^{3N}$ as before. The angle ϕ is given by

$$\phi = \sqrt{2\gamma\Delta t}. \tag{77}$$

Standard Taylor expansion of (76) in the angle ϕ reveals that (76) and (74) agree in the expansion up to terms of $\mathcal{O}(\Delta t^{3/2})$. More importantly, (76) is geometrically ergodic with the Boltzmann distribution (75) as the stationary distribution independent of Δt. To see this, we formally augment (76) by an equation for $\mathbf{R}$, i.e.

$$\mathbf{R}^* = \cos(\phi)\,\mathbf{R}^n - \sin(\phi)\,M^{-1/2}\mathbf{p}^n, \tag{78}$$

and consider the energy values $E_c^n = E_c(\mathbf{p}^n, \mathbf{R}^n)$, $E_c' = E_c(\mathbf{p}^{n+1}, \mathbf{R}^*)$ with

$$E_c(\mathbf{p}, \mathbf{R}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + \frac{1}{2}\mathbf{R}^T\mathbf{R}. \tag{79}$$

It is easy to verify that $E_c' = E_c^n$ under the equations (76)-(78). Since the linear transformation (76)-(78) is also volume conserving (i.e., the determinant of the transformation is equal to one), we may conclude that the distribution

$$\pi_{\mathrm{c}}(\mathbf{p}, \mathbf{R}) \propto e^{-\beta E_c} \tag{80}$$

is invariant, which implies the invariance of (75) in the variable $\mathbf{p}$ alone since (79) is separable.

We propose to view the momentum update (76)-(78) as a numerical approximation to the linear differential equation

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = M^{1/2}\mathbf{R}, \qquad \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}s} = -M^{-1/2}\mathbf{p}. \tag{81}$$

at $s = \phi = \sqrt{2\gamma\Delta t}$ with initial conditions $\mathbf{p}(0) = \mathbf{p}^n$ and $\mathbf{R}(0) = \mathbf{R}^n \sim [\mathrm{N}(0, k_B T)]^{3N}$. We call (81) the generating differential equation for the momentum update. We note that (81) is a Hamiltonian differential equation with conserved energy/Hamiltonian (79). The reader is referred to Arnold (1989); Hairer et al. (2002); Leimkuhler and Reich (2005) for the definition and theory of Hamiltonian differential equations, which essentially is a generalization of the Newtonian equations of motion (3)-(4) of classical mechanics.

Alternative numerical approximations to (73) can now be found by applying different numerical approximations to (81). These numerical approximations should be volume and energy conserving to guarantee the invariance of the momentum Boltzmann distribution (75). As an example, such an approximation is provided by the implicit midpoint rule

$$\mathbf{p}^{n+1} = \mathbf{p}^n + \frac{\Delta s}{2}M^{1/2}(\mathbf{R}^* + \mathbf{R}^n), \tag{82}$$

$$\mathbf{R}^* = \mathbf{R}^n - \frac{\Delta s}{2}M^{-1/2}(\mathbf{p}^{n+1} + \mathbf{p}^n), \tag{83}$$

with $\Delta s = \phi = \sqrt{2\gamma\Delta t}$. Other volume-conserving time-stepping and time-reversible methods, such as Störmer-Verlet, could also be used but need to be combined with a Metropolis-Hasting criterion to achieve exact sampling from the Boltzmann distribution (75) since the energy (79) is not exactly conserved. See Section 3.3 and the following section.

We now compare (82)-(83) to the BBK scheme (21)-(23) with zero potential energy V in terms of momentum propagation. We reformulate (82)-(83) as

$$\frac{\mathbf{p}^{n+1}-\mathbf{p}^n}{\Delta t} = -\gamma\frac{\mathbf{p}^{n+1}+\mathbf{p}^n}{2} + \sqrt{2\gamma/\Delta t}M^{1/2}\mathbf{R}^n \tag{84}$$

and find that (84) agrees with the corresponding BBK scheme except for differences in the treatment of the stochastic contribution. However, these differences are essential to make the scheme (82)-(83) conserve the momentum Boltzmann distribution (75). We also point to Cances et al. (2007) for an alternative modification of the BBK scheme which results in an exact sampling from the Boltzmann distribution under force-free motion.

4.1.2 Dissipative particle dynamics

We now consider the DPD-type force free momentum dynamics

$$\mathrm{d}\mathbf{p} = -\sum_{k=1}^K \nabla_{\mathbf{r}} h_k(\mathbf{r}) \left[\gamma \dot{h}_k(\mathbf{r})\,\mathrm{d}t - \sigma\,\mathrm{d}W_k\right] \tag{85}$$

for fixed position vector $\mathbf{r}$. Note that (85) constitutes a linear SDE. Again the goal is to devise numerical approximations $\{\mathbf{p}^n\}$ at $t_n = n\,\Delta t$, $n = 0, \ldots$, that are geometrically ergodic and possess the Boltzmann distribution as the stationary distribution. Existing approaches, which achieve this goal, rely on a splitting approach in which each contribution

$$\mathrm{d}\mathbf{p} = -\nabla_{\mathbf{r}} h_k(\mathbf{r}) \left[\gamma \dot{h}_k(\mathbf{r})\,\mathrm{d}t - \sigma\,\mathrm{d}W_k\right], \tag{86}$$

$k = 1, \ldots, K$, is treated sequentially in a conservative manner (Lowe, 1999; Peters, 2004; Shardlow, 2003). However, such a technique cannot be implemented efficiently on parallel computers.

Here we instead wish to derive simultaneous momentum update steps with invariant distribution (75). To achieve this goal, we propose to consider the generating differential equation

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = B\mathbf{R}, \qquad \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}s} = -B^T M^{-1}\mathbf{p}, \tag{87}$$

with $\mathbf{R}(0) = \mathbf{R} = (R_1, \ldots, R_K)^T$, $R_i \sim \mathrm{N}(0, k_B T)$ for $i = 1, \ldots, K$, and $B \in \mathbb{R}^{3N\times K}$ an arbitrary matrix. The linear system (87) is Hamiltonian with Hamiltonian function

$$E_c = \frac{1}{2}\left(\mathbf{p}^T M^{-1}\mathbf{p} + \mathbf{R}^T\mathbf{R}\right) \tag{88}$$

and skew-symmetric structure matrix

$$\mathcal{J} = \begin{bmatrix} 0_{3N} & B \\ -B^T & 0_K \end{bmatrix} \tag{89}$$

(Leimkuhler and Reich, 2005).

We define the matrix B for fixed position vector $\mathbf{r}$ through

$$B\mathbf{R} = \sum_{k=1}^K \nabla_{\mathbf{r}} h_k(\mathbf{r})\, R_k \tag{90}$$

and obtain

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r})\, R_k, \tag{91}$$

$$\frac{\mathrm{d}R_k}{\mathrm{d}s} = -\nabla_{\mathbf{r}} h_k(\mathbf{r}) \cdot M^{-1}\mathbf{p}, \qquad k = 1, \ldots, K, \tag{92}$$

as a generating differential equation for the fluctuation-dissipation part of the general DPD formulation (9)-(10).

To obtain a numerical momentum update step, we seek the solution at $s = \sqrt{2\gamma\Delta t}$ for given initial conditions

$$\mathbf{p}(0) = \mathbf{p}^n, \quad R_k(0) = R_k^n = \mathrm{N}(0, k_B T), \quad k = 1, \ldots, K. \tag{93}$$

Let us denote the linear solution operator, generated by the solutions of (91)-(92), by $\mathcal{R}(s) \in \mathbb{R}^{(3N+K)\times(3N+K)}$. The solution operator $\mathcal{R}(s)$ has the following properties:

(a) The solutions of (91)-(92) are volume conserving, i.e., $\det \mathcal{R}(s) = 1$.

