Multi-stage splitting integrators for sampling with modified Hamiltonian Monte Carlo methods

Tijana Radivojević[a,*], Mario Fernández-Pendás[a], Jesús María Sanz-Serna[c], Elena Akhmatskaya[a,b]

[a]*BCAM - Basque Center for Applied Mathematics, Alameda de Mazarredo 14, E-48009 Bilbao, Spain*
[b]*IKERBASQUE, Basque Foundation for Science, María Díaz de Haro 3, E-48013 Bilbao, Spain*
[c]*Departamento de Matemáticas, Universidad Carlos III de Madrid, Avenida de la Universidad 30, E-28911 Leganés (Madrid), Spain*

Abstract

Modified Hamiltonian Monte Carlo (MHMC) methods combine the ideas behind two popular sampling approaches: Hamiltonian Monte Carlo (HMC) and importance sampling. As in the HMC case, the bulk of the computational cost of MHMC algorithms lies in the numerical integration of a Hamiltonian system of differential equations. We suggest novel integrators designed to outperform the Verlet/leapfrog method, the current method of choice. The novel integrators belong to families of splitting algorithms and are therefore easily implemented. We identify optimal integrators within the families by minimizing the energy error or the average energy error. We derive and discuss in detail the modified Hamiltonians of the new integrators, as the evaluation of those Hamiltonians is key to the efficiency of the overall algorithms. Numerical experiments show that the use of the new integrators may enhance very significantly the sampling performance of MHMC methods, in both statistical and molecular dynamics problems.

Keywords: Hamiltonian Monte Carlo, modified Hamiltonian, multi-stage integrators, enhanced sampling

Contents

---

*Contact address: `tijana.radivojevic@gmail.com`

1. Introduction

Hybrid/Hamiltonian Monte Carlo (HMC) [1, 2] and importance sampling [3] algorithms have been effectively used for sampling in molecular simulation and computational statistics. Successful alternatives to these two approaches may be found within the class of modified Hamiltonian Monte Carlo (MHMC) methods that combine HMC and importance sampling, as suggested in [4, 5, 6]. Taking advantage of the fact that symplectic integrators preserve modified Hamiltonians more accurately than they conserve the true Hamiltonian, the authors of those references proposed to sample with respect to modified/shadow Hamiltonians and to recover the target distribution by reweighting. The resulting algorithms are capable of maintaining high acceptance rates and usually exhibit better efficiency than their predecessor HMC, that samples with respect to the true Hamiltonian [7, 8, 9, 10, 6].

The first methods of the MHMC class were derived for atomistic simulations and differed from each other in the way of performing the three main components of the algorithm— momentum updates, computation of modified Hamiltonians and integration of the Hamiltonian dynamics. For example, in the (Separable) Shadow Hybrid Monte Carlo methods introduced in [4, 11] a full momentum update is used, whereas in Targeted Shadow Hybrid Monte Carlo [12] and Generalized Shadow Hybrid Monte Carlo (GSHMC [5]) suitable modifications of the partial momentum update of [13] are advocated in order to better mimic the dynamics and to enhance sampling. More recent MHMC methods aim at specific applications such as computational statistics (Mix&Match Hamiltonian Monte Carlo [6]), or, in molecular problems, multi-scale (MTS-GSHMC [14]), meso-scale (meso-GSHMC [15]) and solid-state (RSM-GSHMC [10]) simulations. As demonstrated in the original papers, for

particular problems, the use of MHMC methods resulted in a sampling efficiency several times higher than that observed with conventional sampling techniques, such as Molecular Dynamics, Monte Carlo (MC) and HMC.

Two elements critically affect the efficiency of MHMC methods. On the one hand, the calculation of modified Hamiltonians at every Monte Carlo step introduces computational overheads when compared to HMC, and hence finding efficient ways of calculating modified Hamiltonians is of interest. On the other hand, the accuracy of the numerical integrator directly influences the acceptance rate of the generated proposals and therefore the sampling efficiency. Verlet/leapfrog has been the integrator of choice for MHMC methods and until recently such a choice has never been challenged. On the contrary, the search for the best performing integrator for HMC has been a live research topic for a decade [16, 17, 18, 19, 20]. In particular, it has been demonstrated that replacing in HMC the standard Verlet integrator with a splitting integrator that includes a specifically chosen parameter value may significantly improve, for a range of step sizes, the conservation of the Hamiltonian and thus the proposal acceptance rate [16, 17].

The goal of this paper is to adapt to MHMC methods the splitting schemes introduced in the literature as successful alternatives to Verlet in the HMC case. In particular, we propose novel multi-stage integrators that may improve the conservation of modified Hamiltonians and, as a consequence, the proposal acceptance rates in the MHMC methods. These integrators are characterized by parameter values resulting from minimization of the (expected) Hamiltonian error introduced by integration. More specifically, we derive two- and three-stage integrators along with their corresponding modified Hamiltonians, and investigate their performance within two specific MHMC methodologies, GSHMC [5] and MMHMC [6], developed for molecular simulation and computational statistics, respectively.

The paper is structured as follows. In Section 2 we present a brief review of modified HMC methods and their main components. In Section 3 we provide new formulations of modified Hamiltonians of 4th and 6th order for splitting integrating schemes that include families of two- and three-stage integrators. Section 4 provides new, optimized, multi-stage integrators arising from two different optimization approaches, namely (i) minimization, in the limit where the step-size approaches 0, of the error in modified Hamiltonian introduced by numerical integration, and (ii) minimization, for finite step-sizes, of the expected value of that error. Section 5 is devoted to testing the new integrators. We give details of the methods and models tested and the performance metrics employed, together with results of comparisons between the approaches suggested here and other popular integration techniques for HMC methods. Our conclusions are summarized in Section 6.

3

2. Modified Hamiltonian Monte Carlo methods

The goal of a modified Hamiltonian Monte Carlo (MHMC) method is to estimate an integral

$$I = \int f(\mathbf{x})\pi(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{1}$$

i.e. the expected value of a function of interest f with respect to the density $\pi(\mathbf{x})$, known up to a multiplicative constant, of a position variable $\mathbf{x}$. Instead of sampling from the canonical distribution

$$\pi(\mathbf{x}, \mathbf{p}) \propto \exp\left(-\beta H(\mathbf{x}, \mathbf{p})\right) \tag{2}$$

(β is the inverse temperature, H the Hamiltonian and $\mathbf{p}$ the auxiliary momentum variable), as is the case of general HMC algorithms, MHMC methods sample from an importance canonical density

$$\tilde{\pi}(\mathbf{x}, \mathbf{p}) \propto \exp\left(-\beta \tilde{H}^{[k]}(\mathbf{x}, \mathbf{p})\right), \tag{3}$$

where $\tilde{H}^{[k]}$ denotes the kth order truncation of the modified Hamiltonian $\tilde{H}$ that is preserved exactly by the symplectic integrator under consideration.

In this paper we consider separable Hamiltonians

$$H(\mathbf{x}, \mathbf{p}) = U(\mathbf{x}) + \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p}, \tag{4}$$

with M a symmetric positive definite matrix (mass matrix) and U the potential function. The dynamics associated to this Hamiltonian is governed by the system of ordinary differential equations

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = M^{-1}\mathbf{p}, \quad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -U_{\mathbf{x}}(\mathbf{x}). \tag{5}$$

We describe now a generic algorithm for an MHMC method. Given a sample $(\mathbf{x}, \mathbf{p})$ from the distribution $\tilde{\pi}$, the next sample $(\mathbf{x}^{\text{new}}, \mathbf{p}^{\text{new}})$ is determined as follows:

- Define the new momentum $\mathbf{p}^*$ by applying a momentum refreshment procedure that preserves the importance density $\tilde{\pi}$ in (3).

- Generate a proposal $(\mathbf{x}', \mathbf{p}')$ by simulating, using a symplectic and reversible numerical integrator, the Hamiltonian system (5) with initial condition $(\mathbf{x}, \mathbf{p}^*)$.

- Accept the proposal as the next sample $(\mathbf{x}^{\text{new}}, \mathbf{p}^{\text{new}})$ with probability

$$\alpha = \min\left\{1, \frac{\tilde{\pi}(\mathbf{x}', \mathbf{p}')}{\tilde{\pi}(\mathbf{x}, \mathbf{p}^*)}\right\}. \tag{6}$$

Otherwise, set $(\mathbf{x}^{\text{new}}, \mathbf{p}^{\text{new}})$ to $(\mathbf{x}, -\mathbf{p}^*)$, i.e., flip the momentum.

Once the samples have been obtained using MHMC, importance reweighting is required in order to estimate (1), since the samples are generated with respect to the importance density (3) on the joint state space. The desired distribution $\pi(\mathbf{x})$ is recovered by marginalizing momenta variables from (2). If $f_n = f(\mathbf{x}^n)$, $n = 1, 2, \ldots, N$, are the values of an observable along a sequence of states $(\mathbf{x}^n, \mathbf{p}^n)$ drawn from $\tilde{\pi}$, (1) is estimated as

$$\hat{I} = \frac{\sum_{n=1}^{N} w_n f_n}{\sum_{n=1}^{N} w_n}, \tag{7}$$

where the *importance weights* are given by

$$w_n = \exp\left(-\beta\big(H(\mathbf{x}^n, \mathbf{p}^n) - \tilde{H}^{[k]}(\mathbf{x}^n, \mathbf{p}^n)\big)\right). \tag{8}$$

3. Modified Hamiltonians for splitting integrators

Our intention is to use within MHMC methods numerical integrators that may offer better conservation properties than the commonly used Verlet/leapfrog integrator. More specifically, we are interested in numerical integrators belonging to the two-stage

$$\psi_h = \varphi_{bh}^B \circ \varphi_{\frac{h}{2}}^A \circ \varphi_{(1-2b)h}^B \circ \varphi_{\frac{h}{2}}^A \circ \varphi_{bh}^B \tag{9}$$

and three-stage

$$\psi_h = \varphi_{bh}^B \circ \varphi_{ah}^A \circ \varphi_{(\frac{1}{2}-b)h}^B \circ \varphi_{(1-2a)h}^A \circ \varphi_{(\frac{1}{2}-b)h}^B \circ \varphi_{ah}^A \circ \varphi_{bh}^B \tag{10}$$

families of splitting methods, that require two or three gradient evaluations per step, respectively. The exact flows φ_h^A and φ_h^B are solutions to the split systems

$$A : \frac{d\mathbf{x}}{dt} = 0, \quad \frac{d\mathbf{p}}{dt} = -U_\mathbf{x}(\mathbf{x}), \tag{11}$$

and

$$B : \frac{d\mathbf{x}}{dt} = M^{-1}\mathbf{p}, \quad \frac{d\mathbf{p}}{dt} = 0, \tag{12}$$

respectively, corresponding to the Hamiltonian (4). A specific integrator is fully characterized by the choice of values for the coefficients $\{b\}$ or $\{a, b\}$.

