

A variational approach to the stochastic aspects of cellular signal transduction

Yueheng Lan

*Department of Chemistry, University of North Carolina at Chapel Hill,
Chapel Hill, North Carolina 27599-3290*

Peter G. Wolynes

*Department of Chemistry and Biochemistry, University of California at San Diego,
La Jolla, California 92093-0371*

Garegin A. Papoian^{a)}

*Department of Chemistry, University of North Carolina at Chapel Hill,
Chapel Hill, North Carolina 27599-3290*

(Received 18 July 2006; accepted 18 August 2006; published online 27 September 2006)

Cellular signaling networks have evolved to cope with intrinsic fluctuations, coming from the small numbers of constituents, and the environmental noise. Stochastic chemical kinetics equations govern the way biochemical networks process noisy signals. The essential difficulty associated with the master equation approach to solving the stochastic chemical kinetics problem is the enormous number of ordinary differential equations involved. In this work, we show how to achieve tremendous reduction in the dimensionality of specific reaction cascade dynamics by solving variationally an equivalent quantum field theoretic formulation of stochastic chemical kinetics. The present formulation avoids cumbersome commutator computations in the derivation of evolution equations, making the physical significance of the variational method more transparent. We propose novel time-dependent basis functions which work well over a wide range of rate parameters. We apply the new basis functions to describe stochastic signaling in several enzymatic cascades and compare the results so obtained with those from alternative solution techniques. The variational *Ansatz* gives probability distributions that agree well with the exact ones, even when fluctuations are large and discreteness and nonlinearity are important. A numerical implementation of our technique is many orders of magnitude more efficient computationally compared with the traditional Monte Carlo simulation algorithms or the Langevin simulations. © 2006 American Institute of Physics.
[DOI: [10.1063/1.2353835](https://doi.org/10.1063/1.2353835)]

I. INTRODUCTION

The life of the cell is regulated by intricate chains of chemical reactions.¹ The whole cell may be viewed as a computing device where information is received, relayed, and processed.² Signal transduction cascades based on protein interactions regulate cell movement, metabolism, and division.^{1,3} Since cells are mesoscopic objects, understanding the role of the intrinsic fluctuations of the biochemical reactions as well as environmental fluctuations is a fundamental part of understanding signaling dynamics.^{4–15} In this regard, the well-organized behavior of cells, which emerges as a result of biochemical reaction dynamics involving hundreds of cross-linked signaling pathways, is remarkable.^{16–22} The problem of how signals can be precisely detected, smoothly transduced, and reliably processed under noisy conditions is a research topic of great current interest that, in turn, should lead to deeper understanding of the origins of the cell's functional responses.^{23,24} Furthermore, these studies can help unravel the design principles for various signaling pathways,

leading, eventually, to better ways to control and efficiently interfere with cellular activity, as would be needed to correct the behavior of diseased cells.^{18,25}

The role of noise in gene regulatory networks has been identified as a key issue and has been intensively studied in recent years.^{10–12,26–31} Linearization of the noise may be acceptable if the dynamics near steady states is being studied.^{10,26,31} When protein numbers are large and, thus, the continuous approximation is valid, time-dependent distributions have been determined using the Langevin or Fokker-Planck equations.^{6,32,33} To account for the discreteness in the linearized equations, the generating function approach has also been used.^{10,26} A variational treatment of steady state stability and switching in nonlinear, discrete gene regulatory processes has been reported.^{29,30}

In cytosolic signal transduction processes, in contrast to gene transcription which involves a unique DNA molecule, all the reacting species are present in multiple copies and participate in unary, binary, or perhaps even higher order reactions. Noise could be multiplicative^{34,35} and the linear description easily breaks down. Moreover, cellular reactions usually take place heterogeneously in space. The localization and compartmentalization of protein organelles require dif-

^{a)}Electronic mail: gpapoian@unc.edu

fusive or active transportation of reacting molecules from one region to another. Spatial coordination combined with temporal coordination generates coherent, yet complex spatiotemporal patterns.^{18–22,36–38}

The extracellular ligands often trigger cascades of chemical reactions which propagate inside a cell and induce responses from various environmental cues. The cell body is a highly heterogeneous entity and never settles to a steady state. To understand cell dynamical processes, an explicitly time-dependent description is required. Within a volume with linear dimensions of the Kuramoto length,^{34,39} diffusion mixes the reagents in a nearly uniform manner. If the reactions are considered in the Kuramoto volume, it is reasonable to neglect the spatial heterogeneity. For many signal transduction networks, however, it is likely that only a few proteins are present in the Kuramoto volume (determined by specific reaction and diffusion rates), and therefore, the continuous description of protein numbers breaks down. To characterize stochastic signaling reactions in this volume, a time-dependent description of a noisy, discrete, nonlinear system is required.⁴⁰ In many situations, such as *Drosophila* oogenesis, the exact shape of the probability distribution profile is very important and determines different developmental paths.^{38,41} In the following, we discuss efficient techniques to compute the time-dependent protein number probability distributions in biochemical reaction networks when the number of protein copies is small.

The Gillespie algorithm provides an effective Monte Carlo technique for simulating stochastic chemical reactions.^{42–44} Each simulation gives a reaction trajectory which is close to the deterministic trajectory in the large particle number limit. To get well-converged statistics, many trajectories may be needed, often on the order of 10^5 . If there is a separation of time scales of the constituent chemical reactions, Gillespie simulations also become exceedingly slow since the reaction events are dominated by the fastest reactions while the observables typically involve the slowest reactions. Although considerable progress has been made in accelerating such simulations,^{6,32,33,45} computational inefficiency continues to be an impediment, especially for the spatially inhomogeneous generalization of the Gillespie algorithm. Furthermore, it is hard to extract the analytical structure of the solution from the numerical results, which can be important for achieving a deeper physical understanding of the system behavior when the rate parameters are widely varied.

Mathematically, a stochastic process may be completely characterized by a master equation—a group of ordinary differential equations (ODEs) describing the evolution of probabilities.^{34,46} The main difficulty in solving a chemical master equation is the enormous number of ODEs involved even for a small reaction cascade. A number of analytical techniques have been developed for solving approximately the master equation.^{10,26,28} In this work, we show how to achieve enormous reduction in the dimensionality of specific reaction cascade dynamics by solving variationally the quantum field theory (QFT) equations of stochastic chemical kinetics.^{47–51} Our present approach is based on mapping the master equation ODEs into a single partial differential equa-

tion (PDE) and applying a variational technique which reduces the PDE into a small number of ODEs. The variational QFT approach has been employed to study steady state stability and switching in gene regulatory networks.^{29,30} In this work we propose novel time-dependent basis functions appropriate for describing protein signaling cascades which work in a wide range of rate parameters. Our method gives probability distributions that agree well with the exact ones, including when the fluctuations are large and discreteness and nonlinearity play important roles.

The paper is organized as follows. In Sec. II, the QFT formulation of the stochastic processes describing chemical reactions^{29,30,47,50–52} and the variational solution technique of Eyink and Alexander and Eyink for solving such field theories^{53,54} are presented. We show that the QFT formulation is equivalent to a generating function approach and also discuss the physical significance of the variational principle in this context. In Sec. III, we apply the new trial functions and the variational technique to a number of two-step, three-step, and four-step enzymatic reaction cascades and compare our results with those found with more traditional methods. We also discuss the more general principles of basis function construction and the limitations of variational approaches. Finally, we provide in the Appendix a pedagogical treatment of a simple binding process to further elucidate the details of the variational approach.

II. QUANTUM FIELD THEORY FORMULATION, VARIATIONAL PRINCIPLE, AND GENERATING FUNCTIONS

In this section, we first discuss briefly the master equation and demonstrate its application to a two-step signal amplification cascade. Next, the master equation is recast into a QFT form in which the probability evolution is governed by a “wave equation.” Then, we show that the field theoretic formulation is equivalent to a generating function approach. To solve these equations, the variational technique of Eyink and Alexander and Eyink and its physical significance are examined. We further explore the practical implementation issues in Sec. III A.

A. The master equation and its solution

Unlike a stochastic simulation which produces an individual random trajectory and generates statistics only after a large number of samplings, master equations directly describe the evolution of probability distributions in the state space of a system based on specific interstate transition rates. For a discrete system, the master equation consists of a set of ODEs (see the following examples), while for a continuous system it becomes an integrodifferential equation such as the Boltzmann equation. Although the master equation is a complete description of a Markovian system, its solution is usually difficult and requires special techniques. This paper presents one variational technique that could be used.

