

# Estimating Position-Dependent and Anisotropic Diffusivity Tensors from Molecular Dynamics Trajectories: Existing Methods and Future Outlook

Tiago S. Domingues, Ronald Coifman, and Amir Haji-Akbari\*



Cite This: <https://doi.org/10.1021/acs.jctc.4c00148>



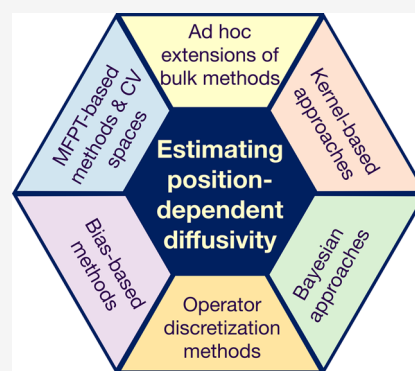
Read Online

ACCESS |

Metrics & More

Article Recommendations

**ABSTRACT:** Confinement can substantially alter the physicochemical properties of materials by breaking translational isotropy and rendering all physical properties position-dependent. Molecular dynamics (MD) simulations have proven instrumental in characterizing such spatial heterogeneities and probing the impact of confinement on materials' properties. For static properties, this is a straightforward task and can be achieved via simple spatial binning. Such an approach, however, cannot be readily applied to transport coefficients due to lack of natural extensions of autocorrelations used for their calculation in the bulk. The prime example of this challenge is diffusivity, which, in the bulk, can be readily estimated from the particles' mobility statistics, which satisfy the Fokker–Planck equation. Under confinement, however, such statistics will follow the Smoluchowski equation, which lacks a closed-form analytical solution. This brief review explores the rich history of estimating profiles of the diffusivity tensor from MD simulations and discusses various approximate methods and algorithms developed for this purpose. Besides discussing heuristic extensions of bulk methods, we overview more rigorous algorithms, including kernel-based methods, Bayesian approaches, and operator discretization techniques. Additionally, we outline methods based on applying biasing potentials or imposing constraints on tracer particles. Finally, we discuss approaches that estimate diffusivity from mean first passage time or committor probability profiles, a conceptual framework originally developed in the context of collective variable spaces describing rare events in computational chemistry and biology. In summary, this paper offers a concise survey of diverse approaches for estimating diffusivity from MD trajectories, highlighting challenges and opportunities in this area.



## 1. INTRODUCTION

Confinement refers to situations in which a material is encapsulated by one or more physical interfaces, and can trigger substantial changes to its physical properties. Such changes arise due to the breaking of both translational and rotational symmetry, which renders all physical properties functions of position. Such position dependence becomes most pronounced within an interfacial region that has a characteristic thickness of a few diameters of the material's building blocks. Whenever the confinement length scale, i.e., the smallest separation between confining interfaces, is also comparable to the interfacial length scale, i.e., the thickness of the interfacial region, materials' properties exhibit the most pronounced deviations from the bulk.<sup>1</sup> Depending on the size of these building blocks and the range of their interactions with the interface, such deviations can be observed at different length scales. Confinement, therefore, is a potent means of fine-tuning the thermodynamic,<sup>2–8</sup> structural,<sup>9–12</sup> and transport<sup>13–20</sup> properties of materials, while also influencing the kinetics and mechanisms of rare events.<sup>21–27</sup> Among the spectrum of physical properties that can be substantially

impacted by confinement, transport properties stand out prominently, as properties such as diffusivity,<sup>19</sup> viscosity,<sup>13,14,16</sup> and thermal<sup>18</sup> and ionic<sup>17,20</sup> conductivity experience substantial alterations under confinement.

Since their advent in the mid-20th century,<sup>28–32</sup> molecular simulations have emerged as indispensable tools for studying confined states of matter,<sup>33</sup> and computational studies characterizing the position dependence of physical properties within confined materials can be traced back to 1970s.<sup>34–39</sup> It is fairly straightforward to determine spatial profiles of the thermodynamic and structural properties that can be unambiguously computed for a particular region within the simulation box from a single snapshot. This is simply achieved by partitioning the simulation box into suitable spatial bins and

**Received:** February 4, 2024

**Revised:** April 22, 2024

**Accepted:** April 30, 2024

estimating the mean of the quantity of interest within each bin via a combination of time- and ensemble-averaging. However, applying this approach to transport properties is not feasible, as those are typically computed using autocorrelations of appropriate mechanical observables, and autocorrelations of such nature cannot be unambiguously defined for open systems. Therefore, generalizing the linear response formalisms, originally devised for translationally isotropic materials, to confined materials is a nontrivial undertaking.

This review is primarily dedicated to exploring methodologies for computing spatial profiles of transport properties from molecular simulations of confined materials, with a particular emphasis on diffusivity—arguably the most widely computed transport property in molecular simulations. Indeed, reports of self-diffusivity calculations can be traced back to the early days of molecular simulations, such as Rahman's pioneering work in 1964.<sup>32</sup> We first describe the fundamental framework employed for computing transport properties in the bulk, and discuss the challenges of generalizing such frameworks to confined geometries. We then overview different classes of strategies aimed at effectively estimating position-dependent anisotropic diffusivity profiles.

This paper is organized as follows. In Section 2, we provide phenomenological definitions of transport coefficients, overview the computational methodologies employed for their estimation from MD, and discuss the specific challenges associated with estimating them in confined geometries. Section 3 describes *ad hoc* extensions of classical methods, with brief discussions of their plausible theoretical foundations. The subsequent sections navigate more rigorous approaches, including kernel-based methods (Section 4), Bayesian techniques (Section 5), and operation discretization approaches (Section 6). A comprehensive overview of methods based on applying biasing potentials to tracer particles is presented in Section 7. In Section 8, we shift focus to using mean first passage time and committor probability profiles to estimate diffusivity, methodologies developed in the context of the mathematically associated problem of characterizing diffusive behavior along collective variables employed for describing rare events. Finally, Section 9 is dedicated to providing a broader perspective and outlining potential avenues for future exploration.

## 2. ESTIMATING TRANSPORT PROPERTIES FROM MOLECULAR DYNAMICS TRAJECTORIES

**2.1. Estimators in the Bulk.** Transport coefficients are phenomenological constants that establish a connection between macroscopic fluxes of physical properties, and external fields or thermodynamic driving forces. More precisely, suppose a system simultaneously exposed to a sequence of sufficiently small gradients,  $\nabla X_1, \nabla X_2, \dots, \nabla X_p$ . Generally, there is always a natural (conjugate) flux associated with each such gradient. For instance, a temperature gradient will result in a heat flux if no other gradient is present. However, in the presence of multiple gradients, the flux associated with the  $k$ -th property,  $J_k$ , will, in principle, be linked to all other gradients through the following relationship:

$$J_k = - \sum_{j=1}^p L_{kj} \cdot \nabla X_j \quad (1)$$

Here,  $L_{kj}$ 's, which are tensors of appropriate ranks, are referred to as *transport coefficients*. Note that for every  $j$  and  $k$ ,  $L_{jk} = L_{kj}$

according to Onsager's reciprocity principle.<sup>40</sup> At a microscopic level, these coefficients can be viewed as rates at which a system responds to microscopic fluctuations.

While transport coefficients can, in principle, be defined for any pair of fluxes and driving forces, certain transport coefficients hold particular significance for physicists and materials scientists, and are widely reported in both experimental and computational studies of materials. One such transport coefficient is *diffusivity*, which is historically defined for multicomponent systems. More precisely,  $D_i$ , the diffusivity of component  $i$ , establishes a connection between its diffusive flux,  $J_i$ , and its concentration gradient:

$$J_i = -D_i \cdot \nabla c_i$$

In general,  $D_i$  is a second-rank symmetric positive-definite tensor, but in the case of bulk simple liquids, it often exhibits isotropic behavior. Similarly, the *self-diffusivity* of a pure material can be defined as the proportionality factor relating the flux and the gradient of the conditional probability of particle displacements. Serving as a proxy for the rate of structural relaxation in materials, self-diffusivity is well-defined but is difficult to measure experimentally. Nevertheless, it stands as one of the most widely computed transport coefficients in molecular simulation studies.

Analogous constitutive relationships govern the relationships between heat flux,  $\mathbf{q}$ , and temperature gradient,  $\nabla T$ , as well as electric current,  $\mathbf{i}$ , and electrostatic potential gradient,  $\nabla \varphi$ :

$$\mathbf{q} = -\mathbf{k} \cdot \nabla T$$

$$\mathbf{i} = -\boldsymbol{\sigma} \cdot \nabla \varphi$$

Here,  $\mathbf{k}$  and  $\boldsymbol{\sigma}$ , both second-rank symmetric positive-definite tensors, denote the *heat conductivity* and *electrical conductivity*, respectively. Finally, a linear relationship can also be postulated between the stress tensor,  $\boldsymbol{\tau}$ , and the symmetric part of the shear rate tensor,  $\nabla \mathbf{v} + \nabla \mathbf{v}^T$ :

$$\tau_{ij} = \mu_{ijkl} \left[ \frac{\partial v_k}{\partial x_l} + \frac{\partial v_l}{\partial x_k} \right]$$

The proportionality constant is referred to as *viscosity* and is naturally a fourth-rank tensor.<sup>41</sup>

In principle, all transport coefficients can be estimated from nonequilibrium MD simulations<sup>42,43</sup> in which an external driving force, such as temperature gradient, shear deformation, or electric field, is applied to the system. The resulting macroscopic fluxes can then be readily computed, providing a means for estimating the relevant proportionality constants.<sup>44–47</sup> Conversely, a constant-flux boundary condition can be imposed on the system, with the transport coefficient estimated from the computed conjugate gradient.<sup>48,49</sup> The major limitation of all such nonequilibrium approaches is the substantial magnitudes of the necessary driving forces (or fluxes), which often surpass their experimental counterparts by several orders of magnitude. This raises profound questions regarding the validity of the linear flux-driving force relationships under such extreme conditions. It is therefore unclear whether the proportionality constants estimated from such simulations will be applicable to experimentally relevant conditions.

An alternative approach, conceptually akin to nonequilibrium techniques in the limit of small gradients, involves utilizing linear response theory<sup>50</sup> to express transport

coefficients in terms of autocorrelations of microscopic fluxes. Linear response theory quantifies a system's response when its Hamiltonian is perturbed from equilibrium by a small external field, i.e.,  $\mathcal{H}(\mathbf{P}, \mathbf{Q}, t) = \mathcal{H}_{\text{eq}}(\mathbf{P}, \mathbf{Q}) - \mathcal{F}(t)\mathcal{A}(\mathbf{P}, \mathbf{Q})$ . Here,  $\mathbf{Q} \equiv (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N)$  and  $\mathbf{P} \equiv (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$  correspond to the positions and momenta of the constituent particles, respectively. The temporal evolution of the expected value of another mechanical observable,  $\mathcal{B}(\mathbf{P}, \mathbf{Q})$ , will be given by

$$\langle \Delta \mathcal{B}(t) \rangle = \int_{-\infty}^t \mathcal{F}(t') \phi_{\mathcal{AB}}(t - t') dt' \quad (2)$$

Here,  $\Delta \mathcal{B}(t) = \mathcal{B}(t) - \langle \mathcal{B} \rangle_{\mathcal{H}_{\text{eq}}}$ , and  $\phi_{\mathcal{AB}}(\tau)$  can be interpreted as a response function and is given by

$$\phi_{\mathcal{AB}}(\tau) = \beta \langle \mathcal{B}(\tau) \dot{\mathcal{A}}(0) \rangle_{\mathcal{H}_{\text{eq}}} \quad (3)$$

For any specific transport coefficient, an external perturbation can be applied to establish a suitably small macroscopic gradient within the system. Equations 2 and 3 can then be employed to relate the corresponding flux of interest to the gradient. This approach leads to a collection of equations known as *Green–Kubo* relationships,<sup>51,52</sup> characterizing transport coefficients in terms of autocorrelations of mechanical observables.

As an illustration, consider self-diffusivity, where a plausible perturbation to the Hamiltonian can be formulated as

$$\mathcal{A}(\mathbf{P}, \mathbf{Q}) = -\alpha \mathbf{w} \cdot \sum_{i=1}^N \mathbf{q}_i \quad (4)$$

Here,  $\alpha$  represents a fixed force pulling the particles along a unit vector  $\mathbf{w}$ , resulting in a net flux of particles along  $\mathbf{w}$ . The response function for momentum flux along another unit vector  $\mathbf{u}$ ,  $\mathcal{B}(\mathbf{P}, \mathbf{Q}) = \mathbf{u} \cdot \sum_{i=1}^N \mathbf{p}_i$ , can be enumerated using eq 3:

$$\begin{aligned} \phi_{\mathcal{AB}}(t) &= -\frac{\alpha\beta}{m} \left\langle \sum_{i,j=1}^N \mathbf{u}^T \mathbf{p}_i(t) \mathbf{p}_j^T(t) \mathbf{w} \right\rangle \\ &\stackrel{(a)}{=} -m\alpha\beta \mathbf{u}^T \left\langle \sum_{i=1}^N \mathbf{v}_i(t) \mathbf{v}_i^T(0) \right\rangle \mathbf{w} \\ &\stackrel{(b)}{=} -mN\alpha\beta \mathbf{u}^T \langle \mathbf{v}(t) \mathbf{v}^T(0) \rangle \mathbf{w} \end{aligned}$$

Here, (a) follows from the fact that momentum degrees of freedom are uncorrelated, while (b) results from the indistinguishability of particles. This expression can be utilized to evaluate the mean velocity along the unit vector  $\mathbf{u}$ :

$$\lim_{t \rightarrow \infty} \mathbf{u}^T \langle \mathbf{v}(t) \rangle = -\alpha\beta \int_0^\infty \mathbf{u}^T \langle \mathbf{v}(t) \mathbf{v}^T(0) \rangle \mathbf{w} dt$$

This observation allows us to compute  $\mathbf{u}^T \mathbf{D} \mathbf{w}$ , given by<sup>50</sup>

$$\mathbf{u}^T \mathbf{D} \mathbf{w} = -\frac{\mathbf{u}^T \langle \mathbf{v}(\infty) \rangle}{\alpha\beta} = \int_0^\infty \mathbf{u}^T \langle \mathbf{v}(t) \mathbf{v}^T(0) \rangle \mathbf{w} dt$$

Choosing  $\mathbf{u}$  and  $\mathbf{w}$  from among the basis vectors in Cartesian coordinates yields the well-known relationship:

$$\mathbf{D} = \int_0^\infty \langle \mathbf{v}(t) \mathbf{v}^T(0) \rangle dt \quad (5)$$

where the integrand is typically referred to as the *velocity autocorrelation function* (VACF). Similar expressions can be

derived for other transport coefficients. For instance, the shear viscosity tensor can be estimated from<sup>53</sup>

$$\mu_{\alpha\beta\gamma\delta} = \frac{V}{k_B T} \int_0^\infty \langle \delta \tau_{\alpha\beta}(t) \delta \tau_{\gamma\delta}(0) \rangle dt \quad (6)$$

where  $\boldsymbol{\tau}$  is the second-rank stress tensor computed from the virial relationship,<sup>54</sup> and  $\delta \boldsymbol{\tau} = \boldsymbol{\tau} - \langle \boldsymbol{\tau} \rangle$ . Likewise, thermal conductivity can be related to autocorrelations of heat flux:<sup>55</sup>

$$\mathbf{k} = \frac{1}{k_B T^2 V} \int_0^\infty \langle \mathbf{J}_h(t) \mathbf{J}_h^T(0) \rangle dt \quad (7)$$

with the instantaneous heat flux,  $\mathbf{J}_h$ , defined as,

$$\mathbf{J}_h(t) = \frac{d}{dt} \sum_{i=1}^N \mathbf{r}_i(t) \epsilon_i(t) \quad (8)$$

Here,  $\epsilon_i(t)$  is the sum of the kinetic and potential energy of particle  $i$ . Finally, electric conductivity can similarly be obtained from autocorrelations of the electric current:<sup>56</sup>

$$\boldsymbol{\sigma} = \frac{1}{k_B T V} \int_0^\infty \langle \mathbf{J}_e(t) \mathbf{J}_e^T(0) \rangle dt \quad (9)$$

with electric current,  $\mathbf{J}_e$ , given by  $\mathbf{J}_e = \sum_{i=1}^N q_i \mathbf{v}_i$ .

A conceptually related class of relationships, developed by Helfand,<sup>57</sup> estimates transport coefficients through the asymptotic slopes of time- and ensemble-averaged generalized displacements. These displacements are time integrals of microscopic fluxes. The most well-known example is the *Einstein relationship*,<sup>58</sup> which links diffusivity to the asymptotic slope of *mean-squared displacement* (MSD):

$$\mathbf{D}_i = \lim_{t \rightarrow \infty} \frac{\langle [\mathbf{r}_i(t) - \mathbf{r}_i(0)][\mathbf{r}_i(t) - \mathbf{r}_i(0)]^T \rangle}{2t} \quad (10)$$

Helfand<sup>57</sup> expanded upon this approach by linearizing the corresponding conservation laws and solving them over an infinite domain. When it comes to the transport of linear momentum, viscosity can be expressed as

$$\begin{aligned} \mu_{\alpha\beta\gamma\delta} &= \lim_{t \rightarrow \infty} \frac{1}{2tk_B T V} \left\langle \left[ \sum_{i=1}^N [p_{i,\alpha}(t) r_{i,\beta}(t) - p_{i,\alpha}(0) r_{i,\beta}(0)] \right. \right. \\ &\quad \left. \left. [p_{i,\gamma}(t) r_{i,\delta}(t) - p_{i,\gamma}(0) r_{i,\delta}(0)] \right]^2 \right\rangle \end{aligned} \quad (11)$$

Similar expressions can be obtained for other transport coefficients. For instance, heat conductivity is given by<sup>59</sup>

$$\mathbf{k} = \lim_{t \rightarrow \infty} \frac{\langle [\mathbf{h}(t) - \mathbf{h}(0)][\mathbf{h}(t) - \mathbf{h}(0)]^T \rangle}{2tV k_B T^2} \quad (12)$$

Here,  $\mathbf{h}(t) = \int_0^t \mathbf{J}_h(\tau) d\tau$  is the integrated heat flux where  $\mathbf{J}_h(\tau)$  is defined by (8).

The Green–Kubo formalism is commonly regarded as the primary method for estimating transport coefficients from MD trajectories. In contrast, the Helfand approach is more frequently applied in the estimation of diffusivity only and is less commonly employed for other transport coefficients. Its limited usage can be partly attributed to the challenges associated with its proper implementation, particularly when dealing with periodic boundary conditions, as highlighted by Viscardy and Gaspard.<sup>60</sup>

**2.2. Challenges in Confined Geometries.** Under confinement, all physical properties become functions of

position. Moreover, transport coefficients also become anisotropic, making it necessary to account for their tensorial nature. (A conceptually related dependence on pair separation can be realized even in translationally isotropic systems, e.g., in the case of the pair distance diffusivity tensor,<sup>61</sup> although such quantities do not strictly depend on the position of the pair in the Cartesian space). The main challenge in computing them, however, lies in the absence of well-defined autocorrelation-based estimators for open systems. Specifically, spatial profiles of mechanical observables that are unambiguously defined for arbitrary regions within the simulation box can be accurately estimated using spatial binning. For instance, the spatial profile of a mechanical observable  $\mathcal{R}$  that is well-defined for every particle can be expressed as

$$\mathcal{R}(\mathbf{r}) = \frac{\langle \sum_{i=1}^N \mathcal{R}_i \delta(\mathbf{r}_i - \mathbf{r}) \rangle}{\langle \sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}) \rangle} \quad (13)$$

In practice,  $\mathcal{R}(\mathbf{r})$  is estimated by partitioning the simulation box into nonoverlapping bins and determining the mean of  $\mathcal{R}(\mathbf{r})$  within each bin. For ergodic systems,  $\mathcal{R}_k$ , the average of  $\mathcal{R}$  over the  $k$ -th bin, can be estimated as

$$\mathcal{R}_k = \frac{\int_0^{t_{\text{sim}}} \mathcal{R}_i(t) \chi_k[\mathbf{r}_i(t)] dt}{\int_0^{t_{\text{sim}}} \chi_k[\mathbf{r}_i(t)] dt} \quad (14)$$

where  $\chi_k(\cdot)$  is the characteristic function of the  $k$ -th bin. However, there exists no natural extension of (14) for quantities such as MSD or VACF, which are autocorrelations of mechanical observables. This limitation arises due to particle exchange among bins and the ambiguity in quantifying the contribution of exchanged particles to the autocorrelation associated with a specific bin. As a consequence, methodologies discussed in Section 2.1 are only robustly applicable in the bulk. Moreover, the statistical behavior of microscopic fluxes in confined geometries, such as particle mobilities, diverges significantly from established conservation laws valid in bulk systems. In the case of self-diffusivity, for instance, the self-part of the van Hove correlation function,<sup>62</sup>  $G_s(\mathbf{r}, t | \mathbf{r}_0, 0)$ , satisfies the Fokker–Planck equation<sup>63,64</sup> in the bulk:

$$\frac{\partial G_s}{\partial t} = \nabla \cdot [\mathbf{D} \cdot \nabla G_s] \quad (15)$$

However, under confinement, both diffusivity and equilibrium density become position-dependent, and  $G_s(\mathbf{r}, t | \mathbf{r}_0, 0)$  will satisfy the Smoluchowski equation:<sup>65</sup>

$$\frac{\partial G_s}{\partial t} = \nabla \cdot [\mathbf{D}(\mathbf{r}) \cdot (\nabla G_s + \beta \nabla G_s \mathcal{F})] = \mathcal{L}_r^\dagger G_s \quad (16)$$

Here,  $\mathcal{F}(\mathbf{r}) = -\beta^{-1} \ln \rho_0(\mathbf{r})$  represents the free energy profile wherein  $\rho_0(\mathbf{r})$  denotes the number density profile. Unlike eq 15, which possesses straightforward analytical solutions in simple geometries, eq 16 lacks a closed-form solution. Consequently, it is not trivial to extract position-dependent diffusivity tensors from MD trajectories. The remainder of this review is dedicated to a comprehensive discussion of various numerical approaches developed for tackling this nontrivial task.

### 3. AD HOC EXTENSIONS OF CLASSICAL METHODS

As discussed above, well-established methodologies based on MSD or VACF cannot be readily applied to confined geometries since the Smoluchowski equation lacks a simple analytical solution. Nevertheless, many researchers have still presumed the local validity of such formalisms, and have accordingly devised *ad hoc* extensions of MSD and VACF for the purpose of estimating position-dependent diffusivities. This section is dedicated to a comprehensive discussion of such efforts, including the common practices and conventions underpinning such *ad hoc* extensions. Moreover, we discuss the merits and limitations of the theoretical arguments that could be made for making such *ad hoc* frameworks more rigorous.

**3.1. Ad Hoc Mean Squared Displacements.** As discussed in Section 2.2, it is not feasible to analytically establish a linear relationship between the asymptotic slope of a localized notion of MSD and local diffusivity. Nonetheless, this has been the most widely adopted approach in the molecular simulations community for estimating position-dependent diffusivity, wherein *ad hoc* localized notions of MSD are constructed, and local diffusivity is extracted through a linear regression between local MSDs and the observation window (i.e., the time lag). As an illustration, in the case of one-dimensional confinement along the  $z$  axis, a lateral MSD for bin  $i$  can be formulated as

$$\mathcal{M}(z_i, t) = \langle [(x_{t+\tau} - x_\tau)^2 + (y_{t+\tau} - y_\tau)^2] \xi_i[z(t')_{\tau \leq t' \leq t+\tau}] \rangle_t \quad (17)$$

Here,  $\xi_i[z(t')]$  specifies the weight assigned to a particle's contribution to the *ad hoc* MSD of bin  $i$ , based on its trajectory  $z(t')$  over the time interval  $\tau \leq t' \leq t + \tau$ . The simplest choice of  $\xi_i[\cdot]$  is given by

$$\xi_i[z(t')] = \chi_i(z_\tau) \quad (18)$$

wherein  $\chi_i(\cdot)$  is the characteristic function of the  $i$ -th bin. In other words, eq 18 only allows particles that are within a particular bin at the beginning of an observation window to contribute to the local MSD of that bin.<sup>66,67</sup> However, this straightforward approach can lead to significant errors over extended timeframes, as particles originating from bin  $i$  may travel to distant bins. To address this issue, alternative definitions have been proposed. For instance, some authors use the average  $z$  along  $z(t')$  to allocate the particle to a specific bin.<sup>68</sup> More restrictive definitions, such as only including particles present within the bin at both the beginning and the end of the observation window,<sup>69–71</sup> or requiring the trajectory to remain within the bin throughout the observation window,<sup>72,73</sup> have also been employed.

Despite inherent limitations of such *ad hoc* approaches, such localized notions of MSD can be modified in creative ways to yield more realistic proxies for position-dependent dynamics. A notable example is the approach proposed by Liu and Berne,<sup>72</sup> who approximate lateral diffusivity as

$$D_{xx} = D_{yy} \approx \frac{\mathcal{M}(z_i, t)}{4tP_i(t)} \quad (19)$$

Here,  $\mathcal{M}(z_i, t)$  is a localized MSD as in eq 18, with the convention that the trajectory should remain within the same bin throughout the entire time interval. Additionally,  $P_i(t)$  denotes the survival probability, i.e., the probability that a

particle starting within the  $i$ -th bin will still remain in that bin after time  $t$ .

Another approach, akin in spirit, is proposed by Nagai and Okazaki,<sup>74</sup> wherein a biasing potential is introduced, which is flat within a designated spatial bin but becomes strongly repulsive outside the bin. The biased Hamiltonian is then employed to launch MD trajectories from which the localized MSD,  $M_{\text{FB}}(t)$ , is computed, with "FB" denoting the "force-biased" nature of these simulations. The authors argue that the distortion introduced within the MSD due to force biasing is independent of the magnitude of the local diffusivity, and is instead determined by a geometric factor:

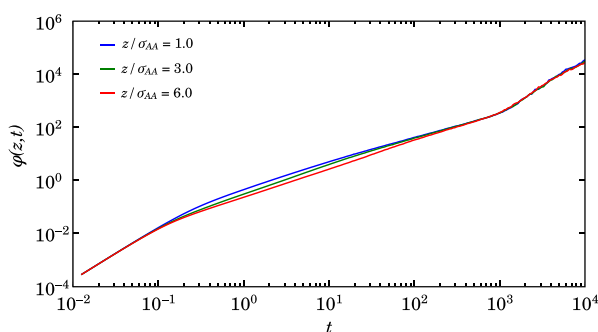
$$\mathcal{G}(\tau) = \frac{\bar{M}_{\text{FB}}(\tau)}{\bar{M}(\tau)}$$

Here,  $M(\cdot)$  is the unbiased MSD in the bulk,  $\tau = tD_0/l_0^2$  represents dimensionless time, and  $\bar{M} = M/l_0^2$  corresponds to dimensionless MSD. After determining the geometric factor, the diffusivity within each bin can be computed by iteratively solving the following equation:

$$\frac{M_{\text{FB},\mathbf{x}_0}(t)}{\mathcal{G}[tD(\mathbf{r}_0)l_0^{-2}]} = 2dD(\mathbf{r}_0)t + C \quad (20)$$

wherein  $C$  is a constant corresponding to the nonzero intercept of MSD.

