**Weighted ensemble: Recent mathematical developments**

D. Aristoff, J. Copperman, G. Simpson, R. J. Webber, and D. M. Zuckerman

1) Mathematics, Colorado State University
2) Biomedical Engineering, Oregon Health Sciences University
3) Mathematics, Drexel University
4) Computing & Mathematical Sciences, California Institute of Technology

The weighted ensemble (WE) method, an enhanced sampling approach based on periodically replicating and pruning trajectories in a set of parallel simulations, has grown increasingly popular for computational biochemistry problems, due in part to improved hardware and the availability of modern software. Algorithmic and analytical improvements have also played an important role, and progress has accelerated in recent years. Here, we discuss and elaborate on the WE method from a mathematical perspective, highlighting recent results which have begun to yield greater computational efficiency. Notable among these innovations are variance reduction approaches that optimize trajectory management for systems of arbitrary dimensionality.

I. INTRODUCTION

Weighted ensemble (WE) is an enhanced sampling method employing multiple trajectories of a stochastic dynamics to estimate mean first passage times (MFPTs) and related statistics. WE can be applied to any stochastic dynamics model, such as Langevin dynamics in a molecular system or continuous-time jump dynamics in a reaction network. In biomolecular systems, WE has enabled estimation of MFPTs that are orders of magnitude larger than the combined lengths of the individual WE trajectories. WE has also recently been used to elucidate the spike opening dynamics in the SARS CoV-2 virus.

WE is based on splitting and merging, as shown in Figure 1. During splitting, “favorable” or “interesting” trajectories, according to a user definition, are replicated. During merging, the “less favorable” or “redundant” trajectories are randomly eliminated from the ensemble. Trajectories are then re-weighted to preserve the statistics of the path ensemble.

The first splitting and merging algorithm was reported in the 1950s and attributed to von Neumann. In 1996, Hu and Kin modified von Neumann’s original approach by grouping the trajectories into bins and applying splitting and merging in each of the bins separately, thus preserving the bin weights. It was recently shown that WE with fixed bin weights leads to convergent estimates, in contrast, von Neumann’s original method can be unstable in the large time limit.

Attention to WE has grown, and the method is now implemented in the widely used WESTPA software package and the more recent wepy package. A Julia language package, used here, is also available. Meanwhile, a growing community of researchers is promoting the method and developing it in different directions.

In modern biochemical applications, WE is often used to estimate the MFPT for a stochastic dynamics to transition from a source state $A$ to a target state $B$. In addition, WE has been used to estimate other transition path statistics, including the distribution of reaction times from $A$ to $B$, the distribution of entry points into $B$, and the characteristics of paths leading from $A$ to $B$. Here, we focus on the estimation of MFPTs, which is a significant and challenging application because the inverse of the MFPT is the reaction rate constant.

Other computational techniques for estimating MFPTs include Markov state models, forward flux sampling, exact milestoning, non-equilibrium umbrella sampling, and transition interface sampling. Like WE, these approaches use large numbers of short, unbiased trajectories to compute the MFPT. However, the following combination of features makes WE especially attractive:

- WE only requires simulation of the stochastic model forward in time, never backward. Thus, WE can be applied to a wide variety of stochastic models arising in chemistry and in other areas, e.g., astronomy, climate science, and systems biology.
- WE is fully parallelizable over time intervals of length $\tau$ between splitting/merging events. Since the trajectories are simulated for the same time interval, the parallelization can be managed simply and efficiently.
- WE provides asymptotically unbiased, convergent estimates in the limit of many time steps for any choice of parameter.

![An illustration of splitting and merging of weighted trajectories. Weights are represented by the widths of trajectories. Two splitting and merging steps are shown, in which the top trajectory is split and the bottom trajectories are merged. Trajectory weights are combined with merging and divided with splitting.](image-url)
• WE has been observed to provide better MFPT estimates than estimates from Markov state models\cite{23} or non-equilibrium umbrella sampling\cite{24}.

Despite progress over the last three decades, theoretical questions about WE have remained unanswered until recently. WE performance is highly dependent on the choice of parameters\cite{23}, including the definition of the bins and the desired number of trajectories in each bin. For a long time it was unclear what the optimal parameter choices would be.

Recent work in the mathematical literature sheds new light on the optimal parameter choices for WE\cite{25,26,27,28}. The optimal merging coordinate is the local MFPT to \( B \) given the current state. The optimal splitting coordinate is the local variance of the MFPT to \( B \), given the current state and the time interval \( \tau \). There is a theoretical limit on the variance reduction achievable through WE, and bins based on the optimal merging and splitting coordinates ensure the optimal variance in the limit of many trajectories and many time steps.

In this article, we aim to communicate these recent mathematical advances in a brief, accessible way. We identify the optimal splitting and merging coordinates in one- and two-dimensional examples. We propose optimized bins based on these coordinates, which differ from the more traditional bins based on the root mean square displacement (RMSD) to \( B \). Our examples show that binning based on the RMSD can give catastrophically wrong results. In such cases, more effective bins are needed, and optimization based on reaction coordinates is helpful.

The rest of the article is organized as follows. Section II discusses the computation of mean first passage times, which is a major problem addressed by WE; Section III introduces the WE method; Section IV identifies the optimal merging and splitting coordinates for WE; Section V recommends variance reduction strategies for WE; and Section VI concludes.

### II. MEAN FIRST PASSAGE TIMES AND THE HILL RELATION

We will study the MFPT of a Markovian stochastic dynamics \( X_t \) from a source state \( A \) to a sink state \( B \). In the biochemistry context, \( X_t \) could represent stochastic molecular dynamics such as Langevin dynamics or constant-energy dynamics generated with a stochastic thermostat. The MFPT could correspond to the characteristic time for folding, binding, or conformational change of a simulated protein.

