

Transition rate theory, spectral analysis, and reactive paths

Cite as: J. Chem. Phys. 156, 134111 (2022); doi: 10.1063/5.0084209

Submitted: 4 January 2022 • Accepted: 9 March 2022 •

Published Online: 4 April 2022



View Online



Export Citation



CrossMark

Benoît Roux^{a)} 

AFFILIATIONS

Department of Biochemistry and Molecular Biology, Department of Chemistry, The University of Chicago, 5735 S Ellis Ave., Chicago, Illinois 60637, USA

^{a)} Author to whom correspondence should be addressed: roux@uchicago.edu

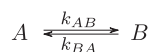
ABSTRACT

The kinetics of a dynamical system dominated by two metastable states is examined from the perspective of the activated-dynamics reactive flux formalism, Markov state eigenvalue spectral decomposition, and committor-based transition path theory. Analysis shows that the different theoretical formulations are consistent, clarifying the significance of the inherent microscopic lag-times that are implicated, and that the most meaningful one-dimensional reaction coordinate in the region of the transition state is along the gradient of the committor in the multidimensional subspace of collective variables. It is shown that the familiar reactive flux activated dynamics formalism provides an effective route to calculate the transition rate in the case of a narrow sharp barrier but much less so in the case of a broad flat barrier. In this case, the standard reactive flux correlation function decays very slowly to the plateau value that corresponds to the transmission coefficient. Treating the committor function as a reaction coordinate does not alleviate all issues caused by the slow relaxation of the reactive flux correlation function. A more efficient activated dynamics simulation algorithm may be achieved from a modified reactive flux weighted by the committor. Simulation results on simple systems are used to illustrate the various conceptual points.

© 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0084209>

I. INTRODUCTION

The dynamics of a complex system with two long-lived metastable states is a classical problem that is often represented phenomenologically by the kinetic model,



with the overall relaxation time $\tau^* = (k_{AB} + k_{BA})^{-1}$ and equilibrium probabilities $p_A = k_{BA}/(k_{AB} + k_{BA})$ and $p_B = k_{AB}/(k_{AB} + k_{BA})$. One of the most important theoretical frameworks to tackle such a problem in complex systems has been Chandler's activated-dynamics reactive flux formalism.¹ Assuming that x is a good "reaction coordinate" for the system of interest, the forward transition rate k_{AB} is then expressed as

$$k_{AB} = \kappa k_{AB}^{\text{TST}}, \quad (1)$$

where k_{AB}^{TST} is the transition state theory (TST) rate evaluated at the transition state x^\ddagger ,²⁻⁴ and the prefactor κ is the transmission

coefficient serving as a correction ($\kappa \leq 1$). In the context of this analysis, κ can be deduced from a time-correlation function evaluated at the molecular time scale, τ_m , expected to be much shorter than the overall relaxation time of the system τ^* . For this reason, one of the most attractive features of the activated-dynamics algorithm is that κ can be determined from the fate of relatively short trajectories of length τ_m initiated at the transition state.^{1,5,6}

Generally, to characterize the dynamics of complex systems with two long-lived metastable states, one may adopt two different fundamental perspectives. One may choose to focus on the global relaxation time τ^* of the system or, alternatively, on the net unidirectional reactive flux J_{AB} from the state A to the state B . For a simple two-state kinetic model, the global relaxation time τ^* incorporates both the influence of forward $A \rightarrow B$ and backward $B \rightarrow A$ transitions, reflecting the natural back and forth dynamics occurring spontaneously in the system under equilibrium conditions. In contrast, the forward transition rate k_{AB} , like the mean first passage time (MFPT),⁷ is more directly associated with a specific unidirectional $A \rightarrow B$ reactive flux, picturing the kinetic of the system in terms of a

“reactant” to a “product” state. In some sense, the first picture seems to simply reflect the system’s natural kinetics, whereas the second picture frames the issue more deliberately by ascribing a directionality to the system’s kinetics.⁷ Of course, both pictures are equivalent because under equilibrium conditions, the forward and backward net unidirectional reactive flux must be equal, $J_{BA} = J_{AB}$, and detailed balance implies that $p_A k_{AB} = p_B k_{BA}$.

Although this discussion is organized around an exceedingly simple two-state system, it helps illustrate the two different computational and theoretical framework that can be used to characterize the kinetics of highly complex systems. One may choose to attack the problem from the point of view of the global timescales within the system by carrying out a spectral decomposition analysis of the dynamical propagator,^{8–12} or one may consider the net forward flux in one specific direction by deliberately identify boundary reactant and product states and determining the steady-state flux of reactive trajectories between them.^{7,13–15} This construct, which aims its attention toward the set of $A \rightarrow B$ reactive trajectories forming the transition path ensemble, is a foundational element of the transition path sampling (TPS) algorithm^{16–20} and transition path theory (TPT).¹⁴ While the spectral analysis is untainted by the somewhat subjective choice of the boundary states, these theoretical frameworks more directly focus on the specific transition that is the object of interest. This is advantageous in highly complex multi-state systems for which a spectral analysis can become obscured by slow processes that are not relevant to the transitions of interest between A and B . The analysis of the conditions establishing a net steady-state flux between two metastable states leads to the observation that the principal lines of reactive probability current from the state A to the state B are largely determined by the equilibrium probability and the variations in the committor probability.^{7,13–15} Originally introduced by Onsager in 1938,²¹ and subsequently re-discovered in the late 1990s,^{22–24} the committor probability is a fundamental building block of TPT.¹⁴ This perspective is a critical insight in the formulation of the string method,^{12,25–29} aimed at determining the “reaction tube” that supports most of the reactive flux from A to B . It is the objective of the present analysis to shed new light on the characterization the kinetics of complex systems by examining the relationship between the reactive flux formalism, the spectral analysis of Markov models, and the TPT framework.

II. THEORETICAL DEVELOPMENTS

A. Reactive flux formalism

For the sake of clarity, it is worth recalling the main elements of the reactive flux formalism for a system with two metastable states A and B . It is assumed that there is a single “reaction channel” between the two states. Having defined a population operator or indicator function H_B that is equal to 1 when the system is in state B and equal to zero when the system is in state A , Chandler started by considering the normalized population time-correlation function,¹

$$C(t) = \frac{\langle \delta H_B(0) \delta H_B(t) \rangle}{\langle \delta H_B(0) \delta H_B(0) \rangle}, \quad (2)$$

where $\delta H_B(t) = H_B(t) - \langle H_B \rangle$ is the deviation from the average. Additional useful relations include $\langle H_B \rangle = p_B$, $\langle H_A \rangle = \langle 1 - H_B \rangle = p_A$, and $\langle \delta H_B \delta H_B \rangle = p_A p_B$. Noting that $C(t) \sim e^{-t/\tau^*}$ at long time according to the phenomenology, Chandler argued that the time derivative $\dot{C}(t)$ must decay from an initial value to a plateau equal to $-1/\tau^*$ in the limit of $t \rightarrow \tau_m$. In the context of this analysis, τ_m represents a molecular time scale that must be much shorter than the overall relaxation time of the system $\tau^* = (k_{AB} + k_{BA})^{-1}$. Traditionally, one must then define a “reaction coordinate,” $x(\mathbf{z})$, as a function of a set of collective variables \mathbf{z} . Once the transition state x^\ddagger is identified, the indicator function H_B can be constructed as

$$H_B(\mathbf{z}) = \theta(x(\mathbf{z}) - x^\ddagger), \quad (3)$$

where $\theta(\cdot)$ is a Heaviside step-function. Assuming that x is a good reaction coordinate for the system of interest, the forward transition rate $k_{AB} = p_B/\tau^*$ is then expressed as

$$\begin{aligned} k_{AB} &= \lim_{t \rightarrow \tau_m} \frac{1}{p_A} \langle \delta(x(0) - x^\ddagger) \dot{x}(0) H_B(t) \rangle \\ &= \frac{\rho_{\text{eq}}(x^\ddagger)}{\int_A \rho_{\text{eq}}(x) dx} \langle \dot{x} \theta(x^\ddagger) \rangle_{(x^\ddagger)} \lim_{t \rightarrow \tau_m} \frac{\langle \delta(x - x^\ddagger) \dot{x} H_B(t) \rangle}{\langle \delta(x - x^\ddagger) \dot{x} \theta(x^\ddagger) \rangle} \\ &= k_{AB}^{\text{TST}} \kappa, \end{aligned} \quad (4)$$

where x^\ddagger is the position of the transition state, $\rho_{\text{eq}}(x)$ is the marginal equilibrium distribution along x , k_{AB}^{TST} is the transition state theory (TST) rate, and κ is the transmission coefficient serving as a correction to the transition state theory rate,

$$\kappa = \lim_{t \rightarrow \tau_m} \frac{\langle \delta(x - x^\ddagger) \dot{x} H_B(t) \rangle}{\langle \delta(x - x^\ddagger) \dot{x} \theta(x^\ddagger) \rangle}. \quad (5)$$

Alternatively, the k_{AB} transition rate can also be formulated as the time derivative of the conditional probability that the system will be found in state B at time t , assuming it was initially in state A at time $t = 0$,³⁰

$$\begin{aligned} k_{AB} &= \lim_{t \rightarrow \tau_m} \frac{d}{dt} \langle H_B(t) \rangle_{(\text{in } A \text{ at } t=0)} \\ &= \lim_{t \rightarrow \tau_m} \frac{\langle H_A(0) \dot{H}_B(t) \rangle}{\langle H_A(0) \rangle} \\ &= \lim_{t \rightarrow \tau_m} \frac{1}{p_A} \langle (1 - H_B(0)) \dot{H}_B(t) \rangle \\ &= \lim_{t \rightarrow \tau_m} -\frac{1}{p_A} \langle H_B(0) \dot{H}_B(t) \rangle, \end{aligned} \quad (6)$$

where $\langle \dot{H}_B(t) \rangle = 0$ was used.

One of the most attractive features of the activated-dynamics algorithm based on Eq. (4) is that κ can be determined from the fate of relatively short trajectories of length τ_m initiated at the transition state x^\ddagger .^{1,5,6} On the other hand, the approach may be

encountering considerable challenges when the time τ_m needed to reach the plateau of the time-correlation function of Eq. (5) is too long. While the formalism provides a sound description of the relaxation of a two-state system from a fundamental point of view, the relative measure of success or failure of the approach depends on the rate of convergence of Eq. (5). Methodologies that can help reduce τ_m and can lead to significantly better statistical properties of the transmission coefficient and the reactive flux estimates are important.^{31,32}

B. Effective propagator and spectral decomposition

The probability density of the system at time t is expressed as $\rho(\mathbf{x}, \mathbf{v}; t)$, where \mathbf{x} and \mathbf{v} represent the set of coordinates x_i and velocities v_i , respectively.¹² Using $\mathbf{u} \equiv (\mathbf{x}, \mathbf{v})$ to represent a point in phase space, the forward propagation step ($\mathbf{u} \rightarrow \mathbf{u}'$) for the probability density from the time t to the time $t + \Delta t$ is

$$\rho(\mathbf{u}'; t + \Delta t) = \int d\mathbf{u} P_{\Delta t}(\mathbf{u}'|\mathbf{u}) \rho(\mathbf{u}; t). \quad (8)$$

