Discrete and Continuous Dynamical Systems - Series S
Vol. 18, No. 10, October 2025, pp. 2769-2790
doi:10.3934/dcdss.2025068

LARGE DEVIATION PRINCIPLE FOR TWO TIME-SCALE REGIME-SWITCHING PROCESSES

Jianrui Li[1] and Jinghai Shao[*1]

[1]Center for Applied Mathematics, Tianjin University,
Tianjin 300072, China

Abstract. This work investigates the large deviation principle for a fully coupled two time-scale system, whose slow process is a diffusion process and fast process is a purely jumping process on a discrete state space. We focus on overcoming the difficulties caused by the infinite countability of the state space of the fast process. To this end, two different drift conditions are proposed separately to deal with two different time-scale ratios.

1. **Introduction.** We study in this work a fully coupled two time-scale stochastic system $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ in $\mathbb{R}^d \times \mathcal{S}$, where $\mathcal{S} = \{1, 2, \dots, N\}$ with $N \leq \infty$. The slow process $(X_t^{\varepsilon,\alpha})$ is described as a solution to the following stochastic differential equation (SDE):

$$\begin{aligned} \mathrm{d}X_t^{\varepsilon,\alpha} &= b(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})\mathrm{d}W_t, \\ X_0^{\varepsilon,\alpha} &= x_0 \in \mathbb{R}^d, \quad Y_0^{\varepsilon,\alpha} = i_0 \in \mathcal{S}, \end{aligned} \tag{1.1}$$

and the fast process $(Y_t^{\varepsilon,\alpha})$ is a jumping-process on $\mathcal{S}$ satisfying

$$\mathbb{P}(Y_{t+\delta}^{\varepsilon,\alpha} = j | Y_t^{\varepsilon,\alpha} = i, X_t^{\varepsilon,\alpha} = x) = \begin{cases} \frac{1}{\alpha} q_{ij}(x)\delta + o(\delta), & \text{if } i \neq j, \\ 1 + \frac{1}{\alpha} q_{ii}(x)\delta + o(\delta), & \text{if } i = j \end{cases} \tag{1.2}$$

for $\delta > 0$, $i, j \in \mathcal{S}$, $x \in \mathbb{R}^d$, and ε, α are small positive parameters. In the existing literature, the system $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ is called *fully coupled* if the diffusion coefficient σ of slow process $(X_t^{\varepsilon,\alpha})$ depends on the fast process $(Y_t^{\varepsilon,\alpha})$ and the transition rates $(q_{ij}(x))_{i,j\in\mathcal{S}}$ of the fast process $(Y_t^{\varepsilon,\alpha})$ depends on $(X_t^{\varepsilon,\alpha})$ as well.

Multi-scale systems arise in many research fields such as in systems biology [7, 17, 18, 20, 36], in mathematical finance [8, 9], etc. Correspondingly, there are many works devoted to the study of averaging principle, central limit theorems, and large deviations of these stochastic models. For a two time-scale system where both slow and fast components are continuous processes given as solutions of SDEs, these problems have been extensively studied, such as, in [1, 19, 20, 21, 28, 29, 33, 34, 37, 40, 41] amongst others. The interaction between the fast component and the slow one makes a fully coupled two time-scale system much complicated, which has been revealed in the works [22, 37, 40, 41].

2020 *Mathematics Subject Classification.* Primary: 60F10; Secondary: 60J74, 60J10.

Key words and phrases. Large deviations, regime-switching, comparison principle, Markov modulated diffusions.

All authors are supported by National Key R&D Program of China (No. 2022YFA1006000) and NNSFs of China (No. 12271397).

*Corresponding author: Jinghai Shao.

In [30] we have studied the averaging principle for $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$, which says that $(X_t^{\varepsilon,\alpha})$ converges strongly or weakly to some limit process $(\bar{X}_t)$ as $\varepsilon, \alpha \to 0$. In [30] we focus on addressing the impact on the limit process $(\bar{X}_t)$ caused by the regularity of invariant probability measure π^x of $(q_{ij}(x))$ when $\mathcal{S}$ is an infinitely countable state space. To establish the averaging principle, we proposed different ergodicity conditions on the Markov chains associated with $(q_{ij}(x))$ for every x, and generalized the coupling method based on Skorokhod's representation theorem for jumping processes.

As a continuation work of [30], we aim to establish the large deviation principle (LDP) associated with the fully coupled system $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ as $\varepsilon, \alpha \to 0$. We mainly want to analyze the difficulties caused by: 1) the infinite countability of the state space $\mathcal{S}$ of the fast process $(Y_t^{\varepsilon,\alpha})$; 2) the different ratios ε/α as $\varepsilon, \alpha \to 0$.

When $\mathcal{S}$ is a finite state space, there are many related works on the LDP of the two time-scale system $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$. In the situation that $(Y_t^{\varepsilon,\alpha})$ is independent of the slow process $(X_t^{\varepsilon,\alpha})$, Eizenberg and Freidlin [6], Freidlin and Lee [12] investigated separately the limit behavior of solutions of PDE systems with Dirichlet boundary associated with $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ when the diffusion coefficient of $X_t^{\varepsilon,\alpha}$ does not depend or depends on $Y_t^{\varepsilon,\alpha}$. These two works reveal that whether the diffusion coefficient of $X_t^{\varepsilon,\alpha}$ depends on $Y_t^{\varepsilon,\alpha}$ or not has an important impact on the method to study the limit behavior of $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$. Moreover, to provide a decisive estimate on the difference between $(X_t^{\varepsilon,\alpha})$ and its limit process, the LDP was established in [14, 15]. In the setting where the fast process $(Y_t^{\varepsilon,\alpha})$ is a jumping process depending on the slow process $(X_t^{\varepsilon,\alpha})$ as well, the averaging principle and the LDP have been studied by Faggionato, Gabrielli, and Crivellari [7] and Budhiraja, Dupuis and Ganguly [3]. [7] considered a simple case without diffusion term for the slow component. Whereas, [3] considered a fully coupled case by using the weak convergence method, and established a process level large deviation principle.

It is known that the infinite countability of the state space $\mathcal{S}$ of $(Y_t^{\varepsilon,\alpha})$ has an important impact on the averaging principle and LDP of $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$. For example, in a simple setting $\alpha \equiv 1$, Bezuidenhout [2] studied the LDP of certain functionals of $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ with the diffusion coefficient of $(X_t^{\varepsilon,\alpha})$ independent of $(Y_t^{\varepsilon,\alpha})$. It showed that the LDP holds when $(Y_t^{\varepsilon,\alpha})$ is in a finite state space. Furthermore, it was shown by a counterexample that when $(Y_t^{\varepsilon,\alpha})$ is a Markov chain in an infinite state space, the LDP may fail.

In order to deal with the difficulty caused by the infinite countability of $\mathcal{S}$, we propose separately two drift conditions in the different situations that the ratio $\varepsilon/\alpha = 1$ or tends to ∞ as $\varepsilon, \alpha \to 0$. We shall apply the nonlinear semigroup method developed in [10] to study the LDP of $\{X_t^{\varepsilon,\alpha}; \varepsilon, \alpha > 0\}$. It is known (cf. [10]) that a crucial point of this method is to establish the comparison principle for the associated Hamilton-Jacobi equation. This task is of great challenging in applications for different stochastic systems. For example, Popovic [36] verified the comparison principle for chemical reaction networks on multiple time-scales. Kraaij and Schlottke [26] established a comparison principle for the Hamilton-Jacobi equation in terms of a Hamiltonian given in a variational form, and applied it in [25] to investigate the LDP for two time-scale interacting particle system on a finite state space coupled to fast drift-diffusion processes on a compact space. In this work we prove the desired comparison principle by checking the conditions of a general result established by Ishii [16], where the continuity of $x \mapsto \pi^x$ established in the study of averaging principle in [30] is needed.

The remainder of this work is organized as follows. In Section 2, we introduce the framework and the nonlinear semigroup method developed in Feng and Kurtz [10] to establish the LDP of $\{X_t^{\varepsilon,\alpha};\varepsilon,\alpha>0\}$. We present the conditions needed to be checked later for our studied processes. In Section 3, we establish the LDP in the case $\varepsilon/\alpha=1$. In Section 4, we deal with the LDP in the case $\varepsilon/\alpha\to\infty$ as $\varepsilon,\alpha\to 0$.

2. **Framework.**

Let us begin with introducing three fundamental conditions on the stochastic system $(X_t^{\varepsilon,\alpha},Y_t^{\varepsilon,\alpha})$, which will be used throughout this work.

(A1) There exist constants K_1, $K_2>0$ such that

$$|b(x,i)-b(y,i)|+\|\sigma(x,i)-\sigma(y,i)\|\le K_1|x-y|,$$
$$|b(x,i)|+\|\sigma(x,i)\|\le K_2,\qquad x,y\in\mathbb{R}^d,\ i\in\mathcal{S}.$$

(A2) For each $x\in\mathbb{R}^d$, $(q_{ij}(x))_{i,j\in\mathcal{S}}$ is a conservative, irreducible transition rate matrix. Assume $\kappa:=\sup_{i\in\mathcal{S}}\sum_{j\in\mathcal{S},j\ne i}\sup_{x\in\mathbb{R}^d}q_{ij}(x)<\infty$.

(A3) There exists a constant $K_3>0$ such that

$$\|Q(x)-Q(y)\|_{\ell_1}:=\sup_{i\in\mathcal{S}}\sum_{j\ne i}|q_{ij}(x)-q_{ij}(y)|\le K_3|x-y|,\quad x,y\in\mathbb{R}^d.$$

Under these conditions (A1)-(A3), the two time-scale system (1.1), (1.2) admit a unique strong solution to any initial value $X_0^{\varepsilon,\alpha}=x_0\in\mathbb{R}^d$ and $Y_0^{\varepsilon,\alpha}=i_0\in\mathcal{S}$; see, e.g. [43] or [38] under certain more general non-Lipschitz conditions.

In this work we shall investigate the LDP for $\{X_t^{\varepsilon,\alpha};\varepsilon,\alpha>0\}$ using the nonlinear semigroup and viscosity solution method. A general method was developed by Feng and Kurtz in [10] to establish the LDP for Markov processes based on nonlinear semigroups and viscosity solutions to HJB equations. Nevertheless, much effort is needed to verify the abstract conditions to apply this method for two time-scale stochastic processes. See, for instance, Peletier and Schlottke [35] and Kumar and Popovic [27], where [35] studied the LDP by this method for two time-scale system (X_t,Y_t) where (X_t) is a diffusion process over torus, and (Y_t) is a jumping process over a finite state space $\mathcal{S}$. [27] studied the LDP for two time-scale jump-diffusion processes. By Skorokhod's representation theorem for jumping processes, the regime-switching processes $(X_t^{\varepsilon,\alpha},Y_t^{\varepsilon,\alpha})$ can be viewed as a degenerate jump-diffusion processes (cf. [30, Section 3]). As noticed in [3, page 3], the requirement of Lipschitz continuity on the jump coefficients prevents the application of the LDP results in [27] to our current setting.

