

# Milestoning: Domain Decomposition for Molecular Dynamics

Ziheng Chen<sup>[0000-0001-9671-3977]</sup>,  
Björn Engquist<sup>[0000-0002-1940-5871]</sup>,  
Laurence Halpern<sup>[0000-0002-7877-7130]</sup>

## 1 Introduction

Molecular dynamics simulations are powerful computational tools with many important applications in physics, chemistry and biology. The evolution of atoms and molecules is typically described by a Hamiltonian system of ordinary differential equations with location and momenta of atoms or molecules as unknowns. Numerical simulations based on this formulation are computationally costly in practice. The challenges can come from complicated Hamiltonians requiring partial differential equation (PDE) computations from density function theory in each time step. The challenges can also come from the multiscale nature of many simulations. Even if the Hamiltonians are relatively simple and based on explicit empirical potentials the computational cost will be high if the individual atom vibrations need to be resolved by short time steps, but the overall time is orders of magnitude longer than in simulations involving protein folding. We will first briefly discuss existing techniques that exploit distributed computing and then focus on milestoning. This is a technique introduced by Ron Elber [9, 7] to allow for accurate simulations over longer time intervals. It is a domain decomposition strategy that aims at reducing the overall computational complexity. The domain boundaries are here called milestones and the coupling is between fluxes and sources. Our aim is to improve the theoretical foundation of milestoning by studying a stochastic model and the related Fokker-Planck equation. We will discuss the potential for application of milestoning in

---

Ziheng Chen  
Department of Mathematics, University of Texas at Austin, 2515 Speedway, Austin, TX 78712,  
USA, e-mail: stokes615@utexas.edu

Björn Engquist  
Department of Mathematics, University of Texas at Austin, 2515 Speedway, Austin, TX 78712,  
USA, e-mail: engquist@oden.utexas.edu

Laurence Halpern  
LAGA Université Sorbonne Paris-Nord, France, e-mail: halpern@math.univ-paris13.fr

solving high-dimensional PDEs and domain decomposition acceleration techniques for milestoning.

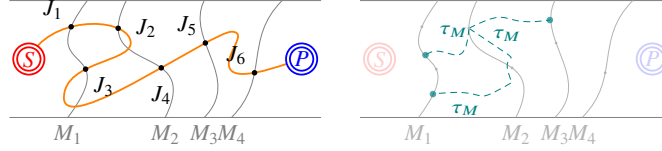
## 2 Domain decomposition for distributed computing in molecular dynamics

Before we discuss milestoning, we briefly review currently common domain decomposition (DD) of the form that just exploits distributed computing for enhancing performance. It has been used for quite some time and is mainly of three types, DD in space, DD in time and finally DD in probability. DD in space is natural and similar to spatial DD of grid based time dependent PDE solvers. Boundary zones near the domain boundaries are used for communication between the domains. This is implemented in standard software packages, as for example in GROMACS [14]. A challenge is here that the interacting forces between molecules are global and not local as is the case for differential operators. The distribution of molecules or atoms is typically far from uniform in space or phase space. The first challenge is overcome by the fact that even if the forces have global reach, in practice it is local. For example, the Lennard-Jones potential decays as distance to power six. The uneven distribution is handled by dynamic load balancing. DD in time relies on the parareal strategy for avoiding the difficulty of causality in distributing the computation between several processors, see, for example [11]. A coarse grid approximation is applied over the full time interval and then more accurate distributed independent simulations correct the coarse approximation. The challenge is that molecular dynamics (MD) is typically in the form of Hamiltonian systems of ordinary differential equations (ODEs) and thus notoriously difficult for parareal techniques [10]. Fortunately, several useful coarse grid or surrogate models have been independently developed in the MD community based on many concepts from clustering to machine learning [11]. DD in probability, which here is called the parallel replica method [17] means that many independent simulations are performed in parallel. Realistic assumption in MD will allow, for example, estimating expected transition times only based on the first arrivals.

## 3 Milestoning Algorithm and Fokker-Planck equation

The milestoning algorithm was initially introduced in [9] to compute the time scales of complex processes by establishing a set of predetermined milestones and estimating the transitional kernel between them. Once the kernel is calculated, it can be used to determine other non-equilibrium quantities such as mean first passage time (MFPT) or reaction rate. However, as highlighted in [16, 13], the vanilla milestoning algorithm relies on the assumption that the trajectories between consecutive milestone hops are statistically independent, which requires careful design of milestone geometry.

To mitigate this issue, the technique of exact milestoning (EM) is developed and studied in [2, 3, 13, 8]. In the context of estimating the MFPT  $\mathbb{E}\tau_P$  to a terminal state  $P$ , the local MFPTs  $\mathbb{E}^\mu \tau_M$  are averaged under the weight of the invariant distribution  $\mu$ , which is estimated through iterative sampling of the transition kernel between milestones [2, Theorem 3.3], as illustrated in fig. 1.



