

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

When computing the frequency of rare events in small systems, 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. The microcanonical density of states accurately preserves equilibrium distributions, but it is an open question whether projecting system dynamics onto the microcanonical ensemble provides accurate kinetics at the macroscopic level. We studied two-state dynamics in a finite kinetic Ising model at subcritical temperatures. 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 provides 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 can be attributed to memory effects. Committor analysis did not confirm the alternative hypothesis that erroneous outcomes were due to a poor choice of collective variables. We were able to empirically relate the value of the rate constant correction factor 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 approach the familiar problem of transition rates in a bistable system by projecting state space dynamics onto the microcanonical ensemble, thereby generating solutions for a range of environmental conditions. Bistability occurs naturally in enzymatic reactions and can also be found in manufactured systems such as MOSFET transistors (Kandiah et al., 1989). The dynamics of bistable systems are 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 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). Efforts have been made to develop a macroscopic picture of ion channel dynamics 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 yields a discrete master equation or diffusion landscape that can be used to calculate transition rate constants (Delemotte et al., 2017). Such solutions are typically incomplete in the sense that they must be repeated for new values of T and Φ . In theory one could obtain a “complete” solution by projecting molecular dynamics onto the microcanonical ensemble, which contains statistical frequencies as a function of collective variables energy (E) and charge (q), and then Laplace-transforming the frequencies to (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 when the full system dynamics is projected onto the combination 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 the electrical isomorph of the traditional two-dimensional Ising model. The existence of an upper bound for both E and q in the finite lattice allows the complete state space to be projected onto a finite microcanonical grid (Ferrenberg and

Swendsen, 1988; Pleimling and Behringer, 2005; Hüller and Pleimling, 2002; Ray and Freléchoz, 1996). We restricted the lattice length L to the range $\{16, 18, 20, 22\}$ corresponding to $N = L^2$ particles, which is comparable to the number of residues in a biological macromolecule. The small N enabled efficient computing, allowing a large number of unbiased simulated transitions to be harvested in a relatively short time. Previous kinetic studies of Ising models (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 findings 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 was the case in our study. Rather our focus was to compute rates for a small system with modest free energy barriers (5 to 9 kT) as a function of T and Φ . In a small system such as this, susceptibilities are softened, resembling response curves seen in regulatory proteins, while 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 and an electrical work term $-q\Phi$ in place of the usual pressure-volume work.

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

The first (energy) term sums over nearest-neighbor interactions $\langle i, j \rangle$ between “gating” particles, which have value $e_i = 0$ or 1 . Adjacent particles with dissimilar e_i 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 in the work term 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 term $U(q) = c(q - q_o)^2$ 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 picked 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 poorly known, but presumably it involves long-range correlations in the secondary structure that change pore selectivity. Figure 1 shows how a small “pore” region in the Ising grid comprised of four selectivity particles generates macroscopic fluctuations in the “ionic current” that are qualitatively similar to those from real single channel recordings, including characteristic voltage- and temperature-dependences. For the Ising model, we focused on “gating” charge q since it is a proper thermodynamic variable, whereas the artificial ionic current i is simply a marker of activation and not energetically linked to gating. We note however, that q and i yield identical random telegraph signals.

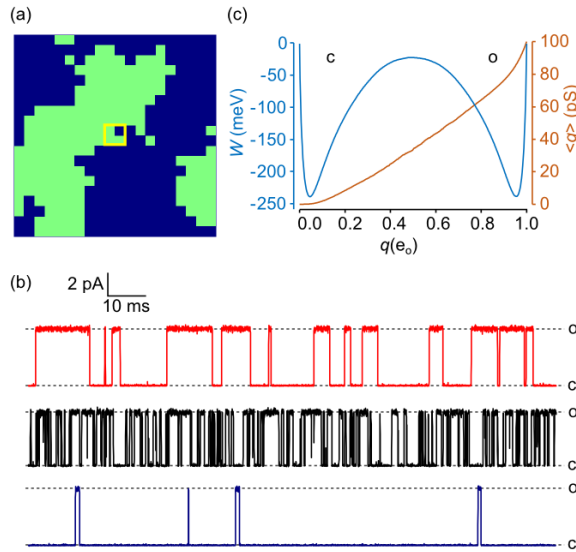


FIG. 1. (a) 20 x 20 electrical Ising model. Activated particles are colored green. An “ionic” current $i(t)$ was generated by specifying a central 2 x 2 “pore” region (yellow square) whose instantaneous conductance g is the sum of four 25 pS partial conductances. (b) Continuous-time Monte Carlo simulations 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 connecting microstates. These are constrained by detailed balance:

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

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. We chose a form based on activated kinetics:

$$\alpha_{u(i)} = \nu \exp\left(\frac{-\Delta\mathcal{E}_{u(i)}}{2kT}\right), \quad (3a)$$

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

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

$$\Delta\mathcal{E}_u = 2(2-u)\varepsilon - \Phi\delta q + U(q + \delta q) - U(q). \quad (4)$$

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

Our choice of microscopic 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

microscopic transitions when using a fixed simulation time step. We avoided this error by employing a continuous-time Monte Carlo simulation algorithm developed by Gillespie (Gillespie, 1977).

We chose for our reference model the 20 x 20 Ising grid ($N = 400$) with “standard” parameters: $T = 22$ °C; $\Phi = 0$ mV; $\varepsilon = 24$ meV; $q_{max} = 1$ eu; and $\nu = 5 \times 10^4$ kHz. The reference model was the basis for all calculations except where explicitly stated.

B. Diffusion states

We assign names to each successive level of coarse-graining. Every level is governed by a master equation with a complete set of (T, Φ) -dependent rate constants. The “Ising” model refers to the full 2^N configurational space (\mathbf{x}) determined by elementary forward and backward rate constants $\{\alpha_{u(i)}, \beta_{u(i)}\}$. The “2D” model is the product of coarse-graining onto the $(N+1)^2$ microcanonical space (E, q) . The 2D rate constants are recognizable as $\{a_{kk'}, b_{kk'}\}$. The “1D” model is the projection of the 2D master equation onto the $(N+1)$ reaction coordinate q with rate constants $\{a_m, b_m\}$. The 1D dynamics can alternatively be described by a continuous diffusion landscape $\{W(q), D(q)\}$, which may be further coarse-grained to a “sawtooth” discrete state model. Finally, the phenomenological two-state model is defined by forward and backward rate constants k and k_{-1} . The original Ising model is “microscopic”. The 1D and 2D representations are “diffusive”, and the sawtooth and two-state models are “macroscopic”, though “mesoscopic” has been used also (Qian, 2001; Kjelstrup et al., 2008). The Ising state space is intractably large and cannot be solved except by Monte Carlo methods, but the diffusive and macroscopic models lend themselves to numerical and analytical solutions of their master equations.

