Noise and critical phenomena in biochemical signaling cycles at small molecule numbers

C. Metzner,* M. Sajitz-Hermstein, M. Schmidberger, and B. Fabry

Biophysics Group, Department of Physics, University of Erlangen, Henkestrasse 91, D-91052 Erlangen, Germany

(Received 21 April 2009; published 17 August 2009)

Biochemical reaction networks in living cells usually involve reversible covalent modification of signaling molecules, such as protein phosphorylation. Under conditions of small molecule numbers, as is frequently the case in living cells, mass-action theory fails to describe the dynamics of such systems. Instead, the biochemical reactions must be treated as stochastic processes that intrinsically generate concentration fluctuations of the chemicals. We investigate the stochastic reaction kinetics of covalent modification cycles (CMCs) by analytical modeling and numerically exact Monte Carlo simulation of the temporally fluctuating concentration. Depending on the parameter regime, we find for the probability density of the concentration qualitatively distinct classes of distribution functions including power-law distributions with a fractional and tunable exponent. These findings challenge the traditional view of biochemical control networks as deterministic computational systems and suggest that CMCs in cells can function as versatile and tunable noise generators.

I. INTRODUCTION

Living cells transduce chemical signals from the environment via trans-membrane receptors to their interior. The activated receptors trigger chains of chemical reactions along the so-called signaling pathways, which can, for example, lead to the expression of selected genes in response to the external stimulus. Complex reaction networks arise when several linear pathways are cross-linked by multiple biochemical interactions. Such signal transduction networks are traditionally thought of as deterministic "computers," in which information is coded by the relative concentration of biochemicals. This study challenges this view and suggests that stochastic concentration fluctuations are the primary mode of operation for most intracellular signaling cascades.

It is well known that the numbers of receptors and signaling molecules fluctuate as a function of time and from cell to cell [1]. The role of these fluctuations, often regarded as noise, is still poorly understood. How can cells properly react to external stimuli when the signals have to pass through noisy channels? Is the degree of noise actively suppressed for certain key signaling proteins? Or is the present understanding of intracellular control, based on mass-action theory, overly simplified?

Using covalent modification cycles (CMCs) as a simple model system, we show that the magnitude of concentration fluctuations, relative to the mean value, can indeed be enormous. We demonstrate that CMCs can be viewed as versatile and tunable noise generators. Depending on the system parameters, qualitatively different classes of probability density functions (PDFs) of concentration fluctuations emerge, including extremely broad and asymmetric distributions with fractional power-law tails. CMCs are a very common motif in cellular reaction networks [2–8]. The typical structure of a CMC is shown in Fig. 1.

In such systems, a substrate protein is found in two different chemical states: an inactive form X_0 and an activated form X (often a phosphorylized version of X_0). The conversion of the two forms into each other is provided by an activating enzyme A (often a kinase) and the deactivation by another enzyme D (often a phosphatase). In the activation process, the catalyst A first binds its substrate X_0. The resulting enzyme-substrate complex AX_0 may decay back into the original components. In the case of a successful conversion, however, a product molecule X is released and the enzyme A is recovered for further use. The deactivation process is analogous.

The CMC can be functionally decomposed into two enzymatic conversion processes. According to Michaelis-Menten kinetics (compare Appendix, Sec. 1), the conversion rate is, in the linear regime, limited by the amount of available substrate. For very high substrate concentration, however, the conversion rate approaches a maximum value, determined only by the amount and efficiency of the enzyme (saturation regime).

As demonstrated in a classical paper by Goldbeter *et al.* [9], the combination of the two enzymatic conversion reactions can lead to interesting behavior if they operate within the saturated regime. In this case, the equilibrium ratio $[X]/[X_0]$ as a function of the ratio of enzyme levels $[A]/[D]$ develops a sigmoidal shape with a sharp transition point (zero-order ultrasensitivity). In the context of biochemical signal networks, CMCs are for this reason understood as switches.
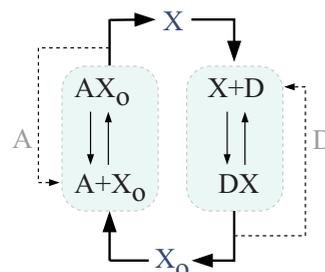
FIG. 1. (Color online) Schematic of a CMC. Substrate X_0 is activated by enzyme A into the modified form X and deactivated by enzyme D. Each shaded submodule denotes an enzymatic conversion reaction (unbinding reactions not shown).

*claus.metzner@gmx.net

The Goldbeter-Koshland theory is based on deterministic (mass-action) rate equations and thus disregards fluctuations entirely. Molecular reactions, however, inevitably generate intrinsic noise due to their discrete and stochastic nature. Even under so-called steady-state conditions, the momentary rates at which reactions proceed are fluctuating around the mean values described by mass-action theory. The corresponding temporal fluctuations of molecule numbers are particularly important in living cells, where the average molecule numbers of many chemical species are low. For this reason, quantitative models of biochemical concentration fluctuations are developed for different types of reaction networks (see, for example, Refs. [10–12]).

Due to their ubiquity in living cells, CMCs are of particular interest. A detailed theoretical investigation of the intrinsic fluctuations of CMCs, their robustness and tunability, was provided by Levine *et al.* [13], who directly solved the master equation for the probability distribution of the number of activated signal molecules. The authors further considered the information transmission properties of the system in the presence of the intrinsic fluctuations by applying a pulselike increase of the kinase activity as an input signal. They find that the noisy CMC can transmit the signal reliably if tuned to an optimal parameter range.

In this paper, we focus on the shape of the stationary probability distributions produced by CMCs in various parameter regimes. The reaction kinetics of this system is simulated using the exact Gillespie algorithm. This simulation yields directly the temporal concentration fluctuations $x(t)$ of the activated signaling molecule.

