**Microcanonical coarse-graining of the kinetic Ising model**

Daniel Sigg,1,a) Vincent A. 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**

We explore the use of the microcanonical ensemble in coarse-graining the dynamics of the 2-D kinetic Ising model. For subcritical temperatures, 2-D and higher-dimensional Ising lattices possess two basins of attraction separated by a free energy barrier. The advantage of the microcanonical ensemble is that the full environmental dependence of crossing rate constants can be obtained from a single Monte Carlo trajectory. Using various methods, we computed the forward rate constant of the coarse-grained Ising model and compared this prediction with the true value obtained from brute force simulation. While coarse-graining preserves detailed balance, the computed rate constant for barrier heights between 5 and 9 *kT* was consistently 50% larger than the true value. Markovianity testing revealed loss of dynamical memory, which we propose accounts for coarse-graining error. Committor analysis did not support the alternative hypothesis that projection onto microcanonical variables fails to produce an optimal reaction coordinate. The correct crossing rate was obtained by spectrally decomposing the diffusion coefficient near the free energy barrier and selecting the slowest (reactive) component. The spectral method also yielded the correct rate constant in the 3-D Ising lattice, where coarse-graining error was 5% and memory effects were diminished. We conclude that microcanonical coarse-graining supplemented by spectral analysis of short-term barrier fluctuations provides a comprehensive kinetic description of barrier crossing in a non-inertial continuous-time jump process.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **Electronic mail:** dansigg@gmail.com

**b) Electronic mail: vincenzo.carnevale@temple.edu**

**I. INTRODUCTION**

Metastability is a ubiquitous phenomenon occurring both in natural (e.g. enzymes, lipid membranes, etc.) and in manufactured systems such as MOSFET transistors1. The dynamics of a bistable system are characterized by random transitions between reactant and product basins. Proportionally little time is spent crossing the dividing free energy barrier, as bottleneck crossings are “rare” events. A biological example of a tunable multistable system is the ion channel, referring to a class of transistor-like membrane proteins2 whose pore components fluctuate (“gate”) between closed and open states with rates that are sensitive to environmental factors such as voltage () and temperature (*T*). Generally speaking, regulatory proteins such as ion channels contain on the order of 102 to 103 interacting residues, which classifies them as “small” from a thermodynamic standpoint3. Recent efforts have produced a diffusion model gating in the (, *T*) ensemble by coarse-graining the very large configuration space **x** of an ion channel onto a more manageable set of collective variables4. This yields a discrete master equation or diffusion landscape that can be used to calculate barrier-crossing rate constants5,6. However, the approach is incomplete in the sense that it must be repeated for new values of  and *T*.

We propose a comprehensive solution by projecting the full molecular dynamics onto the microcanonical ensemble, whose natural variables are charge (*q*) and energy (*E*). The projection was done through Monte Carlo sampling of the configurational space and averaging microscopic frequencies between (*q*, *E*) states. We name the frequencies branching coefficients, since they represent the number of pathways available for a microstate to transition to other microstates in a neighboring (*q*, *E*) state. Branching coefficients are the statistical weights needed to transform the system into a solvable master equation with (, *T*)-dependent rate constants. The gating charge *q* assumes the role of order parameter by distinguishing between closed (reactant) and open (product) states. If *q* is a “good” reaction coordinate7, then projection onto a one-dimensional diffusion landscape described by a *q*-dependent potential of mean force *W*(*q*,, *T*) and associated diffusion coefficient *D*(*q*,, *T*) should theoretically yield the correct value of the forward (or backward) barrier-crossing rate constant.

We chose as a proof-of-concept system a toy lattice model with broken symmetry states, specifically an electrical isomorph of the traditional square Ising model operating at subcritical temperatures. The existence of upper bounds in both *E* and *q* parameters allows the complete state space to be projected onto a finite microcanonical grid8–11. The number of lattice particles *N* = *L*2, where lattice length *L* was restricted to the set of even values: 16, 18, 20, and 22, is comparable to the number of residues in a biological macromolecule. The small *N* permits a large number of brute force transition events to be simulated in a short time. Previous studies12–15 have employed larger systems (*L* = 24 to 100) and lower temperatures (*T*/*Tc* = 0.6 to 0.8, compared to > 0.9 in this study) in order to evaluate Ising kinetics in the context of classical nucleation theory. But classical variables such as nucleation volume, surface area, and anisotropy16 are not well defined if the critical cluster size exceeds system boundaries, as occurs in small systems. Our focus was to compute rate constants across modest free energy barriers (5 to 9 *kT*) as a function of  and *T*. With modest free barriers, susceptibilities are softened, resembling response curves seen in regulatory proteins, but phenomenologically precise two-state kinetics are observed. The main advantage in keeping free energy barriers relatively small is that a large number of unconstrained “brute-force” transition events can be recorded in an effort to obtain good statistics on “true” crossing rates, which can then be compared to predicted rates derived from coarse-grained models.

**II. ELECTRICAL ISING MODEL**

1. **Configurational structure and dynamics**

The “enthalpy” *H* of the electric Ising model, expressed below, is the sum of configurational energy *E* and a work term *q* that couples total charge *q* to an external field .

****, (1)

The energy term, the first term in Eq. 1, sums over nearest-neighbor “gating” particle interactions 〈*i, j*〉, where each particle *i* assumes two flip state values: *ei* = 0, 1. Adjacent particles with differentstates(symbolized by the exclusive-or operator ⊗) increase *E* by a positive interaction energy **. Assuming periodic boundary conditions, *E* increases in increments of 2** from 0 to 2*N*. The “gating” charge *q* (the quantity in parenthesis) changes by an amount *δq* during activation and *δq* for the reverse transition, thus increasing in *δq* increments from 0 to *Nδq*.It is 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 electric Ising model is isomorphic to the traditional magnetic Ising model, all results here apply also to the traditional model. We note that the interaction term ** in the proposed model is twice the value of the traditional *J* factor.

We elected to work with the electric Ising model because of its dynamical similarities to gated ion channels (Fig. 1). Ion-conducting pore units in channels demonstrate weak voltage-sensing behavior even in the absence of canonical voltage sensors17,18. In the Ising model, large-scale fluctuations in a manufactured “ionic current” *i* conducted through a small “pore” region (Fig. 1a) are qualitatively similar to those recorded in single ion channels, including a characteristic voltage- and temperature-dependence (Fig. 1b). The mechanism for intrinsic pore voltage sensitivity in ion channels is incompletely known, but it presumably involves long-range correlations in the secondary structure. In this study, we focused on the gating charge analog *q* because it is a canonical thermodynamic variable whereas *i* is useful mainly as an experimental marker of activation (apart from its substantial physiological role). We note however that, on the millisecond time scale, *q* and *i* produce matching random telegraph signals.

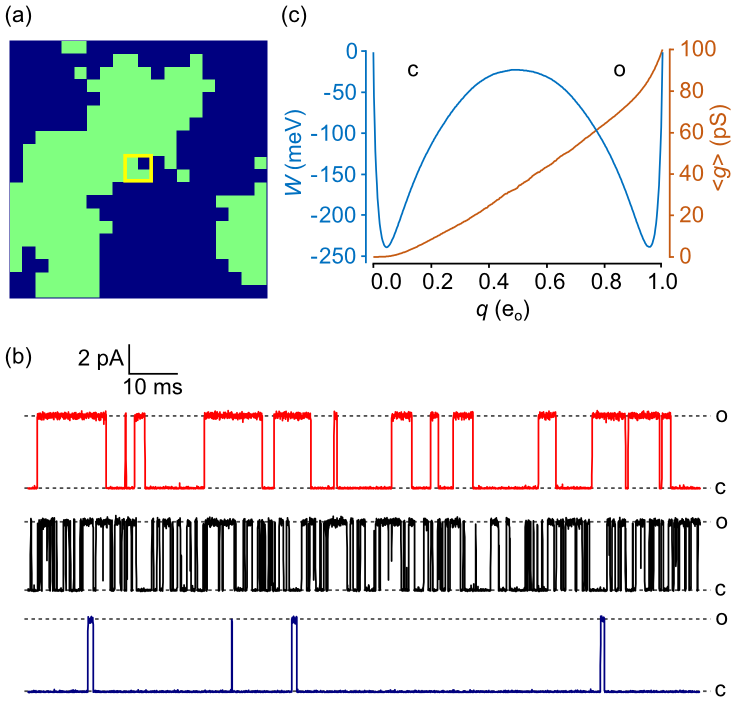


FIG. 1. (a) 20 × 20 electric Ising model with artificial pore. Activated particles are colored green. An “ionic current” *i*(*t*) was generated by specifying a central 2 × 2 square (in yellow) as the “pore” region, whose total conductance *g* depends on the activation states of the four inscribed particles which, when open, contribute each 25 pS of partial conductance. (b) Continuous-time Monte Carlo simulations of the Ising “pore” with ionic current governed by Ohm’s law: *i* = *g*(*V* – *Vrev*). Conductance levels transition between closed (c) and open (o) states. The red trace (top) is the reference model with *rev* (reversal potential) = –60 mV. The current was digitally filtered19 with cutoff frequency *fc* = 10 kHz and sampling rate *fs* = 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 by a decrease of  from 0 mV to -100 mV (blue trace, bottom). (c) 1D free energy landscape *W*(*q*) and mean conductance 〈*g*(*q*)〉 obtained as time-averaged quantities from a 0.4 sec trajectory.

The first step in creating a kinetic Ising model was to define the rate constants between microstates. These are constrained by detailed balance:

, (2)

where *i* and *i* are the forward and backward rate constants for activating the *i*th particle. A variety of expressions for *i* and *i* are compatible with Eq. 2. We based our choice on activation kinetics:

, (3a)

, (3b)

where the energy of activation ** increases linearly with (1) the number *u* of activated neighbor particles, (2) the voltage , and optionally, (3) an applied umbrella potential *U*.

. (4)

The pre-exponential factor ** was assumed to be temperature independent. This establishes a convenient upper limit for diffusion at high temperatures.

An alternative expression to Eqs. 3 attributed to Glauber20 checks microscopic rates from exceeding ** when a large field is applied. The Glauber formulation avoids undersampling fast transitions when using a fixed simulation time step. We circumvented this problem by employing the continuous-time Gillespie21 Monte Carlo algorithm, which generates real-valued transition intervals.

Except when stated, calculations were performed on a reference model with the following parameters: *L* = 20; *T* = 22 °C;  = 0 mV; ** = 24 meV; *qmax* = 1 eo (electronic unit); and ** = 5 × 104 kHz. We note that at 22 °C, *kT* = 25.43 meV. Except when indicated, results were expressed as the mean value ± standard error from 6 sets of simulations.

1. **Coarse graining**

The Ising model and each tier of coarse-grained representations derived from the Ising model are governed by master equations with (, *T*)-dependent rate constants. The kinetics of the full 2*N* Ising space **x** is defined by elementary forward and backward rate constants {*u*(*i*), *u*(*i*)}. The first coarse graining tier is the “2D” microcanonical (*q*, *E*) landscape with (*N* +1)2 states connected by rate constants {*akk′*, *bkk′*}. The next tier, the “1D” landscape, is the projection of the 2D master equation onto the (*N* +1) space of the reaction coordinate *q*, with rate constants {*am*, *bm*}. Alternatively, 1D dynamics are described by a continuous diffusion landscape {*W*(*q*), *D*(*q*)}. The final level of coarse-graining is the two-state model described by phenomenological rate constants *k* and *k*-1. The full trajectory Ising kinetics are “microscopic”. The 1D and 2D representations are “diffusive”, and the two-state model is “mesoscopic” in keeping with usage by some authors for the stochastic kinetics of single macromolecules22,23. The term “macroscopic” is reserved for the deterministic average signal generated by many independent molecules. The full trajectory space ***x*** is intractably large and can be solved only through Monte Carlo methods, but the diffusive and mesoscopic models lend themselves to numerical and analytical solutions of their respective master equations.

