Two-scale large deviations for chemical reaction kinetics through second quantization path integral

Tiejun Li[1] and Feng Lin

Laboratory of Mathematics and Applied Mathematics and School of Mathematical Sciences, Peking University, Beijing 100871, People's Republic of China

E-mail: tieli@pku.edu.cn and math_linfeng@pku.edu.cn

CrossMark

Abstract

Motivated by the study of rare events for a typical genetic switching model in systems biology, in this paper we aim to establish the general two-scale large deviations for chemical reaction systems. We build a formal approach to explicitly obtain the large deviation rate functionals for the considered two-scale processes based upon the second quantization path integral technique. We get three important types of large deviation results when the underlying two timescales are in three different regimes. This is realized by singular perturbation analysis to the rate functionals obtained by the path integral. We find that the three regimes possess the same deterministic mean-field limit but completely different chemical Langevin approximations. The obtained results are natural extensions of the classical large volume limit for chemical reactions. We also discuss its implication on the single-molecule Michaelis–Menten kinetics. Our framework and results can be applied to understand general multi-scale systems including diffusion processes.

Keywords: two-scale large deviations, second quantization path integral, mean-field limit, chemical Langevin approximation, singular perturbation, Michaelis–Menten

(Some figures may appear in colour only in the online journal)

1. Introduction

In recent years there has been a growing interest in studying the rare transitions for fast–slow stochastic dynamics in biology [1, 4, 10, 15, 17, 18, 22, 31, 32]. In computational

---

[1] Author to whom any correspondence should be addressed.
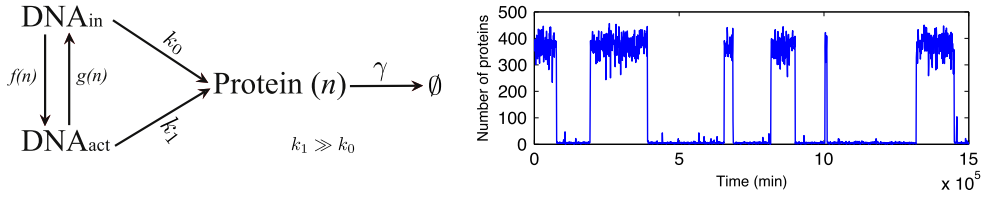
Figure 1. A typical fast–slow genetic switching model considered in systems biology. Left panel: schematics of the chemical reaction schemes, where the switching rates $f(n)$, $g(n)$ are usually large. Right panel: direct Monte Carlo simulations of the genetic switching model. The bi-stability is clearly observed from the time series of protein copy numbers.

neuroscience, the stochastic hybrid system is utilized to model the fast switching of ion channels, and the membrane voltage evolves according to dynamics which depend on the ion channel states. In systems biology, people are interested in the phenotypic switching of the cells modeled by the central dogma, which involves fast switching of DNA states between active and inactive states and the transcriptional and translational processes with different rates depending on the DNA states. In both cases, the transition rates and the most probable transition paths between different stable fixed points are issues being investigated in the literature. The main approaches include the Wentzel–Kramers–Brillouin (WKB) asymptotics and the path integral formulations. However, mathematically it falls in the field of large deviation theory (LDT) [9, 16, 26–30], and the rigorous results for these types of problems are very limited [15]. It is also meaningful to remark that there is a close connection between the LDT and the popular landscape theory for biological systems [17, 18, 33].

In this paper, we will continue our program to study the two-scale large deviations for chemical kinetic systems. To illustrate our points more concretely, let us consider a canonical genetic switching model [10, 32] in systems biology as shown in figure 1. Dynamics of this self-regulating genetic system can be described by the following chemical master equation

$$
\partial_t \boldsymbol{P}(n, t) = \begin{pmatrix} k_0 & 0 \\ 0 & k_1 \end{pmatrix} (\mathbb{E}_n^{-1} - 1) \boldsymbol{P}(n, t) + \gamma (\mathbb{E}_n^1 - 1)[n \boldsymbol{P}(n, t)]
$$
$$
+ \begin{pmatrix} -g(n) & f(n) \\ g(n) & -f(n) \end{pmatrix} \boldsymbol{P}(n, t), \tag{1}
$$

where the two-component vector $\boldsymbol{P}(n, t) = (P_0(n, t), P_1(n, t))^T$, and $P_j(n, t)$ is the probability distribution function that the system has n protein copy numbers in the DNA active ($j = 1$) or inactive state ($j = 0$). The raising or lowering operator $\mathbb{E}_n^k$ is defined through $\mathbb{E}_n^k h(n) = h(n + k)$ for any function h depending on n. k_0 and k_1 are protein synthesis rates, γ is the degradation rate constant, and $f(n)$, $g(n)$ are switching rates between two DNA states.

The biologically relevant parameter setup is $\gamma \sim O(1)$, and k_0/γ, k_1/γ are both large. We will not consider more detailed regimes concerning the magnitudes of k_1 and k_0, although one usually has $k_1 \gg k_0$ in realistic situations. This does not affect the main point in this paper. In this case, the average number of proteins at steady state is of the order k_1/γ. Now let us define the small parameter $\epsilon \approx \gamma/k_1$ or γ/k_0, thus the characteristic number of proteins is $n \sim O(\epsilon^{-1})$. In our fast–slow genetic switching model, we define the switching rates

$f(n)$, $g(n) \sim O(\epsilon^{-\alpha})$, and the realistic situations can be classified into the following three typical regimes:

(1) Case 1: $\alpha > 1$, i.e. the genetic switching process is much faster than the translation process;
(2) Case 2: $\alpha = 1$, i.e. the switching rates are comparable to the translation rates;
(3) Case 3: $0 < \alpha < 1$, i.e. the translation process is much faster than the genetic switching process.

In Case 2, the WKB asymptotics and the rigorous LDT results have been established for a similar model which takes into account the mRNA fluctuation [15]. The obtained LDT rate functional is utilized to find the most probable transition path and characterize the rate of transitions between the high and low expression states. Furthermore, the authors have shown that the Hamiltonian obtained from LDT is convex with respect to the momentum variable, which is one key point in designing robust numerical algorithms. In Case 3, the researchers typically take the continuum limit to the translation process at first since it is even faster than the switching process [10]. With this approach, one obtains a stochastic hybrid system which resembles a similar form to those for ion channels considered in computational neuroscience. So far, the WKB asymptotics and path integral formulations are both proposed for stochastic hybrid systems. Case 1 is also studied with WKB asymptotics applied to the averaged system with respect to the fast switching process.

From the authors' point of view, the approaches employed in [10] are like taking a repeated limit to the switching and translation processes according to their relative magnitudes. More concretely, when DNA switching is much faster than the protein synthesis, the equilibrium pre-averaging of the switching process is taken in [10] at first and one gets a pure translation process with effective translation rates; however, when the protein synthesis is much faster than DNA switching, the large volume limit is taken to the translation process at first and one gets a stochastic hybrid system [10]. Similar ideas and techniques are adopted in [20] and [23] as well, which discussed different timescale issues for the gene expression model. With this understanding, it will be interesting to investigate the double limit of the original process instead of taking an average with respect to the faster process at first. Mathematically it is also desirable to establish the large deviations for the original system with two timescales but different magnitudes. In fact, it is the main motivation of this paper. We will utilize the Doi–Peliti second quantization path integral formalism [7, 24, 32] to study the general two-scale large deviations for the genetic switching models. As we will see, although the second quantization path integral for the spin–boson-type model is formal, it is an effective approach to derive the large deviation results for chemical jump processes. Compared with the classical path integral formalism for diffusion processes, the second quantization path integral for chemical jump processes formulates the weight of each path in an extended space, which involves both coordinate and momentum variables. This means that the large deviation result can be given through a Hamiltonian with an explicit formula, which resolves the dilemma that the Lagrangian in the rate functional does not have a closed form. This is important for further theoretical and numerical studies. Mathematically, rigorously establishing the LDT obtained from the formal approach in this paper is in progress based on our previous analysis [15].

