J Biol Phys (2015) 41:235–246
DOI 10.1007/s10867-015-9375-2

ORIGINAL PAPER

Modeling the effect of transcriptional noise on switching in gene networks in a genetic bistable switch

Srabanti Chaudhury

Received: 22 August 2014 / Accepted: 1 January 2015 / Published online: 1 March 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Gene regulatory networks in cells allow transitions between gene expression states under the influence of both intrinsic and extrinsic noise. Here we introduce a new theoretical method to study the dynamics of switching in a two-state gene expression model with positive feedback by explicitly accounting for the transcriptional noise. Within this theoretical framework, we employ a semi-classical path integral technique to calculate the mean switching time starting from either an active or inactive promoter state. Our analytical predictions are in good agreement with Monte Carlo simulations and experimental observations.

Keywords Gene expression · mRNA noise · Semi classical path integral

1 Introduction

Variability of gene expression can be used to develop a quantitative understanding of the underlying gene regulation [1]. Single-cell observations and stochastic analyses have revealed qualitative and quantitative features of gene regulation in both prokaryotes [2] and eukaryotes [3]. Biochemical noise drives stochastic switching between bistable states in genetic networks [4]. Such switching can hence lead to different phenotypic changes for a cellular population. Theoretical modeling of stochastic gene expression [5–7] has been a subject of intense studies in the last few years and has helped us to develop our understanding of the effect of noise on gene regulation. Such models mostly focus on describing intrinsic fluctuations that arise from randomness in time of individual chemical reactions. Theoretical methods such as generating functions [6–8], Langevin and Fokker-Planck equations [9], linear noise approximation [10], many-body theory [11], as well as stochastic simulations using the Gillespie algorithm [12] have been used to study such models of gene expression. Shahrezaei and Swain [6] developed an analytical theory of gene expression to calculate the steady-state protein distributions for a three-stage model where the gene containing the promoter fluctuates between active and inactive states at a constant rate. By considering the protein burst model, which is based under the assumption that mRNA lifetimes are shorter than protein lifetimes, they could derive the master equation for proteins, which implicitly includes mRNA fluctuations. Recently, it has

S. Chaudhury (✉)
Department of Chemistry, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pune 411008, India
e-mail: srabanti@iiserpune.ac.in

been shown that including mRNA noise in master equations plays a strong role in determining switching rates [13, 14]. Employing the WKB theory [15] to treat the underlying chemical master equations, Assaf et al. [16] provided a theoretical framework to obtain the quasi-stationary probability distributions of mRNA/protein copy numbers for a genetic network with positive feedback and also to calculate the mean switching time from either of these active/inactive states.

In this work, we present a different analytical method to investigate the stability of genetic switches in gene expression. We apply this method to a simple gene expression model where the two-state promoter switches between the active and inactive states and the switching rate depends on the transcription factor, which in this case is the protein. We calculate the noise characteristics for the mRNA/protein distribution as well as the switching time between the active and inactive states of the promoter. Lastly, we extend our calculations to noise in more complex promoter architectures by introducing more than two promoter states using our proposed theoretical recipe.

2 Methods

We start by considering a simple two-state gene expression model as shown in Fig. 1a. In this model, the transition between the inactive and active states of the promoter gene is no longer constant but is controlled by protein number n via positive feedback, the protein thereby inducing its own expression. This two-state model with positive feedback is identical to the model treated by Assaf et al. [16]. This two-state positive feedback switch has been shown to describe biological switching experimentally [17]. The transition rates between active and inactive states are $k_{\mathrm{on}}(n) \equiv f(n)$, $k_{\mathrm{off}}(n) \equiv g(n)$ where n is the protein copy number. $f(n)$ and $g(n)$ are considered to be Hill-type functions, $f(n) = k_0^{\min} + \left(k_0^{\max} - k_0^{\min}\right) n^{h_1} / \left(n_{50}^{h_1} + n^{h_1}\right)$ and $g(n) = k_1^{\max} - \left(k_1^{\max} - k_1^{\min}\right) n^{h_2} / \left(n_{50}^{h_2} + n^{h_2}\right)$ where n_{50} is the curve's midpoint and $h_1 = h_2 = h$. To account for the stochastic behavior of the genetic switches, Assaf et al. [16] used this simple model and employed two coupled chemical master equations that can describe the probability distribution functions, $\dot{P}_{m,n}$ and $\dot{Q}_{m,n}$ of having m mRNAs and n proteins at time t with the promoter DNA being in the inactive and active state, respectively. They used the WKB approximation method [18] to obtain a quasi-stationary solution to the master equation. The WKB ansatz [15] is given as $P_{m,n} \simeq \exp[-S(m,n)]$, where S is the action. This ansatz has been used extensively to study population switches between metastable states [19–21]. Then, by considering stationary distributions $\dot{P}_{m,n} = \dot{Q}_{m,n} = 0$ and employing the WKB ansatz they arrived at the stationary Hamiltonian Jacobi equation, $H = H\left(m, \frac{\partial S}{\partial m}, n, \frac{\partial S}{\partial n}\right) = 0$. Here we can draw an analogy to classical mechanics and introduce momenta and $\chi_n = \frac{\partial S}{\partial n}$ such that $\dot{m} = \frac{\partial H}{\partial \chi_m}$, $\dot{n} = \frac{\partial H}{\partial \chi_n}$, $\dot{\chi_m} = -\frac{\partial H}{\partial m}$, $\dot{\chi_n} = -\frac{\partial H}{\partial n}$.

