Large deviations for the empirical measure and empirical flow of Markov renewal processes with a countable state space[*]

Chen Jia[†] ⬤   Da-Quan Jiang[‡]   Bingjie Wu[§]

Abstract

Here we propose the Donsker-Varadhan-type compactness conditions and prove the joint large deviation principle for the empirical measure and empirical flow of Markov renewal processes (semi-Markov processes) with a countable state space, generalizing the relevant results for continuous-time Markov chains with a countable state space obtained in [Ann. Inst. H. Poincaré Probab. Statist. 51, 867-900 (2015)] and [Stoch. Process. Appl. 125, 2786-2819 (2015)], as well as the relevant results for Markov renewal processes with a finite state space obtained in [Adv. Appl. Probab. 48, 648-671 (2016)]. In particular, our results hold when the flow space is endowed with either the bounded weak* topology or the strong L^1 topology. Even for continuous-time Markov chains, our compactness conditions are weaker than the ones proposed in previous papers. Furthermore, under some stronger conditions, we obtain the explicit expression of the marginal rate function of the empirical flow.

[†]Applied and Computational Mathematics Division, Beijing Computational Science Research Center, Beijing, China. E-mail: chenjia@csrc.ac.cn

[‡]LMAM, School of Mathematical Sciences, Peking University, Beijing, China. Center for Statistical Science, Peking University, Beijing, China. E-mail: jiangdq@math.pku.edu.cn

[§]Applied and Computational Mathematics Division, Beijing Computational Science Research Center, Beijing, China. LMAM, School of Mathematical Sciences, Peking University, Beijing, China. E-mail: wubingjie@pku.edu.cn

1 Introduction

Semi-Markov processes, which can be viewed as a direct extension of discrete-time and continuous-time Markov chains, are one of the most important classes of non-Markov processes. They have attracted considerable attention in recent years and have found wide applications in physics, chemistry, biology, finance, and engineering [25, 3, 34]. The embedded chain of a semi-Markov process is a discrete-time Markov chain, while the waiting times may not be exponentially distributed. Such non-exponential waiting time distributions have been found in many scientific problems such as molecular motors [33, 21], enzyme kinetics [40, 42], gene networks [39, 28], and cell cycle dynmics [26, 29]. The representation of a semi-Markov process in terms of its embedding chain and waiting times is also called a Markov renewal process.

The mathematical theory of large deviations was initiated by Cramér [13], and was later developed by many mathematicians and physicists. In the pioneering work [16, 17, 18, 19], Donsker and Varadhan have established the large deviation principle (LDP) for the empirical measure and for the empirical process associated with a large class of discrete-time and continuous-time Markov processes. The large deviations for the sample mean, empirical measure, and empirical processes are usually said to be at level 1, level 2, and level 3, respectively [15]. For discrete-time and continuous-time Markov chains, the number of jumps along each oriented edge of the transition graph per unit time is called the empirical flow. The large deviations for the empirical flow are often said to be at level 2.5 [31], since it is between level 2 and level 3. For a discrete-time Markov chain $(X_n)_{n \geq 0}$, the large deviations for the empirical flow can be obtained directly from those for the empirical measure since the binary process $(X_n, X_{n+1})_{n \geq 0}$ is also a discrete-time Markov chain. However, things become much more complicated in the continuous-time case. For continuous-time Markov chains, Fortelle [14] proved a weak joint LDP for the empirical measure and empirical flow and obtained the corresponding rate function. Subsequently, Bertini et al. [6, 7] proposed some Donsker-Varadhan-type compactness conditions and proved the full joint LDP for the empirical measure and empirical flow.

Since semi-Markov processes are direct generalizations of discrete-time and continuous-time Markov chains, a natural question is whether the large deviations at different levels can be extended to semi-Markov processes. Along this line, Mariani and Zambotti [35] proved the joint LDP for the empirical measure and empirical flow of semi-Markov processes with a *finite* state space and expressed the corresponding rate function in terms of relative entropy. The large deviations for the empirical flow of Markov and semi-Markov processes have also been applied in statistical mechanics to study the fluctuation relations of thermodynamical systems far from equilibrium [1, 2, 6, 27, 11, 23, 20, 9, 24, 30]. Up till now, it is still unclear whether these finite state space results can be generalized to semi-Markov processes with a *countable* state space. The aim of the present paper is to fill in this gap.

Here we investigate the joint large deviations for the empirical measure and empirical flow of semi-Markov processes with a countable state space. When the state space has an infinite number of states, the choice of the topology of the flow space, i.e. the value space of all empirical flows, will become very important. Following [6, 7], we consider two types of topology of the flow space: the bounded weak* topology and the strong topology. Specifically, we propose two different Donsker-Varadhan-type compactness conditions for the two types of topology, and prove the corresponding LDP for the empirical measure and empirical flow. This is the first main contribution of this paper. In the special case of continuous-time Markov chains, our compactness condition reduces to the one proposed in [7] when the flow space is endowed with the bounded weak* topology. However, when

the flow space is endowed with the strong topology, our compactness condition is even weaker than the one proposed in [6].

In [35], Mariani and Zambotti also provided the explicit expression of the marginal rate function for the empirical measure, but they have not given the explicit expression of the marginal rate function for the empirical flow. Here we propose two strong compactness conditions, one for the embedded chain and one for the waiting time distributions, and we provide a variational representation for the marginal rate function for the empirical flow under these compactness conditions. The relationship between these compactness conditions and the geometric ergodicity of the embedded chain is also clarified. This is the second main contribution of this paper.

The present paper is organized as follows. In Section 2, we recall the definitions of Markov renewal processes and semi-Markov processes, as well as the definitions of the empirical measure and empirical flow. In Section 3, we propose two different compactness conditions when the flow space is endowed with the bounded weak* topology and the strong topology. Moreover, we also state the main results including the joint and marginal LDPs for the empirical measure and empirical flow. The proof of the LDP for the two types of topology will be given in Sections 4 and 5, respectively. In Section 6, we derive the marginal rate function for the empirical flow.

2 Preliminaries

2.1 Markov renewal processes and semi-Markov processes

Let V be a countable set endowed with the discrete topology and the associated Borel σ-algebra is the collection of all subsets of V. The set $[0, \infty]$ is equipped with the topology that is compatible with the natural topology on $[0, \infty)$ so that $(s, \infty]$ is open for any $s \geq 0$. For any Polish space $\mathcal{X}$, let $\mathcal{P}(\mathcal{X})$ denote the collection of Borel probability measures on $\mathcal{X}$. For any $\mu \in \mathcal{P}(\mathcal{X})$ and $f \in L^1(\mu)$, let $\langle \mu, f \rangle$ or $\mu(f)$ denote the integral of f with respect to μ. The set $\mathcal{P}(\mathcal{X})$ is equipped with the topology of weak convergence and the associated Borel σ-algebra.

We first recall the definition of Markov renewal processes, also called $(J\text{-}X)$-processes [34]. Here, we adopt the definition in [35] and [3, Chapter VII.4].

Definition 2.1. *The process* $(X, \tau) = \{(X_k)_{k \geq 0}, (\tau_k)_{k \geq 1}\}$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is called a* Markov renewal process *if*

(a) $X = (X_k)_{k \geq 0}$ *is a discrete-time time-homogeneous Markov chain with countable state space* V *and transition probability matrix* $P = (p_{xy})_{x,y \in V}$.

(b) $\tau = (\tau_k)_{k \geq 1}$ *is a sequence of positive and finite random variables such that conditioned on* $(X_k)_{k \geq 0}$*, the random variables* $(\tau_k)_{k \geq 1}$ *are independent and have distribution*

$$\mathbb{P}\left(\tau_{i+1} \in \cdot \mid (X_k)_{k \geq 0}\right) = \psi_{X_i, X_{i+1}}(\cdot),$$

where $\psi_{xy} \in \mathcal{P}(0, \infty)$ *for any* $x, y \in V$*. The matrix* $\Psi = (\psi_{xy})_{x,y \in V}$ *is called the* waiting time matrix*. Note that it is a matrix of probability measures. The pair* (P, Ψ) *is called the* transition kernel*.*

We next recall the definition of semi-Markov processes.

Definition 2.2. *For any* $t \geq 0$ *and* $n \geq 1$*, let*

$$S_n = \sum_{i=1}^{n} \tau_i, \quad N_t = \sum_{n=1}^{\infty} 1_{(S_n \leq t)} = \inf\{n \geq 0 : S_{n+1} > t\},$$

where $\inf \emptyset := \infty$ *and* 1_A *is the indicator function of the set* A*. The process* $\xi = (\xi_t)_{t \geq 0}$ *with* $\xi_t := X_{N_t}$ *is called the* semi-Markov process *associated with the Markov renewal*

process (X, τ). Clearly, ξ is a jump process whose trajectories are right-continuous on the state space V.

Clearly, S_n represents the nth jump time of the semi-Markov process ξ, N_t represents the number of jumps of ξ up to time t, and $X = (X_k)_{k \geq 0}$ is the embedded chain of ξ. It is easy to see that $N_t = n$ if and only if $S_n \leq t < S_{n+1}$. In particular, if all the waiting times are equal to 1, i.e. $\psi_{xy} = \delta_1$ for any $x, y \in V$ with δ_1 being the point mass at 1, then ξ reduces to a discrete-time Markov chain. If all the waiting times are exponentially distributed, i.e. $\psi_{xy}(\mathrm{d}t) = q_x e^{-q_x t} \mathrm{d}t$ for any $x, y \in V$, then ξ reduces to a continuous-time Markov chain. In the following, we do not distinguish the Markov renewal process (X, τ) and the associated semi-Markov process ξ since they are totally equivalent.

The transition diagram of the embedded chain X is a directed graph (V, E), where the edge set

$$E = \{(x, y) \in V \times V \; : \; p_{xy} > 0\}$$

is composed of all directed edges with positive transition probabilities. Throughout this paper, we impose the following basic assumptions on the Markov renewal process (X, τ).

Assumption 2.3. *The embedded chain X is irreducible.*

Assumption 2.4. *The embedded chain X is recurrent.*

Assumption 2.5. *The waiting time distribution only depends on the current state, i.e. $\psi_{xy} = \psi_x$ for any $x, y \in V$.*

Assumption 2.6. *For any $x \in V$, the number of incoming edges into x of the graph (V, E) and the number of outgoing edges from x are both finite.*

Assumptions 2.3 and 2.4 are standard in the literature [3]. Assumption 2.5 is also common in previous papers [35] which guarantees that the large deviation rate function of the empirical measure and empirical flow has the form of relative entropy. Note that if Assumption 2.5 does not hold for (X, τ), then the process $(Y, \tau) = \{(Y_k)_{k \geq 0}, (\tau_k)_{k \geq 1}\}$ with $Y_k := (X_k, X_{k+1})$ is also a Markov renewal process which satisfies Assumption 2.5 since

$$\mathbb{P}(\tau_{k+1} \in \cdot \mid (X_k)_{k \geq 0}) = \psi_{X_k, X_{k+1}}(\cdot) = \psi_{Y_k}(\cdot).$$

Moreover, we emphasize that Assumption 2.6 is not needed if ξ is a continuous-time Markov chain [6, 7]. Here we need this assumption in order to prove Lemma 4.10 below. In [6, 7], the counterpart of this lemma is proved by using the classical level 3 large deviation results of Donsker and Varadhan [19] and the contraction principle. However, since there are no level 3 large deviation results for Markov renewal processes, we need to impose the above assumption to overcome some technical difficulties.

Recall that the semi-Markov process ξ is called *non-explosive* if the explosion time $S_\infty := \lim_{n \to \infty} S_n$ satisfies

$$\mathbb{P}_x(S_\infty = \infty) = 1$$

for all $x \in V$, where $\mathbb{P}_x(\cdot) = \mathbb{P}(\cdot | X_0 = x)$. In fact, Assumptions 2.3 and 2.4 ensure that ξ is non-explosive. The proof is similar to the one given in [38] for continuous-time Markov chains and thus is omitted.

2.2 Empirical measure and empirical flow

Next we introduce the definitions of the empirical measure and empirical flow for Markov renewal processes [35]. For any $t > 0$, the empirical measure $\mu_t : \Omega \to \mathcal{P}(V \times (0, \infty])$ of (X, τ) is defined by

$$\mu_t = \frac{1}{t} \int_0^t \delta_{(X_{N_s}, \tau_{N_s+1})} \mathrm{d}s, \tag{2.1}$$

where δ. denotes the Dirac delta measure. In other words, for any $x \in V$ and $A \subset (0, \infty]$, we have

$$\mu_t(\{x\}, A) = \frac{1}{t} \int_0^t 1_{(X_{N_s} = x, \tau_{N_s+1} \in A)} \mathrm{d}s.$$

Then μ_t is a random probability measure such that for any Borel measurable function f on $V \times (0, \infty]$,

$$\langle \mu_t, f \rangle = \frac{1}{t} \int_0^t f(X_{N_s}, \tau_{N_s+1}) \mathrm{d}s = \sum_{k=1}^{N_t} \frac{\tau_k}{t} f(X_{k-1}, \tau_k) + \frac{t - S_{N_t}}{t} f(X_{N_t}, \tau_{N_t+1}).$$

Moreover, for any $t > 0$, the empirical flow $Q_t : \Omega \to [0, \infty]^E$ of (X, τ) is defined by

$$Q_t = \frac{1}{t} \sum_{k=1}^{N_t+1} \delta_{(X_{k-1}, X_k)}.$$

In other words, for any $x, y \in V$, we have

$$Q_t(x, y) = \frac{1}{t} \sum_{k=1}^{N_t+1} 1_{(X_{k-1} = x, X_k = y)}.$$

Intuitively, $Q_t(x, y)$ represents the number of times that ξ transitions from x to y per unit time. For any $n \geq 0$, let

$$\mathcal{F}_n = \sigma\left((X_k, \tau_k)_{0 \leq k \leq n}\right), \quad \mathcal{F}_\infty = \sigma\left(\bigcup_{n=1}^\infty \mathcal{F}_n\right).$$

be a filtration. Then $N_t + 1$ is an $\{\mathcal{F}_n\}$-stopping time. It is easy to see that $\langle \mu_t, f \rangle$ and $Q_t(x, y)$ are $\mathcal{F}_{N_t+1}$-measurable random variables.

Remark 2.7. For the semi-Markov process ξ, a more natural definition of the empirical measure $\pi_t : \Omega \to \mathcal{P}(V)$ is given by

$$\pi_t(x) = \frac{1}{t} \int_0^t 1_{(\xi_s = x)} \mathrm{d}s = \mu_t(x, (0, \infty]), \quad x \in V. \tag{2.2}$$

Comparing (2.1) and (2.2), we can see that π_t only focuses on the spatial variable and μ_t focus on both the spatial and temporal variables. The reason why we use μ_t rather than π_t in the study of the joint LDP is that only by using μ_t can we obtain a concise expression of the rate function. It is easy to verify that $(X_k, \tau_{k+1})_{k \geq 0}$ is a Markov process and hence $(X_{N_t}, \tau_{N_t+1})_{t \geq 0}$ is also a semi-Markov process. In fact, the empirical measure μ_t for the process $\xi_t = X_{N_t}$ is exactly the empirical measure π_t for the process $(X_{N_t}, \tau_{N_t+1})_{t \geq 0}$.

Let $L^1(E)$ denote the set of absolutely summable functions on E and let $\|\cdot\|$ denote the associated L^1-norm. The set of nonnegative elements of $L^1(E)$ is denoted by $L^1_+(E)$, which is called the *flow space*. An element in $L^1_+(E)$ is called a *flow*. Since (X, τ) is non-explosive, it is easy to see that $Q_t \in L^1_+(E)$ for any $t > 0$. For any flow $Q \in L^1_+(E)$, let the exit-current $Q^+ : V \to \mathbb{R}$ and entrance-current $Q^- : V \to \mathbb{R}$ be defined by

$$Q^+(x) = \sum_{y: (x,y) \in E} Q(x, y), \quad Q^-(x) = \sum_{y: (y,x) \in E} Q(y, x). \tag{2.3}$$

Intuitively, the exit-current and entrance-current at x are the flows exiting from x and entering into x, respectively. In particular, if $Q^+(x) = Q^-(x)$ for any $x \in V$, then Q is called a *divergence-free flow* and we define

$$Q_x = Q^+(x) = Q^-(x), \quad x \in V. \tag{2.4}$$

Note that both currents map $L^1_+(E)$ into $L^1_+(V)$. Here $L^1_+(V)$ is defined in the same way as $L^1_+(E)$ and let $\|\cdot\|$ denote the associated L^1-norm.

In this paper, we will consider two types of topology on $L^1(E)$: the strong topology generated by the L^1-norm and the bounded weak* topology, which is defined as follows [7]. Let $C_0(E)$ denote the collection of continuous functions $f : E \to \mathbb{R}$ vanishing at infinity, i.e. for any $\epsilon > 0$, there exists a finite set $K \subset E$ such that

$$\sup_{(x,y) \in K^c} |f(x,y)| \leq \epsilon,$$

and it is endowed with the L^∞-norm. It is well-known that the dual space of $C_0(E)$ is $L^1(E)$ endowed with the strong topology. For any $\ell > 0$, let $B_\ell := \left\{ Q \in L^1(E) : \|Q\| \leq \ell \right\}$ denote the closed ball of radius ℓ in $L^1(E)$. In view of the separability of $C_0(E)$ and the Banach-Alaoglu theorem, the closed ball B_ℓ endowed with the weak* topology is a compact Polish space. The bounded weak* topology on $L^1(E)$ is then defined by declaring the set $A \subset L^1(E)$ to be open if and only if $A \cap B_\ell$ is open in the weak* topology of B_ℓ for any $\ell > 0$. In fact, the bounded weak* topology is stronger than the weak* topology and is weaker than the strong topology. Moreover, for each $\ell > 0$, the closed ball B_ℓ is compact with respect to the bounded weak* topology. In particular, the three types of topology on $L^1(E)$ coincide only when E is finite. The proof of the above statements can be found in [36, Section 2.7].

For both the strong topology and the bounded weak* topology, we regard $L^1_+(E)$ as a closed subset of $L^1(E)$ and endow it with the relative topology and the associated Borel σ-algebra. The product space $V \times (0, \infty]$ is equipped with the product topology so that $V \times (0, \infty]$ is a Polish space. The set $\mathcal{P}(V \times (0, \infty])$ of Borel probability measures on $V \times (0, \infty]$ is endowed with the topology of weak convergence. Moreover, the product space $\Lambda = \mathcal{P}(V \times (0, \infty]) \times L^1_+(E)$ is endowed with the product topology. Then for any $t > 0$, the pair (μ_t, Q_t), where μ_t is the empirical measure and Q_t is the empirical flow, can be viewed as a measurable map from Ω to Λ.

3 Main results

3.1 Compactness conditions

The aim of the present paper is to establish the LDP for the empirical measure and empirical flow of semi-Markov processes. Similarly to the LDP of Markov processes [19], we need some compactness conditions to control the convergence rate at infinity when the state space V has an infinite number of states. For Markov processes, the infinitesimal generator is often used to establish the compactness conditions [19]. In the semi-Markov case, the *transition kernel* (P, Ψ) plays the role of the generator.

Before stating the compactness conditions, we need the following notation. Let $\zeta : V \to [0, \infty]$ be the function defined by

$$\zeta(x) = \sup \left\{ \lambda \in \mathbb{R} : \psi_x \left(e^{\lambda \tau} \right) < \infty \right\}, \tag{3.1}$$

where $\psi_x(e^{\lambda \tau}) = \int_{(0,\infty)} e^{\lambda s} \psi_x(\mathrm{d}s)$. For any $x \in V$, let $\theta_x : (0, \infty) \to (-\infty, \infty]$ be the function defined by

$$\theta_x(t) = \sup \left\{ \lambda \in \mathbb{R} : \psi_x \left(e^{\lambda \tau} \right) \leq t \right\}. \tag{3.2}$$

For any $f : V \to (0, \infty)$, let $Pf : V \to (0, \infty]$ be the function defined by

$$Pf(x) = \sum_{y : (x,y) \in E} p_{xy} f(y),$$

and let $Lf : V \to (-\infty, \infty]$ be the function defined by $Lf(x) = \theta_x(f(x))$.

Lemma 3.1. *Let θ_x be the function defined in (3.2). Then*

(a) *θ_x is an increasing continuous function. Moreover, θ_x is strictly increasing on $(0, \psi_x(e^{\zeta(x)\tau}))$.*

(b) *$\theta_x(1) = 0$, $\lim_{t\downarrow 0} \theta_x(t) = -\infty$, and $\lim_{t\uparrow\infty} \theta_x(t) = \zeta(x)$.*

(c) *For any $t > 0$ and $x \in V$, we have*

$$\psi_x\left(e^{\theta_x(t)\tau}\right) = \begin{cases} \psi_x\left(e^{\zeta(x)\tau}\right), & \text{if } t > \psi_x\left(e^{\zeta(x)\tau}\right), \\ t, & \text{otherwise.} \end{cases}$$

Proof. For any $x \in V$, let $\Theta_x(\lambda) = \psi_x(e^{\lambda\tau})$ be a function on $\mathbb{R}$. By the dominated convergence theorem, it is easy to see that $\Theta_x \in C((-\infty, \zeta(x)))$ is a strict increasing function. Moreover, it is clear that $\Theta_x(0) = 1$ and

$$\lim_{\lambda \to -\infty} \Theta_x(\lambda) = 0, \qquad \Theta_x(\lambda) = \infty, \quad \lambda > \zeta(x).$$

Since θ_x is the inverse function of Θ_x for $0 < t < \Theta_x(\zeta(x))$, we immediately obtain (a)-(c). $\qquad\square$

In [18, 19], Donsker and Varadhan proposed a compactness condition for the generator which was then used to prove the LDP for the empirical measure of Markov processes. In [6, 7], Bertini et al. proposed a Donsker-Varadhan-type condition which was then used to prove the LDP for the empirical flow of continuous-time Markov chains. In what follows, we will provide two Donsker-Varadhan-type conditions which are needed for the joint LDP for the empirical measure and empirical flow of semi-Markov processes when the flow space $L^1_+(E)$ is endowed with the bounded weak* topology and strong topology. The following compactness condition is needed for the joint LDP when $L^1_+(E)$ is endowed with the bounded weak* topology.

Condition 3.2. *There exists a sequence of functions $u_n \colon V \to (0, \infty)$ such that*

(a) *For any $x \in V$ and $n \geq 0$, we have $Pu_n(x) < \infty$;*

(b) *There exists a constant $c > 0$ such that $u_n(x) \geq c$ for any $x \in V$ and $n \geq 0$;*

(c) *For any $x \in V$, there exists a constant C_x such that $u_n(x) \leq C_x$ for any $n \geq 0$;*

(d) *The functions u_n/Pu_n converge pointwise to some $\hat{u} : V \to (0, \infty)$;*

(e) *For each $\ell \in \mathbb{R}$, the level set $\{x \in V : L\hat{u}(x) \leq \ell\}$ is finite;*

(f) *There exist $\sigma, C > 0$, $\eta \in (0, 1)$, and a finite set $K \subset V$ such that $L\hat{u}(x) \geq -\sigma\theta_x(\eta) - C1_K(x)$ for any $x \in V$ and $\hat{u}(x) < \psi_x(e^{\zeta(x)\tau})$ for any $x \in K^c$.*

Remark 3.3. For continuous-time Markov chains, the term $C1_K$ in item (f) can be replaced by a constant C and the condition $\hat{u}(x) < \psi_x(e^{\zeta(x)\tau})$ can be removed since $\psi_x(e^{\zeta(x)\tau}) = \infty$ for any $x \in V$. In this case, Condition 3.2 reduces to the compactness condition proposed in [7]. Moreover, it follows from Lemma 3.1 that $L\hat{u}(x) = \theta_x(\hat{u}(x)) \leq \zeta(x)$ for any $x \in V$. Hence item (e) implies that the level sets of ζ are also finite.

Since the strong topology is stronger than the bounded weak* topology, we need to impose a stronger compactness condition, which is essentially the Donsker-Varadhan-type condition for discrete-time Markov chains [18]. Both the following condition and Condition 3.2 are needed for the joint LDP when $L^1_+(E)$ is endowed with the strong topology.

Condition 3.4. *There exists a sequence of functions $u_n \colon V \to (0, \infty)$ such that*

(a) *For any $x \in V$ and $n \geq 0$, we have $Pu_n(x) < \infty$;*

(b) There exists a constant $c > 0$ such that $u_n(x) \geq c$ for any $x \in V$ and $n \geq 0$;

(c) For any $x \in V$, there exists a constant C_x such that $u_n(x) \leq C_x$ for any $n \geq 0$;

(d) The functions u_n/Pu_n converge pointwise to some $\hat{u} : V \to (0, \infty)$;

(e) For each $\ell \in \mathbb{R}$, the level set $\{x \in V : \log \hat{u}(x) \leq \ell\}$ is finite.

It is clear that Condition 3.4 only depends on the embedded chain X of the semi-Markov process and is independent of the waiting time distributions. When $L_+^1(E)$ is endowed with the strong topology, Bertini et al. [6] have proposed another compactness condition for continuous-time Markov chains (see Condition 7.1 below, which is rewritten for semi-Markov process). However, that condition is more complicated than Condition 3.4 and more difficult to verify. In fact, when X is irreducible, Condition 3.4 is not only easier to verify, but also even weaker than Condition 7.1. The proof of this fact can be found in Section 7. This explains why we impose Condition 3.4 rather than Condition 7.1 here.

Remark 3.5. For discrete-time Markov chains, we have $Lf(x) = \theta_x(f(x)) = \log f(x)$ for any $f : V \to (0, \infty)$. Note that item (e) in Condition 3.4 implies that the set $K = \{x \in V : \log \hat{u}(x) \leq -\log \eta\}$ is finite for any $\eta \in (0, 1)$. Hence if we take $\sigma = 1$, then we can always find $C > 0$ such that $\log \hat{u} \geq -\sigma \log \eta - C 1_K$. Furthermore, it is easy to check that $\psi_x(e^{\zeta(x)\tau}) = e^{\zeta(x)} = \infty$ for any $x \in V$. This shows if Condition 3.4 holds, then items (e) and (f) in Condition 3.2 are automatically satisfied. Hence for discrete-time Markov chains, Condition 3.2 is equivalent to Condition 3.4.