To define the MFPT precisely, we must specify the distribution of starting points \( \rho_A \) within the the source state \( A \). Different choices of \( \rho_A \) will lead to different MFPTs, even if \( A \) remains the same. The distribution \( \rho_A \) fully determines the MFPT, which is defined as the averaged length of the trajectories initiated from the source distribution \( \rho_A \) and absorbed upon reaching \( B \).

When computing the MFPT, we “recycle” \( X_t \) according to the source distribution \( \rho_A \) upon arrival at the sink state \( B \). This means that the location immediately changes from \( B \) to \( A \), but the time index \( t \) continues to increase as usual. We assume that the distribution of trajectories for this source-sink recycling dynamics converges as \( t \to \infty \) to a unique steady-state distribution, denoted by \( \pi \).

Under these recycling boundary condition, the Hill relation\cite{29} expresses the MFPT as the inverse of the steady-state flux \( J \) of \( X_t \) into \( B \). To state this more precisely, write \( T_B \) for the first passage time to \( B \), and \( N_t \) for the number of arrivals of \( X_t \) in \( B \) by time \( t \). The Hill relation then states:

\[
\langle T_B \rangle_A = \text{MFPT of } X_t \text{ from } A \text{ to } B \\
= 1/J = (\text{steady-state flux into } B)^{-1} \\
= \left( \frac{d}{dt} \langle N_t \rangle_{\pi} \right)^{-1}.
\]

Here and elsewhere, we use subscripts to indicate the starting distribution for the dynamics \( X_t \). Thus, \( \langle T_B \rangle_A \) indicates the mean first passage time when \( X_0 \) is started from \( \rho_A \), and \( \langle N_t \rangle_{\pi} \) indicates the mean number of arrivals when \( X_0 \) is started from \( \pi \). Equation (2) holds for any \( t \geq 0 \), since \( \pi \) is the steady state.

The application of the Hill relation transforms the computation of a long expected time (the MFPT) into a computation of a small rate (the steady state flux into \( B \)). This transformation makes it possible, in principle, to compute the MFPT using trajectories with lengths much shorter than the MFPT. Moreover, these short trajectories can be run in parallel to reduce the wall-clock time. Yet, a large number of trajectories are required since the flux can be very small.

Fortunately the Hill relation can be combined with enhanced sampling approaches like WE to help accurately estimate the small flux into \( B \). It is necessary that the trajectories approximate the steady-state distribution \( \pi \), but this can frequently be attained, because the timescale for relaxation to steady state can be much shorter than the MFPT in problems of interest\cite{28}.

Figure 2 shows a simple system in which the steady-state flux from \( A \) into \( B \) is small. Like many of the examples to follow, the data in the figure comes from overdamped Langevin

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_t )</td>
<td>underlying Markovian dynamics</td>
</tr>
<tr>
<td>( \beta )</td>
<td>inverse thermal energy ((1/k_B T))</td>
</tr>
<tr>
<td>( U )</td>
<td>potential energy</td>
</tr>
<tr>
<td>( \tau )</td>
<td>evolution time interval or lag time</td>
</tr>
<tr>
<td>( A )</td>
<td>initial (source) set</td>
</tr>
<tr>
<td>( B )</td>
<td>target (sink) set</td>
</tr>
<tr>
<td>( \rho_A )</td>
<td>initial (source) distribution inside ( A )</td>
</tr>
<tr>
<td>( T_B )</td>
<td>first passage time to ( B )</td>
</tr>
<tr>
<td>( N_t )</td>
<td>number of arrivals in ( B ) by time ( t )</td>
</tr>
<tr>
<td>( q )</td>
<td>committer function from ( A ) to ( B )</td>
</tr>
<tr>
<td>( J )</td>
<td>steady-state flux into ( B )</td>
</tr>
<tr>
<td>( \pi )</td>
<td>steady state of source-sink dynamics</td>
</tr>
<tr>
<td>( \langle \cdot \rangle )</td>
<td>mean or average for source-sink dynamics</td>
</tr>
<tr>
<td>( h )</td>
<td>(flux) discrepancy function</td>
</tr>
<tr>
<td>( v^2 )</td>
<td>(flux) variance function</td>
</tr>
<tr>
<td>( J_t )</td>
<td>WE estimate of steady-state flux up to time ( t )</td>
</tr>
<tr>
<td>( \pi, h, v^2 )</td>
<td>( \tau \to 0 ) limits of ( \pi, h, v^2 )</td>
</tr>
</tbody>
</table>
III. WEIGHTED ENSEMBLE: SPLITTING AND MERGING

Throughout this article we consider a simple version of WE in which the splitting and merging of trajectories are reformulated as resampling. For a more general discussion of WE, see the review. The simple WE algorithm is described below and illustrated in Figure 3.

1. **Initialization.** Select starting points for \( N \) trajectories. Assign a weight to each trajectory so that the weights sum to one.

2. **Resampling.** Partition the \( N \) trajectories into bins, and select the desired number of copies for each bin, called the allocation. The bins are simply groupings of the trajectories, and they may change in time. The allocations are positive integers which may change with time but are assumed to always sum to \( N \).

Next, resample trajectories within each bin with probabilities proportional to the weights, and divide the total weight in each bin evenly among the resampled trajectories.

