

## How input fluctuations reshape the dynamics of a biological switching system

Bo Hu,<sup>1</sup> David A. Kessler,<sup>2</sup> Wouter-Jan Rappel,<sup>3</sup> and Herbert Levine<sup>4</sup>

<sup>1</sup>*IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598, USA*

<sup>2</sup>*Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel*

<sup>3</sup>*Center for Theoretical Biological Physics, University of California–San Diego, La Jolla, California 92093-0319, USA*

<sup>4</sup>*Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, USA*

(Received 13 September 2012; published 20 December 2012)

An important task in quantitative biology is to understand the role of stochasticity in biochemical regulation. Here, as an extension of our recent work [*Phys. Rev. Lett.* **107**, 148101 (2011)], we study how input fluctuations affect the stochastic dynamics of a simple biological switch. In our model, the on transition rate of the switch is directly regulated by a noisy input signal, which is described as a non-negative mean-reverting diffusion process. This continuous process can be a good approximation of the discrete birth-death process and is much more analytically tractable. Within this setup, we apply the Feynman-Kac theorem to investigate the statistical features of the output switching dynamics. Consistent with our previous findings, the input noise is found to effectively suppress the input-dependent transitions. We show analytically that this effect becomes significant when the input signal fluctuates greatly in amplitude and reverts slowly to its mean.

DOI: [10.1103/PhysRevE.86.061910](https://doi.org/10.1103/PhysRevE.86.061910)

PACS number(s): 87.18.Cf, 87.18.Tt, 02.50.Le, 05.65.+b

### I. INTRODUCTION

Stochasticity appears to be a hallmark of many biological processes involved in signal transduction and gene regulation [1–11]. Over the past decade, there have been numerous experimental and theoretical efforts to understand the functional roles of noise in various living systems [2–41]. Remarkably, the building block of different regulatory programs is often a simple two-state switch under the regulation of some noisy input signal. For example, a gene network is composed of many interacting genes, each of which is a single switch regulated by specific transcription factors. At a synapse, switching of ligand-gated ion channels is responsible for converting the presynaptic chemical message into postsynaptic electrical signals, while the opening and closing of each channel depends on the binding of certain ligands (such as a neurotransmitter). In bacterial chemotaxis, the cellular motion is powered by multiple flagellar motors, and each motor rotates clockwise or counterclockwise depending on the level of some specific response regulator (e.g., CheY-P in *E. coli*). Irrespective of the context, the input signal, i.e., the number of regulatory molecules, is usually stochastic due to discreteness, diffusion, random birth, and death. How does a biological switching system work in a noisy environment? This is the central question we attempt to address in this paper.

Previous studies on this topic have usually focused on the approximate, static relationship between input and output variations (e.g., the *additive noise rule*) [12–19], while the dynamic details (e.g., dwell time statistics) of the switching system have often been ignored. Our recent work suggests that there is more to comprehend even in the simplest switching system [22]. For example, we showed that increasing input noise does not always lead to an increase in the output variation, disagreeing with the *additive noise rule* as derived from the coarse-grained Langevin approach. Traditional methods often use a single Langevin equation to approximate the joint input-output process (which is only applicable for an ensemble of switches), and they rely on the assumption that the input noise is small enough such that one can linearize

input-dependent nonlinear reaction rates. Our approach to this problem is quite different as we explicitly model how the input stochastic process drives a single output switch, without making any small noise assumption.

In our previous paper [22], the input signal was generated from a discrete birth-death process and regulated the on transition of a downstream switch. By explicitly solving the joint master equation of the system, we found that input fluctuations can effectively reduce the on rate of the switch. In this paper, this problem is revisited in a continuous noise formulation. We propose to model the input signal as a general diffusion process, which is mean-reverting, non-negative, and tunable in Fano factor (the ratio of variance to mean). We employ the Feynman-Kac theorem to calculate the input-dependent dwell time distribution and examine its asymptotic behavior in different scenarios. Within this framework, we recover several findings reported in [22], and we also demonstrate how the noise-induced suppression depends on the noise intensity as well as the relative time scales of the input and output dynamics. Finally, we elaborate on how the diffusion process introduced in this paper can be a reasonable approximation of the discrete birth-death process.

### II. MODEL

The input of our model, denoted by  $X(t)$ , represents a specific chemical concentration at time  $t$  and directly governs the transition rates of a downstream switch. The binary on-off states of the switch in continuous time constitute the output process,  $Y(t)$ . A popular choice for  $X(t)$  is the Ornstein-Uhlenbeck (OU) process due to its analytical simplicity and mean-reverting property [42–44]. However, this process does not rule out negative values, an unphysical feature for modeling chemical concentrations. For both mathematical convenience and biophysical constraints (see Sec. IV for more details), we model  $X(t)$  by a square-root diffusion process [45–48]:

$$dX(t) = \lambda[\mu - X(t)]dt + \sigma\sqrt{X(t)}dW_t, \quad (1)$$

where  $\lambda$  represents the rate at which  $X(t)$  reverts to its mean level  $\mu$ ,  $\sigma$  controls the noise intensity, and  $W_t$  denotes the standard Brownian motion. This simple process is known as the Cox-Ingersoll-Ross (CIR) model of interest rates [45]. The square-root noise term not only ensures that the process never goes negative but also captures a common statistical feature underlying many biochemical processes, that is, the standard deviation of the copy number of molecules scales as the square root of the copy number, as dictated by the central limit theorem. Solving the Fokker-Planck equation for Eq. (1), we obtain the steady-state distribution of the input signal:

$$P_s(X = x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, \quad \alpha \equiv \frac{2\mu\lambda}{\sigma^2}, \quad \beta \equiv \frac{2\lambda}{\sigma^2}, \quad (2)$$

which is a Gamma distribution with stationary variance  $\sigma_X^2 = \mu\sigma^2/(2\lambda)$ . This is another attractive aspect of this model, since the protein abundance from gene expression experiments can often be fitted by a Gamma distribution [8–10]. The parameter  $\alpha$  in Eq. (2) can be interpreted as the *signal-to-noise ratio*, since  $\alpha = \mu^2/\sigma_X^2$ . For  $\alpha \geq 1$ , the zero point is guaranteed to be inaccessible for  $X(t)$ . The other shape parameter  $\beta$  sets the Fano factor as we have  $1/\beta = \sigma_X^2/\mu$ . Using the Itô calculus [44], one can also find the steady-state covariance:  $\lim_{t \rightarrow \infty} \text{cov}[X(t), X(t+s)] = \sigma_X^2 e^{-\lambda|s|}$ . Thus  $X(t)$  is a stationary process with correlation time  $\lambda^{-1}$ .