While diffusivity profiles obtained from these *ad hoc* approaches serve as semiquantitative proxies for spatial dynamic heterogeneities in confined materials, they do not strictly align with the Smoluchowski equation in the sense that the mobility statistics implied by these approaches, in accordance with (16), may not necessarily match actual  $G_s(\mathbf{r}, t)$  obtained from MD simulations. Additionally, these methods provide only lateral components of the diffusivity tensor. There have been attempts to overcome this latter limitation, e.g., by employing dual simulation approaches for estimating the normal component of the diffusivity tensor, as demonstrated by Piu et al.<sup>72</sup> Finally, the *ad hoc* MSDs do not capture diffusive behavior over extremely long time scales due to interbin mixing<sup>70</sup> (Figure 1). Consequently, fitting procedures must be constrained to intermediate time scales to yield reliable estimates.



**Figure 1.** Breakdown of linear scaling of the *ad hoc* MSD with time for MSDs computed at different distances from the substrate within a Kob–Andersen<sup>77</sup> liquid. (Reprinted from ref 70 with the permission of AIP Publishing.)

While *ad hoc* definitions of MSD may be heuristic in nature, they possess inherent fundamental merit. Specifically, the Smoluchowski equation can be readily reformulated as

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\boldsymbol{\mu} \rho) + \nabla \cdot \nabla \cdot [\mathbf{D} \rho], \quad \rho(\mathbf{r}, 0) = \delta(\mathbf{r} - \mathbf{s}) \quad (21)$$

Here,  $\boldsymbol{\mu}$  and  $\mathbf{D}$  represent the drift and diffusivity profiles, respectively, and are related to the formal solution of (21) via Kramers–Moyal<sup>75,76</sup> relations:

$$\boldsymbol{\mu}(\mathbf{s}) = \lim_{\tau \rightarrow 0^+} \frac{1}{\tau} \int (\mathbf{r} - \mathbf{s}) \rho_\tau(\mathbf{r}|\mathbf{s}) d\mathbf{r} \quad (22a)$$

$$\mathbf{D}(\mathbf{s}) = \lim_{\tau \rightarrow 0^+} \frac{1}{2\tau} \int (\mathbf{r} - \mathbf{s})(\mathbf{r} - \mathbf{s})^T \rho_\tau(\mathbf{r}|\mathbf{s}) d\mathbf{r} \quad (22b)$$

Given the interpretation of  $\rho_\tau(\mathbf{r}|\mathbf{s})$  as a probability density, (22a) can be readily recast as

$$\boldsymbol{\mu}(\mathbf{s}) = \lim_{\tau \rightarrow 0^+} \frac{\langle \mathbf{X}_{t+\tau} - \mathbf{X}_t \rangle}{\tau} \Big|_{\mathbf{X}_t=\mathbf{s}} \quad (23a)$$

$$\mathbf{D}(\mathbf{s}) = \lim_{\tau \rightarrow 0^+} \frac{\langle (\mathbf{X}_{t+\tau} - \mathbf{X}_t)(\mathbf{X}_{t+\tau} - \mathbf{X}_t)^T \rangle}{2\tau} \Big|_{\mathbf{X}_t=\mathbf{s}} \quad (23b)$$

Therefore, local diffusivity can be approximated via the following finite difference expression:

$$\mathbf{D}(\mathbf{s}) \approx \hat{\mathbf{D}}^\tau(\mathbf{s}) = \frac{\langle (\mathbf{X}_{t+\tau} - \mathbf{X}_t)(\mathbf{X}_{t+\tau} - \mathbf{X}_t)^T \rangle}{2\tau} \Big|_{\mathbf{X}_t=\mathbf{s}} \quad (24)$$

Likewise, it is possible to propose a finite- $\tau$  estimator for the drift  $\boldsymbol{\mu}$ :

$$\boldsymbol{\mu}(\mathbf{s}) \approx \hat{\boldsymbol{\mu}}^\tau(\mathbf{s}) = \frac{\langle \mathbf{X}_{t+\tau} - \mathbf{X}_t \rangle}{\tau} \Big|_{\mathbf{X}_t=\mathbf{s}} \quad (25)$$

Here, the time scale  $\tau$  is system dependent, and should be chosen in such a manner that single-particle trajectories behave diffusively at and beyond  $\tau$ . Notably, eq 24 bears resemblance to the earlier *ad hoc* MSD definitions, therefore belonging to the broad category of Helfand approaches. A crucial distinction lies in the Helfand approach's consideration of the limiting slope of MSD at  $\tau \rightarrow \infty$  to ensure capturing long-term diffusive behavior. In confined systems, it is instead imperative to select a  $\tau$  that is as small as possible in order to avoid interbin mixing. A large  $\tau$  will introduce considerable discretization errors, compromising the spatial resolution of diffusivity profiles. Thus, selecting an optimal  $\tau$  is pivotal to preserve accurate data without loss due to discretization errors. For systems that behave diffusively across all time scales,  $\tau$  can be made as small as a single time step. MD trajectories, however, are only diffusive beyond the caging regime, and as such there is a strict lower bound on the  $\tau$  that can be used in (24). (An operational procedure for the selection of  $\tau$  is provided in Section 4.)

To comprehend the impact of time discretization introduced by a nonzero  $\tau$ , consider the formal solution of eq 16, expressed as  $\rho_t(\mathbf{y}|\mathbf{x}) = e^{t\hat{\mathcal{L}}^\dagger} \delta(\mathbf{y} - \mathbf{x})$ . This solution allows the reformulation of  $\hat{\mathbf{D}}_\tau(\mathbf{x})$  as follows:

$$\begin{aligned}\hat{\mathbf{D}}^\tau(\mathbf{x}) &= \frac{1}{\tau} \int (\mathbf{y} - \mathbf{x})(\mathbf{y} - \mathbf{x})^T e^{t\mathcal{L}_y^\dagger} \delta(\mathbf{y} - \mathbf{x}) d\mathbf{y} \\ &= \frac{1}{\tau} \int e^{t\mathcal{L}_y} [(\mathbf{y} - \mathbf{x})(\mathbf{y} - \mathbf{x})^T] \delta(\mathbf{y} - \mathbf{x}) d\mathbf{y} \\ &= \int \left[ \sum_{l=1}^{+\infty} \frac{\tau^{l-1}}{l!} \mathcal{L}_y^\dagger [(\mathbf{y} - \mathbf{x})(\mathbf{y} - \mathbf{x})^T] \delta(\mathbf{y} - \mathbf{x}) \right] d\mathbf{y}\end{aligned}$$

where  $\mathcal{L}_y$  is the operator adjoint to the  $\mathcal{L}_y^\dagger$  of eq 16 under the standard  $L^2$  inner product. Notably, the  $l = 1$  term upon integration precisely yields  $\mathbf{D}(\mathbf{x})$ , resulting in

$$\hat{\mathbf{D}}^\tau(\mathbf{x}) = \mathbf{D}(\mathbf{x}) + \int \left[ \sum_{l=2}^{+\infty} \frac{\tau^{l-1}}{l!} \mathcal{L}_y^\dagger [(\mathbf{y} - \mathbf{x})(\mathbf{y} - \mathbf{x})^T] \delta(\mathbf{y} - \mathbf{x}) \right] d\mathbf{y}$$

The second term on the right-hand side represents an error term, which vanishes as  $\tau \rightarrow 0^+$ . For nonzero  $\tau$ , this expression serves as a foundational reference to devise correction schemes aimed at mitigating implicit mixing effects inherent in the error term above. For instance, it has been demonstrated that in one dimension, this series can be approximated to quadratic order as<sup>78,79</sup>

$$\begin{aligned}\langle X_{t+\tau} - X_t \rangle|_{X_t=x} &= \tau\mu(x) + \\ \frac{\tau^2}{2} [\mu(x)\mu'(x) + D(x)\mu''(x)] &+ O(\tau^3)\end{aligned}\quad (26a)$$

$$\begin{aligned}\langle (X_{t+\tau} - X_t)(X_{t+\tau} - X_t)^T \rangle|_{X_t=x} &= 2\tau D(x) + \\ \tau^2 \{ \mu^2(x) + [\mu(x)D(x)]' + D(x)[D'(x) + \mu(x)]' \} &+ O(\tau^3)\end{aligned}\quad (26b)$$

Assuming that all derivatives of  $\mu(x)$  and  $D(x)$  are negligible leads to the following simplified expressions:<sup>79</sup>

$$\langle X_{t+\tau} - X_t \rangle|_{X_t=x} = \tau\mu(x) + O(\tau^3) \quad (27a)$$

$$\begin{aligned}\langle (X_{t+\tau} - X_t)(X_{t+\tau} - X_t)^T \rangle|_{X_t=x} &= 2\tau D(x) + \tau^2 \mu^2(x) + O(\tau^3)\end{aligned}\quad (27b)$$

Note that (27b) can be readily rearranged as

$$\begin{aligned}\langle (\Delta X_{t,\tau} - \tau\mu(x))(\Delta X_{t,\tau} - \tau\mu(x))^T \rangle|_{X_t=x} &= 2\tau D(x) + O(\tau^3)\end{aligned}\quad (28)$$

wherein  $\Delta X_{t,\tau} = X_{t+\tau} - X_t$ . It must be noted that the left-hand side of (28) constitutes a covariance of displacements. The ability to estimate diffusivity from computing a local covariance matrix suggests that the probability density function of  $\mathbf{X}_\tau$ , a particle's position at time  $\tau$ , can be approximated as

$$\mathbf{X}_\tau \sim \mathcal{N}(\mathbf{r}_0 + \tau\boldsymbol{\mu}(\mathbf{r}_0), 2\tau\mathbf{D}(\mathbf{r}_0)) \quad (29)$$

wherein  $\mathbf{r}_0$  is the particle's position at  $t = 0$ , and  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  is a multivariate Gaussian distribution with mean  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}$ . Indeed, eq 24 has been employed for estimating lateral diffusivity of water near interfaces,<sup>80</sup> as well as diffusivity along collective variables (CVs) employed in protein folding

simulations.<sup>81,82</sup> Interestingly, it was shown by Hinczewski et al.<sup>82</sup> that even for collective variable spaces, the estimated diffusivity is acutely sensitive to  $\tau$ . A common strategy<sup>83–85</sup> in protein folding simulations involves using eq 29 to fit Gaussians into empirical histograms obtained around a certain point but at different times, and use the following expression to estimate diffusivity,

$$D(\lambda_0) \approx \frac{\sigma^2(\Lambda_{\tau_2}) - \sigma^2(\Lambda_{\tau_1})}{2(\tau_2 - \tau_1)} \equiv \hat{D}_{\tau_1, \tau_2}(\lambda_0) \quad (30)$$

where  $\Lambda_t$  is a random variable describing the state of the system (within the CV space) at time  $t$ . In the limit of  $\Delta\tau = \tau_2 - \tau_1 \rightarrow 0^+$ , it can be demonstrated that

$$\hat{D}_{\tau_1, \tau_2}(\lambda_0) \xrightarrow{\Delta\tau \rightarrow 0} \int D(\lambda) \rho_{\tau_1}(\lambda|\lambda_0) d\lambda \quad (31)$$

where  $\rho_t$  denotes the solution of the Smoluchowski equation in the collective variable space. Methods designed with this specific application in mind will be detailed in Section 8.

**3.2. Ad Hoc Velocity Autocorrelation Functions.** In confined geometries, suitable *ad hoc* definitions of VACF can be devised in a fashion similar to MSD.<sup>73,86</sup> The local diffusivity within bin  $i$  can then be evaluated as

$$\mathbf{D}_i = \int_0^{+\infty} \mathbf{C}_{v,i}(t) dt$$

with  $\mathbf{C}_{v,i}(t)$ , the localized VACF defined as

$$\mathbf{C}_{v,i}(t) = \langle \mathbf{v}(t + \tau) \mathbf{v}^T(\tau) \xi_i[\mathbf{r}(t')]_{\tau \leq t' \leq \tau+t} \rangle_\tau$$

Here,  $\xi_i[\mathbf{r}(t')]$  serves a similar role as in eq 17.

It is crucial to highlight two notable differences between the *ad hoc* extensions of MSD and VACF. First, we expect *ad hoc* estimates of local diffusivity through VACF to exhibit reduced susceptibility to interbin mixing. This stems from VACF's inherent decay to zero within time scales relevant for such mixing, thereby enhancing their practical utility. Second, as expounded upon later in this section, specific *ad hoc* extensions of VACF can be derived utilizing linear response theory. This entails applying a suitable perturbation term to the entire system while monitoring the response of a locally defined mechanical observable to such a perturbation. The arising mobility profiles can then be linked to local diffusivity in a manner similar to the bulk. As will be discussed later, these approaches are still inherently *ad hoc* in the sense that observed mobility statistics are not guaranteed to conform to predictions based on the Smoluchowski equation<sup>65</sup> or the Kramers-Klein equation.<sup>75</sup>

An illustration, consider Hunter et al.<sup>87</sup> who introduce a perturbation to the Hamiltonian given by

$$\mathcal{A}(\mathbf{P}, \mathbf{Q}) = -F_c \mathbf{w} \cdot \sum_{i=1}^N c_i \mathbf{q}_i$$

Here,  $F_c$  represents the magnitude of the force,  $\mathbf{w}$  is a unit vector, and  $c_i$  signifies the 'color' associated with particle  $i$ , allowing for the adjustment of both the direction and the strength of the biasing force applied to different particles. The authors adopt color currents as an elegant means of handling correlations between momentum degrees of freedom. Such correlations, while typically absent in a strict statistical mechanical sense, are frequently present in MD trajectories where the system's net linear momentum is set to zero. By

using  $c_i = (-1)^i$ , Hunter et al. demonstrate that such correlations readily decay in the thermodynamic limit.

To establish a localized notion of diffusivity in confined geometries, they examine the response of a spatially localized observable given by

$$\mathcal{B}(\mathbf{P}, \mathbf{Q}) = \frac{\mathbf{u} \cdot \sum_{i=1}^N c_i \mathbf{p}_i \chi_A(\mathbf{q}_i)}{m \sum_{i=1}^N \chi_A(\mathbf{q}_i)}$$

Here,  $\chi_A(\cdot)$  represents the characteristic function of set  $A$ . Consequently,  $\mathcal{B}$  signifies a colored momentum flux directed along unit vector  $\mathbf{u}$  and confined to  $A$ . By utilizing eq 3, one can determine the response of  $\mathcal{B}$  to the perturbation given by  $\mathcal{A}$ :

$$\begin{aligned} -\frac{\langle \Delta \mathbf{B}(t) \rangle}{\beta F_c} &= \int_0^{+\infty} \mathbf{u}^T \left\langle \frac{\sum_{i=1}^N \mathbf{v}_i(t) \mathbf{v}_i^T(0) \chi_A(\mathbf{q}_i(t))}{\sum_{i=1}^N \chi_A(\mathbf{q}_i(t))} \right\rangle \mathbf{w} dt \\ &= N \int_0^{+\infty} \mathbf{u}^T \mathbf{C}_{v,A}(t) \mathbf{w} dt \end{aligned}$$

Here,  $\xi_A[\mathbf{r}(t')] = \chi_A[\mathbf{r}(t + \tau)]$ . Subsequently, Hunter et al.<sup>87</sup> propose the mean diffusivity within  $A$  to be associated with  $\mathbf{C}_{v,A}(t)$  as

$$\frac{1}{\text{Vol}(A)} \int_A \mathbf{D}(\mathbf{r}) d\mathbf{r} \stackrel{?}{=} \int_0^{+\infty} \mathbf{C}_{v,A}(t) dt \quad (32)$$

While this quantity is a proxy for local diffusivity, it cannot be directly mapped onto the Smoluchowski formalism. A suitable theoretical framework to assert this proposition is the Kramers-Klein equation,<sup>75</sup> describing Langevin dynamics in the underdamped regime:

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\mathbf{v} \cdot \nabla_{\mathbf{r}} p + \frac{1}{m} \nabla_{\mathbf{v}} [p \nabla_{\mathbf{r}} \mathcal{F}] + \frac{1}{m\beta} \nabla_{\mathbf{v}} \cdot [\mathbf{D}^{-1}(\mathbf{r}) \cdot \mathbf{v} p] \\ &+ \frac{1}{(m\beta)^2} \mathbf{D}^{-1}(\mathbf{r}) : \mathbf{H}_{\mathbf{v}} p = \mathcal{L}_{\mathbf{r},\mathbf{v}}^{\dagger} p \end{aligned} \quad (33)$$

Here,  $m$  represents particle mass,  $\mathcal{F}(\mathbf{r})$  denotes a conservative potential of mean force,  $[\mathbf{H}_{\mathbf{v}} p]_{ij} = \partial^2 p / \partial v_i \partial v_j$  is the Hessian tensor with respect to velocity degrees of freedom, and  $:$  signifies full tensorial contraction. The adjoint operator of  $\mathcal{L}_{\mathbf{r},\mathbf{v}}$  with respect to the standard inner product is given by

$$\begin{aligned} \mathcal{L}_{\mathbf{r},\mathbf{v}} \equiv & \mathbf{v} \cdot \nabla_{\mathbf{r}} - \frac{\nabla_{\mathbf{r}} \mathcal{F}}{m} \cdot \nabla_{\mathbf{v}} - \frac{\mathbf{v}}{m\beta} \cdot \mathbf{D}^{-1}(\mathbf{r}) \cdot \nabla_{\mathbf{v}} + \frac{1}{(m\beta)^2} \mathbf{D}^{-1}(\mathbf{r}) : \mathbf{H}_{\mathbf{v}} \end{aligned} \quad (34)$$

The *ad hoc* VACF of (32) can be formulated as

$$\begin{aligned} \mathbf{C}_{v,A}(t) &= \langle \mathbf{v}(t) \mathbf{v}^T(0) K[\mathbf{r}(t) - \mathbf{r}_0] \rangle \\ &\stackrel{(a)}{=} \langle \mathbf{v}(t) \mathbf{v}^T(0) K[\mathbf{r}(0) - \mathbf{r}_0] \rangle \\ &= \int \mathbf{v} \mathbf{w}^T K(\mathbf{s} - \mathbf{r}_0) p_t(\mathbf{r}, \mathbf{v} | \mathbf{s}, \mathbf{w}) \Theta(\mathbf{s}, \mathbf{w}) d\mathbf{r} d\mathbf{s} d\mathbf{v} d\mathbf{w} \end{aligned} \quad (35)$$

where  $K[\mathbf{r}]$  is a properly normalized kernel function that, in the case of Hunter et al.'s work, is taken as the indicator of set  $A$  (with  $\mathbf{r}_0 \in A$ ). Note that (a) follows from the linear response theory.  $p_t(\cdot)$  is the solution of the Kramers-Klein equation and can be formally expressed as

$$p_t(\mathbf{r}, \mathbf{v} | \mathbf{s}, \mathbf{w}) = e^{t\mathcal{L}_{\mathbf{r},\mathbf{v}}} [\delta(\mathbf{r} - \mathbf{s}) \delta(\mathbf{v} - \mathbf{w})]$$

By using the adjoint operator, the *ad hoc* VACF can be expressed as

$$\mathbf{C}_{v,A}(t) = \int [e^{t\mathcal{L}_{\mathbf{r},\mathbf{v}}} \mathbf{v} \mathbf{v}^T K(\mathbf{r} - \mathbf{r}_0) \Theta(\mathbf{r}, \mathbf{v})] d\mathbf{r} d\mathbf{v}$$

where  $\Theta(\mathbf{r}, \mathbf{v})$  is given by

$$\Theta(\mathbf{r}, \mathbf{v}) = \left( \frac{m\beta}{2\pi} \right)^{d/2} \exp \left[ -\frac{m\beta |\mathbf{v}|^2}{2} \right] \rho_0(\mathbf{r})$$

In the case of a trivial potential of mean force and fixed diffusivity, it is easy to show that

$$\mathcal{L}_{\mathbf{r},\mathbf{v}} \mathbf{v} = -\frac{\mathbf{D}^{-1} \cdot \mathbf{v}}{m\beta}, \quad \Rightarrow [e^{t\mathcal{L}_{\mathbf{r},\mathbf{v}}} \mathbf{v}] = \exp \left[ -\frac{t\mathbf{D}^{-1}}{m\beta} \right] \cdot \mathbf{v}$$

which, upon integrating momenta degrees of freedom, yields

$$\mathbf{C}_{v,A}(t) = \frac{1}{m\beta} \exp \left[ -\frac{t\mathbf{D}^{-1}}{m\beta} \right] \int K(\mathbf{r} - \mathbf{r}_0) \rho_0(\mathbf{r}) d\mathbf{r} \quad (36)$$

By time integrating (36), one can demonstrate the canonical relationship between VACF and diffusivity, namely:

$$\int_0^{+\infty} \langle \mathbf{v}(t) \mathbf{v}^T(0) K(\mathbf{r} - \mathbf{r}_0) \rangle dt = \int \mathbf{D} K(\mathbf{r} - \mathbf{r}_0) \rho_0(\mathbf{r}) d\mathbf{r}$$

for a properly normalized kernel. However, for nontrivial potential of mean force and position-dependent diffusivity,  $\mathcal{L}_{\mathbf{r},\mathbf{v}}$  will possess the following mathematical form:

$$\mathcal{L}_{\mathbf{r},\mathbf{v}} \mathbf{v} = -\frac{\nabla_{\mathbf{r}} \mathcal{F}}{m} - \frac{\mathbf{D}^{-1}(\mathbf{r}) \cdot \mathbf{v}}{m\beta}$$

This makes constructing the  $e^{t\mathcal{L}}$  operator extremely complicated since each successive application of  $\mathcal{L}$  will require computing spatial derivatives of the unknown diffusivity profile as well as the potential of mean force. More precisely, if one denotes  $\mathbf{f}_{\infty}(\mathbf{r}, \mathbf{v}) := \lim_{t \rightarrow \infty} e^{t\mathcal{L}_{\mathbf{r},\mathbf{v}}} \mathbf{v}$ , then one can demonstrate that

$$\begin{aligned} & \int_0^{+\infty} \langle \mathbf{v}(t) \mathbf{v}^T(0) K(\mathbf{r} - \mathbf{r}_0) \rangle dt \\ &= \int \mathbf{g}_{\mathcal{L}}(\mathbf{r}, \mathbf{v}) \mathbf{v}^T K(\mathbf{r} - \mathbf{r}_0) \Theta(\mathbf{r}, \mathbf{v}) d\mathbf{r} d\mathbf{v} \end{aligned}$$

where  $\mathbf{g}_{\mathcal{L}}(\mathbf{r}, \mathbf{v})$  is the solution of the partial differential equation (PDE),  $\mathcal{L}_{\mathbf{r},\mathbf{v}} \mathbf{g}_{\mathcal{L}}(\mathbf{r}, \mathbf{v}) = \mathbf{f}_{\infty} - \mathbf{v}$ , which does not lend itself easily to a solution, and is not definitely consistent with the simplified postulation of (32). Therefore, even *ad hoc* representations constructed using linear response theory fail to yield diffusivity profiles consistent with the Smoluchowski or the Kramers-Klein picture.

Despite this fundamental limitation, linear response theory proves to be a potent framework for crafting effective—albeit *ad hoc*—estimators for various transport coefficients, especially those characterizing the coupling among different thermodynamic driving forces. A good illustration of such capability is presented in the work of Mangaud and Rotenberg,<sup>88</sup> where the authors investigate the transport properties of a solution within a slit pore under simultaneous pressure and chemical potential gradients. In such scenarios, transport coefficients can be appropriately defined utilizing eq 1.

