Gene regulatory networks: A coarse-grained, equation-free approach to multiscale computation

Radek Erban[a)]
Mathematical Institute, University of Oxford, 24-29 St. Giles', Oxford OX1 3LB, United Kingdom

Ioannis G. Kevrekidis
*Department of Chemical Engineering, PACM and Mathematics, Princeton University,
Engineering Quadrangle, Olden Street, Princeton, New Jersey 08544*

David Adalsteinsson
*Applied Mathematics Program, Department of Mathematics, University of North Carolina, Chapel Hill,
North Carolina 27599*

Timothy C. Elston
Department of Pharmacology, University of North Carolina, Chapel Hill, North Carolina 27599

(Received 17 August 2005; accepted 14 November 2005; published online 27 February 2006)

We present computer-assisted methods for analyzing stochastic models of gene regulatory networks. The main idea that underlies this equation-free analysis is the design and execution of appropriately initialized short bursts of stochastic simulations; the results of these are processed to estimate coarse-grained quantities of interest, such as mesoscopic transport coefficients. In particular, using a simple model of a genetic toggle switch, we illustrate the computation of an effective free energy Φ and of a state-dependent effective diffusion coefficient D that characterize an unavailable effective Fokker-Planck equation. Additionally we illustrate the linking of equation-free techniques with continuation methods for performing a form of stochastic "bifurcation analysis"; estimation of mean switching times in the case of a bistable switch is also implemented in this equation-free context. The accuracy of our methods is tested by direct comparison with long-time stochastic simulations. This type of equation-free analysis appears to be a promising approach to computing features of the long-time, coarse-grained behavior of certain classes of complex stochastic models of gene regulatory networks, circumventing the need for long Monte Carlo simulations. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2149854]

I. INTRODUCTION

Various ways to model gene regulatory networks exist, ranging from logical (Boolean) to stochastic (Monte Carlo methods) or deterministic (ordinary differential equations) models (for recent reviews, see Refs. 1–3). Each modeling approach has its advantages and disadvantages. One advantage of stochastic modeling is that it takes into account fluctuations due to the inherently random nature of biochemical reactions. This intrinsic noise gives rise to significant effects when either the molecular abundances of protein or mRNA molecules are small or the kinetics of the transitions between the chemical states of the promoter are slow.[3,4]

The established approach for stochastic modeling of spatially homogeneous chemical systems was introduced by Gillespie.[5] The Gillespie stochastic simulation algorithm (SSA) is based on repeatedly answering two questions: when does the next chemical reaction occur and what kind of reaction is it? Gillespie[5] derived a simple way to answer these two questions that reduces the problem to a continuous-time discrete space Markov process.

The SSA generates exact sample paths of the stochastic process and, for sufficiently large networks, it is computationally more efficient than solving the chemical master equation. However, the large size of naturally occurring gene regulatory networks makes even the SSA computationally intensive and practically impossible to use for computing the long-time behavior of the network. Consequently, an important restriction of stochastic computations for many networks of interest is that we can efficiently run stochastic Gillespie-based simulators for *short times* only. It is therefore natural to look for computational methods that use only short-time simulations (and as few of these as necessary) to compute the required information for the system. Such a computer-assisted approach is presented in this paper.

Model reduction often provides a natural path to efficient simulation of a complicated model. As in other branches of physical modeling, separation of time scales can lead to successful model reduction in gene regulatory network modeling. Separation of time scales is frequently present in this context because synthesis and degradation of new proteins and transcripts usually occur on a slower time scale than processes that change the chemical state of proteins (e.g., multimerization, protein/DNA interactions, and phosphorylation). Theoretical methods for stochastic model reduction that take advantage of separation of time scales are being developed (e.g., Refs. 4 and 6–8). Analytical reduction tech-

[a)] Author to whom correspondence should be addressed. Electronic mail: erban@maths.ox.ac.uk niques assume that fast variables are in quasisteady state with respect to the remaining slow variables. If the quasisteady-state distributions conditioned on the slow variables can be determined, then they can be used to eliminate the fast variables.

Our approach is also based on (and takes advantage of) the separation of time scales and the approximation (computationally, on the fly) of quasisteady marginal distributions (conditioned on the slow variables). The main feature of our approach, as will become apparent through its description and illustration, is that we do not "first reduce and then simulate the reduced model"; our methods come in the form of wrappers around a black box dynamic simulator and could equally well be applied to the most detailed stochastic version of the network model or to its best explicit reduction already available. In our approach, results about the long-term dynamic behavior of a stochastic simulator do not come from long-term simulation; they come from the design, execution, and processing of the results of "intelligently designed" short bursts of direct dynamic simulation.

We believe it is useful to draw here an analogy with the study of nonlinear dynamics in systems of ordinary differential equations (ODEs). Long-term information in the form of detailed bifurcation diagrams can be obtained from long dynamic integration; yet the same information is much more systematically and economically obtained through *different algorithms* using the same model: bifurcation, stability, and continuation methods. It is this alternative to direct, long-term stochastic simulation (whether with the full detailed network model or with any good analytical reduction of it) that our approach makes available to the modeler. Ours is a "design of computational experiments" approach; it is guided by model reduction, but a reduced model is never explicitly obtained.

The remainder of the paper is organized as follows. In Sec. II, we introduce the genetic toggle switch as a simple model to illustrate our methods, and we specify the main questions that one would like to answer with these techniques. In Sec. III, we present the general mathematical framework and main ideas of *equation-free* analysis.[9–13] In Sec. IV, we present an analysis of a deterministic model of the genetic toggle switch to provide insight into this system. We also introduce several stochastic models of increasing complexity that are used to illustrate equation-free analysis. In Sec. V, we compute the effective free energies and the associated stationary distributions for the stochastic models described in Sec. IV. Equation-free bifurcation analysis is then presented, and, in bistable cases, the mean first passage times for the system to switch between apparent stable fixed points are computed. We end with a discussion of the equation-free approach, its strengths, weaknesses, relations to other current methods for the acceleration of SSA-type simulations (e.g., Refs. 4, 6–8, and 14), and its possible extensions in Sec. VI. In particular, we will discuss the applicability of our methods to more complicated gene regulatory networks.

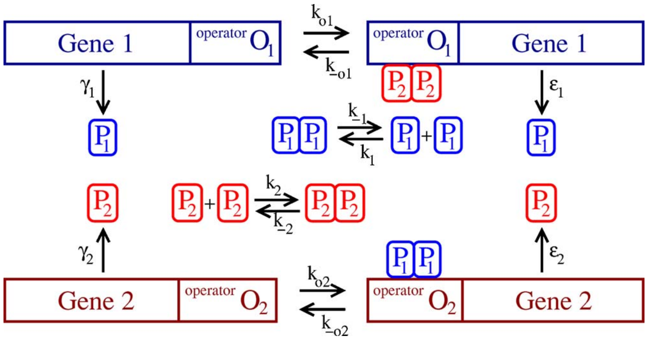

FIG. 1. A schematic diagram of the genetic toggle switch.

II. MODEL DESCRIPTION

Our illustrative example is a two-gene network in which each protein represses the transcription of the other gene (mutual repression). This type of system has been engineered in *E. coli* and is often referred to as a genetic toggle switch.[15,16] The advantage of this simple system is that it allows us to test the accuracy of equation-free methods by direct comparisons with results from long-time stochastic simulations. In Sec. VI, we discuss the applicability of our methods to more complex problems where long direct stochastic simulation is impossible and the accuracy must be checked by online *a posteriori* error estimates.

A simple version of the genetic toggle switch is schematically drawn in Fig. 1. The system contains two proteins P_1 and P_2. The production of P_1 (P_2) depends on the chemical state of the upstream operator O_1 (O_2). If O_1 is empty then P_1 is produced at the rate γ_1 and if O_1 is occupied by a dimer of P_2, then protein P_1 is produced at a rate $\epsilon_1 < \gamma_1$. Similarly, if O_2 is empty then P_2 is produced at the rate γ_2 and if O_2 is occupied by a dimer of P_1, then protein P_2 is produced at a rate $\epsilon_2 < \gamma_2$. Note that for simplicity, transcription and translation are described by a single rate constant. The biochemical reactions and rate constants that correspond to the processes shown in Fig. 1 are

$$\emptyset \underset{\delta_1}{\overset{\gamma_1 O_1 + \epsilon_1 \overline{P_2 P_2 O_1}}{\rightleftarrows}} P_1, \tag{2.1}$$

$$\emptyset \underset{\delta_2}{\overset{\gamma_2 O_2 + \epsilon_2 \overline{P_1 P_1 O_2}}{\rightleftarrows}} P_2, \tag{2.2}$$

$$P_1 + P_1 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \overline{P_1 P_1}, \tag{2.3}$$

$$P_2 + P_2 \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} \overline{P_2 P_2}, \tag{2.4}$$

$$\overline{P_2 P_2} + O_1 \underset{k_{-o1}}{\overset{k_{o1}}{\rightleftarrows}} \overline{P_2 P_2 O_1}, \tag{2.5}$$

$$\overline{P_1 P_1} + O_2 \underset{k_{-o2}}{\overset{k_{o2}}{\rightleftarrows}} \overline{P_1 P_1 O_2}, \tag{2.6}$$

where the overbars denote complexes. Equation (2.1) describes production and degradation of protein P_1, Eq. (2.2) describes production and degradation of protein P_2, Eqs. (2.3) and (2.4) are dimerization reactions, and Eqs. (2.5) and (2.6) represent the binding and dissociation of the dimer and DNA.

Single cell fluorescence measurements can be used to measure intercellular variability in protein expression levels. Therefore it is important to have efficient methods for computing the steady-state distribution of protein abundances from stochastic models similar to the one defined by (2.1)–(2.6). For moderately complex systems using long-time Monte Carlo simulations quickly becomes computationally prohibitive. We will illustrate how equation-free analysis can overcome this difficulty by accelerating the exploration of certain features of the long-term dynamics of the stochastic simulation. For certain values of the model parameters, the genetic toggle switch is bistable. If the system is described in terms of ODEs for the protein concentrations, then standard bifurcation analysis (numerical continuation methods) can be applied to determine the regions of parameter space in which bistability occurs. Using the model described by (2.1)–(2.6) as an example, we show how to extend these techniques to stochastic models. An important quantity that characterizes the dynamics of bistable stochastic systems is the average time for spontaneous transitions between stable steady states to occur. We will illustrate how this mean first passage time can be computed by using only short-time simulations.

III. EQUATION-FREE ANALYSIS: MATHEMATICAL FRAMEWORK

Let us suppose that we have a well-stirred mixture of chemically reacting species; our main assumption is that the evolution of the system can be described in terms of a single, slowly evolving random variable Q (the approach carries through for the case of a small number of slow variables, but in this paper we will focus on the single slow variable case). Q might be the concentration of one of the chemical species or some function of the concentrations. Let $\mathbf{R}$ denote a vector of the other (fast, "slaved" system variables). Our assumption implies that the evolution of the system can be approximately described by the time-dependent probability density function $f(q,t)$ for the values q of the slow variable Q that evolves according to the following effective Fokker-Planck equation:[17]

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial q}\left(\frac{\partial}{\partial q}[D(q)f(q,t)] - V(q)f(q,t)\right). \tag{3.1}$$

If the effective drift $V(q)$ and diffusion coefficient $D(q)$ could be explicitly written down as functions of q, then (3.1) could be used to compute interesting properties of the system (e.g., the steady-state distribution). Note that in addition to the assumption of a single slow variable, the validity of Eq. (3.1) requires sufficiently large molecular abundances and sufficiently fast chemical kinetics for the binding and release of the dimers from the operator sites on the DNA.[4,15] In the limit of fast chemical kinetics these binary transitions at the operator become temporally self-averaging. Assuming that (3.1) provides a good approximation, we make use of the following formulas for the drift and diffusion coefficient[13,18–20]

$$V(q) = \lim_{\Delta t\to 0} \frac{\langle Q(t+\Delta t) - q | Q(t) = q\rangle}{\Delta t}, \tag{3.2}$$

$$D(q) = \frac{1}{2}\lim_{\Delta t\to 0} \frac{\langle [Q(t+\Delta t) - q]^2 | Q(t) = q\rangle}{\Delta t}. \tag{3.3}$$

As described below, estimates of these two quantities can be found by using short-time bursts of appropriately initialized stochastic simulations. The steady solution of (3.1) is proportional to $\exp[-\beta\Phi(q)]$, where the effective free energy $\Phi(q)$ is defined as

$$\frac{\Phi(q)}{k_B T} \equiv \beta\Phi(q) = -\int_0^q \frac{V(q')}{D(q')}dq' + \ln D(q) + \text{constant}. \tag{3.4}$$

Consequently, computing the effective free energy and the steady-state probability distribution can also be accomplished without the need for long-time stochastic simulations.

