Dynamic disorder and the energetic costs of information transduction

Peter Thill

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA

(Received 19 February 2014; accepted 17 June 2014; published online 3 July 2014)

We study a model of dynamic disorder relevant for signal transduction pathways in which enzymatic reaction rates fluctuate over several orders of magnitude. For the simple networks we consider, dynamic disorder drives the system far from equilibrium and imposes an energetic burden for high fidelity signaling capability. We study how the dynamics of the underlying stochastic behavior in the reaction rate process is related to the energetic cost of transmitting information through the network.

© 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4885725>]

I. INTRODUCTION

Single molecule experiments have given scientists unprecedented capabilities to look beyond the static picture of enzymology obtained from ensemble averaged experiments.^{1,2} Scientists have been able to demonstrate that enzymatic rates of catalysis sometimes fluctuate over several orders of magnitude, drawing from heavy-tailed distributions of catalytic activity.^{3,4} In the language of statistical physics, a stochastic process with a fluctuating rate is known as “dynamic disorder.”⁵ Much theoretical work has been done to examine models of dynamic disorder that involve discrete transitions between finitely many rates.^{6–9} A commonly used experimental tool to probe the internal states of single molecules and the transition rates between various states is single molecule FRET spectroscopy (Förster resonance energy transfer).^{10–12}

Chemical reaction networks serve a number of important roles in biological systems, including the problem of signal transduction. Given a signal or input, a biological network must convert this signal to a downstream response. The reaction network can be seen as transmitting mutual information between the signaling molecules that are stimulated first and a downstream response molecule. Due to the stochastic nature of chemical kinetics, information processing will always have some associated error, even without dynamic disorder. However, dynamic disorder can further change the statistics of downstream signaling molecules, and thus fundamentally alter the fidelity of information a network is capable of transmitting.

It is clear that dynamic disorder can drive a system out of equilibrium. A well known consequence of Landauer’s principle is that irreversible computation requires energy consumption.¹³ For chemical reaction networks, how does dynamic disorder influence the energetic requirements (metabolic costs) for signal transduction? In this work, we seek to understand the implications that dynamic disorder can have on the fundamental capacity for biochemical networks to convey information. We will show that there are energetic costs associated with the intrinsic “switch rate” for an enzyme. In other words, when an enzyme transitions very rapidly between internal states thus “self-averaging,” the maximal information between signal and response is

achieved. However, this rapid switching drives the chemical network very far from equilibrium, necessitating greater energy consumption, as measured by the entropy production. Importantly, this energy consumption cannot be explained merely by thermal energy necessary for the enzyme to make conformational transitions.

The paper is organized as follows. In Sec. II, we analyze the limits of fast and slow switching between internal enzyme states. In Sec. III, we review entropy production in non-equilibrium steady states (NESS) and analyze the entropy production for birth death processes in the presence of dynamic disorder. In Sec. IV, we look at a model of a single enzyme with two internal states, for which entropy production and the PDF (probability distribution function) can be solved exactly to illustrate the properties in the first three sections. Finally, we study the information content in network motifs and define the capacity power consumption. We look at two specific examples. The first is a simple birth death network with multiple enzymes each of which has three internal states. The second is a more complicated network known as the Coherent Feedforward Type 1 network, which has been used as a model for the arabinose gene transcription network of *E. coli*.¹⁶ Here, we allow the enzyme rates to draw from various continuous distributions. We show that for these networks, dynamic disorder leads to a trade-off in which maximal information capacity of the network is associated with the greatest energetic costs.

II. EFFECTS OF LIMITING VALUES OF THE SWITCHING RATE

Conventional stochastic chemical kinetics involves the use of a Gillespie algorithm to sample the master equation.^{18–20} For a system of chemical reactions involving K chemical species $\{X_i\}_{i=1}^K$, we let $\mathbf{X}$ represent the vector of numbers of each molecular species. A reaction indexed by ρ changes the vector $\mathbf{X}$ by some stoichiometric vector $\mathbf{n}_\rho$, via $\mathbf{X} \xrightarrow{\rho} \mathbf{X} + \mathbf{n}_\rho$. We assign to each reaction ρ a transition rate $W_\rho(\mathbf{X}|\mathbf{X} + \mathbf{n}_\rho)$, also known in chemical kinetics as a reaction propensity. Transition rates for chemical reactions are given by the product of a fundamental rate constant (denoted by k_ρ) and a combinatorial factor accounting for all the possible

ways the reaction can occur (assuming homogeneously mixed concentrations).

We will consider here a simplified picture of dynamic disorder. We will assume that some of the rate constants k_p are allowed to fluctuate by drawing a value from a prescribed distribution, remaining at this value for an exponentially distributed amount of time as governed by a parameter we will denote λ (the switch rate, with units of s^{-1}), and then redrawing a new value from the same distribution. We note that this is a special case of dynamic disorder in which the transition probability between internal states $i \rightarrow j$ depends only on state j . In general, one can consider that the transition probability for transitions between various internal states is a more complicated function of both the initial state and the final state. However, the qualitative results presented here are likely to hold for more complicated transition probabilities, as we only make arguments appealing to the stationary distribution of the internal enzyme state, or the rate at which the enzyme reaches steady state. (Though we consider the case in which all internal states of the enzyme are interconnected, many have studied how to infer more complicated topologies connecting internal enzyme states from single molecule studies.^{21–24})

As an explicit example, consider the following bimolecular reaction, which may represent phosphorylation of a species Y :

The transition rate for the above reaction is given by $k[X][Y]$. If the rate parameter k changes due to fluctuating structural conformations of the species X , the transition rate will become a stochastic variable. At a particular instant in time, the transition rate will be given by $(\sum_i k_i)[Y]$, where the index i counts the individual rate constants k_i for each molecule of X in the system. We will denote by K the “total rate,” which is the sum of all individual rate constants k_i , such that the transition rate for the above reaction becomes $K[Y]$. We will assume in this work that each molecule will have an individual rate that is independent of the others (i.e., switching events are uncorrelated) but each individual rate draws from the same distribution.

