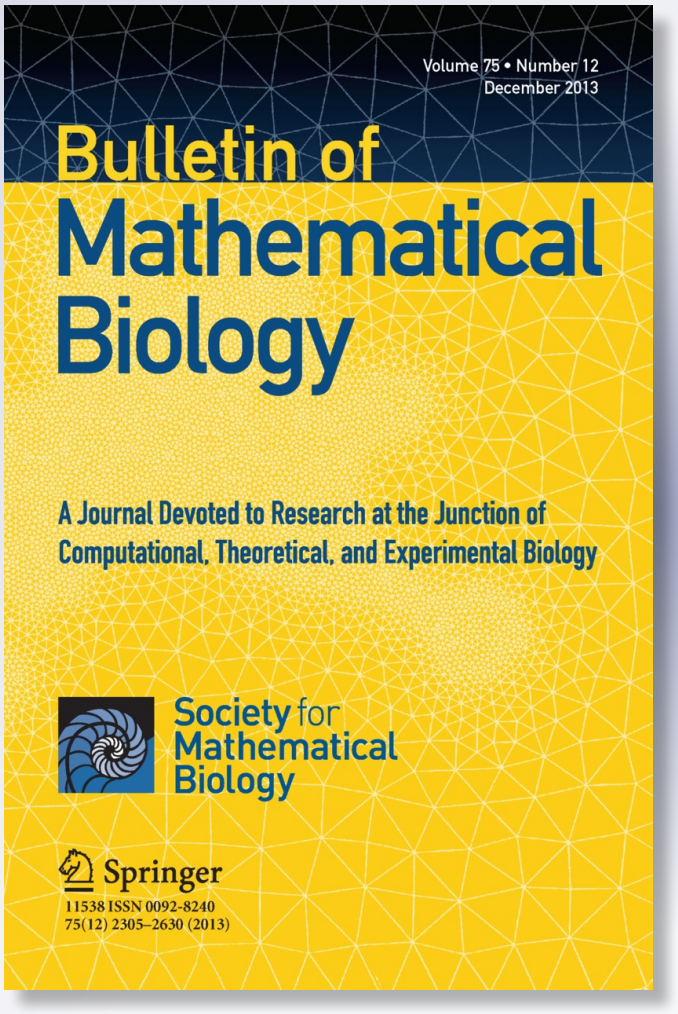
# *Multimodality and Flexibility of Stochastic Gene Expression*

**Guilherme da Costa Pereira Innocentini, Michael Forger, Alexandre Ferreira Ramos, Ovidiu Radulescu & José Eduardo Martinho Hornos**

**Bulletin of Mathematical Biology**
A Journal Devoted to Research at the Junction of Computational, Theoretical and Experimental Biology Official Journal of The Society for Mathematical Biology

ISSN 0092-8240
Volume 75
Number 12

Bull Math Biol (2013) 75:2600-2630
DOI 10.1007/s11538-013-9909-3




Springer

Bull Math Biol (2013) 75:2600–2630
DOI 10.1007/s11538-013-9909-3

Society for Mathematical Biology

ORIGINAL ARTICLE

# Multimodality and Flexibility of Stochastic Gene Expression

**Guilherme da Costa Pereira Innocentini · Michael Forger · Alexandre Ferreira Ramos · Ovidiu Radulescu · José Eduardo Martinho Hornos**

Received: 10 April 2013 / Accepted: 24 September 2013 / Published online: 18 October 2013


**Abstract** We consider a general class of mathematical models for stochastic gene expression where the transcription rate is allowed to depend on a promoter state variable that can take an arbitrary (finite) number of values. We provide the solution of the master equations in the stationary limit, based on a factorization of the stochastic transition matrix that separates timescales and relative interaction strengths, and we express its entries in terms of parameters that have a natural physical and/or biological interpretation. The solution illustrates the capacity of multiple states promoters to generate multimodal distributions of gene products, without the need for feedback. Furthermore, using the example of a three states promoter operating at low, high, and intermediate expression levels, we show that using multiple states operons will typically lead to a significant reduction of noise in the system. The underlying mechanism is that a three-states promoter can change its level of expression from low to high by passing through an intermediate state with a much smaller increase of fluctuations than by means of a direct transition.

**Keywords** Gene expression · Stochasticity · Noise reduction

G.C.P. Innocentini (✉) · M. Forger
Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão, 1010, Cidade Universitária, São Paulo, SP, Brazil, CEP: 05508-090
e-mail: ginnocentini@gmail.com

A.F. Ramos
Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, Av. Arlindo Béttio, 1000, Ermelino Matarazzo, São Paulo, SP, Brazil, CEP: 03828-000

O. Radulescu
IMNP, UMR 5235, Université de Montpellier 2, Pl. E. Bataillon, Bat. 24, 34095 Montpellier Cedex 5, France

J.E.M. Hornos
Instituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador São-Carlense, 400, São Carlos, SP, Brazil, CEP: 13566-590

## 1 Introduction

Modern molecular biology emphasizes the important role of gene networks in the functioning of living organisms. The components of these networks are DNA and mRNA sequences, as well as proteins, that interact and function together in coordination. In particular, gene transcription is governed by the enzymatic action of regulatory proteins, which can enhance or repress the production of mRNA molecules (Monod and Jacob 1961). Moreover, it has been clear ever since the pioneering theoretical work of Delbrück (1940) that in these processes, fluctuations are unavoidable and in fact play a decisive role, not only because biochemical reactions occur on widely varying time scales but also, and more importantly, because at least some of the molecules involved come in small numbers. This expectation has been confirmed experimentally using recent advances in fluorescence microscopy (Ozbudak et al. 2002; Cai et al. 2006; Raj et al. 2006; Elf et al. 2007; Taniguchi et al. 2010; Ferguson et al. 2012), showing that this "gene expression noise" is ubiquitous in both prokaryote and eukaryote biology and produces genetic heterogeneity even from uniform initial conditions, e.g., in clone cell populations. As a result, reliable control of biological functions requires gene networks to be endowed with mechanisms not only for error correction but also for noise reduction.

In order to gain a better understanding of the role that noise plays in genetic networks, various mathematical models have been proposed. Some are based on numerical simulations, taking into account the complete set of chemical reactions involved in the process (Gillespie 1977; Arkin et al. 1998; McAdams and Arkin 1997, 1998; Goss and Peccoud 1998; Kierzek et al. 2001). Others use systems of differential equations for the concentrations of the molecules involved in the process, where the noise is introduced via the Langevin mechanism (Gardner et al. 2000; Hasty et al. 2000; Rao et al. 2002; Thattai and van Oudenaarden 2002). In a third kind of models, often referred to as "microscopic models," one considers the number of each type of molecule as a discrete random variable and establishes a set of master equations for their probability distributions (Paulsson and Ehrenberg 2000; Thattai and van Oudenaarden 2001; Metzler and Wolynes 2002).

Whatever the approach chosen, all information about the system, including that about gene expression noise, is contained in the probability distributions of the gene network variables. Clearly, for computational as well as experimental reasons, it is often easier to analyze just their first two moments. However, in many situations, knowing only the mean value and the standard deviation is insufficient to characterize them. A notable situation of this type is multimodality, where the probability distribution of gene expression has several maxima. Well-known sources of multimodality are positive feedback and multistationarity (see, for instance Huang 2009; Satory et al. 2011 and references therein), but they are not the only ones. Indeed, in what follows, we shall present another source of multimodality: one that can arise without feedback.

The first step in the control of gene expression is at the level of transcription. Gene transcription results from the interaction between regulatory proteins (transcription factors, repressors, activators) and specific regions on the DNA (promoters and cis-regulatory elements such as enhancers, silencers and insulators). Clearly,

treating the transcription regulatory system as a simple "on/off" process is an oversimplification (Cases and de Lorenzo 2005; Vicente et al. 1999; Escolar et al. 1999); a somewhat more realistic model should allow for a larger number of states, each one with a different probability to initiate transcription (Sánchez and Kondev 2008; Coulon et al. 2010; Saiz and Vilar 2008; Garcia and Phillips 2011; Kirkilionis et al. 2011). Moreover, the transcription system can switch stochastically between these states. The multiplicity of discrete transcription states produces multimodal expression distributions, provided that the sojourn time of the system in each of these states is larger than, or of the same order of magnitude as, the time needed for the gene expression to reach equilibrium.

In this paper, we propose a phenomenological model for stochastic gene expression from a promoter with multiple operational states, generalizing the binary models (Kepler and Elston 2001; Sasai and Wolynes 2003) where the gene switches randomly between just an "on" state and an "off" state, characterized by the absence or presence of a repressor protein in the operator site. It is a microscopic model in the sense mentioned above, based on a set of coupled master equations to describe the stochastic dynamics of the system. Contrary to previous papers introducing similar models and solving them numerically (Sánchez and Kondev 2008; Coulon et al. 2010), we provide an explicit solution for the stationary version of the master equations, in terms of recursion relations for the coefficients of an appropriate power series expansion of that solution, allowing to compute these coefficients up to arbitrary order.

An important property of multimodal gene expression that emerges from our analysis is that promoters with multimodal expression can adapt to changes in the environment by a "mode exchange" strategy. In this strategy, the average expression is changed by gradually shifting the bulk of the probability distribution from one mode to another. The existence of intermediate states between the two states of extreme transcription rates allows to perform this mode exchange with little or no increase of the expression noise. By analogy, we may compare this strategy to gear changing in a modern vehicle.

The organization of the paper is as follows. In Sect. 2, we introduce the master equations for a three states model and explain the terms. In Sect. 3, we redefine the parameter space in terms of biological quantities. In Sect. 4, we compute the probability distribution, mean value, fluctuations, and Fano factor for the three states model. In Sect. 5, we show how to formulate the model for an arbitrary number of operational modes and how solutions are obtained in this scenario. Finally, Sect. 6 is dedicated to discussion and conclusions.

## 2 The Model

In what follows, we show how to extend the model considered in Kepler and Elston (2001), Sasai and Wolynes (2003) so as to accommodate three (rather than just two) operational states of different transcriptional efficiencies. A biological paradigm of this situation is the lac operon, responsible for the transport and metabolism of lactose in *Escherichia coli*, as well as in many other bacteria (Monod and Jacob 1961; Muller-Hill 1996). In this system, the availability of glucose inhibits the cAMP-CRP activator

(catabolic repression), whereas the availability of lactose inhibits the LacI repressor. Thus, when glucose is the only carbon source present in the cell, the operon has the lowest level of transcription. When lactose is the only carbon source, the operon transcribes abundantly in order to metabolize the available lactose. An intermediate level of transcription occurs when both glucose and lactose are present in the cellular medium. In general, we suppose that there are stochastic transitions between these three states, with probabilities that are functions of the amounts of glucose and lactose in the system. Although this model is still too simple to cope with all the details of the lac operon behavior, it has the advantage of considering the possibility that lac-induction, i.e., the switching from an "off" to an "on" state, can involve a third state of intermediate transcriptional activity. Furthermore, catabolic repression is a strategy to optimize carbon source utilization, widely employed by bacterial phyla. It is therefore plausible that one can find many other examples of bacterial operons with several states of transcriptional activity.

With this biological paradigm in mind, let us begin by introducing the stochastic variable of our model, $n$, representing the number of mRNA molecules in the cell at any given instant of time, $t$. Then we assign to each operational mode the probability distributions $\phi_n^i(t)$, with $i = 1, 2, 3$ (corresponding to the active, intermediate and basal states of the lac operon). Thus, $\phi_n^i(t)$ is the probability to find the system, at time $t$, in the state $i$ with $n$ molecules of mRNA in the cell. The master equations, of birth and death type, describing the evolution of these probability distributions are:

$$\begin{aligned}
\frac{d\phi_n^1}{dt} &= k_1\left[\phi_{n-1}^1 - \phi_n^1\right] + \rho\left[(n+1)\phi_{n+1}^1 - n\phi_n^1\right] \\
&\quad - h_1\phi_n^1 + h_2\epsilon_2\phi_n^2 + h_3(1-\epsilon_3)\phi_n^3, \\
\frac{d\phi_n^2}{dt} &= k_2\left[\phi_{n-1}^2 - \phi_n^2\right] + \rho\left[(n+1)\phi_{n+1}^2 - n\phi_n^2\right] \\
&\quad + h_1(1-\epsilon_1)\phi_n^1 - h_2\phi_n^2 + h_3\epsilon_3\phi_n^3, \\
\frac{d\phi_n^3}{dt} &= k_3\left[\phi_{n-1}^3 - \phi_n^3\right] + \rho\left[(n+1)\phi_{n+1}^3 - n\phi_n^3\right] \\
&\quad + h_1\epsilon_1\phi_n^1 + h_2(1-\epsilon_2)\phi_n^2 - h_3\phi_n^3.
\end{aligned} \tag{1}$$

Here, the parameters $k_1$, $k_2$ and $k_3$ are the production rates of mRNA in each of the states, which in general will be different from each other since we are interested in processes involving states of different transcriptional activity, while $\rho$ is the degradation rate of mRNA molecules, which is assumed to be the same for all states. The remaining parameters are the transition rates between the states and can be represented as follows:

$$\phi_n^1 \underset{h_2\epsilon_2}{\overset{h_1(1-\epsilon_1)}{\rightleftharpoons}} \phi_n^2, \qquad \phi_n^1 \underset{h_3(1-\epsilon_3)}{\overset{h_1\epsilon_1}{\rightleftharpoons}} \phi_n^3, \qquad \phi_n^2 \underset{h_3\epsilon_3}{\overset{h_2(1-\epsilon_2)}{\rightleftharpoons}} \phi_n^3,$$

with $0 \le h_i < \infty$ and $0 \le \epsilon_i \le 1$, for $i = 1, 2, 3$; see Fig. 1 for a schematic representation.

