Stochasticity in Transcriptional Regulation: Origins, Consequences, and Mathematical Representations

Thomas B. Kepler* and Timothy C. Elston†

*Santa Fe Institute, Santa Fe, New Mexico 87501, and †Biomathematics Graduate Program, Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695 USA

ABSTRACT Transcriptional regulation is an inherently noisy process. The origins of this stochastic behavior can be traced to the random transitions among the discrete chemical states of operators that control the transcription rate and to finite number fluctuations in the biochemical reactions for the synthesis and degradation of transcripts. We develop stochastic models to which these random reactions are intrinsic and a series of simpler models derived explicitly from the first as approximations in different parameter regimes. This innate stochasticity can have both a quantitative and qualitative impact on the behavior of gene-regulatory networks. We introduce a natural generalization of deterministic bifurcations for classification of stochastic systems and show that simple noisy genetic switches have rich bifurcation structures; among them, bifurcations driven solely by changing the rate of operator fluctuations even as the underlying deterministic system remains unchanged. We find stochastic bistability where the deterministic equations predict monostability and vice-versa. We derive and solve equations for the mean waiting times for spontaneous transitions between quasistable states in these switches.

INTRODUCTION

The rate at which proteins are synthesized from individual genes is tightly regulated. In prokaryotes, this regulation is accomplished in part by the binding of regulatory proteins to stretches of DNA upstream (by definition) of the protein-coding region of the gene. Regulatory proteins can either inhibit or facilitate the binding of RNA polymerase to DNA or facilitate the isomerization of the DNA–RNA polymerase complex into a transcriptionally competent state. RNA polymerase processes along the DNA, transcribing the DNA into messenger RNA (mRNA). A ribosome can associate with mRNA and begin translation of the mRNA into an amino acid sequence as soon as a complete ribosome binding site emerges from the RNA polymerase. Translation can occur many times per transcript. The canonical introduction to this subject remains the monograph of Ptashne (1992).

In eukaryotes, regulatory proteins, referred to as transcription factors, are also used as one method of controlling gene expression. Another form of eukaryotic regulation occurs through chromatin–DNA interactions. In general, transcriptional regulation in eukaryotes is more complicated than in prokaryotes. However, we believe the results presented in this manuscript are relevant to both types of cells.

In this manuscript, we adopt the terminology used in studying prokaryotes and use "operator" to refer to upstream regulatory DNA sites. The term "promoter" refers to the nucleotide sequence to which RNA polymerase binds to begin transcription. Single genes may have multiple operators that can overlap with the promoter. The operator is said to be in an occupied state if a regulatory protein is bound to it and unoccupied otherwise. Chemical reactions that change the state of the operator are referred to as operator fluctuations. One of the main goals of this manuscript is to understand the role of operator fluctuations in transcriptional regulation. The mathematical methods developed here also directly apply to eukaryotic systems that regulate gene expression through transcription factors or chromatin–DNA interactions.

Much previous mathematical modeling of transcriptional regulation represents the production of gene product as a deterministic process (for a review, see Hasty et al., 2001b). In these models, both the gene-product concentration and operator state are treated as continuous. The probability of an occupied operator is given by a function of the regulatory protein concentration that is computed using thermodynamic arguments (Shea and Ackers, 1985).

In fact, however, there is now considerable experimental evidence that indicates the presence of significant stochasticity in transcriptional regulation in both eukaryotes and prokaryotes. In particular, several recent experiments on mammalian cells have supported the idea that gene initiation in response to an inductive signal is a stochastic process (Weintraub, 1988; van Roon et al., 1989; Fiering et al., 1990; Ko et al., 1990; Dingemanse et al., 1994; Walters et al., 1995). Additionally, there is evidence that chromatin–regulated gene expression is stochastic (Ahmad and Henikoff, 2001; Wijgerde et al., 1995) and that the initiation and deactivation of pigment expression during melanocyte differentiation proceeds in a random fashion (Bennett, 1983). Finally, studies on engineered gene circuits, which have been designed to act as toggle switches and oscillators, have revealed large stochastic effects (Elowitz and Leibler, 2000; Gardner et al., 2000; Becskei et al. 2001).

Several recent papers have reported theoretical investigations into the effects of fluctuations in gene regulation.

Received for publication 14 May 2001 and in final form 5 September 2001.

Address reprint requests to Thomas B. Kepler, Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501. Tel.: 505-984-8800; Fax: 505-982-0565; E-mail: kepler@santafe.edu.

0006-3495/01/12/3116/21 $2.00

Thattai and van Oudenaarden (2001) used simple models of transcription and translation to derive expressions for the means and variances of mRNA and protein number that compare favorably with experimental results. Using Monte Carlo simulations, McAdams and Arkin (1997) studied a more detailed model of these processes that also takes into account ribosome–RNase binding competition. Neither of these two investigations considered fluctuations in the state of the operator. Simple models that take into account fluctuations between active and inactive genetic states have been studied and used to explain induction-level heterogeneity among individual cells expressing steroid-inducible genes (Ko, 1991, 1992), and to support the conjecture that haploinsufficiency disease arises from stochastic gene expression (Cook et al., 1998). In these models, transcription and translation proceed deterministically while the gene is on. Additionally, these investigations relied on computer simulation and did not fully explore the consequences of fluctuations in regulated systems.

Numerical simulations of complete genetic networks have also been carried out. For instance, Arkin et al. (1998) considered a detailed stochastic model for the initial decision between the two developmental pathways (lysis and lysogeny) of bacteriophage λ. However, in this investigation the chemical kinetics of the operator fluctuations was assumed to be fast. This assumption allowed the operator states to be treated deterministically using a quasi-steady-state approximation. The role of noise has also been considered in engineered gene networks (Hasty et al., 2000, 2001b). In this work, fluctuations were added post-hoc to deterministic rate equations, and, therefore, the noise strength was an adjustable parameter.

Our aim here is to consider the origins of intrinsic fluctuations and to develop appropriate representations for them within the context of de novo stochastic models. We further consider a set of simplifying approximations found in various parameter limits, eventually arriving at deterministic models. Because our method starts from a microscopic description of the process, all the parameters in the approximate models are defined in terms of the underlying chemical rate constants. Our approximate schemes are important for two reasons. First, they provide insight into the dynamics of the system that cannot be gained from Monte Carlo simulations of the full model; and, second, numerical simulations using the approximate schemes can run orders of magnitude faster than Monte Carlo simulations of the full process.

In the first two sections and in the appendices, we present the mathematical techniques for manipulating the models. These methods are then used to examine the consequences of fluctuations in regulatory systems: 1) Noise destabilizes genetic switches. We compute the mean first passage times for a simple single-gene switch and examine their behavior under variations of several parameters, particularly the rate of operator transitions. 2) The destabilization of switches makes it necessary to generalize the notion of bifurcations.

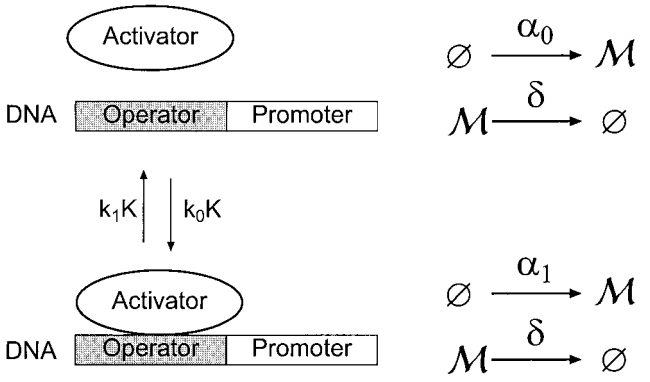

FIGURE 1 A schematic diagram illustrating transcriptional regulation without feedback. The operator has two possible states, occupied and unoccupied, and fluctuates between them. If the operator is empty, protein monomers $\mathcal{M}$ are produced at a rate α_0; if the site is occupied, the production rate is α_1. The operator does not affect the degradation of the protein product, which occurs with rate δ. We assume that the activator concentration is constant.

We examine stochastic bifurcations (Horsthemke and Lefever, 1984), qualitative changes in the probability density function under changes in parameters, again, with special attention to those bifurcations induced solely by changes in the rate of operator transitions.

We start by considering a simple model without feedback to establish our methods, then move on to a switching system consisting of a single self-promoting gene, and finally to a switching system composed of two mutually repressing genes.

SINGLE GENE, NO FEEDBACK

We begin with a model for a gene that has no feedback, direct or indirect, onto its own transcriptional regulation. Similar models have been used to study steroid-inducible genes (Ko, 1991, 1992) and haploinsufficiency disease (Cook et al., 1998). The simplicity of this model allows us to present the techniques we use for the analysis of regulated systems in a direct manner, uncomplicated by nonlinearities. The state space of the model consists of the number of gene product monomers (an integer variable) and the state of the gene's operator (binary). We use the term "gene product" to account for the lumping of mRNA and protein in this treatment.

The events that occur in this model can be represented as biochemical reactions (Fig. 1), and the master equation derived directly from that representation. In this manuscript, we adopt the following notational conventions. Uppercase calligraphic letters denote a molecule of a particular protein species or an operator (e.g., $\mathcal{M}$ for a monomer and $\mathbb{O}$ for an operator). Uppercase letters represent state variables that denote the current number of molecules of a particular protein species or the current chemical state of the operator. Lowercase letters denote

any allowed value of their uppercase counterparts. Because the state variables are pure numbers, all the rate constants have units of inverse time and are reported here as per second. Concentrations are formed by dividing the mean number of a protein species by the relevant volume (e.g., the cell volume). In this case, second-order rates constants would have units of 1/(time × concentration).

The degradation of gene product is written as

$$\mathcal{M} \xrightarrow{\delta} \varnothing, \tag{1}$$

where $\mathcal{M}$ represents the monomer form of the expressed protein, δ is the degradation rate, and $\varnothing$ is used to denote a protein sink. We also use $\varnothing$ to denote a protein source.

We write the spontaneous transitions between operator states, denoted $\mathbb{O}_0$ (occupied) and $\mathbb{O}_1$ (unoccupied), as

$$\mathbb{O}_0 \underset{k_1 K}{\overset{k_0 K}{\rightleftharpoons}} \mathbb{O}_1, \tag{2}$$

where the rate K sets the time scale for this reaction. The constants k_0 and k_1 are dimensionless and constrained to obey $k_0 + k_1 = 1$.

The reaction for the production of protein is written as

$$\varnothing \xrightarrow{\alpha_s} \mathcal{M} \tag{3}$$

where rate α_s ($s = 0$ or 1) depends on the chemical state of the operator. It is important to recognize that this simple reaction is an effective reaction representing a large number of component reactions together making up transcription, translation of mRNA into polypeptides and the folding of these polypeptides into proteins.

In this treatment, we distinguish two sources of stochasticity: operator fluctuations, and the combined action of transcription per se and degradation of protein product. In both cases, the variability results from the inherent discreteness of the state space and the randomness in the dwell time between discrete reactions. (Below, we will consider a further source of variability in the reaction in which gene product monomers interact to form dimers, although our discussion will serve primarily to indicate why these dimer fluctuations can be neglected.)

At any given time t, the state of the system described by reactions 1–3 is specified by the number of monomer proteins $M(t)$ and the chemical state of the operator $S(t)$. $S(t)$ is equal to 0 if the operator is unoccupied and 1 if it is occupied. Thus, the various states of the system can be written as the ordered pair (m, s), where m is a non-negative integer and s is either 0 or 1. $M(t)$ and $S(t)$ are random variables, because the chemical reactions that change the state of the system occur randomly in time. If we assume that the dwell time in any particular chemical state of the system is exponentially distributed, then the system satisfies the Markov property. Physically, this means that the time evolution of the system is determined solely by its current state and is independent of its past. The Markov property allows us to write down a master equation for the time evolution of the probabilities $p^s_m(t) = \Pr[M(t) = m \text{ and } S(t) = s]$ (van Kampen, 1992). Simply put, the master equation is a rate equation for $p^s_m(t)$. Written out explicitly, the master equation for this process has the form

$$\frac{dp^0_m}{dt} = -(Kk_0 + \delta m + \alpha_0)p^0_m + Kk_1 p^1_m + \delta(m+1)p^0_{m+1} + \alpha_0 p^0_{m-1}, \tag{4}$$

$$\frac{dp^1_m}{dt} = -(Kk_1 + \delta m + \alpha_1)p^1_m + Kk_0 p^0_m + \delta(m+1)p^1_{m+1} + \alpha_1 p^1_{m-1}, \tag{5}$$

where we have suppressed the explicit time dependence of $p^s_m(t)$ for readability. The first term on the right-hand side of Eqs. 4 and 5 represents the rate at which probability is flowing out of the state (m, s). It is the product of the average rate, $(Kk_s + \delta m + \alpha_s)$, at which transitions occur out of (m, s) and the probability $p^s_m(t)$ of being in (m, s). Correspondingly, the other three terms on the right-hand side of these equations represent the rates at which probability flows into (m, s) from accessible states.

Although the notation used in Eqs. 4 and 5 makes these equations easy to interpret, it quickly becomes unmanageable as the systems become more complex. Therefore, we will adopt a more compact notation and combine these equations as

$$\frac{dp^s_m}{dt} = \alpha_s(p^s_{m-1} - p^s_m) + \delta[(m+1)p^s_{m+1} - mp^s_m] + K(k_{\hat{s}}p^{\hat{s}}_m - k_s p^s_m), \tag{6}$$

where the hatted index indicates "the other one": $\hat{s} = (s + 1) \bmod 1$.

The partial moments, defined for integer j, are

$$\langle m^j \rangle_s \equiv \sum_m m^j p^s_m. \tag{7}$$

Note that the zeroth moments $\langle m^0 \rangle_s$ are the marginal probabilities for the operator to be in state s. The moments are then the sums over operator states of the partial moments,

$$\langle m^j \rangle = \langle m^j \rangle_0 + \langle m^j \rangle_1. \tag{8}$$

We can use Eq. 6 to derive ordinary differential equations (ODEs) for the time evolution of the partial moments (van Kampen, 1992). Because the reactions are all zeroth- and first-order, the ODEs for the partial moments factorize into independent pairs of linear equations, which can easily be

solved. Here we are interested only in the steady-state values of the first moment and the variance,

$$m_* \equiv \overline{\langle m \rangle} = \frac{1}{\delta}(\alpha_0 k_1 + \alpha_1 k_0) \tag{9}$$

and

$$\overline{\mathrm{Var_m}} = m_* + k_0 k_1 \left(\frac{\alpha_0 - \alpha_1}{\delta}\right)^2 \frac{\delta}{\delta + K}, \tag{10}$$

where the overbar indicates the steady-state value. Under the given circumstances, we expect the first factor in the product on the right side of Eq. 10 to be of order one, the second factor to be of order m_*^2 and the order of the third to depend on the relative rates of product decay, δ and operator transitions K. When product decay is much faster than operator transitions, the last factor is of order one and the variance is dominated by the second term, of order m_*^2. When the operator transitions are much faster, the last factor is very small, and the variance approaches the mean.

