Information-Optimal Transcriptional Response to Oscillatory Driving

Andrew Mugler,[1] Aleksandra M. Walczak,[2] and Chris H. Wiggins[3]

[1]*Department of Physics, Columbia University, New York, New York 10027, USA*
[2]*Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA*
[3]*Department of Applied Physics and Applied Mathematics, Center for Computational Biology and Bioinformatics, Columbia University, New York, New York 10027, USA*

(Received 11 February 2010; published 29 July 2010)

Intracellular transmission of information via chemical and transcriptional networks is thwarted by a physical limitation: The finite copy number of the constituent chemical species introduces unavoidable intrinsic noise. Here we solve for the complete probabilistic description of the intrinsically noisy response to an oscillatory driving signal. We derive and numerically verify a number of simple scaling laws. Unlike in the case of measuring a static quantity, response to an oscillatory signal can exhibit a resonant frequency which maximizes information transmission. Furthermore, we show that the optimal regulatory design is dependent on biophysical constraints (i.e., the allowed copy number and response time). The resulting phase diagram illustrates under what conditions threshold regulation outperforms linear regulation.

DOI: 10.1103/PhysRevLett.105.058101 PACS numbers: 87.10.Mn, 02.70.Hm, 82.20.Fd, 87.10.Vg

It has long been recognized [1] that the ability to measure biochemical quantities, e.g., concentrations, is intrinsically thwarted by the small copy numbers present at the scale of the cell. This observation has launched considerable experimental investigation as to how high-fidelity signal transmission can occur within single cells [2,3], along with an associated literature in mathematical and computational techniques for modeling such noisy information transmission [4–6]. From the perspective of biological design—either to understand the mechanisms which lead to observed biology or to create synthetic systems with desirable properties—these works investigate how regulatory elements which comprise biological systems function in the presence of intrinsic noise [7].

We consider the simplest probabilistic model of a regulatory element with a dynamical input, illustrated in Fig. 1(a), in which a single transcription factor (the ''parent'') with copy number n is driven by an oscillatory creation rate $f(t) = g + \alpha\cos\omega t$ and regulates the expression of a second species (the ''child'') with copy number m; the regulation is modeled via the child's creation rate q_n. This model captures the noisy downstream response to oscillation, e.g., the cell cycle, without limiting the results to a particular mechanism for generating oscillations (e.g., via cell division [8], repressive cycles [3], or activation-repression circuits [9]). We show how the optimal design—i.e., the choice of linear-vs-cooperative and up-vs-down-regulation—is determined by the physical demands in terms of allowed copy number and response time. Furthermore, while our intuition from understanding how best to measure static signals suggests that slower response time is always more accurate [1], we illustrate how oscillatory driving leads to an information-optimal driving frequency and compute how this frequency depends on copy number.

We use the ''spectral method'' [10], which exploits the linearity of the master equation $\dot{p}_{nm} = -\{\mathcal{L}_n[f(t)] + \rho\mathcal{L}_m[q_n]\}p_{nm}$, the equation of motion for the joint probability of observing n and m copies of the parent and child, respectively, by expanding its solution in terms of the natural eigenfunctions of the birth-death process with constant creation and decay. For a birth-death

