

Noise-induced Regime Shifts: A Quantitative Characterization

Sayantari Ghosh,^{*} Amit Kumar Pal,[†] and Indrani Bose[‡]

Department of Physics, Bose Institute, 93/1, A. P. C. Road, Kolkata-700009, India

Diverse complex dynamical systems are known to exhibit abrupt regime shifts at bifurcation points of the saddle-node type. The dynamics of most of these systems, however, have a stochastic component resulting in noise driven regime shifts even if the system is away from the bifurcation points. In this paper, we propose a new quantitative measure, namely, the propensity transition point as an indicator of stochastic regime shifts. The concepts and the methodology are illustrated for the one-variable May model, a well-known model in ecology and the genetic toggle, a two-variable model of a simple genetic circuit. The general applicability and usefulness of the method for the analysis of regime shifts is further demonstrated in the case of the mycobacterial switch to persistence for which experimental data are available.

PACS numbers: 05.10.Gg, 05.40.Ca, 02.30.Oz

I. INTRODUCTION

Dynamical systems are known to undergo sudden regime shifts at critical parameter values, termed the bifurcation points, with the different regimes defined by distinct sets of attractors of the dynamics [1]. Recently, a large number of studies have been devoted to the investigation of regime shifts in complex dynamical systems and processes ranging from ecosystems, climate change, population collapse and financial markets [2–5] to epileptic seizures and asthma attacks [6, 7]. Examples of sudden regime shifts include the collapse of vegetation under semi-arid conditions, the transition from a clear to a turbid lake, catastrophic shifts in fish or wildlife populations [2–5], transitions from one stable gene expression state to another in natural and synthetic genetic circuits [8–10] and the sudden deterioration of complex diseases [11]. The regime shifts, in most of the cases studied, are brought about by abrupt transitions from bistability to monostability via the saddle-node bifurcation [1, 2]. The state of a dynamical system at a specific time t is defined in terms of the magnitudes of one or more key variables at time t . In the steady state, the rates of change in the magnitudes of the variables are zero. In the case of monostability, there is only one single stable steady state. The region of bistability is distinguished by the coexistence of two stable steady states separated by an unstable steady state. The two stable steady states correspond to low and high values of the variable, designated as the L and H states respectively. The bistable region separates two regions of monostability with the L and H states being the respective stable steady states. At a saddle node bifurcation point, one of the stable steady states merges with the unstable steady state leading to a reduction in the number of physical solutions, from three to one, beyond the bifurcation point. Figure 1 illustrates the saddle-node bifurcation in the case of the May model [12], a well-known model in ecology. In the original model, a population of herbivores at a constant density is considered, the population being sustained by vegetation of biomass x . The rate equation describing the changes in the biomass amount is

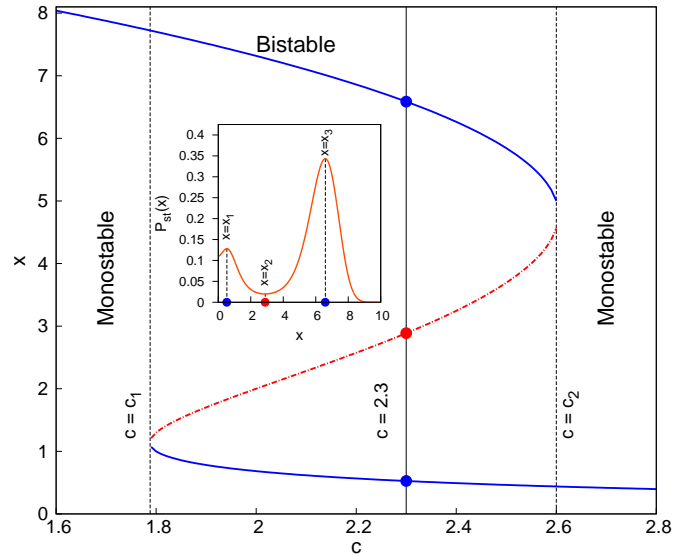


Figure 1. Biomass x versus the maximum grazing rate parameter c in the steady state of the May model the dynamics of which are described by Eq. (1). The solid lines represent the stable steady states and the dot-dashed line the branch of unstable steady states. The points $c = c_1$ and $c = c_2$ represent the lower and upper bifurcation points respectively. The parameter values used are $r = 1$, $K = 10$ and $x_0 = 1$. The inset shows the plot of the steady state probability distribution $P_{st}(x)$ versus x in the presence of additive noise of strength $d_1 = .25$. The maximum grazing rate parameter $c = 2.3$. The extrema points correspond to the stable steady states of the deterministic dynamics, with x_1 and x_3 representing the stable steady states which are separated by an unstable steady state x_2 .

given by

$$\frac{dx}{dt} = rx \left(1 - \frac{x}{K}\right) - \frac{cx^2}{x_0^2 + x^2} \quad (1)$$

The first term on the right hand side represents the logistic growth of the biomass at the rate r , K being the carrying capacity with the growth rate becoming zero when $x = K$. The second term corresponds to the rate of loss of the biomass due to its consumption by the herbivores. The parameter c is the maximum loss rate of the biomass or equivalently the maximum grazing rate. The loss rate

^{*} sayantari@jcbose.ac.in

[†] amit@jcbose.ac.in

[‡] indrani@jcbose.ac.in

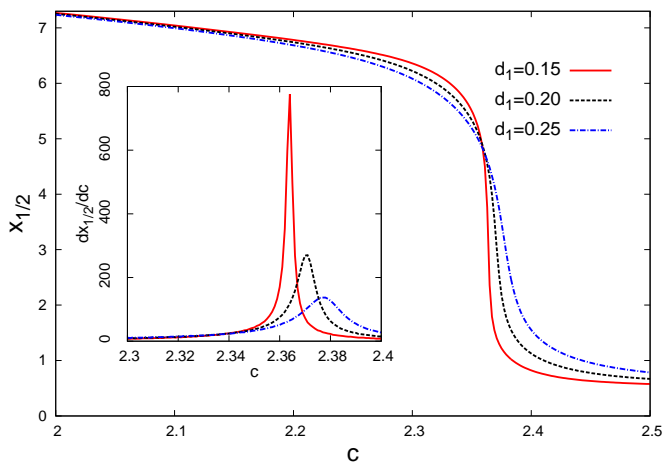


Figure 2. The equipartition point $x_{1/2}$ (Eq. (4)) versus the parameter c for three different strengths of the additive noise, $d_1 = .15$ (solid line), $d_1 = .20$ (dashed line) and $d_1 = .25$ (dot-dashed line). The other parameter values are $r = 1$, $K = 10$ and $x_0 = 1$. The inset shows the plots of $\frac{dx_{1/2}}{dc}$ versus c for the three noise strengths. The peak positions define the propensity transition point c^* .

saturates when x is much larger than a characteristic value x_0 . The May model has been applied to various other ecological problems, e.g., the exploitation of fish populations [13], dynamics of spruce budworms [13] and harvesting of macrophytes [14]. Figure 1 shows the steady state values of the biomass x ($\frac{dx}{dt} = 0$) versus the maximum grazing rate parameter c . The other parameter values are $r = 1$, $K = 10$ and $x_0 = 1$ in appropriate units. The solid lines represent the stable steady states separated by a branch (dot-dashed line) of unstable steady states. The parameters chosen fall in a region of parameter space exhibiting bistability. The lower (c_1) and upper (c_2) bifurcation points separate a region of bistability from two regions of monostability. At the upper bifurcation point c_2 , a saddle node bifurcation takes place and there is an abrupt regime shift from the H (high value of biomass) to the L (low value of biomass) state. The transition to the collapsed biomass state is not reversible with the reverse transition from the L to the H state occurring at the lower bifurcation point c_1 . This is the phenomenon of hysteresis, a characteristic feature of regime shifts of the type shown in Figure 1. At c_1 , a saddle node bifurcation occurs with the L state losing stability.