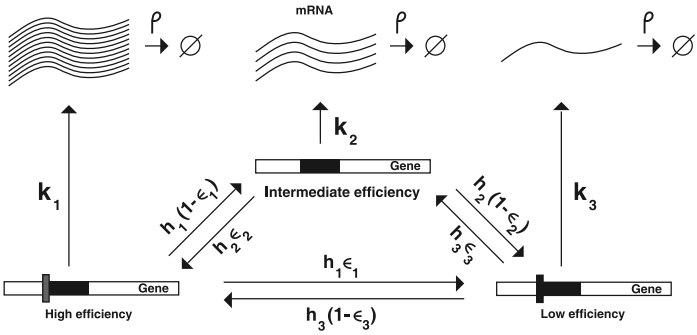



**Fig. 1** Schematic representation of the three states model

The master equations (1) are a system of differential-difference equations, which are linear in $n$. Using the generating functions technique (van Kampen 1992), we can transform it into a system of partial differential equations, setting

$$\phi^1(z,t)=\sum_{n=0}^{\infty}\phi_n^1(t)z^n,\qquad \phi^2(z,t)=\sum_{n=0}^{\infty}\phi_n^2(t)z^n,\qquad \phi^3(z,t)=\sum_{n=0}^{\infty}\phi_n^3(t)z^n, \tag{2}$$

where $z$ is a complex variable belonging to the unitary disk, $|z|\leq 1$. It reads:

$$\begin{aligned}
\frac{\partial\phi^1}{\partial t}&=(z-1)\left[k_1\phi^1-\rho\frac{\partial\phi^1}{\partial z}\right]-h_1\phi^1+h_2\epsilon_2\phi^2+h_3(1-\epsilon_3)\phi^3,\\
\frac{\partial\phi^2}{\partial t}&=(z-1)\left[k_2\phi^2-\rho\frac{\partial\phi^2}{\partial z}\right]+h_1(1-\epsilon_1)\phi^1-h_2\phi^2+h_3\epsilon_3\phi^3,\\
\frac{\partial\phi^3}{\partial t}&=(z-1)\left[k_3\phi^3-\rho\frac{\partial\phi^3}{\partial z}\right]+h_1\epsilon_1\phi^1+h_2(1-\epsilon_2)\phi^2-h_3\phi^3.
\end{aligned} \tag{3}$$

The original probability distributions can be recovered as derivatives of the generating functions, evaluated at $z=0$, as follows:

$$\begin{aligned}
\phi_n^1(t)&=\frac{1}{n!}\frac{\partial^n\phi^1(z,t)}{\partial z^n}\bigg|_{z=0},\\
\phi_n^2(t)&=\frac{1}{n!}\frac{\partial^n\phi^2(z,t)}{\partial z^n}\bigg|_{z=0},\\
\phi_n^3(t)&=\frac{1}{n!}\frac{\partial^n\phi^3(z,t)}{\partial z^n}\bigg|_{z=0}.
\end{aligned} \tag{4}$$

Similarly, their moments are obtained as derivatives of the same generating functions, but evaluated at $z = 1$:

$$\begin{aligned}
\langle n^{1,j}\rangle(t) &= \left(z\frac{\partial}{\partial z}\right)^j \phi^1(z,t)\Big|_{z=1} = \sum_{n=0}^{\infty} n^j \phi_n^1(t),\\
\langle n^{2,j}\rangle(t) &= \left(z\frac{\partial}{\partial z}\right)^j \phi^2(z,t)\Big|_{z=1} = \sum_{n=0}^{\infty} n^j \phi_n^2(t),\\
\langle n^{3,j}\rangle(t) &= \left(z\frac{\partial}{\partial z}\right)^j \phi^3(z,t)\Big|_{z=1} = \sum_{n=0}^{\infty} n^j \phi_n^3(t).
\end{aligned} \tag{5}$$

In this set of equations, the index $j$ indicates the order of the moment considered and the index $i = 1, 2, 3$ is the label of the state.

Before introducing the biological parameter space, which we postpone until the next section, let us rewrite the master equations in vector and matrix notation. The system (1) can be written as a unique vector equation,

$$\frac{d\boldsymbol{\phi}_n}{dt} = K[\boldsymbol{\phi}_{n-1} - \boldsymbol{\phi}_n] + \rho\big[(n+1)\boldsymbol{\phi}_{n+1} - n\boldsymbol{\phi}_n\big] + H\boldsymbol{\phi}_n, \tag{6}$$

where $\boldsymbol{\phi}_n(t) = (\phi_n^1(t), \phi_n^2(t), \phi_n^3(t))^T$ ($T$ denotes transpose). $K$ and $\rho$ are now diagonal $(3 \times 3)$-matrices, where the first contains the production rates of mRNA corresponding to each state of transcriptional efficiency, while $\rho$ is proportional to the identity matrix and contains the degradation rate of mRNA molecules. The matrix $H$ encodes the couplings between the states of the system and has the properties

$$H_{1j} + H_{2j} + H_{3j} = 0, \qquad H_{ii} \le 0, \qquad H_{ij} \ge 0 \quad \text{if } i \ne j. \tag{7}$$

Explicitly,

$$H = \begin{pmatrix} -h_1 & h_2\epsilon_2 & h_3(1-\epsilon_3) \\ h_1(1-\epsilon_1) & -h_2 & h_3\epsilon_3 \\ h_1\epsilon_1 & h_2(1-\epsilon_2) & -h_3 \end{pmatrix}. \tag{8}$$

The master equations in vector notation can also be transformed into a system of partial differential equations by introducing the vector generating function

$$\boldsymbol{\phi}(z,t) = \sum_{n=0}^{\infty} \boldsymbol{\phi}_n(t) z^n. \tag{9}$$

This leads to the system of partial differential equations

$$\frac{\partial\boldsymbol{\phi}}{\partial t} = (z-1)\left[K\boldsymbol{\phi} - \rho\frac{\partial\boldsymbol{\phi}}{\partial z}\right] + H\boldsymbol{\phi}. \tag{10}$$

The vector whose components are the probability distributions of each state is recovered through derivatives according to

$$\boldsymbol{\phi}_n(t) = \frac{1}{n!}\frac{\partial^n \boldsymbol{\phi}(z,t)}{\partial z^n}\bigg|_{z=0}, \tag{11}$$

and the moments are obtained from

$$\langle \boldsymbol{n}^j \rangle = \left(z\frac{\partial}{\partial z}\right)^j \boldsymbol{\phi}(z,t)\bigg|_{z=1}, \tag{12}$$

where $j$ is the order of the moment under consideration.

## 3 Biological Interpretation of the Parameters

The master equations of our model can be interpreted as describing a combination of two types of stochastic processes (van Kampen 1992). The first governs the production and the degradation of mRNA molecules, where the production is controlled by the rates $k_1, k_2, k_3$, and the degradation by the rate $\rho$, which is the same for all states. The second is a telegraphic process ruling the possible transitions between the states of the system. In what follows, we want to discuss the biological interpretation of the various parameters.

The first set of parameters allowing for a direct biological interpretation are the transcriptional efficiency parameters that govern the production-degradation process of mRNA, defined as the ratios $k_1/\rho$, $k_2/\rho$, $k_3/\rho$ between the production rates and the degradation rate. A large value of a certain $k_i/\rho$, in comparison to a certain $k_j/\rho$, say, means that the state $i$ is more efficient for mRNA production than the state $j$. Typically, one will find in the cell dozens of mRNA molecules when the gene is in an efficient state, but just a few or none when it is in an inefficient state.

The telegraphic process is controlled by the transition rates between the states and these are independent of the random variable $n$. However, they do depend on intrinsic properties of the gene, such as operator affinity to a specific regulatory protein, as well as on extrinsic properties, such as the availability of regulatory proteins, for instance. In order to decouple these two sources of dependence, we consider just the "telegraphic part" of the model, obtained by averaging the master equations (1) over the mRNA number $n$, or equivalently, by setting $z = 1$ in (3), resulting in:

$$\begin{aligned}
\frac{dP_1}{dt} &= -h_1 P_1 + h_2\epsilon_2 P_2 + h_3(1-\epsilon_3)P_3,\\
\frac{dP_2}{dt} &= h_1(1-\epsilon_1)P_1 - h_2 P_2 + h_3\epsilon_3 P_3,\\
\frac{dP_3}{dt} &= h_1\epsilon_1 P_1 + h_2(1-\epsilon_2)P_2 - h_3 P_3,
\end{aligned} \tag{13}$$

where each of the time dependent functions $P_i(t) = \phi^i(z=1,t)$ (with $i = 1,2,3$) represents the occupancy probability to find the gene operating in the respective mode

$i$, independently of the mRNA number $n$. Alternatively, we can obtain Eqs. (13) in vector form by setting $z = 1$ in the system (10):

$$\frac{d\boldsymbol{P}}{dt} = H\boldsymbol{P}, \tag{14}$$

where $\boldsymbol{P}(t) = (P_1(t), P_2(t), P_3(t))^T$ and $H$ is the coupling matrix as given in Eq. (8). The solution of (14) and a detailed description of the dynamical behavior of the moments and distributions of the model will be postponed to a second article, already in preparation. Here, we shall explore the properties of the coupling matrix $H$ and reinterpret its parameters in terms of the biological quantities of the model. The decisive feature that allows us to do so is that $H$ can be factorized into the product of two other matrices, namely

$$H = \epsilon h, \tag{15}$$

where

$$\epsilon = \begin{pmatrix} -1 & \epsilon_2 & 1-\epsilon_3 \\ 1-\epsilon_1 & -1 & \epsilon_3 \\ \epsilon_1 & 1-\epsilon_2 & -1 \end{pmatrix} \tag{16}$$

and

$$h = \begin{pmatrix} h_1 & 0 & 0 \\ 0 & h_2 & 0 \\ 0 & 0 & h_3 \end{pmatrix}. \tag{17}$$

Inspection of this decomposition shows that the matrix $\epsilon$ describes the structure and intensity of the couplings in the system under consideration, while the matrix $h$ contains the information about the occupancy of each state.

Indeed, the entries of the matrix $\epsilon$ will define which state is connected to which and how strong the coupling between them is. For instance, if $\epsilon_1 = 1$, state 1 will be strongly coupled to state 3, but will not affect state 2; however, state 2 may affect state 1, depending on the value of $\epsilon_2$. In the limit where all $\epsilon_i$ are zero, the coupling is cyclic ($1 \to 2 \to 3 \to 1$), when all $\epsilon_i$ are one it is anticyclic ($1 \to 3 \to 2 \to 1$), and if all $\epsilon_i$ are equal to $1/2$ all states are coupled symmetrically and with equal intensity.

