Spectral solutions to stochastic models of gene expression with bursts and regulation

Andrew Mugler[*]

Department of Physics, Columbia University, New York, New York 10027, USA

Aleksandra M. Walczak[†]

Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA

Chris H. Wiggins[‡]

Department of Applied Physics and Applied Mathematics, Center for Computational Biology and Bioinformatics,
Columbia University, New York, New York 10027, USA
(Received 20 July 2009; published 20 October 2009)

Signal-processing molecules inside cells are often present at low copy number, which necessitates probabilistic models to account for intrinsic noise. Probability distributions have traditionally been found using simulation-based approaches which then require estimating the distributions from many samples. Here we present in detail an alternative method for directly calculating a probability distribution by expanding in the natural eigenfunctions of the governing equation, which is linear. We apply the resulting *spectral* method to three general models of stochastic gene expression: a single gene with multiple expression states (often used as a model of bursting in the limit of two states), a gene regulatory cascade, and a combined model of bursting and regulation. In all cases we find either analytic results or numerical prescriptions that greatly outperform simulations in efficiency and accuracy. In the last case, we show that bimodal response in the limit of slow switching is not only possible but optimal in terms of information transmission.

I. INTRODUCTION

Signals are processed in cells using networks of interacting components, including proteins, mRNAs, and small signaling molecules. These components are usually present in low numbers [1–6], which means the size of the fluctuations in their copy counts is comparable to the copy counts themselves. Noise in gene networks has been shown to propagate [7] and therefore explicitly accounting for the stochastic nature of gene expression appears important when predicting the properties of real biological networks.

Although summary statistics such as mean and variance are sometimes sufficient for answering questions of biological interest [8], calculating certain quantities, such as information transmission [9–14], requires knowing the full probability distribution. Full knowledge of the probability distribution can also be used to discern different molecular models of the noise sources based on recent exact measurements of probability distributions [15–18].

Much analytical and purely computational effort has gone into detailed models of noise in small genetic switches [8,19–21]. The most general description is based on the master equation describing the time evolution of the joint probability distribution over all copy counts [22]. Some progress has been made by applying approximations to the master equation [24–26]. For example, a wide class of approximations focuses on limits of large concentrations or small switches [8,19,23]. More often, modelers resort to stochastic simulation techniques, the most common of which is based on the varying-step Monte Carlo method [27,28]. This method requires a computational challenge (generating many sample trajectories) followed by an even more difficult statistical challenge (parametrizing or otherwise estimating the probability distribution from which the samples are drawn) [29]. Recently there have been significant advances in simulation-based methods to circumvent these problems [20,21,30]. Simulation techniques are especially useful for more detailed studies of experimentally well-characterized systems, including those incorporating DNA looping, non-specific binding [30], and explicit spatial effects [20,21]. However, it is often beneficial to first gain intuition from simplified analytical models.

In a recent paper [31] we introduce an alternative method for calculating the steady-state distributions of chemical reactants, which we call the spectral method. The procedure relies on exploiting the natural basis of a simpler problem from the same class. The full problem is then solved numerically as an expansion in the natural basis [32]. In the spectral method we use the analytical guidance of a simple birth-death problem to reduce the master equation for a cascade to a set of algebraic equations. We break the problem into two parts: a parameter-independent preprocessing step and the parameter-dependent step of obtaining the actual probability distributions. The spectral method allows huge computational gains with respect to simulations. In prior work [31] we illustrate the method in the example of gene regulatory cascades. We combine the spectral method with a Markov approximation, which exploits the observation that the behavior of a given species should depend only weakly on distant nodes given the proximal nodes.

In this paper we expand upon the application of the spectral method to more biologically realistic models of regula-

---

[*]ajm2121@columbia.edu
[†]awalczak@princeton.edu
[‡]chris.wiggins@columbia.edu

tion: (i) a model of bursting in gene expression and (ii) a model that includes both bursts and explicit regulation by binding of transcription factor proteins. In both cases we demonstrate how the spectral method gives either analytic results or reduced algebraic expressions that can be solved numerically in orders of magnitude less time than stochastic simulations.

We note that although information propagation in biological networks is impeded by numerous mechanisms collectively modeled as noise, here we focus on the inescapable or "intrinsic" noise resulting from the finite copy number of the molecules. One should consider these results, then, as an upper bound on information propagation, further hampered by, for example, cell division [33]; spatial effects [20,21]; active degradation of the constituents (here modeled via a constant degradation rate); and other complicating mechanisms particular to specific systems. Additionally, here we focus on steady-state solutions, but extensions to dynamics are straightforward and currently being pursued.

We begin with a model of a multistate birth-death process, a special case of which has been used to describe transcriptional bursting [17,34]. We illustrate how the spectral method reduces the model to a simple iterative algebraic equation, and in the appropriate limiting case recovers the known analytic results. We also use this section to introduce the basic notation used throughout the paper. Next we explore the problem of gene regulation in detail. The main idea behind the spectral method is the exploitation of an underlying natural basis for a problem which we can solve exactly. We explore four different spectral representations of the regulation model used in previous work [31] that arise from four natural choices of eigenbasis in which to expand the solution (cf. Sec. II A and Fig. 2). All representations reduce the master equation to a set of linear algebraic equations, and one admits an analytic solution by virtue of the tridiagonal matrix algorithm. We compare the efficiencies of the representations' numerical implementations and show that all outperform simulation. Lastly, we apply the spectral method to a model that combines bursting and regulation. We obtain a linear algebraic expression that permits large speedup over simulation and thus admits optimization of information transmission. Optimization reveals two types of solutions: a unimodal response when the rates of switching between expression states are comparable to degradation rates and a bimodal response when switching rates are much slower than degradation rates.

II. BURSTS OF GENE EXPRESSION

We first consider a model of gene expression in which a gene exists in one of Z stochastic "states," i.e., protein production obeys a simple birth-death process but with a state-dependent birth rate. In the special case of $Z=2$, this corresponds to a gene existing in an on or an off state due, for example, to the binding and unbinding of the RNA polymerase. Such a model has been used to describe transcriptional bursting [17,34], and we specialize to this case in Sec. II C.

For a general Z-state process, the master equation

$$\dot{p}_n^z = g_z p_{n-1}^z + (n+1)p_{n+1}^z - (g_z + n)p_n^z + \sum_{z'} \Omega_{zz'} p_n^{z'} \quad (1)$$

describes the time evolution of the joint probability distribution p_n^z, where z specifies the state ($1 \leq z \leq Z$), n is the number of proteins, g_z is the production rate in state z, $\Omega_{zz'}$ is a stochastic matrix of transition rates between states, and $\dot{p}_n^z$ denotes differentiation of the probability distribution with respect to time. Time and all rates have been nondimensionalized by the protein degradation rate. Note that conservation of probability requires

$$\sum_z \Omega_{zz'} = 0. \quad (2)$$

The relationship between the transition rates in $\Omega_{zz'}$ and the probabilities $\pi_z = \Sigma_n p_n^z$ of being in the zth state can be seen by summing Eq. (1) over n; at steady state one obtains

$$\sum_{z'} \Omega_{zz'} \pi_{z'} = 0, \quad (3)$$

and normalization requires

$$\sum_z \pi_z = 1. \quad (4)$$

In the following sections, we introduce the spectral method and demonstrate how it can be used to solve for the full joint distribution p_n^z.

A. Notation and definitions

We begin our solution of Eq. (1) by defining the generating function [22] $G_z(x) = \Sigma_n p_n^z x^n$ over complex variable x [35] (note that superscript z is an index, while superscript n on x^n is a power). It will prove more convenient to rewrite the generating function in a more abstract representation using states indexed by protein number $|n\rangle$,

$$|G_z\rangle = \sum_n p_n^z |n\rangle, \quad (5)$$

with inverse transform

$$p_n^z = \langle n|G_z\rangle. \quad (6)$$

The more familiar form can be recovered by projecting the position space $\langle x|$ onto Eq. (5), with the provision that

$$\langle x|n\rangle = x^n. \quad (7)$$

Concurrent choices of conjugate state

$$\langle n|x\rangle = \frac{1}{x^{n+1}} \quad (8)$$

and inner product

$$\langle f|f'\rangle = \oint \frac{dx}{2\pi i}\langle f|x\rangle\langle x|f'\rangle \quad (9)$$

ensure orthonormality of the states, $\langle n|n'\rangle = \delta_{nn'}$, as can be verified using Cauchy's theorem,

$$\oint \frac{dx}{2\pi i} \frac{f(x)}{(x-x_0)^{n+1}} = \frac{1}{n!} \partial_x^n [f(x)]_{x=x_0} \theta(n+1), \qquad (10)$$

where the convention $\theta(0)=0$ is used for the Heaviside function.

With these definitions, summing Eq. (1) over n against $|n\rangle$ yields

$$|\dot{G}_z\rangle = -(\hat{a}^+ - 1)(\hat{a}^- - g_z)|G_z\rangle + \sum_{z'} \Omega_{zz'}|G_{z'}\rangle, \qquad (11)$$

where the operators $\hat{a}^+$ and $\hat{a}^-$ raise and lower protein number, respectively, i.e.,

$$\hat{a}^+|n\rangle = |n+1\rangle, \qquad (12)$$

$$\hat{a}^-|n\rangle = n|n-1\rangle, \qquad (13)$$

with adjoint operations

$$\langle n|\hat{a}^+ = \langle n-1|, \qquad (14)$$

$$\langle n|\hat{a}^- = (n+1)\langle n+1|. \qquad (15)$$

As in the operator treatment of the simple harmonic oscillator in quantum mechanics, the raising and lower operators satisfy the commutation relation $[\hat{a}^-, \hat{a}^+] = 1$ and $\hat{a}^+\hat{a}^-$ is a number operator, i.e., $\hat{a}^+\hat{a}^-|n\rangle = n|n\rangle$ [36]. This operator formalism for the generating function was introduced in the context of diffusion independently by Doi [37] and Zel'dovich [38] and developed by Peliti [39]. We note that all results can equivalently be obtained by remaining in x space (for example, $\hat{a}^+ \leftrightarrow x$ and $\hat{a}^- \leftrightarrow \partial_x$). However, the raising and lowering operators define an extremely simple algebra which allows us to calculate all projections without explicitly computing overlap integrals (as shown, for example, in Appendix A). A review by Mattis and Glasser [40] introduces and discusses the applications of the formalism for diffusion.

The factorized form of the birth-death operator in Eq. (11) suggests the definition of shifted raising and lowering operators

$$\hat{b}^+ = \hat{a}^+ - 1, \qquad (16)$$

$$\hat{b}_z^- = \hat{a}^- - g_z, \qquad (17)$$

making Eq. (11)

$$|\dot{G}_z\rangle = -\hat{b}^+\hat{b}_z^-|G_z\rangle + \sum_{z'} \Omega_{zz'}|G_{z'}\rangle. \qquad (18)$$

Since $\hat{b}^+$ and $\hat{b}_z^-$ are simply shifted from $\hat{a}^+$ and $\hat{a}^-$ by constants, $\hat{b}^+\hat{b}_z^-$ is a new number operator, and its eigenvalues are nonnegative integers j, i.e.,

$$\hat{b}^+\hat{b}_z^-|j_z\rangle = j|j_z\rangle, \qquad (19)$$

where j indexes (z-dependent) eigenfunctions $|j_z\rangle$. In position space, $\hat{b}^+$ and $\hat{b}_z^-$ are $x-1$ and $\partial_x - g_z$, respectively. Projecting $\langle x|$ onto Eq. (19) therefore yields a first-order ordinary differential equation whose solution (up to normalization) is

$$\langle x|j_z\rangle = (x-1)^j e^{g_z(x-1)}, \qquad (20)$$

with conjugate

$$\langle j_z|x\rangle = \frac{e^{-g_z(x-1)}}{(x-1)^{j+1}}, \qquad (21)$$

such that orthonormality $\langle j_z|j'_z\rangle = \delta_{jj'}$ is satisfied under the inner product in Eq. (9). Note that $\hat{b}^+$ and $\hat{b}_z^-$ raise and lower eigenstates $|j_z\rangle$ as in Eqs. (12)–(15), i.e.,

$$\hat{b}^+|j_z\rangle = |(j+1)_z\rangle, \qquad (22)$$

$$\hat{b}_z^-|j_z\rangle = j|(j-1)_z\rangle, \qquad (23)$$

$$\langle j_z|\hat{b}^+ = \langle (j-1)_z|, \qquad (24)$$

$$\langle j_z|\hat{b}_z^- = (j+1)\langle (j+1)_z|. \qquad (25)$$

As we will see in this and subsequent sections, going between the protein number basis $|n\rangle$ and the eigenbasis $|j_z\rangle$ requires the mixed product $\langle n|j_z\rangle$ or its conjugate $\langle j_z|n\rangle$. There are several ways of computing these objects, as described in Appendix A. Notable special cases are

$$\langle (j=0)_z|n\rangle = 1, \qquad (26)$$

$$\langle n|(j=0)_z\rangle = e^{-g_z}\frac{(g_z)^n}{n!}, \qquad (27)$$

where the latter is the Poisson distribution.

B. Spectral method

We now demonstrate how the spectral method exploits the eigenfunctions $|j_z\rangle$ to decompose and simplify the equation of motion. Expanding the generating function in the eigenbasis,

$$|G_z\rangle = \sum_j G_j^z|j_z\rangle, \qquad (28)$$

and projecting the conjugate state $\langle j_z|$ onto Eq. (18) yields the equation of motion

$$\dot{G}_j^z = -jG_j^z + \sum_{z'} \Omega_{zz'} \sum_{j'} G_{j'}^{z'}\langle j_z|j'_{z'}\rangle \qquad (29)$$

for the expansion coefficients G_j^z [where the dummy index j in Eq. (28) has been changed to j' in Eq. (29)]. Using Eqs. (9), (10), (20), and (21), the product $\langle j_z|j'_{z'}\rangle$ evaluates to

$$\langle j_z|j'_{z'}\rangle = \frac{(-\Delta_{zz'})^{j-j'}}{(j-j')!}\theta(j-j'+1), \qquad (30)$$

where $\Delta_{zz'} = g_z - g_{z'}$. Noting that $\langle j_z|j_z\rangle = 1$ and that $\langle j_z|j'_z\rangle = 0$ for $j' < j$, Eq. (29) becomes

$$\dot{G}_j^z = -jG_j^z + \sum_{z'} \Omega_{zz'} G_j^{z'} + \sum_{z' \neq z} \Omega_{zz'} \sum_{j' < j} G_{j'}^{z'} \frac{(-\Delta_{zz'})^{j-j'}}{(j-j')!}. \tag{31}$$

The last term in Eq. (31) makes clear that each jth term is slaved to terms with $j' < j$, allowing the G_j^z to be computed iteratively in j. The lower-triangular structure of the equation is a consequence of rotating to the eigenspace of the birth-death operator; this structure was not present in the original master equation.