3. **Dynamics.** Evolve the trajectories independently for time \( \tau \) according to the underlying dynamics. The difference between discrete- and continuous-time flux into \( B \) will be small if \( \tau \) is small or the trajectories in \( B \) are very slow to escape. We reserve variables \( t \) and \( \tau \) for the “physical time”, i.e., the time index for the original dynamics \( X_i \).

4. **Convergence.** Repeat steps 2-3 as long as desired or possible. Estimate the steady-state flux \( J \) using the average WE probability flux entering \( B \) from the burn-in time \( t_0 \) to the final time \( t \). That is,

\[
\hat{J}_t = \frac{\text{Total WE weight entering } B}{t - t_0},
\]

where \( t - t_0 \) is the physical time of simulations following the burn-in. The estimate for \( J \) in turn yields a MFPT estimate via the Hill relation.

The crucial step in WE is the resampling step, which leads to the splitting and merging of trajectories. Splitting is the dynamics of replicating trajectories that are “valuable,” while merging is the act of combining trajectories that are “similar.”
Splitting improves the accuracy of the flux estimates, especially when the duplicated trajectories are expected to make large contributions to the future flux. With repeated splitting, merging is needed to limit the number of trajectories.

Although splitting and merging are traditionally applied separately, in the simple WE algorithm above these steps are combined via resampling. In each bin, trajectories are resampled with probabilities proportional to their weights, with the resampled trajectories each receiving an equal proportion of the bin weight. Resampling is easier to formulate and implement than the traditional splitting and merging. Moreover, resampling is optimal when trajectories within a given bin are, except for their weights, indistinguishable.

Splitting and merging must be balanced, and a main contribution of our recent work has been identifying the regions of state space in which splitting or merging should be promoted. Identifying such regions naturally leads to variance reduction strategies. We discuss these developments in Sections IV and V.

Given sufficient resources, the WE method terminates when the estimated flux converges to a near-constant value. However, if steady state cannot be reached during the available simulation time, enhanced methods for WE initialization are needed. For example, incorporating an appropriate burn-in time can reduce transient relaxation effects and adjusting the initial weights can also improve convergence.

WE can be used to estimate other transition path statistics in addition to the steady-state flux. For instance, the WE trajectories – increases local sampling and is of high benefit in regions that are important for the flux computation. We now introduce the scalar reaction coordinate that describes optimal merging function as a normalized difference between the local flux variance for the source-sink dynamics started at various values is appropriate.

Using the Hill relation, the flux discrepancy function can be rewritten in terms of MFPTs initiated from different starting distributions. As shown in Appendix 2, where we recall $T_B$ is the first passage time into $B$ excluding time $t = 0$. Equation (9) allows us to reinterpret the optimal merging function as a normalized difference between the local MFPT to $B$ starting from $x$ and the MFPT for trajectories initiated from the steady state $\pi$. Identity (8) holds independently of both the dimension and the type of Markovian dynamics. It shows that the flux discrepancy function and the local MFPT are equivalent optimal coordinates for merging.

### Optimal splitting

Splitting – the duplication of certain trajectories – increases local sampling and is of high benefit in regions that are important for the flux computation. We now introduce the scalar reaction coordinate that describes optimal splitting behavior, the flux variance function;

$$v(x)^2 = \tau^{-1} \text{Var}_\tau [h_0(X_\tau)], \quad (9)$$

where $h_0$ is a version of the discrepancy function that counts flux at time $t = 0$, i.e., $h_0(x) = \mathbb{E}_0 T_B(x) + h(x)$, and $\text{Var}_\tau$ denotes variance for the source-sink dynamics started at $x$. The flux variance function quantifies the variation in the expected flux over a time interval $\tau$ for trajectories initiated at $x$.

As a heuristic strategy, a favorable allocation should satisfy:

$$\# \text{ trajectories near } x \propto \pi(x) v(x). \quad (10)$$

Allocating in this way minimizes the future contribution to the flux variance, assuming that WE is in steady state. We refer to this as the optimal allocation distribution and provide an intuitive derivation in Appendix 3.

As a more rigorous mathematical result, the following establishes the optimality of $h$ and $v$ as WE reaction coordinates.

**Theorem** If the recycled dynamics $X_t$ is geometrically ergodic, then the WE method satisfies the following:

1. For any choice of bins and allocations, the WE flux estimate $\hat{J}_t$ given in (6) converges with probability one to the inverse MFPT,

$$\lim_{t \to \infty} \hat{J}_t = \frac{1}{\langle T_B \rangle}. \quad (11)$$

2. For any $\varepsilon > 0$ and any choice of bins and allocations, the WE variance satisfies

$$\text{Var} \hat{J}_t \geq \frac{1 - \varepsilon}{N_t} \left( \int v(x) \pi(x) \, dx \right)^2 \quad (12)$$

when $t$ is sufficiently large.
3. For any $\varepsilon > 0$, if the bins are sufficiently small rectangles in $h$ and $v$ coordinates and if the allocation for each bin is proportional to $\int_{a}^{h(x)} v(x) \pi(x) \, dx$, the WE variance satisfies
\[
\text{Var}\,\tilde{h} \leq \frac{1 + \varepsilon}{N t} \left( \int v(x) \pi(x) \, dx \right)^2
\]  
for sufficiently large $t$ and $N$. □

The theorem holds regardless of dimension, temperature, time $\tau$, and type of Markovian dynamics. It highlights the following WE strategy that is optimal in the limit of large $N$ and $\tau$: trajectories with similar $h$ and $v$ values are merged, while splitting enforces $\text{Const} \times \pi(x) v(x)$ trajectories near $x$. The theorem supports the interpretation of $\pi(x) v(x)$ as the optimal distribution for trajectory allocation.

