

Correlation thresholds in the steady states of particle systems and spin glasses

Jacob Calvert^{1,2,*} and Dana Randall^{3,2,†}

¹*Institute for Data Engineering and Science, Georgia Institute of Technology, Atlanta, GA, USA*

²*Santa Fe Institute, Santa Fe, NM, USA*

³*School of Computer Science, Georgia Institute of Technology, Atlanta, GA, USA*

(Dated: August 22, 2025)

A growing body of theoretical and empirical evidence shows that the global steady-state distributions of many equilibrium and nonequilibrium systems approximately satisfy an analogue of the Boltzmann distribution, with a local dynamical property of states playing the role of energy. The correlation between the effective potential of the steady-state distribution and the logarithm of the exit rates determines the quality of this approximation. We demonstrate and explain this phenomenon in a simple one-dimensional particle system and in random dynamics of the Sherrington–Kirkpatrick spin glass by providing the first explicit estimates of this correlation. We find that, as parameters of the dynamics vary, each system exhibits a threshold above and below which the correlation dramatically differs. We explain how these thresholds arise from underlying transitions in the relationship between the local and global “parts” of the effective potential.

I. INTRODUCTION

The steady-state distributions of many physical systems favor a relatively small number of special states. For example, this is true of proteins, which tend to adopt conformations of low energy [1], as well as active matter systems, like swarms of interacting robots that sustain long-lasting, near-periodic “dances” [2]. For proteins, and for systems in thermal equilibrium in general, the Boltzmann distribution explains that states with lower energy are exponentially favored. More precisely, the steady-state probability $\pi(x)$ of a state x satisfies

$$\pi(x) = e^{-\beta H(x)} / Z, \quad (1)$$

in terms of inverse temperature β , energy or Hamiltonian H , and partition function $Z = \sum_x e^{-\beta H(x)}$.

The explanation of order that the Boltzmann distribution provides is powerful because the energy $H(x)$ is a “local” property of state x , in the sense that it can be determined without observing long trajectories of the system to states far from x . For example, the energy of a protein conformation can be estimated from its coordinates using molecular dynamics force fields [3]. In contrast, the steady-state distribution of a nonequilibrium system is “global” because estimating it from a trajectory generally requires the observation of multiple returns to state x , which can entail visits to distant states (e.g., see [4]). For nonequilibrium steady states, there can be no local explanation of order analogous to the one the Boltzmann distribution provides, as Landauer explained [5]. Essentially, the issue with nonequilibrium steady states is that changes to the dynamics in one part of the state space can affect the relative probability of states in another [6].

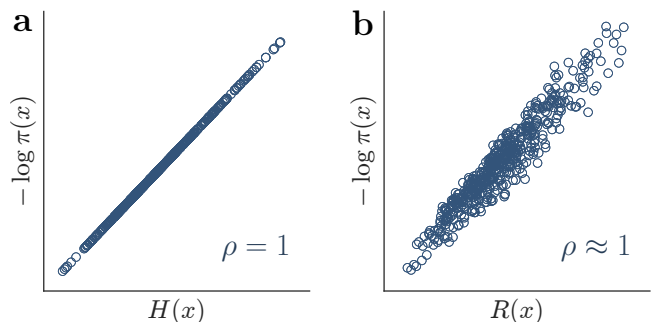


FIG. 1. Local-global correlations. (a) The Boltzmann distribution entails perfect correlation ρ between the effective potential of an equilibrium steady state and the energy of a uniformly random state. (Each mark corresponds to one state.) (b) For many equilibrium and nonequilibrium steady states, a different local property of a state, called rattling, is highly correlated with the effective potential.

Despite this barrier, a local dynamical property of states called “rattling” predicts the steady-state distributions of many nonequilibrium systems [2, 7–10]. Informally, rattling is a measure of how quickly the system exits the vicinity of a state, like a diffusion coefficient in abstract state space, but for processes that are not necessarily diffusive. In their study of robot swarms, Chvykov et al. observed that plotting the effective potential $-\log \pi(x)$ against the rattling $R(x)$ of discretized swarm configurations x produced a roughly linear scatter [2]. This observation is analogous to an approximate Boltzmann distribution because the effective potential and $H(x)$ are exactly collinear in equilibrium steady states (Fig. 1).

These observations motivate the study of correlations between the effective potential of a steady state and local properties of the states. In recent work [11], we derived a general formula for the kind of correlation that Chvykov et al. observed, which Fig. 1b depicts. It applies to Markov jump processes, which are commonly used to

* calvert@gatech.edu

† randall@cc.gatech.edu

model the dynamics underlying equilibrium and nonequilibrium steady states [12–16]. The formula can be used to derive lower bounds on the typical correlation exhibited by broad classes of steady states, such as those of reaction kinetics on disordered energy landscapes [17]. To complement these results, which are somewhat abstract, this paper provides the first concrete calculations of the local–global correlation. Although we focus on two specific models, we expect that our approach will apply to many other models as well.

II. LOCAL–GLOBAL CORRELATION IN MARKOV JUMP PROCESSES

A. Defining local–global correlation

We consider a Markov jump processes on a finite set S of discrete states, defined by a transition rate matrix $(W_{x,y})_{x,y \in S}$. The probability $p_y(t)$ of finding the system in state y after t units of time solves the master equation

$$p_y(t) = \sum_{x \in S} W_{x,y} p_x(t).$$

We assume that the transition rates strongly connect the state space, which guarantees the existence and uniqueness of the steady-state or stationary distribution $\pi = \lim_{t \rightarrow \infty} p(t)$ [18]. Our focus is on the relationship between the effective potential $-\log \pi(x)$ and the logarithm of the exit rates, defined by

$$q(x) = \sum_{y: y \neq x} W_{x,y}.$$

The quantity $\log q(x)$ is the analogue of rattling $R(x)$ for Markov jump processes [11]. Accordingly, to understand the collinearity of scatter plots like Fig. 1b, we analyze the correlation between the effective potential and log exit rate of a uniformly random state X in S :

$$\rho = \mathbf{Corr}(-\log \pi(X), \log q(X)). \quad (2)$$

We assume that π and q are non-constant, to ensure that ρ is well-defined. For reference, the linear correlation of random variables U and V with positive, finite variances is defined as the normalized covariance

$$\mathbf{Corr}(U, V) = \frac{\mathbf{Cov}(U, V)}{\sqrt{\mathbf{Var}(U) \mathbf{Var}(V)}}. \quad (3)$$

The covariance equals $\mathbf{Cov}(U, V) = \mathbf{E}(UV) - \mathbf{E}(U)\mathbf{E}(V)$, in terms of the expected value $\mathbf{E}(U)$ of U and variance $\mathbf{Var}(U) = \mathbf{Cov}(U, U)$.

B. The effective potential’s local and global parts

Associated with the preceding continuous-time Markov jump process is a discrete-time Markov jump process,

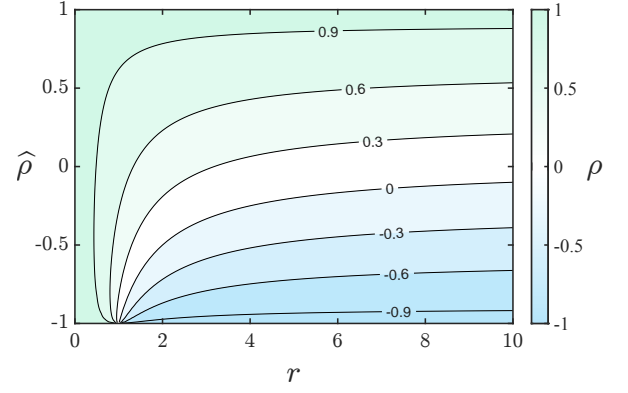


FIG. 2. Contour plot of Eq. (7). Contours and shading indicate the value of ρ for a given pair $(r, \hat{\rho}) \in [0, \infty) \times [-1, 1]$.

defined by normalizing the transition rates out of each state:

$$\widehat{W}_{x,y} = W_{x,y}/q(x). \quad (4)$$

This process also has a unique stationary distribution $\hat{\pi}$, which is related to π and q by

$$\pi(x) = \frac{\hat{\pi}(x)/q(x)}{\sum_{y \in S} \hat{\pi}(y)/q(y)}. \quad (5)$$

Intuitively, the limiting fraction $\pi(x)$ of time spent in state x is proportional to the limiting fraction $\hat{\pi}(x)$ of visits made to x , multiplied by the typical duration of a visit to x , which equals $1/q(x)$. Taking the logarithm of both sides of Eq. (5) shows that

$$-\log \pi(X) = -\log \hat{\pi}(X) + \log q(X) + \text{constant}. \quad (6)$$

Following the discussion in Section I, Eq. (6) justifies the view that $-\log \hat{\pi}(X)$ and $\log q(X)$ are the global and local “parts” of the effective potential.

