Journal of Computational Physics 227 (2008) 8672–8684

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

A numerical scheme for optimal transition paths of stochastic chemical kinetic systems

Di Liu *

Department of Mathematics, Michigan State University, A201 Wells Hall, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:
Received 17 November 2006
Received in revised form 10 June 2008
Accepted 16 June 2008
Available online 24 June 2008

Keywords:
Stochastic chemical kinetic systems
Stochastic processes
Transition paths and transition rates
Large Deviation Theory
Numerical methods
Constraint optimization

ABSTRACT

We present a new framework for finding the optimal transition paths of metastable stochastic chemical kinetic systems with large system size. The optimal transition paths are identified to be the most probable paths according to the Large Deviation Theory of stochastic processes. Dynamical equations for the optimal transition paths are derived using the variational principle. A modified Minimum Action Method (MAM) is proposed as a numerical scheme to solve the optimal transition paths. Applications to Gene Regulatory Networks such as the toggle switch model and the Lactose Operon Model in *Escherichia coli* are presented as numerical examples.

1. Introduction

This paper addresses the important issues on transition paths and transition rates of complex stochastic chemical kinetic systems exhibiting metastability. Metastable biochemical reacting systems influenced by stochastic effects are common and abundant [1,2]. 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 stationary states and dynamic switching between different metastable states. Assuming ergodicity, the transition between different metastable states is guaranteed on the infinite time horizon. Driven by small random perturbations when the system size is large, the time scales between the switchings are usually much longer than the time scales of the relaxation to the deterministic stable states. In this case, transitions between different metastable states are called rare events [3].

The stochastic chemical kinetic system is not only a successful model but also an effective simulation algorithm for chemical reacting systems that takes into account the discreteness of molecular numbers as well as the stochastic effects on reaction events [4–6]. It has found a wide range of applications in many different fields, including computational biology, chemistry, material sciences, and communication networks. A recent motivation on the analysis and computation of stochastic chemical kinetic systems arises from the stochastic modeling of Gene Regulatory Networks in living cells. The synthesis of cellular proteins is a multi-step process. The genetic information stored in DNA directs the production of proteins through a process called gene expression [7]. When a specific gene is expressed, its DNA is first transcribed into a single stranded

* Tel.: +1 517 353 8143; fax: +1 517 432 1562.
E-mail address: richardl@math.msu.edu

0021-9991/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2008.06.010

sequence of mRNA. The mRNA sequence is then translated into a sequence of amino acids as the protein is formed. A special class of protein products called Transcription Factors (TFs) regulates the timing of the transcription of genes into mRNA. A Gene Regulatory Network (GRN), describing a collection of reacting channels and species involved in gene expression, consists of a set of genes, proteins, small molecules and their mutual regulatory interactions. From the point of view of modeling, Gene Regulatory Networks, different from metabolic networks, involve fewer number of species and lower population of molecules contained in a small volume within the cell; therefore stochastic effects have a significant impact on the reaction pathways and stochastic models are particularly well suited to the study of the functionality of GRNs. In the mean time, metastability plays a very important role in the dynamics of gene regulation in the sense that different stable profiles of gene expression determine mutually exclusive phenotypic states and switchings between different stable states can be induced by molecular fluctuations [8]. Classical examples of metastable systems in GRNs include the bacteriophage lambda virus infection [9,10] and the lactose utilization network [11,12] in *E. coli*.

In [13,14], a WKB method is adopted to solve the forward master equation of stochastic chemical kinetic systems assuming the large system size. Although it does provide explicit formulas and quite accurate numerical solutions for the transition paths and transition rates, the WKB method is heuristic in nature. Some recent efforts have been also made based on the Large Deviation Theory (LDT), which gives asymptotic probabilities for rare events. The main idea is that when the system size goes to infinity, the limiting dynamics of the stochastic chemical kinetic system can be described by a stochastic differential equation driven by Brownian noises, which implies that the Large Deviation Theory for diffusion processes can be applied to find the most probable transition paths [15–17]. Even though this has been the common theme in much of the recent works on this subject, the validity of the approach is still questionable. It has been noticed that the LDT action functional takes a different form if we replace the Poisson noise in the system with a Brownian noise [18–20]. Therefore, the key issue of the identification of the optimal transition paths under the most general circumstances has not been fully understood.

Another important issue is the designing of efficient numerical schemes to solve the optimal transition paths. In [17], a Minimum Action Method (MAM) is introduced to obtain the optimal transition paths by numerically minimizing the LDT action functional of the limiting diffusion process. The scheme suggested in [17] is a generalization of the original MAM proposed in [21] for finding optimal transition paths of stochastic PDEs driven by additive space–time white noise. The method in [21] is recently generalized in [22] to handle chemical kinetic systems effectively, assuming the noise driving the system is non-degenerate. Meanwhile, it is realized in [17] that if the chemical reacting system is written in terms of the reaction advancement coordinate, the limiting stochastic differential equation has a diagonal diffusion matrix. This greatly simplifies the evaluation of the LDT action functional, especially when the original diffusion matrix is degenerate. But for general situations without either adopting the diffusion limit or assuming the non-degenerate noise, numerical methods for the optimal transition paths of stochastic chemical kinetic systems are still not well studied.

This paper has two purposes. The first is to identify the optimal transition paths as minimizers of the action functional of the Large Deviation Theory for stochastic processes. The dynamical equation satisfied by the optimal transition paths is derived using the variational principle. It is shown that making use of the special structure of chemical reacting systems to reformulate the system in terms of the reaction advancement coordinate still makes it much easier to compute the LDT action functional and its gradient. Compared with previous works [15–17], the approach proposed here does not assume the limiting diffusion process and avoids the inaccuracy that this assumption may induce. The second purpose of this paper is to suggest a modified version of the Minimum Action Method for numerical solutions of the optimal transition paths. The method minimizes the LDT action functional augmented by a boundary penalty using optimization schemes in the space of all possible transition paths. We introduce a new smoothing technique for the numerical tractability of the LDT action functional based on Taylor expansion. A multiscale numerical technique proposed in [17] is adopted here to handle the inner and boundary terms of the transition paths in a hierarchical fashion, which achieves efficiency while keeping simplicity of implementation. The scheme proposed in this paper does not require the non-degeneracy of the noise driving the system like [22], therefore has a potential for wider applicability.

In the following, after introducing some backgrounds for stochastic chemical kinetic systems and the Large Deviation Theory for chemical kinetic systems with large system size, we will identify the optimal transition paths as minimizers of the LDT action functional and provide its dynamical equations. Then we will discuss the Minimum Action Method and its modifications for chemical reacting systems. Finally, we will illustrate the method through the Toggle Switch Model [10] and the Lactose Operon Model [12] of *E. coli*.