To quantify how fast or slow the switch rate λ is, we assume that there exists some intrinsic rate constant in the network (such as the decay or dephosphorylation rate of active species) that sets the time scales of reaching the stationary distribution. Throughout the paper, we will refer to this parameter as α . To speak about the slow or fast switching limit, we must consider cases in which λ/α is much smaller or much greater than one.

For general systems of chemical reactions, the stationary probability distribution $P(\mathbf{X})$ can be obtained by solving the master equation. Here, we examine how the stationary probability distribution is modified in the limits of $\lambda \rightarrow 0, \infty$. For simplicity, we will first consider the case in which the copy number of the species undergoing conformation fluctuations is constant throughout the course of the reaction. We will later investigate network schemes for which this is not the case. We will denote by $P(\mathbf{X}|K)$ the stationary probability of $\mathbf{X}$ given a system with no dynamic disorder and a total rate given by K .

In the limit that $\lambda \rightarrow 0$, each individual rate k_i remains at a certain value for increasingly long amounts of time. In this limit, we assume that the total rate changes very slowly, and the chemical species $\mathbf{X}$ can adiabatically follow the underlying rate parameter and equilibrate to the current value. At any given time, the distribution of $\mathbf{X}$ should look like $P(\mathbf{X}|K)$. Thus, the stationary distribution over $\mathbf{X}$ measured from a long trajectory in this limit should look like marginalization

$$\lim_{\lambda \rightarrow 0} P(\mathbf{X}) = \int P(\mathbf{X}|K)P(K)dK, \quad (2)$$

where $P(K)$ is the probability that $\sum_i k_i = K$. In general, $P(K)$ will not be analytically tractable, but can be approximated by a normal random variable in the limit of large copy number: $K \approx \mathcal{N}(N\mathbb{E}[k], N\text{Var}[k])$. We have let N represent the total number of the fluctuating chemical species, and $\mathbb{E}[k]$, $\text{Var}(k)$ denote statistics of the distribution that each k_i draws from.

In the limit of $\lambda \rightarrow \infty$, we argue that

$$\lim_{\lambda \rightarrow \infty} P(\mathbf{X}) = P(\mathbf{X}|\mathbb{E}[K]). \quad (3)$$

To see this, we must consider a generic time interval $[t_1, t_2]$ between successive reaction events. For a general time dependent parameter $K(t)$, the time-dependent propensity is proportional to $\int_{t_1}^{t_2} K(t)dt$.¹⁴ When each individual rate constant switches at discrete times, the total rate constant K will take on a discrete set of values between these two time points. The total number of switching events in this period, Ω , is Poisson distributed with parameter $N\lambda(t_2 - t_1)$. Further, the expected time interval dt in between each switching event is of length $(N\lambda)^{-1}$. If we index the intervals between each switching event by i , we can rewrite the time dependent propensity as follows (dropping the combinatoric pre factor):

$$\int_{t_1}^{t_2} K(t)dt = \sum_{i=1}^{\Omega} K(t_i)dt_i \xrightarrow{\lambda \rightarrow \infty} (t_2 - t_1)\mathbb{E}[K]. \quad (4)$$

To determine the limiting form, we have used the Renewal Reward theorem²⁸ (Appendix A). This analysis demonstrates that as the switching rate becomes very large, the time-dependent propensity between any two reaction events will look like the average propensity.

III. POWER CONSUMPTION IN MARKOV CHAINS

The entropy production for a NESS can be computed for a general Markov chain with states indexed by σ and transition rates between the σ given by $W(\sigma, \sigma')$ ^{25,26}

$$\frac{dS}{dt} = \sum_{\sigma, \sigma'} P(\sigma)W(\sigma, \sigma') \log \left(\frac{W(\sigma, \sigma')}{W(\sigma', \sigma)} \right). \quad (5)$$

For certain birth death processes, a simple argument appealing to Jensen’s inequality motivates why the model for dynamic disorder we have considered should result in increased entropy production in the fast switch rate limit. Assume the rate of production is the fluctuating rate K ($W(n, n+1) = K$) and the rate of destruction is $W(n, n-1) = n\alpha$. The entropy production for the process in which there is no dynamic disorder and the parameter K acts only at some average value $\mathbb{E}[K]$

is given by

$$\frac{dS}{dt} = \sum_n P(n|\mathbb{E}[K]) \mathbb{E}[K] \log \left(\frac{\mathbb{E}[K]}{\alpha(n+1)} \right) + \sum_n P(n|\mathbb{E}[K]) \alpha n \log \left(\frac{\alpha n}{\mathbb{E}[K]} \right). \quad (6)$$

In the presence of dynamic disorder and a fast switch rate limit, the entropy production becomes

$$\frac{dS}{dt} = \sum_n P(n|\mathbb{E}[K]) \int P(K) K \log \left(\frac{K}{\alpha(n+1)} \right) dK + \sum_n P(n|\mathbb{E}[K]) \int P(K) \alpha n \log \left(\frac{\alpha n}{K} \right) dK. \quad (7)$$

By the convexity of the functions $x \log x$ and $\log(1/x)$, and invoking Jensen's inequality, the entropy production in Eq. (7) will be greater than in Eq. (6). What's more, because the internal rate switching will equilibrate and the distributions reach a limiting form with increasing switch rate, the entropy production in Eq. (7) represents an asymptotic cost incurred by infinitely fast switching rate. This demonstrates that there is a finite price to pay even as the switching rate increases to infinitely large values.