To clarify the interpretation of the matrix $h$, let us write its entries in terms of the asymptotic occupancy probabilities $p_i = \lim_{t\to\infty} P_i(t)$. As is well known, the general solution of the system (14) has the form

$$\boldsymbol{P}(t) = \exp(tH)\boldsymbol{P}(0), \tag{18}$$

where exp is the matrix exponential, which is defined by the usual power series expansion

$$\exp(tH) = \sum_{n=0}^{\infty} \frac{(tH)^n}{n!}. \tag{19}$$

In particular, when the initial condition is an eigenvector of $H$, with eigenvalue $\lambda$, say, i.e.,

$$H\boldsymbol{P}_\lambda(0) = \lambda \boldsymbol{P}_\lambda(0), \tag{20}$$

the solution becomes

$$\boldsymbol{P}_\lambda(t) = \exp(\lambda t)\boldsymbol{P}_\lambda(0). \tag{21}$$

However, this special type of solution is in general not compatible with the requirement that the total probability should be equal to 1, i.e., that

$$P_1(t) + P_2(t) + P_3(t) = 1 \quad \text{for all } t, \tag{22}$$

except when $\lambda = 0$, in which case the solution is constant. Therefore, the information about the asymptotic state(s) is encoded in the eigenspace corresponding to the zero eigenvalue, or kernel, of the matrix $H$: note that the properties of $H$, as specified in Eq. (7), guarantee that this kernel is nontrivial. In fact, explicit calculation shows that $H$ annihilates the vector

$$\boldsymbol{q} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} h_3 h_2 \delta_1 \\ h_1 h_3 \delta_2 \\ h_2 h_1 \delta_3 \end{pmatrix}, \tag{23}$$

where the $\delta_i$ are the diagonal elements of the cofactor matrix of the matrix $\epsilon$ ($\delta_i = C_{ii}(\epsilon)$ for $i = 1, 2, 3$); explicitly:

$$\begin{aligned} \delta_1 &= 1 - \epsilon_3 + \epsilon_3\epsilon_2, \\ \delta_2 &= 1 - \epsilon_1 + \epsilon_1\epsilon_3, \\ \delta_3 &= 1 - \epsilon_2 + \epsilon_2\epsilon_1. \end{aligned} \tag{24}$$

It may be interesting to observe that the components of the vector $\boldsymbol{q}$ are precisely the diagonal elements of the cofactor matrix of the entire matrix $H$ ($q_i = C_{ii}(H)$ for $i = 1, 2, 3$). When properly normalized, the vector $\boldsymbol{q}$ provides the asymptotic occupancy probability vector

$$\boldsymbol{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} q_1/q \\ q_2/q \\ q_3/q \end{pmatrix}, \tag{25}$$

where the scalar $q$ is the product of the vector $\boldsymbol{q}$ with the left eigenvector $\boldsymbol{u}^T = (1, 1, 1)$ of $H$ corresponding to the zero eigenvalue ($\boldsymbol{u}^T H = 0$), given by

$$q = \boldsymbol{u}^T \boldsymbol{q} = q_1 + q_2 + q_3 = h_3 h_2 \delta_1 + h_1 h_3 \delta_2 + h_2 h_1 \delta_3. \tag{26}$$

Since the matrix $h$ has three entries and the asymptotic occupancy probabilities $p_i$, being subject to the constraint

$$p_1 + p_2 + p_3 = 1, \tag{27}$$

contain only two free parameters, it is clear that we need one more parameter, which allows for a biological interpretation. This parameter will be the trace of $h$, or (up to a sign) of $H$, which we shall denote by $\Sigma$:

$$\Sigma = h_1 + h_2 + h_3. \tag{28}$$

Moreover, we introduce the quantity

$$\Delta = \delta_1 p_2 p_3 + p_1 \delta_2 p_3 + p_1 p_2 \delta_3. \tag{29}$$

Then an elementary algebraic calculation gives the following expressions for the entries of the matrix $h$ and for the scalar $q$:

$$\begin{aligned} h_1 &= \Sigma \delta_1 p_2 p_3 / \Delta, \\ h_2 &= \Sigma p_1 \delta_2 p_3 / \Delta, \\ h_3 &= \Sigma p_1 p_2 \delta_3 / \Delta, \end{aligned} \tag{30}$$

$$q = \frac{\Sigma^2 \delta_1 \delta_2 \delta_3 p_1 p_2 p_3}{\Delta^2}. \tag{31}$$

In this way, the entries of the matrix $h$ are expressed in terms of its trace $\Sigma$, in terms of the parameters $\delta_1, \delta_2, \delta_3$ which, according to Eq. (24), are determined by the entries of the matrix $\epsilon$ that encodes the couplings between the states, and in terms of the parameters $p_1, p_2, p_3$ which, as stated before, are the asymptotic occupancy probabilities of each state. These asymptotic occupancy probabilities depend on the environmental conditions and are classified as extrinsic properties. In the case of the lac operon they are functions of the concentrations of glucose and lactose present in the medium.

In order to discuss the meaning of the parameter $\Sigma$, consider the two nonzero eigenvalues of the matrix $H$, which can be obtained by calculating its characteristic polynomial: in terms of the parameters introduce above, the result is

$$\chi_H(\lambda) = \det(\lambda 1 - H) = \lambda^3 + \Sigma \lambda^2 + q\lambda, \tag{32}$$

so that $\chi_H(\lambda) = \lambda(\lambda + \lambda^+)(\lambda + \lambda^-)$ with

$$\lambda^\pm = \frac{1}{2}\left(\Sigma \pm \sqrt{\Sigma^2 - 4q}\right),$$

or explicitly

$$\lambda^\pm = \frac{\Sigma}{2}\left(1 \pm \sqrt{1 - \frac{4\delta_1 \delta_2 \delta_3 p_1 p_2 p_3}{\Delta^2}}\right). \tag{33}$$

Denoting by

$$\boldsymbol{v}^\pm = \begin{pmatrix} v_1^\pm \\ v_2^\pm \\ v_3^\pm \end{pmatrix} \tag{34}$$

the corresponding (unnormalized) eigenvectors, determined uniquely up to some constant factor by the condition that

$$H\boldsymbol{v}^{\pm} = -\lambda^{\pm}\boldsymbol{v}^{\pm}, \tag{35}$$

we can write down the general solution (18) of the system (14) in the form

$$\boldsymbol{P}(t) = \boldsymbol{p} + \exp\bigl(-\lambda^{+}t\bigr)\boldsymbol{v}^{+} + \exp\bigl(-\lambda^{-}t\bigr)\boldsymbol{v}^{-}. \tag{36}$$

We shall refrain from writing out the components of the eigenvectors $\boldsymbol{v}^{\pm}$ explicitly, since the corresponding expressions are cumbersome and not particularly enlightening, mentioning only that the same argument as in the proof of Eq. (26) shows that they sum up to zero:

$$v_1^{\pm} + v_2^{\pm} + v_3^{\pm} = 0. \tag{37}$$

Together, Eqs. (27) and (37) guarantee that the condition (22) is satisfied.

From this analysis, we see that $\Sigma$ is an important parameter governing the dynamics of the system. In fact, the model has two principal time scales: $1/\rho$, the average lifetime of an mRNA molecule, and $1/\Sigma$, which by the definition of the matrix $H$ (see (8)) provides a typical timescale for the gene switch, i.e., the transition between its different states of transcriptional activity. More precisely, when expressed in purely numerical terms, it is the ratio $\Sigma/\rho$ that measures the speed of switching, as compared to the lifetime of an mRNA molecule, which serves as a reference scale. For large values of this ratio, we say that the switch is *fast*, while for small values, we say that it is *slow*. But to a certain extent, the adjectives "fast" and "slow" also refer to the rate with which the probability distribution, averaged over the mRNA number $n$, approaches its equilibrium configuration. Indeed, that rate is given by the smallest positive eigenvalue of $-H$, i.e., by $\lambda^{-}$, which according to (33) can take any value between 0 and $\Sigma/2$: this means that the system's approach to equilibrium cannot occur on a shorter timescale than $1/\Sigma$, but it may of course occur on a much longer one.

We conclude this section by writing out the explicit form of the matrix $H$ given by (8), factorized according to Eq. (15), in terms of the biological parameters:

$$H = \frac{\Sigma}{\Delta}\begin{pmatrix} -1 & \epsilon_2 & (1-\epsilon_3) \\ (1-\epsilon_1) & -1 & \epsilon_3 \\ \epsilon_1 & (1-\epsilon_2) & -1 \end{pmatrix}\begin{pmatrix} \delta_1 p_2 p_3 & 0 & 0 \\ 0 & p_1\delta_2 p_3 & 0 \\ 0 & 0 & p_1 p_2 \delta_3 \end{pmatrix}. \tag{38}$$

## 4 Asymptotic Properties of the Model

In this section, we want to explore the master equations (3) for the generating functions in a different direction: rather than evaluating them at $z = 1$ but retaining the time dependence in order to obtain information about the time evolution of averages, we shall study the asymptotic limit $t \to \infty$ where we can neglect all time derivatives but retain the $z$-dependence in order to obtain the stationary multi-modal probability distributions and their moments.

### 4.1 Solution of the Steady-State Limit of the Model

In the stationary limit where all time derivatives can be neglected, the master equations (3), written in vector and matrix notation as in Eq. (10), become

$$\rho(z-1)\frac{d\boldsymbol{\phi}}{dz} = (z-1)K\boldsymbol{\phi} + H\boldsymbol{\phi}, \tag{39}$$

where now $\boldsymbol{\phi}(z) = (\phi^1(z), \phi^2(z), \phi^3(z))^T$, $H$ is the coupling matrix and $K$ contains in its diagonal the production rates of mRNA corresponding to each state of transcriptional efficiency, as before. Performing a change of variables by introducing the new variable $w = z - 1$, we obtain

$$\rho w\frac{d\boldsymbol{\phi}}{dw} = wK\boldsymbol{\phi} + H\boldsymbol{\phi}, \tag{40}$$

which we can solve by means of a power series Ansatz

$$\boldsymbol{\phi}(w) = \sum_{n=0}^{\infty} \boldsymbol{\phi}_n w^n. \tag{41}$$

Indeed, substituting in Eq. (40) and collecting the terms, we arrive at the following recursion relations:

$$H\boldsymbol{\phi}_0 = 0, \qquad \boldsymbol{\phi}_n = (n\rho 1 - H)^{-1} K\boldsymbol{\phi}_{n-1} \quad \text{for } n > 0. \tag{42}$$

The first few terms can be computed explicitly. For example, $\boldsymbol{\phi}_0$ is just the value of $\boldsymbol{\phi}(w)$ at $w = 0$, which is equal to the vector $\boldsymbol{p} = \lim_{t\to\infty} \boldsymbol{P}(t)$ of asymptotic occupancy probabilities already introduced in the previous section; note that this does satisfy the first condition in Eq. (42), thus starting the recursion. Similarly, taking the first two derivatives and using Eq. (12), we can express $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ in terms of the moment vectors $\boldsymbol{n}^1$ and $\boldsymbol{n}^2$ to obtain

$$\boldsymbol{\phi}(w) = \boldsymbol{p} + \langle \boldsymbol{n}^1 \rangle w + \frac{1}{2}\big(\langle \boldsymbol{n}^2 \rangle - \langle \boldsymbol{n}^1 \rangle\big)w^2 + \mathrm{O}\big(w^3\big). \tag{43}$$

The same goes for the higher order terms, so the power series expansion in the variable $w$ is just another form of the moment expansion of the asymptotic probability distributions. Moreover, it is clear that the second condition in Eq. (42) completely determines all coefficients $\boldsymbol{\phi}_n$ in terms of $\boldsymbol{p}$. To do so, it is convenient to find a more explicit expression for the inverse of the matrix $U = n\rho 1 - H$: this can be done by using the fact that any matrix $U$ satisfies its own characteristic equation, i.e., we have

$$U^3 - \mathrm{tr}(U)U^2 + \big(C_{11}(U) + C_{22}(U) + C_{33}(U)\big)U - \det(U)1 = 0, \tag{44}$$

where the $C_{ii}(U)$ are the diagonal elements of the cofactor matrix of the matrix $U$. This formula implies that if $\det(U) \neq 0$, then $U$ is invertible, and its inverse is

$$U^{-1} = \frac{1}{\det(U)}\big(U^2 - \mathrm{tr}(U)U + \big(C_{11}(U) + C_{22}(U) + C_{33}(U)\big)\big). \tag{45}$$

For $H$ itself, Eq. (44) becomes $H^3 + \Sigma H^2 + qH = 0$, but since $H$ is not invertible, this does not imply $H^2 + \Sigma H + q1 = 0$; rather, it can be shown by a somewhat tedious calculation, using Eqs. (8) and (23)–(28), that

$$H^2 = -\Sigma H - q(1 - P), \tag{46}$$

where $P$ denotes a matrix obtained by repeating the entries of the vector $\boldsymbol{p}$ along the columns,

$$P = \begin{pmatrix} p_1 & p_1 & p_1 \\ p_2 & p_2 & p_2 \\ p_3 & p_3 & p_3 \end{pmatrix}, \tag{47}$$

so as to guarantee that $HP = 0$. Now taking $U = n\rho 1 - H$, we compute $\det(U) = n\rho(n\rho + \lambda^+)(n\rho + \lambda^-) = n\rho(n^2\rho^2 + n\rho\Sigma + q)$, $\mathrm{tr}(U) = 3n\rho + \Sigma$ and $C_{11}(U) + C_{22}(U) + C_{33}(U) = 3n^2\rho^2 + 2n\rho\Sigma + q$ to obtain

$$(n\rho 1 - H)^{-1} = \frac{n\rho(n\rho + \Sigma)1 + qP + n\rho H}{n\rho\left(n^2\rho^2 + n\rho\Sigma + q\right)}. \tag{48}$$

As a result, the recursion relation (42) assumes the form

$$\boldsymbol{\phi}_0 = \boldsymbol{p}, \qquad \boldsymbol{\phi}_n = \left(\frac{n\rho(n\rho + \Sigma)1 + qP + n\rho H}{n\rho\left(n^2\rho^2 + n\rho\Sigma + q\right)} K\right)\boldsymbol{\phi}_{n-1} \quad \text{for } n > 0. \tag{49}$$

Thus, given the transcriptional efficiency parameters of the gene (entries of the matrix $K/\rho$), the couplings between the states (entries of the matrix $\epsilon$) and the asymptotic occupancy probabilities (components of the vector $\boldsymbol{p}$), we use the recursion relation (49) to compute all the coefficients of the power series expansion of the function $\boldsymbol{\phi}(w)$ around $w = 0$ and, from that, the function $\boldsymbol{\phi}(w)$ itself. Even after that has been done, we must still substitute $w$ by $z - 1$ and evaluate at $z = 0$, since the asymptotic probability distributions of the system are the coefficients of the power series expansion of the function $\boldsymbol{\phi}(z)$ around $z = 0$: more concretely, we may recall that the coefficient of $z^n$ in this expansion is a vector whose $i$th component gives the probability to find the system in the $i$th state of transcriptional activity with $n$ molecules of mRNA in the cell. To perform this last step, we may either use the binomial expansion for $(z - 1)^n$ and collect the terms in powers of $z$ or else apply Eq. (11) to evaluate the derivatives at $z = 0$, which corresponds to $w = -1$. However, in contrast to the asymptotic probability distributions themselves, their moments are much easier to obtain, namely, directly from the original expansion around $w = 0$, as is already evident from Eq. (12) and was used explicitly in the derivation of Eq. (43).

