METROPOLIS INTEGRATION SCHEMES FOR SELF-ADJOINT DIFFUSIONS

NAWAF BOU-RABEE*, ALEKSANDAR DONEV†, AND ERIC VANDEN-EIJNDEN‡

Abstract. We present explicit methods for simulating diffusions whose generator is self-adjoint with respect to a known (but possibly not normalizable) density. These methods exploit this property and combine an optimized Runge-Kutta algorithm with a Metropolis-Hastings Monte-Carlo scheme. The resulting numerical integration scheme is shown to be weakly accurate at finite noise and to gain higher order accuracy in the small noise limit. It also permits to avoid computing explicitly certain terms in the equation, such as the divergence of the mobility tensor, which can be tedious to calculate. Finally, the scheme is shown to be ergodic with respect to the exact equilibrium probability distribution of the diffusion when it exists. These results are illustrated on several examples including a Brownian dynamics simulation of DNA in a solvent. In this example, the proposed scheme is able to accurately compute dynamics at time step sizes that are an order of magnitude (or more) larger than those permitted with commonly used explicit predictor-corrector schemes.

Key words. explicit integrators; Brownian dynamics with hydrodynamic interactions; Metropolis-Hastings algorithm; small noise limit; predictor-corrector schemes; DNA simulations; ergodicity; fluctuation-dissipation theorem;

AMS subject classifications. Primary, 65C30; Secondary, 65C05, 60J05

1. Introduction.

This paper is concerned with the numerical integration of diffusions of the type:

$$\boxed{dY = \underbrace{-M(Y)DU(Y)dt}_{\text{deterministic drift}} + \underbrace{kT\,\mathrm{div}\,M(Y)dt + \sqrt{2kT}B(Y)dW}_{\text{heat bath}}} \tag{1.1}$$

where we have introduced the following.

$Y(t)\in\mathbb{R}^n$	state of the system
$U:\mathbb{R}^n\to\mathbb{R}$	potential energy
$DU:\mathbb{R}^n\to\mathbb{R}^n$	force
$M:\mathbb{R}^n\to\mathbb{R}^{n\times n}$	mobility matrix
$(\mathrm{div}\,M)_i=\sum_{j=1}^n \partial M_{ij}/\partial x_j$	divergence of mobility matrix
$B:\mathbb{R}^n\to\mathbb{R}^{n\times n}$	noise coefficient matrix
$W(t)\in\mathbb{R}^n$	Brownian motion
kT	temperature factor

Let $B(x)^T$ denote the transpose of the real matrix $B(x)$. Assuming that

$$M(x)=B(x)B(x)^T \qquad \text{for all } x\in\mathbb{R}^n, \tag{1.2}$$

this dynamics defines a Markov process whose generator L is *self-adjoint* with respect to the following density

$$\nu(x)=\exp\left(-\frac{1}{kT}U(x)\right). \tag{1.3}$$

*Department of Mathematical Sciences, Rutgers University Camden, 311 N 5th Street, Camden, NJ 08102, USA, (nawaf.bourabee@rutgers.edu). The work of N. B-R. was supported by NSF grant DMS-1212058.

†Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1185, (donev@courant.nyu.edu).

‡Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1185, (eve2@cims.nyu.edu).

Indeed since the action of L on any suitable test function $f(x)$ can be written as

$$(Lf)(x) = kT\nu^{-1}(x)\operatorname{div}(\nu(x)M(x)Df(x))\,, \tag{1.4}$$

it follows from integration by parts that

$$\langle Lf,g\rangle_\nu = \langle f,Lg\rangle_\nu \qquad \text{for all suitable test functions } f(x) \text{ and } G_h(x)\,, \tag{1.5}$$

where $\langle\cdot,\cdot\rangle_\nu$ denotes an L^2-inner product weighted by the density $\nu(x)$:

$$\langle f,g\rangle_\nu = \int_{\mathbb{R}^n} f(x)G_h(x)\nu(x)dx\,.$$

This property implies that the diffusion is *ν-symmetric* [57], in the sense that

$$\nu(x)p_t(x,y) = \nu(y)p_t(y,x) \qquad \text{for all } t>0\,, \tag{1.6}$$

where $p_t(x,y)$ denotes the transition density of $Y(t)$ given that $Y(0)=x$. In fact, (1.5) corresponds to an infinitesimal version of (1.6). We stress that (1.6) can hold even if $\nu(x)$ is not normalizable: in this case, the solution to (1.1) reaches no statistical steady state, i.e., it is not ergodic. If, however, the density (1.3) is normalizable,

$$Z = \int_{\mathbb{R}^n} \exp\left(-\frac{1}{kT}U(x)\right)dx < \infty\,, \tag{1.7}$$

then the diffusion (1.1) is ergodic with respect to $\nu(x)/Z$. This normalized density is called the equilibrium probability density of the diffusion and is also referred to as the *Boltzmann-Gibbs density*. When (1.7) holds, the dynamics observed at equilibrium is time-reversible, and the property (1.6) is referred to as the *detailed balance condition* in the physics literature. Our main purpose here is to exploit the properties above, in particular (1.6), to design stable and accurate numerical integrators for (1.1) that work even in out-of-equilibrium situations when the density $\nu(x)$ in (1.3) is not normalizable.

The importance of this objective stems from the wide range of applications where diffusions of the type (1.1) serve as dynamical models. For example, (1.1) is used to model the evolution of coarse-grained molecular systems in regimes where the details of the microscopic interactions can be represented by a heat bath. Multiscale models of polymers in a Stokesian solvent fit this category, a context in which (1.1) is referred to as Brownian dynamics (BD) with hydrodynamic interactions [22, 25, 86, 108]. BD has been used to quantitatively simulate the non-equilibrium dynamics of biopolymers in dilute solutions [63, 99], and to validate and clarify experimental findings for bacteriophage DNA dynamics in extensional, shear, and mixed planar flows [64, 46, 45, 50, 97, 98]. The dynamics of DNA molecules in confined solutions with complex geometries has also been studied using BD to guide the design of microfluidic devices to manipulate these molecules [12, 51, 52, 53, 109, 110, 35, 34, 33, 36, 29, 16, 111].

Because of the relevance of (1.1) to BD, this is also the context in which most work on the development of numerical schemes to simulate this equation has been devoted [86, 47]. Perhaps the most widely used among these is the so-called Fixman scheme, which is an explicit predictor-corrector scheme that also has the advantage that it avoids the explicit computation of the divergence of the mobility matrix [25]. Unfortunately, explicit BD time integrators like the Fixman scheme are

often insufficient in applications because the interplay between the drift and the noise terms in (1.1) can induce numerical instabilities or artifacts such as systematic drifts [71, 90]. This is a well known problem with explicit discretizations of nonlinear diffusions [102, 39, 84, 38, 48], and it is especially acute if the potential force in (1.1) is stiff, which is typically the case in applications. The standard solution to this problem is to introduce implicitness. In a sequence of works, this strategy was investigated to deal with the steep potentials that enforce finite extensibility of bond lengths in bead-spring models [26, 37, 49, 100, 41]. The leading semi-implicit predictor-corrector scheme emerging from this effort is numerically stable at much larger time step sizes than explicit predictor-corrector schemes but it requires a nonlinear solve at every step, which is time-consuming and raises convergence questions about the methods used to solve this nonlinear system that are not easy to answer, in general.

In this paper, following the strategy introduced in [10], we adopt a probabilistic approach based on the ν-symmetry property in (1.6) to build stable explicit integrators for (1.1) that avoid implicitness and do not assume a specific form of the potential forces. Specifically, we propose an integrator for (1.1) with the following properties.

- **(P1)** it samples the exact equilibrium probability density of the SDE when this density exists (i.e., when $\nu(x)$ in (1.3) is normalizable);
- **(P2)** it generates a weakly accurate approximation to the solution of (1.1) at constant temperature;
- **(P3)** it acquires higher order accuracy in the small noise limit, $kT \to 0$; and,
- **(P4)** it avoids computing the divergence of the mobility matrix.

We stress that Properties (P2)-(P4) hold even when the density $\nu(x)$ in (1.3) is not normalizable, i.e., (1.1) is out-of-equilibrium and not ergodic with respect to any invariant distribution. Thus, even though the scheme we propose involves a Monte-Carlo step, its aim and philosophy are very different from Monte-Carlo methods where a discretized version of (1.1) is used as a proposal step but whose only goal is to sample a target distribution with no concern for the dynamics of (1.1) [82, 32, 95, 17, 40, 56, 69, 3, 2, 66]. Compared to the method proposed in [10], the main novelty of the scheme introduced here stems from Properties (P3) and (P4). Indeed, to achieve (P3) we need to control the rate of rejections in the Monte Carlo step in the small noise limit, $kT \to 0$, which is nontrivial because the density $\nu(x)$ in (1.3) that enters the key relation (1.6) becomes singular in this limit. Similarly, (P4) constrains the type of proposal moves we can use in the scheme, and this property is important in situations where the mobility matrix does not have an explicit expression, e.g., when the solvent is confined [29]. In the sequel, we use a BD model of DNA in a solvent to show that the new scheme achieves the same accuracy as explicit predictor-corrector schemes with time step sizes that are an order of magnitude (or more) larger. Beside BD, this scheme should also be useful in the simulation of polymer conformational transitions in electrophoretic flows [60, 88, 61, 105, 42, 43], in other contexts where the effective dynamics of a set of coarse-grained variables satisfies (1.1) [18, 78, 77, 19, 65], in applications to Bayesian inference [28], etc.

Organization of the paper. The new scheme, with a structure that immediately shows that it satisfies Property (P4) is introduced in §2. In §3, we present numerical examples with comparisons to the Fixman algorithm. A proof of the ergodicity Property (P1) of the scheme is provided in §4. The weak accuracy Property (P2) is proven in §5 using (1.6). The behavior of the scheme in the small noise limit, specifically Property (P3), is investigated in §6. A conclusion is given in §7.

2. The integrator and its main properties. Following [10], the scheme introduced in this paper combines a proposal step obtained via time-discretization of (1.1) with an accept/reject Monte-Carlo step. The detailed algorithm is given below in terms of vector and matrix-valued variables $G_h:\mathbb{R}^n\to\mathbb{R}^n$ and $B_h:\mathbb{R}^n\to\mathbb{R}^{n\times n}$, respectively, whose explicit forms will be specified shortly in such a way that Properties (P1)–(P4) are met.

ALGORITHM 2.1. *Given the current position X_0 at time t the algorithm proposes an updated position $X_1^\star$ at time $t+h$ for some time step size $h>0$ via*

$$\begin{cases} X_1^\star = \tilde{X}_1 + hG_h(\tilde{X}_1) + (\tilde{X}_1 - X_0)\,, \\ \tilde{X}_1 = X_0 + \sqrt{\frac{h}{2}}B_h(X_0)\xi\,. \end{cases} \tag{2.1}$$

Here $\xi\in\mathbb{R}^n$ denotes a Gaussian random vector with mean zero and covariance $\mathbb{E}(\xi_i\xi_j)=kT\delta_{ij}$. The proposal move *$X_1^\star$ is then accepted or rejected by taking as actual update for the position at time $t+h$ the value*

$$X_1 = \gamma X_1^\star + (1-\gamma)X_0\,, \tag{2.2}$$

where γ is a Bernoulli random variable which assumes the value 1 with probability $\alpha_h(X_0,X_1^\star)$ and value 0 with probability $1-\alpha_h(X_0,X_1^\star)$. The function $\alpha_h(X_0,X_1^\star)$ is known as the acceptance probability *and is given by*

$$\begin{aligned} &\alpha_h(X_0,X_1^\star) \\ &\quad = \min\left(1, \frac{\det(B_h(X_0))}{\det(B_h(X_1^\star))}\exp\left(-\frac{1}{kT}\left[\frac{|\eta|^2}{2} - \frac{|\xi|^2}{2} + U(X_1^\star) - U(X_0)\right]\right)\right), \end{aligned} \tag{2.3}$$

where η satisfies:

$$B_h(X_1^\star)\eta = B_h(X_0)\xi + \sqrt{2h}G_h(\tilde{X}_1)\,. \tag{2.4}$$

We remark that the calculation of the acceptance probability in (2.3) usually requires a Cholesky factorization to determine $B_h(X_1^\star)$ and its determinant; and a linear solve to determine η via (2.4). Let us now discuss how to choose $G_h(x)$ and $B_h(x)$ in (2.1) by requiring that Properties (P1)–(P4) are satisfied. It is useful to look at these properties sequentially and see the constraints on $G_h(x)$ and $B_h(x)$ they entail.

2.1. Property (P1): Ergodicity. As long as $B_h(x)B_h(x)^T$ is positive definite for every x at which $\nu(x)>0$, the transition probability distribution induced by Algorithm 2.1 satisfies the ν-symmetry property (1.6) with $t=h$. (Note that this transition probability distribution has no density with respect to Lebesgue measure, so (1.6) must be reinterpreted in terms of distributions: this is a minor technical difficulty on which we will dwell upon in §4.) When $\nu(x)$ is normalizable and the variables $G_h(x)$ and $B_h(x)$ are sufficiently regular, it can be shown that Algorithm 2.1 is ergodic with respect to the equilibrium probability distribution with density $\nu(x)/Z$. Thus, one can stably generate an infinitely long trajectory from the method with the

right stationary distribution. Moreover, the ergodicity theorem implies that

$$\frac{1}{T}\int_0^T f(X_{\lfloor t/h\rfloor})dt \to \frac{1}{Z}\int_{\mathbb{R}^n} f(x)\nu(x)dx\,,\quad \text{as}\quad T\to\infty, \tag{2.5}$$

where $f(x)$ is a suitable test function. Note that it is not surprising that Property (P1) holds for quite general $G_h(x)$ and $B_h(x)$ since the ν-symmetry property (1.6) is enforced by the accept/reject step in Algorithm 2.1 as long as the correct acceptance probability $\alpha_h(X_0,X_1^*)$ is used. This idea is the essence of the Metropolis-Hastings method. More remarkable is the next observation, relevant to Property (P2).

2.2. Property (P2): Weak Accuracy. For every sufficiently regular $G_h(x)$ and $B_h(x)$ satisfying

$$B_h(x)B_h(x)^T = M(x)+\mathcal{O}(h) \qquad \text{for all } x\in\mathbb{R}^n, \tag{2.6}$$

where $M(x)$ is the mobility matrix entering (1.1), Algorithm 2.1 is weakly accurate on finite time intervals:

$$|\mathbb{E}_x(f(Y(\lfloor t/h\rfloor h))-\mathbb{E}_x(f(X_{\lfloor t/h\rfloor}))|\le C(T,G_h)h^{1/2}\,,\ \text{for all } t\in[0,T], \tag{2.7}$$

where $\mathbb{E}_x$ denotes the expectation conditional on the initial state being x. The observation above is remarkable in that it holds for any sufficiently regular $G_h(x)$, including the trivial $G_h(x)=0$. As we will see in §5, this really is a consequence of an (infinitesimal) fluctuation-dissipation theorem: as long as the noise term is handled accurately in the proposal move (2.1) (which it is if (2.6) holds), weak accuracy follows from the fact that the only diffusion satisfying the ν-symmetry property (1.6) (which it does through Property (P1)) is the one with the correct drift term. Note that this immediately opens the door to schemes where the divergence of $M(x)$ does not need to be computed. In fact, with $G_h(x)=0$, the calculation of no part of the drift is necessary. Of course, this trivial choice is not the best (nor even a good) one, and to enforce Property (P3), we will have to use more specific $G_h(x)$ and $B_h(x)$.

2.3. Property (P3): Second-order Accuracy in Small Noise Limit. Let $G_h(x)$ and $B_h(x)$ be the following two-stage Runge-Kutta methods:

$$\begin{cases} G_h(x) = -b_1M(x)DU(x)-b_2M(x)DU(x_1) \\ \qquad\qquad -b_3M(x_1)DU(x)-b_4M(x_1)DU(x_1)\,, \\ x_1 = x-a_{12}hM(x)DU(x)\,, \end{cases} \tag{2.8}$$

$$\begin{cases} B_h(x)B_h(x)^T = d_1M(x)+d_2M(x_1)\,, \\ x_1 = x+c_{12}hM(x)DU(x)\,, \end{cases} \tag{2.9}$$

with parameter values:

$$d_1=1/4\,,\ \ d_2=3/4\,,\ \ b_1=5/8\,,\ \ b_2=b_3=-3/8\,,\ \ b_4=9/8\,,\ \ \&\ \ c_{12}=a_{12}=2/3\,. \tag{2.10}$$

Algorithm 2.1 operated with this choice of $G_h(x)$ and $B_h(x)$ satisfies:

$$\lim_{kT\to 0}(\mathbb{E}_x|Y(\lfloor t/h\rfloor h)-X_{\lfloor t/h\rfloor}|^2)^{1/2}\le C(T)h^2\,, \tag{2.11}$$

for all $t\in[0,T]$.

When the mobility is constant, then $B_h(x) = B(x)$ and $G_h(x)$ reduces to the so-called Ralston Runge-Kutta method [87]. The above statement, which is one of the main results of this paper, is established in §6 using an asymptotic analysis of the rejection rate in the deterministic limit as $kT \to 0$. This analysis reveals that the parameter values in (2.10) are optimal, that is, they are the only choice that yield Property (P3). The validity of this statement requires an assumption to the effect that singularities of the Hessian matrix $D^2U(x)$ can be avoided. The sufficiency of this condition is discussed in the numerical examples in §3, but in short, we found that this assumption is hard to violate in practice. When this condition is violated, Section §6 proposes alternatives choices of $G_h(x)$ and $B_h(x)$ and a general strategy to design them; however, the above choice, we believe, strikes a balance between accuracy and complexity. The small noise limit is relevant when the deterministic drift dominates the dynamics of (1.1), which is the case in BD applications when the Péclet number is moderate to high or when the system is driven by an external flow. (The Péclet number compares the work done by the potential force to the thermal energy kT.)

2.4. Additional remarks on integrator. Note that Property (P4) automatically follows when $G_h(x)$ and $B_h(x)$ are given by (2.8) and (2.9), respectively, since at no point do we need to calculate the divergence of $M(x)$. Note also, that with this choice of $G_h(x)$ and $B_h(x)$, the proposal move in Algorithm 2.1 involves internal stages. Both the mobility matrix and force lack a definition if an internal stage assumes a non-physical value. In this case it is straightforward to show that any extension of these functions results in an algorithm that satisfies the ν-symmetry condition (1.6). Hence, we suggest using the trivial extension where the mobility matrix is set equal to the identity matrix and the force is set equal to zero. Independent of the extension chosen, the energy is taken to be infinite at non-physical states, which implies that non-physical proposal moves are always rejected.