At steady state, G_j^z obeys

$$jG_j^z - \sum_{z'} \Omega_{zz'} G_j^{z'} = \sum_{z' \neq z} \Omega_{zz'} \sum_{j' < j} G_{j'}^{z'} \frac{(-\Delta_{zz'})^{j-j'}}{(j-j')!}, \tag{32}$$

from which it is clear that G_j^z can be computed successively in j. Since

$$G_0^z = \langle (j=0)_z | G_z \rangle = \sum_n p_n^z \langle (j=0)_z | n \rangle = \pi_z \tag{33}$$

[cf. Eq. (26)], the computation is initialized using Eqs. (3) and (4), i.e.,

$$\sum_{z'} \Omega_{zz'} G_0^{z'} = 0, \tag{34}$$

$$\sum_z G_0^z = 1. \tag{35}$$

Recalling Eqs. (6) and (28), the probability distribution is retrieved via

$$p_n^z = \sum_j G_j^z \langle n | j_z \rangle, \tag{36}$$

where the mixed product $\langle n | j_z \rangle$ can be computed as described in Appendix A.

There is an alternative way to decompose the master equation spectrally. Instead of expanding the generating function in eigenfunctions $|j_z\rangle$, which depend on the production rates g_z in each state, we may expand in eigenfunctions parametrized by a single rate $\bar{g}$, i.e.,

$$|G_z\rangle = \sum_j G_j^z |j\rangle, \tag{37}$$

where

$$\langle x | j \rangle = (x-1)^j e^{\bar{g}(x-1)}, \tag{38}$$

$$\langle j | x \rangle = \frac{e^{-\bar{g}(x-1)}}{(x-1)^{j+1}}. \tag{39}$$

The parameter $\bar{g}$ is arbitrary and thus acts as a "gauge" freedom.

We may now partition the birth-death operator as

$$-\hat{b}^+ \hat{b}_z^- = -\hat{b}^+ \bar{b}^- - \hat{b}^+ \Gamma_z \tag{40}$$

where $\bar{b}^- = \hat{a}^- - \bar{g}$ such that the $|j\rangle$ are the eigenstates of $\hat{b}^+ \bar{b}^-$, i.e.,
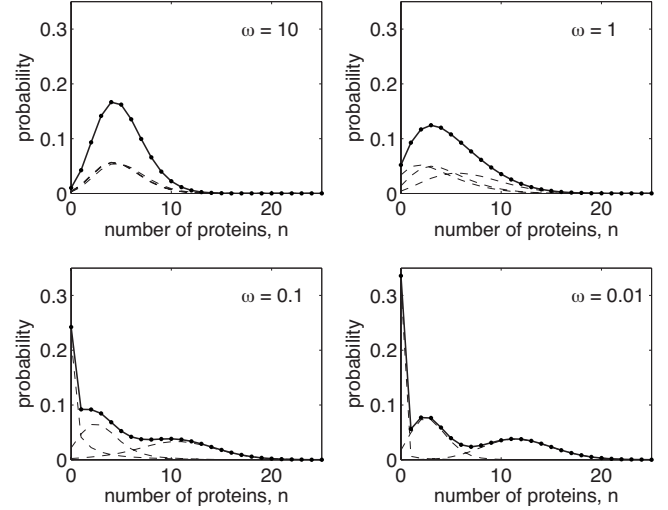

FIG. 1. An example of the spectral method for $Z=3$ states. Dashed curves show the joint distribution p_n^z for each of $z=1$, 2, and 3, and solid curves show the marginal distribution p_n. The stochastic transition matrix is given in Eq. (45), and the setting of ω in each panel is indicated in the upper-right corner. Production rates are $g_1=0$, $g_2=3$, and $g_3=12$ for all panels. Distributions are calculated via the spectral decomposition in Eq. (43) with $\bar{g} = \langle g_z \rangle = 5$.

$$\hat{b}^+ \bar{b}^- |j\rangle = j |j\rangle, \tag{41}$$

and $\Gamma_z = \bar{g} - g_z$ describes the deviation of each state's production rate from the constant $\bar{g}$.

Projecting the conjugate state $\langle j |$ onto Eq. (18) and using Eq. (37) (with dummy index j changed to j') gives

$$\dot{G}_j^z = -jG_j^z + \sum_{j'} \langle j | \hat{b}^+ \Gamma_z | j' \rangle G_{j'}^z + \sum_{z'} \Omega_{zz'} G_j^{z'}. \tag{42}$$

Recalling Eq. (24), Eq. (42) at steady state becomes

$$\sum_{z'} (\Omega_{zz'} - j\delta_{zz'}) G_j^{z'} = \Gamma_z G_{j-1}^z. \tag{43}$$

Equation (43) is subdiagonal in j, meaning computation of the jth term requires only the previous $(j-1)$th term and the inversion of the Z-by-Z matrix $(\mathbf{\Omega} - j\mathbf{I})$ (where $\mathbf{I}$ is the identity matrix). It is initialized with Eqs. (34) and (35) and solved successively in j. The probability distribution is retrieved via

$$p_n^z = \sum_j G_j^z \langle n | j \rangle, \tag{44}$$

where $\langle n | j \rangle$ is computed as described in Appendix A.

As an example of a simple computation employing the spectral method, Fig. 1 shows probability distributions for the case of $Z=3$ states, corresponding to a gene that is either off, producing proteins at a low rate, or producing proteins at a high rate. For simplicity we set the rates of switching among all states equal to a constant ω, making the stochastic matrix

$$\boldsymbol{\Omega} = \begin{pmatrix} -2\omega & \omega & \omega \\ \omega & -2\omega & \omega \\ \omega & \omega & -2\omega \end{pmatrix}. \tag{45}$$

As seen in Fig. 1, when $\omega \ll 1$ (corresponding to a switching rate much slower than the degradation rate) the dwell time in each expression state lengthens. The slow switching gives the protein copy number time to equilibrate in any of the three expression states, resulting in a trimodal marginal distribution p_n. When $\omega \gg 1$ (corresponding to a switching rate much faster than the degradation rate), the system switches frequently among the three expression states, resulting in an average production rate. In this limit, the expression state equilibrates on a faster time scale than the protein number state.

C. On/off gene

For the special case of $Z=2$ states, as when a gene is either "on" ($z=+$) or "off" ($z=-$), it is useful to demonstrate how the spectral method reproduces known analytic results [17,34]. The probability distribution can be written in vector form as

$$\vec{p}_n = \begin{pmatrix} p_n^- \\ p_n^+ \end{pmatrix}, \tag{46}$$

and defining ω_+ and ω_- as the transition rates to and from the on state, respectively, the stochastic matrix takes the form

$$\boldsymbol{\Omega} = \begin{pmatrix} -\omega_+ & \omega_- \\ \omega_+ & -\omega_- \end{pmatrix}. \tag{47}$$

Note that Eq. (3) implies

$$\frac{\pi_-}{\pi_+} = \frac{\omega_-}{\omega_+}, \tag{48}$$

which makes clear that increasing the rate of transition to either state increases the probability of being in that state.

From Eq. (32) the spectral expansion coefficients obey

$$jG_j^{\pm} + \omega_{\mp}G_j^{\pm} - \omega_{\pm}G_j^{\mp} = \omega_{\pm} \sum_{j'<j} G_{j'}^{\mp} \frac{(\mp\Delta)^{j-j'}}{(j-j')!}, \tag{49}$$

where $\Delta = \Delta_{+-} = -\Delta_{-+}$. Initializing with $G_0^{\pm} = \omega_{\pm}/(\omega_+ + \omega_-)$ and computing the first few terms reveals the pattern

$$G_j^{\pm} = \frac{\omega_{\pm}}{\omega_+ + \omega_-} \frac{(\mp\Delta)^j}{j!} \frac{\Pi_{j'=0}^{j-1}(j' + \omega_{\mp})}{\Pi_{j''=0}^{j-1}(j'' + \omega_+ + \omega_- + 1)} \tag{50}$$

$$= \frac{\omega_{\pm}}{\omega_+ + \omega_-} \frac{(\mp\Delta)^j}{j!} \frac{\Gamma(j + \omega_{\mp})}{\Gamma(\omega_{\mp})} \frac{\Gamma(\omega_+ + \omega_- + 1)}{\Gamma(j + \omega_+ + \omega_- + 1)}, \tag{51}$$

where in the second line the products are written in terms of the Gamma function.

For comparison with known results [17,34] we write the total generating function $|G\rangle = \Sigma_{\pm}|G_{\pm}\rangle$ in position space,

$$G(x) = \sum_{\pm} \langle x|G_{\pm}\rangle = \sum_{\pm} \sum_{j} \langle x|j_{\pm}\rangle\langle j_{\pm}|G_{\pm}\rangle \tag{52}$$

$$= \sum_{\pm} \sum_{j} (x-1)^j e^{g_{\pm}(x-1)} G_j^{\pm} \tag{53}$$

$$= \sum_{\pm} \frac{\omega_{\pm}}{\omega_+ + \omega_-} e^{g_{\pm}(x-1)}$$
$$\times \Phi[\omega_{\mp}, \omega_+ + \omega_- + 1; \mp\Delta(x-1)], \tag{54}$$

where

$$\Phi[\alpha, \beta; y] = \sum_{j} \frac{\Gamma(j + \alpha)}{\Gamma(\alpha)} \frac{\Gamma(\beta)}{\Gamma(j + \beta)} \frac{y^j}{j!} \tag{55}$$

is the confluent hypergeometric function. As shown in Appendix B, in the limit $g_- = 0$, Eq. (54) reduces to

$$G(x) = \Phi[\omega_+, \omega_+ + \omega_-; g_+(x-1)], \tag{56}$$

and the marginal p_n is given by

$$p_n = \frac{g_+^n}{n!} \frac{\Gamma(n + \omega_+)}{\Gamma(\omega_+)} \frac{\Gamma(\omega_+ + \omega_-)}{\Gamma(n + \omega_+ + \omega_-)}$$
$$\times \Phi[\omega_+ + n, \omega_+ + \omega_- + n; -g_+], \tag{57}$$

in agreement with the results of Iyer-Biswas *et al.* [34] and Raj *et al.* [17]. We remind the reader that in addition to reducing to known results in the special case of $Z=2$ states with a vanishing off-rate, the spectral method is valid for any number of states with arbitrary production rates.

III. GENE REGULATION

Next we consider a two-gene regulatory cascade, in which the production rate of the second gene is a function of the number of proteins of the first gene. As shown in previous work [31], a cascade of any length can be reduced to such a generalized two-dimensional system using the Markov approximation, which asserts that the probability distribution for a given node of the cascade should depend only weakly on the probability distributions of the distant nodes given the proximal nodes.

In the present section, we consider only the generalized two-dimensional equation and explore different approaches to solving it. The equation describes two genes, each with one expression state, with regulation encoded by a functional dependence of the downstream protein production rate on the upstream protein copy number. In Sec. III D we make an explicit connection between the on/off gene discussed in Sec. II C and the case when the functional dependence is a threshold. Finally, in Sec. III we combine the two types of models and consider a system with regulation and bursts.

A. Representations of the master equation

1. $|n,m\rangle$ basis

Defining n and m as the numbers of proteins produced by the first and second gene, respectively, the master equation

describing the time evolution of the joint probability distribution p_{nm} is [31]

$$\dot{p}_{nm} = g_{n-1}p_{n-1,m} + (n+1)p_{n+1,m} - (g_n + n)p_{nm}$$
$$+ \rho[q_n p_{n,m-1} + (m+1)p_{n,m+1} - (q_n + m)p_{nm}]. \quad (58)$$

The function q_n describes the regulation of the second species by the first, and the function g_n describes the effective autoregulation of the first species due either to a non-Poissonian input distribution or to effects further upstream in the case of a longer cascade [31]. Time is rescaled by the first gene's degradation rate so that each gene's production rate (g_n or q_n) is normalized by its respective degradation rate, and ρ is ratio of the second gene's degradation rate to that of the first. We impose no constraints on the form of g_n or q_n—they can be arbitrary nonlinear functions.

Summing Eq. (58) over m gives a simple recursion relation between g_n and p_n at steady state, from which explicit relations are easily identified. If p_n is known, g_n is found as

$$g_n = (n+1)\frac{p_{n+1}}{p_n}. \quad (59)$$

If on the other hand g_n is known, p_n is found as

$$p_n = \frac{p_0}{n!}\prod_{n'=0}^{n-1} g_{n'}, \quad (60)$$

with p_0 set by normalization. Note that if the first species obeys a simple birth-death process, $g_n = g = $ constant, and Eq. (60) reduces to the Poisson distribution.

In the current representation [Eq. (58)], which we denote the $|n,m\rangle$ basis, finding the steady-state solution for the joint distribution p_{nm} means finding the null space of an infinite (or, effectively for numerical purposes, very large) locally banded tridiagonal matrix. More precisely, defining N as the numerical cutoff in protein number n or m, the problem amounts to inverting an N^2-by-N^2 matrix, which is computationally taxing even for moderate cutoffs N.