A procedure for computationally estimating $V(q)$ and $D(q)$ is as follows:

(A) Given $Q=q$, approximate the conditional density $P(\mathbf{r}|Q=q)$ for the fast variables $\mathbf{R}$. Details of this preparatory step are given below.

(B) Use $P(\mathbf{r}|Q=q)$ from step (A) to determine appropriate initial conditions for the short simulations and run multiple realizations for time Δt. Use the results of these simulations and the definitions (3.2) and (3.3) to estimate the average velocity $V(q)$ and effective diffusion coefficient $D(q)$.

(C) Repeat steps (A) and (B) for sufficiently many values of q and then compute $\Phi(q)$ using formula (3.4) and numerical quadrature.

A very important feature of this algorithm is that it is trivially parallelizable (different realizations of short simulations starting at the "same" q as well as simulation realizations starting at different q values can be run independently on multiple processors).

In order to use the algorithms (A)–(C), we have to specify how step (A) is performed. There are several computational options to approximate the conditional density $P(\mathbf{r}|Q=q)$. The simplest approximation is to estimate (through numerical experiments) the conditional mean $\langle\mathbf{R}|Q=q\rangle$ and approximate $P(\mathbf{r}|Q=q)$ as a Dirac delta function $\delta(\mathbf{r}-\langle\mathbf{R}|Q=q\rangle)$. Then step (A) reads as follows:

(1) Given $Q=q$, pick an initial guess for the conditional mean of $\mathbf{R}$. Denote the initial guess as $\langle\mathbf{R}(0)\rangle$. Run multiple realizations for a short time δt and compute $\langle\mathbf{R}(\delta t)\rangle$. This procedure defines the mapping $\langle\mathbf{R}(0)\rangle \to \langle\mathbf{R}(\delta t)\rangle$. Find the steady state of this mapping using standard numerical methods. The steady state is the required conditional average $\langle\mathbf{R}|Q=q\rangle$.

Initialize $\mathbf{R}(0)$ as $\langle \mathbf{R} | Q=q \rangle$ in all realizations in part (B) of the algorithm.

Another option is to approximate $P(\mathbf{r}|Q=q)$ as a distribution characterized by a few parameters, e.g., as a Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma}$. This can be done as follows:

(2) Given $Q=q$, pick initial guesses for the mean $\boldsymbol{\mu}(0)$ and variance $\boldsymbol{\sigma}(0)$ of the conditional distribution function $P(\mathbf{r}|Q=q)$. Use this distribution to generate many realizations of $\mathbf{R}(0)$. Using these realizations as initial conditions, run stochastic simulations for a short time δt and compute $\mathbf{R}(\delta t)$. Computing the mean and variance of $\mathbf{R}(\delta t)$, we obtain the mapping $[\boldsymbol{\mu}(0), \boldsymbol{\sigma}(0)] \to [\boldsymbol{\mu}(\delta t), \boldsymbol{\sigma}(\delta t)]$. Next use standard numerical methods to find the steady state $[\boldsymbol{\mu}, \boldsymbol{\sigma}]$ of this mapping and approximate $P(\mathbf{r}|Q=q)$ as a Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma}$.

The conditional density $P(\mathbf{r}|Q=q)$ can also be approximated by other basis functions. It is straightforward to generalize (1) or (2) to such a case. The better the approximation of $P(\mathbf{r}|Q=q)$ we have, the shorter the time step Δt required in step (B) to achieve the same accuracy. So, a better approximation of $P(\mathbf{r}|Q=q)$ in step (A) decreases the computational intensity of step (B). On the other hand, step (A) is more computationally intensive if we want to obtain a better approximation of $P(\mathbf{r}|Q=q)$. One possibility for generating a better approximation of $P(\mathbf{r}|Q=q)$ is to use a "run-and-reset" procedure as was done in Ref. 10. This is accomplished as follows:

(3) Given $Q=q$, initialize the other variables $\mathbf{R} \equiv \mathbf{R}(0)$ of the system. Run stochastic simulations for the short time δt. Then reset the value of $Q(\delta t)$ to its original value q keeping $\mathbf{R}$ unchanged. Repeat this procedure for many time steps and compute the conditional density $P(\mathbf{r}|Q=q)$ as a histogram of the recorded values of $\mathbf{R}$.

Approach (3) attempts to compute the $P(\mathbf{r}|Q=q)$ effectively by successive substitution, without resorting to numerical algorithms of the Newton-Raphson-type for locating fixed points of mappings; we will return to this latter issue in the Discussion section. In our illustrative computations in Sec. V, we use step (A) in form (1) or (3) for the simple stochastic models described below. Both give good results for our illustrative example. Since (1) works, there is no need to use (2) or higher-order approximations. For some stochastic simulations of our model problem, we also use slightly modified versions of method (1) or (3), as will be described in Sec. V.

A. Bifurcations

In deterministic problems, we often summarize the parametric dependence of the long-term dynamics in terms of bifurcation diagrams; for example, we may plot the steady states of a deterministic set of ODEs as a function of a distinguished *bifurcation parameter*. Several excellent continuation methods have been developed, implemented, and made available for this purpose over the years, such as AUTO.[21,22]

Here we illustrate how these methods can be extended to stochastic models.[9,11,20,23,24] We assume, as above, that we have a stochastic problem that can be *effectively* described by a single variable Q. Let γ be the bifurcation parameter. The first two steps in the algorithm are as follows:

(i) Given $Q=q$ and the value of the bifurcation parameter γ, compute the conditional density $P(\mathbf{r}|Q=q)$ using step (A) of the previous algorithm.

(ii) Using $P(\mathbf{r}|Q=q)$ from step (i) to determine the initial conditions, run multiple stochastic simulations for a short time Δt and compute the conditional average $\langle Q(\Delta t) | Q(0)=q \rangle$.

Steps (i) and (ii) define the mapping $(Q(0), \gamma) \to \langle Q(\Delta t) \rangle$. We denote this mapping as F, i.e., $F(Q,\gamma) = \langle Q(\Delta t) \rangle$. Our goal is to track the fixed points of F [i.e., $F(Q,\gamma)=Q$] as the bifurcation parameter γ is varied. To do this, we first use a Newton-Raphson algorithm to find two fixed points (Q_1, γ_1) and (Q_2, γ_2) which are sufficiently close to each other [note that one can estimate the derivative of $F(Q,\gamma)$ numerically by evaluating $F(Q,\gamma)$ at different points]. Then, in a parameter continuation context, steps are not taken directly in the parameter but in (pseudo)arclength along the solution branch in $Q \times \gamma$ space[21] (in order to allow the computation to automatically "go around" turning points). A small increment δ is chosen (which can be modified adaptively during the computation), and we find the next steady state Q, as well as the corresponding parameter value γ at distance approximately $\sqrt{\delta}$ on the solution branch from the last solution point by solving the augmented steady-state system of equations

$$Q - F(Q,\gamma) = 0,$$

$$(Q - Q_2)(Q_2 - Q_1) + (\gamma - \gamma_2)(\gamma_2 - \gamma_1) - \delta = 0. \quad (3.5)$$

To find the solution of (3.5), we estimate the Jacobian numerically by evaluating $F(Q,\gamma)$ at several points and then use a Newton-Raphson algorithm. When the number of variables starts becoming large, matrix-free methods of iterative numerical linear algebra [such as Broyden or Newton-Krylov GMRES (Ref. 25)] can be used to solve for the fixed point, as opposed to full numerical Jacobian estimation. The fixed points computed this way provide, under certain conditions, good estimates of the *critical points* (minima and saddles) of the effective potential $\Phi(q)$ as a function of a model parameter γ; this issue is discussed extensively in Refs. 20, 23, 24, and 26, and we will return to it again in the Discussion section.

B. First passage time

Suppose that we have a bistable stochastic system. That is, the effective free energy $\Phi(q)$ has two local minima[27]—see Fig. 2. Then an important quantity characterizing the long-time system dynamics is the mean time for spontaneous transitions to occur between the stable steady states. Let $q_m < q_M$ denote the two stable steady states and let q_u be the unstable state [i.e., local maximum of $\Phi(q)$]. Then

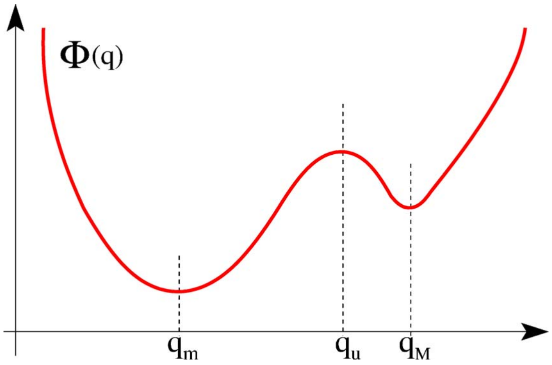

FIG. 2. Potential $\Phi(q)$ of the bistable system.

we define the first passage time for transitions from q_m to q_M as $2\tau_e$, where τ_e is the average time for the system to reach the unstable steady state q_u for the first time given that it starts at q_m. The factor of 2 occurs because once the system reaches the unstable steady state, half the time it returns to the original stable steady state q_m and the other half of the time it transitions to q_M.

Algorithm (A)–(C) gives a procedure to estimate the effective potential $\Phi(q)$ by running short simulations only. Once we have the effective potential, we can compute τ_e as follows:[27]

$$\tau_{e;p} = \int_{q_m}^{q_u} \exp[\beta\Phi(q)] \int_{-\infty}^{q} \frac{1}{D(\xi)} \exp[-\beta\Phi(\xi)] d\xi dq. \tag{3.6}$$

Equation (3.6) can be further simplified if the height of the potential barrier $[\Phi(q_u) - \Phi(q_m)]$ is large compared to the noise strength. In this case, almost all the weight of the function $\exp[-\beta\Phi(\xi)]$ is located at $\xi = q_m$ for $\xi < q_u$, so that the inner integral is essentially constant for $q_m < q < q_u$. Therefore the limit q in this integral can be replaced by q_u, allowing the two integrals to be evaluated separately

$$\tau_{e;p} \approx \int_{q_m}^{q_u} \exp[\beta\Phi(q)] dq \int_{-\infty}^{q_u} \frac{1}{D(q)} \exp[-\beta\Phi(q)] dq. \tag{3.7}$$

The main contribution of the first integral stems from the region around q_u, and the main contribution from the second integral stems from the region around q_m. Consequently, we expand $\Phi(q)$ according to

$$\Phi(q) \approx \Phi(q_u) - \frac{1}{2}|\Phi''(q_u)|(q - q_u)^2,$$

$$\Phi(q) \approx \Phi(q_m) + \frac{1}{2}\Phi''(q_m)(q - q_m)^2$$

for the first and the second integral, respectively.[17] When these expansions are used in Eq. (3.7), the following result is obtained:

$$\tau_{e;k} \approx \frac{4\pi \exp[\beta\Phi(q_u) - \beta\Phi(q_m)]}{\beta[D(q_u) + D(q_m)]\sqrt{\Phi''(q_m)|\Phi''(q_u)|}}, \tag{3.8}$$

which is the generalization of Kramers formula to the case of a state-dependent diffusion coefficient.[13,17] Formulas (3.6) and (3.8) are both used in Sec. V B to estimate τ_e.