**Fig. 1** MFPT estimation via exact milestoning algorithm: breaking down long trajectories into local iterations between milestones.

The EM algorithm is further generalized in [4, Algorithm 2] to sample

$$u(z) := \mathbb{E} e^{-rT} g(X_T) + \int_0^T e^{-rt} [f(X_t) dt + h(X_t) dk_t] \quad (1)$$

along the path driven by the following stochastic process

$$dX_t = b dt + \sigma dW_t + \gamma dk_t, \quad X_0 = z, \quad T := \min_t \{X_t \in \Gamma_D\}, \quad (2)$$

where  $X_t$  reflects on  $\Gamma_N$  in the conormal direction  $\gamma := -\frac{Dn}{|Dn|}$ ,  $D := \frac{1}{2}\sigma\sigma^T$  and  $k_t$  is the local time on the boundary. A sketch of generalized EM is shown in Algorithm 1 for accurate and efficient estimation on  $u(z)$ .

**Given :** Milestones setup (Definition 2), initial guess  $\xi$ , tolerance  $\epsilon > 0$ , running/terminal/reflecting integrands  $f, g, h$ , decay rate  $r$  and point  $z \in \Omega$  to evaluate  $u$ .

**Return:**  $\hat{u}_l(z)$  as an approximation of  $u(z)$  (Eqn (1)).

*Initialize*  $\mu^{(0)} \leftarrow \xi, l \leftarrow 0$ .

**for**  $l = 0, 1, \dots$  *until stopping criterion is met* **do**

**for**  $i$  *iterating through*  $\Lambda$  **do**

        Sample  $N$  independent paths (Eqn (2)) from  $X_0^{(l,i,j)} \sim \mu_i^{(l)}$ .

        Calculate empirical average  $\hat{I}_{l,i} \leftarrow \frac{1}{N} \sum_j I_{l,i,j}$  (Eqn (1)).

**end**

    Collect empirical distribution of local path terminal points as  $\mu^{(l+1)}$ .

    Assemble the numerical estimate  $\hat{u}_l(z) \leftarrow \sum_i \mu_i^{(l)}(M_i) \hat{I}_{l,i}$ .

**end**

**Algorithm 1:** Estimating stochastic path integrals (Eqn (1)) via generalized EM.

The consistency of EM is established by formulating  $u(z)$  as the solution to the Fokker-Planck equation and by validating the corresponding domain-decomposition scheme. In fact, Eqn (1) is the time-independent form of the Feynman-Kac formula with reflexive boundary conditions.

**Theorem 1 (Time-independent Feynman-Kac formula [4, Corollary 10.13])** *The expectation of path integral Eqn (1), driven by the stochastic process Eqn (2), is given by the value at  $z$  of the solution to the stationary Fokker-Planck equation*

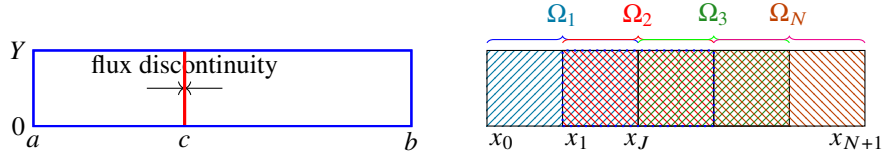
$$\mathcal{L}u + f = 0 \text{ in } \Omega, \quad u = g \text{ on } \Gamma_D, \quad \text{flux } u = |Dn| h \text{ on } \Gamma_N, \quad (3)$$

where  $\mathcal{L} := b \cdot \nabla + D : \nabla^2 - r$ ,  $D := \frac{1}{2} \sigma \sigma^T$ , under the assumption that  $u, f, h \in L^\infty$ ,  $\int_0^\infty e^{-rs} dk_s < \infty$ ,  $r > 0$  or  $r = 0$  with a bounded  $T$ .

The MFPT calculation [2] is a special case of Theorem 1 where  $r = g = h = 0$ ,  $f = 1$  and the Dirichlet boundary condition is set on the terminal milestone  $P$ .

