The effect of coupled stochastic processes in a two-state biochemical switch

Garrett C. Graham · Ovidiu Lipan

Received: 11 October 2010 / Accepted: 2 May 2011 /
Published online: 11 June 2011
© Springer Science+Business Media B.V. 2011

Abstract Cell signaling pathways consist of multiple connections of different types of gene, mRNA and protein networks. It is not a trivial task to follow the signals flowing through these networks. The difficulty comes from considering the entire biological structure as a single network without breaking it into connected modules. The study of these networks simplifies if the complex system is reduced to a hierarchy of interconnected modules. Out of many potential modules, a specific one, namely the Goldbeter–Koshland switch, was encountered by the authors during their study of the Mammalian Heat Shock Response Network (MHSRN) where the switch acts as a stress sensor. Usually, only the steady state behavior of the switch is studied, in which the phosphorylated protein is given as a function of the enzyme concentration. Experimental results show that the heat shock response is still present 20 h after the temperature stress had ended. Thus, it is useful to analyze the transient behavior of the switch that couples the environment to the MHSRN. A stochastic model for the switch is proposed using the Master Equation which is subsequently transformed into an equation for the factorial cumulant generating function. This generating function can be easily read from a graphical representation of the stochastic switch. The second order approximation of the equation for the factorial cumulant generating function is solved and the time dependence of the transient regime of the mean and standard deviation is readily obtained. Using the mean and standard deviation of the switch's output as a function of the stochastic input signals that represent the environment, we classify the switches according to different criteria. The switches differ by the numerical values of the parameters that characterize the switch's chemical reactions. The classifying criteria will distinguish the switches by the levels of the response for a given transition time and by the sensitivity of the

G. C. Graham · O. Lipan (✉)
Department of Physics, University of Richmond, 28 Westhampton Way,
Richmond, VA 23173, USA
e-mail: olipan@richmond.edu

G. C. Graham
e-mail: garrett.c.graham@gmail.com

response to the enzyme levels. It is also found that the environment can drastically change the response of the switch, which has important biological consequences.

Keywords Goldbeter–Koshland switch • Stochastic • Heat shock

1 Introduction

In the study of biological phenomena, scientists have moved from investigating the “who and what” of such phenomena, to the “how” underlying the phenomena, and, now, to the “how much” associated with quantifying such phenomena [1, 2]. One type of biological phenomena that beg quantification are cell signaling pathways, although there are many other biological phenomena worthy of being studied. Cell signaling pathways consist of multiple connections of different varieties of gene, mRNA, and protein networks. It is not a trivial task to follow the signals flowing through these networks [3–6], because the gene, mRNA, and protein interaction networks are often poorly identified.

Experimentally, this need advanced the development of Systems Biology and launched the “-omics” era in molecular biology [7]. At the theoretical level, the complexity of cell signaling pathways has pushed the present computational capabilities to their limits. The difficulty arises in considering the entire biological structure as a single network without breaking it into connected modules. Modularity, besides being used extensively in engineering electronics networks, is employed in the study of complex regulatory networks [8–11]. The problem of a complex system is thus reduced to the study of a hierarchy of interconnected modules.

Another aspect of biological signals is that they are stochastic. Each molecular reaction is associated with a probability of being actually involved in the signal flow [12]. The stochasticity of genetic networks is being studied and experimental results sustain the concept [13–19]. Interconnection of stochastic modules presents itself as a useful approach to understand signaling pathways. Out of many potential modules, a specific one, namely the Goldbeter–Koshland switch, was encountered by the authors during their study of the Mammalian Heat Shock Response Network (MHSRN) [20, 21]. The MHSRN has been the subject of a vast amount of research over the past decades [22–26]. While many of the basic components of the MHSRN were identified by the mid-1990s, it was not until recently that mathematical models of the prokaryotic and eukaryotic heat shock response networks were attempted [20, 21, 27, 28].

Of crucial importance, change in temperature due to the external environment is treated in [20] as a signal that flows into the main heat shock network through a Goldbeter–Koshland [29] switch, Fig. 1. In other words, the coupling of the MHSRN to a different module, be it the environment or another regulatory network, is modeled by the Goldbeter–Koshland switch. However, this coupling is considered as being so rapid that the switch is practically always in a steady state. These different temperature-dependent steady states will influence the downstream behavior of the heat shock response. Experimental results in mammalian cells show that the heat shock response is still present 20 h after the temperature stress has ended [21]. Thus, it is useful to consider a switch that couples the environment to the MHSRN using time-dependent transitory signals and not just sudden jumps between steady states. Also, we need to consider a stochastic coupling, not a deterministic one.

Biological signals are, in general, stochastic, due to their molecular origin and random interactions. However, some biological signals are also very close to deterministic signals

Fig. 1 The coupling of the switch with an exterior input signal

like, for example, in stimulated neurons. In this study, however, our goal was to better understand the Goldbeter–Koshland switch from a stochastic point of view because our measurements for MHSRN show a large cell-to-cell variability in the response to heat shocks [21]. Another example of the importance of stochasticity is presented in [30], where cell-to-cell variability was observed in Ca^{2+} release activity in mammalian cells, which suggests the significance of regenerative mechanisms that amplify protein expression noise. To better understand the stochastic behavior, we also analyzed the effects of parameter variations on a deterministic Goldbeter–Koshland switch.

Our aim is to present a method by which the switch may be stochastically coupled to its environment or to another network and to then study the resulting transitory signaling process. Studying the steady-state levels of two proteins, W and W^* , which are catalyzed one into the other under the influence of enzymes E_1 and E_2 , is the usual approach taken in studying the Goldbeter–Koshland switch [29].

The enzyme E_1 may play the role of a kinase (an enzyme that transfers phosphate groups from high-energy donor molecules to specific substrates) whereas E_2 plays the role of a phosphatase (an enzyme that removes a phosphate group from its substrate). Generally, the behavior studied in this approach is exemplified as follows: if the switch begins in a steady state for which $W \gg W^*$ and if the relative concentrations of the enzymes E_1 and E_2 are modified, then the switch will transition to another steady state in which $W \ll W^*$. By controlling the levels of E_1 and E_2 , the switch flips between states of high and low levels of the molecule W^* . Biological systems have numerous uses for such switches, including registering changes in the extra- or intra-cellular environment.

The deterministic behavior of the switch as it flips between steady states is well studied, and a simple mathematical formula may be deduced for the steady states of W and W^* [29, 31]. The deterministic analysis of the switch's behavior begins with the fundamental biochemical reactions of Fig. 1. Here, beside the four molecular species W , W^* , E_1 , and E_2 , two additional species participate in the switching process. They are the two complexes W/E_1 and W^*/E_2 . From this biochemical diagram, a set of differential equations for the rate of change of all the six species is deduced. By setting the rates equal to zero, a system of

algebraic equations results. These equations determine the switch's steady state. The steady state, however, does not describe the transition of the switch from one state to another. The same deterministic equations can be used to study the transition of the switch, but the deterministic differential equations do not address the possible stochastic variability of the switching process. To cover this last aspect, we describe a stochastic model for the switch in what follows.

2 Reading the differential equation from the molecular diagram

Adding random noise to deterministic differential equations is commonly used to study stochastic models [32]. This approach is especially useful when the number of molecules involved is large. However, many biological systems operate with small populations of molecular species. For such cases, it is more appropriate to use a discrete stochastic model [15, 21]. To build such a model, we define the state of the switch, at any given time, by the set of population numbers for each of the six molecular species:

$$q = (q_1, q_2, q_3, q_4, q_5, q_6) \quad (1)$$

where the components of the vector q are $q_1 = W$, $q_2 = E_1$, $q_3 = W/E_1$, $q_4 = W^*$, $q_5 = E_2$, and $q_6 = W^*/E_2$.