In order to solve Eq. (58) more efficiently we will employ the spectral method. We begin as before by defining the generating function [22] $G(x,y) = \sum_{nm} p_{nm}x^n y^m$ over complex variables x and y or, in state notation,

$$|G\rangle = \sum_{nm} p_{nm}|n,m\rangle, \quad (61)$$

with inverse transform

$$p_{nm} = \langle n,m|G\rangle. \quad (62)$$

Summing Eq. (58) over n and m against $|n,m\rangle$ and employing the same operator notation as in Eqs. (12)–(15) yields

$$|\dot{G}\rangle = -\hat{H}|G\rangle, \quad (63)$$

where

$$\hat{H} = \hat{b}_n^+ \hat{b}_n^-(n) + \rho \hat{b}_m^+ \hat{b}_m^-(n) \quad (64)$$

and
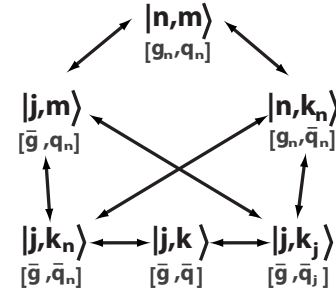
$$\hat{b}_n^+ = \hat{a}_n^+ - 1, \quad (65)$$

FIG. 2. Summary of the bases discussed in Sec. II A (black) and their gauge freedoms (barred parameters in gray). In the top row, neither the m nor the n sector is expanded in an eigenbasis; in the middle row, one sector is expanded; and in the bottom row, both sectors are expanded. The $|j,k\rangle$ basis can be viewed as a special case of the $|j,k_j\rangle$ or $|j,k_n\rangle$ basis with $\bar{q}_j = \bar{q}$ or $\bar{q}_n = \bar{q}$, respectively. The $|j,m\rangle$ basis is not discussed as it is not useful in simplifying the problem.

$$\hat{b}_m^+ = \hat{a}_m^+ - 1, \quad (66)$$

$$\hat{b}_n^-(n) = \hat{a}_n^- - \hat{g}_n, \quad (67)$$

$$\hat{b}_m^-(n) = \hat{a}_m^- - \hat{q}_n. \quad (68)$$

Here the regulation functions have been promoted to operators obeying $\hat{g}_n|n\rangle = g_n|n\rangle$ and $\hat{q}_n|n\rangle = q_n|n\rangle$, subscripts on operators denote the sector (n or m) on which they operate, and the arguments of $\hat{b}_n^-$ and $\hat{b}_m^-$ remind us that both are n dependent.

Equation (64) makes clear that if not for the n dependence of the operators the Hamiltonian $\hat{H}$ would be diagonalizable, or, equivalently, if g_n and q_n were constants the master equation would factorize into two individual birth-death processes. We may still, however, partition the full Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \quad (69)$$

where $\hat{H}_0$ is a diagonalizable part (and $\hat{H}_1$ is the corresponding deviation from the diagonal form), and expand $|G\rangle$ in the eigenbasis of $\hat{H}_0$ to exploit the diagonality.

As with the multistate system in Sec. II B, where we expand the solution in two different bases, there are several natural choices of eigenbasis of $\hat{H}_0$. Figure 2 summarizes these choices diagrammatically: starting in the $|n,m\rangle$ basis (at the top of Fig. 2), one may expand in eigenfunctions either the first species, yielding the $|j,m\rangle$ basis (left), or the second species, yielding the $|n,k_n\rangle$ basis (right; in general we allow the parameter defining the second species' eigenfunctions to depend on n to reflect the regulation of the second species by the first). From either the $|j,m\rangle$ or the $|n,k_n\rangle$ basis, one may expand in eigenfunctions the remaining species, yielding either the $|j,k_n\rangle$ basis (bottom left), in which the second species' eigenfunctions depend on the first species' copy number n or the $|j,k_j\rangle$ basis (bottom right), in which the second species' eigenfunctions depend on the first species' eigenmode number j. Both bases reduce to the $|j,k\rangle$ basis
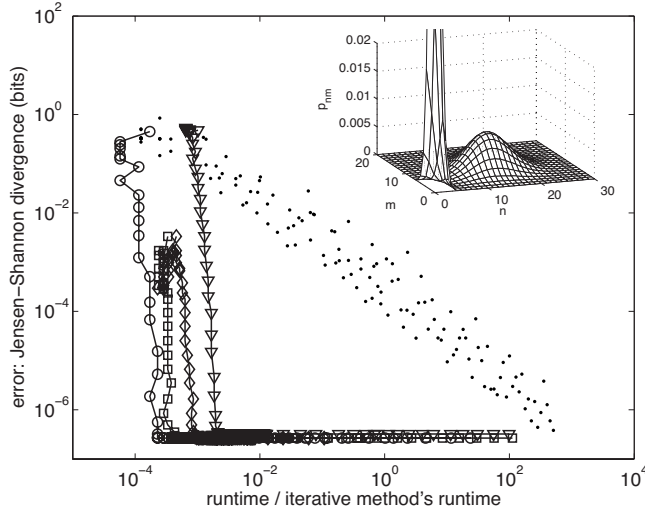
FIG. 3. Error vs runtime for the spectral method and stochastic simulation. Error is the Jensen-Shannon divergence [41] between p_{nm} obtained using the $|n,m\rangle$ basis (via iterative solution of the original master equation) and that obtained using the $|j,k\rangle$ basis [circles; cf. Eq. (80)], the $|j,k_j\rangle$ basis [triangles; cf. Eq. (90)], the $|n,k_n\rangle$ basis [squares; cf. Eq. (100)], the $|j,k_n\rangle$ basis [diamonds; cf. Eq. (114)], or stochastic simulation [28] (dots). Runtimes are scaled by that of the iterative solution, 150 s (in MATLAB). Spectral basis data is obtained by varying K, the cutoff in the eigenmode number k of the second gene; simulation data is obtained by varying the integration time. The input distribution $p_n = \pi_1 e^{-\lambda_1} \lambda_1^n/n! + (1-\pi_1)e^{-\lambda_2}\lambda_2^n/n!$ [from which g_n is calculated via Eq. (59)] is a mixture of two Poisson distributions with $\lambda_1 = 0.5$, $\lambda_2 = 15$, and $\pi_1 = 0.5$. The regulation function $q_n = q_- + (q_+ - q_-)n^\nu/(n^\nu + n_0^\nu)$ is a Hill function with $q_- = 1$, $q_+ = 11$, $n_0 = 7$, and $\nu = 2$. The gauge choices used (cf. Fig. 2) are $\bar{g} = \Sigma_n p_n g_n$, $\bar{q} = \Sigma_n p_n q_n$, $\bar{q}_n = q_n$, and $\bar{q}_j = \Sigma_n \langle j|n\rangle q_n \langle n|j\rangle$. The cutoffs used are $J = 80$ for the eigenmode number j of the first gene and $N = 50$ for the protein numbers n and m. Inset: the joint probability distribution p_{nm}. The peak at low protein number extends to $p_{00} \approx 0.1$.

(bottom center) when the parameter of the second species' eigenfunctions is a constant.

The $|n,m\rangle$ and $|j,m\rangle$ bases are less numerically useful than the other bases: as discussed above and detailed in Fig. 3, the $|n,m\rangle$ basis is numerically inefficient; and the $|j,m\rangle$ basis does not exploit the natural structure of the problem, since, unlike the other bases, it neither retains the tridiagonal structure in n nor gains a lower triangular structure in k (see the sections below). Each of the remaining bases, however, has preferable properties in terms of numerical stability and ability to represent the function sparsely yet accurately (either using a few values of n in the $|n,k_n\rangle$ basis or a few values of j in the $|j,k\rangle$, $|j,k_j\rangle$, or $|j,k_n\rangle$ bases, for example). Moreover, the equation of motion simplifies differently in each of the different bases. We present the derivations of the equations of motion in the following sections, beginning with the $|j,k\rangle$ basis, generalizing to the $|j,k_j\rangle$ basis, moving to the $|n,k_n\rangle$ basis, and ending with the $|j,k_n\rangle$ basis.

2. $|j,k\rangle$ basis

For expository purposes we start by recalling the spectral representation used in previous work [31]. We choose the

diagonal part of the Hamiltonian to correspond to two birth-death process with constant production rates $\bar{g}$ and $\bar{q}$,

$$\hat{H}_0 = \hat{b}_n^+ \bar{b}_n^- + \rho \hat{b}_m^+ \bar{b}_m^-, \tag{70}$$

where $\bar{b}_n^- = \hat{a}_n^- - \bar{g}$ and $\bar{b}_m^- = \hat{a}_m^- - \bar{q}$. As in Sec. II B, the gauge parameters $\bar{g}$ and $\bar{q}$ can be set arbitrarily without affecting the final solution; however, their values can affect the numerical stability of the method. For example, when the regulation q_n is a threshold function, a large discontinuity can narrow the range of $\bar{q}$ for which the method is numerically stable. The nondiagonal part,

$$\hat{H}_1 = \hat{b}_n^+ \hat{\Gamma}_n + \rho \hat{b}_m^+ \hat{\Delta}_n, \tag{71}$$

captures the deviations $\hat{\Gamma}_n = \bar{g} - \hat{g}_n$ and $\hat{\Delta}_n = \bar{q} - \hat{q}_n$ of the regulation functions from the constant rates. We expand the generating function as

$$|G\rangle = \sum_{jk} G_{jk}|j,k\rangle, \tag{72}$$

where $|j,k\rangle$ is the eigenbasis of $\hat{H}_0$, i.e.,

$$\hat{H}_0|j,k\rangle = (j + \rho k)|j,k\rangle. \tag{73}$$

The eigenbasis is parametrized by the rates $\bar{g}$ and $\bar{q}$, meaning in position space $\langle x|j\rangle$ is as in Eq. (38) and similarly for $\langle y|k\rangle$ with $x \to y$, $j \to k$, and $\bar{g} \to \bar{q}$.

With Eqs. (70)–(72), projecting the conjugate state $\langle j,k|$ onto Eq. (63) yields

$$\dot{G}_{jk} = -(j + \rho k)G_{jk} - \sum_{j'k'} \langle j|\hat{b}_n^+\hat{\Gamma}_n|j'\rangle\langle k|k'\rangle G_{j'k'}$$
$$- \rho\sum_{j'k'} \langle k|\hat{b}_m^+|k'\rangle\langle j|\hat{\Delta}_n|j'\rangle G_{j'k'} \tag{74}$$

[where the dummy indices j and k in Eq. (72) have been changed to j' and k', respectively, in Eq. (74)]. Recalling Eq. (24) and restricting attention to steady state, Eq. (74) becomes

$$0 = -(j + \rho k)G_{jk} - \sum_{j'} \Gamma_{j-1,j'}G_{j'k} - \rho\sum_{j'} \Delta_{jj'}G_{j',k-1}, \tag{75}$$

where the deviations have been rotated into the eigenbasis as

$$\Gamma_{jj'} = \langle j|\hat{\Gamma}_n|j'\rangle = \sum_n \langle j|n\rangle(\bar{g} - g_n)\langle n|j'\rangle, \tag{76}$$

$$\Delta_{jj'} = \langle j|\hat{\Delta}_n|j'\rangle = \sum_n \langle j|n\rangle(\bar{q} - q_n)\langle n|j'\rangle. \tag{77}$$

Equation (75) is subdiagonal in k and is therefore similar to Eq. (43) in that the last term acts as a source term. The subdiagonality is a consequence of the topology of the two-gene network: the first species regulates only itself (effectively) and the second species. Although the spectral method is fully applicable to systems with feedback, the subdiagonal structure is not preserved.

Equation (75) is initialized using

$$G_{j0} = \langle j, k=0|G\rangle = \sum_{nm} p_{nm} \langle j|n\rangle \langle 0|m\rangle \qquad (78)$$

$$= \sum_{n} p_n \langle j|n\rangle, \qquad (79)$$

[cf. Eq. (26)] with known p_n [cf. Eq. (60)], then solved at each subsequent k using the result for $k-1$. Equation (75) can be written in linear algebraic notation as

$$\vec{G}_k = -\rho(\mathbf{D}^k + \mathbf{S}^-\mathbf{\Gamma})^{-1}\mathbf{\Delta}\vec{G}_{k-1}, \qquad (80)$$

where $\vec{G}_k$ is a vector over j, bold denotes matrices, and $D_{jj'}^k = (j+\rho k)\delta_{jj'}$ and $S_{jj'}^- = \delta_{j-1,j'}$ are diagonal and subdiagonal matrices, respectively. Equation (80) makes clear that the solution involves only matrix multiplication and the inversion of a J-by-J matrix K times, where J and K are cutoffs in the eigenmode numbers j and k, respectively. In fact, if the first species obeys a simple birth-death process, i.e., $g_n = g =$ constant, setting $\bar{g}=g$ makes $\Gamma_{jj'}=0$, and since $\mathbf{D}^k$ is diagonal, the solution involves only matrix multiplication. The decomposition of the master equation into a linear algebraic equation results in huge gains in efficiency over direct solution in the $|n,m\rangle$ basis; the efficiency of all bases presented in this section is described in Sec. II B and illustrated in Fig. 3.

Recalling Eqs. (62) and (72), the joint distribution is retrieved from G_{jk} via the inverse transform

$$p_{nm} = \sum_{jk} \langle n|j\rangle G_{jk}\langle m|k\rangle, \qquad (81)$$

a computation again involving only matrix multiplication. The mixed product matrices $\langle n|j\rangle$ and $\langle m|k\rangle$ are computed as described in Appendix A.