In reality, the output switching rates may depend on the input  $X(t)$  in complicated ways, depending on the detailed molecular mechanism. For analytical convenience, we assume that the on and off transition rates of the switch are  $k_{\text{on}}X(t)$  and  $k_{\text{off}}$ , respectively (Fig. 1). As a result, the input fluctuations should only affect the chance of the switch exiting the off state; the on-state dwell time distribution is always exponential with rate parameter  $k_{\text{off}}$ . If we mute the input noise, then  $Y(t)$  reduces to a two-state Markov process with transition rates  $k_{\text{on}}\mu$  and  $k_{\text{off}}$ . However, the presence of input noise will generally make  $Y(t)$  a non-Markovian process, because the off-state time intervals may exhibit a nonexponential distribution. To illustrate this point more rigorously, we consider the following first passage time problem [44]. Suppose the switch starts in the off state at  $t = 0$  with the initial input  $X(0) = x$ . Let  $\tilde{\tau}$  be the first time of the switch turning on (Fig. 1). Then the *survival probability*  $f(x, t)$  for the switch staying off up to time  $t$  is given by

$$f(x, t) \equiv P[\tilde{\tau} > t | X(0) = x] = E^x \left[ e^{-\int_0^t k_{\text{on}} X(s) ds} \right], \quad (3)$$

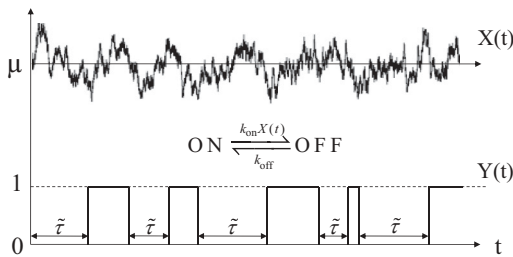


FIG. 1. Illustration of our model:  $X(t)$  represents the input signal which fluctuates around a mean level over time;  $Y(t)$  records the switch process which flips between the off ( $Y = 0$ ) and on ( $Y = 1$ ) states with transition rates  $k_{\text{on}}X(t)$  and  $k_{\text{off}}$ .  $\tilde{\tau}$  is the dwell time in the off state.

where  $E^x[\cdot]$  denotes expectation over all possible sample paths of  $X(s)$  for  $0 \leq s \leq t$ , conditioned on  $X(0) = x$ . In probability theory, Jensen's inequality is generally stated in the following form: if  $X$  is a random variable and  $\varphi$  is a convex function, then  $E[\varphi(X)] \geq \varphi(E[X])$ . Applying Jensen's inequality to Eq. (3) leads to  $f(x, t) \geq \exp(-k_{\text{on}}\mu t)$ , where the equality holds if the input is noise-free. This inequality suggests that the input noise will elongate the mean waiting time for the off-to-on transitions, regardless of the noise intensity or the correlation time. We will elaborate on this noise effect in a more analytical way. The Feynman-Kac formula [44] asserts that  $f(x, t)$  solves the following partial differential equation:

$$\frac{\partial f}{\partial t} = \lambda(\mu - x) \frac{\partial f}{\partial x} + \frac{1}{2} \sigma^2 x \frac{\partial^2 f}{\partial x^2} - k_{\text{on}} x f, \quad (4)$$

with the initial condition  $f(x, 0) = 1$ . As we will show in the next section, the off-state dwell time distribution is not exactly exponential, though it is asymptotically exponential when  $t$  is large.

### III. RESULTS

A similar partial differential equation to Eq. (4) has been solved in Ref. [45] to price zero-coupon bonds under the CIR interest rate model. The closed-form solution for our problem is similar and is found to be

$$f(x, t) = \left[ \frac{\tilde{\lambda} e^{\tilde{\lambda} t/2} / \sinh(\tilde{\lambda} t/2)}{\lambda + \tilde{\lambda} \coth(\tilde{\lambda} t/2)} \right]^\alpha \exp \left[ \frac{-2k_{\text{on}} x}{\lambda + \tilde{\lambda} \coth(\tilde{\lambda} t/2)} \right], \quad (5)$$

where

$$\tilde{\lambda} \equiv \sqrt{\lambda^2 + 2k_{\text{on}}\sigma^2} = \lambda \sqrt{1 + 2k_{\text{on}}\sigma^2/\lambda^2}. \quad (6)$$

Evidently,  $f(x, t)$  remembers the initial input  $x$  and decays with  $t$  in a manner which is not exactly exponential. For  $t \gg \tilde{\lambda}^{-1}$ , Eq. (5) takes the following form:

$$f(x, t) \simeq \left( \frac{2\tilde{\lambda}}{\lambda + \tilde{\lambda}} \right)^\alpha \exp \left( -\frac{\tilde{k}_{\text{on}} x}{\lambda} - \tilde{k}_{\text{on}} \mu t \right). \quad (7)$$

In deriving Eq. (7), we have used the following relationship, which defines the parameter  $\tilde{k}_{\text{on}}$  as below:

$$\tilde{k}_{\text{on}} \equiv \frac{\alpha(\tilde{\lambda} - \lambda)}{2\mu} = \frac{\lambda}{\sigma^2} (\tilde{\lambda} - \lambda) = \frac{2\lambda}{\lambda + \tilde{\lambda}} k_{\text{on}} < k_{\text{on}}. \quad (8)$$

Thus, asymptotically speaking,  $f(x, t)$  decays with  $t$  in an exponential manner at the rate  $\tilde{k}_{\text{on}}\mu$ . It can be easily verified that Eq. (7) is a particular solution to Eq. (4).