We point out that for the specific choice $b = 1/4$, the application of ψ_h in (9) yields the same result as performing two consecutive Verlet steps each with step size $h/2$. Similarly for $b = 1/6$, $a = 1/3$, (10) is equivalent to three Verlet steps of size $h/3$. It is also important to emphasize that since one step of a splitting integrator consists of a sequence of flows of

5

(11)–(12) (some times called kicks and drifts respectively), the implementation of methods of the families (9)–(10) is very similar to that of the Verlet algorithm. Therefore it is easy to incorporate splitting algorithms to existing software based on the Verlet integrator.

To use splitting integrators within an MHMC method, appropriate modified Hamiltonians are required. For the Verlet integrator, one procedure to compute modified Hamiltonians of orders up to 24 is provided in [21, 22]. It is further improved in [23] using Richardson extrapolation. That approach could be generalized to multi-stage integrators. However, it requires a modification of the integrator by introducing into the dynamics an additional scalar variable. We therefore opt here for a different strategy. For the families of splitting integrators above, we begin by writing the expansion of the modified Hamiltonian in terms of Poisson brackets of the partial Hamiltonians A and B of the split systems (11)-(12). We recall that the *Poisson bracket* of two functions $F, G : \mathbb{R}^{2D} \to \mathbb{R}$ is defined as

$$\{F, G\}(\mathbf{z}) = F_{\mathbf{z}}(\mathbf{z})^T \mathbf{J} G_{\mathbf{z}}(\mathbf{z}), \ \mathbf{J} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{13}$$

The expression for the modified Hamiltonian $\tilde{H}$ is found to be

$$\begin{aligned} \tilde{H} = H \ &+ \ h^2\alpha\{A, A, B\} + h^2\beta\{B, B, A\} \\ &+ \ h^4\gamma_1\{A, A, A, A, B\} + h^4\gamma_2\{B, A, A, A, B\} \\ &+ \ h^4\gamma_3\{B, B, A, A, B\} + h^4\gamma_4\{A, A, B, B, A\} + \mathcal{O}(h^6), \end{aligned} \tag{14}$$

where $\alpha, \beta, \gamma_{1-4}$ are polynomials in the integrator coefficients b or a, b [17] and expressions such as $\{A, A, B\}$ refer to iterated Poisson brackets $\{A, \{A, B\}\}$.

The expressions for the expansion of $\tilde{H}$ to arbitrarily high order may be obtained by directly applying the Baker-Campbell-Hausdorff (BCH) formula to the exponentials of the Lie derivatives of the partial Hamiltonians A and B. Recall that the *Lie derivative* with respect to a function $F : \mathbb{R}^{2D} \to \mathbb{R}$ is defined in terms of Poisson bracket as

$$\mathcal{L}_F(\cdot) = \{\cdot, F\}. \tag{15}$$

Unfortunately computations based on the BCH are cumbersome if the order is high [24]. Alternatively, for symmetric composition methods, the coefficients multiplying the Poisson brackets for the 4th, 6th and 8th order truncations of the modified Hamiltonian can be derived from expressions presented in [25]. In the case of general non-symmetric composition methods with an arbitrary number of stages, one can obtain the coefficients α and β using results derived in [26], Lemma III.5.5.

Here we propose two alternative ways to derive the expression for the 4th and 6th order modified Hamiltonians. The first uses derivatives of the potential function, obtained either by automatic differentiation or from analytical expressions, whereas the second relies on numerical time derivatives of the gradient, obtained through quantities available along the simulation.

6

3.1. Modified Hamiltonians in terms of derivatives of the potential

For problems in which derivatives of the potential function can be computed, either from analytical expressions or using automatic differentiation, we derive the 4th and 6th order modified Hamiltonians by first expanding the iterated brackets in (14) using the definition of the Poisson bracket. This yields

$$
\begin{aligned}
\{A, A, B\} &= \mathbf{p}^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} \mathbf{p} \\
\{B, B, A\} &= U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{x}}(\mathbf{x}) \\
\{A, A, A, A, B\} &= U_{\mathbf{xxxx}}(\mathbf{x}) M^{-1} \mathbf{p} M^{-1} \mathbf{p} M^{-1} \mathbf{p} M^{-1} \mathbf{p} \\
\{B, A, A, A, B\} &= -3 U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{xxx}}(\mathbf{x}) M^{-1} \mathbf{p} M^{-1} \mathbf{p} \\
\{B, B, A, A, B\} &= 2 U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} U_{\mathbf{x}}(\mathbf{x}) \\
\{A, A, B, B, A\} &= 2 U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{xxx}}(\mathbf{x}) M^{-1} \mathbf{p} M^{-1} \mathbf{p} \\
&\quad + 2 \mathbf{p}^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} \mathbf{p}
\end{aligned}
$$

and leads to the following 4th and 6th order modified Hamiltonians for splitting integrators

$$
\tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) = H(\mathbf{x}, \mathbf{p}) + h^2 c_{21} \mathbf{p}^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} \mathbf{p} + h^2 c_{22} U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{x}}(\mathbf{x}), \qquad (16)
$$

$$
\begin{aligned}
\tilde{H}^{[6]}(\mathbf{x}, \mathbf{p}) = \tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) &+ h^4 c_{41} U_{\mathbf{xxxx}}(\mathbf{x}) M^{-1} \mathbf{p} M^{-1} \mathbf{p} M^{-1} \mathbf{p} M^{-1} \mathbf{p} \qquad (17) \\
&+ h^4 c_{42} U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{xxx}}(\mathbf{x}) M^{-1} \mathbf{p} M^{-1} \mathbf{p} \\
&+ h^4 c_{43} U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} U_{\mathbf{x}}(\mathbf{x}) \\
&+ h^4 c_{44} \mathbf{p}^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} \mathbf{p},
\end{aligned}
$$

where

$$
c_{21} = \alpha, \quad c_{22} = \beta, \quad c_{41} = \gamma_1, \quad c_{42} = 2\gamma_4 - 3\gamma_2, \quad c_{43} = 2\gamma_3, \quad c_{44} = 2\gamma_4. \qquad (18)
$$

The coefficients $\alpha, \beta, \gamma_{1-4}$ can be derived from expressions given in [25] where the authors analyzed so-called force-gradient integrators for molecular dynamics. In particular, they considered splitting integrators that are extended by an additional higher-order operator into the single-exponential propagators. If the potential function is quadratic, i.e. corresponding to problems of sampling from Gaussian distributions/harmonic oscillators, the 6th order modified Hamiltonian (17) simplifies to

$$
\begin{aligned}
\tilde{H}^{[6]}(\mathbf{x}, \mathbf{p}) = \tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) &+ h^4 c_{43} U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} U_{\mathbf{x}}(\mathbf{x}) \qquad (19) \\
&+ h^4 c_{44} \mathbf{p}^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} \mathbf{p}.
\end{aligned}
$$

Combining (18) with the expressions for $\alpha, \beta, \gamma_{1-4}$, we obtain the following coefficients for the two-stage integrator family (9)

$$
\begin{aligned}
c_{21} &= \frac{1}{24}\left(6b - 1\right) \\
c_{22} &= \frac{1}{12}\left(6b^2 - 6b + 1\right) \\
c_{41} &= \frac{1}{5760}\left(7 - 30b\right) \\
c_{42} &= \frac{1}{240}\left(-10b^2 + 15b - 3\right) \\
c_{43} &= \frac{1}{120}\left(-30b^3 + 35b^2 - 15b + 2\right) \\
c_{44} &= \frac{1}{240}(20b^2 - 1).
\end{aligned}
\tag{20}
$$

For three-stage integrators (10) (a two-parameter family), we get

$$
\begin{aligned}
c_{21} &= \frac{1}{12}\left(1 - 6a(1 - a)(1 - 2b)\right) \\
c_{22} &= \frac{1}{24}\left(6a(1 - 2b)^2 - 1\right) \\
c_{41} &= \frac{1}{720}\left(1 + 2(a - 1)a(8 + 31(a - 1)a)(1 - 2b) - 4b\right) \\
c_{42} &= \frac{1}{240}\left(6a^3(1 - 2b)^2 - a^2(19 - 116b + 36b^2 + 240b^3) + a(27 - 208b + 308b^2) - 48b^2 + 48b - 7\right) \\
c_{43} &= \frac{1}{180}\left(1 + 15a(1 - 2b)(-1 + 2a(2 - 3b + a(4b - 2)))\right) \\
c_{44} &= \frac{1}{240}\left(-1 + 20a(1 - 2b)(b + a(1 + 6(b - 1)b))\right).
\end{aligned}
\tag{21}
$$

Using (20) one can also obtain the modified Hamiltonian for the Verlet integrator, since, as pointed out above, two steps of Verlet integration with step size $h/2$ are equivalent to one step with step size h of the two-stage integrator with $b = 1/4$. In this way the Verlet coefficients are found to be

$$
\begin{aligned}
c_{21} &= \frac{1}{12}, \quad c_{22} = -\frac{1}{24} \\
c_{41} &= -\frac{1}{720}, \quad c_{42} = \frac{1}{120}, \quad c_{43} = -\frac{1}{240}, \quad c_{44} = \frac{1}{60}.
\end{aligned}
\tag{22}
$$

Figure 1 illustrates, for two-stage integrators, the computational overheads, with respect to the HMC method, of MMHMC based on the 4th order modified Hamiltonian (16). The

8

left-hand graph presents the overhead for a model with a tridiagonal Hessian matrix $U_{\mathbf{xx}}(\mathbf{x})$, as appeared in the stochastic volatility model [7], and indicates that, for two different dimensions D of the system, the overhead becomes negligible as the number of integration steps increases. In contrast, for models with a dense Hessian matrix, as appeared in the Bayesian logistic regression model [7], computation of modified Hamiltonians may introduce a significant additional cost, as shown in the right-hand graph.
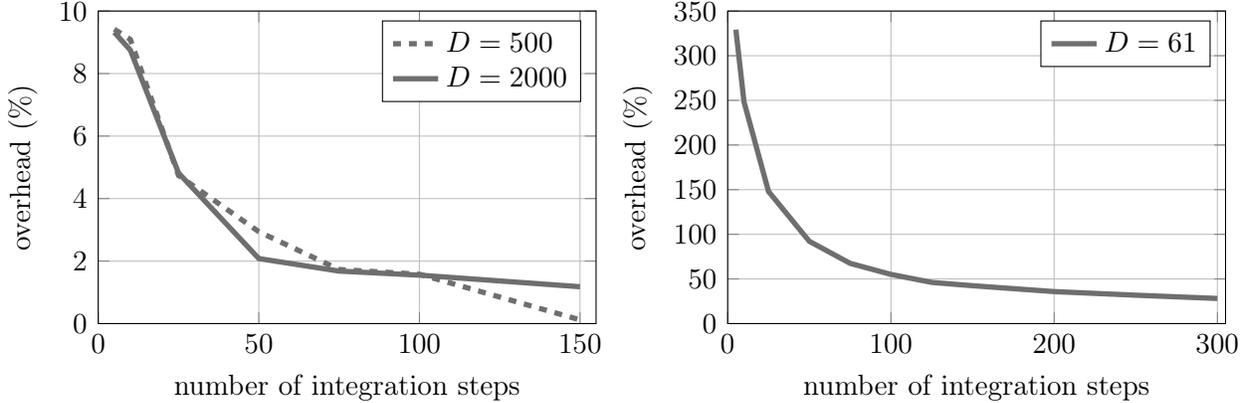

Figure 1: Computational overhead of MMHMC compared to HMC for models with a tridiagonal (left) and a dense Hessian matrix (right) using the 4th order modified Hamiltonian (16) with all required derivatives calculated analytically.