3. $|j,k_j\rangle$ basis

The $|j,k\rangle$ basis treats both genes similarly by expanding each around a constant production rate. We may instead imagine an eigenbasis that more closely reflects the underlying asymmetry imposed by the regulation and make the basis of the second gene a function of that of the first. That is, we expand the first gene in a basis $|j\rangle$ with gauge $\bar{g}$ as before, but now we expand the second gene in a basis $|k_j\rangle$ with a j-dependent local gauge $\bar{q}_j$. We write the generating function as

$$|G\rangle = \sum_{jk} G_{jk}|j,k_j\rangle, \qquad (82)$$

and Eq. (63) at steady state becomes

$$0 = -\sum_{jk} G_{jk}[\hat{H}_0(j) + \hat{H}_1(j)]|j,k_j\rangle, \qquad (83)$$

where we have partitioned the Hamiltonian for each j as

$$\hat{H}_0(j) = \hat{b}_n^+\bar{b}_n^- + \rho\hat{b}_m^+\bar{b}_m^-(j), \qquad (84)$$

$$\hat{H}_1(j) = \hat{b}_n^+\hat{\Gamma}_n + \rho\hat{b}_m^+\hat{\Delta}_n(j), \qquad (85)$$

with $\bar{b}_n^- = \hat{a}_n^- - \bar{g}$ and $\hat{\Gamma}_n = \bar{g} - \hat{g}_n$ as before and now $\bar{b}_m^-(j) = \hat{a}_m^- - \bar{q}_j$ and $\hat{\Delta}_n(j) = \bar{q}_j - \hat{q}_n$. Note that this basis enjoys the eigenvalue equation

$$\hat{H}_0(j)|j,k_j\rangle = (j+\rho k)|j,k_j\rangle. \qquad (86)$$

Projecting the conjugate state $\langle j,k_j|$ onto Eq. (83) yields, after some simplification (cf. Appendix C), the equation of motion

$$0 = (j+\rho k)G_{jk} + \sum_{j'} \Gamma_{j-1,j'}G_{j'k}$$
$$+ \sum_{\ell=1}^{k}\sum_{j'}(\Gamma_{j-1,j'}V_{jj'}^\ell + \rho\Delta_{jj'}V_{jj'}^{\ell-1})G_{j',k-\ell}, \qquad (87)$$

where $\Gamma_{jj'}$ is as in Eq. (76) and

$$\Delta_{jj'} = \langle j|\hat{\Delta}_n(j)|j'\rangle = \sum_n \langle j|n\rangle(\bar{q}_j - q_n)\langle n|j'\rangle, \qquad (88)$$

$$V_{jj'}^\ell = \frac{(-Q_{jj'})^\ell}{\ell!}, \qquad (89)$$

with $Q_{jj'} = \bar{q}_j - \bar{q}_{j'}$. Equation (87) can be written linear algebraically as

$$\vec{G}_k = -(\mathbf{D}^k + \mathbf{S}^-\mathbf{\Gamma})^{-1}\sum_{\ell=1}^{k}[(\mathbf{S}^-\mathbf{\Gamma})*\mathbf{V}^\ell + \rho\mathbf{\Delta}*\mathbf{V}^{\ell-1}]\vec{G}_{k-\ell}, \qquad (90)$$

where $D_{jj'}^k$ and $S_{jj'}^-$ are defined as before [cf. Eq. (80)], and $*$ denotes an element-by-element matrix product. Once again, Eq. (90) is lower-triangular in k and requires only matrix multiplication and the inversion of a J-by-J matrix K times. $\vec{G}_k$ is initialized as in Eq. (79), and the kth term is computed from the previous $k' < k$ terms. The joint distribution is retrieved via the inverse transform

$$p_{nm} = \sum_{jk} \langle n|j\rangle G^{jk}\langle m|k_j\rangle. \qquad (91)$$

4. $|n,k_n\rangle$ basis

Expansion in either the $|j,k\rangle$ or the $|j,k_j\rangle$ basis conveniently turns the master equation into a lower-triangular linear algebraic equation and replaces cutoffs in protein number with cutoffs in eigenmode number (which can be smaller with appropriate choices of gauge). However, these bases sacrifice the original tridiagonal structure of the master equation in the copy number of first gene, n. Therefore we now consider a mixed representation, in which the first gene remains in protein number space $|n\rangle$, and we expand the second gene in an n-dependent eigenbasis $|k_n\rangle$,

$$|G\rangle = \sum_{nk} G_{nk}|n,k_n\rangle. \tag{92}$$

If the rate parameter of the $|k_n\rangle$ basis were the regulation function q_n, it would be natural to make $\hat{H}_0$ the entire Hamiltonian [Eq. (64)]. For generality we will instead allow the rate parameter of the $|k_n\rangle$ basis to be an arbitrary n-dependent local gauge $\bar{q}_n$, such that Eq. (63) at steady state naturally partitions as

$$0 = -\sum_{nk} G_{nk}[\hat{H}_0(n) + \hat{H}_1(n)]|n,k_n\rangle, \tag{93}$$

where

$$\hat{H}_0(n) = \hat{b}_n^+ \hat{b}_n^- + \rho \hat{b}_m^+ \bar{b}_m^-(n), \tag{94}$$

$$\hat{H}_1(n) = \rho \hat{b}_m^+ \hat{\Delta}_n(n), \tag{95}$$

with $\bar{b}_m^-(n) = \hat{a}_m^- - \bar{q}_n$ and $\hat{\Delta}_n(n) = \bar{q}_n - \hat{q}_n$. Note that $|n,k_n\rangle$ is not the eigenbasis of $\hat{H}_0(n)$ but rather $\hat{H}_0(n)|n,k_n\rangle$ retains the original tridiagonal structure in n, i.e.,

$$\hat{H}_0(n)|n,k_n\rangle = (g_n + n)|n,k_n\rangle - g_n|n+1,k_n\rangle$$
$$- n|n-1,k_n\rangle + \rho k|n,k_n\rangle, \tag{96}$$

where Eqs. (12), (13), (65), and (67) are recalled in applying the first term of $\hat{H}_0(n)$.

Projecting the conjugate state $\langle n,k_n|$ onto Eq. (93) yields, after some simplification (cf. Appendix C), the equation of motion

$$g_{n-1}G_{n-1,k} + (n+1)G_{n+1,k} - (g_n + n + \rho k)G_{nk}$$
$$= -g_{n-1}\sum_{\ell=1}^{k} V_{n\ell}^- G_{n-1,k-\ell} - (n+1)$$
$$\times \sum_{\ell=1}^{k} V_{n\ell}^+ G_{n+1,k-\ell} + \rho \Delta_n G_{n,k-1}, \tag{97}$$

where

$$\Delta_n = \bar{q}_n - q_n, \tag{98}$$

$$V_{n\ell}^\pm = \frac{(-Q_n^\pm)^\ell}{\ell!}, \tag{99}$$

and $Q_n^\pm = \bar{q}_n - \bar{q}_{n\pm1}$. Equation (97) can be written linear algebraically as

$$\vec{G}_k = (\mathbf{T}^k)^{-1}\{-(\mathbf{S}^-\vec{g}) * \text{diag}[\mathbf{V}^-(\mathbf{S}^-\widetilde{\mathbf{G}})^T]$$
$$- (\mathbf{S}^+\vec{n}) * \text{diag}[\mathbf{V}^+(\mathbf{S}^+\widetilde{\mathbf{G}})^T] + \rho\vec{\Delta} * \vec{G}_{k-1}\}, \tag{100}$$

where $\mathbf{X}^T$ indicates the transpose of $\mathbf{X}$, $*$ denotes an element-by-element product, $S_{nn'}^\pm = \delta_{n\pm1,n'}$ are super-(+) and subdiagonal (−) matrices, $T_{nn'}^k = g_{n-1}\delta_{n-1,n'} + (n+1)\delta_{n+1,n'} - (g_n + n + \rho k)\delta_{nn'}$ is a tridiagonal matrix, and $\widetilde{\mathbf{G}}$, which is $\mathbf{G}$ with the columns reversed (i.e., $\widetilde{G}_{n\ell} = G_{n,k-\ell}$), is built incrementally in k. As with the $|j,k\rangle$ and $|j,k_j\rangle$ bases, the kth term is slaved to

the previous $k' < k$ terms; in total the solution requires K inversions of an N-by-N matrix. However, here the task of inversion is simplified because the matrix to be inverted is tridiagonal. In fact, using the Thomas algorithm [42], we obtain an analytic solution for the case of constant production of the first gene and threshold regulation of the second gene, as described in Sec. II C.

The solution is initialized at $k=0$ using

$$G_{n0} = \langle n,(k=0)_n|G\rangle = \sum_{n'm} \langle n|n'\rangle\langle 0_n|m\rangle p_{n'm} = p_n \tag{101}$$

[cf. Eq. (26)], where p_n is known [cf. Eq. (60)], and the joint distribution is retrieved via the inverse transform

$$p_{nm} = \langle n,m|G\rangle = \sum_k G_{nk}\langle m|k_n\rangle. \tag{102}$$

5. $|j,k_n\rangle$ basis

We now consider a basis which employs both the constant-rate eigenfunctions $|j\rangle$ in the n sector and the n-dependent eigenfunctions $|k_n\rangle$ in the m sector. Expressing the joint distribution directly in terms of the eigenbasis expansion of the generating function [43], we write

$$p_{nm} = \sum_{jk} G_{jk}\langle n,m|j,k_n\rangle, \tag{103}$$

where, as in the $|j,k\rangle$ and $|j,k_j\rangle$ bases, the $|j\rangle$ are parametrized by the constant rate $\bar{g}$, and, as in the $|n,k_n\rangle$ basis, the $|k_n\rangle$ are parametrized by the arbitrary function $\bar{q}_n$. The inverse of Eq. (103) is

$$G_{jk} = \sum_{nm} p_{nm}\langle j,k_n|n,m\rangle. \tag{104}$$

Substituting Eq. (103) into Eq. (58) at steady state gives, after some simplification (cf. Appendix C),

$$0 = \sum_{jk} G_{jk}\langle n,m|\{-\hat{H}|j,k_n\rangle + \hat{a}_n^+\hat{g}_n|j,\delta_-k_n\rangle + \hat{a}_n^-|j,\delta_+k_n\rangle\}, \tag{105}$$

where $\hat{H}$ is defined as in Eq. (64), $\hat{a}_n^+$ and $\hat{a}_n^-$ act as in Eqs. (12)–(15), and

$$|\delta_\pm k_n\rangle \equiv |k_{n\pm1}\rangle - |k_n\rangle. \tag{106}$$

We may now partition $\hat{H} = \hat{H}_0 + \hat{H}_1$ as

$$\hat{H}_0(n) = \hat{b}_n^+ \bar{b}_n^- + \rho \hat{b}_m^+ \bar{b}_m^-(n), \tag{107}$$

$$\hat{H}_1(n) = \hat{b}_n^+ \hat{\Gamma}_n + \rho \hat{b}_m^+ \hat{\Delta}_n(n), \tag{108}$$

with $\bar{b}_n^- = \hat{a}_n^- - \bar{g}$ and $\hat{\Gamma}_n = \bar{g} - \hat{g}_n$ as in the $|j,k\rangle$ and $|j,k_j\rangle$ bases and $\bar{b}_m^-(n) = \hat{a}_m^- - \bar{q}_n$ and $\hat{\Delta}_n(n) = \bar{q}_n - \hat{q}_n$ as in the $|n,k_n\rangle$ basis. Noting that $|j,k_n\rangle$ is the eigenbasis of $\hat{H}_0(n)$, i.e.,

$$\hat{H}_0(n)|j,k_n\rangle = (j + \rho k)|j,k_n\rangle, \tag{109}$$

Equation (105) becomes, after more simplification (cf. Appendix C),

$$0 = -(j + \rho k)G_{j'k} - \sum_{j'} \Gamma_{j-1,j'}G_{j'k} - \rho\sum_{j'} \Delta_{jj'}G_{j',k-1}$$

$$+ \sum_{\pm} \sum_{j'} \sum_{\ell=1}^{k} \Lambda_{jj'}^{\pm\ell}G_{j',k-\ell}, \tag{110}$$

where $\Gamma_{jj'}$ is as in Eq. (76),

$$\Delta_{jj'} = \sum_n \langle j|n\rangle(\bar{q}_n - q_n)\langle n|j'\rangle, \tag{111}$$

$$\Lambda_{jj'}^{+\ell} = \sum_n \langle j|n\rangle(n+1)\langle n+1|j'\rangle V_{n\ell}^+, \tag{112}$$

$$\Lambda_{jj'}^{-\ell} = \sum_n \langle j|n\rangle g_{n-1}\langle n-1|j'\rangle V_{n\ell}^-, \tag{113}$$

and $V_{n\ell}^{\pm}$ is as in Eq. (99). Linear algebraically,

$$\vec{G}_k = (\mathbf{D}^k + \mathbf{S}^-\boldsymbol{\Gamma})^{-1}\left(-\rho\boldsymbol{\Delta}\vec{G}_{k-1} + \sum_{\pm} \sum_{\ell=1}^{k} \boldsymbol{\Lambda}^{\pm\ell}\vec{G}_{k-\ell}\right),$$

$$\tag{114}$$

with $D_{jj'}^k$ and $S_{jj'}^-$ defined as before [cf. Eq. (80)], revealing once again a lower-triangular equation (i.e., each kth term is slaved to the previous $k' < k$ terms) requiring only matrix multiplication and the inversion of a J-by-J matrix K times. Recalling Eq. (104), the scheme is initialized using

$$G_{j0} = \sum_{nm} p_{nm}\langle j|n\rangle\langle(k=0)_n|m\rangle = \sum_n p_n\langle j|n\rangle \tag{115}$$

[cf. Eq. (26)] with known p_n [cf. Eq. (60)], and the joint distribution is retrieved using Eq. (103).

B. Comparison of the representations

The spectral representations in Sec. II A produce equations of motion with similar levels of numerical complexity. In all cases, the original two-dimensional master equation has been reduced by the lower-triangular structure in the second gene's eigenmode number k to a hierarchy of evaluations of one-dimensional problems. The bases differ in the rate parameters or equivalently gauge freedoms that one is free to choose: the $|j,k\rangle$ basis requires two constants $\bar{g}$ and $\bar{q}$; the $|j,k_j\rangle$ basis requires $\bar{g}$ and a J-valued vector $\bar{q}_j$; the $|n,k_n\rangle$ basis requires a N-valued vector $\bar{q}_n$; and the $|j,k_n\rangle$ basis requires $\bar{g}$ and $\bar{q}_n$.

The bases also differ in the types of problems for which they are most suitable. For example, the $|j,k\rangle$, $|j,k_j\rangle$, and $|j,k_n\rangle$ bases, which all expand the parent species in eigenfunctions $|j\rangle$, are best when a cutoff in j is most appropriate, such as when the parent distribution is a Poisson. The $|n,k_n\rangle$ basis, on the other hand, is useful when a cutoff is n is most appropriate, such as when the parent species is concentrated at low protein number. Different bases are more robust to numerical errors for different regulation functions as well: the $|n,k_n\rangle$ and $|j,k_n\rangle$ bases, which both rely upon repeated manipulation of the object $Q_n^{\pm} = \bar{q}_n - \bar{q}_{n\pm 1}$, are best for smooth regulation functions, for which the differences between $\bar{q}_n$

and $\bar{q}_{n+1}$ are small; the $|j,k\rangle$ and $|j,k_j\rangle$ bases on the other hand, which involve the deviations $\bar{q} - q_n$ and $\bar{q}_j - q_n$ respectively, are less susceptible to numerical error given sharp regulation functions, such as a threshold.