To gain further insight into how the input noise affects the switching dynamics, we first study the “slow switch” limit where  $X(t)$  fluctuates so rapidly that  $\lambda^{-1} \ll \mathcal{T}_Y$ . Here,  $\mathcal{T}_Y \equiv (k_{\text{on}}\mu + k_{\text{off}})^{-1}$  is the output correlation time for the noiseless input model. In this limit, the initial input  $x$  in  $f(x, t)$  at the start of each off period is effectively drawn from the Gamma distribution  $P_s(x)$  defined in Eq. (2); the successive off-state intervals are almost independent of each other (as the input loses memory quickly) and are distributed as

$$P(\tilde{\tau} \leq t) = 1 - P(\tilde{\tau} > t) = 1 - \int_0^\infty f(x, t) P_s(x) dx. \quad (9)$$

By direct integration over  $x$ , we find that

$$P(\tilde{\tau} > t) = \left[ \frac{\beta \tilde{\lambda} e^{\lambda t/2}}{(\beta \lambda + 2k_{\text{on}}) \sinh(\tilde{\lambda} t/2) + \beta \tilde{\lambda} \cosh(\tilde{\lambda} t/2)} \right]^\alpha \simeq \left[ \frac{2\beta \tilde{\lambda} e^{-(\tilde{\lambda}-\lambda)t/2}}{\beta(\lambda + \tilde{\lambda}) + 2k_{\text{on}}} \right]^\alpha \simeq \exp(-\tilde{k}_{\text{on}} \mu t). \quad (10)$$

In the last step above, we have used Eq. (8) as well as the following observation: By introducing  $\theta \equiv k_{\text{on}} \sigma^2 / \lambda^2$ , which reflects the deviation of  $\tilde{\lambda}$  from  $\lambda$ , one can check that as long as  $\lambda \gg k_{\text{on}} \mu$  (as ensured by  $\lambda^{-1} \ll \mathcal{T}_Y$ ) the following holds, regardless of the values of  $\theta$ :

$$\left[ \frac{2\beta \tilde{\lambda}}{\beta(\lambda + \tilde{\lambda}) + 2k_{\text{on}}} \right]^\alpha = \left( \frac{1}{2} + \frac{1 + \theta}{2\sqrt{1 + 2\theta}} \right)^{-\frac{2}{\theta} \frac{k_{\text{on}} \mu}{\lambda}} \simeq 1.$$

Our simulations show that the approximate result,  $P(\tilde{\tau} > t) \simeq e^{-\tilde{k}_{\text{on}} \mu t}$ , is excellent [Fig. 2(a)], independent of the values of  $\theta$ . Thus, the average waiting time for the switch to turn on is  $(\tilde{k}_{\text{on}} \mu)^{-1}$ , longer than the corresponding average time  $(k_{\text{on}} \mu)^{-1}$  for the noiseless input model. This is similar to our previous result [22], and suggests that the input noise will effectively suppress the on state by increasing the average waiting time to exit the off state. Consequently, the probability,  $P_{\text{on}}$ , to find the switch on ( $Y = 1$ ) is less than that in the noiseless input model:

$$P_{\text{on}} \simeq \frac{\mu}{\mu + \tilde{K}_d} < \frac{\mu}{\mu + K_d} = \lim_{\sigma \rightarrow 0} P_{\text{on}}, \quad (11)$$

where  $\tilde{K}_d \equiv k_{\text{off}} / \tilde{k}_{\text{on}}$ , the *effective* equilibrium constant, is larger than the original  $K_d \equiv k_{\text{off}} / k_{\text{on}}$  as per Eq. (8).

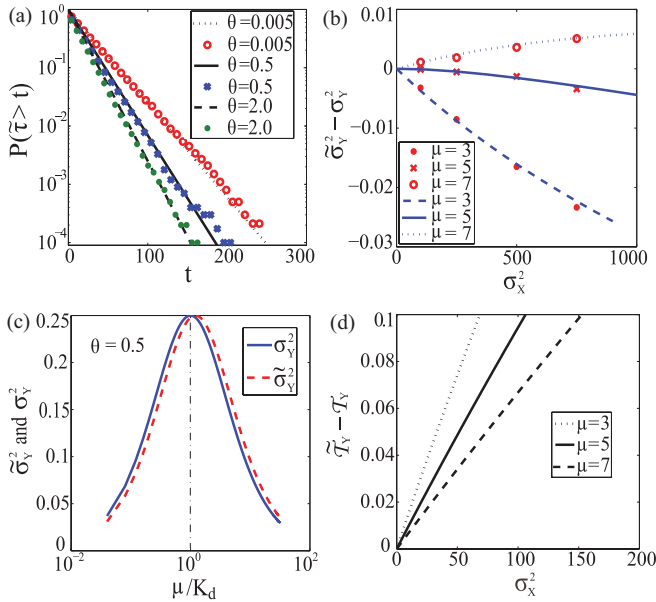


FIG. 2. (Color online) The slow switch case. Here we use  $\lambda = 10$ ,  $k_{\text{on}} = 0.02$ , and  $k_{\text{off}} = 0.1$  (thus  $K_d = 5$ ). (a)  $P(\tilde{\tau} > t)$  vs  $t$  for  $\mu = 3$  and  $\theta = 0.005, 0.5$ , and  $2.0$ , which are achieved by choosing  $\sigma = 5, 50$ , and  $100$ . Symbols represent simulation results, while lines denote  $\exp(-\tilde{k}_{\text{on}} \mu t)$ . (b)  $\tilde{\sigma}_Y^2 - \sigma_Y^2$  vs  $\sigma_X^2$ , where different values of  $\sigma_X^2$  are obtained by tuning  $\sigma$ . (c)  $\tilde{\sigma}_Y^2$  and  $\sigma_Y^2$  vs  $\mu/K_d$  with fixed  $\theta = 0.50$ . (d)  $\tilde{T}_Y - T_Y$  vs  $\sigma_X^2$ .

Therefore, in the slow switch limit ( $\mathcal{T}_Y \gg \lambda^{-1}$ ), the output  $Y(t)$  is approximately a two-state Markov process with transition rates  $\tilde{k}_{\text{on}} \mu$  and  $k_{\text{off}}$ . If we further assume that the noise is modest (i.e.,  $\sigma_X \leq \mu$ ), then

$$\theta \equiv k_{\text{on}} \frac{\sigma^2}{\lambda^2} = \frac{2k_{\text{on}} \mu}{\alpha \lambda} = \frac{2k_{\text{on}} \mu}{\lambda} \left( \frac{\sigma_X^2}{\mu^2} \right) \ll 1. \quad (12)$$

