

A stochastic switch with different phases

Cite as: Chaos **29**, 083107 (2019); <https://doi.org/10.1063/1.5096778>

Submitted: 19 March 2019 . Accepted: 16 July 2019 . Published Online: 07 August 2019

Ovidiu Lipan, and Emily Wu



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Sampling rare events across dynamical phase transitions](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **29**, 083106 (2019); <https://doi.org/10.1063/1.5091669>

[The reliability of recurrence network analysis is influenced by the observability properties of the recorded time series](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **29**, 083101 (2019); <https://doi.org/10.1063/1.5093197>

[Effects of stochastic parametrization on extreme value statistics](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **29**, 083102 (2019); <https://doi.org/10.1063/1.5095756>

AIP Author Services
English Language Editing



A stochastic switch with different phases

Cite as: Chaos 29, 083107 (2019); doi: 10.1063/1.5096778

Submitted: 19 March 2019 · Accepted: 16 July 2019 ·

Published Online: 7 August 2019



View Online



Export Citation



CrossMark

Ovidiu Lipan^{a)} and Emily Wu^{b)}

AFFILIATIONS

Department of Physics, University of Richmond, Richmond, Virginia 23173, USA

^{a)}Electronic mail: olipan@richmond.edu

^{b)}Electronic mail: emilywu@richmond.edu

ABSTRACT

We describe an analog stochastic switch that exhibits three distinct phases as its parameters change. The phases are classified by the mean and variance of the switch's output. A phase change appears if the mean or the variance tends to a finite value or to infinity. The switch can be embedded in a large gene regulatory network for which the moment equations naturally close at the second order. This switch was used to model the response of a heat-shock system.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5096778>

Living systems have to constantly adapt to changes in the external environment. For example, cells exposed to high temperatures use a heat-shock system to survive. To sense the external changes, cells engage gene regulatory networks. A theoretical understanding of gene regulation must account for the fluctuations caused by stochastic interactions of molecules, especially if present in small numbers. The use of the stochastic master equation for reaction kinetics has a long history that starts with Kramers. In the last 60 years, the master equation was successfully solved for a series of basic systems via special functions. However, these special methods are not scalable to large networks, which are at the core of systems biology. Our study focuses on a scalable method that we use here to analytically solve a small but nontrivial stochastic biocircuit. This switch can be used as a building block for a shock-response gene regulatory network.

I. INTRODUCTION

Studying interactions and connectivity among different components of molecular networks is a major challenge in systems biology. Identifying the structure of a network is facilitated at present by technological advances in high-throughput molecular biology.¹⁻³ One aspect of the network's structure is represented by the topology of the network. Different node connectivity leads to different network topologies, which then explain different phenotypical behaviors at the organismal level.⁴⁻⁶ Another aspect is related to the dynamical evolution of the network. Specifically, the time evolution we focus on is described by the master equation introduced by Pauli in

Ref. 7, which was used and described subsequently in Refs. 8–10 and elegantly reviewed by Chandrasekhar.¹¹

The master equation is built upon the intrinsic stochasticity of interactions that take place between the network's nodes. The stochastic processes we point out are discrete and involve only positive integer numbers since they describe changes in the number of molecules. Due to the advent of systems biology, the challenge at present is mainly connected with the use of the master equation for large networks. Large networks are obviously challenging because of the large dimensionality of their master equation. However, the master equation is hard to solve for small networks too, because most often, it leads to an infinite number of ordinary differential equations. Namely, the equations for the variance-covariance matrix depend on the third order moments, which in turn depend on the fourth order and so on.

A variety of procedures were proposed for cutting the system of equations up to the second order. A major problem with the majority of the trimming-down methods, called moment closure methods, is that the trimmed equations they deliver may not describe a true stochastic process. Namely, we start with a stochastic process described by a discrete probability distribution on positive integers, write down an infinite number of differential equations for the moments, and then reduce the system of equations to a finite one. How do we know that there exists a probability distribution, more so a discrete one and on positive integers, that will produce the same finite number of moment equations? A possible solution to this conundrum is to approximate the probability distribution itself so that it generates a finite system for the moment equations instead of cutting or approximating the original infinite system of equations. This way we know from the start that the finite

equations were generated by a discrete positive integer probability distribution.

The present paper is addressing these problems for a simple but nontrivial switch. The switch presented here illustrates that even two elementary components can generate interesting and nontrivial biocircuits. In addition, we choose to work on a switch because turning genes ON and OFF plays a crucial role in living systems.¹² One ubiquitous functionality of a biocircuit is based on transforming a small variation in an input signal into a large variation in the output response. Out of a variety of possible examples, we will explore here a specific switch that was used to describe experimental data on the heat-shock response.^{13,14}

To ensure that the system of moment equations is self-contained at the second order, we need to use a limited number of distinct elementary components for the building blocks of the network. They are listed in Ref. 15. Given that the list of elementary components is narrow, it is useful to add to this list a series of nonelementary but simple biocircuits that can be used to construct stochastic networks solvable at the second-order moment level. This stochastic switch is a biocircuit of this kind.

II. DESCRIPTION OF THE STOCHASTIC SWITCH

The stochastic switch is composed of two distinct molecule species, types 1 and 2, represented by the labels q_1 and q_2 in Fig. 1. Besides being a label in Fig. 1, q_1 and q_2 also represent the number of molecules of types 1 and 2, respectively. This biocircuit behaves like an analog switch because q_2 builds up toward levels higher than the level it starts at time $t = 0$. Analog processes are ubiquitous in living systems. Many processes in biological cells are carried out in an analog fashion rather than in a digital manner.¹⁶

There are two processes represented in Fig. 1. The first process is the autodegradation of the type 1 molecule $(q_1, q_2) \rightarrow (q_1 - 1, q_2)$. During this process, q_2 is not affected. This process is depicted by a box marked $\varepsilon_1 = (-1, 0)$. It may look that a simple process is represented by an unnecessary complex drawing rule. However, if the ε_1 box is eliminated from Fig. 1, the graphical representation is identical with the one that is in widespread use. The line that stops on the type 1 molecule ends in a bar that represents the degradation. The need for

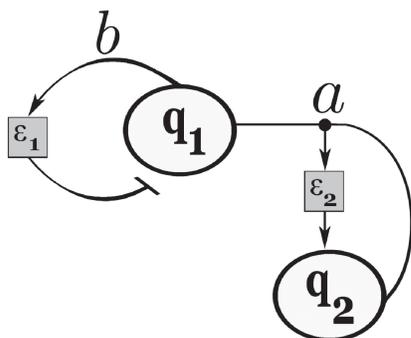


FIG. 1. As molecule q_1 degrades, it creates more q_2 through a complex-formation product rule.

the $\varepsilon_1 = (-1, 0)$ box, usually absent, is to precisely show the change of the state from (q_1, q_2) to $(q_1 - 1, q_2)$ during the process. This way there is no ambiguity between the diagrammatic representation and the mathematical equations.

