Exact non-Markovian permeability from rare event simulations

An Ghysels ⦿,[1] Sander Roet ⦿,[2] Samaneh Davoudi,[1] and Titus S. van Erp ⦿[2,3]

[1]*IBiTech - bioMMeda, Faculty of Engineering and Architecture, Ghent University, 9000 Gent, Belgium*
[2]*Department of Chemistry, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*
[3]*Center for Molecular Modeling (CMM), Ghent University, Technologiepark, 9052 Zwijnaarde, Belgium*

Permeation of compounds through membranes is important in biological and engineering processes, e.g., drug delivery through lipid bilayers, anesthetics, or chemical reactor design. Simulations at the atomic scale can provide insight in the diffusive pathways and they give estimates of the membrane permeability based on counting membrane transitions or on the inhomogeneous solubility-diffusivity model described by the Smoluchowski equation. For many permeants, permeation through a membrane is too slow to gather sufficient statistics with conventional molecular dynamics simulations, i.e., permeation is a rare event. Recent attempts to improve the description of the dynamics of such rare permeation events have been based on milestoning, which allows the study of processes at timescales beyond those achievable by straightforward molecular dynamics. The approach is not relying on an overdamped description, but, still, it uses a Markovian approximation which is only valid for small permeants that are not disruptive to the membrane structure. To overcome this fundamental limitation, we show here how replica exchange transition interface sampling (RETIS) can effectively be used on this problem by deriving an effective set of equations that relate the outcome of RETIS simulations and the permeability coefficient. In addition, we introduce two new path Monte Carlo (MC) moves specifically for permeation dynamics, that are used in combination with the ordinary path generating moves, which considerably increase the efficiency. The advantage of our method is that it gives exact results, identical to brute force molecular dynamics, but orders of magnitude faster.

I. INTRODUCTION

Permeation of compounds through another medium is essential in both biological and engineering processes. In biology, the permeation of molecules, nutrients or nanoparticles through membranes is an integral part of a functioning cell, and understanding the drug delivery process can aid the design of new cancer drugs or anesthetics [1–4]. In chemical engineering, the transport of molecules can play a role in the selectivity of the molecules [5,6]. In highly complex chemical systems, it is valuable to unravel the various diffusion pathways, which can aid the optimization of a chemical reactor setup. Despite the existence of spectroscopic methods, such as fluorescence spectroscopy [7,8] or EPR experiments [9], the insight in diffusive transport at the molecular scale is difficult to obtain experimentally. Especially in inhomogeneous media where the diffusion of permeants is a function of the location, experiments usually only provide a global effective diffusion constant, without discerning local differences at the molecular scale. It is in this respect that molecular dynamics (MD) simulations can play a major role. MD creates molecular

trajectories and transport properties are directly observable at the molecular scale, such as the permeability P.

The permeability of a membrane is the flux through the membrane as a response to a concentration gradient over the membrane. A first standard approach to derive the permeability P from MD simulations is the counting method, which is based on measuring the rate of membrane transitions per unit of time and area [10–12]. Another standard approach is Bayesian analysis (BA) using the Smoluchowski equation, which assumes a position-dependent concentration profile as well as a position-dependent diffusion profile across the membrane [12–17]. The Smoluchowski equation is a pure diffusion model, where memory effects are not modeled. The serious limitation of these approaches is that, when the permeability is low, the statistics provided by MD might be insufficient, even when the trajectories are extended up to 1 microsecond of simulation time.

A more recent approach is the extraction of P from milestoning. Milestoning is based on the sampling of many short trajectories released from equilibrium distributions at hypersurfaces (milestones) [18]. Typically, these hypersurfaces are defined as subsets of configuration space with fixed values of the reaction coordinate (RC). The milestoning method then counts from each originating hypersurface how often the left or right hypersurface is hit first and the time that it takes to let this happen. These two properties for the different milestones are then combined such that the dynamics of the system can be described by a Markovian hopping sequence

from one surface to the other. Rates, mean-first-passage times and other relevant dynamic and thermodynamic data can then be obtained. The central assumption underlying milestoning is that the set of first hitting points, of the trajectories originating from one surface with a hypersurface left or right from it, is again distributed according to the equilibrium distribution. This leads to conflicting requirements. On the one hand, the milestones need to be set closely as this guarantees the highest efficiency boost compared to brute force MD. On the other hand, they need to be well separated for the central assumption to be sufficiently satisfied. In addition, for the assumption to hold, the choice of the RC is crucial and should ideally coincide with the committor [19]. The committor is generally an unknown coordinate that not only describes the position of the permeant relative to the center of mass of the membrane, but also should include nontrivial membrane deformations. Obtaining approximate forms of the committor is generally an immense task [20].

As a way to overcome the limitation of the Markovian assumption, this work will make use of the transition interface sampling (TIS) framework and of its extension, the replica exchange TIS (RETIS) formalism [21–23]. (RE)TIS gives a completely non-Markovian treatment of the interface hopping probabilities, while still being orders of magnitude faster than brute force MD. The overall rate is obtained in a divide-and-conquer mindset, by constructing a series of conditional probabilities to reach the next interface. Here we will show how the permeability, as defined in steady-state nonequilibrium conditions, can be transformed into functional form that depends on the equilibrium path ensemble properties which can be obtained from RETIS (instead of the typical canonical, NVT, or isobaric, NPT, phase space ensembles).

In an extension, we also show how the rates of transition can be defined in a meaningful way for an infinite system or a system with periodic boundaries such that it can be linked to the permeability. When the phase neighboring the membrane is unbounded, particles have a probability to indefinitely diffuse in the opposite direction of the membrane rather than through the membrane. Also when a system has periodic boundary conditions, it is difficult to detect whether particles reach the other side of the membrane through the periodic boundary or through the membrane itself. The permeability formula will be adapted to treat those cases. In addition, two additional moves will be introduced in the MC sampling of the interface ensembles. They will improve the efficiency when the simulation box contains multiple permeant molecules or when the membrane is symmetric with respect to the membrane center.

This article is organized as follows. In Sec. II, we review the existing approaches for obtaining P from MD simulations. In Sec. III, the (RE)TIS formalism is revised and the path ensembles are defined. In Sec. IV, the theoretical derivation of the permeability from (RE)TIS is presented. In Sec. V we derive the needed adaptation to treat systems with periodic boundary conditions. In Sec. VI, the two new MC moves are presented. In Sec. VII, we illustrate the accuracy and effectiveness of our approach by computing the permeability of a few basic model test systems. We end with concluding remarks in Sec. VIII.

II. EXISTING APPROACHES FOR PERMEABILITY CALCULATIONS

A. Direct counting

The permeability is defined as the ratio of the net flux J of particles transiting a membrane in the steady-state regime when imposing a small concentration difference Δc over the membrane,

$$P = \frac{J}{\Delta c}. \tag{1}$$

To compute the permeability, one could consider imposing the concentration difference at the membrane boundaries, as suggested by the definition in Eq. (1). However, at the molecular scale it is troublesome to impose a nonequilibrium steady-state concentration gradient, especially when periodic boundary conditions are applied.

A more common approach is therefore to count transitions in both directions through the membrane in an equilibrium simulation [10]. At first sight, Eq. (1) is no longer adequate for the computation of P. Indeed, note that J in Eq. (1) is the result of both crossings from left to right and from right to left, corresponding to a positive flux J^+ and a negative flux J^-. These fluxes are proportional to the concentrations at the left and right hand side of the membrane, respectively, and bear opposite signs. The net flux is, hence, the result that is left after a partial cancellation of J^+ and J^-, and in an equilibrium situation, where $\Delta c = 0$, the cancellation is complete and the net flux is zero. Whereas the permeability in this case is no longer measurable experimentally, in simulations it is still possible since it is relatively easy to trace the J^+ and J^- fluxes individually from the MD trajectories and one can write

$$P = \frac{|J^+|}{c(-h/2)} = \frac{|J^-|}{c(h/2)} = \frac{|J^+| + |J^-|}{2c_{\text{ref}}}. \tag{2}$$

Here, $z = 0$ is considered to be the center of the membrane of thickness h with borders at $z = \pm h/2$, and $c(z)$ is the concentration profile across the membrane. The reference concentration is the concentration outside the membrane $c_{\text{ref}} = c(-h/2) = c(h/2)$ and corresponds to the concentration of permeants in the bulk liquid.

In an equilibrium run, the sum of fluxes $|J^+| + |J^-|$ is measured by counting full transitions from $-h/2$ to $h/2$ and vice versa. This approach has the advantage that it is nearly model-free. However, it is in practice a challenge to measure the flux at the atomic scale for all but the fastest permeation events [9,10,16,24–33]

B. Smoluchowski equation

A second common approach is to run equilibrium MD simulations and to analyze these assuming the validity of the inhomogeneous solution-diffusivity (ISD) model, where transport is modeled by position-dependent Brownian diffusion (diffusion profile $D(z)$) on a free energy landscape (profile $F(z)$), as governed by the Smoluchowski equation [15]. The free energy is related to the permeant concentration in equilibrium through the Boltzmann probability, $c(z) \sim \exp(-\beta F)$, where $\beta = 1/(k_B T)$ is the inverse temperature, k_B is Boltzmann constant. Given the two profiles F and D, the

permeability follows directly from solving the Smoluchowski equation for its steady-state solution in the presence of a fixed concentration difference Δc over the membrane [13,16],

$$\frac{1}{P} = e^{-\beta F_{\text{ref}}} \int_{-h/2}^{h/2} \frac{1}{e^{-\beta F(z)} D(z)} dz, \qquad (3)$$

where F_{ref} is the free energy at the reference location, usually at $z = -\frac{h}{2}$.

The structure of Eq. (3) shows that the free energy profile is the dominant contribution to the permeability. Several cases can now be thought of for $F(z)$: $F(z)$ has a barrier, is flat, or has a well. The last two cases are rather academic, as realistic membranes often form a combination of free energy barriers and wells, e.g., for O_2 permeation through phospholipid bilayers [16]. First, when the permeation implies crossing a high free energy barrier, the integration range in Eq. (3) can readily be reduced from $[-h/2, h/2]$ to $[-h'/2, h''/2]$ with $h', h'' < h$,

$$\frac{1}{P} \approx e^{-\beta F_{\text{ref}}} \int_{-h'/2}^{h''/2} \frac{1}{e^{-\beta F(z)} D(z)} dz \qquad (4)$$

if the integrand can be neglected for the outer regions $-h/2 < z < -h'/2$ and $h''/2 < z < h/2$. Given the exponential dependence on the free energy barrier, the integration boundaries $-h'/2$ and $h''/2$ can often be chosen rather close to the maximum of the free energy barrier as long as F_{ref} is taken in the bulk region. In the direct counting method, the neglect of the outer regions in the integration relates to the fact that nearly all transitions from $-h/2$ to $h/2$ contain one and only one single $-h'/2$ to $h''/2$ transition.

Second, the permeability in a homogeneous medium where $F(z)$ is a flat profile, equals $P = D/h$ with D the diffusion constant. The permeability halves when doubling the thickness h. This shows that the integration boundaries and reference region need to be chosen with some care whenever the free energy barrier is relatively small. Third, when F shows a free energy well, the permeants may be trapped in the membrane, and the integration boundaries should also be chosen with care.

Besides these conceptual fundamental issues related to the definition of the permeability, there are also practical issues. The danger exists that not all regions inside and outside of the membrane are accurately sampled in the equilibrium MD simulations. Lastly, slow sampling can also originate from a low diffusivity [$D(z)$ in Eq. (3)], e.g., for permeants that are bulky. The challenge of the Smoluchowski approach is to determine the model parameters, i.e., the $F(z)$ and $D(z)$ profiles, which can be extracted *a posteriori* from the MD trajectories in various ways [10,12,15,16,28,29,34]. When MD is inadequate because of its limited timescale, the approach can be combined with rare events simulation techniques like umbrella sampling [35], adaptive bias force [36–38], and biased diffusion [39,40]. All these free energy methods, with the aim to compute the permeability via Eq. (3), have however as a shared limitation that they build on the validity of the Smoluchowski equation. The validity of this equation is questionable for many complex systems in which the transfer of permeants over the barrier involves other types of motion,

like a rotation of the permeant or a local stretch of a membrane opening [38,41].

C. Path sampling approaches

Path sampling methods seem to provide a natural solution to the permeation problem since they are designed to maintain the natural dynamics of the process as much as possible, while still allowing the sampling of events that happen on long timescales. Among the different path sampling methods, applications on the permeation problem have so far mostly adopted the milestoning method [18]. In this technique, phase space is divided in domains that are separated by interfaces, called milestones. Trajectories are initiated at each milestone and run until they cross another milestone. The statistics over this set of short trajectories give the mean first passage times between pairs of milestones, which are then incorporated in a Markovian rate network model to extract the overall rate. Cardenas and Elber [42] proposed the formula for the permeability

$$P = \frac{J_1 \, q_f}{c_{\text{ref}} \, q_1}, \qquad (5)$$

where J_1 is the flux of particles hitting the first milestone per area and per time in equilibrium, q is the absolute flux vector of trajectories crossing the milestones, when solving the Markovian rate model for its steady-state solution with specific boundary conditions. Cardenas and Elber applied this to the permeation of a small peptide [42] or water molecule [43] through a 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) phospholipid bilayer, while Fathizadeh and Elber simulated potassium permeation through this DOPC membrane [44]. Recently, Votapka *et al.* derived an alternative formulation for the permeability based on milestoning [45].

The advantage of milestoning is that the MD trajectories that need to be generated are very short. However, the milestoning relies on the Markovian assumption that the system loses memory when it hits the next interface/milestone [18,46]. This assumption is only correct if the milestones are chosen along the isocommittor lines [19]. Hence, since the milestones are defined by fixed values of the RC, the RC should be the committor [47–50]. The committor is therefore often considered as the ideal RC since the dynamics projected on this one-dimensional coordinate becomes Markovian and models relying on the Markovian assumption, like Smoluchowski and milestoning, become exact. Note that the committor is in principle defined in phase space though by assuming that the dynamics is overdamped the committor can be defined in configuration space alone.

Hence, if milestoning is applied with this *ideal* RC (the committor) then trajectories released from the same milestone have the same probability of reaching $h/2$ before $-h/2$ regardless the point of origin within that milestone. This RC should account for all relevant rotations of the permeant, collective motions, and deformations of the membrane that could be vital for the permeation process. Including these motions within a single coordinate is highly nontrivial even if the committor is assumed to be only configuration space dependent based on the overdamped approximation. In practice, this is generally not even attempted. Rather, a simple intuitive

RC is chosen such as the z coordinate of the permeant that is followed. This pragmatic choice will generally invalidate the Markovian assumption leading to a systematic error that can be mild or severe depending on the system. Another path sampling approach that has similarities with milestoning is the partial path transition interface sampling (PPTIS) [51]. The PPTIS method is a Markovian variant of the TIS method. Still, it uses a less stringent assumption than milestoning by including more memory in the dynamical description. In PPTIS, if the system hits an interface, the chance to move to either its left or right interface depends on the history on the path i.e., from which interface (left or right) it came from. Still, no memory is retained before that point. The systematic error due to the nonideal RC is therefore presumably lower in PPTIS than in milestoning.

Path sampling methods that include the complete history dependence of the dynamics, such as transition path sampling (TPS), transition interface sampling (TIS), forward flux sampling (FFS) [52], adaptive multilevel splitting (AMS) [53], and replica exchange TIS (RETIS), are exact and independent of the RC, which is a big advantage in complex systems. The sampling of complete trajectories will make these methods generally more computationally intensive than milestoning or PPTIS, though a quantitative comparison relies on the trade-off between systematic and statistical error. FFS and AMS are based on a splitting approach which makes it applicable for nonequilibrium dynamics as well. On the other hand, the lack of backward-in-time integration limits the applicability of splitting methods to stochastic dynamics and leads to a relatively high risk of producing nonrepresentative transition trajectories [54]. The other three methods, TPS, TIS, and RETIS, are all based on a MC sampling procedure in path space. The original TPS approach for rate calculations is no longer being used in practice as TIS is both faster and more accurate than TPS. RETIS is even more efficient than TIS, but requires a more complex implementation. By the emergence of open source path sampling libraries like OPS [55,56] and PyRETIS [57,58] the latter aspect has become less of an issue.

In this paper, we show how RETIS can effectively be used to compute permeability coefficients equally exact as the direct counting approach, but orders of magnitude faster. In the next section, we shortly introduce the RETIS approach, then show how it can be amended for permeability calculations in Sec. IV, and show in Sec. VI two new path MC moves that can further enhance ergodic sampling.

III. REPLICA EXCHANGE TRANSITION INTERFACE SAMPLING

TIS and RETIS are rare event techniques that allow to compute rate constants k when transitions have to overcome a high free energy barrier, and thus transitions are unlikely to be observed in a standard MD simulation. Intuitively the rate constant k can be expressed as a number of transitions, from reactant state to product state, per unit time and per amount of reactants. However, the translation of this phenomenological rate constant into a computational measurable property is far from trivial since it requires a microscopic definition of the reactant state and product state. If a single dividing surface is used to assign whether a molecular system is in the reactant
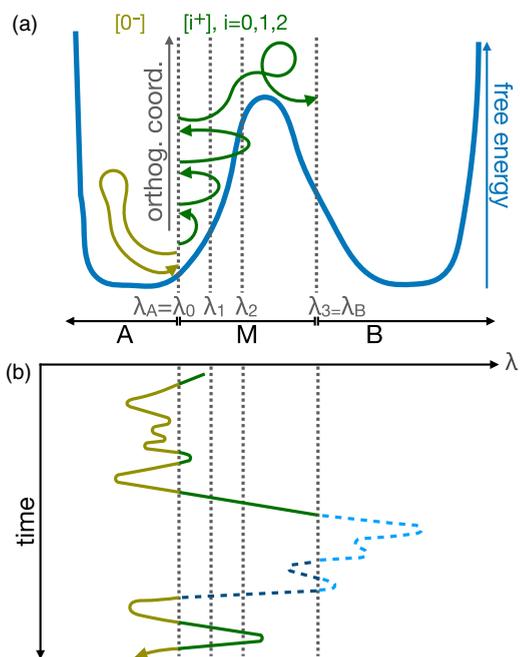
FIG. 1. Path ensembles in RETIS. (a) Paths shown along the RC λ and an arbitrary orthogonal coordinate. The free energy as a function of λ is also shown. All top four paths are part of the $[0^+]$ ensemble containing paths that start at $\lambda_A = \lambda_0$, move in the positive direction and end at λ_A or λ_B. The top three path are part of the $[1^+]$ ensemble. These are like the $[0^+]$ paths but should cross λ_1 in addition. The top two paths are part of the $[2^+]$ path ensemble. The bottom path is a $[0^-]$ path that starts at $\lambda_A = \lambda_0$, moves in the negative direction and ends at λ_A. The stable state regions A and B, and the middle region M (no man's land) are shown. Points in the region M can be part of the larger overall states $\mathcal{A}$ or $\mathcal{B}$. (b) Reaction coordinate as a function of time for a hypothetical long equilibrium MD run. The line is solid green (light: *stable* state A, dark: region M) when the system belongs to *overall* state $\mathcal{A}$. It is dashed blue (bright: *stable* state B, dark: region M) when part of *overall* state $\mathcal{B}$.

state or product state, e.g., based on a geometric observable being lower or higher than a specific value, it is expected to observe many correlated recrossings of the dividing surface. Only when the system has moved far enough beyond the dividing surface, it can be considered as stabilized to the other state. To avoid the issue of correlated recrossings, it can be generally better to consider two separate dividing surfaces left and right of the barrier, λ_A and λ_B, respectively [see Fig. 1(a)]. At the left of λ_A and at the right of λ_B, the system is considered committed to the reactant state and product state respectively. The disadvantage is that the barrier region between the dividing surfaces is not assigned to either state. A transition from reactant state to product state needs to cross the *no man's land* between A and B, which makes it difficult to assign for each transition a specific point in time at which it takes place, which is essential to avoid overcounting.

