Stochastic paths controlling speed and dissipation

Rebecca A. Bone, Daniel J. Sharpe, David J. Wales, and Jason R. Green

Stochastic paths controlling speed and dissipation

Rebecca A. Bone,[1] Daniel J. Sharpe,[2] David J. Wales,[2] and Jason R. Green[1, 3, *]

[1]*Department of Chemistry, University of Massachusetts Boston, Boston, MA 02125*
[2]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW*
[3]*Department of Physics, University of Massachusetts Boston, Boston, MA 02125*

Natural processes occur in a finite amount of time and dissipate energy, entropy, and matter. Near equilibrium, thermodynamic intuition suggests that fast, irreversible processes will dissipate more energy and entropy than slow, quasistatic processes connecting the same initial and final states. For small systems, recently discovered thermodynamic speed limits suggest that faster processes will dissipate more than slower processes. Here, we test the hypothesis that this relationship between speed and dissipation holds for stochastic paths far from equilibrium. To analyze stochastic paths on finite timescales, we derive an exact expression for the path probabilities of continuous-time Markov chains from the path summation solution to the master equation. We present a minimal model for a driven system in which relative energies of the initial and target states control the speed and the nonequilibrium currents of a cycle control the dissipation. While the hypothesis holds near equilibrium, we find that faster processes can dissipate less under far-from-equilibrium conditions because of strong currents. This model serves as a minimal prototype for designing kinetics to sculpt the nonequilibrium path space so that faster paths produce less dissipation.

INTRODUCTION

Biological systems balance the timely formation of structure with the thermodynamic costs of dissipated heat, entropy production, and wasted free energy [1, 2]. For example, cells regulate the speed at which microtubules assemble and dynamically reorganize, which necessarily dissipates energy [3, 4]. If the reorganization of these structures is too slow, or driven by chemical reactions that are too dissipative, the fluctuating dynamics of microtubules would inhibit, not facilitate, cellular functions. A quantitative understanding of the relationship between speed and dissipation is important not only for biological functions, but also for implementing these functions in synthetic, dissipative materials. Particular progress has been made controlling the transient formation of structure in active materials with dissipative cycles of chemical reaction networks [5–7] and accessing the richer set of structures that can be formed outside of chemical, thermal, and diffusive equilibria [8–11]. Despite this progress, when open to external sources of energy and matter, kinetic trapping [12, 13] and dissipative cycles [14, 15] are prevalent, making it an open question how these systems use time and energy in the formation of structure.

The relationship between speed and energy efficiency is particularly challenging for nonstationary, nonequilibrium processes [16, 17]. One expectation that might come from equilibrium thermodynamics is that fast, irreversible processes will dissipate more energy and entropy than slow, quasistatic processes connecting the same initial and final states [18]. Included in near equilibrium processes is the constraint that the system has enough time at each step for local inhomogeneities to relax. In this case, the dissipation, often quantified by the entropy production rate, is roughly the inverse timescale of the

dominant nonequilibrium effects. However, intuition established for large systems near equilibrium need not hold for systems that are small, strongly driven, or undergo transient or fluctuating nonequilibrium processes. Nevertheless, accumulating evidence in stochastic thermodynamics [19] seems to support this tradeoff between speed and dissipation and the idea that more time will be required for processes that are less dissipative [20–25]. For example, some thermodynamic speed limits suggest that rates of dissipation have an upper bound set by the intrinsic timescale of the system, regardless of its size or its "distance" from equilibrium. These results might be taken to suggest that the more time required for a process, the smaller the associated dissipation. However, these global bounds on stochastic processes have not yet been fully analyzed at the level of stochastic paths.

Here, we test the hypothesis that there is a positive correlation between speed and dissipation away from equilibrium by controlling the currents exchanged between a system and its surroundings. Fluctuations in energy, entropy, and concentrations are important for testing this hypothesis for finite-size systems, a regime where stochastic paths are known to provide useful quantitative information [26–28] about both the internal timescales set by the kinetics and dissipation [29–32]. These observables can be extracted from Markov models parameterized based on deterministic trajectories [33, 34] or kinetic theory [35]. However, an open technical challenge is the probability of stochastic paths constrained by time in continuous time Markov chains [36–38]. We overcome this challenge by deriving an explicit form for the path probability of any continuous-time Markov process. This closed form expression provides a direct method for computing the average speed and dissipation of nonequilibrium paths over a fixed period of time. Applying this formula, we quantitatively analyze stochastic paths to

determine when faster processes dissipate more entropy.

CONTRACTED PATH PROBABILITY

Consider a small system with a discrete set of N states, one designated as the initial state and one as the final state. Take the transitions between these states to be a Markov jump process [39]. These stochastic dynamics have been used to describe various processes, including self-assembly [40], quantum dots [41] and molecular motors [42]. Mathematically, they are represented by a collection of time-homogeneous transition rates $w(y|x)$ for a jump from state x to state y with total escape rate $w_x = \sum_{y \neq x} w(y|x)$ from x. The dynamics of occupation probabilities for each state are described by the master equation [39], $\dot{p}(x,t) = \sum_y w(x|y)p(y,t)$. Its path summation solution [45] gives the marginal probability,

$$p(x_f, t) = \sum_{n=0}^{\infty} \sum_{\mathcal{C}_n} \mu(\mathcal{C}_n = x_0, x_1, \ldots, x_f, t), \quad (1)$$

in terms of the joint probability μ that the system takes a path $\mathcal{C}_n = x_0, x_1, \ldots, x_n$ of n jumps that ends in x_f after a time t. Each path, $\mathcal{C}_n$, is a time-ordered sequence of states.

Stochastic thermodynamics has measures of dissipation for these stochastic paths [46, 47]. Assuming local detailed balance [48], the entropy change of the equilibrium reservoirs mediating the fluctuating dynamics puts a constraint on the asymmetry of the transition rates: $-s_e[\mathcal{C}_n]/k_B = \sum_{i=0}^{n-1} \ln w(x_{i+1}|x_i)/w(x_i|x_{i+1})$. This entropy flow is interpreted as the amount of entropy dissipated from the system to the surroundings in traversing a path [49–51]. When the transition rates are exponentially related to the energy, $-s_e[\mathcal{C}_n]/k_B$ is the energy exchanged as heat between the system and surroundings, $k_B T \ln w(x,y)/w(y,x) = q(x,y)$ [52], at a temperature T [19, 46, 47]. However, imposing a fixed observation time $t \geq t_n$ for the process constrains the possible paths included in the path summation [36, 37] and, in turn, the thermodynamic costs. That is, fixing the observation time alters nonequilibrium ensemble averages over paths and the associated dissipation.