In the case of deterministic time evolution of a dynamical system, there are two ways in which regime shifts occur: (i) at the bifurcation points and (ii) on applying large perturbations in the region of bistability. The unstable steady state sets a threshold for switching transitions between the L and H states. Consider the system to be initially in the H state. A sufficiently large perturbation reduces the magnitude of the dynamical variable below the threshold, thereby bringing about a transition to the L state. Another way of bringing about regime shifts is through noise-induced transitions [2, 15–17]. The time evolution in complex dynamical systems exhibiting regime shifts is, in general, stochastic in nature [18] due to the probabilistic nature of the processes associated with the

dynamics. Stochasticity gives rise to fluctuations in the magnitude of the dynamical variable (say, the biomass) which, if sufficiently strong, cause transitions between the L and H states away from the bifurcation points. The occasional switch to an alternative regime in the presence of stochasticity is known as “flickering” [2] and is responsible for the appearance of a bimodal steady state probability distribution in the region of “bistability”. Noise-induced excursions from the stable steady states result in a broadening of the distribution around the steady states. The steady state distribution is achieved when the probability of the H to the L state transition is the same as that of the reverse transition. The bifurcation theory of stochastic dynamical systems is not as well-formulated as that in the case of deterministic dynamics. Kepler and Elston [19] have defined stochastic bifurcation in terms of the number of critical (singular) points in the steady state probability density function. Song et. al. [20] have investigated the stochastic bifurcation structure of cellular networks through specifying four quantities as a function of the bifurcation parameter: (i) the number of distinct subpopulations (or peaks in the steady state probability distributions), (ii) the locations of the peaks in terms of the measurable variable x , (iii) the variability in x for each subpopulation and (iv) the fractions of the whole population represented by the subpopulations. The total probability distribution is expressed as a mixture of component distributions, one for each subpopulation. The term subpopulation is used in a generalized sense, the number of subpopulations being equal to the number of distinct peaks in the probability distribution. For example, in the region of bistability (Figure 1), one has two subpopulations, L and H and the corresponding probability distributions arise due to noise-induced broadening of the steady state levels. Guttal and Jayaprakash [13] have investigated the impact of noise on bistable ecological systems. They have shown that the region of bistability is diminished in the presence of small amounts of noise whereas for noise beyond a critical strength, bistability vanishes altogether. In the latter case, the dynamical system can undergo abrupt regime shifts frequently. In general, however, noise can induce new types of dynamic behaviour with significant differences from the deterministic dynamics [18, 21–24]. To give a few examples, Samoilov et. al. [23] have shown that for the enzymatic futile cycle, a common biomolecular network motif, external noise can generate dynamic bistability through stochastic switchings. Zakharova et. al. [21] have demonstrated that noise can enhance the region of bistability, characterized by a bimodal probability distribution between a steady state and a limit cycle, beyond the bistable region obtained in the deterministic case. In the case of a model based on autoactivating gene expression, Karmakar and Bose [22] have shown that noise-induced bistability is possible in a region of parameter space for which there is no bistability in the deterministic case. To and Maheshri [24] have illustrated this concept in a recent experiment. In this paper, we propose a new framework for the quantitative characterization of noise-induced regime shifts in bistable systems. The methodology has a clear physical interpretation and is easy to implement. We focus on two models to illustrate the methodology, the May model and the genetic toggle [25] describing a simple genetic circuit in which the

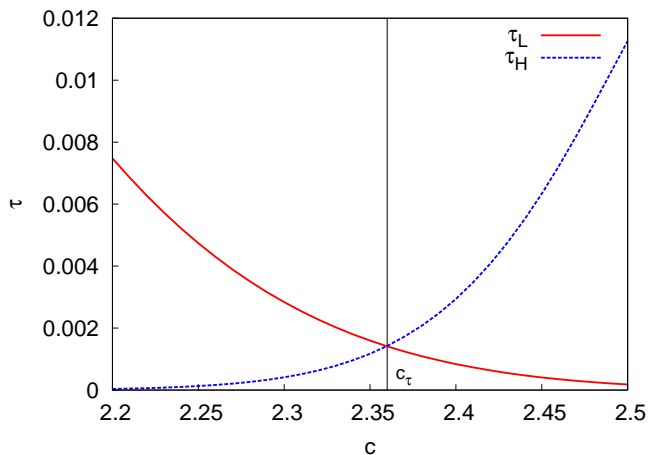


Figure 3. The mean first passage times τ_L (solid line) and τ_H (dashed line) versus the parameter c . The other parameter values are $r = 1$, $K = 10$, $x_0 = 1$ and $d_1 = .15$. The intersection point c_τ ($\tau_L = \tau_H$) of the two curves is approximately equal to c^* , the propensity transition point ($c_\tau = 2.360$, $c^* = 2.364$).

protein products of two genes repress each other's synthesis. The May model is a one-variable model whereas the genetic toggle is a two-variable one, thus demonstrating the general applicability of the methodology proposed in the paper. In Section II, we introduce the concept of the propensity transition point (PTP) which may be considered as an indicator of stochastic regime shifts. As an example, we consider the one-variable May model for which the steady state probability distribution is obtained by solving the appropriate Fokker-Planck equation. In Section III, the steady-state probability distributions are obtained for both the May model and the genetic toggle using the Gillespie simulation algorithm [26]. We further analyze available experimental data to determine the PTP in the case of stochastic regime shifts from the normal to the persister subpopulation in mycobacteria subjected to stress in the form of nutrient depletion [27, 28].

II. EQUIPARTITION AND PROPENSITY TRANSITION POINTS

In the case of stochastic dynamics, the probability distribution $P(x)$ of a random variable x can be determined using a number of analytical and numerical methods [26, 29, 30]. The cumulative distribution function $P_C(x)$ is defined to be

$$P_C(x') = \int_0^{x'} P(x) dx \quad (2)$$

where $0 \leq x' \leq \infty$. Let $x_{1/2}$ be the median of the distribution $P(x)$ marking the equipartition point of the distribution. One can write

$$\int_0^{x_{1/2}} P(x) dx = \int_{x_{1/2}}^\infty P(x) dx = \frac{1}{2} \quad (3)$$

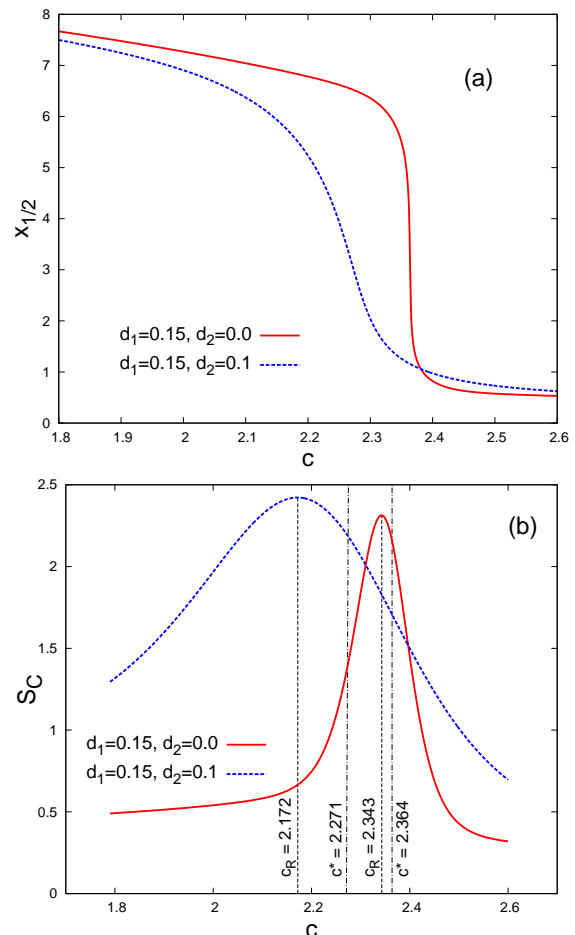


Figure 4. (a) The equipartition point $x_{1/2}$ (Eq. (4)) versus c for different strengths of the multiplicative noise: $d_2 = 0$ (solid line), $d_2 = .1$ (dashed line) with $d_1 = .15$ in each case. The other parameter values are $r = 1$, $K = 10$ and $x_0 = 1$. (b) The cumulative residual entropy, S_C (Eq. (15)) versus c for the two different values of the multiplicative noise strength as in (a). The PTP c^* is included for comparison.

