Flexibility and Robustness of Biochemical Switches

Zahedeh Basharnadesh[1]
[1]*Department of Mathematics. Uppsala University. SE–751 06 Uppsala. Sweden.*

Biochemical switches that can toggle between two discrete molecular states is an ubiquitous mechanism for digital processing of external inputs in cell signaling. A necessary condition for obtaining such bi-stabile behavior is positive and non-linear feedback. In this study we examine the most basic positive feedback process that generates bi-stable dynamics and quantify its properties with respect to chemical noise and delays in the feedback. We find that the Linear Noise Approximation (LNA) as an analytical tools makes it possible to estimate second moments of fluctuations around steady states in a bistable system and due to these fluctuations the system jumps frequently between its stable steady states, the frequency of jumping can be estimated by escape time derived from Chemical Langevin Equation (CLE). If the feedback is indirect, operating early in a series of synthesis events, the high state of the switch is destabilized and hysteresis is no longer apparent.

I. INTRODUCTION

Multistability is the capability of systems achieving multiple steady states in response to an external signal and bistablility is a special case and also of a particular interest in biology as it leads to switch-like behavior. A bistable system can switch between two steady states but not rest in intermediate ones. In cell biology, switch-like behavior is a characteristic of many important biochemical reaction networks which has been investigated: cell-differentiation [1], autophosphorylation [2], multisite phosphorylation [3], complex enzyme-driven reaction networks [4], caspase activation [5], transcriptional regulation [6] and cell-cycle [7, 8]. A property associated with switches is hysteresis, which means that once the system has been switched ”on” by exceeding the stimuli to a threshold, then it cannot be switched ”off” by decreasing it to that specified threshold. In other words, these systems have ”memory”, while the input is withdrawn the system does not return to previous state. The hysteresis is of importance when considering switches that need to be robust to external perturbations, e.g. cell differentiation [1].

Bistability arises in biological systems through many mechanisms, but a commonly accepted underlying mechanism is feedback. Although, in constructing models, feedback is far from being sufficient to present bistability, but as a simplest form of positive feedback, i.e. autocatalysis, it guarantees bistability in very small reaction sets [9].

Once the average behavior of the system should be described a mean-field approach suffices. However, once the fluctuations are considerable, the macroscopic rate equations of the chemical concentrations will not describe the system behavior properly anymore. Since the bistability will be a consequence of chemical reactions it is relevant to ask for what time scales the switch may operate properly since chemical noise will at some point make the system spontaneously switch to the other state. The Chemical Master Equation (CME) describes the time evolution of the probability distribution of the states of the system and can (in principle) be solved analytically. If not solvable analytically one may integrate it numerically from the Stochastic Simulation Algorithm (SSA), a Monte Carlo procedure that generates time trajectories of the molecular populations in exact accordance with the CME [10]. However, except in simple cases [11], an exact closed solution to CME is generally not attainable due to the high dimensionality of the problem. An approximation method to obtain second moments (variances and covariances) from the CME is the Linear Noise Approximation (LNA) [12]. The method is developed by expanding the CME in inverse powers of the system size Ω around the steady states which can be obtained by solving the macroscopic rate equations. The LNA provides information of the chemical noise properties as a function of the parameters and topology of the underlying biochemical reaction network [13]. In the case of a molecular switch, fluctuations may play an important role in changing the steady states in addition to external stimuli—big enough fluctuations might make the system toggle between the two steady states. Almost all feedback mechanisms operate indirectly on molecules upstream in a biochemical pathway which result in a time lag in its response to deviations from the steady state, a time lag that will correspond to the time needed for making the final product. Recent studies on delayed biochemical feedback systems, e.g. delays associated with transcription and translation includes investigations of both discrete stochastic and continuous deterministic modeling [14]. Also an extension to the LNA has been studied for small gene regulatory motifs as time-delayed systems [15]. In a recent paper [16], an analytical expression has been driven to quantify properties of time-delayed negative feedbacks along with comparing two models, step feedback model (SFM) in which every single reaction from the multi-step synthesis is taking into account, and delay feedback model (DFM), in which the intermediate reactions are absorbed into one time delay. Here we study the sensitivity of the bistable systems to delays and investigate the effects of delay on hysteresis.

II. MODEL

A. Non-Delay Process

To promote interpretability and understanding we use the simplest model that displays the characteristics of a switch. The system attains its bistability by means of positive feed-

back regulation. Assume that a intracellular system generates a factor $\mathbf{X}$ as a response to an external stimulation S_E that modulates the external signal of the factor. Furthermore, the factor stimulates its own production through a positive feedback loop $a\Phi(x)$, where a is the maximal self-induced production of the factor ($0 \leq \Phi(x) \leq 1$). Finally, the factor is degraded with the decay rate constant b. The response system can be summarized by the following reactions:

$$\emptyset \xrightarrow{S_E + a\Phi(x)} \mathbf{X}, \qquad \mathbf{X} \xrightarrow{b} \emptyset\,. \tag{1}$$

The time scale of the dynamics is set by the degradation rate constant b. Thus, we can scale the time in terms of b to reduce the number of parameters. The two remaining parameters are then the (normalized) external signal $\sigma_E = S_E/b$ and the (normalized) maximal self-induced rate $\alpha = a/b$.

B. Delay Process

Most metabolites in gene regulatory networks specially those involving macro molecules, such as transcription, translation and degradation processes are synthesized through a series of biochemical steps. Such multi step synthesis are often feedback regulated early in the synthesis path, which imply that the regulation of the product at time t is effectuated after all steps are finalized, on average $t + t_d$ seconds later. This leads to a feedback delay in the regulation which may hamper the fidelity of the feedback response. For a genetic switch, e.g. realized by a protein synthesized in a series of steps with the positive feedback in the form of transcriptional regulation, it is important to quantify the impact of a delayed feedback to the properties of the switch. Homeostatic control through indirect, delayed, negative feedback may be accompanied by direct (negative) feedback later in the synthesis path to obtain precision—with the cost of synthesizing unwanted intermediate molecules [15]. For a genetic switch it is more complicated. It is not obvious how to complement the transcriptional feedback in a similar way and still remain the properties of the switch. Therefore we will consider a switch operating indirectly in a multi-step synthesis path in order to quantify the impact of the feedback delay to properties like hysteresis and stability. The biochemical reactions in N steps leading to the final molecule X_{N+1} can be written as:

$$\xrightarrow{S_E + a\Phi(x_{N+1})} X_1 \xrightarrow{k_1} X_2 \xrightarrow{k_2} \ldots \xrightarrow{k_N} X_{N+1} \qquad \text{and} \qquad X_{N+1} \xrightarrow{b} \emptyset. \tag{2}$$

The re-scaling can be performed just as in the non-delayed counterpart with the same parameters, where the rates k_i in addition are scaled with b, $\kappa_i = k_i/b$ and the average delay $\tau_d = bt_d$.

III. RESULTS

The macroscopic dynamics of the switch is described by an Ordinary Differential Equation (ODE) with a governing law describing how the concentration rate of a species changes. The governing function has its own parameters rising from the underlying biological problem and to study the dynamics of the system one should examine the stability of the system depending on these parameters. Therefore, we use two tools from mathematical studies of dynamical systems, the bifurcation and linear stability analysis. The bifurcation analysis is the study of number of fixed points and their stability depending on parameters changing in the system. The linear stability analysis is examining the evolution of a perturbation from fixed point(s) x_*, which is determined by the eigenvalue $\lambda = f'(x_*)$, where $f(x) = \sigma_E + \alpha\Phi(x) - x$ is the governing law. We model the feedback Φ with a Hill function with Hill coefficient n (see Methods Section).