The coarse-graining procedure to create 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 large enough N , discrete and continuous approaches yield similar results. 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 $N \times N$ 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.

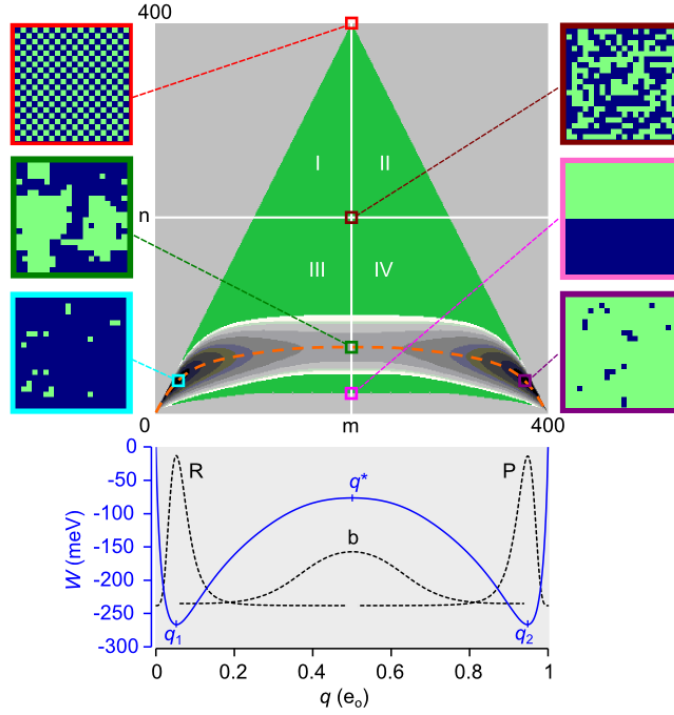


FIG. 2. Equilibrium course-graining of the 20×20 Ising system ($N = 400$). The green triangle is the set of microcanonical state on the (m, n) grid (73,448 states). Positive ε favors the lower portion (quadrants III and IV) of the triangle and positive Φ favors the right half (quadrants II and IV). The 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 sides 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 for 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 following formula:

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

Eq. 5 is the small-system equivalent to the thermodynamic relation $W = E - TS - q\Phi$. Because topologically the 2D Ising model with periodic boundary conditions is a torus, we avoid 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. Projecting the 2D landscape $W_{nm}(T, \Phi)$ onto the 1D potential of mean force $W_m(T, \Phi)$ was performed according to:

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

C. Thermodynamics

The diffusion models (1D and 2D) faithfully preserve system thermodynamics (Fig. 3). From the 2D free energy landscape we can calculate the environmental sensitivities of mean charge displacement $\langle q \rangle$ and the heat capacities C_Φ and C_q . The zero-field values of these quantities escape the discontinuities at the critical temperature $T_c = \varepsilon/k\ln(1 + 2^{1/2})$ experienced by the infinite-particle Ising model (Yang, 1952; Onsager, 1944).

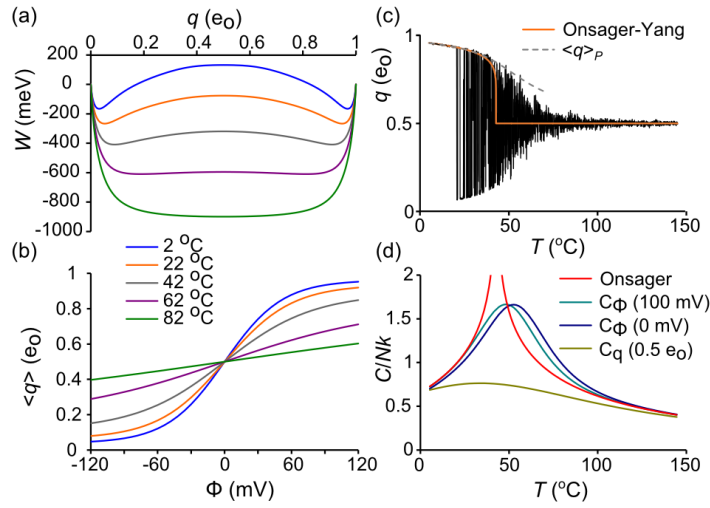


FIG. 3. Ising model thermodynamics. (a) $W(q)$ for different temperatures. The transition from bistable to centrally stable landscape roughly occurs at 73 °C. (b) $\langle q \rangle$ as a function of T and Φ . (c) $q(t)$ trajectory in response to 800 ms temperature ramp (5 °C to 145 °C) (ref). Filtering: $f_s = 10$ 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 reference Onsager solution (Onsager, 1944) demonstrates critical discontinuity at 42.84 °C.

D. Diffusive rate constants

The 2D rate constant from state $k = (m, n)$ to state $k' = (m', n')$ was computed for the (T, Φ) ensemble as the product of two numbers, the microscopic rate (Eq. 3) that contains the T - and Φ -dependence, and a branching coefficient $\langle r_{kk'} \rangle$, which is the statistical average taken over all k microstates of the number of transitions from k to k' (Fig. 4). Each $k \leftrightarrow (m, n)$ state contributes a maximum of five forward $(m + 1)$ transitions corresponding to $u = 0, 1, \dots, 4$:

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

There are also five backward rate constants $b_{kk'}$ 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.

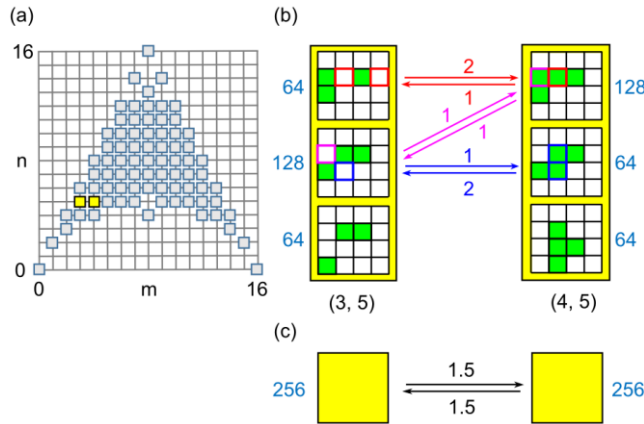


FIG. 4. (a) Microcanonical state space for the 4 x 4 Ising model. The (3, 5) and (4, 5) states are in yellow. (b) Microstates within (3, 5) and (4, 5) diffusion states. Green cells are activated. The number of microstates (Ω) are in shown in blue. Arrows indicate allowed transitions with color references 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 what effect this loss of information has on macroscopic relaxation, Metropolis Monte Carlo (MMC) was used to determine the 2D rate constants. Applying traditional MMC for $T \rightarrow \infty$, $\Phi = 0$ allows every trial move to be accepted, in principle yielding the correct branching coefficients provided the entire configuration space can be sampled. 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 obtain an estimate for $\ln \omega_k$. MMC was then implemented using the following procedure. A random particle is chosen for a trial flip, and 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:

$$r_n \leq \frac{\omega_k}{\omega_{k'}} , \quad (8)$$

If Eq. 8 is satisfied, 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). Initial particle states were randomly chosen (maximum entropy state). After an equilibration period of 2×10^6 time increments, 2D rate constants were time-averaged over the next 2×10^{10} trials using the following formula for forward rates:

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

and the analogous formula for backward rates $\langle b_{u(k)} \rangle$. The weighted sum preceding $\alpha_{u(k)}$ is the branching coefficient $\langle r_{kk} \rangle$. The normalized microcanonical density of states was obtained from:

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

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 $N \times N$ matrices $\ln \mathbf{\Omega}$, \mathbf{A}_u , and \mathbf{B}_u (Fig. 5). These eleven matrices are not independent, as, 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:

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

$$\sum_u B_u(m, n) = m \quad (11b)$$

$$\sum_u u A_u(m, n) = 2n \quad (11c)$$

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

The first line expresses the mirror symmetry between \mathbf{A}_u and \mathbf{B}_{4-u} , which is intrinsic to the structure of the Ising model. The last line describes detailed balance between opposing branching matrices \mathbf{A}_u and \mathbf{B}_u . The middle expressions were discovered empirically.

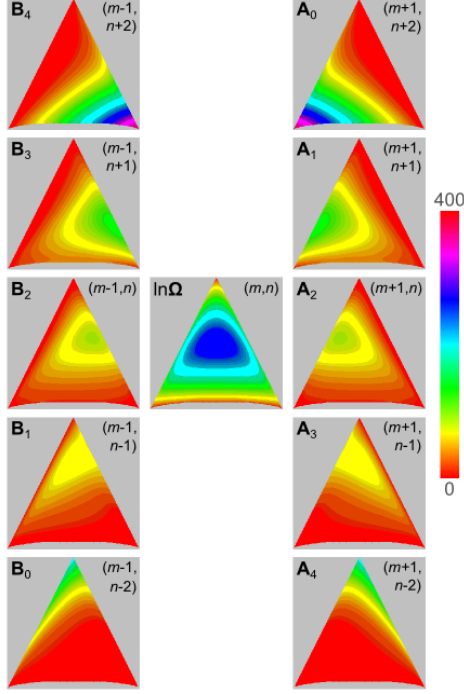


FIG. 5. Branching coefficients for the ten cardinal directions are stored in $N \times N$ matrices \mathbf{A}_u and \mathbf{B}_u . $N = 400$. The center matrix is the density of states $\ln\Omega$. The color scale applies to all matrices.

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:

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

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:

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

For regions of m where a_m and b_m are roughly linear, the expansion ends with Eq. 13. The free energy W_m is determined by detailed balance (Eq. 12). The Smoluchowski equation describes continuous 1D diffusion:

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

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

III. THE MACROSCOPIC RATE CONSTANT

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

To judge the fidelity of coarse-graining to full Ising kinetics, we required the “true” value of the forward rate constant k . We used brute-force simulation to generate a long trajectory with thousands of $R \rightarrow P$ transitions and stored the reactant state dwell times. Experimental dwell times are the intervals between large-scale transition events and are 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 obtained 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 4 seconds of simulation time was completed. The same algorithm was also used to simulate coarse-grained master equation dynamics in 1D and 2D. The first passage times were averaged to calculate the mean first passage time (mfpt), which we named τ . For “rare” event crossings, $k = 1/\tau$ (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 readily apparent when viewing trajectories that have been filtered to simulate real

experimental data (Fig. 8a, b). The phenomenological two-state equation predicts an exponential distribution of reactant dwell times. The distribution of first passage times were binned and transformed to a Sigworth-Sine plot (Sigworth and Sine, 1987). The Sigworth-Sine plot possesses two noteworthy features: (1) it generates uniform fitting residuals; (2) the transformed distribution peaks at τ for easy visualization of the mean first passage time.

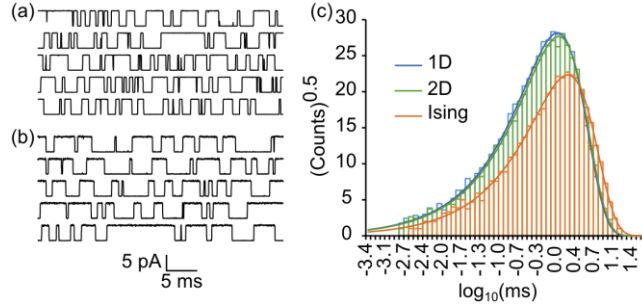


FIG. 6. Dwell time analysis. (a) “Ionic current” trajectories of the Ising model shown for visualization purposes. (b) Same as (a) except starting from 1D master equation. Filtering: $f_c = 10$ kHz, $f_s = 100$ kHz. (c) Sigworth-Sine plot of first passage times demonstrates fewer crossing events and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

In both Ising and coarse-grained models, binned first passage times were exponentially distributed across nearly four orders of magnitude in time. Fitting the single-transition distribution to the experimental dwell time histogram yielded a τ_{fit} that was very close to the statistical τ despite binning error (Table I). The two-state model is evidently a good phenomenological descriptor of macroscopic kinetics. The coarse-grained representations (1D and 2D) generated a significantly 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 over the microcanonical ensemble, though it maintains detailed balance, fails to predict the correct rate of macroscopic relaxation. We assigned the variable γ to the ratio of any coarse-grained value of k to the gold standard value obtained by brute force simulation. Much of the remaining paper concerns itself with validating and exploring the origins of γ , which for the reference model has a value of about 3/2.

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_{it} (ms)	1.675	1.146	1.088

^as.e. (τ) = $\tau/\text{events}^{1/2}$.

^bs.e. (k) = $(k/\tau)\text{s.e.}(\tau)$.