The particular geometry considered in ref 88 is a slit pore of thickness  $h$  that is perpendicular to the  $z$  axis, while pressure and chemical potential gradients are both applied along the  $x$  axis, namely parallel to the walls of the slit pore. They consider two different perturbations to the Hamiltonian, namely,

$$\mathcal{A}_p(\mathbf{P}, \mathbf{Q}) = \frac{V \nabla_x \mathcal{P}}{N} \sum_{i=1}^N x_i$$

$$\mathcal{A}_\mu(\mathbf{P}, \mathbf{Q}) = -\frac{\nabla_x \mu}{1 + \alpha} \left[ \sum_{i=1}^{N_A} x_i - \alpha \sum_{i=N_A+1}^{N_A+N_B} x_i \right]$$

Here,  $N_A$  and  $N_B$  are the number of particles of type A and B, respectively,  $x_i = \mathbf{e}_x \cdot \mathbf{q}_i$  and  $\alpha = \rho_A^b / \rho_B^b$  is the ratio of bulk densities of A and B. (Note that A and B particles are indexed as 1, 2, ...,  $N_A$  and  $N_A + 1$ , ...,  $N_A + N_B$ , respectively.) Moreover,  $\mathcal{A}_\mu$  is formulated so that no net force is exerted on the solution in the bulk. In order to use the formalism of linear response theory, they define the following localized observables that signify volume and concentration fluxes:

$$\mathcal{B}_p(\mathbf{P}, \mathbf{Q}; z) = \frac{h}{N} \sum_{i=1}^{N_A} v_{x,i} \delta(z_i - z)$$

$$\mathcal{B}_\mu(\mathbf{P}, \mathbf{Q}; z) = \frac{h}{V} \left[ \frac{1}{1 + \alpha} \sum_{i=1}^{N_A} v_{x,i} \delta(z_i - z) - \frac{\alpha}{1 + \alpha} \sum_{i=N_A+1}^{N_A+N_B} v_{x,i} \delta(z_i - z) \right]$$

One can then use eq 2 to enumerate the response of each localized observable to the respective global perturbation. More precisely, the coefficients of the matrix  $\mathcal{M}$  given by

$$\lim_{t \rightarrow \infty} \begin{bmatrix} \langle \mathcal{B}_p(z, t) \rangle \\ \langle \mathcal{B}_\mu(z, t) \rangle \end{bmatrix} = - \begin{bmatrix} \mathcal{M}_{p,p} & \mathcal{M}_{p,\mu} \\ \mathcal{M}_{\mu,p} & \mathcal{M}_{\mu,\mu} \end{bmatrix} \begin{bmatrix} \nabla \mathcal{P} \\ \nabla \mu \end{bmatrix}$$

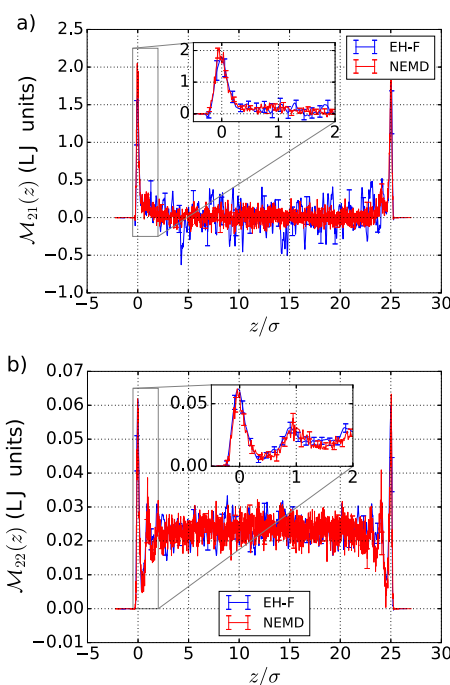
can be evaluated as

$$\mathcal{M}_{ij}(z) = \beta V \int_0^{+\infty} \langle \mathcal{B}_i(t) \dot{\mathcal{A}}_j(0) \rangle dt \quad (37)$$

where  $i, j \in \{p, \mu\}$ . For instance,  $\mathcal{M}_{p\mu}(z)$  and  $\mathcal{M}_{\mu\mu}(z)$  computed using eq 37 for a binary Lennard-Jones<sup>89</sup> (LJ) liquid within a slit pore is depicted in Figure 2, alongside their estimates obtained from NEMD. In particular, the authors relate the diagonal components of  $\mathcal{M}$  to viscosity and diffusivity. More precisely, they interpret  $\mathcal{M}_{\mu\mu}$  as a Fickian binary diffusivity  $D_{AB}$  along the  $x$  axis. By invoking continuum approximation, they assume that  $\mathcal{M}_{pp}(z)$  will take a form consistent with Poiseuille flow:

$$\mathcal{M}_{pp}(z) = \frac{\rho}{2\eta} \left[ \left( \frac{h}{2} \right)^2 - \left( z - \frac{h}{2} \right)^2 \right]$$

which enables them to fit the computed  $\mathcal{M}_{pp}(z)$  to a quadratic profile to estimate viscosity  $\eta$ . Similar to the first example, these estimators are still *ad hoc* in nature, as they do not yield a binary diffusivity profile that is compatible with the Smoluchowski formalism.



**Figure 2.** (a)  $\mathcal{M}_{p,\mu}(z)$  and (b)  $\mathcal{M}_{\mu,\mu}(z)$  computed using NEMD and eq 37 for a binary LJ fluid within a slit pore. (Reprinted from ref 88 with the permission of AIP Publishing.)

#### 4. KERNEL-BASED APPROACHES

Kernel-based methods are a class of methods in which particle positions (and displacements) are processed through the application of a *kernel* function—also referred to as a *filter*. The theoretical foundation of these methods is based on Itô's Lemma,<sup>90</sup> which specifies the temporal evolution of stochastic processes obtained by composing a standard Itô process with a  $C^2$  function. More precisely, suppose that  $\mathbf{X}_t \in \mathbb{R}^k$  is an Itô process, i.e., a stochastic process whose temporal evolution is described by the following stochastic differential equation (SDE):

$$d\mathbf{X}_t = \mathbf{p}(\mathbf{X}_t)dt + \boldsymbol{\pi}(\mathbf{X}_t)d\mathbf{W}_t \quad (38)$$

wherein  $\mathbf{W}_t$  is the  $k$ -dimensional Wiener process. Itô's Lemma states that the stochastic process  $\mathbf{Y}_t = \mathcal{F}(\mathbf{X}_t) \equiv [\mathcal{F}_1(\mathbf{X}_t), \mathcal{F}_2(\mathbf{X}_t), \dots, \mathcal{F}_n(\mathbf{X}_t)] \in \mathbb{R}^n$  will evolve according to the following SDE:

$$d\mathbf{Y}_t = \mathbf{q}(\mathbf{X}_t)dt + \boldsymbol{\sigma}(\mathbf{X}_t)d\mathbf{W}_t \quad (39)$$

with  $\mathbf{q} \in \mathbb{R}^n$  and  $\boldsymbol{\sigma} \in \mathbb{R}^{n \times k}$  given by

$$\mathbf{q}_i(\mathbf{X}_t) = \nabla \mathcal{F}_i^\dagger(\mathbf{X}_t) \mathbf{p}(\mathbf{X}_t) + \frac{1}{2} \boldsymbol{\pi}^\dagger(\mathbf{X}_t) \mathbf{H}_i(\mathbf{X}_t) \boldsymbol{\pi}(\mathbf{X}_t) \quad (40a)$$

$$\boldsymbol{\sigma}_i(\mathbf{X}_t) = \nabla \mathcal{F}_i^\dagger(\mathbf{X}_t) \boldsymbol{\pi}(\mathbf{X}_t) \quad (40b)$$

Here,  $\boldsymbol{\sigma}_i$  is the  $i$ -th column of  $\boldsymbol{\sigma}$  and  $\mathbf{H}_i$  is the Hessian of  $\mathcal{F}_i$ . It follows from the general theory of SDEs that the pointwise covariance matrix  $\boldsymbol{\sigma} \boldsymbol{\sigma}^\dagger$  can be estimated from individual realizations of (39) using<sup>90,91</sup>

$$\boldsymbol{\sigma}(\mathbf{r}) \boldsymbol{\sigma}(\mathbf{r})^\dagger = \lim_{h \rightarrow 0^+} \frac{\langle (\mathbf{Y}_{t+h} - \mathbf{Y}_t)(\mathbf{Y}_{t+h} - \mathbf{Y}_t)^\dagger \rangle}{2h} \bigg|_{\mathbf{X}_t=\mathbf{r}} \quad (41)$$

Note that the Smoluchowski equation can be viewed as a forward Kolmogorov equation associated with the overdamped Langevin SDE given by

$$d\mathbf{X}_t = -[\beta \mathbf{D}(\mathbf{X}_t) \cdot \nabla \mathcal{F}(\mathbf{X}_t) + \nabla \cdot \mathbf{D}(\mathbf{X}_t)]dt + \sqrt{2\mathbf{D}(\mathbf{X}_t)} d\mathbf{W}_t \quad (42)$$

Clearly, eq 42 falls within the broader category of SDEs described by (38) and (39). From Itô's lemma, an expression similar to the one yielding Kramer-Moyal coefficients can also be derived for a filtered trajectory by employing (40a) and (40b). More specifically, suppose that  $\gamma: \mathbb{R}^n \rightarrow \mathbb{C}$  is a piecewise  $C^2$  function. For an  $\mathbf{X}_t$  satisfying (42),  $Y_t = \gamma(\mathbf{X}_t)$  will evolve according to the following SDE:

$$dY_t = [\nabla \gamma(\mathbf{X}_t) \cdot \boldsymbol{\mu}(\mathbf{X}_t) + \mathbf{D}(\mathbf{X}_t) : \mathbf{H}_\gamma(\mathbf{X}_t)]dt + \nabla \gamma(\mathbf{X}_t) \cdot \sqrt{2\mathbf{D}(\mathbf{X}_t)} \cdot d\mathbf{W}_t$$

The associated covariance will thus be given by

$$\nabla \gamma^\dagger(\mathbf{r}) \mathbf{D}(\mathbf{r}) \nabla \gamma(\mathbf{r}) = \lim_{h \rightarrow 0^+} \frac{\langle |\gamma(\mathbf{X}_{t+h}) - \gamma(\mathbf{X}_t)|^2 \rangle}{4h} \bigg|_{\mathbf{X}_t \sim \nu} \quad (43)$$

The expectation given by (43) is proportional to the projection of the diffusivity tensor along the direction given by  $\nabla \gamma$ . The ability to use a filter function gives one an increased level of flexibility to design suitable estimators of diffusivity. It must be noted that eqs 41 and 43 can be easily reformulated if  $\mathbf{X}_t$  is drawn from a probability distribution  $\nu(\cdot)$ , which can be identical to the equilibrium probability distribution  $\rho_0(\cdot)$ ,

$$\int \sigma(\mathbf{r}) \sigma^\dagger(\mathbf{r}) \nu(\mathbf{r}) d\mathbf{r} = \lim_{h \rightarrow 0^+} \frac{\langle (\mathbf{Y}_{t+h} - \mathbf{Y}_t)(\mathbf{Y}_{t+h} - \mathbf{Y}_t)^\dagger \rangle}{2h} \bigg|_{\mathbf{X}_t \sim \nu(\mathbf{r})}$$

One of the first attempts to use general filters (as opposed to characteristic functions that “count” particles in a bin) to design diffusivity estimators in the context of molecular simulations was undertaken by our research group, as presented in a series of papers.<sup>92,93</sup> (Another example of using a similar expression—albeit with a characteristic function of a set as a filter—is the estimation of diffusivity in the collective variable space by Hegger and Stock.<sup>94</sup>)

In the initial paper<sup>92</sup> of the series, our focus was on the analytical derivation of the estimator and its numerical validation using synthetic data obtained from numerical integration of (42) with known *a priori* diffusivity profiles. All employed diffusivity profiles were functions of a single spatial variable. The subsequent paper<sup>93</sup> in this series delves into the application of the method to MD trajectories.

In the first paper,<sup>92</sup> we propose a filter function of the form  $\gamma_{\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{a}\mathbf{k} \cdot \mathbf{r}} G(\mathbf{r})$ . The complex exponential encodes information about the directionality of the diffusivity tensor (by yielding its projection along the unit vector  $\mathbf{k}$ ) while  $G(\cdot)$  is a localization function—also known as a *kernel*—that enables estimating diffusivity around a certain point in space. In practice,  $G(\cdot)$  can be defined as

$$G(\mathbf{r}) = \frac{1}{\epsilon^d} K_\epsilon \left( \frac{\mathbf{r} - \mathbf{r}_0}{\epsilon} \right)$$

wherein  $K_\epsilon(\mathbf{r}) \geq 0$  is chosen in a way that approximates the delta function as  $\epsilon \rightarrow 0^+$ . By applying eq 43 to the filter functions  $\gamma_\pm(\mathbf{r}) = e^{-i\mathbf{a}\mathbf{k} \cdot \mathbf{r}} \left[ G(\mathbf{r}) \pm \frac{1}{4} \right]$  and conducting some algebraic rearrangement, it can be demonstrated that

$$\int \mathbf{k}^T \mathbf{D}(\mathbf{r}) \mathbf{k} G(\mathbf{r}) \nu(\mathbf{r}) d\mathbf{r} = \lim_{h \rightarrow 0^+} \frac{1}{2\alpha^2 h} \times \langle [\gamma_{\mathbf{k}}(\mathbf{X}_{t+h}) - \gamma_{\mathbf{k}}(\mathbf{X}_t)]^* [f_{\mathbf{k}}(\mathbf{X}_{t+h}) - f_{\mathbf{k}}(\mathbf{X}_t)] \rangle_{\mathbf{X}_t \sim \nu(\cdot)}$$

By letting  $\epsilon \rightarrow 0$ , the integral on the left-hand side will converge to the pointwise estimate of  $D_{\mathbf{k}\mathbf{k}}$ , the diffusivity projected along  $\mathbf{k}$ , namely,

$$\hat{D}_{\mathbf{k}\mathbf{k}}(\mathbf{r}_0) \approx \frac{1}{2\alpha^2 h} \frac{\Re[\sum_{i=1}^{N_t} \Delta_h \gamma_{\mathbf{k}}^*(\mathbf{X}_{i,t}) \Delta_h f_{\mathbf{k}}(\mathbf{X}_{i,t})]}{\sum_{i=1}^{N_t} G(\mathbf{X}_{i,t})}$$

wherein  $\Delta_h g(\mathbf{X}_t) = g(\mathbf{X}_{t+h}) - g(\mathbf{X}_t)$  and the summation is conducted over  $N_t$  trajectories. By letting  $\alpha \rightarrow 0$ , one can obtain a limiting estimator given by

$$\hat{D}_{\mathbf{k}\mathbf{k}}^{\alpha \rightarrow 0}(\mathbf{r}_0) = \frac{1}{4h} \frac{\sum_{i=1}^{N_t} [G(\mathbf{X}_{i,t+h}) + G(\mathbf{X}_{i,t})][\mathbf{k} \cdot \Delta_h \mathbf{X}_{i,t}]^2}{\sum_{i=1}^{N_t} G(\mathbf{X}_{i,t})} \quad (44)$$

It is important to note that eq 44 offers a natural means of constructing an *ad hoc* extension of MSD in confined geometries, by assigning equal weight to particles that are present within a designated bin either at the beginning or at the end of an observation window. (Particles present within a bin both at the beginning and at the end of the observation window would contribute twice as much to the *ad hoc* MSD.)

We wish to note that the filter-based estimators proposed by us in ref 92 are closely related to Nadaraya-Watson (NW) estimators<sup>95,96</sup> in the statistics literature. NW estimators are designed to construct a regression function  $m(x)$  that relates two random variables,  $Y$  and  $X$ , with the condition that  $m(x) = \langle Y|X = x \rangle$ , all without resorting to parametric expressions. To create this function, the following Taylor expansion of  $m(x)$  centered at a specific point is introduced:

$$m(x) = \sum_i \sum_{j=0}^p \beta_j (X_i - x)^j K_\epsilon(x - X_i)$$

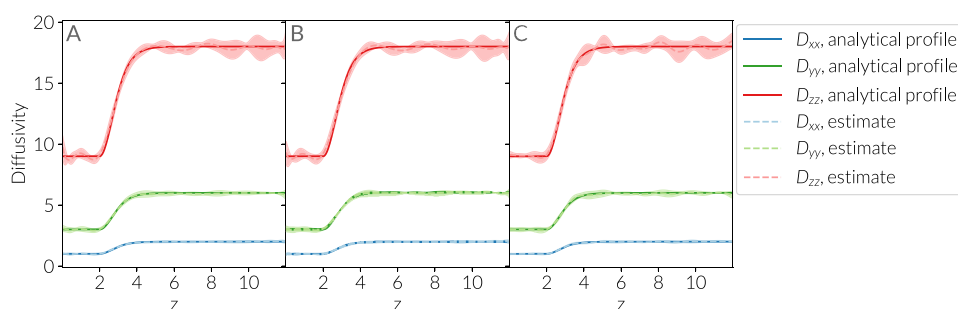
Here,  $p$  is the order of the Taylor expansion, and  $K_\epsilon(x - X_i)$  is a localization function that gives more weight to the  $X_i$ 's that are closer to  $x$ . Subsequently, a weighted least-squares problem is formulated to determine coefficients of the estimator within the Taylor expansion.

$$\hat{\beta}_\epsilon = \arg \min_{\beta \in \mathbb{R}^{p+1}} \sum_i \left[ Y_i - \sum_{j=0}^p \beta_j (X_i - x)^j \right]^2 K_\epsilon(x - X_i) \quad (45)$$

By setting  $p = 0$  in the expression provided above, we obtain the following explicit solution:

$$\hat{m}_\epsilon(x) = \hat{\beta}_{\epsilon,0} = \frac{\sum_i Y_i K_\epsilon(x - X_i)}{\sum_i K_\epsilon(x - X_i)} \quad (46)$$

Drawing inspiration from the pioneering works of Zmirou<sup>97</sup> and Bandi,<sup>98</sup> it has been proposed that a similar approach can be applied to estimate the Kramers-Moyal coefficient:



**Figure 3.** Application of eq 50 to the synthetic data wherein  $g_n(\cdot)$ 's are chosen from (A) Legendre and (B) Chebyshev polynomials, and (C) Fourier series. Solid lines and symbols correspond to the true diffusivity and the projection-based estimates, respectively. (Reproduced with permission from ref 92. Copyright 2023, American Chemical Society.)

$$\hat{\mathbf{D}}_{\epsilon,h}(t, \mathbf{r}) = \frac{1}{2h} \frac{\sum_{i=1}^N (\Delta_h \mathbf{X}_{i,t}) (\Delta_h \mathbf{X}_{i,t})^T K_{\epsilon}(\mathbf{r} - \mathbf{X}_{i,t})}{\sum_i K_{\epsilon}(\mathbf{r} - \mathbf{X}_{i,t})} \quad (47)$$

where  $\Delta_h \mathbf{X}_{i,t}$ 's are displacement samples drawn from realizations of (42). Similar to filtered estimators, eq 47 can be reformulated accordingly if  $\mathbf{X}_{i,t}$ 's (i.e., the starting points of stochastic trajectories) can be drawn from any probability distribution,  $\nu(\mathbf{r})$ :

$$\lim_{h \rightarrow 0^+} \langle \hat{\mathbf{D}}_{\epsilon,h}(t, \mathbf{r}) \rangle = \frac{\int \mathbf{D}(\mathbf{y}) K_{\epsilon}(\mathbf{y} - \mathbf{r}) \nu(\mathbf{y}) d\mathbf{y}}{\int K_{\epsilon}(\mathbf{y} - \mathbf{r}) \nu(\mathbf{y}) d\mathbf{y}} \equiv \tilde{\mathbf{D}}_{\epsilon}(t, \mathbf{r}) \quad (48)$$

Therefore, both filter- and kernel-based estimators can easily accommodate scenarios in which samples are drawn from nonequilibrium distributions. In simulations of equilibrium systems, the expression (47) can be further adapted by averaging it over  $t$ , with  $\nu(\mathbf{x})$  being substituted by  $\rho_0(\mathbf{x})$  in (48). Nevertheless, it is still imperative to observe the same considerations when selecting an appropriate time scale  $\tau$  for this approach. This approach can, in principle, be used for any time series data that can be modeled by eq 39. It has been used successfully in contexts as diverse as financial data,<sup>99</sup> electroencephalographic (EEG) data,<sup>100</sup> and stochastic descriptions of chaotic deterministic systems.<sup>100</sup>

We wish to highlight a subtle—but crucial—distinction between the FCE estimators derived in ref 92 and the NW-based estimators given by eq 48. Specifically, the former method applies the kernel on the particle's position at both the beginning and the end of the observation window, whereas the latter applies the kernel solely to the initial frame. This small contrast potentially affords FCEs a marginal edge in resolving the locality of diffusivity by more effectively attenuating the impact of interbin mixing on displacement statistics.

In ref 92, the majority of numerical tests were conducted using the triangle kernel,  $K_1(z) \propto (1 - |z|)\chi_{|z|<1}(z)$ . However, the choice of kernel did not significantly impact the reliability of diffusivity estimates. This is expected considering the observation that

$$\hat{D}_{\mathbf{k}\mathbf{k}}(\mathbf{r}_0) = D_{\mathbf{k}\mathbf{k}}(\mathbf{r}_0) + e^2 \mathbf{K}_2 : [\mathbf{H}_{D_{\mathbf{k}\mathbf{k}}}(\mathbf{r}_0) + 2\nabla D_{\mathbf{k}\mathbf{k}}(\mathbf{r}_0) \nabla^T \ln \rho_0(\mathbf{r}_0)] \quad (49)$$

where  $\mathbf{H}_f$  is the Hessian of the scalar function  $f$  and  $\mathbf{K}_2 = \frac{1}{2} \int \mathbf{y} \mathbf{y}^T G(\mathbf{y}) d\mathbf{y}$ . In other words, eq 49 illustrates that the variance of the kernel appears as a prefactor in front of the  $e^2$  term, but does not alter the fundamental scaling of systematic

error with  $\epsilon$ . Furthermore, beyond a certain threshold, reducing the value of  $\epsilon$  results in a kernel function with a significantly narrowed support, which adversely affects statistical accuracy, leading to the emergence of large error bars. This observation aligns with the theoretical expectation that, for small  $\epsilon$  values, the variance of the estimator should scale proportionally to  $\propto \epsilon^{-d} \int K_1^2(z) dz$ .

One of the advantages of using kernel-based methods is that the kernel function does not have to be localized, nor does it have to be nonnegative. For instance, one can choose a collection of kernels that belong to a family of orthogonal functions, such as Fourier series, or special polynomials (Figure 3). One can then express

$$\mathbf{D}(\mathbf{r})\nu(\mathbf{r}) = \sum_{n=1}^{+\infty} \mathbf{C}_n g_n(\mathbf{r})$$

The unknown coefficients within this sum can be projected onto  $\mathbf{k}$  in a similar fashion, and determined using

$$\begin{aligned} C_{n,\mathbf{k}\mathbf{k}} &= \mathbf{k}^T \mathbf{C}_n \mathbf{k} = \int D_{\mathbf{k}\mathbf{k}}(\mathbf{r}) g_n(\mathbf{r}) \nu(\mathbf{r}) d\mathbf{r} \\ &\approx \frac{\langle [g_n(\mathbf{X}_{t+h}) + g_n(\mathbf{X}_t)] [\mathbf{k} \cdot \Delta_h \mathbf{X}_t]^2 \rangle}{4h} \end{aligned} \quad (50)$$

Upon evaluating the computational performance of the FCE estimator, we found it to exhibit robust performance across various scenarios, provided that the drift term remains nondivergent. However, significant discretization errors were observed in cases where the drift diverges, particularly in the vicinity of hard boundaries. This phenomenon can be attributed to the increased susceptibility of the Gaussian approximation to temporal discretization in the presence of strong (diverging) drifts. For trajectories that are stochastic at all time scales, such errors can be effectively remedied by choosing a sufficiently small  $h$ .

In the second paper in this series,<sup>93</sup> we adopted the FCE estimator to trajectories generated via MD. This requires identifying a system- and position-dependent time scale  $\tau$  beyond which the system exhibits diffusive behavior. This was achieved by introducing the concept of a *cage escape time*, i.e., the characteristic time scale for a particle to escape the cage formed by its first coordination shell. More precisely, we proposed the following autocorrelation function,

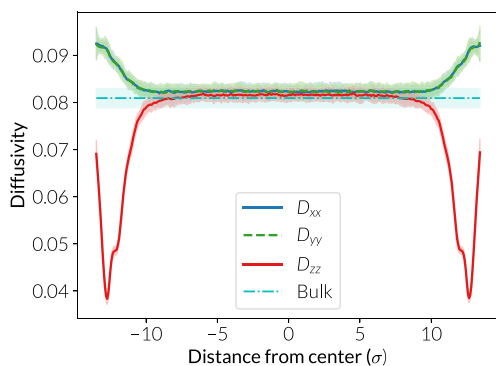
$$C(z, t) = \frac{\left\langle \sum_{i,j=1, i \neq j}^N \delta[z_i(0) - z] \xi(z, \mathbf{r}_{ij}(0)) \xi(z, \mathbf{r}_{ij}(t)) \right\rangle}{\left\langle \sum_{i,j=1, i \neq j}^N \delta[z_i(0) - z] \xi(z, \mathbf{r}_{ij}(0)) \xi(z, \mathbf{r}_{ij}(0)) \right\rangle} \quad (51)$$

Here,  $r_{\text{cage}}$  the first valley of the radial distribution function at  $z$ , and  $\xi(z, \mathbf{r}) = H[r_{\text{cage}}(z) - \|\mathbf{r}\|]$  with  $H(\cdot)$  the Heaviside function. Intuitively,  $C(z, t)$  corresponds to the fraction of the neighboring particles that remain within a distance  $r_{\text{cage}}(z)$  of a central particle after time  $t$  has elapsed.  $C(z, t)$  can thus be computed for particles belonging to each spatial bin and be fitted to a stretched exponential<sup>101</sup>  $C(z, t) = \exp[-(t/\tau_c(z))^{\alpha(z)}]$  to obtain a position dependent time scale  $\tau_c(z)$ . We wish to note that this approach can serve as a systematic means of determining a diffusive time scale in all methods for which the specification of such a time scale is necessary.

Nevertheless, the kernel-based estimators introduced in ref 92, exhibit a small—but systematic—underestimation of diffusivity when applied to MD trajectories, due to the presence of the caging regime that follows the culmination of the ballistic regime. This systematic error can, however, be readily remedied using a slightly modified form of the estimator, namely,

$$\hat{D}_{\text{kk}}^{h_1, h_2} = \frac{h_2 \hat{D}_{\text{kk}}^{h_2} - h_1 \hat{D}_{\text{kk}}^{h_1}}{h_2 - h_1} \quad (52)$$

where  $h_1$  and  $h_2$  constitute two observation windows within the diffusive regime. We applied the estimator given by eq 52 to an LJ fluid confined within a slit-pore, with the resulting diffusivity profiles depicted in Figure 4. An analysis conducted for the



**Figure 4.** Diffusivity estimate from (44) for a LJ fluid confined within a slit pore. (Reproduced with permission from ref 93. Copyright 2023, American Chemical Society.)

purposes of validating the estimator demonstrated that the computed  $D_{xx}$ ,  $D_{yy}$  profiles are accurate, but the predicted  $D_{zz}$  profile loses accuracy in the immediate vicinity of the wall due to the diverging drift. A Bayesian correction scheme based on diffusion maps<sup>102</sup> was then introduced and applied to rectify such inaccuracy.

Kernel-based estimators provide a robust extension of the concept of an *ad hoc* mean squared displacement, due to their conceptual clarity and computational simplicity. Similar to the Kramer-Moyal estimator presented in eq 24, kernel-based estimators offer the distinct advantage of accommodating a local time scale  $\tau$  for each position at which diffusivity is to be estimated. This feature proves advantageous when compared to other methodologies (discussed later in this review) that

require the usage of a fixed observation window for the whole system, as  $\tau_c$ , the time scale required to fully transition into the diffusive behavior might exhibit strong dependence on position in many circumstances. In approaches where a single time scale must be selected for discretizing the Smoluchowski equation across the entire system, it becomes imperative to choose  $\tau_{\text{max}} = \sup_z \tau_c(z)$  to ensure that ballistic effects have been overcome uniformly. It should be noted that, as the time scale  $\tau$  increases, the potential influence of drift effects on the accuracy of the diffusivity estimate  $\mathbf{D}(\mathbf{r})$  also becomes more pronounced.

