Stochastic gene expression conditioned on large deviations

Jordan M Horowitz[1] and **Rahul V Kulkarni**[2]

[1]Physics of Living Systems Group, Department of Physics, Massachusetts Institute of Technology, 400 Technology Square, Cambridge, MA 02139

[2]Department of Physics, University of Massachusetts Boston, Boston, MA 02125

Abstract

The intrinsic stochasticity of gene expression can give rise to large fluctuations and rare events that drive phenotypic variation in a population of genetically identical cells. Characterizing the fluctuations that give rise to such rare events motivates the analysis of large deviations in stochastic models of gene expression. Recent developments in non-equilibrium statistical mechanics have led to a framework for analyzing Markovian processes conditioned on rare events and for representing such processes by conditioning-free *driven* Markovian processes. We use this framework, in combination with approaches based on queueing theory, to analyze a general class of stochastic models of gene expression. Modeling gene expression as a Batch Markovian Arrival Process (BMAP), we derive exact analytical results quantifying large deviations of time-integrated random variables such as promoter activity fluctuations. We find that the conditioning-free driven process can also be represented by a BMAP that has the same form as the original process, but with renormalized parameters. The results obtained can be used to quantify the likelihood of large deviations, to characterize system fluctuations conditional on rare events and to identify combinations of model parameters that can give rise to dynamical phase transitions in system dynamics.

In stochastic systems, it is often of great interest to characterize the fluctuations that give rise to rare events. For example, a recurring theme in current biological research is rare events leading to phenotypic variation in clonal cells [1]. Prominent examples include latency in HIV-1 viral infections [2], sporulation in bacteria [3], and reversible drug tolerance [4] in subpopulations of cancer cells. In several cases, the corresponding rare phenotypic transition is primarily driven by the intrinsic stochasticity of gene expression. These observations provide strong motivation for analyzing rare large deviations in stochastic models of gene expression [5, 6].

The development of a framework for analyzing rare events in stochastic gene expression needs to take into account multiple factors. Single-cell experiments indicate complex mechanisms underlying bursting (*i.e.* long periods of inactivity punctuated by shorter periods of activity) in gene expression [7], motivating the study of general stochastic models with multiple promoter states [8, 9]. Furthermore, the random variables whose rare fluctuations are of interest can be diverse, ranging from promoter activity fluctuations [3] to

the fraction of time spent in specific promoter states [2]. These observations motivate the analysis of a general class of gene expression models that can accommodate arbitrary complexity in promoter dynamics and bursting. The development of an analytical framework for rare events in such models can be used to address several questions of current interest: (1) How do combinations of underlying model parameters control the likelihood of rare events? (2) How can we characterize fluctuations in the system *conditioned* on the occurrence of a rare event? (3) Can we determine the changes in dynamical model parameters that mimic the effects of rare fluctuations?

Recent developments in nonequilibrium statistical mechanics using large deviation theory provide a framework for addressing such issues [10, 11, 12]. In particular, considerable interest has focused on deriving conditioning-free Markov processes, termed *driven* processes, whose statistics reproduce the fluctuations of the original Markov process conditioned on the occurrence of a rare event [13, 14]. So far, nontrivial examples which explicitly characterize the driven process for infinite dimensional systems have been limited. Furthermore, it is of interest to determine systems for which the stochastic generator for the driven process has the same structure as the original process [15]. In the following, we combine large deviation theory framework with tools from queueing theory [16, 17], to obtain analytical formulas for the statistics of rare events in a general class of stochastic models of gene expression. We find that the conditioning-free Markov process that generates rare fluctuations generically maintains the same structure as the original Markov process.

Model

We consider gene expression from a promoter with N internal states, labeled $i = 1, \ldots, N$, as illustrated in figure 1. The promoter makes random transitions among its N states switching from $j \rightarrow i$ with rate a_{ij}. In each state, bursts of gene expression leading to the production of mRNAs occur with rates k_i, and burst sizes n drawn from a state-dependent distribution $b_i(n)$. In the limit that protein degradation rates are much smaller than mRNA degradation rates, a widely-used approximation involves assuming that proteins are created in random instantaneous bursts from each mRNA [18, 19, 20, 21]. Given the validity of this 'bursty protein synthesis' approximation, the model considered in figure 1 can also be used to represent gene expression at the level of proteins.