The elementary propagator for a null time step, $P_0(\mathbf{u}'|\mathbf{u})$, is the identity $\delta(\mathbf{u}' - \mathbf{u})$. The dynamical propagation, which we may formally represent as $\rho(t + \Delta t) = P_{\Delta t} \cdot \rho(t)$, obeys the Chapman–Kolmogorov equation for arbitrary times. The propagator with the microscopic time step may be repeatedly applied an arbitrary number of times as $\rho(t + n\Delta t) = (P_{\Delta t})^n \cdot \rho(t)$. The forward–backward microscopic detailed balance relation, $P_{\Delta t}(\mathbf{u}'|\mathbf{u}) \rho_{\text{eq}}(\mathbf{u}) = P_{\Delta t}^{\dagger}(\mathbf{u}|\mathbf{u}') \rho_{\text{eq}}(\mathbf{u}')$, is satisfied, where $P_{\Delta t}^{\dagger}$ is the backward propagator. While a formulation based on the microscopic propagator $P_{\Delta t}(\mathbf{u}'|\mathbf{u})$ offers the most complete representation of the reactive paths, it is generally necessary to consider the dynamics projected onto a subspace of reduced dimensionality. We define the effective propagator \mathcal{P}_{τ} for the finite lag-time τ within the subspace spanned by a subset of collective variables (CVs) as

$$\begin{aligned} \mathcal{P}_{\tau}(\mathbf{z}'|\mathbf{z}) &= \frac{1}{\rho_{\text{eq}}(\mathbf{z})} \int d\mathbf{u} \\ &\times \int d\mathbf{u}' \delta(\tilde{\mathbf{z}}(\mathbf{x}') - \mathbf{z}') \mathcal{P}_{\tau}(\mathbf{u}'|\mathbf{u}) \rho_{\text{eq}}(\mathbf{u}) \delta(\tilde{\mathbf{z}}(\mathbf{x}) - \mathbf{z}), \end{aligned} \quad (9)$$

where $\tilde{\mathbf{z}}(\mathbf{x}) = (\tilde{z}_1(\mathbf{x}), \dots, \tilde{z}_N(\mathbf{x}))$ is a vector-valued function that maps every microscopic configuration \mathbf{x} of the system on a set of values $\tilde{\mathbf{z}}(\mathbf{x})$. The reduced probability density of the system at time t is expressed as $\rho(\mathbf{z}; t)$. The forward propagation step ($\mathbf{z} \rightarrow \mathbf{z}'$) for the reduced probability density from the time t to the time $t + \tau$ is

$$\rho(\mathbf{z}'; t + \tau) = \int d\mathbf{z} \mathcal{P}_{\tau}(\mathbf{z}'|\mathbf{z}) \rho(\mathbf{z}; t). \quad (10)$$

It is assumed that the dynamics within the reduced subspace of the CVs is Markovian with a finite lag-time τ and that the propagator obeys the Chapman–Kolmogorov equation, $\rho(t + n\tau) = (\mathcal{P}_{\tau})^n \cdot \rho(t)$. It is assumed that the system is in equilibrium and that we have microscopic detailed balance, $\mathcal{P}_{\tau}(\mathbf{z}'|\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}) = \mathcal{P}_{\tau}(\mathbf{z}|\mathbf{z}') \rho_{\text{eq}}(\mathbf{z}')$, where $\rho_{\text{eq}}(\mathbf{z}) = \int d\mathbf{x} d\mathbf{v} \delta(\tilde{\mathbf{z}}(\mathbf{x}) - \mathbf{z}) \rho_{\text{eq}}(\mathbf{x}, \mathbf{v})$ is

the equilibrium probability in the subspace of the CVs. Under these conditions, the effective propagator $\mathcal{P}_{\tau}(\mathbf{z}'|\mathbf{z})$ yields a self-consistent representation of the dynamics of the system within this subspace (closure of the dynamical propagation), built on a trajectory generated via the elementary propagator $P_{\Delta t}(\mathbf{u}'|\mathbf{u})$ with a time step Δt shorter than τ .

In practice, one should seek to determine the smallest possible lag-time τ that achieves Markovity for the effective propagator. An important framework to examine this issue is to rely on a spectral decomposition of the effective dynamical propagator.^{8,9,12} The right-eigenvector $\psi_k^{\text{R}}(\mathbf{z})$ of the operator is defined as^{8,9}

$$\lambda_k(\tau) \psi_k^{\text{R}}(\mathbf{z}') = \int d\mathbf{z} \mathcal{P}_{\tau}(\mathbf{z}'|\mathbf{z}) \psi_k^{\text{R}}(\mathbf{z}), \quad (11)$$

where the eigenvalue $\lambda_k(\tau) = e^{-\mu_k \tau}$. The constants $\mu_k \geq 0$ represents the associated τ -independent intrinsic decay rate of the n th eigenmode. The eigenvector $\psi_1^{\text{R}}(\mathbf{z})$ with the eigenvalue $\lambda_1 = 1$ ($\mu_0 = 0$) corresponds to the invariant equilibrium vector, $\rho_{\text{eq}}(\mathbf{z})$. The eigenvalues are ordered from the slowest to the fastest process, i.e., $1 = \lambda_1 > \lambda_2 > \lambda_3 > \dots$, and $0 = \mu_1 < \mu_2 < \mu_3 < \dots$. There is also a set of associated orthogonal left-eigenvectors,

$$\lambda_k \psi_k^{\text{L}}(\mathbf{z}) = \int d\mathbf{z}' \psi_k^{\text{L}}(\mathbf{z}') \mathcal{P}_{\tau}(\mathbf{z}'|\mathbf{z}), \quad (12)$$

with

$$\delta_{kl} = \int d\mathbf{z} \psi_k^{\text{L}}(\mathbf{z}) \psi_l^{\text{R}}(\mathbf{z}) = (\psi_k^{\text{L}} \cdot \psi_l^{\text{R}}) \quad (13)$$

and $\psi_k^{\text{L}}(\mathbf{z}) = \psi_k^{\text{R}}(\mathbf{z}) \rho_{\text{eq}}(\mathbf{z})^{-1}$. Orthonormalization can be expressed as

$$\delta_{kl} = \int d\mathbf{z} \psi_k^{\text{L}}(\mathbf{z}) \psi_l^{\text{L}}(\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}). \quad (14)$$

The first right-eigenvector is actually the equilibrium distribution, $\psi_1^{\text{R}}(\mathbf{z}) = \rho_{\text{eq}}(\mathbf{z})$, and the first left-eigenvector is equal to unity. The equilibrium time-correlation function of an arbitrary function $v(\mathbf{z})$ is

$$\begin{aligned} \langle v(n\tau) v(0) \rangle &= \int d\mathbf{z} \int d\mathbf{z}' v(\mathbf{z}') \mathcal{P}_{n\tau}(\mathbf{z}'|\mathbf{z}) v(\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}) \\ &= \sum_k (v \cdot \psi_k^{\text{R}})^2 e^{-\mu_k n\tau}, \end{aligned} \quad (15)$$

where

$$(v \cdot \psi_k^{\text{R}}) = \int d\mathbf{z} v(\mathbf{z}) \psi_k^{\text{R}}(\mathbf{z}). \quad (16)$$

C. Spectral decomposition of a two state system

Let us return to the normalized population time-correlation function between a “reactant” state A and a “product” state B of Eq. (2) that is the starting point of the activated-dynamics reactive

flux formalism.¹ The indicator functions, H_A and H_B , are defined such that $H_A = 1$ when the system is in state A and zero otherwise and that $H_B = 1$ when the system is in state B and zero otherwise. By construction, $H_A + H_B = 1$, and the equilibrium probability of the state A and B is $p_A = \langle H_A \rangle$ and $p_B = \langle H_B \rangle$, respectively.

Relying on a spectral analysis of the Markov dynamics, the long-time relaxation of the system displayed by the equilibrium time-correlation function of the indicator function H_B for the state B ,

$$\begin{aligned} \langle \delta H_B(0) \delta H_B(n\tau) \rangle &= \langle H_B(0) H_B(n\tau) \rangle - \langle H_B \rangle \langle H_B \rangle \\ &= \int d\mathbf{z} \int d\mathbf{z}' H_B(\mathbf{z}') \mathcal{P}_{n\tau}(\mathbf{z}'|\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}) H_B(\mathbf{z}) \\ &\quad - \int d\mathbf{z} \rho_{\text{eq}}(\mathbf{z}) H_B(\mathbf{z}) \int d\mathbf{z}' \rho_{\text{eq}}(\mathbf{z}') H_B(\mathbf{z}') \\ &= \sum_{k \geq 1} (H_B \cdot \psi_k^R)^2 e^{-\mu_k n\tau} - (H_B \cdot \psi_1^R)^2 \\ &= \sum_{k > 1} (H_B \cdot \psi_k^R)^2 e^{-\mu_k n\tau}, \end{aligned} \quad (17)$$

where

$$(H_B \cdot \psi_k^R) = \int d\mathbf{z} H_B(\mathbf{z}) \psi_k^R(\mathbf{z}). \quad (18)$$

If the indicator function H_B matches the metastable states of the system accurately, the amplitude of the eigenvector ψ_2^R should be nearly constant for all the states i within the two metastable basins. The integral of the indicator functions with the other eigenvectors should be negligible, i.e., $(H_B \cdot \psi_k^R) \approx 0$ for $k > 2$. This means that the second eigenmode ψ_2^R essentially represents the global transfer of probability between the metastable basins A and B and the normalized population time-correlation function from Eq. (2) is

$$\begin{aligned} C(n\tau) &= \frac{\sum_{k \geq 1} (H_B \cdot \psi_k^R)^2 e^{-\mu_k n\tau}}{\sum_{k > 1} (H_B \cdot \psi_k^R)^2} \\ &= e^{-\mu_2 n\tau} (1 + \dots). \end{aligned} \quad (19)$$

Thus, if there is mainly one slow mode corresponding to the transitions between the metastable states A and B , the long-time relaxation of the correlation function is expected to reflect mainly the second eigenvalue μ_2 . In a system with two metastable states, μ_2 can be related to the overall relaxation time corresponding to $\tau^* = 1/\mu_2$ and the forward and backward transition rates, $k_{AB} = p_B \mu_2$ and $k_{BA} = p_A \mu_2$. If the indicator function overlaps with higher order eigenmodes, the dominant relaxation time can be revealed by considering the correlation function over a longer time $\tau_m = n\tau$,

$$\begin{aligned} \lim_{n\tau \rightarrow \tau_m} \dot{C}(n\tau) &= \lim_{n\tau \rightarrow \tau_m} \frac{\sum_{k \geq 1} (H_B \cdot \psi_k^R)^2 (-\mu_k) e^{-\mu_k n\tau}}{\sum_{k > 1} (H_B \cdot \psi_k^R)^2} \\ &\approx -\mu_2, \end{aligned} \quad (20)$$

which is reasonable as long as $e^{-\mu_2 \tau_m} \approx 1$ while $e^{-\mu_3 \tau_m} \ll 1$. This is possible, in practice, only if there is a large gap between the eigenvalues μ_2 and μ_3 .