Finally, it is pertinent to acknowledge that implementing this method requires determining certain free parameters, such as  $\epsilon$  in the case of localized kernels or the total number of basis functions  $N$  when employing a projection approach. The optimal selection of these parameters may necessitate a systematic exploration to attain the desired level of precision in the analysis. Schemes for choosing  $\epsilon$  have been proposed in the statistics literature in the context of Nadaraya-Watson estimators.<sup>98,103</sup>

## 5. BAYESIAN APPROACHES

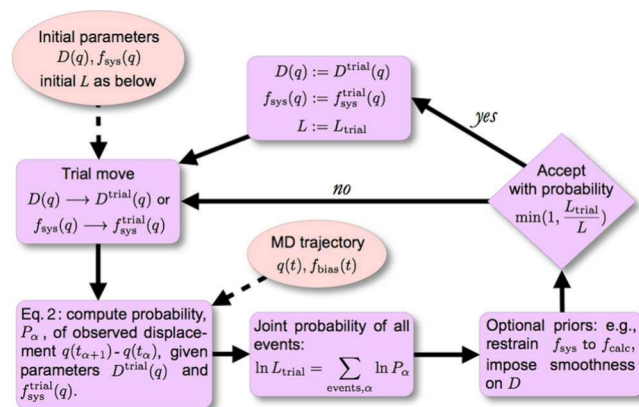
Bayesian approaches attempt to obtain a maximum likelihood estimate (MLE) of diffusivity from MD data, by means of constructing a suitable likelihood function,

$$\mathbf{P}(\mathbf{D}|\text{data}) \propto \mathbf{P}(\text{data}|\mathbf{D})$$

which is then maximized over the space of all plausible diffusivity profiles. While Bayesian approaches can be employed without any prior information about the diffusivity profile, it might, in many cases, be convenient to include preexisting information about diffusivity as a prior distribution, resulting in the following likelihood function:

$$\mathbf{P}(\mathbf{D}|\text{data}) \propto \mathbf{P}(\text{data}|\mathbf{D})\mathbf{P}(\mathbf{D}) \quad (53)$$

Maximizing (53) results in a *maximum a posteriori* (MAP) estimate of diffusivity. A schematic flowchart of Bayesian approaches is depicted in Figure 5.<sup>104</sup>



**Figure 5.** Schematic flowchart of applying Bayesian approaches to estimate position-dependent diffusivity. (Reproduced with permission from ref 104. Copyright 2013, American Chemical Society.)

Indeed, one of the most widely known and popular methods for estimating position-dependent diffusivity is a Bayesian approach proposed by Hummer,<sup>105</sup> which is also based on spatial discretization of the Smoluchowski operator. Let  $\mathbf{Q}$  be a stochastic matrix wherein  $Q_{ij}$  is the expected probability of transitioning from bin  $i$  to bin  $j$ , and let  $\mathbf{P}$  the matrix that

contains the actual transition probabilities obtained from MD. The likelihood that  $\mathbf{Q}$  accurately represents the data is given by

$$P(\text{data}|\mathbf{Q}) = \prod_{i,j=1}^{n_b} Q_{ij}^{P_{ij} n_{w,i}} \quad (54)$$

where  $n_{w,i}$  is the total number of observed transitions starting from the  $i$ -th bin. By taking the logarithm of both sides, and adding  $\pm \sum_{i=1}^{n_b} n_{w,i} \sum_{j=1}^{n_b} P_{ij} \log P_{ij}$  to the right-hand side, the log posterior probability can be expressed as

$$\log P(\mathbf{Q}|\text{data}) = - \sum_{i=1}^{n_b} n_{w,i} \sum_{j=1}^{n_b} P_{ij} \log \frac{P_{ij}}{Q_{ij}} + C_2 \quad (55)$$

Note that maximizing (55) over all stochastic matrices would trivially yield  $\mathbf{Q} = \mathbf{P}$ . One, however, needs to only conduct maximization over matrices that are consistent with diffusive behavior. In Hummer's approach,  $\mathbf{Q}$  is parametrized as  $\mathbf{Q} = e^{\mathbf{R}}$  wherein  $\mathbf{R}$  is a rate matrix with its entries satisfying the following properties:

$$r_{ij} = \begin{cases} r_{ij} & i > j \\ -\sum_{l \neq i} r_{il} & i = j \\ r_{ji} \rho_i / \rho_j & i < j \end{cases} \quad (56)$$

Here,  $\rho_j$  refers to the equilibrium probability of finding a particle at bin  $j$ , and the condition  $r_{ij} \rho_j = r_{ji} \rho_i$  is included to ensure detailed balance. Therefore, the rate matrix  $\mathbf{R}$  will possess  $n_f = \frac{n_b(n_b+1)}{2} - 1$  free entries. The associated optimization problem can be solved using a variety of methods, such as Monte Carlo sampling from a posterior distribution of the rate matrix  $\mathbf{R}$ :

$$\log P(\mathbf{R}|\text{data}) = - \sum_{i=1}^{n_b} n_{w,i} \sum_{j=1}^{n_b} P_{ij} \log \frac{P_{ij}}{[e^{\mathbf{R}}]_{ij}} + C_2 \quad (57)$$

We wish to note that such a posterior will generally be high-dimensional considering the quadratic scaling of  $n_f$  with  $n_b$ . However, since  $Q_{ij}$  approximates  $\approx \rho_i(\mathbf{r}_j|\mathbf{r}_i)$  as a discretized solution of the Smoluchowski equation, the ansatz  $\mathbf{Q} = e^{\mathbf{R}}$  implies that  $\mathbf{R}$  can be regarded as a discrete representation of the  $\mathcal{L}^\dagger$  operator (akin to the operator discretization methods described in Section 6, such as ref 106).

By drawing an analogy with finite difference discretization, constraints can be applied to the matrix  $\mathbf{R}$  to ensure its sparsity, predominantly preserving nonzero elements along few off-diagonal positions. Such constraints would effectively reduce the dimensionality of the posterior distribution, resulting in a linear scaling between  $n_f$  and  $n_b$ . Furthermore, this will allow the utilization of efficient diagonalization techniques for computing the matrix exponential in (57). Much in the same way as the operator discretization scheme of Sicardi et al.<sup>107</sup> (discussed in Section 6), Hummer employs finite differences to discretize  $\mathcal{L}^\dagger$  in one dimension, following Bicout and Szabo,<sup>108</sup> yielding the following relationship:

$$\frac{D_i + D_{i+1}}{2} = \Delta x^2 R_{i,i+1} \left( \frac{\rho_i}{\rho_{i+1}} \right)^{1/2} \quad (58)$$

where  $\Delta x$  is the thickness of each bin. This results in a matrix  $\mathbf{R}$  that is tridiagonal. In other words, within the scope of this approximation, particle exchanges primarily occur between adjacent bins (i.e., the instantaneous rate of exchange between nonadjacent bins is assumed to be zero). This expression provides a means of calculating the diffusivity profile using posterior samples of  $\mathbf{R}$  or through a maximum likelihood estimate of  $\mathbf{R}$ , achieved by optimizing (57). Similarly, (58) enables Monte Carlo sampling of  $\mathbf{D}$ , from which the components of the matrix  $\mathbf{R}$  defining the likelihood can be specified.

In the approach formulated above, the derivative of the diffusivity profile is left unconstrained, potentially leading to the emergence of rapid oscillations in  $D(x)$  due to inherent noise in the underlying MD data. To mitigate this, and consistent with the expectation that physical properties within a single thermodynamic phase are expected to be continuous functions of position, Hummer introduced the following prior distribution within the space of diffusivity profiles:

$$\mathbb{P}(\mathbf{D}) \propto \prod_i e^{-[D_i - D_{i+1}]^2 / 2\gamma^2} \quad (59)$$

We wish to note that there might be a more physically motivated prior than the one described by (59). Specifically, assuming the validity of the Smoluchowski equation, individual realizations of single-particle trajectories can be generated using the SDE given by eq 42 with a drift term given by

$$\boldsymbol{\mu}(\mathbf{r}) = \mathbf{D}(\mathbf{r}) \cdot \nabla \log \rho_0(\mathbf{r}) + \nabla \cdot \mathbf{D}(\mathbf{r}) \quad (60)$$

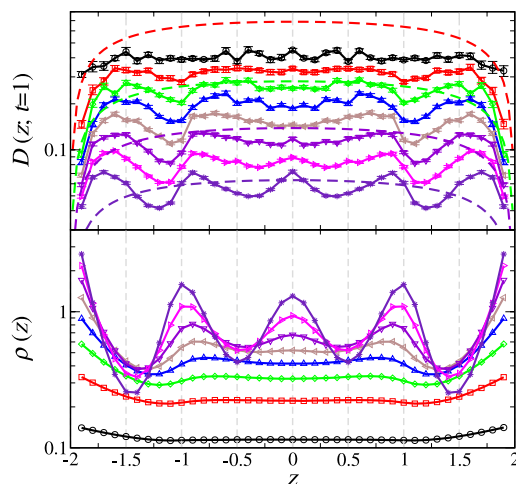
Multiplying both sides of (60) by  $\rho_0$  yields

$$\begin{aligned} \mathbf{D}(\mathbf{r}) \cdot \nabla \rho_0 + \rho_0 \nabla \cdot \mathbf{D}(\mathbf{r}) &= \nabla \cdot [\rho_0 \mathbf{D}(\mathbf{r})] = \rho_0 \boldsymbol{\mu}(\mathbf{r}) \\ \nabla \cdot [\rho_0 \mathbf{D}(\mathbf{r})] &\approx \frac{\rho_0}{\tau} \langle \mathbf{X}_{t+\tau} - \mathbf{X}_t \rangle_{\mathbf{X}_t=\mathbf{r}} + O(\tau^2) \end{aligned} \quad (61)$$

This equation establishes a connection between diffusivity,  $\mathbf{D}(\mathbf{r})$ , and drift,  $\boldsymbol{\mu}(\mathbf{r})$ , within the Smoluchowski framework. It serves as a necessary condition as it offers  $d$  equations, which are fewer than the required  $d(d+1)/2$  independent components of  $\mathbf{D}(\mathbf{r})$ . However, it provides an expression that bounds the spatial derivative of  $\mathbf{D}(\mathbf{r})$ , offering a means to define a prior distribution for  $\mathbf{D}(\mathbf{r})$  as an alternative to the one in eq 59.

As mentioned above, Hummer's method leverages a finite difference discretization of  $\mathcal{L}^\dagger$  in two fundamental ways. First, this discretization serves to reduce the dimensionality of the posterior, thereby enhancing the method's convergence rate. But crucially, the key contribution of the finite difference discretization method—originally introduced by Bicout and Szabo<sup>108</sup>—is its ability to establish a direct connection between the matrix  $\mathbf{R}$  and  $\mathbf{D}(\mathbf{r})$  through eq 58. It is important to note that this scheme is only valid in one dimension, and its extension to more intricate geometries or to alternative curvilinear coordinate systems would necessitate the development of appropriate discretization schemes. This will, in turn, result in relationships between  $\mathbf{R}$  and  $\mathbf{D}$  that are more complicated than (58). Thus, Hummer's approach cannot be readily applied to more complex settings despite its elegance and simplicity. As explained in Appendix A, however, this strategy can still be applied to infer  $D_{zz}(z)$  in situations wherein the diffusivity tensor is axisymmetric.

Hummer applied his Bayesian approach to a simple model system, as well as to the alanine dipeptide<sup>110</sup> model. Since then, however, several authors have employed his approach to compute diffusion constants, e.g., in confined hard sphere fluid<sup>109</sup> (Figure 6), and solutes within lipid bilayers.<sup>111,112</sup>



**Figure 6.** Normal diffusivity and number density profiles for hard sphere fluids of varying packing fractions confined within a slit pore. Diffusivities are computed using Hummer's approach.<sup>105</sup> (Reprinted with permission from ref 109. Copyright 2008, the American Physical Society.)

A conceptually similar approach was introduced by Ghysels et al.<sup>113</sup> who ensured the continuity of diffusivity as a function of position via expanding it using a series of orthogonal functions  $g_q(\mathbf{r})$ , namely  $\mathbf{D}(\mathbf{r}) = \sum_{q=1}^{\infty} a_q g_q(\mathbf{r})$ . By truncating the sum to a finite order, the authors applied Monte Carlo sampling to infer the posterior distribution of the coefficients within this expansion. Additionally, Ghysels et al. extended Hummer's approach to simultaneously infer radial and normal components of the diffusivity tensor in one-dimensional confinement in scenarios where the diffusivity tensor could be decomposed as  $\mathbf{D}(z) = D_{\parallel}(z)[\mathbf{e}_x \mathbf{e}_x^T + \mathbf{e}_y \mathbf{e}_y^T] + D_{\perp}(z)\mathbf{e}_z \mathbf{e}_z^T$ . The differential operator  $\mathcal{L}^{\dagger}$  of eq 16 can then be expressed in cylindrical coordinates:

$$\mathcal{L}^{\dagger} f = \frac{D_{\parallel}(z)}{r} \frac{\partial}{\partial r} \left( r \frac{\partial f}{\partial r} \right) + \frac{\partial}{\partial z} \left[ D_{\perp}(z) \rho_0(z) \frac{\partial}{\partial z} \left( \frac{f}{\rho_0(z)} \right) \right] \quad (62)$$

The second term on the right-hand side corresponds precisely to  $\mathcal{L}^{\dagger}$  in one dimension. Therefore, employing separation of variables, the authors demonstrated the feasibility of using Hummer's method to construct a matrix  $e^{tR}$  and estimate  $D_{\perp}(z)$  by minimizing (57) and utilizing (58). In order to also determine  $D_{\parallel}(z)$ , they devised a Bayesian scheme in which the transition matrix  $W_{i,mj}$  is constructed for a given  $\mathbf{D}(z)$  to yield the probability of transitioning from bin  $i$  to bin  $j$  in the  $z$  direction, while exhibiting a lateral mobility corresponding to  $m$ -th radial bin:

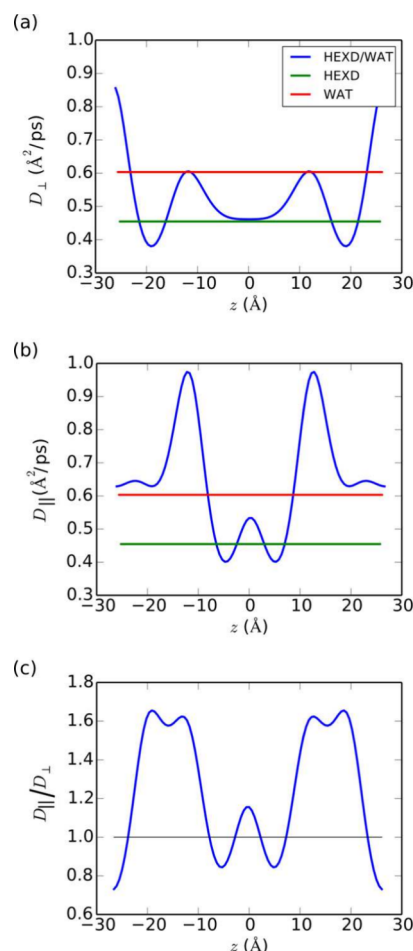
$$W_{i,mj} \approx \sum_{k=1}^{+\infty} \frac{J_0(\alpha_k r_m)}{\pi s^2 J_1^2(x_k)} [e^{[R - \alpha_k \text{diag}(D_{\parallel})]t}]_{ij} \quad (63)$$

Here,  $\text{diag}(D_{\parallel})$  denotes a diagonal matrix, with its entries being  $D_{\parallel}(z_i)$ 's, where  $z_i$  represents the center of the  $i$ -th bin

along the  $z$  direction.  $J_0(\cdot)$  is the zeroth order Bessel function of the first kind. In order to construct a series solution (rather than an integral of Bessel functions), the authors chose a sufficiently large distance  $s$  where they imposed an artificial absorbing boundary condition, i.e.,  $\rho(r = s, z, t | r = 0, z_0, 0) = 0$ . Therefore,  $\alpha_k$ 's are given by  $\alpha_k = \lambda_k/s$  wherein  $\lambda_k$  is the  $k$ -th smallest positive root of  $J_0(x)$ . By also defining circular bins for radial mobility, an empirical equivalent of  $W$ , denoted by  $E_{i,mj}$ , can be estimated from MD. After determining  $D_{\perp}(z)$ , an MLE estimate of  $D_{\parallel}(z)$  can be obtained by minimizing:

$$-\log P(\mathbf{W}|\text{data}) = n_w \sum_{i,j,m=1}^{n_b} E_{i,mj} \log \frac{E_{i,mj}}{W_{i,mj}} + C_2$$

The utilization of a finite-order expansion in terms of orthogonal functions guarantees the smoothness of both  $D_{\perp}(z)$  and  $D_{\parallel}(z)$ . They employed their methodology to characterize oxygen diffusion within organic membranes (Figure 7). Since its development, this approach has found widespread application in molecular simulations, particularly for probing diffusion across membranes.<sup>104,113</sup> It has also been used to predict diffusivity of colloid suspensions<sup>114</sup> as well in collective variable spaces for protein folding.<sup>82,115</sup>



**Figure 7.** (a) Normal and (b) lateral diffusivity profiles of oxygen across a hexadecane/water film alongside (c) the anisotropy ratio computed from the Bayesian approach of Ghysels et al.<sup>113</sup> (Reproduced with permission from ref 113. Copyright 2017, American Chemical Society.)

Motivated by the success of these Bayesian approaches, several other Bayesian methodologies have been developed for estimating diffusivity. One example is a method proposed by Comer et al.,<sup>104</sup> sometimes referred to as the adaptive biasing force (ABF) method. A crucial distinction between their approach and that of Hummer lies in the absence of spatial discretization in the former. Instead, it is based on the observation that  $\mathbf{X}_t$ , the stochastic process associated with the Smoluchowski equation, exhibits a Gaussian distribution over short times. More precisely, for a sufficiently small  $h$ :

$$\mathbf{X}_h - \mathbf{r}_0 \sim \mathcal{N}(\boldsymbol{\mu}(\mathbf{r}_0)h, 2\mathbf{D}(\mathbf{r}_0)h) \quad (64)$$

which is just a restatement of (29) in terms of  $\mathbf{X}_h$ . Using eq 64, it is possible to calculate the probability of observing a particular single-particle trajectory  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  leading to the following expression,

$$P(\{\mathbf{x}_i\}_{i=1}^n) = \prod_{i=1}^n p_h(\mathbf{x}_i|\mathbf{x}_{i-1}) \approx \frac{1}{(4\pi h)^{Nd/2}} \exp\left\{-\sum_{i=1}^n \left[ \frac{(\mathbf{x}_i - \mathbf{x}_{i-1} - h\boldsymbol{\mu}(\mathbf{x}_{i-1}))^2}{4h|\mathbf{D}(\mathbf{x}_{i-1})|^2} + \frac{\log|\mathbf{D}(\mathbf{x}_{i-1})|}{2} \right]\right\}$$

which can be viewed as a likelihood function for a particular diffusivity and drift profile:

$$P(\{\mathbf{x}_i\}_{i=1}^n) \approx \mathbb{P}[\{\mathbf{x}_i\}_{i=1}^n | \mathbf{D}(\mathbf{x}), \boldsymbol{\mu}(\mathbf{x})] \quad (65)$$

Here, the drift term is given by (60). The log-likelihood of  $\mathbf{D}(\mathbf{x})$  given an observed trajectory will thus be given by

$$-\log \mathbb{P}[\{\mathbf{x}_i\}_{i=1}^n | \mathbf{D}(\mathbf{x}), \boldsymbol{\mu}(\mathbf{x})] \approx \frac{1}{2} \sum_{i=1}^n \log |\mathbf{D}(\mathbf{x}_{i-1})| + \sum_{i=1}^n \frac{\|\mathbf{x}_i - \mathbf{x}_{i-1} - h\boldsymbol{\mu}(\mathbf{x}_{i-1})\|^2}{4h|\mathbf{D}(\mathbf{x}_{i-1})|} + C_1 \quad (66)$$

While  $\rho_0$  can be independently estimated from an equilibrium simulation, and the expression above could be considered as a likelihood for  $\mathbf{D}(\mathbf{x})$  alone, the authors choose to treat  $\mathbf{f} = \nabla \log \rho_0$  as a function to be determined through the Bayesian optimization scheme. Ultimately, the total log likelihood is derived by summing (66) over all observed trajectories.

Given that it is not necessary for the trajectories  $\{\mathbf{x}_i\}_{i=1}^n$  entering (66) to be discretized, it becomes more convenient to parametrize both  $\boldsymbol{\mu}(\mathbf{x})$  and  $\mathbf{D}(\mathbf{x})$  using predefined functional forms. The Bayesian scheme will then be used to obtain the unknown parameters. In the particular one-dimensional case considered in ref 104, the authors define a grid with regular spacing  $l$ , namely  $q_i = il + q_0$ . Within the  $i$ th cell,  $D(x)$  is expressed using cubic interpolants of the following mathematical form:

$$D(x) = a_i + \sum_{k=1}^3 a_i^{(k)} \left( \frac{x - q_i}{l} \right)^k \quad (67)$$

with  $a_i^{(j)}$ 's given by

$$\begin{cases} a_i^{(1)} = \frac{a_{i+1} - a_{i-1}}{2} \\ a_i^{(2)} = \frac{-a_{i+2} + 4a_{i+1} - 5a_i}{2} \\ a_i^{(3)} = \frac{a_{i+2} - 3a_{i+1} + 3a_i}{2} \end{cases}$$

A similar approach is utilized for parametrizing  $\boldsymbol{\mu}(\mathbf{x})$ . Upon selecting a parametrization for  $D(x)$ , it will become possible to derive analytical expressions for the gradients in (60) in terms of the unknown parameters. As in the method proposed by Hummer,<sup>105</sup> the authors incorporate prior distributions, which can be applied in tandem with (66) to penalize rapid oscillations in  $D(x)$ . Specifically, they employ the prior (59), alongside:

$$\log \mathbb{P}(\{a_i\}) = -\sum_i \log a_i \quad (68)$$

which ensures that  $\log a_i$  are sampled uniformly.<sup>104</sup> It must be emphasized that the choice of functional forms used for parametrizing  $D(x)$  is by no means unique. For instance,  $D(x)$  and  $\boldsymbol{\mu}(\mathbf{x})$  could be parametrized using a neural network, and (66) could be used as a loss function for its training. The authors present numerical comparisons between their method and Hummer's approach, indicating general qualitative agreement.

We would like to comment on some numerical aspects of this approach. While it circumvents the need for imposing a spatial discretization of the Smoluchowski operator and the associated numerical inaccuracies, the accuracy of the log-likelihood function used therein relies significantly on the Gaussian approximation of (64). This approximation is only valid for short times, and its validity diminishes as the strength of the PMF applied to the particles increases. We anticipate that this could pose challenges, particularly in situations involving hard boundaries close to which  $\nabla \log \rho_0(\mathbf{x})$  will diverge. Instead, methodologies that rely on spatial discretization might offer enhanced accuracy in such circumstances, as they frequently incorporate a zero-flux boundary condition to address the effects of hard boundaries.

We wish to note that one can introduce another Bayesian approach by noting that the logarithm of Hummer's prior, given by eq 59 in one dimension, can be viewed as a Riemann sum. More precisely, by defining  $\gamma = \tilde{\gamma}\sqrt{\Delta x}$ , it can be demonstrated that

$$-\log \mathbb{P}(D) = \sum_i \frac{|D_i - D_{i+1}|^2}{2\gamma^2} = \frac{1}{2\tilde{\gamma}^2} \sum_i \left| \frac{D_i - D_{i+1}}{\Delta x} \right|^2 \Delta x \approx \frac{1}{2\tilde{\gamma}^2} \int \left| \frac{d}{dx} D(x) \right|^2 dx$$

Therefore, in order to find a MAP estimator, one could attempt to minimize the posterior probability by computing its functional derivative with respect to  $D(x)$ :

$$\begin{aligned} \frac{\delta}{\delta D(x)} \log P[D(x)|\text{data}] &= \frac{\delta}{\delta D(x)} \log P[\text{data}|D(x)] \\ &+ \frac{1}{2\tilde{\gamma}^2} \frac{\delta}{\delta D(x)} \int \left| \frac{d}{dx} D(x) \right|^2 dx \end{aligned} \quad (69)$$

Note that the second term will be proportional to  $D''(x)$ . Therefore, setting eq 69 equal to zero leads to a second-order differential equation for  $D(x)$ , which can be solved through a variety of means, provided that one can calculate the functional derivative of the log-likelihood. A somewhat analogous approach, while adhering to the same principles, was introduced by Chang et al.<sup>116</sup> In line with Bayesian approaches in one dimension, the authors perform functional derivatives on  $g(x)$ , which is related to diffusivity through  $D(x) = D_0 e^{g(x)}$ . This ensures that  $D(x)$  remains a non-negative function. In contrast to Hummer's prior, they propose the following prior:

$$\begin{aligned} \log \mathbb{P}(D) &= -\frac{1}{2} \int g(x) \left[ \frac{1}{\beta \sqrt{2\pi\gamma}} e^{-\gamma d^2/dx^2} \right] g(x) dx \\ &\equiv -\frac{1}{2} \int g(x) \mathcal{R}_{\beta,\gamma}[g(x)] dx \end{aligned} \quad (70)$$

The differential operator  $\mathcal{R}_{\beta,\gamma}$ , corresponds to the exponential of the second derivative, with  $\beta$  and  $\gamma$  serving as regularization parameters—akin to  $\gamma$  in eq 59—that control the degree to which rapid oscillations in diffusivity are penalized. Much like the methodology proposed by Comer et al.<sup>104</sup> the log-likelihood is given by eq 66, and a prior with the same structure as eq 70 is applied to the drift term  $\mu(x)$ . This yields a system of coupled differential equations, pertaining to  $\mu(x)$  and  $g(x)$ , arising from the conditions  $\delta \log P(g(x), \mu(x)|\text{data})/\delta g(x) = 0$  and  $\delta \log P(g(x), \mu(x)|\text{data})/\delta \mu(x) = 0$ , which are solved simultaneously. To simplify the presentation, we primarily focus on  $g(x)$  given that  $\mu(x)$  can be readily expressed in terms of  $D(x)$  and  $\rho_0(x)$ . The functional derivative of eq 66 yields

$$\mathcal{R}_{\beta,\gamma}[g(x)] = \sum_{\alpha=1}^{N_{\text{traj}}} f_{\alpha}(x) \quad (71)$$

where summation is over the data coming from distinct trajectories. Each trajectory contributes to the corresponding ODE with a forcing term given by

$$\begin{aligned} f_{\alpha}(x) &= -\frac{1}{2} \sum_{i=1}^n \frac{d}{dx} \{ \delta(x - x_i^{(\alpha)}) [x_i^{(\alpha)} - x_{i-1}^{(\alpha)} - \mu(x_{i-1}^{(\alpha)})h] \} \\ &\quad - \frac{1}{2} \sum_{i=1}^n \delta[x - x_i^{(\alpha)}] \left[ 1 - \frac{|x_i^{(\alpha)} - x_{i-1}^{(\alpha)}|^2 - \mu^2(x_{i-1}^{(\alpha)})h^2}{2hD(x_{i-1}^{(\alpha)})} \right] \end{aligned} \quad (72)$$

Given that the right-hand side of eq 71 comprises a superposition of Dirac masses, and considering the linearity of the equation, we can multiply both sides by a Green's function associated with a singular Dirac mass. Denoting this Green's function as  $G_{\beta,\gamma}(x, y)$  for a source located at  $y$  yields

$$\begin{aligned} g(x) &= \\ &\frac{1}{2} \sum_{i=1}^n \sum_{\alpha=1}^{N_{\text{traj}}} \left\{ \frac{\partial G_{\beta,\gamma}(x, x_i^{(\alpha)})}{\partial x} [x_i^{(\alpha)} - x_{i-1}^{(\alpha)} - \mu(x_{i-1}^{(\alpha)})h] \right. \\ &\quad \left. - G_{\beta,\gamma}[x, x_i^{(\alpha)}] \left[ 1 - \frac{|x_i^{(\alpha)} - x_{i-1}^{(\alpha)}|^2 - \mu^2(x_{i-1}^{(\alpha)})h^2}{2hD(x_{i-1}^{(\alpha)})} \right] \right\} \end{aligned} \quad (73)$$

Since this expression is valid for any point  $x$  within the simulation domain, it should also hold for the  $x_j^{(\alpha)}$ 's, the points

along all trajectories. Evaluating  $g(x)$  at each  $x_j^{(\alpha)}$  results in a large system of nonlinear equations, which can be solved numerically to determine  $g[x_j^{(\alpha)}]$ 's. Once known, eq 73 can be employed to extrapolate  $g(x)$  to all points. To successfully carry out this procedure, prior knowledge of  $G_{\beta,\gamma}(x, y)$  is required. The precise mathematical form of  $G_{\beta,\gamma}(x, y)$  is, however, only known in the absence of hard boundaries. In ref 116, the authors consider  $x \in [0, \infty)$ , which is a semi-infinite domain. As such, the Green function takes the following mathematical form:

$$G_{\beta,\gamma}(x, y) = \beta [e^{-(x-y)^2/2\gamma} - e^{-(x+y)^2/2\gamma}]$$

It is worth mentioning that as part of the procedure to determine this Green's function, the authors introduce a boundary condition at  $x = 0$ , namely  $g(0) = 0$ . Boundary conditions are generally necessary when formulating MAP estimators as differential equations. In cases involving complex geometries where a closed-form expression for  $G_{\beta,\gamma}(x, y)$  may not be readily available, one might start from (71) and (72) and explore alternative means of solving partial differential equations.