Let us end this section by stressing that the idea of using Monte-Carlo to perform BD simulation is not new and goes back at least to [58, 59]. In place of the proposal move (2.1), these papers use the Ermak-McCammon scheme, which corresponds to a forward Euler discretization of (1.1). This scheme reduces to the Metropolis-adjusted Langevin algorithm (MALA) when the mobility matrix is constant [94, 93, 8]. MALA is a special case of the smart and hybrid Monte-Carlo algorithms, which are older and more general sampling methods [95, 17]. However, the Metropolis Ermak-McCammon scheme has two drawbacks: it involves the divergence of the mobility tensor, and worse, as illustrated in the next section, there are important situations where the acceptance probability in the scheme breaks down in the small noise limit. The proposed integrator gets around these problems.

Next, we will use several examples to illustrate these properties of the Metropolis integrator then prove all the statements made in this section in Sections 4, 5 and 6.

3. Numerical Examples. Unless otherwise indicated, in this section we work with Algorithm 2.1 operated with $G_h(x)$ and $B_h(x)$ given by (2.8) and (2.9) with parameter values (2.10). Test problems consist of the following self-adjoint diffusions:

(E1) Brownian particle with a tilted square well potential energy;
(E2) 1D bead-spring chain in a confined solvent;
(E3) 3D bead-spring chain in an unbounded solvent; and,
(E4) Brownian particle with a two-dimensional double-well potential energy.

Example (E1) features a non-normalizable density, and shows that standard explicit integrators may not detect properly features of a potential with jumps, which leads

to unnoticed but large errors in dynamic quantities such as mean first passage times. As shown in (E2) and (E3), standard integrators can also fail when the potential contains a hard-core or Lennard-Jones-type component. Example (E2) is not physically realistic, however, it is simple enough to be allow intensive numerical tests and yet, we believe, complex enough to capture some of the essential features which make (1.1) challenging to simulate including multiplicative noise, nontrivial $(\operatorname{div} M)(x)$, and steep potentials. On the other hand, example (E3) is physically relevant to polymeric fluid simulations but, since the solvent is unbounded $(\operatorname{div} M)(x)=0$ in this case. In sum, (E1) – (E3) demonstrate two main modes of failure of standard explicit integrators. (E4) illustrates the properties of the proposed scheme in the small noise limit.

3.1. Example (E1): Brownian particle with a tilted square well potential energy. The following example emphasizes that the Metropolis integrator, Algorithm 2.1, applies to self-adjoint diffusions even when the density $\nu(x)$ is not normalizable. To introduce this example, it helps to consider simulating a Brownian particle moving in a regularized, periodic square well potential. In order to adequately resolve the jump in the potential, a standard scheme requires a sufficiently small time step size. Without this resolution the scheme's equilibrium distribution will be essentially uniform, and its estimate of dynamic quantities associated to crossing the square potential barrier will be inaccurate. These predictions are borne out in the following numerical experiment. Inspired by [89], we introduce a static tilting force $F>0$ so that the particle drifts to the right intermittently stopping at the jumps in the potential. With this tilting force, Stratonovich was able to derive a formula for the mean first passage time, which we use below to benchmark and compare the Fixman and Metropolis integrator [101]. At this point it is worth mentioning that when the mobility matrix in (1.1) is constant, the Fixman scheme reduces to a second-order trapezoidal discretization of the drift and a first-order approximation to the noise.

To be more precise, let $U(x)$ be a periodic, square well potential given by:

$$U(x)=\tanh\left(\frac{(x \bmod 3)-2}{\epsilon}\right)-\tanh\left(\frac{(x \bmod 3)-1}{\epsilon}\right),$$

where ϵ is a smoothness parameter. The period in this function is selected so that the jumps in $U(x)$ over one cycle $[0,3]$ occur at $x=1$ and $x=2$. A Brownian particle moving in a tilted square well potential satisfies an equation of the form (1.1) with mobility equal to unity and,

$$dY=-\tilde{U}'(Y)dt+\sqrt{2kT}dW\,,\quad Y(0)\in\mathbb{R}\,,$$

where we have introduced the following tilted potential energy function:

$$\tilde{U}(x)=U(x)-Fx\,.$$

Observe that when $F=0$ the potential $\tilde{U}(x)$ reduces to $U(x)$. For every $F\in\mathbb{R}$, it is straightforward to use (1.4) to show that the generator of $Y(t)$ is self-adjoint with respect to the density $\nu(x)=\exp(-\tilde{U}(x)/kT)$. When $F>0$ the static tilting force gives rise to a downward tilt in the potential $\tilde{U}(x)$ as shown in the northwest inset in Figure 3.1, and so, $\nu(x)$ is not normalizable since $\int_{\mathbb{R}}\nu(x)dx=+\infty$. The first passage time of $Y(t)$ from any $x_0\in\mathbb{R}$ to x_0+3 is defined as:

$$\tau=\inf\{t>0\,:\,Y(0)=x_0\,,\quad Y(t)\geq x_0+3\}\,.$$

Parameter	Description	Value
Physical Parameters		
F	static tilting force	0.25
kT	temperature factor	1
ϵ	square well smoothness parameter	0.001
Numerical Parameters		
h	time step size	$\{0.01, 0.05, 0.025, 0.0125, 0.00625, 0.003125\}$
N_s	# of first passage time samples	10^5

TABLE 3.1
Simulation Parameters for a Brownian Particle with a Tilted Square Well Potential.

The numerical experiments that follow consist of launching an integrator with initial condition $X_0 = x_0$ and time step size h, and terminating the simulation at the first step n where $X_n \geq x_0 + 3$. To avoid systematically overestimating the first passage time, we use the approximation, $\tau \approx nh - h/2$. As a side note, we mention that the accuracy of this approximation to the mean first passage time (which goes like $\mathcal{O}(\sqrt{h})$) can be improved upon by accounting for the probability that the particle reaches $x_0 + 3$ in between each discrete step [76].

The numerical and physical parameters used in the numerical experiments are given in Table 3.1. To visualize the long-term behavior of the schemes, it helps to plot the probability density of points produced by each scheme modulo a period of the square well potential. An approximation to this density is shown in the southeast insets in Figure 3.1. At both the coarse and fine time step size tested, the Fixman scheme underestimates the barrier height, and consequently, numerical tests show that it grossly underestimates, e.g., the mean first passage time between wells. This underestimate persists unless its time step size is small enough to resolve the barrier ($h < 10^{-4}$). The Metropolis integrator is able to capture the features of the potential even at the coarse time step size $h = 10^{-1}$, and its approximation to the mean first passage time is about 2% accurate with a time step size $h = 10^{-3}$.

3.2. Example (E2): 1D bead-spring chain with hydrodynamic interactions. This example confirms Properties (P2) and (P3) of the Metropolis integrator. Consider a 1D bead-spring chain consisting of n beads and $n+1$ springs confined to an interval $[0, L]$ and immersed in a 'solvent' with viscosity μ as illustrated in Figure 3.2. We stress that this example is fictitious because incompressibility implies that the velocity of a 1D Stokesian solvent is constant. Still, the example captures some of the key features of hydrodynamic interactions – including a non-trivial $(\operatorname{div} M)(x)$ – while being simple enough to permit detailed numerical studies.

We begin by writing this system as a self-adjoint diffusion of the form (1.1) with a normalizable density $\nu(x)$. Assume that the bead and solvent inertia are negligible. Let q_i and F_i denote the ith bead position and force, respectively; let $u(q) \in \mathbb{R}$ denote the solvent velocity for $q \in [0, L]$; and, let $u_i = u(q_i)$ for $i = 1, \cdots, n$. Order the particle positions so that:

$$0 < q_1 < \cdots < q_n < L\,.$$

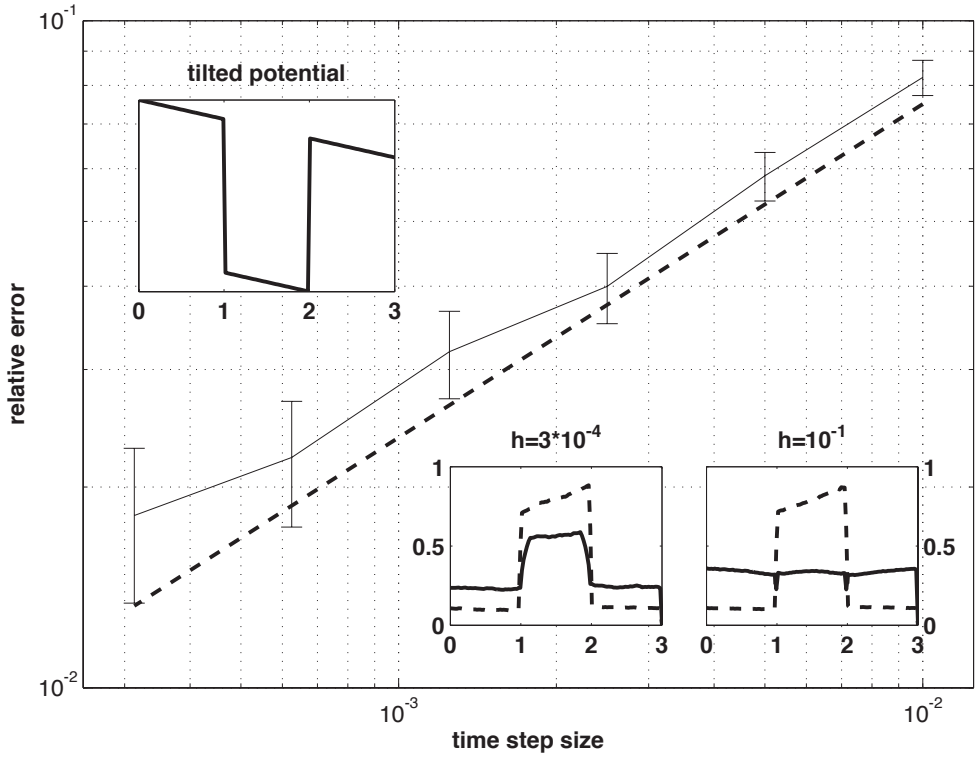


FIG. 3.1. **Brownian Particle with a Tilted Square Well Potential.** *The main figure graphs the relative error in the mean first passage time as a function of the time step size for the Metropolis integrator (solid). The error bars represent a* 99% *confidence interval. The dashed line is an* $\mathcal{O}(h^{1/2})$ *reference slope. The inset on the upper left shows one cycle of the tilted square well potential. The insets on the lower right plot the approximate density of the Fixman (solid) and Metropolis integrator (dashed) obtained by wrapping the points along a trajectory around the boundaries of a period of the square well potential. By comparing the densities at the coarse and fine time step size, notice that the density of the Metropolis integrator is converged at the coarse time step size, while the Fixman scheme at both the coarse and fine time step size is unable to detect properly the jump in the potential. This difference is reflected in the main figure which shows that the Metropolis integrator is about* 2% *accurate with a time step size* $h=10^{-3}$*, while the Fixman scheme is grossly inaccurate at all time step sizes tested.*

This ordering is maintained by the following soft-core spring potential:

$$U_{\mathrm{FENE}}(x) = -\frac{\epsilon}{2}\log\left(1-\left(\frac{x}{2\ell}\right)^2-\left(1-\frac{x}{2\ell}\right)^2\right). \tag{3.1}$$

There are two parameters in this potential: an energy constant ϵ and bond length at rest ℓ. The linear behavior of $U_{\mathrm{FENE}}(x)$ about its resting position $x=\ell$ is given by a Hookean spring with stiffness ϵ/ℓ^2. Moreover, this potential possesses a singularity at $x=0$ that prevents interbead collisions, and a second singularity at $x=2\ell$ to enforce finite extension of the spring length. Potentials of this type play an important role in capturing the right non-Newtonian behavior of a dilute polymer solution using bead-spring models [7, 86].

In addition to neighboring spring interactions, the particles are coupled by non-bonded interactions mediated by a fictitious solvent. Since the bead and solvent

inertia are negligible, a balance between the solvent viscous force and the bead spring forces leads to the following equations for the solvent velocity:

$$\mu\frac{\partial^2 u}{\partial x^2}=-\sum_{i=1}^{n}F_i\delta(q-q_i)\,,\quad u(0)=u(L)=0\,, \tag{3.2}$$

where $\delta(q-q_i)$ is a unit force located at the position of the ith bead q_i. Let $x=(q_1,\cdots,q_n)^T$, and in terms of which, the mobility matrix $M(x)$ in (1.1) comes from solving (3.2) and it describes the linear relationship between the solvent velocities at the bead positions and the bead forces:

$$\sum_{j=1}^{n}M_{ij}(x)F_j=u_i\,. \tag{3.3}$$

The deterministic drift in (1.1) arises from assuming the solvent velocity matches the bead velocities at the bead positions. A heat bath is then added to this drift to ensure that the density of the stationary distribution of the resulting equations is proportional to:

$$\nu(x)=\exp\left(-\frac{1}{kT}U(x)\right)\,,$$

where the total potential energy is just the sum of the $n+1$ spring potential energies:

$$U(x)=\sum_{i=1}^{n+1}U_{\rm FENE}(q_i-q_{i-1})\,,\quad \text{where } q_0=0 \text{ and } q_{n+1}=L\,.$$

Now we show how to solve (3.2) for the solvent velocity, and in the process, derive a procedure for exactly evaluating the mobility matrix in this specific case.

In terms of the friction matrix $\Gamma(x)=M^{-1}(x)$, the linear relationship (3.3) can be written as:

$$F_i=\sum_{j=1}^{n}\Gamma_{ij}u_j\,. \tag{3.4}$$

Given the instantaneous forces and positions of the particles, the solution to (3.2) can be derived as follows. In between the particles, the fluid velocity is linear since the solvent experiences no force there. At the ith particle position, there is a discontinuity in the derivative of the fluid velocity:

$$\frac{\partial u}{\partial q}(q=q_i^+)-\frac{\partial u}{\partial q}(q=q_i^-)=-\frac{F_i}{\mu}\,, \tag{3.5}$$

for $i=1,\cdots,n$. With $u_0=u(q_0)=0$ and $u_{n+1}=u(q_{n+1})=0$, notice that (3.5) implies:

$$\frac{u_{i+1}-u_i}{q_{i+1}-q_i}-\frac{u_i-u_{i-1}}{q_i-q_{i-1}}=-\frac{F_i}{\mu}\,. \tag{3.6}$$

Comparing this equation with (3.4), it is evident that Γ is a tridiagonal matrix with entries given by:

$$\Gamma_{ij}=\begin{cases}\frac{\mu}{q_{i+1}-q_i}+\frac{\mu}{q_i-q_{i-1}}\,, & \text{if } i=j\,,\\ -\frac{\mu}{q_{i+1}-q_i}\,, & \text{if } i-j=1\,,\\ -\frac{\mu}{q_i-q_{i-1}}\,, & \text{if } i-j=-1\,.\end{cases} \tag{3.7}$$

These expressions are explicit and straightforward to implement in the Fixman scheme or the Metropolis integrator, Algorithm 2.1. The friction and mobility matrices are symmetric positive-definite, and the friction matrix is tridiagonal while the mobility matrix is not tridiagonal, in general.

For the numerical experiment, consider an eight bead chain that is initially compressed as shown in the northwest inset of Figure 3.3. The equilibrium positions of the beads are indicated by the tick marks in this inset. We use the Fixman and Metropolis integrator, Algorithm 2.1, to simulate the relaxation of the system from this state. Specifically, the schemes compute the expected value of the average position of the particles as the lattice relaxes towards equilibrium:

$$\mathbb{E}_x\left(\frac{1}{n}\sum_{i=1}^{n} q_i(t)\right), \tag{3.8}$$

for $t\in[0,T]$, where $T=50$ dimensionless time units was found to be sufficient to capture the relaxation dynamics for all parameter values tested. Setting μ_{visc}, L, and ϵ equal to unity is equivalent to rescaling the system by a characteristic length, time and energy scale. This leaves the temperature factor kT as a free physical parameter that we vary in the numerical experiments. These simulation parameter values are provided in Table 3.2.

The relative error of the Metropolis integrator in computing (3.8) is shown in Figure 3.3 for a range of temperatures. Near the deterministic limit, we used a forward Euler scheme operated at a very small time step size to obtain a benchmark solution ($h<10^{-5}$). The Metropolis integrator is accurate for time step sizes that exceed $h=0.05$. As the temperature is increased, and in the absence of an analytical solution, we used Richardson extrapolation to obtain a benchmark solution. Using this benchmark we estimated the error of the Metropolis integrator at various temperatures. These graphs show that although the relative error increases with increasing temperature, the Metropolis integrator remains within 1% error for a time step size $h=0.1$ from the deterministic limit to the moderate temperature $kT=0.1$.

To deal with stochastic instabilities in the Fixman scheme, we use 'the method of rejecting exploding trajectories' [84]. In this stabilization technique, if a nonphysical move is generated by the Fixman scheme the entire sample path is rejected. At the very low temperature of $kT=0.001$, and if $h\leq 0.0125$, the Fixman scheme rejects a negligible amount of trajectories and is as accurate as the Metropolis integrator operated at $h=0.2$. However, the Fixman scheme rejects almost all trajectories generated at the low temperature $kT=0.01$ even if the time step size is reduced to $h=0.0001$, and its performance is much worse at the moderate temperature $kT=0.1$.

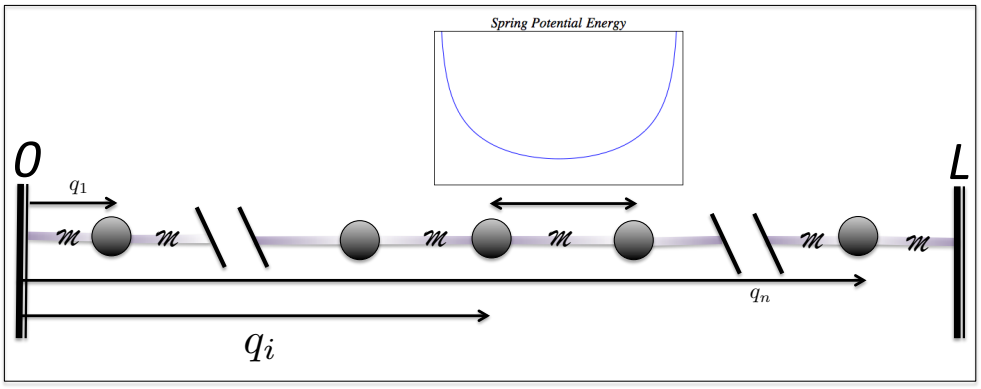


FIG. 3.2. *This figure shows a one-dimensional bead spring chain in a solvent. The chain consists of n beads connected by $n+1$ springs confined to an interval $[0,L]$. The spring potential energies are modeled by a two-sided FENE potential* (3.1) *that is plotted in the inset. The singularities in this potential enforce excluded volume between beads and finite extension of the spring length. The solvent velocity satisfies a steady-state Stokes equation with point sources of drag located at the bead positions.*

Parameter	**Description**	**Value**
Physical Parameters		
μ_{visc}	solvent viscosity	1
L	chain length	1
ϵ	energy constant	1
kT	temperature factor	$\{0.0001, 0.001, 0.01, 0.1\}$
n	# of beads	8
T	time-span for simulation	50
Numerical Parameters		
h	time step sizetime step size	$\{0.2, 0.1, 0.05, 0.025, 0.0125\}$
N_s	# of sample trajectories	10^5
$(q_1(0), \cdots, q_n(0))$	initial positions	$(0.1, 0.11, 0.33, 0.34, 0.56, 0.57, 0.79, 0.81)$
$(\tilde{q}_1, \cdots, \tilde{q}_n)$	equilibrium positions of springs	$(0.1, 0.21, 0.33, 0.44, 0.55, 0.67, 0.79, 0.9)$

TABLE 3.2
Simulation parameters for a one-dimensional bead-spring chain in a solvent.