The main goal is to follow the switch as it transitions through different states q , as the numbers of molecules increase or decrease.

The transition of the switch between different states is graphically represented in the molecular diagram of Fig. 2, where each ϵ -box represents a possible change of state. Each distinct reaction is identified by a variable ϵ_i and a transition probability is associated

Fig. 2 The stochastic biocircuit diagram for the switch coupled to the input signal. The *ovals* represent the molecular species; the *squares* represent the transitions. The *lines* from the transitions to the molecules indicate the effect of the transition on the state of the molecules: increase by *arrowed lines* and decrease by *barred lines*. The *lines* from the molecules to the transitions represent the intensity of the transition. This intensity is a polynomial in the number of molecules

Table 1 Description of the stochastic reactions for the molecular diagram of Fig. 2

Transition	Transition probability
$\epsilon_1 = (-1, -1, 1, 0, 0, 0)$	$T_{\epsilon_1} = a_1 q_1 q_2$
$\epsilon_2 = (1, 1, -1, 0, 0, 0)$	$T_{\epsilon_2} = d_1 q_3$
$\epsilon_3 = (0, 1, -1, 1, 0, 0)$	$T_{\epsilon_3} = k_1 q_3$
$\epsilon_4 = (0, 0, 0, -1, -1, 1)$	$T_{\epsilon_4} = a_2 q_4 q_5$
$\epsilon_5 = (0, 0, 0, 1, 1, -1)$	$T_{\epsilon_5} = d_2 q_6$
$\epsilon_6 = (1, 0, 0, 0, 1, -1)$	$T_{\epsilon_6} = k_2 q_6$
$\epsilon_7 = (0, 1, 0, 0, 0, 0)$	$T_{\epsilon_7} = G(t)$
$\epsilon_8 = (0, 1, 0, 0, 0, 0)$	$T_{\epsilon_8} = G_f(t) q_2$

with each reaction, shown in Table 1, considering the volume of the cell fixed to one unit volume.

To illustrate the rules for building the molecular diagram, we will describe the first reaction from Fig. 1, which represents the binding $W + E_1$, which in turn produces the complex W/E_1 . The state $q = \{q_1, q_2, q_3, q_4, q_5, q_6\} \equiv \{W, E_1, W/E_1, W^*, E_2, W^*/E_2\}$ changes to $q = \{q_1 - 1, q_2 - 1, q_3 + 1, q_4, q_5, q_6\}$, in which one molecule of W and E_1 is lost and one molecule of W/E_1 is produced. The change in the switch's state is written using the transition ϵ_1 , namely $q \mapsto q + \epsilon_1$, Table 1. The transition probability per unit time for this change is proportional with the product of the number of W molecules and E_1 molecules that take part in the complex formation: $T_{\epsilon_1} = a_1 q_1 q_2$. On the molecular diagram of Fig. 2, the transition ϵ_1 from Table 1 is depicted by the lines that originate on the ϵ_1 -box, terminating in an arrow on the increasing molecule and in a bar on the decreasing molecule. The transition probability $T_{\epsilon_1} = a_1 q_1 q_2$ is represented by two lines that originate on q_1 and q_2 , which join, representing the mathematical product $q_1 q_2$, and terminate on the ϵ_1 -box. Similar representation rules are used for all the other transitions from Table 1.

The general schematics of Fig. 2 not only explain the internal structure of the switch, but also take into account the external environment's influence on the switch. To act as a temperature sensor or a protein level detector, for example, the switch must be coupled to other regulatory networks. In this paper, we couple the switch with other regulatory networks by including new transitions in the state of the catalytic enzyme E_1 :

$$\epsilon_7 = \epsilon_8 = (0, 1, 0, 0, 0, 0). \quad (2)$$

Both ϵ_7 and ϵ_8 have the same effect of increasing E_1 by one, but each reaction has a different transition probability associated with it. The transition probabilities are $T_{\epsilon_7} = G(t)$ and $T_{\epsilon_8} = G_f(t) q_2$, respectively. Mathematically, the transition T_{ϵ_7} represents a direct coupling with the variable input signal $G(t)$, whereas T_{ϵ_8} represents a feedback coupling modulated by the time-dependent function $G_f(t)$. We emphasize here that, although the functions $G(t)$ and $G_f(t)$ are deterministic, the input on E_1 is stochastic. This is so because the functions $G(t)$ and $G_f(t)$ represent the transition probabilities that cause the state of E_1 to change randomly.

If only $T_{\epsilon_7} = G(t)$ is used as an input signal to drive E_1 , then the variance of E_1 will depend on the mean of E_1 , being connected to the mean through $G(t)$. In this case, only a limited class of stochastic behavior of E_1 may be taken into consideration. To cover the most general situation, we need to have input signals such that the mean and standard deviation of E_1 can, independently, be varied. A second input is thus necessary. In other words, the gene regulatory network or the fluctuating environment is succinctly represented by $G(t)$ and $G_f(t)$, and this representation is general enough to encompass a wide range of stochastic behavior for E_1 accumulation.

We could also introduce regulation through the second enzyme E_2 . However, because of the symmetry between E_1 and E_2 's switching dynamics, we connect the input signals on E_1 and analyze the switch's response in the presence of different E_2 levels.

Since we are interested not only in the mean numbers of the molecular species that comprise the switch, but also in the standard deviation and correlation functions, we need to find the probability $P(q, t)$ for the switch to be in the state q at time t [12, 33]. This time-dependent probability function is the solution of an equation that is constructed based on the structural schematics of the switch (Fig. 2).

The equation for $P(q, t)$ depends on all transition probabilities and represents the connection between the probabilities of the system at two consecutive times, $P(q, t + dt)$ and $P(q, t)$:

$$P(q, t + dt) = \underbrace{\sum_{\epsilon} [T_{\epsilon}(q - \epsilon, t) P(q - \epsilon, t) dt]}_1 + \underbrace{P(q, t) \left(1 - \sum_{\epsilon} [T_{\epsilon}(q, t)] \right) dt}_2 \quad (3)$$

The equation written above may be described as: (1) the sum of the product of the probability that the switch is in a state $q - \epsilon_i$, represented by $P(q - \epsilon, t)$, and the probability $T_{\epsilon_i}(q - \epsilon_i, t)$ that the reaction ϵ_i will occur during the time-increment dt while the switch is in the state $q - \epsilon_i$ and (2) the product of the probability that the switch is already in a state q at a time t , represented by $P(q, t)$, and the probability that none of the reactions ϵ_i will occur during the time-increment dt while the switch is in the state q , represented by $(1 - \sum_{\epsilon} [T_{\epsilon}(q, t)])dt$. By rearranging the terms and dividing both sides by dt , we obtain what is called the Master Equation for $P(q, t)$, [33].

