# Analytic methods for modeling stochastic regulatory networks

Aleksandra M. Walczak,[1, *] Andrew Mugler,[2, †] and Chris H. Wiggins[3]

[1]*Princeton Center for Theoretical Science, Princeton University, Princeton, NJ 08544*
[2]*Department of Physics, Columbia University, New York, NY 10027*
[3]*Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027*
(Dated: March 8, 2011)

Recent single-cell experiments have revived interest in the unavoidable or *intrinsic* noise in bio-chemical and genetic networks arising from the small number of molecules of the participating species. That is, rather than modeling regulatory networks in terms of the deterministic dynamics of concentrations, we model the dynamics of the probability of a given copy number of the reactants in single cells. Most of the modeling activity of the last decade has centered on stochastic simulation, i.e., Monte-Carlo methods for generating stochastic time series. Here we review the mathematical description in terms of probability distributions, introducing the relevant derivations and illustrating several cases for which analytic progress can be made either instead of or before turning to numerical computation. Analytic progress can be useful both for suggesting more efficient numerical methods and for obviating the computational expense of, for example, exploring parametric dependence.

## Contents

---

*Electronic address: awalczak@princeton.edu
†Electronic address: ajm2121@columbia.edu

## I. INTRODUCTION

Theoretical descriptions of biochemical kinetic processes rely on considering the concentration changes of the molecules involved in the reactions, and specifying the types of regulatory interactions between them. The theoretical tools used to describe many types of biochemical reactions share many similarities. In this chapter we give an overview of the analytical approaches to study biochemical kinetics using the example of small gene regulatory networks.

The regulation of genes by transcription factor proteins is an intrinsically stochastic process, owing to the small numbers of copies of molecules involved. With the development of imaging techniques in molecular biology, we are able to observe directly the fluctuations in the concentrations of proteins and mRNAs and, by measuring the intensity profiles of fluorescence markers, measure full probability distributions [1–3]. Experiments over the last decade have shown that in fact gene regulation is a noisy process, and noise can propagate in gene networks [4].

Many methods for solving the resulting stochastic equations rely on computer simulations. The efficiency of these methods has been greatly advanced in the last several years [5–16] . However numerical simulations are naturally limited to a specific choice of parameters, and changing the parameters requires a completely new calculation. Furthermore simulation suffers from the curse of dimensionality: the computational runtime grows prohibitively as the number of species increases. These problems can be bypassed by developing analytical approaches, which often require certain approximations.

This chapter is intended as a tutorial on theoretical descriptions of biochemical kinetics. For clarity, we start by introducing the simplified kinetic description of the production of proteins. We first consider the deterministic description of the system and then introduce noise. After familiarizing the reader with different levels of description, we discuss models of regulation. We also present a wide spectrum of analytic tools used to find the steady state probability distributions of protein copy numbers in the cell. We point out that while for concreteness we focus on the case of gene regulation, the methods presented in this chapter are very general.

The results presented in this review are not new; some can be found in textbooks [17–19], while others have been derived more recently in the context of gene regulation [3, 20–38]. The goal of this review is to give the reader who is not familiar with analytical methods for modeling stochastic gene regulation an overview of the mathematical tools used in the field. Naturally, we are not able to cover all the developments in the field, but we hope to give the reader a useful starting point. For concreteness we also limit our discussion to the case of small gene networks and do not discuss approximations used to describe larger networks, which is currently an active area of research [10, 11, 15, 16, 28, 29].

## II.   MODEL FOR PROTEIN PRODUCTION AND DEGRADATION: SIMPLE BIRTH–DEATH PROCESS

In this section we describe a simple birth death process for one species with constant production rate, i.e., absent of regulation. We first remind the reader of the deterministic description given by chemical kinetics. Next we introduce the full probabilistic description (given by the master equation obeyed by the probability distribution $p_n(t)$), and, restricting our discussion to the simplest case, show how the deterministic equations arise as the dynamics of the mean $\langle n \rangle \equiv \sum_n p_n$. After calculating the variance, we introduce the generating function formalism, useful in solving the master equation. This formalism may also be described using raising and lowering operators familiar to physicists from quantum mechanics. Finally we relate this description to the Fokker-Plank equation, the analogous master equation for continuous state variables (e.g., real-valued coordinates rather than the integer-valued copy number). These results will be used in later sections when we introduce autoregulation and regulation among different species.

### A.   Deterministic description: the kinetic rate equation

In the simplest case, the number of copies $n$ of a protein species $X$ can change either due to the production of a protein, which occurs at a constant rate $\tilde{g}$, or due to degradation of a protein, which occurs at a constant rate $r$:

$$\emptyset \underset{r}{\overset{\tilde{g}}{\rightleftharpoons}} X. \tag{1}$$

Here we condense the complicated molecular machinery of transcription, translation, and protein modification into a single constant production rate. Similarly we do not specify the molecular processes which have led to degradation; for simplicity we imagine either dilution due to cell division or active degradation with a constant rate.

When $n$ is large, the dynamics for the mean $\langle n \rangle$ are well approximated by the continuous, deterministic description provided by the kinetic rate equation for the concentration $c \equiv \langle n \rangle / V$ in a volume $V$:

$$\frac{dc}{dt} = \frac{\tilde{g}}{V} - rc. \tag{2}$$

The solution of Eqn. 2 is

$$c(t) = \frac{\tilde{g}}{Vr} + e^{-rt} \left[ c(0) - \frac{\tilde{g}}{Vr} \right], \tag{3}$$

where $c(0)$ is the concentration of proteins at initial time $t = 0$. In steady state the mean number of proteins is simply the ratio of the production and degradation rates, $\langle n \rangle = cV = \tilde{g}/r$, which is easily seen by taking either $dc/dt = 0$ in Eqn. 2 or $t \to \infty$ in Eqn. 3.

### B.   Introducing noise: the master equation

The general probabilistic description of chemical reactions, respecting the finite copy number of the reactants, is the master equation, which specifies the rate of change of $p_{\mathbf{n}}$, the probability that there are $\mathbf{n} = (n_1, n_2, \ldots, n_L)$ copies of the $L$ reactants. The macroscopic, deterministic description in terms of chemical kinetics is recovered by considering the dynamics for the mean concentration of these reactants, i.e. the vector $\langle \mathbf{n} \rangle / V = (\langle n_1 \rangle / V, \langle n_2 \rangle / V, \ldots \langle n_L \rangle / V) = (c_1, c_2, \ldots c_L)$. Such a description in terms of a summary statistic necessarily ignores a great amount of information, including all information about fluctuations about these mean values.

Given the transition matrices $w_{nn'}$, with dimensions of rates, the general form of the master equation (c.f. Appendix A), set by conservation of total probability, is

$$\dot{p}_{\mathbf{n}} = \sum_{\mathbf{n}'} \left[ w_{\mathbf{n}\mathbf{n}'}(t) p_{\mathbf{n}'} - w_{\mathbf{n}'\mathbf{n}}(t) p_{\mathbf{n}} \right], \tag{4}$$

For the case of only one species, we need specify only the dynamics of $\mathbf{n} = n$. Restricting further to the case of a simple birth-death process, there are only two nonzero contributions from the transition matrices $w_{nn'}$, given by

$$w_{n+1,n}(t) = \tilde{g}, \tag{5}$$
$$w_{n-1,n}(t) = rn, \tag{6}$$

all other values of $w_{nn'}$ being 0. Under this restriction Eqn. 4 reduces to the familiar

$$\dot{p}_n = -\tilde{g}p_n - rnp_n + \tilde{g}p_{n-1} + r(n+1)p_{n+1}. \tag{7}$$

Qualitatively, the four terms on the right hand side represent how the probability of having $n$ proteins can either (i) decrease in time, if there are $n$ initially and one is either produced (first term) or degraded (second term), or (ii) increase in time, if there are either $n-1$ initially and one is produced (third term) or $n+1$ initially and one is degraded (fourth term).

The dynamics of the mean number of proteins $\langle n \rangle = \sum_{n=0}^{\infty} np_n$ are readily obtained (see Appendix B), as

$$\frac{d\langle n \rangle}{dt} = \tilde{g} - r\langle n \rangle. \tag{8}$$

Eqn. 8 shows that the dynamics of the mean of the protein distribution reproduces the kinetic rate equation, Eqn. 2.

From here on, we will re-scale time $t$ by the degradation rate $r$, such that $rt \to t$, and define $g = \tilde{g}/r$, making the master equation

$$\dot{p}_n = -gp_n - np_n + gp_{n-1} + (n+1)p_{n+1}. \tag{9}$$

### 1.  Steady state solution

The master equation can be rewritten in terms of shift operators $E^+$ and $E^-$ which increase and decrease the number of proteins by one, respectively [17], i.e.

$$E^+ f_n = f_{n+1}, \tag{10}$$
$$E^- f_n = f_{n-1}, \tag{11}$$

for any function $f_n$. We begin by writing Eqn. 9 in terms of only $E^+$ to make clear that, as is the case for any one-dimensional master equation, its steady state can be found iteratively. In terms of $E^+$,

$$\dot{p}_n = (E^+ - 1)(np_n - gp_{n-1}). \tag{12}$$

Setting $\dot{p}_n = 0$ for steady state requires that the second term in parentheses vanish, giving for $n > 0$ the recursive relation

$$p_n = p_{n-1}\frac{g}{n}. \tag{13}$$

Starting with $n = 1$ and computing the first few terms reveals the pattern

$$p_n = \frac{g^n}{n!}p_0 \tag{14}$$

where $p_0 = e^{-g}$ is set by normalization:

$$1 = \sum_{n=0}^{\infty} p_n = p_0 \sum_{n=0}^{\infty} \frac{g^n}{n!} = p_0 e^g. \tag{15}$$

Thus the steady-state probability of having $n$ proteins is the Poisson distribution,

$$p_n = \frac{g^n}{n!}e^{-g}, \tag{16}$$

with parameter $g$, the ratio of production to degradation rates.

We remind the reader that the variance $\sigma^2 = \langle n^2 \rangle - \langle n \rangle^2$ of the Poisson distribution is equal to its mean $\langle n \rangle$ (Appendix C). Therefore the standard deviation $\sigma$ over the mean falls off like

$$\frac{\sigma}{\langle n \rangle} = \frac{\sqrt{\langle n \rangle}}{\langle n \rangle} = \frac{1}{\sqrt{\langle n \rangle}}, \tag{17}$$

demonstrating that the relative effect of fluctuations diminishes for large protein number. In Appendix D we show that in the limit of large protein number the steady state asymptotes to a Gaussian distribution with mean and variance equal to $g$.

For comparison with the other representations of the master equation described below, we now write Eqn. 9 in terms of both shift operators $E^+$ and $E^-$:

$$\dot{p}_n = (E^+ - 1)(n - gE^-)p_n. \tag{18}$$

Eqn. 12 can be rewritten slightly by inserting between the parenthetic terms the unit operator $1 = E^- E^+$ and distributing the $E^-$ to the left and the $E^+$ to the right, giving

$$\dot{p}_n = -(E^- - 1)[(n+1)E^+ - g]p_n. \tag{19}$$

where the negative sign has been factored out of the first parenthetic term.

## 2. Generating function representation

Not all master equations are solvable by straightforward iteration. A more generalizable way to solve a master equation is by the introduction of a generating function [17]. Here we demonstrate the generating function approach on the birth–death master equation, Eqn. 9. The utility of the generating function for solving a master equation is that it turns an infinite set of ordinary differential equations (e.g. Eqn. 9) into a single partial differential equation.

The generating function $G$ is defined as

$$G(x,t) \equiv \sum_{n=0}^{\infty} p_n(t)x^n, \tag{20}$$

a power-series expansion in a continuous variable $x$ whose coefficients are the probabilities $p_n$. Since $x$ is a variable we introduce, we note that defining $x = e^{i\varphi}$ makes clear that the generating function is equivalent to the Fourier transform of $p_n$ in protein number (this point is further developed in Appendix E). The probability distribution may be recovered via the inverse transform

$$p_n(t) = \frac{1}{n!}\partial_x^n G(x,t)|_{x=0}. \tag{21}$$

Additionally we note that the $\ell$th moment may be generated by

$$\langle n^\ell \rangle = \left[(x\partial_x)^\ell G\right]_{x=1}, \tag{22}$$

(which is the reason for the generating function's name).

To obtain an easily-solvable partial differential equation for the generating function, we multiply Eqn. 9 by $x^n$ and sum over $n$ to obtain (see Appendix B for the detailed derivation):

$$\dot{G} = -(x-1)(\partial_x - g)G. \tag{23}$$

We see immediately from comparison of Eqns. 19 and 23 that the representations of operators $E^-$ and $(n+1)E^+$ in $x$ space are $x$ and $\partial_x$, respectively.

In steady state ($\dot{G} = 0$) Eqn. 23 must satisfy $(\partial_x - g)G = 0$, which is solved by

$$G(x) = G(0)e^{gx}, \tag{24}$$

where $G(0) = e^{-g}$ is set by normalization:

$$G(0)e^g = G(1) = \sum_{n=0}^{\infty} p_n(1)^n = \sum_{n=0}^{\infty} p_n = 1. \tag{25}$$

The steady state distribution is recovered via inverse transform:

$$p_n = \frac{1}{n!}\partial_x^n \left[e^{g(x-1)}\right]_{x=0} = e^{-g}\frac{g^n}{n!}, \tag{26}$$

a Poisson distribution with parameter $g$, as before (Eqn. 16).

The time-dependent solution to Eqn. 23 may be obtained using the method of characteristics, in which one looks for characteristic curves along which the partial differential equation becomes an ordinary differential equation. The curves are parameterized by $s$, such that $G(x,t) = G(x(s), t(s)) = G(s)$, and we look for the solution to the ordinary differential equation for $G(s)$, having eliminated $x$ and $t$ from the problem. First Eqn. 23 can be rewritten more explicitly as

$$(x-1)gG = \frac{\partial G}{\partial t} + \frac{\partial G}{\partial x}(x-1). \tag{27}$$

Using the chain rule on $G(s)$ gives

$$\frac{dG}{ds} = \frac{\partial G}{\partial t}\frac{dt}{ds} + \frac{\partial G}{\partial x}\frac{dx}{ds}. \tag{28}$$

Consistency of Eqns. 27 and 28 requires

$$\frac{dt}{ds} = 1, \tag{29}$$

$$\frac{dx}{ds} = x - 1, \tag{30}$$

$$\frac{dG}{ds} = (x-1)gG. \tag{31}$$

Eqn. 29 implies $s = t$ (we set $t_0 = 0$ without loss of generality). Therefore, defining $y \equiv x - 1$, Eqn. 30 implies

$$y = y_0 e^t. \tag{32}$$

Finally, straightforward integration of Eqn. 31 yields

$$G = G_0 \exp\left[gy_0 \int_0^t dt' \, e^{t'}\right] = G_0 \exp\left[gy_0(e^t - 1)\right] = F(y_0)\exp\left[gy_0 e^t\right]. \tag{33}$$

where we define $F(y_0) \equiv G_0 e^{-gy_0}$. We may expand $F(y_0)$ as a function of its argument, i.e. $F(y_0) = \sum_{j=0}^{\infty} A_j y_0^j$ for some coefficients $A_j$, such that

$$G = \sum_j A_j y_0^j \exp\left[gy_0 e^t\right]. \tag{34}$$

Inserting $y_0 = ye^{-t} = (x-1)e^{-t}$ (Eqn. 32), we obtain

$$G(x,t) = \sum_j A_j e^{-jt}(x-1)^j e^{g(x-1)}. \tag{35}$$

For $t \to \infty$ only the $j = 0$ term survives and we recover the steady state generating function (Eqn. 24),

$$G(x) = c_0 e^{g(x-1)} \tag{36}$$

where $c_0 = 1$ by normalization. Eqn. 35 constitutes the full time-dependent solution to the birth death process; the time-dependent distribution can be retrieved using the inverse transform, Eqn. 21, and the coefficients $A_j$ are computable from the initial distribution $p_n(0)$, which will be made explicit in the next section.

Since Eqn. 23 is linear in $G$, one may arrive at its solution in a second, elegant way by expanding in the eigenfunctions of the $x$-dependent operator. That is, writing Eqn. 23 as $\dot{G} = -\mathcal{L}G$, where

$$\mathcal{L} = (x-1)(\partial_x - g), \tag{37}$$

we expand $G$ in eigenfunctions $\phi_j(x)$ with time-dependent expansion coefficients $G_j(t)$,

$$G(x,t) = \sum_j G_j(t)\phi_j(x), \tag{38}$$

where the $\phi_j(x)$ satisfy

$$\mathcal{L}\phi_j = \lambda_j \phi_j \tag{39}$$

for eigenvalues $\lambda_j$. Substituting Eqn. 38 into Eqn. 23 gives $\sum_j \dot{G}_j \phi_j = -\sum_j \lambda_j G_j \phi_j$, which, by orthogonality of the $\phi_j$, yields the set of ordinary differential equations $\dot{G}_j = -\lambda_j G_j$, solved by

$$G_j(t) = A_j e^{-\lambda_j t} \tag{40}$$

for some constants $A_j$. Comparing the result,

$$G(x,t) = \sum_j A_j e^{-\lambda_j t} \phi_j(x), \tag{41}$$

with Eqn. 35 reveals the forms of the eigenvalues

$$\lambda_j = j \in \{0, 1, 2, \dots\} \tag{42}$$

and the eigenfunctions

$$\phi_j(x) = (x-1)^j e^{g(x-1)}, \tag{43}$$

facts that we will confirm in the next section using operator methods.

### 3. Operator representation

We now introduce a representation of the master equation in terms of raising and lowering operators. This representation makes the solution more elegant, yielding a simple algebra which allows calculation of projections between the spaces of protein numbers and eigenfunctions without explicit computation of overlap integrals; it also lays the formal groundwork for solving models of multidimensional regulatory networks. The formalism was first used for diffusion by Doi [39] and Zel'dovich [40] and later developed by Peliti [41].

