Electronic Journal of Probability

Vol. 8 (2003) Paper no. 16, pages 1–46.

Journal URL
http://www.math.washington.edu/~ejpecp/

# Large Deviations for the Empirical Measures of Reflecting Brownian Motion and Related Constrained Processes in $I\!R_+$

**Amarjit Budhiraja**[1]
Department of Statistics, University of North Carolina
Chapel Hill, NC 27599-3260, USA

**Paul Dupuis**[2]
Lefschetz Center for Dynamical Systems
Division of Applied Mathematics, Brown University
Providence, RI 02912, USA

**Abstract:** We consider the large deviations properties of the empirical measure for one dimensional constrained processes, such as reflecting Brownian motion, the M/M/1 queue, and discrete time analogues. Because these processes do not satisfy the strong stability assumptions that are usually assumed when studying the empirical measure, there is significant probability (from the perspective of large deviations) that the empirical measure charges the point at infinity. We prove the large deviation principle and identify the rate function for the empirical measure for these processes. No assumption of any kind is made with regard to the stability of the underlying process.

**Key words and phrases:** Markov process, constrained process, large deviations, empirical measure, stability, reflecting Brownian motion.

**AMS subject classification (2000):** 60F10, 60J25, 93E20.

Submitted to EJP on February 15, 2002. Final version accepted on August 20, 2003.

[1]Research supported in part by the IBM junior faculty development award, University of North Carolina at Chapel Hill.

[2]Research supported in part by the National Science Foundation (NSF-DMS-0072004) and the Army Research Office (ARO-DAAD19-99-1-0223) and (ARO-DAAD19-00-1-0549).

## 1 Introduction

Let $\{X_n, n \in I\!N_0\}$ be a Markov process on a Polish space $S$, with transition kernel $p(x, dy)$. The empirical measure (or normalized occupation measure) for this process is defined by

$$L^n(A) \doteq \frac{1}{n} \sum_{i=0}^{n-1} \delta_{X_i}(A),$$

where $\delta_x$ is the probability measure that places mass 1 at $x$, and $A$ is any Borel subset of $S$. One of the cornerstones in the general theory of large deviations, due to Donsker and Varadhan in [6], was the development of a large deviation principle for the occupation measures for a wide class of Markov chains taking values in a compact state space. This work also studied the empirical measure large deviation principle (LDP) for continuous time Markov processes, where $L^n$ is replaced by

$$L^T(A) \doteq \frac{1}{T} \int_0^T \delta_{X(t)}(A) dt$$

and $X(\cdot)$ is a suitable continuous time Markov process. In subsequent work [7], the results of [6] were extended to Markov processes with an arbitrary Polish state space. These results significantly extended Sanov's theorem, which treats the independent and identically distributed (iid) case and was at that time the state-of-the-art. This work has found many applications since that time, and the general topic has developed into one of the most fertile areas of research in large deviations.

Three main assumptions appear in the empirical measure LDP results proved in [6, 7], and also in most subsequent papers on the same subject. The first is a sort of mixing or transitivity condition, and is the key to the proof of the large deviation lower bound. The second condition is a Feller property on the transition kernel, which is used in the proof of the upper bound. The third condition, and the one of prime interest in the present work, is a strong assumption on the stability of the underlying Markov process. For example, suppose that the underlying process is the solution to the $I\!R^n$-valued stochastic differential equation

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t),$$

where the dimensions of the Wiener process $W$ and $b$ and $\sigma$ are compatible, and $b$ and $\sigma$ are Lipschitz continuous. Then satisfaction of the stability assumption would require the existence of a Lyapunov function $V : I\!R^n \to I\!R$ with the property that

$$\langle b(x), V_x(x)\rangle + \frac{1}{2}\mathrm{tr}\left[\sigma(x)\sigma'(x)V_{xx}(x)\right] \to -\infty \text{ as } \|x\| \to \infty.$$

Here $V_x$ and $V_{xx}$ denote the gradient and Hessian of $V$, respectively, tr denotes trace, and the prime denotes transpose.

A condition of this sort implies a strong restoring force towards bounded sets, and in fact a force that grows without bound as $\|x\| \to \infty$. It is required for a simple reason, and that is to keep the probability that the occupation measure "charges" points at $\infty$ small from a large deviation perspective. Under such conditions, the probability that a sample path wanders out to infinity (i.e., eventually escapes every compact set) on the time interval $[0, T]$ ($\{0, 1, ..., n-1\}$ in discrete time) is super-exponentially small. It is a reasonable condition in many circumstances. For example, in the setting of the linear system

$$dX(t) = AX(t)dt + BdW(t),$$

it simply requires that the matrix $A$ correspond to a stable (deterministic) linear system. Nonetheless, the condition is not satisfied by some very basic processes. Examples include Brownian motion, reflecting Brownian motion, and the Markov process that corresponds to the $M/M/1$ queue.

In the present paper we consider a class of one dimensional reflected (or constrained) processes which includes the last two examples, and obtain the large deviation principle for the empirical measures without any stability assumptions at all. The restriction to $\mathbb{R}_+$ allows us to focus on one main issue: how to deal with the possibility that some of the mass of the occupation measure is placed on the point $\infty$ (asymptotically and from the perspective of large deviations). In a more general setting there are many ways the underlying process can wander out to infinity, and hence the analysis becomes more complex. We defer this general setup to later work.

We begin our study with a discrete time Markov process on $\mathbb{R}_+$. This Markov chain is introduced in Section 2. Once the empirical measure LDP for this family of models is obtained, one can obtain the LDP for the continuous time Markov processes described by a reflected Brownian motion and an $M/M/1$ queue via the standard technique of approximating by suitable super exponentially close processes (cf. [6, Section 3]). This is discussed in greater detail in Section 6.

The removal of the strong stability assumption fundamentally changes the nature of both the large deviation result and the proofs. In particular, given that the empirical measure will put mass on infinity, one must have detailed information on how this happens. We will adopt the weak convergence method of [9]. This approach is natural for the problem at hand, and indeed the combination of appropriately constructed test functions and weak convergence supplies us with exactly the sort of information we need (see, e.g., Lemma 3.9).

As stated earlier in the introduction, the fundamental results on the empirical measure LDP for Markov processes were obtained in [6, 7]. Subsequently, a large amount of work has been done by various authors in refining these basic results. We refer the reader to [4, 5] for a detailed history of the problem. Most of the available work studies Markov processes that satisfy the strong stability assumptions. One notable exception is the work by Ney and Nummelin [13, 14], where large deviation probabilities for additive functionals of Markov chain are considered, essentially assuming only the irreducibility of the underlying Markov chain. However, the goals there are quite different from ours in that the authors obtain *local* large deviation results. Other authors, such as [2, 10, 3], study the large deviation lower bounds under weaker hypothesis than those in [6, 7]. The proof of the lower bound in these papers does not require any stability assumption on the underlying Markov chain. However, in the absence of strong stability, their lower bound is not (in general) the best possible. To illustrate the basic issue we restrict our attention to the discrete time model introduced in Section 2. Let $\mathcal{P}(I\!\!R_+)$ be the space of probability measures on $I\!\!R_+$. Let $F : \mathcal{P}(I\!\!R_+) \mapsto I\!\!R$ be a continuous and bounded map defined as

$$F(\nu) \doteq \int_{I\!\!R_+} f(x)\nu(dx), \quad \nu \in \mathcal{P}(I\!\!R_+).$$

Here $f$ is a real-valued continuous and bounded function on $I\!\!R_+$ such that $f(x)$ converges, to say $f^\infty$, as $x \to \infty$. For such a function the lower bound in the cited papers will imply

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \log I\!\!E \left[ \exp\left( -n \int_{I\!\!R_+} f(x) dL^n(x) \right) \right] \\
&\qquad \geq - \inf_{\nu \in \mathcal{P}(I\!\!R_+)} \left\{ I_1(\nu) + \int_{I\!\!R_+} f(x)\nu(dx) \right\},
\end{aligned} \tag{1.1}$$

where $I_1$ is defined in Section 2 (see (2.6)). In contrast, for the class of one dimensional processes studied in this paper we will show that

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \log I\!\!E \left[ \exp\left( -n \int_{I\!\!R_+} f(x) dL^n(x) \right) \right] \\
&\quad = - \inf_{\nu \in \mathcal{P}(I\!\!R_+),\ \rho\in[0,1]} \left\{ \rho I_1(\nu) + (1-\rho)J + \rho \int_{I\!\!R_+} f(x)\nu(dx) + (1-\rho) f_\infty \right\},
\end{aligned} \tag{1.2}$$

where $J$ is the part of the rate function that accounts for the possibility that mass might wander off to infinity. Clearly, the expression on the right side of (1.2) provides a sharper lower bound than the expression on the right side of (1.1). This paper considers a much more general form of the function $F$, and in fact we prove the full Laplace principle (and hence the large deviation principle) for the empirical measures $\{L^n, n \in I\!\!N\}$ when considered as elements of $\mathcal{P}(\bar{I\!\!R}_+)$, where $\mathcal{P}(\bar{I\!\!R}_+)$ is the space of probability measures on the one point compactification of $I\!\!R_+$.