An appropriate measure of the relative size of the fluctuations is the coefficient of variation, which is defined as the ratio of the variance to the mean squared. The steady-state coefficient of variation can be written as

$$\mathrm{CV} \equiv \frac{\overline{\mathrm{Var_m}}}{m_*^2} = \frac{1}{m_*} + k_0 k_1 \frac{\delta}{\delta + K}\left[\frac{(\alpha_0 - \alpha_1)}{\alpha_0 k_1 + \alpha_1 k_0}\right]^2. \tag{11}$$

Note that, as the average number of monomers increases, the first term in the CV decreases. Even with large protein numbers, however, the CV can still be large due to fluctuations in the operator state. As these operator fluctuations become faster, the second term decreases as well. This stabilizing effect of fast operator fluctuations has been observed in computer simulations of stochastic gene expression (Cook et al., 1998). Below, we construct approximations to the dynamics for cases in which the product number is large or the operator fluctuations are fast.

Small-noise and fast-transition approximations to the master equation

For cases in which there is feedback regulation, we generalize Eq. 6 to include state-dependent transition rates and nonlinearities. When this is done, the moment equations no longer factorize. Numerical solutions can be difficult to obtain as well, especially when several genes are considered as part of a regulatory network. Therefore, we develop approximations to Eq. 6 that can be generalized immediately to the nonlinear cases of feedback regulation and are valid as one or the other source of variability becomes negligible.

Large steady-state gene-product level

When the protein abundance given by Eq. 9 is large compared to one, we can use a diffusion approximation for Eq. 6. In the section, Escape Times, we present an example in which this diffusion approximation is accurate with m_* as small as 25. We start by defining an appropriately scaled continuous variable for the monomer number. This variable is given by

$$X \equiv \frac{M}{m_*}. \tag{12}$$

As defined here, X is dimensionless, though we shall often refer to it as a concentration. To convert to an actual concentration, in Eq. 12 replace m_* by the cell volume in appropriate units. We define the probability density function $\rho_s(x, t)$ such that

$$p_m^s(t) \equiv \int_{(m-1/2)/m^*}^{(m+1/2)/m^*} dx \rho_s(x, t). \tag{13}$$

One way to arrive at the diffusion approximation is to note that, by appropriately rearranging terms in Eq. 6, this equation can be recast in a form that is identical to a second-order finite differencing of the diffusion equation. However, it is possible to quickly arrive at the same result through use of a Taylor series expansion. The Taylor series for a function $g(x)$ about x is

$$g(x + u) = \sum_j \frac{1}{j!}(\partial_x)^j g(x) u^j = e^{u\partial_x} g(x), \tag{14}$$

where ∂_x is shorthand for the partial derivative with respect to x. The last equality in the above expression defines the shift operator $e^{u\partial_x}$, which translates g from x to $x + u$.

Changing variables and using the shift operator in Eq. 6, we get the equations of motion for ρ_s,

$$\begin{aligned}\partial_t \rho_s(x) = {} & \alpha_s(e^{-(1/m^*)\partial_x} - 1)\rho_s(x) \\ & + \delta m_*(e^{+(1/m^*)\partial_x} - 1)x\rho_s(x) \\ & + K[k_{\hat{s}}\rho_{\hat{s}}(x) - k_s\rho_s(x)].\end{aligned} \tag{15}$$

The Taylor series that defines the shift operator can be truncated without incurring large errors when the size of the translation is sufficiently small. In our case, if $1/m_*$ is small enough, we can neglect terms of third order and higher to obtain the diffusion approximation,

$$\begin{aligned}\partial_t \rho_s(x) = {} & -\partial_x\left[\left(\frac{\alpha_s}{m_*} - \partial x\right)\rho_s(x)\right] \\ & + \frac{1}{2m_*}\partial_x^2\left[\left(\frac{\alpha_s}{m_*} + \delta x\right)\rho_s(x)\right] \\ & + K[k_{\hat{s}}\rho_{\hat{s}}(x) - k_s\rho_s(x)].\end{aligned} \tag{16}$$

In the limit, as $m_* \to \infty$, the only stochasticity remaining is that due to the operator fluctuations. The master equation in this limit becomes

$$\partial_t \rho_s(x) = -\partial_x \left[\left(\frac{\alpha_s}{m_*} - \delta x \right) \rho_s(x) \right] + K[k_{\bar{s}} \rho_{\bar{s}}(x) - k_s \rho_s(x)], \tag{17}$$

where the term α_s/m_* is order δ. The above equation can be interpreted in the following way. In each state of the operator, the concentration evolves deterministically according to the equation,

$$\frac{dx}{dt} = \frac{\alpha_s}{m_*} - \delta x. \tag{18}$$

However, the effective rate at which protein is made α_s/m_* fluctuates randomly in time between high ($s = 1$) and low ($s = 0$) levels. A generalization of this system, allowing for multiple operators and operator states is discussed in detail in Appendix B.

Fast operator fluctuations

The variance due to operator transitions decreases as the rate of these transitions increases and their characteristic time becomes much smaller than those of the rest of the system; these fluctuations are effectively averaged out over the longer time scales (Cook et al., 1998). We take advantage of this effect to construct an approximation to the dynamics that is accurate when the fluctuations in the operator state are fast, but finite.

To apply these methods, we re-express Eq. 6 in terms of the marginal probability function $p_m \equiv p_m^0 + p_m^1$ and the difference $\xi_m = k_0 p_m^0 - k_1 p_m^1$,

$$\begin{aligned} \frac{dp_m}{dt} &= (\alpha_0 k_1 + \alpha_1 k_0)(p_{m-1} - p_m) \\ &\quad + \delta[(m+1)p_{m+1} - m p_m] \\ &\quad + (\alpha_0 - \alpha_1)(\xi_{m-1} - \xi_m), \end{aligned} \tag{19}$$

$$\begin{aligned} \frac{d\xi_m}{dt} &= -K\xi_m + (k_0\alpha_0 + k_1\alpha_1)(k_0 - k_1)(\xi_{m-1} - \xi_m) \\ &\quad + \delta(k_0 - k_1)^2[(m+1)\xi_{m+1} - m\xi_m] \\ &\quad + k_0 k_1 (\alpha_1 - \alpha_0)(p_{m-1} - p_m) \end{aligned} \tag{20}$$

When K is very large compared to α_1 and δm_*, ξ reaches a rapid quasi-equilibrium for any value of p. This is realized mathematically by setting the derivative in Eq. 20 equal to zero. The resulting expression for ξ is then substituted back into Eq. 19 with the result being the approximate equation of motion for p_m,

$$\begin{aligned} \frac{dp_m}{dt} &= (\alpha_0 k_1 + \alpha_1 k_0)(p_{m-1} - p_m) \\ &\quad + \delta\{(m+1)p_{m+1} - m p_m\} \\ &\quad + \frac{1}{K} k_0 k_1 (\alpha_0 - \alpha_1)^2 (p_{m-2} - 2p_{m-1} + p_m). \end{aligned} \tag{21}$$

The last term is a second-order finite difference centered on $m - 1$ (rather than on m). It acts as a diffusion term, producing the same dynamics for the mean and variance as the "usual" finite difference centered on m. Although the higher moments differ in these two cases, as m_* increases, these difference are multiplied by steadily decreasing factors and vanish as $m_* \to \infty$. As $K \to \infty$, we obtain a simple Poisson process with degradation, with the instantaneous rate of transcription equal to the equilibrium average $\alpha_0 k_1 + \alpha_1 k_0$.

Simultaneous limits

We can apply these approximations simultaneously. The diffusion approximation applied to Eq. 21 gives the same result as the fast-noise approximation applied to Eq. 16. Taking the marginal density $\rho = \rho_0 + \rho_1$ on x as the dynamical object (see Appendix B for more details), we find

$$\partial_t \rho(x) = -\partial_x [A(x)\rho(x)] + \tfrac{1}{2} \partial_x^2 B(x) \rho(x), \tag{22}$$

where

$$A(x) = \delta(1 - x) + \frac{1}{K}, \tag{23}$$

$$B(x) = 2\frac{1}{K} k_0 k_1 (\alpha_0 - \alpha_1)^2 + \frac{1}{m_*} \delta(1 + x). \tag{24}$$

When the approximations that led to Eq. 22 are appropriate, one advantage of this formulation is that an expression for the steady-state density in terms of a simple quadrature can be found,

$$\bar{\rho}(x) = \frac{\lambda}{B(x)} \exp\left(\frac{1}{2} \int_0^x \frac{A(x')}{B(x')} dx' \right), \tag{25}$$

where the overbar is again used to indicate steady state, and λ is a normalization constant. Another advantage is that sample paths of the process described by Eq. 22 can be

generated from the stochastic differential equation

$$\frac{dX}{dt} = A(X) + \sqrt{B(X)}\zeta(t), \tag{26}$$

where $\zeta(t)$ is a Gaussian white noise process. (We are using the Ito interpretation of the stochastic integral. There is no ambiguity with this choice, because Eq. 22 defines the stochastic process. We choose the Ito interpretation because it is straightforward to implement numerically.) Sample paths generated from this equation can run orders of magnitude faster than Monte Carlo simulations of the full process.

In the limit, as both m_* and K become infinite, the noise term in Eq. 26 goes to zero, and this equation becomes the deterministic rate equation,

$$\frac{dx}{dt} = A(x) = \delta(1 - x). \tag{27}$$

For the simple example we are considering (and indeed for all linear systems), the above expression is identical to the equation for the first moment derived directly from the underlying master equation. For nonlinear systems, the two equations differ. However, from Eq. 26, we can perform a small-noise expansion to find

$$\frac{d\langle X\rangle}{dt} = \langle A(X)\rangle = A(\langle X\rangle) + \frac{1}{2}\frac{d^2A}{dx^2}(\langle X\rangle)\mathrm{Var}(X) + \cdots. \tag{28}$$

So the simple deterministic rate equation is appropriate when $A''(\langle X\rangle)\mathrm{Var}(X)$ is much smaller than $A(\langle X\rangle)$.

REGULATED SYSTEMS I: SELF-PROMOTER

We now consider a system in which the gene product is itself an activating regulatory protein for its gene. This system is regulated by positive feedback. This type of regulation is thought to play a role in the developmental decision pathway of bacteriophage λ (Ptashne, 1992) and has been used to construct a synthetic eukaryotic gene switch in *Saccharomyces cerevisiae* (Becskei et al., 2001). Here we study a minimal version of a system with positive feedback, which serves to illustrate the qualitative features of this type of regulation. For a more biologically complete treatment of the lysis/lysogeny decision pathway, the reader is referred to the pioneering work of Arkin et al. (1998). The system we consider is similar to that shown in Fig. 1, only now the activator is the gene's own product. After developing the models for this system and exploring the consequences of noise for bifurcations and spontaneous transitions in this system, we will come back to explore a second bipolar switch involving a pair of mutually repressing proteins.

Regulatory proteins often bind to their operators as dimers or higher-order oligomers (Ptashne, 1992). We assume that the active form of the protein in our system is a dimer. Thus, we explicitly consider the dimerization reaction and the stochasticity associated with it.

Letting $\mathcal{D}$ represent a protein dimer, the dimerization reaction is

$$\mathcal{M} + \mathcal{M} \underset{\theta\Lambda}{\overset{\Lambda}{\rightleftharpoons}} \mathcal{D}, \tag{29}$$

where θ is the dimensionless equilibrium dissociation constant, and Λ is the forward rate constant.

The operator transition reaction, Eq. 2, is now modified to explicitly include the role of the protein dimer

$$\mathbb{O}_0 + \mathcal{D} \underset{\beta K}{\overset{K}{\rightleftharpoons}} \mathbb{O}_1, \tag{30}$$

where K is now the forward transition rate and β is the dimensionless dissociation constant for this reaction. The other reactions remain as given above, except that we must specify the rate at which dimers are degraded. For simplicity, we have made the arbitrary choice that dimers are not degraded at all. That is, only the monomer form of the protein is unstable. Inclusion of an arbitrary dimer degradation is easily accommodated. Our choice does not have qualitative implications for this analysis within the range of parameter variation considered.

Let $D(t)$ represent the number of dimers and $N(t)$ the total protein number at time t. Then we have $M(t) = N(t) - 2D(t)$. The master equation for this process is

$$\begin{aligned}\frac{dp^s_{n,d}}{dt} = {} & \delta[(n + 1 - 2d)p^s_{n+1,d} - (n - 2d)p^s_{n,d}] \\ & + \alpha_s[p^s_{n-1,d} - p^s_{n,d}] \\ & + \Lambda[(n - 2d + 2)(n - 2d + 1)p^s_{n,d-1} \\ & - (n - 2d)(n - 2d - 1)p^s_{n,d}] \\ & + \theta\Lambda[(d + 1)p^s_{n,d+1} - dp^s_{n,d}] \\ & + (-1)^s K(\beta p^1_{n,d} - dp^0_{n,d}),\end{aligned} \tag{31}$$

where $p^s_{n,d} = \Pr[N(t) = n, D(t) = d, \text{and } S(t) = s]$. It is straightforward to generate sample paths of the "full" process described by the equations given above (see Fig. 6). These Monte Carlo simulations are usually computationally intensive, because the time scales of the various reactions involved can be very different. Therefore, we make use of the limiting cases discussed above to construct approximations to the master equation that are less computationally intensive and provide insight into the dynamics of the system.

