

# Optimal transition paths of stochastic chemical kinetic systems

Di Liu<sup>a)</sup>*Department of Mathematics, Michigan State University, Michigan 48824*

(Received 19 January 2006; accepted 8 March 2006; published online 27 April 2006)

We present a new framework for finding the optimal transition paths of metastable stochastic chemical kinetic systems with large system sizes. The optimal transition paths are identified, in terms of reaction advancement coordinates, to be the most probable paths according to large deviation theory for the limiting dynamics governed by stochastic differential equations. Dynamical equations for the optimal transition paths are obtained using the variational principle. A multiscale minimum action method is proposed as a numerical scheme to solve the optimal transition paths. Applications to the toggle switch model are presented. © 2006 American Institute of Physics. [DOI: 10.1063/1.2191487]

## I. INTRODUCTION

This paper addresses the important issues on transition paths and transition rates of complex stochastic chemical kinetic systems exhibiting metastability. Metastable biochemical systems influenced by stochastic effects are common and abundant.<sup>1-3</sup> In a deterministic model, the system possesses different stable states and the dynamical trajectory converges to one of the steady states depending upon the initial condition. Incorporating stochastic effects into chemical kinetic systems induces random convergence to deterministically stable states and dynamic switching between different metastable states. Assuming ergodicity, the switching between different metastable states is guaranteed on the infinite time horizon. The time scales between the switchings are usually much longer than the time scales for the relaxation to the deterministic stationary states. In this case, transitions between different metastable states are called rare events.<sup>4</sup> Classical examples include the lytic and lysogenic developments of bacteriophage lambda virus infection in *E. coli*<sup>5</sup> and the lactose utilization of *E. coli*.<sup>6</sup>

Some recent efforts have been made on the transition paths and transition rates of metastable chemical kinetic systems<sup>7-10</sup> using large deviation theory,<sup>11</sup> which gives asymptotic probabilities for rare events. The main idea is that when the system size is very large, the limiting dynamics of the stochastic system can be described by a stochastic differential equation.<sup>12,13</sup> Therefore, the Freidlin-Wentzell theory for diffusion processes<sup>14</sup> can be applied to find the most probable transition paths. Even though this has been the common theme in much of the recent works on this subject, the key issues of the identification of the optimal transition paths under the most general circumstances and corresponding efficient numerical schemes have not been fully understood. As a result, the validity and applicability of this strategy have not been satisfactorily demonstrated.

This paper has two purposes. The first is to identify the optimal transition paths as minimizers of the Freidlin-Wentzell action functional after reformulating the system in

terms of reaction advancement coordinates. The dynamical equation satisfied by the optimal transition paths is derived using the variational principle. Compared with previous works,<sup>8,9</sup> the approach proposed here requires less assumptions. It does not need the noise driving the systems to be nondegenerate, or an energy landscape describing the system. The second purpose of this paper is to suggest a multiscale minimum action method for numerical solutions of the optimal transition paths. The method minimizes the Freidlin-Wentzell functional augmented by a boundary penalty along the steepest descent direction in the space of all paths. A multiscale structure is introduced to handle the inner and boundary terms of the transition paths in a hierarchical fashion, thereby achieves efficiency while keeping the simplicity of implementation.

As mentioned above, some recent efforts have been made on the optimal transition paths and transition rates of metastable chemical kinetic systems. In Refs. 7 and 10, a WKB method is introduced to solve the master equation of stochastic chemical kinetic systems assuming the large system size. The optimal transition paths were defined on heuristic arguments. In Refs. 8 and 9, large deviation theory is adopted to study the limiting stochastic differential equations, with applications to the bacteriophage lambda infection of *E. coli*. But the results in Refs. 8 and 9 still have their limitations. The method proposed in Ref. 8 needs the noise driving the system to be nondegenerate, which is only true when the number of reactions is larger than the number of the reacting species. The approach in Ref. 9 requires an energy landscape describing the dynamics, which is not always available for chemical kinetic systems.

In the following, after introducing some backgrounds for stochastic chemical kinetic systems, we will discuss the identification of the optimal transition paths and provide its dynamical equations. Afterwards, we will introduce the minimum action method and discuss its multiscale implementation. Then we will illustrate the method through the toggle switch model proposed in Ref. 15.