It is important to note that even though we consider our underlying Markov chain to evolve in a compact Polish space ($\bar{I\!\!R}_+$), the usual techniques of proving empirical measure LDP for Markov chains with compact state spaces do not apply. One reason is that if one extends the transition probability function of the Markov chain in (2.3) in the natural fashion by setting $p(\infty, dy) \doteq \delta_{\{\infty\}}(dy)$, i.e., by making the point at $\infty$ an absorbing state, then the resulting Markov chain does not satisfy the typical transitivity conditions that are needed for the proof of the lower bound. The proof of the upper bound for compact state space Markov chains, in essence, only uses the Feller property of the Markov chain. It is easy to see that the extended transition probability function introduced above is Feller with respect to the natural topology on $\bar{I\!\!R}_+$ and so one can use the methodology of [6] to obtain an upper bound. This upper bound will be governed by the function

$$I^*(\nu) = \inf_q \left\{ \int_{\bar{I\!\!R}_+} R(q(x,\cdot) \parallel p(x,\cdot))\nu(dx) \right\}, \quad \nu \in \mathcal{P}(\bar{I\!\!R}_+),$$

where infimum is taken over all transition probability kernels $q$ on $\bar{I\!R}_+ \times \bar{I\!R}_+$ with respect to which $\nu$ is invariant and $R(\cdot \parallel \cdot)$ denotes the relative entropy function (see Section 2 for precise definitions). Suppose that $\nu$ puts some mass on $\infty$ and that $I^*(\nu) < \infty$. Let $q(x,dy)$ achieve the infimum in the definition of $I^*(\nu)$. Then the definition of relative entropy implies $q(\infty, dy) = \delta_{\{\infty\}}(dy)$, and hence there is no contribution from $\infty$:

$$I^*(\nu) = \left\{ \int_{I\!R_+} R(q(x,\cdot) \parallel p(x,\cdot))\nu(dx) \right\}.$$

The rate function $I$ we obtain satisfies $I(\nu) \geq I^*(\nu)$, and $I(\nu) > I^*(\nu)$ if $I(\nu) < \infty$ and $\nu(\{\infty\}) > 0$. Thus the point at $\infty$ makes a contribution to the rate function, and in fact a careful analysis of the manner in which mass tends to infinity is needed to properly account for this contribution.

We now give a brief outline of the paper. In Section 2, we present the basic discrete time model and state the empirical measure large deviation result (Theorem 2.9) for this model. Sections 3, 4 and 5 are devoted to the proof of Theorem 2.9. Section 3 deals with the Laplace principle upper bound. In Section 4 we present some useful properties of the rate function. Section 5 proves the Laplace principle lower bound. Finally, in Section 6 we indicate how to obtain the empirical measure large deviations for the $M/M/1$ queue and reflected Brownian motion via super exponentially close approximations by discrete time Markov chains of the form studied in Sections 3-5. In Section 6, (Remark 6.4) we also state a conjecture on the form of the rate function for the LDP of the empirical measure of a one-dimensional Brownian motion. We end the paper with an Appendix which gives details for some of the proofs that are either standard or similar to others in the paper, and a list of notation is collected there for the reader's convenience.

## 2 The Discrete Time Model

In this section we consider a basic discrete time model and study its empirical measure large deviations. Once the empirical measure LDP for this model is obtained, one can obtain the LDP for empirical measures of a reflected Brownian motion and a $M/M/1$ queue via the standard technique of super exponentially close approximations. This is discussed in greater detail in Section 6.

In order to most easily relate the continuous time models of Section 6 with the discrete time models of this section, it is convenient to work with a very general model here. In particular, we will build the discrete time process with time step $T > 0$ by projecting (via the Skorohod map) random processes defined on the time interval $[0, T]$. The "standard" sort of projected discrete time model is a special case (Example A below).

Let $T \in (0, \infty)$ be fixed and let $\{Z_n, n \in I\!N_0\}$ be a sequence of iid $D([0,T] : I\!R_+)$-valued random variables on the space $(\Omega, \mathcal{F}, I\!P)$. We assume that $Z_0(0) = 0$ with probability one. Let $\theta \in \mathcal{P}(D([0,T] : I\!R_+))$ denote the common probability law. Let $\Gamma : D_+([0,T] : I\!R) \mapsto$

$D([0,T]:\mathbb{R}_+)$ be the Skorohod map, given as

$$\Gamma(z)(t) \doteq z(t) - \left(\inf_{0\le s\le t} z(s)\right) \wedge 0, \quad t\in[0,T], \quad z\in D_+([0,T]:\mathbb{R}_+). \tag{2.1}$$

Elementary calculations show that for $z, z' \in D_+([0,T]:\mathbb{R})$

$$\sup_{0\le t\le T} |\Gamma(z)(t) - \Gamma(z')(t)| \le 2 \sup_{0\le t\le T} |z(t) - z'(t)|. \tag{2.2}$$

We recursively define the $\mathbb{R}_+$-valued constrained random walk $\{X_n, n\in\mathbb{N}_0\}$ by setting

$$X_{n+1} \doteq \Gamma(X_n + Z_n(\cdot))(T), \ n\in\mathbb{N}_0. \tag{2.3}$$

If $X_0$ has probability law $\delta_x$, then we denote expectation by $\mathbb{E}_x$. The transition probability function of the Markov chain $X_n$ is denoted by $p(x,dy)$. For $f\in D([0,T]:\mathbb{R})$, let $||f||_T \doteq \sup_{0\le t\le T}|f(t)|$. Since $T$ will be fixed throughout this section, it is suppressed in the notation, and to further simplify we write $D$ in lieu of $D([0,T]:\mathbb{R})$. We are interested in the large deviation properties of the empirical measure associated with $\{X_n, n\in\mathbb{N}_0\}$.

There are two important special cases. The first corresponds to the standard sort of projected discrete time model, while the second will be used to obtain large deviation results for continuous time models.

**Example A.** Let $\theta$ be the probability law of $Z(\cdot)$, where $Z(\cdot)$ is defined as

$$Z(t) \doteq t\xi, \quad 0\le t\le T$$

and $\xi$ is a $\mathbb{R}$-valued random variable with probability law $\zeta$.

**Example B.** Let $\theta$ be the probability law of $Z(\cdot)$, where $Z(\cdot)$ is defined as

$$Z(t) \doteq bt + \sigma W(t), \quad 0\le t\le T,$$

where $W(\cdot)$ is a standard Brownian motion, and $\sigma, b\in\mathbb{R}$.

The following conditions on $\theta$ will be used in various places.

**Assumption 2.1** *For all* $\alpha\in\mathbb{R}$

$$\int_D \exp(\alpha||y||)\theta(dy) < \infty.$$

Let $p^{(k)}(x,dy)$ denote the $k$-step transition probability function of the Markov chain $X_n$.

**Assumption 2.2** *There exist $l_0, n_0 \in I\!N$ such that for all $x_1, x_2 \in I\!R_+$*

$$\sum_{i=l_0}^{\infty} \frac{1}{2^i} p^{(i)}(x_1, dy) \ll \sum_{j=n_0}^{\infty} \frac{1}{2^j} p^{(j)}(x_2, dy). \tag{2.4}$$

**Remark 2.3** Because $p(x, dy)$ is the transition kernel for a constrained random walk driven by iid noise, one can easily pose conditions on $\theta$ which guarantee Assumption 2.2. For Example A one can assume that $\zeta$ is absolutely continuous with respect to $\lambda$, where $\lambda$ denotes Lebesgue measure on $I\!R$, and there exists $\delta > 0$ such that $\frac{d\theta}{d\lambda}(y) > \delta$ for $y \in (-\delta, \delta)$. For both Example A and Example B, the measures on the left and right side of (2.4) will be of the form $\alpha_i \delta_{\{0\}}(dy) + \beta_i m_i(dy)$, $i = 1, 2$ respectively, where $\alpha_i, \beta_i > 0$ and $m_i$ is mutually absolutely continuous with respect to the Lebesgue measure on $(0, \infty)$. Thus Assumption 2.2 is clearly satisfied. If $\theta$ is supported on $D([0, T] : \mathcal{Z})$, where $\mathcal{Z}$ is the set of integers, as would be the case in a discrete time approximation to a queueing model, then one cannot expect a condition such as Assumption 2.2 to hold for all $x_1$ and $x_2$. For example consider the case where $\theta$ is the probability law of the difference of two independent Poisson processes. In this case if one takes $x_2 = 0$ and $x_1 = \frac{1}{2}$, it is easy to see that (2.4) is not satisfied for any $l_0$ and $n_0$. However, Assumption 2.2 will hold when $x_1$ and $x_2$ are restricted to the integers, and this suffices if one wishes to study the large deviations of the empirical measures when only integer-valued initial conditions are considered. This is discussed further in Section 6.

