Erratum

Erratum to "A comparison of generalized hybrid Monte Carlo methods with and without momentum flip" [J. Comput. Phys. 228 (2009) 2256–2265]

Elena Akhmatskaya [a], Nawaf Bou-Rabee [b], Sebastian Reich [c,*]

[a] Fujitsu Laboratories of Europe Ltd (FLE), Hayes Park Central, Hayes End Road, Hayes UB4 8FE, United Kingdom
[b] Department of Mathematics, Freie Universität Berlin, Arnimallee 2-6, 14195 Berlin, Germany
[c] Universität Potsdam, Institut für Mathematik, Am Neuen Palais 10, D-14469 Potsdam, Germany

This note is to point out an error in the theory part of the publication [1]. We will follow the notations and definitions of [1] unless stated otherwise.

Contrary to what is claimed in Section 2.2 of [1], the modified Metropolis–Hastings acceptance criterion (Eq. (6) in [1].) does not satisfy a modified detailed balance condition for the choice of a linear involution $\mathcal{F}$ and for the choice of $\delta(\Gamma', \Gamma)$ suggested in the paper. We will produce the mathematical reasons below, which were kindly pointed out to us by Tony Lelievre and Gabriel Stoltz. We will also produce numerical evidence that despite not being able to preserve a (modified) detailed balance condition, the proposed GHMC method without momentum flip produces better results in terms of sampling accuracy than corresponding results from a method without any Metropolis–Hasting correction step, i.e., a standard discretization of the underlying second-order Langevin dynamics. We also find a good agreement between GHMC with momentum flip (rigorous sampling from canonical distribution) and without momentum flip in terms of sampling accuracy for acceptance rates above 75–80%.

As correctly stated in Section 2.2 of [1], the induced proposal distribution of making a transition from Γ to $\Gamma' \neq \Gamma$ is given by

$$A(\Gamma'|\Gamma) = P(\Gamma'|\Gamma)r(\Gamma', \Gamma) = \rho_{\text{stat}}(\Gamma)^{-1}\delta(\Gamma', \Gamma),$$

which satisfies

$$A(\mathcal{F}\Gamma|\mathcal{F}\Gamma') = P(\mathcal{F}\Gamma|\mathcal{F}\Gamma')r(\mathcal{F}\Gamma, \mathcal{F}\Gamma') = \rho_{\text{stat}}(\Gamma')^{-1}\delta(\Gamma', \Gamma)$$

since

$$\delta(\Gamma', \Gamma) = \delta(\mathcal{F}\Gamma, \mathcal{F}\Gamma') \tag{1}$$

and $\rho_{\text{stat}}(\Gamma) = \rho_{\text{stat}}(\mathcal{F}\Gamma)$. However, one has to also take into account the probability of staying at a given state Γ, i.e., the probability of rejections under the Metropolis–Hastings criterion (see, e.g., [5]). This probability distribution is given by

$$A(\Gamma|\Gamma) = \delta_\Gamma(\Gamma) \int P(\Gamma'|\Gamma)(1 - r(\Gamma', \Gamma))\, d\Gamma' = \delta_\Gamma(\Gamma)\left[1 - \int \rho_{\text{stat}}(\Gamma)^{-1}\delta(\Gamma', \Gamma)\, d\Gamma'\right],$$

where δ_Γ denotes the Dirac delta function centered about Γ. Similarly, we also obtain

$$A(\mathcal{F}\Gamma|\mathcal{F}\Gamma) = \delta_{\mathcal{F}\Gamma}(\mathcal{F}\Gamma) \int P(\mathcal{F}\Gamma'|\mathcal{F}\Gamma)(1 - r(\mathcal{F}\Gamma', \mathcal{F}\Gamma))\, d\Gamma' = \delta_\Gamma(\Gamma)\left[1 - \int \rho_{\text{stat}}(\Gamma)^{-1}\delta(\Gamma, \Gamma')\, d\Gamma'\right],$$

where $\delta_{\mathcal{F}\Gamma}$ denotes now the Dirac delta function centered about $\mathcal{F}\Gamma$. The modified detailed balance condition (Eq. (4) in [1].) implies that $A(\Gamma|\Gamma)$ and $A(\mathcal{F}\Gamma|\mathcal{F}\Gamma)$ should be equal. However that is obviously not the case unless

$$\int \delta(\Gamma', \Gamma)\,\mathrm{d}\Gamma' = \int \delta(\Gamma, \Gamma')\,\mathrm{d}\Gamma',$$

which does not hold for the choice of $\delta(\Gamma', \Gamma)$ considered in [1] on p. 2258 for which we have (1) instead.

We now provide an initial study for the practical impact of the failure of exact conservation of detailed balance under the GHMC method without momentum flip. A more detailed study will be carried out in a separate publication. We investigate two simple model systems.