In addition to ε_1 , we need to specify the transition probability per unit time¹⁰ for this process. This is $T_{\varepsilon_1} = bq_1$, where b is a parameter that multiplies the number of molecules q_1 . The arrow that starts on the type 1 molecule and ends on the ε_1 box is labeled by the parameter b . To avoid any confusion, we mention that from the transition probability per unit time, we obtain the transition probability $T_{\varepsilon_1} dt = bq_1 dt$ for the ε_1 process to take place in a small time interval $(t, t + dt)$. Because $dt \rightarrow 0$, the transition probability is less than 1 for any choice of the parameter b and any value of q_1 . The transition probabilities for events on a finite time interval $(t, t + \tau)$ can be computed from the transition probabilities per unit time.¹⁷

The second process is represented by $\varepsilon_2 = (0, 1)$ and $T_{\varepsilon_2} = aq_1q_2$. In this case, the molecule number q_2 grows by 1, whereas q_1 stays constant. The transition probability depends on the product of the molecule numbers q_1q_2 . This fact has a strong consequence on the production of the type 2 molecule. If $q_1 = 0$ after a sequence of degradations, the molecule 2 stops being produced and stays at a certain level. However, and this is the subtlety of this biocircuit, the complete degradation of q_1 to $q_1 = 0$ may become very slow in comparison with the fast rate at which q_2 increases. In this situation, the transition probability $T_{\varepsilon_2} = aq_1q_2$ becomes much larger than the degradation transition probability $T_{\varepsilon_1} = bq_1$, leaving a very small chance for the degradation process to be completed. The molecule q_2 will continue to be produced without a limit. A switch that permits a limitless production of q_2 is quite different from a switch for which q_2 stops at a certain finite level. We say that these switches are in different phases. The product aq_1q_2 , which is responsible for the emergence of different phases, is depicted as a node labeled a in Fig. 1. There are two lines that meet at this node, one that comes from q_1 and another one that comes from q_2 . The line that emerges from an ε_2 -box reaches only the type 2 molecule because during this second process, only q_2 is affected. The arrow at the end of this line shows a production process. As for the degradation case, this convention is widely used. To conclude, the diagrammatic representation is based in the most general case on two kinds of lines. One, called the control line, represent the transition probability. The control lines start from the molecules that are part of the mathematical formula for the transition probability. The other kind, called the action line, start from an ε box and ends on the molecules that change under the process.

In principle, we may easily avoid drawing a diagram for a simple biocircuit, like the one in Fig. 1. However, for large stochastic networks, a diagram is essential to visualize the connectivity between different parts. A diagram also helps to reduce a large network to a series of connected subnetworks. For this paper, the main reason for constructing the switch's diagram is to visualize a mathematical procedure, described in Sec. III, which helps revealing three distinct phases of this stochastic switch.

In a previous work,¹³ we used the switch in the regime of small coupling parameter $a \simeq 0.1$. For this regime, the approximations that we used were sufficient to explain the experimental data. However, this switch model may potentially be used in regimes that require higher values for the coupling parameter a , so a complete study is

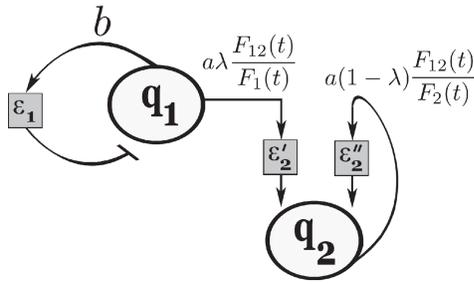


FIG. 2. As molecule q_1 degrades, it creates more q_2 through a complex-formation product rule.

warranted. A way to proceed is through computer simulations. Simulating Monte Carlo paths for a large coupling parameter a was hard to achieve because q_2 quickly accumulates and reaches very high values obstructing the emergence of a clear picture of the switch behavior. The progress came once we realized that the method presented in Ref. 15 leads, for this switch, to a closed-form mathematical solution.

In what follows, we aim (1) to explain the loop-closing (LC) method of splitting the nonlinear node¹⁵ as applied to the stochastic switch from Fig. 1, (2) to show that the LC-model of the switch, Fig. 2, is completely solvable, (3) to study the phase changes of the LC-model of the switch, (4) to use Monte Carlo simulations to support the LC-model predictions, and, finally, (5) to show, for special initial conditions for q_1 and q_2 , that the solution to the original switch (Fig. 1) leads to the same results obtained from the LC-model of the switch (Fig. 2).

III. THE LC-MODEL FOR THE SWITCH

The master equation for the time-dependent probability distribution $P(q_1, q_2, t)$ for the original switch in Fig. 1 is constructed out of two transition probabilities per unit time,

$$T_{\varepsilon_1} = bq_1, \tag{1}$$

$$T_{\varepsilon_2} = aq_1q_2, \tag{2}$$

and reads

$$\begin{aligned} \frac{\partial P(q_1, q_2, t)}{\partial t} &= T_{\varepsilon_1}(q_1 + 1, q_2, t)P(q_1 + 1, q_2, t) \\ &+ T_{\varepsilon_2}(q_1, q_2 - 1, t)P(q_1, q_2 - 1, t) \\ &- (T_{\varepsilon_1}(q_1, q_2, t) + T_{\varepsilon_2}(q_1, q_2, t))P(q_1, q_2, t). \end{aligned} \tag{3}$$

To generate the system of ordinary differential equations for the moments, the master equation is transformed and written for the moment generating function,

$$H(z_1, z_2, t) = \sum_{q_1=0, q_2=0}^{\infty} z_1^{q_1} z_2^{q_2} P(q_1, q_2, t), \tag{4}$$

$$\begin{aligned} s \frac{\partial H(z_1, z_2, t)}{\partial t} &= b(z_1^{-1} - 1)z_1 \frac{\partial H(z_1, z_2, t)}{\partial z_1} \\ &+ a(z_2 - 1)z_1z_2 \frac{\partial^2 H(z_1, z_2, t)}{\partial z_1 \partial z_2}. \end{aligned} \tag{5}$$