or,

$$P_C(x_{1/2}) = 1 - P_C(x_{1/2}) = \frac{1}{2} \quad (4)$$

We consider the steady-state probability distribution (SSPD), which changes as a bifurcation parameter, e.g., the parameter c in the May model, is changed. One can compute the variation of $x_{1/2}$ versus the bifurcation parameter using Eq. (4). We illustrate this in the case of May model. The SSPD, $P_{st}(x)$, is obtained by solving the Fokker-Planck equation in the steady-state. In the case of stochastic dynamics, the one-variable Langevin equation including both additive and multiplicative noise terms is given by

$$\frac{dx}{dt} = f(x) + g(x)\varepsilon(t) + \Gamma(t) \quad (5)$$

where $\Gamma(t)$ (additive noise) and $\varepsilon(t)$ (multiplicative noise) are Gaussian white noises with zero mean and correla-

tions:

$$\begin{aligned}\langle \Gamma(t)\Gamma(t') \rangle &= 2d_1\delta(t-t') \\ \langle \varepsilon(t)\varepsilon(t') \rangle &= 2d_2\delta(t-t')\end{aligned}\quad (6)$$

In Eq. (6), d_1 and d_2 denote the strengths of the additive and multiplicative noises respectively, we further assume that there is no correlation between the two types of noise. The first term in the right hand side of Eq. (5) describes the deterministic dynamics, e.g., for the May model $f(x)$ is given by the right hand side expression in Eq. (1). The Fokker-Planck equation, a rate equation for the probability distribution $P(x, t)$, is obtained from the Langevin equation as [29, 30]:

$$\frac{\partial P(x, t)}{\partial t} = -\frac{\partial}{\partial x}[A(x)P(x, t)] + \frac{\partial^2}{\partial x^2}[B(x)P(x, t)] \quad (7)$$

where

$$A(x) = f(x) + d_2g(x)g'(x) \quad (8)$$

and

$$B(x) = d_1 + d_2[g(x)]^2 \quad (9)$$

The SSPD is given by [29, 31, 32]

$$\begin{aligned}P_{st}(x) &= \frac{N}{B(x)} \exp\left[\int^x \frac{A(x)}{B(x)} dx\right] \\ &= \frac{N}{\{d_2[g(x)]^2 + d_1\}^{\frac{1}{2}}} \\ &\quad \times \exp\left[\int^x \frac{f(x')dx'}{d_2[g(x')]^2 + d_1}\right]\end{aligned}\quad (10)$$

where N is the normalization constant. Equation (10) can be rewritten in the form

$$P_{st}(x) = Ne^{-\phi_F(x)} \quad (11)$$

where

$$\begin{aligned}\phi_F(x) &= \frac{1}{2} \ln [d_2[g(x)]^2 + d_1] \\ &\quad - \int^x \frac{f(y)dy}{d_2[g(y)]^2 + d_1}\end{aligned}\quad (12)$$

defines the ‘‘stochastic potential’’ of the dynamics. The inset of Figure 1 shows the SSPD, $P_{st}(x)$, versus x for the May model in the presence of only additive noise ($g(x) = 0$, $d_2 = 0$) for the parameter values $r = 1$, $K = 10$, $x_0 = 1$, $c = 2.3$ and $d_1 = 0.25$. In this case, $\phi_F(x)$ reduces to

$$\phi_F(x) = \frac{1}{2} \ln d_1 - \frac{1}{d_1} \int^x f(y)dy \quad (13)$$

The deterministic potential $\phi_D(x)$ is given by

$$\phi_D = - \int^x f(y)dy \quad (14)$$

Comparing Eqs. (13) and (14) one finds that the extrema points of $\phi_F(x)$ and $\phi_D(x)$ are identical with the minima (maximum) corresponding to stable (unstable) steady states. In the SSPD $P_{st}(x)$, the stable steady states corre-

spond to the maxima points x_1 and x_3 whereas the minimum point x_2 represents the unstable steady state. We first consider the case when only additive noise is present, i.e., $d_1 \neq 0$ and $d_2 = 0$ in Eq. (6). Figure 2 shows the variation of $x_{1/2}$ with c for the parameter values $r = 1$, $K = 10$, $x_0 = 1$ and different values of the noise strength $d_1 = .15$ (solid line), $d_1 = .2$ (dashed line) and $d_1 = .25$ (dot-dashed line). The deterministic bifurcation points have values $c_1 = 1.788$ and $c_2 = 2.604$. As c is increased from a low value, one notices a transition of $x_{1/2}$ from a high to a low value. The transition is sharp for low values of the additive noise and becomes smeared as the noise strength d_1 is increased. The point $x = x_{1/2}$ provides a quantitative measure of which subpopulation is the preferred (dominant) subpopulation in the steady state. The L subpopulation corresponds to x values in the range $0 \leq x \leq x_2$, the rest of the SSPD is associated with the H subpopulation. For low values of c , $x_{1/2}$ is $> x_2$ and the H subpopulation is more dominant. As c increases, $x_{1/2}$ shifts towards x_2 and the L subpopulation becomes the more preferred one when $x_{1/2}$ is $< x_2$. One can define a propensity transition point (PTP) c^* at which the preferred subpopulation changes its character from H to L . The inset of Figure 2 shows the variation of $\frac{dx_{1/2}}{dc}$ versus c for $d_1 = .15$, $d_1 = .2$ and $d_1 = .25$, with the respective peak positions being $c^* = 2.364$, $c^* = 2.371$ and $c^* = 2.378$. For larger magnitudes of noise below some critical strength, the variation is more smeared but one can still identify a peak position c^* at which $\frac{dx_{1/2}}{dc}$ attains its maximum value. In the presence of large additive noise, flickering becomes prominent and bistability is destroyed. When only additive noise is present and at $c = c^*$, $x_{1/2}$ is equal to x_2 , the minimum of the SSPD. The justification of this statement comes from the definition of $x_{1/2}$. As already discussed, the point $x_{1/2}$ belongs to the preferred subpopulation. At the propensity transition point $c = c^*$, both the subpopulations are equally preferred, i.e., $x_{1/2} = x_2$, the boundary point separating the L and H subpopulations. As already mentioned, in the presence of only additive noise, the point x_2 coincides with the unstable steady state solution in the deterministic case. The cumulative residual entropy defined as [33]

