Supplementary material for 'Automated calculation and convergence of defect transport tensors'

Thomas D Swinburne
*Aix-Marseille Université, CNRS, CINaM UMR 7325, Campus de Luminy, 13288 Marseille, France**

Danny Perez
Theoretical Division T-1, Los Alamos National Laboratory, Los Alamos, NM, 87545, USA
(Dated: October 9, 2020)

CONTENTS

SUPPLEMENTARY NOTE 1: EFFECT OF SAMPLING IN IRREDUCIBLE REPRESENTATION

In previous work,[1] we used an initial version of TAMMBER to study the breakup of a C15-dumbell interstitial cluster in an EAM model of iron. We refer the reader to the original paper for details. We later performed an identical simulation using the new version of TAMMBER that compresses the state space for sampling to a representation irreducible under exchange, space group and translational symmetries, as described in the main text. The compression in the size of state space, and the resultant effect on the model quality, are summarized in figure 1. Interestingly, despite the low degree of space group symmetry for this system, due to the presence of multiple nonlocal exchange events during sampling, the compression of exchange symmetries produced a significant efficiency improvement.

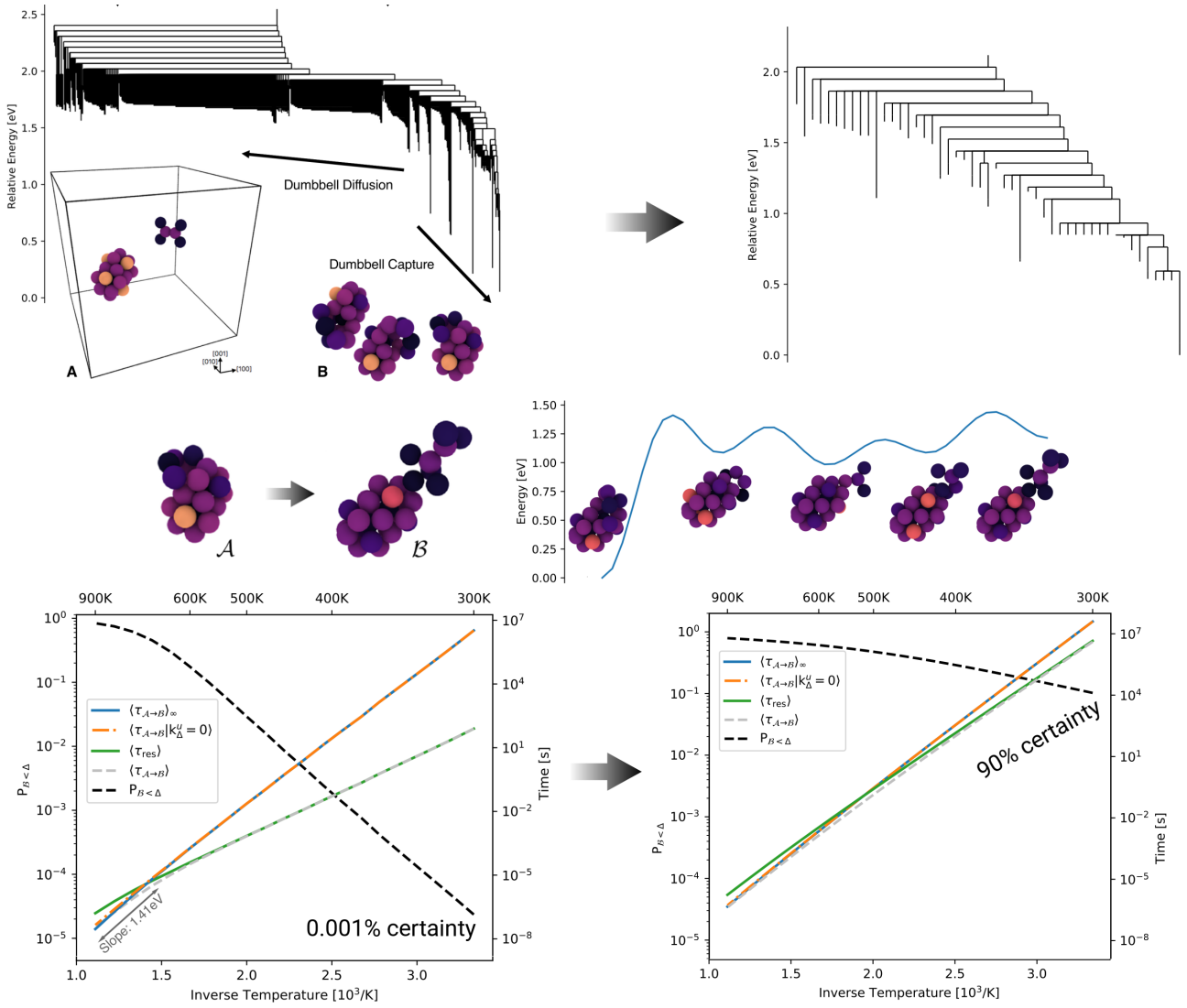

FIG. 1. Supplementary Figure 1: Above: Connectivity graph[2] for the states and transitions found by TAMMBER using the previously presented reducible (left) and new irreducible (right) representation. Middle: Representative structures of the $\mathcal{A}$ and $\mathcal{B}$ basins (left) and the found minimum energy pathway (right). Below: Probability $P_{\mathcal{B}<\Delta}$ that the system executes a $\mathcal{A} \rightarrow \mathcal{B}$ transition before absorption, residence times $\langle \tau_{res} \rangle$ and various approximate projected breakup times[1] at a range of temperatures using a Markov model built using the previously presented reducible (left) or new irreducible (right) representation. We see the irreducible representation achieves a significant compression of phase space and yields much greater certainty, as encoded by $P_{\mathcal{B}<\Delta}$, and that our model has sufficient predictive power to capture the dominant breakup mechanism.