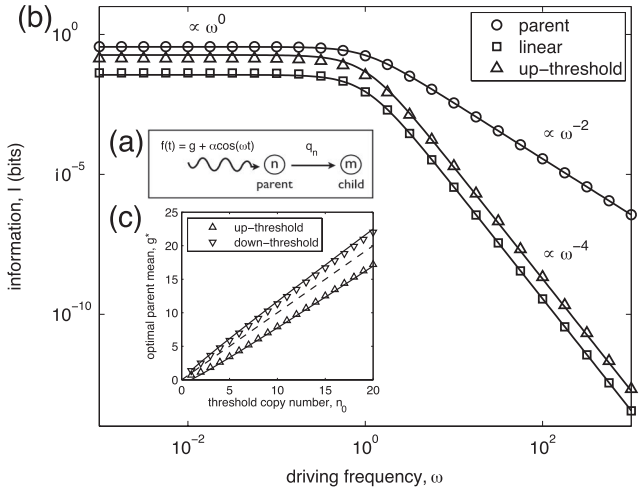

FIG. 1. (a) A transcription factor (the parent) with copy number n is driven by an oscillatory creation rate $f(t) = g + \alpha\cos\omega t$ and regulates via q_n the expression of a second species (the child) with copy number m. (b),(c) Numerical verifications (data points) of analytic expressions (lines) derived in the small-information limit: Eqs. (9) (circles) and (13) top line (squares) and bottom line (triangles) in (b), and Eq. (14) for both up- (up triangles) and down-regulation (down triangles) in (c). Parameters are $g = n_0 = 1$ and $c = 0.1$ for (b), $\omega = 1$ for (c), and $\alpha = \rho = \Delta = 1$ and $q_0 = 0$ for both, yielding small parameters $|\nu_1| \leq 0.5$ [Eq. (5)], $|\mu_1| \leq 0.05$ [linear; Eq. (12)], and $|\mu_1| \leq 0.184$ [threshold; Eq. (12)].

process expressed in terms of an arbitrary creation rate h and copy number s, the positive semidefinite operator $\mathcal{L}_s$ acts as $\mathcal{L}_s[h]p_s = sp_s - (s+1)p_{s+1} + hp_s - hp_{s-1}$; time is normalized via the parent decay rate (in these units ρ is the child decay rate). To study dynamics we also Fourier transform in harmonics of the driving frequency ω:

$$p_{nm}(t) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{z=-\infty}^{\infty} p_{jk}^{z}\langle n \mid j\rangle\langle m \mid k\rangle e^{-iz\omega t}, \quad (1)$$

where the parent and child eigenfunctions (or "spectral modes") enjoy $\mathcal{L}_n[g]\langle n \mid j\rangle = j\langle n \mid j\rangle$ and $\mathcal{L}_m[\bar{q}]\times\langle m \mid k\rangle = k\langle m \mid k\rangle$, respectively. Just as in Refs. [10,11], we introduce a gauge $\bar{q}$ to define the basis; the analytic results below are independent of this choice. The master equation then becomes an algebraic relation among the expansion coefficients p_{jk}^{z}:

$$-i\omega z p_{jk}^{z} = -(j+\rho k)p_{jk}^{z} + \frac{\alpha}{2}\sum_{\pm} p_{j-1,k}^{z\pm 1} - \rho\sum_{j'}\Delta_{jj'}p_{j',k-1}^{z}, \quad (2)$$

where $\Delta_{jj'} = \sum_n \langle j \mid n\rangle(\bar{q} - q_n)\langle n \mid j'\rangle$. Algorithmically we (i) initialize with $G_{00}^{z} = \delta_{z0}$ (set by normalization), (ii) exploit the subdiagonality in k, and (iii) for each k, exploit the subdiagonality in j; no matrices need be inverted.

Efficient computation of $p_{nm}(t)$ allows optimization of the mutual information $I(\phi, n)$ between the input variable—the phase $\phi = \omega t$ of the driving oscillation—and the output variable—the copy number of either the parent or the child:

$$I(\phi, n) = \int_0^{2\pi} d\phi \sum_n p(n \mid \phi)p(\phi)\log\frac{p(n \mid \phi)}{p_n^0}, \quad (3)$$

where $p(n \mid \phi) \equiv p_n(t)$, $p(\phi) = 1/2\pi$, and $p_n^0 = \int_0^{2\pi} d\phi\, p(n \mid \phi)p(\phi)$ is the time-averaged distribution [12].

The dynamics of the parent can be found exactly by using either the method of characteristics or spectral decomposition [13]: $p_n(t)$ is a Poisson distribution with time-dependent mean $\nu(t) = \nu_0 + 2|\nu_1|\cos(\omega t - \gamma)$, with

$$\nu_0 = g, \quad (4)$$

$$|\nu_1| = \frac{\alpha}{2\sqrt{1+\omega^2}}, \quad (5)$$

and phase shift $\gamma = \tan^{-1}\omega$. The Fourier transform coefficients $p_n^{z} = \int_0^{2\pi} d\phi e^{iz\phi} p(n \mid \phi)/(2\pi)$ are then computed by expanding the exponential in $p(n \mid \phi)$ and identifying the modes [13]:

$$p_n^{z} = e^{iz\gamma}\sum_j \frac{|\nu_1|^{2j+|z|}}{j!(j+|z|)!}\langle n \mid 2j + |z|\rangle. \quad (6)$$