First we consider the harmonic oscillator

$$\dot{q} = p, \quad \dot{p} = -q.$$

We implement the GHMC method of [1] with $L = 1$ steps in the molecular dynamics (MD) part, step-size $h \in [0.1, 1.9]$, angle

$$\phi = \sqrt{2\gamma h}, \quad \gamma = 0.05,$$

in the momentum refreshment part, and inverse temperature $\beta = 1$. We compare computed averages of q^2 along samples generated from (i) GHMC without momentum flip, (ii) GHMC with momentum flip, and (iii) GHMC without any Metropolis test in the MD part, i.e., a standard numerical time-stepping method for the underlying Langevin equations. The results can be found in Figs. 1 and 2. One finds that GHMC without momentum flip leads to visible sampling errors for $h \geqslant 1.7$, which are comparable to those for the unadjusted time-stepping method (no Metropolis test) at step-sizes $h \approx 0.8-1.2$. Also note that $h = 2$ is the stability limit for the Störmer–Verlet method. As can be seen from Fig. 1, the decrease in acceptance rate is slightly faster for GHMC without flip resulting in an optimal step-size of about $h = 1.5$, where AR ≈ 0.75. Here optimality is defined as the step-size for which the product of AR and h is largest. The optimal step-size for GHMC with momentum flip is also found to be about $h = 1.5$ and, hence, GHMC with and without flip perform essentially equally well in terms of efficiency and accuracy for this test problem and outperform the unadjusted scheme. This observation is confirmed by numerically computed probability density functions for all three methods at $h = 1.6$. See Fig. 2. We finally assess the impact of momentum reversal upon rejection on the dynamic properties such as velocity autocorrelation functions (ACF) [3]. Results are displayed in Fig. 3 for three different step-sizes $h = 0.7$, 1.0, and 1.5. It can be seen that even for the relatively small step-size of $h = 0.7$ and a low rejection rate the GHMC method with momentum flip changes the velocity ACF significantly (faster decay; i.e. an effectively stronger diffusion). The GHMC method without momentum flip, on the other hand, stays close to the results from the unadjusted scheme for $h = 0.7$ and 1.0 and produces qualitative agreement for $h = 1.5$. Note that neither the unadjusted nor the Metropolis corrected schemes can reproduce the phase of the oscillations correctly for such large step-sizes.

The second example is more challenging. We consider a one-dimensional periodic chain of $N = 10$ particles with positions q_i, $i = 1, \ldots, N$, and mass $m = 0.5$. The nearest neighbor particles interact through a pair-wise Lennard-Jones potential
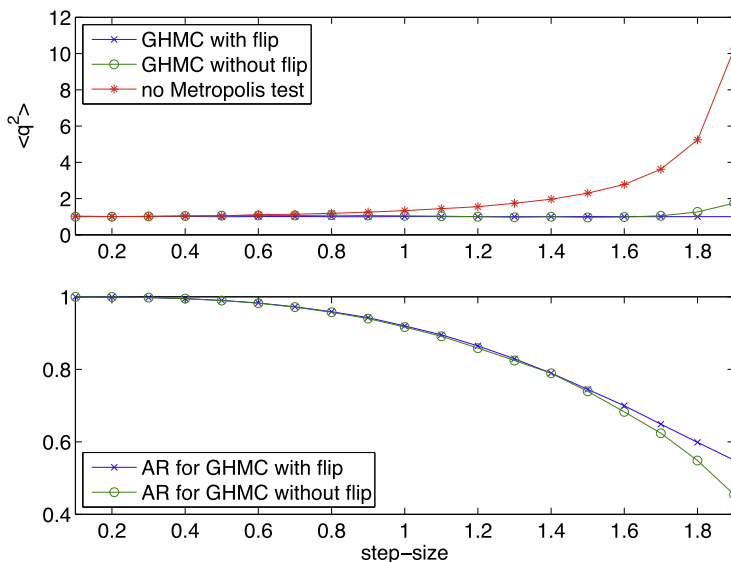