### 4.2 Multimodal Probability Distributions

In this subsection, we show numerical examples and graphical representations of probability distributions calculated according to Eq. (49). In all cases that follow, the transcriptional efficiency parameters follow the hierarchy $k_1 > k_2 > k_3$; the numerical values adopted being $k_1/\rho = 40$, $k_2/\rho = 20$ and $k_3/\rho = 5$. The entries of the coupling matrix $\epsilon$ were all set equal to $1/2$, corresponding to a scheme where all states are

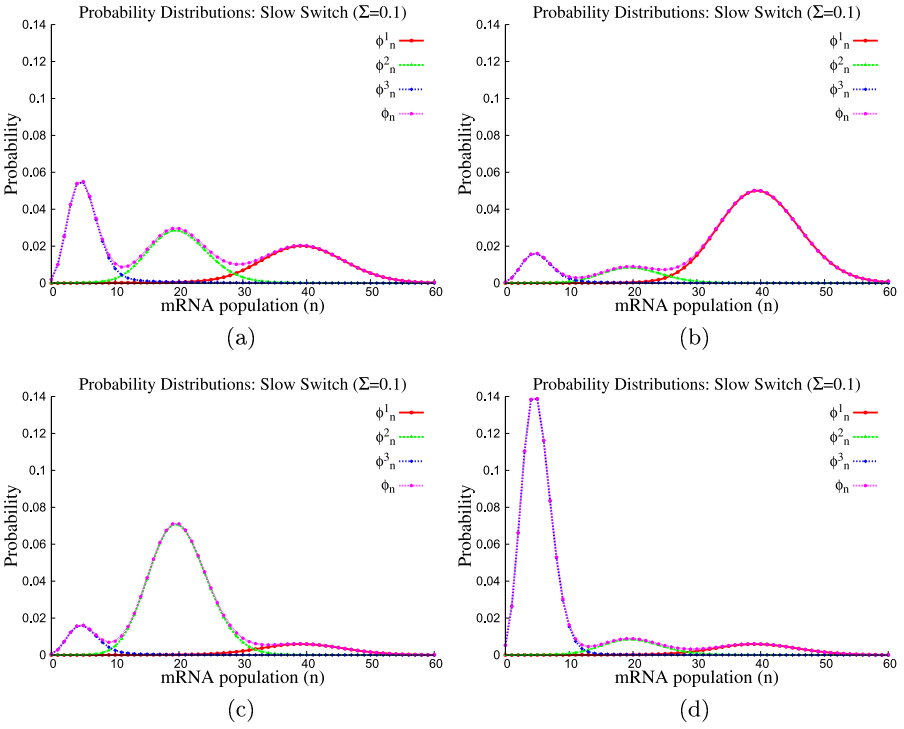



**Fig. 2** Behavior of the partial and total probability distributions in the regime of slow switch, $\Sigma = 0.1$, with mRNA degradation time scale $\rho = 1$, transcriptional efficiency parameters $k_1 = 40$, $k_2 = 20$, $k_3 = 5$ and couplings $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.5$. (**a**) $p_1 = p_2 = p_3 = 1/3$. (**b**) $p_1 = 0.8$, $p_2 = p_3 = 0.1$. (**c**) $p_2 = 0.8$, $p_1 = p_3 = 0.1$. (**d**) $p_3 = 0.8$, $p_1 = p_2 = 0.1$

coupled symmetrically and with equal intensity. All calculations were performed for two flexibility regimes of the gene, namely for slow switch ($\Sigma = 0.1$) and for fast switch ($\Sigma = 10$), using four distinct asymptotic configurations, characterized by the following occupancy probability vectors: $\boldsymbol{p} = (1/3, 1/3, 1/3)^T$, $\boldsymbol{p} = (0.8, 0.1, 0.1)^T$, $\boldsymbol{p} = (0.1, 0.8, 0.1)^T$, and $\boldsymbol{p} = (0.1, 0.1, 0.8)^T$.

In Fig. 2, we exhibit the probability distributions for each mode, $\phi_n^1, \phi_n^2, \phi_n^3$, as well as the total one, $\phi_n = \phi_n^1 + \phi_n^2 + \phi_n^3$, as functions of $n$, for the regime of slow switch: $\Sigma = 0.1$. In the first panel, Fig. 2(a), the occupancy probabilities are all equal ($p_1 = p_2 = p_3 = 1/3$). Each operational mode shows a pronounced peak, which leads to a three peaks structure for the total probability distribution, indicating that the system is operating in all states to produce mRNA. In Fig. 2(b), the values of the occupancy probabilities are $p_1 = 0.8$ and $p_2 = p_3 = 0.1$. One can note a dominant peak centered around $n = 40$ and two residual peaks around $n \approx 20$ and $n \approx 5$, indicating that mRNA is produced predominantly in the state of maximum efficiency. In Fig. 2(c), we have $p_2 = 0.8$ and $p_1 = p_3 = 0.1$: here, the system is more likely to be found producing in its intermediate state, as can be seen by noting the strong peak around $n = 20$. Finally, in Fig. 2(d), we encounter the scenario of basal production, with $p_3 = 0.8$ and $p_1 = p_2 = 0.1$. In comparison, the differences in the behavior of

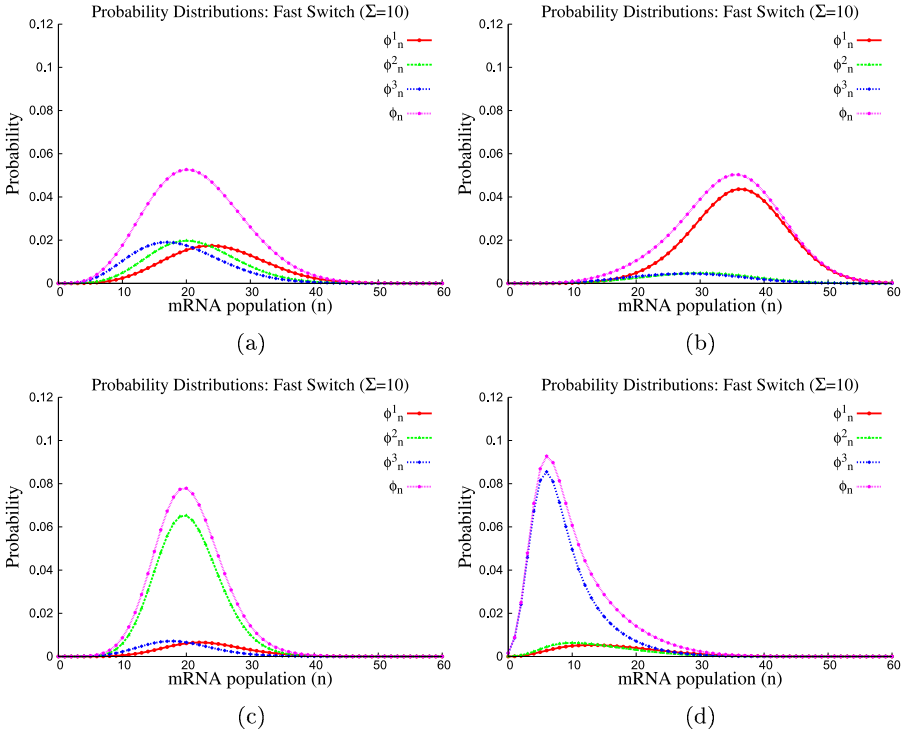



**Fig. 3** Behavior of the partial and total probability distributions in the regime of fast switch, $\Sigma = 10$, with mRNA degradation time scale $\rho = 1$, transcriptional efficiency parameters $k_1 = 40$, $k_2 = 20$, $k_3 = 5$ and couplings $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.5$. (**a**) $p_1 = p_2 = p_3 = 1/3$. (**b**) $p_1 = 0.8$, $p_2 = p_3 = 0.1$. (**c**) $p_2 = 0.8$, $p_1 = p_3 = 0.1$. (**d**) $p_3 = 0.8$, $p_1 = p_2 = 0.1$

the occupancy probabilities as one passes from one panel to the other reveal a transition from a multi-peak probability distribution to a Poisson like structure. Obviously, when multimodality predominates, the moments of the distribution commonly used to describe the system in some approximation scheme, such as mean value and variance, are by themselves insufficient to describe the multi-peak structure.

In Fig. 3, we show the same probability distributions for the same values of the parameters (transcriptional efficiency, couplings, and occupancy probabilities) as in Fig. 2, except that we are now in the regime of fast switch: $\Sigma = 10$. Comparison of the panels in Fig. 3 with the corresponding ones in Fig. 2 reveals that when the ability of the gene to reach the equilibrium configuration is increased, the peaks of the individual modes tend to decrease and to be dislocated toward a common mean. In particular, comparing Fig. 3(a) with Fig. 2(a), we see that $\phi_n^1$ is now centered around $n \approx 28$ ,and $\phi_n^3$ is now centered around $n \approx 12$, the first having been dislocated from $n \approx 40$ and the second from $n \approx 5$; as a result, the pronounced three-peak structure of the total probability distribution has completely disappeared. The remaining panels, representing situations of predominance of one operational mode over the other two, show similar but less pronounced effects.

For extreme values of $\Sigma$, two principal consequences emerge. When $\Sigma \to 0$, the states decouple (this can be seen by inspection of the coupling matrix), while for $\Sigma \to \infty$, the effect is to diminish the fluctuations in the system, which drives all probability distributions toward a Poisson like structure. In order to make this last affirmation more precise, we pass to the next subsection, where we present the moments of the distributions in the stationary regime.

### 4.3 Mean Value and Noise

In this subsection, we shall derive explicit expressions for the first two moments of the probability distributions: these can be inferred directly from the first three terms of the power series expansion (41) of the generating function $\boldsymbol{\phi}(w)$ at $w = 0$ (see Eq. (43)), whose coefficients are determined by the recursion relation (49).