As before the generating function is defined as an expansion in a complete basis indexed by protein number $n$ in which the expansion coefficients are the probabilities $p_n$. Within the operator formalism, this basis is represented as the set of states $|n\rangle$, i.e.

$$|G(t)\rangle = \sum_{n=0}^{\infty} p_n(t)|n\rangle. \tag{44}$$

Here we have adopted state notation commonly used in quantum mechanics [42]; the previous representation (Eqn. 20) is recovered by projecting onto this equation the conjugate states $\langle x|$ with the provisions

$$\langle x|G(t)\rangle \equiv G(x,t), \tag{45}$$
$$\langle x|n\rangle \equiv x^n. \tag{46}$$

In Appendix E we show how the orthonormality of the $|n\rangle$ states,

$$\langle n|n'\rangle = \delta_{nn'}, \tag{47}$$

dictates the form of the conjugate state:

$$\langle n|x\rangle = \frac{1}{x^{n+1}}. \tag{48}$$

Projecting Eqn. 44 onto $\langle n|$ and using Eqn. 47 gives the inverse transform:

$$\langle n|G(t)\rangle = p_n(t). \tag{49}$$

The equation of motion in the operator representation is obtained by summing the master equation (Eqn. 9) over $n$ against $|n\rangle$ (see Appendix B), giving

$$|\dot{G}\rangle = -\hat{\mathcal{L}}|G\rangle, \tag{50}$$

where

$$\hat{\mathcal{L}} \equiv (\hat{a}^+ - 1)(\hat{a}^- - g). \tag{51}$$

Just as in the operator treatment of the quantum harmonic oscillator [42], the operators $\hat{a}^+$ and $\hat{a}^-$ here raise and lower the protein number by 1 respectively, i.e.

$$\hat{a}^+|n\rangle = |n+1\rangle, \tag{52}$$
$$\hat{a}^-|n\rangle = n|n-1\rangle \tag{53}$$

(note, however, that the prefactors here, 1 and $n$, are different than those conventionally used for the harmonic oscillator, $\sqrt{n+1}$ and $\sqrt{n}$, respectively). Comparison of Eqns. 19, 23, and 51 makes clear the following correspondences among the master equation, generating function, and operator representations respectively:

$$E^- \leftrightarrow x \leftrightarrow \hat{a}^+, \tag{54}$$
$$(n+1)E^+ \leftrightarrow \partial_x \leftrightarrow \hat{a}^-. \tag{55}$$

While it might seem strange that the down-shift operator $E^-$ corresponds to the raising operator $\hat{a}^+$ (and the up-shift operator to the lowering operator), in Appendix F we show that, just as with the quantum harmonic oscillator, $\hat{a}^+$ and $\hat{a}^-$ lower and raise protein number respectively when operating to the left, i.e.

$$\langle n|\hat{a}^+ = \langle n-1|, \tag{56}$$
$$\langle n|\hat{a}^- = (n+1)\langle n+1|, \tag{57}$$

which makes the correspondence more directly apparent. Finally, again as with the quantum harmonic oscillator, the raising and lowering operators enjoy the commutation relation (see Appendix F)

$$\left[\hat{a}^-, \hat{a}^+\right] = 1, \tag{58}$$

and $\hat{a}^+\hat{a}^-$ acts as a number operator, i.e. $\hat{a}^+\hat{a}^-|n\rangle = n|n\rangle$.

Eqn. 51 shows that the full operator $\hat{\mathcal{L}}$ factorizes, suggesting the definition of the shifted operators

$$\hat{b}^+ \equiv \hat{a}^+ - 1, \tag{59}$$
$$\hat{b}^- \equiv \hat{a}^- - g. \tag{60}$$

Since $\hat{b}^+$ and $\hat{b}^-$ differ from $\hat{a}^+$ and $\hat{a}^-$, respectively, by scalars, they obey the same commutation relation, i.e.

$$\left[\hat{b}^-, \hat{b}^+\right] = 1. \tag{61}$$

Since the equation of motion (Eqn. 50) is linear, it will benefit from expansion of $|G\rangle$ in the eigenfunctions $|\lambda_j\rangle$ of the operator $\hat{\mathcal{L}}$, where

$$\hat{\mathcal{L}}|\lambda_j\rangle = \lambda_j|\lambda_j\rangle \tag{62}$$

for eigenvalues $\lambda_j$. In Appendix G we show that the commutation relation (Eqn. 61) and the steady state solution $\langle x|G\rangle = e^{g(x-1)}$, for which $\hat{\mathcal{L}}|G\rangle = 0$, completely define the eigenfunctions and eigenvalues of $\hat{\mathcal{L}}$; we summarize the results of Appendix G here. The eigenvalues are nonnegative integers $j$,

$$\lambda_j = j \in \{0, 1, 2, \dots\}, \tag{63}$$

as in Eqn. 42. The eigenvalue equation now reads

$$\hat{\mathcal{L}}|j\rangle = \hat{b}^+\hat{b}^-|j\rangle = j|j\rangle, \tag{64}$$

which is consistent with interpretation of $\hat{b}^+\hat{b}^-$ as a number operator for the eigenstates $|j\rangle$. The eigenfunctions are (in $x$ space)

$$\langle x|j\rangle = (x-1)^j e^{g(x-1)}, \tag{65}$$

as in Eqn. 43 with $\phi_j(x) \equiv \langle x|j \rangle$. The conjugate eigenfunctions are

$$\langle j|x \rangle = \frac{e^{-g(x-1)}}{(x-1)^{j+1}}. \tag{66}$$

The operators $\hat{b}^+$ and $\hat{b}^-$ raise and lower the eigenstates $|j\rangle$, respectively, as $\hat{a}^+$ and $\hat{a}^-$ do the $|n\rangle$ states, i.e.

$$\begin{align}
\hat{b}^+|j\rangle &= |j+1\rangle, \tag{67} \\
\hat{b}^-|j\rangle &= j|j-1\rangle, \tag{68} \\
\langle j|\hat{b}^+ &= \langle j-1|, \tag{69} \\
\langle j|\hat{b}^- &= (j+1)\langle j+1|. \tag{70}
\end{align}$$

Now employing the expansion of $|G\rangle$ in the eigenstates $|j\rangle$,

$$|G(t)\rangle = \sum_j G_j(t)|j\rangle, \tag{71}$$

the equation of motion (Eqn. 50) gives a trivial equation for the expansion coefficients

$$\dot{G}_j = -jG_j, \tag{72}$$

which is solved by

$$G_j(t) = A_j e^{-jt} \tag{73}$$

for some constants $A_j$, as in Eqn. 40. The inverse of Eqn. 71 (obtained by projecting $\langle j|$ and using $\langle j|j'\rangle = \delta_{jj'}$) is

$$\langle j|G(t)\rangle = G_j(t). \tag{74}$$

The probability distribution is retrieved by inverse transform (Eqn. 49),

$$p_n(t) = \langle n|G(t)\rangle = \langle n|\sum_j G_j(t)|j\rangle = \sum_j A_j e^{-jt}\langle n|j\rangle, \tag{75}$$

where the coefficients $A_j$ are computed from the initial distribution $p_n(0)$:

$$A_j = G_j(0) = \langle j|G(0)\rangle = \sum_n p_n(0)\langle j|n\rangle. \tag{76}$$

Eqns. 75 and 76 give the full time-dependent solution to the birth–death process as expansions in the eigenmodes $\langle n|j\rangle$ and conjugate eigenmodes $\langle j|n\rangle$. Because we have decomposed the problem using the eigenbasis, or spectrum, of the underlying operator, we will refer to this as the spectral solution.

The quantities $\langle n|j\rangle$ and $\langle j|n\rangle$ are overlaps between the protein number basis $|n\rangle$ and the eigenbasis $|j\rangle$. Both $\langle n|j\rangle$ and $\langle j|n\rangle$ are readily computed by contour integration or more efficiently by recursive updating; these techniques are presented in Appendix H. Notable special cases are

$$\begin{align}
\langle n|0\rangle &= e^{-g}\frac{g^n}{n!}, \tag{77} \\
\langle 0|n\rangle &= 1, \tag{78}
\end{align}$$

which confirm that Eqn. 75 describes a Poisson distribution in the steady state ($j = 0$).

## C. Fokker-Planck approximation

The previous sections discuss several methods for calculating the steady state and time dependent solutions of the master equation for the birth–death process. For many larger systems full solution of the master equation is not possible. In some of these cases one can make progress by deriving an approximate equation which is valid when protein numbers are large. In the limit of large protein numbers the master equation can be expanded to second

order to yield the Fokker–Plank equation. In this section we derive the Fokker–Planck equation for a general one-dimensional master equation with arbitrary production and degradation rates; the method is easily generalizable to models with regulation.

For arbitrary production rate $g_n$ and degradation rate $r_n$, the master equation reads

$$\partial_t p_n = - (g_n + r_n)\, p_n + g_{n-1} p_{n-1} + r_{n+1} p_{n+1}; \tag{79}$$

setting $g_n = g$ and $r_n = n$ recovers the simple birth–death process with time rescaled by the degradation rate, as in Eqn. 9.

### 1. Large protein number

The Fokker–Plank equation is derived under the assumption that the typical protein number is large ($n \gg 1$), such that $n$ can be approximated as a continuous variable, and a change of 1 protein can be treated as a small change. We will use parentheses when treating $n$ as continuous and a subscript when treating $n$ as discrete. Under this approximation, the function $f(n \pm 1)$, where $f(n) \in \{g(n)p(n), r(n)p(n)\}$, can be expanded to second order as

$$f(n \pm 1) = f(n) \pm \partial_n f(n) + \frac{1}{2} \partial_n^2 f(n). \tag{80}$$

Inserting the results of the expansion into Eqn. 79, we obtain the Fokker–Planck equation

$$\partial_t p(n) = -\partial_n \left[ v(n) p(n) \right] + \frac{1}{2} \partial_n^2 \left[ D(n) p(n) \right], \tag{81}$$

where $v(n) \equiv g(n) - r(n)$, an effective drift velocity, recovers the right-hand side of the deterministic equation, and $D(n) \equiv g(n) + r(n)$, an effective diffusion constant, sets the scale of the fluctuations in protein number. The drift term plays the role of an effective force. The diffusion coefficient plays the role of an effective temperature: the larger it is, the larger the excursions a single trajectory (of particle number versus time) takes from the mean.

The steady state solution of Eqn. 81 is easily obtained by noticing that the Fokker–Planck equation is a continuity equation of the form

$$\partial_t p = -\partial_n j, \tag{82}$$

where $j(n) \equiv v(n)p(n) - \frac{1}{2}\partial_n \left[ D(n)p(n) \right]$ is the current of the probability. In steady state ($\partial_t p = 0$) the current is constant, and since $p(n) \to 0$ and $\partial_n p(n) \to 0$ as $n \to \infty$ (typically more quickly than $v(n)$ or $D(n)$ diverges), the current vanishes for all $n$. The steady state distribution is then found by direct integration, i.e.

$$p(n) = \frac{1}{\mathcal{Z} D(n)} \exp \left[ 2 \int_0^n dn' \frac{v(n')}{D(n')} \right], \tag{83}$$

where $\mathcal{Z}$ is a normalization constant ensuring $\int_0^\infty dn\, p(n) = 1$.

In the simple birth–death process, for which $v(n) = g - n$ and $D(n) = g + n$, Eqn. 83 evaluates to

$$p(n) = \frac{1}{\mathcal{Z} g} \left( 1 + \frac{n}{g} \right)^{4g-1} e^{-2n}. \tag{84}$$

In Appendix D, we show that the exact steady state (the Poisson distribution, Eqn. 16) and Eqn. 84 have the same asymptotic limit for large protein number: a Gaussian distribution with mean and variance equal to $g$, the ratio of production to degradation rates.

### 2. Small noise

In addition to the approximation that the typical protein number is large, one may make the further approximation that the noise is small. This is often referred to as the "linear noise approximation" [17, 37, 38] or "small noise approximation" [32, 33]. Specifically, one assumes that the fluctuations $\eta$ in $n$ are small around the mean $\bar{n}$, i.e.

$$n = \bar{n} + \eta, \tag{85}$$

where $\bar{n} \equiv \langle n \rangle = \sum_n n p_n$ (for brevity we use bar notation in this and several subsequent sections). Since $dn/d\eta = 1$ we have $p(n) = p(\eta)$, and the master equation (Eqn. 81) becomes an equation in $\eta$:

$$\partial_t p(\eta) = -\partial_\eta \left[ v(\bar{n} + \eta) p(\eta) \right] + \frac{1}{2} \partial_\eta^2 \left[ D(\bar{n} + \eta) p(\eta) \right], \tag{86}$$

We use the approximation that $\eta$ is small to expand the drift and diffusion terms to first nonzero order in $\eta$:

$$v(\bar{n} + \eta) = v(\bar{n}) + \eta v'(\bar{n}) + \cdots \approx \eta v'(\bar{n}) \tag{87}$$
$$D(\bar{n} + \eta) = D(\bar{n}) + \eta D'(\bar{n}) + \cdots \approx D(\bar{n}) \tag{88}$$

where prime denotes differentiation with respect to $n$, and the last step in Eqn. 87 recalls the fact that $v(\bar{n}) = g(\bar{n}) - r(\bar{n}) = 0$ in the steady state, as given by the kinetic rate equation (e.g. Eqn. 8). Eqn. 86 is now

$$\partial_t p(\eta) = -v'(\bar{n}) \partial_\eta \left[ \eta p(\eta) \right] + \frac{1}{2} D(\bar{n}) \partial_\eta^2 \left[ p(\eta) \right]. \tag{89}$$

As with Eqn. 83, the steady state of Eqn. 89 is found by direct integration:

$$p(\eta) = \frac{1}{\mathcal{Z} D(\bar{n})} \exp \left[ \frac{2v'(\bar{n})}{D(\bar{n})} \int_0^\eta d\eta' \eta' \right] = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\eta^2/2\sigma^2}, \tag{90}$$

where

$$\sigma^2 \equiv \frac{D(\bar{n})}{-2v'(\bar{n})} \tag{91}$$

(note that $v'(\bar{n})$ is negative for stable fixed points). Eqn. 90 is a Gaussian distribution, meaning that under the linear noise approximation, the steady state probability distribution is a Gaussian centered at the exact mean $\bar{n}$ and with width determined by mean birth and death rates according to Eqn. 91. We note that Eqn. 90 can be equivalently derived by expanding the integrand in Eqn. 83 about its maximum (or "saddle point") $\bar{n}$.

The linear noise approximation is stricter than the large-protein number approximation made in the previous section (Sec. II C 1). While the previous approximation makes no assumption about the form of the probability distribution, the linear noise approximation assumes that the distribution is unimodal and sharply peaked around its mean. In practice, $\bar{n}$ and $\sigma^2$ are obtained by finding the steady state(s) (i.e. the stable fixed point(s)) of the corresponding deterministic rate equation. However, it is easily possible (for processes more complicated than the simple birth–death process) for the deterministic equation to have more than one stable fixed point (see Fig. 2). Although one may make a Gaussian expansion around each fixed point in turn, the linear noise approximation does not describe how these Gaussians might be weighted in a multi-modal distribution. In these cases it is most accurate to use (if solvable) either the large-protein number approximation (Eqn. 83) or the original master equation (Eqn. 79).

In the simple birth–death process, for which $v(n) = g - n$, $D(n) = g + n$, and $\bar{n} = g$, Eqn. 91 gives $\sigma^2 = 2g/[-2(-1)] = g$, and therefore Eqn. 90 reproduces the asymptotic behavior derived in Appendix D.

## D.   Langevin approximation

We now consider a second stochastic approximation: the Langevin equation. The advantage of the Langevin approach is that a large amount can be learned about the process without finding the full distribution, but instead by considering the correlation functions of the concentration, which are readily computed. Starting from the Fokker–Planck equation we can calculate the equation for the mean of the distribution $\langle n \rangle = \sum_{n=0}^{\infty} n p_n$. We arrive at the kinetic rate equation given in Eqn. 2.

We can think about the change in concentration in time as a trajectory in $n$ space. Each realization of a birth-death process, will be described by a certain $n(t)$, and averaging over many experiments, we obtain $\langle n(t) \rangle$, as given by Eqn. 2. If we consider the change of concentrations on time scales longer than the characteristic timescales of the particular reactions, we can assume that there are many birth and death processes in each interval and that the fluctuations of each realization of $n(t)$ around the mean are Gaussian distributed:

$$P[\eta] \sim e^{-\int dn \int dt \frac{\eta^2(n,t)}{4D(n)}} \tag{92}$$

We can include fluctuations around the the mean values by considering an additional noise term $\eta(t)$, such that the equation for the change of the concentration of proteins $n$ becomes:

$$\frac{dn}{dt} = v(n) + \eta(t) = g - n + \eta(t). \tag{93}$$

We require of the noise that:

$$\langle \eta(t) \rangle = 0, \tag{94}$$
$$\langle \eta(t')\eta(t) \rangle = \delta(t - t')D(n) = \delta(t - t')(g + n) \tag{95}$$

where $D(n)$ is in general the diffusion coefficient in the Fokker–Planck equation. In Appendix I we show the equivalence of the Fokker–Planck and Langevin descriptions.

