

Article

A Novel Hybrid Monte Carlo Algorithm for Sampling Path Space

Francis J. Pinski

Department of Physics, University of Cincinnati, Cincinnati, OH 45221, USA; frank.pinski@uc.edu;
Tel.: +1-513-432-8717

Abstract: To sample from complex, high-dimensional distributions, one may choose algorithms based on the Hybrid Monte Carlo (HMC) method. HMC-based algorithms generate nonlocal moves alleviating diffusive behavior. Here, I build on an already defined HMC framework, hybrid Monte Carlo on Hilbert spaces (Beskos, et al. *Stoch. Proc. Applic.* 2011), that provides finite-dimensional approximations of measures π , which have density with respect to a Gaussian measure on an infinite-dimensional Hilbert (path) space. In all HMC algorithms, one has some freedom to choose the mass operator. The novel feature of the algorithm described in this article lies in the choice of this operator. This new choice defines a Markov Chain Monte Carlo (MCMC) method that is well defined on the Hilbert space itself. As before, the algorithm described herein uses an enlarged phase space Π having the target π as a marginal, together with a Hamiltonian flow that preserves Π . In the previous work, the authors explored a method where the phase space π was augmented with Brownian bridges. With this new choice, π is augmented by Ornstein–Uhlenbeck (OU) bridges. The covariance of Brownian bridges grows with its length, which has negative effects on the acceptance rate in the MCMC method. This contrasts with the covariance of OU bridges, which is independent of the path length. The ingredients of the new algorithm include the definition of the mass operator, the equations for the Hamiltonian flow, the (approximate) numerical integration of the evolution equations, and finally, the Metropolis–Hastings acceptance rule. Taken together, these constitute a robust method for sampling the target distribution in an almost dimension-free manner. The behavior of this novel algorithm is demonstrated by computer experiments for a particle moving in two dimensions, between two free-energy basins separated by an entropic barrier.



Citation: Pinski, F.J. A Novel Hybrid Monte Carlo Algorithm for Sampling Path Space. *Entropy* **2021**, *23*, 499. <https://doi.org/10.3390/e23050499>

Academic Editors: Michael Dellnitz, Carsten Hartmann and Feliks Nüske

Received: 23 March 2021

Accepted: 15 April 2021

Published: 22 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Brownian dynamics; stochastic processes; sampling path space; transition paths

1. Introduction

Often, it is important to understand how molecules change conformations. For example, the folding of proteins is of high interest [1]. One approach is to simulate such transitions and generate a thermodynamic-relevant ensemble of paths that start in one free-energy basin and end in another. The molecular motion here is assumed to be driven by the random thermal motions of the surroundings. This is modeled by Brownian dynamics with the thermal noise supplied by a heat reservoir operating at the fixed temperature ϵ .

The evolution of the particle position x is described by the Stochastic Differential Equation (SDE):

$$dx_t = F(x) dt + \sqrt{2\epsilon} dW_t \quad (1)$$

where dW_t is the standard Wiener process. Here, the force (drift) is assumed to be the (negative) gradient of a potential energy function $U(x)$, namely $F(x) = -\partial U/\partial x$. The physical potential $U(x)$ must be bounded from below to be physical.

Integrating the SDE forward over time, one assembles a trajectory: the starting position may be known, but not the ending point. As shown by Onsager and Machlup [2], the path probability can be expressed in terms of the path positions. The (negative) logarithm of this probability has become known as the Onsager–Machlup (OM) functional. As shown

in Appendix A, the form of the functional depends on how the SDE (Equation (1)) is discretized. To understand molecular transitions, it would be useful to extract a thermodynamic distribution of paths that start in one free energy basin and end in another. For such constrained paths, an effective Hamiltonian can be formed and expressed as:

$$H_{eff} = \frac{1}{2} \langle x | L | x \rangle + \Phi(x), \quad (2)$$

with the details given in Appendix A. Using this effective Hamiltonian, standard techniques can be used to sample the probability distribution at an effective temperature $\epsilon_{eff} = 2\epsilon/\Delta t$, with ϵ being the physical temperature and Δt being the time discretization step used in approximating the SDE.

Performing such path sampling is challenging for several reasons. Firstly, as the time step used to solve the SDE becomes small, that is as $\Delta t \rightarrow 0$, contributions from the high-frequency modes lead to a divergence when evaluated using the effective Hamiltonian (Equation (2)). Secondly, at each point along the path, one must have a full description of the molecule. Thirdly, barriers separating different pathways may be insurmountable. Thus, it is extremely important to have efficient algorithms to probe the thermodynamic-relevant distribution of paths. The work by Beskos et al. [3] addressed the first of these issues using a Hybrid Monte Carlo (HMC) [4] method that augments the path space with Brownian bridges. One of the advantages is that HMC generates nonlocal moves, alleviating the diffusive behavior that plagues other methods [5]. However, in that method, the velocities form a Brownian bridge, and thus, they scale with the (time) path length. This growth negatively affects the accuracy of the algorithm. As shown in Section 2, by creating velocities that form an Ornstein–Uhlenbeck (OU) [6] bridge, one can eliminate this dependence on the path length, while continuing to treat the high-frequency modes accurately. In Section 2, the algorithm for the deterministic integration in the Molecular Dynamics (MD) step is described, as well as the associated error in the effective energy. In Section 3, an example in two dimensions is described, as well as how the algorithmic parameters are chosen. In the novel scheme described here, the time step for the deterministic, molecular dynamics integration can be increased by at least an order of magnitude over the previous method [3], leading to a concomitant decrease in computational effort. In other words, the novel method presented here is at least 10 times faster. This is followed in Section 4 by the description of the results for sampling paths that contain a transition from one basin to another. In Section 5, the continuous-time limit of the OM functional is sampled. As found before [7,8], this form produces unphysical results. The paper ends with a short discussion of how this new algorithm can be used in calculations employing the path integral molecular dynamic method and, finally, with some concluding remarks.

2. Time Evolution of the Hamiltonian

In sampling the probability distribution, the effective Hamiltonian, Equation (2), is augmented by including Gaussian distributed variables that can be identified as the momentum, with the aim of using a Hybrid (or Hamiltonian) Monte Carlo (HMC) method. The method is summarized as follows: first, pick the augmented variables from their known distribution; use them to (approximately) evolve the Hamiltonian flow; and then, accept or reject the evolved path using the Metropolis–Hastings criterion. The augmented Hamiltonian is given by:

$$H = \frac{1}{2} \langle p | (L + A_{ou}^2 \mathbf{1})^{-1} | p \rangle + \frac{1}{2} \langle q | L | q \rangle + \Phi(q). \quad (3)$$

The bracket notation, introduced in Appendix A, is used to simplify the look of the equations. Define $x = q + l_t$ as the physical path with l_t being a term linear in time that allows x to have the starting point $x(0) = x_-$ and the ending point $x(T) = x_+$. Note that $\Phi(x) = \Phi(q + l_t)$ and is denoted as $\Phi(q)$. The path l_t is given by the linear (in t) relationship: $l_t = x_- + t(x_+ - x_-)/T$. The path p is the momentum conjugate to the path q . The

distribution of the momentum p is known, but not transparent. However, the distribution of the velocities is easier to understand, as it corresponds to an OU process, with A_{ou} being the OU parameter.

To understand how the momenta p are distributed, one begins with the OU process:

$$dv_t = -A_{ou} v_t dt + \sqrt{2\epsilon} dW_t. \tag{4}$$

The OM functional for this process is:

$$I_{om} \propto \frac{1}{2} \langle v | (L + A_{ou}^2 \mathbb{1}) | v \rangle + \text{unimportant constant}. \tag{5}$$

The mass operator can be defined as $M = (L + A_{ou}^2 \mathbb{1})$ and then the momentum as $|p \rangle = M |v \rangle$, giving the effective Hamiltonian for the momentum variables as:

$$H_{ou} = \frac{1}{2} \langle p | M^{-1} | p \rangle \tag{6}$$

Thus, the velocities will correspond to an OU process with both endpoints being zero, which is an OU bridge. See Appendix B for the numerical details for constructing realizations of OU bridges.

Now return to the original Hamiltonian. Equation (3) with some modifications is:

$$H_\alpha = \frac{1}{2} \langle p | M^{-1} | p \rangle + \frac{1}{2} \langle q | M | q \rangle + \alpha \left(\Phi(q) - \frac{1}{2} \langle q | A_{ou}^2 \mathbb{1} | q \rangle \right). \tag{7}$$

Note that with $\alpha = 1$, this Hamiltonian (Equation (7)) has the term $\frac{1}{2} \langle q | A_{ou}^2 \mathbb{1} | q \rangle$ added and then subtracted. This algebraic manipulation importantly simplifies the numerical procedure, as will be shown below. The role (and value) of the parameter α will be addressed later.