We find an unexpected variety of distribution functions $P(x)$, including Gaussian, exponential, flat, as well as power-law distributions with a fractional and tunable exponent. The type of the emerging distribution function depends on parameters such as the total amount of available enzyme and substrate molecules in their different forms and on reaction-rate coefficients. We speculate that living cells could switch between distinct statistical distributions, on short-time scales, by controlling the overall expression levels of these molecules. In many cases, moreover, the enzymes of a CMC are themselves activated and deactivated by another cycle. In this way, the effective conversion efficiency of an enzyme can be tuned over a wide range with only minimal changes of protein expression levels. This tremendous flexibility of CMCs with respect to their statistical properties suggests a more complex picture of cellular signal processing which is based on the active generation and precise shaping of concentration fluctuations of signaling molecules.

In our paper, we develop analytical approximations of the concentration fluctuations within CMCs based on stochastic differential equations and explicit stationary solutions of the corresponding Fokker-Planck equations. The analytical results are in excellent agreement with the simulations and provide a quantitative understanding of the major statistical features.

II. MODELS AND METHODS

A. Model parameters and assumptions

Let the reactions take place in a container of volume V so that the concentration $[S]$ of a substance corresponds to a molecule number $s=[S]V$. We also assume that the reactor is "well stirred," i.e., diffusion of chemical species is infinitely fast and so spatial effects are disregarded.

We study a CMC of the form

$$X_0 + A \underset{u_1}{\overset{b_1}{\rightleftharpoons}} AX_0 \overset{c_1}{\rightarrow} X + A,$$

$$X + D \underset{u_2}{\overset{b_2}{\rightleftharpoons}} DX \overset{c_2}{\rightarrow} X_0 + D. \tag{1}$$
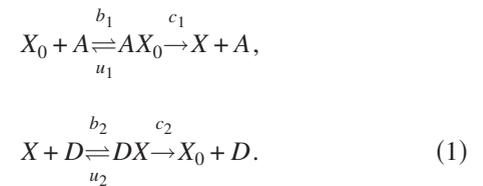
The substrate X_0 is converted into its activated form X by enzyme A. The corresponding deactivation is performed by enzyme D. We thus have to consider six temporally variable molecule numbers, x_0, x, a, d, ax_0, dx, dynamically coupled by six chemical reactions. Within each enzymatic conversion unit, the three reaction coefficients are denoted as b (binding), u (unbinding), and c (conversion). Index 1 is used for the activation and index 2 for the deactivation process. Additional parameters are the total amount of the substrate in its various forms, $x_t = x_0 + ax_0 + x + dx$, as well as the total amounts of enzymes $a_t = a + ax_0$ and $d_t = d + dx$.

B. Analytical and numerical methods

Analytical approximations for chemical reaction networks can be obtained by deriving Langevin equations for the temporal changes of the molecule numbers. These stochastic differential equations contain, besides a deterministic term that corresponds to the mass-action change rates, a stochastic term that accounts for the fluctuations. To make use of the standard methods of stochastic differential calculus, the fluctuation term is approximated by a Gaussian, white-noise random process. This is a critical approximation, since the effective "strength" of the white-noise process has to be chosen with care in order to reflect the true process as faithfully as possible. In the case of chemical Langevin equations, the true process consists of a series of delta peaks arriving with (inhomogeneous) Poisson statistics. It is therefore possible to derive the proper strength of the white-noise process from the fundamental properties of Poisson statistics. This theory of chemical Langevin equations has been developed, for the general case, by Gillespie [14]. In this paper, we take a similar approach, suitable for our specific reaction network.

In order to test our analytical approximations, we shall compare the results to a numerically exact Monte Carlo simulation of the reaction dynamics by implementing the Gillespie algorithm [15]. In this algorithm, the molecule numbers of each species are integers which change abruptly due to elementary reaction events. Statistically, these elementary reactions are Poisson processes with average event rates depending on the momentary molecule numbers according to the chemical rate equations. Therefore, the intrinsic stochastic fluctuations of the reactions are automatically included in a realistic way.

C. Coarse graining of the enzymatic conversion

We first focus on a single enzymatic conversion reaction, for example, the activation process. Our goal is to describe it

in a coarse-grained approximation as a single functional unit with effective statistical properties. Two of these effective units will later be combined (as shown in Fig. 1) to derive a stochastic differential equation for the fluctuating number $x(t)$ of X molecules.

We assume for a moment that the number x_0 of substrate molecules X_0 is constant (ideal reservoir). We are then interested in the average production rate $\bar{R}_{act}(x_0)$ of the activated protein X and in the temporal fluctuations $\Delta R_{act}(x_0,t)$ of this rate. This, in turn, will enable us to write a stochastic rate equation of the production process in the form $\dot{x} = \bar{R}_{act}(x_0) + \Delta R_{act}(x_0,t)$.

As for the average rates, we solve the mass-action rate equations in the stationary state. This follows standard Michaelis-Menten theory, but for completeness, we include the derivation in Appendix, Sec. 1. The result is

$$\bar{R}_{act}(x_0) = v_m \frac{x_0}{x_0 + k_m}, \qquad (2)$$

with the maximum conversion velocity

$$v_m = c a_t \qquad (3)$$

and the Michaelis constant

$$k_m = \frac{c + u}{b}. \qquad (4)$$

Enzymes are sometimes likened to nanomachines, which convert their substrates in a predictable, goal-oriented process. Yet, many enzymes in biological systems are working in a much more imperfect way. Once the enzyme has bound to its substrate, the enzyme-substrate complex often dissociates back into the original two molecules. Each individual enzyme molecule will go through a series of futile binding-dissociation cycles before it actually converts a substrate into the modified form. In the chemical reaction equation (1), this is accounted for by the back reaction with rates u_j (with $j = 1, 2$). The conversion efficiency of an enzyme can be quantified by the fraction of binding events that lead to a successful production and release of the modified substrate molecule. This fraction, in turn, depends on the relative magnitude of the rates u_j and c_j. We can define two limiting regimes. The case $u_j \gg c_j$ corresponds to extremely inefficient enzymes. In the diagram of Fig. 1, almost all activity of the reaction system will then take place within the shaded submodules. The fluxes in and out of these submodules are so weak that within the submodules, a chemical equilibrium is established between the bound and dissociated enzyme-substrate complexes. We therefore call this case the "pre-equilibrium" regime. The opposite case, $c_j \gg u_j$, corresponds to highly efficient enzymes. In the diagram of Fig. 1, the system is running unidirectionally around the cycle for most of the time. We therefore call this case the "sequential" regime.