SUPPLEMENTARY NOTE 2: EQUIVALENCE OF DERIVATIVES OF THE GENERATING FUNCTION TO PATH-WISE AVERAGES

We start by recalling the branching probability matrices for N irreducible states in D dimensions

$$[\mathbf{B}^{u,c}]_p = \frac{\mathrm{k}_p^{u,c}}{\mathrm{k}_p^{tot,c}}, \quad [\mathbf{B}^c(\boldsymbol{\lambda})]_{qp} \equiv \sum_{l\in\mathcal{C}_{qp}} \frac{\mathrm{k}_l^c}{\mathrm{k}_p^{tot,c}} \exp\left(-\boldsymbol{\lambda}\cdot\mathbf{d}_l\right), \quad [\mathbf{B}^c(\mathbf{0})]_{qp} = \frac{\mathrm{k}_{qp}^c}{\mathrm{k}_p^{tot,c}},$$

In the following, we will suppress the c superscript for clarity of presentation, meaning $\mathbf{B}^{u,c}\to\mathbf{B}^u, \mathbf{B}^c\to\mathbf{B}$, etc. The generating function with arbitrary initial condition $\mathbf{P}(0)$ reads

$$Z(\boldsymbol{\lambda}) = \mathbf{B}^u \sum_n \left[\mathbf{B}(\boldsymbol{\lambda})\right]^n \mathbf{P}(0) = \mathbf{B}^u \left[\mathbb{I} - \mathbf{B}(\boldsymbol{\lambda})\right]^{-1} \mathbf{P}(0) = \mathbf{B}^u\mathbf{G}(\boldsymbol{\lambda})\mathbf{P}(0). \tag{1}$$

By conservation of probability, it is clear that $[\mathbf{B}^u]_p + \sum_q [\mathbf{B}(\mathbf{0})]_{qp} = 1$, which implies that $\mathbf{B}^u[\mathbb{I} - \mathbf{B}(\mathbf{0})]^{-1} = \mathbf{1}$ and thus $Z(\mathbf{0}) = \mathbf{1}\mathbf{P}(0) = 1$.

Consider the set $\mathcal{P}_n$ of paths $\xi = \{\xi_i\}_0^n$ which execute n transitions between $n+1$ states before absorption. The probability weight P_ξ of each path and all n-paths is given by

$$P_\xi = [\mathbf{B}^u]_{\xi_n} \left(\prod_{l=1}^n [\mathbf{B}(\mathbf{0})]_{\xi_l\xi_{l-1}}\right) [\mathbf{P}(0)]_{\xi_0}, \quad \Rightarrow \sum_{\xi\in\mathcal{P}_n} P_\xi = \mathbf{B}^u \left[\mathbf{B}(\mathbf{0})\right]^n \mathbf{P}(0) \tag{2}$$

The sum over of path weights over the set of all possible paths $\mathcal{P} = \cup_n \mathcal{P}_n$ can be confirmed to be unity as expected-

$$\sum_{\xi\in\mathcal{P}} P_\xi = \sum_n \sum_{\xi\in\mathcal{P}_n} P_\xi = \sum_n \mathbf{B}^u \left[\mathbf{B}(\mathbf{0})\right]^n \mathbf{P}(0) = \mathbf{B}^u \left[\mathbb{I} - \mathbf{B}(\mathbf{0})\right]^{-1} \mathbf{P}(0) = Z(\mathbf{0}) = 1. \tag{3}$$

The total displacement $\mathbf{x}_\xi = \sum_l \mathbf{d}_{\xi_l\xi_{l-1}}$ across each path can be found by simply constructing the path-wise probability as before, but replacing $\mathbf{B}(\mathbf{0})$ with $\mathbf{B}(\boldsymbol{\lambda})$ to yield

$$P_\xi(\boldsymbol{\lambda}) = [\mathbf{B}^u]_{\xi_n} \left(\prod_{l=1}^n [\mathbf{B}(\mathbf{0})]_{\xi_l\xi_{l-1}}\right) \exp^{-\boldsymbol{\lambda}\cdot\left(\sum_l \mathbf{d}_{\xi_l\xi_{l-1}}\right)} [\mathbf{P}(0)]_{\xi_0} = P_\xi \exp^{-\boldsymbol{\lambda}\cdot\mathbf{x}_\xi} \Rightarrow \mathbf{x}_\xi = -\partial_{\boldsymbol{\lambda}} \ln P_\xi(\boldsymbol{\lambda})\Big|_{\boldsymbol{\lambda}=\mathbf{0}}. \tag{4}$$

The expected value (or first moment) of the total displacement over the set of all possible paths is therefore

$$\langle\mathbf{x}\rangle = \sum_{\xi\in\mathcal{P}} \mathbf{x}_\xi P_\xi = -\partial_{\boldsymbol{\lambda}} \sum_{\xi\in\mathcal{P}} P_\xi(\boldsymbol{\lambda})\Big|_{\boldsymbol{\lambda}=\mathbf{0}} = -\partial_{\boldsymbol{\lambda}} Z(\boldsymbol{\lambda})\Big|_{\boldsymbol{\lambda}=\mathbf{0}}. \tag{5}$$

The second moment matrix of the total displacement follows naturally-

$$\langle\mathbf{x}\otimes\mathbf{x}\rangle = \sum_{\xi\in\mathcal{P}} \mathbf{x}_\xi\otimes\mathbf{x}_\xi P_\xi = \partial_{\boldsymbol{\lambda}}\otimes\partial_{\boldsymbol{\lambda}} \sum_{\xi\in\mathcal{P}} P_\xi(\boldsymbol{\lambda})\Big|_{\boldsymbol{\lambda}=\mathbf{0}} = \partial_{\boldsymbol{\lambda}}\otimes\partial_{\boldsymbol{\lambda}} Z(\boldsymbol{\lambda})\Big|_{\boldsymbol{\lambda}=\mathbf{0}}, \tag{6}$$

where $\otimes$ is the outer (or dyadic) product.