We first introduce some necessary notations before describing the idea of argument. Consider the two time-scale system $(X_t^{\varepsilon,\alpha},Y_t^{\varepsilon,\alpha})$ defined in (1.1), (1.2). For $h\in C_b(\mathbb{R}^d)$, let

$$u_{\varepsilon,\alpha}^h(t,x,i)=\varepsilon\log\mathbb{E}\Big[\exp\Big(\frac{h(X_t^{\varepsilon,\alpha})}{\varepsilon}\Big)\Big|X_0^{\varepsilon,\alpha}=x,Y_0^{\varepsilon,\alpha}=i\Big].\tag{2.1}$$

Let $\mathscr{A}_{\varepsilon,\alpha}$ be the infinitesimal generator of the process $(X_t^{\varepsilon,\alpha},Y_t^{\varepsilon,\alpha})$, that is,

$$\mathscr{A}_{\varepsilon,\alpha}f(x,i)=\langle b(x,i),\nabla f(x,i)\rangle+\frac{\varepsilon}{2}\mathrm{tr}(a(x,i)\nabla^2 f(x,i))+\frac{1}{\alpha}\sum_{j\in\mathcal{S}}q_{ij}(x)(f(x,j)-f(x,i))$$

for $f\in C_c^2(\mathbb{R}^d\times\mathcal{S})$, where $a(x,i)=\sigma(x,i)\sigma^*(x,i)$. Define a nonlinear operator $H_{\varepsilon,\alpha}$ by

$$H_{\varepsilon,\alpha}u(x,i)=\varepsilon\mathrm{e}^{-u/\varepsilon}\mathscr{A}_{\varepsilon,\alpha}\mathrm{e}^{u/\varepsilon}(x,i).\tag{2.2}$$

It can be shown (see [10]) that $u^h_{\varepsilon,\alpha}$ is a viscosity solution of the Cauchy problem:

$$\begin{aligned} \partial_t u &= H_{\varepsilon,\alpha} u, \quad \text{in } (0,T] \times \mathbb{R}^d \times \mathcal{S}, \\ u(0,x,i) &= h(x), \quad \text{for } x \in \mathbb{R}^d,\, i \in \mathcal{S}, \end{aligned} \tag{2.3}$$

where for each $f \in C^2(\mathbb{R}^d)$, $\nabla f(x)$ and $\nabla^2 f(x)$ denote the gradient and the Hessian matrix of f evaluated at x. We refer the reader to, e.g. [9], [11], for notions of viscosity solution, subsolution and supersolution to Hamilton-Jacobi equations.

Along the nonlinear semigroup method of [10] to study the LDP for $\{X^{\varepsilon,\alpha}_t; \varepsilon, \alpha > 0\}$, the proof consists of the following steps.

Step 1. Taking appropriate limits of $u^h_{\varepsilon,\alpha}$ to (2.3), we get upper semi-continuous and lower semi-continuous functions $\bar{u}^h$ and $\underline{u}^h$, respectively.

Step 2. Using an index set Λ, we construct a family of operators $H_0(\,\cdot\,;\beta)$ and $H_1(\,\cdot\,;\beta)$ for $\beta \in \Lambda$, such that $\bar{u}^h$ is a viscosity subsolution to the Cauchy problem for the operator $\inf_{\beta\in\Lambda} H_0(\,\cdot\,;\beta)$, and $\underline{u}^h$ a viscosity supersolution to the Cauchy problem for the operator $\sup_{\beta\in\Lambda} H_1(\,\cdot\,;\beta)$. In our setting, the operators $H_k(\,\cdot\,;\beta)$, $k = 0, 1$, will be constructed differently according to the different ratio ε/α as ε, α tend to 0.

Step 3. Find a suitable operator $\overline{H}_0(\,\cdot\,)$ satisfying

$$\inf_{\beta\in\Lambda} H_0(\,\cdot\,;\beta) \le \overline{H}_0(\,\cdot\,) \le \sup_{\beta\in\Lambda} H_1(\,\cdot\,;\beta)$$

and establish the comparison principle for the Hamilton-Jacobi equations associated with $\overline{H}_0(\,\cdot\,)$, which further implies the convergence of $u^h_{\varepsilon,\alpha}$ to the solution of the Cauchy problem for $\overline{H}_0(\,\cdot\,)$. Together with the exponential tightness of $\{X^{\varepsilon,\alpha}_t; \varepsilon, \alpha > 0\}$, it follows from Bryc's theorem that $\{X^{\varepsilon,\alpha}_t; \varepsilon, \alpha > 0\}$ satisfies a large deviation principle.

As in [9] and [10], the desired results in Step 1 and Step 2 can be proved by checking the following two families of conditions, Condition 3.1 and Condition 3.2. For the convenience of the readers, let us introduce these conditions in a general framework, and in the subsequent sections we shall verify these conditions in respective cases according to $\alpha = \varepsilon$ or $\varepsilon/\alpha \to \infty$.

Let

$$\begin{aligned} D_+ &= \{f : f(x) = \varphi(x) + \log(1 + |x|^2); \varphi \in C^2_c(\mathbb{R}^d)\}, \\ D_- &= \{f : f(x) = \varphi(x) - \log(1 + |x|^2); \varphi \in C^2_c(\mathbb{R}^d)\}. \end{aligned} \tag{2.4}$$

A collection of compact sets in $\mathbb{R}^d \times \mathcal{S}$ is defined by

$$\mathcal{Q} = \big\{K \times \widetilde{K}; K \subset\subset \mathbb{R}^d, \widetilde{K} \subset\subset \mathcal{S}\big\},$$

where $K \subset\subset \mathbb{R}^d$ means that K is a compact subset of $\mathbb{R}^d$. $\widetilde{K} \subset\subset \mathcal{S}$ is defined similarly.

Take the index set

$$\Lambda = \big\{\beta = (\xi, \theta); \xi \in C_c(\mathcal{S}), 0 < \theta < 1\big\}. \tag{2.5}$$

Let

$$\begin{aligned} D_{\varepsilon,+} &= \{f;\, f \in C^2(\mathbb{R}^d \times \mathcal{S}), f \text{ has compact finite level set}\}, \\ D_{\varepsilon,-} &= \{f;\, -f \in D_{\varepsilon,+}\}. \end{aligned} \tag{2.6}$$

Given two functions

$$H_k(x, p; \beta) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \quad \beta \in \Lambda,\, k = 0, 1,$$

define $H_0 f(x) = H_0(x, \nabla f(x))$ for $f \in D_+$, and $H_1 f(x) = H_1(x, \nabla f(x))$ for $f \in D_-$, where

$$H_0(x,p) = \inf_{\beta\in\Lambda} H_0(x,p;\beta), \quad H_1(x,p) = \sup_{\beta\in\Lambda} H_1(x,p;\beta). \tag{2.7}$$

Condition 3.1. For each $f \in D_+$ and $\beta \in \Lambda$, there exists $f_\varepsilon \in D_{\varepsilon,+}$ such that

1. for each $c > 0$, there exists $K \times \widetilde{K} \in \mathcal{Q}$ satisfying
$$\{(x,i);\, H_{\varepsilon,\alpha} f_\varepsilon(x,i) \geq -c\} \cap \{(x,i);\, f_\varepsilon(x,i) \leq c\} \subset K \times \widetilde{K};$$
2. for $K \times \widetilde{K} \in \mathcal{Q}$, $\lim_{\varepsilon\to 0} \sup_{(x,i)\in K\times\widetilde{K}} |f_\varepsilon(x,i) - f(x)| = 0$;
3. whenever $(x_\varepsilon, i) \in K \times \widetilde{K} \in \mathcal{Q}$ satisfies $x_\varepsilon \to x$ as $\varepsilon \to 0$,
$$\limsup_{(\varepsilon,\alpha)\to 0} H_{\varepsilon,\alpha} f_\varepsilon(x_\varepsilon, i) \leq H_0(x, \nabla f(x); \beta).$$

Condition 3.2. For each $f \in D_-$ and $\beta \in \Lambda$, there exists $f_\varepsilon \in D_{\varepsilon,-}$ such that

1. for each $c > 0$, there exists $K \times \widetilde{K} \in \mathcal{Q}$ satisfying
$$\{(x,i);\, H_{\varepsilon,\alpha} f_\varepsilon(x,i) \leq c\} \cap \{(x,i);\, f_\varepsilon(x,i) \geq -c\} \subset K \times \widetilde{K};$$
2. for each $K \times \widetilde{K} \in \mathcal{Q}$, $\lim_{\varepsilon\to 0} \sup_{(x,i)\in K\times\widetilde{K}} |f_\varepsilon(x,i) - f(x)| = 0$;
3. whenever $(x_\varepsilon, i) \in K \times \widetilde{K} \in \mathcal{Q}$ satisfying $x_\varepsilon \to x$,
$$\liminf_{(\varepsilon,\alpha)\to 0} H_{\varepsilon,\alpha} f_\varepsilon(x_\varepsilon, i) \geq H_1(x, \nabla f(x); \beta).$$

To establish the comparison principle, we shall use the following results established in [30]. Let P_t^x be the semigroup associated with the Q-matrix $(q_{ij}(x))$, and π^x its associated invariant probability measure provided it exists.

(A4) There exist constants $c_1, \lambda_1 > 0$ such that
$$\sup_{i\in\mathcal{S}} \|P_t^x(i,\cdot) - \pi^x\|_{\mathrm{var}} \leq c_1 \mathrm{e}^{-\lambda_1 t}, \quad t > 0,\ x \in \mathbb{R}^d.$$

(A5) There exist a function $\theta : \mathcal{S} \to (0,\infty)$, a decreasing function $\eta : [0,\infty) \to [0,2]$ satisfying $\int_0^\infty \eta_s \mathrm{d}s < \infty$ such that
$$\|P_t^x(i,\cdot) - \pi^x\|_{\mathrm{var}} \leq \theta(i)\eta_t, \quad t > 0, x \in \mathbb{R}^d,\ i \in \mathcal{S}.$$

Proposition 1 ([30])**.** (i) *Assume that (A2), (A3) and (A4) hold. Then $\mathbb{R}^d \ni x \mapsto \pi^x \in \mathscr{P}(\mathcal{S})$ is Lipschitz continuous, i.e.*
$$\|\pi^x - \pi^y\|_{\mathrm{var}} \leq C_\pi |x - y|, \quad x, y \in \mathbb{R}^d,$$
where $C_\pi = \frac{4c_1 K_3}{\lambda_1}$.

(ii) *Assume that (A2), (A3), (A5) hold. Then $x \mapsto \pi^x$ is $1/2$-Hölder continuous, i.e.*
$$\|\pi^x - \pi^y\|_{\mathrm{var}} \leq K_4 \sqrt{|x-y|}, \qquad x, y \in \mathbb{R}^d,$$
where $K_4 = \sqrt{K_3 (\inf_{i\in\mathcal{S}} \theta(i)) \int_0^\infty \eta_s \mathrm{d}s}$.

3. **LDP for the switching systems in the case $\varepsilon/\alpha = 1$.** In this section, we consider the LDP of $\{X_t^{\varepsilon,\alpha}; \varepsilon, \alpha > 0\}$ in the situation that $\alpha = \varepsilon \to 0$. We begin with checking Conditions 3.1 and 3.2. To this end, we introduce a drift condition to cope with the case that $\mathcal{S}$ is infinitely countable. Such kind of condition was used in the study of the LDP for Markov chains; see, e.g. [24].

(A6) Let ζ be a function from $\mathcal{S}$ to $[1, \infty)$. If $\mathcal{S}$ is infinitely countable, assume that $\{i \in \mathcal{S}; \zeta_i \le c\}$ is compact for every $c > 0$, and for every $x \in \mathbb{R}^d$, there exist constants $c_2, c_3 > 0$ and a compact set B in $\mathcal{S}$ such that

$$\mathrm{e}^{-\zeta_i} \sum_{j \in \mathcal{S}} q_{ij}(x)(\mathrm{e}^{\zeta_j} - \mathrm{e}^{\zeta_i}) \le -c_2 \zeta_i + c_3 \mathbf{1}_B(i), \quad i \in \mathcal{S}. \tag{3.1}$$

Define

$$\mathbb{T}(x) = \big\{(i,j) \in \mathcal{S} \times \mathcal{S};\ q_{ij}(x) > 0\big\} \tag{3.2}$$

for $x \in \mathbb{R}^d$, and

$$\underline{q}(x) = \inf_{(i,j) \in \mathbb{T}(x)} q_{ij}(x). \tag{3.3}$$

Theorem 3.1 (LDP in the case $\varepsilon/\alpha = 1$)**.** *Assume that conditions (A1)-(A3) and (A6) hold. Suppose that for each $R > 0$, $\inf_{|x| \le R} \underline{q}(x) > 0$ and for each $x \in \mathbb{R}^d$, $i \in \mathcal{S}$, there are finite number of $j \in \mathcal{S}$ such that $q_{ij}(x) > 0$. Then $\{X_t^{\varepsilon,\alpha}; \varepsilon, \alpha > 0\}$ given in (1.1), (1.2) with $\alpha = \varepsilon$ satisfies a LDP with speed $1/\varepsilon$ and rate function*

$$I(x, x_0, t) = \sup_{h \in C_b(\mathbb{R}^d)} \big\{h(x) - u^h(t, x_0)\big\}, \tag{3.4}$$

where u^h is the unique viscosity solution to the Hamilton-Jacobi equation (3.22) below.