B. Numerically solving the master equation to obtain k

To confirm the findings from dwell time analysis, we computed k from coarse-grained master equations directly. In matrix form, the master equation is described by $d\mathbf{p}/dt = \mathbf{p}\mathbf{A}$, where $\mathbf{p}(t)$ is the row vector of state probabilities and \mathbf{A} is the matrix of rate constants a_{ij} for $i \rightarrow j$ transitions, with diagonal entries adjusted so that rows sum to zero. We note that some authors use a column-order format with transposed indices and matrix operations in reverse order. For the two-state model describing macroscopic dynamics, the matrix \mathbf{A} is:

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

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

Unlike $\mathbf{A}_{(\text{macro})}$, 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:

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

where the diagonal blocks **R**, **B**, and **P** represent the reactant, barrier, and product regions, separated on the charge 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** can be inverted since they represent open systems.

We calculated k from 1D and 2D master equations using eigenvalue and mfpt methods. The eigenvalue analysis computes the first non-zero eigenvalue λ_1 of **A**_(diff) and equates it to the macroscopic equivalent $-(k + k_{-1})$. 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 **A**_(diff) at the absorbing boundary q_2 to create an invertible matrix **R'** made from block matrices **R**, **C**, **D**, and **B**. The mfpt τ for $q < q_2$ is obtained (Van Kampen, 1992) by solving

$$\mathbf{R}'\boldsymbol{\tau} = -\mathbf{u}, \quad (17)$$

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

$$\mathbf{B}\boldsymbol{\pi} = -\mathbf{E}\mathbf{u}, \quad (18)$$

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 the reference model) and the flux J is evaluated for any dividing border Γ spanning the bottleneck region using (Berezhkovskii et al., 2009; Metzner et al., 2009):

$$J = \sum_{(u \rightarrow v) \in \Gamma} p_u^{eq} a_{uv} (\pi_v - \pi_u). \quad (19)$$

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

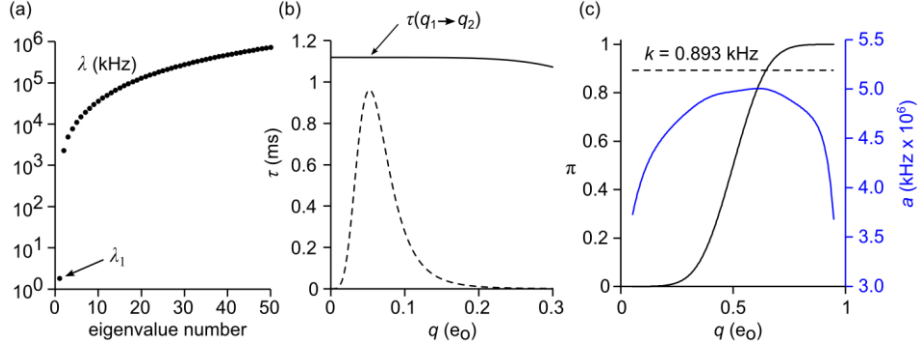


FIG. 7. Methods for calculating k from the 1D master equation. 2D-based methods are analogous. (a) Eigenvalues of \mathbf{A} , demonstrating the large spectral gap between first and second eigenvalues. (b) Mean first passage time (Eq. 17). The value of $\tau(q)$ is nearly constant for q with substantial probability density in the reactant basin (dashed line). (c) Committor analysis (Eqs. 18, 19) The committor probability $\pi(q)$ is given by the black line and the forward rate constant $a(q)$ is the 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 for the 2D model. From the $(N + 1)^2$ microcanonical states on the 2D grid, somewhat less than half the states are physically accessible (Fig 2), but even these are too numerous to handle conventionally. With $L = 20$, the size of the full transition matrix is 73,448 x 73,448. 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 $l \times 2d+1$. Typical values for l and d were $\sim 20,000$ 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 solving the 1D and 2D master equations, tabulated in Table II, are again consistent with a 50% increase in the value of k , or $\gamma \approx 3/2$, confirming the earlier dwell time analysis results shown in Table 1. Despite the agreement between the different numerical methods, it was important to rule out systemic biases that might 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 that could lead to erroneous Monte Carlo results (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 agreement. The MCG algorithm exhibited greater variability and systemic biases compared to the other three generators (Fig. 8), 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.

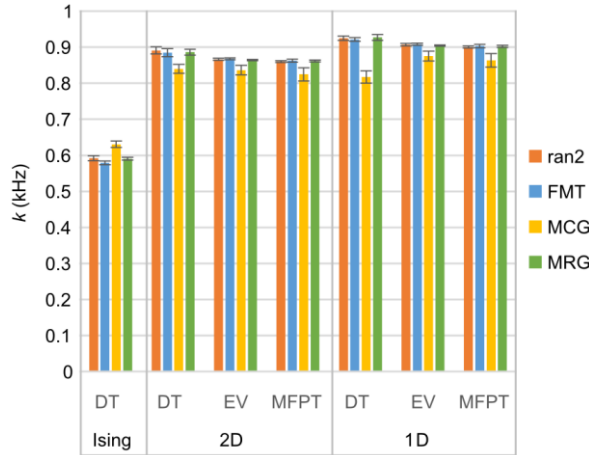


FIG. 8. Forward rate constant k (mean \pm s.e., $n = 6$) calculated for standard model with different methods (DT = dwell time, EV = eigenvalue analysis, MFPT = mean first passage time) and as a function of pseudorandom number generator. Results were consistent between generators except in the case of the multiplicative congruential generator (MCG), which demonstrated lower accuracy and precision compared to the other three generators.

The 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. In order to preserve detailed balance between opposing diffusive rates $a_{kk'}$ and $b_{k'k}$, we applied the following balancing procedure:

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

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

where a and b are the rate constants after balancing, and the equilibrium probability distribution p_k was assumed known. 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 balanced rate is: $(1 + \varepsilon_a + \varepsilon_b)^{1/2} - 1$. Thus systemic bias should be minimized if ε_a and ε_b are centered around zero. In practice, omitting the balancing step prior to dwell time analysis of the 1D model caused the value of k to rise 17% above the expected coarse graining error γ in one experiment, but this dropped to a positive 6% relative error if unbalanced rates were averaged over 30 experiments prior to analysis. The corresponding 2D single experiment relative error was 4%, which paradoxically increased to 7% with prior averaging. The findings suggest that omitting the balancing 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 balancing did have a profound effect on 2D eigenvalue calculations, with a 10-fold increase in unbalanced k estimates, although 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 required detailed balance as a strict requirement for symmetrizing rate matrices (Frehland, 1978). 1D eigenvalue analysis was error-free, as coarse-graining from 2D to 1D ensured detailed balance by design. 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 balancing 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 balancing 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 compared the gold-standard k value against eigenvalue, dwell time, and mfpt calculations for a range of N , T , and Φ values consistent with a 5-9 kT free energy barrier. The relationship between k and applied constraints can be predicted from the change in 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 increased 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).