Let us briefly illustrate our general two-scale LDT results. We will show that the Lagrangian obtained from the second quantization path integral comprises two parts, which correspond to the switching and translation processes, respectively. However, what we are interested in is the LDT only for the concentration of proteins. The different magnitudes of the switching and translation

rates essentially lead to a singularly perturbed variational problem, which has different dominant terms and different scaling limits in the cases of $0 < \alpha < 1$ and $\alpha > 1$. When $\alpha = 1$, the Lagrangians from both parts contribute equally, and we get a result which combines the Donsker–Varadhan-type LDT [28] for the occupation measure of DNA states and the large volume-type LDT [26] for the small noise perturbation [15]. As the LDT gives the sharpest characterization of the considered two-scale chemical kinetic system, we can obtain the deterministic mean-field ordinary differential equations (ODEs) and the chemical Langevin approximation for the system based on the local analysis of the large deviation results [6]. This corresponds to the law of large numbers and the central limit theorem for the process. We found that the three cases possess the same mean-field ODEs. However, the chemical Langevin approximations for them are quite different. If $\alpha > 1$, only the fluctuation from the protein translation process survives. If $0 < \alpha < 1$, only the fluctuation from the genetic switching process survives. And if $\alpha = 1$, fluctuations from both the protein translation and genetic switching processes contribute. Similar results are also valid for single-molecule Michaelis–Menten kinetics [14] with slight modifications (see section 5). Our study extends the insights about the chemical kinetic systems in the classical large volume limit, and the methodology we introduced here can be applied to other multi-scale problems in many fields.

The rest of the paper is organized as follows. In section 2, we introduce the chemical master equation and apply the Doi–Peliti path integral formalism to the considered model. We then rescale the system with system size ϵ^{-1} and get the abstract LDT result based on the singular perturbation analysis in section 3. In section 4, we apply our abstract result to the two-state genetic switching model and present the mean-field limits and chemical Langevin approximations. In section 5, we apply our result to the well-known single-molecule Michaelis–Menten kinetics and mention the implications. Finally we make the conclusion and related discussions in section 6.

2. Transition probability in a path integral form

We start from a more general model rather than equation (1). Assume that the DNA switching could occur among N possible states ($N = 2$ for the model shown in figure 1) and the chemical master equation (CME) for the biological reaction network reads

$$
\begin{aligned}
\frac{\partial}{\partial t} \boldsymbol{P}(n, t) =\ & \boldsymbol{A}[\boldsymbol{P}(n-1, t) - \boldsymbol{P}(n, t)] \\
& + \gamma[(n+1)\boldsymbol{P}(n+1, t) - n\boldsymbol{P}(n, t)] \\
& + \boldsymbol{Q}^{\dagger}\boldsymbol{P}(n, t).
\end{aligned}
\tag{2}
$$

Here $\boldsymbol{P}(n, t) = (P_1(n, t), \ldots, P_N(n, t))^T$, $P_j(n, t)$ is the probability distribution function that the system has n protein copy numbers and the switch is in state j at time t. $\boldsymbol{A}$ is a diagonal matrix with diagonal entry k_j as the protein synthesis rate in state j. γ is the protein degradation rate and $\boldsymbol{Q} = (q_{jk}(n))_{j,k=1}^N$ is the transition rate matrix among different DNA states. Thus $q_{jk}(n) \geqslant 0$ for any $j \neq k$ and $\sum_{k=1}^N q_{jk}(n) = 0$. We assume that the switching process is ergodic.

Now we follow the Doi–Peliti approach to establish the path integral formalism of the CME (2) [7, 24, 31, 32]. We define the creation, annihilation operators $a^{\dagger}$, a and the state function $|\psi\rangle$ as

$$
a^{\dagger}|n\rangle = |n+1\rangle, \quad a|n\rangle = n|n-1\rangle \text{ and } |\psi\rangle = \sum_{n=0}^{\infty}\boldsymbol{P}(n, t)|n\rangle.
$$

Then the CME (2) can be written in a second-quantized form

$$\partial_t \; |\psi\rangle = \Omega \; |\psi\rangle, \tag{3}$$

where the operator

$$\Omega = A(a^\dagger - 1) + \gamma(a - a^\dagger a) + \hat{Q}^\dagger \tag{4}$$

and $\hat{Q}$ is obtained from Q by replacing the transition rates $q_{ij}(n)$ with operators $q_{ij}(a^\dagger a)$.

From equation (3), the transition probability $P(n_f, \tau|n_i, 0)$ of finding a product copy number n_f at time $t = \tau$ starting from n_i at $t = 0$ has the form

$$P(n_f, \tau|n_i, 0) = \langle n_f| \exp(\Omega\tau)|n_i\rangle$$
$$= \lim_{\Delta t \to 0} \langle n_f| \exp(\Omega\Delta t)^{\tau/\Delta t}|n_i\rangle, \tag{5}$$

where $|n_i\rangle = (|n_i\rangle, |n_i\rangle, \cdots, |n_i\rangle)^T$ is an N-dimensional column vector and $\langle n_f| = (\langle n_f|, \cdots, \langle n_f|)$ is an N-dimensional row vector. Following Zhang *et al* [32], we utilize the coherent state representation and a resolution of identity [31, 32] as

$$I_b \otimes I_s = \int_0^\infty dn \int_{-\pi}^\pi \frac{d\beta}{2\pi}|z\rangle\langle\tilde{z}|e^{-n}$$
$$\cdot \frac{1}{N^{N-2}} \prod_{k=1}^{N-1} \int_0^{(N/2)\pi} \sin\frac{2\theta_k}{N}d\theta_k \frac{1}{4\pi} \int_0^{4\pi} d\phi_k|\psi^R\rangle\langle\psi^L|, \tag{6}$$

where

$$|\psi^R\rangle = \begin{bmatrix} \dfrac{2}{N}\cos^2\dfrac{\theta_1}{N}e^{i\phi_1/2} \\ \vdots \\ \dfrac{2}{N}\cos^2\dfrac{\theta_N}{N}e^{i\phi_N/2} \end{bmatrix}, \quad \langle\psi^L| = (e^{-i\phi_1/2}, \cdots, e^{-i\phi_N/2})$$

with

$$|z\rangle = e^{a^\dagger z}\,|0\rangle, \; \langle\tilde{z}| = \langle 0|\,e^{a\tilde{z}}, \; z = ne^{-i\beta}, \; \tilde{z} = e^{i\beta},$$

and i is the imaginary unit. The variable n has the interpretation that it characterizes the mean protein number in the coherent states. We define $c_j = (2/N)\cos^2(\theta_j/N)$. c_j gives the occupation probability of DNA at state j from the probabilistic interpretation of quantum mechanics. They satisfy the normalization condition $\sum_{j=1}^N c_j = 1$. Correspondingly, the phase variable ϕ can be chosen to satisfy $\sum_{j=1}^N \phi_j = 0$ for convenience. These choices are consistent with the resolution of identity. As a consequence, we have only $N - 1$ independent unknowns $(c_1, ..., c_{N-1})$ in $c = (c_1, ..., c_N)$, which is equivalent to using $\theta = (\theta_1, ..., \theta_N)$, and $N - 1$ unknowns $(\phi_1, ..., \phi_{N-1})$ in $\phi = (\phi_1, ..., \phi_N)$.

Inserting (6) into (5), the transition probability density can be represented as a path integral form

$$P(n_f, \tau|n_i, 0) = \text{Const.} \times \int Dn D\beta Dc D\phi \exp\left(-\int_0^\tau L dt\right), \tag{7}$$

where the Lagrangian L is defined as

$$L = i\beta \frac{dn}{dt} + \sum_{j=1}^{N-1} i\varphi_j \frac{d(c_N - c_j)/2}{dt} - H(n, \boldsymbol{c}, i\beta, i\boldsymbol{\varphi}), \tag{8}$$

and the Hamiltonian

$$H(n, \boldsymbol{c}, i\beta, i\boldsymbol{\phi}) = H_1(n, \boldsymbol{c}, i\beta) + H_2(n, \boldsymbol{c}, i\boldsymbol{\phi}) \tag{9}$$

with H_1 for the translation process

$$H_1(n, \boldsymbol{c}, i\beta) = \sum_{j=1}^{N} k_j c_j [\exp(i\beta) - 1] + \gamma n [\exp(-i\beta) - 1] \tag{10}$$

and H_2 for the switching process

$$H_2(n, \boldsymbol{c}, i\boldsymbol{\phi}) = \sum_{m,j}^{N} c_m q_{mj}(n)(e^{(i\phi_m - i\phi_j)/2} - 1). \tag{11}$$

In equation (7), the outer four-fold integral is taken in the path space with respect to $n(t)$, $\beta(t)$, $\boldsymbol{c}(t)$ and $\boldsymbol{\phi}(t)$, which are full trajectories in $[0, \tau]$. The terms involving $|n_i\rangle$ and $\langle n_f|$ have been absorbed to the constant before the integral. The path integral formulation (7) makes the weight of each trajectory explicit. The form of Lagrangian (8) suggests the interpretation that the pairs $i\beta$ and n, $i\phi$ and c are conjugate variables.