<sup>a)</sup>Electronic mail: diliu@math.msu.edu

## II. OPTIMAL TRANSITION PATHS FOR STOCHASTIC CHEMICAL KINETIC SYSTEMS

The stochastic chemical kinetic system<sup>16–18</sup> is the most successful and promising model for mesoscale systems such as genetic regulatory networks, in which reacting species are usually in low concentrations, therefore molecular fluctuations must be incorporated for an accurate account of dynamical features of the reacting system.<sup>19</sup> It describes the time evolution of an isothermal, spatially homogeneous mixture of chemically reacting molecules contained in a fixed volume  $V$ . We take  $N_S$  species of molecules  $S_{i=1,\dots,N_S}$  involved with  $M_R$  reactions  $R_{j=1,\dots,M_R}$ , with  $x_i$  being the number of molecules of species  $S_i$ . The state of the system is given by

$$x = (x_1, \dots, x_{N_S}) \in \mathbb{N}^{N_S}. \quad (1)$$

Each reaction  $R_j$  is characterized by a rate function  $a_j(x)$  and a state change vector  $\nu_j$ . We write

$$R_j = (a_j, \nu_j). \quad (2)$$

Given state  $x$ , the occurrences of the reactions on an infinitesimal time interval  $dt$  are independent of each other and the probability of reaction  $R_j$  during this time interval is given by  $a_j(x)dt$ . The state of the system after reaction  $R_j$  is  $x + \nu_j$ . The time evolution of the probability distribution of the system  $P(x, t)$  is described by the forward Kolmogorov master equation:

$$\frac{\partial P(x, t)}{\partial t} = \sum_j (a_j(x - \nu_j)P(x - \nu_j, t) - a_j(x)P(x, t)). \quad (3)$$

It is well known that when the concentrations of all reacting species in a chemical kinetic system increase to infinity (so does the overall size  $\Omega$  of the system), the forward master equation (3) describing chemical kinetic systems in discrete variables converges to a Fokker-Planck equation describing a diffusion process in a continuous state space.<sup>12</sup> Observing that the rate functions  $\{a_j(x)\}$ 's grow linearly with the system size  $\Omega$ , we first rescale the state space and the reaction rates such that

$$y = \frac{x}{\Omega}, \quad b_j(y) = \frac{a_j(x)}{\Omega}. \quad (4)$$

It is proved in Ref.13 that realizations of the discrete kinetic system (3) converge to the following diffusion process on finite time intervals with an error of order  $O(\ln \Omega / \Omega)$ :

$$\dot{y}_t = \sum_{j=1} \nu_j b_j(y_t) + \frac{1}{\sqrt{\Omega}} \sum_{j=1} \nu_j b_j^{1/2}(y_t) \dot{w}_t^j, \quad (5)$$

where  $\{\dot{w}_t^j\}$ 's are independent Brownian motions.<sup>20</sup> The same equation as above was derived intuitively with a more physical interpretation in Ref. 21 and was verified in Ref. 22 numerically through a simple dimerizing reaction. This convergence can be interpreted as the thermodynamic limit of the molecular systems when the system size  $\Omega$  approaches infinity in such a way that the relative species concentration  $x/\Omega$  remains constant. Equation (5) also justifies the well known rule of thumb in chemical kinetics that relative fluctuations

in chemical systems typically scale as the inverse square root of the system size.

Driven by small noises of magnitude  $1/\sqrt{\Omega}$  in the limiting dynamics (5), the process spends most of the time near the stable states of the deterministic system with rare switches between different metastable states. This means that the metastability of chemical kinetic systems with large system sizes fits into the framework of large deviation theory, which studies the ‘‘small perturbation, large deviation’’ phenomena, giving asymptotic estimates for the small probabilities of rare events.<sup>11</sup> Assuming the noise driving the system to be additive and nondegenerate, the large deviation theory for the sample paths of diffusion processes<sup>14</sup> has been recently combined with optimization techniques<sup>23</sup> to find the most probable transition paths, with applications to the study of nucleation events in the Ginzberg-Landau model<sup>24</sup> driven by thermal fluctuations. We will adopt a similar strategy here to find the optimal transition paths for the large system size limit of chemical kinetics (5). Modifications are needed to deal with the extra difficulty that the noise in (5) is multiplicative and not necessarily to be nondegenerate.

First we need some simplifications of dynamics (5) to overcome the degeneracy of the noise. Notice that the variable  $y$  in (5) only updates along the directions of the state change vectors  $\{\nu_j\}$ 's. We can define a new auxiliary variable  $z$  such that

$$y_t = y_0 + \sum_j z_t^j \nu_j. \quad (6)$$

The auxiliary variable  $z$  introduced here is usually called the reaction advancement coordinate.<sup>18</sup> Notice that  $z \in \mathbb{R}^{M_R}$ . Equations for  $z$  have the following form:

$$\dot{z}_t^j = c_j(z_t) + \frac{1}{\sqrt{\Omega}} c_j^{1/2}(z_t) \dot{w}_t^j, \quad j = 1, \dots, M_R, \quad (7)$$