The bifurcation curves by which the domain of bistability is separated from mono-stability, are the parametric equations defining the external signal σ_E and the maximal self-induced α as a functions of concentration x_*:

$$\alpha(x_*) = \frac{(1 + x_*^n)^2}{n x_*^{n-1}}, \tag{3a}$$

$$\sigma_E(x_*) = x_*\left(1 - \frac{1 + x_*^n}{n}\right). \tag{3b}$$

These parametric equations are derived when both the governing law and its derivative vanishes. By plotting them in parametric space $(\alpha(x_*), \sigma_E(x_*))$ for different n, the range of external signal σ_E, feedback strength α and n for which the system is bistable or monostable is observable, which is done in Fig 1a. For the linear stability analysis we set the Hill coefficient $n = 2$ as it is the smallest value to display bistability and then simply pick $\alpha = 1.8$ in the bistable area. The determining parameter for the eigenvalue λ is hence the external signal σ_E. For values of σ_E where the eigenvalue is negative, $\lambda < 0$, the corresponding fixed point(s) is stable while the small perturbation from the fixed point(s) decays with rate of $e^{\lambda t}$ and where the eigenvalue is positive, $\lambda > 0$, the corresponding fixed point(s) is unstable and a small perturbation from fixed point(s) increases with rate of $e^{\lambda t}$. The result is sketched as bifurcation plot in Fig. 1b.

The bifurcation analysis will reveal the signal strength needed to induce the system to switch to the high state and also at which lower level the external signal no longer hold the switch in its on state. This ability of the system is known as hysteresis. While the system modeled stochastically for the same parameters in deterministic model ($n = 2$ and $\alpha = 1.8$) the hysteresis is only expected to appear for sufficient many molecules as the chemical noise will make the system escape the stable states. Fig. 1c shows the hysteresis effect for three different volume $\Omega = 32, 128, 512$ and compare with deterministic model.

One approach in studying the stochastic behavior of the switch is through stochastic simulation. The Gillespie Algorithm generates trajectories of, in principle, any biochemical reaction network [20] . Fig. 2a is a sample path of the system generated by the Gillespie algorithm for $\Omega = 128$ and $\sigma_E = 0.135$. The trajectory shows how the concentration

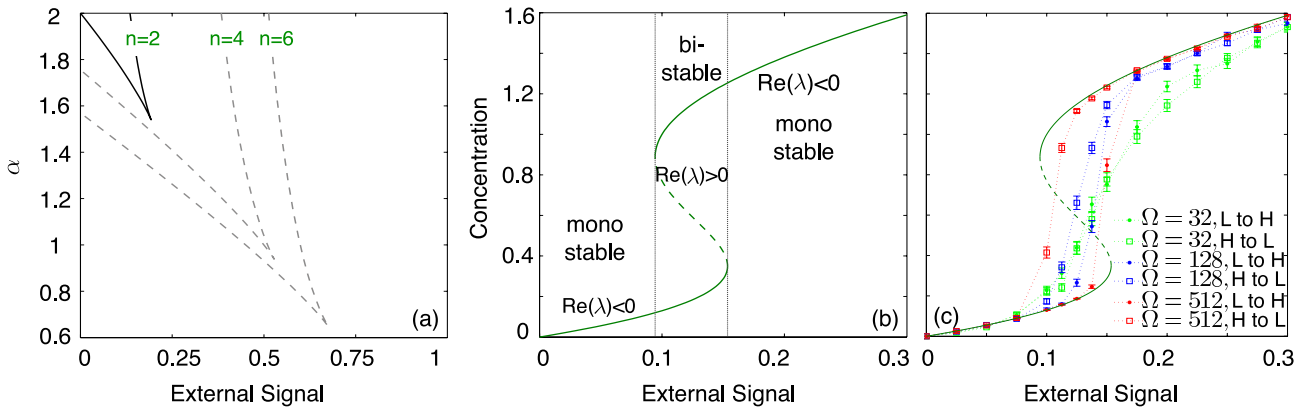

FIG. 1 (a) The domain of bistability for the concentration of molecules for different feedback strengths α as a function of the external signal σ_E for different values of the Hill coefficient n. (b) Bistability plot for the concentration of molecules for feedback strength $\alpha = 1.8$ and Hill coefficient $n = 2$. (c) Hysteresis effect for different volume $\Omega = 32, 128, 512$ as a function of external signal σ_E.

of the system changes with time and it catches both lower and upper steady states and jumps between them. Fig. 2b shows the probability density function of the system for different volumes. It is clear from figure in small copy number of molecules the switching happens more frequently and as the copy numbers of molecules increases the switch behavior converges to the ODE model.

Although the statistical information of the system for specified range of parameters can be obtained by Gillespie Algorithm, an analytical solution to master equation contains more information of the over-all noise characteristics. We can not solve CME exactly for nonlinear feedback reactions, but

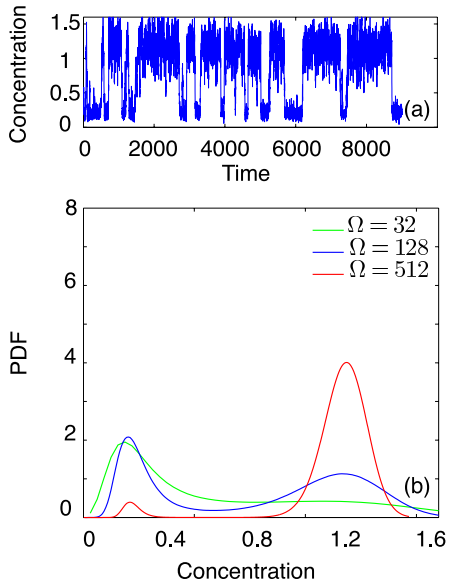

FIG. 2 (a) A sample path of the change of molecules concentration vs. time for external signal $\sigma_E = 0.135$ simulated by Gillespie Algorithm. (b) Probability distribution function for external signal $\sigma_E = 0.135$ for different volumes $\Omega = 32, 128, 512$

the Linear Noise Approximation (LNA) is an approximation method in which the master equation is expanded in powers of the square root of system's volume around the steady state. From the LNA one obtains a fluctuation dissipation relation from which the linearized second moments **C** can be calculated,

$$\mathbf{AC} + \mathbf{CA}^T + \mathbf{V} = 0, \tag{4}$$

where **A** is the Jacobian matrix and **V** is the diffusion matrix. In our model with one species X, the variance C_X is (see Methods Section):

$$C_X = \frac{V}{-2A} = \frac{< X >}{[1 - \alpha\Phi']}. \tag{5}$$

For small x when x is e.g. in the lower state, $x_* \approx \sigma_E$

$$C_X \approx \frac{< X >}{[1 - \alpha x_*]} \approx \frac{\Omega\sigma_E}{1 - 2\alpha\sigma_E}.$$

Fig. 3a shows a comparison between the variance obtained from Eq. (5), solid line, and the simulated CME, marks, for different values of σ_E. Therefore, in studying biochemical switches where number of molecules is large LNA can give valuable analytical expression of the second moments of fluctuations around steady states.

