Microcanonical coarse-graining of the kinetic Ising model

Daniel Sigg,[1,a)] Vincent Voelz,[2] and Vincenzo Carnevale [3,b)]

[1]*dPET, Spokane, WA 99223, USA*

[2]*Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, USA*

[3]*Institute for Computational Molecular Science, College of Science and Technology, Temple University, Philadelphia, Pennsylvania 19122, USA*

ABSTRACT

In theories of escape from a metastable state, the choice of collective variables is paramount. Here we explore the use of microcanonical variables to calculate the transition rate constant across a free energy bottleneck in a finite kinetic Ising model. The microcanonical density of states accurately maintains equilibrium distributions, but projecting the full dynamics onto the microcanonical ensemble does not necessarily preserve accurate two-state kinetics. We found that rate constants calculated from coarse-grained diffusion landscapes were consistently 50% larger than brute force estimates, across environmental constraints. We concluded that, to within a constant correction factor, microcanonical coarse-graining provided a complete solution of phenomenological two-state kinetics as a function of temperature and field strength, and that the systematic error between observed and expected values could be attributed to memory effects. Committor analysis failed to support the alternative hypothesis that poor choice of collective variables led to the erroneous rate. The correction factor was empirically related to the slow reactive component of the energy barrier diffusion coefficient.

[a)] Electronic mail: dansigg@gmail.com

[b)] Electronic mail: vincenzo.carnevale@temple.edu

I. INTRODUCTION

Here we broaden the solution of metastable decay rates to include a range of environmental conditions by performing a one-time projection of the full system dynamics onto the microcanonical ensemble. Metastability in macromolecules occurs naturally in enzymatic reactions and can also be found in manufactured systems such as MOSFET transistors (Kandiah et al., 1989). The dynamics of a bistable system is characterized by random transitions between reactant and product basins, with proportionally little time spent crossing the dividing free energy barrier, as bottleneck crossings are "rare" events. An example of a tunable multistable system in biology is the ion channel, a transistor-like membrane protein (Sigworth, 2003) whose pore component fluctuates ("gates") between closed and open states according to environmental factors such as temperature (T) and voltage (Φ). Generally speaking, enzymatic proteins contain on the order of 10^2 to 10^3 interacting residues, which classifies them as "small" from a thermodynamic standpoint (Hill, 1962). Recent efforts developed a diffusion model of ion channel gating in the (T, Φ) ensemble by coarse-graining the very large configuration space $\mathbf{x}$ spanning activation onto a more manageable set of collective variables (Harpole and Delemotte, 2017). This yielded a discrete master equation or diffusion landscape that was used to calculate transition rate constants (Sigg et al., 1999; Delemotte et al., 2017). These solutions are incomplete in the sense that they must be repeated for new values of T and Φ. In theory one could acquire a "complete" solution by projecting the full molecular dynamics onto the microcanonical ensemble in the form of statistical frequencies as a function of energy (E) and charge (q), and then Laplace-transforming to generate (T, Φ)-dependent rate constants. The gating charge q assumes the role of order parameter by distinguishing between closed (reactant) and open (product) states. A relevant question is whether q is a "good" reaction coordinate if the full system dynamics is projected onto a diffusion landscape that consists of a potential of mean force $W(q, T, \Phi)$ and diffusion coefficient $D(q, T, \Phi)$ (Peters et al., 2013). A straightforward answer to this question is achieved by computing the forward (or backward) rate constant predicted by the coarse-grained diffusion landscape and compare it to experimental rates obtained from brute-force simulations.

For this study we considered a toy lattice model with broken symmetry states, specifically an electrical isomorph of the traditional two-dimensional Ising model. The existence of an upper bound in both E and q values allows the complete state space to be projected onto a finite microcanonical grid (Ferrenberg and Swendsen, 1988;

Pleimling and Behringer, 2005; Hűller and Pleimling, 2002; Ray and Freléchoz, 1996). We restricted the lattice length L to the following set of even values: 16, 18, 20, and 22. The number of lattice particles $N = L^2$ is comparable to the number of residues in a biological macromolecule. The small N permitted a large number of unbiased ("brute force") transition events to be simulated in a short time. Previous studies (Brendel et al., 2005; Pan and Chandler, 2004; Peters and Trout, 2006; Moritz et al., 2017) employed larger systems (L = 24 to 100) and lower temperatures (T/T_c = 0.6 to 0.8, compared to > 0.9 in this study) in order to compare Ising kinetics with classical nucleation theory. But traditional variables such as nucleation volume, surface area, and more recently anisotropy (Leyssale et al., 2007) are not well defined if the critical cluster size exceeds the system boundaries, as may occur in small systems. Our focus instead was to compute rates as a function of T and Φ for small systems with modest free energy barriers (5 to 9 kT). In small systems, susceptibilities are softened, resembling response curves seen in regulatory proteins, while still maintaining phenomenologically precise two-state kinetics.

II. ELECTRICAL ISING MODEL

A. Configurational structure and dynamics

The enthalpy H of the proposed Ising model consists of a configurational energy E, onto which an electrical work term $-q\Phi$ is added in place of the usual pressure-volume work.

$$H = \varepsilon \sum_{\langle i,j \rangle} e_i \otimes e_j - \left(\delta q \sum_i e_i \right) \Phi , \tag{1}$$

The energy term sums over nearest-neighbor interactions $\langle i, j \rangle$ between "gating" particles with values e = 0, 1. Adjacent particles with dissimilar e values (symbolized by the exclusive-or operator $\otimes$) increase E through a positive interaction energy ε. With periodic boundary conditions, E increases in increments of 2ε from 0 to $2N\varepsilon$. The "gating" charge q defined in parenthesis increases by δq for activating particles, and $-\delta q$ for backward transitions. The value of q ranges from 0 to $N\delta q$ in increments of δq. It was sometimes convenient to add an umbrella potential $U(q) = c(q - q_o)^2$ to Eq. 1 in order to constrain the system around the charge displacement q_o. Because the electrical

Ising model is isomorphic to the traditional magnetic Ising model, all results discussed here apply to the traditional model. We note that the interaction term ε has twice the value of the traditional J factor.

We chose to work with the electric Ising model because of its dynamical similarities to gated ion channels whose central ion-conducting pore units demonstrate weak voltage-sensing behavior even in the absence of canonical voltage sensors (Cordero-Morales et al., 2006; Kurata et al., 2010). The mechanism for intrinsic pore voltage sensitivity is incompletely known, but presumably it involves long-range correlations in the secondary structure. Figure 1 shows how a small "pore" region in the Ising grid comprised of four selectivity particles generates large-scale fluctuations in the "ionic current" qualitatively similar to those from real single channel recordings, including a characteristic voltage- and temperature-dependence. In this study, we focused on "gating" charge q because it is a proper thermodynamic variable, whereas the artificial ionic current i serves as a marker for activation and is not energetically linked to gating. We note however, that q and i generate identical random telegraph signals.

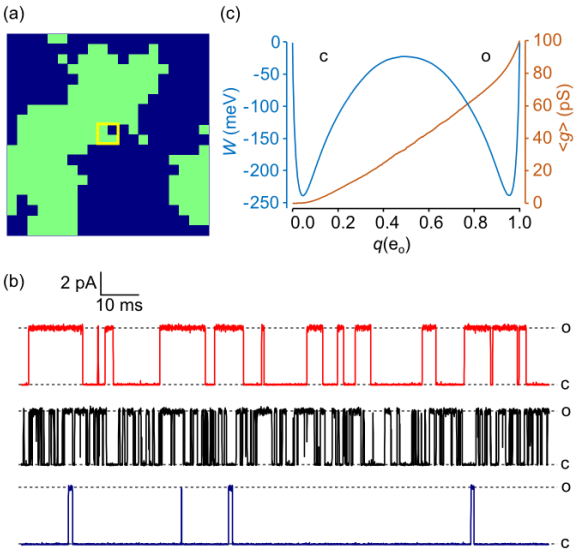

FIG. 1. (a) 20 x 20 electrical Ising model. Activated particles are colored green. An "ionic" current $i(t)$ was generated by specifying a central 2 x 2 "pore" region (yellow square) whose instantaneous conductance g is the sum of four 25 pS partial conductances. (b) Continuous-time Monte Carlo simulations of the Ising "pore" with ionic current governed by Ohm's law: $i = g(V - V_{rev})$. Conductance levels transition between closed (c) and open (o) states. The red trace (top) is the reference model with

Φ_{rev} (reversal potential) = –60 mV. The current was digitally filtered as previously described (Sigg, 2013), with cutoff frequency f_c = 10 kHz and sampling rate f_s = 100 kHz. The black (middle) trace demonstrates a higher frequency of transitions as temperature T is increased from 17 °C to 27 °C, and the closed state is favored in the blue (bottom) trace as Φ is decreased from 0 mV to -100 mV. (c) 1D free energy landscape $W(q)$ and mean conductance $\langle g(q) \rangle$ for the reference model obtained as time-averaged quantities from a 0.4 sec trajectory.

In order to study Ising dynamics we needed to define rate constants between microstates. These are constrained by detailed balance:

$$\frac{\alpha_i}{\beta_i} = \exp\left(-\frac{H(e_i = 1) - H(e_i = 0)}{kT} \right), \tag{2}$$

where α_i and β_i are the forward and backward rate constants associated with activation of the i^{th} particle. A variety of expressions for α_i and β_i are compatible with Eq. 2. Our choice was based on activated kinetics:

$$\alpha_{u(i)} = \nu \exp\left(\frac{-\Delta\varepsilon_{u(i)}}{2kT} \right), \tag{3a}$$

$$\beta_{u(i)} = \nu \exp\left(\frac{\Delta\varepsilon_{u(i)}}{2kT} \right) \quad , \tag{3b}$$

where the energy of activation $\Delta\varepsilon$ increases linearly with the number of activated neighbor particles u, the voltage Φ, and optionally the umbrella potential U, as follows:

$$\Delta\varepsilon_u = 2(2-u)\varepsilon - \Phi\delta q + U(q+\delta q) - U(q) \, . \tag{4}$$

The pre-exponential factor ν was assumed to be temperature-independent. This placed an upper limit on diffusion kinetics for very high temperatures.

Our choice of rate constants is not unique. A different choice attributed to Glauber (Glauber, 1963) prevents microscopic rates from exceeding ν after applying a large external field. This avoids undersampling fast transitions

when using a fixed simulation time step. We avoided undersampling by employing a continuous-time Monte Carlo algorithm developed by Gillespie (Gillespie, 1977) that generated real-valued transition intervals.

We chose as our reference model $L = 20$ with the following parameters: $T = 22$ °C; $\Phi = 0$ mV; $\varepsilon = 24$ meV; $q_{max} = 1$ e_o; and $\nu = 5 \times 10^4$ kHz. Except when stated, calculations were performed using the reference model. At 22 °C, $kT = 25.43$ meV.

B. Coarse-graining

The Ising model and coarse-grained models derived from it are each governed by a master equation with a complete set of (T, Φ)-dependent rate constants. The kinetics of the full 2^N configurational space **x** is determined by elementary forward and backward rate constants $\{\alpha_{u(i)}, \beta_{u(i)}\}$. The first level of coarse-graining is the 2D model residing in the $(N+1)^2$ microcanonical space (E, q). The 2D rate constants are $\{a_{kk'}, b_{kk'}\}$. The 1D model is the projection of the 2D master equation onto the $(N+1)$ reaction coordinate q with rate constants $\{a_m, b_m\}$. Alternatively the 1D dynamics are described by a continuous diffusion landscape $\{W(q), D(q)\}$, which may be condensed to a "sawtooth" discrete state model. Finally, the phenomenological two-state model contains forward and backward rate constants k and k_{-1}. The full Ising kinetics are "microscopic". The 1D and 2D representations are "diffusive", and the sawtooth and two-state models are "mesoscopic" in keeping with usage by some authors to describe the stochastic kinetics of a single macromolecule. (Qian, 2001; Kjelstrup et al., 2008). The full Ising space is intractably large and cannot be solved except by Monte Carlo methods, but the diffusive and mesoscopic models lend themselves to numerical and analytical solutions of their master equations.

The coarse-graining procedure that creates diffusion states is illustrated in Fig. 2. When discussing diffusion, we alternate between continuous (q, E) and discrete (m, n) coordinates, where $q = m\delta q$ and $E = 2n\varepsilon$. Discrete dynamics are governed by a master equation, whereas the continuous model is handled by the Smoluchowski equation. For lattice dimensions considered here, discrete and continuous approaches yield similar outcomes. Coordinates (m, n) are easily determined for any microconfiguration. The order parameter m is the total number of activated particles, and the energy integer n is equal to $2m - b$, where b is the number of shared borders between activated m particles.