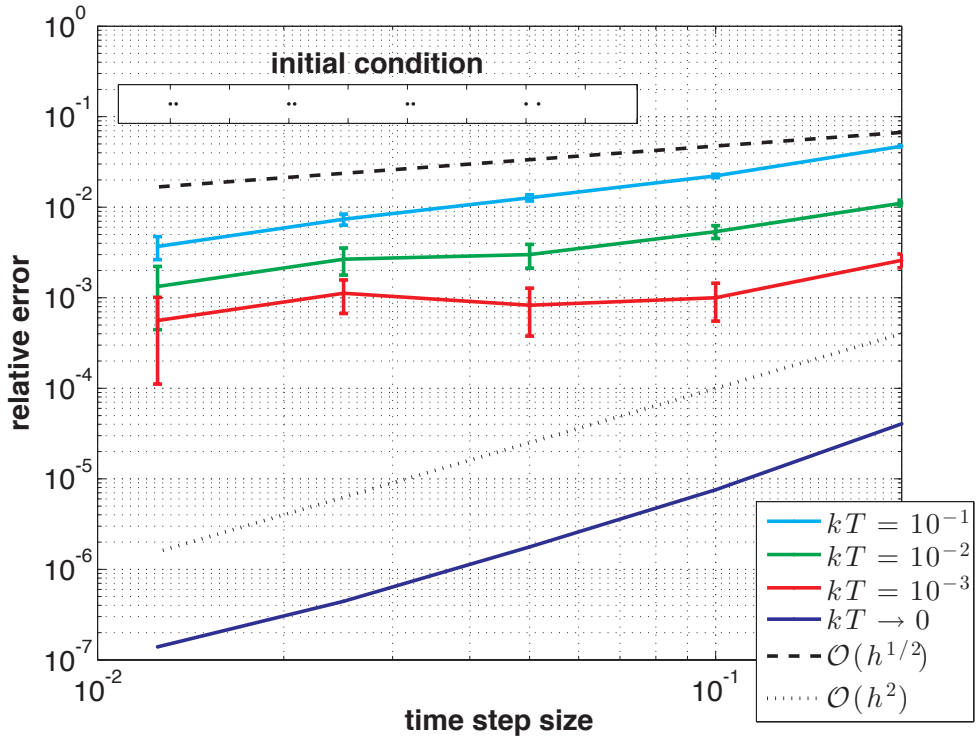


FIG. 3.3. **One-dimensional bead-spring chain in a solvent.** *The main figure shows the relative error in the mean position of the particles* (3.8) *computed using the Metropolis integrator at the kT values shown in the legend. The error bars represent a* 95% *confidence interval. The northwest inset shows the initial condition (dots) and equilibrium positions (tick marks) of the beads. Observe that the Metropolis integrator remains within* 1% *accurate at the large time step size* $h = 0.1$ *for a wide range of temperatures. This figure confirms that the Metropolis integrator is weakly accurate at constant temperature Property (P2), and second-order accurate in the small noise limit Property (P3). The curve* $kT \to 0$ *corresponds to a kT that is small enough – a pico kT – so that the* $\mathcal{O}(h^2)$ *rate of convergence is dominant.*

3.3. Example (E3): DNA dynamics in a solvent. The following example applies the Metropolis integrator to a BD simulation of a DNA molecule with hydrodynamic interactions. Consider a bead-spring model of a bacteriophage DNA molecule with $N_b=11$ spherical beads, where each spring approximates the effect of 4850 base pairs, so that ten springs contain approximately the number of base pairs in a DNA molecule with a contour length of $21\mu m$ [112]. Assume the beads are spherical with radius R_b and move in a Stokesian solvent with viscosity η_s. To be specific, the solvent velocity $u(q)\in\mathbb{R}^3$ and pressure $p(q)\in\mathbb{R}$ satisfy:

$$\eta_s(\nabla^2 u)(q)-(\nabla p)(q)+f(q)=0\,,\quad (\nabla\cdot u)(q)=0\,,\quad \text{for all } q\in\Omega\subset\mathbb{R}^3\quad, \tag{3.9}$$

where Ω is the domain of the solvent and $f(q)\in\mathbb{R}^3$ is the applied force density due to the beads. We augment these equations with the following boundary conditions: the fluid is at rest at infinity and satisfies no-slip conditions on the surfaces of each bead. Let q_i, v_i, and F_i denote the position, translational velocity, and force of the ith bead where $i\in\{1,\cdots,N_b\}$. Introduce the $3N_b$ dimensional vectors of bead positions $x=(q_1,\cdots,q_{N_b})^T$, bead translational velocities $V=(v_1,\cdots,v_{N_b})^T$ and bead forces $F=(F_1,\cdots,F_{N_b})^T$. An immediate consequence of the linearity of the Stokes equation (3.9) is that:

$$V=M(x)F\,,$$

where $M(x)$ is the $3N_b\times 3N_b$ mobility matrix. This linear relationship always holds for bodies moving in a Stokes fluid. Moreover, the matrix $M(x)$ is always symmetric and positive definite for every $x\in\mathbb{R}^{3N_b}$.

Determining the entries of the mobility matrix requires solving the Stokes equation (3.9) for the solvent velocity field which is a very complicated boundary value problem. This difficulty motivates using the Rotne-Pragner-Yamakawa (RPY) approximation of the solvent velocity field, which leads to the following approximate mobility matrix:

$$M(x)=\begin{bmatrix}\Omega_{1,1} & \cdots & \Omega_{1,N_b}\\ \vdots & \ddots & \vdots\\ \Omega_{N_b,1} & \cdots & \Omega_{N_b,N_b}\end{bmatrix}\,,\quad \Omega_{i,j}=\begin{cases}\frac{1}{\zeta}I_{3\times 3}\,, & \text{if } \ i=j\\ \Omega_{RPY}(r_i-r_j)\,, & \text{otherwise}\,,\end{cases} \tag{3.10}$$

for all $x\in\mathbb{R}^{3N_b}$. Here, we have introduced the 3×3 matrix $\Omega_{RPY}(q)$ defined as:

$$\Omega_{RPY}(q)=\frac{1}{\zeta}\left(C_1(q)I_{3\times 3}+C_2(q)\frac{q}{|q|}\otimes\frac{q}{|q|}\right)\,, \tag{3.11}$$

where $C_1(q)$ and $C_2(q)$ are the following scalar-valued functions:

$$C_1(q)=\begin{cases}\frac{3}{4}\left(\frac{R_b}{|q|}\right)+\frac{1}{2}\left(\frac{R_b}{|q|}\right)^3\,, & \text{if } \ |q|>2R_b\,,\\ 1-\frac{9}{32}\left(\frac{|q|}{R_b}\right)\,, & \text{otherwise}\,,\end{cases}$$

and,

$$C_2(q)=\begin{cases}\frac{3}{4}\left(\frac{R_b}{|q|}\right)-\frac{3}{2}\left(\frac{R_b}{|q|}\right)^3\,, & \text{if } \ |q|>2R_b\,,\\ \frac{3}{32}\left(\frac{|q|}{R_b}\right)\,, & \text{otherwise}\,.\end{cases}$$

The quantity $1/\zeta$ is the mobility constant produced by a single bead translating in an unbounded solvent at a constant velocity: $\zeta = 6\pi\eta_s R_b$. The approximation (3.10) preserves the physical property that the mobility matrix is positive semi-definite, satisfies $(\operatorname{div} M)(x) = 0$ for all $x \in \mathbb{R}^{3N_b}$, and is exact up to $\mathcal{O}((R_b/r_{ij})^4)$ where R_b is the bead radius and r_{ij} is the distance between distinct beads i and j [96].

A well-defined characteristic length of DNA is its Kuhn length b_k which represents the distance along the polymer chain over which orientational correlations decay. The bending rigidity of a polymer decreases with increasing b_k. For DNA, the Kuhn length is approximately a tenth of a micrometer. In terms of which, consider a worm like chain (WLC) model for the spring potential energy given by:

$$U_{\mathrm{WLC}}(r) = \frac{kT}{2b_k}\left(\frac{\ell^2}{\ell - r} - r + \frac{2r^2}{\ell}\right), \tag{3.12}$$

where ℓ is the maximum length of each spring. This empirical potential energy captures the entropic elasticity which causes the DNA molecule to be in a tightly coiled state. The linear behavior of this potential is given by a Hookean spring with spring constant: $H_s = 3kT/(b_k\ell)$. The strength of the hydrodynamic interactions can be quantified by using this spring constant to compute the dimensionless bead radius: $a^* = R_b/\sqrt{kT/H_s}$. Using the DNA parameter values provided in Table 3.3, we find that $a^* = 0.291$ which for a dimensionless Rouse model (same bead-spring chain, but with Hookean springs) signifies a moderate strength of hydrodynamic interactions [86]. Since the Rouse chain has the same features as the DNA model, minus the steep potential, a reasonable time step size for the DNA simulation can be determined by simulating a dimensionless Rouse system at this value of a^* and the same non-random initial condition. In particular, it can be shown that a time step size of $h \approx 10^{-4}$ leads to an average acceptance probability of about $98\% - 99\%$ as the Rouse chain transitions from a stretched to a coiled state.

With approximately this time step size $h = (10^{-4})/2$, we use the Metropolis integrator to generate one thousand trajectories of the DNA chain from the initial conformation shown in the inset of Figure 3.4 using the values provided in Table 3.3 over the interval $[0,1]$. This time-span is sufficiently long to capture the relaxation dynamics of the chain. The average acceptance probability is about 98%. From the simulation data, the radius of gyration was estimated and this estimate is plotted in Figure 3.4. Repeating this experiment at higher time resolution led to no noticeable change in this graph. The Metropolis integrator seems able to compute dynamics for this system at a time step size that is about $50\times$ larger than what is possible using the Fixman scheme combined with 'the method of rejecting exploding trajectories' as described in the 1D bead-spring example.

Parameter	Description	Value	Units
Physical Parameters			
N_b	# of beads	11	
b_k	Kuhn length	1×10^{-1}	μm
R_b	bead radius	7.7×10^{-2}	μm
ℓ	maximum spring length	2.1	μm
η_s	solvent viscosity	1×10^{-9}	$\frac{kg}{\mu m\cdot s}$
kT	thermal energy	4.11×10^{-9}	$\frac{kg(\mu m)^2}{s^2}$
T	time-span	1	s
Numerical Parameters			
h	time step size	0.5×10^{-4}	s
N_s	# of sample trajectories	10^3	

TABLE 3.3
DNA Simulation Parameters.

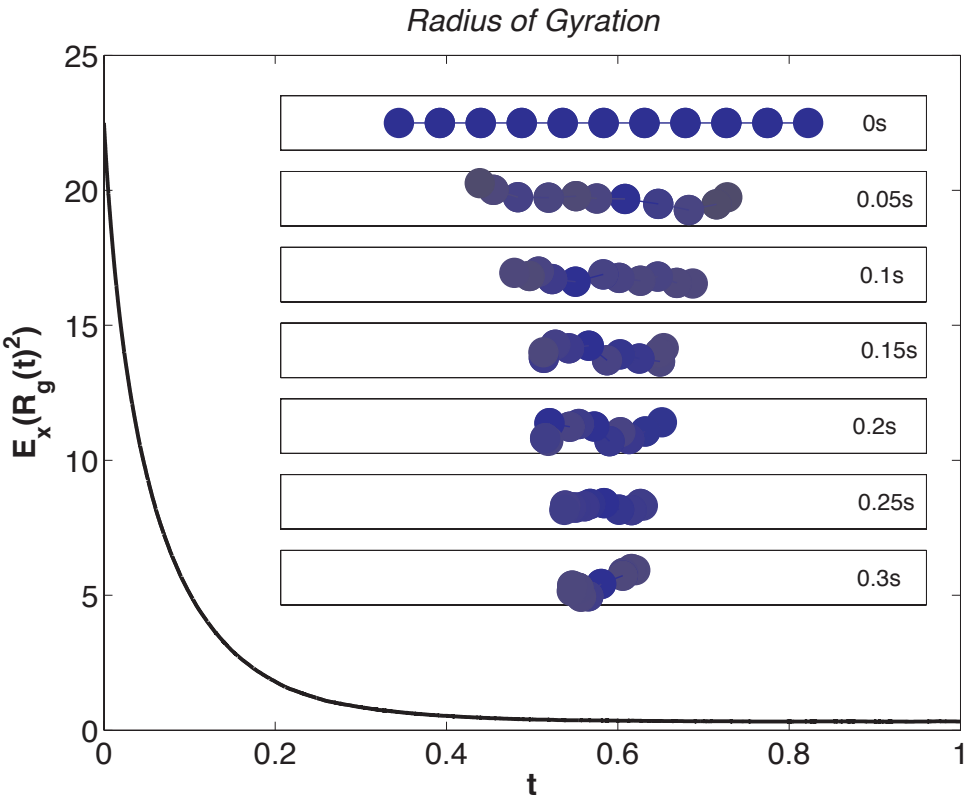


FIG. 3.4. **DNA Simulation.** *The main figure plots the exponential decay in the mean-squared radius of gyration and the inset plots snapshots of the relaxation of a DNA molecule from an initial conformation shown in the top-most panel of the inset. Initially the beads are evenly spaced on the x-axis with a spacing of* $1.5\mu m$. *In this example, the Metropolis integrator is able to take about one to two orders of magnitude larger time step size than conventional explicit schemes.*

3.4. Example (E4): Brownian particle with a double well potential energy. In the small noise limit, the proposal move (2.1) converges to a deterministic update:

$$X_1^\star \to X_0 + hG_h(X_0)\,, \qquad \text{as } kT \to 0\,. \tag{3.13}$$

If this update is a higher order discretization of the zero noise limit of (1.1), then the proposal move will be deterministically accurate to that order. However, it does not follow from this alone that the actual update in (2.2) is deterministically accurate too. Indeed, Property (P3) also requires that

$$\alpha_h(X_0, X_1^\star) \to 1\,, \qquad \text{as } kT \to 0\,. \tag{3.14}$$

This statement may appear to contradict Property (P2): that the algorithm is weakly accurate so long as the noise is handled correctly regardless of the proposal move used, but it does not. To be perfectly clear, both (3.13) and (3.14) are asymptotic statements in kT while keeping the time step size h fixed (and sufficiently small for certain estimates to be valid), whereas Property (P2) is a non-asymptotic result that assumes both kT and h are fixed.

In §6 we show that for general $G_h(x)$ and $B_h(x)$,

$$\alpha_h(X_0, X_1^\star) \sim 1 \wedge \exp\left(-\frac{1}{kT} f_{kT}(X_0, h)\right)\,, \qquad \text{as } kT \to 0\,, \tag{3.15}$$

where we have introduced

$$\begin{cases} f_{kT}(x,h) = U(x + hG_h(x)) - U(x) + hG_h(x)^T M_h(x + hG_h(x))^{-1} G_h(x)\,, \\ M_h(x) = B_h(x) B_h(x)^T\,. \end{cases} \tag{3.16}$$

From (3.15) it follows that if $f_{kT}(x,h) \le 0$ for all $x \in \mathbb{R}^n$, then the asymptotic statement (3.14) is true, and the Metropolis integrator acquires the accuracy of the deterministic update in (3.13). On the other hand, proposal moves from initial conditions $X_0 = x$ where $f_{kT}(x,h) > 0$ will most likely be rejected if the noise is small enough, and therefore, in the small noise limit the Metropolis integrator is not deterministically accurate at these points. In §6 it is proved that when $G_h(x)$ and $B_h(x)$ are given respectively by (2.8) and (2.9) with parameter values given in (2.10), the function $f_{kT}(x,h)$ is negative semi-definite for h small enough, and hence, the Metropolis integrator operated with this choice of $G_h(x)$ and $B_h(x)$ is second-order deterministically accurate. A sufficient condition for this statement to be true is an assumption on the Hessian of $U(x)$ – introduced in Assumption 6.1 – which roughly speaking requires that the integrator does not hit singular points of the Hessian matrix of $U(x)$. The following example illustrates these statements, and in particular, the sufficiency of Assumption 6.1.

Consider a Brownian particle with the following two-dimensional double-well potential energy function:

$$U(x) = 5(x_2^2 - 1) + 1.25(x_2 - x_1/2)^2\,, \qquad x = (x_1, x_2)^T\,,$$

and a mobility matrix set equal to the 2×2 identity matrix for simplicity. The Hessian of this potential energy becomes singular at the two points marked by an 'x' in Figures 3.6 (a) and (b). (In a neighborhood of these points, Assumption 6.1 might not

be satisfied.) For this example, $G_h(x)$ in (2.8) with parameter values given in (2.10) simplifies to the Ralston Runge-Kutta method:

$$\begin{cases} G_h(x) = -\frac{1}{4}DU(x) - \frac{3}{4}DU(\tilde{x}), \\ \tilde{x} = x - \frac{2}{3}hDU(x). \end{cases} \tag{3.17}$$

We will compare this choice against the following choices of $G_h(x)$:

$$G_h(x) = -DU(x), \tag{3.18}$$

$$G_h(x) = -DU\left(x - \frac{h}{2}DU(x)\right), \tag{3.19}$$

which correspond to a first-order forward Euler and second-order midpoint discretization of (1.1) in the zero noise limit, respectively. We also consider a third-order accurate Kutta approximation given by:

$$\begin{cases} G_h(x) = -\frac{1}{6}DU(x) - \frac{2}{3}DU(\tilde{x}) - \frac{1}{6}DU(\bar{x}), \\ \tilde{x} = x - \frac{h}{2}DU(x), \quad \bar{x} = x + hDU(x) - 2hDU(\tilde{x}). \end{cases} \tag{3.20}$$

Figure 3.5 plot the $f_{kT}(x,h)$ function and a sample trajectory for the forward Euler and midpoint schemes. In the gray-shaded regions, the $f_{kT}(x,h)$ function is positive, and therefore, if the noise is small enough, the sample trajectory of the Metropolis scheme will likely get stuck in these regions. In the simulation we pick $kT = 10^{-8}$. The Metropolis integrator based on $G_h(x)$ given by (3.18) or (3.19) stop accepting proposal moves once they enter the shaded region.