Once the biological problem e.g. the cell cycle modeled by a bistable system, one may want to measure the time needed to leave the off state and enter the on state, which is called escape time. This value obtained from Gillespie Algorithm can be compared with its analytical value. This is done by calculating the spending time in the current stable steady state, say x_a, before jumping to the other stable steady state x_c, by passing through the middle unstable state x_b. The estimated time obtained from Chemical Langevin Equation (CLE) is given by (see Methods Section):

$$\tau_{ac} = \frac{2\pi}{\sqrt{U''(x_a)U''(x_b)}} exp[\frac{U(x_b) - U(x_a)}{\theta}]. \tag{6}$$

in which $U(x)$ is the potential function while $U'(x) = -f(x_*)$ is the governing law at fixed point(s) and $\theta = \sigma^2/2$ where σ is the diffusion term in CLE.

Fig. 3b shows the potential, $U(x)$, by integrating $-U'(x) = \sigma_E + \alpha\frac{x^2}{1+x^2} - x$ over x. The positions of the fixed points and their energy levels can be seen in the figure. In Fig. 3c we display the estimated scape time from the lower state x_a to the upper state x_c by passing the middle unstable fixed point given by CLE , red line, and data obtained from simulating CME by Gillespie Algorithm, blue marks.

The effect of delay on the system can be seen from the probability distribution function of the system , Fig 4a ($\Omega = 128$ and $\sigma_E = 0.135$) and its hysteresis plot, Fig 4b. By increasing the feedback delay, the upper fixed point disappears and the hysteresis effect, and hence the memory, vanishes.

IV. DISCUSSION

In this study we have examined the chemical reaction dynamics of switch-like behavior, i.e. all-or-none character in response to an external signal. We examined continuous and deterministic model of the system to calculate the amount of external signal in response to which the system obtains low concentration, off state, or high concentration, on state. The discrete and stochastic description of the system while the change in the external signal leads to bistable mode, reveals that disregarding what the initial copy number of species is the system obtain both high and low states by jumping between them. The time can be estimated from the dynamics of the switch and furthermore the necessary copy number of the participating molecules can be calculated as a function of the desired stability of the particular switch.

From the bistability analysis, the level of the external signal as well as the value of species concentration determine the phase of the system under study. Moreover, looking into hysteresis plot reveals the intrinsic memory of the switch as a function of both volume and also on feedback delays. For small copy numbers, modeled by the system size Ω, the switch has less well defined states, is frequently toggling between the two states and display a weak hysteresis. Also for long delays, modeled by number of intermediate reactions, the stability of higher state is reduced and shows weaker hysteresis this can be understood by stability analysis of fixed point. Therefore, in order to construct a biochemical switch, a significant number of molecules and short feedback delay are needed to generate the associated properties of a switch, all-or-none response and hysteresis.

The probability distribution function together with the potential function of the system reveals that for specific value of external signal the on state is more attractive, when there is large number of molecules and less toggling between states. So, with enough copy number of molecules the desired state of the switch can be controlled by the external signal.

V. METHODS

A. Macroscopic Analysis

1. Ordinary Differential Equation

To see how the concentration of factor **X** in reaction (1), shown by χ, changes with time the reaction rate is obtained by applying a rate law, such as mass action or Michaelis-Menten kinetics, in a form of ordinary differential equation (ODE):

$$\frac{d\chi}{dT} = B + A(\frac{\chi^n}{M^n + \chi^n}) - \kappa\chi \ , \tag{7}$$

the positive feedback can appear in any function showing sigmoidal behavior, here we chose the Hill function. Before analyzing the model it is essential [17] to express it in non-dimensional form:

$$\frac{dx}{dt} = \sigma_E + \alpha(\frac{x^n}{1 + x^n}) - x \ . \tag{8}$$

Here x shows the scaled concentration of the species **X**, t is the scaled time, parameter σ_E presents the external signal, α is the strength parameter of the feedback loop and n is the parameter of Hill function known as Hill coefficient. From now on we choose $n = 2$ which makes the calculation straightforward and also is more meaningful in biological context. Here we refer to $f(x) = \sigma_E + \alpha(\frac{x^n}{1+x^n}) - x$, as governing law of the system.

The study of the conditions under which stability and bistability arise is well established in differential equation modeling by linear stability and bifurcation analysis.

Linear Stability Analysis: Suppose that x_* is the fixed point(s) of the system, which means $f(x_*) = 0$;

$$x_* = \sigma_E + \alpha\frac{(x_*^2)}{1 + x_*^2} \ . \tag{9}$$

And $\delta(t) = x(t) - x_*$ is a small perturbation from x_*. Using Taylor's expansion we obtain:

$$\frac{dx}{dt} = f(x_* + \delta(t)) = f(x_*) + \delta f'(x_*) + O(\delta^2) \ ,$$

the term $O(\delta^2)$ is negligible as we want to see the behavior of the system linearly,

$$\frac{dx}{dt} = \frac{d\delta(t)}{dt} = \delta(t) f'(x_*) \quad \Rightarrow \quad \delta(t) = Ce^{\lambda t}, \quad (\lambda = f'(x_*))$$

λ, the eigenvalue in terms of our model is:

$$\lambda = f'(x_*) = \frac{2\alpha x_*}{(1 + x_*^2)^2} - 1 \tag{10}$$

which should be less than zero ($\lambda < 0$) for each fixed point to guarantee the stability of the system near each of them, on

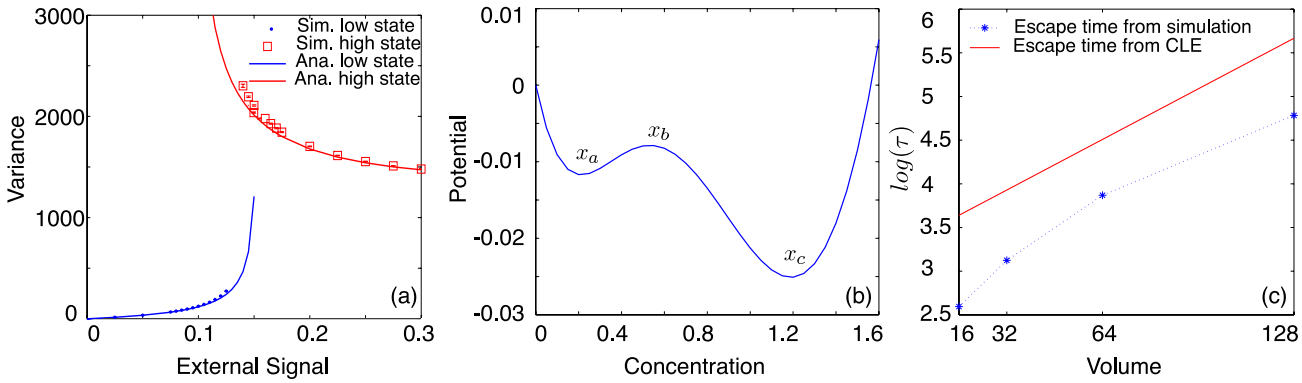

FIG. 3 (a) Comparison between variance obtained from simulation (Sim.) by Gillespie Algorithm and analytical (Ana.) LNA. (b) Potential function $U(x) = \int -\sigma_E - \alpha\frac{x^2}{1+x^2} + x$ as a function of concentration of molecules, $\sigma_E = 0.135$. (c) Comparison between estimated time from theory, CLE and numerical methods vs. volume of the system.

the other hand observing the conditions of bistability is accessible through stability diagram by plotting bifurcation curve