**Definition 2.4 (Relative Entropy)** *For a complete separable metric space $S$ and for each $\nu \in \mathcal{P}(S)$, the relative entropy $R(\cdot \parallel \nu)$ is a map from $\mathcal{P}(S)$ to $\bar{I\!R}_+$ defined as follows. If $\gamma \in \mathcal{P}(I\!R)$ is absolutely continuous with respect to $\nu$ and if $\log \frac{d\gamma}{d\nu}$ is integrable with respect to $\gamma$, then*

$$R(\gamma \parallel \nu) \doteq \int_S \left( \log \frac{d\gamma}{d\nu} \right) d\gamma.$$

*In all other cases, $R(\gamma \parallel \nu)$ is defined to be $\infty$.*

For $n \in I\!N$, the empirical measure $L^n$ corresponding to the Markov chain $\{X_i, i = 1, ..., n\}$ is the $\mathcal{P}(I\!R_+)$-valued random variable defined by

$$L^n(A) \doteq \frac{1}{n} \sum_{j=0}^{n-1} \delta_{X_j}(A), \; A \in \mathcal{B}(I\!R_+).$$

The empirical measure is also called the (normalized) occupation measure. Let $\bar{I\!R}_+ \doteq I\!R_+ \cup \{\infty\}$, the one point compactification of $I\!R_+$. Then $\bar{I\!R}_+$ and $\mathcal{P}(\bar{I\!R}_+)$ (with the weak convergence topology) are compact Polish spaces. With an abuse of notation, a probability measure $\nu \in \mathcal{P}(I\!R_+)$ will be denoted by $\nu$ even when considered as an element of $\mathcal{P}(\bar{I\!R}_+)$. In this paper we are interested in the large deviation properties for the family $\{L^n, n \in I\!N\}$ of random variables with values in the compact Polish space $\mathcal{P}(\bar{I\!R}_+)$. To introduce the rate function that will govern the large deviation probabilities for this family, we need the following notation and definition.

**Definition 2.5 (Stochastic Kernel)** *Let $(\mathcal{V},\mathcal{A})$ be a measurable space and $S$ a Polish space. Let $\tau(dy|x)$ be a family of probability measures on $(S,\mathcal{B}(S))$ parameterized by $x\in\mathcal{V}$. We call $\tau(dy|x)$ a stochastic kernel on $S$ given $\mathcal{V}$ if for every Borel subset $E$ of $S$ the map $x\in\mathcal{V}\mapsto\tau(E|x)$ is measurable. The class of all such stochastic kernels is denoted by $\mathcal{S}(S|\mathcal{V})$.*

If $p^*\in\mathcal{S}(I\!R_+|I\!R_+)$, then $p^*$ is a probability transition function and we will write $p^*(dy|x)$ as $p^*(x,dy)$. We say that a probability measure $\nu\in\mathcal{P}(I\!R_+)$ is invariant with respect to a stochastic kernel $p^*\in\mathcal{S}(I\!R_+|I\!R_+)$, if $\nu(A)=\int_{I\!R_+}p^*(x,A)\nu(dx)$ for all $A\in\mathcal{B}(I\!R_+)$. We will also refer to $\nu$ as a $p^*$-invariant probability measure.

The rate function associated with the family $\{L^n,n\in I\!N\}$ can now be defined. Given $q^*\in\mathcal{S}(D|I\!R_+)$, we associate the stochastic kernel $p^*\in\mathcal{S}(I\!R_+|I\!R_+)$ which is consistent with $q^*$ under the constraint mechanism by setting

$$p^*(x,A)\doteq\int_D 1_{\{\Pi_T(x,z)\in A\}}q^*(dz|x),\quad A\in\mathcal{B}(I\!R_+),\quad x\in I\!R_+, \tag{2.5}$$

where

$$\Pi_T(x,z)\doteq\Gamma(x+z(\cdot))(T),\quad x\in I\!R_+,\quad z\in D.$$

As before, $T$ will be suppressed in the notation.

Note that if a $D$-valued random variable $Z^*$ has the probability law $q^*(dz|x)$ then $p^*(x,dy)$ defined above is the probability law of $\Pi(x,Z^*)$. Thus if one considers the sequence $(X_j^*,Z_j^*)$ of $I\!R_+\times D$-valued random variables defined recursively as

$$\begin{cases} P(Z_j^*\in dz\mid (X_k^*,Z_{k-1}^*),1\le k\le j)\doteq q^*(dz|X_j^*)\\ X_{j+1}^*\doteq\Gamma(X_j^*+Z_j^*(\cdot))(T),\end{cases}$$

then $\{X_j^*\}$ is a Markov chain with transition probability kernel $p^*(x,dy)$.

For $\nu\in\mathcal{P}(I\!R_+)$, let

$$I_1(\nu)\doteq\inf_{\{q^*\in\mathcal{A}_1(\nu)\}}\int_{I\!R_+}R(q^*(\cdot|x)\parallel\theta(\cdot))\nu(dx), \tag{2.6}$$

where $\mathcal{A}_1(\nu)$ is the collection of all $q^*$ for which $\nu$ is $p^*$-invariant.

**Remark 2.6** Let $\nu$ be a probability measure on a complete, separable metric space $S$. It is well known that there is a one-to-one correspondence between probability transition kernels $q$ for which $\nu$ is an invariant distribution, and probability measures $\tau$ on $S\times S$ such that the first and second marginals of $\tau$ equal $\nu$. Indeed, $\nu(dx)q(x,dy)$ is such a measure on $S\times S$, and given any such $\tau$, the "conditional" decomposition $\tau(dx\,dy)=\nu(dx)r(x,dy)$ for some probability transition kernel $r$ and $\int_S\nu(dx)r(x,dy)=\nu(dy)$ imply that $\nu$ is $r$-invariant. We will need an analogous correspondence in the present setting. More precisely, we claim that for any $\nu \in \mathcal{P}(I\!R_+)$, there is a one-to-one correspondence between $q^* \in \mathcal{S}(I\!R_+|D)$ for which $\nu$ is $p^*$-invariant, where $p^*$ is defined via (2.5), and $\tau \in \mathcal{P}(I\!R_+ \times D)$ such that

$$\langle f, \nu \rangle = \langle f, (\tau)_1 \rangle = \int_{I\!R_+ \times D} f(\Pi(x,z))\tau(dx\, dz), \quad \text{for all } f \in C_b(I\!R_+), \tag{2.7}$$

where $(\tau)_1 \in \mathcal{P}(I\!R_+)$ is defined as $(\tau)_1(dx) \doteq \tau(dx \times D)$. Define by $\mathcal{A}(\nu)$ the class of all $\tau \in \mathcal{P}(I\!R_+ \times D)$ such that (2.7) holds and let $\mathcal{A}_1(\nu)$ be as defined below (2.6). The one to one correspondence between $\mathcal{A}(\nu)$ and $\mathcal{A}_1(\nu)$ can be described by the relations that if $\tau \in \mathcal{A}(\nu)$ and $\tau$ is decomposed as $\tau(dx\, dz) = (\tau)_1(dx)q(dz|x)$, then $(\tau)_1 = \nu$ and $q \in \mathcal{A}_1(\nu)$. Conversely, if $q \in \mathcal{A}_1(\nu)$ and $\tau \in \mathcal{P}(I\!R_+ \times D)$ is defined as $\tau(dx\, dz) \doteq \nu(dx)q(dz|x)$, then $\tau \in \mathcal{A}(\nu)$. The above equivalence will be used in the proofs of Theorem 4.1 and Lemma 5.4.

Define by $\mathcal{P}_{\text{tr}}(D)$ the class of all $\sigma \in \mathcal{P}(D)$ for which $\int_D ||z||\sigma(dz) < \infty$ and $\int_D z(T)\sigma(dz) \geq 0$. Here the subscript tr stands for "transient." Strictly speaking, this class includes measures that produce a null recurrent process, but transient is more suggestive of what is intended. Let

$$J \doteq \inf_{\sigma \in \mathcal{P}_{\text{tr}}(D)} R(\sigma \parallel \theta). \tag{2.8}$$

The following mild assumption will be used in the proof of Theorems 4.1 and 5.1. It essentially asserts that there is some positive probability of moving both to the left and the right under $\theta$.