The reaction given by Eq. 29 is generally assumed to be fast compared to all other reactions (McAdams and Arkin, 1997; Hasty et al., 2000, 2001a). Physically, this means that the monomer and dimer concentrations come to quasi-equilibrium before the total amount of protein changes appreciably. In Appendix A, we show how this assumption, along with the assumption that the coefficient of variation of D conditional on N is small, allows us to eliminate the dimer concentration from the problem. In this limit, we get the master equation, written in terms of the marginal density on M (In an abuse of notation, we let the marginal on M be represented as $p^{\mathrm{s}}_{\mathrm{m}}$ and rely on the subscript itself to distinguish this probability distribution from that of N).

$$\frac{\mathrm{d}p^{\mathrm{s}}_{\mathrm{m}}}{\mathrm{d}t} = \delta[(m+1)p^{\mathrm{s}}_{\mathrm{m}+1} - mp^{\mathrm{s}}_{\mathrm{m}}] + \alpha_{\mathrm{s}}[p^{\mathrm{s}}_{\mathrm{m}-1} - p^{\mathrm{s}}_{\mathrm{m}}] + (-1)^{\mathrm{s}}K\left(\beta p^{1}_{\mathrm{m}} - \frac{m(m-1)}{\theta}p^{0}_{\mathrm{m}}\right). \tag{32}$$

To arrive at this simple set of equations, it is necessary to assume that the dissociation constant for dimerization is large. Although this is not necessarily the regime of biological interest, it serves to simplify the presentation considerably and is the regime commonly used, usually implicitly, by other researchers. As discussed in detail in Appendix A, all the analyses presented below can be carried through when this assumption is relaxed; the qualitative nature of the results do not change. A direct comparison with experimental data, however, requires the more general treatment presented in Appendix A.

To move to the diffusion limit, we change to the dimensionless variables $u = \delta t$ and $X = M/m_{\mathrm{o}}$, where $m_{\mathrm{o}} = \alpha_1/\delta$ is the steady-state value of the mean monomer number for the system locked in the occupied state. In terms of these variables, the diffusion limit for large protein number produces

$$\partial_{\mathrm{u}}\rho_{\mathrm{s}}(x) = -\partial_{\mathrm{x}}\left[(a_{\mathrm{s}} - x)\rho_{\mathrm{s}}(x) - \frac{1}{2m_{\mathrm{o}}}\partial_{\mathrm{x}}[a_{\mathrm{s}} + x]\rho_{\mathrm{s}}(x)\right] + (-1)^{\mathrm{s}}\kappa[b\rho_1(x) - x^2\rho_0(x)], \tag{33}$$

where the rescaled parameters are $a_1 = 1$, $a_0 = \alpha_0/\alpha_1$, $\kappa = K\alpha_1^2/(\theta\delta^3)$ and $b = \beta\theta\delta^2/\alpha_1^2$.

As $m_{\mathrm{o}} \to \infty$ the fluctuations in the monomer concentration become negligible; Eq. 33 loses its diffusive terms and can be written as

$$\partial_{\mathrm{u}}\rho_{\mathrm{s}}(x) = -\partial_{\mathrm{x}}(a_{\mathrm{s}} - x)\rho_{\mathrm{s}}(x) + (-1)^{\mathrm{s}}\kappa[b\rho_1(x) - x^2\rho_0(x)]. \tag{34}$$

All of the stochasticity in this system now derives from the operator fluctuations. The key difference between the above equation and Eq. 17 is the appearance of the x^2 factor multiplying $\rho_0(x)$ in Eq. 34. As will be seen, this factor is responsible for the bistable behavior observed in the macroscopic limit of this system.

One useful feature of Eq. 34 is that an explicit expression for the steady-state marginal density $\bar{\rho} = \bar{\rho}_0 + \bar{\rho}_1$ can be found,

$$\bar{\rho}(x) = \lambda \exp\left\{\kappa\left(a_0 x + \frac{x^2}{2}\right)\right\} \cdot (x - a_0)^{\kappa a_0^2 - 1}(1 - x)^{\kappa b - 1}, \tag{35}$$

where λ is a normalization constant. Below, we discuss how this distribution undergoes bifurcations as a result of the fluctuations in the operator state.

Next, we consider the small-noise limit of the fluctuations in the monomer concentration and the fast-noise limit of fluctuations in the operator state. That is, both m_{o} and κ are taken to be large, but finite. In Appendix B, we present a general algorithm for deriving an effective diffusion equation for the marginal density, and find

$$\partial_{\mathrm{u}}\rho(x) = -\partial_{\mathrm{x}}A(x)\rho(x) + \tfrac{1}{2}\partial_{\mathrm{x}}^2 B(x)\rho(x), \tag{36}$$

where

$$A(x) = \frac{ba_0 + x^2}{b + x^2} - x - \frac{2xb(a_0 - 1)[((a_0 - 2) + x)x^2 + b(x - a_0)]}{\kappa(b + x^2)^4} \tag{37}$$

$$B(x) = \frac{1}{m_{\mathrm{o}}}\left(\frac{b(a_0 + x) + x^2(1 + x)}{b + x^2}\right) + \frac{bx^2(a_0 - 1)^2}{\kappa(b + x^2)^3}. \tag{38}$$

The steady-state solution to Eq. 36 is given by Eq. 25, using the above expressions for A and B.

Finally, in the deterministic limit $\kappa, m_{\mathrm{o}} \to \infty$, we are left with the ODE

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{ba_0 + x^2}{b + x^2} - x = -\partial_{\mathrm{x}}\phi(x), \tag{39}$$

where we have introduced the potential

$$\phi(x) = \frac{x^2}{2} - x - (a_0 - 1)\sqrt{b}\arctan\left(\frac{x}{\sqrt{b}}\right). \tag{40}$$

Loosely speaking, $\phi(x)$ can be thought of as an effective free energy for the system. Its local minima represent stable steady states of the concentration. The local maxima are energetic barriers that must be surmounted by thermal activation. In general, such a free energy function does not exist for nonequilibrium systems, as is the case for the mutual repressor system considered below.

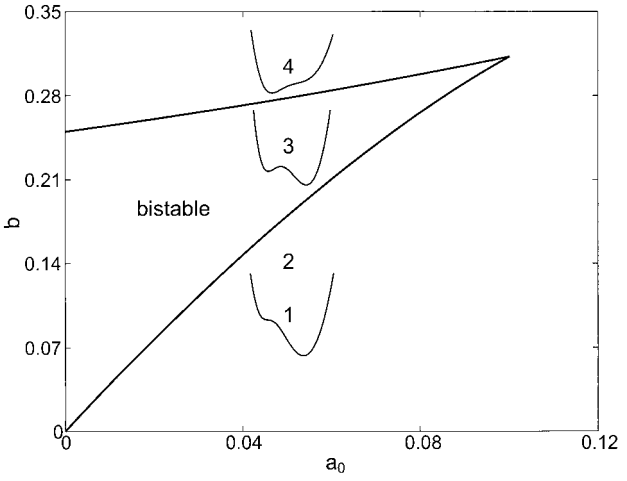

FIGURE 2 The deterministic bifurcation diagram for the self-promoter. The dimensionless parameters are defined as $b = \theta\beta\delta^2/\alpha_1^2$ and $a_0 = \alpha_0/\alpha_1$. The points 1, 3, and 4 indicate the parameter values used to produce the potentials shown in the figure. Points 1 and 2 specify the systems investigated in the Monte Carlo simulations shown in Fig. 4, both of which are monostable.

BIFURCATIONS

The deterministic system given by Eq. 39 acts as a switch in the appropriate parameter regime: the system has two distinct stable fixed points (for a discussion of the necessary conditions required to make a biological switch, see Cherry and Adler, 2000). Figure 2 shows the bifurcation diagram for this system as a function of the two parameters b and a_0. The corresponding potentials $\phi(x)$ are also drawn on the diagram to illustrate the number of stable fixed points in each region. In the region where there are two stable fixed points, the system acts as a genetic switch.

For stochastic systems, the notion of "stable fixed point" in the state space is not well defined. To generalize the idea of bifurcations for applicability to our stochastic models in such a way as to be consistent with the usual meaning as the fluctuations become vanishingly small, we focus our attention on the gain and loss of critical points in the probability density function (Horsthemke and Lefever, 1984). So, for example, a bifurcation in which a single stable fixed point becomes a pair of stable fixed points and a single unstable point corresponds to the transformation of a unimodal probability density function to one that is bimodal.

We now illustrate how fluctuations in the operator state change the dynamics of the system. In particular, these fluctuations can either induce bistability in regions that are deterministically (i.e., in the zero-noise limit) monostable or wash out regions of bistability. Figure 3 *A* is a bifurcation diagram for the stationary distribution given by Eq. 35 as a function of b and a_0. For this figure $\kappa = 50$. The dashed curve corresponds to the deterministic bifurcation diagram shown in Fig. 2. The various regions of this graph are labeled with the numbers 1–6. In Fig. 3 *B*, qualitative features of the steady-state distributions for each region of Fig. 3 *A* are shown. As the vertical line shown in Fig. 3 *A* is crossed from left to right, an integrable singularity occurs at the lower boundary of the distribution. Likewise, as the horizontal line is crossed from top to bottom, an integrable singularity occurs at the upper boundary of the distribution. The points labeled *1* and *2* in Fig. 2 are monostable in the deterministic limit. In Fig. 3 *B*, however, we see that finite operator fluctuations induce a type of bistability. The stationary distribution has a local maximum at high concentrations and the distribution is singular at $x = a_0$. Note that there is small region near the cusp of the deterministic bifurcation curve where the deterministic system predicts bistability, but the fluctuations cause this behavior to be lost. Not surprisingly, the Monte Carlo simulations presented below reveal that fluctuations in the monomer concentration can also wash out bistable behavior. In region *3*, the distribution is bimodal. This region grows and shifts as κ is increased until it coincides with the deterministic limit (*dashed curve*). This effect can be seen in Fig. 3 *C*, which is a bifurcation diagram for the steady-state distribution as a function of b and κ with $a_0 = 0.05$. This value of a_0 is used in the Monte Carlo simulations discussed next.

Figure 4 shows the results of Monte Carlo simulations consistent with Eq. 32. The parameter values in the upper pair of panels correspond to the point labeled *1* in Fig. 2. In the deterministic limit, the system is monostable. From Fig. 3 *A*, however, we see that the finite operator fluctuations have induced bistability. This can be seen from the dashed curve in the top right panel of Fig. 4, representing the steady-state distribution when monomer fluctuations are ignored. In the left panel, a typical time series for the process is shown. There is no indication of bistability. The solid curve shown in the top right panel is the steady-state distribution when the small-and-fast noise approximations are used (i.e., Eqs. 36 and 25). By comparing the dashed and solid curves, we see that, even with this large value of m_o, monomer fluctuations can still be detected. Indeed, these fluctuations are responsible for washing out the bistable nature of the system. We find excellent agreement between this distribution and the histogram from the Monte Carlo simulations.

Comparison of the histograms reveals that bistability can be induced by varying either δ or K. Experimentally, δ could be altered, for example, by the addition of a protease. Comparing the middle panels to Fig. 3 *C*, we expect that this system will be bistable, and, indeed, the time series shown in the middle left panel reveals that this is the case, although low monomer levels are very unlikely. The lower state can be further stabilized by increasing δ. In the bottom panel, K has been reduced while δ is unchanged from the top panel. The bistability of the system is evident. Note that, although both approximations do a pretty good job of reproducing the steady-state distributions, there is some discrepancy at low

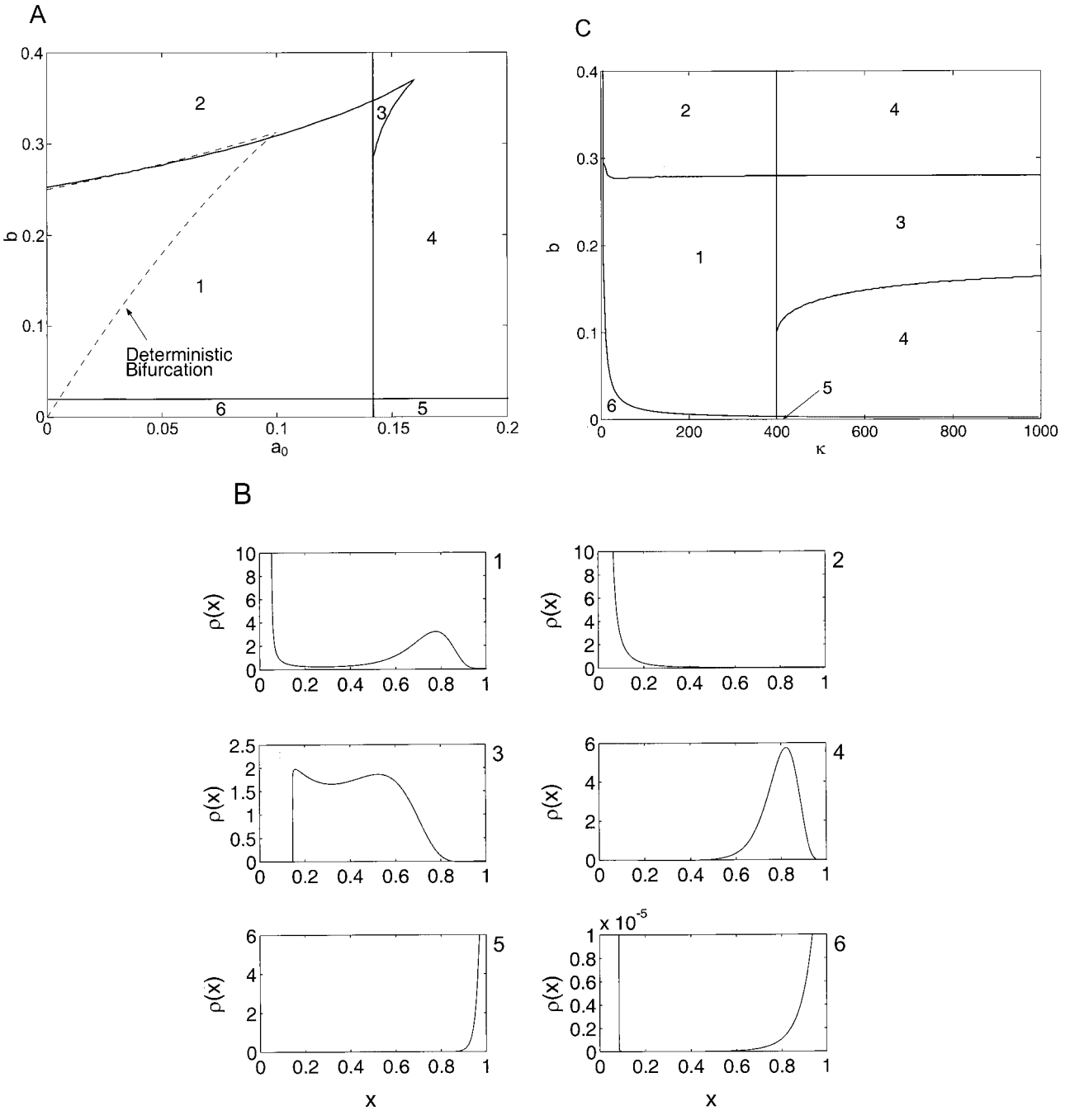

FIGURE 3 (*A*) The bifurcation diagram for Eq. 35 with $\kappa = 50$. The regions numbered 1–6 correspond to qualitatively different steady-state distributions. The dashed curve is the bifurcation diagram in the deterministic limit shown in Fig. 2. (*B*) The steady-state distributions associated with the regions numbered in (*A*) and (*C*). (*C*) The bifurcation diagram for b versus κ with $a_0 = 0.05$.

concentrations. This is exactly where we expect the approximations to break down.