To study the associated LDT, we must have a small parameter ϵ and a deterministic limit as $\epsilon \to 0$. This could be chosen as the inverse of typical system size $\epsilon = \gamma/k_1 = V^{-1}$. As stated in the introduction section, we assume

$$k_j \sim \frac{1}{\epsilon}, \quad q_{ij} \sim \frac{1}{\epsilon^\alpha}, \quad \alpha > 0 \tag{12}$$

and define

$$x = n\epsilon, \quad \tilde{k}_j = k_j \epsilon, \quad \tilde{q}_{jk}(x) = q_{jk}(n)\epsilon^\alpha. \tag{13}$$

With these definitions, equation (7) can be rewritten as

$$P(n_f, \tau | n_i, 0) = \text{Const.} \times \int Dx D\beta D\boldsymbol{c} D\boldsymbol{\varphi} \exp\left(-\frac{1}{\epsilon} \int_0^\tau \tilde{L}_1 dt - \frac{1}{\epsilon^\alpha} \int_0^\tau \tilde{L}_2 dt\right), \tag{14}$$

where the rescaled Lagrangian

$$\tilde{L}_1 = i\beta \frac{dx}{dt} - \tilde{H}_1(x, \boldsymbol{c}, i\beta), \tag{15}$$

$$\tilde{L}_2 = \epsilon^\alpha \sum_{j=1}^{N-1} i\varphi_j \frac{d(c_N - c_j)/2}{dt} - \tilde{H}_2(x, \boldsymbol{c}, i\boldsymbol{\varphi}), \tag{16}$$

and rescaled Hamiltonian

$$\tilde{H}_1(x, \boldsymbol{c}, i\beta) = \sum_{j=1}^{N} \tilde{k}_j c_j [\exp(i\beta) - 1] + \gamma x [\exp(-i\beta) - 1], \tag{17}$$

$$\tilde{H}_2(x, \boldsymbol{c}, i\boldsymbol{\phi}) = \sum_{m,j}^{N} c_m \tilde{q}_{mj}(x)(e^{(i\phi_m - i\phi_j)/2} - 1). \tag{18}$$

Using the method of steepest descent asymptotics, the integration over β and ϕ can be approximated by simply using the value of the integrand at the saddle point [2]. Thus, we get

$$P(n_f, \tau | n_i, 0) \propto \int DxDc \exp\left(-\frac{1}{\epsilon}\int_0^\tau L_1(x, \dot{x}, c)\mathrm{d}t - \frac{1}{\epsilon^\alpha}\int_0^\tau L_2(x, c)\mathrm{d}t\right), \tag{19}$$

where

$$L_1(x, \dot{x}, c) = \sup_p\{p\dot{x} - \tilde{H}_1(x, c, p)\}, \tag{20}$$

$$L_2(x, c) = \sup_\varphi\{-\tilde{H}_2(x, c, \varphi)\}. \tag{21}$$

Note that the term $\varphi \cdot \dot{c}$ does not appear in equation (21) because of the factor ϵ^α in the first term of equation (16). Formally, the functional appearing in the exponential in (19) is a competition between the rate functional $\int L_1 \mathrm{d}t$ which corresponds to the translation process, and the rate functional $\int L_2 \mathrm{d}t$ which corresponds to the switching process. It is interesting to observe that the Lagrangian L_1 corresponds to the large volume-type LDT rate function for the small noise perturbation [26] and L_2 corresponds to the Donsker–Varadhan-type LDT rate function for the occupation measure of DNA states [26, 28]. The second quantization path integral perfectly reveals the intrinsic structure of the considered two-scale chemical kinetic process.

3. Formulation of the LDT in a general setting

The transition probability (19) contains the LDT information about the variables x and c. However, in most cases, one is only interested in slow variables, i.e. the concentration of protein in our case, which is also the observable in experiments. In this sense, we must integrate over c-space. It turns out the final result depends on the value of α and we will have three typical regimes. In what follows, we will discuss different outcomes in different regimes separately.

(i). **Case 1:** $\alpha > 1$. The switching process is much faster than the translation process. In this case, we can rewrite equation (19) as

$$P(n_f, \tau | n_i, 0) \propto \int DxDc \exp\left\{-\frac{1}{\epsilon}\int_0^\tau \mathrm{d}t\left(L_1(x, \dot{x}, c) + \frac{1}{\epsilon^{\alpha-1}}L_2(x, c)\right)\right\}. \tag{22}$$

To integrate over c-space, we take the Laplace asymptotics for each t. The Lagrangian for x has the form

$$L_x(x, \dot{x}) = \inf_c\left\{L_1(x, \dot{x}, c) + \frac{1}{\epsilon^{\alpha-1}}L_2(x, c)\right\} \quad \text{as} \quad \epsilon \to 0+. \tag{23}$$

Since $L_2 \geqslant 0$ and $\epsilon^{1-\alpha} \gg 1$, the term $\epsilon^{1-\alpha}L_2(x, c)$ dominates. From the assumption that the switching process is ergodic, for a given x, $L_2(x, c)$ achieves its minimum $L_2(x, c) = 0$ at a single point $c = c_0(x)$, i.e. the steady state distribution given the concentration x [13, 26]. Thus we get

$$L_x(x, \dot{x}) = L_1(x, \dot{x}, c_0(x)) \tag{24}$$

and

$$P(n_f, \tau | n_i, 0) \propto \int Dx \exp\left\{-\frac{1}{\epsilon}\int_0^\tau \mathrm{d}t L_x(x, \dot{x})\right\}. \tag{25}$$

Although we still leave the factor $\epsilon^{1-\alpha}$ in the Laplace asymptotics (23) in our manipulation and then take the singular perturbation analysis, it is not difficult to establish the final result in

a rigorous way. This result tells us that when $\alpha > 1$, the LDT for the slow variable x is only determined by the effective synthesis rate $\sum_{j=1}^{N} \tilde{k}_j (c_0(x))_j$ and degradation rate γ.

(ii). **Case 3:** $0 < \alpha < 1$. The translation process is much faster than the switching process.

In this case, we rewrite equation (19) as

$$P(n_f, \tau | n_i, 0) \propto \int Dx Dc \, \exp\left\{ -\frac{1}{\epsilon^\alpha} \int dt \left(\frac{1}{\epsilon^{1-\alpha}} L_1(x, \dot{x}, c) + L_2(x, c) \right) \right\}. \tag{26}$$

Taking Laplace asymptotics with respect to the c-integral, we get

$$L_x(x, \dot{x}) = \inf_c \left(\frac{1}{\epsilon^{1-\alpha}} L_1(x, \dot{x}, c) + L_2(x, c) \right). \tag{27}$$

Since $L_1 \geqslant 0$ and $\epsilon^{\alpha-1} \gg 1$, the term $\epsilon^{\alpha-1} L_1(x, \dot{x}, c)$ dominates. We can perform a similar approach to derive L_x as in the previous case. In general, we assume that $L_1(x, \dot{x}, c)$ achieves its minimum $L_1(x, \dot{x}, c) = 0$ at $c = c_x$ for a given x. In our case, c_x satisfies the mean-field ODE by the large volume limit:

$$\dot{x} = \sum_{j=1}^{N} k_j (c_x)_j - \gamma x. \tag{28}$$

By (27) and (28), we have

$$L_x(x, \dot{x}) = \inf_{\{c_x : \dot{x} = \sum_{j=1}^{N} k_j (c_x)_j - \gamma x\}} L_2(x, c_x) \tag{29}$$

and

$$P(n_f, \tau | n_i, 0) \propto \int Dx \, \exp\left\{ -\frac{1}{\epsilon^\alpha} \int_0^\tau dt L_x(x, \dot{x}) \right\}. \tag{30}$$

We want to remark here that from (30) we will expect to get the LDT of the type

$$\lim_{\epsilon \to 0+} \epsilon^\alpha \ln \text{Prob}(X. \in \mathcal{B}) = -\inf_{x \in \mathcal{B}} \int_0^\tau L_x(x, \dot{x}) dt \tag{31}$$

where $\mathcal{B}$ is a Borel set in $D[0, \tau]$ space (functions on $[0, \tau]$ are right continuous with left limits) and X is the sample path of the original jump process. The scaling ϵ^α in (31) is essential to reveal the nontrivial behavior of x. Other choices of the exponent do not give the correct limit which we are interested in for x.

In the considered case $0 < \alpha < 1$, the protein synthesis is much faster than the genetic switching. With this condition, if we neglect the copy-number fluctuation of the protein, we get a reduced stochastic hybrid system:

$$\dot{x} = \sum_{j=1}^{N} k_j I_{\{\xi(t)=j\}} - \gamma x, \tag{32}$$

where $I_{\{\xi(t)=j\}}$ is an indicator function and $\xi(t)$ represents the DNA occupation state. In [3] and [8], the authors established the LDTs for variable x as $\epsilon \to 0+$ for system (32), which is the same as what we derived in equation (29). But we should emphasize that this coincidence is not obvious *a priori*; our result supports the validity of the procedure by taking the repeated limit for two-scale processes in some sense.