D. Perspective from transition path theory

1. Steady-state forward flux and committor probability

A spectral analysis may encounter issues when some of the slowest eigenmodes are uncorrelated with the $A \rightarrow B$ transition of interest. This issue is circumvented by deliberately selecting the boundary states A and B to consider the net steady-state reactive flux for the reaction.^{7,13–15} Within the TPT framework,¹⁴ the net steady-state unidirectional reactive flux from A to B can be constructed from the probability to make a transition from \mathbf{z} to \mathbf{z}' multiplied by the probabilities that the trajectory arriving at \mathbf{z} came from A and that the trajectory will then go on from \mathbf{z}' to reach B , minus the probabilities that the trajectory arriving at \mathbf{z} came from B and that the trajectory will then go on from \mathbf{z}' to reach A . Invoking microscopic detailed balance, this yields

$$J_{AB} = \frac{1}{\tau} \int_{\mathbf{z} \in A'} d\mathbf{z} \int_{\mathbf{z}' \in B'} d\mathbf{z}' (q(\mathbf{z}') - q(\mathbf{z})) \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}), \quad (21)$$

where $q(\mathbf{z})$, called the forward committor, is the probability that a trajectory started at \mathbf{z} will first reach the state B . The bounds on the integral imply that the entire subspace of CVs has been divided in a region A' that includes the state A and a region B' that includes the state B . For the effective dynamics within the subspace \mathbf{z} , $q(\mathbf{z})$ must satisfy the backward propagation condition,

$$q(\mathbf{z}) = \int d\mathbf{z}' q(\mathbf{z}') \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}), \quad (22)$$

with the constraints $q = 0$ when $\mathbf{z} \in A$ and $q = 1$ when $\mathbf{z} \in B$. Equation (21) can be derived by counting only the forward flux $\mathbf{z} \rightarrow \mathbf{z}'$ arising exclusively from reactive $A \rightarrow B$ trajectories members of the transition path ensemble¹⁴ or by considering the net unidirectional reactive flux from A to B under steady-state conditions.¹³ Because one is free to move the boundary between the regions A' and B' , Eq. (21) can be transformed into a convenient unconstrained expression,^{12,14,15}

$$\begin{aligned} J_{AB} &= \frac{1}{2\tau} \int d\mathbf{z} \int d\mathbf{z}' (q(\mathbf{z}') - q(\mathbf{z}))^2 \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}) \\ &= \frac{1}{2\tau} \langle (q(\tau) - q(0))^2 \rangle \\ &= \frac{1}{\tau} (\langle q(0)q(0) \rangle - \langle q(0)q(\tau) \rangle) \\ &= \frac{1}{\tau} C_{qq}(\tau), \end{aligned} \quad (23)$$

where we have defined the committor time-correlation function,

$$C_{qq}(\tau) = \langle q(0)q(0) \rangle - \langle q(0)q(\tau) \rangle. \quad (24)$$

Here, $q(\tau)$ is an implicit short-hand notation for $q(\bar{\mathbf{z}}[\mathbf{x}(\tau)])$. Equations (23) and (24) can also be used to variationally optimize

the committor,^{12,14,33–35} minimizing J_{AB} with respect to a trial function, i.e., $\delta J_{AB}[q(\mathbf{z})]/\delta q(\mathbf{z}) = 0$, recovers the backward propagator condition of Eq. (22). Because the committor to reach state A before state B is simply equal to $(1 - q)$, Eq. (23) makes it clear that the steady-state flux from state B to state A is equal to that from state A to state B , $J_{BA} = J_{AB}$. The forward transition rate can be determined as $k_{AB} = J_{AB}/p_A$, and the mean first passage time (MFPT) from A to B is equal to $1/k_{AB}$.

Additional features about reactive paths can be formulated. For example, the probability $p_{AB}(\mathbf{z})$ for a $A \rightarrow B$ reactive trajectory knowing that the system is at \mathbf{z} is the probability $1 - q(\mathbf{z})$ that the trajectory came from A multiplied by the probability $q(\mathbf{z})$ that the trajectory will lead to B . Averaging over all possible \mathbf{z} weighted by the equilibrium distribution $\rho_{\text{eq}}(\mathbf{z})$ yields the total probability of a reactive $A \rightarrow B$ trajectory, $p_{AB} = \langle q(1 - q) \rangle$, at equilibrium. The latter may be understood as the fraction of time a very long trajectory of length t_{tot} spends on the transition pieces that are purely $A \rightarrow B$ reactive, τ_r/t_{tot} . Since J_{AB} is the total number of $A \rightarrow B$ reactive transitions n_r during the time t_{tot} , we can write $\langle q(1 - q) \rangle = (\tau_r/n_r) (n_r/t_{\text{tot}}) = \langle \tau_r \rangle J_{AB}$. Thus, the mean transit time during a reactive transition $\langle \tau_r \rangle$ is equal to $\langle q(1 - q) \rangle / J_{AB}$.^{13,20} Because $J_{AB} = J_{BA}$, then $p_{AB} = p_{BA}$ and the mean transit time is independent of the direction.

2. Markovian dynamics in the short time limit

While the present analysis is strictly valid if the effective propagator within the reduced subspace of the CVs is Markovian with the finite lag-time τ , one may ask when the reactive flux J_{AB} can be determined in the limit $\tau \rightarrow 0$,

$$J_{AB} = \lim_{\tau \rightarrow 0} \frac{C_{qq}(\tau) - C_{qq}(0)}{\tau} = \left. \frac{d}{d\tau} C_{qq}(\tau) \right|_{\tau=0} \quad (25)$$

$$= - \left. \frac{d}{d\tau} \langle q(0)q(\tau) \rangle \right|_{\tau=0} = - \int d\mathbf{z} \int d\mathbf{z}' q(\mathbf{z}') q(\mathbf{z}) \left(\left. \frac{d}{d\tau} \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \right|_{\tau=0} \right) \rho_{\text{eq}}(\mathbf{z}). \quad (26)$$

The flux is well-defined in the limit of $\tau \rightarrow 0$ if the underlying dynamics arises from a genuine continuous-time Markovian process,

$$\left. \frac{d}{d\tau} \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \right|_{\tau=0} = \mathcal{K}(\mathbf{z}'|\mathbf{z}), \quad (27)$$

where \mathcal{K} is the transition rate matrix, with $\mathcal{P}_\tau \equiv e^{\mathcal{K}\tau}$ and $\mathcal{P}_0(\mathbf{z}'|\mathbf{z}) = [\delta^0]_{(\mathbf{z}'|\mathbf{z})} = \delta(\mathbf{z} - \mathbf{z}')$, yielding

$$J_{AB} = - \int d\mathbf{z} \int d\mathbf{z}' q(\mathbf{z}') q(\mathbf{z}) \mathcal{K}(\mathbf{z}'|\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}). \quad (28)$$

Similarly, the flux is well-defined in the limit of $\tau \rightarrow 0$ if the effective dynamics within the subspace of CVs (\mathbf{z}) is assumed to be diffusive with diffusion matrix \mathbf{D} ; it can be shown that (Appendix A)

$$J_{AB} = \int d\mathbf{z} \nabla q(\mathbf{z}) \cdot \mathbf{D} \rho_{\text{eq}}(\mathbf{z}) \nabla q(\mathbf{z}). \quad (29)$$

Under these conditions, the forward reactive flux J_{AB} is given by the initial slope of the committor correlation function at $\tau = 0$, i.e., the time derivative evaluated at $\tau = 0$. It should be noted, however, that this expression would be identically equal to zero due to time reversibility if the microscopic dynamics were truly inertial.³⁶ Physically, this suggests that the proper limit for τ must be larger than 0 to establish the diffusive dynamics, i.e., $\dot{C}_{qq}(0^+) \approx (C_{qq}(\tau) - C_{qq}(0))/\tau$, with a small but finite $\tau > 0$.

3. Committor and eigenvectors

It is of interest to relate the unidirectional flux J_{AB} to the overall relaxation time determined from the spectral analysis. For this, we need to have a model for the committor. While the committor $q(\mathbf{z})$ is not quite a left-eigenvector of the effective propagator, a useful approximate construction can be written as³⁷

$$q(\mathbf{z}) \approx - \left(\frac{a}{b-a} \right) \psi_1^L(\mathbf{z}) + \left(\frac{1}{b-a} \right) \psi_2^L(\mathbf{z}), \quad (30)$$

where $a = \psi_2^L(\mathbf{z})$ with $\mathbf{z} \in A$ and $b = \psi_2^L(\mathbf{z})$ with $\mathbf{z} \in B$. The constants a and b must have opposite sign to guarantee that the vector is orthogonal to the equilibrium vector. It can be shown (Appendix B) that $a = -(b-a)p_B$, and $1 = (b-a)^2 (p_A p_B - \langle q(1-q) \rangle)$.

This construction makes a function $q(\mathbf{z})$ that is equal to 0 for $\mathbf{z} \in A$ and equal to 1 for $\mathbf{z} \in B$, which approximately satisfies the backward propagation condition of Eq. (22). Using the model committor, it can be shown (Appendix C) that the net unidirectional reactive flux from A to B is

$$J_{AB} = p_A p_B \frac{1}{\tau} (1 - e^{-\mu_2 \tau}) \left(1 + \frac{\langle \tau_r \rangle}{\tau} (1 - e^{-\mu_2 \tau}) \right)^{-1}. \quad (31)$$

This expression is the steady-state flux from A to B for the simple two-state model. For this result to be valid, the lag-time τ must be sufficiently long to yield an effective propagation within the subspace \mathbf{z} that is Markovian. If this condition is met, then the net forward flux J_{AB} is independent of the value of the lag-time. If we assume that $\mu_2 \tau \ll 1$, we have $(e^{-\mu_2 \tau} - 1) \approx (1 - \mu_2 \tau + \dots - 1)$ and that $\langle \tau_r \rangle \mu_2 \ll 1$, then we have $J_{AB} \approx p_A p_B \mu_2 = p_A k_{AB}$. The expression shows that the steady-state flux from A to B remains the same if the lag-time is $n\tau$, as long as $n\mu_2 \tau \ll 1$.

4. Approximate committors

The formal analysis relies on the existence of the effective propagator $\mathcal{P}_\tau(\mathbf{z}'|\mathbf{z})$ representing the Markovian dynamics of the system within the subspace of the CVs, \mathbf{z} . Nonetheless, it is understood that the microscopic trajectory $\mathbf{x}(t)$ is generated from the elementary propagator $P_{\Delta t}(\mathbf{u}'|\mathbf{u})$ with a time step Δt that may be much shorter than the lag-time τ . Formally, one can always fall back on the elementary propagator to represent the time-evolution of the system. In this context, the broadest perspective is offered by treating the function $q(\mathbf{z})$ as a function of the microscopic configuration $\mathbf{x}(t)$ that can be evaluated at any arbitrary time t , i.e., $q(\bar{\mathbf{z}}[\mathbf{x}(t)])$.

The time-correlation function $C_{qq}(t)$ may be considered, even if the underlying dynamics is non-Markovian and comprises inertial aspects at times shorter than τ (this will be the case in Fig. 3). Obviously, the function $q(\mathbf{z})$ associated with a Markovian dynamics within the subspace of the CVs is only an approximation to the committor $q(\mathbf{x}, \mathbf{v})$ in the full phase space. However, this construct does nonetheless provide a useful framework to analyze the true microscopic dynamics of the system. Importantly, the validity of the final rate expression does not require our representation of the effective dynamics within the subspace of the CVs to be exact at all times. A similar strategy was used by Ruiz-Montero *et al.*³² to formulate an efficiency activated-dynamics reactive flux simulations for systems experiencing large dissipative forces (we return to this in Sec. IV).

If the dynamics of the CVs genuinely corresponds to a diffusion process or is governed by a continuous-time Markov process, then the committor time-correlation function $C_{qq}(t)$ is rigorously linear for $t > 0^+$. In this case, there is no short transient at small times and the committor time-correlation function is immediately linear. The forward transition rate follows directly from the forward unidirectional flux J_{AB} in Eq. (31). Whenever any of those conditions are not satisfied, there will be some transient behavior at short time and the variational forward flux relation will be valid only for a lag-time that is sufficient long to establish the linear behavior of $C_{qq}(t)$. For example, the forward flux relation is not valid if the dynamical propagation within the subspace \mathbf{z} is non-Markovian because the lag-time τ is too short or if an approximate committor function $q_*(\mathbf{z})$ is used. In this case, the equilibrium time-correlation function from Eq. (15) is

$$\begin{aligned} \langle q_*(0) q_*(\tau) \rangle &= \int d\mathbf{z} \int d\mathbf{z}' q_*(\mathbf{z}) q_*(\mathbf{z}') \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}) \\ &= \sum_k (q_* \cdot \psi_k^R)^2 e^{-\mu_k \tau}, \end{aligned} \quad (32)$$

where

$$(q_* \cdot \psi_k^R) = \int d\mathbf{z} q_*(\mathbf{z}) \psi_k^R(\mathbf{z}). \quad (33)$$

Such an approximate committor function may overlap with higher order eigenvectors, ψ_k^L with $k > 2$. The correlation function will relax to the dominant rate if some lag-time τ_m exists such that $e^{-\mu_3 \tau_m} \ll 1$ while $e^{-\mu_2 \tau_m} \approx 1$. In this case, the committor time-correlation function is not immediately linear, even if the dynamics within the subspace \mathbf{z} is Markovian. As t increases, $C_{qq}(t)$ should begin to vary linearly with time, a behavior that is valid only up to a certain time scale because $C_{qq}(t) \rightarrow \langle (q_* - \langle q_* \rangle)^2 \rangle$ in the limit $t \rightarrow \infty$. There is an intermediate finite time scale τ_m larger than the Markovian lag-time τ where the correlation function may be expected to vary linearly with time. Because of the expected linear dependence, the net forward unidirectional flux J_{AB} is essentially given by the slope “ s ” of the correlation function $C_{qq}(t)$. This argument is valid if the committor time-correlation function has the form $C_{qq}(t) = c(t) + st$, where $c(t)$ is a small component that rapidly decays to some finite plateau value with $\lim_{t \rightarrow \tau_m} \dot{c}(t) = 0$.