The procedure outlined above can be applied iteratively for different choices of regularization parameters,  $\beta$  and  $\gamma$ . The authors provide a Bayesian framework for sampling these parameters, employing an approximate maximum marginal likelihood approach. This results in a posterior distribution for  $\beta$  and  $\gamma$ , which can be utilized to derive error estimates for the fitting procedure. As mentioned earlier, it is crucial to emphasize that the specific form of the differential operator is contingent on the chosen regularization approach. Consequently, one can, in principle, consider suitable alternatives to (70).

Lastly, it is worth mentioning an intermediate approach proposed by Türkcan et al.<sup>117</sup> bridging elements from both Hummer's<sup>105</sup> and Comer et al.'s<sup>104</sup> methods. This approach is intermediate in the sense that it involves spatial discretization akin to the one used by Hummer, but the transition matrix is constructed using the Gaussian approximation rather than a matrix exponential. More precisely, the observation domain is partitioned into  $n_b$  bins  $\{\mathcal{B}_i\}_{i=1}^{n_b}$ , and the transition matrix is denoted with  $Q$  wherein  $Q_{ij}^h$  is the probability of transitioning for  $\mathcal{B}_i$  to  $\mathcal{B}_j$  over a time increment  $h$ , for a given spatial profile of  $\{\mathbf{D}_i, \boldsymbol{\mu}_i\}_{i=1, \dots, N_{\text{bin}}}$ . The likelihood function is thus expressed as

$$-\log \mathbb{P}(\text{data}|\mathbf{D}_i, \boldsymbol{\mu}_i) = -\sum_{i,j} \log Q_{ij}$$

with  $Q_{ij}$ 's estimated from multiplying contributions from individual trajectories, namely:

$$Q_{ij} = \prod_{\alpha=1}^{n_{\text{traj}}} Q_{ij}^{(\alpha)} \quad (74)$$

Here,  $Q_{ij}^{(\alpha)}$  corresponds to the likelihood associated with trajectory  $\alpha$  and is given by

$$Q_{ij}^{(\alpha)} = \prod_{\mathbf{x}_{k-1}^{(\alpha)} \in \mathcal{B}_i, \mathbf{x}_k^{(\alpha)} \in \mathcal{B}_j} p_h(\mathbf{x}_k^{(\alpha)}|\mathbf{x}_{k-1}^{(\alpha)}; \mathbf{D}_i, \boldsymbol{\mu}_i) \quad (75)$$

Taking logarithms from both sides of (74) yields

$$\log Q_{ij} = \sum_{a=1}^{n_{\text{traj}}} \sum_{k=1}^n \chi_{\mathcal{B}_i}(\mathbf{x}_{k-1}^{(a)}) \chi_{\mathcal{B}_j}(\mathbf{x}_k^{(a)}) \times \log p_h(\mathbf{x}_k^{(a)} | \mathbf{x}_{k-1}^{(a)}; \mathbf{D}_i, \boldsymbol{\mu}_i) \quad (76)$$

with  $\log p_h(\mathbf{y} | \mathbf{x}; \mathbf{D}_i, \boldsymbol{\mu}_i)$  the log of the transition probability calculated from a single trajectory using the following Gaussian approximation:

$$-\log p_h(\mathbf{x}_k^{(a)} | \mathbf{x}_{k-1}^{(a)}; \mathbf{D}_i, \boldsymbol{\mu}_i) \chi_{\mathcal{B}_i}(\mathbf{x}_{k-1}^{(a)}) \approx \begin{cases} \frac{\|\mathbf{x}_k - \mathbf{x}_{k-1} - \boldsymbol{\mu}_i h\|^2}{4h|\mathbf{D}_i|} + \frac{1}{2} \log |\mathbf{D}_i| & \mathbf{x}_{k-1} \in \mathcal{B}_i \\ 0 & \text{otherwise} \end{cases}$$

Note that the indices  $i$  and  $j$  pertain to spatial bins, while  $\mathbf{x}_k$ 's represent points along a discrete-time trajectory, with the index  $k$  tracking time increments.

Akin to the Bayesian methods that are based on Gaussian approximations, this approach does not explicitly include any boundary information. It might therefore exhibit diminished accuracy in the presence of hard boundaries in comparison to Hummer's approach. One notable contribution of Türkcan et al.<sup>117</sup> is to apply their methodology to trajectories obtained from confocal microscopy experiments. In order to account for uncertainties in position measurements, they add an isotropic noise in their definition of the transition matrix, namely:

$$-\log[p_h(\mathbf{x}_k | \mathbf{x}_{k-1})] \chi_{\mathcal{B}_i}(\mathbf{x}_{k-1}) \approx \begin{cases} \frac{\|\mathbf{x}_k - \mathbf{x}_{k-1} - \boldsymbol{\mu}_i h\|^2}{4h|\mathbf{D}_i|} & \mathbf{x}_{k-1} \in \mathcal{B}_i \\ + \frac{1}{2} \log \left| \mathbf{D}_i + \frac{\sigma^2}{h} \mathbf{I} \right| & \\ 0 & \text{otherwise} \end{cases}$$

where  $\sigma$  is the measurement noise, and  $\mathbf{I}$  is the identity matrix. This approach has since been used to probe the diffusive motion of fluorescently tagged proteins at the surface of a cell.<sup>118</sup>

## 6. OPERATOR DISCRETIZATION APPROACHES

Operator discretization approaches are developed based on the recognition that the Smoluchowski equation, eq 16, possesses a formal solution given by  $\rho_t(\mathbf{y} | \mathbf{x}) = e^{t\mathcal{L}^\dagger} \delta(\mathbf{y} - \mathbf{x})$ , and rely on the spatial discretization of either  $\mathcal{L}_y^\dagger$  or  $e^{t\mathcal{L}^\dagger}$  to estimate diffusivity without invoking Bayes's theorem. As an illustration, consider the approach of Sicardi et al.<sup>107</sup> who use Markov state models<sup>119</sup> to estimate position-dependent diffusivity. In this approach, the simulation box is discretized into  $n_b$  bins, and an empirical transition matrix  $M_{ij}^{(\tau)}$  is computed from MD trajectories, recording the number of times a particle moves from bin  $i$  to bin  $j$  over a time window  $\tau$ . This approach utilizes the same approximation as in (24) but expresses it in terms of expectations of the solution of the Smoluchowski equation:

$$\begin{aligned} & \langle (\mathbf{X}_{t+\tau} - \mathbf{X}_t)(\mathbf{X}_{t+\tau} - \mathbf{X}_t)^T \rangle |_{\mathbf{X}_t=\mathbf{r}_i} \\ &= \int (\mathbf{y} - \mathbf{x})(\mathbf{y} - \mathbf{x})^T \rho_t(\mathbf{y} | \mathbf{x}) d\mathbf{y} \end{aligned} \quad (77)$$

$$\approx \sum_j (\mathbf{r}_j - \mathbf{r}_i)(\mathbf{r}_j - \mathbf{r}_i)^T M_{ij}^{(\tau)} \equiv \mathbf{C}_i^{(2),\tau} \quad (78)$$

A similar expression can be obtained for the drift:

$$\begin{aligned} & \langle \mathbf{X}_{t+\tau} - \mathbf{X}_t \rangle |_{\mathbf{X}_t=\mathbf{r}_i} = \int (\mathbf{y} - \mathbf{x}) \rho_t(\mathbf{y} | \mathbf{x}) d\mathbf{y} \\ & \approx \sum_j (\mathbf{r}_j - \mathbf{r}_i) M_{ij}^{(\tau)} \equiv \mathbf{C}_i^{(1),\tau} \end{aligned} \quad (79)$$

Note that the integrals in (77) and (79) are expressed in terms of the nominal solution of the Smoluchowski equation, namely  $\rho_t(\mathbf{y} | \mathbf{x})$ , and are further discretized based on the employed spatial binning. One can then relate eq 78 to diffusivity by employing eqs 24 and 26a. More specifically, it can be demonstrated using the partial correction in (26a) that

$$\mathbf{C}_i^{(2),\tau} - \mathbf{C}_i^{(1),\tau} [\mathbf{C}_i^{(1),\tau}]^T = 2\tau \mathbf{D}(\mathbf{r}) + O(\tau^3) \quad (80)$$

Similar to many other techniques discussed in this review, implementing this approach requires selecting for the entire system a uniform transition time scale,  $\tau$ , which cannot be chosen to be arbitrarily small due to delayed transition into the diffusive regime within MD trajectories. As discussed earlier, this might cause difficulties in probing systems with substantial dynamical heterogeneity. A possible means of resolving this issue is to choose  $\tau$  as the smallest time scale beyond which relaxation times computed from the eigenvalues of the transition matrix become insensitive to  $\tau$ .<sup>107</sup> It is crucial to acknowledge that employing a larger  $\tau$  diminishes the accuracy of the diffusivity estimate not only due to temporal discretization errors but also because of the prevalence of drift effects.

It is essential to highlight that determining the full diffusivity tensor using this methodology requires binning the simulation domain across all dimensions, even in cases where confinement is unidimensional. Similar to what was discussed in Hummer's approach, however, it is feasible to apply this method only for estimating  $D_{zz}$ , the normal component of the diffusivity tensor, while resorting to alternative approaches, such as kernel-based methods, to estimate  $D_{xx}$  and  $D_{yy}$ . Additionally, it is critical to take into account overarching considerations related to system discretization to ensure accurate construction of a Markov state model.

An alternative—but related—approach involves discretizing  $\mathcal{L}^\dagger$  instead of  $e^{t\mathcal{L}}$ , as proposed by Palmer et al.<sup>106</sup> In this method, diffusivity is treated as a free parameter to be optimized by minimizing the following objective function:

$$\chi(\mathbf{D}) = \sum_s \sum_{\{ij\}} [e^{t_s \mathcal{L}(\mathbf{D})} \Delta_{ij} - C_{ij}(t_s)]^2 \quad (81)$$

Here, the index  $s$  runs over time increments, while  $i$  and  $j$  correspond to the bins employed for spatial discretization of the simulation box.  $C_{ij}(t)$  is a time correlation function for the number of particles in  $i$ -th and  $j$ -th bins, and is defined as

$$C_{ij}(t) = \langle N_i(t) N_j(0) \rangle$$

while  $\Delta_{ij}$ 's correspond to equilibrium correlations between bin occupancies:

$$\Delta_{ij} = \langle N_i(0) N_j(0) \rangle$$

The main intuition behind this method is the expectation that the temporal evolution of  $C_{ij}(t)$  will be suitably described by

$e^{tL}\Delta_{ij}$  wherein  $L$  is a discretized version of the continuous operator  $\mathcal{L}^\dagger$  using the same spatial discretization employed in computing  $\Delta_{ij}$  and  $C_{ij}(t)$ . However, instead of using the operator that defines the spatial part of the Smoluchowski equation, the authors consider the following simplified operator:

$$\mathcal{L}_D^\dagger f = \nabla \cdot [\mathbf{D}(\mathbf{r}), \nabla f]$$

which corresponds to a diffusive process with a position-dependent diffusivity tensor but without a drift term. The authors incorporate drift indirectly by means of applying a no-flux boundary condition. In turn, the operator is discretized as a matrix using finite differences, and the boundary conditions appear explicitly in the discretization scheme. The discretized operator  $L(\mathbf{D})$  is a matrix whose entries depend on the diffusivity profile, which can then be used as an independent variable to minimize the objective function (81).

It is worth noting that the approach, as currently described, assumes that the drift term can be satisfactorily represented by a no-flux boundary condition. We wish to note that many confined liquids exhibit strong layering at substrates, even when fluid-wall interactions are short-range in nature. Under such circumstances, the absence of an explicit drift term can cause systematic errors in the estimated diffusivity profiles. Moreover, it is crucial to emphasize that the outlined procedure involves the computation of matrix exponentials, which, in turn, requires efficient matrix diagonalization during each iteration.

Finally, we wish to discuss the work of Schulz et al.,<sup>120</sup> which is strikingly similar to Hummer's Bayesian approach.<sup>105</sup> In this approach, the simulation domain is discretized into bins, and the transition probability between bins is described using a tridiagonal rate matrix,  $\mathbf{R}$ :

$$Q_{ij}(t) = [e^{t\mathbf{R}}]_{ij}$$

It is important to note that this approach is fundamentally an operator discretization approach, as the matrix  $\mathbf{R}$  serves as a discretization of the differential operator  $\mathcal{L}^\dagger$ . However, instead of using Bayes' theorem, as in Hummer,<sup>105</sup> Schulz et al. minimize the mean-squared error (MSE) given by

$$C(Q(\mathbf{R}), P) = \frac{1}{N_T} \sum_{k=1}^{N_T} \sum_{ij} |P_{ij}(t_k) - Q_{ij}(t_k)|^2 \quad (82)$$

Here,  $P_{ij}$  is the observed transition probabilities obtained from MD simulations or experiments. The diffusivity is then inferred from  $\mathbf{R}$  using eq 58. In ref 120, the authors directly apply their methodology to concentration profiles inferred from light absorption experiments to probe drug diffusion over the skin. One potential drawback of minimizing MSE rather than maximizing likelihood is that the latter is based on Kullback–Leibler (KL) divergence, which usually results in stronger gradients that lead to faster convergence.

## 7. BIAS-BASED METHODS

One intriguing approach for estimating diffusivity profiles in confined systems involves explicitly applying a suitable biasing potential to the system's Hamiltonian and utilizing the resulting mobility statistics to infer local diffusivity. At a fundamental level, these methods are exact as long as the underlying assumptions about the approximate mathematical form of the Hamiltonian are satisfied. However, introducing a

bias to the Hamiltonian will inherently alter the free energetics and dynamics of the system in nontrivial ways, potentially impacting the magnitudes and functional forms of transport properties. Furthermore, it typically requires multiple MD simulations (each with biasing potentials centered at different positions) to reconstruct the complete diffusivity profile. This stands in sharp contrast to the methods discussed in previous sections, which infer the full diffusivity profile from a single unbiased MD trajectory.

The very first method of this kind, and one of the earliest methods for estimating position-dependent diffusivity profiles, was introduced by Straub et al.<sup>121,122</sup> This approach is based on the generalized Langevin equation (GLE) given by

$$d\mathbf{X}_t = \mathbf{V}_t dt \\ m d\mathbf{V}_t = - \left[ \nabla \mathcal{F}(\mathbf{X}_t) + \int_0^t \boldsymbol{\zeta}(t-s) \cdot \mathbf{V}_s ds \right] dt + \mathbf{f}_t \quad (83)$$

Here,  $\mathcal{F}$  is the system's Hamiltonian,  $\boldsymbol{\zeta}(\cdot)$  is a memory kernel friction tensor, and  $\mathbf{f}_t$  is a zero-mean random force. In the one-dimensional case, assuming a harmonic PMF,  $\mathcal{F}(x) = \frac{1}{2}m\omega^2(x - x_0)^2$ , the dimensionless velocity autocorrelation function,

$$\bar{C}_v(t) = \frac{\langle \dot{x}(t)\dot{x}(0) \rangle}{\langle \dot{x}^2 \rangle}$$

will satisfy the following differential equation:

$$\frac{d\bar{C}_v}{dt} = - \int_0^t K(\tau) \bar{C}_v(t-\tau) d\tau \quad (84)$$

with

$$K(t) = \omega^2 + \frac{\zeta(t)}{m} \quad (85)$$

Typical decay characteristics of a friction memory kernel computed for a simple model system is depicted in Figure 8. In practice, the Hamiltonian  $\mathcal{F}$  is rarely harmonic in atomic and molecular systems. Therefore, a harmonic biasing potential with a sufficiently large angular velocity is added to the system's Hamiltonian to restraint a tracer particle at  $x_0$ . The time-dependent friction coefficient,  $\zeta(t)$  is then estimated from  $\bar{C}_v(t)$ . The friction coefficient at  $x_0$  is obtained from the following integral:

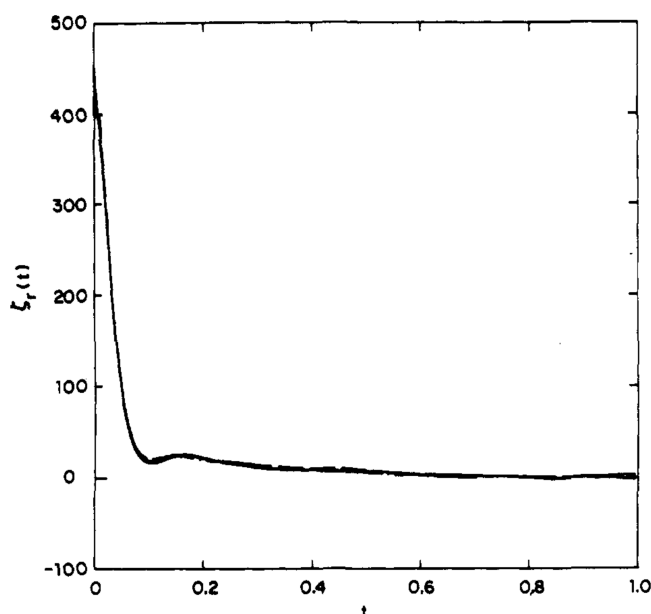
$$\gamma(x_0) = \frac{1}{m} \int_0^\infty \zeta(t) dt \quad (86)$$

The local diffusivity is subsequently related to the local friction coefficient using the Stokes–Einstein relationship:<sup>123</sup>

$$m\beta\gamma(x_0)D(x_0) = 1 \quad (87)$$

As stated above, it is crucial for the force constant of the biasing potential to be sufficiently large in order for the PMF to behave as a harmonic oscillator.<sup>124</sup>

A refined and elegant alternative to this method was introduced by Woolf and Roux.<sup>125</sup> By taking a Laplace transform from both sides of (85), they demonstrate that



**Figure 8.** Typical decay characteristics of the friction kernel of eq 83 computed for a model diatomic molecule within a solvent of LJ particles. (Reproduced with permission from ref 121. Copyright 1987, American Chemical Society.)

$$\begin{aligned} \mathcal{D}_\omega(s, x_0) &= \frac{\bar{C}_v(s) \langle (x - x_0)^2 \rangle \langle \dot{x}^2 \rangle}{\bar{C}_v(s) \left[ s \langle (x - x_0)^2 \rangle + \frac{\langle \dot{x}^2 \rangle}{s} \right] - \langle (x - x_0)^2 \rangle \langle \dot{x}^2 \rangle} \end{aligned} \quad (88)$$

Here,  $\bar{C}_v(s)$  and  $\mathcal{D}_\omega(s, x_0)$  correspond to the Laplace transforms of  $\bar{C}_v(t)$  and  $\mathcal{D}_\omega(t, x_0)$ , respectively, and  $\mathcal{D}_\omega(t, x_0)$  denotes the time-dependent diffusivity obtained by changing the upper limit of integration in (86) from  $\infty$  to  $t$ . (A more accessible step-by-step derivation of (88) is provided by Gaalswyk and Rowley.<sup>126</sup>) It is noteworthy that the infinity limit in (86) corresponds to the  $s \rightarrow 0^+$  limit in (88). In practice,  $\mathcal{D}_\omega(s, x_0)$  can be computed at multiple values of  $s$  and extrapolated to zero.

Finally, Hummer<sup>105</sup> demonstrated that the  $s \rightarrow 0^+$  limit in (88) can be analytically estimated, resulting in the following expression:

$$D_H = \lim_{s \rightarrow 0^+} \mathcal{D}_\omega(s, x_0) = \frac{\langle (x(t) - x_0)^2 \rangle}{\int_0^{+\infty} C_q(t) dt} \quad (89)$$

where  $C_q(t)$  is the position autocorrelation function (PACF)<sup>105,126</sup> defined as

$$C_q(t) \equiv \langle [x(t) - x_0][x(0) - x_0] \rangle \quad (90)$$

Note that eq 89 can be re-expressed as

$$D_H = \frac{\langle (x - x_0)^2 \rangle}{\tau} \quad (91)$$

with the time scale,  $\tau$ , given by

$$\tau = \frac{\int_0^{+\infty} C_q(t) dt}{\langle (x(t) - x_0)^2 \rangle}$$

Upon closer inspection, eq 91 is reminiscent of the concept of mean-squared displacement, specifically the estimators based on Kramers-Moyal coefficients given by (24). However, a notable limitation of these approaches is their assumption that diffusivity is a scalar (i.e., isotropic) position-dependent quantity, a condition almost never met in confined systems. Consequently, it is necessary to adapt these expressions to such circumstances. In Appendix A, we demonstrate the adaptability of this methodology to estimate anisotropic diffusivities that are axisymmetric. Due to its suitability for one-dimensional collective variables, this method has been applied in studies of protein folding to compute diffusivity along a reaction coordinate.<sup>115,124,127</sup>

Despite their differences, these methodologies fall under the broad category of *static restraint* (SR) methods, as per Holland et al.<sup>128</sup> since they all rely on restraining the position of a tracer particle at a fixed location using a harmonic spring. An alternative approach for estimating  $\zeta(t)$  in eq 83 can be devised by applying the fluctuation–dissipation theorem, demonstrating that<sup>68,129,130</sup>

$$\zeta(t) = \beta \langle \mathbf{f}(t) \mathbf{f}^T(0) \rangle \quad (92)$$

Here,  $\mathbf{f}(t) = \mathbf{F}(t) - \langle \mathbf{F}(t) \rangle_t$  is the residual force exerted on the tracer particle at time  $t$ . The diffusivity tensor can then be estimated as (refer to Appendix B)

$$\mathbf{D} = \beta^{-2} \left[ \int_0^{+\infty} \langle \mathbf{f}(\mathbf{X}_t) \mathbf{f}^T(\mathbf{X}_0) \rangle dt \right]^{-1} \quad (93)$$

Note that eq 93 is valid even without a restraining force as long as the diffusivity is constant. It can also be applied, akin to SR methods, to estimate local diffusivity when the restraining force is sufficiently strong to locally restrain the tracer particle. However, since the force autocorrelation function is independent of the tracer's temporal evolution, it can be computed even for a constrained tracer particle. This corresponds to taking the limit of (93) as  $\omega \rightarrow \infty$ . In such a scenario, the diffusivity at  $\mathbf{r}_0$  can be computed by pinning a particle at that position and computing the force autocorrelation function for that particle:<sup>68,131</sup>

$$\mathbf{D}(\mathbf{r}_0) = \beta^{-2} \left[ \int_0^{+\infty} \langle \mathbf{f}(\mathbf{r}_0, t) \mathbf{f}^T(\mathbf{r}_0, 0) \rangle dt \right]^{-1} \quad (94)$$

Eq 94 features a constraint rather than a restraining force. Therefore, following Holland et al.'s<sup>128</sup> terminology, the corresponding method can be labeled as a *static constraint* (SC) method.