In many cases, computing the entropy production is very hard, as it requires enumeration of all possible states and calculation of (or Monte Carlo sampling to determine) the steady state probability for the system to be in each state. However, one can numerically compute the irreversible entropy production for a NESS by looking at the average value of the fluctuating quantity $Z(t)$ defined by Gaspard²⁷ as the log ratio of the probability for a trajectory to the probability of its time-reversed trajectory. On average, this quantity grows linearly with a slope equal to the irreversible entropy or power consumption required to maintain the NESS. For a sequence of reactions inducing the trajectory $\mathbf{X}_0 \rightarrow \mathbf{X}_1 \rightarrow \dots \rightarrow \mathbf{X}_n$, the quantity $Z(t)$ is given by

$$Z(t) = \log \left(\frac{W(\mathbf{X}_0|\mathbf{X}_1)W(\mathbf{X}_1|\mathbf{X}_2) \dots W(\mathbf{X}_{n-1}|\mathbf{X}_n)}{W(\mathbf{X}_1|\mathbf{X}_0)W(\mathbf{X}_2|\mathbf{X}_1) \dots W(\mathbf{X}_n|\mathbf{X}_{n-1})} \right). \quad (8)$$

For the model of dynamic disorder considered in this work, the transition for an individual rate $k_i \rightarrow k'_i$ occurs with a rate given by the switch rate λ times the probability of drawing k'_i . Thus, an individual rate transition will contribute to $Z(t)$ a term of the form

$$\log \frac{p(k'_i)}{p(k_i)}. \quad (9)$$

We note that switch rates in expectation should contribute nothing to irreversible entropy production as the above term in expectation becomes $H(k) - H(k) = 0$ where $H(k)$ is the entropy of the probability distribution that a single enzyme draws from.

IV. SINGLE ENZYME, TWO INTERNAL STATES

To illustrate the relationship between the switch rate and the probability distribution and entropy production, we will look now at a simple example which can be solved exactly

for all values of the switch rate. This is the case of a simple birth death process with a rate of production that fluctuates between two internal states. This can be thought of as a single enzyme with two internal states catalyzing the production of an activated species

where k_i, k_j denote the two rates, and p_j is the steady state probability of the enzyme being in state j . The master equation for the joint PDF $p(n, i)$ where n is the number of Y and i is the internal state is as follows:

$$\frac{\partial}{\partial t} p(n, i) = k_i p(n-1, i) + \alpha(n+1) p(n+1, i) + \lambda p_i p(n, j) - (\alpha n + k_i + \lambda p_j) p(n, i). \quad (11)$$

This equation is mathematically equivalent to that studied by Mehta²⁶ for a model of receptor binding induced conformational change in a protein. It is solved by introducing a probability generating function

$$G_i(z) = \sum_{n=0}^{\infty} p(n, i) z^n \quad (12)$$

and the solution can be shown to be²⁶

$$G_i(z) = p_i e^{(k_i(z-1)/\alpha)} {}_1F_1 \left(\frac{\lambda p_i}{\alpha}, 1 + \frac{\lambda}{\alpha}, \frac{k_j - k_i}{\alpha} (z-1) \right), \quad (13)$$

where ${}_1F_1$ is the confluent hypergeometric function of the first kind.

We can see explicitly that the solution only depends on the ratio λ/α . We plot the probability distribution with $\alpha = 10$ and the two internal rates to be $k = 10, 100$ with probability 0.25, 0.75. In the absence of dynamic disorder, the simple birth death process has a stationary distribution that is Poisson with parameter k/α . In the fast switch rate limit, the distribution with dynamic disorder will approach Poisson with parameter $\mathbb{E}[k]/\alpha$, and in the slow switch rate limit, the distribution will approach a mixture of two Poissons, weighted by the stationary distributions. We show both limits (Figure 1(b)).

We also plot the entropy production associated with the system as a function of the switch rate (Figure 1(a)). In this case, it can be calculated exactly using

$$\frac{dS}{dt} = \sum_{n,i} P(n, i) k_i \log \left(\frac{k_i}{\alpha(n+1)} \right) + \sum_{n,i} P(n, i) \alpha n \log \left(\frac{\alpha n}{k_i} \right). \quad (14)$$

The entropy production should have a sigmoidal shape. For low values of switching rate, it asymptotically approaches a weighted average of entropy productions for two birth death processes, each of which equal zero. For high switch rate limits, it asymptotically approaches a value predicted by Eq. (7), where $P(n|\mathbb{E}[k])$ is Poisson with parameter $\mathbb{E}[k]/\alpha$. Different values of $\mathbb{E}[k]$ change the asymptotic value of the entropy production. In general, the specific topology of the signaling network will determine whether a greater $\mathbb{E}[k]$ will increase

FIG. 1. (a) Here, we plot the entropy production of the birth death process with 2 internal states, as a function of the switch rate between the states. The horizontal black lines are the values of the asymptotic entropy production in the fast switch rate limit as predicted by Eq. (7). Different values of the average rate parameter change the limiting value of entropy production, but not the shape of the curve itself. (b) Here, we plot the probability distribution over the number of Y species for the birth death process with two internal states. Shown for comparison in dark circles are the weighted mixture of Poissons and the Poisson with averaged K in the slow and fast switch rate limits, respectively.

or decrease the limiting value of entropy production. However, importantly, the asymptotic behavior of the curve is not influenced by the choice of $\mathbb{E}[k]$.