The TIS and RETIS approaches circumvent this problem by introducing *overall states* which are history-dependent state definitions. Using the two dividing surfaces, λ_A and λ_B, the system is part of *stable state A* if the value of the chosen RC is below λ_A and it is part of *stable state B* if it is higher

than λ_B. The overall states are denoted by the curly letters $\mathcal{A}$ and $\mathcal{B}$, and they include the stable states A and B, respectively. Further, any instance of the system traveling in no man's land is assigned to *overall state* $\mathcal{A}$ or *overall state* $\mathcal{B}$ based on the stable state that was most recently visited.

The advantage is that the system is always assigned to either overall state $\mathcal{A}$ or $\mathcal{B}$, and the overall state regions are rather insensitive to the placement of the dividing surfaces λ_A and λ_B as long as these are reasonable [23]. In addition, a transition from $\mathcal{A}$ to $\mathcal{B}$ is well defined without hindrance of recrossings and leads to a microscopically measurable net flux without perturbing the equilibrium conditions. As such, the phenomenological rate constant can be expressed as

$$k = \lim_{\Delta t \to 0} \frac{\langle h_{\mathcal{A}}(0) h_{\mathcal{B}}(\Delta t) \rangle}{\langle h_{\mathcal{A}} \rangle \, \Delta t}, \tag{6}$$

where $h_X(t)$ equals 1 when the system is in state X at time t and 0 when the system is not in state X at time t. The interpretation of the ensemble average that involves history dependent functions is given in Appendix A.

The above equation counts the rare crossings from left to right through the surface λ_B of points that actually come from λ_A (when the equations of motion are followed backward in time, λ_A is reached without crossing λ_B). Since we can assume equilibrium, the same number of counts per second will be obtained by considering crossings with λ_A that after crossing that interface reach λ_B without recrossing λ_A. This allows us to write the rate k as

$$k = f_A P_A(\lambda_B | \lambda_A), \tag{7}$$

where f_A is the conditional flux (crossings per time) through λ_A counting all crossings in the positive direction per time spent in state $\mathcal{A}$, and $P_A(\lambda_B|\lambda_A)$ is the crossing probability, i.e., the chance that a crossing with λ_A is followed by a crossing with λ_B before a recrossing with λ_A occurs.

f_A can easily be computed with MD, as is done in TIS, since crossing λ_A is not a rare event. In RETIS, the flux is computed from the average path lengths of two path ensembles. The crossing probability $P_A(\lambda_B|\lambda_A)$ is generally too low to be computed by straightforward MD, but can be recast into the following factorization by defining a set of nonintersecting interfaces:

$$P_A(\lambda_B|\lambda_A) = P_A(\lambda_n|\lambda_0) = \prod_{i=0}^{n-1} P_A(\lambda_{i+1}|\lambda_i). \tag{8}$$

Here, $P_A(\lambda_{i+1}|\lambda_i)$ is the history dependent conditional probability that, given there is a crossing with interface λ_i for the first time since last $\lambda_A = \lambda_0$ crossing, interface λ_{i+1} will be crossed before λ_A. These conditional probabilities are computed in $n-1$ different path sampling simulations. Each simulation samples a different path ensemble using a set of different MC moves to generate new paths in the ensembles. The ensemble $[i^+]$ consists of all possible paths that start at λ_A and end at λ_A or λ_B and have at least one crossing with λ_i. The MC approach is tuned such that the same statistical distribution of paths is generated as the distribution of paths that would result if these are cut out from a hypothetical extremely long MD simulation. The fraction of paths in the $[i^+]$ path ensemble that cross λ_{i+1} equals $P_A(\lambda_{i+1}|\lambda_i)$. Performance of

TIS is optimal when the number of interfaces and the spacing between the interfaces are tuned such that each conditional probability is around 0.2 [59]. The most important MC move is the so-called shooting move in which a random point of the previous path is picked, its velocities are randomly modified to generate a new phase point, and from this phase point the equations of motion are followed backward and forward in time to generate a new path. The new path will be accepted if it fulfills the requirements for the specific ensemble (like crossing λ_i in the $[i^+]$ ensemble) and, depending on the type of velocity randomization procedure, an additional Metropolis acceptance/rejection step will be invoked. In this work, we applied the aimless velocity modification in which velocities are regenerated, independent from the old velocities, from a Maxwell-Boltzmann distribution [60].

Compared to TIS, RETIS has one additional path ensemble called $[0^-]$. Like the other ensembles, this ensemble contains paths starting at λ_0, but from there the paths move in the negative direction away from the barrier. The paths are terminated when they cross λ_0 again. While the flux f_A in Eq. (7) is computed with straightforward MD in TIS, in RETIS it is computed from the average path lengths in the $[0^-]$ and $[0^+]$ ensembles (see Sec. IV A).

Another difference between RETIS and TIS, is that RETIS employs additional MC moves. Since RETIS is purely based on path sampling simulations, instead of MD and path sampling simulations, replica exchange moves between the different path ensembles (parallel path swapping [22]) can be applied throughout the complete RETIS simulation. The swapping moves enhance the sampling in a similar way as parallel tempering [61], since the $[i^+]$ path ensemble for a given i tends to contain trajectories moving on a higher energy surface than the trajectories of the $[(i-1)^+]$ ensemble. The higher energy trajectories are less likely to get trapped in specific reaction channels that are separated by free energy barriers orthogonal to the RC λ [62]. As paths between $[(i-1)^+]$ and $[i^+]$ are sometimes swapped, also the $[(i-1)^+]$ path ensemble will sample these different reaction channels more easily. The combination of TPS and standard parallel tempering can provide a similar effect [63], though in RETIS there are no additional simulations at alleviated temperatures needed. In contrast, the $[(i-1)^+] \leftrightarrow [i^+]$ swapping move is very inexpensive as it does not require any force evaluations. If the move is accepted, it provides a new path for both the $[(i-1)^+]$ and the $[i^+]$ path ensemble. In addition, the accepted swapping moves provide generally paths that are more decorrelated from the previous path than a shooting move.

The swapping moves between the $[0^-]$ and $[0^+]$ path ensembles are done by exchanging the end and start points of the paths and extending those forward and backward in time respectively. Despite not being a free move, the barrierless diffusion within the reactant well of the $[0^-]$ paths, followed by exchange moves, will basically feed the $[0^+]$ ensemble with fresh initializations. This facilitates the decorrelation of the MC sampling even further and orthogonal barriers can be avoided without having to cross them in any of the path ensembles. An additional advantage is that this method works in case where parallel tempering is not effective, that is when barriers are mainly entropic in nature.

IV. PERMEABILITY FROM RETIS SIMULATIONS

The rate k in Eq. (6) (with unit 1/time) describes the kinetics of a process, while the permeability P of a membrane (with unit length/time) describes the transport kinetics through the layer, and clearly a link between k and P is expected. However not only the units, but also the formalism for permeability and rates are somewhat different. This paper will derive the correct connection between the TIS framework for rare events and the permeability.

The progression of permeation is most easily measured by the z coordinate orthogonal to the membrane as the RC. Despite this simple order parameter choice, some details ask for attention. In case of periodic boundary conditions, common in MD simulations of membranes or porous catalytic crystals, the question arises how it may be detected whether a molecule crossed the membrane, or whether it simply circled back to the other side of the membrane through the water phase because of the periodic boundary condition in the z-direction. Therefore we first derive the connection between P and k when the solvent phase is bounded on both sides (e.g., by hard walls as in Fig. 1) in Sec. IV A, and second we derive a relation for the unbounded system when either periodic boundary conditions or an infinite particle bath [64] are applied in Sec. IV B.

A. Connecting permeability and rate

Figure 1(b) shows a hypothetical long MD equilibrium run. The timescales are not very realistic as we would in practice expect that thousands of crossings with λ_A would proceed a transition to state B. Yet, it will be used to show how we can subdivide an ensemble average into different regions A, B or $\mathcal{A}$, $\mathcal{B}$ as follows.

The long trajectory can be seen as a series of visited phase points. Here, we assume that the MD equilibrium run is in fact sampling the distribution of interest. The MD integrator is hence coupled to some kind of thermostat or barostat, whenever this distribution is different from the NVE ensemble. The MD integrator gives a trajectory, which is effectively a series of phase points (time slices) because of the discrete time step Δt in the numerical integration, such as the velocity Verlet integration algorithm [65]. Given ergodicity, the phase points will be visited with the correct relative probability when the trajectory length T goes to infinity. The trajectory phase points can be divided into subsets corresponding to the regions A, M, B in Fig. 1(b) based on the value of the order parameter in each phase point. The ensemble average becomes

$$\langle \ldots \rangle = p_A \langle \ldots \rangle_A + p_M \langle \ldots \rangle_M + p_B \langle \ldots \rangle_B, \quad (9)$$

where M is the membrane region, previously referred to as no man's land in the context of TIS and RETIS. Here, p_X is the probability that the system is in state X. Given that the ergodicity makes the time averages respect the relative probabilities, it follows that $p_X = T_X/T$, where T_X is the total time spent in state X, and $T = \sum_X T_X$. In Eq. (9), the notation $\langle \ldots \rangle_A$ refers to the ensemble average over all trajectory phase points associated to A.

As mentioned earlier in Sec. III, the trajectory phase points can also be assigned to either overall state $\mathcal{A}$ or $\mathcal{B}$. The assignment to $\mathcal{A}$ or $\mathcal{B}$ is generally not based on the evaluation of

the order parameter in a single phase point but rather based on the series of phase points in the trajectory (history dependent). The ensemble average can be divided in two contributions,

$$\langle \ldots \rangle = p_{\mathcal{A}} \langle \ldots \rangle_{\mathcal{A}} + p_{\mathcal{B}} \langle \ldots \rangle_{\mathcal{B}}. \quad (10)$$

The notation $\langle \ldots \rangle_{\mathcal{A}}$ refers to the ensemble average over all trajectory phase points associated to $\mathcal{A}$ (see Appendix A).

Central to the RETIS methodology are the $[i^+]$ path ensembles connected with interfaces at locations λ_i, $i = 0, 1, 2, \ldots$ and the $[0^-]$ path ensemble [see Fig. 1(a)]. Note that any path in the $[i^+]$ ensemble for $i > 0$ is automatically a valid path in the $[0^+]$ ensemble. The $[0^-]$ paths completely lie in state A except for the first and last points of the trajectories. In addition, all phase points inside A must lie on a $[0^-]$ path. We can therefore say that state A is equivalent to $[0^-]$. That is, we can view each path ensemble also as a phase space ensemble by simply collecting all phase space points that lie on corresponding paths except for the terminating points. The $[0^-]$ and $[0^+]$ ensemble (excluding the end points) are disjunct and combined they represent all trajectories that pass through the λ_0 interface while not having reached state B yet. The $\mathcal{A}$ ensemble is therefore equal to the combined $[0^-]$ and $[0^+]$ ensembles.

As is obvious from Fig. 1(b), an equal number of $[0^-]$ as $[0^+]$ trajectories can be cut out from the equilibrium run; each time the end point of a $[0^-]$ path comprises a start point of $[0^+]$ path and, vice versa, each starting point of a $[0^+]$ path relates to an end point of a $[0^-]$ path. When N_X denotes the number of paths in the X ensemble that can be cut out from a long equilibrium MD run, this observation can be summarized as $N_{[0^-]} = N_{[0^+]}$ in thermodynamic equilibrium.

Let us compare the three different flux terms: k in Eq. (6), f_A in Eq. (7) and J^+ in Eq. (2). Here, f_A provides the frequency of state changes from A to M within the overall state ensemble $\mathcal{A}$. J^+ on the other hand measures all crosses from left to right along the full region M of *all* permeants, per time and per membrane surface area σ. The rate constant k finally also measures the frequency of full crossings like J^+ but with the same time normalization $T_{\mathcal{A}}$ as for f_A. In summary,

$$f_A = \frac{\#(A \rightarrow M)_{\text{target}}}{T_{\mathcal{A}}},$$

$$k = \frac{\#(A \rightarrow M \rightarrow B)_{\text{target}}}{T_{\mathcal{A}}},$$

$$J^+ = \frac{\#(A \rightarrow M \rightarrow B)_{\text{all perm}}}{T \sigma}. \quad (11)$$

One could imagine another flux definition, similar to f_A, but where the denominator is the total time T of the long equilibrium run rather than the time spent in $\mathcal{A}$,

$$f = \frac{\#(A \rightarrow M)_{\text{target}}}{T} = p_{\mathcal{A}} f_A. \quad (12)$$

The appearance of the factor $p_{\mathcal{A}}$ gives f_A the interpretation of a "conditional" flux compared to f. The flux f is however not accessible in a RETIS simulation and will not be further discussed.

If all of the N_p permeating particles are identical,

$$\#(A \rightarrow M \rightarrow B)_{\text{all perm.}} = N_p \#(A \rightarrow M \rightarrow B)_{\text{target}} \quad (13)$$

and we can relate J^+ with f_A and k as follows:

$$J^+ = N_p \frac{\#(A \to M \to B)_{\text{target}}}{T_{\mathcal{A}}\sigma} \frac{T_{\mathcal{A}}}{T}$$
$$= \frac{N_p k p_{\mathcal{A}}}{\sigma} = \frac{N_p f_A P_A(\lambda_B|\lambda_A) p_{\mathcal{A}}}{\sigma}, \qquad (14)$$

where $P_A(\lambda_B|\lambda_A)$ is the previously introduced crossing probability. All terms in Eq. (14) are measurable in a RETIS simulation except $p_{\mathcal{A}}$, the probability of overall state $\mathcal{A}$. RETIS simulates the $\mathcal{A}$ ensemble (through $[0^-]$ and $[0^+]$) but not the $\mathcal{B}$ ensemble; hence $T_{\mathcal{B}}$, $T = T_{\mathcal{A}} + T_{\mathcal{B}}$ and ultimately $p_{\mathcal{A}} = T_{\mathcal{A}}/T$ are not known. Luckily, the factor $p_{\mathcal{A}}$ cancels when we include the c_{ref} concentration in order to compute the permeability based on Eq. (2), as we will show now.

First, let ρ_{ref} be the one-dimensional probability density to find a specific permeant (for instance the targeted permeant) at the reference location z_{ref}. We can relate ρ_{ref} to c_{ref} by taking into account the number of permeants N_p in the simulation box and the cross section area σ,

$$\rho_{\text{ref}} = \frac{\sigma c_{\text{ref}}}{N_p}. \qquad (15)$$

Further, we can use the subdivision in global states in Eq. (10), giving

$$\rho_{\text{ref}} = \rho(z_{\text{ref}}) = \langle \delta(z_{\text{ref}} - z_t) \rangle$$
$$= p_{\mathcal{A}} \langle \delta(z_{\text{ref}} - z_t) \rangle_{\mathcal{A}} + p_{\mathcal{B}} \langle \delta(z_{\text{ref}} - z_t) \rangle_{\mathcal{B}}$$
$$= p_{\mathcal{A}} \langle \delta(z_{\text{ref}} - z_t) \rangle_{\mathcal{A}} = p_{\mathcal{A}} (\rho_{\text{ref}})_{\mathcal{A}}, \qquad (16)$$

where z_t is the z coordinate of the target permeant, and $(\rho_{\text{ref}})_{\mathcal{A}}$ refers to the probability density within the $\mathcal{A}$ ensemble. The factor $\langle \delta(z_{\text{ref}} - z_t) \rangle_{\mathcal{B}}$ is zero, since z_{ref} is located in stable state A and z_t does not lie in A for any trajectory phase point that is assigned to $\mathcal{B}$ (see Appendix A). Consequently, we can write

$$c_{\text{ref}} = \frac{N_p \rho_{\text{ref}}}{\sigma} = \frac{N_p (\rho_{\text{ref}})_{\mathcal{A}} p_{\mathcal{A}}}{\sigma}. \qquad (17)$$

Substitution of Eqs. (17) and (14) in Eq. (2) gives

$$P = \frac{k}{(\rho_{\text{ref}})_{\mathcal{A}}} = \frac{f_A P_A(\lambda_B|\lambda_A)}{(\rho_{\text{ref}})_{\mathcal{A}}}. \qquad (18)$$

This is the first theoretical result connecting P and k. Equation (18) shows that the $\mathcal{B}$ path ensemble needs not be simulated to find the permeability. In contrast, the direct counting method of Eq. (2) is based on a long MD run, which includes paths in both the $\mathcal{A}$ and $\mathcal{B}$ ensembles. The paths in the $\mathcal{A}$ ensemble are however sufficient for the computation of the permeability with Eq. (18).

Still, the denominator $(\rho_{\text{ref}})_{\mathcal{A}}$ in Eq. (18) cannot be computed in a single path ensemble since overall state $\mathcal{A}$ comprises both $[0^-]$ and $[0^+]$. Therefore let us first consider how the probability density at the reference location would be computed from a long equilibrium MD simulation. We would need to define a certain interval around z_{ref} with a width Δz, $[z_{\text{ref}} - \Delta z/2, z_{\text{ref}} + \Delta z/2]$, and $(\rho_{\text{ref}})_{\mathcal{A}}$ is the ratio of the average time T_{ref} spent in the reference interval region versus the total time spent in $\mathcal{A}$, divided by Δz,

$$(\rho_{\text{ref}})_{\mathcal{A}} = \frac{T_{\text{ref}}}{T_{\mathcal{A}}} \frac{1}{\Delta z}. \qquad (19)$$

In practice, one would count the number of steps that the z coordinate is inside the interval, and divide this by Δz and by the total number of steps that the system is part of $\mathcal{A}$ paths, assuming a constant MD integration time step Δt. If the reference location is in a region where the free energy profile $F(z)$ is flat, then T_{ref} scales linearly with Δz, and $(\rho_{\text{ref}})_{\mathcal{A}}$ will in principle not be affected by the chosen interval length Δz.

