Large deviations for two scale chemical kinetic processes

Tiejun Li ∗ and Feng Lin†

LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China

ABSTRACT

We formulate the large deviations for a class of two scale chemical kinetic processes motivated from biological applications. The result is successfully applied to treat a genetic switching model with positive feedbacks. The corresponding Hamiltonian is convex with respect to the momentum variable as a by-product of the large deviation theory. This property ensures its superiority in the rare event simulations compared with the result obtained by formal WKB asymptotics. The result is of general interest to understand the large deviations for multiscale problems.

1. INTRODUCTION

We will investigate the large deviations for a class of two scale chemical kinetic processes with the slow variable $z_n \in \mathbb{N}^d / n$ satisfying

$$z_n(t) = z_n(0) + \sum_{i=1}^{S} \frac{1}{n} \mathbb{I}_i \left( n \int_0^t \lambda_i(z_n(s), \xi_n(s)) \, ds \right) u_i \quad (1.1)$$

subject to some fixed initial state $z_n(0) = z^0$, where $\{P_i(t)\}_{i=1,\ldots,S}$ are independent uni-rate Poisson processes, $\lambda_i \in \mathbb{R}^+$ is called the propensity function which characterizes the reaction rate of the $i$th reaction and $u_i \in \mathbb{Z}^d$ is the state change vector. The number $n \in \mathbb{N}$ corresponds to the system volume, thus $z_n$ has the meaning of concentration (number of molecules per volume) for the considered kinetic system. The fast variable $\xi_n \in \mathbb{Z}_D := \{1,2,\cdots,D\}$ is a simple jump process with the time dependent rate $nq_{ij}(z_n(t))$ from state $i$ to $j$ at time $t$. With this mathematical setup, the processes $z_n(t)$ and $\xi_n(t)$ are fully coupled each other.

∗email: tieli@pku.edu.cn
†email: math_linfeng@pku.edu.cn
and the infinitesimal generator $L_n$ of this system has the form

$$L_n h(z, i) = n \sum_{l=1}^{S} \lambda_l(z, i) [h(z + u_l/n, i) - h(z, i)] + n \sum_{j=1}^{D} q_{ij}(z) [h(z, j) - h(z, i)], \quad (1.2)$$

where $z \in \mathbb{N}^d / n$, $i \in \mathbb{Z}^D$ and $h$ is any compactly supported smooth function of $z$ for each $i$. For more about the notations and the backgrounds on the chemical kinetic processes, the readers may be referred to [8, 10].

The above problem is motivated by our recent rare event study in the biological applications [1, 16]. In a cell, the reactions underlying gene expression usually involve low copy number of molecules, such as DNA, mRNAs and transcription factors, so the stochasticity in gene regulation process is inevitable even under constant environmental conditions [7]. When the number of the molecules for all species goes to infinity and the law of mass action holds for the propensity functions, one gets the well-known large volume limit or Kurtz’s limit, which gives the deterministic reaction rate equations for the concentration of the species [13]. The convergence result can be further refined to the large deviation type [19]. Recently, the following typical biological model with positive feedbacks is utilized to investigate the robustness of the genetic switching system [1, 16].

Denote $P_{m,n}$ and $Q_{m,n}$ as the probability distribution functions for the inactive (DNA$_{\text{in}}$) and active DNA (DNA$_{\text{act}}$) states with $m$ mRNAs and $n$ proteins at time $t$, respectively. They satisfy the following forward Kolmogorov equation

$$\partial_t P_{m,n} = g(n)Q_{m,n} - f(n)P_{m,n} + \mathcal{A} P_{m,n},$$

$$\partial_t Q_{m,n} = -g(n)Q_{m,n} + f(n)P_{m,n} + [\mathcal{A} + a(\mathcal{E}_m^{-1} - 1)]Q_{m,n}. \quad (1.3)$$

Here we employ the notation for raising operator $\mathcal{E}_n^j$ acting on $f(n)$ as $\mathcal{E}_n^j f(n) = f(n+j)$, and $\mathcal{A} = (\mathcal{E}_m^{-1} - 1)n + \gamma (\mathcal{E}_m^{-1} - 1)m + \gamma bm(\mathcal{E}_m^{-1} - 1)$ is a birth-death operator related to the DNA active state. In the regime $\gamma, b \sim O(1)$, $a, f(n), g(n) \sim O(V)$ when $m, n \sim O(V)$, one can obtain a mean field ODE
system for $x = m/V$ and $y = n/V$ as

\[
\frac{dx}{dt} = \frac{\hat{a} f(y)}{f(y) + g(y)} - \gamma x, \quad \frac{dy}{dt} = \gamma bx - y \quad \text{with} \quad \hat{a} := \lim_{V \to \infty} a/V, \tag{1.4}
\]

when $V$ goes to infinity through the perturbation analysis for the infinitesimal generator [14, 16, 17]. This problem is a special case of our formulation shown at the beginning of this paper for $d = 2$, $D = 2$ and $S = 4$. With suitable choice for the functions $f(n)$ and $g(n)$, the final mean field ODEs has two stable stationary points and there are noise induced transitions between these two states when $V$ is finite. To understand the robustness of the genetic switching, the biophysicists employed the WKB ansatz [1]

\[
P_{m,n} = P(x, y) \sim \exp[-V S(x, y)] \tag{1.5}
\]

and obtained a steady state Hamilton-Jacobi equation $H(x, y, \nabla S) = 0$. Mathematically the function $S$ resembles the role of the global quasi-potential of the stochastic dynamical system [9] but it is not sure whether it is the case in the current stage. Another related physics approach to study a similar switching system is to utilize the spin-boson path integral formalism in quantum field theory and then take the semiclassical approximation and adiabatic limit [22]. Both approaches are difficult to be rationalized in mathematical sense. So how to formulate this problem in a mathematically rigorous way? To resolve this issue, we have to answer the following two fundamental questions.

(1) Question 1. What is the large deviation principle (LDP) associated with the system (1.2)? Presumably, we can obtain the Lagrangian from the large deviation analysis, then get the Hamiltonian $H$ through the Legendre-Fenchel transform.

(2) Question 2. What is the relation between the rigorously obtained Hamiltonian $H$ in the above question and the Hamiltonian obtained via WKB asymptotics?

The aim of this paper is to make an exploration on these two questions. To do this, we first note that the large volume limit no longer holds in the current example. Although the mRNA and protein copy numbers scale as $V$, we have only one DNA, which switches between the active and inactive states. This fact excludes the direct applicability of the LDP results in [19]. However, the fast switching between the two states of the DNA makes the LDP analysis still feasible by incorporating the Donsker-Varadhan type large deviations. Indeed, similar situation has been nicely
discussed by Liptser in [15] for two scale diffusions as

\[ dX_n(t) = A(X_n(t), \xi_n(t))dt + \frac{1}{\sqrt{n}}B(X_n(t), \xi_n(t))dW_t, \quad (1.6) \]

\[ d\xi_n(t) = nb(\xi_n(t))dt + \sqrt{n}\sigma(\xi_n(t))dV_t. \quad (1.7) \]

The main idea of this paper is to generalize the result in [15] to our two scale chemical kinetic processes. As we will see, although the framework is similar, we have to deal with the technicalities brought by the jump processes and the full coupling between the fast and slow variables ($\xi_n$ is independent of $X_n$ in (1.7)).

To state the main results of this paper, let us introduce the occupation measure $\nu_n$ on $(\mathbb{R}_+ \times \mathbb{Z}_D, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{Z}_D))$ corresponding to $\xi_n$

\[ \nu_n(\Delta \times \Gamma) = \int_0^\infty 1(t \in \Delta, \xi_n \in \Gamma)dt, \quad \Delta \in \mathcal{B}(\mathbb{R}_+), \Gamma \in \mathcal{B}(\mathbb{Z}_D). \quad (1.8) \]

Denote $\mathbb{D} = \mathbb{D}[0, \infty)$ the space of functions on $[0, \infty)$ which are right continuous with left-hand limits, $\mathcal{M} = \mathcal{M}[0, \infty)$ of $\sigma$-finite (locally in $t$) measures $\nu = \nu(dt, i)$ on $(\mathbb{R}_+ \times \mathbb{Z}_D, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{Z}_D))$, and $\mathcal{M}_{ac}$ for $\nu \in \mathcal{M}$ such that $\nu(dt, i)$ is absolutely continuous with respect to $dt$ with the form $\nu(dt, i) = n_v(t, i)dt$ and $n_v(t, i)$ is a probabilistic transition kernel, i.e. $\sum_{i \in \mathbb{Z}_D} n_v(t, i) = 1$. The $\nu_n$ we considered always belongs to $\mathcal{M}_{ac}$. Define metrics $\rho^{(1)}$ and $\rho^{(2)}$ in $\mathbb{D}$ and $\mathcal{M}$ by letting

\[ \rho^{(1)}(z, \tilde{z}) = \sum_{k=1}^\infty \rho^{(1)}_k(z, \tilde{z}) \wedge 1 2^k \quad \text{and} \quad \rho^{(2)}(\nu, \tilde{\nu}) = \sum_{k=1}^\infty \rho^{(2)}_k(\nu, \tilde{\nu}) \wedge 1 2^k, \]

where $\rho^{(1)}_T$ is the Skorohod distance in $\mathbb{D}[0, T]$ and $\rho^{(2)}_T$ is the Levy-Prohorov distance in $\mathcal{M}[0, T]$. Our task is to establish the LDP for the pair $(z_n, \nu_n)$ in metric space $(D, \times \mathcal{M}, \rho^{(1)} \times \rho^{(2)})$.

This paper is organized as follows. In Section 2, we present the main large deviation theorem and give the rate functional of the whole system. By using the contraction principle and the Legendre-Fenchel transform we get the Hamiltonian related to the slow variable $z_n$. As a concrete application, we then study the genetic switching model and compare the difference between the rigorously obtained Hamiltonian and that obtained by WKB ansatz. In Sections 3 and 4, we give the proof of the main theorem. Due to the technicalities to handle the non-negativity constraint for $(x, y)$, we decompose the proof procedure into two steps. In Section 3, we prove the LDT theorem by relaxing the bounded domain condition to the whole space case. The upper bound estimate is standard in some sense. However, the proof of the lower bound is technical because of the full coupling between the fast and slow variables. The resolution is based on the approximation and change-of-measure approach.
The central idea is to make a piecewise linear approximation to any given path and occupation measure \((r, \nu)\) by \((y, \pi)\) at first, and then construct suitable new processes \((\tilde{z}_n, \tilde{\nu}_n)\) such that \(\mathbb{P} - \lim_{n \to \infty} \rho^{(1)}(\tilde{z}_n, y) = 0\) and \(\mathbb{P} - \lim_{n \to \infty} \rho^{(2)}(\tilde{\nu}_n, \pi) = 0\). This turns out to be technical and one key part of the whole paper. In Section 4, we strengthen the result to the half space case. Some details are left in the Appendix.

This paper should be considered as the companion of [16] for studying the rare events in genetic switching system, and it is of general interest to understand the large deviations for multiscale problems [4, 5].

2. Main Result and Its Application

2.1. Main Theorem. We need the following technical assumptions for our main result.

**Assumption 2.1.** Let \(G := \overline{R^d_+}\). Assume the following regularity conditions for the propensity functions and jump rates hold.

1. (a) For each \(i \in \{1, 2, \ldots, S\}, j \in \mathbb{Z}_D\) and all \(z, x \in G\), there is a constant \(K\) such that
   \[
   |\lambda_i(z, j) - \lambda_i(x, j)| \leq K|z - x|.
   \]
   (b) For each \(i \in \{1, 2, \ldots, S\}, j \in \mathbb{Z}_D\) and all \(z \in G^0\), \(\lambda_i(z, j) > 0\).
   (c) For each \(x \in \partial G\) and \(y \in \mathbb{L}\{u_j|\lambda_j(x) > 0\}\), we have \(x + y \in G\), where \(\mathbb{L}\{u_j\}\) is the positive cone spanned by the vectors \(\{u_j\}\) defined as
   \[
   \mathbb{L}\{u_j\} := \{v|\text{there exist } \alpha_j \geq 0 \text{ such that } v = \sum_j \alpha_j u_j\}.
   \]
2. For each \(i, j \in \mathbb{Z}_D\), \(\log q_{ij}(z)\) are bounded and Lipschitz continuous with respect to \(z \in G\).

These assumptions hold in our application example in Section 2.2.

**Theorem 2.2.** Under the Assumption 2.1, the family \((z_n, \nu_n)\) defined by (1.1) and (1.8) obeys the LDP in \((\mathbb{D}^d \times \mathbb{M}, \rho^{(1)} \times \rho^{(2)})\) with a good rate functional \(I(r, \nu) = I_s(r, \nu) + I_f(r, \nu)\), i.e.

0. \(I(r, \nu)\) values in \([0, +\infty]\) and its level sets are compact in \((\mathbb{D}^d \times \mathbb{M}, \rho^{(1)} \times \rho^{(2)})\),
1. for every close set \(F \in \mathbb{D}^d \times \mathbb{M},\)
   \[
   \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}((z_n, \nu_n) \in F) \leq -\inf_{(r, \nu) \in F} I(r, \nu),
   \]  
2. for every open set \(G \in \mathbb{D}^d \times \mathbb{M},\)
   \[
   \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}((z_n, \nu_n) \in G) \geq -\inf_{(r, \nu) \in G} I(r, \nu),
   \]
where the rate functional for the slow variables

\[
I_s(r, v) = \begin{cases} \int_0^\infty L_s(r(t), \dot{r}(t), n_v(t, \cdot)) dt, & \text{otherwise,} \\
\infty, & d r(t) = \dot{r}(t) d t, \quad v(d t, \cdot) = n_v(t, \cdot) d t, 
\end{cases}
\]

Assumption 2.3.

The slow variables Corollary 2.5.

Theorem 2.4.

rate functional

\[
\{ \langle \cdot, \cdot \rangle \}_{i j} = \begin{cases} \sum_{i j = 1}^D \lambda_i(r(t), j) n_v(t, j)(e^{\langle p, i j \rangle} - 1), & (2.3) \\
\infty, & \text{otherwise,}
\end{cases}
\]

and the rate functional for the fast variables

\[
I_f(r, v) = \begin{cases} \int_0^\infty S(r(t), n_v(t, \cdot)) dt, & \text{otherwise,} \\
\infty, & d r(t) = \dot{r}(t) d t, \quad v(d t, \cdot) = n_v(t, \cdot) d t,
\end{cases}
\]

where \( \langle \cdot, \cdot \rangle \) is the inner product in the Euclidean space, \( e_{i j} = e_i - e_j \) and \( \{ e_{i j} \}_{i j = 1}^D \) are canonical basis in Euclidean space \( \mathbb{R}^D \). We take the convention that \( r(t), v(d t, \cdot) \) are absolutely continuous with respect to time when we use the notation \( d r(t) = \dot{r}(t) d t, v(d t, \cdot) = n_v(t, \cdot) d t \).

The proof of Theorem 2.2 relies on first establishing a weaker statement based on the following stronger assumption on the whole space.

Assumption 2.3. Regularity for the propensity functions and jump rates.

1. For each \( i \in [1, 2, \ldots, S] \), \( j \in \mathbb{Z}_D \), \( \log \lambda_i(z, j) \) is bounded and Lipschitz continuous with respect to \( z \in \mathbb{R}^d \).
2. For each \( i, j \in \mathbb{Z}_D \), \( \log q_{i j}(z) \) are bounded and Lipschitz continuous with respect to \( z \in \mathbb{R}^d \).