The (m, n) states form a triangle on the N x N microcanonical grid. The concavity at the base of the triangle has height L and gives rise to the characteristic symmetry-breaking properties of the 2D Ising model.

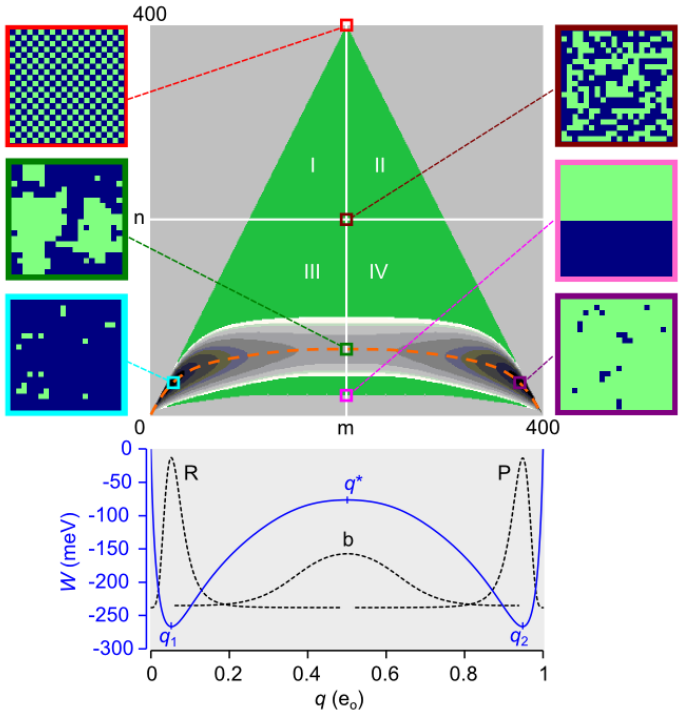

FIG. 2. Equilibrium course-graining of the 20 x 20 Ising system (N = 400). The green triangle is the set of microcanonical state on the (m, n) grid (73,448 states). Positive ε favors the lower portion (quadrants III and IV) of the triangle and positive Φ favors the right half (quadrants II and IV). The free energy landscape (22,149 states) constrained by W_{max} = 150 meV is shown as the 2D contour map near the base of the triangle. The vertical span (spatial bandwidth) of the separatrix (m = N/2) is 68 cells. The dashed orange line is the mean energy $\langle E \rangle$. Representative microscopic configurations shown at the edges of the main plot are the following: maximum energy (red); maximum entropy (maroon); saddle point (green); minimum transition state energy (pink); and the minimum free energies in metastable states R (aqua) and state P (purple). The projected 1D free energy landscape $W(q)$ is plotted in the lower graph. The dotted lines are equilibrium probabilities for the R, P basins and the inverted barrier b. Local energy extrema are q_1, q^* (the transition state), and q_2.

The free energy $W(E, q, T, \Phi)$ depends on the density of states $\Omega(E, q)$ according to the formula:

$$W_{mn}(T,\Phi) = 2n\varepsilon - kT\ln\Omega_{mn} - (m\delta q)\Phi \; . \qquad (5)$$

Eq. 5 is the small-system equivalent to the thermodynamic relation $W = E - TS - q\Phi$. Because topologically the 2D lattice subject to periodic boundary conditions is a torus, we avoid extra terms dealing with end effects. Positive ε populates the states in the lower portion of the free energy triangle (Fig. 2). This corresponds to ferromagnetism/diamagnetism in the traditional Ising model. Anti-ferromagnetism occurs at negative ε, which populates the upper portion of the triangle. In order to limit the accessible (m, n) states to a manageable number, we chose a cut-off value W_{max} for the free energy. This results in a binary 2D landscape with reactant R and product P basins separated by a broad, narrow and slightly curved bottleneck region b. We projected the 2D landscape $W_{nm}(T, \Phi)$ onto the 1D potential of mean force $W_m(T, \Phi)$ by summing over n:

$$\exp\left(-\frac{W_m}{kT}\right) = \sum_n \exp\left(-\frac{W_{mn}}{kT}\right). \tag{6}$$

C. Thermodynamics

The diffusion models (1D and 2D) faithfully preserve system thermodynamics (Fig. 3). From the 2D free energy landscape we calculated the environmental sensitivities of mean charge displacement $\langle q \rangle$ and heat capacities C_Φ and C_q. These quantities are not subject to critical discontinuities experienced by the infinite-particle Ising model (Yang, 1952; Onsager, 1944).

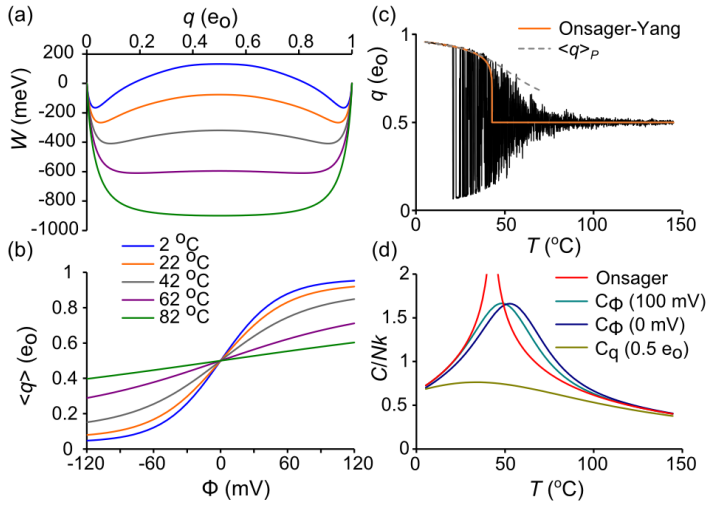

FIG. 3. Ising model thermodynamics. (a) $W(q)$ for different temperatures. The transition from bistable to centrally stable landscape roughly occurs at 73 °C. (b) $\langle q \rangle$ as a function of T and Φ. (c) $q(t)$ trajectory in response to 800 ms temperature ramp (5 °C to 145 °C) (ref). Filtering: f_s = 10 kHz, f_c = 1 kHz. The red line is the Onsager-Yang solution (Yang, 1952) and the dashed curve is $\langle q \rangle_P$. (d) Heat capacities derived from $C/k = \beta^2 \langle (\Delta E)^2 \rangle$. The Onsager solution (Onsager, 1944) demonstrates critical discontinuity at $T_c = \varepsilon / k \ln(1 + 2^{1/2})$, which corresponds to 42.84 °C in the reference model.

D. Diffusive rate constants

The 2D rate constants $\{a_{kk'}, b_{k'k}\}$ connecting states $k = (m, n)$ and $k' = (m', n')$ were computed in the (T, Φ) ensemble as the product of two numbers, the microscopic rate (Eq. 3), which contains the T- and Φ-dependence, and the microcanonical branching coefficient $\langle r_{kk'} \rangle$, which averages the number of available pathways for microscopic transitions between k and k′ (Fig. 4). Each $k \leftrightarrow (m, n)$ state contributes a maximum of five forward $(m + 1)$ transitions corresponding to $u = 0,1,\ldots,4$:

$$a_{kk'}(T, \Phi) = \langle r_{kk'} \rangle \alpha_{u(k)}(T, \Phi) . \tag{7}$$

There are additionally five backward rate constants $b_{k'k}$ in the $(m - 1)$ direction, for a maximum total of ten cardinal rates.

Microcanonical coarse-graining yields an exact kinetic solution only in the case of the mathematically trivial $L = 2$ Ising model, where branching coefficients are strict integers. For $L > 2$, there is an inevitable corruption of kinetic information, though microcanonical detailed balance, defined by $\Omega_k a_{kk'} = \Omega_{k'} b_{k'k}$, is preserved.

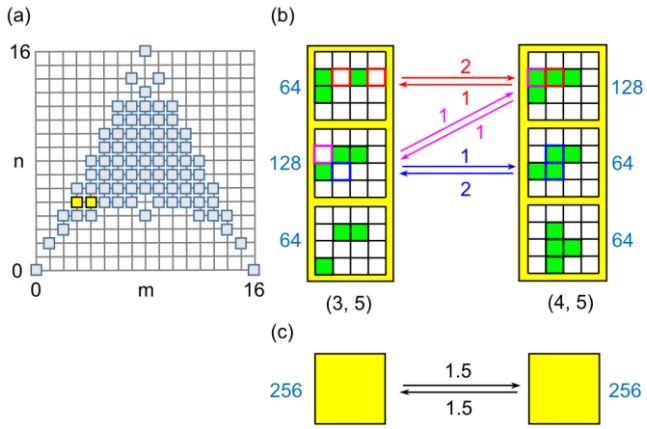

FIG. 4. (a) Microcanonical state space for the 4 x 4 Ising model. The (3, 5) and (4, 5) states are in yellow. (b) Microstates within (3, 5) and (4, 5) diffusion states. Green cells are activated. The number of microstates (Ω) are in shown in blue. Arrows indicate allowed transitions with color references to the state(s) of origin and local branching coefficient. (c) The coarse-grained (3, 5) $\rightarrow$ (4, 5) transition after statistical averaging, with mean branching coefficients $\langle r \rangle$ in black.

To determine how loss of information affects large-scale relaxation, 2D rate constants were acquired using Metropolis Monte Carlo (MMC). Applying traditional MMC for $T \rightarrow \infty$, $\Phi = 0$ allows every trial move to be accepted, in principle yielding the correct branching coefficients. However, the large entropy gradient makes this algorithm impractical for all but the smallest systems, so a bias potential ω_k was applied to "flatten" the gradient and allow uniform sampling over all (m, n) states. The 1/t variant of the Wang-Landau algorithm (Belardinelli and Pereyra, 2007) with endpoint refinement parameter $F_{final} = 10^{-6}$ was used to rapidly estimate $\ln \omega_k$. MMC was then implemented using the following procedure: a random particle is chosen for a trial flip; then, after determining the (m, n) coordinates of the current k state and the trial k' state, a uniform random number $0 < r_n < 1$ is drawn and the following acceptance criterion for the $k \rightarrow k'$ transition is applied:

$$r_n \leq \frac{\omega_k}{\omega_{k'}}, \tag{8}$$

If successful, the particle is flipped. Time is incremented regardless of trial outcome. For each discrete time point s, N total transitions are divided up among the ten cardinal rate constants a_u and b_u. For example, if twelve activated particles contribute to the $u = 3$ backward transition (m, n) $\rightarrow$ ($m - 1$, $n + 1$), then the branching coefficient equals 12

and the instantaneous rate $b_3(s)$ is $12\beta_3$, where β_3 is the backward rate (Eq. 3b). The starting point was the maximum entropy region around (N/2, N/2). After an equilibration period of 2 x 10^6 time increments, 2D rate constants were time-averaged over the next 2 x 10^{10} trials using the following formula for forward rates:

$$\left\langle a_{u(k)} \right\rangle = \frac{\sum_s \omega_k(s) a_{u(k)}(s)}{\sum_s \omega_k(s)} \alpha_{u(k)}\,, \tag{9}$$

with an analogous formula for backward rates $\langle b_{u(k)} \rangle$. The weighted sum that precedes $\alpha_{u(k)}$ in the equation is the branching coefficient $\langle r_{kk'} \rangle$. The microcanonical density of states was obtained from:

$$\ln \Omega_k = N \ln(2) + \ln \frac{\sum_s \omega_k(s)}{\sum_{k,s} \omega_k(s)}\,. \tag{10}$$

The density of states Ω_k and forward and reverse branching coefficients $\langle r_{u(k)} \rangle_a$ and $\langle r_{u(k)} \rangle_b$ were stored as N x N matrices ln$\mathbf{\Omega}$, $\mathbf{A}_u$, and $\mathbf{B}_u$ (Fig. 5). These eleven matrices are not independent since, putting aside numerical error, the entire group can be constructed from any three of its members, for example $\mathbf{A}_0$, $\mathbf{A}_1$, and $\mathbf{A}_3$. The following relations apply:

$$B_{4-u}(N-m,n) = A_u(m,n) \tag{11a}$$

$$\sum_u B_u(m,n) = m \tag{11b}$$

$$\sum_u u A_u(m,n) = 2n \tag{11c}$$

$$A_u(m,n)\Omega(m,n) = B_u(m+1, n+2-u)\Omega(m+1, n+2-u) \tag{11d}$$

The first line expresses the mirror symmetry intrinsic to the Ising model between $\mathbf{A}_u$ and $\mathbf{B}_{4-u}$. The last line describes detailed balance between diametrically opposed matrices $\mathbf{A}_u$ and $\mathbf{B}_u$. The middle expressions were discovered empirically.