We start with the first moment vector $\langle \boldsymbol{n}^1 \rangle$, given by

$$\langle \boldsymbol{n}^1 \rangle = \left. \frac{d\boldsymbol{\phi}}{dw} \right|_{w=0} = \boldsymbol{\phi}_1,$$

or explicitly

$$\langle \boldsymbol{n}^1 \rangle = \frac{\rho(\rho + \Sigma)1 + qP + \rho H}{\rho(\rho^2 + \rho\Sigma + q)} K\boldsymbol{p}. \tag{50}$$

We shall also call it the mean value vector because its components provide the mean mRNA production in each operational mode of the system. It is the sum of three terms: the first term is

$$\frac{\rho + \Sigma}{\rho^2 + \rho\Sigma + q} K\boldsymbol{p} = \frac{\rho + \Sigma}{\rho^2 + \rho\Sigma + q} \begin{pmatrix} k_1 p_1 \\ k_2 p_2 \\ k_3 p_3 \end{pmatrix}, \tag{51}$$

the second term is

$$\frac{q}{\rho(\rho^2 + \rho\Sigma + q)} PK\boldsymbol{p} = \frac{q}{\rho(\rho^2 + \rho\Sigma + q)} (k_1 p_1 + k_2 p_2 + k_3 p_3) \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}, \tag{52}$$

and the third term is

$$\frac{1}{\rho^2 + \rho\Sigma + q} HK\boldsymbol{p}, \tag{53}$$

which cannot be substantially simplified. What is more significant, however, is the total mean value $\langle n^1 \rangle$, since this is a quantity commonly measured in the experiments: it is obtained by summing the components of the mean value vector $\langle \boldsymbol{n}^1 \rangle$, or equivalently, by taking its scalar product with the left eigenvector $\boldsymbol{u} = (1, 1, 1)^T$ of the coupling matrix $H$ corresponding to the zero eigenvalue: $\langle n^1 \rangle = \boldsymbol{u}^T \langle \boldsymbol{n}^1 \rangle$. There is then no contribution from the third term, and taking into account that $p_1 + p_2 + p_3 = 1$, we are left with the following final result:

$$\langle n^1 \rangle = \frac{1}{\rho}(k_1 p_1 + k_2 p_2 + k_3 p_3). \tag{54}$$

Clearly, this quantity is independent of the switch parameter $\Sigma$. However, the components of the vector $\langle \boldsymbol{n}^1 \rangle$ do depend on $\Sigma$. In particular, as was already observed in the last section, the system behaves very differently when we go to the two extreme values of $\Sigma$. When $\Sigma \to 0$ (slow switch), the states decouple completely and the mean values of the modes will be given by the components of the vector

$$\lim_{\Sigma \to 0} \langle \boldsymbol{n}^1 \rangle = \frac{1}{\rho} \begin{pmatrix} k_1 p_1 \\ k_2 p_2 \\ k_3 p_3 \end{pmatrix}. \tag{55}$$

On the other hand, when $\Sigma \to \infty$ (fast switch), they will be given by the components of the vector

$$\lim_{\Sigma \to \infty} \langle \boldsymbol{n}^1 \rangle = \frac{\langle n^1 \rangle}{\rho} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}. \tag{56}$$

Passing to the second moment vector $\langle \boldsymbol{n}^2 \rangle$, we have

$$\boldsymbol{\phi}_2 = \frac{1}{2} \frac{d^2 \boldsymbol{\phi}}{dw^2} \bigg|_{w=0} = \frac{1}{2} \big( \langle \boldsymbol{n}^2 \rangle - \langle \boldsymbol{n}^1 \rangle \big).$$

Taking into account that $\boldsymbol{\phi}_1 = \langle \boldsymbol{n}^1 \rangle$, this can be solved to give

$$\langle \boldsymbol{n}^2 \rangle = 2\boldsymbol{\phi}_2 + \boldsymbol{\phi}_1,$$

Again, we are mainly interested in the total second moment $\langle n^2 \rangle$, which is the sum of the components of the second moment vector $\langle \boldsymbol{n}^2 \rangle$, or equivalently, $\langle n^2 \rangle = \boldsymbol{u}^T \langle \boldsymbol{n}^2 \rangle$, since this determines the total variance $\sigma^2$ of the system, defined as

$$\sigma^2 = \langle n^2 \rangle - \langle n^1 \rangle^2. \tag{57}$$

Thus,

$$\langle n^2 \rangle = 2\boldsymbol{u}^T \boldsymbol{\phi}_2 + \boldsymbol{u}^T \boldsymbol{\phi}_1. \tag{58}$$

The second term in this expression is the total mean value $\langle n^1 \rangle$, which has already been calculated, whereas using the fact that $\boldsymbol{u}^T P = \boldsymbol{u}^T$ (which follows from the condition $p_1 + p_2 + p_3 = 1$) and $\boldsymbol{u}^T H = 0$, we see that the first term becomes

$$2\boldsymbol{u}^T \boldsymbol{\phi}_2 = \boldsymbol{u}^T \frac{2\rho(2\rho + \Sigma)1 + q P + 2\rho H}{\rho\big(4\rho^2 + 2\rho\Sigma + q\big)} K \boldsymbol{\phi}_1 = \frac{1}{\rho} \boldsymbol{u}^T K \boldsymbol{\phi}_1. \tag{59}$$

Inserting the expression (50) for $\boldsymbol{\phi}_1 = \langle \boldsymbol{n}^1 \rangle$, we get

$$\frac{1}{\rho} \boldsymbol{u}^T K \boldsymbol{\phi}_1 = \boldsymbol{u}^T K \frac{\rho(\rho + \Sigma)1 + q P + \rho H}{\rho^2\big(\rho^2 + \rho\Sigma + q\big)} K \boldsymbol{p}. \tag{60}$$

Once again, this is the sum of three terms: the first term is

$$\frac{\rho + \Sigma}{\rho(\rho^2 + \rho\Sigma + q)} \boldsymbol{u}^T K^2 \boldsymbol{p} = \frac{\rho + \Sigma}{\rho(\rho^2 + \rho\Sigma + q)} \big( k_1^2 p_1 + k_2^2 p_2 + k_3^2 p_3 \big), \tag{61}$$

the second term is

$$\frac{q}{\rho^2(\rho^2+\rho\Sigma+q)}\boldsymbol{u}^T K P K \boldsymbol{p} = \frac{q\langle n^1\rangle^2}{\rho^2+\rho\Sigma+q}, \tag{62}$$

and the third term is a rather complicated expression involving the coupling matrix $H$, which no longer vanishes. To calculate it, we use the product decomposition (15) and Eq. (30) to rewrite $H$ in the form

$$H = \frac{\Sigma}{\Delta}\epsilon\delta\pi \tag{63}$$

where $\delta$ and $\pi$ are diagonal matrices defined by

$$\delta = \begin{pmatrix} \delta_1 & 0 & 0 \\ 0 & \delta_2 & 0 \\ 0 & 0 & \delta_3 \end{pmatrix}, \qquad \pi = \begin{pmatrix} p_2p_3 & 0 & 0 \\ 0 & p_1p_3 & 0 \\ 0 & 0 & p_1p_2 \end{pmatrix}. \tag{64}$$

We also set

$$\boldsymbol{k} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}, \tag{65}$$

so that

$$\boldsymbol{u}^T K = \boldsymbol{k}^T, \qquad \pi K \boldsymbol{p} = p_1p_2p_3\boldsymbol{k}, \tag{66}$$

and hence the third term referred to above becomes

$$\frac{1}{\rho(\rho^2+\rho\Sigma+q)}\boldsymbol{u}^T K H K \boldsymbol{p} = \frac{1}{\rho(\rho^2+\rho\Sigma+q)}\frac{\Sigma}{\Delta}p_1p_2p_3\boldsymbol{k}^T\epsilon\delta\boldsymbol{k}, \tag{67}$$

with the scalar product

$$\boldsymbol{k}^T\epsilon\delta\boldsymbol{k} = -\delta_1k_1^2 - \delta_2k_2^2 - \delta_3k_3^2 + \Gamma^3_{1,2}k_1k_2 + \Gamma^2_{1,3}k_1k_3 + \Gamma^1_{2,3}k_2k_3, \tag{68}$$

where the $\Gamma$-coefficients can be expressed entirely in terms of linear combinations of the $\delta$'s:

$$\begin{aligned} \Gamma^3_{1,2} &= \delta_1 + \delta_2 - \delta_3, \\ \Gamma^2_{1,3} &= \delta_1 + \delta_3 - \delta_2, \\ \Gamma^1_{2,3} &= \delta_2 + \delta_3 - \delta_1. \end{aligned} \tag{69}$$

Putting everything together, we arrive at the following formula for the total variance in terms of the biological parameters of the system:

$$\begin{aligned} \sigma^2 = \langle n^1\rangle + \frac{\rho+\Sigma}{\rho(\rho^2+\rho\Sigma+q)}\big[&\big(k_1^2p_1 + k_2^2p_2 + k_3^2p_3\big) \\ &- (k_1p_1 + k_2p_2 + k_3p_3)^2\big] \end{aligned}$$

$$- \frac{1}{\rho(\rho^2 + \rho\Sigma + q)} \frac{\Sigma}{\Delta} p_1 p_2 p_3 \big[\delta_1 k_1^2 + \delta_2 k_2^2 + \delta_3 k_3^2$$
$$- \Gamma^3_{1,2} k_1 k_2 - \Gamma^2_{1,3} k_1 k_3 - \Gamma^1_{2,3} k_2 k_3\big]. \tag{70}$$

The total fluctuation of the system is simply the square root $\sigma$ of this expression.

Another important quantity widely used to classify the noise in gene networks is the *Fano factor*, which is simply the variance normalized by the mean value, i.e., the expression

$$\frac{\sigma^2}{\langle n^1 \rangle} = \frac{\langle n^2 \rangle - \langle n^1 \rangle^2}{\langle n^1 \rangle}. \tag{71}$$

It is used to measure the extent to which a given probability distribution deviates from a Poisson distribution. For values of the Fano factor $<1$, one says that the distribution is *sub-Poissonian* while for values $>1$, one says that the distribution is *super-Poissonian*; finally, if the Fano factor is $= 1$, the distribution is called *Poissonian*.

With the explicit formulas for the mean value and the total fluctuation (hence also for the Fano factor) at our disposal, we are in a position to describe some typical features of the behavior of the system. Consider, for example, the situation where it operates in only one of its states, so we have $p_1 = 1$ or $p_2 = 1$ or $p_3 = 1$ while the other two asymptotic occupancy probabilities vanish. (Mathematically, this situation has to be handled by taking the limit in which one of the three asymptotic occupancy probabilities approaches 1 while the other two approach 0, according to $p_i = (1 - \lambda)p'_i$ and $p_j = \lambda p'_j$ for $j \neq i$, with $p'_i$ and $p'_j$ fixed and $\lambda \to 0$, say, keeping the parameters $\rho$, $\Sigma$, $k_1$, $k_2$, $k_3$, $\epsilon_1$, $\epsilon_2$, $\epsilon_3$—and hence also $\delta_1$, $\delta_2$, $\delta_3$—fixed; this implies that $\Delta$ goes to zero linearly with $\lambda$ while $q$ goes to some finite non-zero value.) Then it is clear that the last two terms in Eq. (70) are zero and the fluctuation is equal to the total mean value: this implies that the Fano factor is equal to 1 and so the distribution is Poissonian. Moreover, in this case, the lowest level of noise is reached when the gene operates in the basal mode ($p_3 = 1$), with fluctuation $k_3/\rho$. More generally, the same argument shows that whenever one of the asymptotic occupancy probabilities vanishes, the last term in Eq. (70) is zero, and hence the fluctuations are governed by the quadratic terms appearing in the second term of the same equation. In a similar fashion, we can study the dependence on the parameter $\Sigma$, assuming now that all other parameters (including $p_1$, $p_2$, $p_3$) are fixed. To do so, note that the denominator of the last two terms in Eq. (70) contains a linear term in $q$ and that according to Eq. (31), $q$ is quadratic in $\Sigma$, whereas the numerator of the last two terms in Eq. (70) is linear in $\Sigma$. Therefore, as $\Sigma \to \infty$, the noise in the system decreases as $1/\Sigma$ down to its limit value, which is $\langle n^1 \rangle$ for the variance or 1 for the Fano factor. This means that increasing the speed with which the gene is able to reach its equilibrium configuration decreases the level of noise in the system and drives the probability distribution to a Poissonian one.