To prove this theorem using the nonlinear semigroup method, we need to make some necessary preparation to check the conditions presented in Section 2. As we consider only $\alpha = \varepsilon$ in this part, we write simply $H_{\varepsilon,\alpha} = H_\varepsilon$ for the operator $H_{\varepsilon,\alpha}$ given in (2.2).

Verifying Condition 3.1. For each $f \in D_+$ and each $\beta = (\xi, \theta) \in \Lambda$ with Λ given in (2.5), let

$$g_i = (1-\theta)\xi_i + \theta \zeta_i, \ \text{ and } f_\varepsilon(x, i) = f(x) + \varepsilon g_i. \tag{3.5}$$

Then,

$$\begin{aligned} H_\varepsilon f_\varepsilon(x,i) = &\langle b(x,i), \nabla f(x)\rangle + \frac{1}{2}\langle a(x,i)\nabla f(x), \nabla f(x)\rangle + \frac{\varepsilon}{2}\mathrm{tr}(a(x,i)\nabla^2 f(x)) \\ &+ \mathrm{e}^{-g_i} Q(x) \mathrm{e}^{g}(i), \end{aligned}$$

where $\mathrm{e}^{-g_i} Q(x)\mathrm{e}^g(i) = \mathrm{e}^{-g_i} \sum_{j \in \mathcal{S}} q_{ij}(x)(\mathrm{e}^{g_j} - \mathrm{e}^{g_i})$. By Hölder's inequality and Young's inequality, as $q_{ii}(x) = -\sum_{j \ne i} q_{ij}(x)$,

$$\begin{aligned} \mathrm{e}^{-g_i} Q(x)\mathrm{e}^g(i) &= \sum_{j \ne i} q_{ij}(x) \mathrm{e}^{(1-\theta)(\xi_j - \xi_i) + \theta(\zeta_j - \zeta_i)} \\ &\le \Big(\sum_{j \ne i} q_{ij}(x)\mathrm{e}^{\xi_j - \xi_i}\Big)^{1-\theta} \Big(\sum_{j \ne i} q_{ij}(x)\mathrm{e}^{\zeta_j - \zeta_i}\Big)^{\theta} + q_{ii}(x) \\ &\le (1-\theta)\sum_{j \ne i} q_{ij}(x)\mathrm{e}^{\xi_j - \xi_i} + \theta \sum_{j \ne i} q_{ij}(x) \mathrm{e}^{\zeta_j - \zeta_i} + q_{ii}(x) \\ &= (1-\theta)\mathrm{e}^{-\xi_i} Q(x) \mathrm{e}^{\xi}(i) + \theta \mathrm{e}^{-\zeta_i} Q(x) \mathrm{e}^{\zeta}(i). \end{aligned}$$

Therefore,

$$H_\varepsilon f_\varepsilon(x,i) \le \langle b(x,i),\nabla f(x)\rangle + \frac{1}{2}\langle a(x,i)\nabla f(x),\nabla f(x)\rangle + \frac{\varepsilon}{2}\mathrm{tr}(a(x,i)\nabla^2 f(x)) \\ + (1-\theta)\mathrm{e}^{-\xi_i}Q(x)\mathrm{e}^{\xi}(i) + \theta\mathrm{e}^{-\zeta_i}Q(x)\mathrm{e}^{\zeta}(i). \tag{3.6}$$

Define the operator

$$H_0(x,p;\beta) = \sup_{i\in\mathcal{S}}\Big\{\langle b(x,i),p\rangle + \frac{1}{2}\langle a(x,i)p,p\rangle + (1-\theta)\mathrm{e}^{-\xi_i}Q(x)\mathrm{e}^{\xi}(i) \\ + \theta\mathrm{e}^{-\zeta_i}Q(x)\mathrm{e}^{\zeta}(i)\Big\}, \qquad \beta=(\xi,\theta)\in\Lambda, \tag{3.7}$$

and it holds that whenever $(x_\varepsilon, i)\in K\times\widetilde{K}\in\mathcal{Q}$ satisfying $x_\varepsilon\to x$,

$$\limsup_{\varepsilon\to 0} H_\varepsilon f_\varepsilon(x_\varepsilon,i) \le H_0(x,\nabla f(x);\beta). \tag{3.8}$$

In addition, by the definition of $f_\varepsilon(x,i)$ and (A6), for each $c>0$, there exists $K\times\widetilde{K}\in\mathcal{Q}$ such that

$$\{(x,i);\ f_\varepsilon(x,i)\le c\}\subset K\times\widetilde{K},$$

and further $\{(x,i); H_\varepsilon f_\varepsilon(x,i)\ge -c\}\cap\{(x,i); f_\varepsilon(x,i)\le c\}\subset K\times\widetilde{K}$. For each $K\times\widetilde{K}\in\mathcal{Q}$,

$$\lim_{\varepsilon\to 0}\sup_{(x,i)\in K\times\widetilde{K}}|f_\varepsilon(x,i)-f(x)| = \lim_{\varepsilon\to 0}\sup_{(x,i)\in K\times\widetilde{K}}\varepsilon|g_i| = 0.$$

Consequently, we have verified Condition 3.1.

Verifying Condition 3.2. For each $f\in D_-$ and $\beta=(\xi,\theta)\in\Lambda$, define

$$\bar{f}_\varepsilon(x,i) = f(x)+\varepsilon\bar{g}_i \quad \text{with } \bar{g}_i=(1+\theta)\xi_i-\theta\zeta_i,$$

then $\bar{f}_\varepsilon\in D_{\varepsilon,-}$ and $\bar{f}_\varepsilon$ converges uniformly to $\bar{f}$ on every $K\times\widetilde{K}\in\mathcal{Q}$. Moreover, (A6) implies that $\{(x,i);\bar{f}_\varepsilon(x,i)\ge -c\}$ is contained in some $K\times\widetilde{K}\in\mathcal{Q}$.

To check the last assertion in Condition 3.2, noting that

$$\frac{1}{1+\theta}\big((1+\theta)(\xi_j-\xi_i)-\theta(\zeta_j-\zeta_i)\big)+\frac{\theta}{1+\theta}(\zeta_j-\zeta_i)=\xi_j-\xi_i,$$

by Young's inequality $ab\le a^p/p+b^q/q$, we have

$$\begin{aligned}\sum_{j\ne i}q_{ij}(x)\mathrm{e}^{\xi_j-\xi_i} &= \sum_{j\ne i}q_{ij}(x)\mathrm{e}^{\frac{1}{1+\theta}\left((1+\theta)(\xi_j-\xi_i)-\theta(\zeta_j-\zeta_i)\right)+\frac{\theta}{1+\theta}(\zeta_j-\zeta_i)}\\ &\le \frac{1}{1+\theta}\sum_{j\ne i}q_{ij}(x)\mathrm{e}^{(1+\theta)(\xi_j-\xi_i)-\theta(\zeta_j-\zeta_i)}+\frac{\theta}{1+\theta}\sum_{j\ne i}q_{ij}(x)\mathrm{e}^{\zeta_j-\zeta_i}.\end{aligned}$$

This yields

$$\mathrm{e}^{-(1+\theta)\xi_i+\theta\zeta_i}\sum_{j\ne i}q_{ij}(x)\mathrm{e}^{(1+\theta)\xi_j-\theta\zeta_j}\ge(1+\theta)\sum_{j\ne i}q_{ij}(x)\mathrm{e}^{\xi_j-\xi_i}-\theta\sum_{j\ne i}q_{ij}(x)\mathrm{e}^{\zeta_j-\zeta_i},$$

and further

$$\begin{aligned}\mathrm{e}^{-(1+\theta)\xi_i+\theta\zeta_i}Q(x)\mathrm{e}^{(1+\theta)\xi-\theta\zeta}(i) &= \mathrm{e}^{-(1+\theta)\xi_i+\theta\zeta_i}\sum_{j\ne i}q_{ij}(x)\mathrm{e}^{(1+\theta)\xi_j-\theta\zeta_j}+q_{ii}(x)\\ &\ge(1+\theta)\mathrm{e}^{-\xi_i}Q(x)\mathrm{e}^{\xi}(i)-\theta\mathrm{e}^{-\zeta_i}Q(x)\mathrm{e}^{\zeta}(i).\end{aligned}$$

Hence,

$$\begin{aligned}H_\varepsilon \bar f_\varepsilon(x,i) &= \langle b(x,i),\nabla f(x)\rangle + \frac12\langle a(x,i)\nabla f(x),\nabla f(x)\rangle + \frac{\varepsilon}{2}\mathrm{tr}(a(x,i)\nabla^2 f(x))\\ &\quad + \mathrm{e}^{-(1+\theta)\xi_i+\theta\zeta_i}Q(x)\mathrm{e}^{(1+\theta)\xi-\theta\zeta}(i)\\ &\geq \langle b(x,i),\nabla f(x)\rangle + \frac12\langle a(x,i)\nabla f(x),\nabla f(x)\rangle + \frac{\varepsilon}{2}\mathrm{tr}(a(x,i)\nabla^2 f(x))\\ &\quad + (1+\theta)\mathrm{e}^{-\xi_i}Q(x)\mathrm{e}^{\xi}(i) - \theta\mathrm{e}^{-\zeta_i}Q(x)\mathrm{e}^{\zeta}(i).\end{aligned} \tag{3.9}$$

Define the operator

$$\begin{aligned}H_1(x,p;\beta) =&\inf_{i\in\mathcal S}\Big\{\langle b(x,i),p\rangle + \frac12\langle a(x,i)p,p\rangle + (1+\theta)\mathrm{e}^{-\xi_i}Q(x)\mathrm{e}^{\xi}(i)\\ &- \theta\mathrm{e}^{-\zeta_i}Q(x)\mathrm{e}^{\zeta}(i)\Big\},\qquad \beta=(\xi,\theta)\in\Lambda.\end{aligned} \tag{3.10}$$

Then it holds that whenever $(x_\varepsilon,i)\in K\times\widetilde K\in\mathcal Q$ with $x_\varepsilon\to x$,

$$\liminf_{\varepsilon\to 0} H_\varepsilon \bar f_\varepsilon(x_\varepsilon,i)\geq H_1(x,\nabla f(x);\beta). \tag{3.11}$$

Consequently, under the assumption (A6), for the constructed operators $H_0(x,p;\beta)$ in (3.7) and $H_1(x,p;\beta)$ in (3.10). Conditions 3.1 and 3.2 have been verified for $(X_t^{\varepsilon,\alpha},Y_t^{\varepsilon,\alpha})$ in the case $\alpha=\varepsilon$.

Let

$$H_0(x,p)=\inf_{\beta\in\Lambda}H_0(x,p;\beta),\qquad H_1(x,p)=\sup_{\beta\in\Lambda}H_1(x,p;\beta). \tag{3.12}$$

After verifying Conditions 3.1 and 3.2, according to [9, Lemma 4.1] or [27, Lemma 6], we can complete Step 1 and Step 2 as mentioned above in applying nonlinear semigroup method. We conclude these assertions in the following lemma.