3.2 Joint LDP for the empirical measure and empirical flow

Let μ and ν be two probability measures on a measurable space $(\mathcal{X}, \mathcal{F})$. Recall that the relative entropy of μ with respect to ν has the following variational expression [19]:

$$H(\mu \,|\, \nu) = \sup_{\varphi \in \mathcal{B}_b(\mathcal{X})} \{\langle \mu, \varphi \rangle - \log \langle \nu, e^\varphi \rangle\} = \begin{cases} \int_{\mathcal{X}} \left(\log \dfrac{\mathrm{d}\mu}{\mathrm{d}\nu} \right) \mathrm{d}\mu, & \text{if } \mu \ll \nu, \\ \infty, & \text{otherwise,} \end{cases} \tag{3.3}$$

where $\mathcal{B}_b(\mathcal{X})$ denotes the space of bounded measurable functions on $\mathcal{X}$. Moreover, if $\mathcal{X}$ is a Polish space and $\mathcal{F}$ is the associated Borel σ-field, then (3.3) still holds when $\mathcal{B}_b(\mathcal{X})$ is replaced by $C_b(\mathcal{X})$, the space of bounded continuous functions on $\mathcal{X}$ [19]. If we set $\varphi' = \varphi - \log \langle \nu, e^\varphi \rangle$, then $\langle \mu, \varphi' \rangle = \langle \mu, \varphi \rangle - \log \langle \nu, e^\varphi \rangle$ and it is easy to see that the relative entropy $H(\mu \,|\, \nu)$ can be represented as

$$H(\mu \,|\, \nu) = \sup_{\{\varphi \in C_b(\mathcal{X}) : \langle \nu, e^\varphi \rangle = 1\}} \langle \mu, \varphi \rangle. \tag{3.4}$$

Let $\mathcal{D}$ be the subset of $\Lambda = \mathcal{P}(V \times (0, \infty]) \times L_+^1(E)$ defined by

$$\mathcal{D} = \left\{ (\mu, Q) \in \Lambda : \int_{(0, \infty]} \frac{1}{t} \mu(x, \mathrm{d}t) = Q^+(x) = Q^-(x), \forall x \in V \right\}, \tag{3.5}$$

where Q^+ and Q^- are the exit-current and entrance-current of the flow Q, respectively. For each $(\mu, Q) \in \mathcal{D}$, we introduce the transition probabilities $(\tilde{Q}_{xy})_{x,y \in V}$ and the waiting time distributions $(\tilde{\mu}_x)_{x \in V}$ as

$$\tilde{Q}_{xy} = \frac{Q(x,y)}{Q_x}, \quad \tilde{\mu}_x(\mathrm{d}t) = \frac{1}{Q_x t} \mu(x, \mathrm{d}t), \tag{3.6}$$

where $Q_x = Q^+(x) = Q^-(x)$ is defined in (2.4) and we set $\tilde{Q}_{xy} = p_{xy}$ and $\tilde{\mu}_x = \psi_x$ if $Q_x = 0$. Let $I : \Lambda \to [0, \infty]$ be the function defined by

$$I(\mu, Q) = \begin{cases} \sum_{x \in V} \left[Q_x H(\tilde{Q}_{x,\cdot} \,|\, p_{x,\cdot}) + Q_x H(\tilde{\mu}_x \,|\, \psi_x) + \zeta(x)\mu(x, \{\infty\}) \right], & \text{if } (\mu, Q) \in \mathcal{D}, \\ \infty, & \text{otherwise.} \end{cases} \tag{3.7}$$

We are now in a position to state the main results of the present paper. The following theorem, whose proof can be found in Section 4, gives the joint LDP for the empirical measure and empirical flow when $L^1_+(E)$ is endowed with the bounded weak* topology.

Theorem 3.6. *Suppose that Assumptions 2.3-2.6 and Condition 3.2 hold. Let $L^1_+(E)$ be endowed with the bounded weak* topology. Then under $\mathbb{P}_x$, the law of (μ_t, Q_t) satisfies an LDP with good and convex rate function $I : \Lambda \to [0, \infty]$. In particular, for any closed set $\mathcal{C} \subset \Lambda$ and open set $\mathcal{A} \subset \Lambda$, we have*

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_x\left((\mu_t, Q_t) \in \mathcal{C}\right) \leq - \inf_{(\mu, Q) \in \mathcal{C}} I(\mu, Q),$$
$$\varliminf_{t \to \infty} \frac{1}{t} \log \mathbb{P}_x\left((\mu_t, Q_t) \in \mathcal{A}\right) \geq - \inf_{(\mu, Q) \in \mathcal{A}} I(\mu, Q). \tag{3.8}$$

The following theorem, whose proof can be found in Section 5, gives the joint LDP when $L^1_+(E)$ is endowed with the strong topology.

Theorem 3.7. *Suppose that Assumptions 2.3-2.6 and Conditions 3.2 and 3.4 hold. Let $L^1_+(E)$ be endowed with the strong topology. Then under $\mathbb{P}_x$, the law of (μ_t, Q_t) satisfies an LDP with good and convex rate function $I : \Lambda \to [0, \infty]$.*

The joint LDP with the strong topology in the flow space allows us to obtain the LDP of some important observables by using the contraction principle. For example, in [6, Theorem 8.1], the authors derived the LDP for the empirical entropy production (Gallavotti-Cohen functional) under some mild conditions.

Remark 3.8. In the above two theorems, we have established the joint LDP when the semi-Markov process starts from a fixed initial state $x \in V$. For any initial distribution $\gamma \in \mathcal{P}(V)$, the joint LDP still holds when making some slight changes to the compactness conditions. In fact, if items (c) in Conditions 3.2 and 3.4 are both replaced by

(c*) There exists a constant C_γ such that $\sum_{x \in V} \gamma(x)u_n(x) \leq C_\gamma$ for any $n \geq 0$,

then the conclusions of Theorems 3.6 and 3.7 remain valid under $\mathbb{P}_\gamma$, where $\mathbb{P}_\gamma(\cdot) = \sum_{x \in V} \gamma(x)\mathbb{P}_x(\cdot)$ is the probability measure under initial distribution γ.

The above two theorems can be applied to obtain the joint LDP for the empirical measure and empirical flow of discrete-time and continuous-time Markov chains with countable state space. For discrete-time Markov chains, we have the following results.

Corollary 3.9. *Let ξ be a discrete-time Markov chain satisfying Assumptions 2.3-2.6 and Condition 3.4. Let $L^1_+(E)$ be endowed with the bounded weak* topology or strong topology. Then under $\mathbb{P}_x$, the law of (μ_t, Q_t) satisfies an LDP with good and convex rate function $I : \Lambda \to [0, \infty]$.*

Proof. In Remark 3.5, we have shown that Conditions 3.2 and 3.4 are equivalent for discrete-time Markov chains. Then the results follow directly from Theorems 3.6 and 3.7. $\square$

We have seen that Condition 3.2 is crucial for the joint LDP of the empirical measure and empirical flow, no matter whether the flow space is endowed with the bounded

weak* topology or the strong topology. In general, Condition 3.2 is difficult to verify because we need to find a sequence of functions u_n satisfying both items (a)-(d), which are related to the embedded chain, and items (e)-(f), we are related to the waiting time distributions. In other words, the conditions imposed on the embedded chain and the conditions imposed on waiting time distributions are intertwined with each other. Next, we provide some novel compactness conditions which can be verified much more easily. In the novel compactness conditions, the ones imposed on the embedded chain and the ones imposed on waiting time distributions can be disassembled and are not intertwined with each other.

The conditions imposed on the embedded chain are as follows.

Condition 3.10. *There exists a sequence of functions* $u_n : V \to (0, \infty)$ *satisfying items (a)-(d) in Condition 3.4 and*

(e) *There exists a constant* $\ell > 0$ *such that the level set* $\{x \in V : \log \hat{u}(x) \leq \ell\}$ *is finite.*

Note that item (e) in Condition 3.10 is weaker than item (e) in Condition 3.4. The conditions imposed on the waiting time distributions are as follows.

Condition 3.11. *There exist a probability measure* $\psi \in \mathcal{P}(0, \infty)$ *and a function* $q : V \to (0, \infty)$ *such that*

(a) *The waiting time distributions* $(\psi_x)_{x \in V}$ *satisfy* $\psi_x(A/q_x) = \psi(A)$ *for any* $x \in V$ *and any Borel set* $A \subset (0, \infty)$*, where* $A/q_x = \{t/q_x : t \in A\}$*;*

(b) *The probability measure* ψ *satisfies* $\psi\left(e^{\zeta \tau}\right) = \infty$*, where* $\zeta = \sup\left\{\lambda \geq 0 : \psi(e^{\lambda \tau}) < \infty\right\}$*;*

(c) *For each* $\ell \in \mathbb{R}$*, the level set* $\{x \in V : q_x \leq \ell\}$ *is finite.*

Remark 3.12. It is easy to check that the following distribution families satisfy items (a) and (b) in Condition 3.11:

(a) Exponential distribution: $\psi_x(\mathrm{d}t) = q_x e^{-q_x t} \mathrm{d}t$;

(b) Gamma distribution: $\psi_x(\mathrm{d}t) = (q_x^\alpha t^{\alpha-1} e^{-q_x t}/\Gamma(\alpha))\mathrm{d}t$ with any $\alpha > 0$;

(c) Dirac distribution: $\psi_x(\mathrm{d}t) = \delta_{1/q_x}(\mathrm{d}t)$;

(d) Rayleigh distribution: $\psi_x(\mathrm{d}t) = (t/q_x^2)e^{-t^2/(2q_x^2)}\mathrm{d}t$.

In particular, items (a) and (b) in Condition 3.11 automatically hold for continuous-time Markov chains.

The following theorem shows that the joint LDP also holds under the new compactness conditions given above.

Theorem 3.13. *Suppose that Assumptions 2.3-2.6 are satisfied.*

(a) *Suppose that Conditions 3.10 and 3.11 both hold. Let* $L_+^1(E)$ *be equipped with the bounded weak* topology. Then under* $\mathbb{P}_x$*, the law of* (μ_t, Q_t) *satisfies an LDP with good and convex rate function* $I : \Lambda \to [0, \infty]$*.*

(b) *Suppose that Conditions 3.4 and 3.11 both hold. Let* $L_+^1(E)$ *be endowed with the strong topology. Then under* $\mathbb{P}_x$*, the law of* (μ_t, Q_t) *satisfies an LDP with good and convex rate function* $I : \Lambda \to [0, \infty]$*.*

Proof. (a) By Theorem 3.6, we only need to check Condition 3.2 for the sequence of functions u_n in Condition 3.10. In fact, items (a)-(d) in Condition 3.2 are trivial. We next prove items (e) and (f) in Condition 3.2.

Let $\theta : (0, \infty) \to (-\infty, \infty]$ be the function defined by

$$\theta(s) = \sup\left\{\lambda \in \mathbb{R} : \psi(e^{\lambda \tau}) \leq s\right\}. \tag{3.9}$$

By item (b) in Condition 3.11, it is easy to see that the function θ_x defined in (3.2) and the function θ defined above are related by $\theta_x(s) = q_x\theta(s)$. Lemma 3.1 implies that θ is an increasing function. It follows from item (e) in Condition 3.10 and item (c) in Condition 3.11 that the set $K := \{x \in V : \log \hat{u}(x) \le \ell\}$ is finite and the set $\{x \in V : q_x \le s\}$ is also finite for any $s \in \mathbb{R}$. For any $\ell' > 0$, we have

$$\left\{x \in V : \hat{u}(x) > e^\ell\right\} \cap \left\{x \in V : q_x > \frac{\ell'}{\theta(e^\ell)}\right\} \subseteq \left\{x \in V : q_x\theta(\hat{u}(x)) > \ell'\right\}.$$

Then we have

$$\{x \in V : L\hat{u}(x) \le \ell'\} \subseteq K \cup \left\{x \in V : q_x \le \frac{\ell'}{\theta(e^\ell)}\right\}.$$

This implies item (e) in Condition 3.2.

Take $\eta = 1/2$ and $\sigma = -\theta(e^\ell)/\theta(1/2)$. By Lemma 3.1, it is easy to check that $\sigma > 0$. Then we have

$$q_x\theta\left(\hat{u}(x)\right) \ge -\sigma q_x\theta(1/2) - C1_K(x),$$

where $C = 1 \vee \max_{x \in K}\{-q_x[\theta(\hat{u}(x)) + \sigma\theta(1/2)]\}$. By item (a) in Condition 3.11, it is easy to see that $\psi_x(e^{\zeta(x)\tau}) = \infty$ for any $x \in V$. This implies item (f) in Condition 3.2.

(b) Note that Condition 3.4 implies Condition 3.10. Then the proof of (b) follows directly from Theorem 3.7. $\qquad\square$

Note that Condition 3.11 is easy to verify and we have given several examples for it to hold in Remark 3.12. Next we will give some criterions for Condition 3.10 to hold. Before doing this, we recall the following geometric ergodic theorem for discrete-time Markov chains, whose proof can be found in [37, Chapter 15].

Lemma 3.14 (geometric ergodic theorem). *Suppose that the embedded chain X is irreducible and aperiodic. Then the following three conditions are equivalent:*

(a) *There exist a finite set $K \subset V$ and constants $\nu_K > 0$, $\rho_K < 1$, and $M_K < \infty$ such that*

$$\sup_{x \in K}\left|\left(\sum_{y \in K}p_{xy}^n\right) - \nu_K\right| \le M_K\rho_K^n.$$

(b) *There exist a finite set $K \subset V$ and a constant $\kappa > 1$ such that*

$$\sup_{x \in K}\mathbb{E}_x\left[\kappa^{T_K}\right] < \infty,$$

where T_K is the first hitting time of X on K (see (4.10)).

(c) *There exist a finite set $K \subset V$, constants $b < \infty$, $\lambda < 1$, and $c > 0$, and a function $u : V \to [c, \infty)$ satisfying the drift condition*

$$Pu(x) \le \lambda u(x) + b1_K(x), \quad x \in V.$$

The following corollary gives a simple criterion for Condition 3.10 to hold.

Corollary 3.15. *If item (c) in Lemma 3.14 holds, then Condition 3.10 also holds. If the embedded chain X is irreducible and aperiodic, then any one of the three conditions in Lemma 3.14 implies Condition 3.10.*

Proof. Let the set K, the constants λ, c, and the function u be as in (c) in Lemma 3.14. We next check Condition 3.10 for the sequence of functions $u_n \equiv u$. Obviously, items (a)-(d) are trivial. Let $\hat{u} = u/Pu$ and $\ell = -(\log \lambda)/2 > 0$. It is easy to see that

$$\{x \in V : \log \hat{u}(x) \le \ell\} \subseteq \{x \in V : Pu(x) > \lambda u(x)\} \subseteq K.$$

This completes the proof of this corollary. $\qquad\square$

The above corollary shows that geometric ergodicity of the embedded chain X implies Condition 3.10. Hence this condition can also be easily verified by using the classical ergodic theory of Markov chains [10, 37, 4, 43].

3.3 Marginal LDP for the empirical measure and empirical flow

Thus far, we have established the joint LDP for the empirical measure μ_t and empirical flow Q_t. However, as discussed in Remark 2.7, a more natural definition of the empirical measure $\pi_t : \Omega \to \mathcal{P}(V)$ is given by

$$\pi_t(x) = \frac{1}{t} \int_0^t 1_{(\xi_s = x)} \mathrm{d}s = \mu_t(x, (0, \infty]), \quad x \in V.$$

The reason why we use μ_t rather than π_t in the study of the joint LDP is that only by using μ_t can we obtain a concise expression of the rate function.

Next we focus on the marginal LDP for the empirical measure π_t and for the empirical flow Q_t. By the contraction principle, the rate function of the marginal LDP can be obtained from the rate function $I : \Lambda \to [0, \infty]$ of the joint LDP as defined in (3.8). In [35], the authors gave a variational expression of the rate function $I_1 : \mathcal{P}(V) \to [0, \infty]$ for the empirical measure π_t. Here we will give the variational expression of the rate function $I_2 : L_+^1(E) \to [0, \infty]$ for the empirical flow Q_t. More importantly, we will also give the explicit expression of I_2 when the waiting time distributions satisfy some additional constraints.

Before stating our results, we introduce some notation. Recall that the rate function $I_{DV} : \mathcal{P}(V) \to [0, \infty]$ for the empirical measure of the embedded chain X is the Donsker-Varadhan functional [16]

$$I_{DV}(\nu) = \sup_{u \in (0, \infty)^V} \sum_{x \in V} \nu_x \log \left(\frac{u}{Pu}(x) \right).$$

Let $G_x(\lambda) = \log(\psi_x(e^{\lambda \tau}))$ be a function on $\mathbb{R}$ and let

$$G_x^*(a) = \sup_{\lambda \in \mathbb{R}} (a\lambda - G_x(\lambda)), \quad a \in \mathbb{R}$$

be the Fenchel-Legendre transform of G_x. Moreover, let $I_1 : \mathcal{P}(V) \to [0, \infty]$ be the functional defined by

$$I_1(\pi) = \inf_{r > 0} \inf_{\nu \in \mathcal{P}(V)} \left(r I_{DV}(\nu) + \sum_{x \in V} r \nu_x G_x^* \left(\frac{\pi_x}{r \nu_x} \right) \right).$$

The following proposition, whose proof can be found in Section 6, gives the marginal LDP for the empirical measure and for the empirical flow.

Proposition 3.16. *Suppose that Assumptions 2.3-2.6 and Condition 3.2 hold.*

(a) *Under $\mathbb{P}_x$, the law of π_t satisfies an LDP with good and convex rate function $I_1 : \mathcal{P}(V) \to [0, \infty]$;*

(b) *Let $L_+^1(E)$ be endowed with the bounded weak* topology. Then under $\mathbb{P}_x$, the law of Q_t satisfies an LDP with good and convex rate function*

$$I_2(Q) = \inf_{(\mu, Q) \in \mathcal{D}} I(\mu, Q), \quad Q \in L_+^1(E);$$

(c) *Let $L_+^1(E)$ be endowed with the strong topology. If Condition 3.4 is also satisfied, then under $\mathbb{P}_x$, the law of Q_t satisfies an LDP with good and convex rate function $I_2 : L_+^1(E) \to [0, \infty]$.*

Moreover, if Condition 3.11 is also satisfied, then the rate function I_2 has the following explicit expression:

$$I_2(Q) = \begin{cases} \displaystyle\sum_{x \in V} Q_x H\left(\tilde{Q}_{x,\cdot} \big| p_{x,\cdot}\right) + \sup_{\lambda < \inf_{x \in V} \zeta(x)} \left\{\lambda - \sum_{x \in V} Q_x \log \psi_x\left(e^{\lambda \tau}\right)\right\}, & \text{if } Q^+ = Q^-, \\ \infty, & \text{otherwise.} \end{cases} \tag{3.10}$$

Note that in [5, Corollary 2.5], the authors gave a variational expression of the rate function I_2 for degenerate jump Markov processes. Our expression is very similar to theirs.

Remark 3.17. This proposition shows that the rate function for the empirical flow has the explicit expression (3.10) when Condition 3.11 is satisfied. In fact, (3.10) may be still valid when Condition 3.11 is broken. For example, if V is finite and $\psi_x(e^{\zeta(x)\tau}) = \infty$ for any $x \in V$, then similarly to the proof of Proposition 3.16, it can be shown that the rate function I_2 is also given by (3.10).

3.4 Examples

Our abstract theorems can be applied to many specific Markov renewal processes. We next focus on two specific examples: birth and death processes and random walks with confining potential and external force. These two examples can be viewed as direct generalizations of the ones studied in [7, 6]. Here we will apply Theorems 3.6 and 3.7 to birth and death processes and will apply Theorem 3.13 to random walks with confining potential and external force.

3.4.1 Birth and death processes

Consider a birth and death Markov renewal process on the set of nonnegative integers $\mathbb{N} = \{0, 1, 2 \cdots\}$ with transition kernel $(P = (p_{xy})_{x,y \in \mathbb{N}}, \Psi = (\psi_x)_{x \in \mathbb{N}})$, where

$$p_{01} = 1, \qquad p_{x,x+1} = p_x > 0, \qquad p_{x,x-1} = q_x := 1 - p_x > 0, \quad x \geq 1.$$

In fact, Assumptions 2.3, 2.5, and 2.6 are trivial. In addition, it is well-known that Assumption 2.4 holds if and only if [41]

$$\sum_{k=1}^{\infty} \frac{q_1 q_2 \cdots q_k}{p_1 p_2 \cdots p_k} = \infty. \tag{3.11}$$

Applying Theorems 3.6 and 3.7 to the above model, we obtain the following proposition.

Proposition 3.18. *Let $p = \overline{\lim}_{x \to \infty} p_x$. Suppose that $\psi_x(e^{\zeta(x)\tau}) = \infty$ for any $x \in V$.*

(a) *Let $L_+^1(E)$ be endowed with the bounded weak* topology. Suppose that $p < 1/2$ and suppose that there exist constants $\kappa < (4p(1-p))^{-1/2}$, $\sigma > 0$, and $\eta \in (0,1)$ such that*

$$\lim_{x \to \infty} \theta_x(\kappa) = \infty, \qquad \theta_x(\kappa) \geq -\sigma \theta_x(\eta), \quad x \in \mathbb{N}. \tag{3.12}$$

Then under $\mathbb{P}_x$, the law of (μ_t, Q_t) satisfies an LDP with good and convex rate function $I : \Lambda \to [0, \infty]$.

(b) *Let $L_+^1(E)$ be endowed with the strong topology. Let $p'_x = \sup_{k \geq x} p_k$ for each x. Suppose that $p = 0$ and suppose that there exist constants $\sigma > 0$ and $\eta \in (0,1)$ such that*

$$\lim_{x \to \infty} \theta_x\left((9p'_{x-1})^{-1/2}\right) = \infty, \qquad \theta_x\left((9p'_{x-1})^{-1/2}\right) \geq -\sigma \theta_x(\eta), \quad x \in \mathbb{N}. \tag{3.13}$$

Then under $\mathbb{P}_x$, the law of (μ_t, Q_t) satisfies an LDP with good and convex rate function $I : \Lambda \to [0, \infty]$.

Proof. (a) Since $\overline{\lim}_{x \to \infty} p_x < 1/2$, it is easy to see that (3.11) holds. For any $x \in \mathbb{N}$ and $n \geq 0$, let $u_n(x) = a^x$, where $a > 1$ is a constant to be chosen later. By Theorem 3.6, we only need to check Condition 3.2 for the sequence of functions u_n. Since u_n do not depend on n, items (a)-(d) in Condition 3.2 are automatically satisfied. Moreover, it is clear that the set $K = \{x \in \mathbb{N} : p_x > p'\}$ is finite. Note that for any $\ell \in \mathbb{R}$,

$$\{x \in \mathbb{N} : L\hat{u}(x) \leq \ell\} \subseteq \{x \in \mathbb{N} : \theta_x(\kappa) \leq \ell\} \cup K.$$

This implies item (e) in Condition 3.2. Letting $C = \max_{x \in K}(L\hat{u}(x) - \sigma\theta_x(\eta))$, it is easy to see that

$$L\hat{u}(x) \geq -\sigma\theta_x(\eta) - C1_K(x), \quad x \in \mathbb{N}.$$

On the other hand, $\hat{u}(x) < \infty = \psi_x(e^{\zeta(x)\tau})$ for any $x \in K^c$. These imply item (g) in Condition 3.2. Thus far, we have check all items in Condition 3.2 and thus the desired result follows from Theorem 3.6.

(b) Let $q'_x = 1 - p'_x$ for each x. For any $n \geq 0$, let $u_n(0) = u_n(1) = 1$ and $u_n(x) = (\prod_{k=1}^{x-1}(q'_k/p'_k))^{1/2}$ for any $x \geq 2$. Similarly, we only need to check item (e) in Condition 3.4 and items (e) and (f) in Condition 3.2 for such u_n. Note that $q'_{x-1} \geq 1/2$ for sufficiently large x. Then we have

$$\hat{u}(x) = \frac{1}{p_x(q'_x/p'_x)^{1/2} + q_x(p'_{x-1}/q'_{x-1})^{1/2}} \geq \frac{1}{(q'_x p'_x)^{1/2} + (2p'_{x-1})^{1/2}} \geq \frac{1}{3(p'_{x-1})^{1/2}}.$$

Since $\lim_{x \to \infty} p'_{x-1} = 0$, item (e) in Condition 3.4 holds. Let $K = \{x \in \mathbb{N} : q'_{x-1} < 1/2\} \cup \{0, 1\}$ be a finite set. The rest of the proof is similar to the proof of (a). $\square$

We emphasize that if ψ_x is chosen to be an exponential distribution for each x, then Proposition 3.18 reduces to the results obtained in [6, 7]. Moreover, if ψ_x is chosen to be the gamma distribution

$$\psi_x(\mathrm{d}t) = \left(q_x^{\alpha_x} t^{\alpha_x - 1} e^{-q_x t}/\Gamma(\alpha_x)\right) \mathrm{d}t$$

for each x, where q_x, α_x are parameters depending on x, then we can give the following more specific characterizations of conditions (3.12) and (3.13).

Lemma 3.19. *(a) Let $\kappa_0 = [1 + (4p(1-p))^{-1/2}]/2$. Suppose that*

$$\varliminf_{x \to \infty} q_x \left(1 - \kappa_0^{-1/\alpha_x}\right) = \infty$$

and suppose that the parameters α_x have a uniform positive lower bound, i.e. $\alpha_x \geq c$ for some constant $c > 0$ and any $x \in \mathbb{N}$. Then condition (3.12) holds.

(b) Let $p'_x = \sup_{k \geq x} p_k$ for each x. Suppose that

$$\varliminf_{x \to \infty} q_x \left(1 - (9p'_{x-1})^{1/(2\alpha_x)}\right) = \infty$$

and suppose that the parameters α_x have a uniform positive lower bound. Then condition (3.13) holds.