(b) Given a fixed position vector $\mathbf{r}$, the solutions of (91)-(92) are time reversible, i.e., $\mathcal{F}\,\mathcal{R}(s)\,\mathcal{F} = \mathcal{R}(-s)$. Here $\mathcal{F}$ denotes the linear involution operator that changes the sign of all R_k's, $k = 1, \ldots, K$.

(c) The solutions of (91)-(92) conserve the extended Hamiltonian/energy (88).

(d) Properties (a) and (c) immediately imply that the solutions of (91)-(92) conserve the extended canonical distribution

$$\pi_{\mathrm{c}}(\mathbf{p}, \mathbf{R}) \propto e^{-\beta E_c}. \tag{94}$$

We now consider the numerical implementation of (91)-(92). To do so we follow the hybrid Monte Carlo (HMC) methodology (Duane et al., 1987; Mehlig et al., 1992) and consider time-reversible and volume conserving propagators for the dynamics in $\mathbf{p}$ and $\mathbf{R}$ with fixed position vector $\mathbf{r}$. Two such methods will be considered: (i) an explicit one, which does not conserve (88), and (ii) an implicit one, which does conserve (88).

Störmer-Verlet method. A first choice is provided by the application of the Störmer-Verlet method to (91)-(92) over J internal steps with internal step-size $\Delta s = \sqrt{2\gamma\Delta t}/J$ and we obtain

$$\mathbf{p}^{j+1/2} = \mathbf{p}^j + \frac{\Delta s}{2} \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}) R_k^j, \tag{95}$$

$$R_k^{j+1} = R_k^j - \Delta s \nabla_{\mathbf{r}} h_k(\mathbf{r}) \cdot M^{-1}\mathbf{p}^{j+1/2}, \qquad k = 1, \ldots, K, \tag{96}$$

$$\mathbf{p}^{j+1} = \mathbf{p}^{j+1/2} + \frac{\Delta s}{2} \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}) R_k^{j+1}. \tag{97}$$

The initial conditions at $j = 0$ are given by (93). The final result, denoted by $\mathbf{p}' = \mathbf{p}^J$ and $R_k' = R_k^J$, $k = 1, \ldots, K$, is accepted with probability

$$r = \min\left(1, \frac{\pi(\mathbf{p}', \mathbf{R}')}{\pi(\mathbf{p}, \mathbf{R})}\right). \tag{98}$$

In case of rejection, we continue with the initial $\mathbf{p}$; i.e., in case of rejection we have $\mathbf{p}^{n+1} = \mathbf{p}^n$. In line with the standard HMC method, the vector $\mathbf{R}'$ is entirely discarded after each completed momentum update step.

Note that the acceptance probability $r \to 1$ as $\Delta s \to 0$. This follows from the convergence of the numerical propagator to the exact $\mathcal{R}(s)$ as $\Delta s \to 0$. Hence, as a rule of thumb, we suggest to pick J large enough such that the rejection rate in (98) becomes negligible (e.g., less than 1%) for given $s = \sqrt{2\gamma\Delta t}$.

Implicit midpoint rule. An alternative propagator is obtained by applying the implicit midpoint rule (see, e.g., Leimkuhler and Reich (2005)) to (91)-(92) over a single step with step-size $\Delta s = \sqrt{2\gamma\Delta t}$ and we obtain

$$\mathbf{p}' = \mathbf{p} + \frac{\Delta s}{2}\sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r})(R_k' + R_k), \tag{99}$$

$$R_k' = R_k - \frac{\Delta s}{2}\nabla_{\mathbf{r}} h_k(\mathbf{r}) \cdot M^{-1}(\mathbf{p}' + \mathbf{p}), \qquad k = 1, \ldots, K. \tag{100}$$

The resulting linear equations in $(\mathbf{p}', \mathbf{R}')$ can be solved by a simple fixed point iteration or some other iterative solver. Since only matrix vector multiplications are involved and the matrices involved are typically very sparse, such iterative methods can be implemented efficiently especially on parallel computers.

An appealing aspect of the implementation (99)-(100) is that it conserves the extended energy (88) exactly and, hence, also the corresponding canonical distribution function (94). Since the method also conserves volume and is time-reversible, the proposed momenta $\mathbf{p}'$ are always accepted; i.e., $\mathbf{p}^{n+1} = \mathbf{p}'$, while the vector $\mathbf{R}'$ is entirely discarded after each momentum update step.

Because of the necessary fixed point iteration, the implicit midpoint method (99)-(100) is more expensive than the Störmer-Verlet method ((95)-(97). However, we nevertheless recommend the implicit midpoint method because of the ideal acceptance probability $r = 1$.

Similar to what has been discussed in Section 4.1.1 for the BBK scheme (21)-(23), we find that (99)-(100) is closely related to the self-consistent DPD scheme (24)-(26) for potential energy $V \equiv 0$. Differences can again be found in the treatment of the stochastic contributions and render (24)-(26) non-conservative in terms of the Boltzmann distribution (75) even for $V \equiv 0$.

4.2 A stochastic composition method

We now describe the complete time-stepping methodology. It consists of a composition of two time-stepping methods; one for the conservative part and one for the fluctuation-dissipation part. Note that none of these methods is necessarily the exact propagator to the corresponding stochastic or deterministic equations of motion.

Using an abstract framework, recall that the state variable is $\Gamma = (\mathbf{r}^T, \mathbf{p}^T)^T$ and denote the exact time-Δt transition probability kernel of the stochastic thermostat by $P_{\Delta t}(\Gamma'|\Gamma)$. We now construct a numerical approximation to $P_{\Delta t}(\Gamma'|\Gamma)$ using a composition method approach.

In a first step we apply the Störmer-Verlet method (18)-(20) to the deterministic and conservative part of the stochastic thermostat. The resulting time-step map is denoted by $\Gamma^{n+1} = \Psi_{\Delta t}(\Gamma^n)$ and gives rise to the (singular) transition probability density kernel

$$P_{\mathrm{CD}}(\Gamma'|\Gamma) = \delta_0(\Gamma' - \Psi_{\Delta t}(\Gamma)), \tag{101}$$

where δ_0 denotes the Dirac delta function, i.e.,

$$\int f(x)\,\delta_0(x)\,\mathrm{d}x = f(0). \tag{102}$$

Next a canonical distribution preserving method of Section 4.1 is being used for the fluctuation-dissipation part, which gives rise to a transition probability kernel $P_{\mathrm{FD}}(\Gamma'|\Gamma)$.[3] The proposed complete time-stepping method is characterized by a simple composition of P_{FD} and P_{CD}, i.e., we obtain the transition probability density

$$\mathcal{P}_{\Delta t}(\Gamma^{n+1}|\Gamma^n) = \int P_{\mathrm{FD}}(\Gamma^{n+1}|\Gamma)\,P_{\mathrm{CD}}(\Gamma|\Gamma^n)\,\mathrm{d}\Gamma. \tag{103}$$

Note that, even though P_{FD} conserves the canonical distribution (75), the complete time-stepping method $\mathcal{P}_{\Delta t}$ generally fails to conserve (28), since the conservative time-stepping method $\Psi_{\Delta t}$ does not simultaneously conserve energy and volume. We will next describe a Metropolis correction approach, which takes $\mathcal{P}_{\Delta t}$ as a proposal step and re-enforces the invariance of (28) through a Metropolis acceptance criterion. We essentially obtain an extension of the GHMC method of Horowitz (1991); Kennedy and Pendleton (2001) to the generalized DPD equations (9)-(10).