To find the average values and standard deviations of the switch's molecular species as a function of time, we may proceed by first solving for $P(q, t)$ and then compute

$$\langle q_i \rangle \equiv \sum_{q=0}^{\infty} q_i P(q, t) \quad (4)$$

$$\langle q_i q_j \rangle \equiv \sum_{q=0}^{\infty} q_i q_j P(q, t) \quad (5)$$

for $i, j = 1 \dots 6$. However, this solution is very difficult and, at times, impossible. An easier method is to take the probability generating function of $P(q, t)$, $F(z, t) = \sum_{q=0}^{\infty} z^q P(q, t)$, [33]. Here, $z = (z_1, z_2, z_3, z_4, z_5, z_6)$ is a list of complex numbers, one for each molecule from the molecular diagram of Fig. 2, and $z^q = \prod_{i=1}^6 z_i^{q_i}$. From $F(z, t)$, the desired mean values are obtained by using partial derivatives at $z = 1$ (which represents the condition $z_i = 1$ for all $i = 1 \dots 6$). To simplify the notation, we will use $\partial_i F(z, t) = \partial_{z_i} F(z, t)$ and similar notations for higher order partial derivatives. In this way, we obtain:

$$\langle q_i \rangle = \partial_i F(z, t)|_{z=1} \quad (6)$$

$$\langle q_i q_j \rangle = \partial_{ij} F(z, t)|_{z=1}, \text{ for } i \neq j \quad (7)$$

$$\langle q_i(q_i - 1) \rangle = \partial_{ii} F(z, t)|_{z=1}. \quad (8)$$

It would be very advantageous to be able to form a closed system of equations that link a finite number of average values $\langle q_i \rangle$, $\langle q_i q_j \rangle$, $\langle q_i q_j q_k \rangle$, $\dots$, $\langle q_{i_1} q_{i_2} \dots q_{i_N} \rangle$. In that case, the mean and standard deviations for each molecular number would be precisely solved for. For the switch, this is not an option, and thus we need to find an approximate finite system

of equations to solve for the mean and standard deviation of the molecular components. This approximation is facilitated by a transformation from the generating function $F(z, t)$ to the function $X(z, t) = \ln(F(z, t))$. The advantage of using $X(z, t)$ instead of $F(z, t)$ is that $\partial_{i_1 i_2 \dots i_N} X(z, t)|_{z=1}$ grows with increasing N more moderately than $\partial_{i_1 i_2 \dots i_N} F(z, t)|_{z=1}$. Whereas eliminating all $\partial_{i_1 i_2 \dots i_N} F(z, t)|_{z=1}$, for N bigger than a chosen threshold, will produce a poor approximation, eliminating $\partial_{i_1 i_2 \dots i_N} X(z, t)|_{z=1}$ will not.

Obviously, the equation for $P(q, t)$, Eq. 3, will translate into an equation for $X(z, t)$. Although the discussion above is necessary to understand the logical steps to obtain $X(z, t)$, in practice the differential equation for $X(z, t)$ can readily be obtained from a visual inspection of the molecular diagram from Fig. 2. We first present the entire equation, which contains the parameters from Table 1, and then explain its relation with the molecular diagram from Fig. 2.

$$\begin{aligned} \frac{\partial X(z, t)}{\partial t} = & (z_1^{-1} z_2^{-1} z_3 - 1) a_1 z_1 z_2 (\partial_{12} X + \partial_1 X \partial_2 X) + (z_1 z_2 z_3^{-1} - 1) d_1 z_3 \partial_3 X \\ & + (z_2 z_3^{-1} z_4 - 1) k_1 z_3 \partial_3 X + (z_4^{-1} z_5^{-1} z_6 - 1) a_2 z_4 z_5 (\partial_{45} X + \partial_4 X \partial_5 X) \\ & + (z_4 z_5 z_6^{-1} - 1) d_2 z_6 \partial_6 X + (z_1 z_5 z_6^{-1} - 1) k_2 z_6 \partial_6 X + (z_2 - 1) G(t) \\ & + (z_2 - 1) G_f(t) z_2 \partial_2 X. \end{aligned} \quad (9)$$

Each molecule in Fig. 2 must be associated with a z -variable, which will become an argument of the $X(z, t)$ in Eq. 9. This association is visible in Fig. 3 where two terms from Eq. 9 are produced. Each ϵ -box from Fig. 2 will generate one term in Eq. 9. Each term contains a factor of the form $(z^\epsilon - 1)$, which in an expanded version is $(z_1^{\epsilon_1} z_2^{\epsilon_2} z_3^{\epsilon_3} z_4^{\epsilon_4} z_5^{\epsilon_5} z_6^{\epsilon_6} - 1)$. For the ϵ_1 -box in Fig. 3, this expression is $(z_1^{-1} z_2^{-1} z_3 - 1)$ and is associated with the lines that originate on the ϵ_1 -box and end on the molecule. There are three such lines. Two repression lines end in a bar on z_1 and z_2 ; from here we get the negative exponents in $z_1^{-1} z_2^{-1}$. The other line ends in an arrow, so z_3 gets a positive exponent. The same is true for the ϵ_2 -box: the line ends with an arrow on E_1 , so the resulting term is $(z_2 - 1)$. Next, we need to associate mathematical terms with the lines that originate on the molecules and end on an ϵ -box. For ϵ_1 -box in Fig. 3, there are two lines that start on W and E_1 and fuse together before they reach the ϵ_1 -box. This fusion represents the product $z_1 z_2$ in the term $a_1 z_1 z_2 (\partial_{12} X + \partial_1 X \partial_2 X)$. The last factor, namely $(\partial_{12} X + \partial_1 X \partial_2 X)$, represents $\partial_{12} F$ transformed into X by $F = \ln(X)$. The fusion of the lines originating from W and E_1 not

Fig. 3 Each transition box in the molecular diagram of Fig. 2 generates a term in the differential equation for $X(z, t)$. Here, the terms for ϵ_1 and ϵ_8 are written below the corresponding transition. Compare with Eq. 9

only gives us the product $z_1 z_2$, but also demands that the partial differentiation is ∂_{12} . For ϵ_2 -box, there is one line with no fusion, so we will have z_2 multiplied by $\partial_2 X$. The numbers a_1 and the function $G_f(t)$ will multiply the respective terms and incorporate the strength of the transition probability.

Once Eq. 9 is available, a system of ordinary differential equations, for variables like Eq. 6, is generated by taking partial derivatives of both sides of Eq. 9. The mean value of W^* equals $m(t) = \partial_4 F(z, t) |_{z=1} = \partial_4 X(z, t) |_{z=1}$, which, for short, is denoted by $X_4(t)$. In general, we use the simple notation $X_{i_1 i_2 \dots i_N}(t) = \partial_{i_1 i_2 \dots i_N} X(z, t) |_{z=1}$. From these functions, the standard deviations and other correlation functions can be obtained. An important quantity for this study is the standard deviation of the molecule W^* , which is given by $\sigma(t) = \sqrt{X_4(t) + X_{44}(t)}$. To obtain the mean $m(t)$ and the standard deviation $\sigma(t)$ as a function of time, we need to solve a system of differential equations for the functions $X_{i_1 i_2 \dots i_N}(t)$. Two such equations are presented below.

$$\frac{d}{dt} X_4(t) = k_1 X_3(t) - a_2 (X_{45}(t) + X_4(t) X_5(t)) + d_2 X_6(t) \quad (10)$$

$$\begin{aligned} \frac{d}{dt} X_{44}(t) = & 2 k_1 X_{34}(t) - 2 a_2 (X_{445}(t) + X_{44}(t) X_5(t) + X_4(t) X_{45}(t)) \\ & + 2 d_2 X_{46}(t). \end{aligned} \quad (11)$$

As mentioned above, we must resort to an approximation in order to work with a finite number of functions. For example, if the third order derivative $X_{445}(t)$ is eliminated from Eq. 10, we obtain a second order approximation. The second order approximation is necessary to obtain the standard deviation of the molecular species. It is also sufficient when the third order variables X_{imn} are small enough that their elimination will not distort the mean and standard deviation computations. We do not know of any general rule that will set the conditions upon which the third and higher orders may be neglected. However, for specific gene regulatory networks the third order variables X_{imn} may be numerically computed. If the values are small compared with the second order variables, they may be neglected. The approximation is validated also by comparing the results with the Monte Carlo simulations for the genetic network for some parameter values. For the present switch, the second order approximation results matched the Monte Carlo simulations. We did not compare the Monte Carlo simulations for all parameter values with the analytic approximation because we do not have extensive computational capabilities.