Assuming that the mRNA lifetimes are much shorter than the protein degradation rates ($\gamma >> 1$), one can eliminate the mRNA degrees of freedom [22] $\left(\dot{m}, \dot{\chi}_m = 0\right)$ and obtain a reduced rather effective Hamiltonian, $H = H(n, \chi_n)$, which describes the evolution of slow variables only. Such a kind of separation of variables has been applied earlier in biochemical networks consisting of reactions operating at multiple time scales [23]. Time-scale separation between variables allows one to coarse-grain the system by integrating over fast degrees of freedom without loss of much information about mesoscopic fluctuations of the slow variables. For example, in the context of modeling viral infection kinetics, we eliminated the virus-related variables and calculated the extinction rate by considering the virus clearance rate to be faster than the infected cell dynamics [24].

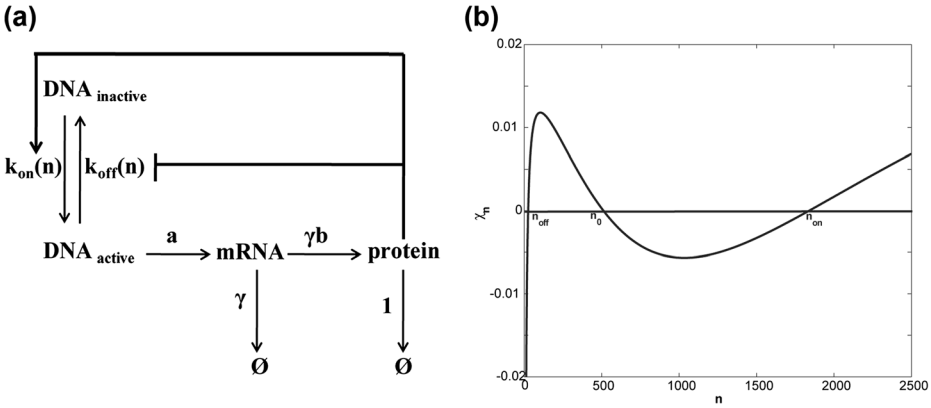

Fig. 1 **a** Model for the positive feedback network derived from [16]. Transcription, translation, mRNA decay, and protein degradation are modeled as first-order processes with rates a, γb, γ, and 1, respectively (rates are rescaled by the protein decay rate). The feedback functions $k_{on}(n)$ and $k_{off}(n)$ regulate promoter transitions. **b** χ_n versus n showing bistability at ab=2,400, b=22.5, h=2, n_{50}=1,000, $k_0^{min}=k_1^{min}=a/100$, $k_0^{max}=k_1^{max}=a$

In our present study, we consider the same model for the positive feedback network as discussed in [16] but propose a different theoretical method that employs the stochastic semi-classical path integral technique to explicitly account for the mRNA noise along with protein fluctuations. This semi-classical method has been used earlier to calculate rare event statistics in reaction diffusion systems [25] and it has been applied to various epidemiological stochastic models [24, 26, 27]. In this method, we start with the eikonal ansatz:

$$Z(\chi,t) = e^{S(\chi,t)} = \sum_{n=0}^{\infty} P_n(t) e^{in\chi} \tag{1}$$

where Z is the generating function for the number of molecules n that are formed during a chemical process in time t. χ is a variable conjugated to n number of molecules and appears as a 'Lagrange multiplier' that takes care of conservation laws during the derivation of path integrals (see supporting information Text S1 of [24]). This eikonal approximation is used to solve master equations and it can recast the problem in terms of an effective classical Hamiltonian system. Here we do not provide the derivation of the semi-classical path integral approach and refer the reader to Text S1 in [24].

By considering the trajectories that dominate the dynamics, it was found that $S(\chi,t)$ is given by [25]:

$$S(\chi_C,t) = \int_0^{t_f} \left(\dot{\chi}(n) + H(n,\chi)\right) dt \tag{2}$$

where H is the Hamiltonian that can be obtained following the general method discussed in [24]. Such a path integral representation has been derived for the case of a Michaelis-Menten enzyme attached to the membrane of a eukaryotic cell and later generalized to a network of reactions [23]. This theory is valid in the limit of short-lived mRNA [16–28]. Our first step is to consider the process of mRNA generation. Let P_m^G and P_m^{G*} be the probabilities to generate m mRNA molecules at time t in the inactive and active state of the promoter, respectively. The master equations for their evolution are [16]:

$$\begin{aligned} \frac{d}{dt}P_m^G &= -\ k_{on}(n)\ P_m^G\ +\ k_{off}(n)\ P_m^{G*} \\ \frac{d}{dt}P_m^{G*} &= k_{on}(n)P_m^G - \left(k_{off}(n)+a\right)P_m^{G*} + aP_{m-1}^{G*} \end{aligned} \tag{3}$$

Assuming that the mRNA is short-lived compared to the protein, we can solve these master equations by introducing the generating function:

$$Z^g(\chi_R) = \sum_{m=0}^{\infty} P_m^g e^{i\chi_R m} \tag{4}$$

with $g \in \{G, G^*\}$ and the total generating function is $Z(\chi_R)=Z_G+Z_{G^*}$. The cumulants are given by:

$$\begin{aligned} \langle m\rangle &= -i\frac{\partial \ln Z(\chi_R)}{\partial \chi_R}\bigg|_{\chi_R=0}, \sigma_m^2 = \langle m^2\rangle - \langle m\rangle^2 = (-i)^2\frac{\partial^2 \ln Z(\chi_R)}{\partial \chi_R^2}\bigg|_{\chi_R=0} \\ F &= \frac{\sigma_m^2}{\langle m\rangle} \end{aligned} \tag{5}$$

The Fano factor F is a ratio of the variance to mean and is a measure of the noise strength due to stochastic switching between different on and off states.