5 Metropolis corrected time-stepping methods

We have already discussed Metropolis corrected time-stepping methods in Section 3 in the context of Brownian dynamics (36) and its standard Euler-Maruyama approximation (37). These corrections lead to the popular MALA (Roberts and Tweedie, 1995)) and HMC (Duane et al., 1987) schemes. In the context of our composition method (103), we need to correct P_{CD} such that the canonical distribution (28) is preserved. This is essentially the approach taken by the generalized hybrid Monte Carlo (GHMC) method of Horowitz (1991); Kennedy and Pendleton (2001). In this section, we demonstrate how to extend the GHMC methodology to DPD-type momentum updates such that the resulting Markov chain Monte Carlo (MCMC) method samples exactly from the canonical distribution (28).

5.1 Generalized hybrid Monte Carlo (GHMC) method for DPD-type thermostats

We begin by recalling that a Markov process will converge to some distribution of configurations if it is constructed out of updates each of which has the desired distribution as a fixed point, and which taken together are ergodic. The GHMC algorithm for the generation of the canonical density function (28) is constructed of two such steps (Kennedy and Pendleton, 2001). The first part consists of numerical propagation of the system under the conservative dynamics. A Metropolis-Hastings acceptance criterion is applied to correct for numerically induced errors in the conservation of total energy. The second step consists in a update of the momenta. This step needs to be modified to become compatible to DPD and we apply the algorithms developed in Section 4.1.2. We now give an algorithmic summary.

[3] Recall that the positions $\mathbf{r}$ are left constant during the application $P_{\mathrm{FD}}(\Gamma'|\Gamma)$ and that the momentum update preserves the Boltzmann distribution (75). Hence $P_{\mathrm{FD}}(\Gamma'|\Gamma)$ also leaves the canonical distribution (28) invariant. But it is clearly not ergodic with respect to (28).

5.1.1 Meso-GHMC: Algorithmic summary

The meso-GHMC method is defined through an energy/Hamiltonian (1), inverse temperature $\beta = 1/k_BT$, a set of position-dependent functions $\{h_k(\mathbf{r})\}_{k=1}^K$, friction constant γ, time-step Δt, and number of time-steps L. The method generates a sequence of states Γ^j, $j = 1, \ldots, J$. We now summarize a single step of the meso-GHMC method.

(i) *Conservative dynamics step.* Given the last accepted state Γ^j with position and momentum vectors $\mathbf{r}$ and $\mathbf{p}$, respectively, we numerically integrate the Hamiltonian equations of motion (3)-(4) over L time-steps with the Störmer-Verlet method, (18)-(20), step-size Δt, and initial conditions $\mathbf{r}^0 = \mathbf{r}$, $\mathbf{p}^0 = \mathbf{p}$. This results in the approximation $(\mathbf{r}^L, \mathbf{p}^L)$. The accepted pair of position and momentum vectors $(\bar{\mathbf{r}}, \bar{\mathbf{p}})$ is obtained via a Metropolis accept/reject test of the form

$$(\bar{\mathbf{r}}, \bar{\mathbf{p}}) = \begin{cases} (\mathbf{r}^L, \mathbf{p}^L) & \text{with probability } \min(1, \exp(-\beta\,\delta E)) \\ (\mathbf{r}^0, -\mathbf{p}^0) & \text{otherwise} \end{cases}, \tag{104}$$

where

$$\delta E := E(\mathbf{r}^L, \mathbf{p}^L) - E(\mathbf{r}^0, \mathbf{p}^0). \tag{105}$$

(ii) *Momentum refreshment step.* A sequence of i.i.d. random numbers $R_k \sim \mathrm{N}(0, k_BT)$, $k = 1, \ldots, K$, is generated. Using the implicit midpoint rule implementation of the momentum refreshment step, the system (99)-(100) is solved for $(\mathbf{p}', \{R_k'\})$ by fixed point iteration with initial momentum $\mathbf{p} = \bar{\mathbf{p}}$ and fixed position vector $\mathbf{r} = \bar{\mathbf{r}}$.

(iii) The newly accepted state Γ^{j+1} is provided by the position vector $\bar{\mathbf{r}}$ (from the conservative dynamics part) and the momentum $\mathbf{p}'$ (from the momentum refreshment step), respectively.

5.1.2 Remark

It should be noted that the Metropolis criterion (104) leads to a trajectory reversal upon rejection of the conservative dynamics proposal step; i.e., the method continues with the previous position vector $\mathbf{r}^0$ and negated momentum vector $\mathbf{p}^0$ upon rejection of the proposal step. While this trajectory reversal is required for detailed balance with respect to the canonical distribution (28) it also implies that rejections interfere strongly with the dynamics. It has been proposed by Akhmatskaya et al. (2009) to replace (104) by

$$(\bar{\mathbf{r}}, \bar{\mathbf{p}}) = \begin{cases} (\mathbf{r}^L, \mathbf{p}^L) & \text{with probability } \min(1, \exp(-\beta\,\delta E)) \\ (\mathbf{r}^0, \mathbf{p}^0) & \text{otherwise} \end{cases}. \tag{106}$$

Unfortunately, detailed balance does no longer holds under (106) (Akhmatskaya and Reich, 2009). However, it has been demonstrated for simple test problems that (106) increases the sampling accuracy compared to (103) and interferes less with the dynamics of the stochastic thermostats than (104) (Akhmatskaya et al., 2009; Akhmatskaya and Reich, 2009). See also Leimkuhler and Reich (2009) for related results on Metropolis corrected stochastic Nosé-Hoover dynamics (Leimkuhler et al., 2009). In this paper, we implement meso-GHMC with the Metropolis criterion (104) to demonstrate exact sampling from the canonical distribution (28).

5.2 Generalized shadow hybrid Monte Carlo (GSHMC) method for DPD-type thermostats

The key idea of the generalized shadow hybrid Monte Carlo (GSHMC) method of Akhmatskaya and Reich (2008) is to assess the Monte Carlo steps of GHMC with regard to a shadow Hamiltonian $\mathcal{E}_{\Delta t}$,

which increases the acceptance rate in the conservative dynamics part of GHMC. A shadow energy can either be found using the methods of Akhmatskaya and Reich (2006, 2008) or Skeel and Hardy (2001). See Sections 3.4 for specific formulations of fourth and eighth-order shadow energies and Section 3.5 for a numerical demonstration of the reduced relative errors in the shadow energies along numerical trajectories from a Störmer-Verlet simulation.

We now outline the generalization of meso-GHMC to GSHMC. We will call the resulting method meso-GSHMC.