Independently from u and c, two other limiting regimes are connected with the amount of substrate x_0 relative to the Michaelis constant k_m. The system is in the "linear" regime for $x_0 \ll k_m$ and in the "saturation" regime for $x_0 \gg k_m$.

Next, we model the fluctuations $\Delta R_{act}(x_0,t)$ of the production rate around the average value $\bar{R}_{act}(x_0)$. The statistical properties of these fluctuations are not obvious even if the substrate molecule number x_0 is artificially held constant. As motivated in Appendix, Sec. 2, we approximate the production process, in a coarse-grained view, as a Poisson process with average event rate $\bar{R}_{act}(x_0)$. Numerical simulations, shown below, confirm that the probability distribution of the waiting time between successive X-production events is indeed exponentially distributed with the expected characteristic time constant. We further approximate the above Poisson process by white Gaussian noise with a proper prefactor (Appendix, Sec. 3). As a result of the above coarse-graining procedure, we obtain

$$\dot{x} = \bar{R}_{act}(x_0) + \sqrt{\bar{R}_{act}(x_0)} \zeta(t), \qquad (5)$$

where $\zeta(t)$ is normalized white Gaussian noise with $\langle \zeta(t)\zeta(t') \rangle = \delta(t - t')$.

D. Stochastic differential equation of a CMC

We next combine the activation and deactivation processes. The molecule numbers $x(t)$ and $x_0(t)$ are now both considered as variables. One obtains

$$\dot{x} = [\bar{R}_{act}(x_0) - \bar{R}_{dea}(x)] + [\sqrt{\bar{R}_{act}(x_0)}\zeta_a(t) + \sqrt{\bar{R}_{dea}(x)}\zeta_d(t)]. \qquad (6)$$

Note that the deactivation rates depend on x, not x_0. To make further progress, we neglect the amount of substrates bound within enzyme-substrate complexes so that $x_0 = x_t - x$. Additionally, we make the simplifying assumption that the noise terms of the activation and deactivation processes fluctuate statistically independent from each other. We can then combine both terms, adding up the variances,

$$\dot{x} = [\bar{R}_{act}(x_t - x) - \bar{R}_{dea}(x)] + [\sqrt{\bar{R}_{act}(x_t - x) + \bar{R}_{dea}(x)}]\zeta(t). \qquad (7)$$

This has the general form of a stochastic differential equation with a multiplicative noise term

$$\dot{x} = f(x) + g(x)\zeta(t). \qquad (8)$$

Here,

$$f(x) = v_a \frac{(x_t - x)}{(x_t - x) + k_a} - v_d \frac{x}{x + k_d} \qquad (9)$$

and

$$g(x) = \sqrt{v_a \frac{(x_t - x)}{(x_t - x) + k_a} + v_d \frac{x}{x + k_d}}, \qquad (10)$$

with obvious definitions of v_a, v_d, k_a, k_d. In the following, we will extract statistical properties of this random process. Note that the Ito interpretation has to be used whenever the true random process (that is to be approximated by Gaussian white noise) consists of a series of δ peaks, such as in our case of intrinsic, chemical noise [16,17].

We define a drift term

$$A(x) = f(x) \tag{11}$$

and a diffusion term

$$B(x) = \frac{1}{2}g^2(x). \tag{12}$$

The time-dependent PDF $P(x,t)$ of the fluctuating variable $x(t)$ approximately satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial t}P(x,t) = -\frac{\partial}{\partial x}[A(x)P(x,t)] + \frac{\partial^2}{\partial x^2}[B(x)P(x,t)]. \tag{13}$$

The stationary solution $P(x)$ of this equation reads

$$P(x) = \frac{N}{B(x)}\exp\left[\int_{x_{\min}}^{x}\frac{A(s)}{B(s)}ds\right]. \tag{14}$$

Here, N is a normalization constant.

E. Symmetric CMC

With v_a, v_d, k_a, k_d, and x_t, there is obviously a large parameter space to explore. In this paper, we shall restrict ourselves to just a few interesting cases. In a symmetric CMC, the activation and deactivation processes have the same parameters, i.e., $v_a = v_d = v$ and $k_a = k_d = k$. We then have

$$f(x) = v\left[\frac{(x_t - x)}{(x_t - x) + k} - \frac{x}{x + k}\right] \tag{15}$$

and

$$g^2(x) = v\left[\frac{(x_t - x)}{(x_t - x) + k} + \frac{x}{x + k}\right]. \tag{16}$$

Because the drift term $A(s)$ and the diffusion term $B(s)$ are both proportional to v, it is clear that the maximum production rate v will not affect the shape of the stationary PDF. Consequently, k and x_t are the only important parameters left.