Accordingly, we can write that the limiting slope of the committor time-correlation function is given by

$$J_{AB} = \lim_{t \rightarrow \tau_m} \dot{C}_{qq}(t). \quad (34)$$

The microscopic time τ_m required for this relation to be valid may need to be longer than the lag-time τ that is needed to yield a Markovian dynamics within the subspace \mathbf{z} .

III. ILLUSTRATIVE SIMULATIONS

For illustrative purposes, we consider three simple double-well one-dimensional systems with the potential,

$$W(z) = \begin{cases} \left(\frac{z}{9.5}\right)^{12} + 3.3 e^{-(z/0.6)^2} - 3.2 & \text{“narrow,”} \\ \left(\frac{z}{9.5}\right)^{12} + 3.3 e^{-(z/4.0)^2} - 3.3 & \text{“medium,”} \\ \left(\frac{z}{9.5}\right)^{12} + 3.2 e^{-(z/7.0)^{16}} - 3.2 & \text{“broad”,} \end{cases} \quad (35)$$

where z is in Å and $W(z)$ is in kcal/mol. The potentials are shown in Fig. 1 (top). The symmetric wells yield equilibrium probabilities $p_A = p_B = 0.5$. In all cases, the double-well system is symmetric with a single free energy barrier, which, in turn, can be characterized as “narrow,” “medium,” and “broad.” The Euler–Lagrange equation determining the committor probability $q(z)$ is

$$\frac{\partial}{\partial z} \left(e^{-W(z)/k_B T} \frac{\partial q(z)}{\partial z} \right) = 0, \quad (36)$$

subject to the constraint $q(z)|_{z \leq z_1} = 0$ and $q(z)|_{z \geq z_2} = 1$. The solution is

$$q(z) = \frac{\int_{z_1}^z e^{W(z')/k_B T} dz'}{\int_{z_1}^{z_2} e^{W(z')/k_B T} dz'}, \quad (37)$$

where $k_B T = 0.5915$ kcal/mol. For the sake of simplicity, the boundaries z_1 and z_2 were set to -7 and $+7$, respectively, for the three cases. The calculated committors are shown in Fig. 1 (bottom). Each system was simulated with Brownian dynamics (BD),³⁰ assuming a diffusion coefficient $D = 1 \text{ \AA}^2 \text{ ps}^{-1}$. A time step of 0.005 ps was used to generate a 1 μs trajectory (200×10^6 steps). Because the BD trajectory is Markovian along the z axis, the time step of 0.005 ps is the same as the lag-time τ .

The system with a broad barrier was also simulated with Langevin dynamics,³⁰ assuming a diffusion coefficient $D = 1 \text{ \AA}^2/\text{ps}$ and a mass m of 20 atomic-mass-unit. For these parameters, the relaxation time of the velocity–velocity correlation function $k_B T/Dm$ is 0.08 ps. At this time scale, the Langevin dynamics becomes essentially equivalent to a Brownian dynamics within a short time τ . A time step of 0.001 ps was used to generate a 1 μs trajectory (10^9 steps).

Three types of time-correlation functions were calculated from the BD trajectories. First, we consider the committor time-correlation function $C_{qq}(t) = \langle q(0)q(0) \rangle - \langle q(0)q(t) \rangle$, which was

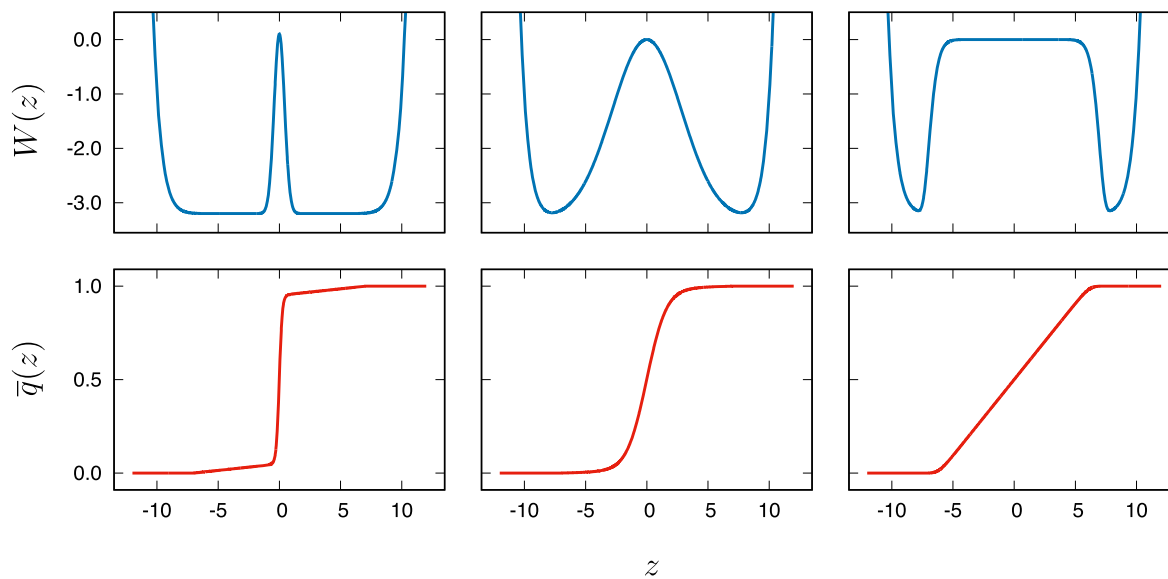


FIG. 1. Schematic illustration of a two-state system comprising a single narrow, medium, and broad free energy barrier. At the top (in blue), the potentials $W(z)$ are defined via Eq. (35) in kcal/mol as a function of z in Å. In the bottom (in red), the committors $q(z)$ for the different potentials are calculated via Eq. (37) with z_1 and z_2 set to -7 and $+7$, respectively.

defined in Eq. (24). We also consider the indicator time-correlation function $C_{BB}(t)$ defined as

$$C_{BB}(t) = p_A p_B \left(1 - \frac{\langle \delta H_B(0) \delta H_B(t) \rangle}{\langle \delta H_B(0) \delta H_B(0) \rangle} \right) = \langle H_B(0) H_B(0) \rangle - \langle H_B(0) H_B(t) \rangle. \quad (38)$$

The ratio in the parentheses is the normalized population time-correlation function introduced in Eq. (2) to develop the reactive flux formulation. It is easy to show that the forward flux $J_{AB} = p_A k_{AB} = \lim_{t \rightarrow \tau_m} \dot{C}_{BB}(t)$. Because $C_{BB}(t)$ is very similar in appearance to the committor-based time-correlation function $C_{qq}(t)$, this form is more convenient to compare the two different formulations. Finally, we define the position time-correlation function as

$$C_{zz}(t) = p_A p_B \left(1 - \frac{\langle \delta z(0) \delta z(t) \rangle}{\langle \delta z(0) \delta z(0) \rangle} \right), \quad (39)$$

with $\delta z(t) = z(t) - \langle z \rangle$. By construction, we have $C_{BB}(t) = C_{zz}(t) = C_{qq}(t) = 0$ at $t = 0$, and all three time-correlation functions are expected to converge to a straight line with the same limiting slope as t increases (but remains much smaller than τ^*). The time derivative of $C_{qq}(t)$ evaluated at the lag-time τ is related to the steady-state reactive flux, J_{AB} according to Eq. (23). Following Eq. (4), the time derivative $\dot{C}_{BB}(t)$ in the limit of $t \rightarrow \tau_m$ also yields the steady-state reactive flux, although the molecular time scale τ_m may be longer than the lag-time τ . Finally, according to Eq. (15), the time derivative of $\dot{C}_{zz}(t)$ is expected to behave similarly in the long time limit.

IV. DISCUSSION

A. Reactive flux and spectral analysis

The activated-dynamics reactive flux time-correlation function of Eq. (4) must decay to a plateau value within a molecular time scale τ_m for the transition rate of the system to be well defined.^{1,5,6} Central to this argument is the assumption that the normalized population time-correlation function of Eq. (2) follows a simple exponential decay with the overall relaxation time τ^* for times $t \gg \tau_m$. The spectral analysis of the effective propagator $\mathcal{P}_\tau(\mathbf{z}'|\mathbf{z})$ in terms of its eigenvalues μ_k leading to Eq. (20) clarifies the context supporting this assumption. Relating the activated-dynamics to the spectral analysis by comparing Eqs. (2) and (20), we can see that to properly extract the overall relaxation time of the two-state system, τ_m must be chosen such that the conditions, $\mu_2 \tau_m \ll 1$ and $\mu_3 \tau_m \gg 1$, are satisfied. This is possible only if there exists a large gap in the eigenvalue spectrum, $\mu_2 \ll \mu_3$. This is a requirement that the two-state system must satisfy; otherwise, the theory based on the fluctuations expressed in Eq. (2) is inappropriate.

The spectral analysis also shed some light on important aspects of the activated-dynamics reactive flux algorithm to efficiently determine a transition rate from computer simulations. In practice, it is well understood that the algorithm is very efficient when the two states are separated by a sharp free energy barrier, the reactive flux time-correlation function in Eq. (4) rapidly within a relatively short time τ_m . However, it is also recognized that the algorithm may not be as effective in the case of slow diffusive motion over a broad flat free energy barrier. In this case, a much longer time τ_m may be needed to reach the plateau of the reactive flux time-correlation function, requiring the simulation of long activated trajectories.

B. Illustrative simulations

To illustrate these points, Brownian dynamics simulations were carried out for simple one-dimensional systems (Fig. 1). The main results are shown in Fig. 2. The indicator time-correlation function $C_{BB}(t)$ is related to the activated-dynamics formalism to calculate the transition rate.¹ Position time-correlation functions, such as $C_{zz}(t)$, appear in the time-lagged independent component analysis (TICA).^{38,39} The committor time-correlation function $C_{qq}(t)$ originates from the TPT framework.^{7,12,14,15,33,40} From Fig. 2, it is clear that all time-correlation functions ultimately converge to the same slope, reflecting the global relaxation time within the system. However, the different time-correlation functions reach the same slope differently. When the barrier is extremely narrow, the indicator time-correlation function $C_{BB}(t)$ converges extremely rapidly to a straight line with constant slope, whereas the position autocorrelation function $C_{zz}(t)$ takes a longer time. There are large fluctuations along z in the two broad wells, but those fluctuations are non-reactive—they are not associated with a $A \rightarrow B$ transition. In contrast, the fluctuations in C_{BB} are faithfully reporting on such transition. In contrast, when the barrier is broad, the position autocorrelation function $C_{zz}(t)$ converges to a straight line more rapidly, whereas the indicator time-correlation function $C_{BB}(t)$ takes a much longer time. Nonetheless, even in the case of the broad barrier, the time-correlation functions relax to its constant slope within less than 1–2 ps. This is shorter

than the time needed to diffuse from the center of the barrier to its edge, estimated as $2Dt = L^2$, with $D = 1 \text{ \AA}^2 \text{ ps}^{-1}$ and $L = 7 \text{ \AA}$, to be on the order of 25 ps. The latter time scale is comparable to the mean transit time during a reactive transition, $\langle \tau_r \rangle = \langle q(1-q) \rangle / J_{AB}$,^{13,20} which is estimated to be on the order of 25 ps in this case. In the present simulations, J_{AB} is estimated to be 4.4×10^{-4} , 1.9×10^{-4} , and $1.1 \times 10^{-4} \text{ ps}^{-1}$ for the narrow, medium, and broad barrier, respectively. The corresponding global relaxation times $\tau^* = (\mu_2)^{-1}$ are estimated to be 568, 1315, and 2272 ps, which is much longer than the Markovian lag-time, τ (0.005 ps), and the mean transit time of reactive transitions, $\langle \tau_r \rangle$.