2. Stochastic chemical kinetic systems with large system size

The stochastic chemical kinetic system, also known as Kinetic Monte Carlo Method (KMC) [4] or Stochastic Simulation Algorithm (SSA) [5,6], is the most successful and promising model for meso-scale kinetic systems in which reacting species are usually in low molecular population therefore molecular fluctuations must be incorporated for an accurate account of the dynamical features of the system. It describes the time evolution of an isothermal, spatially homogeneous mixture of chemically reacting molecules contained in a fixed volume V. Suppose we have N_S species of molecules $S_{i=1,\ldots,N_S}$ involved with M_R reactions $R_{j=1,\ldots,M_R}$. Denote by integer x_i the number of molecules of species S_i. The state of the system is represented by the molecular number of every reacting species:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_S} \end{pmatrix} \in \mathbb{N}^{N_S}. \tag{1}$$

Each reaction R_j is characterized by a rate function $a_j(x)$ and a state change vector $v_j \in \mathbb{N}^{N_S}$. We write

$$R_j = (a_j, v_j). \tag{2}$$

Given system state x at time t, on the infinitesimal time interval $[t, t + \mathrm{d}t)$, the occurrences of the reactions are independent of each other and the probability of each reaction R_j is given by $a_j(x)\mathrm{d}t$. After reaction R_j, the state of the system changes from x to $x + v_j$. The stoichiometric matrix $v \in \mathbb{R}^{N_S \times M_R}$ is defined to be the matrix with the jth column to be v_j, i.e.

$$v = \begin{pmatrix} v_1^1 & v_2^1 & \dots & v_{M_R}^1 \\ v_1^2 & v_2^2 & \dots & v_{M_R}^2 \\ \dots & & & \\ v_1^{N_S} & v_2^{N_S} & \dots & v_{M_R}^{N_S} \end{pmatrix}. \tag{3}$$

The time evolution of the probability distribution of the state variable $P(x,t)$ is described by the forward Kolmogorov master equation:

$$\frac{\partial P(x,t)}{\partial t} = \sum_j (a_j(x - v_j)P(x - v_j, t) - a_j(x)P(x,t)). \tag{4}$$

Let x_t be the state variable at time t. For any arbitrary smooth function f, the observable $u(x,t) = \mathbb{E}_{x_0 = x} f(x_t)$ satisfies the following backward Kolmogorov master equation:

$$\frac{\partial u(x,t)}{\partial t} = Lu(x,t) = \sum_j a_j(x)(u(x + v_j, t) - u(x,t)), \tag{5}$$

where L is defined to be the generator of the Markov process in the state space associated with the chemical kinetic system.

It is well known [23] that when the molecular numbers of all reacting species in the chemical kinetic system increase to infinity, the stochastic process described by (4) converges to the following deterministic ODE on finite time intervals:

$$\dot{y}(t) = \sum_j v_j b_j(y(t)), \tag{6}$$

where we have rescaled the state variable and the reaction rate with the system size Ω such that

$$y = \frac{x}{\Omega}, \quad b(y) = \frac{a(x)}{\Omega}. \tag{7}$$

It can be shown [23] that the strength of the noise diminishes on the scale of $1/\sqrt{\Omega}$ when $\Omega \to \infty$. The limiting dynamics (6) implies that when the system size Ω is large (but not ∞), the state of the system is more likely to be found at the stable stationary states of the deterministic system. If ergodicity is assumed, the precess will be able to escape from any local stable set and switch to different metastable states. Driven by noise of small magnitudes, the time scale between the switchings is much longer than that for the relaxation to the stable states. The Large Deviation Theory (LDT) of stochastic processes [18–20] gives probabilistic estimates for rare events. Now we want to give a brief review for LDT. Suppose $y, y' \in \mathbb{R}^{N_S}$, the local function is defined to be

$$\ell(y, y') = \sup_{\theta \in \mathbb{R}^{N_S}} \left(\langle \theta, y' \rangle - \sum_j b_j(y)(\mathrm{e}^{\langle \theta, v_j \rangle} - 1) \right). \tag{8}$$

For any absolutely continuous function $\varphi : [0,T] \to \mathbb{R}^{N_s}$, the LDT action functional has the following form:

$$I_{[0,T]}[\varphi] = \int_0^T \ell(\varphi(t), \dot{\varphi}(t))\,\mathrm{d}t. \tag{9}$$

We set $I_{[0,T]}[\varphi] = \infty$ when φ is not absolutely continuous. Denote by φ^{Ω} the process generated by the following rescaled equation with the initial state φ_0:

$$\frac{\partial P(y,t)}{\partial t} = \sum_j \Omega\left(b_j\left(y - \frac{v_j}{\Omega}\right) P\left(y - \frac{v_j}{\Omega}, t\right) - b_j(y)P(y,t) \right). \tag{10}$$

Consider the space D_s consisting of bounded functions defined from [0,T] to $\mathbb{R}^{N_s}$ with right hand continuity and left hand limits endowed with the supremum norm. For any measurable set $S \in D_s$, let

$$I_{[0,T]}[S] = \inf_{\{\varphi:\varphi\in S,\varphi(0)=\varphi_0\}} I_{[0,T]}[\varphi]. \tag{11}$$

It is proved in [18–20] that under certain regularity assumptions on the reaction rates and the accessible boundaries of the reactions, the following Large Deviation Principle holds for any closed set $C \in D_s$ and open set $O \in D_s$:

$$\begin{aligned} &\limsup_{\Omega\to\infty} \frac{1}{\Omega} \ln P\{\varphi^\Omega \in C\} \leqslant -I_{[0,T]}[C], \\ &\liminf_{\Omega\to\infty} \frac{1}{\Omega} \ln P\{\varphi^\Omega \in O\} \geqslant -I_{[0,T]}[O]. \end{aligned} \tag{12}$$

The above estimate means that we can assign a probability for each possible reaction path φ in the configuration space such that when $\Omega \gg 1$ and $\delta \ll 1$,

$$P\{\|\varphi^\Omega - \varphi\| < \delta\} \approx \exp\{-\Omega I_{[0,T]}[\varphi]\}. \tag{13}$$

The Large Deviation Theory suggests that the optimal transition paths should be the most probable transition paths that minimize the LDT action functional (9). Numerical schemes using optimization techniques for solving the optimal transition paths can also be proposed along this line. The challenge lies in the fact that the local function is not defined in an explicit form and its computation could be very involved when the dimension of the system is large. First order derivative criterion for the supremum in (8) gives the following equation:

$$y' = \sum_j e^{\langle\theta^*, v_j\rangle} b_j(y) v_j, \tag{14}$$

where θ^* is the value of θ at which the supremum is attained. The cost of solving the above nonlinear equation grows quadratically with the dimension of the system. The dependence of the right hand side of (14) on y makes the pre-computing of its factorization not worthwhile. The possible degeneracy of the stoichiometric matrix also raises the issue of the non-uniqueness of the solution for (14), which may cause confusion on the definition of the optimal paths. In the following, we will introduce the reaction advancement coordinate to handle the computation of the LDT action functional, which will overcome the above difficulties.

3. Optimal transition paths of chemical kinetic systems

In this section, we want to identify the optimal transition paths as minimizers of the LDT action functional and derive the dynamical equations satisfied by the optimal paths. First, we want to do some simplification of the chemical kinetic system (2) using the special structure of chemical reactions to make the LDT action functional easier to work with. Define the auxiliary variable z such that

$$y = y_0 + \sum_j z^j v_j. \tag{15}$$

The variable z introduced here is usually called the reaction advancement coordinate [24]. Each z^j measures the total occurrence of reaction R_j. Notice that $z \in \mathbb{R}^{M_R}$. In terms of z, the reactions take the following form:

$$R_j = (c_j, e_j), \quad j = 1, \ldots, M_R, \tag{16}$$

where

$$c_j(z) = b_j(y), \tag{17}$$

and $\{e_j\}$'s form an orthonormal basis of $\mathbb{R}^{M_R}$ such that

$$e_j^i = \delta_{ij}. \tag{18}$$

In other words, the stoichiometric matrix under the reaction advancement coordinate z becomes an identity matrix.