To account for a time constraint on stochastic processes, we need the probability that a path is traversed by an ensemble of stochastic trajectories; trajectories here are a time-ordered sequence of states and stochastic transition times, $\mathcal{T}_n = x_0, t_0; x_1, t_1; \ldots; x_n, t_n$ [38, 53–55]. This *contracted* path probability, $\mu(\mathcal{C}_n, t)$, depends on the path probability and the probability of a stochastic time sequence through Bayes' theorem:

$$\mu(\mathcal{C}_n, t) = p(t|\mathcal{C}_n)p(\mathcal{C}_n). \quad (2)$$

The probability of a path is relatively straightforward to

compute from the transition rates and the escape rates,

$$p(\mathcal{C}_n) = p(x_0) \prod_{i=1}^{n} \frac{w(x_i|x_{i-1})}{w_{x_{i-1}}}. \quad (3)$$

However, given the path $\mathcal{C}_n$ occurs, the probability it completes in a certain amount of time t,

$$p(t|\mathcal{C}_n) = \rho_0 * \left[\rho_1 * \left[\ldots * \left[\rho_{n-1} * \left[e^{-w_{x_n}(t-t_n)}\right]\right]\right]\right], \quad (4)$$

is more difficult to determine analytically. The nested convolutions, represented by $*$ here, are of the exponential distributions of waiting times along a path. The waiting time, $\Delta t = t_i - t_{i-1}$, in each state, x_i, along a trajectory is exponentially distributed,

$$\rho(\Delta t_i|x_i) = \rho_i = w_{x_i} e^{-\Delta t_i w_{x_i}}, \quad (5)$$

and independent of the other escape rates and states [43, 44]. The last exponential factor in these convolutions is the survival probability of the final state. Because of the stochastic transition times, some trajectories remain in the final state for $t - t_n \geq 0$, but others may not reach x_n or may leave x_n within the chosen time t.

What complicates (or simplifies) the formula for the contracted path probability is the combinatorics of the escape rates along the path: states along the path can have degenerate escape rates that require an accounting for their indistinguishability. In Ref. [38], Sun derived the contracted path probability,

$$\mu(\mathcal{C}_n, t) = p(x_0, t_0) \prod_{i=0}^{n-1} w(x_{i+1}|x_i)$$
$$\times \sum_{j=1}^{n'} \nu_j \frac{\partial^{m_j-1}}{\partial w_{x_j}^{m_j-1}} \left[\frac{e^{-w_{x_j}t}}{\prod_{\substack{k=1, \\ k \neq j}}^{n'} (w_{x_k} - w_{x_j})^{m_k}}\right], \quad (6)$$

where the sum is over the n' unique escape rates. Preceding the sum is the product of the path probability $p(\mathcal{C}_n)$ in Eqn. 3 and the prefactor, $\prod_{i=0}^{n-1} w_{x_i}$, in the convolutions of the waiting time distributions, Eqn. 4. The function ν_j

$$\nu_j = \frac{(-1)^{m_j-1}}{(m_j-1)!}. \quad (7)$$

corrects for the indistinguishability of escape rates using the degeneracy, m_j, of the jth unique escape rate.

Our main analytical result is an exact expression for the contracted path probability, $\mu(\mathcal{C}_n, t)$ (Appendix) [56],

$$\mu(\mathcal{C}_n, t) = p(x_0, t_0) \prod_{i=0}^{n-1} w(x_{i+1}|x_i) \quad (8)$$
$$\times \sum_{j=1}^{n'} \nu_j f_j^{(0)} \sum_{l=1}^{m_j} \binom{m_j-1}{l-1} (-t)^{m_j-l} d_j^{(l-1)},$$

which consists of three functions. The first, ν_j, in Eqn. 7 again accounts for the multiplicity m_j of the jth of the n' unique escape rates along a path. The second is a function proportional to the exponential distribution of waiting times in each state, $f_j^{(0)} \propto e^{-w_{x_j}t}$. The third, $d_j^{(l-1)}$, is a combinatoric function of the escape rate multiplicities that derives from the nested convolutions of the waiting time distributions; these combinatorics were previously pointed out as a particular challenge in deriving the closed-form solution [38, 57, 58]. Equation 8 overcomes this challenge and has advantages for describing the stochastic thermodynamics of processes constrained by time. It allows direct quantification of the relative importance of paths with respect to their stochastic time sequences, making it useful for extracting insights into dynamical mechanisms. For example, the magnitude of the probability can be used to assess the relative importance of competing mechanisms.

To validate this contracted path probability formula, we generated a sample of stochastic trajectories for an arbitrarily chosen path of length $n = 1000$ using kinetic Monte Carlo [70, 71]. From this sample, we found the histogram of occurrence times agreed with the analytical distribution of occurrence times for the trajectories. This theoretical distribution is:

$$p(t|\mathcal{C}_n) = \prod_{i=0}^{n-1} w_{x_i} \sum_{j=1}^{n'} \frac{(-1)^{m_j-1}}{(m_j-1)!} \frac{e^{-w_{x_j}t}}{\prod_{\substack{k=1 \\ k\neq j}}^{n'} (w_{x_k} - w_{x_j})^{m_k}}$$
$$\times \sum_{l=1}^{m_j} \binom{m_j-1}{l-1}(-t)^{m_j-l} d_j^{(l-1)}. \qquad (9)$$

Another point of comparison was the mean occurrence time:

$$\langle\tau\rangle_{\mathcal{C}_n} = \int_{t_0}^{\infty} t\, p(t|\mathcal{C}_n)dt = \sum_{i=0}^{n} \frac{1}{w_{x_i}}, \qquad (10)$$

which can be derived using the statistical independence of the exponentially-distributed waiting times along a path. We found good agreement between the analytical and numerical mean path occurrence time.

