

Multiple-time-stepping generalized hybrid Monte Carlo methods

Bruno Escribano^a, Elena Akhmatskaya^{a,b}, Sebastian Reich^c, Jon M. Azpiroz^d

^a*BCAM - Basque Center for Applied Mathematics, E-48009 Bilbao, Spain*

^b*IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain*

^c*Universität Potsdam, Institut für Mathematik, D-14469 Potsdam, Germany*

^d*Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU) and Donostia International Physics Center (DIPC), P.K. 1072, Donostia, Spain*

Abstract

Performance of the generalized shadow hybrid Monte Carlo (GSHMC) method [1], which proved to be superior in sampling efficiency over its predecessors [2–4], molecular dynamics and hybrid Monte Carlo, can be further improved by combining it with multi-time-stepping (MTS) and mollification of slow forces. We demonstrate that the comparatively simple modifications of the method not only lead to better performance of GSHMC itself but also allow for beating the best performed methods, which use the similar force splitting schemes. In addition we show that the same ideas can be successfully applied to the conventional generalized hybrid Monte Carlo method (GHMC). The resulting methods, MTS-GHMC and MTS-GSHMC, provide accurate reproduction of thermodynamic and dynamical properties, exact temperature control during simulation and computational robustness and efficiency. MTS-GHMC uses a generalized momentum update to achieve weak stochastic stabilization to the molecular dynamics (MD) integrator. MTS-GSHMC adds the use of a shadow (modified) Hamiltonian to filter the MD trajectories in the HMC scheme. We introduce a new shadow Hamiltonian formulation adapted to force-splitting methods. The use of such Hamiltonians improves the acceptance rate of trajectories and has a strong impact on the sampling efficiency of the method. Both methods were implemented in the open-source MD package *ProtoMol* and were tested on a water and a protein systems. Results were compared to those obtained using a Langevin Molly (LM) method [5] on the same systems. The test results demonstrate the superiority of the new methods over LM in terms of stability, accuracy and sampling efficiency. This suggests that putting the MTS approach in the framework of hybrid Monte Carlo and using the natural stochasticity offered by the generalized hybrid Monte Carlo lead to improving stability of MTS and allow for achieving larger step sizes in the simulation of complex systems.

Keywords:

Force splitting, mollification, generalized hybrid Monte Carlo, Molecular dynamics,

1. Introduction

Molecular dynamics (MD) is a useful tool for theoretical investigation of molecular systems. A primary limitation in the application of MD to the study of complex processes involving macromolecules, e.g. biomolecules, is the small time step size of conventional MD. Whereas the latter is typically measured in femtoseconds, dynamical processes of interest happen in microseconds and longer time scales. Bridging the time scale gap between simulations and the phenomena of interest has been an area of active research for more than a decade. One area of intense focus is design of numerical algorithms, which accelerate MD simulations by making use of the multi-scale nature of the macromolecular systems. Multiple-time-stepping (MTS) methods are among the most popular approaches of this type. The methods are based on the idea to build multiple time step integrators for molecular dynamics that allow for time steps of differing lengths according to how rapidly a given type of interaction is evolving in time. When simulating biological systems such as proteins one must take into consideration atomic vibrations in the order of femtoseconds as well as collective motions in the order of milliseconds. MTS methods use different time steps for the integration of different time scale processes, adapting the computational performance to the molecular system being simulated. For example, the most computationally expensive long range electrostatic interactions only contribute to the dynamics on a relatively long time scale and it is not necessary to calculate them as often as other interactions such as bonded or angle interactions. In this way, it is feasible to save computational time by increasing the time step without sacrificing accuracy.

The prototypical algorithm is the Verlet-I/r-RESPA/Impulse integrator [6, 7], which splits the forces into fast and slow components and evaluates the former more frequently than the latter. The ratio between frequencies of evaluation of the slow forces and fast forces, so called the step-size ratio, measures the gain in simulation performance.

Splitting forces can be achieved in different ways and can include several levels [8]. For example initial splitting in bonded and non-bonded forces can be followed by splitting further each group of forces according to their relative intrinsic time scales. In this paper we focus on a single level splitting force scheme.

In theory, MTS methods can dramatically speed up MD simulations by reducing the number of expensive slow force evaluations. For example, in biomolecular applications, the computational complexity of the fast short-range force evaluations scales linearly in the number of atoms in the system, N , while it scales quadratically in N for the slow long-range force evaluations. Furthermore, while the short-range fast forces are easy to compute in parallel, long-range slow forces require global data communication and hence are more difficult to parallelize efficiently.

In practice, however, the impulse MTS methods such as the popular Verlet-I/r-RESPA suffer from severe resonance instabilities that limit practical performance gain [5, 9]. For solvated biomolecular systems, for example, the stability limit corresponds to the step-size ratio to be equal to ~ 4 [9].

Another problem is that common time-stepping methods do not exactly sample from the target temperature even if the simulations are stable and are subject to a thermostat [10, 11]. This error can be controlled with a loss of computational efficiency.

Recent research has partially resolved some of these issues. On the one hand, a group of algorithms which eliminate resonance induced instabilities by constraining the high frequencies [12–14] or by coupling the system to a bath [5, 8, 15–19] has been introduced. Though the reported step-size ratios for such schemes is sometimes an order of magnitude larger than for the original impulse MTS methods, other than resonance issues may still limit the performance of such methods. Thus, some of those algorithms use specific chemical information about the simulated systems and thus cannot be easily generalized. Other algorithms offer a control of the resonance problem at the price of loosing dynamical properties due to the non-Hamiltonian nature of the equations of motion used [15] or making the simulations prohibitively expensive [16]. Also, the ergodicity of the recent MTS algorithms is not always proven.

On the other hand, the incorrect sampling problem has been resolved by generalized hybrid Monte Carlo methods [1, 20–23] which rigorously sample from the canonical ensemble while only weakly interfering with the underlying molecular dynamics model. In addition, these methods offer the potentially faster exploration of conformational space compared to conventional MD simulations.

Here we propose to combine the MTS philosophy with the hybrid Monte Carlo (GHMC) methods [1, 20–23] in a way that the resulting MTS methods benefit from the delayed resonance barrier. This can be achieved by coupling a system to a Langevin thermostat naturally offered by GHMC and also by constraining the high frequencies using MOLLY [12, 13].

This leads us to two novel MTS methods, multi-time stepping generalized Hybrid Monte Carlo (MTS-GHMC) and multi-time stepping generalized shadow Hybrid Monte Carlo (MTS-GSHMC), for constant temperature molecular simulation of large biomolecular systems.

For the MTS-GHMC method, we demonstrate an improved stability compared with the best performed mollified MTS methods which use the similar MTS schemes. Moreover, the method allows for a larger step-size than that permitted by Langevin stabilized MTS methods [5]. However it will be shown that the acceptance rate in the molecular dynamics part of MTS-GHMC quickly deteriorates for very large step-size ratio and the effective efficiency gain (effective step-size ratio) is lower than the gain determined by step-size ratios alone.

A loss of performance due to decreasing the acceptance rate is not a new problem in molecular simulation. It was successfully solved in the past for different hybrid Monte Carlo

methods by Izaguirre and Hampton (SHMC) [23], Akhmatskaya and Reich (TSHMC) [24], Akhmatskaya and Reich (GSHMC) [1] by utilizing shadow / modified Hamiltonians [25].

Therefore, in order to improve the performance of MTS-GHMC we combine the ideas of generalized shadow hybrid Monte Carlo (GSHMC) [1] with the MTS-GHMC method and yield the new mollified multi-time-stepping method MTS-GSHMC.

A key contribution of this paper is the derivation of shadow Hamiltonians for the mollified MTS methods. This step is necessary since the shadow Hamiltonians used in the single time stepping GSHMC [1] are specific to the single time-step Störmer-Verlet method. To overcome this limitation we employ a highly efficient method for evaluating shadow Hamiltonians as proposed by Skeel and Hardy [25] for symplectic splitting methods [26]. However, due to the multi-scale nature of MTS methods, a non-trivial modification of the approach of Skeel and Hardy is required to make it applicable to the MTS-GSHMC method.

We demonstrate that the proposed methods, MTS-GHMC/GSHMC, are stable and thus do not suffer from resonance or non-resonance instabilities. They rigorously sample from the canonical ensemble and can accurately reproduce dynamical properties. In addition, the new methods offer longer time steps and increased computational robustness compared to Langevin stabilized MTS methods.

The paper is organized as follows. In Section 2 we formulate the basic molecular dynamics problem statement, outline key challenges for efficient simulations of biomolecular systems and provide the key ingredients of the proposed new MTS-GHMC/GSHMC methods. Section 3 is devoted to the necessary background material from [13] and [5] on mollified MTS and Langevin stabilized methods. The novel MTS-GHMC method is described in Section 4. In Section 5, we introduce the key theoretical ingredients of the proposed MTS-GSHMC method such as the concept of modified / shadow Hamiltonians and the homogeneous extension [25] of an MTS method. We then extend the general framework of Skeel and Hardy [25] to the efficient evaluation of shadow Hamiltonians for mollified MTS methods and present the algorithmic summary of the new MTS-GSHMC approach.

In Section 6 we give some implementation details; in Section 7 we present numerical results for two benchmark systems. We summarize our main conclusions in Section 8.

2. Problem statement and proposed methodology

Molecular dynamics (MD) requires the solution of Newton’s equations of motion for a classical unconstrained simulation

$$M \frac{d^2}{dt^2} X = -\nabla_X U(X), \quad (1)$$

where M is a diagonal mass matrix of atomic masses, X is the collective atomic position vector, U is the potential energy, typically given by

$$U = U^{bonded} + U^{nonbonded}, \quad (2)$$

$$U^{bonded} = U^{bond} + U^{angle} + U^{dihedral} + U^{improper}, \quad (3)$$

$$U^{nonbonded} = U^{Van\ der\ Waals} + U^{electrostatics}, \quad (4)$$

and the gradient vector $-\nabla_X U(X)$ is the force. Let N denote the number of atoms of the molecular model. Then the computational complexity of bonded interactions is proportional to N while it scales with N^2 for non-bonded interactions. Simple cut-off schemes have been devised to reduce the computational cost of non-bonded interactions. But it has also been found that cut-off schemes lead to poor approximations for highly charged systems such as biomolecular systems [27].

MD simulations are typically performed either under constant number of particles N , volume V and temperature T (NVT ensembles) or under constant number of particles N , pressure $\mathcal{P}$ and temperature T ($N\mathcal{P}T$ ensembles). Various techniques have been proposed to perform such ensemble simulations (see, e.g., [28]). In this report, we will focus on NVT ensemble simulations and will use Langevin dynamics to control temperature. Details will be given in Section 3.2.

Throughout the paper, we use the following conventions. We assume that collective atomic position vector X and the collective atomic momentum vector $P = M \frac{d}{dt} X$ are column vectors of length $3N$. We also introduce the state vector $\Gamma = (X^T, P^T)^T$, which is a column vector of length $6N$, and the collective atomic velocity vector $V = M^{-1}P = \frac{d}{dt}X$. Here Y^T denotes the transpose of a vector Y , i.e., Y^T is a row vector if Y is a column vector and vice versa.

Expectation values from an NVT ensemble are characterized by the canonical distribution

$$\rho_{canon} \propto \exp(-\beta H), \quad (5)$$

where $\beta = 1/k_B T$ is the inverse temperature and H is the total energy, which in turn is defined by

$$H(\Gamma) = \frac{1}{2} P^T M^{-1} P + U(X). \quad (6)$$

We introduce the following notation. Given a function f of the state variable Γ , we denote its expectation value with respect to ρ_{canon} by

$$\mathbb{E}[f] = \int f(\Gamma) \rho_{canon}(\Gamma) d\Gamma. \quad (7)$$

Under the assumption of ergodicity, the ensemble average (7) can be replaced by a time-average along trajectories from a MD simulation under the NVT ensemble and one obtains

$$E[f] = \lim_{t' \rightarrow \infty} \frac{1}{t'} \int_0^{t'} f(\Gamma(t)) dt, \quad (8)$$

which forms the base of computing expectation values from MD simulations. Typical quantities of interest include expectation values for thermodynamic and kinetic properties of the simulated systems [28].