Notice that the total occurrence of a reaction is a non-decreasing function of time. As a consequence, any realistic reaction pathway $\varphi(t) : ([0,T] \to \mathbb{R}^{M_R}$ in the configuration space specified by the reaction advancement coordinate z will have the property of

$$\dot\varphi(t) \geqslant 0. \tag{19}$$

The local function $\ell(\cdot,\cdot)$ has the following form in terms of z:

$$\ell(\varphi, \dot\varphi) = \sup_{\theta\in\mathbb{R}^{M_R}} \left(\langle\theta, \dot\varphi\rangle - \sum_j c_j(\varphi)(e^{\theta_j} - 1) \right). \tag{20}$$

When $c_j(\varphi) = 0$, it is very easy to verify that unless $\dot\varphi^j = 0$, the above local rate function takes the value of positive infinity, which means the path $\varphi(\cdot)$ is unlikely to be followed by the real reaction event according to the LDT estimate (13). Since we

are only interested the realistic paths, the following discussion is confined to the case of $\dot{\varphi}^j = 0$ if $c_j(\varphi) = 0$. Using (19), the local function $\ell(\cdot,\cdot)$ can be explicitly solved such that:

$$\ell(\varphi, \dot{\varphi}) = \sum_j \dot{\varphi}^j \ln\left(\frac{\dot{\varphi}^j}{c_j(\varphi)}\right) - \dot{\varphi}^j + c_j(\varphi). \tag{21}$$

Notice that the above equality is still well defined when $\dot{\varphi}^j$ is equal to zero due to the continuity of function $y = x\ln(x)$ at zero. It can be seen that the introduction of the reaction advancement coordinate greatly simplifies the evaluation of the local function. It avoids solving a large scale optimization problem involving complex matrix computation for Eq. (14) when the dimension of the system is high.

The LDT action functional now has the following form:

$$I_{[0,T]}[\varphi] = \int_0^T \left(\sum_j \dot{\varphi}^j \ln\left(\frac{\dot{\varphi}^j}{c_j(\varphi)}\right) - \dot{\varphi}^j + c_j(\varphi)\right) \mathrm{d}t. \tag{22}$$

Suppose that we have two metastable sets $S_1, S_2 \subset \mathbb{R}^{N_S}$. The optimal transition path ψ between S_1 and S_2 should minimize the LDT functional such that

$$I_{[0,T]}[\psi] = \min_{\varphi} I_{[0,T]}[\varphi], \tag{23}$$

subjecting to the following boundary condition:

$$y_0 + v \cdot \varphi(0) \in S_1, \quad y_0 + v \cdot \varphi(T) \in S_2, \tag{24}$$

where $v = (v_1, \ldots, v_{M_R})$ is the Stoichiometric matrix of the chemical kinetic system. For most cases of interest, the metastable states are attractors of the deterministic system defined by (6). For simplicity, we will focus on the situation when the metastable states are stable fixed points of the above equation. More general situations when the metastable sets are limit cycles or chaotic attractors will be discussed in future. Suppose $A, B \in \mathbb{R}^{N_S}$ are two stationary states of (6). The boundary condition (24) reduces to

$$y_0 + v \cdot \varphi(0) = A, \quad y_0 + v \cdot \varphi(T) = B. \tag{25}$$

It can be seen that the undetermined boundary condition (24) and (25) follows from the definition of the reaction advancement coordinate (15). It represents the fact that at metastable states, the reactions are in equilibrium but not shut off. The reacting species are still being synthesized and degraded, although 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 (22) since written in z, the processes in balance satisfy the following equation:

$$\dot{z}^j(t) = c_j(z(t)), \tag{26}$$

which sets the value of the LDT action functional (22) to zero. The variational principle implies that the optimal transition path ψ should satisfy the following equation subjecting to boundary condition (25):

$$\ddot{\psi}^k = \dot{\psi}^k \sum_\ell \left(\dot{\psi}^\ell \left(\frac{\nabla_\ell c_k(\psi)}{c_k(\psi)} - \frac{\nabla_k c_\ell(\psi)}{c_\ell(\psi)}\right) + \nabla_k c_\ell(\psi)\right) \tag{27}$$

for each $k \in \{1, \ldots, M_R\}$. The details of the derivation cab be found in Appendix.

4. Minimum Action Method for chemical kinetic systems

Now we want to discuss the numerical method for solving the optimal transition paths. The method that we are introducing here is a modification of the Minimum Action Method proposed in [17] for stochastic chemical kinetic systems governed by diffusion processes. The main idea is to proceed by evolving curves in the space of all possible transition paths with a dynamics that relaxes to the most probable path such that a minimizer for the optimization problem (23) subjecting to boundary condition (25) is obtained. The same idea has been applied in the original Minimum Action Method to solve the optimal transition paths of stochastic partial differential equations driven by space–time white noise with fixed boundary conditions [21]. New numerical challenges for chemical reacting systems arise from the non-smoothness of the LDT action functional and the undetermined boundary condition (25). The discontinuity of the LDT action functional $I_{[0,T]}[\varphi]$ is induced by the presence of the logarithmic function, which takes the value of negative infinity when $\dot{\varphi}(t) \leqslant 0$. Although the optimal transition path should have non-negative time derivatives, we can not exclude the possibility of $\dot{\varphi}(t) \leqslant 0$ during the computation, which will blow up the action functional and make any numerical scheme unstable. As shown before, the computation of the LDT action functional becomes much easier after reformulating the system in the reaction advancement coordinate. We will further introduce a smoothing technique based on the Taylor expansion to moderate $I_{[0,T]}[\varphi]$. To deal with the undetermined boundary condition (25), we adopt a standard constraint optimization technique to include penalty terms in the LDT action functional. A multiscale technique will be employed to deal with the stiffness generated by the

boundary penalty terms. To avoid overemphasizing the technical details of the scheme, we simply call the overall scheme discussed here the 'Minimum Action Method'.

First we want to address the issue of smoothing the LDT action functional $I_{[0,T]}[\varphi]$. To mollify the logarithmic function in $I_{[0,T]}[\varphi]$, we perform the following modification for the function $y = \ln(x)$ through Taylor expansion at $x = 1$:

$$\ln^m(x) \stackrel{\text{def}}{=} \begin{cases} (x-1) + \cdots + \frac{(-1)^{m-1}(x-1)^m}{m}, & x \leqslant 1, \\ \ln(x), & x > 1. \end{cases} \tag{28}$$

For any finite positive integer m, function $\ln^m(x)$ defined as above will give us an mth order differentiable function defined on the domain of all the real numbers. If we define the value of $\ln(x)$ to be $-\infty$ for $x \leqslant 0$, $\ln^m(x)$ converges to $\ln(x)$ pointwisely when $m \to \infty$. Replacing $\ln(x)$ with $\ln^m(x)$ in (22), we have the following moderated LDT action functional:

$$I^m_{[0,T]}[\varphi] = \int_0^T \left(\sum_j \dot{\varphi}^j \ln^m \left(\frac{\dot{\varphi}^j}{c_j(\varphi)} \right) - \dot{\varphi}^j + c_j(\varphi) \right) \mathrm{d}t. \tag{29}$$