It is convenient to work with the velocities v given by $|v \rangle = M^{-1} |p \rangle$, which comprise a bridge being zero at both endpoints. When $A_{ou} = 0$, the velocities form a Brownian bridge, with covariance (with $t \geq s$) being $\mathbb{E} (v_s \times v_t) = s(1 - t/T)$ where T is the length of the path. This gives $\mathbb{E} v_{T/2}^2 = T/4$, which grows with the (time) length of the path. In contrast, the covariance of an OU bridge is given by:

$$\mathbb{E} v_s v_t = \frac{2\epsilon}{A_{ou}} \frac{\sinh(A_{ou}s) \sinh(A_{ou}(T-t))}{\sinh(A_{ou}T)} \quad \text{with } t \geq s, \tag{8}$$

with ϵ being the temperature. At the midpoint of the OU bridge, $\mathbb{E} v_{T/2}^2 \approx \epsilon / A_{ou}$ when $A_{ou}T \gg 2$, which is independent of the path length T . In Appendix B, the numerical procedure for constructing an OU bridge is described.

Defining ϕ in terms of Φ as $\phi(\mathbf{q}) = \partial\Phi/\partial\mathbf{q}$, the equations of motion are given by:

$$\begin{pmatrix} \frac{d}{dt} |q \rangle \\ M \frac{d}{dt} |p \rangle \end{pmatrix} = \underbrace{\begin{pmatrix} |v \rangle \\ -M |q \rangle \end{pmatrix}}_A + \alpha \underbrace{\begin{pmatrix} 0 \\ A_{ou}^2 |q \rangle - |\phi(q) \rangle \end{pmatrix}}_B \tag{9}$$

with $|\phi(q) \rangle = \{\phi(q_0), \phi(q_1), \phi(q_3), \dots, \phi(q_{N_t})\}$ and with the A and B groupings being used in the splitting implemented in the numerical algorithms. This method used is symplectic as it corresponds to a symmetric splitting in a Trotter–Strang [9,10] procedure.

2.1. Splitting: ABA

The numerical implementation of the ABA splitting is given in this section. The BAB splitting is relegated to Appendices C and D. Schematically, the integration can be viewed as pictured in Equation (10). The ABA procedure begins with the initial position and velocity (q_0, v_0) , then during the time interval $h/2$, these evolve into (q_H, v_H) using

Equations (11) and (12). Next is the so-called full step, where the velocity evolves over a time h changing to w_H using Equation (13). The final transformation takes (q_H, w_H) and turns them into (q_1, v_1) during a time interval $h/2$ using Equations (14) and (15).

$$\begin{pmatrix} q_0 \\ v_0 \end{pmatrix} \xrightarrow[\text{Half step}]{\mathbf{A}(h/2)} \begin{pmatrix} q_H \\ v_H \end{pmatrix} \xrightarrow[\text{Full step}]{\mathbf{B}(h)} \begin{pmatrix} q_H \\ w_H \end{pmatrix} \xrightarrow[\text{Half step}]{\mathbf{A}(h/2)} \begin{pmatrix} q_1 \\ v_1 \end{pmatrix} \quad (10)$$

Half step, A: $\{q_0, v_0\} \Rightarrow \{q_H, v_H\}$ with $\theta = \frac{h}{2}$:

$$\frac{d}{dt}|q\rangle = |v\rangle \Rightarrow |q_H\rangle = \cos(\theta)|q_0\rangle + \sin(\theta)|v_0\rangle \quad (11)$$

$$\frac{d}{dt}|v\rangle = -|q\rangle \Rightarrow |v_H\rangle = -\sin(\theta)|q_0\rangle + \cos(\theta)|v_0\rangle \quad (12)$$

Full step, B: $\{q_H, v_H\} \Rightarrow \{q_H, w_H\}$ with $M = (L + A_{ou}^2 \mathbb{1})$:

$$M \frac{d}{dt}|v\rangle = -\alpha \left(|\phi(q)\rangle - A_{ou}^2 |q\rangle \right) \Rightarrow M|\Delta\bar{v}\rangle = -\alpha h \left(|\phi(q_H)\rangle - A_{ou}^2 |q_H\rangle \right) \quad (13)$$

with $|\Delta\bar{v}\rangle = |w_H\rangle - |v_H\rangle$.

Half step, A: $\{q_H, w_H\} \Rightarrow \{q_1, v_1\}$ with $\theta = \frac{h}{2}$:

$$\frac{d}{dt}|q\rangle = |v\rangle \Rightarrow |q_1\rangle = \cos(\theta)|q_H\rangle + \sin(\theta)|w_H\rangle \quad (14)$$

$$\frac{d}{dt}|v\rangle = -|q\rangle \Rightarrow |v_1\rangle = -\sin(\theta)|q_H\rangle + \cos(\theta)|w_H\rangle \quad (15)$$

Note that the integrations in the half steps are exact, while the full step integration is approximate. As the numerical algorithm is based on a symmetric splitting of the time evolution operator, it is symplectic and represents the flow of a shadow Hamiltonian [11,12]. Thus, for a small integration time step h , the energy error will be bounded. In the following section, the formula for this error is examined.

2.2. "Energy" Error

The deterministic MD integration is preformed over a total time $N_{MD} h$. For each time increment, the integration corresponds to the ABA method described above. The total energy error for the MD integration can be determined from the sum of the errors of each ABA step. Thus, it is sufficient to find the expression for the energy error in a single ABA step.

The error of the "energy" in the ABA step is due to the middle (full) Step B, since the first and third steps reflect exact integrations. The error is given by:

$$\begin{aligned} \Delta E^{(01)} &= \Phi(q_1) - \Phi(q_0) - \frac{1}{2} \langle q_1 | A_{ou}^2 | q_1 \rangle + \frac{1}{2} \langle q_1 | A_{ou}^2 | q_0 \rangle \\ &+ \frac{1}{2} \langle v_1 | M | v_1 \rangle - \frac{1}{2} \langle v_0 | M | v_0 \rangle + \frac{1}{2} \langle q_1 | M | q_1 \rangle - \frac{1}{2} \langle q_0 | M | q_0 \rangle \end{aligned} \quad (16)$$

which simplifies to:

$$\Delta E^{(01)} = \Phi(q_1) - \Phi(q_0) - \frac{1}{2} \langle q_1 + q_0 | A_{ou}^2 | q_1 - q_0 \rangle + \frac{1}{2} \langle w_H + v_H | M | w_H - v_H \rangle. \quad (17)$$

Substituting for v_H and w_H , one transforms the last term using:

$$\langle w_H + v_H | M | w_H - v_H \rangle = -\frac{h\alpha}{\sin\theta} \langle \phi(q_h) - A_{ou}^2 q_H | q_1 - q_0 \rangle \quad (18)$$

The value of α can be chosen to be $\alpha = \text{sinc}(h/2)$. With this value of α , the energy error is given as:

$$\Delta E^{(01)} = \Phi(q_1) - \Phi(q_0) - \left\langle \phi(q_H) - A_{ou}^2(q_h - \bar{q}_{01}) \middle| q_1 - q_0 \right\rangle \quad (19)$$

with $\bar{q}_{01} = (q_1 + q_0)/2$. Now, it is clear why α was chosen in this way. The approximate equations of motion reproduce the flow of a shadow Hamiltonian [11,12]. Evidently, the flow produced by numerically solving the equations of motion with $\alpha = \text{sinc}(h/2)$ is closer to the actual flow than to the one with unit α .

The full step in the ABA splitting is:

$$M |\Delta v \rangle = -\alpha h \left(|\phi(q_H) \rangle - A_{ou}^2 |q_H \rangle \right) \quad (20)$$

The operator M^{-1} is defined with the boundary conditions such that Δv is zero at both the beginning and end of the path. The above is then a matrix equation: with M being a symmetric, real tridiagonal matrix, with the only nonzero elements given by:

$$M_{i,i} = A_{ou}^2 + \frac{2}{\Delta t^2} \quad M_{i,i+1} = M_{i+1,i} = -\frac{1}{\Delta t^2}. \quad (21)$$

With the vanishing of Δv at both ends of the path, there are no endpoint corrections to M . To solve the matrix equation, the standard Gaussian elimination procedure is used (without pivoting).