B. Maximum gain of WE over direct dynamics

The above mathematical theory enables a quantitative comparison between WE and direct Monte Carlo sampling of first-passage events. Here, we define direct Monte Carlo as independent trajectories without any splitting or merging.

Assuming that trajectories are started from the steady state $\pi$, the direct Monte Carlo variance with $N$ trajectories can be written
\[
\frac{1}{Nt} \int v(x)^2 \pi(x) \, dx.
\]  
This formula reveals that the flux variance function $v^2$ is intrinsic to direct Monte Carlo dynamics as well as WE dynamics, since $v^2$ quantifies the variance in the flux estimates as trajectories are evolved forward over a time interval $\tau$. Dividing $\eqref{eq:flux_variance}$ by $\eqref{eq:flux_variance_wes}$, the ratio of direct Monte Carlo variance to optimal WE variance is then
\[
\text{Optimal gain over direct Monte Carlo} = \frac{\int v^2(x) \pi(x) \, dx}{\left( \int v(x) \pi(x) \, dx \right)^2}.
\]
This ratio quantifies the maximal possible variance reduction achievable by WE in the large time limit, for any number of trajectories.

The optimal gain over direct Monte Carlo has important implications. Analytical and numerical investigation of this quantity can yield insight into how much benefit is possible from using WE. Additionally, knowing the optimal gain over direct Monte Carlo as a reference enables quantitative comparisons of different WE binning strategies against the theoretically optimal performance.

C. Optimal WE for one-dimensional Smoluchowski dynamics

To make the preceding theory more explicit, we consider Smoluchowski (overdamped Langevin) dynamics in a one-dimensional setting with a source state $A = \{x \leq a\}$ and a sink state $B = \{x \geq b\}$, where $a < b$. The recycling distribution is a delta function at the edge of $A$, i.e., $\rho_A = \delta(x-a)$. We consider the continuous-time limit, $\tau \to 0$, which leads to simpler, more interpretable mathematical expressions.

In the limit $\tau \to 0$, the dynamics $X_t$ is observed at all times and recycling occurs immediately upon entry into $B$. The steady-state distribution $\tilde{\pi}$ can be calculated using the relation
\[
\tilde{\pi}(x) \propto [1 - \tilde{q}(x)] e^{\beta U(x)}.
\]  
Here, $\tilde{q}$ is the committor function, the probability for $X_t$ to reach $B$ before $A$ starting from $x$, which is given by
\[
\tilde{q}(x) = \begin{cases} 0 & x \leq a, \\ \int_a^x e^{\beta U(x)} / D(y) \, dy & x < a < b, \\ 1 & x \geq b. \end{cases}
\]  
Using $\eqref{eq:steady_state}$ and $\eqref{eq:committor}$, the steady-state distribution is supported on $\{x < b\}$, and it satisfies
\[
\tilde{\pi}(x) \propto e^{-\beta U(x)} \int_{\max\{a,b\}}^b e^{\beta U(y)} D(y) \, dy.
\]  
Compared to the Boltzmann density $\propto e^{-\beta U(x)}$, the steady-state distribution $\tilde{\pi}$ removes mass from the regions near $B$ due to the recycling of trajectories. See Figure 4 for an illustration.

The flux discrepancy function $h = h_\tau$ and the flux variance function $v = v_\tau$ depend implicitly on the evolution interval $\tau$, and they converge to well-defined limits as $\tau \to 0$:
\[
h(x) = \lim_{\tau \to 0} h_\tau(x), \quad v^2(x) = \lim_{\tau \to 0} v_\tau(x)^2.
\]  
As a result of $\eqref{eq:flux_variance}$, the discrepancy function $\tilde{h}$ satisfies
\[
\tilde{h}(x) = \frac{(\langle T_B \rangle)_{\pi} - (\langle T_B \rangle)_A}{\langle T_B \rangle_A}.
\]  
FIG. 4. The steady state of the source-sink dynamics. Plotted is the steady state $\tilde{\pi}$ of the Smoluchowski dynamics with recycling from $B$ to $A$, a diffusion coefficient $D = 1$, and $\beta = 4$. In this and subsequent figures, $U(x) = 5(x - 1)^2 x^2 + 2x + 0.5x^2 - 0.2x$. The recycling distribution is a delta function at the edge of $A$, i.e., $\rho_A = \delta(x-a)$. We consider the continuous-time limit, $\tau \to 0$, which leads to simpler, more interpretable mathematical expressions.

In the limit $\tau \to 0$, the dynamics $X_t$ is observed at all times and recycling occurs immediately upon entry into $B$. The steady-state distribution $\tilde{\pi}$ can be calculated using the relation
\[
\tilde{\pi}(x) \propto [1 - \tilde{q}(x)] e^{\beta U(x)}.
\]  
Here, $\tilde{q}$ is the committor function, the probability for $X_t$ to reach $B$ before $A$ starting from $x$, which is given by
\[
\tilde{q}(x) = \begin{cases} 0 & x \leq a, \\ \int_a^x e^{\beta U(x)} / D(y) \, dy & x < a < b, \\ 1 & x \geq b. \end{cases}
\]  
Using $\eqref{eq:steady_state}$ and $\eqref{eq:committor}$, the steady-state distribution is supported on $\{x < b\}$, and it satisfies
\[
\tilde{\pi}(x) \propto e^{-\beta U(x)} \int_{\max\{a,b\}}^b e^{\beta U(y)} D(y) \, dy.
\]  
Compared to the Boltzmann density $\propto e^{-\beta U(x)}$, the steady-state distribution $\tilde{\pi}$ removes mass from the regions near $B$ due to the recycling of trajectories. See Figure 4 for an illustration.