**Assumption 2.7** *There exist $\theta_0, \theta_1 \in \mathcal{P}(D)$ such that*

$$\int_D z(T)\theta_0(dz) < 0, \quad \int_D z(T)\theta_1(dz) > 0$$

*and for $i = 0, 1$, the following hold: (a) $\int_D ||z||\theta_i(dz) < \infty$, (b) $\theta$ and $\theta_i$ are mutually absolutely continuous and (c) $R(\theta_i \parallel \theta) < \infty$.*

Note that under Assumption 2.7, $J < \infty$.

**Remark 2.8** For Example A in Remark 2.3, Assumption 2.2 implies that $\zeta(0,\infty) > 0$ and $\zeta(-\infty, 0) > 0$. Using this and Assumption 2.1, one can easily show that Assumption 2.7 holds for Example A. Furthermore, Assumption 2.7 clearly holds for Example B.

For $\nu \in \mathcal{P}(\bar{I\!R}_+)$, let $\hat{\nu} \in \mathcal{P}(I\!R_+)$ be defined as follows. If $\nu(I\!R_+) \neq 0$, then for $A \in \mathcal{B}(I\!R_+)$

$$\hat{\nu}(A) \doteq \frac{\nu(A \cap I\!R_+)}{\nu(I\!R_+)}. \tag{2.9}$$

Otherwise, $\hat{\nu}$ can be taken to be an arbitrary element of $\mathcal{P}(I\!R_+)$. We define the rate function as follows. For $\nu \in \mathcal{P}(\bar{I\!R}_+)$

$$I(\nu) \doteq \nu(I\!R_+)I_1(\hat{\nu}) + (1 - \nu(I\!R_+))J. \tag{2.10}$$

Our main result is the following.

**Theorem 2.9** *Suppose that Assumptions 2.1, 2.2 and 2.7 hold. Then for all $F \in C_b(\mathcal{P}(\bar{I\!\!R}_+))$ and all $x \in I\!\!R_+$*

$$\lim_{n\to\infty} \frac{1}{n} \log I\!\!E_x \left[\exp(-nF(L^n))\right] = - \inf_{\mu \in \mathcal{P}(\bar{I\!\!R}_+)} \{I(\mu) + F(\mu)\}.$$

*Furthermore, $I(\cdot)$ is a rate function on $\mathcal{P}(\bar{I\!\!R}_+)$.*

**Proof:** The fact that the left hand side in the last display is at most the expression in the right hand side (Laplace principle upper bound) is proved in Theorem 3.1. The reverse inequality (Laplace principle lower bound) is proved in Theorem 5.1. The proof that $I(\cdot)$ is a rate function on $\mathcal{P}(\bar{I\!\!R}_+)$ is given in Theorem 4.1(c). ∎

**Remark 2.10** Theorem 2.9 can be summarized by the statement that the family $\{L^n, n \in I\!\!N\}$ satisfies the Laplace principle on $\mathcal{P}(\bar{I\!\!R}_+)$ with the rate function $I(\cdot)$. From Theorem 1.2.3 of [9] it follows that the family $\{L^n, n \in I\!\!N\}$ satisfies the large deviation principle on $\mathcal{P}(\bar{I\!\!R}_+)$ with the rate function $I(\cdot)$.

**Remark 2.11** The convergence in Theorem 2.9 is in fact uniform for $x$ in compact subsets of $I\!\!R_+$. See [9, Section 8.4]

We now present an important corollary of Theorem 2.9. Denote by $\mathcal{S}_0$ the subclass of $C_b(I\!\!R_+)$ consisting of functions $f$ for which $f(x)$ converges as $x \to \infty$, with the limit denoted by $f^\infty$. Such an $f$ can be extended to a function $\bar{f}$ on $\bar{I\!\!R}_+$ by defining $\bar{f}(\infty) \doteq f^\infty$. It is easy to see that there is a one to one correspondence between $\mathcal{S}_0$ and $C_b(\bar{I\!\!R}_+)$, given by $f \in \mathcal{S}_0 \mapsto \bar{f} \ni C_b(\bar{I\!\!R}_+)$ and $g \in C_b(\bar{I\!\!R}_+) \mapsto g|_{I\!\!R_+} \ni \mathcal{S}_0$.

**Corollary 2.12** *Let $F : \mathcal{P}(I\!\!R_+) \mapsto I\!\!R$ be given by*

$$F(\mu) \doteq G(\langle f_1, \mu\rangle, \ldots, \langle f_k, \mu\rangle),$$

*where $G \in C_b(I\!\!R^k)$ and $f_i \in \mathcal{S}_0$, $i = 1, \ldots, k$. Then for all $x \in I\!\!R_+$*

$$\lim_{n\to\infty} \frac{1}{n} \log I\!\!E_x \left[\exp(-nF(L^n))\right] = - \inf_{\nu \in \mathcal{P}(I\!\!R_+),\ \rho \in [0,1]} \{\rho I_1(\nu) + (1-\rho)J + \bar{F}(\rho, \nu)\},$$

*where $\bar{F} \in C_b([0,1] \times \mathcal{P}(I\!\!R_+))$ is defined by*

$$\bar{F}(\rho, \nu) \doteq G(\rho\langle f_1, \nu\rangle + (1-\rho) f_1^\infty, \ldots, \rho\langle f_k, \nu\rangle + (1-\rho) f_k^\infty).$$

The corollary is an immediate consequence of Theorem 2.9. If we set $\rho \doteq \hat{\nu}(I\!\!R_+)$, then for $f \in \mathcal{S}_0$ and $\nu \in \mathcal{P}(\bar{I\!\!R}_+)$

$$\langle f, \nu\rangle = \rho\langle f, \hat{\nu}\rangle + (1-\rho) f^\infty.$$

Thus the unique continuous extension of $F$ to $\mathcal{P}(\bar{I\!\!R}_+)$ equals

$$\bar{F}(\rho, \hat{\nu}),$$

and the corollary follows.

## 3 Laplace Principle Upper Bound

The main result of this section is the following.

**Theorem 3.1** *Suppose that Assumption 2.1 holds, and define $I$ by (2.10). Then for all $F \in C_b(\mathcal{P}(\bar{I\!R}_+))$ and all $x \in I\!R_+$*

$$\limsup_{n\to\infty} \frac{1}{n} \log I\!E_x \left[\exp(-nF(L^n))\right] \le - \inf_{\mu \in \mathcal{P}(\bar{I\!R}_+)} \{I(\mu) + F(\mu)\}. \tag{3.1}$$

Throughout this paper there will be many constructions and results that are analogous to those used in [9, Chapter 8] to study empirical measures under the classical strong stability assumption. While there are some differences in these constructions, in an effort to streamline the presentation we will emphasize those parts of the analysis that are new.

Our first step in the proof will be to give a variational representation for the prelimit expression on the left side of (3.1). This representation will take the form of the value function for a controlled Markov chain with an appropriate cost function. The representation will also be used in the proof of the lower bound in Section 5. We begin with the construction of a controlled Markov chain. It follows a similar construction in Chapter 4 of [9] and thus some details are omitted. In particular, the proof of Lemma 3.2 is not provided. We recall that a table of notation is provided at the end of the paper.

For $n \in I\!N$ and $j = 0, \ldots, n$, let $\nu_j^n(dy|x,\gamma)$ be a stochastic kernel in $\mathcal{S}(D|I\!R_+ \times \mathcal{M}(I\!R_+))$. For $n \in I\!N$ and $x \in I\!R_+$, define a controlled sequence of $I\!R_+ \times \mathcal{M}(I\!R_+) \times D$-valued random variables $\{(\bar{X}_j^n, \bar{L}_j^n, \bar{Z}_j^n), j = 0, \ldots, n\}$ on some probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{I\!P}_x)$ as follows. Set $\bar{X}_0^n \doteq x$ and $\bar{L}_0^n \doteq 0$, and then for $k = 0, \ldots, n-1$ recursively define

$$\begin{aligned} \bar{I\!P}_x(\bar{Z}_k^n \in dy | (\bar{X}_j^n, \bar{L}_j^n), j = 0, \ldots, k) &\doteq \nu_k^n(dy|\bar{X}_k^n, \bar{L}_k^n) \\ \bar{X}_{k+1}^n &\doteq \Pi(\bar{X}_k^n, \bar{Z}_k^n) \\ \bar{L}_{k+1}^n &\doteq \bar{L}_k^n + \frac{1}{n}\delta_{\bar{X}_k^n}. \end{aligned} \tag{3.2}$$

Denote $\bar{L}_n^n$ by $\bar{L}^n$. We now give the variational representation for

$$W^n(x) \doteq -\frac{1}{n} \log I\!E_x \left[\exp(-nF(L^n))\right] \tag{3.3}$$

in terms of the controlled sequences introduced above.