(iii). **Case 2:** $\alpha = 1$. The switching rates are comparable to the translation rates.

When $\alpha = 1$, we have

$$P(n_f, \tau | n_i, 0) = \int Dx Dc \exp\left\{-\frac{1}{\epsilon}\left(\int dt L_1(x, \dot{x}, c) + \int dt L_2(x, c)\right)\right\}. \tag{33}$$

In this case, we have the LDT Lagrangian for variable x:

$$L_x(x, \dot{x}) = \inf_c \{L_1(x, \dot{x}, c) + L_2(x, c)\}. \tag{34}$$

In most cases there is no closed form for L_1, and thus we do not expect to get the closed form of L_x accordingly. This hinders the applicability of the obtained theory. It is more convenient to study the conjugate Hamiltonian of L_x:

$$\begin{aligned}
H_x(x, p) &= \sup_\beta \{p\beta - L_x(x, \beta)\} \\
&= \sup_\beta \{p\beta - \inf_c \{L_1(x, \beta, c) + L_2(x, c)\}\} \\
&= \sup_\beta \sup_c \{p\beta - L_1(x, \beta, c) - L_2(x, c)\} \\
&= \sup_c \sup_\beta \{p\beta - L_1(x, \beta, c) - L_2(x, c)\} \\
&= \sup_c \{\tilde{H}_1(x, p, c) - L_2(x, c)\}.
\end{aligned} \tag{35}$$

As we will show, the dual Hamiltonian may have an explicit expression and it is convex with respect to the momentum variable p. This property makes it competitive for the numerical algorithms for solving the static Hamilton–Jacobi equation through the geometric minimum action method [11, 17].

4. Application to the two-state model

Using the two-state model (1) as an example, we will give the detailed LDT results for different α, and show the mean-field ODE and the chemical Langevin approximation for variable x. Moreover, we will solve the static Hamilton–Jacobi equation for the quasi-potential $\Phi(x)$ in different situations. First, we take the same rescaling (13) for the variables and parameters. We again consider three different cases: (i) $\alpha > 1$, (ii) $0 < \alpha < 1$ and (iii) $\alpha = 1$.

(i). **Case 1:** $\alpha > 1$. In this case, the ergodic limit of DNA occupation probability is

$$c_0(x) = \left(\frac{\tilde{f}(x)}{\tilde{f}(x) + \tilde{g}(x)}, \frac{\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)}\right)$$

for given x. By equation (23), we have

$$L_x(x, \dot{x}) = \sup_p \left\{p\dot{x} - \left[\frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)}(\exp(p) - 1) + \gamma x(\exp(-p) - 1)\right]\right\}, \tag{36}$$

and the dual Hamiltonian

$$H_x(x, p) = \frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)}[\exp(p) - 1] + \gamma x[\exp(-p) - 1]. \tag{37}$$

From the result

$$\frac{\partial H_x}{\partial p}\big|_{p=0} = \frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0 g(x)}{\tilde{f}(x) + \tilde{g}(x)} - \gamma x, \tag{38}$$

we get the mean-field ODE

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H_x}{\partial p}\big|_{p=0} = \frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} - \gamma x. \tag{39}$$

Furthermore, the fact that

$$\frac{\partial^2 H_x}{\partial p^2}\big|_{p=0} = \frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} + \gamma x \tag{40}$$

shows the following chemical Langevin approximation holds

$$\mathrm{d}x = \left(\frac{\tilde{k}_1 f(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} - \gamma x\right)\mathrm{d}t + \sqrt{\epsilon}\left(\sqrt{\frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)}}\,\mathrm{d}w_1 - \sqrt{\gamma x}\,\mathrm{d}w_2\right), \tag{41}$$

where w_1 and w_2 are independent standard Brownian motions.

From classical variational analysis [25], it can be shown that the quasi-potential defined through

$$\Phi(x; x_0) = \inf_{\tau \geqslant 0} \inf_{x(0)=x_0, x(\tau)=x} \int_0^{\tau} L_x(x, \dot{x})\mathrm{d}t$$

in our case satisfies the static Hamilton–Jacobi equation $H(x, \partial_x \Phi) = 0$, where x_0 is a stable fixed point. Based on (37), we have by some algebra

$$\partial_x \Phi = -\log\frac{(\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x))/(\tilde{f}(x) + \tilde{g}(x))}{\gamma x}. \tag{42}$$

This result is consistent with the quasi-potential derived in [10], where the authors neglect the fluctuation of genetic switching and get the result by the WKB ansatz. But of course, there is no hope of obtaining the explicit formula of Φ when the dimension of x is bigger than 1.

(ii). **Case 3:** $0 < \alpha < 1$.

In this case, we have the Lagrangian

$$\begin{aligned}
L_x(x, \dot{x}) &= L_2(x, c_x) \\
&= \sup_{\varphi}\{-[c_1\tilde{g}(x)(e^{\phi_1-\phi_2} - 1) + c_2\tilde{f}(x)(e^{\phi_2-\phi_1} - 1)]\} \\
&= (\sqrt{\tilde{f}(x)c_2} - \sqrt{\tilde{g}(x)c_1})^2,
\end{aligned} \tag{43}$$

where $c_x = (c_1, c_2)$ and $c_1 = (\dot{x} - \tilde{k}_0 + \gamma x)/(\tilde{k}_1 - \tilde{k}_0)$, $c_2 = 1 - c_1$ by the condition $L_1(x, \dot{x}, c_x) = 0$. With the Legendre–Fenchel transform defined by $H_x(x, p) = \sup_{\beta}(p\beta - L_x(x, \beta))$, we get the dual Hamiltonian:

$$H_x(x, p) = p\beta_0 - L_x(x, \beta_0), \tag{44}$$

where $\beta_0 = \tilde{k}_1 s_1 + \tilde{k}_2(1 - s_1) - \gamma x$ and

$$s_1 = \frac{1}{2} + \frac{s_2}{2\sqrt{s_2^2 + 4}}, \quad s_2 = \frac{p(\tilde{k}_1 - \tilde{k}_0) + \tilde{f}(x) - \tilde{g}(x)}{\sqrt{\tilde{f}(x)\tilde{g}(x)}}.$$

Again, we can obtain the deterministic mean-field ODE as

$$\frac{dx}{dt} = \frac{\partial H_x}{\partial p}\Big|_{p=0} = \frac{\tilde{k}_1 \tilde{f}(x) + \tilde{k}_0 \tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} - \gamma x. \tag{45}$$

Similarly, we get

$$\frac{\partial^2 H_x}{\partial p^2}\Big|_{p=0} = \frac{2\tilde{f}(x)\tilde{g}(x)}{(\tilde{f}(x) + \tilde{g}(x))^3}(\tilde{k}_1 - \tilde{k}_0)^2 \tag{46}$$

and thus the chemical Langevin approximation

$$dx = \left(\frac{\tilde{k}_1 \tilde{f}(x) + \tilde{k}_0 \tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} - \gamma x\right)dt + \sqrt{\epsilon^\alpha}\sqrt{\frac{2\tilde{f}(x)\tilde{g}(x)}{(\tilde{f}(x) + \tilde{g}(x))^3}(\tilde{k}_1 - \tilde{k}_0)^2}\, dw. \tag{47}$$

We note here that the fluctuation term has the strength $\sqrt{\epsilon^\alpha}$ since it originates from the fast genetic switching process, and the term $\sqrt{\gamma x}\, dw_2$ disappears because it is in order $\sqrt{\epsilon}$. These are in sharp contrast with the result in (41) and Case 2 below, which has the $O(\sqrt{\epsilon})$ fluctuation.

By equation (44), solving the Hamilton–Jacobi equation $H(x, \partial_x \Phi) = 0$, we have

$$\partial_x \Phi = \frac{\tilde{f}}{\tilde{k}_0 - \gamma x} + \frac{\tilde{g}}{\tilde{k}_1 - \gamma x}. \tag{48}$$

This is consistent with the result in [10], although we have a totally different form of the Hamiltonian H.

(iii). **Case 2:** $\alpha = 1$.

In this case, the genetic switching rate is comparable to the protein synthesis rate. The rigorous LDT result has been obtained in [15]. Now we formally establish the LDT again through the second quantization path integral approach.