The coarse-graining procedure is illustrated in Fig. 2. We alternate between continuous (*q*, *E*) and discrete (*m*, *n*) coordinates, where *q* = *mδq* and *E* = 2*n*. Discrete dynamics are governed by a master equation, while the continuous model is handled by a Smoluchowski equation. We used *N* values that were large enough for discrete and continuous approaches to yield similar outcomes. Coordinates (*m*, *n*) are readily determined for a given microconfiguration through simple formulas. The order parameter *m* is the total number of activated particles, and the energy integer *n* is equal to 2*m* – *b*, where *b* is the number of shared borders between activated *m* particles. The set of (*m*, *n*) states on the *N* × *N* microcanonical grid occupy a triangle-like space. The base of the “triangle” is concave with height *L*. This concavity gives rise to the characteristic symmetry-breaking properties of the square Ising lattice and higher order models, but is lacking in the linear lattice, where the minimum energy *n* = 1 for all *m* except the endpoints.

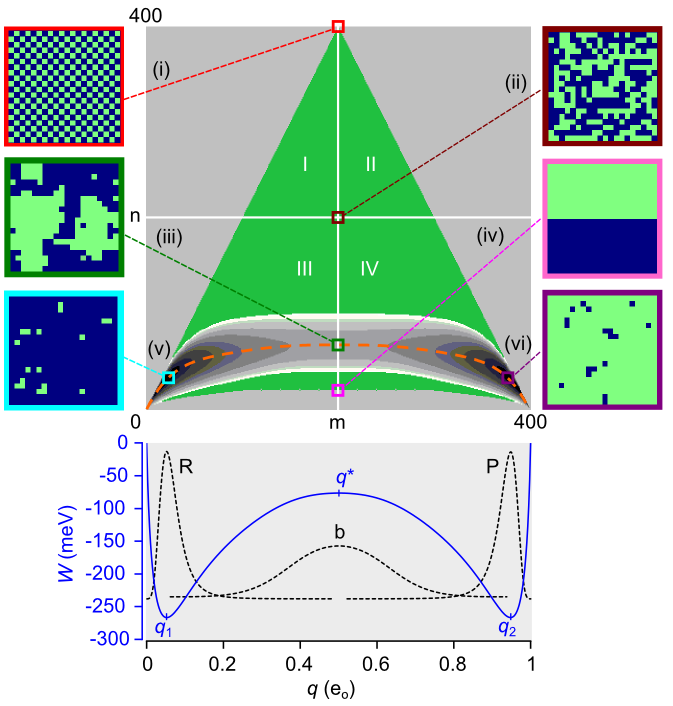


FIG. 2. Equilibrium course-graining of the 20 × 20 Ising system (*N* = 400). The green “triangle” is the set of microcanonical state on the (*m*, *n*) grid (73,448 states). Positive ** favors the lower portion (quadrants III and IV) of the “triangle” and positive  favors the right half (quadrants II and IV). The free energy landscape (22,149 states) limited by *Wmax* = 150 meV is shown as the 2D contour map near the base of the “triangle”. The vertical span (spatial bandwidth) of the free energy saddle point is 68 cells located at the separatrix (*m* = *N*/2). The dashed orange line is the mean energy 〈*E*〉. Representative microscopic configurations shown at the edges of the main plot: maximum energy (i); maximum entropy (ii); saddle point (iii); minimum transition state energy (iv); and the minimum free energies in metastable states R (v) and P (vi). The projected 1D free energy landscape *W*(*q*) is plotted in the lower graph. The dotted lines are equilibrium probabilities computed for the R, P basins, and for the barrier b after inversion. Local energy extrema are *q*1, *q*2, and *q*\* (the transition state).

The free energy *W*(*q*, *E,* , *T*) is a function of the density of states (*q*, *E*). In integer space, this is given by:

. (5)

Eq. 5 is the small-system equivalent to the thermodynamic relation *W* = *E* – *TS* – *q*. The square lattice subject to periodic boundary conditions is topologically a torus, which eliminates end effects. If ** is positive, states in the lower portion of the “triangle” are populated (Fig. 2). This corresponds to ferromagnetism/diamagnetism in the traditional Ising model. Anti-ferromagnetism occurs at negative **, which populates the upper portion of the “triangle”. In order to limit the accessible (*m*, *n*) states to a manageable number (see Supporting Information), we established a cut-off value *Wmax* for the free energy. This resulted in a 2D free energy landscape with reactant R and product P basins separated by a broad, narrow and slightly curved bottleneck region (Fig. 2). We projected the 2D landscape *Wnm*(, *T*) onto the 1D potential of mean force *Wm*(, *T*) by summing over *n*:

. (6)

1. **Thermodynamics**

The 1D and 2D diffusion landscapes faithfully preserve thermodynamic properties of the Ising model (Fig. 3). From the 2D landscape we can calculate the environmental sensitivities of mean charge displacement 〈*q*〉 and heat capacities *C* and *Cq*. These quantities do not suffer from the critical discontinuities that characterize the infinite-particle Ising model24,25.

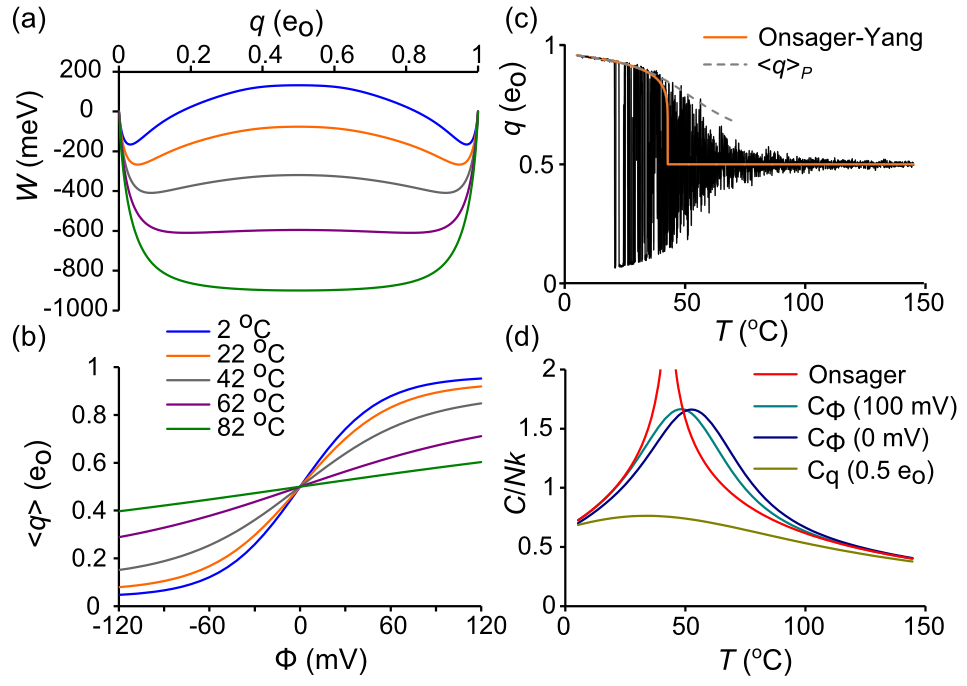


FIG. 3. Ising model thermodynamics. (a) *W*(*q*) for different temperatures. The transition from bistable to centrally stable landscape roughly occurs at 73 °C. (b) 〈*q*〉 as a function of *T* and . (c) *q*(*t*) trajectory in response to an 800 ms temperature ramp (5 °C to 145 °C)26. Filtering: *fs* = 10 kHz, *fc* = 1 kHz. The red line is the Onsager-Yang solution24 and the dashed curve is 〈*q*〉*P*. (d) Heat capacities derived from *C*/*k* = **2〈(*E*)2〉. The Onsager solution25 demonstrates a critical discontinuity at *Tc* = **/*k*ln(1 + 2½), which equals 42.84 °C in our reference model.

1. **Diffusive rate constants**

The 2D diffusion rate constants{*akk′*, *bk′k*} connecting states *k* = (*m*, *n*) and *k*′ = (*m*′, *n*′) were computed in the (, *T*) ensemble as the product of two numbers: the microcanonical branching coefficient 〈*rkk′*〉, which averages the number of available pathways for microscopic transitions between *k* and *k*′ (Fig. 4), and the microscopic rate (Eq. 3), which contains the - and *T*-dependence. Each *k* ↔ (*m*, *n*) state contributes up to five forward (*m* → *m*  1) transitions corresponding to the number *u* of activated-neighbors:

. (7)

For each forward transition there is a corresponding backward (*m* → *m*  1) rate constant *bk′k* = 〈*rk*′*k*〉*u*(*k*′).

Microcanonical coarse-graining yields an exact kinetic solution only in the case of the mathematically trivial case of *L* = 2, where averaged branching coefficients are integers. For *L* > 2, there is inevitable loss of kinetic information, though detailed balance, defined by *k akk′* = *k′**bk′k*, remains satisfied.

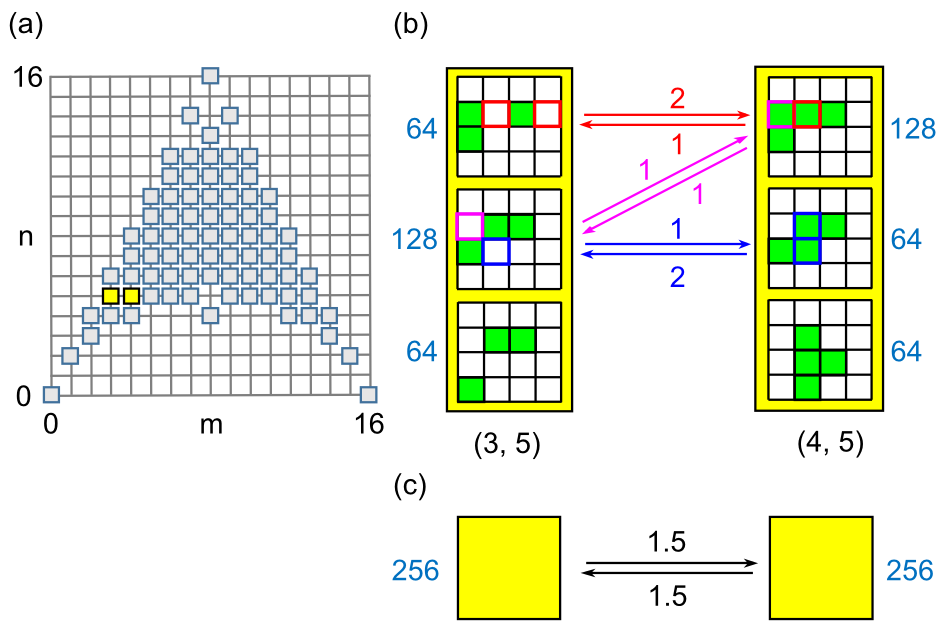


FIG. 4. (a) Microcanonical state space for the 4 × 4 Ising model (*N* = 16). The diffusion states (3, 5) and (4, 5) are highlighted in yellow. The highlighted states are expanded in (b) to show their microstate structures. Activated cells are colored green. The microstate density  for each configuration is in written in blue. Allowed transitions between the (3, 5) and (4, 5) states are labeled with branching coefficients color-matched to the transitioning particle(s). (c) The coarse-grained (3, 5) ↔ (4, 5) transition after statistical averaging. Here, averaged opposing branching coefficients 〈*r*〉 have the same value, consistent with detailed balance between states with equal .