Theorem 2.4. The large deviation result in Theorem 2.2 hold for \( z_n, r \in \mathbb{R}^d \) under the Assumption 2.3.

As a straightforward application of the contraction principle, we have

Corollary 2.5. The slow variables \( z_n \) obeys the LDP in \( (\mathbb{D}^d, \rho^{(1)}) \) with the rate functional

\[
I(r) = \inf_{v \in \mathcal{M}} (I_s(r, v) + I_f(r, v)).
\]

For \( v \in \mathcal{M}_{ac} \), we denote its Radon-Nikodym derivative as \( n_v(t, \cdot) \) and we have \( n_v(t, \cdot) \geq 0, \sum_{i \in \mathbb{Z}_D} n_v(t, i) = 1 \). Define the set of probabilistic transition kernels as \( \Delta_D = \{ w : w_1, w_2, \ldots, w_D \geq 0, \sum_{i = 1}^D w_i = 1 \} \) where \( w = (w_1, w_2, \ldots, w_D) \). We will take the notation \( n_v \in \mathcal{M}_{ac} \) when \( v \in \mathcal{M}_{ac} \) and \( n_v(t, \cdot) \in \Delta_D \) for any \( t \geq 0 \) in later texts.
Lemma 2.6.

\[
\inf_{n_v \in \mathcal{M}_{ac}} \int_0^\infty L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \, dt \\
= \int_0^\infty \inf_{n_v \in \mathcal{M}_{ac}} \{ L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \} \, dt. \tag{2.4}
\]

Proof. It is straightforward to have that

\[
\inf_{n_v \in \mathcal{M}_{ac}} \int_0^\infty L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \, dt \\
\geq \int_0^\infty \inf_{n_v \in \mathcal{M}_{ac}} \{ L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \} \, dt.
\]

Let us show the converse part. Given \( \epsilon > 0 \), for any \( t \geq 0 \), there exist \( n_{v_t} \in \Delta_D \) such that

\[
\inf_{n_v \in \mathcal{M}_{ac}} \{ L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \} \\
\geq L_s(r(t), \dot{r}(t), n_{v_t}(\cdot)) + S(r(t), n_{v_t}(\cdot)) - 2^{-[t]} \epsilon. \tag{2.5}
\]

where \([t]\) represents the smallest integer which is larger than \( t \).

Define the probabilistic transition kernel \( n_{\hat{\nu}} \) as

\[
n_{\hat{\nu}}(t, \cdot) = n_{\nu_t}(\cdot)
\]

and define

\[
\hat{\nu}(dt, \cdot) = n_{\nu_t}(t, \cdot) \, dt.
\]

Then \( \hat{\nu} \in \mathcal{M}_{ac} \) and

\[
\int_0^\infty \inf_{n_v \in \mathcal{M}_{ac}} \{ L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \} \, dt \\
\geq \int_0^\infty L_s(r(t), \dot{r}(t), n_{\nu_t}(t, \cdot)) + S(r(t), n_{\nu_t}(t, \cdot)) \, dt - \sum_{k=1}^{\infty} 2^{-k} \epsilon \\
\geq \inf_{n_v \in \mathcal{M}_{ac}} \int_0^\infty L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \, dt - \epsilon.
\]

The proof is completed. \( \square \)

Define

\[
L(r, \beta) = \inf_{n_v \in \mathcal{M}_{ac}} \{ L_s(r, \beta, n_v(t, \cdot)) + S(r, n_v(t, \cdot)) \}. \tag{2.6}
\]
By Lemma 2.4, we have
\[
I(r) = \inf_{\nu \in \mathbb{R}} (I_s(r, \nu) + I_f(r, \nu))
\]
\[
= \inf_{n_v \in \mathbb{M}_{ac}} \int_0^\infty L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \, dt
\]
\[
= \int_0^\infty \inf_{n_v \in \mathbb{M}_{ac}} \left\{ L_s(r(t), \dot{r}(t), n_v(t, \cdot)) + S(r(t), n_v(t, \cdot)) \right\} \, dt
\]
\[
= \int_0^\infty L(r(t), \dot{r}(t)) \, dt.
\]

**Lemma 2.7.** $L(r, \beta)$ is convex in $\beta$.

**Proof.** By Lemma 5.3,
\[
L(r, \beta) = \inf_{n_v \in \mathbb{M}_{ac}} (L_s(r, \beta, n_v(t, \cdot)) + S(r, n_v(t, \cdot)))
\]
\[
= \inf_{n_v \in \mathbb{M}_{ac}} \sup_{\theta \in \mathbb{R}^d} (\langle \theta, \beta \rangle - H_s(r, \theta, n_v(t, \cdot)) + S(r, n_v(t, \cdot)))
\]
\[
= \sup_{\theta \in \mathbb{R}^d} \inf_{n_v \in \mathbb{M}_{ac}} (\langle \theta, \beta \rangle - H_s(r, \theta, n_v(t, \cdot)) + S(r, n_v(t, \cdot))).
\]

It is easy to see that $\inf_{n_v \in \mathbb{M}_{ac}} (\langle \theta, \beta \rangle - H_s(r, \theta, n_v(t, \cdot)) + S(r, n_v(t, \cdot))$ is convex in $\beta$, thus $L(r, \beta)$ is convex in $\beta$ according to Lemma 5.2. \qed

It is well-known that the Lagrangian $L_s$ does not have a closed form for the standard chemical reaction kinetic system, instead it is more convenient to investigate its dual Hamiltonian $H_s$ by Legendre-Fenchel transform. The explicit form of the Hamiltonian is important for the numerics to study the rare events in systems biology [12]. With similar idea, we have

\[
H(r, \theta) = \sup_{\beta \in \mathbb{R}^d} (\langle \theta, \beta \rangle - L(r, \beta))
\]
\[
= \sup_{\beta \in \mathbb{R}^d} \left\{ \langle \theta, \beta \rangle - \inf_{n_v \in \mathbb{M}_{ac}} (L_s(r, \beta, n_v(t, \cdot)) + S(r, n_v(t, \cdot))) \right\}
\]
\[
= \sup_{\beta \in \mathbb{R}^d} \sup_{n_v \in \mathbb{M}_{ac}} \left\{ \langle \theta, \beta \rangle - L_s(r, \beta, n_v(t, \cdot)) - S(r, n_v(t, \cdot)) \right\}
\]
\[
= \sup_{n_v \in \mathbb{M}_{ac}} \sup_{\beta \in \mathbb{R}^d} \left\{ \langle \theta, \beta \rangle - L_s(r, \beta, n_v(t, \cdot)) - S(r, n_v(t, \cdot)) \right\}
\]
\[
= \sup_{n_v \in \mathbb{M}_{ac}} (H_s(r, \theta, n_v(t, \cdot)) - S(r, n_v(t, \cdot))).
\] (2.7)

A consequence about $H$ from its definition is that $H$ is convex with respect to $\theta$ from the convexity of $L$ and the Legendre-Fenchel transform [6]. Furthermore if the matrix $Q = (q_{ij})_{D \times D}$ is symmetrizable, $S(r, n_v(t, \cdot))$
2.2. Application to the genetic switching model. The formula (2.7) has a nice application in the genetic switching model introduced before. In this model, we have $d = 2, D = 2$ and $S = 4$. Denote $x = (x, y) = (m, n) / V$ the slow variables after large volume scaling. For better use of notation, here we take the fast variable $\xi \in [0, 1]$ instead of $[1, 2]$ to represent that the DNA is at inactive ($\xi = 0$) or active state ($\xi = 1$), respectively. By taking into account the scaling of parameters

$$a \sim V b^{-1}, \, f(n) \sim V f(y), \, g(n) \sim V g(y)$$

considered in [1], we have the jump rates for DNA

$$q_{01}(x, y) = f(y), \quad q_{10}(x, y) = g(y),$$

and the following list of reactions associated with slow variables.

**Table 1.** Reaction schemes and parameters

| Reaction scheme          | Propensity function | State change vector |
|--------------------------|---------------------|---------------------|
| DNA$_{\text{act}} \rightarrow$ mRNA | $\lambda_1(x, y, \xi) = b^{-1} \xi$ | $u_1 = (1, 0)$ |
| mRNA $\rightarrow$ $\varnothing$               | $\lambda_2(x, y, \xi) = \gamma x$ | $u_2 = (-1, 0)$ |
| mRNA $\rightarrow$ Protein            | $\lambda_3(x, y, \xi) = \gamma b x$ | $u_3 = (0, 1)$ |
| Protein $\rightarrow$ $\varnothing$     | $\lambda_4(x, y, \xi) = y$ | $u_4 = (0, -1)$ |

With this setup, we have

$$H_s(x, p, n_v) = b^{-1} n_v(\cdot, 1)(e^{p_x} - 1) + A(x, y, p_x, p_y),$$

(2.9)

where $A(x, y, p_x, p_y) = \gamma x(e^{-p_x} - 1) + \gamma b x(e^{p_y} - 1) + y(e^{-p_y} - 1)$ is the corresponding coordinate form of the operator $\mathcal{A}$ in (1.3), and

$$S(x, n_v) = \left(\sqrt{n_v(\cdot, 0) f(y)} - \sqrt{n_v(\cdot, 1) g(y)}\right)^2.$$ 

Applying (2.7) with the constraints $n_v(\cdot, 0) + n_v(\cdot, 1) = 1$ and $n_v(\cdot, 0), n_v(\cdot, 1) \geq 0$, we obtain the final Hamiltonian

$$H(x, p) = b^{-1} s(e^{p_x} - 1) - \left(\sqrt{(1 - s) f(y) - \sqrt{s g(y)}}\right)^2 + A(x, y, p_x, p_y),$$

(2.10)

where

$$s = \frac{1}{2} + \frac{s_1}{2\sqrt{s_1^2 + 4}}, \quad s_1 = \frac{b^{-1}(e^{p_x} - 1) + f(y) - g(y)}{\sqrt{f(y) g(y)}},$$

$$s_1 = \frac{b^{-1}(e^{p_x} - 1) + f(y) - g(y)}{\sqrt{f(y) g(y)}}.$$
It is instructive to compare this Hamiltonian with that obtained via WKB asymptotics by plugging the ansatz (1.5) into (1.3) and keeping only the lowest order terms. This procedure gives a new form of the Hamiltonian

\[ \tilde{H}(x, p) = A + g(y)^{-1}[A + b^{-1}(e^{p_x} - 1)][f(y) - A], \quad (2.11) \]

where \( A = A(x, y, p_x, p_y) \). The relation between the Hamiltonian \( \tilde{H} \) and \( H \) is not clear so far. But one crucial difference is that \( H \) is convex with respect to the momentum variable \( p \) from the form (2.7), while \( \tilde{H} \) is not. It turns out this property is crucial for the numerical computations, especially for computing the transition path in geometric minimum action method (gMAM) [12]. It is also interesting to observe that the quasi-potential \( S(x, y) \) obtained from

\[ H(x, \nabla S) = 0 \quad \text{or} \quad \tilde{H}(x, \nabla S) = 0 \]

is the same even \( H \) and \( \tilde{H} \) are so different [16]. It can be also verified that \( H \) is not the convex hull of \( \tilde{H} \) with respect to \( p \). From the Hamilton-Jacobi theory, one may speculate that these two Hamiltonians are connected through some canonical transformation. But it is only a plausible answer which is difficult to be verified even for this concrete example.

As the large deviation results give the sharpest characterization of the considered two-scale chemical kinetic system, we can obtain the deterministic mean field ODEs and the chemical Langevin approximation for the system based on the large deviations [3], which corresponds to the law of large numbers (LLN) and the central limit theorem (CLT) for the process. Taking advantage of (2.10), we get

\[ \frac{\partial H}{\partial p_x} \bigg|_{p=0} = b^{-1} f(y) \frac{1}{f(y) + g(y)} - \gamma x, \quad \frac{\partial H}{\partial p_y} \bigg|_{p=0} = \gamma bx - y. \quad (2.12) \]

The mean field ODEs defined by

\[ \frac{dx}{dt} = \left. \frac{\partial H}{\partial p_x} \right|_{p=0} \quad \text{and} \quad \frac{dy}{dt} = \left. \frac{\partial H}{\partial p_y} \right|_{p=0} \]

are exactly (1.4) with \( \tilde{a} = \lim_{V \rightarrow \infty} a/V = b^{-1} \). Furthermore, we have

\[ \left. \frac{\partial^2 H}{\partial p_x^2} \right|_{p=0} = b^{-1} f(y) \frac{1}{f(y) + g(y)} + \frac{2b^{-1} f(y)g(y)}{(f(y) + g(y))^3} + \gamma x, \]

\[ \left. \frac{\partial^2 H}{\partial p_y^2} \right|_{p=0} = \gamma bx + y. \]
This naturally leads to the following chemical Langevin approximation

\[
\begin{align*}
    dx_t &= \left[ \frac{b^{-1}f}{f + g} - \gamma x \right] dt + \frac{1}{\sqrt{V}} \left[ \sqrt{\frac{b^{-1}f}{f + g} + \frac{2b^{-1}fg}{(f + g)^3}} dw_t^1 - \sqrt{\gamma x_t} dw_t^2 \right], \\
    dy_t &= \left[ \gamma bx_t - y_t \right] dt + \frac{1}{\sqrt{V}} \left[ \sqrt{\gamma bx_t dw_t^3 - \gamma y_t dw_t^4} \right],
\end{align*}
\]

(2.14)

where \( f, g \) are abbreviations of functions \( f(y) \) and \( g(y) \), and \( w_t^i (i = 1, \ldots, 4) \) are independent standard Brownian motions. It is instructive to compare (2.14) with a granted formulation by naively transplanting the Langevin approximation from the simple large volume limit [11], where the equation for \( x_t \) reads

\[
\begin{align*}
    dx_t &= \left[ \frac{b^{-1}f}{f + g} - \gamma x \right] dt + \frac{1}{\sqrt{V}} \left[ \sqrt{\frac{b^{-1}f}{f + g} dw_t^1 - \sqrt{\gamma x_t} dw_t^2} \right], \\
\end{align*}
\]

(2.15)

and the equation for \( y_t \) is the same. It is remarkable that the Eq. (2.14) has an additional term related to the noise \( dw_t^1 \). This additional fluctuation is induced by the fast switching of DNA states. Similar situation will also occur when we derive the chemical Langevin equations for enzymatic reactions, whereas we should take the fluctuation effect of the fast switching into consideration. However, this point does not seem to be paid much attention in previous research.