3.2. Modified Hamiltonians in terms of time derivatives of the gradient

For applications where the potential function has a dense Hessian matrix (and tensors of higher derivatives), the overhead resulting from the computation of modified Hamiltonians reduces the advantages of MHMC methods. In order to implement such computations in an efficient manner, we propose to express modified Hamiltonians in terms of quantities that are available during the simulation, as done in [5]. However, instead of making use of time derivatives of the position vectors, as in the original GSHMC method [5], here we employ time derivatives of the gradient of the potential function, as follows (superscripts in brackets indicate the order of the time derivative)

$$
\begin{aligned}
U_{\mathbf{x}}^{(1)} &= U_{\mathbf{xx}}(\mathbf{x})M^{-1}\mathbf{p} \\
U_{\mathbf{x}}^{(2)} &= U_{\mathbf{xxx}}(\mathbf{x})M^{-1}\mathbf{p}M^{-1}\mathbf{p} - U_{\mathbf{xx}}(\mathbf{x})M^{-1}U_{\mathbf{x}}(\mathbf{x}) \\
U_{\mathbf{x}}^{(3)} &= U_{\mathbf{xxxx}}(\mathbf{x})M^{-1}\mathbf{p}M^{-1}\mathbf{p}M^{-1}\mathbf{p} - 3U_{\mathbf{xxx}}(\mathbf{x})M^{-1}U_{\mathbf{x}}(\mathbf{x})M^{-1}\mathbf{p} \\
&\quad - U_{\mathbf{xx}}(\mathbf{x})M^{-1}U_{\mathbf{xx}}(\mathbf{x})M^{-1}\mathbf{p}.
\end{aligned}
\tag{23}
$$

9

Substituting the time derivatives (23) into the analytical expressions (16)-(17) for the 4th and 6th order modified Hamiltonians, respectively, one obtains

$$\tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) = H(\mathbf{x}, \mathbf{p}) + h^2 k_{21} \mathbf{p}^T M^{-1} U_{\mathbf{x}}^{(1)} + h^2 k_{22} U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{x}}(\mathbf{x}), \tag{24}$$

$$\begin{aligned} \tilde{H}^{[6]}(\mathbf{x}, \mathbf{p}) &= \tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) + h^4 k_{41} \mathbf{p}^T M^{-1} U_{\mathbf{x}}^{(3)} + h^4 k_{42} U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{x}}^{(2)} \\ &+ h^4 k_{43} U_{\mathbf{x}}^{(1)T} M^{-1} U_{\mathbf{x}}^{(1)} + h^4 k_{44} U_{\mathbf{x}}(\mathbf{x})^T M^{-1} U_{\mathbf{xx}}(\mathbf{x}) M^{-1} U_{\mathbf{x}}(\mathbf{x}), \end{aligned} \tag{25}$$

where the coefficients are

$$\begin{aligned} k_{21} &= c_{21}, \quad k_{22} = c_{22}, \\ k_{41} &= c_{41}, \quad k_{42} = 3c_{41} + c_{42}, \quad k_{43} = c_{41} + c_{44}, \quad k_{44} = 3c_{41} + c_{42} + c_{43}. \end{aligned} \tag{26}$$

We note that the expression (24) does not include the Hessian of the potential and thus, allows the computation of $\tilde{H}^{[4]}$ using quantities available from the integration of the dynamics. However, this is not the case for the 6th order Hamiltonians. The last term in (25), arising from an expansion of the Poisson bracket $\{B, B, A, A, B\}$, cannot be computed using time derivatives of available quantities and requires explicit calculation of the Hessian matrix of the potential function. Only for the Verlet integrator does this term vanish and the resulting coefficients are

$$\begin{aligned} k_{21} &= \frac{1}{12}, \quad k_{22} = -\frac{1}{24}, \\ k_{41} &= -\frac{1}{720}, \quad k_{42} = \frac{1}{240}, \quad k_{43} = \frac{11}{720}, \quad k_{44} = 0. \end{aligned} \tag{27}$$

One can now write explicit expressions for the coefficients k_{ij} by substituting the expressions for the coefficients c_{ij} (20) or (21) into (26) for two- and three-stage integrators, respectively.

In the original GSHMC method, an interpolating polynomial of the positions $\mathbf{x}(t_i) = \mathbf{x}^i$, $i = n - k, \ldots, n, \ldots, n + k$, $n \in \{0, L\}$ is constructed from a numerical trajectory $\{\mathbf{x}^i\}_{i=-k}^{L+k}$, where $k = 2$ and $k = 3$ for the 4th and 6th order modified Hamiltonian, respectively. This requires four or six additional gradient calculations in order to compute $\tilde{H}^{[4]}$ or $\tilde{H}^{[6]}$, respectively. Here we choose a different strategy and calculate the polynomial in terms of the gradient of the potential function

$$\mathbf{U}(t_i) = U_{\mathbf{x}}(\mathbf{x}^i), \ i = n - k, \ldots, n, \ldots, n + k.$$

With this approach, $k = 1$ for the 4th order and $k = 2$ for the 6th order modified Hamiltonian, which implies that the evaluation of $\tilde{H}^{[4]}$ and $\tilde{H}^{[6]}$ requires fewer additional gradient computations, namely two and four, respectively. Note that k corresponds to a multiple of

10

the full integration step only in the case of the Verlet integrator; for other integrators it is the number of stages performed (e.g. $k = 2$ corresponds to a half integration step of a four-stage method). Also note that an efficient implementation does not include the unnecessary integration sub-step that would update momentum at the very beginning and very end of the numerical trajectory $\{U_{\mathbf{x}}(\mathbf{x}^i)\}_{i=-k}^{L+k}$.

Time derivatives of the gradient of the potential function are approximated using central finite differences of second order of accuracy for the 4th order modified Hamiltonian

$$U_{\mathbf{x}}^{(1)} \approx \frac{\mathbf{U}(t_{n+1}) - \mathbf{U}(t_{n-1})}{2\varepsilon} =: \mathbf{U}^{(1)}, \tag{28}$$

where $\varepsilon = h$ for the Verlet, $\varepsilon = h/2$ for two-stage and $\varepsilon = ah$ for three-stage integrators (h is the integration step size and a the coefficient in (10). The 6th order modified Hamiltonian, here considered only for the Verlet and two-stage integrators, is calculated using centered differences of fourth order accuracy for the first derivative and second order accuracy for the second and third derivatives

$$\begin{aligned}
U_{\mathbf{x}}^{(1)} &\approx \frac{\mathbf{U}(t_{n-2}) - 8\mathbf{U}(t_{n-1}) + 8\mathbf{U}(t_{n+1}) - \mathbf{U}(t_{n+2})}{12\varepsilon} =: \mathbf{U}^{(1)} \\
U_{\mathbf{x}}^{(2)} &\approx \frac{\mathbf{U}(t_{n-1}) - 2\mathbf{U}(t_n) + \mathbf{U}(t_{n+1})}{\varepsilon^2} =: \mathbf{U}^{(2)} \\
U_{\mathbf{x}}^{(3)} &\approx \frac{-\mathbf{U}(t_{n-2}) + 2\mathbf{U}(t_{n-1}) - 2\mathbf{U}(t_{n+1}) + \mathbf{U}(t_{n+2})}{2\varepsilon^3} =: \mathbf{U}^{(3)},
\end{aligned}$$

where ε depends on the integrator as before.

The final expressions for the modified Hamiltonians computed in this way are

$$\tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) = H(\mathbf{x}, \mathbf{p}) + hk_{21}\mathbf{p}^T M^{-1}P_1 + h^2 k_{22}U_{\mathbf{x}}(\mathbf{x})^T M^{-1}U_{\mathbf{x}}(\mathbf{x}) \tag{29}$$

$$\begin{aligned}
\tilde{H}^{[6]}(\mathbf{x}, \mathbf{p}) &= \tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) + hk_{41}\mathbf{p}^T M^{-1}P_3 + h^2 k_{42}U_{\mathbf{x}}(\mathbf{x})^T M^{-1}P_2 \\
&+ h^2 k_{43}P_1^T M^{-1}P_1 + h^4 k_{44}U_{\mathbf{x}}(\mathbf{x})^T M^{-1}U_{\mathbf{xx}}(\mathbf{x})M^{-1}U_{\mathbf{x}}(\mathbf{x}),
\end{aligned} \tag{30}$$

where $P_i = \mathbf{U}^{(i)} \cdot h^i$. Note that the term with coefficient k_{22} is calculated exactly, i.e. avoiding finite difference approximation; this improves the approximation of the modified Hamiltonian when compared to the strategy originally used in GSHMC. Also note that, compared to the expressions with analytical derivatives (16) and (17), in the formulations (29) and (30) the terms involving the coefficients c_{21}, c_{41}, c_{42} and c_{44} are approximated by P_i. The order of accuracy provided by the modified Hamiltonians (29) and (30), however, is not affected by these approximations.

The computational overhead of the MMHMC method [6] using the modified Hamiltonians (29) or (30), when compared to HMC is shown in Figure 2 for models with a tridiagonal (left-hand graph) and dense Hessian matrix (right-hand graph) of the potential, as described for Figure 1. Compared to Figure 1, where all derivatives are calculated analytically, we note that for models with a sparse Hessian (left-hand graphs), the 4th order modified Hamiltonian (16) with analytical derivatives introduces less computational overhead than (29) with a numerical approximation of the time derivative. This is due to the additional forward and backward integration steps, which do not counterbalance the inexpensive Hessian computation. For models with a dense Hessian matrix (right-hand graphs) we recommend always using (29), which significantly reduces the overhead. The 6th order modified Hamiltonian (30) clearly requires additional computational effort, due to two extra gradient calculations per MC iteration. In the following sections we show that using modified Hamiltonians of order higher than 4 can be avoided by introducing accurate multi-stage integrators specifically tuned for the MHMC methods.
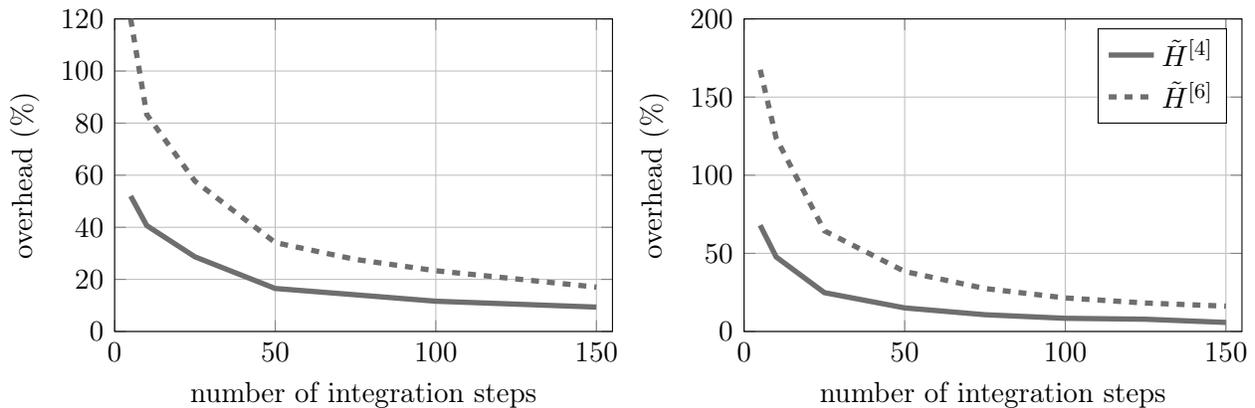

Figure 2: Computational overhead of MMHMC compared to HMC for models with a tridiagonal (left) and a dense (right) Hessian matrix, using 4th and 6th order modified Hamiltonians with numerical approximation of the time derivatives.