3. Numerical Experiments

To explore the efficacy of the HMC method described above, I chose a two-dimensional example. A particle moving in a potential is given by:

$$U(x, y) = \exp\left(-2\left(x + \frac{1}{2}\right)^2 - 3(y + 1)^2\right) + \left(x^2 + y^{16} - 1\right)^2 \quad (22)$$

where x and y are the usual Cartesian coordinates. The first term above is a unit Gaussian centered at $x_g = -0.5$ and $y_g = -1$. The second term is a trough, with a minimum of zero and somewhat squarish in shape. A contour plot of this potential is given in Figure 1. In this figure, note that the white contour corresponds to $U(x, y) = 2$ and the dashed line designates the contour $U(x, y) = 0.05$, which encloses the solid black contour $U(x, y) = 0.001$. A saddle point exists near $x_s \approx -0.5$ and $y_s \approx -1$ with $U(x_s, y_s) = 1$. A local maximum exists near $x_m \approx -0.225$ and $y_m \approx -0.794$ with $U(x_m, y_m) \approx 1.6$. Connecting the free-energy basins is an energy-barrier-free path that becomes severely narrowed as it crosses the x -axis. Thus, one can view this potential as consisting of two basins separated by an entropic barrier if the temperature is not too high.

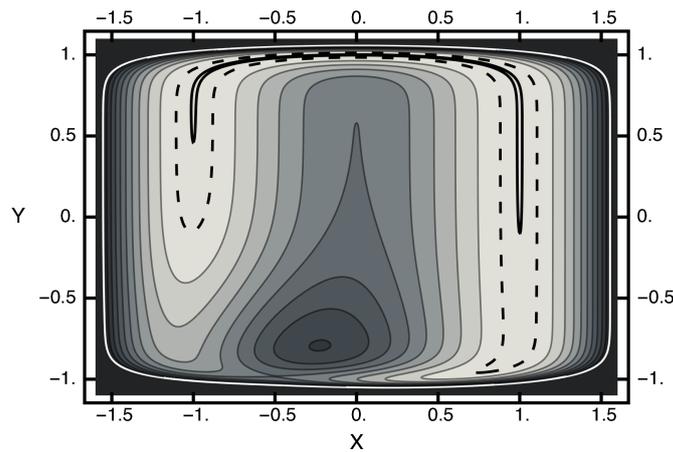


Figure 1. Contour plot of the two-dimensional potential. The horizontal axis corresponds to the x value; the vertical axis, the y axis. The value of the white contour is two and of the dashed contour is 0.05. The solid black contour enclosed by the dashed contour has a value of 0.001. The potential at the saddle point is approximately unity. The potential at the local maximum is ≈ 1.6 .

3.1. Equilibrium Distribution

In this work, the computer experiments were performed with a temperature $\epsilon = 0.05$. The temperature is much smaller than the lowest energy barrier, which is unity at the saddle point. The equilibrium (Boltzmann) distribution at a temperature $\epsilon = 0.05$ is displayed in the next two figures. In Figure 2, this distribution is shown in terms of a density plot. The narrow channel at the top of the figure connects the smaller basin on the left with the larger basin (on the right). The lack of an energy barrier is reflected in the (equal) values of the distribution along the channel.

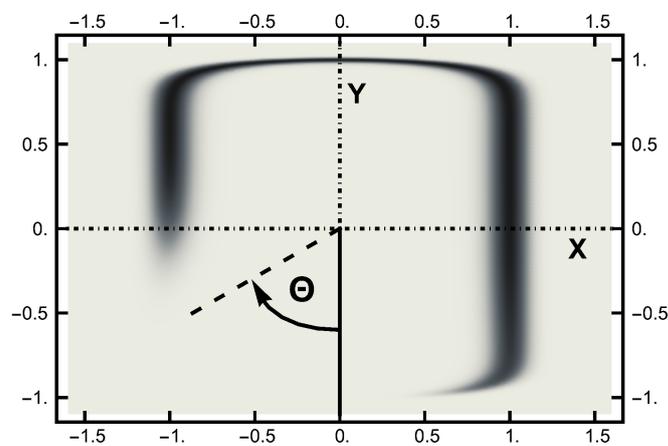


Figure 2. The equilibrium (Boltzmann) distribution for a temperature $\epsilon = 0.05$ plotted as a “density” plot. The black areas denote the higher probability areas. The narrow channel connects the left and right probability basins. The lack of an energy barrier is seen by the lack of variation in the shading along the channel. The angle Θ is pictured here.

To quantify the equilibrium distribution, I looked at its values as a function of angle Θ , which is defined in a clockwise sense, measured with respect to the negative y axis, as defined in Figure 2. At a temperature $\epsilon = 0.05$, the left basin has 36% of the weight, and the remaining 64% is in the right basin. A function \bar{P} is defined as

$$\bar{P}(\Theta) = (Z \delta)^{-1} \int_{\Theta-\delta/2}^{\Theta+\delta/2} d\theta \int_0^\infty r dr \exp\left(-\frac{U(-r \sin \theta, r \cos \theta)}{\epsilon}\right), \quad (23)$$

and is plotted in Figure 3 to emphasize the entropic barrier between the two basins.

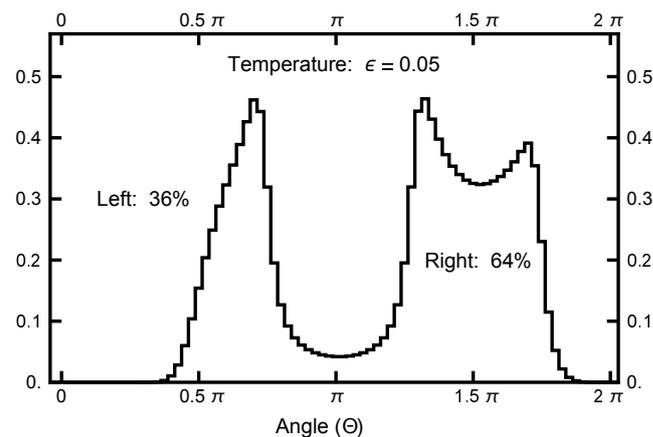


Figure 3. The function $\bar{P}(\Theta)$ plotted as a function of the angle Θ . See Equation (23) for its definition. The value of $\delta = \pi/40$ was used. Notice that the single-peak structure on the left differs from the twin peaks on the right. This structure is caused by the geometric factor, which arises when the radius line slices the distribution function.

3.2. Parameter Tuning

Unless stated otherwise, the form of the OM functional used was that corresponding to the midpoint integration of the SDE as written in Equation (A9). To investigate the two-dimensional paths that contain a transition between the two basins, one must pick several parameters; some are physical: the temperature and the path length. For this problem, the temperature was chosen to be $\epsilon = 0.05$. This temperature is small compared to the unit energy barrier. Thus, at this temperature, the narrowness of the channel will be the limiting factor that impedes basin hopping. The path length T must not be too long nor too short. For a very short value of T , the motion will become ballistic and uninteresting. For a very large T , the path will be dominated by intra-basin motion, which is not the focus of the study. The remaining parameters are those of the algorithm: the time length of the deterministic integration, the OU bridge parameters, A_{ou}^x and A_{ou}^y , and finally, the value of the algorithmic time step h . The last of these will be determined by requiring the Metropolis–Hastings acceptance rate to be sufficiently large.

3.2.1. Path Length

The next point to address is how quickly the particle moves from the left to right basins. To make an estimate of this transition time, I collected information from 10,000 realizations of Brownian dynamics by performing forward integrations of the two-dimensional SDE. Each integration started at the same place in the left basin. The results for the number of attempts residing in the right basin are tabulated as a function of time and plotted in Figure 4. As one sees, the curve approaches its equilibrium value as the total time reaches 1000. Furthermore, when the elapsed time is 125, over a third of the realizations are in the right basin. Thus, this latter value is a reasonable time (path length) to use for exploring doubly constrained paths that contain a transition. In the following sections, paths are examined that start in the left basin and end in the right basin, while time $T = 125$ has elapsed. The path is then described by 125,001 ordered pairs with $\Delta t = 1.0 \times 10^{-3}$.

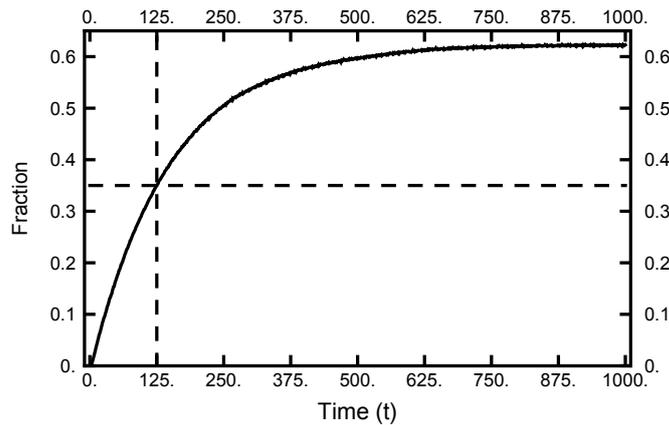


Figure 4. The results of 10,000 forward integrations of Brownian dynamics with temperature $\epsilon = 0.05$. All had the same initial conditions (in the left basin). The plot gives the fraction of paths with positive values of x as a function of time. As the dotted lines indicate, after $T = 125$, over 35% of the integrations have ending points in the right well. As the integration time exceeds 1000, the fraction approaches the equilibrium distribution.