**Lemma 3.2** *Fix $F \in C_b(\mathcal{P}(\bar{I\!R}_+))$ and let $W^n(x)$ be defined via (3.3). Then for all $n \in I\!N$ and $x \in I\!R_+$,*

$$W^n(x) = \inf_{\{\nu_j^n\}} \bar{I\!E}_x \left[ \frac{1}{n} \sum_{j=0}^{n-1} R\left(\nu_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)\right) + F(\bar{L}^n) \right], \tag{3.4}$$

*where the infimum is taken over all possible sequences of stochastic kernels $\{\nu_j^n, j = 0, \ldots, n\}$ in $\mathcal{S}(D|I\!\!R_+ \times \mathcal{M}(I\!\!R_+))$.*

The proof of this lemma is similar to the proof of Theorem 4.2.2 of [9]. The main difference between the two results is that the latter gives a representation which involves the transition probability function, $p(x, dy)$, of the Markov chain rather than the probability law, $\theta$, of the noise sequence. In the representation, the original empirical measures $L^n$ are replaced by $\bar{L}^n$, which are the empirical measures for the process generated by the stochastic kernels $\nu_j^n$. We pay a cost of relative entropy for "twisting" the distribution from $\theta$ to $\nu_j^n$, plus a cost of $\bar{I\!\!E}_x F(\bar{L}^n)$ that depends on where the controlled empirical measure ends up. The representation exhibits $W^n(x)$ as the minimum expected total cost.

Let $\epsilon > 0$ be arbitrary. In view of the preceding lemma, for each $n \in I\!\!N$ we can find a sequence of "$\epsilon$-optimal" stochastic kernels $\{\bar{\nu}_j^n, j = 0, \ldots, n\}$, such that

$$W^n(x) + \epsilon \geq \bar{I\!\!E}_x \left[ \frac{1}{n} \sum_{j=0}^{n-1} R(\bar{\nu}_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)) + F(\bar{L}^n) \right]. \tag{3.5}$$

Since $F$ is bounded, both $|F(\bar{L}^n)|$ and $|W^n(x)|$ are bounded above by $\|F\|_\infty$. Thus

$$\Delta \doteq \sup_n \bar{I\!\!E}_x \left[ \frac{1}{n} \sum_{j=0}^{n-1} R(\bar{\nu}_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)) \right] < \infty. \tag{3.6}$$

Define $\bar{\nu}^n \in \mathcal{P}(\bar{I\!\!R}_+ \times D^2)$ as follows. For $A \in \mathcal{B}(\bar{I\!\!R}_+)$ and $B, C \in \mathcal{B}(D)$,

$$\bar{\nu}^n(A \times B \times C) \doteq \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\bar{X}_j^n}(A) \bar{\nu}_j^n(B|\bar{X}_j^n, \bar{L}_j^n) \delta_{\bar{Z}_j^n}(C). \tag{3.7}$$

**Lemma 3.3** *Suppose Assumption 2.1 holds. Then $\{\bar{\nu}^n, n \in I\!\!N\}$, defined on the probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{I\!\!P}_x)$, is a tight family of $\mathcal{P}(\bar{I\!\!R}_+ \times D^2)$-valued random variables.*

The result is a consequence of (3.6) and follows along the lines of the proof of Proposition 8.2.5 of [9]. We remark that unlike in the cited Proposition, we do not need to make any stability assumption on the underlying Markov chain. This is one of the key advantages of working with the representation in Lemma 3.2 given in terms of the probability law of the noise sequence ($\theta$) rather than the transition probability function of the Markov chain ($p(x, dy)$). For sake of completeness the proof is included in the Appendix.

Take a convergent subsequence of $\bar{\nu}^n$ and denote the limit point by $\bar{\nu}$. To simplify the notation, we retain $n$ to denote this convergent subsequence, so that

$$\bar{\nu}^n \Rightarrow \bar{\nu} \quad \text{as} \quad n \to \infty. \tag{3.8}$$

*For the rest of this section, we will assume that Assumption 2.1 is satisfied, and that (3.8) holds.*

For $\gamma \in \mathcal{P}(\bar{\mathbb{R}}_+ \times D^2)$ and $1 \le i < j \le 3$, denote by $(\gamma)_i$ and $(\gamma)_{i,j}$ the $i$-th marginal and the $(i,j)$-th marginal of $\gamma$, respectively. For example, $(\gamma)_{2,3}$ is the element of $\mathcal{P}(D \times D)$ defined as

$$(\gamma)_{2,3}(A \times B) \doteq \gamma(\bar{\mathbb{R}}_+ \times A \times B), \quad A, B \in \mathcal{B}(D).$$

**Lemma 3.4** *Let $\bar{\nu}$ be as in (3.8). Then*

$$(\bar{\nu})_{1,2} = (\bar{\nu})_{1,3}, \quad a.s.$$

**Proof:** It suffices to show that for all $g \in C_b(\bar{\mathbb{R}}_+)$ and $h \in C_b(D)$,

$$\left| \int_{\bar{\mathbb{R}}_+ \times D} g(x)h(y)(\bar{\nu}^n)_{1,2}(dx\, dy) - \int_{\bar{\mathbb{R}}_+ \times D} g(x)h(y)(\bar{\nu}^n)_{1,3}(dx\, dy) \right| \tag{3.9}$$

converges to 0, as $n \to \infty$, in probability. Observe that

$$\int_{\bar{\mathbb{R}}_+ \times D} g(x)h(y)(\bar{\nu}^n)_{1,2}(dx\, dy) = \frac{1}{n} \sum_{j=0}^{n-1} g(\bar{X}^n_j) \int_D h(y) \bar{\nu}^n_j (dy | \bar{X}^n_j, \bar{L}^n_j)$$

and

$$\int_{\bar{\mathbb{R}}_+ \times D} g(x)h(y)(\bar{\nu}^n)_{1,3}(dx\, dy) = \frac{1}{n} \sum_{j=0}^{n-1} g(\bar{X}^n_j) h(\bar{Z}^n_j).$$

Thus the expression in (3.9) can be rewritten as

$$\Lambda \doteq \left| \frac{1}{n} \sum_{j=0}^{n-1} g(\bar{X}^n_j) \left( h(\bar{Z}^n_j) - \int_D h(y) \bar{\nu}^n_j (dy | \bar{X}^n_j, \bar{L}^n_j) \right) \right|.$$

From (3.2) we have that the conditional distribution of $\bar{Z}^n_k$ given $\{\bar{Z}^n_j, \bar{X}^n_{j+1},\ j = 0, 1, \ldots, k-1\}$ is $\nu^n_k(dy|\bar{X}^n_k, \bar{L}^n_k)$. Therefore, for $0 \le j < k \le n-1$ and an arbitrary real bounded and measurable function $\psi$,

$$\bar{\mathbb{E}}_x \left[ \psi(\bar{X}^n_j, \bar{L}^n_j, \bar{Z}^n_j, \bar{X}^n_k, \bar{L}^n_k) \left( h(\bar{Z}^n_k) - \int_D h(y) \bar{\nu}^n_k (dy | \bar{X}^n_k, \bar{L}^n_k) \right) \right] = 0.$$

This implies that $\bar{\mathbb{E}}_x[\Lambda^2]$ is $O(1/n)$ and thus the expression in (3.9) converges to 0 in probability, as $n \to \infty$. This proves the lemma. ∎

**Lemma 3.5** *Let $\{\bar{\nu}^n\}$ and $\bar{\nu}$ be as in (3.8). Then*

$$\lim_{C \to \infty} \sup_n \bar{\mathbb{E}}_x \int_{||z|| > C} ||z||\ (\bar{\nu}^n)_3(dz) = 0, \tag{3.10}$$

$$\bar{\mathbb{E}}_x \int_D ||z|| (\bar{\nu})_3(dz) < \infty. \tag{3.11}$$

**Proof:** The proof follows along the lines of Lemma 5.3.2 and Theorem 5.3.5 of [9], however for the sake of completeness we provide the details. We begin by showing that (3.11) follows easily, once (3.10) is proven. So now suppose that (3.10) holds. In order to show (3.11), it is enough to show that

$$\limsup_{C\to\infty} \bar{I\!\!E}_x \int_{||z||>C} ||z||(\bar{\nu})_3(dz) = 0. \tag{3.12}$$

Recall that $(\bar{\nu}^n)_3$ converges in distribution to $(\bar{\nu})_3$. By the Skorohod representation theorem, we can assume that $(\bar{\nu}^n)_3$ converges almost surely to $(\bar{\nu})_3$. Then by the lower semi-continuity of the map $\mathcal{P}(D) \ni \gamma \mapsto \int_{||z||>C} ||z||\gamma(dz) \in \bar{I\!\!R}_+$ we have that

$$\begin{aligned}
\bar{I\!\!E}_x \int_{||z||>C} ||z||(\bar{\nu})_3(dz) &\leq \bar{I\!\!E}_x \left[\liminf_{n\to\infty} \int_{||z||>C} ||z||(\bar{\nu}^n)_3(dz)\right] \\
&\leq \liminf_{n\to\infty} \bar{I\!\!E}_x \left[\int_{||z||>C} ||z||(\bar{\nu}^n)_3(dz)\right] \\
&\leq \sup_n \bar{I\!\!E}_x \left[\int_{||z||>C} ||z||(\bar{\nu}^n)_3(dz)\right],
\end{aligned}$$

where the second inequality above follows from Fatou's lemma. This proves (3.12) and hence (3.11).