The approximate computation of expectation values using (8) requires performing MD simulations over time intervals $[0, t']$ with t' as large as possible. The length of a MD simulation is, on the other hand, limited by the length of the time-step that can be used. MD simulations can be accelerated by the use of multiple-time-stepping (MTS) methods, such as the Verlet-I/r-RESPA method [6, 7], which is based on approximating slow forces as widely separated impulses. One derives impulse MTS methods by first rewriting (1) as

$$M \frac{d^2}{dt^2} X = F^{fast}(X) + F^{slow}(X), \quad (9)$$

where $F^{fast} = -\nabla_X U^{fast}$ and $F^{slow} = -\nabla_X U^{slow}$ subject to $U = U^{fast} + U^{slow}$. The partitioning of the potential U into a fast part U^{fast} and a slow part U^{slow} is done such that

- an appropriate outer time-step Δt for the slow part is significantly larger than an inner time-step δt for the fast part,
- evaluations of the fast force field F^{fast} are computationally much less expensive than evaluations of F^{slow} .

Given an integer $p > 1$ such that the outer time-step Δt and the inner time-step δt satisfy $\Delta t = p\delta t$, an impulse MTS method can now be stated as

$$M \frac{d^2}{dt^2} X(t) = \sum_{m=0}^L c_m \Delta t \delta_x(t - m\Delta t) F^{slow}(X(t)) + \sum_{n=0}^{pL} d_n \delta t \delta_x(t - n\delta t) F^{fast}(X(t)), \quad (10)$$

for $t \in [0, t' = L\Delta t]$, δ_x is the Dirac delta function, and $c_m = d_n = 1$ except when $m = n = 0$ or $m = L$, $n = pL$, respectively, in which case $c_m = d_n = 1/2$. Analytic solutions to (10) (i.e., numerical approximations to (1)) can be found by integrating (10) twice with respect to time. See *Appendix A* for more details.

However, the impulse MTS method (10) suffers from resonance instabilities. This limits large step-sizes possible. For solvated biomolecular systems, one, for example, has to restrict

the time step-ratio $p = \Delta t / \delta t$ to $p \leq 4$ with an inner time-step of $\delta t = 1$ fs [13]. These step-size limitations can be partially overcome by the mollified MTS methods of Garca-Archilla et al. [12] and Izaguirre et al. [5].

In the following section, we summarize the improved MTS methods of Izaguirre et al. [5, 13].

3. Mollified multiple-time-stepping methods

We first summarize the basic results of Izaguirre et al. [13] and then suggest an improved force field mollification based on the concept of soft constraints, as introduced by Zhou et al. [29].

3.1. Constant energy mollified MTS method

Let $r_{ij} \geq 0$ denote the distance between atom i and atom j . The switching function S serves to split non-bonded interactions into slow and fast parts. It is defined by

$$S(r_{ij}) = \begin{cases} 0 & \text{if } r_{ij} > r_{cutoff}, \\ 1 & \text{if } r_{ij} < r_{on}, \\ \frac{(r_{cutoff}^2 - r_{ij}^2)^2 (r_{cutoff}^2 + 2r_{ij}^2 - 3r_{on}^2)}{(r_{cutoff}^2 - r_{on}^2)^3} & \text{if } r_{on} \leq r_{ij} \leq r_{cutoff}, \end{cases} \quad (11)$$

for given cutoff radius r_{cutoff} and transition interval defined by $[r_{on}, r_{cutoff}]$ [30]. Given a potential energy U with entries defined by (2) – (4) we now define the “fast” potential by

$$U^{fast} = U^{bonded} + U^{nonbonded} S \quad (12)$$

and the slow contribution by

$$U^{slow} = U^{nonbonded} - U^{nonbonded} S. \quad (13)$$

We have to stress that such a splitting scheme is not unique, we use it here for simplicity. To retain the Hamiltonian, the switch must be applied to the energy rather than the forces. In commonly used MD software packages, switches often have C^1 continuity, which yields a C^0 vector field. In our study we use the scheme of C^2 continuity as implemented in ProtoMol. This guarantees the sufficient accuracy of calculated shadow Hamiltonians to provide high acceptance rates.

Mollified MTS methods are then defined by an operator

$$\bar{X} = A(X) \quad (14)$$

which assigns a filtered, averaged position $\bar{X}$ to an instantaneous collective atomic position vector X . The averaging operator is then applied to the slow potential to yield a mollified slow potential

$$U_{molly}^{slow}(X) = U^{slow}(A(X)). \quad (15)$$

The mollified impulse MTS method of Izaguirre et al., 1999 is now given by

$$M \frac{d^2}{dt^2} X(t) = \sum_{m=0}^L c_m \Delta t \delta_x(t - m\Delta t) F_{molly}^{slow}(X(t)) + \sum_{n=0}^{pL} d_n \delta t \delta_x(t - n\delta t) F^{fast}(X(t)), \quad (16)$$

for $t \in [0, t' = L\Delta t]$ with the slow forces defined by

$$F_{molly}^{slow}(X) = -\nabla U_{molly}^{slow}(X) = A_X(X)^T F^{slow}(A(X)), \quad (17)$$

where $A_X(X)$ denotes the Jacobian matrix of partial derivatives. The coefficients c_m and d_n take values as defined for (10).

One notes that (16) can be viewed as a standard impulse MTS method (10) applied to a “mollified” molecular system defined by the energy functional

$$H_{molly}(\Gamma) = \frac{1}{2} P^T M^{-1} P + U^{fast}(X) + U_{molly}^{slow}(X). \quad (18)$$

The assumption is, of course, that the difference between (6) and (18) is small. Nevertheless the difference between H and H_{molly} has to be kept in mind and we will come back to this issue when putting (16) into the context of generalized hybrid Monte Carlo (GHMC) methods [1, 20, 21].

Two particular instances of averaging operators A have been presented in [13]. While both operators can be used in the context of the newly proposed MTS-GHMC/GSHMC methods, we focus here on the *Equilibrium* method, which has been shown to lead to the most stable variant of (16) [5, 13]. Let us write the bond stretching and bond bending potentials as

$$U^{bond}(X) + U^{angle}(X) = \frac{1}{2} g(X)^T K g(X), \quad (19)$$

where K is a diagonal matrix of force constants and $g(X)$ is a vector of independent bond constraints (stretches and bends). Let us denote the Jacobian matrix of partial derivatives $g_X(X)$ by $G(X)$. The averaging operator $\bar{X} = A(X)$ for the method *Equilibrium* is defined by the implicit system of equations

$$\bar{X} = X - M^{-1} G(X)^T \mu \quad (20)$$

$$0 = g(\bar{X}) \quad (21)$$

in the vector of Lagrange multipliers μ . Details of the implementation of *Equilibrium* can be found in [13].

It has been demonstrated by Zhou et al. [29] that the averaged positions $\bar{X}$ are better characterized by soft constraints, i.e., (21) should be replaced by

$$0 = G(X)M^{-1}\nabla_X U^{fast}(X). \quad (22)$$

In other words, $\bar{X}$ is now defined as the minimizing state for the potential energy U^{fast} along $M^{-1}G(X)^T$.

3.2. Langevin stabilization

The stochastically stabilized and mollified MTS methods of [5] are based on the regular Langevin equations

$$dP = -\nabla_X U_{molly}(X)dt - \gamma Pdt + \sqrt{2\gamma k_B T} M^{1/2} dW(t), \quad dX = M^{-1} Pdt, \quad (23)$$

applied to the mollified potential energy

$$U_{molly} = U^{fast} + U_{molly}^{slow}. \quad (24)$$

Here γ is the collision frequency, $W(t)$ is a vector of independent standard Wiener processes, k_B is the Boltzmann constant, and T is the target temperature.

The collision frequency γ should be chosen sufficiently small to not alter dynamic properties of the molecular system. It has been demonstrated in [5] that $\gamma = 0.2 \text{ ps}^{-1}$ is suitable for simulations of solvated biomolecular systems. Further implementation details can be found in the original paper [5]. For our numerical experiments we use a particular implementation described in Section 4.2, which is closely related to the newly proposed MTS-GHMC method. The MTS-GHMC method is described in the next section.

4. Multiple-time-stepping generalized hybrid Monte Carlo method

The proposed simulation method takes the hybrid Monte Carlo (GHMC) method [20, 21] as a starting point. GHMC is a rigorous sampling tool for molecular systems. As such, GHMC can be used, for example, to compute free energies and to find molecular conformations with applications to drug design.

Throughout this paper we apply the “reduced-flipping” method upon rejection of trajectories [31]. This method decreases autocorrelation times compared to the standard momentum flip, while rigorously satisfying the detailed balance condition and minimizing the interference with the system’s dynamics.

We now demonstrate how to combine GHMC with a mollified MTS method. To do so, we assume that a splitting of the potential energy function U into a fast contribution (12) and a slow contribution (13) is given. We also assume that an averaging operator (14) has been defined, which implies a mollified slow potential (15) and a mollified total energy (18).

4.1. MTS-GHMC: Algorithmic summary

The generalized hybrid Monte Carlo (GHMC) algorithm of Horowitz [21] and Kennedy and Pendleton [20] for sampling from the canonical ensemble (5) is defined as the concatenation of two Markov Chain Monte Carlo (MCMC) steps: a molecular dynamics Monte Carlo (MDMC) and a partial momentum refreshment Monte Carlo (PMMC) step. The key novel step of the proposed MTS-GHMC method is the implementation of a mollified impulse MTS method in the MDMC step. We now summarize the proposed method.

The accepted states of the MTS-GHMC method are denoted by $\Omega_i = (\Gamma_i^T, t_i)^T$, $i = 0, \dots, I$, where $\Gamma_i = (X_i^T, P_i^T)^T$, X_i is a collective vector of atomic positions, P_i is a collective vector of atomic momenta, and t_i is time.

4.1.1. Partial momentum refreshment Monte Carlo (PMMC)

Given an accepted state $\Omega_i = (\Gamma_i^T, t_i)^T$, $\Gamma_i = (X_i^T, P_i^T)^T$, its collective atomic momentum vector P_i is now mixed with an independent and identically distributed normal (Gaussian) noise vector Ξ_i of dimension $3N$ and the partial momentum refreshment step is given by

$$\begin{aligned}\bar{P}_i &= \cos(\phi)P_i + \sin(\phi)\Xi_i, \\ \bar{\Xi}_i &= \cos(\phi)\Xi_i - \sin(\phi)P_i,\end{aligned}\tag{25}$$

where $0 < \phi \leq \pi/2$ is a given angle and $\Xi_i \sim \mathcal{N}[0, \beta M^{-1}]$. Here $\mathcal{N}[0, \beta M^{-1}]$ denotes the $(3N)$ -dimensional normal distribution with zero mean and covariance matrix βM^{-1} , M is the diagonal mass matrix of the molecular system, $\beta = 1/k_B T$ is the inverse temperature. Denote the new state vector by $\bar{\Omega}_i = (\bar{\Gamma}_i^T, t_i)^T$, $\bar{\Gamma}_i = (X_i^T, \bar{P}_i^T)^T$.

4.1.2. Molecular dynamics Monte Carlo (MDMC)

This step consists of the following two sub-steps.

- *Molecular dynamics* (MD). We apply the mollified impulse MTS method (16) to the current state $\bar{\Omega}_i$. As shown in *Appendix A*, (16) gives rise to a time-reversible and volume conserving mapping Ψ_τ , $\tau = L\Delta t$, and a proposal state is defined by

$$\hat{\Omega}_i = (\hat{\Gamma}_i^T, t_i + \tau)^T, \quad \hat{\Gamma}_i = \Psi_\tau(\bar{\Gamma}_i),\tag{26}$$

where $L > 0$ is a given integer.