The statistical state of the system is specified by a vector $\mathbf{p}_t(M) = \left\{ p_t^i(M) \right\}_{i=1}^N$ whose elements are the probabilities $p_t^i(M)$ for the promoter to be in state i at time t having produced a total of M mRNAs. The dynamics is an example of what is known in queueing theory as a Batch Markovian Arrival Process (BMAP) [22], whose evolution is specified by the master equation

$$\dot{\mathbf{p}}_t(M) = \sum_{M'} \mathbf{D}_{M-M'} \mathbf{p}_t(M'), \tag{1}$$

where the $N \times N$ matrix $\mathbf{D}_0$ is the transition rate matrix for intra-promoter transitions, whereas $\mathbf{D}_n$ (n 1) is an $N \times N$ diagonal matrix whose nonzero elements specify the rate of creating n mRNA in a burst:

$$D_0^{ij} = \begin{cases} -\sum_j \alpha_{ij} - k_j, & i = j \\ \alpha_{ij}, & i \neq j \end{cases} \; ; \; D_n^{ij} = k_j b_j(n) \delta_{ij}. \tag{2}$$

The diagonal elements D_0^{ii} enforce probability conservation. While the representation of the dynamics in terms of the matrices $\mathbf{D}_n$ is convenient, the full dynamics occur on the infinite-dimensional space spanned by the states (i, M). Thus, the generator matrix for the master equation (1) is infinite dimensional,

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_0 & 0 & 0 & \cdots \\ \mathbf{D}_1 & \mathbf{D}_0 & 0 & \\ \mathbf{D}_2 & \mathbf{D}_1 & \mathbf{D}_0 & \\ \vdots & & & \ddots \end{pmatrix}, \tag{3}$$

which is a formulation of the dynamics we will need in the following.

Finally, note that if we focus on promoter dynamics alone, the state of the system is specified by the vector $\mathbf{P}_t = \sum_M \mathbf{p}_t(M)$. The corresponding master equation has the generator $\mathscr{D} = \sum_n \mathbf{D}_n$. In this case, the dynamics will be referred to as promoter-only dynamics, whereas the dynamics in (1) is the full system dynamics. In the following, unless otherwise stated, we will assume that the matrix $\mathscr{D}$ is irreducible.

Large deviations

Large deviation theory for Markov processes specifies the relative likelihood to observe large fluctuations in trajectory observables [12, 14]. We will be interested in the fluctuations of the class of trajectory random variables

$$A_t = \sum_{j \to i} g_{ij}^0 + \sum_{(i,M) \to (i,M+n)} g_{ii}^n + \int_0^t ds \, f_{i\,(s)}, \tag{4}$$

where the sums extend over all transitions along a random realization of the BMAP, with each transition weighted by the parameters g and the time spent in each state weighted by f. Notable examples are the promoter activity, that is the number of mRNAs/proteins produced up to time t ($f_i = 0$ and $g_{ij}^n = n\,\delta_{ij}$), or the total time spent in promoter state i ($f_j = \delta_{ij}$ and $g_{ij}^n = 0$).

Now, the law of large numbers tells us that for large t, the *rate* $a = A_t/t$, will approach its average value $= \lim_{t \to \infty} A_t/t$ with near unit probability. However, rare large deviations

from this mean value are possible, and can be quantified, because the probability density of a satisfies a large deviation principle for large t [11, 23],

$$P(A_t=at) \sim e^{-tI(a)} \quad (5)$$

with large deviation rate function $I(a)$. One can view this as an extension of the central limit theorem, quantifying not just the Gaussian fluctuations about the typical value (given by the minimum of I), but also the relative likelihood of rare fluctuations.

The theory of large deviations establishes that one way of obtaining the rate function $I(a)$ is by evaluating the (scaled) cumulant generating function (SCGF) [23, 11]

$$\psi(\lambda) = - \lim_{t \to \infty} \frac{1}{t} \ln \left\langle e^{-\lambda A_t} \right\rangle t, \quad (6)$$

where $\langle \cdot \rangle_t$ denotes an average over full system trajectories of duration t. The rate function can then be recovered using the Legendre-Fenchel transform [23, 11]

$$I(a) = \sup_{\lambda}(\psi(\lambda) - \lambda a). \quad (7)$$

The function $I(a)$ ($\psi(\lambda)$) is the nonequilibrium analog of the entropy (free energy) in equilibrium systems. It is important to note that (7) only gives the convex hull of I, that is the smallest convex set that encompasses I [11]. When I is not strictly convex, that is having multiple local minima or having a linear part, the SCGF ψ is not differentiable everywhere, and is characterized by a nonanalyticity such as a kink. The appearance of such a structure in ψ is called a *dynamical phase transition* [11, 24, 25] in analogy with the nonanalytic behavior of the free energy characterizing phase transitions in equilibrium statistical mechanics.