5.2.1 Momentum refreshment step

To put the meso-GHMC method in the context of the GSHMC method of Akhmatskaya and Reich (2008), we need to modify the partial momentum update step, as defined by (99)-(100) (in case of the implicit midpoint implementation). Given a shadow energy $\mathcal{E}_{\Delta t}$, the key idea is to replace the extended canonical density (94) by

$$\widehat{\pi}(\mathbf{r},\mathbf{p},\mathbf{R}) \propto \exp\left(-\beta\mathcal{E}_{\Delta t}(\mathbf{r},\mathbf{p}) - \beta/2\sum_{k=1}^{K}(R_k)^2\right), \tag{107}$$

and the acceptance probability (98) by

$$r = \min\left(1, \frac{\widehat{\pi}(\mathbf{r},\mathbf{p}',\mathbf{R}')}{\widehat{\pi}(\mathbf{r},\mathbf{p},\mathbf{R})}\right). \tag{108}$$

Recall that $r = 1$ for the implicit midpoint implementation (99)-(100) and $\mathcal{E}_{\Delta t} = E$. However, this is no longer the case for $\mathcal{E}_{\Delta t} \neq E$. This is a drawback of GSHMC, which can be partially overcome by GS2HMC (see Akhmatskaya and Reich (2008) for details).

Since the meso-GSHMC method samples with respect to a modified canonical ensemble, it is necessary to re-weight the computed samples $\{\Omega_j\}$ of an observable $\Omega = \Omega(\mathbf{r},\mathbf{p})$. See Akhmatskaya and Reich (2006, 2008) and the algorithmic summary below for details.

5.2.2 Meso-GSHMC: Algorithmic summary

The meso-GSHMC method is defined through an energy/Hamiltonian (1), a shadow energy $\mathcal{E}_{\Delta t}$, inverse temperature $\beta = 1/k_BT$, a set of position-dependent functions $\{h_k(\mathbf{r})\}_{k=1}^K$, friction constant γ, time-step Δt, and number of time-steps L. The method generates a sequence of states Γ^j, $j = 1,\ldots,J$. We now summarize a single step of the meso-GSHMC method.

(i) *Conservative dynamics step.* Given the last accepted state Γ^j with a pair of position and momentum vectors $(\mathbf{r},\mathbf{p})$, we numerically integrate the Hamiltonian equations of motion (3)-(4) over L time-steps with the Störmer-Verlet method, (18)-(20), step-size Δt, and initial conditions $\mathbf{r}^0 = \mathbf{r}$, $\mathbf{p}^0 = \mathbf{p}$. This results in the approximation $(\mathbf{r}^L,\mathbf{p}^L)$. The accepted pair of position and momentum vectors $(\bar{\mathbf{r}},\bar{\mathbf{p}})$ is obtained via the Metropolis accept/reject test

$$(\bar{\mathbf{r}},\bar{\mathbf{p}}) = \begin{cases} (\mathbf{r}^L,\mathbf{p}^L) & \text{with probability } \min(1,\exp(-\beta\,\delta\mathcal{E}_{\Delta t})) \\ (\mathbf{r}^0,-\mathbf{p}^0) & \text{otherwise} \end{cases}, \tag{109}$$

where

$$\delta\mathcal{E}_{\Delta t} := \mathcal{E}_{\Delta t}(\mathbf{r}^L,\mathbf{p}^L) - \mathcal{E}_{\Delta t}(\mathbf{r}^0,\mathbf{p}^0). \tag{110}$$

(ii) *Momentum refreshment step.* A sequence of i.i.d. random numbers $R_k \sim \mathrm{N}(0, \beta^{-1})$, $k = 1, \ldots, K$, is generated. Using the implicit midpoint rule implementation of the momentum refreshment step, the system (99)-(100) is solved for $(\mathbf{p}', \mathbf{R}')$ by fixed point iteration and initial momentum $\mathbf{p} = \bar{\mathbf{p}}$ and fixed position $\mathbf{r} = \bar{\mathbf{r}}$. The accepted momentum vector $\mathbf{p}''$ is obtained via the Metropolis criterion

$$\mathbf{p}'' = \begin{cases} \mathbf{p}' & \text{with probability } \min(1, \exp(-\beta\, \delta\mathcal{E}_{\Delta t,\mathrm{ext}})) \\ \bar{\mathbf{p}} & \text{otherwise} \end{cases}, \tag{111}$$

where

$$\delta\mathcal{E}_{\Delta t,\mathrm{ext}} := \left[\mathcal{E}_{\Delta t}(\bar{\mathbf{r}}, \mathbf{p}') + \frac{1}{2}\sum_{k=1}^{K}(R_k')^2\right] - \left[\mathcal{E}_{\Delta t}(\bar{\mathbf{r}}, \bar{\mathbf{p}}) + \frac{1}{2}\sum_{k=1}^{K}(R_k)^2\right]. \tag{112}$$

(iii) The newly accepted pair of position and momentum vectors is provided by $\bar{\mathbf{r}}$ (from the conservative dynamics part) and $\mathbf{p}''$ (from the momentum refreshment step), respectively, and give rise to Γ^{j+1}.

Under the assumption of ergodicity of the induced Markov chain, the ensemble average of an observable $\Omega(\Gamma) = \Omega(\mathbf{r}, \mathbf{p})$ with respect to the canonical ensemble (28) is approximated as

$$\langle\Omega\rangle = \frac{\sum_{j=1}^{J} w_j \Omega(\Gamma^j)}{\sum_{j=1}^{J} w_j} \tag{113}$$

where

$$w_j = \exp\left(\beta\left[\mathcal{E}_{\Delta t}(\Gamma^j) - E(\Gamma^j)\right]\right). \tag{114}$$

6 Numerical results

6.1 DPD test systems

Numerical results from the meso-GHMC/GSHMC methods are compared to the MD-VV implementation (Vattulainen et al., 2002) of DPD (for simplicity of presentation, we start with half a time-step in the positions):

$$\mathbf{q}^{n+1/2} = \mathbf{q}^n + \frac{\Delta t}{2} M^{-1}\mathbf{p}^n, \tag{115}$$

$$\begin{aligned}\mathbf{p}^{n+1} = {} & \mathbf{p}^n - \Delta t \nabla_{\mathbf{r}} V(\mathbf{r}^{n+1/2}) \\ & - \sum_{k=1}^{K} \nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1/2}) \left[\gamma\Delta t \nabla_{\mathbf{r}} h_k(\mathbf{r}^{n+1/2}) \cdot M^{-1}\mathbf{p}^n + (2\gamma\Delta t)^{1/2} R_k\right],\end{aligned} \tag{116}$$

$$\mathbf{q}^{n+1} = \mathbf{q}^{n+1/2} + \frac{\Delta t}{2} M^{-1}\mathbf{p}^{n+1}, \tag{117}$$

where $R_k \sim \mathrm{N}(0, k_B T)$, $k = 1, \ldots, K$ are i.i.d. random numbers.

Numerical experiments are conducted for Model A and Model C of Vattulainen et al. (2002). For the chosen units, we have $k_B T = 1$ in dimensionless variables.

6.1.1 Model A

We consider a total of $N = 4000$ particles with mass $m = 1$ in a cubic domain of size $10\times10\times10$ with periodic boundary conditions. The conservative forces are set equal to zero, i.e., the Hamiltonian (1) reduces to

$$\mathcal{H} = \frac{1}{2}\mathbf{p}^T\mathbf{p} \tag{118}$$

and the resulting equations of motion can be solved exactly. Hence we formally use $\widehat{\mathcal{H}}_{\Delta t} = \mathcal{H}$ as a modified energy for meso-GSHMC and it is sufficient to only implement meso-GHMC for Model A. We always set $L = 1$ in the conservative dynamics part of meso-GHMC, i.e., $\tau = \Delta t$ and perform experiments for different values of the step-size Δt.