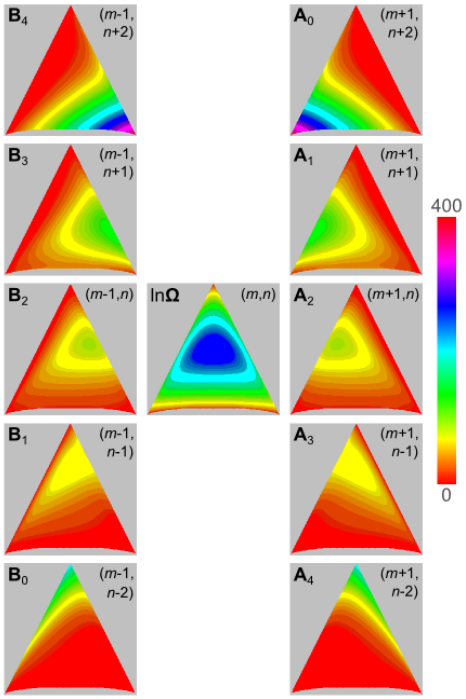

FIG. 5. Branching coefficients for the ten cardinal directions are stored in N x N matrices $\mathbf{A}_u$ and $\mathbf{B}_u$. N = 400. The center matrix is the density of states ln$\mathbf{\Omega}$. The color scale applies throughout.

The 1D rate constants a_m and b_m were obtained in the (T, Φ) ensemble by projecting the 2D rate constants onto the q-axis using Boltzmann averaging. In 1D, detailed balance is described by:

$$\frac{a_m}{b_{m+1}} = \exp\left(-\frac{W_{m+1} - W_m}{kT}\right). \tag{12}$$

For small increments δq, the 1D master equation can be approximated by a continuum diffusion equation (Van Kampen, 1992). The correspondence between discrete {a_m, b_m} and diffusive {W_m, D_m} variables arises from the Kramers-Moyal expansion of the master equation. The diffusion coefficient equals the second Kramers-Moyal coefficient:

$$D_m = \frac{1}{2}\left(a_m + b_m\right)\delta q^2. \tag{13}$$

The free energy W_m is determined by detailed balance (Eq. 12). For regions of m where a_m and b_m are roughly linear, such as near the barrier, the expansion effectively ends at second order and the Smoluchowski equation exactly describes continuous 1D diffusion:

$$\frac{\partial p(q,t)}{\partial t} = \frac{\partial}{\partial q} D(q) \exp\left(-\frac{W(q)}{kT} \right) \frac{\partial}{\partial q} \exp\left(\frac{W(q)}{kT} \right) p(q,t) \,, \tag{14}$$

where $p(q,t)$ is the gating charge probability distribution. Parallel methods for computing mesoscopic rates exist for master and Smoluchowski equations. We mostly we worked with the master equation since it is the fundamental dynamic equation for the Ising model. The diffusion equation is an approximation, but yielded similar results and plays a significant role later in the paper.

III. THE MESOSCOPIC RATE CONSTANT

A. Dwell time analysis reveals systematic error after coarse-graining

To gauge the loss of kinetic information with coarse-graining at the mesoscopic level, we needed accurate values of the transition rate k for the Ising and 1D/2D models. To this end, we used brute-force simulation to generate a long trajectory with thousands of R $\rightarrow$ P transitions and stored the reactant state dwell times. Generally speaking, dwell times are the intervals between large-scale transition events and may be subject to experimental error due to noise and bandwidth limitations (Roux and Sauvé, 1985), but by employing the continuous-time Gillespie Monte Carlo algorithm (Gillespie, 1977), we could obtain exact first passage times from q_1 to q_2. Briefly, at the beginning of a Monte Carlo step, state-dependent flip rates (α_i or β_i) were determined for each particle i, and sums of rates were grouped in the forward ($\sum \alpha_i$) and backward ($\sum \beta_i$) directions. After drawing a uniform random number r_n, the step interval was calculated using the formula $-\ln r_n/(\sum \alpha_i + \sum \beta_i)$, and a second random number determined the transitioning particle weighted by its microscopic rate constant. The process was repeated until the completion of 4 seconds of simulation time, which included thousands of transition events. The same algorithm was used to simulate transition events in 1D and 2D. The accumulated first passage times were averaged as the mean first passage time (mfpt), which we named τ. If transitions are "rare" events, we can equate k to τ^{-1} (Reimann et al., 1999).

The disparity in reactant dwell times (first passage times from q_1 to q_2) between Ising and coarse-grained model trajectories $q(t)$ is already apparent by eye for trajectories that have been filtered to simulate real experimental data (Fig. 8a, b). To quantify the disparity, distribution of first passage times were log-transformed and binned using a Sigworth-Sine plot (Sigworth and Sine, 1987). A two-state model predicts an exponential distribution of reactant dwell times. The Sigworth-Sine plot contains two features of interest: (1) it generates uniform fitting residuals; (2) the transformed distribution peaks at τ for easy visualization of the mean first passage time.

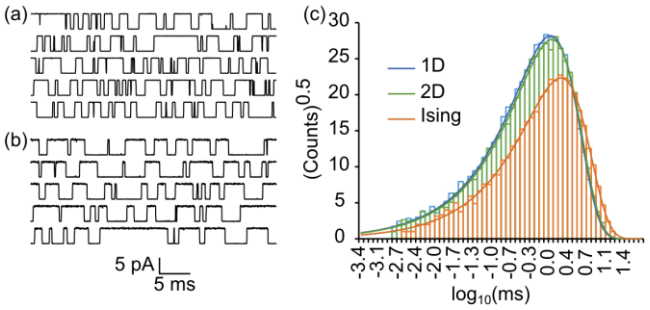

FIG. 6. Dwell time analysis. (a) "Ionic current" trajectories of the Ising model shown for visualization purposes. (b) Same as (a) except starting from 1D master equation. Filtering: f_c = 10 kHz, f_s = 100 kHz. (c) Sigworth-Sine plot of first passage times demonstrates fewer crossing events and a longer average dwell time for Ising versus 1D and 2D dynamics.

First passage times from Ising and coarse-grained trajectories were exponentially distributed across nearly four orders of magnitude in time. Fitting the theoretical distribution to the experimental dwell time histogram yielded τ_{fit}, which was very close to the mfpt τ despite binning error (Table I). The two-state model is evidently a good phenomenological descriptor of mesoscopic kinetics. The coarse-grained models (1D and 2D) generated a smaller τ than the Ising model, resulting in a roughly 50% larger rate constant k. The 2D model was only marginally more accurate than the 1D model. Thus coarse-graining, though it maintained detailed balance, failed to predict the correct rate constant. The statistics for the brute force determination of k through dwell time analysis were sufficiently accurate for the method to be considered the gold standard when applied to the full (Ising) configuration space. Other methods were judged by the ratio γ of their estimate of k to the gold standard value. Much of the remaining paper concerns itself with validating and exploring the origins of γ, which we have thus far determined to be about 3/2 if evaluating k is preceded by coarse-graining.

TABLE I. Reactant state dwell times for 4 s simulation (reference model).

	Ising	2D	1D
events	6,776	10,397	10,791
τ (ms)	1.672 ± 0.020[a]	1.145 ± 0.011[a]	1.086 ± 0.010[a]
k (kHz)	0.5982 ± 0.0073[b]	0.8735 ± 0.0085[b]	0.9210 ± 0.0088[b]
τ_{fit} (ms)	1.675	1.146	1.088

[a] s.e.(τ) = τ/events$^{1/2}$.
[b] s.e.(k) = (k/τ)s.e.(τ).

B. Numerically solving the master equation to obtain *k*

To confirm the findings from dwell time analysis, we obtained k from coarse-grained master equations using direct calculation. In matrix form, the master equation is described by $d\mathbf{p}/dt = \mathbf{pA}$, where $\mathbf{p}$(t) is the row vector of state probabilities and $\mathbf{A}$ is the rate constant matrix with diagonal entries ensuring each row sums to zero. The row-major convention is consistent with the notation a_{ij} for $i \to j$ transitions. We note that some authors use a column-order format with transposed indices for the rate constants and matrix operations in reverse order. For the two-state mesoscopic model, the matrix $\mathbf{A}$ is:

$$\mathbf{A}_{(\text{meso})} = \begin{bmatrix} -k & k \\ k_{-1} & -k_{-1} \end{bmatrix}. \tag{15}$$

The first non-zero eigenvalue of $\mathbf{A}_{(\text{meso})}$ is $\lambda_1 = -(k + k_{-1})$. If $\Phi = 0$, then $k = k_{-1}$, so each rate constant equals $-\lambda_1/2$. Even if applying a field Φ, knowing λ_1 determines k, since the ratio k/k_{-1} can be determined from detailed balance.

Unlike $\mathbf{A}_{(\text{meso})}$, rate matrices for the diffusive (1D, 2D) master equations are quite large, requiring a numerical solution. In block form, $\mathbf{A}_{(\text{diff})}$ looks like:

$$\mathbf{A}_{(\mathrm{diff})} = \begin{bmatrix} \mathbf{R} & \mathbf{C} & \mathbf{0} \\ \mathbf{D} & \mathbf{B} & \mathbf{E} \\ \mathbf{0} & \mathbf{F} & \mathbf{P} \end{bmatrix}, \tag{16}$$

where the diagonal blocks **R**, **B**, and **P** represent the reactant, barrier, and product regions, demarcated on the q-axis by q_1 and q_2, and **C**, **D**, **E**, and **F** are off-diagonal transition matrices. Though **A** is singular, **R**, **B**, and **P** are invertible since they represent open systems.

We obtained k through direct calculation from 1D and 2D master equations using eigenvalue and mfpt methods. The eigenvalue analysis computes the first non-zero eigenvalue λ_1 of $\mathbf{A}_{(\mathrm{diff})}$ and equates it to the mesoscopic form $-(k + k_{-1})$. The large spectral gap between λ_1 and higher eigenvalues (Fig. 7a) indicated good temporal separation of inter- and intra-state events, consistent with two-state kinetics (Eq. 15). For the reference model, the ratios λ_2/λ_1 were 1150 (2D) and 1260 (1D).

The mfpt method involves truncating $\mathbf{A}_{(\mathrm{diff})}$ at the absorbing boundary q_2 to create an invertible matrix $\mathbf{R}'$ made from block matrices **R**, **C**, **D**, and **B**. The mfpt τ is obtained (Van Kampen, 1992) by solving

$$\mathbf{R}'\boldsymbol{\tau} = -\mathbf{u}, \tag{17}$$

where $\boldsymbol{\tau}$ is the column vector containing $\tau(q \rightarrow q_2)$ and **u** is the unit vector. As in dwell time analysis, we equated the forward rate constant k to $\tau^{-1}(q_1 \rightarrow q_2)$. The value of k obtained this way is insensitive to the precise choice of q within the reactant basin (Figure 7b), a reliable indicator of two-state kinetics. A close relation to Eq. 17 is the following (Van Kampen, 1992):

$$\mathbf{B}\boldsymbol{\pi} = -\mathbf{E}\mathbf{u}, \tag{18}$$

where $\boldsymbol{\pi}$ is the column vector of splitting or "committor" probabilities from q located within **B** to q_2. To compute k from $\boldsymbol{\pi}$, we used the flux-over-probability expression $k = J/p_R$, where p_R is the equilibrium probability of residing in the reactant basin ($p_R = 0.5$ for $\Phi = 0$) and the flux J is evaluated for any dividing border Γ spanning the bottleneck region using (Berezhkovskii et al., 2009; Metzner et al., 2009):

$$J = \sum_{(u \to v) \in \Gamma} p_u^{eq} a_{uv} \left(\pi_v - \pi_u \right). \tag{19}$$

The mfpt and committor methods computed numerically identical values of k, underscoring the common origin of the two algorithms, which is that both Eq. 17 and Eq. 18 are related to the Kolmogorov backward equation.