where

$$c_j(z) = b_j(y). \quad (8)$$

We then define the following Freidlin-Wentzell action functional

$$I_T[\varphi] = \frac{1}{2} \int_0^T \sum_j c_j^{-1}(\varphi_t) (\dot{\varphi}_t^j - c_j(\varphi_t))^2 dt, \quad (9)$$

for each  $\varphi_t \in \{\varphi_0=0, \varphi_t \in C([0, T]: \mathbb{R}^{M_R})\}$ . We set  $I_T[\varphi]=\infty$  when the above integral is divergent. Notice that when  $c_j=0$ , the reaction  $R_j$  is essentially frozen. So we also set  $c_j^{-1}=0$  when  $c_j=0$ , which means the stationary state does not contribute in the action functional. The main result of the Freidlin-Wentzell theory<sup>11,14</sup> is that for sufficiently large  $\Omega$ , a probability can be assigned for each path  $\varphi$  such that

$$P\{\|\varphi - \varphi^*\| < \delta\} \approx \exp\{-\Omega I_T[\varphi^*]\}, \quad (10)$$

where  $\delta$  is a sufficiently small number. The transformations (6) and (7) change the dynamics into a stochastic differential equation with a diagonal diffusion matrix that is more likely to be nondegenerate, which makes the large deviation theory for diffusion processes more applicable. Without this transformation, to define the Freidlin-Wentzell action functional

(9) for the original diffusion process (5) as in Ref. 8, we would have to assume the diffusion matrix  $a = \sigma\sigma'$  with  $\sigma = \{b_j^{1/2} \nu_j\}$  to be nongenerate, which is only true when the number of reactions is larger than the number of reacting species.

The probabilistic estimate (10) suggests that we can find the most probable path by minimizing the rate functional  $I_T[\varphi]$  over all possible transition pathways. Suppose that we have two metastable states  $A$  and  $B$ . The optimal transition path between  $A$  and  $B$  should minimize the Freidlin-Wentzell functional such that

$$I_T[\psi] = \min_{\varphi} I_T[\varphi]. \quad (11)$$

The variational principle implies that  $\psi$  should satisfy the following equation:

$$\begin{aligned} \ddot{\psi}_i^k &= \dot{\psi}_i^k \sum_j \dot{\psi}_i^j \nabla_j \ln c_k(\psi_i) + \frac{1}{2} c_k(\psi_i) \sum_j (\dot{\psi}_i^j)^2 \nabla_k c_j^{-1}(\psi_i) \\ &+ \frac{1}{2} c_k(\psi_i) \sum_j \nabla_k c_j(\psi_i), \quad k = 1, \dots, M_R, \end{aligned} \quad (12)$$

with the boundary condition

$$\nu \cdot \varphi_0 = 0, \quad \nu \cdot \varphi_T = B - A, \quad (13)$$

where  $\nu = \{\nu^1, \dots, \nu^{M_R}\}$ . The derivation of the above equation is provided in the Appendix. Following Ref. 24, we call the optimal transition path satisfying Eqs. (12) and (13) the “minimum action path.” The undetermined boundary condition (13) is from the fact that at metastable states, the reactions are in equilibrium but not shut off. The reacting species are still being synthesized and degraded, though the processes are in balance with each other. It can also be seen from the definition that the equilibrium states do not contribute to the action functional (9) since the processes in balance satisfy the deterministic equation without random perturbation,

$$\dot{z}_i^j = c_j(z). \quad (14)$$

Notice that (12) and (13) may possess multiple solutions. Provided that  $\psi_T$  is a global minimizer of  $I_T$  satisfying (12) and (13), the mean exit time for the process to switch from metastate  $A$  to metastate  $B$  can be given<sup>11,14</sup> as the following:

$$\tau \approx \exp\{\Omega \inf_T I_T[\psi_T]\}. \quad (15)$$

### III. THE MULTISCALE MINIMUM ACTION METHOD

Now we want to introduce a numerical scheme solving Eqs. (12) and (13) satisfied by the optimal transition paths. The standard scheme using optimization techniques to solve Eq. (12) with fixed boundary values is the so-called “minimum action method.”<sup>24</sup> The method proceeds by evolving curves in the space of all transition paths with a dynamics that relaxes to the most probable transition path. Here we will propose a modified version of the minimum action method which will include penalty terms in the curve dynamics to handle the undetermined boundary condition (13).

A multiscale structure of the scheme to deal with the stiffness generated by the boundary penalty terms will also be introduced.