The moments are generated from partial derivatives. For example,

$$H_2(t) = \left. \frac{\partial H(z_1, z_2, t)}{\partial z_2} \right|_{z_1, z_2 \rightarrow 1} \tag{6}$$

and

$$H_{122}(t) = \left. \frac{\partial^3 H(z_1, z_2, t)}{\partial z_1 \partial z_2 \partial z_2} \right|_{z_1, z_2 \rightarrow 1}. \tag{7}$$

The equations for the mean values (first order moments) are

$$\frac{dH_1}{dt} = -bH_1, \tag{8}$$

$$\frac{dH_2}{dt} = aH_{12}, \tag{9}$$

whereas for the second order moments, they are

$$\frac{dH_{11}}{dt} = -2bH_{11}, \tag{10}$$

$$\frac{dH_{12}}{dt} = aH_{12} + aH_{112} - bH_{12}, \tag{11}$$

$$\frac{dH_{22}}{dt} = 2aH_{12} + 2aH_{122}. \tag{12}$$

Molecule q_1 obeys trivial solutions,

$$H_1(t) = H_1(0)e^{-bt}, \tag{13}$$

$$H_{11}(t) = H_{11}(0)e^{-2bt}. \tag{14}$$

However, for molecule q_2 , the system is incomplete requiring equations for the third order moments H_{112}, H_{122} , which will require fourth order moments and so on.

A way to break this infinite chain of equations is to split the product node that represents the aq_1q_2 transition probability. The details of the procedure were reported in Ref. 15. Here, we present its implementation for the switch. The first step is to transform the topology of the network from the original version in Fig. 1 into a new version in Fig. 2. Notice that the original topology is almost the same with the exception of the product node, which is split. The splitting creates two ε -boxes out of the original one. The second step is to associate new transition probabilities to each of the new ε -boxes. These new transition probabilities depend on the generating function for the split-node version from Fig. 2. To avoid confusion with $H(z_1, z_2, t)$, we denote by $F(z_1, z_2, t)$ the generating function of the split-node biocircuit from Fig. 2. The results from Ref. 15 require that the coupling constant a from Fig. 1 splits into two time-dependent couplings $\lambda a \frac{F_{12}(t)}{F_1(t)}$ and $(1 - \lambda)a \frac{F_{12}(t)}{F_2(t)}$. These time-dependent couplings are not known in advance since they depend on the unknown moments $F_1(t), F_2(t)$, and $F_{12}(t)$. However, they are found from solving the system of equations for the moments generated from (15). In addition, a new parameter $0 \leq \lambda \leq 1$ is needed. The parameter λ behaves like a weight parameter for the two new control lines. The weight parameter for the control line that starts from q_1 is λ , whereas for the control line from q_2 is $1 - \lambda$.

The third and the last step of the procedure is to solve the time evolution of the moments of the master equation for the split-node

version. This is possible either by analytical or numerical methods because the LC-procedure leads to a closed system of equations at the second order moments. It is worth mentioning that the splitting node procedure is not restricted to a two-molecule system. The procedure generates equations that are closed at second order moments for genetic networks that contain any number of molecules. This property allows any two split-node networks to be interconnected and maintain the solvability at the second order moment and also higher orders if those moments are needed.

The LC-master equation for the split-node switch is

$$\begin{aligned} \frac{\partial F(z_1, z_2, t)}{\partial t} &= b(z_1^{-1} - 1)z_1 \frac{\partial F(z_1, z_2, t)}{\partial z_1} \\ &+ a\lambda(z_2 - 1) \frac{F_{12}(t)}{F_1(t)} z_1 \frac{\partial F(z_1, z_2, t)}{\partial z_1} \\ &+ a(1 - \lambda)(z_2 - 1) \frac{F_{12}(t)}{F_2(t)} z_2 \frac{\partial F(z_1, z_2, t)}{\partial z_2}. \end{aligned} \quad (15)$$

To continue the analysis of Fig. 2, notice that the action ε_2 in Fig. 1 is controlled by both q_1 and q_2 , which is visible in the term $z_1 z_2 \partial^2 / \partial z_1 \partial z_2$ from (5). After splitting, Fig. 2, the action ε'_2 is controlled by q_1 , which is represented by $z_1 \partial / \partial z_1$ in (15). The other action, ε''_2 , is represented by $z_2 \partial / \partial z_2$.

As stated above, a feature of the splitting-node approach is the appearance of a new parameter denoted here by λ , which is absent in Fig. 1. The value of the λ -parameter depends on the global behavior of the system. Namely, for this switch, $\lambda = 1$ places the entire action under the control of q_1 , whereas $\lambda = 0$ highlights q_2 as the driver of the switch. In general, the driver is identified by the difference in the number of molecules.¹⁵ If $q_1 < q_2$, then q_1 is the driver and conversely for $q_2 < q_1$. For this switch, $\lambda = 1$, which signifies that the driver of the switch is q_1 because q_1 only decays in time. For some networks, a bistable biocircuit, for example, no molecule can be designated as a driver¹⁵ since the stochastic paths meander from one attraction point to another. For such situations, the λ -parameters, which can be many since each product node creates one, acquire values that are not the extreme values of 0 or 1.

For a general biocircuit, the split-node procedure renders numerically solvable systems of equations but not analytically solvable ones as we may expect, especially for large networks. However, for this switch, the system of equations can be reduced to a manageable analytical form. The value of these analytical expressions does not lie solely in their precise time-dependency. Anyway, the split switch is meant to be embedded into a large network for which an analytical solution is most likely unachievable. The formulas that will be discussed below are valuable for their asymptotic behavior. They show that the split switch has three phases. For weak coupling constant a , both the mean and the variance of q_2 grow toward a finite steady state equilibrium. This is phase 1. Increasing the coupling constant, the switch enters phase 2, where the mean of q_2 still reaches a finite steady state, but the variance grows in time to infinity. Once the coupling becomes large, in phase 3, both the mean and the variance continue to grow in time without reaching a steady state. In view of this stochastic switch being initially employed for a heat shock model,^{13,14} we see that the transition between different phases may

be used as a response to heat shocks of different degrees of intensities. For weak temperature shocks, in the vicinity of the physiological temperatures, the system will produce a relatively small amount of heat-shock proteins that will level out at a stationary state. If the shock is above a first threshold, the heat-shock system may enter phase 2. Here, since the variance is growing without an upper bound, in some cells, the stochastic paths reach much larger levels than the average mean of the population of cells. In phase 3, for strong shocks above the second threshold, the heat-shock system responds fast producing large amounts of proteins needed to protect the cells. The way the cell is turning off the heat-shock system is not part of the present switch. Other feedback mechanisms are in place to detect the ever growing levels of q_2 in phase 3 and turn it off.