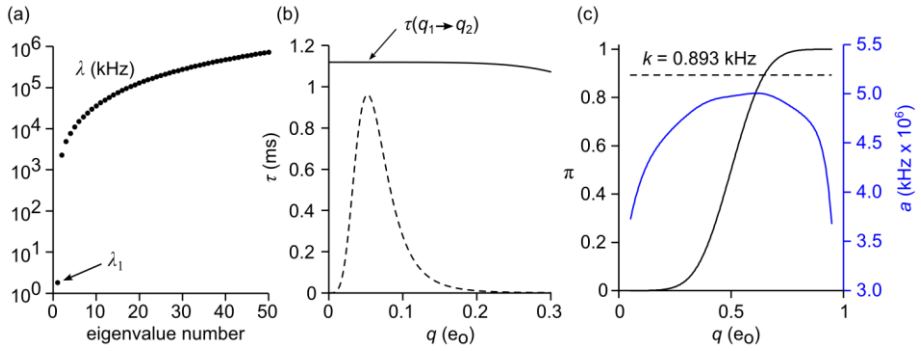

FIG. 7. Methods for calculating k from the 1D master equation. 2D-based methods are analogous. (a) Eigenvalues of **A**, demonstrating the large spectral gap between first and second eigenvalues. (b) Mean first passage time (Eq. 17). The value of $\tau(q)$ is nearly constant across the reactant basin (dashed line showing equilibrium probability distribution). (c) Committor analysis (Eqs. 18 and 19). The committor probability $\pi(q)$ is plotted as a black line and the forward rate constant $a(q)$ as a blue line. The computed value of k (dashed line) is constant for all q in the range $[q_1, q_2]$.

A computational challenge arises from the size of $\mathbf{A}_{(\text{diff})}$, particularly in the case of the 2D model. From the $(N+1)^2$ microcanonical states in 2D-space, somewhat less than half are physically accessible (Fig 2), but even these are too numerous to handle conventionally. The number of accessible states l was substantially reduced by implementing a free energy cutoff W_{max}. To avoid an artificially constraining bottleneck in the barrier region, W_{max} was assigned a value at least 150 meV greater than the saddle point energy. Converting $\mathbf{A}_{(2D)}$ to band form further increased packing efficiency. We numbered states in column order of increasing m and $n(m)$. This resulted in a maximum spatial bandwidth d equal to the length of the separatrix. The final dimensions of the banded 2D rate matrix $\mathbf{A}_{(2D)}$ were l x $2d+1$. Typical values for l and d were ~20,000 and ~70. By taking advantage of eigenvalue and linear routines designed for symmetric band matrices, k could be computed in seconds on a desktop computer. The same methods were used for 1D calculations, where $\mathbf{A}_{(1D)}$ is tridiagonal ($d = 1$) and $l = N + 1$.

The results from direct solutions of the 1D and 2D master equations, tabulated in Table II, confirmed the earlier dwell time analysis results that $\gamma \approx 3/2$. Despite the agreement between the different numerical methods, it was important to rule out any systemic biases that could falsely increase the value of *k*.

C. Error Analysis

We identified two potential sources of bias. The first is a poor choice of random number generator leading to erroneous Monte Carlo simulation (Coddington, 1994). To address the possibility of biased sampling, we tested several pseudorandom number generators: (1) Numerical Recipes© long-period L'Ecuyer generator with Bays-Durham shuffle (ran2) (Press et al., 1992); (2) Intel® MKL Library SIMD-oriented Fast Mersenne Twister (FMT) (Saito and Matsumoto, 2008); (3) Intel® MKL Library multiplicative congruential generator (MCG) (L'Ecuyer, 1999a); and (4) Intel® MKL Library combined multiple recursive generator (MRG) (L'Ecuyer, 1999b). Rate constant values from three of the four generators were in close agreement across all methods (Fig. 8). The MCG algorithm exhibited greater variability and systemic biases compared to the other three generators and is not recommended for use in our model. Most simulations employed the ran2 algorithm. In the small number of cases with large barriers where the performance of the random number generator was in doubt, repeating the simulations with the FMT and MRG algorithms confirmed the ran2 results.

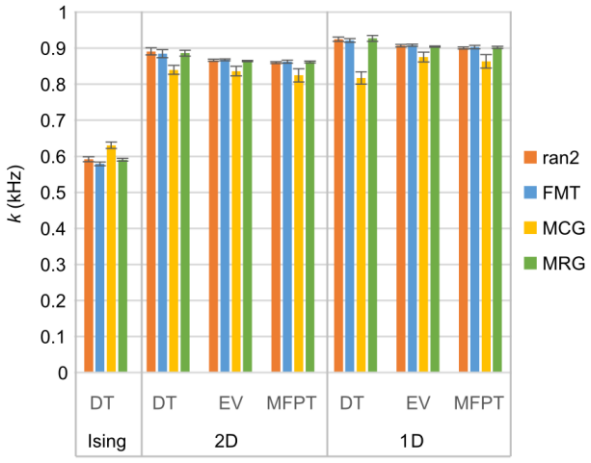

FIG. 8. Forward rate constant *k* (mean ± s.e., *n* = 6) calculated for standard model with different methods (DT = dwell time, EV = eigenvalue analysis, MFPT = mean first passage time) and as a function of pseudorandom number generator. Results were

consistent between generators except in the case of the multiplicative congruential generator (MCG), which demonstrated lower accuracy and precision compared to the other three generators.

A second source of potential error related to the coarse-graining of rate constants. Since branching coefficients were obtained from MMC simulations, a certain amount of statistical error was unavoidable. After coarse-graining, in order to re-establish detailed balance between opposing diffusive rates $a_{kk'}$ and $b_{k'k}$ in accordance with equilibrium probabilities p_k, we applied the following geometric mean-preserving rebalancing procedure:

$$a = \sqrt{a_{kk'} b_{k'k}} \sqrt{\frac{p_{k'}}{p_k}}, \tag{20a}$$

$$b = \sqrt{a_{kk'} b_{k'k}} \sqrt{\frac{p_k}{p_{k'}}}, \tag{20b}$$

where a and b are the rate constants after rebalancing. For small relative errors ε_a and ε_b in the unbalanced rates $a_{kk'}$ and $b_{k'k}$, the first order relative error in each rebalanced rate is: $(1 + \varepsilon_a + \varepsilon_b)^{1/2} - 1$. Thus systemic bias would be minimized if ε_a and ε_b are centered around zero. In practice, omitting the rebalancing step prior to dwell time analysis of the 1D model caused the value of k in one experiment to rise 17% above the expected coarse graining error γ, but this dropped to a 6% increased error if unbalanced rates were averaged from 30 experiments prior to analysis. The corresponding 2D single experiment relative error was 4%, which paradoxically increased to 7% with averaging. The findings suggest that omitting the rebalancing procedure prior to generating trajectory data increased variability in single experiments but lead to a small but consistent overestimation (6–7%) in pooled data. There was no measurable effect on the mfpt method, but omitting rebalancing did have a profound effect on 2D eigenvalue calculations, with a 10-fold increase in unbalanced k estimates, though relative error diminished to a 30% increase with averaging. The excessive error in eigenvalue analysis can be attributed to eigenvalue routines designed for symmetric matrices, which require detailed balance as a strict requirement for symmetrizing rate matrices (Frehland, 1978). The 1D master equation satisfies detailed balance by design, so its eigenvalues were error-free. To summarize, unbalanced diffusive rate constants increased k estimates by a variable amount that could be reduced by using pooled data for $a_{kk'}$ and $b_{k'k}$, though not eliminated. We note that omitting the rebalancing process did not bring

the value of k closer to the gold standard value, instead it went the opposite direction. This, and the consistency of balanced k values across all methods suggests that rebalancing is an essential step in coarse-graining, and does not account for non-unitary γ values.

D. *k* as a function of *N*, *T*, and Φ

We examined whether the coarse-graining error γ was constant across environmental factors. The relationships between k and thermodynamic constraints N, T, and Φ are predictable from changes to the 1D diffusion landscape. Reaction rate theory predicts that k should decrease exponentially with increasing barrier height $\Delta W = W_b - W_R$ and increase linearly with D_b. In Figure 9, we see that ΔW increases in proportion to N and inversely to T and Φ. The critical barrier value of the diffusion coefficient (D_b) increased substantially with rising N, less dramatically with rising T, and was fairly insensitive to Φ. The net effect was that k exponentially decreased with N and increased with both T and Φ (Fig. 10). We considered barrier heights in the range 5-9 kT. These are modest energy barriers, allowing us to obtain a statistically significant number of simulated transition events and achieve accurate numerics while maintaining strict two-state dynamics.

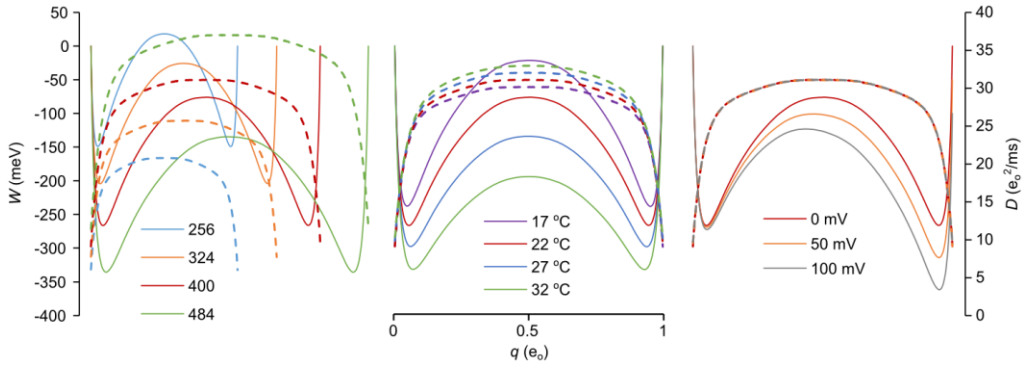

FIG. 9. Free energy (W, solid lines) and diffusion (D, dashed lines) landscapes as a function of N (left), T (middle), and Φ (right). The red curve in each plot corresponds to the reference model. Landscapes were calculated from matrices in Fig. 5, though identical results were obtained by time-averaging Monte Carlo trajectories.

The value of γ was fairly constant across environmental conditions (Fig. 10). 1D rate constant estimates were again slightly larger than their 2D counterparts, and both models yielded k values about 50% greater than the Ising

model. There were some mild discrepancies under conditions where the free energy barrier exceeded 8 kT (T = 17 °C and N = 484). For these taller barriers, the results from dwell time analysis yielded slightly larger k values than the other two methods. However, even these minor outliers did not significantly increase averages across all three methods.

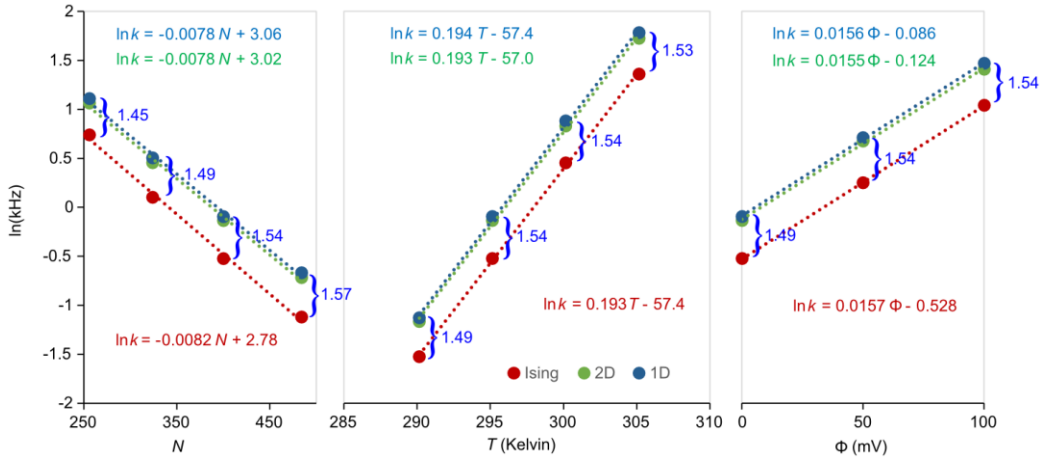

FIG. 10. Linear regression analysis of ln$k$ with environmental conditions same as in Fig 9. The Ising rate constants were calculated as the inverse means of reactant state dwell times. The 1D and 2D rate constants were averaged from their respective dwell time, mfpt, and eigenvalue calculations. Error bars are not visible as they do not exceed the markers. The linear regression equations are color-matched to their respective data. Values of γ are in blue.

