Supporting Material: Scaling of Energy Dissipation in Nonequilibrium Reaction Networks

Qiwei Yu,[1] Dongliang Zhang,[1] and Yuhai Tu[2]

[1] *School of Physics, Peking University, Beijing 100871, China*
[2] *IBM T. J. Watson Research Center, Yorktown Heights, NY 10598*

CONTENTS

I. SUPPLEMENTARY ANALYTIC DERIVATION

A. Derivation of the scaling exponent λ

To derive the expression for the scaling exponent λ in the main text, we will first study the square lattice, where the derivation is more intuitive, and then generalize it to generic reaction networks.

1. The square lattice

The steady-state dissipation rate of the system is given by

$$\dot{W} = \sum_{i<j} (J_{i,j} - J_{j,i}) \ln \frac{J_{i,j}}{J_{j,i}} = \sum_{i<j} \left(k_{i,j} P_i^{ss} - k_{j,i} P_j^{ss}\right) \ln \frac{k_{i,j} P_i^{ss}}{k_{j,i} P_j^{ss}}, \tag{S1}$$

where P_i^{ss} is the stead-state probability of state i and $J_{i,j}$ is the steady-state flux from state i to state j [1]. The transition fluxes can be decomposed into symmetric and antisymmetric components:

$$J_{i,j} = \frac{1}{2}(S_{i,j} + A_{i,j}), \tag{S2}$$

where $S_{i,j} = J_{i,j} + J_{j,i}$ and $A_{i,j} = J_{i,j} - J_{j,i}$. The antisymmetric component $A_{i,j}$ is actually the net current from state i to j. The dissipation rate as function of A and S reads

$$\dot{W} = \sum_{i<j} A_{i,j} \ln \frac{S_{i,j} + A_{i,j}}{S_{i,j} - A_{i,j}} \tag{S3}$$

At the microscopic scale, we take the continuum limit that the net flux is an infinitesimal flux compared to the symmetric flux, i.e. $|A_{i,j}| \ll |S_{i,j}|$. This leads to

$$\dot{W} = \sum_{i<j} A_{i,j} \ln \frac{S_{i,j} + A_{i,j}}{S_{i,j} - A_{i,j}} = \sum_{i<j} A_{i,j} \ln \frac{1 + \frac{A_{i,j}}{S_{i,j}}}{1 - \frac{A_{i,j}}{S_{i,j}}} \approx 2 \sum_{i<j} \frac{A_{i,j}^2}{S_{i,j}} = 2L \langle \frac{A_{i,j}^2}{S_{i,j}} \rangle, \tag{S4}$$

where L is the number of links and $\langle \cdot \rangle$ denotes averaging over all links. Numeric justifications for this approximation will be provided in section II.A of the SI, which demonstrates that asymptotically $|A|$ is a lower-order term compared to $|S|$.

As demonstrated in Fig. 1A of the main text, renormalization in the square lattice involves merging two adjacent fluxes:

$$J_{i,j} = J_{i_2,j_1} + J_{i_4,j_3}. \tag{S5}$$

Hence, their symmetric and antisymmetric fluxes are merged accordingly:

$$S_{i,j} = S_{i_2,j_1} + S_{i_4,j_3}, \;\; A_{i,j} = A_{i_2,j_1} + A_{i_4,j_3}. \tag{S6}$$

This leads to

$$\begin{aligned} &\langle S_{i,j} \rangle = \langle S_{i_2,j_1} \rangle + \langle S_{i_4,j_3} \rangle = 2 \langle S_{i_\alpha,j_\beta} \rangle, \\ &\langle A_{i,j}^2 \rangle = \langle A_{i_2,j_1}^2 \rangle + \langle A_{i_4,j_3}^2 \rangle + 2 \langle A_{i_2,j_1} A_{i_4,j_3} \rangle = 2 \langle A_{i_\alpha,j_\beta}^2 \rangle (1 + C) \end{aligned} \tag{S7}$$

where $C = \mathrm{corr}(A_{i_2,j_1}, A_{i_4,j_3})$ is the Pearson correlation coefficient between the adjacent fluxes, explicitly given by

$$C = \frac{\langle A_{i_2,j_1} A_{i_4,j_3} \rangle}{\sqrt{\langle A_{i_2,j_1}^2 \rangle \langle A_{i_4,j_3}^2 \rangle}} = \frac{\langle (J_{i_2,j_1} - J_{j_1,i_2})(J_{i_4,j_3} - J_{j_3,i_4}) \rangle}{\sqrt{\langle (J_{i_2,j_1} - J_{j_1,i_2})^2 \rangle \langle (J_{i_4,j_3} - J_{j_3,i_4})^2 \rangle}}, \tag{S8}$$

Quantities with subscripts i_α and j_β are averaged over microscopic links while those with subscripts i and j are averaged over macroscopic (coarse-grained) links. The link-averaged term $\langle A^2/S \rangle$ in Eq. S4 is thus given by

$$\frac{\left\langle \frac{A_{i,j}^2}{S_{i,j}} \right\rangle}{\left\langle \frac{A_{i_\alpha,j_\beta}^2}{S_{i_\alpha,j_\beta}} \right\rangle} \approx \frac{\langle A_{i,j}^2 \rangle}{\langle A_{i_\alpha,j_\beta}^2 \rangle} \frac{\langle S_{i_\alpha,j_\beta} \rangle}{\langle S_{i,j} \rangle} = 1 + C. \tag{S9}$$

Note that taking the ratio out of the averaging $\langle\cdot\rangle$ is justified by assuming that the correlation between A^2 and S^{-1} does not change in different coarse-grained layers. Even if the correlation between A^2 and S^{-1} does change after renormalization, its cumulative contribution to the dissipation is bounded, so the average effect per round goes to zero in the limit of infinite rounds of renormalization (which leads to the RG fixed point). We will provide further numeric justifications for this approximation in the later sections of the SI. The ratio of dissipation rate in the coarse-grained system to the original one is

$$\frac{\dot{W}_1}{\dot{W}_0} = \frac{L_1}{L_0}(1 + C_0). \tag{S10}$$

Here we add subscript 0 to C to indicate that it is the correlation at the finest-grained level. The dissipation rate at the s-th coarse-grained level is given by

$$\frac{\dot{W}_s}{\dot{W}_0} = \frac{L_s}{L_0}\prod_{i=0}^{s-1}(1 + C_i) \to \frac{L_s}{L_0}(1 + C^\star)^s, \tag{S11}$$

where $C^\star$ is the RG fixed point of C:

$$C^\star = \lim_{s\to\infty} C_s \tag{S12}$$

The block size and the number of links in the system after s rounds of coarse-graining is given by:

$$\frac{n_0}{n_s} = 4^s, \;\; \frac{L_s}{L_0} = \frac{1}{4^s} \tag{S13}$$

Therefore, the dissipation scaling exponent is

$$\lambda_{2\mathrm{d}} = -\lim_{s\to\infty} \frac{\ln \frac{\dot{W}_s}{\dot{W}_0}}{\ln \frac{n_0}{n_s}} = \lim_{s\to\infty} \frac{s\ln 4 - \sum_{i=1}^{s}\ln(1 + C_i)}{s\ln 4} = 1 - \log_4(1 + C^\star), \tag{S14}$$

thus deriving Eq. 9 of the main text.

2. *The general expression*

In cubic lattice or even more general reaction networks, coarse-graining would involve the merging of more than two fluxes. Suppose the renormalized flux $J_{i,j}$ comprises t microscopic fluxes denoted by J_{i_1,j_1}, J_{i_2,j_2}, …, J_{i_t,j_t}:

$$J_{i,j} = \sum_{\alpha=1}^{t} J_{i_\alpha,j_\alpha}. \tag{S15}$$

Their symmetric and antisymmetric fluxes are merged accordingly:

$$S_{i,j} = \sum_{\alpha=1}^{t} S_{i_\alpha,j_\alpha}, \;\; A_{i,j} = \sum_{\alpha=1}^{t} A_{i_\alpha,j_\alpha}. \tag{S16}$$

We assume that all microscopic fluxes J_{i_α,j_α} follow the same statistics. The summation over α can thus be rewritten as the averaging over the distribution of microscopic fluxes $\langle\cdot\rangle$. This leads to

$$\begin{aligned}
\langle S_{i,j}\rangle &= \left\langle \sum_{\alpha=1}^{t} S_{i_\alpha,j_\alpha} \right\rangle = t\langle S_{i_\alpha,j_\alpha}\rangle, \\
\langle A_{i,j}^2\rangle &= \left\langle \sum_{\alpha=1}^{t} A_{i_\alpha,j_\alpha}^2 + 2\sum_{\alpha=1}^{t}\sum_{\beta=\alpha+1}^{t} A_{i_\alpha,j_\alpha}A_{i_\beta,j_\beta} \right\rangle = t\langle A_{i_\alpha,j_\alpha}^2\rangle(1 + C)
\end{aligned} \tag{S17}$$

where $C = \mathrm{corr}(A_{i_\alpha,j_\alpha}, A_{i,j} - A_{i_\alpha,j_\alpha})$ is the Pearson correlation coefficient between the microscopic net flux A_{i_α,j_α} and the sum of all the microscopic net fluxes that will be merged with it to form the macroscopic flux $A_{i,j}$. Following

the same argument used in the square lattice, the link-averaged term $\langle A^2/S\rangle$ is thus multiplied by $(1+C)$ after coarse-graining. The dissipation rate at the s-th level is given by

$$\frac{\dot{W}_s}{\dot{W}_0} = \frac{L_s}{L_0}\prod_{i=1}^{s}(1+C_i) \to \frac{L_s}{L_0}(1+C^\star)^s, \tag{S18}$$

which has a form identical to the square lattice but involves a more general definition of the correlation $C^\star$. The block size and the number of links in the system after s rounds of coarse-graining are given by:

$$\frac{n_0}{n_s} = r^s, \;\; \frac{L_s}{L_0} = \left(\frac{n_0}{n_s}\right)^{-d_L} = r^{-sd_L} \tag{S19}$$

where d_L is the link exponent. Therefore, the dissipation scaling exponent is given by

$$\lambda = -\lim_{s\to\infty}\frac{\ln\frac{\dot{W}_s}{\dot{W}_0}}{\ln\frac{n_0}{n_s}} = \lim_{s\to\infty}\frac{sd_L\ln r - \sum_{i=1}^{s}\ln(1+C_i)}{s\ln r} = d_L - \log_r(1+C^\star), \tag{S20}$$

recovering Eq. 6 of the main text. The square lattice is a special case with $r=4$ and $d_L=1$.

B. Derivation of the link exponent d_L in the 2D-embedded scale-free network

In this part, we estimate the link exponent in the 2D-embedded scale-free network (Eq. 11 of the main text) by calculating the number of links in the coarse-grained network. The degree distribution of the initial network is

$$p(k) = Ak^{-\alpha}, \;\; k\in[k_{\min}, k_{\max}], \tag{S21}$$

where $k_{\min}$ is the minimum degree and $k_{\max}$ is a sufficiently large cutoff. The normalization constant A is given by

$$A = \left(\sum_{k=k_{\min}}^{k_{\max}} k^{-\alpha}\right)^{-1} \approx \left(\int_{k=k_{\min}}^{+\infty} k^{-\alpha}\,\mathrm{d}k\right)^{-1} = (\alpha-1)k_{\min}^{\alpha-1}. \tag{S22}$$

The embedding in 2D allows us to use the same coarse-graining method employed in the square lattice. Suppose that we coarse-grain states i_1, i_2, i_3, and i_4 to form a new state i. The self-similarity property of the scale-free network leads to the following degree relation

$$k_i = l_B^{-d_k}\max_{\alpha=1,2,3,4} k_{i_\alpha}, \tag{S23}$$

where $l_B = 2$ is the linear size of the coarse-grained state and d_k is the degree-scaling fractal dimension of the network [2]. Since k_{i_α} follows a power-law distribution with minimum degree $k_{\min}$ and exponent α, its cumulative degree distribution is

$$P(k_{i_\alpha}\geq k_0) = A\sum_{k=k_0}^{k_{\max}} k^{-\alpha} \approx A\int_{k=k_0}^{+\infty} k^{-\alpha}\,\mathrm{d}k = \left(\frac{k_{\min}}{k_0}\right)^{\alpha-1}. \tag{S24}$$

The cumulative degree distribution of k_i is therefore:

$$\begin{aligned} P(k_i\geq k_0) &= P\left(\max_{\alpha=1,2,3,4} k_{i_\alpha} \geq 2^{d_k}k_0\right) \\ &= 1 - \prod_{\alpha=1}^{4} P\left(k_{i_\alpha} < 2^{d_k}k_0\right) \\ &= 1 - \prod_{\alpha=1}^{4}\left[1 - \left(\frac{k_{\min}}{2^{d_k}k_0}\right)^{\alpha-1}\right] \\ &\approx 4\left(\frac{k_{\min}}{2^{d_k}k_0}\right)^{\alpha-1} = \left(\frac{k'_{\min}}{k_0}\right)^{\alpha-1}. \end{aligned} \tag{S25}$$

Hence, the coarse-grained network is also scale-free, with the same exponent α but a different minimum degree

$$k'_{\min} = k_{\min} 2^{\frac{2}{\alpha-1}-d_k}. \tag{S26}$$

The average degree in the scale-free network is

$$\langle k\rangle = \sum_{k=k_{\min}}^{k_{\max}} p(k)k \approx A\int_{k=k_{\min}}^{+\infty} k^{1-\alpha}\,\mathrm{d}k = \frac{\alpha-1}{\alpha-2}k_{\min} \propto k_{\min}. \tag{S27}$$

Therefore, the ratio of the number of links in the coarse-grained network to the original one is

$$\frac{L'}{L} = \frac{n'\langle k'\rangle}{n\langle k\rangle} = \frac{n'}{n}\frac{k'_{\min}}{k_{\min}} = 2^{\frac{2}{\alpha-1}-d_k-2}, \tag{S28}$$

which leads to the link exponent

$$d_L = -\frac{\ln\frac{L'}{L}}{\ln\frac{n}{n'}} = 1 - \frac{1}{\alpha-1} + \frac{d_k}{2}. \tag{S29}$$

To link d_L back to the structural parameters of the original network (the finest-grained level), we utilize the following relation to replace the index d_k with the box-covering dimension d_B:

$$\alpha = 1 + \frac{d_B}{d_k}. \tag{S30}$$

d_B can be numerically determined in the original network with a well-established box-covering algorithm [2–4]. The final expression of the link exponent is:

$$d_L = 1 - \frac{1}{\alpha-1} + \frac{d_B}{2(\alpha-1)} = 1 + \frac{d_B-2}{2(\alpha-1)}, \tag{S31}$$

recovering Eq. 11 of the main text.