To further validate the explicit contracted path probability formula, we compared symbolic and numerical calculations for paths with length $n \le 10$. We evaluated the convolution integrals in Eqn. 4, the path probability $p(\mathcal{C}_n)$ in Eqn. 3, the non-explicit contracted path probability formula in [38], and the explicit contracted path probability in Eqn. 8 for paths of lengths $n \le 10$. We also analyzed paths with $n \le 10$ with all possible combinations of degeneracies; for example, for $n = 3$ we checked this formula for degeneracies of (1,1,2), (1,3), and (2,2) as well as (1,1,1,1), the fully non-degenerate case, and 4 in the fully degenerate case. Analytically, we confirmed

that this formula also simplifies to known expressions [36] when the escape rates are distinct,

$$\mu(\mathcal{C}_n, t) = p(x_0, t_0) \prod_{i=1}^{n} w(x_i|x_{i-1}) \times$$
$$\sum_{j=0}^{n} \frac{e^{-w_{x_j}t}}{\prod_{\substack{k=1 \\ k\neq j}}^{n}(w_{x_k} - w_{x_j})^{m_k}}, \qquad (11)$$

and when the escape rates along a path are identical,

$$\mu(\mathcal{C}_n, t) = p(x_0, t_0)\frac{t^n}{n!}e^{-wt}\prod_{i=1}^{n} w(x_i|x_{i-1}). \qquad (12)$$

In the former case, the sum is over the $f_j^{(0)}$ for each escape rate j along the path, Appendix. This formula can be evaluated for any path with known escape rates and a given observation time, regardless of the path length n or the size of the system.

Using the general expression for the path probability, $\mu(\mathcal{C}_n, t)$, we consider two ensemble average observables, one for speed and one for dissipation, for the paths from an initial state x_0 to a target state x_f. The amount of time it takes a stochastic trajectory on average to follow a path $\mathcal{C}_n$ is the path occurrence time $\langle\tau\rangle_{\mathcal{C}_n} = \sum_{i=0}^{n} w_{x_i}^{-1}$, the cumulative mean of the independent and exponentially distributed waiting times along the path. Its inverse $1/\langle\tau\rangle_{\mathcal{C}_n}$ is a measure of the "speed" at which the system traverses a single path. We first analyze paths of length n connecting two states, measuring the speed with the ensemble average for a given n: $\langle1/\tau\rangle_n :=$ $\sum_{\mathcal{C}_n}\langle\tau\rangle_{\mathcal{C}_n}^{-1}\mu(\mathcal{C}_n, \langle\tau\rangle_{\mathcal{C}_n})$, and evaluating the path probability at the path occurrence time, $\langle\tau\rangle_{\mathcal{C}_n}$. For the set of paths with length n, we also analyze the average entropy dissipated, $-\Delta_e S_n/k_B = -\sum_{\mathcal{C}_n}(s_e[\mathcal{C}_n]/k_B)\mu(\mathcal{C}_n, \langle\tau\rangle_{\mathcal{C}_n})$.

MARKOV MODEL FOR RELATIONSHIP BETWEEN PATH SPEED AND DISSIPATION

Equipped with the contracted path probability, we built a minimal model to control these measures of speed and dissipation and to test the hypothesis that faster paths will dissipate more, Fig. 1(a). At small length and time scales, if systems dissipate more to actuate structure formation on a specified timescale [59–62] they must often evolve through a set of intermediate states that separate the initial and target states [63, 64]. To represent this physical scenario, we adapt Onsager's three-state cycle, which he used for illustrating detailed balance (breaking) [65], by adding additional states. This expanded model is a discrete state Markov model of dissipative self-assembly in which paths that lead downhill when $\epsilon_b > 0$ (or uphill if $\epsilon_b < 0$) in energy connect an initial state to a final state, with an impeding dissipative cycle.
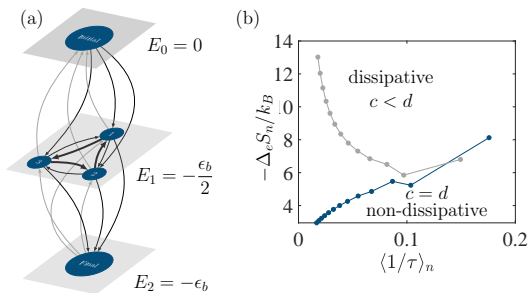
FIG. 1. *Completing a nonequilibrium process more quickly can result in more or less dissipated entropy on average.* (a) Model with three energy levels in which a potentially dissipative cycle is intermediate between initial (high energy) and final (low energy) states. Transition rates up the energy gradient (gray) are a function $\exp(-\beta\epsilon_b/2)$ of the effective binding energy $\beta\epsilon_b$ in units of $k_B T$. Those down the energy gradient (thin black) represent the concentration of monomers, c. Transitions around the cycle can dominate the entropy flow. Clockwise (thick black) transitions with rate constant d can be tuned relative to counterclockwise (thin black) with rate constant c to control the dissipation. (b) Dissipation $-\Delta_e S_n/k_B$ versus speed $\langle\tau^{-1}\rangle_n$, both conditioned on path length n, for the model in (a). When $c = d$ (blue), dissipation increases with the average speed of paths that connect unassembled and assembled states. However, when $c < d$ (gray), dissipation decreases with the average rate of path completion.

Stochastic paths are often numerically sampled or explicitly enumerated [28, 66–69]. Computationally, we evaluate the path ensemble averages of speed and dissipation by enumerating paths up to $n = 15$ that connect the high and low energy states. Taking the initial marginal probability $p(x_0 = 1) = 1$, we compute the probability of each with our exact expression, but we could also compute the probability of preferred paths (e.g., high probability paths accounting for specified percentage of probability flux) for larger networks using a depth-first search procedure with path-culling criteria [37].

The model has three parameters representing common experimental control variables: the relative stability of the monomers and assembled structure (the binding energy, ϵ_b), the concentration of the monomeric units, c, and a dissipation parameter d. The relationship between c and d controls the dissipation associated with traversal of a step of the cycle, $\pm\ln c/d$, which is zero when $c = d$. In a physical system, the parameter d could be the concentration of a fuel consumed to drive the process, the diffusion of matter in space, or some energy input driving the system through the cycle states.