The authors in [11] used Eq. (6) to derive the formula

$$\rho = \frac{1 + \hat{\rho}r}{\sqrt{1 + 2\hat{\rho}r + r^2}}, \quad (7)$$

in terms of

$$\hat{\rho} = \mathbf{Corr}(-\log \hat{\pi}(X), \log q(X)),$$

the correlation between the global and local parts, and

$$r = \sqrt{\frac{\mathbf{Var}(\log \hat{\pi}(X))}{\mathbf{Var}(\log q(X))}},$$

the ratio of their standard deviations. Note that $\hat{\rho} \in [-1, 1]$ and $r \in [0, \infty)$.

In principle, it should be as difficult to calculate $\hat{\rho}$ and r as it is to directly calculate ρ using Eq. (2). The value of Eq. (7) is therefore not in more easily calculating ρ . It does, however, imply a useful lower bound of ρ in terms

of r alone, which is tractable for broad classes of dynamics [17]. The virtue of Eq. (7) is that it explains how high correlation arises, in terms of the relationship between the parts of the effective potential.

Fig. 2, which visualizes Eq. (7), shows that there are two ways for the correlation ρ to be close to 1. First, if the global part of the effective potential varies far less than the local part (i.e., $r \ll 1$), then $\rho \approx 1$, regardless of the value of $\hat{\rho}$. Second, if the parts of the effective potential are highly correlated (i.e., $\hat{\rho} \approx 1$), then $\rho \approx 1$, regardless of the value of r . For ρ to be negative, the parts of the effective potential must be sufficiently anti-correlated: $\hat{\rho} < -1/r$.

III. PARTICLES ON A RING

We first demonstrate the calculation of ρ for a toy model of two particles moving clockwise around a discrete ring. We show that the local-global correlation is either close to 1 or close to -1 depending on a parameter that biases the particles' motion.

Consider two identical particles that move clockwise around a ring with $L+2$ discrete sites, where $L \geq 1$ is an integer. Each particle jumps clockwise to the next open site, at a rate that depends on the number of open sites “ahead” of it. Specifically, if there are $y \geq 1$ open sites in the clockwise direction from one particle to the next, then the first particle jumps at a rate of α^y , where $\alpha > 0$ is a parameter of the dynamics. If there are no open sites ahead of it, then the particle does not move. We arbitrarily choose one particle to be the “first” particle, and use x to denote the number of open sites to the second. Note that x , which takes values in $S = \{0, \dots, L\}$, completely determines the state of the system.

Based on the preceding description, the transition rate $W_{x,y}$ from x to y satisfies

$$W_{x,y} = \begin{cases} \alpha^x & y = x - 1, \\ \alpha^{L-x} & y = x + 1, \\ 0 & y \notin \{x - 1, x + 1\}, \end{cases} \quad (8)$$

for $x, y \in S$. As before, we denote the stationary distribution of the Markov chain with these transitions rates by π . The corresponding exit rates are

$$q(x) = \sum_{y: y \neq x} W_{x,y} = \begin{cases} \alpha^L & x \in \{0, L\}, \\ \alpha^x + \alpha^{L-x} & x \notin \{0, L\}. \end{cases}$$

Note that although $W_{x,y}$, π , and q depend on the parameter α , we omit it from the notation.

The main result of this section is an estimate of the local-global correlation that the particle system exhibits.

Theorem 1. *For every $\alpha \neq 1$, the correlation satisfies*

$$\rho(\alpha) = \frac{\sqrt{15}}{4} \text{sign}(\alpha - 1) (1 + O(L^{-1})), \quad (9)$$

as $L \rightarrow \infty$.

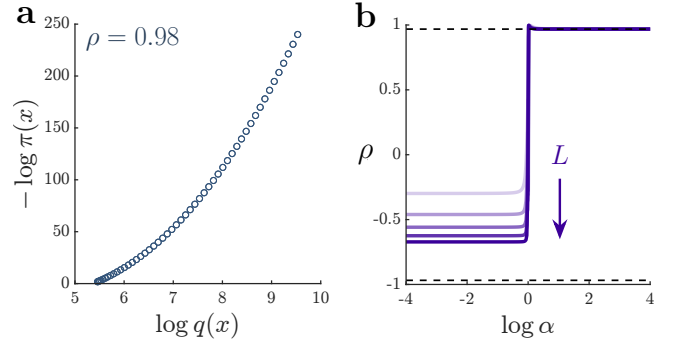


FIG. 3. Local-global correlations of particles on a ring. (a) The effective potential scattered against the log exit rate of each state for the dynamics in Eq. (8) with $L = 100$ and $\alpha = 1.1$. (b) The correlation as a function of L and α for $L = 25k$, $k \in \{2, \dots, 6\}$. Dashed lines indicate $\rho = \pm \sqrt{15}/4$.

Theorem 1 says that $\rho(\alpha)$ is close to $\sqrt{15}/4 \approx 0.97$ when $\alpha > 1$ and close to $-\sqrt{15}/4$ when $\alpha < 1$, for all sufficiently large L in terms of α . (In Eq. (9), we use the standard asymptotic notation $O(L^{-1})$ to denote a quantity that is at most a constant multiple of L^{-1} in absolute value for all sufficiently large L .) Fig. 3a shows a representative scatter plot of the effective potential against the log exit rates when $\alpha > 1$. For such values of α , the correlation is already close to its limiting value when $L = 100$. For $\alpha < 1$, ρ approaches $-\sqrt{15}/4$ more slowly (Fig. 3b).

The error term in Eq. (9) arises in part from approximating the exit rate $q(x)$ by $\alpha^{\max\{x, L-x\}}$ when $\alpha > 1$, and $\alpha^{\min\{x, L-x\}}$ when $\alpha < 1$. The proof of Theorem 1 in Appendix B finds that this approximation is sufficiently accurate when L is large relative to the quantity $1/\sqrt{|\alpha^4 - 1|}$. Hence, if $\alpha \geq 2$ or $\alpha \leq 1/2$, for example, then L only needs to be larger than some constant.

A. The correlation of a birth-death process

We outline the calculation of ρ in this subsection, deferring some details to Appendix B. We start by observing that the number of open sites x is a birth-death process, because a jump from x to y is only possible if $y = x \pm 1$. It is well known that the stationary distribution of such a process has the form

$$\frac{\pi(x)}{\pi(0)} = \prod_{y=1}^x \frac{W_{y-1,y}}{W_{y,y-1}}. \quad (10)$$

Alternatively, one can check that this π satisfies detailed balance, which verifies that it is the stationary distribution. Substituting the rates from Eq. (8) into Eq. (10), we find that the stationary distribution satisfies

$$\frac{\pi(x)}{\pi(0)} = \prod_{y=1}^x \frac{\alpha^{L-(y-1)}}{\alpha^y} = \alpha^{\sum_{y=1}^x (L-2y+1)} = \alpha^{x(L-x)}.$$

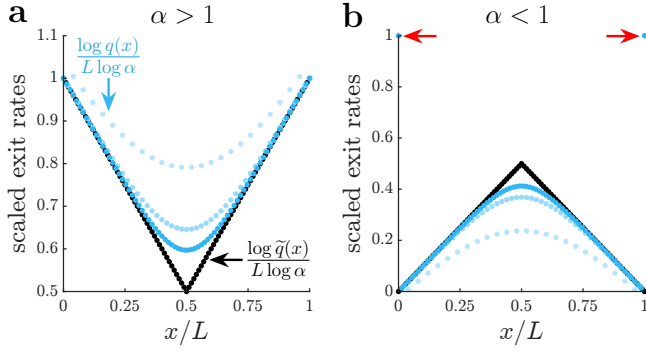


FIG. 4. Comparison of $\log q(x)$ (blue points) and $\log \tilde{q}(x)$ (black points), scaled by $1/L \log \alpha$, for (a) $\alpha = 1.1$ and (b) $\alpha = 0.9$. The blue scatters correspond to $L \in \{25, 50, 75\}$ (darker blue indicates higher L); the black scatter has $L = 100$. Note the points in (b) indicated by red arrows, which all blue scatters share.