By (34) and (35), we have the dual Hamilton:

$$\begin{aligned}
H_x(x, p) &= \sup_{c} \{\tilde{H}_1(x, \boldsymbol{c}, p) - L_2(x, \boldsymbol{c})\} \\
&= \sup_{c} \{(\tilde{k}_1 c_1 + \tilde{k}_0 c_2)[\exp(p) - 1] + \gamma x[\exp(-p) - 1] \\
&\quad - (\sqrt{\tilde{f}(x)c_2} - \sqrt{\tilde{g}(x)c_1})^2\}.
\end{aligned} \tag{49}$$

Since $c_1 + c_2 = 1$ and $c_j \geqslant 0$, we can obtain the explicit expression of $H_x(x, p)$:

$$\begin{aligned}
H_x(x, p) &= (\tilde{k}_1 s + \tilde{k}_0(1 - s))[\exp(p) - 1] + \gamma x[\exp(-p) - 1] \\
&\quad - (\sqrt{\tilde{f}(1 - s)} - \sqrt{\tilde{g}s})^2,
\end{aligned} \tag{50}$$

where

$$s = \frac{1}{2} + \frac{s_1}{2\sqrt{s_1^2 + 4}}, \qquad s_1 = \frac{(\tilde{k}_1 - \tilde{k}_0)(e^p - 1) + \tilde{f}(x) - \tilde{g}(x)}{\sqrt{\tilde{f}(x)\tilde{g}(x)}}.$$

As before, we get the mean-field ODE

$$\frac{dx}{dt} = \frac{\partial H_x}{\partial p}\Big|_{p=0} = \frac{\tilde{k}_1 \tilde{f}(x) + \tilde{k}_0 \tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} - \gamma x. \tag{51}$$

The second-order expansion to p

$$\frac{\partial^2 H_x}{\partial p^2}\Big|_{p=0} = \frac{2\tilde{f}(x)\tilde{g}(x)}{(\tilde{f}(x)+\tilde{g}(x))^3}(\tilde{k}_1 - \tilde{k}_0)^2 + \frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} + \gamma x \tag{52}$$

yields the following chemical Langevin approximation

$$dx = \left(\frac{\tilde{k}_1\tilde{f} + \tilde{k}_0\tilde{g}}{\tilde{f} + \tilde{g}} - \gamma x\right)dt + \sqrt{\epsilon}\left(\sqrt{\frac{\tilde{k}_1\tilde{f} + \tilde{k}_0\tilde{g}}{\tilde{f} + \tilde{g}} + \frac{2\tilde{f}\tilde{g}}{(\tilde{f} + \tilde{g})^3}(\tilde{k}_1 - \tilde{k}_0)^2}\,dw_1 - \sqrt{\gamma x}\,dw_2\right), \tag{53}$$

where $\tilde{f}$, $\tilde{g}$ are abbreviations of $\tilde{f}(x)$ and $\tilde{g}(x)$, and w_1, w_2 are independent standard Brownian motions.

It is worth discussing the relationship between the Hamiltonian (50) and that obtained by WKB asymptotics. To get a Hamiltonian via WKB asymptotics, we follow the procedures in [21] and sketch its outline. We assume that the stationary solution of (2) has the form

$$P_i^s(x) \sim r_i(x)k(x)\exp\left[-\frac{1}{\epsilon}\Phi(x)\right], \quad i = 1, 2. \tag{54}$$

Substituting (54) into (2) and collecting leading-order terms, we get $M(x, p) \cdot (r_1(x), r_2(x))^T = 0$, where

$$M(x, p) = \begin{pmatrix} \tilde{k}_1(e^p - 1) + \gamma x(e^{-p} - 1) - \tilde{g} & \tilde{f} \\ \tilde{g} & \tilde{k}_0(e^p - 1) + \gamma x(e^{-p} - 1) - \tilde{f} \end{pmatrix}.$$

Now there is a subtlety to get the correct Hamilton in LDT. If one assumes that $H(x, p)$ is defined as the largest eigenvalue of $M(x, p)$, it can be shown that it is equivalent to (50). However, if one assumes that it is defined as the determinant of $M(x, p)$ [1, 21], then H is not convex in the momentum variable p and the equivalence is lost. This resembles the issue of the choice of Hamiltonians for the parametrized curve problem in classical mechanics (see p. 40 in [5]).

Let us make some comments on the obtained mean-field limit and Langevin approximations. Recall that there are two parts of noise in the original dynamics: one is from the translation process and the other from the DNA switching process. Our result shows that different diffusion approximations arise according to the magnitudes of residual noise in different reaction channels. If $\alpha > 1$, the dominant part of noise is from the translation, so only the fluctuation from the translation process survives. If $0 < \alpha < 1$, the dominant part is from the switching process, so only the fluctuation from the DNA switching process survives. And if $\alpha = 1$, both fluctuations from the protein translation and genetic switching contribute. A similar situation occurs in the LDT analysis, where the singular perturbation is performed for variational minimizations. The obtained results show the validity of the procedure by taking the limit for the faster process first and then performing the corresponding analysis for the slower process. Although we only consider the two-state models, the essential structure and results hold for general cases. It is a natural extension of the classical large volume limit for chemical reaction processes. We summarize our discussions for the three regimes in table 1.

Table 1. Comparison of LDTs, mean-field limits and Langevin approximations for three regimes

	LDT Hamiltonian	Deterministic Drift	Noise in Langevin Approximation
Case 1: ($\alpha > 1$)	$H_x(x, p)$ in (37)		$\sqrt{\epsilon}\left(\sqrt{\frac{\tilde{k}_1\tilde{f} + \tilde{k}_0\tilde{g}}{\tilde{f} + \tilde{g}}}\,\mathrm{d}w_1 - \sqrt{\gamma x}\,\mathrm{d}w_2\right)$
Case 2: ($\alpha = 1$)	$H_x(x, p)$ in (50)	$\frac{\tilde{k}_1\tilde{f}(x) + \tilde{k}_0\tilde{g}(x)}{\tilde{f}(x) + \tilde{g}(x)} - \gamma x$	$\sqrt{\epsilon}\left(\sqrt{\frac{\tilde{k}_1\tilde{f} + \tilde{k}_0\tilde{g}}{\tilde{f} + \tilde{g}} + \frac{2\tilde{f}\tilde{g}}{(\tilde{f} + \tilde{g})^3}(\tilde{k}_1 - \tilde{k}_0)^2}\,\mathrm{d}w_1 - \sqrt{\gamma x}\,\mathrm{d}w_2\right)$
Case 3: ($\alpha < 1$)	$H_x(x, p)$ in (44)		$\sqrt{\epsilon^\alpha}\sqrt{\frac{2\tilde{f}\tilde{g}}{(\tilde{f} + \tilde{g})^3}(\tilde{k}_1 - \tilde{k}_0)^2}\,\mathrm{d}w_1$

5. Application to single-molecule Michaelis–Menten kinetics

Our approach and observation have interesting implications on the single-molecule Michaelis–Menten system [14], in which a substrate S binds reversibly with an enzyme E to form an enzyme-substrate complex ES that decomposes to form a product P. The reaction schemes can be schematically shown as

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\longrightarrow} E + P, P \overset{\gamma}{\longrightarrow} \varnothing. \tag{55}$$
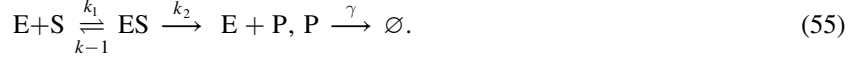
In the case of the single-molecule enzyme setup, the reaction system (55) falls within the framework considered in this paper. As in [14], we assume that the substrate is abundant enough and there is essentially no depletion of substrate by a single-enzyme molecule. That is, we assume the concentration of substrate is a constant, which will be denoted as $[S]$. It is well known that the rate of product formulation v has the following form in the quasi-steady state approximation [12, 19]

$$v = \frac{k_2[S]}{[S] + k_M}, \tag{56}$$

where $k_M = (k_{-1} + k_2)/k_1$. In [14], the statistics of enzymatic turn-over time and dynamical disorder are considered. Here we are interested in deriving the Langevin approximations of the Michaelis–Menten system in different regimes.

In [14], k_{-1} ranges from $0\ s^{-1}$ to $2000\ s^{-1}$, k_1 is usually taken as $10^7\ M^{-1}s^{-1}$, $k_2 = 250\ s^{-1}$, and $[S]$ ranges from the order $0.001mM$ to $0.1mM$, where $1\ M = 1molL$. Some specific choices of these parameters include

(1) Case 1: $k_{-1} = 2000\ s^{-1}$, $k_1[S] = 10^7\ M^{-1}s^{-1} \times 0.30mM = 3000\ s^{-1}$, $k_2 = 250\ s^{-1}$;
(2) Case 2: $k_{-1} = 200\ s^{-1}$, $k_1[S] = 10^7\ M^{-1}s^{-1} \times 0.02mM = 200\ s^{-1}$, $k_2 = 250\ s^{-1}$;
(3) Case 3: $k_{-1} = 50\ s^{-1}$, $k_1[S] = 10^7\ M^{-1}s^{-1} \times 0.005mM = 50\ s^{-1}$, $k_2 = 250\ s^{-1}$.