In terms of this generating function, the master equation can be written as [29]:

$$\begin{aligned} \frac{d}{dt}Z^G &= -\ k_{on}(n)\ Z^G\ +\ k_{off}(n)\ Z^{G*} \\ \frac{d}{dt}Z^{G*} &= k_{on}(n)Z^G - \left(k_{off}(n)+a\right)Z^{G*} + ae^{i\chi_R}Z^{G*} \end{aligned} \tag{6}$$

Equation (6) can be considered as a Schrödinger equation:

$$\frac{d}{dt}\begin{pmatrix} Z^G \\ Z^{G*} \end{pmatrix} = \hat{H}(\chi_R)\begin{pmatrix} Z^G \\ Z^{G*} \end{pmatrix} \tag{7}$$

where

$$H(\chi_R) = \begin{pmatrix} -k_{on}(n) & k_{off}(n) \\ k_{on}(n) & a\left(e^{i\chi_R}-1\right)-k_{off}(n) \end{pmatrix}. \tag{8}$$

Assuming that the switching rates between the inactive and active transcriptional state are faster than the time of a substantial change of protein number n or order of $\sqrt{n}$, Eq. (7) can be solved as an eigenvalue problem with:

$$Z(\chi_R) = e^{\lambda_0 t} \tag{9}$$

where λ_0 is the eigenvalue of the least negative real part. By diagonalizing, $|H-\lambda I|=0$, we have:

$$Z(\chi_R) = e^{\frac{t}{2}\left\{K(\chi_R)+\sqrt{K(\chi_R)+4k_{on}H_0(\chi_R)}\right\}}. \tag{10}$$

where $K(\chi_R) = H_0(\chi_R) - \left(k_{on}(n) + k_{off}(n)\right)$

We also assume that when the gene is in the on state, mRNAs are generated according to a Poisson process and hence:

$$H_0(\chi_R) = a\left(e^{i\chi_R}-1\right) \tag{11}$$

Following Eq. (5), for this two-state promoter model, we can compute the first two moments of the mRNA distribution and calculate the Fano factor:

$$\langle m\rangle = \frac{ak_{on}(n)}{\gamma\left(k_{on}(n)+k_{off}(n)\right)}, \sigma_m^2 = \langle m^2\rangle - \langle m\rangle^2 = \langle m\rangle\left(1+\frac{2ak_{off}(n)}{\gamma\left(k_{off}(n)+k_{on}(n)\right)^2}\right) \quad (12)$$
$$F_{mRNA} = 1 + \frac{a}{\gamma}\Delta.$$

where γ is the degradation rate of each mRNA molecule and $\Delta = \frac{2k_{off}(n)}{\left(k_{on}(n)+k_{off}(n)\right)^2}$

This Fano factor is larger than unity and depends on the concentration of the regulatory factor binding to the promoter, which here is the protein copy number n.

The next step is to combine Eq. (10) with the generating function of proteins translated from a single mRNA.

Let $P(\tau)=\gamma e^{-\gamma\tau}$ be the probability for the mRNA to live during time τ. Then the probability $P(n|1_R)$ to have n proteins generated by one mRNA is:

$$P(n|1_R) = \int_0^\infty d\tau P(\tau)p(n|\tau). \quad (13)$$

where $p(n|\tau) = \int d\chi_0 e^{-in\chi_0+\tau H(\chi_0)}$

The total generating function of probabilities to have n proteins generated by any single mRNA is given by:

$$Z(\chi_1|1_R) = \sum_{n=-\infty}^{\infty} P(n|1_R)e^{in\chi_1} = \frac{1}{1-H(\chi_1)/\gamma} \quad (14)$$

where we have performed the integration over τ. Summation over n then produces a delta function $\delta(\chi_1-\chi_0)$, which is then removed by integration over χ_0.

If we assume that all mRNAs generate proteins independently, then the generating function of protein distribution due to m mRNA is $[Z(\chi_1|1_R)]^m$. Then the probability to have n proteins generated during time t, where $t>>\tau$ is given by:

$$P(n,t) = \frac{1}{2\pi}\sum_{m=0}^{\infty}\int_{-\pi}^{\pi} d\chi_1 \int_{-\pi}^{\pi} d\chi_R\, [Z(\chi_1|1_R)]^m e^{-in\chi_1-im\chi_R+tH(\chi_R)}$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} d\chi_1 \int_{-\pi}^{\pi} d\chi_R \frac{1}{1-e^{-i\chi_R}Z(\chi_1|1_R)} e^{in\chi_1+tH(\chi_R)} \quad (15)$$

Then the moment generating function to generate n proteins in time $t>>\tau$ is as follows:

$$Z_p(\chi_n) = \sum_{n=-\infty}^{\infty} P(n|t)e^{in\chi_n}. \quad (16)$$

Substituting Eq. (15) into Eq. (16) and writing $\sum_{n=-\infty}^{\infty} e^{in(\chi_n-\chi_1)}$ as a delta function $\delta(\chi_n-\chi_1)$, we have:

$$Z_p(\chi_n) = \int_{-\pi}^{\pi} d\chi_1 \int_{-\pi}^{\pi} d\chi_R \frac{1}{1-e^{-i\chi_R}Z(\chi_1|1_R)} e^{tH(\chi_R)}\delta(\chi_n-\chi_1) \quad (17)$$