V. MULTIPLE ENZYMES WITH THREE INTERNAL STATES

A. Information in a chemical network

To quantify the ability of a signaling network to transduce signals, we will study the mutual information between a signal input and a response, $I(S; R)$.¹⁵ This measures the Kullback Leibler (KL) divergence between the joint distribution $P(R, S)$ and a completely independent distribution $P(R)P(S)$. The KL divergence is an information theoretic measure of the distance between two distributions. For the purposes of this paper, we will work with simplified pictures in which the signal species copy number can only take on a finite number of values, and the mutual information will quantify the ability to determine what the signal strength is based on our measurement of the response. We first look at the information in the birth death network

We allow the rate of production k to fluctuate, meaning the quantity $k[X]$ becomes $\sum_{i=1}^{n_x} k_i$ which we hereafter refer to as K (note that $\mathbb{E}[K] = \mathbb{E}[k_i]n_x$). We study the system using the Gillespie algorithm.^{18–20} For these calculations, the internal states of the enzyme are chosen to be 0.1, 1, 10. First, for fixed n_x , we verify that the probability distributions over Y tend to the limiting values predicted in Sec. II. The probability distribution in the absence of rate fluctuation is Poisson with parameter $k[X]/\alpha$, and we compare in the fast switch rate limit to $\mathbb{E}[k_i]n_x/\alpha$ for $\alpha = 10$ and multiple values of $\mathbb{E}[k_i]$ (Figure 2(a)).

Next, we study the information in this network by allowing X to take the values of 10 or 50. To analyze the information in the network, we must specify the prior probability that $X = 50$, which we will denote by p . We study the information $I(X; Y)$ for all priors by computing separate steady states for each value of n_x (Figure 2(c)). The information for this network is maximized in the fast switch rate limit, and decreases as the switch rate decreases.

B. Capacity power consumption

One way to quantify the mutual information is to examine the channel capacity. For a generic channel, the channel capacity is defined as the maximal mutual information between a signal and a response, where the maximization is done with respect to the prior distribution over the signal. In this case, the channel capacity is $I(p^*)$, where $p^* = \arg \max_p I(X; Z)$.

To quantify the relationship between energy consumption and information content, we can define the capacity power consumption as below

$$p^* \left. \frac{dS}{dt} \right|_{n_x=50} + (1 - p^*) \left. \frac{dS}{dt} \right|_{n_x=10}. \quad (16)$$

For a given p^* that optimizes the information content of a network, there is an associated energetic cost that corresponds to a weighted average of the entropy productions in each of the NESS of n_x . In this sense, the capacity power consumption measures the energetic cost incurred for a network to transmit at its capacity. It is important to note that in general, the value of p^* depends on the switch rate.

For the parameters considered here, the value of p^* is 0.49. For the case of multiple enzymes with three internal states, we see that the capacity entropy production increases with the switch rate, but saturates at the high switch rate limit, as predicted before (Figure 2(b)).

FIG. 2. (a) For $n_x = 50$, we plot the probability distribution of species Y in the birth death network. The curve labeled $E[K]$ indicates the Poisson distribution evaluated at the average K , for comparison with the fast switch rate limit. (b) The capacity power consumption for the birth death network with multiple enzymes shows the characteristic sigmoidal shape. In this case, larger values of $E[k]$ (expected rate for a single enzyme) lead to lower power consumption in the large switch rate limit. (c) The information in the birth death network increases with increasing switch rate, and reaches a limiting value predicted for the network in the absence of dynamic disorder and evaluated for the Poisson distribution at the average value of K , labeled here as $E[K]$.

VI. COHERENT FEEDFORWARD TYPE 1 NETWORK

A. Model and probability distributions

We now test these results in the context of a more complicated network scheme, and using various choices of continuous distribution for the fluctuating rate parameter. The network we will analyze here will be the Coherent Feedforward Type 1 (CFF1) network.

This particular network has been used as a model of gene transcription of arabinose in *E. coli*.¹⁶ The network is chosen to make the analysis tractable. The reactions we consider are

An interesting property of the CFF1 network is its ability to filter noise.¹⁶ The reactions above are chosen such that X is constant throughout the course of a trajectory. We first wish to demonstrate that the two limits of the switch rate are captured by marginalization and expectation over the random total rate. We will study the effects of allowing k_1 to fluctuate. As such, the combination $k_1[X]$ is replaced by $\sum_{i=1}^{n_x} k_i$, hereafter denoted by K . We study the system using the Gillespie algorithm.^{18–20} We will do the analysis for various choices of distribution over the k_i . To ensure that rates are always positive, we first study the system using a distribution that is proportional to the positive part of a normal PDF. In other words, $k_i \sim \mathcal{N}(0, 1)|(k_i > 0)$. Toward a better analytical understanding of the network, we first write down a system size expansion (Appendix B) to obtain an approximate stationary multivariate normal distribution $P(Y, Z)$.¹⁷ We compare the proba-

bility distribution over Z in the fast and slow switch rate limits, and analytically compute the marginal distribution over Z by using the system size expansion to find $P(Z|K)$ and integrating out K assuming that the copy number n_x is large enough that K is approximately normal (Figure 3(a)).