2.3. Some properties of the Hamiltonian \( H \). The Hamiltonian \( H(r, \theta) \) has some nice properties which can be utilized to simplify the computations in many cases. According to (2.7), we have

\[
H(r, \theta) = \sup_{w \in \Delta_D} h(r, \theta, w),
\]

where

\[
h(r, \theta, w) = H_r(r, \theta, w) - S(r, w)
\]

\[
= \sum_{i=1}^D \sum_{j=1}^D \lambda_i(r, j) w_j(e^{(\theta, w_i)} - 1)
\]

\[-\frac{1}{2} \sum_i \sum_{j \neq i} \left[ \sqrt{w_i q_{ij}(r)} - \sqrt{w_j q_{ji}(r)} \right]^2.
\]

We will show that the supremum of \( h \) in \( \Delta_D \) can be only taken in the interior \( \Delta_D^0 \) of \( \Delta_D \). To do this, we first note that \( h \) is continuous in \( \Delta_D \) and differentiable in \( \Delta_D^0 \). For any \( w_0 \in \partial(\Delta_D) \), define \( v = w_0 - c_0 \) where \( c_0 = (1, 1, \ldots, 1)/D \) is the center of \( \Delta_D \). It is easy to check that

\[
\frac{\partial h}{\partial v}(r, \theta, w_0) = -\infty.
\]
This means that the supremum of $h$ can not be taken in $\partial(\Delta_D)$. Furthermore, since $h$ is strictly concave in $w$, there exists only one point $w^*$ in $\Delta^o_D$, such that

$$w^* = \operatorname{arg sup}_{w \in \Delta_D} h(r, \theta, w).$$

An important consequence of this fact is that we can get the derivative

$$\frac{\partial H(r, \theta)}{\partial \theta} = \frac{dh(r, \theta, w^*)}{d\theta} \bigg|_{w^*} = \frac{\partial h(r, \theta, w^*)}{\partial \theta} + \frac{\partial h(r, \theta, w^*)}{\partial w} \bigg|_{w^*} \frac{dw^*}{d\theta},$$

$$\frac{\partial H_s(r, \theta, w^*)}{\partial \theta} = \frac{\partial h(r, \theta, w^*)}{\partial \theta}.$$

This is very useful to simplify the derivations when utilizing the gMAM algorithm to explore the transition paths.

3. PROOF OF THEOREM 2.4

We will mainly follow the framework as in [19, 15] to make the proof. To prove the LDP for the family $(z_n, \nu_n)$ in the space $(D^d \times M, \rho^1 \times \rho^2)$, we apply the Dawson-Gärtner theorem (see [3]). The LDP in $(D^d \times M, \rho^1 \times \rho^2)$ can be implied by the LDPs in the spaces $(D^d[0, T] \times M[0, T], \rho^1_T \times \rho^2_T)$. The definition of the LDP in $(D^d[0, T] \times M[0, T], \rho^1_T \times \rho^2_T)$ is given in terms of Theorem 2.4 with obvious modifications, and we denote the finite time rate functionals as $I^T_s(r, \nu) = I^T_{s}(r, \nu) + I^T_f(r, \nu)$. From Dawson-Gärtner theorem, the LDP for $(z_n, \nu_n)$ in $(D^d \times M, \rho^1 \times \rho^2)$ is given as

$$\sup_{T > 0} I^T_s(r, \nu).$$

Hence only the LDP in $(D^d[0, T] \times M[0, T], \rho^1_T \times \rho^2_T)$ has to be checked for any $T > 0$. For notational ease, we will omit the superscript $T$ in the finite-time rate functionals in later text. This will not bring confusion since we only consider the LDP in the finite interval $[0, T]$ later on.

First we prove the upper bound and then the lower bound.

3.1. Upper Bound. The proof of upper bound (2.1) is standard in some sense. It is difficult to estimate the probability of $(z_n, \nu_n) \in F$ directly. We proceed with the following steps. Firstly, we approximate $(z_n, \nu_n)$ by $(\tilde{z}_n, \tilde{\nu}_n)$, where $\tilde{z}_n$ is an absolutely continuous path and $\tilde{\nu}_n(dt, \cdot)$ is absolutely continuous with respect to $dt$. Secondly, for a given compact set, we can get an upper bound for $(\tilde{z}_n, \tilde{\nu}_n)$. Thirdly, we prove that after excluding a set of exponentially small probability, $\tilde{z}_n$ and $\tilde{\nu}_n$ stay in compact sets, which means that $\tilde{z}_n$ and $\tilde{\nu}_n$ are exponentially tight sequence.
And finally, we get the desired result by combing the previous steps with further estimates.

To construct the approximation, we subdivide the time interval $[0, T]$ into $n$ pieces with nodes $t^n_j = T j / n$, $j = 0, 1, \ldots, n$. Define the piecewise linear interpolation $\tilde{z}_n(t)$ of $z_n(t)$ as

$$\tilde{z}_n(t) = (1 - \gamma_j(t))z_n(t^n_j) + \gamma_j(t)z_n(t^n_{j+1}), \quad t \in [t^n_j, t^n_{j+1}],$$

where $\gamma_j(t) = (t - t_j)n/T$ if $T \in [0, 1]$. Denote $F_n(t,i) = v_n([0, t], i)$ and define

$$\tilde{F}_n(t, i) = (1 - \gamma_j(t))F_n(t^n_j, i) + \gamma_j(t)F_n(t^n_{j+1}, i), \quad t \in [t^n_j, t^n_{j+1}].$$

Then we construct the new occupation measure $\tilde{\nu}_n(dt, i)$ with density

$$n_{\tilde{\nu}_n}(t, i) = \frac{d}{dt} \tilde{F}_n(t, i) = (F_n(t^n_{j+1}, i) - F_n(t^n_j, i))\frac{n}{T}, \quad t \in [t^n_j, t^n_{j+1}].$$

It is obvious that $\tilde{\nu}_n \in \mathcal{M}_{ac}$.

We have that $(\tilde{z}_n, \tilde{\nu}_n)$ is exponentially equivalent to $(z_n, \nu_n)$.

**Lemma 3.1.** For each $\delta > 0$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(\rho(1)(z_n, \tilde{z}_n) > \delta) = -\infty$$

and

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(\rho(2)(\nu_n, \tilde{\nu}_n) > \delta) = -\infty.$$

The proof of Lemma 3.1 is left in the Appendix.

For given compact sets, the following quasi-LDP upper bound for $(\tilde{z}_n, \tilde{\nu}_n)$ holds.

**Lemma 3.2.** Fix step functions $\theta(t) \in \mathbb{R}^d$ and $\alpha(t) \in \mathbb{R}^D$. For each $\delta > 0$ and given compact sets $\mathcal{K} \in \mathbb{D}^d[0, T]$ and $\mathcal{S} \in \mathbb{M}(0, T)$, we have

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}((\tilde{z}_n, \tilde{\nu}_n) \in \mathcal{K} \times \mathcal{S}) \leq -\inf_{(r, v) \in \mathcal{K} \times \mathcal{S}} \left( I^\delta_s(r, v, \theta) + I_f(r, v, \alpha) \right),$$

where

$$I^\delta_s(r, v, \theta) = \left\{ \int_0^T L^\delta_s(r(t), \dot{r}, n_v(t, \cdot), \theta(t)) dt, \quad dr(t) = \dot{r}(t) dt, v(dt, \cdot) = n_v(t, \cdot) dt, \right. \quad \text{otherwise,}$$

$$L^\delta_s(r(t), \dot{r}, n_v(t, \cdot), \theta(t)) = \langle \dot{r}(t), \theta(t) \rangle - H^\delta_s(r(t), \theta(t), n_v(t, \cdot)), \quad H^\delta_s(r(t), \theta(t), n_v(t, \cdot)) = \sup_{|x-r(t)| \leq \delta} H_s(x, \theta(t), n_v(t, \cdot)),$$

and

$$I_f(r, v, \alpha) = \left\{ \int_0^T S(r(t), n_v(t, \cdot), \alpha(t)) dt, \quad v(dt, \cdot) = n_v(t, \cdot) dt, \right. \quad \text{otherwise,}$$

$$S(r(t), n_v(t, \cdot), \alpha(t)) = \sup_{|x-r(t)| \leq \delta} S(x, \alpha(t), n_v(t, \cdot)).$$
Combining this with (3.1), we get

\[ S(r(t), n_v(t, \cdot), \alpha) = -\sum_{i,j=1}^D n_v(t, i)q_{ij}(r(t)) \left( e^{\langle \alpha(t), e_j \rangle} - 1 \right). \]

**Proof.** It is obvious that we only need to consider absolutely continuous functions \( r(t) \) and occupation measures \( v \in \mathcal{M}_{\text{ac}} \). For any \( r(t) \) and \( v \), define the sum

\[ J_n(r, \theta, v, \alpha) = \sum_{j=0}^{n-1} \left( \langle r(t^n_j) - r(t^n_{j+1}) \rangle, \theta(t^n_j) \right) \]

\[- \int_{t^n_j}^{t^n_{j+1}} H_{\bar{s}} \left( \langle r(t^n_j), \theta(t^n_j), n_v(t, \cdot) \rangle \right) dt + \int_{t^n_j}^{t^n_{j+1}} S(r(t), n_v(t, \cdot), \alpha(t^n_j)) dt.\]

By Corollary 5.8 in Appendix, we have

\[ \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E} \exp \left( n J_n(z_n, \theta, \tilde{v}_n, \alpha) \right) \leq 0. \] (3.1)

For \((z_n, \tilde{v}_n) \in \mathcal{K} \times \mathcal{S}\), it is obvious that

\[ J_n(z_n, \theta, \tilde{v}_n, \alpha) - \inf_{(r, v) \in \mathcal{K} \times \mathcal{S}} J_n(r, \theta, v, \alpha) \geq 0. \] (3.2)

So we have

\[ \exp \left( J_n(z_n, \theta, \tilde{v}_n, \alpha) - \inf_{(r, v) \in \mathcal{K} \times \mathcal{S}} J_n(r, \theta, v, \alpha) \right) \geq 1 \]

and

\[ \mathbb{P}((z_n, \tilde{v}_n) \in \mathcal{K} \times \mathcal{S}) \leq \mathbb{E} \exp \left( n \left[ J_n(z_n, \theta, \tilde{v}_n, \alpha) - \inf_{(r, v) \in \mathcal{K} \times \mathcal{S}} J_n(r, \theta, v, \alpha) \right] \right). \]

Combining this with (3.1), we get

\[ \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}((z_n, \tilde{v}_n) \in \mathcal{K} \times \mathcal{S}) \leq -\liminf_{n \to \infty} \left( \inf_{(r, v) \in \mathcal{K} \times \mathcal{S}} J_n(r, \theta, v, \alpha) \right). \] (3.3)

We now represent the sum on the right-hand side of (3.3) as an integral. Since \( \mathcal{K} \) is compact, the absolutely continuous functions \( r \in \mathcal{K} \) are thus bounded. Let \( C_1 \) be a compact set in \( \mathbb{R}^d \) such that

\[ \{x : x = r(t) \text{ for some } r \in \mathcal{K} \text{ and } t \in [0, T]\} \subset C_1. \]

For step function \( \theta \), let us investigate an interval in which \( \theta \) takes constant value \( \theta_0 \), say, the interval \([0, T]\) without loss of generality. Then

\[ \sum_{j=0}^{n-1} \chi_{(t^n_j \leq t)} \left\langle r(t^n_{j+1}) - r(t^n_j), \theta_0 \right\rangle = \int_0^T \langle \dot{r}, \theta_0 \rangle dt + e_n, \]
where the error \( e_n \) takes into account the fact that \( \tau \) may not match any of \( t_j^n \). It goes to zero uniformly for \( r \in \mathcal{K} \) when \( n \) goes to infinity from the bound

\[
|e_n| \leq \frac{T}{n} |\theta_0| \sup_{x \in C_t} |x|.
\]

Now \( H^\delta_s(x, \theta(t), n_v(t, \cdot)) \) is continuous in \( x, \theta \) and \( n_v \) from the smoothness of \( H_s \) on \( r, \theta \) and \( n_v \). What’s more, \( x, \theta \) and \( n_v \) live in bounded sets. So we have

\[
\left| H^\delta_s \left( r(t^n_j), \theta(t^n_j), n_v(t, \cdot) \right) - H^\delta_s \left( r(t), \theta(t^n_j), n_v(t, \cdot) \right) \right|, \quad t^n_j \leq t \leq t^n_{j+1}
\]

goes to zero uniformly in \( j \) for \( r \in \mathcal{K} \) and \( v \in \mathcal{S} \). Therefore,

\[
\sum_{j=0}^{n-1} \int_{t^n_j}^{t^n_{j+1}} H^\delta_s \left( r(t^n_j), \theta(t^n_j), n_v(t, \cdot) \right) \, dt
\]

\[
= \int_0^T \left( H^\delta_s \left( r(t), \theta(t^n_j), n_v(t, \cdot) \right) \right) \, dt + e_n,
\]

with \( e_n \) converging to zero uniformly in \( (r, v) \in \mathcal{K} \times \mathcal{S} \).

With the same manner we can estimate for the part \( S(r(t), n_v(t, \cdot), \alpha(t^n_j)) \) and repeat the argument on the finite number of intervals on which \( \theta \) and \( \alpha \) are constants. Thanks to the uniformity in \( (r, v) \in \mathcal{K} \times \mathcal{S} \), we obtain

\[
\liminf_{n \to \infty} \left( \inf_{(r, v) \in \mathcal{K} \times \mathcal{S}} \int_{n} (r, \theta, v, \alpha) \right) = \inf_{(r, v) \in \mathcal{K} \times \mathcal{S}} (I^\delta_s(r, v, \theta) + I_f(r, v, \alpha)).
\]

Together with (3.3), the proof is completed.

Next we show the exponential tightness of the sequence \( (\tilde{z}_n, \tilde{v}_n) \). Define the modulus of continuity of a function \( f \) as

\[
V_0(f) = \sup \{|f(t) - f(s)| : 0 \leq s \leq t \leq T, |t - s| < \delta\}
\]

and the set

\[
\mathcal{K}(M) = \bigcap_{m=M}^{\infty} \left\{ z \in C[0, T] : z(0) = z^0, V_{2^{-m}}(z) \leq \frac{1}{\log m} \right\}. \tag{3.4}
\]

**Lemma 3.3** (Exponential tightness for \( \tilde{z}_n \)). For each \( B < \infty \), there is a compact set \( \mathcal{K} \subset \mathbb{D}^T \) such that

\[
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(\tilde{z}_n \notin \mathcal{K}) \leq -B.
\]

**Proof.** It is not difficult to see that the set \( \mathcal{K}(M) \) is closed and the functions in \( \mathcal{K}(M) \) are equicontinuous. Thus \( \mathcal{K}(M) \) is compact by the Arzelà-Ascoli Theorem. If \( 2^{-m} < T/n \), we have

\[
V_{2^{-m-1}}(\tilde{z}_n) = \frac{1}{2} V_{2^{-m}}(\tilde{z}_n).
\]
since $\tilde{z}_n$ is piecewise linear. Therefore, to check whether $\tilde{z}_n$ is in $K(M)$, we only need to consider a finite intersection, for values of $m$ up to

$$M(n) = \max \left\{ M, \left\lfloor \log(n/T) \right\rfloor \right\}.$$ 

Using Corollary 5.6 in Appendix, we have for any $n$ with $M(n) > M$,

$$\mathbb{P}(\tilde{z}_n \notin K(M)) \leq \sum_{m=M}^{M(n)} \mathbb{P} \left( V_{2^{-m}}(\tilde{z}_n) > \frac{1}{\log m} \right) \leq \sum_{m=M}^{M(n)} \sum_{j=0}^{n-1} \mathbb{P} \left( \sup_{0 \leq t \leq 2^{-m}} |z_n(t^n_j + t) - z_n(t^n_j)| > \frac{1}{\log M} \right) \leq nM(n) \cdot \exp \left( -n \frac{c_1}{\log M} \log \left( \frac{2^M c_2}{\log M} \right) \right)$$

for positive constants $c_1$ and $c_2$. Thus

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(\tilde{z}_n \notin K(M)) \leq -c \frac{M}{\log M}$$

for some positive constant $c$ when $M \gg 1$. \qed

**Lemma 3.4.** The measure space $\mathbb{M}[0, T]$ is compact.

**Proof.** Since $[0, T] \times \{1, 2, \ldots, D\}$ is compact, $\mathbb{M}[0, T]$ is tight. By Prohorov’s theorem, $\mathbb{M}[0, T]$ is relatively compact, thus is evidently compact. \qed

The straightforward consequence of Lemma 3.4 is that $\tilde{v}_n$ is also exponentially tight.