As an example, let us consider the following set of equations that represents the simplest enzymatic signal amplification process (Fig. 1). In this simple reaction scheme, without feedback loops, R represents an inactive receptor, which be-

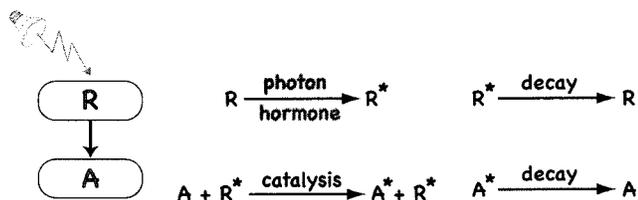


FIG. 1. An inactive receptor R , when activated by a signal, activates downstream protein A .

comes activated into R^* upon binding of an external ligand (stimulus). When the receptor is activated, it acts as an enzyme, catalyzing the phosphorylation of the next kinase downstream ($A + R^* \rightarrow A^* + R^*$) with a rate μ . A^* spontaneously decays to A with a rate λ and R^* to R with a rate k . In the absence of R^* , $A \rightarrow A^*$ may occur naturally, however, with a much lower rate, so that it can be ignored when we introduce the catalyst R^* .

Although the R^* reaction is unary and independent of the A reaction, the latter one is binary, making the system non-linear, thus, different from those considered in a number of prior works on the gene regulatory networks.^{10,11,26,27} To write the master equation, we denote by $P(m, n)$ the probability of having m R^* 's and n A 's, then

$$\begin{aligned} \frac{dP}{dt}(m, n) = & \mu[-mmP(m, n) + m(n+1)P(m, n+1)] \\ & + \lambda[-(N-n)P(m, n) + (N-n+1)P(m, n-1)] \\ & + g[-P(m, n) + P(m-1, n)] \\ & + k[-mP(m, n) + (m+1)P(m+1, n)], \end{aligned} \quad (1)$$

where N is the total number of A and A^* . In Eq. (1), the first two terms describe the $A-A^*$ reaction and the rest the $R-R^*$ reaction. This simple two-step cascade is commonly found embedded in the onset of a reaction pathway of many important signaling cascades.^{55,56} If a large number of inactive receptors R are present the rate of conversion depends on the arrival times of the external cue and the reaction becomes Poissonian. We assume that this is the case in all the following calculations. If the $R \rightarrow R^*$ reaction is the usual birth-death problem, our formalism still applies with only minor changes. The master equation (1) actually contains infinitely many coupled ODEs.

B. The QFT formulation

The differential-difference equations, such as Eq. (1), are well represented in the QFT formulation by introducing creation and annihilation operators a, a^\dagger and states $|n\rangle$.^{29,30,47-51} In analogy to quantum mechanics, the operators satisfy the commutation relation that

$$[a, a^\dagger] = 1.$$

As usual, the “vacuum state” $|0\rangle$ and its conjugate $\langle 0|$ are defined to satisfy

$$\langle 0|a^\dagger = a|0\rangle = 0, \quad \langle 0|0\rangle = 1.$$

Other states are built up from the vacuum state, such as the n -particle state $|n\rangle$,

$$|n\rangle = a^{\dagger n}|0\rangle.$$

It is easy to check with the help of the commutation relation

$$a|n\rangle = n|n-1\rangle, \quad a^\dagger|n\rangle = |n+1\rangle, \quad a^\dagger a|n\rangle = n|n\rangle.$$

Hence, $a^\dagger a$ is the “particle number operator.” Notice that the states $|n\rangle$ are not normalized in the usual sense since

$$\langle n|n\rangle = \langle 0|a^n|n\rangle = n!,$$

but they are orthogonal,

$$\langle m|n\rangle = \langle 0|a^m|n\rangle = 0 \quad \text{for } m \neq n.$$

The state that corresponds to a probability distribution $P(n)$ is

$$|\Psi\rangle = \sum_n P(n)|n\rangle.$$

The probabilities are thus encoded into the coefficients of different particle number states superimposed into the “wave function” $|\Psi\rangle$. In order to compute physical observables, the *harvesting state* $\langle \phi| = \langle 0|e^a$ is introduced. It is easy to check that

$$\langle \phi|n\rangle = 1, \quad \langle \phi|\Psi\rangle = 1, \quad \langle \phi|(a^\dagger a)^m|\Psi\rangle = \langle n^m \rangle.$$

The first equation shows the particular normalization of an n -particle state. The second equation corresponds to the probability conservation $\sum_n P(n) = 1$. The third equation may be used to calculate the m th moment of the particle number. The evolution of probabilities is governed by a wave equation for Ψ :

$$\frac{d|\Psi\rangle}{dt} = \Omega|\Psi\rangle. \quad (2)$$

The original large sets of ODEs are now compacted into just one equation. Applied to the two-step cascade (Fig. 1), Eq. (2) is characterized by the following operator Ω :

$$\begin{aligned} \Omega = & (1 - a^\dagger)(\mu b^\dagger b a - \lambda N + \lambda a^\dagger a) + g(b^\dagger - 1) \\ & + k(b - b^\dagger b), \end{aligned} \quad (3)$$

where b^\dagger, b are the creation and annihilation operators associated with R^* and a^\dagger, a with A . In this case,

$$|\Psi\rangle = \sum_{m,n} P(m, n)|m, n\rangle,$$

where

$$|m, n\rangle = a^{\dagger m} b^{\dagger n}|0\rangle.$$

Equation (3) is readily verified by substituting into Eq. (2) and comparing the coefficients of each (m, n) -particle state. In contrast to ordinary quantum mechanics, the operator Ω is non-Hermitian, so the inner products between the states are not conserved. This was the reason in introducing earlier the *harvesting state*. Nevertheless, many QFT techniques may be profitably applied, albeit with some modifications.^{29,30,48,50,57-59} We will not discuss those and instead will translate the above field theoretic formulation to the familiar differential equation language.

C. Differential operators and the generating function

In the field theoretic form, the computations are carried out by commutator manipulations that sometimes are awkward. Fortunately, it turns out that we may convert the operator equation (2) into a PDE. To accomplish that, we explore the analogy between a , a^\dagger and $d/dx, x$. Not only do they have the same commutator

$$[a, a^\dagger] = 1 \Leftrightarrow \left[\frac{d}{dx}, x \right] = 1,$$

but a more comprehensive correspondence is found:

$$aa^\dagger|0\rangle = |0\rangle \Leftrightarrow \frac{d}{dx}x = 1,$$

$$a|0\rangle = 0 \Leftrightarrow \frac{d}{dx}1 = 0,$$

$$aa^\dagger n|0\rangle = na^{\dagger n-1}|0\rangle \Leftrightarrow \frac{d}{dx}x^n = nx^{n-1},$$

$$a^m a^\dagger n|0\rangle = \frac{n!}{m!} a^{\dagger a-m}|0\rangle \Leftrightarrow \left(\frac{d}{dx}\right)^m x^n = \frac{n!}{m!} x^{n-m}$$

for $n \geq m$.