Figure 3.6 plot the $f_{kT}(x,h)$ function of the Ralston and Kutta schemes, which as expected are negative semi-definite when h is small enough. The inset in Figures 3.6 (a) and (b) zoom into neighborhoods of the two points where the matrix $D^2U(x)$ is singular, showing that the $f_{kT}(x,h)$ function of the Ralston scheme can be positive in these neighborhoods. The insets also show that these regions become smaller as the time step size is reduced. (On the other hand, the $f_{kT}(x,h)$ function of the midpoint scheme can be positive in regions that do not shrink with decreasing time step size – underscoring the optimality of the Ralston scheme among RK2 methods.) The Kutta scheme does not have such a problem at these singular points of the Hessian. Indeed, in §6 we prove that the $f_{kT}(x,h)$ function of any third-order Runge-Kutta approximation is always negative semi-definite if h is small enough regardless of whether Assumption 6.1 is satisfied.

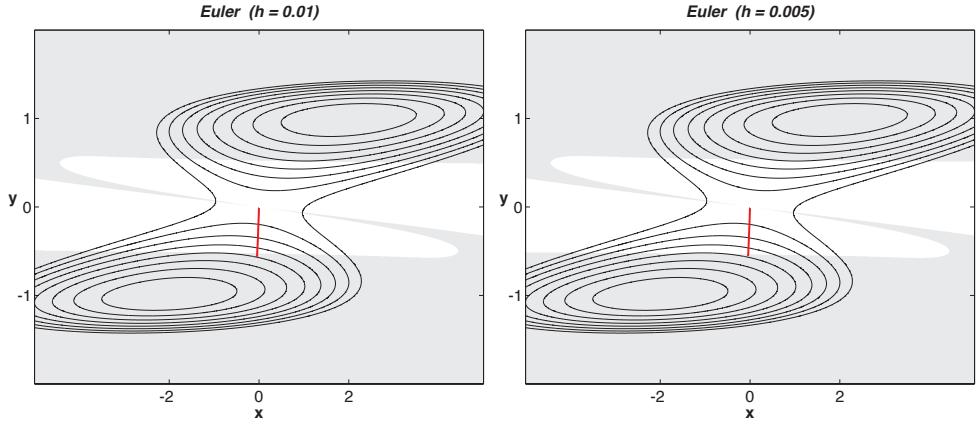
We have also considered the same system, but with the non-constant mobility matrix given by:

$$M(x) = \begin{bmatrix} x_1^2 + x_2^2 + 1 & 0 \\ 0 & x_1^2 + x_2^2 + 1 \end{bmatrix}.$$

Experiments revealed that choosing $B_h(x) = B(x)$ leads to an $f_{kT}(x,h)$ function that is not negative definite, whereas using $B_h(x)$ given by (2.9) gives the desired property. Moreover, as shown in §6, $G_h(x)$ given by a third-order accurate Runge-Kutta scheme (6.22), and $B_h(x)$ correspondingly selected (6.24), also leads to a deterministically third-order accurate scheme.

Parameter	Description	Value
Physical Parameters		
kT	temperature factor	10^{-8}
Numerical Parameters		
h	time step size	$\{0.01, 0.005\}$
$Y(0)$	initial condition	$(0,-0.01)^T$

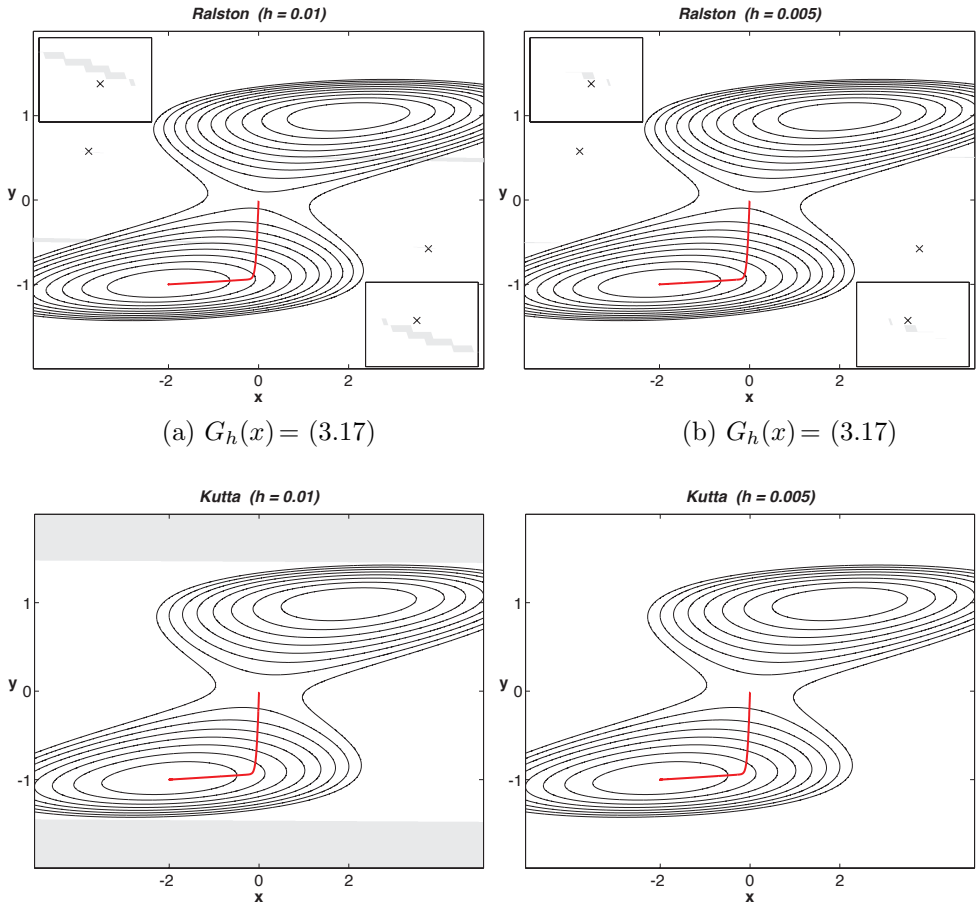
TABLE 3.4
Simulation Parameters for a Brownian Particle with a 2D Double Well Potential.



(a) $G_h(x) =$ (3.18) (b) $G_h(x) =$ (3.18)

(c) $G_h(x) =$ (3.19) (d) $G_h(x) =$ (3.19)

FIG. 3.5. **Deterministically Inaccurate Metropolis Integrators.** *The shaded regions in (a) - (d) are areas where the $f_{kT}(x,h)$ function of the Metropolis integrator is positive. In the background are contours of the potential energy in black. Observe that the $f_{kT}(x,h)$ functions of the Euler and Midpoint schemes are positive in a region which persists even if the time step size is halved. Thus, sample trajectories produced by these schemes – starting at $(0,-0.01)^T$ with $kT = 10^{-8}$ – incorrectly terminate once they enter the shaded regions.*



(a) $G_h(x) =$ (3.17) (b) $G_h(x) =$ (3.17)

(c) $G_h(x) =$ (3.20) (d) $G_h(x) =$ (3.20)

FIG. 3.6. **Deterministically Accurate Metropolis Integrators.** *Gray shading represents regions where the f_{kT} function of the Metropolis integrator is positive. In the background are contours of the potential energy in black. The singular points of the Hessian of $U(x)$ are x-marked in (a) and (b). Moreover, the insets in (a) and (b) show that the $f_{kT}(x,h)$ function of the Ralston scheme can be positive in neighborhoods of these singular points. When the time step size is reduced in the Ralston scheme, the gray regions become smaller. The Kutta scheme – being third-order accurate – does not have a problem at these points, confirming the theory provided in §6. As shown sample trajectories – starting at $(0,-0.01)^T$ with $kT = 10^{-8}$ – are accurately represented by both schemes.*

4. Ergodicity for Normalizable $\nu(x)$. In this section it is shown that Algorithm 2.1 satisfies Property (P1) when the density $\nu(x)$ in (1.3) is normalizable, with normalization constant Z given in (1.7). We begin by showing that $\nu(x)/Z$ is a stationary density for the method by viewing the algorithm first as a Metropolis method in an enlarged space, and second, as a Metropolis-Hastings method in the original space. By using tools from the theory of Metropolis-Hastings methods, we use the second viewpoint to prove that this stationary density is unique, and that the k-step transition probability distribution of the algorithm converges to the equilibrium distribution $\mu(dx)=(\nu(x)/Z)dx$ as $k\to\infty$ for arbitrary initial distributions.

4.1. Integrator as a Metropolis algorithm in an enlarged space. Interestingly, even though the solution to (1.1) is not differentiable, the proposal move in Algorithm 2.1 can be derived from a discretization of a second-order differential equation. This derivation can then be used to prove that the scheme is a Metropolis algorithm. In this way the steps that follow resemble those used in the derivation of the hybrid Monte-Carlo method [95, 17].

Let $G_h:\mathbb{R}^n\to\mathbb{R}^n$ and $B_h:\mathbb{R}^n\to\mathbb{R}^{n\times n}$ be the functions appearing in Algorithm 2.1. In terms of which, consider the following extended dynamics:

$$\begin{cases}\dot{Q}=V\,,\\ \dot{V}=G_h(Q)\,,\end{cases} \tag{4.1}$$

with initial conditions:

$$Q(0)=X_0\,,\quad V(0)=B_h(X_0)\xi\,,$$

where $\xi\in\mathbb{R}^n$ denotes a Gaussian random vector with mean zero and covariance $\mathbb{E}(\xi_i\xi_j)=kT\delta_{ij}$. A position Verlet discretization of this system is given by [30]:

$$\begin{cases}Q_{1/2}=Q_0+\frac{\delta t}{2}V_0\,,\\ V_1=V_0+\delta t G_h(Q_{1/2})\,,\\ Q_1=Q_{1/2}+\frac{\delta t}{2}V_1\,,\end{cases} \tag{4.2}$$

where δt is an artificial time step size. The proposal move (2.1) is obtained by setting:

$$\delta t=\sqrt{2h}\,,\quad X_1^\star=Q_1\,, \tag{4.3}$$

where h is the physical time step size. Since position Verlet is volume-preserving and time reversible in the enlarged position-velocity space, it is a valid proposal move within a Metropolis algorithm targeted to the equilibrium distribution with density:

$$\nu_{\text{extended}}(x,v)=\frac{(2\pi kT)^{-n/2}}{\det(B_h(x))}\exp\left(-\frac{1}{kT}\frac{v^TM_h(x)^{-1}v}{2}\right)\frac{\nu(x)}{Z}\,, \tag{4.4}$$

where we have introduced the matrix $M_h(x)=B_h(x)B_h(x)^T$. Notice that the marginal probability density function of $\nu_{\text{extended}}(x,v)$ in the original position space is the correct one: $\nu(x)/Z$. The acceptance probability (2.3) is obtained from:

$$\alpha_h(X_0,X_1^\star)=1\wedge\frac{\nu_{\text{extended}}(Q_1,V_1)}{\nu_{\text{extended}}(Q_0,V_0)}\,,$$

which is the standard Metropolis ratio in the enlarged space. We emphasize that nowhere in this derivation did we assume a specific form of $G_h(x)$, and that by casting the integrator as a Metropolis algorithm, it immediately follows that the algorithm preserves the probability density $\nu(x)/Z$.

4.2. Integrator as a Metropolis-Hastings algorithm. Here it is shown that the Markov chain produced by Algorithm 2.1 can be put in the frame of a Metropolis-Hastings method. A key idea in this framework is the notion of Markov chain reversibility: a Markov chain with transition probability distribution $P(x,dy)$ is said to be *reversible* with respect to a measure $\mu(dx)$ if and only if the following equality of measures holds,

$$P(x,dy)\mu(dx) = P(y,dx)\mu(dy)\,. \tag{4.5}$$

This condition is a generalization of the ν-symmetry condition (1.6) to measures. Indeed, setting $P(x,dy) = p_t(x,y)dy$ and $\mu(dy) = \nu(y)dy$, (4.5) implies (1.6). In other words, (1.6) is a special case of (4.5), when both the transition probability distribution and the measure $\mu(dx)$ have a common dominating measure, like Lebesgue measure. A reversible Markov chain automatically admits $\mu(dx)$ as an invariant measure since:

$$\int\int f(y)P(x,dy)\mu(dx) = \int\int f(y)P(y,dx)\mu(dy) = \int f(y)\mu(dy)\,,$$

for arbitrary test function $f(x)$.

The Metropolis-Hastings algorithm constructs a Markov chain with a specified stationary distribution $\mu(dx)$ by enforcing condition (4.5) at every step of the chain [82, 32]. The method is made up of a proposal move with probability distribution $Q(x,dy)$ and an acceptance probability $\alpha(x,y)$. If the current state is x, the algorithm updates this state in two sub-steps: first, a proposal move is generated from $Q(x,dy)$; and second, this proposal move is accepted with probability $\alpha(x,y)$, and otherwise, the proposal move is rejected and the chain remains at x. Thus, the transition probability distribution of a Metropolis-Hastings chain can be written as:

$$P(x,dy) = Q(x,dy)\alpha(x,y) + \delta_x(dy)\int(1-\alpha(x,z))Q(x,dz)\,, \tag{4.6}$$

where $\delta_x(dy)$ denotes the Dirac-delta distribution on $\mathbb{R}^n$ concentrated at the current state $x\in\mathbb{R}^n$. The Metropolis-Hastings method requires that $Q(x,dy)$ and $\alpha(x,y)$ are selected so that $P(x,dy)$ is reversible with respect to $\mu(dx)$. The necessary and sufficient conditions on the acceptance probability and proposal distribution for this to be true is:

$$Q(x,dy)\alpha(x,y)\mu(dx) = Q(y,dx)\alpha(y,x)\mu(dy)\,, \tag{4.7}$$

which is an identity statement about measures [104]. In the case where $Q(x,dy) = q(x,y)dy$ and $\mu(dy) = \nu(y)dy$, it can be shown that this requirement is fulfilled by:

$$\alpha(x,y) = \min\left(1, \frac{q(y,x)\nu(y)}{q(x,y)\nu(x)}\right). \tag{4.8}$$

A quick glance at Algorithm 2.1 suggests that the scheme is a Metropolis-Hastings method, but in order to establish this, we derive the probability transition density of the proposal move in (2.1), and use this density to obtain the acceptance probability in (2.3) from (4.8). Once we have written the Metropolis integrator as a Metropolis-Hastings method, it immediately follows that it is reversible with respect to $\mu(dy) = \nu(y)dy$, and with a few additional steps, it can be shown that this stationary density is unique, and that the scheme is ergodic. The proof of the following theorem fleshes out the details in this sketch.

THEOREM 4.1. *Assume that $\nu(x)$ in (1.3) is normalizable with normalization constant $Z>0$ given in (1.7). Let $P_h(x,dy)$ denote the transition probability distribution of the integrator Algorithm 2.1. For sufficiently regular $G_h(x)$, $B_h(x)$, and $\nu(x)$, the Markov chain induced by Algorithm 2.1 preserves the equilibrium distribution $\mu(dx)=(\nu(x)/Z)dx$, and the k-step transition probability distribution of the scheme converges to μ in the total variation norm:*

$$\lim_{k\to\infty}\|P_h^k(x,\cdot)-\mu\|_{\mathrm{TV}}=0\,,\qquad \textit{for all } x\in\mathbb{R}^n\,. \tag{4.9}$$

Here, we have used the total variation (TV) distance between two probability measures, which is defined as

$$\|\mu_1-\mu_2\|_{\mathrm{TV}}=2\sup_A|\mu_1(A)-\mu_2(A)|\,, \tag{4.10}$$

where the supremum runs over all measurable sets. In particular, the total variation distance between two probability measures is two if and only if they are mutually singular.

Proof. A formula for the transition probability density of the proposal move (2.1), which we denote by $q_h(x,y)$, can be derived as follows. Consider the transformation:

$$\varphi_{(x,y)}(z)=y-2z+x-hG_h(z)\,.$$

Since $\tilde{X}_1$ appearing in (2.1) is a multivariate Gaussian with mean X_0 and covariance matrix $kTh/2M_h(X_0)=kTh/2B_h(X_0)B_h(X_0)^T$, a simple change of variables under φ implies that:

$$q_h(x,y)=\int_{\mathbb{R}^n}\frac{(\pi kTh)^{-n/2}}{\det(B_h(x))}\exp\left(-\frac{(z-x)^TM_h(x)^{-1}(z-x)}{kTh}\right)\delta(\varphi_{(x,y)}(z))dz\,. \tag{4.11}$$

A second change of variables yields,

$$\begin{aligned}q_h(x,y)=&\frac{(\pi kTh)^{-n/2}}{\det(B_h(x))}\exp\left(-\frac{(a_{(x,y)}-x)^TM_h(x)^{-1}(a_{(x,y)}-x)}{kTh}\right)\\&\times\frac{1}{|\det(D\varphi_{x,y}(a_{(x,y)}))|}\,,\end{aligned} \tag{4.12}$$

where $a_{(x,y)}$ satisfies: $\varphi_{(x,y)}(a_{(x,y)})=0$. Given $q_h(x,y)$ in (4.12) and $\nu(x)$ in (1.3), the acceptance probability (2.3) can be derived from the Metropolis-Hastings ratio (4.8). Specifically,

$$\begin{aligned}\frac{q_h(y,x)\nu(y)}{q_h(x,y)\nu(x)}=&\frac{\det(B_h(x))}{\det(B_h(y))}\\&\times\exp\left(-\frac{(a-y)^TM_h(y)^{-1}(a-y)-(a-x)^TM_h(x)^{-1}(a-x)}{hkT}-\frac{U(y)-U(x)}{kT}\right),\end{aligned}$$

where we have introduced $a=a_{(x,y)}=a_{(y,x)}$, which satisfies: $y-2a+x-hG_h(a)=0$. (Since $\varphi_{(x,y)}=\varphi_{(y,x)}$ it follows that $a_{(x,y)}=a_{(y,x)}$ and the Jacobian determinant of the transformation φ appearing in (4.12) drops out of this ratio.) The probability

transition distribution of the algorithm can now be written in the Metropolis-Hastings form (4.6):

$$P_h(x,dy) = p_h(x,y)dy + r_h(x)\delta_x(dy)\,, \tag{4.13}$$

where $p_h(x,y)$ is the off-diagonal transition density,

$$p_h(x,y) = q_h(x,y)\alpha_h(x,y)\,,$$

and $r_h(x)$ is the probability of remaining at the same point,

$$r_h(x) = 1 - \int_{\mathbb{R}^n} p_h(x,z)dz\,.$$

Thus, we conclude that the Metropolis integrator, Algorithm 2.1, is reversible, and therefore, preserves the stationary distribution $\mu(dx) = (\nu(x)/Z)dx$.