Lemma 3.2. *Suppose (A6) holds. Let $u^h_{\varepsilon,\alpha}$ given by (2.1) be the viscosity solution to the Cauchy problem (2.3) for $h\in C_b(\mathbb R^d)$. Define*

$$\begin{aligned}u^h_\dagger &= \sup\big\{\limsup_{\varepsilon\to0}u^h_{\varepsilon,\varepsilon}(t_\varepsilon,x_\varepsilon,i);\ \exists\,(t_\varepsilon,x_\varepsilon,i)\in[0,T]\times K\times\widetilde K,\\ &\qquad (t_\varepsilon,x_\varepsilon)\to(t,x), K\times\widetilde K\in\mathcal Q\big\},\\ u^h_\ddagger &= \inf\big\{\liminf_{\varepsilon\to0}u^h_{\varepsilon,\varepsilon}(t_\varepsilon,x_\varepsilon,i);\ \exists\,(t_\varepsilon,x_\varepsilon,i)\in[0,T]\times K\times\widetilde K,\\ &\qquad (t_\varepsilon,x_\varepsilon)\to(t,x), K\times\widetilde K\in\mathcal Q\big\},\end{aligned}$$

and $\bar u^h$ the upper semicontinuous regularization of $u^h_\dagger$ and $\underline{u}^h$ the lower semicontinuous regularization of $u^h_\ddagger$. Then, $\bar u^h$ is a viscosity subsolution to

$$\partial_t u(t,x)\leq H_0(x,\nabla u(t,x)),\quad u(0,x)=h(x),$$

and $\underline{u}^h$ is a viscosity supersolution to

$$\partial_t u(t,x)\geq H_1(x,\nabla u(t,x)),\quad u(0,x)=h(x).$$

Proof. The proof is completely similar to [9, Lemma 4.1] after checking Conditions 3.1 and 3.2 under assumption (A6). The uniform boundedness of $u^h_{\varepsilon,\varepsilon}$, i.e. $\sup_{\varepsilon>0}\|u^h_{\varepsilon,\varepsilon}\|_\infty<\infty$, follows easily from the definition of $u^h_{\varepsilon,\alpha}$ and $h\in C_b(\mathbb R^d)$. □

Next, we proceed to Step 3 in applying the nonlinear semigroup method to establish the LDP for $\{X_t^{\varepsilon,\alpha};\alpha=\varepsilon>0\}$. Introduce the operator

$$\overline{H}_0(x,p)=\sup_{\mu\in\mathscr{P}(\mathcal{S})}\Big\{\sum_{i\in\mathcal{S}}\mu_i\langle b(x,i),p\rangle+\sum_{i\in\mathcal{S}}\mu_i\langle a(x,i)p,p\rangle-I_Q(\mu;x)\Big\}, \tag{3.13}$$

where

$$I_Q(\mu;x)=-\inf_{g\in D^{++}(Q^x)}\sum_{i\in\mathcal{S}}\frac{\mu_i Q(x)g(i)}{g_i}, \tag{3.14}$$

and $D^{++}(Q^x)$ denotes the subset of domain of operator $Q(x)$ which is strictly bounded from below by a positive constant; $\mathscr{P}(\mathcal{S})$ stands for the space of all probability measures over $\mathcal{S}$.

Proposition 2. *Assume (A6) holds. Then it holds*

$$H_0(x,p)=\overline{H}_0(x,p),\qquad x,\,p\in\mathbb{R}^d. \tag{3.15}$$

Proof. When $\mathcal{S}$ is a finite state space, clearly $\mathrm{e}^{\zeta}\in D^{++}(Q^x)$. When $\mathcal{S}$ is an infinite state space, by (A6), ζ has compact finite level set, and hence $\mathrm{e}^{\zeta}\in D^{++}(Q^x)$. Then,

$$\inf_{g\in D^{++}(Q^x)}\sum_{i\in\mathcal{S}}\frac{\mu_i Q(x)g(i)}{g_i}\le\sum_{i\in\mathcal{S}}\frac{\mu_i Q(x)\mathrm{e}^{\zeta}(i)}{\mathrm{e}^{\zeta_i}}.$$

Thus, the desired conclusion follows from [10, Lemma 11.35]. □

We proceed to show $H_1(x,p)\ge\overline{H}_0(x,p)$. For simplicity of notation, set

$$V_{x,p}(i)=\langle b(x,i),p\rangle+\frac{1}{2}\langle a(x,i)p,p\rangle. \tag{3.16}$$

Lemma 3.3. *Let $(\widetilde{Y}_t^x)_{t\ge0}$ be a continuous time Markov chain on $\mathcal{S}$ with transition rate matrix $(q_{ij}(x))$. For each $x,p\in\mathbb{R}^d$, if there exists a constant $C(x,p)$ such that*

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}_i\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]=C(x,p),\quad\forall\, i\in\mathcal{S}, \tag{3.17}$$

then

$$H_1(x,p)\ge\overline{H}_0(x,p),\qquad x,p\in\mathbb{R}^d.$$

Proof. According to [10, Theorem B.12], if there exists a constant C_V such that

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}_\nu\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]=C_V \tag{3.18}$$

for each $\nu\in\mathscr{P}(\mathcal{S})$ satisfying $\nu(\psi)>-\infty$ with $\psi_i:=\mathrm{e}^{-\zeta_i}Q(x)\mathrm{e}^{\zeta}(i)$, then

$$H_1(x,p)\ge\overline{H}_0(x,p),\quad x,p\in\mathbb{R}^d.$$

To prove this lemma, we only need to show that condition (3.17) implies condition (3.18).

Denote by P_t^V the Feynman-Kac semigroup

$$P_t^V f(i)=\mathbb{E}_i\Big[f(\widetilde{Y}_t^x)\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big],\qquad f\in C_b(\mathcal{S}). \tag{3.19}$$

Firstly, let us show that (3.18) holds with $C_V=C(x,p)$ for $\nu\in\mathscr{P}(\mathcal{S})$ in the form

$$\nu=\sum_{k=1}^m\lambda_k\delta_{i_k},\qquad\text{satisfying }\lambda_k\in[0,1],\ \sum_{k=1}^m\lambda_k=1,\text{and } i_k\in\mathcal{S},\ m\in\mathbb{N}.$$

Indeed,

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}_\nu\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]&=\lim_{t\to\infty}\frac{1}{t}\log\sum_{k=1}^m\lambda_k P_t^V\mathbf{1}(i_k)\\&\le\lim_{t\to\infty}\max_{1\le k\le m}\frac{1}{t}\log P_t^V\mathbf{1}(i_k)\\&=C(x,p).\end{aligned}$$

On the other hand,

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}_\nu\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]\ge\lim_{t\to\infty}\min_{1\le k\le m}\frac{1}{t}\log P_t^V\mathbf{1}(i_k)=C(x,p).$$

Secondly, for a general $\nu\in\mathscr{P}(\mathcal{S})$, there exists a sequence of $\mu_n\in\mathscr{P}(\mathcal{S})$ in the form

$$\mu_n=\sum_{k=1}^{m_n}\lambda_k^n\delta_{i_k^n},\quad \lambda_k^n\in[0,1],\ \sum_{k=1}^{m_n}\lambda_k^n=1,\ i_k^n\in\mathcal{S}$$

such that μ_n converges weakly to ν. For each fixed $x,p\in\mathbb{R}^d$, $i\mapsto V_{x,p}(i)$ is a bounded function on $\mathcal{S}$ due to (A1), then it follows from the weak convergence of μ_n to ν that

$$\lim_{n\to\infty}\mathbb{E}_{\mu_n}\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]=\lim_{n\to\infty}\sum_{i\in\mathcal{S}}\mu_n(i)P_t^V\mathbf{1}(i)=\mathbb{E}_\nu\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big].$$

Moreover, for any $\epsilon>0$, there exists $K>0$ such that for all $n>K$,

$$(1-\epsilon)\mathbb{E}_{\mu_n}\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]\le\mathbb{E}_\nu\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]\le(1+\epsilon)\mathbb{E}_{\mu_n}\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big].$$

Therefore, by the arbitrariness of ϵ, we obtain

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}_\nu\Big[\mathrm{e}^{\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s}\Big]=C(x,p),$$

which completes the proof. □

Remark 1. (i) In Lemma 3.3, we assume the convergence of log-moment generating functions in (3.17) for the Markov chains $(\widetilde{Y}_t^x)_{t\ge0}$ associated with $(q_{ij}(x))_{i,j\in\mathcal{S}}$ for every $x\in\mathbb{R}^d$. This condition can be used to prove large deviations estimates for $\int_0^t F(\widetilde{Y}_s^x)\mathrm{d}s$ by standard large deviations techniques (cf. [5, Chapter 6]). Especially, when $\mathcal{S}$ is a finite state space, the limit (3.17) always exists for conservative, irreducible Markov chains on $\mathcal{S}$ (cf. [5, Theorem 6.3.8, Remark p.275]).

(ii) When $\mathcal{S}$ is an infinite state space, the limit (3.17) corresponds to a type of multiplicative mean ergodic theorem, and many works have been devoted to this topic. For example, Wu [42] proved the limit exists supposing the logarithmic Sobolev inequality holds. Kontoyiannis and Meyn [23, 24] studied the multiplicative ergodic theory using the spectral theory, which can yield the existence of the limit in (3.17). Their tools rely on the geometric ergodicity or drift conditions. When V is unbounded, [13] and [39] investigated this limit under the help of dimension-free Harnack inequality.

Next, when $\mathcal{S}$ is an infinite state space, based on assumption (A6), we shall introduce a proposition to guarantee the existence of limit in (3.17) according to the investigation in [24]. The results in [24] are stated in the setting of discrete-time Markov chains, which are also valid for the continuous-time Markov chains analogous to the discussion in [23].

Lemma 3.4. *Assume (A2) and (A6) hold. Then, there exist constants $\varepsilon_0 > 0$, $0 < \eta_0 \le 1$, $\Lambda_{x,p}$, a function $\tilde f : \mathcal{S} \to \mathbb{R}$, constants b_0, $b_1 > 0$ such that*

$$\Big|\mathbb{E}_i \exp\Big[\int_0^t V_{x,p}(\widetilde{Y}_s^x)\mathrm{d}s - t\Lambda_{x,p}\Big] - \tilde f(i)\Big| \le \varepsilon_0 \mathrm{e}^{\eta_0\zeta_i + b_1 - b_0 t}, \quad i \in \mathcal{S}. \tag{3.20}$$

This yields that (3.17) holds with the constant $C(x,p) = \Lambda_{x,p}$.

Proof. Notice first that condition (A2) means that the Markov chain $(\widetilde{Y}_t^x)$ is irreducible and aperiodic. For any function $W : \mathcal{S} \to (0,\infty]$, define the Banach space

$$L_\infty^W = \Big\{g : \mathcal{S} \to \mathbb{R};\ \sup_{i\in\mathcal{S}} \frac{|g_i|}{W_i} < \infty\Big\}$$

endowed with the norm

$$\|F\|_W = \sup_{i\in\mathcal{S}} \frac{|F_i|}{W_i}.$$

By virtue of [24, Theorem 3.4], if there is a function $W : \mathcal{S} \to [1,\infty)$, a small set $C \subset \mathcal{S}$, and constants $\delta > 0$, $b < \infty$, such that

$$\mathrm{e}^{-\zeta} Q(x) \mathrm{e}^{\zeta} \le -\delta W + b\mathbf{1}_C, \tag{3.21}$$

then there exist constants $\varepsilon_0 > 0$, $0 < \eta_0 \le 1$, b_0, $b_1 > 0$ and a function $\tilde f : \mathcal{S} \to \mathbb{R}$ such that for any $F \in L_\infty^W$ satisfying $\|F\|_W \le \varepsilon_0$ we have

$$\Big|\mathbb{E}_i\Big[\exp\Big(\int_0^t F(\widetilde{Y}_s^x)\mathrm{d}s - t\Lambda(F)\Big)\Big] - \tilde f_i\Big| \le \|F\|_W \mathrm{e}^{\eta_0\zeta_i + b_1 - b_0 t}, \quad i \in \mathcal{S},$$

where $\Lambda(F)$ is a constant depending only on F and independent of the initial state $i \in \mathcal{S}$. Therefore, to establish (3.20), we only need to find suitable function W on $\mathcal{S}$, a small set $C \subset \mathcal{S}$, constants $\delta > 0$, $b < \infty$ such that (3.21) holds and the function $V_{x,p}$ satisfying $\|V_{x,p}\|_W < \varepsilon_0$.