Fits using the formula ln$k$ = cX + d, with X representing N, T, and Φ, resulted in nearly identical slopes c for the coarse-grained and gold standard data (Fig. 10). The slope of the ln$k$ vs. N plot is consistent with a chemical potential of activation of 0.2 meV per added gating particle. The ln$k$ vs. T plot yielded an activation energy E_a of 1450 meV. The value of E_a was found to be very close to ΔE = 1485 meV determined from Boltzmann averaging $\langle E\rangle_b - \langle E\rangle_R$ over reactant and barrier regions (Fig. 2). The large activation energy (~57 kT) is interesting from the standpoint of enzyme reactions, since it suggests why temperature studies can sometimes yield large E_a extracted from an Arrhenius plot, while the actual free energy barrier ΔW may be quite small. A large positive ΔS of activation accounts for the discrepancy. Likewise, the ln$k$ vs. Φ plot yielded an activation charge q_a = 0.40 e_o, which compares to $\Delta q \equiv \langle q\rangle_b - \langle q\rangle_R$ = 0.43 e_o. As discussed later, both Δq and ΔE vary substantially for a wider range of temperatures than the "physiological" range considered in Fig.10.

E. Diffusion methods

All but one of preceding numerical methods were based on the 1D or 2D coarse-grained master equation. The exception was brute-force Monte Carlo simulation of Ising dynamics that provided the gold standard value of k. The agreement among the master equation methods suggests coarse-graining is responsible for elevated γ values. To evaluate this further, we employed methods that did not rely on intermediate coarse-graining apart from determining the equilibrium variable $W(q)$. The new quantities of interest are the reactive flux f and diffusion coefficient D, both evaluated at the barrier saddle point region. Because barrier relaxation is fast compared to k, these are efficient calculations.

1. Reactive flux

The reactive flux method originated with molecular dynamics simulations (Chandler, 1978) but was expanded to included discrete-state models (Borkovec and Talkner, 1990). The method distinguishes between starting forward- and backward-directed trajectory ensembles initiated at the barrier separatrix, then keeps track of trajectories residing on the product side of the barrier. At time zero, only the positive flux ensemble is reactive, so $f(0)$ equals the rate constant from transition state theory (TST). At about time T, which is the time required for the system to decay to a metastable state after recrossing the separatrix multiple times, $f(t)$ decays to a plateau value equal to k. The ratio $f(T)/f(0)$ is the transmission coefficient κ. Thus $k = \kappa k_{TST}$. Based on the gold standard value $k = 0.592$ kHz and the 1D model prediction of $k_{TST} = p^* a^* = 55.2$ kHz, where p^* and a^* are the equilibrium probability and 1D forward rate constant at the separatrix, we expect κ to be about 1%. This is a small fraction, requiring a rather large number of reactive trajectories to obtain good statistics.

The reactive flux method applied to the kinetic Ising model, which is formally a continuous-time Markov jump process, is implemented as follows:

$$f(t \geq 0^+) = \frac{p_{eq}^* \left[\left\langle \left(\Sigma \alpha^* \right) h^*(t) \right\rangle_{(+)} - \left\langle \left(\Sigma \beta^* \right) h^*(t) \right\rangle_{(-)} \right]}{p_R^{eq}} \tag{21}$$

where $\Sigma\alpha^*(0)$ and $\Sigma\beta^*(0)$ are summed microscopic forward and backward rates evaluated at time zero (defined as the moment the barrier-confined trajectory lands on the separatrix, after a brief equilibration time) and $h^*(t)$ is the Helmholtz function: $h(q(t) - q^*) = 1$ if $q(t) \geq q^*$, else 0. The subscripts (+) and (–) refer to the forward and backward ensembles of starting trajectories. The (+) ensemble supplies the TST rate for $t = 0^+$. After releasing the barrier confinement for $t > 0$ and applying absorbing boundaries at q_1 and q_2, Eq. 26 approaches at long times:

$$f_{\text{absorbing}}(t \to \infty) = \frac{p_{eq}^* \left[a\left(q^*\right) \pi_{(+)}^* - b\left(q^*\right) \pi_{(-)}^* \right]}{p_R^{eq}}, \tag{22}$$

which may be compared to Eq. 20.

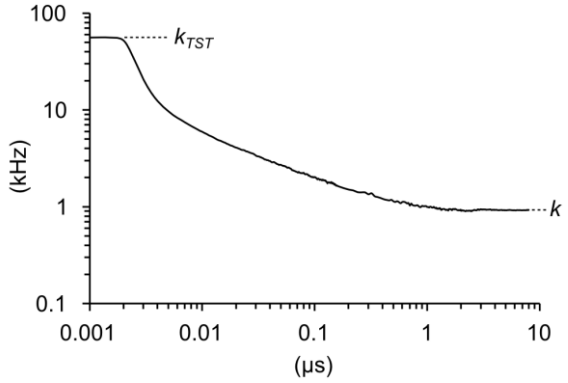

FIG. 11. Reactive flux $f(t)$ for reference model. Independent trajectories began with an 800 ns equilibration phase at the barrier peak, culminating with the crossing of the separatrix, followed by an 8 μs test run from which k was obtained by averaging the last 3 μs.

The result of six experiments, each containing 2 x 10^5 runs, is $k = 0.92 \pm 0.07$ kHz, which is numerically indistinguishable from earlier coarse-grained estimates and again reveals a systematic error $\gamma \approx 3/2$ between accelerated methods and the brute-force gold standard. The error in the reactive rate methodology lies not with the

transmission coefficient, which has value κ = 0.017, but with the statistically averaged TST rate, whose experimental value of 55.8 kHz matches the prediction by the 1D coarse-grained model.

2. Nyquist Analysis

The second of the two diffusion methods measures the local barrier diffusion coefficient D_b from large bandwidth fluctuations confined to the barrier region. The relationship between the landscape variable $D(q)$ and k derives from the Smoluchowski equation (Eq. 14) giving rise to the well-known double integral for the mfpt from q_1 to q_2 (Ansari, 2000):

$$k^{-1} = \int_0^{q_2} dq \frac{1}{D(q)} \exp\left(\frac{W(q)}{kT}\right) \int_{q_1}^{q'} dq' \exp\left(\frac{-W(q)}{kT}\right). \tag{23}$$

Eq. 23 can be simplified in the case of large transition barriers that minimize the overlap between probability distributions in the reactant and barrier regions (see Fig. 2). The double integral reduces to a product of two integrals serving as local partition functions (Sigg, 2014). Since $D(q)$ in the Ising model is nearly constant in the critical barrier region, the mfpt is insensitive to the spatial variation of $D(q)$ and depends largely on the barrier value D_b. The simplified expression is:

$$k = \frac{D_b}{Z_R Z_b}, \tag{24}$$

where $Z_R = \int_R \exp(-W(q)/kT)dq$ is the reactant state partition function evaluated from q = 0 to q^*, and $Z_b = \int_b \exp(W(q)/kT)dq$ is the partition function of the inverted barrier, evaluated from q_1 to q_2. We could derive D_b from the 1D model using Eq. 13, but this simply recapitulates the coarse-grained value obtained using the mfpt method. Instead, we determined D_b independently by applying Nyquist's formula for current noise: $4D_bB = \langle i_g^2 \rangle_b$ to brief, barrier-confined stationary trajectories of the "gating" current $i_g = dq/dt$. Since transitions in the Ising model are discontinuous, careful attention was paid to filtering and sampling i_g trajectories as previously described (Sigg et al.,

1999). The recording bandwidth for a Gaussian filter is $B = 1.064 f_c$, where f_c is the cut-off frequency (Colquhoun and Sigworth, 1983). With increasing bandwidth D_b approaches the correct limiting value (Fig 12a). We expanded Nyquist's theorem to two dimensions to evaluate diffusion along charge and energy axes:

$$4D_{mn}B = \frac{1}{T}\int_{t=0}^{T} m'(t)n'(t)dt \ . \tag{25}$$

D_{mn} is the discrete-variable diffusion matrix evaluated at the saddle point. We confined trajectories to a hard window between 0.35 e_o and 0.65 e_o. The charge diffusion coefficient is $D_q = D_{mm}\delta q^2$. Averaging over 6 x 500 Ising trajectories of 1.6 μs duration with $f_c = 10^7$ kHz, we obtained $D_q = 30.1$ e_o^2/ms for the reference model. Applying Eq. 24, we calculated $k = 0.893 \pm 0.004$ kHz, which, like the reactive flux method, is consistent with earlier results. The complete diffusion matrix was:

$$\mathbf{D} = \begin{bmatrix} 1 & 0.009 \\ 0.009 & 1.775 \end{bmatrix} D_{mm} \ . \tag{26}$$

It is possible for the rate constant error γ to exceed unity if diffusion across the reactive q-coordinate (D_{mm}) is substantially faster than the stable E-coordinate (D_{nn}) (Berezhkovskii and Zitserman, 1990). Eq. 26 rules out diffusion anisotropy error in the 2D model, as the energy diffusion coefficient D_{nn} is roughly three-quarters larger than charge diffusion D_{mm}. The exact relaxation time depends also on the curvature of the free energy profile. The curvature matrix **C** (Hessian) at the saddle point, obtained by fitting W_{mn} to a second order polynomial in m and n, is:

$$\mathbf{C} = \begin{bmatrix} -1 & 0.002 \\ 0.002 & 30.19 \end{bmatrix} \left| C_{mm} \right| , \tag{27}$$

where $C_{mm} = -9.72 \times 10^{-3}$ meV. Using only diagonal terms in **C** and **D**, we expect $\langle E \rangle$ to decay about 54 (1.775 x 30.19) times more rapidly than $\langle q \rangle$. The observed ratio (Fig. b) is 39, which we attribute to fitting error and small deviations from parabolic potentials. In any case, q is clearly the slower variable.

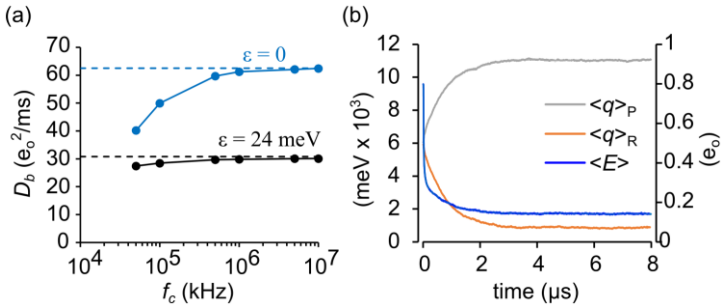

FIG. 12. Diffusion coefficients. (a) Evaluating D_q as a function of filter cut-off frequency. The reference model corresponds to black circles, which approach the peak value of $D(q)$ (dashed line) for large bandwidth. The blue circles represent the control case of non-interacting particles, where $D_q = 0.5N\nu\delta q^2$ (dashed line) is derived analytically. (b) $\langle E\rangle$ and $\langle q\rangle$ decay averaged from 500 simulations starting from an initial random distribution. $\langle E\rangle$ decay was bi-exponential, with a fast decay (19.5 ns) reflecting rapid equilibration at the saddle point, followed by a slower (745 ns) decay as the system relaxed to one of the stable states. The slower decay matched the single time constant of $\langle q\rangle$ relaxation (761 ns).

We tested the appropriateness of q as reaction coordinate in the 2D model by applying results from variational transition state theory (Berezhkovskii and Szabo, 2005). The theory determines the direction **e** of the unstable diffusive mode by solving the eigenvalue equation:

$$\mathbf{CDe} = -\lambda_+ \mathbf{e}\,, \tag{28}$$

where λ_+ is the only positive eigenvalue of $-\mathbf{CD}$. Given very small value of cross terms in **C** and **D**, it is not surprising that **e** = (0.9999998, 0.0062) aligns almost exactly with the q-axis, confirming that q is a good reaction coordinate for (E, q) space. However, in full configuration space there might be additional collective variables that shift **e** away from the q-axis.