Proof. Here we only give the proof of (a); the proof of (b) is similar. Taking $\kappa = \kappa_0$ in Proposition 3.18, straightforward computations show that

$$\theta_x(t) = q_x \left(1 - t^{-1/\alpha_x}\right), \quad t > 0.$$

This shows that

$$\varliminf_{x \to \infty} \theta_x(\kappa) = \varliminf_{x \to \infty} q_x(1 - \kappa_0^{-1/\alpha_x}) = \infty.$$

Moreover, letting $\sigma = \kappa_0^{-1/c} > 0$ and $\eta = 1/\kappa_0 \in (0,1)$, we have

$$\theta_x(\kappa_0) = q_x \left(1 - \kappa_0^{-1/\alpha_x}\right) = -\kappa_0^{-1/\alpha_x} q_x \left(1 - \eta^{-1/\alpha_x}\right) \geq -\sigma \theta_x(\eta).$$

This completes the proof of this lemma. □

3.4.2   Random walks with confining potential and external force

We now apply our main theorems to a nearest neighbor random walk on $\mathbb{Z}^d$ with confining potential and external force, whose transition kernel $(P = (p_{xy})_{x,y \in \mathbb{Z}^d}, \psi = (\psi_x)_{x \in \mathbb{Z}^d})$ has the form of

$$p_{xy} = \frac{1}{C_x} \exp \left\{ -\frac{1}{2}(U(y) - U(x)) + \frac{1}{2}F(x,y) \right\}, \quad (x,y) \in E,$$

where $E = \{(x,y) \in \mathbb{Z}^d \times \mathbb{Z}^d : |x - y| = 1\}$ is the collection of nearest neighbours in $\mathbb{Z}^d$, $U : \mathbb{Z}^d \to \mathbb{R}$ is a potential function satisfying $\sum_{y \in \mathbb{Z}^d} e^{-U(y)} < \infty$, $F \in L^\infty(E)$ represents the external force, and

$$C_x = \sum_{y:|x-y|=1} e^{-(U(y)-U(x))/2 + F(x,y)/2}, \quad x \in \mathbb{Z}^d$$

are normalization constants. It is clear that U has compact level sets, i.e. for any $\ell \in \mathbb{R}$, the set $\{x \in \mathbb{Z}^d : U(x) \leq \ell\}$ is finite. We emphasize that if $\psi_x(dt) = C_x e^{-C_x t} dt$ is chosen to be an exponential distribution for each x, then the above model reduces to the Markov chain model studied in [6, Section 10.2]. We next focus on general waiting time distributions. Applying Theorem 3.13 to the above model, we obtain the following proposition.

Proposition 3.20. Let $r(x) = \sum_{y:|x-y|=1} e^{-(U(y)-U(x))/2}$.

(a) Let $L_+^1(E)$ be endowed with the bounded weak* topology. If

$$\varliminf_{|x| \to \infty} r(x) > 2d e^{\|F\|_\infty},$$

then Condition 3.10 holds. Moreover, if Condition 3.11 also holds, then under $\mathbb{P}_x$, the law of (μ_t, Q_t) satisfies an LDP with good and convex rate function $I : \Lambda \to [0, \infty]$.

(b) Let $L_+^1(E)$ be endowed with the strong topology. If

$$\lim_{|x| \to \infty} r(x) = \infty,$$

then Condition 3.4 holds. Moreover, if Condition 3.11 also holds, then under $\mathbb{P}_x$, the law of (μ_t, Q_t) satisfies an LDP with good and convex rate function $I : \Lambda \to [0, \infty]$.

Proof. Here we only give the proof of (a); the proof of (b) is similar. We first check Condition 3.10. Let $u_n(x) = e^{U(x)/2}$ for any $x \in \mathbb{Z}^d$ and $n \geq 0$. Since u_n do not depend on n, items (a)-(d) in Condition 3.4 are automatically satisfied. Note that

$$\hat{u}(x) = \frac{e^{U(x)/2}}{\sum_{y:(x,y) \in E} p_{xy} e^{U(y)/2}} = \frac{\sum_{y:|x-y|=1} e^{-(U(y)-U(x))/2 + F(x,y)/2}}{\sum_{y:|x-y|=1} e^{F(x,y)/2}} \geq \frac{r(x)}{2d e^{\|F\|_\infty}}.$$

Since $\underline{\lim}_{|x|\to\infty} \hat{u}(x) > 1$, there exists $\kappa > 1$ such that $\underline{\lim}_{|x|\to\infty} \hat{u}(x) \geq \kappa$. Letting $\ell = \log \kappa > 0$, it is easy to see that the set $\{x \in \mathbb{Z}^d : \log \hat{u}(x) \leq \ell\}$ is finite. This implies item (e) in Condition 3.10.

We next prove the joint LDP. It is easy to see that Assumptions 2.3, 2.5, and 2.6 are trivial. Similarly to the above proof, we can also validate item (c) in Lemma 3.14 with $u(x) = e^{U(x)/2}$. It then follows from [37, Theorems 8.0.2 and 15.0.1] that Assumption 2.4 holds. By Theorem 3.13, we complete the proof of this lemma. $\qquad\square$

Note that if $\psi_x(\mathrm{d}t) = C_x e^{-C_x t}\mathrm{d}t$ is chosen to be an exponential distribution for each x, then the above model reduces to the continuous-time Markov chain model studied in [6, Section 10.2]. It is easy to check that

$$e^{-\|F\|_\infty/2} r(x) \leq C_x \leq e^{\|F\|_\infty/2} r(x).$$

Then $\lim_{|x|\to\infty} r(x) = \infty$ if and only if $\lim_{|x|\to\infty} C_x = \infty$. Hence our compactness conditions in item (a) exactly coincide with the ones proposed in [6] when $L^1_+(E)$ is endowed with the bounded weak* topology. However, since we do not need to verify Condition 7.1, our compactness conditions in item (b) is weaker than the ones proposed in [6] when $L^1_+(E)$ is endowed with the strong topology.

Remark 3.21. When $U \in C^1(\mathbb{R}^d)$, we consider the orthogonal decomposition

$$\nabla U(x) = \langle \nabla U(y), \hat{y}\rangle \hat{y} + W(y), \quad y \in \mathbb{R}^d \setminus \{0\},$$

where $\hat{y} = y/|y|$, $\langle y, W(y)\rangle = 0$, and $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^d$. We say that the potential $U \in C^1(\mathbb{R}^d)$ has diverging radial variation which dominates the transversal variation [6] if there exist $\alpha \in [0, 1)$ and $C > 0$ such that

$$\lim_{|x|\to\infty} \langle \nabla U(y), \hat{y}\rangle = \infty, \qquad |W(y)| \leq \frac{\alpha}{d^{1/2}} \langle \nabla U(y), \hat{y}\rangle + C.$$

In fact, it follows from [6, Lemma 10.3] that if $U \in C^1(\mathbb{R}^d)$ has diverging radial variation which dominates the transversal variation, then $\lim_{|x|\to\infty} r(x) = \infty$.

4 Proof of Theorem 3.6

Note that $L^1(E)$ endowed with the bounded weak* topology is a locally convex, complete linear topological space and a completely regular space, i.e. for every closed set $C \subset L^1(E)$ and every element $Q \in L^1(E) \setminus C$, there exists a continuous function $f : L^1(E) \to [0, 1]$ such that $f(Q) = 1$ and $f(Q') = 0$ for any $Q' \in C$) [36, Theorem 2.7.2].

It is a well-known result that if an exponentially tight family of probability measures satisfies a weak LDP with rate function I, then I is good and the (full) LDP holds [15, Lemma 1.2.18]. Hence to prove Theorem 3.6, we will first prove the exponential tightness for the empirical measure and empirical flow under Condition 3.2, and then prove the weak joint LDP without any compactness conditions. We will directly consider the case where ξ starts from a general initial distribution γ (see Remark 3.8).

4.1 Exponential local martingales

We start by considering the change of probability measures for Markov renewal processes. Let Γ be the set of measurable functions (F, h) defined by

$$\Gamma = \left\{ (F, h) : F : E \to \mathbb{R} \text{ such that } \sum_{z \in V} p_{xz}\, e^{F(x,z)} < \infty, \right.$$

$$\left. h : V \times (0, \infty] \to \mathbb{R} \text{ such that } \int_{(0,\infty)} e^{sh(x,s)}\psi_x(\mathrm{d}s) < \infty \text{ for any } x \in V \right\}.$$

$$\tag{4.1}$$

For any $(F, h) \in \Gamma$, let $g_{F,h} : V \to \mathbb{R}$ be the function given by

$$g_{F,h}(x) = \log \sum_{z \in V} p_{xz} e^{F(x,z)} + \log \int_{(0,\infty)} e^{sh(x,s)} \psi_x(\mathrm{d}s).$$

To proceed, we define a new transition kernel (P^F, Ψ^h) as

$$p_{xy}^F = \frac{p_{xy}\, e^{F(x,y)}}{\sum_{z \in V} p_{xz}\, e^{F(x,z)}}, \qquad \psi_x^h(\mathrm{d}u) = \frac{e^{u\, h(x,u)}\, \psi_x(\mathrm{d}u)}{\int_{(0,\infty)} e^{s\, h(x,s)}\, \psi_x(\mathrm{d}s)}, \tag{4.2}$$

and let $\mathbb{P}_x^{F,h}$ be the probability measure under which (X, τ) is a Markov renewal process with transition kernel (P^F, Ψ^h) and initial state $X_0 = x$. Note that the semi-Markov process ξ may be explosive under $\mathbb{P}_x^{F,h}$. As a result, we need to consider $\mathbb{P}_x^{F,h}$ and $\mathbb{P}_x$ restricted to the set $\{N_t < \infty\}$, i.e.

$$\mathbb{P}_{x,t}^{F,h}(A) = \mathbb{P}_x^{F,h}(A \cap \{N_t < \infty\}), \quad \mathbb{P}_{x,t}(A) = \mathbb{P}_x(A \cap \{N_t < \infty\}), \quad A \in \mathcal{F}_\infty.$$

Moreover, we denote by $\mathbb{P}_{x,t}^{F,h}|_{\mathcal{F}_{N_t+1}}$ and $\mathbb{P}_{x,t}|_{\mathcal{F}_{N_t+1}}$ the restrictions of $\mathbb{P}_{x,t}^{F,h}$ and $\mathbb{P}_{x,t}$ to $\mathcal{F}_{N_t+1}$, respectively. It is easy to verify that $\mathbb{P}_{x,t}^{F,h}|_{\mathcal{F}_{N_t+1}}$ is absolutely continuous with respect to $\mathbb{P}_{x,t}|_{\mathcal{F}_{N_t+1}}$ and has the Radon-Nykodim derivative

$$\frac{1}{t} \log \frac{\mathrm{d}\mathbb{P}_{x,t}^{F,h}|_{\mathcal{F}_{N_t+1}}}{\mathrm{d}\mathbb{P}_{x,t}|_{\mathcal{F}_{N_t+1}}}$$

$$= \frac{1}{t} \sum_{i=1}^{N_t+1} \log \frac{e^{F(X_{i-1}, X_i)}}{\sum_{z \in V} p_{X_{i-1}, z}\, e^{F(X_{i-1}, z)}} + \frac{1}{t} \sum_{i=1}^{N_t+1} \log \frac{e^{\tau_i\, h(X_{i-1}, \tau_i)}}{\int_{(0,\infty)} e^{s\, h(X_{i-1}, s)}\, \psi_{X_{i-1}}(\mathrm{d}s)}$$

$$= \sum_{(x,y) \in E} Q_t(x,y) \left[F(x,y) - \left(\log \sum_{z \in V} p_{xz} e^{F(x,z)} + \log \int_{(0,\infty)} e^{sh(x,s)} \psi_x(\mathrm{d}s) \right) \right]$$

$$+ \sum_x \int_{(0,\infty]} h(x,u) \mu_t(x, \mathrm{d}u) + \frac{\tau_{N_t+1} - t + S_{N_t}}{t} h(X_{N_t}, \tau_{N_t+1})$$

$$= \langle Q_t, F - g_{F,h} \rangle + \langle \mu_t, h \rangle + \frac{S_{N_t+1} - t}{t} h(X_{N_t}, \tau_{N_t+1}) \tag{4.3}$$

where $\langle Q_t, F \rangle = \sum_{(x,y) \in E} Q_t(x,y) F(x,y)$ and $\langle Q_t, g_{F,h} \rangle = \sum_{(x,y) \in E} Q_t(x,y) g_{F,h}(x)$.

For convenience, for any Borel measurable function f on $V \times (0, \infty]$, let

$$\langle \hat{\mu}_t, f \rangle = \langle \mu_t, f \rangle + \frac{S_{N_t+1} - t}{t} f(X_{N_t}, \tau_{N_t+1}) = \frac{1}{t} \sum_{k=1}^{N_t+1} \tau_k f(X_{k-1}, \tau_k).$$

Then as an immediate consequence of the Radon-Nykodim derivative (4.3), we obtain the following result.

Lemma 4.1. *For any $(F, h) \in \Gamma$ and $t \geq 0$, let $\mathcal{M}_t^{F,h} : \Omega \to (0, \infty)$ be the function defined by*

$$\mathcal{M}_t^{F,h} = \exp \left\{ t \big[\langle Q_t, F - g_{F,h} \rangle + \langle \hat{\mu}_t, h \rangle \big] \right\}. \tag{4.4}$$

Then for each $x \in V$ and $t \geq 0$, we have $\mathbb{E}_x\big(\mathcal{M}_t^{F,h} \big) \leq 1$.

Proof. Note that $\{N_t < \infty\} \in \mathcal{F}_{N_t+1}$. It follows from (4.3) that $\mathbb{E}_x(\mathcal{M}_t^{F,h}) = \mathbb{P}_x^{F,h}(N_t < \infty) \leq 1$. $\square$

In fact, it is easy to check that under $\mathbb{P}_x$, the process $\mathcal{M}^{F,h}$ is a positive local martingale and a supermartingale with respect to $(\mathcal{F}_{N_t+1})_{t \geq 0}$. The next statement can be deduced from Lemma 4.1 by choosing specific F and h.

Lemma 4.2. *Let $u\colon V \to (0,\infty)$ be a function satisfying $Pu(x) < \infty$ for any $x \in V$. Let $A \subset V$ be a set satisfying $u(x)/Pu(x) \leq \psi_x(e^{\zeta(x)\tau})$ for any $x \in A^c$. Let $F(x,y) = \log[u(y)/u(x)]$ for any $(x,y) \in E$ and*

$$h(x,s) = h^{u,A}(x,s) := \frac{\log \frac{u}{Pu}(x)}{s} 1_A(x) + L\frac{u}{Pu}(x) 1_{A^c}(x), \quad (x,s) \in V \times (0,\infty]. \quad (4.5)$$

For any $t \geq 0$, let $\mathcal{M}_t^{u,A}\colon \Omega \to (0,\infty)$ be the function defined as

$$\mathcal{M}_t^{u,A} = \frac{u(X_{N_t+1})}{u(X_0)} \exp\left\{ t \left\langle \hat{\mu}_t, h^{u,A} \right\rangle \right\}.$$

Then $(F,h) \in \Gamma$ and $\mathbb{E}_x\big(\mathcal{M}_t^{u,A}\big) \leq 1$ for any $x \in V$ and $t \geq 0$.

Proof. For any $x \in V$, it is clear that $\sum_{y \in V} p_{xy} e^{F(x,y)} = Pu(x)/u(x)$ and

$$\int_{(0,\infty)} e^{sh(x,s)} \psi_x(\mathrm{d}s) = \frac{u(x)}{Pu(x)} 1_A(x) + \psi_x\left(e^{\theta_x(u(x)/Pu(x))\tau} \right) 1_{A^c}(x).$$

Therefore, by item (c) in Lemma 3.1, we have $\int_{(0,\infty)} e^{sh(x,s)} \psi_x(\mathrm{d}s) = u(x)/Pu(x)$. This implies that $(F,h) \in \Gamma$ and $g_{F,h} \equiv 0$. On the other hand, it is easy to check that $\langle Q_t, F \rangle = u(X_{N_t+1})/u(X_0)$. This completes the proof of this lemma. $\qquad\square$

4.2 Exponential tightness

We will next prove the exponential tightness of the empirical measure and empirical flow under Condition 3.2 with item (c) is replaced by item (c*) in Remark 3.8.

Let the function $\hat{u}$, the sequence of functions u_n, the set K, and the constants $c, C_\gamma, \sigma, \eta$ be as in Condition 3.2 and item (c*) in Remark 3.8. Recall the definition of $h^{u,A}$ in (4.5). Let $h^{\hat{u}}\colon V \times (0,\infty] \to \mathbb{R}$ be the function defined by

$$h^{\hat{u}}(x,s) = \lim_{n \to \infty} h^{u_n,K}(x,s) = \frac{\log \hat{u}(x)}{s} 1_K(x) + L\hat{u}(x) 1_{K^c}(x), \quad (x,s) \in V \times (0,\infty],$$

where we have used the condition that u_n/Pu_n converges pointwise to $\hat{u}$.

Lemma 4.3. *Assume Condition 3.2 to hold. Then*

$$\mathbb{E}_\gamma\left(e^{t\langle \hat{\mu}_t, h^{\hat{u}} \rangle} \right) \leq \frac{C_\gamma}{c}, \qquad \mathbb{E}_\gamma\left(e^{t\langle \mu_t, h^{\hat{u}} \rangle} \right) \leq \frac{C_\gamma e^N}{c}, \quad (4.6)$$

where $N = 1 \vee (-\inf_{x \in K} \log \hat{u}(x))$.

Proof. By Fatou's lemma, we have

$$\mathbb{E}_\gamma\left(e^{t\langle \hat{\mu}_t, h^{\hat{u}} \rangle} \right) \leq \sum_{x \in V} \gamma(x) \varliminf_{n \to \infty} \mathbb{E}_x\left(e^{t\langle \hat{\mu}_t, h^{u_n,K} \rangle} \right). \quad (4.7)$$

For any $x \in K^c$, it follows from items (d) and (f) in Condition 3.2 that there exists $n_x \in \mathbb{N}$ such that $u_n(x)/Pu_n(x) < \psi_x(e^{\zeta(x)\tau})$ for any $n \geq n_x$. By Lemma 4.2 and item (b) in Condition 3.2, we have

$$\mathbb{E}_x\left(e^{t\langle \hat{\mu}_t, h^{u_n,K} \rangle} \right) = \mathbb{E}_x\left(\frac{u_n(X_0)}{u_n(X_{N_t+1})} \mathcal{M}_t^{u_n,K} \right) \leq \frac{u_n(x)}{c}, \quad n \geq n_x. \quad (4.8)$$

Combining (4.7), (4.8), and item (c*) in Remark 3.8, we have

$$\mathbb{E}_\gamma\left(e^{t\langle \hat{\mu}_t, h^{\hat{u}} \rangle} \right) \leq \sum_{x \in V} \gamma(x) \frac{u_n(x)}{c} \leq \frac{C_\gamma}{c}.$$

Moreover, by item (f) in Condition 3.2, we obtain

$$h^{\hat{u}}(x, s) \geq -\frac{N}{s} 1_K(x) - \sigma L \eta(x) 1_{K^c}(x). \tag{4.9}$$

Since $N > 0$ and $L\eta(x) < 0$ for any $x \in K^c$, we have

$$\begin{aligned}
\langle \hat{\mu}_t, h^{\hat{u}} \rangle &= \langle \mu_t, h^{\hat{u}} \rangle + \frac{S_{N_t+1} - t}{t} h^{\hat{u}}(X_{N_t}, \tau_{N_t+1}) \\
&\geq \langle \mu_t, h^{\hat{u}} \rangle - \frac{S_{N_t+1} - t}{t} \frac{N}{\tau_{N_t+1}} \\
&\geq \langle \mu_t, h^{\hat{u}} \rangle - \frac{N}{t}.
\end{aligned}$$

This implies the second inequality in (4.6). $\qquad\square$

For any $B \subset V$, let $(T_B^k)_{k \geq 0}$ be the sequence of stopping times defined by

$$T_B^0 = 0, \qquad T_B^k = \inf\left\{ n \geq T_B^{k-1} + 1 : X_n \in B \right\}, \quad k \geq 1. \tag{4.10}$$

Clearly, T_B^k is the kth hitting time of $(X_n)_{n \geq 0}$ on the set B. The first hitting time T_B^1 is always abbreviated as T_B in what follows. If we only focus on the behavior of (X, τ) in the set B, we obtain a new process $(\bar{X}, \bar{\tau}) = \{(\bar{X}_k)_{k \geq 0}, (\bar{\tau}_k)_{k \geq 1}\}$, which is defined by

$$\bar{X}_k = X_{T_B^k}, \qquad \bar{\tau}_{k+1} = \sum_{i=T_B^k+1}^{T_B^{k+1}} \tau_i, \quad k \geq 0. \tag{4.11}$$

Proposition 4.4. *Suppose that (X, τ) satisfies Assumptions 2.3 and 2.4. Then $\{(\bar{X}_k)_{k \geq 1}, (\bar{\tau}_k)_{k \geq 2}\}$ is a Markov renewal process and also satisfies Assumptions 2.3 and 2.4. In other words, we have*

(a) *$(\bar{X}_k)_{k \geq 1}$ is an irreducible and recurrent discrete-time Markov chain with state space B and transition probability matrix $(\bar{p}_{xy})_{x,y \in B}$ given by $\bar{p}_{xy} = \mathbb{P}_x\left(X_{T_B} = y\right)$.*

(b) *$(\bar{\tau}_k)_{k \geq 2}$ is a sequence of positive and finite random variables such that conditioned on $(\bar{X}_k)_{k \geq 1}$ the random variables $(\bar{\tau}_k)_{k \geq 2}$ are independent and the waiting time matrix $(\bar{\psi}_{xy})_{x,y \in B}$ is given by*

$$\bar{\psi}_{xy}(\cdot) = \mathbb{P}\left(\sum_{i=1}^{T_B} \tau_i \in \cdot \,\middle|\, X_0 = x, X_{T_B} = y \right), \quad x, y \in B,$$

where the definition of the waiting time matrix can be found in Definition 2.1.

Proof. (a) Since $(X_k)_{k \geq 0}$ is irreducible and recurrent, we have $T_B^k < \infty$, $\mathbb{P}_x$-a.s. for any $x \in V$ and $k \geq 0$. For any $n \geq 0$ and $x_0, x_1, \cdots, x_{n+1} \in B$,

$$\mathbb{P}\left(\bar{X}_{n+1} = x_{n+1} \middle| \bar{X}_n = x_n, \cdots, \bar{X}_0 = x_0 \right) = \mathbb{P}\left(X_{T_B^{n+1}} = x_{n+1} \middle| X_{T_B^n} = x_n, \cdots, X_{T_B^0} = x_0 \right)$$
$$= \mathbb{P}_{x_n}\left(X_{T_B} = x_{n+1} \right),$$

where the last step follows from the strong Markov property. In fact, we can obtain the recurrence of $(\bar{X}_k)_{k \geq 1}$ directly from the recurrence of $(X_k)_{k \geq 0}$. For any $x, y \in B$, since $(X_k)_{k \geq 0}$ is irreducible, there exist a positive integer $n \geq 1$ and a sequence of states $x_0, x_1, \cdots, x_n \in V$ with $x_0 = x$ and $x_n = y$ such that $p_{x_0, x_1} p_{x_1, x_2} \cdots p_{x_{n-1}, x_n} > 0$. Select all states in $\{x_k\}_{0 \leq k \leq n}$ in the set B and write them as $x_{i_1}, \cdots, x_{i_s}$ with $0 = i_1 < i_2 < \cdots < i_s = n$. Then

$$\bar{p}_{x_{i_j} x_{i_{j+1}}} \geq p_{x_{i_j} x_{i_j+1}} \cdots p_{x_{i_{j+1}-1} x_{i_{j+1}}} > 0, \quad 1 \leq j \leq s - 1.$$

This implies that $(\bar{X}_k)_{k \geq 1}$ is irreducible.

(b) Since $T_B^k < \infty$, $\mathbb{P}_x$-a.s. for any $x \in V$ and $k \geq 0$, it is easy to see that $(\bar{\tau}_{k+1})_{k \geq 0}$ is a sequence of positive and finite random variables. Note that conditioned on $(X_k)_{k \geq 0}$ the random variables $(\tau_k)_{k \geq 1}$ are independent and have distribution $\mathbb{P}(\tau_{i+1} \in \cdot \mid (X_k)_{k \geq 0}) = \psi_{X_i, X_{i+1}}(\cdot)$. For any $f \in \mathcal{B}_b(0, \infty)$, we have

$$\mathbb{E}\left[f(\bar{\tau}_k) | (X_k)_{k \geq 0}\right] = \int f\left(\sum_{i=T_B^{k-1}+1}^{T_B^k} t_i\right) \prod_{i=T_B^{k-1}+1}^{T_B^k} \psi_{X_{i-1}X_i}(\mathrm{d}t_i), \tag{4.12}$$

where we use the fact that $(T_B^k)_{k \geq 0}$ is $\sigma((X_k)_{k \geq 0})$-measurable. For any $n \geq 2$ and $f_1, \cdots, f_n \in \mathcal{B}_b(0, \infty)$, similarly to (4.12), we have

$$\mathbb{E}\left[f_1(\bar{\tau}_1) \cdots f_n(\bar{\tau}_n) | (X_k)_{k \geq 0}\right] = \int f_1\left(\sum_{i=T_B^0+1}^{T_B^1} t_i\right) \cdots f_n\left(\sum_{i=T_B^{n-1}+1}^{T_B^n} t_i\right) \prod_{i=T_B^0+1}^{T_B^n} \psi_{X_{i-1}X_i}(\mathrm{d}t_i)$$

$$= \mathbb{E}\left[f_1(\bar{\tau}_1) | (X_k)_{k \geq 0}\right] \cdots \mathbb{E}\left[f_n(\bar{\tau}_n) | (X_k)_{k \geq 0}\right]. \tag{4.13}$$

Note that $\bar{X}_k = X_{T_B^k}$. By the strong Markov property of $(X_k)_{k \geq 0}$, it is clear that conditioned on $\{\bar{X}_{k-1} = x, \bar{X}_k = y\}$ the random variable $\mathbb{E}\left[f(\bar{\tau}_k) | (X_k)_{k \geq 0}\right]$ is independent of $(X_k)_{0 \leq k \leq T_B^{k-1}}$ and $(X_k)_{k \geq T_B^k}$. Moreover, for any $x, y \in B$, the distribution of $\mathbb{E}[f(\bar{\tau}_k) | (X_k)_{k \geq 0}]$ conditioned on $\{\bar{X}_{k-1} = x, \bar{X}_k = y\}$ is equal to the distribution of $\mathbb{E}[f(\bar{\tau}_1) | (X_k)_{k \geq 0}]$ conditioned on $\{\bar{X}_0 = x, \bar{X}_1 = y\}$. It follows from (4.13) that

$$\mathbb{E}\left[f_1(\bar{\tau}_1) \cdots f_n(\bar{\tau}_n) \Big| (\bar{X}_k)_{k \geq 0}\right] = \mathbb{E}\left[\mathbb{E}\left[f_1(\bar{\tau}_1) \cdots f_n(\bar{\tau}_n) | (X_k)_{k \geq 0}\right] \Big| (\bar{X}_k)_{k \geq 0}\right]$$

$$= \mathbb{E}\left[f_1(\bar{\tau}_1) \Big| (\bar{X}_k)_{k \geq 0}\right] \cdots \mathbb{E}\left[f_n(\bar{\tau}_n) \Big| (\bar{X}_k)_{k \geq 0}\right].$$

Moreover, for any $x, y \in B$, we have

$$\mathbb{E}\left[f(\bar{\tau}_k) \Big| \bar{X}_{k-1} = x, \bar{X}_k = y\right] = \mathbb{E}\left[\mathbb{E}\left[f(\bar{\tau}_k) | (X_k)_{k \geq 0}\right] \Big| \bar{X}_{k-1} = x, \bar{X}_k = y\right]$$

$$= \mathbb{E}\left[f(\bar{\tau}_1) \Big| \bar{X}_0 = x, \bar{X}_1 = y\right].$$

This completes the proof of this proposition. $\qquad \square$

Remark 4.5. In the above proposition, we have proved that conditioned on $(\bar{X}_k)_{k \geq 0}$, the random variables $(\bar{\tau}_k)_{k \geq 1}$ are independent. This implies that $(X, \tau) = \{(\bar{X}_k)_{k \geq 0}, (\bar{\tau}_k)_{k \geq 1}\}$ is a delayed Markov renewal process, whose transition probability matrix and waiting time matrix of the first step is different from the remaining steps [25, Chapter 4.12].