A key insight from large deviation theory is that $\psi(\lambda)$ can be obtained from the largest eigenvalue of a modified or twisted generator matrix (cf. (3))

$$\tilde{\mathbf{D}}(\lambda) = \begin{pmatrix} \tilde{\mathbf{D}}_0(\lambda) & 0 & \cdots \\ \tilde{\mathbf{D}}_1(\lambda) & \tilde{\mathbf{D}}_0(\lambda) & \\ \vdots & & \ddots \end{pmatrix}, \quad (8)$$

with elements [13, 14]

$$\tilde{D}_0^{ij}(\lambda) = \begin{cases} -\sum_j \alpha_{ij} - k_j - \lambda f_j, & i=j \\ \alpha_{ij} e^{-\lambda g_{ij}^0}, & i \neq j \end{cases} \quad ; \quad \tilde{D}_n^{ij}(\lambda) = k_j b_j(n) e^{-\lambda g_{jj}^n} \delta_{ij}. \quad (9)$$

To obtain ψ from the largest eigenvalue of $\tilde{\mathbf{D}}(\lambda)$, we have to diagonalize an infinite-dimensional matrix. As we will show, the structure of BMAPs greatly simplifies this calculation, and reduces the computation to finding the dominant eigenvalue of an N-dimensional matrix, where N is the number of promoter states.

The basic idea is to investigate the generating function $G_t(\lambda) = \left\langle e^{-\lambda A_t} \right\rangle_t$ for A_t, from which, in the long-time limit, the SCGF can be obtained as $\psi(\lambda) = -\lim_{t\to\infty} (1/t) \ln G_t(\lambda)$ (by definition (6)). We can express $G_t(\lambda)$ in terms of the joint probability at time t to be at (i, M) having accumulated A of the trajectory observable, $\mathbf{p}_t(M, A)$:

$$G_t(\lambda) = \left\langle e^{-\lambda A_t} \right\rangle_t = \int dA \sum_M \mathbf{1} \cdot \mathbf{p}_t(M, A) e^{-\lambda A} \equiv \sum_{M'} \mathbf{1} \cdot \mathbf{G}_t(M'; \lambda), \tag{10}$$

where we have introduced the state-dependent generating function $\mathbf{G}_t(M; \lambda) = \int dA\, \mathbf{p}_t(M, A) e^{-\lambda A}$. The key insight from large deviation theory is that $\mathbf{G}_t(M; \lambda)$ is the solution of the twisted dynamics [13]

$$\dot{\mathbf{G}}_t(M; \lambda) = \sum_{M'} \tilde{\mathbf{D}}_{M-M'}(\lambda) \mathbf{G}_t(M'; \lambda), \tag{11}$$

and its long-time behavior is controlled by the largest eigenvalue of its generator, the infinite-dimensional matrix $\tilde{\mathbf{D}}$.

The structure of BMAPs allows us to simplify this calculation significantly. To this end, let us solve for the long-time dynamics of $\mathbf{G}_t$ by introducing $\boldsymbol{\Gamma}_t(z; \lambda) = \sum_M z^M \mathbf{G}_t(M; \lambda)$, which from (11) evolves according to the simplified N-dimensional linear equation

$$\dot{\boldsymbol{\Gamma}}_t(z; \lambda) = \tilde{\mathscr{D}}(z; \lambda) \boldsymbol{\Gamma}_t(z; \lambda), \tag{12}$$