$$S_C = - \int_x P_c(x) \ln P_c(x) dx \quad (15)$$

with $P_c(x)$ as given in Eq. (2), attains its maximum value at $c = c_R$ which is close to c^* . For $d_1 = .15$, $c_R = 2.343$. Let P_L and P_H be the probabilities of belonging to subpopulations L and H respectively. P_L and P_H can be written as $P_L(t) = \int_0^{x_2} P(x, t) dx$ and $P_H(t) = \int_{x_2}^\infty P(x, t) dx$. The time evolutions of $P_L(t)$ and $P_H(t)$ are given by

$$\begin{aligned}\frac{dP_L(t)}{dt} &= -k_L P_L(t) + k_H P_H(t) \\ \frac{dP_H(t)}{dt} &= -k_H P_H(t) + k_L P_L(t)\end{aligned}\quad (16)$$

where k_L and k_H are the stochastic transition rates for the $L \rightarrow H$ and $H \rightarrow L$ transitions respectively. In the steady state, $\frac{dP_L(t)}{dt} = 0$, $\frac{dP_H(t)}{dt} = 0$ and one gets

$$\frac{P_{HS}}{P_{LS}} = \frac{k_L}{k_H} \quad (17)$$

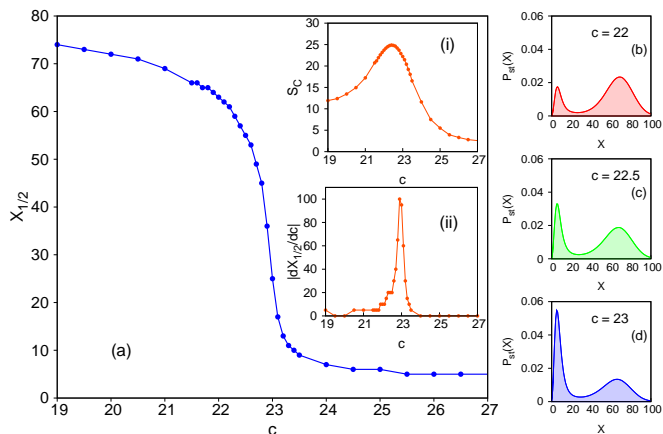


Figure 5. (a) Variation of the equipartition point $X_{1/2}$ with the bifurcation parameter c . The graph shows a sharp fall in the values of $X_{1/2}$ in a small range of c values. The insets (i) and (ii) show the variations of the cumulative residual entropy, S_C , and the first derivative of $X_{1/2}$ w.r.t. the bifurcation parameter c , $\frac{dX_{1/2}}{dc}$, respectively, as functions of c . $\frac{dX_{1/2}}{dc}$ attains a sharp maximum at $c^* = 22.9$, the PTP. The cumulative residual entropy also becomes maximum near c^* . The SSPDs for three values of c spanning the PTP are shown in (b),(c) and (d).

where ‘S’ in the suffix indicates the steady state probabilities. In the case of weak noise, one can write [34]

$$\begin{aligned} \tau_L &\simeq \frac{1}{k_L} \\ \tau_H &\simeq \frac{1}{k_H} \end{aligned} \quad (18)$$

where τ_L and τ_H denotes the mean first passage times (MFPTs) for exits from the domains of the L and H subpopulations respectively. The exits are brought about by noise and a higher value of the MFPT indicates a greater stability of the steady state from which the exit occurs. At the PTP c^* , $P_{HS} = P_{LS}$ which leads to the equality $\tau_L = \tau_H$. For weak noise, k_H and k_L are given by the approximate expressions [34],

$$\begin{aligned} k_H &= \frac{d_1}{2\pi} \sqrt{|\phi_F''(x_2)| |\phi_F''(x_3)|} e^{\phi_F(x_3) - \phi_F(x_2)} \\ k_L &= \frac{d_1}{2\pi} \sqrt{|\phi_F''(x_2)| |\phi_F''(x_1)|} e^{\phi_F(x_1) - \phi_F(x_2)} \end{aligned} \quad (19)$$

where $\phi_F''(x_2)$ denotes the double derivative of ϕ_F with respect to x . Figure 3 shows the plots of the MFPTs τ_L and τ_H (determined using Eqs. (13), (18) and (19)) versus the bifurcation parameter c . The other parameter values are $r = 1$, $K = 10$, $x_0 = 1$ and $d_1 = .15$. The intersection point of the two curves, c_τ , at which $\tau_L = \tau_H$, is approximately equal to c^* ($c_\tau = 2.360$, $c^* = 2.364$). The numerical result supports the analytic argument, using Eqs. (16) and (18), that $\tau_L = \tau_H$ at the PTP $c = c^*$.

We next consider the general case when additive and multiplicative noise terms are present in the Langevin equation, Eq. (5). We assume that the multiplicative noise is associated with the rate constant r in Eq. (1) with $g(x)$ (Eq. (5)) given by

$$g(x) = x \left(1 - \frac{x}{K}\right) \quad (20)$$

In this case, the extrema points x_1 , x_2 and x_3 of the SSPD no longer coincide with the deterministic steady state values. As before, the equipartition point $x_{1/2}$ can be determined as a function of c . Figure 4 (a) shows the variation of $x_{1/2}$ versus c for the parameter values $r = 1$, $K = 10$ and $x_0 = 1$. Figure 4 (b) shows the cumulative residual entropy (Eq. (15)) versus the bifurcation parameter c . The peak position c_R is close to the PTP c^* in the presence of only additive noise. Entropy-like quantities are measure of uncertainty about the values of random variables. In the case of a bimodal probability distribution, the uncertainty is maximal when the component subpopulations are equally preferred. In the presence of only additive noise and at $c = c^*$, the equipartition point $x_{1/2} = x_2$, the boundary point separating the ranges of x values corresponding to the two subpopulations. In this case, both the subpopulations are equally preferred leading to a maximal value of the cumulative residual entropy (maximal uncertainty) close to c^* . When both additive and multiplicative noise terms appear in the Langevin equation (Eq. (5)), the difference between c_R and c^* increases. The differences between the values of x_2 and c^* , the PTP, and c_R and c^* provide a measure of the multiplicative noise present in the system.

III. GILLESPIE SIMULATION

The stochastic time evolution of a dynamical system can be studied through computer simulation based on the Gillespie algorithm [26]. We now discuss briefly the salient features of the Gillespie algorithm, more detailed discussions can be obtained from Refs. [26, 35–37]. Let M be the number of reactions controlling the time evolution of molecular numbers in a dynamical system. Each reaction is characterized by a reaction propensity a_μ ($\mu = 1, 2, \dots, M$) with $a_\mu(t)dt$ defining the probability that the reaction μ occurs in volume V in the time interval $(t, t+dt)$ given the state of the system at time t . The propensity $a_\mu(t)$ is a product of two parts, the reaction rate c_μ for reaction μ and the number of possible reactions μ in volume V . One also defines a reaction stoichiometry matrix \mathbf{S} which is an $N \times M$ matrix where N is the number of reactants (different species of molecules) in the dynamical system. The element S_{ij} ($i = 1, \dots, N$, $j = 1, \dots, M$) denotes the change in the number of reactant i molecules, from X_i to $X_i + S_{ij}$, when the j th reaction takes place. The deterministic rate equation can be written in a compact form

$$\frac{d\mathbf{X}}{dt} = \mathbf{S} \cdot \nu \quad (21)$$

The Gillespie algorithm monitors time evolution by obtaining information firstly on the time of occurrence of the next reaction given the state of the system in terms of molecular numbers at time t and secondly the reaction type. Let us define a quantity a_0 as $a_0 = \sum_{\mu=1}^M a_\mu$. With knowledge of the state of the system at time t , the probability that the next reaction occurs in the time interval