The flux discrepancy function $h = h_\tau$ and the flux variance function $v = v_\tau$ depend implicitly on the evolution interval $\tau$, and they converge to well-defined limits as $\tau \to 0$:
\[
h(x) = \lim_{\tau \to 0} h_\tau(x), \quad v^2(x) = \lim_{\tau \to 0} v_\tau(x)^2.
\]  
As a result of $\eqref{eq:flux_variance}$, the discrepancy function $\tilde{h}$ satisfies
\[
\tilde{h}(x) = \frac{(\langle T_B \rangle)_{\pi} - (\langle T_B \rangle)_A}{\langle T_B \rangle_A}.
\]
The ideal reaction coordinates for WE. Plotted are the discrepancy function $\tilde{h}$ the committor $\tilde{q}$ and the variance function $\tilde{v}$ for the Smoluchowski dynamics with recycling from $B$ to $A$, a diffusion coefficient $D = 1$, and $\beta = 10$.

Here, $\tilde{T}_B$ is the exact MFPT function to the target state $B$, which is given for $x \leq b$ by:

$$\langle \tilde{T}_B \rangle_x = \int_x^b \frac{e^{\beta U(z)}}{D(z)} \int_z^{\infty} e^{-\beta U(y)} dy dz. \quad (21)$$

The variance function $\tilde{v}$ satisfies the relations

$$\tilde{v}^2(x) = \lim_{\tau \to 0} \frac{\text{Var}_x[\tilde{h}(X_t)]}{\tau} = 2D(x) \left| \frac{d}{dx} \tilde{h}(x) \right|^2, \quad (22)$$

where the first equality follows from the definition (4) and the second equality comes from applying Itô’s lemma. Formula (22) suggests that there should be more splitting in regions of high $\tilde{h}$ variability, or, equivalently, high variability in the local MFPT $\langle \tilde{T}_B \rangle_x$.

Note that the committor $\tilde{q}$ is usually understood to be the relevant coordinate for computing the MFPT and similar problems. We have shown, however, that two different scalar coordinates, namely $\tilde{h}$ and $\tilde{v}^2$, are the relevant ones for splitting and merging in the context of WE simulation. The difference between these coordinates is exhibited by a model problem with an asymmetric two-barrier system in Figure 5. For this problem, the committor $\tilde{q}$ would not be an ideal reaction coordinate since it poorly resolves the largest barrier on the forward path from $A$ to $B$. In contrast, $\tilde{h}$ and $\tilde{v}$ resolve the largest forward barrier, making them appropriate for WE.

Lastly, a simple formula for the optimal WE allocation holds in the low-temperature limit as $\beta \to \infty$. Assume that the largest energy increase on the forward path to $B$ occurs over the interval $[x_-, x_+]$, i.e.,

$$\Delta U = U(x_+) - U(x_-) = \max_{a \leq y \leq b, x < y} (U(y) - U(x)). \quad (23)$$

Note that the interval $[x_-, x_+]$ can include multiple energy barriers. Then, as $\beta \to \infty$, an application of Laplace’s method (see Appendix 4) yields

$$\frac{\mathcal{P}(x) \tilde{v}(x)}{\int \mathcal{P}(y) \tilde{v}(y) dy} \rightarrow \begin{cases} 0 & x \leq x_-, \\ \frac{D^{-1/2}(x)}{\int_{x_-}^x D^{-1/2}(y) dy} & x_+ < x < x_+, \\ 0 & x \geq x_+. \end{cases} \quad (24)$$

Thus, the optimal allocation is proportional to $1/\sqrt{D}$ over the interval $[x_-, x_+]$, and it is vanishingly small elsewhere as illustrated in Figure 6.

In the low-temperature limit $\beta \to \infty$, the optimal gain over direct Monte Carlo sampling grows exponentially with the size of the largest energy increase (see Appendix 4):

$$\begin{aligned} \text{Gain over direct Monte Carlo} &= \beta \to \infty \quad \frac{\pi / \beta}{\left( \int_{x_-}^{x_+} \sqrt{D(x) / D(x)} dx \right)^2 \sqrt{U'(x_-)U'(x_+)}}. \quad (25) \end{aligned}$$

This provides a formal explanation of earlier numerical finding that WE’s advantage over direct Monte Carlo grows dramatically as $\beta \Delta U$ increases.

V. A PRINCIPLED WE STRATEGY

The theoretical results above can serve as guidelines for optimizing WE. For example, we describe one principled WE strategy called “MFPT binning” in Section V A and we compare MFPT binning to a more traditional WE binning strategy in Section V B.
A. MFPT binning

From the theory in Section IV A, merging has a low cost among trajectories with similar \( h \) values. This motivates the following MFPT binning strategy in which bins are comprised of similar \( h \) values or equivalently based on \( v \), similar values of the local MFPT to \( B \):

1. Define the bins to be intervals in the \( h \) coordinate. If there are \( K \) bins, choose the endpoints \( h_0 < h_1 < \ldots < h_K \) such that
   \[
   \int_{h_k \leq x < h_{k+1}} \pi(x)v(x)dx = \text{Const.} \tag{26}
   \]

2. Allocate trajectories uniformly among these \( K \) bins, so that approximately \( N/K \) trajectories are assigned to each bin.

This approach agrees with strategy (10) while also satisfying the traditional WE rule of keeping the number of trajectories per bin nearly constant. See Figure 7 for an illustration.