By design, these parameters provide control over both the speed of paths connecting the high (low) energy initial state and the low (high) energy final state and the associated dissipation. With increasing path length, n, the entropy dissipated $-\Delta_e S_n/k_B$ can increase or decrease, depending on the exact values of the parameters. However, the speed $\langle 1/\tau\rangle_n$ decreases with n because

longer paths have more terms $(1/w_{x_i})$ contributing to the mean path occurrence time; a longer amount of time to complete results in a slower speed, hence the decrease with n. (We note, however, that this correlation may not hold in networks whose escape rates are not similar in magnitude.) Varying these control parameters also modulates the competition between the speed of path completion and the associated entropy dissipated, showing that faster paths do not necessarily dissipate more. Figure 1(b) shows that in kinetic networks of this type, the dissipation $-\Delta_e S_n/k_B$ can increase with the speed $\langle 1/\tau\rangle_n$, as one would expect from equilibrium thermodynamics, when there is no preferred direction among the intermediate states. However, when there are nonequilibrium currents and a strongly preferred direction among the intermediate states, then faster paths dissipate less, Fig. 1(b) (gray).

To understand this behavior, first suppose that the concentration parameters are nearly equal, $c \approx d$. In this case, the energy gradient determines the dependence of dissipation on the timescale of path completion, Fig. 1(b)(blue). Only the sequence of states down the energy gradient contribute to the path entropy flow $-s_e[\mathcal{C}_2]/k_B = 2\ln(c) + \beta\epsilon_b$. [The behavior is similar when cycle transitions are removed, Supplemental Material (SM).] Longer paths, which necessarily include additional jumps around the cycle, take longer, but do not dissipate more. Consequently, the ensemble averages $-\Delta_e S_n/k_B$ and $\langle 1/\tau\rangle_n$ are positively correlated because they both decrease with path length n. Now, when $c \neq d$ and paths are long enough that jumps around the cycle are significant, each transition adds $\pm\ln c/d$ to the entropy flow. Contributions to the entropy flow in one direction of the cycle can be negated by subsequent transitions in the reverse direction, so only net transitions on the cycle contribute. Paths that make K transitions around the cycle in a particular direction will dissipate an additional $\pm K\ln c/d$. The average entropy dissipated increases with path length n. As a result, it is negatively correlated with the mean path occurrence time $\langle\tau\rangle_n$, which increases linearly with n, Fig. 1(b)(gray).

Because of this balance between dissipation (traversal of the cycle) and speed (traversal of the energy gradient), there are regions of parameter space where the dissipation is an increasing or decreasing function of the rate of path completion, Fig. 2. Transitions between these regimes, where equilibrium thermodynamic intuition does and does not hold, are controlled by the relative magnitudes of the cycle transition rates, c/d, and the relative binding energy, $\beta\epsilon_b$. Dissipation and speed are positively correlated for $c/d \approx 1$, when jumps around the cycle do not significantly increase dissipation. Overall, increasing the speed only increases the dissipation when the cycle is weakly dissipative, $c \approx d$.

Further, the thermodynamic stability of the assembled state (determined by the binding energy $\beta\epsilon_b$) has
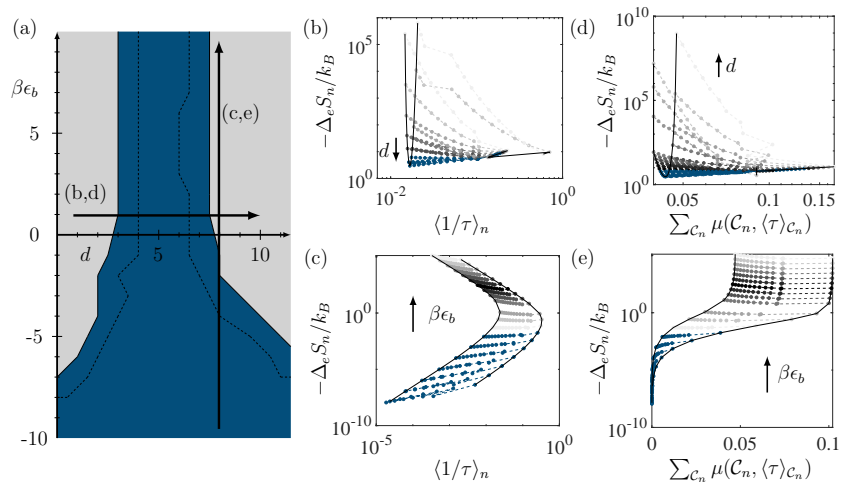
FIG. 2. *Model parameters control whether faster paths dissipate more or less entropy on average.* (a) Diagram mapping the sign of the slope from $-\Delta_e S_n/k_B$ as a function of $\langle 1/\tau \rangle_n$ for a cross section of the parameter space $(\beta \epsilon_b, d \mid c = 5)$. For $\beta \epsilon_b > 0$, the dissipation increases with speed when $c \approx d$ (blue). For $\beta \epsilon_b < 0$, the width of this region grows with increasingly negative $\beta \epsilon_b$. Arrows indicate the parameter sweeps shown in (b-e). (b) When paths are resolved by the number of jumps between the initial and final states (and averaged), the entropy dissipated to the surroundings, $-\Delta_e S_n/k_B$, can increase or decrease as a function of the mean rate of path completion $\langle 1/\tau \rangle_n$. The slope is controlled by the current around the cycle, which we tune by sweeping $d \in [0.5, 10]$ in increments of 0.5 at fixed $c = 5$ and $\beta \epsilon_b = 1$. Dissipation increases with the rate of path completion (blue) for $c \approx d$ but decreases sharply when $|c - d| \gg 0$ (grayscale). Solid lines connecting the slowest paths for the parameter sweep show a parabolic trend (black). (c) Dissipation, $-\Delta_e S_n/k_B$, as a function of the mean rate of path completion $\langle 1/\tau \rangle_n$ resolved (averaged) by the path length n. The slope is controlled by sweeping $\beta \epsilon_b \in [-10, 10]$ in increments of 1 with fixed $c = 5$ and $d = 8$. (d) The yield $\sum_{\mathcal{C}_n} \mu(\mathcal{C}_n, \langle \tau \rangle_{\mathcal{C}_n})$ has the same relationship to the dissipation $-\Delta_e S_n/k_B$ as the speed $\langle 1/\tau \rangle_n$. (e) Dissipation, $-\Delta_e S_n/k_B$, as a function of the yield $\sum_{\mathcal{C}_n} \mu(\mathcal{C}_n, \langle \tau \rangle_{\mathcal{C}_n})$ resolved (averaged) by the path length n. The yield plateaus above $\beta \epsilon_b = 1$.