To show that the results hold for the case in which X fluctuates stochastically, we next analyze the network with the addition of the following two reactions:

In this case, we allow each X molecule to draw a new rate k_i when it is formed, and the rate then switches according to the above protocol. The slow switching limit is now set by the timescale of $1/\alpha_x$ since it should not matter if the rate k_i persists for longer than the lifetime of an individual X . We can approximately calculate the marginal distribution in the slow limit for the case of fluctuating X species by an adiabatic approximation, and by choosing the distribution of the k_i to be closed to addition. As such, we let the k_i draw from the chi-squared distribution with one degree of freedom (the distribution of the square of a normal random variable). The details are outlined in Appendix C. The results are shown in Figure 3(b).

B. Channel capacity and rate variance

We would like to understand the relationship between the capacity for a biochemical network to convey information, and the noisiness of the underlying distribution for the k_i (as measured by the variance).

We will analyze the Coherent Feedforward system as before treating X as the signal/input, and Z as the response/output. We will consider a two-state system where signal strength is high with probability p , and X can be either 500 or 100. In other words, $P(X = 500) = p, P(X = 100) = 1 - p$. For these calculations, we let the k_i draw from a

FIG. 3. (a) We simulate the coherent feedforward network for $n_x = 100$, $k_2 = 5$, $\alpha = 10$ and the distribution over the k_i is proportional to the standard $\mathcal{N}(0, 1)$ density when $k_i > 0$ and zero otherwise. We compute the distribution over downstream species Z for fast switch rate (50 s^{-1}) and slow switch rate (0.1 s^{-1}) limits. We analytically compute $P(z)$ by making a system size expansion to compute $P(Z|K)$ when K does not fluctuate and integrating out K by assuming $P(K)$ is approximately normal, with average $n_x \mathbb{E}[k_i]$ and variance $n_x \text{Var}[k_i]$. The slow switch rate limit approaches that computed in the SSE after marginalizing the stochastic rate parameter. (b) The coherent feedforward network is simulated with additional reactions to degrade and produce X with parameters $\alpha_x = 1$, $k_f = 500$. Gillespie simulations are performed for a switch rate of $0.1, 30 \text{ s}^{-1}$, and also for the network in the absence of dynamic disorder in which each k_i is equal to the average value in the corresponding dynamic disordered simulations (No fluc). We compute the marginalization of $P(z)$ using the distributions for $P(K)$, $P(Z|K)$ and performing the integration with the software package Mathematica. As the switch rate increases, it approaches the distribution predicted in the absence of rate fluctuation. As the switch rate slows, the distribution approaches that predicted by marginalizing the stochastic switch rate out of the distribution predicted by the SSE.

lognormal distribution. To adjust the variance while holding the mean constant requires adjusting both the μ and σ parameters of the lognormal distribution.

The graph of $I(X; Z)$ in Figure 4 as a function of p shows that the information content of the network decreases as the variance of the underlying rate distribution increases. This is consistent with the intuition that the noisier an individual parameter becomes, the harder it is to convey information. Similar metrics related to binary decision making such as the area under the receiver operating characteristic (ROC) curve show similar trends.

FIG. 4. Here, we plot the mutual information between X and Z as a function of p , the probability of high signal strength. Different curves represent different values of the variance in distribution for an individual rate constant. Increasing variance leads to less information transduction in the network.

C. Power consumption and channel capacity

From the approximate distribution $P(Y, Z)$, one can readily show that (Appendix B)

$$\text{Var}[Z|K] \propto K, \quad \mathbb{E}[Z|K] \propto K. \quad (19)$$

From the law of total variance and the functional form of the variance for this example, we know that

$$\text{Var}[Z] \geq \mathbb{E}[\text{Var}[Z|K]] = \text{Var}[Z|\mathbb{E}[K]]. \quad (20)$$

This suggests that in the slow and fast switch rate limits, the average value of Z is the same, and the distribution over Z will be roughly centered at the same place (if we assume a symmetric distribution in Z). However, in the slow switch rate limit, the width of the distribution over Z will be larger than in the fast switch rate limit. In our two state system, this means that peaks for low and high signal strength will be located in the same place regardless of the switch rate being fast or slow, but will begin to overlap to a greater degree as the switch rate becomes slower. This makes it more difficult to discriminate a weak or strong signal if the switching rate is slow.

In the information theoretic framework, a slower switch rate corresponds to an increasing overlap between distributions $P(Z|X = 500)$ and $P(Z|X = 100)$, and thus it appears that slow switching should correspond to lower mutual information between a signal and response.

We can analytically compute the information using the distribution $P(Y, Z)$ obtained from the system size expansion for the limits that $\lambda \rightarrow 0, \infty$. Intermediate values are computed using the Gillespie algorithm. Indeed, we see that

FIG. 5. (a) Plot of power consumption required for the NESS corresponding to the maximal information content of the network motif as a function of switch rate. Faster switch rates require more power consumption. (b) A plot of mutual information between X and Z as a function of p . Different curves represent different switch rates, and the curve labeled “No fluc” is calculated in the absence of dynamic disorder. High switch rates approach the limit of information predicted by the system size expansion (SSE) assuming the average value of the rate. The information evaluated at the expected value of the rate parameter and evaluated after marginalizing out the rate parameter assuming the central limit theorem (labeled SSE, Marg.) set bounds for the information.

increasing switch rate increases the information content of the network (Figure 5(b)).