By integration over χ_1 and using properties of the delta function, we have a path integral over the χ_n variable that can be written as:

$$Z_p(\chi_n) = \oint d\frac{\left(e^{i\chi_{\{R\}}}\right) \quad -i}{2\pi \quad e^{-i\chi_R} - Z(\chi_n|1_R)} e^{tH(\chi_R)} = e^{tH(\chi_R)}\left|e^{i\chi R} = Z(\chi_n|1_R).\right. \tag{18}$$

Using the above equations, the Hamiltonian for protein generation and degradation can be written as:

$$H_p(\chi_n) = \frac{1}{2}\left\{K_1(\chi_n) \pm \sqrt{K_1(\chi_n)^2 + 4k_{on}H_1(\chi_n)}\right\} + n\left(e^{-i\chi_n} - 1\right) \tag{19}$$

where $H_1(\chi_n) = \frac{a}{\gamma}\frac{H(\chi_n)}{1-H(\chi_n)/y}$ and $H(\chi_n) = b\gamma(e^{i\chi_n} - 1)$ where b is the burst size, i.e., the average number of proteins translated from one mRNA molecule. It is assumed that generation of a protein from a single mRNA is also a Poisson process.

From Eqs. (5) and (19) and using the relation $Z_p(\chi_n) = e^{tH_p(\chi_n)}$, the Fano factor for protein distribution is given by:

$$F_{protein} = 1 + 2b + b\Delta' \tag{20}$$

where $\Delta' = \frac{a}{\gamma}\Delta$.

The switching between the on and off states occurs along a trajectory in the phase space of the classical Hamiltonian known as the optimal path trajectory. The mean switching time along the Hamiltonian trajectories that start from the metastable state and end either in the off or on state has the form:

$$\tau_{on/off \to off/on} = \sigma \exp\left[S_{on/off \to off/on}\right] \tag{21}$$

where S is known as '*action*' in classical physics. A precise estimation of the prefactor σ is a hard task, although it can be obtained for one-dimensional systems; however it is not important because Eq. (21) is dominated by the exponent, when the system is not too close to the bifurcation point [25, 30]. Strong fluctuations drive the system from its equilibrium metastable state to the on/off states along an optimal path that minimizes the action S. The optimal Hamiltonian trajectory that describes S starts from the metastable state and ends either at the on or off state. Since the probability of extinction is found by minimizing the action, we compute such trajectories satisfying the zero energy condition, i.e., $H_p(\chi_n)=0$ as discussed in [24]. This simplifies our calculations because, from $H_p(\chi_n)=0$, one can find how χ_n depends on n along this trajectory and then the action has the form:

$$S = \int_{n_{off}/n_{on}}^{n_0} \chi_n dn. \tag{22}$$

If n_{on} and n_{off} are the average protein number in the on and off states respectively, then at the metastable state these points are separated by a repeller n_0 such that $n_{\text{off}}<n_0<n_{\text{on}}$ and is obtained from the solution:

$$\dot{n}_0 = \frac{\partial H_P(\chi_n)}{\partial \chi_n}|\chi_n = 0 \tag{23}$$

and χ_n as a function of n along the zero energy trajectory of the Hamiltonian can be obtained at $H_p(\chi_n)=0$ such that:

$$\chi_n = \ln\left[\frac{-B + \sqrt{B^2 - 4AC}}{2A}\right] \tag{24}$$

where $B=-n\{n(1+2b)+ab+(1+b)(k_{\text{off}}(n)+k_{\text{on}}(n))\}$, $A=nbk_{\text{off}}(n)+(ab+bn(n+k_{\text{on}}(n)))$, $C=n^2(1+b)$. Figure 1b shows this trajectory as a function of the protein number n.

The probability of switching from the off to on state and vice versa along this trajectory is given by $P_{\text{on/off}\to\text{off/on}}(n)\sim\exp[-S_{\text{on/off}\to\text{off/on}}(n)]$.

3 Results and discussion

To check our theoretical predictions, we compare our results for MST prediction to Monte Carlo simulations (MC) using the Gillespie algorithm [12]. Figure 2 shows the functional dependence of $\tau_{\text{off}\to\text{on}}$ on the translation rate b at small and large gamma values. The numerical results are fitted with MC simulations using suitable values of the prefactor σ in Eq. (21). The MST has a superexponential dependence on b. In Fig. 2a, the burst size ranges from 1 to 4 for large values of gamma, whereas in Fig. 2b the burst size varies from 10 to 35 for small values of gamma. This suggests that the most important parameter in this theory is bγ instead of γ. This is the reason for obtaining good agreement with the simulation results for small values of gamma because in that case b>>1. Figure 3 shows the MST from on to off states for a high value of gamma and off to on states for a low value of gamma as a function of n_{50}. In Fig. 3a, as n_{50} increases, k_{off} becomes smaller. This indicates that the switching rate from the on to off state decreases. Conversely, in Fig. 3b as n_{50} increases, k_{on} increases and the mean switching time from the off to on state becomes faster. In both Figs. 2 and 3, the numerical results are fitted with MC simulation by adjusting the value of the prefactor σ.