an effect on where exactly this border between dissipative and weakly/non dissipative cycles lies in parameter space. Widening the energy-level gap when $\beta \epsilon_b < 0$ increases the width of the area in $|\log c/d|$ parameter space where the hypothesis holds, Fig. 2(a)(blue). Under these conditions, paths are uphill in energy ($\beta \epsilon_b < 0$), so an increasing amount of energy is required to force the system into the energetically unfavorable assembled state. These paths are then generally lower in probability with less entropy dissipation than their counterparts with higher $\beta \epsilon_b$, Fig. 2(e). Further, the probability decays (exponentially with n) more quickly than the entropy dissipation increases (linearly) with n. Therefore, the average dissipation decreases with n and, as a result, speed and dissipation are positively correlated for sufficiently negative $\beta \epsilon_b$, Fig. 2(b). We see the same correlation when the cycle transitions are removed and the binding energy is negative (SM).

These observations for paths conditioned on their length translate into the average speed, $\langle 1/\tau \rangle$, and dissipation, $-\Delta_e S/k_B$, over the entire ensemble of paths, Fig. 3. For example, the dissipation can be averaged over paths of various lengths $0 \leq n \leq n_{\max}$, $-\Delta_e S/k_B = -\sum_{n=0} \sum_{\mathcal{C}_n} (s_e[\mathcal{C}_n]/k_B) \mu(\mathcal{C}_n, \langle \tau \rangle_{\mathcal{C}_n})$. We systematically varied the model parameters and identified conditions where speed is maximal and this measure of dissipation is minimal. Scanning values of $\beta \epsilon_b$ with fixed c and d, there is a maximum average speed located at $c = e^{-\beta \epsilon_b}$,

when jumps up and down the energy gradient have the same transition rate, Fig. 3(a). Scanning d with fixed values of $\beta \epsilon_b$ and c, there is a minimum average entropy dissipation located at $c = d$, Fig. 3(b). These extrema are also apparent in Fig. 2(b-e). Their locations mark the transition of ensemble averages between regimes of positive and negative correlation between speed and dissipation.

Extremizing with respect to the model parameters, we find that minimizing dissipation does not simultaneously maximize speed or vice-versa in this model. Hence, dissipation cannot be at its global minimum when speed is at its global maximum. The maximal speed is dictated by the thermodynamic stability of the final state controlled by the binding energy $\beta \epsilon_b$, Fig. 3(a). The minimal dissipation is largely dictated by the nonequilibrium currents around the intermediate states controlled by the kinetic coefficients d/c, Fig. 3(b). To simultaneously minimize dissipation and maximize speed, we use the fact that the parameters provide independent control and extremize each observable in sequence. Fig. 3(d-bottom) shows that the minimum average entropy dissipated increases exponentially with $\beta \epsilon_b$, and with different growth rates on either side of $\beta \epsilon_b = -\ln c$. We see in Fig. 3(b) that this minimum entropy dissipated occurs for a given value of $\beta \epsilon_b$ at $c = d$. At approximately the same relative binding energy, the associated average rate has a maximum, Fig. 3(d-top). This maximum suggests that the most op-
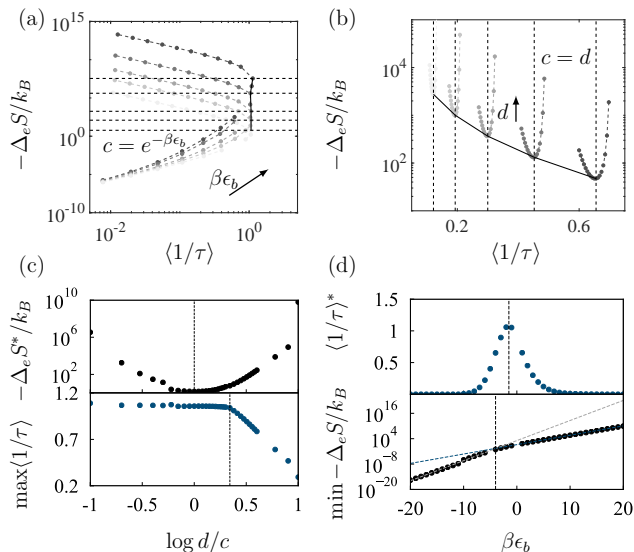
FIG. 3. *Speed and dissipation extrema in the parameter space.* (a) Entropy dissipated $-\Delta_e S/k_B$ versus speed $\langle 1/\tau \rangle$ averaged over all paths varying $\beta\epsilon_b \in [-10, 10]$ with fixed $c = 5$ and $d = \{0.5, 0.6, 1, 1.5, 4\}$. Darker gray indicates larger values of d. Dashed horizontal lines indicate $c = e^{-\beta\epsilon_b}$ for each d. Solid black line connects vertices for different d values. (b) Ensemble level dissipation $-\Delta_e S/k_B$ versus speed $\langle 1/\tau \rangle$ for $d \in [1.5, 10]$ in increments of 0.5 and other parameters fixed at $c = 5$ and $\beta\epsilon_b \in [1, 4]$ in increments of 1 (darker gray indicates lower value of $\beta\epsilon_b$). Dashed vertical lines indicate $c = d$ for each fixed $\beta\epsilon_b$ value. Solid black line connects vertices for different $\beta\epsilon_b$ values. (c) Maximum average speed (bottom) and associated average entropy dissipated $-\Delta_e S^*/k_B := -\Delta_e S/k_B|_{\max\langle 1/\tau \rangle}$ (top) for $d \in [0.5, 10]$ and fixed $c = 5$. (d) Minimum entropy dissipated on average (bottom) and associated speed $\langle 1/\tau \rangle^* := \langle 1/\tau \rangle|_{\min(-\Delta_e S/k_B)}$ (top) for $\beta\epsilon_b \in [-10, 10]$ for fixed $c = 5$. Dashed lines are exponential trends for $\beta\epsilon_b > -\ln c$ (blue) and $\beta\epsilon_b < -\ln c$ (gray).