3.2.2. Deterministic Integration Time

One of the advantages of the HMC method is its ability to generate nonlocal moves. To explore the nonlocality of a proposed move, the correlation function $d^{(x)}(nh)$ is defined by:

$$d^{(x)}(nh) = \sqrt{\frac{\sum_{i=1}^{N_t} (x_i(nh) - x_i(0))^2}{N_t}} \tag{24}$$

with a similar definition for $d^{(y)}(nh)$, where n is the number of the molecular dynamics steps and N_t is the number of points along the path. In Figure 5, $d^{(x)}(\tau)$ and $d^{(y)}(\tau)$ are plotted as functions of τ for two cases, $A_{ou}^{(x)} = A_{ou}^{(y)} = 0$ and $A_{ou}^{(x)} = A_{ou}^{(y)} = 1$. The time step parameter h was adjusted so that the (energy) errors were comparable, as seen in Figure 6. Such energy errors will be examined more closely later in the paper. As seen in Figure 5, both $d^{(x)}(\tau)$ and $d^{(y)}(\tau)$ level out for $\tau > 2$. Thus, the MD integration time should be on the order of $\pi/2$ to maximize the nonlocality of the move. To avoid any resonances that might occur if a fixed integration time should be chosen, the length of the integration is chosen to be uniformly distributed in the interval $(\pi/4, 3\pi/4)$ designated by the gray shaded area in Figure 5.

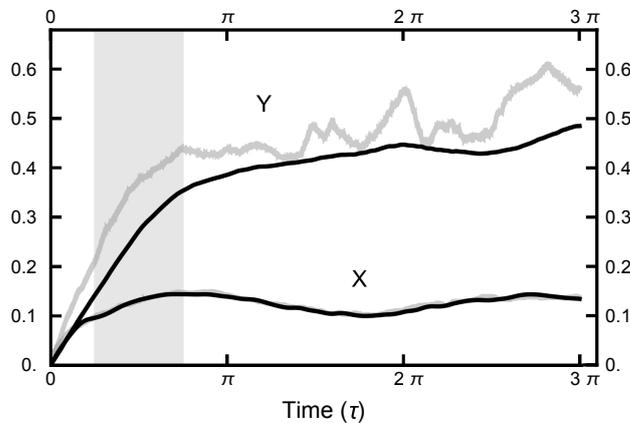


Figure 5. The correlation functions $d^{(x)}(\tau)$ and $d^{(y)}(\tau)$ for two runs. For the black curves, the parameters were $A_{Ou}^x = A_{Ou}^y = 1$; for the gray curves, $A_{Ou}^x = A_{Ou}^y = 0$. The functions $d^{(x)}(\tau)$ are almost identical for the two runs: the lower black curve lies on top of the lower gray curve, hiding it.

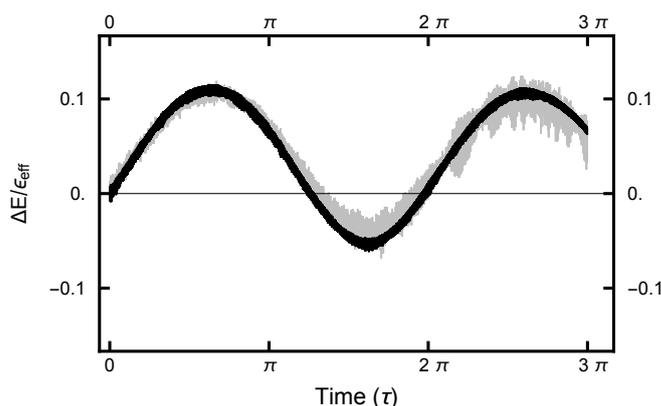


Figure 6. The energy error plotted as a function of MD integration time. The black curves designate the error for the case $A_{ou}^{(x)} = A_{ou}^{(y)} = 1$; the gray curves for $A_{ou}^{(x)} = A_{ou}^{(y)} = 0$. The time step parameter for the former case is $h = 6.667 \times 10^{-4}$ and for the latter $h = 6.667 \times 10^{-5}$. In the latter case, the smaller h meant that an order of magnitude more computing resources were required to generated the gray curve, as compared to the black one.

3.2.3. OU Bridge Parameter $A_{ou}^{(x,y)}$

To understand how the choice of A_{ou}^x and A_{ou}^y affect the sampling behavior, consider the plots of the correlation functions in Figure 7. As shown in the this figure and in Figure 5, the y-degree of freedom shows the floppiness that corresponds to the low energy cost of movement in that direction in either basin. The relative stiffness of motion in the x direction is a direct result of the restriction due to the narrow channel connecting the two basins. The floppiness (in the y direction) can be mathematically reduced by increasing the value of A_{ou}^y . For the chosen integration times (as designated as a gray rectangle in Figure 7), values A_{ou}^y between four and eight put the correlation function $d^{(y)}(\tau)$ in the same range as $d^{(x)}(\tau)$. However, decreasing A_{ou}^y below unity (not shown in the figures) does not alter the stiffness as its origin is due to the geometry of the connecting channel.

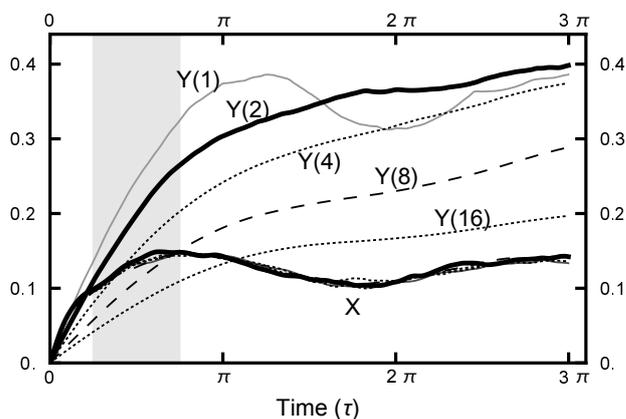


Figure 7. The correlation functions $d^{(x)}(\tau)$ and $A_{ou}^{(y)}$ for various runs. For all the runs, $A_{ou}^{(x)} = 1$. From the top, the curves for $d^{(y)}(\tau)$ have the values of $A_{ou}^{(y)}$ of 1, 2, 4, 8, and 16, respectively. The curves for $d^{(x)}(\tau)$ all lie on top of one other; only one is plotted.

3.2.4. MD Time Step Size (h)

For the deterministic integration in the Molecular Dynamics (MD) step, a choice must be made for the size of the time increment h . The size of h must not be too large nor too small. In the first case, the errors become large and the acceptance rate becomes unacceptably small. In the second case, computational resources are wasted. Since the numerical algorithm is based on a Trotter–Strang [9,10] splitting, for small values of h , the energy error is bounded, as the energy of a “nearby” Hamiltonian (the shadow Hamiltonian) is conserved. As shown in Figure 8, for small h , the energy error oscillates as a function of

integration time τ . In addition, because of the stochastic nature of the integrals, the curves are noisy, with the noise increasing as h increases.

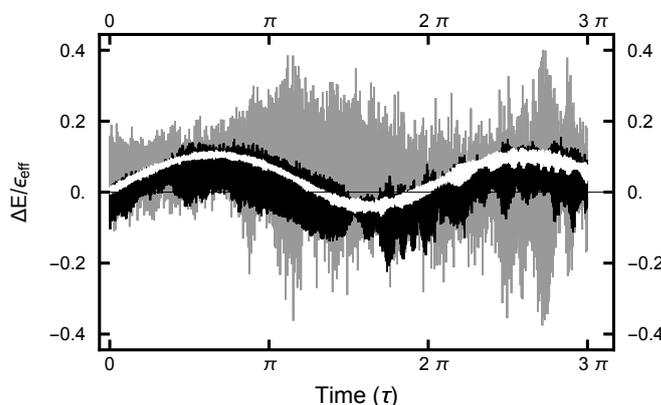


Figure 8. The effective-energy error plotted as a function of the MD integration time (τ). The three curves correspond to runs with values $A_{ou}^{(x)} = A_{ou}^{(y)} = 1$. The gray curve gives the error for the run with $h = 0.002$; the black curve, $h = 0.001$; and the white curve, $h = 0.00067$.

The value for h can be chosen by requiring the Metropolis–Hastings acceptance rate to be on the order of 80%. In Figure 9, this acceptance rate is plotted as a function of the time step h . As can be seen in the plot, the rate is a very steep function of h . Furthermore, when $A_{OU}^x = A_{OU}^y = 0$, h must be less than 0.0002 to have a substantial acceptance rate, while when $A_{OU}^x = 1$ and $A_{OU}^y = 8$, h can be chosen almost two orders of magnitude larger. Evidently the acceptance rate is sensitive to h as the values of A_{OU}^x and A_{OU}^y change.