As indicated in Fig. 2, the $|j,k\rangle$ basis can be viewed as a special case of either the $|j,k_j\rangle$ basis with $\bar{q}_j = \bar{q}$ [in which case Eq. (87) reduces to Eq. (75)] or of the $|j,k_n\rangle$ basis with $\bar{q}_n = \bar{q}$ [in which case Eq. (110) reduces to Eq. (75)]. Although possible in principle, expanding in the $|j,m\rangle$ basis does not exploit the natural structure of the problem, since it neither retains the tridiagonal structure in n nor gains the lower triangular structure in k. This example explicitly shows that not all bases are good candidates for simplifying the master equation.

The strength of all the spectral bases discussed in this section and of the proposed spectral method in general is that it allows for a fast and accurate calculation of full steady-state probability distributions of the number of protein molecules in a gene regulatory network. In Fig. 3 we demonstrate this property for the two-gene system by plotting error versus computational runtime for each spectral basis, as well as for a stochastic simulation using a varying step Monte Carlo procedure [28]. For error we use the Jensen-Shannon divergence [41] (a measure in bits between two probability distributions) between the distribution p_{nm} computed in the $|n,m\rangle$ basis (via iterative solution of the original master equation) and the distribution computed either via the spectral formulas in this section or by stochastic simulation. We plot this measure against the runtime of each method scaled by the runtime of the iterative solution in the $|n,m\rangle$ basis (all numerical experiments are performed in MATLAB). We find that the computations via the spectral bases achieve accuracy up to machine precision $\sim 10^3$–10^4 times faster than the iterative method's runtime and $\sim 10^7$–10^8 times faster than the runtime necessary for the stochastic simulation to achieve the same accuracy. Computation in the $|j,k\rangle$ basis is most efficient since its equation of motion is simplest [cf. Eq. (80)]; the $|j,k_j\rangle$ and $|j,k_n\rangle$ bases tend to be slightly less efficient since they require inner loops over ℓ [cf. Eqs. (90) and (114)]. Figure 3 demonstrates the tremendous gain in performance achieved by the joint analytic-numerical spectral method over traditional simulation approaches.

C. Analytic solution

In general, the equations of motion in the spectral representations [Eqs. (75), (87), (97), and (110)] need to be evaluated numerically. In the case of the $|n,k_n\rangle$ basis, however, we can exploit the tridiagonal structure of Eq. (97) to find an exact analytic solution. Specifically, in the case of a Poisson parent ($g_n = g = $ const) and for threshold regulation, i.e.,

$$q_n = \begin{cases} q_- & \text{for } n \leq n_0 \\ q_+ & \text{for } n > n_0, \end{cases} \tag{116}$$

setting $\bar{q}_n = q_n$ makes Eq. (97)

$$gG_{n-1,k} + (n+1)G_{n+1,k} - (\rho k + g + n)G_{nk}$$

$$= -g\phi_k^-\delta_{nn_1} - n_1\phi_k^+\delta_{nn_0}, \tag{117}$$

where

$$\phi_k^- = \sum_{\ell=1}^{k} \frac{(-\Delta)^\ell}{\ell!} G_{n_0, k-\ell}, \qquad (118)$$

$$\phi_k^+ = \sum_{\ell=1}^{k} \frac{\Delta^\ell}{\ell!} G_{n_1, k-\ell}, \qquad (119)$$

$\Delta = q_+ - q_-$, and $n_1 = n_0 + 1$. Equation (117) is solved using the tridiagonal matrix algorithm (also called the Thomas algorithm [42]), as described in detail in Appendix D. The result is an analytic expression for the kth column of G_{nk} in terms of its previous columns (i.e., the matrix inversion has been performed explicitly),

$$G_{nk} = \frac{n_1}{g} \frac{(n_0-1)!}{n!} \frac{\eta_n^k}{\eta_{n_0-1}^k}$$
$$\times \begin{cases} (\phi_k^+ + f_k F_{n_1}^k)/(\epsilon_{n_0}^k - 1) & n \le n_0 \\ f_k F_n^k / \prod_{i=n_0}^{n-1} (\epsilon_i^k - 1) & n > n_0, \end{cases} \qquad (120)$$

where

$$f_k = \phi_k^+ - \frac{g n_0}{n_1} \frac{\eta_{n_0-1}^k}{\eta_{n_0}^k} (\epsilon_{n_0}^k - 1) \phi_k^-, \qquad (121)$$

$$F_n^k = \sum_{i=0}^{N-n} \prod_{\ell=n}^{n+i} \frac{1}{\epsilon_\ell^k - 1}, \qquad (122)$$

$$\epsilon_n^k = \frac{\rho k + g + n}{g n} \frac{\eta_n^k}{\eta_{n-1}^k}, \qquad (123)$$

$$\eta_n^k = \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} g^{n-i} \prod_{\ell=0}^{i-1} (\rho k + \ell) \qquad (124)$$

$$= \frac{1}{\Gamma(\rho k)} \int_0^t dt\, e^{-t} t^{(\rho k-1)} (g+t)^n, \qquad (125)$$

and N is the cutoff in protein number n. Along with the analytic form of the mixed product

$$\langle m | k_n \rangle = (-1)^k e^{-q_n} (q_n)^m k!$$
$$\times \sum_{\ell=0}^{\min(m,k)} \frac{1}{\ell!(m-\ell)!(k-\ell)!(-q_n)^\ell} \qquad (126)$$

(cf. Appendix A), Eq. (120) in the limit $N \to \infty$ constitutes an exact analytic solution for the joint distribution p_{nm} as calculated using Eq. (102).

D. Threshold-regulated gene approximates the on/off gene

If a gene is regulated via a threshold function [cf. Eq. (116)], its steady-state protein distribution p_m can be well approximated by the two-state process discussed in Sec. II C. To make the connection clear, we first observe that the

off-state ($z = -$) corresponds to the first gene expressing the same or fewer proteins n than the threshold n_0, i.e.,

$$p_m^- = \sum_{n \le n_0} p_{nm}, \qquad (127)$$

and the on state ($z = +$) corresponds to the first gene expressing more proteins than the threshold, i.e.,

$$p_m^+ = \sum_{n > n_0} p_{nm}. \qquad (128)$$

The dynamics of $p_m^\pm$ are then obtained by summing the master equation for two-gene regulation [Eq. (58)] over either all $n \le n_0$ or all $n > n_0$, giving

$$\dot{p}_m^\pm = \rho[q_\pm p_{m-1}^\pm + (m+1)p_{m+1}^\pm - (q_\pm + m)p_m^\pm]$$
$$\mp n_1 p_{n_1 m} \pm g_{n_0} p_{n_0 m}, \qquad (129)$$

where $n_1 = n_0 + 1$. Making the approximations

$$\frac{p_m^-}{\pi_-} = p(m|-) \approx p(m|n_0) = \frac{p_{n_0 m}}{p_{n_0}}, \qquad (130)$$

$$\frac{p_m^+}{\pi_+} = p(m|+) \approx p(m|n_1) = \frac{p_{n_1 m}}{p_{n_1}}, \qquad (131)$$

where

$$\pi_- = \sum_m p_m^- = \sum_{n \le n_0} p_n, \qquad (132)$$

$$\pi_+ = \sum_m p_m^+ = \sum_{n > n_0} p_n \qquad (133)$$

are the total probabilities of being in the off and on states, respectively, and noting from Eq. (59) that $g_{n_0} = n_1 p_{n_1}/p_{n_0}$, Eq. (129) at steady state becomes

$$0 = q_z p_{m-1}^z + (m+1)p_{m+1}^z - (q_z + m)p_m^z + \sum_{z'} \Omega_{zz'} p_m^{z'}, \qquad (134)$$

with $z = \pm$ and

$$\mathbf{\Omega} = \begin{pmatrix} -\omega_+ & \omega_- \\ \omega_+ & -\omega_- \end{pmatrix}, \qquad (135)$$

where

$$\omega_\pm = \frac{n_1 p_{n_1}}{\rho \pi_\mp}. \qquad (136)$$

Equations (134) and (135) have the same form as Eqs. (1) and (47) at steady state with $n \to m$ and $g \to q$, and Eq. (136) relates the effective switching rates $\omega_\pm$ to input and regulation parameters p_{n_1}, $\pi_\pm$, and n_1, and the ratio ρ of the degradation rate of the second gene to that of the first. Note that Eq. (136) satisfies
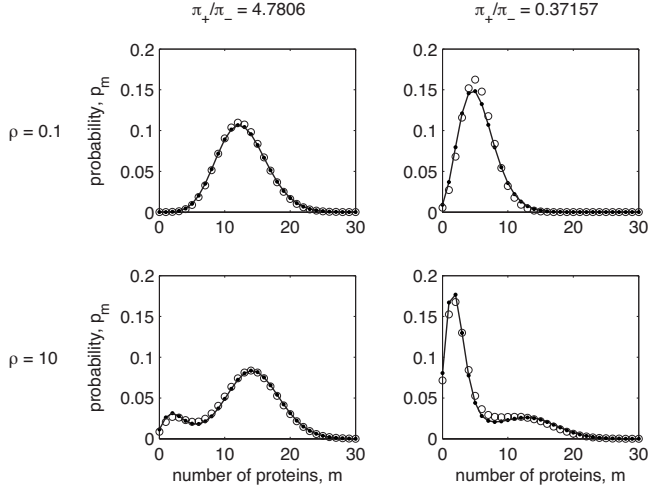
FIG. 4. Protein distributions for a gene regulated by a threshold function [dots; calculated via Eq. (75)] and a gene with two stochastic states [circles; calculated via Eq. (43)]. The relationship between regulation parameters and state transition rates is given by Eq. (136). In all panels the input to the regulation is a Poisson distribution with mean $g=7$, and the regulation rates [cf. Eq. (116)] are $q_-=2$ and $q_+=15$. In the first column the threshold is $n_0=4$ making $\pi_+=0.827 > \pi_-=0.173$; in the second column $n_0=8$ making $\pi_+=0.271 < \pi_-=0.729$. In the first row the ratio of the second gene's degradation rate to that of the first is $\rho=0.1$; in the second row $\rho=10$.

$$\frac{\pi_-}{\pi_+} = \frac{\omega_-}{\omega_+}, \tag{137}$$

in agreement with Eq. (48), and exhibits the intuitive behavior that increasing ρ (i.e., decreasing the relative response rate of the first gene) is equivalent to decreasing the switching rates $\omega_\pm$.

A comparison of the distributions of a threshold-regulated gene with those of an on/off gene for various parameter settings reveals that Eqs. (130) and (131) are a good approximation. Figure 4 shows a demonstration for a threshold-regulated system with a Poisson input distribution. In the first column, the mean g of the input lies above the threshold n_0, making the output more likely to be in the on-state, i.e., $\pi_+ > \pi_-$; in the second column, $g < n_0$, making $\pi_+ < \pi_-$. In the first row $\rho < 1$; in the second row $\rho > 1$, corresponding to lower effective switching rates $\omega_\pm$ and producing bimodal distributions with peaks near the on/off rates $q_\pm$. In all examples, the approximation as a two-state process with switching rates given by Eq. (136) agrees well with the actual output from threshold regulation.

IV. REGULATION WITH BURSTS

The final system we consider combines the multistate process used to model bursts of expression in Sec. I with gene regulation as discussed in Sec. II. Specifically we consider a system of two species, with protein numbers n and m, existing in Z possible states, distinguished by the settings of the two production rates g_z and q_z respectively, where $1 \le z \le Z$. Regulation is achieved by allowing the rates of transition among states affecting the production of the second gene to depend on the number n of proteins expressed by the first gene. Recalling Eqs. (1) and (58), the master equation describing the evolution of the joint probability distribution p_{nm}^z reads

$$\dot{p}_{nm}^z = g_z p_{n-1,m}^z + (n+1)p_{n+1,m}^z - (g_z+n)p_{nm}^z + \rho[q_z p_{n,m-1}^z \\ + (m+1)p_{n,m+1}^z - (q_z+m)p_{nm}^z] + \sum_{z'} \Omega_{zz'}(n)p_{nm}^{z'}, \tag{138}$$

where the dependence of the stochastic matrix $\Omega_{zz'}$ on n incorporates the regulation.