The main goal of this manuscript is to study the influence of stochastic input signals ($G(t)$, ϵ_7) and ($G_f(t)$, ϵ_8). This influence propagates on the molecule W^* through the complex W/E_1 , which is dependent upon the levels of the enzyme E_1 . This is visible below, where the mean value of E_1 is X_2 and of W/E_1 is X_3 .

$$\begin{aligned} \frac{d}{dt} X_2(t) = & -a_1 (X_{12}(t) + X_1(t) X_2(t)) + d_1 X_3(t) + k_1 X_3(t) \\ & + G(t) + G_f(t) X_2(t) \end{aligned} \quad (12)$$

$$\frac{d}{dt} X_3(t) = a_1 (X_{12}(t) + X_1(t) X_2(t)) - d_1 X_3(t) - k_1 X_3(t). \quad (13)$$

The influence of the coupling is not restricted to the mean values only. The standard deviation of the enzyme E_1 , $\sqrt{X_2(t) + X_{22}(t)}$, depends on the fluctuating environment not only through $X_2(t)$ but also through $X_{22}(t)$, which the feedback $G_f(t)$ directly influences:

$$\begin{aligned} \frac{d}{dt} X_{22}(t) = & -2a_1(X_{122}(t) + X_{12}(t)X_2(t) + X_1(t)X_{22}(t)) + 2d_1X_{23}(t) \\ & + 2k_1X_{23}(t) + 2G_f(t)X_2(t) + 2G_f(t)X_{22}(t). \end{aligned} \quad (14)$$

Therefore, the fluctuating environment or the gene regulatory network is stochastically coupled with the switch.

The partial derivatives of $F(z, t)$ are referred to as factorial moments, and those of $X(z, t)$ as factorial cumulants. Thus, by using the equation for $X(z, t)$, we find the time evolution of the factorial cumulants of the system. Numerically solving for the factorial cumulants of this system up to a chosen order gives the time evolution of the mean behavior, along with the standard deviation.

Each specific switch has its own numerical values for the parameters $a_1, d_1, a_2, d_2, k_1, k_2$, as shown in Fig. 2. These parameters specify the intensity of the transition probabilities for each switch. The dynamical response of the switch to the input signal that enters through E_1 drastically depends on these parameters. An aim of this study is to understand the stochastic response of the switch as a function of these six parameters. We reduce the number of parameters from six to five by scaling the time variable:

$$t_{\text{new}} = k_1 t \quad (15)$$

Any one of the six parameters can be used to scale the time. We choose k_1 since it governs the production of W^* , which is modulated by the input signal coupled to E_1 . In other words, the unit of time will be scaled relative to the order of magnitude of the transition probability T_{e3} . For simplicity, in what follows we use the symbol t for t_{new} .

The time scaling implies that all equations be divided by k_1 . Thus, five parameters are sufficient to distinguish between any two switches. So as to explore a wide swath of the five-dimensional space of parameters, we use the logarithm scale and, thus, identify a switch by the following set:

$$\left\{ \mathbf{a}_1 = \log_{10} \left(\frac{a_1}{k_1} \right), \mathbf{d}_1 = \log_{10} \left(\frac{d_1}{k_1} \right), \mathbf{a}_2 = \log_{10} \left(\frac{a_2}{k_1} \right), \right. \\ \left. \mathbf{d}_2 = \log_{10} \left(\frac{d_2}{k_1} \right), \mathbf{k}_2 = \log_{10} \left(\frac{k_2}{k_1} \right) \right\} \quad (16)$$

The region covered in the five-dimensional space is a grid of $5^5 = 3,125$ points, by letting each parameter in the set $(\mathbf{a}_1, \mathbf{d}_1, \mathbf{a}_2, \mathbf{d}_2, \mathbf{k}_2)$ assume the following values: $-2, -1, 0, 1$, and 2 . Thus, the range of the unscaled parameters covers four orders of magnitude: $0.01, 0.1, 1, 10$, and 100 .

As stated, we are interested in the stochastic behavior of the switch at low numbers of molecules. A specific feature of the switch is that the molecule W cannot be created or destroyed. Instead, it is either transformed into W^* or bound in the complex W/E_1 . Consequently, we can study the switch's behavior with low numbers of molecules W and W^* by simply constraining the initial conditions of the switch to a low number of W and W^* at $t = 0$. However, this constraint on the molecular population does not necessarily hold for the enzyme E_1 because it is coupled with another input network. The range of values for E_1 can widely vary, depending on the strength of the input signals $G(t)$ and $G_f(t)$. Since

the transitions governed by these input signals can only increase E_1 and W/E_1 , we begin with: $E_1 = 0$ and $W/E_1 = 1$ at $t = 0$. We then allow E_1 and W/E_1 to accumulate as time progresses. If the initial number of W is greater than W^* , then the accumulation of E_1 and W/E_1 will flip the switch from one state to another. We choose $W = 100$ and $W^* = 0$ at time $t = 0$ in order that a flipping of the switch may be observed and in order that the condition for a small total molecular population may be fulfilled.

The enzyme E_2 limits the accumulation of W^* by catalyzing the reaction that converts W^* back to W . We regard the levels of E_2 as regulating the speed at which the switch transitions to a state with high W^* . In order to observe the regulatory effects of E_2 , we study three scenarios: $E_2 = 0, 1$ or 10 . The case $E_2 = 0$ is unique in the sense that the reaction converting W^* back to W does not exist. This case, which is a simple catalytic process with no switching effects, will be contrasted with the situation $E_2 = 1$ to emphasize the dynamical changes when both enzymatic reactions occur in the same switch. The $E_2 = 10$ situation will show the power of E_2 to modulate the response of the switch to the input signals.

The dynamical transition of the switch from low numbers of W^* to high numbers depends on the strengths of the input signals $G(t)$ and $G_f(t)$. Notice here that we work with scaled inputs: $G(t)/k_1$ is simply denoted by $G(t)$; we choose $G_f(t)$ to denote the equivalent scaled input for the feedback. The method presented above works for any positive functions $G(t)$ and $G_f(t)$. However, in order to compare a large class of different switches, we fix the input signals. We choose one of the most simple yet useful of inputs, the step input. For the step input, the signal is zero before $t = 0$ and constant for positive time; i.e., $G(t) = G$ and $G_f(t) = G_f$. The step input is intended to mimic the temperature impulses delivered to mammalian cells in experimental work.

The level of the input signals dictates the switching time interval, during which the molecule W^* population transitions from 0 to 100 molecules. We will measure this switching time in terms of k_1 because this parameter governs the production of W^* . In terms of the real time (not the time scaled as in Eq. 15), the switching interval can be short, and span a range up to $1/k_1$, or larger, and go beyond $1,000/k_1$. For this study, we choose to analyze switches with moderate transition times of up to $50/k_1$. The switching range is controlled by the level of the input signals $G(t) = G$ and $G_f(t) = G_f$. In terms of the unitless time (Eq. 15), the transition range is 0 to 50 and depends on the levels of the scaled input signals G/k_1 and G_f/k_1 . We found, through numerical studies, that the unitless switching time spans a range from 0 to 50 if the scaled input levels G/k_1 and G_f/k_1 fall between 0 and 0.4.