Define the quasi-rate-functional

$$I^\delta_s(r, \nu) = \left\{ \int_0^T L_s^\delta(r(t), \dot{r}(t), n_\nu(t, \cdot))dt, \quad dr(t) = \dot{r}(t)dt, \quad \nu(dt, \cdot) = n_\nu(t, \cdot)dt, \right.$$ otherwise.

where

$$L_s^\delta(r(t), \dot{r}(t), n_\nu(t, \cdot)) = \sup_{\theta \in \mathbb{R}^d} L_s^\delta(r(t), \dot{r}(t), n_\nu(t, \cdot), \theta).$$

We have the following approximation lemma.

**Lemma 3.5.** For any $\epsilon > 0$, the absolutely continuous functions $r \in \mathcal{K}$ and $\nu \in \mathbb{M}_{ac}[0, T]$, there exist neighborhoods $N_r \subset \mathbb{R}^d[0, T]$ and $N_\nu \subset \mathbb{M}[0, T]$ for $r$ and $\nu$, step functions $\theta_{r_\nu} \subset \mathbb{R}^d$ and $\alpha_{r_\nu} \subset \mathbb{R}^D$, such that for any $p \in N_r$ and $\mu \in N_\nu$ which are both absolutely continuous, we have

$$I^\delta_s(p, \mu, \theta_{r_\nu}) + I_f(p, \mu, \alpha_{r_\nu}) \geq I^\delta_s(r, \nu) + I_f(r, \nu) - \epsilon.$$
**Lemma 3.6.** Given a compact set $\mathcal{K} \in \mathbb{D}^d[0, T]$. For any pair $(r, v) \in \mathcal{K} \times \mathbb{M}[0, T]$ and $M_0 > 0$, if either $r$ or $v$ is not absolutely continuous, there exist neighborhood $N_{r,v} \in \mathbb{D}^d[0, T] \times \mathbb{M}[0, T]$ and step functions $\theta_{r,v} \in \mathbb{R}^d$ and $\alpha_{r,v} \in \mathbb{R}^d$, such that for any $(p, \mu) \in N_{r,v}$, we have

$$I_\delta^s(p, \mu, \theta_{r,v}) + I_f(p, \mu, \alpha_{r,v}) \geq M_0.$$

The Lemmas 3.5 and 3.6 are direct consequences of Lemmas 5.11 and 5.12 in the appendix.

Denote the product metric $\rho^{(1)} \times \rho^{(2)}$ on $\mathbb{D}^d[0, T] \times \mathbb{M}[0, T]$ as $d(\cdot, \cdot)$ as well and define the closed sets

$$\Phi(K) = \{(r, v) \in \mathbb{D}^d[0, T] \times \mathbb{M}[0, T] : I_s(r, v) + I_f(r, v) \leq K\}$$

and

$$\Phi^\delta(K) = \{(r, v) \in \mathbb{D}^d[0, T] \times \mathbb{M}[0, T] : I_\delta^s(r, v) + I_f(r, v) \leq K\}.$$

**Proposition 3.7.** For each $K > 0$, $\delta > 0$ and $\epsilon > 0$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}((\hat{z}_n, \hat{\nu}_n) \in \Phi^\delta(K) > \epsilon) \leq -(K - \epsilon).$$

**Proof.** From the exponential tightness, we can find a compact set $\mathcal{K}^N$ for each $N > 0$ such that

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(\hat{z}_n \in \mathcal{K}^N) \leq -N.$$

Define the set

$$\mathcal{K}^{N,\epsilon} = \{(r, v) \in \mathbb{D}^d[0, T] \times \mathbb{M}[0, T] : d((r, v), \Phi^\delta(K) > \epsilon) \leq (\mathcal{K}^N \times \mathbb{M}[0, T])\}.$$

For any $(r, v) \in \mathcal{K}^{N,\epsilon}$, we can find the neighborhood $N_{r,v}$ either satisfying Lemma 3.5 if $r$ and $v$ are both absolutely continuous, or satisfying Lemma 3.6 if one of them are not absolutely continuous. This forms a covering of $\mathcal{K}^{N,\epsilon}$. By compactness, we can choose a finite subcover \{N_{r_i,v_j}\}_{i,j} for $\mathcal{K}^{N,\epsilon}$. Define

$$\mathcal{K}_{ij} = \overline{N_{r_i,v_j} \cap \mathcal{K}^{N,\epsilon}}.$$

Applying Lemma 3.2, Lemma 3.5, Lemma 3.6 and letting $M_0$ in Lemma 3.6 larger than $K$, we have for any $i, j$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}((\hat{z}_n, \hat{\nu}_n) \in \mathcal{K}_{ij}) \leq -(K - \epsilon).$$
Then we have
\[
\limsup_{n \to \infty} \frac{1}{n} \log P\left( d\left( (\tilde{z}_n, \tilde{\nu}_n), \Phi^\delta(K) \right) > \epsilon \right)
\]
\leq \limsup_{n \to \infty} \frac{1}{n} \log \left[ P(\tilde{z}_n \notin \mathcal{K}^N) + \sum_{(i,j)} P((\tilde{z}_n, \tilde{\nu}_n) \in \mathcal{K}_{ij}) \right]
\leq - \min\{N, K - \epsilon\}

Choosing \(N\) large enough, we complete the proof. \(\square\)

We are now ready to establish the upper bound.

**Lemma 3.8.** Given \(K > 0\) and \(\epsilon > 0\), there exist \(\delta > 0\) such that
\[
\Phi^\delta(K - \epsilon) \subset \{ (r, \nu) : d((r, \nu), \Phi(K)) \leq \epsilon \}.
\]

The proof of Lemma 3.8 can be referred to the proof of Lemma 5.48 in [19].

**Theorem 3.9.** For each closed set \(F \subset \mathbb{D}^{d}([0, T] \times \mathbb{M}([0, T])

\[
\limsup_{n \to \infty} \frac{1}{n} \log P((z_n, \nu_n) \in F) \leq - \inf_{(r, \nu) \in F} (I_s(r, \nu) + I_f(r, \nu)).
\]

**Proof.** Suppose \(\inf_{(r, \nu) \in F} (I_s(r, \nu) + I_f(r, \nu)) = K < \infty\). Since \(F\) and \(\Phi(K - \epsilon)\) are both closed sets, we assume the distance between them is \(\eta_0 > 0\). For any \(\eta \leq \eta_0\),
\[
P((z_n, \nu_n) \in F) \leq P\left( d\left( (\tilde{z}_n, \tilde{\nu}_n), F \right) \leq \frac{\eta}{2} \right) + P\left( d\left( (\tilde{z}_n, \tilde{\nu}_n), (z_n, \nu_n) \right) \geq \frac{\eta}{2} \right)
\]
\[
\leq P\left( d\left( (\tilde{z}_n, \tilde{\nu}_n), \Phi(K - \epsilon) \right) \geq \frac{\eta}{2} \right) + P\left( \rho^{(1)}((\tilde{z}_n, z_n) \geq \frac{\eta}{4} \right)
\]
\[
+ P\left( \rho^{(2)}(\tilde{\nu}_n, \nu_n) \geq \frac{\eta}{4} \right). \tag{3.5}
\]

By Lemma 3.8, we can choose \(\delta\) and \(\eta\) small enough such that
\[
d\left( (\tilde{z}_n, \tilde{\nu}_n), \Phi(K - \epsilon) \right) \geq \frac{\eta}{2} \text{ implies } d\left( (\tilde{z}_n, \tilde{\nu}_n), \Phi^\delta(K - \epsilon - \eta/4) \right) \geq \frac{\eta}{4}.
\]

From Proposition 3.7 we have
\[
\limsup_{n \to \infty} \frac{1}{n} \log P\left( d\left( (\tilde{z}_n, \tilde{\nu}_n), \Phi(K - \epsilon) \right) \geq \frac{\eta}{2} \right)
\leq \limsup_{n \to \infty} \frac{1}{n} \log P\left( d\left( (\tilde{z}_n, \tilde{\nu}_n), \Phi^\delta(K - \epsilon - \eta/4) \right) \geq \frac{\eta}{4} \right)
\leq - (K - \epsilon - \eta/2). \tag{3.6}
\]

Combining (3.5), (3.6) and Lemma 3.1 for \(\delta = \eta/4\), we obtain
\[
\limsup_{n \to \infty} \frac{1}{n} \log P((z_n, \nu_n) \in F) \leq - (K - \epsilon - \eta/2).
\]
The case for
\[
\inf_{(r, v) \in F} \left( I_s(r, v) + I_f(r, v) \right) = \infty
\]
can be established similarly by choosing \( K \) arbitrarily large. \( \square \)

### 3.2. Lower bound.

The proof of the lower bound is based on the change of measure formula. From [3], it suffices to prove that for any \((r, v) \in \mathbb{D}^d[0, T] \times \mathbb{M}[0, T]\) and arbitrarily small \( \epsilon > 0 \) we have

\[
\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P} (z_n \in N_\epsilon(r), v_n \in N_\epsilon(v)) \geq - \left( I_s(r, v) + I_f(r, v) \right). \tag{3.7}
\]

For given \( r \in \mathbb{D}^d[0, T] \) and \( v \in \mathbb{M}[0, T] \), if they are not absolutely continuous with respect to \( dt, I_s(r, v) + I_f(r, v) = \infty \), thus nothing needs to be proved. Below we will exclude this case. For convenience, we further assume that \( n_\nu (t, i) \) is continuous in \( t \), and the case that \( n_\nu (t, i) \) is not continuous will be discussed in Theorem 3.16 in this section. To prove the lower bound, we perform the following steps. Firstly, we approximate \( r(t) \) by a piecewise linear path \( y \), and the occupation measure \( v \) by \( \pi \) with \( n_\pi (t, \cdot) \) piecewise constant in \( t \). Secondly, we construct new processes \( \tilde{z}_n \) and \( \tilde{\xi}_n \) with occupation measure \( \tilde{\nu}_n \) such that

\[
\mathbb{P} - \lim_{n \to \infty} \rho^{(1)}(\tilde{z}_n, y) = 0, \quad \mathbb{P} - \lim_{n \to \infty} \rho^{(2)}(\tilde{\nu}_n, \pi) = 0. \tag{3.8}
\]

Moreover, we ask \( \tilde{z}_n \) and the jump rates of \( \tilde{\xi}_n \) satisfy the conditions required by Lemmas 3.10 and 3.11. Finally, based on the change of measure formula related to \((z_n, \xi_n)\) and \((\tilde{z}_n, \tilde{\xi}_n)\), we get the limit and the proof is then finished.

As promised in the above procedure, we approximate \( r \) by a path \( y \), and \( v \) by a occupation measure \( \pi \) in the first step. For a given \( f \), define \( \Delta = T / f \) and let \( t_m = m \Delta \). On each interval \([t_m, t_{m+1}]\), define \( \Delta r_m = r(t_{m+1}) - r(t_m) \). Take \( \mu_m = \{ \mu_{im}, i = 1, \cdots, k \} \) so as to satisfy

\[
\sum_{i=1}^d \mu_{im} u_i = \frac{\Delta r_m}{\Delta}.
\]

If \( \Delta r_m \) are in the positive cone generated by the \( \{ u_i \} \) for all \( m \), such a choice of \( \mu_{im} \) is possible. If at least one of \( \Delta r_m \) is not in the positive cone generated by the \( \{ u_i \} \), it is easy to check that for all \( v \in \mathbb{M} \), \( I_s(r, v) = +\infty \) (see the Remark of Lemma 5.21 in [19]) and nothing needs to be proved.

Now we construct the piecewise linear interpolation \( y(t) \) of \( r(t) \) such that in each time interval \([t_m, t_{m+1}]\)

\[
\frac{d}{dt} y(t) = \sum_{i=1}^d \mu_{im} u_i, \quad y(t_0) = r(t_0)
\]

and thus \( y(t_m) = r(t_m) \) for each \( m \).
Also, for given \( \nu \in \mathbb{M}[0, T] \) with \( n_{\nu}(t, i) \) exists and continuous in \( t \). We define a piecewise constant interpolation \( n_\pi(t, \cdot) \) of \( n_{\nu}(t, \cdot) \), i.e. for each \( i \in \{1, 2, \cdots, D\} \), we ask
\[
n_\pi(t, i) = n_{\nu}(t_m, i), \quad t \in [t_m, t_{m+1}).
\]
Then define \( \pi \) as the approximation of \( \nu \) by letting
\[
\pi(t, i) = n_\pi(t, i)dt.
\]
Since \( r \) is absolutely continuous and \( n_{\nu}(t, i) \) is continuous in \( t \), for each \( \epsilon > 0 \), we can choose \( J \) large enough such that
\[
\rho^{(1)}(y, r) < \epsilon/4, \quad \rho^{(2)}(\pi, \nu) < \epsilon/4.
\]
In the second step, we construct new processes \( \tilde{z}_n \) and \( \tilde{\xi}_n \) with occupation measure \( \bar{v}_n \) such that (3.8) holds. Define
\[
S_m = \{ \eta = (\eta_{ij})_{D \times D} : \eta_{ij} \geq 0 \text{ and } \eta_{ii} = 0; \sum_{i=1}^{D} n_\pi(t_m, i) \sum_{j=1}^{D} \eta_{ij}e_{ij} = 0 \}.
\]
and
\[
K_\eta = \{ \mu = (\mu_{im})_{S \times J} : \mu_{im} \geq 0, \sum_{i=1}^{S} \mu_{im}u_i = y(t_{m+1}) - y(t_m) \}.
\]
We also take the frequently used notation \( \lambda_i^\pi \) in later text as the expectation of \( \lambda_i \) with respect to the distribution \( n_\pi \)
\[
\lambda_i^\pi(y(s)) = \sum_{j=1}^{D} \lambda_i(y(s), j) n_\pi(s, j).
\]

**Lemma 3.10.** For any \( \epsilon > 0 \) and large enough \( J \), there exists a further subdivision of time interval \( [t_m, t_{m+1}] \) for each \( m \in \{0, 1, \cdots, J-1\} \) (i.e., \( t_m = t_{m0} < t_{m1} < \cdots < t_{mk_m} = t_{m+1} \)) and related matrices \( \eta^{mk} \in S_m \) (\( m = 0, 1, \cdots, J-1; k = 0, 1, \cdots, K_m \)), such that
\[
\sum_{m=0}^{J-1} \sum_{k=0}^{K_m-1} \int_{t_{mk}}^{t_{mk+1}} \sum_{i=1}^{D} n_\pi(t, i) \sum_{j=1}^{D} \left( \eta^{mk}_{ij} \log \frac{\eta^{mk}_{ij}}{q_{ij}(y(t))} + q_{ij}(y(t)) - \eta^{mk}_{ij} \right) dt \leq s_f(r, v) + \epsilon.
\]

The proof of Lemma 3.10 can be found in the Appendix.