II. SUPPLEMENTARY SIMULATION RESULTS

This section supplies additional numeric results that further support discussion in the main text and rationalize the approximations used in the analytical derivation.

A. The square lattice

1. *Finite size and boundary effects*

The scaling behavior analyzed above assumes an infinitely large system which can approach a RG fixed point after infinite rounds of coarse-graining. In reality, nevertheless, we can only start with a finite system and do finite rounds of renormalization. Fortunately, the total dissipation rate does not depend on the lattice size N as long as the total probability is normalized to 1 and periodic boundary conditions are imposed on all sides (Fig. S1a).

To evaluate the flux correlation $C^\star$ at the RG fixed point, we need to extrapolate the correlation from finite rounds of coarse-graining to the limit $s \to \infty$. As shown in Fig. S1b, C decreases with the block size following the relation:

$$C\left(\frac{n_0}{n_s}\right) = \frac{x+p}{x+q}C^\star, \quad x = \log_2\frac{n_0}{n_s}, \tag{S32}$$

with fitting parameters $C^\star = -0.50$, $p = 2.19$, $q = 3.74$, and $R^2 = 0.9992$. Apparently, C converges to $C^\star$ at the RG fixed point. To make the convergence more salient, we demonstrate in Fig. S1c that the inverse of the distance between C and $C^\star$ increases linearly with $x = \log_2\frac{n_0}{n_s}$ and will eventually go to infinity. This correspond to first-order convergence with respect to x.

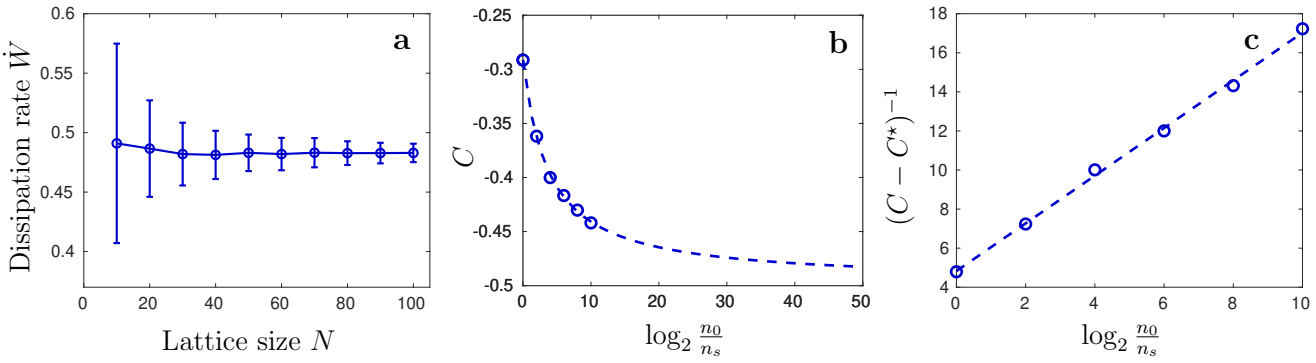

FIG. S1. (a) The steady-state dissipation rate for square lattices with different side length N. Error bar: one standard deviation. The parameters used to generate the lognormal distribution of the transition rates $\ln k \sim \mathcal{N}(\mu, \sigma)$ are $\mu = 0$, $\sigma = 0.5$. (b) The convergence of correlation C in the square lattice. Dashed line: numeric fit which extrapolates to the fixed point $C^\star = -0.5$. (c) To make the convergence in (b) more salient, we plot $\frac{1}{C-C^\star}$ against the block size. Its linear increase demonstrates a first-order convergence to $C^\star = -0.5$ (also see discussion in the SI text).

2. *Statistics of various quantities and justification for approximations*

It was demonstrated that the transition rates at all coarse-grained levels all follow lognormal distribution (Fig. 2A, main text). The fact that the distributions at different coarse-grained levels have the same form manifests the system's self-similarity. Here we present similar evidences concerning the local fluxes and dissipation rates of the system to further demonstrate the self-similar property.

Let us start by looking at the fluxes. Similar to the transition rates discussed in Fig. 2 of the main text, the transition fluxes J also follow lognormal distributions at all coarse-grained levels with an increasing mean and decreasing standard deviation (Fig. S2a–c). Remarkably, the histograms of normalized $\ln J$ all collapse perfectly to a standard Gaussian distribution S2a inset), namely

$$\frac{\ln J - \mu}{\sigma} \sim \mathcal{N}(0, 1). \tag{S33}$$

The same behaviors are observed in the symmetric flux $S_{i,j} = J_{i,j} + J_{j,i}$ (Fig. S2d–f), including the collapsing of the normalized $\ln S$ to a standard Gaussian distribution (Fig. S2d inset). This is not completely surprising since the sum of two random variables independently drawn from identical lognormal distirbutions is approximately lognormal-distributed, per Fenton-Wilkinson approximation [5, 6]. In contrast, the antisymmetric flux A no longer follows lognormal distribution. As demonstrated in Fig. S2g, it is symmetrically distributed with respect to $A = 0$ and approximately decays exponentially on both sides. The probability density function is the Laplace distribution:

$$p(A) = \frac{1}{2A_0} \exp(-|A|/A_0). \tag{S34}$$

The parameter A_0 can be determined by max-likelihood estimation $\hat{A}_0 = \langle |A| \rangle$. It increases nearly linearly with the logarithm of the block size $\log_2 (n_0/n_s)$ (Fig. S2i). The distribution of the scaled net fluxes A/A_0 at all coarse-grained levels collapse to a (unit) Laplace distribution $p(x) = \frac{1}{2} \exp(-|x|)$, except for deviations near $x = 0$ (Fig. S2h).

Notably, the growth of μ and A_0 for the fluxes in Fig. S2 entails that S grows as fast as the block size n_0/n_s while $|A|$ merely grows as fast as the logarithm of the block size, namely $|A| \sim \log_2 (n_0/n_s)$. Therefore, $|A/S|$ must decay to zero as the block size goes to infinity, further justifying the approximation $|A/S| \ll 1$ employed in Eq. S4.

We also numerically evaluate the validity of approximations taken in Eq. S9. Concretely, it includes two layers of approximations:

$$\frac{\left\langle \frac{A_{i,j}^2}{S_{i,j}} \right\rangle}{\left\langle \frac{A_{i_\alpha,j_\beta}^2}{S_{i_\alpha,j_\beta}} \right\rangle} \approx \frac{\left\langle A_{i,j}^2 \right\rangle \left\langle S_{i,j}^{-1} \right\rangle}{\left\langle A_{i_\alpha,j_\beta}^2 \right\rangle \left\langle S_{i_\alpha,j_\beta}^{-1} \right\rangle} \approx \frac{\langle A_{i,j}^2 \rangle}{\langle A_{i_\alpha,j_\beta}^2 \rangle} \frac{\langle S_{i_\alpha,j_\beta} \rangle}{\langle S_{i,j} \rangle} = 1 + C \tag{S35}$$

The first approximation concerns the correlation between A^2 and S^{-1}. The second approximation concerns the difference between $1/\langle S \rangle$ and $\langle S^{-1} \rangle$. For the first one, we directly calculate the ratio of $\langle A^2 \rangle \langle S^{-1} \rangle$ to $\langle A^2 S^{-1} \rangle$ and

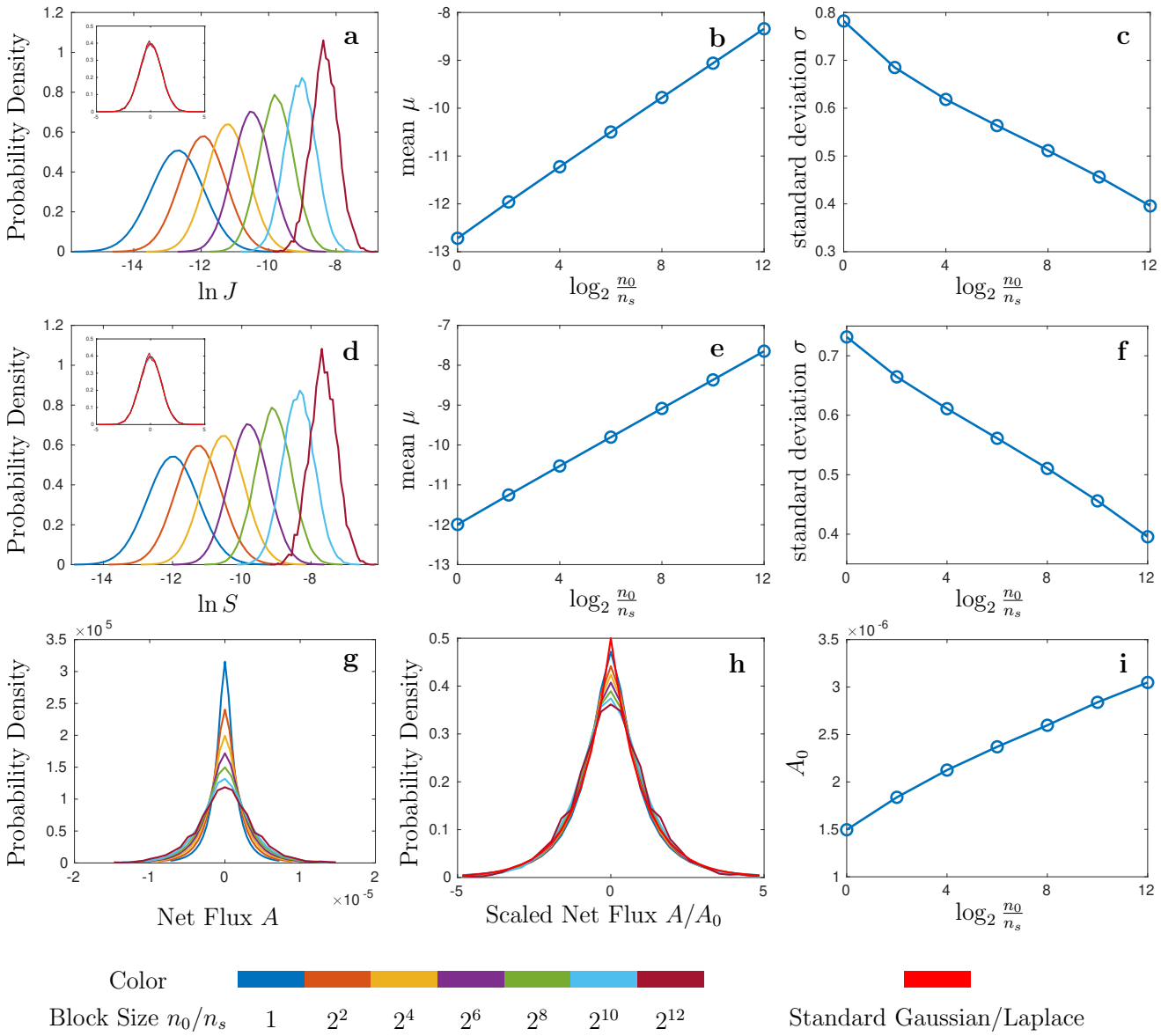

FIG. S2. Statistics of various fluxes (J, S, A) at different coarse-grained levels. (a) The probability density function for $\ln J$. Inset: the distributions at all levels collapse to a standard Gaussian distribution $p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\right)$ (red curve) after shifting the mean to 0 and scaling the standard deviation to 1. (b)–(c) The parameters μ and σ of the lognormal distribution $\ln J \sim \mathcal{N}(\mu, \sigma)$ as a function of the block size n_0/n_s. (d)–(f) depicts the statistics of $\ln S$ in the exact same format as (a)–(c). (g) The probability density function for the net flux $\ln A$. (h) the distributions of the fluxes scaled by A_0. The red curve is Laplace distribution $p(x) = \frac{1}{2}\exp(-|x|)$. (i) The distribution parameter A_0 as a function of the block size n_0/n_s. The corresponding relation between the colors in histograms (a), (d), (g), (h) and the block sizes of coarse-grained levels are given at the bottom.

plot it in Fig. S3. The deviation from 1 indicates negative correlation between A^2 and S^{-1}. The ratio effectively adds another term to the dissipation scaling exponent. The magnitude of change to the scaling exponent is

$$\delta\lambda = \frac{\Delta \log_2 \left(\langle A^2\rangle\langle S^{-1}\rangle / \langle A^2 S^{-1}\rangle\right)}{\Delta \log_2 (n_0/n_s)} = \frac{\ln (3.40/1.54)}{16 \ln 2} = 0.07, \tag{S36}$$

which is small compared to the $\log_4(1+C)$ term due to correlation.