In the general case when there are many types of proteins in the system, we can define a time dependent correlation function

$$C_{ij}(t) = \langle \delta N_i(0)\delta N_j(t) \rangle, \tag{96}$$

where the average implies a time average and $\delta N_i(t) = N_i(t) - \langle \delta N_i(t) \rangle$ are the variances of each type of protein. In one dimension $\delta n^2$ is the variance $\sigma^2$ (in subsequent sections we use $\delta n^2$ and $\sigma^2$ interchangeably). We note that in the steady state the time average can be replaced by an ensemble average. For the case of the single species birth–death process we have already computed the means and variances in Sec. II B 1. In case of the birth–death process, we calculate the time dependent correlation function to obey the equation:

$$\frac{dC(t)}{dt} = -C(t). \tag{97}$$

In the steady state the solution must reduce to the previously calculated variance (Appendix C): $\delta n^2 = \langle (n - \langle n \rangle)((n - \langle n \rangle)) \rangle = \langle n^2 \rangle - \langle n \rangle^2 = g$. Therefore the solution of Eqn. 97 for a simple birth–death process is

$$C(t) = ge^{-t}. \tag{98}$$

We can also consider the correlation functions in Fourier space

$$C_{ij}(\omega) = \int_0^\infty dt e^{-i\omega t} C_{ij}(t). \tag{99}$$

Often it is easier to calculate the Fourier transform of the correlation function directly from the Fourier transform of the Langevin equation, and then invert the transform back. In case of the simple birth–death process we can vary the Langevin equation around its mean values $n(t) = \langle n \rangle + \delta n(t)$ (that is simply linearize the equation around its mean) and consider the resulting equation in Fourier space to obtain

$$-i\omega\delta\tilde{n}(\omega) = -\delta\tilde{n}(\omega) + \tilde{\eta}(\omega), \tag{100}$$

where $\delta\tilde{n}(\omega) = \int_0^\infty dt e^{-i\omega t}\delta n(t)$ and

$$\tilde{\eta}(\omega) = \int_0^\infty dt e^{-i\omega t}\eta(t) \tag{101}$$
$$\langle \tilde{\eta}(\omega)\tilde{\eta}(\omega') \rangle = 2\pi\delta(\omega - \omega')(g + \langle n \rangle). \tag{102}$$

Solving Eqn. 100 for $\delta\tilde{n}(\omega)$ we find

$$\delta\tilde{n}(\omega) = \frac{\tilde{\eta}(\omega)}{1 - i\omega} \tag{103}$$

from which we may calculate the correlation function:

$$\langle \delta\tilde{n}^*(\omega)\delta\tilde{n}(\omega') \rangle = 2\pi\delta(\omega - \omega')\frac{g + \langle n \rangle}{\omega^2 + 1} = 2\pi\frac{2g}{\omega^2 + 1}, \tag{104}$$

where $\delta\tilde{n}^*(\omega)$ denotes the complex conjugate of $\delta\tilde{n}(\omega)$. Inverting the Fourier transform reproduces the real time correlation function.

$$C(t) = \int_0^\infty d\omega e^{i\omega t}\frac{2g}{\omega^2 + 1} = 2\pi i e^{-t}\frac{2g}{2i} = ge^{-t}, \tag{105}$$

where we reproduce the result of Eqn.99. The real part of the auto correlation function in Fourier space is also called the power spectrum

$$\mathcal{N}(\omega) = \langle \delta \tilde{n}^*(\omega) \delta \tilde{n}(\omega') \rangle = \int_0^\infty dt e^{-i\omega t} C_{ii}(t),$$
(106)

because it tells us which frequency modes contribute most to the form of the noise. We note that the power spectrum $\mathcal{N}(\omega)$ may be written as integral of the correlation function in real time.

### E. Comparison of the descriptions

For the birth–death process, the steady state of the kinetic rate equation agrees with the mean of the steady state probability distribution (under both the exact and approximate stochastic descriptions): $\langle n \rangle = g$. In the limit of large protein number, the Poisson distribution is well approximated by a Gaussian, and the Fokker–Planck (and therefore the Langevin) approximation is a good description. To investigate the validity of the Fokker–Planck approximation at small and intermediate protein number, in Fig. 1 we compare for a range of protein numbers the probability distribution obtained directly from the master equation (i.e. the Poisson distribution, Eqn. 16) and from the Fokker–Planck approximation (Eqn. 83). We quantify the disagreement between two distributions using the Kullback–Leibler divergence,

$$D_{KL} = \sum_n p_n \log \frac{p_n}{\tilde{p}_n},$$
(107)

where $p_n$ corresponds to the master equation and $\tilde{p}_n$ corresponds to the Fokker-Planck approximation. The Kullback–Leibler divergence is not symmetric and is therefore appropriate for a comparison between a "true" distribution ($p_n$ here) and its approximation ($\tilde{p}_n$ here). As the mean protein number $\langle n \rangle = g$ increases, the accuracy of the approximation increases, and the divergence $D_{KL}$ decreases. We plot explicitly three sample pairs of exact and approximate distributions, at small, intermediate, and large protein numbers. As expected the Fokker-Planck distribution deviates from the Poisson distribution at small protein number, and agrees well at large protein number.

### III. AUTOREGULATION

We now begin to turn our attention from the simple one-dimensional birth–death process to more realistic models of gene regulation (see [43, 44] for a discussion of regulation functions of gene expression). We start with the description of autoregulation of one gene, in which its protein production rate is an arbitrary function of its own protein number.

### A. Deterministic model

As before the mean dynamics are captured by the kinetic rate equation. The kinetic rate equation for a birth–death process in which the production rate is an arbitrary function $g(\bar{n})$ of mean protein number $\bar{n}$ is

$$\frac{d\bar{n}}{dt} = g(\bar{n}) - \bar{n}.$$
(108)

The problem becomes potentially much harder, since the autoregulation function $g(\bar{n})$ can be nonlinear.

The form of the autoregulation function depends specifically on the molecular model of regulation which is being considered. For example, the Hill function,

$$g(n) = \frac{g_- K^h + g_+ n^h}{K^h + n^h},$$
(109)

can be derived by considering a gene with two production rates $g_+$ and $g_-$ (corresponding to the states in which a transcription factor protein is bound and unbound to the DNA, respectively) for which the rate of switching to the bound state depends on the protein number (see Sec. IV and Appendix J). The derivation assumes that the rates of binding and unbinding are faster than the protein degradation rate (which sets the time-scale for changes in protein number), such that equilibrium is reached, with equilibrium binding constant $K$. We remind the reader that, as we
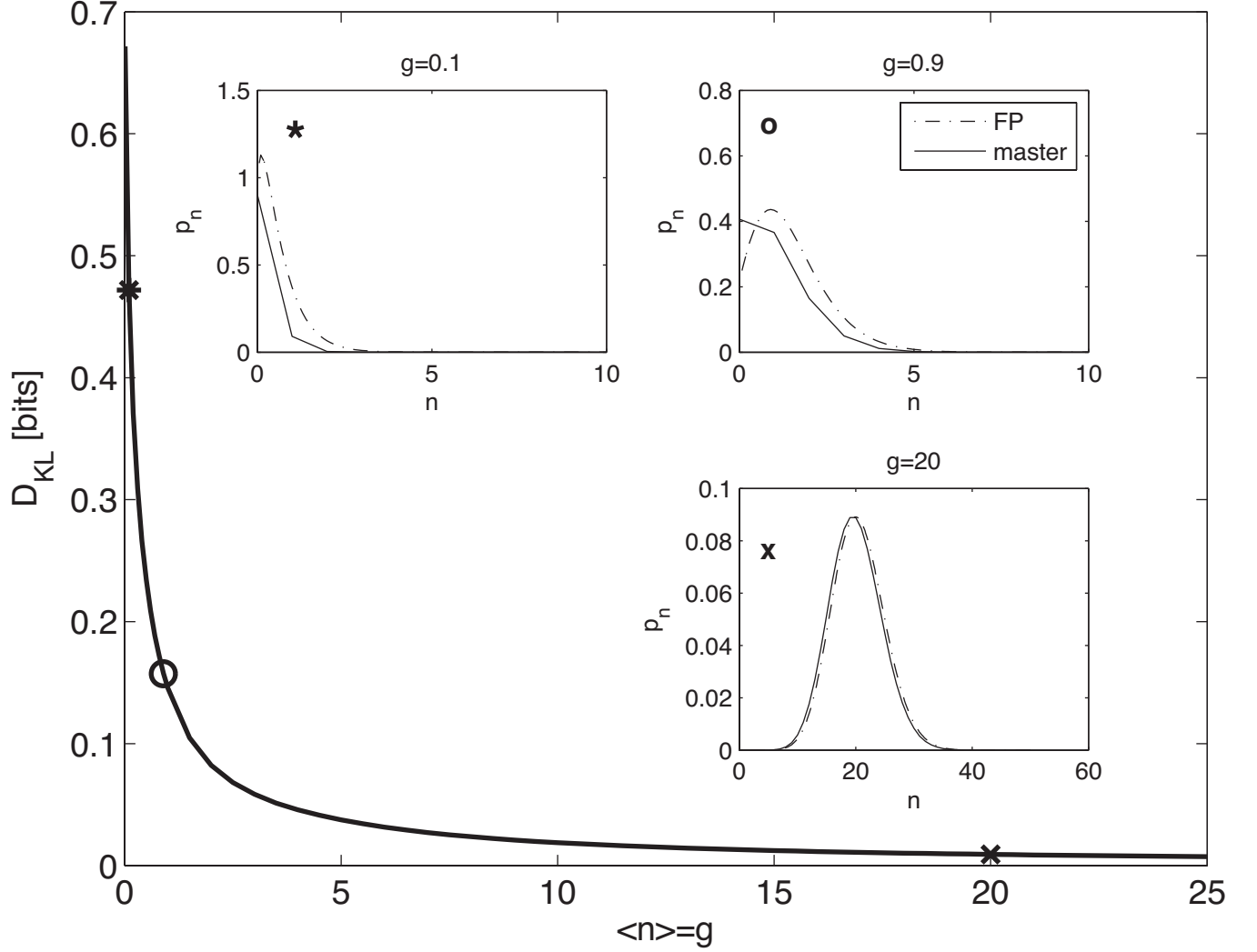
FIG. 1: Comparison of the distributions obtained from the master equation (Eqn. 16) and from the Fokker–Planck approximation (Eqn. 83), for the simple birth–death process. The Kullback–Leibler divergence $D_{KL}$ (Eqn. 107) between the two distributions is plotted as a function of mean protein number $\langle n \rangle = g$. Insets show distributions obtained from the master equation (solid) and from the Fokker–Planck equation (dashed) for $g = 0.1$ (star), $g = 0.9$ (circle), and $g = 20$ (cross), where the symbols correspond to points on the $D_{KL}$ curve.

are modeling autoregulation, there is only one species of protein in this problem: the product of the gene acts as that gene's transcription factor.

The parameter $h$ describes the cooperativity of protein binding. When $g_+ > g_-$, the case $h > 0$ corresponds to activation and the case $h < 0$ corresponds to repression; the special case $h = 0$ reproduces the simple birth–death process. For $|h| \geq 2$, Eqn. 108 has two stable fixed points for certain parameter regimes; for $|h| \geq 2$ fixed points must be found numerically.

Although the Hill equation is often used in models of gene regulation, we note that other functional forms are derivable from other biochemical processes. Many of the following results are valid for arbitrary $g(n)$.

## B. The master equation

The full stochastic description corresponding to the deterministic Eqn. 108 is given by the birth–death master equation in which the production rate is described by the arbitrary autoregulation function $g_n$ (recall that we replace

parentheses with subscript when treating $n$ as discrete):

$$\dot{p}_n = -g_n p_n - n p_n + g_{n-1} p_{n-1} + (n+1) p_{n+1}, \tag{110}$$

We may easily generalize the solution in Sec. II B 1 to find the steady state probability distribution,

$$p_n = \frac{p_0}{n!} \prod_{n'=0}^{n-1} g_{n'}, \tag{111}$$

with $p_0$ set by normalization. Except for special cases of the regulation function we cannot find a closed form solution for the distribution, but the product is easily evaluated.

The results in Sec. II for the simple birth–death process can be generalized in the case of autoregulation. We will pay particular attention to the generalization of the eigenfunction expansion in operator notation for arbitrary $g_n$, when discussing two genes in Sec. V.

## C. Bistability and noise

Auto-regulation can affect the statistics of the steady state distribution. Even before specifying the form of $g(n)$, we may obtain a general statistical result in the limit of large protein number by using the linear noise approximation (Sec. II C 2). The Fokker-Planck Eqn. 91 describes the variance $\sigma^2$ of fluctuations around the steady state mean $\bar{n}$. Here we compute the ratio $\sigma^2/\bar{n}$ for autoregulation in order to compare with the Poisson distribution, for which $\sigma^2/\bar{n} = 1$.

For autoregulation the diffusion coefficient becomes $D(\bar{n}) = g(\bar{n}) + \bar{n} = 2\bar{n}$, where we have used the fact that $0 = g(\bar{n}) - \bar{n}$ at steady state. The derivative of the drift term evaluated at the mean is $v'(\bar{n}) = g'(\bar{n}) - 1$, where the prime denotes differentiation with respect to $n$. Thus Eqn. 91 becomes

$$\frac{\sigma^2}{\bar{n}} = \frac{1}{1 - g'(\bar{n})}. \tag{112}$$

Eqn. 112 diverges when $g'(\bar{n}) = 1$. This corresponds to a bifurcation point for the self-activating gene, where there are two possible solutions to the steady state equation. In this case the ratio of the variance to the mean is not very informative about the distribution because the distribution is bimodal. The self-repressing gene, in contrast, cannot be bistable, and the ratio of the variance to the mean is always a good description of noise. In Fig. 2 we demonstrate that the exact distribution becomes bimodal when the deterministic equation crosses a bifurcation point.

The ratio in Eqn. 112 shows that self-activation ($g'(\bar{n}) > 0$) leads to super-Poissonian noise ($\sigma^2/\bar{n} > 1$), and self-repression ($g'(\bar{n}) < 0$) leads to sub-Poissonian noise ($\sigma^2/\bar{n} < 1$); when there is no regulation we have $g'(\bar{n}) = 0$, and we recover the Poisson result, $\sigma^2/\bar{n} = 1$. Fig. 3 demonstrates this behavior by computing the exact distribution (Eqn. 111) for Hill function autoregulation (Eqn. 109) with various values of the cooperativity parameter $h$.

Eqn. 112 can be be equivalently obtained by computing the power spectrum from the linearized Langevin equation. Linearizing the Langevin equation and going to Fourier space we obtain

$$-i\omega \delta \tilde{n}(\omega) = g'(\bar{n}) \tilde{n}(\omega) - \tilde{n}(\omega) + \tilde{\eta}(\omega). \tag{113}$$

The power spectrum becomes

$$\mathcal{N} = \frac{2\bar{n}}{\omega^2 + [1 - g'(\bar{n})]^2} \tag{114}$$

The steady state auto–correlation function is therefore

$$\delta n^2 = \frac{\bar{n}}{|1 - g'(\bar{n})|}, \tag{115}$$

and the ratio of the variance to the mean is

$$\frac{\delta n^2}{\bar{n}} = \frac{1}{|1 - g'(\bar{n})|}. \tag{116}$$

We point the reader to the paper of Warren, Tanase-Nicola and ten Wolde [35] for a pedagogical discussion of power spectra in gene expression.
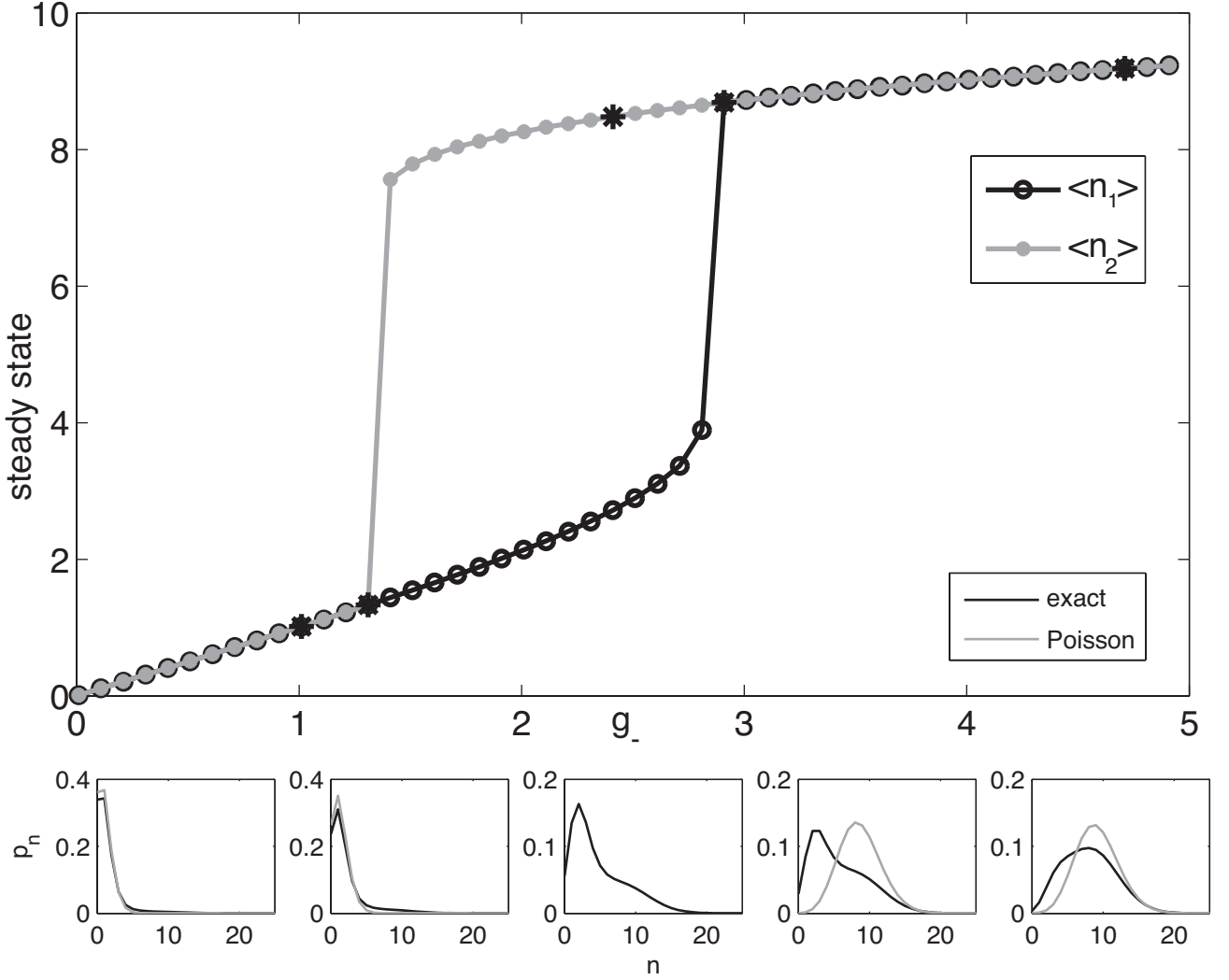
FIG. 2: The steady state solution(s) for a self-activating gene as a function of the basal rate $g_-$. The figure shows regimes which are monostable (high and low $g_-$) and bistable (intermediate $g_-$). The probability distributions below (solid black lines), obtained from Eq. 111, correspond consecutively to the points above marked by stars. In the monostable regime the distributions are compared to Poisson distributions (solid gray lines). The autoregulation function is a Hill function (Eqn. 109) with $h = 4$, $g_+ = 10$, and $K = 6$.