1. Linear regime

The limit of a large Michaelis constant, $k \gg x_t$, corresponds to the linear regime of the two enzymatic conversion reactions. In this case, the terms x and $(x_t - x)$ can be neglected in Eqs. (9) and (10). This leaves us with

$$f(x) = (vx_t/k) - (2v/k)x \tag{17}$$

and

$$g^2(x) = (vx_t/k). \tag{18}$$

A straightforward calculation of the PDF yields a Gaussian, centered at $\bar{x} = \frac{x_t}{2}$, with a variance $\sigma_x^2 = \frac{x_t}{4}$,

$$P(x) \propto \exp\left\{-\frac{2[x-(x_t/2)]^2}{x_t}\right\}. \tag{19}$$

The stochastic differential equation of a symmetric, linear CMC corresponds to an Ornstein-Uhlenbeck process. Besides the Gaussian PDF, we therefore expect an exponentially decaying autocorrelation function

$$C_{xx}(\tau) = \langle\Delta x(\tau)\Delta x(0)\rangle = \left(\frac{x_t}{4}\right)e^{-(2v/k)\tau}. \tag{20}$$

The characteristic time constant is $\tau_c = \frac{k}{2v}$.

2. Saturation regime

Next, we consider the opposite case of a small Michaelis constant, i.e., $k \ll x_t$, corresponding to the saturation regime. We then have

$$f(x) = v\left[1 - \frac{x}{x + k}\right] \rightarrow \frac{vk}{x} \quad \text{for } x \gg k \tag{21}$$

and

$$g^2(x) = v\left[1 + \frac{x}{x + k}\right] \rightarrow 2v \quad \text{for } x \gg k. \tag{22}$$

The asymptotic drift and diffusion terms are $A(x) = \frac{vk}{x}$, $B(x) = v$, and $A(x)/B(s) = \frac{k}{x}$. Therefore,

$$\int_{x_{\min}}^{x}\frac{A(s)}{B(s)}ds = k\log(x/x_{\min}) \tag{23}$$

and

$$P(x) \propto e^{k\log(x/x_{\min})} \propto (x/x_{\min})^k. \tag{24}$$

Hence, we expect an increasing power-law tail for the asymptotic PDF in the saturation regime of the symmetric CMC. The exponent of the power law can be fractional and is equal to the dimensionless Michaelis constant [Eq. (4)]. The above analytical approximations will break down when x approaches the limits 0 or x_t.

F. Asymmetric CMC

We now allow the activation parameters k_a and v_a to differ from the corresponding deactivation parameters k_d and v_d. Under saturation conditions ($x_t \gg k_a$, $x_t \gg k_d$) and in the limit of large x, one obtains $f(x) \rightarrow (v_a - v_d)$ and $g^2(x) \rightarrow (v_a + v_d)$, so that

$$\frac{A(s)}{B(s)} \rightarrow \lambda = 2\frac{v_a - v_d}{v_a + v_d}. \tag{25}$$

This results in a stationary PDF with an exponential tail

$$P(x) \propto e^{\lambda x}. \tag{26}$$

The decay constant λ is positive for $v_a > v_d$ and negative for $v_a < v_d$.

III. RESULTS

A. Validation of Poisson statistics

We first investigate the statistics of the enzymatic activation process, with artificially fixed number x_0 of substrate molecules. For this purpose, we perform direct Monte Carlo simulations in different parameter regimes. All rates and times are presented in dimensionless numbers.
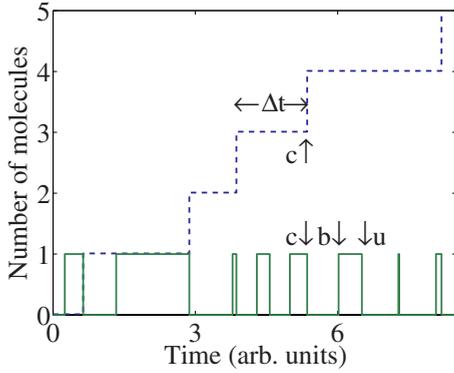
FIG. 2. (Color online) Monte Carlo simulation of enzymatic conversion. Molecule numbers of the enzyme-substrate complex (solid) and of the activated product (dashed) in the case of only one enzyme molecule. Parameters: $b=u=c=1.0$, $e_t=1$. Vertical arrows denote conversion (c), binding (b), and unbinding (u) processes. Δt is the time interval between two successive conversion events.

The stochastic time evolution of the enzymatic activation process is characterized by abrupt changes of the various molecule numbers by integer amounts (Fig. 2). A single enzyme molecule sometimes undergoes binding (b) and unbinding (u) without conversion (c) to a product molecule. The time interval Δt between two successive conversion events is fluctuating around the inverse of the average production rate.

Since the production process involves a sequence of elementary reaction steps, the distribution function $P(\Delta t)$ of this waiting time is not expected to be exponential for an individual enzyme molecule. However, the superposition of many such multistep processes running independently from each other can closely mimic a Poisson process (Fig. 3).

B. Monte Carlo simulation of the CMC

Next, we discuss the statistical properties of CMCs in selected parameter regimes, as obtained by Monte Carlo simulation of the reaction dynamics. We shall mainly focus on CMCs with symmetric parameters for the activation and

TABLE I. Parameter space explored in Monte Carlo simulations.

Fig.	Subf.	b	u	c	Asym.	k_m	SP	EP
4,5	b	$5\cdot10^{-4}$	0.1	10		$2\cdot10^4$	0.01	0.01
	c	$5\cdot10^{-4}$	10	0.1		$2\cdot10^4$	0.01	100
	d	0.5	0.1	10		20	10	0.01
	e	0.5	10	0.1		20	10	100
6	a	1	0.1	2		2.1	95	0.05
	b	1	0.1	1.5		1.6	125	0.07
	c	1	0.1	1		1.1	180	0.1
	d	1	0.1	0.4		0.5	400	0.25
	e	1	0.1	0.1		0.2	1000	1
7	a	1	0.1	1	$c_1=1.5$			
	b	1	0.1	1	$c_1=1.25$			
	c	1	0.1	1	$c_1=1.125$			
	d	1	0.1	1		1.1	180	0.1

deactivation processes. The total number of substrate molecules x_t was 200 in all cases. Our analytic theory was based on the assumption that the amount of substrate bound in complexes is small compared to x_t. We have therefore chosen a small number of enzyme molecules, $a_t=d_t=10$. The (rounded) parameters for all the following simulations are listed in Table I. We have also included a saturation parameter (SP), defined as $SP=x_t/k_m$, and an equilibrium parameter (EP), defined as $EP=c/u$. For instance, $SP \gg 1$, $EP \gg 1$ would indicate that the system is in the saturated, pre-equilibrium regime.