Thus in order to complete the proof of the lemma we need to prove (3.10). We begin by observing that

$$\begin{aligned}
\bar{I\!\!E}_x \int_{||z||>C} ||z||(\bar{\nu}^n)_3(dz) &= \bar{I\!\!E}_x \left[\frac{1}{n}\sum_{j=0}^{n-1} ||\bar{Z}_j^n||\, 1_{||\bar{Z}_j^n||>C}\right] \\
&= \bar{I\!\!E}_x \left[\frac{1}{n}\sum_{j=0}^{n-1} \bar{I\!\!E}_x \left[||\bar{Z}_j^n||\, 1_{||\bar{Z}_j^n||>C} | \bar{X}_j^n, \bar{L}_j^n\right]\right] \\
&= \bar{I\!\!E}_x \left[\frac{1}{n}\sum_{j=0}^{n-1} \int_{||z||>C} ||z||\, \bar{\nu}_j^n(dz|\bar{X}_j^n, \bar{L}_j^n)\right].
\end{aligned} \tag{3.13}$$

From (3.6) we have that for each $j \in \{0, \ldots, n-1\}$, $R(\bar{\nu}_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)) < \infty$, a.s. $[\bar{P}_x]$. Let $f_j^n$ be a measurable map from $\bar{\Omega} \times D \to I\!\!R_+$ such that

$$f_j^n(\bar{\omega}, z) = \frac{d\bar{\nu}_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n)(\bar{\omega})}{d\theta(\cdot)}(z), \quad a.s.\ [\bar{P}_x \otimes \theta].$$

Henceforth we will suppress $\bar{\omega}$ in the notation when writing $f_j^n(\bar{\omega}, z)$. Thus

$$\begin{aligned}
&\bar{I\!\!E}_x \left[\int_{||z||>C} ||z||\bar{\nu}_j^n(dz \mid \bar{X}_j^n, \bar{L}_j^n)\right] \\
&= \bar{I\!\!E}_x \left[\int_{||z||>C} ||z||f_j^n(z)d\theta\right] \\
&\leq \int_{||z||>C} e^{\alpha||z||}\theta(dz) + \frac{1}{\alpha}\bar{I\!\!E}_x \left[\int_D \left(f_j^n(z)\log(f_j^n(z)) - f_j^n(z) + 1\right)\theta(dz)\right],
\end{aligned}$$

where the inequality above follows from the well known inequality

$$ab \le e^{\alpha a} + \frac{1}{\alpha}(b \log b - b + 1); \quad a, b \in [0, \infty) \ \ \alpha \in (0, \infty).$$

Next observing that

$$\int_D f_j^n(z) \log(f_j^n(z)) \theta(dz) = R(\bar{\nu}_j^n(\cdot \mid \bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot))$$

and

$$\int_D f_j^n(z) \theta(dz) = \bar{\nu}_j^n(D \mid \bar{X}_j^n, \bar{L}_j^n) = 1,$$

we have that

$$\begin{aligned}
&\limsup_{C\to\infty} \sup_n \frac{1}{n} \sum_{j=0}^{n-1} \bar{E}_x \left[ \int_{||z||>C} ||z|| \bar{\nu}_j^n(dz \mid \bar{X}_j^n, \bar{L}_j^n) \right] \\
&\le \limsup_{C\to\infty} \sup_n \left( \int_{||z||>C} e^{\alpha ||z||} \theta(dz) + \frac{1}{\alpha} \bar{E}_x \left[ \frac{1}{n} \sum_{j=0}^{n-1} R(\bar{\nu}_j^n(\cdot \mid \bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)) \right] \right) \\
&\le \limsup_{C\to\infty} \int_{||z||>C} e^{\alpha ||z||} \theta(dz) + \frac{\Delta}{\alpha} \\
&\le \frac{\Delta}{\alpha},
\end{aligned}$$

where the next to last inequality follows from (3.6) and the last inequality follows from Assumption 2.1. Finally, (3.10) follows on taking limit as $\alpha \to \infty$ in the above inequality. ∎

One of the key steps in the proof of the upper bound is to show that $\int_{\bar{\mathbb{R}}_+ \times D} z(T) 1_{\{\infty\}}(x)(\bar{\nu})_{1,3}(dx\, dz) \ge 0$. Heuristically, this says that the drift of the controlled chain is non-negative when the point at $\infty$ is charged and the chain is far from the origin. To prove this, the following test function will be useful.

For $c \in (0, \infty)$, define a real continuously differentiable function $F_c$ on $\mathbb{R}_+$ by

$$F_c(x) \doteq \begin{cases} \dfrac{1}{2} & \text{if } x \in [0, c] \\ \dfrac{x^2}{2c^2} - \dfrac{x}{c} + 1 & \text{if } x \in (c, c^2 + c] \\ x - \dfrac{c^2}{2} - c + \dfrac{1}{2} & \text{if } x \in (c^2 + c, \infty). \end{cases} \tag{3.14}$$

The function $F_c$ has the property that $F_c' \in \mathcal{S}_0$ and $|F_c'(x)| \le 1$. Retaining $F_c'$ to denote the continuous extension of $F_c'$ to $\bar{\mathbb{R}}_+$, we have that $F_c'(x) \to \delta_{\{\infty\}}(x)$ for all $x \in \bar{\mathbb{R}}_+$ as $c \to \infty$.

We now present an elementary lemma concerning the function $F_c$.

**Lemma 3.6** *For all $x, y \in \mathbb{R}_+$*

$$F_c(y) - F_c(x) = (y - x) F_c'(x) + R_c(x, y),$$

*where $F_c'$ denotes the derivative of $F_c$ and the remainder $R_c(x,y)$ satisfies the inequality*

$$R_c(x,y) \leq \frac{|y-x|}{c} + |y-x| 1_{|y-x| \geq c}. \tag{3.15}$$

**Proof:** We will only prove the result for the case when $x \leq y$. The result for $y \leq x$ follows in a similar fashion. Note that the derivative $F_c'$ is given as

$$F_c'(x) = \begin{cases} 0 & \text{if } x \in [0,c] \\ \frac{x}{c^2} - \frac{1}{c} & \text{if } x \in (c, c^2+c] \\ 1 & \text{if } x \in (c^2+c, \infty). \end{cases}$$

Since $F_c'(x)$ is a continuous function, we have

$$\begin{aligned} F_c(y) - F_c(x) &= \int_x^y F_c'(u) du \\ &= \int_x^y (F_c'(u) - F_c'(x)) du + (y-x) F_c'(x). \end{aligned}$$

Now define

$$R_c(x,y) \doteq \int_x^y (F_c'(u) - F_c'(x)) du.$$

Since $F_c'$ is an increasing function bounded above by 1,

$$R_c(x,y) \leq \left( \int_x^y (F_c'(u) - F_c'(x)) du \right) 1_{|y-x| \leq c} + |y-x| 1_{|y-x| \geq c}. \tag{3.16}$$

Now let $x, y$ be such that $|y-x| \leq c$. Then

$$\begin{aligned} \int_x^y (F_c'(u) - F_c'(x)) du &\leq (y-x)(F_c'(y) - F_c'(x)) \\ &\leq \frac{(y-x)^2}{c^2} \\ &\leq \frac{|y-x|}{c}, \end{aligned}$$

where the first inequality follows on noting that $F_c'(u) - F_c'(x) < F_c'(y) - F_c'(x)$ for all $u \leq y$, the second inequality is a consequence of the fact that for all $0 \leq x < y < \infty$, $F_c'(y) - F_c'(x) \leq (y-x)/c^2$ and the final inequality is obtained on using that $|y-x| \leq c$. Using the last inequality in (3.16), we have (3.15). This proves the lemma. ∎