where $\tilde{\mathscr{D}}(z; \lambda) = \sum_n z^n \tilde{\mathbf{D}}_n(\lambda)$. Let us define $\tilde{\mathscr{D}}(\lambda) = \tilde{\mathbf{D}}_n(1; \lambda) = \sum_n \tilde{\mathbf{D}}_n(\lambda)$. Notably, $\tilde{\mathscr{D}}(\lambda)$ is the $N \times N$ generator for promoter-only twisted dynamics. The largest eigenvalue of $\tilde{\mathscr{D}}(\lambda)$ controls the long-time evolution of $\boldsymbol{\Gamma}_t(z = 1, \lambda)$ and thus gives us the SCGF. Indeed, denoting the largest eigenvalue of $\tilde{\mathscr{D}}(\lambda)$ by $-\psi(\lambda)$ (anticipating the conclusion), and substituting the solution of (12) into (10) gives us

$$G_t(\lambda) = \left\langle e^{-\lambda A_t} \right\rangle_t = \mathbf{1} \cdot \boldsymbol{\Gamma}_t(z = 1; \lambda) \sim e^{-t\psi(\lambda)}. \tag{13}$$

While the generator for the full system dynamics is infinite dimensional, the SCGF is obtained from the dominant eigenvalue of the N-dimensional matrix $\tilde{\mathscr{D}}(\lambda)$ for a model with N promoter states, a substantial simplification.

Conditioning on rare fluctuations

Suppose we want to know what typical trajectories of the system look like, conditioned on observing a (possibly rare) rate a for the random variable A_t. Recent research in nonequilibrium statistical mechanics has shown that such information is encoded in the left eigenvector $\tilde{\mathbf{L}}(\lambda)$ corresponding to the dominant eigenvalue $\psi(\lambda)$ of the twisted generator $\tilde{\mathbf{D}}(\lambda)$ [14, 25, 26]. Again, the structure of BMAPs leads to a significant simplification in determining this left eigenvector. Denoting $\tilde{\mathbf{l}}_\lambda$ as the left eigenvector corresponding to ψ for $\tilde{\mathcal{D}}(\lambda)$, we have that the infinite-dimensional left eigenvector $\tilde{\mathbf{L}}(\lambda) = \left\{ \tilde{\mathbf{l}}(\lambda), \ldots, \tilde{\mathbf{l}}(\lambda) \right\}$ is formed from repeated blocks of $\tilde{\mathbf{l}}_\lambda$, due to the repeated block structure of $\mathbf{D}$ (3). Using this in combination with recent results for the driven process, we find that the conditioning-free Markov process that reproduces the statistics of the original process conditioned on a has a generator with the same structure as (3), with the corresponding transition rate matrices given by [14, 26]

$$\Delta_n^{ij}(\lambda^*) = l_i(\lambda^*)\tilde{D}_n^{ij}(\lambda^*)l_j^{-1}(\lambda^*) + \psi(\lambda^*)\delta_{ij}\delta_{n0}, \quad (14)$$

with $\lambda^* = -I'(a)$. This surprising fact crucially depends on the class of random variables A_t considered in (4), which count promoter transitions with a weight that is independent of the number of mRNAs. Furthermore, given that the random variable A_t is derived solely from dynamics related to the production of mRNAs/proteins, the results for the driven process are independent of the decay dynamics for mRNAs/proteins. In particular, we note that the results for parameters defining the driven model will be the same for the case where the decay rate is $\mu = 0$ (which does not have a well-defined steady-state) and for the case for finite μ (which does have a well-defined steady-state).

By analyzing the equations determining the left eigenvector $\tilde{\mathbf{l}}_\lambda$ we find, remarkably, that the driven process corresponds to another BMAP with modified rates

$$\tilde{\alpha}_{ij} = l_i\alpha_{ij}l_j^{-1}, \quad \tilde{k}_i = k_i\sum_n b_i(n)e^{-\lambda g_{ii}^n}. \quad (15)$$

and renormalized burst distribution

$$\tilde{b}_i(n) = \frac{b_i(n)e^{-\lambda g_{ii}^n}}{\sum_n b_i(n)e^{-\lambda g_{ii}^n}}. \quad (16)$$

The above results, in combination with the equations determining the SCGF $\psi(\lambda)$ and rate function $I(a)$, constitute an analytical framework for characterizing rare events in a general class of stochastic models of gene expression.

Applications

In the following, we will focus on applications of the preceding theoretical framework to the case of activity fluctuations (*i.e.* $f_i = 0$ and $g_{ij}^n = n\delta_{ij}$).