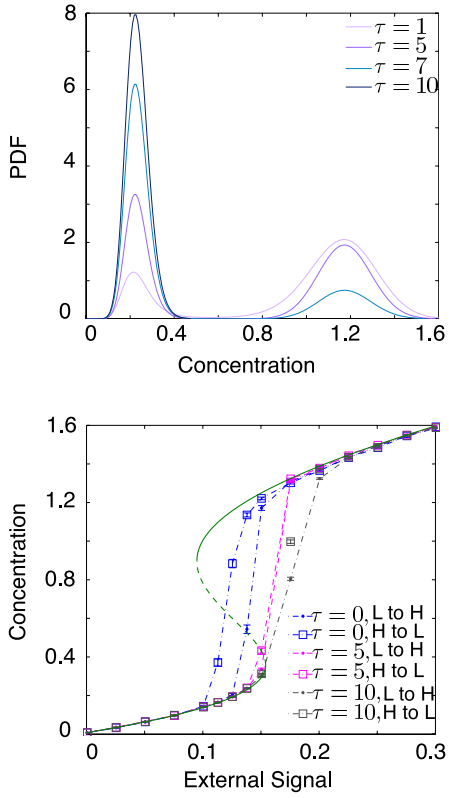

FIG. 4 (a) Probability distribution function for external signal $\sigma_E = 0.135$ for different delay length $\tau = 1, 5, 7, 10$. (b) Hysteresis effect for different delay length $\tau = 0, 5, 10$ as a function of external signal σ_E.

in $(\alpha(x_*), \sigma_E(x_*))$ plane with parametric forms:

$$\alpha(x_*) = \frac{(1+x_*^2)^2}{2x_*}\,, \tag{11a}$$

$$\sigma_E(x_*) = x_*(1 - \frac{1+x_*^2}{2})\,. \tag{11b}$$

For a fixed $\mathring{\alpha}$ and $\mathring{\sigma_E}$ in bistable area, found from parametric equations (11) one can find the value of λ for different fixed points x_*'s. By a simple calculation one obtains $Re(\lambda_1) < 0$, $Re(\lambda_2) >$ and $Re(\lambda_3) < 0$, respectively for $x_{*1} < x_{*2} < x_{*3}$, which means x_{*1} and x_{*3} are stable while x_{*2} is unstable fixed points.

2. Delay Differential Equation

The macroscopic dynamics of the delayed process can be govern by a set of differential equations described the chain of reactions (2) by the following differential equations:

$$\begin{aligned}
\frac{dx_1}{dt} &= S_E + a\Phi(x_{N+1}) - k_1 x_1\,,\\
\frac{dx_2}{dt} &= k_1 x_1 - k_2 x_2\,,\\
&\vdots\\
\frac{dx_{N+1}}{dt} &= k_N x_N - b x_{N+1}\,.
\end{aligned}$$

Where x_i's are the concentration of components X_i's. By knowing the rate of intermediate reactions, one can absorb all medial reactions with one finite delay τ_d. After non-dimesionalizing the equations with b, the life time of the last component, the delay differential equation of the last component $N+1$ is:

$$\frac{dx_{N+1}}{d\tau} = S_E + \alpha\Phi[x_{N+1}(t-\tau)] - x_{N+1}\,. \tag{12}$$

The equation below is a delayed counterpart of Eq.(8) of our model:

$$\frac{dx_t}{dt} = \sigma_E + \alpha(\frac{x_{t-\tau}^2}{1+x_{t-\tau}^2}) - x_t \, . \tag{13}$$

Linear Stability Analysis: First we find the stationary solution to the Eq. ((13)). For less complexity in appearance let $x_t = x$ and $x_{t-\tau} = x_\tau$

$$\frac{dx}{dt} = \sigma_E + \alpha(\frac{x_\tau^2}{1+x_\tau^2}) - x = 0 \Rightarrow x = \sigma_E + \alpha(\frac{x_\tau^2}{1+x_\tau^2})$$

as there is no time dependency in stationary points, $\tau = 0$, the steady state(s) x_* meets the following statement:

$$\sigma_E(1+x_*^2) + \alpha x_*^2 = x_*(1+x_*^2) \tag{14}$$

Now assume a small perturbation δ from x_*

$$x(t) = x_* + \delta(t)$$

δ defines the perturbation from fixed point at time t and δ_τ the perturbation from fixed point at time $t-\tau$.

$$\frac{dx}{dt} = \frac{d\delta}{dt} = \sigma_E + \alpha(\frac{(x_*+\delta_\tau)^2}{1+(x_*+\delta_\tau)^2}) - (x_*+\delta)$$

$$\begin{aligned}\dot{\delta}(1+x_*^2) = \sigma_E(1+x_*^2+2x_*\delta_\tau+\delta_\tau^2) + \alpha(x_*^2+2x_*\delta_\tau+\delta_\tau^2) \\ - x_*(1+x_*^2+2x_*\delta_\tau+\delta_\tau^2) - \delta(1+x_*^2+2x_*\delta_\tau+\delta_\tau^2)\end{aligned}$$

in which, the terms $\delta\delta_\tau$ and δ_τ^2 are negligible. So,

$$\begin{aligned}\dot{\delta}(1+x_*^2) = \sigma_E(1+x_*^2) + \alpha(x_*^2) - x_*(1+x_*^2) \\ + \delta_\tau 2x_*(\sigma_E + \alpha - x_*) - \delta(1+x_*^2)\end{aligned}$$

using the stationary solution ((14)),

$$\frac{d\delta(t)}{dt} = \frac{2\alpha x_*}{(1+x_*^2)^2}\delta(t-\tau) - \delta(t) \, .$$

Now we make the ansatz ;

$$\begin{cases}\delta(t) = e^{\lambda t} \\ \delta(t-\tau) = e^{\lambda(t-\tau)} = \delta(t)e^{-\lambda\tau}\end{cases}$$

then,

$$\lambda\delta(t) = \frac{2\alpha x_*}{(1+x_*^2)^2}\delta(t)e^{-\lambda\tau} - \delta(t)$$

$$\lambda = \frac{2\alpha x_*}{(1+x_*^2)^2}e^{-\lambda\tau} - 1 \, . \tag{15}$$

if $\tau = 0$, $\lambda = (\frac{2\alpha x_*}{(1+x_*^2)^2} - 1)$ should be less than zero to guarantee the stability, hence:

$$\lambda < 0 \quad \Rightarrow \quad \frac{2\alpha x_*}{(1+x_*^2)^2} < 1$$

if $\tau \neq 0$, define $\upsilon = \frac{2\alpha x_*}{(1+x_*^2)^2}$,

$$\lambda = \upsilon e^{-\lambda\tau} - 1 < 0$$

Set $\lambda = r + i\omega$,

$$r + i\omega = \upsilon e^{-r\tau}\cos(\omega\tau) - 1 - i\upsilon e^{-r\tau}\sin(\omega\tau)$$

Separating real and imaginary part, we obtain:

$$\begin{aligned} r &= \upsilon e^{-r\tau}\cos(\omega\tau) - 1 \\ \omega &= -\upsilon e^{-r\tau}\sin(\omega\tau)\end{aligned}$$

if $r = 0$,

$$\begin{aligned} 1 &= \upsilon\cos\omega\tau \quad (I) \\ \omega &= -\upsilon\sin\omega\tau \quad (II)\end{aligned}$$

$$(I)^2 + (II)^2 = 1 + \omega^2 = \upsilon^2(\cos^2(\omega\tau) + \sin^2(\omega\tau))$$

if $\upsilon^2 < 1$ there is no oscillation. If $\upsilon^2 > 1$:

$$\begin{aligned}\omega\tau &= \arccos(-\frac{1}{\upsilon}) \\ \tau^* &= \frac{1}{\sqrt{\upsilon^2-1}}\arccos(-\frac{1}{\upsilon})\end{aligned}$$

when $\tau = \tau^*$ the switch produce stable oscillation, for $\tau > \tau^*$ the oscillation grow and for $\tau < \tau^*$ the oscillation dies off.