We can use our analytical model to determine a relationship between gene expression and noise, which is defined as the ratio of variance to the square of the mean, $\mu_{mRNA}=\sigma_m^2/\langle m\rangle^2$ or $\mu_n=\sigma_n^2/\langle n\rangle^2$. From Eqs. (3) and (15), we have:

$$\mu_n = \frac{1+2b}{\langle n\rangle} + \frac{b\Delta'}{\langle n\rangle}. \tag{25}$$

The second term in Eq. (25) captures the effect of the promoter architecture. The same form of noise is also obtained for mRNA distributions, $\mu_{mRNA} = \frac{1}{\langle m\rangle} + \frac{\Delta}{\langle m\rangle}$, where both Δ and Δ' have been defined earlier. In Fig. 4, we plot the protein noise μ_n as a function of the gene expression $\langle n\rangle$. We find that for this protein feedback loop where the on-off switching rate is

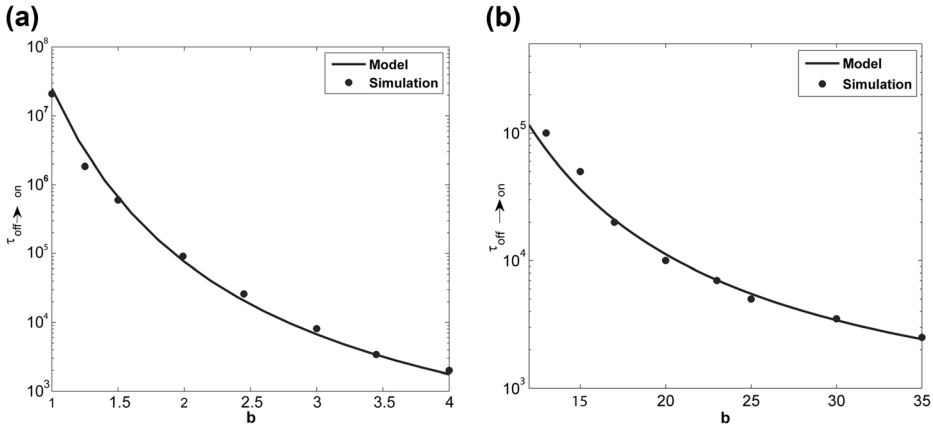

Fig. 2 MST from off to on state as a function of b for (a) γ=50 (b) γ=2. The *solid lines* are the numerical results with preexponents while (●) are from MC simulations. Prefactors are 17.08 and 300 for **a** and **b**, respectively. Here ab=2,400, h=2, $k_0^{\min}=k_1^{\min}=a/100$, $k_0^{\max}=k_1^{\max}=a$. For ,a, n_{50}=720, b n_{50}=1,000

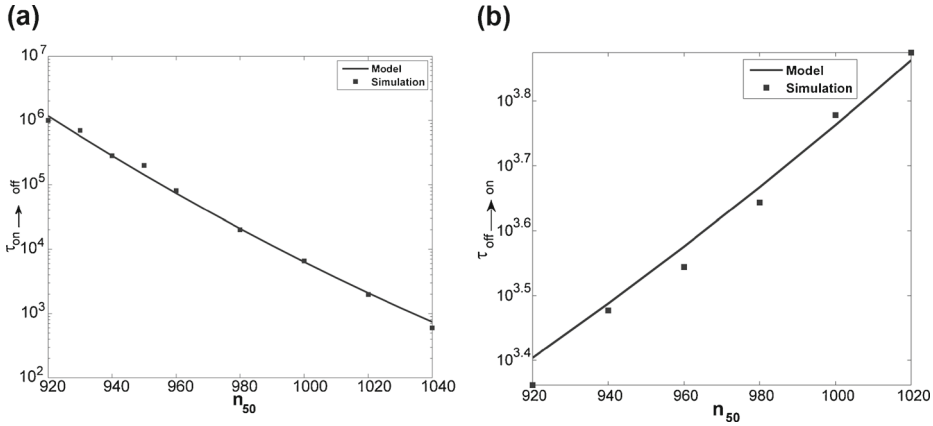

Fig. 3 **a** MST as a function of Hill-type function curve's midpoint n_{50} for **a** on to off state, γ=50, **b** off to on state, γ=2. The *solid lines* are the numerical results with preexponents while (●) are from MC simulations. Prefactors are 4.8 and 230 for **a** and **b**, respectively. Here ab=2,400, h=2, $k_0^{min}=k_1^{min}=a/100$, $k_0^{max}=k_1^{max}=a$. For a, b=15, b b=22.5

controlled by the unbinding and binding of the protein to the promoter region, the expression increases and the noise decreases with the increase in the protein copy number n. This relationship between mean expression and noise in response to binding of the transcription factor (TF) to binding sites containing the promoter has also been observed experimentally by Segal et al. [31]. The gene expression and noise were measured as a function of the TF Zap1, which can act both as an activator and a repressor. For the experimental scenario of a Zap1-