Let us consider rare events corresponding to low activity a from a given promoter with a typical activity . This correspond to a fold-reduction in the mean activity of $r = a/$. Let $\tilde{\psi}_{\lambda^*}$ with $\lambda^* = -I'(a)$ denote the SCGF for the driven process. Then, by construction $a = \tilde{\psi}'_{\lambda^*}(0)$. As a result, the fold-change r can be expressed in terms of the driven process as,

$$r = \frac{\tilde{\psi}'_{\lambda^*}(0)}{\psi'(0)}. \quad (17)$$

However, we can flip this equation around and consider it as an equation for λ^* in terms of the fold change r. Once λ^* is obtained by solving this equation, the driven process for a given fold-change r is completely determined.

The driven process represents the most likely way the rare event in activity fluctuations occurs and thus is an ideal choice for generating distributions that can be used in importance sampling. Recent work, using a control theory perspective, has also shown that the driven process is the optimal way to to achieve the fold-reduction in activity while minimizing the 'distance', in the sense of relative entropy, from the original process [27, 26]. Thus explicit characterization of the driven process is also critical in designing regulatory strategies for reducing promoter activity (thereby reducing mean mRNA/protein levels) that give rise to processes that are closest to the original process.

To illustrate the theoretical framework outlined above, we consider a widely used model of gene expression at the protein level [18, 19, 20]: geometrically distributed bursts (with mean b) of proteins arriving according to a Poisson process with rate k_m. In this case, the burst generating function is $g_b(z) = \dfrac{z}{b - z(b-1)}$ and correspondingly, we obtain that the SCGF $\psi(\lambda) = k_m - k_m g_b(e^{-\lambda})$. For a rare event corresponding to a r-fold reduction in mean activity, using (17) we see that the corresponding value of λ^* is obtained by solving

$$\frac{e^{-\lambda^*}}{\left(b - e^{-\lambda^*}(b-1)\right)^2} = r \quad (18)$$

Interestingly, we find that the renormalized burst distribution for the driven process continues to be a geometric distribution with the corresponding mean given by

$$\tilde{b} = \frac{b}{b - e^{-\lambda*}(b-1)} \quad (19)$$

This indicates that the optimal way to reduce activity involves arrival of geometric bursts with a reduced mean and a modified arrival rate. Interestingly, in previous work [28], we have shown that post-transcriptional regulation (e.g by miRNAs) can alter bursts of protein expression by reducing the mean but keeping the burst distribution geometric. These results indicate that that optimal control (in the sense discussed above) can be achieved using a combination of transcriptional regulation and post-transcriptional regulation by miRNAs.

As a second illustration, we consider the two-state promoter model [29] depicted in figure 2 with burst size one. The SCGF for mRNA activity is readily obtained by diagonalizing the twisted transition rate matrix:

$$\psi(\lambda) = \frac{1}{2}\left(\alpha + \beta + k_{\mathrm{m}}(1 - e^{-\lambda}) - \sqrt{4\alpha\beta + (\beta - \alpha + k_{\mathrm{m}}(1 - e^{-\lambda}))^2}\right). \quad (20)$$

Remarkably, ψ develops a nonanalyticity or kink in either the $\alpha \to 0$ or $\beta \to 0$ limit. Specifically for $\beta \to 0$, we plot in figure 2

$$\psi(\lambda)\frac{1}{2}(\alpha + k_{\mathrm{m}}(1 - e^{-\lambda}) - |\alpha - k_{\mathrm{m}}(1 - e^{-\lambda})|), \quad (21)$$

which is nonanalytic for $\alpha < k_{\mathrm{m}}$. The observed kink in the SCGF is a signature of a dynamical phase transition, akin to the multiphase behavior observed in photon counting statistics of a quantum two-level system [30]. Here, the phase transition occurs because, as the activity level corresponding to the large deviation is changed, there is a qualitative change in the fluctuations that give rise to the rare event. In particular, below a critical activity level, these fluctuations involve long sojourns in the 0 (OFF) state (with no activity) that correspond to a nonvanishing finite fraction of the total time t. We do note, however, that the nonanalytic behavior only arises when the matrix for promoter-only dynamics $\mathscr{D}$ becomes reducible. A detailed characterization of fluctuations in this limit will be presented in future work.

