# Path-accelerated molecular dynamics: Parallel-in-time integration using path integrals

Jorge L. Rosa-Raíces, <sup>1</sup> Bin Zhang, <sup>2</sup> and Thomas F. Miller III<sup>1, a)</sup>

Massively parallel computer architectures create new opportunities for the performance of long-timescale molecular dynamics (MD) simulations. Here, we introduce the path-accelerated molecular dynamics (PAMD) method that takes advantage of distributed computing to reduce the wall-clock time of MD simulation via parallelization with respect to MD timesteps. The marginal distribution for the time evolution of a system is expressed in terms of a path integral, enabling the use of path sampling techniques to numerically integrate MD trajectories. By parallelizing the evaluation of the path action with respect to time and by initializing the path configurations from a non-equilibrium distribution, the algorithm enables significant speedups in terms of the length of MD trajectories that can be integrated in a given amount of wall-clock time. The method is demonstrated for Brownian dynamics, although it is generalizable to other stochastic equations of motion including open systems. We apply the method to two simple systems, a harmonic oscillator and a Lennard-Jones liquid, and we show that in comparison to the conventional Euler integration scheme for Brownian dynamics, the new method can reduce the wall-clock time for integrating trajectories of a given length by more than three orders of magnitude in the former system and more than two in the latter. This new method for parallelizing MD in the dimension of time can be trivially combined with algorithms for parallelizing the MD force evaluation to achieve further speedup.

#### I. INTRODUCTION

Molecular dynamics  $(MD)^{1,2}$  is the central tool for simulating chemical, biological, and materials systems, with new algorithms and hardware expanding the range of accessible timescales and length scales  $^{3-5}$ . Faster processors have played an important role in this expansion, although the most dramatic improvements in recent years have come from the number of available processors, rather than the clock-speed of the individual cores<sup>6,7</sup>. In particular, highly multi-threaded computer architectures have been used to parallelize the MD force evaluation, greatly reducing the wall-clock time needed to perform an individual MD step<sup>8-13</sup>. However, despite this progress in the parallelization of MD simulations with respect to the force evaluations (i.e., in space), less attention has been dedicated to the notion of parallelization with respect to the MD timesteps (i.e., in time).

The sequential nature of MD (i.e., the need to have access to a given timestep before the next timestep can be computed) would seem to discount the possibility of exploiting parallelization in time; nonetheless, methods for parallel-in-time integration are being developed and applied to MD simulation. Most approaches 14-17 are based on a prediction-correction paradigm that combines fine (i.e., accurate and expensive) and coarse (i.e., inaccurate and inexpensive) solvers to iteratively refine approximations of a trajectory in a convergent and parallel-intime fashion. A range of coarse solvers and iteration schemes have been employed to evaluate MD trajectories

of molecular systems with parallelization in the time domain  $^{18-22}$ , leading to order-of-magnitude reductions in the wall-clock time-to-solution with respect to sequential integration at the fine level of accuracy. Schemes for approximate long-timescale integration via trajectory splicing are an alternative route to parallelization in time, yielding accurate time evolution for systems that exhibit strong timescale separation on well-characterized regions of the potential energy landscape  $^{23}$ .

The current work takes a different approach to parallelizing MD in time. We demonstrate that by working with ensembles of trajectories in a path-integral framework, multiple processors can be employed to reduce the wall-clock time needed to evolve an MD trajectory of arbitrary length, without resorting to parallelization of the MD force evaluation. This method of parallelization for MD trajectories is independent of, and thus entirely complementary to, parallelization of the MD force evaluations, and it creates new opportunities to harness large numbers of available computer processors for the generation of long-timescale MD trajectories.

## II. METHOD

#### A. MD integration based on path distributions

In this work, we focus on the MD equation of motion governing Brownian (i.e., overdamped Langevin) dynamics under potential V at temperature  $\beta^{-1}$ ,

$$\dot{x}(t) = -\gamma^{-1} V'(x(t)) + \sqrt{2D} \,\dot{w}(t),\tag{1}$$

<sup>&</sup>lt;sup>1)</sup>Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, 91125

<sup>&</sup>lt;sup>2)</sup> Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, 02139

a) Correspondence: tfm@caltech.edu

where the diffusion coefficient D and the friction coefficient  $\gamma$  are related by the Einstein relation  $D=(\beta\gamma)^{-1}$ , and w(t) is the standard Wiener process. MD trajectories can be generated by discretizing Eq. 1 with various numerical integration schemes<sup>24–28</sup>, such as the forward Euler algorithm<sup>2</sup>

$$x(t+dt) - x(t) = -\gamma^{-1} V'(x(t)) dt + \sqrt{2Ddt} \xi,$$
 (2)

where dt is the discretization timestep, and  $\xi$  a standard Gaussian random variate. The marginal distribution associated with time evolution of the system by dt according to Eq. 2 is<sup>29</sup>

$$K(x(t+dt)|x(t);dt) \propto \exp \left\{ -\frac{dt}{4D} \left( \frac{x(t+dt) - x(t)}{dt} + \frac{V'(x(t))}{\gamma} \right)^2 \right\}, \quad (3)$$

such that the likelihood of a MD trajectory of length T = N dt that evolves the system along positions  $\mathbf{X} = \{x(t_0), x(t_1), \dots, x(t_N)\}$  at times  $t_n = t + n dt$  is

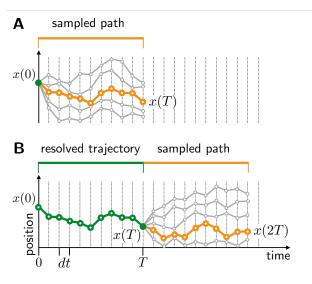
$$\prod_{n=0}^{N-1} K(x(t_{n+1})|x(t_n); dt) \equiv e^{-S[X]}, \tag{4}$$

where  $S[\boldsymbol{X}]$  is the action associated with the MD trajectory. From Eq. 4, the position of the time-evolved system at time T has a marginal distribution given by the path integral

$$K(x(T+t)|x(t);T) \propto \int_{\mathbb{R}} dx_1 \cdots \int_{\mathbb{R}} dx_{N-1} e^{-S[\mathbf{X}]}, (5)$$

where  $x_n = x(t_n)$ . It is clear that this path-integral formulation of the ensemble of MD trajectories provides an equivalent description of the time evolution of the system as Eq. 2. Numerous studies have explored this path-integral formulation with variations of the underlying equation of motion and of the discretization of the action<sup>30–34</sup>.

Setting aside issues of efficiency until section IIB, we note that the path-integral formulation of the marginal distribution for the time-evolved system offers a simple MD integration scheme, illustrated in Fig. 1. First, sampling from the distribution of paths of length T, with likelihood given by Eq. 4, is performed using Monte Carlo (MC) or related methods (Fig. 1A)<sup>35–37</sup>; by drawing a realization from this distribution, we obtain a segment of MD trajectory from time 0 to time T (illustrated by the heavy orange path in Fig. 1A). Then, by shifting from x(0) to x(T) along the sampled path, we resolve a trajectory from x(0) to x(T) (represented by the heavy green path in Fig. 1B) that is statistically equivalent to a realization from the Euler algorithm defined in Eq. 2. After shifting the tail of the path to x(T), we restart the path sampling to extend the trajectory from time T to time 2T. Iteration of this scheme will lead to the numerical integration of a MD trajectory of arbitrary length in time.