Now we turn to the second approximation. Since we know that S follows a lognormal distribution $\mathcal{N}(\mu, \sigma)$, the

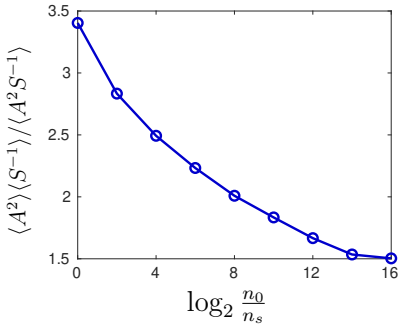

FIG. S3. The ratio of $\langle A^2\rangle\langle S^{-1}\rangle$ to $\langle A^2S^{-1}\rangle$ as a function of the block size n_0/n_s.

difference between $1/\langle S\rangle$ and $\langle S^{-1}\rangle$ can be computed analytically:

$$\begin{aligned}\frac{1}{\langle S\rangle} &= \left[\int_0^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x-\mu)}{2\sigma^2}\right) \mathrm{d}x\right]^{-1} = e^{-\mu-\frac{\sigma^2}{2}},\\ \langle\frac{1}{S}\rangle &= \int_0^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x-\mu)}{2\sigma^2}\right) \frac{\mathrm{d}x}{x^2} = e^{\frac{\sigma^2}{2}-\mu}.\end{aligned} \tag{S37}$$

The ratio is

$$\langle\frac{1}{S}\rangle / \frac{1}{\langle S\rangle} = e^{\sigma^2} \tag{S38}$$

Therefore, the second approximation

$$\frac{\frac{\langle A^2_{i,j}\rangle\langle S^{-1}_{i,j}\rangle}{\langle A^2_{i_\alpha,j_\beta}\rangle\langle S^{-1}_{i_\alpha,j_\beta}\rangle}}{\frac{\langle A^2_{i,j}\rangle}{\langle A^2_{i_\alpha,j_\beta}\rangle}\frac{\langle S_{i_\alpha,j_\beta}\rangle}{\langle S_{i,j}\rangle}} = \exp\left(\sigma_s^2-\sigma_{s-1}^2\right). \tag{S39}$$

The change of σ^2 can be estimated from Fig. S2f. σ decreases from 0.73 to 0.40 as the block size increases from 1 to 2^{12}. Therefore, the influence of this approximation to the scaling exponent is

$$\delta\lambda = \frac{\log_2 e^{\sigma_0^2-\sigma_{12}^2}}{\log_2(n_0/n_{12})} = \frac{0.73^2-0.40^2}{12\ln 2} = 0.04, \tag{S40}$$

which is negligible.

B. The spatial profile of dissipation rate

Next, we investigate the spatial profile of dissipation. The local dissipation rate at site (i,j) is defined by the average of the dissipation in all four links associated with it, namely

$$w_{i,j} = \frac{1}{4}\sum_{\alpha=1}^{4}\left(J_{(i,j)\to(i_\alpha,j_\alpha)} - J_{(i_\alpha,j_\alpha)\to(i,j)}\right)\ln\frac{J_{(i,j)\to(i_\alpha,j_\alpha)}}{J_{(i_\alpha,j_\alpha)\to(i,j)}}. \tag{S41}$$

(i_α,j_α) $(\alpha=1,2,3,4)$ are the four neighbors adjacent to (i,j), namely $(i,j+1)$, $(i,j-1)$, $(i+1,j)$, and $(i-1,j)$. The local dissipation rate is self-similar in the sense of both its spatial profile and overall distribution. Fig. S4 demonstrates the spatial profile of $\ln w_{i,j}$ at the first three levels. The arrows indicate the direction of coarse-graining. The overall profiles are similar.

Fig. S5 presents the statistical distribution of $\ln w_{i,j}$. The distribution at all scales collapse to a single distribution function after shifting the mean to zero and scaling the standard deviation to one (Fig. S5b), which stands as another example of self-similarity. Unlike the common distribution function for $\ln J$, which is Gaussian, the one found here undoubtedly deviates from the standard Gaussian distribution. The mean and standard deviation of $\ln w$ are presented in Fig. S5c–d.

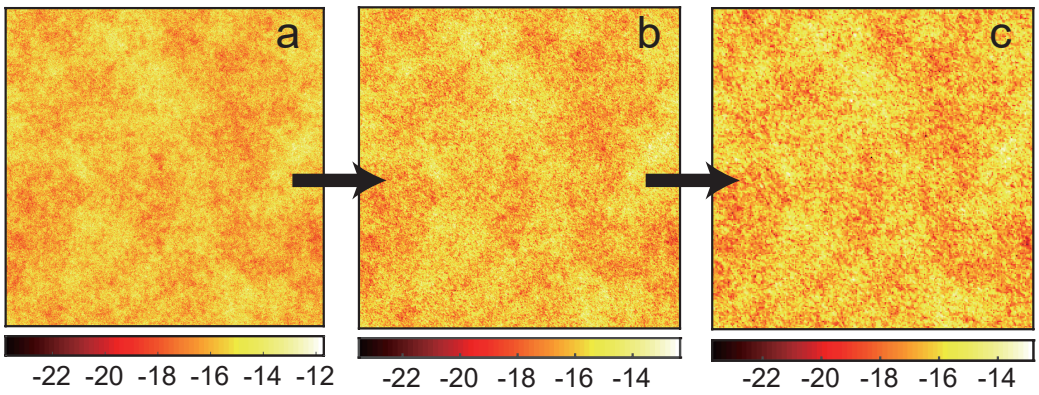

FIG. S4. Local dissipation rate profile in the square lattice (in log scale, plotting $\ln w_{i,j}$). (a) The original network of size $N = 1024$. The transition rates follow a lognormal distribution with mean $\mu = 0$ and standard deviation $\sigma = 0.5$. (b) The lattice resulting from the coarse-graining of (a), of size $N = 512$. (c) The lattice resulting from the coarse-graining of (b), of size $N = 256$.

C. The dissipation scaling law does not depend on the rate distribution.

It is apparent from the derivation in section IA of the SI that the dissipation scaling relation does not depend on the exact form of initial transition rate distribution. To demonstrate this point, we performed all the studies in the square lattice with transition rates sampled from three other distributions: Weibull distribution, gamma distribution, and exponential distribution. The probability density function for these distributions are:

$$\begin{aligned}
&\text{Weibull: } f(x;\lambda,k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}\exp\left(-\left(\frac{x}{\lambda}\right)^{k}\right) && (x>0)\\
&\text{gamma: } f(x;\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} && (x>0)\\
&\text{exponential: } f(x;\mu) = \frac{1}{\mu}\exp\left(-\frac{x}{\mu}\right) && (x>0)
\end{aligned} \tag{S42}$$

The last one is a special case of the gamma distribution (with $\alpha = 1$).

Fig. S6 presents the results with these three distributions. The parameters are chosen as

$$k = 5, \lambda = \frac{1}{\Gamma(1+1/k)};\ \ \alpha = \beta = 2;\ \ \mu = 1, \tag{S43}$$

such that the mean of the transition rates is 1 in all three cases. In all three systems, the dissipation scaling behaviors are identical to that displayed in Fig. 2c of the main text, with scaling exponents close to 1.35 (Fig. S6a–c). However, the distribution of transition rates evolves differently (Fig. S6d–f). If the original transition rate distribution is lognormal, it stays lognormal at all coarse-grained levels (Fig. 2a, main text). On the contrary, the distributions here start with a considerable deviation from lognormal at the finest level but gradually converge to lognormal as the system is coarse grained (Fig. S6g–i). In other words, the lognormal distribution is an stable attractor for the transition rates distribution, i.e. the rates distribution at the RG fixed point. Since we are mainly concerned with the scaling properties at the RG fixed point of the reaction network, it is reasonable to start with a lognormal distribution and study the convergence or scaling of other properties as the system is coarse-grained. Therefore, it suffices to consider only lognormal-distributed rates in all networks studied here.

D. The dissipation scaling law does not require the landscape to be flat.

In the main text, we focused on the flat landscape assumption to simplify the problem. However, the scaling behavior reported in the text does not rely on this assumption. It is also applicable to energy landscapes that are not flat and even spatially correlated.

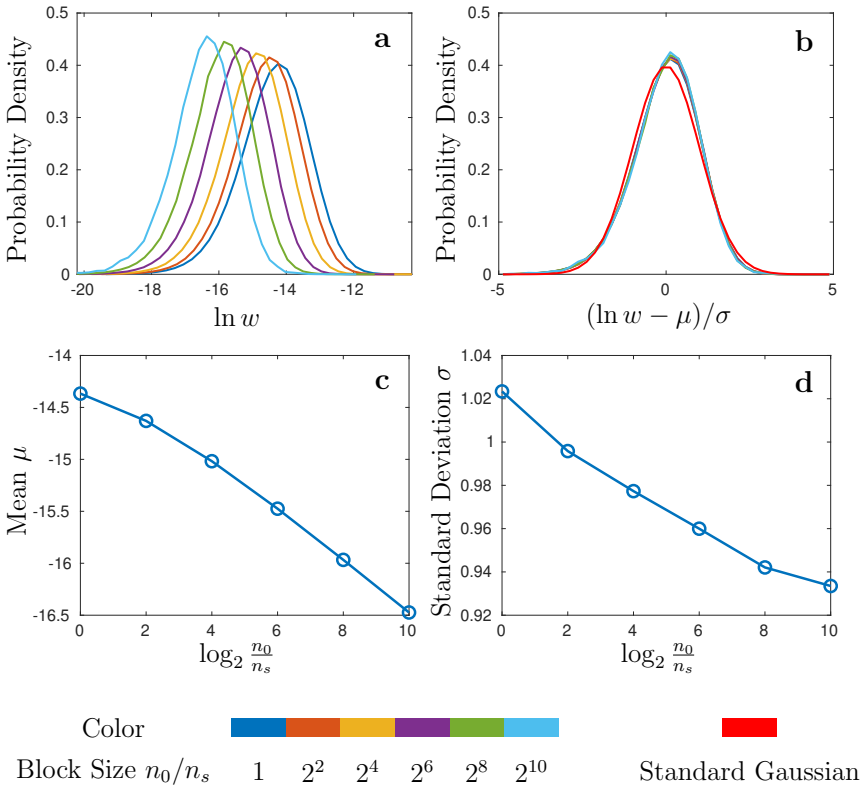

FIG. S5. Statistics of the local dissipation rate $w_{i,j}$ (in log scale). (a) The probability density function of $\ln w$ at different coarse-grained levels of the network (see the bottom of the figure for color denotation). (b) After shifted to zero mean and scaled to unit standard deviation, distributions of $\ln w$ at all levels collapse to a single distribution. However, this distribution deviates from the standard Gaussian distribution (red curve). (c)–(d) mean μ and σ of $\ln w$ at different coarse-grained levels, as functions of the block size.

1. *Scaling law on a random rough landscape without spatial correlation*

To begin with, we consider a rough energy landscape that has no spatial correlation: the energy levels of individual states are independent and identically distributed (i.i.d.) Gaussian variables, whose probability distribution function is given by:

$$E_i \sim \mathcal{N}(0,\sigma_E) \Leftrightarrow P(E_i) = \frac{1}{\sqrt{2\pi\sigma_E^2}} e^{-\frac{E_i^2}{2\sigma_E^2}}, \quad P(E_i,E_j) = P(E_i)P(E_j) \tag{S44}$$

The mean of the energy landscape is set to zero since it only sets the reference point and does not affect the transition dynamics. The roughness of the landscape is characterized by the standard deviation σ_E. It has the same dimension as the parameter σ in the lognormal distribution for $\gamma_{i,j}$, which characterizes the energy scale of the nonequilibrium driving. For the sake of clarity, we use σ_γ to denote this latter σ in the following discussion. The reaction rates are related to the energy landscape through Eq. 1 of the main text.