In conclusion, we have presented a framework for the quantitative analysis of the probability of large deviations and for characterizing system fluctuations conditional on rare events during gene expression. The framework developed provides explicit analytical formulae determining the driven process for a general class of stochastic models corresponding to BMAPs. Our results demonstrate that the driven process corresponding to a BMAP is another BMAP with renormalized parameters. This property may also be present in a related class of renewal processes, which, while distinct, share a similar structure in the dynamics [31]. Since BMAPs are used to model a wide range of applications in science and engineering (*e.g.* computer and communications networks [22]), the results derived are expected to have diverse applications. In the context of models of gene expression, the results derived can be used to: (1) determine the probability of rare events corresponding to a broad class of random variables for general promoter models with arbitrary mRNA/protein burst distributions; (2) directly simulate rare system trajectories corresponding to a large deviation in the random variable of interest and to predict the corresponding mRNA/protein

distributions conditioned on the rare event; (3) optimally control gene expression to achieve, for example, a specific fold-regulation while minimizing deviation from the unregulated process; and (4) determine regions in parameter space corresponding to nonanalytic behavior of the SCGF $\psi(\lambda)$. The results obtained have thus opened new avenues for analyzing rare events in gene expression and have multiple applications ranging from importance sampling to optimal strategies for cellular regulation of gene expression.

Acknowledgments

References

1. Raj A, van Oudenaarden A. Cell. 2008; 135:216–226. [PubMed: 18957198]

2. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV. Cell. 2005; 122:169–182. [PubMed: 16051143]

3. Mirouze N, Prepiak P, Dubnau D. PLoS Genet. 2011; 7:e1002048. [PubMed: 21552330]

4. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA, et al. Cell. 2010; 141:69–80. [PubMed: 20371346]

5. Assaf M, Roberts E, Luthey-Schulten Z. Physical review letters. 2011; 106:248102. [PubMed: 21770603]

6. Newby JM. Physical biology. 2012; 9:026002. [PubMed: 22473173]

7. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Science. 2011; 332:472–474. [PubMed: 21415320]

8. Sanchez A, Garcia HG, Jones D, Phillips R, Kondev J. PLoS Comput Biol. 2011; 7:e1001100. [PubMed: 21390269]

9. Zhang J, Zhou T. Biophysical journal. 2014; 106:479–488. [PubMed: 24461023]

10. Lecomte V, Appert-Rolland C, van Wijland F. Journal of Statistical Physics. 2007; 127:51–106. ISSN 0022-4715.

11. Touchette H. Physics Reports. 2009; 478:1–69.

12. Touchette, H., Harris, RJ. Nonequilibrium statistical physics of small systems: Fluctuation relations and beyond. Wiley-VCH; Weinheim: 2013. chap 11

13. Jack RL, Sollich P. Prog Theor Phys. 2010; 184:304–317.

14. Chetrite R, Touchette H. Phys Rev Lett. 2013; 111:120601. [PubMed: 24093237]

15. Torkaman P, Jafarpour FH. Phys Rev E. 2015; 92:062104.

16. Williams RJ. Annual Review of Statistics and Its Application. 2016; 3:323–345.

17. Arazi A, Ben-Jacob E, Yechiali U. Physica A: Statistical Mechanics and its Applications. 2004; 332:585–616.

18. Friedman N, Cai L, Xie XS. Physical review letters. 2006; 97:168302. [PubMed: 17155441]

19. Shahrezaei V, Swain PS. Proceedings of the National Academy of Sciences. 2008; 105:17256–17261.

20. Elgart V, Jia T, Fenley AT, Kulkarni R. Physical biology. 2011; 8:046001. [PubMed: 21490380]

21. Jia T, Kulkarni RV. Physical review letters. 2011; 106:058102. [PubMed: 21405439]

22. Cordeiro JD, Kharoufeh JP. Wiley Encyclopedia of Operations Research and Management Science. 2011

23. Ellis, R. Entropy, large deviations, and statistical mechanics (Grundlehren der mathematischen wissenschaften). Vol. 271. Springer-Verlag; New York: 1985.