First, we want to give a brief review of the minimum action method. Suppose we want to find a minimizer  $\psi$  for the action functional  $I_T[\varphi]$  with fixed boundary conditions at  $A_0$  and  $A_1$ , i.e.,

$$\begin{cases} I_T[\psi] = \min_{\varphi} I_T[\varphi], \\ \text{subject to } \varphi_0 = A_0, \varphi_T = A_1. \end{cases} \quad (16)$$

Denote by  $\varphi(\alpha, s)$  the evolving curve with fixed boundary at  $A_0$  when  $\alpha=0$  and at  $A_1$  when  $\alpha=T$ . After an appropriate time discretization for  $\alpha$ , the minimum action method evolves the following equation:

$$\frac{\partial \varphi}{\partial s} = - \frac{\delta I_T[\varphi]}{\delta \varphi}. \quad (17)$$

Direct time discretizations for  $s$  in the above equation with Euler method will amount to the steepest descent method for the minimization of  $I_T[\varphi]$ . More sophisticated optimization techniques are available to gain extra efficiency.<sup>23,24</sup>

To deal with the undetermined boundary condition (13), we propose a modified version of minimum action method. Notice that (13) can be seen as a constraint imposed on the boundary of the transition paths. This suggests standard techniques of constraint optimization. We first introduce a new action functional with boundary penalty terms:

$$I_T[\varphi, \mu] = I_T[\varphi] + \frac{1}{2\mu} (|\nu \cdot \varphi_0|^2 + |\nu \cdot \varphi_T - (B - A)|^2), \quad (18)$$

where  $\mu > 0$  is the *penalty parameter*. By driving  $\mu$  to zero, we penalize constraint violations at the boundary with increasing severity. In practice, we choose a sequence of values  $\{\mu_k\}$  with  $\mu_k \downarrow 0$  as  $k \rightarrow \infty$  and seek approximate minimizers for  $I_T[\varphi, \mu_k]$ . As  $\mu_k \rightarrow 0$ , the solutions should converge to the minimizer of the action functional with the undetermined boundary condition (13):

$$\begin{cases} I_T[\psi] = \min_{\varphi} I_T[\varphi], \\ \text{subject to } \nu \cdot \varphi_0 = 0, \nu \cdot \varphi_T = B - A. \end{cases} \quad (19)$$

The procedure of the algorithm can be described as the following:

(S1) Initiate  $k=0$ , choose  $\mu_0 > 0$ ,  $\lambda_0 > 0$ , and starting curve  $\varphi^0$ ;

(S2) Solve with starting curve  $\varphi^k$  the following gradient dynamics for each  $k$ :

$$\frac{\partial \varphi}{\partial s} = - \frac{\delta I_T[\varphi, \mu_k]}{\delta \varphi}, \quad (20)$$

to reach an approximate solution  $\psi^k$  such that the absolute value of the right hand side of (20) is smaller than a threshold  $\lambda_k$ ; and

(S3) Reset starting curve  $\varphi_s^{k+1} = \psi^k$  and choose

$$\mu_{k+1} < \mu_k, \quad \lambda_{k+1} < \lambda_k. \quad (21)$$

Repeat (S2).

We call this scheme the “minimum action method with boundary penalty.”

Now we want to discuss efficient numerical schemes for solving (20). We first discuss the time discretization of (20). Then we will introduce a multiscale scheme to deal with the stiffness generated by the smallness of  $\mu$ . We discretize the time domain  $[0, T]$  with a mesh of size  $\Delta\alpha = 1/L$  and define the grid point  $\alpha_\ell$  to be

$$\alpha_\ell = \ell\Delta\alpha, \quad \ell = 0, 1, \dots, L. \quad (22)$$

The numerical approximation to  $\varphi(\alpha_\ell)$  is denoted by  $\Phi_\ell$ . We use the midpoint rule to discretize the time integral in the action functional  $I_T[\varphi]$ , which will give us

$$I_T(\Phi) = \sum_j \sum_{\ell=0}^{L-1} c_j^{-1} \left( \frac{\Phi_{\ell+1} + \Phi_\ell}{2} \right) \times \left( \frac{\Phi_{\ell+1}^j - \Phi_\ell^j}{\Delta\alpha} - c_j \left( \frac{\Phi_{\ell+1} + \Phi_\ell}{2} \right) \right)^2. \quad (23)$$