To determine how loss of information affects crossing rates, 2D rate constants were acquired using Metropolis Monte Carlo (MMC). Applying traditional MMC for *T* → ∞,  = 0 allows every trial move to be accepted. This would in principle yield the correct branching coefficients, but the large entropy gradient makes this algorithm impractical for all but the smallest systems, so a bias potential *k* was applied to “flatten” the gradient and allow uniform sampling over all (*m*, *n*) states. The “1/t” variant of the Wang-Landau algorithm27 with endpoint refinement parameter *Ffinal* = 10-6 was used to rapidly estimate ln*k*. MMC was then implemented in the following manner. At each integer time *s*, a random particle was chosen for a trial flip. After determining the (*m*, *n*) coordinates of the current *k* state and the trial *k*′ state, a uniform random number 0 < *rn* < 1 was drawn and the following acceptance criterion for the *k* → *k*′ transition was applied:

, (8)

If successful, the particle was flipped, and regardless of outcome, *s* was incremented and the cycle repeated. For every iteration, prior to choosing the trial particle, branching coefficients *rkk*′(*s*) were summed by flipping every particle in turn, categorizing each particle’s transition into one of the ten cardinal directions (*m*, *n*) → (*m* ± 1, *n* ± (2 – *u*)), *u* = 0…4, before returning it to its original position. If, for example, at time *s* there are *b*/*N* trial flips that contribute to the *k* → *k*′ transition, then *rkk*′(*s*) = *b*, and so forth for the remaining nine directions. After an equilibration period of 2 × 106 steps starting from a random distribution, branching coefficients were time-averaged over the next 2 × 1010 trials using the formula:

 (9)

The microcanonical density of states was obtained from:

. (10)

The density of states and branching coefficients were stored in *N* × *N* matrices ln****, **A**u, and **B**u (Fig. 5). These eleven matrices are not independent since, neglecting numerical error, the entire group can be constructed from any three of its members, for example **A**0, **A**1, and **A**3. The following relations apply:

 (11a)

 (11b)

 (11c)

 (11d)

The first line expresses the mirror symmetry between **A** and **B** matrices as evident from Fig. 5. The last line is a statement of detailed balance. The two middle lines express relationships between summed branching coefficients and their (*m*, *n*) coordinates.

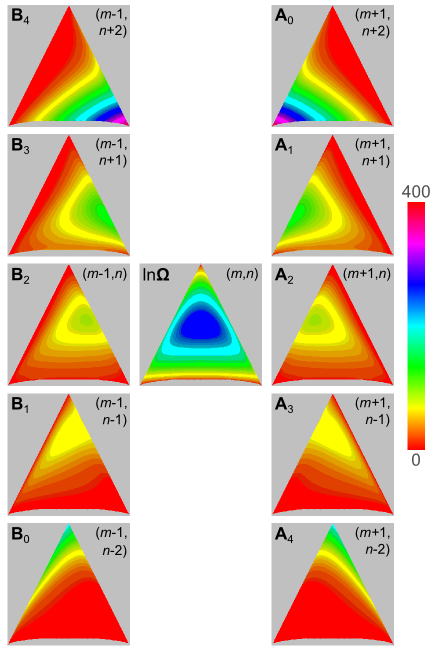


FIG. 5. Branching coefficients for the ten cardinal directions originating from an (*m*, *n*) state stored in *N* × *N* matrices **A***u* and **B***u*, where *u* is the number of neighboring particles with flip state 1. *N* = 400. The center matrix is the density of states ln****.

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

. (12)

For small increments *q*, the 1D master equation can be approximated by a continuum diffusion equation28. The relationship between discrete {*am*, *bm*} and diffusive {*Wm*, *Dm*} variables arises from Kramers-Moyal expansion of the master equation. The diffusion coefficient corresponds to the second Kramers-Moyal coefficient:

. (13)

The free energy *Wm* is determined by detailed balance (Eq. 12). In regions where *am* and *bm* are roughly linear, such as near the barrier, the expansion ends at second order and 1D diffusion is described by the Smoluchowski equation:

, (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 worked with the master equation since it is fundamental to Ising dynamics, but the diffusion equation, which assumes significance later, yielded similar outcomes.

**III. METHODS FOR COMPUTING TRANSITION RATE CONSTANTS**

1. **Dwell time analysis reveals systematic error after coarse-graining**

To judge whether coarse-graining conserves relaxation kinetics, we needed accurate values of the crossing rate *k* derived from the Ising and 1D/2D models. To this end, we used brute-force simulation to generate long trajectories for each model type. Reactant state dwell times from thousands of reactant-to-product (R → P) transitions were averaged to obtain the mean first passage time (mfpt), which we named **. If transitions are “rare” events, *k* = ** -1 29. In laboratory measurements, dwell times are typically subject to measurement error due to noise and bandwidth limitations30. However, using the continuous-time Gillespie Monte Carlo algorithm21, we obtained exact first passage times between *q*1 and *q*2. Briefly, to initiate a Monte Carlo step, state-dependent flip rates (*i* or *i*) were determined for each particle *i*, and sums of rates were grouped in the forward (∑*i*) and backward (∑*i*) directions. After drawing a uniform random number *rn*, the step interval was calculated from ln*rn*/(∑*i* + ∑*i*) A second random number selected the transitioning particle with probability proportional to its transition rate. The process was repeated until 4 s of simulation was completed, at which time several thousand first passage times were recorded. The Gillespie algorithm was also used to simulate 1D and 2D dynamics from their respective master equations.

The disparity in reactant dwell times between Ising and coarse-grained model trajectories *q*(*t*) is visually apparent in sample trajectories (Fig. 8a, b), but to accentuate the different distributions we log-transformed and binned first passage times into a Sigworth-Sine plot31. The Sigworth-Sine plot is helpful in two ways: (1) it generates uniform residuals when fitting data, and (2) the transformed distribution peaks at ** for easy identification of the mean first passage time.

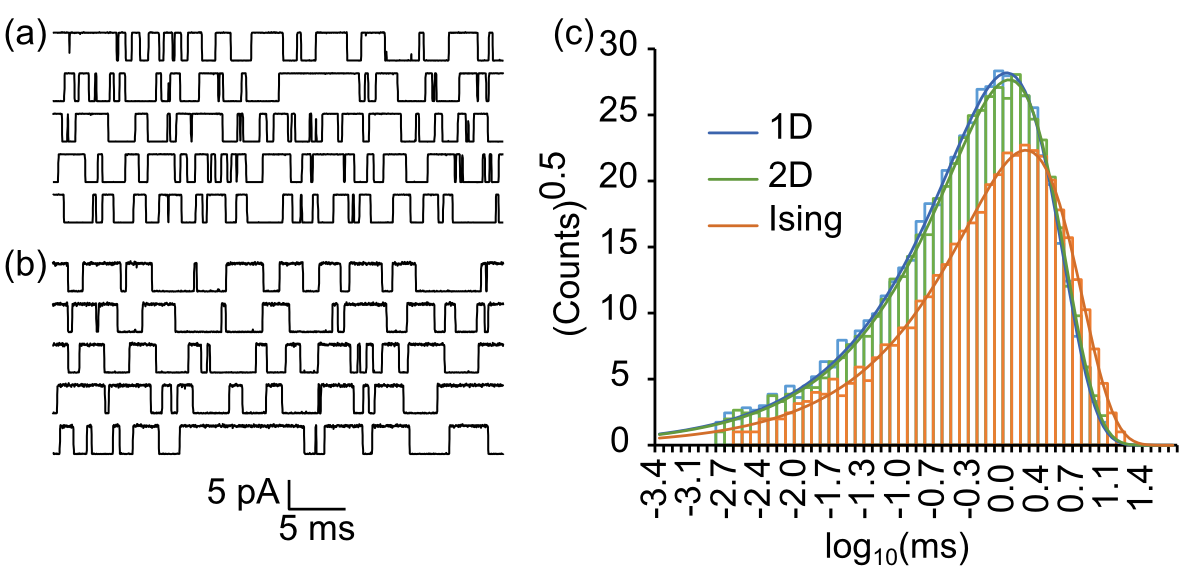


FIG. 6. Dwell time analysis. (a) “Ionic current” trajectories of the reference Ising model. (b) Same as (a) except starting from the 1D master equation. Filtering: *fc* = 10 kHz, *fs* = 100 kHz. (c) Sigworth-Sine plot of first passage times demonstrates fewer total crossing events and a longer average 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. Fitting the theoretical distribution to the experimental dwell time histogram yielded a *fit* that was very close to the average ** (Table I). The trajectories are evidently well-described by a two-state kinetic scheme. The coarse-grained models (1D/2D) generated a smaller ** than the Ising model by about a 2:3 ratio, thus overestimating the true rate constant *k* by 50%. When comparing the two coarse-grained models, the 2D model was only slightly more accurate than the 1D model. Therefore coarse-graining, while it preserves equilibrium, fails to predict the correct transition rate constant. The statistics of brute force dwell time analysis are sufficiently good for the method to be considered the gold standard for the true crossing rate. Other methods were judged by computing the ratio ** of their estimate for *k* to the true crossing rate. Much of the remaining paper concerns itself with validating and exploring the origins of the ** ratio, which calculated from dwell times has values 1.46 ± 0.02 (2D) or 1.54 ± 0.02 (1D).

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

|  |  |  |  |
| --- | --- | --- | --- |
|  | Ising | 2D | 1D |
| events | 6,776 | 10,397 | 10,791 |
| ** (ms) | 1.672 ± 0.020a | 1.145 ± 0.011a | 1.086 ± 0.010a |
| *k* (kHz) | 0.5982 ± 0.0073b | 0.8735 ± 0.0085b | 0.9210 ± 0.0088b |
| *fit* (ms) | 1.675 | 1.146 | 1.088 |

as.e.(**) = **/events½. s.e. = standard error

bs.e.(*k*) = (*k*/**)s.e.(**).

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

To validate the findings from dwell time analysis, we calculated *k* directly from the 1D and 2D master equations. The matrix formulation of the master equation is *d***p**/*dt* = **pA**, where **p**(t) is the row vector of state probabilities and **A** is the rate matrix whose diagonal entries ensure each row sums to zero. The row-major notation designates *aij* as the rate constant for the *i* → *j* transition (the reader should be aware that some authors use a column-order convention with transposed indices).

The **A** matrix for the mesoscopic two-state model is:

. (15)

The only non-zero eigenvalue of **A**(meso) is **1 = –(*k* + *k*-1). If  = 0, then *k* = *k*-1, and so both equal –**1/2. In cases where  ≠ 0, knowing **1 will still determine *k*, since *k*/*k*-1 can be found from detailed balance.

Unlike **A**(meso), rate matrices **A**(diff) for the diffusive (1D, 2D) master equations are large (~*N* × *N* or ~*N*2 × *N*2), requiring a numerical solution (see Supporting Information for details of the calculation). In block form, **A**(diff) looks like:

, (16)

where the diagonal blocks **R**, **B**, and **P** represent the reactant, barrier, and product regions, and **C**, **D**, **E**, and **F** are off-diagonal transition matrices. The basin minima *q*1 and *q*2 (Fig. 2) determine the partition boundaries. Though the rate matrix **A** for the entire system is singular, the regional matrices **R**, **B**, and **P** are invertible.