24. Garrahan JP, Jack RL, Lecomte V, Pitard E, van Duijvendijk K, van Wijland F. Physical Review Letters. 2007; 98:195702. [PubMed: 17677633]

25. Vaikuntanathan S, Gingrich TR, Geissler PL. Phys Rev E. 2014; 89:062108.

26. Jack RL, Sollich P. The European Physical Journal Special Topics. 2015; 224:2351–2367.

27. Chetrite R, Touchette H. J Stat Mech: Theor Exp. 2015:P12001.

28. Jia T, Kulkarni RV. Physical review letters. 2010; 105:018101. [PubMed: 20867481]

29. Sanchez A, Golding I. Science. 2013; 342:1188–1193. [PubMed: 24311680]

30. Garrahan JP, Lesanovsky I. Phys Rev Lett. 2010; 104:160601. [PubMed: 20482036]

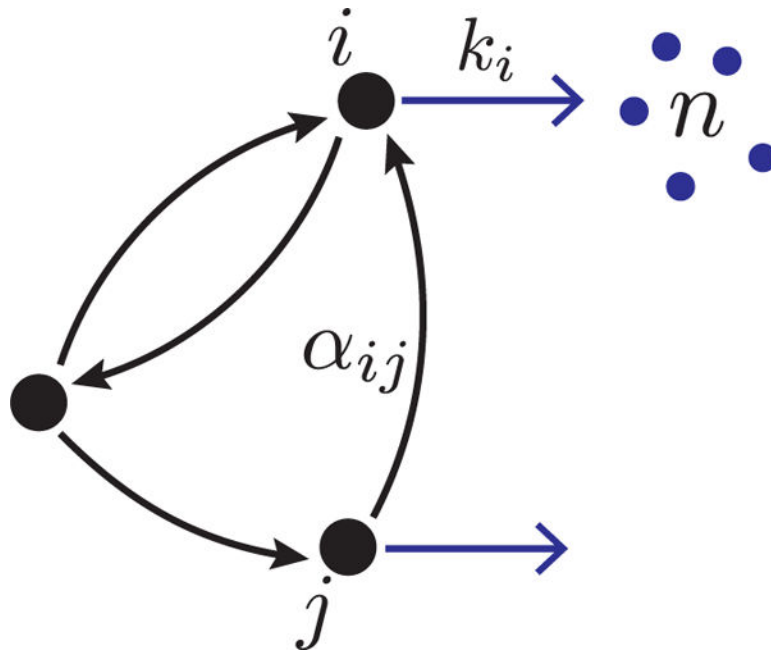31. Harris RJ, Touchette H. J Phys A: Math Theor. 2017; 50:13.

Figure 1.
Schematic of generic promoter model for bursty mRNA transcription. The promoter makes random transitions among its $N = 3$ states with rates a_{ij}. Bursts of mRNA of size n are drawn from a state-dependent burst size distribution $b_i(n)$ and arrive with rate k_i.

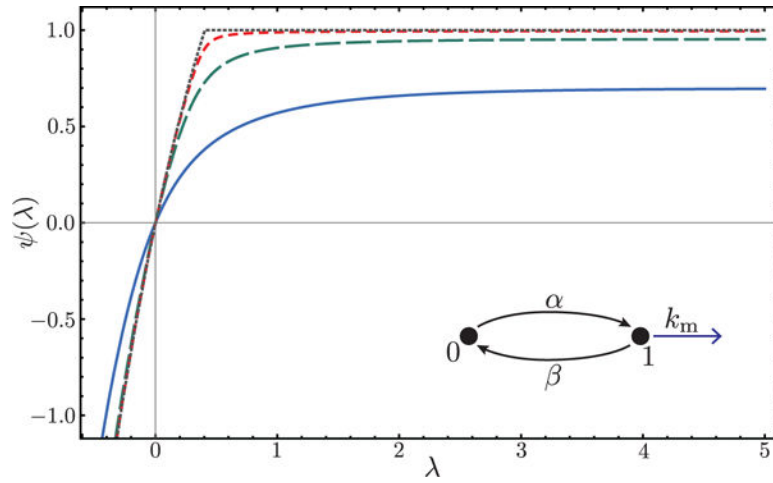**Figure 2.**
Cumulant generating function $\psi(\lambda)$ for the two-state promoter model pictured inset.
Different lines correspond to decreasing β with the limiting value $\beta \to 0$ given by the dotted
line. All graphs have $\alpha < k_m$ and thus are in the nonanalytic phase for $\beta \to 0$. Values: $\alpha = 1$,
$k_m = 3$, $\beta = 1, 0.1, 0.01, 0$.