The MFPT binning strategy is similar to the one described in Section IV A, which attains the lowest possible variance in the limit of many trajectories and time steps. However, instead of bins that are rectangles in \( h \) and \( v \) coordinates, we define bins as intervals of \( h \) values alone. This alteration still performs well in our model problem in Section IV B even with a small number of trajectories (100–1000) and just a few bins (5–10).

The MFPT binning strategy requires an approximation to the flux discrepancy function \( h \) and flux variance function \( v \). Strategies for estimating the functions will be explored in future work. We note that even crude estimates of \( h \) and \( v \) may be helpful, compared to naive or ad-hoc binning strategies, and WE remains unbiased regardless of the approximation quality of \( h \) and \( v \). As a concrete strategy, initial simulations can be performed, and discrete Markov state models (MSMs) can be used to approximate \( h \) and \( v \) based on the relationships to the MFPT described in Section IV.

In contrast to this MFPT binning strategy, the most common WE approach involves bins defined using root mean square displacement (RMSD) to \( B \) in some feature space, with the same number of trajectories allocated to each bin. This approach would be optimal if the dynamics in the RMSD coordinate were well-described by the one-dimensional Smoluchowski equation with large \( \beta \) and constant \( D \); see Section IV C and Figure 5. However, the RMSD coordinate rarely behaves in this way for complex systems, and our results below show that RMSD binning can catastrophically fail when the RMSD coordinate is very different from the MFPT.

B. Numerical comparison

In this section, we numerically compare two strategies for binning and allocation within WE:

1. The MFPT binning strategy discussed in Section V A

2. Traditional WE bins based on RMSD to \( B \) with allocations proportional to \( \int_\text{bin} v(x)\pi(x)dx \).

In both strategies, the allocations enforce the heuristic \( h \) of having \( \leq \pi(x)v(x) \) trajectories near \( x \), which is derived mathematically in Section IV.

We will apply these WE strategies to the two-dimensional Smoluchowski model system pictured in Figure 8, with full specifications in Appendix 5. We begin our study of this model by constructing a fine-grained MSM approximation of functions \( \pi, h, \) and \( v \). Next, we use these approximations to construct the MFPT bins in Figure 8. Observe that the MFPT bins resolve the largest energy barriers on the forward path from \( A \) to \( B \), whereas the RMSD bins have concentric circle bin boundaries that do not account for the energy landscape at all. See Appendix 5 for further bin construction details. These computations were performed using 104 independent runs to produce these variance constants.

To evaluate the precision – and hence, efficiency – of the flux estimates from MFPT and RMSD binning, we define a compute-cost-independent variance constant:

\[
\text{Variance Constant} = Nt \times \text{Var} \left( \frac{J}{Nt} \right). \tag{27}
\]

Here, the variance of the empirical WE flux has been scaled to account for total simulation cost \( Nt \). The variance constants for MFPT and RMSD binning are compared in Figure 9. The figure also shows estimates of the optimal WE variance constant \( \left( \int \pi v \right)^2 \) from the fine-grained MSM model (see equation (15)), as well as the variance constant for direct Monte Carlo, without any splitting or merging. We used \( N = 10^4 \) trajectories and averaged over \( 10^2 \) independent runs to produce these variance constants.

Figure 9 reveals a dramatic difference between MFPT and RMSD binning. RMSD binning leads to worse-quality results than direct Monte Carlo simulation for this non-trivial
FIG. 8. Different binning strategies in a challenging two-dimensional model. Pictured are the potential \( U \) (top), MFPT bins (middle), and RMSD bins (bottom), with the source \( A \) depicted as a filled dot and the sink \( B \) depicted as a dashed rectangle. Different shades indicate different bins. The model is fully specified in Appendix 5: see the potential (50).

model, for any number of bins. Meanwhile, MFPT binning leads to a major improvement over direct Monte Carlo simulation, particularly when the number of bins is large (20–40 bins). With only 5 bins, we see a modest gain with the MFPT binning strategy. Increasing the number of bins with MFPT binning systematically improves the output such that with 40 bins, we are within an order of magnitude of the estimated optimal constant. This is two and a half orders of magnitude of improvement over direct Monte Carlo.

To encourage a fair comparison, we have applied the optimal allocation rule (10) to both MFPT and RMSD binning. However, analogous tests (not pictured) show that RMSD binning performs even worse with uniform allocations, leading to higher variance constants. Our MFPT binning method here also outperformed related strategies in our earlier work (41).

Lastly, the data in Figure 9 was generated using \( N = 10^4 \) trajectories, which may exceed the computational resources available for problems of interest. Figure 10 presents comparisons for \( 10^2, 10^3, \) and \( 10^4 \) trajectories, indicating that the variance constant typically has no more than an order of magnitude in variability. In particular, for the MFPT bins the results for \( 10^3 \) trajectories are nearly indistinguishable from those obtained with \( 10^4 \) trajectories.
VI. CONCLUSION

This work has attempted to assemble, expand upon, and make accessible to a chemical physics audience, recent mathematical results of fundamental importance to weighted ensemble (WE) simulation. The mathematical theory exposes the fundamental capabilities of WE and guiding principles behind WE optimization. It leads to the identification of optimal reaction coordinates for merging and splitting. The theory also identifies the minimal possible variance for WE in the limit of many trajectories and time steps, and we demonstrate a simple binning strategy for approaching that optimum that works well in a two-dimensional model problem. Testing these approaches in complex systems will be pursued in future work.

To illustrate the mathematical theory in a physically intuitive way, explicit formulas were derived for the optimal coordinates and optimal WE variance in the case of the Smoluchowski dynamics (5) in one dimension. These formulas, presented for the first time here, demonstrate that the variance reduction from WE can be exponentially large in the size of the largest forward energy barrier from A to B.