As with the previously discussed models, Eq. (138) benefits from spectral expansion, and for simplicity we present only the formulation in the $|j,k\rangle$ basis, parametrized by constant rates $\bar{g}$ and $\bar{q}$ respectively, as in Secs. II B and III A 2. As before the first step is to define the generating function

$$|G_z\rangle = \sum_{nm} p_{nm}^z |n,m\rangle, \tag{139}$$

with which Eq. (138), upon summing over n and m against $|n,m\rangle$, becomes

$$|\dot{G}_z\rangle = -\hat{H}_z|G_z\rangle + \sum_{z'} \hat{\Omega}_{zz'}|G_{z'}\rangle, \tag{140}$$

where

$$\hat{H}_z = \hat{b}_n^+ \hat{b}_{nz}^- + \rho \hat{b}_m^+ \hat{b}_{mz}^- \tag{141}$$

$$\hat{b}_n^+ = \hat{a}_n^+ - 1, \tag{142}$$

$$\hat{b}_m^+ = \hat{a}_m^+ - 1, \tag{143}$$

$$\hat{b}_{nz}^- = \hat{a}_n^- - g_z, \tag{144}$$

$$\hat{b}_{mz}^- = \hat{a}_m^- - q_z, \tag{145}$$

and $\hat{\Omega}_{zz'}$ is $\Omega_{zz'}(n)$ with every instance of n replaced by the number operator $\hat{a}_n^+ \hat{a}_n^-$. Defining $\bar{b}_n^- = \hat{a}_n^- - \bar{g}$ and $\bar{b}_m^- = \hat{a}_m^- - \bar{q}$, we partition the Hamiltonian as $\hat{H}_z = \hat{H}_0 + \hat{H}_1^z$, with

$$\hat{H}_0 = \hat{b}_n^+ \bar{b}_n^- + \rho \hat{b}_m^+ \bar{b}_m^- \tag{146}$$

as the operator of which $|j,k\rangle$ is the eigenbasis, i.e.,

$$\hat{H}_0|j,k\rangle = (j+\rho k)|j,k\rangle \tag{147}$$

and

$$\hat{H}_1^z = \hat{b}_n^+ \Gamma_z + \rho \hat{b}_m^+ \Delta_z \tag{148}$$

capturing the deviations $\Gamma_z = \bar{g} - g_z$ and $\Delta_z = \bar{q} - q_z$ of the constant rates from the state-dependent rates. Upon expanding the generating function in the eigenbasis,

$$|G_z\rangle = \sum_{jk} G^z_{jk}|j,k\rangle, \tag{149}$$

and taking dummy indices $j \to j'$ and $k \to k'$, projecting the conjugate state $\langle j,k|$ onto Eq. (140) gives

$$\dot{G}^z_{jk} = -(j + \rho k)G^z_{jk} - \Gamma_z G^z_{j-1,k} - \Delta_z G^z_{j,k-1}$$
$$+ \sum_{z'} \sum_{j'} \langle j|\hat{\Omega}_{zz'}|j'\rangle G^{z'}_{j'k}, \tag{150}$$

where the components of $\hat{\Omega}_{zz'}$ need only be evaluated in the j sector, not the k sector, because the transition rates depend on only n, not m [cf. Eq. (138)]. Like Eqs. (43) and (75), Eq. (150) is subdiagonal in k and thus far more efficient to solve than the original master equation [Eq. (138)] as we demonstrate for a special case in the next section. The joint distribution is retrieved from G^z_{jk} via inverse transform,

$$p^z_{nm} = \sum_{jk} \langle n|j\rangle G^z_{jk}\langle m|k\rangle, \tag{151}$$

with the mixed products calculated as in Appendix A.

A. Four-state process

As a simple example of the model in Eq. (138), we consider a system in which each of the two species has an on state and an off state, and the transition rate of the second species to its on state is a function of the number of copies of the first species. This system models both (i) a single gene for which the production of proteins depends on the number of transcripts, and each is produced in on and off states by the binding and unbinding of ribosomes and RNA polymerase, respectively, and (ii) one gene regulating another with each undergoing burstlike expression.

There are a total of $Z=4$ states, i.e.,

$$p^z_{nm} = (p^{--}_{nm}, p^{+-}_{nm}, p^{-+}_{nm}, p^{++}_{nm}), \tag{152}$$

where the first signed index denotes the state of the first gene (with protein count n) and the second signed index denotes the state of the second gene (with protein count m). Defining $g_\pm$ as the production rates of the first species in its on (+) and off states (−), and similarly $q_\pm$ for the second species, the production rates of the $Z=4$ states are

$$g_z = (g_-, g_+, g_-, g_+), \tag{153}$$

$$q_z = (q_-, q_-, q_+, q_+). \tag{154}$$

Defining $\omega_\pm$ as the transition rates of the first species to (+) and from (−) its on state, and similarly $\alpha_\pm$ for the second species, the transition matrix takes the form

$$\Omega_{zz'}(n) = \begin{pmatrix} -\omega_+ - \alpha_+(n) & \omega_- & \alpha_- & 0 \\ \omega_+ & -\omega_- - \alpha_+(n) & 0 & \alpha_- \\ \alpha_+(n) & 0 & -\omega_+ - \alpha_- & \omega_- \\ 0 & \alpha_+(n) & \omega_+ & -\omega_- - \alpha_- \end{pmatrix}. \tag{155}$$

The simple form $\alpha_+(n) = cn^\nu$ for constant c and integer ν corresponds to the first species activating the second as a multimer, with ν the order of the multimerization. In the limit of fast switching this description reduces to a Hill function with cooperativity ν [23]. Recalling that $\hat{b}^+_n = \hat{a}^+_n - 1$ and $\bar{b}^-_n = \hat{a}^-_n - \bar{g}$, the n-dependent terms of $\langle j|\hat{\Omega}_{zz'}|j'\rangle$ are evaluated as

$$\langle j|\alpha_+(\hat{a}^+_n \hat{a}^-_n)|j'\rangle = c\langle j|[(\hat{b}^+_n + 1)(\bar{b}^-_n + \bar{g})]^\nu|j'\rangle. \tag{156}$$

Since $\hat{b}^+_n$ and $\bar{b}^-_n$ raise and lower $|j'\rangle$ states respectively [cf. Eqs. (22) and (23)], the modified transition matrix $\langle j|\hat{\Omega}_{zz'}|j'\rangle$ is nearly diagonal, with nonzero terms only for $|j-j'| \leq \nu$.

Equation (150) at steady state,

$$-(j + \rho k)G^z_{jk} - \Gamma_z G^z_{j-1,k} + \sum_{z'} \sum_{j'} \langle j|\hat{\Omega}_{zz'}|j'\rangle G^{z'}_{j'k} = \Delta_z G^z_{j,k-1}, \tag{157}$$

is solved successively in k, requiring the inversion of a $4J$-by-$4J$ matrix K times. It is initialized at $k=0$ by comput-

ing the null space of the left-hand side and normalizing with $\Sigma_z G^z_{00} = 1$ [cf. Eq. (35)]. The joint distribution p^z_{nm} is retrieved via inverse transform [Eq. (151)].

With $\nu=2$, a typical solution of Eq. (157) takes a few seconds (in MATLAB), which, depending on the cutoff N, is $\sim 10^2 - 10^3$ times faster than direct solution of the master equation [Eq. (138)] by iteration for equivalent accuracy. The advantage of such a large efficiency gain is that it allows repeated evaluations of the governing equation, necessary for parameter inference or optimization [31]. We demonstrate this possibility in the next section by finding and interpreting the solutions that optimize the information flow from the first to the second species.

B. Information-optimal solution

Cells use regulatory processes to transmit relevant information from one species to the next [44–48]. Information processing is quantified by the mutual information I, which, between the first and second species in the four-state process, is
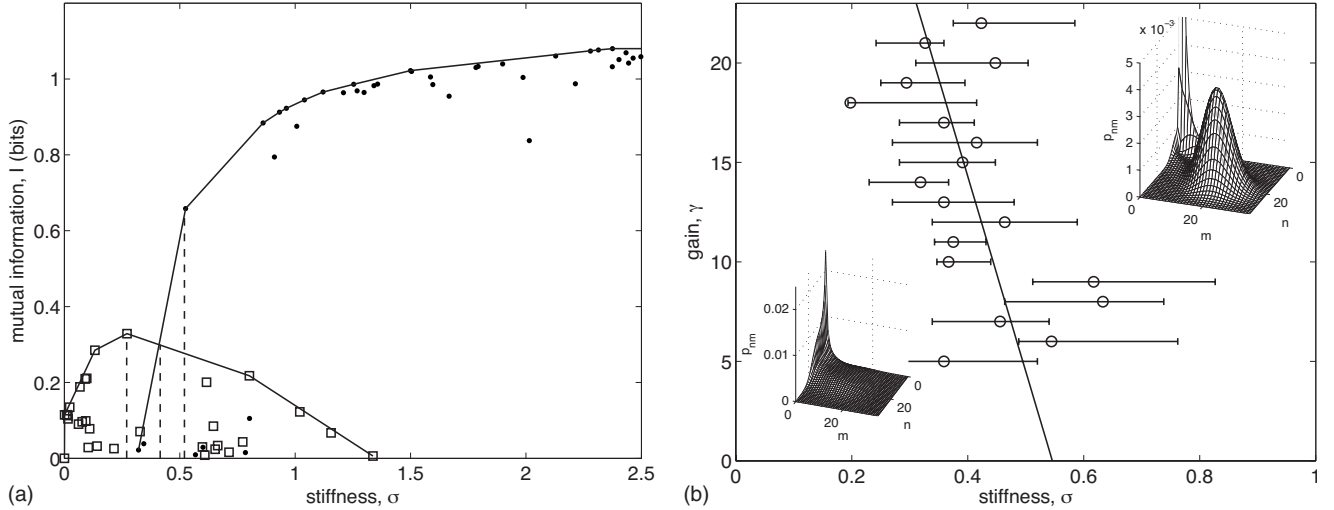
FIG. 5. (A) Mutual information I [cf. Eq. (158)] versus stiffness σ [cf. Eq. (160)] for fixed gain [$\gamma=16$, cf. Eq. (159)] obtained by optimizing Eq. (161) for λ values between 10^{-3} and 10^{1}. Squares denote solutions whose joint distribution p_{nm} has one peak (cf. B, lower left inset), and dots denote solutions for which p_{nm} has two peaks (cf. B, upper right inset). Solid lines show the convex hulls of the one- and two-peaked solutions. Dotted lines indicate the stiffness value at which the hulls intersect and the stiffness values of the hull points to the left and right of the intersection. (B) Phase diagram between one- and two-peaked optimal solutions in the gain-stiffness plane. Circles and left and right error bars at each gain are determined by the stiffness values at the intersection of the one- and two-peak convex hulls and at the hull points to the left and right of the intersection, respectively (see dotted lines for the example case in A). Solid line shows a line of best fit. Insets show examples of one-(lower left) and two-peaked (upper right) optimal distributions p_{nm}.

$$I = \sum_{nm} p_{nm} \log_2 \frac{p_{nm}}{p_n p_m}, \qquad (158)$$

where the distributions p_{nm}, p_n, and p_m are obtained from summing the joint distribution p_{nm}^z [cf. Eq. (151)], and the log is taken with base 2 to give I in bits.

Upon optimization of I for the four-state process, two distinct types of optimal solutions become clear: those in which the distribution p_{nm} has one peak and those in which p_{nm} has two peaks. The former occur when copy number is constrained to be low, and switching rates are constrained to be near the decay rates of both species, producing a single peak at low copy number [see lower left inset of Fig. 5(B)]. As these constraints are lifted, it is optimal for the switching rates of the parent species to become much less than the decay rate. The slow switching produces a second peak whose location is specified by the on rate of each species [see upper right inset of Fig. 5(B)].

To quantify the transition between the two types of solutions, we numerically optimized mutual information over parameters g_+, q_+, ω_-, ω_+, α_-, and c (the off-rates g_- and q_- were fixed at 0; the cooperativity ν was fixed at 2; and the decay rate ratio ρ was fixed at 1). Information may always be trivially optimized by allowing infinite copy number or arbitrary separation of relevant time scales. We limit copy number by constraining the gain

$$\gamma = \frac{\Gamma + \Delta}{2}, \qquad (159)$$

defined as the average of the parent gain $\Gamma = g_+ - g_-$ and the child gain $\Delta = q_+ - q_-$. Since $g_- = q_- = 0$, the maximum number of particles is dictated by the on rates g_+ and q_+ and thus

constraining γ limits the copy number. We limit separation between the switching time scales and the decay time scales by constraining the stiffness

$$\sigma = \frac{1}{4}(|\log_{10} \omega_-| + |\log_{10} \omega_+| + |\log_{10} \alpha_-| + |\log_{10}[\alpha_+ \langle n^\nu \rangle]|), \qquad (160)$$

where the average $\langle n^\nu \rangle$ is taken over p_n. Stiffness σ is the average of the absolute deviation of (the logs of) all four switching rates from the (unit) decay rates so constraining σ prevents fast or slow switching. Gain is fixed by varying g_+ and q_+ such that γ is a constant, and stiffness is constrained by optimizing the objective function

$$\mathcal{L} = I - \lambda \sigma \qquad (161)$$

for a given value of the Lagrange multiplier λ.

As shown in Fig. 5(A), one-peaked solutions are more informative at low stiffness, while two-peaked solutions are more informative at high stiffness. We compute the convex hulls of the one- and two-peaked data to remove suboptimal solutions, and the transition occurs at the stiffness value at which the convex hulls intersect [cf. Fig. 5(A)]. Repeating this procedure for many choices of gain allows one to trace out the phase transition shown Fig. 5(B), which makes clear that one-peaked solutions are most informative at low stiffness, two-peaked solutions are most informative at high stiffness, and the critical stiffness decreases weakly with increasing gain.

Examining the two-peaked solutions, which are optimal at high stiffness, we find that the marginal probability to be in the $(-,+)$ state, where the first gene is off and the second gene is on, $\pi^{-,+} = \Sigma_{nm} p_{nm}^{-,+}$, is much smaller than the marginal

probabilities to be in the other three states $[\pi^{-,+}/\pi^z < 0.01,$ where $z = \{(+,+),(+,-),(-,-)\}]$. This result states that it is unlikely for the regulated gene to be expressing proteins at a high level, if the regulator protein or mRNA (depending on the interpretation of the model) is not being expressed. Optimizing information, we find rates which result in this intuitive solution.

V. CONCLUSIONS

The presented spectral method exploits the linearity of the master equation to solve for probability distributions directly by expanding in the natural eigenfunctions the linear operator. We demonstrate the method on three models of gene expression: a single gene with multiple expression states, a gene regulatory cascade, and a model that combines multistate expression with explicit regulation through binding of transcription factor proteins.

The spectral method permits huge computational gains over simulation. As demonstrated for all spectral expansions of the two-gene cascade (cf. Fig. 3), directly solving for the distribution via the spectral method is $\sim 10^7$–10^8 times faster than building the distribution from samples using a simulation technique. This massive speedup makes possible optimization and inference problems requiring full probability distributions that were not computationally feasible previously. For example, by optimizing information flow in a two-gene cascade in which both parent and child undergo two-state production, we reveal a transition from a one-peaked to a two-peaked joint probability distribution when constraints on protein number and time scale separation are relaxed. We emphasize that this optimization would not have been possible without the efficiency of the spectral method.