## IV.  BURSTS

The previous section assumes that the processes of transcriptional and translational regulation can be described by one deterministic regulation function. Recent experiments have observed a feature of gene regulation that is not captured by such a model: protein production often occurs in bursts [3, 45, 46]. Bursts can result from the gene transitioning among two or more states with different rates of protein production (for example modeled in Eqn. 117, below), for example when a transcription factor is bound or unbound, or, in higher eukaryotes, when different production states are introduced by chromatin remodeling [45]. Bursts can also arise during translation, in which a single mRNA transcript produces many copies of the protein in a short amount of time. Here we present simple models of the first type, which we call transcriptional bursts, and the second type, which we call translational bursts. We then discuss a model which combines bursting with autoregulation. We refer the reader to the literature for discussion of other burst models including those which capture both transcriptional and translational bursts [25, 47].
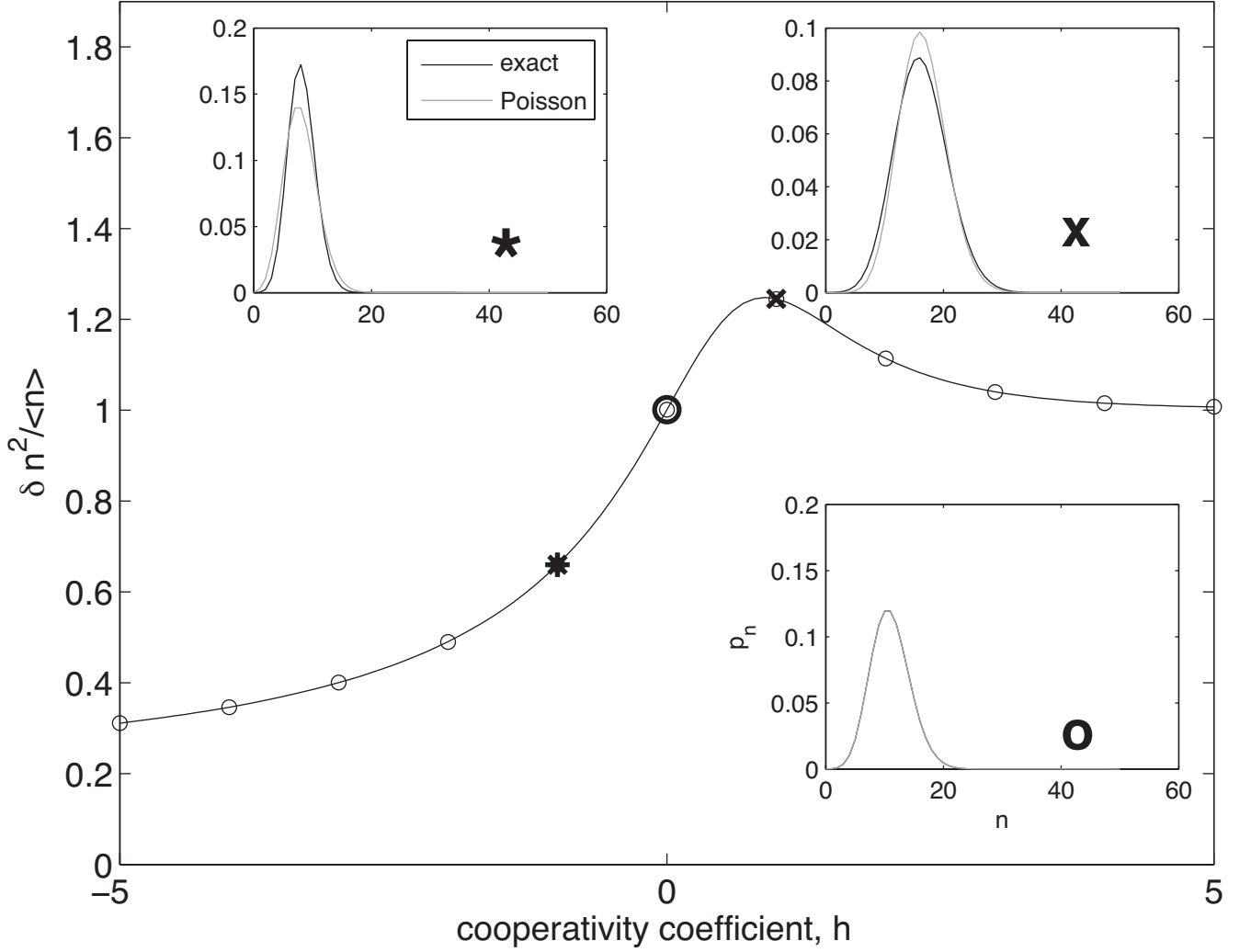
FIG. 3: The ratio of the variance $\delta n^2$ to the mean $\langle n \rangle$ as a function of the cooperativity coefficient $h$, for a gene undergoing Hill function (Eqn. 109) autoregulation. For $h < 0$ the gene represses its own production; for $h > 0$ the gene activates its own production. The case $h = 0$ recovers the simple birth-death process result $\delta n^2/\langle n \rangle = 1$. Points are plotted for integer values of $h$; the line shows Eqn. 109 for continuous $h$. Insets show probability distributions (black solid lines) from the master equation in Eq. 110, compared with Poisson distributions with the same mean (gray solid lines), at $h = -1$ (star), $h = 0$ (circle) and $h = 1$ (cross), where symbols correspond to points on the $\delta n^2/\langle n \rangle$ curve. Parameters are $g_- = 2$, $g_+ = 20$, and $K = 4$, which restrict solutions to the monostable regime, yielding unimodal distributions.

## A. Transcriptional bursts

In this section we consider a gene that can exist in multiple states, each with its own protein production rate. We first specialize to the case of two states, which is commonly used as a model for bursting [3, 34]. We solve this case using a generating function. The conventional solution is limited to two states; for an arbitrary number of states we present the spectral solution using operator methods. Finally we present a solution to a simple model with both bursting and autoregulation.

### 1. The master equation

We begin by considering the general process in which a gene has $Z$ different states. In each state $z$ the protein production is described by a simple birth-death process (Sec. II) with production rate $g_z$. Transitions between the

states occur at rates given by the transition matrix $\Omega_{zz'}$, such that

$$\sum_{z=1}^{Z} \Omega_{zz'} = 0, \tag{117}$$

which follows from conservation of probability. The rates are constant here; in Sec. IV A 4 we extend the model such that the rates depend on protein number.

The system is described by a $Z$-state probability vector with elements given by $p_n^z$. The particular states evolve according to the coupled master equation

$$\dot{p}_n^z = -g_z p_n^z - n p_n^z + g_z p_{n-1}^z + (n+1) p_{n+1}^z + \sum_{z'} \Omega_{zz'} p_n^{z'}. \tag{118}$$

where all but the last term describes the birth–death process for state $z$, and the last term describes transitions among states. The probabilities of being in state $z$ regardless of the number of proteins are given by $\pi_z = \sum_n p_n^z$; summing the master equation over $n$ in the steady state gives

$$\sum_{z'} \Omega_{zz'} \pi_{z'} = 0, \tag{119}$$

and normalization requires

$$\sum_z \pi_z = 1. \tag{120}$$

### 2. The two-state gene

We first consider the case in which the gene has only two states ($Z = 2$), an active state ($z = +$) and an inactive state ($z = -$):

$$p_n^z = \begin{pmatrix} p_n^+ \\ p_n^- \end{pmatrix}. \tag{121}$$

The transition matrix takes the form

$$\Omega_{zz'} = \begin{pmatrix} -\omega_- & \omega_+ \\ \omega_- & -\omega_+ \end{pmatrix}. \tag{122}$$

where $\omega_+$ and $\omega_-$ are the transition rates to the active and to the inactive states, respectively.

As in Sec. II B 2, we define the generating function $G_\pm(x) = \sum_n p_n^\pm x^n$ and sum the master equation (Eqn. 118) over $x^n$, yielding at steady state

$$0 = -y(\partial_y - g_\pm)G_\pm \pm \omega_+ G_- \mp \omega_- G_+, \tag{123}$$

where $y \equiv x - 1$. Writing $G_\pm(y) = e^{g_\pm y} H_\pm(y)$ yields a simpler equation for $H_\pm(y)$; writing out the $+$ and $-$ cases explicitly,

$$0 = -y e^{y\Delta} \partial_y H_+ + \omega_+ H_- - \omega_- e^{y\Delta} H_+, \tag{124}$$

$$0 = -y e^{-y\Delta} \partial_y H_- + \omega_- H_+ - \omega_+ e^{-y\Delta} H_-, \tag{125}$$

where $\Delta \equiv g_+ - g_-$. Eqns. 124-125 are a coupled pair of first-order ordinary differential equations; combining them yields a second-order differential equation. Specifically, solving Eqn. 124 (125) for $H_-$ ($H_+$) and substituting it into Eqn. 125 (124), one obtains

$$0 = u \partial_u^2 H_\pm + (\beta - u)\partial_u H_\pm - \alpha H_\pm, \tag{126}$$

where $u \equiv \mp y\Delta$, $\alpha \equiv \omega_\mp$, and $\beta \equiv \omega_+ + \omega_- + 1$. Eqn. 126 is the canonical equation for the confluent hypergeometric function. Only one of its two solutions satisfies $p_n \to 0$ as $n \to \infty$; it is $H_\pm(x) = N_\pm \Phi[\alpha, \beta; u]$, where

$$\Phi[\alpha, \beta; u] = \sum_{\ell=0}^{\infty} \frac{\Gamma(\ell + \alpha)}{\Gamma(\alpha)} \frac{\Gamma(\beta)}{\Gamma(\ell + \beta)} \frac{u^\ell}{\ell!} \tag{127}$$

is the confluent hypergeometric function of the first kind, and the scalars $N_\pm$ are found by normalization: $N_\pm = G_\pm(1) = \sum_n p_n^\pm(1)^n = \pi^\pm = \omega_\pm/(\omega_+ + \omega_-)$ (the last step uses Eqns. 119-120). Thus,

$$G_\pm(x) = \frac{\omega_\pm}{\omega_+ + \omega_-} e^{g_\pm(x-1)} \Phi[\omega_\mp, \omega_+ + \omega_- + 1; \mp(g_+ - g_-)(x-1)]. \tag{128}$$

As before, the probability distribution can be recovered from $G(x) = \sum_\pm G_\pm(x)$ by differentiation (Eqn. 21).

As shown in Appendix K, in the limit $g_- = 0$, Eqn. 128 reduces to the slightly simpler expression

$$G(x) = \Phi[\omega_+, \omega_+ + \omega_-; g_+(x-1)], \tag{129}$$

for which the probability distribution has the explicit form

$$p_n = \frac{g_+^n}{n!} \frac{\Gamma(n+\omega_+)}{\Gamma(\omega_+)} \frac{\Gamma(\omega_+ + \omega_-)}{\Gamma(n+\omega_+ + \omega_-)} \Phi[\omega_+ + n, \omega_+ + \omega_- + n; -g_+], \tag{130}$$

in agreement with the results of Iyer-Biswas et al. [34] and Raj et al. [3].

### 3. The spectral solution

It is clear that the method presented in the previous section results in a $Z$th order ordinary differential equation, and therefore its utility is generally limited to $Z = 2$ states. In contrast, here we show that the spectral solution, in which we expand in eigenfunctions of the birth–death operator, can be used for an arbitrary number of states.

We begin as in Sec. II B 3 by defining $|G_z\rangle = \sum_n p_n^z |n\rangle$ and writing the master equation in terms of raising and lowering operators,

$$|\dot{G}_z\rangle = -\hat{b}^+ \hat{b}_z^- |G_z\rangle + \sum_{z'} \Omega_{zz'} |G_{z'}\rangle, \tag{131}$$

where now $\hat{b}_z^- = \hat{a}^- - g_z$ is $z$-dependent. We expand the generating function in the $z$-dependent eigenfunctions of the birth–death process as

$$|G_z\rangle = \sum_j G_j^z |j_z\rangle, \tag{132}$$

where the eigenvalue relation reads

$$\hat{b}^+ \hat{b}_z^- |j_z\rangle = j |j_z\rangle. \tag{133}$$

This gives the following equation for the expansion coefficients:

$$\dot{G}_j^z = -j G_j^z + \sum_{z'} \Omega_{zz'} \sum_{j'} G_{j'}^{z'} \langle j_z | j'_{z'} \rangle. \tag{134}$$

Here the transition matrix $\Omega_{zz'}$ couples the otherwise independent birth–death processes, and

$$\langle j_z | j'_{z'} \rangle = \frac{(-\Delta_{zz'})^{j-j'}}{(j-j')!} \theta(j-j'), \tag{135}$$

where $\Delta_{zz'} = g_z - g_{z'}$, and the convention $\theta(0) = 1$ is used for the Heaviside function. Eqn. 135 is derived by evaluating the inner product using the $x$ space representations of the eigenfunctions,

$$\langle x | j_z \rangle = (x-1)^j e^{g_z(x-1)}, \tag{136}$$

$$\langle j_z | x \rangle = \frac{e^{-g_z(x-1)}}{(x-1)^{j+1}}, \tag{137}$$

and Cauchy's theorem, as in Appendix E.

The Heaviside function makes Eqn. 134 lower-triangular. This is explicitly clear, for example, in the steady state,

$$j G_j^z - \sum_{z'} \Omega_{zz'} G_j^{z'} = \sum_{z' \neq z} \Omega_{zz'} \sum_{j' < j} G_{j'}^{z'} \frac{(-\Delta_{zz'})^{j-j'}}{(j-j')!}, \tag{138}$$
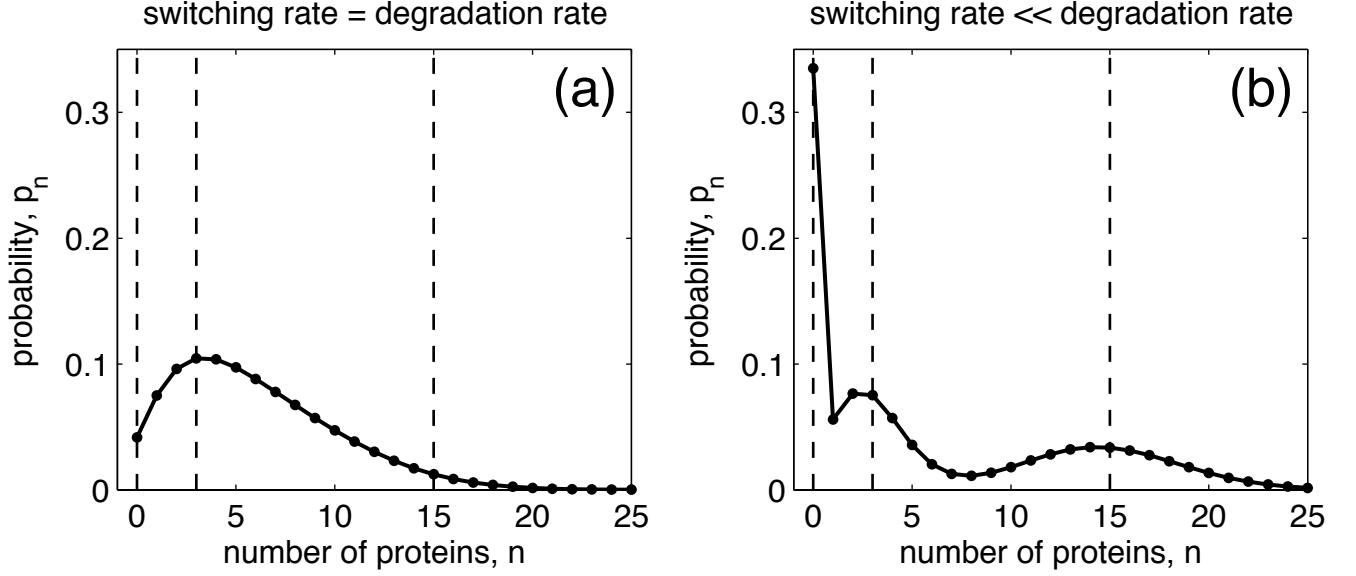
FIG. 4: Steady state probability distribution for transcriptional bursting with $Z = 3$ states, when the transition rate among states is (a) equal to the degradation rate ($\omega = 1$), and (b) much less than the degradation rate ($\omega = 0.01$). Production rates are $g_z = (0, 3, 15)$ and the rate of transition $\omega$ between any two states is equal (i.e. all off-diagonal elements of $\Omega_{zz'}$ are $\omega$ and all three diagonal elements are $-2\omega$).

where one observes that the $j$th component is a function only of components $j' < j$. The lower-triangular structure allows $G_j^z$ to be computed iteratively, which makes the spectral solution numerically efficient. The lower-triangular structure is a consequence of rotating to the eigenbasis of the birth-death operator; this structure was not present in the original master equation. Other spectral decompositions are possible by exploiting other eigenbases; for example by expanding in a single eigenbasis parameterized by a constant (not $z$-dependent) production rate, one obtains an equation that is sub-diagonal in $j$ [29].