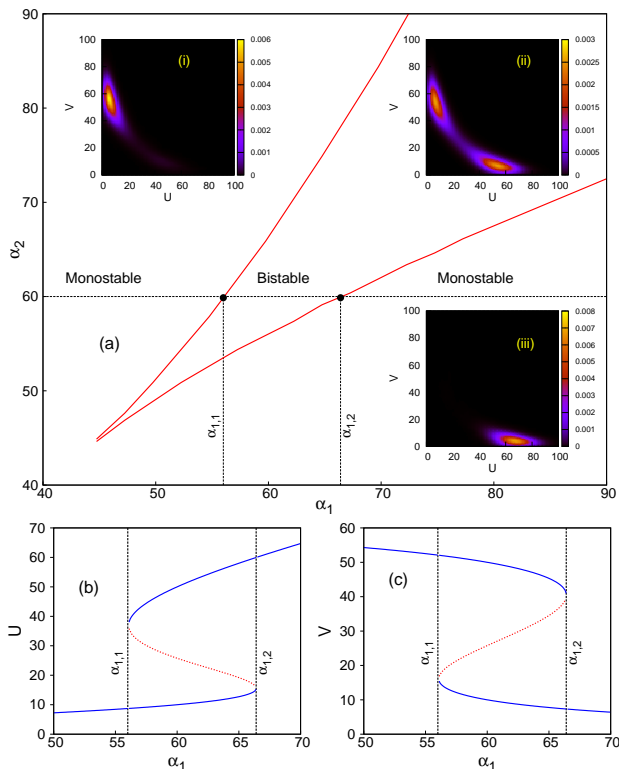


Figure 6. (a) The α_1 - α_2 phase diagram for the genetic toggle. The other parameter values are $K = 500$, $\beta_1 = \beta_2 = 2$ and $\gamma = 1$. The steady-state probability distributions $P_{st}(U, V)$ for three different values of α_1 are shown in the insets (i) $\alpha_1 = 50$, (ii) $\alpha_1 = 60$ and (iii) $\alpha_1 = 70$ with α_2 kept fixed at 60. The variations of U and V versus the parameter α_1 for $\alpha_2 = 60$ are shown in (b) and (c) respectively. The solid lines in each plot represent the stable steady states and the dotted line the branch of unstable steady states.

$t + \tau$ and $t + \tau + d\tau$ and is of type μ is $P_\mu(\tau)d\tau$ where

$$\begin{aligned} P_\mu(\tau) &= a_\mu \exp(-a_0\tau) \\ &= \frac{a_\mu}{a_0} a_0 \exp(-a_0\tau) \\ &= P_2(\mu)P_1(\tau) \end{aligned} \quad (22)$$

where $P_1(\tau) = a_0 \exp(-a_0\tau)$ and $P_2(\mu) = \frac{a_\mu}{a_0}$. The Gillespie algorithm generates two random numbers r_1 and r_2 using a standard uniform random number generator. The time τ is then given by [26]

$$\tau = \frac{1}{a_0} \ln\left(\frac{1}{r_1}\right) \quad (23)$$

The reaction type μ is taken to be the integer for which the condition

$$\sum_{\nu=1}^{\mu-1} a_\nu < r_2 a_0 \leq \sum_{\nu=1}^{\mu} a_\nu \quad (24)$$

is satisfied. The random number τ obtained from Eq. (23) is generated according to the probability distribution $P_1(\tau)$ whereas Eq. (24) generates the random integer μ according to the probability distribution $P_2(\mu)$. Once the pair (τ, μ) is determined, the time t is advanced by τ , i.e., $t \rightarrow t + \tau$ and the molecular number X_i 's ($i = 1, \dots, N$)

are adjusted according to the reaction μ . The SSPD $P(X_i)$ of a molecular type can be computed by combining the data over a sufficiently large interval of time after the steady state conditions are achieved. A rigorous approach to the study of the stochastic time evolution of a system of N chemical species participating in M chemical reactions is based on the chemical master equation [26, 35]. The chemical master equation is a time-evolution equation for the probability $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ that the system is in the state $\mathbf{x} = (x_1, \dots, x_N)$ where x_i , $i = 1, \dots, N$, is the number of molecules of the i th species at time t . The chemical master equation involves the reaction propensity a_μ ($\mu = 1, 2, \dots, M$) defined earlier. An exact, analytic solution of the chemical master equation is, however, possible only in a few cases. The Gillespie algorithm simulates the temporal trajectories of $\mathbf{x}(t)$ starting from a given initial state. The probability distribution $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ is computed from the knowledge of an ensemble of sample trajectories, with the computed value approaching the exact solution in the limit of large ensemble size. The 'exact' Gillespie algorithm is, however, computationally prohibitive when the number of chemical species, N , and the number of reactions, M , become large. Approximate versions of the Gillespie algorithm have been developed [35, 38] to reduce the computational complexity. The Langevin and Fokker-Planck equations provide approximate approaches to the solution of the master equation and are less rigorous than the simulation approach based on the Gillespie algorithm. Gillespie has shown [38, 39] how to derive the chemical Langevin equation from the chemical master equation. The "white-noise" form of the chemical Langevin equation is

$$\begin{aligned} \frac{dx_i(t)}{dt} &= \sum_{j=1}^M S_{ij} a_j(\mathbf{x}(t)) \\ &+ \sum_{j=1}^M S_{ij} \sqrt{a_j(\mathbf{x}(t))} N_j(0, 1), \quad (i = 1, \dots, N) \end{aligned} \quad (25)$$

where \mathbf{S} is the reaction stoichiometry matrix defined earlier and $N_j(0, 1)$ represents a unit normal random variable with mean zero and variance one. While the chemical Langevin equation provides an explicit structure to the noise terms in terms of the reaction propensity, the majority of studies utilizing the Langevin formalism start with equations of the type shown in (5). The choice is dictated by the simplicity of the calculational scheme with specific focus on the separate effects of additive and multiplicative noise. The Langevin equation can be demonstrated to be mathematically equivalent to the Fokker-Planck equation [40]. The latter is further obtained through a Taylor series expansion of the master equation (Kramers-Moyal expansion) and retaining terms only upto the second derivative term.

We next report the results of the Gillespie simulation in the cases of the May model and the genetic toggle. The simulation tracks stochastic time evolution without distinguishing between additive and multiplicative types of noise. The rigorous Gillespie simulation approach is valid only for unimolecular and bimolecular reactions [38].

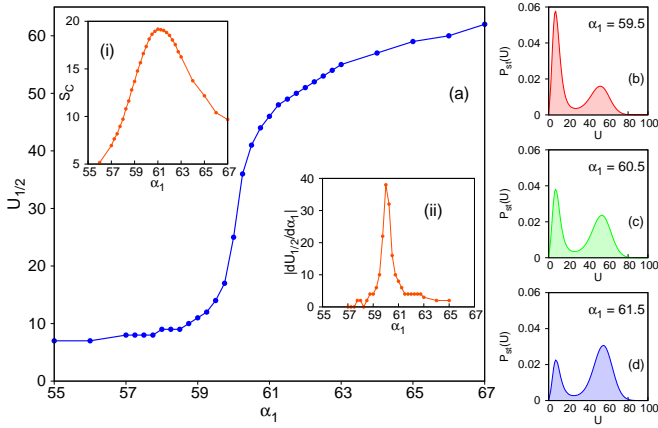


Figure 7. (a) Variation of the equipartition point $U_{1/2}$ obtained from the SSPD of U with the bifurcation parameter α_1 . The other parameter values are $\alpha_2 = 60$, $K = 500$, $\beta_1 = \beta_2 = 2$ and $\gamma = 1$. The graph shows a sharp fall in the values of $U_{1/2}$ in a small region of α_1 values with decreasing α_1 . The insets (i) and (ii) show the variations of the cumulative residual entropy, S_c , and the first derivative of $U_{1/2}$ w.r.t. the bifurcation parameter α_1 , $\frac{dU_{1/2}}{d\alpha_1}$, respectively, as functions of α_1 . $\frac{dU_{1/2}}{d\alpha_1}$ attains a sharp maximum at $\alpha_1^* = 60$, the PTP. The cumulative residual entropy also becomes maximum near α_1^* . The SSPDs for three values of α_1 spanning the PTP are shown in (b),(c) and (d).