Lemma 4.6. *Let K be a finite subset of V. Then for any $\ell \in \mathbb{N}$, there exists a real sequence $A_\ell \uparrow \infty$ such that*

$$\overline{\lim_{t \to \infty}} \frac{1}{t} \log \mathbb{P}_\gamma(\langle Q_t, 1_K \rangle > A_\ell) \leq -\ell, \tag{4.14}$$

where $\langle Q_t, 1_K \rangle = \sum_{(x,y) \in E} Q_t(x, y) 1_K(x)$.

Proof. Let $(\bar{X}, \bar{\tau})$ be defined as in (4.11) with $B = K$. By Remark 4.5, we have seen that $(\bar{X}, \bar{\tau})$ is a delayed Markov renewal process. Similarly to Markov renewal processes, we can also define the nth jump time $\bar{S}_n$, the number $\bar{N}_t$ of jumps up to time t, and the

empirical flow $\bar{Q}_t$ for $(\bar{X}, \bar{\tau})$. Note that $\sum_{i=1}^{N_t} 1_K(X_i)$ is the number of times that ξ jumps into the set K in the time interval $(0, t]$. It is easy to see that

$$\bar{N}_t + 1 = 1 + \sum_{i=1}^{N_t} 1_K(X_i) \geq \sum_{i=1}^{N_t+1} 1_K(X_{i-1}) = t\langle Q_t, 1_K \rangle.$$

By the exponential Chebyshev inequality, we have

$$\mathbb{P}_\gamma(\langle Q_t, 1_K \rangle > A_\ell) \leq \mathbb{P}_\gamma(\bar{N}_t + 1 > A_\ell t) = \mathbb{P}_\gamma(\bar{S}_{\lfloor A_\ell t \rfloor} \leq t) \leq e^t \mathbb{E}_\gamma\left(e^{-\bar{S}_{\lfloor A_\ell t \rfloor}}\right), \quad (4.15)$$

where $\lfloor A_\ell t \rfloor$ denotes the integer part of $A_\ell t$. By Proposition 4.4 and Remark 4.5, it is clear that conditioned on $(\bar{X}_k)_{k\geq 0}$, the random variables $(\bar{\tau}_k)_{k\geq 1}$ are independent. Then we obtain

$$\mathbb{E}_\gamma\left(e^{-\bar{S}_{\lfloor A_\ell t \rfloor}}\right)$$
$$= \sum_{x_1,\cdots,x_{\lfloor A_\ell t \rfloor} \in K} \mathbb{P}_\gamma\left(\bar{X}_1 = x_1, \cdots, \bar{X}_{\lfloor A_\ell t \rfloor} = x_{\lfloor A_\ell t \rfloor}\right) \mathbb{E}_\gamma\left[e^{-\bar{S}_{\lfloor A_\ell t \rfloor}} \Big| \bar{X}_1 = x_1, \cdots, \bar{X}_{\lfloor A_\ell t \rfloor} = x_{\lfloor A_\ell t \rfloor}\right]$$
$$= \sum_{x_1,\cdots,x_{\lfloor A_\ell t \rfloor} \in K} \mathbb{P}_\gamma\left(\bar{X}_1 = x_1\right) \mathbb{E}_\gamma\left[e^{-\bar{\tau}_1} \Big| \bar{X}_1 = x_1\right] \prod_{i=2}^{\lfloor A_\ell t \rfloor} \left(\bar{p}_{x_{i-1}x_i} \bar{\psi}_{x_{i-1}x_i}\left(e^{-\tau}\right)\right)$$
$$\leq \left(\sup_{x,y \in K} \bar{\psi}_{xy}(e^{-\tau})\right)^{\lfloor A_\ell t \rfloor - 1} \mathbb{E}_\gamma\left[e^{-\bar{\tau}_1}\right],$$
$$(4.16)$$

where $\bar{\psi}_{xy}(e^{-\tau}) = \int_{(0,\infty)} e^{-s} \bar{\psi}_{xy}(\mathrm{d}s)$. Since K is finite, it is clear that $\sup_{x,y \in K} \bar{\psi}_{xy}(e^{-\tau}) < 1$. Combining (4.15) and (4.16), we have

$$\varlimsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\gamma(\langle Q_t, 1_K \rangle > A_\ell) \leq 1 + (A_\ell - 1)\log C.$$

This completes the proof by choosing $A_\ell = 1 - (1+\ell)/\log C$. $\qquad\square$

The following proposition states the exponential tightness of the empirical measure and empirical flow.

Proposition 4.7. *Assume Condition 3.2 to hold. Then there exists a sequence $\{\mathcal{K}_\ell\}$ of compact sets in $\mathcal{P}(V \times (0,\infty])$ and a real sequence $A_\ell \uparrow \infty$ such that for any $\ell \in \mathbb{N}$*

$$\varlimsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\gamma(\mu_t \notin \mathcal{K}_\ell) \leq -\ell, \quad (4.17)$$

$$\varlimsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\gamma(\|Q_t\| > A_\ell) \leq -\ell. \quad (4.18)$$

In particular, the empirical measure and empirical flow are exponentially tight.

Proof. We first prove (4.18). We consider the exponential local martingale in Lemma 4.1 by choosing $F \equiv 0$ and

$$h(x,s) = h^\eta(x,s) := \frac{N}{\sigma s} 1_K(x) + L\eta(x) 1_{K^c}(x), \quad (x,s) \in V \times (0,\infty],$$

where N is as in Lemma 4.3. By Lemma 3.1, it is easy to check that $(F, h) \in \Gamma$ and $g_{F,h} = (N/\sigma)1_K + \log \eta 1_{K^c}$. It then follows form (4.4) that

$$\mathcal{M}_t^{F,h} = \exp\left\{t\left[\left\langle Q_t, -\frac{N}{\sigma}1_K - \log\eta 1_{K^c}\right\rangle + \langle \hat{\mu}_t, h^\eta \rangle\right]\right\}.$$

Combining (4.9) and the first inequality in Lemma 4.3, we have

$$\mathbb{E}_\gamma\left(e^{t\langle\hat{\mu}_t,-\sigma h^\eta\rangle}\right) = \mathbb{E}_\gamma\left(e^{t\langle\hat{\mu}_t,-\frac{N}{s}1_K-\sigma L\eta1_{K^c}\rangle}\right) \le \mathbb{E}_\gamma\left(e^{t\langle\hat{\mu}_t,h^{\hat{u}}\rangle}\right) \le \frac{C_\gamma}{c}.$$

By exponential Chebyshev inequality, we obtain

$$\mathbb{P}_\gamma\left(\langle\hat{\mu}_t,-\sigma h^\eta\rangle > \ell\right) \le \frac{C_\gamma}{c}e^{-t\ell}.$$

On the other hand, by Lemma 4.6, there exists a real sequence $A'_\ell\uparrow\infty$ such that

$$\varlimsup_{t\to\infty}\frac{1}{t}\log\mathbb{P}_\gamma\left(\langle Q_t,1_K\rangle > A'_\ell\right) \le -\ell. \tag{4.19}$$

Hence it is enough to show that there exists a sequence $A_\ell\uparrow\infty$ such that for any $t > 0$ and $\ell\in\mathbb{N}$,

$$\mathbb{P}_\gamma\left(\|Q_t\| > A_\ell,\,\langle\hat{\mu}_t,-\sigma h^\eta\rangle\le\ell,\,\langle Q_t,1_K\rangle\le A'_\ell\right) \le e^{-t\ell}. \tag{4.20}$$

Since $\log\eta < 0$, we have

$$\mathbb{P}_\gamma\left(\|Q_t\| > A_\ell,\,\langle\hat{\mu}_t,-\sigma h^\eta\rangle\le\ell,\langle Q_t,1_K\rangle\le A'_\ell\right)$$
$$= \mathbb{E}_\gamma\left(e^{t[\log\eta\|Q_t\|+\langle Q_t,(\frac{N}{\sigma}-\log\eta)1_K\rangle-\langle\hat{\mu}_t,h^\eta\rangle]}\,\mathcal{M}_t^{F,h}\,1_{\{\|Q_t\|>A_\ell\}}\,1_{\{\langle\hat{\mu}_t,-\sigma h^\eta\rangle\le\ell\}}\,1_{\{\langle Q_t,1_K\rangle\le A'_\ell\}}\right)$$
$$\le \exp\left\{t\left[\log\eta A_\ell + \left(\frac{N}{\sigma}-\log\eta\right)A'_\ell + \frac{1}{\sigma}\ell\right]\right\},$$

where we have used Lemma 4.1 in the last step. The proof of (4.20) is now completed by choosing

$$A_\ell = \frac{(N/\sigma - \log\eta)A'_\ell + (1+1/\sigma)\ell}{-\log\eta}.$$

Recall that the closed ball in $L^1_+(E)$ is compact with respect to the bounded weak* topology. Then the exponential tightness of the empirical flow follows from (4.18).

We next prove (4.17). For a sequence of constants $a_m\uparrow\infty$ to be chosen later, set $W_m = \{x\in V : L\hat{u}(x)\le a_m\}\cup K$. In view of items (e) and (f) in Condition 3.2, it is clear that W_m is a finite subset of V. Now set

$$\mathcal{K}^1_\ell = \bigcap_{m\ge\ell}\left\{\mu\in\mathcal{P}(V\times(0,\infty]) : \mu\left(W_m^c\times(0,\infty]\right)\le\frac{1}{m}\right\},$$
$$\mathcal{K}^2_\ell = \left\{\mu\in\mathcal{P}(V\times(0,\infty]) : \left\langle\mu,\frac{1}{s}\right\rangle\le A_\ell\right\},$$

where $\langle\mu,1/s\rangle = \sum_{x\in V}\int_{(0,\infty]}\mu(x,\mathrm{d}s)/s$. Then for any $\mu\in\mathcal{K}_\ell = \mathcal{K}^1_\ell\cap\mathcal{K}^2_\ell$, we have

$$\mu(V\times(0,\epsilon)) = \sum_{x\in V}\int_{(0,\epsilon)}\frac{s}{s}\mu(x,\mathrm{d}s) \le \epsilon\left\langle\mu,\frac{1}{s}\right\rangle \le \epsilon A_\ell,$$

$$\mu\left(\left(W_m\times[\epsilon,\infty]\right)^c\right) \le \mu(W_m^c\times(0,\infty]) + \mu(V\times(0,\epsilon)) \le \frac{1}{m} + \epsilon A_\ell, \quad m\ge\ell.$$

Since $W_m\times[\epsilon,\infty]$ a compact subset of $V\times(0,\infty]$, it follows from Prokhorov's theorem that $\mathcal{K}_\ell$ is a compact subset of $\mathcal{P}(V\times(0,\infty])$. Since

$$\varlimsup_{t\to\infty}\frac{1}{t}\log\mathbb{P}_\gamma\left(\mu_t\notin\mathcal{K}_\ell\right) \le \varlimsup_{t\to\infty}\frac{1}{t}\log\left(P_\gamma\left(\mu_t\notin\mathcal{K}^1_\ell\right) + P_\gamma\left(\mu_t\notin\mathcal{K}^2_\ell\right)\right),$$

we only need to prove

$$\varlimsup_{t\to\infty} \frac{1}{t}\log\mathbb{P}_\gamma\left(\mu_t \notin \mathcal{K}_\ell^1\right) \leq -\ell \quad \text{and} \quad \varlimsup_{t\to\infty} \frac{1}{t}\log\mathbb{P}_\gamma\left(\mu_t \notin \mathcal{K}_\ell^2\right) \leq -\ell. \tag{4.21}$$

Note that

$$\left\langle \mu_t, \frac{1}{s} \right\rangle = \frac{1}{t}N_t + \frac{t - S_{N_t}}{t\tau_{N_t+1}} \leq \frac{1}{t}(N_t + 1) = \|Q_t\|.$$

Then we can obtain the second inequality in (4.21) from (4.18), i.e.

$$\begin{aligned}
\varlimsup_{t\to\infty} \frac{1}{t}\log\mathbb{P}_\gamma\left(\mu_t \notin \mathcal{K}_\ell^2\right) &= \varlimsup_{t\to\infty} \frac{1}{t}\log\mathbb{P}_\gamma\left(\left\langle \mu_t, \frac{1}{s}\right\rangle > A_\ell\right) \\
&\leq \varlimsup_{t\to\infty} \frac{1}{t}\log\mathbb{P}_\gamma\left(\|Q_t\| > A_\ell\right) \\
&\leq -\ell.
\end{aligned}$$

We next prove the first inequality in (4.21). By item (f) in Condition 3.2, it is easy to see that $L\hat{u}(x) \geq -\sigma L\eta(x) \geq 0$ for any $x \in K^c$. Recall the definition of W_m, $h^{\hat{u}}$, and N. It is easy to see that

$$h^{\hat{u}}(x,s) = \frac{\log\hat{u}(x)}{s}1_K(x) + L\hat{u}(x)1_{K^c}(x) \geq -\frac{N}{s}1_K(x) + a_m 1_{W_m^c}(x).$$

Note that

$$\left\langle \mu_t, \frac{1}{s}1_K \right\rangle \leq \left\langle \hat{\mu}_t, \frac{1}{s}1_K \right\rangle = \langle Q_t, 1_K \rangle.$$

By the exponential Chebyshev inequality and Lemma 4.3, we obtain

$$\begin{aligned}
\mathbb{P}_\gamma\left(\mu_t\left(W_m^c \times (0,\infty]\right) > \frac{1}{m}\right) &\leq \mathbb{P}_\gamma\left(\left\langle \mu_t, h^{\hat{u}} + \frac{N}{s}1_K\right\rangle > \frac{a_m}{m}\right) \\
&\leq \mathbb{P}_\gamma\left(\langle \mu_t, h^{\hat{u}}\rangle + N\langle Q_t, 1_K\rangle > \frac{a_m}{m}\right) \\
&\leq \mathbb{P}_\gamma\left(\langle \mu_t, h^{\hat{u}}\rangle > \frac{a_m}{2m}\right) + \mathbb{P}_\gamma\left(N\langle Q_t, 1_K\rangle > \frac{a_m}{2m}\right) \\
&\leq \frac{C_\gamma e^N}{c}\exp\left\{-t\frac{a_m}{2m}\right\} + \mathbb{P}_\gamma\left(\langle Q_t, 1_K\rangle > \frac{a_m}{2Nm}\right).
\end{aligned} \tag{4.22}$$

From (4.19), the proof is now easily concluded by choosing $a_m = 2m^2 + 2NA_m'm$. $\qquad\square$

Remark 4.8. In fact, the empirical measure π_t defined in (2.2) is also exponentially tight and the proof is similar to the first inequality in (4.21). Moreover, at this time we only need item (f) in Condition 3.2 with $\sigma = 0$. Note that we will not resort to the compactness conditions anymore in the following proof of upper bound and lower bound. This means that the marginal LDP for the empirical measure π_t in Proposition 3.16 still holds under Condition 3.2 with $\sigma = 0$.

4.3 Upper bound

We next prove the upper bound of the LDP. Since we have proved the exponential tightness for the empirical measure and empirical flow, to prove the upper bound of the LDP for closed sets, we only need to prove the upper bound for compact sets [15, Lemma 1.2.18]. Before stating the upper bound, we introduce the following notation. For any Polish space $\mathcal{X}$, let $C_c(\mathcal{X})$ be the collection of all continuous functions $f : X \to \mathbb{R}$ with compact supports. For any functions $M : V \to [0,\infty)$, $\varphi \in C_c(V \times (0,\infty))$, and $c \in C_c(V)$

such that $0 \le c \le \zeta$ and $c(x) < \zeta(x)$ for any $x \in V$ with $\zeta(x) > 0$, let $h^{\varphi,c,M} : V \times (0, \infty] \to \mathbb{R}$ be the function defined by

$$h^{\varphi,c,M}(x,s) = \frac{\varphi_x(s)}{s} + c(x) \, 1_{(M(x),\infty]}(s), \quad (x,s) \in V \times (0, \infty].$$

Recall the definition of Γ in (4.1). Let Γ_0 be the subset of Γ defined by

$$\Gamma_0 = \left\{ (F,h) \in \Gamma : F \in C_c(E), h = h^{\varphi,c,M} \text{ for some } \varphi, c, M \text{ as above, and } g_{F,h} \equiv 0 \right\}.$$

For any $(F,h) \in \Gamma_0$, let $I_{F,h} : \Lambda \mapsto \mathbb{R}$ be the functional defined by

$$I_{F,h}(\mu, Q) = \langle Q, F \rangle + \langle \mu, h \rangle. \tag{4.23}$$

Based on the proof of Lemma 4.10, it is easy to see that $I_{F,h}$ is a lower semicontinuous function on Λ. For any $\delta > 0$, set

$$\mathcal{C}_\delta = \left\{ (\mu, Q) \in \Lambda : \max_{x \in V} |Q^+(x) - Q^-(x)| \le \delta \right\}. \tag{4.24}$$

Let $I_{F,h,\delta} : \Lambda \to \mathbb{R}$ be the functional defined by

$$I_{F,h,\delta}(\mu, Q) = \begin{cases} I_{F,h}(\mu, Q), & \text{if } (\mu, Q) \in \mathcal{C}_\delta, \\ \infty, & \text{otherwise.} \end{cases} \tag{4.25}$$

Lemma 4.9. *For any $(F,h) \in \Gamma_0$, $\delta > 0$, and measurable $\mathcal{O} \subset \Lambda$, we have*

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma((\mu_t, Q_t) \in \mathcal{O}) \le -\inf_{(\mu,Q) \in \mathcal{O}} I_{F,h,\delta}(\mu, Q). \tag{4.26}$$

Proof. We first prove that for any measurable $\mathcal{B} \subset \Lambda$,

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma\left((\mu_t, Q_t) \in \mathcal{B} \right) \le -\inf_{(\mu,Q) \in \mathcal{B}} I_{F,h}(\mu, Q). \tag{4.27}$$

Since $c \ge 0$ and $\varphi \in C_c(V \times (0, \infty))$, we have

$$(S_{N_t+1} - t) h(X_{N_t}, \tau_{N_t+1}) \ge (S_{N_t+1} - t) \frac{\varphi_{X_{N_t}}(\tau_{N_t+1})}{\tau_{N_t+1}} \ge -\|\varphi\|_\infty, \quad \mathbb{P}_\gamma\text{-a.s.}, \tag{4.28}$$

where $\|\cdot\|_\infty$ denotes the standard L^∞-norm. Recall the definition of the semimartingale $\mathcal{M}^{F,h}$ in Lemma 4.1. For each $t > 0$ and measurable set $\mathcal{B} \subset \Lambda$, it follows from (4.23) and (4.28) that

$$\mathbb{P}_\gamma\left((\mu_t, Q_t) \in \mathcal{B} \right)$$
$$= \mathbb{E}_\gamma\left(\exp\left\{ -t \, I_{F,h}(\mu_t, Q_t) - (S_{N_t+1} - t) h(X_{N_t}, \tau_{N_t+1}) \right\} \mathcal{M}_t^{F,h} \, 1_{\mathcal{B}}(\mu_t, Q_t) \right)$$
$$\le \exp\{\|\varphi\|_\infty\} \sup_{(\mu,Q) \in \mathcal{B}} \exp\{ -t \, I_{F,h}(\mu, Q) \} \, \mathbb{E}_\gamma\left(\mathcal{M}_t^{F,h} \, 1_{\mathcal{B}}(\mu_t, Q_t) \right).$$

Hence we have proved (4.27). It is easy to see that $\mathbb{P}_\gamma((\mu_t, Q_t) \in \mathcal{C}_\delta) = 1$ for any $t \ge 1/\delta$. Finally, taking $\mathcal{B} = \mathcal{O} \cap \mathcal{C}_\delta$ in (4.27), we obtain

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma((\mu_t, Q_t) \in \mathcal{O}) = \varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma((\mu_t, Q_t) \in \mathcal{O} \cap \mathcal{C}_\delta)$$
$$\le -\inf_{(\mu,Q) \in \mathcal{O} \cap \mathcal{C}_\delta} I_{F,h}(\mu, Q)$$
$$= -\inf_{(\mu,Q) \in \mathcal{O}} I_{F,h,\delta}(\mu, Q).$$

This completes the proof of this lemma. $\qquad\square$

Lemma 4.10. *Suppose that Assumption 2.6 holds. For any $(F,h) \in \Gamma_0$ and $\delta > 0$, $I_{F,h,\delta}$ is a convex and lower semicontinuous function on Λ, where $L^1_+(E)$ is endowed with the bounded weak* topology.*

Proof. Since a linear functional on $L^1_+(E)$ is continuous with respect to the bounded weak* topology if and only if it is continuous with respect to the weak* topology [36, Theorem 2.7.8], the weak* topology on $L^1(E)$ is the smallest topology such that the maps $Q \in L^1(E) \to \langle Q, f \rangle \in \mathbb{R}$ with $f \in C_0(E)$ being continuous. Let

$$f_x^1(y,z) = 1_x(y), \qquad f_x^2(y,z) = 1_x(z), \quad (y,z) \in E.$$

By Assumption 2.6, the graph (V, E) is locally finite. It is easy to check that $f_x^1, f_x^2 \in C_0(E)$ for any $x \in V$. Note that

$$\mathcal{C}_\delta = \bigcap_{x \in V} \left(\left\{ (\mu, Q) \in \Lambda : Q^+(x) - Q^-(x) \leq \delta \right\} \bigcap \left\{ (\mu, Q) \in \Lambda : Q^-(x) - Q^+(x) \leq \delta \right\} \right)$$

$$= \bigcap_{x \in V} \left(\left\{ (\mu, Q) \in \Lambda : \langle Q, f_x^1 - f_x^2 \rangle \leq \delta \right\} \bigcap \left\{ (\mu, Q) \in \Lambda : \langle Q, f_x^2 - f_x^1 \rangle \leq \delta \right\} \right).$$

This implies that $\mathcal{C}_\delta$ is a closed subset of Λ. Moreover, it is easy to see that $\mathcal{C}_\delta$ is convex. Thus we only need to prove that $I_{F,h}$ is a convex and lower semicontinuous function on Λ.

Since $I_{F,h}$ is a linear functional, it must be convex. Since $F \in C_c(E) \subset C_0(E)$, the map $Q \mapsto \langle Q, F \rangle$ is continuous. On the other hand, since $\varphi \in C_c(V \times (0, \infty))$, there exists a finite set $K \subset V$ such that $\varphi_x \equiv 0$ for $x \in K^c$ and $\varphi_x \in C_c(0, \infty)$ for $x \in K$. Then $h^{\varphi,c,M}$ is a bounded lower semicontinuous function. For any $\mu \in \mathcal{P}(V \times (0, \infty])$, let $\mu_n \in \mathcal{P}(V \times (0, \infty])$ be a sequence of probability measures such that μ_n weakly converges to μ. It then follows from the Portmanteau theorem [8, Corollary 8.2.5] that

$$\varliminf_{n \to \infty} \left\langle \mu_n, h^{\varphi,c,M} \right\rangle \geq \left\langle \mu, h^{\varphi,c,M} \right\rangle.$$

Since $\mathcal{P}(V \times (0, \infty])$ is a metric space with the topology of weak converge, it is clear that $\langle \mu, h \rangle$ is lower semicontinuous with respect to μ. This completes the proof of this lemma. $\qquad \square$

Lemma 4.11. *For any $(\mu, Q) \in \Lambda$ satisfying $Q^+ = Q^-$, we have*

$$I(\mu, Q) = \sup_{(F,h) \in \Gamma_0} I_{F,h}(\mu, Q).$$

Proof. The proof for $I(\mu, Q) \geq \sup_{(F,h) \in \Gamma_0} I_{F,h}(\mu, Q)$ is similar to [35, Proposition 2.1]. Here only prove the converse inequality

$$I(\mu, Q) \leq \sup_{(F,h) \in \Gamma_0} I_{F,h}(\mu, Q). \tag{4.29}$$

Case 1: $(\mu, Q) \notin \mathcal{D}$. Then there exists $x \in V$ such that

$$\mu\left(x, \frac{1}{\tau}\right) := \int_{(0, \infty]} \frac{1}{s} \mu(x, \mathrm{d}s) \neq Q_x.$$

Without loss of generality, we assume that $\mu(x, 1/\tau) < Q_x$. For any $C > 0$, $(y, z) \in E$, and $(y, s) \in V \times (0, \infty]$, set

$$F_C(y, z) = C 1_x(y), \qquad h_C(y, s) = -\frac{C}{s} 1_x(y).$$

It is easy to check that $g_{F_C, h_C} \equiv 0$ and

$$I_{F_C, h_C}(\mu, Q) = C\left[Q_x - \mu\left(x, \frac{1}{\tau}\right)\right].$$

Here $(F_C, h_c) \notin \Gamma_0$, but we still define $I_{F_C, h_c}(\mu, Q)$ as in (4.23). As $C \to \infty$, we have $I_{F_C, h_C}(\mu, Q) \to \infty$. Now for any $C > 0$, we construct a sequence $(F_n, h_n) \in \Gamma_0$ such that

$$\lim_{n \to \infty} I_{F_n, h_n}(\mu, Q) = I_{F_C, h_C}(\mu, Q). \tag{4.30}$$

For any $y \in V$, it is easy to see that there exists a non-negative function $g_y \in C_c(0, \infty)$ such that $\psi_y(\{t \in (0, \infty) : g_y(t) > 0\}) > 0$. Let

$$F_n(y, z) = C 1_x(y) 1_{K_n}(z), \quad h_n(y, s) = \frac{C}{s}(f_n(s) + c_n g_x(s)) 1_x(y), \tag{4.31}$$

where $f_n : (0, \infty] \to \mathbb{R}$ is a sequence of continuous functions satisfying $f_n(s) = 0$ if $s \in (0, 1/(n+1)) \cup (n+1, \infty]$ and $f_n(s) = -1$ if $s \in (1/n, n)$, c_n is a sequence of constants to be chosen later such that $g_{F_n, h_n} \equiv 0$, and K_n is a sequence of finite sets such that $K_n \uparrow V$ and

$$\left[1 + (e^C - 1) \sum_{y \in K_n} p_{xy}\right] \langle \psi_x, e^{C f_n} \rangle \le 1.$$

Since $\lim_{n \to \infty} \langle \psi_x, e^{C f_n} \rangle = -C$, such K_n must exist. It is easy to check that $g_{F_n, h_n}(y) = 0$ for any $y \ne x$ and

$$g_{F_n, h_n}(x) = \log\left(1 + (e^C - 1) \sum_{y \in K_n} p_{xy}\right) + \log\left\langle \psi_x, e^{C(f_n + c_n g_x)} \right\rangle. \tag{4.32}$$

Note that

$$\lim_{c_n \to \infty} g_{F_n, h_n}(x) = \infty, \quad \lim_{c_n \to 0} g_{F_n, h_n}(x) = \log\left(1 + (e^C - 1) \sum_{y \in K_n} p_{xy}\right) + \log\langle \psi_x, e^{C f_n} \rangle \le 0.$$

Since $g_{F_n, h_n}(x)$ is continuous with respect to c_n, by the intermediate value theorem, there exists a sequence of constants c_n such that $g_{F_n, h_n}(x) \equiv 0$, which implies that $(F_n, h_n) \in \Gamma_0$. Taking $n \to \infty$ on both sides of (4.32), it is easy to see that $c_n \to 0$. Thus in the sense of pointwise convergence, we have $(F_n, h_n) \to (F_C, h_C)$. By the dominated convergence theorem, we obtain (4.30). This further implies (4.29).