## 4 PDE interpretation of the Milestoning algorithm

In this section we understand the Milestoning algorithm as a sequence of solutions of Laplace equations with an internal flux jump and Dirichlet boundary data. The geometry of the problem is defined by  $\Omega = (a, b) \times (0, 1)$ , with boundary  $\partial\Omega$ . The



**Fig. 2** Initial problem and domain decomposition in bands

interface  $\Gamma = \{c\} \times (0, 1)$  carries a flux discontinuity, and  $u$  is continuous across the interface, such that

$$-\Delta u = f(y)\delta_c \text{ in } \Omega = (a, b) \times (0, Y), \quad u = 0 \text{ on } \partial\Omega. \quad (4)$$

The domain  $\Omega$  is split into overlapping subdomains, using a decomposition of  $(a, b)$  with a mesh  $x_j$ ,  $x_0 = a$ ,  $x_{N+1} = b$ ,  $H_j = x_j - x_{j-1}$  for  $j = 1, \dots, N+1$ , and  $\Omega_j = (x_{j-1}, x_{j+1}) \times (0, 1)$ . We suppose that the mesh contains  $c$ :  $c = x_J$ , with  $J \geq 2$ . Figure 2 provides a graphical illustration of the decomposition setup of interest.

**Definition 1** For each subdomain  $\Omega_j$  and input  $f_j$  defined on  $\Omega_j$ , we define the local problem

$$-\Delta u_j = f_j(y)\delta_{x_j} \text{ in } \Omega_j = (x_{j-1}, x_{j+1}) \times (0, 1), \quad u_j = 0 \text{ on } \partial\Omega_j, \quad (5)$$

and extend  $u_j$  by 0 on  $\Omega \setminus \Omega_j$ . Since  $u_j$  depends linearly on  $f_j$ , the same linearity holds for its boundary derivatives. We therefore define the operators  $A_{j,j\pm 1}$  such that

$$A_{j,j-1}f_j = -d_x u_j(x_{j-1}), \quad A_{j,j+1}f_j = d_x u_j(x_{j+1}). \quad (6)$$

The closed form of  $A_{j,j\pm 1}$  is computed in the proof of the theorem 2 below, and given in Eqn (9).

**Theorem 2** *If the division is equidistant, there exists a unique  $(f_1, \dots, f_N)$  such that  $u = u_1 + \dots + u_N$ .*

*Proof.* By the distribution theory, we know that  $-\Delta u = f(y)\delta_c$  if and only if  $[d_x u](c, \cdot) = -f$ , where  $[\phi](c) = \phi(c^+) - \phi(c^-)$ . The relation between the functions  $u_j$  is equivalent to a linear system on the functions  $f_j$ , obtained by writing the jumps of derivatives on any  $\Gamma_j = \{x_j\} \times (0, 1)$  to be  $f\delta_{jJ}$ .

$$-[d_x u_1 + \dots + d_x u_N](x_j, y) = f\delta_{jJ}\delta_c, \quad j = 1, \dots, N. \quad (7)$$

The system is block linear tridiagonal with matrix  $A$ . We identify it to a tridiagonal linear system in the  $f_j$ , since each  $u_j$  is a linear function of  $f_j$ , therefore  $-[d_x u_{j-1}](x_j, \cdot) = d_x u_{j-1}(x_j, \cdot) := A_{j,j-1}f_{j-1}$  and similarly  $-[d_x u_{j+1}](x_j, \cdot) = -d_x u_{j+1}(x_j, \cdot) := A_{j,j+1}f_{j+1}$ , and  $-[d_x u_j](x_j, \cdot) = f_j$ . In order to identify the operator  $A$ , we use Fourier series in the  $y$  direction to solve Problem (Eqn (4)).  $u(x, y) = \sum_{k=1}^{+\infty} \hat{u}(x, k) \sin \omega_k y$ , with  $\omega_k = k\frac{\pi}{Y}$ . For fixed  $k$ ,  $\hat{u}(x, k)$  is solution of  $-d_{xx}\hat{u}(x, k) - \omega_k^2 \hat{u}(x, k) = 0$  in  $((a, c) \times (0, Y)) \cup ((c, b) \times (0, Y))$ , with  $[\hat{u}'_k](c) = -\hat{f}(k)$ . This is solvable in closed form:

$$\begin{cases} \hat{u}(x, k) = \alpha_k \sinh \omega_k (x - a)\chi_{(a,c)} + \beta_k \sinh \omega_k (b - x)\chi_{(c,b)}, \\ \alpha_k = \frac{f_k \sinh(\omega_k(b - c))}{\omega_k \sinh(\omega_k(b - a))}, \quad \beta_k = \frac{f_k \sinh(\omega_k(c - a))}{\omega_k \sinh(\omega_k(b - a))}, \end{cases} \quad (8)$$

where  $\chi$  is the characteristic function of the interval. This applies to the  $u_j$  as well, from which we can compute the derivatives at the endpoints of the interval and obtain the form of the linear operators  $A_{j,j\pm 1}$ ,

$$A_{j,j-1} = -\frac{\sinh(\omega_k H_{j-1})}{\sinh(\omega_k(H_{j-1} + H_j))}, \quad A_{j,j+1} = -\frac{\sinh(\omega_k H_{j+2})}{\sinh(\omega_k(H_{j+2} + H_{j+1}))}. \quad (9)$$