The functions $\{h_k(\mathbf{r})\}$ for the momentum refreshment step are defined via

$$\phi'(r) = \begin{cases} 1 - r/r_c, & \text{for } r \le r_c, \\ 0, & \text{for } r > r_c \end{cases} \tag{119}$$

in (12). The cutoff distance is set equal to $r_c = 1$. We also use $\alpha = (2\gamma\Delta t)^{1/2}$ for all Monte Carlo simulations with $\gamma = 4.5$.

The reference experiments with the DPD method (115)-(117) use the same parameter settings.

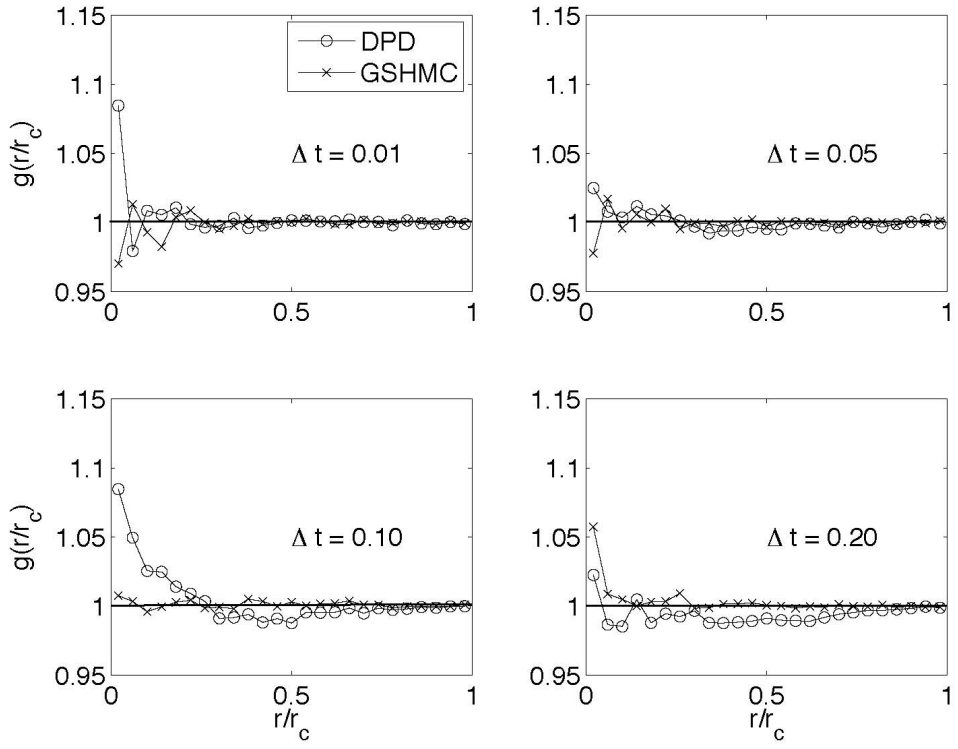

Figure 4: Radial distribution function $g(r/r_c)$ for different values of the step-size Δt in model A. The analytic value is $g(r/r_c) = 1$ and is indicated by a solid reference line. Numerical results are obtained from a standard DPD integration scheme and the newly proposed meso-GHMC/GSHMC method for different values of the step-size Δt.

The numerical results from the meso-GHMC and traditional DPD simulations can be found in Figs. 4-6. Following the argument of Peters (2004), Fig. 5 demonstrates that the deviations of the numerically computed radial distribution function $g(r/r_c)$ from its exact value $g = 1$ is of purely statistical origin (finite sample size) for the meso-GHMC method. The same does not hold for the

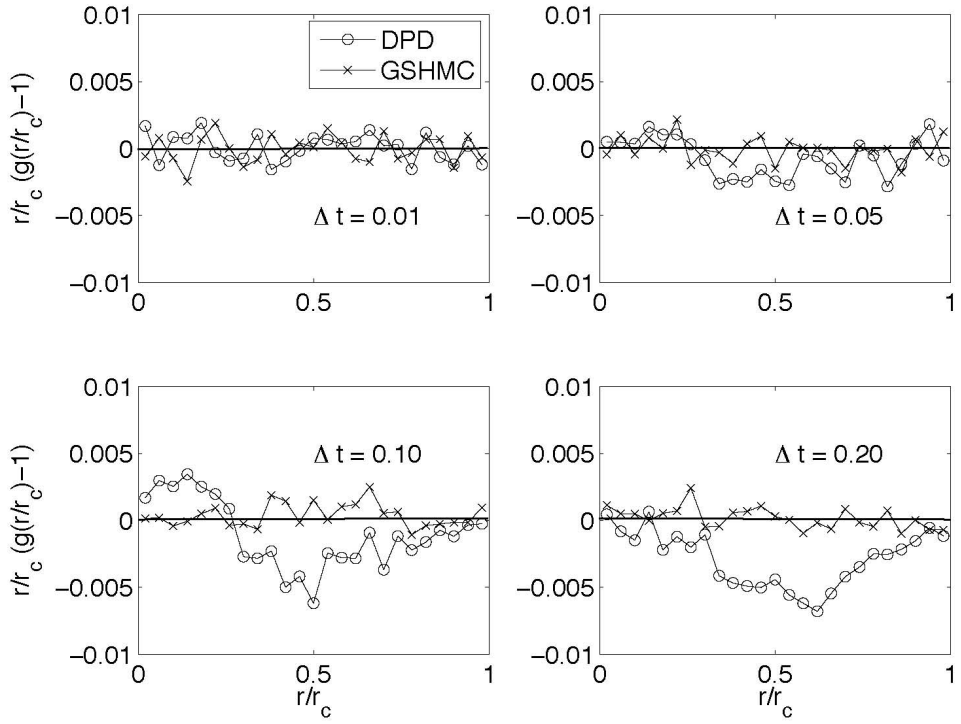

Figure 5: To distinguish statistically from numerically induced deviations of $g(r/r_c)$ we also provide $(r/r_c)(g(r/r_c)-1)$ for the different step-sizes and methods. The exact value is indicated by a solid reference line. It can be concluded that the deviations in the radial distribution function are entirely of statistical nature for the meso-GHMC/GSHMC method.

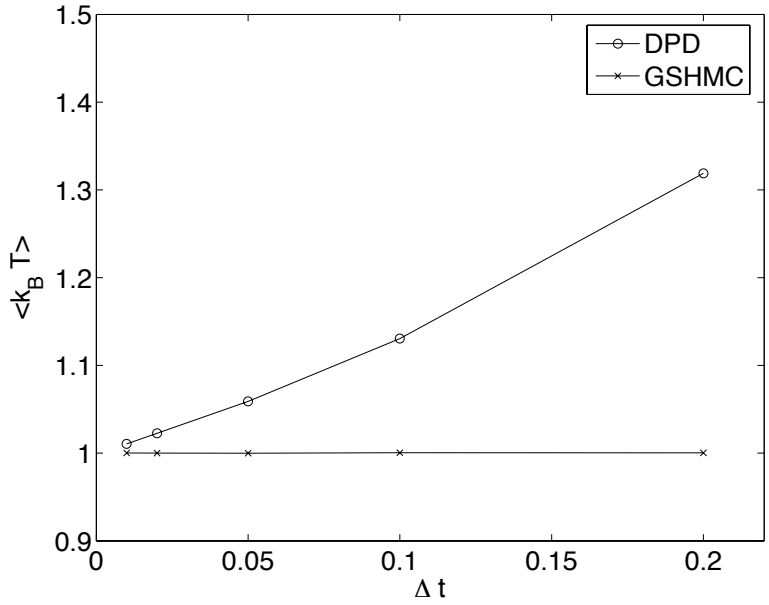

Figure 6: Numerically observed temperature $\langle k_B T\rangle$ vs the step-size Δt in Model A. Results are obtained from a standard DPD integration scheme and the newly proposed meso-GHMC/GSHMC method. The correct result is $\langle k_B T\rangle = 1$.