SUPPLEMENTARY NOTE 3: EVALUATING DERIVATIVES OF THE GENERATING FUNCTION

Consider the following application of the chain rule, using the above definition $\mathbf{G}(\boldsymbol{\lambda}) = [\mathbb{I} - \mathbf{B}(\boldsymbol{\lambda})]^{-1}$, but writing $\mathbf{G}_{\boldsymbol{\lambda}}$ instead of $\mathbf{G}(\boldsymbol{\lambda})$ for readability-

$$\partial_{\lambda_\alpha}\left(\mathbf{G}\left[\mathbb{I} - \mathbf{B}_{\boldsymbol{\lambda}}\right]\right) = \partial_{\lambda_\alpha}\mathbf{G}_{\boldsymbol{\lambda}}\left[\mathbb{I} - \mathbf{B}_{\boldsymbol{\lambda}}\right] - \mathbf{G}_{\boldsymbol{\lambda}}\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}} = \mathbf{0} \quad \Rightarrow \partial_{\lambda_\alpha}\mathbf{G}_{\boldsymbol{\lambda}} = \mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}} \tag{7}$$

. The second derivative of $\mathbf{G}(\boldsymbol{\lambda})$ follows by induction as

$$\begin{aligned}\partial_{\lambda_\alpha}\partial_{\lambda_\beta}\mathbf{G}_{\boldsymbol{\lambda}} &= \left(\partial_{\lambda_\beta}\mathbf{G}_{\boldsymbol{\lambda}}\right)\left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}} + \mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\beta}\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}} + \mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\left(\partial_{\lambda_\beta}\mathbf{G}_{\boldsymbol{\lambda}}\right) &(8)\\ &= \mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\beta}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}} + \mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\beta}\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}} + \mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\beta}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}}. &(9)\end{aligned}$$

Using the above identity $\mathbf{B}^u\mathbf{G}_{\boldsymbol{\lambda}}|_{\boldsymbol{\lambda}=\mathbf{0}} = \mathbf{1}$ we find

$$\left[\langle\mathbf{x}\rangle\right]_\alpha = \mathbf{1}\left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\mathbf{GP}(0) \tag{10}$$

and

$$\left[\langle\mathbf{x}\otimes\mathbf{x}\rangle\right]_{\alpha\beta} = \mathbf{1}\left[\left(\partial_{\lambda_\beta}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right) + \left(\partial_{\lambda_\beta}\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right) + \left(\partial_{\lambda_\alpha}\mathbf{B}_{\boldsymbol{\lambda}}\right)\mathbf{G}_{\boldsymbol{\lambda}}\left(\partial_{\lambda_\beta}\mathbf{B}_{\boldsymbol{\lambda}}\right)\right]\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\mathbf{GP}(0), \tag{11}$$

where the vector and matrix-valued first and second derivatives of $\mathbf{B}_{\boldsymbol{\lambda}}$ write

$$\left[\partial_{\boldsymbol{\lambda}}\mathbf{B}_{\boldsymbol{\lambda}}|_{\boldsymbol{\lambda}=\mathbf{0}}\right]_{qp} = \sum_{l\in\mathcal{C}_{qp}}\frac{\mathrm{k}_l}{\mathrm{k}_p^{tot}}\mathbf{d}_l \equiv \left[\overline{\mathbf{B}\otimes\mathbf{d}}\right]_{qp}, \quad \left[\partial_{\boldsymbol{\lambda}}\otimes\partial_{\boldsymbol{\lambda}}\mathbf{B}(\boldsymbol{\lambda})|_{\boldsymbol{\lambda}=\mathbf{0}}\right]_{qp} = \sum_{l\in\mathcal{C}_{qp}}\frac{\mathrm{k}_l}{\mathrm{k}_p^{tot}}\mathbf{d}_l\otimes\mathbf{d}_l \equiv \left[\overline{\mathbf{B}\otimes\mathbf{d}\otimes\mathbf{d}}\right]_{qp}. \tag{12}$$

SUPPLEMENTARY NOTE 4: TRANSPORT COEFFICIENTS FROM TRAJECTORIES OF VARIABLE DURATION

Consider an average $\langle\dots\rangle_\tau$ over the set of all trajectories $\mathcal{P}_\tau$ with $\tau_{res}\in[\tau,\tau+\mathrm{d}\tau]$. By the definition of drift and diffusivity, the average displacement estimators $\langle\mathbf{x}\rangle_\tau$ and $\langle\mathbf{x}\otimes\mathbf{x}\rangle_\tau$ have limits

$$\langle\mathbf{x}\rangle_\tau = \sum_{\xi\in\mathcal{P}_\tau}\mathbf{x}_\xi P_\xi \to \tau\boldsymbol{\mu}, \quad \langle\mathbf{x}\otimes\mathbf{x}\rangle_\tau = \sum_{\xi\in\mathcal{P}_\tau}\mathbf{x}_\xi\otimes\mathbf{x}_\xi P_\xi \to 2\tau\mathbf{D} + \tau^2\boldsymbol{\mu}\otimes\boldsymbol{\mu}, \quad (\tau\to\infty) \tag{13}$$

which defines the true drift and diffusion coefficients $\boldsymbol{\mu}$ and $\mathbf{D}$. The behavior of $\langle\dots\rangle_\tau$ at small τ is challenging to quantify for *general* initial conditions, due to the possible presence of multiple transient modes. To see how a well defined estimates of $\boldsymbol{\mu}$ and $\mathbf{D}$ emerge, we remove initial transients by preparing the system in the longest lived eigenmode of $\mathbf{K}^{tot} - \mathbf{K}$, known as the quasistationary distribution (QSD)[3], with right eigenvector $\mathbf{v}_0 = \boldsymbol{\pi}^{QSD}$ and (smallest) eigenvalue $\nu_0 > 0$. In the complete sampling limit $\nu_0 \to 0$, the QSD limits to the Boltzmann distribution, $\boldsymbol{\pi}^{QSD} \to \hat{\boldsymbol{\pi}}$, and the left eigenvector $\mathbf{w}_0 = \mathbf{1}^{QSD} \to \mathbf{1}$. With $\mathbf{P}(0) = \hat{\boldsymbol{\pi}}^{QSD}$ the probability density in the known state space becomes the single exponential decay