The aforementioned approaches, whether based on pinning the particle to a fixed position or employing a restraining force, prove valuable in capturing diffusivity profiles within regions of the simulation box with low probabilities of being visited by certain solutes. Indeed, these approaches were historically developed for the study of membrane permeation, and have been extensively utilized in estimating the diffusivity of permeants across lipid membranes<sup>68,126,131–134</sup> and ion transport through pores.<sup>130</sup> A detailed review of such approaches is given by Shinoda.<sup>135</sup>

One can devise an alternative<sup>68</sup> method for estimating the friction coefficient (and diffusivity) starting with the SDE describing underdamped Langevin dynamics:

$$\begin{cases} d\mathbf{X}_t = \mathbf{V}_t dt \\ d\mathbf{V}_t = -(\nabla \tilde{\mathcal{F}}(\mathbf{X}_t) + \gamma(\mathbf{X}_t) \cdot \mathbf{V}_t) dt + \sqrt{2\tilde{\mathbf{D}}(\mathbf{X}_t)} d\mathbf{W}_t \end{cases} \quad (95)$$

Here,  $m$  is the mass of the particle,  $\tilde{\mathcal{F}}(\mathbf{r}) = m^{-1}\mathcal{F}(\mathbf{r})$  is the reduced PMF, and  $\tilde{\mathbf{D}}$  is the ‘velocity diffusivity’ (i.e., with units of velocity squared divided by time). Taking the expectation value of both sides of the velocity equation yields

$$m\langle d\mathbf{V}_t \rangle = [\langle \mathbf{F}_{\text{ext}} \rangle + m\gamma(\mathbf{r}) \cdot \langle \mathbf{V}_t \rangle] \quad (96)$$

where  $\langle \mathbf{F}_{\text{ext}} \rangle$  is the mean external force exerted on the tracer (due to both the PMF and possibly nonequilibrium external forces). Note that  $\langle \mathbf{F}_{\text{ext}} \rangle$  does not include any random forces, which average out to zero. Assuming that the external force can be kept sufficiently small for the change in velocity to be negligible, one can use the magnitude of the net force to compute the friction coefficient. In the simplest case of one-dimensional diffusivity, this will result in

$$D(z) \approx \frac{\langle V_t \rangle}{\beta \langle F_{\text{ext}} \rangle} \quad (97)$$

One way of implementing (97), originally proposed by Cicotti and Jacucci,<sup>136</sup> is to apply vanishingly small external forces such that the friction can overcome  $F_{\text{ext}}$ . However, it must be emphasized that  $F_{\text{ext}}$  also encompasses the effects of the PMF and is not solely comprised of the external force exerted during nonequilibrium MD. This may cause some issues in applying this methodology in the vicinity of hard boundaries.

The most interesting applications of this expression come from the works of McKinnon et al.<sup>137</sup> and Holland et al.<sup>128</sup> Both use steered molecular dynamics in which a harmonic potential is applied to the tracer particle with a moving minimum  $z(t)$ . In McKinnon et al.’s work,<sup>137</sup> this minimum moves with a constant velocity  $v_d$  and the instantaneous magnitude of the external force is calculated such that the left-hand side of (96) remains zero at all times. The forcing term can therefore be replaced by a term involving the work of the restraining force:

$$D[z(t)] = \frac{v_d}{\beta d\langle W_{\text{ext}} \rangle / dz} \approx v_d \frac{\langle l \rangle}{\beta \langle W_{\text{ext}} \rangle} \approx \frac{\langle l \rangle^2}{t\beta \langle W_{\text{ext}} \rangle} \quad (98)$$

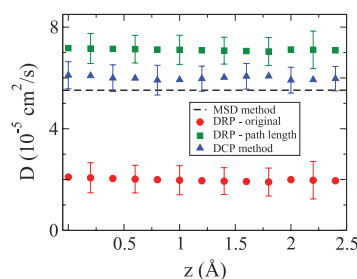
Using Holland et al.’s terminology,<sup>128</sup> this method can be categorized as a *dynamic restraints* (DR) method. Clearly, neglecting the effect of PMF will affect the validity of eq 98. Indeed, a more rigorous derivation, which explicitly accounts for the effect of the PMF and uses overdamped Langevin dynamics, is given by Park and Schulten.<sup>138</sup>

Holland et al.<sup>128</sup> highlight several practical issues in numerically implementing the method proposed by McKinnon et al.<sup>137</sup> and introduce a new method, which they call a *dynamic constraint* (DC) method, to address some of those issues. They consider the limiting case in which the tracer particles are constrained to follow a prescribed path  $z(t)$  rather than being subjected to a harmonic force centered at  $z(t)$ . Intuitively, this could be thought of as a special case of the latter approach in the limit of very stiff springs. The external force,  $F_{\text{ext}}(t)$ , needed to drive the particle along  $z(t)$  is then back-calculated by estimating the force exerted on the tracer by the environment. The work from this force is then subtracted from the change in PMF and employed in (98).

They also remark on proper choices of  $\langle l \rangle$  and  $v_d$  in eq 98 when the DR approach is employed. They argue that the observation window  $t$  needs to be partitioned into smaller windows of duration  $\Delta t \ll t_s$  where  $t_s$  is the characteristic oscillation period of the harmonic spring. They then argue that  $\langle l \rangle$  and  $v_d$  need to be chosen as

$$\langle l \rangle = \sum_{i=1}^{n_w} \langle |\Delta z_i| \rangle, \quad v_d = \frac{1}{t} \left| \sum_{i=1}^{n_w} \langle \Delta z_i \rangle \right|$$

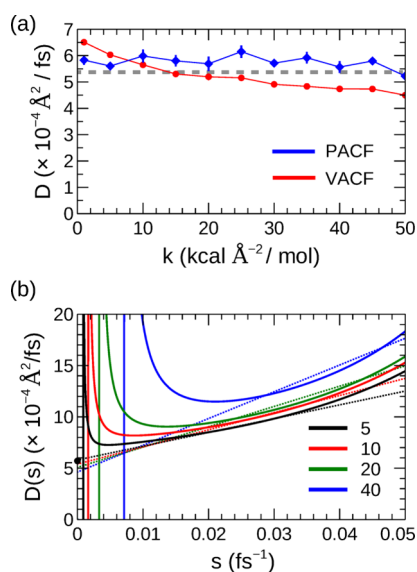
In other words,  $\langle l \rangle$  needs to be chosen as the average arc length of tracer paths, while  $v_d$  should be estimated from the net displacement. Failing to do so will result in a systematic underestimation of diffusivity, as shown in Figure 9. In contrast, the DC approach, in which the particle follows a prescribed path, is not impacted by such uncertainties and provides more accurate estimates of diffusivity.



**Figure 9.** Effect of the particular choice of  $\langle l \rangle$  on the performance of the DR method described by eq 98. Using the path length results in better agreement with the true value of diffusivity (estimated from MSD). Overall, the DC method exhibits a superior performance in comparison with different implementations of the DR method. (Reprinted with permission from ref 128. Copyright 2012, American Physical Society.)

Since restraining force-based methods have been in use for a long time, known numerical issues with their implementation have been reported. For instance, it is known that Langevin-based thermostats, such as the Berendsen thermostat,<sup>139</sup> introduce extra unphysical friction despite correctly sampling the equilibrium Boltzmann distribution. This additional friction results in a systematic underestimation of diffusivity,<sup>126</sup> which is absent from deterministic thermostats, such as the Nosé–Hoover thermostat.<sup>140,141</sup> Another obvious issue is that using stiff springs might necessitate employing a smaller MD time step.

An interesting implementation question is the sensitivity of the estimated friction coefficient to the particular choice of the spring constant. Gallschwyk et al.<sup>126</sup> investigated this issue, comparing the VACF-based method of Woolf and Roux,<sup>125</sup> with its reformulation by Hummer,<sup>105</sup> given by eqs 88 and 89, respectively. They observed that for moderate values of  $k$ , the VACF-based approach of eq 88 is more sensitive to  $k$ . They conjecture that this increased sensitivity arises from the uncertainties in the numerical extrapolation of (88) to  $s \rightarrow 0^+$  for larger values of  $k$  (Figure 10). Despite the reduced sensitivity of the PACF-based method of Hummer to  $k$ , it can suffer from slow decay of PACF in some circumstances.<sup>126</sup> In bulk liquids, a few picoseconds might be sufficient for the position autocorrelation function to decay to zero at room temperature. In contrast, a solute restrained deep inside a lipid bilayer might exhibit long-lived oscillations that last up to

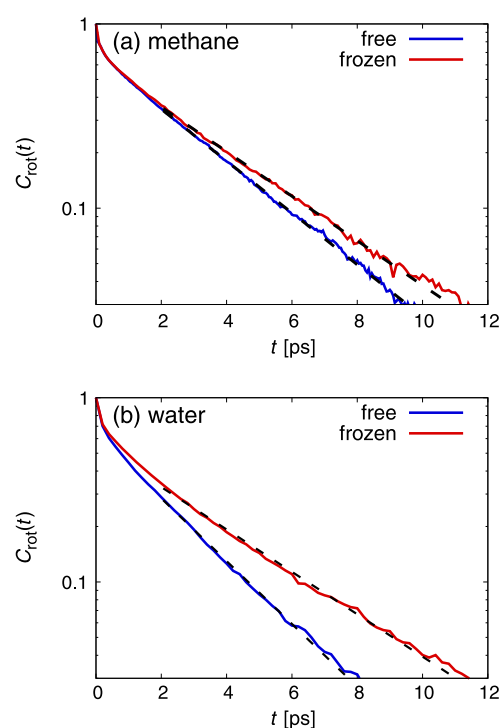


**Figure 10.** Effect of spring constant on (a) estimates of diffusivity obtained using eqs 88 and 89 and (b) the qualitative behavior of  $D_\omega(s)$  in eq 88. (Reproduced with permission from ref 126. Copyright 2016, American Chemical Society.)

hundreds of picoseconds, resulting in slow decay of PACF. This observation is corroborated by Daldrop and Netz,<sup>142</sup> who demonstrate a 100 ps decay for the PACF of a methane molecule restrained within water.

Daldrop and Netz<sup>142</sup> also analyze the effect of larger spring constants on the performance of Hummer's PACF-based approach, something not considered by Gaalswyk et al.<sup>126</sup> They demonstrate that while larger  $k$ 's lead to faster decay of PACF, they might make it highly oscillatory, which could also make its accurate numerical integration more challenging. Moreover, they observe that larger spring constants lead to a systematic overestimation of the friction coefficient, and hence an underestimation of diffusivity. For instance, they report that the friction coefficient of a methane solute within SPC/E water<sup>143</sup> increases by 60% in the limit of  $k \rightarrow \infty$ . They attribute the source of this systematic error to the alteration of the dynamics of the first hydration shell by a frozen molecule. For instance, they compute an orientational autocorrelation function for the water molecules in the first hydration shell and find it to decay more slowly when the solute is frozen (Figure 11). This discrepancy becomes larger when stronger interactions are present between the solute and the solvent. Similarly, they observe an increase in the mean escape time of water molecules around the solute. In other words, solvent molecules tend to be less mobile around a pinned particle. This, once again, underscores the fundamental issue with bias-based methods in general that can alter the dynamics of the underlying system in nontrivial ways, and highlights the merits of utilizing alternative approaches, e.g., based on path sampling techniques,<sup>144</sup> to study membrane permeation.

Various other details in the implementation of bias-based methods have been examined in the literature. For instance, Fujimoto et al.<sup>145</sup> examined the impact of finite size effects, especially in the context of long-range electrostatic interactions. Additionally, Holland et al.<sup>128</sup> illustrated that increasing the speed of the moving spring might adversely impact the accuracy of diffusivity estimates. These findings collectively emphasize the importance of exercising careful



**Figure 11.** Sensitivity of the orientational autocorrelation functions of water molecules within the first hydration shell of a fixed and a freely moving (a) methane and (b) water molecules. (Reproduced with permission from ref 142. Published 2017 by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license.)

consideration when selecting parameters for the implementation of bias-based methods.

## 8. COLLECTIVE VARIABLE-BASED APPROACHES

In this section, we will discuss methodologies originally developed in the context of collective variables, i.e., mechanical observables carefully chosen to accurately represent the free energy landscapes of physical and biological systems, particularly with the aim of characterizing the kinetics and mechanisms of rare events.<sup>146,147</sup> These variables can be formulated through various approaches, including physical intuition, experimental insights, or data science methodologies such as principal component analysis,<sup>148,149</sup> diffusion maps,<sup>150,151</sup> and machine learning.<sup>152–155</sup> From a mathematical standpoint, CVs can be seen as embeddings or projections that map the high-dimensional configuration space onto a lower-dimensional space. The projection formalism introduced by Mori and Zwanzig<sup>156,157</sup> demonstrates that the temporal evolution of a collective variable can be modeled as a stochastic process with memory, even in cases where the underlying equations of motion are Hamiltonian. In situations with negligible memory effects, or just as a first approximation, a Smoluchowski-type equation is postulated within the CV space:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot [\mathbf{D}(\lambda) \cdot [\nabla_\lambda \rho + \beta \rho \nabla_\lambda F(\lambda)]] = \mathcal{L}^\dagger \rho \quad (99)$$

Here,  $\lambda$  represents a vectorial CV that exhibits diffusive behavior, and  $F(\lambda)$  denotes the Landau free energy profile<sup>158</sup> with respect to  $\lambda$ .

Assuming the validity of this description, algorithms can be developed to estimate  $\mathbf{D}(\lambda)$  from the CV time series,  $\lambda_t$ . Indeed, several of the algorithms<sup>82,105,107,115,125,127</sup> discussed in prior sections have also been employed for characterizing diffusivity variations within a CV space. In this section, we want to focus on two classes of conceptually distinct methodologies, both historically rooted in computational chemistry and biology, which are based on estimating the mean first passage time (MFPT) and committor analysis, respectively. We discuss both approaches within the framework of transition path theory<sup>159</sup> (TPT), specifically focusing on transitions between  $A$  and  $B$ , two open sets within  $\mathbb{R}^n$ , which are also (meta)stable basins of attractions within the free energy landscape.

**8.1. MFPT-Based Approaches.** Consider an open set  $C \subset \mathbb{R}^n$ . The first passage time to leave  $C$  (while starting at  $\lambda_0 \in C$ ) is defined as

$$\tau_{\text{fp}}(\text{Cl}\lambda_0) = \inf_t \{t > 0: \lambda_t \notin C\}$$

First passage times can be estimated by imposing absorbing boundary conditions at  $C$ 's boundary, i.e.,

$$\rho(\lambda, t) = 0, \quad \forall \lambda^* \in C^* \subset \partial C \quad (100)$$

Note that the  $\rho(\lambda, t)$  that satisfies (100) will not be properly normalized, and will instead represent a joint probability density, i.e., the probability that the particle is found at  $\lambda$  at time  $t$  and that it has not crossed the absorbing boundary  $C^*$  up to time  $t$ . Within the Smoluchowski framework,  $\tau_{\text{fp}}(\text{Cl}\lambda_0)$  can be calculated as<sup>91</sup>

$$\begin{aligned} \tau_{\text{fp}}(\text{Cl}\lambda_0) &= \int_0^\infty -t \frac{\partial}{\partial t} \left[ \int_C \rho(\lambda, t|\lambda_0) d\lambda \right] dt \\ &= \int_C \left[ \int_0^\infty \rho(\lambda, t|\lambda_0) dt \right] d\lambda \equiv \int_C G(\lambda|\lambda_0) d\lambda \end{aligned} \quad (101)$$

The second equality is derived through the inversion of the order of integration and the application of integration by parts. The function  $G(\lambda|\lambda_0) = \int_0^\infty \rho(\lambda, t|\lambda_0) dt$  represents a Green's function that satisfies the following PDE:

$$\begin{aligned} \nabla \cdot [\mathbf{D}(\lambda) \cdot [\nabla_\lambda G(\lambda|\lambda_0) + \beta G(\lambda|\lambda_0) \nabla_\lambda F(\lambda)]] \\ = \rho_0(\lambda) - \delta(\lambda - \lambda_0) \end{aligned} \quad (102)$$

Eqs 101 and 102 are valid for any open set  $C$  within a multidimensional CV space. Following Weiss,<sup>91</sup> it is possible to derive an analytical expression for the first passage time for a scalar (i.e., one-dimensional) CV. First, one can define:

$$u(\lambda, t) = \int_C \rho(\mathbf{y}, t|\lambda) d\mathbf{y} \quad (103)$$

Note that  $u(\lambda, t)$  is the probability that  $\tau_{\text{fp}}(\text{Cl}\lambda) \geq t$  and satisfies the adjoint evolution equation given by

$$\frac{\partial u}{\partial t} = e^{\beta F(\lambda)} \nabla_\lambda \cdot [e^{-\beta F(\lambda)} \mathbf{D}(\lambda) \cdot \nabla_\lambda u] = \mathcal{L}u \quad (104)$$

where  $\mathcal{L}$  is the adjoint of the operator on the right-hand side of (99). As discussed previously,  $\tau_{\text{fp}}(\text{Cl}\lambda)$  can be calculated as

$$\tau_{\text{fp}}(\text{Cl}\lambda) = - \int_0^\infty t \frac{\partial u}{\partial t} dt$$

It can be demonstrated that  $\tau_{\text{fp}}(\text{Cl}\lambda)$  will satisfy the following PDE:

$$e^{\beta F(\lambda)} \nabla_\lambda \cdot [e^{-\beta F(\lambda)} \mathbf{D}(\lambda) \cdot \nabla_\lambda \tau_{\text{fp}}(\text{Cl}\lambda)] = -1 \quad (105)$$

All these assertions are valid for any CV space, irrespective of its dimensionality. If the CV is scalar, however, (105) will turn into an ordinary differential equation (ODE), which can be solved analytically for an arbitrary open interval,  $C = (a, b)$  to yield:

$$\begin{aligned} \tau_{\text{fp}}(\text{Cl}\lambda) &= C_0 + \\ &C_1 \int_a^\lambda \frac{1}{\rho_0(y)D(y)} dy - \int_a^\lambda \frac{1}{\rho_0(y)D(y)} \left[ \int_a^y \rho_0(z) dz \right] dy \end{aligned} \quad (106)$$

Here,  $\rho_0(\lambda) = e^{-\beta F(\lambda)} / \int_a^b d\bar{\lambda} e^{-\beta F(\bar{\lambda})}$  represents the conditional equilibrium distribution of  $\lambda$  within the interval  $(a, b)$ . The constants  $C_0$  and  $C_1$  depend on the specific boundary conditions imposed at  $a$  and  $b$ .

We now consider the solutions of eq 106 when reflective boundary conditions are imposed at  $a$  and  $b$ . This scenario is widely explored in the literature, representing the case where  $a$  and  $b$  correspond to two (local) minima in the free energy landscape, indicative of transitions between two (meta)stable basins. Additionally, we introduce an intermediate point  $\lambda_f \in (a, b)$  where an absorbing boundary condition is applied. Consequently, trajectories can originate from two distinct sets:  $\lambda \in [a, \lambda_f]$  or  $\lambda \in [\lambda_f, b]$ , each featuring a reflective and an absorbing boundary condition. The mean first passage time of reaching  $\lambda_f$  can be readily estimated as<sup>82</sup>

$$\tau_{\text{fp}}(\lambda_f|\lambda) = \begin{cases} \int_{\lambda_f}^\lambda \frac{d\xi e^{\beta F(\xi)}}{D(\xi)} \int_\xi^b d\zeta e^{-\beta F(\zeta)} & \lambda \in [\lambda_f, b] \\ \int_\lambda^{\lambda_f} \frac{d\xi e^{\beta F(\xi)}}{D(\xi)} \int_a^\xi d\zeta e^{-\beta F(\zeta)} & \lambda \in [a, \lambda_f] \end{cases} \quad (107)$$

Note that the system can only exit the starting set at  $\lambda_f$  as reflective boundary conditions are applied at the other end. This is why we denote the first passage time with  $\tau_{\text{fp}}(\lambda_f|\lambda)$  instead of  $\tau_{\text{fp}}(\text{Cl}\lambda)$ . Eq 107 can be subsequently used to derive the following expressions for  $D(\lambda)$  in terms of the derivative of  $\tau$  with respect to  $\lambda_0$ :

$$D(\lambda) = \begin{cases} \frac{e^{\beta F(\lambda)}}{\partial \tau_{\text{fp}} / \partial \lambda} \int_\lambda^b e^{-\beta F(\xi)} d\xi & \lambda > \lambda_f \\ -\frac{e^{\beta F(\lambda)}}{\partial \tau_{\text{fp}} / \partial \lambda} \int_a^\lambda e^{-\beta F(\xi)} d\xi & \lambda < \lambda_f \end{cases} \quad (108)$$

Alternative, one can define the notion of a round-trip time as<sup>82</sup>

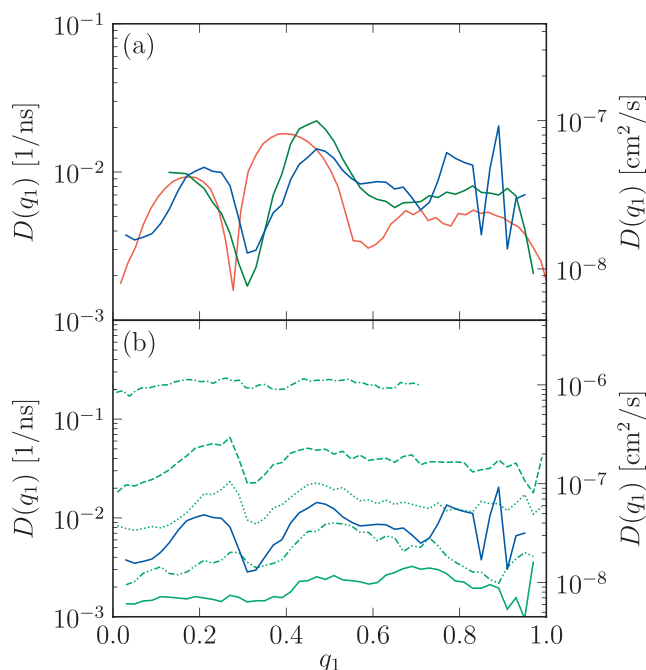
$$\begin{aligned} \tau_{\text{rt}}(\lambda, \lambda_f) &= \text{sign}(\lambda - \lambda_f) [\tau_{\text{fp}}(\lambda, \lambda_f) + \tau_{\text{fp}}(\lambda_f, \lambda)] \\ &= \int_{\lambda_f}^\lambda d\xi \frac{e^{\beta F(\xi)}}{D(\xi)} \int_a^b d\zeta e^{-\beta F(\zeta)} \end{aligned} \quad (109)$$

By differentiating eq 109 with respect to  $\lambda$ , it is easy to demonstrate that

$$D(\lambda) = \frac{1}{\rho_0(\lambda) \partial \tau_{\text{rt}} / \partial \lambda} \quad (110)$$

Here, we adopt the notation employed by Hinczewski et al.<sup>82</sup> who use eqs 108 and 110 to compute diffusivity along different

CVs associated with folding of a helix-forming short peptide into an  $\alpha$ -helix (Figure 12). However, similar expressions for



**Figure 12.** Computing diffusivity along  $q_1$ , the root-mean-square deviation from a perfect helix for an  $\alpha$ -helix forming short peptide. (a)  $D(q_1)$  computed from eq 108 (green curve), eq 110 (blue curve) and using Hummer's Bayesian approach<sup>105</sup> (red curve). (b)  $D(q_1)$  estimates from eq 110 (blue curve) as well as using eq 30 using different lag times (green curves). (Reprinted from ref 82, with the permission of AIP Publishing.)

the case of  $\lambda \in [a, \lambda_f]$  have also been utilized by Chahine et al.,<sup>124</sup> Sedlmeier et al.<sup>80</sup> and Bollinger and Truskett.<sup>160</sup> It is crucial to underscore that these analytical expressions are exact when applied to one-dimensional CV spaces. Consequently, they have found widespread applications in characterizing transitions within biomolecular systems that can be adequately represented by a scalar CV. Both Hinczewski et al.<sup>82</sup> and Chahine et al.<sup>124</sup> employed these expressions for probing protein folding, while Sedlmeier et al.<sup>80</sup> and Bollinger and Truskett<sup>160</sup> applied them to estimate actual diffusivity profiles within an axisymmetric geometry—specifically for SPC/E water<sup>143</sup> confined within a slit pore and the hard sphere fluid exposed to a one-dimensional sinusoidal potential, respectively. In both cases, diffusivity was solely a function of  $z$ , and the methodology was applied to estimate  $D_{zz}$ , which is justified due to the axisymmetric geometry of the system (see Appendix A).

It is important to acknowledge that the numerical implementation of this approach may pose challenges, even in cases that are truly one-dimensional. This challenge arises from the necessity to estimate the derivatives of either the mean first passage time or the round-trip time with respect to  $\lambda$ . Importantly, these derivatives appear in the denominators of (108) and (110). As such, substantial instabilities might arise if such derivatives are small in magnitude. Alternatively, one can introduce appropriate approximations to estimate these derivatives, thereby mitigating the reliance on numerical differentiation.