Now we want to make a comparison with the Minimum Action Method for chemical kinetic systems proposed in [17]. Written in reaction advancement coordinate z, the limiting diffusion process for the chemical kinetic system when the system size Ω is large has the form:

$$\dot{z}^j(t) = c_j(z(t)) + \frac{\sqrt{c_j(z(t))}}{\sqrt{\Omega}} \dot{w}^j_t, \quad j = 1, \ldots, M_R. \tag{30}$$

It is well known [25] that when $\Omega \to \infty$, the limiting diffusion process (30) satisfies the Large Deviation Principle with the Freidlin–Wentzell action functional:

$$I^D_{[0,T]}[\varphi] = \frac{1}{2} \int_0^T \sum_j c_j^{-1}(\varphi)(\dot{\varphi}_j - c_j(\varphi))^2 \mathrm{d}t. \tag{31}$$

If we approximate the logarithmic function in the LDT action functional with the first order term in the Taylor expansion on the whole real domain:

$$\ln(1 + x) \approx x, \tag{32}$$

we find that

$$I_{[0,T]} \approx 2 I^D_{[0,T]}. \tag{33}$$

In other words, the Freidlin–Wentzell action functional for the diffusion limit of stochastic chemical kinetic systems is a first order approximation for the LDT action functional. Consequently, the Minimum Action Method proposed in [17], which finds the optimal transition path by minimizing the above Freidlin–Wentzell action functional, can be also seen as a first order approximation for the scheme proposed here.

To handle the undetermined boundary condition (25), we adopt the same approach in [17] to introduce the following augmented LDT action functional that includes boundary penalty terms:

$$I^m_{[0,T]}[\varphi, \mu] = I^m_{[0,T]}[\varphi] + \frac{1}{2\mu} \{ |y_0 + v \cdot \varphi(0) - A|^2 + |y_0 + v \cdot \varphi(T) - B|^2 \}, \tag{34}$$

where $\mu > 0$ is the penalty parameter. By driving μ to zero and seeking approximate minimizers for $I^m_{[0,T]}[\varphi, \mu]$, we penalize constraint violations at the boundary with increasing severity. As $\mu \to 0$, the solutions should converge to the minimizer of the action functional $I^m_{[0,T]}[\varphi]$ with the constraint of boundary condition (25) being satisfied. In practice, we choose a sequence of values with $\mu_k \to 0$ as $k \to \infty$ and solve the following minimization problem:

$$I^m_{[0,T]}[\psi_k, \mu_k] = \min_{\varphi} I^m_{[0,T]}[\varphi, \mu_k]. \tag{35}$$

As mentioned before, the main idea of the Minimum Action Method is to evolve curves in the space of transition paths and search for minimizers of the LDT action functional. The simplest dynamics for this relaxation will be the following gradient flow in the path space:

$$\frac{\partial \varphi}{\partial s} = -\frac{\delta I^m_{[0,T]}[\varphi, \mu]}{\delta \varphi}, \tag{36}$$

where $\varphi(\alpha, s)$, $(0 \leqslant \alpha \leqslant T)$ denotes the evolving curve in the path space. The direct time discretization of the above equation with an adaptive time step such that a line minimization along the action gradient is reached at each step will amount to the Steepest Descent method. More efficient methods like Conjugate Gradient method or Quasi-Newton method [26] can also be adopted to evolve the curves and solve (35). Special methods tailored for the optimization in the path space will be addressed in future. The basic procedure of this algorithm can be described as the following:

S1. Initiate $k = 0$, choose penalty parameter $\mu_0 > 0$, error threshold $\lambda_0 > 0$, and starting curve φ^0

S2. Solve (35) with starting curve φ^k to reach an approximate solution ψ^k such that the absolute value of the action gradient on the right hand side of (36) is smaller than λ_k;

S3. Reset starting curve $\varphi^{k+1} = \psi^k$ and reduce parameters $0 < \mu_{k+1} < \mu_k$, $0 < \lambda_{k+1} < \lambda_k$. Repeat **S2**.

The space discretization of the LDT action functional is straightforward. We partition the domain $[0,T]$ with a mesh of size $\Delta\alpha = 1/L$ and define the grid points:

$$\alpha_\ell = \ell\Delta\alpha, \quad \ell = 0, 1, \ldots, L. \tag{37}$$

We use the midpoint rule to discretize the integral in the LDT action functional, which will give us

$$I^m_{[0,T]}(\Phi) = \sum_j \sum_{\ell=0}^{L-1} \frac{\Phi^j_{\ell+1} - \Phi^j_\ell}{\Delta\alpha} \ln^m \left(\frac{\Phi^j_{\ell+1} - \Phi^j_\ell}{\Delta\alpha} c_j^{-1} \left(\frac{\Phi_{\ell+1} + \Phi_\ell}{2} \right) \right) - \frac{\Phi^j_{\ell+1} - \Phi^j_\ell}{\Delta\alpha} + c_j \left(\frac{\Phi_{\ell+1} + \Phi_\ell}{2} \right), \tag{38}$$

where Φ_ℓ denotes the numerical approximation to $\varphi(\alpha_\ell)$. The LDT action functional with boundary penalty can be defined accordingly to be

$$I^m_{[0,T]}(\Phi, \mu) = I^m_{[0,T]}(\Phi) + \frac{1}{2\mu}\{|y_0 + v \cdot \Phi_0 - A|^2 + |y_0 + v \cdot \Phi_T - B|^2\}. \tag{39}$$

The introduction of the penalty parameter μ can generate stiffness when evolving the curve in the path space. This can be seen from dynamics (36). After space discretization, it takes the form:

$$\begin{cases} \dot{\Phi}_\ell = -\nabla_{\Phi_\ell} I^m_{[0,T]}(\Phi), \quad \ell = 1, \ldots, L-1, \\ \dot{\Phi}_0 = -\nabla_{\Phi_0} I^m_{[0,T]}(\Phi) - \frac{v^T}{\mu}\{y_0 + v \cdot \Phi_0 - A\}, \\ \dot{\Phi}_L = -\nabla_{\Phi_L} I^m_{[0,T]}(\Phi) - \frac{v^T}{\mu}\{y_0 + v \cdot \Phi_L - B\}, \end{cases} \tag{40}$$

where $\Phi_{\ell=1,\ldots,L-1}$ denotes the inner points and (Φ_0,Φ_L) denotes the boundary points of the transition path. There is a time scale separation measured by μ in Eq. (40). The slow variables $\Phi_{\ell=1,\ldots,L-1}$ evolve on a slow time scale of order O(1), while the fast variables (Φ_0,Φ_L) evolve on a fast time scale of order $O(\mu)$. The slow dynamics consists of the first $L-1$ equations in (40) while the fast dynamics is given by the last two equations. Direct simulation of system (40) will entail very small time steps prescribed by μ 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 $\Phi_{\ell=1,\ldots,L-1}$, which describe the optimal transition paths. Here we want to provide a multiscale scheme originally proposed in [17], which consists 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**, the nested scheme does the following to solve the minimization problem (35):

M1. *Inner solver*: Fix slow variables $\Phi_{\ell=1,\ldots,L-1}$ as parameters. Solve the following minimization problem with respect to the fast variables (Φ_0,Φ_L):

$$I^m_{[0,T]}(\Psi_0, \Phi_{1,\ldots,L-1}, \Psi_L, \mu_k) = \min_{(\Phi_0,\Phi_L)} I^m_{[0,T]}(\Phi_0, \Phi_{1,\ldots,L-1}, \Phi_L, \mu_k), \tag{41}$$

stop when an approximate stationary solution (Ψ_0,Ψ_L) is reached such that for $\ell = 1,\ldots,L-1$,

$$\max\left\{ \left| \nabla_{\Phi_0} I^m_{[0,T]}(\Phi, \mu_k) \right|, \left| \nabla_{\Phi_L} I^m_{[0,T]}(\Phi, \mu_k) \right| \right\} \leqslant \left| \nabla_{\Phi_\ell} I^m_{[0,T]}(\Phi, \mu_k) \right|. \tag{42}$$

M2. *Outer solver*: Minimize $I^m_{[0,T]}(\Psi_0, \Phi_{1,\ldots,L-1}, \Psi_L, \mu_k)$ with respect to the slow variables $\Phi_{\ell=1,\ldots,L-1}$ for one step. Repeat **M1**.