**Figure 1**. Illustration of a simple path-based MD integration scheme. **A** Sampling of the distribution of paths for the period of time from 0 to T. A particular path drawn from this distribtion is indicated in orange. **B** Shifting along a sampled path (orange path in **A**) from x(0) to x(T), thereby resolving the segment of MD trajectory indicated in green, and then resuming the sampling of the paths for the period of time from T to 2T.

Fig. 2 illustrates a generalized version of the integration scheme presented in Fig. 1. Fig. 2A repeats Fig. 1A; we first sample a path of length T that is discretized into  $N_{\rm path}$  timesteps (where  $N_{\rm path} = T/dt$ ) to obtain a realization of the path that is consistent with the marginal distribution of the time-evolved system for each time  $\Delta t \leq T$ . Then, in Fig. 2B, we shift the tail of the sampled path (indicated in orange) by  $N_{\text{shift}}$  timesteps (where  $N_{\text{shift}} = \Delta t/dt$ ) to the position  $x(\Delta t)$ . With the remaining segment of the path now located at positions  $\{x(\Delta t), x(\Delta t + dt), \dots, x(T)\}\$ , we grow the path out of x(T) by  $N_{\text{shift}}$  timesteps to regenerate the original number of timesteps in the path. The positions of the system at the regenerated timesteps can be drawn from any distribution (and in Fig. 2B they are obtained via straight-line extrapolation). Finally, as illustrated in Fig. 2C, sampling is again performed to generate a path consistent with evolution from time  $\Delta t$  to time  $T + \Delta t$ ; this sampling removes any artifacts introduced by the arbitrary distribution used to grow the shifted path. As for the scheme in Fig. 1, iteration of the scheme in Fig. 2 yields an MD trajectory of arbitrary length in time that is statistically equivalent to a realization from the Euler algorithm. The only difference between these two pathbased integration schemes is that Fig. 1 involves shifting along the full length of the sampled path, whereas Fig. 2 involves shifting only a fraction of the way along the sampled path.

Just like the Euler scheme in Eq. 2, the schemes illustrated in Figs. 1 and 2 enable the numerical integration of MD trajectories. Each of these integration schemes

consist of sequential iterations of an elementary step that predicts the state of the system at some later time. In the Euler scheme, the prediction can be conducted analytically based on the distribution defined in Eq. 3. In pathbased integration schemes employing path lengths longer that dt, however, no such analytical expression exists for general systems; path sampling is therefore needed before each shifting event to generate time-evolved system positions consistent with the correct marginal distribution.

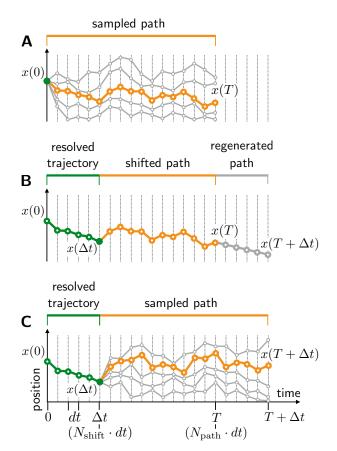


Figure 2. Illustration of a path-accelerated molecular dynamics (PAMD). A Sampling of the distribution of paths for the period of time from 0 to T. A particular path drawn from this distribtion is indicated in orange. B Shifting along a sampled path (orange path in A) from x(0) to  $x(\Delta t)$ , thereby resolving the segment of MD trajectory indicated in green, and regenerating the full length of the path by drawing positions for the system from time T+dt to  $T+\Delta t$  from an arbitrary distribution. C Sampling of the distribution of paths for the period of time from  $\Delta t$  to  $\Delta t + T$ .

For the scheme in Fig. 2, it is assumed that the path distributions in parts A and C are well sampled. For MC path sampling algorithms, this implies that the number of configurations of the path that are sampled in parts A and C,  $N_{\text{sample}}$ , is large in comparison to the number that is needed to generate uncorrelated realizations of the path. If  $N_{\text{sample}}$  is smaller than this decorrelation number, then the distribution of paths that is generated

in part C may be biased by the way in which the path was regenerated in part B. However, the only requirement for generating accurate MD trajectories using the scheme in Fig. 2 is accurate sampling of paths consistent with the marginal distribution  $K(x(\Delta t)|x(0);\Delta t)$ ; it is not essential that the marginal distribution associated with the full path, K(x(T)|x(0);T), be sampled without error. Recalling that  $T=N_{\rm path}\,dt$  and  $\Delta t=N_{\rm shift}\,dt$ , this suggests that for a given path sampling algorithm, there is an interplay between parameters  $N_{\rm sample},\,N_{\rm path}$  and  $N_{\rm shift}$ ; for given values of  $N_{\rm path}$  and  $N_{\rm shift}$ , there is an associated number of path configurations  $(N_{\rm sample})$  that must be sampled in order to generate a sufficiently accurate marginal distribution  $K(x(\Delta t)|x(0);\Delta t)$ .

This interplay between  $N_{\rm sample}$ ,  $N_{\rm path}$  and  $N_{\rm shift}$  is illustrated in Figs. 3A-C, which plot the error in the marginal distribution generated using the scheme in Fig. 2 for the Brownian dynamics of a harmonic oscillator. Full calculation details are provided in Section III. The error plotted in Figs. 3A-C corresponds to the Kullback-Leibler divergence,

$$D_{\mathrm{KL}}(t) = \left\langle \int_{\mathbb{R}} \mathrm{d}x_t \, P(x_t|x_0;t) \log \frac{P(x_t|x_0;t)}{Q(x_t|x_0;t)} \right\rangle_{x_0}, \quad (6)$$

where  $Q(x_t|x_0;t)$  is the marginal distribution estimated using sampled paths from the scheme in Fig. 2, and  $P(x_t|x_0;t)$  is the exact marginal distribution. The angled brackets denote averaging with respect to the Boltzmann distribution of positions that is sampled by the exact dynamics,  $P(x_0) = Z^{-1}e^{-\beta V(x_0)}$ , where  $Z = \int_{\mathbb{R}} \mathrm{d}x_0 \, e^{-\beta V(x_0)}$  is the partition function. For a harmonic oscillator with potential  $V(x) = \frac{1}{2}kx^2$  38,

$$P(x_t|x_0;t) \propto \exp\left\{-\frac{\beta k}{2} \frac{(x_t - e^{-\gamma^{-1}kt}x_0)^2}{(1 - e^{-2\gamma^{-1}kt})}\right\},$$
 (7)

and we employ k=1 for the oscillator force constant,  $\beta^{-1}=1$  for the temperature and  $\gamma=1$  for the friction coefficient.  $D_{\mathrm{KL}}(t)$  returns non-negative values that approach 0 as  $Q(x_t|x_0;t)$  more accurately reproduces  $P(x_t|x_0;t)$ . As a function of time t along the sampled paths,  $D_{\mathrm{KL}}(t)$  is plotted in Fig. 3A-C for seven simulations that employ the scheme in Fig. 2 with different values of  $N_{\mathrm{sample}}$ ,  $N_{\mathrm{path}}$  and  $N_{\mathrm{shift}}$ . The results correspond to sampled paths of length  $T\geq 1$  that are discretized into timesteps of dt=1/32, and  $D_{\mathrm{KL}}(t)$  is evaluated for the numerically generated marginal distributions at times  $dt\leq t\leq 1$ .