We turn to analyze the analytical solutions of the split switch and its three-phase behavior. The equations for q_2 are

$$\frac{dF_2}{dt} = aF_{12}, \quad (16)$$

$$\frac{dF_{12}}{dt} = -bF_{12} + a\lambda F_{12} + a\lambda \frac{F_{11}}{F_1} F_{12} + a(1 - \lambda) \frac{F_{12}^2}{F_2}, \quad (17)$$

$$\frac{dF_{22}}{dt} = 2a(1 - \lambda)F_{12} + 2a\lambda \frac{F_{12}^2}{F_1} + 2a(1 - \lambda) \frac{F_{12}}{F_2} F_{22}. \quad (18)$$

The solution for the mean value $\langle q_2 \rangle = F_2$ is written in terms of the incomplete Gamma function,

$$\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt, \quad (19)$$

which is as follows:

$$\begin{aligned} F_2(t) &= F_2(0) \left(\frac{F_2(0)}{aF_{12}(0)} \right)^{-\frac{1}{\lambda}} \\ &\times \left(\frac{F_2(0)}{aF_{12}(0)} + \frac{e^\eta \eta^{-\zeta}}{b} (\Gamma(\zeta, \eta e^{-bt}) - \Gamma(\zeta, \eta)) \right)^{\frac{1}{\lambda}}, \end{aligned} \quad (20)$$

where

$$\zeta = \frac{b - a\lambda}{b}, \quad (21)$$

$$\eta = \lambda \frac{a F_{11}(0)}{b F_1(0)}. \quad (22)$$

The correlation F_{12} can again be solved in terms of the incomplete Gamma function,

$$\begin{aligned} F_{12}(t) &= F_{12}(0) e^{-bt + \eta(1 - e^{-bt}) + a\lambda t} \left(\frac{F_2(0)}{aF_{12}(0)} \right)^{1 - \frac{1}{\lambda}} \\ &\times \left(\frac{F_2(0)}{aF_{12}(0)} + \frac{e^\eta \eta^{-\zeta}}{b} (\Gamma(\zeta, \eta e^{-bt}) - \Gamma(\zeta, \eta)) \right)^{-1 + \frac{1}{\lambda}}, \end{aligned} \quad (23)$$

whereas the second order moment for the molecule q_2 is reduced to an integral form,

$$F_{22}(t) = F_2(t)^{2-2\lambda} \left(F_{22}(0)F_2(0)^{2\lambda-2} + \frac{2(1-\lambda)}{2\lambda-1} (F_2(t)^{2\lambda-1} - F_2(0)^{2\lambda-1}) + \frac{2\lambda}{aF_1(0)} \int_0^t e^{b\tau} F_2(\tau)^{2\lambda-2} F_2'(\tau)^2 d\tau \right). \quad (24)$$

All these three solutions allow the study of the asymptotic behavior as $t \rightarrow \infty$. The behavior of the mean value $F_2(t)$ is revealed by the asymptotic of $\Gamma(\zeta, \eta\varepsilon)$ for $\varepsilon = e^{-bt} \rightarrow 0$,

$$\Gamma(\zeta, \eta\varepsilon) \sim \Gamma(\zeta) + \varepsilon^\zeta \left(\frac{\varepsilon\eta^{\zeta+1}}{\zeta+1} - \frac{\eta^\zeta}{\zeta} \right). \quad (25)$$

The term that decides the asymptotic behavior of $F_2(t)$ in (20) is thus

$$\varepsilon^{\frac{\zeta}{\lambda}} = e^{-t\frac{b-a\lambda}{\lambda}}. \quad (26)$$

If $b - a\lambda > 0$, the mean value for the switch attains a steady state, whereas if $b - a\lambda < 0$, the mean continues to raise. For $b - a\lambda = 0$, the asymptotic expansion contains $-\ln(\varepsilon)$, and so the mean value tends to infinity as $t \rightarrow \infty$.

The asymptotic analysis for $F_{22}(t)$ goes as follows. If $F_2(t)$ grows to infinity, it is easy to see that $F_{22}(t)$ grows to infinity too. However, if $F_2(t)$ tends to a constant, then $F_{22}(t)$ may tend to a constant too or to infinity. Indeed, from the asymptotic behavior of $F_2(t)$, we obtain the asymptotic behavior of the integral term to be $e^{b(1-2\zeta)t}$. If

$$1 - 2\zeta = -\frac{b - 2a\lambda}{b} < 0, \quad \text{then } F_{22} \rightarrow \text{const}, \quad (27)$$

$$1 - 2\zeta = -\frac{b - 2a\lambda}{b} > 0, \quad \text{then } F_{22} \rightarrow \infty. \quad (28)$$

For $\zeta = 1/2$, the integral grows linearly with t and F_{22} diverges.

To summarize, we find that the split-node switch can be placed in three phases:

phase 1 : $\lambda \frac{a}{b} < \frac{1}{2}$,
 both q_2 mean and variance \rightarrow constants; (29)

phase 2 : $\frac{1}{2} \leq \lambda \frac{a}{b} < 1$,
 q_2 mean \rightarrow constant, q_2 variance $\rightarrow \infty$; (30)

phase 3 : $1 \leq \lambda \frac{a}{b}$,
 both q_2 mean and variance $\rightarrow \infty$. (31)

In principle, we could stop at this point because only the split-node switch is useful as a building block for a large network so that the equations close at the second order moment. However, it is useful to explore the behavior of the switch from Fig. 1 and show that it also swipes three phases as the coupling parameter a increases. This means that the node-splitting procedure preserves the phase properties.

IV. PHASE TRANSITION FROM MONTE CARLO SIMULATIONS

We ran a series of Monte Carlo simulations using the Gillespie algorithm.¹⁸ For each parameter setting, we ran 10^4 paths. The Monte Carlo simulations suggest that, indeed, the switch moves from one phase to another as the coupling a crosses the values of $1/2$ and then 1 (Figs. 3–5). The $1/2$ and 1 are the thresholds we obtain for $b = 1$ and $\lambda = 1$. The set value $b = 1$ defines the unit of time, whereas $\lambda = 1$ is set by the driver,¹⁵ which for this biocircuit is q_1 . This expresses the fact that at some point in time, q_1 decreases below q_2 and so q_1 becomes the sole driver of q_2 . In addition, we confirm the choice of $\lambda = 1$ in Sec. V, where we approach the original switch from a different perspective. It is worth mentioning that q_1 may not be the driver if the switch is embedded into a larger network. In this case, another molecule from the large network may influence the dynamics of the switch, which requires $\lambda \neq 1$. For this reason, the split-node switch is a biocircuit by its own and can be used for any λ . A variable λ offers an additional flexibility vs the original switch, which has fixed phase thresholds.