On the other hand, the conditional flux is the number of crossings through λ_A in the positive direction divided by the time spent in $T_{\mathcal{A}}$,

$$f_A = \frac{N_{[0^+]}}{T_{\mathcal{A}}}, \qquad (20)$$

where, as said earlier, $N_{[0^+]}$ is the number of $[0^+]$ trajectories that can be cut out from the long equilibrium trajectory. Substitution of the two previous equations into the permeability in Eq. (18) makes the $T_{\mathcal{A}}$ drop out, and gives a practical expression for P,

$$P = \frac{N_{[0^+]}}{T_{\text{ref}}} P_A(\lambda_B|\lambda_A) \, \Delta z. \qquad (21)$$

This is the second expression linking P with RETIS quantities. It gives the insight that P depends on time spent in the reference region, but not explicitly on time spent in the $[0^+]$ nor $[0^-]$ ensemble.

However, Eq. (21) still refers to the quantities $N_{[0^+]}$ and T_{ref}, which are obtained from a long equilibrium simulation. In the next last step, the conversion to path ensemble averages is made. With T_X the time spent in a path ensemble X and N_X the number of trajectories in the path ensemble X, the average path length is given by $\tau_X = T_X/N_X$. An advantage of τ_X is that it is in principle independent of the simulation computer time, i.e., if the number of simulated paths in the ensemble is doubled, the average τ_X will not change, which gives τ_X an intrinsic meaning. We would like to stress that a path ensemble average like τ_X is a property that is averaged over all paths in the path ensemble, and it differs from the common phase space average, denoted as $\langle \ldots \rangle_Y$, where the average is taken over all phase points within the ensemble Y.

The conversion to path averages is now simply executed by introducing factors $N_{[0^-]}$, using $T_{\mathcal{A}} = T_{[0^-]} + T_{[0^+]}$, and exploiting the fact that $N_{[0^-]} = N_{[0^+]}$ in an equilibrium run. For the probability density, this gives

$$(\rho_{\text{ref}})_{\mathcal{A}} = \frac{T_{\text{ref}}}{T_{[0^-]} + T_{[0^+]}} \frac{1}{\Delta z}$$
$$= \frac{T_{\text{ref}}/N_{[0^-]}}{T_{[0^-]}/N_{[0^-]} + T_{[0^+]}/N_{[0^+]}} \frac{1}{\Delta z}$$
$$= \frac{\tau_{\text{ref},[0^-]}}{\tau_{[0^-]} + \tau_{[0^+]}} \frac{1}{\Delta z}, \qquad (22)$$

while the conditional flux is converted to (see Ref. [22])

$$f_A = \frac{N_{[0^+]}}{T_{[0^-]} + T_{[0^+]}}$$
$$= \frac{1}{T_{[0^-]}/N_{[0^-]} + T_{[0^+]}/N_{[0^+]}}$$
$$= \frac{1}{\tau_{[0^-]} + \tau_{[0^+]}}. \qquad (23)$$

Here, $\tau_{[0^-]}$ and $\tau_{[0^-]}$ are the average lengths of paths in the $[0^-]$ path ensemble and $[0^+]$ path ensemble, respectively, while $\tau_{\mathrm{ref},[0^-]}$ is the average time spent in the $[z_{\mathrm{ref}} - \Delta z/2 : z_{\mathrm{ref}} + \Delta z/2]$ interval per path in the $[0^-]$ ensemble.

Again, substituting the previous equations into the permeability in Eq. (18) gives another expression for P,

$$P = \frac{k\Delta z}{\tau_{\mathrm{ref},[0^-]} f_A} = \frac{P_A(\lambda_B|\lambda_A)\Delta z}{\tau_{\mathrm{ref},[0^-]}}. \qquad (24)$$

This is the third theoretical result expressing P in terms of intrinsic RETIS quantities. Equation (24) is an expression that can be fully computed in a RETIS simulation since $P_A(\lambda_B|\lambda_A)$ is standard output of this approach, while $\tau_{\mathrm{ref},[0^-]}$ can easily be obtained from the paths generated in the $[0^-]$ ensemble by histogramming the z coordinate.

Figure 1 is however not typical for an actual membrane system that has no natural confining energy barrier that prohibits the permeants to drift far away from the membrane. To deal with the situation of an unconfined system, we derive in Sec. IV B how a λ_{-1} interface can be used to constrain the system in a way that the permeability coefficient can still be determined exactly, as if the system were unrestrained.

B. The λ_{-1} interface

The previous section linked the rate constant to a permeability, though the system depicted in Fig. 1 is not exemplary for a permeation system where we imagine a membrane in a near-infinite solvent. The rate constant k that follows a single particle will naturally decrease with the amount of solvent added to the model system since the target particle will spend more time away from the membrane. The permeability P is insensitive to this since it does not depend on the size of the simulation box. Increasing the solvent while maintaining constant concentration automatically implies an increase in the number of permeants. This increase cancels the decrease of the rate for permeation of the individual particles. Still, as RETIS computes rates rather than permeation directly, some confinement is required in practice.

This confinement is achieved by introducing an extra interface λ_{-1}. With the introduction of this interface the overall state $\mathcal{A}$ is reduced to the $\lambda > \lambda_{-1}$ region, and is denoted $\mathcal{A}'$. The same can be done at the product region side, but this is not essential for the $A \to B$ rate calculation. The new state division is depicted in Fig. 2, where the outer regions are here called *freeze-time zones* which will not be accessed in this adaptation of the RETIS algorithm. A prime is added to show that the states A', B', $\mathcal{A}'$, and $\mathcal{B}'$ contain fewer phase points than A, B, $\mathcal{A}$, and $\mathcal{B}$, respectively. The ensemble average in Eq. (10) is updated to reflect the additional *freeze-time* zones,

$$\langle \ldots \rangle = p_{\mathrm{fr\text{-}ti}}\langle \ldots \rangle_{\mathrm{fr\text{-}ti}} + p_{\mathcal{A}'}\langle \ldots \rangle_{\mathcal{A}'} + p_{\mathcal{B}'}\langle \ldots \rangle_{\mathcal{B}'}. \qquad (25)$$

The path ensemble $[0^{-\prime}]$ resembles the $[0^-]$ ensemble, but contains paths that can start and end at λ_{-1} in addition to λ_0. The time slices of paths in path ensembles $[0^{-\prime}]$ and $[0^+]$ fully build up the overall state $\mathcal{A}'$. The freeze-time zone on the product side does not affect the $[0^+]$ ensemble nor the other $[i^+]$ ensembles. The (RE)TIS crossing probability $P_A(\lambda_B|\lambda_A)$ remains unaffected by the λ_{-1} interface.
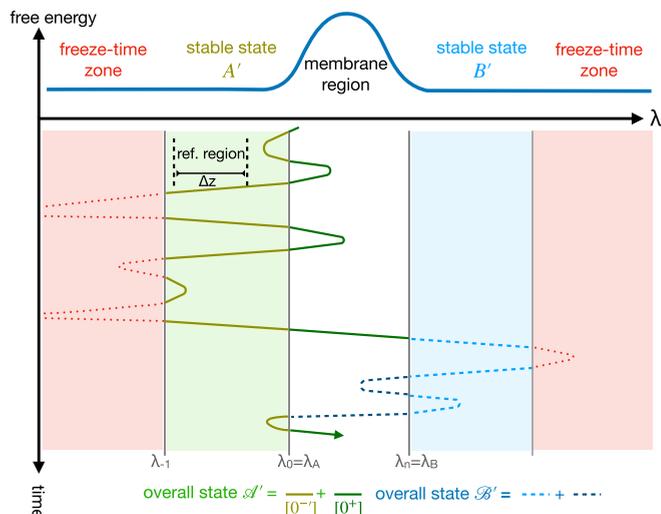
FIG. 2. Illustration of how the *freeze-time zones* in red are introduced by extra interfaces left of λ_A (called λ_{-1}) and right of λ_B (unnamed). The green and blue regions refer to states A' and B', respectively. The line represents a long equilibrium run (RC λ versus time T). Note that in the membrane region (previously called *no man's land*) the color is given by the stable states last visited [like in Fig. 1(b)]. Whenever the system enters a freeze-time zone, the time T' is stopped and continued whenever it exits this zone. The trajectories that can be cut out from the $[\lambda_{-1}, \lambda_0]$ interval constitute the $[0^{-\prime}]$ path ensemble. This ensemble is different from the $[0^-]$ ensemble since its paths can start and end at λ_{-1} in addition to λ_0.

Figure 2 shows again a hypothetical unrestrained equilibrium MD run just like in Fig. 1(b) with a free energy surface that is flat at either side of the membrane. Conceptually, the existence of an equilibrium in such an infinite system is not obvious as the ergodicity hypothesis implies that for instance ensemble averages are identical to time averages of an infinite equilibrium run. However, if the partition function diverges, even an infinitely long equilibrium run might not visit all the relevant phase space regions. This issue can be solved conceptually by taking the infinite limits for space and time in a controlled way. That is, we consider the potential of Fig. 2 as a special case of the potential shown in Fig. 1, but with the vertical-like increase of the free energy occurring at $z = -W$ and $+W$. By letting $T \to \infty$ and $W \to \infty$ such that $W^2/(TD) \to 0$ with D the average diffusion constant, it can be shown that the ergodicity hypothesis holds. While this solves the conceptual problem whether we can assume equilibrium statistics, it does not solve the practical problem that the rate k is still zero in this limit.

The additional λ_{-1} interface, however, ensures that the overall state $\mathcal{A}'$ becomes finite and that anything that happens in the freeze-time zone can be ignored as if the stopwatch, measuring T', is paused and statistics are not updated each time that the trajectory enters that region.

Now that the rate k' and the flux $f_{A'}$ can be computed from this long equilibrium run using the new definition for the overall state $\mathcal{A}'$, the same counting strategy applies: it is a number of crossings/transitions divided by the time spent in overall state $\mathcal{A}'$. The only difference is that, besides time spent in overall state $\mathcal{B}'$, also the freeze-time zone is ignored in the

normalization. The path ensemble $[0^-]$ is however changed to $[0^{-\prime}]$, which statistically corresponds to the same ensemble one would obtain by cutting out the trajectory segments between λ_{-1} and λ_0 from the equilibrium run. Specifically, it is no longer valid that $N_{[0^{-\prime}]}$ and $N_{[0^+]}$ are equal. We will therefore introduce the parameter ξ to measure the mismatch between number of paths in these two path ensembles,

$$\xi = \frac{N_{[0^+]}}{N_{[0^{-\prime}]}} = \frac{N_{\to R,[0^{-\prime}]}}{N_{[0^{-\prime}]}} = \overline{h_{\to R}[0^{-\prime}]}. \quad (26)$$

Here $N_{\to R,[0^{-\prime}]}$ is the number of $[0^{-\prime}]$ paths ending at the right side (at λ_0) that can be cut out of the long equilibrium run. It is obvious that $N_{\to R,[0^{-\prime}]} = N_{[0^+]}$ from observing the λ_A interface in Fig. 2. It follows that $\xi < 1$.

Similarly, $h_{\to R}$ is the characteristic function that for each path provides the output 1 if the path ends at the right and 0 if the path ends at the left, irrespective of the starting point. Finally, $\overline{h_{\to R}[0^{-\prime}]}$ is the average of this function over all paths in the $[0^{-\prime}]$ ensemble. This parameter ξ can hence be obtained from a RETIS simulation by the analysis of the $[0^{-\prime}]$ path ensemble.

In order to reformulate the expressions for $f_{A'}$ and P, we first update the expressions for $T_{A'}$ and $T_{\text{ref}'}$ to link them to average path lengths τ_X, which are intrinsic quantities of a given path ensemble.

The time $T_{A'}$ can be written to be proportional to $N_{[0^+]}$ as

$$T_{A'} = T_{[0^{-\prime}]} + T_{[0^+]} = N_{[0^{-\prime}]}\tau_{[0^{-\prime}]} + N_{[0^+]}\tau_{[0^+]}$$
$$= N_{[0^+]}(\xi^{-1}\tau_{[0^{-\prime}]} + \tau_{[0^+]}). \quad (27)$$

The time $T_{A'}$ is smaller than T_A since $\mathcal{A}'$ comprises fewer phase points than $\mathcal{A}$ (the clock is stopped). The time T_{ref} spent in the reference interval is not affected by the presence of the λ_{-1} as the reference interval is located between λ_{-1} and λ_A. The associated intrinsic quantity $\tau_{\text{ref},[0^{-\prime}]}$ can nevertheless change as $N_{[0^{-\prime}]}$ might differ from $N_{[0^-]}$ (e.g., seven paths versus four paths in Fig. 2),

$$T_{\text{ref}} = N_{[0^{-\prime}]}\tau_{\text{ref},[0^{-\prime}]}$$
$$= N_{[0^+]}\xi^{-1}\tau_{\text{ref},[0^{-\prime}]} \quad (28)$$

These two equations have implications for $f_{A'}$ and P. For $f_{A'}$ in Eq. (20), the number of paths $N_{[0^+]}$ remains unaltered according to Fig. 2, but the time T_A needs to be updated with $T_{A'}$, leading to

$$f_{A'} = \frac{N_{[0^+]}}{T_{A'}} = \frac{\xi}{\tau_{[0^{-\prime}]} + \xi\,\tau_{[0^+]}}. \quad (29)$$

For P in Eq. (21), the number of paths $N_{[0^+]}$ and the crossing probability $P_A(\lambda_B|\lambda_A)$ remain unaltered, and the time T_{ref} is updated with Eq. (28), leading to the generalized version of Eq. (24),

$$P = \frac{\xi\,\Delta z}{\tau_{\text{ref},[0^{-\prime}]}} P_A(\lambda_B|\lambda_A). \quad (30)$$

The last expression is the central expression that links the permeability with thermodynamic averages that can be computed in a RETIS path sampling simulation. The parameter ξ is obtained from analyzing the end points of the $[0^{-\prime}]$ paths, the crossing probability $P_A(\lambda_B|\lambda_A)$ is a direct output of a (RE)TIS simulation, Δz is a chosen bin width of the reference region

in the flat free energy region to the left of λ_0, and $\tau_{\text{ref},[0^{-\prime}]}$ is the corresponding time spent per path in the $[0^{-\prime}]$ ensemble in this reference bin. The latter can be rewritten as

$$\tau_{\text{ref},[0^{-\prime}]} = \frac{T_{\text{ref}}}{T_{A'}}\frac{T_{A'}}{N_{[0^{-\prime}]}} = (\rho_{\text{ref}})_{A'}\Delta z\,\tau_{[0^{-\prime}]}, \quad (31)$$

which leads to an alternative expression for the permeability

$$P = \frac{\xi P_A(\lambda_B|\lambda_A)}{(\rho_{\text{ref}})_{A'}\tau_{[0^{-\prime}]}}. \quad (32)$$

Equation (32) combines path ensemble averages (ξ, $P_A(\lambda_B|\lambda_A)$, $\tau_{[0^{-\prime}]}$) with a phase space average ($(\rho_{\text{ref}})_{A'}$), while Eq. (30) is only based on path ensemble averages.

All quantities in Eqs. (30) and (32) are intrinsic, meaning they are expressed as averages over the paths. This means that the same expression is applicable when doubling the number of paths, i.e., running the MC steps in the path ensembles longer. This will not affect the absolute value of any of the terms provided in Eq. (30), but will naturally increase the accuracy of their numerical estimates.

V. RC FOR PERMEATION WITH PERIODIC BOUNDARY CONDITIONS

In the previous section, we solved the problem to link the permeability with the rate constant in an infinite system by introducing the interface λ_{-1} and the rate constant k', which is nonzero unlike k. In most practical simulations the infinite system is represented by a system with periodic boundary conditions (PBC) which poses the need to properly determine the relative position of the permeant with respect to the center of the membrane, which defines the RC, λ. The simple minimum-image convention will generally not work since the RETIS trajectories can span the full $[\lambda_{-1}, \lambda_B]$ region and a permeant located at λ_{-1} might actually be closer to the left periodic image of the membrane than to the membrane in the central image.

In order to properly deal with PBC, we first map the z_i coordinates of all particles i in the system within the $[-L_z/2, L_z/2]$ interval, where $z = 0$ is matched by convention at the center of mass of the membrane and L_z is the box length in the z dimension,

$$z_i = (z_i' - z_{\text{mem.}}') - \text{round}\left(\frac{(z_i' - z_{\text{mem.}}')}{L_z}\right)L_z, \ \forall i. \quad (33)$$

Here, z_i' is the z coordinate of particle i provided by the molecular simulation program. By this operation, the center of mass of the membrane is set at 0, while the center of the solvent slab is at $\pm L_z/2$. Here we assume that original coordinates z_j' of the membrane particles j are constructed such that the center of mass at the z' axis can be computed without the need to add or subtract L_z or multiples of it to any of the membrane particles. As the membrane is stable, we can assume that the above remains valid during the full simulation. That is, membrane particles might move over large distances only if all membrane particles move in cohort.

The RC is given by the relative position z_i of the tagged permeant i (target) with respect to the membrane. This implies that if we want to compute the permeability of oxygen through a membrane and our model system contains N_p

oxygen molecules, one of those will be selected and considered as our target permeant.

In some cases, it can be advantageous to select a collective RC such as the maximum value of the z coordinates of all permeants. This has the advantage that the rate increases which makes it less of a rare event and is therefore easier to compute. In addition, the collective RC facilitates the decorrelation of the sampling since the target permeant defining the RC can switch during the simulation. This strategy was for instance applied to study water dissociation where the RC was defined as the largest OH bond in the system [66]. Also in a recent paper on permeation by some of us [67] such an approach was applied to compute escape rates of permeants being trapped in a membrane.

The reason that we nevertheless choose here a RC based on a single target permeant is because the permeation problem is in some applications less of a rare event than for example water dissociation. This implies that for a system with many permeants there is almost always one of them in the membrane region. In fact, the membrane often does not correspond to a single peaked free energy barrier, but may have a well in the middle where permeants get temporary trapped. This makes a collective RC impracticable for describing stable state A. In addition, there are technical and theoretical problems associated to such a collective RC for the calculation of $\tau_{\mathrm{ref},[0^{-\prime}]}$ that is needed for Eq. (30). The implementation of the collective RC would also be more cumbersome with periodic boundaries than when the RC just depends on a single target permeant. Instead, we can recover the advantages of the collective RC by the introduction of the new MC moves discussed in Sec. VI.