DPD method unless $\Delta t \leq 0.05$. We also note that the meso-GHMC method exactly reproduces the target inverse temperature $\beta = 1$ for all values of Δt, while the DPD method (115)-(117) leads to a

nearly linear increase in the numerically observed temperature with respect to the step-size Δt (see Fig. 6).

6.1.2 Model C

Model C is a simple interacting Lennard-Jones fluid with truncated pairwise interaction potential

$$U(\mathbf{r}_{ij}) = \begin{cases} 4\left[\left(\frac{l}{r_{ij}}\right)^{12} - \left(\frac{l}{r_{ij}}\right)^{6}\right], & r_{ij} \leq r_c, \\ 0, & r_{ij} > r_c, \end{cases} \tag{120}$$

with $r_c = 1$ and $l = 2^{-1/6}$. The simulation box is of size $16\times16\times16$ with a total of $N = 2867$ particles. This corresponds to a density of $\rho = 0.7$.

The conservative dynamics part is implemented with $\tau = 0.05$ and varying values for Δt and, hence, $L = \tau/\Delta t$, while the momentum refreshment step is implemented as for Model A with the only difference being that $\alpha = (2\gamma\tau)^{1/2} \approx 4.4721$ for all Monte Carlo simulations with $\gamma = 200$ ($\sigma = 20$, respectively).

The meso-GSHMC method is implemented with a fourth-order accurate shadow energy as $\mathcal{E}_{\Delta t}$. The numerical experiments are conducted with the fourth-order shadow energy of Akhmatskaya and Reich (2008). An implementation of (71) would not alter the results significantly. Note that the truncated interaction potential (120) leads to a continuous only force field. Higher regularity of the force fields is required to achieve a fourth-order accuracy in the shadow energy. The fourth-order behavior is indeed not observed in our numerical experiments. See Hafskjold et al. (2004) for the use of smoother truncation schemes in the context of DPD. The reference experiments with the DPD method (115)-(117) use the same parameter settings.

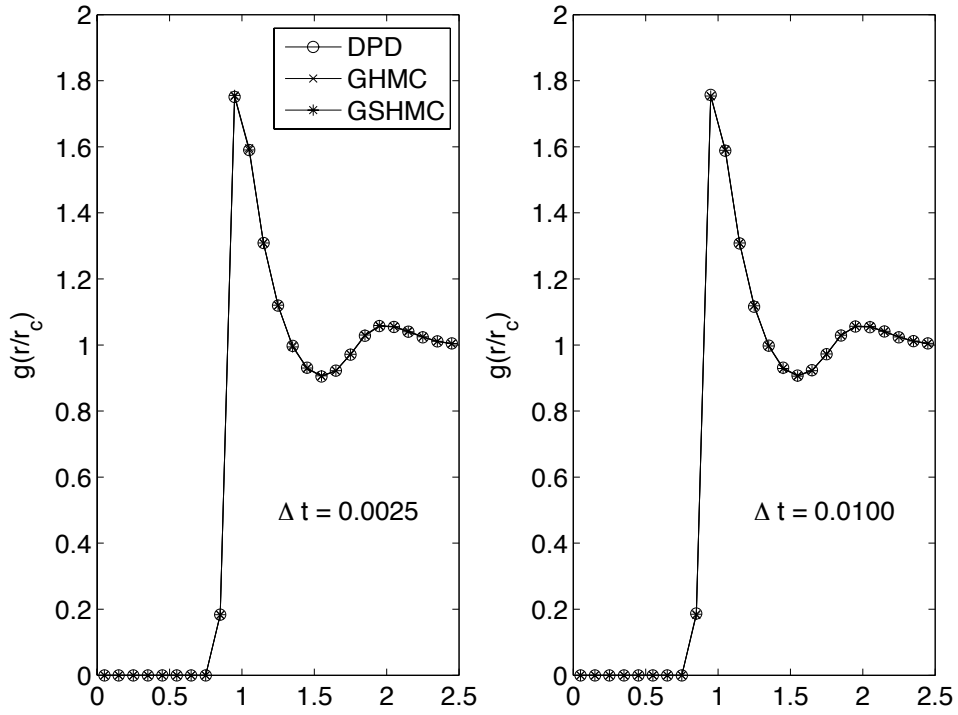

Figure 7: Radial distribution function $g(r/r_c)$ for different values of the step-size Δt in model C. Results are obtained from a standard DPD integration scheme and the newly proposed meso-GHMC/GSHMC methods.

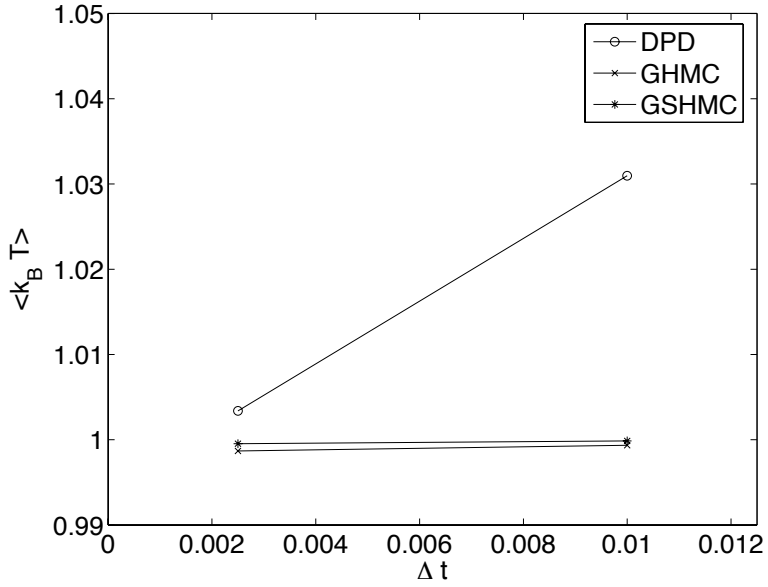

Figure 8: Numerically observed temperature $\langle k_B T \rangle$ vs the step-size Δt in Model C. Results are obtained from a standard DPD integration scheme and the newly proposed meso-GHMC/GSHMC methods. The correct result is $\langle k_B T \rangle = 1$.

method	step-size	rejection rate in CD step	rejection rate in MR step
meso-GHMC	0.0025	1.65%	0%
meso-GSHMC	0.0025	1.28%	0.11%
meso-GHMC	0.0100	26.51%	0%
meso-GSHMC	0.0100	20.64%	1.64%

Table 1: Rejection rates in the conservative dynamics (CD) and momentum refreshment (MR) steps of meso-GHMC and meso-GSHMC, respectively, for different values of the step-size Δt and constant $\tau = L\Delta t = 0.05$.

We conclude from Fig. 7 that the computed radial distribution functions are in nearly perfect agreement for all simulation methods and step-sizes Δt. Note, however, that Vattulainen et al. (2002) do *not* specify the number of particles for Model C and, hence, our numerical results differ slightly from the findings in Vattulainen et al. (2002). We also confirm in Fig. 8 that the Monte Carlo methods reproduce the target value $k_B T = 1$.