Since the simulation approach is computationally exhaustive, approximation versions of the algorithm are often in use. One such approximation involves the study of a reduced model for which each reaction is a composite one representing the combined effects of multiple elementary reactions [36]. The reaction propensity a_μ ($\mu = 1, 2, \dots, M$) and the $N \times M$ stoichiometry matrix \mathbf{S} are defined for a system of M composite reactions in which N chemical species participate. The approximate approach, though less rigorous, reproduces qualitatively the results obtained using the rigorous simulation approach [36]. For the May model, we have two composite reactions each of which effectively represents a number of more elementary reactions. Rewriting the biomass molecular number as X , the rate equation (1) can be recast as

$$\frac{dx}{dt} = F_1(X) - F_2(X) \quad (26)$$

with

$$\begin{aligned} F_1(X) &= rX \\ F_2(X) &= \frac{rX^2}{K} + \frac{cX^2}{X_0^2 + X^2} \end{aligned} \quad (27)$$

There are thus two composite reactions $X \rightarrow X + 1$ and $X \rightarrow X - 1$ with the reaction propensities given by $F_1(X)$ and $F_2(X)$ respectively. The stoichiometry matrix \mathbf{S} is of size (1×2) with $S_{11} = +1$ and $S_{12} = -1$. Figure 5(a) shows the variation of the equipartition point $X_{1/2}$ with the bifurcation parameter c . The other parameter values are $r = 1$, $K = 100$ and $X_0 = 10$. One notes the sharp fall in the values of $X_{1/2}$ in a small region of c values. The required SSPD's are computed by collecting together the data in the Gillespie simulation over a large period of time. The insets (i) and (ii) of Figure 5(a) show

the variations of the cumulative residual entropy, S_c , and the first derivative $\frac{dX_{1/2}}{dc}$ respectively as functions of the bifurcation parameter c . The first derivative attains a sharp maximum at $c^* = 22.9$, the PTP. The cumulative residual entropy, S_c , has its maximum value close to c^* . Figures 5(b), (c) and (d) show the SSPDs for three values of the bifurcation parameter spanning the PTP.

We next describe the results for the genetic toggle, a two-variable model. The toggle circuit consists of two genes synthesizing proteins with molecular numbers U and V . The proteins repress each others' synthesis. The dynamics of the model are described by the set of equations [25]

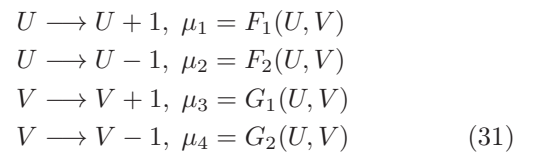
$$\begin{aligned} \frac{dU}{dt} &= \frac{\alpha_1}{1 + \frac{V^{\beta_1}}{K}} - \gamma U \\ &= F_1(U, V) - F_2(U, V) \end{aligned} \quad (28)$$

$$\begin{aligned} \frac{dV}{dt} &= \frac{\alpha_2}{1 + \frac{U^{\beta_2}}{K}} - \gamma V \\ &= G_1(U, V) - G_2(U, V) \end{aligned} \quad (29)$$

The parameters α_1, α_2 denote the effective rates of synthesis and γ the degradation rate constant, assumed to be the same for the two proteins. The parameters β_1 and β_2 are the indices indicating cooperativity in repression and K is related to the binding constant of proteins. The functions $F_i(U, V)$ and $G_i(U, V)$ ($i = 1, 2$) are:

$$\begin{aligned} F_1(U, V) &= \frac{\alpha_1}{1 + \frac{V^{\beta_1}}{K}}, \quad F_2(U, V) = \gamma U \\ G_1(U, V) &= \frac{\alpha_2}{1 + \frac{U^{\beta_2}}{K}}, \quad G_2(U, V) = \gamma V \end{aligned} \quad (30)$$

There are now four composite reactions ($M = 4$) with reaction propensities μ_i , $i = 1, \dots, 4$. The stoichiometry matrix \mathbf{S} is of size 2×4 . The composite reaction scheme is:



Also,

$$\mathbf{S} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \quad (32)$$

Figure 6 shows the α_1 - α_2 phase diagram of the genetic toggle exhibiting two regions of monostability and one region of bistability. The other parameter values are: $K = 500$, $\beta_1 = \beta_2 = 2$ and $\gamma = 1$. The SSPDs $P_{st}(U, V)$ for three different values of α_1 and with $\alpha_2 = 60$ are shown in the insets (i) $\alpha_1 = 50$, (ii) $\alpha_1 = 60$ and (iii) $\alpha_1 = 70$. These have been computed using the Gillespie simulation algorithm. The variations of the steady state values of U and V versus the parameter α_1 ($\alpha_2 = 60$) are shown in Figures 6(b) and 6(c) respectively. The branches of stable steady states (solid lines) are separated by the branches of unstable steady states (dotted lines). Figure 7 exhibits the results obtained from the Gillespie simulation. Figure

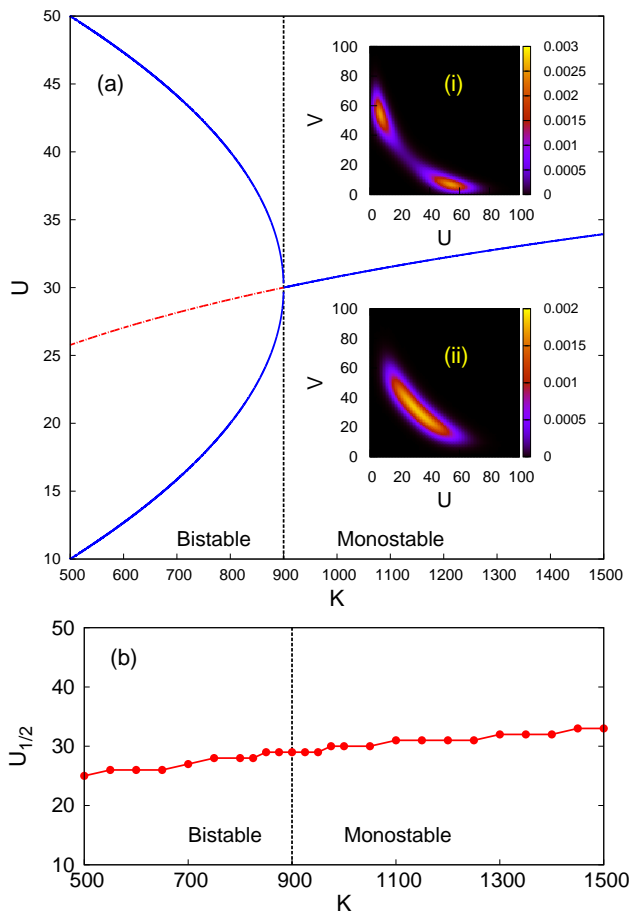


Figure 8. (a) The genetic toggle model exhibits pitchfork bifurcation with respect to the parameter K when $\alpha_1 = \alpha_2$. The parameter values are $\alpha_1 = \alpha_2 = 60$, $\beta_1 = \beta_2 = 2$ and $\gamma = 1$. The pitchfork bifurcation point is $K = 900$. The model shows bistability for $K < 900$ and monostability for $K > 900$. The insets show SSPDs for the cases (i) $K = 500$ and (ii) $K = 1200$. (b) The variation of $U_{1/2}$ against the bifurcation parameter K . The plot does not provide a signature of stochastic bifurcation.