To obtain the phase transition thresholds from the Monte Carlo simulations for Fig. 1, we look into the tail of the distribution of q_2 after some time away from the start $t = 0$ (Fig. 6).

We compare the statistical behavior of the tail with the power-law distribution. The reason we chose a power-law distribution is because it stands out for its mathematical properties and for its surprising materialization in physics and other fields. The power-law exponent is usually connected with the dynamical processes that generate the power-law distribution like in the case of the phase transitions in thermodynamic systems. For the switch, the dynamics is represented by the transition probabilities. The power-law exponent depends on how fast the degradation of q_1 takes place and on how fast q_2 accumulates through the coupling parameter a .

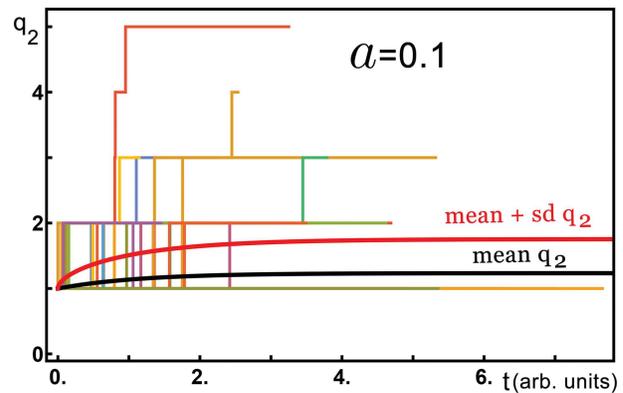


FIG. 3. Paths for $q_1 = 2$, $q_2 = 1$, $b = 1$. In phase 1, for which, $a < 1/2$, both the mean and the variance tend to a constant value. The stairlike curves illustrate stochastic paths of q_2 . The continuous black curve represents the mean value, whereas the red curve represents the mean plus one standard deviation for q_2 . The same meanings are associated with paths and the continuous curves for Figs. 4 and 5.

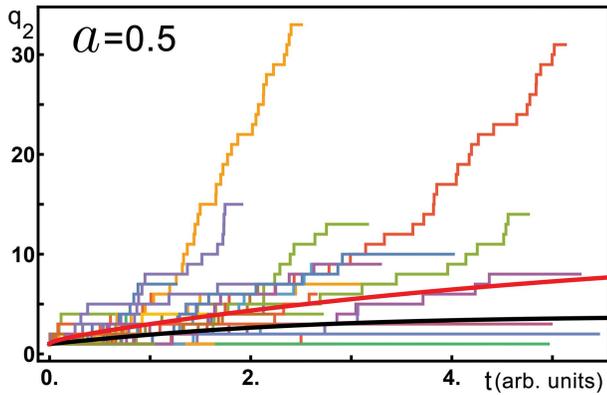


FIG. 4. Paths for $q_1 = 2, q_2 = 1, b = 1$. At the first threshold $a = 1/2$, the switch transits from phase 1 into phase 2. Here, the mean value remains finite, but the variance grows to infinity. For practical applications, for which time does not reach infinity, this means that at the end of the process, a good fraction of stochastic paths reaches high q_2 values. The arbitrary units for the time are set by $b = 1$.

Mathematically, a quantity x obeys a power law if it is drawn from a probability distribution $p(x) \propto x^{-\alpha}$, where α is the power-law exponent or scaling parameter. Few empirical phenomena obey power laws for all values of x . Usually, the power law applies only for values greater than some minimum x_{\min} . In such cases, only the tail of the distribution follows a power law,

$$p(x) = \frac{\alpha - 1}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\alpha}, \quad x \geq x_{\min}. \quad (32)$$

The moments of the power-law distribution,

$$\langle x^m \rangle = \int_{x_{\min}}^{\infty} x^m p(x) dx = \frac{\alpha - 1}{\alpha - 1 - m} x_{\min}^m, \quad (33)$$

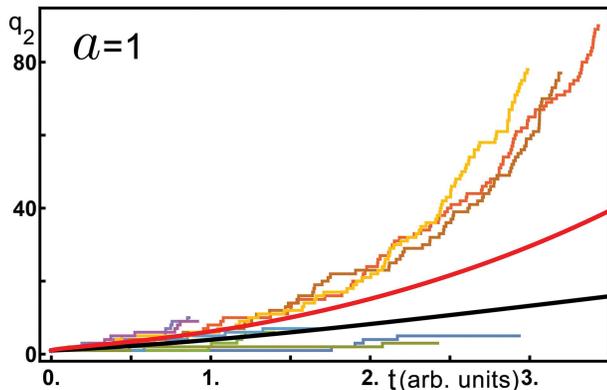


FIG. 5. Paths for $q_1 = 2, q_2 = 1, b = 1$. Once $a = 1$, the switch enters from phase 2 into phase 3 where both the mean value and the variance grow to infinity. In this case, after a finite amount of time, most of the stochastic paths reached high levels of q_2 .

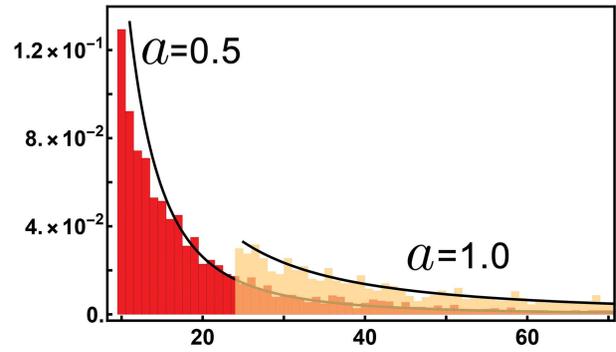


FIG. 6. Tails for $q_1 = 2, q_2 = 3$. The estimated power-law exponents are $\hat{\alpha} = 2.72$ for $a = 0.5$ and $\hat{\alpha} = 1.85$ for $a = 1.0$ instead of the theoretical $\alpha = 3$ and $\alpha = 2$, respectively. The data correspond to the time $t = 15$ in units set by $b = 1$. For $a = 0.5, x_{\min} = 10$ and for $a = 1, x_{\min} = 24$.

diverge if

$$m \geq \alpha - 1. \quad (34)$$

We estimated the power-law exponent, $\hat{\alpha}$, from the Monte Carlo simulations of the original switch (Fig. 1). The results, Fig. 7, show the desired relation: as the coupling constant increases, the power-law exponent decreases. The estimation of the power-law exponent was computed using the powerLaw package following the procedure from Ref. 19. To ensure that the switch is not in a transient regime, we estimated the power-law exponent for two times, $t = 10$ and $t = 15$, in time units specified by $b = 1$. For both times, the estimated power-law exponents came out the same.