Fig. S7 presents the dissipation scaling relation in two example systems where the energy landscapes have different roughness ($\sigma_\gamma = 0.5$, $\sigma_E = 0.5$; $\sigma_\gamma = 0.5$, $\sigma_E = 2.5$). This demonstrates that the scaling law reported in the main text exists in systems with a rough landscape. The scaling exponent λ depends on the specific realization of the system (because the rates are randomly drawn and the system is finite). Therefore, we tested the scaling laws for multiple realizations for each different σ_E and confirmed that the scaling law remained valid. Moreover, the scaling exponents λ also does not change with the roughness of the landscape (characterized by σ_E, see Fig. S8). The small fluctuation in λ is only due to the sampling of $\gamma_{i,j}$ in a finite system ($N = 256$).

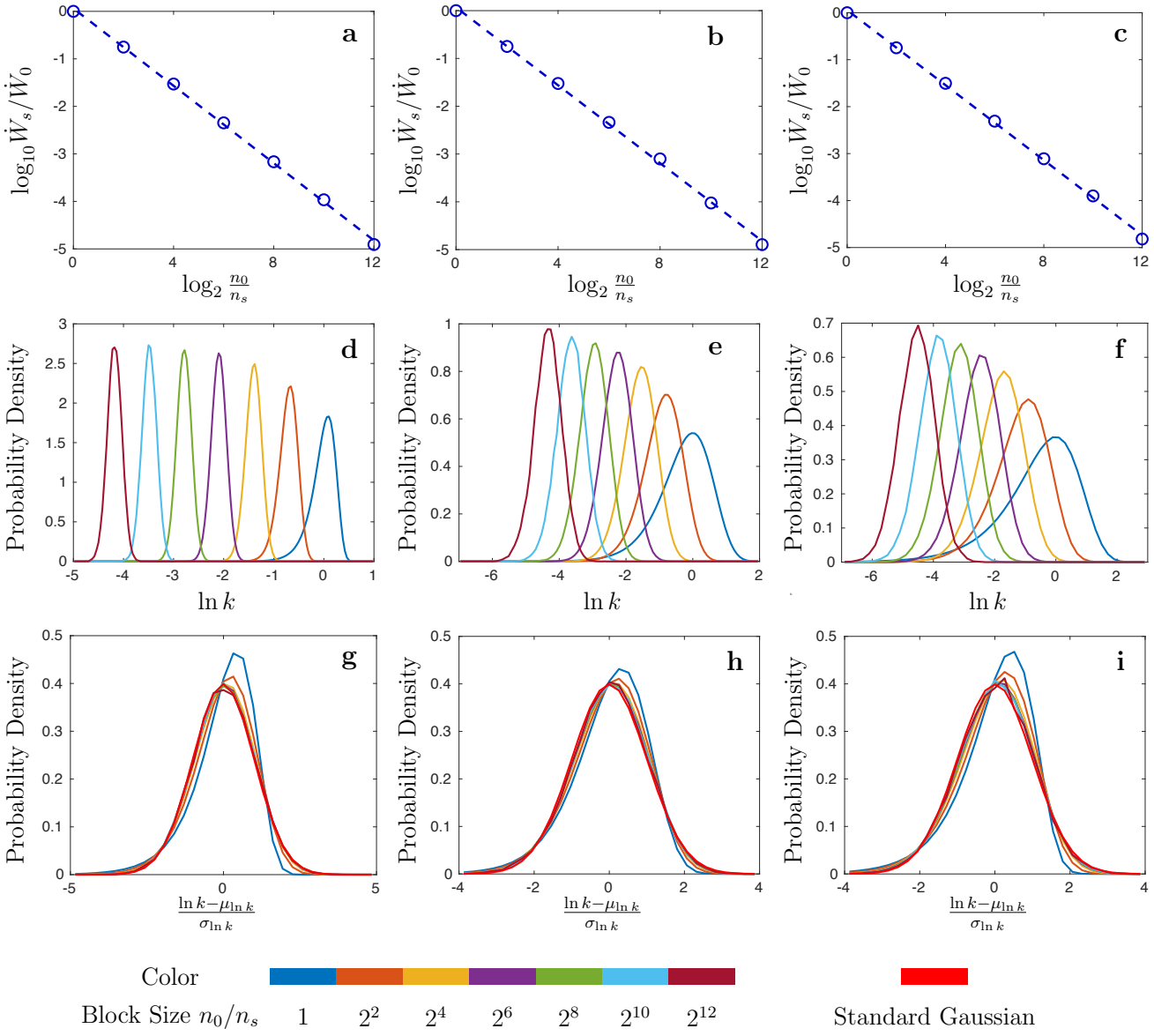

FIG. S6. Results in the square lattice where transitions follow Weibull distribution (left panel), gamma distribution (middle panel), or exponential distribution (right panel). (a)–(c): The dissipation rate decreases in a power-law relation with the block size. The scaling exponents from the least-square fit are 1.35, 1.35, and 1.33. (d)–(f): The distribution of $\ln k$ (k is the transition rate) at different coarse-grained levels. (g)–(f) The distribution of $\frac{\ln k-\mu_{\ln k}}{\sigma_{\ln k}}$ at different coarse-grained levels, where $\mu_{\ln k}$ and $\sigma_{\ln k}$ are the mean and the standard deviation of $\ln k$, respectively. The transformation shifts the mean to zero and scales the standard deviation to 1 to allow the comparison between the shape of the probability density functions. The red curve is the standard Guassian distribution $p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\right)$. The corresponding relation between the colors in histograms (d)–(i) and the block sizes of coarse-grained levels are given at the bottom.

2. *Scaling law on a random rough landscape with spatial correlation*

Next, we show that the scaling relation is valid in an energy landscape that has a finite correlation length ξ_0. To demonstrate the effect of spatial correlation. We consider the following energy distribution:

$$E_i \sim \mathcal{N}(s_i\epsilon, \sigma_E) \tag{S45}$$

where s_i are taken from a snapshot of spins in an equilibrium Ising model at a finite temperature T. $\{s_i\}$ are therefore correlated with one another, and so does the energy landscape $\{E_i\}$. The correlation length is determined by fitting

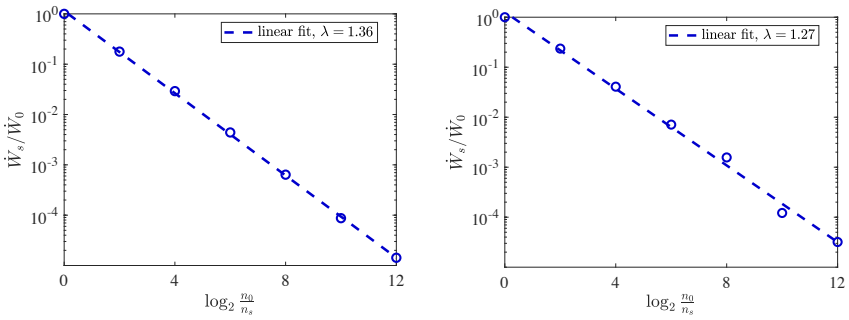

FIG. S7. Two examples of the dissipation scaling behavior on a non-flat energy landscape following identical independent distributions (Eq. S44). Left: $\sigma_E = 0.5$; Right: $\sigma_E = 2.5$. The nonequilibrium parameters $\gamma_{i,j}$ follow a lognormal distribution with $\mu = 0$ and $\sigma_\gamma = 0.5$.

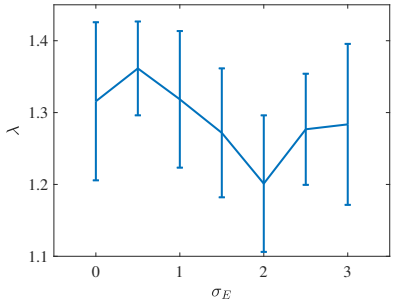

FIG. S8. Scaling exponents on a rough energy landscape. σ_γ was evenly sampled in $(0, 3]$ and has been averaged out. Errorbar: 1 standard deviation.

the correlation function to the following function

$$C_E(r) = \langle E_i E_{i+r} \rangle = A x^{-1/4} \exp\left(-\frac{x}{\xi}\right). \tag{S46}$$

The energy landscape, correlation function, and dissipation scaling relation of an example of such system are presented in Fig. S9, where the correlation length is about $\xi = 28$. Indeed, the landscape is non-flat and has some local correlated structure. However, as shown in the last panel, the dissipation scaling relation still holds. To demonstrate the generality of our results, we extensively sampled the temperature T of the Ising model to examine the dissipation scaling law in systems with correlation lengths ranging from ~ 1 to those comparable to the system size. As shown in Fig. S11, the correlation length ξ and the absolute dissipation rate $\dot{W}_0$ changes with temperature T, but the scaling law holds in all these systems. This is not hard to imagine that the scaling law still holds in landscapes with long-range correlation since increasing ξ has a similar effect of reducing σ_E. When ξ reaches the system size, we effectively get a flat landscape, and the energy dissipation follows the same scaling law reported in the main text.

To conclude, we have shown numerically that the scaling law reported in the main text is applicable to random energy landscapes that are rough and even spatially correlated.

3. Scaling law on a periodic energy landscape

Here we examine the dissipation scaling law in another form of energy landscape which is non-flat but regular. More specifically, we consider a square lattice with the following landscape, which resembles the profile of standing wave:

$$E(x, y) = A_E \sin\frac{2\pi x}{\Lambda_x} \sin\frac{2\pi y}{\Lambda_y} \tag{S47}$$

This landscape is fundamentally different from the random landscape due to the existence of a characteristic length scale $\Lambda = \Lambda_x = \Lambda_y$. We demonstrate that the scaling relation is valid after the coarse-graining process goes across this characteristic length scale Λ.

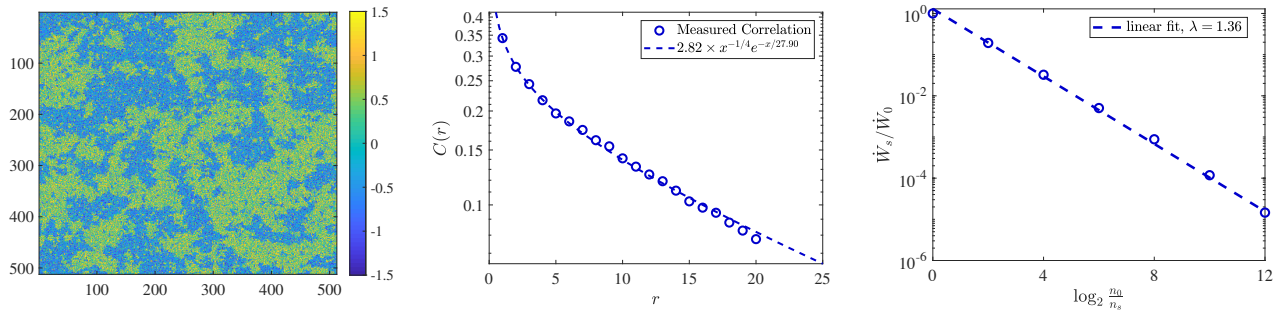

FIG. S9. An example of the scaling behavior of energy dissipation on a landscape that has a correlation length of $\xi = 28$. Left: energy landscape; middle: correlation of energy landscape; right: scaling of energy dissipation The parameters are $T = T_c$ (T_c is the critical tempearture; due to the finite size, the correlation does not diverge at T_c), $\mu = 0$. $\epsilon = \sigma_E = \sigma = 0.5$.

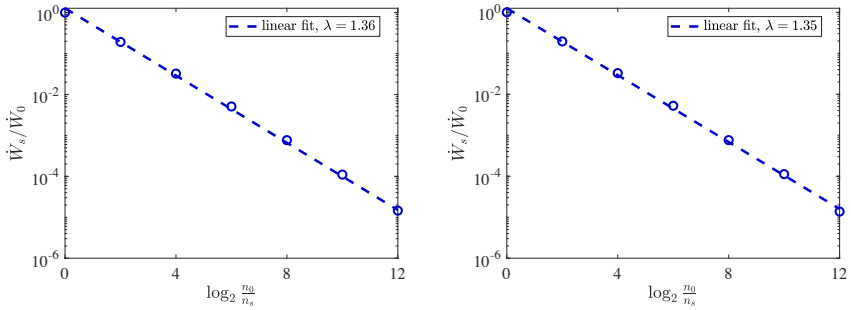

FIG. S10. Scaling of energy dissipation at $T = 0.95T_c$ (left) and $T = 1.20T_c$ (right). The landscape for the left panel has a larger correlation length. Parameters: $\mu = 0$. $\epsilon = \sigma_E = \sigma = 0.5$.