Having defined the constraining factors for molecular populations and the structure of the input signals, we can now study the switch's transition profiles.

3 Results

3.1 Mutation-inducing stochastic survival strategies displayed by high-variability switch $\{-2, 2, -2, 0, 0\}$

Since our study focuses on the stochastic characteristics of the switch, we searched for a combination of parameters $(\mathbf{a}_1, \mathbf{d}_1, \mathbf{k}_1, \mathbf{a}_2, \mathbf{d}_2, \mathbf{k}_2)$ demonstrating a strong variation in the switch's response. Such a "high-variability switch" is found to have the following values for its log-scaled parameters:

$$\{\mathbf{a}_1 = -2, \mathbf{d}_1 = 2, \mathbf{a}_2 = -2, \mathbf{d}_2 = 0, \mathbf{k}_2 = 0\}. \quad (17)$$

The meaning of high-variability switch becomes clear when the responses of a large class of switches are compared, as is done later in this paper. First, we define the output of the switch as being the mean value of W^* , denoted $m(t)$, and its standard deviation, denoted $\sigma(t)$, as a function of time along the transition process. We implemented a program in Mathematica to compute the switch's output by solving for the factorial cumulants up to second order. The mean value of $m(t)$, along with $m(t) + \sigma(t)$ and $m(t) - \sigma(t)$ are displayed in Fig. 4 for a typical response together with a set of Monte Carlo simulations. We notice that Eq. 9 produces the correct band into which the majority of Monte Carlo simulations fall. It is worth mentioning here that the mean value and the standard deviations were produced extremely fast by solving the second order approximation to Eq. 9. The Monte Carlo simulations take a longer time and necessitate subsequent data analysis to obtain the mean and standard deviation for the response. The inset in Fig. 4 shows the rapid accumulation of the enzyme E_1 at the end of the time course. This phenomenon will be analyzed below in connection with Fig. 7.

The time evolution from a low level of W^* to a high level depends drastically on the switch's parameters and on the strength of the input signals, as the results from Fig. 5 show. This time evolution depends on the input signal G and G_f , so that the response will be a function of three variables, (G, G_f, t) . For a fixed value of G_f , the response represents a 3-dimensional surface in the variables (G, t) . In Fig. 5, the upper and the lower surfaces represent $m(G, t) + \sigma(G, t)$ and $m(G, t) - \sigma(G, t)$, respectively, whereas the middle surface is $m(G, t)$. For a low G_f the response is close to deterministic, as in Fig. 5a, in which case $\sigma(t)$ is small and the band between the upper, $m(t) + \sigma(t)$, and lower, $m(t) - \sigma(t)$, becomes very narrow. However, increasing the value of G_f , the band becomes very large, implying a high heterogeneity in the cell population. The departure from a deterministic case depends not only on the input signals, but also on the parameters that describe the switch.

Fig. 4 A typical stochastic response for the W^* enzyme. The mean and the standard deviation computed from $X(z, t)$ characterize very well the Monte Carlo simulations. The inset shows the accumulation of E_1 and its stochastic character. The switch's parameters and the input signals are $\{a_1 = -2, d_1 = 2, a_2 = -2, d_2 = 0, k_2 = 0\}$ and $G = 0.05, G_f = 0.2$

Fig. 5 Response depends strongly on the input signal and on the feedback strength. A 100-fold increase in G_f from (a) to (b) changes the shape of the response. c shows a perspective of (b) viewed at higher values on the G -axis. The plateaus for the plots for E_1 on the second row do not represent saturation phenomena, but rather the intersection of the E_1 accumulation with the upper horizontal plane

It is interesting to note the importance of the feedback. A change in magnitude in the feedback drastically changes the response profile, as seen when comparing Fig. 5a, b. The high variability of this specific switch refers not only to the huge standard deviation for mild G inputs, but also to the abrupt change in the standard deviation when the G input goes beyond a threshold, as seen in Fig. 5b, c. The time evolution of a low-feedback circuit does not display the abrupt dependence of a high-feedback switch to the input signal G , as seen in Fig. 5a.

Observing these characteristics of high- and low-feedback circuits, it is possible to consider that an organism could employ a high-feedback switch to its advantage.

The enormous standard deviation of the switch, Fig. 5b, c, under mild G inputs allows this high-feedback circuit to assume nearly any behavior within the range of possible behaviors. Such variability, over the long term, would create a broad range of morphologies among an otherwise homogenous cell population.

When considering the response to severe inputs, depicted by high values of G in Fig. 5b, c, the mean value immediately begins climbing towards and quickly attains the maximum possible value for W^* ; the mean's curve is closely guarded by the upper and lower one-standard deviation bounds. This type of response from the switch guarantees that, no matter how much its behavior has varied under mild conditions, under severe conditions the switch always adheres to nearly exactly the same response. These observations suggest that the high-feedback, high-variability switch maximizes the possibility for divergent behaviors under mild conditions, while limiting the system's behavior to a narrow range of fast responses under severe conditions, which guarantees the survival of a maximum number of the switch's host organism's population. Thus, the switch promotes both adaptation and survival while maximizing the chances for success of either of these goals, depending upon the current input conditions. Whereas the molecule W^* has the above described behavior, the enzyme E_1 does not behave similarly. There are no abrupt changes in E_1 as a function of the input G for both low and high-feedback inputs. The only difference is that the number of E_1 molecules reaches higher levels for high values of G , Fig. 5a, b. In relation to this analysis it worth mentioning the study [34] on noise-induced bistability in biomolecular

systems as well as the study [35] where it was shown that increased variability in gene expression can be beneficial after an acute change in environmental conditions.

3.2 Classification of switches: fast and slow

For a better understanding of possible responses to input signals of a large class of switches, the behavior of the high-variability switch must be compared with the responses of a wide set of switches. To this end, we simulated the responses of $5^5 = 3,125$ different switches to the same stochastic input step functions that were used above to investigate the behavior of the high-variability switch.

To this extent, compare the different switches depicted in Fig. 6, which was created with the help of dChip [36]. Each row represents the output of a different switch (different values for the switch's parameters $\{a_1, d_1, a_2, d_2, k_2\}$), which dChip has clustered and collectively displayed. Different shades indicate different severities in response: a darker hue indicates higher values of W^* than the lighter and near-white shades, a few rows of which appear in the upper rows of the figure. A subset of the 5^5 switches are displayed in Fig. 6 to show that the response depends on the switch's parameters. Thus, we observe that the darker rows marked "Fast" in the figure indicate a stronger, more immediate response than the nearly blank rows at the top, marked "Slow". Responses that gradually interpolate between the fast and slow switches are also present in Fig. 6. At this point we do not have a mathematical definition of "Fast" and "Slow" circuits because we use the dChip as an exploratory tool to reveal different possible classes of switches' behavior. The dChip study, as appears in Fig. 6, is to be taken as a qualitative source for future quantitative definitions.