Concerning the exit rates, we note that when L and $|\log \alpha|$ are large, q is roughly equal to

$$\tilde{q}(x) = \begin{cases} \alpha^{\max\{x, L-x\}} & \alpha > 1, \\ \alpha^{\min\{x, L-x\}} & \alpha < 1, \end{cases}$$

the logarithm of which has a simpler form than $\log q(x)$. Fig. 4 compares q and \tilde{q} . In particular, it highlights the fact that $\tilde{q}(x)$ underestimates $q(x)$ for some values of x and α , and overestimates it for others. However, aside from when $\alpha < 1$ and $x \in \{0, L\}$ (indicated by red arrows in Fig. 4b), $\tilde{q}(x)$ approximates $q(x)$ increasingly well as L increases. For the moment, we treat this approximation as exact and consider the case of $\alpha > 1$.

By substituting the expressions for $\pi(x)$ and $\tilde{q}(x)$ into Eq. (2), and by dropping the common factors of $\log \alpha$, which make no difference to the correlation, we find that ρ is approximately equal to

$$\tilde{\rho} = \mathbf{Corr}(U, \tilde{V}), \quad (11)$$

where $U = -X(L-X)$ and $\tilde{V} = \max\{X, L-X\}$. We can calculate this correlation using Eq. (3) and the following simple estimates:

$$\begin{aligned} \mathbf{Cov}(U, \tilde{V}) &= \frac{1}{96}L^3 + O(L^2), \\ \mathbf{Var}(U) &= \frac{1}{180}L^4 + O(L^3), \quad \text{and} \\ \mathbf{Var}(\tilde{V}) &= \frac{1}{48}L^2 + O(L). \end{aligned}$$

The covariance, for example, can be calculated to leading order in L by treating X as a continuous uniform random

variable on $[0, L]$. The estimate follows from

$$\begin{aligned} \mathbf{E}(U\tilde{V}) &\approx -\frac{1}{L} \int_0^L x(L-x) \max\{x, L-x\} dx = -\frac{11}{96}L^3, \\ \mathbf{E}(U) &\approx -\frac{1}{L} \int_0^L x(L-x) dx = -\frac{1}{6}L^2, \quad \text{and} \\ \mathbf{E}(\tilde{V}) &\approx \frac{1}{L} \int_0^L \max\{x, L-x\} dx = \frac{3}{4}L. \end{aligned}$$

Substituting the covariance and variance estimates into Eq. (3) shows that

$$\tilde{\rho} = \frac{(L^3/96)(1 + O(L^{-1}))}{\sqrt{(L^4/180)(L^2/48)}} = \frac{\sqrt{15}}{4}(1 + O(L^{-1})).$$

In the $\alpha < 1$ case, we instead set $\tilde{V} = \min\{X, L-X\}$. The same argument shows that $\tilde{\rho} = -\frac{\sqrt{15}}{4}(1 + O(L^{-1}))$ because $\mathbf{Cov}(U, \tilde{V})$ changes sign, while $\mathbf{Var}(\tilde{V})$ is unchanged.

As we prove in Appendix B, $\tilde{\rho}$ is a close approximation to ρ when L is sufficiently large as a function of α . The error of this approximation, combined with the preceding estimates of $\tilde{\rho}$, establishes Theorem 1.

B. A threshold in the correlation of the parts

Recall that Eq. (7) explains how the correlation ρ arises in terms of $\hat{\rho}$, the correlation between the global and local parts of the effective potential, and r , the ratio of their standard deviations. In particular, Eq. (7) implies that ρ approximately equals $\hat{\rho}$ when $r \gg 1$ (i.e., the global part varies far more than the local part). As we now explain, for any $\alpha \neq 1$, the particle system has $r \gg 1$ for all sufficiently large L . The threshold in $\rho(\alpha)$ that Theorem 1 identifies therefore arises from an underlying threshold in the correlation between the parts of the effective potential.

The quantities $\hat{\rho}$ and r depend on the stationary distribution $\hat{\pi}$ of the discrete-time Markov jump process with transition rates \hat{W} (4). We could calculate $\hat{\pi}$ directly, as we calculated π , using the fact this process is also of birth-death type. However, it is simpler to combine the fact that $\hat{\pi}(x)$ is proportional to $\pi(x)q(x)$ (see Eq. (5)) with the approximation of q by \tilde{q} :

$$\hat{\pi}(x) \propto \pi(x)q(x) \approx \begin{cases} \alpha^{x(L-x) + \min\{x, L-x\}} & \alpha < 1, \\ \alpha^{x(L-x) + \max\{x, L-x\}} & \alpha > 1. \end{cases}$$

Using the approximation of q by \tilde{q} a second time, we find that r satisfies

$$r = \sqrt{\frac{\mathbf{Var}(\log \hat{\pi}(X))}{\mathbf{Var}(\log q(X))}} \approx \sqrt{\frac{\mathbf{Var}(U - \tilde{V})}{\mathbf{Var}(\tilde{V})}}.$$

As in the previous subsection, we denote $U = -X(L-X)$ and $\tilde{V} = \text{sign}(\alpha - 1) \max\{X, L-X\}$ for $\alpha \neq 1$. The

variance estimates for U and \tilde{V} then imply that

$$\begin{aligned}\sqrt{\frac{\text{Var}(U - \tilde{V})}{\text{Var}(\tilde{V})}} &= \sqrt{\frac{L^4/180}{L^2/48}} (1 + O(L^{-1})) \\ &= \frac{2L}{\sqrt{15}} (1 + O(L^{-1})).\end{aligned}$$

We find that, for any fixed $\alpha \neq 1$, $r \gg 1$ for all sufficiently large L . As a consequence of Eq. (7), ρ is primarily determined by the correlation $\hat{\rho}$ between the parts of the effective potential.

To establish that $\hat{\rho}$ exhibits a threshold in α , we use the preceding expressions for $\hat{\pi}$ and \tilde{q} :

$$\hat{\rho} = \text{Corr}(-\log \hat{\pi}(X), \log q(X)) \approx \text{Corr}(U - \tilde{V}, \tilde{V}).$$

Since the variance of U grows faster with L than the variance of \tilde{V} does, the correlation of $U - \tilde{V}$ with \tilde{V} is approximately that of U and \tilde{V} for all sufficiently large L . The proof of Theorem 1 then shows that

$$\text{Corr}(U - \tilde{V}, \tilde{V}) \approx \frac{\sqrt{15}}{4} \text{sign}(\alpha - 1) (1 + O(L^{-1})),$$

for all $\alpha \neq 1$. We conclude that the threshold in ρ arises because the local and global parts of the effective potential themselves undergo a transition in correlation as α varies.

IV. SPIN-GLASS DYNAMICS

For our second example, we analyze a significantly more complicated family of dynamics associated with the Sherrington–Kirkpatrick (SK) spin glass. The dynamics will be random, so the correlation ρ will be random as well. Like Theorem 1, the main result of this section identifies a threshold in ρ , as a parameter of the dynamics varies. Although the threshold in ρ is similar to that of the particle system, our subsequent analysis of $\hat{\rho}$ and r shows that the underlying relationship between the local and global parts of the effective potential is entirely different.

For an integer $N \geq 1$ and inverse temperature $\beta > 0$, the SK model is the random probability distribution π on the N -dimensional hypercube $S = \{-1, 1\}^N$ that assigns to a configuration $x \in S$ the Boltzmann probability $\pi(x) \propto \exp(-\beta H(x))$, in terms of a random Hamiltonian H . The Hamiltonian of x is defined as

$$H(x) = \sum_{i,j \in [N]} g_{ij} x_i x_j,$$

in terms of random couplings $(g_{ij})_{i,j \in [N]}$ of the coordinates, which are independent and identically distributed (i.i.d.) normal random variables with mean 0 and variance $1/N$.