B. Stochastic Analysis

1. Chemical Langevin Equation

As mentioned earlier the deterministic approach to biochemical reactions fails to take into account their stochastic nature of such processes. Therefore an additive noise term to the Eq.(8) might refine our model.

$$dx(t) = f(x)dt + g(x)dW(t). \tag{16}$$

Eq. (16) is called Langevin Equation or more generally Stochastic Differential Equation (SDE). The solutions of Langevin Eq. give different random trajectories of the process. However, by applying Ito's formula [18] to a SDE one

can instead of all different trajectories, calculate the probability density function as a solution to the following equation,

$$\partial_t p(x,t|x_0,t_0) = -\partial_x[f(x)p(x,t|x_0,t_0)] + \frac{1}{2}\partial_x^2[g(x)^2 p(x,t|x_0,t_0)] \,. \quad (17)$$

Eq. (17) is called Fokker Planck Equation (FPE), in which the deterministic functions $f(x)$ and $g(x)$ are drift and diffusion terms, respectively. Although we are not taking this approach to study stochastic behavior of the system, we use it in estimating escape time. Escape time is the time the system needs to leave the current steady states and jump to the other, in our case the transition from middle unstable fixed point could determines the jump happens. Consider three steady states of the system as a, b and c, where $a < b < c$ and correspond to the lower, middle and upper steady states, respectively. Consider the SDE:

$$dx = -U'(x)dt + \sigma dW(t)$$

where $-U'(x)$ is the vector field of ODE. (8), ($U'(x) = x - \sigma_E - \alpha(\frac{x^2}{1+x^2})$) and its corresponding FPE is :

$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial(U'(x)P(x,t))}{\partial x} + \theta\frac{\partial^2 P(x,t)}{\partial x^2}$$

where $\theta = \sigma^2/2$. Now the estimated time [12] is given by:

$$\tau_{ac} = \frac{2\pi}{\sqrt{U''(a)U''(b)}} exp[\frac{U(b)-U(a)}{\theta}] \,. \quad (18)$$

θ can be inserted by calculating σ from Chemical Langevin Equation (CLE) of the system :

$$dx = \sum_{j=1}^{2} S_j f_j(x) + \frac{1}{\sqrt{\Omega}}\sum_{j=1}^{2} S_j \sqrt{f_j} dW_j \,,$$

since W_j's are independent we have

$$dx = -U'(x)dt + \sigma dW(t) \,,$$

where

$$\sigma^2 = \frac{1}{\Omega}\sum_j S_j^2 f_j$$

and W is a Wiener process, $dW \sim \mathcal{N}(0, dt)$. So $\sigma = \frac{1}{\sqrt{\Omega}}\sqrt{\sigma_E + \frac{\alpha x_a^2}{1+x_a^2} + x_a}$.

2. Chemical Master Equation [19]

Consider a set of biochemical reactions with N different chemical components homogeneously distributed in a living cell with volume Ω. The state of such a homogenous system is defined by the number of molecules of each components $\mathbf{n} = [n_1, \ldots, n_N]$, in which n_i shows the current number of molecules of species i. A state change takes place by any one of R reactions. When reaction j occurs the chemical component number i changes from n_i to $n_i + S_{ij}$ molecules. The integers S_{ij}, $i = 1, 2, \ldots, N$; $j = 1, 2, \ldots, R$; are the elements of the $N \times R$ stoichiometric matrix $\mathbf{S}$ of the reaction network. The probability that the reaction j occurs in a small time interval δt is given by $\Omega\tilde{f}_j(\mathbf{n}, \Omega)\delta t$, where $\tilde{f}_j(\mathbf{n}, \Omega)$ is the transition rate for reaction j.

The probability that a system of chemical reactions is in state $\mathbf{n}_m$ at time $t + dt$, $P(\mathbf{n}_m, t + dt)$, will depend on the probability that it was in state $\mathbf{n}_m$ at time t, $P(\mathbf{n}_m, t)$, as well on the probability that the system will either reach $\mathbf{n}_m$ from or leave $\mathbf{n}_m$ to any other state in the system during the interval dt:

$$P(\mathbf{n}_m, t+dt) \approx P(\mathbf{n}_m, t) + dt.\sum_{k\neq m} W(\mathbf{n}_k, \mathbf{n}_m)P(\mathbf{n}_k, t) - dt.\sum_{k\neq m} W(\mathbf{n}_m, \mathbf{n}_k)P(\mathbf{n}_m, t) \quad (19)$$

$W(\mathbf{n}_k, \mathbf{n}_m).dt$ is the probability of a transition from state $\mathbf{n}_k$ to state $\mathbf{n}_m$ in any time interval of length dt. Accordingly, the first sum in Eq. (19) is the total probability that the system reaches the state $\mathbf{n}_m$ and the second sum is the total probability that the system leaves $\mathbf{n}_m$ in the time interval $(t, t+dt)$. Moving $P(\mathbf{n}_m, t)$ to the left hand side then dividing by dt and taking the limit $dt \to 0$ lead to the Chemical Master Equation (CME).

$$\frac{dP(\mathbf{n}_m, t)}{dt} = \sum_{k\neq m} W(\mathbf{n}_k, \mathbf{n}_m)P(\mathbf{n}_k, t) - \sum_{k\neq m} W(\mathbf{n}_m, \mathbf{n}_k)P(\mathbf{n}_m, t) \,. \quad (20)$$

The chemical reaction system discussed here is characterized by the transition rates $\Omega\tilde{f}_j(\mathbf{n}, \Omega)$, which can be used to formulate its corresponding master equation with the help of stoichiometric matrix $\mathbf{S}$. To do this, identify for each reaction j the rate $\Omega\tilde{f}_j(\mathbf{n}, \Omega)$ for the system's transition from state $\mathbf{n}$ to state $\mathbf{n} + S_{.j}$ with $W(\mathbf{n}, \mathbf{n} + S_{.j})$ in Eq. (20) and the rate $\Omega\hat{f}_j(\mathbf{n} - S_{.j}, \Omega)$ for the system's transition from state $\mathbf{n} - S_{.j}$ to state $\mathbf{n}$ with $W(\mathbf{n} - S_{.j}, \mathbf{n})$. This gives the time evolution of the probability for any state $\mathbf{n}$ as

$$\frac{dP(\mathbf{n}, t)}{dt} = \Omega\sum_{j=1}^{R}(\mathbb{E}^{-S_j} - 1)\tilde{f}_j(\mathbf{n}, \Omega)P(\mathbf{n}, t) \,. \quad (21)$$

$\mathbb{E}^{-S_j}$ is a step operator defined from $\mathbb{E}^{-S_j}g(\mathbf{n}) = g(\mathbf{n} - S_{.j})$, where g is an arbitrary function of the state.