Case 2: $(\mu, Q) \in \mathcal{D}$. Then for any $x \in V$, we have

$$\mu\left(x, \frac{1}{\tau}\right) = Q_x < \infty.$$

Recall the definition of $\tilde{Q}_{xy}$ and $\tilde{\mu}_x$ in (3.6). Since V and $(0, \infty)$ are both Polish spaces, it follows from (3.4) that for any $x \in V$,

$$H\left(\tilde{Q}_{x, \cdot} \mid p_{x, \cdot}\right) = \sup_{\{F_x \in C_b(V) : \sum_{y \in V} p_{xy} e^{F_x(y)} = 1\}} \sum_{y \in V} \tilde{Q}_{xy} F_x(y),$$

$$H\left(\tilde{\mu}_x \mid \psi_x\right) = \sup_{\{\varphi_x \in C_b(0, \infty) : \langle \psi_x, e^{\varphi_x} \rangle = 1\}} \int_{(0, \infty)} \varphi_x(s) \tilde{\mu}_x(ds).$$

Recall the definition of I in (3.7). Then for any $F \in C_b(E)$, $\varphi \in C_b(V \times (0, \infty))$ satisfying $\sum_{y \in V} p_{xy} e^{F(x, y)} = 1$ and $\langle \psi_x, e^{\varphi_x} \rangle = 1$ for any $x \in V$, and $c : V \to [0, \infty)$ satisfying

$0 \leq c \leq \zeta$ and $c(x) < \zeta(x)$ for any $x \in V$ with $\zeta(x) > 0$, we only need to find a sequence $(F_n, h_n) \in \Gamma_0$ such that

$$\varliminf_{n \to \infty} I_{F_n, h_n}(\mu, Q) \geq \sum_{x \in V} \left[Q_x \sum_{y \in V} \tilde{Q}_{xy} F(x, y) + Q_x \int_{(0, \infty)} \varphi_x(s) \tilde{\mu}_x(\mathrm{d}s) + c_x \mu(x, \{\infty\}) \right]. \tag{4.33}$$

For any $(x, y) \in E$ and $(x, s) \in V \times (0, \infty]$, let

$$F_n(x, y) = (F(x, y) - a_x^n) 1_{K_n}(x) 1_{K_n}(y),$$

$$h_n(x, s) = \left[\frac{1}{s} (\varphi_x^n(s) - b_x^n g_x(s)) + c_x 1_{(n+1, \infty]}(s) \right] 1_{K_n}(x),$$

where $\varphi_x^n : (0, \infty] \to \mathbb{R}$ is a sequence of continuous functions satisfying $\varphi_x^n(s) = 0$ if $s \in (0, 1/(n+1)) \cup (n+1, \infty)$ and $\varphi_x^n(s) = \varphi_x(s)$ if $s \in (1/n, n)$, a_x^n and b_x^n are two sequences of constants satisfying $\sum_{y \in V} p_{xy} e^{F_n(x,y)} = 1$ as well as $\langle \psi_x, e^{sh_n(x,s)} \rangle = 1$, g_x is the function used in (4.31), and K_n is a sequence of finite sets such that $K_n \uparrow V$. For any $x \in V$ and sufficiently large n that may depend on x, straightforward computations show that

$$a_x^n = \log \frac{\sum_{y \in K_n} p_{xy} e^{F(x,y)}}{\sum_{y \in K_n} p_{xy}}, \qquad \int_{(0, n+1]} e^{\varphi_x^n(s) - b_x^n g_x(s)} \psi_x(\mathrm{d}s) + \int_{(n+1, \infty)} e^{c_x s} \psi_x(\mathrm{d}s) = 1.$$

Moreover, it is easy to see that such sequence of constants b_x^n exists, and $a_x^n \to 0$ and $b_x^n \to 0$ as $n \to \infty$. For all $x \in V$, we have

$$\varliminf_{n \to \infty} \left[\sum_{y \in V} Q(x, y) F_n(x, y) + \int_{(0, \infty]} h_n(x, s) \mu(x, \mathrm{d}s) \right]$$

$$\geq \sum_{y \in V} Q(x, y) F(x, y) + \int_{(0, \infty)} \frac{1}{s} \varphi_x(s) \mu(x, \mathrm{d}s) + c_x \mu(x, \{\infty\}).$$

Then (4.33) follows from Fatou's lemma. This completes the proof of this lemma. $\qquad \square$

We are now in a position to prove the upper bound of the LDP.

Proposition 4.12. *Suppose that Assumptions 2.3-2.6 are satisfied. Let $L_+^1(E)$ be endowed with the bounded weak* topology. Then I is a convex and lower semicontinuous function, and for each compact set $\mathcal{K} \subset \Lambda$, we have*

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma \left((\mu_t, Q_t) \in \mathcal{K} \right) \leq - \inf_{(\mu, Q) \in \mathcal{K}} I(\mu, Q). \tag{4.34}$$

Moreover, if Condition 3.2 is satisfied, then the above equation also holds for any closed set $\mathcal{K} \subset \Lambda$.

Proof. We first prove that I is convex and lower semicontinuous. Since we have proved $I_{F, h, \delta}$ is convex and lower semicontinuous in Lemma 4.10, we only need to verify that

$$I(\mu, Q) = \sup_{(F, h) \in \Gamma_0, \delta > 0} I_{F, h, \delta}(\mu, Q), \quad (\mu, Q) \in \Lambda.$$

Case 1: $Q^+ \neq Q^-$. Let δ be a constant satisfying $0 < \delta < \max_{x \in V} |Q^+(x) - Q^-(x)|$. Since $(\mu, Q) \in \Lambda \setminus \mathcal{C}_\delta$, we have

$$I(\mu, Q) = \infty = \sup_{(F, h) \in \Gamma_0, \delta > 0} I_{F, h, \delta}(\mu, Q).$$

Case 2: $Q^+ = Q^-$. By Lemma 4.11, we have

$$I(\mu, Q) = \sup_{(F,h) \in \Gamma_0} I_{F,h}(\mu, Q).$$

Recall the definition of C_δ in (4.24) and the definition of $I_{F,h,\delta}$ in (4.25). It is easy to check that for any $\delta > 0$, we have $(\mu, Q) \in C_\delta$ and $I_{F,h}(\mu, Q) = I_{F,h,\delta}(\mu, Q)$. Optimizing over $(F, h) \in \Gamma_0$ and $\delta > 0$, we obtain

$$I(\mu, Q) = \sup_{(F,h) \in \Gamma_0} I_{F,h}(\mu, Q) = \sup_{(F,h) \in \Gamma_0, \delta > 0} I_{F,h,\delta}(\mu, Q).$$

We next prove (4.34). Minimizing (4.26) over $(F, h) \in \Gamma_0$ and $\delta > 0$, we obtain

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma((\mu_t, Q_t) \in \mathcal{O}) \leq - \sup_{(F,h) \in \Gamma_0, \delta > 0} \inf_{(\mu,Q) \in \mathcal{O}} I_{F,h,\delta}(\mu, Q).$$

Note that $I_{F,h,\delta}$ is lower semicontinuous for any $(F, h) \in \Gamma_0$ and $\delta > 0$. For any compact set $\mathcal{K} \subset \Lambda$, it follows from the min-max lemma [32, Lemma 3.3 in Appendix 2.3] that

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma((\mu_t, Q_t) \in \mathcal{K}) \leq - \inf_{(\mu,Q) \in \mathcal{K}} \sup_{(F,h) \in \Gamma_0, \delta > 0} I_{F,h,\delta}(\mu, Q) = - \inf_{(\mu,Q) \in \mathcal{K}} I(\mu, Q).$$

Finally, under Condition 3.2, it follows from Proposition 4.7 that (μ_t, Q_t) is exponentially tight. Hence the upper bound in (4.34) also holds for any closed set $\mathcal{K} \subset \Lambda$. □

4.4 Lower bound

We next prove the lower bound of the LDP. Before giving the proof, we recall the following lemma, which can be found in [7, Lemma 5.2].

Lemma 4.13. *Let $\{P_t\}$ be a family of probability measures on a completely regular topological space $\mathcal{X}$ and let $J : \mathcal{X} \to [0, \infty]$ be a function. Assume that for each $x \in \mathcal{X}$, there exists a family of probability measures $\{\tilde{P}_t^x\}$ weakly convergent to δ_x such that*

$$\varlimsup_{t \to \infty} \frac{1}{t} H(\tilde{P}_t^x | P_t) \leq J(x). \tag{4.35}$$

Then the family $\{P_t\}$ satisfies the large deviation lower bound with rate function $\mathrm{sc}^- J$. Here $\mathrm{sc}^- J$ is the lower semicontinuous envelope of J, i.e.

$$(\mathrm{sc}^- J)(x) = \sup_{U \in \mathcal{N}_x} \inf_{y \in U} J(y),$$

where $\mathcal{N}_x$ denotes the collection of the open neighborhoods of x.

In order to prove the lower bound, we will apply Lemma 4.13 for completely regular topological space Λ. Recall the definitions of $\mathcal{D}$ and $\tilde{Q}$ in (3.5) and (3.6), respectively. Let $\mathcal{D}_2 \subset \mathcal{D}_1 \subset \mathcal{D}$ be defined by

$$\mathcal{D}_1 = \big\{ (\mu, Q) \in \mathcal{D} : I(\mu, Q) < \infty, \mu(x, (0, \infty)) > 0, \ \forall x \in V,$$

$$(\tilde{Q}_{xy})_{x,y \in V} \text{ defines an irreducible transition matrix on } V \big\},$$

$$\mathcal{D}_2 = \{ (\mu, Q) \in \mathcal{D}_1 : \mu(x, \{\infty\}) = 0, \ \forall x \in V \}.$$

Moreover, let $J : \Lambda \to [0, \infty]$ be the restriction of the rate function I to $\mathcal{D}_2$, i.e.

$$J(\mu, Q) = \begin{cases} I(\mu, Q), & \text{if } (\mu, Q) \in \mathcal{D}_2, \\ \infty, & \text{otherwise.} \end{cases} \tag{4.36}$$

In the following lemma, we will construct a family of probability measures $\{\tilde{P}_t^x\}$ on Λ and prove (4.35) with the upper bound given by (4.36). The proof is in the spirit of that given in [35, Proposition 5.1], but some details are supplemented.

Lemma 4.14. *Let $P_t = \mathbb{P}_\gamma \circ (\mu_t, Q_t)^{-1}$. For each $(\mu, Q) \in \Lambda$ there exists a family of probability measures $\{\tilde{P}_t^{(\mu,Q)}\}$ on Λ weakly convergent to $\delta_{(\mu,Q)}$ such that*

$$\varlimsup_{t \to \infty} \frac{1}{t} H\left(\tilde{P}_t^{(\mu,Q)} \big| P_t\right) \leq J(\mu, Q).$$

Proof. By the definition of J in (4.36), we may restrict the proof to $(\mu, Q) \in \mathcal{D}_2$. When $(\mu, Q) \in \mathcal{D}_2$, we have

$$Q_x = \mu\left(x, \frac{1}{\tau}\right) > 0, \qquad \mu(x, \{\infty\}) = 0, \quad x \in V,$$

and it is easy to check that

$$\sum_{x \in V} Q_x H\left(\tilde{Q}_{x,\cdot} \mid p_{x,\cdot}\right) < \infty, \qquad \sum_{x \in V} Q_x H\left(\tilde{\mu}_x \mid \psi_x\right) < \infty.$$

It then follows from [19, Theorem 2.1 and the remark after Theorem 2.1] that $\tilde{Q}_{x,\cdot} \ll p_{x,\cdot}$, $\tilde{\mu}_x \ll \psi_x$ and

$$\sum_{(x,y) \in E} Q(x,y) \left| \log\left(\frac{Q(x,y)}{Q_x p_{xy}}\right) \right| < \infty, \tag{4.37}$$

$$\sum_{x \in V} \int_{(0,\infty)} \frac{1}{s} \left| \log\left(\frac{\mu(x, \mathrm{d}s)}{Q_x s \psi_x(\mathrm{d}s)}\right) \right| \mu(x, \mathrm{d}s) < \infty. \tag{4.38}$$

Let $(F, h) \in \Gamma$ be defined by

$$F(x,y) = \log\left(\frac{Q(x,y)}{Q_x p_{xy}}\right), \quad h(x,s) = \frac{1}{s} \log\left(\frac{\mu(x, \mathrm{d}s)}{Q_x s \psi_x(\mathrm{d}s)}\right). \tag{4.39}$$

In this way, the transition probability kernel (P^F, Ψ^h) defined in (4.2) are given by

$$p_{xy}^F = \frac{1}{Q_x} Q(x,y) = \tilde{Q}_{xy}, \qquad \psi_x^h(\mathrm{d}s) = \frac{1}{Q_x s} \mu(x, \mathrm{d}s) = \tilde{\mu}_x(\mathrm{d}s).$$

Let $\mathbb{P}_\gamma^{F,h}$ be the probability measure under which (X, τ) is a Markov renewal process with transition kernel (P^F, Ψ^h) and initial distribution γ. Since $(\mu, Q) \in \mathcal{D}_2$, it is easy to see that $p^F = \tilde{Q}$ is irreducible. It then follows from (3.5) that

$$\sum_{x \in V} \nu_x^F p_{xy}^F = \frac{1}{\|Q\|} \sum_{x \in V} Q(x,y) = \frac{Q_y}{\|Q\|} = \nu_y^F,$$

where $\nu_x^F = Q_x/\|Q\|$ for any $x \in V$. This means that ν^F is the unique invariant distribution for P^F. Since $\mu(x, \{\infty\}) = 0$ for any $x \in V$, we have

$$\mathbb{E}_{\nu^F}^{F,h}(\tau_1) = \sum_{x \in V} \nu_x^F \int_{(0,\infty)} s \psi_x^h(\mathrm{d}s) = \sum_{x \in V} \frac{Q_x}{\|Q\|} \int_{(0,\infty)} s \frac{1}{Q_x s} \mu(x, \mathrm{d}s) = \frac{1}{\|Q\|} < \infty. \tag{4.40}$$

By the strong law of large numbers for semi-Markov processes, for any $f \in C_b(V \times (0, \infty])$ and $g \in C_b(E)$, we have

$$\lim_{t \to \infty} \langle \mu_t, f \rangle = \frac{1}{\mathbb{E}_{\nu^F}^{F,h}(\tau_1)} \sum_{x \in V} \nu_x^F \int_{(0,\infty)} s f(x,s) \psi_x^h(\mathrm{d}s) = \langle \mu, f \rangle, \quad \mathbb{P}_\gamma^{F,h}\text{-a.s.} \tag{4.41}$$

$$\lim_{t \to \infty} \langle Q_t, g \rangle = \frac{1}{\mathbb{E}_{\nu^F}^{F,h}(\tau_1)} \sum_{(x,y) \in E} \nu_x^F p_{xy}^F g(x,y) = \langle Q, g \rangle, \quad \mathbb{P}_\gamma^{F,h}\text{-a.s.} \tag{4.42}$$

We now construct the family of probability measures $\{\tilde{P}_t^{(\mu,Q)}\}$. For any $\epsilon > 0$ and $t \geq 0$, let $T_t = \lfloor \|Q\|(1 + \epsilon)t \rfloor$ and let $\mathbb{P}_{\gamma,t,\epsilon}$ be the probability measure under which the law of the process $(X, \tau) = \{(X_k)_{k \geq 0}, (\tau_k)_{k \geq 1}\}$ satisfies the following requirements:

(a) $\{(X_k)_{0 \le k \le T_t}, (\tau_k)_{1 \le k \le T_t}\}$ is a Markov renewal process with transition kernel (P^F, Ψ^h) and initial distribution γ.

(b) Conditioned on X_{T_t}, the processes $\{(X_k)_{0 \le k \le T_t}, (\tau_k)_{1 \le k \le T_t}\}$ and $\{(X_k)_{k \ge T_t}, (\tau_k)_{k \ge T_t+1}\}$ are independent of each other. Moreover, $\{(X_k)_{k \ge T_t}, (\tau_k)_{k \ge T_t+1}\}$ is a Markov renewal process with transition kernel (P, Ψ) and initial distribution $\delta_{X_{T_t}}$.

Intuitively, under $\mathbb{P}_{\gamma,t,\epsilon}$, the process (X, τ) has the transition kernel (P^F, Ψ^h) before time T_t and has the transition kernel (P, Ψ) after time T_t. Set $\tilde{P}_{t,\epsilon} = \mathbb{P}_{\gamma,t,\epsilon} \circ (\mu_t, Q_t)^{-1}$ and let $\epsilon(t) \downarrow 0$ to be chosen later such that $\tilde{P}_t^{(\mu,Q)} := \tilde{P}_{t,\epsilon(t)}$ weakly converges to $\delta_{(\mu,Q)}$ as $t \to \infty$. In other words, for any $G \in C_b(\Lambda)$, we have

$$\lim_{t \to \infty} \int_\Lambda G \mathrm{d}\tilde{P}_{t,\epsilon(t)} = G(\mu, Q). \tag{4.43}$$

For any $\epsilon > 0$, it follows from (4.40) and the strong law of large numbers for Markov renewal processes [34, Theorem 3.13] that

$$\lim_{t \to \infty} \frac{S_{T_t}}{t} = \lim_{t \to \infty} \frac{S_{T_t}}{T_t} \frac{T_t}{t} = \mathbb{E}_{\nu^F}^{\mu,Q}(\tau_1) \|Q\| (1+\epsilon) = 1 + \epsilon, \quad \mathbb{P}_\gamma^{F,h}\text{-a.s.}$$

For any $t > 0$, let

$$D_{t,\epsilon} = \{S_{T_t} > t\} = \{N_t + 1 \le T_t\} \in \mathcal{F}_{T_t},$$

where $\mathcal{F}_{T_t} = \sigma((X_k, \tau_k)_{0 \le k \le T_t})$. It is easy to see that $\mathbb{P}_\gamma^{F,h}|_{\mathcal{F}_{T_t}} = \mathbb{P}_{\gamma,t,\epsilon}|_{\mathcal{F}_{T_t}}$. Then we have

$$\lim_{t \to \infty} \mathbb{P}_{\gamma,t,\epsilon}(D_{t,\epsilon}) = \lim_{t \to \infty} \mathbb{P}_{\gamma,t,\epsilon}(S_{T_t} > t) = \lim_{t \to \infty} \mathbb{P}_\gamma^{F,h}\left(\frac{S_{T_t}}{t} > 1\right) = 1. \tag{4.44}$$

Before determining $\epsilon(t)$, we give an estimation of the relative entropy. We observe that for any $\epsilon > 0$,

$$\begin{aligned}
H\left(\tilde{P}_{t,\epsilon}\big|P_t\right) &\le H\left(\mathbb{P}_{\gamma,t,\epsilon}\big|\mathbb{P}_\gamma\right) = \mathbb{E}_\gamma^{F,h}\left[\log \frac{\mathrm{d}\mathbb{P}_\gamma^{F,h}|_{\mathcal{F}_{T_t}}}{\mathrm{d}\mathbb{P}_\gamma|_{\mathcal{F}_{T_t}}}\right] \\
&= \mathbb{E}_\gamma^{F,h}\left[\sum_{i=1}^{T_t}\left(F(X_{i-1}, X_i) + \tau_i h(X_{i-1}, \tau_i)\right)\right].
\end{aligned} \tag{4.45}$$

Indeed, the first inequality follows from the variational characterization of the relative entropy [19, Section 2] and the last equality is a straightforward computation of the Radon-Nikodym density (similarly to (4.3)).

Combining (4.37), (4.38), and (4.39), we have $\langle Q, |F|\rangle < \infty$ and $\langle \mu, |h|\rangle < \infty$. Note that $(X_k, X_{k+1}, \tau_{k+1})_{k \ge 0}$ is a Markov process. By the ergodic theorem of Markov processes, we have

$$\begin{aligned}
\lim_{t \to \infty} \frac{1}{t}\mathbb{E}_\gamma^{F,h}\left[\sum_{i=1}^{T_t}\left(F(X_{i-1}, X_i) + \tau_i h(X_{i-1}, \tau_i)\right)\right] &= (1+\epsilon)\|Q\|\frac{1}{\|Q\|}\left[\langle Q, F\rangle + \langle \mu, h\rangle\right] \\
&\le (1+\epsilon)I(\mu, Q).
\end{aligned} \tag{4.46}$$

We next construct $\epsilon(t) \downarrow 0$. Let $\epsilon(t) = 1/n$ for any $t_{n-1} < t \le t_n$ be a step function, where t_n is an increasing sequence such that

$$\mathbb{P}_{\gamma,t,1/n}\left(D_{t,1/n}\right) \ge 1 - \frac{1}{n} \quad \text{and} \quad \frac{1}{t}H\left(\tilde{P}_{t,1/n}\big|P_t\right) \le \left(1 + \frac{1}{n-1}\right)I(\mu, Q), \quad t > t_{n-1}.$$

It follows from (4.44), (4.45), and (4.46) that such $t_n \uparrow \infty$ exist. Then we have

$$\lim_{t \to \infty} \mathbb{P}_{\gamma,t,\epsilon(t)}\left(D_{t,\epsilon(t)}\right) = 1, \qquad \overline{\lim_{t \to \infty}} \frac{1}{t}H\left(\tilde{P}_{t,\epsilon(t)}\big|P_t\right) \le I(\mu, Q). \tag{4.47}$$

Finally, we prove (4.43). Note that

$$
\int_\Lambda G \mathrm{d}\tilde{P}_{t,\epsilon(t)} = \int_\Omega G(\mu_t, Q_t) \mathrm{d}\mathbb{P}_{\gamma,t,\epsilon(t)}
$$

$$
= \int_\Omega G(\mu_t, Q_t) 1_{D_{t,\epsilon(t)}} \mathrm{d}\mathbb{P}_{\gamma,t,\epsilon(t)} + \int_\Omega G(\mu_t, Q_t) 1_{D^c_{t,\epsilon(t)}} \mathrm{d}\mathbb{P}_{\gamma,t,\epsilon(t)}.
$$

It thus follows from the first equality of (4.47) that

$$
\left| \int_\Omega G(\mu_t, Q_t) 1_{D^c_{t,\delta(t)}} \mathrm{d}\mathbb{P}_{\gamma,t,\delta(t)} \right| \leq \|G\|_\infty \mathbb{P}_{\gamma,t,\epsilon(t)} \left(D^c_{t,\epsilon(t)} \right) \to 0, \qquad \text{as } t \to \infty.
$$

On the other hand, since $D_{t,\epsilon(t)} = \{N_t + 1 \leq T_t\}$, we have $G(\mu_t, Q_t) 1_{D_{t,\epsilon(t)}} \in \mathcal{F}_{T_t}$. Note that $\mathcal{P}(V \times (0, \infty])$ is endowed with the weak convergence topology and $L^1_+(E)$ is endowed with the bounded weak* topology. It is a direct consequence of (4.41) and (4.42) that

$$
\lim_{t \to \infty} G(\mu_t, Q_t) = G(\mu, Q), \quad \mathbb{P}^{F,h}_\gamma\text{-a.s.}
$$

Moreover, it follows from the first equality of (4.47) that

$$
\lim_{t \to \infty} \mathbb{P}^{F,h}_\gamma \left(D_{t,\epsilon(t)} \right) = \lim_{t \to \infty} \mathbb{P}_{\gamma,t,\epsilon(t)} \left(D_{t,\epsilon(t)} \right) = 1.
$$

Then we obtain

$$
\int_\Omega G(\mu_t, Q_t) 1_{D_{t,\epsilon(t)}} \mathrm{d}\mathbb{P}_{\gamma,t,\epsilon(t)} = \int_\Omega G(\mu_t, Q_t) 1_{D_{t,\epsilon(t)}} \mathrm{d}\mathbb{P}^{F,h}_\gamma
$$
$$
\to \int_\Omega G(\mu, Q) \mathrm{d}\mathbb{P}^{F,h}_\gamma = G(\mu, Q), \qquad \text{as } t \to \infty,
$$
(4.48)

where the convergence in (4.48) follows from the dominated convergence theorem. $\square$

To prove the lower bound of the LDP, we only need to prove that the rate function I coincides with $\mathrm{sc}^- J$. Before giving the desired results, we need the following two lemmas.