1. Symmetric CMC in the linear and weakly saturated regimes

In the linear regime, we expect for the substrate X a Gaussian distribution, peaked at $\bar{x}=x_t/2$ and with variance $x_t/4$. The autocorrelation of the random variable $x(t)$ should decay exponentially with time constant $\tau_c=k/2v$. The agreement of the Monte Carlo results with this analytic theory is excellent (see Fig. 4). In the weakly saturated regime, we find a decrease of the average molecule number and a con-

FIG. 3. (Color online) Waiting time distributions. Simulated PDF of the time intervals between successive conversion events. Parameters: $b=u=c=1.0$, $x_0=1=$const. Case (a): Only one enzyme molecule. Case (b): Ten independent enzyme molecules. Inset shows the same data in a semilogarithmic plot.
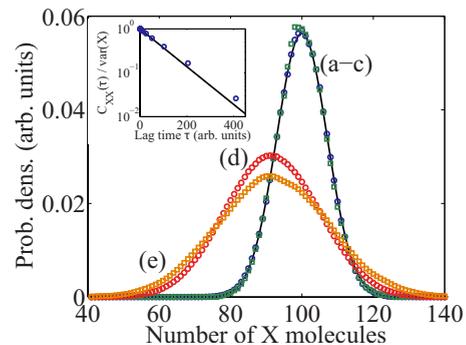
FIG. 4. (Color online) X distributions in the [(b) and (c)] linear regime and in the [(d) and (e)] weakly saturated regime. Solid line (a) is the analytical solution to the linear case. Inset: Normalized autocorrelation function for linear case (symbols) with analytical solution (solid line). For parameters, see Table I.
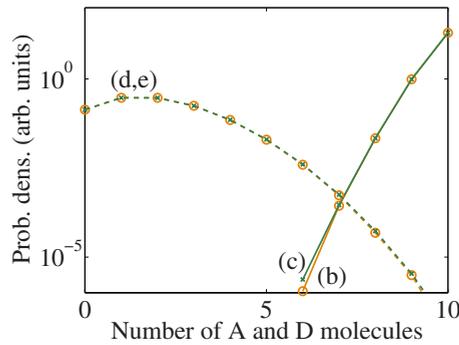
FIG. 5. (Color online) Semilogarithmic plots of the simulated enzyme distributions in the [(b) and (c)] linear regime and in the [(d) and (e)] weakly saturated regime. For parameters, see Table I.

siderable broadening of the distribution, while the shape of the PDF remains approximately Gaussian. The distributions do not change dramatically when the parameter regime is changed from sequential to pre-equilibrium conditions as long as the ratio of enzyme to substrate molecules is small (see [29]).

The Monte Carlo simulations also yield the distributions of the enzyme molecule number $E=a=d$ (see Fig. 5). The effect of sequential or pre-equilibrium conditions is almost invisible for the particular parameters chosen.

2. Symmetric CMC in the saturation regime

Next, we turn to CMCs operating within the saturation regime, which corresponds to the hypersensitive "switchlike" mode of the cycle. In the simulations, k_m was indirectly changed via the conversion rate c. While small conversion rates result in a Gaussian PDF, the distributions become extremely asymmetric as the system runs into the saturation regime (Fig. 6). The double-logarithmic plot reveals a power-law wing at the "left" side of the peak. The positive exponent of the power-law tail is fractional in the general case. It is determined by the Michaelis constant, as expected from the analytical theory above. For a very small Michaelis constant, one obtains an almost flat distribution, which can cover several decades of concentration. Of course, the PDF
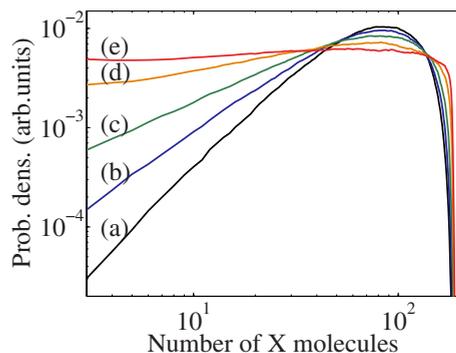
FIG. 6. (Color online) Power-law tails in the strongly saturation regime. Double-logarithmic plots of the X distribution. In cases (a)–(e), the conversion rate c has been gradually increased. For parameters, see Table I.
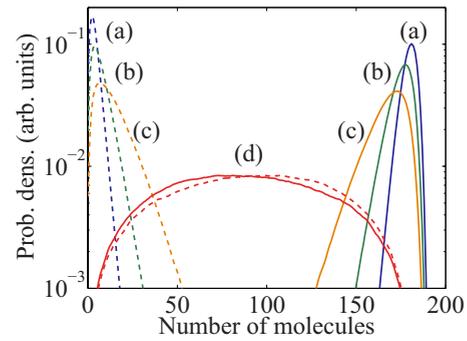
FIG. 7. (Color online) Collapse of exponential distributions at the critical point of parametric symmetry. Distributions of X_0 (dashed lines) and X (solid lines). In cases (a)–(d), the conversion rate of the activating reaction only has been gradually increased. The left wing of (d) corresponds to Fig. 6(c) when plotted double logarithmically. For parameters, see Table I.

has sharp cutoffs at the maximum particle number $x=x_t$ and close to $x=0$ (not shown).