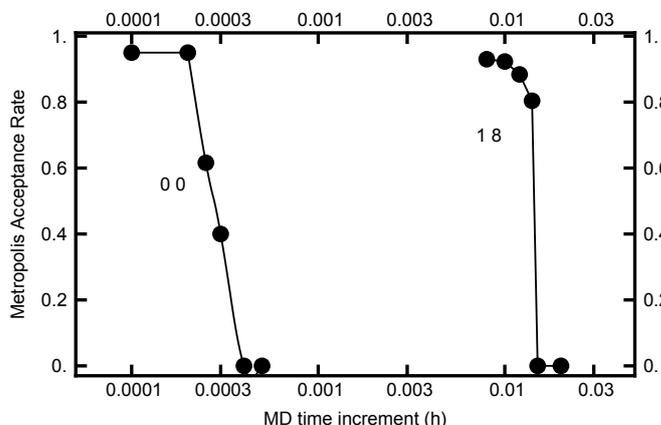


Figure 9. The Metropolis–Hastings acceptance rate for two runs. For the curve on the left, the parameters were $A_{OU}^x = A_{OU}^y = 0$ and for the curve on the right, $A_{OU}^x = 1$ and $A_{OU}^y = 8$. For both runs, the number of molecular dynamics steps were chosen to be $(1 + \eta)\pi/(4h)$ with η being a uniformly distributed random number in the unit interval.

The above highlights the advantage of the novel algorithm presented here. By choosing the OU parameters, A_{OU}^x and A_{OU}^y , to be unity, reasonably high acceptance rates can be achieved by using a value of h 10 times larger than when both OU parameters are zero. As shown in Figure 5, the nonlocality generated in the MD integration is only weakly dependent on the OU parameters, but does depend on the total MD integration time. Taken together, this then leads to a corresponding factor of 10 decrease in the computational effort by choosing the OU parameters to be unity. As shown in Figure 9, larger values of h can be used when larger OU parameters are used, leading to even fewer MD steps being necessary to attain the same MD integration time. This translates into a speedup of a factor of 10 or more, showing the power of picking nonzero values for the OU parameters and using the corresponding “optimal” value for h .

4. Path Sampling

As seen in Figures 2 and 3, in equilibrium, a particle will spend a small fraction of time in the narrow channel. Thus, the expectation is that a transition path would consist of motion in one basin or the other with a short transition while in the narrow channel. In Figure 10, a typical path is displayed: it starts in the left basin, remaining in it for under 25% of the path length, making a transition, remaining in the right basin for approximately the last 75% of the path. The boundary conditions influence the shape of the path, although the time spent in the narrow channel can change during the sampling procedure. The shown path is the end result of a simulation using $A_{OU}^x = 1$ and $A_{OU}^y = 8$ after 100,000 Metropolis–Hastings steps.

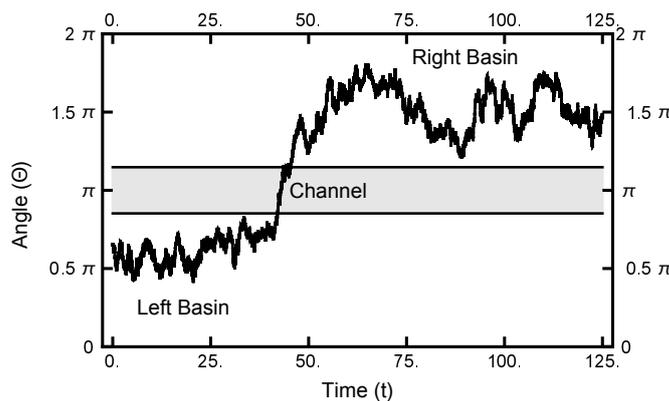


Figure 10. The one-dimensional representation of a typical path. The angle Θ plotted as a function of time. See Figure 2 for the definition of Θ . At $t = 0$, the particle starts out in the left basin, makes its way through the narrow channel at $t \approx 50$, and ends in the right basin.

In Figure 11, a summary of the evolution of this simulation is plotted. For these values of A_{OU} , on average, the particle spends about 70% of the time in the right basin. The variation of this percentage changes very little (only a few percent) over the course of the simulation run. In Figure 12, a similar plot is shown for the parameters $A_{OU}^x = 1$ and $A_{OU}^y = 4$ and 200,000 Metropolis–Hastings steps, twice as long as used for the results shown in Figure 11. The pictures are similar, except that the mean percentages are slightly shifted, and the variation about the mean is larger in the second case. The smaller variation in the first case is due to the larger value of A_{OU}^y , which dampens the oscillations, the floppiness, in the y component. With either parameter set, the path sampling calculations correctly infer that the free energy of the right basin is larger than that of the left basin.

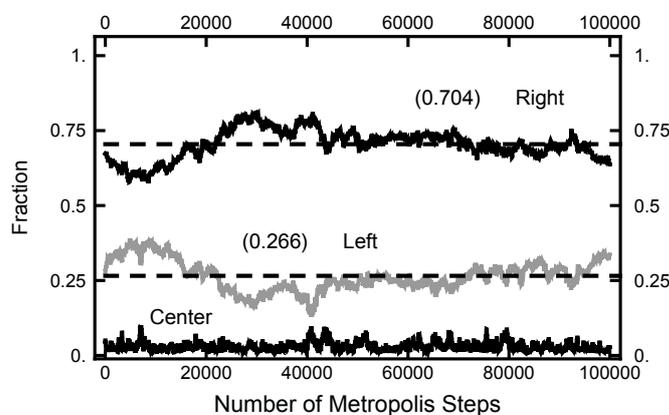


Figure 11. Results for a calculation with $A_{OU}^x = 1$ and $A_{OU}^y = 8$. The bottom (black) curve is the fraction of the path with $|x| < 1/2$; the middle (gray) curve is the fraction of the path with $x < -1/2$; and the top (black) curve is the fraction of the path with $x > 1/2$.

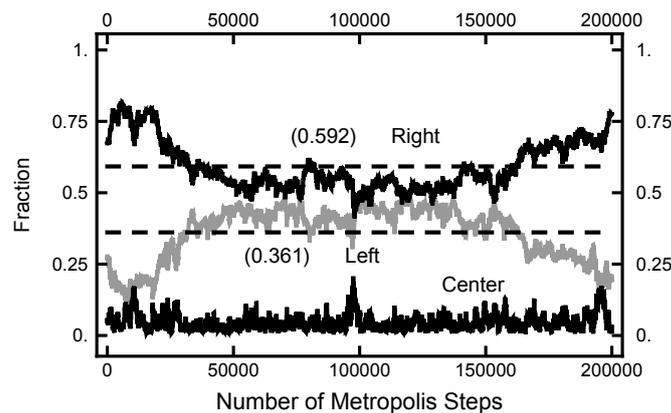


Figure 12. Results for a calculation with $A_{OU}^x = 1$ and $A_{OU}^y = 4$. The bottom (black) curve is the fraction of the path with $|x| < 1/2$; the middle (gray) curve is the fraction of the path with $x < -1/2$; and the top (black) curve is the fraction of the path with $x > 1/2$.

Now, consider Figure 13, where the information for the second simulation run is summarized. Here, the normalized histogram is plotted for the run with $A_{OU}^x = 1$ and $A_{OU}^y = 4$. Note that the comparison with the equilibrium histogram is quite close. This closeness indicates that the calculated ensemble of paths accurately probes the potential shape of each basin.

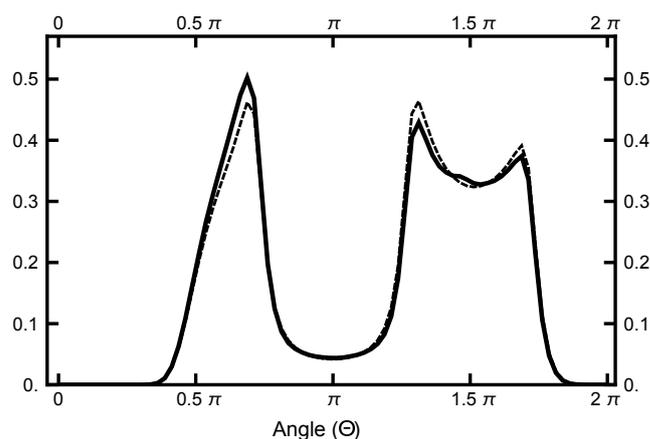


Figure 13. Results for a calculation with $A_{OU}^x = 1$ and $A_{OU}^y = 4$. The solid black curve is the histogram of the $\bar{P}(\Theta)$ for the 200,000 Metropolis steps pictured above in Figure 12. The dashed line represents the histogram that corresponds to the equilibrium distribution.