We consider a family of Glauber dynamics for the SK model, like those commonly used in Monte Carlo simulations thereof [19]. The dynamics are defined in terms of a parameter $\lambda \in [0, 1]$ as

$$W_{x,y}^\lambda = e^{\beta(\lambda H(x) - (1-\lambda)H(y))}, \quad y \sim x. \quad (12)$$

Note that $y \sim x$ indicates that y neighbors x in the hypercube, meaning they differ in exactly one coordinate. This family interpolates between two extremes, in a sense. The first extreme entails transitions from x to y at a rate of $W_{x,y}^0 = e^{-\beta H(y)}$. In other words, jumps occur at a rate that depends on the energy of the destination. The second extreme entails rates that depend on the originating state: $W_{x,y}^1 = e^{\beta H(x)}$.

We are again interested in the relationship between the exit rates

$$q^\lambda(x) = \sum_{y: y \sim x} W_{x,y}^\lambda,$$

and the stationary distribution, which is the Boltzmann distribution π for every $\lambda \in [0, 1]$. To verify that π is the stationary distribution, note that it satisfies detailed balance for a generic λ :

$$\begin{aligned}\pi(x) W_{x,y}^\lambda &= e^{-\beta H(x) + \beta(\lambda H(x) - (1-\lambda)H(y))} / Z \\ &= e^{-\beta H(y) + \beta(\lambda H(y) - (1-\lambda)H(x))} / Z = \pi(y) W_{y,x}^\lambda.\end{aligned}$$

Since the dynamics depend on the random Hamiltonian H , the correlation defined by

$$\rho(\lambda) = \text{Corr}(-\log \pi(X), \log q^\lambda(X)),$$

is a random variable. The main result of this section is an estimate of ρ that holds at high temperatures, with a probability that tends to 1 as the dimension N grows.

Theorem 2. *Let $N \geq 3$ be an integer and set $\lambda_*(N) = \frac{1-4/N}{2-4/N}$. As $\beta \rightarrow 0$, the correlation satisfies*

$$\rho(\lambda) = \text{sign}(\lambda - \lambda_*) \left(1 + \frac{\sqrt{N}}{|\lambda - \lambda_*|} \cdot O(\beta) \right) \quad (13)$$

for every $\lambda \neq \lambda_*$, with a probability of at least $1 - O(N^{-2})$ as $N \rightarrow \infty$.

Informally, Theorem 2 means that, in high enough dimensions and at high enough temperatures, the correlation is approximately $\text{sign}(\lambda - \lambda_*)$ for every $\lambda \neq \lambda_*$, with a probability close to 1 over the randomness of the couplings. According to Eq. (13), the correlation abruptly increases from -1 to 1 as λ increases past λ_* (Fig. 5). Note that λ_* lies in $[0, 1]$ only if $N \geq 5$. In particular, Eq. (13) implies that the correlation is nonnegative regardless of the value of λ when the dimension N is at most 4 (Fig. 5c). As in Theorem 1, our estimate of the correlation features an error term that arises from an approximation of the exit rates by a function that is easier to analyze.

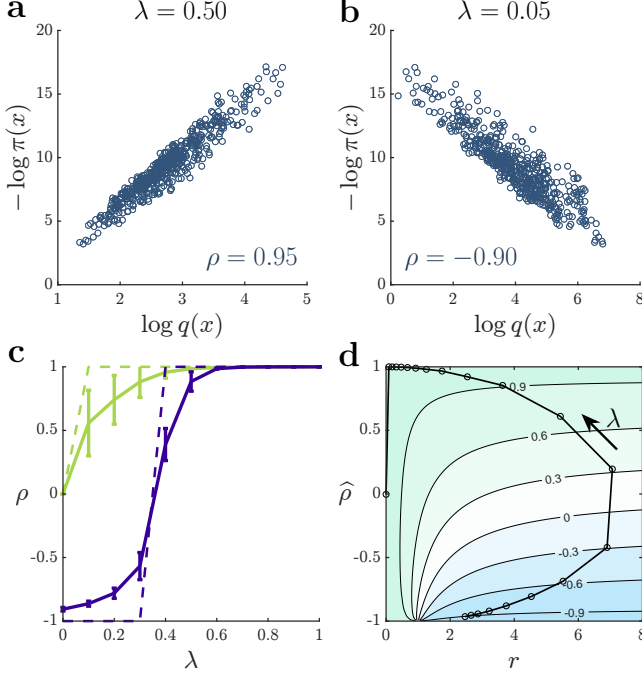


FIG. 5. Local-global correlations of SK model dynamics. The effective potential scattered against the log exit rate of each state for (a) $\lambda = 0.50$ and (b) $\lambda = 0.05$. (c) The correlation as a function of λ for $N = 2^4$ (green curve) and $N = 2^{10}$ (blue curve). The curves show the mean correlation over 50 and 10 independent trials, respectively, with corresponding error bars of ± 1 standard deviation. Dashed lines linearly interpolate the formula in Eq. (13). (d) Plot of $(r, \hat{\rho})$ pairs for $\lambda \in [0, 1]$ with $N = 2^{10}$ fixed. Each mark represents the average over 10 independent trials of $\hat{\rho}$ and r for a particular λ . Shading and contours indicate the corresponding value of ρ . The arrow indicates the direction of increasing λ . In all cases, $\beta = 1$.

A. Calculating the correlation at high temperature

For the moment, assume that the random Hamiltonian H is given, in which case the only randomness is in the uniformly random state $X \in S$. We begin by expressing the exit rates as

$$q^\lambda(x) = \sum_{k \in [N]} e^{\beta(\lambda H(x) - (1-\lambda)H(x^{(k)}))},$$

in terms of $x^{(k)}$, the element of S that agrees with x except in the k th coordinate:

$$x^{(k)} = (x_1, \dots, x_{k-1}, -x_k, x_{k+1}, \dots, x_N).$$

We then separate a factor of $e^{\beta(2\lambda-1)H(x)}$ from the rates, which results in

$$q^\lambda(x) = e^{\beta(2\lambda-1)H(x) + \beta A(x)}, \quad (14)$$

in terms of the quantity

$$A(x) = \frac{1}{\beta} \log \sum_{k \in [N]} e^{-\beta(1-\lambda)(H(x^{(k)}) - H(x))}.$$

By substituting Eq. (14) and the Boltzmann distribution $\pi(x) \propto e^{-\beta H(x)}$ into Eq. (2), we find that the correlation satisfies

$$\rho = \text{Corr}(H(X), (2\lambda - 1)H(X) + A(X)). \quad (15)$$

The correlation is therefore determined by the variances of $H(X)$ and $A(X)$, as well as their covariance.

While the variance of $H(X)$ can be calculated exactly, the quantities involving A are more complicated. We will resort to estimates of these quantities in the high-temperature setting, based on the following idea. A short calculation shows that, since the couplings are i.i.d. normal random variables with mean 0 and variance $1/N$, the $N2^N$ random energy differences of the form $H(x^{(k)}) - H(x)$ that define A will rarely exceed \sqrt{N} in absolute value. Hence, if β is sufficiently small in terms of N , then $A(x)$ will be close to

$$\tilde{A}(x) = \frac{1}{\beta} \log N - \frac{1-\lambda}{N} \sum_{k \in [N]} (H(x^{(k)}) - H(x)),$$

uniformly in x . We will show that the closeness of A and \tilde{A} implies that ρ is close to

$$\tilde{\rho} = \text{Corr}(H(X), (2\lambda - 1)H(X) + \tilde{A}(X)).$$

Accordingly, instead of calculating ρ directly, we will calculate $\tilde{\rho}$ and then bound the error incurred by doing so.