The degradation rate constant γ is not essential and we assume it is $O(1\ s^{-1})$. The above choices underlie the rationale to study different regimes in previous sections since we can make the assumption

$$k_2 \sim \frac{1}{\epsilon}, \quad k_1[S], k_{-1} \sim \frac{1}{\epsilon^{\alpha}}, \quad \alpha > 0 \tag{57}$$

if we define $\epsilon = 1/250$.

The chemical master equation of the system (55) can be written as

$$\partial_t \boldsymbol{P}(n, t) = \begin{pmatrix} 0 & k_2 \\ 0 & 0 \end{pmatrix}(\mathbb{E}_n^{-1} - 1)\boldsymbol{P}(n, t) + \gamma(\mathbb{E}_n^{1} - 1)[n\boldsymbol{P}(n, t)]$$
$$+ \begin{pmatrix} -k_1[S] & k_{-1} + k_2 \\ k_1[S] & -(k_{-1} + k_2) \end{pmatrix}\boldsymbol{P}(n, t). \tag{58}$$

Denote c_1 the occupation probability of the free enzyme molecule state E and c_2 the probability of the complex state ES. Using a similar approach to deriving (7), the transition probability can be obtained with Lagrangian L

$$L = i\beta\frac{\mathrm{d}n}{\mathrm{d}t} + i\phi_1\frac{\mathrm{d}(-c_1)}{\mathrm{d}t} - H(n, \boldsymbol{c}, i\beta, i\boldsymbol{\phi}), \tag{59}$$

where the Hamiltonian

$$H(n, c, i\beta, i\phi) = k_2 c_2 [\exp(i\beta) - 1] + \gamma n [\exp(-i\beta) - 1]$$
$$+ c_2(k_{-1} + k_2 e^{i\beta})(e^{i\phi_2} - 1) + c_1 k_1 [S](e^{-i\phi_2} - 1). \tag{60}$$

According to (57), we make the rescaling

$$x = n\epsilon \text{ (or } x_\alpha = n\epsilon^\alpha), \quad \tilde{k}_2 = k_2 \epsilon, \quad \tilde{k}_1 = k_1 [S] \epsilon^\alpha, \quad \tilde{k}_{-1} = k_{-1} \epsilon^\alpha. \tag{61}$$

Then the Cases 1, 2 and 3 correspond to $\alpha > 1$, $\alpha = 1$ and $0 < \alpha < 1$, respectively. Next, let us study the three cases separately. The order of discussion will be from easy to difficult, which may be slightly different from previous sections.

(i). **Case 2:** $\alpha = 1$.

With the steepest descent asymptotics as in (19), we have

$$P(n_f, \tau | n_i, 0) \propto \int Dx Dc \exp\left(-\frac{1}{\epsilon} \int_0^\tau L(x, \dot{x}, c) dt\right). \tag{62}$$

The Lagrangian L has the form

$$L(x, \dot{x}, c) = \sup_p \sup_\varphi \{p\dot{x} - \tilde{H}(x, c, p, \varphi)\}$$
$$= \sup_p \{p\dot{x} - H(x, c, p)\}, \tag{63}$$

where $H(x, c, p) = \inf_\varphi \{\tilde{H}(x, c, p, \varphi)\}$ and

$$\tilde{H}(x, c, p, \varphi) = \tilde{k}_2 c_2 (e^p - 1) + \gamma x (e^{-p} - 1) + c_2(\tilde{k}_{-1} + \tilde{k}_2 e^p)(e^\varphi - 1)$$
$$+ c_1 \tilde{k}_1 (e^{-\varphi} - 1). \tag{64}$$

In this case, we have the LDT Lagrangian for variable x by applying Laplace asymptotics

$$L_x(x, \dot{x}) = \inf_c \{L(x, \dot{x}, c)\}. \tag{65}$$

Following the approaches in deriving (35), we get the conjugate Hamiltonian of L_x:

$$H_x(x, p) = \sup_\beta \{p\beta - L_x(x, \beta)\} = \sup_\beta \{p\beta - \inf_c \{L(x, \beta, c)\}\}$$
$$= \sup_\beta \sup_c \{p\beta - L(x, \beta, c)\} = \sup_c \sup_\beta \{p\beta - L(x, \beta, c)\}$$
$$= \sup_c \{H(x, c, p)\} = \sup_c \inf_\varphi \{\tilde{H}(x, c, p, \varphi)\} \tag{66}$$

$$= \tilde{k}_2 s(e^p - 1) + \gamma x(e^{-p} - 1) - (\sqrt{\tilde{k}_1(1-s)} - \sqrt{(\tilde{k}_{-1} + \tilde{k}_2 e^p)s})^2, \tag{67}$$

where

$$s = \frac{1}{2} + \frac{s_1}{2\sqrt{s_1^2 + 4}}, \quad s_1 = \frac{\tilde{k}_1 - (\tilde{k}_{-1} + \tilde{k}_2)}{\sqrt{\tilde{k}_1(\tilde{k}_{-1} + \tilde{k}_2 e^p)}}.$$

We have the mean-field ODE by local analysis

$$\frac{dx}{dt} = \frac{\partial H_x}{\partial p}\Big|_{p=0} = \frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2} - \gamma x. \tag{68}$$

This is consistent with the Michaelis–Menten law shown in (56). To see this, we first note that the reaction rate v should be rescaled with ϵ^{-1} since (68) is for the concentration variable x instead of n. We have

15

$$\lim_{\epsilon \to 0+} \epsilon \frac{k_2 [S]}{[S] + k_M} = \lim_{\epsilon \to 0+} \epsilon \frac{\tilde{k}_1 \epsilon^{-1} \tilde{k}_2 \epsilon^{-1}}{\tilde{k}_1 \epsilon^{-1} + \tilde{k}_{-1} \epsilon^{-1} + \tilde{k}_2 \epsilon^{-1}} = \frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2}.$$

Furthermore, the second-order expansion of H_x with respect to p

$$\frac{\partial^2 H_x}{\partial p^2}|_{p=0} = \frac{2\tilde{k}_1^2 \tilde{k}_2^2}{(\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2)^3} + \frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2} + \gamma x \tag{69}$$

yields the following chemical Langevin approximation

$$\begin{aligned}
dx = {} & \left(\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2} - \gamma x \right) dt \\
& + \sqrt{\epsilon} \left(\sqrt{\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2} + \frac{2\tilde{k}_1^2 \tilde{k}_2^2}{(\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2)^3}} \, dw_1 - \sqrt{\gamma x} \, dw_2 \right).
\end{aligned} \tag{70}$$

Although the above result is quite natural based on our derivations in previous sections, the application to the Michaelis–Menten system again tells us that the strict correspondence between the drift and diffusion terms in the classical large volume limit is lost.

(ii). **Case 1:** $\alpha > 1$.

In this regime, $k_1 [S]$ and k_{-1} are much larger than k_2. Similar to in Case 2, we have the transition probability density (62) with Lagrangian (63), and thus the Hamiltonian (66) for the slow variable x. We get

$$\begin{aligned}
H_x(x, p) = {} & \sup_c \inf_\varphi \{ \tilde{k}_2 c_2 (e^p - 1) + \gamma x (e^{-p} - 1) \\
& + c_2 \left(\frac{\tilde{k}_{-1}}{\epsilon^{\alpha-1}} + \tilde{k}_2 e^p \right) (e^\varphi - 1) + c_1 \frac{\tilde{k}_1}{\epsilon^{\alpha-1}} (e^{-\varphi} - 1) \} \\
= {} & \sup_c \{ \tilde{k}_2 c_2 (e^p - 1) + \gamma x (e^{-p} - 1) \\
& - \frac{1}{\epsilon^{\alpha-1}} (\sqrt{c_2 (\tilde{k}_{-1} + \tilde{k}_2 e^p \epsilon^{\alpha-1})} - \sqrt{c_1 \tilde{k}_1})^2 \}.
\end{aligned} \tag{71}$$

As $\epsilon \to 0+$, the singular perturbation analysis suggests the term involving $1/\epsilon^{\alpha-1}$ to be 0, which gives $c_2 = \tilde{k}_1 / (\tilde{k}_{-1} + \tilde{k}_1)$. We obtain

$$H_x(x, p) = \frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1}} (e^p - 1) + \gamma x (e^p - 1). \tag{72}$$

The mean-field limit

$$\frac{dx}{dt} = \frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1}} - \gamma x. \tag{73}$$

Its consistency with the Michaelis–Menten law is straightforward by checking

$$\lim_{\epsilon \to 0+} \epsilon \frac{k_2 [S]}{[S] + k_M} = \lim_{\epsilon \to 0+} \epsilon \frac{\tilde{k}_1 \epsilon^{-\alpha} \tilde{k}_2 \epsilon^{-1}}{\tilde{k}_1 \epsilon^{-\alpha} + \tilde{k}_{-1} \epsilon^{-\alpha} + \tilde{k}_2 \epsilon^{-1}} = \frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1}}$$

for $\alpha > 1$.