Fig. 3A addresses the case where  $N_{\rm sample}$  and  $N_{\rm path}$  are held fixed and various values of  $N_{\rm shift}$  are used. Comparison of the blue  $(N_{\rm shift}=4)$ , green  $(N_{\rm shift}=8)$  and red  $(N_{\rm shift}=16)$  curves shows that for a given value of  $N_{\rm sample}$ , smaller values of  $N_{\rm shift}$  lead to smaller errors in the numerically generated marginal distribution. Using the scheme in Fig. 2, a given segment of the path is sampled  $N_{\rm path} \cdot N_{\rm sample}/N_{\rm shift}$  times before it is used to generate the marginal distribution for the integration of

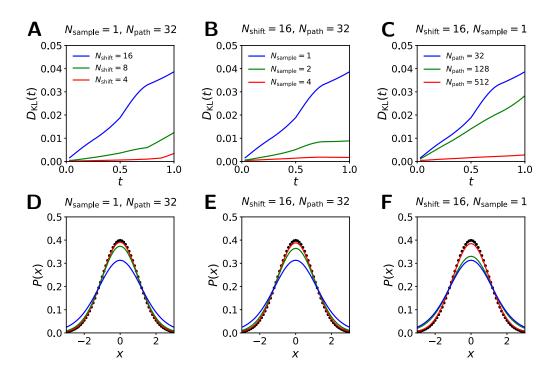


Figure 3. The interplay of the parameters  $N_{\rm shift}$ ,  $N_{\rm sample}$  and  $N_{\rm path}$  in determining the accuracy of the PAMD integration scheme. Panels  $\bf A$ ,  $\bf B$  and  $\bf C$  show the Kullback-Leibler divergence  $D_{\rm KL}(t)$  defined in Eq. 6, from the marginal distribution generated using the scheme in Fig. 2 to the exact marginal distribution for the dynamics of an overdamped harmonic oscillator. Panels  $\bf D$ ,  $\bf E$  and  $\bf F$  plot the position distributions P(x) sampled by the MD trajectories integrated using the numerically generated marginal distributions from panels  $\bf A$ ,  $\bf B$  and  $\bf C$ , respectively, in comparison to the exact position distribution shown in black dots. It is seen that the accuracy of the numerically generated marginal distribution dictates that of the integrated MD trajectory, and improves with decreasing  $N_{\rm shift}$  (panels  $\bf A$  and  $\bf D$ ), increasing  $N_{\rm sample}$  (panels  $\bf B$  and  $\bf E$ ) or increasing  $N_{\rm path}$  (panels  $\bf C$  and  $\bf F$ ) for given values of the remaining parameters (indicated at the top of each panel).

the MD trajectory; therefore, smaller values of  $N_{\rm shift}$  lead to better sampling of the path distribution and smaller errors in the marginal distribution.

Fig. 3B illustrates a second scenario where  $N_{\rm shift}$  and  $N_{\rm path}$  are held fixed and increasing values of  $N_{\rm sample}$  are used. Comparison of the blue ( $N_{\rm sample}=1$ ), green ( $N_{\rm sample}=2$ ) and red ( $N_{\rm sample}=4$ ) curves shows that for a given value of  $N_{\rm shift}$ , larger values of  $N_{\rm sample}$  (i.e., more sampling per shifting event) lead to smaller errors in the numerically generated marginal distribution for integrating the MD trajectory. This result is intuitive, as more sampling leads to elimination of the bias associated with the arbitrary distribution used in the regeneration of the full length of the path.

In Fig. 3C,  $N_{\rm shift}$  and  $N_{\rm sample}$  are held fixed as the length of the sampled path  $(N_{\rm path})$  is increased while keeping the path discretization timestep unchanged. Comparison of the blue  $(N_{\rm path}=32)$ , green  $(N_{\rm path}=128)$  and red  $(N_{\rm path}=512)$  curves demonstrates that increasing the total length  $N_{\rm path} \cdot dt$  of the sampled paths improves the accuracy of the numerically generated marginal distribution. Like decreasing  $N_{\rm shift}$  for a given  $N_{\rm path}$  (as in Fig. 3A), increasing  $N_{\rm path}$  for a given  $N_{\rm shift}$  allows for more sampling of each segment of the path employed to generate the marginal distribution as-

sociated with the MD time evolution.

While Figs. 3A-C illustrate the errors in the marginal distribution generated using the scheme in Fig. 2, Figs. 3D-F illustrate the corresponding errors in the equilibrium distribution that is sampled by the integrated MD trajectories. For the various employed parameters, the results from path-based MD integration are compared to the exact Boltzmann distribution (dots) and, as expected, the errors in the marginal distribution with given values of  $N_{\rm sample}, N_{\rm path}$  and  $N_{\rm shift}$  are reflected in the distribution of positions that are visited in the MD trajectories.

In summary, Fig. 3 demonstrates that decreasing  $N_{\rm shift}$ , increasing  $N_{\rm sample}$ , or increasing  $N_{\rm path}$  leads to greater accuracy in the integrated MD trajectories; as will be shown in Section IIB, the interplay between these three parameters is also critical for determining the computational efficiency of MD integration using the scheme in Fig. 2.

Before addressing efficiency, however, Fig. 4 illustrates that the integration scheme in Fig. 2 is a non-equilibrium relaxation process for the segments of the sampled path. For the case of the harmonic oscillator, Figs. 3A-C indicate that errors in the numerically generated marginal distributions are typically larger at the nose of the sam-

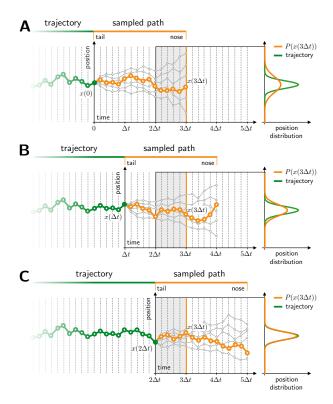


Figure 4. PAMD integrates equilibrium trajectories by relaxing non-equilibrium path segments. The state of the sampled path is shown for three consecutive iterations of the pathbased integration scheme applied to a Brownian harmonic oscillator on the left of panels A, B, and C. The gray box in each panel highlights the configuration of a particular segment of the path after each iteration of the integration scheme. On the right of each panel are plotted the distributions  $P(x(3\Delta t))$  of positions  $x(3\Delta t)$  sampled by the boxed path segment at time  $3\Delta t$  (orange curves), along with the distribution of configurations sampled by the integrated trajectory (green curve). The path segment inside the gray box in each panel reaches equilibrium by undergoing sampling as it shifts from the nose to the tail of the sampled path; accordingly, the distribution of sampled positions approaches that sampled by the integrated trajectory.