Still, also in the target permeant approach, care has to be taken with periodic boundaries when the relative position of the target permeant with respect to the membrane has to be determined. As a start, we obey the convention $\lambda_{-1} < \lambda_0 < \lambda_n$.

In addition, the RC should change continuously along the sequence of time slices of a complete path, i.e., it should not suddenly jump by a value equal to L_z which could lead to untrue transitions between states. Trajectories end and start with a time slice outside the $[\lambda_{-1}, \lambda_n]$ region and the RC of these points define the states at which they start and end, so the 'jump-free' interval needs to be extended slightly beyond the $[\lambda_{-1}, \lambda_n]$ interval. Based on the above, the safest option is to allow the jump in the RC to occur in the mid-point between λ_n and the periodic image of λ_{-1} (at $\lambda_{-1} + L_z$). This yields for the final RC

$$\lambda = \begin{cases} z & \text{if } -\frac{L_z}{2} < z < \frac{\lambda_B + \lambda_{-1} + L_z}{2} \\ z - L_z & \text{if } \frac{\lambda_B + \lambda_{-1} + L_z}{2} \leqslant z \leqslant \frac{L_z}{2} \end{cases} \quad (34)$$

where $z \in [-L_z/2, L_z/2]$ is the z coordinate of the target permeant following the convention of Eq. (33).

The RC as a function of the target permeant's position within a periodic system is shown in Fig. 3. For NPT simulations with fluctuating box dimensions it might be convenient to define the RC relative to the box size along z: $\lambda^{\mathrm{NPT}} = \lambda/L_z$ where λ is still defined by Eq. (34). The only difference is that L_z is the instantaneous box length, which is a variable instead of a constant.
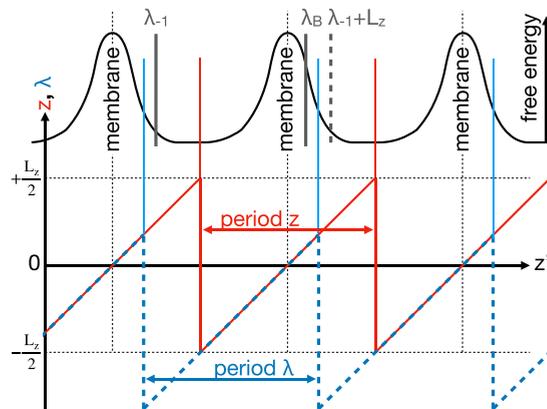
FIG. 3. Definition z and λ. Here z' is the unbounded coordinate of the target permeant. z is the same coordinate adjusted for the PBC such that it is 0 at the membrane center and restricted to $[-L_z/2, L_Z/2]$. The grey dashed line at $\lambda_{-1} + L_z$ is the mirror image of λ_{-1}. The mid-point between the grey dashed line and λ_B is located at $(\lambda_B + \lambda_{-1} + L_z)/2$ (thin blue line) and sets the switch for the λ definition such that λ can have values $\in [(\lambda_B + \lambda_{-1} - L_z)/2, (\lambda_B + \lambda_{-1} + L_z)/2]$.

VI. NEW MC MOVES IN PATH SPACE

The choice to select a single target permeant instead of a more collective RC and allowing the target permeant to cross the membrane in just one direction (left to right) can lead to a somewhat restrictive sampling speed in comparison with a more collective RC. In the next two sections, we show how to remove these restrictions by introducing two new MC moves for the $[0^{-\prime}]$ ensemble without the need to alter the definition of the RC or the setup of interfaces.

A. Target swap move

As discussed above, the RC is determined by the z coordinate of a single target permeant. The other permeants basically contribute to the environment around the target permeant like any of the other nonpermeating particles in the system. The occurrences of these particles crossing the membrane do not have to be counted as they are part of the natural fluctuations in the environment.

Especially when the membrane is not uniform containing different channels through which the permeants could transfer, it would still be advantageous to utilize the contributions of all the permeants. Some regions in the membrane that are easier to penetrate could be blocked by a nontagged permeant. Waiting for a swap through diffusion of both permeants within the bulk might take a long time. In order to speed up this process, we design a MC move in path space that allows for a swap without diffusion, but by simply reassigning the target.

The target swap move is explained in Fig. 4 and a stepwise description of the algorithm is given below.

(1) Assume the old path (upper panel in Fig. 4) has length $L^{(o)}$ (including start and end points) and is represented by time frames numbered from 1 to $L^{(o)}$.

(2) For each frame (1 to $L^{(o)}$), count the number of non-target permeants inside the $[\lambda_{-1}, \lambda_0]$ interval. The sum over
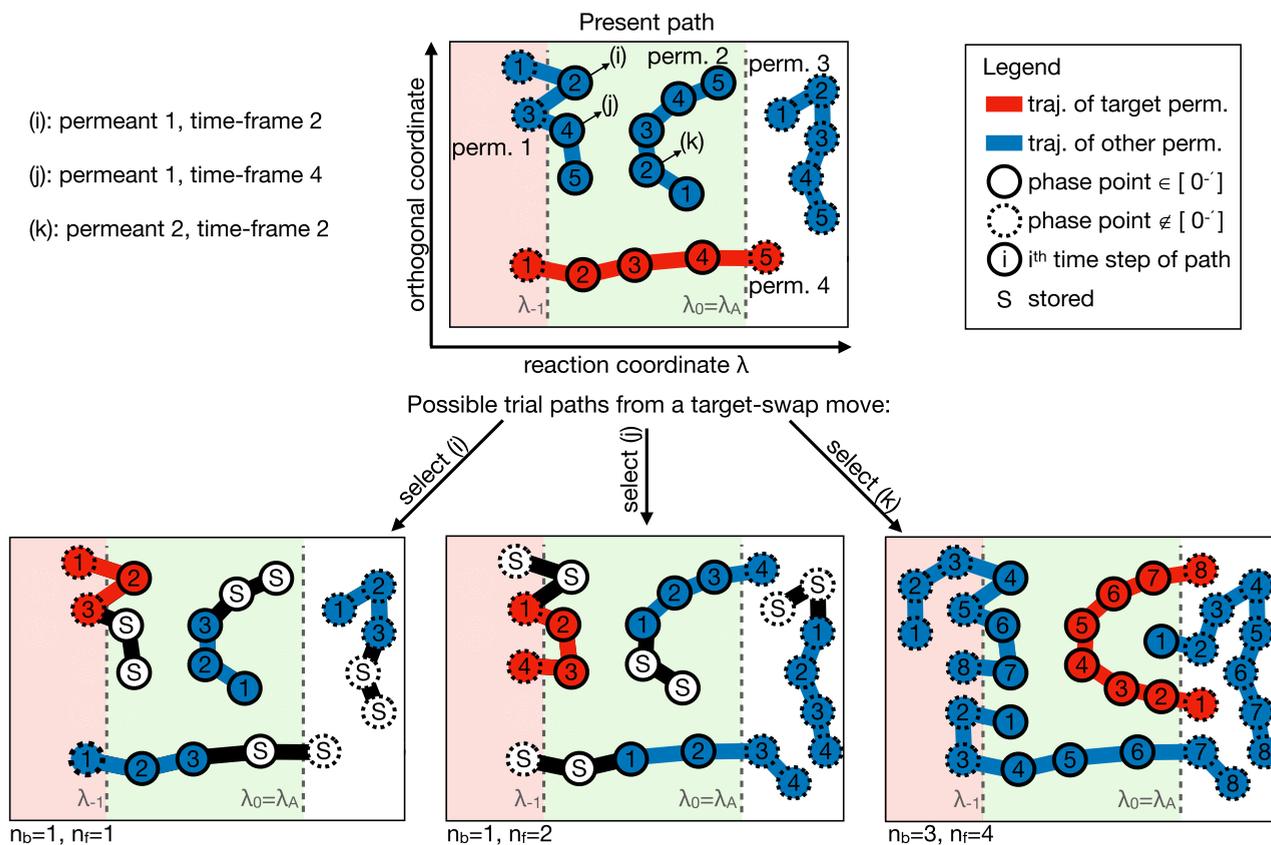
FIG. 4. Illustration of target swap move from an old path (shown in top panel) to a new path. Here, three possible new paths that could have been generated via the target swap move are shown in the lower panels. The numbers inside the circles indicate the frame numbers (time slices) of the path. Red circles represent the target permeant's positions at the different time frames, while blue circles represent the nontargeted permeants. The top panel shows a path of 5 time slices (including start and end points). The position of the target permeant at the first and last frame lies outside the $[\lambda_{-1}, \lambda_0]$ interval while the other time frames lie inside, in agreement with the path ensemble's criteria. All permeant/time-frames outside the interval are shown by dashed contour lines and cannot be selected in the target swap moves. Possible selectable permeant/time-frames are: time frames 2, 4, 5 of permeant 1 and time frames 1-5 of permeant 2. Permeant 3 is outside the interval at all time frames. The present target, permeant 4, cannot be selected. Each of these permeanant/time-frames have an equal probability to be selected, i.e., with a chance 1/8 in this example. Three possible selections are indicated by (i), (j), and (k) in the top figure and the resulting new trial paths are shown in the lower panels. After a permeant/time frame is selected, the corresponding permeant is the new target and the path is either lengthened or shortened by going backward and forward in time starting from the selected time frame until the new target has a frame outside the interval. Deleted time frames can be temporarily stored (indicated by "S") such that they could be reused if the next MC move requires extending the path.

all frames will be called $Z_t^{(o \to n)}$. For example, in Fig. 4 top, we have 5 frames to sum over, of which permeant 1 is 3 times and permeant 2 is 5 times inside the interval, leading to $Z_t^{(o \to n)} = 8$.

(3) Pick a random integer i from 1 to $Z_t^{(o \to n)}$.

(4) Select the permeant/time-frame combination corresponding to the i-th count at step 2. We will call this the new target permeant and the frame index at which this count was registered is from now on called j.

(5) Starting from time slice j, the path is *followed* (as detailed below) backward until we detect a frame in which the new target has a position outside the $[\lambda_{-1} : \lambda_0]$ interval. We call the number of backward steps n_b.

(6) Starting from time slice j, the path is *followed* (as detailed below) forward until we detect a frame in which the new target has a position outside the $[\lambda_{-1} : \lambda_0]$ interval. We call the number of forward steps n_f.

(7) The new path length is $n_b + n_f + 1$ which starts with frame index $j - n_b$ (which can be negative) and ends with frame index $j + n_f$. Renumber all frame indices by adding $-j + n_b + 1$ to each frame index. The new indices now run from 1 to $n_b + n_f + 1$.

(8) Compute $Z_t^{(n \to o)}$ just as in step 2, but using the new time region and new target permeant. For example, for the three bottom panels in Fig. 4, it is 5, 5, and 8 (from left to right).

(9) Compute the number of selectable permeant/time-frames $n_s^{(n)}$ [see Eq. (35) below] from which the same new trajectory can be obtained from the old path, and the number of selectable permeant/time frames $n_s^{(o)}$ at the new path from which the old path could be reobtained [see Eq. (36) below].

(10) Accept the move with a probability [see Eq. (37) below]: $P_{\text{acc}} = \min\left(1, \dfrac{n_s^{(o)}}{n_s^{(n)}} \dfrac{Z_t^{(o \to n)}}{Z_t^{(n \to o)}}\right)$.

At steps 4 and 5, the path is followed backward and forward in time starting from the selected time frame. In order to minimize the number of expensive force evaluations, this is done via one of the three possibilities that are listed here in order of preference. (1) Take the previous/next time slice from the old path, (2) if all time slices of the old path along a time direction are used, check for possible stored time slices in that time direction, and (3) create a new time slice by an actual MD step if no reusable time slice is available at (1) or (2).

At step 9, the calculation of $n_s^{(n)}$ and $n_s^{(o)}$ proceeds as follows. $n_s^{(n)}$ equals the number of selectable permeant/time frames combinations by which the same new path can be obtained. The way how it was actually generated, by selecting the new target permeant and time frame j, is one of the possible realisations. However, the exact same path might be generated by selecting time frames earlier and/or later. To clarify this, we discuss $n_s^{(n)}$ for the three cases shown in the bottom panels of Fig. 4. In the bottom-left panel, the new path can only be obtained by selecting time slice number $j = 2$ and consequently $n_s^{(n)} = 1$. The path in the bottom-middle panel can be obtained in two ways: selecting permeant 1 and either time slice $j = 4$ or $j = 5$ (renumbered as 2 and 3). Hence $n_s^{(n)} = 2$ in this case. Finally, the path in the bottom-right panel could be obtained by any of the 5 time slices when permeant 2 is selected ($n_s^{(n)} = 5$).

An expression for $n_s^{(n)}$ and $n_s^{(o)}$ is now derived. The number of time slices earlier than j that, if selected, would result into the same path, is restricted by either the old path, $j - 1$ time slices, or the new path, $n_b - 1$ backward steps (the last backward step is outside the $[\lambda_{-1}, \lambda_n]$ interval and cannot be selected). So this gives a contribution $\min(j - 1, n_b - 1)$ to $n_s^{(n)}$. Forward in time these numbers are $L^{(o)} - j$ and $n_f - 1$ for restriction by either the old path or the new path, respectively. This yields a contribution of $\min(L^{(o)} - j, n_f - 1)$. Including the time slice j itself, this gives

$$n_s^{(n)} = \min(j - 1, n_b - 1) + \min(L^{(o)} - j, n_f - 1) + 1$$
$$= \min(j, n_b) + \min(L^{(o)} - j, n_f - 1). \quad (35)$$

For the reverse move, i.e., selecting the old path from the new one, the same reasoning applies with the roles of the new and old paths switched. Hence, $j - 1$ and $n_b - 1$ are replaced by n_b and $j - 2$, respectively, when computing the selectable time slices before j. Further, $L^{(o)} - j$ and $n_f - 1$ are replaced by, respectively, n_f and $L^{(o)} - j - 1$. This gives for $n_s^{(o)}$

$$n_s^{(o)} = \min(n_b, j - 2) + \min(n_f, L^{(o)} - j - 1) + 1$$
$$= \min(n_b + 1, j - 1) + \min(n_f, L^{(o)} - j - 1). \quad (36)$$

Then, using Metropolis-Hastings rule [68,69], the acceptance probability can be written as

$$P_{\text{acc}}(o \to n) = \min\left(1, \frac{p(n)P_{\text{gen}}(n \to o)}{p(o)P_{\text{gen}}(o \to n)}\right)$$
$$= \min\left(1, \frac{n_s^{(o)}/Z_t^{(n \to o)}}{n_s^{(n)}/Z_t^{(o \to n)}}\right), \quad (37)$$

where $p(o)$ and $p(n)$ are the probabilities of the old and new path, respectively, and $P_{\text{gen}}(X \to X')$ is the generation probability to generate path X' from X. As the identity of the target has no effect on the path probabilities, the probabilities $p(o)$ and $p(n)$ are essentially the same except for stochastic force terms related to extending or shortening of the path. However, these terms cancel in Eq. (37) as they are also part of the generation probabilities [54]. The only remaining terms that do not cancel are, hence, the selection probabilities for selecting the permeant/time slice.

B. Mirror move

In the case that the membrane is symmetric, transitions through the membrane from left to right and from right to left are statistically indistinguishable within an equilibrium sampling. It is then favorable to count transitions in both directions [10,12]. For instance, the direct counting method described in Sec. II uses this strategy to improve statistics.

One obvious way to include two-directional transitions could be achieved by defining the RC as the absolute distance $|z|$ between the target permeant and the center of the solvent slab (see also Ref. [67]). However, as explained in Sec. V, this can lead to an overlap in the z-coordinate space. While this could still be solved by letting the RC value depend on the history of the path, i.e., the solvent slab's periodic image to be considered is determined by the minimum distance image at the start of the path, this becomes problematic when the path's history is not yet fully determined. For instance, this is the case when a shooting move is carried out or when the target swap move implies that some backward in time integration is required.

The mirror move in path space (see Fig. 5) achieves the same versatility of the two-directional approach in the counting method without having the problems discussed above. The mirror move in $[0^{-\prime}]$, which is always accepted, mirrors the whole system with respect to the membrane center. The z coordinates of every particle are mirrored and the z-component of the velocities are multiplied with -1. Because of the periodicity, this mirror move is equivalent to mirroring the whole system with respect to the midpoint between the membrane and its periodic image, which is, loosely speaking, the midpoint of the solvent slab. By construction, this midpoint of the solvent slab lies in the middle of the $[\lambda_{-1}, \lambda_0]$ interval related to $[0^{-\prime}]$.

The mirror move swaps the roles of the λ_0 and λ_{-1} interfaces. This requires that λ_0 and λ_{-1} are placed at the same distance from the mirror plane. To achieve this, consider the membrane's center-of-mass position at $z = z_{\text{mem}}$ and its left periodic image at $z = z_{\text{mem}} - L_z$. The distance between λ_0 and membrane should equal the distance between λ_{-1} and this left periodic image membrane. In other words, λ_{-1} should be placed such that

$$z_{\text{mem}} - \lambda_0 = \lambda_{-1} - (z_{\text{mem}} - L_z). \quad (38)$$

Equivalently, the interfaces λ_{-1} and λ_0 should have the same distance to the midpoint of the solvent slab at $z = z_{\text{mem}} - L_z/2$. Since we applied the periodic coordinates z as defined in Eq. (33), we have $z_{\text{mem}} = 0$ and therefore $\lambda_{-1} = -(\lambda_0 + L_z)$.