Although Fig. 7 confirms that the split switch maintains the phase properties of the original switch from Fig. 1, the inherent

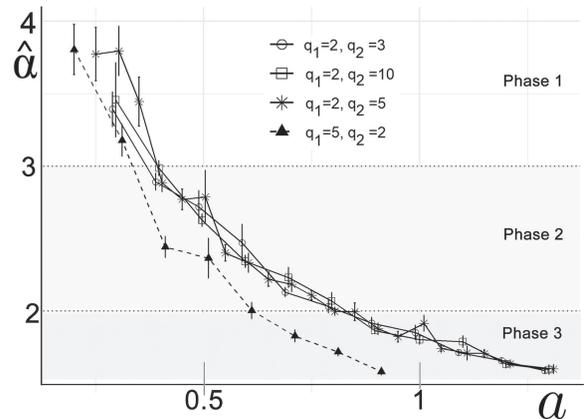


FIG. 7. The estimated power-law exponent shows that the switch from Fig. 1 traverses all three phases as the coupling parameter a increases. For all simulations, the time scale is set by $b = 1$, which places the first phase threshold at $a = 0.5$ and the second at $a = 1$. The data correspond to the time $t = 15$.

limitations associated with the statistical estimation and Monte Carlo simulations obscure the fact that the phase transition thresholds are $a = 1/2$ and $a = 1$. Section V will show that the thresholds are exactly those numbers, albeit the proof is only for special initial conditions.

V. THE PHASES OF THE ORIGINAL SWITCH FOR THE SPECIAL INITIAL CONDITION $q_1 = 2$ AND $q_2 = 1$

It would be useful to have an analytic solution for the original switch in the limit $t \rightarrow \infty$ for any initial condition. Such a solution can be used to judge the LC method and also the statistical methods used for the estimation of the power-law exponent, like the powerLaw package.¹⁹ Although we do not have a solution for any initial condition, we found one for the special case $q_1 = 2$ and $q_2 = 1$ at $t = 0$. This solution, presented below, shows that, for this special initial condition, the LC method is accurate, whereas the power-law estimation is not.

The goal is to obtain the level of q_2 after $q_1 = 0$. Once $q_1 = 0$, the switch stops and so the q_2 -attained level remains flat for the rest of the time as $t \rightarrow \infty$. This way we obtain the probability distribution of q_2 at $t = \infty$ from which we will obtain the phase thresholds. The accumulation of q_2 follows the Monte Carlo stochastic process which runs in two stages. Following the Gillespie algorithm,¹⁸ first, a uniform random number $0 \leq U_T \leq 1$ decides the length between two consecutive reactions,

$$\tau = \frac{1}{bq_1 + aq_1q_2} \ln\left(\frac{1}{1 - U_T}\right). \tag{35}$$

Second, after the time interval τ selected above ends, another uniform random number $0 \leq U_R \leq 1$ selects one reaction out of the only two reactions available for this switch. If

$$\frac{T_{\varepsilon_1}}{T_{\varepsilon_1} + T_{\varepsilon_2}} \leq U_R \leq 1, \tag{36}$$

then reaction $\varepsilon_2, q_2 \rightarrow q_2 + 1$ takes place. Otherwise, the other reaction $q_1 \rightarrow q_1 - 1$ is active. In what follows, we scale the time to get $b = 1$, similar to the procedure explained for the Monte Carlo simulations.

Each stochastic path can exhibit only two ε_1 reactions, given that $q_1 = 2$. Once $q_1 = 0$, the switch stops and the level of q_2 remains the same as $t \rightarrow \infty$. We will use two parameters to uniquely describe a path. One is the total number of reactions J that were active along the path, including the last one $q_1 = 1 \rightarrow 0$. A second parameter, J_1 , represents the position of the reaction $q_1 = 2 \rightarrow 1$ along the path. The final level attained by q_2 is $J - 1$. The time length of the path from $t = 0$ until the end of reaction J is a random variable built up as a sum of J random variables, (35), each one following a different exponential distribution. A time-dependent study of the stochastic nonlinear switch needs to take into consideration the stochasticity of time until $q_1 \neq 0$. However, for $t \rightarrow \infty$, this is not necessary since the relevant probability distribution is $P(J_1, J)$ for the path (J_1, J) to be generated.

The quantities of interest at $t \rightarrow \infty$ are

$$\langle q_2 \rangle = \sum_{J=2}^{\infty} \sum_{J_1=1}^{J-1} (J - 1) P(J_1, J), \tag{37}$$

$$\langle (q_2)^2 \rangle = \sum_{J=2}^{\infty} \sum_{J_1=1}^{J-1} (J - 1)^2 P(J_1, J), \tag{38}$$

where $J \geq 2$ means that a path is based on at least two reactions and $J_1 \leq J - 1$ is specifying that the q_1 -reaction ($q_1 = 1 \rightarrow q_1 - 1 = 0$) is the last, the J th, reaction. Given that U_R is uniformly sampled, we get that $P(J_1, J)$ is composed of four factors. The first factor contains the contribution of consecutive $J_1 - 1$ reactions of the type $q_2 \rightarrow q_2 + 1$,

$$\prod_{j=1}^{J_1-1} \left(1 - \frac{1}{1 + aj}\right). \tag{39}$$

Note that the probability for the reaction number $J_1 - 1$ is set up by the level of $q_2 = J_1 - 1$ reached by q_2 before reaction number $J_1 - 1$.