The mathematical theory is important because it enables principled variance reduction strategies for WE. Here, we have proposed and tested one such variance reduction strategy, called MFPT binning. After pilot runs that estimate $h$ and $v$ (or equivalently, the local MFPT to B and its variance), this strategy defines the bins to be intervals in the $h$ coordinate, chosen in a way that optimizes the variance. In a model problem, MFPT binning significantly outperforms a naive WE approach that constructs bins using the RMSD to B, and these results remain robust even for relatively small N.

ACKNOWLEDGEMENTS

D. Aristoff and G. Simpson gratefully acknowledge support from the National Science Foundation via the awards DMS 2111277 and DMS 1818726. J. Copperman is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRQ-09-20). R. J. Webber was supported by the Office of Naval Research through BRC award N00014-18-1-2363 and the National Science Foundation through FRG award 1952777, under the aegis of Joel A. Tropp. D. M. Zuckerman was supported by NIH grant GM115805. Computational resources were provided by Drexel’s University 673 Research Computing Facility. The authors are grateful to Mats Johnson for preliminary numerical work related to the results in Section V.

APPENDIX

1. Computing other statistics using WE

WE can be used to compute other statistics, in addition to the MFPT from A to B. Namely, any integral $\int f(x) \pi(x) dx$ involving an observable $f$ can be estimated using $N$ WE trajectories. In the long time limit, the trajectories satisfy

$$\int f(x) \pi(x) dx = \lim_{t \to \infty} \frac{1}{t} \sum_{i=1}^{N} w_i f(\xi_{i,t}^\dagger), \quad (28)$$

where $(\xi_{i,t}^\dagger, w_i)$ denotes the position and weight of trajectory $i$ at time $t$. Therefore, as a practical estimator, we can terminate the simulations at a finite time $t$ and approximate

$$\int f(x) \pi(x) dx \approx \frac{1}{t - t_0} \sum_{s=t_0}^{t} \sum_{i=1}^{N} w_i f(\xi_{i,s}^\dagger), \quad (29)$$

where $t_0 > 0$ is a suitable burn-in time. Here we have assumed $f$ is recorded only at $t$ intervals, although WE permits saving points more frequently.

The estimator (29) is readily extended to history-dependent observables, such as transition-path times or mechanistic pathways from A to B. This requires defining the trajectories as paths starting from A, which are recycled upon reaching B, and defining $\pi$ as a stationary measure on path space. The recent mathematical literature provides a set of error bounds and convergence guarantees for the estimator (29).
2. Relationship between flux discrepancy and MFPT

To establish the relationship

$$\lim_{t \to \infty} \left[ \langle N_t \rangle_x - \langle N_t \rangle_\pi \right] = \frac{\langle T_B \rangle_x - \langle T_B \rangle_\pi}{\langle T_B \rangle_A},$$

(30)

we first introduce the following terminology:

- $T_B$ is the first passage time into $B$ for a dynamics started at $X_0 = x$.
- $T_B^s$ is the first passage time into $B$ for a dynamics started at $X_0 \sim \pi$, i.e., started at a point chosen randomly from the steady-state distribution.
- $N_t^A$ is the number of arrivals into $B$ by time $t$ for a dynamics started at $X_0 \sim \rho_A$.

Next we recall the renewal theorem from classical probability which states that for any $s > 0$,

$$\lim_{t \to \infty} \left[ \langle N_t^A \rangle - \langle N_t^A \rangle_s \right] = \frac{s}{\langle T_B \rangle_A},$$

(31)

Lastly, we verify

$$\lim_{t \to \infty} \left[ \langle N_t \rangle_x - \langle N_t \rangle_\pi \right]$$

$$= \lim_{t \to \infty} \left[ \langle N_t \rangle_x - \langle N_t \rangle_{T_B^s} \right] - \lim_{t \to \infty} \left[ \langle N_t \rangle_{T_B^s} - \langle N_t \rangle_{T_B} \right]$$

$$= \lim_{t \to \infty} \left[ \langle N_t \rangle_{T_B^s} - \langle N_t \rangle_{T_B} \right] - \lim_{t \to \infty} \left[ \langle N_t \rangle_x - \langle N_t \rangle_{T_B^s} \right]$$

$$= \frac{\langle T_B^s \rangle - \langle T_B \rangle}{\langle T_B \rangle_A},$$

(35)

where we have applied the renewal theorem after conditioning on $s = T_B^s$ and $s = T_B$.

3. Derivation of optimal allocation rule

In this section we provide an intuitive derivation of the optimal allocation rule

$$\# \text{ trajectories near } x \propto \pi(x) \nu(x).$$

(36)

We introduce the following notation:

- $\text{weight}_{\text{bin}}$ and $\text{alloc}_{\text{bin}}$ indicate a bin weight and bin allocation.
- $\langle \rangle_{\text{bin}}$ and $\text{Var}_{\text{bin}}$ denote the mean and variance with respect to the weighted trajectory distribution in a bin, so $\text{Var}_{\text{bin}}(f) = \langle f^2 \rangle_{\text{bin}} - \langle f \rangle_{\text{bin}}^2$.
- Finally, $\langle \rangle$ denotes the average over the WE ensemble up to time $t$.