This remarkable result demonstrates that the notion of deterministic biomolecular networks, with well-defined average levels of concentration and negligibly small Gaussian fluctuations, dramatically fails in certain parameter ranges. Concentration fluctuations with a power-law wing are scale free and therefore arbitrarily large deviations from the average value occur with nonnegligible probability.

3. Asymmetric CMC in the saturation regime

From a systems biology point of view, an interesting question is the sensitivity of the CMC with respect to its parameter values. In particular, we investigated the effects of tuning the system slightly away from the completely symmetric parameter settings considered so far. The most dramatic effects are expected for a CMC in the hypersensitive saturation regime.

For this purpose, we start again with the parameters of the symmetric saturated CMC, which produced a PDF with a power-law tail of slope 1.1 [compare Fig. 6(e)]. Now, however, we fine tune the conversion rate c_1 of the activation reaction, while leaving the corresponding parameter c_2 at its former value 1.

As expected, if $c_2 < c_1$, the PDF of X_0 is peaked around a small average concentration, while X has a high average concentration (Fig. 7). The average concentrations are drastically different even for rather similar c parameters due to the hypersensitive response of the saturated CMC. We find PDFs with exponential tails for all cases, except in a very narrow range around perfect parametric symmetry. This is in agreement with the analytical theory presented in Sec. II F. In the narrow symmetrical regime, the two PDFs collapse to one. They are mirror symmetric with respect to the average molecule number in this case.

This behavior somewhat resembles critical phenomena in physics, where fluctuations of arbitrary size occur when a control parameter is precisely tuned to a critical value. In biological systems, it would be extremely improbable to find a CMC where all the microscopic parameters of the activa-

tion and deactivation reaction are precisely identical. However, Eqs. (9) and (10) show that effective dynamical symmetry can be achieved under the much weaker conditions $v_a \approx v_d$ and $k_a \approx k_d$. In terms of the microscopic parameters, this translates into $c_1 a_t \approx c_2 d_t$ and $\frac{c_1 + u_1}{b_1} \approx \frac{c_2 + u_2}{b_2}$.

In order to demonstrate that the dynamics is only controlled by the conversion velocities, the Michaelis constants, and the total amount of substrate x_t, we have performed another Monte Carlo simulation for a CMC with $x_t = 200$ and the microscopically nonsymmetric parameters $a_t = 50$, $b_1 = 1$, $u_1 = 0.1$, $c_1 = 1$ for the activation reaction and $d_t = 10$, $b_2 = 10$, $u_2 = 6$, $c_2 = 5$ for the deactivation reaction. These parameters are nevertheless symmetrical on the coarse-grained level of k and v. The simulation results indeed show a power-law behavior, thus confirming the analytical prediction (see Fig. 8). Note that the total amount of enzymes a_t and d_t can be easily varied in a living cell, for example, by changing the expression levels or the activity of the enzymes. This offers a way to tune the CMC through the critical point. If, for instance, we detune a_t away from the critical value $a_t^{(crit)} = 50$ by $\pm 10\%$, we find that one of the distributions $P(X)$ and $P(X_0)$ is loosing its power-law behavior. Yet, the (respective) complementary form of substrate still shows a very steep power-law tail under these conditions of disturbed symmetry.

IV. DISCUSSION AND OUTLOOK

The statistical properties of concentration fluctuations produced by CMCs reveal an extremely rich behavior. A variety of qualitatively different probability distributions has been found for the molecule numbers of the activated substrate, depending on the parameter settings. A particularly remarkable result for symmetric CMCs operated in the saturation regime was the emergence of an extremely broad PDF with a power-law tail. These fluctuations are driven by purely intrinsic noise, originating from the stochastic arrival times of the molecular reaction events.

We note that in biological systems, there are additional extrinsic sources of noise as well. For example, we have considered the total number of enzyme molecules, a_t and d_t, as being strictly constant in this paper. In biological systems, the enzymes are themselves subject to production and consumption processes and will therefore undergo concentration fluctuations. When these enzymes serve a CMC in the saturation regime, the steady-state activation level x/x_0 of the substrate will depend hypersensitively on the momentary ratio of enzyme concentrations a_t/d_t. Small (and sufficiently slow) fluctuations of the enzyme concentrations will therefore be amplified, leading to an additional extrinsic broadening of the PDF of $x(t)$.

At first glance, it seems that such extreme concentration fluctuations would compromise the function of biochemical networks [18,19]. However, recent reports have suggested that large biochemical fluctuations can also be beneficial for organisms, ranging from bacteria to humans [20–23]. In a recent review article [24], Losick *et al.* summarized a number of studies showing that certain cells choose one or an-
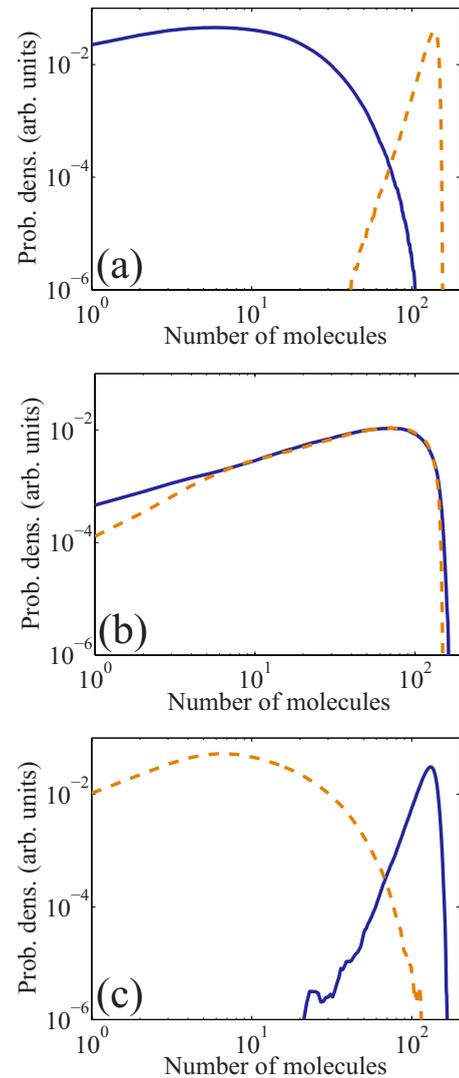
FIG. 8. (Color online) Tuning the CMC through the critical point by changing the enzyme concentration a_t. The simulated CMC has asymmetric rate constants, but becomes symmetric with respect to the effective coarse-grained parameters k_m and v_m for $a_t = 50$. Parts (a)–(c) correspond to $a_t = 45$, 50, and 55. Shown are double-logarithmic distributions of the activated (solid lines) and deactivated (dashed lines) substrates. For parameters, see text.