The delineated process is a special case of Markov process known as birth-death process with multi variables. Our model is a single variable of this type defined by two biochemical reactions with one chemical component $\mathbf{X}$ in a well-stirred distributed system with volume Ω. The state of the system is

defined by the number of molecules of species **X**, $\{N\}$ which shows the current number of molecules in the system. The creation and annihilation of component **X** take place by the following reactions, respectively:

$$\emptyset \xrightarrow{k_1} \mathbf{X}\,, \qquad \mathbf{X} \xrightarrow{k_2} \emptyset \tag{22}$$

where, k_1 and k_2 are rate constants. When the left reaction takes place the number of molecules **X** changes from N to $N+1$ molecules and when the right reaction happens, it changes from N to $N-1$. The integers 1 and -1 are the elements of stoichiometry vector $S = [1, -1]$. The probability that the first or second reaction happens in a small time interval δt is given by $p_i = \Omega \tilde{f}_i(n, \Omega)\delta t$, $i = 1, 2$. Where $\tilde{f}_i(n, \Omega)$ is the rate transition for each reaction or, equivalently, the probability of reaction per unit time and unit volume. The mesoscopic transition rate is approximated by its macroscopic counterpart $f_i(x) = \lim_{\Omega\to\infty} \tilde{f}_i(X = \Omega x, \Omega)$, where x and X are respectively the average concentration and number of molecules of species **X**. So from Eq.(8), $f(x)$ the governing law can be split into two backward and forward rate equations, $f_1 = \sigma_E + \alpha(\frac{x^n}{1+x^n})$ governing the rate of creation and $f_2 = x$ governing the rate of annihilation. Therefore, the corresponding transition rates are:

$$\tilde{f}_1 = (\sigma_E + \alpha\frac{X^2}{1+X^2})\,, \tag{23a}$$

$$\tilde{f}_2 = X\,. \tag{23b}$$

The Master equation for the problem can be written as:

$$\frac{dP(X,t)}{dt} = \Omega \sum_{j=1}^{2} (\mathbb{E}^{-S_j} - 1)\tilde{f}_j(X, \Omega)P(X, t)\,. \tag{24}$$

$\mathbb{E}^{-S_j}$ is a step operator which removes S_j molecules from component **X**.

3. Analytical solution to CME [13]

Taylor Expansion of Master Equation: When the step operator $\mathbb{E}^{-S_j}$ in the master Eq. (21) operates on a function like $g_j(\mathbf{n}) = \tilde{f}_j(\mathbf{n}, \Omega)P(\mathbf{n}, t)$, the function is evaluated in a shifted state from **n** to $\mathbf{n} - S_{.j}$. Now if the displacement is small and the function varies smoothly, the displaced function may be approximated by a Taylor expansion around the state **n**. For a general function $g(\mathbf{n})$ such a Taylor expansion looks like:

$$\mathbb{E}^{S_{.j}} g(\mathbf{n}) \approx \Big[1 - \sum_i S_{ij}\frac{\partial}{\partial n_i} + \frac{1}{2}\sum_{i,k} S_{ij}S_{kj}\frac{\partial^2}{\partial n_i \partial n_k} + \dots\Big] g(\mathbf{n})\,. \tag{25}$$

Insertion of Eq. (25) in the Master Eq. (21) followed by truncation after the second order term leads to the Fokker-Planck (FP) approximation of the Master Equation:

$$\frac{dP(\mathbf{n},t)}{dt} = \Omega \sum_{j=1}^{R}\Big(- \sum_i S_{ij}\frac{\partial \tilde{f}_j(\mathbf{n},\Omega)P(\mathbf{n},t)}{\partial n_i} + \frac{1}{2}\sum_{i,k} S_{ij}S_{kj}\frac{\partial^2 \tilde{f}_j(\mathbf{n},\Omega)P(\mathbf{n},t)}{\partial n_i \partial n_k}\Big)\,. \tag{26}$$

For complicated reaction schemes, the FP-equation is almost as complicated to work with as the original master equation.

The Fokker-Planck approximation is very exact when the jumps in state space are infinitely small, but sizes of jumps in the state space of chemical reactions are fixed in size. A solution to this problem is offered by the Ω expansion method [12], where Ω is the volume of the system. The method is known as Linear Noise Approximation (LNA). LNA is a Taylor expansion of CME in powers of $\Omega^{1/2}$ around the steady states of the macroscopic rate equation. The rationale behind this approach is that Ω is the parameter in system that governs the size of fluctuations and therefore the size of jumps. In the size expansion, a new stochastic variable ξ_i is defined from the relation $n_i = \Omega x_i + \Omega^{1/2}\xi_i$, where n_i is the copy number of component i, x_i is a deterministic function of time. In LNA the kinetics of the system is described in the limit of an infinitely large, well stirred volume. By this assumption, the stochastic fluctuation in the state vector **n** is negligible, so the state can be approximated by macroscopic average concentration $\mathbf{x} = [x_1, \dots, x_N]$ and each $\tilde{f}_j(\mathbf{n}, \Omega)$ by its macroscopic rate law counterpart $f_j(\mathbf{x})$. The deterministic rate equation $\dot{\mathbf{x}} = \mathbf{S}\mathbf{f}(\mathbf{x})$ governs the time evolution of the macroscopic concentration vector **x** where **S** is the stoichiometric matrix and $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_R(\mathbf{x})]$. The probability distribution $P(\mathbf{n}, t)$ for $\mathbf{n} = (n_1, \dots, n_N) = \Omega\mathbf{x}$ is related to the probability distribution $\Pi(\xi, t)$ for $\xi = (\xi_1, \dots, \xi_N)$ through

$$P(\mathbf{n}, t) = P(\Omega\phi + \Omega^{1/2}\xi, t) = \Pi(\xi, t) \tag{27}$$

For the probability density function $\Pi(\xi, t)$, the fluctuations are characterized by the linear Fokker-Planck equations: Differentiating Eq. (27) with respect to time at constant molecule numbers gives:

$$\frac{\partial P(\mathbf{n},t)}{\partial t} = \frac{\partial \Pi(\xi,t)}{\partial t} + \sum_{i=1}^{N}\frac{\partial \xi_i}{\partial t}\frac{\partial \Pi(\xi,t)}{\partial \xi_i} = \frac{\partial \Pi(\xi,t)}{\partial t} - \Omega^{1/2}\sum_{i=1}^{N}\frac{\partial \varphi_i}{\partial t}\frac{\partial \Pi(\xi,t)}{\partial \xi_i}$$

$\partial n_i/\partial t = 0$ implies that $\partial \xi_i/\partial t = -\Omega^{1/2}\partial\varphi_i/\partial t$.

Taylor expansion of the transition rates $\tilde{f}_j(\mathbf{n})$ around the macroscopic value $f_j(\phi)$ gives:

$$\tilde{f}_j(\mathbf{x}) = \tilde{f}_j(\phi + \Omega^{-1/2}\xi) = f_j(\phi) + \Omega^{-1/2}\sum_{i=1}^{N}\frac{\partial f_j(\phi)}{\partial \varphi_i}\xi_i + O(\Omega^{-1}) \,. \tag{28}$$

Taylor expansion of the step operator around the steady states gives the approximation:

$$\mathbb{E}^{-S.j} \approx 1 - \Omega^{-1/2}\sum_i S_{ij}\frac{\partial}{\partial \xi_i} + \frac{\Omega^{-1}}{2}\sum_i\sum_k S_{ij}S_{kj}\frac{\partial^2}{\partial \xi_i \partial \xi_k} + \dots \tag{29}$$

Inserting Eqs. (27), (28) and (29) into CME yields:

$$\frac{\partial \Pi(\xi,t)}{\partial t} - \Omega^{1/2}\sum_{i=1}^{N}\frac{\partial \varphi_i}{\partial t}\frac{\partial \Pi(\xi,t)}{\partial \xi_i} =$$
$$\Omega\sum_{j=1}^{R}\Big(-\Omega^{-1/2}\sum_i S_{ij}\frac{\partial}{\partial \xi_i} +$$
$$\frac{\Omega^{-1}}{2}\sum_i\sum_k S_{ij}S_{kj}\frac{\partial^2}{\partial \xi_i \partial \xi_k} + \dots\Big)\times$$
$$\Big(f_j(\phi) + \Omega^{-1/2}\sum_i \frac{\partial f_j(\phi)}{\partial \varphi_i}\xi_i + \dots\Big)\Pi(\xi,t) \,.$$