By (A6) and the boundedness of $V_{x,p} : \mathcal{S} \to \mathbb{R}$, we have $\lim_{i\to\infty} |V_{x,p}(i)|/\zeta_i = 0$, and hence there exists an $N_0 \in \mathbb{N}$ such that $\sup_{i\in\mathcal{S}, i>N_0} |V_{x,p}(i)|/\zeta_i < \varepsilon_0$. Moreover, there exists a constant $k_0 > 0$ such that

$$\max\big\{|V_{x,p}(i)|/(\zeta_i + k_0);\ 1 \le i \le N_0\big\} < \varepsilon_0.$$

Define our desired function W by

$$W_i = \zeta_i + k_0 \mathbf{1}_{\widetilde{C}}(i), \quad i \in \mathcal{S},$$

where $\widetilde{C} = \{1, \dots, N_0\} \subset \mathcal{S}$. Then

$$\|V_{x,p}\|_W < \varepsilon_0.$$

Thanks to [31, Theorem 6.0.1] and [31, Theorem 5.5.7], the irreducible and aperiodic property of the Markov chain $(\widetilde{Y}_t^x)$ yield that the compact set $\widetilde{C} \cup B$ is a small set, and it follows from (A6) that

$$\mathrm{e}^{-\zeta} Q(x) \mathrm{e}^{\zeta} \le -c_2 W + c_2 k_0 \mathbf{1}_{\widetilde{C}} + c_3 \mathbf{1}_B \le -c_2 W + (c_2 k_0 + c_3)\mathbf{1}_{\widetilde{C}\cup B}.$$

Consequently, (3.21) and further (3.20) hold by [24, Theorem 3.4]. □

Invoking Lemma 3.4 and Lemma 3.3, we obtain that:

Proposition 3. *Let conditions (A2) and (A6) hold. Then,*

$$\overline{H}_0(x,p) \le H_1(x,p), \quad x, p \in \mathbb{R}^d.$$

Now we go to establish the comparison principle for the Hamilton-Jacobi equation

$$\partial_t u = \overline{H}_0(x, \nabla u), \qquad u(0,x) = h(x), \tag{3.22}$$

i.e. for any viscosity subsolution $u_0(t,x)$ and any viscosity supersolution $u_1(t,x)$ to (3.22), it holds

$$u_0(t,x) \le u_1(t,x), \qquad (t,x) \in [0,T] \times \mathbb{R}^d.$$

To this end, we shall use the comparison principle established in [16], and provide appropriate conditions on the coefficients of $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ to verify the required conditions in [16]. According to [16, Theorem 2] and the discussion following [16, Theorem 1], we introduce a general comparison principle for a Hamilton-Jacobi equation. The conditions and the Hamiltonian H there have been simplified to adapt to our current setting.

Lemma 3.5 (Ishii [16]). *Consider the Hamilton-Jacobi equation*

$$\partial_t u + H(x, \nabla u(t,x)) = 0, \quad \text{in } (0,T) \times \mathbb{R}^d, \tag{3.23}$$

where $T > 0$ is a given constant. Assume the following conditions hold.

(H1) The function $(x,p) \mapsto H(x,p)$ is continuous on $\mathbb{R}^d \times \mathbb{R}^d$.

(H2) For each $R > 0$ it holds

$$\lim_{r\to\infty} \inf \big\{ H(x,p); x \in B_R(0), |p| \ge r \big\} = \infty,$$

where $B_R(0)$ denotes the closed ball centered at 0 with radius R.

(H3) The function $p \mapsto H(x,p)$ is convex and

$$H(x,0) \le 0, \quad \text{and} \quad H\Big(x, -\frac{\delta x}{|x|^2+1}\Big) \le C$$

for all x, t and for some $\delta, C > 0$.

Let u and v be respectively a viscosity subsolution and a viscosity supersolution to (3.23). Assume that $u(0,x) \le v(0,x)$ for all $x \in \mathbb{R}^d$, and that $\inf_{\mathbb{R}^d \times [0,T)} v > -\infty$. Then $u \le v$ on $\mathbb{R}^d \times [0,T)$, that is, the comparison principle holds for the equation (3.23).

Theorem 3.6 (Comparison principle). *Assume (A1)-(A3) hold and $\inf_{|x|\le R} \underline{q}(x) > 0$ for every $R > 0$. Suppose that for each $x \in \mathbb{R}^d$, each $i \in \mathcal{S}$, there are only finite number of $j \in \mathcal{S}$ such that $q_{ij}(x) > 0$. In addition, suppose that there exists $\eta : [0,\infty) \to (0,\infty)$ such that for each $R > 0$*

$$\langle a(x,i)p, p\rangle \ge \eta(R)|p|^2, \quad \forall\, x \in B_R(0),\ i \in \mathcal{S},\ p \in \mathbb{R}^d.$$

Then the comparison principle holds for the Hamilton-Jacobi equation (3.22), which implies that the viscosity solution to (3.22) is unique.

Proof. We shall prove this theorem by Lemma 3.5, and hence we need to verify the conditions (H1)-(H3) of Lemma 3.5.

By virtue of the expression of $\overline{H}_0(x,p)$ in (3.13), due to (A1), it is clear that $p \mapsto \overline{H}_0(x,p)$ is convex, $\overline{H}_0(x,0) = 0$ by noting $Q(x)\mathbf{1} = 0$, and

$$\overline{H}_0\Big(x, \frac{-2x}{|x|^2+1}\Big) \le 2K_2.$$

Therefore, condition (H3) of Lemma 3.5 is satisfied.

In order to check (H1) and (H2) of Lemma 3.5, the key point is to show the uniform continuity of $x \mapsto I_Q(\mu; x)$ w.r.t. $\mu \in \mathscr{P}(S)$. For $R > 0$, $B_R(0)$ stands for the closed ball in $\mathbb{R}^d$ centered at 0 with radius R. By the hypothesis of this theorem, for every $x \in \mathbb{R}^d$ and every $i \in \mathcal{S}$, there are only finite number of $j \in \mathcal{S}$ such that $q_{ij}(x) > 0$, which means that for every measurable function h on $\mathcal{S}$,

$$\lim_{t \downarrow 0} \frac{P_t^x h(i) - h(i)}{t} = Q(x)h(i), \qquad i \in \mathcal{S},$$

and hence h is in the weak domain $D(Q^x)$ of generator $Q(x)$. This means that $D(Q^x) = \mathscr{B}(\mathcal{S})$, and hence

$$\begin{aligned} I_Q(\mu; x) &= - \inf_{g \in D^{++}(Q^x)} \sum_{i,j \in \mathcal{S}, i \neq j} \mu_i q_{ij}(x) \Big(\frac{g_j}{g_i} - 1 \Big) \\ &= \sup_{g \in \mathscr{B}(\mathcal{S}), g > 0} \sum_{i,j \in \mathcal{S}, i \neq j} \mu_i q_{ij}(x) \Big(1 - \frac{g_j}{g_i} \Big). \end{aligned} \tag{3.24}$$

This deduces that $D^{++}(Q^x) = D^{++}(Q^y)$ for $x, y \in \mathbb{R}^d$ and $x \neq y$. For each $x \in B_R(0)$, for any $\epsilon > 0$, there exists a measurable function g^ϵ with $g^\epsilon > 0$ such that

$$0 \leq I_Q(\mu; x) \leq \sum_{i,j \in \mathcal{S}, i \neq j} \mu_i q_{ij}(x) \Big(1 - \frac{g_j^\epsilon}{g_i^\epsilon} \Big) + \epsilon.$$

This yields

$$\sum_{i,j \in \mathcal{S}, i \neq j} \mu_i q_{ij}(x) \frac{g_j^\epsilon}{g_i^\epsilon} \leq \sum_{i,j \in \mathcal{S}, i \neq j} \mu_i q_{ij}(x) + \epsilon \leq \kappa + \epsilon,$$

where constant κ is given in (A2), and further that

$$\sum_{(i,j) \in \mathbb{T}(x)} \mu_i \frac{g_j^\epsilon}{g_i^\epsilon} \leq \frac{1}{\underline{q}(x)} (\kappa + \epsilon). \tag{3.25}$$

Then, for any $y \in B_R(0)$, by (A3) and (3.25),

$$\begin{aligned} & I_Q(\mu; x) - I_Q(\mu; y) \\ & \leq \epsilon + \sum_{i,j \in \mathcal{S}, i \neq j} |q_{ij}(x) - q_{ij}(y)| \mu_i \big(1 + \frac{g_j^\epsilon}{g_i^\epsilon} \big) \\ & \leq \epsilon + K_3 |x - y| \Big(1 + \sum_{(i,j) \in \mathbb{T}(x)} \mu_i \frac{g_j^\epsilon}{g_i^\epsilon} \Big) \\ & \leq \epsilon + K_3 |x - y| + K_3 \frac{\kappa + \epsilon}{\inf_{|x| \leq R} \underline{q}(x)} |x - y|. \end{aligned}$$

By the arbitrariness of ϵ and the symmetric position of x and y, there exists a positive constant $C(R)$ independent of $\mu \in \mathscr{P}(\mathcal{S})$ such that

$$|I_Q(\mu; x) - I_Q(\mu; y)| \leq C(R) |x - y|, \qquad x, y \in B_R(0). \tag{3.26}$$

By virtue of the definition of $\overline{H}_0(x, p)$ in (3.13), for x, y, p, $p' \in \mathbb{R}^d$ and for any $\epsilon > 0$ there exists $\mu_\epsilon \in \mathscr{P}(\mathcal{S})$ such that

$$\begin{aligned} & \overline{H}_0(x, p) - \overline{H}_0(y, p') \\ & \leq \sum_{i \in \mathcal{S}} \mu_i^\epsilon \big[\langle b(x, i), p \rangle + \langle a(x, i) p, p \rangle - \langle b(y, i), p' \rangle - \langle a(y, i) p', p' \rangle \big] \\ & \quad - I_Q(\mu^\epsilon; x) + I_Q(\mu^\epsilon; y) + \epsilon. \end{aligned}$$

By (A1), (3.26), the arbitrariness of ϵ, and the symmetric position of x, y, we can derive from this estimate that $(x,p)\mapsto \overline{H}_0(x,p)$ is continuous, i.e. (H1) of Lemma 3.5 holds.

By (A2) and (3.24), $I_Q(\mu,x)\le\kappa$, and hence

$$\begin{aligned}&\lim_{r\to\infty}\inf\big\{\overline{H}_0(x,p);\ x\in B_R(0),|p|\ge r\big\}\\&\ge\lim_{r\to\infty}\inf\big\{\sup_{\mu\in\mathscr{P}(\mathcal{S})}\big\{-K_2|p|+\eta(R)|p|^2-I_Q(\mu;x)\big\};\ x\in B_R(0),|p|\ge r\big\}\\&=\infty.\end{aligned}$$

So we have verified (H2) of Lemma 3.5. Consequently, we obtain from Lemma 3.5 that the comparison principle holds for the Hamilton-Jacobi equation (3.22). □

Proof of Theorem 3.1. Invoking Lemma 3.2, Proposition 2, Proposition 3 and Theorem 3.6, and applying [9, Lemma 4.2], we obtain that

$$\bar u^h=\underline{u}^h=u^h,\ \text{and}\ \lim_{\varepsilon\to0}\sup_{t\in[0,T]}\sup_{(x,i)\in K\times\widetilde{K}}|u^h_{\varepsilon,\varepsilon}(t,x,i)-u^h(t,x)|=0 \tag{3.27}$$

for any $K\times\widetilde{K}\in\mathcal{Q}$, where u^h is the unique viscosity solution to the Hamilton-Jacobi equation (3.22).