We now turn to proving that the k-step transition probability of the integrator converges to this stationary distribution in the limit $k \to \infty$. Since the target density (1.3) and the transition density (4.11) are strictly positive and smooth everywhere, the algorithm is irreducible with respect to Lebesgue measure and aperiodic; see, e.g., Lemma 1.2 of [81] and references therein. According to Corollary 2 of [103], a Metropolis-Hastings algorithm that is irreducible with respect to the same measure it is designed to preserve is positive Harris recurrent. Consequently, the algorithm is irreducible, aperiodic, and positive Harris recurrent. According to Proposition 6.3 of [85], these properties are equivalent to ergodicity of the chain. □

We conclude this section by remarking that while Theorem 4.1 assumes that $\nu(x)$ is normalizable, neither the ν-symmetry condition (1.6) – which is a special case of (4.5) – nor the Metropolis-Hastings algorithm require that $\nu(x)$ is normalizable. In fact, the identity (4.5) does not require that $\mu(dx)$ is a probability measure, and is automatically satisfied by a Metropolis-Hastings chain with transition probability distribution (4.6). In the next section, we discuss the benefits of using an integration scheme that exactly preserves the ν-symmetry condition (1.6).

5. Weak Accuracy at Constant Temperature. By casting Algorithm 2.1 as a Metropolis-Hastings algorithm, the previous section showed that the algorithm satisfies the ν-symmetry condition (1.6) by design. For the diffusion process $Y(t)$ that solves (1.1), this ν-symmetry property is equivalent to the self-adjoint property (1.5), which imposes a fairly strong constraint on the dynamics. To see this, expand the generator of the process $Y(t)$ acting on a test function $f(x)$ in (1.4) as follows:

$$(Lf)(x) = (-M(x)DU(x) + kT\,\mathrm{div}\,M(x))^T Df(x) + kT\,\mathrm{trace}(M(x)D^2f(x))\,, \tag{5.1}$$

and consider another diffusion process $\tilde{Y}(t)$ driven by the same noise, yet with a different drift:

$$d\tilde{Y} = a(\tilde{Y})dt + \sqrt{2kT}B(\tilde{Y})dW\,, \qquad \tilde{Y}(0) \in \mathbb{R}^n\,. \tag{5.2}$$

Let $\tilde{L}$ represent the generator of $\tilde{Y}(t)$ whose action on a test function is given by:

$$(\tilde{L}f)(x) = a(x)^T Df(x) + kT\,\mathrm{trace}(M(x)D^2f(x))\,.$$

If the generator of this diffusion is self-adjoint with respect to $\nu(x)$, then its drift is *uniquely determined*:

$$a(x) = -M(x)DU(x) + kT(\mathrm{div}\,M)(x)\,, \tag{5.3}$$

and is identical to the drift appearing in (1.1). To prove this statement, note that self-adjointness of the differential operator $\tilde{L}$ implies that:

$$\begin{aligned}\langle f,\tilde{L}g\rangle_\nu = \langle \tilde{L}f,g\rangle_\nu &= \int (\tilde{L}f)(x)g(x)\nu(x)dx\,,\\ &= \int a_i(x)\frac{\partial f}{\partial x_i}(x)g(x)\nu(x)dx + kT\int M_{ij}(x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x)g(x)\nu(x)dx\,,\\ &= \int\left(a_i(x) - kT\frac{\partial M_{ij}}{\partial x_j}(x) + M_{ij}(x)\frac{\partial U}{\partial x_j}(x)\right)\frac{\partial f}{\partial x_i}(x)g(x)\nu(x)dx\\ &\quad - kT\int M_{ij}(x)\frac{\partial g}{\partial x_j}(x)\frac{\partial f}{\partial x_i}(x)\nu(x)dx\,,\end{aligned}$$

where the Einstein notation has been used to indicate a sum over repeated indices. The last step in this calculation involves an integration by parts with vanishing boundary terms due to suitable boundary conditions on the test functions. Since $g(x)$ above is arbitrary, we can choose $g(x)=1$ to obtain:

$$\int\left(a_i(x) - kT\frac{\partial M_{ij}}{\partial x_j}(x) + M_{ij}(x)\frac{\partial U}{\partial x_j}(x)\right)\frac{\partial f}{\partial x_i}(x)\nu(x)dx = 0\,.$$

Arbitrariness of $f(x)$ implies (5.3). This constraint on the dynamics seems to motivate using approximations which preserve the ν-symmetry property of the continuous process, like the Metropolis integrator. We must consider, however, what precisely is gained – if anything – from a discretization that satisfies ν-symmetry especially when $\nu(x)$ is not normalizable.

We thus take this analysis one step further. Consider a process $\tilde{X}(t)$ that satisfies the SDE:

$$d\tilde{X} = \tilde{a}(\tilde{X})dt + \sqrt{2kT}B_h(\tilde{X})dW\,, \qquad \tilde{X}(0)\in\mathbb{R}^n\,, \tag{5.4}$$

for all $t\in[0,h]$. The infinitesimal generator L_h of this process is given by:

$$(L_hf)(x) = \tilde{a}(x)^T Df(x) + kT\,\mathrm{trace}(M_h(x)D^2f(x))\,,$$

where $M_h(x)=B_h(x)B_h(x)^T$. Assuming that the noise in (5.4) is a single step approximation to the exact noise appearing in (1.1) and that L_h is self-adjoint with respect to $\nu(x)$, it immediately follows from the derivation of (5.3) that $\tilde{a}(x)$ is a single step approximation to the true drift in (1.1):

$$\tilde{a}(x) = -M(x)DU(x) + kT(\mathrm{div}\,M)(x) + \mathcal{O}(h)\,.$$

We interpret this result as saying that:

noise accuracy is sufficient for accuracy of a ν-symmetric integrator.

This result enables the design of integrators that satisfy Property (P4) – avoid computing the divergence of the mobility matrix – and is one of the main advantages of using a ν-symmetric integrator to simulate a self-adjoint diffusion.

The following theorem specializes this result to a Metropolis integrator. It requires that $B_h(x)$ is chosen so that the Brownian force appearing in (1.1) is approximated to leading order. We emphasize that nowhere in the statement of this theorem is there a similar requirement on $G_h(x)$, and just to be clear, $G_h(x)$ does not need to approximate any part of the drift appearing in (1.1).

THEOREM 5.1 (Weak Accuracy). *Assume that*

$$B_h(x)B_h(x)^T = M(x) + \mathcal{O}(h) \qquad \text{for all } x \in \mathbb{R}^n .$$

For every sufficiently regular $B_h(x)$, $G_h(x)$, and time interval $T>0$, there exists a $C(T,G_h,B_h)>0$ such that the Metropolis integrator, Algorithm 2.1, satisfies:

$$|\mathbb{E}_x(f(Y(\lfloor t/h \rfloor h)) - \mathbb{E}_x(f(X_{\lfloor t/h \rfloor}))| \le C(T,G_h,B_h)h^{1/2} ,$$

for all $t \in [0,T]$.

In order to prove Theorem 5.1, we analyze the transition probability distribution of the Metropolis integrator. To this end let us fix some notation. Let $\mathcal{P}_t$ and P_h respectively denote the transition probabilities of the true solution and Metropolis integrator, whose actions on a test function $f(x)$ are given by:

$$(\mathcal{P}_t f)(x) = \mathbb{E}_x f(Y(t)) , \quad \text{with } Y(0) = x , \tag{5.5}$$

$$(P_h f)(x) = \mathbb{E}_x(f(X_1^\star)\alpha_h(X_0,X_1^\star)) + f(X_0)\mathbb{E}_x(1-\alpha_h(X_0,X_1^\star)) , \quad \text{with } X_0 = x . \tag{5.6}$$

To express the derivative of the mobility tensor, we adopt the following shorthand:

$$DM(x)(u,v,w) = \frac{\partial M_{ij}}{\partial x_k} u_k v_j w_i .$$

In this notation there are three vectorial inputs to the tensor $DM(x)$. Notice that the first input gives the direction of the derivative of $M(x)$. Also notice that since $M(x)$ is symmetric, $DM(x)(u,v,w) = DM(x)(u,w,v)$. When evaluated at two vectors, this third-order tensor returns a vector with ith component given by:

$$(DM(x)(u,v))_i = \frac{\partial M_{ij}}{\partial x_k} u_k v_j .$$

With this notation consider the following approximation to the acceptance probability:

$$\tilde{\alpha}_h(x,\xi) = 1 \wedge e^{-\frac{\sqrt{2h}}{kT}\Gamma(x,\xi)} , \tag{5.7}$$

where we have introduced:

$$\begin{aligned} \Gamma(x,\xi) &= (DU(x) + M(x)^{-1}G_h(x))^T B_h(x)\xi \\ &\quad + \frac{kT}{2}\operatorname{trace}(M(x)^{-1}DM(x)(B_h(x)\xi)) - \frac{1}{2}DM(x)(B_h(x)\xi, B_h(x)^{-T}\xi, B_h(x)^{-T}\xi) . \end{aligned} \tag{5.8}$$

Recall, $\xi \in \mathbb{R}^n$ denotes a Gaussian random vector with mean zero and covariance $\mathbb{E}(\xi_i\xi_j) = kT\delta_{ij}$. It is straightforward to show that this approximation satisfies:

$$\left\{ \tilde{\alpha}_h(x,\xi) = \alpha_h(x,y) + \mathcal{O}(h) , \quad \text{with } y = x + \sqrt{2h}B_h(x)\xi + hG_h\left(x + \sqrt{\tfrac{h}{2}}B_h(x)\xi\right) .\right.$$

Notice that when $G_h(x) = -M(x)DU(x)$ the first term in (5.8) vanishes which may lead to an improvement in the acceptance rate, especially if the mobility matrix is constant. We emphasize that this choice of $G_h(x)$ is not required by Theorem 5.1. It

is also apparent that adding $(\operatorname{div} M)(x)$ to $G_h(x)$ does not yield a similar improvement in the acceptance rate.

Proof. [Proof of Theorem 5.1] The proof focuses on deriving the following upper bound on the difference between $\mathcal{P}_h$ in (5.5) and P_h in (5.6):

$$|(\mathcal{P}_h f)(x) - (P_h f)(x)| \leq C h^{3/2}\,, \tag{5.9}$$

for both sufficiently small h and regular test functions $f(x)$. Standard results in numerical analysis for SDEs then imply the algorithm converges weakly on finite time intervals with global order 1/2; see for instance [83, Chapter 2.2]. For a detailed treatment of the technical issues involved in extending this one-step error estimate to a global error estimate when the potential force appearing in (1.1) is not globally Lipschitz, the reader is referred to [8].

Using Taylor's theorem and (5.6), one can write:

$$\mathbb{E}_x f(X_1) = f(x) + I_1 + I_2 + I_3 + I_4 + I_5 + I_6\,, \tag{5.10}$$

where we have introduced

$$\begin{aligned}
I_1 &= \mathbb{E}_x\left(Df(x)(X_1^\star - x)\right),\\
I_2 &= \frac{1}{2}\mathbb{E}_x\left(D^2 f(x)(X_1^\star - x, X_1^\star - x)\right),\\
I_3 &= \mathbb{E}_x\left(Df(x)(X_1^\star - x)(\tilde{\alpha}_h(x,\xi) - 1)\right),\\
I_4 &= \mathbb{E}_x\left(Df(x)(X_1^\star - x)(\alpha_h(x, X_1^\star) - \tilde{\alpha}_h(x,\xi))\right),\\
I_5 &= \frac{1}{2}\mathbb{E}_x\left(D^2 f(x)(X_1^\star - x, X_1^\star - x)(\alpha_h(x, X_1^\star) - 1)\right),\\
I_6 &= \frac{1}{2}\mathbb{E}_x\left(\alpha_h(x, X_1^\star)\int_0^1 (1-s)^2 D^3 f(x + s(X_1^\star - x))(X_1^\star - x)^3 ds\right),
\end{aligned}$$

where $X_1^\star$ denotes the proposal move (2.1) with $X_0 = x$. (Here we interpret $D^3 f(x) y^3$ as being the trilinear form $D^3 f(x)$ applied to the triple (y,y,y).) The terms $f(x)$, I_1, I_2 and I_3 contribute to the weak accuracy of the method. In what follows we describe how to treat these terms. It is straightforward to show that the remaining terms I_4, I_5 and I_6 in (5.10) are $\mathcal{O}(h^{3/2})$.

Notice that I_1 simplifies to,

$$\begin{aligned}
I_1 &= h\mathbb{E}_x\left(Df(x)^T G_h(\tilde{X}_1)\right),\\
&= hDf(x)^T G_h(x) + \mathcal{O}(h^{3/2})\,.
\end{aligned} \tag{5.11}$$

The leading term in I_2 involves a quadratic form in ξ, and hence,

$$\begin{aligned}
I_2 &= h\mathbb{E}\left(D^2 f(x)(B_h(x)\xi, B_h(x)\xi)\right) + \mathcal{O}(h^2)\,,\\
&= hkT\operatorname{trace}(M(x)D^2 f(x)) + \mathcal{O}(h^2)\,.
\end{aligned} \tag{5.12}$$

By referring to (5.1), let's take stock of the analysis so far. The term I_2 contributes the last term in (5.1) – the so-called 'Ito correction term.' Choosing $G_h(x) = -M(x)DU(x) + \mathcal{O}(h)$, the term I_1 accurately represents the contribution of the deterministic drift to (5.1). However, this choice of $G_h(x)$ is insufficient for accuracy since the term involving $(\operatorname{div} M)(x)$ in (5.1) has not been accounted for. While

this missing term can be added to $G_h(x)$, an explicit formula for $(\operatorname{div} M)(x)$ is not typically available in practical BD problems. This discussion foretells the importance of the term I_3. In fact, it will be shown that – independent of the precise form of $G_h(x)$ – the sum of I_1 and I_3 represent the leading order effect of the drift in (5.1).

Now we show how to estimate I_3. To simplify these calculations, introduce

$$\begin{cases} c(z,z,z) = \frac{1}{2}\sqrt{2hkT}DM(B_h z, B_h^{-T} z, B_h^{-T} z)\,, \\ b(z) = -\sqrt{\frac{2h}{kT}}(DU + M^{-1}G_h)^T B_h z - \frac{\sqrt{2hkT}}{2}\operatorname{trace}(M(x)^{-1}DM(B_h z))\,, \\ a = \sqrt{2hkT}B_h^T Df\,, \end{cases} \tag{5.13}$$

where $z \in \mathbb{R}^n$. For the sake of clarity, the dependence of $Df(x)$, $G_h(x)$, $M(x)$ and $B_h(x)$ on x is suppressed. In terms of these variables I_3 can be written as,

$$\begin{aligned} I_3 &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} a^T z e^{-\frac{|z|^2}{2}}\left(1 \wedge e^{b(z)+c(z,z,z)} - 1\right)dz + \mathcal{O}(h^{3/2})\,, \\ &= -(2\pi)^{-n/2}\int_{\mathbb{R}^n} \operatorname{div}_z(e^{-\frac{|z|^2}{2}}a)\left(1 \wedge e^{b(z)+c(z,z,z)} - 1\right)dz + \mathcal{O}(h^{3/2})\,. \end{aligned}$$

Let $\Omega = \{z \in \mathbb{R}^n \,|\, b(z) + c(z,z,z) \le 0\}$, and reduce this integral into,

$$\begin{aligned} I_3 &= -(2\pi)^{-n/2}\int_{\Omega} \operatorname{div}_z(e^{-\frac{|z|^2}{2}}a)\left(e^{b(z)+c(z,z,z)} - 1\right)dz + \mathcal{O}(h^{3/2})\,, \\ &= (2\pi)^{-n/2}\int_{\Omega} e^{-\frac{|z|^2}{2}}(b(a) + c(a,z,z) + 2c(z,a,z))\left(e^{b(z)+c(z,z,z)}\right)dz + \mathcal{O}(h^{3/2})\,. \end{aligned}$$

The last step applies the following multidimensional integration by parts formula

$$\int_{\partial\Omega} g_1(z)\vec{g}_2(z)^T\vec{\nu}dz = \int_{\Omega} g_1(z)(\operatorname{div}_z \vec{g}_2)(z)dz + \int_{\Omega}\vec{g}_2(z)^T(\nabla_z g_1)(z)dz\,,$$

by setting: $g_1(z) = e^{b(z)+c(z,z,z)} - 1$ and $\vec{g}_2(z) = e^{-\frac{|z|^2}{2}}a$. Here, $\partial\Omega = \{z \,|\, b(z) + c(z,z,z) = 0\}$ and $\vec{\nu}$ is the outward pointing normal to the region Ω. Since the map $z \mapsto b(a) + c(a,z,z) + 2c(z,a,z)$ is symmetric (or even: $f(-z) = f(z)$) and $z \mapsto b(z) + c(z,z,z)$ is anti-symmetric (or odd: $f(-z) = -f(z)$), it follows that

$$\begin{aligned} I_3 &= (2\pi)^{-n/2}\int_{\Omega} e^{-\frac{|z|^2}{2}}(b(a) + c(a,z,z) + 2c(z,a,z))dz + \mathcal{O}(h^{3/2}) \\ &\quad + (2\pi)^{-n/2}\int_{\Omega} e^{-\frac{|z|^2}{2}}(b(a) + c(a,z,z) + 2c(z,a,z))(e^{b(z)+c(z,z,z)} - 1)dz\,, \\ &= \frac{1}{2}(2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-\frac{|z|^2}{2}}(b(a) + c(a,z,z) + 2c(z,a,z))dz + \mathcal{O}(h^{3/2}) \\ &\quad + (2\pi)^{-n/2}\int_{\Omega} e^{-\frac{|z|^2}{2}}(b(a) + c(a,z,z) + 2c(z,a,z))(e^{b(z)+c(z,z,z)} - 1)dz\,, \\ &= \frac{1}{2}(b(a) + c_{ijj}a_i + 2c_{iji}a_j) + \mathcal{O}(h^{3/2})\,. \end{aligned}$$