After reaction $J_1 - 1$, the level of q_2 is J_1 and so the second factor that contributes to $P(J_1, J)$ is

$$\frac{1}{1 + aJ_1}, \tag{40}$$

which is the probability for the first reaction ($q_1 = 2 \rightarrow q_1 - 1 = 1$). Then comes again a series of consecutive $q_2 \rightarrow q_2 + 1$ reactions, which gives the third part of $P(J_1, J)$,

$$\prod_{j=J_1}^{J-2} \left(1 - \frac{1}{1 + aj}\right). \tag{41}$$

Finally, the probability for the final reaction ($q_1 = 1 \rightarrow q_1 - 1 = 0$) is

$$\frac{1}{1 + a(J - 1)}, \tag{42}$$

which does not change the level $q_2 = J - 1$. Altogether, these factors give

$$P(J_1, J) = \prod_{j=1}^{J_1-1} \left(\frac{aj}{1 + aj}\right) \frac{1}{1 + aJ_1} \frac{1}{1 + a(J - 1)}. \tag{43}$$

The sum over J_1 in (37) is

$$\sum_{J_1=1}^{J-1} \frac{1}{1 + aJ_1} = a^{-1}(\Psi(a^{-1} + J) - \Psi(a^{-1} + 1)), \tag{44}$$

where $\Psi(z)$ is the digamma function. Next, using the product

$$\prod_{j=1}^{J-1} \frac{aj}{1 + aj} = (J - 1)! \frac{\Gamma[a^{-1} + 1]}{\Gamma[a^{-1} + J]}, \tag{45}$$

we arrive at

$$\langle q_2 \rangle = \frac{1}{a^2} \sum_{J=2}^{\infty} (\Psi(a^{-1} + J) - \Psi(a^{-1} + 1))(J - 1)! \frac{\Gamma[a^{-1} + 1]}{\Gamma[a^{-1} + J]} \tag{46}$$

TABLE I. Comparison between the LC method and the solution to the nonlinear switch for $q_1 = 2, q_2 = 1, b = 1, t \rightarrow \infty$. For the LC method, $\lambda = 1$.

a	$\langle q_2 \rangle$	$\langle q_2 \rangle$ LC	$\langle (q_2)^2 \rangle$	$\langle (q_2)^2 \rangle$ LC
1/8	64/49	1.31	2.25	2.08
1/4	16/9	1.77	56/9	4.59
1/3	9/4	2.23	63/4	8.93
5/12	2.98	2.87	69.06	23.65
7/16	3.16	3.08	124.84	34.00
31/64	3.76	3.61	2044.24	163.13
511/1024	3.98	3.81	524 284.01	2123.39
1/2	4	3.82	∞	∞
5/6	36	21.78	∞	∞
9/10	100	42.36	∞	∞
99/100	1000	529.7	∞	∞
999/1000	10^6	5422.38	∞	∞
1	∞	∞	∞	∞

TABLE II. The power-law exponent $\hat{\alpha}$ and x_{\min} for $q_1 = 2, q_2 = 1$. The data correspond to the time $t = 15$ in units set by $b = 1$.

a	$\hat{\alpha}$	x_{\min}
0.1	5.15	3
0.2	4.28	5
0.3	3.62	8
0.4	3.03	13
0.5	2.69	11
0.6	2.30	9
0.7	2.09	11
0.8	1.93	15
0.9	1.94	44
1.0	1.76	24
1.1	1.67	11
1.2	1.69	94
1.3	1.60	24

and

$$\langle (q_2)^2 \rangle = \frac{1}{a^2} \sum_{j=2}^{\infty} (j-1) \left(\Psi \left(j + \frac{1}{a} \right) - \Psi \left(1 + \frac{1}{a} \right) \right) \times (j-1)! \frac{\Gamma[1 + \frac{1}{a}]}{\Gamma[j + \frac{1}{a}]} \tag{47}$$

By Raabe Duhamel’s test, we obtain that the series for $\langle q_2 \rangle$ converges if $a < 1$, whereas the series for $\langle (q_2)^2 \rangle$ converges for $2a < 1$. These conclusions are identical with the LC-method results for $\lambda = 1$.

Since for $q_1 = 2$ and $q_2 = 1$ we have a series solution, it is useful to compare the original switch from Fig. 1 with the split-switch. Table I shows this comparison from which we recognize that although the phases are the same, the split-switch lags behind the original switch in terms of the accumulated number of q_2 molecules.

In connection with the results of Table I, it is interesting to observe the behavior of $\langle q_1 q_2 \rangle$. For the original switch and for $t \rightarrow \infty$, it follows that $\langle q_1 q_2 \rangle = 0$ if $\langle q_2 \rangle \rightarrow \text{const}$. This is so because, for $\langle q_2 \rangle$ to reach a steady state, the process needs to stop, which happens only when $q_1 = 0$. If the process of accumulation of q_2 never stops, we have $\langle q_2 \rangle \rightarrow \infty$, and, at the same time, $q_1 \neq 0$. This implies that $\langle q_1 q_2 \rangle \rightarrow \infty$. The conclusion is that in the limit $t \rightarrow \infty$, the correlation $\langle q_1 q_2 \rangle$ takes only two values, 0 or ∞ . The same result comes out of the LC-version of the switch. Equation (23) can be written as

$$F_{12}(t) = F_{12}(0) e^{-bt + \eta(1 - e^{-bt}) + a\lambda t} \left(\frac{F_2(t)}{F_2(0)} \right)^{1-\lambda} \tag{48}$$

From here, we see that $F_{12}(t) \rightarrow 0$ as $t \rightarrow \infty$ if $F_2(t) \rightarrow \text{const}$, which happens in phases 1 and 2. However, $F_{12}(t) \rightarrow \infty$ if $F_2(t) \rightarrow \infty$, which is the case for phase 3.

From the complete solution to the original switch for $q_1 = 2$ and $q_2 = 1$, it is also useful to better understand Fig. 7 and the accuracy of the statistical method to estimate the power-law exponent. From Table II, we see that, for $a = 0.5$, the estimated value is $\hat{\alpha} = 2.69$.