**Lemma 3.7** *For $c \in (0, \infty)$, let $F_c$ be given via (3.14). For $n \in \mathbb{N}$, let $\bar{\nu}^n$ be defined via (3.7). Then*

$$\begin{aligned} \int_{\bar{\mathbb{R}}_+ \times D} y(T) F_c'(x) (\bar{\nu}^n)_{1,3}(dx\, dy) &\geq -5 \int_D ||y|| 1_{||y|| \geq c/2} (\bar{\nu}^n)_3(dy) \\ &\quad - \frac{2}{c} \int_D ||y|| (\bar{\nu}^n)_3(dy) - \frac{F_c(\bar{X}_0^n)}{n}. \end{aligned} \tag{3.17}$$

**Proof:** For $j \in \{0, \ldots, n-1\}$, define $\xi_j^n \doteq \bar{X}_{j+1}^n - \bar{X}_j^n$. We begin by observing that

$$
\begin{aligned}
&\int_{\bar{I\!R}_+ \times D} y(T) F_c'(x) (\bar{\nu}^n)_{1,3}(dx\, dy) \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \bar{Z}_i^n(T) F_c'(\bar{X}_i^n) \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \bar{Z}_i^n(T) F_c'(\bar{X}_i^n) 1_{||\bar{Z}_i^n|| \geq c} + \frac{1}{n} \sum_{i=0}^{n-1} \bar{Z}_i^n(T) F_c'(\bar{X}_i^n) 1_{||\bar{Z}_i^n|| < c} \\
&\geq -\frac{1}{n} \sum_{i=0}^{n-1} ||\bar{Z}_i^n|| 1_{||\bar{Z}_i^n|| \geq c} + \frac{1}{n} \sum_{i=0}^{n-1} \bar{Z}_i^n(T) F_c'(\bar{X}_i^n) 1_{||\bar{Z}_i^n|| < c} \\
&\geq -\frac{1}{n} \sum_{i=0}^{n-1} ||\bar{Z}_i^n|| 1_{||\bar{Z}_i^n|| \geq c} + \frac{1}{n} \sum_{i=0}^{n-1} \xi_i^n F_c'(\bar{X}_i^n) 1_{||\bar{Z}_i^n|| < c} \\
&\geq -3 \frac{1}{n} \sum_{i=0}^{n-1} ||\bar{Z}_i^n|| 1_{||\bar{Z}_i^n|| \geq c} + \frac{1}{n} \sum_{i=0}^{n-1} \xi_i^n F_c'(\bar{X}_i^n).
\end{aligned}
\tag{3.18}
$$

The first inequality follows on recalling that $F_c'$ is non-negative and bounded above by 1. The second equality is a consequence of the fact that $F_c'(x)$ equals 0 for $x \leq c$ and that on the set $\{\bar{X}_i^n > c\} \cap \{||\bar{Z}_i^n|| \leq c\}$, $\bar{X}_i^n$ is far enough from the origin that $\xi_i^n$ equals $\bar{Z}_i^n(T)$. The third inequality is a consequence of the fact that in all cases $|\xi_i^n| \leq 2||\bar{Z}_i^n||$, and that $F_c'(x) \in [0, 1]$. Next observe that, from Lemma 3.6

$$
\begin{aligned}
0 &\leq \frac{1}{n} F_c(\bar{X}_n^n) \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \left( F_c(\bar{X}_{i+1}^n) - F_c(\bar{X}_i^n) \right) + \frac{F_c(\bar{X}_0^n)}{n} \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \xi_i^n F_c'(\bar{X}_i^n) + \frac{1}{n} \sum_{i=0}^{n-1} R_c(\bar{X}_i^n, \bar{X}_{i+1}^n) + \frac{F_c(\bar{X}_0^n)}{n}.
\end{aligned}
\tag{3.19}
$$

Also, note that

$$
R_c(\bar{X}_i^n, \bar{X}_{i+1}^n) \leq \frac{|\xi_i^n|}{c} + |\xi_i^n| 1_{|\xi_i^n| \geq c}.
$$

Combining the last two observations we have

$$
\begin{aligned}
\frac{1}{n} \sum_{i=0}^{n-1} \xi_i^n F_c'(\bar{X}_i^n) &\geq -\frac{1}{n} \sum_{i=0}^{n-1} |\xi_i^n| 1_{|\xi_i^n| \geq c} - \frac{1}{c} \frac{1}{n} \sum_{i=0}^{n-1} |\xi_i^n| - \frac{F_c(\bar{X}_0^n)}{n} \\
&\geq -\frac{2}{n} \sum_{i=0}^{n-1} ||Z_i^n|| 1_{||Z_i^n|| \geq \frac{c}{2}} - \frac{2}{c} \frac{1}{n} \sum_{i=0}^{n-1} ||Z_i^n|| - \frac{F_c(\bar{X}_0^n)}{n}.
\end{aligned}
\tag{3.20}
$$

Finally, substituting (3.20) in (3.18) gives

$$
\int_{\bar{I\!R}_+ \times D} y(T) F_c'(x) (\bar{\nu}^n)_{1,3}(dx\, dy)
$$

$$\begin{aligned}
&\geq -5\frac{1}{n}\sum_{i=0}^{n-1}||\bar{Z}_i^n||1_{||\bar{Z}_i^n||\geq\frac{c}{2}} - \frac{2}{c}\frac{1}{n}\sum_{i=0}^{n-1}||\bar{Z}_i^n|| - \frac{F_c(\bar{X}_0^n)}{n} \\
&= -5\int_D ||y||1_{||y||\geq\frac{c}{2}}(\bar{\nu}^n)_3(dy) - \frac{2}{c}\int_D ||y||(\bar{\nu}^n)_3(dy) - \frac{F_c(\bar{X}_0^n)}{n}.
\end{aligned}$$

This proves the lemma. ∎

**Lemma 3.8** *Let $\bar{\nu}$ be as in (3.8). For $c \in (0,\infty)$, let $F_c$ be defined via (3.14). Let $\{c_k\}_{k\in\mathbb{N}}$ be a sequence in $(0,\infty)$ such that $c_k \to \infty$ as $k \to \infty$, and for all $k \in \mathbb{N}$*

$$\bar{\mathbb{E}}_x\left[(\bar{\nu})_3\left\{z(\cdot) : ||z|| = \frac{c_k}{2}\right\}\right] = 0.$$

*Then*

$$\liminf_{k\to\infty}\int_{\bar{\mathbb{R}}_+\times D} z(T)F'_{c_k}(x)(\bar{\nu})_{1,3}(dx\,dz) \geq 0, \quad a.s. \tag{3.21}$$

**Proof:** Note that from Lemma 3.7

$$\begin{aligned}
&\int_{\bar{\mathbb{R}}_+\times D} z(T)F'_{c_k}(x)(\bar{\nu}^n)_{1,3}(dx\,dz) \\
&\quad\geq -5\int_D ||z||1_{||z||\geq\frac{c_k}{2}}(\bar{\nu}^n)_3(dz) - \frac{2}{c_k}\int_D ||z||(\bar{\nu}^n)_3(dz) - \frac{F_{c_k}(\bar{X}_0^n)}{n},
\end{aligned} \tag{3.22}$$

for all $k \in \mathbb{N}$, a.s. We would like to show that the analogous inequality holds with $\bar{\nu}^n$ replaced by $\bar{\nu}$. By using the Skorohod representation theorem, we can assume without loss of generality that $\bar{\nu}^n$ converges to $\bar{\nu}$ almost surely. The term $F_c(\bar{X}_0^n)/n$ disappears in the limit $n \to \infty$, since $\bar{X}_0^n$ takes a fixed deterministic value for all $n$. We will now show that if $\bar{\nu}^n \to \bar{\nu}$ almost surely then all other terms in the above inequality converge to the corresponding terms with $\bar{\nu}^n$ replaced by $\bar{\nu}$.