timal combination of extremizing both the entropy dissipated and the rate is obtained with a homogeneous kinetic network, $c = d = e^{-\beta\epsilon_b}$. Similarly, the maximal average rate occurs for $\log d/c < 0.4$, Fig. 3(c-bottom). We see from Fig. 3(a) that the maximum average rate for a given d occurs at $c = e^{-\beta\epsilon_b}$. Within this range the associated entropy dissipated has a minimum at $c \approx d$, Fig. 3(c-top). So, for this model, the minimization of the mean entropy dissipated and maximization of the mean speed at this level commute.

CONCLUSIONS

Natural and synthetic systems often balance the speed of transitioning between states and the associated cost of dissipation. Understanding this balance has the potential to benefit the rational design of synthetic systems that are adaptive and responsive to their environment. As a step in this direction, we derived an analytical formula for the occurrence probability of stochastic paths through the path summation solution of the master equation. This formula provides a means to identify and assess the kinetic relevance of paths for stochastic processes over a specified time frame and without necessarily constraining endpoints. This contracted path probability is necessary to directly calculate ensemble-level observables, such as entropy production and flow, from a set of paths explicitly constrained to complete the process in a fixed time. Applying this formula to a model for self-assembly showed that increasing speed need not be accompanied by increasing dissipation and, in some cases, dissipation cannot be minimized when speed is maximized. Based on these results, networks governing the dynamics of other physical and chemical systems, such as biochemical reaction cycles and the dissipative cycles of chemically-active materials, might be tuned to control whether faster paths dissipate less than slower paths on average, counter to the equilibrium thermodynamic intuition that "more haste" brings "more waste".

* jason.green@umb.edu
[1] R. Marsland and J. England, Rep. Prog. Phys. **81**, 016601 (2017).
[2] B. Mishra and M. E. Johnson, J. Chem. Phys. **154**, 194101 (2021).
[3] A. Desai and T. J. Mitchison, Annu. Rev. Cell Dev. Biol. **13** (1997), 10.1146/annurev.cellbio.13.1.83.
[4] D. A. Head, W. J. Briels, and G. Gompper, Phys. Rev. E **89**, 032705 (2014).
[5] J. Boekhoven, W. E. Hendriksen, G. J. M. Koper, R. Eelkema, and J. H. van Esch, Science **349**, 1075 (2015).
[6] K. J. M. Bishop, C. E. Wilmer, S. Soh, and B. A. Grzybowski, Small **5**, 1600 (2009).
[7] E. Penocchio, R. Rao, and M. Esposito, Nat. Commun. **10**, 3865 (2019).
[8] S. Whitelam and R. L. Jack, Annu. Rev. Phys. Chem. **66**, 143 (2015).
[9] T. Yasui, N. Kaji, R. Ogawa, S. Hashioka, M. Tokeshi, Y. Horiike, and Y. Baba, Nano Letters **15**, 3445 (2015).
[10] M. Tagliazucchi, E. A. Weiss, and I. Szleifer, Proc. Natl. Acad. Sci. U.S.A. **111**, 9751 (2014).

[11] M. Nguyen and S. Vaikuntanathan, Proc. Natl. Acad. Sci. U.S.A. **113**, 14231 (2016).

[12] R. Rao and M. Esposito, New J. Phys. **20**, 023007 (2018).

[13] W. M. Jacobs and D. Frenkel, J. Am. Chem. Soc. **138**, 2457 (2016).

[14] Z. Zhou, Y. Ouyang, J. Wang, and I. Willner, J. Am. Chem. Soc. **143**, 5071 (2021).

[15] S. A. P. van Rossum, M. Tena-Solsona, J. H. van Esch, R. Eelkema, and J. Boekhoven, Chem. Soc. Rev. **46**, 5519 (2017).

[16] S. J. Bryant and B. B. Machta, Proc. Natl. Acad. Sci. U.S.A. **117**, 3478 (2020).

[17] Y. Song and C. Hyeon, J. Phys. Chem. Lett. **11**, 3136 (2020).

[18] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (John Wiley & Sons, Inc., 1985).

[19] L. Peliti and S. Pigolotti, *Stochastic Thermodynamics* (Princeton Univ. Press, 2021).

[20] N. Shiraishi, K. Funo, and K. Saito, Phys. Rev. Lett. **121**, 070601 (2018).

[21] G. Falasco and M. Esposito, Phys. Rev. Lett. **125**, 120604 (2020).

[22] G. Falasco, M. Esposito, and J. C. Delvenne, New J. Phys. **22**, 053046 (2020).

[23] S. Ito and A. Dechant, Phys. Rev. X **10**, 021056 (2020).

[24] J. M. Horowitz and T. R. Gingrich, Nat. Phys. **16**, 15 (2020).

[25] S. B. Nicholson, L. P. García-Pintos, A. del Campo, and J. R. Green, Nat. Phys. **16**, 1211 (2020).

[26] B. Altaner, S. Grosskinsky, S. Herminghaus, L. Katthän, M. Timme, and J. Vollmer, Phys. Rev. E **85**, 041133 (2012).

[27] J. R. Green, A. B. Costa, B. A. Grzybowski, and I. Szleifer, Proc. Natl. Acad. Sci. U.S.A. **110**, 16339 (2013).

[28] S. B. Nicholson, R. A. Bone, and J. R. Green, J. Phys. Chem. B **123**, 4792 (2019).

[29] S. B. Nicholson, M. Alaghemandi, and J. R. Green, J. Chem. Phys. **145**, 084112 (2016).