C. Reactive flux and approximate committor

The activated-dynamics reactive flux formalism requires the definition of a dividing surface between the states A and B , whereas the TPT framework requires the definition of two boundary regions associated with the states A and B . Yet, it is interesting to note that the expressions derived above for the forward transition rate k_{AB} from the activated-dynamics reactive flux formalism bears a certain similarity to that from the committor time-correlation function. Equations (4) and (34) are equivalent if one imagines that the indicator function H_B is substituted for the committor q . Specifically, we have that $\dot{C}_{qq}(\tau) = \dot{C}_{BB}(\tau_m) = J_{AB}$, with the substitution $\langle q(0)\dot{q}(t) \rangle \rightarrow \langle H_B(0)\dot{H}_B(t) \rangle$. Like the committor q , H_B is equal to

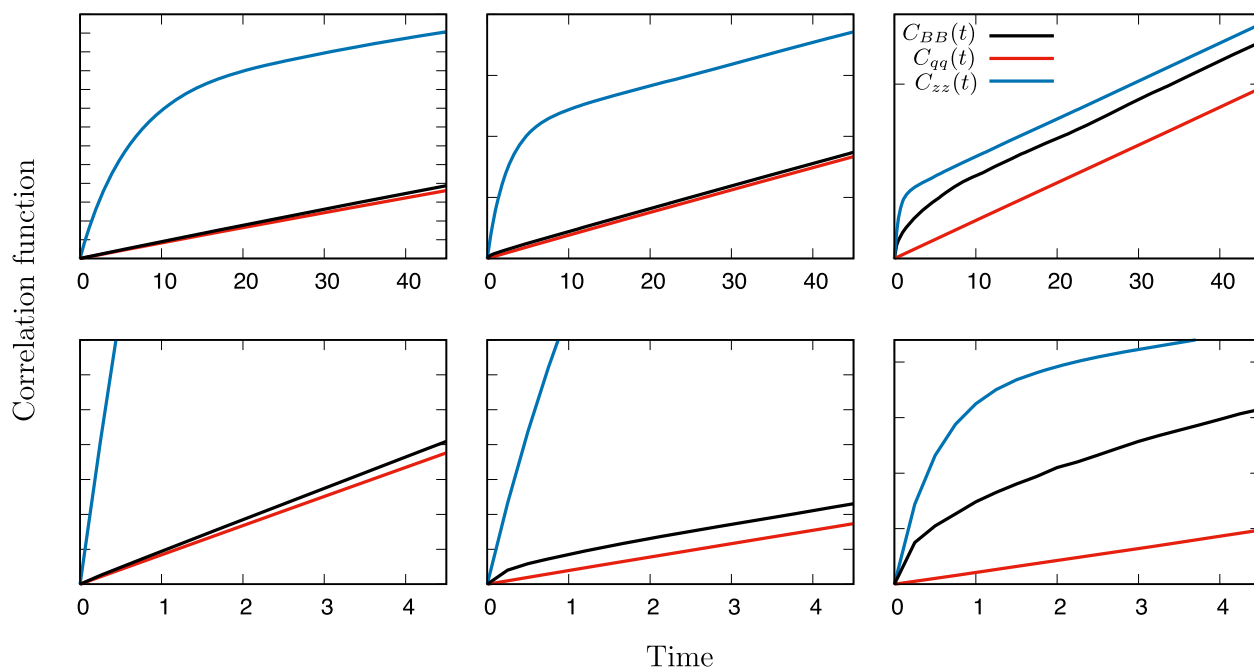


FIG. 2. Time-correlation functions calculated from Brownian dynamics simulations for the simple two-state system with a narrow, medium, and broad free energy barrier (shown in Fig. 1). Shown is the indicator autocorrelation function $C_{BB}(t)$ (in black), the committor correlation function $C_{qq}(t)$ (in red), and the position autocorrelation function $C_{zz}(t)$ (in blue). The bottom plots show a zoom of the plots at the top. For a given potential, all the correlation functions converges to the same slope, which is related to the net steady-state unidirectional reactive flux J_{AB} . For a narrow, medium, and broad free energy barrier, J_{AB} is estimated to be 4.4×10^{-4} , 1.9×10^{-4} , and $1.1 \times 10^{-4} \text{ ps}^{-1}$, respectively. The tick marks along the y-axis are evenly separated by 0.005 in the top plot and by 0.0005 in the bottom plot (the y labels are not displayed for the sake of clarity). The time axis is in ps.

zero in state A and one in state B , and in that sense, the Heaviside step-function H_B stands as some kind of crude approximation to the exact committor. By extension, when there are multiple collective variables \mathbf{z} , one could imagine defining the reaction coordinate as $x = (\mathbf{z} - \mathbf{z}^\ddagger) \cdot \mathbf{n}$ and the indicator function as

$$H_B(x) = \theta\left((\mathbf{z} - \mathbf{z}^\ddagger) \cdot \mathbf{n}\right). \quad (40)$$

For $H_B(x)$ to serve as an effective (albeit crude) approximation to the committor $q_*(\mathbf{z})$, the vector \mathbf{n} should be parallel to the gradient of the committor, $\nabla q(\mathbf{z}^\ddagger)$, with \mathbf{z}^\ddagger in the vicinity of the transition state region. For example, \mathbf{z}^\ddagger can be defined as

$$\mathbf{z}^\ddagger = \frac{\int d\mathbf{z} \mathbf{z} \delta(q(\mathbf{z}) - 0.5) \rho_{\text{eq}}(\mathbf{z})}{\int d\mathbf{z} \delta(q(\mathbf{z}) - 0.5) \rho_{\text{eq}}(\mathbf{z})}, \quad (41)$$

where $q(\mathbf{z}) = 0.5$ is the separatrix. Because the indicator function in Eq. (40) serves as an approximate committor $q_*(\mathbf{z})$, a time τ_m longer than τ may be needed in this case to reach the linear regime of the time-correlation function. It is also interesting to consider Eq. (40) as a functional form depending parametrically on \mathbf{n} and \mathbf{z}^\ddagger . Recalling that Eqs. (23) and (24) for the reactive flux J_{AB} and time-correlation function $C_{qq}(\tau)$ provide a variational principle to determine the committor from a trial function,^{12,14,33–35} one would minimize the reactive flux J_{AB} as a function of \mathbf{n} and \mathbf{z}^\ddagger to obtain variationally optimized parameters. This sheds new light on the significance of variational TST, which seeks to optimize the reaction coordinate by minimizing the transition rate.⁴¹ It also suggests that the most meaningful reactive direction at the transition state in the multidimensional space of the CVs is parallel to $\nabla q(\mathbf{z}^\ddagger)$, consistent with the multidimensional Kramers–Langer theory.⁴² Indeed, as noted in TPS studies,¹⁷ configurations sampled according to $\delta(\mathbf{z} - \mathbf{z}^\ddagger) \cdot \mathbf{n} \rho_{\text{eq}}(\mathbf{z})$ lead to a broad distribution of committor probabilities when using a suboptimal reaction coordinate with a vector \mathbf{n} that is not parallel to $\nabla q(\mathbf{z}^\ddagger)$.

Approximating the committor as a Heaviside step-function is likely to be more accurate when a sharp free energy barrier separates the states A and B . However, in the case of diffusive dynamics occurring on the top of a wide flat barrier, such a Heaviside step-function is not a very good representation of the correct committor. This is precisely the situation where the activated-dynamics reactive flux algorithm encounters difficulties to converge rapidly (Fig. 2). This suggests that designing improved model committors may help formulate more efficient computational algorithms to determine the transition rate.

The idea of exploiting an approximate committor as a framework to improve the convergence of a transition rate calculation bares a certain similarity with a very interesting strategy previously designed by Ruiz-Montero, Frenkel, and Brey.³² Their method, which also starts from the idea that the Heaviside step functions in the high friction limit should be replaced with the smooth function akin to the committor, was shown to dramatically improve the statistical convergence of the reactive flux algorithm for diffusive barriers. While they used an initial perturbation and characteristic functions that resemble the steady state concentration profile, it is well known in TPT that the latter equal to $\rho_{\text{eq}}(\mathbf{z})q(\mathbf{z})(1 - q(\mathbf{z}))$.¹³

As with the idea of using an approximate committor, the high friction steady-state picture served only to better formulate the problem in their method, and the validity of the final rate expression did not require this picture to be exact.³²

D. Committor as reaction coordinate

A striking observation from Fig. 2 is that the committor time-correlation function $C_{qq}(t)$ yields the correct slope at very short time in all cases, independent of the shape of the free energy barrier. This is in strong contrast with the time-correlation functions $C_{BB}(t)$ and $C_{zz}(t)$. The convergence of the committor time-correlation function to the correct slope within the lag-time τ is expected. If the underlying dynamics in the subspace \mathbf{z} is governed by a continuous-time Markov process as in Eq. (28) or by a diffusion process as in Eq. (A1), then the committor time-correlation function is immediately linear and the transition rate can be determined in the limit of $t \rightarrow 0^+$ from Eq. (28). When those conditions are not quite satisfied, there will be some transient behavior over the time-lag τ needed to achieve Markovity for the effective propagator \mathcal{P}_τ within the subspace \mathbf{z} .¹² Establishing the Markovity of \mathcal{P}_τ is generally non-trivial, although the present observations suggest that a simple criteria might be to verify whether the time derivative $\dot{C}_{qq}(t)$ reaches a plateau as $t \rightarrow \tau$. Whether this conjecture can be demonstrated rigorously is unclear. Once this is done, the net steady-state reactive flux J_{AB} and the transition rate k_{AB} evaluated at the lag-time τ can be obtained on the basis of Eq. (23).

It is sometimes tempting to imagine that the dynamics along the committor is Markovian with the lag-time τ . To elaborate on this point, C_{qq} could be formally expressed as

$$C_{qq}(\tau) = \frac{1}{2} \int dq \int dq' (q' - q)^2 \mathcal{P}_\tau(q'|q) \rho_{\text{eq}}(q), \quad (42)$$

where $\mathcal{P}_\tau(q'|q)$ represents an effective reduced propagator for the forward step ($q \rightarrow q'$) for the time step τ , which can be defined via Eqs. (9) and (10). Let us recall that τ is the lag-time that achieves Markovity for the effective propagator \mathcal{P}_τ within the subspace \mathbf{z} .¹² However, there is no guaranty that the dynamics of the system projected onto the one-dimensional coordinate q ought to be Markovian with the same lag-time τ . Therefore, while Eq. (42) is rigorously correct, it does not imply that the effective propagator $\mathcal{P}_\tau(q'|q)$ obeys the Chapman–Kolmogorov equation. In fact, achieving a self-consistent representation of the dynamics of the system projected onto the committor with the lag-time τ (closure of the dynamical propagation) is likely to require non-Markovian memory effects.^{43–47}

Nonetheless, there are reasons to believe that the committor might serve as a useful guide to define a reaction coordinate. Indeed, it has been shown that in the case of a multidimensional activated process controlled by diffusion, using the vector normal to the iso-committor plane separatrix ($q^\ddagger = 0.5$) to construct a one-dimension reaction coordinate yields a rate constant that is identical to that predicted by the multidimensional Kramers–Langer theory.⁴² Similarly, as made clear by Eq. (40), the most effective 1D reaction coordinate in a multidimensional space \mathbf{z} is parallel to the gradient of the committor in the region of the transition state, $\nabla q(\mathbf{z}^\ddagger)$. However, simply treating the committor as a one-dimensional

reaction coordinate in the activated-dynamics formalism does not necessarily alleviate the issues caused by slow relaxation and intermediate time scales. In practice, redefining the population operator H_B on the basis of the separatrix q^\ddagger would return the same normalized population time-correlation function from Eq. (2) and the same reactive flux time-correlation function. In this case, the committor-based reactive flux of Eq. (45) will converge more rapidly than the activated-dynamics reactive flux time-correlation function Eq. (4).