- *Monte Carlo* (MC). The next accepted state Ω_{i+1} is found through a Metropolis accept / reject criterion

$$\Omega_{i+1} = \begin{cases} \hat{\Omega}_i & \text{with probability } P_A(\bar{\Gamma}_i, \hat{\Gamma}_i) \\ \bar{\Omega}_i & \text{with probability } P_S(\bar{\Gamma}_i | \bar{\Gamma}_{i-1}, \hat{\Gamma}_i) \\ \mathcal{F}\bar{\Omega}_i & \text{otherwise} \end{cases},\tag{27}$$

where

$$P_A(\bar{\Gamma}_i, \hat{\Gamma}_i) = \min(1, \exp(-\beta(H_{molly}(\hat{\Gamma}_i) - H_{molly}(\bar{\Gamma}_i)))), \quad (28)$$

is the probability of acceptance and H_{molly} is the Hamiltonian defined in (18). $P_S(\bar{\Gamma}_i | \bar{\Gamma}_{i-1}, \hat{\Gamma}_i)$ is the self-transition probability defined in [31], in other words, it is the probability of avoiding the momentum flip after a rejected trajectory. The operator $\mathcal{F}$ symbolizes the momentum flip.

4.2. Remarks

We wish to point out that the collective vector of atomic positions X_i as well as time t_i remain unchanged from Ω_i to Ω_{i+1} in case of rejection of the MDMC proposal step.

The free parameters of the MTS-GHMC scheme include the angle ϕ in (25), the inner and outer step-sizes, δt and Δt , respectively and the number of outer time-steps L . We will always assume that

$$\phi = \sqrt{2\gamma\tau} \ll 1. \quad (29)$$

Here $\gamma > 0$ is the collision frequency of an underlying Langevin model (23) and $\tau = \Delta t L$. Applying the condition (29) ensures that our method stays close to Langevin dynamics, which will guarantee a fair comparison with LM.

We note that (25) and (26) without a Metropolis accept / reject step (i.e., with all MD proposal steps being accepted) and with $L = 1$ yield a standard multiple time-stepping method for the underlying Langevin equations (23). The Metropolis test (27) turns (25) - (26) into a thermodynamically consistent implementation of stochastically stabilized and mollified MTS methods. See also [22] for a more detailed discussion on the relation between Langevin dynamics (23) and GHMC.

4.3. Data analysis

Let $\{\Omega_i\}_{i=1}^I$ denote a sequence of accepted states from a MTS-GHMC simulation with $\Omega_i = (\Gamma_i^T, t_i)^T$ and $\Gamma_i = (X_i^T, P_i^T)^T$. Expectation values of a function $f(\Gamma)$ with respect to the canonical density (5) are computed according to the formula

$$\langle f \rangle = \frac{\sum_{i=1}^I w_i f(\Gamma_i)}{\sum_{i=1}^I w_i} \quad (30)$$

with weight factors

$$w_i = \exp(-\beta(H(\Gamma_i) - H_{molly}(\Gamma_i))). \quad (31)$$

Provided the MTS-GHMC method generates an ergodic Markov chain for the given molecular system, we have

$$\lim_{I \rightarrow \infty} \langle f \rangle = \mathbb{E}[f] \quad (32)$$

independent of the specific parameter values of the implemented MTS-GHMC method [32].

5. Multiple-time-stepping generalized shadow hybrid Monte Carlo (MTS-GSHMC) method

It is a well-known problem of hybrid Monte Carlo (HMC) and generalized hybrid Monte Carlo (GHMC) methods that the acceptance rate of the molecular dynamics proposal step deteriorates with increasing time-steps and increasing system sizes [20]. One option to counteract this effect is to apply higher-order symplectic time-stepping methods. However this also increases the computational cost of HMC/GHMC. An alternative, less expensive approach has been proposed by Izaguirre and Hampton [23] for HMC and by Akhmatskaya and Reich [1, 24] for GHMC and is based on the concept of modified shadow Hamiltonians for symplectic time-stepping methods. The key observation is that symplectic time-stepping methods (see [26, 33] for an introduction to symplectic methods and modified equation analysis) conserve modified energy (shadow Hamiltonians) to much higher accuracy than the accuracy of the method itself. This suggests an implementation of HMC/GHMC in the framework of importance sampling with respect to such shadow Hamiltonians. The feasibility of the generalized shadow hybrid Monte Carlo (GSHMC) approach [1] for large biomolecular systems has been demonstrated by Wee et al. [2] and Escibano et al. [3, 4].

To be able to extend MTS-GHMC to MTS-GSHMC we need to establish easily computable shadow Hamiltonians for the mollified MTS methods of Izaguirre et al. [13]. In the remainder of this section, we describe the key construct for the shadow Hamiltonians for the mollified MTS methods and formulate the multiple-time stepping generalized shadow Hybrid Monte Carlo algorithm.

5.1. Shadow Hamiltonians

5.1.1. Skeel & Hardy construction of 4th and 8th order shadow Hamiltonians

Newton's equations of motion (1) can be put into the Hamiltonian dynamics framework by introducing the collective atomic momentum vector $P = M \frac{d}{dt}X$ and the Hamiltonian (energy)

$$H(X, P) = \frac{1}{2}P^T M^{-1}P + U(X). \quad (33)$$

The associated canonical Hamiltonian equations of motion are given by

$$\frac{d}{dt}X = \nabla_P H(X, P) = M^{-1}P, \quad \frac{d}{dt}P = -\nabla_X H(X, P) = -\nabla_X U(X). \quad (34)$$

We assume that there are N atoms and, hence, $X, P \in \mathbb{R}^{3N}$.

Following Skeel and Hardy [25], we define the homogeneous of degree two extension of a Hamiltonian $H(X, P)$ by

$$\tilde{H}(X, a, P, b) = a^2 H(a^{-1}X, a^{-1}P), \quad (35)$$

where $a \in \mathbb{R}$ and $b \in \mathbb{R}$ are two additional degrees of freedom. Setting $a = 1$ we obtain the simplified augmented system of equations of motion for the Hamiltonian defined in (33).

$$\frac{d}{dt}X = M^{-1}P, \quad \frac{d}{dt}P = -\nabla_X U(X), \quad \frac{d}{dt}b = X^T \nabla_X U(X) - 2U(X). \quad (36)$$

Let $Y(t) = (X(t)^T, 1, P(t)^T, b(t)^T)$ denote a solution of (36). A key result of Skeel and Hardy [25] states that

$$H(X(t), P(t)) = \frac{1}{2} \frac{d}{dt} Y(t)^T J Y(t), \quad (37)$$

which, using (36), is easy to verify. Here J is the skew-symmetric matrix

$$J = \begin{bmatrix} 0_{3N+1} & I_{3N+1} \\ -I_{3N+1} & 0_{3N+1} \end{bmatrix} \in \mathbb{R}^{(6N+2) \times (6N+2)} \quad (38)$$

and I_{3N+1} denotes the identity matrix of dimension $3N + 1$.

Another important observation is that the above arguments hold true for any Hamiltonian $H(X, P)$ and in particular for the modified Hamiltonians $H_{\Delta t}(X, P)$ of a symplectic time-stepping method, which defines a mapping

$$\Gamma_{n+1} = \Psi_{\Delta t}(\Gamma_n), \quad (39)$$

$\Gamma = (X^T, P^T)^T \in \mathbb{R}^{6N}$ [26, 33]. Modified Hamiltonians are defined such that they are conserved to an order in the step-size Δt which is higher than the order of the method itself. Recursive formulas for computing modified Hamiltonians are known; but are typically not viable for practical computations [26, 33]. The homogeneous augmented approach on the other hand allows for a relatively simple computation of modified Hamiltonians along numerical trajectories.

Before giving the details of the construction we outline the basic idea. Let $\hat{\Psi}_{\Delta t}$ denote the natural extension of a symplectic time-stepping method (39) to the extended system (36), i.e.,

$$Y_{n+1} = \hat{\Psi}_{\Delta t}(Y_n). \quad (40)$$

Let $\Pi(t)$ denote an interpolation polynomial of order $2k$, $k > 0$, through discrete solution points Y_n , $n = -k, \dots, 0, \dots, +k$. Then we define a modified energy value at $t = t_0$ by

$$H_{\Delta t}(X_0, P_0) = \frac{1}{2} \frac{d}{dt} \Pi(t)^T J \Pi(t), \quad (41)$$

with skew-symmetric matrix J defined by (38) and $Y_0 = (X_0^T, 1, P_0^T, 0)^T$. The accuracy of the modified Hamiltonian increases with the number of time-steps $2k > 0$. With $k = 2$ one can achieve fourth-order while $k = 4$ leads to an eighth-order modified Hamiltonian in Δt .

Assume that we wish to determine the value of the shadow Hamiltonian about $Y_i = (X_i^T, 1, P_i^T, b_i)^T$. Then $k = 2$ (fourth-order) or $k = 4$ (eighth-order) time-steps forward and backward in time are performed with (40) (unless those steps have already been performed as part of the simulation). Hence we have $2k + 1$ discrete approximations Y_j , $j = i - k, \dots, i, \dots, i + k$, centered about Y_i available and, for $k = 2$, we define

$$A_0 = Y_i, \quad (42)$$

$$A_1 = \frac{1}{2}(Y_{i+1} - Y_{i-1}), \quad (43)$$

$$A_2 = Y_{i+1} - 2Y_i + Y_{i-1}, \quad (44)$$

$$A_3 = \frac{1}{2}[Y_{i+2} - 2Y_{i+1} + 2Y_{i-1} - Y_{i-2}], \quad (45)$$

$$A_4 = Y_{i+2} - 4Y_{i+1} + 6Y_i - 4Y_{i-1} + Y_{i-2}. \quad (46)$$

as well as

$$A_{l,m} = \frac{1}{2\Delta t} A_l^T J A_m, \quad l, m = 0, \dots, 4, \quad (47)$$

with the skew-symmetric matrix J defined by (38). In case of $k = 2$ we need only to compute A_0, A_1, A_2 and $A_{l,m}$ for $m, l = 0, \dots, 2$.

The fourth-order shadow Hamiltonian at Y_i is now defined by

$$H_{\Delta t}^{[4]}(X_i, P_i) = A_{1,0} - \frac{1}{6}A_{1,2} \quad (48)$$

and the eighth-order shadow Hamiltonian by

$$H_{\Delta t}^{[8]}(X_i, P_i) = A_{1,0} - \frac{2}{7}A_{1,2} + \frac{5}{42}A_{3,0} + \frac{13}{105}A_{3,2} - \frac{19}{210}A_{1,4} - \frac{1}{140}A_{3,4}, \quad (49)$$

respectively.

Skeel and Hardy [25] use the homogeneous augmenting approach to develop shadow Hamiltonians for the Störmer-Verlet method and these shadow Hamiltonians have been implemented in the shadow hybrid Monte Carlo (SHMC) method of Izaguirre and Hampton [23].

5.1.2. Augmented mollified MTS method

In order to use the just described technique for constructing shadow Hamiltonians for mollified MTS methods, an appropriate extension of a symplectic time-stepping method has to be formulated.

As suggested in Section 5.1.1, we introduce an additional variable b and the extended mollified MTS method (16) can be expressed compactly as follows:

$$M \frac{d^2}{dt^2} X(t) = \sum_{m=0}^L c_m \Delta t \delta_x(t - m\Delta t) F_{molly}^{slow}(X(t)) + \sum_{n=0}^{pL} d_n \delta t \delta_x(t - n\delta t) F^{fast}(X(t)), \quad (50)$$

$$\begin{aligned} \frac{d}{dt} b(t) = & - \sum_{m=0}^L c_m \Delta t \delta_x(t - m\Delta t) [X(t)^T F_{molly}^{slow}(X(t)) + 2U_{molly}^{slow}(X(t))] \\ & - \sum_{n=0}^{pL} d_n \delta t \delta_x(t - n\delta t) [X(t)^T F^{fast}(X(t)) + 2U^{fast}(X(t))]. \end{aligned} \quad (51)$$

Note that (50) is identical to (16) and recall that (16) can be integrated exactly. The same applies to (51) once $X(t)$ has been determined from (50). Hence (50)-(51) define a one-step method in $Y = (X^T, 1, P^T, b)^T$, which we denote by

$$Y_{i+1} = \hat{\Psi}_{L\Delta t}(Y_i). \quad (52)$$

Finally note that (51) corresponds to the third equation in (36).

Once an appropriate extension of a symplectic time-stepping method has been formulated, Skeel and Hardy [25] provide a straightforward approach for evaluating higher-order shadow Hamiltonians (48)-(49).