Written in discretized variables, dynamics (20) has the following form for the inner points of the transition path:

$$\frac{\partial\Phi_\ell}{\partial s} = -\frac{\partial I_T(\Phi)}{\partial\Phi_\ell}, \quad \ell = 1, \dots, L-1. \quad (24)$$

coupled with the equation for the boundary points

$$\begin{aligned} \frac{\partial\Phi_0}{\partial s} &= -\frac{\partial I_T(\Phi)}{\partial\Phi_0} - \frac{1}{\mu} \nu' \nu \Phi_0, \\ \frac{\partial\Phi_L}{\partial s} &= -\frac{\partial I_T(\Phi)}{\partial\Phi_L} - \frac{1}{\mu} \nu' (\nu\Phi_L - (B-A)). \end{aligned} \quad (25)$$

Equations (24) and (25) have a multiscale nature due to the stiffness generated by  $\mu$ , which represents the time scale separation of the system. Equation (24) describes the time evolution of the slow dynamics for the slow variables  $\Phi_{\ell=1, \dots, L-1}$ , while Eq. (25) describes the time evolution of the fast dynamics for the fast variables ( $\Phi_0, \Phi_L$ ). Direct simulation of (24) and (25) will entail very small time steps prescribed by  $\mu$  and most of the computing time will be spent on the simulation of the fast dynamics. Meanwhile, what is more of interest is the slow variables  $\Phi_{\ell=1, \dots, L-1}$ , which give the optimal transition paths. Here we propose a multiscale scheme consisting of two solvers organized with one nested in the other: An outer solver for the slow variables only, with the coefficients of the slow dynamics being computed in an inner solver for the fast dynamics only. At each iteration of step (S2) of the minimum action method with boundary penalty, the nested scheme does the following to solve (20):

(M1) *Inner solver.* Solve fast dynamics (25), with the slow variables  $\Phi_{\ell=1, \dots, L-1}$  fixed as parameters, until an approximate stationary solution ( $\Psi_0, \Psi_L$ ) is reached such that the absolute value of the right hand side of (25) is smaller than  $\lambda_k$ , where  $\lambda_k$  is the same threshold in (S2).

(M2) *Outer solver.* Solve slow dynamics (24) for one time step, with the right hand side of (24) evaluated using the quasistationary solution ( $\Psi_0, \Psi_L$ ) obtained from the inner solver. Repeat (M1).

As before, we could use the steepest descent method or other optimization algorithms as the inner and outer solvers. The rationale behind the above scheme is that fast dynamics

(25) is still a gradient flow driven by a potential in which the slow variables are fixed parameters. The fast variables should reach a quasiequilibrium on a time scale of  $O(1/\mu)$ , which is much faster than the  $O(1)$  time scale on which the slow variables advance. The same idea has been applied to the stochastic simulation of chemical kinetic systems with multiple time scales (see Refs. 25 and 26 and the reference therein). As shown in Refs. 25 and 26, if we choose appropriate numerical parameters, the above scheme can achieve an increasing accuracy for smaller  $\mu$  with a computational cost independent of  $\mu$ . We call the overall scheme given as (S1)–(S3) and (M1) and (M2) the “multiscale minimum action method.”

#### IV. NUMERICAL EXAMPLE: THE TOGGLE SWITCH MODEL

We consider the following example of toggle switch model,<sup>15</sup> which artificially realizes that a switch consists of two genes repressing each other's expression, placed in a high copy plasmid in *E. coli*. Once expressed, each protein can bind particular DNA sites upstream of the gene which codes for the other protein, thereby represses its expression. The deterministic equation describing the system has the following form:

$$\begin{aligned} \dot{u} &= \frac{\alpha_1}{1+v^\beta} - u, \\ \dot{v} &= \frac{\alpha_2}{1+u^\gamma} - v, \end{aligned} \quad (26)$$

where  $u$  is the concentration of repressor 1,  $v$  is the concentration of repressor 2, and  $\alpha_{i=1,2}$  are the effective rates of synthesis of repressor 1 and repressor 2, respectively. There are four reactions in the system, namely, the synthesis and degradations of the repressor  $u$  and repressor  $v$ . The chemical Langevin equation, consistent with (5), has the following form:

$$\begin{aligned} \dot{u} &= \frac{\alpha_1}{1+v^\beta} + \sqrt{\frac{\alpha_1}{\Omega(1+v^\beta)}} \dot{w}_t^1 - u - \sqrt{\frac{u}{\Omega}} \dot{w}_t^2, \\ \dot{v} &= \frac{\alpha_2}{1+u^\gamma} + \sqrt{\frac{\alpha_2}{\Omega(1+u^\gamma)}} \dot{w}_t^3 - v - \sqrt{\frac{v}{\Omega}} \dot{w}_t^4, \end{aligned} \quad (27)$$

where  $w_t^i$  ( $i=1, 2, 3, 4$ ) are independent standard Brownian motions. The transition paths of slightly different versions of the same system have been studied in Refs. 8–10. The reaction advancement coordinate  $z$  defined as in (6) can be given according to the state change vectors:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + z_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + z_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + z_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + z_4 \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \quad (28)$$