We obtained *k* by numerically solving the 1D and 2D master equations using eigenvalue and mfpt methods. In eigenvalue analysis, the first non-zero eigenvalue **1 of **A**(diff) was computed and equated to –(*k* + *k*-1). The large spectral gap between **1 and higher eigenvalues (Fig. 7a) indicated good temporal separation of inter- and intra-state events, consistent with two-state kinetics. The values of **2/**1 under reference conditions was 1150 for the 2D model and 1260 for the 1D model.

The mfpt method involved creating an absorbing boundary at *q*2 by truncating **A**(diff) to the invertible matrix **R**′ composed of block matrices **R**, **C**, **D**, and **B**. The mfpt ** is obtained28 by solving

**R**′**** = **u**, (17)

where **** is the column vector containing **(*q*→ *q*2) and **u** is the unit vector. As with the previous dwell time analysis, we equated *k* with**(*q*1→ *q*2). The insensitivity of *k* to the precise value of *q*within the reactant basin (Fig. 7b) is also a reliable indicator of two-state kinetics. A close relative to Eq. 17 is the splitting equation28:

**B** = **Eu**, (18)

where **** is the column vector of splitting or “committor” probabilities from the open region in **B** to the right-hand absorbing boundary determined by *q*2. To compute *k* from **** we employed the flux-over-probability expression *k* = *J*/*pR*, where *pR* is the equilibrium probability in the reactant basin (e.g. *pR* = 0.5 when  = 0) and the flux *J* across any dividing border  spanning the bottleneck region is given by32,33:

. (19)

The mfpt and committor methods computed numerically identical values of *k*, underscoring the fact that Eq. 17 and Eq. 18 are both connected to the Kolmogorov backward equation.

The *k* estimates from solving the master equation in 1D and 2D are tabulated in Table II. They confirm the earlier results from dwell time analysis, again yielding ** ≈ 3/2.

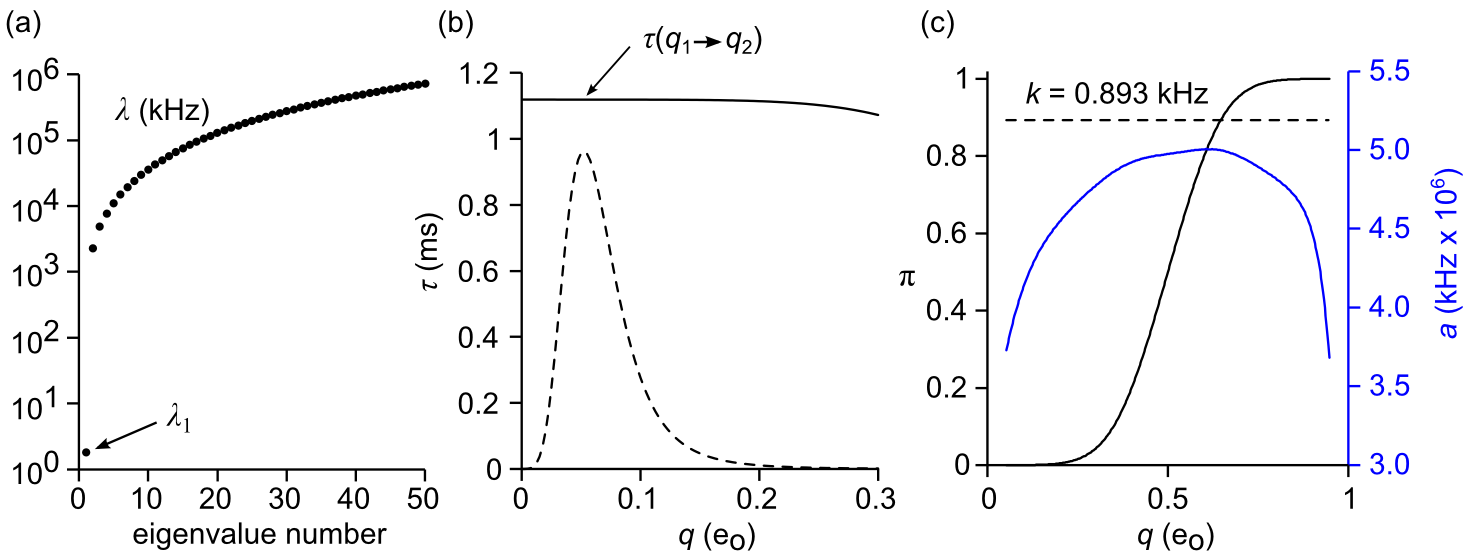


FIG. 7. Methods for calculating *k* from the 1D master equation. Analogous methods were used for 2D. (a) Eigenvalues of **A**, demonstrating the large spectral gap between first and second eigenvalues. (b) Mean first passage time (Eq. 17). The value of **(*q*) is nearly constant where there is probability in the reactant basin (dashed line is the equilibrium probability distribution *peq*). (c) Committor analysis (Eqs. 18 and 19). The computed value of *k* (dashed line) computed from **(*q*) (black line), *a*(*q*) (blue line), and *peq* using Eq. 19 is constant for all *q* in the range [*q*1, *q*2].

1. ***k* as a function of *N*, *T*, and **

We examined whether the ratio ** varied with environmental factors. The relationship between *k* and variables *N*, *T*, and  can be inferred by observing how the environmental alters the 1D diffusion landscape. Reaction rate theory predicts *k* should decrease exponentially with increasing barrier height *W*‡ = *Wb* – *WR* and increase linearly with the diffusion coefficient *D*, particularly in the critical barrier region (*Db*). In Figure 8, we see that *W*‡ increases with increasing *N* and with decreasing *T* and . The value of *Db* increases rapidly with increasing *N*, more slowly with increasing *T*, and is insensitive to . The net effect is that *k* decreases with increasing *N* and increases with increasing *T* and  (Fig. 9). We considered only barrier heights in the range 5-9 *kT*. These are modest energy barriers that are sufficiently small for unbiased simulations to generate a large number of transitions in a reasonable time, but large enough that two-state dynamics are observed (Figs. 1 and 6).

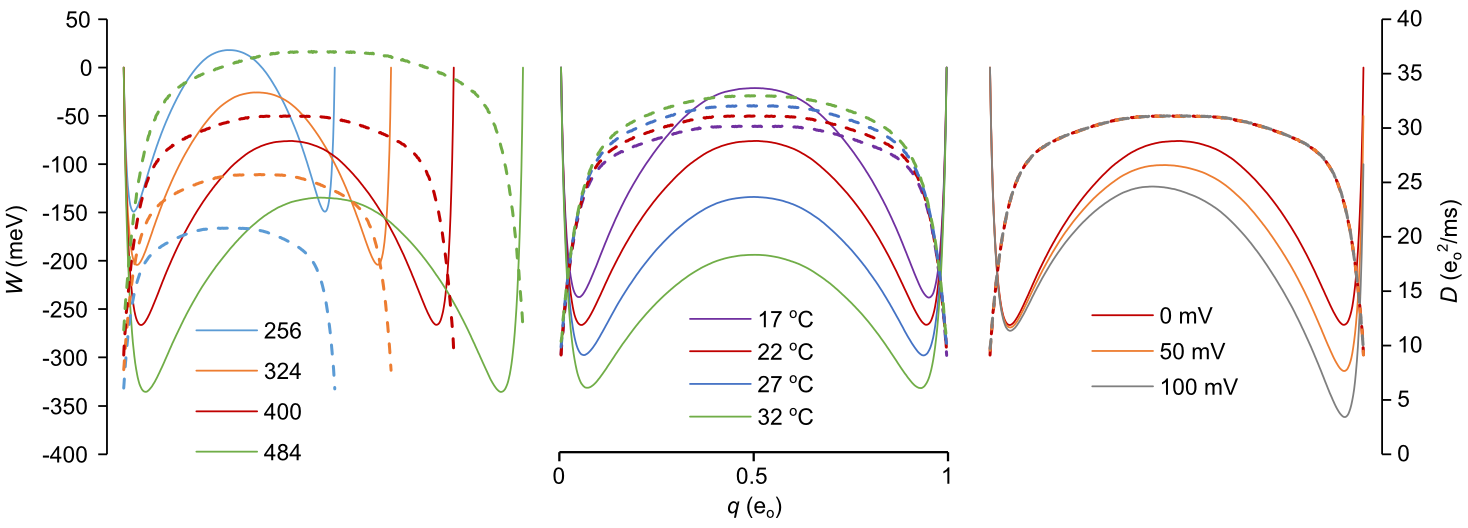


FIG. 8. Free energy (*W*, solid lines) and diffusion (*D*, dashed lines) landscapes as a function of *N* (left), *T* (middle), and right). The red curve in each plot corresponds to the reference model. Landscapes were calculated from the coarse-grained matrices in Fig. 5, though identical results were obtained by time-averaging Monte Carlo trajectories.

The value of ** was fairly constant across environmental conditions (Fig. 9). Consistent with earlier results, 1D rate constant estimates were slightly larger than comparable 2D values, but both landscapes yielded *k* values 50% greater than the Ising model. There were small exceptions under conditions where the free energy barrier exceeded 8 *kT* (*T* = 17 °C and *N* = 484). For these larger barriers, the results from dwell time analysis yielded slightly greater *k* values than the other two methods (mfpt and eigenvalue). However, even these minor outliers did not significantly increase averages taken over all three methods.

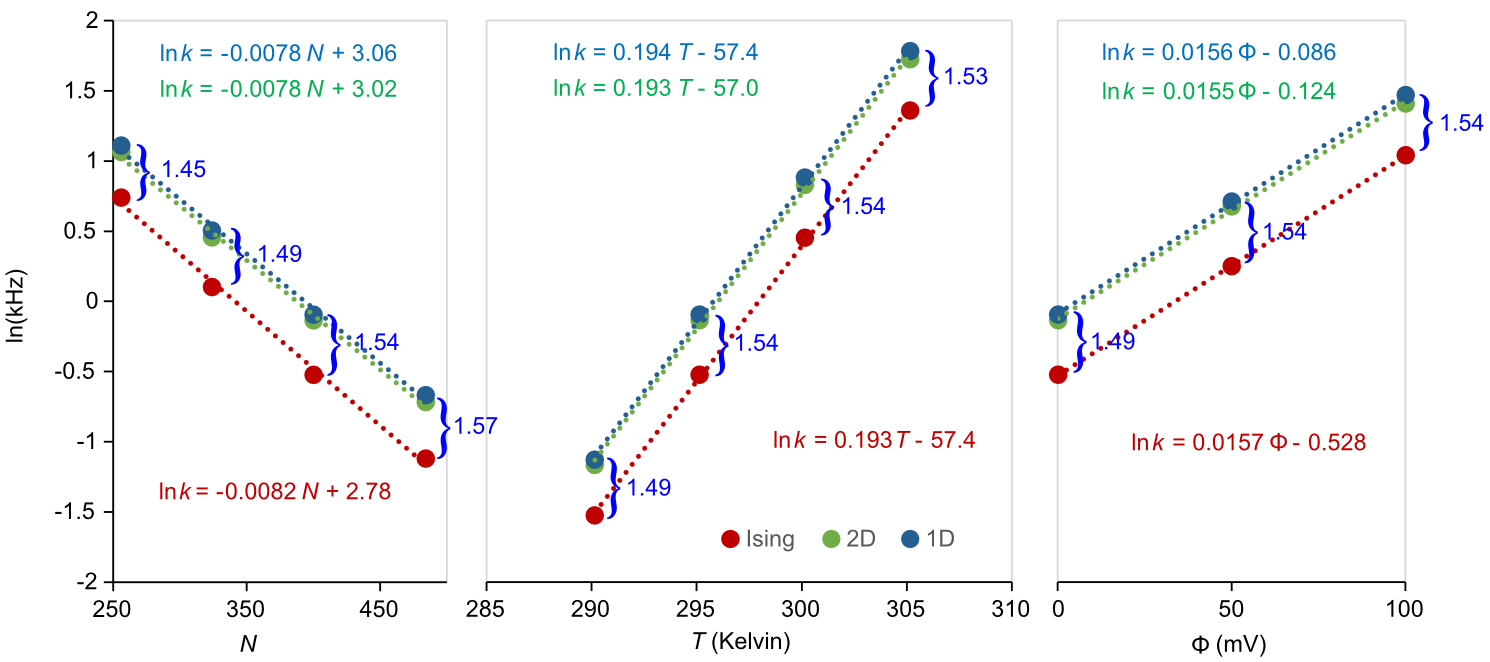


FIG. 9. Linear regression analysis of ln*k* under environmental conditions that are same as in Fig 8. The Ising rate constants were calculated as the inverse means of reactant state dwell times. The 1D and 2D rate constants were averaged from their respective dwell time, mfpt, and eigenvalue calculations. Error bars are buried within the markers. Values for ** are written in blue.