other pathway of differentiation stochastically without regard to environment or history.

Another example of stochastic signal processing is provided by the well-understood bacterial chemotaxis network. The flagellar motor of the bacterium is normally rotating in the counterclockwise (CCW) direction, but shows stochastic intervals of clockwise (CW) rotation. This gives rise to distinct phases of straight swimming motion of the bacterium, separated by random tumbling phases. Cell-membrane receptors detect the concentration of attractant molecules in the surrounding medium of the bacterium. Over several intermediate steps, the activation level of the receptors affects the distribution of CCW interval length and, thereby, the run length distribution of the bacterium's random walk in the medium. A statistical analysis of the CCW intervals revealed a power-law distribution [25], which has been related to mo-

lecular noise in the reaction network [26]. Interestingly, such random walks with power-law-distributed run lengths (Levy flights) are known to generate trajectories which are the optimum strategies to search efficiently for randomly located objects [27]. This example shows how the shaping of molecular noise and the modulation of the noise parameters in response to environmental stimuli can be used by cells for complex tasks such as foraging behavior.

We note that similar ideas of stochastic signal processing have recently emerged in the field of neuroscience [28]. In the new concept of "reservoir computing," a network of (randomly) connected neurons generates a so-called transient state dynamics, where the trajectory of the system state is temporally fluctuating between various unstable attractors. This autonomously active "reservoir" network is only weakly coupled to the "input" and "output" units. As simulations have shown, the mapping of low-dimensional input signals onto the high-dimensional state space of the reservoir network can be advantageous for the signal processing.

Finally, in this report, we have discussed the stationary behavior of a single CMC in which the total number of molecules is fixed. In living cells, however, multiple CMCs are connected in linear and branched signaling networks. Moreover, the total number of molecules fluctuates as new proteins are expressed or old proteins are recycled. If already a single CMC under stationary conditions gives rise to such highly complex, bizarre, and nondeterministic behavior as described in this paper, we argue that concentration fluctuations in living cells are even less predictable by classical mass-action theory.

APPENDIX

1. Average production rates

We consider an enzymatic conversion reaction of the general form

$$X + E \underset{u}{\overset{b}{\rightleftharpoons}} Y \overset{c}{\rightarrow} Z + E. \quad (A1)$$
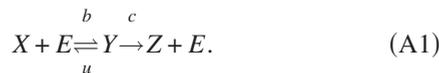
Using mass-action rate theory, we obtain for the temporal change of the concentration $y(t)$ of the enzyme-substrate complex

$$\dot{y} = bxe - uy - cy. \quad (A2)$$

We make the simplifying approximations that $x(t)$ is held constant. After a certain relaxation time, the system will reach a steady state, in which also $y(t) = $ const. The condition $\dot{y} = 0$ then leads to

$$y = xe\frac{b}{u + c}. \quad (A3)$$

The expression $\frac{b}{u+c} = \frac{1}{k_m}$ is defined as the inverse Michaelis constant, so that $y = \frac{xe}{k_m}$. Since the enzyme can either be free or bound in the complex, $e_t = e + y$, one obtains $y = \frac{x(e_t - y)}{k_m}$. Solving for y yields

$$y = e_t \frac{x}{x + k_m}. \quad (A4)$$

For the quantity of interest, the steady-state generation rate $\dot{z} = cy$ of the product, we finally obtain

$$\dot{z} = (ce_t)\frac{x}{x + k_m} = v_m \frac{x}{x + k_m}. \quad (A5)$$

2. Enzymatic conversion as an effective Poisson process

In general, an individual A-enzyme molecule can undergo a series of binding/unbinding events with the (nonexhaustible) substrate X_0 before the substrate is finally converted into a new X molecule. Therefore, even though each elementary reaction step, i.e., binding, unbinding, and conversion, is a Poisson process, the same is not true for the multistep production process [30].

However, many individual A-enzyme molecules, dispersed throughout the volume of the container, are simultaneously active, with independent temporal statistics. Our numerical simulations show that the superposition of many independent non-Poisson processes can resemble an effective Poisson process very closely. As expected, the characteristic time constant of this effective Poisson process is given by the inverse of the average total production rate $\bar{R}_{act}(x_0)$.

In our CMC system, the substrate molecule number x_0 and therefore $\bar{R}_{act}(x_0)$ are not constant. The resulting Poisson process is therefore not stationary but has a time-varying rate. We conclude that in systems with many independent enzyme molecules, the overall conversion process can be approximated by an inhomogeneous Poisson process.