In summary, we have provided two alternative formulations of the 4th and 6th order modified Hamiltonians corresponding to multi-stage integrators (9)–(10) with arbitrary coefficients. For the cases when derivatives of the potential function are available and inexpensive to compute, the modified Hamiltonians can be calculated using (16)–(22). For problems in which this is not the case, we provided formulations of modified Hamiltonians which mainly rely on quantities available from the simulation. Both approaches can be used with any multi-stage integrator (9)–(10), including the Verlet integrator.

In the following section, we choose the coefficients in the families (9)–(10) so as to obtain methods specifically aimed at sampling with modified Hamiltonians.

4. Multi-stage integrators with optimized coefficients

Until now, the Verlet/leapfrog integrator has been the integrator of choice for MHMC methods. In this section, we consider alternative integrators and investigate their competitiveness with the Verlet integrator.

Our focus is on multi-stage integrators belonging to families (9)–(10). There are two reasons for an interest in these integrators. One is their potential to achieve, at a given computational cost, higher accuracy than Verlet. Here it is important to emphasize that, in comparisons, different integrators have to be applied with the same computational effort, rather than with the same step length; an r-stage integrator requires r gradient evaluations per time step and to be compared with Verlet has to be used with a step length correspondingly longer. More accurate integrations imply higher acceptance rates and thus better space exploration. A second possible benefit of the integrators of this class is that, due to the extra accuracy, they may avoid the need for computationally expensive, higher order modified Hamiltonians.

Our goal is to derive new multi-stage integrators to be used in the methods which sample with modified Hamiltonians and compare their impact on the performance of such methods with the efficiency of advanced integrators for HMC [16, 17] and the Verlet integrator.

In MHMC methods, the equations of motion of the Hamiltonian dynamics are the same as in HMC methods. However, MHMC are based on different Metropolis tests where the acceptance rate depends on the capability of the integrator to conserve the value of a modified Hamiltonian. Indeed, the sampling performance of MHMC is controlled not by the energy error with respect to the true Hamiltonian as in HMC, but by the energy error with respect to the modified Hamiltonian. Thus, inspired by the ideas of [16] and [17] for improving HMC performance by minimizing (expected) energy error through the appropriate choice of parameters of the integrator, in order to enhance the performance of MHMC, we design the new integrators by considering either the error in the modified Hamiltonians $\tilde{H}^{[l]}$ of order l

$$\Delta = \tilde{H}^{[l]}(\Psi_{h,L}(\mathbf{x}, \mathbf{p})) - \tilde{H}^{[l]}(\mathbf{x}, \mathbf{p}), \tag{31}$$

or the expected values of such errors $\mathbb{E}_{\tilde{\pi}}(\Delta)$ taken with respect to the modified canonical density $\tilde{\pi}$ given by (3). Here $\Psi_{h,L}(\mathbf{x}, \mathbf{p})$ is the hL-time map of the integrator. In order to distinguish the new minimum error and minimum expected error integrators for sampling with modified (M) Hamiltonians from the corresponding ones designed for the HMC method, we use the prefix M-; for instance M-ME will denote minimum error for sampling with modified Hamiltonians.

4.1. Minimum error (M-ME) integrators

We first construct the minimum error integrators for the 4th order modified Hamiltonian.

The Taylor expansion of the 4th order modified Hamiltonian after one integration step with the method Ψ_h can be written as [24]

$$
\begin{aligned}
\tilde{H}^{[4]}(\mathbf{x}', \mathbf{p}') &= \tilde{H}^{[4]}(\Psi_h(\mathbf{x}, \mathbf{p})) = \exp\left(h\mathcal{L}_{\tilde{H}}\right)\tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) \\
&= \tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) + h\mathcal{L}_{\tilde{H}}\tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) + \frac{1}{2}h^2\mathcal{L}_{\tilde{H}}^2\tilde{H}^{[4]}(\mathbf{x}, \mathbf{p}) + \ldots,
\end{aligned}
$$

where $\tilde{H}$ is the modified Hamiltonian (14) and we have used the Lie derivative (15). The error Δ in $\tilde{H}^{[4]}$ after one integration step reads

$$
\begin{aligned}
\Delta(\mathbf{x}, \mathbf{p}) = h^5 \big(&\gamma_1\{A, A, A, A, A, B\}(\mathbf{x}, \mathbf{p}) + \gamma_1\{B, A, A, A, A, B\}(\mathbf{x}, \mathbf{p}) \\
&+\gamma_2\{A, B, A, A, A, B\}(\mathbf{x}, \mathbf{p}) + \gamma_2\{B, B, A, A, A, B\}(\mathbf{x}, \mathbf{p}) \\
&+\gamma_3\{A, B, B, A, A, B\}(\mathbf{x}, \mathbf{p}) + \gamma_3\{B, B, B, A, A, B\}(\mathbf{x}, \mathbf{p}) \\
&+\gamma_4\{A, A, A, B, B, A\}(\mathbf{x}, \mathbf{p}) + \gamma_4\{B, A, A, B, B, A\}(\mathbf{x}, \mathbf{p})\big) + \mathcal{O}(h^6).
\end{aligned}
\tag{32}
$$

Thus, as h approaches 0 and the $\mathcal{O}(h^6)$ remainder may be ignored, an error metric for the 4th order modified Hamiltonian can then be defined by the following function of the integrating coefficients

$$
E = \sqrt{\gamma_1^2 + \gamma_2^2 + \gamma_3^2 + \gamma_4^2}.
\tag{33}
$$

The expressions for γ_{1-4} derived from (18) are

$$
\gamma_1 = c_{41}, \quad \gamma_2 = \frac{1}{3}(c_{44} - c_{42}), \quad \gamma_3 = \frac{1}{2}c_{43}, \quad \gamma_4 = \frac{1}{2}c_{44}.
$$

The coefficients c_{ij} are calculated from (20) and (21) for two- and three-stage integrators, respectively. For quadratic potential and kinetic energies, corresponding to the problem of sampling from a Gaussian distribution with a covariance matrix Σ, or harmonic oscillators, the error (32) simplifies to

$$
\Delta(\mathbf{x}, \mathbf{p}) = h^5(\gamma_4 - \gamma_3)(4\mathbf{p}^T M^{-1} M^{-1} M^{-1} \Sigma^{-1} \Sigma^{-1} \Sigma^{-1} \mathbf{x}) + \mathcal{O}(h^6,
$$

as the other terms involve third (or higher) order partial derivatives of A and B and are equal to zero. Therefore, the error metric can be defined as

$$
E^Q = |\gamma_4 - \gamma_3|.
\tag{34}
$$

In contrast to this approach, the error metric for the minimum error integrator derived in [16] for sampling with the true Hamiltonian, i.e. the HMC method, is defined through the Hamiltonian truncation error $H - \tilde{H}$ at the state $(\mathbf{x}, \mathbf{p})$, rather than the error in Hamiltonian after numerical integration, as suggested here.

14

In order to obtain values for integrating coefficients for the MHMC methods, we minimized the metrics E or E^Q on the interval $(0, 0.5)$ using *Mathematica*. For the three-stage family of integrators for problems with quadratic potential and kinetic function, i.e. for E^Q, we take into account the analysis from [27], where the authors provide the condition for which the stability limit is the highest for these kind of integrators. In particular, the integrators that lie on the hyperbola

$$6ab - 2a - b + \frac{1}{2} = 0, \tag{35}$$

have considerably longer stability limit than others. Therefore, we enforce the relationship between a and b from (35) in the minimization of E^Q. The resulting values of coefficients for two- and three-stage minimum error integrators for quadratic (M-ME2, M-ME3) and general problems (M-ME2gen, M-ME3gen) are given in Table 1. We note that the difference between the coefficients for the two versions of the two-stage methods is minor.

The error metric definitions (33) and (34) are based on the assumption that the different iterated brackets that feature in (32) contribute equally to the Hamiltonian error. While this assumption is reasonable, in problems where information on the relative size of the iterated brackets is available one may modify accordingly the error metric so as to give more weight to the error coefficients with larger iterated bracket; of course that change in metric would entail a change in the coefficient values resulting from the minimization procedure.

4.2. Minimum expected error integrators (M-BCSS)

The modified Hamiltonians we consider here are of order 4 or 6. We adopt a strategy similar to the one proposed in [17], namely to consider the problem of a one-dimensional quadratic potential and to find the parameters of integrators that minimize the expected value of the error for a suitable range of finite values of h; note that this is different from the procedure above based on the behaviour of the algorithms as h approaches 0. In our case, the error (31), resulting from numerical integration, is in terms of the modified Hamiltonian and the expected value is taken with respect to the modified canonical density $\tilde{\pi}$.

As in the case when considering the error in the true Hamiltonian, one may prove that the expected error in the modified Hamiltonian $\mathbb{E}_{\tilde{\pi}}(\Delta)$ is also positive. The objective is, therefore, to find a function $\rho(h, \boldsymbol{\xi})$ that bounds $\mathbb{E}_{\tilde{\pi}}(\Delta)$, i.e.

$$0 \leq \mathbb{E}_{\tilde{\pi}}(\Delta) \leq \frac{1}{\beta} \rho(h, \boldsymbol{\xi}).$$

Here $\boldsymbol{\xi}$ is a parameter vector, i.e. $\boldsymbol{\xi} = \{b\}$ for two-stage integrators, $\boldsymbol{\xi} = \{a, b\}$ for three-stage integrators, and h is a dimensionless step size. We omit here the derivation of $\rho(h, \boldsymbol{\xi})$ as it can be found in [28] and provide the expression for ρ for the families of two- and three-stage

integrators when sampling with 4th and 6th order modified Hamiltonians. This expression if of the form

$$\rho(h, \boldsymbol{\xi}) = \frac{\left(SB_h + C_h\right)^2}{2S(1 - A_h^2)}. \tag{36}$$

The symbols in the right-hand side are as follows. For the 4th order modified Hamiltonian

$$S = \frac{1 + 2h^2 c_{22}}{1 + 2h^2 c_{21}}$$

and for the 6th order modified Hamiltonian

$$S = \frac{1 + 2h^2 c_{22} + 2h^4 c_{43}}{1 + 2h^2 c_{21} + 2h^4 c_{44}},$$

where coefficients c_{ij} depend on the integrator and were derived in Section 3. The h-dependent quantities A_h, B_h, C_h come from the matrix $\tilde{M}_h$ that advances the numerical solution over a single time step in the integration of the harmonic oscillator:

$$\begin{bmatrix} x_{n+1} \\ p_{n+1} \end{bmatrix} = \tilde{M}_h \begin{bmatrix} x_n \\ p_n \end{bmatrix}, \quad \tilde{M}_h = \begin{bmatrix} A_h & B_h \\ C_h & A_h \end{bmatrix}.$$