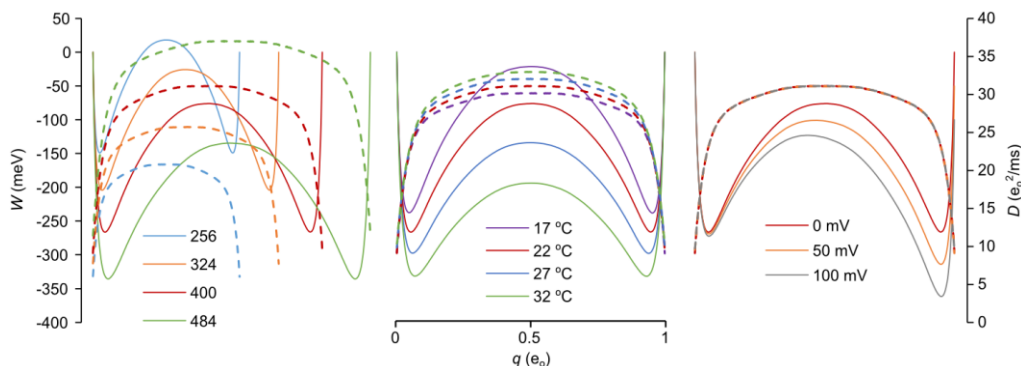


FIG. 9. Free energy (W , solid lines) and diffusion (D , dashed lines) landscapes as a function of N (left), T (middle), and Φ (right). The red curve in each plot corresponds to the reference model. Landscapes were calculated from matrices in Fig. 5, though identical results were obtained by time-averaging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

We examined whether the coarse-graining error γ was constant across environmental factors. This proved to be true within experimental error. 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.

Linear fits describing the relationships $\ln k = cX + d$, with X representing N , T , and Φ , resulted in nearly identical slopes c for the coarse-grained and gold standard data (Fig. 10). These slopes have physical significance. 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 $\Delta E = 1485$ meV determined from Boltzmann averaging $\langle E \rangle_b - \langle E \rangle_R$ of regional energy (Fig. 2). The fairly large activation energy 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. The difference of course is due to a large positive ΔS of activation. Similarly, 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 . We will see later that both Δq and ΔE vary substantially when computed for temperatures exceeding the limited range consistent with a 5-9 kT barrier height.

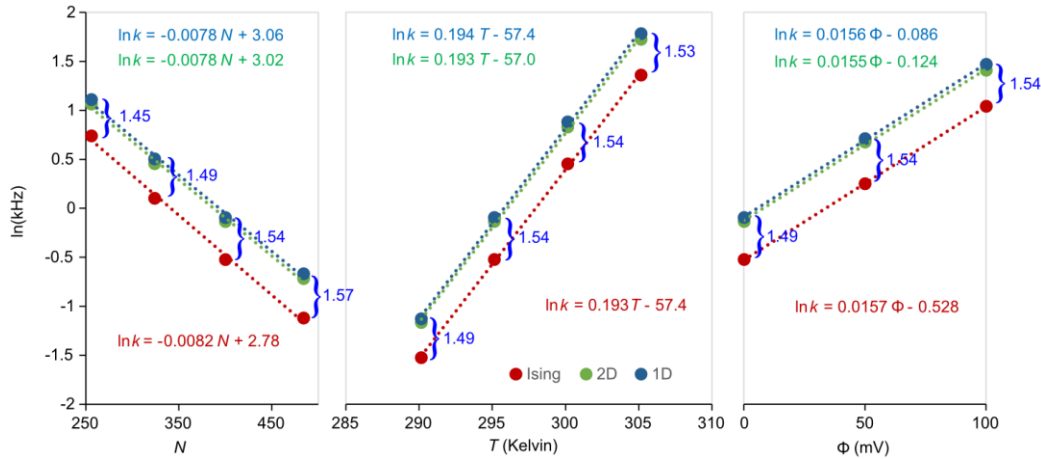


FIG. 10. Linear regression analysis of semi-log plot of k with environmental conditions same as in Fig 9. The Ising rate constants were calculated as the inverse means of reactant state dwell times, as described in the text. The 1D and 2D rate constants were averaged from their respective dwell time, 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 fitted data. Values of γ are in blue.

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 applied to the full Ising dynamics, which provided the gold standard value of k . The general agreement among the master equation methods strongly suggests that coarse-graining rather than numerical error is responsible for elevated γ values. To further evaluate this, we employed additional methods that did not rely on intermediate coarse-graining apart from determining the equilibrium variable $W(q)$. The quantities of interest in the new methods 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 very efficient calculations.

1. Reactive flux

The reactive flux method originated with molecular dynamics simulations (Chandler, 1978) but was expanded to include discrete-state models (Borkovec and Talkner, 1990). The method distinguishes between forward- and backward-directed trajectory ensembles initiated at the barrier separatrix, and follows the fractions of trajectories residing on the product side of the barrier. At time zero, only the positive flux ensemble is reactive, so that $f(0)$ equals the rate constant from transition state theory (TST). However, as a result of short-time barrier recrossings from both ensembles, $f(t)$ rapidly decays to a plateau value equal to k after a time T , which is the time required for the system to decay to one of the stable states. 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 for the kinetic Ising model, which is a continuous-time Markov jump process, is implemented as follows:

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

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 forward and backward ensembles of starting trajectories. The (+) ensemble reduces to 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:

$$f_{\text{absorbing}}(t \rightarrow \infty) = \frac{p_{eq}^* [a(q^*)\pi_{(+)}^* - b(q^*)\pi_{(-)}^*]}{P_R^{eq}}, \quad (22)$$

which may be compared to Eq. 20.

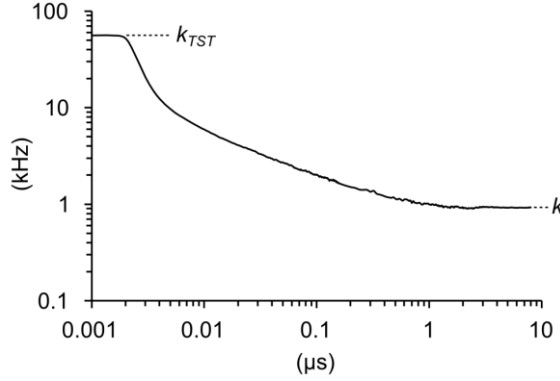


FIG. 11. Reactive flux $f(t)$ for reference model. Independent trajectories began with an 800 ns equilibration phase at the barrier peak, culminating with the crossing of the separatrix, followed by an 8 μ s test run from which k was obtained by averaging the last 3 μ s.