The probability distribution is recovered via inverse transform,

$$p_n^z = \sum_j G_j^z \langle n | j_z \rangle, \tag{139}$$

as in Sec. II B 3. The spectral solution is valid for any number of states $Z$, while direct solution of the differential equation is generally limited to $Z = 2$. As expected, for $Z = 2$ the spectral solution recovers the confluent hypergeometric function solution (Eqn. 128), as shown in Appendix K.

Fig. 4 demonstrates the spectral solution for $Z = 3$ states when the transition rate among states is (a) equal to, and (b) much less than, the degradation rate. In (a) the bursts give rise to a long tail in the distribution (compared to the Poisson distribution); in (b) the system spends long times in each production state before transitioning, resulting in a trimodal distribution.

### 4. Bursting and autoregulation

Here we consider the simplest model that combines bursting with autoregulation [27]. The solution follows that of Sec. IV A 2. The gene has two states, an active state in which the regulatory protein is bound ($p_n^+$) and an inactive state in which the regulatory protein is unbound ($p_n^-$). The regulation is incorporated by making the rate at which the gene switches to the active state proportional to the protein number:

$$\Omega_{zz'} = \begin{pmatrix} -\omega_- & \omega_+ n \\ \omega_- & -\omega_+ n \end{pmatrix}. \tag{140}$$

In this case the analog of Eqn. 123 for the generating function is

$$0 = -(x - 1)(\partial_x - g_\pm)G_\pm \pm \omega_+ x \partial_x G_- \mp \omega_- G_+, \tag{141}$$

As in Sec. IV A 2, the $+$ and $-$ equations can be combined to give a second-order differential equation, e.g. for $G_-(x)$,

$$0 = \partial_x^2 G_- + C_1(x)\partial_x G_- + C_0(x)G_-, \tag{142}$$

with

$$C_1(x) \equiv \frac{g_- + g_+ + \omega_- + \omega_+ + 1 - x[g_+(1+\omega_+) + g_-]}{(1+\omega_+)x - 1}, \tag{143}$$

$$C_0(x) \equiv \frac{g_- g_+ x - g_-(g_+ + \omega_- + 1)}{(1+\omega_+)x - 1}, \tag{144}$$

whose solution $G_-(x) = N_- e^{g_+(x-1)}\Phi[\alpha, \beta; u]$ is proportional to the confluent hypergeometric function of the first kind (Eqn. 126) with

$$\alpha = 1 + \frac{\omega_-}{1+\omega_+}\left(1 + \frac{\omega_+ g_-}{g_- - (1+\omega_+)g_+}\right), \tag{145}$$

$$\beta = 1 + \frac{\omega_-}{1+\omega_+} + \frac{\omega_+ g_-}{(1+\omega_+)^2}, \tag{146}$$

$$u = -\frac{[g_+(1+\omega_+) - g_-][(1+\omega_+)x - 1]}{(1+\omega_+)^2}. \tag{147}$$

Again $N_-$ is set by normalization; $G_+(x)$ is computed from $G_-(x)$ using Eqn. 141 (bottom signs).

In Appendix J we consider an extension of this model, in which the transition rate to the active state is proportional to $n^h$, where $h$ is the cooperativity of the binding. We show that in the limit of fast transitions this model yields an effective autoregulation function of the Hill form (Eqn. 109). Full solution of this model using the method presented here for $h = 0$ and $h = 1$ is of limited utility for $h > 1$, because the high-order differential equation is not solvable.

## B. Translational bursts

We now turn our attention to translational bursts, in which a single mRNA transcript produces many proteins in a short amount of time. In this case the model includes just one protein production rate, but each production event results in an instantaneous burst of multiple proteins instead of one. This approach does not explicitly consider mRNAs; instead it models the translational step as an effective burst of proteins.

With production occurring in bursts of (integer) size $B$, the birth–death master equation becomes

$$\dot{p}_n = -gp_n - np_n + gp_{n-B} + (n+1)p_{n+1}. \tag{148}$$

This equation is easily solved by spectral decomposition. In terms of raising and lowering operators (Sec. II B 3), Eqn. 148 reads

$$|\dot{G}\rangle = \left[-g - \hat{a}^+\hat{a}^- + g(\hat{a}^+)^B + \hat{a}^-\right]|G\rangle \tag{149}$$

$$= \left(-\hat{b}^+\hat{b}^- + \Gamma\right)|G\rangle, \tag{150}$$

where $\hat{b}^+ = \hat{a}^+ - 1$ and $\hat{b}^- = \hat{a}^- - g$ as before, and $\Gamma \equiv (\hat{a}^+)^B - \hat{a}^+$. Just as in Sec. II B 3, Eqn. 150 benefits from a spectral expansion in the eigenfunctions $|j\rangle$ of the simple birth–death process, i.e.

$$|G\rangle = \sum_j G_j |j\rangle, \tag{151}$$

where

$$\hat{b}^+\hat{b}^-|j\rangle = j|j\rangle. \tag{152}$$

Eqn. 150 then becomes an equation for the expansion coefficients:

$$\dot{G}_j = -jG_j + \sum_{j'}\langle j|\Gamma|j'\rangle G_{j'}. \tag{153}$$

The simplification

$$\langle j|\Gamma|j'\rangle = \langle j|\left[\left(\hat{b}^+ + 1\right)^B - \left(\hat{b}^+ + 1\right)\right]|j'\rangle \tag{154}$$

$$= \langle j|\left[\sum_{\ell=0}^{B}\binom{B}{\ell}\left(\hat{b}^+\right)^\ell - \left(\hat{b}^+ + 1\right)\right]|j'\rangle \tag{155}$$

$$= \langle j|\left[(B-1)\hat{b}^+ + \sum_{\ell=2}^{B}\binom{B}{\ell}\left(\hat{b}^+\right)^\ell\right]|j'\rangle \tag{156}$$

$$= (B-1)\delta_{j,j'-1} + \sum_{\ell=2}^{B}\binom{B}{\ell}\delta_{j,j'+\ell}. \tag{157}$$

gives

$$\dot{G}_j = -jG_j + (B-1)G_{j-1} + \sum_{\ell=2}^{B}\binom{B}{\ell}G_{j-\ell}, \tag{158}$$

The equation is lower-triangular, which makes a recursive solution simple. Note that setting $B = 1$ recovers Eqn. 72 for the simple birth–death process. The probability distribution is retrieved by inverse transform,

$$p_n = \sum_j G_j\langle n|j\rangle. \tag{159}$$

with $\langle n|j\rangle$ computed as in Appendix H.

Translational bursts increase noise relative to the simple birth–death process. For example, we can easily compute the variance and mean of the steady state distribution for translational bursts (see Appendix C); their ratio is

$$\frac{\sigma^2}{\bar{n}} = \frac{1+B}{2}, \tag{160}$$

which for $B > 1$ is (potentially much) greater than the birth–death result $\sigma^2/\bar{n} = 1$.

We note that within the Langevin approach we can model translational bursts as:

$$\frac{dn}{dt} = Bg - n + \eta(t) \tag{161}$$

$$\langle \eta(t)\eta(t')\rangle = \delta(t - t')\left(Bg + n\right). \tag{162}$$

We obtain a power spectrum of the form:

$$\mathcal{N} = \frac{B+1}{\omega^2 + 1}, \tag{163}$$

which again is a signature of a more noisy process.

## V. TWO GENES

At this point we have discussed models of a single gene, including the simple birth–death process, autoregulation, and bursts. Here we extend the discussion to include multiple genes with regulation. We focus on the case of two genes, and we describe how the method may be extended to larger regulation networks. We refer the reader to the literature for more detailed discussion of the application of the spectral method to cascades [28] and to systems with both regulation and bursts [29].

Here we consider two genes, the first with protein number $n$, and the second with protein number $m$. The first gene regulates its own expression via arbitrary function $g_n$, and it regulates the expression of the second gene via arbitrary function $q_n$. This system can be thought of as the simplest possible regulatory network; it is also the minimum system necessary for completely describing a regulatory cascade of arbitrary length [28]. Because the model is general, the two degrees of freedom could correspond to two species of proteins, or to a protein and an mRNA species.

The master equation for this system is

$$
\begin{aligned}
\dot{p}_{nm} = {} & -g_n p_{nm} - n p_{nm} + g_{n-1} p_{n-1,m} + (n+1) p_{n+1,m} \\
& + \rho \left[ -q_n p_{nm} - m p_{nm} + q_n p_{n,m-1} + (m+1) p_{n,m+1} \right],
\end{aligned}
\tag{164}
$$

where time is rescaled by the degradation rate of the first gene and $\rho$ is the ratio of the degradation rate of the first gene to that of the second.

## A. Deterministic model

As always, the kinetic rate equations can be obtained by deriving the mean dynamics from the master equation. Summing the master equation over both indices against either $n$ or $m$ (and performing index shifts as in Appendix B) gives

$$
\frac{d\bar{n}}{dt} = g(\bar{n}) - \bar{n},
\tag{165}
$$

$$
\frac{d\bar{m}}{dt} = q(\bar{n}) - \bar{m},
\tag{166}
$$

respectively. Here we have assumed that $\sum_n g_n p_n \approx g(\bar{n})$ and $\sum_n q_n p_n \approx q(\bar{n})$ by taking $n = \bar{n} + \eta$ and keeping the first terms of the Taylor expansions $g(n) = g(\bar{n}) + \eta g'(\bar{n}) + \ldots$ and $q(n) = q(\bar{n}) + \eta q'(\bar{n}) + \ldots$, respectively.

Solving for the steady state of the deterministic system requires finding the roots of two coupled nonlinear equations, which can only be done numerically. Solving the deterministic system is essential for finding the power spectrum and calculating the correlation functions in the Langevin approach. However, as we show in the next section, the full stochastic description can be solved quite efficiently using the spectral method, i.e. by expanding in the eigenfunctions of the simple birth–death process.

## B. The spectral method

Although it is usually not possible to find an exact solution at either the master and Fokker–Planck equation levels, here we show that efficient numerical computation of the full steady state distribution is possible by expanding in the eigenfunctions of the simple birth–death process (Sec. II B 3). This approach is general and does not depend on the algebraic forms of the regulation functions.

As in Sec. II B 3 we define the generating function $|G\rangle = \sum_{nm} p_{nm} |n, m\rangle$, which makes the master equation

$$
|\dot{G}\rangle = -\hat{\mathcal{L}} |G\rangle,
\tag{167}
$$

where the full linear operator

$$
\hat{\mathcal{L}} = \hat{b}_n^+ \hat{b}_n^-(n) + \rho \hat{b}_m^+ \hat{b}_m^-(n)
\tag{168}
$$

is written in terms of raising and lowering operators defined analogously to Eqns. 59-60:

$$
\hat{b}_n^+ = \hat{a}_n^+ - 1,
\tag{169}
$$

$$
\hat{b}_m^+ = \hat{a}_m^+ - 1,
\tag{170}
$$

$$
\hat{b}_n^-(n) = \hat{a}_n^- - \hat{g}(n),
\tag{171}
$$

$$
\hat{b}_m^-(n) = \hat{a}_m^- - \hat{q}(n).
\tag{172}
$$

The difference here is that both lowering operators are $n$-dependent because of the regulation functions. This suggests that we separate the full operator as $\hat{\mathcal{L}} = \hat{\mathcal{L}}_0 + \hat{\mathcal{L}}_1(n)$, where we have defined a constant part,

$$
\hat{\mathcal{L}}_0 \equiv \hat{b}_n^+ \bar{b}_n^- + \rho \hat{b}_m^+ \bar{b}_m^-,
\tag{173}
$$

with $\bar{b}_n^- \equiv \hat{a}_n^- - \bar{g}$ and $\bar{b}_m^- \equiv \hat{a}_m^- - \bar{q}$, and an $n$-dependent part

$$
\hat{\mathcal{L}}_1(n) \equiv \hat{b}_n^+ \hat{\Gamma}(n) + \rho \hat{b}_m^+ \hat{\Delta}(n),
\tag{174}
$$

with $\hat{\Gamma}(n) \equiv \bar{g} - \hat{g}(n)$ and $\hat{\Delta}(n) \equiv \bar{q} - \hat{q}(n)$. The operator $\hat{\mathcal{L}}_0$ describes two independent birth-death processes with constant production rates $\bar{g}$ and $\bar{q}$ respectively, and the operator $\hat{\mathcal{L}}_1$ captures how the regulated processes differ from the constant rate processes. The constants $\bar{g}$ and $\bar{q}$ act as gauges: they can be chosen freely, and the final solution is independent of this choice.

We have already solved for the eigenfunctions $|j, k\rangle$ of the constant rate birth–death processes in Sec. II B 3. The spectral method exploits this fact by expanding the solution of the full problem in these eigenfunctions:

$$|G\rangle = \sum_{jk} G_{jk} |j, k\rangle. \tag{175}$$

Eqn. 167 then gives an algebraic relation among the expansion coefficients:

$$\dot{G}_{jk} = -(j + \rho k)G_{jk} - \sum_{j'} \Gamma_{j-1,j'} G_{j'k} - \rho \sum_{j'} \Delta_{jj'} G_{j',k-1}, \tag{176}$$

where

$$\Gamma_{jj'} = \langle j|\hat{\Gamma}(n)|j'\rangle = \sum_n \langle j|n\rangle(\bar{g} - g_n)\langle n|j'\rangle, \tag{177}$$

$$\Delta_{jj'} = \langle j|\hat{\Delta}(n)|j'\rangle = \sum_n \langle j|n\rangle(\bar{q} - q_n)\langle n|j'\rangle. \tag{178}$$

Eqn. 176 enjoys the property that it is sub-diagonal in the second gene's eigenvalue $k$. This feature comes from the fact that both regulation functions are dependent on the first gene's protein number $n$ (and not $m$). The sub-diagonal structure allows the matrix $G_{jk}$ to be computed iteratively by columns in $k$, which makes the method numerically efficient. The initial condition

$$G_{j0} = \langle j, k = 0|G\rangle = \sum_{nm} p_{nm}\langle j|n\rangle\langle 0|m\rangle = \sum_n p_n\langle j|n\rangle \tag{179}$$

requires the marginal probability distribution of the first gene $p_n$, which can be found recursively (Eqn. 111) as described in Sec. III B on one-dimensional auto regulation. The joint probability distribution is retrieved by inverse transform:

$$p_{nm} = \sum_{jk} G_{jk}\langle n|j\rangle\langle m|k\rangle. \tag{180}$$

To demonstrate the capability of the spectral method, Fig. 5 plots the steady state joint probability distribution for a particular choice of regulation function, for cases (a) without and (b) with autoregulation. To demonstrate the accuracy of the spectral method, Fig. 6 compares steady state marginal distributions obtained by the spectral method with those obtained both by direct iterative solution of Eqn. 164 and by standard Gillespie simulation [48]. The spectral method attains excellent agreement with both methods in orders of magnitude less runtime than either other method; full details on the efficiency and accuracy of the spectral method are available in [28].

The spectral decomposition presented here constitutes an expansion in only one of many possible eigenbases: that with constant production rates. Other spectral expansions are possible, as discussed in detail in [29]; we do not repeat the discussion here.

## VI.   SUMMARY

In this pedagogical chapter, we have attempted to introduce the reader to the phenomenon of intrinsic noise and to the simplest possible mathematical tools for modeling. We have avoided detailed discussion of a vast literature in Monte Carlo simulation of such phenomena in favor of simple calculations which can be performed analytically or which are useful in numerical solutions of the master equation itself (rather than random generation of sample trajectories).