F. Committor testing

We attempted to detect the existence of other reactive collective variables with committor testing (Du et al., 1998; Geissler et al., 1999). It is known that the optimal reaction coordinate in configuration space is the committor or splitting probability $\pi(\mathbf{x})$ (Hummer, 2004; Banushkina and Krivov, 2016; E et al., 2005; Berezhkovskii and

Szabo, 2013). The diffusion landscape using the committor as reaction coordinate $\{W(\pi(\mathbf{x})), D(\pi(\mathbf{x}))\}$ yields the correct value of k, but constructing the committor landscape from the full configuration space may be onerous (reviewed by (Peters, 2016)) and does not fulfill our stated aim of achieving a complete solution for the (T, Φ) ensemble. Using the coarse-grained 1D or 2D model as a starting point, extracting the committor function $\pi(q)$ or $\pi(E, q)$ by solving Eq. 18 is straightforward (Fig. 7c), but computing k by incorporating these committor values into Eq. 19 yields precisely the same value of γ as the mfpt method. This suggests that coarse-graining distorts the committor function. We can test this directly by invoking the committor test:

$$\pi(\mathbf{x}) = \pi\left(q(\mathbf{x}), E(\mathbf{x})\right) \tag{29}$$

Eq. 29 states the committor of a configuration **x** should equal the committor of **x** projected onto the coarse-grained (E, q) space. We approached the test in two ways. First, we mined configurations **x** for which $\pi(\mathbf{x}) = 0.5$ and plotted them on the (E, q) landscape, where the symmetry of the reference model requires $\pi(q^*) = 0.5$. Second, we tested configurations located on the separatrix (E, q^*) to see if the distribution of committor probabilities deviated significantly from the expected value of 0.5.

To apply the first test, we harvested 34 unbiased reactive trajectories starting from $q = 0.3$ e_o and ending at $q = 0.7$ e_o. Trajectories ranged from 3,029 to 21,595 Monte Carlo moves. Using the method described in Pan and Chandler (Pan and Chandler, 2004), we tested each configuration for the condition $\pi(\mathbf{x}) = 0.5$ by launching n test runs that terminated after one or the other stable basins (q_1 or q_2) was reached. We then determined if the cumulative value of $\pi(\mathbf{x}) = n_2/n$ was within an acceptable range. Conditions were met if, after 10^3 runs, a candidate configuration landed within the 95% confidence interval of a coin flip: 0.5 ± 0.031. A total of 14,837 isocommittor configurations were mined in this fashion and projected onto the (E, q) free energy surface. The isocommittors clustered around the separatrix with a single peak (Fig. 13a). The absence of a multimodal pattern argues against the need to invoke additional collective variables in defining the reaction coordinate.

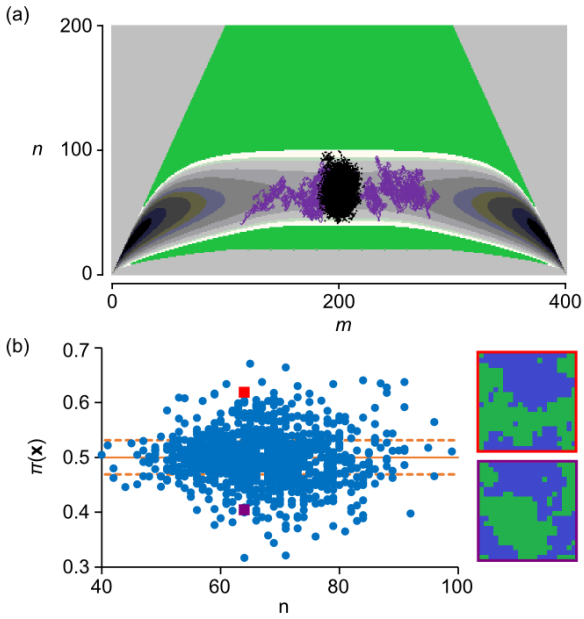

FIG. 13. Committor analysis. (a) 2D distribution of $\pi = 0.5$ configurations (black dots) with 95% confidence interval [0.469, 0.532]. A typical crossing trajectory is shown in purple. (b) Committor probabilities of configurations harvested from the separatrix ($m = 200$), evaluated from 10^3 test runs. The dashed lines demonstrate the 95% confidence interval for 10^3 random coin flips. Red and violet squares represent independent configurations (shown at right) in the same state ($m = 200$, $n = 65$) but with different π values (red square, 0.407 ± 0.003; violet square, 0.623 ± 0.002).

We implemented the second test by constraining the Ising dynamics to the barrier region with hard walls and launching 10^3 unconstrained test runs every time the system reached q^*. The resulting $\pi(\mathbf{x}^*)$ values were compared to the expected distribution of 10^3 coin flips. A greater than expected number of $\pi(\mathbf{x}^*)$ values exceeded the 95% confidence interval (Fig 13b). Absent a multimodal committor distribution, failure of coarse-grained states to predict the exact committor probability fundamentally violates the Markov (memoryless) condition, since one cannot predict the future even in a probabilistic sense; therefore past events assume importance.

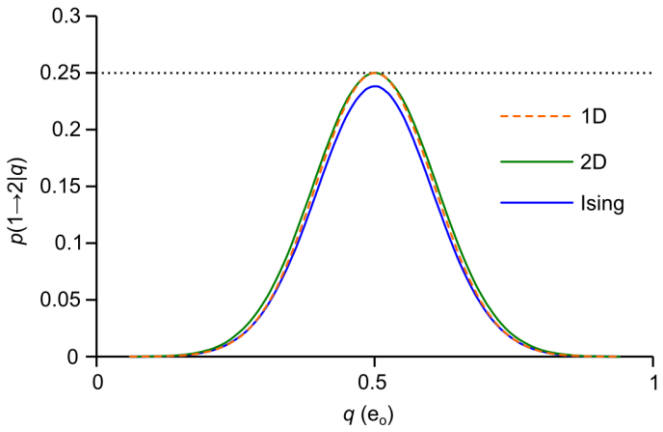

FIG. 14. Markovianity analysis demonstrating memory in Ising kinetics projected onto 1D, but not projections from 1D (control) or 2D kinetics.

To confirm that coarse-graining induces loss of dynamical memory, we took advantage of a simple test for Markovianity published recently (Berezhkovskii and Makarov, 2018). The test evaluates conditional probabilities of the sort $p(a \to b \mid q)$ for an ensemble of trajectories that begin at q_a and are absorbed by q_b after crossing the test point q at least once. From a single unbiased trajectory containing 5,000 barrier transitions, using the free energy minima q_1 and q_2 as our endpoints, we accumulated statistics for each permutation $p(1 \to 1 \mid q)$, $p(1 \to 2 \mid q)$, $p(2 \to 1 \mid q)$, and $p(2 \to 2 \mid q)$. Counts were tallied concurrently for every q spanning the interval q_1 to q_2. This lead to correlations between neighboring q, but running the simulation multiple times (n = 14) yielded good statistics for the entire curve. The Markovianity condition requires $p(1 \to 2 \mid q)_{max} = 0.25$. A smaller peak probability implies memory effects. A larger value is consistent with ballistic dynamics, which do not play a role here. We simulated Ising, 1D, and 2D trajectories. The 1D trajectory serves as control, since 1D projects onto its own space. Both 1D and 2D peak values were statistically indistinguishable from the critical 0.25 value (1D: 0.2499 ± 0.0006; 2D: 0.2497 ± 0.0008), but the Ising trajectory peaked at a lower value of 0.2381 ± 0.0004 (Fig. 14). This 4.8% reduction from the Markov prediction is small. We propose nevertheless that the memory effect accounts for the observed coarse-graining error γ. A bonus to performing the Markovianity test is that, since $\pi(q) = p(1 \to 2 \mid q) + p(2 \to 2 \mid q)$, event tallies can be used to generate committor probabilities. Measuring $\pi(q)$ from the Markovianity analysis, we calculated k through the use of Eq. 20. Unsurprisingly the results were consistent with earlier findings of $\gamma \approx 3/2$ (see Table II).

G. Summary of *k* calculations

In Table II, we grouped outcomes according to degree of coarse-graining. Only one method, brute force determination of dwell times, was a pure configurational process (**x**) lacking any coarse-graining; therefore this produced the numerical gold standard value $k = 0.592 \pm 0.007$ kHz. The remaining eleven methods relied on coarse-grained (1D or 2D) models or nominally "model-free" diffusion methods (reactive flux and Nyquist analysis). All eleven methods predicted $\gamma \approx 3/2$. The pooled average value of k was 0.896 ± 0.007 kHz, yielding $\gamma = 1.51 \pm 0.02$. The γ error was fairly constant across conditions consistent with two-state dynamics (Fig. 10). There were only minor differences between 1D and 2D methods. Committor and Markovianity analysis suggest that most of information loss incurred by coarse-graining occurs while projecting Ising kinetics onto the (E, q) plane.

TABLE II. Summary of k_{12} averaged from 6 experiments (standard model).

Method	Process	k_{12} (kHz)	standard error
Dwell time (Ising)	**x**	0.59159	0.006895
Dwell time(2D)	(E, q)	0.891089	0.009467
Dwell time (1D)	q	0.92513	0.005992
eigenvalue (2D)	(E, q)	0.866373	0.003352
eigenvalue (1D)	q	0.906609	0.003512
mfpt (2D)	(E, q)	0.859934	0.002575
mfpt (1D)	q	0.90043	0.002721
Nyquist (Ising)	q(**x**)	0.892715	0.003941
reactive flux (Ising)	q(**x**)	0.921411	0.073493
1D committor (Ising)	q(**x**)	0.907928	0.000948
1D committor (2D)	q(E, q)	0.886759	0.000979
1D committor (1D)	q(q)	0.903906	0.000579

IV. DISCUSSION

In this study we numerically solved the forward rate constant k of a kinetic Ising model for a range of temperatures and field strengths compatible with a modest (5-9 kT) energy barrier. The solution is "complete" in the sense that the initial projection of Ising dynamics onto the microcanonical ensemble served as a kernel for subsequent choices of environmental constraints. Brute force simulations were used to obtain the gold standard value of k. The q-coordinate proved to be a good reaction coordinate within the context of the coarse-grained (E, q) space, and committor analysis failed to uncover additional reactive variables. Nevertheless, coarse-grained kinetics consistently overestimated the gold standard value by about 50% ($\gamma \approx 3/2$). The value of γ did not vary significantly across experimental conditions. Markovianity testing applied to reactive trajectories were consistent with loss of dynamical memory after coarse-graining—presumably this is the major source of coarse-graining error.

A. WHICH MEMORY MODEL?

The positive Markovianity test raises the question of which memory model is appropriate for a non-inertial jump system such as the Ising model. The model is a step-in, step-out scheme without intra-state transitions (Fig. 4). Memory therefore resides in the sequence of events leading up to an occupied state, ruling out a class of generalized master equations in which a time-dependent rate constant arises from intra-state thermalization (Hänggi and Talkner, 1981). To our knowledge there is no analytically derived coarse-grained memory kernel that phenomenologically describes the kinetic Ising model.

The two diffusion methods, reactive flux and Nyquist analysis, are "model-free" in the sense that no conditions are placed on the shape of the free energy landscape, provided the reaction coordinate is a good one (i.e. closely aligns with the committor function). We considered adapting these methods to incorporate memory effects. Memory in an inertial system governed by molecular dynamics, such as a protein molecule in a solvent bath, can be modeled by a general Langevin equation with a time-dependent friction kernel (Mori, 1965; Zwanzig, 1961). In such a system featuring a single reaction coordinate with harmonic barrier frequency ω, the forward rate constant is given by $k = \kappa k_{TST}$, where k_{TST} is determined by equilibrium statistical mechanics and the transmission coefficient κ equals the (Grote-Hynes) ratio λ_1/ω, where λ_1 is the smallest positive root of the transfer function between the Langevin

force and the reaction coordinate (Grote and Hynes, 1980; Hanggi and Mojtabai, 1982; Pollak, 1986; Kohen and Tannor, 1999). Thus in inertial systems, memory is reflected in the transmission coefficient. In the overdamped Ising model, reactive flux experiments revealed a different dynamic. The coarse-graining error γ was accounted for entirely by the value of k_{TST}, since the latter is predicted by the 1D coarse-grained formula $k_{TST} = a^* p^*$. Thus, unlike the Grote-Hynes formalism, memory effects in the Ising model do not appear in the transmission coefficient.