The reaction rates in terms of  $z$  are



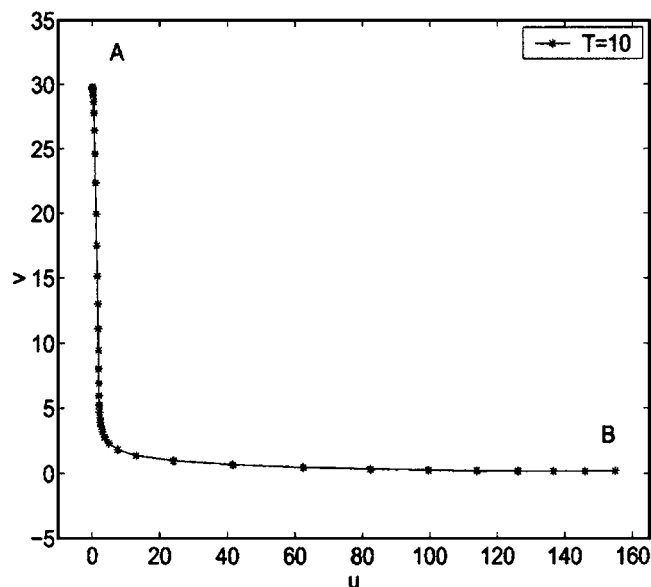


FIG. 1. Optimal transition path for the toggle switch model from metastable state A to metastable state B when  $T=10$ .

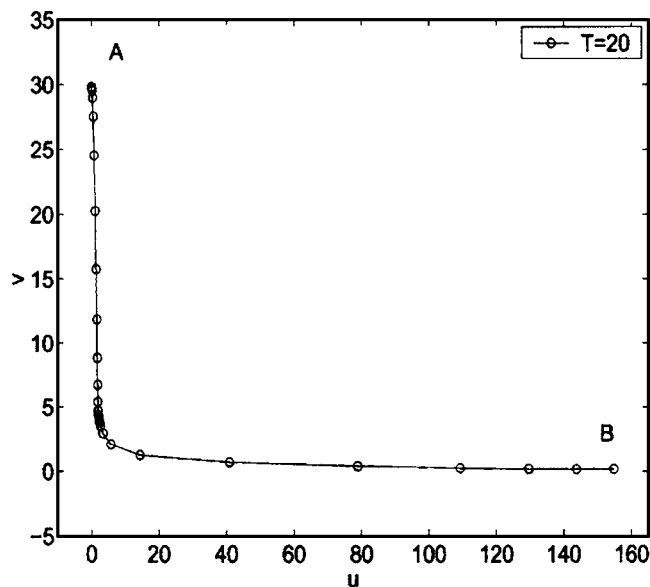


FIG. 2. Optimal transition path for the toggle switch model from metastable state A to metastable state B when  $T=20$ .

$$\begin{aligned}
 c_1(z) &= \frac{\alpha_1}{1 + (v_0 + z_3 - z_4)^\beta}, \\
 c_2(z) &= u_0 + z_1 - z_2, \\
 c_3(z) &= \frac{\alpha_2}{1 + (u_0 + z_1 - z_2)^\gamma}, \\
 c_4(z) &= u_0 + z_3 - z_4.
 \end{aligned}
 \tag{29}$$

We choose the parameters in the toggle switch model (27) to be the same as in Ref. 10 such that

$$\alpha_1 = 156, \quad \alpha_2 = 30, \quad \beta = 3, \quad \gamma = 1. \tag{30}$$

The metastable states in the system are

$$A = (0.005\ 88, 29.825), \quad B = (154.897, 0.192). \tag{31}$$

The evolving path is discretized with  $L=40$  nodes equally distributed as in (22). We choose the parameters in the multiscale minimum action method to be

$$\mu_k = \lambda_k = 2^{-k}, \quad k = 0, 1, \dots \tag{32}$$

The steepest descent method is chosen as both inner and outer solvers. In Figs. 1 and 2, we show the optimal transition paths for different time horizons  $T=10$  and  $T=20$ , respectively. The results qualitatively agree with those in Ref. 10. Figure 3 gives the value of the Freidlin-Wentzell action functional (9) for the optimal transition paths on different time horizons.