To examine the effect of this particular length scale, we construct landscapes on an $N \times N$ ($N = 512$) square lattice with different characteristic length scale Λ and computed the dissipation rate at different coarse-grained levels (Fig. S12). As the system is coarse-grained, the dissipation rate experiences a sudden drop and then collapse to the same scaling relation (black dashed line) as reported in the flat landscape. We call the first point after the sudden decrease in dissipation the onset of scaling. The results in Fig. S12 demonstrate that the onset is delayed by one coarse-graining iteration every time the length scale Λ is doubled. This indicates that the onset of scaling takes place when the scale of a coarse-grained state is comparable to the characteristic length scale Λ, and the scaling law holds in all the following coarse-graining iterations.

Therefore, for regular systems with a characteristic length scale, the dissipation scaling law holds after coarse-

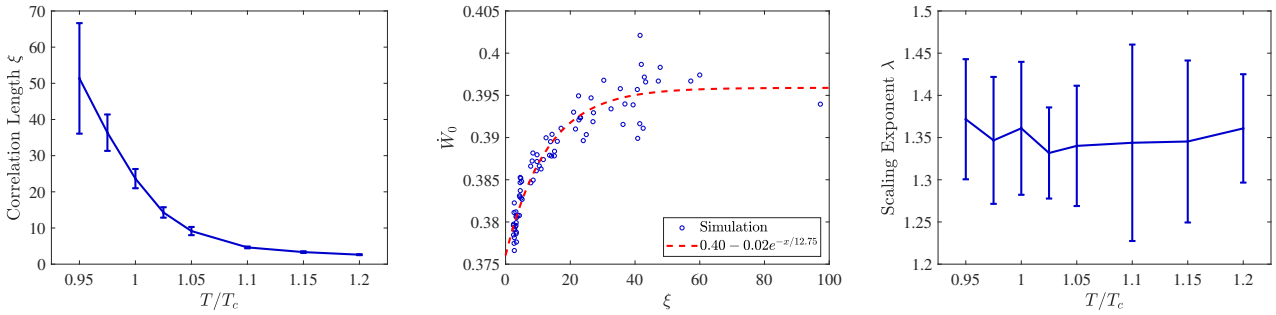

FIG. S11. Summary of the calculations on a rough energy landscape. Left: the correlation length can be tuned by the temperature parameter T. Middle: The total dissipation rate $\dot{W}_0$ increases with the correlation length ξ and saturates to the value on a flat energy landscape. Right: The scaling exponent does not depend on T, and therefore does not depend on ξ. The errorbar stands for one standard deviation. The spread of λ is due to the finite size of our system ($N = 512$). The parameters are $\mu = 0$. $\epsilon = \sigma_E = \sigma = 0.5$.

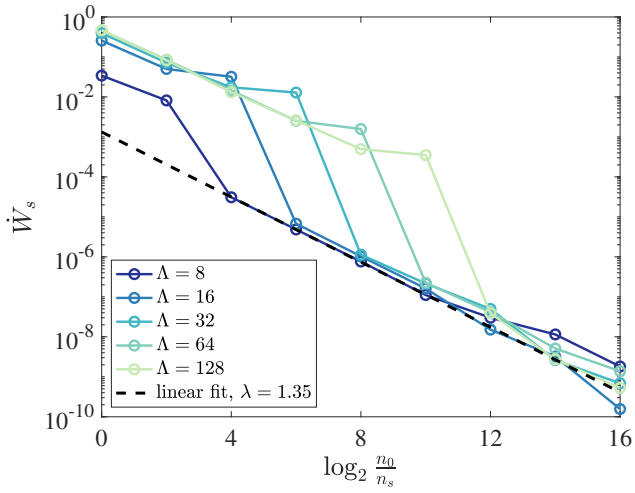

FIG. S12. Scaling of energy dissipation on the periodic landscape given by Eq. S47 with different length scale parameters Λ. Parameters of the energy landscape: $A_E = 10$; for the lognormal distribution of γ: $\mu = 0$, $\sigma = 0.5$.

d	m	λ
4	4	1.057 ± 0.027
6	4	1.068 ± 0.021
6	6	1.070 ± 0.022
8	6	1.075 ± 0.020
10	6	1.077 ± 0.021
8	8	1.055 ± 0.015

TABLE S1. The dissipation scaling exponents of the random hierarchical network (mean $\pm$ one standard deviation, over 100 replica). The scaling exponent λ is close to 1, independent of the average degree d and the number of states with each macroscopic state m.

graining goes across the characteristic length scale.

E. The random-hierarchical network

Table S1 presents the scaling exponents for different random hierarchical networks constructed with parameters d (the average degree) and m (the average cluster size). The last column shows that the dissipation scaling exponent λ is in close proximity to 1 regardless of d and m, corroborating the conclusion $\lambda_{\mathrm{RHN}} = 1$ in the main text.

F. The scale-free network

First, we briefly summarize the construction of the 2D-embedded scale-free network, which was used in refs. [7, 8]. Briefly, the network is constructed by assigning degrees to all sites on a square lattice according to a power-law distribution and fulfilling the degree requirements by considering the sites in a random order. For each site, we examine its neighbors from close to distant and create links whenever possible, until the number of its links reaches its pre-assigned degree or the search radius reaches an upper limit. The preference to connecting with nearer neighbors allows us to use the coarse-graining method that group neighboring states together. Although not all degree requirements can be satisfied, the resulted network is indeed scale-free, consistent with previous work [8]. The 2D-embedding allows us to directly use the coarse-graining procedure employed in the square lattice, which preserves the probability fluxes between coarse-grained states.

Next, we demonstrate the determination of α and d_B with the example in Fig. S13. After creating the scale-free network, we plot the cumulative degree distribution $P(k) \sim \left(\frac{k_{\min}}{k}\right)^{\alpha-1}$ (see Eq. S24) against k on a log-log scale and

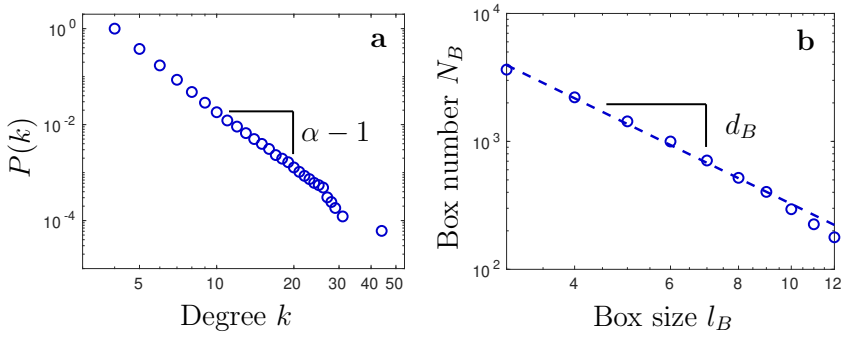

FIG. S13. (a) The cumulative degree distribution is power-law ($P(k) \propto k^{1-\alpha}$). The slope is $\alpha - 1 = 4.00$ in this particular example. (b) The power-law relation between N_B and l_B reveals a fractal dimension d_B. The slope is $d_B = 1.98$ in this particular example.

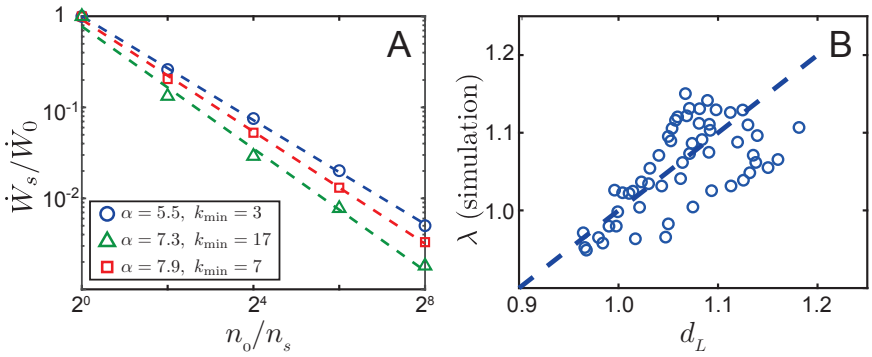

FIG. S14. The scaling of energy dissipation in different scale-free networks. (A) The scaled dissipation rate $\dot{W}_s/\dot{W}_0$ in three different scale-free networks. (B) The dissipation scaling exponent in simulation is positively correlated with the link exponent d_L, in 56 different scale-free networks. The dashed line corresponds to perfect agreement $\lambda = d_L$. The correlation coefficient is 0.65.

determine the slope $\alpha - 1$ by linear fitting and therefore determine α. The index d_B is obtained similarly by fitting the power-law relation between the box size l_B and the minimum number of boxes containing states whose pairwise distance does not exceed l_B that are needed to cover the entire graph, N_B. Since finding the exact minimum $N_B(l_B)$ is NP-hard, we instead use a greedy coloring algorithm that is known to provide good approximations to the value of $N_B(l_B)$ [3].

The dissipation rate in the 2D-embedded SFN also scales with block size as a power law with the exponent λ depending on the network structure (Fig. S14A). To test the estimation of $\lambda_{\rm SF}$ (Eq. 11 in the main text), we constructed 56 scale-free networks with $\alpha \in [5, 7]$ and $k_{\min} \in [3, 22]$ and computed the energy dissipation scaling exponent. As shown in Fig. S14B, there is indeed a positive correlation between d_L and $\lambda_{\rm SF}$.

G. Networks without the dissipation scaling relation

Finally, we introduce two networks in which the dissipation does not obey a power-law relation with respect to the block size after coarse-graining and discuss why this scaling relation breaks down.

1. The Erdős-Rényi random network

We create an Erdős-Rényi random network with a binomial model that is equivalent to the method used in the classic paper by Erdős and Rényi [9]. To generate a network with average degree d, we connect each pair of states on a N_0 by N_0 square lattice with a constant probability $p = \frac{d}{N_0^2(N_0^2-1)}$. The coarse-graining is conducted in the exact same way as that used in the 2D-embedded scale-free network. We find that the dissipation decreases with the block

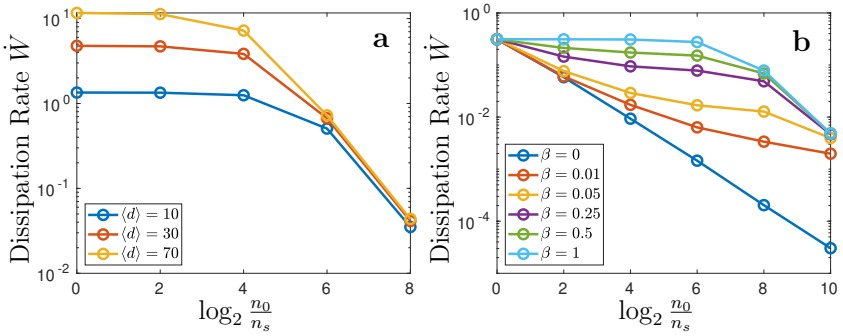

FIG. S15. The dissipation rate in the Erdős-Rényi random network and Watts-Strogatz small-world network. The initial transition rates are drawn from lognormal distribution $\ln k \sim \mathcal{N}(\mu, \sigma)$ with $\mu = 0$ and $\sigma = 0.4$. (a) Results in Erdős-Rényi random network with different average degrees $\langle d \rangle$ for $N_0 \times N_0$ nodes ($N_0 = 128$). (b) Results in the Watts-Strogatz small-world network with different rewiring probabilities β (square lattice with 256×256 nodes).

size but does not follow a power-law relation (Fig. S15a). This is not unexpected since the coarse-graining completely disregards the connectivity of the network. In the scale-free network, states are preferentially linked to neighbors by construction, which relates the network topology with the geography of the square lattice. Hence, grouping together geographically adjacent states during coarse-graining also means that states that are more interconnected (i.e., neighbors) are more likely to belong to the same macrostate. In stark contrast, a state in an Erdős-Rényi random network has equal probabilities of being connected to a state that belong to the same macrostate or a state that does not. Therefore, the first few iterations of coarse-graining of the sparse network almost does not involve any link merger, while the system approaches a complete graph at later rounds of coarse-graining with the number of links rapidly decreasing. The discrepancy between early and later stages of coarse-graining reveals the lack of self-similarity in the Erdős-Rényi random network, which inevitably leads to the absence of dissipation scaling.

2. *The Watts-Strogatz small-world network*

The Watts-Strogatz network can be created by randomly rewiring the links in a regular square lattice with probability $\beta \in (0, 1)$ [10]. Coarse-graining can be carried out in the same fashion as scale-free network and Erdős-Rényi random network. The network approaches a square lattice as $\beta \to 0$ and approaches the Erdős-Rényi random network as $\beta \to 1$. Therefore, we expect the dissipation to scale with the block size as β tends to 0 and to breaks down as β is increased. Indeed, this is what we find numerically (Fig. S15b). The same arguments used in the random network can be employed to account for the breakdown of the scaling property here.

III. DISSIPATION SCALING IN BIOCHEMICAL SYSTEMS

In this section, we utilize the coarse-graining method to study three realistic biochemical reaction systems discussed in the main text and elucidate the scaling bahavior of dissipation in these systems.