If the mesh is equidistant the coefficients are all equal to  $1/2 \cosh(\omega_k H)$ . For each frequency  $k$ , the matrix  $A(k)$  is strictly diagonally dominant, so invertible by Hadamard's Lemma. Hence the theorem is proved.  $\square$

We recall the Jacobi algorithm for the linear system  $AX = b$ . The usual splitting of the matrix  $A = D - L - U$  gives the Jacobi algorithm  $DX^{n+1} = (L + U)X^n + b$ . Here the diagonal is the identity, and the algorithm to approximate  $\mathbf{f} = (f_1, \dots, f_N)$  reads

$$f_j^{n+1} = \delta_{Jj} f - A_{j,j-1} f_{j-1}^n - A_{j,j+1} f_{j+1}^n.$$

which is exactly the multigrid algorithm which, written in terms of the  $u_j$  in the subdomains, is for  $j = 1, \dots, N$ ,

$$-[d_x u_j^{n+1}](x_j, :) = [d_x u_{j-1}^n](x_j, :) + [d_x u_{j+1}^n](x_j, :) + f \delta_{jJ} \delta_c. \quad (10)$$

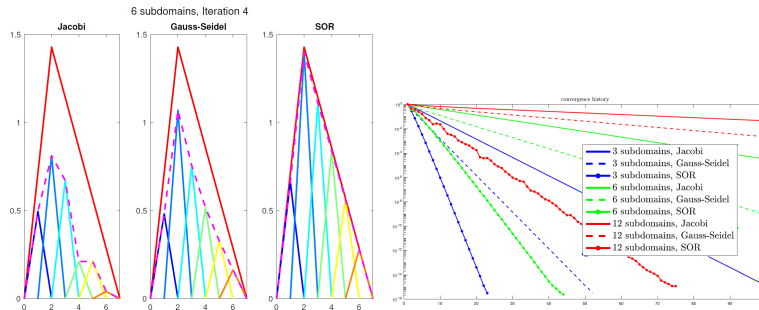
**Theorem 3** *The multigrid algorithm is a Jacobi algorithm for system (Eqn (7)). If the division is equidistant, it is convergent, with a convergence factor smaller than  $\frac{\cos \frac{\pi}{N+1}}{\cosh(\omega_1 H)}$ .*

*Proof.* Use Fourier series. If the  $x_j$  are equidistant, the extradiagonal coefficients of the matrix  $A(k)$  are all equal to  $\frac{1}{\cosh(\omega_k H)}$ , and the eigenvalues of the Jacobi matrix  $J = I - A$  are equal to  $\frac{\cos \frac{\ell \pi}{N+1}}{\cosh(\omega_k H)}$  for  $\ell = 1, \dots, N+1$ . We conclude by Parseval Theorem.  $\square$

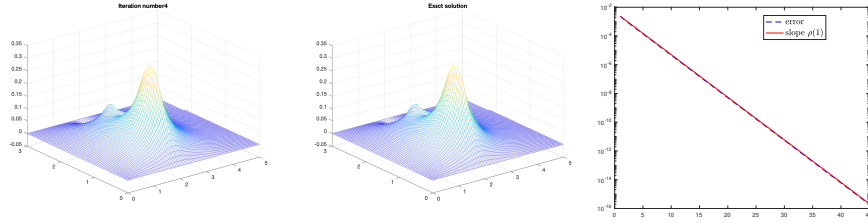
*Remark 1* In dimension 1, the solution of (Eqn (4)) is piecewise affine, equal to

$$u(x) = \frac{f}{b-a} ((b-c)(x-a)\chi_{(a,c)} + (c-a)(b-x)\chi_{(c,b)}).$$

Then for equidistant mesh, the matrix  $A$  is strongly diagonally dominant and irreducible, hence invertible. The Jacobi algorithm is similarly convergent, and the convergence factor from Theorem 3 is equal to  $\cos(\frac{\pi}{N+1})$  which is close to 1 for a large number of subdomains. The Jacobi algorithm is in general very slow. It can be accelerated by Gauss-Seidel algorithm, which converges twice as fast (for tridiagonal matrices), and can even be improved by a successive over-relaxation (SOR) algorithm, that is  $(D - \omega L)X^{n+1} = \omega b + (\omega U + (1 - \omega)D)X^n$ . This algorithm is convergent in our particular case, with an optimal parameter  $\omega^* = \frac{2}{1 + \sqrt{1 - \rho^2(J)}}$ , where  $\rho(J)$  is the spectral radius of  $J$  equal to  $\cos \pi/(N+1)$ , so  $\omega^* = \frac{2}{1 + \sin \pi/(N+1)}$ . The Gauss-Seidel algorithm corresponds to  $\omega = 1$ . These results are numerically verified as shown in figs. 3 and 4.