The inner and outer solvers can be implemented with any optimization algorithm. The rationale behind the above scheme can be illustrated through (40). When the slow variables are fixed, the LDT action functional is still a well defined potential for the fast variables. The fast dynamics should reach a quasi-equilibrium on a time scale of $O(\mu)$, which is much faster than the O(1) time scale on which the slow dynamics advances. At the same time, the quasi-equilibrium state of the fast dynamics depends continuously on the slow variables and only need to be updated when there is a change in the latter. The same idea has been applied to the stochastic simulation of chemical kinetic systems with multiple time scales (see [27,28] and the reference therein). As shown in [27,28], if we choose appropriate parameters in the inner and outer solvers, the above scheme can achieve an increasing accuracy for smaller μ with a computational cost independent of μ. Notice that we have proposed a stopping criterion for step **M1** different from that proposed in [17], where it is required that the absolute value of the action gradient with respect to the fast variables to be smaller than the same threshold λ_k as in the outer solver. Numerical experiments have shown that the new stopping criterion is more efficient.

5. Numerical Example I: the toggle switch model

For a numerical example, we now consider the following Toggle Switch Model [10], which artificially realizes a genetic switch consisting 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 that codes for the other protein, thereby represses its expression. The transition paths of slightly different versions of the same system have been studied in [14–17]. The deterministic equation describing the system has the following form:

$$\begin{aligned}\dot{u} &= \tfrac{\alpha_1}{1+v^\beta} - u,\\ \dot{v} &= \tfrac{\alpha_2}{1+u^\gamma} - v,\end{aligned} \tag{43}$$

where u is the concentration of repressor 1, v is the concentration of repressor 2. $\alpha_{i=1,2}$ are the effective rates of synthesis of repressor 1 and repressor 2, respectively. There are 4 reactions in the system, namely the synthesis and degradations of repressor 1 and repressor 2. We choose the parameters in the Toggle Switch Model (43) to be the same as in [14] such that

$$\alpha_1 = 156, \quad \alpha_2 = 30, \quad \beta = 3, \quad \gamma = 1. \tag{44}$$

Incorporating the stochastic effects, the backward master equation has the following form:

$$\frac{\partial v}{\partial t} = \Omega\left\{\frac{\alpha_1}{1+v^\beta}\left(v\left(u+\frac{1}{\Omega},v,t\right)-v\right) + u\left(v\left(u-\frac{1}{\Omega},v,t\right)-v\right) + \frac{\alpha_2}{1+u^\gamma}\left(v\left(u,v+\frac{1}{\Omega},t\right)-v\right) + v\left(v\left(u,v-\frac{1}{\Omega},t\right)-v\right)\right\}. \tag{45}$$

The reaction advancement coordinate z can be given according to the state change vectors:

$$\begin{pmatrix}u\\ v\end{pmatrix} = \begin{pmatrix}u_0\\ v_0\end{pmatrix} + z_1\begin{pmatrix}1\\ 0\end{pmatrix} + z_2\begin{pmatrix}-1\\ 0\end{pmatrix} + z_3\begin{pmatrix}0\\ 1\end{pmatrix} + z_4\begin{pmatrix}0\\ -1\end{pmatrix}. \tag{46}$$

The reaction rates in terms of z are

$$\begin{aligned}c_1(z) &= \frac{\alpha_1}{1+(v_0+z_3-z_4)^\beta},\\ c_2(z) &= u_0+z_1-z_2,\\ c_3(z) &= \frac{\alpha_2}{1+(u_0+z_1-z_2)^\gamma},\\ c_4(z) &= u_0+z_3-z_4.\end{aligned} \tag{47}$$

The metastable states in the system, as plotted in Fig. 1, are

$$A = (0.00588, 29.825), \quad B = (154.897, 0.192). \tag{48}$$

In the Minimum Action Method, the evolving path is discretized with $L = 40$ nodes equally distributed as in (37). The truncation number in (29) for smoothing the action functional is chosen to be

$$m = 1, 2, 3, \ldots \tag{49}$$

We choose the computational parameters for the numerical iterations according to

$$\mu_k = \lambda_k = \lambda^{-k}, \quad k = 0, 1, \ldots, \tag{50}$$

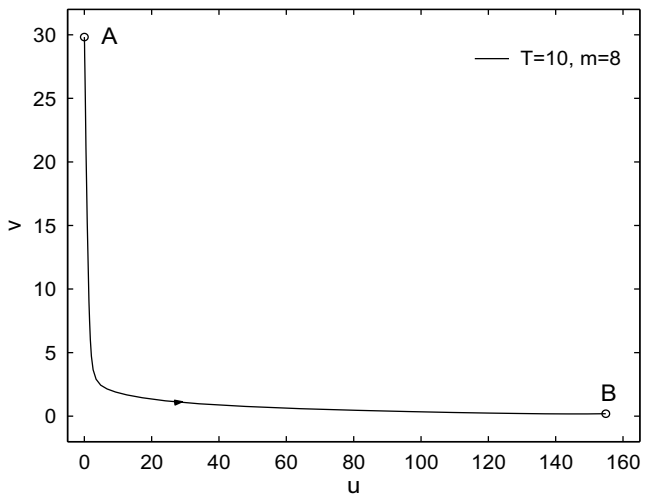

Fig. 1. Optimal transition path for the Toggle Switch Model from metastable state A to metastable state B when $T = 10$.

for some $\lambda < 1$. The Conjugate Gradient method is adopted for both inner and outer solvers. In Fig. 1, we show the optimal transition paths for the time horizon of $T = 10$. The results qualitatively agree with those in [14] obtained using the WKB method. Fig. 2 gives the values of the LDT action functional of the optimal transition paths for different values of truncation number m. The transition path we get here is also highly consistent with that in [17] computed using the Freidlin–Wentzell action functional for the limiting diffusion process. In other words, for this example, the higher order corrections in the LDT action functional does not really change the optimal transition path.