From these, we can also deduce for any smooth function f that

$$af(a^\dagger)|0\rangle \Leftrightarrow \frac{d}{dx}f(x),$$

$$a^m f(a^\dagger)|0\rangle \Leftrightarrow \left(\frac{d}{dx}\right)^m f(x).$$

The analogy $|\Psi\rangle = \sum_n P(n) a^{\dagger n}|0\rangle \Leftrightarrow \Psi(x) = \sum_n P(n) x^n$ converts a wave function to a generating function. The inner product with the harvesting state corresponds to evaluation at $x=1$. It is easy to check the following relations:

$$\langle 0|e^a|\Psi\rangle = \Psi(1),$$

$$\langle 0|e^a f(a^\dagger)|\Psi\rangle = f(1)\Psi(1),$$

$$\text{as } e^{d/dx}f(x) = f(x+1).$$

The wave equation (2) becomes then a PDE for the generating function after all the necessary conversions are done. For the two-step cascade, this expression is

$$\begin{aligned} \frac{\partial \Psi}{\partial t} = & (1-x) \left(\mu y \frac{\partial^2}{\partial x \partial y} - \lambda N + \lambda x \frac{\partial}{\partial x} \right) \Psi \\ & + g(y-1)\Psi - k(y-1) \frac{\partial \Psi}{\partial y}, \end{aligned} \quad (4)$$

where the first term describes the $A-A^*$ reaction and the rest the $R-R^*$ reaction. Generating functions were previously used to treat unary reactions in the gene regulatory network.^{10,60} The second order derivative term $\partial^2 \Psi / \partial x \partial y$ in Eq. (4) is characteristic of a binary reaction, indicative of nonlinear kinetics. This higher derivative changes the order

of the PDE and adds significant difficulty to solving Eq. (4). In the generating function formalism, the equation $\Psi(1)=1$ encodes the conservation of probability and the first two moments are given by

$$\begin{aligned} \langle n \rangle &= \left. \frac{\partial \Psi}{\partial x} \right|_{x=1}, \\ \langle n^2 \rangle &= \left. \left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial \Psi}{\partial x} \right) \right|_{x=1}, \end{aligned} \quad (5)$$

where $|_{x=1}$ means evaluation at $x=1$. Therefore, the moments are obtained when the generating function is expanded at $x=1$, while the probability distribution is obtained from the Taylor coefficients when the generating function Ψ is expanded at $x=0$.

D. The variational solution

In the QFT formulation of the stochastic processes, a variational principle may be derived which is equivalent to the evolution equation (2). This principle indicates that the physical solution of Eq. (2) is given by the stationary points of the following functional:^{29,30,53}

$$H[\Phi_L, \Phi_R] = \int_0^\infty dt \langle \Phi_L | \partial_t - \Omega | \Phi_R \rangle, \quad (6)$$

where Φ_L and Φ_R are arbitrary quantum states under quite general constraints consistent with the positivity of probabilities and the fixed boundary conditions. In practice, we take a finite-dimensional subset of the infinite-dimensional function space and apply the variational principle in this subspace to get closed equations that may be subsequently solved by simple numerical calculation. If the essential qualitative properties of the system are known, good approximations of the original problem can be achieved through an informed choice of time-dependent basis functions that define the relevant subset in the function space.

Because Ω is not Hermitian, the right and left eigenvectors are different. To characterize the system, we, therefore, need two sets of vectors Φ_L and Φ_R . The stationary variation condition for Φ_L restores the original equation (2) and that for Φ_R defines an equation satisfied by Φ_L . If we view the operator $\partial_t - \Omega$ as a large matrix parametrized by t , the Φ_L and Φ_R generated by the stationary variation condition correspond to its singular vectors^{61,62} and the extremum values of Eq. (6) are the singular values. Physically, from the Schrödinger picture point of view, Φ_R is the evolving quantum state and Φ_L represents the measurable quantities in which we are interested. Equation (6) serves to find the most significant state and physical observables. Alexander and Eyink originally applied this variational principle to Fokker-Planck equations.⁵⁴ Subsequently, Sasai and Wolynes used this variational approach in the field theoretic form and obtained moments in a toggle-switch gene regulatory problem.²⁹ In this paper, we show how the variational principle may be applied, instead, to the generating functions. We introduce novel basis functions to obtain the time-dependent probability distributions in signal transduction cascades. Another novelty of the present formulation is our

avoidance of cumbersome commutator computations in the derivation of the evolution equations, making the physical significance of the variational method more transparent.

There are many ways to choose the time-dependent basis functions. We follow the approach of Sasai and Wolynes:²⁹

$$|\Phi_R\rangle = \Phi_R(a^\dagger, \{f_i(t)\}_{i=1}^n)|0\rangle, \quad (7)$$

$$\langle\Phi_L| = \langle 0|\exp(a)\left(1 + \sum_{i=1}^m c_i(t)a^i\right), \quad (8)$$

where n is the number of unknown functions in $|\Phi_R\rangle$ and m is the number of parameters in Eq. (8). The exponential factor in $\langle\Phi_L|$ acting on $|0\rangle$ gives the harvesting state. If we substitute this “Ansatz” into Eq. (6) and carry out the variations with respect to c_i , a finite set of ODEs for the evolution of $\{f_i(t)\}$ is obtained, which then determines the evolution of the probability distribution.

As mentioned, the variational method can also be recast into the generating function language using the conversion scheme discussed previously. Now Φ_L becomes a differential operator and Φ_R a guess function of variable x . For example, Eqs. (7) and (8) correspond to

$$\begin{aligned} \Phi_R &= \Phi_R(x, \{f_i(t)\}), \\ \Phi_L &= 1 + \sum_{i=1}^m c_i(t) \frac{d^i}{dx^i}. \end{aligned} \quad (9)$$

The function H simply becomes

$$H = \int_0^\infty dt \Phi_L(\partial_t - \Omega) \Phi_R|_{x=1}. \quad (10)$$

In the new picture, we have much simpler mathematical operations, e.g., the variational principle becomes simply a function extremization condition,

$$\left. \frac{\delta H}{\delta c_i(t)} \right|_{\{c_j(t)=0\}_{j \leq m}, x=1} = 0 \quad \text{for } i = 1, 2, \dots, m \quad (11)$$

or equivalently

$$\frac{d^i}{dx^i} (\partial_t - \Omega) \Phi_R|_{x=1} = 0 \quad \text{for } i = 1, 2, \dots, m. \quad (12)$$

The evolution of the generating function should always conserve the total probability. As in Eq. (4), the total probability $\Psi(1, 1)$ does not change with time. The proper choice of Φ_R should also guarantee this invariance, satisfying $(\partial_t - \Omega) \Phi_R|_{x=1} = 0$. Now, Eq. (12) tells us that the higher derivatives of the expression $(\partial_t - \Omega) \Phi_R$ evaluated at $x=1$ are also zero. Therefore, in the limit of $m \rightarrow \infty$, Eq. (12) leads to the PDE $\partial_t \Psi_R = \Omega \Psi_R$. For finite m , this PDE is approximately satisfied in the neighborhood of $x=1$.

III. NUMERICAL APPLICATIONS

In this part, we discuss the implementation of the PDE version of the variational method and apply it to several simple, yet important enzymatic cascades. Before proceeding to the individual examples, we emphasize our motivation for

selecting the time-dependent basis functions. We also briefly discuss several alternative methods also used to solve the master equation.

A. Computational details

It is reasonable to require the following constraints on the right basis function $\Phi_R(x, y)$.

- (1) The total probability should be equal to 1, i.e., $\Phi_R(1, 1) = 1$.
- (2) The probability should be positive, i.e., the coefficients of the Taylor expansion of Φ_R around $(x, y) = (0, 0)$ should be non-negative.
- (3) The time rate of the unknown functions, $\dot{f}_i(t)$, should be obtainable by solving Eq. (12) derived from the variational principle.

In the following, we introduce two sets of basis functions. One set is simple but is of limited applicability, while the other is in a more complex integral form and can be applied very generally.

We use simple left basis functions, $\Phi_L(x, y)$. As suggested in Eq. (11), they are represented by differential operators and can be easily extended to multivariate cases. For the two-step cascade, we use

$$\Phi_{L,1} = 1 + c_1 \partial_x + c_2 \partial_y, \quad (13)$$

with a simple right basis function [see Eq. (18) below], and

$$\Phi_{L,2} = \Phi_{L,1} + c_3 \partial_{yy}, \quad (14)$$

with a more complicated right basis [see Eq. (21) below]. For the three-step cascade discussed below, we use

$$\Phi_{L,3} = \Phi_{L,2} + c_4 \partial_z + c_5 \partial_{zz}. \quad (15)$$

Following a similar pattern, we can simply write the left basis function for the four-step cascade discussed below as

$$\Phi_{L,4} = \Phi_{L,3} + c_6 \partial_w + c_7 \partial_{ww}. \quad (16)$$

In all the above equations, $\partial_x, \partial_{xx}$ denote the first and second derivatives with respect to x , and so on. Other choices of the left basis function are, of course, possible. The current choice is simple and gave the best results among different basis functions which we tried. MAPLE symbolic software was used to derive the time evolution equations and, subsequently, MATLAB numerical software was used to carry out the evolution of equations of motion and for plotting the computation results.