Like reactive flux, Nyquist analysis, which determines the value of the diffusion coefficient D_b in the critical barrier region, overestimates the true rate constant k by the factor $\gamma \approx 3/2$ The source of error in this case resides in the value of D_b, which equals the coarse-grained diffusion coefficient $D^* = (1/2)(a^* + b^*)\delta q^2$ at the transition state. We hypothesized that a spectral decomposition of D_b would reveal the reactive component leading to the correct value of k. To this end, we examined the stationary autocorrelation function of the gating current $R(t) = \langle i_g(0) i_g(t) \rangle_{eq}$. $R(t)$ can be computed by eigenvalue decomposition of the rate matrix **A** (Sigg et al., 1999):

$$R(t) = \chi\delta(t) + \sum_r \sigma_r \exp(\lambda_r t)\,, \tag{30}$$

where $\sigma_r = -\lambda_r^2 \langle \mathbf{u}_0 \mathbf{Q} \mathbf{v}_r \rangle \langle \mathbf{u}_r \mathbf{q} \rangle$ are spectral amplitudes containing eigenvalues (λ_r) of **A** and the corresponding left- and right- eigenvectors ($\mathbf{u}_r$, $\mathbf{v}_r$) multiplied by q in diagonal matrix (**Q**) and vector (**q**) form. The spectral amplitudes are real-valued and negative, offset at time zero by the positive-valued delta term. This is in contrast to inertial systems whose velocity autocorrelation function decays with positive real amplitudes and may also contain dampened oscillations. The delta weight χ adds no additional information not already present in the exponential term, but ensures that the filtered one-sided power spectrum calculated from the Wiener-Khintchine theorem is positive (Sigg et al., 1999):

$$S(\omega) = 4\sum_r \left(\frac{\sigma_r}{\lambda_r} \right) \frac{\omega^2}{\omega^2 + \lambda_r^2} |H|^2\,, \tag{31}$$

where H is the Fourier-transformed filter response function. For cut-off frequencies f_c much larger than the largest significant eigenvalue $\lambda_{\max}$, integrating Eq. 31 across positive frequencies yields a gating current variance that is proportional to the filter bandwidth:

$$\left\langle i_g^2 \right\rangle = 4\sum_{r=1}^{r_{\max}} \frac{\sigma_r}{\lambda_r} B\,, \tag{32}$$

Comparing Eq. 32 to the Nyquist current formula $\langle i_g^2 \rangle = 4DB$, we can express the diffusion coefficient as a sum of spectral components:

$$D = \sum_{r=1}^{r_{\max}} \frac{\sigma_r}{\lambda_r}\,. \tag{33}$$

We performed spectral analysis by applying a harmonic umbrella potential with curvature c centered at the transition state q^* and generating 10^5 runs of 410 ns trajectories with fc = 10^7 kHz bandwidth. The function $R(t)$ was calculated by averaging $i_g(t_1)i_g(t_1 + t)$ over every possible t_1 in each run, and then averaging over all runs. The number of paired currents included in a correlation time t was proportional to $(410 - t)$. This number was used as a weighting factor to generate a uniform residual for exponential fits of the autocorrelation function (Fig. 15a, b). The 1D and 2D autocorrelations were fit to a single exponential ($r_{max} = 1$) as expected from a process that is stationary, quasi-harmonic, and Markovian, thus modeled by the Ornstein-Uhlenbeck process (Doob, 1942). The area beneath the 1D curve equaled $4D_b$ = 124.0 e_o^2/ms (n = 6) consistent with the coarse-grained prediction of $D_b = (a^* + b^*)\delta q^2/2$ = 31.1 e_o^2/ms. The decay rate λ of the 1D autocorrelation function depended linearly on c, with extrapolation back to c = 0 yielding the negative curvature of the original barrier potential. We chose $c = 10^4$ mV/e_o since this value struck the right balance between constraining the system to the barrier region but being small enough to separate eigenvalues and still satisfy $f_c >> \lambda_{max}$. The outcome was insensitive to the value of c as long as the preceding conditions were satisfied. The full configuration Ising process yielded three exponential components (r_{max} = 3), with total area again equal to $4D_b$. The slowest Ising component decayed with nearly the same rate λ_1 as the 1D model (Fig.15c, d), a result that held across all temperatures (Fig. 15e), but the corresponding area was $4D_1$ = 20.0 e_o^2/ms (n = 6), about 2/3 of the 1D prediction. This suggests that the slow Ising decay matching the 1D model decay is the reactive component, since it is consistent with $\gamma \approx 3/2$. We are empirically lead to the relation:

$$k = \frac{D_1}{Z_R Z_b}\,, \tag{34}$$

where $D_b/D_1 \approx 3/2$. However, this ratio, which equals the experimental γ, cannot hold for all conditions, since for very large temperatures the full configuration Ising model behaves as system of independent particles approaching the Ornstein-Uhlenbeck process, where $D_b = D_1$ (Fig. 12a). But for conditions consistent with a modest free energy barrier, the spectral ratio (D_1/D_b) roughly coincides with experimental γ values (Fig. 15f).

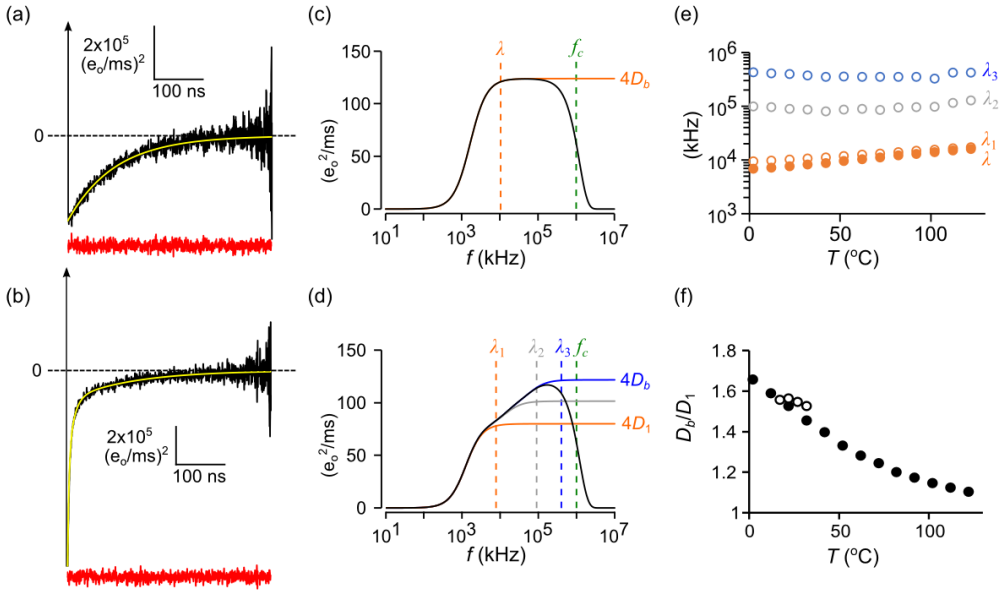

FIG. 15. Spectral decomposition of the barrier diffusion coefficient. (a) Gating current autocorrelation function for the 1D model (black line) fit to a single exponential (yellow line). The red trace is the residuals plot after normalizing for count statistics. Sampling and filtering cut-off frequencies were: $f_s = 10^7$ kHz; $f_c = 10^6$ kHz. (b) Same as (a) except with full configuration Ising kinetics, requiring a three-exponential fit. (c) Power spectrum of the 1D model calculated from the fitted data. Dashed vertical lines indicate values for the decay rate (λ) and the cut-off frequency (f_c). (d) Same as (c) except with Ising kinetics. (e) Ising (solid circles) and 1D (open circles) decay rates as a function of temperature. (f) D_b/D_1 as a function of temperature (filled circles) compared to experimental values of γ (open circles). Each data point in (e) and (f) is averaged from 6 experiments.

B. NUMERICAL STRATEGY FOR COMPUTING $k(T, \Phi)$

We conclude by proposing a general strategy for computing k as a function of T and Φ that is summarized by Eq. 34. First, one projects, using Monte Carlo sampling, the full Ising dynamics onto the microcanonical space (E, q). Then, after solving the coarse-grained master equation in the (T, Φ) ensemble, one compensates for the ensuing systematic error γ by finding the slowest (reactive) spectral component D_1 of the barrier diffusion coefficient D_b.

The last part is empirical but makes intuitive sense and gives numerically correct answers for small barriers. We provided evidence that most of the coarse-graining error was due to projecting Ising dynamics onto (E, q)-space, and that the source of the error was attributable to loss of system memory rather than obviously poor choices in collective variables. Projecting (E, q)-space dynamics onto q-space introduced only minor error; therefore q was considered to be a good reaction coordinate once γ was factored into estimates of k.

The equilibrium factors Z_R^{-1} and Z_b^{-1} in Eq. 34 deserve further discussion. As explained previously, the Z are local partition functions summing over individual basin and barrier regions. Their product can be rewritten as $\exp(-\Delta W/kT)$, where the activation free energy $\Delta W = W_b - W_R$ is the difference between barrier and reactant energies $W_R = -kT\ln Z_R$ and $W_b = kT\ln Z_b$. The activation free energy can be further broken down into its thermodynamic components: $\Delta W = \Delta E - T\Delta S - \Phi\Delta q$, where $\Delta E = E_b - E_R$ and so forth for the other extensive variables S and q. The regional values of extensive variables are obtained by Boltzmann averaging over the region of interest. Thus $E_x = \langle E\rangle_x$, $q_x = \langle q\rangle_x$, and $S_x = -k\langle \ln p\rangle_x$, where $x = \{\mathrm{R, b, P}\}$ defines the region of summation. In our study, ensemble averaging was performed over microcanonical ensemble states $(m, n)_x$ with regional probability distribution functions $p_{mn} = Z_x^{-1}\exp[\pm(E_m - k\ln\Omega_{mn}T - q_m\Phi)/kT]$, where the positive sign in the exponent corresponds to barriers and the negative sign to basins. We defined the reactant basin as in the range $[0, q^*]$ and the barrier region in the range $[q_1, q_2]$. The reactant and barrier ranges overlap, but for well-separated regional probability distributions (Fig. 2), this has little consequence for the calculation of ΔW. The regional averaging described above reduces the 2D pmf to a discrete "saw-tooth" landscape consisting of only three (q, W) coordinates in a way that preserves the thermodynamic contribution to the rate constant, and does not require basins or barriers to be parabolic or symmetric (Fig. 16a).

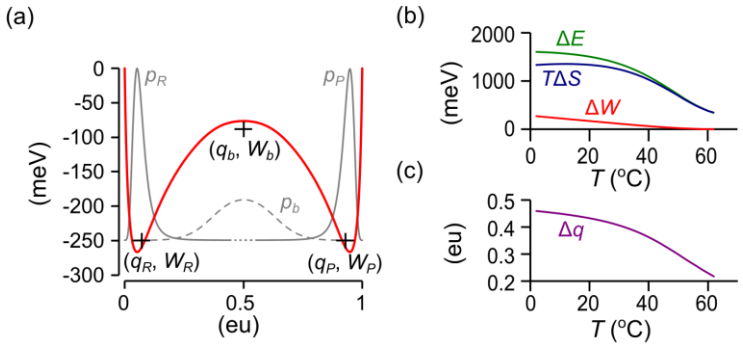

FIG. 16. Discretization of 1D diffusion landscape (reference model). (a) Black crosses marking regional reaction coordinates (q_x, W_x) form the basis of a "saw-tooth" landscape. The coordinates are superimposed onto $W(q)$ (red curve) and the respective segmental equilibrium probability distributions (x = R, P: solid gray curves; x = b: dashed gray curve). Note that the discrete (q, W) coordinates do not necessarily coincide with the extrema of $W(q)$. (b) Temperature dependence of forward reaction energies. (c) Temperature dependence of forward reaction charge.

The kinetic term D_1 in Eq. 34 is the reactive Nyquist component whose decay matches the single component decay of the 1D coarse-grained model, and whose amplitude relative to D_b predicts the experimentally determined coarse-graining error γ. Therefore, we equated $D_1(T, \Phi)$ with $D^*(T, \Phi)/\gamma$, where $\gamma \approx 3/2$ was shown to be roughly constant for the range of (T, Φ) consistent with a modest-sized barrier. We should note that for the Ising model, the 1D diffusion coefficient D_q was found to be nearly constant across the barrier region (Fig. 9). A model in which D_q varies strongly across the barrier region would add a "spurious drift" term $-kT\ln(D_q/D_1)$ to the exponent of Z_b (Van Kampen, 1992; Peters et al., 2013). This would not change global equilibrium parameters since Z_R and Z_P are unaffected.

It is worth repeating in the context of the saw-tooth free energy landscape that the primary advantage of the microcanonical approach is that one-time coarse-graining yields D, ΔE, ΔS, and Δq as functions of T and Φ (Fig. 9 and Fig. 16b, c). This knowledge, and the correction factor γ, yields a complete picture of Ising two-state dynamics. The hope is that analogous microcanonical coarse-graining of molecular dynamics simulations applied to proteins and other macromolecules can be used to expand the range of environmental sensitivities that are currently limited to one set of conditions per simulation.