The discussion also hopefully illustrates how different problem settings motivate different analytic methods, many of which have close parallels in the physics canon. These include approximation of differences with derivatives, yielding a description very similar to that used in Fokker-Planck descriptions; as well as the algebraic relations among the different eigenfunctions of the birth-death process itself, yielding an algebra of "ladder operators" similar to that used in quantum mechanics.
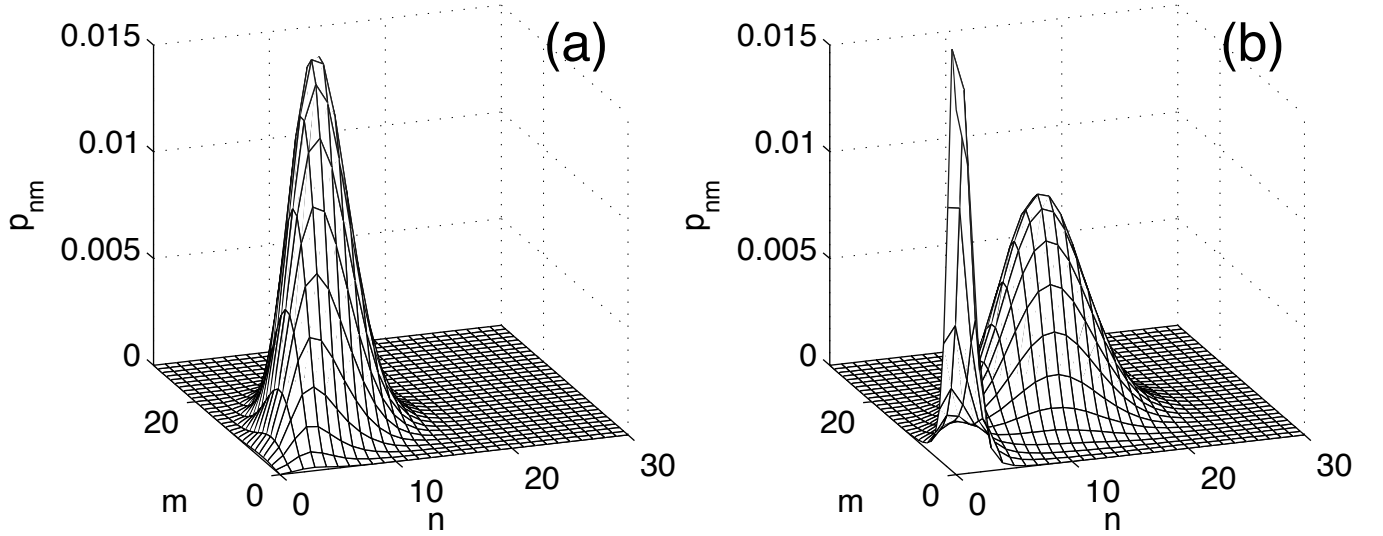
FIG. 5: Demonstration of the spectral method with two genes. Regulation is of Hill form, $q_n = (q_- K^h + q_+ n^h)/(K^h + n^h)$, with $q_- = 1$, $q_+ = 12$, $K = 4$, and $h = 4$. In (a) the first gene undergoes a simple birth–death process with $g_n = 7$; in (b) the first gene regulates itself with $g_n = q_n$. Degradation rates are equal: $\rho = 1$.
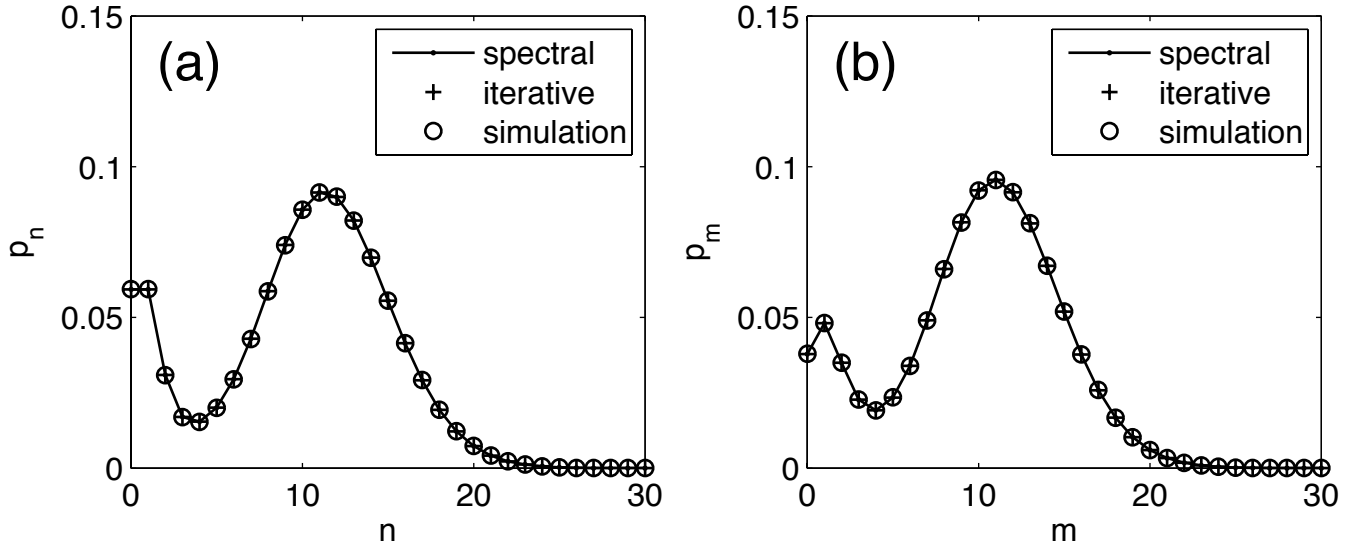


FIG. 6: Marginal distributions for the (a) first and (b) second gene in a two-gene network, summed from the joint distribution in Fig 5(b). Agreement is demonstrated among distributions obtained by the spectral method (solid), by direct iterative solution of the master equation (Eqn. 164; plus signs), and by standard Gillespie simulation [48] (circles).

While the literature in simulation of intrinsic noise has grown immensely since the introduction of time-varying Monte Carlo methods over thirty years ago [48], it is our hope that for model systems these calculations will help point out how the linearity of master equations may be exploited to yield either solutions, numerical methods, or analytical approximations which give more direct insight into the design principles of stochastic regulatory networks.

## Appendix A: Deriving the master equation

Here we derive the master equation in one dimension from the laws of probability. The probability of having $n$ proteins at a time $t + \tau$ (where $\tau$ is small) is equal to the probability of having had $n'$ proteins at time $t$ multiplied

by the (possibly time-dependent) probability of transitioning from $n'$ to $n$ in time $\tau$, summed over all $n'$:

$$p_n(t+\tau) = \sum_{n'} p_{n'}(t)p_{n|n'}(t,\tau). \tag{A1}$$

The transition probability $p_{n|n'}(t,\tau)$ has two contributions: (i) the probability of transitioning from state $n' \neq n$ to state $n$, equal to a transition rate $w_{nn'}(t)$ times the transition time $\tau$, and (ii) the probability of starting and remaining in state $n' = n$, which we call $\pi_n(t)$:

$$p_{n|n'}(t,\tau) = (1-\delta_{nn'})w_{nn'}(t)\tau + \delta_{nn'}\pi_n(t) \tag{A2}$$

Applying the normalization condition $\sum_n p_{n|n'}(t,\tau) = 1$ to Eqn. A2, we obtain

$$\pi_n(t) = 1 - \tau \sum_{n' \neq n} w_{n'n}(t), \tag{A3}$$

making Eqn. A1

$$p_n(t+\tau) = \sum_{n'} p_{n'}(t)\left[(1-\delta_{nn'})w_{nn'}(t)\tau + \delta_{nn'}\left(1 - \tau \sum_{n'' \neq n} w_{n''n}(t)\right)\right] \tag{A4}$$

$$= \tau \sum_{n'} p_{n'}(t)w_{nn'}(t) - \tau p_n(t)w_{nn} + p_n(t) - \tau p_n(t)\sum_{n'' \neq n} w_{n''n}(t) \tag{A5}$$

$$= p_n(t) + \tau\left[\sum_{n'} p_{n'}(t)w_{nn'}(t) - p_n(t)\sum_{n'} w_{n'n}(t)\right]. \tag{A6}$$

Taking the limit $\tau \to 0$, we obtain

$$\frac{dp_n}{dt} = \sum_{n'} p_{n'}w_{nn'}(t) - p_n\sum_{n'} w_{n'n}(t), \tag{A7}$$

as in Eqn. 4.

## Appendix B: Summation of the birth–death master equation

Here we sum the one-dimensional birth–death master equation over $n$ against (i) $n$, to derive the mean dynamics, (ii) $x^n$, to derive the equation of motion in generating function space, and (iii) $|n\rangle$, to find the equation of motion in operator notation.

First we sum the master equation (Eqn. 9) over $n$ against $n$ to derive the dynamics of the mean $\langle n \rangle = \sum_n np_n$:

$$\partial_t \sum_n np_n = -g\sum_n np_n - \sum_n n^2 p_n + g\sum_n np_{n-1} + \sum_n n(n+1)p_{n+1}. \tag{B1}$$

The left-hand side (LHS) is the time derivative of the mean. On the right-hand side (RHS), recalling that $n$ runs from 0 to $\infty$, we may define for the third term $n' \equiv n-1$ and for the fourth term $n'' \equiv n+1$, giving

$$\partial_t \langle n \rangle = -g\sum_n np_n - \sum_n n^2 p_n + g\sum_{n'=-1}^{\infty}(n'+1)p_{n'} + \sum_{n''=1}^{\infty}(n''-1)n''p_{n''}. \tag{B2}$$

The lower limits on all sums can be set to 0 once more without changing the expression (in the third term we impose $p_{-1} \equiv 0$ since protein number cannot be negative); simplifying gives

$$\partial_t \langle n \rangle = -g\sum_n np_n - \sum_n n^2 p_n + g\sum_{n'} n'p_{n'} + g\sum_{n'} p_{n'} + \sum_{n''} n''^2 p_{n''} - \sum_{n''} n'' p_{n''} \tag{B3}$$

$$= g\sum_{n'} p_{n'} - \sum_{n''} n'' p_{n''} \tag{B4}$$

$$= g - \langle n \rangle, \tag{B5}$$

as in Eqn. 8.

Next we sum the master equation (Eqn. 9) over $n$ against $x^n$ to derive the equation of motion of the generating function $G(x,t) = \sum_n p_n(t)x^n$:

$$\partial_t \sum_n p_n x^n = -g \sum_n p_n x^n - \sum_n n p_n x^n + g \sum_n p_{n-1} x^n + \sum_n (n+1)p_{n+1} x^n. \tag{B6}$$

The LHS and the first term on the RHS reduce directly to functions of $G$. The third and fourth terms on the RHS benefit from the same index shifts applied to obtain Eqn. B2, giving

$$\partial_t G = -gG - \sum_n p_n n x^n + g \sum_{n'} p_{n'} x^{n'+1} + \sum_{n''} n'' p_{n''} x^{n''-1}. \tag{B7}$$

We may now eliminate bare appearances of $n$ in the second and fourth terms on the RHS using $nx^n = x\partial_x x^n$ and $nx^{n-1} = \partial_x x^n$ respectively, giving

$$\partial_t G = -gG - \sum_n p_n x\partial_x x^n + g \sum_{n'} p_{n'} x^{n'+1} + \sum_{n''} p_{n''} \partial_x x^{n''} \tag{B8}$$

$$= -gG - x\partial_x G + gxG + \partial_x G \tag{B9}$$

$$= -(x-1)(\partial_x - g)G \tag{B10}$$

as in Eqn. 23.

Finally we sum the master equation (Eqn. 9) over $n$ against $|n\rangle$ to derive the equation of motion of the generating function in operator notation $|G\rangle = \sum_n p_n |n\rangle$:

$$\partial_t \sum_n p_n |n\rangle = -g \sum_n p_n |n\rangle - \sum_n n p_n |n\rangle + g \sum_n p_{n-1} |n\rangle + \sum_n (n+1)p_{n+1} |n\rangle. \tag{B11}$$

Again we may reduce the LHS and the first term on the RHS directly, and we may apply index shifts to the third and fourth terms on the RHS:

$$\partial_t |G\rangle = -g|G\rangle - \sum_n n p_n |n\rangle + g \sum_{n'} p_{n'} |n'+1\rangle + \sum_{n''} n'' p_{n''} |n''-1\rangle. \tag{B12}$$

Now defining operators $\hat{a}^+$ and $\hat{a}^-$ that raise and lower the protein number by 1 respectively, i.e.

$$\hat{a}^+|n\rangle \equiv |n+1\rangle, \tag{B13}$$

$$\hat{a}^-|n\rangle \equiv n|n-1\rangle, \tag{B14}$$

Eqn. B11 can be written

$$\partial_t |G\rangle = -g|G\rangle - \sum_n p_n \hat{a}^+ \hat{a}^- |n\rangle + g \sum_{n'} p_{n'} \hat{a}^+ |n'\rangle + \sum_{n''} p_{n''} \hat{a}^- |n''\rangle \tag{B15}$$

$$= -g|G\rangle - \hat{a}^+ \hat{a}^- |G\rangle + g\hat{a}^+ |G\rangle + \hat{a}^- |G\rangle \tag{B16}$$

$$= -(\hat{a}^+ - 1)(\hat{a}^- - g)|G\rangle \tag{B17}$$

as in Eqn. 50.

### Appendix C: Statistics of the birth–death distribution

In this appendix we calculate the means of a birth–death distribution. In the simplest case, starting directly from the solution in Eqn. 16 we obtain:

$$\langle n \rangle = \sum_n n p_n = \sum_n n \frac{g^n}{n!} e^{-g} = e^{-g} g \partial_g \sum_n \frac{g^n}{n!} = g. \tag{C1}$$

Similarly we can calculate

$$\langle n(n-1) \rangle = \sum_n n(n-1) \frac{g^n}{n!} e^{-g} = e^{-g} g^2 \partial_g^2 \sum_n \frac{g^n}{n!} = g^2, \tag{C2}$$

which gives us the variance

$$\langle n^2 \rangle - \langle n \rangle^2 = g^2 + g - g^2 = g. \tag{C3}$$

These results can also be obtained in the operator formalism, and the fact that for steady state $|G\rangle = |j = 0\rangle$. For example:

$$\langle n \rangle = \langle x | \hat{a}^+ \hat{a}^- | j = 0 \rangle |_{x=1} = \langle x | \hat{a}^+ \hat{a}^- | G \rangle |_{x=1} = \langle x | \hat{a}^+ \hat{a}^- \sum_n p_n | n \rangle |_{x=1} = \sum_n n p_n x^n |_{x=1} = \sum_n n p_n. \tag{C4}$$

To use the number operator on the $|j = 0\rangle$ state we need to rewrite it in terms of $\hat{a}^+ = \hat{b}^+ + 1$ and $\hat{a}^- = \hat{b}^- + g$. Acting to the left we obtain:

$$\langle n \rangle = \langle x | \left( \hat{b}^+ \hat{b}^- + g + \hat{b}^- + g\hat{b}^+ \right) |0\rangle_{x=1} = \langle x | g | 0 \rangle_{x=1} + \langle x | g | 1 \rangle_{x=1} = g. \tag{C5}$$

We can also calculate the mean of the bursty process. Multiplying Eqn. 148 by $n$ and summing over $n$, we use,

$$\sum_{n=0}^{\infty} n p_{n-B} = \sum_{n=0}^{\infty} (n + B) p_n = \langle n \rangle + B, \tag{C6}$$

to obtain:

$$g(\langle n \rangle + B) - g\langle n \rangle + \langle n^2 \rangle - \langle n \rangle - \langle n^2 \rangle = 0, \tag{C7}$$

which is solved by:

$$\langle n \rangle = gB. \tag{C8}$$

The second moment $\langle n^2 \rangle = \sum n^2 p_n$ is calculated the same way:

$$\langle n^2 \rangle = \frac{1}{2} \left( 2g\langle n \rangle B + \langle n \rangle + gB^2 \right) = (gB)^2 + \frac{gB^2 + gB}{2}, \tag{C9}$$

which yields the variance as

$$\delta n^2 = \langle n^2 \rangle - \langle n \rangle^2 = \frac{gB}{2}(1 + B). \tag{C10}$$

## Appendix D: Asymptotic distributions for large protein number

Here we show that in the limit of large protein number $n$, the steady state of the birth–death process (the Poisson distribution, Eqn. 16) approaches a Gaussian distribution with mean and variance equal to $g$, the ratio of production to degradation rates. We then show that in the same limit, the steady state solution to the birth–death Fokker–Planck equation (Eqn. 84) also approaches a Gaussian distribution with mean and variance equal to $g$.

The large-$n$ limit of the Poisson distribution is most conveniently evaluated by first taking the log, which allows one to make use of Stirling's approximation, $\log n! \approx n \log n - n + \log \sqrt{2\pi n}$ for large $n$:

$$\log p(n) \approx -g + n \log g - n \log n + n - \log \sqrt{2\pi n}. \tag{D1}$$

The derivative

$$\partial_n \log p(n) = \log \frac{g}{n} + \mathcal{O}(1/n) \tag{D2}$$

vanishes at the maximum

$$n = g \tag{D3}$$

about which we Taylor expand to second order:

$$\log p(n) \approx \log p(g) + \frac{1}{2}(n - g)^2 \partial_n^2 \left[ \log p(n) \right]_g \tag{D4}$$

$$= -\log \sqrt{2\pi g} - \frac{(n - g)^2}{2g}. \tag{D5}$$

The result is a Gaussian distribution with mean and variance equal to $g$:

$$p(n) = \frac{1}{\sqrt{2\pi g}} e^{-(n-g)^2/(2g)}. \tag{D6}$$

The large-$n$ limit of the Fokker–Planck distribution is similarly evaluated: the log and its derivatives are

$$\log p(n) = \log\frac{1}{\mathcal{Z}g} + (4g-1)\log\left(1 + \frac{n}{g}\right), \tag{D7}$$

$$\partial_n \log p(n) = \frac{4g-1}{n+g} - 2, \tag{D8}$$

$$\partial_n^2 \log p(n) = \frac{1-4g}{(n+g)^2}. \tag{D9}$$

The first derivative vanishes at $n = g - 1/2 \approx g$ for large $g$, at which the second derivative evaluates to $\partial_n^2 \left[\log p(n)\right]_g = -1/g + 1/(4g^2) \approx -1/g$. Taylor expanding to second order and exponentiating then give

$$p(n) = \frac{1}{\sqrt{2\pi g}} e^{-(n-g)^2/(2g)}, \tag{D10}$$

where $\mathcal{Z}$ is eliminated by normalization.

## Appendix E: Orthonormality and the inner product

Here we show that interpreting the generating function as a Fourier transform motivates a particular choice of inner product and conjugate state in the protein number basis.