Identifying terms of order Ω^0:

$$\frac{\partial \Pi(\xi,t)}{\partial t} = -\sum_{i,k} A_{ik}\frac{\partial(\xi_k \Pi)}{\partial \xi_i} + \frac{1}{2}\sum_{i,k} V_{ik}\frac{\partial^2 \Pi}{\partial \xi_i \partial \xi_k} \,, \tag{30}$$

in which,

$$A_{ik} = \sum_{j=1}^{R} S_{ij}\frac{\partial f_i}{\partial x_k}, \quad V_{ik} = \sum_{j=1}^{R} S_{ij}S_{kj}f_j(\mathbf{x}) \,.$$

The matrix **A** is the Jacobian and **V** is the diffusion matrix evaluated in the state $\mathbf{x}(t)$ as determined by differential equation $\dot{\mathbf{x}} = \mathbf{Sf(x)}$.

Eq. (30) is the Fokker-Plank equation for the probability density function $\Pi(\xi,t)$. Its stationary solution is known, a Gaussian with zero average and the covariance **C** of the fluctuations around the average that is given by the Lyapunov equation:

$$\mathbf{AC} + \mathbf{CA}^T + \mathbf{V} = 0 \,. \tag{31}$$

The covariance **C** are large if the fluctuations **V** are large and the eigenvalues of the Jacobian **A** are small. Large fluctuation means that many random events occur per time unit and small eigenvalues of **A** means that the forces that bring the system back to steady state are weak. In our case with one variable,

$$C_X = \frac{V}{-2A} \,. \tag{32}$$

For the switch the Jacobian and Diffusion matrix are:

$$\mathbf{A} = \sum_{j=1}^{2} S_j\frac{\partial f_j}{\partial x_*} = \frac{\partial \mathbf{S.f}}{\partial x_*} = -1 + \alpha\frac{2x_*}{(1+x_*^2)^2} \tag{33a}$$

$$\mathbf{V} = \sum_{j=1}^{2} S_j S_j f_j(x_*) = (-1)^2(x_*) + (1)^2(\alpha\frac{x_*^2}{1+x_*^2} + \sigma_E) \tag{33b}$$

where x_* is the steady state of Eq.(8),

$$f(x_*) = 0 \quad \Rightarrow \quad x_* = \sigma_E + \alpha(\frac{x_*^2}{1+x_*^2})$$

therefore the variance of X is:

$$C_X = \frac{< X >}{[1 - \alpha\frac{2x_*}{(1+x_*^2)^2}]} \,. \tag{34}$$

Define a new variable $\phi = -\alpha(\Phi')$ to simplify variance yields:

$$C_X = < X > \frac{1}{1+\phi} \,. \tag{35}$$

For small x when x is e.g. in the lower state, $x_* \approx \sigma_E$

$$C_X \approx \frac{< X >}{[1 - \alpha x_*]} \approx \frac{\Omega\sigma_E}{1 - 2\alpha\sigma_E} \,.$$

For the variance to be defined, $\phi > -1$. Consequently, one can calculate the Fano factor,

$$F = \frac{C_X}{< X >} = \frac{1}{1+\phi} \,. \tag{36}$$

4. Delayed Chemical Master Equation [15]

In a general form a system with R non-delayed reactions, and D delayed reactions has the following time evolution of its states:

$$\frac{\partial p(\mathbf{n},t)}{\partial t} = \sum_{r=1}^{R}\tilde{f}_r(\mathbf{n}-\nu_r)p(\mathbf{n}-\nu_r,t) - \sum_{r=1}^{R}\tilde{f}_r(\mathbf{n})p(\mathbf{n},t)$$
$$+ \sum_{d=1}^{D}\sum_{\tilde{\mathbf{n}}}\tilde{g}_d(\tilde{\mathbf{n}})p(\mathbf{n}-\upsilon_d,t;\tilde{\mathbf{n}},t-\tau_d)$$
$$- \sum_{d=1}^{D}\sum_{\tilde{\mathbf{n}}}\tilde{g}_d(\tilde{\mathbf{n}})p(\mathbf{n},t;\tilde{\mathbf{n}},t-\tau_d) \,. \tag{37}$$

The system's state $\mathbf{n}$ is a time dependent vector of length equal to components' total number and the i'th component n_i denotes the number of molecules of i'th species. The functions $\tilde{f}_r$ are the transition probability function of the R non-delayed reactions with state change vector ν_r, $(r = 1, \ldots, R)$. The functions $\tilde{g}_d$ are the transition probability functions of the D delayed reactions with state change vector υ_d, $(d = 1, \ldots, D)$. The stoichiometry matrix S here is the matrix with columns of ν's and υ's.

The joint probability $p(\mathbf{n}, t; \tilde{\mathbf{n}}, t - \tau_d)$ can be defined:

$$\begin{aligned} p(\mathbf{n}, t; \tilde{\mathbf{n}}, t - \tau_d) &= p(\mathbf{n}, t|\tilde{\mathbf{n}}, t - \tau_d) p(\tilde{\mathbf{n}}, t - \tau_d) \\ &= p(\tilde{\mathbf{n}}, t - \tau_d|\mathbf{n}, t) p(\mathbf{n}, t) \, . \end{aligned} \tag{38}$$

Inserting equality (38) into statement (37), yields:

$$\begin{aligned} \frac{\partial p(\mathbf{n}, t)}{\partial t} &= \sum_{r=1}^{R} \tilde{f}_r(\mathbf{n} - \nu_r) p(\mathbf{n} - \nu_r, t) - \sum_{r=1}^{R} \tilde{f}_r(\mathbf{n}) p(\mathbf{n}, t) \\ &+ \sum_{d=1}^{D} \sum_{\tilde{\mathbf{n}}} \tilde{g}_d(\tilde{\mathbf{n}}) p(\tilde{\mathbf{n}}, t - \tau_d|\mathbf{n} - \upsilon_d, t) p(\mathbf{n} - \upsilon_d, t) \\ &- \sum_{d=1}^{D} \sum_{\tilde{\mathbf{n}}} \tilde{g}_d(\tilde{\mathbf{n}}) p(\tilde{\mathbf{n}}, t - \tau_d|\mathbf{n}, t) p(\mathbf{n}, t) \end{aligned}$$

the inner summation can be summarized into one function summed over $\tilde{\mathbf{n}}$ as h_d:

$$\begin{aligned} \frac{\partial p(\mathbf{n}, t)}{\partial t} &= \sum_{r=1}^{R} \tilde{f}_r(\mathbf{n} - \nu_r) p(\mathbf{n} - \nu_r, t) - \sum_{r=1}^{R} \tilde{f}_r(\mathbf{n}) p(\mathbf{n}, t) \\ &+ \sum_{d=1}^{D} h_d(\tilde{\mathbf{n}} - \upsilon_2, t) p(\tilde{\mathbf{n}} - \upsilon_d, t) - \sum_{d=1}^{D} h_d(\tilde{\mathbf{n}}, t) p(\mathbf{n}, t) \, . \end{aligned}$$

And the associated Master Equation for the delayed process in our model is:

$$\begin{aligned} \frac{dP(X, t)}{dt} &= [\tilde{f}_1(X - 1) P(X - 1, t) - \tilde{f}_1(X) P(X, t)] \\ &+ [h_1(X + 1, t) P(X + 1, t) - h_1(X, t) P(X, t)] \, . \end{aligned}$$

C. Numerical Methods

1. Numerical solution to the non-delayed CME

A simple and very useful way to estimate the properties of a master equation is to simulate realizations of the corresponding Markov process using Monte Carlo (MC) methods, known as the Stochastic Simulation Algorithm (SSA). The first well-established SSA was done by Gillespie [20], he suggested a method to simulate trajectories of chemical reactions modeled by a discrete Markov process in continuous time. Although the algorithm is fully equivalent to the master equation, it should not be considered the numerical solution to it.