Defining stochastic processes as being either "on" or "off" is nontrivial. Sometimes the mean of a switch's behavior might never grow large enough to trigger a threshold-based reaction within the cell, while its mean plus one standard deviation may be higher than the

Fig. 6 Stochastic responses for fast and slow switches

trigger threshold. Moreover, the terms “on” or “off” are suitable to express the steady states of the switch for deterministic models. Since we are interested in the entire transition of the switch as a time variable process, we classify the switches as “fast” or “slow” and not “on” or “off.” This classification depends on the entire dynamical behavior of the switch during its transition. The intuitive conception of a fast stochastic switch is that the lower bound $m(t) - \sigma(t)$ grows quickly, as contrasted with a slow stochastic switch, whose upper bound $m(t) + \sigma(t)$ grows slowly. Precise definitions for the classes of fast and slow stochastic switches will be given below, after we introduce some mathematical measures.

Although the hierarchical clusterings, as seen in (see Fig. 6), can give us a global view of the evolution of the entire class of switches, it is desirable to capture the response with a simple metric, which could then be used as a comparison tool [37]. One such metric is the l^2 -Norm of the mean $m(t)$ and standard deviation $\sigma(t)$. For practical purposes, the numerical format of the response consists of sets of samples $\{m(t_0), m(t_1), m(t_2), \dots, m(t_n)\}$ and $\{\sigma(t_0), \sigma(t_1), \sigma(t_2), \dots, \sigma(t_n)\}$ taken at discrete time points $(t_0, t_1, t_2, \dots, t_n)$. We will thus use the discrete version of the l^2 -Norm and arrive at the following measure of the response:

$$\|f\| = \frac{\sqrt{\sum_{k=1}^n f(t_k)^2}}{\sqrt{\sum_{k=1}^n f_{\max}^2}}. \quad (18)$$

Here, $f(t)$ may represent $m(t)$ or $2\sigma(t)$. The results that follow do not depend on the normalization by $f_{\max}$ or on the sampling length n . We divide by the measure of the maximum value to obtain a scaled norm that is less dependent on the model-specific numerical conditions. For our chosen initial conditions, the maximum value is 100 in both cases and the length of the sampling sequence is $n = 47$. With the help of this measure, the intuitive idea about fast and slow switches becomes precise.

We define a fast switch from a population of switches to be that subset of switches for which $\|m - \sigma\|$ is in the greatest 10% for that population. Similarly, for the subset of slow switches, $\|m + \sigma\|$ is in the lowest 10% for that population.

3.3 Stall-based strategies for fast switches

The obvious disparities between the responses of fast and slow switches, as shown in Fig. 6, begs the question: why are the fast switches fast and the slow switches slow? We investigate the strategies employed by fast switches, which are the switches most likely provide the greatest chance of survival under threatening conditions. To hypothesize potential fast-switch strategies, examine Fig. 1. For a switch to be fast, it must produce as much W^* as rapidly as possible. An obvious strategy to accomplish this goal would be to make the binding rate in the $W + E_1 \rightarrow W/E_1$ reaction, a_1 , as large as possible, while minimizing its inverse reaction rate, d_1 . In this way, the substrate would be catalyzed by the enzyme at a high rate, with little or no chance for the reaction to fail. Thus, this strategy would ostensibly produce a high level of W^* .

However, an analysis of the fast switches in the population reveals a drastically different story. An overwhelming majority of the switches, 77%, have values of $\frac{a_1}{d_1} \leq 1$, while 90% have values of $\frac{a_1}{d_1} \leq 10$. Recall that since $\frac{a_1}{d_1} \in \{10^{-4}, \dots, 10^4\}$, $\frac{a_1}{d_1}$ could equal as much as 10,000, which is a value our initial hypothesis for fast switch strategies suggests would be more likely to be used in fast switches. To understand such a surprising result, we must investigate the strategies of these fast switches, especially since nearly 51% of the fast switches have values of $\frac{a_1}{d_1} \leq 0.10$.

Fig. 7 This diagram shows the reactions of an averagely constructed fast switch in chronological order, as they occur. It reveals the degree to which fast switches rely upon feedback to accelerate their response. Each dot represents one reaction with the corresponding reaction listed on the vertical axis. Note the tremendous number of alternations between the $W + E_1 \rightarrow W/E_1$ and $W/E_1 \rightarrow W + E_1$ reactions, as marked by ϵ_1 and ϵ_2 on the vertical axis. This stalling tactic, designed to allow the feedback loop to gain strength and represented by ϵ_8 on the y axis, is what makes this switch one of the fastest switches in the population

To find what strategy these switches use to boost their rate of activation, we run Gillespie Monte Carlo simulations of the switch. Examination of the results of the Monte Carlo simulations shows that the majority of fast circuits rely heavily upon the switch's positive feedback and use a counter-intuitive strategy of stalling in order to open more quickly.

Figure 7 shows the step-by-step reactions from a simulation for an archetypically structured fast circuit, with $\frac{a_1}{d_1} \leq 0.10$. Each dot represents a reaction associated with the corresponding reaction rate on the vertical axis. The dots associated with a_1 and d_1 are clustered so thickly that one could mistake them for thick lines. Here, the switch rapidly jumps between producing W/E_1 complexes and dissociating them back into $W + E_1$. The jump between these two inverse reactions allows the positive feedback loop, ϵ_8 , which produces E_1 , to gain strength; the feedback transition probability $T_{\epsilon_8} = G_f q_2$ gains strength the longer the input signal lasts because q_2 has enough time to grow (compare with the inset in Fig. 4). Thus, if the stalling behavior keeps the activated product from forming, the feedback loop gains power and after a time, enough of the activating enzymes exist in the circuit that huge quantities of W^* are produced.

3.4 Response plane description of the switch behavior under parameter variations and E_2 fluctuation

The set of parameters that define a specific switch $\{a_1, d_1, a_2, d_2, k_2\}$ may not be precisely known in an experimental setting. Moreover, these parameters may vary under the influence of the morphological changes in the cell or simply under the influence of factors that couple the switch with other signaling pathways. Thus, the question of how susceptible the fast switches' reaction rates are to environmental fluctuations is of critical biological importance. Extreme susceptibility is a frightening prospect for any living organism to whom the susceptible switch is crucial for survival. If the reaction rate of such a circuit varies under thermal stress, for instance, the fast circuit might suddenly become a slow circuit and a potentially lethal weakness to the organism. Thus, we desire to determine

Fig. 8 The mean values of W^* for every switch in our population of 3,125. An explosive divergence of behavior occurs when the levels of E_2 increase from $E_2 = 0$ to $E_2 = 1$, for both stochastic and deterministic models. The deterministic mean values reach higher levels in a shorter time than the mean values predicted by the stochastic model

if the fast and slow switches are easily separated in the five-dimensional space in which all of these switches are located by their identification set $\{a_1, d_1, a_2, d_2, k_2\}$. In this space of parameters, we define the nearest neighbors to $\{a_1, d_1, a_2, d_2, k_2\}$ as the points $\{a_1 \pm 1, d_1 \pm 1, a_2 \pm 1, d_2 \pm 1, k_2 \pm 1\}$.

In determining whether fast and slow switches are sufficiently separated as to be impervious to environmental instability, we seek to guarantee that nothing but a twofold magnitude or greater shift in reaction rates could cause a fast switch to morph into a slow switch.

We trained a support vector machine (SVM) to recognize the fast and slow switches in our population. The SVM was able to construct a hyperplane that separated the fast and slow switches.