Given the positioning of λ_{-1} and λ_0 according to Eq. (38), the mirror move simply mirrors the z coordinates of every particle in the system at every time slice with respect to the
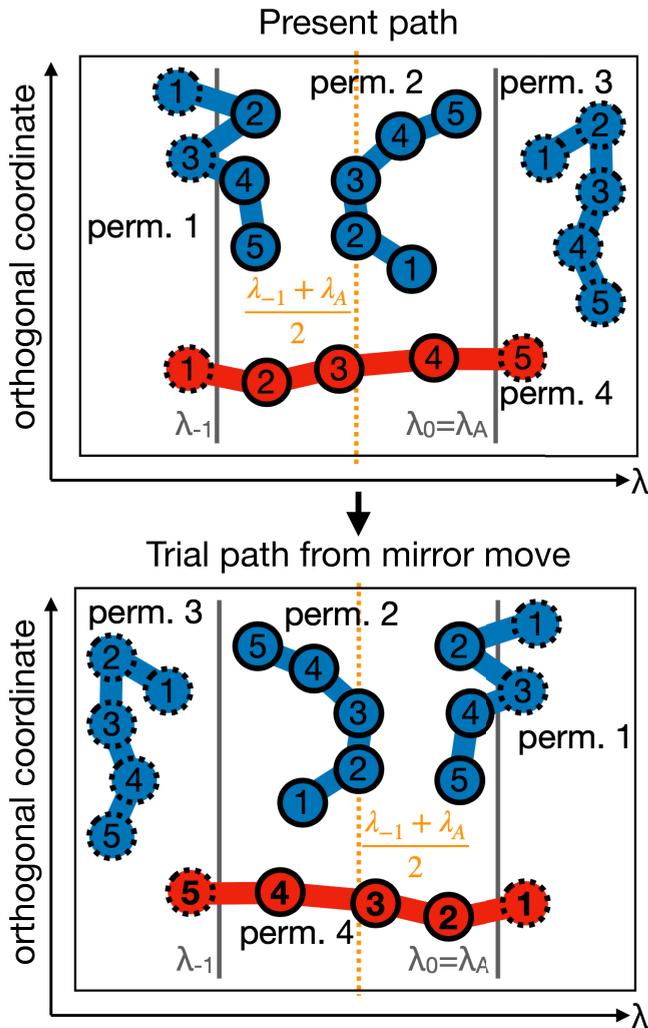
FIG. 5. Illustration of the mirror move. Numbers, line style and color scheme are the same as in Fig. 4. The move requires a specific definition of the λ_{-1} interface that must be symmetrically positioned with respect to λ_A, implying that these are equidistant to the center of the solvent slab. The top panel shows the old path. The bottom panel shows the new path in which the positions for each time slice are inverted in the mirror plane centered at the solvent slab (orange dashed line at $(\lambda_{-1} + \lambda_A)/2$). The mirroring is applied to all particles in the system, not only the permeants.

plane $z = -L_z/2$. Hence, it proceeds according to the following steps.

(1) Let k run over all time slices of the path. This includes the start and end point of the path and might include stored time slices (see Sec. VI A).

(2) For the given time slice, let j run over all point-particles in the system (atoms/coarse-grained particles describing the permeants, nonpermeating solvent particles, membrane particles, ...).

(3) Consider the z' coordinate of permeant j and its velocity component at discrete time step k: $z'_j(k)$ and $v_{z,j}(k)$.

(4) Mirror operation: replace $z'_j(k)$ with $2z'_{\text{mem.}} - z'_j(k) - L_z$ and replace $v_{z,j}(k)$ with $-v_{z,j}(k)$.

(5) Accept the new path. If the old path started/ended at λ_{-1} then the new path will start/end at λ_0 and vice versa.

For code-technical reasons we implemented a slightly different approach where we did not alter the coordinates or velocities, but instead the definition of the RC function [Eq. (34) with z replaced by $-z$]. The system was hence assigned an additional flag which indicates whether Eq. (34) has to be used with the plain z coordinate of the target permeant or with $-z$. This pragmatic choice made it easier to use PYRETIS [57,58] with external MD engines, as these might have very different ways of altering the coordinates and velocities. The flag is also exchanged in the replica exchange moves of the RETIS algorithm.

The new moves are only implemented for the $[0^{-\prime}]$ ensemble. For the mirror move, this is because this is the only ensemble where paths can start at both the left or the right hand side. In addition, the target swap move is not expected to give a high acceptance for the $[i^+]$ ensemble when crossing λ_i is a rare event.

VII. NUMERICAL RESULTS

The theoretical derivation for the permeability calculation from RETIS has been implemented in the python based open-source code PYRETIS [57,58]. First, a one-dimensional toy system was constructed where a Langevin, Brownian, or deterministic Newtonian particle permeates through a medium with or without barrier. For some limiting cases, an analytical expression for the permeability is available, which can serve as a validation of the new RETIS permeability formula. Second, a two-dimensional membrane was simulated with periodic boundary conditions, where permeants can pass the membrane through two different permeation channels. This last system is used to illustrate efficiency of the new Monte Carlo moves.

A. One-dimensional system setup

For simplicity, the membrane is located symmetrically around $z = 0$, in the region $|z| < a$, with $h = 2a$ as the membrane height. The effect of the membrane is modeled by an external cosine-shaped potential that acts on a single permeant particle,

$$V(z) = \begin{cases} \frac{1}{2}V_0\left(\cos\frac{\pi z}{a} + 1\right), & |z| \leqslant a \\ 0, & a < |z| \leqslant b \\ \frac{1}{2}k_{\text{harm}}(|z| - b)^2, & |z| > b \end{cases} \quad (39)$$

Here, V_0 is barrier height. This membrane model ensures that the force on the particle is continuous at the membrane borders $z = \pm a$. The harmonic potential $\frac{1}{2}k_{\text{harm}}(|z| - b)^2$ is added only to allow the system also to be studied by a reference simulation without λ_{-1} interface. In most simulation setups, the k_{harm} parameter is set to zero, which reflects the real physical situation for a permeation system that is unbounded at either side of the membrane. The dynamics of the permeant is governed by either Langevin dynamics, Brownian motion, or deterministic dynamics.

In the Langevin dynamics, the permeant experiences both friction, inertia effects, and random collisions with a degree that is controlled by the friction parameter γ. The friction constant γ (unit 1/time) of the particle relates to the particle's diffusion constant as $D = k_B T/(m\gamma)$. Note that some other

textbooks use a friction coefficient with unit mass/time. This refers to an alternative definition of the friction coefficient $\tilde{\gamma} = m\gamma$. The two other types of dynamics can be viewed as limiting cases of the Langevin dynamics.

A Brownian particle propagates in discrete steps without memory under influence of a random Gaussian force and the force $-dV/dz$ exerted by the potential. It can be considered as the overdamped limit of the Langevin particle, when $\gamma \to \infty$.

A Newtonian particle on the other hand propagates deterministically in time according to the equations of Newton. It has inertia but undergoes no friction nor random collisions, and energy is conserved. In the RETIS algorithm, the effect of temperature is then only present in the MC moves when a new constant-energy path is created from an old constant-energy path. The detailed balance MC procedure allows the energy to change between paths such that the overall path ensembles are canonically distributed. This simulation set up reflects a system that is so weakly coupled to a thermostat that the dynamics of a single crossing basically occurs at a constant energy (NVE, microcanonical). However, at the much longer timescale between crossing events, the energy could change. The Newtonian particle can be seen as the low friction limit of the Langevin particle, when $\gamma \to 0$.

The driver for solving the equations of motion was the internal MD-engine of PYRETIS in our toy system [57]. It is also possible to let PYRETIS manage the path ensembles book keeping, while it calls simulation programs, such as GROMACS, OPENMM, or LAMMPS, to drive the molecular dynamics [58]. Throughout Sec. VII, reduced units are used in which mass m, the Boltzmann constant k_B, and temperature T are equal to unity ($m = k_B = T = 1$ in reduced dimensionless units). Potentially, several physically realistic systems could be mapped on the model presented in Eq. (39) by tuning appropriate units of energy, length and mass.

We examined the model system Eq. (39) using the following parameters: $a = 0.1$ ($h = 0.2$), $k_{harm} = 100$ or 0, $m = 1$, $T = 1$, and integration time step $\Delta t = 0.002$. The friction coefficient γ had values 0.1, 5, 10, 20, 40, 60, 80, or 100. The Brownian dynamics propagates through configuration space at discrete steps with a displacement that is governed by a step-size parameter. This step-size parameter can be associated to a Δt in an equivalent Langevin simulation for a given mass and friction. In our case, the step-size parameter was set such that the time between steps was the same as Δt of the corresponding Langevin simulation with $\gamma = 100$.

In the RETIS simulations, the number of cycles was set to 20 000 (this is the number of MC moves in the path ensembles), the RETIS swapping move frequency was 0.1, the shooting frequency 0.45, and the time reversal move frequency 0.45. The position z was used as the order parameter λ. Three interfaces were used, $\lambda_0, \lambda_1, \lambda_2$, which were chosen at $\lambda = -0.1$, 0, and 0.1. The additional interface λ_{-1} was chosen at $\lambda = -0.2$ and varied to a few other locations in the calculations of Table I. The chosen reference region is $[-0.12, -0.1]$.

B. Analysis of permeability

For the flat potential membrane [$V_0 = 0$ in Eq. (39)], we examined the effect of the λ_{-1} interface versus a system that

TABLE I. Numerical results (reduced units) for permeability P and the three contributing factors ξ, $\tau_{ref}/\Delta z$, and $P_A(\lambda_B | \lambda_A)$. RETIS simulation of Langevin particle with $\gamma = 5$ and flat potential membrane ($V_0 = 0$). On first line, τ_{ref} refers to $[0^-]$, on other lines to $[0^{-\prime}]$. The reported error is based on block averaging and error propagation rules assuming independence of the different path ensemble simulations. As the latter assumption is not fully valid due to the replica exchange moves (Ref. [22]), we also estimated the error on P via 10 independent realizations, shown in the last column.

λ_{-1}	ξ		$\tau_{ref}/\Delta z$	$P_A(\lambda_B|\lambda_A)$	P	
-	1.000	(0%)	2.49 (1%)	0.662 (2%)	0.266	(2.2,2.0%)
-0.2	0.493	(0.3%)	1.22 (1%)	0.674 (2%)	0.274	(2.4,2.9%)
-0.15	0.504	(0.2%)	1.26 (1%)	0.641 (2%)	0.256	(2.1,2.4%)
-0.3	0.507	(1%)	1.28 (1%)	0.661 (2%)	0.261	(2.3,2.9%)

is bounded by a harmonic potential [$k_{harm} = 100$ in Eq. (39)] for Langevin dynamics with $\gamma = 5$. Table I shows that the change in $\tau_{ref}/\Delta z$ is compensated by the ξ factor. The permeability further remains fairly unaffected when changing the λ_{-1} position. The time spent per path per length is about 1.22 to 1.28 in $[0^{-\prime}]$, which lies close to the analytical value 1.253 of the deterministic particle (see Appendix B). The time spent per path per length would become larger when the friction increases. The crossing probability is lower than for the deterministic particle, which makes the Langevin permeability lower than the deterministic value 0.399 (see Appendix B).

Figure 6 compares the computed permeability for the potential Eq. (39) with $V_0 = 0$ (flat), 0.5, and 1 (cosine barrier membrane) for the three types of particle dynamics. In the Appendix we derived analytical expressions for P based on the Smoluchowski equation and Kramers' expression for Brownian dynamics and Langevin dynamics, respectively. These theoretical curves are shown in the same graphs. The analytical result for deterministic dynamics can be obtained by taking the limit of Kramers' expression for $\gamma \to 0$. The validity of these theoretical results relies on different kind of approximations. The Smoluchovski expression [Eq. (B1)] is reliable for high friction and low to high barriers, while Kramers' theoretical result [Eq. (B6)] is the reliable reference for high barriers and low to high friction. There is henceforth a blind spot in system parameter space: dynamics with low friction and low barriers is poorly described by both theories.

Indeed, consider the flat potential membrane ($V_0 = 0$) in Fig. 6(a). It shows good agreement between the theoretical Smoluchowski curve, $P = D/h$, and the simulated Langevin results for the large γ values. Also the computed values for Brownian and Langevin at $\gamma = 100$ agree, as Brownian dynamics can be seen as the high friction limit of Langevin dynamics. However, for low friction, the Smoluchowski curve and the numerical Langevin results deviate. The Kramers prediction of the permeability is zero for any $\gamma > 0$ and thus also fails to approximate the Langevin numerical results. Only the limiting case $\gamma \to 0$ with $\gamma^2/V_0 \to 0$ has a nonzero solution equal to $\sqrt{k_B T/(2\pi m)}$. The Kramers curve in Fig. 6(a) actually shows the theoretical results for $V_0 = 10^{-4}$ to visualize this limiting case. The deterministic $\sqrt{k_B T/(2\pi m)}$ limit to the permeability is indicated by the dashed horizontal line. Our numerical data on deterministic dynamics and the Langevin result with $\gamma = 0.1$ agree with this limit.
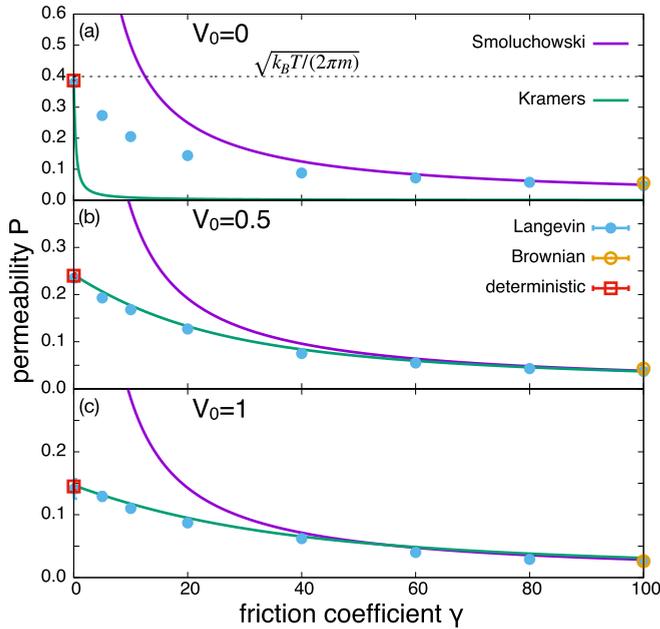
FIG. 6. Permeability P versus friction coefficient γ for $V_0 = 0$ (a), 0.5 (b), and 1 (c). The purple and green solid lines refer to the theoretical Smoluchowski expression Eq. (B1) and Kramers expression Eq. (B6), respectively. The Kramers curve in (a) actually shows the function for $V_0 = 10^{-4}$ instead of $V_0 = 0$ to visualize a limiting case of this expression in which $V_0 \to 0$, $\gamma \to 0$ and with $\gamma^2/V_0 \to 0$. The dashed horizontal line in the top panel shows the deterministic limit for the permeability in the flat potential. The filled blue circles refer to the numerical Langevin results. The open red square and open gold circle at $\gamma = 0$ and 100 show the results for deterministic dynamics and Brownian dynamics, respectively. Error bars based on a single standard deviation are mostly within symbol size.

Figured 6(b) and 6(c) show the numerical and theoretical curves for systems with a membrane barrier, $V_0 = 0.5$ and 1.0, respectively. For these membranes, the two theories agree in the large friction regime. In the low friction regime, the Smoluchowski expression is not a good approximation of the Langevin dynamics. The numerical Langevin simulations agree with the Kramers' curve for all values of γ. The deterministic and Brownian dynamics simulations also agree with Kramers' expression in the limiting $\gamma = 0$ and $\gamma = 100$, respectively.

C. Two-channel membrane system setup

A two-dimensional system is constructed that mimics a membrane barrier through which particles can permeate through two competing pathways. It could for instance represent a membrane with two transmembrane protein channels. Three noninteracting Langevin particles are subjected to the potential

$$V(y, z) = e^{-cz^2}\left(V_1 + A + A \sin \frac{2\pi y}{L_y} + B + B \cos \frac{4\pi y}{L_y}\right),$$
$$A = (V_2 - V_1)/2, \tag{40}$$
$$B = V_{max}/2 - V_1/4 - V_2/4.$$
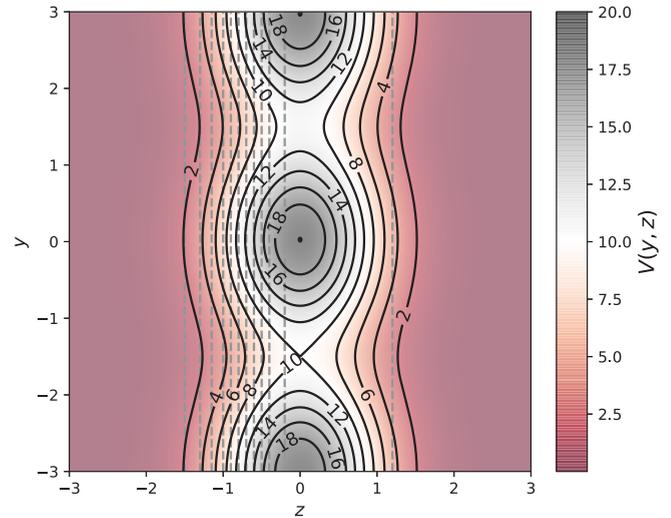
FIG. 7. The potential energy $V(y, z)$ represents a two-channel membrane. Interfaces $\lambda_0, \ldots, \lambda_{11}$ indicated with vertical dashed lines. Reduced units.

The membrane is located in the center of the unit cell around $z = 0$, while $V(y, z)$ is approximately zero far away from the center due to the factor e^{-cz^2} (see Fig. 7). Periodic boundary conditions are applied, where the system is made periodic in the z direction with a period $[-L_z/2, L_z/2]$ and in the y direction the period is L_y. Particles can permeate the membrane through two channels: one channel at about $y = -0.25L_y$ with barrier height V_1 and another channel at about $y = 0.25L_y$ with barrier height V_2. The maximum barrier height is V_{max}.

Reduced units are used as in the one-dimensional case. The parameters in our simulations were $V_1 = 10$, $V_2 = 11$, $V_{max} = 20$, $c = 1$, and $L_z = L_y = 6$. The PYRETIS simulations were run with three Langevin particles with settings $\Delta t = 0.02$, $\gamma = 5$, $T = 1$, and $m = 1$.

The order parameter is the reduced z coordinate of the target permeant: $\lambda = z_j$ if permeant j is tagged ($j = 1, 2, 3$). While the z coordinates lie in the interval $[-3, 3]$, the periodicity of λ is shifted to the interval $[-4.65, 1.35]$ (see Fig. 3). Twelve interfaces are located at $\lambda = -1.5, -1.3, -1.15, -1, -0.9, -0.8, -0.7, -0.6, -0.5, -0.4, -0.2$, and 1.2. The λ_{-1} interface is located at $\lambda = -4.5$. The reference region for $\tau_{ref,[0^{-\prime}]}$ is chosen as $[-3.2, -2.8]$. Three simulations are performed.

(1) TIS: In TIS, there are no swapping moves between the path ensembles. The MC moves are the shooting move and time reversal move, with equal frequency 0.5. Differently to standard TIS, the sampling of state A was done using the $[0^{-\prime}]$ path ensemble simulation and not via MD.

(2) RETIS: In standard RETIS, the swapping move of paths between the path ensembles is also allowed as an MC move. The swapping move frequency was 0.5, shooting move frequency 0.25, and time reversal move frequency 0.25.

(3) RETIS*: In the last simulation, RETIS is performed with swapping moves and the newly implemented MC moves in the $[0^{-\prime}]$ ensemble. The mirror plane was located at $\lambda = -3$, which is indeed midway between $\lambda_{-1} = -4.5$ and $\lambda_0 = -1.5$.

TABLE II. Numerical results (reduced units) for permeability P of 3 Langevin particles permeating through a two-channel membrane. Standard error from block averaging and error propagation between brackets.