The result of six experiments, each containing 2×10^5 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 $\kappa = 0.017$, 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), from which the mfpt from q_1 to q_2 is described by a well-known double integral (Ansari, 2000):

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

Eq. 23 can be simplified if the transition barrier is large, since this minimizes the overlap between probability distributions in the reactant and barrier regions (see Fig. 2). As a result, the double integral reduces to a product of two integrals behaving like 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 final expression is:

$$k = \frac{D_b}{Z_R Z_b}, \quad (24)$$

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 earlier coarse-grained result using the mfpt method. Instead, we determined D_b independently by applying Nyquist's formula for current noise, $4D_b B = \langle i_g^2 \rangle_b$, to short, 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 was $B = 1.064f_c$, where f_c is the filter cut-off frequency. 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:

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

D_{mm} is the discrete-variable diffusion matrix evaluated at the saddle point region. We confined simulations to a hard window bordered by q -values 0.35 e_o and 0.65 e_o . The charge component of the diffusion coefficient was computed from the diffusion matrix through $D_q = D_{mm}\delta q^2$. Averaging over 6 x 500 Ising trajectories of 1.6 μs duration with $B = 1.06 \times 10^7$ 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 coarse-grained results.

The complete diffusion matrix was found to be:

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

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 actual relaxation time depends on the curvature of the free energy profile. The curvature matrix \mathbf{C} (Hessian) at the saddle point, obtained by fitting W_{mn} to a second order polynomial in m and n , is:

$$\mathbf{C} = \begin{bmatrix} -1 & 0.002 \\ 0.002 & 30.19 \end{bmatrix} |C_{mm}| , \quad (27)$$

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

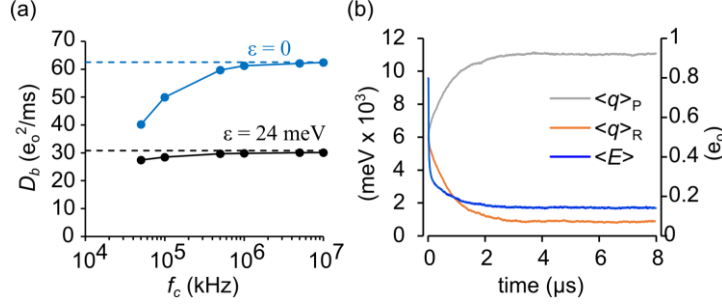


FIG. 12. Diffusion coefficients. (a) Evaluating D_q as a function of filter cut-off frequency. The reference model corresponds to black circles, which approach the peak value of $D(q)$ (dashed line) for large bandwidth. The blue circles represent the control case of non-interacting particles, where $D_q = 0.5N\nu\delta q^2$ (dashed line) is theoretically known. (b) Reference model $\langle E \rangle$ 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 \mathbf{e} of the unstable diffusive mode by solving the eigenvalue equation:

$$\mathbf{C}\mathbf{D}\mathbf{e} = -\lambda_+ \mathbf{e}, \quad (28)$$

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

F. Committor testing

In an attempt to detect the existence of other reactive collective variables, we performed committor analysis (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 $\{W(\pi(\mathbf{x})), D(\pi(\mathbf{x}))\}$ yields the correct value of k , but constructing the committor landscape from 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. Starting from the coarse-grained 1D or 2D models, extracting committors $\pi(q)$ and $\pi(q, E)$ by solving Eq. 18 is straightforward (Fig. 7c), but computing k by incorporating these committor values into Eq. 19 yields precisely the same incorrect value as the mfpt method. This suggests that coarse-graining distorts the committor function. We can test this directly by invoking the committor test for lossless coarse-graining:

$$\pi(\mathbf{x}) = \pi(q(\mathbf{x}), E(\mathbf{x})) \quad (29)$$

Eq. 29 states the committor of a configuration \mathbf{x} should equal the committor of \mathbf{x} projected onto the coarse-grained (q, E) space. We approached the problem in two ways. First, we mined configurations \mathbf{x} for which $\pi(\mathbf{x}) = 0.5$ and plotted them on the (q, E) landscape, where the symmetry of the reference model requires that $\pi(q^*) = 0.5$. Second, we tested configurations located on the separatrix (q^*, E) 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 $q = 0.3 \text{ e}_o$ and ending at $q = 0.7 \text{ e}_o$. 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 (0.469 to 0.532). A total of 14,837 isocommittor configurations were mined in this fashion and projected onto the (q, E) free energy surface. The isocommittors clustered around the saddle point region with a unimodal distribution (Fig. 13a). The absence of a multimodal pattern argues against additional reactive collective variables being critical to reactivity.

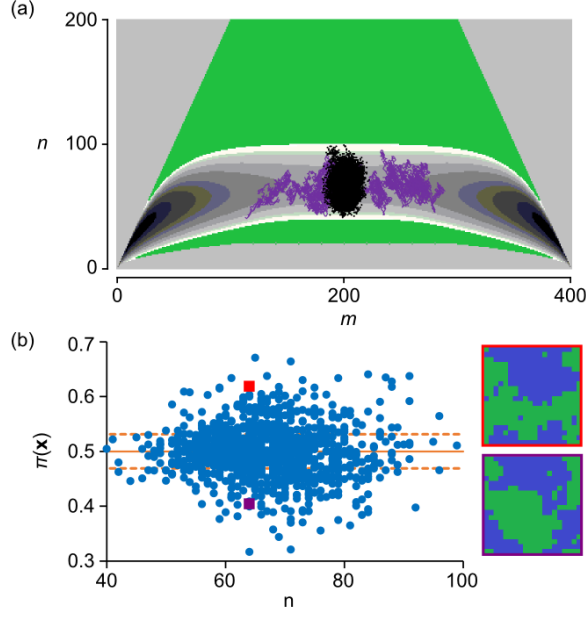


FIG. 13. Committor analysis. (a) 2D distribution of $\pi = 0.5$ configurations (black dots) with 95% confidence interval [0.469, 0.532]. A typical crossing trajectory is shown in purple. (b) Committor probabilities of configurations harvested from the separatrix ($m = 200$), evaluated from 10^3 test 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 to q_1 or q_2 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). In outlying configurations, an excess of short-lived cellular islands within the dominant cluster might explain the tendency to decay to the dominant side despite starting from the separatrix. The failure of coarse-grained states to predict the exact committor probability fundamentally violates the Markov condition since one cannot predict the future even in a probabilistic sense; therefore past events assume importance. Simply stated, coarse-graining introduces memory.