5.1.3. Shadow Hamiltonians for the MTS-GSHMC method

A straightforward application of (48) or (49) to the MTS-GHMC method (see Section 4) led however, to an unsatisfactory behavior of the resulting shadow MTS-GSHMC method. While the acceptance rate in the MDMC part increased as expected, a large drift in kinetic energy was observed in the course of the Monte Carlo simulation. An explanation for this phenomenon is that the shadow Hamiltonian $H_{\Delta t}^{[4],MTS}$ is nearly conserved along numerical trajectories under the MDMC step but, at the same time, it is found that the difference between the original Hamiltonian H and the shadow Hamiltonian $H_{\Delta t}^{[4],MTS}$ becomes large as momenta P get refreshed under the PMMC part of the MTS-GSHMC method.

To overcome this large drift in the shadow Hamiltonian $H_{\Delta t}^{[4],MTS}$, we investigated an alternative derivation of shadow Hamiltonians from the standard BCH-formula for splitting methods [26, 33]. The associated fourth-order shadow Hamiltonian is given by

$$\begin{aligned} H_{\Delta t}^{[4]} = & H_{molly} - \frac{\delta t^2}{24} \left[(F^{fast})^T M^{-1} F^{fast} - 2P^T M^{-1} U_{XX}^{fast} M^{-1} P \right] \\ & - \frac{\Delta t^2}{24} \left[(F_{molly}^{slow} + 2F^{fast})^T M^{-1} F_{molly}^{slow} - 2P^T M^{-1} (U_{molly}^{slow})_{XX} M^{-1} P \right] + O(\delta t^4, \Delta t^4). \end{aligned} \quad (53)$$

Here U_{XX} denotes the Hessian matrix of second-order partial derivatives of a scalar-valued function U . We note that (53) is the sum of the original Hamiltonian plus the corrections from a Störmer-Verlet method applied to the fast and slow systems alone plus a mixed term

$-\frac{\Delta t^2}{12}(F^{fast})^T M^{-1} F_{molly}^{slow}$. Next we write (53) in the equivalent form (up to higher order terms in δt and Δt)

$$H_{\Delta t}^{[4]} = H_{\delta t}^{[4],SV,fast} + H_{\Delta t}^{[4],SV,slow} - \frac{1}{2}P^T M^{-1}P - \frac{\Delta t^2}{12}(F^{fast})^T M^{-1} F_{molly}^{slow}. \quad (54)$$

Here $H_{\delta t}^{[4],SV,fast}$ denotes a fourth-order shadow Hamiltonian for a Störmer-Verlet integration of the fast system with Hamiltonian $H^{fast} = P^T M^{-1}P/2 + U^{fast}(X)$ and step-size δt . Similarly, $H_{\Delta t}^{[4],SV,slow}$ denotes a fourth-order shadow Hamiltonian for a Störmer-Verlet integration of the slow system with Hamiltonian $H^{slow} = P^T M^{-1}P/2 + U_{molly}^{slow}(X)$ and step-size Δt . Explicit formulas for $H_{\delta t}^{[4],SV,fast}$ and $H_{\Delta t}^{[4],SV,slow}$ can be found in *Appendix B* and follow from the results stated in [25] for the standard Störmer-Verlet method. Alternatively, one could use the shadow Hamiltonians proposed in [1].

We further note that

$$H_{\Delta t}^{[4],SV,slow} - \frac{1}{2}P^T M^{-1}P - \frac{\Delta t^2}{12}(F^{fast})^T M^{-1} F_{molly}^{slow} = H_{\Delta t}^{[4],MTS} - H_{\Delta t}^{[4],MTS,fast} + O(\delta t^4, \Delta t^4), \quad (55)$$

where $H_{\Delta t}^{[4],MTS,fast}$ denotes the fourth-order shadow Hamiltonian (48) for the MTS method (16) with $F_{molly}^{slow} \equiv 0$. Hence, we finally propose the following fourth-order shadow Hamiltonian

$$H_{\Delta t}^{[4]} = H_{\Delta t}^{[4],MTS} - H_{\Delta t}^{[4],MTS,fast} + H_{\delta t}^{[4],SV,fast}. \quad (56)$$

We used this shadow Hamiltonian in the numerical experiments of Section 6. The key observation is that (56) contains approximations on both time-step levels (outer and inner) while (48) relies on the outer time-step Δt only.

We now generalize (56) to arbitrary order. Let $H_{\Delta t}^{[q],MTS}$ denote the q th order shadow Hamiltonian for a MTS method according to the construction of Skeel and Hardy [25]. Let $H_{\Delta t}^{[q],MTS,fast}$ denote the q th order shadow Hamiltonian for the same MTS method with F_{molly}^{slow} set equal to zero. Finally, let $H_{\delta t}^{[q],SV,fast}$ denote the q th order shadow Hamiltonian for the Störmer-Verlet method applied to the fast system with step-size δt . Then the shadow Hamiltonian for use in the newly proposed MTS-GSHMC method has a form:

$$H_{\Delta t}^{[q]} = H_{\Delta t}^{[q],MTS} - H_{\Delta t}^{[q],MTS,fast} + H_{\delta t}^{[q],SV,fast}. \quad (57)$$

5.2. Multiple-time-stepping generalized shadow hybrid Monte Carlo (MTS-GSHMC) algorithm

The proposed simulation method takes the hybrid Monte Carlo (GHMC) method [20, 21] and its extension to shadow Hamiltonians by Akhmatskaya and Reich [1] as a starting point.

As in the case of MTS-GHMC, we will use the reduced-flipping method in our implementation [31].

We now demonstrate how to combine GSHMC with a mollified MTS method. To do so, we assume that a splitting of the potential energy function U into a fast contribution U^{fast} and a slow contribution U^{slow} is given. We also assume that an averaging operator (14) has been defined, which implies a mollified slow potential (15). Finally we suppose that an extended mollified MTS method (52) as well as a shadow Hamiltonian (57) of order $q \geq 4$ have been implemented. For simplicity of notation, we denote this shadow Hamiltonian by $H_{\Delta t}$.

5.2.1. MTS-GSHMC: Algorithmic summary

The generalized shadow hybrid Monte Carlo (GSHMC) algorithm of [1] and [22] for sampling with respect to the shadow canonical distribution

$$\tilde{\rho}_{\text{canon}} \propto \exp(-\beta H_{\Delta t}) \quad (58)$$

is defined as the concatenation of two MCMC steps: a molecular dynamics Monte Carlo (MDMC) and a partial momentum refreshment Monte Carlo (PMMC) step. The key novel steps of the proposed MTS-GSHMC method are the implementation of an extended mollified impulse MTS method (50)-(51) and importance sampling with respect to a shadow Hamiltonian $H_{\Delta t}$. We now summarize the proposed MTS-GSHMC method.

The accepted states of the MTS-GSHMC method are denoted by $\Omega_i = (Y_i^T, t_i)^T$, $i = 0, \dots, I$, where $Y_i = (X_i^T, 1, P_i^T, b_i)^T$, X_i is a collective vector of atomic positions, P_i is a collective vector of atomic momenta, b_i is a scalar, and t_i is time. Note that the definition of Ω_i is different from the one used for the MTS-GHMC method (Section 4).

5.2.2. Partial momentum refreshment Monte Carlo (PMMC)

The PMMC step of MTS-GSHMC consists of a proposal step and a Metropolis accept / reject criterion.

- *Partial momentum refreshment* (PM). Given an accepted state $\Omega_i = (Y_i^T, t_i)^T$, $Y_i = (X_i^T, 1, P_i^T, b_i)^T$, its collective atomic momentum vector P_i is mixed with an independent and identically distributed normal (Gaussian) noise vector Ξ_i of dimension $3N$ and the partial momentum refreshment step is given by

$$\begin{aligned} \tilde{P}_i &= \cos(\phi)P_i + \sin(\phi)\Xi_i, \\ \tilde{\Xi}_i &= \cos(\phi)\Xi_i - \sin(\phi)P_i, \end{aligned} \quad (59)$$

where $0 < \phi \leq \pi/2$ is a given angle and $\Xi_i \sim N[0, \beta M^{-1}]$. Here $N[0, \beta M^{-1}]$ denotes the $(3N)$ -dimensional normal distribution with zero mean and covariance matrix βM^{-1} ,

M is the diagonal mass matrix of the molecular system, $\beta = 1/k_B T$ is the inverse temperature. Denote the proposed state vector by $\tilde{\Omega}_i = (\tilde{Y}_i^T, t_i)^T$, $\tilde{Y}_i = (X_i^T, 1, \tilde{P}_i^T, b_i)^T$.

- *Monte Carlo* (MC). The accepted state $\bar{\Omega}_i$ is found through a Metropolis accept/reject criterion:

$$\bar{\Omega}_i = \begin{cases} \tilde{\Omega}_i & \text{with probability } \min(1, \exp(-\beta \Delta H_{\Delta t}^e)) \\ \Omega_i & \text{otherwise} \end{cases}, \quad (60)$$

with

$$\Delta H_{\Delta t}^e := H_{\Delta t}(X_i, \tilde{P}_i) + \frac{1}{2} \tilde{\Xi}_i^T M^{-1} \tilde{\Xi}_i - H_{\Delta t}(X_i, P_i) - \frac{1}{2} \Xi_i^T M^{-1} \Xi_i. \quad (61)$$

5.2.3. Molecular dynamics Monte Carlo (MDMC)

This step consists of the following two sub-steps.

- *Molecular dynamics* (MD). We apply the extended mollified MTS method (52) to the current state $\bar{\Omega}_i = (\bar{Y}_i^T, t_i)^T$ with $\bar{Y}_i = (X_i^T, 1, \bar{P}_i^T, b_i)^T$. The proposal state is defined by

$$\hat{\Omega}_i = (\hat{Y}_i^T, t_i + L\Delta t)^T, \quad \text{with} \quad \hat{Y}_i = \hat{\Psi}_\tau(\bar{Y}_i) \quad (62)$$

for $\tau = \Delta t L$ and given integer $L \geq 1$. $\hat{\Psi}_\tau$ is defined as in (52).

- *Monte Carlo* (MC): The next accepted state Ω_{i+1} is found through a Metropolis accept/reject criterion

$$\Omega_{i+1} = \begin{cases} \hat{\Omega}_i & \text{with probability } P_A(\bar{Y}_i, \hat{Y}_i) \\ \bar{\Omega}_i & \text{with probability } P_S(\bar{Y}_i | \bar{Y}_{i-1}, \hat{Y}_i) \\ \mathcal{F}\bar{\Omega}_i & \text{otherwise} \end{cases}, \quad (63)$$

with acceptance probability

$$P_A(\bar{Y}_i, \hat{Y}_i) = \min(1, \exp(-\beta(H_{\Delta t}(\hat{Y}_i) - H_{\Delta t}(\bar{Y}_i)))), \quad (64)$$

and self-transition probability $P_S(\bar{Y}_i | \bar{Y}_{i-1}, \hat{Y}_i)$ as defined in [31]. The operator $\mathcal{F}$ represents the momentum flip.

5.2.4. Remarks

We wish to point out that the collective vector of atomic positions X_i as well as time t_i remain unchanged from Ω_i to Ω_{i+1} in case of rejection of the MDMC proposal step.

The free parameters of the MTS-GSHMC scheme include the angle ϕ in (59), the inner and outer step-sizes δt and Δt , respectively, as well as the number of outer time-steps L and the order of the shadow Hamiltonian. We will always assume that

$$\phi = \sqrt{2\gamma\tau} \ll 1. \quad (65)$$

Here $\gamma > 0$ is the collision frequency of an underlying Langevin model [22] and $\tau = \Delta t L$.

Note that the variable b is only required for computing shadow Hamiltonians. Furthermore, only differences in b appear in the formulas for shadow Hamiltonians and, hence, one can set $b = 0$ at the beginning of each MDMC and PMMC step.