REFERENCES

Ansari, A. 2000. Mean first passage time solution of the Smoluchowski equation: Application to relaxation dynamics in myoglobin. *J. Chem. Phys.* 112:2516–2522. doi:10.1063/1.480818.

Banushkina, P. V., and S. V. Krivov. 2016. Optimal reaction coordinates. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 6:748–763. doi:10.1002/wcms.1276.

Belardinelli, R.E., and V.D. Pereyra. 2007. Fast algorithm to calculate density of states. *Phys. Rev. E*. 75:046701.

doi:10.1103/PhysRevE.75.046701.

Berezhkovskii, A., G. Hummer, and A. Szabo. 2009. Reactive flux and folding pathways in network models of coarse-grained protein dynamics. *J. Chem. Phys.* 130:205102. doi:10.1063/1.3139063.

Berezhkovskii, A., and A. Szabo. 2005. One-dimensional reaction coordinates for diffusive activated rate processes in many dimensions. *J. Chem. Phys.* 122:14503.

Berezhkovskii, A.M., and D.E. Makarov. 2018. Single-Molecule Test for Markovianity of the Dynamics along a Reaction Coordinate. *J. Phys. Chem. Lett.* 9:2190–2195. doi:10.1021/acs.jpclett.8b00956.

Berezhkovskii, A.M., and A. Szabo. 2013. Diffusion along the Splitting/Commitment Probability Reaction Coordinate. *J. Phys. Chem. B*. 117:13115–13119. doi:10.1021/jp403043a.

Berezhkovskii, A.M., and V.Y. Zitserman. 1990. Activated rate processes in a multidimensional case. A new solution of the Kramers problem. *Phys. A Stat. Mech. its Appl.* 166:585–621. doi:10.1016/0378-4371(90)90075-4.

Borkovec, M., and P. Talkner. 1990. Generalized reactive flux method for numerical evaluation of rate constants. *J. Chem. Phys.* 92:5307–5310. doi:10.1063/1.458535.

Brendel, K., G.T. Barkema, and H. van Beijeren. 2005. Nucleation times in the two-dimensional Ising model. *Phys. Rev. E*. 71:031601. doi:10.1103/PhysRevE.71.031601.

Chandler, D. 1978. Statistical mechanics of isomerization dynamics in liquids and the transition state approximation. *J. Chem. Phys.* 68:2959. doi:10.1063/1.436049.

Coddington, P.D. 1994. Analysis of random number generators using Monte Carlo simulation. *Int. J. Mod. Phys. C*. 05:547–560. doi:10.1142/S0129183194000726.

Colquhoun, D., and F.J. Sigworth. 1983. Fitting and Statistical Analysis of Single-Channel Records. *In* Single-Channel Recording. Springer US, Boston, MA. 191–263.

Cordero-Morales, J.F., L.G. Cuello, and E. Perozo. 2006. Voltage-dependent gating at the KcsA selectivity filter. *Nat. Struct. Mol. Biol.* 13:319–22. doi:10.1038/nsmb1070.

Delemotte, L., M.A. Kasimova, D. Sigg, M.L. Klein, V. Carnevale, and M. Tarek. 2017. Exploring the Complex Dynamics of an Ion Channel Voltage Sensor Domain via Computation. *bioRxiv*. 108217. doi:10.1101/108217.

Doob, J.L. 1942. The Brownian Movement and Stochastic Equations. *Ann. Math.* 43:351. doi:10.2307/1968873.

Du, R., V.S. Pande, A.Y. Grosberg, T. Tanaka, and E.S. Shakhnovich. 1998. On the transition coordinate for protein folding. *J. Chem. Phys.* 108:334–350. doi:10.1063/1.475393.

E, W., W. Ren, and E. Vanden-Eijnden. 2005. Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes. *Chem. Phys. Lett.* 413:242–247. doi:10.1016/j.cplett.2005.07.084.

Ferrenberg, A.M., and R.H. Swendsen. 1988. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* 61:2635–2638. doi:10.1103/PhysRevLett.61.2635.

Frehland, E. 1978. Current noise around steady states in discrete transport systems. *Biophys. Chem.* 8:255–65.

Geissler, P.L., C. Dellago, and D. Chandler. 1999. Kinetic Pathways of Ion Pair Dissociation in Water. *J. Phys. Chem. B.* 103:3706–3710. doi:10.1021/jp984837g.

Gillespie, D.T. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81:2340–2361. doi:10.1021/j100540a008.

Glauber, R.J. 1963. Time-Dependent Statistics of the Ising Model. *J. Math. Phys.* 4:294–307. doi:10.1063/1.1703954.

Grote, R.F., and J.T. Hynes. 1980. The stable states picture of chemical reactions. II. Rate constants for condensed and gas phase reaction models. *J. Chem. Phys.* 73:2715–2732. doi:10.1063/1.440485.

Hanggi, P., and F. Mojtabai. 1982. Thermally activated escape rate in presence of long-time memory. *Phys. Rev. A.* 26:1168–1170. doi:10.1103/PhysRevA.26.1168.

Hänggi, P., and P. Talkner. 1981. Non-Markov processes: The problem of the mean first passage time. *Zeitschrift für Phys. B Condens. Matter*. 45:79–83. doi:10.1007/BF01294279.

Harpole, T.J., and L. Delemotte. 2017. Conformational landscapes of membrane proteins delineated by enhanced sampling molecular dynamics simulations. *Biochim. Biophys. Acta - Biomembr.* 0–1.

doi:10.1016/j.bbamem.2017.10.033.

Hill, T.L. 1962. Thermodynamics of Small Systems. *J. Chem. Phys.* 36:3182–3197. doi:10.1063/1.1732447.

Hűller, A., and M. Pleimling. 2002. Microcanonical Determination of the Order Parameter Critical Exponent. *Int. J. Mod. Phys. C*. 13:947–956. doi:10.1142/S0129183102003693.

Hummer, G. 2004. From transition paths to transition states and rate coefficients. *J. Chem. Phys.* 120:516–523. doi:10.1063/1.1630572.

Van Kampen, N.G. 1992. Stochastic Processes in Physics and Chemistry. 2nd ed. North-Holland Personal Library, Amsterdam.

Kandiah, K., M.O. Deighton, and F.B. Whiting. 1989. A physical model for random telegraph signal currents in semiconductor devices. *J. Appl. Phys.* 66:937–948. doi:10.1063/1.343523.

Kjelstrup, S., G.P. Beretta, A. Ghoniem, and G. Hatsopoulos. 2008. Mesoscopic Non-Equilibrium Thermodynamics and Biological Systems. *In* AIP Conference Proceedings. AIP. 223–228.

Kohen, D., and D.J. Tannor. 1999. Advances in Chemical Physics. 111. I. Prigogine and S.A. Rice, editors. John Wiley & Sons, Inc., Hoboken, NJ, USA. 219–398 pp.

Kurata, H.T., M. Rapedius, M.J. Kleinman, T. Baukrowitz, and C.G. Nichols. 2010. Voltage-Dependent Gating in a "'Voltage Sensor-Less'" Ion Channel. *PLoS Biol*. 8:1000315. doi:10.1371/journal.pbio.1000315.

L'Ecuyer, P. 1999a. Tables of linear congruential generators of different sizes and good lattice structure. *Math. Comput.* 68:249–261. doi:10.1090/S0025-5718-99-00996-5.

L'Ecuyer, P. 1999b. Good Parameters and Implementations for Combined Multiple Recursive Random Number Generators. *Oper. Res.* 47:159–164. doi:10.1287/opre.47.1.159.

Leyssale, J.-M., J. Delhommelle, and C. Millot. 2007. Hit and miss of classical nucleation theory as revealed by a molecular simulation study of crystal nucleation in supercooled sulfur hexafluoride. *J. Chem. Phys.* 127:044504. doi:10.1063/1.2753147.

Metzner, P., C. Schütte, and E. Vanden-Eijnden. 2009. Transition Path Theory for Markov Jump Processes.

Multiscale Model. Simul. 7:1192–1219. doi:10.1137/070699500.

Mori, H. 1965. Transport, Collective Motion, and Brownian Motion. *Prog. Theor. Phys.* 33:423–455. doi:10.1143/PTP.33.423.

Moritz, C., A. Tröster, and C. Dellago. 2017. Interplay of fast and slow dynamics in rare transition pathways: The disk-to-slab transition in the 2d Ising model. *J. Chem. Phys.* 147:152714. doi:10.1063/1.4997479.

Onsager, L. 1944. Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition. *Phys. Rev.* 65:117–149. doi:10.1103/PhysRev.65.117.

Pan, A.C., and D. Chandler. 2004. Dynamics of Nucleation in the Ising Model †. *J. Phys. Chem. B.* 108:19681–19686. doi:10.1021/jp0471249.

Peters, B. 2016. Reaction Coordinates and Mechanistic Hypothesis Tests. *Annu. Rev. Phys. Chem.* 67:669–690. doi:10.1146/annurev-physchem-040215-112215.

Peters, B., P.G. Bolhuis, R.G. Mullen, and J.-E. Shea. 2013. Reaction coordinates, one-dimensional Smoluchowski equations, and a test for dynamical self-consistency. *J. Chem. Phys.* 138:054106. doi:10.1063/1.4775807.

Peters, B., and B.L. Trout. 2006. Obtaining reaction coordinates by likelihood maximization. *J. Chem. Phys.* 125:054108. doi:10.1063/1.2234477.

Pleimling, M., and H. Behringer. 2005. Microcanonical analysis of small systems. *Phase Transitions*. 78:787–797. doi:10.1080/01411590500288999.

Pollak, E. 1986. Theory of activated rate processes: A new derivation of Kramers' expression. *J. Chem. Phys.* 85:865. doi:10.1063/1.451294.

Press, W., S. Teukolsky, W. Vetterling, B. Flannery, E. Ziegel, W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. 1992. Numerical Recipes in C. 2nd ed. Cambridge University Press, New York.

Qian, H. 2001. Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation. *Phys. Rev. E*. 65:016102. doi:10.1103/PhysRevE.65.016102.

Ray, J.R., and C. Freléchoz. 1996. Microcanonical ensemble Monte Carlo method for discrete systems. *Phys. Rev.*

E. 53:3402–3407. doi:10.1103/PhysRevE.53.3402.

Reimann, P., G.J. Schmid, and P. Hänggi. 1999. Universal equivalence of mean first-passage time and Kramers rate. *Phys. Rev. E*. 60:R1–R4. doi:10.1103/PhysRevE.60.R1.

Roux, B., and R. Sauvé. 1985. A general solution to the time interval omission problem applied to single channel analysis. *Biophys. J.* 48:149–158. doi:10.1016/S0006-3495(85)83768-1.

Saito, M., and M. Matsumoto. 2008. SIMD-Oriented Fast Mersenne Twister: a 128-bit Pseudorandom Number Generator. *In* Monte Carlo and Quasi-Monte Carlo Methods 2006. Springer-Verlag, Berlin, Heidelberg. 607–622.

Sigg, D. 2013. A linkage analysis toolkit for studying allosteric networks in ion channels. *J. Gen. Physiol.* 141:29–60. doi:10.1085/jgp.201210859.

Sigg, D. 2014. Modeling ion channels: Past, present, and future. *J. Gen. Physiol.* 144:7–26. doi:10.1085/jgp.201311130.

Sigg, D., H. Qian, and F. Bezanilla. 1999. Kramers' Diffusion Theory Applied to Gating Kinetics of Voltage-Dependent Ion Channels. *Biophys. J.* 76:782–803. doi:10.1016/S0006-3495(99)77243-7.

Sigworth, F.J. 2003. Life's transistors. *Nature*. 423:21–22. doi:10.1038/423021a.

Sigworth, F.J., and S.M. Sine. 1987. Data transformations for improved display and fitting of single-channel dwell time histograms. *Biophys. J.* 52:1047–1054. doi:10.1016/S0006-3495(87)83298-8.

Yang, C.N. 1952. The Spontaneous Magnetization of a Two-Dimensional Ising Model. *Phys. Rev.* 85:808–816. doi:10.1103/PhysRev.85.808.

Zwanzig, R. 1961. Memory Effects in Irreversible Thermodynamics. *Phys. Rev.* 124:983–992. doi:10.1103/PhysRev.124.983.