Figure 5 is the same as Fig. 4 except that, instead of performing Monte Carlo simulations of the discrete process, we have generated sample paths based on the stochastic differential equation associated with Eq. 36. This method runs roughly an order of magnitude faster than the discrete Monte Carlo simulations. As can be seen, the agreement between the diffusion approximation and the full simulation is good, and it looks as if the diffusion approximation is faithfully capturing the dynamics as well as the steady-state distribution. We expand on this point in our discussion of the mean first passage time below.

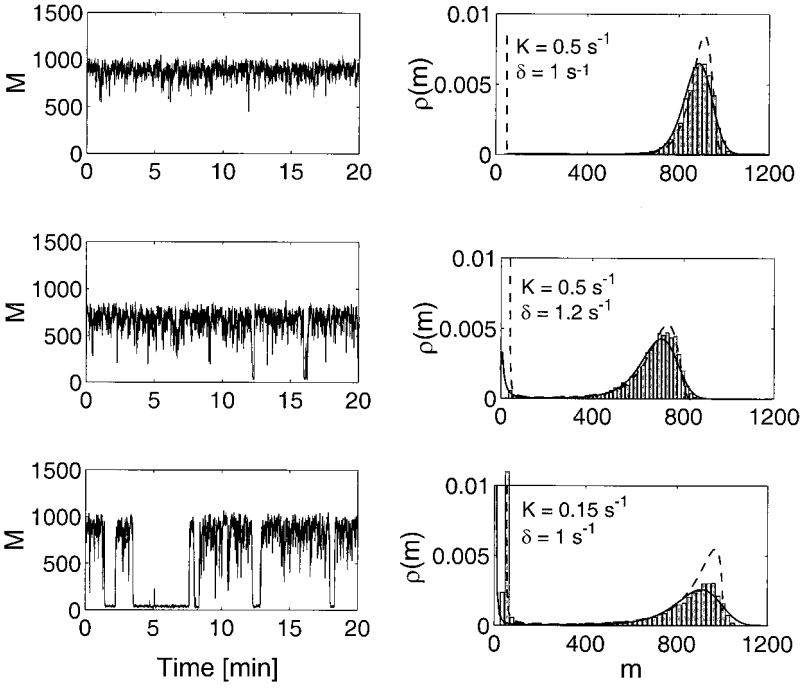

FIGURE 4 Sample paths and distributions for the self promoter. The dimer concentration has been eliminated using the quasiequilibrium approximation (Eq. 32). In all the panels, $\theta = 10{,}000$, $\alpha_1 = 1000\ s^{-1}$, $\alpha_0 = 50\ s^{-1}$, and $\beta = 10$ ($a_0 = 0.05$). Each histogram corresponds to the time series on its left. Together, they illustrate the bifurcations that occur as δ or K are varied. The bifurcation that occurs as K is varied is due solely to fluctuations in the operator state and does not occur in the macroscopic limit. The distributions shown as solid lines are the results of the small-and-fast noise approximation. The dashed curves are the steady-state distributions given by Eq. 35.

Finally, we illustrate the validity of ignoring the inherent dimer concentration fluctuations. Figure 6 *A* shows sample paths for $M(t)$ and $D(t)$ generated by Monte Carlo simulation of the process described by Eq. 31. In the deterministic limit, the system is bistable and this is indeed evident in the time series. The intrinsic fluctua-

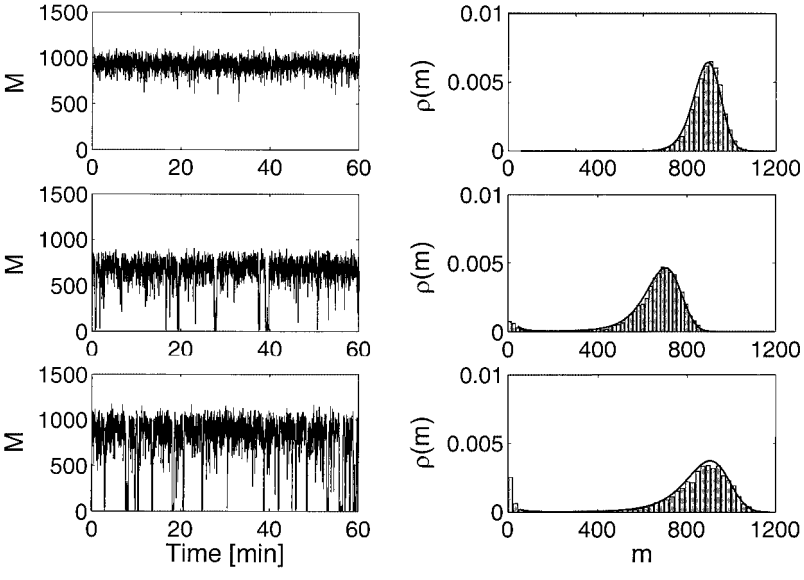

FIGURE 5 Sample paths of the stochastic differential equation associated with the small-and-fast noise approximation and corresponding steady-state distributions. The solid curves are the same as in Fig. 4.

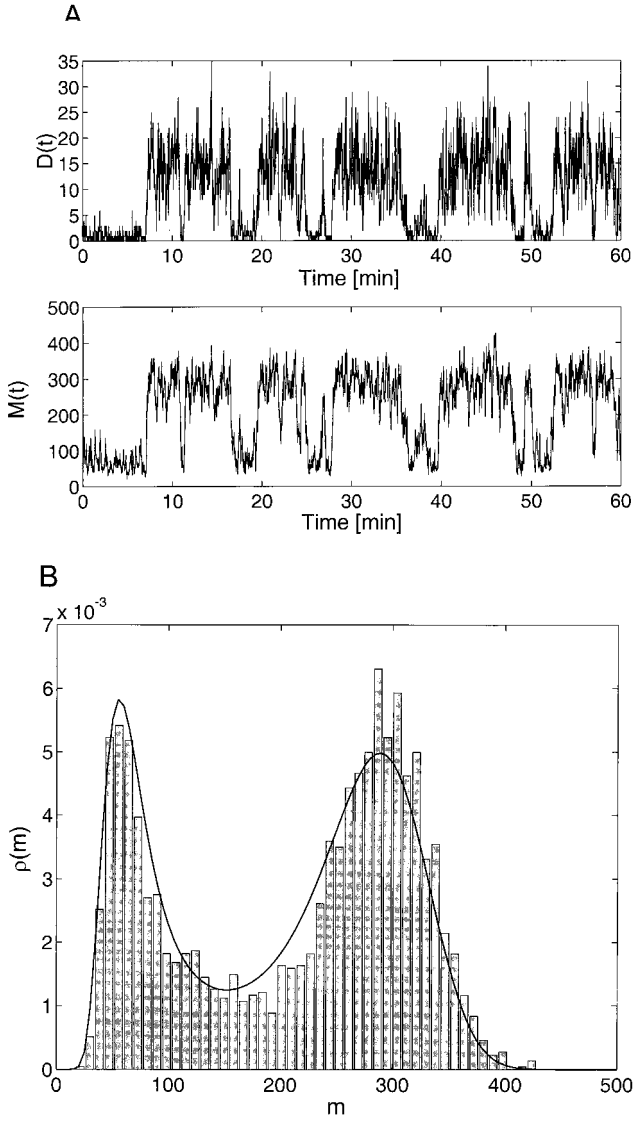

FIGURE 6 (*A*) Sample paths from Monte Carlo simulations that include independently both dimer and monomer number (i.e., the processes described by Eq. 31) with $\Lambda = 0.1\ s^{-1}$, $\theta = 6000$, $k_0 = 12\ s^{-1}$, $\beta = 11.7$, $\alpha_1 = 500\ s^{-1}$, $\alpha_0 = 40\ s^{-1}$, and $\delta = 1$ ($m_o = 500$, $\kappa = 500$, b = 0.28 and $a_0 = 0.08$). In the macroscopic limit, the system is bistable, as is evident in the time series. (*B*) The steady-state monomer distribution. The histogram was generated from the time series shown in (*A*). The solid curve is the steady-state distribution using the small-and-fast noise approximation together with the quasiequilibrium approximation for the dimer number.

tions, however, allow for transitions between the high and low protein levels. Figure 6 *B* shows a comparison between the corresponding histogram and the steady-state distributions found using simultaneous application of the small-and-fast noise and quasi-equilibrium approximations. Note the excellent agreement between the numerical and analytical results.

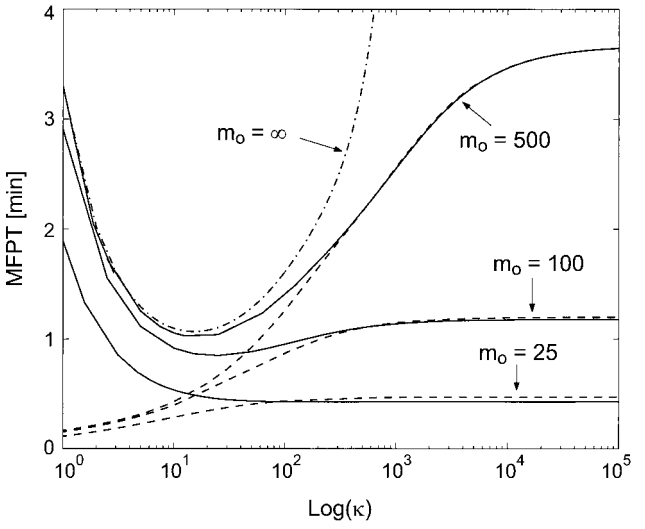

FIGURE 7 The MFPT as a function of log(κ). The solid lines are the results for the discrete process Eq. 32 and the dashed lines are the small-and-fast noise approximation. The dot-dashed line is the limit m_0 goes to infinity. To produce this curve, the process described by Eq. 34 was used. The parameters used to produce this figure were $\delta = 1\ s^{-1}$ and $\theta = 6000$. The parameters α_1, α_0, and b were adjusted to vary m_o with $b = 0.28$ and $a_0 = 0.08$ fixed. The system is deterministically bistable.

ESCAPE TIMES

When the system is stochastically bistable, a quantity of interest is the time between switches from low to high concentration and vice versa. This time is a random variable and is often referred to as the first-passage time. Here we present results for the mean first-passage time (MFPT). In Appendix D, we present the mathematical details needed to compute the MFPT. For illustration, we consider cases in which the corresponding deterministic system is bistable. We let the initial dimensionless concentration X_0 equal its lower bound, a_0, and compute the average time for the concentration to reach 0.65, which is close to the high concentration steady state of Eq. 39. It is assumed that, initially, the probabilities for the operator states take their equilibrium values. Figure 7 is a plot of the MFPT versus log κ for various values of m_o. The solid curves shown in this figure are the MFPT as calculated from the discrete process Eq. 32. The dot-dashed curve is the limiting case in which $m_o \to \infty$. In this limit, the MFPT goes to infinity as κ is increased, because there are no fluctuations to induce switching. This approximation is valid for situations in which the operator fluctuations are slow and the total concentration is large. In this limit, the first-passage time again becomes large due to the long times spent in the unoccupied state. The dashed curves are the approximations in which both operator and concentration fluctuations are treated using the small noise approximation. This approximation is poor when κ is small. Surprisingly, with sufficiently large κ, the approximation is good when m_o is even as small as 25. The fact that the approximations accurately reproduce the MFPT indicates fur-

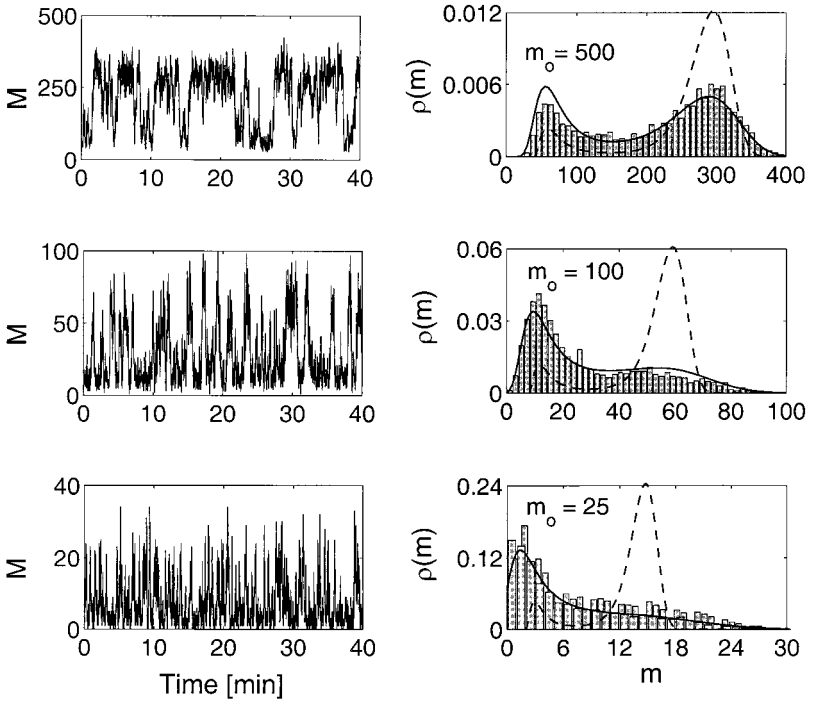

FIGURE 8 Time series and steady-state distributions for the various cases illustrated in Fig. 7 with $\kappa = 500$. The time series and histograms show the results of Monte Carlo simulation of the discrete process described by Eq. 32. The solid lines are the steady-state distribution obtained using the small-and-fast noise approximation. The dashed lines show the steady-state distribution in the limit $m_o \to \infty$ (Eq. 34). The transition to bistability for increasing m_o is evident.

ther that they are accurately depicting the dynamics and the steady-state distribution of the system.