The rate matrices \( \eta^{mk} \in S_m \) in Lemma 3.10 is not sufficient for the continued Lemmas, which require ergodicity property. So we make further approximations. Thanks to Lemma 8.61 of [19], for any jump rate matrix \( \eta^{mk} \in S_m \) and \( \epsilon > 0 \), there exists a rate matrix \( \tilde{\eta}^{mk} \) with an unique invariant measure \( \Gamma^{mk}(i) \) such that
\[
|n_\pi(t_m, i) - \Gamma^{mk}(i)| < \epsilon/4DT \quad \text{for all } i,
\]
\[ |n_{\pi}(t_m,i)\eta_{ij}^{mk} - \Gamma_{mk}(i)\tilde{\eta}_{ij}^{mk}| < \epsilon/2 \quad \text{for all } i \text{ and } j, \]
\[ \tilde{\eta}_{ij}^{mk} > 0 \quad \text{for all } i \neq j, \]
and
\[ \sum_{i=1}^{D} \Gamma_{mk}(i) \sum_{j=1}^{D} \left( \tilde{\eta}_{ij}^{mk} \log \frac{\tilde{\eta}_{ij}^{mk}}{q_{ij}(y(t))} + q_{ij}(y(t)) - \eta_{ij}^{mk} \right) \leq \sum_{i=1}^{D} n_{\pi}(t_m,i) \sum_{j=1}^{D} \left( \eta_{ij}^{mk} \log \frac{\eta_{ij}^{mk}}{q_{ij}(y(t))} + q_{ij}(y(t)) - \eta_{ij}^{mk} \right) + \epsilon. \quad (3.9) \]

For each \( m \) and \( i \in \{1, 2, \cdots, D\} \), define
\[ n_{\tilde{\pi}}(t_m,i) = \Gamma_{mk}(i), \quad t \in [t_{mk}, t_{m,k}+1) \]
and let \( \tilde{\pi}(dt,i) = n_{\tilde{\pi}}(t,i)dt \). So we have \( \rho^{(2)}(\tilde{\pi}, \pi) < \epsilon/4 \) and thus \( \rho^{(2)}(\tilde{\pi}, \nu) < \epsilon/2 \).

With similar idea as proving Lemma 3.10, we can show

**Lemma 3.11.** For any \( \epsilon > 0 \) and large enough \( J \), there exists \( \mu \in K_y \) such that
\[ \sum_{m=0}^{J-1} \int_{t_m}^{t_{m+1}} \sum_{i=1}^{S} \left( \lambda_{i}^{\tilde{\pi}}(y(t)) - \mu_{im} + \mu_{im} \log \frac{\mu_{im}}{\lambda_{i}^{\tilde{\pi}}(y(t))} \right) dt \leq I_s(r, \nu) + \epsilon. \]

With the constructed matrices \( \{\tilde{\eta}_{ij}^{mk}\} \), we define the process \( \tilde{\xi}_n \) with jump rate \( n\eta_{ij}(t) \) where \( \eta_{ij}(t) = \tilde{\eta}_{ij}^{mk}, \ t \in [t_{mk}, t_{m,k}+1) \). Similarly, we take \( \mu \) constructed from Lemma 3.11 and define \( \tilde{z}_n \) with jump rate
\[ n\mu_i(t) \frac{\lambda_i(\tilde{z}_n(t), \tilde{\xi}_n(t))}{\lambda_i^{\tilde{\pi}}(y(t))} \]
for its \( i \)th component, where \( \mu_i(t) = \mu_{im}, \ t \in [t_m, t_{m+1}) \).

We have the following convergence result for the constructed approximations for \( \tilde{\pi} \) and \( \pi \).

**Lemma 3.12.** Convergence of the approximation \( \tilde{\nu}_n \)
\[ \mathbb{P} - \lim_{n \to \infty} \rho^{(2)}(\tilde{\nu}_n, \tilde{\pi}) = 0. \]

**Lemma 3.13.** Convergence of the approximation \( \tilde{z}_n \)
\[ \mathbb{P} - \lim_{n \to \infty} \rho^{(1)}(\tilde{z}_n, \pi) = 0. \]

The proof of Lemmas 3.12, 3.13 will be given in the Appendix.

As we have finished the construction of \( \tilde{z}_n \) and \( \tilde{\xi}_n \), we now perform the change of measure. Denote \( Q_n \) and \( \tilde{Q}_n \) the distributions of \( (z_n(t), \xi_n(t))_{t \leq T} \)
and \((\tilde{z}_n(t), \tilde{\xi}_n(t))_{t \leq T}\), respectively. We have

\[
\frac{dQ_n}{dQ_n}(\tilde{z}_n(t), \tilde{\xi}_n(t)) = \exp \left\{ - \int_0^T \frac{1}{n} \sum_{i=1}^D \left( \lambda_i(t) \tilde{z}_n(t) - \mu_i(t) \frac{\lambda_i}{\lambda_i}(y(t)) \right) dt \right. \\
- \left. \int_0^T \sum_i \log \frac{\mu_i(t^-)}{\lambda_i(y(t^-))} dY_i^i - \int_0^T \frac{1}{n} \sum_{i,j=1}^D \left( q_{ij}(\tilde{z}_n(t)) - \eta_{ij}(t^-) \right) dt \right. \\
- \left. \int_0^T \sum_{i,j} \log \frac{\eta_{ij}(t^-)}{q_{ij}(\tilde{z}_n(t^-))} dM_{ij}^i \right\},
\]

(3.10)

where \(Y_i^i\) is the counting process induced by \(\tilde{z}_n(t)\) that will increase by one each time when a jump occurs in the \(u_i\) direction and \(M_{ij}^i\) is the counting process induced by \(\tilde{\xi}_n(t)\) that will increase by one each time when a jump occurs from state \(i\) to state \(j\).

The next lemma shows that in the limit \(n \to \infty\) the right hand side becomes simple.

**Lemma 3.14.**

\[
\lim_{n \to \infty} \mathbb{E} \left[ \int_0^T \frac{\lambda_i(\tilde{z}_n(t), \tilde{\xi}_n(t))}{\lambda_i(y(t))} dt \right] = 1.
\]

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \int_0^T \sum_i \log \frac{\mu_i(t^-)}{\lambda_i(y(t^-))} dY_i^i \right] = \sum_{m=0}^{J-1} \sum_{t_m+1}^T \sum_{i=1}^D \mu_{im} \log \frac{\mu_{im}}{\lambda_i(y(t))} dt.
\]

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \int_0^T \sum_{i,j} \log \frac{\eta_{ij}(t^-)}{q_{ij}(\tilde{z}_n(t^-))} dM_{ij}^i \right] = \sum_{m=0}^{J-1} \sum_{k=1}^{K_m} \sum_{t_{mk}+1}^{t_{m+1}} \sum_{i=1}^D n_{i}(t) \sum_{j=1}^D \eta_{ij}^{mk} \log \frac{\tilde{\eta}_{ij}^{mk}}{q_{ij}(y(t))} dt.
\]

The proof of Lemma 3.14 can be obtained following the same approach as Lemma 7.6 in [19].

**Lemma 3.15.** For given \(r \in \mathbb{D}^d[0, T]\) and \(\nu \in \mathbb{M}[0, T]\), assume that \(r\) and \(\nu\) are absolutely continuous with respect to \(dt\) and \(n_v(t, \cdot)\) is continuous in \(t\). Then for arbitrarily small \(\epsilon > 0\) we have

\[
\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}(z_n \in N_{\epsilon}(r), \nu_n \in N_{\epsilon}(\nu)) \geq - \left( I_s(r, \nu) + I_f(r, \nu) \right).
\]
Thus, according to (3.10) and Lemma 3.14, we have

\[ \liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_\epsilon(r), \nu_n \in N_\epsilon(v)) \geq \liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_{\epsilon/2}(y), \nu_n \in N_{\epsilon/2}(\tilde{\nu})) \]

By Jensen’s inequality, for any \( \nu \) are absolutely continuous with respect to \( d \mu \), we have

\[ \rho(\nu, \tilde{\nu}) \geq \rho(\nu, \tilde{\nu}) \geq \rho(\nu, \tilde{\nu}) = -\rho(\nu, \tilde{\nu}) \]

which is the proof.

Combining Eqs. (3.11), (3.12), (3.9), Lemma 3.10 and Lemma 3.11, we finish the proof.

We can construct a sequence of measures \( \nu^{(k)} \) such for any \( k \), \( \nu^{(k)} \) is continuous in \( t \) and \( \rho^{(2)}(\nu, \nu^{(k)}) \to 0 \). From Lemma 5.10, \( I_s(r, \nu) + I_f(r, \nu) \) is lower semi-continuous in \( \nu \). Thus, we can choose \( k_0 \) large enough such that for any \( \delta > 0 \) and \( \epsilon > 0 \),

\[ I_s(r, \nu^{(k_0)}) + I_f(r, \nu^{(k_0)}) \geq I_s(r, \nu) + I_f(r, \nu) - \delta \]

and

\[ \rho(\nu, \nu^{(k_0)}) < \epsilon/2. \]

Proof. By Jensen’s inequality, for any \( \epsilon > 0 \)

\[ \liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_\epsilon(r), \nu_n \in N_\epsilon(v)) \geq \liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_{\epsilon/2}(y), \nu_n \in N_{\epsilon/2}(\tilde{\nu})) \]

By Lemmas 3.12 and 3.13, we know that

\[ \liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_{\epsilon/2}(y), \nu_n \in N_{\epsilon/2}(\tilde{\nu})) = 0. \]

Thus, according to (3.10) and Lemma 3.14, we have

\[ \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{E}_{\tilde{\nu}_n} \left[ \chi_{z_n \in N_{\epsilon/2}(y), \nu_n \in N_{\epsilon/2}(\tilde{\nu})} \right] \]

Combining Eqs. (3.11), (3.12), (3.9), Lemma 3.10 and Lemma 3.11, we finish the proof.

In the final theorem, we remove the continuity assumption on \( n, r \) to get the desired lower bound estimate.

Theorem 3.16. For given \( r \in \mathbb{R}^d \) and \( \nu \in \mathcal{M}[0, T] \), assume that \( r \) and \( \nu \) are absolutely continuous with respect to \( d \mu \), we have

\[ \liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_r(r), \nu_n \in N_r(v)) \geq -\left( I_s(r, \nu) + I_f(r, \nu) \right) \]

Proof. We can construct a sequence of measures \( \nu^{(k)} \) such for any \( k \), \( n^{(k)} \) is continuous in \( t \) and \( \rho^{(2)}(\nu, \nu^{(k)}) \to 0 \). From Lemma 5.10, \( I_s(r, \nu) + I_f(r, \nu) \) is lower semi-continuous in \( \nu \). Thus, we can choose \( k_0 \) large enough such that for any \( \delta > 0 \) and \( \epsilon > 0 \),

\[ I_s(r, \nu^{(k_0)}) + I_f(r, \nu^{(k_0)}) \geq I_s(r, \nu) + I_f(r, \nu) - \delta \]

and

\[ \rho(\nu, \nu^{(k_0)}) < \epsilon/2. \]
Thanks to Lemma 3.15, we have
\[
\liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_\epsilon(r), \nu_n \in N_\epsilon(\nu)) \geq \liminf_{n \to \infty} \frac{1}{n} \log P\left(z_n \in N_\epsilon(r), \nu_n \in N_{\epsilon/2}(\nu^{(k_0)})\right) \\
\geq - \left(I_s(r, \nu^{(k_0)}) + I_f(r, \nu^{(k_0)})\right) \\
\geq - (I_s(r, \nu) + I_f(r, \nu)) - \delta.
\]
The proof is completed. □

3.3. Goodness of the rate functional. The rate functional \(I_s(r, \nu) + I_f(r, \nu)\) is lower semicontinuous by Lemma 5.10. The goodness of the rate functional is a direct consequence of the following result.

Lemma 3.17. For any \(K > 0\), the level set \(\Phi_K = \{(r, \nu) \in D^d[0, T] \times \mathbb{M}[0, T] : I_s(r, \nu) + I_f(r, \nu) \leq K\}\) is a compact set.

Proof. By Lemma 3.4, \(\mathbb{M}[0, T]\) is a compact set. By Lemma 5.9, the functions \(r \in \Phi_K\) are equicontinuous. Combining with the fact that \(r(0) = z^0\), we have that \(\Phi_K\) is pre-compact. By Lemma 5.10, \(I_s(r, \nu) + I_f(r, \nu)\) is lower semicontinuous. Consequently, \(\Phi_K\) is closed and thus compact. □

4. PROOF OF THEOREM 2.2

Now we prove Theorem 2.2 under the consideration \((x, y) \in G = (\mathbb{R}^+)^d\) instead of the whole space. The main clue of the proof is the same as the proof of Theorem 2.4 except some technicalities to understand the behavior of jumps near the boundary of \(G\). We will only focus on the key parts which is different from the proof of Theorem 2.4.

The difficulty in the proof of lower bound is that we can not use the change of measure formula directly, since some of the jump rates may diminish on the boundary. Mainly following [20], We overcome this issue by carefully analyzing the boundary behavior of the dynamics.