In Figs. 4 and 5 below, we exhibit graphically the behavior of the fluctuation and of the Fano factor in the regimes of slow switch ($\Sigma = 0, 1$) and fast switch ($\Sigma = 10$), respectively, as functions of the asymptotic occupancy probabilities $p_1$ and $p_3$, which are indicated along the axes. The analytical expressions underlying these graphs are

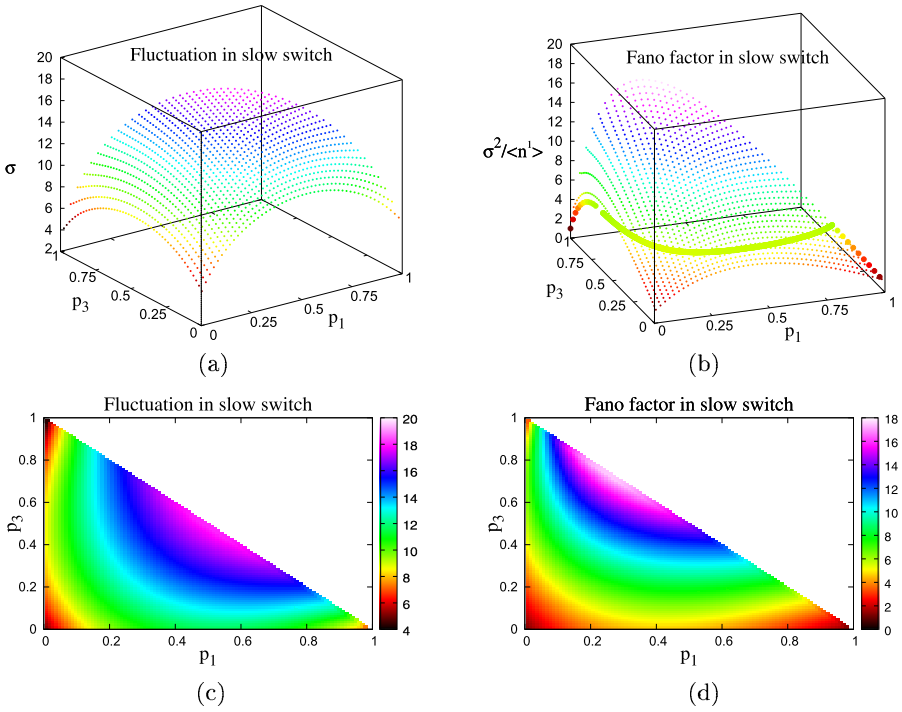



**Fig. 4** Behavior of the total fluctuation $\sigma$ and the Fano factor $\sigma^2/\langle n^1\rangle$ in the regime of slow switch, $\Sigma = 0.1$, with mRNA degradation time scale $\rho = 1$, transcriptional efficiency parameters $k_1 = 40$, $k_2 = 20$, $k_3 = 5$ and couplings $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.5$. (**a**) 3-d plot of fluctuation. (**b**) 3-d plot of Fano factor. (**c**) Density plot of fluctuation. (**d**) Density plot of Fano factor. A path connecting the state $p_3 = 1$ to the state $p_1 = 1$ and avoiding the "mountain pass problem" is shown in (**b**). Along this path, the Fano factor remains $< 6$, while along a direct path that does not use the intermediate state (i.e., with $p_2$ remaining equal to zero), it would reach values between 17 and 18

given by Eqs. (70) and (71), with the same choices for the mRNA degradation time scale ($\rho = 1$), the transcriptional efficiency parameters ($k_1 = 40$, $k_2 = 20$, $k_3 = 5$) and the entries of the coupling matrix $\epsilon$ (all $\epsilon_i = 1/2$) as in the preceding section.

In Figs. 4(a) and 5(a), we show the fluctuation and in Figs. 4(b) and 5(b) the Fano factor, in the form of 3-d plots. Similarly, Figs. 4(c) and 5(c) exhibit the fluctuation and Figs. 4(d) and 5(d) the Fano factor in the form of 2-d density plots, following the color scheme indicated to the right of each panel. In all cases, the graphs show clearly that the region of highest noise level is the one where the gene operates predominantly in the two extreme states (1 or 3). Remarkably, the possibility of occupying an intermediate state (2), that is, increasing $p_2$, leads to a significant noise reduction! Moreover, comparing the panels in Fig. 4 with the corresponding ones in Fig. 5 shows that this behavior is more pronounced in the slow switch regime, since in the fast switch regime, noise is globally reduced and all probability distributions become closer to Poissonian ones.

Finally, it should not go unnoticed that even though, as mentioned before, the lowest noise level is reached when the gene operates in the basal mode, this goes

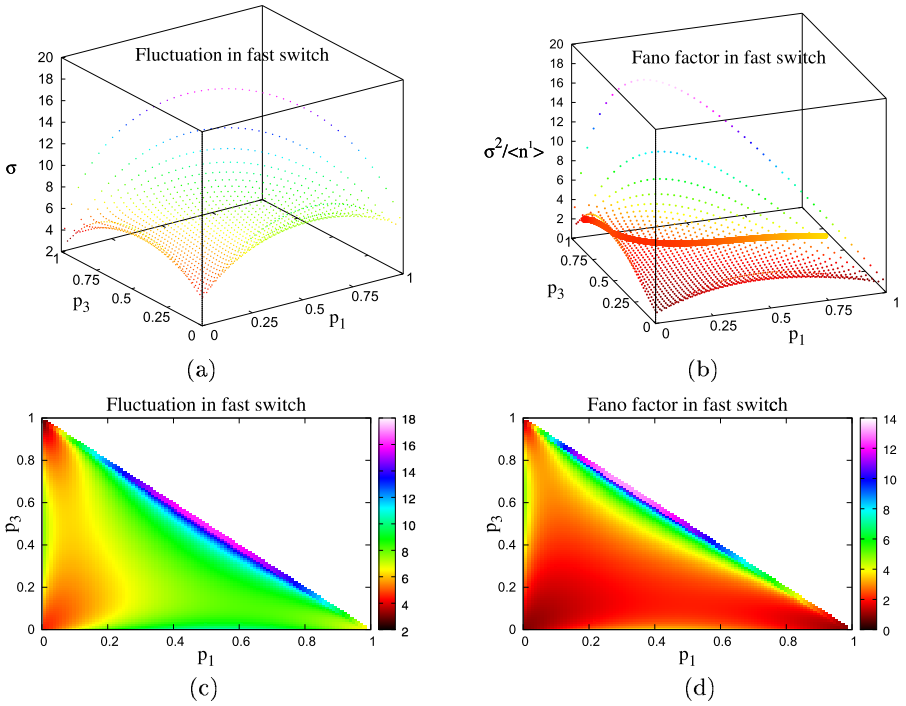



**Fig. 5** Behavior of the total fluctuation $\sigma$ and the Fano factor $\sigma^2/\langle n^1\rangle$ in the regime of fast switch, $\Sigma = 10$, with mRNA degradation time scale $\rho = 1$, transcriptional efficiency parameters $k_1 = 40$, $k_2 = 20$, $k_3 = 5$ and couplings $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.5$. (**a**) 3-d plot of fluctuation. (**b**) 3-d plot of Fano factor. (**c**) Density plot of fluctuation. (**d**) Density plot of Fano factor. A path connecting the state $p_3 = 1$ to the state $p_1 = 1$ and avoiding the "mountain pass problem" is shown in (**b**). Along this path, the Fano factor remains $< 7$, while along a direct path that does not use the intermediate state (i.e., with $p_2$ remaining equal to zero), it would reach values between 17 and 18

along with an almost insignificant rate of molecule production (small value of $k_3/\rho$). Indeed, noise reduction is not a goal just by itself: it should happen for states with a production rate that is sufficient to guarantee functionality of the produced molecules in the cell. Thus, we see that multimodal systems, such as the one proposed here, provide a good strategy to achieve noise reduction while maintaining an acceptable level of gene expression: this is achieved by allowing the system to operate in an intermediate state of efficiency with non-negligible occupancy probability.

### 4.4 Probabilities and Noise for a Fixed Mean Value

In this section, we analyze the shape of the distributions and the noise amplitude when the system is forced to operate at a fixed average level of production.

Figure 6(a) exhibits a density plot for the fluctuation of the system as a function of the mean value $\langle n^1\rangle$ and the asymptotic occupancy probability $p_2$ for the intermediate state. As a general rule, valid for any fixed mean value $\langle n^1\rangle$ of the production level, the possibility of the system to operate in an intermediate state ($p_2 \neq 0$) causes the total noise to decrease. In the case where $p_2$ can vary all the way from zero to one (which

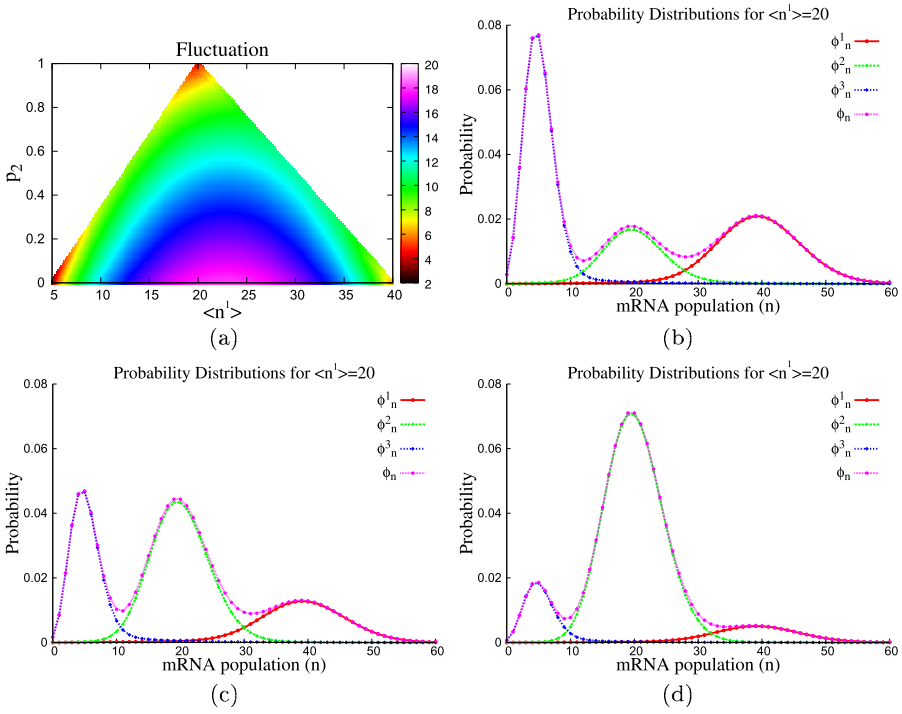



**Fig. 6** (**a**) Density plot for fluctuation as a function of mean value $\langle n^1 \rangle$ and intermediate state occupancy $p_2$ in the slow switch regime ($\Sigma = 0.1$). Remaining figures: partial and total probability distributions as functions of the mean RNA population number $n$, in the slow switch regime ($\Sigma = 0.1$), with mean value fixed at $\langle n^1 \rangle = 20$, for different values of the occupancy probabilities. (**b**) $p_1 = 0.35$, $p_2 = 0.2$, $p_3 = 0.45$. (**c**) $p_1 = 0.22$, $p_2 = 0.5$, $p_3 = 0.28$. (**d**) $p_1 = 0.09$, $p_2 = 0.8$, $p_3 = 0.11$. As before, in all figures, $\rho = 1$, $k_1 = 40$, $k_2 = 20$, $k_3 = 5$, $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.5$

happens for $\langle n^1 \rangle \approx 20$), we observe that the fluctuations vary from a maximum value of about 20 (for $p_2 = 0$) to a minimal value of about 4 (for $p_2 = 1$). The remaining panels exhibit the different shapes that the probability distributions can assume, even when all of them are associated with the same mean value $\langle n^1 \rangle$ (again, $\approx 20$), but with noise levels that decrease systematically as $p_2$ increases, from Fig. 6(b) via Fig. 6(c) to Fig. 6(d).

### 4.5 Changing the Mean Value at Fixed Maximum Noise

Changes in the environment (for instance, the replacement of a carbon source by another one, in the case of a bacterial metabolism) induce changes in the average gene expression. In our model, this can be taken into account by allowing the occupancy probabilities $p_i$ $(i = 1, 2, 3)$ to depend on some external variable.

Concretely, let us consider a continuous transition of the average gene expression from its maximum value ($p_1 = 1$, $p_2 = p_3 = 0$) to its minimum value ($p_1 = p_2 = 0$, $p_3 = 1$). In the initial and the final states, the probability distributions are Poissonian, with Fano factor equal to one. In the intermediate states, the Fano factor is greater

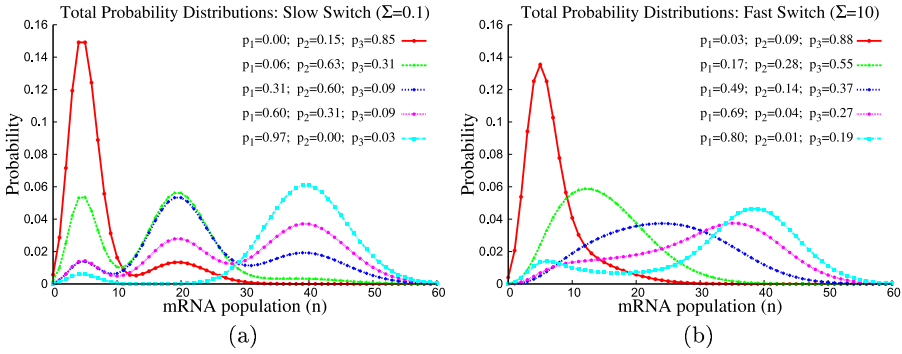



**Fig. 7** Behavior of the total probability distribution for different values of the occupancy probabilities, as indicated, forcing the Fano factor to remain below a fixed value, chosen as small as possible, and allowing the mean value to vary. As before, in all figures, $\rho = 1$, $k_1 = 40$, $k_2 = 20$, $k_3 = 5$, $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.5$

than one, but we want to keep it below a given fixed value $F^*$, say, chosen as small as possible.