**Fig. 3** 1D Simulation + error: Comparison Jacobi, Gauss-Seidel and successive over-relaxation (SOR).



**Fig. 4** 2D Simulation (left and center), error as a function of the iterations (right).

## 5 Decomposition of Primal and Dual Equations

Inspired by the coupled estimation formula  $\widehat{u}_l(z) = \sum_i \mu_i^{(l)}(M_i) \widehat{I}_{l,i}$  (Algorithm 1), we decompose the solution of Eqn (3) into the distribution part  $\mu_i^{(l)}$  and the local estimate part  $\widehat{I}_{l,i}$ . As shown in the later paragraph, the former maps to the dual-flux of the local dual solutions (Eqn (15) and definition 3) while the latter solves the local primal equation (Eqn (14)). The setup for a milestoning-style decomposition of domain is described as follows (fig. 5).

**Definition 2** Given  $\Omega \subset \mathbb{R}^d$  with boundary  $\partial\Omega$  partitioned into  $\Gamma_N$  and  $\Gamma_D$ , we consider the set of milestones  $\{M_i\}_{i \in \Lambda}$  and the associated compartments  $\{Y_i\}_{i \in \Lambda}$  under the index set  $\Lambda$  ( $0 \notin \Lambda$ ) satisfying:

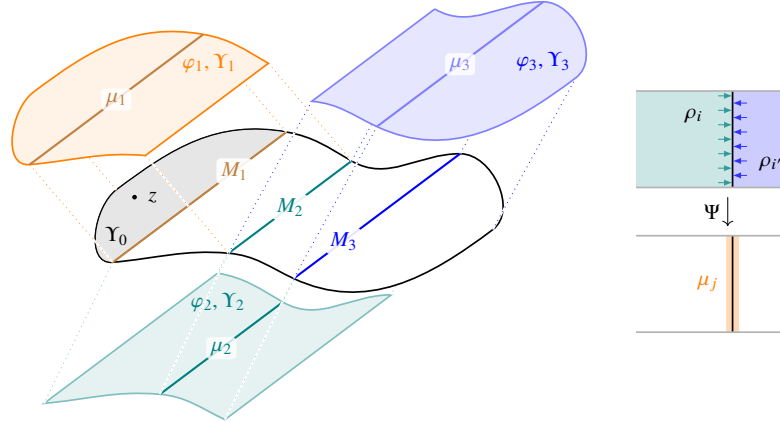
1.  $M_i$  is the image of a  $C^1$ -mapping from a closed set in  $\mathbb{R}^{d-1}$ .
2.  $Y_i$  is an open subset of  $\Omega$  that contains  $M_i$  and  $\{Y_i\}$  covers  $\Omega$
3. The boundary follows a decomposition of  $\Gamma_{M,i} := \sqcup_{j \in \mathcal{N}_i} M_j$ ,  $\Gamma_{D,i} := \Gamma_{M,i} \sqcup \widetilde{\Gamma}_{D,i}$  where  $\Gamma_{N,i} \subset \Gamma_N$ ,  $\widetilde{\Gamma}_{D,i} \subset \Gamma_D$  and  $\partial Y_i = (\sqcup_{j \in \mathcal{N}_i} M_j) \sqcup \Gamma_{N,i} \sqcup \widetilde{\Gamma}_{D,i}$  with its inherited normal direction  $n^{j,i}$ .
4.  $M_i$  has a positive distance to  $\Gamma_{D,i}$ , i.e. there exists  $\delta > 0$ , independent of  $i$ , s.t.

$$\text{dist}(M_i, \Gamma_{D,i}) \geq \delta, \forall i \in \Lambda. \quad (11)$$

For the simplicity of notations, we introduce a virtual milestone  $M_0 = \{z\}$  and define the extended index set  $\widetilde{\Lambda} := \{0\} \cup \Lambda$ . In addition, we require that  $0 \notin \cup_{j \in \Lambda} \mathcal{N}_j$ , i.e. we do not consider the hitting event on  $M_0$ .

**Definition 3** The global/local and primal/dual equations are listed in the 2-by-2 table (table 1). The probability flow operator is defined as  $\mathcal{J}\rho := \nabla \cdot (D\rho) - b\rho$ , followed by the dual of  $\mathcal{L}$  as  $\mathcal{L}^*\rho := (\nabla \cdot \mathcal{J} - r)\rho$ . The boundary flux operator and its dual counterpart is defined as flux  $u := (\nabla u \otimes n) : D$  and flux\*  $\rho := (n \cdot \mathcal{J})\rho$ . For each subdomain  $Y_i$ , the local versions of flux operators use the corresponding normal direction.