At this point, as it is clear that a faster switch rate should correspond to greater information content, we would like to comment on the energetic costs associated with a faster switch rate. Macroscopically, it can be shown that the CFF1 network is at equilibrium at steady state (Appendix B), but does dynamic disorder drive the network away from equilibrium?

Given the distribution $P(Y, Z)$ for the CFF1 network considered above, the entropy production in the absence of dynamic disorder can be calculated easily. It is not exactly zero (as would be expected macroscopically) due to finite size effects, but is very close to equilibrium. For cases of rate fluctuation, we must compute this quantity numerically. Nonetheless, we can see that dynamic disorder drives the system away further from equilibrium, with faster switching rates driving the system further away; the capacity power consumption increases with the switch rate (see Figure 5(a)). For these parameters, the value of p^* that maximizes the mutual information is 0.5021 for all switch rates.

The enzyme transitions satisfy detailed balance, and indeed would even for a more complicated model in which the transition probabilities were more complex functions of the current and final state. This suggests that the energetic cost of this irreversibility which we document above cannot be explained by merely thermal energy required to allow the enzyme to transition between internal states.

VII. CONCLUSION

In this paper, we have analyzed a form of dynamic disorder in the context of signal transduction. We have analyzed the limits of the switching rate governing the time intervals in between transitions amongst rate values, and shown numerically and with analytical arguments that the infinitely fast

limit leads to a system which is identical to that in the absence of dynamic disorder with all rate parameters at their average value. As the switch rate tends to zero, the system looks as if we have marginalized or integrated out the stochastic rate parameter. We analyzed a birth death network with multiple enzymes and three internal states, and the Coherent Feedforward Type 1 network, and demonstrated that the fast switch rate is most conducive to information transduction, but also leads to the most expensive energetic costs.

In certain networks, noise is exploited by a phenomenon known as stochastic resonance.^{29–31} While it seems likely that in some contexts, dynamic disorder may reduce the signal to noise ratio that is characteristic of this behavior, in other cases where internal noise itself is used as the dynamical parameter that induces stochastic resonance, perhaps dynamic disorder can help facilitate noise induced coherent resonance. While the metrics considered in this paper (entropy production, channel capacity) varied monotonically with the switch rate and variance of the underlying rate distribution, it is worth examining whether or not in certain regimes known to exhibit stochastic resonance some of these metrics are non-monotonic. However, for the purposes of signal transduction considered in this work, it appears that dynamic disorder can only hinder the ability of a signaling network to convey information, while at the same time incurring added energetic costs.

In the limit of small copy number, any cellular network will be subject to stochastic noise. The probabilistic nature of chemical kinetics implies that any such network will have a fundamental limit in terms of its ability to transduce a signal reliably (quantified in this work by the mutual information between a signal and a response). However, this work suggests that for at least some signaling network motifs, even relatively simple ones, dynamic disorder serves to further decrease the channel capacity. What’s more, for an enzyme to self

average by making very rapid transitions between internal states seems to require greater energetic consumption. The asymptotic cost and channel capacity ultimately depend strongly on the parameters of the dynamic disorder, including the value of the average of the rate, the variance in the rate and the switching rate between internal states.

Ultimately, this implies that the evolution of certain network motifs required for cellular computation must take into account dynamic disorder to optimally transmit information.

ACKNOWLEDGMENTS

We thank Arup K. Chakraborty (A.K.C.) for helpful discussions. We would like to acknowledge financial support from P01 AI09158001 (awarded to A.K.C. and collaborators).

APPENDIX A: THE INFINITELY FAST SWITCHING LIMIT

Renewal theory involves processes for which inter-arrival times are IID. An important result is the Renewal Reward Theorem. A renewal reward process is a process for which a reward r_i is given in each inter-arrival interval dt_i . If we define $\Omega(T)$ to be the counter function which equals the number of arrivals up to time T , and define $R(T) = \sum_{i=1}^{\Omega(T)} r_i$, the Renewal Reward theorem states that, with probability one,

$$\lim_{T \rightarrow \infty} \frac{R(T)}{T} \rightarrow \frac{\mathbb{E}[r_i]}{\mathbb{E}[dt_i]}. \quad (\text{A1})$$

In the time-dependent propensity for the stochastic chemical network, the rewards are given by $r_i = K_i dt_i$. To use the theorem, we will define a related process denoted with $'$ such that $dt'_i = \lambda dt_i$, $T' = \lambda T$ and a counter function defined such that $\Omega'(T') = \Omega(T)$, such that the limit of $\lambda \rightarrow \infty$ implies $T' \rightarrow \infty$ and vice versa for finite T . Note that

$$\frac{R(T')}{T'} = \frac{\sum_{i=1}^{\Omega'(T')} K_i dt'_i}{T'} = \frac{\sum_{i=1}^{\Omega(T)} K_i dt_i}{T} = \frac{R(T)}{T}. \quad (\text{A2})$$

The theorem says that

$$\lim_{T' \rightarrow \infty} \frac{R(T')}{T'} = \lim_{\lambda \rightarrow \infty} \frac{R(T)}{T} = \mathbb{E}[K] \quad (\text{A3})$$

which is the result stated in the text.