We finally state the rejection rates for meso-GHMC and meso-GSHMC in Table 1. The meso-GSHMC slightly reduces the rejection rate in the conservative dynamics part. This is at the expense of a non-zero rejection rate in the momentum refreshment step. It appears that the meso-GHMC method is optimal for force fields with a non-smooth cut-off and that improvements through the use of shadow energies are limited due to the non-smooth nature of the cut-off. A smooth truncation of the force field has been discussed by Hafskjold et al. (2004) in the context of DPD and it has been demonstrated that smooth cut-offs leads to better energy conservation.

6.2 Membrane protein

Meso-scale simulations techniques are increasingly applied to problems from biophysical chemistry such as the interaction of macromolecules with the interface between an aqueous solution and bio-

logical membranes; i.e., phospholipid bilayer. These simulations are based on coarse-grained (CG) approaches which replace the all atom classical mechanics force-fields of traditional molecular dynamics. As an example we mention the studies of Bond and Sansom (2006); Bond et al. (2007); Shih et al. (2006) on CG simulations of simple peptides and integral membranes interacting with a lipid bilayer. In addition to finding more accurate CG force fields, there also remains the need to improve accuracy and efficiency of sampling to guarantee convergence of CG simulations. Since CG simulations are typically performed in a NVT ensemble setting, stochastic thermostats such as those considered in this paper can be applied. A comparison between standard constant temperature CG-MD simulations using a Berendsen thermostat with the GSHMC method, as described in this paper, has been conducted in Wee et al. (2008). The position-dependent functions $\{h_k(\mathbf{r})\}_{k=1}^K$ are defined by (13) and the test application is a gating-modifier peptide toxin in a membrane with environment. See Wee et al. (2008) for a detailed description of the model system, simulation set-up, and numerical results. In Figure 9, the evolution of the toxin is shown from an initial position within the membrane to its preferred position at the surface of the bilayer. The GSHMC method leads to a much more rapid repositioning of the toxin compared to a CG-MD simulation with Berendsen thermostat.

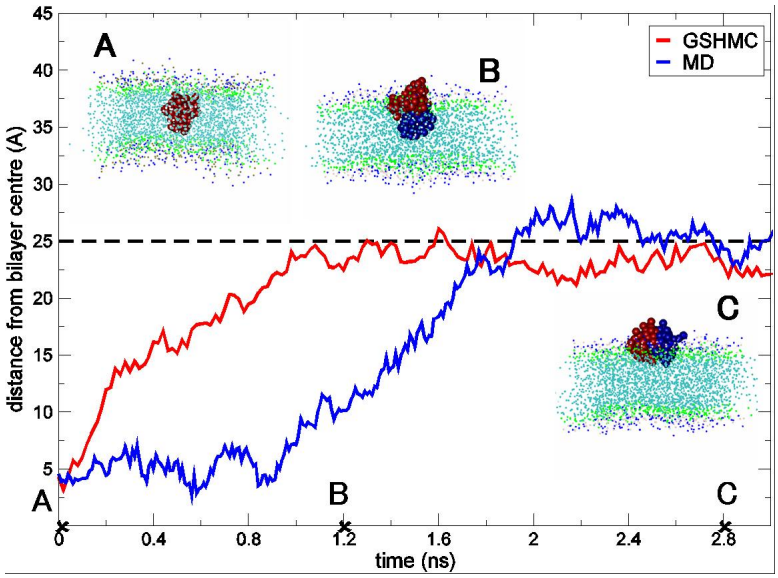

Figure 9: Displayed is the evolution of peptide toxin from an initial position within the membrane to its preferred position at the surface of the membrane. More specifically, the distance of the toxin to an energy minimized reference position at the surface of the membrane is shown for a CG-MD and a GSHMC simulation. While both simulation methods correctly identify the preferred location of the toxin, the toxin equilibrates much faster under the GSHMC simulation.

7 Concluding remarks

In this paper, we have investigated local stochastic thermostats such as Langevin and dissipative particle dynamics. We have proposed simple modifications to existing methods which sample exactly from the Boltzmann distribution (75) for position-dependent fluctuation-dissipation terms and force free motion. Alternative methods with the same property have been proposed in the past but rely on a splitting into pairwise interactions (Lowe, 1999; Shardlow, 2003; Peters, 2004; Koopman and Lowe, 2006) and can therefore not easily be applied to very large systems. As a completely

novel contribution, we have also shown how to put stochastic thermostats with position-dependent fluctuation-dissipation terms, for example DPD, within the framework of Markov chain Monte Carlo methods which implies rigorous sampling from the canonical distribution (28) regardless of the chosen step-size Δt. We have verified this property numerically for Model A and C from Vattulainen et al. (2002). We have also demonstrated for a membrane protein system that Metropolis corrected time-stepping methods are efficient for applications which require a careful and rapid equilibration with respect to a given target temperature.

The proposed methods can be viewed as Metropolis corrected time-stepping methods similar to the MALA/HMC scheme for first-order Brownian dynamics. However, it should be noted that the Metropolis test induces non-trivial changes to the dynamics in case of rejection of the proposal step. A possible compromise between sampling accuracy and reduced interference with dynamics is to remove the momentum reversal upon rejection in (104) and (109), respectively, from the meso-GHMC/GSHMC methods and to use a Metropolis test of type (106). See Akhmatskaya et al. (2009); Akhmatskaya and Reich (2009) for some initial results, which indicate that GHMC/GSHMC without momentum flip improves the accuracy of standard time-stepping methods while interfering little with dynamic properties such as autocorrelation functions provided the acceptance rate is kept sufficiently high. This aspect requires further investigations.

References

E. Akhmatskaya and S. Reich. The targeted shadowing hybrid Monte Carlo (TSHMC) method. In B. Leimkuhler et al, editor, *New Algorithms for Macromolecular Simulations*, volume 49 of *Lecture Notes in Computational Science and Engineering*, pages 145–158, Berlin, 2006. Springer-Verlag.

E. Akhmatskaya and S. Reich. GSHMC: An efficient method for molecular simulations. *J. Comput. Phys.*, 227:4934–4954, 2008.

E. Akhmatskaya and S. Reich. Erratum for "generalized hybrid Monte Carlo methods without momentum flip". *J. Comput. Phys.*, submitted, 2009.

E. Akhmatskaya, N. Bou-Rabee, and S. Reich. Generalized hybrid Monte Carlo methods without momentum flip. *J. Comput. Phys.*, 228:2256–2265, 2009.

M.P. Allen. Algorithms for Brownian dynamics. *Mol. Phys.*, 47:599–601, 1982.

M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1987.

H.C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384–2393, 1980.

V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag, New York, 2nd edition, 1989.

G. Besold, I. Vattulainen, M. Karttunen, and J.M. Polson. Towards better integrators for dissipative particle dynamics. *Phys. Rev. E*, 62:R7611–R7614, 2000.

P. J. Bond and M. S. P. Sansom. Insertion and assembly of membrane proteins via simulation. *J. Amer. Chem. Soc.*, 128:2697–2704, 2006.

P.J. Bond, J. Holyoake, A. Ivetac, S. Khalid, and M. S. P. Sansom. Coarse-grained molecular dynamics simulations of membrane proteins and peptides. *J. Struct. Biol.*, 157:593–605, 2007.