Finding a continuous path connecting the maximum and the minimum expression states and along which the Fano factor $F$ is as small as possible is a "mountain pass" problem for the function $F = f(p_1, p_3)$. Some solutions to this problem are shown in Figs. 4(b) and 5(b), from which one may infer that paths of low Fano factor connecting the two extreme states always pass through states with $p_2 > 0$, while a direct transition along the straight line corresponding to $p_2 = 0$ leads to a much higher increase of $F$. This means that, typically, the presence of intermediate states leads to noise reduction. In order to gain more insight into the mechanism behind this effect, one can calculate the behavior of the probability distribution for gene expression level along such "mountain pass" paths. For small $\Sigma$ (slow switch), the probability distribution is multimodal all along the path and the change of average is performed by "mode exchange," namely by the transfer of probability between modes. Noise reduction is obtained by performing this transfer sequentially, first from the high expression mode to the intermediate expression mode and then from there to the low expression mode; see Fig. 7(a). For large $\Sigma$ (fast switch), only the first step of the mode exchange mechanism is employed, producing a moderate decrease of the average expression; at the same time, the Fano factor increases to its maximum value, corresponding to a saddle point of the function $f(p_1, p_3)$; see Fig. 7(b). A further decrease of the average expression is achieved by keeping a single mode and gradually shifting its position. Thus, the presence of an intermediate state is noise reducing also in the fast switching case, allowing the mode shift to be performed with no further increase of the Fano factor.

## 5 Generalization of the Model to an Arbitrary Number of States

In this section, we briefly discuss the generalization of the model to an arbitrary number $N$ of operational states of different transcriptional efficiencies, modeling the mRNA production of a gene controlled by multiple factors. Situations of this type

abound in biology, a classical example from the bacterial world being the λ-phage, with $N = 8$ (Ackers et al. 1982; Ptashne 1992); see also Cases and de Lorenzo (2005) for further examples. Here, it is essential to write the master equations as a unique vector equation,

$$\frac{d\boldsymbol{\phi}_n}{dt} = K[\boldsymbol{\phi}_{n-1} - \boldsymbol{\phi}_n] + \rho\big[(n+1)\boldsymbol{\phi}_{n+1} - n\boldsymbol{\phi}_n\big] + H\boldsymbol{\phi}_n, \tag{72}$$

where $\boldsymbol{\phi}_n(t) = (\phi_n^1(t), \ldots, \phi_n^N(t))^T$ ($T$ denotes transpose). $K$ and $\rho$ are now diagonal $(N \times N)$-matrices, where as before the first contains the mRNA production rates and the second is proportional to the identity matrix, containing the mRNA molecule degradation rate. Again, the matrix $H$ encodes the couplings between the states of the system and has the properties

$$\sum_{i=1}^{N} H_{ij} = 0, \qquad H_{ii} \le 0, \qquad H_{ij} \ge 0 \quad \text{if } i \ne j. \tag{73}$$

The vector generating function is defined as in Eq. (9),

$$\boldsymbol{\phi}(z,t) = \sum_{n=0}^{\infty} \boldsymbol{\phi}_n(t) z^n, \tag{74}$$

and it satisfies the same system of partial differential equations (10),

$$\frac{\partial \boldsymbol{\phi}}{\partial t} = (z-1)\left[K\boldsymbol{\phi} - \rho\frac{\partial \boldsymbol{\phi}}{\partial z}\right] + H\boldsymbol{\phi}. \tag{75}$$

The probability distributions of each state and the moments are also recovered as before (see Eqs. (11) and (12)), and the same goes for the occupancy probabilities $P_i(t) = \phi^i(z=1,t)$ (with $i = 1, \ldots, N$): their dynamics is governed by the same system, Eqs. (14),

$$\frac{d\boldsymbol{P}}{dt} = H\boldsymbol{P}, \tag{76}$$

where $\boldsymbol{P}(t) = (P_1(t), \ldots, P_N(t))^T$, that is, $\boldsymbol{P}(t) = \boldsymbol{\phi}(z,t)|_{z=1}$. As before, the decisive property of the coupling matrix $H$ is that it can be factorized according to

$$H = \epsilon h, \tag{77}$$

where $h = \mathrm{diag}(h_1, \ldots, h_N)$ with $h_i = -H_{ii}$, so that the matrix $\epsilon$ satisfies the conditions

$$\sum_{i=1}^{N} \epsilon_{ij} = 0, \qquad \epsilon_{ii} = -1, \qquad 0 \le \epsilon_{ij} \le 1 \quad \text{if } i \ne j. \tag{78}$$

Generically (that is, except for special choices of the matrix $\epsilon$), $H$ will have $N$ distinct eigenvalues with eigenvectors $\boldsymbol{p}$, corresponding to the eigenvalue 0, and

$\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(N-1)}$, corresponding to the non-zero eigenvalues $-\lambda^{(1)}, \ldots, -\lambda^{(N-1)}$, respectively: then the general solution of the system (76) can be written in the form

$$\boldsymbol{P}(t) = \boldsymbol{p} + \exp\bigl(-\lambda^{(1)}t\bigr)\boldsymbol{v}^{(1)} + \cdots + \exp\bigl(-\lambda^{(N-1)}t\bigr)\boldsymbol{v}^{(N-1)}. \tag{79}$$

Moreover, the condition that the total probability should be equal to 1, i.e.,

$$\sum_{i=1}^{N} P_i(t) = 1, \tag{80}$$

is guaranteed by requiring that

$$\sum_{i=1}^{N} p_i = 1, \tag{81}$$

since as long as the $\lambda^{(1)}, \ldots, \lambda^{(N-1)}$ are all non-zero, it follows from the eigenvalue equations $H\boldsymbol{v}^{(1)} = \lambda^{(1)}\boldsymbol{v}^{(1)}, \ldots, H\boldsymbol{v}^{(N-1)} = \lambda^{(N-1)}\boldsymbol{v}^{(N-1)}$, together with the properties of $H$ that

$$\sum_{i=1}^{N} v_i^{(1)} = 0, \qquad \ldots, \qquad \sum_{i=1}^{N} v_i^{(N-1)} = 0. \tag{82}$$

In order to determine the components of the null eigenvector $\boldsymbol{p}$, we make use of the following fact from linear algebra, which does not seem to be widely known and whose proof will therefore be given here.

**Lemma 1** *Let $A$ be an $(N \times N)$-matrix and $A^c$ its cofactor matrix, i.e., the matrix whose entry in the $i$th row and $j$th column is $(-1)^{i+j}$ times the determinant of the $(N-1) \times (N-1)$-matrix obtained from $A$ by deleting the $j$th row and $i$th column.*

$$A^c_{i,j} = (-1)^{i+j} \det \begin{pmatrix} A_{1,1} & \ldots & A_{1,i-1} & A_{1,i+1} & \ldots & A_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{j-1,1} & \ldots & A_{j-1,i-1} & A_{j-1,i+1} & \ldots & A_{j-1,N} \\ A_{j+1,1} & \ldots & A_{j+1,i-1} & A_{j+1,i+1} & \ldots & A_{j+1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{N,1} & \ldots & A_{N,i-1} & A_{N,i+1} & \ldots & A_{N,N} \end{pmatrix}. \tag{83}$$

*Assume that the row vectors of $A$ sum up to zero, i.e.,*

$$\sum_{i=1}^{N} A_{i,j} = 0 \quad \textit{for } 1 \le j \le N. \tag{84}$$

*Then every row of $A^c$ is a multiple of the covector $\boldsymbol{u}^T = (1, \ldots, 1)$, and the vector*

$$\boldsymbol{q} = \begin{pmatrix} A^c_{1,1} \\ \vdots \\ A^c_{N,N} \end{pmatrix} \tag{85}$$

*belongs to the kernel of $A$.*

*Proof* The first assertion can be reformulated as stating that if $A$ satisfies Eq. (84), then its cofactor matrix satisfies

$$A^c_{i,j} = A^c_{i,k} \quad \text{for } 1 \le i, j, k \le N.$$

Note that this conclusion is clearly invariant under permutations both of the rows and the columns of $A$, so it suffices to prove it when $i = 1$, $j = 1$ and $k = 2$, say. But

$$A^c_{1,1} - A^c_{1,2} = \det \begin{pmatrix} A_{2,2} & \dots & A_{2,N} \\ A_{3,2} & \dots & A_{3,N} \\ \vdots & \ddots & \vdots \\ A_{N,2} & \dots & A_{N,N} \end{pmatrix} + \det \begin{pmatrix} A_{1,2} & \dots & A_{1,N} \\ A_{3,2} & \dots & A_{3,N} \\ \vdots & \ddots & \vdots \\ A_{N,2} & \dots & A_{N,N} \end{pmatrix}$$
$$= -\sum_{i=3}^{N} \det \begin{pmatrix} A_{i,2} & \dots & A_{i,N} \\ A_{3,2} & \dots & A_{3,N} \\ \vdots & \ddots & \vdots \\ A_{N,2} & \dots & A_{N,N} \end{pmatrix} = 0\,,$$

where we have used Eq. (84) together with the fact that the determinant of a matrix is a linear function in each of its row vectors when all other row vectors are held fixed. Now it is a standard fact from linear algebra that the cofactor matrix $A^c$ satisfies $AA^c = \det(A)1$, and since in the present case, $\det(A) = 0$, we get

$$\sum_{k=1}^{N} A_{i,k} A^c_{k,j} = 0 \quad \text{for } 1 \le i, j \le N,$$

which in view of the fact that, as has just been shown, the $A^c_{k,j}$ are independent of $j$ and hence $A^c_{k,j}$ can be replaced by $A^c_{k,k}$, means precisely that $A$ annihilates the vector $\boldsymbol{q}$ as defined in (85). □

Here, we shall apply this lemma to the matrices $H$ and $\epsilon$ and use the fact that the factorization (77) with $h$ diagonal implies $H_{i,j} = h_i \epsilon_{i,j}$, and hence

$$H^c_{i,i} = \prod_{\substack{k=1 \\ k \neq i}}^{N} h_k \epsilon^c_{i,i}. \tag{86}$$

Denoting the diagonal elements of the cofactor matrix of the matrix $\epsilon$ by $\delta_i$ ($\delta_i = \epsilon^c_{i,i}$ for $i = 1, \dots, N$), as before, we may thus conclude that the asymptotic occupancy probabilities $p_i$, being the components of a vector $\boldsymbol{p}$ satisfying $H\boldsymbol{p} = 0$, must have the form

$$h_i p_i = \alpha \delta_i$$

with some proportionality constant $\alpha$. To determine it, multiply this equation with the product of all $p_k$ except $p_i$ and sum over $i$; then using that the $p_i$ sum up to 1, we get

$$h_i = \frac{\Sigma}{\Delta} \prod_{\substack{k=1 \\ k\neq i}}^{N} p_k \delta_i, \tag{87}$$

where $\Sigma$ is the trace of $h$ (or up to a sign, of $H$),

$$\Sigma = \sum_{i=1}^{N} h_i = \mathrm{tr}(h) = -\mathrm{tr}(H), \tag{88}$$

and

$$\Delta = \sum_{i=1}^{N} \prod_{\substack{k=1 \\ k\neq i}}^{N} p_k \delta_i. \tag{89}$$

As a result, the factorization (77) can be rewritten in the form

$$H = \frac{\Sigma}{\Delta} \epsilon \delta \pi, \tag{90}$$

where $\delta$ and $\pi$ are diagonal matrices defined by

$$\delta = \begin{pmatrix} \delta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \delta_N \end{pmatrix}, \qquad \pi = \begin{pmatrix} p_2 \dots p_N & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & p_1 \dots p_{N-1} \end{pmatrix}. \tag{91}$$

As an example, consider the case $N = 4$. Here, we can factorize the coupling matrix $H$ as in Eq. (90), with

$$\epsilon = \begin{pmatrix} -1 & \epsilon_3\epsilon_4 & \epsilon_5(1-\epsilon_6) & 1-\epsilon_7 \\ 1-\epsilon_1 & -1 & \epsilon_5\epsilon_6 & \epsilon_7(1-\epsilon_8) \\ \epsilon_1(1-\epsilon_2) & 1-\epsilon_3 & -1 & \epsilon_7\epsilon_8 \\ \epsilon_1\epsilon_2 & \epsilon_3(1-\epsilon_4) & 1-\epsilon_5 & -1 \end{pmatrix} \tag{92}$$

and

$$\delta = \begin{pmatrix} \delta_1 & 0 & 0 & 0 \\ 0 & \delta_2 & 0 & 0 \\ 0 & 0 & \delta_3 & 0 \\ 0 & 0 & 0 & \delta_4 \end{pmatrix}, \qquad \pi = \begin{pmatrix} p_2p_3p_4 & 0 & 0 & 0 \\ 0 & p_1p_3p_4 & 0 & 0 \\ 0 & 0 & p_1p_2p_4 & 0 \\ 0 & 0 & 0 & p_1p_2p_3 \end{pmatrix}. \tag{93}$$

Similarly, there are now four transcriptional efficiency parameters, coded into a diagonal matrix

$$K = \begin{pmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{pmatrix}. \tag{94}$$