Linear regression analysis using ln*k* = *cX* + *d*, where *X* represented *N*, *T*, and , resulted in very similar slopes *c* between coarse-grained and Ising data (Fig. 9). The slope of the ln*k* vs. *N* plot is consistent with a chemical potential of activation of 0.2 meV per added particle. The ln*k* vs. *T* plot yielded an activation energy *Ea* of 1450 meV. The value of *Ea* was found to be very close to *E*‡ = 1485 meV determined from Boltzmann averaging of the reactant and barrier regions (*E*‡ = 〈*E*〉*b* – 〈*E*〉*R*, see Fig. 2). This large activation energy (~57 *kT*) is interesting from the standpoint of enzyme reactions, since it illustrates how a large positive *S* of activation allows a large *Ea* to be observed on an Arrhenius plot whereas the actual free energy barrier *W* can be much smaller. Plotting ln*k* vs.  yielded the activation charge *qa* = 0.40 eo, which is comparable to *q*‡ ≡ 〈*q*〉*b* – 〈*q*〉*R* = 0.43 eo. Although *q*‡ and *E*‡ are fairly constant for temperatures consistent with modest free energy barriers, we show in Supporting Information that their values vary significantly outside this range.

1. **Methods with no explicit coarse-graining**

All but one numerical methods considered so far were based on the 1D or 2D coarse-grained master equation, the exception being brute-force Monte Carlo simulation of full Ising dynamics. The good agreement among the other methods suggests that coarse-graining consistently overestimates *k* values. To evaluate this further, we employed methods that do not rely on intermediate coarse-graining apart from determining *W*(*q*), which is an equilibrium quantity and therefore is not subject to coarse-graining error. The new quantities of interest are the reactive flux *f* and the diffusion coefficient *D*, both of which were obtained by observing trajectories near the saddle point of the transition barrier. Because barrier relaxation is fast compared to unconstrained brute force simulation of crossing events, these are efficient calculations.

1. **Reactive flux**

The reactive flux method originated with molecular dynamics simulations34 but was expanded to included discrete-state models35. The method makes a distinction between forward- and backward-directed trajectory ensembles initiated at the barrier separatrix, and then keeps a continuous tally of the number of trajectories located on the product side of the barrier. At time zero, only the positive flux ensemble is reactive, so *f*(0) equals the transition state theory (TST) rate constant. At time *T*, the time it takes the system to relax to a metastable state, usually after multiple crossings of the separatrix, *f*(*t*) decays to a plateau value equal to *k*. The ratio *f*(*T*)/*f*(0) is the transmission coefficient **. Thus *k* = *kTST*. We expect ** to be about 1% based on the gold standard value *k* = 0.592 kHz and the 1D model prediction of *kTST*, which is *p*\**a*\* = 55.2 kHz, where *p*\* is the equilibrium probability at the separatrix and *a*\* is the corresponding forward rate constant. The predicted ** value is small, requiring a rather large number of reactive trajectories to obtain good statistics.

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

, (20)

where *\**(0) and *\**(0) are summed microscopic forward and backward rates evaluated at time zero (operationally defined as the exact time a barrier-confined, equilibrated trajectory lands on the separatrix) and *h*\*(*t*) is the Helmholtz function (*h*[*q*(*t*) – *q*\*] = 1 if *q*(*t*) ≥ *q*\*, else 0). The subscripts () and () refer to the forward and backward ensembles of starting trajectories. The () ensemble is the sole determiner of the TST rate at *t* = 0+. Upon relaxing barrier confinement and applying absorbing boundaries at *q*1 and *q*2, the longtime expression of Eq. 26 approaches:

, (21)

which has similarities to Eq. 19.

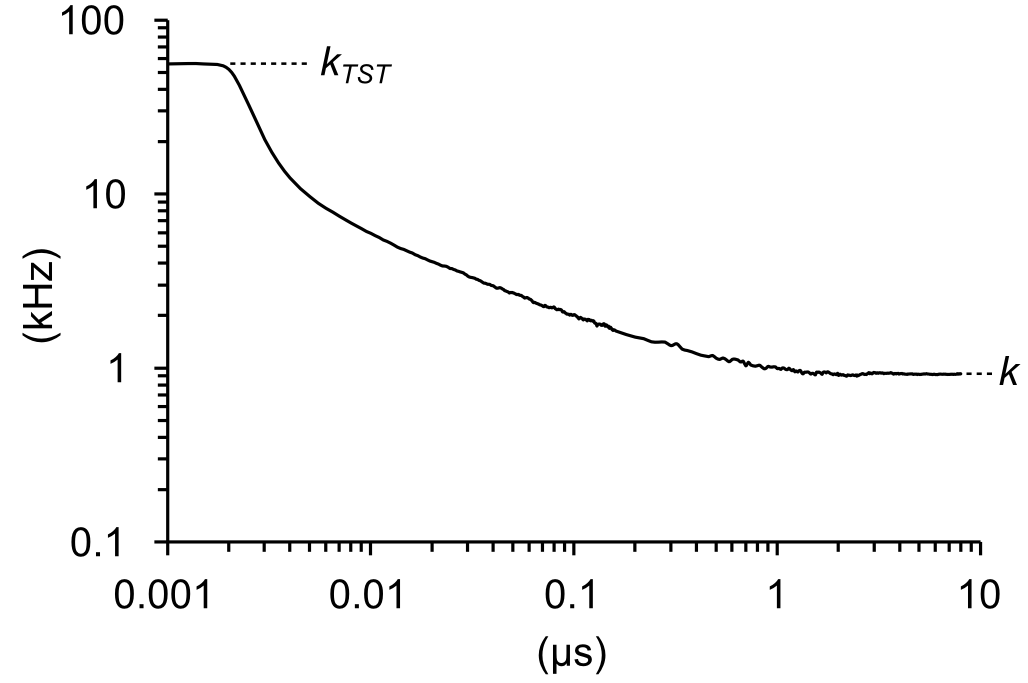


FIG. 10. Reactive flux *f*(*t*) plotted on a log-log scale. Trajectories were run for 8 s after an 800 ns equilibration phase. The value of *k* was determined by averaging the last 3 s.

The outcomes from six simulations, each containing 2 x 105 runs, was *kTST* = 55.8 ± 0.05 kHz and *k* = 0.92 ± 0.07 kHz, yielding ** = 0.016 ± 0.001. Like earlier coarse-grained estimates, *k* overestimated the true transition rate by ** ≈ 3/2.

1. **Nyquist Analysis**

The second of the two nominally non-coarse-graining methods uses the regional barrier diffusion coefficient *Db* obtained from measuring large bandwidth barrier fluctuations. The relationship between the coarse-grained *D*(*q*) from Eq. 13 and the rate constant *k* derives from the Smoluchowski equation (Eq. 14), giving rise to the well-known double integral computing the mfpt from *q*1 to *q*2 36:

. (22)

The first equality in Eq. 22 requires a large free energy barrier, which causes separation of reactant and barrier probability distributions (Fig. 2) and allows the double integral to be factored and Eq. 22 to be reduced to the product of two regional partition functions37. Since *D*(*q*) in the Ising model has nearly constant value within the critical barrier region, the mfpt is insensitive to the spatial variation of *D*(*q*) and depends largely on the barrier value *Db*. The simplified expression is:

, (23)

where *Z*R = ∫Rexp(*W*(*q*)/*kT*)*dq* is the reactant state partition function evaluated from *q* = 0 to *q*\* and *Zb* = ∫bexp(*W*(*q*)/*kT*)*dq* is the partition function of the inverted barrier, evaluated from *q*1 to *q*2. Deriving *Db* from the 1D model using Eq. 13 simply returns the earlier mfpt result. Instead, we independently determined *Db* using Nyquist’s formula for current noise 4*DbB* = 〈*ig*2〉*b*, which we applied to short, barrier-confined trajectories of the “gating” current *ig* = *dq*/*dt*. Since transitions in the Ising model are discontinuous, careful attention was paid to filtering and sampling of *ig* trajectories, as previously described5. The recording bandwidth used to filter the gating current impulses is *B* = 1.064*fc* for a digital Gaussian filter, where *fc* is the cutoff frequency38. With a sufficiently large bandwidth, *Db* approaches the correct limiting value (Fig 11a). We expanded Nyquist’s theorem to two dimensions in order to simultaneously evaluate charge and energy diffusion:

, (24)

where primes indicate differentiation. *Dmn* is the discrete-variable diffusion matrix evaluated at the saddle point. We confined trajectories to a hard window between *q* = 0.35 eo and *q* = 0.65 eo. The charge diffusion coefficient is *Dq* = *Dmmq*2. Averaging over 6 x 500 Ising trajectories lasting 1.6 s with *fc* = 107 kHz, we obtained for the reference model: *Dq* = 30.1 eo2/ms. Applying Eq. 23, we calculated *k* = 0.893 ± 0.004 kHz, which, like the reactive flux method, is consistent with the outcomes from coarse-grained models. The complete diffusion matrix was:

. (25)

One mechanism by which *k* can be overestimated explanation is diffusion anisotropy39, where diffusion across the reactive *q*-coordinate (*Dmm*) is substantially faster than for the stable *E*-coordinate (*Dnn*). Eq. 25 excludes diffusion anisotropy in the 2D model, as the energy diffusion coefficient *Dnn* is only 78% larger than charge diffusion *Dmm*. The actual relaxation time depends on the curvature of the free energy profile. The curvature matrix **C** (Hessian) at the saddle point, obtained by fitting *Wmn* to a second order polynomial in *m* and *n*, is:

, (26)

where *Cmm* = 9.72 x 10-3 meV. Using only the diagonal terms in **C** and **D**, we expect 〈*E*〉 to decay about 54 (1.775 x 30.19) times more rapidly than 〈*q*〉. The observed ratio of 39 (Fig. 11b) is somewhat smaller, which we attribute to fitting error and small deviations from parabolic potentials. In any case, *q* is clearly the slower variable.