We have defined the generating function in terms of a continuous variable $x$ as $G(x) = \sum_n p_n x^n$. We could equivalently write $x \equiv e^{i\varphi}$ and define

$$G(\varphi) = \sum_n p_n e^{in\varphi}, \tag{E1}$$

which makes clear that the generating function is simply the Fourier transform of the probability distribution in protein number $n$. In the state notation commonly used in quantum mechanics [42], Eqn. E1 is written $\langle \varphi | G \rangle = \sum_n p_n \langle \varphi | n \rangle$, where $\langle \varphi | G \rangle \equiv G(\varphi)$ and

$$\langle \varphi | n \rangle \equiv e^{in\varphi}. \tag{E2}$$

Eqn. E2 is the representation of $|n\rangle$ in $\varphi$ space. In order to compute projections of $|n\rangle$ onto other states or itself using this representation, we must define a conjugate state $\langle n | \varphi \rangle$ and a consistent inner product. With complex exponentials, it is common to make the conjugate state the complex conjugate,

$$\langle n | \varphi \rangle \equiv \langle \varphi | n \rangle^* = e^{-in\varphi}, \tag{E3}$$

and the inner product the integral

$$\langle f | h \rangle \equiv \int_0^{2\pi} \frac{d\varphi}{2\pi} \langle f | \varphi \rangle \langle \varphi | h \rangle, \tag{E4}$$

for any $f$ and $h$. Given the definition of conjugate state, this choice of inner product ensures the orthonormality of the $|n\rangle$ states:

$$\langle n | n' \rangle = \int_0^{2\pi} \frac{d\varphi}{2\pi} \langle n | \varphi \rangle \langle \varphi | n' \rangle = \int_0^{2\pi} \frac{d\varphi}{2\pi} e^{i(n'-n)\varphi} = \delta_{nn'}. \tag{E5}$$

We may now reinterpret these definitions in terms of our original variable $x$. Since $x = e^{i\varphi}$ we have $dx = ie^{i\varphi}d\varphi = ixd\varphi$, and in the orthonormality condition (Eqn. E5) the integration from 0 to $2\pi$ along the real $\varphi$ line segment becomes contour integration along the unit circle in the complex $x$ plane:

$$\langle n | n' \rangle = \int_0^{2\pi} \frac{d\varphi}{2\pi} e^{-in\varphi} e^{in'\varphi} = \oint \frac{dx}{2\pi ix} x^{-n} x^{n'} = \oint \frac{dx}{2\pi i} \frac{1}{x^{n+1}} x^{n'} = \delta_{nn'}. \tag{E6}$$

Since we have defined

$$\langle x|n \rangle = x^n, \tag{E7}$$

Eqn. E6 suggests the definition of conjugate state

$$\langle n|x \rangle = \frac{1}{x^{n+1}} \tag{E8}$$

and inner product

$$\langle f|h \rangle = \oint \frac{dx}{2\pi i} \langle f|x \rangle \langle x|h \rangle \tag{E9}$$

in $x$ space.

With these definitions, we will find Cauchy's theorem,

$$\oint \frac{dx}{2\pi i} \frac{f(x)}{(x-a)^{n+1}} = \frac{1}{n!} \partial_x^n [f(x)]_{x=a} \, \theta(n), \tag{E10}$$

(with the convention $\theta(0) = 1$ for the Heaviside function) very useful in evaluating projections. For example, we may immediately use it to confirm the orthonormality condition,

$$\langle n|n' \rangle = \oint \frac{dx}{2\pi i} \frac{1}{x^{n+1}} x^{n'} = \frac{1}{n!} \partial_x^n \left[ x^{n'} \right]_{x=0} \theta(n) = \delta_{nn'}, \tag{E11}$$

and in Appendix H we use it to compute analytic expressions for the projections between protein number states $|n\rangle$ and birth–death eigenstates $|j\rangle$.

## Appendix F: Properties of the raising and lowering operators

Just as in the operator treatment of the quantum harmonic oscillator [42], in this paper we have defined operators $\hat{a}^+$ and $\hat{a}^-$ that act on $|n\rangle$ states by raising and lowering the protein number $n$ by 1 respectively, i.e.

$$\hat{a}^+|n\rangle = |n+1\rangle, \tag{F1}$$
$$\hat{a}^-|n\rangle = n|n-1\rangle \tag{F2}$$

(note, however, that the prefactors here, 1 and $n$, are different than those conventionally used for the harmonic oscillator, $\sqrt{n+1}$ and $\sqrt{n}$, respectively). In this appendix we derive the actions of these operators when acting to the left, as well as their commutation relation.

The actions of $\hat{a}^+$ and $\hat{a}^-$ to the left can be found by projecting onto Eqns. F1 and F2 the conjugate state $\langle n'|$:

$$\langle n'|\hat{a}^+|n\rangle = \langle n'|n+1\rangle = \delta_{n',n+1} = \delta_{n'-1,n} = \langle n'-1|n\rangle, \tag{F3}$$
$$\langle n'|\hat{a}^-|n\rangle = \langle n'|n|n-1\rangle = n\delta_{n',n-1} = (n'+1)\delta_{n'+1,n} = (n'+1)\langle n'+1|n\rangle. \tag{F4}$$

The first and last expressions in each case imply

$$\langle n'|\hat{a}^+ = \langle n'-1|, \tag{F5}$$
$$\langle n'|\hat{a}^- = (n'+1)\langle n'+1|, \tag{F6}$$

as in Eqns. 56-57.

The commutation relation between $\hat{a}^-$ and $\hat{a}^+$ is defined as $[\hat{a}^-, \hat{a}^+] \equiv \hat{a}^-\hat{a}^+ - \hat{a}^+\hat{a}^-$. It is evaluated by considering its action on a state $|n\rangle$:

$$\left[\hat{a}^-, \hat{a}^+\right]|n\rangle = \hat{a}^-\hat{a}^+|n\rangle - \hat{a}^+\hat{a}^-|n\rangle = \hat{a}^-|n+1\rangle - \hat{a}^+n|n-1\rangle = (n+1)|n\rangle - n|n\rangle = |n\rangle. \tag{F7}$$

The first and last expressions imply

$$\left[\hat{a}^-, \hat{a}^+\right] = 1, \tag{F8}$$

just as with the quantum harmonic oscillator.

## Appendix G: Eigenvalues and eigenfunctions in the operator representation

Here we derive the eigenfunctions $|\lambda_j\rangle$ (also called eigenstates in the operator representation) and eigenvalues $\lambda_j$ of the operator $\hat{\mathcal{L}} = \hat{b}^+\hat{b}^-$, for which

$$\hat{\mathcal{L}}|\lambda_j\rangle = \lambda_j|\lambda_j\rangle, \tag{G1}$$

as well as the actions of the individual operators $\hat{b}^+$ and $\hat{b}^-$ on the eigenstates. We will find that the commutation relation

$$\left[\hat{b}^-, \hat{b}^+\right] = 1 \tag{G2}$$

and the existence of the steady state solution $\langle x|G\rangle = e^{q(x-1)}$, for which

$$\hat{\mathcal{L}}|G\rangle = 0, \tag{G3}$$

are all that are necessary to completely define the eigenvalues and eigenstates of $\hat{\mathcal{L}}$. The treatment here finds many parallels with the derivation of the eigenvalue spectrum of the quantum harmonic oscillator [42].

First it is useful to compute the commutation relation of $\hat{\mathcal{L}}$ with each of its components $\hat{b}^+$ and $\hat{b}^-$:

$$\left[\hat{b}^+, \hat{\mathcal{L}}\right] = \left[\hat{b}^+, \hat{b}^+\hat{b}^-\right] = \hat{b}^+\left[\hat{b}^+, \hat{b}^-\right] + \left[\hat{b}^+, \hat{b}^+\right]\hat{b}^- = -\hat{b}^+, \tag{G4}$$

$$\left[\hat{b}^-, \hat{\mathcal{L}}\right] = \left[\hat{b}^-, \hat{b}^+\hat{b}^-\right] = \hat{b}^+\left[\hat{b}^-, \hat{b}^-\right] + \left[\hat{b}^-, \hat{b}^+\right]\hat{b}^- = \hat{b}^- \tag{G5}$$

(here we have used Eqn. G2 and the properties $[f, gh] = g[f, h] + [f, g]h$ and $[f, f] = 0$ for any $f$, $g$, and $h$). We now consider a particular eigenvalue $\lambda$, and evaluate the action of $\hat{\mathcal{L}}\hat{b}^+$ on $|\lambda\rangle$:

$$\hat{\mathcal{L}}\hat{b}^+|\lambda\rangle = \hat{b}^+\hat{\mathcal{L}}|\lambda\rangle - \left[\hat{b}^+, \hat{\mathcal{L}}\right]|\lambda\rangle = \hat{b}^+\lambda|\lambda\rangle + \hat{b}^+|\lambda\rangle = (\lambda+1)\hat{b}^+|\lambda\rangle. \tag{G6}$$

The equality of the first and last expressions,

$$\hat{\mathcal{L}}\left(\hat{b}^+|\lambda\rangle\right) = (\lambda+1)\left(\hat{b}^+|\lambda\rangle\right) \tag{G7}$$

reveals two results: (i) the existence of a state with eigenvalue $\lambda$ implies the existence of a state with eigenvalue $\lambda + 1$, and (ii) the state with eigenvalue $\lambda + 1$ is proportional to $\hat{b}^+|\lambda\rangle$. By induction the first result means that the eigenvalues are spaced by 1; moreover, the existence of a state with eigenvalue zero (the steady state solution; Eqn. G3) anchors the eigenvalues to the integers, i.e.

$$\lambda_j = j \tag{G8}$$

for integer $j$. The second result can now be written explicitly

$$\hat{b}^+|j\rangle = |j+1\rangle, \tag{G9}$$

where we are free by normalization to set the proportionality constant to 1. Eqn. G9 demonstrates that $\hat{b}^+$ is a raising operator for the eigenstates $|j\rangle$.

Evaluating the action of $\hat{\mathcal{L}}\hat{b}^-$ on eigenstate $|j\rangle$ similarly reveals that $\hat{b}^-|j\rangle$ is proportional to $|j-1\rangle$ (cf. Eqn. G6); consistency with the eigenvalue equation (Eqn. G1) requires that the proportionality constant be $j$, giving

$$\hat{b}^-|j\rangle = j|j-1\rangle. \tag{G10}$$

Eqn. G10 demonstrates that $\hat{b}^-$ is a lowering operator for the eigenstates $|j\rangle$. The operators $\hat{b}^+$ and $\hat{b}^-$ raise and lower $|j\rangle$ as $\hat{a}^+$ and $\hat{a}^-$ do $|n\rangle$; therefore they act to the left as (cf. Eqns. F3-F4)

$$\langle j|\hat{b}^+ = \langle j-1|, \tag{G11}$$

$$\langle j|\hat{b}^- = (j+1)\langle j+1|. \tag{G12}$$

Eqn. G10 imposes a floor on the eigenvalue spectrum, since $\hat{b}^-|0\rangle = 0$, and no lower eigenstates can be generated. Therefore the eigenvalues are limited to nonnegative integers:

$$j \in \{0, 1, 2, 3, \dots\}. \tag{G13}$$

Eqn. G9 describes how any state can be obtained from the $j = 0$ state:

$$|j\rangle = \left(\hat{b}^+\right)^j |0\rangle. \tag{G14}$$

Recalling that $\hat{b}^+ = \hat{a}^+ - 1$ and that $|0\rangle$ is the steady state solution, Eqn. G14 can be used to derive the representation of the eigenfunctions in $x$ space, $\langle x|j\rangle$. Projecting $\langle x|$ onto Eqn. G14 and recalling that $\hat{a}^+$ corresponds to $x$ (Eqn. 54) and that $\langle x|0\rangle = e^{g(x-1)}$ (Eqn. 24), we obtain

$$\langle x|j\rangle = (x-1)^j e^{g(x-1)}. \tag{G15}$$

We may now define the conjugate state $\langle j|x\rangle$ such that the $|j\rangle$ are orthonormal under the inner product defined in Eqn. E9:

$$\delta_{jj'} = \langle j|j'\rangle = \oint \frac{dx}{2\pi i} \langle j|x\rangle\langle x|j'\rangle = \oint \frac{dx}{2\pi i} \langle j|x\rangle (x-1)^{j'} e^{g(x-1)} = \oint \frac{dy}{2\pi i} y^{j'} f_j(y), \tag{G16}$$

where $y \equiv x - 1$ and $f_j(x-1) \equiv e^{g(x-1)}\langle j|x\rangle$. Eqn. G16 is equivalent to Eqn. E6, from which we identify $f_j(y) = 1/y^{j+1}$ and thus

$$\langle j|x\rangle = \frac{e^{-g(x-1)}}{(x-1)^{j+1}}. \tag{G17}$$

We have shown that the operator $\hat{\mathcal{L}} = \hat{b}^+\hat{b}^-$ has nonnegative integer eigenvalues $j$ and eigenfunctions (in $x$ space) given by Eqns. G15 and G17.

## Appendix H: Overlaps between protein number states and eigenstates

Here we describe two methods for computing the overlaps $\langle n|j\rangle$ and $\langle j|n\rangle$ between the protein number states $|n\rangle$ and the eigenstates $|j\rangle$: by contour integration and by recursive updating.

The first method evaluates the overlaps using the inner product defined in Eqn. E9 and Cauchy's theorem (Eqn. E10). Recalling the representations in $x$ space of the protein number states and eigenstates (Eqns. 46, 48, and 65-66), the first overlap becomes

$$\langle n|j\rangle = \oint \frac{dx}{2\pi i} \langle n|x\rangle\langle x|j\rangle = \oint \frac{dx}{2\pi i} \frac{e^{g(x-1)}(x-1)^j}{x^{n+1}} = \frac{1}{n!}\partial_x^n \left[e^{g(x-1)}(x-1)^j\right]_{x=0}. \tag{H1}$$

Repeated derivatives of a product follow a binomial expansion, i.e.

$$\langle n|j\rangle = \frac{1}{n!}\sum_{\ell=0}^n \frac{n!}{\ell!(n-\ell)!}\partial_x^{n-\ell}\left[e^{g(x-1)}\right]_{x=0} \partial_x^\ell \left[(x-1)^j\right]_{x=0} \tag{H2}$$

$$= \sum_{\ell=0}^n \frac{1}{\ell!(n-\ell)!}\left[g^{n-\ell}e^{-g}\right]\left[\frac{j!}{(j-\ell)!}(-1)^{j-\ell}\theta(j-\ell)\right] \tag{H3}$$

$$= (-1)^j e^{-g} g^n j! \xi_{nj}, \tag{H4}$$

where we define

$$\xi_{nj} \equiv \sum_{\ell=0}^{\min(n,j)} \frac{1}{\ell!(n-\ell)!(j-\ell)!(-g)^\ell}. \tag{H5}$$

Following similar steps, the conjugate overlap evaluates to

$$\langle j|n\rangle = n!(-g)^j \xi_{nj}, \tag{H6}$$

with $\xi_{nj}$ as in Eqn. H5. For the special case $j = 0$, Eqns. H4 and H6 reduce to Eqns. 77-78.

It is more computationally efficient to compute the overlaps recursively using rules that can be derived from the raising and lowering operations (Eqns. 52-53, 56-57, and 67-70). For example, using the raising operators,

$$\langle n|j+1\rangle = \langle n|\hat{b}^+|j\rangle = \langle n|(\hat{a}^+ - 1)|j\rangle = \langle n-1|j\rangle - \langle n|j\rangle, \tag{H7}$$

which can be initialized using $\langle n|0\rangle = e^{-g}g^n/n!$ (Eqn. H4) and updated recursively in $j$. Eqn. H7 makes clear that in $n$ space the $(j+1)$th mode is simply the (negative of the) discrete derivative of the $j$th mode. The lowering operators give an alternative update rule,

$$(n+1)\langle n+1|j\rangle = \langle n|\hat{a}^-|j\rangle = \langle n|(\hat{b}^- + g)|j\rangle = j\langle n|j-1\rangle + g\langle n|j\rangle, \tag{H8}$$

which can be initialized using $\langle 0|j\rangle = (-1)^j e^{-g}$ (Eqn. H4) and updated recursively in $n$. One may similarly derive recursion relations for $\langle j|n\rangle$, i.e.

$$\langle j|n+1\rangle = \langle j-1|n\rangle + \langle j|n\rangle, \tag{H9}$$
$$(j+1)\langle j+1|n\rangle = n\langle j|n-1\rangle - g\langle j|n\rangle, \tag{H10}$$

initialized with $\langle j|0\rangle = (-g)^j/j!$ or $\langle 0|n\rangle = 1$ respectively (Eqn. H6) and updated recursively in $n$ or $j$ respectively. Two-term recursion relations can be similarly derived from the full operator $\hat{b}^+\hat{b}^-$ [29].