7(a) shows the variation of the equipartition point $U_{1/2}$, obtained from the SSPD of U , versus the bifurcation parameter α_1 . The other parameter values are $\alpha_2 = 60$, $K = 500$, $\beta_1 = \beta_2 = 2$ and $\gamma = 1$. The SSPDs are obtained by collecting data over 10^8 time points in the steady state. One notes a sharp fall in the values of $U_{1/2}$ in a small region of α_1 values when α_1 is decreased from high to low values. The insets (i) and (ii) exhibit the variations of the cumulative residual entropy, S_c , and $\frac{dU_{1/2}}{d\alpha_1}$ respectively versus the bifurcation parameter α_1 . $\frac{dU_{1/2}}{d\alpha_1}$ attains a sharp maximum at $\alpha_1^* = 60$, the PTP. The cumulative residual entropy also attains its maximum values for α_1 close to α_1^* . The SSPDs for three values of α_1 , spanning the PTP, are shown in Figures 7(b), (c) and (d).

The sharp fall of the equipartition point (median) $U_{1/2}$ in the vicinity of the PTP is a unique characteristic of the saddle-node bifurcation. The pitchfork bifurcation is another type of bifurcation involving a transition between bistability and monostability. The genetic toggle exhibits pitchfork bifurcation as a function of the parameter K when the effective synthesis rate constants α_1 and

α_2 are kept equal. Figure 8(a) demonstrates the bifurcation with the steady state value of U plotted against K . The solid lines represent stable steady state and the dot-dashed line the unstable steady state. The parameter values are $\alpha_1 = \alpha_2 = 60$, $\beta_1, \beta_2 = 2$ and $\gamma = 1$. The pitchfork bifurcation point $K = 900$ separates a region of bistability ($K < 900$) from a region of monostability ($K > 900$). The insets of Figure 8(a) show the SSPDs determined using the Gillespie algorithm for (i) $K = 500$ and (ii) $K = 1200$. Figure 8(b) shows the variation of $U_{1/2}$ against K . In this case, there is no distinguishing feature in the plot providing a signature of stochastic bifurcation. The numerical evidence of a propensity transition thus enables one to distinguish between the pitchfork and saddle node bifurcations as the basis for noise-induced regime shifts. We note here that since the propensity transition point as an indicator of stochastic regime shifts is not an appropriate indicator of the pitchfork bifurcation, it cannot also distinguish between the cases of pitchfork bifurcation and the absence of bifurcation. In the latter case, since only one population is present, the variation of the equipartition point as a function of the bifurcation parameter is expected to be featureless, a trait shared with the pitchfork bifurcation.

There is recent experimental evidence that microorganisms take recourse to noise-induced regime shifts as a strategy for survival under stress [27, 28, 41, 42]. In the case of mycobacteria, stresses like nutrient depletion activate the stringent response pathway involving the two component system MprAB, the sigma factor SigE and the stringent response regulator Rel. The enzyme polyphosphate kinase 1 (PPK1) regulates the pathway by catalysing the synthesis of polyphosphate required for the activation of MprB. Multiple positive feedback loops and molecular sequestration of the sigma factor [27, 43] create the potential for bistability, i.e., alternative attractor states. Stochastic gene expression gives rise to fluctuations in the levels of key regulatory proteins like Rel which, if sufficiently strong, can bring about regime shifts from one attractor state (low Rel level) to another (high Rel level). The total mycobacterial population divides into two distinct subpopulations L and H corresponding to low and high Rel levels respectively. The stringent response pathway, activated by Rel proteins, is initiated in the H subpopulation. It is this subpopulation which adapts to stress and is designated as the persister subpopulation. The other subpopulation, termed the normal subpopulation, is unable to survive under stress. When exposed to antibiotic drugs, the normal subpopulation is killed while the persisters are able to survive. The persisters have slow growth rates and low rates of metabolic activity. They stay dormant for long periods of time waiting for the opportune moment to revive and restart the mycobacterial infection. Single cell analysis via flow cytometry provides experimental evidence of regime shifts from the L to the H subpopulation in a mycobacterial population subjected to nutrient depletion till the stationary phase is reached [27, 28]. The experiments were carried out on *M. smegmatis* which shares similar genetic circuitry with *M. tuberculosis*, the pathogen responsible for tubercular infection. The existence of persisters acts against the total eradication of the tubercular infection. The major goal of an effective drug treatment is to eliminate the

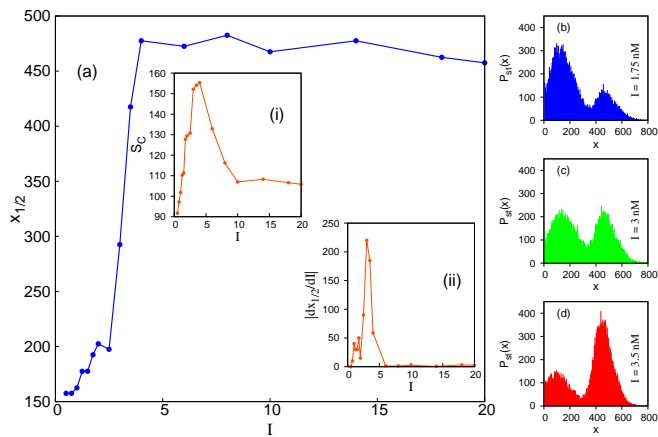


Figure 9. (a) Variation of equipartition point $x_{1/2}$ versus inducer concentration I (in nM). The insets show the variation of (i) the cumulative residual entropy, S_C , and (ii) $\left|\frac{dx_{1/2}}{dI}\right|$ versus the inducer concentration I . (b)–(d) Stationary phase probability distributions of Rel levels for $I = 1.75$ nM, 3 nM and 3.5 nM respectively.

population of persisters. In this context, it is of interest to identify the PTP above which the subpopulation of persisters becomes the dominant subpopulation.

The stochastic bifurcation analysis is carried out on available experimental data [27, 44] with the *ppk1* gene designed to be tetracycline-inducible. Figure 9(b), (c) and (d) shows the distribution of Rel levels in the stationary phase of a population of *M.smegmatis* cells for three different inducer concentrations I (in nM). Figure 9(a) shows the variation of the equipartition point $x_{1/2}$ as a function of I (in nM), the inducer concentration. The insets (i) and (ii) show the variations of $\left|\frac{dx_{1/2}}{dI}\right|$ and the cumulative residual entropy, S_C , respectively versus the inducer concentration I . The propensity transition to the persister subpopulation is found to occur at low values of I .