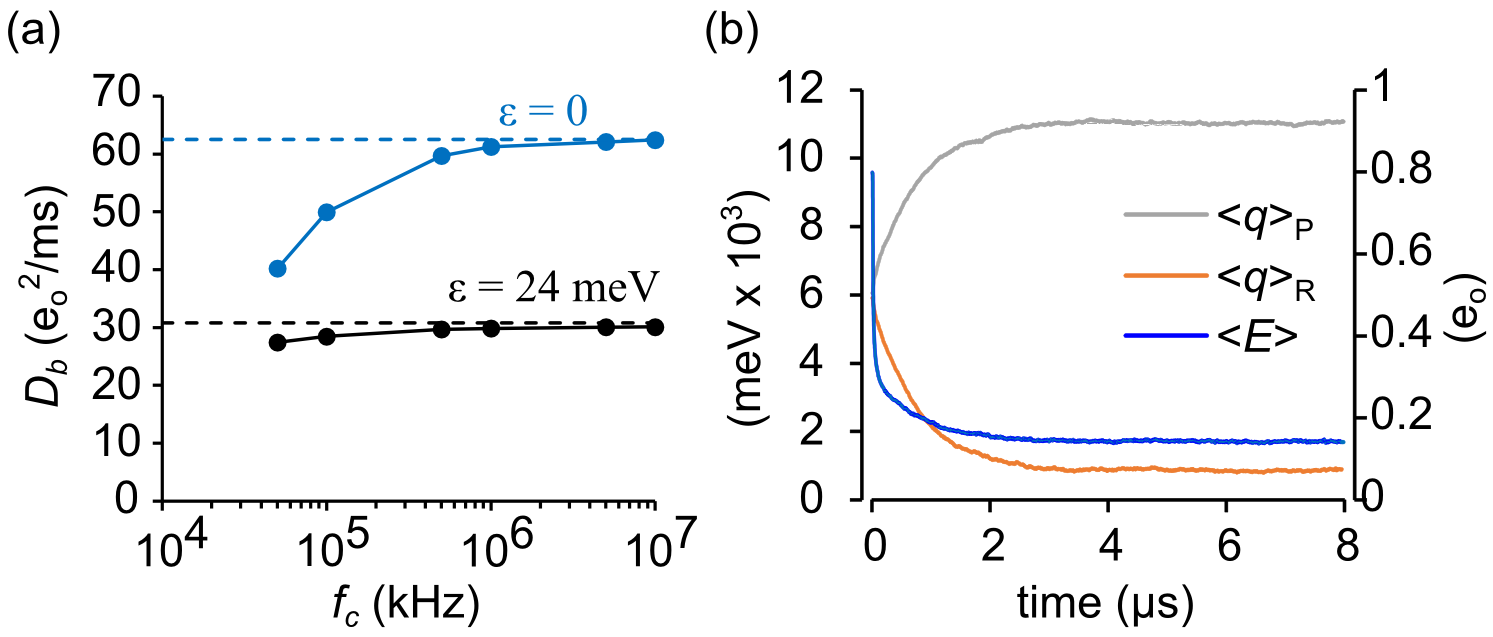


FIG. 11. Diffusion coefficients. (a) *Dq* as a function of filter cut-off frequency for the reference model (black circles) and a system of non-interacting particles (blue circles) serving as a control. Dashed lines indicate limiting values, which in the case of the control system can be solved analytically: *Dq* = 0.5*Nq*2. (b) 〈*E*〉 and 〈*q*〉 decay averaged from 500 simulations starting from a random distribution. 〈*E*〉 decay was bi-exponential, with a fast component (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 〈*q*〉 relaxation (761 ns).

We tested whether *q* can be used as a reaction coordinate in (*E*, *q*) space by applying results from variational transition state theory40. The theory determines the direction **e** of the unstable diffusive mode by solving the eigenvalue equation:

, (27)

where **+ is the only positive eigenvalue of **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 when applied to the 2D coarse-grained model. However, in full configuration space there could be additional collective variables that shift **e** away from the *q*-axis.

1. **Committor testing**

We attempted to detect the existence of other reactive collective variables using committor testing41,42. It is known that the optimal reaction coordinate in configuration space is the committor, or splitting probability **(**x**)43–46. A diffusion landscape that uses the committor as reaction coordinate {*W*(**(**x**)), *D*(**(**x**))} yields the correct value of *k*, but constructing the committor landscape from the full configuration space can be onerous (reviewed by 47) and does not fulfill our stated aim of a comprehensive solution for the (, *T*) ensemble. Using the coarse-grained 1D or 2D model as a starting point, extracting the committor function **(*q*) or**(*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 ** as the mfpt method. This suggests that coarse-graining distorts the committor function. We investigated this by invoking the committor test:

 (28)

Eq. 28 expresses the objective that the committor of the full configuration space **x** should equal the committor of the projected **x** onto the coarse-grained (*q*, *E*) space. We approached the test in two ways. First, we mined configurations for which **(**x**) = 0.5 and plotted them on the (*E*, *q*) landscape, where the symmetry of the Ising model with zero applied field also requires that **(*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 from *q* = 0.3 to 0.7 eo. Trajectory lengths ranged from 3,029 to 21,595 Monte Carlo trials. Using the method described in Pan and Chandler13, we tested each configuration along the trajectory for the condition **(**x**) = 0.5 by launching *n* test runs, terminating each after one or the other stable basins (*q*1 or *q*2) was reached. We then determined whether the cumulative value of **(**x**) = *n*2/*n* was within an acceptable range. Conditions were met if, after 103 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 in a tight locus (Fig. 12a). The presence of “hidden” collective variables should have yielded significant populations of points on either side of the separatrix. The absence of a significant multimodal pattern argues against the need to search for hidden variables in order to construct a reaction coordinate.

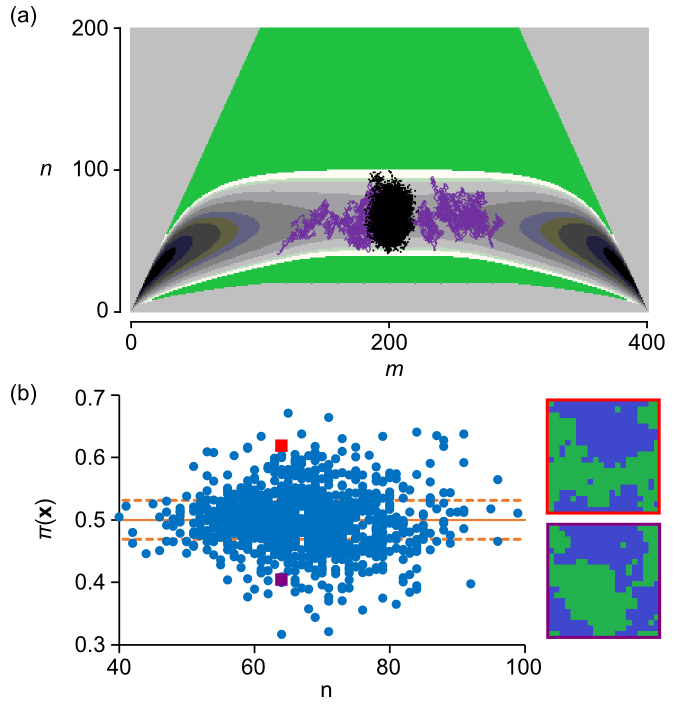


FIG. 12. Committor analysis. (a) 2D distribution of ** = 0.5 configurations (black dots) as a function of energy integer *n* with dashed red lines indicating 95% confidence intervals for 103 random coin flips. A typical crossing trajectory is shown in purple. (b) Committor probabilities of configurations harvested from the separatrix (*m* = 200), evaluated from 103 test runs. The dashed lines demonstrate the 95% confidence interval for 103 random coin flips. The right side of the panel displays two micro-configurations with significantly different ** values (red square, 0.407 ± 0.003; violet square, 0.623 ± 0.002) corresponding to the same coarse-grained state (*m* = 200, *n* = 65).

We implemented the second test by constraining the Ising dynamics to the barrier region between 0.4 eo and 0.6 eo, and launching 103 unconstrained test runs every time *q*\* was crossed The resulting distribution of **(**x**\*) values (Fig 12b) exceeded the expected 95% confidence interval of 103 coin flips. The best explanation for the failure of the coarse-grained (*q\**, *E*) state to predict the (future) committor probability—given the demonstrated lack of evidence for hidden variables—is that memory of prior visitations to neighboring diffusion states is lost when coarse-graining to a memoryless (Markovian) master equation.

1. **Markovianity Testing**

To confirm that coarse-graining causes loss of dynamical memory, we took advantage of a simple test for Markovianity that was recently published48. The test evaluates probabilities of the type *p*(*a* → *b* | *q*) for ensembles of trajectories starting from *q*a and absorbed at *q*b, conditioned on the trajectory crossing *q* located between *q*a and *q*b. From a single unbiased trajectory containing 5,000 forward and backward transitions, we accumulated statistics for the four possible trajectories with free energy minima *q*1 and *q*2 as endpoints, namely *p*(1→ 1 | *q*), *p*(1→ 2 | *q*), *p*(2→ 1 | *q*), and *p*(2→ 2 | *q*). Counts were tallied concurrently for all *q*. This led to correlations between neighboring *q* values, though running the simulation multiple times (*n* = 14) yielded excellent statistics for the entire curve. The Markov assumption is consistent with *p*(1→ 2 | *q*\*) attaining a peak value of 0.25 48. A smaller probability implies a loss of system memory with coarse-graining. We simulated Ising, 1D, and 2D trajectories. The 1D trajectory served as a 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, *n* = 14), but the Ising trajectory peaked at a statistically significant lower value of 0.2381 ± 0.0004 (Fig. 13). This 4.8% reduction from the Markov prediction is small. Nevertheless, we propose that it is the primary determinant of coarse-graining error. A bonus feature of the Markovianity test is that, since *p*(1→ 2 | *q*) + *p*(2→ 2 | *q*) = **(*q*), event tallies can be used to generate committor probabilities. Measuring **(*q*) in this fashion, we calculated *k* by using Eq. 19. Unsurprisingly, the results were consistent with earlier findings of ** ≈ 3/2 (see Table II).

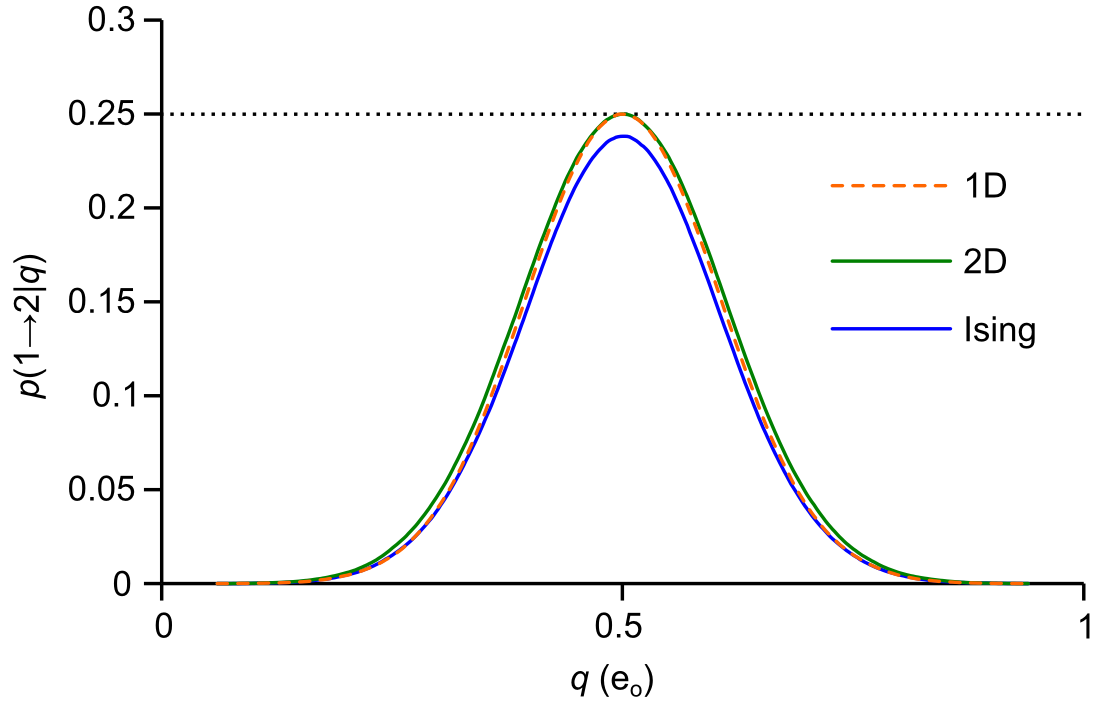


FIG. 13. Markovianity analysis demonstrating memory in Ising kinetics projected onto 1D, but not projections from 1D (control) or 2D kinetics.