The spectral method also makes explicit the linear algebraic structure underlying the master equation. In many cases, such as in two-state bursting and the two-gene threshold regulation problem, this leads to analytic solutions. In general, such as shown in the case of the linear cascade, this leads to a set of natural bases for expansion of the generating function and reveals the features of each basis that are better suited to different types of problems. Specifically, bases in which the parent species is expanded in eigenfunctions are best when the parent distribution is Poissonian, and bases in which the parent is left in protein number space are best when the parent distribution is concentrated at low protein number. As well, bases in which the eigenfunctions of the child depend on the number of copies of the parent's protein are best suited for smooth regulation functions, whereas a basis in which the eigenfunctions of the child are parametrized by a constant is more numerically robust for sharp regulation functions such as thresholds. In all cases the linear algebraic structure of the spectral decomposition yields numerical prescriptions that greatly outperform simulation techniques. We anticipate that the computational speedup of the method, as well as the removal of the statistical obstacle of density estimation inherently limiting simulation-based approaches, will make spectral methods such as those demonstrated here useful in addressing a wide variety of biological questions regarding accurate and efficient modeling of noisy information transmission in biological systems.

APPENDIX A

In this appendix we describe two ways to compute the mixed products $\langle n|j \rangle$ and $\langle j|n \rangle$ between the protein number states $|n\rangle$ and the eigenstates $|j\rangle$: by direct evaluation and by recursive updating.

The direct evaluation follows from Eqs. (8)–(10) and (20), and the fact that repeated derivatives of a product follow a binomial expansion. Introducing g as the rate parameter for the $|j\rangle$ states,

$$\langle n|j \rangle = \oint \frac{dx}{2\pi i} \langle n|x \rangle \langle x|j \rangle \tag{A1}$$

$$= \oint \frac{dx}{2\pi i} \frac{e^{g(x-1)}(x-1)^j}{x^{n+1}} \tag{A2}$$

$$= \frac{1}{n!} \partial_x^n [e^{g(x-1)}(x-1)^j]_{x=0} \tag{A3}$$

$$= \frac{1}{n!} \sum_{\ell=0}^{n} \frac{n!}{\ell!(n-\ell)!} \partial_x^{n-\ell}[e^{g(x-1)}]_{x=0}$$
$$\times \partial_x^{\ell}[(x-1)^j]_{x=0} \tag{A4}$$

$$= \sum_{\ell=0}^{n} \frac{1}{\ell!(n-\ell)!}[g^{n-\ell}e^{-g}]$$
$$\times \left[\frac{j!}{(j-\ell)!}(-1)^{j-\ell}\theta(j-\ell+1) \right] \tag{A5}$$

$$= (-1)^j e^{-g} g^n j! \, \xi_{nj}, \tag{A6}$$

where

$$\xi_{nj} = \sum_{\ell=0}^{\min(n,j)} \frac{1}{\ell!(n-\ell)!(j-\ell)!(-g)^{\ell}}. \tag{A7}$$

Similarly, noting Eqs. (7) and (21),

$$\langle j|n \rangle = n!(-g)^j \xi_{nj}, \tag{A8}$$

with ξ_{nj} as in Eq. (A7). Equations (A6) and (A8) clearly reduce to Eqs. (26) and (27) for the special case $j=0$.

It is more computationally efficient to take advantage of the selection rules in Eqs. (12)–(15) and (22)–(25) to compute the mixed products recursively. For example, using Eqs. (14), (16), and (22),

$$\langle n|j+1\rangle = \langle n|\hat{b}^+|j\rangle = \langle n|(\hat{a}^+ - 1)|j\rangle = \langle n-1|j\rangle - \langle n|j\rangle, \tag{A9}$$

which can be initialized using $\langle n|0\rangle = e^{-g}g^n/n!$ [cf. Eq. (A6)] and updated recursively in j. Equation (A9) makes clear that in n space the $(j+1)$th mode is simply the (negative of the) discrete derivative of the jth mode.

Alternatively, Eqs. (15), (17), and (23) give

$$(n+1)\langle n+1|j\rangle = \langle n|\hat{a}^-|j\rangle = \langle n|(\hat{b}^- + g)|j\rangle$$
$$= j\langle n|j-1\rangle + g\langle n|j\rangle, \tag{A10}$$

which can be initialized using $\langle 0|j\rangle = (-1)^j e^{-g}$ [cf. Eq. (A6)] and updated recursively in n.

One may similarly derive recursion relations for $\langle j|n\rangle$, i.e.,

$$\langle j|n+1\rangle = \langle j-1|n\rangle + \langle j|n\rangle, \tag{A11}$$

$$(j+1)\langle j+1|n\rangle = n\langle j|n-1\rangle - g\langle j|n\rangle, \tag{A12}$$

initialized with $\langle j|0\rangle = (-g)^j/j!$ or $\langle 0|n\rangle = 1$, respectively [cf. Eq. (A8)] and updated recursively in n or j, respectively.

One may also use the full birth-death operator $\hat{b}^+\hat{b}^-$ to derive the recursion relations

$$(n+1)\langle n+1|j\rangle = (g+n-j)\langle n|j\rangle - g\langle n-1|j\rangle, \tag{A13}$$

$$g\langle j|n+1\rangle = (g+n-j)\langle j|n\rangle - n\langle j|n-1\rangle, \tag{A14}$$

initialized with $\langle 0|j\rangle = (-1)^j e^{-g}$ and $\langle 1|j\rangle = (-1)^j e^{-g}(g-j)$ [cf. Eq. (A6)], and $\langle j|0\rangle = (-g)^j/j!$ and $\langle j|1\rangle = (-g)^j(1-j/g)/j!$ [cf. Eq. (A8)], respectively, and updated recursively in n. We find Eqs. (A13) and (A14) are more numerically stable than Eqs. (A9)–(A12), as the former are two-term recursion relations while the latter are one-term recursion relations.

APPENDIX B

In the limit $g_- = 0$, Eq. (54) reads

$$G(x) = \frac{\omega_+}{\omega_+ + \omega_-} e^y \Phi[\omega_-, \omega_+ + \omega_- + 1; -y]$$
$$+ \frac{\omega_-}{\omega_+ + \omega_-} \Phi[\omega_+, \omega_+ + \omega_- + 1; y], \tag{B1}$$

where $y = e^{g_+(x-1)}$. Using the fact that [49]

$$e^y \Phi[\alpha, \beta; -y] = \Phi[\beta - \alpha, \beta; y], \tag{B2}$$

Eq. (B1) can be written

$$G(x) = \frac{\omega_+}{\omega_+ + \omega_-} \Phi[\omega_+ + 1, \omega_+ + \omega_- + 1; y]$$
$$+ \frac{\omega_-}{\omega_+ + \omega_-} \Phi[\omega_+, \omega_+ + \omega_- + 1; y], \tag{B3}$$

or noting Eq. (55) and the fact that $\Gamma(z+1) = z\Gamma(z)$ for any z,

$$G(x) = \sum_j \left(\frac{\omega_+}{\omega_+ + \omega_-} \frac{\Gamma(j+\omega_+ + 1)}{\Gamma(\omega_+ + 1)} + \frac{\omega_-}{\omega_+ + \omega_-} \frac{\Gamma(j+\omega_+)}{\Gamma(\omega_+)} \right)$$
$$\times \frac{\Gamma(\omega_+ + \omega_- + 1)}{\Gamma(j+\omega_+ + \omega_- + 1)} \frac{y^j}{j!} \tag{B4}$$

$$= \sum_j \left(\frac{\omega_+}{\omega_+ + \omega_-} \frac{(j+\omega_+)\Gamma(j+\omega_+)}{\omega_+ \Gamma(\omega_+)} \right.$$
$$\left. + \frac{\omega_-}{\omega_+ + \omega_-} \frac{\Gamma(j+\omega_+)}{\Gamma(\omega_+)} \right)$$
$$\times \frac{(\omega_+ + \omega_-)\Gamma(\omega_+ + \omega_-)}{(j+\omega_+ + \omega_-)\Gamma(j+\omega_+ + \omega_-)} \frac{y^j}{j!} \tag{B5}$$

$$= \sum_j \frac{\Gamma(j+\omega_+)}{\Gamma(\omega_+)} \frac{\Gamma(\omega_+ + \omega_-)}{\Gamma(j+\omega_+ + \omega_-)} \frac{y^j}{j!} \tag{B6}$$

$$= \Phi[\omega_+, \omega_+ + \omega_-; y], \tag{B7}$$

as in Eq. (56).

The marginal p_n is given by

$$\langle n|G\rangle = \frac{1}{n!} \partial_x^n [G(x)]_{x=0} \tag{B8}$$

[cf. Eq. (10)]. Using Eq. (56) and the derivative of the confluent hypergeometric function,

$$\partial_y^n \Phi[\alpha, \beta; y] = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \frac{\Gamma(\beta)}{\Gamma(n+\beta)} \Phi[\alpha + n, \beta + n; y], \tag{B9}$$

one obtains Eq. (57).

APPENDIX C

In this appendix, we fill in the details of the derivations of the equations of motion for the latter three of the four spectral bases discussed in Sec. II A.

1. $|j, k_j\rangle$ basis

Projecting the conjugate state $\langle j, k_j|$ onto Eq. (83) (in which dummy indices j and k are changed to j' and k', respectively) gives

$$0 = \sum_{j'k'} (j' + \rho k')\langle j|j'\rangle \langle k_j|k'_{j'}\rangle G_{j'k'} + \sum_{j'k'} \langle j|\hat{b}_n^+ \hat{\Gamma}_n|j'\rangle$$
$$\times \langle k_j|k'_{j'}\rangle G_{j'k'} + \rho \sum_{j'k'} \langle k_j|\hat{b}_m^+|k'_{j'}\rangle \langle j|\hat{\Delta}_n(j')|j'\rangle G_{j'k'}. \tag{C1}$$

From the orthonormality of states, the first term of Eq. (C1) simplifies to

$$\sum_{k'} (j + \rho k')\langle k_j|k'_j\rangle G_{jk'} = (j + \rho k)G_{jk}. \tag{C2}$$

Recalling Eq. (30), the product $\langle k_j|k'_{j'}\rangle$ simplifies to

$$\langle k_j | k'_{j'} \rangle = \frac{(-Q_{jj'})^{k-k'}}{(k-k')!} \theta(k-k'+1), \tag{C3}$$

with $Q_{jj'} = \bar{q}_j - \bar{q}_{j'}$, whereupon Eq. (C1), separating the part of its second term which is diagonal in k from that which is subdiagonal and applying Eq. (24) to its third term, becomes

$$0 = (j + \rho k)G_{jk} + \sum_{j'} \Gamma_{j-1,j'} G_{j'k}$$

$$+ \sum_{k'<k} \sum_{j'} \Gamma_{j-1,j'} \frac{(-Q_{jj'})^{k-k'}}{(k-k')!} G_{j'k'}$$

$$+ \rho \sum_{k'<k} \sum_{j'} \Delta_{jj'} \frac{(-Q_{jj'})^{k-k'-1}}{(k-k'-1)!} G_{j'k'}, \tag{C4}$$

with $\Gamma_{jj'}$ as in Eq. (76) and

$$\Delta_{jj'} = \langle j | \hat{\Delta}_n(j') | j' \rangle = \langle j | (\bar{q}_{j'} - \hat{q}_n) | j' \rangle \tag{C5}$$

$$= \langle j | (\bar{q}_j - \hat{q}_n) | j' \rangle \tag{C6}$$

$$= \sum_n \langle j | n \rangle (\bar{q}_j - q_n) \langle n | j' \rangle \tag{C7}$$

[where the orthonormality of $|j\rangle$ states is used in going from Eq. (C5) to Eq. (C6)]. Defining $\ell = k - k'$ and

$$V_{jj'}^\ell = \frac{(-Q_{jj'})^\ell}{\ell!}, \tag{C8}$$

Eq. (C4) can be written more compactly as Eq. (87).

2. $|n,k_n\rangle$ basis

Projecting the conjugate state $\langle n,k_n |$ onto Eq. (93) (in which dummy indices n and k are changed to n' and k', respectively) gives

$$0 = \sum_{n'k'} (g_{n'} + n' + \rho k') \langle n | n' \rangle \langle k_n | k'_{n'} \rangle G_{n'k'} - \sum_{n'k'} g_{n'} \langle n | n' + 1 \rangle$$

$$\times \langle k_n | k'_{n'} \rangle G_{n'k'} - \sum_{n'k'} n' \langle n | n' - 1 \rangle \langle k_n | k'_{n'} \rangle G_{n'k'}$$

$$+ \rho \sum_{n'k'} \Delta_{n'} \langle n | n' \rangle \langle (k-1)_n | k'_{n'} \rangle G_{n'k'}, \tag{C9}$$

where $\Delta_n = \bar{q}_n - q_n$. Noting that, as in Eq. (30),

$$\langle k_n | k'_{n\pm1} \rangle = \frac{(-Q_n^\pm)^{k-k'}}{(k-k')!} \theta(k-k'+1), \tag{C10}$$

where $Q_n^\pm = \bar{q}_n - \bar{q}_{n\pm1}$, Eq. (C9) becomes

$$0 = (g_n + n + \rho k)G_{nk} - g_{n-1} \sum_{k'\le k} \frac{(-Q_n^-)^{k-k'}}{(k-k')!} G_{n-1,k'}$$

$$- (n+1) \sum_{k'\le k} \frac{(-Q_n^+)^{k-k'}}{(k-k')!} G_{n+1,k'} + \rho \Delta_n G_{n,k-1}. \tag{C11}$$

Separating the parts of the second and third term that are diagonal in k and defining $\ell = k - k'$ and

$$V_{n\ell}^\pm = \frac{(-Q_n^\pm)^\ell}{\ell!}, \tag{C12}$$

Eq. (C11) becomes Eq. (97).

3. $|j,k_n\rangle$ basis

Substituting Eq. (103) into Eq. (58) at steady state gives

$$0 = \sum_{jk} G_{jk}\{g_{n-1}\langle n-1,m|j,k_{n-1}\rangle + (n+1)\langle n+1,m|j,k_{n+1}\rangle$$

$$- (g_n + n)\langle n,m|j,k_n\rangle + \rho[q_n\langle n,m-1|j,k_n\rangle$$

$$+ (m+1)\langle n,m+1|j,k_n\rangle - (q_n+m)\langle n,m|j,k_n\rangle]\}, \tag{C13}$$

or in terms of raising and lowering operators [cf. Eqs. (14)]

$$0 = \sum_{jk} G_{jk}\langle n,m|\{\hat{a}_n^+\hat{g}_n|j,k_{n-1}\rangle + \hat{a}_n^-|j,k_{n+1}\rangle - (\hat{g}_n + \hat{a}_n^+\hat{a}_n^-)|j,k_n\rangle$$

$$+ \rho[\hat{a}_m^+\hat{q}_n|j,k_n\rangle + \hat{a}_m^-|j,k_n\rangle - (\hat{q}_n + \hat{a}_m^+\hat{a}_m^-)|j,k_n\rangle]\}. \tag{C14}$$

Using the definitions in Eqs. (64)–(68), Eq. (C14) can be written as Eq. (105).