The results shown in Figures 11 and 12 used an initial path where the fraction spent in the left basin was less than the fraction spent in the right. This was chosen because the relative fractions were approximately what was expected from their equilibrium values. For the next calculation, the starting path had a value for the fraction spent in the left well of $\approx 80\%$, much larger than its equilibrium value. In Figure 14, the evolution of the right and left fractions is shown as a function of the algorithmic steps. In this simulation, $A_{OU}^x = A_{OU}^y = 1$ and the MD time step was $h = 0.003$. As indicated, it took about 10^5 steps before the left and right fractions became roughly equal and twice that many steps before the right fraction became significantly larger than the left. The change in the fraction was due to the movement of the time at which the transition took place and not the creation of a new pair of transitions. One does see an incipient attempt at the creation of a new pair of transitions at the $\approx 325,000$ th step of the simulation. This attempt is signaled by the spike in the fraction of the path that is spent in the narrow channel, shown as the dark grey curve.

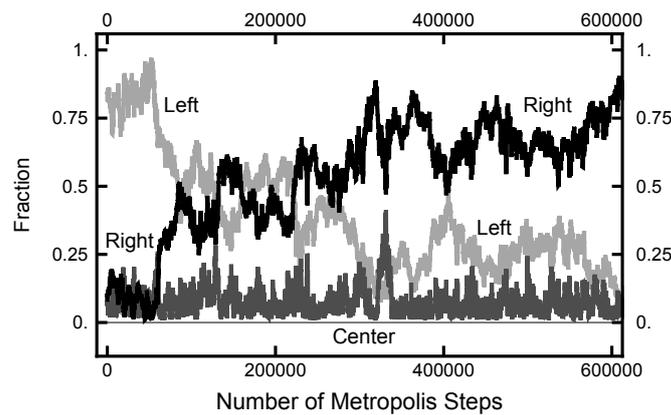


Figure 14. Results for a calculation with $A_{OU}^x = A_{OU}^y = 1$. The black curve is the fraction of the path with $x > 1/2$; the light gray curve is the fraction of the path with $x < -1/2$; and the bottom (dark gray) curve is the fraction of the path with $|x| < 1/2$.

5. Continuous-Time Limit

Now, I turn to the calculations using the continuous-time limit of the OM functional [13,14]. The Radon–Nikodym derivative is manipulated using the Girsanov theorem and Ito’s lemma. Displayed in Equation (A11), this is denoted as the Ito–Girsanov measure. As shown in Figure 15, the results of a simulation using the Ito–Girsanov measure gave unphysical results. As shown in Figure 16, the paths converge very quickly to objects that look similar to noisy versions of the Most Probable Path (MPP). Notice that this behavior happens after only 5000 hybrid steps. The MPP reflects the maximum of the measure and is dominated by the maxima of the Laplacian (of the physical potential). In the narrow channel, the Laplacian reflects the large curvature due to the confining sides. For calculations using the discrete form of the OM functional, Equation (A4), a very different behavior is found, as shown in Figure 11. This unphysical behavior is similar to what was found before [7,8], where the cause was traced back to a buildup of correlations that were inconsistent with the assumption in the original SDE that the noise is white and uncorrelated with the particle position.

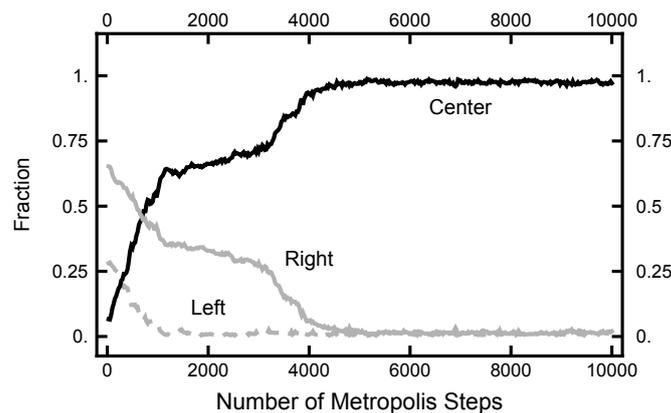


Figure 15. Results for a calculation with $A_{OU}^x = 1$ and $A_{OU}^y = 4$ using Equation (A11) as the effective Hamiltonian. The black curve labeled “Center” is the fraction of the path with $|x| < 1/2$; the solid gray curve is the fraction of the path with $x > 1/2$; and the dashed gray curve is the fraction of the path with $x < -1/2$.

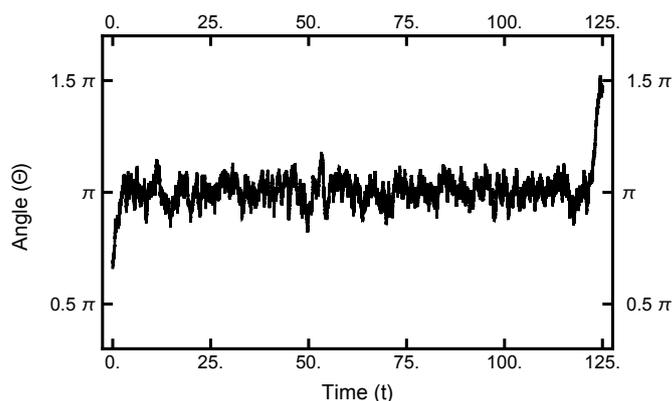


Figure 16. Results for a calculation with $A_{OU}^x = 1$ and $A_{OU}^y = 4$ using Equation (A11) as the effective Hamiltonian. This is qualitatively different from the path plotted in Figure 10.

6. Discussion

This work is related to the recent work of Korol et al. [15], who explored using a similar approach and applied it to path-integral molecular dynamics. In that work, the authors suggested that one should use mollified forces. This force mollification is a frequency-dependent factor that limits the effects of high-frequency internal modes and yet keeps intact the low-frequency motion. In this current work, this task is performed by the mass operator, $M = L + A_{ou}^2 \mathbb{1}$. Here, A_{ou} can be adjusted to change the crossover between low- and high-frequency behavior. At high frequencies, the modes revert to the free-particle motion. This reversion is due to the effects of the L in the mass operator, which dominates at high frequencies. In the novel method described here, this “force mollification” is folded into a unified framework, incorporating its effects in a self-consistent manner.

If, instead of the dynamics, one is interested in highlighting particular modes, one can choose A_{ou} differently for each mode. In particular, when working in normal mode coordinates, one can choose A_{ou} individually to minimize the size of Δv in Equation (13), thereby reducing the numerical errors in the integration.

7. Conclusions

A novel HMC procedure for sampling paths was presented here. This new procedure is much more efficient than previous methods, resulting in an order of magnitude speedup for the two-dimensional problem explored. The decrease in computational effort comes from increasing the size of the MD time increment h by at least an order of magnitude. The novelty in the procedure lies in the choice of the mass matrix and concomitantly the velocities that form an OU bridge. Using this new procedure, I successfully examined transitions across an entropic barrier and found that the method can determine the relative free energy of the two basins that are not separated by an energy barrier.

Again, using the Ito–Girsanov measure was found to give unphysical results. The sampling procedure, for the two-dimensional problem examined, converged to paths that look like noisy versions of the MPP [16] (most probable path), where the vast majority of the path was trapped in the narrow channel that connected the probability basins. Clearly, the source for this unphysical feature is the Laplacian (of the physical potential) in the formula for the measure.

This new algorithm is also useful for calculations using path-integral molecular dynamics [15] where the effective Hamiltonian has a similar form to that studied here. The novel procedure presented here naturally folds the “mollification” of forces into the HMC formalism by a modification of the mass matrix. Thus, this novel procedure includes this mollifying effect in a self-consistent manner. The parameters $\{A_{ou}\}$ in this novel method are able to be tuned individually for each degree of freedom to emphasize different physical effects.

In addition, this method is able to be adapted for use with the OBABO integration scheme [17] or with the second-order Langevin method [18].