IV. ANALYSIS OF THE MODEL PROBLEM

In this section, we study the behavior of the model given by Eqs. (2.1)–(2.6). To provide insight into the problem, we start by analyzing *the deterministic system*. In Secs. IV B and IV C, we introduce two stochastic models that are simplified versions of the model defined by (2.1)–(2.6). We use these models because of the relative ease in performing long-time stochastic simulations with them; this allows the results from the equation-free analysis to be validated by direct comparisons with Monte Carlo simulations. We will also verify that the equation-free methods can be applied to the full model. As discussed below, for this case the long-time Monte Carlo simulations become computationally very expensive.

A. The deterministic model

To simplify the deterministic analysis, we make the assumption that Eqs. (2.3)–(2.6) are at quasiequilibrium and derive deterministic rate equations for the protein concentrations. Let x_1 and x_2 denote the average monomer concentrations of P_1 and P_2, respectively, and let d_1 and d_2 denote the respective dimer concentrations. Also, let o_1 and o_2 denote the probabilities that the operators O_1 and O_2 are not occupied. For the dimerization process the assumption of quasiequilibrium implies that

$$d_1 = \frac{k_1}{k_{-1}} x_1^2 \quad \text{and} \quad d_2 = \frac{k_2}{k_{-2}} x_2^2. \tag{4.1}$$

Similarly, the quasiequilibrium assumption for the operators implies that

$$o_1 = \frac{k_{-o1}}{k_{-o1} + k_{o1} d_2} \quad \text{and} \quad o_2 = \frac{k_{-o2}}{k_{-o2} + k_{o2} d_1}. \tag{4.2}$$

The total concentration of P_1 is given by $y_1 = x_1 + 2d_1$. The total concentration y_1 evolves according to the following ordinary differential equation:

$$\frac{dy_1}{dt} = \gamma_1 o_1 + \varepsilon_1 (1 - o_1) - \delta_1 x_1, \tag{4.3}$$

where δ_1 is the degradation rate of the monomers, and it has been assumed that dimers are protected from degradation. Substituting $y_1 = x_1 + 2d_1 = x_1 + 2[k_1/(k_{-1})]x_1^2$ into (4.3), we obtain

$$\left(1 + 4\frac{k_1}{k_{-1}} x_1\right)\frac{dx_1}{dt} = \gamma_1 o_1 + \varepsilon_1(1 - o_1) - \delta_1 x_1.$$

Finally, using (4.1) and (4.2) produces

$$\frac{dx_1}{dt} = \frac{1}{1 + \kappa_1 x_1}\left[\gamma_1 \frac{1}{1 + \omega_1 x_2^2} + \varepsilon_1 \frac{\omega_1 x_2^2}{1 + \omega_1 x_2^2} - \delta_1 x_1\right], \tag{4.4}$$

where the parameters κ_1 and ω_1 are defined as follows:

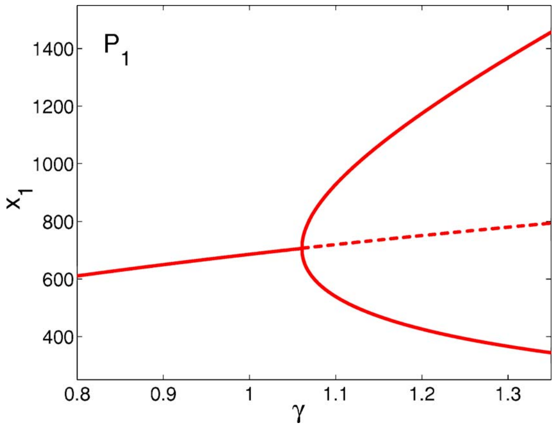

FIG. 3. The dependence of the steady-state values of x_1 on γ. The solid lines denote stable fixed points and the dashed line corresponds to unstable fixed points. In this figure and throughout the paper $\delta = 7.5\times 10^{-4}$ and $\omega = 2\times 10^{-6}$.

$$\kappa_1 = 4\frac{k_1}{k_{-1}} \quad \text{and} \quad \omega_1 = \frac{k_{o1}}{k_{-o1}}\frac{k_2}{k_{-2}}. \tag{4.5}$$

Using similar reasoning an analogous equation for x_2 can be derived.

For simplicity, we will present the symmetric case in which the rate constants for processes involving P_1 are identical to those involving P_2. That is, we assume $\kappa \equiv \kappa_1 = \kappa_2$, $\gamma \equiv \gamma_1 = \gamma_2$, $\omega \equiv \omega_1 = \omega_2$, and $\delta \equiv \delta_1 = \delta_2$. Moreover, we assume that the production rate is zero if an operator is occupied, i.e., $\varepsilon_1 = \varepsilon_2 = 0$. Making these assumptions, (4.4) simplifies to

$$\frac{dx_1}{dt} = \frac{1}{1+\kappa x_1}\left[\frac{\gamma}{1+\omega x_2^2} - \delta x_1\right], \tag{4.6}$$

and the equation for x_2 is obtained by alternating the subscripts in the above equation. Hence, the problem has been reduced to a system of two equations with four parameters. Note that the value of κ does not influence the steady-state behavior of the system. In this paper, we fix the values of δ and ω to be 7.5×10^{-4} and 2×10^{-6}, respectively.

The steady-state values of x_1 as a function of γ are shown in Fig. 3. In this figure, solid lines denote stable steady states and dashed lines denote unstable steady states. For $\gamma < 1.06$ there is a single steady state. At $\gamma = 1.06$ a pitchfork bifurcation occurs, and for $\gamma > 1.06$, there exist three steady states. The steady state with $x_1 = x_2$ is unstable and the other two steady states are stable.

Due to separation of time scales, the long-term dynamics of this problem lie on a lower-dimensional (here one-dimensional) *slow manifold*; this suggests that one may be able to construct an effective one-dimensional dynamical system describing the long-term evolution of the model on (near) this slow manifold. In constructing such a reduced model, an important question even in the simple deterministic case is the choice of the right *observable*—the variable in terms of which the long-term dynamics will be expressed. An extensive discussion of the choice of such a "right observable" for the deterministic case can be found, for example, in Ref. 28; as discussed there, even if we do not know the *exact* slow variables, any set of variables that parametrizes the slow manifold can be practically used to reduce the system in an equation-free context. For the stochastic case, a good early illustration and discussion of manifold parametrization can be found in Ref. 12. Choosing the right observable is an important issue in the implementation of equation-free computations and the subject of intense current research which we will briefly comment on in Sec. VI.

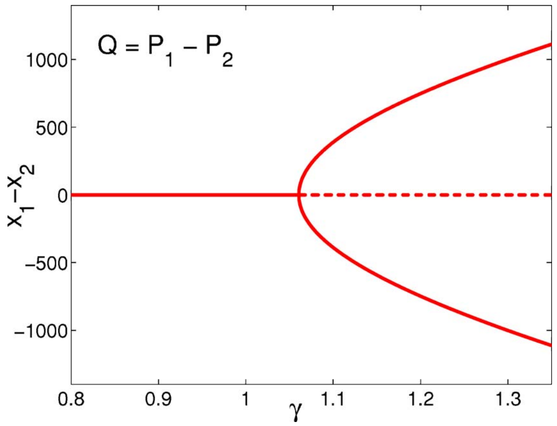

FIG. 4. The dependence of steady-state values of the symmetric variable $s = x_1 - x_2$ on γ.

In this paper, and for this example, our equation-free analysis assumes that the problem can be described in terms of a single variable. Consequently, it becomes important to select a good observable that further simplifies the two-dimensional problem to one dimension. A tempting (and obvious) choice for the one-dimensional observable is the molecular abundance of P_1 (or P_2). However, we also make use of the symmetric variable defined as the difference in the protein abundances $Q = P_1 - P_2$. In terms of the rate equations the symmetric variable is $s = x_1 - x_2$. The bifurcation diagram in terms of s is shown in Fig. 4. The symmetry of the diagram suggests that Q might be a more natural observable than P_1 (which also produces good results, as we will see below).

B. Stochastic model I

To start our investigations in the equation-free framework, we constructed a very simple stochastic model of the system. We use this simple model to benchmark equation-free computations, since the results can be tested against Monte Carlo simulations easily. Results for the full system are also presented below. The simple stochastic model consists only of reactions for the synthesis and degradation of proteins P_1 and P_2, but the following effective rate constants are used:

$$\emptyset \underset{\frac{\delta}{1+\kappa P_1}}{\overset{\frac{1}{1+\kappa P_1}\frac{\gamma}{1+\omega P_2^2}}{\rightleftarrows}} P_1, \tag{4.7}$$

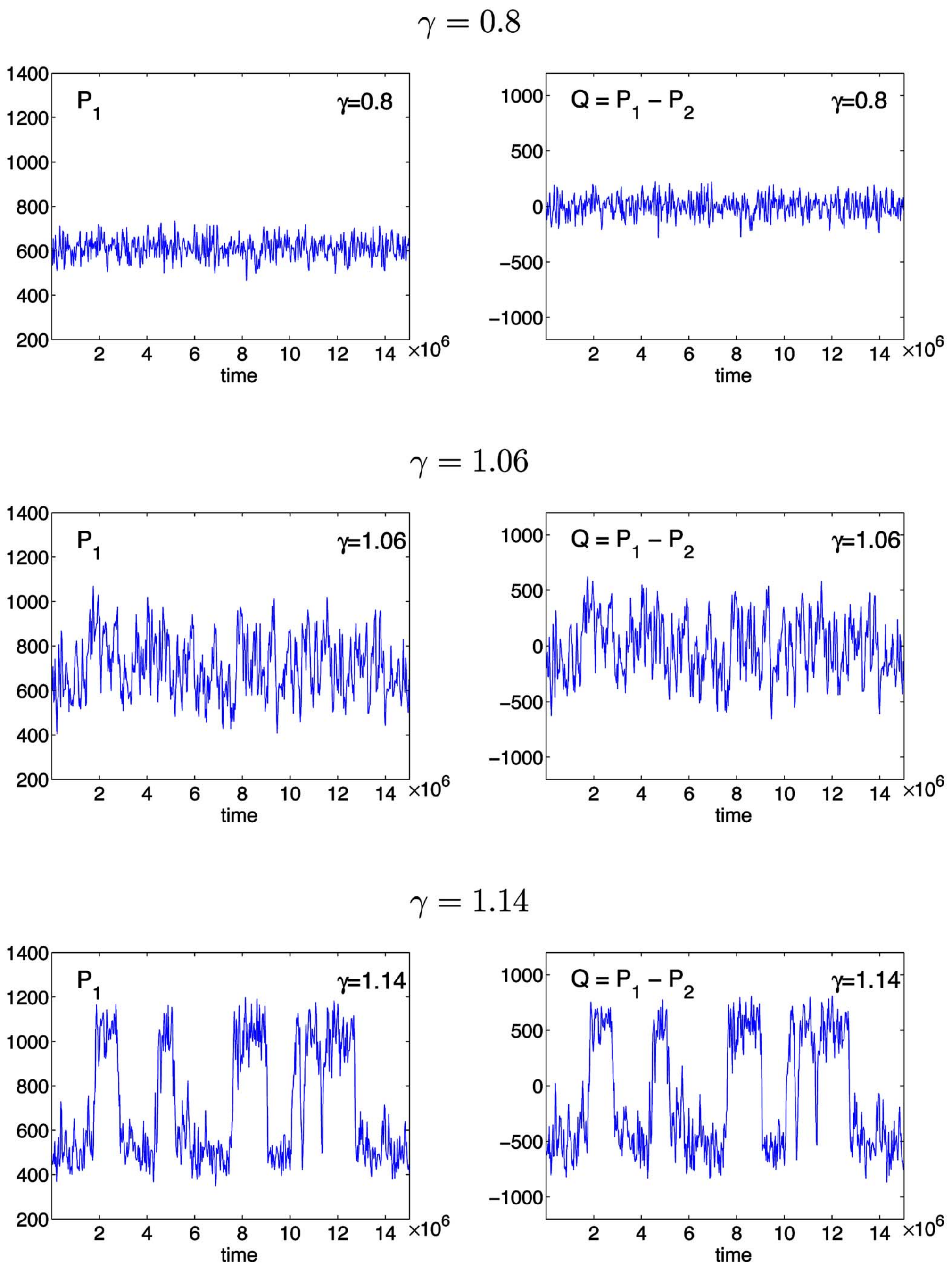

FIG. 5. Stochastic model I. Plots of P_1 and $Q=P_1-P_2$ as a function of time for different values of γ. The parameter values used to produce these figures are $\delta=7.5\times10^{-4}$, $\omega=2\times10^{-6}$, and $\kappa=2\times10^{-4}$.

$$\varnothing \underset{\frac{\delta}{1+\kappa P_2}}{\overset{\frac{1}{1+\kappa P_2}\frac{\gamma}{1+\omega P_1^2}}{\rightleftarrows}} P_2. \tag{4.8}$$

The above reactions are consistent with the deterministic model but, in general, do not preserve the noise structure of the full stochastic model.