Let a \(d\)-dimensional unit vector \(\boldsymbol{v} := (1, 1, \ldots, 1)/\sqrt{d}\) and define the shifting \(r_\delta(t) = r(t) + \delta \boldsymbol{v}\). With the same approach in proving Lemma 5.1 in [20], we can show that
\[
\limsup_{\delta \to 0^+} \left(I_s(r_\delta, \nu) + I_f(r_\delta, \nu)\right) \leq I_s(r, \nu) + I_f(r, \nu). \tag{4.1}
\]
Next we will prove
\[
\liminf_{n \to \infty} \frac{1}{n} \log P(z_n \in N_\delta(r), \nu_n \in N_\delta(\nu)) \geq - \left(I_s(r, \nu) + I_f(r, \nu)\right).
\]
Denote by $\tilde{\eta}$ the modulus of continuity of $r$ and set $\eta(a) = \max\{\tilde{\eta}(a), a\}$ so that $\eta^{-1}(a) \leq a$. Now, fix $\delta$ and set $t_\delta = \eta^{-1}(\delta/3)$. Then, $t_\delta \leq \delta/3$ and for $t \leq t_\delta$,
\[
\sup_{0 \leq t \leq t_\delta} |r(0) + t \cdot v - r(t)| \leq t_\delta \cdot |v| + \eta(t_\delta) \leq 2\delta/3.
\]
Therefore, for $0 < \alpha < 1/6$,
\[
\mathbb{P}(z_n \in N_{\delta}(r), \nu_n \in N_{\delta}(v)) \geq \mathbb{P}\left(|z_n(t) - r(0) - t \cdot v| \leq \alpha \delta \text{ on } t \in [0, t_\delta], z_n \in N_{\delta}(r); \nu_n \in N_{\delta}(v)\right),
\]
where $N_{\delta}(r; [t_\delta, T])$ is the $\delta$-neighborhood of $r$ restricted on $t \in [t_\delta, T]$. Now, on this time interval
\[
\sup_{t_\delta \leq t \leq T} |r(t) - r_{t_\delta}(t)| \leq \delta/3
\]
and, moreover, $d(r_{t_\delta}(t), \partial G) \geq t_\delta / \sqrt{d}$. Therefore, for any function $u$ on $t \in [t_\delta, T]$, $||u - r_{t_\delta}|| \leq t_\delta / 2\sqrt{d}$ implies that $||u - r|| \leq 5\delta / 6$ and $d(r_{t_\delta}(t), \partial G) \geq t_\delta / 2\sqrt{d}$. Now define $A_\delta$ the $\alpha\delta$-neighborhood of $r_0 + t_\delta v$, i.e. $A_\delta := B_{\alpha\delta}(r_0 + t_\delta v)$ and let $r_{t_\delta}'$ be the shift of $r_{t_\delta}$ such that $r_{t_\delta}'(t_\delta) = y$. Then,
\[
\mathbb{P}(z_n \in N_{\delta}(r), \nu_n \in N_{\delta}(v)) \geq \mathbb{P}\left(|z_n(t) - r(0) - t \cdot v| \leq \alpha \delta \text{ on } t \in [0, t_\delta], \nu_n \in N_{\epsilon}(v; [0, t_\delta])\right) \times \inf_{y \in A_\delta} \mathbb{P}_{y}\left(z_n \in N_{t_\delta}(r_{t_\delta}'; [t_\delta, T]); \nu_n \in N_{\epsilon}(v; [t_\delta, T])\right).
\]
The first term satisfies a large deviation lower bound
\[
\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left(|z_n(t) - r(0) - t \cdot v| \leq \alpha \delta \text{ on } t \in [0, t_\delta], \nu_n \in N_{\epsilon}(v; [0, t_\delta])\right) \geq -C t_\delta
\]
by estimating the probability of a specific path $z_n$ lying in the $\alpha\delta$-neighborhood of the curve $r(0) + tv$. Because the paths in $N_{t_\delta/2\sqrt{d}}(r_{t_\delta}'; [t_\delta, T])$ are bounded away from the boundary uniformly for $y \in A_\delta$, by Theorem 3.16, we have
\[
\liminf_{n \to \infty} \frac{1}{n} \log \inf_{y \in A_\delta} \mathbb{P}_{y}\left(z_n \in N_{t_\delta/2\sqrt{d}}(r_{t_\delta}'; [t_\delta, T]); \nu_n \in N_{\epsilon}(v; [t_\delta, T])\right) \geq -\left(I_s^{[t_\delta, T]}(r_{t_\delta}, v) + I_f^{[t_\delta, T]}(r_{t_\delta}, v)\right)
\]
where $I_s^{[t_\delta, T]}(r_{t_\delta}, v)$ and $I_f^{[t_\delta, T]}(r_{t_\delta}, v)$ are rate functionals defined on the integration interval $[t_\delta, T]$. According to (4.1), (4.2) and (4.3), we proved the lower bound.
Next let us consider the upper bound. At first we note that since the rates \( \lambda_i(z, j) \) satisfies the linear growth condition

\[
\lambda_i(z, j) \leq K(1 + |z|),
\]

it is easy to show that

\[
\lim_{K \to \infty} \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}( \sup_{0 \leq t \leq T} |z_n(t)| > K) = -\infty
\]

by simple moment estimates and Doob’s martingale inequality. Consequently, it suffices to prove the large deviation estimates for bounded sets and we can assume \( \lambda_i(z, j) \) are bounded.

We only need to recheck Lemma 3.8 and Lemma 5.12, since the other lemmas in upper bound estimates can be verified easily under the assumption that \( \lambda_i(z, j) \) are bounded. Thanks to Corollary 4.2 and Lemma 4.6 in [20], we can obtain that Lemma 3.8 and Lemma 5.12 are also correct under Assumption 2.1. Thus the upper bound is also established.

The goodness of the rate functional trivially holds under Assumption 2.1. So we complete the proof of Theorem 2.2.

5. Appendix

**Lemma 5.1.** Let \( \{f_\alpha\} \) be a collection of lower semi-continuous functions on a metric space. Then the function \( f \) define by \( f(x) = \sup_\alpha f_\alpha(x) \) is lower semicontinuous.

**Lemma 5.2.** Let \( \{f_\alpha\} \) be a collection of convex functions on a metric space. Then the function \( f \) define by \( f(x) = \sup_\alpha f_\alpha(x) \) is convex.

**Lemma 5.3.** Let \( K(x, y) \) be a real-valued function, continuous in \((x, y)\) on metric space \( \chi_1 \times \chi_2 \), convex in \( x \) for each \( y \), and concave in \( y \) for each \( x \). Let two closed convex sets \( U \) and \( V \) be given, at least one of which is bounded. Then

\[
\inf_{x \in U} \sup_{y \in V} K(x, y) = \sup_{y \in V} \inf_{x \in U} K(x, y).
\]

The proof of Lemma 5.3 may be referred to [18].

**Lemma 5.4.** Let \( z \in \mathbb{R}^d \) be a random vector. Suppose there exist numbers \( a \) and \( \delta \) such that for each \( \theta \in \mathbb{R}^d \) with \( |\theta| = 1 \),

\[
\mathbb{P}(|\langle \theta, z \rangle| \geq a) \leq \delta.
\]

Then

\[
\mathbb{P}(|z| \geq a \sqrt{d}) \leq \delta.
\]
The above lemma is directly borrowed from Lemma 5.7 in [19].
In later texts, we will take an abused notation $\xi_n(t) = e_i \in \mathbb{R}^D$ when $\xi_n(t) = i \in \mathbb{Z}_D$. This will not bring confusion since $\xi_n(t)$ is considered as a multidimensional vector only when we take inner product with other vectors.

**Lemma 5.5.** There exists a function $K : \mathbb{R}^+ \to \mathbb{R}^+$ with
\[ \lim_{a \to \infty} K(a)/a = +\infty, \]
such that
\[ \mathbb{P} \left( \sup_{0 \leq t \leq T} |z_n(t) - z_n(0)| \geq a \right) \leq 2d \exp \left( -nT \frac{a}{T} \right). \] (5.1)

**Proof.** The inequality (5.1) holds trivially whenever $K(a/T) = 0$. It suffice to prove the lemma when $a$ is large. For $\theta \in \mathbb{R}^d$, $\alpha \in \mathbb{R}^D$ and any $\rho > 0$, with the form of infinitesimal generator $L_n (1.2)$, we define a mean one exponential martingale
\[ M_t^\alpha = \exp \left( \langle z_n(t) - z_n(0), \rho \theta \rangle - n \int_0^t \sum_{i=1}^d \lambda_i(z_n(s),\xi_n(s))(e^{\langle \rho \theta, u_i/n \rangle} - 1)ds \right. \]
\[ \left. + \langle \xi_n(t) - \xi_n(0), \alpha \rangle - n \int_0^t \sum_{i=1}^D \chi_{\xi_n(s)=i} \sum_{j=1}^D q_{ij}(z_n(s))(e^{\langle \alpha, e_{ij} \rangle} - 1)ds \right). \]

Define
\[ \Lambda = \sup \{\lambda_i(z,j) \} \quad \text{for } z \in \mathbb{R}^d, 1 \leq i \leq k, j \in \mathbb{Z}_D \]
and $E = \max_{1 \leq i \leq d} |u_i|$. Fix $|\theta| = 1$, we have
\[ n \int_0^t \sum_{i=1}^d \lambda_i(z_n(s),\xi_n(s))(e^{\langle \rho \theta, u_i/n \rangle} - 1)ds \leq tnd\Lambda e^{Ep/n} = R(t, \rho) \]
by Assumption 2.3.
Hence we obtain
\[ \mathbb{P} \left( \sup_{0 \leq t \leq T} \langle z_n(t) - z_n(0), \theta \rangle \geq a \right) \]
\[ = \mathbb{P} \left( \sup_{0 \leq t \leq T} \exp \langle \rho (z_n(t) - z_n(0), \theta) \rangle \geq \exp(\rho a) \right) \]
\[ \leq \mathbb{P} \left( \sup_{0 \leq t \leq T} M_t^{a=0} \geq \exp(\rho a) - R(t, \rho) \right) \]
\[ \leq \exp \left( nT \left[k\Lambda e^{Ep/n} - \frac{\rho a}{nT} \right] \right), \]
where the inequality follows from Doob’s martingale inequality. Take
\[ \rho = \frac{n}{E} \log \frac{a}{Tk\Lambda E} > 0. \]
Then it is not difficult to show that if we set
\[ K(a) = \frac{a}{E} \left( \log \frac{a}{kAE} - 1 \right) \]
for \( a \) large and \( K(a) = 0 \) otherwise, then
\[ \mathbb{P}\left( \sup_{0 \leq t \leq T} \langle z_n(t) - z_n(0), \theta \rangle \geq a \right) \leq \exp\left( -nkTK\left( \frac{a}{T} \right) \right). \]
Applying Lemma 5.4, the lemma is established. \( \square \)

**Corollary 5.6.** There are positive constants \( c_1 \) and \( c_2 \), such that for any \( t, \Delta \in [0, T] \) with \( 0 \leq t + \Delta \leq T \),
\[ \mathbb{P}(\sup_{t \leq s \leq t + \Delta} |z_n(s) - z_n(t)| \geq a) \leq \exp\left(-nac_1 \log\left( \frac{ac_2}{\Delta} \right) \right). \]

**Proof of Lemma 3.1.** Consider a typical interval \([t^n_j, t^n_{j+1}]\). Since \( z_n(t) \) and \( \tilde{z}_n(t) \) agree at the endpoints of this interval, it is obvious that
\[ |\tilde{z}_n(t^n_j) - \tilde{z}_n(t^n_{j+1})| > \frac{\delta}{2} \quad \text{implies} \quad |z_n(t^n_{j+1}) - z_n(t^n_j)| > \frac{\delta}{2}. \]
On the other hand, we have
\[ |z_n(t) - z_n(t^n)| \geq |z_n(t) - \tilde{z}_n(t)| - |\tilde{z}_n(t^n_{j+1}) - \tilde{z}_n(t^n_j)| \]
since \( \tilde{z}_n \) is piecewise linear and \( \tilde{z}_n(t^n_j) = z_n(t^n_j) \). Therefore, if \( |z_n(t) - \tilde{z}_n(t)| > \delta \) for some \( t \) in the \( j \)th interval, we must have
\[ \sup_{t^n_j \leq t \leq t^n_{j+1}} |z_n(t) - z_n(t^n)| \geq \delta/2. \]
Applying Corollary 5.6 with \( a = \delta/2 \) and \( \Delta = T/n \) we obtain
\[ \mathbb{P}\left( \sup_{t^n_j \leq t \leq t^n_{j+1}} |z_n(t) - z_n(t^n)| \geq \delta/2 \right) \leq \exp\left(-n\frac{\delta c_1}{2} \log\left( \frac{n\delta c_3}{2} \right) \right), \]
where \( c_3 = c_2/T \). Thus,
\[ \mathbb{P}(\rho^{(1)}(z_n, \tilde{z}_n) > \delta) \leq \sum_{j=0}^{n-1} \mathbb{P}\left( \sup_{t^n_j \leq t \leq t^n_{j+1}} |z_n(t) - \tilde{z}_n(t)| > \delta \right) \leq \sum_{j=0}^{n-1} \mathbb{P}\left( \sup_{t^n_j \leq t \leq t^n_{j+1}} |z_n(t) - z_n(t^n)| > \delta/2 \right) \leq n \exp\left(-n\frac{\delta c_1}{2} \log\left( \frac{n\delta c_3}{2} \right) \right). \]
The result follows since \( c_1 \) and \( c_3 \) are positive constants.

By the construction of \( \tilde{v}_n \), for any \( i \in \mathbb{Z}_D \), we have
\[ |\tilde{F}_n(t, i) - F_n(t, i)| \leq F_n(t^{m+1}, i) - F_n(t_m, i), \ t \in [t_m, t_{m+1}]. \]
Thus
\[ \int_0^T |\tilde{F}_n(t, i) - F_n(t, i)| \, dt \leq \frac{T}{n} \sum_{m=0}^{T-1} (F_n(t_{m+1}, i) - F_n(t_m, i)) \]
\[ = \frac{T}{n} F_n(T, i) \leq \frac{T^2}{n}. \]

So we have
\[ \mathbb{P}(\rho^{(2)}(\tilde{\nu}_n, \nu_n) > \delta) \leq \mathbb{P} \left( \sum_{i=1}^D \int_0^T |\tilde{F}_n(t, i) - F_n(t, i)| > \delta^2 \right) \]
\[ \leq \mathbb{P} \left( \frac{DT^2}{n} > \delta^2 \right). \]

It means that if \( n > DT^2/\delta^2, \) \( \mathbb{P}(\rho^{(2)}(\tilde{\nu}_n, \nu_n) > \delta) \) is exact 0. \( \square \)

Lemma 5.7. Uniformly in \( x \in \mathbb{R}^d \) and in \( \theta \) and \( \alpha \) in bounded sets,
\[
\limsup_{n \to \infty} \mathbb{E}_x \exp \left\{ n \left( \langle \tilde{z}_n \left( \frac{T}{n} \right) - \tilde{z}_n(0), \theta \rangle - \int_0^{T/n} H_x^\delta(x, \theta, n\tilde{\nu}_n(t, \cdot)) \, dt \right) \right. 
+ \left. \left( \langle \xi_n \left( \frac{T}{n} \right) - \xi_n(0), \alpha \rangle + n \int_0^{T/n} \tilde{S}(\tilde{z}_n(t), n\tilde{\nu}_n(t, \cdot), \alpha) \, dt \right) \right\} \leq 1.
\]

where \( \mathbb{E}_x \) means the expectation with respect to the paths of \( z_n \) starting from \( x \) at \( t = 0 \).

Proof. Define a mean one martingale
\[
M_t = \exp \left\{ n \left( \langle z_n(t) - z_n(0), \theta \rangle - \int_0^t \sum_{i=1}^d \lambda_i(z_n(s), i) (e^{\langle \theta, u_i \rangle} - 1) \, ds \right) \right. 
+ \left. \left( \langle \xi_n(t) - \xi_n(0), \alpha \rangle - n \int_0^t \sum_{i=1}^D \chi_{\langle \xi_n(s) \rangle = i} \sum_{j=1}^D q_{ij}(z_n(s)) (e^{\langle \alpha, u_{ij} \rangle} - 1) \, ds \right) \right\}.
\]

Since \( \tilde{z}_n(t^n_j) = z_n(t^n_j) \), for any \( \theta \) we have
\[
1 = \mathbb{E}_x \exp \left\{ n \left( \langle \tilde{z}_n \left( \frac{T}{n} \right) - \tilde{z}_n(0), \theta \rangle 
- \int_0^{T/n} \sum_{i=1}^d \sum_{j=1}^D \lambda_i(z_n(s), j) (e^{\langle \theta, u_i \rangle} - 1) \nu_n(ds, j) \right) \right. 
+ \left. \left( \langle \xi_n \left( \frac{T}{n} \right) - \xi_n(0), \alpha \rangle - n \int_0^{T/n} \sum_{i,j=1}^D q_{ij}(z_n(s)) (e^{\langle \alpha, e_{ij} \rangle} - 1) \nu_n(ds, i) \right) \right\}.
\]
For convenience, we introduce the probabilistic Dirac measure $K_{\nu_n}$ with $\nu_n(d,i) = d t K_{\nu_n}(t, i)$ for $i = 1, \ldots, D$. Then the term
\[
\sum_{i=1}^{d} \sum_{j=1}^{D} \lambda_i(z_n(s), j)(e^{(\theta, u_i)} - 1)\nu_n(ds, j)
\]
can be formally written as $H_s(z_n(s), \theta, K_{\nu_n}(s, \cdot))ds$ and
\[
-\sum_{i=1}^{D} \sum_{j=1}^{D} q_{ij}(z_n(s))(e^{(\alpha, e_{ij})} - 1)\nu_n(ds, i)
\]
can be formally written as $S(z_n(s), K_{\nu_n}(s, \cdot), \alpha)ds$.