It can be seen from Figs. 1 and 2 that there is no significant change in the coordinates that the optimal paths for different time horizons visit. But for the longer time horizon of  $T=20$ , the optimal path spends more time around the critical point in the middle of the path. This makes the minimum action method very undesirable for longer time scales because more nodes are needed to capture the nucleation events near the end points of the path, which will increase the com-

putational load. On the other hand, to estimate the mean transition rates using (15) involves the minimizers of the action functional on the infinite time horizon. Approximating (15) with the minimizers for  $T=80$  will give a transition rate much smaller than the experimentally observed rate. This could be partially due to the inadequacy of the model<sup>10</sup> or due to the limitation of the minimum action method on infinite time horizon.

On the theoretical side, an important issue is the convergence and accuracy of the multiscale minimum action method in terms of the system size  $\Omega$  and parameters of the numerical discretization. Without making use of the auxiliary diffusion process (5), Shwartz and Weiss provided the action functional for the rescaled variable  $y$  defined as in (4) in continuous state space in a different form.<sup>27</sup> But the action

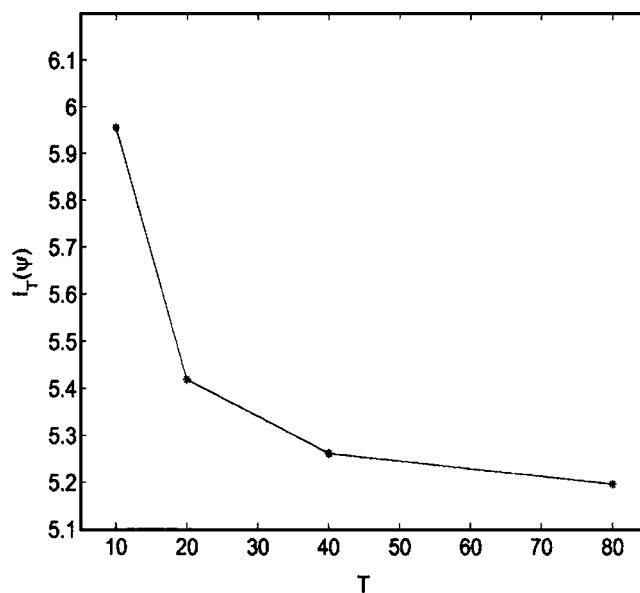


FIG. 3. Freidlin-Wentzell action functionals for the optimal paths on different time horizons.

functional proposed in Ref. 27 does not have the seemingly numerical tractability that the Freidlin-Wentzell functional possesses. A future thorough mathematical analysis for the relation between the two action functionals and the error induced by adopting the Freidlin-Wentzell functional is needed for the rigorous justification of the multiscale minimum action method.

## V. CONCLUSION

We propose a new framework for the optimal transition paths of stochastic chemical kinetic systems by minimizing the Freidlin-Wentzell functional in reaction advancement coordinates using the multiscale minimum action method. Future investigations may involve (1) identification and numerical methods for optimal transition paths on the infinite time horizon, (2) convergence analysis for the multiscale minimum action method, and (3) applications to more complex systems.

## ACKNOWLEDGMENTS

We are grateful to Weinan E and Eric Vanden-Eijnden for stimulating discussions and insightful comments. We want to thank Princeton Institute for Computational Science and Engineering (PICSciE) for providing their computing resources.

## APPENDIX: MINIMIZERS OF THE ACTION FUNCTIONAL

Denote by  $\psi$  a minimizer of the action functional  $I_T$  as defined in (9). By definition, we have

$$I_T[\psi] = \frac{1}{2} \int_0^T \sum_j ((\dot{\psi}_t^j)^2 c_j^{-1}(\psi_t) - 2\dot{\psi}_t^j + c_j(\psi_t)) dt. \quad (\text{A1})$$

Suppose  $\phi$  is a smooth function with  $\phi_0 = \phi_T = 0$ . Differentiating  $I_T[\psi + \epsilon\phi]$  with respect to  $\epsilon$  and taking the first order derivative to be zero, we have

$$\begin{aligned} \frac{1}{2} \int_0^T \sum_j \left( 2\dot{\psi}_t^j \dot{\phi}_t^j c_j^{-1}(\psi_t) - \left( \frac{\dot{\psi}_t^j}{c_j(\psi_t)} \right)^2 \sum_\ell \nabla_\ell c_j(\psi_t) \phi_t^\ell \right. \\ \left. - 2\dot{\phi}_t^j + \sum_\ell \nabla_\ell c_j(\psi_t) \phi_t^\ell \right) dt = 0. \end{aligned} \quad (\text{A2})$$