Figure 8 shows the time series and the steady-state distributions for the three vales of m_o shown in Fig. 7 with $\kappa = 500$. The time series represent Monte Carlo simulations of the discrete process ignoring intrinsic dimer fluctuations. When $m_o = 25$, the bistable nature of the system is washed out by the fluctuations in monomer concentration. Again, we see that the small noise approximation accurately captures the steady-state distribution (*solid curve*). As m_o is increased, bistability becomes apparent, and, for very large m_o (data not shown), the small noise distribution becomes indistinguishable from the distribution that ignores intrinsic concentration fluctuations, but explicitly includes operator-induced fluctuations (*dashed curve*).

We can also see the emergence of bistability if we plot the MFPT verses the initial concentration X_0 (Fig. 9). The emergence of a step in MFPT at large values of m_o indicates a potential barrier at around $X_0 = 0.3$ that fluctuations must surmount.

REGULATED SYSTEMS II: MUTUAL REPRESSORS

Another type of switch is formed when two proteins, $\mathcal{M}_1$ and $\mathcal{M}_2$, act as mutual repressors; each binds to the operator of the other and represses its transcription. Recently, such a regulatory network has been engineered and shown to act as

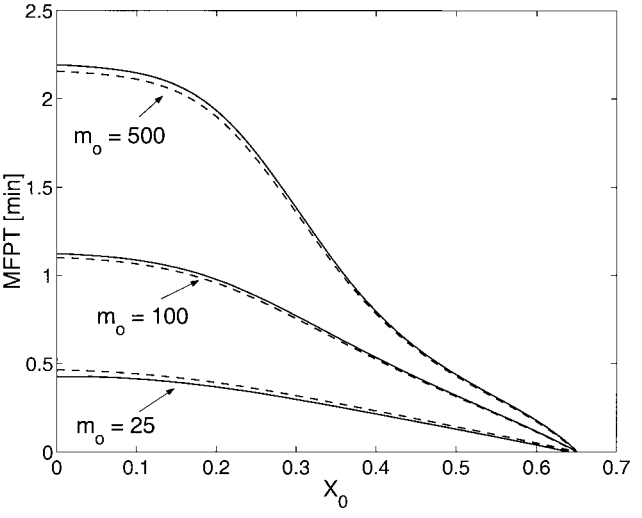

FIGURE 9 The MFPT as a function of the initial concentration X_0. The parameters used to produce this plot are the same as in Fig. 8. As before, the solid lines show the results for the discrete process and the dashed lines represent the small-and-fast noise approximation. With this value of κ, the approximation is a good even when $m_o = 25$.

a toggle switch (Gardner et al., 2000). To illustrate the mathematical techniques and highlight the main features of this system, we make several simplifying assumptions. Even for this oversimplified system, the effective diffusion equation for the system is quite complicated and the dynamics, nontrivial.

Our first simplifying assumption is that the two genes that code for $\mathcal{M}_1$ and $\mathcal{M}_2$ share the same operator. Although our motivation is simplification, similar arrangements do occur in nature (Ptashne, 1992). This assumption reduces the number of operator states from 4 to 3: $s = 0$ (empty), $s = 1$ (occupied by $\mathcal{M}_1$), and $s = 2$ (occupied by $\mathcal{M}_2$). Again we assume that the proteins bind to the operator as dimers, $\mathcal{D}_1$ and $\mathcal{D}_2$. For simplicity, we assume that the proteins differ only in the gene they repress—all other biophysical parameters are identical. Furthermore, we consider only the limiting case in which the dimerization reaction is fast as compared to all other processes. As before, we assume that the dimeric form of the protein is stable and the dissociation constant θ is large.

We assume that, if $\mathcal{M}_1$ is bound to the operator, then the production rate of $\mathcal{M}_2$ is $\alpha_{0,2} = 0$, and vice versa. If there is no repressor bound both proteins are produced at rate $\alpha_{1,i} \equiv \alpha_1$, where $i = 1$ or 2. All these considerations lead to the following set of biochemical reactions

$$\mathcal{M}_i \xrightarrow{\delta} \varnothing, \tag{41}$$

$$\varnothing \xrightarrow{\alpha_s^i} \mathcal{M}_i, \tag{42}$$

$$\mathcal{M}_i + \mathcal{M}_i \underset{\theta\Lambda}{\overset{\Lambda}{\rightleftharpoons}} \mathcal{D}_i, \tag{43}$$

$$\mathcal{O}_0 + \mathcal{D}_i \underset{\beta K}{\overset{K}{\rightleftharpoons}} \mathcal{O}_i. \tag{44}$$

It is straightforward to write down the master equation for the reaction scheme given above and use the quasi-equilibrium approximation to eliminate the dimer numbers. The resulting master equation for the monomer abundances is not enlightening and will not be presented here.

As is shown in Appendix C, the diffusion equation in the small noise limit has the form

$$\begin{aligned}\partial_u \rho(x_1, x_2, s) = &-\partial_{x_1} A_1(x_1, x_2)\rho - \partial_{x_2} A_2(x_1, x_2)\rho \\ &+ \tfrac{1}{2}\partial_{x_1}^2 B_1(x_1, x_2) + \tfrac{1}{2}\partial_{x_2}^2 B_2(x_1, x_2) \\ &- \partial_{x_1}\partial_{x_2} B_{12}(x_1, x_2)\rho.\end{aligned} \tag{45}$$

The explicit forms of A and B are given in Appendix C. The occurrence of a cross term with a negative coefficient in the above equation indicates that the fluctuations in the two concentrations are anticorrelated, as expected, given the inhibitory interactions between the two protein species. In the deterministic limit, the ODEs for the dimensionless concentrations are

$$\frac{dx_1}{dt} = \frac{1}{1 + x_2^2/(b + x_1^2)} - x_1, \tag{46}$$

$$\frac{dx_2}{dt} = \frac{1}{1 + x_1^2/(b + x_2^2)} - x_2, \tag{47}$$

where, again, $b = \theta\beta\delta^2/\alpha_1^2$.

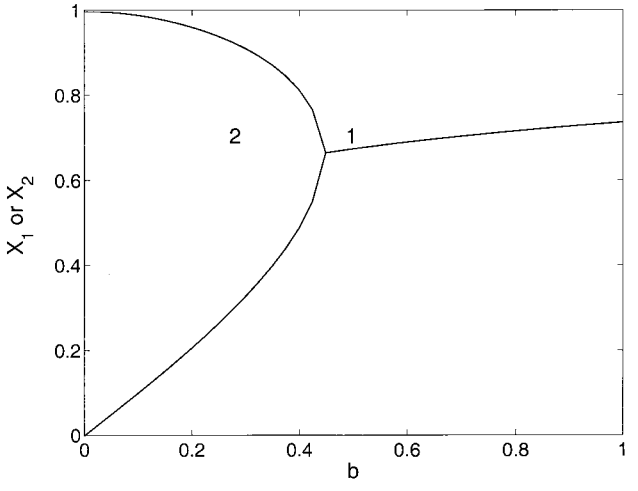

FIGURE 10 The bifurcation diagram for the mutual repressor. The solid curves indicate the stable fixed points of the system. The system undergoes a bifurcation at $b = 4/9$. The point marked *1* indicates the value of b used in the top and bottom panels of Figs. 11 and 12, and the point marked *2* indicates the value of b used in the middle panel of these figures.

Results for the mutual repressor

The bifurcation diagram for Eqs. 46 and 47 (Fig. 10) shows that, when $b < 4/9$, there are two stable fixed points. In the absence of a simple form for the steady-state distribution, we examine the effectiveness of the approximations by direct comparison of realizations of the discrete process to the continuous diffusion approximation (i.e., realizations from the stochastic differential equations corresponding to Eq. 45).

Figure 11 shows time series and histograms for Monte Carlo simulations of the discrete process. The top three panels correspond to the point marked *1* in Fig. 10. In the deterministic limit, the system is monostable. The values of m_o and κ are 1000 and 50, respectively; we expect the small noise approximation to be valid. In the middle three panels, δ has been reduced. They correspond to the point marked *2* in Fig. 10; we expect the system to be bistable. This behavior is clearly seen in the time series and histogram. Finally, in the bottom three panels, the same parameter values are

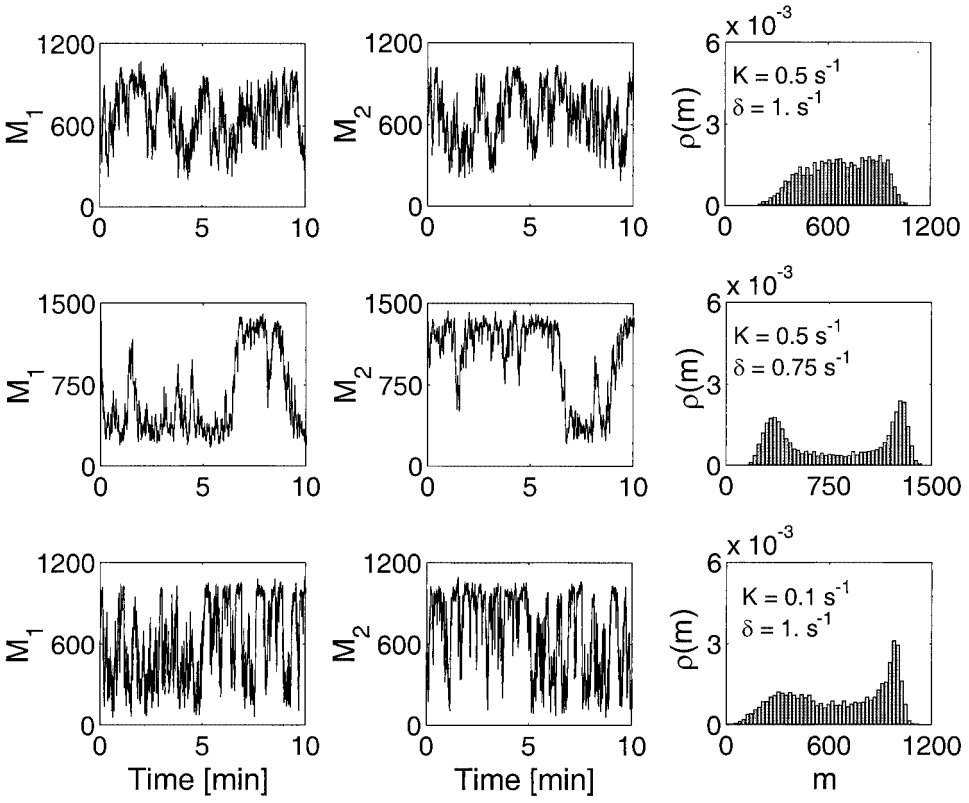

FIGURE 11 Monte Carlo simulations for the mutual repressor with $\theta = 10{,}000$, $\beta = 50$, $\alpha_1 = 1000\ s^{-1}$. In the top three panels, $k_0 = 0.5\ s^{-1}$, $\delta = 1\ s^{-1}$, and $b = 0.5$. The dimensionless parameters are: (*upper*) $b = 0.5$ and $\kappa = 50$, (*middle*) $b = 0.28$ and $\kappa = 118.5$, and (*lower*) $b = 0.5$ and $\kappa = 10$. Note that, in the lower panels, κ and κb are not large, so we expect fluctuations in the operator state to have a significant effect. The dimer concentration has been approximated using the quasi-steady-state approximation.

used as in the top panels except that K has been reduced one-hundred-fold. The deterministic description predicts that the system is monostable. However, as can be seen from the figure with finite fluctuations, the system is bistable. Due to the small value of κ and κb used in the bottom panel, we may not be justified in using the diffusion approximation for this case.

Figure 12 illustrates results obtained using the small-and-fast noise approximations. That is, the time series shown in the figure were generated using the stochastic differential equations associated with Eq. 45. Very good agreement is seen between the top and middle panels of this figure and Fig. 11, verifying the validity of these approximations. Some discrepancies are noticeable, however, in the bottom panels of the two figures where κb is only 5. To accurately capture the dynamics of the system with this value of κ, the fluctuations in the operator state must be explicitly included in the model. That is, Eqs. C1–C3 should be used. When this is done, there is good are agreement for all three cases (data not shown).

DISCUSSION

Genetic regulation is a topic of central importance in biology. With the advent of new techniques for the simultaneous determination of expression levels of tens of thousands of genes, many of its key issues are likely to be dealt with comprehensively in the next several years. Given the extraordinary quantities of data that will be necessary to accomplish these goals and the inherent complexity of the systems involved, it is inevitable that these gains will require significant use of novel mathematical and statistical tools. Furthermore, the nature of transcription—small transcript numbers and discrete operator states—dictates that stochasticity be explicitly treated and understood in the basic models. This contention is buttressed by the existence of several macroscopic gene-regulatory phenomena in which stochastic effects play a major role (Weintraub, 1988; van Roon et al., 1989; Fiering et al., 1990; Dingemanse et al., 1994; Walters et al., 1995; Wijgerde et al., 1995; Ahmad and Henikoff, 2001).