$$\dot{\mathrm{P}}(\tau) = \mathbf{1}\dot{\mathbf{P}}(\tau) = \sum_l \nu_l\exp(-\nu_l\tau)\left(\mathbf{1}\mathbf{v}_l\right)\left(\mathbf{w}_l\boldsymbol{\pi}^{QSD}\right) = \nu_0\exp(-\nu_0 t). \tag{14}$$

The residence time thus has moments $\langle\tau_{res}\rangle = \nu_0^{-1}$ and $\langle\tau_{res}^2\rangle = 2\nu_0^{-2}$. Using these initial conditions we now consider averages over an ensemble of trajectories of variable duration, defined through

$$\langle\mathbf{x}\rangle \equiv \int_0^\infty\langle\mathbf{x}\rangle_\tau\dot{\mathrm{P}}(\tau)\mathrm{d}\tau, \quad \langle\mathbf{x}\otimes\mathbf{x}\rangle \equiv \int_0^\infty\langle\mathbf{x}\otimes\mathbf{x}\rangle_\tau\dot{\mathrm{P}}(\tau)\mathrm{d}\tau, \tag{15}$$

where $\dot{\mathrm{P}}(\tau)$ is the probability density of trajectory duration.

We have seen that the average displacement $\langle\mathbf{x}\rangle_\tau$ and displacement variance are all expected to scale linearly with trajectory duration τ. When $\dot{\mathrm{P}}(\tau)$ is a simple exponential decay, averages $\langle f(\tau)\rangle$ of functions that scale approximately linearly with τ, $f(\tau)\simeq A\tau$, will have a contribution of $\tau/\langle\tau_{\mathrm{res}}\rangle\exp(-\tau\langle\tau_{\mathrm{res}}\rangle)$ from trajectories of length τ. This contribution is negligible for trajectories shorter than $\langle\tau_{\mathrm{res}}\rangle$- for example, removing all contributions from trajectories with $\tau\leq 0.1\langle\tau_{\mathrm{res}}\rangle$ affects the final result by around 0.5%.

We therefore know by (13) that in the limit $\langle\tau_{\rm res}\rangle \to \infty$, ensemble averages $\langle\mathbf{x}\rangle, \langle\mathbf{x}\otimes\mathbf{x}\rangle$ will be swiftly dominated by trajectory averages where the drift and diffusion tensors are independent of the trajectory length, giving limits

$$\langle\mathbf{x}\rangle \to \langle\tau_{res}\rangle\boldsymbol{\mu}, \quad \langle\mathbf{x}\otimes\mathbf{x}\rangle \to 2\langle\tau_{res}\rangle\mathbf{D} + \langle\tau_{res}^2\rangle\boldsymbol{\mu}\otimes\boldsymbol{\mu}. \quad (\langle\tau_{\rm res}\rangle \to \infty, \quad \mathbf{P}(0) = \boldsymbol{\pi}^{QSD}) \tag{16}$$

This qualitative statement on convergence will be demonstrated in more detail in , where we formally take the $\langle\tau_{\rm res}\rangle \to \infty$ limit of our drift and diffusion estimators to demonstrate convergence to standard results for diffusion in periodic media[4,5]. Equation (16) is a key result, as it shows that averages over an ensemble of trajectories with exponentially distributed duration yield well defined transport coefficients when the ensemble *average* duration $\langle\tau_{\rm res}\rangle$ becomes very large, in close analogy to the typical ensemble of uniform duration. We therefore *define* our drift and diffusion estimators for general $\langle\tau_{\rm res}\rangle$ as

$$\boldsymbol{\mu}(\mathbf{k}^u) \equiv \frac{\langle\mathbf{x}\rangle}{\langle\tau_{res}\rangle}, \quad \mathbf{D}(\mathbf{k}^u) \equiv \frac{1}{2\langle\tau_{res}\rangle}\left[\langle\mathbf{x}\otimes\mathbf{x}\rangle - 2\langle\mathbf{x}\rangle\otimes\langle\mathbf{x}\rangle\right], \tag{17}$$

which we show in yield exact results as $\langle\tau_{\rm res}\rangle \to \infty$. Using the above relations (5) and (6), we thus obtain

$$\boldsymbol{\mu}(\mathbf{k}^u) = \langle\tau_{res}\rangle^{-1}\partial_{\boldsymbol{\lambda}}\left[1/Z(\boldsymbol{\lambda})\right]\Big|_{\boldsymbol{\lambda}=\mathbf{0},\mathbf{P}(0)=\hat{\boldsymbol{\pi}}^{QSD}}, \quad \mathbf{D}(\mathbf{k}^u) = \frac{1}{2}\langle\tau_{res}\rangle^{-1}\partial_{\boldsymbol{\lambda}}\otimes\partial_{\boldsymbol{\lambda}}\left[1/Z(\boldsymbol{\lambda})\right]\Big|_{\boldsymbol{\lambda}=\mathbf{0},\mathbf{P}(0)=\hat{\boldsymbol{\pi}}^{QSD}}, \tag{18}$$

as presented in the main text. We have shown that transport coefficients can be estimated from trajectories of variable length when the system is prepared in the QSD, as the influence of short trajectories, where transport coefficients are ill-defined, vanishes as $\langle\tau_{res}\rangle \to \infty$. We note that the diffusion tensor is no longer the covariance of the displacement as $\langle\tau_{res}^2\rangle \neq \langle\tau_{res}\rangle^2$ when trajectory duration are variable. In we provide a demonstration of convergence for a simple toy system.