In the limit of weak ($\alpha \ll 1$) or fast ($\omega \gg 1$) driving, an approximation for $I(\phi, n)$ may be obtained by expanding in the small parameter $|\nu_1|$. We first express Eq. (3) in terms of the Fourier transform $p(n \mid \phi) = \sum_z p_n^{z} e^{-iz\phi}$:

$$I(\phi, n) = \sum_{n,z} p_n^{z}\int_0^{2\pi}\frac{d\phi}{2\pi}e^{-iz\phi}\log\left(1 + \sum_{z'\neq 0}\frac{p_n^{z'}}{p_n^0}e^{-iz'\phi}\right). \quad (7)$$

Then we note that, for small $|\nu_1|$, Eq. (6) is dominated by the $j = 0$ term, i.e., $p_n^{z} \approx e^{iz\gamma}|\nu_1|^{|z|}\langle n \mid |z|\rangle/|z|!$. Since this is itself small for $z \neq 0$, we expand the log in Eq. (7) as $\log(1+x) = x - x^2/2 + \cdots$ for small x. The first two terms in the log expansion [13] contain the leading-order behavior in p_n^{z} (proportional to $p_n^1 p_n^{-1} = |p_n^1|^2$); employing $\int_0^{2\pi} d\phi e^{i(z-z')\phi} = 2\pi\delta_{zz'}$ one obtains

$$I(\phi, n) \approx \sum_n \frac{|p_n^1|^2}{p_n^0} \approx |\nu_1|^2\sum_n \frac{\langle n \mid 1\rangle^2}{\langle n \mid 0\rangle} = \frac{|\nu_1|^2}{\nu_0} \quad (8)$$

$$= \frac{\alpha^2}{4g}\frac{1}{1+\omega^2}, \quad (9)$$

where the second to last step uses $\langle n \mid 0\rangle = e^{-g}g^n/n!$ and $\langle n \mid 1\rangle = \langle n \mid 0\rangle(n-g)/g$ [11] to evaluate the sum. Equation (8) shows that mutual information asymptotes to the square of the amplitude of the oscillation over the mean. Equation (9) scales like ω^0 at low frequency and ω^{-2} at high frequency, demonstrating that the parent acts as a low-pass filter of information; this scaling is numerically verified in Fig. 1(b).

Although the child distribution $p_m(t)$ is not analytically accessible in general, its mean $\mu(t)$ is exactly calculable: Summing the master equation over both indices against m and Fourier transforming yields $\mu(t) = \sum_z \mu_z e^{-iz\omega t}$, where

$$\mu_z = \frac{1}{1 - iz\omega/\rho}\sum_n q_n p_n^{z}. \quad (10)$$

In the limit of weak regulation (i.e., when q_n is near constant) we may approximate $p_m(t)$ as a Poisson distribution with oscillatory mean parameterized by the first and second Fourier modes of the exact mean, i.e., $\mu(t) \approx \mu_0 \pm 2|\mu_1|\cos(\omega t - \theta)$ for up- (down-) regulation, where $\theta = \text{phase}(\mu_1) = \tan^{-1}\omega/\rho + \tan^{-1}\omega$. Under this approximation, as in Eq. (8), the information between the phase of the driving oscillation and the copy number of the child is the oscillation amplitude squared over the mean, i.e., $I(\phi, m) = |\mu_1|^2/\mu_0$ for small $|\mu_1|$.

To compare the transmission properties of both non- and highly cooperative regulation, we study both the linear function $q_n = q_0 + cn$ and the threshold function $q_n = q_0 + \Delta\chi$ ($n \in \Omega_{\pm}$), respectively, where χ is a characteristic function equal to 1 when n is in the set $\Omega_+ = \{n > n_0\}$ (up-regulation) or $\Omega_- = \{n \leq n_0\}$ (down-regulation) and 0 otherwise. In these cases, the mean of the child distribution oscillates about the point

$$\mu_0 = \sum_n q_n p_n^0 = q_0 + \begin{cases} cg & \text{linear} \\ \Delta p_{\pm}^0 & \text{threshold.} \end{cases} \quad (11)$$