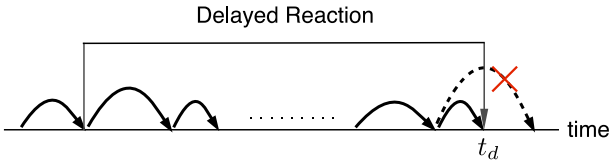

FIG. 5 Illustration to modified Gillespie Algorithm of delayed reactions

The direct method is based on sampling two random numbers, the first one determines the next time, $(t+\tau)$, at which the next reaction takes place and the second one determines the probable reaction j. That is, for a system in state $\mathbf{n}$ at time t, the probability that the next reaction event occurs, $\delta\tau p(\tau, j|\mathbf{n}, i)$ in the time interval $(t + \tau, t + \tau + \delta\tau)$ and is of type j is given by

$$\begin{aligned} p(\tau, j|\mathbf{n}, t) &= \underbrace{a(\mathbf{n}) e^{-a(\mathbf{n})\tau}}_{I} \underbrace{\Omega \tilde{f}_j(\mathbf{n}, \Omega)/a(\mathbf{n})}_{II} \\ &= \Omega \tilde{f}_j(\mathbf{n}, \Omega) e^{-a(x)\tau} \end{aligned} \tag{39}$$

where $a(\mathbf{n}) = \Omega \sum_{j=1}^{R} \tilde{f}_j(\mathbf{n}, \Omega)$. The factor I is the probability density for the time $t + \tau$ of any next reaction given that the system was in state $\mathbf{n}$ at time t. The factor II is the probability that the reaction is of type j.

In our model there are two reactions, $R = 2$, and the state of the system is defined by one varibale, X.

$$p(\tau, j|X, t) = a(X) e^{-a(X)\tau} \Omega \tilde{f}_j(X, \Omega)/a(X) = \Omega \tilde{f}_j(X, \Omega) e^{-a(x)\tau}$$

where $a(X) = \Omega \sum_{j=1}^{2} \tilde{f}_j(X, \Omega)$, which is the sum of propensity functions.

2. Numerical solution to delayed CME

The Delay Stochastic Simulation Algorithm (DSSA) is different from its non-delay algorithm counterpart SSA in the sense that the chosen reaction maybe one of the delayed reactions. Consider a system with N components reacting through $R + D$ reactions among which R reactions are non-delayed and D reactions are delayed. When the next time step is determined to be t^* and the chosen reaction is delayed, so non-Markovian, the reaction will be completed as time advances to $t^* + \tau$, but if the selected reaction is non-delayed one the time of the next reaction t^* is compared with the time of previously scheduled delayed reactions. If none of those are to occur before t^* the time advances to t^* and the number of molecules changes according to the chosen non-delayed reaction. If there is a delayed reaction scheduled to occur at some time $t_d < t^*$ then the selected time step t^* is ignored and time advances to t_d and the scheduled delayed reaction completed with state change according to the scheduled reaction. Fig. 5 is the schema of the algorithm [21].

REFERENCES

[1] L. Wang, B.L. Walker, S. Iannaccone, D. Bhatt, P.J. Kennedy, and W.T. Tse. Bistable switches control memory and plasticity in cellular differentiation. *Proceedings of the National Academy of Sciences*, 106(16):6638, 2009.

[2] M.N. Steijaert, H.M.M. ten Eikelder, A.M.L. Liekens, D. Bošnacki, and P.A.J. Hilbers. Stochastic switching behavior of a bistable auto-phosphorylation network. *transition*, 10:5.

[3] N.I. Markevich, J.B. Hoek, and B.N. Kholodenko. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *The Journal of cell biology*, 164(3):353, 2004.

[4] G. Craciun, Y. Tang, and M. Feinberg. Understanding bistability in complex enzyme-driven reaction networks. *Proceedings of the National Academy of Sciences*, 103(23):8697, 2006.

[5] T. Eissing, H. Conzelmann, E.D. Gilles, F. Allgöwer, E. Bullinger, and P. Scheurich. Bistability analyses of a caspase activation model for receptor-induced apoptosis. *Journal of Biological Chemistry*, 279(35):36892, 2004.

[6] E.M. Ozbudak, M. Thattai, H.N. Lim, B.I. Shraiman, and A. Van Oudenaarden. Multistability in the lactose utilization network of escherichia coli. *Nature*, 427(6976):737–740, 2004.

[7] S.D.M. Santos and J.E. Ferrell. Systems biology: On the cell cycle and its switches. *Nature*, 454(7202):288–289, 2008.

[8] W. Sha, J. Moore, K. Chen, A.D. Lassaletta, C.S. Yi, J.J. Tyson, and J.C. Sible. Hysteresis drives cell-cycle transitions in xenopus laevis egg extracts. *Proceedings of the National Academy of Sciences of the United States of America*, 100(3):975, 2003.

[9] N. Ramakrishnan, U.S. Bhalla, and H.M. Sauro. Memory switches in chemical reaction space. *PLoS Computational Biology*, 4(7), 2008.

[10] D.T. Gillespie. Stochastic simulation of chemical kinetics. *Annual review of physical chemistry*, 58(1):35–55, 2007.

[11] T. Jahnke and W. Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of mathematical biology*, 54(1):1–26, 2007.

[12] N.G. van Kampen. *Stochastic processes in physics and chemistry*. North Holland, 1992.

[13] J. Elf and M. Ehrenberg. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome research*, 13(11):2475, 2003.

[14] M. Barrio, K. Burrage, A. Leier, and T. Tian. Oscillatory regulation of hes1: discrete stochastic delay modelling and simulation. *PLoS computational biology*, 2(9):1017–1030, 2006.

[15] A. Grönlund, P. Lötstedt, and J. Elf. Costs and constraints from time-delayed feedback in small gene regulatory motifs. *Proceedings of the National Academy of Sciences*, 107(18):8171, 2010.

[16] A. Grönlund, P. Lötstedt, and J. Elf. Delay-induced anomalous fluctuations in intracellular regulation. *Nature Communications*, 2:419, 2011.

[17] L.A. Segel. Simplification and scaling. *SIAM review*, 14(4):547–571, 1972.

[18] C.W. Gardiner. *Handbook of stochastic methods*. Springer Berlin, 1985.

[19] J. Elf, J. Paulsson, O. Berg, and M. Ehrenberg. Mesoscopic kinetics and its applications in protein synthesis. *Systems Biology*, pages 13–17, 2005.

[20] J.H. Gillespie. A general model to account for enzyme variation in natural populations. iii. multiple alleles. *Evolution*, 31(1):85–90, 1977.

[21] D. Bratsun, D. Volfson, L.S. Tsimring, and J. Hasty. Delay-induced stochastic oscillations in gene regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 102(41):14593, 2005.