To validate our calculations, we used the Gillespie simulation^{5,42–44} results as the reference (exact) solutions. 10^5 stochastic trajectories were sampled to derive the distributions and other statistical quantities. At the same time, the variational method was also compared with two commonly used methods—the Langevin equation³⁴ and the Ω expansion.^{34,63–65} In the Langevin equation simulation, we also used 10^5 realizations. In addition, to prevent the appearance of negative particle numbers, we applied the selection procedure commonly used.⁶⁶ It is awkward and time con-

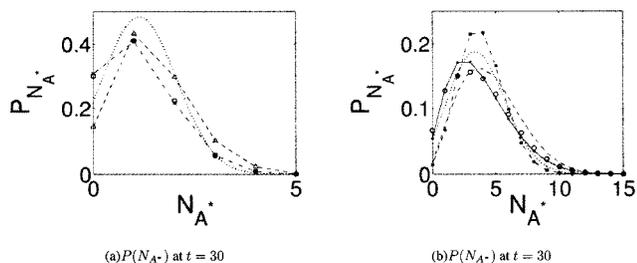


FIG. 2. Comparison of the computed distributions for N_A^* at $t=30$ for the two-step cascade: Gillespie simulation (circles), one-term basis (dash dotted line), integral form basis (solid line), Ω expansion (dotted line), and Langevin equation (dashed line). (a) $g=10$, $k=5$, $\mu=0.02$, and $\lambda=0.15$ with initial conditions $(N_R, N_{R^*}, N_A, N_{A^*})=(20, 0, 5, 0)$ and (b) $g=0.2$, $k=0.1$, $\mu=0.02$, and $\lambda=0.15$ with $(N_R, N_{R^*}, N_A, N_{A^*})=(20, 0, 20, 0)$.

suming to use the Ω -expansion method to compute the distributions. We only applied it to the simplest two-step cascade.

B. Application to a two-step amplification cascade

In a previous paper,⁴⁰ approximate analytical solutions for (4) in certain parameter range were obtained using the method of characteristics. If the initial conditions correspond to zero R^* 's and N A 's, then a generating function solution reads

$$\Psi(x, y) = \left[1 + \left(\frac{\lambda}{\lambda + \mu m(t)} + \frac{\mu m(t)}{\lambda + \mu m(t)} e^{-(\lambda + \mu m(t))t} \right) (y - 1) \right]^w \times \exp(m(t)(x - 1)), \quad (17)$$

where $m(t) = (g/k)(1 - e^{-kt})$ is the average number of R^* at time t . We make use of this specific functional form and try the following Ansatz:

$$\Phi_R = (1 + f_2(t)(y - 1))^{N_2} e^{(x-1)f_1(t)}. \quad (18)$$

This results in the following two-dimensional (2D) ODEs:

$$\begin{aligned} \dot{f}_1 &= g - kf_1, \\ \dot{f}_2 &= \lambda(1 - f_2) - \mu f_1 f_2. \end{aligned} \quad (19)$$

These equations have a particularly simple physical explanation—they correspond to the deterministic chemical kinetics equations since f_1 and $N^* f_2$ are equal to the average numbers of R^* and A , respectively. But now, we may obtain probability distributions through Eq. (18). For example, the variance of A can be easily calculated as $\sigma^2 = \langle n^2 \rangle - \langle n \rangle^2 = f_2(t) - f_2^2(t)$.

These ODEs can be solved exactly and we show in Fig. 2 the probability distribution of A^* at $t=30$ for two sets of parameter values. Also shown in the figure are results obtained from calculations using more traditional techniques. The first set of reaction rate parameters was chosen as $g=10$, $k=5$, $\mu=0.02$, and $\lambda=0.15$, with the initial conditions $(N_R, N_{R^*}, N_A, N_{A^*})=(20, 0, 5, 0)$. Since the $R-R^*$ reaction is much faster than the $A-A^*$ reaction, one expects Eq. (17) to be a good approximation.⁴⁰ Indeed, in Fig. 2(a), the variational Ansatz Eq. (18) leads to a result that overlaps significantly better with the exact Gillespie calculation, compared

with the results from the Ω expansion and the Langevin equation. The Ω -expansion result turns out to be more concentrated than the exact result, while the Langevin equation does not work well near the left boundary, shifting the average to the right.

For other parameter values, as long as the $R-R^*$ reaction is fast, the Ansatz Eq. (18) works fine as expected.⁴⁰ However, if the first reaction is considerably slower than the second one, this Ansatz becomes less useful, as shown in Fig. 2(b) for $g=0.2$, $k=0.1$, $\mu=0.02$, and $\lambda=0.15$, with $(N_R, N_{R^*}, N_A, N_{A^*})=(20, 0, 20, 0)$. The variational result gives a too narrow distribution. The Langevin equation is still not accurate on the left boundary, the average being shifted to the right.

In general, the Ansatz (18) tends to generate a distribution narrower than the exact one, which is also shown in Fig. 4(b). This can be explained as follows. The Ansatz (18) is a product of functions of x and y and hence only the average particle number f_1 appears in the second equation of (19). Therefore, the fluctuation generated in the $R-R^*$ reaction is absent in the treatment of the $A-A^*$ reaction. Physically, if the first reaction is fast, then the second reaction only “sees” an average number of R^* , with its fluctuation averaged out, and the Ansatz (18) produces accurate results [Fig. 2(a)]. If the first reaction is slow, however, then the fluctuations in the number of R^* strongly influence the $A-A^*$ reaction and the mere average f_1 is not capable of passing this information. The distribution computed from Ansatz (18) only accounts for the internal fluctuation of the $A-A^*$ reaction and hence has a narrower profile than the exact result. On the other hand, despite the apparent simplicity, this Ansatz allows one to estimate fluctuations in a reaction network in a semiquantitative way, with an extremely low computational cost, similar to solving the ordinary deterministic kinetics equations.

It is straightforward to generalize Ansatz (18) to longer cascades. For example, for the three-step cascades considered next, we may write the right Ansatz as

$$\Phi_R = (1 + f_3(t)(z - 1))^{N_2} (1 + f_2(t)(y - 1))^{N_1} e^{f_1(t)(x-1)}. \quad (20)$$

The resulting ODEs for $f_i(t)$'s are similar to Eq. (19) and have a physical interpretation related to the chemical kinetics equations, as discussed above.

To get more accurate results, we have to convolute the number fluctuation of R^* with the number fluctuation of A . Since the Ansatz based on simple separation of variables does not work, we need an equation in which x, y are explicitly entangled. To facilitate the computation, we use the following integral form representation:

$$\begin{aligned} \Phi_R(x, y) &= \int_{-\infty}^{\infty} ds \frac{e^{-s^2}}{\sqrt{\pi}} (1 + f_2(t) e^{-(s - f_3(t))^2} (y - 1) \\ &\quad + f_1(t)(x - 1))^N, \end{aligned} \quad (21)$$

where $f_1(t)$ is related to the $R-R^*$ reaction and $f_2(t)$ and $f_3(t)$ are related to the $A-A^*$ reaction. Note that $\Phi_R(1, 1)=1$. For $y=1$, we get the expected generating function for R^* ,

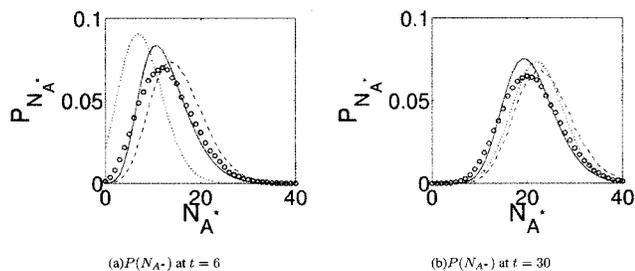


FIG. 3. Comparison of the computed distributions for N_A^* at $t=6$ and $t=30$ for the *two-step* cascade: Gillespie simulation (circles), integral form basis (solid line), Ω expansion (dotted line), and Langevin equation (dashed line). $g=2$, $k=1$, $\mu=0.02$, and $\lambda=0.15$ with initial conditions $(N_R, N_{R^*}, N_A, N_{A^*})=(20, 0, 100, 0)$.

$$\Phi_R(x, 1) = (1 + f_1(t)(x-1))^N \approx e^{Nf_1(t)(x-1)}, \quad (22)$$

the above approximation being valid when $f_1(t)$ is small, which is true in all simulations below. We could have used

$$\frac{e^{-s^2}}{\sqrt{\pi}} (1 + f_2(t)e^{-(s-f_3(t))^2}(y-1))^N \exp((s-f_1(t))^2(x-1))$$

in the integrand of (21) to achieve a larger range of f_1 . But when the number of R^* is small, *Ansatz* (21) produces better results, probably due to its more convoluted form.