Let
\[
S_\delta = \left\{ \omega: \sup_{0 \leq t \leq T/n} |z_n(t) - x| < \frac{\delta}{2} \right\},
\]
we have
\[
1 \geq \mathbb{E}_x \chi_{S_\delta} \exp \left\{ n \left( \frac{\hat{z}_n(T/n)}{n} - z_n(0), \theta \right) - n \int_0^{T/n} H_s^\delta (x, \theta, K_{\nu_n}(t, \cdot)) dt \right\}
+ \left( \xi_n \left( \frac{T}{n} \right) - \xi_n(0), \alpha \right) + n \int_0^{T/n} S(z_n(t), K_{\nu_n}(t, \cdot), \alpha) dt
\]
\[
= \mathbb{E}_x \exp \left\{ n \left( \frac{\hat{z}_n(T/n)}{n} - z_n(0), \theta \right) - n \int_0^{T/n} H_s^\delta (x, \theta, K_{\nu_n}(t, \cdot)) dt \right\}
+ \left( \xi_n \left( \frac{T}{n} \right) - \xi_n(0), \alpha \right) + n \int_0^{T/n} S(z_n(t), K_{\nu_n}(t, \cdot), \alpha) dt
\]
\[
- \mathbb{E}_x \chi_{S_\delta} \exp \left\{ n \left( \frac{\hat{z}_n(T/n)}{n} - z_n(0), \theta \right) - n \int_0^{T/n} H_s^\delta (x, \theta, K_{\nu_n}(t, \cdot)) dt \right\}
+ \left( \xi_n \left( \frac{T}{n} \right) - \xi_n(0), \alpha \right) + n \int_0^{T/n} S(z_n(t), K_{\nu_n}(t, \cdot), \alpha) dt \right\}
\]

From the Assumption 2.3 and the boundedness of $\theta$ and $\alpha$, we have
\[
\mathbb{E}_x \chi_{S_\delta} \exp \left\{ n \left( \frac{\hat{z}_n(T/n)}{n} - z_n(0), \theta \right) - n \int_0^{T/n} H_s^\delta (x, \theta, K_{\nu_n}(t, \cdot)) dt \right\}
+ \left( \xi_n \left( \frac{T}{n} \right) - \xi_n(0), \alpha \right) + n \int_0^{T/n} S(z_n(t), K_{\nu_n}(t, \cdot), \alpha) dt \right\}
\leq \mathbb{E}_x \chi_{S_\delta} \exp \left\{ n \left( \frac{\hat{z}_n(T/n)}{n} - z_n(0), \theta \right) + 3M \right\}
\leq \sum_{k=1}^{\infty} \exp \left\{ n(K+1) - \frac{\delta}{2} | \theta | + 3M \right\} \times \mathbb{P} \left( \frac{K\delta}{2} \leq \sup_{0 \leq t \leq T/n} |z_n(t) - x| \leq \frac{(K+1)\delta}{2} \right)
\leq \sum_{k=1}^{\infty} \exp \left\{ n \left( \frac{(K+1)\delta}{2} B_1 - \frac{K\delta c_1}{2} \log \left( \frac{K\delta c_2 n}{2T} \right) \right) \right\} \times e^{3M} \to 0 \quad (5.3)
as $n \to \infty$ for all $x \in \mathbb{R}^d$ with $|\theta| \leq B_1$ and $|\alpha| \leq B_2$. Combining (5.3), (5.2) and Lemma 3.1, we complete the proof. \hfill \Box

**Corollary 5.8.** There exist constants $C > 0$ and $n_0$, such that

$$
\mathbb{E} \exp\{n J_n(\tilde{z}_n, \theta, \tilde{\nu}_n, \alpha)\} \leq C
$$

for all $n > n_0$.

**Proof.** By definition

$$
\exp\{n J_n(\tilde{z}_n, \theta, \tilde{\nu}_n, \alpha)\}
= \exp\left\{ \sum_{j=0}^{n-1} \left( n \left\langle \tilde{z}_n(t^n_{j+1}), \theta(t^n_j) \right\rangle - n \int_{t^n_j}^{t^n_{j+1}} H_s^{\delta}(\tilde{z}_n(t^n_j), \theta(t^n_j), \nu_n(t, \cdot)) \, dt \right) \right\}
$$

$$
- n \int_{t^n_0}^{t^n_n} H_s^{\delta}(\tilde{z}_n(t^n_0), \theta(t^n_0), \nu_n(t, \cdot)) \, dt
$$

$$
+ n \int_{t^n_0}^{t^n_n} S(\tilde{z}_n(t, \cdot), \nu_n(t, \cdot), \alpha(t^n_0)) \, dt
$$

$$
+ \left( \xi_n(t^n_{j+1}) - \xi_n(t^n_j), \alpha(t^n_j) \right) + n \int_{t^n_j}^{t^n_{j+1}} S(\tilde{z}_n(t, \cdot), \nu_n(t, \cdot), \alpha(t^n_0)) \, dt \right) \right\}
$$

$$
- \sum_{j=0}^{n-1} \left( \xi_n(t^n_{j+1}) - \xi_n(t^n_j), \alpha(t^n_j) \right) \right\}.
$$

(5.4)

Now $\alpha$ is a step function, so it is constant on an interval, say $[0, \tau]$. We have

$$
\sum_{j=0}^{n-1} \chi_{\{t^n_{j+1} \leq \tau\}} \left\langle \xi_n(t^n_{j+1}), \alpha(0) \right\rangle
= \left\langle \xi_n \left( \frac{|n\tau|}{n} \right) - \xi_n(0), \alpha(0) \right\rangle
$$

where $|a|$ is the largest integer smaller than $a$. Since $\xi_n$ and $\alpha$ are bounded,

$$
\left| \left\langle \xi_n \left( \frac{|n\tau|}{n} \right) - \xi_n(0), \alpha(0) \right\rangle \right|
$$

is uniformly bounded. Repeating this argument on the finite number of intervals on which $\alpha$ are constants and by the Markov property of $\xi_n(t)$, $\left| \sum_{j=0}^{n-1} \left\langle \xi_n(t^n_{j+1}) - \xi_n(t^n_j), \alpha(t^n_j) \right\rangle \right|$ is bounded.
Thus by (5.4), Lemma 5.7 and the Markov property of $z_n(t)$,
\[
\limsup_{n \to \infty} \mathbb{E} \exp \{nJ_n(z_n, \theta, \bar{\nu}_n, \alpha)\} \leq C
\]
where $C$ is a positive constant. \hfill \Box

**Lemma 5.9.** If $(r_n, \nu_n) \in \Phi_K = \{(r, \nu) \in \mathbb{D}^d[0, T] \times \mathbb{M}[0, T] : I_s(r, \nu) + I_f(r, \nu) \leq K\}$ for all $n$, then $(r_n, \nu_n)$ are both uniformly absolutely continuous.

The proof of Lemma 5.9 can be made similarly as Lemma 5.18 in [19].

**Lemma 5.10.** The rate functional is lower semicontinuous, i.e., if $(r_n, \nu_n) \to (r, \nu)$ as $n \to \infty$, then
\[
\liminf_{n \to \infty} I_s(r_n, \nu_n) + I_f(r_n, \nu_n) \geq I_s(r, \nu) + I_f(r, \nu).
\]

**Proof.** We only need to consider the sequences of $r_n$ and $\nu_n$ which are absolutely continuous since it will be trivial otherwise. Let $(r_n, \nu_n) \to (r, \nu)$ as $n \to \infty$. We may assume that $I_s(r_n, \nu_n) + I_f(r_n, \nu_n)$ is bounded, say by a constant $K$. By Lemma 5.9, we know that $r$ and $\nu$ are also absolutely continuous.

We will first prove that $I_f(r, \nu)$ is lower semicontinuous. Recall the notation defined in Theorem 2.4,
\[
S(r(t), n_\nu(t, \cdot)) = \sup_{\alpha \in \mathbb{R}^D} \left\{ - \sum_{i,j} n_\nu(t, \cdot) q_{ij}(r(t)) \left( e^{\langle \alpha, e_i \rangle(J)} - 1 \right) \right\}.
\]

By Lemmas 5.1 and 5.2, $S(r, n_\nu)$ is lower semicontinuous in $r$ and convex in $n_\nu$. Thus for any $\epsilon > 0$ and small enough $\Delta$, when $n$ is large enough, we have
\[
\begin{align*}
& \int_0^T S(r_n(t), n_{\nu_n}(t, \cdot)) dt \\
& \geq \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} S(r(t_j), n_{\nu_n}(t, \cdot)) dt - \epsilon \\
& \geq \sum_{j=0}^{J-1} \Delta \cdot S \left( r(t_j), \frac{\int_{t_j}^{t_{j+1}} n_{\nu_n}(t, \cdot) dt}{\Delta} \right) dt - \epsilon \\
& = \sum_{j=0}^{J-1} \Delta \cdot S \left( r(t_j), \frac{F_n(t_{j+1}, \cdot) - F_n(t_j, \cdot)}{\Delta} \right) - \epsilon. \quad (5.5)
\end{align*}
\]

Define the functions $r_j, F_j$ and $F'_j$ as
\[
r_j(t) = r(t_j), \quad \text{for } t_j \leq t < t_{j+1}, \quad j = 0, 1, \cdots, J - 1,
\]
and
\[
F_j(t, \cdot) = F(t_j, \cdot), \quad F'_j(t, \cdot) = F(t_{j+1}, \cdot), \quad \text{for } t_j \leq t < t_{j+1}, \quad j = 0, 1, \cdots, J - 1.
\]
By (5.5) and Fatou’s Lemma,
\[
\liminf_{n \to \infty} \int_0^T S(r_n(t), n_v(t, \cdot)) \, dt \\
\geq \sum_{j=0}^{T-1} \int_{t_j}^{t_{j+1}} \liminf_{n \to \infty} S \left( r(t), \frac{F_n(t_{j+1}, \cdot) - F_n(t_j, \cdot)}{\Delta} \right) \, dt - \epsilon \\
= \sum_{j=0}^{T-1} \int_{t_j}^{t_{j+1}} S \left( r(t), \frac{F^j(t, \cdot) - F(t, \cdot)}{\Delta} \right) \, dt - \epsilon \\
= \int_0^{T-\Delta} S \left( r(t), \frac{F^j(t, \cdot) - F(t, \cdot)}{\Delta} \right) \, dt - \epsilon.
\]

Now we choose a sequence of nested partition, say \( J_k = 2^k \) and \( \Delta_k = T/2^k \). Again by Fatou’s Lemma and lower semicontinuity of \( S \),
\[
\liminf_{k \to \infty} \int_0^{T-\Delta_k} S \left( r_{J_k}(t), \frac{F^{J_k}(t, \cdot) - F(t, \cdot)}{\Delta} \right) \, dt \\
\geq \int_0^T \liminf_{k \to \infty} \chi_{\{t \leq T-\Delta_k\}} S \left( r_{J_k}(t), \frac{F^{J_k}(t, \cdot) - F(t, \cdot)}{\Delta} \right) \, dt \\
\geq \int_0^T S(r(t), n_v(t, \cdot)) \, dt.
\]

Thus we obtain the lower semicontinuity of \( I_f(r, v) \). The lower semicontinuity of \( I_s(r, v) \) can be established similarly.

With the same manner, we can obtain the following lemma:

**Lemma 5.11.** For any fix step functions \( \theta(t) \in R^d \) and \( \alpha(t) \in R^D \), if \( (r_n, v_n) \to (r, v) \) as \( n \to \infty \), then
\[
\liminf_{n \to \infty} I_s^\delta (r_n, v_n, \theta) + I_f(r_n, v_n, \alpha) \geq I_s(r, v, \theta) + I_f(r, v, \alpha).
\]

**Lemma 5.12.** Given \( r \in D^d[0, T] \) and \( v \in M[0, T] \), for any \( \epsilon > 0 \), there exist step functions \( \theta(t) \in R^d \) and \( \alpha(t) \in R^D \), such that
\[
I_s^\delta (r, v, \theta) \geq I_s^\delta (r, v) - \epsilon \quad (5.6)
\]
and
\[
I_f(r, v, \alpha) \geq I_f(r, v) - \epsilon. \quad (5.7)
\]
The proof of (5.6) can be referred to Lemma 5.43 in [19] and the proof of (5.7) is similar.

**Proof of Lemma 3.10.** The goal is to prove that for each $m$, there exist an integer $K_m > 0$ and matrices $\eta^{mk} \in S_m$, $k = 0, 1, \cdots, K_m$, such that

\[
\sum_{k=0}^{K_m-1} \int_{t_{mk}}^{t_{m+1}} \sum_{i=1}^D n_i(t, i) \sum_{j=1}^D \left( \eta_{ij}^m \log \frac{\eta_{ij}^m}{q_{ij}(y(t))} + q_{ij}(y(t)) - \eta_{ij}^m \right) dt 
\leq \int_{t_m}^{t_{m+1}} \sup_{\alpha \in R^D} \left( - \sum_{i,j=1}^D n_v(t, i) q_{ij}(r(t))(e^{(\alpha, e_{ij})} - 1) \right) dt + (t_{m+1} - t_m)\epsilon.
\]

For any $\epsilon > 0$, we can choose $J$ large enough such that for any $t, s \in [t_m, t_{m+1}]$, $|t - s|$ can be arbitrarily small. Also, $\rho^{(1)}(y, \nu)$ and $\rho^{(2)}(\pi, \nu)$ can be arbitrarily small. According to Assumption 2.3 and Lemma 5.1, we have that for $J$ large enough

\[
\sup_{\alpha \in R^D} \left( - \sum_{i,j=1}^D n_v(t, i) q_{ij}(r(t))(e^{(\alpha, e_{ij})} - 1) \right) + \epsilon 
\leq \sup_{\alpha \in R^D} \left( - \sum_{i,j=1}^D n_v(t, i) q_{ij}(r(t))(e^{(\alpha, e_{ij})} - 1) \right) + 2\epsilon
\]

for all $t, s \in [t_m, t_{m+1}]$. Define $\psi(i) = n_i(t, i)$ for $i = 1, \cdots, D$, then $\psi \in R^D$ with $\sum_{i=1}^D \psi(i) = 1$. Along the lines of the proof of Theorem 8.19 in [19], we have for any fixed $t$

\[
\inf_{\eta \in S_m} \sum_{i=1}^D \psi(i) \sum_{j=1}^D \left( \eta_{ij} \log \frac{\eta_{ij}}{q_{ij}(y(t))} + q_{ij}(y(t)) - \eta_{ij} \right) = \sup_{\alpha \in R^D} \left( - \sum_{i,j=1}^D \psi(i) q_{ij}(y(t))(e^{(\alpha, e_{ij})} - 1) \right).
\]

Thus for each $s \in [t_m, t_{m+1}]$, there exists a matrix $\eta^m(s) \in S_m$ such that

\[
\sum_{i=1}^D n_i(s, i) \sum_{j=1}^D \left( \eta_{ij}^m(s) \log \frac{\eta_{ij}^m(s)}{q_{ij}(y(s))} + q_{ij}(y(s)) - \eta_{ij}^m(s) \right) 
\leq \sup_{\alpha \in R^D} \left( - \sum_{i,j=1}^D n_v(s, i) q_{ij}(y(s))(e^{(\alpha, e_{ij})} - 1) \right) + \epsilon.
\]
Substitute (5.8) into (5.9), we have

\[
\sum_{i=1}^{D} n_{\pi}(s, i) \sum_{j=1}^{D} \left( \frac{\eta_{ij}^m(s)}{q_{ij}(y(t))} \right) + q_{ij}(y(t)) - \eta_{ij}^m(s) \right) \right) (e^{(\alpha_e, e_i)} - 1) + 3\epsilon \tag{5.10}
\]

for all \( t \in [t_m, t_{m+1}] \). Since (5.10) is continuous with respect to \( n_{\pi} \) and \( q_{ij} \), there exists \( \delta_s > 0 \) such that for each fixed \( s \in [t_m, t_{m+1}] \)

\[
\sum_{i=1}^{D} n_{\pi}(t, i) \sum_{j=1}^{D} \left( \frac{\eta_{ij}^m(s)}{q_{ij}(y(t))} \right) + q_{ij}(y(t)) - \eta_{ij}^m(s) \right) \right) (e^{(\alpha_e, e_i)} - 1) + 4\epsilon \tag{5.11}
\]

holds for any \( t \in O_s = (s - \delta_s, s + \delta_s) \cap [t_m, t_{m+1}] \). By Heine-Borel theorem, we can choose finite number of \( O_s \) in \( \{O_s\}_{s\in[t_m, t_{m+1}]} \) to cover \( [t_m, t_{m+1}] \).