However, from the complete solution we get that for $a = 0.5$, the value for α actually is $\alpha = 3$. Indeed, from (34) for $\alpha = 3$, the mean value converges, whereas the variance diverges. The other threshold that comes from the complete solution is at $a = 1$ for which both the mean and the variance diverge. The power-law exponent, from (34) applied to this case, comes out as $\alpha = 2$. However, the statistically estimated power-law exponent from Table II that corresponds to $a = 1$ is $\hat{\alpha} = 1.76$. Both thresholds are underestimated for the initial conditions $q_1 = 2, q_2 = 1$ (Fig. 8). The same underestimation is visible in Fig. 7 for every initial condition. We do not have a general proof for the behavior of the original switch for any initial conditions so that we cannot claim that the underestimation appears for any initial conditions. However, we do have a general solution for the LC-switch that gives precise thresholds, which are independent of the initial conditions. Moreover, for $q_1 = 2$ and $q_2 = 1$, the LC-switch thresholds and the original switch thresholds coincide for $\lambda = 1$, which means that the LC method works better than the statistical estimation procedure. A justification for choosing $\lambda = 1$ is based on the work.¹⁵ Namely, the value of $0 \leq \lambda \leq 1$ takes its extreme values 0 or 1 when one of the complex-formation molecule is present in a low number than the other one. For this switch, $q_1 \rightarrow 0$, and so q_1 is the molecule that is lower in number than q_2 . In Ref. 15, it was found that the lower-number-molecule controls the process. Specifically, by placing $\lambda = 1$ in (15), only the term $a\lambda(z_2 - 1) \frac{F_{12}(t)}{F_1(t)} z_1 \frac{\partial F(z_1, z_2, t)}{\partial z_1}$ survives, whereas the other term, $a(1 - \lambda)(z_2 - 1) \frac{F_{12}(t)}{F_2(t)} z_2 \frac{\partial F(z_1, z_2, t)}{\partial z_2}$, is eliminated. This means that the master equation is driven only by q_1 since the control factor in the surviving term is $z_1 \frac{\partial F(z_1, z_2, t)}{\partial z_1}$. It is worth mentioning that the λ -parameter need not be equal to its extreme values of 0 and 1 when the LC-switch is embedded into a larger network. In this case, q_1 can be pumped up by interactions with other molecules and so it is not the only driver of the switch. Examples of this kind are presented in Ref. 15. Another reason for working with $\lambda \neq 0, 1$ is when experimental data are used for a statistical inference of the parameters of the network described by the LC method. In this case, the statistical inference may predict λ -parameters that are different from 0 or 1.

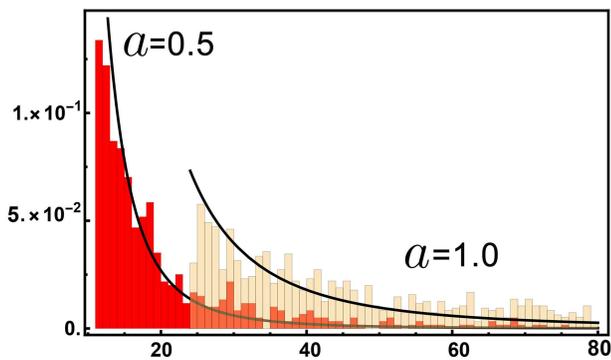


FIG. 8. Tails for $q_1 = 2$, $q_2 = 1$. The estimated power-law exponents, Table II, are $\hat{\alpha} = 2.69$ for $a = 0.5$ and $\hat{\alpha} = 1.76$ for $a = 1.0$ instead of the theoretical $\alpha = 3$ and $\alpha = 2$, respectively.

VI. CONCLUSION

We study the effects of a specific approximation procedure, called the LC method, on a molecular switch that transits through different phases as its coupling constant increases. The approximation, which produces a different switch, the split-node switch, maintains the phases of the original switch. Moreover, unlike the original switch, the split-node switch can be used as a building block biocircuit to assemble large networks. It is also worth emphasizing that the deterministic version of the switch loses completely the phase transition property, allowing only phase 1 to manifest itself for any strength of the coupling constant. The stochastic switch has a richer behavior than its deterministic counterpart.

REFERENCES

¹W. Z. Ouma, K. Pogacar, and E. Grotewold, "Topological and statistical analyses of gene regulatory networks reveal unifying yet quantitatively different emergent properties," *PLoS Comput. Biol.* **14**, e1006098 (2018).

- ²L. T. MacNeil and A. J. M. Walhout, "Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression," *Genome Res.* **21**(5), 645–665 (2011).
- ³C. Carré, A. Mas, and G. Krouk, "Reverse engineering highlights potential principles of large gene regulatory network design and learning," *NPJ Syst. Biol. Appl.* **3**, 17 (2017).
- ⁴J. J. Tyson, K. C. Chen, and B. Novak, "Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell," *Curr. Opin. Cell Biol.* **15**, 221–231 (2003).
- ⁵U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman and Hall/CRC, 2006).
- ⁶U. Alon, "Network motifs: Theory and experimental approaches," *Nat. Rev. Genet.* **8**, 450 (2007).
- ⁷A. Sommerfeld and P. Debye, *Probleme Der Modernen Physik* (Hirzel, Leipzig, 1928).
- ⁸W. H. Furry, "On fluctuation phenomena in passage of high energy electrons through lead," *Phys. Rev.* **52**, 569–581 (1937).
- ⁹A. Nordsieck, W. E. Lamb, Jr., and G. E. Uhlenbeck, "On the theory of cosmic-ray showers," *Physica* **7**, 344–360 (1940).
- ¹⁰N. G. V. Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier Science B.V., 1992).
- ¹¹S. Chandrasekhar, "Stochastic problems in physics and astronomy," *Rev. Mod. Phys.* **15**, 1 (1943).
- ¹²L. Arbiza, I. Gronau, B. Aksoy, M. Hubisz, B. Gulko, A. Keinan, and A. Siepel, "Genome-wide inference of natural selection on human transcription factor binding sites," *Nat. Genet.* **45**, 723–729 (2013).
- ¹³O. Lipan, J.-M. Navenot, Z. Wang, L. Huang, and S. C. Peiper, "Heat shock response in CHO mammalian cells is controlled by a nonlinear stochastic process," *PLoS Comput. Biol.* **3**, e187 (2007).
- ¹⁴J. Cates, G. C. Graham, N. Omattage, E. Pavesich, I. Setliff, J. Shaw, C. L. Smith, and O. Lipan, "Sensing the heat stress by mammalian cells," *BMC Biophys.* **4**, 16 (2011).
- ¹⁵C. Ferwerda and O. Lipan, "Splitting nodes and linking channels: A method for assembling biocircuits from stochastic elementary units," *Phys. Rev. E* **94**, 1–15 (2016).
- ¹⁶H. M. Sauro and K. Kim, "Synthetic biology: It's an analog world," *Nature* **497**, 572 (2013).
- ¹⁷S. Achimescu and O. Lipan, "Signal propagation in nonlinear stochastic gene regulatory networks," *IEE Proc. Syst. Biol.* **153**, 120–134 (2006).
- ¹⁸D. T. Gillespie, "A general method for numerically simulating the stochastic evolution of coupled chemical reactions," *J. Comput. Phys.* **22**, 403–434 (1976).
- ¹⁹A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM Rev.* **51**, 661–703 (2009).