Lemma 4.15. *Suppose that Assumption 2.3 is satisfied. Then for any $\mu = (\mu_x)_{x \in V} \in \mathcal{P}(V)$ satisfying $\mu_x > 0$ for any $x \in V$, there exist a constant $C > 0$ and an irreducible, positive recurrent Markov chain $(\hat{X}_k)_{k \geq 0}$ with state space V, transition probability matrix $\hat{P} = (\hat{p}_{xy})_{x,y \in V}$, and invariant distribution $\hat{\nu} = (\hat{\nu}_x)_{x \in V}$ such that $\hat{E} \subset E$ and $\hat{\nu}_x \leq C\mu_x$ for any $x \in V$, where $\hat{E} = \{(x, y) \in V \times V : \hat{p}_{xy} > 0\}$.*

Proof. We first construct a sequence of finite subsets $V_n \uparrow V$ by induction. Without loss of generality, we assume that $V = \{z_i\}_{i \geq 1}$. Since Assumption 2.3 holds, it is clear that (V, E) is connected. Thus, there exists a self-avoiding cycle $(x_1(:= z_1), x_2 \cdots, x_{k_0})$ of elements of V such that $(x_i, x_{i+1}) \in E$ when $i = 1, \cdots, k_0$ and the sum in the indices is modulo k_0 (cycle $(x_1, \cdots, x_{k_0})$ is called self-avoiding if $x_i \neq x_j$ for any $1 \leq i \neq j \leq k_0$). Set

$$
V_0 = \{x_1, x_2, \cdots, x_{k_0}\}.
$$

Suppose that we have constructed V_n and $z_1, \cdots, z_q \in V_n$, $z_{q+1} \notin V_n$. Since (V, E) is connected, it is easy to see that there exist $x_r \in V_n$ for some $1 \leq r \leq k_n = |V_n|$ and a sequence of distinct states $w_1, \cdots, w_{m_1 - 1} \in V \setminus V_n$ such that $(w_i, w_{i+1}) \in E$ when $i = 0, \cdots, m_1 - 1$, where $w_0 := x_r$ and $w_{m_1} := z_{q+1}$. Similarly, there exist $x_l \in V_n$ and a sequence of distinct states $w_{m_1+1}, \cdots, w_{m_2} \in V \setminus V_n$ such that $(w_i, w_{i+1}) \in E$ when $i = m_1, \cdots, m_2$, where $w_{m_2+1} := x_l$.

If $\{w_1, \cdots, w_{m_1}\} \cap \{w_{m_1+1}, \cdots, w_{m_2}\} = \emptyset$, let $m = m_2$ and $y_i = w_i$ for $i = 1, \cdots, m$.

If $\{w_1, \cdots, w_{m_1}\} \cap \{w_{m_1+1}, \cdots, w_{m_2}\} \neq \emptyset$, let κ_1 be the minimum integer such that $w_{\kappa_1} \in \{w_{m_1+1}, \cdots, w_{m_2}\}$. Then there exists $m_1 + 1 \leq \kappa_2 \leq m_2$ such that $w_{\kappa_1} = w_{\kappa_2}$. Moreover, let $m = \kappa_1 + m_2 - \kappa_2$ and $y_1 = w_1, \cdots, y_{\kappa_1} = w_{\kappa_1}, y_{\kappa_1+1} = w_{\kappa_2+1}, \cdots, y_m = w_{m_2}$, it is clear that $y_1, \cdots, y_m$ is a sequence of distinct states of $V \setminus V_n$.

Set $k_{n+1} = k_n + m$ and

$$V_{n+1} = \{x_1, \cdots x_{k_n}, x_{k_n+1}(:= y_1), \cdots, x_{k_{n+1}}(:= y_m)\}.$$

Then we have constructed a sequence of subsets $\{V_n\}_{n \geq 0}$ such that $|V_n| < \infty$ and $V_n \subset V_{n+1}$. Moreover, following from the above constructions, it is easy to see that for any $z_q \in V$, there exists V_n such that $z_q \in V_n$. This implies that $\cup_n V_n = V$.

We next construct a sequence of matrices $P^n = (p_{xy}^n)_{x,y \in V}$ base on the previous construction of $\{V_n\}_{n \geq 0}$. Let

$$p_{x_i, x_{i+1}}^0 = 1, \quad 1 \leq i \leq k_0 - 1, \qquad p_{x_{k_0}, x_1}^0 = 1, \qquad p_{xy}^0 = 0, \quad \text{otherwise.}$$

Suppose that we have constructed $P^n = (p_{xy}^n)_{x,y \in V}$. Let

$$p_{x_r, x_{k_{n+1}}}^{n+1} = \alpha_{n+1}, \qquad p_{x_i, x_{i+1}}^{n+1} = 1, \quad k_n + 1 \leq i \leq k_{n+1} - 1, \qquad p_{x_{k_{n+1}}, x_l}^{n+1} = 1,$$
$$p_{x_r, y}^{n+1} = (1 - \alpha_{n+1})p_{x_r, y}^n, \quad y \in V_n, \qquad p_{xy}^{n+1} = p_{xy}^n, \quad \text{otherwise,} \tag{4.49}$$

where $x_r, x_l \in V_n$ are defined the same as in the above construction of $\{V_n\}_{n \geq 0}$ and α_{n+1} is a positive constant remaining to decide. Then the sequence of matrices P^n is constructed by induction. It is easy to see that P^n can be considered as a transition probability matrix with state space V_n, which corresponds to an irreducible Markov chain. We let $\nu^n \in \mathcal{P}(V)$ be the invariant distribution for P^n. Now we decide $\{\alpha_n\}_{n \geq 0}$ and C such that $\nu_x^n < C\mu_x$ for any $n \in \mathbb{N}$ and $x \in V$ by induction. Let $C = 1 + (\max_{x \in V_0}(1/\mu_x))/k_0$ and $\alpha_0 = 0$. It is clear that

$$\nu_x^0 = \frac{1}{k_0} < C\mu_x, \quad x \in V_0, \qquad \nu_x^0 = 0 < C\mu_x, \quad x \notin V_0.$$

Suppose that we have decided $(\alpha_k)_{0 \leq k \leq n}$ such that $\nu_x^n < C\mu_x$ for any $x \in V$. Note that ν^{n+1} is continuous with respect to α_{n+1}. In other words,

$$\lim_{\alpha_{n+1} \downarrow 0} \nu_x^{n+1} = \nu_x^n, \quad x \in V.$$

Since V is countable, the strong convergence of ν^{n+1} to ν^n in $\mathcal{P}(V)$ is equivalent to the pointwise convergence of ν_x^{n+1} to ν_x^n for any $x \in V$. This implies $\lim_{\alpha_{n+1} \downarrow 0} \|\nu^{n+1} - \nu^n\| = 0$. Hence, there exists $0 < \alpha_{n+1} \leq 1/2^{n+1}$ such that

$$\nu_x^{n+1} < C\mu_x, \quad x \in V_{n+1}, \qquad \|\nu^{n+1} - \nu^n\| \leq \frac{1}{2^n}. \tag{4.50}$$

Since $\nu_x^{n+1} = 0$ for any $x \notin V_{n+1}$, it is clear that $\nu_x^{n+1} < C\mu_x$ for any $x \in V$.

Finally, we construct $\hat{P}$. For any $x \in V$, since $V_n \uparrow V$, there exists N such that $x \in V_N$. It follows from (4.49) that

$$\sum_{y \in V} |p_{xy}^{n+1} - p_{xy}^n| \leq 2\alpha_{n+1} \leq \frac{1}{2^n}, \qquad n \geq N.$$

This shows that $\hat{P}$ can be defined as $\hat{p}_{xy} = \lim_{n \to \infty} p_{xy}^n \geq 0$. By Fatou's lemma, we have

$$\sum_{y \in V} |p_{xy}^n - \hat{p}_{xy}| \leq \varliminf_{m \to \infty} \sum_{y \in V} |p_{xy}^n - p_{xy}^m| \leq \varliminf_{m \to \infty} \sum_{k=n}^{m-1} \frac{1}{2^k} = \frac{1}{2^{n-1}}, \quad n \geq N, \tag{4.51}$$

which implies $\sum_{y \in V} \hat{p}_{xy} = 1$ for any $x \in V$. In other words, $\hat{P}$ is a transition probability matrix. Note that if there exists n such that $p_{xy}^n > 0$, then

$$\hat{p}_{xy} = \lim_{k \to \infty} p_{xy}^k \geq \left[\prod_{k=n+1}^{\infty} (1 - \alpha_k) \right] p_{xy}^n \geq \left(1 - \frac{1}{2^n} \right) p_{xy}^n > 0.$$

Since P^n corresponds to an irreducible Markov chain with state space V_n and $V_n \uparrow V$, it is easy to see that $\hat{P}$ is irreducible.

On the other hand, it follows from (4.50) that we can set $\hat{\nu} = \lim_{n \to \infty} \nu^n$. Moreover, it is clear that $\hat{\nu} \in \mathcal{P}(V)$ and $\hat{\nu}_x \leq C\mu_x$ for any $x \in V$. It follows from (4.51) that for any $y \in V$ and $N \in \mathbb{N}$,

$$\sum_{x \in V} \left| \nu_x^n p_{xy}^n - \hat{\nu}_x \hat{p}_{xy} \right| \leq \sum_{x \in V} |\nu_x^n - \hat{\nu}_x| p_{xy}^n + \sum_{x \in V_N} \hat{\nu}_x \left| p_{xy}^n - \hat{p}_{xy} \right| + \sum_{x \in V \setminus V_N} \hat{\nu}_x \left| p_{xy}^n - \hat{p}_{xy} \right|$$

$$\leq \| \nu^n - \hat{\nu} \| + \frac{1}{2^{n-1}} + 1 - \sum_{x \in V_N} \hat{\nu}_x, \quad n \geq N.$$

Taking $n \to \infty$ and then taking $N \to \infty$ on both two sides of the above inequality, we have

$$\hat{\nu}_y = \lim_{n \to \infty} \nu_y^n = \lim_{n \to \infty} \sum_{x \in V} \nu_x^n p_{xy}^n = \sum_{x \in V} \hat{\nu}_x \hat{p}_{xy}, \quad y \in V.$$

Obviously, we have $\hat{E} \subset E$. Then we finish the proof of this lemma. $\quad \square$

Lemma 4.16. *Suppose that Assumptions 2.3-2.6 are satisfied. Then $\mathcal{D}_1 \neq \emptyset$.*

Proof. We first construct a transition kernel $(\hat{P} = (\hat{p}_{xy})_{x,y \in V}, \hat{\Psi} = (\hat{\psi}_x)_{x \in V})$ satisfying the following four requirements:

(a) $\hat{P}$ is an irreducible transition probability matrix with an unique invariant probability measure $\hat{\nu}$.

(b) $\hat{\psi}_x \ll \psi_x$ and $\hat{p}_{x,\cdot} \ll p_{x,\cdot}$ for any $x \in V$.

(c) $\sum_{x \in V} \hat{\nu}_x \int_{(0,\infty)} s \hat{\psi}_x(\mathrm{d}s) < \infty$.

(d) $\sum_{x \in V} \hat{\nu}_x [H(\hat{p}_{x,\cdot} | p_{x,\cdot}) + H(\hat{\psi}_x | \psi_x)] < \infty$.

Without loss of generality, we assume that $V = \mathbb{N}$ is the set of nonnegative integers. Let $m_x = \sup\{c \geq 0 : \psi_x((0,c)) = 0\}$, $A_x = [m_x, m_x + 1]$, and $\hat{\psi}_x(\mathrm{d}s) = \psi_x(\mathrm{d}s | A_x) := 1_{A_x}(s)\psi_x(\mathrm{d}s)/\psi_x(A_x)$. Then A_x is a bounded Borel subset of $(0, \infty)$ satisfying $\psi_x(A_x) > 0$ and

$$\int_{(0,\infty)} s \hat{\psi}_x(\mathrm{d}s) \leq m_x + 1, \qquad H(\hat{\psi}_x | \psi_x) = -\log \psi_x(A_x), \quad x \in \mathbb{N}. \tag{4.52}$$

For any $x \in \mathbb{N}$, let $C_x = -\sum_{y \in \mathbb{N}} \log p_{xy}$. By Assumption 2.6, it is clear that $0 \leq C_x < \infty$. Let $\mu \in \mathcal{P}(\mathbb{N})$ be defined by

$$\mu_x = \frac{1}{M} \frac{2^{-x}}{\max\{m_x + 1, -\log \psi_x(A_x), C_x\}}, \quad x \in \mathbb{N}, \tag{4.53}$$

where

$$M = \sum_{y \in \mathbb{N}} \frac{2^{-y}}{\max\{m_y + 1, -\log \psi_y(A_y), C_y\}}$$

is a normalization constant. By Lemma 4.15, there exist a constant $C > 0$ and a Markov chain $\hat{X}$ with transition probability matrix $\hat{P}$ and invariant distribution $\hat{\nu}$ such that $\hat{E} \subset E$ (i.e. $\hat{p}_{x,\cdot} \ll p_{x,\cdot}$ for any $x \in \mathbb{N}$) and $\hat{\nu}_x \leq C\mu_x$ for any $x \in \mathbb{N}$. Note that

$$H(\hat{p}_{x,\cdot} | p_{x,\cdot}) = \sum_{y \in \mathbb{N}} \hat{p}_{xy} \log \frac{\hat{p}_{xy}}{p_{xy}} \leq \sum_{y:(x,y) \in E} \log \frac{1}{p_{xy}} \leq C_x. \tag{4.54}$$

Combining (4.52), (4.53), and (4.54), it is clear that $(\hat{P}, \hat{\Psi})$ satisfies items (a)-(d) of the above requirements.

Finally, let $Z = \sum_{x \in V} \hat{\nu}_x \int_{(0,\infty)} s \hat{\psi}_x(\mathrm{d}s)$ and

$$\mu^0(x, \mathrm{d}s) = \frac{\hat{\nu}_x \, s \, \hat{\psi}_x(\mathrm{d}s)}{Z}, \qquad Q^0(x, y) = \frac{\hat{\nu}_x \, \hat{p}_{xy}}{Z}.$$

It is easy to check that $(\mu^0, Q^0) \in \mathcal{D}_1$. This completes the proof of this lemma. $\qquad\square$

We are now in a position to finish the proof of the lower bound of the LDP.

Proposition 4.17. *Suppose that Assumptions 2.3-2.6 are satisfied. Let $L^1_+(E)$ be endowed with the bounded weak* topology. Then under $\mathbb{P}_\gamma$, the law of $(\mu_t, Q_t) \in \Lambda$ satisfies an LDP lower bound with convex rate function I. Moreover, if Condition 3.2 is satisfied, then I is a good rate function.*

Proof. Let J be the functional defined in (4.36). By Lemmas 4.13 and 4.14, we only need to prove $I = \mathrm{sc}^- J$. By Proposition 4.12, I is convex and lower semicontinuous on Λ. It is then easy to see that $\mathrm{sc}^- J \geq I$.

We next prove the converse inequality. In fact, we only need to prove that for any $(\mu, Q) \in \Lambda$ with $I(\mu, Q) < \infty$, there exists a sequence $(\mu_n, Q_n)_{n \geq 0}$ in $\mathcal{D}_2$ such that $(\mu_n, Q_n) \to (\mu, Q)$ in Λ and

$$\varlimsup_{n \to \infty} I(\mu_n, Q_n) \leq I(\mu, Q). \tag{4.55}$$

Here, we only prove that there exists a sequence $(\mu_n, Q_n)_{n \geq 0}$ in $\mathcal{D}_1$ such that $(\mu_n, Q_n) \to (\mu, Q)$ in Λ and (4.55) holds. The rest of the proof is similar to [35, Lemma 2.5].

By Lemma 4.16, there exists $(\mu^0, Q^0) \in \mathcal{D}_1$. For any $(\mu, Q) \in \mathcal{D}$ with $I(\mu, Q) < \infty$, let

$$\mu_n = \left(1 - \frac{1}{n}\right) \mu + \frac{1}{n} \mu^0, \qquad Q_n = \left(1 - \frac{1}{n}\right) Q + \frac{1}{n} Q^0, \qquad n \geq 1.$$

Obviously, $(\mu_n, Q_n) \to (\mu, Q)$ in Λ in the sense of the strong topology. Since I is convex on Λ, we have

$$I(\mu_n, Q_n) \leq \left(1 - \frac{1}{n}\right) I(\mu, Q) + \frac{1}{n} I(\mu^0, Q^0).$$

Then it is easy to see that $(\mu_n, Q_n) \in \mathcal{D}_1$ and (4.55) holds. This completes the proof of the lower bound.

Finally, under Condition 3.2, it follows from Proposition 4.7 that (μ_t, Q_t) is exponentially tight. This fact, together with the lower bound of the LDP implies that I is a good rate function [15, Lemma 1.2.18]. $\qquad\square$

Remark 4.18. In fact, we can further prove that for any $(\mu, Q) \in \mathcal{D}$ with $I(\mu, Q) < \infty$, there exists a sequence $(\mu_n, Q)_{n \geq 0}$ in Λ such that $\mu_n(x, \cdot) = \mu(x, \cdot)$ whenever $Q_x = 0$, and $\mu_n(x, \{\infty\}) = 0$ whenever $Q_x > 0$, $(\mu_n, Q) \to (\mu, Q)$ in Λ, and

$$\varlimsup_{n \to \infty} I(\mu_n, Q) \leq I(\mu, Q).$$

The proof is similar to that given in [35, Lemma 2.5].

5 Proof of Theorem 3.7

Next we will prove the joint LDP for the empirical measure and empirical flow when $L^1_+(E)$ is endowed with the strong topology. Note that the bounded weak* topology is weaker than the strong topology [36, Theorem 2.7.2]. In other words, any open (closed) subset of $L^1_+(E)$ under the bounded weak* topology is also open (closed) under

the strong topology. Since we have established the joint LDP when $L_+^1(E)$ is endowed with the bounded weak* topology in Section 4, we only need to prove the exponential tightness of the empirical flow when $L_+^1(E)$ is endowed with the strong topology [15, Corollary 4.2.6]. Before proving the exponential tightness, we introduce some notation.

Recall the definition of exit-current and entrance-current in (2.3). For any $t > 0$, we define the associate empirical currents $Q_t^+ : \Omega \to [0, \infty]^V$ and $Q_t^- : \Omega \to [0, \infty]^V$ as

$$Q_t^+(x) = \sum_{y \in V} Q_t(x, y) = \frac{1}{t} \sum_{k=1}^{N_t+1} 1_{(X_{k-1}=x)}, \quad Q_t^-(x) = \sum_{y \in V} Q_t(y, x) = \frac{1}{t} \sum_{k=1}^{N_t+1} 1_{(X_k=x)}.$$
(5.1)

It is clear that $Q_t^+, Q_t^- \in L_+^1(V)$, $\mathbb{P}_\gamma$-a.s. For any $J \in L_+^1(V)$ and $f : V \to \mathbb{R}$, we set $\langle J, f \rangle = \sum_{x \in V} J(x) f(x)$.

The exponential tightness of the empirical currents is stated in the following proposition. Here we also consider the case where the semi-Markov process starts from the general initial distribution γ (see Remark 3.8).

Proposition 5.1. *Assume Conditions 3.2 and 3.4 to hold. Then there exists a sequence $\{\mathcal{K}_\ell\}$ of compact sets in $L_+^1(V)$ such that for any $\ell \in \mathbb{N}$,*

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma(Q_t^+ \notin \mathcal{K}_\ell) \leq -\ell,$$
(5.2)

where $L_+^1(V)$ is endowed with the strong topology. In particular, the empirical exit-current is exponentially tight.

Proof. We first construct the compact sets in $L_+^1(V)$ under the strong topology. Let $W_m \uparrow V$ be an invading sequence of finite subsets of V, for any positive integer ℓ, we let

$$\mathcal{K}_\ell = \left\{ J \in L_+^1(V) \ : \ \|J\| \leq A_\ell \,, \ \langle J, 1_{W_m^c} \rangle \leq \frac{1}{m}, \ \ \forall m \geq \ell \right\}.$$

where A_ℓ is defined as in Proposition 4.7. Similarly to the proof in [6, Theorem 5.2], it is easy to prove that $\mathcal{K}_\ell \subset L_+^1(V)$ is compact under the strong topology.

Now we prove (5.2). It is easy to see that

$$\mathbb{P}_\gamma(Q_t^+ \notin \mathcal{K}_\ell) \leq \mathbb{P}_\gamma(\|Q_t^+\| \geq A_\ell) + \sum_{m \geq \ell} \mathbb{P}_\gamma\left(\langle Q_t^+, 1_{W_m^c} \rangle > \frac{1}{m} \right).$$
(5.3)

Note that $\|Q_t\| = \sum_{x,y \in V} Q_t(x, y) = \sum_{x \in V} Q_t^+(x) = \|Q_t^+\| = (N_t + 1)/t$. It thus follows from (4.18) that

$$\varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma(\|Q_t^+\| \geq A_\ell) = \varlimsup_{t \to \infty} \frac{1}{t} \log \mathbb{P}_\gamma(\|Q_t\| \geq A_\ell) \leq -\ell.$$

On the other hand, let the function $\hat{u}$, the sequence of functions u_n, and the constants c, C_γ be as in Condition 3.4 and item (c*) in Remark 3.8. Taking $A = V$ in Lemma 4.2, we obtain the local martingale

$$\mathcal{M}_t^{u_n, V} = \frac{u_n(X_{N_t+1})}{u_n(X_0)} \exp\left\{ t \left\langle \hat{\mu}_t, h^{u_n, V} \right\rangle \right\} = \frac{u_n(X_{N_t+1})}{u_n(X_0)} \exp\left\{ t \left\langle Q_t^+, \log \frac{u_n}{Pu_n} \right\rangle \right\}.$$
(5.4)

Note that u_n/Pu_n converges pointwise to $\hat{u}$. By Fatou's lemma, we have

$$\mathbb{E}_\gamma\left(\exp\left\{t\left\langle Q_t^+, \log\hat{u}\right\rangle\right\}\right) \leq \sum_{x\in V}\gamma(x)\varliminf_{n\to\infty}\mathbb{E}_x\left(\exp\left\{t\left\langle Q_t^+, \log\frac{u_n}{Pu_n}\right\rangle\right\}\right)$$

$$= \sum_{x\in V}\gamma(x)\varliminf_{n\to\infty}\mathbb{E}_x\left(\frac{u_n(X_0)}{u_n(X_{N_t+1})}\mathcal{M}_t^{u_n,V}\right)$$

$$\leq \sum_{x\in V}\gamma(x)\frac{u_n(x)}{c} \leq \frac{C_\gamma}{c},$$

where the second inequality follows from item (b) in Condition 3.4 and the last inequality follows from item (c*) in Remark 3.8.

Let $\{a_m\}_{m\geq 0}$ be a sequence of constants with $a_m\uparrow\infty$ to be chosen later and let $W_m = \{x\in V: \log\hat{u}(x)\leq a_m\}$ be an invading sequence of V. In view of item (e) in Condition 3.4, W_m are finite sets. Let

$$C = 1\vee\left(-\inf_{x\in V}\log\hat{u}(x)\right) < \infty.$$

It is easy to see that $\log\hat{u}\geq a_m 1_{W_m^c} - C$. By the exponential Chebyshev inequality, we obtain

$$\mathbb{P}_\gamma\left(\langle Q_t^+, 1_{W_m^c}\rangle > \frac{1}{m}\right) \leq \mathbb{P}_\gamma\left(\langle Q_t^+, \log\hat{u}\rangle + C\|Q_t^+\| > \frac{a_m}{m}\right)$$

$$\leq \mathbb{P}_\gamma\left(\langle Q_t^+, \log\hat{u}\rangle > \frac{a_m}{2m}\right) + \mathbb{P}_\gamma\left(C\|Q_t\| > \frac{a_m}{2m}\right)$$

$$\leq \frac{C_\gamma}{c}\exp\left\{-t\frac{a_m}{2m}\right\} + \mathbb{P}_\gamma\left(\|Q_t\| > \frac{a_m}{2Cm}\right).$$

By choosing $a_m = 2m^2 + 2CmA_m$, the proof is now easily concluded from (5.3). $\quad\square$

Corollary 5.2. *Assume Conditions 3.2 and 3.4 to hold. Then the empirical entrance-current Q_t^- with $L_+^1(V)$ endowed with the strong topology is exponentially tight.*

Proof. For any $\eta > 0$, it is easy to see that

$$\varlimsup_{t\to\infty}\frac{1}{t}\log\mathbb{P}_\gamma(\|Q_t^+ - Q_t^-\| > \eta) = -\infty.$$

The result of this corollary now immediately follows from [22, Lemma 3.13]. $\quad\square$

We are now in a position to prove Theorem 3.7.

Proof of Theorem 3.7. Let $(Z,\tau) = \{(Z_k)_{k\geq 0}, (\tau_k)_{k\geq 1}\}$ be a Markov renewal process with the initial distribution γ^Z, where $Z_k = (X_{k-1}, X_k)$ and X_{-1} can be any random variables such that $\gamma^Z\in\mathcal{P}(E)$. Note that the empirical entrance-current for (Z,τ) is exactly the empirical flow for (X,τ), i.e.

$$Q_t(x,y) = \frac{1}{t}\sum_{k=1}^{N_t+1}1_{(X_{k-1}=x,X_k=y)} = \frac{1}{t}\sum_{k=1}^{N_t+1}1_{(Z_k=(x,y))} = Q_t^{-,Z}(x,y).$$

Next we will apply Corollary 5.2 to (Z,τ). It is easy to verify that (Z,τ) satisfies Assumptions 2.3-2.6. In order to apply Corollary 5.2, we need to prove that (Z,τ) satisfies Conditions 3.2 and 3.4. Here we only verify Condition 3.2 for (Z,τ) and the proof of Condition 3.4 is similar.