The equality above shows that  $\theta$  is a characteristic parameter determined by the ratio of the input to output time scales and the relative noise strength. For  $\theta \ll 1$ , we have  $\tilde{\lambda} = \lambda \sqrt{1 + 2\theta} \simeq \lambda(1 + \theta)$  by Eq. (6), and the effective on rate defined in Eq. (8) becomes

$$\tilde{k}_{\text{on}} = \frac{2k_{\text{on}}}{1 + \sqrt{1 + 2\theta}} \simeq \frac{k_{\text{on}}}{1 + \frac{\theta}{2}} = k_{\text{on}} \left( 1 + \frac{k_{\text{on}} \sigma_X^2}{\lambda \mu} \right)^{-1}. \quad (13)$$

In [22], the input signal was taken to be a Poisson birth-death process for which the variance is equal to the mean ( $\sigma_X^2 = \mu$ ), and the time scale has been normalized by putting the death rate equal to 1 (which amounts to setting  $\lambda = 1$  here). These two constraints reduce Eq. (13) to  $\tilde{k}_{\text{on}} \simeq k_{\text{on}} / (1 + k_{\text{on}})$ , recovering the result we derived in [22]. The consistency indicates that our key findings are general and independent of the specific model we choose. The continuous diffusion model here, however, is more flexible as it allows the Fano factor to differ from 1, that is,  $\sigma_X^2 \neq \mu$ .

For small  $\theta$ , the stationary variance of  $Y(t)$  can be expanded as follows:

$$\begin{aligned} \tilde{\sigma}_Y^2 &= P_{\text{on}}(1 - P_{\text{on}}) \simeq \sigma_Y^2 + \frac{\mu K_d (\mu - K_d)}{2(\mu + K_d)^3} \theta + O(\theta^2), \\ &= \sigma_Y^2 + \frac{\mu - K_d}{(\mu + K_d)^3} \frac{k_{\text{off}}}{\lambda} \sigma_X^2 + O(\sigma_X^4), \end{aligned} \quad (14)$$

where  $\sigma_Y^2 \equiv \mu K_d / (\mu + K_d)^2$  is the output variance of the noiseless input model. Equation (14) indicates that the input noise  $\sigma_X^2$  does not always contribute positively to the output variance  $\tilde{\sigma}_Y^2$ . In fact, the contribution is negligible when  $\mu$  is near  $K_d$  and even negative for  $\mu < K_d$  [Fig. 2(b)]. The explanation is that the input noise will effectively suppress the on transition rate and thus defines an effective equilibrium constant  $\tilde{K}_d$  which, as implied by Eq. (11) and shown in Fig. 2(c), is larger than the original dissociation constant  $K_d$ . Finally, with the effective on rate  $\tilde{k}_{\text{on}}$ , the correlation time of  $Y(t)$  becomes  $\tilde{T}_Y = (\tilde{k}_{\text{on}} \mu + k_{\text{off}})^{-1}$ , and can likewise be expanded as

$$\tilde{T}_Y = \mathcal{T}_Y \left( 1 - \frac{\tilde{\lambda} - \lambda}{\tilde{\lambda} + \lambda} \frac{\mu}{\mu + K_d} \right)^{-1} \simeq \mathcal{T}_Y + \frac{1}{\lambda} \frac{\sigma_X^2}{(\mu + K_d)^2}. \quad (15)$$

Thus,  $\tilde{T}_Y$  increases almost linearly with the input noise in this small noise limit [Fig. 2(d)].

We now examine the “fast switch” limit where the switch flips much faster than the input reverts to its mean ( $\mathcal{T}_Y \ll \lambda^{-1}$ ). In this scenario, the initial input values  $\{x_i, i = 1, 2, \dots\}$  for successive first-passage time intervals  $\{\tilde{\tau}_i, i = 1, 2, \dots\}$  are correlated due to the slow relaxation of  $X(t)$ . This memory makes the sequence  $\{\tilde{\tau}_i, i = 1, 2, \dots\}$  correlated as well, as confirmed by our Monte Carlo simulations [Fig. 3(a)]. For the same reason, the autocorrelation function (ACF) of the

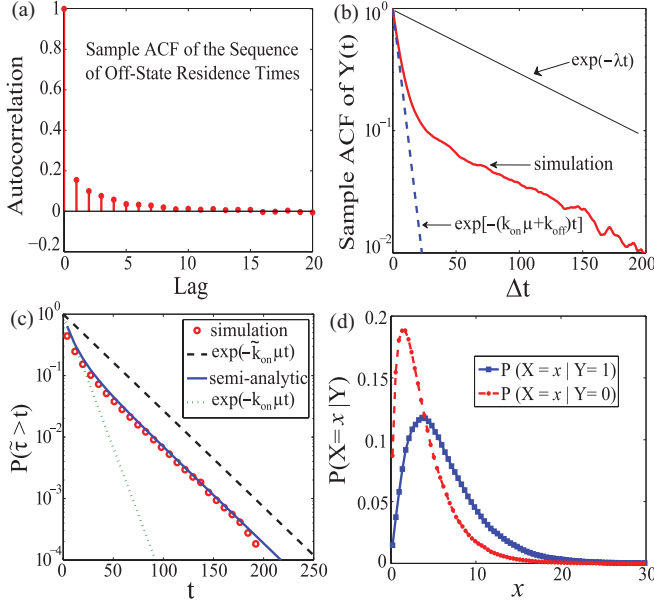


FIG. 3. (Color online) The fast switch case. Here we choose  $\lambda = 0.0125$ ,  $\mu = 5$ ,  $k_{\text{off}} = 0.1$ ,  $k_{\text{on}} = 0.02$ , and  $\theta = 10$  (such that  $\sigma \simeq 0.28$ ). (a) Sample autocorrelation function (ACF) of successive off-state time intervals. (b) Sample ACF of a simulated sample path of  $Y(t)$  in the semilogarithmic scale. (c) Distribution of the off-state intervals,  $P(\tilde{\tau} > t)$ . Symbols are from Monte Carlo simulations and the solid blue line is the semianalytical prediction described in the main text. (d) Input distributions conditioned on the output.

output  $Y(t)$  exhibits two exponential regimes [Fig. 3(b)]: over short time scales, it is dominated by the intrinsic time  $\bar{T}_Y$  of the switch; over long time scales, however, it decays exponentially at the input relaxation rate  $\lambda$ . This demonstrates that the long-term memory in  $X(t)$  is inherited by the output process  $Y(t)$ . For a fast switch, the distribution  $P(\tilde{\tau} > t)$  is not fully exponential [Fig. 3(c)], though it decays exponentially at the rate  $\tilde{k}_{\text{on}}\mu$  for large  $t$ , as predicted by the asymptotic Eq. (7). We can still use the closed-form solution of  $f(x, t)$  in Eq. (5) to fit the simulation data. Specifically, we calculate  $P(\tilde{\tau} > t) = \int_0^\infty f(x, t) P_{\tilde{\tau}}(x) dx$ , where  $P_{\tilde{\tau}}(x)$  is the distribution of the initial  $x$  for each switching event (defining the “first-passage” time  $\tilde{\tau}$ ) and can be obtained from the same Monte Carlo simulations. Clearly, this semianalytical approach [solid blue line in Fig. 3(c)] provides a nice fit to the simulation results (open circles).