6. Numerical example II: the *E. Coli* Lactose Operon

Now we consider another numerical example, namely the Lac Operon Induction Model in *E. coli* [12]. The Lac Operon is a set of genes that encodes proteins required to import and digest the disaccharide lactose. At the molecular level, induction of the operon in a non-glucose containing medium by a non-metabolizable inducer is controlled by a positive feedback loop involving lac permease, lac repressor and lac inducer. In the absence of inducer, transcription of the operon is repressed by the binding of repressor to the lac operator. When the inducer is added to the culture medium, it enters the cell where it may indirectly initiate transcription by binding to the repressor so as to reduce repressor's affinity to the operator. Once the repressor is disengaged from the operator, transcription of the operon may be initiated. Lac permease, a gene product of the operon, serves to import inducer, therefore positively affects its own expression.

In [12], one version of the Lactose Operon Induction Model is proposed in the form of the following ODE:

$$\begin{aligned} \dot{Y}_t &= k_1 O_T \frac{1 + K_1 I_t^2}{1 + K_1 I_t^2 + K_2 R_T} - k_2 Y_t, \\ \dot{I}_t &= \frac{\alpha I_{ex} Y_t}{\beta + I_{ex}} - \delta(I_t - I_{ex}) - k_2 I_t, \end{aligned} \tag{51}$$

where Y and I represent the cell-associated permease and intracellular inducer concentrations, respectively. R_T and O_T denote the total concentrations, bound or unbound, of repressor and operator. I_{ex} is the extracellular concentration of inducer. There are totally 5 reactions represented by the above equation. The first reaction is the generation of permease through operon transcription:

$$R_{GEN} = k_1 O_T \frac{1 + K_1 I^2}{1 + K_1 I^2 + K_2 R_T}. \tag{52}$$

The second reaction is the active transport of inducer via permease:

$$R_{ACTIVE} = \frac{\alpha I_{ex} Y}{\beta + I_{ex}}, \tag{53}$$

where α and β denote the permease turnover number and the permease saturation constant, respectively. The turnover number represents the maximum number of inducer molecules that can be transported into the cell per permease in unit time. The third reaction is the transport of inducer via processes independent of lactose permease. The rate function is given as

$$R_{FACILITATED} = \delta(I - I_{ex}), \tag{54}$$

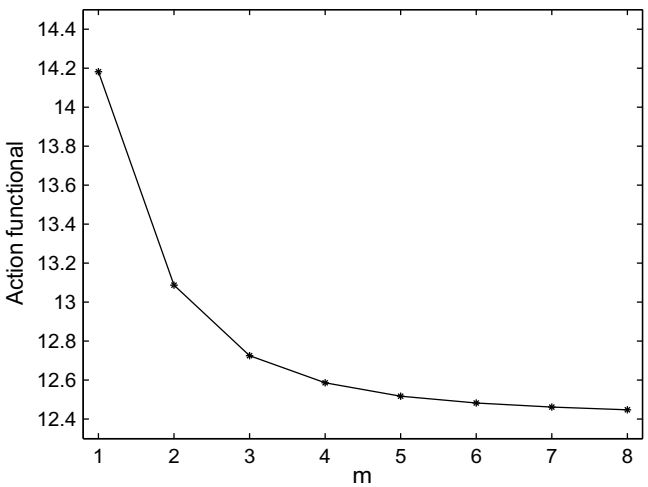

Fig. 2. LDT action functionals of the optimal transition paths in the Toggle Switch Model for different values of the truncation number m.

where δ represents the transport coefficient. This functionality allows the concentration dependent nature of the facilitated transport to be represented and naturally leads to an equilibration between I and I_{ex} under balanced growth conditions. The fourth and fifth reactions are the dilutions of Y and I due to growth, which is represented by the common reaction constant k_2. The parameters listed in Table 1 are chosen to be the same as in [12], in which the details on the sources of the parameters can also be found.

The metastability of the Lactose Operon Induction Model is illustrated in Figs. 3 and 4 through the hysteresis curves of the permease and intracellular inducer populations. In these hysteresis curves, the population of extracellular inducer I_{ex} is increased and decreased quasi-statically. For each I_{ex}, we solve system (51) with an adaptive time step until a stable state is reached. There are two types of stable states for (51), the first one is the uninduced state in which both the permease and the intracellular inducer are in low molecular numbers, and the second one is the induced state with permease and intracellular inducer being in large molecular numbers. It can be seen that when the population of the extracellular inducer is low, the only stable state is the uninduced state, and when the population of the extracellular inducer is high, the only stable state is the induced state. For the extracellular inducer populations between certain critical values, both the induced and uninduced states are stable. Due to this separation of the stable states, the system exhibits a memory effect that the current state depends on its history.

It can also be seen from Figs. 3 and 4 that there is a population scale separation between permease and intracellular inducer. In the uninduced states, the molecular number of permease is of the order of $O(10^0)$ and the molecular number of intracellular inducer is in the range of $O(10^4)$. In the induced states, the orders of the molecular numbers of permease and intracellular inducer become $O(10^3)$ and $O(10^7)$, respectively. In other words, there is a $O(10^4)$ scale separation in population between the two reacting species, which will induce a stiffness in the direct minimization of the LDT action functional. Using the fact that stochastic effects have a less significant impact on reacting species in high molecular numbers, a common simulation technique dealing with the stiffness of stochastic chemical kinetic systems with multiple population scales is to either omit the fluctuations in high population species, or moderate the perturbation with a Brownian noise using the limit diffusion process [29–31]. As shown in [23], this methodology is equivalent to approximating the original system with lower order terms in the Taylor series, both in the backward equation (5) and the LDT action functional (29). Adopting the same approach, we add the random perturbation only to the evolution of permease in equation (51), which will give the following stochastic integral equation for the Lac Operon Model:

Table 1
Parameters in the Lactose Operon Induction Model

Parameter	Value
k_1	$9\ \text{min}^{-1}$
k_2	$0.0055\ \text{min}^{-1}$
O_T	0.6022
K_1	3.309×10^{-8}
$K_2 R_T$	1×10^5
α	$6 \times 10^4\ \text{min}^{-1}$
β	3.011×10^5
δ	$0.82\ \text{min}^{-1}$

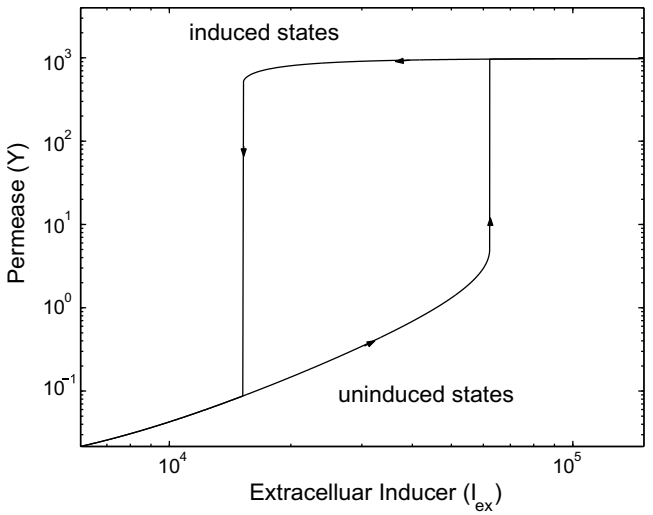

Fig. 3. Hysteresis curves depicting the equilibrium molecular numbers of permease Y at different populations of extracellular inducer I_{ex}.