E. Indicator-restricted committor time-correlation function

One of the most attractive features of the activated-dynamics reactive flux formalism as expressed by Eq. (4) is that it is based on a local sampling with trajectories initiated at the transition state, x^\ddagger , along some one-dimensional reaction coordinate.^{1,5,6} This is in contrast with the reactive flux expression based on the committor time-correlation function Eq. (23), which is based on an unconstrained equilibrium average over all possible initial positions within the subspace \mathbf{z} . Seeking to recover such a local sampling, we consider the steady-state reactive flux from A to B based on Eq. (21) expressed in terms of transitions from the point \mathbf{z} with committor $q(\mathbf{z}) < q^\ddagger$ to the point \mathbf{z}' with committor $q(\mathbf{z}') > q^\ddagger$,

$$J_{AB}^\ddagger = \frac{1}{\tau} \int d\mathbf{z} \int d\mathbf{z}' \theta(q^\ddagger - q(\mathbf{z})) \theta(q(\mathbf{z}') - q^\ddagger) \times (q(\mathbf{z}') - q(\mathbf{z})) \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}) = \frac{1}{\tau} \left\langle \theta(q(\tau) - q^\ddagger) \theta(q^\ddagger - q(0)) (q(\tau) - q(0)) \right\rangle. \quad (43)$$

This expression is formally correct for any value of $0 < q^\ddagger < 1$. For the sake of convenience, we choose the separatrix where $q^\ddagger = 0.5$. By identification, we can then write the state indicator functions as $\theta(q - q^\ddagger) = H_A$ and $\theta(q^\ddagger - q) = H_B$. We follow similar steps from Eq. (4), leading to the development of the activated-dynamics reactive flux formalism. Defining the indicator-restricted committor time-correlation function,

$$C_{AB-qq}(t) = \langle H_A(0) H_B(t) (q(t) - q(0)) \rangle, \quad (44)$$

the reactive flux at the boundary between the A and B regions can be written as

$$J_{AB}^\ddagger = \lim_{t \rightarrow \tau} \dot{C}_{AB-qq}(t) = \lim_{t \rightarrow \tau} \left\langle H_A(0) \frac{d}{dt} [H_B(t) (q(t) - q(0))] \right\rangle = - \lim_{t \rightarrow \tau} \left\langle \dot{H}_A(0) H_B(t) (q(t) - q(0)) \right\rangle = \left\langle \delta(q(0) - q^\ddagger) \dot{q}(0) H_B(\tau) (q(\tau) - q(0)) \right\rangle. \quad (45)$$

Like the activated-dynamics reactive flux formalism of Eq. (4), this expression is based on a local sampling of trajectories initiated at a given interface, with $q(\tilde{\mathbf{z}}[\mathbf{x}]) = q^\ddagger$. From this point of view, the expression focuses the calculation of the forward transition rate

constant at a transition state. By analogy with activated-dynamics, one may note that the forward rate in Eq. (4) could also be expressed from the flux at the interface between the state A and B as $k_{AB} = \langle H_A(0) \dot{H}_B(\tau_m) \rangle / p_A$. However, it is noteworthy that this form is not equivalent to simply substituting $x \rightarrow q$ in Eq. (4). Equation (45) is not the same as simply using the committor as a “reaction coordinate” in the reactive flux formalism. What is different is the term $[q(\tau) - q(0)]$, weighting differently the dynamical excursion away from the dividing surface between the states A and B , which greatly affects the convergence of the reactive flux. This committor-based term is present here because the fundamental expression for the reactive flux at the boundary between the A and B regions Eq. (21) is derived by including only local transitions $\mathbf{z} \rightarrow \mathbf{z}'$ that belong to reactive paths with genuine $A \rightarrow B$ transitions.¹²

To illustrate these ideas and further examine the convergence of the activated-dynamics reactive flux formalism and transition state theory,^{1,5,6} we consider again the time-correlation functions calculated from Langevin simulations in the case of the broad free energy barrier. The parameters were chosen such that the Langevin dynamics is in a high friction limit and becomes equivalent to Brownian dynamics within the short time-lag $\tau = 0.08$ ps. The time-correlation functions are shown in Fig. 3. The indicator function time-correlation function $C_{BB}(t)$ and its time-derivative $\dot{C}_{BB}(t)$ corresponding to the reactive flux are shown at the top and bottom, respectively. The initial value of $\dot{C}_{BB}(t)$ evaluated at $t = 0^+$ corresponds to p_A times the transition state rate k_{AB}^{TST} . The plateau value of $\dot{C}_{BB}(\tau_m)$ is equal to p_A times the forward transition rate k_{AB} , and the transmission coefficient $\kappa(\tau_m)$ is equal to $\dot{C}_{BB}(\tau_m) / \dot{C}_{BB}(0^+)$ via Eq. (4).

The indicator-restricted committor time-correlation function $C_{AB-qq}(t)$ is shown in Fig. 3 (top). It essentially mirrors the evolution of the committor time-correlation function $C_{qq}(t)$, rapidly converging toward the same slope. The reactive flux (bottom) shows that, like $\dot{C}_{qq}(t)$, $\dot{C}_{AB-qq}(t)$ reach a plateau within the lag-time τ of ~ 0.08 ps (see the inset). In contrast, the activated-dynamics reactive flux $\dot{C}_{BB}(t)$ decays to reach its plateau value after a considerably longer time τ_m (~ 12 ps in Fig. 3). The position time-correlation function $\dot{C}_{zz}(t)$ rises sharply and reaches its plateau faster than $\dot{C}_{BB}(t)$ but still more slowly than $\dot{C}_{qq}(t)$ and $\dot{C}_{AB-qq}(t)$. The behavior of $C_{AB-qq}(t)$ displays some very interesting similarity with the TPS study of van Erp *et al.*,¹⁸ who showed that the time-derivative of the population correlation function constructed from indicator functions that depends on the history of the trajectories (past and future) also reached the correct plateau value at very short time. Whereas a trajectory’s history (past and future) is explicitly visible in TPS,^{16,17} such information in TPT is incorporated through the committor probability function.¹⁴ Accordingly, it is the term $(q(\tau) - q(0))$ in the TPT-based expression for $C_{AB-qq}(t)$ of Eq. (45), which carries information about a trajectory’s history and insures that only members of the transition paths ensemble are included, that dramatically affects the convergence of the reactive flux. While further analysis would be needed to formally relate the present analysis to TPS simulations,¹⁸ this highlights the importance of attaching some information about the reactive or non-reactive “fate” of a given trajectory to determine a transition rate, a concept that has been well understood, from the early days of transition state theory (TST)²⁻⁴ to more modern developments of the reactive flux formalism.³¹

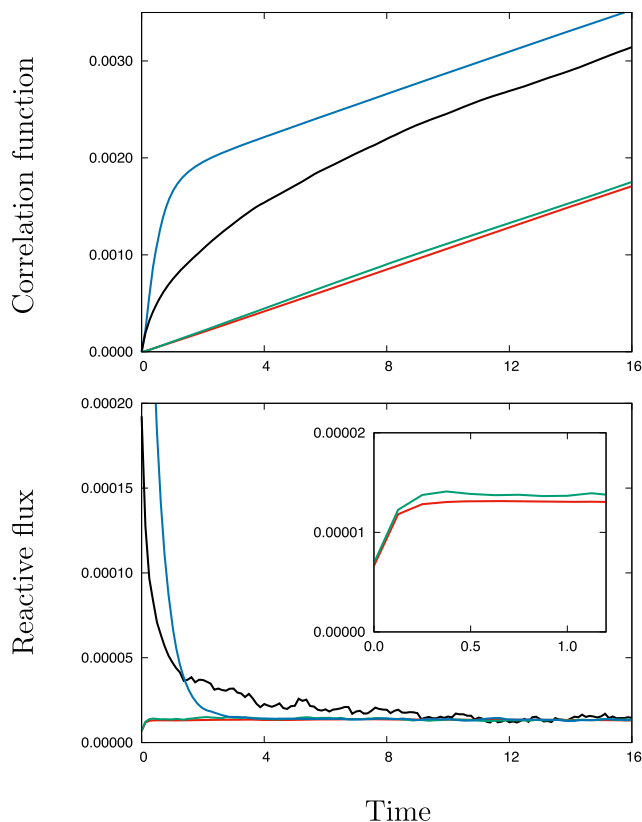


FIG. 3. Time-correlation functions for the simple two-state system with a broad free energy barrier (shown in Fig. 1). In the top, the indicator autocorrelation function $C_{BB}(t)$ (in black), the position autocorrelation function $C_{zz}(t)$ (in blue), the committor time-correlation function $C_{qq}(t)$ (in red), and the indicator-restricted committor time-correlation function $C_{AB-qq}(t)$ (in green) is shown. The time axis is in ps. All the time-correlation functions converge to the same slope, which is related to the steady-state forward reactive flux J_{AB} (top). The time-derivative of the time-correlation functions yielding the reactive flux are plotted (bottom) to illustrate how the slope relaxes to a constant plateau value corresponding to $J_{AB} = p_A k_{AB} = 1.05 \times 10^{-4} \text{ ps}^{-1}$, close to the value for the broad barrier obtained by Brownian dynamics reported in Fig. 2.

The rapid convergence of the indicator-restricted committor time-correlation function $C_{AB-qq}(t)$ suggests that it may be possible to formulate an algorithm based on Eq. (45) that would rely on a local sampling like the activated-dynamics reactive flux formalism,^{1,5,6} but which would converge more rapidly toward the correct transition rate constant. This advantage, however, needs to be qualified. Although such an algorithm would converge within a short lag-time, an accurate estimate of the committor is required. The latter can be calculated for a given configuration by shooting trajectories to determine the probability of reaching state B before state A , although the most practical approach is to optimize a trial function $q(\mathbf{z})$ via the variational principle based on Eqs. (23) and (24).^{12,34,35} The important question is whether the usage of an approximate committor, which introduces additional factors associated high order eigenvalues as shown by Eq. (32), can be utilized productively.

V. CONCLUSION

The kinetic relaxation of a complex dynamical system dominated by two metastable states was examined from a Markov state eigenvalue spectral decomposition,^{8–11} a committor-based steady-state reactive flux,^{7,13,14} and the activated-dynamics reactive flux formalism.^{1,5,6} The present analysis was based on the assumption that the metastable states A and B were already known. Generally, discovering, identifying, and locating such metastable states in complex systems, however, can be a difficult challenge requiring specialized methods.^{48–51} Importantly, the robustness of the analysis and the overall shape of the committor function is typically not affected by the precise definition of the boundary states—this is one reason why the same definition was used for all three examples shown in Fig. 1 with very different free energy barriers.

The analysis showed that the different theoretical perspectives are consistent. This is most clearly displayed by the close correspondence between the eigenmode relaxation of Eq. (20), the steady-state reactive flux via Eq. (31), and the activated-dynamics formalism via Eq. (2). The spectral analysis of the effective propagator \mathcal{P}_τ yields the eigenvalue $\lambda_k(\tau) = e^{-\mu_k \tau}$, with the second eigenvalue μ_2 corresponding to the global relaxation time $\tau^* = (k_{AB} + k_{BA})^{-1} = (\mu_2)^{-1}$ of the two-state system, while the net steady-state reactive flux from A to B is $J_{AB} = p_A k_{AB}$.