Another possible fluctuation that may influence the switch is the one connected with the second enzyme E_2 . We choose to connect the switch to an external regulatory network through the enzyme E_1 . The model may be changed to also include a stochastic, time-dependent connection through E_2 . However, instead of increasing the complexity of the connections, we choose instead to preserve E_1 as the coupled circuit component, allowing input signals to enter the circuit schematic only through that enzyme. In order to

Fig. 9 All circuits in response space. **a** Generated using half of the sampling time points used to generate **(b)**. The organization of the switches in the response plane does not depend on the sampling length n used in the definition of the global measure (Eq. 18). The limited varieties of switch responses are readily apparent for $E_2 = 0$. The presence of the second enzyme, $E_2 = 1$, has a major impact on the variety of behaviors generated. Notice the slight shift in behaviors, with more switches migrating toward the origin for $E_2 = 10$

study the stochastic time-dependent connection through E_2 , we examine the influence of different levels of E_2 on the circuit's output.

The case $E_2 = 0$ is unique: when $E_2 = 0$, the inverse reaction that transforms W^* back to W is eliminated. The presence of a single molecule E_2 reinstates the inverse process and the presence of ten molecules of E_2 strengthens it. In order to remain within the range of low molecular numbers, we consider three levels for the enzyme $E_2 = 0, 1, 10$, in what follows. Considering the mean value of the response, we observe the rapid divergence from a small number of relatively similar behaviors (Fig. 8, $E_2 = 0$) to an immense range of profiles (Fig. 9, $E_2 = 1$). A good opportunity appears here for a comparison between the stochastic model and the deterministic one. Besides the fact that the deterministic model cannot generate the standard deviations, it generates the mean value with some errors. The deterministic mean values in Fig. 8, for both cases, $E_2 = 0$ and $E_2 = 1$, reach higher levels in a shorter time than the mean values predicted by the stochastic model. The stochastic model has thus an influence on the mean values and is to be preferred when an accurate description of the switch is needed.

Instead of viewing the circuits' responses from the somewhat overwhelming perspective of the total time evolution (Fig. 8), it is helpful to view it in the space of the l^2 -Norm of the mean of each circuit plotted against the l^2 -Norm of twice that circuit's standard deviation. We use the doubled value of the standard deviation instead of just the standard deviation, to show the band in which the majority of the stochastic paths will fall in a Monte Carlo simulation (Fig. 4). This 2-dimensional plane of global measures $||m||$ and $||2\sigma||$ of W^* , giving an instantaneous picture of how fast and how precise a switch's response is, will be called the response plane. In principle, instead of the global measures we could have used other parameters to represent the switch's response. For example, if the parametric plot $(\sigma(t), m(t))$ would have been a straight line, we could have used the slope of that line as a measure. However, this is not the case, the parametric plot $(\sigma(t), m(t))$ is not linear and highly depends on the biochemical parameters of the switch. We emphasize that the global representation of a switch in the response plane is computed here through the Master Equation and not using Monte Carlo simulations. Through this procedure we drastically reduce the computation time. Each switch is represented as a point in this response plane. Figure 9 clearly shows the extreme sensitivity of these circuits to changes in enzymatic concentrations.

Fig. 10 Migration of the high-variability switch. As the circuit migrates from its original position at $E_2 = 0$ to its position when $E_2 = 1$, it attains the highest $||2\sigma||$ -value reached by any switch in this population. The plot diagrams its further migrations to its positions at $E_2 = 10$ and $E_2 = 100$

Fig. 11 Migration of fast (a) and slow (b) switches

The presence or absence of the enzyme E_2 has a strong impact on the response. Few dots in Fig. 9 represent the responses for $E_2 = 0$, and they are located mainly on the right side. From these locations, the representative points migrate and spread to occupy an extended region down to the origin, as E_2 increases. The global migration of the representative points toward the origin is intensified once E_2 reaches the level of ten molecules. Many features previously discussed may now be visualized in this response plane.

The high-variability switch $\{-2, 2, -2, 0, 0\}$ occupies the apex position in the $E_2 = 1$ response plane of Fig. 10. Its standard deviation measure is the highest among the switches with a mean located in the center of the plane. The migration of this switch for increasing values of E_2 traverses a path along the exterior boundary of the region of the representative points. For $E_2 = 0$, the switch still has a high standard deviation among its peers, a position maintained for increasing E_2 enzyme concentration. The power of the response plane is that the stochastic behavior of a particular switch or a subclass of switches may be visualized against the whole set of switches as one or more parameters vary.

Compared with the high-variability switch, $\{-2, 2, -2, 0, 0\}$, the subclass of fast switches, which is concentrated on the right side of the response plane, is less displaced for increasing E_2 , as shown in Fig. 11a. The fast switches occupy the leftmost islands on the $E_2 = 1$ response plane. For $E_2 = 10$, they diverge towards a lower mean value, keeping though almost the same standard deviation. For slow switches, the migration is dramatic, as shown in Fig. 11b. Starting at $E_2 = 0$ from the right side, where all switches are located, they

Fig. 12 Migration of stable (a) and unstable (b) switches

Fig. 13 Nearest neighbors of fast and slow switches

collapse together near the origin at $E_2 = 1$ and continue their migration and compression towards the origin for $E_2 = 10$.

These kinds of circuits may be employed by different organisms as detectors for low levels of the second enzyme E_2 . Their response shuts down for a trace presence of this enzyme and remains inoperable for increasing values. The fast circuits are less dependent on E_2 and may be used as a robust switch against E_2 variation.

This difference in stability between the fast and slow switches raised questions on how to classify the switches according to their stability with regards to E_2 fluctuation. Switches migrate in the response plane from their initial positions at $E_2 = 0$ to two other positions at $E_2 = 1$ and $E_2 = 10$. These three positions form a triangle, which is small for switches that do not depend on E_2 . We will call such switches stable. As a measure of stability, we take the perimeter of this triangle and sort the switches according to this stability measure. For small perimeters, the points in the response plane will be clustered close to each other. The most stable switches, (Fig. 12a), which occupy the lowest 10% of the stability measure, are located on the right-most islands on the response plane, with one exception. The most unstable circuits, which occupy the highest 10% of the stability measure, have a wider spread than the slow switches (Fig. 12b).

In terms of stability, it is interesting to locate the nearest neighbors of the fast and slow switches to visualize the separation between these two classes in the response plane. The SVM analysis mentioned above is supported by Fig. 13 where the neighbors of the fast switches are located mostly in the right region, whereas the neighbors of the slow switches are located to the left.

4 Conclusions

The coupling of the Goldbeter–Koshland switch to an external stochastic gene regulatory network or fluctuating environment gives rise to a variety of behaviors. The switch may be used as a sensor for stress detection or as a node where different signaling pathways intersect. For the MHSRN, the Goldbeter–Koshland switch was used to model the coupling

of the environment to the biochemical network [20]. Our measurements of the MSHRN, show that mammalian cells have an extensive response (20 h) to a heat shock and that the response depends on the heat shock temperature and its duration [21]. Also, our measurements show a large cell-to-cell variation in the response. This study proves that a Goldbeter–Koshland is capable of an extended response and high variability in its response. We thus cannot rule out the presence of a Goldbeter–Koshland as a temperature sensor, although its importance as a temperature sensor still needs to be demonstrated through experimental measurements. Both the input signal and the output signal are stochastic and represent the fluctuating nature of the biological systems. The methods presented in this paper allow one to obtain the stochastic response in terms of the stochastic input variables. The numerical implementation of the method is much faster than a Monte Carlo simulation and offers an analytical perspective of the signal flow. These advantages come with the price of having only an approximate analytical model, although quite close to the simulated one. The numerical implementation of the method is fast and easy coded in Mathematica software. The method also encompasses the response plane representation. The stochastic nature of the switch is visually represented as a point in the response plane, which is based on the global measure of the mean and standard deviation of the switches' output signal. Each representative point contains information about the time evolution of the switch as it is guided by its fluctuating input. Unlike other studies, here we are interested not in the steady-state behavior of a gene regulatory network. On the contrary, we study its transitory evolution, which is time dependent. The point in the response plane incorporates this transitory evolution. The advantage of the response plane representation is its ability to represent a vast quantity of different switches under different conditions, opening a new way to classify such switches.