Now we can control both the average and the variance of A by manipulating $f_2(t)$ and $f_3(t)$. Roughly speaking, $f_2(t)$ controls the average and $f_3(t)$ controls the variance. For the same parameter set shown in Fig. 2(b), we did the computation by using *Ansatz* Eq. (21) and displayed the result in the same figure (solid line). It matches closely with the exact result, better than all other computations.

To show the effectiveness of the *Ansatz* (21), we use it to do one more computation with $g=2$, $k=1$, $\mu=0.02$, $\lambda=0.15$, and the initial conditions $(N_R, N_{R^*}, N_A, N_{A^*})=(20, 0, 100, 0)$. In Fig. 3, the distributions of A^* are displayed at $t=6$ and $t=30$. Although, the result from *Ansatz* (21) is slightly narrower than the Gillespie computation, they match very well both at $t=6$ and at $t=30$. Actually, this is true for all times as can be seen in Fig. 4 where the time evolution of the average and the variance are depicted. In Figs. 3 and 4, the average of A^* from Langevin equation is always greater than the exact result as explained before, although the variance is computed accurately. Curiously, the average from the Ω expansion is

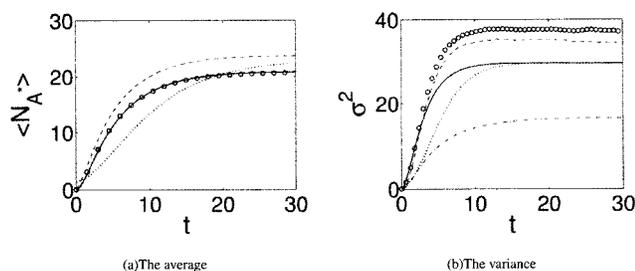


FIG. 4. Comparison of the N_A^* average and variance computed for the *two-step* cascade in the time interval $t=[0,30]$: Gillespie simulation (circles), one-term basis (dash dotted line), integral form basis (solid line), Ω expansion (dotted line), and Langevin equation (dashed line). $g=2$, $k=1$, $\mu=0.02$, and $\lambda=0.15$ with initial conditions $(N_R, N_{R^*}, N_A, N_{A^*})=(20, 0, 100, 0)$.

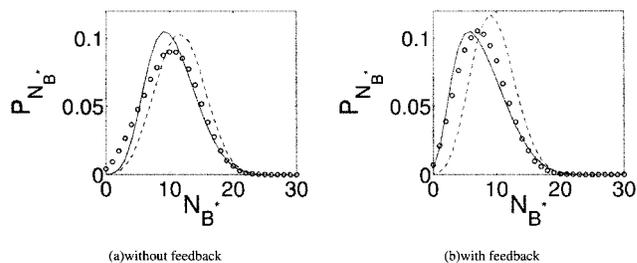


FIG. 5. Comparison of the computed distributions for N_B^* at $t=60$ for the *three-step* cascade without (a) and with (b) negative feedback: Gillespie simulation (circles), integral form basis (solid line), and Langevin equation (dashed line). $g=0.2$, $k=0.1$, $\mu=0.02$, $\lambda=0.15$, $\mu_2=0.01$, $\lambda_2=0.07$, and $\mu_3=0.01$ with initial conditions $(N_R, N_{R^*}, N_A, N_{A^*}, N_B, N_{B^*})=(20, 0, 20, 0, 30, 0)$.

smaller initially but later grows larger than the exact one. We also plotted the computation results from Eq. (18). It gives a smaller variance even though the average is quite accurate. In this case, the R^* fluctuation is important.

C. Application to a three-step amplification cascade

It is not hard to write *Ansätze* similar to Eq. (21) for longer or more complicated cascades. In this section, we demonstrate the use of the variational method for a three-step cascade with and without feedback loop. In the next section, we will write the equation for a four-step cascade.

Assume that A^* catalyzes a subsequent enzyme activation/deactivation reaction $B \rightleftharpoons B^*$ with a forward rate μ_2 and a backward decay rate λ_2 . The total number N_2 of B and B^* is a constant during the reaction. Following similar procedures as before, we found that the generating function $\Psi(x, y, z)$ satisfies

$$\begin{aligned} \frac{\partial \Psi}{\partial t} = & (1-z) \left(-\mu_2 y \frac{\partial^2}{\partial y \partial z} - \lambda_2 N_2 + (\lambda_2 z + \mu_2 N) \frac{\partial}{\partial z} \right) \Psi \\ & + (1-y) \left(\mu x \frac{\partial^2}{\partial x \partial y} - \lambda N + \lambda y \frac{\partial}{\partial y} \right) \Psi \\ & + g(x-1)\Psi - k(x-1) \frac{\partial \Psi}{\partial x}, \end{aligned} \quad (23)$$

where the first term describes the $B-B^*$ reaction. The *Ansatz* similar to Eq. (21) reads

$$\begin{aligned} \Phi_R(x, y, z) = & \int_{-\infty}^{\infty} ds \frac{e^{-s^2}}{\sqrt{\pi}} (1 + f_4(t)e^{-(s-f_3(t))^2}(z-1))^{N_2} \\ & \times (1 + f_2(t)e^{-(s-f_3(t))^2}(y-1) + f_1(t)(x-1))^N, \end{aligned} \quad (24)$$

where $f_4(t)$ and $f_5(t)$ describe the $B-B^*$ reaction. The calculation results from this *Ansatz* are shown in Figs. 5(a), 6(a), and 7(a).

Shown in Fig. 5(a) is the B^* distribution computed from different methods. *Ansatz* (25) computation matches very well with the exact solution while the Langevin profile is shifted to the right. On the left boundary, both *Ansatz* (25) and Langevin equation approach zero while the exact solution has a finite value there. Interestingly, the variance shows

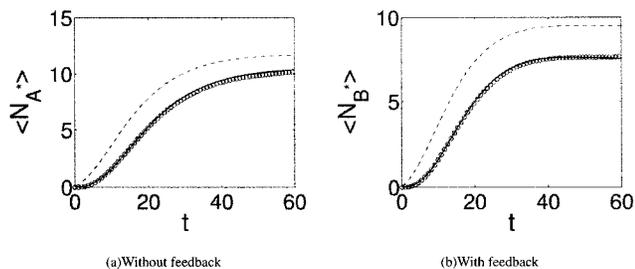


FIG. 6. Comparison of the N_B^* average computed for the *three-step* cascade without (a) and with (b) negative feedback in the time interval $t=[0,60]$: Gillespie simulation (circles), integral form basis (solid line), and Langevin equation (dashed line). $g=0.2$, $k=0.1$, $\mu=0.02$, $\lambda=0.15$, $\mu_2=0.01$, $\lambda_2=0.07$, and $\mu_3=0.01$ with initial conditions $(N_R, N_R^*, N_A, N_A^*, N_B, N_B^*)=(20, 0, 20, 0, 30, 0)$.

a maximum value during the evolution as displayed in Fig. 7(a). The computation from *Ansatz* (25) captures this nonmonotonous behavior accurately which is not obvious at all in the Langevin computation.

Next, we consider a three-step signaling cascade with a feedback loop. For example, we can imagine a reaction in which B^* turns off the R^* signaling, by catalyzing the $R^* \rightarrow R$ decay at a rate μ_3 (Fig. 8). Mathematically, this corresponds to adding an extra term

$$\mu_3(1-y)(N_2 \partial \Psi / \partial x - z \partial^2 \Psi / \partial x \partial z)$$

to the right hand side of PDE (23). We may still use the same right *Ansatz* (25) and the results are displayed in Figs. 5(b), 6(b), and 7(b). Surprisingly, despite the time scale mixing and nonlinearity, the variational computation matches even better with the exact result than without feedback [compare Figs. 5(a) and 5(b)]. The relative shift of the average computed from the Langevin equation increases. The maximum in the variance still exists but its height decreases with the variance itself. In this case, it seems that the negative feedback sharpens the signal.