J. Phys. A: Math. Theor. **49** (2016) 135204

T Li and F Lin

In this regime the chemical Langevin approximation takes the form

$$
dx = \left(\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1}} - \gamma x \right) dt + \sqrt{\epsilon} \left(\sqrt{\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1}}} dw_1 - \sqrt{\gamma x} dw_2 \right) \tag{74}
$$

and we formally recover the correspondence between the drift and diffusion terms in this case.

(iii). **Case 3:** $0 < \alpha < 1$.

In this regime, we need to select the scaling parameter for n as ϵ^α, namely to define $x_\alpha = n\epsilon^\alpha$. We have

$$
P(n_f, \tau | n_i, 0) \propto \int Dx_\alpha Dc \exp\left(-\frac{1}{\epsilon^\alpha} \int_0^\tau L(x_\alpha, \dot{x}_\alpha, c) dt \right), \tag{75}
$$

where

$$
L(x_\alpha, \dot{x}_\alpha, c) = \sup_p \sup_\varphi \left\{ p\dot{x}_\alpha - \frac{\tilde{k}_2}{\epsilon^{1-\alpha}} c_2 (e^p - 1) - \gamma x_\alpha (e^{-p} - 1) \right.
$$
$$
\left. - c_2 \left(\tilde{k}_{-1} + \frac{\tilde{k}_2 e^p}{\epsilon^{1-\alpha}} \right) (e^\varphi - 1) - c_1 \tilde{k}_1 (e^{-\varphi} - 1) \right\}. \tag{76}
$$

The Hamiltonian corresponding to variable x_α has the form

$$
H_x(x_\alpha, p) = \sup_c \inf_\varphi \left\{ \frac{\tilde{k}_2}{\epsilon^{1-\alpha}} c_2 (e^p - 1) + \gamma x_\alpha (e^{-p} - 1) \right.
$$
$$
+ c_2 \left(\tilde{k}_{-1} + \frac{\tilde{k}_2 e^p}{\epsilon^{1-\alpha}} \right) (e^\varphi - 1) + c_1 \tilde{k}_1 (e^{-\varphi} - 1) \right\}
$$
$$
= \sup_c \left\{ \frac{\tilde{k}_2}{\epsilon^{1-\alpha}} c_2 (e^p - 1) + \gamma x_\alpha (e^{-p} - 1) \right.
$$
$$
\left. - \left(\sqrt{c_2 \left(\tilde{k}_{-1} + \frac{\tilde{k}_2 e^p}{\epsilon^{1-\alpha}} \right)} - \sqrt{c_1 \tilde{k}_1} \right)^2 \right\}.
$$

As $\epsilon \to 0+$, the singular perturbation analysis gives $c_2 = \epsilon^{1-\alpha} \tilde{k}_1 e^p / \tilde{k}_2$, and the final Hamiltonian

$$
H_x(x_\alpha, p) = \tilde{k}_1 (e^p - 1) + \gamma x_\alpha (e^{-p} - 1). \tag{77}
$$

We get the mean-field ODE

$$
\frac{dx_\alpha}{dt} = \tilde{k}_1 - \gamma x_\alpha. \tag{78}
$$

The consistency with the Michaelis–Menten law can be verified by checking

$$
\lim_{\epsilon \to 0+} \epsilon^\alpha \frac{k_2 [S]}{[S] + k_M} = \lim_{\epsilon \to 0+} \epsilon^\alpha \frac{\tilde{k}_1 \epsilon^{-\alpha} \tilde{k}_2 \epsilon^{-1}}{\tilde{k}_1 \epsilon^{-\alpha} + \tilde{k}_{-1} \epsilon^{-\alpha} + \tilde{k}_2 \epsilon^{-1}} = \tilde{k}_1
$$

for $0 < \alpha < 1$. In this regime we have the chemical Langevin approximation as

$$
dx_\alpha = (\tilde{k}_1 - \gamma x_\alpha) dt + \sqrt{\epsilon^\alpha} (\sqrt{\tilde{k}_1} dw_1 - \sqrt{\gamma x_\alpha} dw_2). \tag{79}
$$

Again, the formal correspondence between the drift and diffusion terms is recovered but with special concentration variable definition and rescaling. It is instructive to compare (79) and
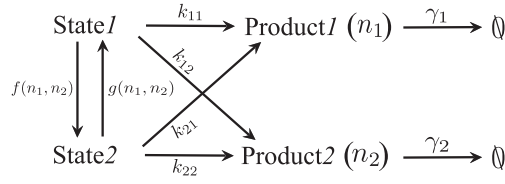
Figure 2. Schematics of a two-scale kinetic model with two kinds of products.

Table 2. LDTs, mean-field limits and Langevin approximations for single-molecule Michaelis–Menten kinetics

	LDT Hamiltonian	Deterministic Drift	Noise in Langevin Approximation
Case 1: ($\alpha > 1$)	$H_x(x, p)$ in (72)	$\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1}} - \gamma x$	$\sqrt{\epsilon}\left(\sqrt{\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1}}}\,dw_1 - \sqrt{\gamma x}\,dw_2\right)$
Case 2: ($\alpha = 1$)	$H_x(x, p)$ in (67)	$\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2} - \gamma x$	$\sqrt{\epsilon}\left(\sqrt{\frac{\tilde{k}_1 \tilde{k}_2}{\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2} + \frac{2\tilde{k}_1^2 \tilde{k}_2^2}{(\tilde{k}_1 + \tilde{k}_{-1} + \tilde{k}_2)^3}}\,dw_1 - \sqrt{\gamma x}\,dw_2\right)$
Case 3: ($\alpha < 1$)	$H_x(x, p)$ in (77)	$\tilde{k}_1 - \gamma x_\alpha$	$\sqrt{\epsilon^\alpha}\,(\sqrt{\tilde{k}_1}\,dw_1 - \sqrt{\gamma x_\alpha}\,dw_2)$

(47) in the case $0 < \alpha < 1$. We have an additional term $\sqrt{\gamma x_\alpha}\,dw_2$ in (79) because of the utilized scaling $x_\alpha = n\epsilon^\alpha$ instead of $x = n\epsilon$ in the two-state model. This reveals the difference between the single-molecule Michaelis–Menten and the two-state genetic switching model. Indeed, in this regime, the reactions can be simplified to

$$ S \xrightarrow[1]{k} P \xrightarrow{\gamma} \varnothing $$
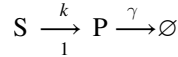
since the production rate is limited by the rate of formation of complex *ES*.

We summarize our findings for the single-molecule Michaelis–Menten kinetics in three regimes in table 2.

6. Discussions and conclusion

The methods and LDT results we proposed in this paper are not limited to the two-state model, single-molecule Michaelis–Menten and single kind of product case. They are indeed general for a class of two-scale kinetic systems. To show this, let us consider the following extension as shown in figure 2.