A. Brusselator

a. The Brusselator model and state-space coarse-graining. The Brusselator is one of the biochemical network motifs that generate sustained oscillation [12, 13]. To incorporate dissipation calculation, we take the reversible Brusselator model studied in a previous work [14] and study the thermodynamic effect of coarse-graining the state space. The reactions are described by

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X; \quad B + X \underset{k_{-2}^0}{\overset{k_2^0}{\rightleftharpoons}} D + Y; \quad 2X + Y \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} 3X, \tag{S48}$$

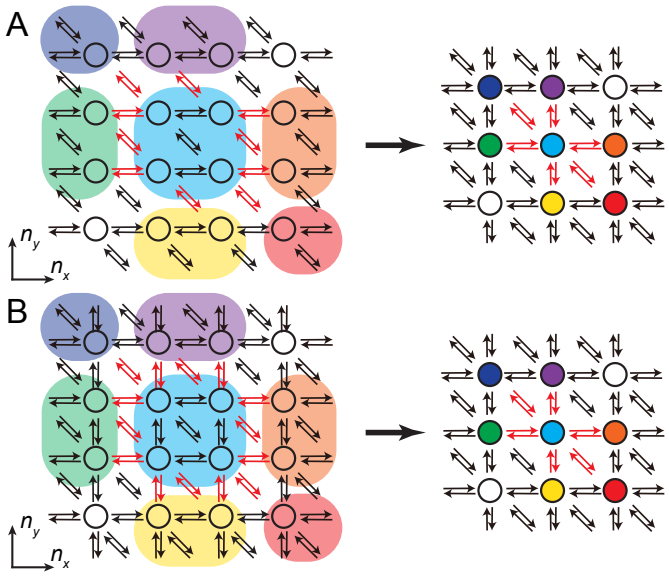

FIG. S16. Illustration of the coarse-graining of the Brusselator model. The states in the colored shades in the left panel are merged to states with the same color in the right panel. The red links are those connecting the light blue coarse-grained state with the six neighboring coarse-grained states (red, orange, yellow, green, purple, dark blue). (A) The first coarse-graining operation changes the network topology by creating links along the y direction. (B) All subsequent coarse-graining operations do not change the topology of the network.

We construct the state space on a 2D lattice with coordinates (n_X, n_Y) denoting a system having n_X molecules of X and n_Y molecules of Y. The concentrations of all molecular species other than X and Y are considered constants and absorbed into the reaction rates. Following ref. [14], we choose parameter values $[A] = 1$, $k_{\pm 1} = 1$, $k_2 = k_2^0[B] = \frac{1}{\gamma_1}$, $k_{-2} = k_{-2}^0[D] = \gamma_2$, and $k_{\pm 3} = 1$. The irreversibility parameter $\gamma = \gamma_1\gamma_2 \neq 1$ indicates that the system is out of equilibrium and hence dissipative. We consider the system to be well-stirred with a fixed volume V. The master equation which governs the dynamics of the probability distribution $P_{n_x,n_y}(t)$ in the state space reads:

$$\begin{aligned}\frac{\mathrm{d}P_{n_x,n_y}(t)}{\mathrm{d}t} =& k_1 P_{n_x-1,n_y}(t) + k_{-1}\frac{n_x+1}{V}P_{n_x+1,n_y}(t) + k_2\frac{n_x+1}{V}P_{n_x+1,n_y-1}(t) \\ &+ k_{-2}\frac{n_y+1}{V}P_{n_x-1,n_y+1}(t) + k_3\frac{(n_x-1)(n_x-2)(n_y+1)}{V^3}P_{n_x-1,n_y+1}(t) \\ &+ k_{-3}\frac{(n_x+1)n_x(n_x-1)}{V^3}P_{n_x+1,n_y-1}(t) \\ &- \left[k_1 + k_{-1}\frac{n_x}{V} + k_2\frac{n_x}{V} + k_{-2}\frac{n_y}{V} + k_3\frac{n_x(n_x-1)n_y}{V^3} + k_{-3}\frac{n_x(n_x-1)(n_x-2)}{V^3}\right]P_{n_x,n_y}(t),\end{aligned} \tag{S49}$$

The steady-state probability distribution can be solved for from the master equation $\frac{\mathrm{d}P_{n_x,n_y}(t)}{\mathrm{d}t} = 0$, and the coarse-graining can be done by combining four neighboring states together while conserving the net fluxes, which is similar to the protocol used in the square lattice except for the extra diagonal link. Fig. S16 illustrates the reaction dynamics in the state space, including how this model is coarse-grained into a regular lattice with horizontal, vertical, and diagonal links. In the first (original) network, there are only horizontal and diagonal links. The first coarse-graining process combines four adjacent states together, creating the vertical links in the coarse-grained network (Fig. S16A, left $\rightarrow$ right). The subsequent coarse-graining operations do not change the structure of the network (Fig. S16B, left $\rightarrow$ right). This remains as the structures of all coarse-grained networks (self-similarity).

For numeric studies, we construct this state space with cutoffs $n_x \leq n_{x,\max} = 500$ and $n_y \leq n_{y,\max} = 500$. The volume is set to $V = 50$.

b. Scaling of energy dissipation due to coarse-graining the state space of Brusselator. To elucidate the effect of state-space coarse-graining in Brusselator, we study a specific system with $\gamma_1 = 0.2$ and $\gamma_2 = 0.1$, whose results are summarized in Fig. S17. The left panel illustrates the evolution of number of molecules X and Y governed by the

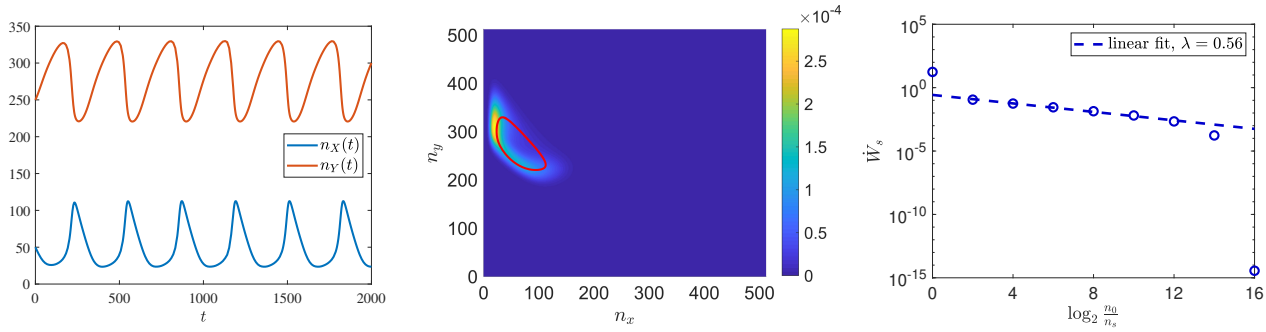

FIG. S17. Scaling of dissipation in an oscillating Brusselator with nonequilibrium parameters $\gamma_1 = 0.2$ and $\gamma_2 = 0.1$. Left: oscillation of the number of molecules n_X and n_Y calculated from the deterministic equations; middle: steady-state probability distribution in the state space determined by solving the chemical master equation, and the deterministic limit cycle (red circle); right: scaling of energy dissipation as a result of coarse-graining. The parameter values used are: $[A] = 1$, $k_{\pm 1} = 1$, $k_2 = k_2^0[B] = \frac{1}{\gamma_1}$, $k_{-2} = k_{-2}^0[D] = \gamma_2$, and $k_{\pm 3} = 1$. These same parameters are used for Fig. 3B in the main text.

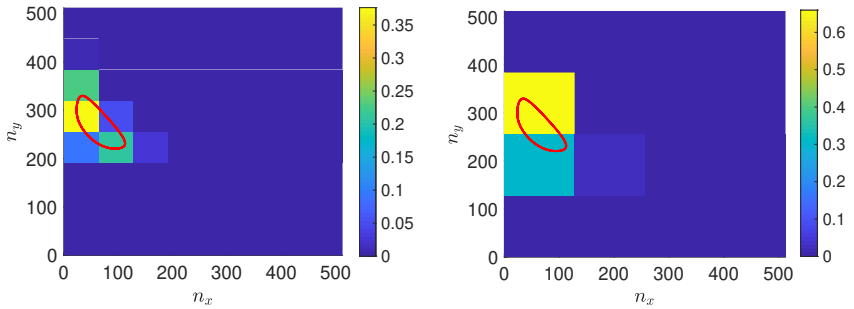

FIG. S18. Termination of the dissipation scaling in an oscillating Brusselator with nonequilibrium parameters $\gamma_1 = 0.2$ and $\gamma_2 = 0.1$. Left and right: the probability distribution in the coarse-grained state space with the deterministic limit cycle (red circle) for $\log_2 \frac{n_0}{n_s} = 12$ (left) and $\log_2 \frac{n_0}{n_s} = 14$ (right).

deterministic ordinary differential equations[1]:

$$\begin{aligned} \frac{\mathrm{d}n_X(t)}{\mathrm{d}t} &= k_1 - k_{-1}\frac{n_X}{V} - k_2\frac{n_X}{V} + k_{-2}\frac{n_Y}{V} + k_3\frac{n_X^2 n_Y}{V^3} - k_{-3}\frac{n_X^3}{V^3} \\ \frac{\mathrm{d}n_Y(t)}{\mathrm{d}t} &= k_2\frac{n_X}{V} - k_{-2}\frac{n_Y}{V} - k_3\frac{n_X^2 n_Y}{V^3} + k_{-3}\frac{n_X^3}{V^3} \end{aligned} \tag{S50}$$

Indeed, the dynamics of the number of molecules indicate that the Brusselator is in oscillation. The middle panel plots the state-space probability distribution overlaid with the limit cycle of the deterministic system (red circle), which is in agreement with Fig. 2b of ref. [14]. The center of the limit cycle is located at $(\langle n_x \rangle, \langle n_y \rangle) = (50, 283)$. Finally, the right panel demonstrates the dissipation rate at different coarse-grained levels. The first coarse-graining operation changes the structure of the network (see Fig. S16A) and results in a significant decrease in the dissipation rate. The subsequent coarse-graining operations preserve the structure of the network (see Fig. S16B), and the dissipation rate scales following a power-law as reported for self-similar networks in the main text. This illustrates the general applicability of the scaling law to the coarse-graining of the state space of realistic biochemical reaction networks. For the Brusselator model, the link exponent is $d_L = 1$. Thus, the observed exponent $\lambda < 1$ implies that the correlation between fluxes is positive: $C^\star > 0$.

After a certain number of coarse-graining operations, the dissipation rate drops drastically and exits the regime described by the scaling law (the last two points fall below the dashed line in Fig. S17, right panel). This can be understood by looking at the state space of the coarse-grained system. The significant decrease in dissipation rate

[1] These are in principle the equations that govern the evolution of concentrations $[X]$ and $[Y]$ in the $V \to \infty$ limit. Here we apply it to our finite system with $n_X = V[X]$ and $n_Y = V[Y]$ to illustrate the location of the limit cycle.

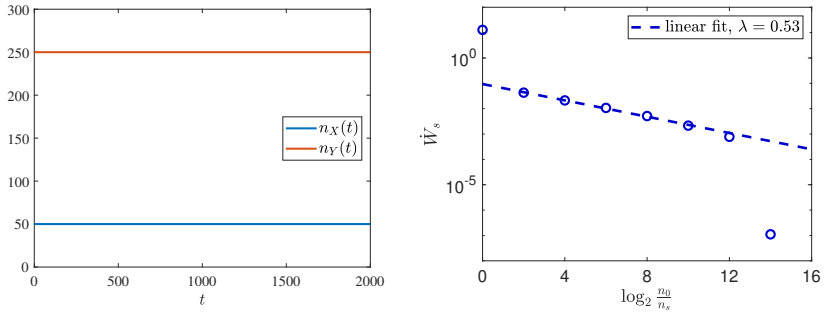

FIG. S19. Scaling of dissipation in a non-oscillating Brusselator with nonequilibrium parameters $\gamma_1 = 0.2$ and $\gamma_2 = 0.2$. Left: oscillation of the number of molecules n_X and n_Y; right: scaling of energy dissipation as a result of coarse-graining.

takes place between $\log_2 \frac{n_0}{n_s} = 12$ and $\log_2 \frac{n_0}{n_s} = 14$ (Fig. S17, right panel), which correspond to coarse-graining the state space from the left panel to the right panel of Fig. S18. After this coarse-graining iteration, there are only two states with a significant probability (yellow and cyan). Since there are only one link between these two states, they are almost in equilibrium, and the dissipation becomes very small. Therefore, there is an upper limit of the coarse-graining iterations which terminates the scaling relation. After crossing the upper limit, the system has too few states, and the dissipation rates becomes very small.

c. Scaling does not rely on oscillation. The dissipation in Brusselator stems from the breaking of detailed balance ($\gamma \neq 1$) and does not depend on the system to be in the oscillation phase. Even when the system is not oscillating, the existence of net probability fluxes still leads to a positive dissipation, which then follows the scaling law reported in the main text.