pled path. This trend emerges because the integration scheme regenerates path segments in configurations that are out of equilibrium (Fig. 2B). The light gray paths in Fig. 4 correspond to independent realizations of the sampled path (orange) obtained while integrating a harmonic oscillator trajectory. Since the segment at the nose of the path has undergone little sampling after regeneration, it is far from equilibrium with respect to the distribution of segments of an equilibrium harmonic oscillator trajectory (Fig. 4A; distributions at right). However, as that segment works its way from the nose to the tail of the path, it is sampled with increasing accuracy (Figs. 4B) and C). This relaxation process is illustrated by the distribution of positions,  $P(x(3\Delta t))$ , sampled by the foremost end  $(x(3\Delta t))$  of the path segment in the grey box throughout Fig. 4; as the orange curves indicate, this

distribution approaches that sampled by the harmonic oscillator trajectory (green curve) as the segment relaxes toward equilibrium and simultaneously approaches the tail of the sampled path.

#### B. An opportunity for speedup

At face value, the path-based integration scheme in Fig. 2 may appear to be inefficient, given the difficulties of sampling uncorrelated paths<sup>33,35,39,40</sup>. Yet, it has several potential advantages: Firstly, there is an opportunity for parallelization, given that typical expressions for the path action incur a dominant source of computational cost from the evaluation of the forces in the system along the path (V'(x)) in Eq. 3). These forces can be evaluated independently, enabling straightforward parallelization of the action with respect to time. Secondly, regeneration of the path to its full length following shifting (Fig. 2B) can be performed using an arbitrary distribution to obtain the system positions for the regenerated timesteps; consequently, it is possible to carry out this operation at a cost that is negligible relative to evaluation of the MD forces. Thirdly, MC path sampling provides a numerically more stable way for generating trajectories than integration of the discretized equations of motion<sup>31</sup>; thus, a sufficiently accurate MD trajectory may be obtained with the path-based integration scheme at a larger timestep than a conventional Brownian dynamics integrator would allow.

The above considerations suggest that the scheme in Fig. 2 could lead to reduction of the wall-clock time associated with MD integration, in comparison with standard methods. To quantify the speedup achieved with the new scheme, we introduce a measure  $\chi$ , with  $\chi^{-1}$ defined as the number of force evaluations per processor per step of time  $dt_{\rm E}$ , where  $dt_{\rm E}$  is the timestep used by the Euler algorithm to integrate the Brownian dynamics. The wall-clock speedup of the path-based integration scheme is thus  $\chi$ , assuming that (i) evaluation of the MD forces dominates the cost of the evaluation of the path action, (ii) parallel computer processors are used to independently evaluate the forces along the discretized path, and (iii) regeneration of the full length of the path following shifting (Fig. 2B) is performed without evaluating the MD forces. It is clear that for the Euler algorithm,  $\chi = 1$ , such that this measure provides a simple basis of comparison of the wall-clock time for the proposed pathbased integration scheme (which employs parallelization in time) versus the wall-clock time for a conventional MD integration scheme (which does not). In the current work, we set aside the complementary issue of speeding up MD integration via parallelization within the force evaluation.

For a general implementation of the integration scheme in Fig. 2, the expression for  $\chi$  is obtained as follows. Recalling previously introduced notation, we employ sampled paths of length T that are discretized with a

timestep of dt, which may be different (and is typically larger) than the numerically stable timestep for the Euler algorithm,  $dt_{\rm E}$ . Let  $N_{\rm force}$  be the number of MD force evaluations that are required during path sampling per shifting event, which depends on both  $N_{\rm sample}$  and the details of the path sampling algorithm, and let  $N_{\rm procs}$  be the number of employed parallel processors. Since the number of force evaluations per processor per shifting event is given by  $N_{\rm force}/N_{\rm procs}$ , the speedup is

$$\chi = N_{\text{shift}} \cdot \frac{N_{\text{procs}}}{N_{\text{force}}} \cdot \frac{dt}{dt_{\text{E}}}.$$
 (8)

Eq. 8 shows that the path-based integration scheme in Fig. 2 offers the possibility for reduction of the wallclock time needed to compute MD trajectories, relative to conventional MD. Factors that enable this speedup include the increase in the discretization timestep (dt)relative to that possible for conventional MD ( $dt_{\rm E}$ ), maximization of the number of integrated timesteps per shifting event  $(N_{\text{shift}})$ , maximization of the number of parallel processors to perform the independent force evaluations associated with the calculation of the path action  $(N_{\rm procs})$ , and minimization of the number of force evaluations needed per shifting event  $(N_{\text{force}})$ . As will be shown in Section IV, this approach indeed enables substantial speedups in the integration of MD trajectories while preserving the accuracy of the dynamics, and we henceforth refer to the method as path-accelerated molecular dynamics (PAMD).

## III. CALCULATION DETAILS

In the current work, we implement the PAMD method with sampling of the path distribution via the multilevel sliding and sampling algorithm for stochastic dynamics<sup>33</sup>. For a path of  $N_{\rm path}$  timesteps, a total of  $L = \log_2 N_{\text{path}}$  levels are defined (Fig. 5A); finer levels (smaller values of the level index  $1 \le l \le L$ ) correspond to partitions of the path into fragments of increasingly smaller length where the local configuration of the path is sampled. In accordance with the sliding and sampling algorithm, neighboring path fragments share endpoints that are chosen randomly such that the length of the fragments varies from 1 to  $2^l$  timesteps; we call this random fragmentation. For all fragmentations of the path at level l, internal fragments of the path are of length  $2^{l}$  and fragments at the termini of the path have a combined length of  $2^{l}$ . During a MC step for a given fragmentation of the path (Fig. 5B), the system positions at shared endpoints of neighboring path fragments are held fixed to permit mutually independent updates of the fragment configurations. Furthermore, the position of the system at the tail endpoint of the path is always fixed throughout the MC step, whereas that at the nose endpoint of the path undergoes sampling together with the nose fragment. Path fragment configurations are updated according to the Metropolis-Hastings criterion<sup>41,42</sup>,

with trial configurations drawn from a distribution that satisfies the boundary conditions at the fragment endpoints. Random fragmentation of the path is performed between MC steps, so that fixed system positions at previous fragment endpoints can be sampled during subsequent steps (Fig. 5C).

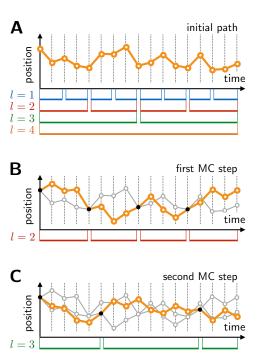


Figure 5. Illustration of the sliding and sampling algorithm. A Multilevel representation of a path with  $N_{\rm path}=16$  timesteps (L=4 levels), shown for one Cartesian coordinate. An initial configuration of the path is shown in orange. B Update of the path configuration in A via a MC step at level l=2. Fixed system positions along the path are represented with black dots, configurations of the path before the update in gray, and the configuration of the path after the update in orange. C MC step at level l=3 following that shown in B. The new fragmentation of the path allows for updates of system positions that were held fixed in previous MC steps.