APPENDIX B: SYSTEM SIZE EXPANSION OF CFF1 NETWORK

The master equation for the reactions considered can be written in terms of step operators $\mathbb{E}$ as

$$\dot{P} = \left\{ \begin{array}{l} (k_1 n_x [\mathbb{E}_y^{-1} - 1]) + (k_1 n_x + k_2 n_y) [\mathbb{E}_z^{-1} - 1] \\ + \alpha [\mathbb{E}_y - 1] n_y + \alpha [\mathbb{E}_z - 1] n_z \end{array} \right\} P. \quad (\text{B1})$$

Define macroscopic functions ϕ and stochastic functions ξ, η such that

$$\begin{aligned} n_y &= \Omega \phi_y(t) + \Omega^{1/2} \xi, \\ n_z &= \Omega \phi_z(t) + \Omega^{1/2} \eta, \end{aligned} \quad (\text{B2})$$

where Ω is the volume. Expanding the step operators ($\mathbb{E}_y = \exp(\Omega^{-1/2} \partial_\xi)$) to linear order in $1/\Omega$ yields the following macroscopic equations:

$$\begin{aligned} \dot{\phi}_y &= k_1 \phi_x - \alpha \phi_y, \\ \dot{\phi}_z &= k_1 \phi_x + k_2 \phi_y - \alpha \phi_z, \end{aligned} \quad (\text{B3})$$

and the following Fokker Planck equation for probability distribution $\Pi(\xi, \eta)$:

$$\begin{aligned} \frac{\partial \Pi}{\partial t} &= \alpha \frac{\partial}{\partial \xi} (\xi \Pi) + \frac{\partial}{\partial \eta} [(\alpha \eta - k_2 \xi) \Pi] + (k_1 \phi_x + \alpha \phi_y) \frac{1}{2} \frac{\partial^2 \Pi}{\partial \xi^2} \\ &+ (k_1 \phi_x + k_2 \phi_y + \alpha \phi_z) \frac{1}{2} \frac{\partial^2 \Pi}{\partial \eta^2}. \end{aligned} \quad (\text{B4})$$

The steady state macroscopic concentrations can easily be shown to balance fluxes for all reactions, and thus the macroscopic steady state is at equilibrium. The above allows us to compute stationary means and stationary variances, leading to the stated results

$$\begin{aligned} \mathbb{E}[Z] &= \frac{k_1 n_x}{\alpha} \left(1 + \frac{k_2}{\alpha} \right), \\ \text{Var}[Z] &= \frac{k_1 n_x}{\alpha} \left(1 + \frac{k_2}{\alpha} + \frac{k_2^2}{2\alpha^2} \right). \end{aligned} \quad (\text{B5})$$

APPENDIX C: MARGINAL DISTRIBUTION OF $P(Z)$

Here, we seek to find the marginal distribution of $P(Z)$ for the coherent feedforward type 1 network for comparison with the slow switch rate limit. We assume that $\mathbb{E}[K]$ is sufficiently large such that downstream species Y, Z equilibrate at time scales much faster than the equilibration times of a single molecule of X . This sort of adiabatic approximation in Y, Z will make calculations more tractable, and allow us to demonstrate the behavior of both limits in λ . We can then make the claim that the probability distribution of a downstream species such as Z is given by

$$P(z) = \int P(z|K) P(K) dK, \quad (\text{C1})$$

where $P(z|K)$ is obtained from the system size expansion for the reaction scheme in which X is constant. Because the copy number n_x is now fluctuating, and we know that the value of K depends on the current copy number of X , we can compute the stationary distribution over K by the following:

$$P(K) = \sum_{n_x} P(K|n_x) P(n_x), \quad (\text{C2})$$

where we know that $P(n_x) \sim \text{Poisson}(\gamma)$ by properties of the birth death process, with $\gamma \equiv k_f/\alpha_x$. For purposes of calculation, we can choose a distribution over k_i such that the n_x -fold convolution is easy to compute, and we can write down the distribution of the sum of n_x random variables k_i . To ensure that the k_i are purely positive, we use the square of a standard $\mathcal{N}(0, 1)$ normal random variable, such that $P(K|n_x) \sim \chi_{n_x}^2$, the chi-squared distribution.

The distribution $P(K = x)$ is then

$$P(K = x) = \sum_{n=0}^{\infty} \left(\frac{\gamma^n}{n!} e^{-\gamma} \right) \left(\frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2} \right) \\ = \frac{1}{x} e^{-(\gamma+x/2)} \sum_{n=0}^{\infty} \frac{\alpha^n}{n! \Gamma(n/2)}, \quad \alpha \equiv \gamma \sqrt{\frac{x}{2}}. \quad (\text{C3})$$

The difficult sum required to compute the marginalization can be written in terms of Hypergeometric functions by first splitting the sum into even and odd powers

$$\sum_{n=0}^{\infty} \frac{\alpha^n}{n! \Gamma(n/2)} = \frac{\alpha^2}{2} \left(\sum_{n=0}^{\infty} \frac{\alpha^{2n}}{(2n+2)! \Gamma(n/2)} \right) \\ + \frac{\alpha}{\Gamma(1/2)} \left(\sum_{n=0}^{\infty} \frac{\alpha^{2n}}{\Gamma(2n+2) \Gamma(n/2)} \right), \quad (\text{C4})$$

where we have used the Pochhammer symbol

$$(\gamma)_k = \frac{\Gamma(\gamma + k)}{\Gamma(\gamma)} \quad (\text{C5})$$

and the fact that for integer argument, $\Gamma(n) = (n-1)!$. Using the fact that $\Gamma(3/2) = \sqrt{\pi}/2$, $\Gamma(2) = 1$, $\Gamma(1/2) = \sqrt{\pi}$, and the following property of gamma functions:

$$\Gamma(z) \Gamma(z + 1/2) 2^{2z-1} = \sqrt{\pi} \Gamma(2z), \quad (\text{C6})$$

we can rewrite the above as

$$\frac{\alpha^2}{2} \sum_{n=0}^{\infty} \frac{(\alpha^2/4)^n}{(3/2)_n (2)_n n!} + \frac{\alpha}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(\alpha^2/4)^n}{n! (3/2)_n (1/2)_n}. \quad (\text{C7})$$