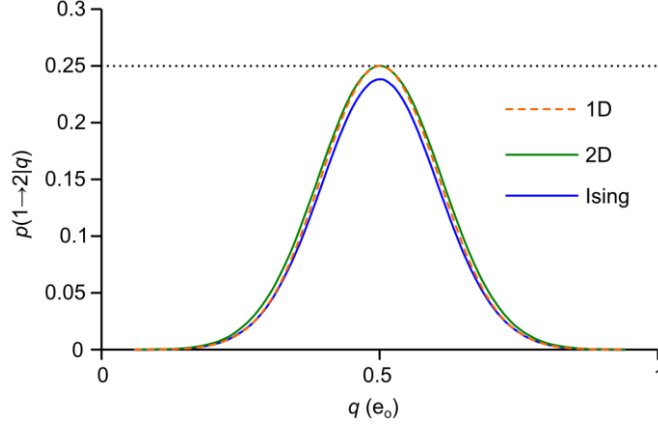


FIG. 14. Markovianity analysis demonstrating memory in Ising kinetics projected onto 1D, but not projections from 1D (control) or 2D kinetics.

To confirm the presence of dynamical memory, we took advantage of a simple test for Markovianity recently published (Berezhkovskii and Makarov, 2018). It involves evaluating probabilities of the sort $p(a \rightarrow b | q)$ for an ensemble of trajectories beginning at q_a and absorbed by q_b , crossing the test point q at some intermediate time. Using the basin minima q_1 and q_2 as our endpoints, we accumulated statistics for the possible pathways: $p(1 \rightarrow 1 | q)$, $p(1 \rightarrow 2 | q)$, $p(2 \rightarrow 1 | q)$, and $p(2 \rightarrow 2 | q)$, using a single unbiased trajectory containing 5,000 barrier transitions. We concurrently recorded event tallies for every q spanning the interval q_1 to q_2 . Although the $p(\dots | q)$ are correlated between neighboring q , running the simulation multiple times ($n = 14$) yielded good statistics for the entire curve. The Markovianity condition requires $p(1 \rightarrow 2 | q)_{\max} = 0.25$. A smaller peak value implies memory effects. A larger peak 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 the 1D model 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. Nevertheless, we propose that the observed memory accounts for the elevated γ value. An added bonus of the Markovianity test is that, since $\pi(q) = p(1 \rightarrow 2 | q) + p(2 \rightarrow 2 | q)$, tallied event numbers can be used to generate committor probabilities. Measuring $\pi(q)$ from 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 the level of coarse-graining. Only one method, brute force determination of dwell times, was a pure configurational process (\mathbf{x}) akin to performing an experiment; 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 “model-free” diffusion methods (reactive flux and Nyquist analysis). Their combined average was 0.896 ± 0.007 kHz, yielding a γ value of 1.51 ± 0.02 . This ratio 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 the information loss is incurred by projecting Ising kinetics onto the (q, E) plane.

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

Method	Process	k_{12} (kHz)	standard error
Dwell time (Ising)	\mathbf{x}	0.59159	0.006895
Dwell time(2D)	(q, E)	0.891089	0.009467
Dwell time (1D)	q	0.92513	0.005992
eigenvalue (2D)	(q, E)	0.866373	0.003352
eigenvalue (1D)	q	0.906609	0.003512
mfpt (2D)	(q, E)	0.859934	0.002575
mfpt (1D)	q	0.90043	0.002721
Nyquist (Ising)	$q(\mathbf{x})$	0.892715	0.003941
reactive flux (Ising)	$q(\mathbf{x})$	0.921411	0.073493
1D committor (Ising)	$q(\mathbf{x})$	0.907928	0.000948
1D committor (2D)	$q(q, E)$	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 ($5\text{--}9\text{ }kT$) 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 (q, E) 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 preceding raises the question of what memory model can be used 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 in the Ising system 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 equation that phenomenologically describes the kinetic Ising model.

The two diffusion methods that do not rely on prior coarse-graining (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 memory effects in inertial systems reside in the transmission coefficient. In the non-inertial, large-friction Ising model, reactive flux experiments revealed a different dynamic. The error in k 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 in the Grote-Hynes formalism, memory effects in the Ising model do not appear in the transmission coefficient.

Like the reactive flux method, Nyquist analysis, which determines the value of the diffusion coefficient D_b in the critical barrier region, overestimates the true rate constant k , and by the same empirical factor γ . The error in this case resides in the value of D_b , which coincides precisely with the coarse-grained estimate of the transition state diffusion coefficient $D^* = (1/2)(a^* + b^*)\delta q^2$. We hypothesized that spectral decomposition of D_b might reveal a reactive component leading to the correct value of k . To this end, we examined the stationary autocorrelation function of the gating current, which can be computed from the master equation rate matrix \mathbf{A} (Sigg et al., 1999):

$$\langle \dot{q}(0)\dot{q}(t) \rangle_{eq} = \chi \delta(t) + \sum_r \sigma_r \exp(\lambda_r t), \quad (30)$$

where $\sigma_r = -\lambda_r^2 \langle \mathbf{u}_0 \mathbf{Q} \mathbf{v}_r \rangle \langle \mathbf{u}_r \mathbf{q} \rangle$ is the spectral amplitude made up of eigenvalues (λ_r) and left- and right- eigenvectors (\mathbf{u}_r , \mathbf{v}_r) of \mathbf{A} multiplied by charge q in diagonal matrix (\mathbf{Q}) and vector (\mathbf{q}) form. The autocorrelation function is the sum of real negative amplitude decay components—except for the zero time delta term, which is positive-valued. This is in contrast to inertial systems whose velocity autocorrelation function is composed of positive-value real components, with the possibility of memory-induced imaginary components that generate dampened oscillations. The delta weight χ adds no information not already present in the exponential term and so may be ignored. The parameters σ_r and λ_r combine to determine the filtered one-sided power spectrum (Sigg et al., 1999):

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

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

$$\langle i_g^2 \rangle = 4 \sum_{r=1}^{r_{\max}} \frac{\sigma_r}{\lambda_r} B. \quad (32)$$

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

$$D = \sum_{r=1}^{r_{\max}} \frac{\sigma_r}{\lambda_r}. \quad (33)$$

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

$$k = \frac{D_1}{Z_R Z_b}, \quad (34)$$

where $D_b/D_1 \approx 3/2$. However, this ratio cannot hold for all conditions, since for very large temperatures the Ising model resembles a purely entropic system resembling the Ornstein-Uhlenbeck process, with γ approaching unity (Fig. 12a). But within the narrow range of experimental (T, Φ) consistent with modest-size free energy barriers, the spectral ratio (D_1/D_b) roughly coincides with the experimental γ value (Fig. 15f).