At each MC step, a level is randomly selected between  $l_{\rm min}$  and  $l_{\rm max}$ , with  $1 \leq l_{\rm min} \leq l_{\rm max} \leq L$ . The calculations reported here employ  $l_{\rm min} > 1$  and  $l_{\rm max} < L$ , such that not all levels are directly sampled. The choice of  $l_{\rm max} < L$  corresponds to excluding the direct sampling of levels associated with long path fragments, on the basis of negligible acceptance. The choice of  $l_{\rm min} > 1$  corresponds to excluding the direct sampling of levels associated with short path fragments, as these are trivially updated via the direct sampling of longer fragments at coarser levels. Despite these choices, the sampling remains ergodic due to the random fragmentation of the path that occurs between MC steps. <sup>33</sup>

The distribution of paths used to generate trial configurations in each application is chosen to maximize

**Table I.** Summary of notation employed to describe the path-based MD integration scheme introduced in the current work. The gray region corresponds to the definitions of parameters specific to the path sampling algorithm used.

-dt	timestep for discretization of the sampled path
$N_{ m path}$	number of timesteps in the sampled path
$N_{ m procs}$	number of processors for parallel-in-time force evaluations
$N_{ m force}$	number of force evaluations per shifting event
$N_{ m sample}$	number of MC steps per shifting event
$N_{ m shift}$	number of timesteps shifted
L	total number of levels in the sampled path
$l_{\min}$	finest sampled level of the path
$l_{ m max}$	coarsest sampled level of the path

the statistical efficiency of the sampling (i.e., minimize  $N_{\text{sample}}$ ) without requiring evaluation of the MD forces. For the harmonic oscillator, trials are drawn from the distribution of free particle paths. For simulations of the Lennard-Jones liquid, trials are drawn from the path distribution of a fluid of hard spheres with diameter  $\sigma_{\rm HS}$ ; this strategy reduces the number of force evaluations needed to obtain likely Lennard-Jones path configurations by excluding those with high interparticle overlap from the ensemble of trial paths. The likelihood of a path at the hard-sphere level is evaluated using an approximation of the pair propagator for diffusive hard spheres.<sup>43</sup> For both applications, regeneration of the sampled path after shifting is performed with the same distribution used to generate trial configurations for the path sampling.

In total, each MC step involves a total number of  $N_{\rm path}$  MD force evaluations. Since  $N_{\rm force}$  is defined as the number of force evaluations per shifting event, and since  $N_{\rm sample}$  is the number of MC steps per shifting event, we have

$$N_{\text{force}} = N_{\text{sample}} \cdot N_{\text{path}}.$$
 (9)

As we seek to maximize the wall-clock speedup via parallelization of these independent force evaluations, we employ one processor per force evaluation, and thus

$$N_{\text{procs}} = N_{\text{path}}.$$
 (10)

Additional parallelization within the force evaluation is of course possible, but is not considered in the current work. Thus, we insert Eqs. 9 and 10 into Eq. 8 to arrive at the following expression for the PAMD speedup:

$$\chi = \frac{N_{\text{shift}}}{N_{\text{sample}}} \cdot \frac{dt}{dt_{\text{E}}}.$$
 (11)

See Table I for a summary of terms.

### IV. RESULTS

We now apply the PAMD algorithm to two model systems, considering time evolution according to Brownian

dynamics (Eq. 1) with  $\gamma=1$  in appropriately reduced units. Like the Euler algorithm (Eq. 2), PAMD is a rigorous and formally exact way to integrate the dynamics of Eq. 1, yet the numerical accuracy of the trajectories depends on the parameters employed. In each application, we examine the relationship between the number of parallel processors employed and the speedup in the PAMD algorithm relative to the Euler algorithm ( $\chi$  in Eq. 11), subject to the requirement that the MD trajectories integrated using both PAMD and the Euler algorithm preserve well-defined measures of accuracy.

In the current work, we focus exclusively on wall-clock speedups achieved via parallelization of the MD integration in time, setting aside the separate and complementary issue of parallelizing the force evaluation at each timestep. All reported speedups for PAMD in the current study are theoretical; they are obtained from Eq. 11 under the stated assumptions.

#### A. Harmonic oscillator

Here, we consider the example of an overdamped harmonic oscillator, with potential  $V(x) = \frac{1}{2}x^2$ . Two measures of the accuracy of the integrated MD trajectories are considered. The first reports on the degree to which the trajectories sample the correct equilibrium distribution.

$$E_{\rm eq} = \frac{\sqrt{\int_{\mathbb{R}} dx \, |P(x) - P_{\rm s}(x)|^2}}{Z} \tag{12}$$

where P(x) is the exact Boltzmann distribution, Z is the associated partition function, and  $P_{\rm s}(x)$  is the equilibrium distribution of positions sampled by the numerical integration schemes. The second measure of error reports on the accuracy of the MD time evolution. Specifically, we consider the autocovariance function

$$C(t) = \langle x(t')x(t'+t)\rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt' \, x(t')x(t'+t),$$
(13)

which is a simple exponential function for the overdamped harmonic oscillator<sup>38</sup>,

$$C(t) = \langle x^2 \rangle \exp(-\kappa t),$$
 (14)

where the angled brackets indicate Boltzmann averaging. The second measure of error is thus

$$E_{\rm dyn} = \frac{\sqrt{(\kappa - \kappa_{\rm s})^2}}{k},\tag{15}$$

where  $\kappa = \gamma^{-1} = 1$  is the exact decay constant and  $\kappa_s$  is the decay constant obtained by fitting the exponential decay of the autocovariance from the numerically integrated MD trajectories. Specifically,  $\kappa_s$  is obtained by averaging over 100 independent trajectories of length  $10^5$  time units that are divided into  $10^3$  non-overlapping time series, for which the log-autocovariance is linearly fit in the range  $t \in [0, 4]$ . Simulation parameters for integration of the MD trajectories in this application are chosen to ensure that both measures of error remain below 3%.

Table II indicates that with a timestep of  $dt_{\rm E}=0.025$ , the Euler algorithm yields error values of  $E_{\rm eq}=0.3\%$  and  $E_{\rm dyn}=1.0\%$ . Also shown in the table are parameters for three separate PAMD simulations that obtain speedups of  $\chi=16$ , 128 and 1024 with respect to the Euler algorithm. For both the equilibrium distribution and the autocovariance function, Fig. 6 shows the comparison of the exact results and those obtained using PAMD with the aforementioned speedups. The accuracy of the PAMD trajectories is clearly preserved in all simulations, as indicated by the plotted results and the reported values of  $E_{\rm eq}$  and  $E_{\rm dyn}$ .

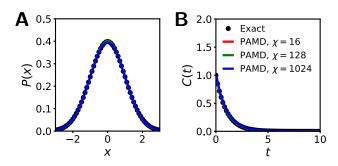


Figure 6. For the harmonic oscillator, comparison of PAMD results (colored lines) with exact results (black dots) for **A** the Boltzmann distribution P(x) and **B** the autocovariance function C(t). The PAMD results correspond to Simulation 1 (red), Simulation 2 (green) and Simulation 3 (blue) in Table II, which respectively achieve speedups of  $\chi=16,\ 128,\ and\ 1024$  relative to the Euler algorithm.