D. Application to a four-step amplification cascade

Our last demonstration of the variational method is concerned with a four-step cascade. We append a further enzymatic reaction $C \rightleftharpoons C^*$ to our three-step cascade without feedback. In this reaction, the protein C is switched on with a rate μ_3 by B^* and decays at a rate λ_3 . Again, the total

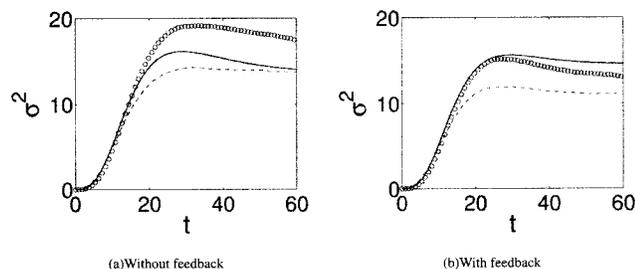


FIG. 7. Comparison of the N_B^* variance computed for the *three-step* cascade without (a) and with (b) negative feedback in the time interval $t=[0,60]$: Gillespie simulation (circles), integral form basis (solid line), and Langevin equation (dashed line). $g=0.2$, $k=0.1$, $\mu=0.02$, $\lambda=0.15$, $\mu_2=0.01$, $\lambda_2=0.07$, and $\mu_3=0.01$ with initial conditions $(N_R, N_R^*, N_A, N_A^*, N_B, N_B^*)=(20, 0, 20, 0, 30, 0)$.

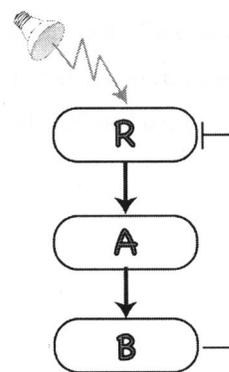


FIG. 8. An inactive receptor R , when activated by a signal, activates downstream protein A , which in turn activates protein B . In a negative feedback loop, B downregulates the R activation.

number N_3 of C and C^* is a constant during the reaction. Routinely, we add the corresponding extra term

$$(1-w) \left(-\mu_3 z \frac{\partial^2}{\partial z \partial w} - \lambda_3 N_3 + (\lambda_3 w + \mu_3 N_2) \frac{\partial}{\partial w} \right) \Psi$$

to the right hand side of Eq. (23). The right *Ansatz* is also postulated following the previous pattern,

$$\begin{aligned} \Phi_R(x, y, z, w) = & \int_{-\infty}^{\infty} ds \frac{e^{-s^2}}{\sqrt{\pi}} (1 + f_6(t) e^{-(s-f_7(t))^2} (w-1))^{N_3} \\ & \times (1 + f_4(t) e^{-(s-f_5(t))^2} (z-1))^{N_2} \\ & \times (1 + f_2(t) e^{-(s-f_3(t))^2} (y-1) + f_1(t)(x-1))^N, \end{aligned} \quad (25)$$

where $f_6(t)$ and $f_7(t)$ describe the $C-C^*$ reaction. The computations for a particular set of parameters were carried out and the results are depicted in Figs. 9 and 10.

In Fig. 9, the distributions of C^* at $t=100$ from different calculations agree with each other very well. The variational profile is slightly narrower than the exact one but the averages overlap at all times [see Fig. 10(a)]. Now, the maximum in the variance becomes more pronounced. Even the Langevin computation clearly displays this feature in Fig. 10(b) though its peak is considerably smaller than the exact one.

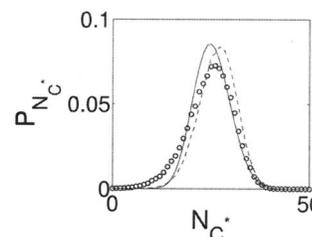


FIG. 9. Comparison of the computed distributions for N_{C^*} at $t=100$ for the *four-step* cascade: Gillespie simulation (circles), integral form basis (solid line), and Langevin equation (dashed line). $g=0.2$, $k=0.1$, $\mu=0.02$, $\lambda=0.15$, $\mu_2=0.01$, $\lambda_2=0.07$, $\mu_3=0.005$, and $\lambda_3=0.05$ with initial conditions $(N_R, N_R^*, N_A, N_A^*, N_B, N_B^*, N_C, N_{C^*})=(20, 0, 20, 0, 30, 0, 50, 0)$.

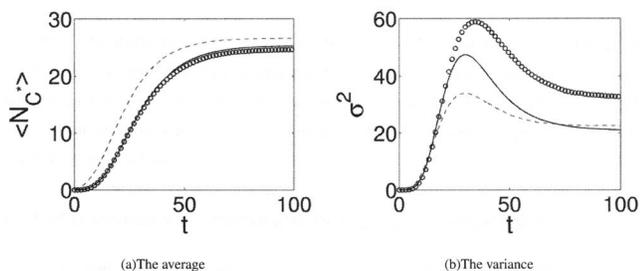


FIG. 10. Comparison of the N_C^* average and variance computed for the *four-step* cascade in the time interval $t=[0,100]$: Gillespie simulation (circles), integral form basis (solid line), and Langevin equation (dashed line). $g=0.2$, $k=0.1$, $\mu=0.02$, $\lambda=0.15$, $\mu_2=0.01$, $\lambda_2=0.07$, $\mu_3=0.005$, and $\lambda_3=0.05$ with initial conditions $(N_R, N_{R^*}, N_A, N_{A^*}, N_B, N_{B^*}, N_C, N_{C^*}) = (20, 0, 20, 0, 30, 0, 50, 0)$.

IV. THE ADVANTAGES AND DRAWBACKS OF THE VARIATIONAL PRINCIPLE

The enzymatic reaction cascades of various lengths considered in Sec. III are a common occurrence, for instance, in the MAPK family of signaling cascades.^{67–70} It is straightforward to extend the use of our variational scheme to more complex cases, to cascades or networks with complex topology. In general, the generating function is to be postulated in an integral form, as demonstrated earlier for the two-, three-, and four-step cascades, with the time-dependent parameter functions being determined by a set of ODEs derived from the variational principle. Our scheme may be used to treat both the small and large particle number systems. It is many orders of magnitude computationally more efficient in computing the distributions compared with alternative numerical simulation techniques such as the Gillespie algorithm or the Langevin equation.

However, in the current form, the *Ansatz* has a number of practical limitations, discussed next. It is difficult to represent efficiently distributions with multiple peaks, for example, or to directly compute transition rates between two deterministically stable states, a common scenario in a gene switch modeling. Another problem is that the derived set of ODEs is quite complicated; thus, symbolic algebra software is necessary to carry out the necessary manipulations. The method accuracy may also depend on the choice of the left *Ansatz*. We chose the current left *Ansatz* form from several trials for simplicity and efficiency. Occasionally when Eq. (12) is used, the time evolution of the unknown functions may result in a possible singularity, requiring work-arounds. For reaction types other than the enzymatic one discussed here, such as the binding reactions, the current basis functions may not work properly, since the total particle number of one species (including both the activated and inactive ones) is required to be constant. This may not be true for some arbitrary reaction, necessitating development of new basis functions. However, this is straightforward, and the general principles and considerations that were discussed are expected to apply to those cases as well.

In general, sufficient accuracy may be achieved with a large number of basis functions. The probability distributions are obtained when Φ_R is expanded at $x=0$ and the moments are obtained when Φ_R is expanded at $x=1$. Therefore, Ψ

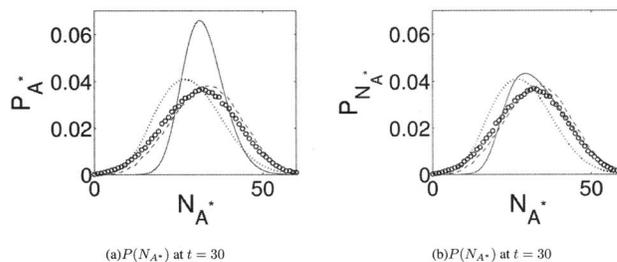


FIG. 11. Comparison of the computed distributions for N_{A^*} at $t=30$ for the *two-step* cascade: Gillespie simulation (circles), integral form basis (solid line), Ω expansion (dotted line), and Langevin equation (dashed line). $g=0.4$, $k=0.1$, $\mu=0.02$, and $\lambda=0.15$ with initial conditions $(N_R, N_{R^*}, N_A, N_{A^*}) = (20, 0, 100, 0)$. (a) $f_2(t=30)=0.7304$ and $f_3(t=30)=0.1413$ calculated by integrating the ODEs derived from Eq. (20). (b) $f_2(t=30)=0.78$ and $f_3(t=30)=0.3$ estimated by best fitting the exact solution. Shown in the picture is only the distribution profile on $[0, 60]$ with other part close to zero.

together with its derivatives is needed to be well approximated in the whole interval $x \in [0, 1]$. However, the variational equations (11) only consider the validity of Eq. (2) in the neighborhood of $x=1$. It is difficult to estimate the error bounds of Ψ and especially its derivatives near $x=0$, though we know that it generally decreases with increasing accuracy at $x=1$. The choice of basis function, therefore, is essential for the variational technique to be successful.