Note that  $P_{\tilde{\tau}}(x) \neq P_s(x)$  due to memory effects in the fast switch limit. To show this, we illustrate the input-output interdependence by plotting the input distributions conditioned on the output state in Fig. 3(d). In fact, we have  $P(X = x | Y = 1) = P_{\tilde{\tau}}(x)$ . This is because  $x$  in  $P_{\tilde{\tau}}(x)$  denotes  $x = X(t_0^+)$ , where  $t_0$  is the last time of off-transition;  $X(t_0^+) = X(t_0^-)$  due to continuity and  $Y(t_0^-) = 1$  by definition; and finally, all the on-state intervals are memoryless with the same Poisson rate  $k_{\text{off}}$ . This simple relation  $P(X = x | Y = 1) = P_{\tilde{\tau}}(x)$  has been confirmed by our simulations (results not shown). It is also obvious from Fig. 3(d) that the expectation value of the input conditioned on  $Y = 1$  is larger than when conditioned on  $Y = 0$ . Thus the mean of  $X(t)$  should lie in between, i.e.,  $E[X | Y = 1] > E[X] > E[X | Y = 0]$ , an interesting feature

of the fast switch limit [22]. Since  $f(x, t)$  is a decreasing function of  $x$  and the initial input  $x$  is likely to be larger under  $P_{\tilde{\tau}}(x) = P(X = x | Y = 1)$  than under the measure  $P_s(X = x)$ , we should have

$$P(\tilde{\tau} > t) = \int_0^\infty f(x, t) P_{\tilde{\tau}}(x) dx < \int_0^\infty f(x, t) P_s(x) dx \simeq e^{-\tilde{k}_{\text{on}}\mu t}. \quad (16)$$

The above inequality explains why the distribution of  $\tilde{\tau}$  is below the single exponential  $e^{-\tilde{k}_{\text{on}}\mu t}$  (dashed black line) in Fig. 3(c). All the above results [Figs. 3(a)–3(d)] indicate that the output process  $Y(t)$  is non-Markovian in the fast switch limit and, again, confirm the more general applicability of our findings reported in [22].

#### IV. DIFFUSION APPROXIMATION

In this paper, we have used the square-root diffusion (or CIR) process to model biochemical fluctuations. Here, we will show that the CIR process is well motivated and gives an excellent approximation to the birth-death process adopted in our earlier work [22]. We will also demonstrate that the non-negativity of the CIR process is advantageous as it avoids unphysical results.

We first argue that the CIR process is inspired by the fundamental nature of general biochemical processes. Many biochemical signals are subject to counteracting effects: synthesis and degradation, activation and deactivation, transport in and out of a cellular compartment, etc. As a result, these signals tend to fluctuate around their equilibrium values. A simple yet realistic model to capture these phenomena is the birth-death process, which we adopted to model biochemical noise in [22]. Remarkably, a birth-death process can be approximated by a Markov diffusion process [42]. The standard procedure is to employ the Kramers-Moyal expansion to convert the master equation into a Fokker-Planck equation (terminating after the second term). This connection allows the use of a Langevin equation to approximate the birth-death process. We will explain this in a more intuitive way.

Assume that the birth and death rates for the input signal  $X(t)$  are  $\nu$  and  $\lambda X(t)$ . Then the stationary distribution of  $X(t)$  is a Poisson distribution, with its mean, variance, and skewness given by  $\mu \equiv \nu/\lambda$ ,  $\sigma_x^2 = \mu$ , and  $\mu^{-1/2}$ , respectively. The corresponding Langevin equation that approximates this birth-death process is

$$\frac{dX(t)}{dt} = \nu - \lambda X(t) + \eta(t), \quad (17)$$

where the stochastic term  $\eta(t)$  represents a white noise with  $\langle \eta(t) \rangle = 0$  and  $\delta$  correlation,

$$\langle \eta(t)\eta(t') \rangle = (\nu + \lambda X)\delta(t - t') = \lambda(\mu + X)\delta(t - t'). \quad (18)$$

Physically, the Langevin approximation Eq. (17) holds when the copy number of molecules is large and the time scale of interest is longer than the characteristic time ( $\lambda^{-1}$ ) of the birth-death process. Equation (18) indicates that the instantaneous variance of the noise term  $\eta(t)$  equals the sum of the birth rate  $\nu$  and the death rate  $\lambda X(t)$ . An intuitive interpretation is that since both the birth and death events follow independent Poisson



processes, the total variance of the increment  $X(t + \Delta t) - X(t)$  over a short time  $\Delta t$  must be equal to  $v\Delta t + \lambda X(t)\Delta t$ . As  $\mu \equiv v/\lambda$ , we can rewrite the Langevin Eq. (17) as the following Itô-type stochastic differential equation (SDE):

$$dX(t) = \lambda[\mu - X(t)]dt + \sqrt{\lambda\mu + \lambda X(t)}dW_t, \quad (19)$$

which is similar to Eq. (1) introduced at the beginning. A transformation  $X'(t) \equiv X(t) + \mu$  is convenient for exploiting our existing results, as the SDE for  $X'(t)$  is

$$dX' = \lambda(2\mu - X')dt + \sqrt{\lambda X'}dW_t, \quad (20)$$

which is a particular CIR process. It is easy to check that the stationary variance of  $X'(t)$  equals  $\mu$ , and so does the variance of  $X(t)$ . In other words, we still have  $\sigma_X^2 = \mu$  for  $X(t)$  under Eq. (19). As a CIR process,  $X'$  in equilibrium follows a Gamma distribution, the skewness of which is found to be  $\mu^{-1/2}$ . Therefore,  $X = X' - \mu$  follows a “shifted” Gamma distribution with its mean, variance, and skewness given by  $\mu$ ,  $\mu$ , and  $\mu^{-1/2}$ , respectively, equal to those of the Poisson distribution. This matching of moments suggests that Eq. (19) is indeed a good approximation to the original birth-death process. However,  $X(t)$  in Eq. (19) can take negative values, because  $X = X' - \mu$  is bounded below by  $-\mu$  due to its “shift” in distribution. This becomes a limitation of the Langevin approximation Eq. (17), which could fail if the noise is large (i.e., the number of molecules is small).