A. Brünger, C.B. Brooks, and M. Karplus. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.*, 105, 1984.

G. Bussi and M. Parrinello. Accurate sampling using Langevin dynamics. *Physical Review E*, 75: 056707, 2007.

G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126:014101, 2007.

E. Cances, F. Legoll, and G. Stoltz. Theoretical and numerical comparison of some sampling methods. *M2AN*, 41:351–390, 2007.

C.J. Cotter and S. Reich. An extended dissipative particle dynamics model. *Europhys. Lett.*, 64: 723–729, 2003.

S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid Monte-Carlo. *Phys. Lett. B*, 195: 216–222, 1987.

R.D. Engle, R.D. Skeel, and M. Drees. Monitoring energy drift with shadow hamiltonians. *J. Comput. Phys.*, 206:432–452, 2005.

P. Español. Dissipative Particle Dynamics. In V.M. Harik and M.D. Salas, editors, *Trends in nanoscale mechanics: Analysis of nanostructure materials and multi-scale modeling*, pages 1–23. Kluwer, 2003.

P. Español and P.B. Warren. Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.*, 30:191–196, 1995.

G. De Fabritis, M. Serrano, P. Español, and P.V. Coveney. Efficient numerical integrators for stochastic models. *Physica A*, 361:429–440, 2006.

S.E. Feller, Y. Zhang, R.W. Pastor, and B.R. Brooks. Constant pressure molecular dynamics: The Langevin piston method. *J. Chem. Phys.*, 103:4613–4621, 1995.

B. Forrest and U. Suter. Hybrid Monte Carlo simulations of dense polymer systems. *J. Chem. Phys.*, 101:2616–2629, 1994.

D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, New York, 2nd edition, 2001.

W.F. Van Gunsteren and H.J.C. Berendsen. Algorithms for Brownian dynamics. *Mol. Phys.*, 45: 637–647, 1982.

B. Hafskjold, C.C. Liew, and W. Shinoda. Can such long time steps really be used in Dissipative Particle Dynamics simulations. *Mol. Sim.*, 30:879–885, 2004.

E. Hairer, Ch. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer-Verlag, Berlin Heidelberg, 2002.

E. Hairer, Ch. Lubich, and G. Wanner. Geometric numrical integration by the Störmer-Verlet method. *Acta Numerica*, 12:399–450, 2003.

D.J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43:525–546, 2001.

P.J. Hoogerbrugge and J.M.V.A. Koelman. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett*, 19:155–160, 1992.

A.M. Horowitz. A generalized guided Monte-Carlo algorithm. *Phys. Lett. B*, 268:247–252, 1991.

J.A. Izaguirre and S.S. Hampton. Shadow Hybrid Monte Carlo: An efficient propagator in phase space of macromolecules. *J. Comput. Phys.*, 200:581–604, 2004.

J.A. Izaguirre, D.P. Catarello, J.M. Wozniak, and R.D. Skeel. Langevin stabilization of molecular dynamics. *J. Chem. Phys.*, 114:2090–2098, 2001.

A.D. Kennedy and B. Pendleton. Cost of the generalized hybrid Monte Carlo algorithm for free field theory. *Nucl. Phys. B*, 607:456–510, 2001.

E.A. Koopman and C.P. Lowe. Advantages of a Lowe-Andersen thermostat in molecular dynamics simulations. *J. Chem. Phys.*, 124:204103, 2006.

B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge University Press, Cambridge, 2005.

B. Leimkuhler and S. Reich. A Metropolis adjusted Nosé-Hoover thermostat. *M2AN*, 2009.

B. Leimkuhler, E. Noorizadeh, and F. Theil. A gentle ergodic thermostat for molecular dynamics. *J. Stat. Phys.*, page in print, 2009.

J.S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer-Verlag, New York, 2001.

C.P. Lowe. An alternative approach to dissipative particle dynamics. *Europhys. Lett.*, 47:145–151, 1999.

J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for sdes and approximations: locally lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

B. Mehlig, D.W. Heermann, and B.M. Forrest. Hybrid Monte Carlo method for condensed-matter systems. *Phys. Rev. B*, 45:679–685, 1992.

S. P. Meyn and R.L. Tweedie. Markov chains and stochastic stability. *Springer-Verlag, London,*, 1993.

P. Nikunen, M. Karttunen, and I. Vattulainen. How would you integrate the equations of motion in dissipative particle dynamics. *Computer Physics Communications*, 153:407–423, 2003.

B. Oksendal. *Stochastic Differential Equations*. Springer-Verlag, Berlin-Heidelberg, 5th edition, 2000.

I. Pagonabarraga, M.H.J. Hagen, and D. Frenkel. Self-consistent dissipative particle dynamics. *Europhys. Lett.*, 42:377–382, 1998.

R.W. Pastor, B.R. Brooks, and A. Szabo. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.*, 65:1409–1419, 1988.

E.A.J.F. Peters. Elimination of time step effects in DPD. *Europhys. Lett.*, 66:311–317, 2004.

Gareth. O. Roberts and Richard. L. Tweedie. Exponential convergence of langevin diffusions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1995.

A. Samoletov, M. A. J. Chaplain, and C. P. Dettmann. Thermostats for "slow" configurational modes. *J. Stat. Phys.*, 128:1321–1336, 2007.

A. Scemama, T. Leliévre, G. Stoltz, E. Cancés, and M Caffarel. An efficient sampling algorithm for variational Monte Carlo. *J. Chem. Phys.*, 125:114105, 2006.

M. Serrano, G. De Fabritiis, P. Español, and P.V. Coveney. A stochastic Trotter integration scheme for dissipative particle dynamics. *Mathematics and Computers in Simulation*, 72:190–194, 2006.

T. Shardlow. Splitting for dissipative particle dynamics. *SIAM J. Sci. Comput.*, 24:1267–1282, 2003.

T. Shardlow and Y. Yan. Geometric ergodicity for dissipative particle dynamics. *Stochastics and Dynamics*, 6(1):123–154, 2006.

A.Y. Shih, A. Arkhipov, P.L. Freddolino, and K. Schulten. Coarse grained protein-lipid model with application to lipoprotein particles. *J. Phys. Chem. B*, 110(8):3674–3684, 2006.

R.D. Skeel. Integration schemes for molecular dynamics and related applications. In M. Ainsworth, J. Levesley, and M. Marletta, editors, *The Graduate Student's Guide to Numerical Analysis*, volume 4 of *SSCM*, pages 119–176, Berlin, 1999. Springer-Verlag.

R.D. Skeel and D.J. Hardy. Practical construction of modified Hamiltonians. *SIAM J. Sci. Comput.*, 23:1172–1188, 2001.

R.D. Skeel and J.A. Izaguirre. An impulse integrator for Langevin dynamics. *Mol. Phys.*, 100: 3885–3891, 2002.

I. Vattulainen, M. Karttunen, B. Besold, and J.M. Polson. Integration schemes for dissipative particle dynamics simulations: From softly interacting systems towards hybrid models. *J. Chem. Phys.*, 116:3967–3979, 2002.

L. Verlet. Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Lett.*, 159:98–103, 1967.

W. Wang and R.D. Skeel. Analysis of a few numerical integration methods for the Langevin equation. *Mol. Phys.*, 101:2149–2156, 2003.

C.L. Wee, M.S.P. Sansom, S. Reich, and E. Akhmatskaya. Improved sampling for simulations of interfacial membrane proteins: Application of GSHMC to a peptide toxin/bilayer system. *J. Phys. Chem. B*, 112:5710–5717, 2008.