1. **Summary of *k* calculations**

In Table II, *k* estimates are grouped with respect to the starting level of coarse-graining. Only one method, brute force determination of dwell times, employed the full configuration space (**x**) without intermediate projection onto *q* or *E*; therefore we considered the resulting value *k* = 0.592 ± 0.007 kHz to be the numerical gold standard. The other eleven methods operated from coarse-grained (1D or 2D) models or were nominally “model-free”. The latter (reactive flux and Nyquist analysis) consisted of rapid diffusion methods projected onto *q*-space confined to the barrier region. The pooled *k* value from the eleven “projection” methods was 0.896 ± 0.007 kHz, yielding ** = 1.51 ± 0.02. Earlier we demonstrated that ** was nearly constant for modest energy barriers (Fig. 9), with 1D and 2D methods producing similar outcomes.

TABLE II. Summary of *k* averaged from 6 experiments (reference model).

|  |  |  |  |
| --- | --- | --- | --- |
| Method | Process | *k* (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. PREDICTING COARSE-GRAINING ERROR**

1. **Which Memory Model?**

The results from the Markovianity test raises the following question: which coarse-grained memory model is appropriate for a non-inertial jump system such as the Ising model? Ising dynamics at every level of coarse-graining can be described by a step-in, step-out scheme without possibility of vertical 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 thermalization49. To our knowledge there is no analytical memory model describing Ising kinetics.

The two diffusion methods, reactive flux and Nyquist analysis, are “model-free” in the sense that no assumptions are made on the shape of the free energy landscape, and they do not rely on intermediate coarse-grained master equations. We considered modifying these methods to incorporate memory effects, starting with the reactive flux. The dynamics of an inertial system, such as a protein molecule in a solvent bath, can be modeled by a generalized Langevin equation with time-dependent friction kernel50,51. Provided the inertial system has a single reaction coordinate and a harmonic barrier, applying reactive flux methods to the general Langevin equation52 leads to the Grote-Hynes transmission coefficient53–55. Since we are not aware of an analogous method for obtaining a memory kernel from a non-inertial system, we did not pursue this line of reasoning further.

1. **Spectral Nyquist Analysis**

Nyquist analysis, like the reactive flux method, overestimates the true rate constant *k* by the ratio ** ≈ 3/2. The source of error was shown to be the value of barrier diffusion coefficient *Db*, which matches the coarse-grained diffusion coefficient evaluated at the separatrix: *D*\* = (1/2)(*a*\* + *b*\*)*q*2. We hypothesized that spectral decomposition of *Db* could 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*) = 〈*ig*(0)*ig*(*t*)〉*eq*. In a non-inertial system governed by a master equation, *R*(*t*) can be computed by eigenvalue decomposition of the rate matrix **A** 5:

, (29)

where *r* = *r*2〈**q***eq***v***r*〉〈**u***r***q**〉 are spectral amplitudes constructed from eigenvalues (*r*) and corresponding left and right eigenvectors (**u***r*, **v***r*). The angle brackets represent inner products between eigenvectors and the gating charge state vector **q** = {*q*} or charge equilibrium distribution **q***eq* = {*peqq*}. The spectral amplitudes of *R*(*t*) are real and negative, but the delta term gives rise to a positive variance. This is in contrast to inertial systems whose velocity autocorrelation function decays with positive real amplitudes and is subject to oscillations. The delta term is the primary contributor to the one-sided power spectrum, which after filtering is expressed as5:

, (30)

where *H*(*f*) is the Fourier-transformed filter response function. For cutoff frequencies *fc* much larger than the largest contributing eigenvalue **max, Eq. 30 integrated across positive frequencies *f* yields a gating current variance that is proportional to the filter bandwidth:

, (31)

Comparing Eq. 31 to the Nyquist current formula 〈*ig*2〉 = 4*DB*, we can write the diffusion coefficient as the sum of spectral components:

. (32)

We performed spectral analysis of the gating current variance by adding to the system energy a harmonic umbrella potential with curvature *c* centered at the transition state *q*\*, which confines **A** to the barrier region. After generating 105 runs of 410 ns trajectories with *fc* = 107 kHz bandwidth, the function *R*(*t*) was calculated by averaging *ig*(*t*1)*ig*(*t*1 + *t*) for each run as a function of *t*1, then averaging over all runs. The number of paired data time points for a correlation time *t* was proportional to (410 – *t*). This number was used as a weighting factor to generate uniform residuals for exponential fits of the autocorrelation function (Fig. 14a, b).

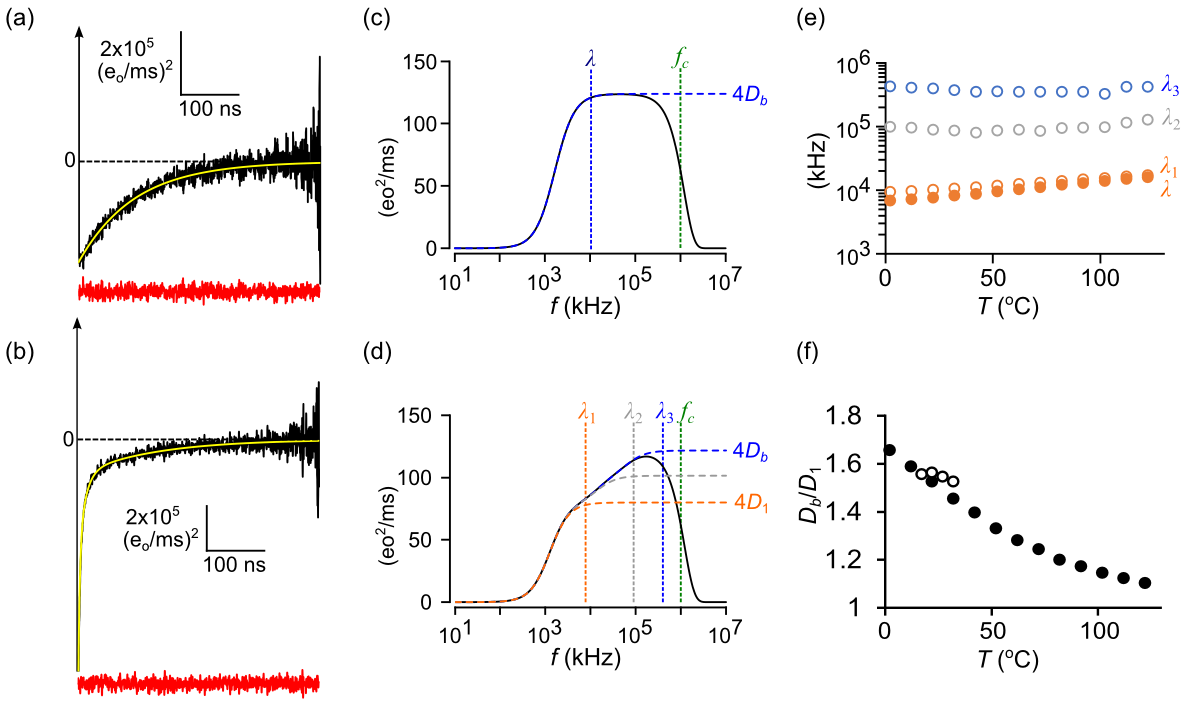


FIG. 14. Spectral decomposition of the barrier diffusion coefficient. (a) Gating current autocorrelation function for the 1D model (black line) fit to a single exponential (yellow line). The red trace is residuals after normalizing for pair numbers. Sampling and filtering cut-off frequencies were: *fs* = 107 kHz; *fc* = 106 kHz. (b) Same as (a) except with Ising kinetics, requiring a three-exponential fit. (c) Power spectrum of 1D dynamics calculated from the fitted data. Dashed vertical lines indicate values for the decay rate (**) and the filter cutoff frequency (*fc*). The decomposition of the Ising power spectrum into three components is shown in (d). (e) Ising (solid circles) and 1D (open circles) decay rates as a function of temperature. (f) *Db*/*D*1 (filled circles) as it approaches unity for increasing temperatures compared to simulation values for ** taken from dwell time analysis (open circles). Each data point in (e) and (f) is averaged over 6 sets of simulations.

The 1D and 2D autocorrelations were fit to a single exponential (*rmax* = 1) as expected from a process that is stationary, quasi-harmonic, and Markovian—and therefore closely related to the Ornstein-Uhlenbeck process56. The area beneath the 1D curve equaled 4*Db* = 124.0 eo2/ms (*n* = 6), consistent with the prediction after coarse-graining: *D\**1D = (*a*\* + *b*\*)*q*2/2 = 31.1 eo2/ms. The decay rate ** of the 1D autocorrelation function depended linearly on *c*, and back-extrapolation to *c* = 0 yielded the negative curvature of the original barrier potential, as expected. We chose *c* = 104 mV/eo since this value strikes the right balance between constraining the system to the barrier region and adequately separating the eigenvalues while also satisfying the integration requirement *max* << *fc*. The outcome (*Db*) was insensitive to the value of *c* as long as these conditions were satisfied.

The full configuration Ising process yielded three exponential components (*rmax* = 3), with a total area 4*Db* = 121.8 eo2/ms (*n* = 6) that matched the 1D value. The slowest Ising component decayed with nearly the same rate **1 as the 1D model (Fig.14c, d), and in temperature-independent fashion (Fig. 14e). However, the corresponding area was 4*D*1 = 20.0 eo2/ms (*n* = 6), about 2/3 the value of *Db*. Since *Db*/*D*1 ≈ **, it suggests that *D*1 is the reactive diffusion coefficient. We are empirically led to the following equation that corrects the error in Eq. 23:

. (33)

However, *D*1/*Db* cannot be constant at all temperatures. Like the 1D barrier fluctuations encountered earlier, the high-temperature Ising model resembles a single-component Ornstein-Uhlenbeck process, though for a different reason: the high-temperature model reduces to a system of independent particles (Fig. 11a). In the high-temperature limit *Db* and *D*1 are identical, whereas we showed that at subcritical temperatures—for modest-sized transition barriers—the spectral ratio (*Db*/*D*1) coincides with ** values (Figs. 9 and 14f). The fact that *Db*/*D*1 ≈ ** for only a narrow range of temperatures raises the possibility that this equality occurred by chance. To investigate this possibility we needed a system with a different ** to see if the equality holds true.

1. **Extending the Ising model to three dimensions**

In order to test the validity of Eq. 33 in a different system, we chose the three-dimensional Ising model. The 3-D Ising model is analogous to the two-dimensional system except that every cell has 6 rather than 4 neighboring interactions. As a result, the number of cardinal transition matrices increases from 10 to 14 and the number of discrete energies increases from *N* to 3*N*/2 (Fig. 15a). Comparing dwell times from Ising and coarse-grained 1D dynamics in the 3-D model using brute force simulations in an 8 x 8 x 8 (*N* = 512) lattice with ** = 12.2 meV, *T* = 22 deg C, *V* = 0 mV, and ** = 104 kHz, the coarse-graining error was found to be small (** = 1.055 ± 0.009). This is in keeping with general observation that signal averaging is often more reliable in systems with a greater number of interaction. A statistically indistinguishable number (*Db*/*D*1 = 1.063 ± 0.003) was obtained through spectral Nyquist analysis (Fig. 15c,d), demonstrating that at least in Ising lattices *Db*/*D*1 ≈ ** for conditions leading to a modest-sized energy barrier. Furthermore, Markovianity testing revealed that the 3-D Ising model is only marginally non-Markovian (p(1 → 2 | q\*) = 0.2469 ± 0.0015, and noticeably less so than the 2-D model (Fig. 15e), providing further evidence, albeit still circumstantial, that coarse-graining error in Ising lattices is attributable to memory effects.