The remainder of this section is devoted to evaluating explicit expressions for these estimators via (10) and (11), which we use to evaluate the diffusivity and perform our Monte Carlo uncertainty quantification. In we show how these estimators reduce to known expressions for drift and diffusion coefficients in periodic media[4,5] in the complete sampling limit $\langle\tau_{\rm res}\rangle \to \infty$ (equivalently $|\mathbf{k}^u| \to 0$), whilst in we demonstrate how various terms in the limiting expression simplify for systems where detailed balance holds.

Expressions for drift and diffusion estimators

To evaluate $\boldsymbol{\mu}(\mathbf{k}^u)$ and $\mathbf{D}(\mathbf{k}^u)$, we expand the Greens function matrix $\mathbf{G} = \mathbf{K}^{tot}\left[\mathbf{K}^{tot} - \mathbf{K}\right]^{-1}$ in the eigenbasis $\{\mathbf{w}_l, \mathbf{v}_l\}$ of $\mathbf{K}^{tot} - \mathbf{K}$, where $\mathbf{w}_l \cdot \mathbf{v}_m = \delta_{lm}$. This yields

$$\mathbf{G} = \mathbf{K}^{tot}\left[\mathbf{K}^{tot} - \mathbf{K}\right]^{-1} \equiv \mathbf{G}_+ + \langle\tau_{res}\rangle\mathbf{K}^{tot}\boldsymbol{\pi}^{QSD}\otimes\mathbf{1}^{QSD}, \tag{19}$$

where only the relation $\langle\tau_{res}\rangle = \nu_0^{-1}$ requires the system to be prepared in the QSD. The definition of $\mathbf{G}_+$ in the limit $\nu_0 \to 0$ is discussed in . Using these identities in the general expression (10), the drift estimator emerges as

$$\boldsymbol{\mu}(\mathbf{k}^u) \equiv \frac{\langle\mathbf{x}\rangle}{\langle\tau_{res}\rangle} = \mathbf{1}\overline{\mathbf{B}\otimes\mathbf{d}}\,\mathbf{K}^{tot}\hat{\boldsymbol{\pi}}^{QSD}. \tag{20}$$

For the second moment we use the expansion (19) of $\mathbf{G}$ in (11) to write $\langle\mathbf{x}\otimes\mathbf{x}\rangle = \mathbf{M} + \mathbf{M}^\top$ where

$$\mathbf{M} = \langle\tau_{res}\rangle\mathbf{1}\left[\frac{1}{2}\overline{\mathbf{B}\otimes\mathbf{d}\otimes\mathbf{d}} + \overline{\mathbf{B}\otimes\mathbf{d}}\left[\mathbf{G}_+ + \langle\tau_{res}\rangle\mathbf{K}^{tot}\boldsymbol{\pi}^{QSD}\otimes\mathbf{1}^{QSD}\right]\overline{\mathbf{B}\otimes\mathbf{d}}\right]\mathbf{K}^{tot}\hat{\boldsymbol{\pi}}^{QSD}. \tag{21}$$

Multiplying out this expression and noting that $\langle\tau_{res}^2\rangle = 2\langle\tau_{res}\rangle^2$ when we prepare the system in the QSD, the diffusivity estimator reads

$$\mathbf{D}(\mathbf{k}^u) \equiv \frac{1}{2\langle\tau_{res}\rangle}\left[\langle\mathbf{x}\otimes\mathbf{x}\rangle - 2\langle\mathbf{x}\rangle\otimes\langle\mathbf{x}\rangle\right] = \frac{1}{2}\left[\widetilde{\mathbf{D}}(\mathbf{k}^u) + \widetilde{\mathbf{D}}^\top(\mathbf{k}^u)\right], \tag{22}$$

$$\widetilde{\mathbf{D}}(\mathbf{k}^u) = \mathbf{1}\left[\frac{1}{2}\overline{\mathbf{B}\otimes\mathbf{d}\otimes\mathbf{d}} + \overline{\mathbf{B}\otimes\mathbf{d}}\mathbf{G}_+\overline{\mathbf{B}\otimes\mathbf{d}}\right]\mathbf{K}^{tot}\hat{\boldsymbol{\pi}}^{QSD}. \tag{23}$$

Infinite sampling limit and connection to previous work

In the limit of complete sampling $|\mathbf{k}^u| \to 0$, $\nu_0 \to 0$ and thus $\langle \tau_{res} \rangle \to \infty$ and $\hat{\boldsymbol{\pi}}^{QSD} \to \hat{\boldsymbol{\pi}}$, with the Greens function matrix $\mathbf{G}$ becoming singular, as can be seen from the expansion (19), with $\mathbf{G}_+$ thus the *pseudoinverse* of $[\mathbb{I} - \mathbf{B}] = [\mathbf{K}^{tot} - \mathbf{K}]\,[\mathbf{K}^{tot}]^{-1}$. We take the $\nu_0 \to 0$ limit by setting the unknown rate to be very small and the same for each state, namely $\mathbf{k}^u \to \mathrm{k}^u \mathbf{1}^\top$, where $\mathrm{k}^u \to 0$. Combining this with the eigenvalue relation for ν_0, we find $\mathbf{1}\,[\mathbf{K}^{tot} - \mathbf{K}]\,\hat{\boldsymbol{\pi}}^{QSD} = \mathrm{k}^u = \nu_0$. We can then split the diagonal total rate matrix $\mathbf{K}^{tot}$ into two components, $\mathbf{K}^{tot} = \mathbf{K}^{tot,0} + \nu_0 \mathbb{I}$, meaning $\mathbf{K}^{tot,0} - \mathbf{K}$ is now a standard rate matrix with no absorbing rates. To compare with previous work that starts directly with $\mathbf{K}^{tot,0} - \mathbf{K}$, we rewrite the estimators in terms of $\mathbf{W}^+ = \mathbf{K}^{tot}\mathbf{G}_+$, where $\mathbf{W}^+$ will become the pseudoinverse of $\mathbf{K}^{tot,0} - \mathbf{K}$ in the limit $\nu_0 \to 0$. Our choice of notation follows the recent work by Trinkle[4]. Omitting the full expansion in orders of $\nu_0 = 1/\langle \tau_{res} \rangle \ll 1$ for brevity, the expressions for the transport coefficients in the limit of comprehensive sampling then read