For two-stage integrators, the matrix $\tilde{M}_h$ has the expression

$$\tilde{M}_h = B\left(b\right) \cdot A\left(\frac{1}{2}\right) \cdot B\left(1 - 2b\right) \cdot A\left(\frac{1}{2}\right) \cdot B\left(b\right),$$

where

$$A(a) = \begin{bmatrix} 1 & ah \\ 0 & 1 \end{bmatrix}, \quad B(b) = \begin{bmatrix} 1 & 0 \\ -bh & 1 \end{bmatrix}$$

correspond to mappings φ_h^A and φ_h^B, respectively. The resulting elements of $\tilde{M}_h$ are then

$$
\begin{aligned}
A_h &= \frac{h^4}{4} b(1 - 2b) - \frac{h^2}{2} + 1, \\
B_h &= -\frac{h^3}{4}(1 - 2b) + h, \\
C_h &= -\frac{h^5}{4} b^2(1 - 2b) + h^3 b(1 - b) - h.
\end{aligned}
$$

Similarly, for the three-stage family we compute

$$\tilde{M}_h = B(b) \cdot A(a) \cdot B(\frac{1}{2} - b) \cdot A(1 - 2a) \cdot B(\frac{1}{2} - b) \cdot A(a) \cdot B(b)$$

16

and obtain

$$
\begin{aligned}
A_h &= \frac{h^6}{4}a^2(2a-1)(1-2b)^2b + \frac{h^4}{4}a\left(1-4b^2-a(1-4b)\right) - \frac{h^2}{2} + 1 \\
B_h &= \frac{h^5}{4}a^2(1-2a)(1-2b)^2 - h^3a(1-a)(1-2b) + h \\
C_h &= \frac{h^7}{4}a^2(1-2a)(1-2b)^2b^2 + \frac{h^5}{2}a(2a(1-b)-1)b(1-2b) + \\
&\quad \frac{h^3}{4}\left(1-2a(1-2b)^2\right) - h.
\end{aligned}
$$

Note that the true Hamiltonian can be recovered by setting the coefficients c_{ij} to zero. Doing so, we obtain exactly the same function derived in [17]

$$
\rho_{\mathrm{HMC}}(h, \boldsymbol{\xi}) = \frac{(B_h + C_h)^2}{2(1 - A_h^2)}. \tag{37}
$$

The coefficients $\boldsymbol{\xi}$ can be found by minimizing the function

$$
\|\rho\|_{(\bar{h})} = \max_{0 < h < \bar{h}} \rho(h, \boldsymbol{\xi}), \tag{38}
$$

where $\bar{h}$ is equal to the number of stages in the integrator [17]. Thus we obtain the parameter $b = 0.238016$ for the two-stage M-BCSS2 integrator derived for sampling with the MHMC methods. We note the difference in value for the coefficient of the original two-stage BCSS2 integrator, $b = 0.21178$, introduced for HMC and obtained by minimizing the function (38) using (37).

Using again the stability analysis from [27] for three-stage integrators, namely enforcing the condition (35), we obtain the coefficients $b = 0.1441153, a = (1-2b)/4(1-3b)$ for the M-BCSS3 integrator for sampling with MHMC.

In Figure 3 $\|\rho_{\mathrm{HMC}}\|_{(\bar{h})}$ from (38) and (37) is plotted as a function of the maximal step size $\bar{h}$ (here normalized to the three-stage schemes, i.e. $\bar{h}_{r\text{-stage}} = r \cdot \bar{h}/3, r = 1, 2, 3$) for the two- and three-stage integrators for the HMC method (dashed lines), and the corresponding function $\|\rho\|_{(\bar{h})}$ from (36) and (38) for the two- and three-stage integrators, derived in this section for sampling with MHMC (solid lines). The functions $\|\rho_{\mathrm{HMC}}\|_{(\bar{h})}$ and $\|\rho\|_{(\bar{h})}$ for the Verlet integrator are also plotted. We note that the upper bound of the expected error in Hamiltonian, or modified Hamiltonian, and thus the error of the method, is lower for integrators derived for MHMC than in the case of the HMC specific integrators, which confirms a better conservation of modified Hamiltonians than true Hamiltonians by symplectic integrators. As follows from Figure 3, the multi-stage integrators derived for HMC and MHMC provide better accuracy than Verlet for step sizes less or equal to a half stability limit of Verlet, i.e.

17

$\bar{h} = 3$, with three-stage integrators being superior to two-stage class. Please notice that $\bar{h}$ in Figure 3 refers to a step size for a three-stage integrator. If Verlet is viewed as a single stage integrator, its half stability limit corresponds to $\bar{h} = 1$. The integrators derived for MHMC guarantee a better accuracy than other integrators for $\bar{h}$ even bigger than 3, which implies their efficiency for longer step sizes compared with Verlet and multi-stage integrators for HMC. A logarithmic scale version of the left-hand graph, shown in the right-hand graph, gives a better insight into the behavior of the functions.
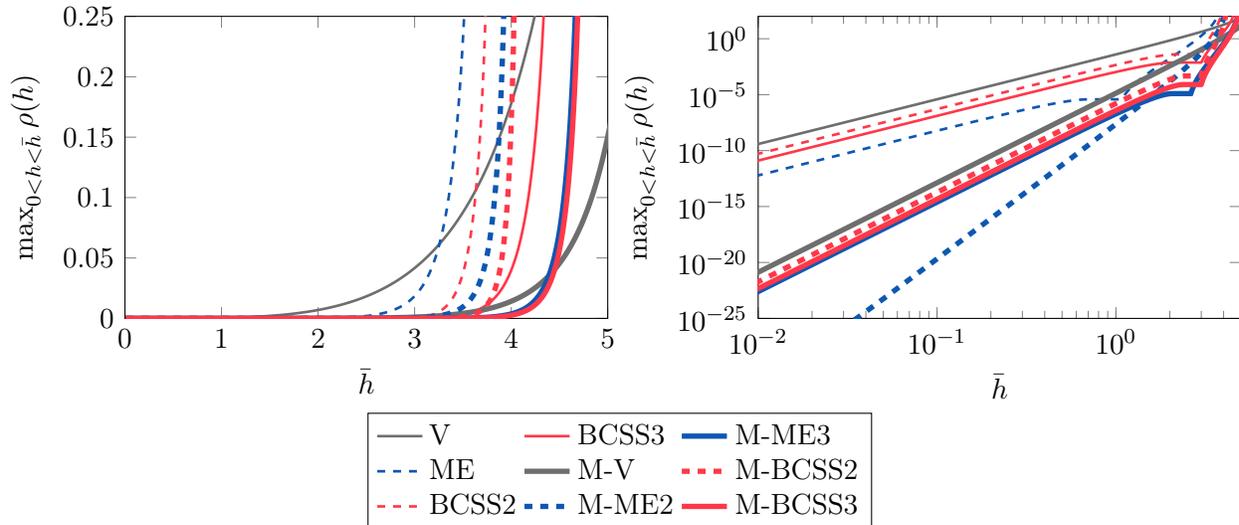
Figure 3: Upper bound for the expected energy error for the two- and three-stage (M-)BCSS, (M-)ME and Verlet integrators for sampling with the true Hamiltonian (dashed) and 4th order modified Hamiltonian (solid). Right-hand graph shows the same functions on a logarithmic scale.

It is important to note that the Verlet integrator has the largest stability interval[1] among other multi-stage integrators, and due to this, care should be taken of the choice of the step size. The stability intervals computed for the considered model problem and for each of the examined integrators are given in Table 1 in terms of the three-stage family. We note that the trends of the stability limit $h_{\max}$ for each integrator are in agreement with the corresponding upper bound functions. Nevertheless, as Figure 3 suggests, the accuracy is degrading with $\bar{h}$ approaching the stability limit. It is the characteristics of the sampling problem (such as the dimension of the system, number of observations, nature of the underlying model) that determine the optimal step size and therefore the integrator which would provide the best

---

[1]The stability interval of an integrator is defined as the largest interval $(0, h_{\max})$ such that the integrator is stable for all $h \in (0, h_{\max})$.

performance.

Integrator	Application	N. of stages	Coefficients	$h_{\max}$	Ref.
Verlet	HMC, MHMC	1	$-$	6.000	[1, 2, 4, 5]
BCSS2	HMC	2	$b = 0.21178$	3.951	[17]
M-BCSS2	MHMC	2	$b = 0.238016$	4.144	this work
ME	HMC	2	$b = 0.193183$	3.830	[16]
M-ME2	MHMC	2	$b = 0.230907$	4.089	this work
M-ME2gen	MHMC	2	$b = 0.23061$	4.0871	this work
BCSS3	HMC	3	$a = (1 - 2b)/4(1 - 3b)$ $b = 0.11888$	4.662	[17, 27]
M-BCSS3	MHMC	3	$a = (1 - 2b)/4(1 - 3b)$ $b = 0.1441153$	4.902	this work
M-ME3	MHMC	3	$a = (1 - 2b)/4(1 - 3b)$ $b = 0.142757$	4.887	this work
M-ME3gen	MHMC	3	$a = 0.355423$ $b = 0.184569$	4.916	this work

Table 1: The splitting integrators for sampling with the true or 4th order modified Hamiltonians developed or tested in this study. Stability limit $h_{\max}$ is computed for problems with a quadratic potential and here presented in terms of the three-stage family.

5. Numerical experiments

In this section we examine the performance of the novel schemes on two benchmark models and compare them against Verlet and integrators suggested in the literature to be used within HMC algorithms.