To simulate the mechanism contained in models (4.7) and (4.8), we use the standard Gillespie SSA.[5] The results for different values of the parameter γ are plotted in Fig. 5. For each γ we plot the time evolution P_1 (left panel) and $Q=P_1-P_2$ (right panel). We see that for small γ, the solution fluctuates around the stable deterministic steady state with relatively small noise amplitude. When $\gamma=1.06$, the noise amplitude has increased substantially, which is typical of stochastic systems near a "bifurcation"; the word bifurcation is put here in quotes to denote that (in contrast to the deterministic case) there is no isolated parameter value marking the onset of bistability—no clear bifurcation point exists for the *stochastic dynamics*. Yet one can still claim that a clear bifurcation point exists for the critical points of the potential $\Phi(q,\gamma)$ in the stochastic model; furthermore, depending on the *time horizon* of our observation of a stochastic simulation, one may still appear to see an apparent bifurcation point for its averaged *statistics* (see the discussion in Refs. 13 and 26). If γ is increased further, then $Q=0$ is no longer a stable steady state and the system clearly shows bistability. All plots are computed for the same time interval $[0,15\times 10^6]$. For $\gamma=1.25$ the steady states are sufficiently stable so that no transitions occurred in this time interval (data not shown). Therefore, for this case, determining the steady-state probability distribution from long-term Monte Carlo simulations would be very time consuming.

C. Stochastic model II

Stochastic model I considers only two variables P_1 and P_2. Here, we introduce a stochastic model that also takes into account the biochemical states of the operators, while maintaining the assumption that dimerization reactions (2.3) and (2.4) are at equilibrium. That is, we consider the four variables P_1, P_2, O_1, and O_2. The model is defined in terms of the following reaction steps:

$$\emptyset \quad \overset{\frac{\gamma}{1+\kappa P_1}O_1}{\underset{\frac{\delta}{1+\kappa P_1}}{\rightleftarrows}} \quad P_1, \tag{4.9}$$

$$\emptyset \quad \overset{\frac{\gamma}{1+\kappa P_2}O_2}{\underset{\frac{\delta}{1+\kappa P_2}}{\rightleftarrows}} \quad P_2, \tag{4.10}$$

$$\text{“}O_1=0\text{”} \overset{K}{\underset{K\omega P_2^2}{\rightleftarrows}} \text{“}O_1=1\text{”}, \tag{4.11}$$

$$\text{“}O_2=0\text{”} \overset{K}{\underset{K\omega P_1^2}{\rightleftarrows}} \text{“}O_2=1\text{”}, \tag{4.12}$$

and contains an extra parameter, $K\equiv k_{-o1}=k_{-o2}$. Note that “$O_1=0$” means that the operator O_1 has a dimer of P_2 bound to it and therefore is “off” and “$O_1=1$” means that the operator O_1 is empty and therefore is “on.” The same is true for O_2. This implies that the random variables O_1 and O_2 are binary, whereas the variables P_1 and P_2 can take on any non-negative integer value. Stochastic model I is recovered from stochastic model II in the limit $K\to\infty$. We thus expect the models to produce similar results for large values of K.

Again, we use the standard Gillespie SSA (Ref. 5) to simulate model (4.9)–(4.12). The results for different values of K for $\gamma=1.14$ are plotted in Fig. 6. Comparing Fig. 6 and corresponding panel from Fig. 5, we can confirm that stochastic model II produces the same behavior as stochastic model I for large K. However, in general, different values of K can change the bifurcation structure of the system and affect the first passage times between the two stable steady states of the bistable system.[4]

Because stochastic models I and II do not explicitly take into account dimerization, which in general is a fast process, they run much more efficiently than the full model given by (2.1)–(2.6). However, they do not in general preserve the noise structure of the full system. In the next section we use all three models to highlight the computational features (and potential benefits) of equation-free analysis.

V. RESULTS OF EQUATION-FREE ANALYSIS

In our approach, we want to study the stochastic models presented above using only short bursts of appropriately initialized stochastic simulations; the goal is to design these bursts and process their results so as to determine long-time properties of the system (e.g., steady-state distributions, bifurcations, and mean first passage times) efficiently. We use (and compare) the different algorithms discussed in Sec. III.

A. The effective potential and steady-state distribution

In this section, we use equation-free analysis to evaluate the effective potential (an “effective free energy”) and the steady-state distribution for stochastic models I and II and the full system. We start with stochastic model I. First, we will consider the slow variable $Q\equiv P_1-P_2$ and the fast (slaved) variable $R\equiv P_1+P_2$. Initially the preparatory step (A) of the algorithm presented in Sec. III was done using the method outlined in (1), i.e., we used the conditional mean $\langle R|Q=q\rangle$ to initialize the computations in step (B). A good approximation to this average can be found using the deterministic equations, and this was the number used in our preliminary computations to initialize the simulations in step (B). That is, for a given Q, we initialized all realizations in step (B) with the same value of R. Then we chose Δt equal to 100 time steps of the Gillespie SSA. Note that this implies that the actual value of Δt varies for each realization and depends on the values of the rate constants. However, the computer (CPU) time is the same for all the results presented for this case. We averaged over 2×10^6 of realizations in part (B). Hence, we used 200×10^6 realizations for a given Q.

The equation-free results for the effective potential for different values of γ are given in Fig. 7. These results are in good agreement with the long-term stochastic simulations presented in Sec. IV B. The potential has a single minimum $\gamma<1.06$. As γ is increased the potential broadens implying that the system becomes “noisier.” When $\gamma>1.06$, the potential shows two local minima and the system is bistable.

Since we are using a very simple stochastic model, it is not computationally expensive to compute the steady-state distributions directly by long-time simulations. We used the Gillespie SSA to generate 10^{11} time steps of the stochastic process and recorded the value of Q at each time step. The resulting time series was binned to produce the steady-state distribution of the system. Figure 8 presents a comparison of the two computed steady-state distributions. The results obtained by long-time simulations are shown as histograms and the steady-state distributions computed from the effective potential $C\exp[-\beta\Phi(Q)]$ are given by the thick lines. We see that equation-free analysis gives very good results.

In Sec. III, we introduced three possible methods, (1)–(3), to perform the preparatory step (A) (typically called the “lifting” step in the equation-free framework). We have shown that approach (1) produces good results for stochastic model I. Since (1) works, there is no need to improve the results by considering (2). Instead, we discuss approach (3). In this approach given $Q=q$ we run the simulations for a short time δt and record the value of R. Then we reset Q

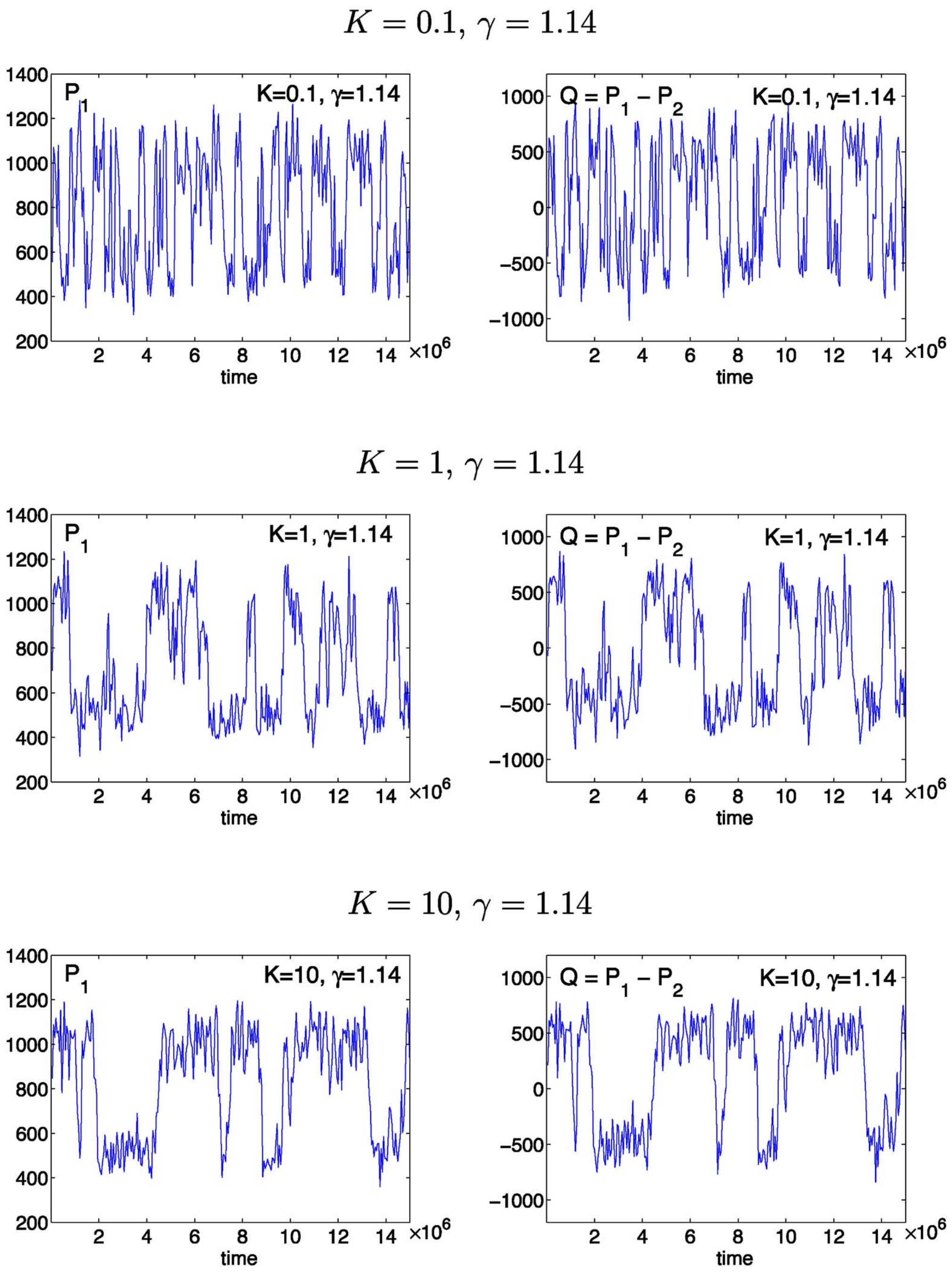

FIG. 6. Stochastic model II. Plots of P_1 and $Q=P_1-P_2$ as a function of time for different values of K and $\gamma=1.14$. The other model parameter values are the same as in Fig. 5.