Notice that the primal and dual equations shall be understood in the weak sense. The well-posedness of solutions  $u, \rho, \varphi_i, \phi_i$  is established in [4, Proposition 10.2].



**Fig. 5** Left: Milestoning setup (Definition 2). Right: Flux mapping  $\Psi$  (Eqn (16)).

	Primal	Dual
Global	$\mathcal{L}u + f = 0 \quad (12)$ $u _{\mathcal{D}} = g$ $\text{flux } u _{\mathcal{N}} = h$	$\mathcal{L}^* \rho = \delta \quad (13)$ $\rho _{\mathcal{D}} = 0$ $\text{flux}^* \rho _{\mathcal{N}} = 0$
Local	$\mathcal{L}\varphi_i + f = 0 \quad (14)$ $\varphi_i _{\tilde{\Gamma}_{\mathcal{D},i}} = g$ $\varphi_i _{\Gamma_{\mathcal{M},i}} = 0$ $\text{flux}_i \varphi_i _{\Gamma_{\mathcal{N},i}} = h$	$\mathcal{L}^* \phi_i = \delta \quad (15)$ $\phi_i _{\Gamma_{\mathcal{D},i}} = 0$ $\text{flux}_i^* \phi_i _{\Gamma_{\mathcal{N},i}} = 0$

**Table 1** Primal and Dual equations, listed in global and local version.

In the limit of convergence of the milestoning algorithm, there is a balance between the inward and outward probability flux on each milestone. The inward flux can be described by aggregating the dual-flux of the Green's function  $\rho_{i;z}$  on a given milestone, while the outward flux is distributed to neighboring milestones via the local-dual equation (Eqn (15)). Thus, the collection of Green's functions serve as a bridge between the distribution on the milestone and the balancing condition. A rigorous definition is demonstrated as follows.

**Definition 4** The space of particle distributions is defined as the tensor product  $V := L^1(\bigcup_{i \in \tilde{\Lambda}} M_i) = \otimes_{i \in \tilde{\Lambda}} L^1(M_i)$ , endowed with the  $L^1$  norm  $\|\mu\|_1 := \sum_{i \in \tilde{\Lambda}} \|\mu_i\|_{L^1(M_i)}$  for  $\mu := (\mu_i)_{i \in \tilde{\Lambda}} \in V$ . The space of Green's functions  $\rho_z := (\rho_{i;z})_{i \in \tilde{\Lambda}}$  is given as the tensor product of  $\mathcal{D}'(\Upsilon_i)$ . The flux mapping  $\Psi_z : W \rightarrow V$  is defined as

$$\Psi_z(\rho_z) := \left( \sum_{i \in \mathcal{N}_j^* \neq 0} \text{flux}^* \rho_{i;z} + \mathbf{1}_{0 \in \mathcal{N}_j^*} \text{flux}^* \phi_{0;z} \right)_{j \in \tilde{\Lambda}} \quad (16)$$

where  $\mathcal{N}_j^* := \{i : j \in \mathcal{N}_i\}$  is the co-neighbor set. The balancing equation is thus defined as follows.

$$\begin{aligned} \mathcal{L}^* \rho_{i;z}(x) &= \delta(x \in M_i) [\Psi_z \rho_z(x)]_i \text{ in } Y_i, \\ \rho_{i;z} &= 0 \text{ on } \Gamma_{D,i}, \quad \text{flux}^* \rho_{i;z} = 0 \text{ on } \Gamma_{N,i} \end{aligned} \quad (17)$$

Given  $\rho_z$  solving the balancing equation (Eqn (17)),  $\mu_z = \Psi_z(\rho_z)$  satisfies

$$\rho_{i;z}(x) = \int_{M_i} \mu_{i;z}(y) \phi_{i;y}(x) \, d\sigma(y). \quad (18)$$

This can be understood as a decomposition of sampled paths within a milestoning sub-region: under the equilibrium, the collection of paths is a superposition of those from individual points along the milestone weighted by  $\mu_z$  [4, Equation 10.31].

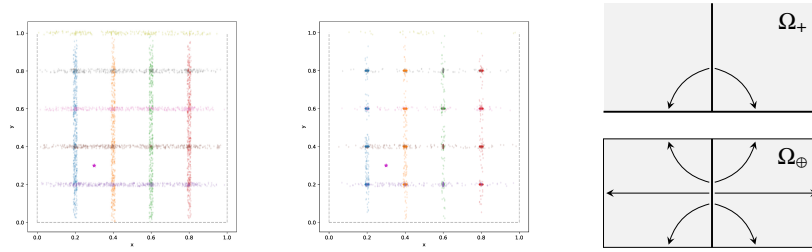
Now we move onto the decomposition of the global equations.