5.2.5. Data analysis

Let $\{\Omega_i\}_{i=1}^I$ denote a sequence of accepted states from a MTS-GSHMC simulation with $\Omega_i = (Y_i^T, t_i)^T$ and $Y_i = (X_i^T, 1, P_i^T, b_i)^T$. Expectation values of a function $f(\Gamma)$, $\Gamma = (X^T, P^T)^T$, with respect to the canonical density

$$\rho_{\text{canon}} \propto \exp(-\beta H) \quad (66)$$

are computed according to the formula

$$\langle f \rangle = \frac{\sum_{i=1}^I w_i f(\Gamma_i)}{\sum_{i=1}^I w_i} \quad (67)$$

with weight factors

$$w_i = \exp(-\beta(H(X_i, P_i) - H_{\Delta t}(X_i, P_i))) \quad (68)$$

and state variables $\Gamma_i = (X_i^T, P_i^T)^T$. Provided the MTS-GSHMC method generates an ergodic Markov chain for the given molecular system, we have

$$\lim_{I \rightarrow \infty} \langle f \rangle = \mathbb{E}[f] \quad (69)$$

independent of the specific parameter values of the implemented MTS-GSHMC method [32].

The predictive performance analysis suggested for HMC-based methods [1, 23, 34] is applicable to the new methods introduced in this paper.

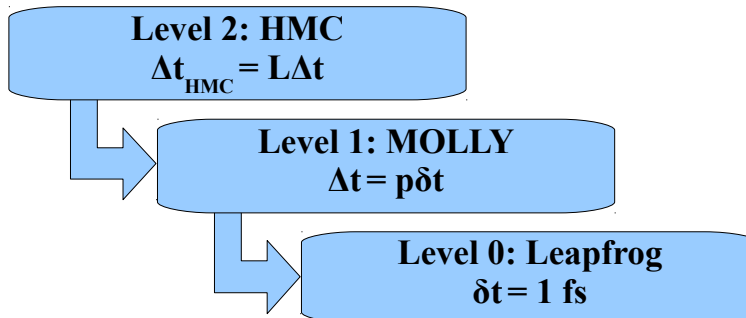


Figure 1: Example of a ProtoMol MTS-HMC simulation. Several integrator levels can be defined as nested loops and the forces can be split between these levels.

6. Numerical experiments: Methodology

We have implemented MTS-GHMC/GSHMC in the molecular dynamics software ProtoMol [30], version 2.1. We chose ProtoMol instead of other more popular optimized software packages for its flexibility in adopting new numerical schemes and algorithms, which saved much implementation time at the expense of losing some computational power. We prefer to use version 2.1 because the documentation available for this version is much more extensive compared with the latest version. ProtoMol 2.1 includes a hybrid Monte Carlo scheme and the procedure for calculation of shadow Hamiltonians (see for details [23]). However, it is worth saying at this point that the implementation strategy chosen by ProtoMol requires the additional integration of several molecular dynamics time steps, which in the case of macromolecular systems can add a significant amount of computational time. It is becoming even more crucial for the method discussed in this paper as three different shadow Hamiltonians have to be calculated per one Monte Carlo step. A more efficient implementation approach has been presented recently [3] which does not require integration of any additional steps during the MD part of GSHMC, although the momentum update does require additional short trajectories. Overall, in an efficient implementation of shadow Hamiltonians, the computational overhead can be negligible.

In ProtoMol every numerical method is called an *integrator*, regardless of whether it actually integrates differential equations, like a leapfrog scheme, or it samples velocities, like a Monte Carlo algorithm. The structure of ProtoMol allows the user to combine several integrator levels at runtime using MTS and organize them as nested loops. Figure 1 shows the example of an HMC simulation performing Metropolis test every L steps. It is combined with a mollification for slow forces every p steps and a leapfrog integrator scheme every single step.

For the implementation of MTS-GHMC, we used the HMC integrator already available in ProtoMol and replaced the velocity resetting between MC steps with a partial velocity update. The only additional modifications required were the introduction of a new parameter for the angle of the velocity update ($0 < \phi \leq \pi/2$) and a new weight term in the output files for re-weighting observables to account for the mollification of the positions.

For the implementation of MTS-GSHMC we used the SHMC integrator in ProtoMol as a template. We introduced a similar partial velocity update as in MTS-GHMC, in this case followed by a Metropolis test. However, the method for obtaining shadow Hamiltonians in SHMC is only valid for single-time-stepping (STS) integrators, so we included a new method that uses the three-step scheme:

$$H_{\Delta t}^{[q]} = H_{\Delta t}^{[q],MTS} - H_{\Delta t}^{[q],MTS,fast} + H_{\delta t}^{[q],SV,fast}, \quad (70)$$

as discussed in Section 5.1.

The first term is calculated employing the method already implemented in ProtoMol using backward differences [25] but for full MTS cycles with a step-size $\Delta t = p\delta t$, $p > 1$. The second term is calculated in the same way but without slow forces ($F^{slow}=0$). The third term is calculated by using only the STS integrator (Störmer-Verlet) with step-size δt . We also added the weight term to the output for re-weighting observables. In this case it is necessary to account for the mollification of the positions and the approximation made by the shadow Hamiltonians.

7. Numerical results

7.1. Description of model systems

A very basic implementation of these methods was previously tested on a one-dimensional chain of diatomic molecules interacting through Lennard-Jones potentials [4, 35, 36]. Here we present results from our ProtoMol implementation.

We consider two different systems to test the numerical accuracy of the new integrators MTS-GHMC and MTS-GSHMC. First we consider a system of 423 flexible TIP3P [37] water molecules simulated using periodic boundary conditions. As a second system we consider a small solvated protein such as BPTI comprising 1101 atoms. We used the CHARMM22 force field for proteins [38]. We will refer to this system as protein further through the paper. Similar test systems have been widely used to test MTS algorithms (see, for example, [39]). The objective was to compare our results with a Langevin stabilization method, Langevin Molly (LM), which demonstrated superior performance over conventional MTS methods [5] and which is also available in ProtoMol.

All simulations are performed in constant NVT ensemble at $T=300$ K. The potential energy includes bonded and unbonded terms. Bonded terms are bond and angle interactions

in the water system, as well as dihedral and improper interactions in the protein system. Non-bonded terms include Lennard-Jones and Coulomb potentials. The force splitting is done in accordance with (12) and (13), using the switching function (11). As a mollification scheme we chose the *Equilibrium* method for its proven stability [5, 13], implemented in ProtoMol 2.1.

7.2. Accuracy of the new methods for single time stepping integration

We first check that the new methods produce accurate results in STS simulations. We produce reference values with an STS molecular dynamics simulation using the standard Störmer-Verlet leapfrog integrator for 1 ns, $\delta t=1$ fs and $T=300$ K. Also for reference, we ran an STS Langevin dynamics simulation using a Langevin impulse integrator [40] with the same parameters and $\gamma=0.2$ ps⁻¹. From here on we will refer to these reference simulations as MD and LD respectively. We performed similar STS simulations using HMC, SHMC, GHMC and GSHMC, with no force splitting at the moment. HMC and SHMC results correspond to the standard implementations available in ProtoMol, while GHMC and GSHMC results were produced with the new implementation. For each of these methods, we tested different values for the trajectory length ($L=1, 10, 100, 500, 1000$) and for the step-size ($\delta t=0.25$ fs, 0.5 fs, 1 fs). As we can see from Table 1 the new methods produce results for the average potential and total energies in good agreement with the reference values within the statistical error.

It is worth clarifying that the error bars presented in Table 1 represent the standard deviation taking into account the effect of the correlation between samples in MCMC simulations. The estimation of accurate error bars in Markov chain Monte Carlo simulations is not a trivial matter as a consequence of the potentially high correlation between samples. A full discussion on the topic can be found elsewhere [41]. In our results the correlation was removed by using a binning analysis in which we first calculate averages over ‘bins’ of 5 consecutive samples and then average over the bins, which are less correlated than the initial samples [42, 43].

These results were used for choosing the optimal values of inner step-size δt and trajectory length L for the subsequent MTS simulations. Based on the acceptance rate observed, we found for the new methods the optimal values to be $\delta t=1$ fs and $L=10$. This choice satisfies $0 < \phi \leq \pi/2$ and keeps $\phi \ll 1$, considering $\phi = \sqrt{2\gamma L \Delta t}$. Note that for SHMC the results shown correspond to a step-size $\delta t=0.5$ fs. For $\delta t=1.0$ fs the integrator was trapped in the momentum update stage. As a consequence, SHMC calculations with $\delta t=1.0$ fs require long computation times and provide poor statistics.

7.3. Efficiency of the new MTS methods in reproducing dynamical properties

Though Hybrid Monte Carlo cannot reproduce in full kinetic properties of simulated systems, the generalized hybrid Monte Carlo (GHMC) was designed specifically for keeping more

Table 1: Comparison of STS methods for the water and protein systems. Potential energy U and total energy E given in kJ/mol. Acceptance rate Acc given as a percentage over total simulated time of 1 ns, with step-size $\delta t=1$ fs, $\gamma=0.2$ ps $^{-1}$ and trajectory length $L=10$. With $\delta t=1$ fs the SHMC integrator was trapped in the momentum update stage and we were not able to collect the statistics. For sake of comparison, the results obtained with $\delta t=0.5$ fs are shown.

System	Method	Acc	$U(kJ/mol)$	$E(kJ/mol)$
Water	MD	–	-1191 $\pm$ 12	-828.8 $\pm$ 0.5
	LD	–	-1185 $\pm$ 22	-807 $\pm$ 27
	HMC	89%	-1182 $\pm$ 20	-804 $\pm$ 25
	GHMC	90%	-1179 $\pm$ 22	-803 $\pm$ 28
	SHMC*	100%	-1193 $\pm$ 19	-817 $\pm$ 23
	GSHMC	100%	-1191 $\pm$ 20	-816 $\pm$ 25
Protein	MD	–	-1891 $\pm$ 20	-871.1 $\pm$ 0.8
	LD	–	-1910 $\pm$ 33	-926 $\pm$ 41
	HMC	61%	-1894 $\pm$ 27	-912 $\pm$ 36
	GHMC	76%	-1901 $\pm$ 31	-915 $\pm$ 40
	SHMC*	100%	-1879 $\pm$ 24	-898 $\pm$ 33
	GSHMC	99%	-1898 $\pm$ 37	-900 $\pm$ 22

*Results obtained with step-size $\delta t=0.5$ fs.

dynamical information in simulations. Momentum flips introduced in GHMC to satisfy the detailed balanced conditions may potentially damage the simulated dynamics. To overcome this problem, either a number of rejected proposals has to be minimized [22] or a frequency of momentum flips should be reduced [31] or both ideas can be implemented at once. In this work we apply the both approaches to MTS-GSHMC and use the reduced flip technique [31] in MTS-GHMC. The resulting methods still rigorously sample canonical distribution and satisfy the detailed balance condition.

To guarantee the fair comparison of dynamical properties achieved by three tested methods, the range of parameters providing the accurate dynamics in the generalized hybrid Monte Carlo methods has to be identified. Once it is done (see subsection 7.2), our main objective is to find the maximum effective step-size for each of the new methods that still produces good dynamics. We define effective step-size as $\Delta t_{eff} = p\delta t Acc$, where Acc is the acceptance rate for proposed trajectories in MTS-GHMC/MTS-GSHMC and p is the number of steps between an integration of the slow forces. This gives an intuitive idea about the real performance improvement and provides the fair comparison with the non-HMC methods. We test the accuracy of the dynamical results by calculating the self-diffusion coefficient of water molecules $D(\gamma)$ in both tested systems for small γ and comparing it with results from a standard MD simulation and with mollified Langevin stabilization (LM). We calculate the self-diffusion coefficient D from the definition $D = R(\tau)/6\tau$, where $R = \langle |r(\tau) - r(0)|^2 \rangle$ is

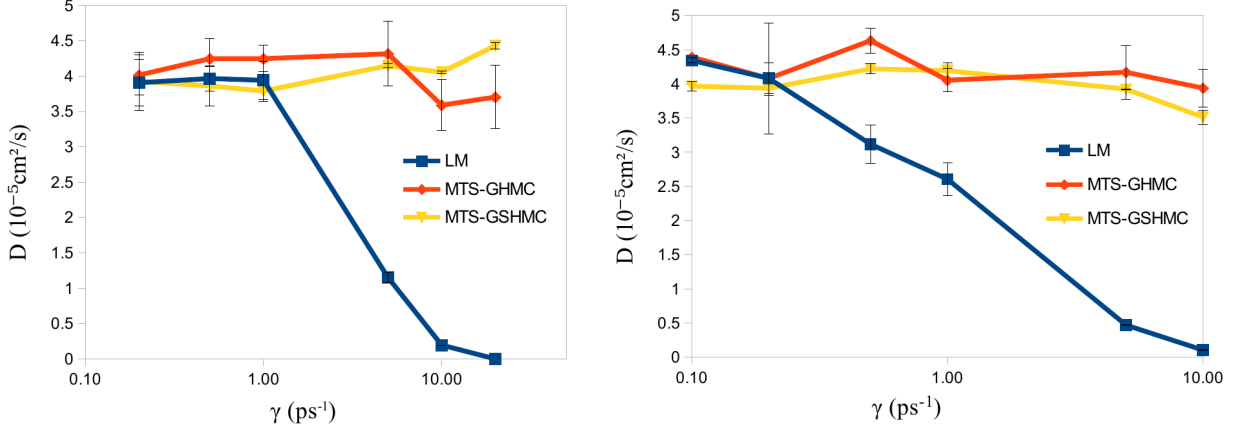


Figure 2: Variation of the water self-diffusion coefficient for the different MTS methods with increasing γ for the water system (left) and protein system (right).

the mean square displacement averaged over the particles in the system and τ is the total length of the simulation [44].