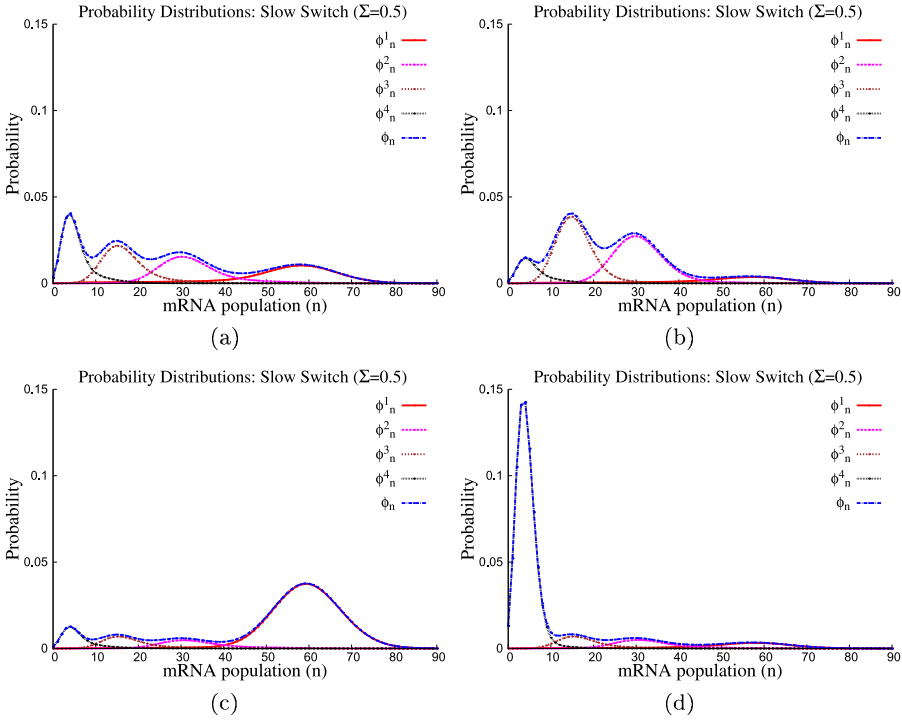



**Fig. 8** Probability distributions for the example of a system with four different states of transcriptional efficiency in the case of slow switch ($\Sigma = 0.5$). (**a**) $p_1 = 1/4,\ p_2 = 1/4,\ p_3 = 1/4,\ p_4 = 1/4$; (**b**) $p_1 = 1/10,\ p_2 = 4/10,\ p_3 = 4/10,\ p_4 = 1/10$; (**c**) $p_1 = 3/4,\ p_2 = 1/12,\ p_3 = 1/12,\ p_4 = 1/12$; (**d**) $p_1 = 1/12,\ p_2 = 1/12,\ p_3 = 1/12,\ p_4 = 3/4$. In all figures, $\rho = 1$, $k_1 = 60$, $k_2 = 30$, $k_3 = 15$, $k_4 = 4$ and $\epsilon_i = 0$ (which is the cyclic coupling) for $i = 1, \ldots, 8$

In terms of these data, we can again calculate probability distributions, by solving the recursion relations (42). An example is shown in Fig. 8, the values of the parameters chosen being listed in the figure caption.

## 6 Conclusions

In this paper, we have presented an analytical method for computing steady state probability density functions and their first two moments for a gene promoter model with multiple states.

The parametrization of the transition matrix of the model adopted here is based on its factorization into the product of two components: a matrix $\epsilon$, which encodes the couplings between the different states of transcriptional efficiency, representing an intrinsic property of the system, and a matrix $h$, which gives the strengths of the couplings. As a result, we have been able to reinterpret the mathematical constants that appear in the model in terms of biological parameters. Moreover, this approach decouples intrinsic aspects from extrinsic ones: the couplings ($\epsilon_i$) and the switch decay parameter ($\Sigma$) reflect intrinsic properties of the system, whereas the asymptotic

occupancy probabilities for the transcriptional efficiency states are controlled by external agents.

In particular, in our approach the desired probability distributions are calculated directly in terms of the asymptotic occupancy probabilities. This choice facilitates the analysis of noise, revealing the differences between the noise patterns for slow and fast switch, at a fixed value of mean expression and also when the mean expression changes under the influence of external agents such as inducers, repressors, carbon source, etc.

Already in the case of a three-states system, the three-peak probability distributions resulting from our calculations show that, in such a configuration, the knowledge of just a few moments of the distribution may not be enough. Changes in gene expression triggered by external agents can use a mode exchange mechanism to redistribute the probabilities among the modes. In a repressed condition, the promoter will operate in a low expression basal mode. When external conditions become permissive, the promoter will first function in the intermediate expression state and then in the maximum expression state. Our theoretical results indicate that the study of such probability distributions can provide more information about details of transcription regulation than just its moments. This suggests undertaking a more accurate experimental determination of such distributions. Scanning two photon fluorescence microscopy, combined with a number and brightness analysis, is well adapted for this purpose and has already provided hints on how to reconcile observed noise patterns with mechanisms of transcription control (Ferguson et al. 2012).

Finally, our setup for the transition matrix allows for a straightforward generalization of the model to an arbitrary number of states, and we show how to calculate stationary probability distributions in this general case. All information about the steady state statistics of the general model, including the moments, is contained in this distribution. This is important to the extent that biological complexity provides strong motivation for constructing models of gene expression that admit states of intermediate transcriptional efficiency, rather than just treating it as a simple "on/off" process (Cases and de Lorenzo 2005), and there is then no logical reason to assume that there is just one such intermediate state. This actually happens for numerous prokaryote promoters, but it also applies to eukaryote gene expression. As an example, we may mention the transcriptional control of segmentation in *Drosophila melanogaster*, which involves 34 binding sites for 7 regulatory proteins (Janssens et al. 2006). In fact, multiple intermediate states have already been invoked by biologists to understand the relation between chromatin state regulation and the transcription readout (Chen et al. 2012). Of course, gene expression in eukaryotic cells is much more complex than in prokaryotes, due to the presence of a wealth of additional factors that modify the basic process, and only comparison with experiment can show whether predictions of a (still fairly simple) mathematical model such as the one discussed in this paper will be able to catch at least some of its salient features. But we do expect that one of these feature will be multi-modality, as exemplified in recent work in synthetic biology (Murphy et al. 2007, 2010).

**Acknowledgements** Work supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil) and by the USP-COFECUB 2008-2012 program.

## References


Ackers, G. K., Johnson, A. D., & Shea, M. A. (1982). Quantitative model for gene regulation by λ phage repressor. *Proc. Natl. Acad. Sci. USA*, *79*, 1129–1133.

Arkin, A., Ross, J., & McAdams, H. H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected escherichia coli cells. *Genetics*, *149*, 1633–1648.

Cai, L., Friedman, N., & Xie, X. (2006). Stochastic protein expression in individual cells at the single molecule level. *Nature*, *440*(7082), 358–362.

Cases, I., & de Lorenzo, V. (2005). Promoters in the environment: transcriptional regulation in its natural context. *Nat. Rev. Microbiol.*, *3*, 105–118.

Chen, H., Monte, E., Parvatiyar, M. S., Rosa-Garrido, M., Franklin, S., & Vondriska, T. M. (2012). Structural considerations for chromatin state models with transcription as a functional readout. *FEBS Lett.*, *586*, 3548–3554.

Coulon, A., Gandrillon, O., & Beslon, G. (2010). On the spontaneous stochastic dynamics of a single gene: complexity of the molecular interplay at the promoter. *BMC Syst. Biol.*, *4*(1), 2–18.

Delbrück, M. (1940). Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.*, *8*, 120–124.

Elf, J., Li, G., & Xie, X. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, *316*(5828), 1191–1194.

Escolar, L., Pérez-Martín, J., & de Lorenzo, V. (1999). Opening the iron box: transcriptional metalloregulation by the fur protein. *J. Bacteriol.*, *181*, 6223–6229.

Ferguson, M., Le Coq, D., Jules, M., Aymerich, S., Radulescu, O., Declerck, N., & Royer, C. (2012). Reconciling molecular regulatory mechanisms with noise patterns of bacterial metabolic promoters in induced and repressed states. *Proc. Natl. Acad. Sci.*, *109*(1), 155–160.

Garcia, H., & Phillips, R. (2011). Quantitative dissection of the simple repression input–output function. *Proc. Natl. Acad. Sci.*, *108*(29), 12,173–12,178.

Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of genetic toggle switch in Escherichia coli. *Nature*, *403*, 339–342.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, *81*, 2340–2361.

Goss, P., & Peccoud, J. (1998). Quantitative modeling of stochastic systems in molecular biology by using stochastic petri nets. *Proc. Natl. Acad. Sci.*, *95*(12), 6750–6755.

Hasty, J., Pradines, J., Dolnik, M., & Collins, J. J. (2000). Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. USA*, *97*, 2075–2080.

Huang, S. (2009). Non-genetic heterogeneity of cells in development: more than just noise. *Development*, *136*(23), 3853–3862.

Janssens, H., Hou, S., Jaeger, J., Kim, A. R., Myasnikova, E., Sharp, D., & Reinitz, J. (2006). Quantitative and predictive model of transcriptional control of the Drosophila melanogaster even skipped gene. *Nat. Genet.*, *38*, 1159–1165.

Kepler, T., & Elston, T. (2001). Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.*, *81*(6), 3116–3136.

Kierzek, A., Zaim, J., & Zielenkiewicz, P. (2001). The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *J. Biol. Chem.*, *276*, 8165–8172.

Kirkilionis, M., Janus, U., & Sbano, L. (2011). Multi-scale genetic dynamic modelling ii: application to synthetic biology. *Theory Biosci.*, *130*(3), 183–201.

McAdams, H. H., & Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA*, *94*, 814–819.

McAdams, H. H., & Arkin, A. (1998). Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.*, *27*, 199–224.

Metzler, R., & Wolynes, P. G. (2002). Number fluctuations and the threshold model of kinetic switches. *Chem. Phys.*, *284*, 469–479.

Monod, J., & Jacob, F. (1961). Genetic regulatory mechanisms in synthesis of proteins. *J. Mol. Biol.*, *3*, 318–356.

Muller-Hill, B. (1996). *The lac operon: a short history of a genetic paradigm*. Berlin: de Gruyter.

Murphy, K. F., Balázsi, G., & Collins, J. J. (2007). Combinatorial promoter design for engineering noisy gene expression. *Proc. Natl. Acad. Sci. USA*, *104*, 12,726–12,731.

Murphy, K. F., Adams, R. M., Wang, X., Balázsi, G., & Collins, J. J. (2010). Tuning and controlling gene expression noise in synthetic gene networks. *Nucleic Acids Res.*, *38*, 2712–2726.

Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., & van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nat. Genet.*, *31*, 69–73.

Paulsson, J., & Ehrenberg, M. (2000). Random signal fluctuations can reduce random fluctuations in regulated components of chemical regulatory networks. *Phys. Rev. Lett.*, *84*, 5447–5450.

Ptashne, M. (1992). *A genetic switch: phage λ and higher organisms*. Cambridge: Cell Press/Blackwell.

Raj, A., Peskin, C., Tranchina, D., Vargas, D., & Tyagi, S. (2006). Stochastic mrna synthesis in mammalian cells. *PLoS Biol.*, *4*(10), e309.

Rao, C. V., Wolf, D. M., & Arkin, A. P. (2002). Control, exploitation and tolerance of intracellular noise. *Nature*, *420*, 231–237.

Saiz, L., & Vilar, J. (2008). Ab initio thermodynamic modeling of distal multisite transcription regulation. *Nucleic Acids Res.*, *36*(3), 726–731.

Sánchez, Á., & Kondev, J. (2008). Transcriptional control of noise in gene expression. *Proc. Natl. Acad. Sci.*, *105*(13), 5081–5086.

Sasai, M., & Wolynes, P. G. (2003). Stochastic gene expression as a many-body problem. *Proc. Natl. Acad. Sci. USA*, *100*, 2374–2379.

Satory, D., Gordon, A., Halliday, J., & Herman, C. (2011). Epigenetic switches: can infidelity govern fate in microbes? *Curr. Opin. Microbiol.*, *14*(2), 212–217.

Taniguchi, Y., Choi, P., Li, G., Chen, H., Babu, M., Hearn, J., Emili, A., & Xie, X. (2010). Quantifying e. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, *329*(5991), 533–538.

Thattai, M., & van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA*, *98*, 8614–8619.

Thattai, M., & van Oudenaarden, A. (2002). Attenuation of noise in ultrasensitives signaling cascades. *Biophys. J.*, *82*, 2943–2950.

van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry*. Amsterdam: North-Holland.

Vicente, M., Chater, K. F., & de Lorenzo, V. (1999). Bacterial transcription factors involved in global regulation. *Mol. Microbiol.*, *33*, 8–17.