	ξ		$\dfrac{\tau_{\mathrm{ref},[0^-{}']}}{\Delta z}$		$P_A(\lambda_B|\lambda_A)$ $\times 10^{-5}$		P $\times 10^{-6}$	
TIS	0.498	(1%)	5.66	(1%)	1.10	(12%)	0.97	(12%)
RETIS	0.540	(1%)	6.19	(1%)	1.20	(14%)	1.05	(14%)
RETIS*	0.507	(1%)	5.93	(1%)	1.23	(13%)	1.06	(13%)

Each of the three particles could be selected as the target permeant when performing a target swap move. The swapping move frequency was 0.5, time reversal move frequency 0.25, mirror move frequency 0.05, target swap move frequency 0.05, and shooting move frequency 0.15.

After an equilibration run of about 1600 MC moves, the analysis was performed based on a production run of 35 000 MC moves.

D. Two-channel membrane: analysis

Table II shows the permeability together with the calculated variables that enter in Eq. (30), ξ, $\tau_{\mathrm{ref},[0^-{}']}/\Delta z$, and $P_A(\lambda_B|\lambda_A)$. Note that the RETIS results on ξ and $\tau_{\mathrm{ref},[0^-{}']}/\Delta z$ are somewhat off compared to TIS and RETIS*. Naturally, for this symmetric system $\xi = 0.5$ is the exact result which agrees with TIS and RETIS* while RETIS is 8% too high. This is due to the relatively wide region between λ_0 and λ_{-1} and the lower frequency of shooting moves in the RETIS simulation compared to TIS. In RETIS, 50% of the moves are swapping moves (replica exchange moves between path ensembles), which are very useful to improve the sampling of the barrier region, but not necessarily help the exploration of the water phase. Both the swapping and time-reversal moves are unable to generate a $\lambda_{-1} \rightarrow \lambda_{-1}$ path from a $\lambda_0 \rightarrow \lambda_0$ path and vice versa. In RETIS*, the target swap move and the mirror move repair this weakness even if these moves only represent 10% of the executed MC cycles.

The crossing probabilities $P_A(\lambda_B|\lambda_A)$ and the permeabilities in Table II give quantitative good agreement within about 10%. Given a barrier of at least 10 k_BT this is a notable result. However, even more difficult challenges for a simulation method in this system, are (i) the ability to sample transitions through both channels, and (ii) to achieve this with the correct ratio. Since the channels' barriers only differ by 1 k_BT, both permeation routes are competing, but successful permeation transitions are expected to proceed via the lowest barrier channel in about 73% of the cases. Getting this ratio right is extremely challenging for any rare event method.

Figure 8 shows the distribution $P(y^*)$ of the orthogonal coordinate at the first crossing with λ_i, y^*, for different path ensembles, $[i^+]$, $i = 0, 1, \ldots, 10$. The crossing point y^* of a path is indicative of the channel visited by that path. The distributions show that for TIS, all y^* crossing points in the ensembles $[6^+]$ and higher are in the V_2 channel, while for RETIS and RETIS* the other channel is visited as well in all ensembles. This clearly demonstrates the deficit of the shooting move. The chance for this move to generate an acceptable
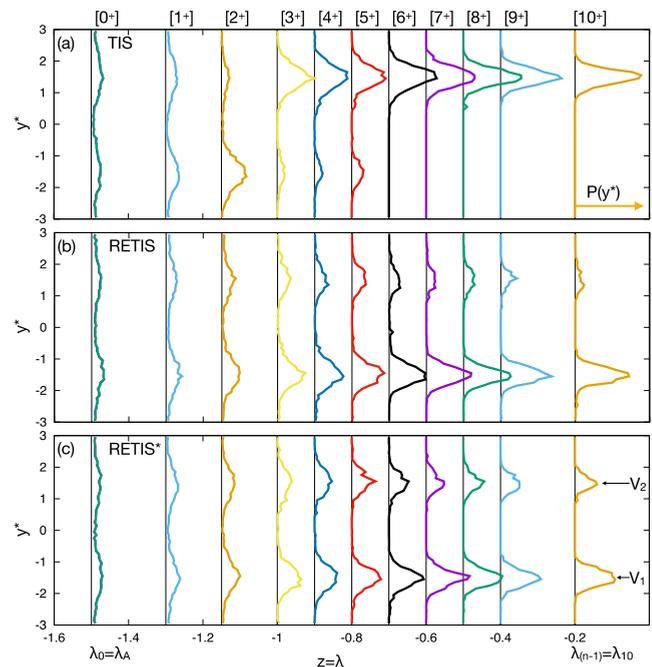
FIG. 8. Distributions $P(y^*)$ of first crossing point with λ_i along the y-direction for the different $[i^+]$ ensembles. Results are shown for (a) TIS, (b) RETIS, and (c) RETIS*. Note that the TIS simulations only sample the high-energy barrier V_2 for ensembles $[6^+]$ and higher.

path is highest when the shooting point is chosen on the barrier region, close to λ_i for ensemble $[i^+]$. However, switching between channels can practically only occur if the shooting is initiated from the well region. The fact that TIS got stuck in the high-energy channel, rather than the low-energy channel is purely accidental reflecting the memory of the initial path that was used to bootstrap the simulation. The TIS result (Table II) is lower than the other values, as one might expect based on its bias towards the high-barrier channel. Yet, since the crossing probability up to λ_6 is based on the progress through both channels, the TIS permeability is still rather close to the RETIS and RETIS* results.

Provided ergodic sampling, TIS and RETIS should be capable to sample nontrivial multiple-channel systems where splitting based methods, like FFS and AMS, would fail. An example of such a case is a system with two channels in which the lowest-barrier channel goes initially much steeper uphill than the channel with the higher barrier [62]. However, as is clear from Fig. 8, the TIS simulation is not ergodic since it is not able to switch between channels for ensembles $[6^+]$ and higher with just the shooting move.

For this academic model, this aspect could be repaired using nonlocal shooting moves in which not only the velocities, but also the configuration point is changed by a nonlocal displacement. Such a move, however, would have vanishingly low acceptance in a realistic condensed matter system as nearly every attempt will lead to a molecular overlap.

The RETIS and RETIS* simulations, however, are able to sample both channels and get the ratio between low- and high-barrier pathways at least qualitatively correct. The replica exchange swapping moves allow the exchange of paths be-

tween the different path ensembles. Consequently, sampling in $[0^{-\prime}]$ and $[i^+]$ with low i can facilitate the sampling in $[i^+]$ with high i. Especially the $[0^{-\prime}]$ and $[0^+]$ path ensembles are very effective to sample the direction that is orthogonal to the RC enabling the entrance of both channels. Figuratively speaking, this improved sampling of the orthogonal coordinate can then trickle down into the other ensembles by means of the swapping moves.

If we examine the height of the distributions in Fig. 8, we see that both RETIS and RETIS* predict that the majority of transitions will pass via the low-barrier channel. The RETIS simulation, however, seems to overestimate the preference of the V_1 channel, especially when the $[10^+]$ ensemble is considered. Integration of $\exp(-\beta V(y, \lambda_{10}))$ over y along positive and negative values indicates a 2.54 higher probability to be in the negative y-range. Even if these relative probabilities deviate a bit from the distribution of first crossing points y^*, this deviation is expected to be marginal for this model system.

To further analyze the effectiveness of the MC schemes we analyzed the number of channels switches observed in the path ensemble simulations. For path ensemble $[i^+]$ each path is assigned to channel V_1, channel V_2, or neither of the two, based on the first crossing point y^* with λ_i. It is assigned to belong to the V_1-channel if $-2.5 < y^* < -0.5$ and to the V_2 channel if $0.5 < y^* < 2.5$. A channel switch is counted when the MC move produces a V_2 channel path while the V_1 channel was most recently visited, and vice versa.

Figure 9(a) shows the number of channel switches that are calculated via this approach for different path ensembles. The TIS results are magnified by a factor 20 for visualization as this approach shows dramatically less channel switches than RETIS and RETIS*. This shows that replica exchange (swaps between path ensembles) is absolutely necessary for efficient sampling. From $[0^+]$ to $[5^+]$ ($\lambda_0 = -1.5$ to $\lambda_5 = -0.8$), the number of channel switches drops from 144 to 1. From $[6^+]$ to $[10^+]$ ($\lambda_6 = -0.7$, $\lambda_{10} = -0.2$), there is not a single channel switch observed.

The difference in channel switches seems negligible between RETIS and RETIS* up to $\lambda_i = -1$. After that point, RETIS* seems to produce significantly more switches. This is remarkable since the extra moves, the mirror and target swap move, are only executed in the $[0^{-\prime}]$ ensemble and its effect on the $[10^+]$ path ensemble is only indirect via the replica exchange moves. It requires at least 10 path ensemble swaps to process any information from $[0^{-\prime}]$ up to $[10^+]$. Still, the effect is most noticeable for the last seven path ensembles, but hardly before. Our conjecture is that the channel switches due to the mirror move and certainly due to the target swap move are more effective in decorrelating the ensemble. A channel switch is likely more effective if it enters the new channel along its central line and indeed this happens more often with the target swap move than with the shooting move. Also, the number of channel switches does not tell the full story. If two path ensemble $[i^+]$ and $[(i+1)^+]$ are at different channels and then make a lot of successful replica exchange moves solely between each other, this will yield a lot of channel switches. However, the effectiveness in decorrelating the sampling will be modest.

To examine this decorrelation, Fig. 9(b) plots the ratio of the number of paths in the V_1 and V_2 channels as a function
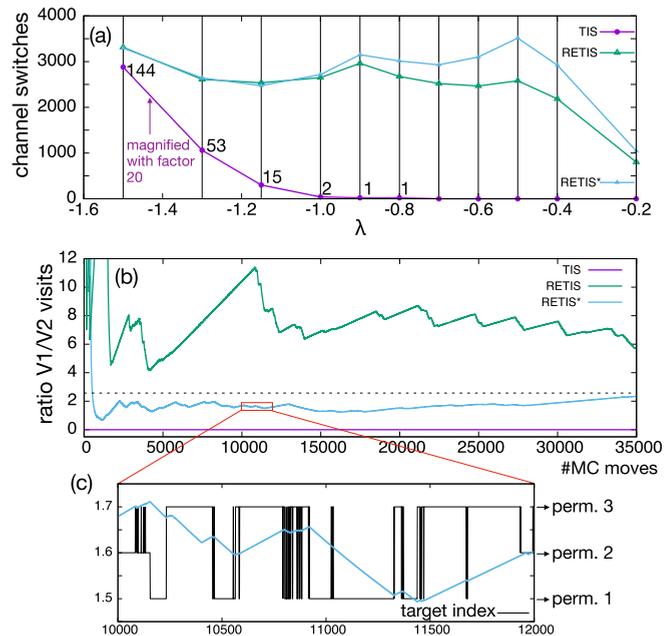
FIG. 9. (a) Number of channel switches observed in $[i^+]$ as a function of λ_i for TIS, RETIS, and RETIS*. The TIS results are magnified for visibility and shown with explicit numbers for the nonzero values. (b) Running average of ratio between generated V_1 and V_2 trajectories for the $[10^+]$ ensemble. Dashed line is the predicted value based on the numerical integration of $\exp(-\beta V(y, \lambda_{10}))$. (c) A zoom of the upper curve together with target index on the right vertical axis. It shows that channel switches (indicated by a change in direction of the blue line) occur often when the target permeant is reassigned.

of the number of MC moves in the $[10^+]$ ensemble. The TIS curve is a flat zero line as it is stuck in the V_2 channel for the full simulation. Comparing RETIS and RETIS*, we see that RETIS* converges much faster and approaches the predicted value of 2.54 that was obtained from the numerical integration. Figure 9(c) shows a zoom of the curve together with the index for the target permeant. These running averages for RETIS and RETIS* have a typical sawtooth shape though the latter is able to flip much more frequently the slope of the curve. The zoom in Fig. 9(c) shows that such a flip often coincides with a change of the index for the target permeant. This indicates that the target swap move has a strong influence in the overall sampling even if it is only applied in the $[0^{-\prime}]$ ensemble.

VIII. CONCLUSION

In this work, we derived a formula for the permeability based on path sampling quantities that can be determined in a RETIS simulation. As the idealized permeation model represents a membrane inside an infinite solution, the RETIS path ensembles require an adaptation via the introduction of an additional interface prior to λ_0, called λ_{-1}, and a newly defined path ensemble $[0^{-\prime}]$ that replaces the $[0^-]$ path ensemble. The resulting approach is exact and does not depend on the positions of the interfaces including λ_{-1}. Their positions are therefore set to optimize efficiency.

In addition to this theoretical derivation, we introduce a few algorithmic developments such as a consistent way to define the reaction coordinate for permeability whenever periodic boundary conditions are applied and two additional MC moves that mainly operate in the new $[0^{-\prime}]$ path ensemble. One of these new MC moves is the mirror move which can be applied whenever the membrane is symmetric. The other MC move is the target swap move and can be used when more than one permeant is present in the simulation model system.

Our new theoretical formulation and algorithmic developments have been implemented in the open-source PYRETIS code [57,58], and it was successfully tested on a one-dimensional Langevin system for which analytical results exist. After this, a challenging two-dimensional model membrane with two competing permeation channels was simulated to test the effectiveness of the new MC moves. These simulations show that the replica exchange moves are essential to simulate this system as the plain TIS method gets trapped inside a single channel. The inclusion of the two new MC moves considerably improves the sampling efficiency even further, as is clear when inspecting the relative transmission through the two channels. This noticeable difference is surprising given the fact that the new moves only operate in the $[0^{-\prime}]$ ensemble and it takes at least 10 replica exchange moves to transfer the effect of these moves up to the last path ensemble $[10^+]$. Still, the direct relation between the improved efficiency and the new MC moves was demonstrated by a correlation between the channel-switches and changes of the target permeant's identity.

The theoretical derivation in this paper is valid for all kinds of microscopically reversible dynamics (e.g., deterministic Newtonian dynamics, Langevin, Brownian, Nosé-Hoover, etc.). Besides the standard ergodicity hypothesis, it does not rely on any further assumption nor approximations. This implies that our approach will in principle give the same value for the permeability as the direct counting method based on brute force simulations, but orders of magnitude faster.

Our approach has the great advantage that the Markovian assumption of memoryless hopping between interfaces (see Sec. I) is not needed like the approaches based on milestoning [42–45]. Especially for large molecules the permeation process is often driven by nontrivial membrane fluctuations such that the projected dynamics on a one-dimensional coordinate gets a memory dependent character. Since RETIS is inherently non-Markovian in its description, it allows a much broader range of applications. An interesting route of thought could be the combination of our facilitating RETIS framework with the high-throughput methods that efficiently scan chemical space [70]

On the other hand, a milestoning type approach avoids the creation of full transition trajectories which can be computationally demanding when the transit time through the membrane is long. In this case, PPTIS [51] could be an interesting alternative. PPTIS avoids the sampling of complete transition paths like milestoning, but still maintains some of the history dependence. Alternatively, the exact non-Markovian character could be kept by alternating between short and long paths by means of stone skipping/web throwing [71]. Both PPTIS and stone skipping/web throwing can straightforwardly be implemented in our theoretical framework.

In conclusion, our permeability method presents a model-free approach for the computation of permeability and it is expected to become a valid standard method when membrane crossings are rare events.

APPENDIX A: ENSEMBLE AVERAGES IN TRAJECTORY SPACE

At several instances in our article [e.g., Eqs. (6), (10), and (16)], we refer to phase space ensemble averages with the remark that these should actually be viewed as an average over "trajectory phase points." Even if this point has mainly conceptual importance, we will outline here its mathematical interpretation since it is yet underreported in literature. One convenient way is to refer to path space ensemble averages instead of phase space ensemble averages [72]. Here, the path $X = \{x_0, x_1, \dots, x_L\}$ can be viewed as a *"chain of states"* [73] with x_i the phase point that is visited after i MD steps, at time $t = i\Delta t$ with Δt the time step. From this, the path probability follows as

$$P[X] = \rho(x_0) \prod_{i=0}^{L-1} p(x_i \to x_{i+1}), \qquad (A1)$$

where $\rho(x_0)$ is the probability density of the initial state (x_0 at $t = 0$) of the path, and $p(x_i \to x_{i+1})$ are the single time step transition probabilities. The latter are dependent on the type of dynamics. Mostly, we assume that $\rho(\cdot)$ is the equilibrium phase space density given by the Boltzmann distribution: $\rho(x) \propto \exp(-\beta E(x))$ with $E(x)$ the total energy of phase point x. Actually, while $P[X]$ and $p(x_i \to x_{i+1})$ are commonly referred to as a type of probabilities, it would have been more accurate to call these probability densities as well.

Now, by expressing an observable f as a functional of X, the path ensemble average can be formally written as

$$\langle f \rangle = \int dX \, P[X] f[X] \text{ with } dX = \prod_{i=0}^{L} dx_i \qquad (A2)$$

As we assume microscopically time-reversible dynamics, we can write [72]

$$\rho(x_i) p(x_i \to x_{i+1}) = p(\bar{x}_{i+1} \to \bar{x}_i) \rho(x_{i+1}), \qquad (A3)$$

where $\bar{x}$ refers to the momenta-reversed phase point: if $x = (r, v)$ with r the configuration and v the particles' velocities, then $\bar{x} = (r, -v)$. Here, it is also assumed that $\rho(x) = \rho(\bar{x})$ for any phase point x. Applying Eq. (A3) multiple times on Eq. (A1), allows us to write alternative expressions for the path probability [59]:

$$
\begin{aligned}
P[X] &= \rho(x_0) p(x_0 \to x_1) p(x_1 \to x_2) p(x_2 \to x_3) \dots \\
&= p(\bar{x}_1 \to \bar{x}_0) \rho(x_1) p(x_1 \to x_2) p(x_2 \to x_3) \dots \\
&= p(\bar{x}_1 \to \bar{x}_0) p(\bar{x}_2 \to \bar{x}_1) \rho(x_2) p(x_2 \to x_3) \dots \\
&= p(\bar{x}_1 \to \bar{x}_0) p(\bar{x}_2 \to \bar{x}_1) p(\bar{x}_3 \to \bar{x}_2) \rho(x_3) \dots
\end{aligned}
$$

In the TIS and RETIS theoretical framework, the path concept is extended by including time slices *before* x_0. In this view, x_0 is considered the present state, the principle phase point, while x_i is a state in the future or in the past whenever i is, respectively, positive or negative. Hence, for a path $X = \{x_{-M}, x_{-M+1}, \dots, x_{-1}, x_0, x_1, \dots x_{L-1}, x_L, \}$, we can

write

$$P[X] = \rho(x_0)\left(\prod_{i=0}^{L-1} p(x_i \to x_{i+1})\right)\left(\prod_{i=0}^{M-1} p(\bar{x}_{-i} \to \bar{x}_{-i-1})\right).$$

(A4)

Using Eq. (A4), we can in principle redefine all phase space ensemble averages, in which one integrates over x, as path ensemble averages in which one integrates over x_0 and additional phase points $x_{i\neq 0}$ with both positive and negative index via Eq. (A2) with $dX = \prod_{i=-M}^{L} dx_i$. Whenever the value of f is instantaneously available from the present phase point x_0, the integrals over the additional phase points can be ignored, since they are unity.