To see $\{X_t^{\varepsilon,\alpha};\varepsilon,\alpha>0\}$ is exponential tight, let $f(x)=\log(1+|x|^2)$. Then $f(x)\to\infty$ as $|x|\to\infty$, and $\sup_{x\in\mathbb{R}^d}\{|\nabla f(x)|_\infty+\|\nabla^2 f(x)\|\}<\infty$. For any $c>0$ there exists a compact set $K_c\subset\mathbb{R}^d$ such that $f(x)>c$ for any $x\notin K_c$. Taking $f_\varepsilon(x,i)=f(x)$ for $x\in\mathbb{R}^d$, $i\in\mathcal{S}$, it holds

$$\begin{aligned}H_\varepsilon f_\varepsilon(x,i)&=\varepsilon\mathrm{e}^{-f_\varepsilon/\varepsilon}\mathscr{A}_\varepsilon\mathrm{e}^{f_\varepsilon/\varepsilon}(x,i)\\&=\langle b(x,i),\nabla f(x)\rangle+\frac12\langle a(x,i)\nabla f(x),\nabla f(x)\rangle+\frac{\varepsilon}{2}\mathrm{tr}(a(x,i)\nabla^2 f(x)).\end{aligned}$$

By (A1), there exists a constant $\tilde C>0$ such that

$$\sup_{x\in\mathbb{R}^d,i\in\mathcal{S}}H_\varepsilon f_\varepsilon(x,i)\le\tilde C.$$

Since

$$M_t:=\exp\Big\{f(X_t^{\varepsilon,\alpha})-f(X_0^\varepsilon)-\int_0^t H_\varepsilon f(X_s^{\varepsilon,\alpha},Y_s^{\varepsilon,\alpha})\mathrm{d}s\Big\}$$

is a nonnegative local martingale, it follows from the choice of K_c and the estimate above that

$$\mathbb{P}(X_t^{\varepsilon,\alpha}\notin K_c)\mathrm{e}^{(c-f(x_0)-\tilde Ct)/\varepsilon}\le\mathbb{E}\Big[\mathrm{e}^{f(X_t^{\varepsilon,\alpha})-f(X_0^\varepsilon)-\int_0^t H_\varepsilon f(X_s^{\varepsilon,\alpha},Y_s^{\varepsilon,\alpha})\mathrm{d}s}\Big]\le1.$$

Hence, $\varepsilon\log\mathbb{P}(X_t^{\varepsilon,\alpha}\notin K_c)\le t\tilde C+f(x_0)-c$. Since $\tilde C$ is fixed and independent of c, this means that $\{X_t^{\varepsilon,\alpha};\varepsilon,\alpha>0\}$ is exponentially tight.

Consequently, applying Bryc's theorem, $\{X_t^{\varepsilon,\alpha};\varepsilon=\alpha>0\}$ satisfies the LDP with rate $1/\varepsilon$ and good rate function $I(x,x_0,t)$ given in (3.4). □

Remark 2. In Theorem 3.1, when $\mathcal{S}$ is finite, (A6) is a trivial condition, and one only needs to pay attention to the restriction $\inf_{|x|\le R}\underline{q}(x)>0$ for each $R>0$. This type of condition on $\underline{q}(x)$ is also used in [3, Assumption 2.3], where it is assumed that $\inf_{x\in\mathbb{R}^d}\underline{q}(x)>0$, which is a little stronger than ours.

4. **The LDP in the case $\varepsilon/\alpha \to \infty$.** In this section, we go to investigate the LDP for the system $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ when the time scales satisfying $\lim_{(\varepsilon,\alpha)\to 0} \varepsilon/\alpha = \infty$. We still use the notations Λ, D_+, D_-, $D_{\varepsilon,+}$, $D_{\varepsilon,-}$ given in (2.4), (2.5) and (2.6). Instead of (A6), we shall use the following drift condition in this part.

(A7) Let $\tilde\zeta : \mathcal{S} \to [0,\infty)$. If $\mathcal{S}$ is infinitely countable, suppose that $\{i; \tilde\zeta_i \le c\}$ is compact for every $c > 0$, and for each $x \in \mathbb{R}^d$ there are constants $c_4, c_5 > 0$ and a compact set $\tilde B \subset \mathcal{S}$ such that

$$Q(x)\tilde\zeta(i) \le -c_4\tilde\zeta_i + c_5\mathbf{1}_{\tilde B}(i), \quad \text{for any } i \in \mathcal{S}. \tag{4.1}$$

According to [24, Proposition 2.1], the condition (4.1) in (A7) is weaker than the condition (3.1) in (A6).

Verifying Condition 3.1. For $\beta = (\xi,\theta) \in \Lambda$, and $f \in D_+$, let

$$f_\alpha(x,i) = f(x) + \alpha g_i, \quad \text{with } g_i = \xi_i + \theta\tilde\zeta_i, \;\; x \in \mathbb{R}^d, i \in \mathcal{S}. \tag{4.2}$$

Then $f_\alpha \in D_{\varepsilon,+}$ and f_α converges to f as $\alpha \to 0$ uniformly on compact sets. Moreover, (A7) yields that for every $c > 0$, $\{(x,i); f_\alpha(x,i) \le c\}$ is contained in some $K \times \widetilde K \in \mathcal{Q}$.

Next, a direct calculation yields

$$\begin{aligned} H_{\varepsilon,\alpha} f_\alpha(x,i) &= \varepsilon \mathrm{e}^{-\frac{f_\alpha}{\varepsilon}} \mathscr{A}_{\varepsilon,\alpha} \mathrm{e}^{\frac{f_\alpha}{\varepsilon}}(x,i) \\ &= \langle b(x,i), \nabla f(x)\rangle + \frac12 \langle a(x,i)\nabla f(x), \nabla f(x)\rangle + \frac{\varepsilon}{2}\mathrm{tr}(a(x,i)\nabla^2 f(x)) \\ &\quad + \frac{\varepsilon}{\alpha}\sum_{j\ne i} q_{ij}(x)\big(\mathrm{e}^{\frac{\alpha}{\varepsilon}(g_j-g_i)} - 1\big). \end{aligned} \tag{4.3}$$

Under the assumption (A2), for each $i \in \mathcal{S}$ and $x \in \mathbb{R}^d$, there is only finite number of items in the summation $\sum_{j\ne i} q_{ij}(x)\big(\mathrm{e}^{\frac{\alpha}{\varepsilon}(g_j-g_i)} - 1\big)$. Noting further that $\frac{\alpha}{\varepsilon} \to 0$ as $(\varepsilon,\alpha) \to 0$, we obtain

$$\lim_{\Gamma\ni(\varepsilon,\alpha)\to 0} \frac{\varepsilon}{\alpha}\sum_{j\ne i} q_{ij}(x)\big(\mathrm{e}^{\frac{\alpha}{\varepsilon}(g_j-g_i)} - 1\big) = \sum_{j\ne i} q_{ij}(x)(g_j - g_i). \tag{4.4}$$

Here and in the remainder of this subsection, we use the notation $\Gamma \ni (\varepsilon,\alpha) \to 0$ to represent taking the limit as $(\varepsilon,\alpha) \to 0$ satisfying at the same time $\varepsilon/\alpha \to \infty$.

Different to $H_0(x,p;\beta)$ defined in (3.7) in the case $\varepsilon = \alpha$, now we define the operator $H_0(x,p;\beta)$ for $\beta = (\xi,\theta) \in \Lambda$ by

$$\begin{aligned} &H_0(x,p;\beta) \\ &= \sup_{i\in\mathcal{S}} \big\{\langle b(x,i),p\rangle + \frac12\langle a(x,i)p,p\rangle + Q(x)\xi(i) - \theta c_4\tilde\zeta_i + \theta c_5 \mathbf{1}_{\tilde B}(i)\big\}, \end{aligned} \tag{4.5}$$

where $\tilde\zeta$ is given in (A7). Then, we derive from (A7) and (4.4) that

$$\begin{aligned} \limsup_{\Gamma\ni(\varepsilon,\alpha)\to 0} H_{\varepsilon,\alpha} f_\alpha(x,i) &\le \langle b(x,i), \nabla f(x)\rangle + \frac12\langle a(x,i)\nabla f(x), \nabla f(x)\rangle \\ &\quad + \sum_{j\ne i} q_{ij}(x)(\xi_j - \xi_i) - \theta c_4\tilde\zeta_i + \theta c_5\mathbf{1}_{\tilde B}(i) \\ &\le H_0(x, \nabla f(x); \beta). \end{aligned} \tag{4.6}$$

Therefore, all the three assertions in Condition 3.1 have been checked.

Verifying Condition 3.2. For $\beta=(\xi,\theta)\in\Lambda$, let $\tilde f\in D_-$ and $g_i=\xi_i-\theta\tilde\zeta_i$, $i\in\mathcal S$ with $\xi\in C_c(\mathcal S)$. Let $\tilde f_\alpha(x,i)=\tilde f(x)+\alpha g_i$. Define

$$\begin{aligned} &H_1(x,p;\beta)\\ &=\inf_{i\in\mathcal S}\big\{\langle b(x,i),p\rangle+\frac12\langle a(x,i)p,p\rangle+Q(x)\xi(i)+\theta c_4\tilde\zeta_i-\theta c_5\mathbf 1_{\tilde B}(i)\big\}. \end{aligned}\tag{4.7}$$

Then,

- $\tilde f_\alpha\in D_{\varepsilon,-}$, and $\tilde f_\alpha$ converges uniformly to $\tilde f$ on every $K\times\widetilde K\in\mathcal Q$ as $\alpha\to0$.
- For $c>0$, $\{(x,i);\tilde f_\alpha(x,i)\ge -c\}$ is contained in some $K\times\widetilde K\in\mathcal Q$ due to (A7).
- Completely similar to (4.6), one gets

$$\liminf_{\Gamma\ni(\varepsilon,\alpha)\to0}H_{\varepsilon,\alpha}\tilde f_\alpha(x,i)\ge H_1(x,\nabla\tilde f(x);\beta).\tag{4.8}$$

Hence, Condition 3.2 has been checked.

Lemma 4.1. *When $\mathcal S$ is infinitely countable, under (A2), either condition (A7) or condition (A6) implies condition (A5).*

Proof. As mentioned above, (A7) is weaker than (A6), hence we only need to prove (A5) from (A7). By virtue of [32, Theorem 6.1], it follows from (A7) that there exist constants $R,\lambda>0$ such that

$$\|P_t^x(i,\cdot)-\pi^x\|_{\rm var}\le R(1+\tilde\zeta_i)\mathrm e^{-\lambda t},\quad t>0,i\in\mathcal S,x\in\mathbb R^d.$$

The constants R,λ may depend on the function $\tilde\zeta$, and c_4, c_5 in (A7). It is clear that $\mathrm e^{-\lambda t}$ is integrable over $[0,\infty)$, and hence (A5) holds with $\theta_i=R(1+\tilde\zeta_i)$ and $\eta_t=\mathrm e^{-\lambda t}$. □

Lemma 4.2. *Suppose that (A1)-(A3) and (A7) hold. Then the Markov chain associated with $(q_{ij}(x))$ admits a unique invariant probability measure π^x, which is 1/2-Hölder continuous from $\mathbb R^d$ to $\mathscr P(\mathcal S)$.*

Proof. When $\mathcal S$ is finite, since $(q_{ij}(x))$ is irreducible, $(q_{ij}(x))$ admits a unique invariant measure π^x. When $\mathcal S$ is infinitely countable, by Lemma 4.1, the Markov chain associated with $(q_{ij}(x))$ is ergodic and has a unique invariant probability measure π^x. Moreover, as shown in the argument of Lemma 4.1, (A7) yields that (A5) holds with $\theta_i:=R(1+\tilde\zeta_i)$. Furthermore, we can even get from (A7) that

$$\begin{aligned} Q(x)\theta(i)&=RQ(x)\tilde\zeta(i)\\ &\le -Rc_4\tilde\zeta_i+Rc_5\mathbf 1_{\tilde B}(i)\\ &=-c_4\theta(i)+Rc_4+Rc_5\mathbf 1_{\tilde B}(i), \end{aligned}$$

which implies that the condition (2.7) in [30, Theorem 2.4] is also satisfied. Consequently, according to Proposition 1(ii), we have that $x\mapsto\pi^x$ is 1/2-Hölder continuous. □