The significance of the different time scales that appear in the analysis of the kinetic relaxation time of a two-state system was clarified. While Δt is the microscopic time step used to generate the trajectory of the system, $\tau \gg \Delta t$ is the shortest lag-time that achieves Markovity for the effective propagator \mathcal{P}_τ within the subspace \mathbf{z} . The analysis also revealed the existence of an intermediate time scale τ_m , which is associated with additional molecular processes that must decay to display the true global relaxation time of the two-state system. The intermediate time scale τ_m is larger than the lag-time τ but much smaller than the global relaxation time τ^* . According to the spectral analysis, a good separation of time scales can be achieved if there is a large gap between the eigenvalues μ_2 and μ_3 such that $e^{-\mu_3 \tau_m} \ll 1$ while $e^{-\mu_2 \tau_m} \approx 1$. In particular, this intermediate time scale τ_m is displayed by the decay of the activated-dynamics reactive flux time-correlation function to a constant plateau corresponding to the transmission coefficient κ . However, it can also appear when an imperfect approximation for the committor is used in the expression for the steady-state reactive flux. In fact, there is a close correspondence between these two situations because the activated-dynamics reactive flux formalism is akin to using a Heaviside indicator function (population operator) to approximate the committor. This explains why the reactive flux with the indicator function is an effective route for a narrow sharp barrier but much less so in the case of a broad flat barrier. In this case, for example, the activated-dynamics reactive flux time-correlation function decays slowly to its plateau because the indicator function is affected by the diffusion process on top of the flat barrier.

An attractive feature of the activated-dynamics reactive flux formalism is the ability to calculate the forward transition rate constant from a local sampling of trajectories initiated at a given interface.^{1,5,6} While this provides an effective route to calculate the transition rate in the case of a narrow sharp barrier, it is much less effective in the case of a broad flat barrier. In contrast, the committor-based steady-state reactive flux rapidly converges to the

correct value at the lag-time τ for both sharp and broad free energy barriers when the exact committor is used. The rapid convergence to the correct reactive flux within the time-lag τ and the absence of intermediate time scale τ_m is an attractive and robust feature of a committor-based steady-state reactive flux formulation. A tantalizing conjecture is whether this observation can serve to ascertain the Markovity of the effective propagator \mathcal{P}_τ in the reduced subspace of the CVs.

It was shown that the most effective choice to define a one-dimensional reaction coordinate in the region of the transition state is along the gradient of the committor, $\nabla q(\mathbf{z})$, in the multi-dimensional subspace of collective variables \mathbf{z} . However, while the committor can be exploited as a useful reaction coordinate, simply treating the committor as a one-dimensional reaction coordinate in the activated-dynamics formalism does not alleviate the issues caused by slow relaxation and intermediate time scales. In that sense, the idea that the committor is the “perfect” reaction coordinate, as is sometimes suggested, needs to be interpreted carefully. It is possible to express the reactive flux as a local average via the indicator-restricted committor time-correlation function $C_{AB-qq}(t)$ in Eq. (45), which is one of the most attractive features of the activated-dynamics reactive flux formalism.^{1,5,6} Like Eq. (4), the forward transition rate constant can be calculated from a local sampling of very short trajectories initiated at a separating interface.

There is a large body of methods seeking to identify optimal reaction coordinates from complex dynamics,^{11,33,47,52,53} including dynamical self-consistency,⁴⁶ memory reduction,⁵⁴ multi-dimensional spectral gap optimization of order parameters (SGOOP),^{55,56} maximally predictive one-dimensional projection,¹⁰ and variational committor-based steady-state reactive flux.^{12,34,35} Also related are the dynamical Galerkin approximation (DGA) formulated to predict the long-timescale behavior from short-trajectory^{34,35} and nonparametric variational optimization of reaction coordinates.^{57,58} While challenging, the determination of an optimal reaction coordinate remains of great value, and a long term objective is to use these ideas to provide further guidance to theoretical frameworks built upon an optimized pathway.^{29,59,60}

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) through Grant No. MCB-1517221. The author would like to acknowledge useful discussions and feedback from Attila Szabo, Giovanni Ciccotti, Sergei Krivov, Jonathan Weare, Aaron Dinner, Spencer Guo, Ziwei Ge, Tony Lelièvre, and Chris Chipot.

AUTHOR DECLARATIONS

Conflict of Interest

The author has no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available within the article.

APPENDIX A: DIFFUSIVE DYNAMICS AND COMMITTOR

If the dynamics is diffusive, we have⁶¹

$$\begin{aligned} J_{AB} &= - \int d\mathbf{z} \int d\mathbf{z}_0 q(\mathbf{z}) \left(\frac{d}{d\tau} \mathcal{P}_\tau(\mathbf{z}|\mathbf{z}_0)|_{\tau=0} \right) q(\mathbf{z}_0) \rho_{\text{eq}}(\mathbf{z}_0) \\ &= - \int d\mathbf{z} q(\mathbf{z}) \int d\mathbf{z}_0 \nabla \cdot \left[\mathbf{D} \rho_{\text{eq}}(\mathbf{z}) \nabla \left(\frac{\delta(\mathbf{z} - \mathbf{z}_0)}{\rho_{\text{eq}}(\mathbf{z})} \right) \right] q(\mathbf{z}_0) \rho_{\text{eq}}(\mathbf{z}_0) \\ &= - \int d\mathbf{z} q(\mathbf{z}) \nabla \cdot \left[\mathbf{D} \rho_{\text{eq}}(\mathbf{z}) \nabla \left(\int d\mathbf{z}_0 \frac{\delta(\mathbf{z} - \mathbf{z}_0)}{\rho_{\text{eq}}(\mathbf{z})} q(\mathbf{z}_0) \rho_{\text{eq}}(\mathbf{z}_0) \right) \right] \\ &= - \int d\mathbf{z} q(\mathbf{z}) \nabla \cdot [\mathbf{D} \rho_{\text{eq}}(\mathbf{z}) \nabla q(\mathbf{z})], \end{aligned} \quad (\text{A1})$$

where the identity

$$\frac{d}{d\tau} \mathcal{P}_\tau(\mathbf{z}|\mathbf{z}_0)|_{\tau=0} = \nabla \cdot \left[\mathbf{D} \rho_{\text{eq}}(\mathbf{z}_0) \nabla \left(\frac{\delta(\mathbf{z} - \mathbf{z}_0)}{\rho_{\text{eq}}(\mathbf{z}_0)} \right) \right] \quad (\text{A2})$$

obtained from

$$\frac{d}{d\tau} \mathcal{P}_\tau(\mathbf{z}|\mathbf{z}_0) = \nabla \cdot \left[\mathbf{D} \rho_{\text{eq}}(\mathbf{z}_0) \nabla \left(\frac{\mathcal{P}_\tau(\mathbf{z}|\mathbf{z}_0)}{\rho_{\text{eq}}(\mathbf{z}_0)} \right) \right] \quad (\text{A3})$$

and $\mathcal{P}_0(\mathbf{z}|\mathbf{z}_0) = \delta(\mathbf{z} - \mathbf{z}_0)$, has been used. It should be noted that the integrals are over the whole space rather than just the intermediate region to account for the discontinuity at the boundaries.

APPENDIX B: MODEL COMMITTOR

The equilibrium population of the states A and B is given by

$$p_B \approx \int d\mathbf{z} \rho_{\text{eq}}(\mathbf{z}) q(\mathbf{z}) \quad (\text{B1})$$

and $p_B = (1 - p_A)$. When a very long trajectory is “colored” based on transitions to the two boundary regions, Eq. (B1) correctly provides the probability of the color ascribed to B .⁷ Using the model committor from Eq. (30), the backward propagation condition is^{7,37}

$$\begin{aligned} \int d\mathbf{z}' q(\mathbf{z}') \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) &= - \left(\frac{a}{b-a} \right) \int d\mathbf{z}' \psi_1^L(\mathbf{z}') \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \\ &\quad + \left(\frac{1}{b-a} \right) \int d\mathbf{z}' \psi_2^L(\mathbf{z}') \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \\ &= - \left(\frac{a}{b-a} \right) \psi_1^L(\mathbf{z}) + \left(\frac{1}{b-a} \right) \lambda_2 \psi_2^L(\mathbf{z}) \\ &= q(\mathbf{z}) + \left(\frac{\lambda_2 - 1}{b-a} \right) \psi_2^L(\mathbf{z}) \\ &\approx q(\mathbf{z}), \end{aligned} \quad (\text{B2})$$

where $|(\lambda_2 - 1)/(b - a)| \ll 1$ (see below). We note that the vector $\psi_1^L(\mathbf{z}) = 1$ and the vector $\psi_1^R(\mathbf{z}) = \rho_{\text{eq}}(\mathbf{z})$. From the construction,

$$\psi_2^L(\mathbf{z}) = (b - a)q(\mathbf{z}) + a\psi_1^L(\mathbf{z}). \quad (\text{B3})$$

The coefficients a and b are determined by the orthogonality and normalization conditions, $(\psi_1^L \cdot \psi_2^R) = 0$ and $(\psi_2^L \cdot \psi_2^R) = 1$. By virtue of orthogonality, we have

$$\begin{aligned} 0 &= \int d\mathbf{z} \psi_2^L(\mathbf{z}) \psi_1^R(\mathbf{z}) \\ &= (b - a) \int d\mathbf{z} q(\mathbf{z}) \psi_1^R(\mathbf{z}) + a \int d\mathbf{z} \psi_1^L(\mathbf{z}) \psi_1^R(\mathbf{z}) \\ &= (b - a)p_B + a. \end{aligned} \quad (\text{B4})$$

Hence, $a = -(b - a)p_B$. By virtue of normalization, we have

$$\begin{aligned} 1 &= \int d\mathbf{z} \psi_2^L(\mathbf{z}) \psi_2^R(\mathbf{z}) \\ &= \int d\mathbf{z} \psi_2^L(\mathbf{z}) \rho_{\text{eq}}(\mathbf{z}) \psi_2^L(\mathbf{z}) \\ &= \int d\mathbf{z} \rho_{\text{eq}}(\mathbf{z}) ((b - a)q(\mathbf{z}) + a\psi_1^L(\mathbf{z}))^2 \\ &= (b - a)^2 \langle qq \rangle + 2(b - a)a \langle q \rangle + a^2 \\ &= (b - a)^2 (p_B - \langle q(1 - q) \rangle) + 2(b - a)a p_B + a^2 \\ &= (b - a)^2 (p_A p_B - \langle q(1 - q) \rangle). \end{aligned} \quad (\text{B5})$$

This shows that $|(\lambda_2 - 1)/(b - a)| \approx \mu_2 \tau \sqrt{p_A p_B} \ll 1$.