$=q$ but leave R unchanged. We repeat this procedure many times and approximate the conditional density $P(r|Q=q)$ as a histogram of recorded values of R. In our simulations, we chose δt equal to one SSA step. To compute the conditional density $P(r|Q=q)$, we used 11×10^6 SSA steps. First we let the system run for a million time steps to remove the transient in R, and then used the remaining 10×10^6 time steps to compute the conditional density. In part (B) of the algorithm, we used Δt equal to 100 SSA time steps and we averaged over 2×10^6 of realizations, similarly as before. Thus, for each Q, we used 11×10^6 SSA time steps in step (3) and 200×10^6 SSA time steps in step (B) which means that step (3) did not significantly change the computational cost of the program.

The graphs of $P(r|Q=q)$ for $\gamma=0.98$ and $\gamma=1.14$ are given in Fig. 9. The left panel in these figures shows $P(r|Q=q)$ for five values of Q. The right panels show $P(r|Q=q)$ as a function of r and q. Next, we can use the computed conditional density $P(r|Q=q)$ to initialize R in step (B). Doing this produces results which are virtually identical to results from Fig. 8 (graphs not shown).

We now repeat the previous computations using the more complicated stochastic model II. The results are shown in Figs. 10 and 11. In Fig. 10, we choose $\gamma=1.14$ and compute the steady-state distribution for Q for three values of K. The results are compared with direct simulations of stochastic model II and with each other. The results from Fig. 10 can

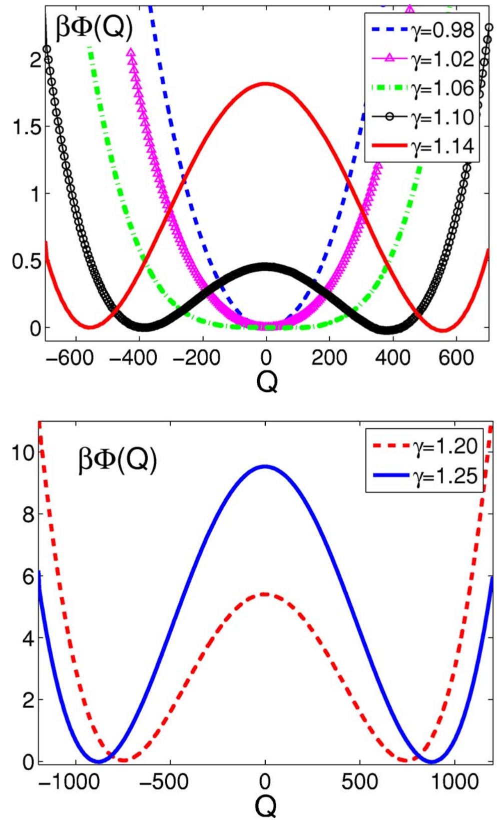

FIG. 7. The effective free energy Φ for different values of γ computed by our procedure. The other model parameter values are the same as in Fig. 5.

also be compared to the corresponding plot with $\gamma=1.14$ in Fig. 8, which can be viewed as the limit $K\rightarrow\infty$. As can be seen, the results given by stochastic model II for $K=10$ are already in good agreement with the corresponding results obtained by stochastic model I. Figure 11 shows similar results for $\gamma=1.20$. Again, we obtained accurate results using the equation-free method.

Up to now we have used the symmetric variable $Q\equiv P_1-P_2$ as our observable. However, we often do not have *a priori* knowledge of the slow variable or, more generally, of a good observable to parametrize the long-time system dynamics. To investigate the sensitivity of our results to the choice of observable, we repeated the computations on stochastic model I using P_1 instead of Q. To use P_1 as our observable, we modify step (1) so that we simply initialize P_2 using $P_2=\gamma/(\delta+\delta\omega P_1^2)$. The numerical results for different values of γ are given in Fig. 12. Again good agreement is seen between the equation-free method and the Monte Carlo simulations. Because P_1 has both a slow and a fast component, this result illustrates that equation-free methods may be attempted even when the slow variable is unknown. An extensive discussion of this point in a deterministic context can be found in Ref. 28: one does not necessarily need *the correct* slow variable—one needs an observable that *parametrizes* the slow manifold, a quantity in terms of which the slow manifold can be expressed as the graph of a function.

Encouraged by the success of our computational framework for the simple stochastic models I and II considered above, we next investigated how well these methods would work on the full system described by Eqs. (2.1)–(2.6). We first performed long-time Monte Carlo simulations using BIONETS.[14] A two-dimensional histogram for the total protein numbers $T_1=P_1+2\overline{P_1P_1}$ and $T_2=P_2+2\overline{P_2P_2}$ is shown in Fig.

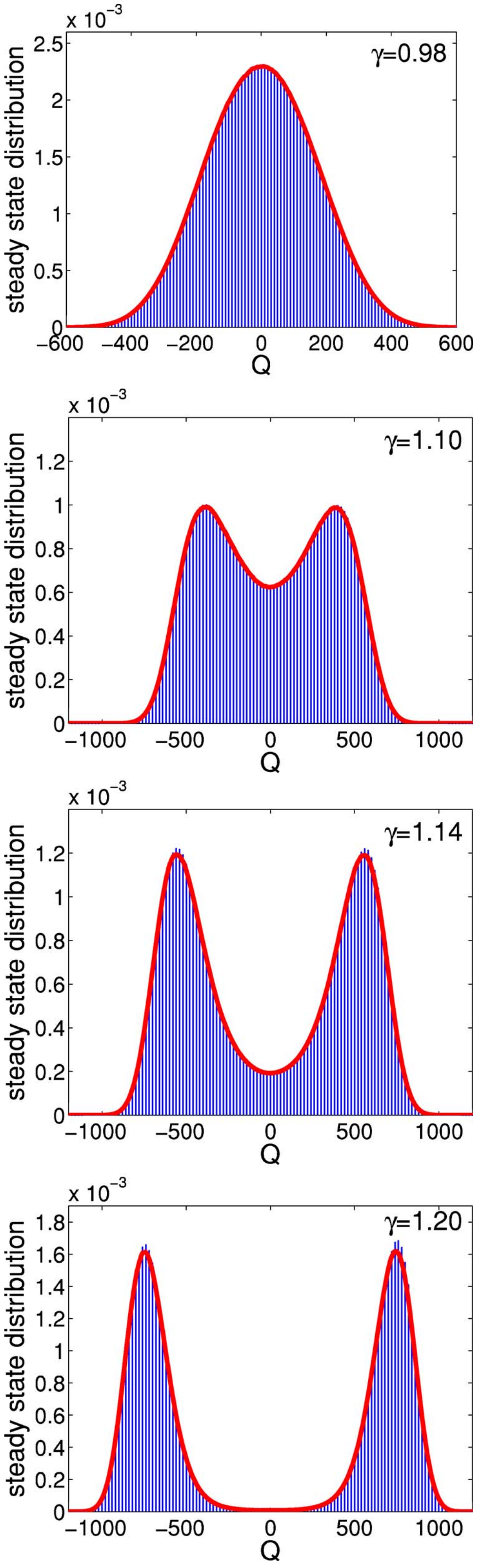

FIG. 8. Comparison of steady-state distributions obtained from the effective free energies shown in Fig. 7 (thick lines) with histograms obtained by long-time simulations.

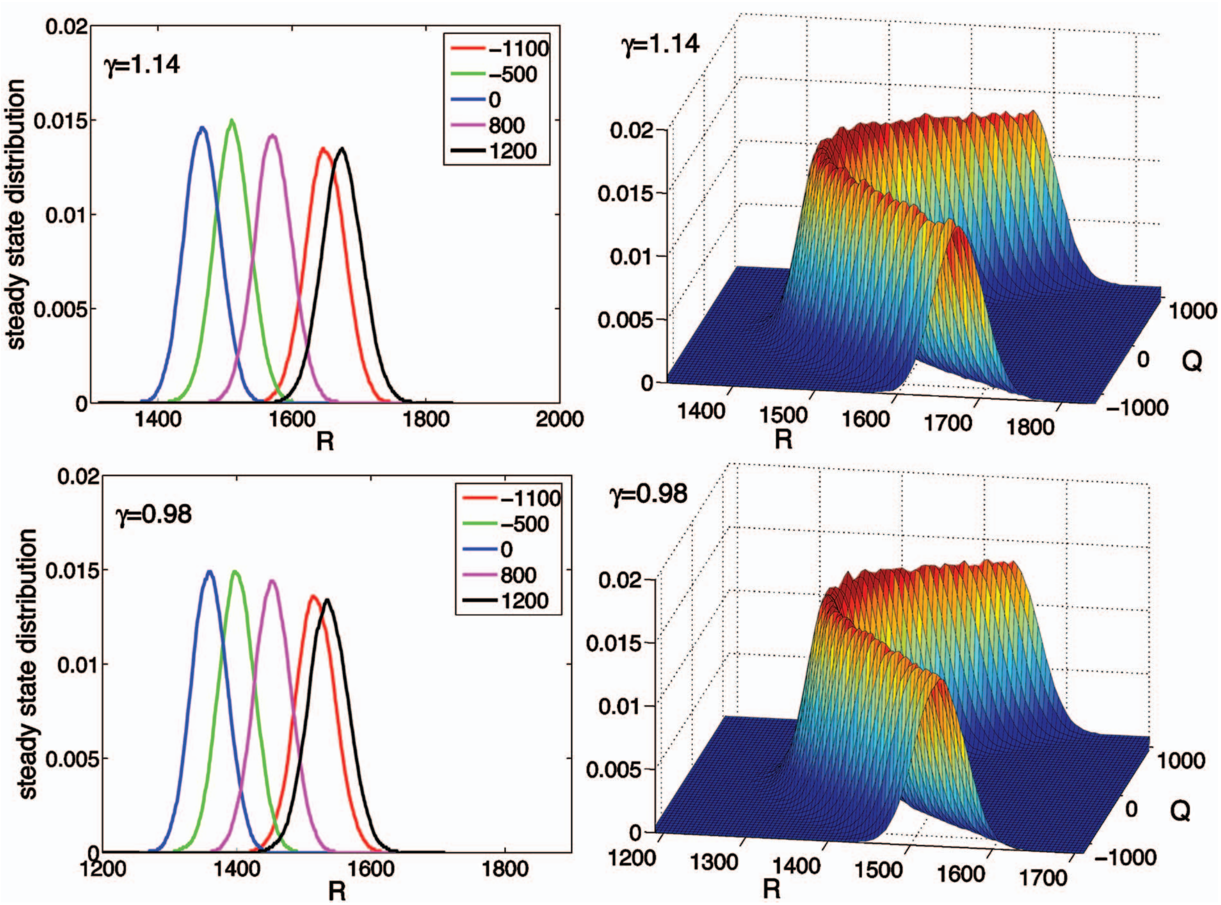

FIG. 9. (Color) Conditional distribution $P(r|Q=q)$ for stochastic model I. Pictures on the left show $P(r|Q=q)$ for selected values of q. Pictures on the right show $P(r|Q=q)$ as a function of r and q.