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: I wish to acknowledge many conversations with Patrick Malsom, Andrew Stuart, Kody Law, Gideon Simpson, and Robin Ball. In addition, I wish to thank the Department of Physics, University of Cincinnati, for the funds to cover the publication costs associated with this article.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Onsager–Machlup Functionals

In this Appendix, the algebraic forms of the OM functionals are delineated. For simplicity, the one-dimensional forms of the equations are written below; the multi-dimensional generalization is straight forward. First, recognize that different discretizations of the SDE (Equation (1)) result in different forms of the OM function [19]. Here, two are considered: the Euler–Maruyama and midpoint methods. The former is given by:

$$x_{n+1} = x_n + F(x_n) \Delta t + \sqrt{2 \epsilon \Delta t} \zeta_{n+1} \tag{A1}$$

where ζ_n is a random Gaussian variate with mean zero and unit variance and Δt is the time increment. The original Onsager–Machlup functional is based on this expression. If one iterates Equation (A1) N_t times, one obtains a sequence of points $|x \rangle = \{x_0, x_1, x_2, x_3, x_4, \dots x_{N_t}\}$, with a probability:

$$P_{EM} = (2 \pi)^{-N_t/2} \prod_{n=1}^{N_t} \exp\left(-\frac{1}{2} \zeta_n^2\right) \tag{A2}$$

Onsager and Machlup recognized that one can rewrite this probability distribution in terms of the path positions as $P_{EM} \propto \exp(-I_{OM})$ with:

$$I_{OM} = \frac{\Delta t}{2\epsilon} \sum_{n=1}^{N_t} \frac{1}{2} \left(\frac{x_{n+1} - x_n}{\Delta t} - F(x_n) \right)^2 \tag{A3}$$

Defining $\epsilon_{eff} = 2 \epsilon / \Delta t$ as an effective temperature:

$$P_{EM} \propto \exp\left(-\frac{H_{EM}}{\epsilon_{eff}}\right)$$

with:

$$H_{EM} = \sum_{n=1}^{N_t} \frac{1}{2} \left(\frac{x_{n+1} - x_n}{\Delta t} \right)^2 - \left(\frac{x_{n+1} - x_n}{\Delta t} \right) F(x_n) + \frac{1}{2} F(x_n)^2$$

which can be rewritten as:

$$H_{EM} = \frac{1}{2 \Delta t^2} \sum_{n=2}^{N_t-1} x_n (2 x_n - x_{n+1} - x_{n-1}) + \sum_{n=1}^{N_t} \left(\frac{1}{2} F(x_n)^2 - \left(\frac{x_{n+1} - x_n}{\Delta t} \right) F(x_n) \right) \tag{A4}$$

Here, it is convenient to use the bra-ket mathematical notation. Specifically:

$$\sum_n \sum_m x_n B_{n,m} x_m = \langle x | B | x \rangle.$$

By rearranging the order in the sum, the first term in Equation (A4) can be transformed to:

$$\frac{1}{2} \sum_{n=1}^{N_t} \left(\frac{x_{n+1} - x_n}{\Delta t} \right)^2 = \frac{1}{2 \Delta t^2} \sum_{n=2}^{N_t-1} x_n (2 x_n - x_{n+1} - x_{n-1}) = \frac{1}{2} \langle x | L | x \rangle, \tag{A5}$$

where L is a symmetric, tridiagonal, positive definite matrix. The last term in Equation (A4) is denoted as $\Phi_{EM}(\{x\})$. The effective Hamiltonian can be written as:

$$H_{EM} = \frac{1}{2} \langle x | L | x \rangle + \Phi_{EM}(\{x\}) \quad (\text{A6})$$

Another discretization of the SDE (Equation (1)) uses the Stratonovich midpoint [20] integration and is written as:

$$x_{n+1} = x_n + F(x_{mp}) \Delta t + \sqrt{2\epsilon \Delta t} \xi_{n+1} \quad (\text{A7})$$

with $x_{mp} = (x_{n+1} + x_n)/2$. The effective Hamiltonian is a bit more complicated because of the Jacobian and again can be put in the form:

$$H_{mp} = \frac{1}{2} \langle x | L | x \rangle + \Phi_{mp}(\{x\}). \quad (\text{A8})$$

The second term in the equation above is:

$$\Phi_{mp}(\{x\}) = -\frac{2\epsilon}{\Delta t} \sum_{n=1}^{n_t} \ln(\det J_n) + \sum_{n=1}^{n_t} \left(\frac{1}{2} F(x_{mp})^2 - \left(\frac{x_{n+1} - x_n}{\Delta t} \right) F(x_{mp}) \right) \quad (\text{A9})$$

where J_n is the Jacobian matrix, whose elements are:

$$J_{(\alpha,\beta)} \propto \delta_{\alpha,\beta} + \frac{\Delta t}{2} \frac{\partial^2 U}{\partial x_\alpha \partial x_\beta} \quad (\text{A10})$$

and where the Jacobian J_n is evaluated at the midpoint.

Lastly, the continuous-time limit of Equations (A6) and (A8) can be written as:

$$H_{IG} = \frac{1}{2} \langle x | L | x \rangle + \Phi_{IG}(\{x\}) \quad (\text{A11})$$

with:

$$\Phi_{IG}(\{x\}) = \sum_{n=1}^{n_t} G(x_n)$$

where:

$$G(x_n) = \frac{1}{2} F^2(x_n) - \epsilon U''(x_n) \quad (\text{A12})$$

in one dimension. In higher dimensions, the Laplacian of U replaces $U''(x_n)$.

Appendix B. Constructing Ornstein–Uhlenbeck Bridges

The following describes the method used to construct Ornstein–Uhlenbeck (OU) bridges. An OU bridge is an OU process that begins and ends at the origin. Consider the standard OU process for a variable Z_i given by the following SDE:

$$dZ_t = -A Z_t dt + \sqrt{2\epsilon} dW_t \quad (\text{A13})$$

with A being the OU parameter and ϵ being the temperature.

To form a realization of an OU process, we use the close relationship between HMC and the Metropolis Adjusted Langevin Algorithm (MALA). To sample the Boltzmann distribution for a harmonic potential, the Hamiltonian is augmented by auxiliary variables, which can be viewed as momenta, conjugate to the position variables. In one dimension, this is given by:

$$H = \frac{1}{2} p^2 + \frac{1}{2} A x^2 \quad (\text{A14})$$

The equations of motion are then:

$$\frac{dx}{dt} = p \quad \text{and} \quad \frac{dp}{dt} = -A x \tag{A15}$$

The HMC procedure is first to form the momentum $p_0 = \sqrt{\epsilon} \zeta$ where ϵ is the temperature and ζ is a Gaussian random variate with mean zero and unit variance.

If one integrates the equations of motion over a time h and uses the velocity-Verlet approximation, one obtains:

$$x_1 = x_0 + p_0 h - \frac{A}{2} x_0 h^2 \tag{A16}$$

which becomes:

$$x_1 - x_0 = -A x_0 \Delta t + \sqrt{2 \epsilon \Delta t} \zeta \tag{A17}$$

with $\Delta t = h^2/2$. This is then the Euler–Maruyama [21] formula for the OU process. Within the one-step HMC, one uses the Metropolis–Hastings step to accept or reject the proposed move:

$$\{x_0, p_0\} \rightarrow \{x_1, p_1 = p_0 - \sqrt{\frac{\Delta t}{2}} A (x_0 + x_1)\}. \tag{A18}$$

Such a move may be rejected if the energy error is too large.

However, the equations of motion above can be integrated exactly, which gives:

$$x_1 = x_0 \cos \sqrt{2 A \Delta t} + \zeta_0 \sqrt{\frac{\epsilon}{A}} \sin \sqrt{2 A \Delta t}. \tag{A19}$$

In such a method, no proposed moves will be rejected in the Metropolis–Hastings step. Thus, these steps can be combined to obtain

$$x_{i+1} = x_i \cos \sqrt{2 A \Delta t} + \zeta_i \sqrt{\frac{\epsilon}{A}} \sin \sqrt{2 A \Delta t}. \tag{A20}$$

To first order in Δt , the above agrees with other expressions for forming a realization of an OU process. The advantage of the above form is that for any size of Δt , the path asymptotically approaches the target Boltzmann (Gaussian) distribution.

Now, the ending point of the OU bridge needs to be addressed. The ending point of an OU bridge is also zero, but this is not the case for a typical realization of an OU process. To form a bridge, two OU processes can be combined by the following procedure. Assume that $Z^{(1)}$ and $Z^{(2)}$ are two different finite realizations of Equation (A13); both starting at the origin with endpoints $z_N^{(1)}$ and $z_N^{(2)}$, respectively. The linear combinations:

$$Z_{\pm}^{(3)} = \pm \left(Z^{(1)} \cos \kappa + Z^{(2)} \sin \kappa \right) \tag{A21}$$

are also realizations of the same OU process. Then, the angle κ can be chosen by $\tan \kappa = -z_N^{(1)} / z_N^{(2)}$, which makes the endpoints of $Z_{\pm}^{(3)}$ be $z_N^{(3)} = 0$. With this choice of κ , the OU processes $Z_{\pm}^{(3)}$ are OU bridges. Another way exists in the literature [22] for forming OU bridges. However, I found it to be numerically unstable due to the need to calculate the ratios of hyperbolic sines with large arguments.