Substituting (5.13) back into this expression and simplifying yields,

$$I_3 = h(kT\operatorname{div} M - MDU - G_h)^T Df + \mathcal{O}(h^{3/2})\,. \tag{5.14}$$

The estimates (5.11), (5.12) and (5.14) imply that,

$$(P_h f)(x) = f(x) + h(Lf)(x) + \mathcal{O}(h^{3/2})\,,$$

which agrees with an Ito-Taylor expansion of $(\mathcal{P}_h f)(x)$ up to $\mathcal{O}(h^{3/2})$, and hence, the scheme has the desired single step accuracy. □

6. Small Noise Limit.

6.1. General Principle. For every $t>0$, the solution to (1.1) with non-random initial condition $Y(0)=x\in\mathbb{R}^n$ converges in the small noise limit to the solution of the following ordinary differential equation:

$$\dot{\mathcal{Y}}=-M(\mathcal{Y})DU(\mathcal{Y})\,,\quad \mathcal{Y}(0)=x\in\mathbb{R}^n\,, \tag{6.1}$$

in the L^2-norm, meaning that:

$$\mathbb{E}_x\left\{|Y(t)-\mathcal{Y}(t)|^2\right\}\to 0\,,\quad \text{as } kT\to 0\,,$$

which we denote as $Y(t)\overset{L^2}{\to}\mathcal{Y}(t)$. At the same time, the proposal move in (2.1) satisfies,

$$X_1^\star\overset{L^2}{\to}X_0+hG_h(X_0)\,. \tag{6.2}$$

Higher order accuracy of Algorithm 2.1 necessitates that the deterministic update (6.2) approximates $\mathcal{Y}(h)$ to higher-order. However, this condition is not sufficient because the actual update in Algorithm 2.1 involves a Bernoulli random variable (2.2), which imposes an additional requirement for deterministic accuracy on the acceptance probability $\alpha_h(x,y)$ in (2.3):

$$\alpha_h(X_0,X_1^\star)\to 1\,,\quad \text{as } kT\to 0\,. \tag{6.3}$$

Otherwise, the proposal move will not be accepted, and consequently, the integrator will be dynamically inaccurate. Finding necessary and sufficient conditions for deterministic accuracy motivates an asymptotic analysis of the Metropolis integrator. This analysis reveals the following asymptotic relationship between $\alpha_h(x,y)$ and a deterministic function $f_{kT}(x,h)$,

$$\alpha_h(X_0,X_1^\star)\sim 1\wedge\exp\left(-\frac{1}{kT}f_{kT}(X_0,h)\right)\quad\text{as}\quad kT\to 0\,, \tag{6.4}$$

where we have introduced

$$\begin{cases} f_{kT}(x,h)=U(x+hG_h(x))-U(x)+hG_h(x)^TM_h(x+hG_h(x))^{-1}G_h(x)\,,\\ M_h(x)=B_h(x)B_h(x)^T\,. \end{cases} \tag{6.5}$$

From this asymptotic relation, it follows that:

$$f_{kT}(X_0,h)<0\Longrightarrow\alpha_h(X_0,X_1^\star)\to 1\,,$$

and the Metropolis integrator acquires the deterministic accuracy of its proposal move.

Thus, our design philosophy is to pick $G_h(x)$ and $B_h(x)$ in Algorithm 2.1 so that:

(1) the update $x\mapsto x+hG_h(x)$ generates a higher order accurate approximation to the solution of (6.1); and simultaneously,
(2) $f_{kT}(x,h)$ is negative definite.

We emphasize that the time step size h is held fixed in the above statements. These ideas are formulated precisely in the following theorem.

THEOREM 6.1. *Consider the solutions $Y(t)$ and X produced by (1.1) and Algorithm 2.1, respectively. Let $\mathcal{Y}(t)$ denote the exact solution to (6.1) with non-random initial condition $\mathcal{Y}(0)=x$. Let $\mathcal{X}$ denote the discrete path defined by:*

$$\mathcal{X}_{k+1}=\mathcal{X}_k+hG_h(\mathcal{X}_k)\,,\qquad \mathcal{X}_0=x\in\mathbb{R}^n\,. \tag{6.6}$$

Assume that $G_h(x)$ and $B_h(x)$ in Algorithm 2.1 are selected so that the following hold.

(A1) *For every $T>0$ there exists $\tilde{C}(T)>0$ such that:*

$$|\mathcal{Y}(\lfloor t/h\rfloor h)-\mathcal{X}_{\lfloor t/h\rfloor}|\le\tilde{C}(T)h^p\,,\qquad \text{for all } t\in[0,T]\,.$$

(A2) *The function $f_{kT}(x,h)$ in (6.5) is negative definite for all $x\in\mathbb{R}^n$.*

Then there exists a constant $C(T)>0$ such that

$$\lim_{kT\to 0}(\mathbb{E}_x|Y(\lfloor t/h\rfloor h)-X_{\lfloor t/h\rfloor}|^2)^{1/2}\le C(T)h^p\,,$$

for every $t\in[0,T]$ and initial condition $x\in\mathbb{R}^n$.

A Metropolis integrator that satisfies these assumptions has the interesting property that even when the noise is infinitesimal the scheme will be both accurate and ν-symmetric. In practice, to meet Assumption (A1) we select $G_h(x)$ and $B_h(x)$ in Algorithm 2.1 to be n-stage Runge-Kutta methods, and pick the parameters in these methods to ensure that both Assumption (A1) and (A2) in Theorem 6.1 are satisfied.

Proof. By the triangle inequality, the desired error can be decomposed into

$$\begin{aligned}(\mathbb{E}_x|Y(\lfloor t/h\rfloor h)-X_{\lfloor t/h\rfloor}|^2)^{1/2}\le\;& \overbrace{|\mathcal{Y}(\lfloor t/h\rfloor h)-\mathcal{X}_{\lfloor t/h\rfloor}|}^{\le\tilde{C}(T)h^p\text{ by Assumption (A1)}}\\ &+\underbrace{(\mathbb{E}_x|Y(\lfloor t/h\rfloor h)-\mathcal{Y}(\lfloor t/h\rfloor h)|^2)^{1/2}}_{\overset{L^2}{\to}0\text{ by Lemma 8.1}}+\underbrace{(\mathbb{E}_x|X_{\lfloor t/h\rfloor}-\mathcal{X}_{\lfloor t/h\rfloor}|^2)^{1/2}}_{\substack{\overset{L^2}{\to}0\text{ by Lemma 8.2}\\ \text{using Assumption (A2)}}}.\end{aligned} \tag{6.7}$$

By Assumption (A1), $\mathcal{X}$ is a *pth*-order approximation of $\mathcal{Y}(t)$, and so, the first term in the upper bound in (6.7) is bounded by $C(T)h^p$. The remaining terms vanish in the small noise limit. Indeed, by applying the Ito-Taylor formula to $|Y(t)-\mathcal{Y}(t)|^2$, and using Gronwall's lemma, it can be shown that

$$\mathbb{E}_x|Y(t)-\mathcal{Y}(t)|^2\le kTC_2\exp(C_1T)\,,\quad\forall\quad t\in[0,T]\,. \tag{6.8}$$

From this somewhat crude estimate, it follows that the second term appearing in the upper bound in (6.7) vanishes in the small noise limit. The third term in this upper bound vanishes because the mean-squared acceptance probability equals one in the small noise limit as a consequence of Assumption (A2). More details are provided in Lemmas 8.1 and 8.2 provided in the Appendix. □

6.2. Necessity of Second-Order Accuracy. Here we theoretically show what can go wrong with the simple choice:

$$G_h(x) = -M(x)DU(x)\,.$$

This $G_h(x)$ corresponds to a forward Euler approximation to (6.1). For the sake of clarity, assume that the mobility is constant and set $B_h = B$. Using Taylor's theorem (6.5) can be written as:

$$f_{kT}(x,h) = h^2\left(\int_0^1 (1-s)D^2U(x_s)(MDU(x),MDU(x))ds\right)\,,\quad x_s = x - shMDU(x)\,.$$

From this expression it is clear that the function $f_{kT}(x,h)$ can become positive in regions where $U(x)$ is convex, no matter how small the time step size h is made. In this case the rejection rate of the Metropolis integrator, Algorithm 2.1, will tend to one in the small noise limit. In other words, even though the proposal move is first-order accurate, it is never accepted, and hence, the Metropolis integrator fails to be first-order accurate. This prediction was verified in the numerical experiments provided in §3, and motivates using proposal moves that are deterministically second or higher-order accurate.

6.3. Second-Order Accuracy. We will now use Theorem 6.1 to prove that the Metropolis integrator, Algorithm 2.1, based on the following choice of $G_h(x)$ and $B_h(x)$ acquires second-order accuracy in the small noise limit. Let $G_h(x)$ be a two-stage Runge-Kutta method:

$$\begin{cases} G_h(x) = -b_1M(x)DU(x) - b_2M(x)DU(x_1) \\ \qquad\qquad -b_3M(x_1)DU(x) - b_4M(x_1)DU(x_1)\,; \\ x_1 = x - ha_{21}M(x)DU(x)\,, \end{cases} \tag{6.9}$$

with parameters b_1, b_2, b_3, b_4, and a_{21}. Second-order accuracy requires that:

$$b_1 + b_2 + b_3 + b_4 = 1\,,\quad (b_2+b_4)a_{21} = 1/2\,,\quad b_3 = b_2\,, \tag{6.10}$$

and so, we can choose a_{21} and b_4 as free parameters. Define the matrix $B_h(x)$ through:

$$\begin{cases} B_h(x)B_h(x)^T = d_1M(x) + d_2M(x_1)\,; \\ x_1 = x + hc_{21}M(x)DU(x)\,. \end{cases} \tag{6.11}$$

First-order accuracy at constant temperature requires that $d_1 + d_2 = 1$ which leaves d_2 and c_{21} as free parameters. To summarize, choosing $G_h(x)$ and $B_h(x)$ given by (6.9) and (6.11), respectively, leads to a Metropolis integrator that satisfies Assumption (A1) of Theorem 6.1 with $p=2$. The following assumption on $U(x)$ will be sufficient to select these free parameters so that Assumption (A2) is also satisfied.

ASSUMPTION 6.1. *Assume the potential energy function $U:\mathbb{R}^n \to \mathbb{R}$ satisfies:*

$$\det D^2U(y) \neq 0\,,$$

for all $y \in \mathbb{R}^n$.

With this assumption we are now in position to prove the following statement which optimizes the free parameters appearing in (6.9) and (6.11) so that Assumption (A2) is satisfied.

THEOREM 6.2. *Consider the solution $\mathcal{Y}(t)$ to (6.1) with the non-random initial condition $x\in\mathbb{R}^n$ and assume that $U(x)$ satisfies Assumption 6.1. Let X denote the approximation produced by Algorithm 2.1 with $G_h(x)$ in (6.9) and $B_h(x)$ in (6.11) with parameter values:*

$$b_1=5/8\,,\quad b_2=b_3=-3/8\,,\quad b_4=9/8\,,\quad d_1=1/4\,,\quad d_2=3/4\,,\quad a_{21}=c_{21}=2/3\,. \tag{6.12}$$

For every $T>0$ there exists a constant $C(T)>0$ such that

$$\lim_{kT\to 0}(\mathbb{E}_x|Y(\lfloor t/h\rfloor h)-X_{\lfloor t/h\rfloor}|^2)^{1/2}\le C(T)h^2\,,$$

for every $t\in[0,T]$.

Proof. According to Assumptions (A1) and (A2) of Theorem 6.1, to prove this statement it suffices to show that Assumption (A2) is satisfied by Algorithm 2.1 operated with $G_h(x)$ in (6.9), $B_h(x)$ in (6.11), and the parameter values given in (6.12). To illustrate how we do this, let us assume for simplicity that: M is constant and (6.10) holds. Under these assumptions (6.5) simplifies to:

$$\begin{aligned} f_{kT}(x,h)=&-\frac{h^3}{4}\left(M-\frac{h}{2}MA(x)M\right)(A(x)MDU(x),A(x)MDU(x)) \\ &+\frac{h^2}{2}(B(x)-A(x))(G_h(x),G_h(x))\,, \end{aligned} \tag{6.13}$$

where we have introduced:

$$A(x)=\int_0^1 D^2U(x+sha_{21}k_1)ds\,,\quad\text{and}\quad B(x)=2\int_0^1(1-s)D^2U(x+shG_h)ds\,. \tag{6.14}$$

Assumption (A2) in Theorem 6.1 requires that the leading order term arising in (6.13) is strictly negative and dominates the indefinite terms appearing at higher order. (In situations where $D^2U(x)MDU(x)=0$, this might not be the case no matter how small h is made, but such situations are excluded by requiring Assumption 6.1 to hold.) If $U(x)$ is quadratic, then $B=A$, the second term in (6.13) vanishes, and $f_{kT}(x,h)$ is negative definite if $h<2/\|MA\|$. In the non-quadratic case, $\|MA\|$ in this condition is replaced by a local Lipschitz constant on the vector field $MDU(x)$ and the second term in (6.13) is bounded as follows. A Taylor expansion of $B(x)-A(x)$ yields:

$$B(x)-A(x)=h\left(\frac{1}{3}-\frac{a_{21}}{2}\right)D^3U(x)(MDU(x))+O(h^2)\,.$$

Choosing $a_{21}=2/3$ ensures that the second term in (6.13) is $O(h^4)$, and therefore, that $f_{kT}(x,h)$ is negative definite for h small enough. This choice of a_{21} corresponds to the so-called Ralston Runge-Kutta scheme [87].

When the mobility matrix is not constant, the choice $M_h(x)=M(x)$ does not always lead to a negative semi-definite $f_{kT}(x,h)$ function. This motivates the approximation (6.11). We will choose the remaining free parameters to remove the indefinite terms appearing at $\mathcal{O}(h^2)$ and $\mathcal{O}(h^3)$ that involve the derivatives of the mobility. By eliminating these terms, $f_{kT}(x,h)$ will be dominated by the leading term appearing in (6.13), and hence, the argument presented when the mobility is constant carries

over to the non-constant case. In addition to the notation introduced in the proof of Theorem 5.1, we use the following shorthand for the second derivative of the mobility:

$$D^2M(x)(s,u,v,w)=\frac{\partial M_{ij}}{\partial x_k\partial x_l}s_lu_kv_jw_i\,.$$

Symmetry of second derivatives implies that $D^2M(x)(s,u,v,w)=D^2M(x)(u,s,v,w)$. For brevity, if a function is evaluated at x, the function's input is omitted.

Based on the preceding discussion, assume the second order conditions (6.10) hold and $a_{21}=2/3$. These conditions leave b_4, d_2, and c_{21} as free parameters. Rewrite $f_{kT}(x,h)$ as,

$$\begin{aligned}f_{kT}(x,h)=&hG_h^TM^{-1}G_h+U(x+hG_h)-U+hG_h^TM^{-1}(M-M_h(x+hG_h))M^{-1}G_h\\&+hG_h^TM^{-1}(M-M_h(x+hG_h))M^{-1}(M-M_h(x+hG_h))M^{-1}G_h+\mathcal{O}(h^4)\,.\end{aligned}\tag{6.15}$$

A Taylor expansion of the mobility matrix implies $G_h(x)$ can be written as,

$$\begin{aligned}G_h=&-MDU+\frac{h}{2}(MAMDU+DM(MDU,DU))\\&-\frac{h^2}{6}D^2M(MDU,MDU,DU)-h^2b_4a_{21}^2DM(MDU,AMDU)+\mathcal{O}(h^3)\,.\end{aligned}\tag{6.16}$$

Substituting (6.16) into (6.15) yields,

$$\begin{aligned}f_{kT}(x,h)=&-\frac{h^3}{4}\left(M-\frac{h}{2}MAM\right)(AMDU,AMDU)+\frac{h^2}{2}(B-A)(G_h,G_h)\\&-\frac{h^2}{2}DM(MDU,DU,DU)+\frac{h^3}{4}DM(MDU,DU)^TM^{-1}DM(MDU,DU)\\&+h^3b_4a_{21}^2DM(MDU,DU,AMDU)+\frac{h^3}{6}D^2M(MDU,MDU,DU,DU)\\&+hG_h^TM^{-1}(M-M_h(x+hG_h))M^{-1}G_h\\&+hG_h^TM^{-1}(M-M_h(x+hG_h))M^{-1}(M-M_h(x+hG_h))M^{-1}G_h+\mathcal{O}(h^4)\,.\end{aligned}\tag{6.17}$$

Observe that the leading term in (6.17) includes (6.13) and additional terms involving the derivatives of the mobility. Using (6.11) expand $M_h(x+hG_h)$ as:

$$M_h(x+hG_h)=M+h(d_2c_{21}-1)DM(MDU)+\mathcal{O}(h^2)\,.\tag{6.18}$$

Substitute (6.18) into (6.17) to obtain:

$$\begin{aligned}f_{kT}(x,h)=&-\frac{h^3}{4}\left(M-\frac{h}{2}MAM\right)(AMDU,AMDU)+\frac{h^2}{2}(B-A)(G_h,G_h)\\&+h^2(\frac{1}{2}-d_2c_{21})DM(MDU,DU,DU)+\mathcal{O}(h^3)\,.\end{aligned}\tag{6.19}$$

If $d_2c_{21}=1/2$ the derivatives of the mobility up to $\mathcal{O}(h^3)$ are eliminated in this expression. Under this condition

$$M_h(x+hG_h)=M-\frac{h}{2}DM(MDU)+c_{21}\frac{h^2}{4}D^2M(MDU,MDU)+\mathcal{O}(h^3)\,.\tag{6.20}$$

Substituting (6.20) into (6.17) yields:

$$f_{kT}(x,h) = -\frac{h^3}{4}\left(M - \frac{h}{2}MAM\right)(AMDU, AMDU) + \frac{h^2}{2}(B-A)(G_h, G_h) \qquad (6.21)$$
$$+ h^3(b_4 a_{21}^2 - \frac{1}{2})DM(MDU, DU, AMDU)$$
$$+ h^3(\frac{1}{6} - \frac{c_{21}}{4})D^2M(MDU, MDU, DU, DU) + \mathcal{O}(h^4)\,.$$

Eliminating the terms up to $\mathcal{O}(h^3)$ requires the conditions: $b_4 a_{21}^2 = 1/2$ and $c_{21} = 2/3$. The above conditions uniquely specify the parameters appearing in (6.9) and (6.11). Moreover, they imply that the leading term in (6.21) is the same as (6.13), and by the same dominant asymptotic argument, $f_{kT}(x,h)$ is negative semi-definite when h is small enough. □

6.4. Third-Order Accuracy. Define $G_h(x)$ to be the following three-stage Runge-Kutta method:

$$\begin{cases} G_h(x) = -b_1 M(x)DU(x) - b_2 M(x_1)DU(x_1) - b_3 M(x_2)DU(x_2)\,; \\ x_1 = x - h a_{21} M(x)DU(x)\,; \\ x_2 = x - h a_{31} M(x)DU(x) - h a_{32} M(x_1)DU(x_1)\,. \end{cases} \qquad (6.22)$$

Third-order accuracy requires that:

$$\begin{cases} b_1 + b_2 + b_3 = 1\,, \\ b_2 a_{21} + b_3(a_{31} + a_{32}) = 1/2\,, \\ b_2 a_{21}^2 + b_3(a_{31} + a_{32})^2 = 1/3\,, \\ a_{21} a_{32} b_3 = 1/6\,, \end{cases} \qquad (6.23)$$

and so, we can choose a_{31} and a_{32} as free parameters. In a corresponding manner, the covariance matrix of the approximation to the noise is defined as:

$$\begin{cases} B_h(x)B_h(x)^T = d_1 M(x) + d_2 M(x_1) + d_3 M(x_2)\,; \\ x_1 = x + h c_{21} M(x)DU(x)\,; \\ x_2 = x + h c_{31} M(x)DU(x) + h c_{32} M(x_1)DU(x_1)\,. \end{cases} \qquad (6.24)$$

First-order accuracy at constant temperature requires that $d_1 + d_2 + d_3 = 1$ which leaves five free parameters. Interestingly, every third-order Runge-Kutta method yields a deterministically third-order Metropolis algorithm provided the parameters in (6.24) are coupled to the parameters in (6.22) in the manner prescribed in the following theorem.