13(a). This simulation consisted of approximately 10^{13} Gillespie SSA steps and took over 500 total CPU hours (800 runs distributed over 18 CPUs). Each run resulted in approximately 150 transitions between the stable steady states. Therefore, Fig. 13(a) is a result of roughly 120 000 transitions. This number of transitions is probably excessive for most cases. However, if we reduce this number by a factor of 100, so that on the order of 1000 transitions are used to construct the distribution, it would still require 10^{11} Gillespie SSA steps and 5 h of CPU time. The dashed curve in Fig. 13(b) is the projection of the histogram shown in Fig. 13(a) onto the T_1 axis. The parameter values used to compute Fig. 13 are $\gamma_1=\gamma_2=1.14$, $\delta_1=\delta_2=7.5\times10^{-4}$, $\epsilon_1=\epsilon_2=0$, $k_1=k_2=5\times10^{-4}$, $k_{-1}=k_{-2}=1$, $k_{o1}=k_{o2}=0.004$, and $k_{-o1}=k_{-o2}=0.1$. These values are consistent with the parameter values $K=0.1$, $\delta=7.5\times10^{-4}$, and $\omega=2\times10^{-6}$ used in stochastic models I and II. Note that Fig. 13(b) is a plot of the total protein number, whereas the distributions shown in Fig. 12 are for the monomer number. Therefore these two figures are not directly comparable.

We next performed equation-free computations for the system. As our single observable, Q, we used the total protein number T_1, because this is a quantity that can be measured using single cell fluorescent techniques. We used a slightly modified version of step (3) to compute the conditional density $P(\mathbf{r}|T_1=t_1)$. For a given value of T_1, we set the rate constants for synthesis and degradation of this protein equal to zero. We then ran the simulations for a time of 1×10^5 to remove any transients. Next still keeping T_1 fixed, 10 000 samples of the other variables were collected at evenly distributed intervals over a time period of 2×10^5 and used to generate the conditional density. A time step of $\Delta t=15$ was used in step (B) of the algorithm. To compute the steady-state distribution, polynomials were fitted to the average velocity and effective diffusion coefficient computed from the equation-free analysis and then used to compute the effective free energy. The red solid curve shown in Fig. 13(b) is the result of the equation-free analysis. Very good agreement between the equation-free method and Monte Carlo simulation is seen. The simulations used to produce the equation-free result consisted of approximately 10^{10} Gillespie SSA steps and took less than an hour of CPU time. Note that in the equation-free method the number of Gillespie steps needed to compute the effective free energy is independent of the parameter γ. In contrast, as γ is increased the average transition time between the stable steady states grows exponentially (see Table I). Therefore, using direct simulation to compute the steady-state distribution for large values of γ quickly becomes impractical. Our investigations into these methods revealed that whereas the drift $V(q)$ is robust to changes in Δt, the effective diffusion coefficient $D(q)$ is quite sensitive and needs to be treated with care. Also, because of the exponential in the integral for the effective potential, small changes in the average drift or effective

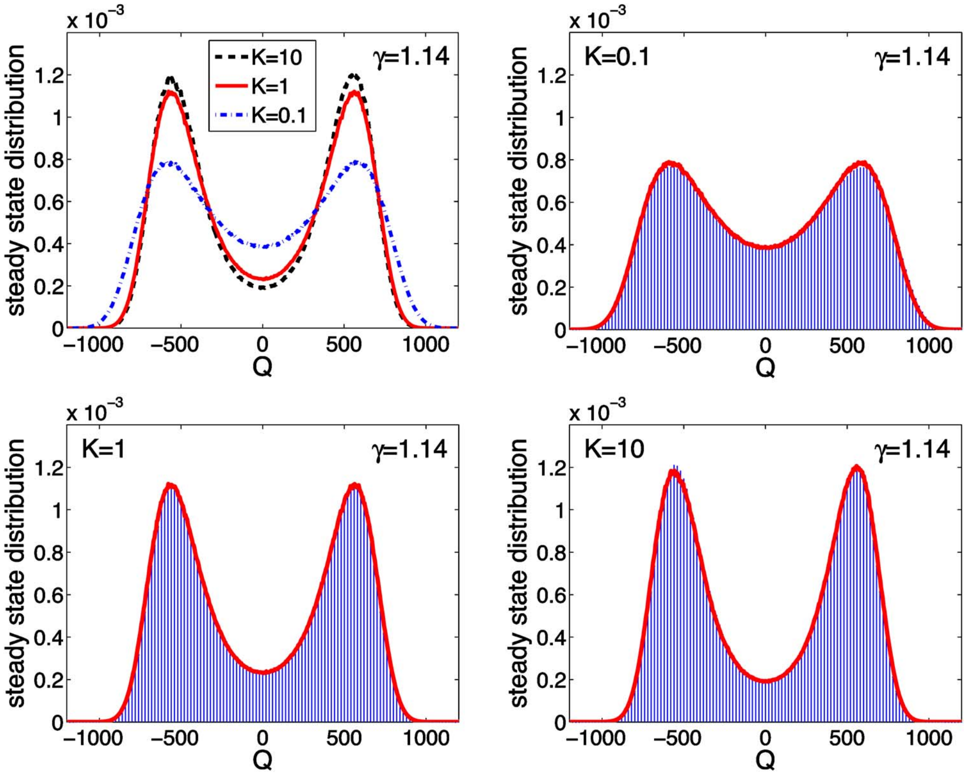

FIG. 10. Comparison of steady-state distributions from stochastic model II. The top left panel are results from the equation-free analysis for $K=0.1$, $K=1$, $K=10$ and $\gamma=1.14$. The remaining three panels compare these results (thick lines) to the steady-state distribution computed from long-time Monte Carlo simulation (histograms). The other model parameter values are the same as in Fig. 5.

diffusion coefficient can have large effects on the steady-state distribution. Therefore, it is important to average over sufficiently many realizations to ensure convergence of average drift and effective diffusion coefficient. Better estimation techniques, such as those developed by Aït-Sahalia using maximum likelihood,[29] should be incorporated in the data-processing step of the algorithms. Even with these caveats, the results presented in this section demonstrate the feasibility and high potential of equation-free methods for analyzing stochastic models of genetic networks.

B. First passage time

When γ is sufficiently large the system is bistable. An important characterization of bistable systems is the average time for noise-induced transitions between the stable states. Here we make use of the definition of the first passage time from Sec. III B. For the results presented in this section we use $Q=P_1-P_2$ as our observable and stochastic model I. The system is bistable for $\gamma > 1.06$. Let the deterministic stable steady states of P_1 be denoted as p_m and p_M with $p_m < p_M$. Because of the symmetry of our problem, p_m and p_M are also the stable steady states of P_2. Let the random variable $\mathcal{T}_e$ be defined as the first time when $P_1=P_2$ given the initial conditions $P_1=p_m$ and $P_2=p_M$. In terms of Q, this means that $\mathcal{T}_e$ denotes the time to reach $Q=0$ when the process starts with Q equal to the negative steady state $q_m \equiv p_m - p_M$. Let τ_e denote the average of $\mathcal{T}_e$. Then, direct Monte Carlo simulations can be used to compute the value of τ_e. The results of such simulations for three different values of γ are presented in Table I.

As expected, the computational time needed to compute the mean first passage time increases rapidly with γ. In Sec. III B, we introduced two formulas (3.6) and (3.8) to compute τ_e. Both formulas make use of the effective free energy com-

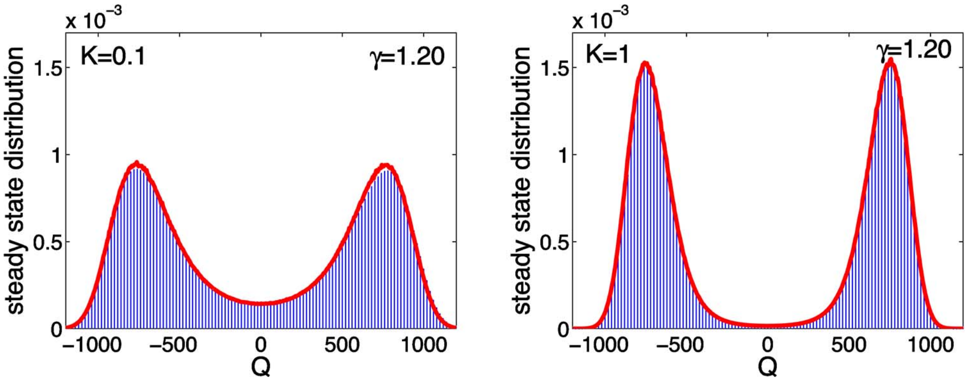

FIG. 11. Stochastic model II. Comparison of steady-state distributions obtained by equation-free analysis (thick line) with histograms obtained by long-time stochastic simulations. In this figure $\gamma=1.2$ and $K=0.1$ in the left panel and $K=1$ in the right. The other model parameter values are the same as in Fig. 5.

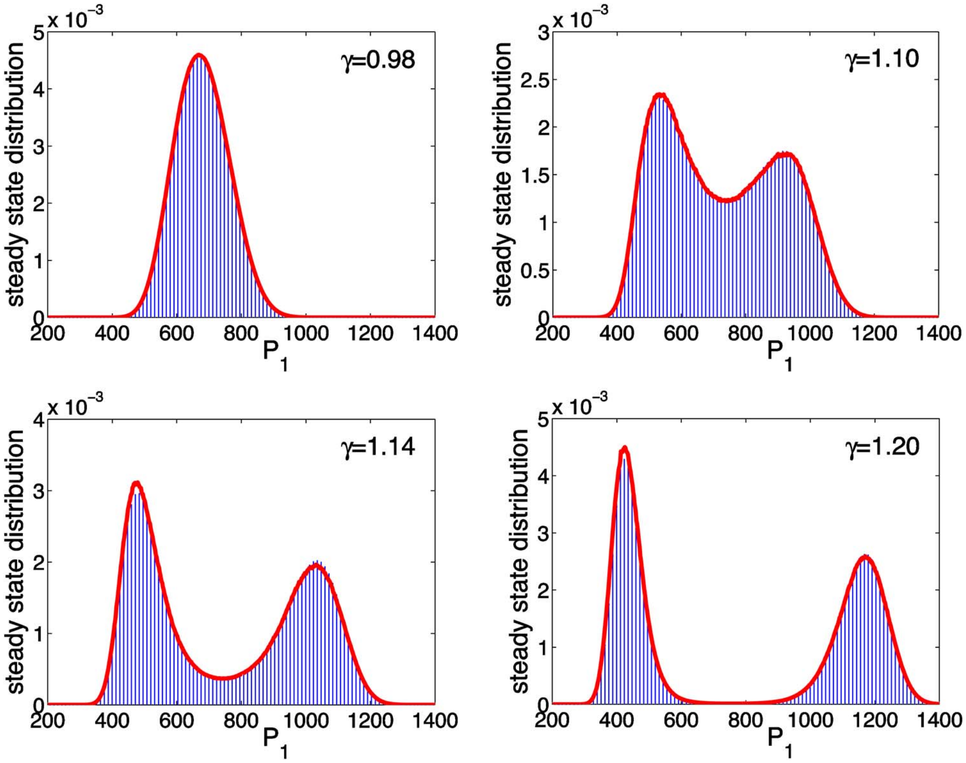

FIG. 12. Comparison of steady-state distributions using the variable P_1 as the observable for various values of γ. The other model parameter values are the same as in Fig. 5. Again, the thick lines are the results of equation-free analysis and the histograms are obtained by the long-time stochastic simulations.

puted by the equation-free algorithm. These potentials for $\gamma=1.14$, $\gamma=1.20$, and $\gamma=1.25$ are given in Fig. 7. Consequently, we can compare the results obtained by the long simulations with the results found from formulas (3.6) and (3.8) for $\tau_{e;p}$ and $\tau_{e;k}$, respectively. The results are shown in Table II.

Not surprisingly, the results given by $\tau_{e;p}$ are better than results given by the Kramers approximation $\tau_{e;k}$. However, both methods produce results that are within a factor of 2 of the waiting times estimated from Monte Carlo simulations. As γ becomes large the Monte Carlo simulations become computationally expensive. Therefore only 250 realizations were used to estimate the mean first passage time for $\gamma=1.25$, and we expect that the discrepancy between the Monte Carlo simulations and equation-free analysis for this case is due to finite sampling errors. Initializing the simulation at conditions that are rarely visited by the direct simulation itself constitutes a form of bias; this bias is designed to give faster computational estimates of the effective potential and—through this—of the first passage times. Clearly, this approach hinges on knowledge of a good observable and, in principle, does not depend strongly on the value of the parameter γ; therefore, the larger the parameter γ the higher the computational speedup in the first passage time estimation that will result. A quantitative study of this speedup is underway and will be reported elsewhere; it does not lie within the scope of this paper. We stress, however, that (as in molecular-dynamics simulations) knowledge of a good observable (a good "reaction coordinate") is crucial for the success of the approach.