We implement this strategy in Appendix C, where we further show that the variance of \tilde{A} and its covariance with H are exactly

$$\begin{aligned} \text{Var}(\tilde{A}(X)) &= (4(1-\lambda)/N)^2 \text{Var}(H(X)) \quad \text{and} \\ \text{Cov}(H(X), \tilde{A}(X)) &= (4(1-\lambda)/N) \text{Var}(H(X)). \end{aligned}$$

Using these estimates, for $\lambda \neq \lambda_*$, we calculate $\tilde{\rho}$ to be

$$\begin{aligned}\tilde{\rho}(\lambda) &= \frac{(2\lambda - 1) \mathbf{Var} H(X) + \mathbf{Cov}(H(X), A(X))}{\sqrt{\mathbf{Var} H(X)((2\lambda - 1)^2 \mathbf{Var} H(X) + \mathbf{Var} A(X) + 2(2\lambda - 1) \mathbf{Cov}(H(X), A(X)))}} \\ &= \frac{(2\lambda - 1) + 4(1 - \lambda)/N}{\sqrt{(2\lambda - 1)^2 + 8(2\lambda - 1)(1 - \lambda)/N + (4(1 - \lambda)/N)^2}} = \text{sign}((2\lambda - 1) + 4(1 - \lambda)/N) = \text{sign}(\lambda - \lambda_*). \quad (16)\end{aligned}$$

This explains the formula in Theorem 2, aside from the error term, the control of which addresses the randomness of the couplings. We provide a full proof in Appendix C.

B. Interpreting the correlation threshold

As in Section III B, we turn to the relationship between the parts of the effective potential. By Eq. (5), $\hat{\pi}^\lambda(x)$ is proportional to $\pi(x)q^\lambda(x)$, hence

$$\hat{\pi}^\lambda(x) \propto e^{-\beta(2(1-\lambda)H(x) - A(x))}.$$

The correlation $\hat{\rho}$ therefore equals

$$\mathbf{Corr}(2(1 - \lambda)H(X) - A(X), (2\lambda - 1)H(X) + A(X)),$$

while the parameter r satisfies

$$r = \sqrt{\frac{\mathbf{Var}(2(1 - \lambda)H(X) - A(X))}{\mathbf{Var}((2\lambda - 1)H(X) + A(X))}}.$$

We could calculate $\hat{\rho}$ and r directly, like we did in Section III B. Instead, we note that Eq. (5) implies an analogue of Eq. (7) for $\hat{\rho}$:

$$\hat{\rho} = -\frac{1 - \rho s}{\sqrt{1 - 2\rho s + s^2}} \quad \text{where} \quad s = \sqrt{\frac{\mathbf{Var}(\log \pi(X))}{\mathbf{Var}(\log q(X))}}.$$

Note that $\hat{\rho} \approx \rho$ when $s \gg 1$. Approximating A by \tilde{A} and then using the covariance and variance estimates from the preceding subsection, we find that

$$\begin{aligned}s &= \sqrt{\frac{\mathbf{Var} H(X)}{\mathbf{Var}((2\lambda - 1)H(X) + A(X))}} \\ &\approx \frac{1}{|(2\lambda - 1) + 4(1 - \lambda)/N|} = \frac{1}{|\lambda - \lambda_*|}.\end{aligned}$$

Heuristically, when λ is sufficiently close to the threshold at λ_* , s is much larger than 1, in which case $\hat{\rho}$ is close to ρ . We conclude that, as in Theorem 1, the transition from negative to positive ρ is mediated primarily through a similar change in $\hat{\rho}$ (Fig. 5d).

V. DISCUSSION

Steady states that exhibit high local–global correlation satisfy an approximate analogue of the Boltzmann distribution (Fig. 1). In this case, a simple local property of

a state—the logarithm of the exit rate—plays the role of energy. The virtue of this approximation is that exit rates are often easy to estimate. Indeed, the exit rate of a state is the reciprocal of the average holding time in it. To estimate the exit rate, it therefore suffices to repeatedly initialize the system in the state and average the times it takes to leave. This means that it is unnecessary to observe long trajectories of the system to faraway states, which is generally required of estimating the steady-state distribution [4].

While a major goal of local–global correlation is to explain order in broad classes of nonequilibrium steady states, the concept applies equally well to equilibrium steady states. For this reason, and to facilitate our calculations of ρ , we analyzed reversible dynamics with stationary distributions that are explicitly known. For some equilibrium steady states, the Hamiltonian and the logarithm of exit rates coincide. For example, when $\lambda = 1$, the exit rates of the SK model dynamics (12) satisfy $q(x) = e^{\beta H(x)}$ and therefore $\rho = 1$. More generally, high local–global correlation can be valuable for understanding equilibrium steady states because exit rates may be readily estimable even when the Hamiltonian is complicated or unknown.

Clearly, many steady states cannot exhibit high local–global correlation. It is therefore important to understand precisely why and how typically ρ is close to 1. The general formula for ρ in Eq. (7) somewhat abstractly explains that there are two “ways” for ρ to be close to 1 (Section II B). These ways are defined in terms of key parameters $\hat{\rho}$ and r , which together characterize the relationship between the local and global parts of the effective potential (Fig. 2). Then, to understand how typically ρ is close to 1, one could estimate the typical values of $\hat{\rho}$ and r under a probability distribution on a class of steady states. However, these quantities are generally difficult to calculate, and our recent work instead resorts to bounding their expected values [17].

In this context, Theorems 1 and 2 serve two purposes. First, they provide explicit estimates of the local–global correlation in two families of dynamics, one deterministic and one random. The proofs of these estimates demonstrate an approach to calculating ρ that we expect will apply to many other dynamics. Second, they show that the property of a steady state exhibiting high local–global correlation can depend sensitively on parameters of the dynamics (Fig. 3b and Fig. 5c). In both cases, the thresholds in ρ arise from underlying thresholds in the correlation $\hat{\rho}$ between the parts of the effective potential. Such

a threshold can also arise when r , which measures the relative variance of the parts, crosses 1 (e.g., see Fig. 6 of [11]).

An important goal for future work is to develop means of estimating or bounding ρ in situations where the stationary distribution π is inexplicit, as is generally true of dynamics that violate detailed balance. For example, recent work in this direction compares the stationary distributions of nonreversible Markov chains with random transition rates to the probability distribution that is inversely proportional to the exit rates [20]. One possible strategy is to use the Markov chain tree theorem [21] to express and manipulate π , which has been integral to several recent results in stochastic thermodynamics [15, 22, 23].

ACKNOWLEDGMENTS

J.C. thanks Sid Redner for valuable input. This work was supported by U.S. Army Research Office award MURI W911NF-19-1-0233 and by the National Science Foundation award CCF-2106687.

Appendix A: Approximation of correlation

The correlations ρ in Theorems 1 and 2 are difficult to calculate exactly, due to the form of each model's exit rates. Instead, we calculate the correlations $\tilde{\rho}$ resulting from a close approximation of the exit rates, which is simpler to analyze. The next lemma states the intuitive fact that, so long as this approximation is sufficiently accurate, $\tilde{\rho}$ and ρ are close. In the statement and its proof, we write $1 \pm a$ to denote the interval $[1 - a, 1 + a]$ and $b = 1 \pm a$ to mean that b belongs to this interval.