We are particularly interested in examining the gene-product concentration variability due to internal fluctuations in the discrete states of the operator, because, to our knowledge, a theoretical treatment of these fluctuations does not exist. However, computer simulations of simple models of inducible gene expression have been studied (Ko, 1991; Cook et al., 1998). We are not aware of any estimates of reaction rates for operator fluctuations and therefore cannot

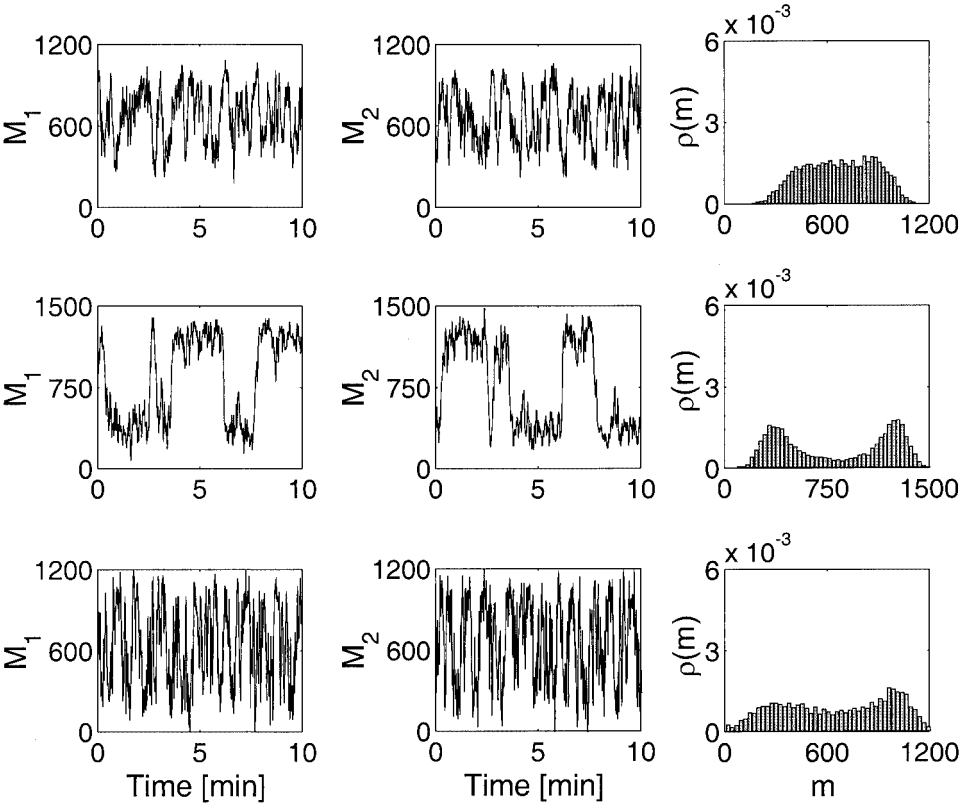

FIGURE 12 Sample paths of the stochastic differential equations associated with the small-and-fast noise approximation of the mutual repressor. Good agreement between this figure and Fig. 11 is evident in the upper and middle panels, through differences are clearly visible in the lower panels. These disparities arise because the fluctuations in the operator state are not fast enough to warrant the fast noise approximation.

say in advance just how large the associated effects will be. We have derived expressions for the gene-product variance attributable to these operator fluctuations that could help to estimate these effective rates, or conversely, given these effective rates, estimate the size of the associated variance.

Even the more complicated of the models discussed here leave out potentially important features. The most egregious such omission is that of the distinction between transcription and translation. In other words, the models, read literally, are models of direct translation from DNA into protein. This simplification clearly has significant impact on phenomena in some systems. The artificial three-gene cycling construct (Elowitz and Leibler, 2000), for example, would not oscillate at all were it not for the delay between transcription and translation.

Other researchers have developed models that treat transcription and translation separately. In a model of the tryptophan operon, Santillan and Mackey (2001a,b) used delay differential equations to take into account time delays associated with these two processes, whereas others have constructed stochastic models of transcription and translation that explicitly account for delays in these processes (McAdams and Arkin, 1997; Thattai and van Oudenaarden, 2001). In all these investigations, operator fluctuations were ignored. One aim of this manuscript is to understand the role of these fluctuations in transcriptional regulation. This is the motivation for simplifying transcription and translation into a single kinetic step. A serious concern with this assumption is that our model allows a finite probability for the instantaneous production of protein. We do not expect including explicit models of transcription and translation to affect the qualitative features of our results. However, to fully understand the combined effects of all the relevant processes requires further investigation.

We have attempted here to strike a balance between those models that are based on discrete-object simulation (Endy and Brent, 2001; McAdams and Arkin 1998) and those that are derived directly in terms of macroscopic state variables and either neglect randomness or add it by hand (Shea and Ackers, 1985; Hasty et al., 2001a,b; Hasty 2000; Santillan and Mackey, 2001a,b). Both these approaches are useful and have provided insight into genetic networks. However, the former can be very difficult to analyze, or even understand adequately, and becomes computationally intractable for large networks, whereas the latter may fail to represent the phenomena quantitatively, or indeed, as we show here, qualitatively.

We have presented a set of stochastic models of differing levels of temporal and, effectively, spatial resolution, derived in various parameter limits. We have used these models to explore some of the basic consequences of stochasticity in transcriptional regulation in two simple models exhibiting stable switching behavior in the deterministic limit. Because our approach starts from a microscopic description, all the macroscopic parameters are defined in terms of the underlying chemical processes. Even for situations in which intrinsic fluctuations are negligible, it is important to derive the macroscopic rate equations directly from the underlying master equation, because phenomenological treatments do not reliable capture the dynamics of the system (cf. Eq. A12).

In the presence of noise, the switches are destabilized, and, correspondingly, the bifurcation diagrams that provide insight into the nature of the switches must be generalized. We have used the appearance of critical points in the steady-state probability density function (rather than of singular points of the deterministic dynamics) to characterize the qualitative behavior of these systems, which have rich bifurcation structures, including bifurcations associated with changes in the operator fluctuation rates alone. In other words, the qualitative behavior of these switches changes when the characteristic time for operator fluctuations is assumed zero, the limit in which the deterministic rate equations are derived.

The stochasticity of these switches causes spontaneous transitions to occur. In the context of our models, we can compute the mean first-passage times. We find that these mean waiting times increase exponentially with the operator fluctuation characteristic rate. Thus, we might expect to find that evolution has tuned these rates to just over the value necessary to prevent spontaneous transitions within the lifetime of the cell, except in the case where spontaneous transitions are an intrinsic component of the functional behavior of the system.

Beyond the results reported here, our primary concern is that the effects of intrinsic noise may actually become increasingly important as more and more components are assembled into the large regulatory networks that clearly comprise the basic apparatus of the cell. Thus far, we have only examined the case of very small networks, and do not now have any sense for how the effects in question scale with network size.

There surely are sources of variability we have not yet considered (fluctuations in cell volume, ionic environment, DNA accessibility, etc.) that may have profound consequences for genetic regulation. There are several DNA chemical modifications, such as methylation (Yeivin and Razin, 1993) and acetylation (Grunstein, 1997) unique to multicellular eukaryotes that result in longer term, more stable regulatory changes. Even here, we expect that the initiation of these changes has great variability from cell to cell. We are anxious to learn, among other things, the role of stochasticity coupled to more macroscopic intercellular processes in driving the extraordinary development of a complete organism from a single cell.

APPENDIX A: QUASI-EQUILIBRIUM DIMER FLUCTUATIONS

In this appendix, we show how the quasi-equilibrium assumption can be used to eliminate the dimer number as a state variable from the problem. This procedure depends on two characteristics of the system: First, that the coefficient of variation of the dimer number conditional on the total protein concentration is small; Second, that the rate constants for the dimerization reactions are fast compared to other rates in the system.

We write the joint probability function as the product of the marginal in n and the conditional of d on n.

$$p^{s}_{n,d} = p^{s}_{n} p^{s}_{d|n}. \tag{A1}$$

Summing Eq. 31 over d gives the differential equation for p^{s}_{n},

$$\frac{dp^{s}_{n}}{dt} = \delta[(n + 1 - 2\langle d|n+1\rangle_{s})p^{s}_{n+1} - (n - 2\langle d|n\rangle_{s})p^{s}_{n}] + \alpha_{s}(p^{s}_{n-1} - p^{s}_{n}) + (-1)^{s}K(\beta p^{1}_{n} - \langle d|n\rangle_{s} p^{0}_{n}). \tag{A2}$$

Substitution of Eq. A1 into Eq. 31 yields that for $p^{s}_{d|n}$. This latter equation is somewhat complicated and not particularly transparent. The dominant term in this equation, however, is the dimer-related probability flux, or "dimer flux," which, for both operator states, s is given by

$$j^{s}_{n,d}(t) = \Lambda[(n - 2d + 2)(n - 2d + 1)p^{s}_{d-1|n} - \theta d p^{s}_{d|n}]p^{s}_{n}. \tag{A3}$$

As Λ tends toward infinity, this flux must remain finite for all d: quasi-equilibrium corresponds to setting the expression inside the square brackets to zero. Equations for the conditional moments can then be found by multiplying this expression by d^{q} and summing over d. The result for $q = 1$ is

$$\theta\overline{\langle d|n\rangle} = (n - 2\overline{\langle d|n\rangle})(n - 2\overline{\langle d|n\rangle} - 1) + 4\overline{\mathrm{Var}[d|n]}. \tag{A4}$$

When the coefficient of variation of d conditional on n is small (we will determine shortly when it will be small), the variance term can be neglected, and the mean is given simply by

$$\overline{\langle d|n\rangle} = \frac{1}{\theta} m(n)(m(n) - 1), \tag{A5}$$

where $m(n) = n - 2\overline{\langle d|n\rangle}$. Under the quasi-steady-state approximation, Eq. A2 becomes

$$\frac{dp^{s}_{n}}{dt} = \delta\{(n + 1 - 2\overline{\langle d|n+1\rangle})p^{s}_{n+1} - (n - 2\overline{\langle d|n\rangle})p^{s}_{n}\} + \alpha_{s}(p^{s}_{n-1} - p^{s}_{n}) + (-1)^{s}K(\beta p^{1}_{n} - \overline{\langle d|n\rangle} p^{0}_{n}), \tag{A6}$$

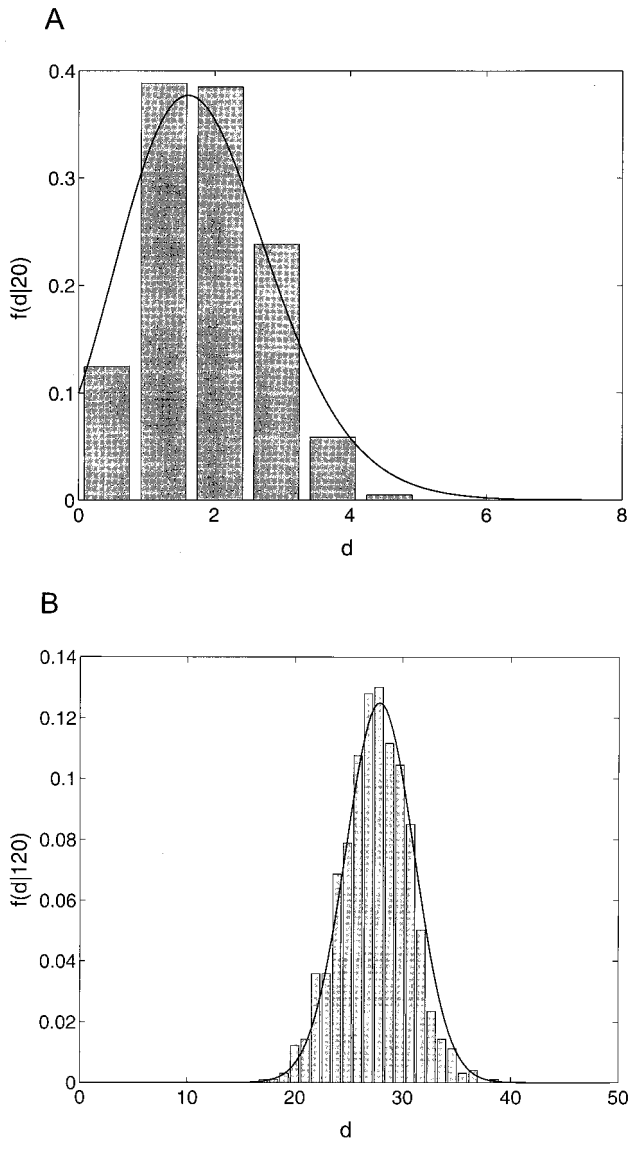

FIGURE A1 The equilibrium dimer distribution with $\theta = 150$: (*A*) $N = 20$ and (*B*) $N = 120$. The histograms are the results of Monte Carlo simulations and the solid lines are the diffusion approximation, which treats the number of dimers as a continuous variable.

where $\overline{\langle d|n\rangle}$ is given by Eq. A4. For the above equation to be useful, however, we need the functional dependence of $\overline{\langle d|n\rangle}$ on n. This relation can be found easily if the coefficient of variation is small, in which case Eq. A5 can be used to solve for $\overline{\langle d|n\rangle}$ in terms of n. If, additionally, the dissociation constant is large, then $M \approx N$. Making this change of variable in Eq. A6 produces Eq. 32 of the text.

We can approximate the coefficient of variation under the assumption that the third central moment is small compared to the mean cubed, and find

$$\overline{CV\,[d|n]} = \frac{1}{\overline{\langle d|n\rangle}}\,\frac{1}{1 + 4\overline{\langle d|n\rangle}} \tag{A7}$$

giving a condition for the self-consistency of the approximation. If the dimer number is large enough to warrant the diffusion approximation, then an analytic expression for the equilibrium dimer probability density for fixed N can be found and used to examine this assumption explicitly. The equilibrium probability density is given by

$$\log \overline{\rho_{D|n}}(d|n) = \lambda + 2d + \theta(\theta - 4n)(8\theta n - \theta^2)^{-1/2} \times \arctan\left(\frac{8d + \theta - 4n}{\sqrt{8\theta n - \theta^2}}\right) - \left(1 + \frac{\theta}{2}\right)\log(4^2 m_0^2 + d(\theta - 4n) + n^2), \tag{A8}$$

where λ is a normalization constant. Figure A1, *A* and *B*, show plots of this distribution with $\theta = 150$. In Fig. A1, *A*, $N = 20$. With this value, the average number of dimers is between 1 and 2. Surprisingly, the continuum limit works relatively well even for this small number of dimers. In Fig. A1, *B*, $N = 120$. For this case, the mean and the variance are 27.5 and 10.0, respectively. Note that, even with this small number of dimers, the distribution looks nearly Gaussian and the coefficient of variation ($\sim$0.01) is small.