3. Poisson process as white Gaussian noise

Assume now a Poisson process with constant average event rate $\bar{k} = \bar{R}_{act}(x_0)$. We express the temporal change of the number $x(t)$ of product molecules in the form

$$\dot{x} = \bar{k} + \Delta k(t). \quad (A6)$$

For later convenience, we want to approximate the fluctuation term by Gaussian white noise

$$\langle \Delta k(t) \Delta k(t') \rangle = \Gamma \delta(t - t'). \quad (A7)$$

What is the proper choice for the prefactor Γ so that the major statistical properties of a Poisson process are consistently reproduced?

To answer this question, we consider the number $n(T)$ of X molecules which are produced during an interval of length T,

$$n(T) = \int_0^T \dot{x}(t)dt = \bar{k}T + \int_0^T \Delta k(t)dt = \bar{n} + \Delta n. \quad (A8)$$

In the ensemble average, a Poisson process must fulfill

$$\langle (\Delta n)^2 \rangle = \bar{n} \qquad (A9)$$

or

$$\left\langle \left(\int_0^T \Delta k(t) dt \right)^2 \right\rangle = \bar{k}T. \qquad (A10)$$

The left side of the above equation can be reduced to ΓT. Using Eq. (A7), we therefore obtain $\Gamma = \bar{k}$ and therefore

$$\langle \Delta k(t) \Delta k(t') \rangle = \bar{k} \delta(t - t'). \qquad (A11)$$

Dividing this equation by $\bar{k}$ leads to

$$\left\langle \frac{\Delta k(t)}{\sqrt{\bar{k}}} \frac{\Delta k(t')}{\sqrt{\bar{k}}} \right\rangle = \delta(t - t'). \qquad (A12)$$

We now define a new stochastic process by

$$\zeta(t) = \frac{\Delta k(t)}{\sqrt{\bar{k}}}. \qquad (A13)$$

It is also normally distributed, but shows the desired property of δ autocorrelation with unit strength

$$\langle \zeta(t) \zeta(t') \rangle = \delta(t - t'). \qquad (A14)$$

We conclude that a proper description of a Poisson process by a stochastic differential equation should have the form

$$\dot{x} = \bar{k} + \sqrt{\bar{k}} \zeta(t). \qquad (A15)$$

---

[1] C. Furusawa *et al.*, Biophysics (Engl. Transl.) **1**, 25 (2005).

[2] H. Kitano, Science **295**, 1662 (2002).

[3] B. N. Kholodenko, Nat. Rev. Mol. Cell Biol. **7**, 165 (2006).

[4] L. H. Hartwell *et al.*, Nature (London) **402**, C47 (1999).

[5] D. E. Koshland, Jr., Science **280**, 852 (1998).

[6] E. Shacter *et al.*, J. Biol. Chem. **259**, 12252 (1984).

[7] E. G. Krebs, Curr. Top. Cell. Regul. **18**, 401 (1981).

[8] E. R. Stadtman *et al.*, Proc. Natl. Acad. Sci. U.S.A. **74**, 2761 (1977).

[9] A. Goldbeter *et al.*, Proc. Natl. Acad. Sci. U.S.A. **78**, 6840 (1981).

[10] J. C. Nacher *et al.*, Phys. Lett. A **360**, 174 (2006).

[11] P. B. Warren *et al.*, J. Chem. Phys. **125**, 144904 (2006).

[12] H. Qian *et al.*, PNAS Early Edition **99**, 10376 (2002).

[13] J. Levine *et al.*, Biophys. J. **92**, 4473 (2007).

[14] D. T. Gillespie, J. Chem. Phys. **113**, 297 (2000).

[15] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

[16] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1984).

[17] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992).

[18] M. Thattai *et al.*, Proc. Natl. Acad. Sci. U.S.A. **98**, 8614 (2001).

[19] M. Thattai *et al.*, Biophys. J. **82**, 2943 (2002).

[20] D. W. Austin *et al.*, Nature (London) **439**, 608 (2006).

[21] C. V. Rao *et al.*, Nature (London) **420**, 231 (2002).

[22] J. Hasty *et al.*, Proc. Natl. Acad. Sci. U.S.A. **97**, 2075 (2000).

[23] J. Paulsson *et al.*, Proc. Natl. Acad. Sci. U.S.A. **97**, 7148 (2000).

[24] R. Losick *et al.*, Science **320**, 65 (2008).

[25] E. Korobkova *et al.*, Nature (London) **428**, 574 (2004).

[26] Y. Tu and G. Grinstein, Phys. Rev. Lett. **94**, 208101 (2005).

[27] G. M. Viswanathan, Nature (London) **401**, 911 (1999).

[28] C. Gros, Cogn. Comput. **1**, 77 (2009).

[29] Note that in Sec. II E, we have neglected the amount of substrate which is bound in complexes. In order to refine the theory, let us define a new dynamic variable $\alpha = x + dx$ (the complementary variable $\beta = x_0 + ax_0$ is unnecessary since $\beta = x_t - \alpha$). This variable α defines the macrostate of the system in our coarse-grained view. It is changed only by activation or deactivation processes. On the other hand, binding and unbinding processes only affect the microstate of the system. The latter is defined by the numbers dx and ax_0, each of which can vary between 0 and the respective number of enzyme molecules. Thus, each macro state α can be subdivided into several microstates (dx, ax_0). The fluctuations of our variable of interest, $x(t)$, are determined by changes of the macro- and of the microstate. In the pre-equilibrium regime, for each momentary macrostate α, we expect that equilibrium distributions $P_{eq}(dx|\alpha)$ [and $P_{eq}(ax_0|\beta)$] of microstates are building up. The probability of having x activated substrate molecules is under such conditions given by $P(x) = \Sigma_{\alpha \geq x} P(\alpha) P_{eq}(dx = \alpha - x|\alpha)$.

[30] For a simple example, consider a sequence of one binding and one conversion step. The PDF of each elementary Poisson step is exponential. The PDF of the sequence is a convolution of two exponential functions, i.e., a gamma distribution with shape parameter $k = 2$.