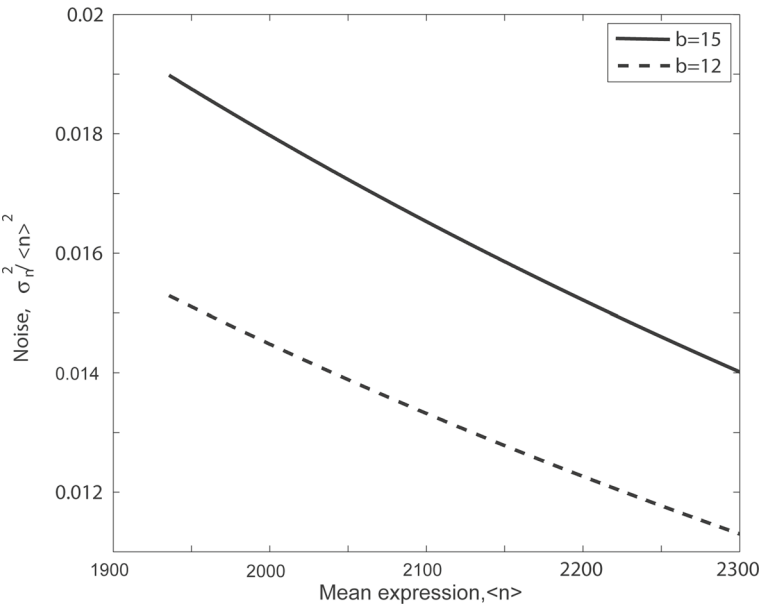

Fig. 4 Mean expression versus noise for the positive feedback model as shown in Fig. 1a for b=15 (*solid line*) and b=12 (*dashed line*), ab=2,400, h=2, $k_0^{min}=k_1^{min}=a/100$, $k_0^{max}=k_1^{max}=a$, n_{50}= *720,* γ=1

activated target gene, we can easily make a direct comparison with our two-state model (Fig. 1a) where the positive protein feedback loop essentially acts as an activator. Our calculations show that having a positive protein feedback loop causes an increase in gene expression along with a decrease of noise. This in fact agrees with the experimental predictions [31] of increased gene expression and decreased noise with the increased binding of TF Zap1 to the promoter region of the gene. The observed parallelism between our calculated results and reported experimental observations clearly implies the practical utility of our proposed theoretical methodology.

To test the robustness of our theoretical framework, we used our theory of stochastic transcriptional regulation to handle further complexity in promoter architectures. Such complexity can be introduced by varying the strength, number, and position of the transcription factor binding sites and also the strength of the repressors or activators. It has been reported earlier both in experiments and in theoretical studies that the promoter architecture of the gene regulatory network affects cell-to-cell variability in gene expression [32–35]. Thermodynamic models of gene expression have been used earlier to calculate the mean mRNA number and the protein copy number per cell [36–38]. However, these simple thermodynamic models can only compute the mean and not the higher moments of protein/mRNA distributions. As shown in the previous section, we can not only use our approach to calculate the mean and the variance of the protein/mRNA distributions but to also obtain the mean switching time between active and inactive promoter states for a complex promoter architecture such as the three-state promoter model [39–41] as shown in Fig. 5. State 1 refers to the inaccessible promoter (repressor bound) that makes a reversible transition to the DNA in the inactive state. The DNA then undergoes stochastic switching between the on and off states. For this birth-death mechanism the master equation for the mRNA probability distribution in terms of generating functions can be written as:

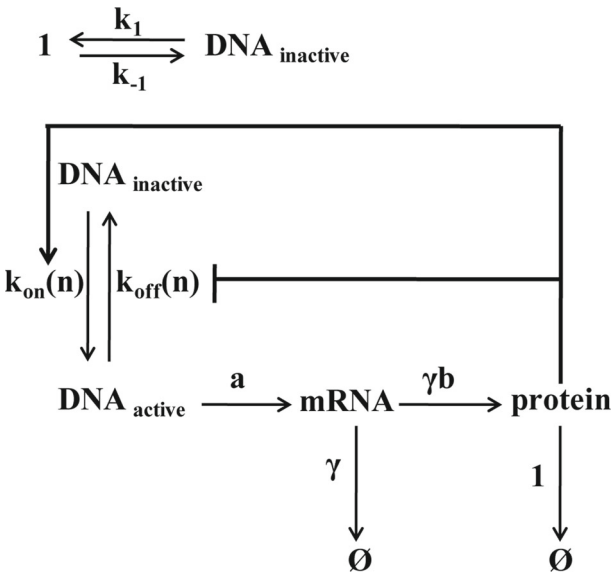

Fig. 5 Three-state promoter model for positive feedback. State 1 is the inaccessible promoter that may be bound to a repressor and k_1 and k_{-1} are the transition rates between state 1 and inactive state of the promoter, DNA_{inactive}. The rest is the same as in Fig. 1a

$$\begin{aligned}\frac{\partial}{\partial t}Z^1 &= -k_1Z^1 + k_{-1}Z^G\\ \frac{\partial}{\partial t}Z^G &= -(k_{-1}+k_{on}(n))Z^G + k_1Z^1 + k_{off}(n)Z^{G*}\\ \frac{\partial}{\partial t}Z^{G*} &= -\left(a+k_{off}(n)\right)Z^{G*} + ae^{i\chi_R}Z^{G*} + k_{on}(n)Z^G\end{aligned} \tag{26}$$

As shown previously in Eq. (6), one can recast this in the form of a Schrödinger equation such that:

$$\frac{d}{dt}\begin{pmatrix} Z^1\\ Z^G\\ Z^{G*}\end{pmatrix} = H(\chi_R)\begin{pmatrix} Z^1\\ Z^G\\ Z^{G*}\end{pmatrix} \tag{27}$$

where

$$H(\chi_R) = \begin{pmatrix} -k_1 & k_{-1} & 0\\ k_1 & -(k_{-1}+k_{on}(n)) & k_{off}(n)\\ 0 & k_{on}(n) & -\left(a+k_{off}(n)-ae^{i\chi_R}\right)\end{pmatrix}$$