For a large gene-product pool, we can use the diffusion approximation to Eq. A6. To do this, we convert to the dimensionless variables $u = \delta t$ and $Y = N/m_o$. Then the marginal density for $X = M/m_o$ is found by using the change of variables $Y = X + 2X^2/\theta$. The result is

$$\partial_u \rho_s = -\partial_x\left(\left[\frac{\theta(a_s - x)}{\theta + 4xm_o} - \frac{2\theta^2(a_s + x)}{(\theta + 4xm_o)^3}\right]\rho_s - \frac{1}{2m_0}\,\partial_x \frac{\theta^2(a_s + x)}{(\theta + 4xm_o)^2}\,\rho_s\right) + (-1)^s \kappa(b\rho_1 - x^2\rho_0). \tag{A9}$$

In the limit $m_0 \to \infty$, the above equation becomes

$$\partial_u \rho_s = -\partial_x \frac{\theta(a_s - x)}{\theta + 4xm_o}\,\rho_s + (-1)^s \kappa(b\rho_1 - x^2\rho_0). \tag{A10}$$

The steady-state marginal density $\bar{\rho} = \bar{\rho}_0 + \bar{\rho}_1$ for the above equation is

$$\log \bar{\rho}(x) = \lambda + \kappa\left(\frac{x^2}{2} + a_0 x\right) + \frac{2}{3}\frac{\kappa m_o x}{\theta}(6b + 6a_0^2 + 3a_0 x + 2x^2) + \log\left(1 + \frac{4m_o x}{\theta}\right) + \left(b\kappa - 1 + \frac{4m_o b\kappa}{\theta}\right)\log(1 - x) + \left(\kappa a_0^2 - 1 + \frac{4a_0^3 m_o \kappa}{\theta}\right)\log(x - a_0), \tag{A11}$$

which reduces to Eq. 35 in the limit $\theta \gg m_0$. Next, taking the limit $\kappa \to$

∞ in Eq. A10 results in the Liouville equation, equivalent to the ODE,

$$\frac{dx}{dt} = \frac{1}{1 + 4m_o x/\theta}\left(\frac{ba_0 + x^2}{b + x^2} - x\right), \qquad \text{(A12)}$$

which reduces to Eq. 39 in the limit $\theta \gg m_o$. Note that the prefactor multiplying the right-hand side of Eq. A12 does not affect the fixed points of the system. It does, however, play a significant role in the dynamics.

APPENDIX B: EFFECTIVE DIFFUSION EQUATION FROM OPERATOR FLUCTUATIONS: GENERAL CASE

We often expect that the fluctuations in the operator state occur on a faster time scale than the rate of production and degradation of protein. In this limit, it is possible to derive an effective diffusion equation for the marginal probability density. We start from the small noise approximation for the protein concentrations and make use of the dimensionless variables discussed in the text. Let the elements of the q-dimensional vector $\mathbf{X}(t)$ denote the dimensionless concentrations of the q protein species involved in the process at time t. The single time density for the operator state and protein concentrations is denote by $\rho_i(\mathbf{x}, t)$. Therefore, $\boldsymbol{\rho}$ is a c-dimensional vector, where c is the number of chemical states of the operator. The single-time densities satisfy the equation

$$\partial_t \boldsymbol{\rho} = \mathbf{L}(x)\boldsymbol{\rho} + \mathbf{K}\boldsymbol{\rho}, \qquad \text{(B1)}$$

where $\mathbf{K}$ is the $c \times c$ transition matrix that contains the reaction rates for transitions between chemical states of the operator. The diagonal matrix operator $\mathbf{L}$ has the form

$$L_{ii} = -\sum_{j=1}^{q} \partial_{x_j} g_{ji}(\mathbf{x}) + \frac{1}{2}\sum_{j=1}^{q} \partial_{x_j}^2 h_{ji}(\mathbf{x}). \qquad \text{(B2)}$$

The matrix $\mathbf{g}$ is a $q \times c$ matrix, whose jth column contains the net production rates of the q protein species when the operator is in the jth chemical state. The matrix $\mathbf{h}$ is likewise a $q \times c$ matrix. The columns of $\mathbf{h}$ are the diffusion coefficients for each protein species in that particular chemical state of the operator. To make explicit our assumption that the chemical kinetics of the operator are fast, we scale $\mathbf{K}$ as $\mathbf{K}/\epsilon$ and write

$$\partial_t \boldsymbol{\rho} = \mathbf{L}(x)\boldsymbol{\rho} + \frac{1}{\epsilon}\mathbf{K}\boldsymbol{\rho}. \qquad \text{(B3)}$$

Now, because probability is conserved, the matrix $\mathbf{K}(\mathbf{x})$ must have one zero eigenvalue for all values of $\mathbf{x}$. We assume that $\mathbf{K}$ has exactly one zero eigenvalue (all the rest must be negative) at each point. Furthermore, the left eigenvector corresponding to the eigenvalue zero is the row vector $\mathbf{1}^T \equiv (1, 1, \ldots, 1)$; i.e.,

$$\mathbf{1}^T\mathbf{K} = 0. \qquad \text{(B4)}$$

We designate the corresponding right eigenvector $\mathbf{r}_0(\mathbf{x})$ and normalize it to satisfy $\mathbf{1}^T\mathbf{r}_0 = 1$. Thus, the elements of $\mathbf{r}_0$ are the steady-state probabilities of the chemical states of the operator for fixed $\mathbf{x}$. The projection operator,

$$\boldsymbol{\pi} \equiv \mathbf{I} - \mathbf{r}_0\mathbf{1}^T, \qquad \text{(B5)}$$

projects out the dynamics of the system that does not lie in the null space of $\mathbf{K}$. The marginal density for X is $f \equiv \mathbf{1}^T\boldsymbol{\rho}$. The joint density $\boldsymbol{\rho}$ can be decomposed as

$$\boldsymbol{\rho} = \boldsymbol{\xi} + f\mathbf{r}_0, \qquad \text{(B6)}$$

where $\boldsymbol{\xi} = \boldsymbol{\pi}\boldsymbol{\rho}$. In terms of this decomposition, Eq. B3 becomes

$$\partial_t f = \mathbf{1}^T\mathbf{L}(\boldsymbol{\xi} + f\mathbf{r}_0), \qquad \text{(B7)}$$

$$\partial_t \boldsymbol{\xi} = \boldsymbol{\pi}\mathbf{L}(\boldsymbol{\xi} + f\mathbf{r}_0) + \frac{1}{\epsilon}\mathbf{K}\boldsymbol{\xi}, \qquad \text{(B8)}$$

where the last term in the second equation takes account of the fact that $\boldsymbol{\pi}\mathbf{K} = \mathbf{K}$.

We next make the quasi-equilibrium approximation. That is, we assume that the probabilities for the chemical states of the operator reach their steady-state values, before X changes appreciably. This amounts to setting the left-hand side of Eq. B8 equal to zero. Remembering that, by construction, $\boldsymbol{\xi}$ does not have a component that lies in the null space of $\mathbf{K}$, we have

$$\boldsymbol{\xi} = -\epsilon\mathbf{K}^+\mathbf{L}f\mathbf{r}_0 + O(\epsilon^2), \qquad \text{(B9)}$$

where $\mathbf{K}^+$ is a pseudo-inverse of $\mathbf{K}$ defined by

$$\mathbf{K}^+\mathbf{K} = \mathbf{K}\mathbf{K}^+ = \boldsymbol{\pi} \quad \text{and} \quad \mathbf{K}^+\boldsymbol{\pi} = \mathbf{K}^+. \qquad \text{(B10)}$$

An explicit formula for $\mathbf{K}^+$ is

$$\mathbf{K}^+ = \mathbf{E}\boldsymbol{\Lambda}^*\mathbf{E}^{-1}, \qquad \text{(B11)}$$

where $\mathbf{E}$ is the matrix whose columns are the right eigenvectors of $\mathbf{K}$, and $\boldsymbol{\Lambda}^*$ is the diagonal matrix whose entries on the diagonal are the inverses of the eigenvalues of $\mathbf{K}$, except that the entry corresponding to the null eigenvalue is itself zero. One can see that the matrix thus defined satisfies Eq. B10. We then substitute Eq. B9 into Eq. B7 to give the diffusion equation for the marginal density,

$$\partial_t f = \mathbf{1}^T\mathbf{L}f\mathbf{r}_0 - \epsilon\mathbf{1}^T\mathbf{L}\mathbf{K}^+\mathbf{L}f\mathbf{r}_0. \qquad \text{(B12)}$$

EXAMPLE Two chemical states (self-promoter). As an example of the algorithm described above, we treat the case of one protein species and two operator states, a special case of which is the self-promoter discussed in the text.

Let $k_0 = -K_{11} = K_{21}$ and $k_1 = -K_{22} = K_{12}$. In this case, the matrices $\mathbf{g}$ and $\mathbf{h}$ are 1×2. Let $g_{11} = g_0$ and $g_{12} = g_1$. Likewise, let $h_{11} = h_0$ and $h_{12} = h_1$. Then we have

$$\mathbf{r}_0 = \frac{1}{k_0 + k_1}\begin{pmatrix} k_1 \\ k_0 \end{pmatrix}, \qquad \text{(B13)}$$

$$K^+ = \frac{1}{(k_0 + k_1)^2}\begin{pmatrix} -k_0 & k_1 \\ k_0 & -k_1 \end{pmatrix}, \qquad \text{(B14)}$$

$$\mathbf{g}\mathbf{r}_0 = \frac{g_0k_1 + g_1k_0}{k_0 + k_1}, \qquad \text{(B15)}$$

$$\mathbf{h}\mathbf{r}_0 = \frac{h_0k_1 + h_1k_0}{k_0 + k_1}. \qquad \text{(B16)}$$

So Eq. B1 becomes

$$\partial_t f(x, t) = -\partial_x(A(x)f(x, t)) + \tfrac{1}{2}\partial_x^2(B(x)f(x, t)), \qquad \text{(B17)}$$

where

$$A(x) = \frac{k_1 g_0 + k_0 g_1}{k_0 + k_1} + \epsilon \frac{1}{(k_0 + k_1)}\left[k_0 g_0 \partial_x\left(k_0 \frac{g_0 - g_1}{(k_0 + k_1)^2}\right) - k_0 g_1 \partial_x\left(k_1 \frac{g_0 - g_1}{(k_0 + k_1)^2}\right)\right] \quad \text{(B18)}$$

and

$$B = 2\epsilon \frac{k_0 k_1 (g_0 - g_1)^2}{(k_0 + k_1)^3} + \frac{h_0 k_1 + h_1 k_0}{k_0 + k_1}. \quad \text{(B19)}$$

Note that, when $\epsilon \to 0$ and when within-operator state fluctuations are negligible, we recover the deterministic ODE for the system,

$$\frac{dx}{dt} = \frac{k_1 g_0 + k_0 g_1}{k_0 + k_1}. \quad \text{(B20)}$$

The correspondence with the self-promoter system discussed in the text is established with $g_0(x) = 1 - x$, $g_1(x) = a_0 - x$, $h_0 = 1 + x$, $h_1 = a_0 + x$, $K_{10}(x) = b$, $K_{01}(x) = x^2$, and $\epsilon = 1/\kappa$, leading to the expressions for $A(x)$ and $B(x)$ given by Eqs. 37 and 38.

APPENDIX C: MUTUAL REPRESSORS

Converting to dimensionless variables, the diffusion limit for fluctuations in the monomer concentration is

$$\partial_u \rho_1(x_1, x_2) = -\partial_{x_1}(1 - x_1)\rho_1 + \partial_{x_2} x_2 \rho_1 + \frac{1}{2m_o}(\partial_{x_1}^2(1 + x_1) + \partial_{x_2}^2 x_2)\rho_1 - \kappa(b\rho_1 - x_1^2\rho_0), \quad \text{(C1)}$$

$$\partial_u \rho_2(x_1, x_2) = -\partial_{x_2}(1 - x_2)\rho_2 + \partial_{x_1} x_1 \rho_2 + \frac{1}{2m_o}(\partial_{x_1}^2 x_1 + \partial_{x_2}^2(1 + x_2))\rho_2 - \kappa(b\rho_2 - x_1^2\rho_0), \quad \text{(C2)}$$

$$\partial_u \rho_0(x_1, x_2) = -\partial_{x_2}(1 - x_2)\rho_0 - \partial_{x_1}(1 - x_1)\rho_0 + \frac{1}{2m_O}(\partial_{x_1}^2(1 + x_1) + \partial_{x_2}^2(1 + x_2))\rho_0 - \kappa((x_1^2 + x_2^2)\rho_0 - b\rho_1 - b\rho_2), \quad \text{(C3)}$$

where $X_1 = M_1/m_o$ and $X_2 = M_2/m_o$.

In the limit of fast operator fluctuations, the diffusion approximation for the marginal density $\rho = \rho_0 + \rho_1 + \rho_2$ has the form

$$\partial_s \rho(x_1, x_2, s) = -\partial_{x_1} A_1(x_1, x_2)\rho - \partial_{x_2} A_2(x_1, x_2)\rho + \frac{1}{2}(\partial_{x_1}^2 B_1(x_1, x_2) + \partial_{x_2}^2 B_2(x_1, x_2)) - \partial_{x_1}\partial_{x_2} B_{12}(x_1, x_2)\rho, \quad \text{(C4)}$$

where

$$A_1(x_1, x_2) = \frac{1}{1 + x_2^2/(b + x_1^2)} - x_1 - \frac{1}{\kappa}(2x_1 x_2(x_1 + x_2)) \times \left(\frac{(x_1 - 1)(b + x_1^2)(2b + x_1^2) + x_1 x_2^2(3b + x_1(2x_1 - 1)) + x_1 x_2^4}{b(b + x_1^2 + x_2^2)^4}\right),$$

$$A_2(x_1, x_2) = \frac{1}{1 + x_1^2/(b + x_2^2)} - x_2 - \frac{1}{\kappa}(2x_1 x_2(x_1 + x_2)) \times \left(\frac{(x_2 - 1)(b + x_2^2)(2b + x_2^2) + x_2 x_1^2(3b + x_2(2x_2 - 1)) + x_2 x_1^4}{b(b + x_2^2 + x_1^2)^4}\right),$$

$$B_1(x_1, x_2) = \frac{1}{m_o}\left(\frac{1}{1 + x_2^2/(b + x_1^2)} + x_1\right) + \frac{1}{\kappa}\left(\frac{x_2^2(b^2 + 2bx_1^2 + x_1^2 x_2^2 + x_1^4)}{b(b + x_1^2 + x_2^2)^3}\right),$$

$$B_2(x_1, x_2) = \frac{1}{m_o}\left(\frac{1}{1 + x_1^2/(b + x_2^2)} + x_2\right) + \frac{1}{\kappa}\left(\frac{x_1^2(b^2 + 2bx_2^2 + x_1^2 x_2^2 + x_2^4)}{b(b + x_1^2 + x_2^2)^3}\right),$$

$$B_{12}(x_1, x_2) = \frac{1}{\kappa}\left(\frac{x_1^2 x_2^2(2b + x_1^2 + x_2^2)}{b(b + x_1^2 + x_2^2)^3}\right).$$

APPENDIX D: THE MEAN FIRST PASSAGE TIME

Here we derive equations that govern the mean time to switch from one quasi-stable state to another. For simplicity, we will only consider the case in which there is one protein species. For more complicated systems, one is, in general, forced to use numerical techniques to compute the mean first-passage time.