Under Eq. (19) for the input  $X(t)$ , we can still evaluate the analog of Eq. (3):

$$f(x, t) = E^x \left[ e^{-\int_0^t k_{\text{on}} X(s) ds} \right] = E^{x+\mu} \left[ e^{-\int_0^t k_{\text{on}} X'(s) ds} \right] e^{k_{\text{on}} \mu t}, \quad (21)$$

where  $E^{x+\mu}[\dots]$  denotes expectation over all possible sample paths of  $X'$  over  $[0, t]$ , conditioned on  $X'(0) = x + \mu$ . Since  $X'$  follows the CIR process Eq. (20), the expectation  $E^{x+\mu}[\dots]$  has an expression similar to Eq. (5). In fact, the survival probability  $f(x, t)$  in Eq. (21) equals

$$\left[ \frac{\tilde{\lambda} e^{\tilde{\lambda} t/2} / \sinh(\tilde{\lambda} t/2)}{\lambda + \tilde{\lambda} \coth(\tilde{\lambda} t/2)} \right]^\alpha \exp \left[ \frac{-2k_{\text{on}}(x + \mu)}{\lambda + \tilde{\lambda} \coth(\tilde{\lambda} t/2)} + k_{\text{on}} \mu t \right],$$

with  $\tilde{\lambda} \equiv \sqrt{\lambda^2 + 2k_{\text{on}}\lambda}$  and  $\alpha = 4\mu$ . At large  $t$ , this is

$$f(x, t) \sim \exp \left[ -\frac{k'_{\text{on}}(x + \mu)}{\lambda} - (2k'_{\text{on}} - k_{\text{on}})\mu t \right], \quad (22)$$

where  $k'_{\text{on}} \equiv 2k_{\text{on}}/(1 + \sqrt{1 + 2\theta'})$  and  $\theta' \equiv k_{\text{on}}/\lambda$ . Thus, the dwell time distribution behaves asymptotically as

$$P(\tilde{\tau} > t) \sim e^{-\tilde{k}_{\text{on}} \mu t}, \quad \text{where } \tilde{k}_{\text{on}} = 2k'_{\text{on}} - k_{\text{on}}. \quad (23)$$

For  $\theta' \ll 1$ , the asymptotic rate above becomes

$$\tilde{k}_{\text{on}} = \left( \frac{3 - \sqrt{1 + 2\theta'}}{1 + \sqrt{1 + 2\theta'}} \right) k_{\text{on}} \simeq \left( \frac{2 - k_{\text{on}}/\lambda}{2 + k_{\text{on}}/\lambda} \right) k_{\text{on}}. \quad (24)$$

The first equality above shows that  $\tilde{k}_{\text{on}}$  can be negative when  $\theta' > 4$  or  $k_{\text{on}} > 4\lambda$ . This is a consequence of the possibility that  $X(t)$  becomes negative under Eq. (19).

When dealing with the Langevin approximation Eq. (17), one usually assumes that the input noise is small enough (given a large copy number) such that the random variable  $X$  could be replaced by its mean  $\mu$  in Eq. (18). This results in an OU approximation:

$$dX(t) = \lambda[\mu - X(t)]dt + \sqrt{2v}dW_t, \quad (25)$$

which takes a Gaussian distribution in steady state with  $\sigma_X^2 = \mu$ . Compared to the shifted Gamma distribution derived from Eq. (19), the Gaussian model of  $X(t)$  has zero skewness and is unbounded below. Thus, when  $\mu$  is small, the OU approximation becomes an inappropriate choice. By the Feynman-Kac theorem, the survival probability  $f(x, t)$  under Eq. (25) satisfies

$$\frac{\partial f}{\partial t} = (v - \lambda x) \frac{\partial f}{\partial x} + v \frac{\partial^2 f}{\partial x^2} - k_{\text{on}} x f, \quad (26)$$

which can also be exactly solved. We omit the solution here, but later will show that  $P[X(t) < 0] > 0$  for the OU process will lead to  $\tilde{k}_{\text{on}} < 0$  in certain regimes.

We can propose an alternative fix for the Langevin approximation, replacing the mean  $\mu$  by its random counterpart  $X(t)$  in Eq. (18). This yields a CIR process:

$$dX(t) = \lambda[\mu - X(t)]dt + \sqrt{2\lambda X(t)}dW_t, \quad (27)$$

under which  $X(t)$  follows a Gamma distribution in steady state, with  $\sigma_X^2 = \mu$  as in all the previous diffusion approximations. The skewness in this model is found to be  $2\mu^{-1/2}$ , which is twice the skewness in the (birth-death) Poisson distribution. However, it is better than the OU model, which gives zero skewness. Figure 4 plots a comparison of the Poisson, Gamma, shifted Gamma, and Gaussian distributions, all satisfying  $\sigma_X^2 = \mu$ . One can see that the shifted Gamma distribution, which follows from Eq. (19), gives the closest approximation to the (birth-death) Poisson distribution, while the Gamma and Gaussian distributions are both good enough to approximate the Poisson. Nonetheless, only the Gamma density from the CIR model is non-negative, like the original Poisson distribution.

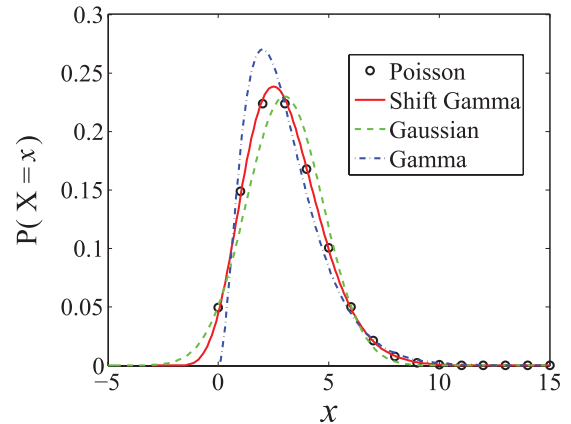


FIG. 4. (Color online) Comparison of the Poisson, “shifted” Gamma, Gaussian, and Gamma distributions, all of which have the same mean  $\mu = 3$  and the same variance  $\sigma_X^2 = \mu$ .