Note that formula $\tau_{e;k}$ requires estimates of the second derivative of the potential at points q_u and q_m. To do this, we

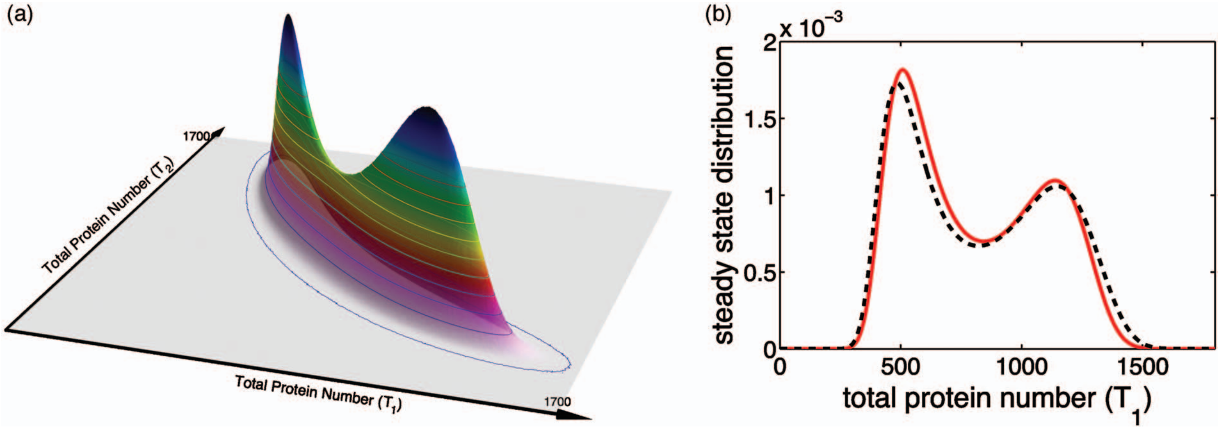

FIG. 13. (Color) (a) The steady-state distribution for the total protein numbers computed from long-time Monte Carlo simulations of the full model (2.1)–(2.6). (b) The projection of the two-dimensional (2D) distribution onto the T_1 axis (dashed curve). The red solid curve is the result of the equation-free analysis. The model parameter values are given in the text.

TABLE I. The mean first passage time computed from long-time stochastic simulations, averaging over N transitions. The results are expressed in the form ([sample mean]±[sample variance]/$\sqrt{N}$).

γ	p_m	p_M	$q_m = p_m - p_M$	Computed τ_e from simulations	N
1.14	481.1	1038.6	−557.5	$7.0\times10^5 \pm 6.7\times10^3$	10 000
1.20	425.8	1174.2	−748.4	$1.6\times10^7 \pm 1.6\times10^5$	10 000
1.25	392.4	1274.3	−881.9	$1.0\times10^9 \pm 6.3\times10^7$	250

fitted $\beta\Phi(q)$ locally to a polynomial and used the derivatives of the polynomial at the required points; once again, maximum-likelihood techniques (e.g., Ref. 29) should be used for better results. The formula for $\tau_{e;p}$ requires the evaluation of an indefinite integral. The integral was approximated by considering only a finite interval that neglected contributions from the region of sufficiently small q where the potential Φ is very large.

C. Bifurcations

In this section, our goal is to run the simulations for short times only and compute a form of "stochastic bifurcation diagram" using continuation methods, as an extension of the deterministic bifurcation computations. We use stochastic model I and study the dependence of the "steady states" on γ; the steady states we report are the fixed points of the algorithm from Sec. III A with the conditional density $P(R|Q=q)$ approximated by the Dirac delta function in (i) and (ii), similar to approach (1) from Sec. III. Numerical results are given in Fig. 14. For comparison we also plot the steady states of the corresponding deterministic equation (compare with Fig. 4). The plot in Fig. 14 was computed by initializing on different branches far from the bifurcation point and continuing from these different initializations (our simple arclength continuation algorithm did not include a "pitchfork detection" component).

The accuracy of the numerical results depends on several factors: the estimation technique for the Jacobian elements, the tolerance of the error for Newton-Raphson iterations, the number of realizations which are used to evaluate F, the time interval Δt, and the steepness of the underlying potential Φ. As can be seen in Fig. 14, stochasticity along with all these numerical factors have slightly perturbed the pitchfork bifurcation; this could be exacerbated by our choice of (symmetric or asymmetric) observable. It is easy to follow any branch of steady states far from the bifurcation point. For obvious reasons this becomes more complicated when we are close to the "bifurcation point" at $\gamma=1.06$. The main problem is that the potential becomes "flat" close to the bifurcation point—see Fig. 7. One way to improve the results is to adaptively change the number of realizations in (ii). That is, if the Newton-Raphson iterations of (3.5) do not converge to a desired tolerance, then more realizations are added. Another approach is to estimate directly a local polynomial model of the underlying diffusion process from discrete SSA data using maximum-likelihood tools and then search for the bifurcations of the critical points of the effective potential. Indeed, one can plot the zeros of the estimated drift, or—in the case of a state-dependent diffusion coefficient—one can correct them to report the maxima of the steady-state distribution;[19,20] both of these are good candidate bifurcation diagrams for the stochastic case. When the potential is steep and the equilibrium is "less noisy" it is not necessary to use many realizations; the relation between computational effort (in terms of number of replicas, simulation time horizon, and estimation method) and resulting accuracy is, again, a subject of current investigation beyond the scope of this paper.

TABLE II. Comparison of the mean first passage time computed from equation-free analysis with long-time stochastic simulations.

γ	τ_e from Table I	$\tau_{e;p}$ given by (3.6)	$\tau_{e;k}$ given by (3.7)
1.14	7.0×10^5	6.1×10^5	1.3×10^6
1.20	1.6×10^7	1.4×10^7	2.6×10^7
1.25	1.0×10^9	6.7×10^8	1.2×10^9

Finally, the results using P_1 instead of Q as the observable are shown in Fig. 15. In this case, the asymmetry of our observable and the perturbation it causes on the initialization process make the perturbation of the pitchfork bifurcation stronger. We used symmetric rate constants in our model as a means to simplify the system by reducing the number of model parameters. However, biological systems are very unlikely to possess such symmetries, in which case the artificial broken symmetry seen in Fig. 15 that results from using a numerical method that does not preserve this symmetry would not be an issue. Of course, the results also depend on the initialization procedure, our estimation technique, the error tolerance, the number of realizations, the length of time step Δt as well as the type of continuation algorithm we are using (here we used a very simple one, without bifurcation detection, in order to demonstrate what is possible). Accurate bifurcation detection depends on accurate Jacobians and even higher derivatives; estimating these from dynamic (and

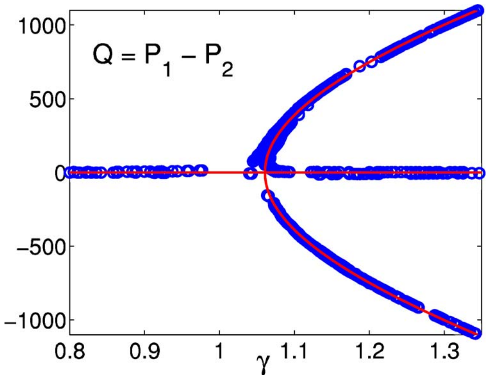

FIG. 14. A plot of the steady states obtained by equation-free analysis (3.5) (circles). Also shown are the deterministic steady states from Fig. 4 (solid line).

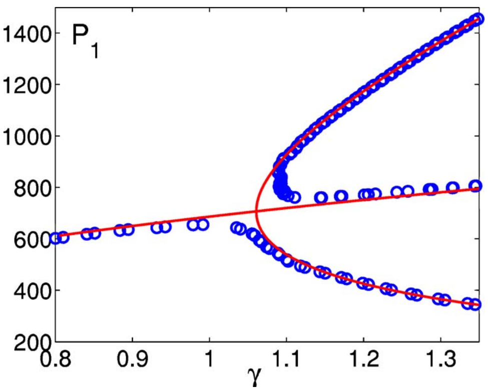

FIG. 15. Plot of the steady states obtained by using P_1 as the observable (circles). Again for comparison the deterministic case is shown as the solid line.

noisy) data is notoriously difficult. While conceptually we do have the tools to "hone in" the more accurate detection of bifurcation points, careful quantitative work is necessary to pin down the trade-offs between computational effort, model estimation accuracy, and bifurcation point estimation accuracy.

VI. DISCUSSION

In this paper we discussed and illustrated the use of certain equation-free numerical techniques that have the potential to accelerate the computer-assisted analysis of stochastic models of regulatory networks. There is a clear current need for accelerating such simulations; even for modestly complex regulatory networks, stochastic models rapidly become computationally expensive. Computational acceleration is usually based on model reduction; theoretical methods for stochastic model reduction that take advantage of a separation of time scales are the focus of intense current research.[4,6–8,30,38] As we discussed in the Introduction, many important gene regulatory networks do satisfy this assumption of a separation of time scales because synthesis and degradation of new proteins and transcripts usually occur on a slower time scale than processes that change the chemical state of proteins. Analytical model reduction techniques assume that the fast variables are in quasisteady state with respect to the slow variables and use the quasi-steady-state distributions conditioned on the slow variables to eliminate the fast variables by averaging. These methods have been successfully applied to simple models, but the theory is not as well established as the deterministic counterpart. Having an explicit model lies often at the basis of such stochastic reduction methods.

In the equation-free approach many of the same elements (separation of time scales and approximation of conditional quasisteady distributions) also underpin computational efficiency; but the basic premise is that the model is available in the form of a "black box" simulation code. We do not try to first reduce and then simulate the reduced surrogate; we try to design the smallest number of "intelligent" short computational experiments with the full stochastic model to find the quantities of interest, whether these are steady-state probability distributions, their maxima, or transition rates in the bistable case. In that sense, the approaches we described here *do not* hinge upon the "inner," detailed simulator as being a Gillespie SSA one—the methods are equally applicable to any inner simulator, stochastic or deterministic, as long as the main assumption of a low-dimensional *effective* stochastic model is a good one for the long-term system dynamics. Indeed, if another reduction method can be used to produce a good approximate dynamic simulator, our algorithms can be "wrapped" around this surrogate simulator rather than the full model for further acceleration.

Another important point has to do with the type of computation we are interested in—do we want to accelerate the *direct simulation* of the model, or do we want to accelerate the computation of certain features of its long-term dynamics (e.g., of the maxima of the steady-state distribution)? These latter quantities can also be obtained from long-term direct simulation, but one of the points that we want to stress is that we can link direct simulation to *different* numerical algorithms (such as contraction mappings and continuation methods) to obtain these quantities, often faster than with direct simulation alone. In the same way that bifurcation diagrams for dynamical systems are usually *not computed* through direct ODE integration, but through bifurcation algorithms, the parametric dependence of the long-term dynamics of stochastic models does not have to be computed through long-time direct simulation only. This "alternative" acceleration, not through accelerating the direct simulation itself, but through linking it to different numerical algorithms, lies at the basis of the equation-free framework.