Continuous and discrete processes

We begin by considering the general situation in which the stochastic process has both a continuous and discrete component. This corresponds to a case in which the monomer concentration is considered to be continuous, but the states of the operator are discrete. Below, we specialize the treatment to consider purely discrete or continuous systems.

The starting point for these considerations is the backward equation (corresponding to the forward equation Eq. B1) for the conditional or "transition" densities $\rho_{ij}(y, t|x, 0)$ (Gardiner, 1990)

$$\frac{\partial}{\partial t}\boldsymbol{\rho} = (\mathbf{L}^{\dagger}(x)\boldsymbol{\rho})^{\mathrm{T}} + \boldsymbol{\rho}\mathbf{K}(x), \quad \text{(D1)}$$

where

$$L^{\dagger}_{ii} = \sum_{j=1}^{n} g_{ji}(x)\partial_{x_j} + \frac{1}{2}\sum_{j=1}^{n} h_{ji}(x)\partial^2_{x_j}. \quad \text{(D2)}$$

If we consider the closed interval $[a, b]$ with an absorbing barrier at one or both ends, the mean first-passage time, $T_i(x)$ for the concentration to leave this interval, given that it started at x with the operator in state i at $t = 0$, is

$$T_i(x) = \int_0^{\infty}\int_a^b \sum_j^m \rho_{ji}(y|x)\mathrm{d}y\mathrm{d}t = \sum_j^m \tau_{ji}(x). \quad \text{(D3)}$$

Using Eq. D1, it is straightforward to show that the τ satisfy the equation

$$-\mathbf{I} = (\mathbf{L}^{\dagger}(x)\boldsymbol{\tau})^{\mathrm{T}} + \boldsymbol{\tau}\mathbf{K}(x). \quad \text{(D4)}$$

Let $\mathbf{T}$ denote the vector whose elements consist of $T_i(x)$. Summing the above matrix equation over the rows produces the vector equation for $\mathbf{T}$,

$$-\mathbf{1} = (\mathbf{L}^{\dagger}\mathbf{T})^{\mathrm{T}} + \mathbf{T}\mathbf{K}, \quad \text{(D5)}$$

where $\mathbf{1}$ is an m-dimensional row vector of ones. The above equation represents a set of m nonhomogeneous-coupled second-order ODEs. The boundary conditions are $T_i = 0$ at an absorbing boundary and $\partial T_i/\partial x = 0$ at a reflecting boundary condition. The mean first-passage time for the process is found from

$$\langle T(x)\rangle = \sum_{i=1}^{n} p_i T_i(x), \quad \text{(D6)}$$

where p_i denotes the probability of being in state i at $t = 0$.

Even for the simple two-state system discussed above, analytic solutions to Eq. D5 are unavailable. If we ignore fluctuations in the concentration, the equations for the two-state system become

$$-1 = g_0(x)\frac{\mathrm{d}T_0}{\mathrm{d}x} - k_0(x)(T_0 - T_1), \quad \text{(D7)}$$

$$-1 = g_1(x)\frac{\mathrm{d}T_1}{\mathrm{d}x} - k_1(x)(T_1 - T_0). \quad \text{(D8)}$$

Notice that the order of the equations has been reduced from second to first. Therefore, we only need two boundary conditions. If the deterministic flow is toward an absorbing barrier, then T_i for that state must vanish at the boundary. If the flow is toward a reflecting boundary or stable fixed point, then $\mathrm{d}T_i/\mathrm{d}x$ vanishes at that point. There are no boundary conditions at the points where the flow is away from absorbing or reflecting boundary or unstable fixed point.

To solve Eqs. D7 and D8, we use the change of variables $\Delta = T_0 - T_1$. This produces

$$\frac{\mathrm{d}\Delta}{\mathrm{d}x_0} = \frac{g_1k_0 + g_0k_1}{g_0g_1}\Delta - \frac{g_1 - g_0}{g_0g_1}, \quad \text{(D9)}$$

whose solution can be written explicitly. The remaining differential equation is

$$\frac{\mathrm{d}T_1}{\mathrm{d}x_0} = -\frac{1}{g_1(x_0)} - \frac{k_1(x_0)}{g_1(x_0)}\Delta, \quad \text{(D10)}$$

which, again, has solution given by quadrature involving the solution for Δ.

Discrete processes

Here we consider the case in which the monomer number is treated as a discrete random variable. We restrict ourselves to the self-promoting system discussed in the text. Let $p_{b|a}(j, t|m, 0)$ denote the transition density for the operator to be in state b with j monomers present at time t given that, at time 0, the operator was in state a with m monomers present. If there are l possibilities for the number of monomers, the transition densities can be arranged in a $2l \times 2l$ matrix $\mathbf{P}$, where the 2 comes from the fact that the operator has 2 states. The backward equation then has the form

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \mathbf{P}\mathbf{W}, \quad \text{(D11)}$$

where $\mathbf{W}$ is the transition matrix for the entire process. The elements of W come from the underlying master equation for the process, e.g., Eq. 32.

Let $T^{\mathrm{a}}_{\mathrm{m}}$ be the MFPT for a system starting in state (a, m) with an absorbing barrier placed at $N \geq m$ and a reflecting barrier at $n \leq m$. Using Eqs. D11 and 32 and following the same reasoning as described above for the mixed case, we have

$$\begin{aligned} -1 = {} & \alpha_0(T^0_{\mathrm{m}+1} - T^0_{\mathrm{m}}) \\ & + \delta m(T^0_{\mathrm{m}-1} - T^0_{\mathrm{m}}) \\ & + \frac{k_0}{\theta}m(m-1)(T^1_{\mathrm{m}} - T^0_{\mathrm{m}}), \end{aligned} \quad \text{(D12)}$$

$$\begin{aligned} -1 = {} & \alpha_1(T^1_{\mathrm{m}+1} - T^1_{\mathrm{m}}) \\ & + \delta(m-1)(T^1_{\mathrm{m}-1} - T^1_{\mathrm{m}}) \\ & + k_0\beta(T^0_{\mathrm{m}} - T^1_{\mathrm{m}}). \end{aligned} \quad \text{(D13)}$$

The boundary conditions are as follows. At the absorbing boundary we have $T^{\mathrm{a}}_{\mathrm{N}} = 0$. At the reflecting boundary we have

$$0 = \alpha_0(T^0_{\mathrm{n}+1} - T^0_{\mathrm{n}}) + \frac{k_0}{\theta}m(m-1)(T^1_{\mathrm{n}} - T^0_{\mathrm{n}}), \quad \text{(D14)}$$

$$0 = \alpha_1(T^1_{\mathrm{n}+1} - T^1_{\mathrm{n}}) + k_0\beta(T^0_{\mathrm{n}} - T^1_{\mathrm{n}}). \quad \text{(D15)}$$

Eqs. D2–D15 must be solved numerically.

Continuous processes

For completeness, we note that an expression for the mean first-passage time can be constructed in the full diffusion limit (Gardiner, 1990). In this case, the mean first-passage time $T(x)$ satisfies the equation,

$$-1 = A(x)\frac{\mathrm{d}T(x)}{\mathrm{d}x} + \frac{B(x)}{2}\frac{\mathrm{d}^2T}{\mathrm{d}x^2}. \quad \text{(D16)}$$

The mean first-passage time for the concentration to leave the interval $[0, a]$ with a reflecting boundary at $x = 0$ and an absorbing boundary at $x = a$ is

$$T(x) = 2\int_x^a \psi^{-1}(x')\mathrm{d}x' \int_0^{x'} \frac{\psi(y)}{B(y)}\,\mathrm{d}y, \quad \text{(D17)}$$

where

$$\psi(x) = \exp\left[2\int_0^x \frac{A(y)}{B(y)}\,dy\right]. \qquad \text{(D18)}$$

Because of the multiple integrals involved in the above expression, in general, it is not particularly useful. In fact, the results for this case presented in the manuscript were obtained by numerically solving Eq. D16 using a shooting method.

The authors would like to thank Bard Ermentrout for useful discussions during the early work on this project and Jeff Hasty for his critical reading of the manuscript.

T.C.E. was supported by the National Science Foundation under research grant DMS-0075821. T.B.K. was partially supported by National Science Foundation award MCB 9357637.

REFERENCES

Ahmad, K., and S. Henikoff. 2001. Modulation of a transcription factor counteracts heterochromatic gene silencing in *Drosophila*. *Cell*. 104: 839–847.

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

Becskei, A., B. Seraphin, L. Serrano. 2001. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J*. 20:2528–2535.

Bennett, D. C. 1983. Differentiation in mouse melanoma cells: initial reversibility and an on–off stochastic model. *Cell*. 34:445–453.

Cherry, J. L., and F. R. Adler. 2000. How to make a biological switch. *J. Theor. Biol*. 203:117–133.

Cook, D. L., A. N. Gerber, and S. J. Tapscott. 1998. Modeling stochastic gene expression: implications for haploinsufficiency. *Proc. Natl. Acad. Sci. U.S.A*. 95:15641–15646.

Dingemanse, M. A., P. A. J. de Boer, A. F. M. Moorman, R. Charles, W. H. Lamers. 1994. The expression of liver-specific genes within rat embryonic hepatocytes is a discontinuous process. *Differentiation*. 56: 153–162.

Elowitz, M. B., and S. Leibler. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature*. 403:335–338.

Endy, D., and R. Brent. 2001. Modeling cellular behavior. *Nature*. 409: 391–395.

Fiering, S., J. P. Northrop, G. P. Nolan, P. S. Mattila, G. R. Crabtree, L. A. Herzenberg. 1990. Single cell assay of a transcription factor reveals a threshold in transcription activated by signals emanating from the T-cell antigen receptor. *Genes Dev*. 4:1823–1834.

Gardiner, C. 1990. Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences. Springer-Verlag, Berlin. 136–142.

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

Grunstein, M. 1997. Histone acetylation in chromatin structure and transcription. *Nature*. 389:349–352.

Hasty, J., F. Isaacs, M. Dolnik, D. McMillen, and J. J. Collins. 2001a. Designer gene networks: towards fundamental cellular control. *Chaos*. 11:207–220.

Hasty, J., D. McMillen, F. Isaacs, and J. J. Collins. 2001b. Computational studies of gene regulatory networks: in numero molecular biology. *Nat. Rev. Gen*. 2:268–278.

Hasty, J., J. Pradines, M. Dolnik, and J. J. Collins. 2000. Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. U.S.A*. 97:2075–2080.

Horsthemke, W., and R. Lefever. 1984. Noise Induced Transitions. Theory and Applications in Physics, Chemistry and Biology. Springer-Verlag, Berlin. 258–292.

Ko, M. S., H. Nakauchi, and N. Takahashi. 1990. The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. *EMBO J*. 9:2835–2842.

Ko, M. S. H. 1991. A stochastic model for gene induction. *J. Theor. Biol*. 153:181–194.

Ko, M. S. H. 1992. Induction mechanism of a single gene molecule: stochastic or deterministic? *Bioessays*. 14:341–346.

McAdams, H. H., and A. Arkin. 1997. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. U.S.A*. 94:814–819.

McAdams, H. H., and A. Arkin. 1998. Simulation of prokaryotic genetic circuits. *Ann. Rev. Biophys. Biomol. Struct*. 27:199–224.

Ptashne, M. 1992. A Genetic Switch: Phage λ and Higher Organisms. 2nd edition. Cell Press and Blackwell Scientific Publications, Cambridge, MA.

Santillan, M. S., and M. C. Mackey. 2001a. Dynamic regulation of the tryptophan operon: a modeling study and comparison with experimental data. *Proc. Natl. Acad. Sci. U.S.A*. 98:1364–1369.

Santillan, M. S., and M. C. Mackey. 2001b. Dynamic behavior in mathematical models of the tryptophan operon. *Chaos*. 11:261–268.

Shea, M. A., and G. K. Ackers. 1985. The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. *J. Mol. Biol*. 181:211–230.

Thattai, M., and A. van Oudenaarden. 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. U.S.A*. 98:8614–8619.

van Kampen, N. G. 1992. Stochastic Processes in Chemistry and Physics. North-Holland, Amsterdam. 139–208.

van Roon, M. A., J. A. Aten, C. H. van Oven, R. Charles, and W. H. Lamers. 1989. The initiation of hepatocyte-specific gene expression within embryonic hepatocytes is a stochastic event. *Dev. Biol*. 136: 508–516.

Walters, M. C., S. Fiering, J. Eidemiller, W. Magis, M. Groudine, and D. I. K. Martin. 1995. Enhancers increase the probability but not the level of gene expression. *Proc. Natl. Acad. Sci. U.S.A*. 92:7125–7129.

Weintraub, H. 1988. Formation of stable transcription complexes as assayed by analysis of individual templates. *Proc. Natl. Acad. Sci. U.S.A*. 85:5819–5823.

Wijgerde, M., F. Grosveld, and P. Fraser. 1995. Transcription complex stability and chromatin dynamics in vivo. *Nature*. 377:209–213.

Yeivin, A., and A. Razin. 1993. DNA methylation: molecular biology and biological significance. *In* DNA Methylation: Molecular Biology and Biological Significance. J. P. Jost, H. P. Saluz, editors. Birhauser Verlag, Basel, Switzerland. 523–568.