The diffusion approximations we have discussed so far, including Eqs. (1), (19), (25), and (27), are all special cases of the following general Itô SDE:

$$dX(t) = \lambda[\mu - X(t)]dt + \sqrt{\sigma_0^2 + \sigma_1^2 X(t)}dW_t. \quad (28)$$

Under Eq. (28), the survival probability  $f(x, t)$  satisfies

$$\frac{\partial f}{\partial t} = \lambda(\mu - x)\frac{\partial f}{\partial x} + \frac{1}{2}(\sigma_0^2 + \sigma_1^2 x)\frac{\partial^2 f}{\partial x^2} - k_{\text{on}}xf. \quad (29)$$

Again, a shortcut for solving Eq. (29) is to introduce  $X' \equiv X + \sigma_0^2/\sigma_1^2$ , which will evolve as a CIR process. This will allow us to make use of the existing results and obtain a similar expression for  $f(x, t)$  as before. However, our main interest is the effective rate  $\tilde{k}_{\text{on}}$  in the asymptotic behavior of  $P(\tilde{\tau} > t) \simeq \Omega \exp(-\tilde{k}_{\text{on}}\mu t)$ , where  $\Omega$  is some constant. Inspired by Eqs. (7) and (22), we guess that when  $t$  is sufficiently large,  $f(x, t) \sim \exp(-Ax - \tilde{k}_{\text{on}}\mu t)$ , for some constant coefficients  $A$  and  $\tilde{k}_{\text{on}}$  (to be determined). Plugging this expression into Eq. (29) yields

$$-\tilde{k}_{\text{on}}\mu f = -A\lambda(\mu - x)f + \frac{1}{2}A^2(\sigma_0^2 + \sigma_1^2 x)f - k_{\text{on}}xf, \quad (30)$$

which holds only when  $A$  and  $\tilde{k}_{\text{on}}$  jointly solve the following two algebraic equations:

$$A\lambda\mu - \frac{1}{2}A^2\sigma_0^2 = \tilde{k}_{\text{on}}\mu, \quad (31)$$

$$A\lambda + \frac{1}{2}A^2\sigma_1^2 = k_{\text{on}}. \quad (32)$$

Given  $\sigma_1^2 > 0$ , the solution of Eq. (32) is

$$A = \frac{-\lambda + \sqrt{\lambda^2 + 2k_{\text{on}}\sigma_1^2}}{\sigma_1^2} = \frac{2k_{\text{on}}}{\lambda + \sqrt{\lambda^2 + 2k_{\text{on}}\sigma_1^2}}. \quad (33)$$

Thus, by defining  $\theta' \equiv k_{\text{on}}\sigma_1^2/\lambda^2$ , Eq. (33) becomes

$$A\lambda = \frac{2k_{\text{on}}}{1 + \sqrt{1 + 2\theta'}} \equiv k'_{\text{on}}. \quad (34)$$

Eliminating  $A^2$  in Eqs. (31) and (32), we find that

$$\tilde{k}_{\text{on}} = k'_{\text{on}} \frac{\sigma_0^2 + \mu\sigma_1^2}{\mu\sigma_1^2} - k_{\text{on}} \frac{\sigma_0^2}{\mu\sigma_1^2}. \quad (35)$$

Clearly, when  $\sigma_0^2 = 0$ , Eq. (28) becomes a CIR process and Eq. (35) reduces to  $\tilde{k}_{\text{on}} = k'_{\text{on}} = 2k_{\text{on}}/(1 + \sqrt{1 + 2\theta'})$ , recovering our result in Sec. III. When  $\sigma_0^2/\sigma_1^2 = \mu$ , Eq. (35) is  $\tilde{k}_{\text{on}} = 2k'_{\text{on}} - k_{\text{on}}$ , coinciding with Eq. (23).

Finally, if  $\sigma_1^2 = 0$ , the diffusion Eq. (28) becomes an OU process and  $A = k_{\text{on}}/\lambda$  by Eq. (32). In this case,

$$\tilde{k}_{\text{on}} = k_{\text{on}} \left( 1 - \frac{k_{\text{on}}\sigma_0^2}{2\mu\lambda^2} \right) = k_{\text{on}} \left[ 1 - \frac{k_{\text{on}}\mu}{\lambda} \left( \frac{\sigma_X^2}{\mu^2} \right) \right], \quad (36)$$

where  $\sigma_X^2 \equiv \sigma_0^2/(2\lambda)$  is the variance of the OU process. We can define  $\theta$  in the same way as in Eq. (12), such that  $\tilde{k}_{\text{on}} = k_{\text{on}}(1 - \theta/2)$  in Eq. (36). This becomes negative if  $\theta > 2$ . It is again a consequence of the finite probability that  $X(t)$  takes negative values. Note that the OU approximation Eq. (25) is a particular OU process with  $\sigma_X^2 = \mu$ . This leads to  $\tilde{k}_{\text{on}} = k_{\text{on}}(1 - k_{\text{on}}/\lambda)$  in Eq. (36). So for small  $\theta$  or  $\lambda \gg k_{\text{on}}$ , we have  $\tilde{k}_{\text{on}} \simeq k_{\text{on}}/(1 + k_{\text{on}}/\lambda)$ , consistent with our result in [22].

In sum, the general diffusion process defined by Eq. (28) may take negative values with finite probability unless  $\sigma_0^2 = 0$ , which corresponds to the CIR process. Negative biochemical input is unrealistic and may lead to unphysical results (such as  $\tilde{k}_{\text{on}} < 0$ ) if the input noise is large. For this reason, we conclude that the CIR process Eq. (1) is more suitable for modeling biochemical noise in the continuous framework. As shown, it is analytically tractable and possesses desirable statistical features, including stationarity, mean reversion, Gamma distribution, and a tunable Fano factor.

## V. CONCLUSION

In this paper, we have extended our previous work on the role of noise in biological switching systems. We propose that a square-root diffusion process can be a more reasonable model for biochemical fluctuations than the commonly used OU process. We employ standard tools in stochastic processes to solve a well-defined fundamental biophysical problem. Consistent with our earlier results, we find that the input noise acts to suppress the input-dependent transitions of the switch. Our analytical results in this paper indicate that this suppression increases with the input noise level as well as the input correlation time. The statistical features uncovered in this basic problem can provide us with insights to understand various experimental observations in gene regulation and signal transduction systems. The current modeling framework may also be generalized to incorporate other biological features such as ultrasensitivity and feedbacks. Work along these lines is underway.