One such approximate approach was proposed by Belousov et al.,<sup>161</sup> who consider a sufficiently narrow interval containing

$\lambda_0$ , namely  $C = [\lambda_0 - l, \lambda_0 + l]$ . By applying absorbing boundary conditions at both ends of  $C$ , (106) will yield:

$$\tau_{\text{fp}}(C|\lambda_0) = \int_{\lambda_0-l}^{\lambda_0} [\bar{M} - M(y)] \frac{dy}{D(y)\rho_0(y)}$$

where

$$\begin{aligned} M(y) &= \int_{\lambda_0-l}^y \rho_0(z) dz \\ \bar{M} &= \frac{\int_{\lambda_0-l}^{\lambda_0+l} \frac{M(y) dy}{D(y)\rho_0(y)}}{\int_{\lambda_0-l}^{\lambda_0+l} \frac{dy}{D(y)\rho_0(y)}} \end{aligned} \quad (111)$$

and  $\rho_0(\lambda) \propto e^{-\beta F(\lambda)}$ . The approximation works by obtaining the Taylor expansion of  $F(\lambda)$  and log diffusivity around the fixed point  $\lambda_0$ :

$$F(\lambda) = F(\lambda_0) + F'(\lambda_0)(\lambda - \lambda_0) + O(|\lambda - \lambda_0|^2)$$

$$\frac{D(\lambda)}{D(\lambda_0)} = \exp \left[ \frac{D'(\lambda_0)}{D(\lambda_0)} (\lambda - \lambda_0) + O(|\lambda - \lambda_0|^2) \right]$$

which can then be used to estimate mean first passage and roundtrip times over short displacements within the CV space. For instance, the MFPT of starting at  $\lambda_0$  and reaching  $\lambda_0 \pm l$  will be given by<sup>161</sup>

$$\tau_{\text{fp}}(C|\lambda_0) \approx \frac{1}{\beta \kappa F'(\lambda_0) D(\lambda_0)} \left[ \frac{\cosh\{\beta F'(\lambda_0) + \kappa\} l/2}{\cosh\{\beta F'(\lambda_0) - \kappa\} l/2} - 1 \right] \quad (112)$$

Here,  $\kappa = D'(\lambda_0)/D(\lambda_0)$ . If  $\kappa \ll 1$ , eq 112 can be further simplified to yield:

$$\tau_{\text{fp}}(C|\lambda_0) \approx \frac{l \tanh[\beta F'(\lambda_0)/2]}{\beta D(\lambda_0) F'(\lambda_0)} \quad (113)$$

Therefore,  $D(\lambda)$  can be directly estimated from the first derivative of  $F(\lambda)$ , i.e., the mean force exerted along the scalar CV, and the mean first passage time of reaching  $\lambda \pm l$ . Further approximations along the same lines can be made, particularly if the Smoluchowski equation is replaced by a related Fokker–Planck equation.<sup>162</sup>

**8.2. Committor-Based Methods.** Here, we discuss a method<sup>163</sup> proposed by Berezhkovskii and Makarov that estimates position-dependent diffusivity from committor analysis. Assuming the existence of two basins of attractions,  $A$  and  $B$ , within the free energy landscape,  $q_B(x)$ , the committor probability of reaching  $B$ , is defined as

$$q_B(\lambda) = \mathbb{P}(T_B < T_A)$$

where  $T_C$  is the first passage time of reaching set  $C$ :

$$T_C = \inf_t \{t > 0: \lambda_t \in C\}$$

Assuming that the evolution of  $\lambda_t$  within the CV space follows (99), it can be demonstrated that  $q_B(\lambda)$  will satisfy the following elliptic PDE:<sup>164</sup>

$$\begin{aligned}\nabla \cdot [\mathbf{D}(\lambda) e^{-\beta F(\lambda)} \cdot [\nabla_{\lambda} q_B]] &= 0 \quad \lambda \notin A \cup B \\ q_B &\equiv 1 \quad \lambda \in \partial B \\ q_B &\equiv 0 \quad \lambda \in \partial A\end{aligned}\quad (114)$$

It is necessary to note that (114) is only valid when  $\lambda \notin A \cup B$ , otherwise  $q_B(\lambda) = 0$  and 1 for  $\lambda \in \partial A$  and  $\lambda \in \partial B$ , respectively. In the case of a scalar CV, one can obtain an analytical expression for  $q_B(\lambda)$ . More precisely, for  $A = (-\infty, a)$  and  $B = (b, \infty)$ , it can be demonstrated that

$$q_B(\lambda) = \frac{\int_a^{\lambda} \frac{e^{\beta F(s)}}{D(s)} ds}{\int_a^b \frac{e^{\beta F(s)}}{D(s)} ds} \quad (115)$$

Upon differentiation and rearrangement, (115) can be re-expressed as<sup>163</sup>

$$D(\lambda) = \frac{\int_a^b q_B(z)(1 - q_B(z))\rho_0(z) dz}{\rho_0(\lambda) q_B'(\lambda) \langle \tau_{tr}(a, b) \rangle} \quad (116)$$

Here,  $\tau_{tr}(a, b)$ , the *transition time*, is the earliest time that a trajectory initiated at  $a$  or  $b$  leaves the interval  $(a, b)$ . Note that the transition time is smaller than the first passage time. For a one-dimensional CV space, the mean transition time is given by

$$\langle \tau_{tr}(a, b) \rangle = \int_a^b \frac{e^{\beta F(s)}}{D(s)} ds \int_a^b q_B(u)[1 - q_B(u)]e^{-\beta F(u)} du \quad (117)$$

Note that the only derivative appearing in eq 116 is  $q_B'(\lambda)$ , which can be evaluated using recent algorithms for the parametrization of the committor probabilities, e.g., through the application of neural networks.<sup>165</sup> While eq 116 can, in principle, be used to estimate  $D(\lambda)$ , we are not aware of any instance of its application in the literature, possibly due to numerical instabilities emerging from the exceedingly small values of  $q_B'(\lambda)$  within regions that are too far from the transition state.

## 9. CONCLUSIONS

In this work, we provide a brief overview of computational attempts to estimate position-dependent diffusivity tensors (and other related transport coefficients) from MD trajectories of confined systems. At a fundamental level, these methodologies involve solving the inverse problem of inferring diffusivity profiles within the Smoluchowski framework from the observed mobility statistics of individual particles. We classify these attempts based on their underlying theoretical foundations. In addition to *ad hoc* extensions (Section 3) of rigorous algorithms developed for bulk simulations (Section 2), we discuss more rigorous methodologies, such as kernel-based methods (Section 4), Bayesian approaches (Section 5), operation discretization methods (Section 6), and bias-based methods (Section 7). We also discuss the related problem of estimating diffusivity profiles in collective variable spaces through estimating mean first passage times and committor probabilities (Section 8).

In order to effectively implement these algorithms, it is imperative to make certain design choices. First and foremost, most such methods rely on partitioning the simulation domain

into discrete bins, necessitating the specification of the size of each bin. On one hand, it would be preferable for the bins to be as small as possible to resolve subtle spatial variations of the diffusivity tensor. On the other hand, utilizing bins that are too narrow could compromise the statistical precision of these estimates due to insufficient sampling within each bin. Therefore, choosing the optimal bin size can be viewed as an optimization problem, balancing a trade-off between resolution and precision. It is crucial to note that the optimal bin size could vary across different algorithms, even if applied to the same system. For instance, the MFPT-based methods of Section 8 are more susceptible to binning due to their reliance on determining the first derivative of MFPT, which can be fairly noisy if the bins are too small.

In practice, bin sizes are usually determined through a combination of physical intuition and sensitivity analysis (i.e., observing the qualitative impact of altering the bin width on the diffusivity estimate). The latter approach can be made more systematic by splitting the MD trajectories into smaller subsets and choosing the smallest bin size for which the variance among estimates obtained from different subsets remains smaller than a prespecified threshold. For some methods, more rigorous frameworks can be utilized. For instance, in kernel-based methods of Section 4, one could use the leave-one-out cross validation (LOOCV) framework<sup>166</sup> to formulate an optimization problem for choosing  $\epsilon$ , the thickness of the localization kernel. In methods based on Markov state models, approaches such as time-lagged independent component analysis<sup>167</sup> (tICA) can be used for approximating the eigenfunctions of  $e^{tL_r}$  and use the second eigenfunction, corresponding to the PMF, to devise suitable binning schemes.<sup>119</sup>

For methods that rely on projecting diffusivity profiles onto a finite set of orthogonal basis functions, such as refs 92 and 113, no binning is necessary. Instead, the critical design parameter is the number of basis functions utilized in the projection, which is typically determined by progressively including more basis functions until the representation of the diffusivity profile ceases to change meaningfully. Since many classical dictionaries of basis functions (such as Fourier series) become increasingly oscillatory, the optimal number of basis functions is connected to the decay characteristics of the diffusivity profile in the Fourier domain, which are often *a priori* unknown.

While we primarily focus on methods developed and utilized in the context of molecular simulations, we wish to note that the Smoluchowski equation can be equivalently expressed as the forward Kolmogorov equation associated with a stochastic process. The task of deducing the PMF and position-dependent diffusivity from observed mobility statistics can thus be perceived as an inference problem in stochastic processes, which has a rich history within the statistics community and remains an active area of research.<sup>168–170</sup> Our kernel-based method, as elaborated in our previous works,<sup>92,93</sup> and detailed in Section 4, represents an endeavor to adapt successful estimators from the statistics community to the problem of determining transport coefficients in molecular simulations. We contend that numerous unexplored opportunities exist in this realm. An interesting example is the study of electromagnetic wave propagation in highly scattering media, a process that is described by a PDE very similar to the Smoluchowski equation. Consequently, methodologies in the optics community have been developed to estimate

diffusion tensors,<sup>171</sup> and there exists potential in adapting such approaches to molecular dynamics trajectories.

On a broader—but related—note, the question of inferring a term (or parameter) within a PDE from observations of its solutions—generally referred to *inverse problem*—is at the heart of applied mathematics. Indeed, the applied mathematics literature is replete with many more strategies for discretizing the Smoluchowski operator, often validated using synthetic data, as in the work of Crommelin and Vanden-Eijnden.<sup>172</sup> Adapting such strategies to be applicable to MD data could provide further opportunities to develop effective operator discretization algorithms for diffusivity estimation, beyond those already considered in this review.

In addition to developing new diffusivity estimators, it is also imperative to establish standard benchmarks for validating and assessing the performance of—new and existing—estimators. Establishing such benchmarks is essential due to the challenges in ensuring the numerical accuracy and stability of the employed algorithms, even when adhering to best practices in numerical analysis. Traditionally, the validation of new estimators has relied on their ability to accurately recover diffusivity profiles used for generating synthetic stochastic trajectories or to precisely estimate diffusivity within bulk systems. We contend that such traditional benchmarks are necessary, but not sufficient, for guaranteeing the robustness and reliability of an estimator, and more rigorous validation criteria are necessary. One such approach, as demonstrated in our previous work,<sup>93</sup> involves feeding the predictions of the estimator into a Langevin-based SDE, and comparing the mobility statistics of the arising stochastic trajectories, with the van Hove correlation functions obtained from MD simulations. This will not only enable one to evaluate the estimator's capability to generate internally consistent diffusivity profiles, but will also make it feasible to identify and flag deviations from the Smoluchowski picture within specific systems.

It is essential to acknowledge that the Smoluchowski equation serves as an approximation for describing particle mobility over extended time scales and may not be universally applicable to all molecular systems. A key assumption within the Smoluchowski formalism pertains to the Gaussian nature and the absence of temporal correlation for the random force acting on each particle. This assumption can be relaxed by extending the GLE formalism of eq 83, or through the incorporation of colored noise. Such extensions accommodate scenarios where the dynamics of a typical particle is non-Markovian. The GLE framework also exhibits improved agreement with VACFs computed from MD.<sup>122</sup> Notably, there have been endeavors<sup>173,174</sup> within the molecular simulations community, including machine learning approaches, to fit data from MD simulations to a GLE framework without explicitly aiming to estimate diffusivity. Adapting such methodologies for the estimation of position-dependent diffusivity could be a promising avenue for future investigations.

An interesting category of systems and processes, not addressed in this discussion, include those exhibiting anomalous diffusion.<sup>175–177</sup> One notable example is systems comprised of interacting Brownian particles governed by the McKean-Vlasov equation.<sup>178</sup> Unlike the standard and generalized Langevin formalisms, which both rely on a 'mean field' treatment of a typical particle within a bath, the McKean-Vlasov equation makes it possible to account for multiparticle effects. The estimation of transport coefficients within such

generalized frameworks poses an intriguing question that has not been addressed in this review. It must be noted that such complicated scenarios, such as those involving anomalous diffusion or systems described by the McKean-Vlasov equation, can still be tackled by modifying some of the methodologies discussed earlier. One viable option is to employ Bayesian or operator discretization approaches, known for their adaptability to diverse settings. In cases where the PDE governing the spatiotemporal evolution of probabilities is associated with a stochastic process, such as the McKean-Vlasov equation, kernel-based methods offer a valuable avenue. Depending on the specific characteristics of the underlying stochastic process, adapting autocorrelation-based techniques is also conceivable, although their generalization might prove more challenging.

We also do not discuss the problem of estimating transport properties other than diffusivity. Given the mathematical similarity between mass and charge transport, we expect some of the techniques described here to be applicable to the estimation of position-dependent and anisotropic electrical conductivity. Indeed, methodologies such as the one proposed by Mangaud and Rotenberg<sup>88</sup> have been employed by Helms et al.<sup>179</sup> to estimate the response matrix  $\mathcal{M}(z)$  in the presence of an external electrical potential. It is, however, far more challenging to treat momentum and heat transfer in a similar fashion. Although frameworks akin to the one discussed in ref 88 can be utilized alongside closed-form solutions of macroscopic fluid mechanics problems to estimate quantities such as viscosity, their applicability to more intricate geometries remains uncertain.

We contend that more systematic approaches for estimating transport coefficients, such as viscosity, can be formulated by positing that  $f(t, \mathbf{r}, \mathbf{v})$ , the probability density of a particle being at  $\mathbf{r}$  and having a velocity  $\mathbf{v}$  will satisfy the following kinetic ansatz:

$$\frac{\partial}{\partial t}f(t, \mathbf{r}, \mathbf{v}) + \mathbf{v} \cdot \nabla_{\mathbf{r}}f + \nabla_{\mathbf{r}} \tilde{U} \cdot \nabla_{\mathbf{v}}f = O[f] \quad (118)$$

Here, the operator  $O[f]$  encapsulates the effective interactions among particles in the system. For instance, in the case of the Kramers-Klein equation,<sup>75</sup>  $O[\cdot]$  is given by

$$O[f] = \nabla_{\mathbf{v}} \cdot \left( \frac{1}{m\beta} \mathbf{D}^{-1}(\mathbf{x}) \cdot \mathbf{v} f \right) + \frac{1}{(m\beta)^2} \mathbf{D}^{-1}(\mathbf{x}) : \mathbf{H}_{\mathbf{v}} f$$

In the context of a master equation framework,  $O[\cdot]$  can have the following structure:

$$O[f] = \int [W(\mathbf{r}', \mathbf{r}, \mathbf{v}', \mathbf{v})f(t, \mathbf{r}', \mathbf{v}') - W(\mathbf{r}, \mathbf{r}', \mathbf{v}, \mathbf{v}')f(t, \mathbf{r}, \mathbf{v})] d\mathbf{r}' d\mathbf{v}' \quad (119)$$

Multiplying (118) by  $\mathbf{v}$  and integrating over velocity space allows the derivation of a momentum balance equation. A similar approach can be employed to derive an energy conservation equation. Data-driven techniques can then be applied to fit the statistics obtained from MD simulations to such a kinetic description. Transport coefficients would be implicit to the choice of the operator, and could potentially be extracted from it. These concepts present avenues for future exploration.

## ■ APPENDIX A: AXISYMMETRIC SOLUTION TO THE SMOLUCHOWSKI EQUATION

Here, we consider a scenario in which both the diffusivity and PMF are functions of  $z$  only, and that the diffusivity tensor is axisymmetric, i.e.,  $D_{xx}(z) \equiv D_{yy}(z) \equiv D_{rr}(z)$ . This will imply translational invariance within the  $xy$  plane, a condition met in systems wherein a fluid is sandwiched between chemically uniform parallel plates. Under these conditions, the Smoluchowski equation takes the following form:

$$\frac{\partial \rho}{\partial t} = D_{rr}(z) \left[ \frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} \right] + \frac{\partial}{\partial z} \left[ D_{zz}(z) \rho_0 \frac{\partial}{\partial z} \left( \frac{\rho}{\rho_0} \right) \right] \quad (\text{A1})$$

with the operator acting on  $D_{zz}(z)$  denoted as

$$\mathcal{L}_z^\dagger f \equiv \frac{\partial}{\partial z} \left[ D_{zz}(z) \rho_0(z) \frac{\partial}{\partial z} \left( \frac{f}{\rho_0} \right) \right]$$

which is the differential operator on the right-hand side of the one-dimensional Smoluchowski equation. It can be shown that

$$\begin{aligned} \frac{\partial \hat{\rho}}{\partial t} &= -4\pi^2 D_{rr}(z) (k_x^2 + k_y^2) \hat{\rho} + \mathcal{L}_z^\dagger \hat{\rho} \\ \hat{\rho}(k_x, k_y, z, 0) &= \delta(z - z_0) \end{aligned} \quad (\text{A2})$$

with  $\hat{\rho}(k_x, k_y, z, t)$ , the partial Fourier transform of  $\rho(\cdot)$  defined as

$$\hat{\rho}(k_x, k_y, z, t) = \int_{\mathbb{R}^2} e^{-2\pi i [k_x(x-x_0) + k_y(y-y_0)]} \rho(x, y, z, t | x_0, y_0, z_0) dx dy \quad (\text{A3})$$

In other words,  $\hat{\rho}(k_x, k_y, z, t)$  satisfies a diffusion-reaction equation in one dimension. Considering the translational invariance within the  $xy$  plane for  $\rho(x, y, z, t | x_0, y_0, z_0)$ , which is solely a function of  $\Delta x = x - x_0$  and  $\Delta y = y - y_0$ , we observe that the marginal probability density:

$$p(z, t | z_0) = \int_{\mathbb{R}^2} \rho(\Delta x, \Delta y, z, t | z_0) d\Delta x d\Delta y$$

is identical to  $\hat{\rho}(0, 0, z, t | z_0)$ , as given in (A3). By setting  $k_x = k_y = 0$  in (A2), the reactive term vanishes, and  $p(z, t | z_0) = \hat{\rho}(0, 0, z, t | z_0)$  satisfies the one-dimensional Smoluchowski equation. This implies that for any system accurately described by (A1) or (62), the time series of  $z$  coordinates will adhere to the one-dimensional Smoluchowski picture. As such, methods developed for scalar diffusive coordinates, such as the ones discussed in Section 8, can be applied to estimate  $D_{zz}(z)$ .

## ■ APPENDIX B: RELATIONSHIP BETWEEN DIFFUSIVITY AND FORCE AUTOCORRELATION FUNCTION

Consider the SDE given by eq 95, which describes underdamped Langevin dynamics. By imposing the Boltzmann distribution as the steady-state distribution of (95), it can be demonstrated that  $\tilde{\mathbf{D}}(\mathbf{r}) = (m\beta)^{-2} \mathbf{D}^{-1}(\mathbf{r})$  and  $\gamma(\mathbf{r}) = (m\beta)^{-1} \mathbf{D}^{-1}(\mathbf{r})$ . The second term on the right-hand side of the velocity equation can be interpreted as a random acceleration term resulting from interactions with the surrounding environment. Denoting this random acceleration as  $\delta \mathbf{a}(\mathbf{X}_t)$ , Kubo's framework can be used to demonstrate that<sup>129</sup>

$$\begin{aligned} \tilde{\mathbf{D}} &= \int_0^\infty \langle \delta \mathbf{a}(\mathbf{X}_t) \delta \mathbf{a}^T(\mathbf{X}_0) \rangle dt \\ &= \frac{1}{m^2} \int_0^\infty \langle \delta \mathbf{f}(\mathbf{X}_t) \delta \mathbf{f}^T(\mathbf{X}_0) \rangle dt \end{aligned}$$

where  $\delta \mathbf{f}_t = m \delta \mathbf{a}_t$  represents the random forcing terms over the time interval  $dt$ . Eq 93 directly follows from the relationship between  $\mathbf{D}$  and  $\tilde{\mathbf{D}}$ . It is important to note that forces in molecular dynamics are continuous functions of time, so the assumption of the validity of underdamped Langevin dynamics is an approximation.

## ■ AUTHOR INFORMATION

### Corresponding Author

Amir Haji-Akbari – Department of Chemical and Environmental Engineering, Yale University, New Haven, Connecticut 06520, United States; [orcid.org/0000-0002-2228-6957](https://orcid.org/0000-0002-2228-6957); Email: [amir.hajiakbaribalou@yale.edu](mailto:amir.hajiakbaribalou@yale.edu)

### Authors

Tiago S. Domingues – Department of Chemical and Environmental Engineering, Yale University, New Haven, Connecticut 06520, United States

Ronald Coifman – Department of Mathematics and Department of Computer Science, Yale University, New Haven, Connecticut 06520, United States

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jctc.4c00148>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

A.H.-A. gratefully acknowledges the support from the National Science Foundation Grants CBET-1751971 (CAREER Award) and CBET-2024473. This work was supported as part of the Center for Enhanced Nanofluidic Transport (CENT), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award #DE-SC0019112.

## ■ REFERENCES

- (1) Alcoutlabi, M.; McKenna, G. B. Effects of confinement on material behaviour at the nanometre size scale. *J. Phys.: Condens. Mater.* **2005**, *17*, R461.
- (2) Jackson, C. L.; McKenna, G. B. The melting behavior of organic materials confined in porous solids. *J. Chem. Phys.* **1990**, *93*, 9002–9011.
- (3) Swallen, S. F.; Kearns, K. L.; Mapes, M. K.; Kim, Y. S.; McMahon, R. J.; Ediger, M. D.; Wu, T.; Yu, L.; Satija, S. Organic glasses with exceptional thermodynamic and kinetic stability. *Science* **2007**, *315*, 353–356.
- (4) Giovambattista, N.; Rossky, P. J.; Debenedetti, P. G. Phase transitions induced by nanoconfinement in liquid water. *Phys. Rev. Lett.* **2009**, *102*, No. 050603.
- (5) Zhang, C.; Guo, Y.; Priestley, R. D. Glass transition temperature of polymer nanoparticles under soft and hard confinement. *Macromolecules* **2011**, *44*, 4001–4006.
- (6) Chaban, V. V.; Prezhdo, V. V.; Prezhdo, O. V. Confinement by carbon nanotubes drastically alters the boiling and critical behavior of water droplets. *ACS Nano* **2012**, *6*, 2766–2773.
- (7) Moore, E. B.; Allen, J. T.; Molinero, V. Liquid-ice coexistence below the melting temperature for water confined in hydrophilic and hydrophobic nanopores. *J. Phys. Chem. C* **2012**, *116*, 7507–7514.

- (8) Rodriguez-Larrea, D.; Bayley, H. Multistep protein unfolding during nanopore translocation. *Nature Nanotechnol.* **2013**, *8*, 288–295.
- (9) Hayes, R.; El Abedin, S. Z.; Atkin, R. Pronounced structure in confined aprotic room-temperature ionic liquids. *J. Phys. Chem. B* **2009**, *113*, 7049–7052.
- (10) Fumagalli, L.; Esfandiar, A.; Fabregas, R.; Hu, S.; Ares, P.; Janardanan, A.; Yang, Q.; Radha, B.; Taniguchi, T.; Watanabe, K.; et al. Anomalously low dielectric constant of confined water. *Science* **2018**, *360*, 1339–1342.
- (11) Feng, X.; Kawabata, K.; Cowan, M. G.; Dwulet, G. E.; Toth, K.; Sixdenier, L.; Haji-Akbari, A.; Noble, R. D.; Elimelech, M.; Gin, D. L.; et al. Single crystal texture by directed molecular self-assembly along dual axes. *Nat. Mater.* **2019**, *18*, 1235–1243.
- (12) Le, T.; Morita, A.; Tanaka, T. Refractive index of nanoconfined water reveals its anomalous physical properties. *Nanoscale Horiz.* **2020**, *5*, 1016–1024.
- (13) Hu, H.-W.; Carson, G. A.; Granick, S. Relaxation time of confined liquids under shear. *Phys. Rev. Lett.* **1991**, *66*, 2758.
- (14) Demirel, A. L.; Granick, S. Glasslike transition of a confined simple fluid. *Phys. Rev. Lett.* **1996**, *77*, 2261.
- (15) Zhang, Z.; Sun, X.; Dresselhaus, M.; Ying, J. Y.; Heremans, J. Electronic transport properties of single-crystal bismuth nanowire arrays. *Phys. Rev. B* **2000**, *61*, 4850.
- (16) Raviv, U.; Laurat, P.; Klein, J. Fluidity of water confined to subnanometre films. *Nature* **2001**, *413*, 51–54.
- (17) Kim, O.; Jo, G.; Park, Y. J.; Kim, S.; Park, M. J. Ion transport properties of self-assembled polymer electrolytes: The role of confinement and interface. *J. Phys. Chem. Lett.* **2013**, *4*, 2111–2117.
- (18) Pourali, M.; Maghari, A. Non-equilibrium molecular dynamics simulation of thermal conductivity and thermal diffusion of binary mixtures confined in a nanochannel. *Chem. Phys.* **2014**, *444*, 30–38.
- (19) Berrod, Q.; Ferdeghini, F.; Judeinstein, P.; Genevaz, N.; Ramos, R.; Fournier, A.; Dijon, J.; Ollivier, J.; Rols, S.; Yu, D.; et al. Enhanced ionic liquid mobility induced by confinement in 1D CNT membranes. *Nanoscale* **2016**, *8*, 7845–7848.
- (20) Tu, C.-H.; Veith, L.; Butt, H.-J.; Floudas, G. Ionic conductivity of a solid polymer electrolyte confined in nanopores. *Macromolecules* **2022**, *55*, 1332–1341.
- (21) Lucent, D.; Vishal, V.; Pande, V. S. Protein folding under confinement: a role for solvent. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10430–10434.
- (22) Mittal, J.; Best, R. B. Thermodynamics and kinetics of protein folding under confinement. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20233–20238.
- (23) Jiang, Q.; Ward, M. D. Crystallization under nanoscale confinement. *Chem. Soc. Rev.* **2014**, *43*, 2066–2079.
- (24) Haji-Akbari, A.; Debenedetti, P. G. Computational investigation of surface freezing in a molecular model of water. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 3316–3321.
- (25) Altabet, Y. E.; Haji-Akbari, A.; Debenedetti, P. G. Effect of material flexibility on the thermodynamics and kinetics of hydrophobically induced evaporation of water. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E2548–E2555.
- (26) Hussain, S.; Haji-Akbari, A. Role of nanoscale interfacial proximity in contact freezing in water. *J. Am. Chem. Soc.* **2021**, *143*, 2272–2284.
- (27) Shoemaker, B. A.; Khalifa, O.; Haji-Akbari, A. Correlations in Charged Multipore Systems: Implications for Enhancing Selectivity and Permeability in Nanoporous Membranes. *ACS Nano* **2024**, *18*, 1420–1431.
- (28) Rosenbluth, M. N.; Rosenbluth, A. W. Further results on Monte Carlo equations of state. *J. Chem. Phys.* **1954**, *22*, 881–884.
- (29) Alder, B.; Frankel, S.; Lewinson, V. Radial Distribution Function Calculated by the Monte-Carlo Method for a Hard Sphere Fluid. *J. Chem. Phys.* **1955**, *23*, 417–419.
- (30) Alder, B. J.; Wainwright, T. E. Phase transition for a hard sphere system. *J. Chem. Phys.* **1957**, *27*, 1208–1209.
- (31) Alder, B. J.; Wainwright, T. E. Studies in molecular dynamics. I. General method. *J. Chem. Phys.* **1959**, *31*, 459–466.
- (32) Rahman, A. Correlations in the motion of atoms in liquid argon. *Phys. Rev.* **1964**, *136*, A405.
- (33) Gubbins, K. E.; Liu, Y.-C.; Moore, J. D.; Palmer, J. C. The role of molecular modeling in confined systems: impact and prospects. *Phys. Chem. Chem. Phys.* **2011**, *13*, 58–85.
- (34) Croxton, C.; Ferrier, R. Statistical mechanical calculation of surface properties of simple liquids. IV. Molecular dynamics. *J. Phys. C Solid State* **1971**, *4*, 2447.
- (35) Lee, J. K.; Barker, J.; Pound, G. Surface structure and surface tension: Perturbation theory and Monte Carlo calculation. *J. Chem. Phys.* **1974**, *60*, 1976–1980.
- (36) Liu, K. Phase separation of Lennard-Jones systems: A film in equilibrium with vapor. *J. Chem. Phys.* **1974**, *60*, 4226–4230.
- (37) Abraham, F. F.; Schreiber, D. E.; Barker, J. On the structure of a free surface of a Lennard-Jones liquid: A Monte Carlo calculation. *J. Chem. Phys.* **1975**, *62*, 1958–1960.
- (38) Toxvaerd, S.; Praetgaard, E. Molecular dynamics calculation of the liquid structure up to a solid surface. *J. Chem. Phys.* **1977**, *67*, 5291–5295.
- (39) Subramanian, G.; Davis, H. Molecular dynamics of a hard sphere fluid in small pores. *Mol. Phys.* **1979**, *38*, 1061–1066.
- (40) Onsager, L. Reciprocal relations in irreversible processes. I. *Phys. Rev.* **1931**, *37*, 405.
- (41) Avron, J. Odd viscosity. *J. Stat. Phys.* **1998**, *92*, 543–557.
- (42) Hoover, W. G. Nonequilibrium molecular dynamics. *Annu. Rev. Phys. Chem.* **1983**, *34*, 103–127.
- (43) Evans, D. J.; Morriss, O. Non-Newtonian molecular dynamics. *Comp. Phys. Rep.* **1984**, *1*, 297–343.
- (44) Vogelsang, R.; Hoheisel, G.; Luckas, M. Shear viscosity and thermal conductivity of the Lennard-Jones liquid computed using molecular dynamics and predicted by a memory function model for a large number of states. *Mol. Phys.* **1988**, *64*, 1203–1213.
- (45) Svishchev, I.; Kusalik, P. Nonequilibrium molecular dynamics of a dense ionic fluid. *Phys. Chem. Liq.* **1994**, *26*, 237–246.
- (46) Wang, Y.; Ruan, X.; Roy, A. K. Two-temperature non-equilibrium molecular dynamics simulation of thermal transport across metal-nonmetal interfaces. *Phys. Rev. B* **2012**, *85*, 205311.
- (47) Jadhao, V.; Robbins, M. O. Probing large viscosities in glass-formers with nonequilibrium simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 7952–7957.
- (48) Müller-Plathe, F. A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity. *J. Chem. Phys.* **1997**, *106*, 6082–6085.
- (49) Tenney, C. M.; Maginn, E. J. Limitations and recommendations for the calculation of shear viscosity using reverse nonequilibrium molecular dynamics. *J. Chem. Phys.* **2010**, *132*, 014103.
- (50) Marconi, U. M. B.; Puglisi, A.; Rondoni, L.; Vulpiani, A. Fluctuation–dissipation: response theory in statistical physics. *Phys. Rep.* **2008**, *461*, 111–195.
- (51) Green, M. S. Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids. *J. Chem. Phys.* **1954**, *22*, 398–413.
- (52) Kubo, R. Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. *J. Phys. Soc. Jpn.* **1957**, *12*, 570–586.
- (53) Bradlyn, B.; Goldstein, M.; Read, N. Kubo formulas for viscosity: Hall viscosity, Ward identities, and the relation with conductivity. *Phys. Rev. B* **2012**, *86*, 245309.
- (54) Shi, K.; Smith, E. R.; Santiso, E. E.; Gubbins, K. E. A perspective on the microscopic pressure (stress) tensor: History, current understanding, and future challenges. *J. Chem. Phys.* **2023**, *158*, No. 040901.
- (55) Schelling, P. K.; Phillpot, S. R.; Keblinski, P. Comparison of atomic-level simulation methods for computing thermal conductivity. *Phys. Rev. B* **2002**, *65*, 144306.
- (56) Ma, Q.; Kang, D.; Zhao, Z.; Dai, J. Directly calculated electrical conductivity of hot dense hydrogen from molecular dynamics

simulation beyond Kubo-Greenwood formula. *Phys. Plasmas* **2018**, *25*, No. 012707.