Then the following variance formula holds in the limit as $t \to \infty$:

$$r \text{Var}(\hat{J}_t) \sim \sum_{\text{bin}} \frac{\text{weight}_{\text{bin}}^2}{\text{alloc}_{\text{bin}}} \left[ \text{Var}_{\text{bin}}(h) + \text{Var}_{\text{bin}}(v) + \langle v \rangle_{\text{bin}}^2 \right].$$

(37)

We consider a greedy minimization strategy for minimizing the variance formula. The variance terms $\text{Var}_{\text{bin}}(h)$ and $\text{Var}_{\text{bin}}(v)$ are minimized by choosing bins in which $h$ and $v$ values are nearly constant. The other term, namely,

$$\frac{\text{weight}_{\text{bin}}^2}{\text{alloc}_{\text{bin}}} \langle v \rangle_{\text{bin}}^2,$$

is minimized when the allocation is chosen using

$$\text{alloc}_{\text{bin}} \propto \text{weight}_{\text{bin}} \times \langle v \rangle_{\text{bin}},$$

(38)

which reduces to (36) in the limit of many trajectories.

4. Low temperature limit

Here we derive the optimal splitting and merging coordinates in the low-temperature limit $\beta \to \infty$. First, we combine the expression for the steady state

$$\tilde{\pi}(x) \propto e^{-\beta U(x)} \int_{\max\{x,a\}}^{b} \frac{e^{\beta U(y)} dy}{D(y)}$$

(39)

with the expression for the flux variance function

$$\bar{\nu}(x) = \frac{\sqrt{2} e^{\beta U(x)}}{D(x)^{1/2} \langle T_B \rangle_A} \int_{-\infty}^{x} \frac{dy}{e^{\beta U(y)}},$$

(40)

to obtain the optimal allocation rule

$$\bar{\pi}(x) \bar{\nu}(x) \propto \frac{1}{D(x)^{1/2}} \int_{\max\{x,a\}}^{b} \frac{e^{\beta U(y)} dy}{D(y)} \int_{-\infty}^{x} \frac{dy}{e^{\beta U(y)}},$$

(41)

Next, we assume the pair of points $(x_-, x_+)$ is the unique solution to the maximization

$$\max_{a \leq y < b, x < y} (U(y) - U(x)),$$

(42)

and $U''(x_+) < 0 < U''(x_-)$. Using Laplace’s method, we obtain, for any $x_- < x < x_+$,

$$\frac{1}{D(x)^{1/2}} \int_{\max\{x,a\}}^{b} \frac{e^{\beta U(y)} dy}{D(y)} \int_{-\infty}^{x} \frac{dy}{e^{\beta U(y)}} \sim \frac{C}{D(x)^{1/2}} \beta e^{\beta U(x_+)},$$

(43)

where the prefactor $C$ is explicitly

$$C = \frac{2\pi}{D(x_+) U''(x_-)^{1/2} |U''(x_+)|^{1/2}}.$$

(44)

For $x$ outside of the interval $[x_-, x_+]$, Laplace’s method dictates that the optimal allocation $\bar{\pi}(x) \bar{\nu}(x)$ is exponentially smaller as $\beta \to \infty$, whence we recover the optimal allocation.
From the above calculations, we find that the optimal gain from using WE takes the form
\[
\text{Gain over direct Monte Carlo} = \frac{QR}{S^2},
\]
where
\[
Q = \int_{-\infty}^{b} e^{B(U(x))} \frac{D(x)}{D(y)} \left( \int_{-\infty}^{\infty} e^{B(U(y))} dy \right)^2 dx,
\]
\[
R = \int_{-\infty}^{b} e^{-B(U(x))} \frac{D(x)}{D(y)} dx,
\]
\[
S = \int_{-\infty}^{b} \frac{1}{D(x)^{1/2}} \left( \int_{-\infty}^{\infty} e^{B(U(y))} dy \right) dx.
\]
Applying the Laplace principle to (46)–(48) verifies the expression
\[
\text{Gain over direct Monte Carlo} \sim \frac{\pi/\beta}{\left( \int_{x_{\min}}^{x_{\max}} \sqrt{D(x+)} / D(x) dx \right)^2} \sqrt{\left[ U'(x_{\min}) U'(x_{\max}) \right]}.
\]

5. Details of numerical simulations

In the 2D problem, the landscape is given by the expression
\[
U(x, y) = U_1(x, y) + U_2(x, y) + 0.5U_3(x, y),
\]
where
\[
\begin{align*}
\log U_1(x, y) &= -c_1 (x - 0.25)^2 - c_1 (y - 0.75)^2 - 2c_2 (x - 0.25) (y - 0.75) \\
\log U_2(x, y) &= -c_3 x^2 (1 - x)^2 y^2 (1 - y)^2 \\
\log U_3(x, y) &= -c_4 x^2 - c_4 y^2 + 2c_5 x y
\end{align*}
\]
and the constants are given by
\[
(c_1, c_2, c_3, c_4, c_5) = (50.5, 49.5, 10^5, 51, 49).
\]
The diffusivity is \( D = 1 \), and the inverse temperature is \( \beta = 30 \).

To simulate the Smoluchowski dynamics, we apply Euler-Maruyama integration with an integration step of size 0.001. Each evolution step is composed of ten such steps, so that \( \tau = 0.01 \). We constrain the trajectories to reside in \([0, 1]^2\).

The functions \( \pi, h, v \) are approximated using an MSM model generated from “microbins” \([413] \) with Voronoi cell centers that are spaced \( \Delta x = \Delta y = 0.02 \) units apart on a regular Cartesian grid. This spacing results in a total of \( 49 \times 49 = 2401 \) microbins, which are the small rectangles visible in Figure 8.


42T. L. Hill, Free energy transduction and biochemical cycle kinetics (Courier Corporation, 2005).