It means that there exists a further subdivision of interval \( [t_m, t_{m+1}] \) (i.e., \( t_m = t_{m0} < t_{m1} < \cdots < t_{mK_m} = t_{m+1} \)) and related matrices \( \eta^m(s_k) \in S_m \) such that for all \( t \in [t_{mk}, t_{mk+1}] \)

\[
\sum_{i=1}^{D} n_{\pi}(t, i) \sum_{j=1}^{D} \left( \frac{\eta_{ij}^m(s_k)}{q_{ij}(y(t))} \right) + q_{ij}(y(t)) - \eta_{ij}^m(s_k) \right) \right) (e^{(\alpha_e, e_i)} - 1) + 4\epsilon.
\]

To simplify the notation, we rewrite \( \eta^m(s_k) \) as \( \eta^{mk} \). Thus we have

\[
\sum_{k=0}^{K_{m-1}} \int_{t_{mk}}^{t_{mk+1}} \sum_{i=1}^{D} n_{\pi}(t, i) \sum_{j=1}^{D} \left( \frac{\eta_{ij}^{mk}}{q_{ij}(y(t))} \right) + q_{ij}(y(t)) - \eta_{ij}^{mk} \right) dt \leq \int_{t_{m}}^{t_{m+1}} \sup_{a \in R^d} \left( - \sum_{i,j=1}^{D} n_{\pi}(t, i) q_{ij}(r(t))(e^{(\alpha_e, e_i)} - 1) \right) dt + 4(t_{m+1} - t_m)\epsilon.
\]

The proof is completed. \(\square\)

**Proof of Lemma 3.12.** We need to prove that for any bounded continuous function \( h(t, z) \),

\[
\lim_{n \to \infty} \int_{0}^{T} h(t, \xi_n(t)) dt = \int_{0}^{T} \sum_{i=1}^{D} h(t, i) n_{\pi}(t, i) dt.
\]
It suffices to prove that for each time interval \([t_{mk}, t_{m,k+1}]\),
\[
\lim_{n \to \infty} \int_{t_{mk}}^{t_{m,k+1}} h(t, \bar{\xi}_n(t)) \, dt = \int_{t_{mk}}^{t_{m,k+1}} \sum_{i=1}^{D} h(t, i) n_{\bar{\pi}}(t, i) \, dt.
\]
Since \(\bar{\xi}_n\) lives on only finite states, then for any \(\epsilon > 0\), there exists \(\delta > 0\) such that for \(|t_k - t| < \delta\)
\[
|h(t, \bar{\xi}_n(t)) - h(t_k, \bar{\xi}_n(t))| < \epsilon,
\]
for all \(t \in [t_{mk}, t_{m,k+1}]\).

Take an integer \(L\) large enough and define \(\tilde{\delta} = (t_{m,k+1} - t_{mk}) / L < \delta\). Let \(\tau_l = t_{mk} + l\tilde{\delta}\) for \(l = 0, 1, \ldots, L\). We have
\[
\limsup_{n \to \infty} \int_{t_{mk}}^{t_{m,k+1}} h(t, \bar{\xi}_n(t)) \, dt = \limsup_{n \to \infty} \sum_{l=0}^{L-1} \int_{\tau_l}^{\tau_{l+1}} h(t, \bar{\xi}_n(t)) \, dt + T\epsilon
\]
\[
\leq \sum_{k=0}^{L-1} \int_{\tau_l}^{\tau_{l+1}} \sum_{i=1}^{D} h(t, i) n_{\bar{\pi}}(t, i) \, dt + T\epsilon
\]
\[
= \sum_{k=0}^{L-1} \int_{\tau_l}^{\tau_{l+1}} \sum_{i=1}^{D} h(t, i) n_{\bar{\pi}}(t, i) \, dt + 2T\epsilon
\]
\[
= \int_{t_{mk}}^{t_{m,k+1}} \sum_{i=1}^{D} h(t, i) n_{\bar{\pi}}(t, i) \, dt + 2T\epsilon.
\]
In (5.12) we utilized the ergodicity of the process \(\bar{\xi}_n\). Similarly, we can prove
\[
\liminf_{n \to \infty} \int_{0}^{T} h(t, \bar{\xi}_n(t)) \, dt \geq \int_{t_{mk}}^{t_{m,k+1}} \sum_{i=1}^{D} h(t, i) n_{\bar{\pi}}(t, i) \, dt - 2T\epsilon.
\]
The proof is completed.

**Proof of Lemma 3.13.** The goal is to prove that for any \(\epsilon > 0\),
\[
\lim_{n \to \infty} \mathbb{P} \left( \sup_{0 \leq t \leq T} |\bar{z}_n(t) - y(t)| \geq \epsilon \right) = 0.
\]
For any \( \theta \in \mathbb{R}^d \) and \( \rho > 0 \), we have the martingale

\[
M_t = \exp \left\{ \langle \tilde{z}_n(t) - y(t), \rho \theta \rangle \right. \\
- \int_0^t \sum_{i=1}^d \left( n \mu_i(s) \frac{\lambda_i(\tilde{z}_n(s), \tilde{\xi}_n(s))}{\lambda_i^\theta(y(s))} (e^{\langle \rho \theta, u_i/n \rangle} - 1) \\
- \mu_i(s) \langle \rho \theta, u_i \rangle \right) ds \}
\]

\[
= \exp \left\{ \langle \tilde{z}_n(t) - y(t), \rho \theta \rangle \right. \\
- \int_0^t \sum_{i=1}^d \left( \mu_i(s) \frac{\lambda_i(\tilde{z}_n(s), \tilde{\xi}_n(s)) - \lambda_i^\theta(y(s))}{\lambda_i^\theta(y(s))} \langle \rho \theta, u_i \rangle \\
+ \mu_i(s) \frac{\lambda_i(\tilde{z}_n(s), \tilde{\xi}_n(s))}{\lambda_i^\theta(y(s))} (n(e^{\langle \rho \theta, u_i/n \rangle} - 1) - \langle \rho \theta, u_i \rangle) \right) ds \}
\]

Recall the Assumption 2.3 and \( \mu_i(t) \) is piecewise constant and bounded, we can perform similar estimate as in Lemma 5.5 to obtain

\[
\mathbb{P} \left( \sup_{0 \leq t \leq T} |\tilde{z}_n(t) - y(t) - \int_0^t \sum_{i=1}^d \mu_i(s) \frac{\lambda_i(\tilde{z}_n(s), \tilde{\xi}_n(s)) - \lambda_i^\theta(y(s))}{\lambda_i^\theta(y(s))} dsu_i| \geq \epsilon \right)
\leq \exp \left( -n \epsilon c_1 \log(\epsilon c_2) \right),
\]  
(5.13)

where \( c_1 \) and \( c_2 \) are positive constants. By Lemma 3.12, we have

\[
\int_0^t \lambda_i(\tilde{z}_n(s), \tilde{\xi}_n(s)) ds - \int_0^t \lambda_i^\theta(y(s)) ds
= \left( \int_0^t \lambda_i(\tilde{z}_n(s), \tilde{\xi}_n(s)) ds - \int_0^t \lambda_i(y(s), \tilde{\xi}_n(s)) ds \right)
+ \left( \int_0^t \sum_{j=1}^D \lambda_i(y(s), j)n_{\tilde{v}_n}(s, j) ds - \int_0^t \lambda_i^\theta(y(s)) ds \right)
\leq K \int_0^t |\tilde{z}_n(s) - y(s)| ds + F_n(t),
\]  
(5.14)

where

\[
F_n = \sup_{t \in [0, T]} \left| \sum_{j=1}^D \int_0^t \lambda_i(y(s), j) \left( n_{\tilde{v}_n}(s, j) - n_\theta(s, j) \right) ds \right| \rightarrow 0
\]  
(5.15)

as \( n \) goes to infinity for \( t \leq T \).
From the Assumption 2.3 that \( \log \lambda_i \) is bounded, we can find \( M > 0 \) such that
\[
\inf_{s \in [0, T]} \lambda_i^\bar{\pi}(y(s)) \geq \frac{1}{M}. \tag{5.16}
\]
Let \( C = dMABT \), where \( A = \max_{t \in [0, T]} \max_{i=1, \ldots, k} \mu_i(t) \) and \( B = \max_{i=1, \ldots, k} |u_i| \). Combining (5.13), (5.14) and (5.16), we have
\[
\mathbb{P}\left( \sup_{0 \leq t \leq T} \left| \bar{z}_n(t) - y(t) \right| - CK \int_0^t \left| \bar{z}_n(s) - y(s) \right| ds - CF_n \right) \geq \epsilon \leq \mathbb{P}\left( \sup_{0 \leq t \leq T} \left| \bar{z}_n(t) - y(t) - \int_0^t \sum_{i=1}^d \mu_i(t) \frac{\lambda_i(\bar{z}_n(s), \bar{\xi}_n(s)) - \lambda_i^\bar{\pi}(y(s))}{\lambda_i^\bar{\pi}(y(s))} ds u_i \right| \right) \geq \epsilon \leq \exp\left( -nc_1 \log(\epsilon c_2) \right). \tag{5.17}
\]
From (5.17) and Gronwall's inequality, we obtain
\[
\mathbb{P}\left( \sup_{0 \leq t \leq T} \left| \bar{z}_n(t) - y(t) \right| \geq (\epsilon + CF_n)e^{CKT} \right) \leq \mathbb{P}\left( \sup_{0 \leq t \leq T} \left| \bar{z}_n(t) - y(t) - \int_0^t \sum_{i=1}^d \mu_i(t) \frac{\lambda_i(\bar{z}_n(s), \bar{\xi}_n(s)) - \lambda_i^\bar{\pi}(y(s))}{\lambda_i^\bar{\pi}(y(s))} ds u_i \right| \right) \leq \exp\left( -nc_1 \log(\epsilon c_2) \right).
\]
Combining the condition (5.15) and the inequality
\[
\mathbb{P}\left( \sup_{0 \leq t \leq T} \left| \bar{z}_n(t) - y(t) \right| \geq 2\epsilon e^{CKT} \right) \leq \mathbb{P}\left( \sup_{0 \leq t \leq T} \left| \bar{z}_n(t) - y(t) \right| \geq (\epsilon + CF_n)e^{CKT} \right) + \mathbb{P}\left( F_n \geq \frac{\epsilon}{C} \right),
\]
we finish the proof. \( \square \)

**Acknowledgement**

T. Li acknowledges the support of NSFC under grants 11171009, 11421101, 91130005 and the National Science Foundation for Excellent Young Scholars (Grant No. 11222114). They also thank Weinan E, Yong Liu, and Xiaoguang Li for helpful discussions.

**References**

[1] M. Assaf, E. Roberts, and Z. Luthey-Schulten, *Determining the Stability of Genetic Switches: Explicitly Accounting for mRNA Noise*, Phys. Rev. Lett. 106 (2011), 2048102.

[2] M. Chen and Y. Lu, *On evaluating the rate functions of large deviations for jump processes*, Acta. Math. Sin. 6 (1990), 206-219.
[3] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Springer-Verlag, New York, 2nd Edition, 1998.
[4] W. E, D. Liu and E. Vanden-Eijnden, Analysis of multiscale methods for stochastic differential equations, Comm. Pure App. Math. 58 (2005), 1544-1585.
[5] W. E, D. Liu and E. Vanden-Eijnden, Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales, J. Comp. Phys. 221 (2007), 158-180.
[6] R.S. Ellis, *Entropy, large deviations and statistical mechanics*, Springer-Verlag, New York, 1985.
[7] M. Elowitz, A. Levine, E. Siggia and P. Swain, *Stochastic gene expression in a single cell*, Science 297 (2002), 1183-1186.
[8] S.N. Ethier and T.G. Kurtz, *Markov processes: characterization and convergence*, John Wiley and Sons, New York, 1986.
[9] M.I. Freidlin and A.D. Wentzel, *Random perturbations of dynamical systems*, Springer-Verlag, New York, 2nd Edition, 1998.
[10] D.T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem. 81 (1977), 2340-2361.
[11] D.T. Gillespie, *The chemical Langevin equation*, J. Chem. Phys. 113 (2000), 297-306.
[12] M. Heymann and E. Vanden-Eijnden, *The geometric minimum action method: A least action principle on the space of curves*, Comm. Pure Appl. Math. 61 (2008), 1052-1117.
[13] T.G. Kurtz, *The relationship between stochastic and deterministic models for chemical reactions*, J. Chem. Phys. 57 (1972), 2976 - 2978.
[14] T.G. Kurtz, *A limit theorem for perturbed operator semigroups with applications to random evolutions*, J. Funct. Anal. 12 (1973), 55 - 67.
[15] R. Liptser, *Large deviations for two scaled diffusions*, Prob. Theory Relat. Fields 106 (1996), 71-104.
[16] C. Lv, X. Li, F. Li and T. Li, *Constructing the energy landscape for genetic switching system driven by intrinsic noise*, PLoS ONE 9 (2014), e88167.
[17] G. Papanicolaou, *Introduction to the asymptotic analysis of stochastic equations*, Lectures in Mathematics, Vol. 16, American Mathematical Society, Rode Island, 1977.
[18] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.
[19] A. Shwartz and A. Weiss, *Large deviations for performance analysis: queues, communications and computing*, Chapman and Hall, London, 1995.
[20] A. Shwartz and A. Weiss, *Large deviations with diminishing rates*, Math. Oper. Res. 30 (2005), 281-310.
[21] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier, Amsterdam, 3rd Edition, 2007.
[22] K. Zhang, M. Sasai and J. Wang, *Eddy current and coupled landscapes for nonadiabatic and nonequilibrium complex system dynamics*, Proc. Nat. Acad Sci. USA 110 (2013), 14930-14935.