Using Eqs. (107)–(109), Eq. (105) can be written

$$0 = -\langle n,m|\sum_{j'k'} (j' + \rho k')|j',k'_n\rangle G_{j'k'}$$

$$- \langle n,m|\sum_{j'k'} \hat{b}_n^+\hat{\Gamma}_n|j',k'_n\rangle G_{j'k'} - \langle m|\rho \sum_{j'k'} (\bar{q}_n - q_n)$$

$$\times \langle n|j'\rangle \hat{b}_m^+|k'_n\rangle G_{j'k'} + \langle m|\sum_{j'k'} g_{n-1}\langle n-1|j'\rangle|\delta_- k'_n\rangle G_{j'k'}$$

$$+ \langle m|\sum_{j'k'} (n+1)\langle n+1|j'\rangle|\delta_+ k'_n\rangle G_{j'k'}, \tag{C15}$$

where

$$|\delta_\pm k_n\rangle \equiv |k_{n\pm1}\rangle - |k_n\rangle, \tag{C16}$$

and the dummy indices j and k have been changed to j' and k' respectively. Using Eq. (C10) to note that

$$\langle k_n | \delta_\pm k'_n \rangle = \frac{(-Q_n^\pm)^{k-k'}}{(k-k')!} \theta(k-k'), \tag{C17}$$

where $Q_n^\pm = \bar{q}_n - \bar{q}_{n\pm1}$, we multiply Eq. (C15) by $\langle k_n | m \rangle$ and sum over m to obtain

$$0 = -\langle n| \sum_{j'} (j' + \rho k)|j'\rangle G_{j'k} - \langle n| \sum_{j'} \hat{b}_n^+ \hat{\Gamma}_n |j'\rangle G_{j'k} - \rho \sum_{j'} (\bar{q}_n$$

$$- q_n)\langle n|j'\rangle G_{j',k-1} + \sum_{j'} g_{n-1}\langle n-1|j'\rangle \sum_{\ell=1}^{k} V_{n\ell}^- G_{j',k-\ell}$$

$$+ \sum_{j'} (n+1)\langle n+1|j'\rangle \sum_{\ell=1}^{k} V_{n\ell}^+ G_{j',k-\ell}, \tag{C18}$$

in which we exploit the completeness of $|m\rangle$ states, i.e., $\sum_m |m\rangle\langle m| = 1$, and $V_{n\ell}^{\pm}$ is as in Eq. (99). Multiplying Eq. (C18) by $\langle j|n\rangle$, summing over n, and exploiting $\sum_n |n\rangle\langle n| = 1$ for the first two terms, we obtain Eq. (110).

APPENDIX D

In this appendix we explicitly solve for G_{nk} in Eq. (117) using the tridiagonal matrix or Thomas [42] algorithm. We start by identifying the subdiagonal, diagonal, superdiagonal, and right-hand side elements of Eq. (117), respectively, as

$$A_n = g \quad (n = 1, \dots, N), \tag{D1}$$

$$B_n = -(\rho k + g + n) \quad (n = 0, \dots, N), \tag{D2}$$

$$C_n = n + 1 \quad (n = 0, \dots, N-1), \tag{D3}$$

$$R_n = -g\phi_k^- \delta_{nn_1} - n_1 \phi_k^+ \delta_{nn_0} \quad (n = 0, \dots, N), \tag{D4}$$

where N is the cutoff in protein count n and $n_1 = n_0 + 1$. Auxiliary variables are defined iteratively as

$$C_0' = \frac{C_0}{B_0}, \tag{D5}$$

$$C_n' = \frac{C_n}{B_n - C_{n-1}' A_n} \quad (n = 1, \dots, N-1), \tag{D6}$$

$$R_0' = \frac{R_0}{B_0}, \tag{D7}$$

$$R_n' = \frac{R_n - R_{n-1}' A_n}{B_n - C_{n-1}' A_n} \quad (n = 1, \dots, N), \tag{D8}$$

and the solution is obtained by backward iteration with

$$G_N^k = R_N', \tag{D9}$$

$$G_{n-1}^k = R_{n-1}' - C_{n-1}' G_n^k \quad (n = N, \dots, 1) \tag{D10}$$

(where k has been moved from subscript to superscript for ease of reading).

Computing the first few terms of Eq. (D6) reveals the pattern

$$C_n' = -(n+1)\frac{\eta_n^k}{\eta_{n+1}^k}, \tag{D11}$$

where

$$\eta_n^k = \sum_{i=0}^{n} \frac{n!}{i!\,(n-i)!} g^{n-i} \prod_{\ell=0}^{i-1} (\rho k + \ell), \tag{D12}$$

with the convention that $\Pi_a^b[\cdot] = 1$ if $a > b$. Note that since $\Pi_{\ell=0}^{i-1}(\rho k + \ell) = \Gamma(\rho k + i)/\Gamma(\rho k)$, we may also use the integral representation of the Gamma function to write

$$\eta_n^k = \frac{1}{\Gamma(\rho k)} \sum_{i=0}^{n} \frac{n!}{i!\,(n-i)!} g^{n-i} \int_0^t dt\, e^{-t} t^{(\rho k+i-1)} \tag{D13}$$

$$= \frac{1}{\Gamma(\rho k)} \int_0^t dt\, e^{-t} t^{(\rho k-1)} \sum_{i=0}^{n} \frac{n!}{i!\,(n-i)!} g^{n-i} t^i \tag{D14}$$

$$= \frac{1}{\Gamma(\rho k)} \int_0^t dt\, e^{-t} t^{(\rho k-1)} (g+t)^n. \tag{D15}$$

Using Eq. (D8) it is immediately clear that

$$R_{n<n_0}' = 0. \tag{D16}$$

The first nonzero term is

$$R_{n_0}' = \frac{n_1 \phi_k^+}{g n_0} \frac{\eta_{n_0}^k}{\eta_{n_0-1}^k} \frac{1}{\epsilon_{n_0}^k - 1}, \tag{D17}$$

where we have defined

$$\epsilon_n^k = \frac{\rho k + g + n}{g n} \frac{\eta_n^k}{\eta_{n-1}^k}. \tag{D18}$$

Further iteration of Eq. (D8) makes clear that

$$R_{n>n_0}' = \frac{n_1}{g} f_k \prod_{i=n_0}^{n} \left(\frac{1}{i} \frac{\eta_i^k}{\eta_{i-1}^k} \frac{1}{\epsilon_i^k - 1} \right) \tag{D19}$$

$$= \frac{n_1}{g} \frac{(n_0-1)!}{n!} \frac{\eta_n^k}{\eta_{n_0-1}^k} f_k \prod_{i=n_0}^{n} \frac{1}{\epsilon_i^k - 1}, \tag{D20}$$

where

$$f_k = \phi_k^+ - \frac{g n_0}{n_1} \frac{\eta_{n_0-1}^k}{\eta_{n_0}^k} (\epsilon_{n_0}^k - 1)\phi_k^-. \tag{D21}$$

Computing the first few terms of Eq. (D10) reveals that

$$G_{n>n_0}^k = (\epsilon_n^k - 1)R_{n>n_0}' F_n^k \tag{D22}$$

$$= \frac{n_1}{g} \frac{(n_0-1)!}{n!} \frac{\eta_n^k}{\eta_{n_0-1}^k} f_k F_n^k \prod_{i=n_0}^{n-1} \frac{1}{\epsilon_i^k - 1}, \tag{D23}$$

where

$$F_n^k = \sum_{i=0}^{N-n} \prod_{\ell=n}^{n+i} \frac{1}{\epsilon_\ell^k - 1}. \tag{D24}$$

At the threshold Eq. (D10) gives

$$G_{n_0}^k = \frac{n_1}{g n_0} \frac{\eta_{n_0}^k}{\eta_{n_0-1}^k} \frac{1}{\epsilon_{n_0}^k - 1} (\phi_k^+ + f_k F_{n_1}^k) \qquad (D25)$$

and since $R'_{n<n_0} = 0$, the solution is easily completed using Eq. (D10), giving

$$G_{n<n_0} = G_{n_0} \prod_{i=1}^{n_0-n} (-C'_{n_0-i}) \qquad (D26)$$

$$= \frac{n_1}{g} \frac{(n_0-1)!}{n!} \frac{\eta_n^k}{\eta_{n_0-1}^k} \frac{\phi_k^+ + f_k F_{n_1}^k}{\epsilon_{n_0}^k - 1}. \qquad (D27)$$

These results are summarized in Eqs. (120)–(125).

[1] S. Hooshangi, S. Thiberge, and R. Weiss, Proc. Natl. Acad. Sci. U.S.A. **102**, 3581 (2005).

[2] M. Thattai and A. van Oudenaarden, Biophys. J. **82**, 2943 (2002).

[3] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science **297**, 1183 (2002).

[4] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, Nat. Genet. **31**, 69 (2002).

[5] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Proc. Natl. Acad. Sci. U.S.A. **99**, 12795 (2002).

[6] J. Paulsson, O. G. Berg, and M. Ehrenberg, Proc. Natl. Acad. Sci. U.S.A. **97**, 7148 (2000).

[7] J. M. Pedraza and A. van Oudenaarden, Science **307**, 1965 (2005).

[8] J. Paulsson, Nature (London) **427**, 415 (2004).

[9] G. Tkačik, C. G. Callan, and W. Bialek, Phys. Rev. E **78**, 011910 (2008).

[10] G. Tkacik, C. G. Callan, Jr.,, and W. Bialek, Proc. Natl. Acad. Sci. U.S.A. **105**, 12265 (2008).

[11] E. Ziv, I. Nemenman, and C. H. Wiggins, PLoS ONE **2**, e1077 (2007).

[12] A. Mugler, E. Ziv, I. Nemenman, and C. H. Wiggins, IET Sys. Bio. **3**, 379 (2009).

[13] E. Emberly, Phys. Rev. E **77**, 041903 (2008).

[14] F. Tostevin and P. R. ten Wolde, Phys. Rev. Lett. **102**, 218101 (2009).

[15] J. Elf, G.-W. Li, and X. S. Xie, Science **316**, 1191 (2007).

[16] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, Science **322**, 442 (2008).

[17] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, PLoS Biol. **4**, e309 (2006).

[18] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Cell **123**, 1025 (2005).

[19] S. Tănase-Nicola, P. B. Warren, and P. Rein ten Wolde, Phys. Rev. Lett. **97**, 068102 (2006).

[20] J. S. van Zon, M. J. Morelli, S. Tănase-Nicola, and P. R. ten Wolde, Biophys. J. **91**, 4350 (2006).

[21] J. S. van Zon and P. R. ten Wolde, J. Chem. Phys. **123**, 234910 (2005).

[22] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[23] A. M. Walczak, M. Sasai, and P. G. Wolynes, Biophys. J. **88**, 828 (2005).

[24] Y. Lan and G. A. Papoian, J. Chem. Phys. **125**, 154901 (2006).

[25] Y. Lan, P. G. Wolynes, and G. A. Papoian, J. Chem. Phys. **125**, 124106 (2006).

[26] Y. Lan and G. A. Papoian, Phys. Rev. Lett. **98**, 228301 (2007).

[27] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).

[28] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

[29] R. E. Bellman, *Adaptive Control Processes* (Princeton University Press, Princeton, NJ, 1961).

[30] M. J. Morelli, P. R. ten Wolde, and R. J. Allen, Proc. Natl. Acad. Sci. U.S.A. **106**, 8101 (2009).

[31] A. M. Walczak, A. Mugler, and C. H. Wiggins, Proc. Natl. Acad. Sci. U.S.A. **106**, 6529 (2009).

[32] A master equation in which the coordinate appears explicitly in the rates [e.g. g_n or q_n in Eq. (58)] is sometimes mistakenly termed "nonlinear" in the literature, perhaps discouraging calculations which exploit its inherent linear algebraic structure. We remind the reader that the master equation is perfectly linear.

[33] S. Klumpp and T. Hwa, Proc. Natl. Acad. Sci. U.S.A. **105**, 20245 (2008).

[34] S. Iyer-Biswas, F. Hayot, and C. Jayaprakash, Phys. Rev. E **79**, 031911 (2009).

[35] Note that setting $x = e^{ik}$ makes clear that the generating function is simply the Fourier transform.

[36] Note here, however, the difference with respect to the normalization convention commonly used in quantum mechanics in the prefactors of the creation and annihilation operations.

[37] M. Doi, J. Phys. A **9**, 1465 (1976).

[38] Ya. B. Zel'dovich and A. A. Ovchinnikov, JETP **47**, 829 (1978).

[39] L. Peliti, J. Phys. A **19**, L365 (1986).

[40] D. C. Mattis and M. L. Glasser, Rev. Mod. Phys. **70**, 979 (1998).

[41] J. Lin, IEEE Trans. Inf. Theory **37**, 145 (1991).

[42] L. Thomas, *Watson Sci* (Columbia University, New York, 1949).

[43] Because this basis is a function of three coordinates (j, k, and n), we cannot abstractly define the generating function as in Eq. (61); we must instead work directly with the joint probability distribution.

[44] B. D. Gomperts, I. M. Kramer, and P. E. R. Tantham, *Signal Transduction* (Academic Press, San Diego, 2002).

[45] A. Y. Ting and D. Endy, Science **298**, 1189 (2002).

[46] P. B. Detwiler, S. Ramanathan, A. Sengupta, and B. I. Shraiman, Biophys. J. **79**, 2801 (2000).

[47] H. Bolouri and E. H. Davidson, Proc. Natl. Acad. Sci. U.S.A. **100**, 9371 (2003).

[48] B. L. Bassler, Curr. Opin. Microbiol. **2**, 582 (1999).

[49] W. Koepf, *Hypergeometric Summation: An Algorithmic Approach to Summation and Special Function Identities* (Vieweg, Braunschweig, Germany, 1998).