Fig. 1. Displayed are the computed expectation value $\langle q^2 \rangle$ for the three different implementation of GHMC. The analytic value is $\langle q^2 \rangle = 1$. Acceptance rates (AR) are provided for the GHMC methods with and without momentum flip. Here AR = 1 corresponds to a 100% acceptance rate. Results are displayed for step-sizes $h \in [0.1, 1.9]$. While the GHMC method without momentum flip leads to noticeable sampling errors for $h \geqslant 1.7$, the errors are significantly smaller than those for the unadjusted time-stepping method (no Metropolis test).
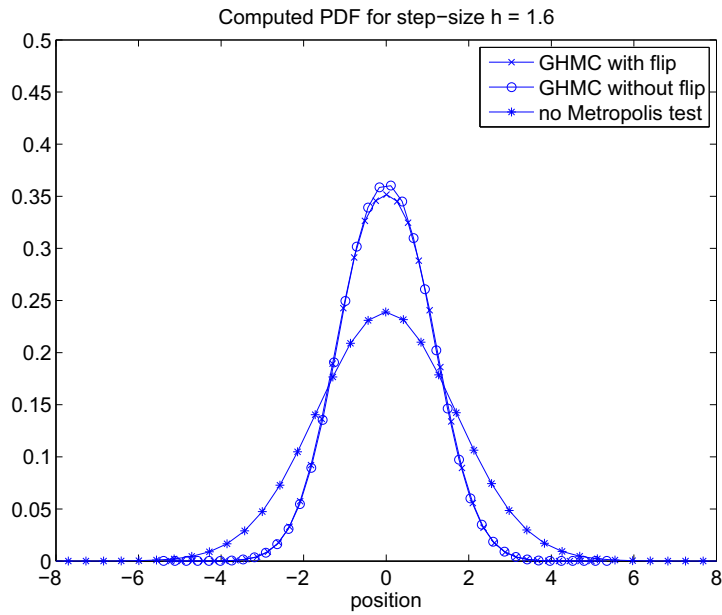
Fig. 2. Displayed are the computed probability density functions for step-size $h = 1.6$ and for the three different implementations of GHMC. The agreement between GHMC with and without momentum flip is remarkable compared to the large distortion in the distribution for the unadjusted scheme.
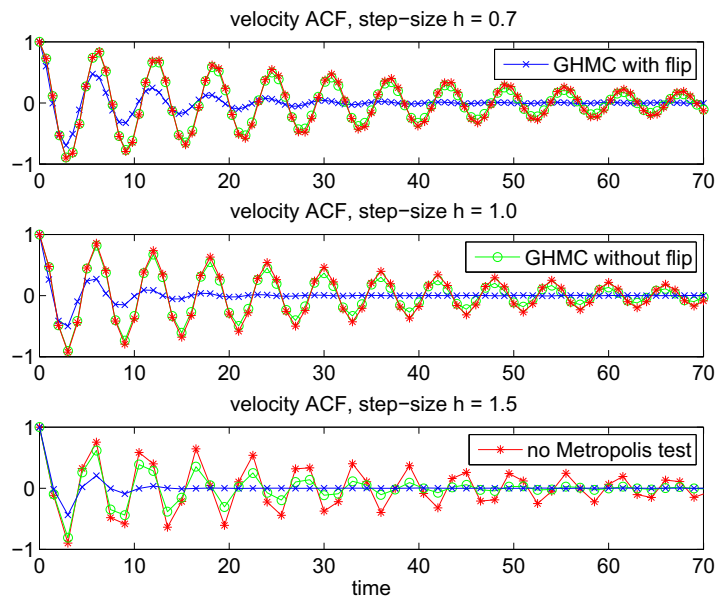

Fig. 3. Displayed are the computed velocity autocorrelation functions (ACF) for the harmonic oscillator for all three methods and for three different step-sizes ($h = 0.7$, 1.0, and 1.5). The results demonstrate the large impact of the momentum flip on dynamic properties, while GHMC without momentum flip qualitatively reproduces the results from the unadjusted method.

$$V(r) = \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}, \quad \sigma = (1/2)^{1/6},$$

with $r = q_{j+1} - q_j$, $j = 1, \ldots, N$ and the convention $q_{N+1} = q_1$. The length of the domain is $l = 10$ and the system is simulated at inverse temperature $\beta = 1$. We compute the average of

$$S = \frac{1}{N}\left(\sum_{i=1}^{N}(q_{i+1} - q_i - 1)^2\right)^{1/2} \tag{2}$$
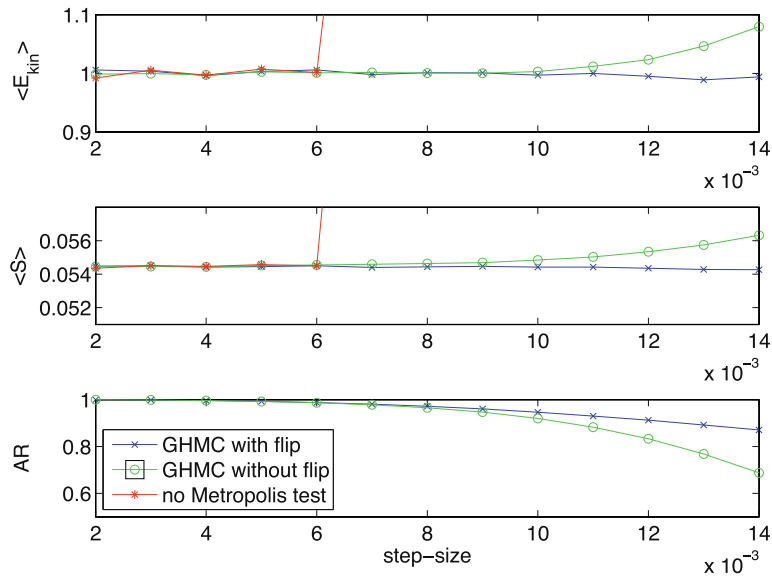
Fig. 4. Displayed are the computed expectation value of the kinetic energy per degree of freedom and the computed expectation value $\langle S \rangle$ of (2) for the three implementations of the GHMC method. The exact reference value for the kinetic energy per degree of freedom is one. Acceptance rates (AR) are provided for the GHMC methods with and without momentum flip. The implementation without any Metropolis test resulted in unstable simulations for all $h > 0.006$. A significant systematic drift in the computed expectation values can be observed for the GHMC method without momentum flip for step-sizes $h > 0.012$.