## Appendix I: The equivalence of the Fokker–Planck and Langevin descriptions

In this appendix we start from the Langevin equation defined in Eqns. 93 and 94 and derive the Fokker–Planck equation in Eqn. 81. As described in Sec. II D the observed trajectory is one realization $r_\eta(t)$ of the random stochastic process with Gaussian noise as presented in Eqn. 92:

$$p_n(t) = \int \mathcal{D}\eta\, \delta\left(n - r_\eta(t)\right) P[\eta] \equiv \langle \delta\left(n - r_\eta(t)\right)\rangle. \tag{I1}$$

Consider the evolution of the probability distribution:

$$p_n(t + \Delta t) - p_n(t) = \langle \delta\left(n - r_\eta(t+\Delta t)\right) - \delta\left(n - r_\eta(t)\right)\rangle, \tag{I2}$$

and expand the increment in the trajectory as $r(t + \Delta t) = r(t) + \Delta r(t)$. We can now Taylor expand the difference in delta functions:

$$p_n(t + \Delta t) - p_n(t) = \langle (-\Delta r(t))\delta'\left(n - r_\eta(t)\right) + \frac{1}{2}(-\Delta r(t))^2 \delta''\left(n - r_\eta(t)\right)\rangle, \tag{I3}$$

$$= -\partial_n \langle \Delta r(t)\delta\left(n - r_\eta(t)\right)\rangle + \frac{1}{2}\partial_n^2 \langle (\Delta r(t))^2 \delta\left(n - r_\eta(t)\right)\rangle, \tag{I4}$$

where the primes denote derivatives in $n$. Using the Langevin equation to calculate the increments in the trajectories:

$$\Delta r(t) = v\left[r(t)\right]\Delta t + \eta(t)\Delta t \tag{I5}$$

we obtain:

$$\langle \Delta r(t)\delta\left(n - r_\eta(t)\right)\rangle = \langle v\left[r(t)\right]\Delta t\delta\left(n - r_\eta(t)\right)\rangle + \langle \eta(t)\Delta t\delta\left(n - r_\eta(t)\right)\rangle \tag{I6}$$

$$= v\left[r(t)\right]\Delta t\langle \delta\left(n - r_\eta(t)\right)\rangle = v(n)\,\Delta t p_n(t) \tag{I7}$$

$$\langle (\Delta r(t))^2 \delta\left(n - r_\eta(t)\right)\rangle = (\Delta t)^2 \langle (v\left[r(t)\right])^2 \delta\left(n - r_\eta(t)\right)\rangle + 2\langle v\left[r(t)\right]\eta(t)\delta\left(n - r_\eta(t)\right)\rangle + \langle \eta^2(t)\delta\left(n - r_\eta(t)\right)\rangle \tag{I8}$$

$$= (\Delta t)^2 (v(n))^2 p_n(t) + \langle \eta^2(t)\rangle p_n(t), \tag{I9}$$

where we have used (assuming a discretized process):

$$\langle \eta(t)\delta\left(n - r_\eta(t)\right)\rangle = \langle \eta(t)\rangle\langle \delta\left(n - r_\eta(t)\right)\rangle = 0. \tag{I10}$$

We identify

$$D(n) = \frac{\Delta t}{2}\langle \eta^2(t)\rangle. \tag{I11}$$

Putting together all elements and keeping only leading terms in $\Delta t$:

$$\frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} = -\partial_n \left[ v(n) \Delta t p_n(t) \right] + \partial_n^2 \left[ D(n) p_n(t) \right]. \tag{I12}$$

Taking the limit $\Delta t \to 0$ we recover the Fokker–Planck equation:

$$\partial_t p_n(t) = -\partial_n \left[ v(n) \Delta t p_n(t) \right] + \partial_n^2 \left[ D(n) p_n(t) \right]. \tag{I13}$$

## Appendix J: Derivation of the Hill function

The Hill function can be derived as an effective production rate for an autoregulating gene with two production states (Sec. IV A 4). The gene is found either in an inactive state $(-)$, in which the production rate is a constant $g_-$, or in an active state $(+)$, in which the production rate $\omega_+ n^h$ depends on the number of proteins $n$, which incorporates the autoregulation; $h$ describes the cooperativity, with $h > 0$ corresponding to activation and $h < 0$ corresponding to repression. The probability $p_n^\pm$ of the gene being in a given state $+$ or $-$ and there being $n$ proteins evolves in time according to the master equation in Eqn. 118 with $\Omega_{zz'}$ as in Eqn. 140; in the steady state:

$$0 = -\mathcal{L}^\pm p_n^\pm \pm \omega_+ n^h p_n^- \mp \omega_- p_n^+, \tag{J1}$$

where $\mathcal{L}^\pm$ describe simple birth–death terms with constant production rates $g_\pm$ in each of the two states. We define moments of the master equation as

$$\pi^\pm \equiv \sum_n p_n^\pm, \tag{J2}$$

$$\pi^\pm \mu_\ell^\pm \equiv \sum_n p_n^\pm n^\ell \quad \text{for } \ell \geq 1. \tag{J3}$$

with

$$\pi^+ + \pi^- = 1 \tag{J4}$$

by normalization.

Summing Eqn. J1 (top signs) over $n$ (and recalling that the birth–death terms sum to zero) gives

$$0 = \omega_+ \sum_n n^h p_n^- - \omega_- \sum_n p_n^+ \tag{J5}$$

$$= \omega_+ \pi^- \mu_h^- - \omega_- \pi^+ \tag{J6}$$

which, with Eqn. J4, becomes an expression for $\pi^+$, the probability of being in the active state:

$$\pi^+ = \frac{\mu_h^-}{\mu_h^- + \omega_+/\omega_-}. \tag{J7}$$

We solve for $\mu_h^-$ using two approximations. The first is that higher moments can be be decoupled, i.e.

$$\mu_{h+1}^- \approx \mu_h^- \mu_1^- \quad \text{for } h \geq 1, \tag{J8}$$

which implies that

$$\mu_h^- \approx (\mu_1^-)^h. \tag{J9}$$

This approximation allows one to simplify the mean equation in the $+$ state (obtained by summing the Eqn. J1, top signs, against $n$ over $n$):

$$0 = \pi^+ g_+ - \sum_n n p_n^+ + \omega_+ \sum_n n^{h+1} p_n^- - \omega_- \sum_n n p_n^+ \tag{J10}$$

$$= \pi^+ g_+ - \pi^+ \mu_1^+ + \omega_+ \pi^- \mu_{h+1}^- - \omega_- \pi^+ \mu_1^+ \tag{J11}$$

$$\approx \pi^+ g_+ - \pi^+ \mu_1^+ + \omega_+ \pi^- \mu_h^- \mu_1^- - \omega_- \pi^+ \mu_1^+ \tag{J12}$$

$$= \pi^+ g_+ - \pi^+ \mu_1^+ + \omega_- \pi^+ \left( \mu_1^- - \mu_1^+ \right), \tag{J13}$$

where the last step uses Eqn. J6. The second approximation is that transitions between states are fast, i.e. $\omega_+ \sim \omega_- \gg 1$. This approximation allows us to neglect the first two terms of Eqn. J13 compared to the third term, which implies that $\mu_1^- \approx \mu_1^+$. Summing Eqn. J3 over $\pm$ for $\ell = 1$ then gives

$$\pi^+ \mu_1^+ + \pi_- \mu_1^- \approx \left(\pi^+ + \pi^-\right) \mu_1^- = \mu_1^- = \sum_n p_n n = \bar{n}, \tag{J14}$$

which shows that $\mu_1^-$ approximates the mean of the distribution. Therefore with Eqn. J9 the probability that the gene is active (Eqn. J7) can be written

$$\pi^+ = \frac{\bar{n}^h}{\bar{n}^h + K}, \tag{J15}$$

with equilibrium constant $K \equiv \omega_+/\omega_-$. The effective production rate is the sum of the production rates in each state times the corresponding probabilities of being in each state:

$$g(\bar{n}) = g_- \pi^- + g_+ \pi^+ \tag{J16}$$

$$= g_- \left(1 - \frac{\bar{n}^h}{\bar{n}^h + K}\right) + g_+ \frac{\bar{n}^h}{\bar{n}^h + K} \tag{J17}$$

$$= \frac{g_- K + g_+ \bar{n}^h}{\bar{n}^h + K}, \tag{J18}$$

which is the Hill function (Eqn. 109).

## Appendix K: Limiting case of the two-state gene

Here we show that the spectral solution to transcriptional bursting (Eqn. 138) reduces to the hypergeometric form (Eqn. 128) in the limit of $Z = 2$ production states. We also derive the slightly simpler expression for the special case of zero production in the inactive state.

In the case of two states (Eqns. 121-122; $z = \pm$), Eqn. 138 becomes

$$jG_j^\pm + \omega_\mp G_j^\pm - \omega_\pm G_j^\mp = \omega_\pm \sum_{j' < j} G_{j'}^\mp \frac{(\mp\Delta)^{j-j'}}{(j-j')!}, \tag{K1}$$

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

$$G_j^\pm = \frac{\omega_\pm}{\omega_+ + \omega_-} \frac{(\mp\Delta)^j}{j!} \frac{\prod_{j'=0}^{j-1}(j' + \omega_\mp)}{\prod_{j''=0}^{j-1}(j'' + \omega_+ + \omega_- + 1)} \tag{K2}$$

$$= \frac{\omega_\pm}{\omega_+ + \omega_-} \frac{(\mp\Delta)^j}{j!} \frac{\Gamma(j + \omega_\mp)}{\Gamma(\omega_\mp)} \frac{\Gamma(\omega_+ + \omega_- + 1)}{\Gamma(j + \omega_+ + \omega_- + 1)}, \tag{K3}$$

where in the second line the products are written in terms of the Gamma function. Writing the total generating function $|G\rangle = \sum_\pm |G_\pm\rangle$ in position space recovers the hypergeometric form (Eqn. 128):

$$G(x) = \sum_\pm \langle x|G_\pm\rangle \tag{K4}$$

$$= \sum_\pm \sum_j \langle x|j_\pm\rangle\langle j_\pm|G_\pm\rangle \tag{K5}$$

$$= \sum_\pm \sum_j (x-1)^j e^{g_\pm(x-1)} G_j^\pm \tag{K6}$$

$$= \sum_\pm \frac{\omega_\pm}{\omega_+ + \omega_-} e^{g_\pm(x-1)} \Phi[\omega_\mp, \omega_+ + \omega_- + 1; \mp\Delta(x-1)], \tag{K7}$$

where

$$\Phi[\alpha, \beta; u] = \sum_{j=0}^{\infty} \frac{\Gamma(j+\alpha)}{\Gamma(\alpha)} \frac{\Gamma(\beta)}{\Gamma(j+\beta)} \frac{u^j}{j!} \tag{K8}$$

is the confluent hypergeometric function of the first kind.

In the limit $g_- = 0$, Eqn. K7 reads

$$G(x) = \frac{\omega_+}{\omega_+ + \omega_-} e^u \Phi[\omega_-, \omega_+ + \omega_- + 1; -u] + \frac{\omega_-}{\omega_+ + \omega_-} \Phi[\omega_+, \omega_+ + \omega_- + 1; u], \tag{K9}$$

where $u \equiv g_+(x-1)$. Using the fact that [49]

$$e^u \Phi[\alpha, \beta; -u] = \Phi[\beta - \alpha, \beta; u], \tag{K10}$$

Eqn. K9 can be written

$$G(x) = \frac{\omega_+}{\omega_+ + \omega_-} \Phi[\omega_+ + 1, \omega_+ + \omega_- + 1; u] + \frac{\omega_-}{\omega_+ + \omega_-} \Phi[\omega_+, \omega_+ + \omega_- + 1; u], \tag{K11}$$

or, noting Eqn. K8 and the fact that $\Gamma(s+1) = s\Gamma(s)$ for any $s$,

$$G(x) = \sum_j \left( \frac{\omega_+}{\omega_+ + \omega_-} \frac{\Gamma(j + \omega_+ + 1)}{\Gamma(\omega_+ + 1)} + \frac{\omega_-}{\omega_+ + \omega_-} \frac{\Gamma(j + \omega_+)}{\Gamma(\omega_+)} \right) \frac{\Gamma(\omega_+ + \omega_- + 1)}{\Gamma(j + \omega_+ + \omega_- + 1)} \frac{u^j}{j!}, \tag{K12}$$

$$= \sum_j \left( \frac{\omega_+}{\omega_+ + \omega_-} \frac{(j + \omega_+)\Gamma(j + \omega_+)}{\omega_+ \Gamma(\omega_+)} + \frac{\omega_-}{\omega_+ + \omega_-} \frac{\Gamma(j + \omega_+)}{\Gamma(\omega_+)} \right) \frac{(\omega_+ + \omega_-)\Gamma(\omega_+ + \omega_-)}{(j + \omega_+ + \omega_-)\Gamma(j + \omega_+ + \omega_-)} \frac{u^j}{j!}, \tag{K13}$$

$$= \sum_j \frac{\Gamma(j + \omega_+)}{\Gamma(\omega_+)} \frac{\Gamma(\omega_+ + \omega_-)}{\Gamma(j + \omega_+ + \omega_-)} \frac{u^j}{j!} \tag{K14}$$

$$= \Phi[\omega_+, \omega_+ + \omega_-; u], \tag{K15}$$

as in Eqn. 129. The marginal $p_n$ is obtained by $p_n = \partial_x^n[G(x)]_0/n!$; using Eqn. K15 and the derivative of the confluent hypergeometric function,

$$\partial_u^n \Phi[\alpha, \beta; u] = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \frac{\Gamma(\beta)}{\Gamma(n+\beta)} \Phi[\alpha + n, \beta + n; u], \tag{K16}$$

one obtains

$$p_n = \frac{g_+^n}{n!} \frac{\Gamma(n + \omega_+)}{\Gamma(\omega_+)} \frac{\Gamma(\omega_+ + \omega_-)}{\Gamma(n + \omega_+ + \omega_-)} \Phi[\omega_+ + n, \omega_+ + \omega_- + n; -g_+], \tag{K17}$$

as in Eqn. 130.

---

[1] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science **297**, 1183 (2002).
[2] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, Nat Genet **31**, 69 (2002).
[3] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, Plos Biol **4**, e309 (2006).
[4] J. M. Pedraza and A. van Oudenaarden, Science **307**, 1965 (2005).
[5] J. S. van Zon and P. R. ten Wolde, Physical Review Letters **94**, 128103 1 (2005).
[6] J. S. van Zon and P. R. ten Wolde, Journal of Chemical Physics **123**, 234910 1 (2005).
[7] R. J. Allen, D. Frenkel, and P. R. ten Wolde, Journal of Chemical Physics **124**, 194111 1 (2006).
[8] C. Valeriani, R. Allen, M. Morelli, D. Frenkel, and P. R. ten Wolde, Journal of Chemical Physics **127**, 114109 (2007).
[9] M. J. Morelli, R. J. Allen, S. Tanase-Nicola, and P. R. ten Wolde, Journal of Chemical Physics **128**, 045105 1 (2008).
[10] B. Munsky and M. Khammash, Journal of Chemical Physics **124**, 044104 (2006).
[11] B. Munsky and M. Khammash, Journal of Computational Physics **226**, 818 (2007).
[12] H. El-Samad, M. Khammash, L. Petzold, and D. Gillespie, Int. J. Robust and Nonlinear Control **15**, 691 (2005).
[13] S. Lampoudi, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **130**, 094104 (2009).
[14] D. T. Gillespie, M. Roh, and L. R. Petzold, J. Chem. Phys. **130**, 174103 (2009).

[15] M. Chevalier and H. El-Samad, Journal of Chemical Physics **131**, 054102 (2009).
[16] F. Didier, T. A. Henzinger, M. Mateescu, and V. Wolf, HiBi09 - High Performance Computational Systems Biology (2009).
[17] N. G. van Kampen, *Stochastic processes in physics and chemistry* (Amsterdam: North-Holland, 1992).
[18] C. Gardiner, *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences* (Springer; 3rd edition, 2004).
[19] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, USA; 1 edition, 2001).
[20] T. B. Kepler and T. C. Elston, Biophys. J. **81**, 3116 (2001).
[21] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Proc Natl Acad Sci USA **99**, 12795 (2002).
[22] J. Hasty, F. Issacs, M. Dolnik, D. McMillen, and J. J. Collins., Chaos **11**, 207 (2001).
[23] J. Hasty, J. Pradines, M. Dolnik, and J. J. Collins, Proc. Nat. Acad. Sci. USA **97**, 2075 (2000).
[24] W. Bialek, Advances in Neural Information Processing 13, TK Leen, TG Dietterich and V Tresp **13**, 103 (2001).
[25] P. Mehta, R. Mukhopadhyay, and N. S. Wingreen, Physical Biology **5**, 026005 (2008).
[26] M. Thattai and A. van Oudenaarden, Proc Natl Acad Sci USA **98**, 8614 (2001).
[27] J. E. M. Hornos, D. Schultz, G. C. P. Innocentini, J. Wang, A. M. Walczak, J. N. Onuchic, and P. G. Wolynes, Phys. Rev. E **72**, 051907 (2005).
[28] A. M. Walczak, A. Mugler, and C. H. Wiggins, Proc Natl Acad Sci USA **106**, 6529 (2009).
[29] A. Mugler, A. M. Walczak, and C. H. Wiggins, Phys. Rev. E **80**, 041921 (2009).
[30] A. Mugler, E. Ziv, I. Nemenman, and C. H. Wiggins, arXiv **q-bio.MN** (2008), 0811.2834v1.
[31] W. Bialek and S. Setayeshgar, Proc Natl Acad Sci USA **102**, 10040 (2005).
[32] G. Tkacik, A. M. Walczak, and W. Bialek, Phys Rev E **80**, 031920 (2009).
[33] A. M. Walczak, G. Tkacik, and W. Bialek, Phys Rev E p. 041905 (2010).
[34] S. Iyer-Biswas, F. Hayot, and C. Jayaprakash, Phys. Rev. E **79**, 31911 (2009).
[35] P. B. Warren, S. Tanase-Nicola, and P. R. ten Wolde, Journal of Chemical Physics **125**, 1449041 (2006).
[36] S. Tănase-Nicola, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **97**, 068102 (2006).
[37] J. Paulsson, Nature **427**, 415 (2004).
[38] J. Elf and M. Ehrenberg, Genome Res. **13**, 2475 (2003).
[39] M. Doi, Journal of Physics A: Mathematical and General **9**, 1465 (1976).
[40] Y. B. Zel'Dovich and A. A. Ovchinnikov, Soviet Journal of Experimental and Theoretical Physics **47**, 829 (1978).
[41] L. Peliti, Journal of Physics A: Mathematical and General **19**, L365 (1986).
[42] J. Sakurai, *Modern quantum mechanics* (Pearson Education India, 1985).
[43] L. Bintu, N. E. Buchler, H. G. Garcia, and et al, Current Opinion in Genetics and Development **15**, 116 (2005).
[44] L. Bintu, N. E. Buchler, H. G. Garcia, and et al., Current Opinion in Genetics and Development **15**, 125 (2005).
[45] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).
[46] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Cell **123**, 1025 (2005).
[47] A. M. Walczak, M. Sasai, and P. Wolynes, Biophys. J. **88**, 828 (2005).
[48] D. T. Gillespie, J Phys Chem **81**, 2340 (1977).
[49] W. Koepf, *Hypergeometric summation: an algorithmic approach to summation and special function identities* (Braunschweig, Germany: Vieweg, 1998).