To illustrate this point, we consider another parameter combination $\gamma_1 = \gamma_2 = 0.2$, which does not lead to oscillation (Fig. S19, left panel). Subjecting the system to the same coarse-graining process reveals a dissipation scaling relation (Fig. S19, middle panel) that's qualitatively the same as that for the oscillating Brusselator (Fig. S17, right panel). These results highlight the applicability of the scaling law to general (not necessarily oscillating) nonequilibrium systems, especially open biochemical reaction systems.

d. Connecting mesoscopic probability fluxes with microscopic dissipation. Nonequilibrium systems are characterized by broken detailed balance, which can be experimentally quantified by measuring the probability fluxes through chemical states at the mesoscopic scale [15, 16]. However, the results reported here indicate that the dissipation calculated from mesoscopic fluxes will be an underestimate of the real dissipation, which should in principle be calculated from the microscopic fluxes. While such microscopic measurements might not be experimentally feasible (e.g. due to the uncertainty in the number of molecules n_x and n_y), our scaling law provides a useful tool for scaling the dissipation from the mesoscopic level back to the microscopic level in a nonequilibrium system.

B. Microtubule-kinesin active flow system

We consider an active matter mixture composed of microtubule filaments and kinesin motors. Despite the macroscopic nematic order, individual MT filaments are polar entities, which can be transported by kinesin only in one direction. In this section, we construct a kinetic model of the MT-kinesin mixture that falls within the scope of the random flux model and demonstrate that the scaling law reported in the main text holds regardless the details of the model.

a. The microscopic model. Inspired by the passive scalar behavior in turbulence [17], we study the dynamics of a single microtubule filament as a proxy for the dynamics at the population level. The MT filament is a polar object that undergoes two processes: rotation and transport. For illustrative purposes, we focus on the 2D model, which can be easily generalized to 3D. The microscopic model and its coarse-graining are demonstrated in Fig. S20. Essentially, we discretize the 2D space into $N \times N$ states on a square lattice and the polarity of microtubule into four discrete orientations. Fig. S20A illustrates the dynamics on a single spatial location. The four colors encode four discrete microscopic states with orientations left (yellow), right (red), up (green), and down (blue). Within each spatial locations, the MT can change its orientation by 90 degrees with the switching rate ω (for example, red to green, red to blue). However, the MT cannot reverse its orientation in one step (red cannot directly go to yellow). The MT can also be transported along its orientation with rates $k_{i,j,p}$, where $p = \pm 1, \pm 2, \pm 3, \pm 4$ encodes the MT

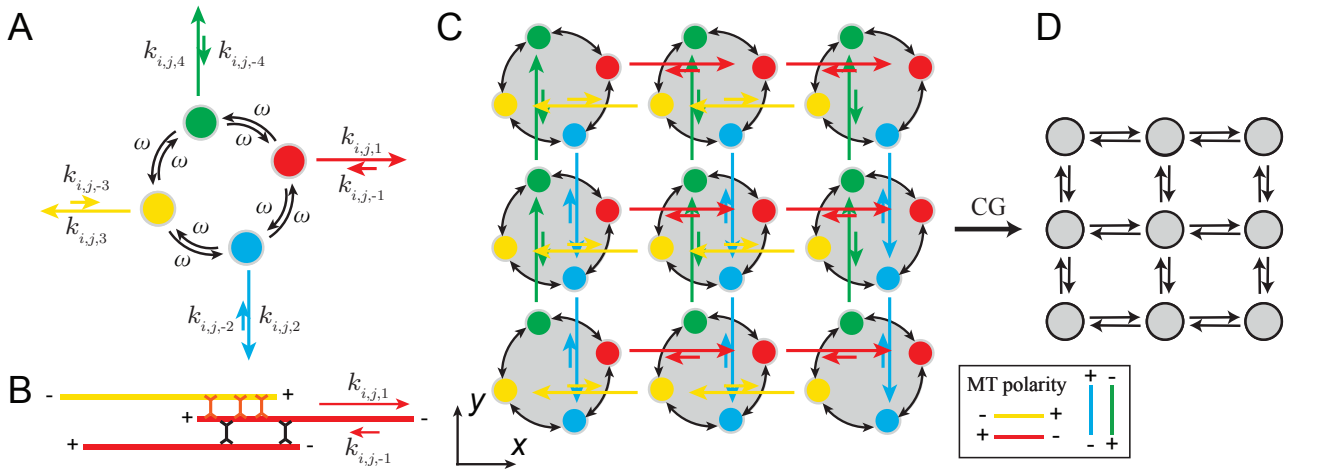

FIG. S20. Illustration of the MT-kinesin mixture model. (A) the microscopic model of a single spatial location. The four colors (red, yellow, green, blue) correspond to four spatial orientations of the microtubule. The microtubule can change its orientation by switching to a neighboring orientational state with rate ω (black arrows), or be transported by kinesin motors along its orientation to an adjacent spatial location (colored arrows). (B) nonequilibrium transport of microtubules by kinesin. The orange kinesins are attached to microtubules with different polarities and contribute to the active transport. The black kinesin is attached to microtubules with the same polarity and does not contribute to their active transport. For the microtubule in the middle, the number of kinesin motors facilitating the transport to the right is $n_{i,j,1} = 3$. The forward and backward rates $k_{i,j,\pm 1}$ are given in the text. (C) the microscopic model for 9 adjacent spatial locations (each denoted by a gray-shaded circle with bidirectional links representing rotation). The orientation of microtubules is preserved after transport. (D) the model after the first coarse-graining process which integrates out the orientational degree of freedom. The states on the same circle in (C) are now combined into a single state. The network structure is identical to the square lattice reported in the main text, and the subsequent coarse-graining processes preserve self-similarity.

polarity. The forward transport rates along direction p ($p = 1, 2, 3, 4$) are chosen as

$$k_{i,j,p} = k_d + \tilde{k}_{i,j,p}, \tag{S51}$$

where k_d corresponds to the rate for thermal diffusive transport and the term $\tilde{k}_{i,j,p} = k_0 n_{i,j,p}$ corresponds to active transport facilitated by kinesin motors. k_0 is a coefficient that's proportional to the concentration of ATP. $n_{i,j,p}$ denotes the number of kinesin motors facilitating the transport along the p direction at site (i, j). For thermodynamic consistency, the backward transport rates are given by

$$k_{i,j,-p} = k_d + \tilde{k}_{i,j,p} \exp(-n_{i,j,p} \Delta\mu_0), \tag{S52}$$

where $\Delta\mu_0$ is the free energy dissipation in ATP hydrolysis, which is coupled with kinesin stepping. Only the kinesin motors between microtubules of opposite polarities contribute to the active transport. Fig. S20B illustrates the details of the transport process which involves kinesin motors (orange and black linkers) walking on MTs (red and yellow segments). Following the denotation in Fig. S20A, the red segments and yellow segment have the opposite polarity. For the red MT in the middle, only orange motors contribute to the active transport, namely $n_{i,j,1} = 3$. The number of motors follows a Gaussian distribution $n_{i,j,p} \sim \mathcal{N}(\mu, \sigma)$. It is nonetheless not necessary for the rates to take this form in order to have the scaling behavior of dissipation rates since a previous SI section has demonstrated that the scaling behavior does not depend on the exact form of the rate distribution.

Fig. S20C illustrates the full dynamics of rotation and transport between a few neighboring spatial locations. Note that the transport does not change the MT orientation. The central red state, for example, corresponds to a MT oriented to the right, which can only be transported to another red state to the left or right.

To sum up, the MT dynamics is characterized by the dynamics a reaction network with $N \times N \times 4$ microscopic states, whose dynamics and energy dissipation rate are fully described by the chemical master equation formalism described in the main text.

b. Coarse-graining and dissipation scaling. Next, we consider the coarse-graining of this model and elucidate the dissipation scaling behavior. The first coarse-graining operation effectively merge the four orientational states within the same spatial location (Fig. S20C→D). The resulting network is exactly the regular lattice network studied in the main text. The subsequent coarse-graining operations correspond to the merger of spatially neighboring states in the exact same way as Fig. 2A of main text. Fig. S21 plots the dissipation rate at different coarse-grained levels for three

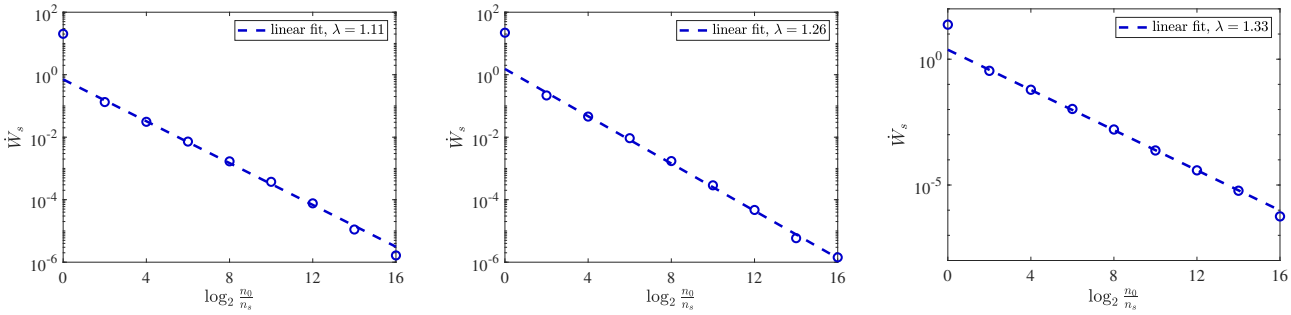

FIG. S21. The scaling of dissipation rate in the MT-kinesin model with different switching rates: $\omega = 0.1$ (left), $\omega = 1$ (middle), and $\omega = 10$ (right). The parameters are: $k_0 = 1$, $k_d = 0.1$, and $\mu = 6$, $\sigma = 2$ for the Gaussian distribution of the number of motors. These parameters and $\omega = 10$ are used for Fig. 3D in the main text.

different switching rates ω. Indeed, they exhibit a power-law scaling behavior with the exception of the first point. This is consistent with our theory since the first coarse-graining operation changes the topology of the network by merging the four orientation states, while the subsequent coarse-graining procedures preserve the self-similarity of the network (square lattice). The scaling exponent λ depends on the switching rate ω. For sufficiently large ω, the four orientation states achieve fast equilibrium and can be considered equivalent. In other words, a quarter of the MT point towards each of the four directions at the population level. In this limit, the orientation degree of freedom becomes trivial. The scaling exponent approaches that of the random flux model reported in the main text (Fig. S21, right panel). The scaling exponent λ is smaller for smaller ω.

c. Correlation in $n_{i,j,p}$ does not change the scaling behavior. In the realistic MT-kinesin system, both the number of kinesins $n_{i,j,p}$ and the MT density (probability density of discrete states $P_{i,j,p}$) are spatially correlated. Here, we demonstrate that the spatial correlation does not affect the existence of a dissipation scaling relation. In particular, the correlation is introduced by $n_{i,j,p} = s_{i,j,p} n_0$, where n_0 is a constant and $s_{i,j,p}$ is taken from a snapshot of an equilibrium Ising model on an $N \times N$ lattice. The number of kinesins transporting MT in different directions are not correlated (i.e. $\langle n_{i,j,p_1} n_{i,j,p_2} \rangle - \langle n_{i,j,p_1} \rangle \langle n_{i,j,p_2} \rangle = 0$ if $p_1 \neq p_2$). Fig. S22 demonstrates that the scaling relation exists for systems with different correlation lengths of $n_{i,j,p}$.

To show that our model reproduces the experimentally observed MT dynamics (alignment and spatially correlated active flow), we present the typical orientation field and coarse-grained flow field of our model in Fig. S23. For each spatial location (i, j), the probability of MT being in horizontal and vertical orientations are $P_{i,j,1} + P_{i,j,3}$ and $P_{i,j,2} + P_{i,j,4}$. The alignment ordering is visualized by coloring each location to the orientation whose probability is larger than 50%. (Namely, the site is colored yellow if $P_{i,j,1} + P_{i,j,3} > P_{i,j,2} + P_{i,j,4}$ and blue if otherwise.) As shown in the left panel of Fig. S23, the orientation of microtubules are spatially correlated, forming the nematic ordering that was experimentally observed [18]. The flow of microtubules is characterized by the probability flow in our system, which is visualized in the right panel of Fig. S23. The probability flow is spatially correlated and resembles the spatial MT flow that was experimentally observed [18].

The above model can be easily generalized to 3D, where each spatial location splits to 6 microscopic state. The analysis here indicates that the dissipation scaling behavior after the first coarse-graining (merger of orientation states) should resemble that of the cubic lattice. The analysis related to the 3D experiment system is presented in the main text.