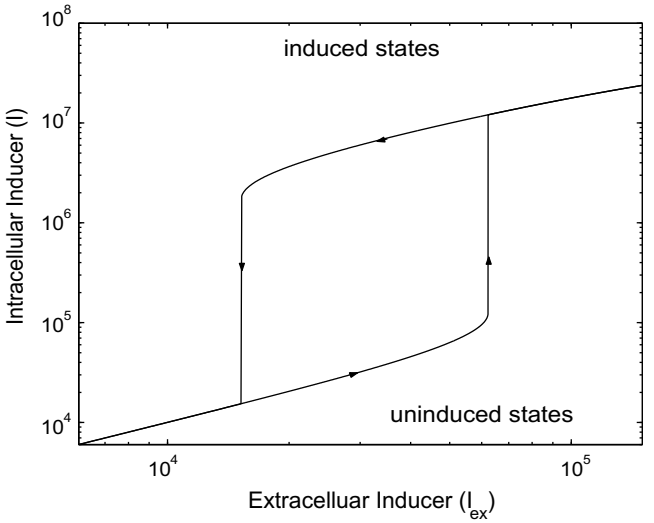

Fig. 4. Hysteresis curves depicting the equilibrium molecular numbers of intracellular inducer I at different populations of extracellular inducer I_{ex}.

$$\begin{aligned} Y_t &= Y_0 + X^1\left(\int_0^t k_1 O_T \frac{1+K_1 I_s^2}{1+K_1 I_s^2 + K_2 R_T}\,\mathrm{d}s\right) - X^2\left(\int_0^t k_2 Y_s\,\mathrm{d}s\right), \\ I_t &= I_0 + \int_0^t \left(\frac{\alpha I_{ex} Y_s}{\beta + I_{ex}} - \delta(I_s - I_{ex}) - k_2 I_s\right)\mathrm{d}s, \end{aligned} \tag{55}$$

where $X^{i=1,2}$ are independent Poisson processes with

$$\mathbb{E}(X^i(u)) = u. \tag{56}$$

For simplicity, we are writing the above equation in original variables without rescaling with the system size Ω.

To apply the Minimum Action Method, we first need to identify the reaction advancement coordinate z as

$$\begin{pmatrix} Y \\ I \end{pmatrix} = Z_1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + Z_2\begin{pmatrix} 0 \\ 1 \end{pmatrix} + Z_3\begin{pmatrix} 0 \\ -1 \end{pmatrix} + Z_4\begin{pmatrix} -1 \\ 0 \end{pmatrix} + Z_5\begin{pmatrix} 0 \\ -1 \end{pmatrix}. \tag{57}$$

We solve the optimal transition paths for the time horizon

$$T = 80. \tag{58}$$

The space discretization parameter for the transition path is fixed such that

$$L = 100. \tag{59}$$

The computational parameters μ_k and λ_k are chosen to be the same as in (50). Fig. 5 gives the transition path when $I_{ex} = 62188$. The values of the LDT action functional (31) for the optimal transition paths obtained by using different truncation number m is provided in Fig. 6, with the system size Ω chosen to be 100.

7. Conclusion and future directions

We proposed a modified Minimum Action Method (MAM) for finding the optimal transition paths of chemical kinetic systems with large system size based on the Large Deviation Theory for Markov processes. The optimal transition paths are solved using standard constraint optimization techniques. A smoothing technique is introduced to moderate the LDT action functional. To handle the numerical stiffness generated by the boundary constraints, a multiscale scheme is employed to handle the inner and boundary points in a nested structure. The method is a modification of the previously proposed Minimum Action Method for chemical Langevin equations. By incorporating the higher order terms in the LDT action functional, the accuracy of the optimal paths is improved. New implementation of stopping criterion in the inner solver also speeds up the algorithm.

The major limitation of the Minimum Action Method discussed here is that it turns very inefficient when the time horizon T becomes large, since more space grid points (37) are needed for an accurate representation of the transition paths. At the same time, an asymptotic estimate for the mean transition time can be given as the following [19,20]:

$$\tau \approx \exp\left\{\Omega \inf_T I_{[0,T]}[\psi_T]\right\}, \tag{60}$$

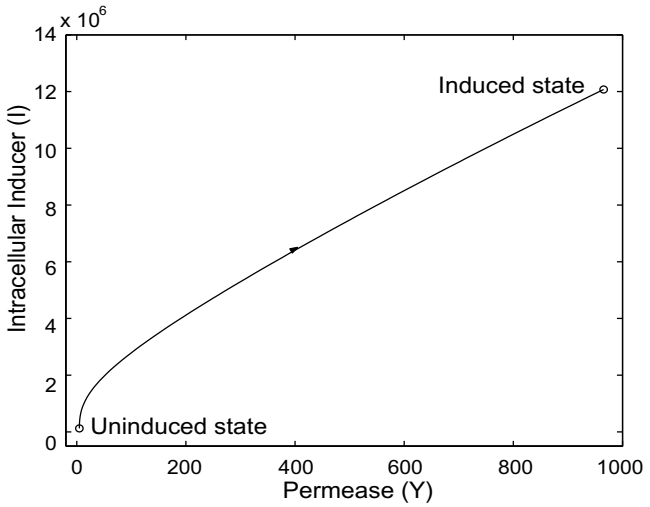

Fig. 5. Transition path in terms of molecular numbers of Permease and Intracellular inducers from uninduced state to induced state for $T = 80$ when the extracellular inducer has the population of $I_{ex} = 62188$.

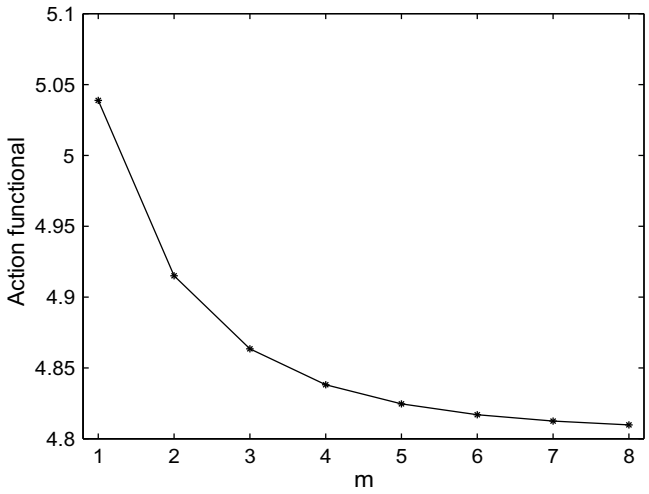

Fig. 6. Values of the LDT action functional for different truncation number m when the population of the extracellular inducer is $I_{ex} = 62188$.

which means that the evaluation of the mean transition time entails the computation of the optimal transition path when $T \to \infty$. On the other hand, for systems like the Lac Operon induction, different versions of mathematical models with a wide range of parameters have been proposed to explain the behaviors of the system [11,12]. Substituting the minimum LDT action functionals obtained for finite time intervals into the above estimate (60) will give a transition time much longer than the real biological processes. This could be a result of either the inefficiency of MAM on the infinite time horizon or the inaccuracy of the model and its parameters. Notice that the situation when $T \to \infty$ has been handled successfully in [22,32] under the assumption of the non-degeneracy of the noise. Therefore, future work will involve designing efficient numerical methods for finding optimal transition paths when $T \to \infty$ for systems driven by general noises and using the methods as a model validation tool.