Using the same argument used earlier, we can solve this eigenvalue problem and obtain the generating function for mRNA production:

$$Z(\chi_R) = e^{\frac{t\left(K_2(\chi_R)+\sqrt{K_2(\chi_R)+4H_2(\chi_R)}\right)}{2}} \tag{28}$$

where $K_2(\chi_R) = -(k_1+A+B)$, $H_2(\chi_R) = k_{\text{on}}(n)k_{\text{off}}(n) + k_1AB - k_1(A+B+k_{\text{on}}(n)k_{\text{off}}(n)) - AB$ and A and B is given by $A = \left(k_{off}(n) + a(1-e^{i\chi_R})\right)$ and $B=k_{\text{on}}(n)+2k_{-1}$, respectively.

Using Eq. (5) and from Eq. (28) we can compute the Fano factor that is given by:

$$F_{mRNA} = 1 + \frac{a}{\gamma}\Delta_1 \tag{29}$$

where:

$$\Delta_1 = 2\frac{k_{on}(n)/k_{off}(n)}{1+k_{on}(n)/k_{off}(n)+k_{-1}/k_1}\left(\left(\frac{1/k_{on}(n)}{1+k_{on}(n)/k_{off}(n)+k_{-1}/k_1}\right) + \left(\frac{k_{-1}/k_1}{1+k_{on}(n)/k_{off}(n)+k_{-1}/k_1}\right)\left(\frac{1}{k_1} + \frac{2+k_{-1}/k_1}{k_{on}(n)}\right)\right)$$

In the absence of a repressor, this noise is identical to Eq. (10), which corresponds to a two-state promoter model. We find that Δ_1 is a function of the switching rates between different promoter states. This clearly shows that although the expression of the Fano factor has the same structural form, it will be different for different promoter architectures and is a measure of the variability of the gene expression level in different cells. Similar findings were reported in [40] where it was found that DNA-looping that acts as a third transcriptional state changes the bistability properties of this stochastic model of gene expression in the lac operon. Also it was shown in [41] that the transmission of genetic information was affected for generic promoter models with multiple internal states including the one described in [40]. Further, we can use our semi-classical path integral technique to obtain the on-off switching times between different promoter states in this particular three-state promoter model.

4 Conclusions

In this work we presented an analytical method to understand the role of mRNA fluctuations in switching stability. We find that that this semi-classical path integral is a suitable method to

coarse-grain the system and obtain an effective model for protein/mRNA fluctuations. Also, our technique can be used to calculate switching times, in the limit of short-lived mRNA. The considered model is only the minimal model of protein production. However, our theoretical model is quite unique and diverse in the sense that it can be applied to more complex models of gene expression with multiple promoter sites. We illustrated the effective use of our theoretical framework for a complex system, namely the three-state promoter model as shown in Fig. 5. As reported earlier [39–41], our analysis of the three-state promoter model shows that the gene expression level is strongly affected by the promoter architecture. Also, in the two-state promoter model, one can introduce a non-exponential protein degradation step instead of a simple Poisson process. It is practically impossible to capture such variations in the model using previously known theoretical techniques but we show that these can be easily incorporated into our proposed theoretical method. As reported experimentally in the literature, we find that promoter architecture has the same qualitative effect on cell-to-cell variability in the noise and gene expression for both mRNA and protein distributions. In this work we only focus on describing intrinsic noise but, in the high copy-number regime, extrinsic noise that arises from the interactions of the system of interest with its environment is equally significant. Extension of this theory to describe the combined effect of intrinsic and extrinsic noise on the dynamics of stochastic genetic switches is underway.

Acknowledgments The author acknowledges Dr. Nikolai Sinitsyn for his useful comments and suggestions. S.C. acknowledges support from IISER Pune.

References

1. Munsky, B., Nuert, G., van Oudenaarden, A.: Using gene expression noise to understand gene regulation. Science **336**, 183–187 (2012)
2. Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S.: Stochastic gene expression in a single cell. Science **297**, 1183–1186 (2002)
3. Raser, J.M., O'Shea, E.K.: Control of stochasticity in eukaryotic gene expression. Science **304**, 1811–1814 (2004)
4. Dubnau, D., Losick, R.: Bistability in bacteria. Mol. Microbiol. **61**, 562–572 (2006)
5. Paulsson, J.: Models of stochastic gene expression. Phys. Life Rev. **2**, 157–175 (2005)
6. Shahrezaei, V., Swain, P.: The stochastic nature of biochemical networks. Proc. Natl. Acad. Sci. U. S. A. **105**, 17256–17261 (2008)
7. Raj, A., van Oudenaarden, A.: Nature, nurture, or chance: stochastic gene expression and its consequences. Cell **135**, 216–226 (2008)
8. Thattai, M., van Oudenaarden, A.: Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. U. S. A. **98**, 8614–8619 (2001)
9. Hasty, J., Pradines, J., Dolnik, M., Collins, J.J.: Noise-based switches and amplifiers for gene expression. Proc. Natl. Acad. Sci. U. S. A. **97**, 2075–2080 (2000)
10. Paulsson, J.: Summing up the noise in gene networks. Nature **427**, 415–418 (2004)
11. Sasai, M., Wolynes, P.G.: Stochastic gene expression as a many-body problem. Proc. Natl. Acad. Sci. U. S. A. **100**, 2374–2379 (2003)
12. Gillespie, G.T.: Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. **81**, 2340–2361 (1977)
13. Mehta, P., Mukhopadhyay, R., Wingreen, N.S.: Exponential sensitivity of noise-driven switching in genetic networks. Phys. Biol. **5**, 026005 (2005)
14. Zong, C., So, L.-H., Sepúlveda, L.A., Skinner, S.O., Golding, I.: Lysogen stability is determined by the frequency of activity bursts from the fate-determining gene. Mol. Syst. Biol. **6**, 440 (2010)
15. Bender, V.M., Orszag, S.A.: Advanced Mathematical Methods for Scientists and Engineers. Springer (1999)
16. Assaf, M., Roberts, E., Luthey Schulten, Z.: Determining the stability of genetic switches: explicitly accounting for mRNA noise. Phys. Rev. Lett. **106**, 248102 (2011)