APPENDIX C: REACTIVE FLUX FROM THE MODEL COMMITTOR

Using the model committor from Eq. (30) and the conditions established in Appendix B, the steady-state flux from A to B is

$$\begin{aligned} J_{AB} &= -\frac{1}{\tau} \left(\int d\mathbf{z} \int d\mathbf{z}' q(\mathbf{z}') \rho_{\text{eq}}(\mathbf{z}) \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) q(\mathbf{z}) - \langle qq \rangle \right) \\ &= -\frac{1}{\tau} \left(\frac{a^2}{(b - a)^2} \int d\mathbf{z} \int d\mathbf{z}' \psi_1^L(\mathbf{z}') \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \psi_1^R(\mathbf{z}) \right. \\ &\quad \left. + \frac{1}{(b - a)^2} \int d\mathbf{z} \int d\mathbf{z}' \psi_2^L(\mathbf{z}') \mathcal{P}_\tau(\mathbf{z}'|\mathbf{z}) \psi_2^R(\mathbf{z}) \right. \\ &\quad \left. - p_B + \langle q(1 - q) \rangle \right) \\ &= -\frac{1}{\tau} (p_B^2 + (p_A p_B - \langle q(1 - q) \rangle) \lambda_2 - p_B + \langle q(1 - q) \rangle) \\ &= \frac{1}{\tau} (p_A p_B (1 - \lambda_2) - \langle q(1 - q) \rangle (1 - \lambda_2)). \end{aligned} \quad (\text{C1})$$

Since $\langle q(1 - q) \rangle = \langle \tau_r \rangle J_{AB}$, we have

$$J_{AB} = \frac{1}{\tau} p_A p_B (1 - \lambda_2) \left(1 + \frac{\langle \tau_r \rangle}{\tau} (1 - \lambda_2) \right)^{-1} \quad (\text{C2})$$

if $\mu_2 \tau \ll 1$, leading to Eq. (31).

REFERENCES

- D. Chandler, "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation," *J. Chem. Phys.* **68**, 2959–2970 (1978).
- W. Wynne-Jones and H. Eyring, "The absolute rate of reactions in condensed phases," *J. Chem. Phys.* **3**(8), 492–502 (1935).
- E. Wigner, "The transition state method," *Trans. Faraday Soc.* **34**(1), 29–41 (1938).
- K. J. Laidler and M. C. King, "The development of transition-state theory," *J. Phys. Chem.* **87**(15), 2657–2664 (1983).
- B. J. Berne, "Molecular dynamics and Monte Carlo simulations of rare events," in *Multiple Time Scales*, edited by J. U. Brackbill and B. I. Cohen (Academic Press, New York, 1985), pp. 419–436.
- B. J. Berne, M. Borkovec, and J. E. Straub, "Classical and modern methods in reaction rate theory," *J. Phys. Chem.* **92**, 3711–3725 (1988).
- A. M. Berezhkovskii and A. Szabo, "Committers, first-passage times, fluxes, Markov states, milestones, and all that," *J. Chem. Phys.* **150**(5), 054106 (2019).
- F. Nueske, B. G. Keller, G. Perez-Hernandez, A. S. J. S. Mey, and F. Noe, "Variational approach to molecular kinetics," *J. Chem. Theory Comput.* **10**(4), 1739–1752 (2014).
- H. Wu, F. Nüske, F. Paul, S. Klus, P. Koltai, and F. Noé, "Variational Koopman models: Slow collective variables and molecular kinetics from short off-equilibrium simulations," *J. Chem. Phys.* **146**(15), 154104 (2017).
- R. T. McGibbon, B. E. Husic, and V. S. Pande, "Identification of simple reaction coordinates from complex dynamics," *J. Chem. Phys.* **146**(4), 044109 (2017).
- S. V. Krivov, "Protein folding free energy landscape along the committor—The optimal folding coordinate," *J. Chem. Theory Comput.* **14**(7), 3418–3427 (2018).
- B. Roux, "String method with swarms-of-trajectories, mean drifts, lag time, and committor," *J. Phys. Chem. A* **125**(34), 7558–7571 (2021).
- A. Berezhkovskii, G. Hummer, and A. Szabo, "Reactive flux and folding pathways in network models of coarse-grained protein dynamics," *J. Chem. Phys.* **130**(20), 205102 (2009).
- W. E and E. Vanden-Eijnden, "Transition-path theory and path-finding algorithms for the study of rare events," *Annu. Rev. Phys. Chem.* **61**, 391–420 (2010).
- A. M. Berezhkovskii and A. Szabo, "Diffusion along the splitting/commitment probability reaction coordinate," *J. Phys. Chem. B* **117**(42), 13115–13119 (2013).
- C. Dellago, P. Bolhuis, F. Csajka, and D. Chandler, "Transition path sampling and the calculation of rate constants," *J. Chem. Phys.* **108**(5), 1964–1977 (1998).
- P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, "Transition path sampling: Throwing ropes over mountain passes, in the dark," *Annu. Rev. Phys. Chem.* **53**, 291 (2002).
- T. S. van Erp, D. Moroni, and P. G. Bolhuis, "A novel path sampling method for the calculation of rate constants," *J. Chem. Phys.* **118**(17), 7762–7774 (2003).
- D. Moroni, P. G. Bolhuis, and T. S. van Erp, "Rate constants for diffusive processes by partial path sampling," *J. Chem. Phys.* **120**(9), 4055–4065 (2004).
- G. Hummer, "From transition paths to transition states and rate coefficients," *J. Chem. Phys.* **120**(2), 516–523 (2004).
- L. Onsager, "Initial recombination of ions," *Phys. Rev.* **54**(8), 554–557 (1938).
- R. Du, V. Pande, A. Grosberg, T. Tanaka, and E. Shakhnovich, "On the transition coordinate for protein folding," *J. Chem. Phys.* **108**(1), 334–350 (1998).
- P. Geissler, C. Dellago, and D. Chandler, "Kinetic pathways of ion pair dissociation in water," *J. Phys. Chem. B* **103**(18), 3706–3710 (1999).
- P. G. Bolhuis, C. Dellago, and D. Chandler, "Reaction coordinates of biomolecular isomerization," *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5877–5882 (2000).
- W. E, W. Ren, and E. Vanden-Eijnden, "String method for the study of rare events," *Phys. Rev. B* **66**(5), 052301 (2002).
- W. E, W. Ren, and E. Vanden-Eijnden, "Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes," *Chem. Phys. Lett.* **413**(1–3), 242–247 (2005).
- W. E, W. Ren, and E. Vanden-Eijnden, "Simplified and improved string method for computing the minimum energy paths in barrier-crossing events," *J. Chem. Phys.* **126**(16), 164103 (2007).

- ²⁸L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, "String method in collective variables: Minimum free energy paths and isocommittor surfaces," *J. Chem. Phys.* **125**, 024106 (2006).
- ²⁹A. C. Pan, D. Sezer, and B. Roux, "Finding transition pathways using the string method with swarms of trajectories," *J. Phys. Chem. B* **112**, 3432–3440 (2008).
- ³⁰B. Roux, *Computational Modeling and Simulations of Biomolecular Systems* (World Scientific Publishing, Singapore, 2021).
- ³¹J. Anderson, "Predicting rare events in molecular dynamics," *Adv. Chem. Phys.* **91**, 381–431 (1995).
- ³²M. Ruiz-Montero, D. Frenkel, and J. Brey, "Efficient schemes to compute diffusive barrier crossing rates," *Mol. Phys.* **90**(6), 925–941 (1997).
- ³³S. V. Krivov, "On reaction coordinate optimality," *J. Chem. Theory Comput.* **9**(1), 135–146 (2013).
- ³⁴E. H. Thiede, D. Giannakis, A. R. Dinner, and J. Weare, "Galerkin approximation of dynamical quantities using trajectory data," *J. Chem. Phys.* **150**(24), 244111 (2019).
- ³⁵C. Lorpaiboon, E. H. Thiede, R. J. Webber, J. Weare, and A. R. Dinner, "Integrated variational approach to conformational dynamics: A robust strategy for identifying eigenfunctions of dynamical operators," *J. Phys. Chem. B* **124**(42), 9354–9364 (2020).
- ³⁶D. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).
- ³⁷A. Berezhkovskii and A. Szabo, "Ensemble of transition states for two-state protein folding from the eigenvectors of rate matrices," *J. Chem. Phys.* **121**(18), 9186–9187 (2004).
- ³⁸G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, "Identification of slow molecular order parameters for Markov model construction," *J. Chem. Phys.* **139**(1), 015102 (2013).
- ³⁹C. R. Schwantes and V. S. Pande, "Improvements in Markov state model construction reveal many non-native interactions in the folding of NTL9," *J. Chem. Theory Comput.* **9**(4), 2000–2009 (2013).
- ⁴⁰E. Vanden-Eijnden, *Transition Path Theory*, Computer Simulations in Condensed Matter: From Materials to Chemical Biology Vol. 2 (Springer, 2006), p. 439.
- ⁴¹D. Truhlar, B. Garrett, and S. Klippenstein, "Current status of transition-state theory," *J. Phys. Chem.* **100**(31), 12771–12800 (1996).
- ⁴²A. Berezhkovskii and A. Szabo, "One-dimensional reaction coordinates for diffusive activated rate processes in many dimensions," *J. Chem. Phys.* **122**(1), 014503 (2005).
- ⁴³R. Zwanzig, "Memory effects in irreversible thermodynamics," *Phys. Rev.* **124**(4), 983 (1961).
- ⁴⁴H. Mori, "Transport, collective motion, and Brownian motion," *Prog. Theor. Phys.* **33**, 423–455 (1965).
- ⁴⁵R. Kubo, "The fluctuation-dissipation theorem," *Rev. Mod. Phys.* **29**, 255–284 (1966).
- ⁴⁶B. Peters, P. G. Bolhuis, R. G. Mullen, and J. E. Shea, "Reaction coordinates, one-dimensional Smoluchowski equations, and a test for dynamical self-consistency," *J. Chem. Phys.* **138**(5), 054106 (2013).
- ⁴⁷P. V. Banushkina and S. V. Krivov, "Optimal reaction coordinates," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **6**(6), 748–763 (2016).
- ⁴⁸P. Deuffhard and M. Weber, "Robust Perron cluster analysis in conformational dynamics," *Linear Algebra Appl.* **398**(Suppl. C), 161–184 (2005).
- ⁴⁹S. Röblitz and M. Weber, "Fuzzy spectral clustering by PCCA+: Application to markov state models and data classification," *Adv. Data. Anal. Classif.* **7**(2), 147–179 (2013).
- ⁵⁰A. Mardt, L. Pasquali, H. Wu, and F. Noé, "VAMPnets for deep learning of molecular kinetics," *Nat. Commun.* **9**(1), 5 (2018).
- ⁵¹D. Wang and P. Tiwary, "State predictive information bottleneck," *J. Chem. Phys.* **154**(13), 134111 (2021).
- ⁵²W. Li and A. Ma, "Recent developments in methods for identifying reaction coordinates," *Mol. Simul.* **40**(10–11), 784–793 (2014).
- ⁵³B. Peters, "Reaction coordinates and mechanistic hypothesis tests," *Annu. Rev. Phys. Chem.* **67**, 669–690 (2016).
- ⁵⁴N. Guttenberg, J. F. Dama, M. G. Saunders, G. A. Voth, J. Weare, and A. R. Dinner, "Minimizing memory as an objective for coarse-graining," *J. Chem. Phys.* **138**(9), 094111 (2013).
- ⁵⁵P. Tiwary and B. J. Berne, "Spectral gap optimization of order parameters for sampling complex molecular systems," *Proc. Natl. Acad. Sci. U. S. A.* **113**(11), 2839–2844 (2016).
- ⁵⁶Z. Smith, D. Pramanik, S.-T. Tsai, and P. Tiwary, "Multi-dimensional spectral gap optimization of order parameters (SGOOP) through conditional probability factorization," *J. Chem. Phys.* **149**(23), 234105 (2018).
- ⁵⁷P. V. Banushkina and S. V. Krivov, "Nonparametric variational optimization of reaction coordinates," *J. Chem. Phys.* **143**(18), 184108 (2015).
- ⁵⁸S. V. Krivov, "Blind analysis of molecular dynamics," *J. Chem. Theory Comput.* **17**(5), 2725–2736 (2021).
- ⁵⁹A. C. Pan and B. Roux, "Building Markov state models along pathways to determine free energies and rates of transitions," *J. Chem. Phys.* **129**(6), 064107 (2008).
- ⁶⁰A. K. Faradjian and R. Elber, "Computing time scales from reaction coordinates by milestoning," *J. Chem. Phys.* **120**, 10880–10889 (2004).
- ⁶¹A. Szabo, personal communication (2021).