From the above considerations one may expect that the variational principle itself may still be improved. In Fig. 11, we show the distributions of A^* calculated with different methods with a parameter set $g=0.4$, $k=0.1$, $\mu=0.02$, $\lambda=0.15$, and initial conditions $(N_R, N_{R^*}, N_A, N_{A^*}) = (20, 0, 100, 0)$. In this case, the $R-R^*$ reaction is unusually slow compared with the $A-A^*$ reaction so that the large fluctuations in the first reaction are retained in the second one. To obtain a highly accurate solution in this parameter regime, a special convolution form was used to solve the generating function PDE in our previous work.⁴⁰ The current variational scheme, however, underestimates the A^* distribution variance. By manually adjusting the f_2 and f_3 , we may obtain a much better fit [solid line in Fig. 11(b)], demonstrating that the variational calculation does not necessarily provide an optimal solution. However, these results suggest that the present time-dependent basis sets are powerful enough to account for these extremely broad distributions. From the experience of numerical solution of ODEs and the conventional variational method in quantum mechanics, a better variational strategy may be to consider simultaneously the validity of Eq. (2) at all points on the interval $[0, 1]$. We are currently developing an improved variational approach to address some of the shortcomings discussed above.

V. SUMMARY

Cells live in a fluctuating environment in which signals and noise keep bombarding the cell receptors.^{1,71,72} Noisy signals propagate inside the cell via microscopic chemical reaction events. Cells have evolved to adapt to or even exploit the seemingly deleterious effect of fluctuations on signaling dynamics within a mesoscopic size object. Thus, it is important to develop a qualitative picture, based on math-

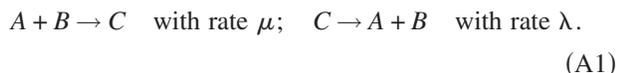
emational modeling of stochastic chemical kinetics, of how signaling networks process noisy signals. In this paper, we applied a variational principle to the solution of the master equation which describes the noisy signal propagation.

The essential difficulty associated with the master equation approach is the enormous number of ODEs involved. To compactly encode information, we use a QFT formulation in which the evolution of probability distributions is governed by one “quantum” wave equation. We have explicitly demonstrated the equivalence of the field theoretic formalism with the generating function approach, greatly facilitating the practical application of the variational technique proposed by Eyink.⁵³ We further examined the significance of the variational principle in this context. According to our previous investigation,⁴⁰ we suggest two novel classes of time-dependent basis functions: one is in simple algebraic form and the other is in an integral convolution. These basis functions are a key to the successful application of the variational method to various signaling pathways. We applied the new basis functions to describe stochastic signaling in two-step, three-step, and four-step enzymatic cascades and compared the obtained results with alternative solution techniques. The variational scheme presented here works favorably in a large parameter range. It treats effectively both the small and the large particle numbers and is orders of magnitude faster to compute compared with various Monte Carlo simulation algorithms.

However, the current scheme has also some limitations. The resulting evolution equations may be complicated and their derivation requires considerably symbolic manipulation, somewhat ameliorated by using modern computer algebra software. We also showed that the variational principle itself in this context is not the most optimal. Despite these shortcomings, the present variational approach may already be profitably applied to various signal transduction pathways, allowing one to obtain quantitative and semiquantitative solutions to stochastic signaling dynamics in a broad range of parameters. The technique may be further improved to extend its limits of applicability, which is a work in progress.

APPENDIX: APPLICATION OF THE VARIATIONAL METHOD TO A SIMPLE BINDING REACTION

To further illustrate the use of the variational method, we consider the following binding-dissociation reactions:



There are two independent constants in the above reactions,

$$N_A - N_B = \text{const}, \quad N_A + N_C = \text{const}.$$

For simplicity, we consider the case $N_A = N_B$. The probability of having n A's in the system, $P(n)$, satisfies the following master equation:

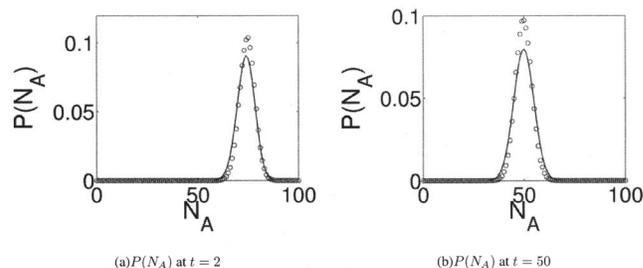


FIG. 12. Comparison of the computed distributions for N_A at $t=2$ (a) and $t=50$ (b) for the binding-dissociation reactions given by Eq. (A1): Gillespie simulation (circles), variational approach with simple basis, Eq. (A4), (solid line). $\mu=0.002$ and $\lambda=0.1$ with initial condition $(N_A, N_B, N_C) = (100, 100, 0)$.

$$\frac{dP(n)}{dt} = \mu(n+1)^2 P(n+1) - \mu n^2 P(n) + \lambda(N-n+1)P(n+1) - \lambda(N-n)P(n), \quad (\text{A2})$$

where N is the initial number of A or B. The generating function $\Psi = \sum_0^N P(n)x^n$ satisfies

$$\frac{\partial \Psi}{\partial t} = (1-x) \left(\mu \frac{\partial}{\partial x} x \frac{\partial}{\partial x} - \lambda N + \lambda x \frac{\partial}{\partial x} \right) \Psi. \quad (\text{A3})$$

Equation (A3) can be solved exactly through eigenfunction expansion in terms of the Gegenbauer polynomials.⁷³ The obtained expression is cumbersome and nonintuitive. Instead, our variational method gives a simple approximate solution.

As discussed in the main text, the formal procedure for applying the variational technique is clear, but hinges on the good choice of the basis functions. Here, if the number of B were not to change much, the reaction (A1) would be essentially linear and the characteristics method would give a nearly exact binomial solution. However, in the current case, the number of B is highly correlated with the number of A, since $N_A = N_B$, such that the binomial solution is not exact. Nevertheless, we can still use the following binomial expression as the right basis function in the variational treatment:

$$\Psi_R = (1 - f(t)(x-1))^N. \quad (\text{A4})$$

This trial function guarantees that the total probability is conserved during the evolution and the probability is zero outside the physically sensible region $[0, N]$. The ensuing variational scheme leads to an equation of motion for $f(t)$,

$$\frac{df}{dt} = \mu(f^2 - f - Nf^2) + \lambda(1 - f), \quad (\text{A5})$$

which is quite similar to the deterministic equation and can be solved exactly. In Fig. 12, we show the probability distribution of N_A at an early time $t=2$ and at a later time $t=50$, when the system is equilibrated, with the parameter values $(\mu, \lambda) = (0.002, 0.1)$. The variational approximation from Eqs. (A4) and (A5) agrees closely with the exact solution obtained through the Gillespie simulations. The exact profile is insignificantly narrower than the approximate one, which is presumably due to the correlation between the numbers of A and B.

If the first two terms in the parentheses that multiplies μ in Eq. (A5) were deleted, the resulting equation would correspond to the deterministic equation. For the current parameter values, they are small compared to the third term in the parentheses but their effect can already be seen from the small difference in the averages. The variational average $\bar{N}_v=49.83$ is very close to the Gillespie average $\bar{N}_g=49.88$ and both have a perceivable difference from the deterministic average $\bar{N}_d=50$. For smaller N , this difference would be more pronounced.

To a large extent, the success of the variational method depends on the proper choice of basis functions. Unlike the usual linear expansion, the unknown functions enter the basis in a nonlinear way; thus, the projection exhibits different properties from the commonly used linear projection. The nonlinear set of basis functions represents a curved hypersurface in the functional space and is essential in compact description of the stochastic dynamics in terms of a few reaction coordinates. On the other hand, we have no *a priori* control of the approximation error and no systematic way of choosing adequate basis functions. These are important issues for future research.