**Theorem 4 ([4, Proposition 10.11])** *Under mild conditions, if  $\rho_z$  solves the balancing equation (Eqn (17)), then  $u(z) = \sum_{i \in \Lambda} \int_{M_i} \varphi_i(y) \mu_{i;z}(y) \, d\sigma(y)$  solves the primal equation (Eqn (12)) and  $\rho = \sum_i \rho_i$  solves the dual equation (Eqn (13)).*

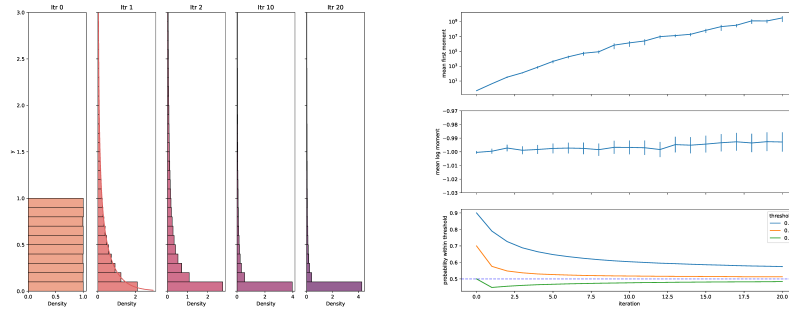
## 6 Crystallization

Through numerical experiments, we observe the so-called crystallization effect where the sampled distributions on the milestone tend to concentrate over iterations (Figure 6). Intuitively, neighboring milestones pull away a significant portion of probabilities due to close adjacency to the source milestone. This imbalance develops over time and eventually leads to crystallization. In this section, we analyze the crystallization effect in the context of isotropic diffusion models and give a rationale on the requirement of positive distance between milestones and Dirichlet boundaries (Definition 2, Item 4). We study this effect under two domain setups: a half-plane setup  $\Omega_+ = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}$  and a finite-domain setup  $\Omega_\oplus = \{(x_1, x_2) \in [-1, 1] \times [0, 1]\}$ . In both cases, the source milestone is defined as the vertical line segment in the middle of the domain, as illustrated in Figure 6.

Let  $\mu(x)$  denote the distribution of sampled points along the milestone and  $\mathcal{K}$  be the iterative operator mapping distributions over milestones. Without loss of generality, we assume the initial distributions are the same on all milestones, so the intermediate iteration distributions shall remain the same on every milestone. Motivated by Eqn (18),  $\mathcal{K}$  possesses a convolutional form  $\mathcal{K}\mu(x) = \int_I \mu(y) p(x; y) dy$  where the kernel  $p$  characterizes the probabilistic transitional structure between milestones.



**Fig. 6** Development of crystallization effect under a 1-by-2 domain setup. Left and middle: snapshot of distribution in iteration at initial state and after the second iteration. The purple star location serves as the  $z$  point where the solution  $u$  is evaluated. Right: Crystallization domain setup (upper: half-plane; lower: finite-domain).



**Fig. 7** Half-crystallization for half-plane setup  $\Omega_+$ . Left: distribution in iterations. Right: first and logarithmic moments in expectation and probability below given thresholds; the error bars stand for a standard deviation among all test results.

- For half-plane setup,  $I = \mathbb{R}^+$  and  $p(y; x) = \frac{1}{\pi y} \cdot \frac{1}{(x/y)^2 + 1}$ .
- For finite-domain setup,  $I = [0, \frac{1}{2}]$  with  $p$  derived from the normal derivative of the associated Green's function [4, Definition 10.21].

We now state the singularity of the iteration sequence  $\{\mathcal{K}^k \mu\}$  in both setups.

**Theorem 5 ([4, Remark 10.4, Proposition 10.37])** *For the half-plane setup,  $\lim_{k \rightarrow \infty} \int_0^t \mathcal{K}^k \mu(x) dx = \frac{1}{2}$  for any given  $t > 0$ . For the finite domain setup, the integral series  $\int_0^{1/2} \log x \mathcal{K}^k \mu(x) dx$  monotonically diverges to  $-\infty$  as  $k \rightarrow \infty$ .*

The proof of Theorem 5 exploits the fact that log modulus of a 2-D Brownian motion is a martingale, thus the sampled intersections under the half-plane setup is a random walk under the log scale (referred to as the “half-crystallization” phenomenon, fig. 7). On the other hand, the log moment decreases (in negative scale) under the finite domain setup due to replication of singularities. It immediately

implies that  $\{\mathcal{K}^k \mu\}$  is not uniformly bounded in either setup, thus posing numerical difficulties.