For all PAMD simulations reported in Table II, a significant component of the speedup comes from the 16-fold larger timestep that can be employed in the path-based scheme (dt = 0.4 vs.  $dt_{\rm E} = 0.025$ ). The larger speedups achieved in Simulation 2 ( $\chi = 128$ ) and Simulation 3 ( $\chi = 1024$ ), in comparison to Simulation 1 ( $\chi = 16$ ), arise

from the larger ratios of  $N_{\rm shift}$  to  $N_{\rm sample}$  that are used in these simulations ( $N_{\rm shift}/N_{\rm sample}=8$  for Simulation 2 and  $N_{\rm shift}/N_{\rm sample}=64$  for Simulation 3) in comparison to Simulation 1 (for which  $N_{\rm shift}/N_{\rm sample}=1$ ). The higher frequency of shifting events associated with larger values of the ratio  $N_{\rm shift}/N_{\rm sample}$  places greater demand on the efficiency of the path sampling, and a larger number of timesteps in the sampled path ( $N_{\rm path}$ ; hence, a larger number of parallel processors,  $N_{\rm procs}$ , per Eq. 10) is needed to allow path segments to undergo a sufficient number of MC steps before they are used to generate the marginal distribution for the MD trajectories, as discussed in connection with Figs. 3 and 4.

**Table II.** Summary of PAMD simulation parameters used for the harmonic oscillator application. The gray region corresponds to the parameters specific to the path sampling algorithm used.

	Euler		
$dt_{ m E}$	0.025		
$E_{ m eq}$	0.3%		
$E_{ m dyn}$	1.0%		
	PAMD		
	Simulation 1	Simulation 2	Simulation 3
dt	0.4	0.4	0.4
$N_{ m path}$	16	256	4096
$N_{ m sample}$	2	1	1
$N_{ m shift}$	2	8	64
L	4	8	12
$l_{ m min}$	3	4	4
$l_{ m max}$	3	5	8
${\chi}$	16	128	1024
$E_{ m eq}$	0.3%	0.3%	0.5%
$E_{ m eq} \ E_{ m dyn}$	0.3%	1.1%	2.9%

## B. Lennard-Jones liquid

Here, we apply PAMD to a model for a molecular liquid. The pairwise interaction between particles is described using the standard cut-and-force-shifted Lennard-Jones potential,<sup>2</sup>

$$U(r) = \begin{cases} u(r) - u(r_{\rm c}) - (r - r_{\rm c}) u'(r_{\rm c}), & r \le r_{\rm c} \\ 0, & r > r_{\rm c} \end{cases}, (16)$$

where  $u(r) = 4\epsilon \left\{ (\sigma/r)^{12} - (\sigma/r)^6 \right\}$ ; throughout, we take  $\epsilon = 1$  and  $\sigma = 1$ . The system consists of 27 particles placed in a cubic box at reduced density  $\rho\sigma^3 = 0.50$  and at constant reduced inverse temperature  $\beta\epsilon = 0.74$ . Simulations are performed with periodic boundary conditions at constant volume, and the cutoff distance  $r_{\rm c}$  corresponds to half of the simulation box-length.

As described in Section III, the reported simulations for the Lennard-Jones liquid employ trial configurations drawn from a distribution of paths for a fluid of hard spheres with diameter  $\sigma_{\rm HS}$ . To prevent the pathsampling bias from affecting the accuracy of the integrated trajectories, we employ a hard-sphere schedule that varies as a function of the path-time  $\tau$ ,  $\sigma_{\rm HS}(\tau)$ ; path configurations are then sampled in accordance with the path-time dependent potential

$$V(r; \sigma_{\rm HS}(\tau)) = U(r) + U_{\rm HS}(r; \sigma_{\rm HS}(\tau)), \tag{17}$$

where U(r) is defined in Eq. 16, and  $U_{\rm HS}(r;\sigma_{\rm HS})$  is the hard-sphere potential

$$U_{\rm HS}(r;\sigma_{\rm HS}) = \begin{cases} +\infty, & r \le \sigma_{\rm HS} \\ 0, & r > \sigma_{\rm HS} \end{cases} . \tag{18}$$

The schedule  $\sigma_{\rm HS}(\tau)$  is chosen such that configurational volume is excluded at the nose of the path for enhanced sampling efficiency, and no volume is excluded at the tail of the path where the marginal distribution for MD integration is sampled (Fig. 7); in this way, path segments regenerated from the hard-sphere distribution are subsequently relaxed into the Lennard-Jones distribution as they shift from the nose to the tail of the path.

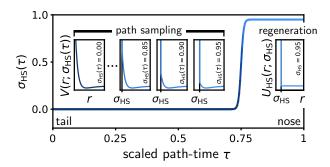


Figure 7. Hard-sphere schedule  $\sigma_{\rm HS}(\tau)$ , versus scaled pathtime  $\tau$ , employed in the reported Lennard-Jones simulations. Insets show slices of the path-time dependent potential  $V(r;\sigma_{\rm HS}(\tau))$  employed to sample path configurations (Eq. 17) at schedule values corresponding to biased ( $\sigma_{\rm HS} \simeq 1$ ) and unbiased ( $\sigma_{\rm HS} = 0$ ) sampling of the equilibrium path distribution, and the hard-sphere potential  $U(r;\sigma_{\rm HS})$  used to regenerate path segments (Eq. 18).

The accuracy of the integrated MD trajectories is evaluated in terms of the radial distribution function g(r) and the self-diffusion coefficient D, using the respective error measures

$$E_{\rm eq} = \frac{\sqrt{\int_0^{r_c} dr \, |g(r) - g_s(r)|^2}}{\int_0^{r_c} dr \, g(r)}$$
(19)

and

$$E_{\rm dyn} = \frac{\sqrt{(D - D_{\rm s})^2}}{D},\tag{20}$$

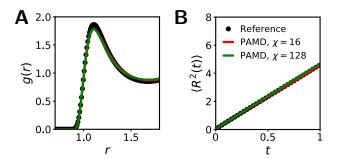
where g(r) and D are reference quantities obtained using the Euler algorithm with a small timestep  $(5 \times 10^{-5})$ 

Lennard-Jones time units), and  $g_s(r)$  and  $D_s$  are obtained using PAMD and the Euler algorithm with larger timesteps. The diffusion coefficient is given by  $D = \frac{1}{6} \lim_{t \to \infty} \partial_t \langle R^2(t) \rangle$ , where

$$\langle R^2(t) \rangle = \lim_{T \to \infty} \int_0^T dt' \frac{1}{N} \sum_{i=1}^N |\mathbf{r}_i(t'+t) - \mathbf{r}_i(t')|^2 \quad (21)$$

is the mean-square displacement,  $\mathbf{r}_i$  the position of the ith particle, and N the number of particles<sup>1</sup>.  $\langle R^2(t) \rangle$  is obtained by averaging over 100 independent trajectories that are divided into 100 non-overlapping time series of length 1, in Lennard-Jones time units, and a linear fit is performed in the range  $t \in [0.2, 1]$  to evaluate  $D_{\rm s}$  for the PAMD and Euler simulations.