Integrating by parts, the arbitrariness of  $\phi$  gives

$$\begin{aligned} -2\dot{\psi}_t^k c_k^{-1}(\psi_t) + 2\dot{\psi}_t^k \frac{1}{c_k(\psi_t)} \sum_\ell \nabla_\ell c_k(\psi_t) \dot{\psi}_t^\ell \\ - \sum_j \left( \frac{\dot{\psi}_t^j}{c_j(\psi_t)} \right)^2 \nabla_k c_j(\psi_t) + \sum_j \nabla_k c_j(\psi_t) = 0, \end{aligned} \quad (\text{A3})$$

or equivalently

$$\begin{aligned} \ddot{\psi}_t^k &= \frac{\dot{\psi}_t^k}{c_k(\psi_t)} \sum_\ell \nabla_\ell c_k(\psi_t) \dot{\psi}_t^\ell \\ &\quad - \frac{1}{2} c_k(\psi_t) \sum_j \left( \frac{\dot{\psi}_t^j}{c_j(\psi_t)} \right)^2 \nabla_k c_j(\psi_t) \\ &\quad + \frac{1}{2} c_k(\psi_t) \sum_j \nabla_k c_j(\psi_t), \end{aligned} \quad (\text{A4})$$

which gives (12).

By the definition of the reaction advancement coordinate  $z$ , setting  $y_0 = A$  in (6), we have at equilibrium states  $A$  and  $B$

$$A = A + \nu \cdot z, \quad B = A + \nu \cdot z, \quad (\text{A5})$$

which is equivalent to the boundary condition (13):

$$\nu \cdot z = 0, \quad \nu \cdot z = B - A. \quad (\text{A6})$$

- <sup>1</sup>J. E. Ferrel and E. M. Machleder, *Science* **280**, 339 (1998).
- <sup>2</sup>A. Bescke, B. Seraphin, and L. Serrano, *EMBO J.* **20**, 2528 (2001).
- <sup>3</sup>P. Smolen, D. Baxter, and J. Byrne, *Am. J. Physiol.: Cell Physiol.* **274**, 531 (1998).
- <sup>4</sup>P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251341 (1990).
- <sup>5</sup>A. Arkin, J. Ross, and H. MacAdams, *Genetics* **149**, 1633 (1998).
- <sup>6</sup>E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden, *Nature (London)* **427**, 737 (2004).
- <sup>7</sup>M. Dykman, E. Mori, J. Ross, and P. Hunt, *J. Chem. Phys.* **100**, 5735 (1994).
- <sup>8</sup>E. Aurell and K. Sneppen, *Phys. Rev. Lett.* **88**, 048101 (2002).
- <sup>9</sup>X. Zhu, L. Yin, L. Hood, and P. Ao, *Funct. Integr. Genomics* **4**, 188 (2004).
- <sup>10</sup>D. Roma, R. O'Flanagan, A. Ruckenstein, A. Sengupta, and R. Mukhopadhyay, *Phys. Rev. E* **71**, 011901 (2005).
- <sup>11</sup>A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. (Springer, New York, 1998).
- <sup>12</sup>C. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer-Verlag, Berlin, 2003).
- <sup>13</sup>T. Kurtz, *Stochastic Proc. Appl.* **6**, 223 (1978).
- <sup>14</sup>M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, 2nd ed. (Springer, New York, 1998).
- <sup>15</sup>T. S. Gardner, C. R. Cantor, and J. J. Collins, *Nature (London)* **403**, 339 (2000).
- <sup>16</sup>D. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).
- <sup>17</sup>D. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- <sup>18</sup>N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
- <sup>19</sup>H. E. Samad, M. Khammash, L. Petzold, and D. Gillespie, *Int. J. Robust Nonlinear Control* **15**, 691 (2005).
- <sup>20</sup>B. Oksendal, *Stochastic Differential Equations*, 5th ed. (Springer-Verlag, Berlin, 1991).
- <sup>21</sup>D. Gillespie, *J. Chem. Phys.* **113**, 297 (2000).
- <sup>22</sup>D. Gillespie, *J. Phys. Chem.* **106**, 5063 (2002).
- <sup>23</sup>R. Fletcher, *Practical Methods of Optimization*, 2nd ed. (Wiley, New York, 2000).
- <sup>24</sup>W. E, W. Ren, and E. Vanden-Eijnden, *Commun. Pure Appl. Math.* **57**, 637 (2004).
- <sup>25</sup>W. E, D. Liu, and E. Vanden-Eijnden, *J. Chem. Phys.* **123**, 194107 (2005).
- <sup>26</sup>W. E, D. Liu, and E. Vanden-Eijnden, *J. Comput. Phys.* (to be published).
- <sup>27</sup>A. Shwartz and A. Weiss, *Large Deviations for Performance Analysis: Queues, Communications, and Computing* (Chapman and Hall, London, 1995).