In all the following results we used a fixed trajectory length $L=10$, a simulation time of 1 ns, cubic periodic boundary conditions and the potential switching function defined by (11). The same switching function is used for splitting the Coulomb potential between fast and slow and for bringing the Lennard-Jones potential to zero smoothly. It is important to use a C^2 function (with continuous second order derivative) in order to avoid energy drift [13], which would affect the accuracy of shadow Hamiltonians and hence the acceptance rate.

The strategy to follow begins from calculating a reference value for D using a 1 ns MD (leapfrog) simulation. We then run similar simulations with MTS-GHMC, MTS-GSHMC and LM, gradually increasing γ until the value of D diverges by 3% or more of the reference value, and find in this way γ_{max} . We then gradually increase Δt until D diverges again, finding Δt_{max} for each integrator. The process is repeated for the water and protein systems.

It is worth noting that for proper calculation of diffusion coefficients it is necessary to perform much longer simulations and with larger systems. The values calculated here do not exactly correspond to experimental measurements, but we can still use them to benchmark the new methods by comparing them to a reference value calculated with standard MD, as it was suggested in [5].

In Figure 2 we show the results for the self-diffusion coefficient D calculated for the two benchmark systems, water and protein, using the three tested methods with increasing γ . All simulations ran for 1 ns with $\delta t=1$ fs and $\Delta t=1$, so with no effective force separation

Table 2: Maximum values of Δt that produce good dynamics measurements for the water system. Step-size $\delta t=1$ fs, trajectory length $L=10$ and a total simulated time of 1 ns. Δt (fs) outer step-size, D diffusion coefficient, Acc acceptance rate, Δt_{eff} (fs) effective step-size.

$\gamma(\text{ps}^{-1})$	LM		MTS-GHMC			MTS-GSHMC		
	Δt	ΔD	Acc	Δt_{eff}	ΔD	Acc	Δt_{eff}	ΔD
0.0	1	–	88%	0.88	–	97%	0.97	–
0.2	16	3%	69%	17.25	2%	83%	23.12	1%
0.5	12	3%	70%	12.60	1%	77%	17.05	3%
1.0	12	1%	71%	12.72	2%	78%	14.11	1%

Table 3: Maximum step-sizes Δt that produce good dynamics measurements for the protein system. Step-size $\delta t=1$ fs, trajectory length $L=10$ and a total simulated time of 1 ns. Δt (fs) outer step-size, D diffusion coefficient, Acc acceptance rate, Δt_{eff} (fs) effective step-size.

$\gamma(\text{ps}^{-1})$	LM		MTS-GHMC			MTS-GSHMC		
	Δt	ΔD	Acc	Δt_{eff}	ΔD	Acc	Δt_{eff}	ΔD
0.0	1	–	76%	0.76	–	99%	0.99	–
0.2	12	1%	67%	13.30	3%	88%	19.36	3%
0.5	8	3%	69%	11.01	2%	96%	15.39	2%
1.0	2	38%	67%	8.01	2%	95%	15.25	3%

yet. In the case of LM, a slight increase of the collision frequency γ produced a severe loss of accuracy, while both MTS-GHMC/GSHMC obtained reasonable values of D for all $\gamma \leq 20$.

We now choose the γ_{max} for each system that shows good dynamics results and gradually increase the cycle-length for the slow forces, Δt . In this way we find the maximum step-size for each method that does not corrupt the dynamics of the system.

Tables 2 and 3 show a summary of the largest step-sizes achieved for each method. For MTS-GHMC and MTS-GSHMC larger step-sizes can be used without affecting the accuracy of the diffusion coefficient, which does not diverge more than 3% from the reference value measured with an STS simulation. In particular, for $\gamma=0.2 \text{ ps}^{-1}$ MTS-GSHMC obtains reliable results using an effective step-size approximately 50% longer than in the best LM simulation, an improvement that becomes even more evident as we increase γ .

7.4. Efficiency of the new methods in reproducing thermodynamical properties

The objective here is to check the accuracy of the new methods by calculating average potential energies while using $\gamma=0.2 \text{ ps}^{-1}$, which allows the comparison with LM. We use a similar approach and same conditions as in the previous section, keeping $L = 10$ to satisfy

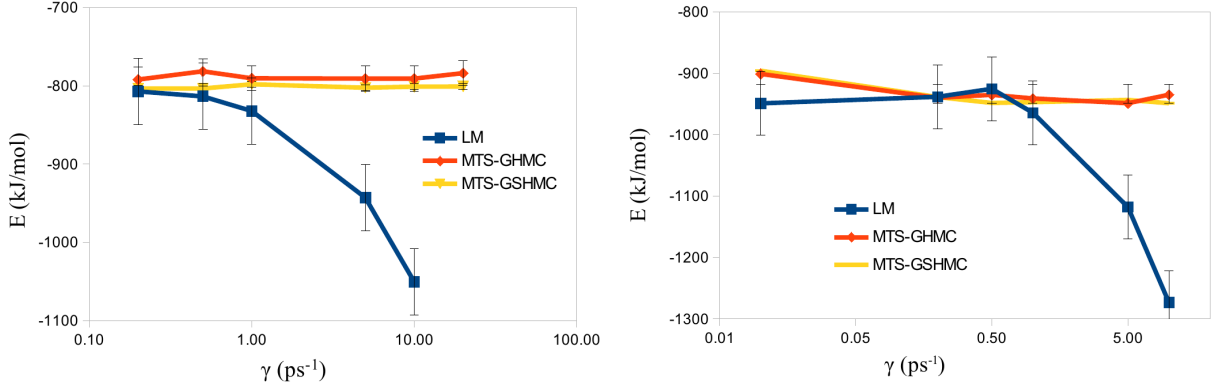


Figure 3: Total energy values as a function of γ for the water system (left) and protein system (right).

Table 4: Maximum step-sizes Δt (fs) that produce good energy measurements. $\gamma=0.2 \text{ ps}^{-1}$ for both water and protein systems. Step-size $\delta t=1 \text{ fs}$, trajectory length $L=10$ and a total simulated time of 1 ns.

System	LM		MTS-GHMC			MTS-GSHMC		
	Δt	ΔE	Acc	Δt_{eff}	ΔE	Acc	Δt_{eff}	ΔE
Water	8	0.1%	72%	11.59	0.1%	86%	18.85	0.4%
	10	0.5%	71%	14.21	0.4%	85%	20.32	0.9%
	12	0.6%	71%	15.57	0.9%	80%	22.50	0.6%
Protein	8	0.2%	68%	10.91	0.1%	91%	18.15	0.5%
	10	0.5%	68%	12.27	0.1%	84%	18.37	0.4%
	12	1.0%	67%	13.32	1.0%	89%	22.26	0.3%

that $\phi \leq \pi/2$, considering $\phi = \sqrt{2\gamma L \Delta t}$.

Our criterion for considering an accurate result is that the divergence from the reference value is not bigger than 1%, so that:

$$\frac{\langle E \rangle - E_{ref}}{E_{ref}} < 0.01,$$

where E_{ref} is the averaged total energy calculated in a similar simulation using the STS method. The procedure we follow is to find the maximum Δt for the three MTS methods that produces an acceptable error of $\Delta E < 0.01 E_{ref}$.

MTS-GHMC proves to be more efficient than LM, achieving longer effective step sizes for the water and protein systems, with a maximum of 15 fs and 13 fs respectively. MTS-GSHMC performs even better, reaching an effective step-size of over 22 fs for both systems. In

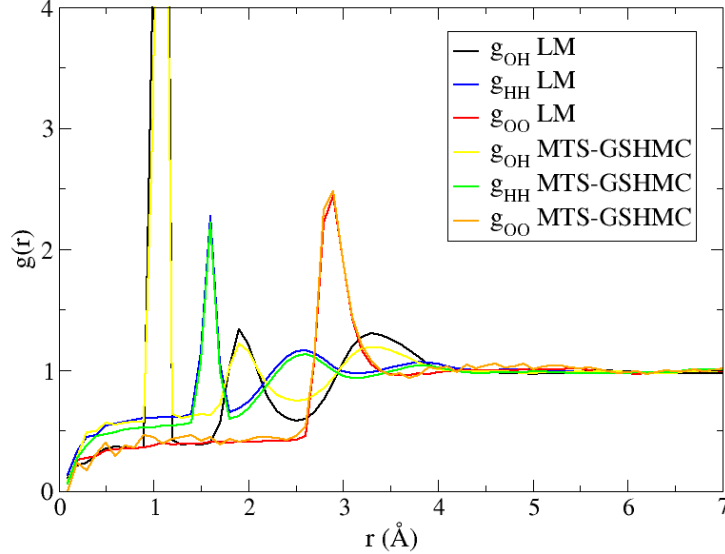


Figure 4: Radial distribution function comparison for the water system with $\gamma=0.2 \text{ ps}^{-1}$.

Table 4 we show the biggest achievable step-sizes for each method considering only accuracy limitations.

Even larger step-sizes can be achieved in all three methods by further increasing Δt at the expense of losing numerical precision. We found it is possible to run LM simulations with $\gamma=1 \text{ ps}^{-1}$ and $\Delta t=30 \text{ fs}$, although the average energies measured with this step-size are hardly reliable. An increase in the dissipation coefficient γ further damages the accuracy of energy measurements in LM (see Figure 3). On the other hand, MTS-GHMC/GSHMC methods could achieve larger step-sizes, but the subsequent decrease of the acceptance rate and the poor efficiency of shadow Hamiltonian calculation in ProtoMol makes them impractical, even if the simulation is still stable. We can conclude that the maximum achievable step-size for slow forces in both MTS-GHMC and MTS-GSHMC is only determined by accuracy considerations.

As an additional comparison and to further check the accuracy of our simulations, Figure 4 shows radial distribution functions extracted from the simulations in Table 2 for $\gamma=0.2 \text{ ps}^{-1}$. We can see that both LM and MTS-GSHMC reproduce the well known radial distribution function for liquid water [45].

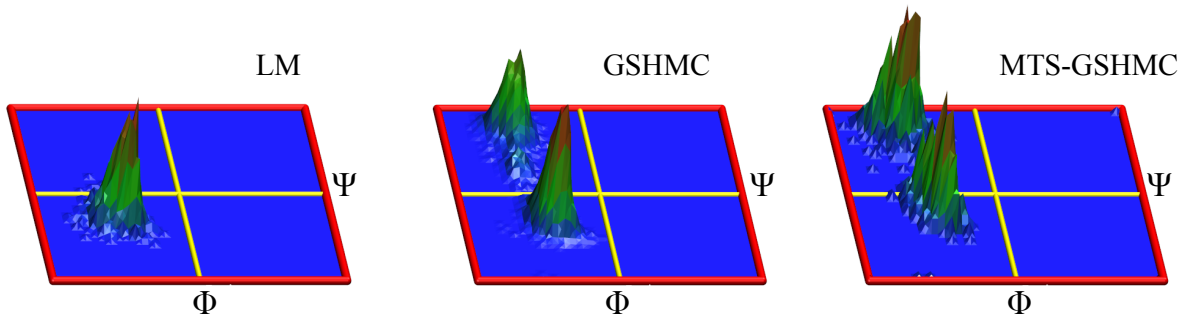


Figure 5: Ramachandran plot for the Φ/Ψ angles of the Gly⁵⁶ residue of BPTI.