$$\boldsymbol{\mu}(\mathbf{k}^u) \to \mathbf{1}\overline{\mathbf{K} \otimes \mathbf{d}}\boldsymbol{\pi} + \mathcal{O}(\nu_0), \quad \widetilde{\mathbf{D}}(\mathbf{k}^u) \to \mathbf{1}\left[\frac{1}{2}\overline{\mathbf{B} \otimes \mathbf{d} \otimes \mathbf{d}} + \overline{\mathbf{B} \otimes \mathbf{d}}\,\mathbf{W}_+\overline{\mathbf{K} \otimes \mathbf{d}}\right]\hat{\boldsymbol{\pi}} + \mathcal{O}(\nu_0), \quad (\nu_0 \to 0) \tag{24}$$

where as before $\mathbf{D} = \frac{1}{2}\left[\widetilde{\mathbf{D}} + \widetilde{\mathbf{D}}^\top\right]$. We thus see that the limiting error in the transport coefficients is expected to scale as $\nu_0 = 1/\langle \tau_{res} \rangle$, i.e. inversely with $\langle \tau_{res} \rangle$. Taking the limiting expressions, elementary algebraic manipulations show that $\mathrm{Tr}\mathbf{D} = \mathrm{Tr}\widetilde{\mathbf{D}}$ gives exactly the form for the self diffusion constant (diagonal terms in the Onsager matrix) as derived recently by Trinkle[4,6]. We also refer the reader to many other works on diffusion in periodic media[5]. We have thus shown that our transport coefficient estimators agree with existing results on diffusive transport in periodic media, and the limiting form for the error is $\mathcal{O}(1/\langle \tau_{res} \rangle)$, which we have observed empirically, as shown in figure 2.

Symmetries under detailed balance

With a steady state Boltzmann distribution $\lim_{\nu_0 \to 0}[\mathbf{K}^{tot} - \mathbf{K}]\,\boldsymbol{\pi} = \mathbf{0}$, and a Boltzmann matrix $[\boldsymbol{\Pi}]_{pq} = \delta_{pq}\pi_q$, the detailed balance condition reads

$$\mathbf{K}\boldsymbol{\Pi} = [\mathbf{K}\boldsymbol{\Pi}]^\top = \boldsymbol{\Pi}\,[\mathbf{K}]^\top, \quad \Rightarrow \quad \boldsymbol{\Pi}^{-1}\mathbf{K} = [\mathbf{K}]^\top\boldsymbol{\Pi}^{-1}, \quad \Rightarrow \quad \mathbf{W}^+\boldsymbol{\Pi} = \boldsymbol{\Pi}\left[\mathbf{W}^+\right]^\top. \tag{25}$$

We define the transpose of $\overline{\mathbf{K} \otimes \mathbf{d}}$ and $\overline{\mathbf{K} \otimes \mathbf{d} \otimes \mathbf{d}}$ as

$$\left[\overline{\mathbf{K} \otimes \mathbf{d}}^\top\right]_{qp} \equiv \sum_{l \in \mathcal{C}_{pq}} \mathrm{k}_l \mathbf{d}_l, \quad \left[\overline{\mathbf{K} \otimes \mathbf{d} \otimes \mathbf{d}}^\top\right]_{qp} \equiv \sum_{l \in \mathcal{C}_{pq}} \mathrm{k}_l \mathbf{d}_l \otimes \mathbf{d}_l. \tag{26}$$

As every jump vector $\mathbf{d}_l, l \in \delta\mathcal{C}_{pq}$ will have a one-to-one correspondence with a vector $\mathbf{d}_{l'} = -\mathbf{d}_l$ for some $l' \in \delta\mathcal{C}_{qp}$, the detailed balance symmetries imply in particular that

$$\overline{\mathbf{K} \otimes \mathbf{d}}\boldsymbol{\Pi} = -\boldsymbol{\Pi}\overline{\mathbf{K} \otimes \mathbf{d}}^\top, \quad \overline{\mathbf{K} \otimes \mathbf{d} \otimes \mathbf{d}}\boldsymbol{\Pi} = \boldsymbol{\Pi}\overline{\mathbf{K} \otimes \mathbf{d} \otimes \mathbf{d}}^\top, . \tag{27}$$

These relations imply that the drift vector is always zero in equilibrium- using $\boldsymbol{\pi} = \boldsymbol{\Pi}\mathbf{1}^\top$ and the fact that $\mathbf{1}\mathbf{M}\mathbf{1}^\top = \mathbf{1}\mathbf{M}^\top\mathbf{1}^\top$ for any matrix $\mathbf{M}$ we have

$$\boldsymbol{\mu} = \mathbf{1}\overline{\mathbf{K} \otimes \mathbf{d}}\boldsymbol{\Pi}\mathbf{1}^\top = \mathbf{1}\left[\overline{\mathbf{K} \otimes \mathbf{d}}\boldsymbol{\Pi}\right]^\top\mathbf{1}^\top = -\mathbf{1}\overline{\mathbf{K} \otimes \mathbf{d}}\boldsymbol{\Pi}\mathbf{1}^\top = -\boldsymbol{\mu}, \quad \Rightarrow \boldsymbol{\mu} = \mathbf{0}. \tag{28}$$