Let the functions $\hat{u}$, the sequence of functions u_n, the set K, and the constants c,σ,C,η,C_γ be as in Condition 3.2 and item (c*) in Remark 3.8. By choosing $u_n^Z(x,y) =$

$u_n(y)$, we immediately obtain item (b). Note that Z has the transition probability $p_{(x,y),(z,w)}^Z = \delta_y(z)p_{zw}$. Then for any $(x,y) \in E$ and $n \geq 0$, we have

$$P^Z u_n^Z(x,y) = \sum_{(z,w) \in E} \delta_y(z)p_{zw}u_n^Z(z,w) = \sum_{w \in V} p_{yw}u_n(w) = Pu_n(y) < \infty,$$

which implies item (a). For any $n \geq 0$, we have

$$\sum_{(x,y) \in E} \gamma^Z(x,y)u_n^Z(x,y) = \sum_{(x,y) \in E} \gamma^Z(x,y)u_n(y) = \sum_{y \in V} \gamma(y)u_n(y) \leq C_\gamma.$$

By choosing $C_\gamma^Z = C_\gamma$, we obtain item (c). Note that

$$\lim_{n \to \infty} \frac{u_n^Z(x,y)}{P^Z u_n^Z(x,y)} = \lim_{n \to \infty} \frac{u_n(y)}{Pu_n(y)} = \hat{u}(y).$$

By choosing $\hat{u}^Z(x,y) = \hat{u}(y)$, we obtain item (d). It is easy to check that $L^Z \hat{u}^Z(x,y) = L\hat{u}(y)$. Then for each $\ell \in \mathbb{R}$, we have

$$\left\{ (x,y) \in E : L^Z \hat{u}^Z(x,y) \leq \ell \right\} = \left\{ (x,y) \in E : L\hat{u}(y) \leq \ell \right\}.$$

Since (V,E) is locally finite, item (e) also holds. Note that $\psi_{(x,y)}^Z = \psi_y$ and $\zeta^Z(x,y) = \zeta(y)$. By choosing $\sigma^Z = \sigma$, $C^Z = C$, $\eta^Z = \eta$, and $K^Z = (V \times K) \cap E$, we obtain $\hat{u}^Z(x,y) < \psi_{(x,y)}^Z(e^{\zeta^Z(x,y)\tau})$ for any $(x,y) \in (K^Z)^c$ and

$$L^Z \hat{u}(x,y) = L\hat{u}(y) \geq -\sigma L\eta(y) - C1_K(y) \geq -\sigma^Z L^Z \eta^Z(x,y) - C^Z 1_{K^Z}(x,y).$$

This completes the proof of item (f). $\qquad \square$

6 Proof of Proposition 3.16

Here we consider the marginal LDP for the empirical flow Q_t. Before giving the proof of Proposition 3.16, we need some notation and lemmas. Let μ and ν be two positive σ-finite measures on a measurable space $(\mathcal{X}, \mathcal{F})$. For any sequence of non-negative measurable functions $(f_i)_{i \geq 1}$ on $\mathcal{X}$ and any sequence of non-negative constants $(b_i)_{i \geq 0}$ satisfying $0 < g := b_0 + \sum_{i=1}^\infty b_i f_i < \infty$, μ-a.s., let the generalized relative entropy between μ and ν be defined by

$$H_g(\mu|\nu) = \begin{cases} \int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \right) \mathrm{d}\mu, & \text{if } \mu \ll \nu, \\ \infty, & \text{otherwise.} \end{cases}$$

Similarly to the maximum entropy principle [12, Theorem 12.1.1], we have the following lemma.

Lemma 6.1. *Suppose that there exists a positive σ-finite measure μ^* satisfying $\mu^* \ll \nu$, $\nu \ll \mu^*$, and*

$$g \log \frac{\mathrm{d}\mu^*}{\mathrm{d}\nu} = \lambda_0 + \sum_{i=1}^\infty \lambda_i f_i, \tag{6.1}$$

where the sequence of constant $(\lambda_i)_{i \geq 0}$ are chosen so that μ^* satisfies the following constraints:

$$\mu(\mathcal{X}) = a_0, \quad \int_{\mathcal{X}} f_i d\mu = a_i, \quad i \geq 1, \tag{6.2}$$

where $(a_i)_{i \geq 0}$ is a sequence of constants satisfying $0 < \sum_{i=0}^{\infty} a_i b_i < \infty$. Then μ^* uniquely minimizes $H_g(\cdot|\nu)$ over all positive σ-finite probability measures satisfying (6.2). Moreover, the minimum is given by

$$H_g(\mu^*|\nu) = \sum_{i=0}^{\infty} a_i \lambda_i.$$

Proof. Let μ be a positive σ-finite measure satisfying (6.2). Note that $\int_{\mathcal{X}} g \mathrm{d}\mu = \int_{\mathcal{X}} g \mathrm{d}\mu^* = \sum_{i=0}^{\infty} a_i b_i < \infty$. Then we have

$$
\begin{aligned}
\int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \right) \mathrm{d}\mu &= \int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu}{\mathrm{d}\mu^*} \right) \mathrm{d}\mu + \int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu^*}{\mathrm{d}\nu} \right) \mathrm{d}\mu \\
&= \left(\int_{\mathcal{X}} g \mathrm{d}\mu \right) H \left(\frac{g \mathrm{d}\mu}{\int_{\mathcal{X}} g \mathrm{d}\mu} \Big| \frac{g \mathrm{d}\mu^*}{\int_{\mathcal{X}} g \mathrm{d}\mu^*} \right) + \int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu^*}{\mathrm{d}\nu} \right) \mathrm{d}\mu \qquad (6.3) \\
&\geq \int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu^*}{\mathrm{d}\nu} \right) \mathrm{d}\mu,
\end{aligned}
$$

where the last inequality follows from the nonnegativity of the relative entropy. It then follows from (6.1) that

$$
\begin{aligned}
\int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu^*}{\mathrm{d}\nu} \right) \mathrm{d}\mu &= \int_{\mathcal{X}} \left(\lambda_0 + \sum_{i=1}^{\infty} \lambda_i f_i \right) \mathrm{d}\mu = \sum_{i=0}^{\infty} \lambda_i a_i = \int_{\mathcal{X}} \left(\lambda_0 + \sum_{i=1}^{\infty} \lambda_i f_i \right) \mathrm{d}\mu^* \\
&= \int_{\mathcal{X}} g \left(\log \frac{\mathrm{d}\mu^*}{\mathrm{d}\nu} \right) \mathrm{d}\mu^*.
\end{aligned}
$$

Hence we have proved that $H_g(\mu|\nu) \geq H_g(\mu^*|\nu)$. Here the equality holds if and only if $\mathrm{d}\mu/\mathrm{d}\mu^* = 1$. This shows that except for a set of measure zero, μ^* is unique. $\qquad \square$

Recall the definition of $\zeta(x)$ in (3.1). We also need the following lemma.

Lemma 6.2. *For any $x \in V$, let $G_x(\lambda) = \log(\psi_x(e^{\lambda\tau}))$ and $F_x(\lambda) = \psi_x(\tau e^{\lambda\tau})/\psi_x(e^{\lambda\tau})$, where $\psi_x(\tau e^{\lambda\tau}) = \int_{(0,\infty)} s e^{\lambda s} \psi_x(\mathrm{d}s)$. Suppose that $G_x(\zeta(x)) = \infty$ for any $x \in V$. Then $G_x \in C^2(-\infty, \zeta(x))$ is a strictly increasing function, $F_x = \mathrm{d}G_x/\mathrm{d}\lambda$ is an increasing function, and*

$$\lim_{\lambda \to -\infty} F_x(\lambda) = m_x := \sup \left\{ \lambda \geq 0 : \psi_x(0, \lambda) = 0 \right\}, \qquad (6.4)$$

$$\lim_{\lambda \to \zeta(x)} F_x(\lambda) = M_x := \inf \left\{ \lambda \geq 0 : \psi_x(\lambda, \infty) = 0 \right\}. \qquad (6.5)$$

Moreover, if ψ_x is not a Dirac measure, then F_x is strictly increasing.

Proof. Without loss of generality, we can drop the dependence on x in the proof. We first prove that $G \in C^1((-\infty, \zeta))$ and $\mathrm{d}G/\mathrm{d}\lambda = F$. Note that $f_\epsilon(s) := (e^{(\lambda+\epsilon)s} - e^{\lambda s})/\epsilon$ converge pointwise to $s e^{\lambda s}$ as $\epsilon \to 0$, $|f_\epsilon(s)| \leq h(s) := e^{\lambda s}(e^{\eta s} - 1)/\eta$ for every $\epsilon \in (-\eta, \eta)$, and $\langle \psi, |h| \rangle < \infty$ for $\eta > 0$ small enough. By dominated convergence theorem, we immediately obtain the results. Moreover, since $F > 0$, it is easy to see that G is strictly increasing.

We next prove that $F \in C^1((-\infty, \zeta))$ is increasing. The proof of differentiability is similar to the above. Direct computations show that

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda}(\lambda) = \frac{\psi(\tau^2 e^{\lambda\tau})\psi(e^{\lambda\tau}) - (\psi(\tau e^{\lambda\tau}))^2}{(\psi(e^{\lambda\tau}))^2} \geq 0,$$

where the last inequality follows from Cauchy-Schwarz inequality. Moreover, the above equality holds if and only if ψ is a Dirac measure.

Finally, we prove (6.4) and (6.5). We only prove (6.5) and the proof of (6.4) is similar. Note that F is an increasing function. Let $L = \lim_{\lambda \to \zeta} F(\lambda)$.

Case 1: $\zeta < \infty$. It is easy to see that $M = \infty$. For any $0 \le \lambda \le \zeta$ and $N > 1$, we have

$$\psi\left(e^{\lambda \tau}\right) = \int_{(0,N]} e^{\lambda s} \psi(\mathrm{d}s) + \int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s) \le e^{\lambda N} + \int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s). \tag{6.6}$$

Note that $\lim_{\lambda \to \zeta} \psi(e^{\lambda \tau}) = e^{G(\zeta)} = \infty$. Taking $\lambda \to \zeta$ on both sides of (6.6), we obtain

$$\lim_{\lambda \to \zeta} \int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s) = \infty.$$

On the other hand, we have

$$F(\lambda) = \frac{\psi(\tau e^{\lambda \tau})}{\psi(e^{\lambda \tau})} \ge \frac{\int_{(N,\infty)} s e^{\lambda s} \psi(\mathrm{d}s)}{e^{\lambda N} + \int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s)} \ge \frac{N \int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s)}{e^{\lambda N} + \int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s)}. \tag{6.7}$$

Taking $\lambda \to \zeta$ on both sides of (6.7), we obtain $L \ge N$. Since $N > 1$ is arbitrary, it is clear that $L = \infty = M$.

Case 2: $\zeta = \infty$. Note that

$$F(\lambda) = \frac{\int_{(0,M]} s e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(0,M]} e^{\lambda s} \psi(\mathrm{d}s)} \le \frac{\int_{(0,M]} M e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(0,M]} e^{\lambda s} \psi(\mathrm{d}s)} = M. \tag{6.8}$$

Taking $\lambda \to \infty$ on both sides of (6.8), we obtain $L \le M$.

If $M < \infty$, for any $0 < \epsilon < M/2$ and $\lambda > 0$, we have

$$\begin{aligned}
\frac{1}{F(\lambda)} &= \frac{\int_{(0,M-2\epsilon]} e^{\lambda s} \psi(\mathrm{d}s) + \int_{(M-2\epsilon,M]} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(0,M]} s e^{\lambda s} \psi(\mathrm{d}s)} \\
&\le \frac{\int_{(0,M-2\epsilon]} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(M-\epsilon,M]} s e^{\lambda s} \psi(\mathrm{d}s)} + \frac{\int_{(M-2\epsilon,M]} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(M-2\epsilon,M]} s e^{\lambda s} \psi(\mathrm{d}s)} \\
&\le \frac{e^{\lambda(M-2\epsilon)}}{(M-\epsilon) e^{\lambda(M-\epsilon)} \psi((M-\epsilon,M])} + \frac{\int_{(M-2\epsilon,M]} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(M-2\epsilon,M]} (M-2\epsilon) e^{\lambda s} \psi(\mathrm{d}s)} \\
&= \frac{e^{-\lambda \epsilon}}{(M-\epsilon) \psi((M-\epsilon,M])} + \frac{1}{M-2\epsilon}.
\end{aligned} \tag{6.9}$$

Taking $\lambda \to \infty$ on both sides of (6.9), we obtain $L \ge M - 2\epsilon$. Since $0 < \epsilon < M/2$ is arbitrary, it is clear that $L \ge M$.

If $M = \infty$, for any $\epsilon > 0$, $N > 1$, and $\lambda > 0$, we have

$$\begin{aligned}
\frac{1}{F(\lambda)} &= \frac{\int_{(0,N]} e^{\lambda s} \psi(\mathrm{d}s) + \int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(0,\infty)} s e^{\lambda s} \psi(\mathrm{d}s)} \\
&\le \frac{\int_{(0,N]} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(N+\epsilon,\infty)} s e^{\lambda s} \psi(\mathrm{d}s)} + \frac{\int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(N,\infty)} s e^{\lambda s} \psi(\mathrm{d}s)} \\
&\le \frac{e^{\lambda N}}{(N+\epsilon) e^{\lambda(N+\epsilon)} \psi((N+\epsilon,\infty))} + \frac{\int_{(N,\infty)} e^{\lambda s} \psi(\mathrm{d}s)}{\int_{(N,\infty)} N e^{\lambda s} \psi(\mathrm{d}s)} \\
&= \frac{e^{-\lambda \epsilon}}{(N+\epsilon) \psi((N+\epsilon,\infty))} + \frac{1}{N}.
\end{aligned} \tag{6.10}$$

Taking $\lambda \to \infty$ on both sides of (6.10), we obtain $L \ge N$. Since $N > 1$ is arbitrary, it is clear that $L = \infty = M$. $\square$

By Lemma 6.2, we immediately obtain the following corollary.

Corollary 6.3. *For any $Q \in L^1_+(E)$ satisfying $Q^+ = Q^-$ and $V_+ := \{x \in V : Q_x > 0\} \neq \emptyset$, let $G_Q(\lambda) = \sum_{x \in V} Q_x G_x(\lambda)$ and $F_Q(\lambda) = \sum_{x \in V} Q_x F_x(\lambda)$. Suppose that $G_Q(\zeta_Q) = \infty$. Then $G_Q \in C^2(-\infty, \zeta_Q)$ is a strictly increasing function, $F_Q = \mathrm{d}G_Q/\mathrm{d}\lambda$ is an increasing function, and*

$$\lim_{\lambda \to -\infty} F_Q(\lambda) = m_Q := \sum_{x \in V} Q_x m_x, \qquad \lim_{\lambda \to \zeta_Q} F_Q(\lambda) = M_Q := \sum_{x \in V} Q_x M_x, \qquad (6.11)$$

where $\zeta_Q = \sup\{\lambda \geq 0 : G_Q(\lambda) < \infty\}$. Moreover, if $m_Q < M_Q$, then F_Q is strictly increasing.

Proof. We only prove the second equality of (6.11). The proof of the first equality is similar.

Case 1: $\zeta_Q = \infty$. By the monotone convergence theorem, we immediately obtain the desired result.

Case 2: $\zeta_Q < \infty$. Note that $\zeta_Q \leq \zeta(x)$ for any $x \in V_+$. By Lemma 6.2, we have $F_x(\zeta_Q) \leq F_x(\zeta(x)) = M_x$. It then follows from the monotone convergence theorem that

$$\lim_{\lambda \to \zeta_Q} F_Q(\lambda) = \lim_{\lambda \to \zeta_Q} \sum_{x \in V_+} Q_x F_x(\lambda) \leq \sum_{x \in V_+} Q_x M_x.$$

On the other hand, it is easy to see that

$$G_Q(\zeta_Q) \leq F_Q(\zeta_Q)\zeta_Q + G_Q(0).$$

Since $G_Q(\zeta_Q) = \infty$, we have $F_Q(\zeta_Q) = \infty$. This completes the proof of this corollary. $\square$

For any $Q \in L^1_+(E)$ satisfying $Q^+ = Q^-$ and $V_+ \neq \emptyset$, let $\mathcal{X} = V_+ \times (0, \infty)$ be endowed with the product topology and let $\mathcal{F}$ be the associated Borel σ-field. Let $\nu(x, \mathrm{d}s) = sQ_x\psi_x(\mathrm{d}s)$ be a σ-finite measure on $(\mathcal{X}, \mathcal{F})$. Here we take the sequence of functions $(f_x)_{x \in V_+}$ and the sequence of constants $(b_0, (b_x)_{x \in V_+})$ as

$$f_x(y, s) = \delta_x(y)\frac{1}{s}, \quad (y, s) \in \mathcal{X}, \qquad b_0 = 0, \quad b_x = 1.$$

Then the associated function g is given by $g(y, s) = b_0 + \sum_{x \in V_+} b_x f_x(y, s) = 1/s$ and the associated generalized relative entropy is given by

$$H_g(\mu|\nu) = \sum_{x \in V_+} \int_{(0,\infty)} \frac{1}{s}\left(\log\frac{\mu(x, \mathrm{d}s)}{sQ_x\psi_x(\mathrm{d}s)}\right)\mu(x, \mathrm{d}s).$$

Lemma 6.4. *Let $a > 0$ be a constant and let $Q \in L^1_+(E)$ be a flow satisfying $Q^+ = Q^-$ and $V_+ \neq \emptyset$. If $G_Q(\zeta_Q) = \infty$, then we have*

$$\inf_\mu H_g(\mu|\nu) = G_Q^*(a) = \begin{cases} a\lambda^* - G_Q(\lambda^*), & \text{if } m_Q < a < M_Q, \\ -\sum_{x \in V} Q_x \log\psi_x(\{m_x\}), & \text{if } a = m_Q, \\ -\sum_{x \in V} Q_x \log\psi_x(\{M_x\}), & \text{if } a = M_Q, \\ \infty, & \text{otherwise,} \end{cases}$$

where μ in the infimum ranges over all positive σ-finite measures on $(\mathcal{X}, \mathcal{F})$ satisfying

$$\mu(\mathcal{X}) = \sum_{x \in V_+} \mu(x, (0, \infty)) = a, \qquad \int_\mathcal{X} f_x \mathrm{d}\mu = \int_{(0,\infty)} \frac{1}{s}\mu(x, \mathrm{d}s) = Q_x, \quad x \in V_+, \quad (6.12)$$

λ^* is any solution of the equation $F_Q(\lambda) = a$, and $G_Q^*(a) = \sup_{\lambda \in \mathbb{R}}\{a\lambda - G_Q(\lambda)\}$, $a \in \mathbb{R}$ is the Fenchel-Legendre transform of G_Q.

Proof. We will prove this lemma in three different cases.

Case 1: $m_Q < a < M_Q$. Let $\lambda_0 = \lambda^*$ and $\lambda_x = -G_x(\lambda^*)$. Let $\mu^*(x, \mathrm{d}s) = sQ_x e^{\lambda^* s + \lambda_x} \psi_x(\mathrm{d}s)$ be a σ-finite measure μ^* on $(X, \mathcal{F})$. Then (6.1) holds and

$$\frac{1}{s} \log \frac{\mu^*(x, \mathrm{d}s)}{sQ_x \psi_x(\mathrm{d}s)} = \lambda^* + \lambda_x \frac{1}{s}, \quad (x, s) \in \mathcal{X}.$$

Moreover, it is easy to check that μ^* satisfies (6.12). By Lemma 6.1, we have

$$\inf_\mu H_g(\mu|\nu) = a\lambda^* + \sum_{x \in V_+} \lambda_x Q_x = a\lambda^* - G_Q(\lambda^*).$$

On the other hand, it follows from Corollary 6.3 and the intermediate value theorem that there exists a solution $\lambda^* \in (-\infty, \zeta_Q)$ of the equation $F_Q(\lambda) = a$. Straightforward computations show that

$$\sup_{\lambda \in \mathbb{R}} \{a\lambda - G_Q(\lambda)\} = a\lambda^* - G_Q(\lambda^*).$$

Case 2: $a = m_Q$ or $a = M_Q$. Here we only consider the case of $a = m_Q$. The proof in the case of $a = M_Q$ is similar. Let μ be a σ-finite measure satisfying (6.12). Then we have

$$a = m_Q = \sum_{x \in V_+} m_x \int_{[m_x, \infty)} \frac{1}{s} \mu(x, \mathrm{d}s) \leq \mu(X) = a.$$

The above equality hold if and only if $\mu(x, \mathrm{d}s) = Q_x m_x \delta_{m_x}(\mathrm{d}s)$ for any $x \in V_+$. Thus we have

$$\inf_\mu H_g(\mu|\nu) = -\sum_{x \in V_+} Q_x \log \psi_x(\{m_x\}).$$

On the other hand, direct computations show that

$$\sup_{\lambda \in \mathbb{R}} \{a\lambda - G_Q(\lambda)\} = \lim_{\lambda \to -\infty} \{a\lambda - G_Q(\lambda)\}$$

$$= \lim_{\lambda \to -\infty} \sum_{x \in V_+} Q_x \log \frac{e^{\lambda m_x}}{e^{\lambda m_x} \psi_x(\{m_x\}) + \int_{(m_x, \infty)} e^{\lambda s} \psi(\mathrm{d}s)}.$$

By the dominated convergence theorem, it is easy to see that

$$\lim_{\lambda \to -\infty} \int_{(m_x, \infty)} e^{\lambda(s - m_x)} \psi_x(\mathrm{d}s) = 0$$

for any $x \in V_+$. This implies that

$$\sup_{\lambda \in \mathbb{R}} \{a\lambda - G_Q(\lambda)\} = -\sum_{x \in V_+} Q_x \log \psi_x(\{m_x\}).$$

Case 3: $a < m_Q$ or $a > M_Q$. Here we only consider the case of $a < m_Q$. The proof in the case of $a > M_Q$ is similar. Note that there is no σ-finite measure μ satisfying (6.12). Then we have $\inf_\mu H_g(\mu|\nu) = \infty$. On the other hand, take a sequence of constants a_x such that $0 < a_x < m_x$ for any $x \in V_+$ and $a = \sum_{x \in V_+} Q_x a_x$. Straightforward computations show that

$$\sup_{\lambda \in \mathbb{R}} \{a\lambda - G_Q(\lambda)\} = \lim_{\lambda \to -\infty} \{a\lambda - G_Q(\lambda)\}$$

$$= \lim_{\lambda \to -\infty} \sum_{x \in V_+} Q_x \log \frac{1}{\int_{[m_x, \infty)} e^{\lambda(s - a_x)} \psi(\mathrm{d}s)} = \infty,$$

where the last equality follows from the dominated convergence theorem. $\qquad \square$

The following lemma gives the properties of G_Q^*.

Lemma 6.5. *Let $Q \in L_+^1(E)$ be a flow satisfying $Q^+ = Q^-$ and $V_+ \neq \emptyset$. If $m_Q < M_Q$ and $G_Q(\zeta_Q) = \infty$, then $G_Q^* \in C^2(m_Q, M_Q)$, $\lambda^* = \mathrm{d}G_Q^*/\mathrm{d}a$ is a strictly increasing function, and*

$$\lim_{a \to m_Q} \lambda^*(a) = -\infty, \qquad \lim_{a \to M_Q} \lambda^*(a) = \zeta_Q. \qquad (6.13)$$

where $\lambda^*(a)$ is any solution of the equation $F_Q(\lambda) = a$.

Proof. It follows from Lemma 6.3 that $F_Q \in C^1(-\infty, \zeta_Q)$ is strictly increasing. Hence $\lambda^* = \mathrm{d}G_Q^*/\mathrm{d}a$ is the inverse function of F_Q. Moreover, $\lambda^* \in C^1(m_Q, M_Q)$ is strictly increasing and (6.13) follows from (6.11). By Lemma 6.4, it is easy to see that $G_Q^* \in C^2(m_Q, M_Q)$ and

$$\frac{dG_Q^*}{da}(a) = \lambda^*(a) + a\frac{d\lambda^*}{da}(a) - \frac{d\lambda^*}{da}(a)F_Q(\lambda^*(a)) = \lambda^*(a).$$

This completes the proof of this lemma. $\square$

Finally, we also need the following lemma to ensure $G_Q(\zeta_Q) = \infty$.

Lemma 6.6. *Let $Q \in L_+^1(E)$ be a flow satisfying $Q^+ = Q^-$ and $V_+ \neq \emptyset$. If Condition 3.11 is satisfied, then we have $\zeta_Q = \inf_{x \in V_+} \zeta(x)$ and $G_Q(\zeta_Q) = \infty$.*

Proof. Let ψ, ζ, and q_x be as in Condition 3.11. It is easy to see that

$$\zeta(x) = q_x \zeta, \qquad G_x(\lambda) = G\left(\frac{\lambda}{q_x}\right),$$

where $G(\lambda) = \log \psi(e^{\lambda \tau})$. By item (c) in Condition 3.11, it is easy to see that there exists $y \in V_+$ such that $q_y = \inf_{x \in V_+} q_x > 0$. By Lemma 6.2, it is clear that G is increasing. Then for any $\lambda_0 < \inf_{x \in V_+} \zeta(x) = q_y \zeta$, we have

$$G_Q(\lambda_0) = \sum_{x \in V_+} Q_x G\left(\frac{\lambda_0}{q_x}\right) \leq \sum_{x \in V_+} Q_x G\left(\frac{\lambda_0}{q_y}\right) < \infty.$$

By the definition of ζ_Q, it is easy to see that $\zeta_Q \geq \lambda_0$. Since $\lambda_0 < \inf_{x \in V_+} \zeta(x)$ is arbitrary, we have $\zeta_Q \geq \inf_{x \in V_+} \zeta(x)$. On the other hand, we have

$$\varliminf_{\lambda \to \inf_{x \in V_+} \zeta(x)} G_Q(\lambda) \geq Q_y \varliminf_{\lambda \to q_y \zeta} G_y(\lambda) = Q_y \varliminf_{\lambda \to q_y \zeta} G\left(\frac{\lambda}{q_y}\right) = \infty.$$

This implies the proof of this lemma. $\square$

We are now in a position to prove Proposition 3.16.

Proof of Proposition 3.16. The proof of the marginal LDP for the empirical measure π_t is similar to that given in [35, Proposition 1.2]. Here we only focus on the marginal LDP for the empirical flow Q_t. Note that $(\mu, Q) \mapsto Q$ is a continuous map from Λ to $L_+^1(E)$. By the contraction principle [15, Theorem 4.2.1], the law of $Q_t \in L_+^1(E)$ satisfies an LDP with good rate function

$$\hat{I}_2(Q) = \inf\{I(\mu, Q) : (\mu, Q) \in \mathcal{D}\}.$$

We next prove $\hat{I}_2 = I_2$ under Condition 3.11. By Remark 4.18, we obtain

$$\hat{I}_2(Q) = \inf\{I(\mu, Q) : (\mu, Q) \in \mathcal{E}(Q)\},$$

where $\mathcal{E}(Q) = \{(\mu, Q) \in \mathcal{D} : \mu(x, \{\infty\}) = 0 \text{ for any } x \in V_+\}$. Recall the definition of I in (3.7). For any $Q \in L_+^1(E)$ satisfying $Q^+ = Q^-$, we have

$$\hat{I}_2(Q) = \sum_{x \in V} Q_x H\left(\tilde{Q}_{x,\cdot}\middle|p_{x,\cdot}\right) + \inf_{(\mu,Q)\in\mathcal{E}(Q)} \sum_{x\in V} \left[Q_x H\left(\tilde{\mu}_x\middle|\psi_x\right) + \zeta(x)\mu(x, \{\infty\})\right]$$

$$= \sum_{x \in V} Q_x H\left(\tilde{Q}_{x,\cdot}\middle|p_{x,\cdot}\right) + \inf_{(\mu,Q)\in\mathcal{E}(Q)} \left(\sum_{x\in V_+} Q_x H\left(\tilde{\mu}_x\middle|\psi_x\right) + \sum_{x\notin V_+} \zeta(x)\mu(x, \{\infty\})\right).$$

If $V_+ = \emptyset$, then we have

$$\inf_{(\mu,Q)\in\mathcal{E}(Q)} \left(\sum_{x\in V_+} Q_x H\left(\tilde{\mu}_x\middle|\psi_x\right) + \sum_{x\notin V_+} \zeta(x)\mu(x, \{\infty\})\right) = \inf_{\sum_{x\in V}\mu(x,\{\infty\})=1} \zeta(x)\mu(x, \{\infty\})$$

$$= \inf_{x\in V} \zeta(x),$$

which implies (3.10). If $V_+ \neq \emptyset$, then we have

$$\mathcal{E}(Q) = \bigcup_{0<a\leq 1} \mathcal{E}(Q, a),$$

where $\mathcal{E}(Q, a) = \{(\mu, Q) \in \mathcal{E}(Q) : \sum_{x\in V_+} \mu(x, (0,\infty)) = a \text{ and } \sum_{x\notin V_+} \mu(x, \{\infty\}) = 1 - a\}$. It then follows from Lemmas 6.4 and 6.6 that

$$\inf_{(\mu,Q)\in\mathcal{E}(Q)} \left(\sum_{x\in V_+} Q_x H\left(\tilde{\mu}_x\middle|\psi_x\right) + \sum_{x\notin V_+} \zeta(x)\mu(x, \{\infty\})\right)$$

$$= \inf_{0<a\leq 1} \inf_{(\mu,Q)\in\mathcal{E}(Q,a)} \left(\sum_{x\in V_+} Q_x H\left(\tilde{\mu}_x\middle|\psi_x\right) + \sum_{x\notin V_+} \zeta(x)\mu(x, \{\infty\})\right)$$

$$= \inf_{0<a\leq 1} \left[\inf_{(\mu,Q)\in\mathcal{E}(Q,a)} \left(\sum_{x\in V_+} \int_{(0,\infty)} \frac{1}{s}\left(\log \frac{\mu(x, ds)}{sQ_x\psi_x(ds)}\right)\mu(x, ds)\right) + (1-a)\inf_{x\notin V_+} \zeta(x)\right]$$

$$= \inf_{0<a\leq 1} \left[G_Q^*(a) + (1-a)\inf_{x\notin V_+} \zeta(x)\right].$$

We next calculate the last term of the above equation in five different cases.