7.5. Sampling efficiency

We can compare the sampling efficiency of MTS-GSHMC and LM by counting the number of different geometric conformations visited during a fixed simulated time, as suggested in [46]. The idea is to divide the energy profiles of each dihedral angle into wells, with divisions between wells placed at the maxima of the dihedral potential energy. At each integration time step we create a conformation string based on which well each dihedral occupies. At the end of the simulation we count how many different conformation strings have been produced. We applied this method to a 1 ns simulation of the protein system with step-size for slow forces $\Delta t=16$ fs, inner step-size $\delta t=1$ fs and $\gamma=0.2$ ps⁻¹. We obtained 665 different conformations with LM and 1147 with MTS-GSHMC. This great improvement in sampling efficiency is also observed in STS simulations using GSHMC, which finds 741 different conformations. This improvement is due to the beneficial combination of partial momentum updates and the use of shadow Hamiltonians in Metropolis tests, which improves the sampling by helping the MD discern which trajectories are worth following. As an example of the benefits of better sampling, in Figure 5 we show the Ramachandran plot for the Φ/Ψ angles of the Gly⁵⁶ residue of BPTI, where we can see that MTS-GSHMC and GSHMC both discover a second stable conformation not found by LM. The case of this residue is specially significant because LM does not find that second conformation even when running longer simulated times. Ramachandran plots of other residues are not as significant, but in all cases we observe broader histograms, meaning that we are exploring a larger range of the Φ/Ψ space.

7.6. Parameters

The two new methods introduce several new parameters that deserve special attention, namely, the angle of the partial momentum update ϕ , the MD trajectory length L and the order of the shadow Hamiltonian k (only for MTS-GSHMC).

As we discussed in Section 4.2, the angle ϕ can be related to the collision frequency γ in a Langevin model through the expression:

$$\phi = \sqrt{2\gamma\Delta t L},$$

and can take values between 0 and $\pi/2$. In order to not interfere with the dynamics of the system, γ –and hence ϕ – must be kept as small as possible, although bigger γ enables longer step-sizes for the slow forces. We found that the optimal values for γ are between 0.2 and 1 ps⁻¹.

The length of the MD trajectory L has a strong impact on the sampling efficiency of the methods and in their computational cost. A short L means that the Metropolis test and the partial momentum update have to be performed too often and the computational overhead becomes important. On the other hand, a very long trajectory length implies that the benefits in sampling efficiency are diminished, as the simulation becomes closer to standard MD. Furthermore, if the acceptance rate is not close to 100%, much computational time is wasted integrating trajectories that will be discarded. In consequence, finding an optimal value for L usually requires a trial and error approach. Most commonly the values will be between 10 and 1000 steps, but one must remember that the value of ϕ is influenced by L and ϕ can not be bigger than $\pi/2$.

In the case of MTS-GSHMC we have an additional new parameter, the order of the shadow Hamiltonian k . In *Appendix B* we detail the expression for a fourth-order shadow Hamiltonian of Skeel and Hardy [25] for the Störmer-Verlet method. In most cases fourth order is enough to achieve the desired accuracy. Higher orders might provide better acceptance rates, although the additional computation required for their calculation is rarely justified, and not compensated by the number of additional accepted trajectories.

7.7. Performance

The straightforward way to measure the performance of the new methods would normally be to compare computational times with respect to the corresponding STS method. For example, MTS-GSHMC with an outer step-size of $\Delta t=20$ fs can save approximately 60% of the computational time compared to a similar STS simulation. However, one would expect a much higher gain considering that we are avoiding 19 of every 20 slow-force integrations. This is a negative consequence of ProtoMol’s modular design and in particular its implementation of Shadow Hamiltonians.

In order to measure the real performance gain of the new methods when compared with LM we must take into consideration not only the maximum achievable step-size of each method, but also the acceptance rates, as bigger step-sizes often come at the expense of more rejected trajectories. However, ultimately what limits the performance gain is the accuracy of the measurements obtained at those step-sizes. In Table 5 we show the maximum effective

Table 5: Maximum effective step-size Δt_{eff}^{max} in fs for the three tested methods in terms of accuracy limitations for both simulated systems with $\gamma=0.2 \text{ ps}^{-1}$. D ($10^{-5} \text{ cm}^2/\text{s}$) water self-diffusion, E (kJ/mol) total energy.

System	Method	Kinetic	Thermodyn.	All		Relative gain	
		Δt_{eff}^{max}	Δt_{eff}^{max}	Δt_{eff}^{max}	D		E
Water	LM	16	12	12	4.3±0.3	-804±24	1.00
	MTS-GHMC	17.25	15.57	15.57	4.1±0.4	-786±27	1.30
	MTS-GSHMC	23.12	22.50	22.50	4.1±0.3	-793±26	1.88
Protein	LM	12	10	10	4.1±0.5	-927±39	1.00
	MTS-GHMC	13.30	13.32	13.30	4.2±0.2	-938±36	1.33
	MTS-GSHMC	19.36	22.26	19.36	4.1±0.2	-946±30	1.94

step-sizes for all three methods on both simulated systems taking into account the accuracy limitations when calculating kinetic and thermodynamic properties. For the water system we found that $\Delta t=12$ fs was the biggest step-size that produces good measurements for diffusion coefficients and energy averages using LM. Using MTS-GHMC, we can increase the effective step-size to 15.57 fs, and in the case of MTS-GSHMC, the performance gain is even higher, reaching a step-size of 22.50 fs. For the protein system we found a maximum of $\Delta t_{eff}=13.30$ fs for MTS-GHMC and $\Delta t_{eff}=19.36$ fs for MTS-GSHMC while still producing accurate values, compared with 10 fs for LM. However, any larger step-sizes did not compensate the decrease in acceptance rate.

We can measure the relative gain in performance by comparing the maximum step-size that reproduces both good dynamics and thermodynamical properties in each system. As we can see in Table 5, MTS-GSHMC obtains a relative performance gain of 1.88 over LM in the water system and 1.94 in the protein system. There is clearly an important advantage in using MTS-GSHMC, although MTS-GHMC also beats LM by factors of 1.30 and 1.33, while still avoiding the added complexity of using shadow Hamiltonians.

8. Conclusions

We have introduced two new methods for molecular simulations, MTS-GHMC and MTS-GSHMC. Both methods use force splitting techniques for larger step-sizes and force field mollification combined with the naturally possessed weak stochastic stabilization for avoiding numerical resonances. First results show that these methods can be used for obtaining reliable dynamical information as well as accurate sampling data, improving the performance and efficiency of their STS counterparts, GHMC and GSHMC. We have compared our new methods with a mollified Langevin dynamics implementation (LM) known as an efficient MTS technique, and found that the HMC methods outperform LM in terms of accuracy and

sampling efficiency, allowing for larger step-size in calculating slow forces.

Both MTS-GHMC and MTS-GSHMC produce results comparable in accuracy, improving that of LM for $\gamma > 1 \text{ ps}^{-1}$. Both methods converge rigorously to the NVT ensemble, conserving temperature without requiring any additional thermostats. The generalized partial momentum update ensures temperature conservation without interfering with the accuracy of dynamics for all tested collision frequencies. With the proper choice of parameters, both methods produce accurate measurements of kinetic and thermodynamic data. Values for the self-diffusion coefficient of water molecules remain accurate when using dissipation coefficients ($\gamma > 1 \text{ ps}^{-1}$) that spoil LM simulations. The outer step-size of MTS-GHMC can be made at least as long as that of LM for the same γ , while in MTS-GSHMC it can be increased by almost a factor of two for both the water and protein systems.

However, an increase of step-sizes leads to an increase in Metropolis rejection rates. Hence, the effective sampling efficiency gain is lower than the gain determined by step-size ratios alone. A better measure for efficiency is given by the effective outer step-size $\Delta t_{eff} = p \delta t Acc$, where Acc is the acceptance rate for proposed trajectories and p the number of steps between slow force integrations. Using this concept we have compared the maximum effective step-sizes obtained by each method in terms of accuracy limitations. Throughout our simulations, we have found that MTS-GHMC consistently outperforms LM with an effective gain between 10% and 20%. But MTS-GSHMC has shown that it can easily improve those results with an effective step-size that is typically 50% longer than that of LM, while not losing accuracy in the simulation results.

We have also measured the increase in sampling performance by monitoring the number of different dihedral conformations explored by MTS-GSHMC compared to LM. As expected, the partial momentum update in combination with the use of shadow Hamiltonians lead to the better sampling efficiency compared with LM. We found that some dihedral conformations are completely missed by LM even in relatively long simulations ($> 1 \text{ ns}$). Indeed, with the optimal choice of parameters, the number of different conformations per simulated time is twice bigger in MTS-GSHMC than in LM simulations.

Throughout this paper we have used the reduced-flipping method [31] upon rejection of proposal trajectories. This method ensures rigorous sampling in the canonical ensemble by fulfilling the detailed balance condition. At the same time we found that this approach only slightly interferes with the dynamics of the systems. Other ways of treating the momentum update are discussed in [22, 47].

We found that a straightforward application of the shadow Hamiltonian formulations of Skeel and Hardy [25] to MTS methods leads to a large drift in energy values along MC samples. An improved formulation (57) was suggested and successfully implemented within the molecular dynamics software ProtoMol. The use of shadow Hamiltonians in the Metropolis tests results in larger step-sizes for slow force integration, although the algorithm used in

ProtoMol for the calculation of these Hamiltonians limits the computational gain. A more efficient implementation, similar to the one suggested in [3], will be necessary to properly measure the potential advantages in terms of computational time.

In summary, we showed that the generalized Hybrid Monte Carlo methods based on a mollified multiple time stepping molecular dynamics and the sampling with respect to a shadow Hamiltonian specifically defined for a multi-scale integrator, demonstrate superior sampling performance not only over its progenitor GSHMC but also over Langevin dynamics and hybrid Monte Carlo methods which use similar force splitting schemes.

Acknowledgements

This research is supported by MICINN under grant MTM2011-24766 and also by the Basque Government through the BERC 2014-2017 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013-0323. JMA would like to thank the Spanish Ministry of Education for funding through a FPU fellowship (AP2009-1514). The SGI/IZO-SGIker UPV/EHU and the i2BASQUE academic network are acknowledged for computational resources. Fujitsu Laboratories of Europe U.K. is acknowledged for support and inspiration at the initial stage of this project.

Appendix A. Impulse time-stepping methods

In this appendix, we outline how (16) can be solved exactly to give rise to a numerical approximation for the differential equation

$$M \frac{d^2}{dt^2} X = -\nabla U^{fast}(X) - \nabla U_{molly}^{slow}(X). \quad (\text{A.1})$$

We note that equation (A.1) is conservative with energy (18).

We recall that the ratio between the inner step-size δt and the outer step-size Δt is denoted by $p = \Delta t / \delta t$ and L is the number of outer step-sizes. We introduce the notation $P(t' + \varepsilon)$ and $P(t' - \varepsilon)$ to denote the left and right hand side, respectively, limits of a time-dependent collective atomic momentum vector $P(t)$ with a discontinuity at $t = t'$. We also introduce $t_n = n\delta t$, $n = 0, \dots, pL$.