We now split $\widetilde{\mathbf{D}}$ into uncorrelated and correlated terms as

$$\widetilde{\mathbf{D}} = \widetilde{\mathbf{D}}_{uc} + \widetilde{\mathbf{D}}_c, \quad \widetilde{\mathbf{D}}_{uc} = \frac{1}{2}\mathbf{1}\overline{\mathbf{K} \otimes \mathbf{d} \otimes \mathbf{d}}\boldsymbol{\Pi}\mathbf{1}^\top, \quad \widetilde{\mathbf{D}}_c = \mathbf{1}\overline{\mathbf{K} \otimes \mathbf{d}}\,\mathbf{W}^+\overline{\mathbf{K} \otimes \mathbf{d}}\boldsymbol{\Pi}\mathbf{1}^\top. \tag{29}$$

As all rates are positive, it is clear that $\mathrm{Tr}\widetilde{\mathbf{D}}_{uc} > 0$. For the correlated contribution we first use the symmetry of $\mathbf{W}^+$ to write

$$\boldsymbol{\Pi}^{-1}\mathbf{W}^+ = \left[\boldsymbol{\Pi}^{-1}\mathbf{W}^+\right]^\top \equiv \widetilde{\mathbf{W}}^\top\widetilde{\mathbf{W}} \tag{30}$$

where we introduce $\widetilde{\mathbf{W}}$ to write the symmetric matrix $\boldsymbol{\Pi}^{-1}\mathbf{W}^+$ as $\widetilde{\mathbf{W}}^\top\widetilde{\mathbf{W}}$. We then use detailed balance symmetries to rewrite the correlated contribution as

$$\left[\widetilde{\mathbf{D}}_c\right]_{\alpha\beta} = -\sum_p \mathrm{A}_{p\alpha}\mathrm{A}_{p\beta}, \quad \mathbf{A} = \widetilde{\mathbf{W}}\overline{\mathbf{K} \otimes \mathbf{d}}\boldsymbol{\Pi}\mathbf{1}^\top = -\widetilde{\mathbf{W}}\boldsymbol{\Pi}\overline{\mathbf{K} \otimes \mathbf{d}}^\top\mathbf{1}^\top, \quad \Rightarrow \left[\widetilde{\mathbf{D}}_c\right]_{\alpha\alpha} = -\sum_p \mathrm{A}_{p\alpha}^2 < 0, \quad \Rightarrow \mathrm{Tr}\widetilde{\mathbf{D}}_c < 0, \tag{31}$$

which shows that the correlated contribution always reduces the diffusivity for a system obeying detailed balance.

SUPPLEMENTARY NOTE 5: SENSITIVITY OF THE DIFFUSIVITY TENSOR TO ADDITIONAL SAMPLING DATA

Under additional sampling, the unknown rates $\mathbf{k}^u$ will decrease, additional transition rates will be found, and new states may be discovered. We first focus on the case where no new state is discovered, only additional rates, and assume the QSD $\hat{\boldsymbol{\pi}}^{QSD} \simeq \hat{\boldsymbol{\pi}}$ is unchanged. The rate matrix perturbations can be grouped into additional sets of connections $\mathcal{C}_{pq} \to \mathcal{C}_{pq} + \delta\mathcal{C}_{pq}$ and changes in the unknown rates $\delta \mathrm{k}^u$, satisfying

$$[\delta\mathbf{K}]_{pq} = \sum_{l \in \delta\mathcal{C}_{pq}} \mathrm{k}_l > 0, \quad [\delta\mathbf{k}^u]_p \leq -[\mathbf{1}\delta\mathbf{K}]_p, [\delta\mathbf{k}^u]_p \leq -[\mathbf{k}^u]_p, \quad \Rightarrow \quad \left[\delta\mathbf{K}^{tot}\right]_{pp} \leq 0, \tag{32}$$

which assumes that the unknown rate estimates $\mathbf{k}^u$ are accurate and therefore an upper bound, meaning that the total additional escape rates $\mathbf{1}\delta\mathbf{K}$ from each state cannot exceed the negative change in unknown rates $-\delta\mathbf{k}^u$. To generate candidate $\delta\mathbf{K}$ we first define the symmetric matrix $\delta\widetilde{\mathbf{K}}$ which satisfies

$$\delta\widetilde{\mathbf{K}} = \delta\widetilde{\mathbf{K}}^\top, \quad \delta\widetilde{\mathbf{K}} = \delta\mathbf{K}\boldsymbol{\Pi}, \quad \left[\mathbf{1}\delta\widetilde{\mathbf{K}}\right]_p \leq [\boldsymbol{\Pi}\mathbf{k}^u]_p \tag{33}$$

The Monte Carlo procedure discussed in the main text created candidate $\delta\mathbf{K}$ in the following procedure-

1. $\mathcal{L}$ = randomly shuffled list of pairs p, q, where $p \leq q$
2. Iterate through $\mathcal{L}$, filling $\delta\widetilde{\mathrm{K}}_{pq} = \delta\widetilde{\mathrm{K}}_{qp}$ with a random number drawn uniformly from $[0, \min(\mathrm{k}^u_q\pi_q - \sum_p \widetilde{\mathrm{K}}_{pq}, \mathrm{k}^u_p\pi_p - \sum_q \widetilde{\mathrm{K}}_{qp}))$, to ensure the above inequality is satisfied
3. Pick a random (possibly zero) lattice vector for each new transition add this to the difference in intercell positions to form $\mathbf{d}_{pq} = -\mathbf{d}_{qp}$ for each new transition
4. Add $\delta\mathbf{K} = \delta\widetilde{\mathbf{K}}\boldsymbol{\Pi}^{-1}$ to existing $\mathbf{K}$, remove $\mathbf{1}\delta\mathbf{K}$ from $\mathbf{k}^u$ then calculate new $\mathbf{D} + \delta\mathbf{D}$
5. Repeat until convergence

For the example calculations presented in the main text we found 200-400 samples per temperature were adequate.