Another way to generalize the ensemble average, which in some cases could be arguably more intuitive, can be derived from another perspective on the path object as stated by Crooks and Chandler [73]: "A stochastic trajectory can be defined by the chain of states that the system visits, but it can also be represented by the initial state and the set of random numbers, the noise history, that was used to generate the trajectory." [73]. As an example they show that the probability density of a one-dimensional Brownian dynamics path X consisting of L time slices can be written as

$$P[X] = P[\{x_0, x_1, \ldots, x_L\}] = \rho(x_0) \prod_{i=1}^{L} \frac{1}{\sqrt{2\pi\epsilon}} \exp(-\xi_i^2/2\epsilon),$$

where each ξ_i is a Gaussian random number of zero mean and ϵ variance, the stochastic force acting at time between $t = (i-1)\Delta t$ to $t = i\Delta t$. Here, we deliberately shifted the indexing of the noise terms from 1 to L instead of the original [73] indexing from 0 to $L-1$. The reason becomes clear when we introduce Eq. (A8).

Since the phase point of the system at $t = \Delta t$, x_1, is fully determined by the first phase point and first stochastic noise term, we can write $x_1 = \phi(x_0, \xi_1)$ with ϕ being the MD time-step integrator. Likewise, $x_2 = \phi(x_1, \xi_2) = \phi(\phi(x_0, \xi_1), \xi_2)$ etc. It is thus apparent that, when we add to x_0 all the information of the random noise sequence $\xi_1, \xi_2, \ldots, \xi_L$ to make an "extended phase point" or "trajectory phase point," $\tilde{x} = \{x_0, \xi_1, \ldots, \xi_L\} = \{x_0, \xi^L\}$, basically every property of the system between $t = 0$ and $L\Delta t$ becomes a function of $\tilde{x}$, as if the dynamics would be deterministic.

Henceforth, for a general type of stochastic dynamics that proceeds via random noises that are drawn from a distribution $p_\xi(\cdot)$, we can define phase space density of an extended phase point $\tilde{x}$ as

$$\rho(\tilde{x}) = P[X(\tilde{x})] = \rho(x_0) \prod_{i=1}^{L} p_\xi(\xi_i).$$

(A5)

So equivalently to Eq. (A2), by expressing an observable f as a function of an extended phasepoint $\tilde{x}$, we write its ensemble average as

$$\langle f \rangle = \int d\tilde{x}\, \rho(\tilde{x}) f(\tilde{x}) \text{ with}$$

$$d\tilde{x} = dx_0 \prod_{i=1}^{L} d\xi_i = dx_0 d\xi^L,$$

$$\rho(\tilde{x}) = \rho(x_0) \prod_{i=1}^{L} p_\xi(\xi_i) = \rho(x_0) p_\xi(\xi^L).$$

(A6)

This automatically becomes a standard phase space ensemble average with x instead of $\tilde{x}$ when f is not noise-dependent since all integrals over $d\xi_i$ become 1.

While the concept of an extended phase point is generally not explicitly referred to, it is often implicitly used. For instance, time-correlation functions are often casually introduced as $C(t) = \langle a(0)b(t) \rangle$ without being specific about the noise dependence. Based on Eqs. (A5) and (A6), we can rigorously define the ensemble average as an integral over extended phase space:

$$C(t) = \int \rho(\tilde{x}) a(x_0) b(x_L) d\tilde{x} \quad\quad (A7)$$

$$= \int \rho(x_0) p_\xi(\xi^L) a(x_0) b(x_L(x_0, \xi^L)) dx_0\, d\xi^L$$

with $L = t/\Delta t$,

where a and b are functions of the phase point of the system at the time under consideration, at $t = 0$ and $L\Delta t$, respectively. The absolute timescale is irrelevant here since we generally assume we are at an equilibrium distribution at $t = 0$ and the dynamics conserves this distribution, i.e $\langle a(0)b(t) \rangle = \langle a(t')b(t+t') \rangle$ or any arbitrary t'. Hence, the correlation function $C(t)$ becomes an ensemble average $\langle a(\tilde{x})b(\tilde{x};t) \rangle$ where one just integrates over $\tilde{x}$ and b is parametrically dependent on t in addition to its dependence on $\tilde{x}$.

Comparing Eqs. (A1) and (A5), it is apparent that $p(x_i \to x_{i+1}) = p_\xi(\xi_{i+1})$ for stochastic dynamics with ξ_{i+1} being the noise that forces the dynamics to produce x_{i+1} from x_i; $x_{i+1} = \phi(x_i, \xi_{i+1})$. For deterministic dynamics, we can write $p(x_i \to x_{i+1}) = \delta(x_{i+1} - \phi(x_i))$. In addition, it is clear that the path interpretation and the extended phase point interpretation are equivalent; if one knows the initial phase point and the noise sequence, one knows the path X and vice versa.

As stated before, the TIS and RETIS theoretical framework requires the description of phase points before x_0. This means that the "noise history" term by Crooks and Chandler to denote $\xi_+^L = \{\xi_1, \xi_2, \ldots, \xi_L\}$ is now recoined as noise future while $\xi_-^M = \{\xi_{-1}, \xi_{-2}, \ldots, \xi_{-M}\}$ is the actual noise history or noise past.

Then, equivalent to Eq. (A4), we can define the phase space density of $\tilde{x} = \{\xi_{-M}, \ldots, \xi_{-1}, x_0, \xi_1, \ldots, \xi_L\}$ as

$$\rho(\tilde{x}) = P[X(\tilde{x})] = \rho(x_0)\left(\prod_{i=1}^{L} p_\xi(\xi_i)\right)\left(\prod_{i=1}^{M} p_\xi(\xi_{-i})\right)$$

$$= \rho(x_0) p_\xi(\xi_+^L) p_\xi(\xi_-^M),$$

(A8)

where the noise terms have a slightly different interpretations depending on the index being positive or negative. For $i > 0$, ξ_i is the noise needed for ϕ to produce x_i given x_{i-1}, while ξ_{-i} is the noise needed for ϕ to produce $\bar{x}_i$ given $\bar{x}_{-i+1}$: $\phi(x_{i-1}, \xi_i) = x_i$ and $\phi(\bar{x}_{-i}, \xi_i) = \bar{x}_i$. Hence, the history of the path X follows from the negative noise terms as: $x_{-1} = \overline{\phi(\overline{x_0}, \xi_{-1})}$, $x_{-2} = \overline{\phi(\overline{x_1}, \xi_{-2})}$, etc, again showing that there is a one-to-one relation between X and $\tilde{x}$.

Based on Eq. (A8), it is now possible to define the probability of overall state $\mathcal{A}$ as: $p_\mathcal{A} = \langle h_\mathcal{A} \rangle$. Here $h_\mathcal{A} = h_\mathcal{A}(\tilde{x}) = $

$h_\mathcal{A}(X)$ equals

$$h_\mathcal{A}(\tilde{x}) = \sum_{n=0}^{\infty} h_\mathcal{A}^{-n}(\tilde{x}) \tag{A9}$$

with

$$h_\mathcal{A}^{-n}(\tilde{x}) = h_A(x_{-n}) \prod_{i=0}^{n-1} (1 - h_A(x_{-i}) - h_B(x_{-i})) \tag{A10}$$

such that $h_\mathcal{A}^{-n}$ is a function of $(x_0, \xi_{-1}, \xi_{-2} \ldots \xi_{-n})$ which is a part of $\tilde{x}$. The product term is simply 1 if none of the points x_i with index $-n < i \leqslant 0$ is inside A or B. Otherwise it is 0. Further, $h_\mathcal{A}^0(\tilde{x})$ is simply $h_A(x_0)$. Likewise, we can define

$$h_\mathcal{B}(\tilde{x}) = \sum_{n=0}^{\infty} h_\mathcal{B}^{-n}(\tilde{x}) \text{ with} \tag{A11}$$

$$h_\mathcal{B}^{-n}(\tilde{x}) = h_B(x_{-n}) \prod_{i=0}^{n-1} (1 - h_A(x_{-i}) - h_B(x_{-i})).$$

In principle, we consider the path or random noise sequence to extend to infinite in both time directions [$L \to \infty$, $M \to \infty$ in Eq. (A8)]. However, as $h_\mathcal{A}^{-n}(\tilde{x})$ does not depend on ξ_i with $i > -1$ nor $i < n$, many noise integrals are simply 1 and therefore

$$p_\mathcal{A} = \langle h_\mathcal{A} \rangle = \int \rho(\tilde{x}) h_\mathcal{A}(\tilde{x}) d\tilde{x} \tag{A12}$$

$$= \int dx_0 \, \rho(x_0) \sum_{n=0}^{\infty} \int d\xi_-^n \, p_\xi(\xi_-^n) h_\mathcal{A}^{-n}(x_0, \xi_-^n)$$

with $d\xi_-^n = d\xi_{-1} d\xi_{-2} \cdots d\xi_{-n}$.

Now, suppose f is a function of phase space: $f = f(x)$. Then, the ensemble average of f does not require the integration of any noise terms

$$\langle f \rangle = \int dx \, f(x) \rho(x) \tag{A13}$$

though the conditional ensemble average $\langle f \rangle_\mathcal{A}$ does as

$$\langle f \rangle_\mathcal{A} = \frac{\langle f h_\mathcal{A} \rangle}{\langle h_\mathcal{A} \rangle} \tag{A14}$$

$$= \frac{\int dx_0 \, f(x_0) \rho(x_0) \sum_{n=0}^{\infty} \int d\xi_-^n \, p_\xi(\xi_-^n) h_\mathcal{A}^{-n}(x_0, \xi_-^n)}{\langle h_\mathcal{A} \rangle}.$$

The division used in Eq. (10) can now be understood with this extended phase point picture in mind as

$$p_\mathcal{A} \langle f \rangle_\mathcal{A} + p_\mathcal{B} \langle f \rangle_\mathcal{B} = \langle h_\mathcal{A} \rangle \langle f \rangle_\mathcal{A} + \langle h_\mathcal{B} \rangle \langle f \rangle_\mathcal{B}$$

$$= \int dx_0 \, f(x_0) \rho(x_0) \sum_{n=0}^{\infty}$$

$$\times \int d\xi_-^n \, p(\xi_-^n) \big[h_\mathcal{A}^{-n}(x_0, \xi_-^n) + h_\mathcal{B}^{-n}(x_0, \xi_-^n) \big]. \tag{A15}$$

If we consider the third line of Eq. (A15) separately, we can identify this, for a given phase point x_0, as the chance that the stochastic dynamics needed exactly n steps backward in time to move outside *no man's land*, i.e., enter either stable state A or B. Since we assume that no point x_0 can be trapped into no

man's land forever, the sum over n of this probability equals 1. Hence,

$$p_\mathcal{A} \langle f \rangle_\mathcal{A} + p_\mathcal{B} \langle f \rangle_\mathcal{B} = \int dx_0 \, f(x_0) \rho(x_0) = \langle f \rangle. \tag{A16}$$

Note that the we can use integration over x [Eq. (A13)] or x_0 [Eq. (A16)] interchangeably since $\rho(\cdot)$ refers to the equilibrium phase density that is time-invariant. So indeed, $\langle \ldots \rangle = p_\mathcal{A} \langle \ldots \rangle_\mathcal{A} + p_\mathcal{B} \langle \ldots \rangle_\mathcal{B}$ like stated in Eq. (10).

As a special case, we can take $f(x; z) = \delta(z - z_t)$ with z_t being the z coordinate of a specific particle (the target permeant). Here, z_t is a part of the system's phase point x that on its turn can be viewed as x_0 which a part of $\tilde{x}$ (recall that the phase space density is time-invariant). In addition, z is a parameter that specifies a reference region in configuration space. Note that the parametric dependence of f on z is not vanishing when taking the ensemble average since it is not a part of $\tilde{x}$ and therefore not integrated out. Therefore, we can write

$$\rho(z) = \langle \delta(z - z_t) \rangle = \langle f(x; z) \rangle \tag{A17}$$

$$= p_\mathcal{A} \langle f(\tilde{x}; z) \rangle_\mathcal{A} + p_\mathcal{B} \langle f(\tilde{x}; z) \rangle_\mathcal{B}$$

$$= p_\mathcal{A} \frac{\langle f(x_0; z) h_\mathcal{A}(\tilde{x}) \rangle}{\langle h_\mathcal{A} \rangle} + p_\mathcal{B} \frac{\langle f(x_0; z) h_\mathcal{B}(\tilde{x}) \rangle}{\langle h_\mathcal{B} \rangle}$$

$$= p_\mathcal{A} \frac{\langle \delta(z - z_t(x_0)) h_\mathcal{A}(\tilde{x}) \rangle}{\langle h_\mathcal{A} \rangle} + p_\mathcal{B} \frac{\langle \delta(z - z_t(x_0)) h_\mathcal{B}(\tilde{x}) \rangle}{\langle h_\mathcal{B} \rangle}.$$

All ensemble averages in Eq. (A17) are in principle integrals over $\tilde{x}$, though in the first line an integral over configuration space would be sufficient since the integrals over momenta and noise terms are unity. In the last line of Eq. (A17), the integrals need to be carried out on the principle phase point x_0 and the backward noise terms $\xi_{-1}, \xi_{-2}, \ldots$ The integrals over the forward noises $\xi_1, \xi_2, \ldots$ are still unity.

The delta-function is only nonzero whenever $z_t(x_0)$ equals z. This implies that if z is inside stable state A, then the product $\delta(z - z_t(x_0)) h_\mathcal{B}(\tilde{x})$ is by definition zero; if $z_t(x_0) = z \in A$ then $x_0 \in A$ and, therefore, $h_\mathcal{B}(\tilde{x}) = 0$. This explains the last equality of Eq. (16) where $z = z_\text{ref} \in A$.

Finally, for the rate in Eq. (6), the product of $h_\mathcal{A}(0)$ and $h_\mathcal{B}(\Delta t)$ should be evaluated. By defining the principle phase point to be x_0, we have $\tilde{x} = (\ldots, \xi_{-2}, \xi_{-1}, x_0, \xi_1, \ldots)$ and $x_1 = \phi(x_0, \xi_1)$ with ϕ the Δt time step integrator. Further, the product can only be nonzero if both $h_\mathcal{A}(\tilde{x})$ and $h_\mathcal{B}(x_1)$ are equal to 1, and $h_\mathcal{B}(\Delta t)$ can be replaced by $h_\mathcal{B}(\phi(x_0, \xi_1))$ in the product,

$$h_\mathcal{A}(0) h_\mathcal{B}(\Delta t) = h_\mathcal{A}(\ldots, \xi_{-1}, x_0, \xi_1, \ldots) h_\mathcal{B}(\phi(x_0, \xi_1))$$

$$= h_\mathcal{A}(\tilde{x}) h_\mathcal{B}(\phi(x_0, \xi_1)).$$

The ensemble average in Eq. (6) can be written as

$$k = \lim_{\Delta t \to 0} \frac{\langle h_\mathcal{A}(0) h_\mathcal{B}(\Delta t) \rangle}{\langle h_\mathcal{A} \rangle \, \Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{1}{\langle h_\mathcal{A} \rangle \, \Delta t} \int dx_0 \rho(x_0) \int d\xi_1 \, p_\xi(\xi_1) h_\mathcal{B}(\phi(x_0, \xi_1))$$

$$\times \sum_{n=0}^{\infty} \int d\xi_-^n \, p_\xi(\xi_-^n) h_\mathcal{A}^{-n}(x_0, \xi_-^n) \tag{A18}$$

where noise history and noise future appear within one equation. The limit $\Delta t \to 0$ only exists formally since Δt will be taken equal to the typical MD step in any practical case.

APPENDIX B: THEORETICAL PERMEABILITY FOR 1D TOY SYSTEM

In the 1D toy system, we can express the permeability in an analytical shape for the following three situations: for a Brownian particle (based on the Smoluchowski equation), for a Langevin particle crossing a high barrier (based on the Kramers equation), and for a deterministic Newtonian particle.

The Smoluchowski equation for a Brownian particle leads to the permeability expression in Eq. (3). In the one-dimensional case, the absence of any other degrees of freedom implies that the free energy $F(z)$ and the potential energy $V(z)$ are the same. Hence, for overdamped dynamics, we can derive a theoretical expression for P by inserting $V(z)$ of Eq. (39) as $F(z)$ into Eq. (3),

$$P = \frac{D}{h} \frac{e^{-\beta V_0/2}}{I_0(\beta V_0/2)}. \tag{B1}$$

Here, $h = 2a$ and $I_0(x) = (1/\pi) \int_0^\pi \exp(\cos\theta) d\theta$ is the 0th order modified Bessel function of the first kind. When $V_0 = 0$, then $I_0(0) = 1$, and the resulting flat potential yields $P = D/h$. For large V_0, the cosine barrier can be approximated by a second order Taylor expansion about $z = 0$, $F(z) = V(z) \approx V_0(1 - (\frac{\pi z}{h})^2)$ and it can be assumed that $\exp(-\beta V(z))$ rapidly decays when moving away from the membrane. This can be inserted into Eq. (3) and the integration boundaries can be moved from $\pm h/2$ to $\pm\infty$. Solving the resulting Gaussian integral yields an approximation of P for large V_0,

$$P = \frac{D}{h} \sqrt{\frac{\pi V_0}{k_B T}} e^{-\beta V_0}. \tag{B2}$$

An alternative approach to the Smoluchowski approach is to use Kramer's relation for the rate constant k instead. The permeability P is then obtained via Eq. (18) by first computing the rate k while assuming a hard wall at $z = -W$ that can be taken to infinite. Using a harmonic approximation and a high barrier assumption, this rate constant k can be written as [74]

$$k = \kappa \sqrt{\frac{k_B T}{2\pi m}} \frac{\exp(-\beta V(0))}{\int_{-W}^0 \exp(-\beta V(z)) \, dz}, \tag{B3}$$

where κ is the transmission coefficient that can be approximated using Kramers' relation

$$\kappa = \frac{1}{\omega_+} \left(-\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \omega_+^2} \right). \tag{B4}$$

Here, ω_+ is the frequency associated to the curvature at the top of the barrier: $\omega_+ = \sqrt{k_+/m}$ with $V(z) \approx V_0 - \frac{1}{2} k_+ z^2$. From the above Taylor expansion, we have $k_+ = 2V_0\pi^2/h^2$ and $\omega_+ = (\pi/h)\sqrt{2V_0/m}$.