Integration of (16) once with respect to time reveals that $P(t) = \text{const.}$ whenever $t \neq t_n$, which implies

$$X(t_{n+1}) = X(t_n) + \delta t M^{-1} P(t_n + \varepsilon), \quad P(t_{n+1} - \varepsilon) = P(t_n + \varepsilon). \quad (\text{A.2})$$

For all $t = t_n$, $n = 0, \dots, pL$, an “impulse / kick” is applied to the velocities while the positions remain constant, i.e., $X(t_n + \varepsilon) = X(t_n - \varepsilon) = X(t_n)$ since $X(t)$ is continuous. The

magnitude of the impulse / kick depends on whether there is an integer m , $m = 0, \dots, L$, such that $n = mp$ (outer step-size) or not (inner step-size). In case of an outer step-size we have

$$P(t_n + \varepsilon) = P(t_n - \varepsilon) + c_m \Delta t F_{molly}^{slow}(X(t_n)) + d_n \delta t F^{fast}(X(t_n)), \quad (n = pm) \quad (\text{A.3})$$

while an inner step-size leads to

$$P(t_n + \varepsilon) = P(t_n - \varepsilon) + d_n \delta t F^{fast}(X(t_n)). \quad (\text{A.4})$$

The constant coefficients c_n and d_m take values as defined for (10).

Given an initial molecular state $\Gamma(0)$, we formally set $P(-\varepsilon) = P(0)$ to initiate the algorithm. Similarly, at final time $\tau = \Delta t L$, we formally set $P(\tau) = P(\tau + \varepsilon)$ to define the molecular state vector $\Gamma(\tau)$ at $t = \tau$. Hence we have constructed a mapping $\Psi_\tau : \Gamma(0) \rightarrow \Gamma(\tau)$ which maps an initial molecular state $\Gamma(0)$ into the desired numerical approximation $\Gamma(\tau)$ at time $\tau = L\Delta t = pL\delta t$. The mapping Ψ_τ is time-reversible, symplectic and volume conserving [13].

Appendix B. Shadow Hamiltonian definition

The fourth-order shadow Hamiltonian of Skeel and Hardy [25] for the Störmer-Verlet method applied to a system with Hamiltonian

$$H(X, P) = \frac{1}{2} P^T M^{-1} P + U(X) \quad (\text{B.1})$$

and step-size h is given by

$$\begin{aligned} H_h^{[4],SV}(X_i, P_i) = & H(X_i, P_i) + \frac{1}{4} [U(X_{i+1}) - 2U(X_i) + U(X_{i-1})] \\ & + \frac{h}{6} P_i^T M^{-1} [F(X_{i+1}) - F(X_{i-1})] + \frac{5h^2}{24} F(X_i)^T M^{-1} F(X_i) \\ & + \frac{h^2}{12} F(X_i)^T M^{-1} [F(X_{i+1}) - 2F(X_i) + F(X_{i-1})], \end{aligned} \quad (\text{B.2})$$

with $F = -\nabla U$. We obtain $H_{\delta t}^{[4],SV,fast}$ by setting $h = \delta t$, $U = U^{fast}$ and $F = F^{fast}$ and $H_{\Delta t}^{[4],SV,slow}$ by setting $h = \Delta t$, $U = U_{molly}^{slow}$ and $F = F_{molly}^{slow}$, respectively. Alternatively, one could use the fourth-order shadow Hamiltonian proposed in [1] for the Störmer-Verlet method.

- [1] E. Akhmatskaya, S. Reich, GSHMC: An efficient method for molecular simulation, Journal of Computational Physics 227 (10) (2008) 4934 – 4954. doi:10.1016/j.jcp.2008.01.023.

- [2] C. Wee, M. Sansom, S. Reich, E. Akhmatskaya, Improved sampling for simulations of interfacial membrane proteins: Application of GSHMC to a peptide toxin / bilayer system, *J. Phys. Chem. B* 112 (2008) 5710–5717.
- [3] B. Escribano, E. Akhmatskaya, J. Mujika, Combining stochastic and deterministic approaches within high efficiency molecular simulations, *Central European Journal of Mathematics* 11 (4) (2013) 787–799. doi:10.2478/s11533-012-0164-x.
- [4] E. Akhmatskaya, S. Reich, New hybrid Monte Carlo methods for efficient sampling: from physics to biology and statistics, *Progress in NUCLEAR SCIENCE and TECHNOLOGY* 2 (2011) 447–462.
- [5] J. Izaguirre, D. Catarello, J. Wozniak, R. Skeel, Langevin stabilization of molecular dynamics, *Journal of Chemical Physics* 114 (5) (2001) 2090–2098.
- [6] H. Grubmiller, H. Heller, A. Windemuth, K. Schulten, Generalized verlet algorithm for efficient molecular dynamics simulations with long-range interactions, *Molecular Simulation* 6 (1-3) (1991) 121–142. doi:10.1080/08927029108022142.
- [7] M. Tuckerman, B. J. Berne, G. J. Martyna, Reversible multiple time scale molecular dynamics, *The Journal of Chemical Physics* 97 (3) (1992) 1990–2001.
- [8] J. A. Morrone, T. E. Markland, M. Ceriotti, B. J. Berne, Efficient multiple time scale molecular dynamics: Using colored noise thermostats to stabilize resonances, *The Journal of Chemical Physics* 134 (1) (2011) –.
- [9] Q. Ma, J. A. Izaguirre, R. D. Skeel, Nonlinear instability in multiple time stepping molecular dynamics, in: *Proceedings of the 2003 ACM symposium on Applied computing, SAC '03*, ACM, New York, NY, USA, 2003, pp. 167–171.
- [10] R. Pastor, B. Brooks, A. Szabo, An analysis of the accuracy of Langevin and molecular dynamics algorithms, *Mol. Phys.* 65 (1988) 1409–1419.
- [11] S. Bond, B. Leimkuhler, GSHMC: An efficient method for molecular simulation, *Acta Numerica* 16 (2007) 1–65.
- [12] B. Garca-Archilla, G. M. Sanz-Serna, R. D. Skeel, Long-time-step methods for oscillatory differential equations, *SIAM J. Sci. Comput.* 20[3] (1998) 930–963.
- [13] J. Izaguirre, S. Reich, R. Skeel, Longer time steps for molecular dynamics, *Journal of Chemical Physics* 110 (1999) 9853–9864.

- [14] G. Han, Y. Deng, J. Glimm, G. Martyna, Error and timing analysis of multiple time-step integration methods for molecular dynamics, *Comput. Phys. Comm.* 176 (2007) 271–291.
- [15] P. Minari, M. E. Tuckerman, G. J. Martyna, Long time molecular dynamics for enhanced conformational sampling in biomolecular systems, *Phys. Rev. Lett.* 93 (2004) 150201.
- [16] C. Sweet, P. Petrone, V. Pande, J. Izaguirre, Normal mode partitioning of Langevin dynamics for biomolecules, *J. Chem. Phys.* 128 (2008) 145101.
- [17] E. Barth, T. Schlick, Overcoming stability limitations in biomolecular dynamics. I. Combining force splitting via extrapolation with Langevin dynamics in LN, *J. Chem. Phys.* 109 (1998) 1617.
- [18] X. Qian, T. Schlick, Efficient multiple-time-step integrators with distance-based force splitting for particle-mesh-Ewald molecular dynamics simulations, *J. Chem. Phys.* 116 (2002) 5971.
- [19] M. T. B. Leimkuhler, D.T. Margul, Stochastic, resonance-free multiple time-step algorithm for molecular dynamics with very large time steps, *Mol. Phys.* 111 (2013) 3579–3594.
- [20] A. Kennedy, B. Pendleton, Cost of the generalised hybrid Monte Carlo algorithm for free field theory, *Nuclear Physics B* 607 (3) (2001) 456 – 510. doi:10.1016/S0550-3213(01)00129-8.
- [21] A. M. Horowitz, A generalized guided Monte Carlo algorithm, *Physics Letters B* 268 (2) (1991) 247 – 252. doi:10.1016/0370-2693(91)90812-5.
- [22] E. Akhmatskaya, N. Bou-Rabee, S. Reich, Generalized hybrid Monte Carlo methods without momentum flip, *Journal of Computational Physics* 228 (2009) 2256–2265.
- [23] J. A. Izaguirre, S. S. Hampton, Shadow hybrid Monte Carlo: an efficient propagator in phase space of macromolecules, *Journal of Computational Physics* 200 (2) (2004) 581 – 604. doi:10.1016/j.jcp.2004.04.016.
- [24] E. Akhmatskaya, S. Reich, The targeted shadowing hybrid Monte Carlo (TSHMC) method, in B. L. et al., ed., *“New Algorithms for Macromolecular Simulations”*, Vol. 49 of *Lecture Notes in Computational Science and Engineering*, Springer-Verlag, Berlin, 2006.

- [25] R. Skeel, D. Hardy, Practical construction of modified Hamiltonians, *SIAM J. Sci. Comput* 23 (2001) 1172–1188.
- [26] E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration*, Springer-Verlag, Berlin Heidelberg, 2002.
- [27] T. Schlick, *Molecular Modeling and Simulation: An Interdisciplinary Guide*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2002.
- [28] M. P. Allen, D. J. Tildesley, *Computer simulation of liquids*, Clarendon Press, New York, NY, USA, 1989.
- [29] J. Zhou, S. Reich, B. Brooks, Elastic molecular dynamics with flexible constraints, *Journal of Chemical Physics* 112 (2000) 7919–7929.
- [30] T. Matthey, T. Cickovski, S. Hampton, A. Ko, Q. Ma, M. Nyerges, T. Raeder, T. Slabach, J. A. Izaguirre, Protomol, an object-oriented framework for prototyping novel algorithms for molecular dynamics, *ACM Trans. Math. Softw.* 30 (3) (2004) 237–265. doi:10.1145/1024074.1024075.
- [31] J. A. Wagoner, V. S. Pande, Reducing the effect of Metropolization on mixing times in molecular dynamics simulations, *J. Chem. Phys.* 137 (2012) 214105.
- [32] J. Liu, *Monte Carlo strategies in scientific computing*, Springer-Verlag, New York, 2001.
- [33] B. Leimkuhler, S. Reich, *Simulating Hamiltonian Dynamics*, Cambridge University Press, Cambridge, UK, 2005.
- [34] A. Beskos, N. Pillai, G. Roberts, J. Sanz-Serna, A. Stuart, Optimal tuning of the hybrid Monte Carlo algorithm, *Bernoulli Journal* 19 (2013) 1501–1534.
- [35] E. V. Akhmatskaya, S. Reich, A Method, Apparatus and Computer Program for Multiple Time Stepping Simulation of a Thermodynamic System Using Shadow Hamiltonians (08 2013).
- [36] E. V. Akhmatskaya, S. Reich, A Method, Apparatus and Computer Program for Multiple Time Stepping Simulation of a Thermodynamic System, Patent, uS20110010145(A1) (01 2011).
- [37] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water, *The Journal of Chemical Physics* 79 (2) (1983) 926–935. doi:10.1063/1.445869.

- [38] A. D. MacKerell, D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wirkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, *The Journal of Physical Chemistry B* 102 (18) (1998) 3586–3616. doi:10.1021/jp973084f.
- [39] P. Batcho, D. Case, T. Schlick, Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations, *Journal of Chemical Physics* 115 (9) (2001) 4003–4018.
- [40] R. D. Skeel, J. A. Izaguirre, An impulse integrator for Langevin dynamics, *Molecular Physics* 100 (24) (2002) 3885–3891.
- [41] C. J. Geyer, Practical Markov Chain Monte Carlo, *Statistical Science* 7 (4) (1992) 473–483.
- [42] B. Hess, Determining the shear viscosity of model liquids from molecular dynamics simulations, *The Journal of chemical physics* 116 (1) (2002) 209–217.
- [43] V. Ambegaokar, M. Troyer, Estimating errors reliably in Monte Carlo simulations of the Ehrenfest model, *American Journal of Physics* 78 (2) (2010) 150–157.
- [44] A. R. Leach, *Molecular modelling: principles and applications*, Addison-Wesley, Essex, 1996.
- [45] A. Soper, The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa, *Chemical Physics* 258 (23) (2000) 121–137.
- [46] P. D. Kirchhoff, M. B. Bass, B. A. Hanks, J. Briggs, A. Collet, J. A. McCammon, Structural fluctuations of a cryptophane host: a molecular dynamics simulation, *J. Am. Chem. Soc.* 118 (1996) 3237–3246.
- [47] M. Fernández-Pendás, B. Escibano, T. Radivojevic, E. Akhmatkaya, S. Reich, Momentum flip in hybrid Monte Carlo applications, (in preparation).