[30] S. B. Nicholson, M. Alaghemandi, and J. R. Green, J. Chem. Phys. **148**, 044102 (2018).

[31] D. J. Sharpe and D. J. Wales, J. Chem. Phys. **151**, 124101 (2019).

[32] T. D. Swinburne, D. Kannan, D. J. Sharpe, and D. J. Wales, J. Chem. Phys. **153**, 134115 (2020).

[33] J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope, J. Chem. Phys. **126**, 04B616 (2007).

[34] J. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé, J. Chem. Phys. **134**, 174105 (2011).

[35] D. Wales, *Energy landscapes: Applications to clusters, biomolecules and glasses* (Cambridge University Press, 2003).

[36] D. Helbing, Phys. Lett. A **195**, 128 (1994).

[37] D. Helbing and R. Molini, Phys. Lett. A **212**, 130 (1996).

[38] S. X. Sun, Phys. Rev. Lett. **96**, 210602 (2006).

[39] D. T. Gillespie, *Markov processes: An introduction for physical scientists* (Academic Press, 1991).

[40] M. R. D'Orsogna, G. Lakatos, and T. Chou, J. Chem. Phys. **136**, 084110 (2012).

[41] K. Head-Marsden, J. Flick, C. J. Ciccarino, and P. Narang, Chem. Rev. **121**, 3061 (2021).

[42] D. Seiferth, P. Sollich, and S. Klumpp, Phys. Rev. E **102**, 062149 (2020).

[43] J. van der Meer, B. Ertel, and U. Seifert, arXiv.org (2022), 10.485550/arXiv.2203.12020.

[44] D. J. Skinner and J. Dunkel, Phys. Rev. Lett. **127**, 198101 (2021).

[45] M. F. Weber and E. Frey, Rep. Prog. Phys. **80**, 046601 (2017).

[46] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).

[47] C. Van den Broeck and M. Esposito, Physica A **418**, 6 (2015).

[48] C. Maes, SciPost Phys. Lect. Notes , 32 (2021).

[49] U. Seifert, Phys. Rev. Lett **95**, 040602 (2005).

[50] G. Falasco and M. Esposito, Phys. Rev. E **103**, 042114 (2021).

[51] K. H. Hoffmann, B. Andresen, and P. Salamon, Phys. Rev. A **39**, 3618.

[52] J. M. Horowitz, J. Chem. Phys. **143**, 044111 (2015).

[53] S. Bai, D. Zhou, M. J. Davis, and R. T. Skodje, J. Phys. Chem. Lett. **6**, 183 (2015).

[54] S. Bai, M. J. Davis, and R. T. Skodje, J. Phys. Chem. A **119**, 11039 (2015).

[55] S. Bai and R. T. Skodje, J. Phys. Chem. Lett. **8**, 3826 (2017).

[56] R. Bone and J. Green, "contracted-paths," (2022), code is available at the Github repository https://doi.org/10.5281/zenodo.6823402.

[57] O. Flomenbom, J. Klafter, and R. J. Silbey, Phys. Rev. Lett. **97**, 178901 (2006).

[58] S. X. Sun, Phys. Rev. Lett. **97**, 178902 (2006).

[59] B. A. Grzybowski, C. E. Wilmer, J. Kim, K. P. Browne, and K. J. M. Bishop, Soft Matter **5**, 1110 (2009).

[60] S. Whitelam, C. Rogers, A. Pasqua, C. Paavola, J. Trend, and P. L. Geissler, Nano. Lett. **9**, 292 (2009).

[61] A. Arvind, D. A. Huse, and S. Leibler, Proc. Natl. Acad. Sci. U.S.A. **109**, 12034 (2012).

[62] X. Fang, K. Kruse, T. Lu, and J. Wang, Rev. Mod. Phys. **91**, 045004 (2019).

[63] S. Komine, S. Takahashi, T. Kojima, H. Sato, and S. Hiraoka, J. Am. Chem. Soc. **141**, 3178 (2019).

[64] B. G. P. van Ravensteijn, I. K. Voets, W. K. Kegel, and R. Eelkema, Langmuir **36**, 10639 (2020).

[65] L. Onsager, Phys. Rev. **37**, 405 (1931).

[66] S. A. Trygubenko and D. J. Wales, Mol. Phys. **104**, 1497 (2006).

[67] J. Wang, J. Onuchic, and P. Wolynes, Phys. Rev. Lett. **76**, 4861 (1996).

[68] H. Feng, B. Han, and J. Wang, J. Phys. Chem. Lett. **1**, 1836 (2010).

[69] S. B. Nicholson, J. S. Greenberg, and J. R. Green, Phys. Rev. E **97**, 012146 (2018).

[70] M. Athénes and V. V. Bulatov, Phys. Rev. Lett. **113**, 230601 (2014).

[71] A. F. Voter, Radiat. Eff. Solids **235**, 1 (2007).

APPENDIX

We derived an explicit, closed form expression for the contracted path probability:

$$\mu(\mathcal{C}_n, t) = p(x_0, t_0) \prod_{i=0}^{n-1} w(x_{i+1}|x_i) \sum_{j=1}^{n'} \left(\frac{(-1)^{m_j-1}}{(m_j-1)!} \right) \left(\frac{e^{-w_{x_j}t}}{\prod_{\substack{k=1\\k\neq j}}^{n'} (w_{x_k} - w_{x_j})^{m_k}} \right) \sum_{l=1}^{m_j} \binom{m_j-1}{l-1} (-t)^{m_j-l} d_j^{(l-1)}. \quad (13)$$

One way to derive this result is to explicitly evaluate the successive derivatives in Eqn. 6. These are derivatives of the function

$$f_j^{(0)} = \frac{e^{-w_{x_j}t}}{\prod_{\substack{k=1\\k\neq j}}^{n'} (w_{x_k} - w_{x_j})^{m_k}} \quad (14)$$

with respect to the jth unique escape rate w_{x_j}. Here, $f^{(l)}$ indicates the lth derivative of f. These derivatives can be evaluated using the product rule, where $f_j^{(0)} = g(w_{x_j})h(w_{x_j})$ is composed of the functions $g(w_{x_j}) = e^{-w_{x_j}t}$ and $h(w_{x_j}) = \left[\prod_{k=1,k\neq j}^{n'} (w_{x_k} - w_{x_j})^{m_k} \right]^{-1}$. Because of the product in $h(w_{x_j})$, each derivative of f_j results in a series of terms, all of which have a factor $f_j^{(0)}$. Each successive derivative (with respect to the same chosen unique escape rate) also accumulates a series of terms t^0 to t^{m_j-1}. These terms have an alternating sign from the qth derivative of $g(w_{x_j})$: $(-t)^q e^{-w_{x_j}t}$. The terms of order t^0 and t^{m_j-1} correspond to the pure derivatives of the h and g, respectively.