17. Choi, P.J., Cai, L., Frieda, K., Xie, X.S.: A stochastic single-molecule event triggers phenotype switching of a bacterial cell. Science **322**, 442–446 (2008)
18. Kubo, R., Matsuo, K., Kitahara, K.: Fluctuation and relaxation of macrovariables. J. Stat. Phys. **9**, 51–96 (1973)
19. Assaf, M., Meerson, B.: Extinction of metastable stochastic populations. Phys. Rev. E **81**, 021116 (2010)
20. Dykman, M.I., Mori, E., Ross, J., Hunt, P.M.: Large fluctuations and optimal paths in chemical kinetics. J. Chem. Phys. **100**, 5735–5750 (1994)
21. Ovaskainen, O., Meerson, B.: Stochastic models of population extinction. Trends Ecol. Evol. **25**, 643–652 (2010)
22. Assaf, M., Meerson, B.: Noise enhanced persistence in a biochemical regulatory network with feedback control. Phys. Rev. Lett. **100**, 058105 (2008)
23. Sinitsyn, N.A., Hengartner, N., Nemenman, I.: Adiabatic coarse-graining and simulations of stochastic biochemical networks. Proc. Natl. Acad. Sci. U. S. A. **106**, 10546–10551 (2009)
24. Chaudhury, S., Perelson, A.S., Sinitsyn, N.A.: Spontaneous clearance of viral infections by mesoscopic fluctuations. PLoS ONE **7**, e38549 (2012)
25. Elgart, V., Kamenev, A.: Rare event statistics in reaction-diffusion systems. Phys. Rev. E **70**, 041106 (2004)
26. Kamenev, A., Meerson, B.: Extinction of an infectious disease: A large fluctuation in a nonequilibrium system. Phys. Rev. E **77**, 061107 (2008)
27. Khasin, M., Dykman, M.I.: Extinction rate fragility in population dynamics. Phys. Rev. Lett. **103**, 068101 (2009)
28. Shahrezaei, V., Swain, P.S.: Analytical distributions for stochastic gene expression. Proc. Natl. Acad. Sci. U. S. A. **105**, 17256 (2008)
29. de Rhonde, W.H., Daniels, B.C., Mugler, A., Sinitsyn, N.A.: Mesoscopic statistical properties of multistep enzyme-mediated reactions. IET Syst. Biol. **3**, 429–437 (2009)
30. Escudero, C., Kamenev, A.: Switching rates of multistep reactions. Phys. Rev. E **79**, 041149 (2009)
31. Carey, L.B., Dijk, D., Sloot, P.M.A., Kaandorp, J.A., Segal, E.: Promoter sequence determines the relationship between expression level and noise. PLOS Biol. **11**, e1001528 (2013)
32. Sanchez, A., Kondev, J.: Transcriptional control of noise in gene expression. Proc. Natl. Acad. Sci. U. S. A. **105**, 5081–5086 (2008)
33. Sanchez, A., Garcia, H.G., Jones, D., Phillips, R., Kondev, J.: Effect of promoter architecture on the cell-to-cell variability in gene expression. PLOS Comput. Biol. **7**, e1001100 (2011)
34. Murphy, K.F., Balázsi, G., Collins, J.J.: Combinatorial promoter design for engineering noisy gene expression. Proc. Natl. Acad. Sci. U. S. A. **104**, 12726–12731 (2007)
35. Blake, W.J., Balázsi, G., Kohanski, M.A., Isaacs, F.J., Murphy, K.F., Kuang, Y., Cantor, C.R., Walt, D.R., Collins, J.J.: Phenotypic consequences of promoter-mediated transcriptional noise. Mol Cell **24**,853–865
36. Vilar, J.M.G., Leibler, S.: DNA looping and physical constraints on transcription regulation. J. Mol. Biol. **331**, 981–989 (2003)
37. Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., Phillips, R.: Transcriptional regulation by the numbers: models. Curr. Op. Genet. Dev. **15**, 116–124 (2005)
38. Shea, M.A., Ackers, G.K.: The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. J. Mol. Biol. **181**, 211–230 (1985)
39. Earnest, T.M., Roberts, E., Assaf, M., Dahmen, K., Luthey Schulten, Z.: DNA looping increases the range of bistability in a stochastic model of the lac genetic switch. Phys. Biol. **10**, 026002 (2013)
40. Newby, J.M.: Isolating intrinsic noise sources in a stochastic genetic switch. Phys. Biol. **9**, 026002 (2012)
41. Rieckh, G., Tkačik, G.: Noise and information transmission in promoters with multiple internal states. Biophys. J. **106**, 1194–1204 (2014)