5.1. Algorithms and performance metrics used

For the numerical experiments we employ the GSHMC [5] and MMHMC [6] algorithms, introduced for sampling in molecular simulation and computational statistics problems, respectively. The main feature of these two algorithms that distinguishes them from other MHMC methods is the momentum update step, called Partial Momentum Monte Carlo (PMMC). In this step, the partial momentum refreshment is combined with a modified Metropolis test. Namely, for the current momentum $\mathbf{p}$ and a noise vector $\mathbf{u} \sim \mathcal{N}(0, \beta^{-1}M)$ a proposal in GSHMC is defined as

$$\begin{aligned}
\mathbf{p}^* &= \cos(\phi)\mathbf{p} + \sin(\phi)\mathbf{u} \\
\mathbf{u}^* &= -\sin(\phi)\mathbf{p} + \cos(\phi)\mathbf{u},
\end{aligned} \tag{39}$$

19

where the parameter $\phi \in (0, \pi/2]$ controls the amount of noise introduced in every iteration. In MMHMC, the momentum update is defined through the noise parameter $\varphi \in (0, 1]$ ($\varphi = \sin^2(\phi)$) as

$$
\begin{aligned}
\mathbf{p}^* &= \quad \sqrt{1 - \varphi}\mathbf{p} + \sqrt{\varphi}\mathbf{u} \\
\mathbf{u}^* &= -\sqrt{\varphi}\mathbf{p} + \sqrt{1 - \varphi}\mathbf{u}.
\end{aligned}
\tag{40}
$$

In the modified Metropolis test, the proposal $(\mathbf{p}^*, \mathbf{u}^*)$ is accepted according to

$$
(\bar{\mathbf{p}}, \bar{\mathbf{u}}) = \begin{cases} (\mathbf{p}^*, \mathbf{u}^*) & \text{with probability } \mathcal{P} \\ (\mathbf{p}, \mathbf{u}) & \text{otherwise,} \end{cases}
\tag{41}
$$

where

$$
\mathcal{P} = \min \left\{ 1, \frac{\exp\left(-(\tilde{H}(\mathbf{x}, \mathbf{p}^*) + \frac{1}{2}(\mathbf{u}^*)^T M^{-1}\mathbf{u}^*)\right)}{\exp\left(-(\tilde{H}(\mathbf{x}, \mathbf{p}) + \frac{1}{2}\mathbf{u}^T M^{-1}\mathbf{u})\right)} \right\}.
\tag{42}
$$

This step can be considered as a standard HMC method in which the vector $\mathbf{x}$ is fixed, the vector $\mathbf{p}$ plays a role of the "position" and the noise vector $\mathbf{u}$ becomes the "conjugate momenta". The extended Hamiltonian

$$
\hat{H}(\mathbf{x}, \mathbf{p}, \mathbf{u}) = \tilde{H}(\mathbf{x}, \mathbf{p}) + \frac{1}{2}\mathbf{u}^\mathsf{T} M^{-1}\mathbf{u}
\tag{43}
$$

defines the extended reference density $\hat{\pi}(\mathbf{x}, \mathbf{p}, \mathbf{u}) \propto \exp(-\beta\hat{H}(\mathbf{x}, \mathbf{p}, \mathbf{u}))$.

Note that in the case of the 4th order modified Hamiltonian with analytical derivatives of the potential available, there is no need for additional integration steps forward and backward nor calculation of gradients within the PMMC step (see details in [6]). Otherwise, additional integration steps need to be performed in order to evaluate the modified Hamiltonian $\tilde{H}(\mathbf{x}, \mathbf{p}^*)$ with the proposed momenta $\mathbf{p}^*$. The noise vectors $\mathbf{u}$ and $\bar{\mathbf{u}}$ are discarded once the momentum has been updated.

The GSHMC and MMHMC methods have been implemented in the in-house software packages MultiHMC-GROMACS [28] and HaiCS (Hamiltonians in Computational Statistics) [6], respectively, both written in C and targeted to computers running UNIX certified operating systems.

To assess the performance of the tested methods we use the following metrics:

- Acceptance rate (AR);

- Effective Sample Size (ESS)—indicating the number of effectively uncorrelated samples out of N collected samples;

- Monte Carlo Standard Error (MCSE)—reflects how much error is in the estimate (7) due to the use of a Monte Carlo method. It is related to ESS as $MCSE = \sqrt{\hat{\sigma}^2/ESS}$, where $\hat{\sigma}^2$ is the sample variance;

- Total distance from the mean, defined as $\|\mathbf{x} - \boldsymbol{\mu}\| = \sum_{d=1}^{D} |\hat{x}_d - \mu_d|$ for the mean $\boldsymbol{\mu}$.

The ESS and MCSE metrics are calculated as proposed in [6] for methods that produce samples that are both correlated and weighted. In all the numerical results we present the relative ESS and MCSE with respect to Verlet; a bigger numbers signal improvements on Verlet.

To make the comparison among the schemes fair, the following issues have been taken into account while producing the numerical results. The step size h and the number of integration steps L were adjusted to the number of stages in the integrator in such a way that the computational cost is equal for all tested integrators, i.e. for an r-stage integrator we set $h_{r\text{-stage}} = rh_{\text{Verlet}}$ and $L_{r\text{-stage}} = L_{\text{Verlet}}/r$. In all plots of numerical results, values of step sizes correspond to Verlet and imply step sizes r times bigger for r-stage integrators. Each individual test has been repeated 10 times; the results reported are obtained by averaging over the 10 runs to reduce statistical errors.

5.2. Benchmarks

Since we are expecting improvements over the Verlet integrator in problems where harmonic oscillations, or quadratic terms, are dominating, we choose to test the novel multi-stage integrators on benchmarks falling in this category. One obvious choice is a molecular simulation application, namely a benchmark describing a realistic coarse-grained system that corresponds to a spider venom toxin in a bilayer. In the following, we shall refer to this systems as *toxin*. Gating-modifier toxins, such as spider venom toxin, are of interest as tools for probing channel-structure functions and the biophysical mechanisms of toxin blockade. The system has been previously studied in [29, 30, 31] and served as a benchmark in [9, 15, 32, 19, 28]. The other benchmark selected for this study consists of sampling from multivariate Gaussian distributions. This is a model highly relevant to computational statistics where many important realistic problems involve distributions that are almost Gaussian or mixtures of Gaussians, see [33, 34, 35, 36] among many others.

We sample these two benchmarks with the GSHMC and MMHMC methods, respectively.

5.2.1. Toxin

Toxin is a coarse-grained system describing a VSTx1 toxin in a POPC bilayer [37]. Four heavy particles are represented on average as one sphere [38, 39], which produces a total number of 7810 particles. In the simulations performed, Coulomb and van der Waals interactions were solved using the shift algorithm. Both potential energies were shifted to 0 kJ mol^{-1} at

the radius of 1.2 nm. Periodic boundary conditions were considered in all directions. The target temperature was chosen to be 310 K. No constraints were defined for this system. The tests were run over a range of time steps h. The total length of all simulations was 20 ns, which was sufficient for equilibration of the system for those choices of time steps that provided a stable integration. Different lengths of MD trajectories L were also tested. For the sake of clarity, in all tests presented here the length of MD trajectories was fixed to 4000 steps for Verlet and scaled correspondigly for two- and three-stage integrators. These values were found to be good choices for GSHMC with different integration schemes [28]. The angle ϕ used for the momentum refreshment was set to 0.2 and the modified Hamiltonian (29) was used for all tests.

We start by measuring the acceptance rates in the GSHMC simulations with different multi-stage integration schemes. A fundamental feature of the GSHMC method is that it maintains very high acceptance rates. This is confirmed in Figure 4 (left), where the effect of various multi-stage integrators and the standard Verlet on the acceptance rates in GSHMC simulations is presented. For small time steps, all integrators guarantee high acceptance rates, but the situation changes as the time step increases and the shorter stability intervals of the different multi-stage methods (cf. Table 1) result in acceptance rates well below those achieved with Verlet. We observe that the integrators derived specifically for sampling with modified Hamiltonians in general show better acceptance rates than their non-modified counterparts. Moreover, the M-BCSS3 integrator provides the best conservation of the modified Hamiltonians and thus the highest acceptance rates. The exception occurs for the biggest time steps tested, where Verlet leads to the highest acceptance rates due to its better stability. All the trends presented in Figure 4 (left) are in a good agreement with the theoretical predictions shown in Figure 3.

The averages of simulated temperatures T calculated over whole GSHMC simulations, were used for evaluating the accuracy provided by the tested integration schemes. From Figure 4 (right) it can be observed that all the methods are able to produce the desired averaged temperature with the exception of the two-stage methods derived for HMC, which for the biggest time steps obtained unrealistically high temperatures as a result of the very low acceptance rates observed during the simulations in these cases.

We shall see next how the integrators impact the sampling efficiency of GSHMC, measured in terms of ESS of the toxin drift to the preferred interfacial location over the equilibration and production periods. Figure 5 presents the relative ESS (i.e. ESS normalized with respect to the values obtained with Verlet) calculated from GSHMC simulations using different integrators and time steps. In the left-hand graph, ESS is calculated for equilibration, during which the toxin is moving towards its desired position in the bilayer membrane. Clearly, M-BCSS3 provided the highest values of ESS and thus the best sampling for all choices of time steps but the last one, for which the acceptance rates decay due to a lower stability
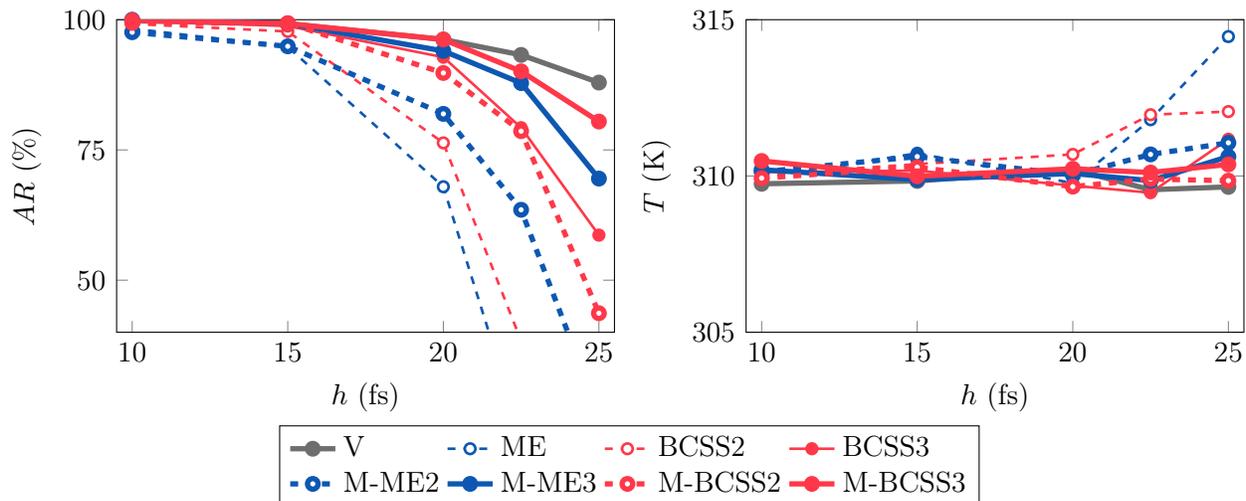
Figure 4: Toxin. Acceptance rates (left) and temperatures (right) as functions of the time step h. Comparison of the two-stage (M-)BCSS2, (M-)ME(2), three-stage (M-)BCSS3, (M-)ME(3), and Verlet integrators.

limit than for Verlet. In the right-hand graph, ESS is calculated for the production phase of the simulations, i.e. once the toxin has reached its equilibrium position. For each time step, ESS has been calculated once all the simulations with the different integrators have reached the equilibrium configuration. As it was the case for equilibration, M-BCSS3 provided the highest values of ESS and thus the best sampling for all choices of time steps but the last one. Importantly, the highest absolute ESS over the range of studied time steps has been also observed with the M-BCSS3 integrator.

5.2.2. Multivariate Gaussian distribution

In this section we test the performance of multi-stage integrators in MMHMC with the modified Hamiltonian (16) used for sampling of multivariate Gaussian distributions.

The goal of this benchmark, proposed in [40], is to sample from a D-dimensional Gaussian $\mathcal{N}(0, \Sigma)$, where the precision matrix Σ^{-1} is generated from a Wishart distribution with D degrees of freedom and the D-dimensional identity scale matrix.