THEOREM 6.3. *Let X denote the approximation Algorithm 2.1 with $G_h(x)$ given by* (6.22)*, $B_h(x)$ given by* (6.24)*, and parameter values satisfying:*

$$d_1 = b_1\,,\quad d_2 = b_2\,,\quad d_3 = b_3\,,\quad c_{21} = a_{21}\,,\quad c_{31} = a_{31}\,,\quad c_{32} = a_{32}\,. \qquad (6.25)$$

For every $(b_1, b_2, b_3, a_{21}, a_{31}, a_{32})$ satisfying (6.23)*, and for every $T > 0$, there exists a constant $C(T) > 0$ such that*

$$\lim_{kT \to 0} (\mathbb{E}_x |Y(\lfloor t/h \rfloor h) - X_{\lfloor t/h \rfloor}|^2)^{1/2} \le C(T) h^3\,, \qquad \text{for every } t \in [0,T]\,.$$

Proof. According to Theorem 6.1, to prove this statement it suffices to show that Assumption (A2) is met by Algorithm 2.1 operated using $G_h(x)$ in (6.22), $B_h(x)$ in (6.24), and parameter values satisfying (6.23) and (6.25). A Taylor expansion of $G_h(x)$ yields:

$$G_h = -MDU + \frac{h}{2}(MAMDU + DM(MDU,DU)) + R_1 h^2 + \mathcal{O}(h^3)\,. \tag{6.26}$$

where we have introduced

$$\begin{aligned} R_1 = &-\frac{h^2}{6}MA_2MA_1MDU - \frac{h^2}{6}D^2M(MDU,MDU,DU)\,, \\ &-\frac{h^2}{2}DM(MDU,D^2UMDU) - \frac{h^2}{6}DM(MD^2UMDU + DM(MDU,DU),DU)\,, \end{aligned} \tag{6.27}$$

and the following matrices:

$$\begin{aligned} A_1 &= \int_0^1 D^2U(x - sha_{21}MDU)ds\,, \\ A_2 &= \int_0^1 D^2U(x - sh(a_{31}MDU + a_{32}M(x_1)DU(x_1)))ds\,, \quad x_1 = x - ha_{21}MDU\,, \\ A &= 2(b_2a_{21}A_1 + b_3(a_{31}+a_{32})A_2)\,. \end{aligned}$$

Observe that (6.26) matches (6.16) up to $\mathcal{O}(h^2)$. It follows that

$$\begin{aligned} f_{kT}(x,h) = &-\frac{h^3}{4}\left(M - \frac{h}{2}MAM\right)(AMDU,AMDU) + \frac{h^2}{2}(B-A)(G_h,G_h) \\ &-\frac{h^2}{2}DM(MDU,DU,DU) + \frac{h^3}{4}DM(MDU,DU)^TM^{-1}DM(MDU,DU) \\ &-h^3R_1^TDU + hG_h^TM^{-1}(M - M_h(x+hG_h))M^{-1}G_h \\ &+hG_h^TM^{-1}(M - M_h(x+hG_h))M^{-1}(M - M_h(x+hG_h))M^{-1}G_h + \mathcal{O}(h^4)\,, \end{aligned} \tag{6.28}$$

where as before $B = 2\int_0^1 (1-s)D^2U(x+shG_h)ds$. In the case when U is quadratic:

$$A = B = A_1 = A_2\,,$$

and (6.28) simplifies to:

$$f_{kT} = -h^3\left(\frac{1}{12}M - \frac{h}{8}MAM\right)(AMDU,AMDU)\,, \quad \text{(if } U \text{ is quadratic)}\,.$$

This expression is negative if $h < 2/31/\|MA\|$. In the non-quadratic case, and as before, the norm of MA is replaced with a local Lipschitz constant on DU.

Assuming (6.25) an expansion of M_h yields:

$$\begin{aligned} M_h(x+hG_h) = &M - \frac{h}{2}DM(MDU) \\ &+\frac{h^2}{6}\left(D^2M(MDU,MDU) + DM(MD^2UMDU + DM(MDU,DU))\right)\,. \end{aligned} \tag{6.29}$$

Substituting (6.27) and (6.29) into (6.28) and simplifying yields:

$$f_{kT}(x,h) = -\frac{h^3}{4}\left(M - \frac{h}{2}MAM\right)(AMDU, AMDU) + \frac{h^2}{2}(B-A)(G_h, G_h) \quad (6.30)$$
$$+ \frac{h^3}{6} DU^T M A_2 M A_1 M DU\,.$$

This function is negative semi-definite provided h is small enough. □

Although all of these third-order Runge Kutta methods lead to a negative semi-definite $f_{kT}(x,h)$ function in the deterministic limit for sufficiently small time step, the following theorem specifies an optimal choice of free parameters.

PROPOSITION 6.1. *Let X denote the approximation induced by Algorithm 2.1 with $G_h(x)$ in* (6.22) *with parameter values given by*

$$b_1 = 1/6\,,\quad b_2 = 2/3\,,\quad b_3 = 1/6\,,\quad a_{21} = 1/2\,,\quad a_{31} = -1\,,\quad a_{32} = 2\,. \quad (6.31)$$

When M is constant, and h sufficiently small, this choice of parameters minimizes the indefinite remainder terms in (6.28).

Proof. When M is constant, (6.28) simplifies to:

$$f_{kT}(x,h) = -\frac{h^3}{4}\left(M - \frac{h}{2}MAM\right)(AMDU, AMDU) + \frac{h^2}{2}(B-A)(G_h, G_h)\,. \quad (6.32)$$

A Taylor expansion of the indefinite remainder in (6.32) yields

$$\begin{aligned} B - A = {} & h(b_2 a_{21}^2 + b_3(a_{31}+a_{32})^2 - \frac{1}{3})D^3U(x)(MDU) \\ & + h^2(\frac{1}{6} - \frac{a_{31}+a_{32}}{6})D^3U(x)(D^2UMDU) \\ & + h^2(\frac{1}{12} - \frac{b_2 a_{21}^3}{3} - \frac{b_3(a_{31}+a_{32})^3}{3})D^4U(x)(MDU, MDU) + O(h^3)\,. \end{aligned}$$

The third-order conditions (6.23) imply the $\mathcal{O}(h)$ term in $B-A$ vanishes. Within this two-parameter family of schemes it appears that one is optimal. In particular, to eliminate the $O(h^2)$ error in $B-A$, the following additional conditions must be satisfied:

$$\begin{cases} a_{31} + a_{32} = 1\,, \\ b_2 a_{21}^3 + b_3(a_{31}+a_{32})^3 = \frac{1}{4}\,. \end{cases} \quad (6.33)$$

The six equations in (6.23) and (6.33) uniquely specify a three stage Runge-Kutta scheme. It is given by the following choice of coefficients:

$$b_1 = 1/6\,,\quad b_2 = 2/3\,,\quad b_3 = 1/6\,,\quad a_{21} = 1/2\,,\quad a_{31} = -1\,,\quad a_{32} = 2\,. \quad (6.34)$$

This scheme is known as Kutta's third order method [62]. □

7. Conclusion. This paper presented a ν-symmetric integrator for self-adjoint diffusions. This integrator is a Metropolis-Hastings algorithm with an optimized Runge-Kutta based proposal move and target density set equal to $\nu(x)$. Since the Metropolis-Hastings ratio does not involve the normalization constant of $\nu(x)$, the algorithm is well-defined even in situations where this density is not normalizable (i.e., it is not a probability density). In the context of non-normalizable $\nu(x)$, the paper proved that the algorithm is weakly accurate for finite noise as a direct consequence of its ν-symmetry and its consistent approximation of the noise. Through an asymptotic analysis of the integrator's rejection rate in the small noise limit, second-order deterministic accuracy was also established. For applications to BD, the scheme shares the nice properties of the Fixman scheme (explicitness, finite-time accuracy, second-order deterministic accuracy and avoids divergence of the mobility). In addition, it is ergodic if $\nu(x)$ is normalizable. These properties imply that the scheme is able to stably calculate dynamic quantities at reasonable time step sizes and generate long trajectories. The paper verified these claims on a collection of low-dimensional toy problems and a more realistic simulation of DNA in an unbounded solvent. These features make the scheme appealing for the simulations of realistic BD applications, which will be the topic of future investigations.

8. Appendix. Here we provide Lemmas required in the proof of Theorem 6.1.

LEMMA 8.1. *Consider the solutions $Y(t)$ to (1.1) and $\mathcal{Y}(t)$ to (6.1) with initial condition: $Y(0)=\mathcal{Y}(0)=x\in\mathbb{R}^n$. For every $T>0$, then*

$$Y(t)\xrightarrow{L^2}\mathcal{Y}(t)\,,\quad \textit{for all } t\in[0,T] \textit{ and } x\in\mathbb{R}^n\,.$$

Proof. By the Ito-Taylor formula,

$$\begin{aligned}|Y(t)-\mathcal{Y}(t)|^2 \le{}& 2\int_0^t \langle Y(s)-\mathcal{Y}(s), -M(Y(s))DU(Y(s))+M(\mathcal{Y}(s))DU(\mathcal{Y}(s))\rangle ds\\ &+2kT\int_0^t \langle Y(s)-\mathcal{Y}(s), \operatorname{div} M(Y(s))\rangle ds + 2kT\int_0^t \operatorname{trace} M(Y(s))ds\\ &+2\sqrt{2kT}\int_0^t \langle Y(s)-\mathcal{Y}(s), B(Y(s))dW(s)\rangle\end{aligned}$$

Using bounds on $M(x)$ and $U(x)$ and their derivatives, and taking the expectation of both sides of this inequality, it follows from Gronwall's Lemma that

$$\begin{aligned}\mathbb{E}_x|Y(t)-\mathcal{Y}(t)|^2 &\le C_1\int_0^t \mathbb{E}_x|Y(s)-\mathcal{Y}(s)|^2 ds + kT\,C_2\,T\,,\\ &\le kT\exp(C_1T)\,C_2\,T\,.\end{aligned}$$

Passing to the small noise limit produces the desired L^2 convergence. □

LEMMA 8.2. *Let X and $\mathcal{X}$ denote the numerical solutions generated by Algorithm 2.1 and (6.6) respectively, with: $X_0=\mathcal{X}_0=x\in\mathbb{R}^n$. If the $f_{kT}(x,h)$ function of Algorithm 2.1 satisfies Assumption (A2) in Theorem 6.1, then for every $T>0$,*

$$X_{\lfloor t/h\rfloor}\xrightarrow{L^2}\mathcal{X}_{\lfloor t/h\rfloor}\,,\quad \textit{for all } t\in[0,T] \textit{ and } x\in\mathbb{R}^n\,.$$

Proof. The proof goes by induction over the number of steps, so it suffices to consider the difference between a single step of both schemes conditioned on the previous time step:

$$\begin{aligned}\mathbb{E}|X_1-\mathcal{X}_1|^2 &= \mathbb{E}|X_1^\star-\mathcal{X}_1|^2\,\alpha_h(X_0,X_1^\star)+|X_0-\mathcal{X}_1|^2\,\mathbb{E}\,(1-\alpha_h(X_0,X_1^\star))\,,\\ &\le (1+C_1\,h)\,|X_0-\mathcal{X}_0|^2+C_2\,h^{3/2}\,kT+C_3\,h\,\mathbb{E}\,(1-\alpha_h(X_0,X_1^\star))\,.\end{aligned}$$

The desired L^2 convergence follows from applying Discrete Gronwall's Lemma to this recurrence inequality, passing to the small noise limit, and invoking Lemma 8.3. ∎

LEMMA 8.3. *Let $\alpha_h(X_0,X_1^\star)$ denote the acceptance probability in* (2.3)*. If the $f_{kT}(x,h)$ function of Algorithm 2.1 satisfies Assumption (A2) in Theorem 6.1, then*

$$\mathbb{E}\,\alpha_h(X_0,X_1^\star)\to 1\,,\qquad \text{as } kT\to 0\,,$$

for all $X_0\in\mathbb{R}^n$.

Proof. Let $\mathcal{X}_1=X_0+hG_h(X_0)$ and $M_h(x)=B_h(x)B_h(x)^T$. Expand the exponent of $\alpha_h(X_0,X_1^\star)$ in powers of kT to obtain:

$$\alpha_h(X_0,X_1^\star)= 1\wedge\exp\left(-\frac{1}{kT}\left[f_{kT}(X_0,h)+\frac{kT}{2}\log\left(\frac{\det M_h(\mathcal{X}_1)}{\det M_h(X_0)}\right)+g_{kT}(X_0,\xi,h)\right]\right) \tag{8.1}$$

where $f_{kT}(x,h)$ is given by (6.5) and $g_{kT}(x,\xi,h)$ is exactly:

$$\begin{aligned}g_{kT}(x,\xi,h) &= U(X_1^\star)-U(\mathcal{X}_1)\\ &\quad+\sqrt{2h}(B_h(X_0)\xi)^T M_h(\mathcal{X}_1)^{-1}G_h(\mathcal{X}_1)\\ &\quad+h(G_h(\tilde{X}_1)^T M_h(\mathcal{X}_1)^{-1}G_h(\tilde{X}_1)-G_h(X_0)^T M_h(\mathcal{X}_1)^{-1}G_h(X_0))\\ &\quad+\frac{1}{2}\xi^T B_h(X_0)^T(M_h(\mathcal{X}_1)^{-1}-M_h(X_0)^{-1})B_h(X_0)\xi\\ &\quad+\frac{1}{2}\eta^T B_h(X_1^\star)^{-1}(M_h(\mathcal{X}_1)-M_h(X_1^\star))M_h(\mathcal{X}_1)^{-1}B_h(X_1^\star)\eta\\ &\quad+\frac{kT}{2}\log\left(\frac{\det M_h(X_1^\star)}{\det M_h(\mathcal{X}_1)}\right).\end{aligned}$$

Recall that, $\xi\in\mathbb{R}^n$ denotes a Gaussian random vector with mean zero and covariance $\mathbb{E}(\xi_i\xi_j)=kT\delta_{ij}$. Since:

$$\begin{cases}X_1^\star-\mathcal{X}_1=2(\tilde{X}_1-X_0)+h(G_h(\tilde{X}_1)-G(X_0))\,,\\ \tilde{X}_1-X_0=\sqrt{\frac{h}{2}}B_h(X_0)\xi\,,\end{cases} \tag{8.2}$$

it follows from bounds on the derivatives of both $G_h(x)$ and $M_h(x)$ that,

$$g_{kT}(x,\xi,h)\le C_1\,|\xi|^2\,h^{1/2}\,.$$

Thus,

$$\tilde{\alpha}_h(X_0,\xi)\le\alpha_h(X_0,X_1^\star)\le 1 \tag{8.3}$$

where we have introduced:

$$\tilde{\alpha}_h(X_0,\xi) = 1 \wedge \exp\left(-\frac{1}{kT}\left[f_{kT}(X_0,h) + \frac{kT}{2}\log\left(\frac{\det M_h(\mathcal{X}_1)}{\det M_h(X_0)}\right) + C_1\,|\xi|^2\,h^{1/2})\right]\right).$$

Taking the expectation of the lower bound in (8.3) yields:

$$\begin{aligned}\mathbb{E}\,\tilde{\alpha}_h(X_0,\xi) = &\int_{\Omega_{kT}} \exp\left(-\frac{|z|^2}{2}\right)(2\pi)^{-n/2}dz \\ &+ \exp\left(-\frac{1}{kT}\left[f_{kT}(X_0,h) + \frac{kT}{2}\log\left(\frac{\det M_h(\mathcal{X}_1)}{\det M_h(X_0)}\right)\right]\right) \\ &\times \int_{\Omega_{kT}^c} \exp\left(-\frac{|z|^2}{2}(1+C_1 h^{1/2})\right)(2\pi)^{-n/2}dz\end{aligned}$$

where

$$\Omega_{kT} = \left\{ y \in \mathbb{R}^n \;:\; C_1\,h^{1/2}\,|y|^2 \leq -\frac{f_{kT}(X_0,h)}{kT} - \frac{1}{2}\log\left(\frac{\det M_h(\mathcal{X}_1)}{\det M_h(X_0)}\right)\right\}.$$

Passing to the small noise limit $kT \to 0$, Assumption (A2) in Theorem 6.1 implies:

$$\Omega_{kT} \to \mathbb{R}^n, \quad \text{as } kT \to 0\,.$$

The desired limit statement follows from applying the Squeeze Lemma to (8.3). □

REFERENCES

[1] C. F Abrams and E. Vanden-Eijnden, *Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics*, Proceedings of the National Academy of Sciences **107** (2010), no. 11, 4961–4966.

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

[3] E. Akhmatskaya and S. Reich, *GSHMC: An efficient method for molecular simulation*, J Comput Phys **227** (2008), 4937–4954.

[4] R. J. Allen, D. Frenkel, and P. R. Ten Wolde, *Simulating rare events in equilibrium or nonequilibrium stochastic systems*, J Chem Phys **124** (2006), 024102.

[5] A. Beskos, N. S. Pillai, G. O. Roberts, J. M. Sanz-Serna, and A. M. Stuart, *Optimal tuning of hybrid Monte-Carlo algorithm*, To appear in Bernoulli.