Lemma A.1. *Let U , V , and \tilde{V} be random variables with finite, positive variances, that satisfy*

$$|\text{Cov}(U, \tilde{V})| \geq c\sqrt{\text{Var}(U)} \quad \text{and} \quad \text{Var}(\tilde{V}) \geq c^2,$$

for some $c > 0$. If $\text{Var}(V - \tilde{V}) < c^2/4$, then

$$\text{Corr}(U, V) = \text{Corr}(U, \tilde{V}) \left(1 \pm \frac{5}{c} \sqrt{\text{Var}(V - \tilde{V})} \right).$$

Proof. Let $\epsilon = V - \tilde{V}$ and $a = (\text{Var}(\epsilon)/\text{Var}(\tilde{V}))^{1/2}$. The Cauchy-Schwarz inequality implies that

$$\begin{aligned} \text{Cov}(U, V) &= \text{Cov}(U, \tilde{V}) + \text{Cov}(U, \epsilon) \\ &= \text{Cov}(U, \tilde{V})(1 \pm a). \end{aligned}$$

The variance of V analogously equals

$$\begin{aligned} \text{Var}(V) &= \text{Var}(\tilde{V}) + \text{Var}(\epsilon) + 2\text{Cov}(\tilde{V}, \epsilon) \\ &= \text{Var}(\tilde{V}) (1 + a^2 \pm 2a). \end{aligned}$$

Substituting these expressions into Eq. (3) shows that

$$\text{Corr}(U, V) = \text{Corr}(U, \tilde{V}) \left(\frac{1 \pm a}{1 + a^2 \pm 2a} \right).$$

The assumed bounds on $\text{Var}(\tilde{V})$ and $\text{Var}(\epsilon)$ imply that $a < 1/4$. As a consequence, the parenthetical expression is always positive, and some algebra further shows that the interval it represents is contained in $[1 - 3a, 1 + 5a]$. This completes the proof, since $a \leq \sqrt{\text{Var}(\epsilon)}/c$ by the assumed lower bound of $\text{Var}(\tilde{V})$. \square

Appendix B: Proof of Theorem 1

The proof outline in Section III A was incomplete because, instead of calculating ρ , we calculated

$$\tilde{\rho} = \text{Corr}(-\log \pi(X), \log \tilde{q}(X)),$$

where

$$\tilde{q}(x) = \begin{cases} \alpha^{\max\{x, L-x\}} & \alpha > 1, \\ \alpha^{\min\{x, L-x\}} & \alpha < 1. \end{cases}$$

We now use Lemma A.1 to obtain the estimate of ρ in Theorem 1 from our estimate of $\tilde{\rho}$.

Proof of Theorem 1. Recall that $\rho = \text{Corr}(U, V)$, where

$$U = \frac{-\log \pi(X)}{\log \alpha} \quad \text{and} \quad V = \frac{\log q(X)}{\log \alpha}. \quad (\text{B1})$$

We analogously define $\tilde{\rho} = \text{Corr}(U, \tilde{V})$, in terms of

$$\tilde{V} = \frac{\log \tilde{q}(X)}{\log \alpha},$$

and set $\epsilon = \tilde{V} - V$.

Eq. (11) and the covariance estimates that immediately follow it show that

$$\tilde{\rho} = \text{Corr}(U, \tilde{V}) = \frac{\sqrt{15}}{4} \text{sign}(\alpha - 1) (1 + O(L^{-1})).$$

The same estimates show that U and \tilde{V} satisfy the hypotheses of Lemma A.1 with $c = \delta L$, for a small positive number δ . To apply the conclusion of Lemma A.1, it therefore suffices to show that ϵ has variance of less than $\delta^2 L^2/4$. In fact, $\text{Var}(\epsilon)$ is much smaller as a function of L . We consider the cases of $\alpha > 1$ and $\alpha < 1$ separately.

By definition, when $\alpha > 1$,

$$\epsilon(x) = \begin{cases} 0 & x \in \{0, L\}, \\ \log(1 + \alpha^{\min\{x, L-x\} - \max\{x, L-x\}}) & x \notin \{0, L\}. \end{cases}$$

Since $\log(1 + x) \leq x$ for all $x > -1$, the second moment of ϵ satisfies

$$\mathbf{E}(\epsilon^2) \leq \frac{1}{L+1} \sum_{x=1}^{L-1} \alpha^{2(\min\{x, L-x\} - \max\{x, L-x\})}.$$

Simplifying the geometric sum, and using the fact that $\mathbf{Var}(\epsilon) \leq \mathbf{E}(\epsilon^2)$, we find that

$$\mathbf{Var}(\epsilon) \leq \frac{2\alpha^4}{L(\alpha^4 - 1)}.$$

This bound shows that, as long as L is sufficiently large as a function of α , then $\mathbf{Var}(\epsilon) < \delta^2 L^2/4$. Lemma A.1 then implies that

$$\rho = \tilde{\rho} \left(1 \pm \frac{5\sqrt{2}}{\delta L^{3/2}} \sqrt{\frac{\alpha^4}{\alpha^4 - 1}} \right).$$

When $\alpha < 1$, the only difference is that the minimum and maximum that appear in the preceding expression for ϵ are swapped. Analogous reasoning then shows that $\mathbf{Var}(\epsilon) \leq 2/(L(1 - \alpha^4))$. Hence, $\mathbf{Var}(\epsilon) < \delta^2 L^2/4$ for all sufficiently large L as a function of α , and so Lemma A.1 implies that the correlations are related by

$$\rho = \tilde{\rho} \left(1 \pm \frac{5\sqrt{2}}{\delta L^{3/2}} \frac{1}{\sqrt{1 - \alpha^4}} \right).$$

In both the $\alpha > 1$ and $\alpha < 1$ cases, by taking L larger in terms of α if necessary, we can ensure that

$$\rho = \tilde{\rho} (1 + O(L^{-1})).$$

Together with the estimate of $\tilde{\rho}$, this proves that

$$\rho = \frac{\sqrt{15}}{4} \text{sign}(\alpha - 1) (1 + O(L^{-1})).$$

□

Appendix C: Proof of Theorem 2

To complete the proof of Theorem 2 outlined in Section IV A, we need to bound the difference of ρ and $\tilde{\rho}$, and justify the expressions for the variance of $\tilde{A}(X)$ and its covariance with $H(X)$, which we used to calculate $\tilde{\rho}$. For the former task, we aim to apply Lemma A.1 with $\epsilon = A(X) - \tilde{A}(X)$, which requires a bound on $\mathbf{Var}(\epsilon)$. We emphasize that $\mathbf{Var}(\epsilon)$ is a random variable, because the variance is taken with respect to the uniformly random state X , not the randomness of the couplings $(g_{ij})_{i,j \in [N]}$ that define the energy.

We will bound above $\mathbf{Var}(\epsilon)$ on the event that energy differences of the form $D_k(x) = H(x^{(k)}) - H(x)$ are at most $O(\sqrt{N})$ in absolute value, across all states $x \in S$ and coordinates $k \in [N]$. Recall that

$$A(x) = \frac{1}{\beta} \log \sum_{k \in [N]} e^{-\beta(1-\lambda)D_k(x)} \quad \text{and} \\ \tilde{A}(x) = \frac{1}{\beta} \log N - \frac{1-\lambda}{N} \sum_{k \in [N]} D_k(x).$$

We will assume that β is small enough as a function of N so that, when this even occurs, the summands that appear in $A(x)$ approximately satisfy

$$e^{-\beta(1-\lambda)D_k(x)} \approx 1 - \beta(1-\lambda)D_k(x), \quad (\text{C1})$$

uniformly for $x \in S$. Approximating the summands of $A(x)$ in this way results in $\tilde{A}(x)$, hence ϵ is the error associated with Eq. (C1).

1. Control of energy differences

The next lemma uses the fact that the $D_k(x)$ are centered normal random variables to prove that all energy differences are $O(\sqrt{N})$ in absolute value, with probability 1 as $N \rightarrow \infty$. Throughout this section, as in the statement of Theorem 2, we assume that $N \geq 2$ is an integer.

Lemma C.1. *The largest absolute energy difference satisfies*

$$\max_{x \in S, k \in [N]} |D_k(x)| \leq 4\sqrt{N},$$

with a probability of at least $1 - N^{-2}$.