Time independence of the Kullback-Leibler divergence

The fundamental solution $\rho(\mathbf{x}, t)$ to the three dimensional diffusion equation $\partial_t\rho = (\boldsymbol{\nabla}\mathbf{D}\boldsymbol{\nabla})\rho$ with initial condition $\rho(\mathbf{x}, 0) = \delta(\mathbf{x})$ reads

$$\rho(\mathbf{x}, t) = \exp\left(-\mathbf{x}^\top\mathbf{D}^{-1}\mathbf{x}/(4t)\right)/\sqrt{(4\pi t)^3 \mathrm{Det}(\mathbf{D})}, \tag{34}$$

giving a Kullback-Leibler divergence of

$$\mathcal{R}_\pm = \int \mathrm{d}\mathbf{x}\rho(\mathbf{x}, t) \ln\left|\frac{\rho(\mathbf{x}, t)}{\rho_\pm(\mathbf{x}, t)}\right| \tag{35}$$

$$= -\ln\sqrt{\mathrm{Det}\left(\mathbf{D}^{-1}\mathbf{D}_\pm\right)} + \int \mathrm{d}\mathbf{x}\rho(\mathbf{x}, t)\mathbf{x}^\top\left(\mathbf{D}_\pm^{-1} - \mathbf{D}^{-1}\right)\mathbf{x}/(4t) \tag{36}$$

$$= -\ln\sqrt{\mathrm{Det}\left(\mathbf{D}^{-1}\mathbf{D}_\pm\right)} + \frac{1}{2}\mathrm{Tr}\left(\mathbf{D}^{-1}\mathbf{D}_\pm\right) - \frac{3}{2}, \tag{37}$$

which is manifestly time independent.

Empirical study of convergence behavior

In figure 2 we illustrate how our estimate for the diffusion constant and convergence measure $\delta\mathcal{R}$ behave with increasing expected residence time $\langle\tau_{res}\rangle$ for a simple one-dimensional system with two irreducible states and two irreducible transitions. After the discovery of the lower barrier $1 \to 2 \to 1$ pathway, a diffusion constant is estimated along with the Kullback-Leibler spread $\delta\mathcal{R}$ using the Monte Carlo procedure. The expected diffusivity and MC-UQ

bounds change sharply when the direct $1 \to 1$ path is discovered, leading to a sharp jump in the predicted diffusivity. However, this is well within the MC-UQ bounds. The convergence measure $\delta\mathcal{R}$ is monotonically decreasing, with a convergence rate $\delta\mathcal{R} \to c/\langle\tau_{\text{res}}\rangle$, as expected from the analysis in . As a result, when $\delta\mathcal{R}$ is small we can be confident that our estimate of the diffusivity of the discovered states is accurate to an error of order $\mathcal{O}(1/\langle\tau_{\text{res}}\rangle)$, as discussed in the main text.

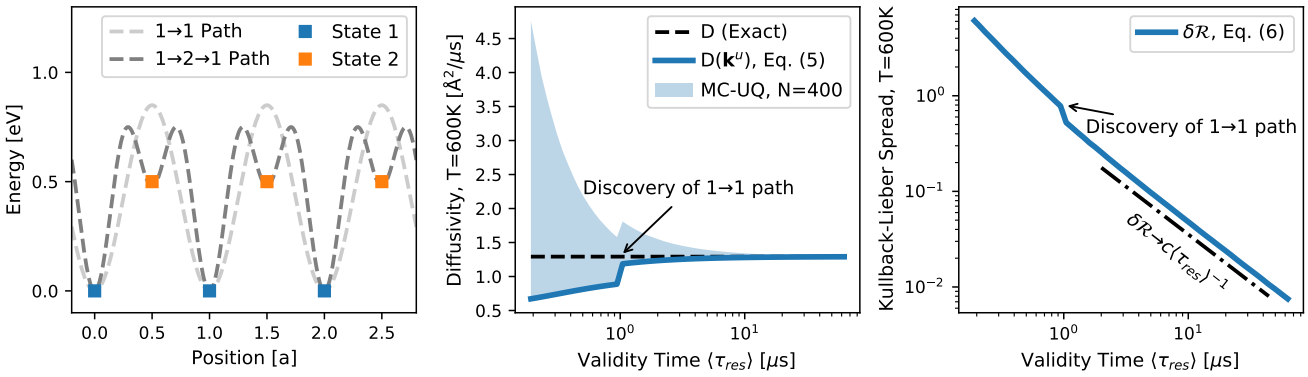

FIG. 2. Supplementary Figure 2: Demonstration of diffusivity convergence with increasing computational effort for a simple one dimensional periodic system. The validity time is used as a proxy as it is monotonically increasing. Left: Energy landscape of three unit cells, with the two irreducible states and two irreducible transitions shown. Center: The exact diffusivity D along with estimates $\mathrm{D}(\mathbf{k}^u)$ (Eq. (5) in the main text) and error bounds from our Monte Carlo procedure with 400 samples. Right: The Kullback-Leibler spread $\delta\mathcal{R}$, Eq. (6) in the main text, with increasing validity time. The convergence rate of $1/\langle\tau_{\text{res}}\rangle$ can be seen.

SUPPLEMENTARY REFERENCES

* swinburne@cinam.univ-mrs.fr

[1] T. D. Swinburne and D. Perez, Phys. Rev. Materials **2**, 053802 (2018).
[2] D. J. Wales, *Energy Landscapes*, edited by C. U. Press (Cambridge, 2003).
[3] C. Le Bris, T. Lelievre, M. Luskin, and D. Perez, Monte Carlo Methods and Applications **18**, 119 (2012).
[4] D. R. Trinkle, Physical review letters **121**, 235901 (2018).
[5] U. Landman and M. F. Shlesinger, Physical Review B **19**, 6207 (1979).
[6] See equation S15 of the supplementary material for the most direct comparison.