The pure derivative of the denominator, h, has a form that is harder to evaluate; each factor the product within $h(w_{x_j})$ requires application of the product rule and chain rule. The lth derivative for the kth factor is:

$$\frac{\partial^l}{\partial w_{x_j}^l} \frac{1}{(w_{x_k} - w_{x_j})^{m_k}} = \frac{(-1)^l \prod_{u=0}^{l-1} (m_k + u)}{(w_{x_k} - w_{x_j})^{m_k+l}}. \quad (15)$$

This formula is sufficient to determine the first derivative of $h(w_{x_j})$,

$$h^{(1)}(w_{x_j}) = \frac{-1}{\prod_{\substack{k=1\\k\neq j}}^{n'} (w_{x_k} - w_{x_j})^{m_k}} \sum_{\substack{l=1\\l\neq j}}^{n'} \frac{m_l}{w_{x_l} - w_{x_j}}, \quad (16)$$

where the sum accounts for the derivative of each factor in the product.

For higher order derivatives of $h(w_{x_j})$ with respect to a unique escape rate, we must take into account all combinations of orders of derivatives among the terms in the product. In each combination, the order of the derivatives of each term in the product must sum to the total order of the derivative of $h(w_{x_j})$. The original function $h(w_{x_j})$ is an eigenfunction with respect to the derivative operator, just as every derivative of $g(w_{x_j})$ contains $g(w_{x_j})$: $\partial^k/\partial w_{x_j}^k \left(e^{-w_{x_j}t} \right) = (-t)^k e^{-w_{x_j}t}$. So, $f_j^{(0)}$ is also an eigenfunction of $\partial^k/\partial w_{x_j}^k$ with eigenvalues that account for the multiplicities of the unique escape rates. The eigenvalues have m_j terms, where m_j is the degeneracy of the unique escape rate chosen. One of the terms is the pure k^{th} derivative of h (multiplied by g) and another is the pure k^{th} derivative of g (multiplied by h). The remaining $m_j - 2$ terms of index l account for mixed derivatives of h and g. The number of occurrences of each of these mixed terms is represented by the binomial coefficient $\binom{m_j-1}{l-1}$ appearing in Eqn. 13. Each mixed term is the product of the $(l-1)^{\text{th}}$ derivative of $g(w_{x_j})$ and the $(m_j - l)^{\text{th}}$ derivative of $h(w_{x_j})$.

The closed form expression of the contracted path probability in Eqn. 13 has one remaining term, $d_j^{(l-1)}$. This term has the form of the $h(w_{x_j})$ portion of the eigenvalue of f_j (i.e., the portion of $f_j^{(m_j-1)}/f_j^{(0)}$ corresponding to the derivative(s) of the denominator in Eqn. 14). For the first derivative, $l - 1 = 1$, this function is:

$$d_j^{(1)} = \sum_{\substack{\alpha=1\\\alpha\neq j}}^{n'} \frac{m_\alpha}{w_{x_\alpha} - w_{x_j}}, \quad (17)$$

which is what we expect from the first derivative of $h(w_{x_j})$ in Eqn. 16. The key to finding the explicit expression for the function $d_j^{(l-1)}$ is recognizing that their structure is related to the unrestricted partitions U_{l-1} of $l - 1$. The number of terms in $d_j^{(l-1)}$ is the number of unrestricted partitions $|U_{l-1}|$ of $l - 1$. For example, if

$l - 1 = 2$ then:

$$d_j^{(2)} = \sum_{\substack{\alpha=1 \\ \alpha \neq j}}^{n'} \frac{m_\alpha(m_{\alpha+1})}{(w_{x_\alpha} - w_{x_j})^2}$$

$$+ \sum_{\substack{\beta=1 \\ \beta \neq j}}^{n'} \sum_{\substack{\gamma=1 \\ \gamma \neq j,\beta}}^{n'} \frac{m_\beta m_\gamma}{(w_{x_\beta} - w_{x_j})(w_{x_\gamma} - w_{x_j})}, \quad (18)$$

with two terms corresponding to the two partitions $|U_2| = 2$ of $l - 1 = 2$. The sums are over the n' unique escape rates where $w_{x_\alpha} \neq w_{x_j}$; the multiplicity of the unique escape rate α is m_α. We can recognize that each term is related to an unrestricted partition: the first term corresponds to the partition (2) and the second term corresponds to the partition (1,1). More generally, we can label each term as the qth partition, U_{k-1}^q.

Also noteworthy is that each term has numerical coef-

ficients. In the examples above, these coefficients are all one. However, for the qth partition, the number of coefficients, χ, in the partition $|U_{k-1}^q|$ is the number of sums for that term. The coefficient of the kth term in $d_j^{(l-1)}$ is also related to the unrestricted partitions of $l - 1$ by:

$$C(l-1, k) = \frac{(l-1)!}{\prod_{v=1}^{|U_{l-1}^k|} (\chi_v)! \prod_{w=1}^{|U_{l-1}^k|'} (\lambda_w)!} \quad (19)$$

where there are $|U_{l-1}^q|'$ unique values in the qth partition of $l - 1$ and the vth unique coefficient χ_v in the qth partition of $l - 1$ has a multiplicity of λ_v. With these observations, we can construct any $d_j^{(l-1)}$. For example, in Eqn. 18 we can recognize that the first term corresponds to the partition (2) and has coefficient $C(2,2) = 1$, while the second term corresponds to the partition (1,1) and has coefficient $C(2,1) = 1$. We confirmed these coefficients for $l - 1 \leq 12$.