along samples generated from (i) GHMC without momentum flip, (ii) GHMC with momentum flip, and (iii) GHMC without any Metropolis test in the MD part (unadjusted time-stepping scheme). We again use $L = 1$ in the MD part of GHMC. The step-sizes satisfy $h \in [0.002, 0.014]$ and $\phi = \sqrt{2\gamma h}$ with $\gamma = 0.2$.

Numerical simulations without any Metropolis correction lead to numerical blow-up for $h > 0.006$ in simulations over a time interval $t \in [0, 40000]$. This is due to the fact, that the Lennard-Jones force field is not globally Lipschitz continuous. In fact, numerical blow-up is possible for all step-sizes $h > 0$ but with rapidly decaying probability as $h \to 0$. See [4,5] for a theoretical study of such instability phenomena for numerical approximations to Langevin dynamics. Simulations with and without momentum flip, on the other hand, lead to stable simulations with a decreasing acceptance rate as the step-size increases. The decrease in acceptance rate is faster for GHMC without flip resulting in an optimal step-size of about $h = 0.012/0.013$, where AR ≈ 0.8. As before, optimality is defined as the step-size for which the product of AR and h is largest. As can also be seen in Fig. 4, GHMC without momentum flip reproduces expectation values within 5% of their reference value for the kinetic energy per particle and the observable S at $h = 0.012/0.013$. The optimal step-size for the GHMC method with momentum flip is $h = 0.014$ and, hence, GHMC with momentum flip is slightly more efficient and accurate in terms of computing expectation values than GHMC without momentum flip for this test problem.

We summarize our findings. Provided that the acceptance rate in a GHMC implementation without momentum flip does not drop below a certain threshold (between 75% and 80% in our two experiments, which should provide a good benchmark value), GHMC without flip improves the accuracy as well as the stability of a time-stepping method without Metropolis correction. In particular, numerical stability is achieved for non-global Lipschitz force fields.

On the other hand, strict sampling from the canonical distribution requires the application of a momentum flip upon rejection which also improves acceptance rates as already noted in [1] and, hence, GHMC with momentum flip is more efficient and accurate than GHMC without momentum flip in terms of pure sampling. However, momentum flips interfere strongly with the dynamics of the underlying Langevin equations, as demonstrated in [1], and dynamic properties need to be computed by other means (e.g., by computing finite time trajectories originating from Monte Carlo samples). Alternatively, one can increase the acceptance rate in the molecular dynamics part of GHMC by sampling with respect to a high-order shadow energy. See [2] for a description of the generalized shadow hybrid Monte Carlo (GSHMC) method and [6] for numerical results from a membrane protein simulation.

References

[1] E. Akhmatskaya, N. Bou-Rabee, S. Reich, A comparison of generalized hybrid Monte Carlo methods with and without momentum flip, J. Comput. Phys. 228 (2009) 2256–2265.

[2] E. Akhmatskaya, S. Reich, GSHMC: an efficient method for molecular simulations, J. Comput. Phys. 227 (2008) 4934–4954.

[3] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Clarendon Press, Oxford, 1987.

[4] J.C. Mattingly, A.M. Stuart, D.J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerated noise, Stoch. Process. Appl. 101 (2002) 185–232.

[5] G.O. Roberts, R.L. Tweedie, Exponential convergence of Langevin distributions and their discrete approximations, Bernoulli 2 (1996) 341–363.

[6] C.L. Wee, M.S.P. Sansom, S. Reich, E. Akhmatskaya, Improved sampling for simulations of interfacial membrane proteins: applications of generalized shadow hybrid Monte Carlo to a peptide toxin/bilayer system, Phys. Chem. B 112 (2008) 5710–5717.