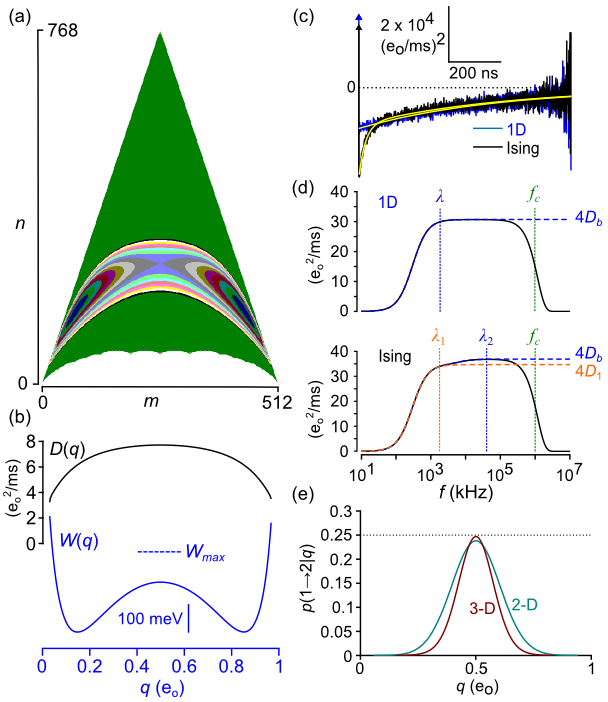


FIG. 15. Three-dimensional Ising model (*N* = 83). (a) Populated states in (*m*, *n*) space. The transition pathway shown as a free energy contour truncated 200 meV above the saddle point energy. (b) Potential of mean force *W*(*q*) and coarse-grained diffusion coefficient *D*(*q*) projected from the truncated energy landscape in (a) and from 14 cardinal transition matrices. The dashed line is *W* plus the umbrella potential *U* used for Nyquist analysis. (c) Umbrella-sampled gating current autocorrelation function (average of 6 simulations) obtained under same conditions as Fig. 14a,b. Yellow lines represent one- and two-exponential fits. (d) Power spectra. The slower of the two resolvable Ising components matched the single 1D component in frequency but was 1/** as large in amplitude. (e) Markovianity testing of the two-dimensional (same as Fig. 13) and three-dimensional Ising models.

**V. DISCUSSION**

In this study, using numerical methods, we computed the forward rate constant *k* of the electric analog of the kinetic 2-D Ising model for a range of temperatures andfield strengths consistent with a modest (5-9 *kT*) free energy barrier. The procedure is “comprehensive” in the sense that projection of Ising dynamics onto the microcanonical ensemble provided, in addition to the density of states, a matrix set of “branching coefficients” that served as the kernel for computing *k* in the (, *T*) ensemble. The basis for the calculation is Eq. 33, and specifics of the numerical strategy for computing *k* are summarized in the Supporting Information. The gating charge *q* proved to be a good reaction coordinate in (*q*, *E*) space, and committor analysis of configuration-space Ising dynamics failed to uncover additional reactive variables. However, coarse-grained kinetics consistently overestimated the gold standard value by about 50% (** ≈ 1.5). The error ** did not vary substantially for free energy barriers of modest height (5-9 *kT*). Markovianity testing of Ising dynamics projected was consistent with a loss of dynamical memory—presumably this is the major source of coarse-graining error, though the origin is unknown. Decomposition of the stationary gating current autocorrelation function *R*(*t*) in the barrier vicinity successfully predicted ** by relating the slowest integrated component of *R*(*t*) to the reactive diffusion coefficient *D*1.

Outcomes for the 3-D Ising model, where coarse-graining error was shown to be smaller (** ≈ 1.06), followed the patterns observed in the 2-D system. Notably, the reactive component frequencies of the Ising and 1D coarse-grained models were closely aligned and the 3-D implementation of “spectral Nyquist analysis” successfully predicted the smaller ** value. The expectation was that the 3-D model should exhibit less memory loss with coarse-graining, and Markovianity testing proved this to be true, though more work is needed to better characterize the link between Markovianity and **.

The success of spectral Nyquist analysis in correctly coarse-graining error in 2-D and 3-D Ising systems leads to the following conjecture, applicable only to barrier crossing events. The reactive diffusion coefficient defined in Eq. 33 is the slowest spectral component of the integrated autocorrelation function of the barrier-confined gating current (or more generally the reaction coordinate velocity). The conjecture simplifies to the following relation between the proposed reactive diffusion coefficient and the slowest eigencomponent from umbrella-sampled barrier dynamics:

, (34)

where **1 is the slow eigenvalue and **u***r*, **v***r* are the corresponding left and right eigenvectors. If the conjecture can be proved and generalized to systems other than the Ising model, Eq. 33 may prove useful in computing barrier crossing rate constants in systems where the best available reaction coordinate is suspected to be non-Markovian. In cases where it is also possible to project the system dynamics onto the microcanonical ensemble, a comprehensive solution can be generated that is valid for a range of environmental conditions.

**ACKNOWLEDGEMENTS**

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (R01GM093290, R01GM131048 and S10OD020095), and by National Science Foundation (CNS-1625061). We thank the anonymous reviewers for their helpful comments, in particular their suggestions for committor testing and including three-dimensional Ising models.

**REFERENCES**

1 K. Kandiah, M.O. Deighton, and F.B. Whiting, J. Appl. Phys. **66**, 937 (1989).

2 F.J. Sigworth, Nature **423**, 21 (2003).

3 T.L. Hill, J. Chem. Phys. **36**, 3182 (1962).

4 T.J. Harpole and L. Delemotte, Biochim. Biophys. Acta - Biomembr. 0 (2017).

5 D. Sigg, H. Qian, and F. Bezanilla, Biophys. J. **76**, 782 (1999).

6 L. Delemotte, M.A. Kasimova, D. Sigg, M.L. Klein, V. Carnevale, and M. Tarek, BioRxiv 108217 (2017).

7 B. Peters, P.G. Bolhuis, R.G. Mullen, and J.-E. Shea, J. Chem. Phys. **138**, 054106 (2013).

8 A.M. Ferrenberg and R.H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).

9 M. Pleimling and H. Behringer, Phase Transitions **78**, 787 (2005).

10 A. Hűller and M. Pleimling, Int. J. Mod. Phys. C **13**, 947 (2002).

11 J.R. Ray and C. Freléchoz, Phys. Rev. E **53**, 3402 (1996).

12 K. Brendel, G.T. Barkema, and H. van Beijeren, Phys. Rev. E **71**, 031601 (2005).

13 A.C. Pan and D. Chandler, J. Phys. Chem. B **108**, 19681 (2004).

14 B. Peters and B.L. Trout, J. Chem. Phys. **125**, 054108 (2006).

15 C. Moritz, A. Tröster, and C. Dellago, J. Chem. Phys. **147**, 152714 (2017).

16 J.-M. Leyssale, J. Delhommelle, and C. Millot, J. Chem. Phys. **127**, 044504 (2007).

17 J.F. Cordero-Morales, L.G. Cuello, and E. Perozo, Nat. Struct. Mol. Biol. **13**, 319 (2006).

18 H.T. Kurata, M. Rapedius, M.J. Kleinman, T. Baukrowitz, and C.G. Nichols, PLoS Biol **8**, 1000315 (2010).

19 D. Sigg, J. Gen. Physiol. **141**, 29 (2013).

20 R.J. Glauber, J. Math. Phys. **4**, 294 (1963).

21 D.T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

22 H. Qian, Phys. Rev. E **65**, 016102 (2001).

23 S. Kjelstrup, G.P. Beretta, A. Ghoniem, and G. Hatsopoulos, in *AIP Conf. Proc.* (AIP, 2008), pp. 223–228.

24 C.N. Yang, Phys. Rev. **85**, 808 (1952).

25 L. Onsager, Phys. Rev. **65**, 117 (1944).

26 D. Sigg and F. Bezanilla, J. Gen. Physiol. **109**, 27 (1997).

27 R.E. Belardinelli and V.D. Pereyra, Phys. Rev. E **75**, 046701 (2007).

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

29 P. Reimann, G.J. Schmid, and P. Hänggi, Phys. Rev. E **60**, R1 (1999).

30 B. Roux and R. Sauvé, Biophys. J. **48**, 149 (1985).

31 F.J. Sigworth and S.M. Sine, Biophys. J. **52**, 1047 (1987).

32 A. Berezhkovskii, G. Hummer, and A. Szabo, J. Chem. Phys. **130**, 205102 (2009).

33 P. Metzner, C. Schütte, and E. Vanden-Eijnden, Multiscale Model. Simul. **7**, 1192 (2009).

34 D. Chandler, J. Chem. Phys. **68**, 2959 (1978).

35 M. Borkovec and P. Talkner, J. Chem. Phys. **92**, 5307 (1990).

36 A. Ansari, J. Chem. Phys. **112**, 2516 (2000).

37 D. Sigg, J. Gen. Physiol. **144**, 7 (2014).

38 D. Colquhoun and F.J. Sigworth, in *Single-Channel Rec.* (Springer US, Boston, MA, 1983), pp. 191–263.

39 A.M. Berezhkovskii and V.Y. Zitserman, Phys. A Stat. Mech. Its Appl. **166**, 585 (1990).

40 A. Berezhkovskii and A. Szabo, J. Chem. Phys. **122**, 14503 (2005).

41 R. Du, V.S. Pande, A.Y. Grosberg, T. Tanaka, and E.S. Shakhnovich, J. Chem. Phys. **108**, 334 (1998).

42 P.L. Geissler, C. Dellago, and D. Chandler, J. Phys. Chem. B **103**, 3706 (1999).

43 G. Hummer, J. Chem. Phys. **120**, 516 (2004).

44 P. V. Banushkina and S. V. Krivov, Wiley Interdiscip. Rev. Comput. Mol. Sci. **6**, 748 (2016).

45 W. E, W. Ren, and E. Vanden-Eijnden, Chem. Phys. Lett. **413**, 242 (2005).

46 A.M. Berezhkovskii and A. Szabo, J. Phys. Chem. B **117**, 13115 (2013).

47 B. Peters, Annu. Rev. Phys. Chem. **67**, 669 (2016).

48 A.M. Berezhkovskii and D.E. Makarov, J. Phys. Chem. Lett. **9**, 2190 (2018).

49 P. Hänggi and P. Talkner, Zeitschrift Für Phys. B Condens. Matter **45**, 79 (1981).

50 H. Mori, Prog. Theor. Phys. **33**, 423 (1965).

51 R. Zwanzig, Phys. Rev. **124**, 983 (1961).

52 D. Kohen and D.J. Tannor, *Advances in Chemical Physics* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 1999).

53 E. Pollak, J. Chem. Phys. **85**, 865 (1986).

54 R.F. Grote and J.T. Hynes, J. Chem. Phys. **73**, 2715 (1980).

55 P. Hanggi and F. Mojtabai, Phys. Rev. A **26**, 1168 (1982).

56 J.L. Doob, Ann. Math. **43**, 351 (1942).