(57) Helfand, E. Transport coefficients from dissipation in a canonical ensemble. *Phys. Rev.* **1960**, *119*, 1.

(58) Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **1905**, *322*, 549–560.

(59) Kinaci, A.; Haskins, J. B.; Çağın, T. On calculation of thermal conductivity from Einstein relation in equilibrium molecular dynamics. *J. Chem. Phys.* **2012**, *137*, No. 014106.

(60) Viscardi, S.; Gaspard, P. Viscosity in molecular dynamics with periodic boundary conditions. *Phys. Rev. E* **2003**, *68*, No. 041204.

(61) Mittal, J.; Hummer, G. Pair diffusion, hydrodynamic interactions, and available volume in dense fluids. *J. Chem. Phys.* **2012**, *137*, No. 034110.

(62) Van Hove, L. Correlations in space and time and Born approximation scattering in systems of interacting particles. *Phys. Rev.* **1954**, *95*, 249.

(63) Fokker, A. D. Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld. *Ann. Phys.* **1914**, *348*, 810–820.

(64) Planck, M. Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie. *Sitzber. Preuss. Akad.* **1917**, *24*, 324–341.

(65) Smoluchowski, P. M. Drei vorträge über diffusion, brownische molekularbewegung und koagulation von kolloidteilchen. *Phys. Z.* **1916**, *2*, 530–594.

(66) Teboul, V.; Simionescu, C. A. Properties of a confined molecular glass-forming liquid. *J. Phys.: Condens. Mater.* **2002**, *14*, 5699.

(67) Desai, T.; Koblinski, P.; Kumar, S. K. Molecular dynamics simulations of polymer transport in nanocomposites. *J. Chem. Phys.* **2005**, *122*, 134910.

(68) Marrink, S.-J.; Berendsen, H. J. Simulation of water transport through a lipid membrane. *J. Phys. Chem.* **1994**, *98*, 4155–4168.

(69) Lançon, P.; Batrouni, G.; Lobry, L.; Ostrowsky, N. Brownian walker in a confined geometry leading to a space-dependent diffusion coefficient. *Physica A* **2002**, *304*, 65–76.

(70) Haji-Akbari, A.; Debenedetti, P. G. The effect of substrate on thermodynamic and kinetic anisotropies in atomic thin films. *J. Chem. Phys.* **2014**, *141*, No. 024506.

(71) Haji-Akbari, A.; Debenedetti, P. G. Thermodynamic and kinetic anisotropies in octane thin films. *J. Chem. Phys.* **2015**, *143*, 214501.

(72) Liu, P.; Harder, E.; Berne, B. On the calculation of diffusion coefficients in confined fluids and interfaces with an application to the liquid-vapor interface of water. *J. Phys. Chem. B* **2004**, *108*, 6595–6602.

(73) Shi, Z.; Debenedetti, P. G.; Stillinger, F. H. Properties of model atomic free-standing thin films. *J. Chem. Phys.* **2011**, *134*, 114524.

(74) Nagai, T.; Tsurumaki, S.; Urano, R.; Fujimoto, K.; Shinoda, W.; Okazaki, S. Position-Dependent Diffusion Constant of Molecules in Heterogeneous Systems as Evaluated by the Local Mean Squared Displacement. *J. Chem. Theory Comput.* **2020**, *16*, 7239–7254.

(75) Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **1940**, *7*, 284–304.

(76) Moyal, J. Stochastic processes and statistical physics. *J. R. Stat. Soc. B Met.* **1949**, *11*, 150–210.

(77) Kob, W.; Andersen, H. C. Scaling behavior in the  $\beta$ -relaxation regime of a supercooled Lennard-Jones mixture. *Phys. Rev. Lett.* **1994**, *73*, 1376.

(78) Friedrich, R.; Renner, C.; Siefert, M.; Peinke, J. Comment on “Indispensable Finite Time Corrections for Fokker-Planck Equations from Time Series Data. *Phys. Rev. Lett.* **2002**, *89*, 149401.

(79) Gottschall, J.; Peinke, J. On the definition and handling of different drift and diffusion estimates. *New J. Phys.* **2008**, *10*, No. 083034.

(80) Sedlmeier, F.; von Hansen, Y.; Mengyu, L.; Horinek, D.; Netz, R. R. Water Dynamics at Interfaces and Solutes: Disentangling Free Energy and Diffusivity Contributions. *J. Stat. Phys.* **2011**, *145*, 240–252.

(81) Oliveira, R. J. d. Coordinate-Dependent Drift-Diffusion Reveals the Kinetic Intermediate Traps of Top7-Based Proteins. *J. Phys. Chem. B* **2022**, *126*, 10854–10869.

(82) Hinczewski, M.; von Hansen, Y.; Dzubiella, J.; Netz, R. R. How the diffusivity profile reduces the arbitrariness of protein folding free energies. *J. Chem. Phys.* **2010**, *132*, 245103.

(83) Yang, S.; Onuchic, J. N.; Garcia, A. E.; Levine, H. Folding Time Predictions from All-atom Replica Exchange Simulations. *J. Mol. Biol.* **2007**, *372*, 756–763.

(84) Yang, S.; Onuchic, J. N.; Levine, H. Effective stochastic dynamics on a protein folding energy landscape. *J. Chem. Phys.* **2006**, *125*, No. 054910.

(85) Freitas, F. C.; Lima, A. N.; Contessoto, V. d. G.; Whitford, P. C.; Oliveira, R. J. d. Drift-diffusion (DrDiff) framework determines kinetics and thermodynamics of two-state folding trajectory and tunes diffusion models. *J. Chem. Phys.* **2019**, *151*, 114106.

(86) Mamonov, A. B.; Kurnikova, M. G.; Coalson, R. D. Diffusion constant of K<sup>+</sup> inside Gramicidin A: A comparative study of four computational methods. *Biophys. Chem.* **2006**, *124*, 268–278.

(87) Hunter, M. A.; Demir, B.; Petersen, C. F.; Searles, D. J. New Framework for Computing a General Local Self-Diffusion Coefficient Using Statistical Mechanics. *J. Chem. Theory Comput.* **2022**, *18*, 3357–3363.

(88) Mangaud, E.; Rotenberg, B. Sampling mobility profiles of confined fluids with equilibrium molecular dynamics simulations. *J. Chem. Phys.* **2020**, *153*, No. 044125.

(89) Lennard-Jones, J. E. On the Determination of Molecular Fields. *Proc. R. Soc. London A* **1924**, *106*, 463–477.

(90) Schuss, Z. *Theory and Applications of Stochastic Processes: An Analytical Approach*; Applied Mathematical Sciences; Springer-Verlag: New York, 2010.

(91) Weiss, G. H. *Advances in Chemical Physics*; John Wiley & Sons, Ltd., 1967; pp 1–18.

(92) Domingues, T. S.; Coifman, R. R.; Haji-Akbari, A. Robust Estimation of Position-Dependent Anisotropic Diffusivity Tensors from Stochastic Trajectories. *J. Phys. Chem. B* **2023**, *127*, 5273–5287.

(93) Domingues, T. S.; Coifman, R. R.; Haji-Akbari, A. Robust Estimation of Position-Dependent Anisotropic Diffusivity Tensors from Molecular Dynamics Trajectories. *J. Phys. Chem. B* **2023**, *127*, 8644–8659.

(94) Hegger, R.; Stock, G. Multidimensional Langevin modeling of biomolecular dynamics. *J. Chem. Phys.* **2009**, *130*, No. 034106.

(95) Nadaraya, E. A. On Estimating Regression. *Theor. Probab. Appl.* **1964**, *9*, 141–142.

(96) Watson, G. S. Smooth regression analysis. *Sankhya Ser. A* **1964**, *26*, 359–372.

(97) Florens-Zmirou, D. On Estimating the Diffusion Coefficient from Discrete Observations. *J. Appl. Probab.* **1993**, *30*, 790–804.

(98) Bandi, F. M.; Phillips, P. C. B. Fully Nonparametric Estimation of Scalar Diffusion Models. *Econometrica* **2003**, *71*, 241–283.

(99) Jiang, G. J.; Knight, J. L. A Nonparametric Approach to the Estimation of Diffusion Processes, with an Application to a Short-Term Interest Rate Model. *Economet. Theor.* **1997**, *13*, 615–645.

(100) Lamouroux, D.; Lehnertz, K. Kernel-based regression of drift and diffusion coefficients of stochastic processes. *Phys. Lett. A* **2009**, *373*, 3507–3512.

(101) Phillips, J. Stretched exponential relaxation in molecular and electronic glasses. *Rep. Prog. Phys.* **1996**, *59*, 1133.

(102) Coifman, R. R.; Lafon, S. Diffusion maps. *Appl. Comput. Harmon. Anal.* **2006**, *21*, 5–30.

(103) Silverman, B. W. *Density estimation for statistics and data analysis*; Routledge, 2018.

(104) Comer, J.; Chipot, C.; González-Nilo, F. D. Calculating Position-Dependent Diffusivity in Biased Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 876–882.

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

- (106) Palmer, B. J.; Chun, J.; Morris, J. F.; Mundy, C. J.; Schenter, G. K. Correlation function approach for diffusion in confined geometries. *Phys. Rev. E* **2020**, *102*, No. 022129.
- (107) Sicard, F.; Koskin, V.; Annibale, A.; Rosta, E. Position-Dependent Diffusion from Biased Simulations and Markov State Model Analysis. *J. Chem. Theory Comput.* **2021**, *17*, 2022–2033.
- (108) Bicout, D. J.; Szabo, A. Electron transfer reaction dynamics in non-Debye solvents. *J. Chem. Phys.* **1998**, *109*, 2325–2338.
- (109) Mittal, J.; Truskett, T. M.; Errington, J. R.; Hummer, G. Layering and Position-Dependent Diffusive Dynamics of Confined Fluids. *Phys. Rev. Lett.* **2008**, *100*, 145901.
- (110) Hermans, J. The amino acid dipeptide: small but still influential after 50 years. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3095–3096.
- (111) Ghorbani, M.; Wang, E.; Krämer, A.; Klauda, J. B. Molecular dynamics simulations of ethanol permeation through single and double-lipid bilayers. *J. Chem. Phys.* **2020**, *153*, 125101.
- (112) Krämer, A.; Ghysels, A.; Wang, E.; Venable, R. M.; Klauda, J. B.; Brooks, B. R.; Pastor, R. W. Membrane permeability of small molecules from unbiased molecular dynamics simulations. *J. Chem. Phys.* **2020**, *153*, 124107.
- (113) Ghysels, A.; Venable, R. M.; Pastor, R. W.; Hummer, G. Position-Dependent Diffusion Tensors in Anisotropic Media from Simulation: Oxygen Transport in and through Membranes. *J. Chem. Theory Comput.* **2017**, *13*, 2962–2976.
- (114) Beltran-Villegas, D. J.; Edwards, T. D.; Bevan, M. A. Self-Consistent Colloidal Energy and Diffusivity Landscapes in Macromolecular Solutions. *Langmuir* **2013**, *29*, 12337–12341.
- (115) Best, R. B.; Hummer, G.; Wolynes, P. G. Coordinate-Dependent Diffusion in Protein Folding. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1088–1093.
- (116) Chang, J. C.; Fok, P.-W.; Chou, T. Bayesian Uncertainty Quantification for Bond Energies and Mobilities Using Path Integral Analysis: Biophysical Journal. *Biophys. J.* **2015**, *109*, 966–974.
- (117) Türkcan, S.; Alexandrou, A.; Masson, J.-B. A Bayesian Inference Scheme to Extract Diffusivity and Potential Fields from Confined Single-Molecule Trajectories. *Biophys. J.* **2012**, *102*, 2288–2298.
- (118) Masson, J.-B.; Dionne, P.; Salvatico, C.; Renner, M.; Specht, C. G.; Triller, A.; Dahan, M. Mapping the Energy and Diffusion Landscapes of Membrane Proteins at the Cell Surface Using High-Density Single-Molecule Imaging and Bayesian Inference: Application to the Multiscale Dynamics of Glycine Receptors in the Neuronal Membrane. *Biophys. J.* **2014**, *106*, 74–83.
- (119) Husic, B. E.; Pande, V. S. Markov state models: From an art to a science. *J. Am. Chem. Soc.* **2018**, *140*, 2386–2396.
- (120) Schulz, R.; Yamamoto, K.; Klossek, A.; Flesch, R.; Hönzke, S.; Rancan, F.; Vogt, A.; Blume-Peytavi, U.; Hedtrich, S.; Schäfer-Korting, M.; Rühl, E.; Netz, R. R. Data-based modeling of drug penetration relates human skin barrier function to the interplay of diffusivity and free-energy profiles. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 3631–3636.
- (121) Straub, J. E.; Borkovec, M.; Berne, B. J. Calculation of dynamic friction on intramolecular degrees of freedom. *J. Phys. Chem.* **1987**, *91*, 4995–4998.
- (122) Straub, J. E.; Berne, B. J.; Roux, B. Spatial dependence of time-dependent friction for pair diffusion in a simple fluid. *J. Phys. Chem.* **1990**, *93*, 6804–6812.
- (123) Miller, C. C. The Stokes-Einstein law for diffusion in solution. *Proc. R. Soc. London A* **1924**, *106*, 724–749.
- (124) Chahine, J.; Oliveira, R. J.; Leite, V. B. P.; Wang, J. Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14646–14651.
- (125) Woolf, T. B.; Roux, B. Conformational Flexibility of o-Phosphorylcholine and o-Phosphorylethanolamine: A Molecular Dynamics Study of Solvation Effects. *J. Am. Chem. Soc.* **1994**, *116*, 5916–5926.
- (126) Gaalswyk, K.; Awoonor-Williams, E.; Rowley, C. N. Generalized Langevin Methods for Calculating Transmembrane Diffusivity. *J. Chem. Theory Comput.* **2016**, *12*, 5609–5619.
- (127) Succi, N. D.; Onuchic, J. N.; Wolynes, P. G. Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* **1996**, *104*, 5860–5868.
- (128) Holland, B. W.; Gray, C. G.; Tomberli, B. Calculating diffusion and permeability coefficients with the 39oscillating forward-reverse method. *Phys. Rev. E* **2012**, *86*, No. 036707.
- (129) Kubo, R. The fluctuation-dissipation theorem. *Rep. Prog. Phys.* **1966**, *29*, 255.
- (130) Roux, B.; Karplus, M. Ion transport in a gramicidin-like channel: dynamics and mobility. *J. Phys. Chem.* **1991**, *95*, 4856–4868.
- (131) Saito, H.; Shinoda, W. Cholesterol Effect on Water Permeability through DPPC and PSM Lipid Bilayers: A Molecular Dynamics Study. *J. Phys. Chem. B* **2011**, *115*, 15241–15250.
- (132) Carpenter, T. S.; Kirshner, D. A.; Lau, E. Y.; Wong, S. E.; Nilmeier, J. P.; Lightstone, F. C. A Method to Predict Blood-Brain Barrier Permeability of Drug-Like Compounds Using Molecular Dynamics Simulations. *Biophys. J.* **2014**, *107*, 630–641.
- (133) Sugii, T.; Takagi, S.; Matsumoto, Y. A molecular-dynamics study of lipid bilayers: Effects of the hydrocarbon chain length on permeability. *J. Chem. Phys.* **2005**, *123*, 184714.
- (134) Shinoda, W.; Mikami, M.; Baba, T.; Hato, M. Molecular Dynamics Study on the Effects of Chain Branching on the Physical Properties of Lipid Bilayers: 2. Permeability. *J. Phys. Chem. B* **2004**, *108*, 9346–9356.
- (135) Shinoda, W. Permeability across lipid membranes. *BBA - Biomembranes* **2016**, *1858*, 2254–2265.
- (136) Ciccotti, G.; Jacucci, G. Direct Computation of Dynamical Response by Molecular Dynamics: The Mobility of a Charged Lennard-Jones Particle. *Phys. Rev. Lett.* **1975**, *35*, 789–792.
- (137) McKinnon, S. J.; Whittenburg, S. L.; Brooks, B. Non-equilibrium molecular dynamics simulation of oxygen diffusion through hexadecane monolayers with varying concentrations of cholesterol. *J. Phys. Chem.* **1992**, *96*, 10497–10506.
- (138) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946–5961.
- (139) Berendsen, H. J.; Postma, J. v.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (140) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.
- (141) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (142) Daldrop, J. O.; Kowalik, B. G.; Netz, R. R. External Potential Modifies Friction of Molecular Solutes in Water. *Phys. Rev. X* **2017**, *7*, No. 041065.
- (143) Berendsen, H. J.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (144) Malmir, H.; Epsztein, R.; Elimelech, M.; Haji-Akbari, A. Induced charge anisotropy: A hidden variable affecting ion transport through membranes. *Matter* **2020**, *2*, 735–750.
- (145) Fujimoto, K.; Nagai, T.; Yamaguchi, T. Momentum removal to obtain the position-dependent diffusion constant in constrained molecular dynamics simulation. *J. Comput. Chem.* **2021**, *42*, 2136–2144.
- (146) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **2016**, *67*, 159–184.
- (147) Hussain, S.; Haji-Akbari, A. Studying rare events using forward-flux sampling: Recent breakthroughs and future outlook. *J. Chem. Phys.* **2020**, *152*, No. 060901.
- (148) Sittel, F.; Stock, G. Perspective: Identification of collective variables and metastable states of protein dynamics. *J. Chem. Phys.* **2018**, *149*, 150901.
- (149) Kacirani, A.; Uralcan, B.; Domingues, T. S.; Haji-Akbari, A. Effect of Pressure on the Conformational Landscape of Human  $\gamma$ D-

Crystallin from Replica Exchange Molecular Dynamics Simulations. *J. Phys. Chem. B* **2024**, *128* (20), 4931–4942.

(150) Ferguson, A. L.; Panagiotopoulos, A. Z.; Debenedetti, P. G.; Kevrekidis, I. G. Systematic determination of order parameters for chain dynamics using diffusion maps. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 13597–13602.

(151) Evans, L.; Cameron, M. K.; Tiwary, P. Computing committors in collective variables via Mahalanobis diffusion maps. *Appl. Comput. Harmon. Anal.* **2023**, *64*, 62–101.

(152) Nadler, B.; Lafon, S.; Coifman, R. R.; Kevrekidis, I. G. Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. *Appl. Comput. Harmon. Anal.* **2006**, *21*, 113–127.

(153) Appeldorn, J. H.; Lemcke, S.; Speck, T.; Nikoubashman, A. Employing artificial neural networks to identify reaction coordinates and pathways for self-assembly. *J. Phys. Chem. B* **2022**, *126*, 5007–5016.

(154) Beyerle, E. R.; Zou, Z.; Tiwary, P. Recent advances in describing and driving crystal nucleation using machine learning and artificial intelligence. *Curr. Opin. Solid St. M* **2023**, *27*, 101093.

(155) Domingues, T.; Hussain, S.; Haji-Akbari, A. Divergence between local structure, dynamics and nucleation outcome in heterogeneous nucleation of close-packed crystals. *J. Phys. Chem. Lett.* **2024**, *15*, 1279–1287.

(156) Zwanzig, R. Ensemble method in the theory of irreversibility. *J. Chem. Phys.* **1960**, *33*, 1338–1341.

(157) Mori, H. Transport, collective motion, and Brownian motion. *Prog. Theor. Phys.* **1965**, *33*, 423–455.

(158) Mendels, D.; Piccini, G.; Parrinello, M. Collective variables from local fluctuations. *J. Phys. Chem. Lett.* **2018**, *9*, 2776–2781.

(159) E, W.; Vanden-Eijnden, E. Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* **2010**, *61*, 391–420.

(160) Bollinger, J. A.; Jain, A.; Truskett, T. M. Structure, Thermodynamics, and Position-Dependent Diffusivity in Fluids with Sinusoidal Density Variations. *Langmuir* **2014**, *30*, 8247–8252.

(161) Belousov, R.; Qaisrani, M. N.; Hassanali, A.; Roldán, É. First-passage fingerprints of water diffusion near glutamine surfaces. *Soft Matter* **2020**, *16*, 9202–9216.

(162) Belousov, R.; Hassanali, A.; Roldán, É. Statistical physics of inhomogeneous transport: Unification of diffusion laws and inference from first-passage statistics. *Phys. Rev. E* **2022**, *106*, No. 014103.

(163) Berezhkovskii, A. M.; Makarov, D. E. Communication: Coordinate-dependent diffusivity from single molecule trajectories. *J. Chem. Phys.* **2017**, *147*, 201102.

(164) E, W.; Vanden-Eijnden, E. Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events. *Annu. Rev. Phys. Chem.* **2010**, *61*, 391–420.

(165) Khoo, Y.; Lu, J.; Ying, L. Solving for high-dimensional committor functions using artificial neural networks. *Res. Math. Sci.* **2019**, *6*, 1.

(166) Molinaro, A. M.; Simon, R.; Pfeiffer, R. M. Prediction error estimation: a comparison of resampling methods. *Bioinformatics* **2005**, *21*, 3301–3307.

(167) Schwantes, C. R.; Pande, V. S. Modeling molecular kinetics with tICA and the kernel trick. *J. Chem. Theory Comput.* **2015**, *11*, 600–608.

(168) Yuecai, H.; Dingwen, Z. Nadaraya-Watson estimator for reflected stochastic processes driven by Brownian motions. *arXiv Preprint* **2022**, 00141.

(169) Ganguly, A. Infinite-dimensional optimization and Bayesian nonparametric learning of stochastic differential equations. *J. Mach. Learn. Res.* **2022**, *24*, 1–39.

(170) Ren, Y.; Lu, Y.; Ying, L.; Rotskoff, G. M. Statistical Spatially Inhomogeneous Diffusion Inference. *arXiv Preprint* **2023**, 05793.

(171) Markel, V. A.; Schotland, J. C. Inverse problem in optical diffusion tomography. I. Fourier–Laplace inversion formulas. *J. Opt. Soc. Am.* **2001**, *18*, 1336–1347.

(172) Croomlin, D.; Vanden-Eijnden, E. Diffusion estimation from multiscale data by operator eigenpairs. *Multiscale Model. Simul.* **2011**, *9*, 1588–1623.

(173) Vroylandt, H.; Goudenège, L.; Monmarché, P.; Pietrucci, F.; Rotenberg, B. Likelihood-based non-Markovian models from molecular dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, No. e2117586119.

(174) Xie, P.; Car, R.; E, W. Ab Initio Generalized Langevin Equations. *arXiv Preprint* **2022**, 06558 DOI: 10.48550/arXiv.2211.06558.

(175) Jeon, J.-H.; Monne, H. M.-S.; Javanainen, M.; Metzler, R. Anomalous diffusion of phospholipids and cholesterol in a lipid bilayer and its origins. *Phys. Rev. Lett.* **2012**, *109*, 188103.

(176) von Hansen, Y.; Gekle, S.; Netz, R. R. Anomalous anisotropic diffusion dynamics of hydration water at lipid membranes. *Phys. Rev. Lett.* **2013**, *111*, 118103.

(177) Krott, L. B.; Gavazzoni, C.; Bordin, J. R. Anomalous diffusion and diffusion anomaly in confined Janus dumbbells. *J. Chem. Phys.* **2016**, *145*, 244906.

(178) McKean, H. P., Jr A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *56*, 1907–1911.

(179) Helms, P.; Poggioli, A. R.; Limmer, D. T. Intrinsic Interface Adsorption Drives Selectivity in Atomically Smooth Nanofluidic Channels. *Nano Lett.* **2023**, *23*, 4226–4233.