Case 1: $m_Q > 1$. Since $G_Q^*(a) = \infty$ for any $0 < a \leq 1$, we have

$$\inf_{0<a\leq 1} \left[G_Q^*(a) + (1-a)\inf_{x\notin V_+} \zeta(x)\right] = \infty.$$

On the other hand, similarly to the proof of Case 3 in Lemma 6.4, we have

$$\sup_{\lambda < \inf_{x\in V} \zeta(x)} \left\{\lambda - \sum_{x\in V} Q_x \log \psi_x\left(e^{\lambda\tau}\right)\right\} = \lim_{\lambda\to-\infty} \left\{\lambda - \sum_{x\in V} Q_x \log \psi_x\left(e^{\lambda\tau}\right)\right\} = \infty.$$

Case 2: $m_Q \leq 1 < M_Q$ and $\lambda^*(1) \leq \inf_{x\notin V_+} \zeta(x)$. Straightforward computations show that

$$\inf_{0<a\leq 1} \left[G_Q^*(a) + (1-a)\inf_{x\notin V_+} \zeta(x)\right] = G_Q^*(1).$$

On the other hand, it follows from Lemma 6.5 that

$$\lambda^*(1) < \lim_{a\to M_Q} \lambda^*(a) = \inf_{x\in V_+} \zeta(x).$$

This implies that $\lambda^*(1) \leq \inf_{x \in V} \zeta(x)$. Direct computations show that

$$\sup_{\lambda < \inf_{x \in V} \zeta(x)} \left\{ \lambda - \sum_{x \in V} Q_x \log \psi_x \left(e^{\lambda \tau} \right) \right\} = \lambda^*(1) - \sum_{x \in V} Q_x \log \psi_x \left(e^{\lambda^*(1)\tau} \right) = G_Q^*(1).$$

Case 3: $m_Q \leq 1 < M_Q$ and $\lambda^*(1) > \inf_{x \notin V_+} \zeta(x)$. It follows from Lemma 6.5 that there exists $\eta \in (m_Q, 1)$ such that $\lambda^*(\eta) = \inf_{x \notin V_+} \zeta(x)$. Straightforward computations show that

$$\inf_{0 < a \leq 1} \left[G_Q^*(a) + (1-a) \inf_{x \notin V_+} \zeta(x) \right] = G_Q^*(\eta) + (1-\eta) \inf_{x \notin V_+} \zeta(x)$$

$$= \eta \lambda^*(\eta) - \sum_{x \in V} Q_x \log \psi_x \left(e^{\lambda^*(\eta)\tau} \right) + (1-\eta)\lambda^*(\eta)$$

$$= \lambda^*(\eta) - \sum_{x \in V} Q_x \log \psi_x \left(e^{\lambda^*(\eta)\tau} \right)$$

$$= \sup_{\lambda < \inf_{x \in V} \zeta(x)} \left\{ \lambda - \sum_{x \in V} Q_x \log \psi_x \left(e^{\lambda \tau} \right) \right\}. \tag{6.14}$$

Note that $\inf_{x \notin V_+} \zeta(x) < \lambda^*(1) \leq \lim_{a \to M_Q} \lambda^*(a) = \inf_{x \in V_+} \zeta(x)$. Then the last equability in (6.14) follows from the fact that $\inf_{x \in V} \zeta(x) = \inf_{x \notin V_+} \zeta(x) = \lambda^*(\eta)$.

Case 4: $M_Q \leq 1$ and $\inf_{x \notin V_+} \zeta(x) = \infty$. Since $M_Q < \infty$, it is easy to see that $\inf_{x \in V_+} \zeta(x) = \infty$. Then we have $\inf_{x \in V} \zeta(x) = \infty$ and

$$\inf_{0 < a \leq 1} \left[G_Q^*(a) + (1-a) \inf_{x \notin V_+} \zeta(x) \right] = G^*(1) = \sup_{\lambda < \inf_{x \in V} \zeta(x)} \left\{ \lambda - \sum_{x \in V} Q_x \log \psi_x \left(e^{\lambda \tau} \right) \right\}.$$

Case 5: $M_Q \leq 1$ and $\inf_{x \notin V_+} \zeta(x) < \infty$. Note that $G_Q^*(a) = \infty$ for any $a > M_Q$. Then we have

$$\inf_{0 < a \leq 1} \left[G_Q^*(a) + (1-a) \inf_{x \notin V_+} \zeta(x) \right] = \inf_{0 < a \leq M_Q} \left[G_Q^*(a) + (1-a) \inf_{x \notin V_+} \zeta(x) \right].$$

By Lemma 6.5, we have $\lambda^*(M_Q) = \infty$. Since $\inf_{x \in V_+} \zeta(x) = \infty$, there exists $\eta \in (m_Q, M_Q)$ such that $\lambda^*(\eta) = \inf_{x \notin V_+} \zeta(x)$. Similarly to the proof in Case 3, we have

$$\inf_{0 < a \leq M_Q} \left[G_Q^*(a) + (1-a) \inf_{x \notin V_+} \zeta(x) \right] = G_Q^*(\eta) + (1-\eta) \inf_{x \notin V_+} \zeta(x)$$

$$= \sup_{\lambda < \inf_{x \in V} \zeta(x)} \left\{ \lambda - \sum_{x \in V} Q_x \log \psi_x \left(e^{\lambda \tau} \right) \right\}.$$

This completes the proof of this proposition. □

7 Another compactness condition

In [6], the authors imposed an alternative compactness condition, which is different from Condition 3.4 proposed in the present paper, and proved the joint LDP for the empirical measure and empirical flow of continuous-time Markov chains when $L_+^1(E)$ is endowed with the strong topology. In fact, we can also prove the joint LDP for semi-Markov processes under a similar compactness condition. To see this, we introduce some notation. For any set $\hat{E} \subset E$, we define an $\hat{E}$-dependent function $R : V \to \mathbb{R}$ by

$$R(y) = \sum_{z:(y,z) \in \hat{E}} p_{yz}, \quad y \in V. \tag{7.1}$$

Condition 7.1. *Suppose there exists a set $\hat{E} \subset E$ satisfying the following three requirements:*

(a) *For any $y \in V$, there exists $z \in V$ such that $(y, z) \in \hat{E}$;*

(b) *The function $R : V \to (0, \infty)$ defined in (7.1) vanishes at infinity;*

(c) *For any $x \in V$, there exist constants $a_0, \lambda > 0$ such that for any $a < a_0$ one can find a set $W = W(x, a) \subset E$ satisfying the following properties:*

(i) *The complement $E \setminus W$ is finite;*

(ii) *If $(y, z) \in W$, then $R(y) < a$;*

(iii) *For each path exiting from x, the number of edges in $\hat{E} \cap W$ is at least λ-times the total number of edges in W. In other words, for any path $x_0, x_1, \dots, x_k$ with $x_0 = x$ and $(x_i, x_{i+1}) \in E$, we have*

$$\sharp \left\{ i : (x_{i-1}, x_i) \in \hat{E} \cap W, 1 \le i \le k \right\} \ge \lambda \sharp \left\{ i : (x_{i-1}, x_i) \in W, 1 \le i \le k \right\}.$$

Obviously, we have $\lambda < 1$.

Note that both Conditions 3.4 and 7.1 only depend on the embedded chain X of the semi-Markov process. The following proposition shows that Condition 3.4 is weaker than Condition 7.1. Due to this reason, we impose the former rather than the latter in Section 3.1.

Proposition 7.2. *Suppose that Assumption 2.3 and Condition 7.1 are satisfied. Then Condition 3.4 holds.*

Proof. Let the constants $a_0, \lambda > 0$ be as in item (c) of Condition 7.1. For any $x_0 \in V$ and $m \in \mathbb{N}$ satisfying $1/m < a_0$, we will next construct the sequence of functions u_n, $n \ge m$ such that Condition 3.4 holds. We first construct auxiliary functions $F_n : E \to (0, \infty)$ by induction. Let $W_n \subset E$ be the sequence of sets defined as

$$W_m = W\left(x_0, \frac{1}{m}\right), \qquad W_n = W\left(x_0, \frac{1}{n}\right) \cap \left(W_{n-1} \cup \hat{E}\right), \quad n \ge m+1.$$

It is easy to see that W_n also satisfies item (c) in Condition 7.1 for each $n \ge m$. Set

$$F_m(y, z) = \begin{cases} m^{1-\lambda}, & \text{if } (y, z) \in \hat{E} \cap W_m, \\ m^{-\lambda}, & \text{if } (y, z) \in \hat{E}^c \cap W_m, \\ 1, & \text{if } (y, z) \in E \setminus W_m, \end{cases}$$

and for $n \ge m+1$, set

$$F_n(y, z) = \begin{cases} n^{1-\lambda}, & \text{if } (y, z) \in \hat{E} \cap W_n, \\ n^{-\lambda}, & \text{if } (y, z) \in \hat{E}^c \cap W_n, \\ F_{n-1}(y, z), & \text{if } (y, z) \in E \setminus W_n. \end{cases}$$

Obviously, $F_n(x, y) \le n^{1-\lambda}$ for any $(x, y) \in E$. For any $x \in V$, let G_x be the collection of all paths in (V, E) with initial state x_0 and terminal state x, i.e.

$$G_x = \left\{ (x_0, x_1, \cdots, x_k) : k \in \mathbb{N}, x_k = x, x_i \in V, \text{ and } (x_i, x_{i+1}) \in E \right\}.$$

It then follows from Assumption 2.3 that (V, E) is connected. This implies that $G_x \ne \emptyset$ for any $x \in V$. For each $n \ge m$, let $u_n : V \to (0, \infty)$ be the function defined by

$$u_n(x) = \inf_{\mathbf{x} \in G_x} \prod_{i=1}^{k} F_n(x_{i-1}, x_i), \quad x \in V.$$

Now we verify Condition 3.4 for the sequence of functions u_n.

(a) Obviously, if $(y, z) \in E$, then we have $u_n(z) \leq u_n(y)F_n(y, z) \leq u_n(y)n^{1-\lambda}$. For any $x \in V$ and $n \geq m$, we obtain

$$Pu_n(x) = \sum_{y:(x,y)\in E} p_{xy}u_n(y) \leq \sum_{y:(x,y)\in E} p_{xy}u_n(x)n^{1-\lambda} = u_n(x)n^{1-\lambda} < \infty.$$

(b) Recall the definition of auxiliary functions F_n given above. For any path $\mathbf{x} = (x_0, x_1, \cdots, x_k) \in G_x$, it follows from item (c)-(iii) in Condition 7.1 that

$$\prod_{i=1}^{k} F_m(x_{i-1}, x_i) = m^{[(1-\lambda)\#\{i:(x_{i-1},x_i)\in \hat{E}\cap W_m, 1\leq i\leq k\}-\lambda\#\{i:(x_{i-1},x_i)\in \hat{E}^c\cap W_m, 1\leq i\leq k\}]} \tag{7.2}$$
$$= m^{[\#\{i:(x_{i-1},x_i)\in \hat{E}\cap W_m, 1\leq i\leq k\}-\lambda\#\{i:(x_{i-1},x_i)\in W_m, 1\leq i\leq k\}]} \geq 1.$$

Similarly, for $n \geq m + 1$, we have

$$\prod_{i=1}^{k} \frac{F_n(x_{i-1}, x_i)}{F_{n-1}(x_{i-1}, x_i)} = \prod_{i:(x_{i-1},x_i)\in \hat{E}\cap W_n} \frac{n^{1-\lambda}}{F_{n-1}(x_{i-1}, x_i)} \prod_{i:(x_{i-1},x_i)\in \hat{E}^c\cap W_n} \frac{n^{-\lambda}}{F_{n-1}(x_{i-1}, x_i)}$$
$$\geq \left(\frac{n}{n-1}\right)^{[(1-\lambda)\#\{i:(x_{i-1},x_i)\in \hat{E}\cap W_n, 1\leq i\leq k\}-\lambda\#\{i:(x_{i-1},x_i)\in \hat{E}^c\cap W_n, 1\leq i\leq k\}]} \geq 1. \tag{7.3}$$

Indeed, $\hat{E}^c \cap W_n = \hat{E}^c \cap W(x_0, 1/n) \cap W_{n-1}$, which implies the first inequality in (7.3). For any $x \in V$, optimizing over G_x on both sides of (7.2), we have $u_m(x) \geq 1$. Similarly, it follows from (7.3) that $u_n(x) \geq u_{n-1}(x)$ for $n \geq m + 1$. This implies that u_n is an increasing sequence of functions and $u_n(x) \geq 1$ for any $x \in V$ and $n \geq m$.

(c) For any $y \in V$, it follows from item (a) in Condition 7.1 that $R(y) > 0$. For any $x \in V$, since $G_x \neq \emptyset$, there exists $\mathbf{x} = (x_0, x_1 \cdots, x_k) \in G_x$. Hence, we can let $N = m \vee (1 + \max_{0\leq i\leq k}\lfloor 1/R(x_i)\rfloor)$. By item (c)-(ii) in Condition 7.1, it is easy to see that $(x_{i-1}, x_i) \notin W_N$ for any $0 \leq i \leq k$. Then for any $n \geq N$, we have

$$u_n(x) \leq \prod_{i=1}^{k} F_n(x_{i-1}, x_i) = \prod_{i=1}^{k} F_N(x_{i-1}, x_i) \leq N^{(1-\lambda)k}.$$

Obviously, N and k only depend on x.

(d) For any $x \in V$, let $\mathbf{x}$ and N be defined as in (c). Since $R(x) > 1/N$, it is easy to see that $(x, y) \notin W_N$ for any $(x, y) \in E$. Then for any $n \geq N$, we have

$$Pu_n(x) \leq \sum_{y:(x,y)\in E} p_{xy}u_n(x)F_n(x, y) = \sum_{y:(x,y)\in E} p_{xy}u_n(x)F_N(x, y) \leq N^{1-\lambda}u_n(x). \tag{7.4}$$

Note that u_n is an increasing sequence of functions with an upper bound. Then we define $u^*(x) = \lim_{n\to\infty} u_n(x)$. It follows from (7.4) and the monotone convergence theorem that

$$Pu^*(x) = \lim_{n\to\infty} Pu_n(x) \leq N^{1-\lambda}u^*(x).$$

Thus u_n/Pu_n converges pointwise to the function $\hat{u} = u^*/Pu^* : V \to (0, \infty)$.

(e) Let V_n and V_n' be two sequences of subsets of V defined by

$$V_n = \{x \in V : (x, y) \in E \text{ implies } (x, y) \in W_n \text{ for any } y \in V\},$$

$$V_n' = \left\{x \in V : R(x) < \frac{1}{n}\right\}, \quad n \geq m.$$

Obviously, we have $V_n \subset V'_n$. For any $x \in V_j \subset V'_j = \cup_{i=j}^\infty (V'_i \setminus V'_{i+1})$, there exists $i \geq j$ such that $x \in V'_i \setminus V'_{i+1}$. If $(x,y) \in E$, it is easy to see that $(x,y) \notin W_n$ for $n \geq i+1$. This implies that $F_n(x,y) = F_i(x,y)$ for $n \geq i$. Then we have

$$u^*(y) = \lim_{n\to\infty} u_n(y) \leq \lim_{n\to\infty} u_n(x) F_n(x,y) = u^*(x) F_i(x,y).$$

Note that for any $(x,y) \in E$, we have $(x,y) \in W_j$. Then we have

$$\frac{Pu^*(x)}{u^*(x)} \leq \sum_{y:(x,y)\in E} p_{xy} F_i(x,y) = \sum_{y:(x,y)\in \hat{E}} p_{xy} F_i(x,y) + \sum_{y:(x,y)\in \hat{E}^c} p_{xy} F_i(x,y)$$
$$\leq R(x) i^{1-\lambda} + j^{-\lambda} \leq i^{-\lambda} + j^{-\lambda} \leq 2j^{-\lambda}.$$

This means that $\log \hat{u}(x) \geq \log(j^\lambda/2)$ for any $x \in V_j$. By item (c)-(i) in Condition 7.1, it is easy to see that

$$V \setminus V_j = \{x \in V : \exists\, y \text{ such that } (x,y) \in E \setminus W_j\}$$

is a finite set. For any $\ell > 0$, select j such that $\log(j^\gamma/2) > \ell$. Then $\{x \in V : \log \hat{u}(x) \leq \ell\} \subset V \setminus V_j$ is a finite set. $\qquad\square$

Corollary 7.3. *Suppose that Assumptions 2.3-2.6 and Conditions 3.2 and 7.1 are satisfied. Then the results of Theorem 3.7 remain valid.*

Proof. The results follow directly from Theorem 3.7 and Proposition 7.2. $\qquad\square$

References

[1] David Andrieux and Pierre Gaspard, *Fluctuation theorem for currents and Schnakenberg network theory*, J. Stat. Phys. **127** (2007), no. 1, 107–131. MR2313063

[2] David Andrieux and Pierre Gaspard, *The fluctuation theorem for currents in semi-Markov processes*, J. Stat. Mech.: Theory Exp. **2008** (2008), no. 11, P11007. MR2461405

[3] Soeren Asmussen, *Applied probability and queues*, 2 ed., vol. 51, Springer, New York, 2003. MR1978607

[4] Srinivasan Balaji and Sean P. Meyn, *Multiplicative ergodicity and large deviations for an irreducible Markov chain*, Stoch. Proc. Appl. **90** (2000), no. 1, 123–144. MR1787128

[5] Giada Basile and Lorenzo Bertini, *Donsker-Varadhan asymptotics for degenerate jump Markov processes*, ALEA Lat. Am. J. Probab. Math. Stat. **12** (2015), no. 1, 1–23. MR3333733

[6] Lorenzo Bertini, Alessandra Faggionato, and Davide Gabrielli, *Flows, currents, and cycles for Markov chains: large deviation asymptotics*, Stoch. Proc. Appl. **125** (2015), no. 7, 2786–2819. MR3332855

[7] Lorenzo Bertini, Alessandra Faggionato, and Davide Gabrielli, *Large deviations of the empirical flow for continuous time Markov chains*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 3, 867–900. MR3365965

[8] Vladimir I. Bogachev, *Measure theory*, vol. 1, Springer, Berlin, Heidelberg, 2007. MR2267655

[9] Amarjit Budhiraja, Yong Chen, and Lihu Xu, *Large deviations of the entropy production rate for a class of Gaussian processes*, J. Math. Phys. **62** (2021), no. 5, Paper No. 052702, 25. MR4252725

[10] Mu-Fa Chen, *Eigenvalues, inequalities, and ergodic theory*, Springer, London, 2006. MR2105651

[11] Yong Chen, Hao Ge, Jie Xiong, and Lihu Xu, *The large deviation principle and steady-state fluctuation theorem for the entropy production rate of a stochastic process in magnetic fields*, J. Math. Phys. **57** (2016), no. 7, 073302, 16. MR3528662

[12] Thomas M. Cover and Joy A.Thomas, *Elements of information theory*, 2 ed., John Wiley & Sons, Inc., Hoboken, New Jersey, 2005. MR2239987

[13] Harald Cramér, *Sur un nouveau théorème-limite de la théorie des probabilités*, Actual. Sci. Ind. **736** (1938), 5–23.

[14] Arnaud de La Fortelle, *Large deviation principle for Markov chains in continuous time*, Probl. Inf. Transm. **37** (2001), no. 2, 120–139. MR2099897

[15] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, 2 ed., vol. 98, Springer, New York, 1998. MR1619036

[16] Monroe D. Donsker and Srinivasa R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time, I*, Commun. Pure Appl. Math. **28** (1975), no. 1, 1–47. MR0386024

[17] Monroe D. Donsker and Srinivasa R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time, II*, Commun. Pure Appl. Math. **28** (1975), no. 2, 279–301. MR0386024

[18] Monroe D. Donsker and Srinivasa R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time. III*, Comm. Pure Appl. Math. **29** (1976), no. 4, 389–461. MR428471

[19] Monroe D. Donsker and Srinivasa R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time. IV*, Comm. Pure Appl. Math. **36** (1983), no. 2, 183–212. MR690656

[20] Alessandra Faggionato, *Large deviation principles and fluctuation theorems for currents in semi-Markov processes*, arXiv preprint arXiv:1709.05653 (2017). MR3606734

[21] Alessandra Faggionato and Vittoria Silvestri, *Fluctuation theorems for discrete kinetic models of molecular motors*, J. Stat. Mech.: Theory Exp. **2017** (2017), no. 4, 043206. MR3652725

[22] Jin Feng and Thomas G Kurtz, *Large deviations for stochastic processes*, vol. 131, American Mathematical Society, 2006. MR2260560

[23] Hao Ge, Chen Jia, and Da-Quan Jiang, *Cycle symmetry, limit theorems, and fluctuation theorems for diffusion processes on the circle*, Stoch. Proc. Appl. **127** (2017), no. 6, 1897–1925. MR3646435

[24] Hao Ge, Chen Jia, and Xiao Jin, *Martingale structure for general thermodynamic functionals of diffusion processes under second-order averaging*, J. Stat. Phys. **184** (2021), no. 2, Paper No. 17, 41. MR4292485

[25] Jacques Janssen and Raimondo Manca, *Applied semi-Markov processes*, Springer, New York, 2006. MR2175472

[26] Chen Jia and Ramon Grima, *Frequency domain analysis of fluctuations of mRNA and protein copy numbers within a cell lineage: theory and experimental validation*, Phys. Rev. X **11** (2021), 021032.

[27] Chen Jia, Da-Quan Jiang, and Min-Ping Qian, *Cycle symmetries and circulation fluctuations for discrete-time and continuous-time Markov chains*, Ann. Appl. Probab. **26** (2016), no. 4, 2454–2493. MR3543902

[28] Chen Jia and Youming Li, *Analytical time-dependent distributions for gene expression models with complex promoter switching mechanisms*, SIAM J. Appl. Math. **83** (2023), no. 4, 1572–1602. MR4623289

[29] Chen Jia, Abhyudai Singh, and Ramon Grima, *Concentration fluctuations in growing and dividing cells: Insights into the emergence of concentration homeostasis*, PLoS Comput. Biol. **18** (2022), no. 10, e1010574.

[30] Yuhao Jiang, Bingjie Wu, and Chen Jia, *Large deviations and fluctuation theorems for cycle currents defined in the loop-erased and spanning tree manners: a comparative study*, Phys. Rev. Res. **5** (2023), no. 1, 013207.

[31] George Kesidis and Jean Walrand, *Relative entropy between Markov transition rate matrices*, IEEE Trans. Inform. Theory **39** (1993), no. 3, 1056–1057. MR1237729

[32] Claude Kipnis and Claudio Landim, *Scaling limits of interacting particle systems*, vol. 320, Springer, Berlin, Heidelberg, 1999. MR1707314

[33] Anatoly B Kolomeisky and Michael E Fisher, *Extended kinetic models with waiting-time distributions: exact results*, J. Chem. Phys. **113** (2000), no. 24, 10867–10877. MR1797136

[34] Nikolaos Limnios and Gheorghe Oprisan, *Semi-Markov processes and reliability*, Birkhäuser Boston, MA, 2012. MR1843923

[35] Mauro Mariani and Lorenzo Zambotti, *Large deviations for the empirical measure of heavy-tailed Markov renewal processes*, Adv. in Appl. Probab. **48** (2016), no. 3, 648–671. MR3568885

[36] Robert E. Megginson, *An introduction to banach space theory*, vol. 183, Springer, New York, 2012. MR1650235

[37] Sean P. Meyn and Richard L. Tweedie, *Markov chains and stochastic stability*, Springer, London, 2012. MR2509253

[38] James Norris, *Markov chains*, Cambridge University Press, Cambridge, 1997. MR1600720

[39] Juan M. Pedraza and Johan Paulsson, *Effects of molecular memory and bursting on fluctuations in gene expression*, Science **319** (2008), no. 5861, 339–343.

[40] Hong Qian and X. Sunney Xie, *Generalized Haldane equation and fluctuation theorem in the steady-state cycle kinetics of single enzymes*, Phys. Rev. E **74** (2006), no. 1, 010902.

[41] Min-Ping Qian, Guanglu Gong, Dayue Chen, and Fuxi Zhang, *Applied stochastic processes (chinese)*, Higher Education Press, Beijing, 2011.

[42] Hongyun Wang and Hong Qian, *On detailed balance and reversibility of semi-Markov processes and single-molecular enzyme kinetics*, J. Math. Phys. **48** (2007), no. 1, 013303, 15. MR2292615

[43] Qingda Wei and Xian Chen, *Nonzero-sum expected average discrete-time stochastic games: the case of uncountable spaces*, SIAM J. Control Optim. **57** (2019), no. 6, 4099–4124. MR4040788

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System: https://vtex.lt/services/ejms-peer-review/
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: https://imstat.org/shop/donation/