## ACKNOWLEDGMENTS

We would like to thank Ruth J. Williams, Yuhai Tu, Jose Onuchic, and Wen Chen for stimulating discussions. This work has been supported by the NSF-sponsored CTBP Grant No. PHY-0822283.

- [1] C. V. Rao, D. M. Wolf, and A. Arkin, *Nature (London)* **420**, 231 (2002).  
 [2] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. D. Swain, *Science* **207**, 1183 (2002).  
 [3] P. S. Swain, M. B. Elowitz, and E. D. Siggia, *Proc. Natl. Acad. Sci. USA* **99**, 12795 (2002).

- [4] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, *Nature (London)* **422**, 633 (2003).  
 [5] M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins, *Nat. Rev. Genetics* **6**, 451 (2005).  
 [6] J. M. Raser and E. K. O'Shea, *Science* **309**, 2010 (2005).

- [7] J. M. Pedraza and A. van Oudenaarden, *Science* **307**, 1965 (2005).
- [8] L. Cai, N. Friedman, and X. S. Xie, *Nature (London)* **440**, 358 (2006).
- [9] N. Friedman, L. Cai, and X. S. Xie, *Phys. Rev. Lett.* **97**, 168302 (2006).
- [10] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, *Science* **322**, 442 (2008).
- [11] A. Eldar and M. B. Eolwitz, *Nature (London)* **467**, 167 (2010).
- [12] J. Paulsson, *Nature (London)* **427**, 415 (2004).
- [13] M. L. Simpson, C. D. Cox, and G. S. Saylor, *J. Theor. Biol.* **229**, 383 (2004).
- [14] W. Bialek and S. Setayeshgar, *Proc. Natl. Acad. Sci. USA* **102**, 10040 (2005).
- [15] W. Bialek and S. Setayeshgar, *Phys. Rev. Lett.* **100**, 258101 (2008).
- [16] G. Tkačik, T. Gregor, and W. Bialek, *PLoS ONE* **3**, e2774 (2008).
- [17] G. Tkačik and W. Bialek, *Phys. Rev. E* **79**, 051901 (2009).
- [18] T. Shibata and K. Fujimoto, *Proc. Natl. Acad. Sci. USA* **102**, 331 (2005).
- [19] S. Tănase-Nicola, P. B. Warren, and P. Reintjes Wolde, *Phys. Rev. Lett.* **97**, 068102 (2006).
- [20] F. Tostevin and P. Reintjes Wolde, *Phys. Rev. Lett.* **102**, 218101 (2009).
- [21] E. Levine and T. Hwa, *Proc. Natl. Acad. Sci. USA* **104**, 9224 (2007).
- [22] B. Hu, D. A. Kessler, W.-J. Rappel, and H. Levine, *Phys. Rev. Lett.* **107**, 148101 (2011).
- [23] B. Hu, W. Chen, W.-J. Rappel, and H. Levine, *Phys. Rev. E* **83**, 021917 (2011).
- [24] B. Hu, W. Chen, W.-J. Rappel, and H. Levine, *J. Stat. Phys.* **142**, 1167 (2011).
- [25] B. Hu, W. Chen, W.-J. Rappel, and H. Levine, *Phys. Rev. Lett.* **105**, 048104 (2010).
- [26] B. Hu, D. Fuller, W. F. Loomis, H. Levine, and W.-J. Rappel, *Phys. Rev. E* **81**, 031906 (2010).
- [27] K. Wang, W.-J. Rappel, R. Kerr, and H. Levine, *Phys. Rev. E* **75**, 061905 (2007).
- [28] W.-J. Rappel and H. Levine, *Phys. Rev. Lett.* **100**, 228101 (2008).
- [29] W.-J. Rappel and H. Levine, *Proc. Natl. Acad. Sci. USA* **105**, 19270 (2008).
- [30] D. Fuller, W. Chen, M. Adler, A. Groisman, H. Levine, W.-J. Rappel, and W. F. Loomis, *Proc. Natl. Acad. Sci. USA* **107**, 9656 (2010).
- [31] R. G. Endres and N. S. Wingreen, *Proc. Natl. Acad. Sci. USA* **105**, 15749 (2008).
- [32] R. G. Endres and N. S. Wingreen, *Phys. Rev. Lett.* **103**, 158101 (2009).
- [33] P. Cluzel, M. Surette, and S. Leibler, *Science* **287**, 1652 (2000).
- [34] E. A. Korobkova, T. Emonet, J. M. G. Vilar, T. S. Shimizu, and P. Cluzel, *Nature (London)* **428**, 574 (2004).
- [35] E. A. Korobkova, T. Emonet, H. Park, and P. Cluzel, *Phys. Rev. Lett.* **96**, 058105 (2006).
- [36] T. Emonet and P. Cluzel, *Proc. Natl. Acad. Sci. USA* **105**, 3304 (2008).
- [37] Y. Tu and G. Grinstein, *Phys. Rev. Lett.* **94**, 208101 (2005).
- [38] Y. Tu, *Proc. Natl. Acad. Sci. USA* **105**, 11737 (2008).
- [39] T. S. Shimizu, Y. Tu, and H. C. Berg, *Mol. Syst. Biol.* **6**, 382 (2010).
- [40] N. Bostani, D. A. Kessler, N. M. Shnerb, W. J. Rappel, and H. Levine, *Phys. Rev. E* **85**, 011901 (2012).
- [41] J. E. M. Hornos, D. Schultz, G. C. P. Innocentini, J. Wang, A. M. Walczak, J. N. Onuchic, and P. G. Wolynes, *Phys. Rev. E* **72**, 051907 (2005).
- [42] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985).
- [43] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007).
- [44] B. K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer, Berlin, 2000).
- [45] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *Econometrica* **53**, 385 (1985).
- [46] L. Pechenik and H. Levine, *Phys. Rev. E* **59**, 3893 (1999).
- [47] T. Shibata, *Phys. Rev. E* **67**, 061906 (2003).
- [48] S. Azaele, J. R. Banavar, and A. Maritan, *Phys. Rev. E* **80**, 031916 (2009).