Having said this, we briefly mention that equation-free methods for accelerating the direct simulation itself also exist. *Coarse projective integration* which uses short bursts of direct simulation to estimate time derivatives of evolving probability densities and then passes them to standard numerical integration algorithms, has been successfully used in many contexts.[9–11,31] Coarse-projective integration has a strong relation to the direct simulation acceleration methods in Ref. 32; it has not been discussed in this paper, because we chose to focus on *very long-term* features of the network dynamics; it might interest the reader that the method can be used to also integrate backward in time and solve "effective boundary-value problems" to find "coarse" limit cycles.[33]

In the equation-free methods for analyzing stochastic models of gene regulation that we discussed in this paper, we have tried to circumvent the difficulties encountered by direct simulation (in this case SSA) through the design of short bursts of *appropriately initialized* computational experiments with the full simulator. In a sense, we "resign ourselves" to the fact that the direct simulator is expensive; we ask what is the shortest amount of running of this expensive direct simulator in order to obtain the quantities we are interested in. The "design of experiments" protocols are templated on traditional continuum numerical methods, such as the fixed-point and continuation algorithms to compute bifurcation points, or quadrature to estimate Kramers formula. The only difference is that the quantities (residuals, actions of Jacobians, and values of the integrand) that are required for numerical computation are not given by a closed formula, but rather through direct numerical simulation of the full model and estimation. We reiterate once more that these techniques can be wrapped around the full direct simulator, or our best available reduction of it, without change.

Knowing appropriate coarse-grained observables (the variables in terms of which the unavailable effective model would be written) is an important feature of the algorithms. Extensive experience with the problem, intuition, or analytical work may often suggest such observables; we did take advantage of such knowledge in this paper. We did already demonstrate an important point: more than one observables are capable of doing a good job as the parametrizing variables in an equation-free context; one does not need to know *the exact* slow variables. This issue is discussed extensively for the deterministic context in Ref. 28. It is, however, important to note that algorithms for the detection of low dimensionality in high-dimensional data can be vital in *suggesting* such observables from simulations. Principal component analysis is an established linear method for the detection of appropriate lower-order observables from simulation data; numerically estimated eigenmodes of the problem may also provide good observables (see the discussion in Ref. 12 about estimating gaps between eigenvalues and using them to decide whether we should include more observables as independent variables). There are, however, some important developments in this area: the recent use of harmonic analysis (geometric diffusion) on graphs constructed from high-dimensional data shows great promise in detecting good observables (reaction coordinates) for complex, high-dimensional systems.[34–36] This "variable-free" approach can be naturally linked to equation-free computation (one designs computational experiments both to detect the appropriate observables *and* to do computations with them);[34] we are currently working on demonstrating this link for gene regulatory network modeling.

It is interesting to observe that the dimensionality of the fine scale model does not, in principle, affect the complexity of the equation-free protocols; what is important is the dimension of the effective free-energy surface, and not how many variables the detailed model contains. This is analogous to the dimensionality of normal forms close to bifurcations—it is the number of "slow eigenvalues" that matter, and not the dimensionality of the problem in which the bifurcation occurs. Similarly, what is important here is the number of coarse variables (one in our case, both for the detailed and the simplified models) and not the detailed model dimension.

It is clear that, in certain cases, an equation-free computational approach is expected to have advantages over direct simulation. For steep potentials and low noise, for example, the way equation-free computation uses a good observable to bias the simulation will sample the effective potential and give a good estimate of the transition rates much faster than direct simulation. Also, parametric analysis methods should be able to explore parametric transitions faster and more systematically than direct simulation, in analogy with the use of bifurcation techniques rather than direct simulation in deterministic dynamical systems (e.g., by writing augmented algorithms that converge on marginally stable or unstable solutions). The complexity of the computation depends crucially on the dimensionality of the unavailable *reduced* model, and not so crucially on the dimensionality of the detailed, full model. We are currently working on the quantification of these computational benefits; this work is complicated by the fact that—lacking explicit formulas from which to obtain derivative information—errors must be computed online through *a posteriori* estimates.

This brings us to a final, yet vital issue: estimation. Given the noisy nature of the data, estimating the numerical quantities of interest lies at the heart of the accuracy (and thus the viability) of the computation. For our gene networks, these quantities included the effective potential $\Phi(q)$ and the effective diffusion coefficient $D(q)$. Preliminary investigations revealed that the effective diffusion coefficient $D(q)$ is quite sensitive to the time step Δt and needs to be treated with care. Also, small changes in the average velocity or effective diffusion coefficient can have relatively large effects on the steady-state distribution. Even though some computations are "embarrassingly parallel" (one short, fine scale realization per processor, running independently) variance reduction becomes an important feature (see, e.g., Ref. 37). Maximum-likelihood estimation techniques (e.g., Ref. 29) take the place of simple formulas such as (3.2) and (3.3); one can envision certain hypothesis testing computations (is our model locally well approximated by a diffusion process?) becoming part of the overall computational scheme. Until these elements, and their computational cost, are analyzed and tested, there will be no firm guarantees for the computational efficiency of equation-free methods. Yet, even with these caveats, as we computationally demonstrated in this paper, we believe that the equation-free framework provides a promising approach to gene regulatory network modeling, alternative to long-direct simulation. It links directly with powerful and tested traditional continuum numerical algorithms (such as numerical integration, fixed-point algorithms, and matrix-free iterative linear algebra) and with system theory techniques such as filtering and estimation. These techniques are, in some sense, "off the shelf" and do not need to be redeveloped. In our opinion, it is the linking of equation-free techniques with novel data reduction/clustering techniques (such as the use of the graph Laplacian to detect good reaction coordinates[34]) that holds the most promise in the computational study of complicated stochastic systems in general and of gene regulatory networks and their models in particular.

Finally, we return to the important issue of the efficiency of the equation-free computations. We have not touched this issue much in this paper—we had the luxury to be able to use enough brute force SSA time steps to obtain accurate results, to be compared with those of equation-free methods. Decreasing the number of SSA time steps/number of realizations of the stochastic process, we will increase the error of computations. For example, let us consider the problem from Fig. 11 (panel on the right). Let us define the error of the long-time or equation-free computations as the L^1 norm of the difference between the computed steady-state distribution and the exact steady-state distribution which can be obtained by using a very long stochastic simulation (we used 2^{40} SSA time steps). To get the error of long-time stochastic simulations below 10%, we found out that we need about $2^{32} = 4.3 \times 10^9$ SSA time steps. On the other hand, we can achieve an error of 6% through equation-free computation using only 2.4×10^8 time steps. So, the tenfold gain of the method can be seen for the simulation of the problem from Fig. 11 (panel on the right).

In general, the computational efficiency of equation-free methods depends both on the model under consideration (e.g., rate constants of the stochastic model) and on "numerical" parameters associated with the approach (e.g., discretization of the effective potential, number of realizations for variance reduction, estimation technique, and quality of the random number generator). It also depends on the type of computational task we are considering (equilibrium distribution computation, escape time computation, and optimal transition path computation). A careful study of these issues is beyond the scope of this paper and the subject of ongoing research. It is clear that in the presence of high barriers between deep effective wells, the approach we have discussed has the potential to significantly accelerate the computation of macroscopic quantities of interest. We believe that linking this approach with parametric dependence analysis (along the lines of the coarse-grained bifurcation algorithms discussed here) holds additional promise for the use of equation-free methods.

ACKNOWLEDGMENTS

One of the authors (I.G.K.) acknowledges partial support by DARPA and a NSF/ITR grant. Another author (R.E.) acknowledges partial support by the Biotechnology and Biological Sciences Research Council, University of Oxford and University of Minnesota. Another author (T.C.E.) also acknowledges support by DARPA (F30602-01-2-0579).

[1] T. Schlitt and A. Brazma, FEBS Lett. **579**, 1859 (2005).
[2] J. Hasty, D. McMillen, F. Isaacs, and J. Collins, Nat. Rev. Genet. **2**, 268 (2001).
[3] M. Kaern, T. Elston, W. Blake, and J. Collins, Nat. Rev. Genet. **6**, 451 (2005).
[4] T. Kepler and T. Elston, Biophys. J. **81**, 3116 (2001).
[5] D. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[6] Y. Cao, D. Gillespie, and L. Petzold, J. Chem. Phys. **122**, 14116 (2005).
[7] E. Haseltine and J. Rawlings, J. Chem. Phys. **117**, 6959 (2002).
[8] C. Rao and A. Arkin, J. Chem. Phys. **118**, 4999 (2003).
[9] I. Kevrekidis, C. Gear, J. Hyman, P. Kevrekidis, O. Runborg, and K. Theodoropoulos, Commun. Math. Sci. **1**, 715 (2003).
[10] R. Erban, I. Kevrekidis, and H. Othmer, Physica D (in press).
[11] C. Gear, I. Kevrekidis, and C. Theodoropoulos, Comput. Chem. Eng. **26**, 941 (2002).
[12] C. Siettos, M. Graham, and I. Kevrekidis, J. Chem. Phys. **118**, 10149 (2003).
[13] M. Haataja, D. Srolovitz, and I. Kevrekidis, Phys. Rev. Lett. **92**, 160603 (2004).
[14] D. Adalsteinsson, D. McMillen, and T. Elston, BMC Bioinf. **5**, 1 (2004).
[15] T. Gardner, C. Cantor, and J. Collins, Nature (London) **403**, 339 (2000).
[16] J. Hasty, D. McMillen, and J. Collins, Nature (London) **420**, 224 (2002).
[17] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer-Verlag, Berlin, 1989).
[18] G. Hummer and I. Kevrekidis, J. Chem. Phys. **118**, 10762 (2003).
[19] D. Kopelevich, A. Panagiotopoulos, and I. Kevrekidis, J. Chem. Phys. **122**, 044908 (2005).
[20] S. Sriraman, I. Kevrekidis, and G. Hummer, Phys. Rev. Lett. **95**, 130603 (2005).
[21] E. Doedel, H. Keller, and J. Kernevez, Int. J. Bifurcation Chaos Appl. Sci. Eng. **1**, 493 (1991).
[22] E. Doedel, H. Keller, and J. Kernevez, Int. J. Bifurcation Chaos Appl. Sci. Eng. **1**, 745 (1991).
[23] A. Makeev, D. Maroudas, and I. Kevrekidis, J. Chem. Phys. **116**, 10083 (2002).
[24] A. Makeev, D. Maroudas, A. Panagiotopoulos, and I. Kevrekidis, J. Chem. Phys. **117**, 8229 (2002).
[25] C. Kelley, *Solving Nonlinear Equations with Newton's Method* (SIAM, Philadelphia, PA, 2003).
[26] D. Barkley, I. Kevrekidis, and A. Stuart, SIAM J. Appl. Dyn. Syst. (submitted).
[27] D. Gillespie, *Markov Processes; An Introduction for Physical Scientists* (Academic, New York, 1992).
[28] C. Gear, T. Kaper, I. Kevrekidis, and A. Zagaris, SIAM J. Appl. Dyn. Syst. **4**, 711 (2005); e-print physics/0405074.
[29] Y. Aït-Sahalia, Econometrica **70**, 223 (2002).
[30] H. Salis and Y. Kaznessis, J. Chem. Phys. **122**, 054103 (2005).
[31] C. Gear, Report No. NEC TR 2001-130, 2001 (unpublished), pp. 1–9.
[32] D. Gillespie, J. Chem. Phys. **115**, 1716 (2001).
[33] R. Rico-Martinez, C. Gear, and I. Kevrekidis, J. Comput. Phys. **196**, 474 (2004).
[34] B. Nadler, S. Lafon, R. Coifman, and I. Kevrekidis Appl. Comput. Harmon. Anal. (in press).
[35] M. Belkin and P. Niyogi, Neural Comput. **15**, 1373 (2003).
[36] B. Nadler, S. Lafon, R. Coifman, and I. Kevrekidis, NIPS 2005 Proceedings; e-print math-ph/0506090.
[37] M. Melchior and H. Öttinger, J. Chem. Phys. **103**, 9506 (1995).
[38] W. E, D. Liu, and E. Vanden-Eijnden, J. Chem. Phys. **123**, 194107 (2005).