## 7 Discussion

The primary contribution of this work is a domain decomposition based interpretation of the exact-milestoning algorithm. We discuss the connection to the original proposed method in Section 3 and describe the setup in Definition 2. Theorem 1 serves as a key component that bridges the stochastic path integral and the Fokker-Planck problem. Then, we translate the milestoning algorithm as a domain decomposition method (Theorem 4) and derives the balancing equation in Eqn (17).

As a side target, we also study the so-called crystallization phenomenon which may occur for improper milestone setups. We confirm the half crystallization effect for a half-plane setup while the logarithmic moment blows up for finite domains Theorem 5. In either case, the singularities in the distributions on the milestones pose a great difficulty for numerical methods to work efficiently.

In the context of computational chemistry, the family of milestoning algorithms is popular due to its easy setup and capability of accurate solving wide range of problems, in comparison with other approaches such as minimal energy path method [5], finite temperature string method [6], transition interface sampling [15], and forward flux sampling [1]. We also plan to extend the range of problems that can be solved by exact milestoning by investigating the non-linear Feynman-Kac formula.

It is a prevalent idea to develop coupled systems that integrate macroscopic and microscopic dynamics. For instance, the lattice kinetic Monte Carlo method [12] implements a jump process at the microscopic (lattice) level, while the overall effect is modeled as a diffusion/drift process using density functional theory. However, it is important to note that the coupling mechanism in the milestoning algorithm functions differently; specifically, a class of continuous processes is simulated at the microscopic level to provide transitional probabilities for the macroscopic jump process. We aim to delve deeper to identify potential connections with other numerical methods used in molecular dynamics.

## References

1. Allen, R.J., Warren, P.B., ten Wolde, P.R.: Sampling rare switching events in biochemical networks. *Phys. Rev. Lett.* **94**(1), 018104 (2005)
2. Aristoff, D., Bello-Rivas, J.M., Elber, R.: A mathematical framework for exact milestoning. *MMS* **14**(1), 301–322 (2016)
3. Bello-Rivas, J.M., Elber, R.: Exact milestoning. *J. Chem. Phys.* **142**(9), 1–19 (2015)
4. Chen, Z.: Numerical multiscale methods: from homogenization to milestoning. Ph.D. thesis, University of Texas at Austin (2024)
5. E, W., Ren, W., Vanden-Eijnden, E.: String method for the study of rare events. *Phys. Rev. B* **66**, 052301 (2002)

6. E, W., Ren, W., Vanden-Eijnden, E.: Finite temperature string method for the study of rare events. *J. Phys. Chem. B* **109**(14), 6688–6693 (2005)
7. Elber, R.: A milestoning study of the kinetics of an allosteric transition: Atomically detailed simulations of deoxy Scapharca hemoglobin. *Biophys. J.* **92**(9), L85–L87 (2007)
8. Elber, R., Fathizadeh, A., Ma, P., Wang, H.: Modeling molecular kinetics with Milestoning. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **11**(4), 1–26 (2021)
9. Faradjian, A.K., Elber, R.: Computing time scales from reaction coordinates by Milestoning. *J. Chem. Phys.* **120**(23), 10880–10889 (2004)
10. Gander, M.J., Hairer, E.: Analysis for parareal algorithms applied to Hamiltonian differential equations. *J. Comput. Appl. Math.* **259**, 2–13 (2014)
11. Gorynina, O., Legoll, F., Lelievre, T., Perez, D.: Combining machine-learned and empirical force fields with the parareal algorithm: application to the diffusion of atomistic defects. *arXiv preprint arXiv:2212.10508* (2022)
12. Hoffmann, M.J., Bligaard, T.: A Lattice Kinetic Monte Carlo Solver for First-Principles Microkinetic Trend Studies. *J. Chem. Theory Comput.* **14**(3), 1583–1593 (2018)
13. Lin, L., Lu, J., Vanden-Eijnden, E.: A mathematical theory of optimal milestoning (with a detour via exact milestoning). *Commun. Pure Appl. Math.* **71**(6), 1149–1177 (2018)
14. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J.C.: Gromacs: Fast, flexible, and free. *J. Comput. Chem.* **26**(16), 1701–1718 (2005)
15. Van Erp, T.S., Moroni, D., Bolhuis, P.G.: A novel path sampling method for the calculation of rate constants. *J. Chem. Phys.* **118**(17), 7762–7774 (2003)
16. Vanden-Eijnden, E., Venturoli, M., Ciccotti, G., Elber, R.: On the assumptions underlying milestoning. *J. Chem. Phys.* **129**(17), 1–13 (2008)
17. Voter, A.F.: Parallel replica method for dynamics of infrequent events. *Phys. Rev. B* **57**(22), R13985 (1998)