Table III indicates that at a timestep of  $dt_{\rm E} = 2.5 \times$  $10^{-4}$ , the Euler algorithm yields error values of  $E_{\rm eq} =$ 1.4% and  $E_{\rm dyn} = 0.1\%$ ; larger timesteps were found to lead to unstable Euler trajectories. Also shown in Table III are two PAMD simulations that lead to 16-fold  $(\chi = 16; \text{ Simulation 1}) \text{ and } 128\text{-fold } (\chi = 128; \text{ Simu-}$ lation 2) reductions of the wall-clock time required to generate equivalently accurate MD trajectories for the Lennard-Jones liquid via the Euler algorithm, using simulation parameters that kept error values below 5%. The radial distribution functions and mean-square displacements obtained from these two simulations are plotted with the corresponding reference quantities in Figs. 8A and B. Excellent agreement between the PAMD and reference quantities is evident in the plots and from the values of  $E_{\rm eq}$  and  $E_{\rm dyn}$  reported in Table III.



**Figure 8.** For the Lennard-Jones liquid at  $\beta\epsilon=0.74$  and  $\rho\sigma^3=0.50$ , comparison of PAMD results (colored lines) with numerically exact results (black dots) for **A** the radial distribution function and **B** the mean-square displacement. The PAMD results correspond to Simulation 1 (red) and Simulation 2 (green) in Table III, which respectively achieve speedups of  $\chi=16$  and 128 relative to the Euler algorithm.

As in the harmonic oscillator application, the speedups reported in Table III for the Lennard-Jones liquid are partially enabled by the use of a larger timestep in PAMD  $(dt=5\times 10^{-4})$  than is possible for stable numerical integration via the Euler method  $(dt_{\rm E}=2.5\times 10^{-4})$ . The remaining speedup in both simulations comes from using shift lengths that integrate  $N_{\rm shift}=8$  (Simulation 1) and

 $N_{\rm shift}=64$  (Simulation 2) timesteps of MD trajectory at a rate of  $N_{\rm sample}=1$  path-sampling steps per shifting event. Accurate integration at these speedups requires efficient sampling of path modes that are commensurate with the shifting timescale (8 dt in Simulation 1 and 64 dt in Simulation 2); accordingly, long paths ( $N_{\rm path}=128$  in Simulation 1 and  $N_{\rm path}=4096$  in Simulation 2) are employed in both simulations.

**Table III.** Summary of PAMD simulation parameters used for the application to the Lennard-Jones liquid. The gray region corresponds to the parameters specific to the path sampling algorithm used.

	Euler			
$dt_{\rm E}$	$2.5 \times 10^{-4}$			
$E_{ m eq}$	1.4%			
$E_{ m dyn}$	0.1%			
	PAMD			
	Simulation 1	Simulation 2		
dt	$5 \times 10^{-4}$	$5 \times 10^{-4}$		
$N_{ m path}$	128	4096		
$N_{ m sample}$	1	1		
$N_{ m shift}$	8	64		
L	7	12		
$l_{\min}$	3	4		
$l_{ m max}$	5	7		
${\chi}$	16	128		
$E_{ m eq}$	3.5%	3.7%		
$E_{\rm dyn}$	1.0%	1.9%		

## V. CONCLUSIONS

The field of MD simulation faces important challenges in harnessing massively parallel computer architectures. Although successful parallelization of the the force evaluation can be expected as the system size grows (i.e., weak scaling), there exists a much more difficult challenge of employing ever-larger numbers of parallel processors to accelerate the simulation of systems of a fixed size (i.e., strong scaling). Remarkable success has been achieved in this vein<sup>8–13</sup>, but fundamental limitations are inevitable.

The current work suggests that parallelization in the dimension of time via path integrals offers a promising avenue for future progress. We introduce the PAMD approach, which enables significant speedups over conventional Brownian dynamics algorithms via parallelization of the path-action with respect to time. Proof-of-principle applications demonstrate that the algorithm can be applied straightforwardly to the harmonic oscillator and the Lennard-Jones liquid, where speedups of up to three orders of magnitude over the conventional Euler integration scheme for Brownian dynamics are achieved. For a large class of systems including the two examples studied here, we suspect that even greater speedups are possible with the use of larger numbers of parallel pro-

cessors and enhancement of the MC path sampling efficiency.

Although promising for the systems presented here, the PAMD approach will likely require additional methodological developments to become applicable to long-timescale, large-scale simulations. Central to this effort will be the refinement of path sampling methodologies that lead to the reduction in the number of parallel processors that are needed for a given amount of speedup with the method. Regardless, we feel that the natural parallelization of path-integral formulations, combined with the increasing availability of massively parallel computer resources, should motivate increased attention to the opportunities of parallelizing molecular dynamics simulation in time.

#### **ACKNOWLEDGMENTS**

We gratefully acknowledge stimulating discussions with Matthew G. Welborn, Eric Vanden-Eijnden and Gavin E. Crooks. This work was supported in part by the Department of Energy under Award No. DE-FOA-0001912 and the Office of Naval Research under Award No. N00014-10-1-0884.

- 1. D. Frenkel and B. Smit, *Understanding molecular simulation:* From algorithms to applications (Academic Press, 2002).
- M. P. Allen and D. J. Tildesley, Computer simulation of liquids (Clarendon Press, 2017).
- J. D. Durrant and J. A. McCammon, "Molecular dynamics simulations and drug discovery," BMC Biology 9, 71 (2011).
- R. O. Dror, R. M. Dirks, J. P. Grossman, H. Xu, and D. E. Shaw, "Biomolecular simulation: A computational microscope for molecular biology," Annual Review of Biophysics 41, 429– 452 (2012).
- A. F. Voter, F. Montalenti, and T. C. Germann, "Extending the time scale in atomistic simulation of materials," Annual Review of Materials Research 32, 321–346 (2002).
- D. Keyes, "Petaflops, seriously," in High Performance Computing HiPC 2007, edited by S. Aluru, M. Parashar, R. Badrinath, and V. K. Prasanna (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007) pp. 2–3.
- J. Dongarra, P. Beckman, T. Moore, P. Aerts, G. Aloisio, J.-C. Andre, D. Barkai, J.-Y. Berthou, T. Boku, B. B, F. Cappello, B. Chapman, X. Chi, A. Choudhary, S. Dosanjh, T. Dunning, S. Fiore, A. Geist, B. Gropp, B. Harrison, M. Hereld, M. Heroux, A. Hoisie, K. Hotta, Z. Jin, Y. Ishikawa, F. Johnson, S. Kale, R. Kenway, D. Keyes, B. Kramer, J. Labarta, A. Lichnewsky, T. Lippert, B. Lucas, B. Maccabe, S. Matsuoka, P. Messina, P. Michielse, M. B., M. S. Mueller, W. E. Nagel, H. Nakashima, M. E. Papka, D. Reed, M. Sato, E. Seidel, J. Shalf, D. Skinner, M. Snir, T. Sterling, R. Stevens, F. Streitz, B. Sugar, S. Sumimoto, W. Tang, J. Taylor, R. Thakur, A. Trefethen, M. Valero, A. van der Steen, J. Vetter, P. Williams, R. Wisniewski, and K. Yelick, "The International Exascale Software Project roadmap," The International Journal of High Performance Computing Applications 25, 3–60 (2011).
- S. Plimpton, "Fast parallel algorithms for short-range molecular-dynamics," Journal of Computational Physics 117, 1–19 (1995).
- J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," Journal of Computational Chemistry 26, 1781–1802 (2005).