Here the linear result exploits the fact that the mean of the time-averaged parent distribution p_n^0 is g (which can be seen from the relationship between distribution moments and spectral modes [13]). In the threshold result we define $p_\pm^0 \equiv \sum_{n\in\Omega_\pm} p_n^0 = \pi_\pm \pm \sum_{j>0} |\nu_1|^{2j} \langle n_0 \mid 2j-1\rangle / (j!)^2 \approx \pi_\pm$, where $\pi_\pm \equiv \sum_{n\in\Omega_\pm} \langle n \mid 0\rangle$; the second to last step exploits the result $\sum_{n\in\Omega_\pm} \langle n \mid j\rangle = \pm\langle n_0 \mid j-1\rangle$ for $j > 0$ [13], and the last step takes $j = 0$ in the small $|\nu_1|$ limit. The amplitude of the oscillation of the child mean is

$$|\mu_1| = \frac{\sum_n q_n |p_n^1|}{\sqrt{1+(\omega/\rho)^2}} = \frac{1}{\sqrt{1+(\omega/\rho)^2}} \times \begin{cases} c|\nu_1| \\ \Delta |p_\pm^1|, \end{cases} \tag{12}$$

where once more the linear result (top) uses the relationship between moments and modes and in the threshold result (bottom) we define and approximate $|p_\pm^1| \equiv \sum_{n\in\Omega_\pm} |p_n^1| = \sum_j |\nu_1|^{2j+1} \langle n_0 \mid 2j\rangle / [j!(j+1)!] \approx |\nu_1| \langle n_0 \mid 0\rangle$ [13]. Equations (11) and (12) yield the following approximations for linear (top) and threshold (bottom) regulation:

$$I(\phi, m) \approx \frac{g I(\phi, n)}{1+(\omega/\rho)^2} \times \begin{cases} c^2/(q_0 + cg) \\ \Delta^2 \langle n_0 \mid 0\rangle^2 / (q_0 + \Delta \pi_\pm), \end{cases} \tag{13}$$

where $I(\phi, n)$ is as in Eq. (9). Equation (13) shows that the child $I(\phi, m)$ is a sharper low-pass filter than the parent $I(\phi, n)$, falling off like ω^{-4} at high frequency instead of ω^{-2}; this scaling is verified numerically in Fig. 1(b). We note that since t, n, and m are not Markov related, i.e., $p(m \mid t) \neq \sum_n p(m \mid n) p(n \mid t)$, Eq. (13) is not bound by the data-processing inequality [14], and it is possible for the child to transmit more information than the parent, i.e., $I(\phi, m) > I(\phi, n)$ (e.g., for linear regulation with $\omega \to 0$, $q_0 = 0$, and $c > 1$), which we have verified numerically [13].

Equation (13) also offers analytic intuition about the optimal placement of the parent distribution with respect to a threshold regulation function. The derivative of Eq. (13) (bottom) with respect to g vanishes at g^*, the information-optimal mean of the parent distribution:

$$g^* = \frac{n_0}{1 \pm \Delta \langle n_0 \mid 0\rangle / [2(q_0 + \Delta \pi_\pm)]}. \tag{14}$$

As verified in Fig. 1(c), Eq. (14) shows that the parent distribution is shifted below the threshold for up-regulation and above the threshold for down-regulation. These shifts account for the ability of up-regulation to outperform down-regulation when the copy number is highly constrained (see Fig. 3), an effect we observed previously [10] when numerically optimizing steady-state information between the first and last species in a regulatory cascade.

The above analytic approximations offer guidance during a full optimization of $I(\phi, m)$ via numerical integration of Eq. (3). As suggested by Eq. (13), numerical optimization confirms that $I(\phi, m)$ increases when (i) the amplitude of the driving oscillation is maximal ($\alpha = g$) and (ii) the dynamic range is maximal ($q_0 = 0$ and $c \to \infty$ or $\Delta \to \infty$). The slope c or discontinuity Δ, however, is constrained by the average copy number of the child μ_0 [Eq. (11)]. Therefore, for a fixed driving frequency and fixed total average copy number $N = \langle n\rangle + \langle m\rangle = g + \mu_0$, we optimize over the single parameter g by setting $\alpha = g$, $q_0 = 0$, and $c = \mu_0/g = (N-g)/g$ or $\Delta = \mu_0/p_\pm^0 = (N-g)/p_\pm^0$; additionally, we set $\rho = 1$ for equal parent and child decay rates [10]. For threshold regulation, an optimization over g is done at each of a set of values of the (discrete) parameter n_0, and the global optimum is selected.