Appendix C. BAB Splitting: Numerical Integration

In this section, the numerical algorithm is written down for the case when the BAB splitting is used. The method is schematically described in Equation (A22).

$$\begin{pmatrix} q_0 \\ v_0 \end{pmatrix} \xrightarrow[\text{Half step}]{\mathbf{B}(h/2)} \begin{pmatrix} q_0 \\ v_H \end{pmatrix} \xrightarrow[\text{Full step}]{\mathbf{A}(h)} \begin{pmatrix} q_1 \\ w_H \end{pmatrix} \xrightarrow[\text{Half step}]{\mathbf{B}(h/2)} \begin{pmatrix} q_1 \\ v_1 \end{pmatrix} \tag{A22}$$

Half step, B: $\{q_0, v_0\} \Rightarrow \{q_0, v_H\}$ with $M = (L + A_{ou}^2 \mathbb{1})$

$$M \frac{d}{dt} |v\rangle = - \left(|\phi(q)\rangle - A_{ou}^2 |q\rangle \right) \Rightarrow M |\Delta v_0\rangle = -\frac{h}{2} \text{Big}(|\phi(q_0)\rangle - A_{ou}^2 |q_0\rangle) \quad (\text{A23})$$

with $|\Delta v_0\rangle = |v_H\rangle - |v_1\rangle$.

Full step, A: $\{q_0, v_H\} \Rightarrow \{q_1, w_H\}$ with $\theta = h$

$$\frac{d}{dt} |q\rangle = |v\rangle \Rightarrow |q_1\rangle = \cos(\theta) |q_0\rangle + \sin(\theta) |v_H\rangle \quad (\text{A24})$$

$$\frac{d}{dt} |v\rangle = -|q\rangle \Rightarrow |w_H\rangle = -\sin(\theta) |q_0\rangle + \cos(\theta) |v_H\rangle \quad (\text{A25})$$

Half step, B: $\{q_1, w_H\} \Rightarrow \{q_1, v_1\}$ with $M = (L + A_{ou}^2 \mathbb{1})$

$$M \frac{d}{dt} |v\rangle = - \left(|\phi(q)\rangle - A_{ou}^2 |q\rangle \right) \Rightarrow M |\Delta v_1\rangle = -\frac{h}{2} \left(|\phi(q_1)\rangle - A_{ou}^2 |q_1\rangle \right) \quad (\text{A26})$$

with $|\Delta v_1\rangle = |v_1\rangle - |w_H\rangle$.

Appendix D. BAB Splitting: Energy Error

In the BAB splitting, the error in the effective energy is:

$$\begin{aligned} \Delta E^{(01)} &= \Phi(q_1) - \Phi(q_0) - \frac{1}{2} \langle q_1 | A_{ou}^2 | q_1 \rangle + \frac{1}{2} \langle q_1 | A_{ou}^2 | q_0 \rangle \\ &+ \frac{1}{2} \langle v_1 | M | v_1 \rangle - \frac{1}{2} \langle v_0 | M | v_0 \rangle + \frac{1}{2} \langle q_1 | M | q_1 \rangle - \frac{1}{2} \langle q_0 | M | q_0 \rangle \end{aligned} \quad (\text{A27})$$

which simplifies to:

$$\begin{aligned} \Delta E^{(01)} &= \Phi(q_1) - \Phi(q_0) - \frac{1}{2} \langle q_1 + q_0 | A_{ou}^2 | q_1 - q_0 \rangle \\ &+ \frac{1}{2} \langle v_1 + w_H | M | v_1 - w_H \rangle + \frac{1}{2} \langle v_H + v_0 | M | v_H - v_0 \rangle. \end{aligned} \quad (\text{A28})$$

One can express this in terms of q_0 and q_1 , using Equations (A23) and (A26) and:

$$|v_H\rangle = -|q_0\rangle \cot \theta + |q_1\rangle \csc \theta \quad (\text{A29})$$

$$|w_H\rangle = |q_1\rangle \cot \theta - |q_0\rangle \csc \theta \quad (\text{A30})$$

In the work by Beskos et al. [3], a similar method was constructed for the case where $A_{ou} = 0$. In the middle step of the numerical integration (Full Step A), the Crank–Nicolson method was used. This changes the rotation angle in Equations (A24) and (A25) to $\theta = \sin^{-1} \left(4h / (4 + h^2) \right)$. This in turn modifies the solutions for v_H and w_H used in Beskos et al., namely:

$$|v_H\rangle = \frac{1}{h} \left(|q_1\rangle - |q_0\rangle \right) + \frac{h}{4} \left(|q_1\rangle + |q_0\rangle \right) \quad (\text{A31})$$

$$|w_H\rangle = \frac{1}{h} \left(|q_1\rangle + |q_0\rangle \right) - \frac{h}{4} \left(|q_1\rangle - |q_0\rangle \right) \quad (\text{A32})$$

References

1. Englander, S.W.; Mayne, L. The case for defined protein folding pathways. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8253–8258. [[CrossRef](#)] [[PubMed](#)]
2. Onsager, L.; Machlup, S. Fluctuations and irreversible processes. *Phys. Rev.* **1953**, *91*, 1505. [[CrossRef](#)]
3. Beskos, A.; Pinski, F.; Sanz-Serna, J.; Stuart, A. Hybrid Monte Carlo on Hilbert spaces. *Stoch. Process. Their Appl.* **2011**, *121*, 2201–2230. [[CrossRef](#)]

4. Duane, S.; Kennedy, A.; Pendleton, B.J.; Roweth, D. Hybrid Monte Carlo. *Phys. Lett. B* **1987**, *195*, 216–222. [[CrossRef](#)]
5. Neal, R.M. In *Handbook of Markov Chain Monte Carlo*; Brooks, S., Gelman, A., Jones, G., Meng, X., Eds.; CRC Press: Boca Raton, FL, USA, 2011; Chapter 5, pp. 113–162.
6. Uhlenbeck, G.E.; Ornstein, L.S. On the Theory of the Brownian Motion. *Phys. Rev.* **1930**, *36*, 823–841. [[CrossRef](#)]
7. Malsom, P.J.; Pinski, F.J. Role of Ito's lemma in sampling pinned diffusion paths in the continuous-time limit. *Phys. Rev. E* **2016**, *94*, 042131. [[CrossRef](#)] [[PubMed](#)]
8. Malsom, P. Rare Events and the Thermodynamic Action. Ph.D. Thesis, University of Cincinnati, Cincinnati, OH, USA, 2015.
9. Trotter, H.F. On the product of semi-groups of operators. *Proc. Am. Math. Soc.* **1959**, *10*, 545–551. [[CrossRef](#)]
10. Strang, G. On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.* **1968**, *5*, 506–517. [[CrossRef](#)]
11. Rowlands, G. A numerical algorithm for Hamiltonian systems. *J. Comp. Phys.* **1991**, *97*, 235–239. [[CrossRef](#)]
12. Toxvaerd, S. Hamiltonians for discrete dynamics. *Phys. Rev. E* **1994**, *50*, 2271. [[CrossRef](#)] [[PubMed](#)]
13. Øksendal, B. *Stochastic Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2003.
14. Stuart, A.; Voss, J.; Wiberg, P. Conditional Path Sampling of SDEs And The Langevin MCMC Method. *Comm. Math. Sci.* **2004**, *2*, 685. [[CrossRef](#)]
15. Korol, R.; Rosa-Raíces, J.L.; Bou-Rabee, N.; Miller, T.F., III. Dimension-free path-integral molecular dynamics without preconditioning. *J. Chem. Phys.* **2020**, *152*, 104102. [[CrossRef](#)] [[PubMed](#)]
16. Pinski, F.J.; Stuart, A.M. Transition paths in molecules at finite temperature. *J. Chem. Phys.* **2010**, *132*, 184104. [[CrossRef](#)]
17. Leimkuhler, B.; Matthews, C. Efficient molecular dynamics using geodesic integration and solvent–solute splitting. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2016**, *472*, 20160138. [[CrossRef](#)] [[PubMed](#)]
18. Ottobre, M.; Pillai, N.S.; Pinski, F.J.; Stuart, A.M. A function space HMC algorithm with second order Langevin diffusion limit. *Bernoulli* **2016**, *22*, 60–106. [[CrossRef](#)]
19. Lavenda, B. Thermodynamic criteria governing the stability of fluctuating paths in the limit of small thermal fluctuations: critical paths in the limit of small thermal fluctuations: critical paths and temporal bifurcations. *J. Phys. A Math. Gen.* **1984**, *17*, 3353. [[CrossRef](#)]
20. Stratonovich, R. On the probability functional of diffusion processes. *Selected Trans. Math. Stat. Prob.* **1971**, *10*, 273–286.
21. Maruyama, G. Continuous Markov processes and stochastic equations. *Rend. del Circ. Mat. di Palermo* **1955**, *4*, 48–90. [[CrossRef](#)]
22. Barczy, M.; Kern, P. Sample path deviations of the Wiener and the Ornstein–Uhlenbeck process from its bridges. *Braz. J. Probab. Stat.* **2013**, *27*, 437–466. [[CrossRef](#)]