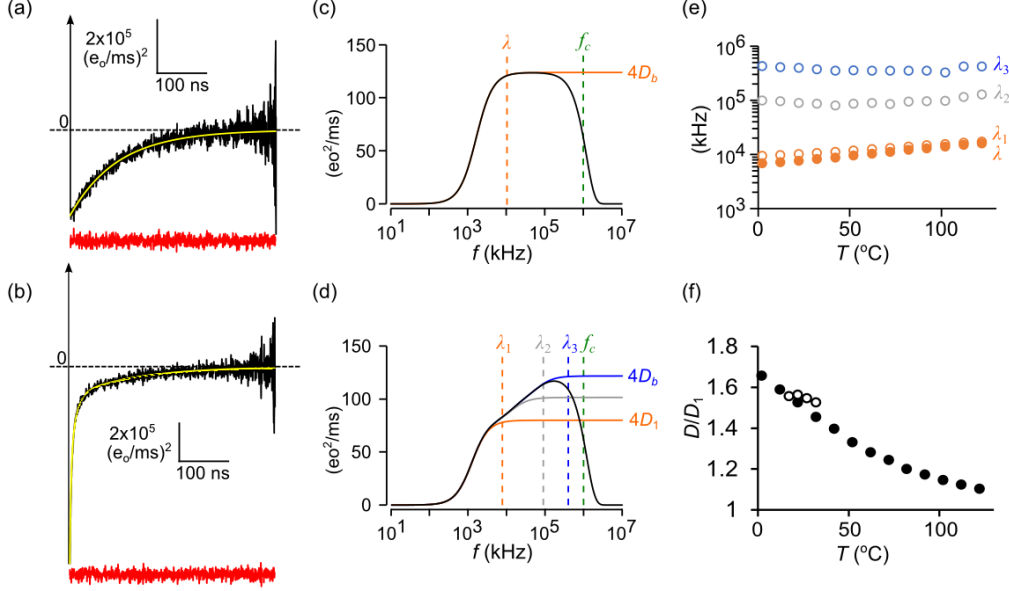


FIG. 15. Spectral decomposition of the diffusion coefficient for the transition state. (a) Gating current autocorrelation function for the 1D model (black line) fit to a single exponential (yellow line). The red trace is the fitting residual after normalizing for count statistics. Sampling and filtering cut-off frequencies were: $f_s = 10^7$ kHz; $f_c = 10^6$ kHz. (b) Same as (a) except with Ising kinetics requiring a three-exponential fit. (c) Power spectrum of the 1D model calculated from the fitted data. Values for the decay rate (λ) and the cut-off frequency (f_c) are marked by dashed vertical lines. (d) Same as (c) except for Ising kinetics, where three decay rates shape the power spectrum. (e) Ising (solid) circles and 1D (open circles) decay rates as a function of temperature. (f) Ratio D_1/D_b as a function of temperature (filled circles). This is compared to experimental values of γ (open circles). Each data point in (e, f) is the average of 6 experiments or simulations.

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

We end by proposing a general strategy, motivated by the preceding developments, for numerically compute k as a function of T and Φ . The procedure involves projecting Ising dynamics onto the microcanonical space (q, E) , then solving the coarse-grained master equation, and compensating for the ensuing systematic error γ by finding the

slowest (reactive) spectral component of the barrier diffusion coefficient. The last part is speculative but makes intuitive sense and gives numerically correct answers for small barrier (T, Φ) values. The overall strategy is summarized by Eq. 34, which is derived from the mfpt of rare-event crossings of a barrier and includes the reactive diffusion component D_1 . We provided evidence that the error in coarse-graining occurred primarily while projecting Ising dynamics onto (q, E) space, and the source of the error was due to loss of system memory rather than obviously poor choices in collective variables. Further projection of (q, E) space onto q -space introduced only minor error; therefore q was considered to be a good reaction coordinate for a 1D diffusion landscape 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. The product of reactant and barrier factors can be rewritten: $\exp(-\Delta W/kT)$, 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 remaining 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 $x = \{R, b, P\}$. 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 exponent corresponds to barriers and the negative exponent to basins. We defined the reactant basin to have range $[0, q^*]$ and the barrier region with range $[q_1, q_2]$. The react and barrier ranges overlap, but for well-separated regional probability distributions (Fig. 2), this is of little consequence when computing ΔW . The above procedure describes the coarse-graining of the 2D pmf to a discrete “saw-tooth” landscape described by (q, W) coordinates in a way that preserves the thermodynamic contribution to the rate constant, but does not require basins or barriers to be parabolic or symmetric (Fig. 16a).

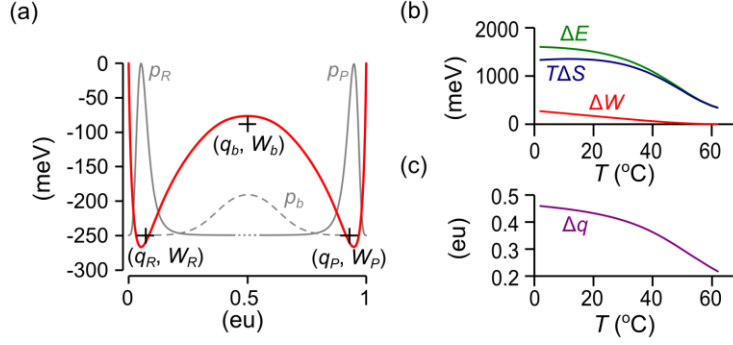


FIG. 16. Discretization of 1D diffusion landscape (reference model). (a) Black crosses marking regional reaction coordinates (q_x, W_x) form the basis of a “saw-tooth” landscape. The coordinates are superimposed onto $W(q)$ (red curve) and the respective segmental equilibrium probability distributions ($x = R, P$: solid gray curves; $x = b$: dashed gray curve). Note that the discrete coordinates do not necessarily coincide with the extrema of W . (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 tentatively equate $D_1(T, \Phi)$ to $D^*(T, \Phi)/\gamma$, 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 is nearly constant across the barrier region (Fig. 9), and therefore is constant for the purpose of rate equations such as Eq. 34. 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 have no effect on equilibrium 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 coarse-graining performed one time 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.
- 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.
- Hüller, 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.