- 10. D. E. Shaw, R. O. Dror, J. K. Salmon, J. P. Grossman, K. M. Mackenzie, J. A. Bank, C. Young, M. M. Deneroff, B. Batson, K. J. Bowers, E. Chow, M. P. Eastwood, D. J. Ierardi, J. L. Klepeis, J. S. Kuskin, R. H. Larson, K. Lindorff-Larsen, P. Maragakis, M. A. Moraes, S. Piana, Y. Shan, and B. Towles, "Millisecond-scale molecular dynamics simulations on Anton," Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, 1–11 (2009).
- R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, and R. C. Walker, "Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald," Journal of Chemical Theory and Computation 9, 3878–3888 (2013).
- 12. S. Páll, M. J. Abraham, C. Kutzner, B. Hess, and E. Lindahl, "Tackling exascale software challenges in molecular dynamics simulations with GROMACS," in *Solving Software Challenges* for Exascale, edited by S. Markidis and E. Laure (Springer International Publishing, 2015) pp. 3–27.
- J. P. Grossman, B. Towles, B. Greskamp, and D. E. Shaw, "Filtering, reductions and synchronization in the Anton 2 network," Parallel and Distributed Processing Symposium, 860– 870 (2015).
- J.-L. Lions, Y. Maday, and G. Turinici, "A parareal discretization in time of pde's," Comptes Rendus de l'Académie des Sciences - Series I - Mathematics 332, 661–668 (2001).
- C. Farhat and M. Chandesris, "Time-decomposed parallel timeintegrators: Theory and feasibility studies for fluid, structure, and fluid-structure applications," International Journal for Numerical Methods in Engineering 58, 1397–1434 (2003).
- I. D. Garrido, B. Lee, G. E. Fladmark, and M. S. Espedal, "Convergent iterative schemes for time parallelization," Mathematics of Computation 75, 1403–1428 (2006).
- M. Emmett and M. L. Minion, "Toward an efficient parallel in time method for partial differential equations," Communications in Applied Mathematics and Computational Science 7, 105–132 (2012).
- L. Baffico, S. Bernard, Y. Maday, G. Turinici, and G. Zérah, "Parallel-in-time molecular-dynamics simulations," Physical Review E 66, 057701 (2002).
- Y. Yanan, A. Srinivasan, and N. Chandra, "Scalable timeparallelization of molecular dynamics simulations in nano mechanics," Proceedings of the 2006 International Conference on Parallel Processing, 119–126 (2006).
- R. Speck, D. Ruprecht, R. Krause, M. Emmett, M. Minion, M. Winkel, and P. Gibbon, "A massively space-time parallel N-body solver," in Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '12 (IEEE Computer Society Press, Los Alamitos, CA, USA, 2012) pp. 92:1–92:11.
- E. J. Bylaska, J. Q. Weare, and J. H. Weare, "Extending molecular simulation time scales: Parallel in time integrations for high-level quantum chemistry and complex force representations," Journal of Chemical Physics 139, 074114 (2013).
- A. L. Blumers, Z. Li, and G. E. Karniadakis, "Supervised parallel-in-time algorithm for long-time Lagrangian simulations of stochastic dynamics: Application to hydrodynamics," Journal of Computational Physics 393, 214–228 (2019).
- D. Perez, E. D. Cubuk, A. Waterland, E. Kaxiras, and A. F. Voter, "Long-time dynamics through parallel trajectory splicing," Journal of Chemical Theory and Computation 12, 18–28 (2016).
- A. Brünger, C. L. Brooks, and M. Karplus, "Stochastic boundary conditions for molecular dynamics simulations of ST2 water," Chemical Physics Letters 105, 495–500 (1984).
- A. C. Brańka and D. M. Heyes, "Algorithms for Brownian dynamics simulation," Physical Review E 58, 2611–2615 (1998).
- A. Ricci and G. Ciccotti, "Algorithms for Brownian dynamics," Molecular Physics 101, 1927–1931 (2003).
- G. Bussi and M. Parrinello, "Accurate sampling using Langevin dynamics," Physical Review E 75, 056707 (2007).

- N. Bou-Rabee, "Time integrators for molecular dynamics," Entropy 16, 138 (2014).
- H. Risken and T. Frank, The Fokker-Planck equation: Methods of solution and applications (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996).
- L. R. Pratt, "A statistical method for identifying transition states in high dimensional problems," Journal of Chemical Physics 85, 5045–5048 (1986).
- R. Olender and R. Elber, "Calculation of classical trajectories with a very large time step: Formalism and numerical examples," Journal of Chemical Physics 105, 9299–9315 (1996).
- 32. P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, "Transition path sampling: Throwing ropes over rough mountain passes, in the dark," Annual Reviews of Physical Chemistry 53, 291–318 (2002).
- T. F. Miller III and C. Predescu, "Sampling diffusive transition paths," Journal of Chemical Physics 126 (2007).
- 34. D. A. Sivak, J. D. Chodera, and G. E. Crooks, "Time step rescaling recovers continuous-time dynamical properties for discrete-time Langevin integration of nonequilibrium systems," Journal of Physical Chemistry B 118, 6466-6474 (2014).
- D. M. Ceperley, "Path integrals in the theory of condensed helium," Reviews of Modern Physics 67, 279–355 (1995).
- A. M. Stuart, J. Voss, and P. Wilberg, "Conditional path sampling of SDEs and the Langevin MCMC method," Communications in Mathematical Sciences 2, 685–697 (2004).
- G. Stoltz, "Path sampling with stochastic dynamics: Some new algorithms," Journal of Computational Physics 225, 491–508 (2007).
- 38. C. W. Gardiner, Stochastic methods: A handbook for the natural and social sciences, Springer Series in Synergetics (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009).
- 39. C. Dellago, P. G. Bolhuis, and P. L. Geissler, "Transition path sampling," in *Advances in Chemical Physics* (John Wiley & Sons, Inc., 2003) pp. 1–78.
- T. R. Gingrich and P. L. Geissler, "Preserving correlations between trajectories for efficient path sampling," Journal of Chemical Physics 142, 234104 (2015).
- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," Journal of Chemical Physics 21, 1087– 1092 (1953).
- W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," Biometrika 57, 97–109 (1970).
- 43. H. Behringer and R. Eichhorn, "Brownian dynamics simulations with hard-body interactions: Spherical particles," Journal of Chemical Physics 137, 164108 (2012).