[6] A. Beskos, G. O. Roberts, and A. M. Stuart, *Optimal scalings for local Metropolis-Hastings chains on non-product targets in high dimensions*, Ann Appl Probab **19** (2009), 863–898.

[7] R. B. Bird, R. C. Armstrong, O. Hassager, and C. F. Curtiss, *Dynamics of polymeric liquids*, Wiley, 1987.

[8] N. Bou-Rabee and M. Hairer, *Non-asymptotic mixing of the MALA algorithm*, IMA J of Numer Anal **33** (2013), 80–110.

[9] N. Bou-Rabee and H. Owhadi, *Long-run accuracy of variational integrators in the stochastic context*, SIAM J. of Numer. Anal. **48** (2010), 278297.

[10] N. Bou-Rabee and E. Vanden-Eijnden, *Pathwise accuracy and ergodicity of Metropolized integrators for SDEs*, Comm Pure and Appl Math **63** (2010), 655–696.

[11] ———, *A patch that imparts unconditional stability to explicit integrators for Langevin-like equations*, J Comput Phys **231** (2012), 2565–2580.

[12] M. Chopra and R. G. Larson, *Brownian dynamics simulations of isolated polymer molecules in shear flow near adsorbing and nonadsorbing surfaces*, J Rheol **46** (2002), 831.

[13] B. Cichocki, B. U. Felderhof, K. Hinsen, E. Wajnryb, and J. Bławzdziewicz, *Friction and mobility of many spheres in Stokes flow*, J Chem Phys **100** (1994), 3780.

[14] B. Cichocki, R. B. Jones, R. Kutteh, and E. Wajnryb, *Friction and mobility for colloidal spheres in Stokes flow near a boundary: The multipole method and applications*, J Chem Phys **112** (2000), 2548.

[15] G. E. Crooks and D. Chandler, *Efficient transition path sampling for nonequilibrium stochastic dynamics*, Phys Rev E **64** (2001), no. 2, 026109–026109.
[16] J. J. de Pablo, *Coarse-grained simulations of macromolecules: From DNA to nanocomposites*, Annu Rev Phys Chem **62** (2011), 555–574.
[17] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Hybrid Monte-Carlo*, Phys Lett B **195** (1987), 216–222.
[18] W. E and E. Vanden-Eijnden, *Metastability, conformation dynamics, and transition pathways in complex systems*, Multiscale Modelling and Simulation (S. Attinger and P. Koumoutsakos, eds.), Lecture Notes in Computational Science and Engineering, Springer, 2004, pp. 35–68.
[19] ———, *Transition-path theory and path-finding algorithms for the study of rare events*, Annual Review of Physical Chemistry **61** (2010), 391–420.
[20] J.-P. Eckmann and M. Hairer, *Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators*, Comm Math Phys **212** (2000), no. 1, 105–164.
[21] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures*, Comm Math Phys **201** (1999), no. 3, 657–697.
[22] D. L. Ermak and J. A. McCammon, *Brownian dynamics with hydrodynamics interactions*, J Chem Phys **69** (1978), 1352–1360.
[23] G. S. Ezra, *Reversible measure-preserving integrators for non-Hamiltonian systems*, J Chem Phys **125** (2006), 034104.
[24] K. Feng and D.-L. Wang, *Dynamical systems and geometric construction of algorithms*, Computational Mathematics in China (Z.-C. Shi and C.-C. Yang, eds.), Contemporary Mathematics Series, vol. 163, AMS, 1994, pp. 1–32.
[25] M. Fixman, *Simulation of polymer dynamics*, J Chem Phys **69** (1978), 1527–1545.
[26] ———, *Implicit algorithm for Brownian dynamics of polymers*, Macromolecules **19** (1986), 1195.
[27] A. Gelman, W. R. Gilks, and G. O. Roberts, *Weak convergence and optimal scaling of random walk metropolis algorithms*, Ann Appl Probab **7** (1997), 110–120.
[28] M. Girolami and B. Calderhead, *Riemann manifold Langevin and Hamiltonian Monte Carlo methods*, J R Statist Soc B **73** (2011), 123–214.
[29] M. D. Graham, *Fluid dynamics of dissolved polymer molecules in confined geometries*, Annu Rev Fluid Mech **43** (2011), 273–298.
[30] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration*, Springer, 2010.
[31] M. Hairer and J. C. Mattingly, *Slow energy dissipation in anharmonic oscillator chains*, Comm Pure and Appl Math **62** (2009), no. 8, 999–1032.
[32] W. K. Hastings, *Monte-Carlo methods using Markov chains and their applications*, Biometrika **57** (1970), 97–109.
[33] J. P. Hernández-Ortiz, J. J. de Pablo, and M. D. Graham, *Fast computation of many-particle hydrodynamic and electrostatic interactions in a confined geometry*, Phys Rev Lett **98** (2007), 140602.
[34] J. P. Hernández-Ortiz, H. Ma, J. J. de Pablo, and M. D. Graham, *Cross-stream-line migration in confined flowing polymer solutions: theory and simulation*, Phys Fluids **18** (2006), 123101.
[35] ———, *$N \log N$ method for hydrodynamic interactions of confined polymer systems: Brownian dynamics*, J Chem Phys **125** (2006), 164906.
[36] ———, *Concentration distributions during flow of confined flowing polymer solutions at finite concentration: slit and grooved channel*, Korea-Aust Rhelo J **20** (2008), 143.
[37] M. Herrchen and H. C. Öttinger, *A detail comparison of various FENE dumbbell models*, J Non-Newton Fluid **68** (1997), 17.
[38] D. J. Higham, *Stochastic ordinary differential equations in applied and computational mathematics*, IMA J Appl Math **76** (2011), 449–474.
[39] D. J. Higham, X. Mao, and A. M. Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equations*, IMA J Num Anal **40** (2002), 1041–1063.
[40] A. M. Horowitz, *A generalized guided Monte-Carlo algorithm*, Phys Lett B **268** (1991), 247–252.
[41] C.-C. Hsieh, L. Li., and R. G. Larson, *Modeling hydrodynamic interaction in Brownian dynamics: simulations of extensional flows of dilute solutions of DNA and polystyrene*, J Non-Newton Fluid **113** (2003), 147.
[42] C.-C. Hsieh and T.-H. Lin, *Simulation of conformational preconditioning strategies for electrophoretic stretching of DNA in a microcontraction*, Biomicrofluidics **5** (2011), no. 4, 044106.

[43] C.-C. Hsieh, T.-H. Lin, and C.-D. Huang, *Simulation guided design of a microfluidic device for electrophoretic stretching of DNA*, Biomicrofluidics **6** (2012), 044105.
[44] I. Huopaniemi, K. Luo, T. Ala-Nissila, and S.-C. Ying, *Langevin dynamics simulations of polymer translocation through nanopores*, J Chem Phys **125** (2006), 124901.
[45] J. S. Hur, E. S. G. Shaqfeh, H. P. Babcock, and S. Chu, *Dynamics and configurational fluctuations of single DNA molecules in linear mixed flows*, Phys Rev E **66** (2002), 011915.
[46] J. S. Hur, E. S. G. Shaqfeh, and R. G. Larson, *Brownian dynamics simulations of single DNA molecules in shear flow*, J Rheol **44** (2000), 713.
[47] M. Hütter and H. C. Öttinger, *Fluctuation-dissipation theorem, kinetic stochastic integral, and efficient simulations*, J Chem Soc, Faraday Trans **94** (1998), 1403–1405.
[48] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, *Strong convergence of an explicit numerical method for sdes with non-globally Lipschitz continuous coefficients*, Ann Appl Probab **22** (2012), 1611–1641.
[49] R. M. Jendrejack, J. J. de Pablo, and M. D. Graham, *Hydrodynamic interactions in long chain polymers: application of the chebyshev polynomial approximation in stochastic simulations*, J Chem Phys **113** (2000), 7752.
[50] ———, *Stochastic simulations of DNA in flow: Dynamics and the effects of hydrodynamic interaction*, J Chem Phys **116** (2002), 7752.
[51] R. M. Jendrejack, E. T. Dimalanta, D. C. Schwartz, M. D. Graham, and J. J. de Pablo, *DNA dynamics in a microchannel*, Phys Rev Lett **91** (2003), 038102.
[52] ———, *Effect of confinement on DNA dynamics in microfluidic devices*, J Chem Phys **119** (2003), 2894.
[53] R. M. Jendrejack, D. C. Schwartz, J. J. de Pablo, and M. D. Graham, *Shear-induced migration in flowing polymer solutions: simulation of long-chain Deoxyribose Nucleic Acid in micro channels*, J Chem Phys **120** (2004), 2513.
[54] S. Jiang, Z. Liang, and J. Huang, *A fast algorithm for Brownian dynamics simulation with hydrodynamic interactions*, Mathematics of Computation **82** (2013), no. 283, 1631–1645.
[55] M. Kedzierski and E. Wajnryb, *Precise multipole method for calculating many-body hydrodynamic interactions in a microchannel*, J Chem Phys **133** (2010), 154105.
[56] A. D. Kennedy and B. Pendleton, *Cost of the generalized hybrid Monte Carlo algorithm for free field theory*, Nucl. Phys. B **607** (2001), 456–510.
[57] J. Kent, *Time-reversible diffusions*, Adv. Appl. Prob. **10** (1978), 819–835.
[58] K. Kikuchi, M. Yoshida, T. Maekawa, and H. Watanabe, *Metropolis Monte Carlo method as a numerical technique to solve the Fokker-Planck equation*, Chem Phys Lett **185** (1991), 335–338.
[59] ———, *Metropolis Monte Carlo method for Brownian dynamics simulation generalized to include hydrodynamic interactions*, Chem Phys Lett **196** (1992), 57–61.
[60] J. M. Kim and P. S. Doyle, *A Brownian dynamics-finite element method for simulating DNA electrophoresis in nonhomogeneous electric fields*, J Chem Phys **125** (2006), 074906.
[61] ———, *Design and numerical simulation of a DNA electrophoretic stretching device*, Lab on a Chip **7** (2007), no. 2, 213–225.
[62] J. D. Lambert, *Computational methods in ordinary differential equations*, Wiley New York, 1973.
[63] R. G. Larson, *The rheology of dilute solutions of flexible polymers: Progress and problems*, J Rheol **49** (2005), 1.
[64] R. G. Larson, H. Hu, D. E. Smith, and S. Chu, *Brownian dynamics simulations of a DNA molecule in an extensional flow field*, J Rheol **43** (1999), 267.
[65] T. Lelievre and F. Legoll, *Effective dynamics using conditional expectations*, Nonlinearity **23** (2010), 2131–2163.
[66] T. Lelièvre, M. Rousset, and G. Stoltz, *Free energy computations: A mathematical perspective*, 1st ed., Imperial College Press, 2010.
[67] Z. Liang, Z. Gimbutas, L. Greengard, J. Huang, and S. Jiang, *A fast multipole method for the Rotne-Prager-Yamakawa tensor and its applications*, J Comp Phys **234** (2012), 133–139.
[68] K. K. Lin and L.-S. Young, *Nonequilibrium steady states for certain Hamiltonian models*, J Stat Phys **139** (2010), no. 4, 630–657.
[69] J. S. Liu, *Monte Carlo strategies in scientific computing*, 2nd ed., Springer, 2008.
[70] H. C. Loebl, R. Randel, S. P. Goodwin, and C. C. Matthai, *Simulation studies of polymer translocation through a channel*, Phys Rev E **67** (2003), no. 4, 041913.
[71] J. J. López Cascales and J. Garca de la Torre, *Shear-rate dependence of the intrinsic viscosity of bead-and-spring chains: hydrodynamic interaction and excluded-volume effects*, Polymer **32** (1991), 3359.
[72] D. K. Lubensky and D. R. Nelson, *Driven polymer translocation through a narrow pore*,

Biophysical Journal **77** (1999), no. 4, 1824–1838.
[73] R. S. Maier and D. L. Stein, *Transition-rate theory for nongradient drift fields*, Phys Rev Lett **69** (1992), no. 26, 3691.
[74] ———, *Escape problem for irreversible systems*, Phys Rev E **48** (1993), no. 2, 931.
[75] ———, *A scaling theory of bifurcations in the symmetric weak-noise escape problem*, J Stat Phys **83** (1996), no. 3-4, 291–357.
[76] R. Manella, *Absorbing boundaries and optimal stopping in a stochastic differential equation*, Physics Letters A **254** (1999), 257–262.
[77] L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, *String method in collective variables: Minimum free energy paths and isocommittor surfaces*, J Chem Phys **125** (2006), 024106.
[78] L. Maragliano and E. Vanden-Eijnden, *A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations*, Chem Phys Lett **426** (2006), 168175.
[79] J. C. Mattingly, N. S. Pillai, and A. M. Stuart, *Diffusion limits of the random walk Metropolis algorithm in high dimensions*, Ann Appl Probab **22** (2012), 881–930.
[80] S. Matysiak, A. Montesi, M. Pasquali, A. B. Kolomeisky, and C. Clementi, *Dynamics of polymer translocation through nanopores: Theory meets experiment*, Phys Rev Lett **96** (2006), no. 11, 118103.
[81] K. L. Mengersen and R. L. Tweedie, *Rates of convergence of the Hastings and Metropolis algorithms*, Ann Stat **24** (1996), 101–121.
[82] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equations of state calculations by fast computing machines*, J Chem Phys **21** (1953), 1087–1092.
[83] G. N. Milstein and M. V. Tretyakov, *Stochastic numerics for mathematical physics*, Springer, Berlin, 2004.
[84] ———, *Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients*, IMA J Num Anal **43** (2005), 1139–1154.
[85] E. Nummelin, *General irreducible Markov chains and non-negative operators*, Cambridge University Press, New York, NY, 1984.
[86] H. C. Öttinger, *Stochastic processes in polymeric fluids*, Springer-Verlag, 1996.
[87] A. Ralston, *Runge-kutta methods with minimum error bounds*, Mathematics of computation **16** (1962), no. 80, 431–437.
[88] G. C. Randall, K. M. Schultz, and P. S. Doyle, *Methods to electrophoretically stretch DNA: microcontractions, gels, and hybrid gel-microcontraction devices*, Lab on a Chip **6** (2006), no. 4, 516–525.
[89] P. Reimann, C. Van den Broeck, H. Linke, P. Hänggi, J.M. Rubi, and A. Pérez-Madrid, *Giant acceleration of free diffusion by use of tilted periodic potentials*, Phys Rev Lett **87** (2001), 010602.
[90] A. Rey, J. J. Freire, and J. Garca de la Torre, *Brownian dynamics simulation of flexible polymer chains with excluded volume and hydrodynamic interactions. a comparison with Monte Carlo and theoretical results*, Polymer **33** (1992), 3477–3481.
[91] L. Rey-Bellet and L. Thomas, *Exponential convergence to non-equilibrium stationary states in classical statistical mechanics*, Comm Math Phys **255** (2001), no. 2, 305–329.
[92] G. O. Roberts and J. S. Rosenthal, *Optimal scaling of discrete approximations to Langevin diffusions*, J Roy Statist Soc Ser B **60** (1998), 255–268.
[93] G. O. Roberts and R. L. Tweedie, *Exponential convergence of Langevin distributions and their discrete approximations*, Bernoulli **2** (1996), 341–363.
[94] ———, *Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms*, Biometrika **1** (1996), 95–110.
[95] P. J. Rossky, J. D. Doll, and H. L. Friedman, *Brownian dynamics as smart Monte Carlo simulation*, J Chem Phys **69** (1978), 4628.
[96] J. Rotne and S. Prager, *Variational treatment of hydrodynamic interaction in polymers*, J Chem Phys **50** (1969), 4831.
[97] C. M. S. Schroeder, H. B. Babcock, E. S. G. Shaqfeh, and S. Chu, *Observation of polymer conformation hysteresis in extensional flow*, Science **301** (2003), 1515.
[98] C. M. S. Schroeder, E. S. G. Shaqfeh, and S. Chu, *The effect of hydrodynamic interactions on DNA dynamics in extensional flow: simulation and single molecule experiment*, Macromolecules **37** (2004), 9242.
[99] E. S. G. Shaqfeh, *The dynamics of single-molecule DNA in flow*, J Non-Newton Fluid **130** (2005), 1.
[100] M. Somasi, B. Khomami, N. J. Woo, J. S. Hur, and S. G. Shaqfeh, *Brownian dynamics simulations of bead-rod and bead-spring chains: numerical algorithms and coarse-graining*

issues, J Non-Newtonian Fluid Mech **108** (2002), 227–255.

[101] R. L. Stratonovich, *Synchronization of a self-excited oscillator in the presence of noise*, Radiotekh Electron **3** (1958), 497.

[102] D. Talay, *Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme*, Markov Processes and Related Fields **8** (2002), 1–36.

[103] L. Tierney, *Markov chains for exploring posterior distributions*, Ann Stat **22** (1994), 1701–1728.

[104] ———, *A note on metropolis-hastings kernels for general state spaces*, Ann Appl Probab (1998), 1–9.

[105] D. W. Trahan and P. S. Doyle, *Simulation of electrophoretic stretching of DNA in a microcontraction using an obstacle array for conformational preconditioning*, Biomicrofluidics **3** (2009), 012803.

[106] C. Valeriani, R. J. Allen, M. J. Morelli, D. Frenkel, and P. R. ten Wolde, *Computing stationary distributions in equilibrium and nonequilibrium systems with forward flux sampling*, J Chem Phys **127** (2007), 114109.

[107] Eric Vanden-Eijnden, *Some recent techniques for free energy calculations*, J Comput Chem **30** (2009), no. 11, 1737–1747.

[108] E Weinan, *Principles of multiscale modeling*, Cambridge University Press, 2011.

[109] N. Woo, E. S. G. Shaqfeh, and B. Khomami, *The effect of confinement on the dynamics and rheology of dilute DNA solutions. i. entropic spring force under confinement and numerical algorithm*, J Rheol **48** (2004), 281.

[110] ———, *The effect of confinement on the dynamics and rheology of dilute DNA solutions. ii. effective rheology and single chain dynamics*, J Rheol **48** (2004), 299.

[111] Y. Zhang, J. J. de Pablo, and M. D. Graham, *An immersed boundary method for Brownian dynamics simulation of polymers in complex geometries: Application to DNA flowing through a nano slit with embedded nanopits*, J Chem Phys **136** (2012), 014901.

[112] Y. Zhang, A. Donev, T. Weisgraber, B. J. Alder, M. D. Graham, and J. J. de Pablo, *Tethered DNA dynamics in shear flow*, J Chem Phys **130** (2009), 234902.