Let

$$H_0(x,p)=\inf_{\beta\in\Lambda}H_0(x,p;\beta),\qquad H_1(x,p)=\sup_{\beta\in\Lambda}H_1(x,p;\beta).\tag{4.9}$$

Define

$$\overline H_0(x,p)=\langle\bar b(x),p\rangle+\frac12\langle\bar a(x)p,p\rangle,\tag{4.10}$$

where

$$\bar{b}(x) = \sum_{i\in\mathcal{S}} \pi_i^x b(x,i), \quad \bar{a}(x) = \sum_{i\in\mathcal{S}} \pi_i^x a(x,i), \quad x\in\mathbb{R}^d.$$

Lemma 4.3. *Suppose (A1)-(A3) and (A7) hold. If $\mathcal{S}$ is infinitely countable, assume further that for every $x,p\in\mathbb{R}^d$ the solution $(\tilde{\xi}_j(x,p))_{j\in\mathcal{S}}$ to the following Poisson equation is bounded:*

$$\sum_{j\neq i} q_{ij}(x)\big(\tilde{\xi}_j(x,p)-\tilde{\xi}_i(x,p)\big) = \eta_i(x,p), \qquad i\in\mathcal{S}, \tag{4.11}$$

where

$$\eta_i(x,p) = \langle \bar{b}(x),p\rangle + \frac{1}{2}\langle \bar{a}(x)p,p\rangle - \big(\langle b(x,i),p\rangle + \frac{1}{2}\langle a(x,i)p,p\rangle\big).$$

Then,

$$H_0(x,p)\le \overline{H}_0(x,p)\le H_1(x,p), \qquad x,p\in\mathbb{R}^d. \tag{4.12}$$

Proof. (i) Consider first the case that $\mathcal{S}$ is a finite state space. By the definition of $\bar{b}(x)$ and $\bar{a}(x)$, it holds

$$\sum\nolimits_{i\in\mathcal{S}} \pi_i^x \eta_i(x,p) = 0, \quad x,p\in\mathbb{R}^d. \tag{4.13}$$

The Fredholm alternative theorem yields that for every $x,p\in\mathbb{R}^d$ there exists a solution $(\tilde{\xi}_i(x,p))_{i\in\mathcal{S}}$ to the Poisson equation (4.11), which is unique up to adding a constant. The solution $(\tilde{\xi}_i(x,p))_{i\in\mathcal{S}}$ to (4.11) obviously belongs to $C_c(\mathcal{S})$ since $\mathcal{S}$ is compact itself. Due to the definitions in (4.9), we have

$$\begin{aligned}
H_0(x,p) &\le \inf_{0<\theta<1} H_0\big(x,p;(\tilde{\xi}(x,p),\theta)\big)\\
&= \inf_{0<\theta<1}\sup_{i\in\mathcal{S}}\big\{\langle \bar{b}(x),p\rangle + \frac{1}{2}\langle \bar{a}(x)p,p\rangle - \theta c_4\tilde{\zeta}_i + \theta c_5 \mathbf{1}_{\tilde{B}}(i)\big\}\\
&\le \inf_{0<\theta<1}\sup_{i\in\mathcal{S}}\big\{\langle \bar{b}(x),p\rangle + \frac{1}{2}\langle \bar{a}(x)p,p\rangle + \theta c_5 \mathbf{1}_{\tilde{B}}(i)\big\}\\
&= \overline{H}_0(x,p).
\end{aligned}$$

Similarly,

$$H_1(x,p)\ge \sup_{0<\theta<1} H_1\big(x,p;(\tilde{\xi}(x,p),\theta)\big)\ge \overline{H}_0(x,p).$$

As a consequence,

$$H_0(x,p)\le \overline{H}_0(x,p)\le H_1(x,p), \qquad x,p\in\mathbb{R}^d.$$

(ii) Now consider the case $\mathcal{S}=\{1,2,\ldots\}$. The equality (4.13) still holds in this case, and hence the solution $(\tilde{\xi}_i(x,p))_{i\in\mathcal{S}}$ to (4.11) exists. However, a new difficulty comes from the fact that $(\tilde{\xi}_i(x,p))_{i\in\mathcal{S}}$ may not belong to $C_c(\mathcal{S})$. Therefore, we need to use the truncation method to define

$$\tilde{\xi}_i^m(x,p) = \tilde{\xi}_i(x,p)\mathbf{1}_{i\le m} \quad \text{for } m\in\mathbb{N}.$$

Then,

$$Q(x)\tilde{\xi}^m(x,p)(i) = \begin{cases} Q(x)\tilde{\xi}(x,p)(i) - \sum_{j>m} q_{ij}(x)\tilde{\xi}_j(x,p), & \text{if } i\le m,\\ Q(x)\tilde{\xi}(x,p)(i) + q_i(x)\tilde{\xi}_i(x,p) - \sum\limits_{j>m,j\neq i} q_{ij}(x)\tilde{\xi}_j(x,p), & \text{if } i>m. \end{cases}$$

According to the assumption that $(\tilde\xi_i(x,p))_{i\in\mathcal S}$ is bounded, there exists a positive constant $C_0(x,p)$ such that

$$\sup_{i\in\mathcal S}|\tilde\xi_i(x,p)|\le C_0(x,p),\quad i\in\mathcal S. \tag{4.14}$$

Then, by (A2) and (A7), for all $m\ge 1$,

$$\begin{aligned}&\lim_{i\to\infty}Q(x)\tilde\xi^m(x,p)(i)-\theta c_4\tilde\zeta_i\\ &\le\lim_{i\to\infty}\sum_{j\neq i}q_{ij}(x)\sup_{k,l\in\mathcal S}|\tilde\xi_k(x,p)-\tilde\xi_l(x,p)|-\theta c_4\tilde\zeta_i\\ &=-\infty.\end{aligned} \tag{4.15}$$

Combining this with the estimate

$$\langle b(x,i),p\rangle+\frac12\langle a(x,i)p,p\rangle\le K_2|p|+\frac12K_2^2|p|^2,$$

there exists $i'\in\mathcal S$ such that for all $m\ge1$ and for all $i\in\mathcal S$ with $i\ge i'$,

$$\begin{aligned}&\langle b(x,i),p\rangle+\frac12\langle a(x,i)p,p\rangle+Q(x)\tilde\xi^m(x,p)(i)-\theta c_4\tilde\zeta_i\\ &\le\langle b(x,1),p\rangle+\frac12\langle a(x,1)p,p\rangle+Q(x)\tilde\xi^m(x,p)(1)-\theta c_4\tilde\zeta_1\end{aligned}$$

for $x,p\in\mathbb R^d$, $\theta\in(0,1)$. Hence, for all $m\ge1$,

$$\begin{aligned}&\sup_{i\in\mathcal S}\big\{\langle b(x,i),p\rangle+\frac12\langle a(x,i)p,p\rangle+Q(x)\tilde\xi^m(x,p)(i)-\theta c_4\tilde\zeta_i+\theta c_5\mathbf 1_{\tilde B(i)}\big\}\\ &=\sup_{i\le i'}\big\{\langle b(x,i),p\rangle+\frac12\langle a(x,i)p,p\rangle+Q(x)\tilde\xi^m(x,p)(i)-\theta c_4\tilde\zeta_i+\theta c_5\mathbf 1_{\tilde B(i)}\big\}.\end{aligned}$$

Invoking the definition of $H_0(x,p)$,

$$\begin{aligned}&H_0(x,p)\\ &\le\liminf_{m\to\infty}\inf_{0<\theta<1}H_0\big(x,p;(\tilde\xi^m(x,p),\theta)\big)\\ &\le\liminf_{m\to\infty}\inf_{0<\theta<1}\sup_{i\le i'}\big\{\langle b(x,i),p\rangle+\frac12\langle a(x,i)p,p\rangle+Q(x)\tilde\xi^m(x,p)(i)+\theta c_5\mathbf 1_{\tilde B(i)}\big\}\\ &\le\liminf_{m\to\infty}\inf_{0<\theta<1}\sup_{i\le i'}\big\{\langle\bar b(x),p\rangle+\frac12\langle\bar a(x)p,p\rangle-\sum_{j>m}q_{ij}(x)\tilde\xi_j(x,p)+\theta c_5\mathbf 1_{\tilde B(i)}\big\}\\ &\le\langle\bar b(x),p\rangle+\frac12\langle\bar a(x)p,p\rangle\\ &=\overline H_0(x,p).\end{aligned} \tag{4.16}$$

Similarly, we can show that

$$H_1(x,p)\ge\limsup_{m\to\infty}\sup_{0<\theta<1}H_1\big(x,p;(\tilde\xi^m(x,p),\theta)\big)\ge\overline H_0(x,p). \tag{4.17}$$

Combining (4.17) with (4.16), we finally get (4.12) and complete the proof. □

Lemma 4.4 (Comparison principle)**.** *Assume (A1)-(A3) and (A7) hold. In addition, suppose that there exists $\eta:[0,\infty)\to(0,\infty)$ such that for each $R>0$*

$$\langle a(x,i)p,p\rangle\ge\eta(R)|p|^2,\quad\forall\,x\in B_R(0),\ i\in\mathcal S,\ p\in\mathbb R^d.$$

Then the comparison principle holds for the Hamilton-Jacobi equation

$$\begin{aligned} \partial_t u &= \overline{H}_0(x, \nabla u) \quad \text{in } (0,T)\times\mathbb{R}^d, \\ u(0,x) &= h(x) \qquad\quad\; \text{in } \mathbb{R}^d, \end{aligned} \tag{4.18}$$

where $h \in C_b(\mathbb{R}^d)$.

Proof. By Lemma 4.3, $x \mapsto \pi^x$ is 1/2-Hölder continuous. Together with (A1), we obtain that $\bar{b}$, $\bar{a}$ are also 1/2-Hölder continuous, and further

$$(x,p) \mapsto \overline{H}_0(x,p) = \langle \bar{b}(x), p\rangle + \frac{1}{2}\langle \bar{a}(x)p, p\rangle \;\text{ is continuous.}$$

Clearly, $p \mapsto \overline{H}_0(x,p)$ is convex by the nonnegative definiteness of $\bar{a}$. Hence, (H1) of Lemma 3.5 has been checked. (H3) of Lemma 3.5 is obvious by the boundedness of $\bar{b}$ and $\bar{a}$ due to (A1). By (A1), $|\bar{b}(x)| \le K_2$,

$$\begin{aligned} &\lim_{r\to\infty} \inf\{\overline{H}(x,p);\; x \in B_R(0), |p| \ge r\} \\ &\ge \lim_{r\to\infty} \inf\{-K_2|p| + \eta(R)|p|^2; |p| \ge r\} \\ &= \infty. \end{aligned}$$

This implies that condition (H2) of Lemma 3.5 holds. Consequently, the comparison principle holds for the Hamilton-Jacobi equation (4.18). □

Theorem 4.5 (LDP in the case $\varepsilon/\alpha \to \infty$)**.** *Suppose that the assumptions of Lemma 4.3 hold and that* $\varepsilon/\alpha \to \infty$ *as* $\varepsilon, \alpha \to 0$. *In addition, suppose there exists* $\eta : [0,\infty) \to (0,\infty)$ *such that for each* $R > 0$

$$\langle a(x,i)p, p\rangle \ge \eta(R)|p|^2, \quad \forall\, x \in B_R(0),\; i \in \mathcal{S},\; p \in \mathbb{R}^d.$$

Then $\{X_t^{\varepsilon,\alpha}; \varepsilon, \alpha > 0\}$ *satisfies the LDP with rate* $1/\varepsilon$ *and good rate function*

$$I(x, x_0, t) = \sup_{h \in C_b(\mathbb{R}^d)} \big\{h(x) - u^h(t, x_0)\big\}, \tag{4.19}$$

where u^h *is the unique viscosity solution to* (4.18).