Acknowledgements

The research is supported by NSF grant DMS-0609315. The author is grateful to Adam Shwartz, Weinan E and Eric Vanden–Eijnden for stimulating discussions and insightful comments. The author also wants to thank Princeton Institute for Computational Science and Engineering (PICSciE) and High Performance Computing Center at Michigan State University for providing their computing resources.

Appendix A. Minimizers of the LDT action functional

Denote by ψ a minimizer of the action functional $I_{[0,T]}$ as defined in (22). Suppose η is a smooth function with $\eta_0 = \eta_T = 0$. Differentiating $I_{[0,T]}[\psi + \varepsilon\eta]$ with respect to ε at zero and taking the first order derivative to be zero, we have

$$\int_0^T \sum_j \left(\dot{\eta}^j \ln\left(\frac{\dot{\psi}^j}{c_j(\psi)}\right) + \left(1 - \frac{\dot{\psi}^j}{c_j(\psi)}\right) \sum_\ell \nabla_\ell c_j(\psi)\eta^\ell \right) \mathrm{d}t = 0. \tag{61}$$

For any $k \in \{1,\ldots,M_R\}$, the arbitrariness of η allows us to set $\eta^j = 0$ for $j \neq k$. Then the above equality becomes

$$\int_0^T \left(\dot{\eta}^k \ln\left(\frac{\dot{\psi}^k}{c_k(\psi)}\right) + \sum_j \left(1 - \frac{\dot{\psi}^j}{c_j(\psi)}\right) \nabla_k c_j(\psi)\eta^k \right) \mathrm{d}t = 0. \tag{62}$$

Integrating by parts and the arbitrariness of η will give

$$-\frac{\ddot{\psi}^k}{\dot{\psi}^k} + \frac{\sum_\ell \nabla_\ell c_k(\psi)\dot{\psi}^\ell}{c_k(\psi)} + \sum_j \left(1 - \frac{\dot{\psi}^j}{c_j(\psi)}\right) \nabla_k c_j(\psi) = 0, \tag{63}$$

or equivalently

$$\ddot{\psi}^k = \frac{\dot{\psi}^k \sum_\ell \nabla_\ell c_k(\psi)\dot{\psi}^\ell}{c_k(\psi)} + \dot{\psi}^k \sum_j \left(1 - \frac{\dot{\psi}^j{}_t}{c_j(\psi)}\right) \nabla_k c_j(\psi) = \frac{\dot{\psi}^k \sum_\ell \nabla_\ell c_k(\psi)\dot{\psi}^\ell}{c_k(\psi)} + \dot{\psi}^k \sum_\ell \left(1 - \frac{\dot{\psi}^\ell{}_t}{c_\ell(\psi)}\right) \nabla_k c_\ell(\psi), \tag{64}$$

which gives (27).

References

[1] H.H. MacAdams, A. Arkin, It's a noisy business! Genetic regulation at the nanomolar scale, Trends Genet. 15 (1999) 65–69.
[2] N. Fedoroff, W. Fontana, Small numbers of big molecules, Science 297 (2002) 1129–1131.
[3] P. Hänggi, P. Talkner, M. Borkovec, Reaction-rate theory: fifty years after Kramers, Rev. Mod. Phys. 62 (1990) 251–341.
[4] A.B. Bortz, M.H. Kalos, J.L. Lebowitz, A new algorithm for Monte Carlo simulation of Ising spin systems, J. Comp. Phys. 17 (1975) 10–18.
[5] D. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, J. Comp. Phys. 22 (1976) 403–434.
[6] D. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. 81 (1977) 2340–2361.
[7] B. Alberts et al, Essential Cell Biology, second ed., Garland Science, New York, 2004.
[8] S. Wang, Multistability and multicellularity: cell fates as high-dimensional attractors of Gene Regulatory Networks, in: A. Kriete, R. Eils (Eds.), Computational System Biology, Elsevier Academic Press, New York, 2006, pp. 293–326.
[9] A. Arkin, J. Ross, H. MacAdams, Stochastic kinetic analysis of the developmental pathway bifurcation in phase λ-infected *Escherichia coli* cells, Genetics 149 (1998) 1633–1648.
[10] T.S. Gardner, C.R. Cantor, J.J. Collins, Construction of a genetic toggle switch in *Escherichia coli*, Nature 403 (2000) 339–342.
[11] A. Novick, M. Winer, Enzyme induction as an all-or-none phenomenon, PNAS 43 (1957) 553–567.
[12] J.D. Chung, G. Stephanopoulos, On physiological multiplicity and population of biological systems, Chem. Eng. Sci. 51 (1996) 1509–1521.
[13] M. Dykman, E. Mori, J. Ross, P. Hunt, Large fluctuations and optimal paths in chemical kinetics, J. Chem. Phys. 100 (1994) 5735–5750.
[14] D. Roma, R. O'Flanagan, A. Ruckenstein, A. Sengupta, R. Mukhopadhay, Optimal path to epigenetic switching, Phys. Rev. E 71 (2005) 011901.
[15] E. Aurell, K. Sneppen, Epigenetics as a first exit problem, Phy. Rev. Lett. 88 (2002) 048101.
[16] X. Zhu, L. Yin, L. Hood, P. Ao, Calculating biological behaviors of epigenetic states in the phage λ life cycle, Funct. Integr. Cemonics 4 (2004) 188–195.
[17] D. Liu, Optimal transition paths of stochastic chemical kinetic systems, J. Chem. Phys. 124 (2006) 164104.
[18] A.D. Wentzell, Limit Theorems on Large Deviations for Markov Stochastic Processes, Kluwer Academic Publishers., Boston, London, 1990.
[19] A. Shwartz, A. Weiss, Large Deviations for Performance Analysis: Queues, Communications, and Computing, Chapman and Hall, London, New York, 1995.
[20] A. Shwartz, A. Weiss, Large deviations with diminishing rates, Math. Operat. Res. 30 (2005) 281–310.
[21] W. E, W. Ren, E. Vanden-Eijnden, Minimum action method for the study of rare events, Comm. Pure Appl. Math. 57 (2004) 637–656.
[22] M. Heymann, E. Vanden-Eijnden, The geometric minimum action method: a lest action principle on the space of curves, Comm. Pure Appl. Math. LXI (2008) 1052–1117.
[23] T. Kurtz, Strong approximation theorems for density dependent Markov chains, Stoc. Proc. Appl. 6 (1978) 223–240.
[24] N.G. Van Kampen, Stochastic Processes in Physics and Chemistry, Elsevier North-Holland Publishing, New York, 1981.
[25] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems, second ed., Springer, New York, 1998.
[26] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, New York, 1999.
[27] W. E, D. Liu, E. Vanden-Eijnden, Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates, J. Chem. Phys. 123 (2005) 194107.
[28] W. E, D. Liu, E. Vanden-Eijnden, Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales, J. Comput. Phys. 221 (2007) 158–180.
[29] E.L. Haseltine, J.B. Rawlings, Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics, J. Chem. Phys. 117 (2002) 6959–6969.
[30] J. Puchalka, A.M. Kierzek, Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks, Biophys. J. 86 (2004) 1357–1372.
[31] H. Salis, Y. Kaznessis, Accurate hybrid stochastic simulation of a system of coupled chemical or biological reactions, J. Chem. Phys. 122 (2005) 054103.
[32] Z. Zhou, W. Ren, W. E, Adaptive minimum action method for the study of rare events, J. Chem. Phys. 128 (2008) 104111.