The above sums represent the generalized Hypergeometric functions ${}_pF_q$, and thus we get the identity

$$\sum_{n=0}^{\infty} \frac{\alpha^n}{n! \Gamma(n/2)} = \frac{\alpha^2}{2} {}_0F_2([3/2, 2], \alpha^2/4) \\ + \frac{\alpha}{\sqrt{\pi}} {}_0F_2([3/2, 1/2], \alpha^2/4). \quad (\text{C8})$$

Therefore,

$$P(K = x) = \frac{1}{x} e^{-(\gamma+x/2)} \left[\frac{\gamma^2 x}{4} {}_0F_2 \left(\left[\frac{3}{2}, 2 \right], \frac{\gamma^2 x}{8} \right) \right. \\ \left. + \gamma \sqrt{\frac{x}{2\pi}} {}_0F_2 \left(\left[\frac{3}{2}, \frac{1}{2} \right], \frac{\gamma^2 x}{8} \right) \right]. \quad (\text{C9})$$

From this, we compute the marginal distribution $P(z)$ using the software package Mathematica. Again, the distribution

over downstream species Z is studied numerically using the Gillespie simulation in both the fast and slow switch rate limits. The fast switch rate limit is compared to the case of no switching (i.e., all rates are always at the average value), and the slow switch rate limit is compared to the analytical approximation in which we use the exact result for the stationary distribution of $P(K)$ and compute the distribution $P(z|K)$ assuming that Z adiabatically follows the species X (Figure 3(b)).

¹W. E. Moerner and M. Orrit, *Science* **283**, 1670–1676 (1999).

²S. Weiss, *Science* **283**, 1676–1683 (1999).

³A. M. van Oijen, P. C. Blainey, D. J. Crampton, C. C. Richardson, T. Ellenberger, and X. S. Xie, *Science* **301**, 1235–1238 (2003).

⁴O. Flomenbom, K. Velonia, D. Loos, S. Masuo, M. Cotlet, Y. Engelborghs, J. Hofkens, A. E. Rowan, R. J. M. Nolte, M. V. der Auweraer, F. C. de Schryver, and J. Klafter, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2368–2372 (2005).

⁵R. Zwanzig, *Acc. Chem. Res.* **23**, 148–152 (1990).

⁶W. Min, B. P. English, G. Luo, B. J. Cherayil, S. C. Kou, and X. Sunney Xie, *Acc. Chem. Res.* **38**, 923–931 (2005).

⁷S. C. Kou, B. J. Cherayil, W. Min, B. P. English, and X. Sunney Xie, *J. Phys. Chem. B* **109**, 19068–19081 (2005).

⁸G. K. Schenter and H. P. Lu, *J. Phys. Chem. A* **103**, 10477–10488 (1999).

⁹J. Cao and R. J. Silbey, *J. Phys. Chem. B* **112**, 12867–12880 (2008).

¹⁰H. Yang and X. Sunney Xie, *J. Chem. Phys.* **117**, 10965 (2002).

¹¹I. V. Gopich and A. Szabo, *J. Phys. Chem. B* **114**, 15221 (2010).

¹²I. V. Gopich and A. Szabo, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7747 (2012).

¹³R. Landauer, *IBM J. Res. Dev.* **5**, 183–191 (1961).

¹⁴D. F. Anderson, *J. Chem. Phys.* **127**, 214107 (2007).

¹⁵R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, *Science* **334**, 354–358 (2011).

¹⁶U. Alon, *Nat. Rev.* **8**, 450–461 (2007).

¹⁷N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Oxford, 1981).

¹⁸D. T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).

¹⁹D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).

²⁰D. T. Gillespie, *Physica A* **188**, 404–425 (1992).

²¹O. Flomenbom and R. J. Silbey, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10907 (2006).

²²O. Flomenbom and R. J. Silbey, *J. Chem. Phys.* **128**, 114902 (2008).

²³W. J. Bruno, J. Yang, and J. E. Pearson, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6326 (2005).

²⁴F. Qin, A. Auerbach, and F. Sachs, *Proc. R. Soc. London, Ser. B* **264**, 375 (1997).

²⁵J. Lebowitz and H. Spohn, *J. Stat. Phys.* **95**, 333 (1999).

²⁶P. Mehta and D. J. Schwab, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17978–17982 (2012).

²⁷P. Gaspard, *J. Chem. Phys.* **120**, 8898 (2004).

²⁸R. G. Gallagher, *Stochastic Processes Theory for Applications* (Cambridge University Press, New York, 2013).

²⁹D. S. Leonard and L. E. Reichl, *Phys. Rev. E* **49**, 1734 (1994).

³⁰Z. Hou and H. Xin, *J. Chem. Phys.* **119**, 11508 (2003).

³¹C. Schmitt, B. Dybiec, P. Hanggi, and C. Bechinger, *Europhys. Lett.* **74**, 937–943 (2006).