We compare the performance of the standard Verlet integrator, and multi-stage schemes derived for problems with a quadratic potential, summarized in Table 1, for sampling from a multivariate Gaussian distribution of dimensions $D = 100, 1000, 2000$. We have identified a range of reasonable values for the parameters L, h and φ and performed the comparisons for this range. For each MC iteration the number of integration steps is drawn randomly from a uniform distribution on $\{1, \ldots, L\}$ and the step size is uniformly distributed from $(0.8h, 1.2h)$. We report only results obtained with those choices of φ and L that provided
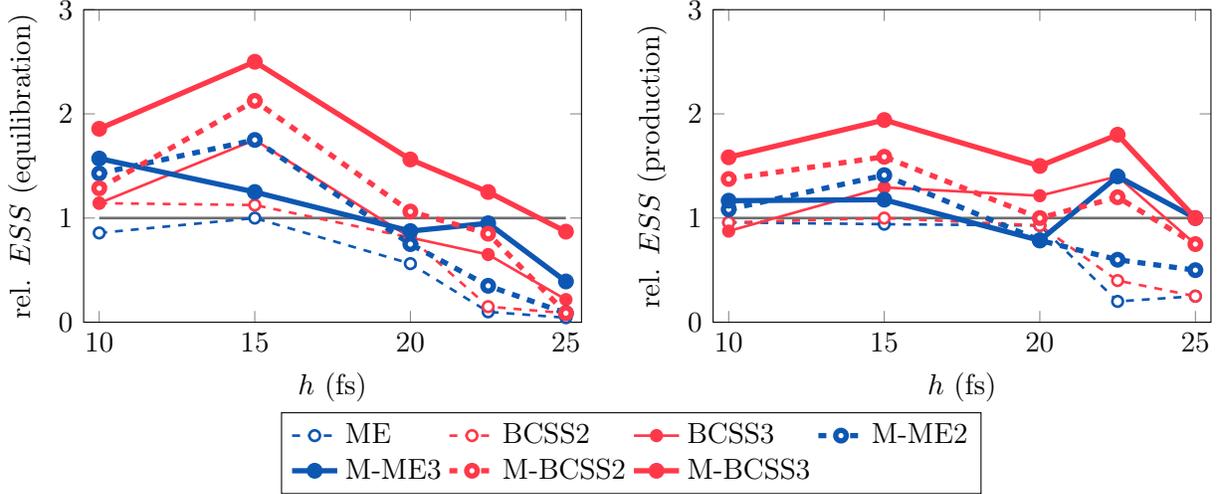
23

Figure 5: Toxin. Relative ESS (with respect to Verlet) for the equilibration (left) and production (right) phases of the simulations. Comparison of the two-stage (M-)BCSS2, (M-)ME(2), three-stage (M-)BCSS3, M-ME3, and Verlet integrators.

the best performance of MMHMC regardless of the choice of integrator. All the experiments here are carried out with the identity mass matrix for MMHMC. The number of production samples generated is 10000 after a warm-up of 2000 discarded samples.

Figure 6 presents the resulting acceptance rates as functions of the step size h. MHMC-specific integrators always lead to higher AR than their counterparts derived for the HMC method. We note that for the smallest dimension ($D = 100$) the Verlet integrator performs better than all two-stage integrators, due to its larger stability limit. However, the new three-stage methods outperform Verlet even for this dimension. For bigger dimensions, which require smaller step sizes, the better conservation of the modified Hamiltonian by all multi-stage integrators (see Figure 3) implies higher acceptance rates. In this case, all the newly derived multi-stage integrators show improvement over Verlet, with M-BCSS3 performing best.

The relative sampling performance with respect to the Verlet integrator, in terms of minimum ESS (top) and maximum MCSE (bottom) over variates for the mean estimate (i.e. $\min_{d \in D} ESS(\hat{x}_d)$ and $\max_{d \in D} MCSE(\hat{x}_d)$), is presented in Figure 7. Values below 1 correspond to cases of sampling efficiency lower than Verlet's and values above 1 correspond to integrators that outperform Verlet. As in the case of the acceptance rates, for the smallest dimension, the Verlet integrator demonstrates a better performance than all two-stage methods. We note that for the smallest step sizes there is no difference among integrators. For bigger step sizes, the novel three-stage integrators outperform all other schemes and im-
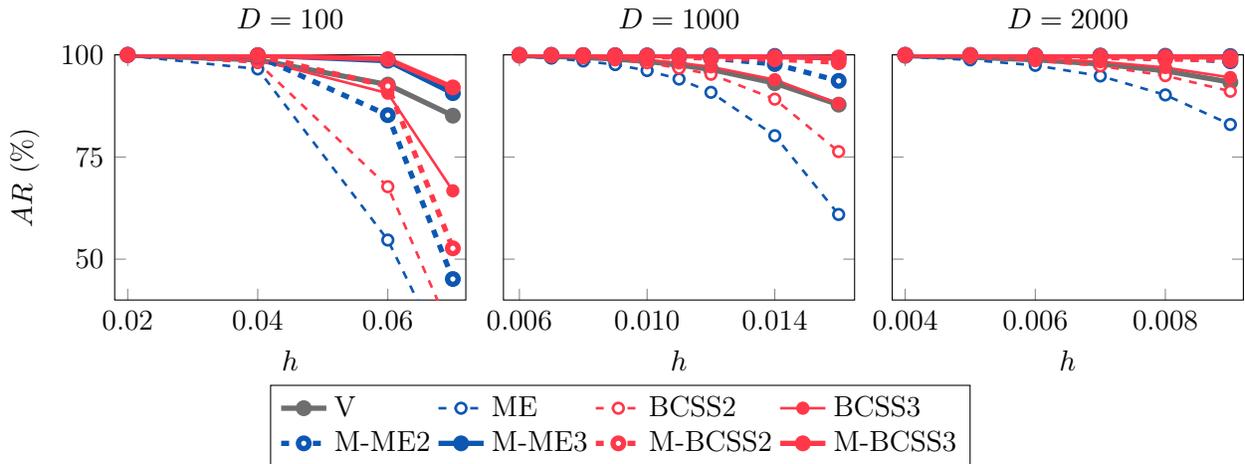
24

Figure 6: Acceptance rates as functions of the step size h for sampling from a D-dimensional Gaussian distribution. Comparison of the two-stage (M-)BCSS2, (M-)ME(2), three-stage (M-)BCSS3, M-ME3, and Verlet integrators.

prove sampling efficiency over Verlet up to 8 and 3 times for ESS and MCSE, respectively. The improvement increases with dimension; therefore we believe that for high dimensional problems the new multi-stage integrators provide crucial ingredients of efficient samplers. In addition, Figure 8 presents a comparison in terms of total distance from the mean of the target distribution; lower values correspond to better accuracy. For high dimensions, the MHMC-specific integrators demonstrate higher accuracy than the commonly used Verlet.

In summary, the tests performed on multivariate Gaussian distribution demonstrated the superiority of the integrators specifically designed for the MHMC methods over both the standard Verlet and sophisticated splitting schemes for the HMC method. The three-stage modified integrators clearly outperform all tested integrators for this problem.

6. Conclusions

We have introduced new multi-stage integrators for enhanced sampling with modified Hamiltonian Monte Carlo (MHMC) methods. The proposed two- and three-stage integration methods provide better conservation of modified Hamiltonians than the Verlet integrator commonly used in MHMC. Each of the methods derived is characterized by its coefficients, which were obtained from the minimization of the (expected) error in modified Hamiltonians introduced by numerical integration. Accordingly, we proposed computationally efficient expressions for modified Hamiltonians of order 4 and 6 for the multi-stage splitting integrating schemes. The new methods were tested and compared with Verlet and also with sophisticated splitting integrators previously suggested for sampling with HMC. The comparisons use
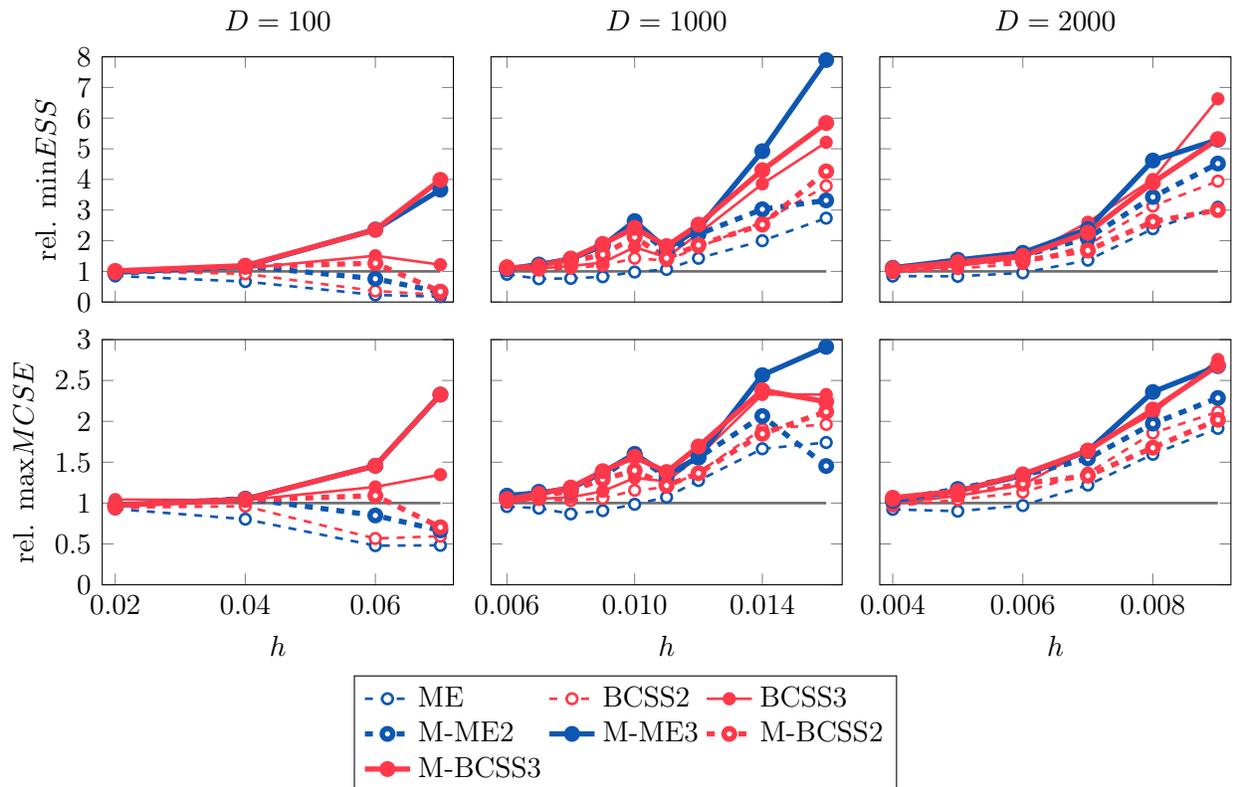
Figure 7: Relative sampling performance with respect to the Verlet integrator, in terms of minimum ESS (top) and maximum MCSE (bottom) over variates, as functions of the step size h for sampling from a D-dimensional Gaussian distribution. Comparison of the two-stage (M-)BCSS2, (M-)ME(2), three-stage (M-)BCSS3, M-ME3, and Verlet integrators.