Proof. Following the approach of [10], it follows from Lemmas 4.1 and 4.3 that

$$\lim_{\Gamma \ni (\varepsilon,\alpha)\to 0} \sup_{t\in[0,T]} \sup_{(x,i)\in K\times\widetilde{K}} |u^h_{\varepsilon,\alpha}(t,x,i) - u^h(t,x)| = 0$$

for any $K \times \widetilde{K} \in \mathcal{Q}$, where u^h is the unique viscosity solution to (4.18). Since the exponential tightness of $\{X_t^{\varepsilon,\alpha}; \varepsilon, \alpha > 0\}$ has been proved in Theorem 4.5, we get that $\{X_t^{\varepsilon,\alpha}; \varepsilon, \alpha > 0\}$ satisfying $\lim_{(\varepsilon,\alpha)\to 0} \varepsilon/\alpha = \infty$ satisfies the LDP with rate $1/\varepsilon$ and rate function $I(x, x_0, t)$ given in (4.19) by virtue of Bryc's theorem. □

In Theorem 4.5, we impose a condition on the boundedness of the solution to the Poisson equation (4.11) when $\mathcal{S} = \{1, 2, \ldots\}$, which is needed to establish the comparison principle. There are a lot of investigations on Poisson equations, but it is not an easy task to provide explicit conditions to ensure the boundedness of the solution. This is an interesting topic to be studied later.

In the following, for a special class of Markov chains, single-birth processes, we provide an explicit solution to the Poisson equation, hence its boundedness can be checked immediately. Moreover, see [4, Theorem 1.1] for the explicit solutions to more general Poisson equations for single-birth processes.

Example 1. Assume that $(q_{ij}(x))$ is a conservative Q-matrix satisfying $q_{i(i+1)}(x) > 0$, $q_{ik}(x) = 0$ for all $i \in \mathcal{S}$ and $k \geq i+2$. Let

$$F_i^{(i)} = 1, \quad F_n^{(i)} = \frac{1}{q_{n(n+1)}(x)} \sum_{k=i}^{n-1} q_n^{(k)}(x) F_k^{(i)}, \quad n > i \geq 1,$$

$$\text{where} \quad q_n^{(k)}(x) := \sum_{j=1}^{k} q_{nj}(x), \quad 1 \leq k < n.$$

Then for a function f on $\mathcal{S}$, the solution g to the Poisson equation

$$Q(x)g = f$$

has the following explicit representation in terms of $(q_{ij}(x))$

$$g_n = g_1 + \sum_{1 \leq k \leq n-1} \sum_{1 \leq j \leq k} \frac{F_k^{(j)} f_j}{q_{j(j+1)}(x)}, \qquad n \geq 1.$$

Example 2. Consider the following two time-scale system:

$$\begin{aligned} \mathrm{d}X_t^{\varepsilon,\alpha} &= b(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})\mathrm{d}t + \sqrt{\varepsilon}\mathrm{d}W_t, \\ X_0^{\varepsilon,\alpha} &= x_0 \in \mathbb{R}, \qquad Y_0^{\varepsilon,\alpha} = i_0 \in \mathcal{S}, \end{aligned} \tag{4.20}$$

where

$$b(x,i) = \mathrm{e}^{-i} \sin x, \qquad i \in \mathcal{S},\ x \in \mathbb{R},$$

$\mathcal{S} = \{1, 2, \ldots\}$ is an infinite state space, $(Y_t^{\varepsilon,\alpha})$ is a jumping process on $\mathcal{S}$ satisfying (1.2) and for every $i \geq 2$, $q_{ii}(x) = -\sum_{j \in \mathcal{S}} q_{ij}(x)$,

$$q_{ii+1}(x) = 1 + \cos x, \quad q_{ii-1}(x) = i^2 - \cos x, \quad q_{ij}(x) = 0,\ \forall\, j \notin \{i, i-1, i+1\},$$

and $-q_{11}(x) = q_{12}(x) = 1 + \cos x$. It is clear that conditions (A1)-(A3) hold.

Taking

$$\zeta_i = \log(i+3), \quad i \in \mathcal{S},$$

we have that for some $C > 0$, $K \in \mathbb{N}$,

$$\begin{aligned} &\mathrm{e}^{-\zeta_i} \sum_{j \in \mathcal{S}} q_{ij}(x)(\mathrm{e}^{\zeta_j} - \mathrm{e}^{\zeta_i}) \\ &= \frac{1}{i+3}\big(q_{ii+1}(x) - q_{ii-1}(x)\big) \\ &= \frac{1}{i+3}\big(1 + \cos x - i^2 + \cos x\big) \\ &\leq -\log(i+3) + C\mathbf{1}_{\{i \leq K\}}, \qquad i \in \mathcal{S}. \end{aligned}$$

Consequently, (A6) has been checked. Therefore, Theorem 3.1 implies that $(X_t^{\varepsilon,\alpha})$ satisfies a LDP when $\alpha = \varepsilon \to 0$.

REFERENCES

[1] J. Bao, Q. Song, G. Yin and C. Yuan, Ergodicity and strong limit results for two-time-scale functional stochastic differential equations, *Stoch. Anal. Appl.*, **35** (2017), 1030-1046.

[2] C. Bezuidenhout, A large deviations principle for small perturbations of random evolution equations, *Ann. Probab.*, **15** (1987), 646-658.

[3] A. Budhiraja, P. Dupuis and A. Ganguly, Large deviations for small noise diffusions in a fast Markovian environment, *Electron. J. Probab.*, **23** (2018), Paper No. 112, 33 pp.

[4] M. Chen and Y. Zhang, Unified representation of formulas for single birth processes, *Front. Math. China*, **9** (2014), 761-796.

[5] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications, 2nd ed.*, Springer, New York, 1998.

[6] A. Eizenberg and M. Freidlin, Large deviations for Markov processes corresponding to PDE systems, *Ann. Probab.*, **21** (1993), 1015-1044.

[7] A. Faggionato, D. Gabrielli and M. Crivellari, Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors, *Markov Process. Relat. Fields*, **16** (2010), 497-548.

[8] J. Feng, M. Forde and J. Fouque, Short maturity asymptotics for a fast mean reverting Heston stochastic volatility model, *SIAM J. Finac. Math.*, **1** (2010), 126-141.

[9] J. Feng, J. Fouque and R. Kumar, Small-time asymptotics for fast mean-reverting stochastic volatility models, *Ann. Appl. Probab.*, **22** (2012), 1541-1575.

[10] J. Feng and T. G. Kurtz, *Large Deviation for Stochastic Processes, in: Mathematical Surveys and Monographs*, vol. 131, American Mathematical Society, 2006.

[11] W. Fleming and H. Soner, *Controlled Markov Processes and Viscosity Solutions, 2nd ed.*, Stochastic Modelling and Applied Probability 25, Springer, New York, 2006.

[12] M. Freidlin and T. Lee, Large deviation principle for the diffusion-transmutation processes and dirichlet problem for PDE systems with small parameter, *Probab. Theory Relat. Fields*, **105** (1996), 227-254.

[13] F. Gao, Long time asymtotics of unbounded additive functionals of Markov processes, *Electron. J. Probab.*, **22** (2017), Paper No. 94, 21 pp.

[14] Q. He, G. Yin and Q. Zhang, Large deviations for two-time-scale systems driven by nonhomogeneous Markov chains and associated optimal control problems, *SIAM J. Control Optim.*, **49** (2011), 1737-1765.

[15] G. Huang, M. Mandjes and P. Spreij, Large deviations for Markov-modulated diffusion processes with rapid switching, *Stoch. Proc. Appl.*, **126** (2016), 1785-1818.

[16] H. Ishii, Comparison results for Hamilton-Jacobi equations without growth condition on solutions from above, *Applicable Analysis*, **67** (1997), 357-372.

[17] H. Kang and T. Kurtz, Separation of time-scales and model reduction for stochastic reaction networks, *Ann. Appl. Probab.*, **23** (2013), 529-583.

[18] H. Kang, T. Kurtz and L. Popovic, Central limit theorems and diffusion approximations for multiscale Markov chain models, *Ann. Appl. Probab.*, **24** (2014), 721-759.

[19] R. Khasminskii, On an averaging principle for Itô stochastic differential equations, *Kibernetica*, **4** (1968), 260-279 (in Russian).

[20] R. Khasminskii and G. Yin, On averaging principles: An asymptotic expansion approach, *SIAM J. Math. Anal.*, **35** (2004), 1534-1560.

[21] R. Khasminskii and G. Yin, Limit behavior of two-time-scale diffusions revisited, *J. Differ. Equations*, **212** (2005), 85-113.

[22] Y. Kifer, *Large Deviations and Adiabatic Transitions for Dynamical Systems and Markov Processes in Fully Coupled Averaging*, Memoirs of the Amer. Math. Soc., 944, AMS, Providence, RI, 2009.

[23] I. Kontoyiannis and S. Meyn, Spectral theory and limit theorems for geometrically ergodic Makrov processes, *Ann. Appl. Probab.*, **13** (2003), 304-362.

[24] I. Kontoyiannis and S. Meyn, Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes, *Electron. J. Probab.*, **10** (2005), 61-123.

[25] R. Kraaij and M. Schlottke, A large deviation principle for Markovian slow-fast systems, `arXiv:2011.05686v2`, 2021.

[26] R. Kraaij and M. Schlottke, Comparison principle for Hamilton-Jacobi-Bellman equations via a bootstrapping procedure, *Nonlinear Differ. Equ. Appl.*, **28** (2021), Paper No. 22, 45 pp.

[27] R. Kumar and L. Popovic, Large deviations for multi-scale jump-diffusion processes, *Stoch. Proc. Appl.*, **127** (2017), 1297-1320.

[28] R. Liptser, Large deviations for two scaled diffusions, *Probab. Theory Relat. Fields,* **106** (1996), 71-104.

[29] W. Liu, M. Rockner, X. Sun and Y. Xie, Averaging principle for slow-fast stochastic differential equations with time dependent locally Lipschitz coefficients, *J. Differ. Equations*, **268** (2020), 2910-2948.

[30] Y. Mao and J. Shao, Averaging principle for two time-scale regime-switching processes, *Electron. J. Probab.*, **29** (2024), Paper No. 14, 21 pp.

[31] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability,* Springer Verlag, 1993.

[32] S. Meyn and R. Tweedie, Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes, *Adv. Appl. Probab.*, **25** (1993), 518-548.

[33] E. Pardoux and A. Veretennikov, On the Poisson equation and diffusion approximation. I, *Ann. Probab.*, **29** (2001), 1061-1085.

[34] E. Pardoux and A. Veretennikov, On Poisson equation and diffusion approximation 2, *Ann. Probab.*, **31** (2003), 1166-1192.

[35] M. Peletier and M. Schlottke, Large-deviation principles of switching Markov processes via Hamilton-Jacobi equations, `arXiv:1901.08478v3`, 2021.

[36] L. Popovic, Large deviations of Markov chains with multiple time-scales, *Stoch. Proc. Appl.*, **129** (2019), 3319-3359.

[37] A. Puhalskii, On large deivations of coupled diffusions with time scale separation, *Ann. Probab.*, **44** (2016), 3111-3186.

[38] J. Shao, Strong solutions and strong Feller properties for regime-switching diffusion processes in an infinite state space, *SIAM J. Control Optim.*, **53** (2015), 2462-2479.

[39] J. Shao and S. Wang, Harnack inequality and long time asymptotics of unbounded additive functionals of regime-switching diffusion processes, *Potential Analysis*, **56** (2022), 549-570.

[40] A. Veretennikov, On large deviations in the averaging principle for SDEs with a "full depedence", *Ann. Probab.*, **27** (1999), 284-296.

[41] A. Veretennikov, On large deviations in the averaging principle for SDEs with a "full depedence" revisited, *Discrete Contin. Dyn. Syst. Ser. B*, **18** (2013), 523-549.

[42] L. Wu, Uniformly integrable operators and large deviations for Markov processes, *J. Funct. Anal.*, **172** (2000), 301-376.

[43] G. Yin and C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Stochastic Modeling and Applied Probability, Vol. 63, Springer, New York, 2010.

Received February 2025; revised April 2025; early access May 2025.