IV. HIDDEN FREE ENERGY COSTS IN THE COARSE-GRAINED NETWORK

In this section, we elucidate the physical origin of the difference between dissipation at fine-grained and coarse-grained levels of the network. For the sake of illustration, let's first consider the regular lattice in Fig. 1A of the main text. Let $\dot{W}_0$ be the dissipation rate at the finest-grained level. This dissipation rate can be decomposed into three components

$$\dot{W}_0 = \dot{W}_0^{(1)} + \dot{W}_0^{(2)} + \dot{W}_0^{(3)}, \tag{S53}$$

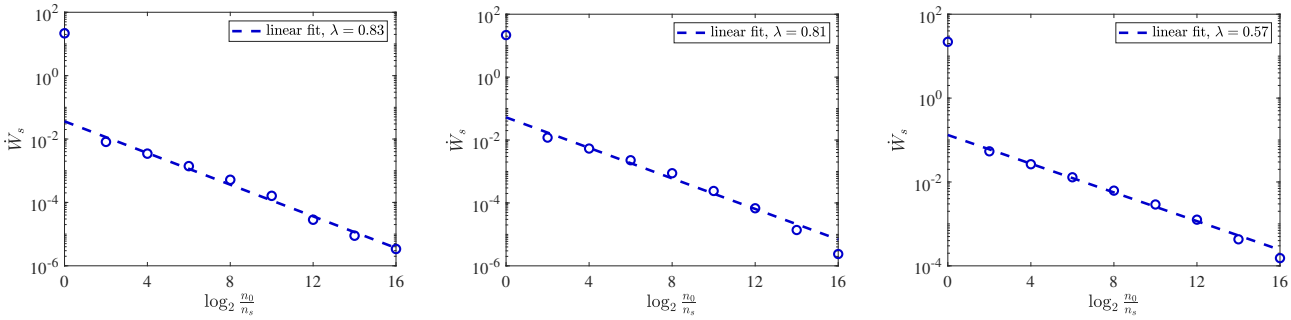

FIG. S22. The scaling of dissipation rate in the MT-kinesin model where the kinesin numbers $n_{i,j,p}$ are spatially correlated. The correlation lengths are $\xi \sim 2.6$ (left), $\xi \sim 4.7$ (middle), $\xi \sim 24$ (right). The parameters are: $k_0 = 1$, $k_d = 0.1$, $\omega = 0.01$, and $\mu = 6$, $\sigma = 2$ for the Gaussian distribution of the number of motors.

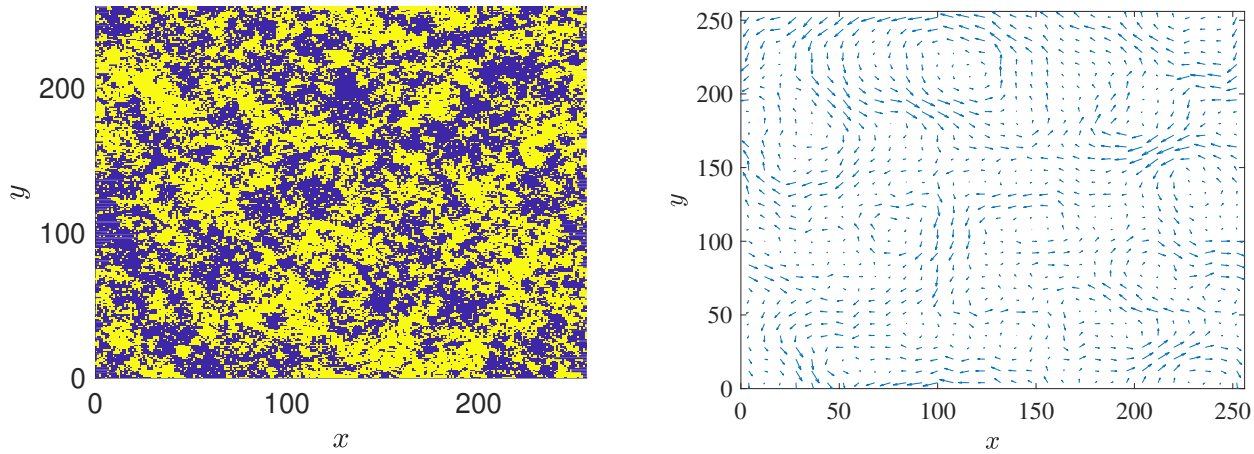

FIG. S23. Simulation results where the number of kinesins $n_{i,j,p}$ is spatially correlated (correlation length $\xi \sim 30$). Left: The stochastic orientation field of the MT-kinesin model. Yellow and blue denote horizontal and vertical orientations, respectively. Right: The coarse-grained (averaged in 8×8 squares) fluxes form a spatially correlated pattern.

all of which are nonnegative [11]. The first term $\dot{W}_0^{(1)}$ represents the dissipation associated with the coarse-grained transitions between CG states, which is exactly the total dissipation rate of the first CG level $\dot{W}_1$:

$$\dot{W}_0^{(1)} = \dot{W}_1 = 4^{-\lambda}\dot{W}_0 = \frac{1+C^*}{4}\dot{W}_0. \tag{S54}$$

The second term $\dot{W}_0^{(2)}$ is the dissipation caused by transitions within each CG state, which necessary for sustaining the internal nonequilibrium steady state. In the homogeneous lattice we study, this is proportional to the number of internal links:

$$\dot{W}_0^{(2)} = \frac{L_{\text{internal}}}{L_{\text{total}}}\dot{W}_0 = \frac{1}{2}\dot{W}_0, \tag{S55}$$

where $L_{\text{internal}} = N^2$ is the number of internal links and and $L_{\text{total}} = 2N^2$ is the total number of links in a regular lattice of size N.

The third term $\dot{W}_0^{(3)}$ is associated with merging multiple (in this case, two) reaction pathways into one during coarse-graining. As shown in the example Fig. 1A in the main text, the two pathways $i_2 \leftrightarrow j_1$ and $i_4 \leftrightarrow j_3$ are merged into a single reaction $i \leftrightarrow j$. The randomness in which of the two microscopic transitions actually happened when a (macroscopic) transition between CG states i and j is observed contributes to this extra term of dissipation. Concretely, the total dissipation rate in the two microscopic pathways is

$$\sigma_0(i,j) = (J_{i_2,j_1} - J_{j_1,i_2}) \ln \frac{J_{i_2,j_1}}{J_{j_1,i_2}} + (J_{i_4,j_3} - J_{j_3,i_4}) \ln \frac{J_{i_4,j_3}}{J_{j_3,i_4}}. \tag{S56}$$

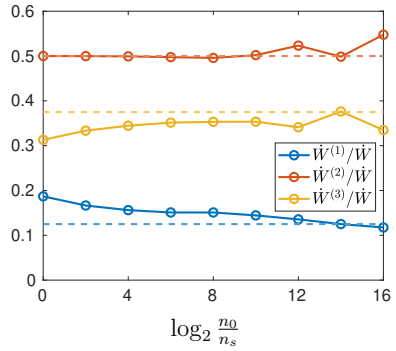

FIG. S24. The partitioning ratios $\dot{W}^{(i)}/\dot{W}$ $(i = 1, 2, 3)$ in the square lattice. The circles are data points from simulation and the horizontal dashed lines are their expected values at the RG fixed point: blue, $\dot{W}^{(1)}/\dot{W} \to \frac{1+C^\star}{4} = 0.125$ (Eq. S54); orange, $\dot{W}^{(2)}/\dot{W} \to \frac{1}{2}$ (Eq. S55); yellow, $\dot{W}^{(3)}/\dot{W} \to \frac{1-C^\star}{4} = 0.375$ (Eq. S61). The deviation of the last few points of $\dot{W}^{(2)}$ should be attributed to the fluctuation due to the small system size.

The total dissipation for the coarse-grained pathway is

$$\sigma_1(i,j) = (J_{i_2,j_1} + J_{i_4,j_3} - J_{j_1,i_2} - J_{j_3,i_4}) \ln \frac{J_{i_2,j_1} + J_{i_4,j_3}}{J_{j_1,i_2} + J_{j_3,i_4}}. \tag{S57}$$

Their difference leads to $\dot{W}_0^{(3)}$. Following the same method in section I.A.1 of the SI, we decompose the fluxes into symmetric and antisymmetric components to determine the average ratio of σ_1 to σ_0, which reads

$$\left\langle \frac{\sigma_1(i,j)}{\sigma_0(i,j)} \right\rangle = \frac{1}{2} \left\langle \frac{(A_{i_2,j_1} + A_{i_4,j_3})^2}{A_{i_2,j_1}^2 + A_{i_4,j_3}^2} \right\rangle = \frac{1 + C^\star}{2} \tag{S58}$$

Since $C^\star < 0$, the dissipation in the coarse-grained pathway is always smaller than the sum of dissipation in the two microscopic pathways. Their difference $(\sigma_0(i,j) - \sigma_1(i,j))$ makes up the third term $\dot{W}_0^{(3)}$. We sum up such terms over the entire network:

$$\dot{W}_0^{(3)} = \sum_{i<j} (\sigma_0(i,j) - \sigma_1(i,j)). \tag{S59}$$

On the other hand, the total dissipation rate of microscopic transitions between different CG states is

$$\dot{W}_0 - \dot{W}_0^{(2)} = \frac{1}{2}\dot{W}_0 = \sum_{i<j} \sigma_0(i,j). \tag{S60}$$

Therefore $\dot{W}_0^{(3)}$ can be simplified to

$$\dot{W}_0^{(3)} = \frac{1}{2}\dot{W}_0 \left(1 - \left\langle \frac{\sigma_1(i,j)}{\sigma_0(i,j)} \right\rangle\right) = \frac{1 - C^\star}{4}\dot{W}_0. \tag{S61}$$

It can be verified that the sum of the three components in Eq. S54, S55 and S61 exactly makes up the total dissipation $\dot{W}_0$ (Eq. S53).

The decomposition suggests that the dissipation that is missing at the coarse-grained levels consists of two components: $\dot{W}^{(2)}$, which arises from ignoring transitions within coarse-grained states, and $\dot{W}^{(3)}$, which stems from merging the transition pathways between coarse-grained states. The former maintains the internal nonequilibrium state of CG states, and the latter is associated with the multiplicity of ways in which a macroscopic transition can occur. Each of the microscopic pathways has a different nonequilibrium driving force γ.

The partitioning ratios $\dot{W}^{(i)}/\dot{W}$ $(i = 1, 2, 3)$ can also be numerically determined for comparison with the analytic results in Eq. S54, S55 and S61. The result is presented in Fig. S24, which shows that the ratios converge to their predicted value (dashed lines) as the system is coarse-grained.

REFERENCES

[1] T. L. Hill, *Free energy transduction in biology* (Academic Press, 1977).
[2] C. Song, S. Havlin, and H. A. Makse, Self-similarity of complex networks, Nature **433**, 392 (2005).
[3] C. Song, L. K. Gallos, S. Havlin, and H. A. Makse, How to calculate the fractal dimension of a complex network: the box covering algorithm, J. Stat. Mech. Theory Exp. **2007**, P03006 (2007).
[4] J. S. Kim, K.-I. Goh, B. Kahng, and D. Kim, Fractality and self-similarity in scale-free networks, New J. Phys. **9**, 177 (2007).
[5] L. Fenton, The sum of log-normal probability distributions in scatter transmission systems, IRE Trans. Commun. Syst. **8**, 57 (1960).
[6] R. L. Mitchell, Permanence of the log-normal distribution, J. Opt. Soc. Am. **58**, 1267 (1968).
[7] A. F. Rozenfeld, R. Cohen, D. ben Avraham, and S. Havlin, Scale-free networks on lattices, Phys. Rev. Lett. **89**, 218701 (2002).
[8] B. J. Kim, Geographical coarse graining of complex networks, Phys. Rev. Lett. **93**, 168701 (2004).
[9] P. Erdős and A. Rényi, On random graphs, Publ. Math. (Debrecen) **6**, 290 (1959).
[10] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature **393**, 440 (1998).
[11] M. Esposito, Stochastic thermodynamics under coarse graining, Phys. Rev. E **85**, 041125 (2012).
[12] G. Nicolis and I. Prigogine, *Self-organization in non-equilibrium systems* (Wiley, New York, 1977).
[13] H. Qian, S. Saffarian, and E. L. Elson, Concentration fluctuations in a mesoscopic oscillating chemical reaction system, Proc. Natl. Acad. Sci. U. S. A. **99**, 10376 (2002).
[14] C. Fei, Y. Cao, Q. Ouyang, and Y. Tu, Design principles for enhancing phase sensitivity and suppressing phase fluctuations simultaneously in biochemical oscillatory systems, Nat. Commun. **9**, 1434 (2018).
[15] C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, and F. C. Mackintosh, Broken detailed balance at mesoscopic scales in active biological systems, Science **352**, 604 (2016).
[16] F. S. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, Broken detailed balance and non-equilibrium dynamics in living systems: a review, Rep. Prog. Phys. **81**, 066601 (2018).
[17] Z. Warhaft, Passive Scalars in Turbulent Flows, Annu. Rev. Fluid Mech. **32**, 203 (2000).
[18] T. Sanchez, D. T. N. Chen, S. J. DeCamp, M. Heymann, and Z. Dogic, Spontaneous motion in hierarchically assembled active matter, Nature **491**, 431 (2012).