- ¹B. D. Gomperts, I. M. Kramer, and P. E. R. Tatham, *Signal Transduction* (Academic, San Diego, 2002).
- ²D. Bray, *Nature* (London) **376**, 307 (1995).
- ³B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. (Garland Science, New York, 2002).
- ⁴J. Paulsson, O. G. Berg, and M. Ehrenberg, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7148 (2000).
- ⁵T. C. Meng, S. Somani, and P. Dhar, *In Silico Biology* **4**, 0024 (2004).
- ⁶T. Shibata and K. Fujimoto, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 331 (2005).
- ⁷N. Barkai and S. Leibler, *Nature* (London) **403**, 267 (1999).
- ⁸K. Wiesenfeld and F. Moss, *Nature* (London) **373**, 33 (1995).
- ⁹M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Nature* (London) **297**, 1183 (2002).
- ¹⁰P. S. Swain, M. B. Elowitz, and E. D. Siggia, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12795 (2002).
- ¹¹E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, *Nat. Genet.* **31**, 69 (2002).
- ¹²W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins, *Nature* (London) **422**, 633 (2003).
- ¹³L. S. Weinberger, J. C. Burnett, J. E. Toettcher, A. P. Arkin, and D. V. Schaffer, *Cell* **122**, 169 (2005).
- ¹⁴M. Thattai and A. van Oudenaarden, *Genetics* **167**, 523 (2004).
- ¹⁵C. V. Rao, D. M. Wolf, and A. P. Arkin, *Nature* (London) **420**, 231 (2002).
- ¹⁶D. Hansel and G. Mato, *Phys. Rev. Lett.* **86**, 4175 (2001).
- ¹⁷N. I. Markevich, J. B. Hoek, and B. N. Kholodenko, *J. Cell Biol.* **164**, 353 (2004).
- ¹⁸K. C. Huang, Y. Meir, and N. S. Wingreen, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12724 (2003).
- ¹⁹E. Cohen, D. A. Kessler, and H. Levine, *Phys. Rev. Lett.* **94**, 158302 (2005).
- ²⁰J. P. Keener, *Bull. Math. Biol.* **63**, 625 (2001).
- ²¹C. Lemerle, B. D. Ventura, and L. Serrano, *FEBS Lett.* **579**, 1789 (2005).
- ²²R. V. Kulkarni, K. C. Huang, M. Kloster, and N. S. Wingreen, *Phys. Rev. Lett.* **93**, 228103 (2004).
- ²³R. Heinrich, B. G. Neel, and T. A. Rapoport, *Mol. Cell* **9**, 957 (2002).
- ²⁴M. Chaves, E. D. Sontag, and R. J. Dinerstein, *J. Phys. Chem. B* **108**, 15311 (2004).
- ²⁵J. P. Keener, *J. Theor. Biol.* **234**, 263 (2005).
- ²⁶M. Thattai and A. van Oudenaarden, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8614 (2001).
- ²⁷A. M. Kierzek, J. Zaim, and P. Zielenkiewicz, *J. Biol. Chem.* **276**, 8165 (2001).
- ²⁸T. B. Kepler and T. C. Elston, *Biophys. J.* **81**, 3116 (2001).
- ²⁹M. Sasai and P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 2374 (2003).
- ³⁰A. M. Walczak, M. Sasai, and P. G. Wolynes, *Biophys. J.* **88**, 828 (2005).
- ³¹J. R. Pirone and T. C. Elston, *J. Theor. Biol.* **226**, 111 (2004).
- ³²M. Thattai and A. van Oudenaarden, *Biophys. J.* **82**, 2943 (2002).
- ³³P. S. Swain, *J. Mol. Biol.* **344**, 965 (2004).
- ³⁴N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland Personal Library, Amsterdam, 1992).
- ³⁵R. F. Fox, *Phys. Rev. A* **33**, 467 (1986).
- ³⁶D. Holcman and Z. Schuss, *J. Chem. Phys.* **122**, 114710 (2005).
- ³⁷H. Kuthan, *Prog. Biophys. Mol. Biol.* **75**, 1 (2001).
- ³⁸S. Y. Shvartsman, C. B. Muratov, and D. A. Lauffenburger, *Development* **129**, 2577 (2002).
- ³⁹Y. Kuramoto, *Prog. Theor. Phys.* **52**, 711 (1974).
- ⁴⁰Y. Lan and G. A. Papoian, *J. Chem. Phys.* (to be published).
- ⁴¹N. Yakoby, C. A. Bristow, I. Gouzman, M. P. Rossi, Y. Gogotsi, T. Schüpbach, and S.Y. Shvartsman, *IEE Proceedings—Systems Biology* **152**, 276 (2005).
- ⁴²D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- ⁴³D. T. Gillespie, *J. Chem. Phys.* **115**, 1716 (2001).
- ⁴⁴J. S. van Zon and P. R. ten Wolde, *Phys. Rev. Lett.* **94**, 128103 (2005).
- ⁴⁵W. E. D. Liu, and E. Vanden-Eijnden, *J. Chem. Phys.* **123**, 194107 (2005).
- ⁴⁶C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, New York, 2002).
- ⁴⁷M. Doi, *J. Phys. A* **9**, 1479 (1976).
- ⁴⁸L. Onsager and S. Machlup, *Phys. Rev.* **91**, 1505 (1953).
- ⁴⁹D. C. Mattis and M. L. Glasser, *Rev. Mod. Phys.* **70**, 979 (1998).
- ⁵⁰J. Z. Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon, Oxford, 2002).
- ⁵¹M. Doi, *J. Phys. A* **9**, 1465 (1976).
- ⁵²Y. B. Zel'dovich and A. A. Ovchinnikov, *Sov. Phys. JETP* **47**, 829 (1977).
- ⁵³G. L. Eyink, *Phys. Rev. E* **54**, 3419 (1996).
- ⁵⁴F. J. Alexander and G. L. Eyink, *Phys. Rev. Lett.* **78**, 1 (1997).
- ⁵⁵J. E. N. Pugh and T. D. Lamb, *Biochim. Biophys. Acta* **1141**, 111 (1993).
- ⁵⁶B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, and G. Müller, *Nat. Biotechnol.* **20**, 370 (2002).
- ⁵⁷J. Wang and P. Wolynes, *Chem. Phys.* **180**, 141 (1994).
- ⁵⁸N. M. Shnerb, E. Bettelheim, Y. Louzoun, O. Agam, and S. Solomon, *Phys. Rev. E* **63**, 021103 (2001).
- ⁵⁹S. Krishnamurthy, E. Smith, D. Krakauer, and W. Fontana, e-print q-bio.MN/0312020.
- ⁶⁰O. G. Berg, J. Paulsson, and M. Ehrenberg, *Biophys. J.* **79**, 2944 (2000).
- ⁶¹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1992).
- ⁶²G. Golub and C. van Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, MD, 1996).
- ⁶³J. Elf and M. Ehrenberg, *Genome Res.* **13**, 2475 (2003).
- ⁶⁴F. Hayot and C. Jayaprakash, *Phys. Biol.* **1**, 205 (2004).
- ⁶⁵Y. Tao, Y. Jia, and T. G. Dewey, *J. Chem. Phys.* **122**, 124108 (2005).
- ⁶⁶R. F. Fox and Y. Lu, *Phys. Rev. E* **49**, 3421 (1994).
- ⁶⁷J. J. E. Ferrell and R. R. Bhatt, *J. Biol. Chem.* **272**, 19008 (1997).
- ⁶⁸C.-Y. F. Huang and J. J. E. Ferrell, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10078 (1996).
- ⁶⁹W. R. Burack and T. W. Sturgill, *Biochemistry* **36**, 5929 (1997).
- ⁷⁰R. Seger and E. G. Krebs, *FASEB J.* **99**, 726 (1995).
- ⁷¹N. Barkai and S. Leibler, *Nature* (London) **393**, 18 (1998).
- ⁷²V. N. Smelyankiy, D. G. Luchinsky, A. Stefanovska, and P. V. E. McClintock, *Phys. Rev. Lett.* **94**, 098101 (2005).
- ⁷³L. E. Reichl, *A Modern Course in Statistical Physics* (Wiley, New York, 1998).