At low copy number, optimal information $I^*(\phi, m)$ behaves as one might expect from the small-oscillation limit [Eq. (13)]: It decreases monotonically with frequency [Fig. 2(a), bottom curves]. At high copy number, $I^*(\phi, m)$ decreases monotonically with frequency for linear regulation but for threshold regulation exhibits a maximum at a resonant frequency [Fig. 2(a), top curves]. Careful examination of the child distribution at different phases [Figs. 2(b)–2(d)] (or simply its mean [13]) reveals the origin of this maximum as follows. As the parent oscillates about the threshold, the child distribution is switchlike, with two long-lived switch states centered at the threshold's low and high rates and brief intermediate states in between. At high copy number, the threshold rates

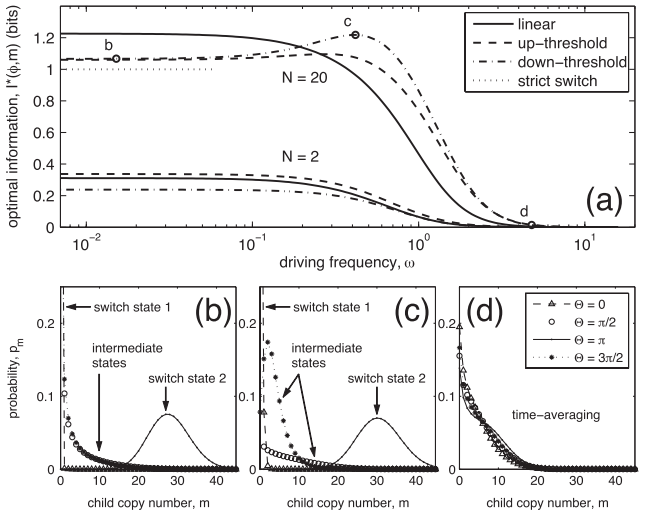

FIG. 2. (a) At high copy number ($N = 20$, top curves) the optimal information $I^*(\phi, m)$ exhibits a resonant driving frequency (point c) for up- (dashed line) and down-threshold (dot-dashed line) regulation but not for linear (solid line) regulation; at low copy number ($N = 2$, bottom curves), there is no resonant frequency, and slowest ($\omega \to 0$) is best. (b)–(d) correspond to marked points in (a) and show the optimal child distribution for down-threshold regulation at phases $\Theta \equiv \omega t - \theta = 0$, $\pi/2$, π, and $3\pi/2$ [legend in (d) applies to (b)–(d)]: (b) Slow driving produces switchlike behavior, with long-lived low- ($\Theta = 0$) and high-copy-number ($\Theta = \pi$) states and brief intermediates ($\Theta = \pi/2$, $3\pi/2$) in between; (c) moderate driving produces switchlike behavior with distinguishable intermediates, transmitting the most information; and (d) fast driving time-averages the parent, and thus the child, distribution.

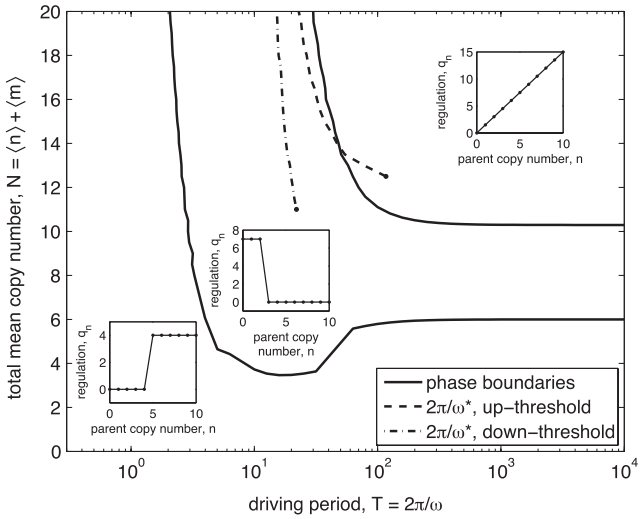

FIG. 3. Phase diagram showing information-optimal regulation among linear, up-threshold, and down-threshold regulation in the space of driving period and copy number. Phases are separated by solid lines and marked by sample regulation functions (insets). Also shown is (2π over) resonant frequency ω^* [see point c in Fig. 2(a)] as a function of copy number for both up- (dashed line) and down-regulation (dot-dashed line).

are far apart (one is zero and the other is large), making the switch states well separated and transmitting (slightly more than, due to the intermediate states) the strict switch limit [10] of $I^*(\phi, m) \sim 1$ bit. For slow oscillations the intermediate states are symmetric [Fig. 2(b)], but for faster oscillations there is a lag in transitioning from one switch state to the other, making the intermediate states distinguishable [Fig. 2(c)] and transmitting more information about phase. Thus the resonant frequency ω^* balances the slowness required to avoid time averaging [Fig. 2(d)] with the speed required for distinguishable intermediate states. The high-copy-number behavior described here has been verified for N as high as 50 [13].