We assume similar scaling to that considered in (12):

$$ k_{ij} \sim \frac{1}{\epsilon}, \quad f, g \sim \frac{1}{\epsilon^\alpha}, \quad \alpha > 0, \quad (i, j = 1, 2). $$

Define $x_j = n_j \epsilon$, $\tilde{k}_{ij} = k_{ij}\epsilon$ for $i, j = 1, 2$ and $\tilde{f}(x_1, x_2) = f(n_1, n_2)\epsilon^\alpha$, $\tilde{g}(x_1, x_2) = g(n_1, n_2)\epsilon^\alpha$. Performing the same approach as in section 2, we get the transition probability

$$ P(n_f, \tau | n_i, 0) \propto \int Dx Dc \exp\left(-\frac{1}{\epsilon}\int dt L_1(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{c}) - \frac{1}{\epsilon^\alpha}\int dt L_2(\boldsymbol{x}, \boldsymbol{c})\right). $$

Here the Lagrangian

$$L_1(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{c}) = \sup_{\boldsymbol{p}}\{\boldsymbol{p} \cdot \dot{\boldsymbol{x}} - \tilde{H}_1(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{p})\}, \quad L_2(\boldsymbol{x}, \boldsymbol{c}) = \sup_{\boldsymbol{\varphi}}\{-\tilde{H}_2(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{\varphi})\},$$

where $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{c} = (c_1, c_2)$, $\boldsymbol{p} = (p_1, p_2)$, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$ and

$$\tilde{H}_1(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{p}) = (\tilde{k}_{11}c_1 + \tilde{k}_{21}c_2)(e^{p_1} - 1) + \gamma_1 x_1[e^{-p_1} - 1]$$
$$+ (\tilde{k}_{12}c_1 + \tilde{k}_{22}c_2)(e^{p_2} - 1) + \gamma_2 x_2(e^{-p_2} - 1),$$
$$\tilde{H}_2(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{\varphi}) = c_1 \tilde{g}(x)(e^{\varphi_1 - \varphi_2} - 1) + c_2 \tilde{f}(x)(e^{\varphi_2 - \varphi_1} - 1).$$

All of the analyses performed for the two-state model can be applied here to obtain the LDTs for variable $\boldsymbol{x} = (x_1, x_2)$ with different α.

One can also employ the WKB ansatz $P_j(x) \sim \exp(-\epsilon^{-1}\Phi_j(x))$ for the stationary distribution of the stochastic hybrid system (32), where $x = n\epsilon$ and j is a state of DNA. In the asymptotics, one gets a static Hamilton–Jacobi equation for the quasi-potential Φ_j and it turns out Φ_j does not depend on the specific choice of j. However, if not handled appropriately, the WKB approximation may lead to totally different forms of Hamiltonian [15] as mentioned at the end of section 4. This non-uniqueness is due to the lack of variational selection in LDT, which gives a unique Hamiltonian dual to the obtained Lagrangian in rate functional. And this Hamiltonian has the superiority that it is convex with respect to the momentum variable as the by-product of Legendre–Fenchel transform and LDT. This property is important for the nice behavior of numerical discretization.

In this paper, we assume the switching rates between different DNA states are in the order $\epsilon^{-\alpha}$. It is not necessary and could be more general. As long as the switching rates between different DNA states are in $O(\lambda(\epsilon))$, and the cases $\lim_{\epsilon \to 0+}\epsilon\lambda(\epsilon) = \infty$, $O(1)$ and 0 are considered, we will get similar results. In particular, readers may easily verify that if we assume $\lambda(\epsilon) = K\epsilon^{-1}$, then the two-scale LDT Lagrangian with ϵ-scaling in front of $-\ln P$ has the form

$$L_x(x, \dot{x}) = \inf_{\boldsymbol{c}}\{L_1(x, \dot{x}, \boldsymbol{c}) + KL_2(x, \boldsymbol{c})\}, \tag{80}$$

and the LDT Lagrangian with $\lambda(\epsilon)$-scaling in front of $-\ln P$ has the form

$$L_x(x, \dot{x}) = \inf_{\boldsymbol{c}}\{K^{-1}L_1(x, \dot{x}, \boldsymbol{c}) + L_2(x, \boldsymbol{c})\}. \tag{81}$$

When K goes to 0, 1 or ∞, the appropriate choices of scaling recover the desired results shown in the paper.

In conclusion, we established the two-scale LDTs for a class of chemical reaction kinetics through the second quantization path integral formulation. Although not rigorous, we showed that this formal approach is very effective and transparent to understand the two-scale LDTs associated with different reaction channels. This provides essential insights to rigorously prove the corresponding LDTs, which is our ongoing research. We discussed its implication on single-molecule Michaelis–Menten kinetics as well. The proposed framework and results also shed light on our understanding of general multi-scale systems including diffusion processes. It will be interesting to investigate the application of two-scale LDTs in other systems.

Acknowledgments

11222114). The authors also thank Weinan E, Yong Liu and Xiaoguang Li for helpful discussions.

References

[1] Assaf M, Roberts E and Luthey-Schulten Z 2011 Determining the stability of genetic switches: explicitly accounting for mRNA noise *Phys. Rev. Lett.* **106** 2048102

[2] Bender C M and Orszag S A 1999 *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory* (New York: Springer)

[3] Bressloff P C and Faugeras O 2014 On the Hamiltonian structure of large deviations in stochastic hybrid systems (arXiv:1410.2152v1)

[4] Bressloff P C and Newby J M 2013 Metastability in a stochastic neural network modeled as a velocity jump Markov process *SIAM Appl. Dyn. Syst.* **12** 1394

[5] Bühler O 2006 *A Brief Introduction to Classical, Statistical and Quantum Mechanics* (Providence, RI: American Mathematical Society)

[6] Dembo A and Zeitouni O 1998 *Large deviations techniques and applications* 2nd edition (New York: Springer)

[7] Doi M 1976 Second quantization representation for classical many-particle system *J. Phys. A: Math. Gen.* **9** 1465

[8] Faggionato A, Gabrielli D and Crivellari M R 2010 Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors *Markov Process. Relat. Fields* **16** 497

[9] Freidlin M I and Wentzell A D 1998 *Random perturbations of dynamical systems* 2nd edition (New York: Springer)

[10] Ge H, Qian H and Xie X S 2015 Stochastic phenotype transition of a single cell in an intermediate region of gene state switching *Phys. Rev. Lett.* **114** 078101

[11] Heymann M and Vanden-Eijnden E 2008 The geometric minimum action method: A least action principle on the space of curves *Comm. Pure Appl. Math.* **61** 1052

[12] Keener J and Sneyd J 1998 *Mathematical Physiology* (New York: Springer)

[13] Kemeny J G and Snell J L 1960 *Finite Markov chains* (New York, Berlin and Heidelberg: Springer)

[14] Kou S C, Cherayil B J, Min W, English B P and Xie X S 2005 Single-molecule Michaelis-Menten Equations *J. Phys. Chem.* B **109** 19068

[15] Li T and Lin F 2015 Large deviations for two-scale chemical kinetic processes arXiv:1504.03781

[16] Liptser R 1996 Large deviations for two scaled diffusions *Prob. Theory Relat. Fields* **106** 71

[17] Lv C, Li X, Li F and Li T 2014 Constructing the energy landscape for genetic switching system driven by intrinsic noise *PLoS ONE* **9** e88167

[18] Lv C, Li X, Li F and Li T 2015 Energy landscape reveals that the budding yeast cell cycle is a robust and adaptive multi-stage process *PLoS Comp. Biol.* **11** e1004156

[19] Michaelis L and Menten M L 1913 Die Kinetik der Invertinwerkung *Biochem. Z.* **49** 333

[20] Newby J M 2012 Isolating intrinsic noise sources in a stochastic genetic switch *Phys. Biol.* **9** 026002

[21] Newby J M 2014 Spontaneous Excitability in the Morris–Lecar Model with Ion Channel Noise *SIAM J. Appl. Dyn. Syst.* **13** 1756

[22] Newby J M and Bressloff P C 2010 Local synaptic signaling enhances the stochastic transport of motor-driven cargo in neurons *Phys. Biol.* **7** 036004

[23] Newby J M and Chapman J 2014 Metastable behavior in Markov processes with internal states *J. Math. Biol.* **69** 941

[24] Peliti L 1985 Path integral approach to birth-death processes on a lattice *J. Phys.* **46** 1469

[25] Rockafellar R T and Wets R J B 1998 *Variational Analysis* (Berlin: Springer)

[26] Shwartz A and Weiss A 1995 *Large deviations for performance analysis: queues, communications and computing* (London: Chapman and Hall)

[27] Touchette H 2009 The large deviation approach to statistical mechanics *Phys. Rep.* **478** 1

[28] Varadhan S R S 1984 *Large deviations and applications* (Philadelphia, PA: SIAM)

[29] Veretennikov A Y 2000 On large deviations for SDEs with small diffusion and averaging *Stoch. Process. Appl.* **89** 69

[30] Veretennikov A Y 1999 On large deviations in the averaging principle for SDE's with a full dependence *Ann. Prob.* **27** 284

[31] Zhang B and Wolyness P G 2014 Stem cell differentiation as a many-body problem *Proc. Natl. Acad. Sci. U.S.A* **111** 10185

[32] Zhang K, Sasai M and Wang J 2013 Eddy current and coupled landscapes for nonadiabatic and nonequilibrium complex system dynamics *Proc. Natl. Acad. Sci. U.S.A* **110** 14930

[33] Zhou P and Li T 2015 Realization of Waddington's metaphor: Potential landscape, quasi-potential, A-type integral and beyond (arXiv:1511.02088)