Thus, we presented the classes of fast and slow switches together with their stability with regards to changes in the quantity of the second enzyme. The fast switches may be used by organisms interested in a fast and strong time-dependent response to the first enzyme, E_1 , while being simultaneously robust to fluctuations in the second enzyme, E_2 . The fast circuits use a stall-based strategy to attain a fast response, waiting in a limbo state for the coupled feedback loop to gain strength and afterwards exploding with a rapid response.

The slow switches may be used to detect the presence or absence of the enzyme E_2 . Their response dramatically changes when E_2 jumps from zero to one. Such a switch may be used as a node of intersection of signaling pathways.

Another class of switches is those with a high standard deviation. These switches are located on the upper part of the response plane, where the high-variability switch is positioned. Using this switch, the organism may use the feedback loop G_f and the direct connection G to adapt.

Under a strong G_f input, the high-variability switch maximizes the possibility for divergent behaviors under mild G signals (high standard deviation in the response) while limiting the system's behavior to a narrow range of fast responses under severe G signals, guaranteeing the survival of a maximum number the switch's host organism's population. The switch promotes both adaptation and survival while maximizing the chances for success of either of these goals, depending upon the current input conditions.

It will be interesting to use the method described in this paper to classify the stochastic behavior of other coupled switches and sort their behavior into larger classes. It is possible that a functional criteria may emerge, a criteria less based on the biochemical structure of the switch but more oriented on its response.

Acknowledgments This work was supported by the National Science Foundation grant NSF-0827193. OL is grateful to Stephen Peiper, Zi-Xuan Wang, and Jean-Marc Navenot for their kind hospitality at the Thomas Jefferson Medical College during which part of this work was done.

References

1. Koshland, J.D.: The era of pathway quantification. *Science* **280**, 852–853 (1998)
2. Koshland, J.D.: The seven pillars of life. *Science* **295**, 2215–2216 (2002)
3. Kholodenko, B.N.: Cell-signaling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.* **7**, 165–167 (2006)
4. Alon, U.: An introduction to systems biology: design principles of biological circuits. Chapman & Hall, New York (2007)
5. Lipan, O., Wing, W.H.: Is the future biology Shakespearean or Newtonian? *Molecular Biosystems* **2**, 411–416 (2006)
6. Lipan, O.: Enlightening rhythms. *Science* **319**(5862), 417–418 (2008)
7. Kitano, H.: Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002)
8. Davidson, E.H.: The regulatory genome: gene regulatory networks in development and evolution. Academic Press, San Diego (2006)
9. Rives, A.W., Galitski, T.: Modular organization of cellular networks. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1128–1133 (2003)
10. Newman, M.E.J.: Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8577–8582 (2006)
11. Zhang, H., Ning, X.M., Ding, C., Zhang, X.S.: Determining modular organization of protein interaction networks by maximizing modularity density. *BMC Syst. Biol.* **2**(Suppl), S10 (2010)
12. Achimescu, S., Lipan, O.M.: Signal propagation in nonlinear stochastic gene regulatory networks. *IEE Syst. Biol.* **153**(3), 120–134 (2006)
13. Kaern, M., Elston, T.C., Blake, W.J., Collins, J.J.: Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* **6**, 451–464 (2005)
14. Hasty, J., Pradines, J., Dolnik, H., Collins, J.J.: Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2075–2080 (2000)
15. Thattai, M., van Oudenaarden, A.: Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8614–8619 (2001)
16. McAdams, H.H., Arkin, A.: Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 814–819 (1997)
17. Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O’Shea, E., Pilpel, Y., Barkai, N.: Noise in protein expression scales with natural protein abundance. *Nat. Genet.* **38**, 636–643 (2006)
18. Choi, P.J., Cai, L., Frieda, K., Xie, X.S.: A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* **322**, 442–446 (2008)
19. Arkin, A., Ross, J., McAdams, H.H.: Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* **149**, 1633–1648 (1998)
20. Rieger, T., Morimoto, R., Hatzimanikatis, V.: Mathematical modeling of the eukaryotic heat shock response: dynamics of the hsp70 promoter. *Biophys. J.* **88**, 1646–1658 (2005)
21. Lipan, O., Navenot, J.M., Wang, Z., Huang, L., Peiper, S.C.: Heat shock response in CHO mammalian cells is controlled by a nonlinear stochastic process. *PLoS Computational Biology* **3**(10), 1859–11870 (2007)
22. Morimoto, R.: Cells in stress: transcriptional activation of heat shock genes. *Science* **259**, 11409–11410 (1993)
23. Ellis, R.J., Hartl, F.U.: Principles of protein folding in the cellular environment. *Curr. Opin. Struct. Biol.* **9**, 102–110 (1999)
24. Hartl, F.U., Hayer-Hartl, M.: Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852–1858 (2002)
25. Tavaria, M., Gabriele, T., Kola, I., Anderson, R.L.: A hitchhiker’s guide to the human hsp70 family. *Cell Stress Chaperones* **1**, 23–28 (1996)
26. Morimoto, R.I., Tissieres, A., Georgopoulos, C.: The stress response, function of the proteins, and perspectives. In: Morimoto, R.I., Tissieres, A., Georgopoulos, C. (eds.) *Stress Proteins in Biology and Medicine*, vol. 19, 1st edn., pp. 1–36. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1990)

27. Srivastava, R., Peterson, M., Bentley, W.: Stochastic kinetic analysis of the Escherichia coli stress circuit using σ^{32} -targeted antisense. *Biotechnol. Bioeng.* **75**, 120–129 (2001)
28. El-Samad, H., Kurata, H., Doyle, J.C., Gross, C.A., Khammash, M.: Surviving heat shock: Control strategies for robustness and performance. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2736–2741 (2005)
29. Goldbeter, A., Koshland, D.E., Jr.: An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6840–6844 (1981)
30. Nakamura, N., Yamazawa, T., Okubo, Y., Iino, M.: Temporal switch and cell-to-cell variability in Ca^{2+} release activity in mammalian cells. *Mol. Syst. Biol.* **5**, 247:1–247:10 (2009)
31. Ciliberto, A., Capuani, F., Tyson, J.J.: Modeling networks of coupled enzymatic reactions using a total quasi-steady state approximation. *PLoS Computational Biology* **3**, 463–472 (2007)
32. Yang, Q., Pando, B.F., Dong, G., Golden, S.S., van Oudenaarden, A.: Circadian gating of the cell cycle revealed in single cyanobacterial cells. *Science* **327**, 1522–1526 (2010)
33. van Kampen, N.G.: *Stochastic Processes in Physics and Chemistry*. North Holland, Amsterdam (1992)
34. Samoilov, M., Plyasunov, S., Arkin, A.P.: Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2310–2315 (2005)
35. Blake, W.J., Balázsi, G., Kohanski, M.A., Isaacs, F.J., Murphy, K.F., Kuang, Y., Cantor, C.R., Walt, D.R., Collins, J.J.: Phenotypic consequences of promoter-mediated transcriptional noise. *Mol. Cell.* **24**, 853–865 (2006)
36. Li, C., Wong, W.H.: Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 31–36 (2001)
37. Nevozhay, D., Adams, R.M., Murphy, K.F., Josic, K., Balázsi, G.: Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5123–5128 (2009)