Proof. By definition, $H(x) = \sum_{i,j \in [N]} g_{ij} x_i x_j$. The energy difference $D_k(x)$ therefore satisfies

$$D_k(x) = \sum_{i,j \in [N]} g_{ij} x_i^{(k)} x_j^{(k)} - \sum_{l,m \in [N]} g_{lm} x_l x_m \\ = \sum_{j \in [N] \setminus \{k\}} -2(g_{jk} + g_{kj}) x_j x_k. \quad (\text{C2})$$

The couplings g_{jk} are i.i.d. $\mathcal{N}(0, 1/N)$ random variables, so the preceding expression shows that $D_k(x)$ is a sum of $N - 1$ i.i.d. $\mathcal{N}(0, 4/N)$ random variables. Hence, if $\sigma^2 = 4(1 - \frac{1}{N})$, then $D_k(x)/\sigma$ is a standard normal random variable, which satisfies the tail bound

$$\mathbf{P}(|D_k(x)/\sigma| > t) \leq 2e^{-t^2/2}, \quad t \geq 0. \quad (\text{C3})$$

Anticipating a sum of the tail bound over $N2^N$ possible choices of state x and coordinate k , we set

$$t = \sqrt{(2 \log 2)N + 8 \log N},$$

and define $\mathcal{E}_N(x, k)$ to be the event that $D_k(x)/\sigma$ exceeds t in absolute value:

$$\mathcal{E}_N(x, k) = \{|D_k(x)/\sigma| > t\}.$$

We use \mathcal{E}_N to denote the event that some $\mathcal{E}_N(x, k)$ occurs, and bound the probability that it occurs as

$$\mathbf{P}(\mathcal{E}_N) \leq \sum_{x \in S, k \in [N]} \mathbf{P}(\mathcal{E}_N(x, k)) \\ \leq N2^N \cdot \mathbf{P}(|D_1(\mathbf{1})/\sigma| > t) \\ \leq N2^N \cdot 2e^{-(\log 2)N - 4 \log N} = 2N \cdot N^{-4} \leq N^{-2}.$$

The first inequality is a union bound over x and k . The second bound holds by the definition of the event $\mathcal{E}_N(x, k)$ and the fact that the k th energy difference of state x has the same distribution as the first energy difference of state $\mathbf{1} = (1, \dots, 1) \in \{-1, 1\}^N$. The third inequality is due to the tail bound in Eq. (C3). The last inequality holds by the assumption that $N \geq 2$. Hence, except with a probability of $1 - N^{-2}$,

$$\max_{x \in S, k \in [N]} |D_k(x)| \leq \sigma t \leq 4\sqrt{N}.$$

□

2. Approximation of A by \tilde{A}

Following the discussion at the beginning of Appendix C, we proceed to bound the variance of $A(X) - \tilde{A}(X)$. We do so using the approximation in Eq. (C1), which is justified for small β due to Lemma C.1. We continue to assume that $N \geq 2$ is an integer.

Lemma C.2. *As $\beta \rightarrow 0$, the quantity $\epsilon = A(X) - \tilde{A}(X)$ satisfies*

$$\frac{1}{N} \sqrt{\mathbf{Var}(\epsilon)} = O(\beta),$$

with a probability of at least $1 - N^{-2}$.

Proof. Suppose that the largest energy difference $D_k(x)$ is at most $4\sqrt{N}$ in absolute value. According to Lemma C.1, this event occurs with a probability of at least $1 - N^{-2}$. When it occurs, we can take $\beta \rightarrow 0$ and expand A as

$$\begin{aligned} A(x) &= \frac{1}{\beta} \log \sum_{k \in [N]} (1 - \beta(1 - \lambda)D_k(x) + N \cdot O(\beta^2)) \\ &= \frac{1}{\beta} \log N - \frac{1 - \lambda}{N} \sum_{k \in [N]} D_k(x) + N \cdot O(\beta). \end{aligned}$$

Hence, $\epsilon = N \cdot O(\beta)$ and therefore

$$\frac{1}{N} \sqrt{\mathbf{Var}(\epsilon)} \leq \frac{1}{N} \sqrt{\mathbf{E}(\epsilon^2)} = O(\beta).$$

□

3. Variance and covariance calculations

The next inputs to the proof of Theorem 2 are the following expressions for the variances and covariance of $H(X)$ and $\tilde{A}(X)$. We will use these expressions both to calculate $\tilde{\rho}$ and to show that $\tilde{\rho}$ is close to ρ .

Lemma C.3. *Given couplings $(g_{ij})_{i,j \in [N]}$, the variances and covariance of $H(X)$ and $\tilde{A}(X)$ satisfy*

$$\begin{aligned} \mathbf{Var}(H(X)) &= \sum_{1 \leq i < j \leq N} (g_{ij} + g_{ji})^2, \\ \mathbf{Var}(\tilde{A}(X)) &= b^2 \mathbf{Var}(H(X)), \quad \text{and} \\ \mathbf{Cov}(H(X), \tilde{A}(X)) &= b \mathbf{Var}(H(X)), \end{aligned}$$

where $b = 4(1 - \lambda)/N$.

We note that, since H is a pseudo-Boolean function, i.e., a real-valued function on the hypercube $\{-1, 1\}^N$, it is possible to calculate $\mathbf{Var} H(X)$ using Parseval's theorem [24]. Instead, we calculate the moments directly, both for the sake of exposition and to facilitate the subsequent calculation of $\mathbf{Var}(\tilde{A}(X))$.

Proof of Lemma C.3. We start with the variance of $H(X) = \sum_{i,j \in [N]} g_{ij} X_i X_j$. Its first moment equals

$$\mathbf{E}(H(X)) = \sum_{i,j \in [N]} g_{ij} \mathbf{E}(X_i X_j) = \sum_{i \in [N]} g_{ii},$$

because $\mathbf{E}(X_i X_j)$ equals 1 when $i = j$, and 0 otherwise. The second moment satisfies

$$\mathbf{E}(H(X)^2) = \sum_{i,j,k,l \in [N]} g_{ij} g_{kl} \mathbf{E}(X_i X_j X_k X_l).$$

The quantity $\mathbf{E}(X_i X_j X_k X_l)$ equals 0 unless $\{i, j, k, l\}$ contains one or two distinct coordinates, in which case it equals 1. By considering the various cases, we find that

$$\mathbf{E}(H(X)^2) = \sum_{i \in [N]} g_{ii}^2 + \sum_{i \neq j} (g_{ii} g_{jj} + g_{ij}^2 + g_{ij} g_{ji}).$$

Hence, the variance satisfies

$$\mathbf{Var}(H(X)) = \sum_{i \neq j} (g_{ij}^2 + g_{ij} g_{ji}) = \sum_{1 \leq i < j \leq N} (g_{ij} + g_{ji})^2.$$

Next, we calculate $\mathbf{Var}(\tilde{A}(X))$, as follows:

$$\begin{aligned} \mathbf{Var}(\tilde{A}(X)) &= \mathbf{Var} \left(\frac{1}{\beta} \log N - \frac{1 - \lambda}{N} \sum_{k \in [N]} D_k(X) \right) \\ &= \frac{(1 - \lambda)^2}{N^2} \mathbf{Var} \left(\sum_{k \in [N]} D_k(X) \right) \\ &= \frac{4(1 - \lambda)^2}{N^2} \mathbf{Var} \left(\sum_{i \neq j} (g_{ij} + g_{ji}) X_i X_j \right) \\ &= \frac{16(1 - \lambda)^2}{N^2} \mathbf{Var} \left(\sum_{i \neq j} g_{ij} X_i X_j \right) \\ &= b^2 \mathbf{Var}(H(X)). \end{aligned}$$

The first equality is due to the definition of $\tilde{A}(X)$. The second holds because the constant $(\log N)/\beta$ does not affect the variance. The third follows from the formula for $D_k(x)$ in Eq. (C2). The fourth holds because the sums of $g_{ij}X_iX_j$ and $g_{ji}X_iX_j$ are equal. The final equality is due to the fact that the quantities $\sum_{i \neq j} g_{ij}X_iX_j$ and $H(X)$ differ only by the constant $\sum_{i \in [N]} g_{ii}$, which does not affect the variance.

By simply repeating these steps for the covariance, we find that

$$\mathbf{Cov}(H(X), \tilde{A}(X)) = b \mathbf{Var}(H(X)),$$

which concludes the proof. \square

Note that, as a consequence of Lemma C.3, we know the distribution of $\mathbf{Var}(H(X))$ exactly.