IV. CONCLUDING REMARKS

The quantitative measure of regime shifts, namely, the PTP, proposed in this Letter is of broad applicability and can be computed both when the model dynamics are known or when only the time series data are available. The physical interpretation of the PTP in terms of a tipping of the balance towards one regime or the other helps in identifying the parameter regime to be avoided in order to forestall disastrous regime shifts. The variation of the equipartition point $x_{1/2}$ versus a bifurcation parameter provides advanced knowledge of the approach to the PTP. A number of signatures of impending regime shifts have been proposed so far in the scenario of the saddle-node bifurcation [1, 2, 15, 16]. These include the critical slowing down, rising variance and skewness of a subpopulation probability distribution and the rising lag-1 auto-correlation function, as a function of the bifurcation parameter. If the model governing the dynamics of the system of interest is known, one can calculate the return time (the time taken by a system to regain its stable steady state after the system is weakly perturbed) using

the standard procedure of linear stability analysis [1, 45]. The return time diverges (critical slowing down) as a bifurcation point at which the steady state loses its stability is approached. A variety of techniques, e.g., the linear noise approximation [30] can be used to calculate quantities like the variance and the lag-1 autocorrelation function. The demonstration of the skewness of a subpopulation probability distribution as the bifurcation point is approached requires the use of somewhat ad-hoc procedures for obtaining a component probability distribution from the total distribution [13, 15]. One of these is to introduce suitable cut-off procedures to isolate a component distribution which exhibits rising skewness as a bifurcation point is approached. The formalism developed in this paper for a quantitative characterization of noise-induced regime shifts is applicable even when only time series data are available, without an adequate knowledge of the underlying model. To make the analysis possible, one has to identify a parameter θ which is a relevant bifurcation parameter. If time series data are available over a sufficiently long interval of time, for which θ remains constant, one can construct the probability distribution of the relevant variable from the data itself. Similar probability distributions can be computed for other values of θ if it is a changing function of time. Several studies have recently been undertaken on the early signatures of sudden regime shifts [2, 3, 17] when only time series data are available. As already discussed in the Introduction, regime shifts are often noise-induced. Such transitions, depending on the magnitude of the noise, may occur even when the system is not close to a bifurcation point. The frequency of the transitions is expected to increase in the vicinity of the bifurcation point. The method developed in the present study is based on the concept of the PTP which precedes the bifurcation point. The sharp fall of the equipartition point $x_{1/2}$ in the vicinity of the PTP also serves as an indicator of the system approaching a bifurcation point apart from signifying a stochastic shift in the propensity of the system to be in one of two alternative regimes. The utility of the PTP formalism has been demonstrated for both one-variable and two-variable models. The $x_{1/2}$ versus bifurcation parameter plot further carries a distinctive signature of the saddle-node bifurcation when the plot exhibits a sharp fall. The sharp fall of $x_{1/2}$ in the vicinity of the PTP is an indicator of a stochastic regime shift. Regime shifts are of common occurrence in wide ranging dynamical systems. The shifts may be due to changing parameters resulting in a bifurcation or due to a large perturbation causing a switch between one attractor state to another across the border separating the basins of attraction of the attractors or due to noisy dynamics with the fluctuations in key variables driving the regime shift. A large number of studies [2, 3, 17] provide a quantitative characterization of sudden regime shifts at the bifurcation points. The present study focuses on a quantitative indicator of stochastic regime shifts which has a simple physical interpretation and is straightforward to compute.

Acknowledgements

SG acknowledges the support of CSIR, India, under Grant No. 09/015(0361)/2009-EMR-I.

-
- [1] Strogatz H. 1994 *Nonlinear dynamics and chaos - with applications to physics, biology, chemistry and engineering* (Addison-Wesley)
- [2] Scheffer M. *et. al* 2009 Nature 461 53
- [3] Scheffer M. 2009 *Critical Transitions in Nature and Society* (Princeton University Press)
- [4] Scheffer M. and Carpenter S. R. 2003 Trends in Ecology and Evolution 18 648
- [5] Lenton T. M. 2011 Nature Climate Change 1, 201
- [6] Venegas J. G. *et. al* 2005 Nature 434, 777
- [7] McSharry P. E., Smith L. A. and Tarassenko L. 2003 *Nature Medicine* 9(3), 241
- [8] Ozbudak E. M., Thattai M., Lim H. N., Shraiman B. I. and Van Oudenaarden A. 2004 Nature 427, 737
- [9] Pomeroy J. R. 2008 *Curr. Opin. Biotechnol.* 19, 381
- [10] Ferrell J. E. and Xiong W. 2001 Chaos 11, 227
- [11] Chen L., Liu R., Liu Z-P., Li M. and Aihara K. 2012 Scientific Reports 2, 342
- [12] May R. M. 1977 Nature, 269, 471
- [13] Guttal V. and Jayprakash C. 2007 Ecological Modelling 201, 420
- [14] van Nes E., Scheffer M., Van der Bers M. and Coops H. 2002 Aquatic Botany 72, 275
- [15] Guttal V. and Jayprakash C. 2008 Ecology Letters 11(5) 450
- [16] Dakos V., Van Nes E. H., D'Odorico P. and Scheffer M. 2012 Ecology 93(2), 264
- [17] Scheffer M. *et. al* 2012 Science 338, 344
- [18] Horsthemke V. and Lefevre R. 1984 *Noise Induced Transitions* (Springer-Verlag, Berlin)
- [19] Kepler T. B. and Elston T. C. 2001 Biophysical Journal 81, 3116
- [20] Song C. *et. al* 2010 PLoS Computational Biology 6, e1000699
- [21] Zakharova A., Vadvivasova T., Anishchenko V., Koseska A. and Kurths J. 2010 Phys. Rev. E 81, 011106
- [22] Karmakar R. and Bose I. 2007 Phys. Biol. 4, 29
- [23] Samoilov M., Plyasunov S. and Arkin A. P. 2005 Proc. Natl. Acad. Sci. 102(7), 2310
- [24] To T. L. and Maheshri N. 2010 Science 327(5969), 1142
- [25] Gardner T. S., Cantor C. R. and Collins J. J. 2000 Nature 403, 339
- [26] Gillespie D. T. 1977 J. Chem. Phys. 81, 2340
- [27] Sureka K., Ghosh B., Dasgupta A., Basu J., Kundu M. and Bose I. 2008 Plos One 3, e1771
- [28] Ghosh S., Sureka K., Ghosh B., Bose I., Basu J. and Kundu M. 2011 BMC Syst. Biol. 5, 18
- [29] Gardiner C. W. 1983 *Handbook of Stochastic Methods* (Springer-Verlag, Berlin)
- [30] van Kampen N. G. 1992 *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam)
- [31] Risken H. 1984 *The Fokker-Planck Equation* (Springer-Verlag, Berlin)
- [32] Jin W. D., Li C. and Zhi K. S. 1994 Phys. Rev. E 50, 2496
- [33] Rao M., Chen Y., Vemuri B. C. and Fei Wang 2004 Information Theory, IEEE Transactions 50, 1220
- [34] Zheng X-D., Yan X-Q. and Tau Y. 2011 PLoS One 6, e17104
- [35] Gillespie D. T. 1992 *Markov Processes: An Introduction for Physical Scientists* (Academic Press, London)
- [36] Scott M., 2006 *Tutorial: Genetic Circuits and Noise*, available at [M. Scott's website](#)
- [37] Erban R., Chapman J. and Maini P. 2007 arXiv:0704.1908[q-bio.SC]
- [38] Gillespie D. T., Hellander A. and Petzold L. R. 2013 J. Chem. Phys. 138, 170901
- [39] Gillespie D. T. (2007) Annu. Rev. Phys. Chem. 58, 35
- [40] Gillespie D. T. 2000 J. Chem. Phys. 113, 297
- [41] Smits W. K., Kuipers O. P. and Veening J. -W. 2006 Nat. Rev. Microbiol. 4, 259
- [42] Maamar H., Raj A. and Dubnau D. 2007 Science 317, 526
- [43] Tiwari A., Bálazsi G., Gennaro M. L. and Igoshin O. A. 2010 Phys. Biol. 7, 036005
- [44] Sureka K., Ghosh B., Dasgupta A., Basu J., Kundu M. and Bose I., *unpublished data*
- [45] Wissel C. 1984 Oecologia 65, 101