The phase diagram (Fig. 3) shows the information-optimal regulation among linear, up-threshold, and down-threshold regulation across a range of copy numbers N and periods $T = 2\pi/\omega$. Linear regulation is best when both N and T are large, ultimately surpassing threshold regulation's limit of $\sim$1 bit. Down-threshold regulation is best at values of T near $2\pi/\omega^*$ because its intermediate states are more distinguishable (i.e., have a larger Jensen-Shannon divergence) than those of a similarly parameterized up threshold. Up-threshold regulation is best at low N due to its tendency, as discussed above [Eq. (14)] and in Ref. [10], to require fewer proteins to match the transmission across a similarly parameterized down threshold.

Finally, we turn to comparing our predictions with real biological parameters. In simpler systems, where decay rates are bounded from below by dilution from cell division, the ratio of the cell cycle frequency to a species' decay rate is $\omega \leq 2\pi/\ln 2 \approx 9$. Our results predict that slightly smaller ratios are information-optimal ($\omega^* \sim 2\pi/10 \approx 0.6$ at high N; see Fig. 3), as in, e.g., yeast glycolysis oscillations ($\omega \sim 0.5$ [15]) or bacterial circadian rhythms, when, through either active degradation [16] or rapid cell division [17], decay rates are several times larger than the circadian frequency, making ω order 1.

The low-pass behavior revealed in Eqs. (9) and (13) is consistent with our intuition from measuring static quantities in the presence of intrinsic noise [1]: The longer we wait, the more accurate our estimate. However, in the presence of an oscillatory driving signal, we find that threshold regulation can lead to an information-optimal frequency, and waiting longer is not necessarily the optimal strategy. Furthermore, we have shown that, at a fixed allowed copy number and allowed integration time, one may find that a different regulation strategy (linear, threshold up-regulation, or threshold down-regulation) is optimal for responding to oscillatory driving. Absent from this analysis are intriguing questions such as whether the diversity of other network topologies observed in nature—including cascades and feedback circuits—are consistent with these observations.

[1] H. C. Berg and E. M. Purcell, Biophys. J. **20**, 193 (1977).
[2] J. Mettetal, D. Muzzey, C. Gomez-Uribe, and A. van Oudenaarden, Science **319**, 482 (2008).
[3] M. B. Elowitz and S. Leibler, Nature (London) **403**, 335 (2000).
[4] F. Tostevin and P. R. ten Wolde, Phys. Rev. Lett. **102**, 218101 (2009).
[5] S. Tănase-Nicola, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **97**, 068102 (2006).
[6] G. Tkacik, C. G. Callan, and W. Bialek, Proc. Natl. Acad. Sci. U.S.A. **105**, 12 265 (2008).
[7] W. Bialek and S. Setayeshgar, Proc. Natl. Acad. Sci. U.S.A. **102**, 10 040 (2005).
[8] A. Csikász-Nagy, D. Battogtokh, K. C. Chen, B. Novák, and J. J. Tyson, Biophys. J. **90**, 4361 (2006).
[9] N. Cookson, L. Tsimring, and J. Hasty, FEBS Lett. **583**, 3931 (2009).
[10] A. M. Walczak, A. Mugler, and C. H. Wiggins, Proc. Natl. Acad. Sci. U.S.A. **106**, 6529 (2009).
[11] A. Mugler, A. M. Walczak, and C. H. Wiggins, Phys. Rev. E **80**, 041921 (2009).
[12] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 632 (1948).
[13] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevLett.105.058101 for complete analytic derivations and additional numerical results.
[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[15] A. Boiteux, A. Goldbeter, and B. Hess, Proc. Natl. Acad. Sci. U.S.A. **72**, 3829 (1975).
[16] Y. Xu, T. Mori, and C. H. Johnson, EMBO J. **22**, 2117 (2003).
[17] T. Kondo, T. Mori, N. V. Lebedeva, S. Aoki, M. Ishiura, and S. S. Golden, Science **275**, 224 (1997).