Lemma C.4. *Let $N \geq 2$ be an integer. The distribution of $\mathbf{Var}(H(X))$ is $\text{Gamma}(\frac{N(N-1)}{4}, \frac{4}{N})$.*

Proof. Since the couplings are i.i.d. $\mathcal{N}(0, \frac{1}{N})$ random variables, the quantity $(g_{ij} + g_{ji})^2$ has a $\text{Gamma}(\frac{1}{2}, \frac{4}{N})$ distribution. By Lemma C.3, $\mathbf{Var}(H(X))$ is a sum of $\frac{N(N-1)}{2}$ independent such random variables, hence it has a $\text{Gamma}(\frac{N(N-1)}{4}, \frac{4}{N})$ distribution. \square

A consequence of Lemma C.4 is that $\mathbf{Var}(H(X))$ has a mean of $N-1$ and a variance of $4(1 - \frac{1}{N})$ with respect to the randomness of the couplings, so it rarely takes values less than $(N-1)/2$. Specifically, the standard Chernoff bound for gamma random variables implies that

$$\mathbf{P}\left(\mathbf{Var}(H(X)) < \frac{N-1}{2}\right) \leq e^{-cN(N-1)}, \quad (\text{C4})$$

where $c = (\log 2 - 1/2)/4 \approx 0.048$.

4. Conclusion

We now combine the results from the two preceding subsections to prove Theorem 2.

Proof of Theorem 2. For the moment, assume that the couplings are given. We write ρ as $\mathbf{Corr}(U, V)$, where

$$U = H(X) \quad \text{and} \quad V = (2\lambda - 1)H(X) + A(X).$$

Analogously, we write $\tilde{\rho} = \mathbf{Corr}(U, \tilde{V})$, in terms of

$$\tilde{V} = (2\lambda - 1)H(X) + \tilde{A}(X).$$

By Lemma C.3, the corresponding covariance satisfies

$$\begin{aligned} \mathbf{Cov}(U, \tilde{V}) &= (2\lambda - 1) \mathbf{Var}(H(X)) + \mathbf{Cov}(H(X), \tilde{A}(X)) \\ &= ((2\lambda - 1) + b) \mathbf{Var}(H(X)), \end{aligned}$$

in terms of $b = 4(1 - \lambda)/N$. Similarly, the variance of \tilde{V} satisfies

$$\begin{aligned} \mathbf{Var}(\tilde{V}) &= (2\lambda - 1)^2 \mathbf{Var}(H(X)) \\ &\quad + 2(2\lambda - 1) \mathbf{Cov}(H(X), \tilde{A}(X)) + \mathbf{Var}(\tilde{A}(X)), \end{aligned}$$

so Lemma C.3 implies that

$$\begin{aligned} \mathbf{Var}(\tilde{V}) &= ((2\lambda - 1)^2 + 2(2\lambda - 1)b + b^2) \mathbf{Var}(H(X)) \\ &= ((2\lambda - 1) + b)^2 \mathbf{Var}(H(X)). \end{aligned}$$

These expressions for $\mathbf{Cov}(U, \tilde{V})$ and $\mathbf{Var}(\tilde{V})$ show that

$$|\mathbf{Cov}(U, \tilde{V})| \geq c\sqrt{\mathbf{Var}(U)} \quad \text{and} \quad \mathbf{Var}(\tilde{V}) \geq c^2, \quad (\text{C5})$$

in terms of $c = |2\lambda - 1 + b|\sqrt{\mathbf{Var}(H(X))}$. Note that

$$2\lambda - 1 + b = 2\left(1 - \frac{2}{N}\right)(\lambda - \lambda_*),$$

so c depends on the absolute distance of λ from the location of the threshold λ_* in Theorem 2.

Next, to verify the hypotheses of Lemma A.1, we address the randomness of the couplings. Lemma C.2 and Eq. (C4) together imply that the event consisting of

$$\begin{aligned} \frac{1}{N}\sqrt{\mathbf{Var}(\epsilon)} &= O(\beta) \text{ as } \beta \rightarrow 0 \quad \text{and} \\ \mathbf{Var}(H(X)) &\geq \frac{N-1}{2} \end{aligned}$$

occurs with a probability of at least $1 - O(N^{-2})$ as $N \rightarrow \infty$. When it does, c is positive for all $\lambda \neq \lambda_*$ because

$$c \geq 2\left(1 - \frac{2}{N}\right)|\lambda - \lambda_*|\sqrt{\frac{N-1}{2}} \geq \frac{2}{3\sqrt{3}}|\lambda - \lambda_*|\sqrt{N}.$$

Note that the second inequality follows from the simple bounds $(1 - 2/N) \geq 1/3$ and $\sqrt{(N-1)/2} \geq \sqrt{N/3}$, which hold for all $N \geq 3$. By Eq. (C5), U and \tilde{V} satisfy the hypotheses of Lemma A.1 for this choice of c , for all $\lambda \neq \lambda_*$. Lemma A.1 then implies that, as $\beta \rightarrow 0$,

$$\rho = \tilde{\rho} \left(1 \pm \frac{5}{c}\sqrt{\mathbf{Var}(\epsilon)}\right) = \tilde{\rho} \left(1 + \frac{\sqrt{N}}{|\lambda - \lambda_*|} \cdot O(\beta)\right),$$

for all such λ . This completes the proof, since $\tilde{\rho} = \text{sign}(\lambda - \lambda_*)$ for all $\lambda \neq \lambda_*$ by Eq. (16). \square

-
- [1] K. A. Dill, *Biochemistry* **29**, 7133 (1990).
 - [2] P. Chvykov, T. A. Berrueta, A. Vardhan, W. Savoie, A. Samland, T. D. Murphey, K. Wiesenfeld, D. I. Goldman, and J. L. England, *Science* **371**, 90 (2021).
 - [3] F. Noé, S. Olsson, J. Köhler, and H. Wu, *Science* **365**, eaaw1147 (2019).
 - [4] C. E. Lee, A. Ozdaglar, and D. Shah, in *Advances in Neural Information Processing Systems*, Vol. 26 (Curran Associates, Inc., 2013).
 - [5] R. Landauer, *Physical Review A* **12**, 636 (1975).
 - [6] R. Landauer, *Journal of Statistical Physics* **53**, 233 (1988).
 - [7] P. Chvykov and J. England, *Physical Review E* **97**, 032115 (2018).
 - [8] J. Gold, *Self-organized fine-tuned response in a driven spin glass*, Ph.D. thesis, Massachusetts Institute of Technology (2021).
 - [9] J. F. Yang, T. A. Berrueta, A. M. Brooks, A. T. Liu, G. Zhang, D. Gonzalez-Medrano, S. Yang, V. B. Koman, P. Chvykov, L. N. LeMar, M. Z. Miskin, T. D. Murphey, and M. S. Strano, *Nature Communications* **13**, 5734 (2022).
 - [10] H. Kedia, D. Pan, J.-J. Slotine, and J. L. England, *The Journal of Chemical Physics* **159**, 214106 (2023).
 - [11] J. Calvert and D. Randall, *Proceedings of the National Academy of Sciences* **121**, e2411731121 (2024).
 - [12] J. Schnakenberg, *Reviews of Modern Physics* **48**, 571 (1976).
 - [13] R. K. P. Zia and B. Schmittmann, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, P07012 (2007).
 - [14] A. C. Barato and U. Seifert, *Physical Review Letters* **114**, 158101 (2015).
 - [15] J. A. Owen, T. R. Gingrich, and J. M. Horowitz, *Physical Review X* **10**, 011066 (2020).
 - [16] G. Falasco and M. Esposito, *Reviews of Modern Physics* **97**, 015002 (2025).
 - [17] J. Calvert and D. Randall, *Local–global correlations of dynamics on disordered energy landscapes* (2025), arXiv:2508.04501 [math.PR].
 - [18] J. R. Norris, *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, 1997).
 - [19] P. Mathieu, *Communications in Mathematical Physics* **215**, 57 (2000).
 - [20] J. Calvert, F. den Hollander, and D. Randall, *A note on the asymptotic uniformity of Markov chains with random rates* (2025), arXiv:2505.01608 [math.PR].
 - [21] F. Leighton and R. Rivest, *IEEE Transactions on Information Theory* **32**, 733 (1986).
 - [22] S. Dal Cengio, V. Lecomte, and M. Polettini, *Physical Review X* **13**, 021040 (2023).
 - [23] C. Floyd, A. R. Dinner, A. Murugan, and S. Vaikuntanathan, *Nature Communications* **16**, 7184 (2025).
 - [24] R. O'Donnell, *Analysis of Boolean Functions* (Cambridge University Press, Cambridge, 2014).