The conditional probability appearing in Eq. (18) is expressed as

$$(\rho_{\text{ref}})_{\mathcal{A}} = \frac{1}{\int_{-W}^0 \exp(-\beta V(z)) \, dz} \tag{B5}$$

where we assumed that overall state $\mathcal{A}$ condition is statistically equivalent to the condition $z < 0$ for this case, which is a valid assumption for a high barrier.

Inserting Eq. (B4) in Eq. (B3) and inserting Eqs. (B3) and (B5) in Eq. (18) gives the permeability for a high barrier,

$$P = \kappa \sqrt{\frac{k_B T}{2\pi m}} e^{-\beta V_0} \tag{B6}$$

$$= \frac{h}{\pi} \sqrt{\frac{m}{2V_0}} \left(-\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + \frac{2V_0\pi^2}{mh^2}} \right) \sqrt{\frac{k_B T}{2\pi m}} e^{-\beta V_0}.$$

In the high friction limit where $\gamma \gg \omega_+$, Eq. (B4) reduces to $\kappa = \omega_+/\gamma$, and P for large V_0 becomes

$$P = \frac{\pi}{h\gamma} \sqrt{\frac{2V_0}{m}} \sqrt{\frac{k_B T}{2\pi m}} e^{-\beta V_0}$$

$$= \frac{k_B T}{h\gamma m} \sqrt{\frac{\pi V_0}{k_B T}} e^{-\beta V_0}. \tag{B7}$$

Since $D = k_B T/(m\gamma)$, this equation is equal to the Smoluchowski equation Eq. (B2) within the harmonic approximation for the high barrier. A Langevin particle with high friction is indeed well described by the overdamped dynamics of a Brownian particle.

In the low friction limit $\gamma \ll \omega_+$, Eq. (B4) reduces to $\kappa = 1$, and P in Eq. (B6) becomes, for any V_0,

$$P = \sqrt{\frac{k_B T}{2\pi m}} e^{-\beta V_0}. \tag{B8}$$

This friction-less limit is exactly the permeability of the deterministic particle. It can also be obtained from Eq. (24). Here, $\xi = 1/2$, since a deterministic particle in a flat free energy region either moves to the right (velocity positive), either to the left (velocity negative), which have equal Boltzmann probability. A particle moving to the right will reach the barrier top with a probability $\exp(-\beta V_0)$ and it will not recross, and therefore $P_A(\lambda_B | \lambda_A) = \exp(-\beta V_0)$. The time spent per path in $[0^{-'}]$ in a reference region of size Δz can be computed from the flux-weighted velocity distribution as

$$\tau_{\text{ref},[0^{-'}]} = \Delta z \sqrt{\pi \beta m/2}. \tag{B9}$$

Inserting these three factors into Eq. (24) gives Eq. (B8).

Let us recap the case of a flat potential membrane ($V_0 = 0$). For the Brownian particle, the permeability is $P = D/h$. If the particle has low friction $\gamma \to 0$, then $D \to \infty$, and the permeability diverges. For the Langevin particle, the high friction limit of P in Eq. (B7) based on Kramer's relation vanishes when $V_0 = 0$, which is not an adequate approximation of a flat potential's permeability. Nevertheless, again considering a Langevin particle and Kramer's equation, the low friction limit in Eq. (B8) converges to $P = \sqrt{k_B T/(2\pi m)}$, which is finite.

In conclusion, the two theoretical expressions for the one-dimensional case, Eq. (B1) based Smoluchowski and

Eq. (B6) based on Kramers, use respectively an over-damped assumption or a harmonic approximation to describe the top of the barrier. For high friction and low barriers, Eq. (B1) will be more accurate than Eq. (B6). For high barriers and low friction Eq. (B6) will prevail over Eq. (B1). In the case that both the friction and the barrier is high, both converge to the same value. In the case that both the friction and the barrier is low, neither Eq. (B1) nor Eq. (B6) will be accurate descriptions of a Langevin particle.

---

[1] E. Awoonor-Williams and C. N. Rowley, Molecular Simulation of Nonfacilitated Membrane Permeation, Biochim. Biophys. Acta, Biomembr. **1858**, 1627 (2016).

[2] P. Berben, A. Bauer-Brandl, M. Brandl, B. Faller, G. E. Flaten, A. C. Jacobsen, J. Brouwers, and P. Augustijns, Drug permeability profiling using cell-free permeation tools: Overview and applications, Eur. J. Pharm. Sciences **119**, 219 (2018).

[3] B. J. Bennion, N. A. Be, M. W. McNerney, V. Lao, E. M. Carlson, C. A. Valdez, M. A. Malfatti, H. A. Enright, T. H. Nguyen, F. C. Lightstone, and T. S. Carpenter, Predicting a drugs membrane permeability: A computational model validated with in vitro permeability assay data, J. Phys. Chem. B **121**, 5228 (2017).

[4] E. N. Petersen, H. W. Chung, A. Nayebosadri, and S. B. Hansen, Kinetic disruption of lipid rafts is a mechanosensor for phospholipase D, Nat. Commun. **7**, 13873 (2016).

[5] B. Smit and T. L. M. Maesen, Molecular simulations of zeolites: Adsorption, diffusion, and shape selectivity, Chem. Rev. **108**, 4125 (2008).

[6] A. Ghysels, S. L. C. Moors, K. Hemelsoet, K. De Wispelaere, M. Waroquier, G. Sastre, and V. Van Speybroeck, Shape-selective diffusion of olefins in 8-ring solid acid microporous zeolites, J. Phys. Chem. C **119**, 23721 (2015).

[7] E. L. Elson, Fluorescence correlation spectroscopy: Past, present, future, Biophys. J. **101**, 2855 (2011).

[8] M. Przybylo, A. Olzynska, S. Han, A. Ozyhar, and M. Langner, A fluorescence method for determining transport of charged compounds across lipid bilayer, Biophys. Chem. **129**, 120 (2007).

[9] W. K. Subczynski, M. Pasenkiewicz-Gierula, R. N. McElhaney, J. S. Hyde, and A. Kusumi, Molecular dynamics of 1-palmitoyl-2-oleoylphosphatidylcholine membranes containing transmembrane α-helical peptides with alternating leucine and alanine residues, Biochemistry **42**, 3939 (2003).

[10] R. M. Venable, A. Krämer, and R. W. Pastor, Molecular dynamics simulations of membrane permeability, Chem. Rev. **119**, 5954 (2019).

[11] R. J. Dotson, K. Smith, Bueche, G. Angles, and S. C. Pias, Influence of cholesterol on the oxygen permeability of membranes: Insight from atomistic simulations, Biophys. J. **112**, 2336 (2017).

[12] A. Krämer, A. Ghysels, E. Z. Wang, R. M. Venable, J. B. Klauda, B. Brooks, and R. W. Pastor, Membrane permeability of small molecules from unbiased molecular dynamics simulations, J. Chem. Phys. **153**, 124107 (2020).

[13] S. Marrink and H. J. C. Berendsen, Simulation of water transport through a lipid membrane, J. Phys. Chem. **98**, 4155 (1994).

[14] D. J. Bicout and A. Szabo, Electron transfer reaction dynamics in non-Debye solvents, J. Chem. Phys. **109**, 2325 (1998).

[15] G. Hummer, Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations, New J. Phys. **7**, 34 (2005).

[16] A. Ghysels, R. M. Venable, R. W. Pastor, and G. Hummer, Position-dependent diffusion tensors in anisotropic media from simulation: Oxygen transport in and through membranes, J. Chem. Theory Comput. **13**, 2962 (2017).

[17] Z. Yue, C. Li, G. A. Voth, and J. M. J. Swanson, Dynamic protonation dramatically affects the membrane permeability of drug-like molecules, J. Am. Chem. Soc. **141**, 13421 (2019).

[18] A. K. Faradjian and R. Elber, Computing time scales from reaction coordinates by milestoning, J. Chem. Phys. **120**, 10880 (2004).

[19] E. Vanden-Eijnden, M. Venturoli, G. Ciccotti, and R. Elber, On the assumptions underlying milestoning, J. Chem. Phys. **129**, 174102 (2008).

[20] B. Peters, *Reaction Rate Theory and Rare Events* (Elsevier, Amsterdam, Netherlands, 2017).

[21] T. van Erp, D. Moroni, and P. Bolhuis, A novel path sampling method for the calculation of rate constants, J. Chem. Phys. **118**, 7762 (2003).

[22] T. S. van Erp, Reaction Rate Calculation by Parallel Path Swapping, Phys. Rev. Lett. **98**, 268301 (2007).

[23] R. Cabriolu, K. M. S. Refsnes, P. G. Bolhuis, and T. S. van Erp, Foundations and latest advances in replica exchange transition interface sampling, J. Chem. Phys. **147**, 152722 (2017).

[24] I. I. Ivanov, G. E. Fedorov, R. A. Guskova, K. I. Ivanov, and A. B. Rubin, Permeability of lipid membranes to dioxygen, Biochem. Biophys. Res. Commun. **322**, 746 (2004).

[25] J. Widomska, M. Raguz, and W. K. Subczynski, Oxygen permeability of the lipid bilayer membrane made of calf lens lipids, BBA - Biomembranes **1768**, 2635 (2007).

[26] M. Möller, J. Lancaster, and A. Denicola, The interaction of reactive oxygen and nitrogen species with membranes, Curr. Top. Membr. **61**, 23 (2007).

[27] M. N. Möller, Q. Li, C. Chinnaraj, H. C. Cheung, J. R. Lancaster Jr., and A. Denicola, Solubility and diffusion of oxygen in phospholipid membranes, Biochim. Biophys. Acta **1858**, 2923 (2016).

[28] O. De Vos, T. Van Hecke, and A. Ghysels, Effect of chain unsaturation and temperature on oxygen diffusion through lipid membranes from simulations, Oxygen Transport to Tissue XL, Advances in Experimental Medicine and Biology **1072**, 399 (2018).

[29] A. Ghysels, A. Krämer, R. Venable, W. Teague, E. Lyman, K. Gawrisch, and R. W. Pastor, Permeability of membranes in the liquid ordered and liquid disordered phases, Nat. Commun. **10**, 5616 (2019).

[30] S. Davoudi and A. Ghysels, Sampling efficiency of the counting method for permeability calculations estimated with the inhomogeneous solubility-diffusion model, J. Chem. Phys. **154**, 054106 (2021).

[31] E. Sezgin, I. Levental, S. Mayor, and C. Eggeling, The mystery of membrane organization: composition, regulation and roles of lipid rafts, Nat. Rev. Mol. Cell Biol. **18**, 361 (2017).

[32] C. Dietrich, L. A. Bagatolli, Z. N. Volovyk, N. L. Thompson, M. Levi, K. Jacobson, and E. Gratton, Lipid rafts reconstituted in model membranes, Biophys. J. **80**, 1417 (2001).

[33] A. V. Samsonov, I. Mihalyov, and F. S. Cohen, Characterization of cholesterol-sphingomyelin domains and their dynamics in bilayer membranes, Biophys. J. **81**, 1486 (2001).

[34] O. De Vos, R. M. Venable, T. Van Hecke, G. Hummer, R. W. Pastor, and A. Ghysels, Membrane permeability: Characteristic times and lengths for oxygen and a simulation-based test of the inhomogeneous solubility-diffusion model, J. Chem. Theory Comput. **14**, 3811 (2018).

[35] G. Torrie and J. Valleau, Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella samping, J. Comput. Phys. **23**, 187 (1977).

[36] E. Darve and A. Pohorille, Calculating free energies using average force, J. Chem. Phys. **115**, 9169 (2001).

[37] J. Comer, C. Chipot, and F. D. González-Nilo, Calculating position-dependent diffusivity in biased molecular dynamics simulations, J. Chem. Theory Comput. **9**, 876 (2013).

[38] C. T. Lee, J. Comer, C. Herndon, N. Leung, A. Pavlova, R. V. Swift, C. Tung, C. N. Rowley, R. E. Amaro, C. Chipot, Y. Wang, and J. C. Gumbart, Simulation-Based Approaches for Determining Membrane Permeability of Small Compounds. J. Chem. Inf. Model. **56**, 721 (2016).

[39] M. Badaoui, A. Kells, C. Molteni, C. J. Dickson, V. Hornak, and E. Rosta, Calculating kinetic rates and membrane permeability from biased simulations, J. Phys. Chem. B **122**, 11571 (2018).

[40] A. Krämer, R. W. Pastor, and A. Ghysels, Membrane permeability of small molecules from biased molecular dynamics simulations (unpublished).

[41] J. Comer, K. Schulten, and C. Chipot, Calculation of lipid-bilayer permeabilities using an average force, J. Chem. Theory Comput. **10**, 554 (2014).

[42] A. E. Cardenas and R. Elber, Computational study of peptide permeation through membrane: Searching for hidden slow variables, Mol. Phys. **111**, 3565 (2013).

[43] A. E. Cardenas and R. Elber, Modeling kinetics and equilibrium of membranes with fields: Milestoning analysis and implication to permeation, J. Chem. Phys. **141**, 054101 (2014).

[44] A. Fathizadeh and R. Elber, Ion permeation through a phospholipid membrane: Transition state, path splitting, and calculation of permeability, J. Chem. Theory Comput. **1**, 720 (2019).

[45] L. W. Votapka, C. T. Lee, and R. E. Amaro, Two relations to estimate membrane permeability using milestoning, J. Phys. Chem. B **120**, 8606 (2016).

[46] A. M. Berezhkovskii and A. Szabo, Committors, first-passage times, fluxes, markov states, milestones, and all that, J. Chem. Phys. **150**, 054106 (2019).

[47] P. L. Geissler, C. Dellago, and D. Chandler, Kinetic pathways of ion pair dissociation in water, J. Phys. Chem. B **103**, 3706 (1999).

[48] Weinan E, W. Ren, and E. Vanden-Eijnden, Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes, Chem. Phys. Lett. **413**, 242 (2005).

[49] B. Peters, G. T. Beckham, and B. L. Trout, Extensions to the likelihood maximization approach for finding reaction coordinates, J. Chem. Phys. **127**, 034109 (2007).

[50] R. Best and G. Hummer, Reaction coordinates and rates from transition paths, Proc. Natl. Acad. Sci. U.S.A. **102**, 6732 (2005).

[51] D. Moroni, P. G. Bolhuis, and T. S. van Erp, Rate constant for diffusive processes by partial path sampling, J. Chem. Phys. **120**, 4055 (2004).

[52] R. J. Allen, P. B. Warren, and P. R. ten Wolde, Sampling Rare Switching Events in Biochemical Networks, Phys. Rev. Lett. **94**, 018104 (2005).

[53] D. Aristoff, T. Lelièvre, C. G. Mayne, and I. Teo, Adaptive multilevel splitting in molecular dynamics simulations, ESAIM: Proc. **48**, 215 (2015).

[54] T. S. van Erp, Dynamical rare event simulation techniques for equilibrium and nonequilibrium systems, *Kinetics and Thermodynamics of Multistep Nucleation and Self-Assembly in Nanoscale Materials* (John Wiley & Sons, Ltd, 2012), Chap. 2, pp. 27–60.

[55] D. W. H. Swenson, J.-H. Prinz, F. Noe, J. D. Chodera, and P. G. Bolhuis, OpenPathSampling: A Python Framework for Path Sampling Simulations. 1. Basics, J. Chem. Theory Comput. **15**, 813 (2019).

[56] D. W. H. Swenson, J.-H. Prinz, F. Noe, J. D. Chodera, and P. G. Bolhuis, OpenPathSampling: A python framework for path sampling simulations. 2. Building and customizing path ensembles and sample schemes, J. Chem. Theory Comput. **15**, 837 (2019).

[57] A. Lervik, E. Riccardi, and T. S. van Erp, PyRETIS: A well-done, medium-sized python library for rare events, J. Comput. Chem. **38**, 2439 (2017).

[58] E. Riccardi, A. Lervik, S. Roet, O. A. en, and T. S. van Erp, PyRETIS 2: An improbability drive for rare events, J. Comput. Chem. **41**, 370 (2020).

[59] T. S. van Erp and P. Bolhuis, Elaborating transition interface sampling methods, J. Comput. Phys. **205**, 157 (2005).

[60] B. Peters and B. L. Trout, Obtaining reaction coordinates by likelihood maximization, J. Chem. Phys. **125**, 054108 (2006).

[61] D. Earl and M. Deem, Parallel tempering: Theory, applications, and new perspectives, Phys. Chem. Chem. Phys. **7**, 3910 (2005).

[62] T. S. van Erp, Efficient path sampling on multiple reaction channels, Comput. Phys. Commun. **179**, 34 (2008).

[63] T. J. H. Vlugt and B. Smit, On the efficient sampling of pathways in the transition path ensemble, Phys. Chem. Comm. **4**, 11 (2001).

[64] M. Grunwald, E. Rabani, and C. Dellago, Mechanisms of the Wurtzite to Rocksalt Transformation in CdSe Nanocrystals, Phys. Rev. Lett. **96**, 255701 (2006).

[65] L. Verlet, Computer experiments on classical fluids. i. thermodynamical properties of lennard-jones molecules, Phys. Rev. **159**, 98 (1967).

[66] M. Moqadam, A. Lervik, E. Riccardi, V. Venkatraman, B. K. Alsberg, and T. S. van Erp, Local initiation conditions for water autoionization, Proc. Natl. Acad. Sci. U.S.A. **115**, E4569 (2018).

[67] E. Riccardi, A. Krämer, T. van Erp, and A. Ghysels, Permeation rates of oxygen transport through POPC membrane using replica exchange transition interface sampling, J. Phys. Chem. B **125**, 193 (2021).

[68] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. **21**, 1087 (1953).

[69] W. Hastings, Monte-Carlo sampling methhods using Markov chains and their applications, Biometrika **57**, 97 (1970).

[70] R. Roberto Menichetti, K. H. Kanekal, and T. Tristan Bereau, Drug-membrane permeability across chemical space, ACS Cent. Sci. **5**, 290 (2019).

[71] E. Riccarfdi, O. Dahlen, and T. S. van Erp, Fast decorrelating monte carlo moves for efficient path sampling, J. Phys. Chem. Lett. **8**, 4456 (2017).

[72] C. Dellago, P. G. Bolhuis, and P. L. Geissler, Transition path sampling, *Advances in Chemical Physics* (John Wiley & Sons, Ltd., 2002), Chap. 1, pp. 1–78 .

[73] G. E. Crooks and D. Chandler, Efficient transition path sampling for nonequilibrium stochastic dynamics, Phys. Rev. E **64**, 026109 (2001).

[74] D. Frenkel and B. Smit, *Understanding Molecular Simulations from Algorithms to Applications* (Academic, San Diego, California, U.S.A., 2002).