benchmarks in molecular simulation and computational statistics problems, sampled with the GSHMC [5] and MMHMC [6] methods, respectively. Both method GSHMC and MMHMC belong to the MHMC class and had previously been shown to provide good performance when sampling in applications envisaged here. The tests of the new integration schemes reveal that the novel three-stage integrators lead to an outstanding improvement over the Verlet integrator for problems in which the potential function is (approximately) quadratic. The improvement, which for the tested systems is of up to 8 times, comes both in terms of acceptance rate and sampling efficiency over a range of simulation parameters. For such application problems, all new integrators specifically derived for MHMC methods outperform their counterparts proposed previously for HMC.
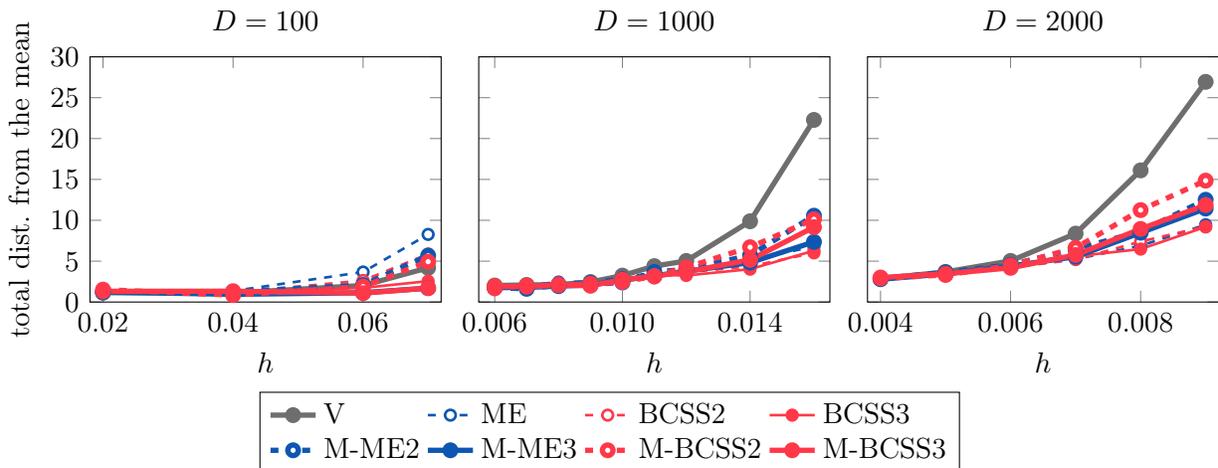
Figure 8: Total distance from the mean as function of the step size h for sampling from a D-dimensional Gaussian distribution. Comparison of the two-stage (M-)BCSS2, (M-)ME(2), three-stage (M-)BCSS3, M-ME3, and Verlet integrators.

Acknowledgements

[1] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Hybrid Monte Carlo, Physics Letters B 195 (2) (1987) 216–222.

[2] R. M. Neal, MCMC using Hamiltonian dynamics, in: Handbook of Markov Chain Monte Carlo, Vol. 2, Chapman & Hall / CRC Press, 2011, pp. 113–162.

[3] H. Kahn, A. W. Marshall, Methods of Reducing Sample Size in Monte Carlo Com-

putations, Journal of the Operations Research Society of America 1 (5) (1953) 263 – 278.

[4] J. A. Izaguirre, S. S. Hampton, Shadow hybrid Monte Carlo: an efficient propagator in phase space of macromolecules, Journal of Computational Physics 200 (2004) 581–604.

[5] E. Akhmatskaya, S. Reich, GSHMC: An efficient method for molecular simulation, Journal of Computational Physics 227 (10) (2008) 4934–4954. doi:http://dx.doi.org/10.1002/andp.19053221004.

[6] T. Radivojević, E. Akhmatskaya, Mix & Match Hamiltonian Monte Carlo, arXiv:1706.04032v1.

[7] T. Radivojević, Enhancing Sampling in Computational Statistics Using Modified Hamiltonians, Ph.D. thesis, UPV-EHU (2016).

[8] 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 (2012) 447–462.

[9] C. L. Wee, M. S. Sansom, S. Reich, E. Akhmatskaya, Improved sampling for simulations of interfacial membrane proteins: Application of generalized shadow hybrid Monte Carlo to a peptide toxin/bilayer system, The Journal of Physical Chemistry B 112 (18) (2008) 5710–5717.

[10] B. Escribano, A. Lozano, T. Radivojević, M. Fernández-Pendás, J. Carrasco, E. Akhmatskaya, Enhancing sampling in atomistic simulations of solid state materials for batteries: a focus on olivine NaFePO4, Theor. Chem. Acc. 136 (43). doi:10.1007/s00214-017-2064-4.

[11] C. R. Sweet, S. S. Hampton, R. D. Skeel, J. A. Izaguirre, A separable shadow Hamiltonian hybrid Monte Carlo method, The Journal of Chemical Physics 131. doi:doi:10.1063/1.3253687.

[12] E. Akhmatskaya, S. Reich, The Targeted Shadowing Hybrid Monte Carlo (TSHMC) Method, in: New Algorithms for Macromolecular Simulation, Lecture Notes in Computational Science and Engineering, Vol. 49, Springer-Verlag, Berlin, 2006, pp. 141–153.

[13] A. M. Horowitz, A generalized guided Monte Carlo algorithm, Physics Letters B 268 (1991) 247–252.

[14] B. Escribano, E. Akhmatskaya, S. Reich, J. M. Azpiroz, Multiple-time-stepping generalized hybrid Monte Carlo methods, Journal of Computational Physics 280 (2015) 1–20. doi:10.1016/j.jcp.2014.08.052.

[15] E. Akhmatskaya, S. Reich, Meso-gshmc: A stochastic algorithm for meso-scale constant temperature simulations, Procedia Computer Science 4 (2011) 1353–1362.

[16] R. I. McLachlan, On the numerical integration of ordinary differential equations by symmetric composition methods, SIAM Journal on Scientific Computing 16 (1) (1995) 151–168.

[17] S. Blanes, F. Casas, J. M. Sanz-Serna, Numerical integrators for the Hybrid Monte Carlo method, SIAM Journal on Scientific Computing 36 (4) (2014) A1556–A1580.

[18] W.-L. Chao, J. Solomon, D. L. Michels, F. Sha, Exponential Integration for Hamiltonian Monte Carlo, in: Proceedings of the 32nd International Conference on Machine Learning, Vol. 37, 2015.

[19] M. Fernández-Pendás, E. Akhmatskaya, J. M. Sanz-Serna, Adaptive multi-stage integrators for optimal energy conservation in molecular simulations, Journal of Computational Physics 327 (2016) 434–449. doi:10.1016/j.jcp.2016.09.035.

[20] N. Bou-Rabee, J. M. Sanz-Serna, Geometric integrators and the Hamiltonian Monte Carlo method, Acta Numerica (to appear).
URL https://arxiv.org/abs/1711.05337

[21] R. D. Skeel, D. J. Hardy, Practical construction of modified Hamiltonians, SIAM Journal on Scientific Computing 23 (4) (2001) 1172–1188.

[22] R. D. Engle, R. D. Skeel, M. Drees, Monitoring energy drift with shadow Hamiltonians, Journal of Computational Physics 206 (2005) 432–452. doi:doi:10.1016/j.jcp.2004.12.009.

[23] P. C. Moan, J. Niesen, On an asymptotic method for computing the modified energy for symplectic methods, Dynamical Systems 34 (3) (2014) 1105–1120.

[24] J. M. Sanz-Serna, M. Calvo, Numerical Hamiltonian Problems, Chapman and Hall, London, 1994.

[25] I. P. Omelyan, I. M. Mryglod, R. Folk, Construction of high-order force-gradient algorithms for integration of motion in classical and quantum systems, Physical Review E 66 (2).

[26] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration, 2nd Edition, Springer-Verlag, 2006.

[27] C. M. Campos, J. M. Sanz-Serna, Palindromic 3-stage splitting integrators, a roadmap, Journal of Computational Physics 346 (2017) 340–355.

[28] E. Akhmatskaya, M. Fernández-Pendás, T. Radivojević, J. M. Sanz-Serna, Adaptive Splitting Integrators for Enhancing Sampling Efficiency of Modified Hamiltonian Monte Carlo Methods in Molecular Simulation, Langmuir 33 (42) (2017) 11530–11542. `doi:10.1021/acs.langmuir.7b01372`.

[29] K. J. Swartz, R. MacKinnon, An inhibitor of the Kv2.1 potassium channel isolated from the venom of a Chilean tarantula, Neuron 15 (1995) 941–949.

[30] Y. Li-Smerin, K. J. Swartz, Gating modifier toxins reveal a conserved structural motif in voltage-gated Ca^{2+} and K^+ channels, Proceedings of the National Academy of Sciences of the United States of America 95 (15) (1998) 8585–8589.

[31] S. Y. Lee, R. MacKinnon, A membrane-access mechanism of ion channel inhibition by voltage sensor toxins from spider venom, Nature 430 (2004) 232–235.

[32] M. Fernández-Pendás, B. Escribano, T. Radivojević, E. Akhmatskaya, Constant pressure hybrid monte carlo simulations in gromacs, Journal of Molecular Modeling 20 (12) (2014) 1–10.

[33] N. Tripuraneni, M. Rowland, Z. Ghahramani, R. Turner, Magnetic hamiltonian monte carlo, in: Proceedings of the 34th International Conference on Machine Learning, 2017. URL `arXiv:1607.02738v2`

[34] A. Nishimura, D. B. Dunson, Geometrically Tempered Hamiltonian Monte Carlo, arXiv:1604.00872v2. URL `arXiv:1604.00872v2`

[35] S. Lan, T. Bui-Thanh, M. Christie, M. Girolami, Emulation of Higher-Order Tensors in Manifold Monte Carlo Methods for Bayesian Inverse Problems, Journal of Computational Physics 308 (2016) 81–101. `doi:doi.org/10.1016/j.jcp.2015.12.032`.

[36] C. E. Rasmussen, Evaluation of gaussian processes and other methods for non-linear regression, Ph.D. thesis, Department of Computer Science, University of Toronto (1996).

[37] H. J. Jung, J. Y. Lee, S. H. Kim, Y.-J. Eu, S. Y. Shin, M. Milescu, K. J. Swartz, J. I. Kim, Solution Structure and Lipid Membrane Partitioning of VSTx1, an Inhibitor of the KvAP Potassium Channel, Biochemistry 44 (16) (2005) 6015–6023. `doi:10.1021/bi0477034`.

[38] E. J. Wallace, M. S. P. Sansom, Carbon nanotube/detergent interactions via coarse-grained molecular dynamics, Nano Lett. 7 (7) (2007) 1923–1928. doi:10.1021/nl070602h.

[39] A. Y. Shih, A. Arkhipov, P. L. Freddolino, K. Schulten, Coarse grained protein-lipid model with application to lipoprotein particles, J. Phys. Chem. B 110 (8) (2006) 3674–3684. doi:10.1021/jp0550816.

[40] M. D. Hoffman, A. Gelman, The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, Journal of Machine Learning Research 15 (2014) 1593—1623.