We begin by considering $\int_{\bar{\mathbb{R}}_+\times D} z(T)F'_{c_k}(x)(\bar{\nu}^n)_{1,3}(dxdz)$. Note that for all $g \in C_b(\mathbb{R})$,

$$\int_{\bar{\mathbb{R}}_+\times D} g(z(T))F'_{c_k}(x)(\bar{\nu}^n)_{1,3}(dxdz) \to \int_{\bar{\mathbb{R}}_+\times D} g(z(T))F'_{c_k}(x)(\bar{\nu})_{1,3}(dxdz)$$

a.s. We can approximate the identity by a bounded continuous function $g(u)$ which equals $u$ when $|u| \leq C$ for a large constant $C$. The uniform integrability expressed in Lemma 3.5 then justifies the replacement of $g(z(T))$ in the last display by $z(T)$, giving

$$\int_{\bar{\mathbb{R}}_+\times D} z(T)F'_{c_k}(x)(\bar{\nu}^n)_{1,3}(dxdz) \to \int_{\bar{\mathbb{R}}_+\times D} z(T)F'_{c_k}(x)(\bar{\nu})_{1,3}(dxdz) \tag{3.23}$$

a.s. See, e.g., the proof of Lemma 5.3.6 in [9]. In exactly the same way, we have that

$$\int_D ||z||\,(\bar{\nu}^n)_3(dz) \to \int_D ||z||\,(\bar{\nu})_3(dz), \tag{3.24}$$

a.s. as $n \to \infty$. Finally we consider $\int_{||z|| \geq \frac{c_k}{2}} ||z||(\bar{\nu}^n)_3(dz)$. The condition assumed of the sequence $c_k$ implies that the function $z \mapsto ||z|| 1_{||z|| \geq \frac{c_k}{2}}$ is continuous w.p.1 under $\bar{\nu}$, a.s. When combined with the uniform integrability stated in Lemma 3.5, it follows that

$$\int_{||z|| \geq \frac{c_k}{2}} ||z||(\bar{\nu}^n)_3(dz) \to \int_{||z|| \geq \frac{c_k}{2}} ||z||(\bar{\nu})_3(dz), \tag{3.25}$$

a.s. Thus taking the limit as $n \to \infty$ in (3.22), we have from (3.23), (3.24) and (3.25) that

$$\begin{aligned} \int_{\bar{I\!\!R}_+ \times D} z(T) F'_{c_k}(x)(\bar{\nu})_{1,3}(dx\, dz) \quad &\geq \quad -5 \int_D ||z|| 1_{||z|| \geq \frac{c_k}{2}} (\bar{\nu})_3(dz) \\ &\quad - \frac{2}{c_k} \int_D ||z||(\bar{\nu})_3(dz). \end{aligned}$$

Finally the proof is completed by letting $k \to \infty$ in the last display and using Lemma 3.5. ∎

Let $\bar{\nu}$ be as in (3.4) and let $\bar{q}$ be a stochastic kernel in $\mathcal{S}(D|\bar{I\!\!R}_+)$ such that

$$(\bar{\nu})_{1,2}(dx\, dy) = (\bar{\nu})_1(dx) \otimes \bar{q}(dy|x), \tag{3.26}$$

i.e. for $A \in \mathcal{B}(\bar{I\!\!R}_+)$, $B \in \mathcal{B}(D)$,

$$(\bar{\nu})_{1,2}(A \times B) = \int_A \bar{q}(B|x)(\bar{\nu})_1(dx).$$

As an immediate consequence of Lemma 3.8, we have the following result.

**Lemma 3.9** *For $\bar{I\!\!P}_x$ - a.e. $\bar{\omega}$ in the set $\{(\bar{\nu})_1(I\!\!R_+) < 1\}$,*

$$\int_D ||y|| \bar{q}(dy|\infty) < \infty \tag{3.27}$$

*and*

$$\int_D y(T) \bar{q}(dy|\infty) \geq 0. \tag{3.28}$$

**Proof:** The inequality in (3.27) follows from the fact that

$$\begin{aligned} (1 - (\bar{\nu})_1(I\!\!R_+)) \int_D ||y|| q(dy|\infty) \quad &\leq \quad \int_{\bar{I\!\!R}_+ \times D} ||y||(\bar{\nu})_{1,2}(dx\, dy) \\ &= \quad \int_{\bar{I\!\!R}_+ \times D} ||z||(\bar{\nu})_{1,3}(dx\, dz) \\ &= \quad \int_D ||z||(\bar{\nu})_3(dz) \\ &< \quad \infty \end{aligned}$$

a.s., where the last step follows from Lemma 3.5. In order to see (3.28), note that if the sequence $\{c_k\}$ is chosen as in Lemma 3.8, then (3.21) holds. Furthermore, note that for all $(x, z) \in \bar{I\!\!R}_+ \times D$, $|z(T) F'_{c_k}(x)| \leq ||z||$ and $z(T) F'_{c_k}(x) \to z(T) 1_{\{\infty\}}(x)$, as $k \to \infty$. Thus observing that $\int_{\bar{I\!R}_+\times D}||z||(\bar{\nu})_{1,3}(dx\,dz) < \infty$, we have via an application of the dominated convergence theorem that

$$\begin{aligned}
\lim_{k\to\infty}\int_{\bar{I\!R}_+\times D} z(T)F'_{c_k}(x)(\bar{\nu})_{1,3}(dx\,dz) &= \int_{\bar{I\!R}_+\times D} z(T)1_{\{\infty\}}(x)(\bar{\nu})_{1,3}(dx\,dz) \\
&= \int_{\bar{I\!R}_+\times D} y(T)1_{\{\infty\}}(x)(\bar{\nu})_{1,2}(dx\,dy) \\
&= (\bar{\nu})_1\{\infty\}\int_D y(T)\bar{q}(dy|\infty), \qquad (3.29)
\end{aligned}$$

where the second equality follows from Lemma 3.4. Combining (3.21) and (3.29), we have the result. ∎

Let $\bar{\nu}$ and $\bar{q}$ be related as in (3.26). Define a stochastic kernel $\bar{p} \in \mathcal{S}(I\!R_+|I\!R_+)$ via

$$\bar{p}(x,A) \doteq \int_D 1_{\{\Pi(x,z)\in A\}}\bar{q}(dz|x).$$

Recall that $(\bar{\nu})_1$ can be decomposed as

$$(\bar{\nu})_1(\cdot) = (\bar{\nu})_1(I\!R_+)\widehat{(\bar{\nu})_1}(\cdot) + (1-(\bar{\nu})_1(I\!R_+))\delta_\infty(\cdot),$$

where $\widehat{(\bar{\nu})_1}$ is given via (2.9) with $\nu$ there replaced by $(\bar{\nu})_1$.

The following lemma characterizes $\widehat{(\bar{\nu})_1}$ as a $\bar{p}$-invariant probability measure.

**Lemma 3.10** *Let $\bar{\nu}$, $\bar{q}$ and $\bar{p}$ be as above. For $\bar{I\!P}_x$ - a.e. $\bar{\omega}$ in the set $\{(\bar{\nu})_1(I\!R_+) > 0\}$, $\widehat{(\bar{\nu})_1}$ is $\bar{p}$-invariant, i.e.*

$$\widehat{(\bar{\nu})_1}(A) = \int_{I\!R_+}\bar{p}(x,A)\widehat{(\bar{\nu})_1}(dx), \quad \textit{for all } A\in\mathcal{B}(I\!R_+).$$

**Proof:** It suffices to show that for all $g \in C_c(I\!R_+)$

$$\int_{\bar{I\!R}_+} g(x)(\bar{\nu})_1(dx) = (\bar{\nu})_1(I\!R_+)\int_{I\!R_+}\left(\int_{I\!R_+} g(y)\bar{p}(x,dy)\right)\widehat{(\bar{\nu})_1}(dx),$$

a.s. We begin by observing that

$$\begin{aligned}
&(\bar{\nu})_1(I\!R_+)\int_{I\!R_+}\left(\int_{I\!R_+} g(y)\bar{p}(x,dy)\right)\widehat{(\bar{\nu})_1}(dx) \\
&= (\bar{\nu})_1(I\!R_+)\int_{I\!R_+}\left(\int_D g(\Pi(x,y))\bar{q}(dy|x)\right)\widehat{(\bar{\nu})_1}(dx) \\
&= (\bar{\nu})_1(I\!R_+)\int_{I\!R_+\times D} g(\Pi(x,y))\widehat{(\bar{\nu})_{1,2}}(dx\,dy) \\
&= \int_{\bar{I\!R}_+\times D} H_g(x,y)(\bar{\nu})_{1,2}(dx\,dy) \\
&= \int_{\bar{I\!R}_+\times D} H_g(x,z)(\bar{\nu})_{1,3}(dx\,dz) \qquad (3.30)
\end{aligned}$$

a.s., where $\widehat{(\bar{\nu})_{1,2}} \in \mathcal{P}(\mathbb{R}_+ \times D)$ is defined as

$$\widehat{(\bar{\nu})_{1,2}}(A \times B) \doteq \frac{(\bar{\nu})_{1,2}(A \times B)}{(\bar{\nu})_{1,2}(\mathbb{R}_+ \times D)}, \quad A \in \mathcal{B}(\mathbb{R}_+), \quad B \in \mathcal{B}(D),$$

and $H_g \in C_b(\bar{\mathbb{R}}_+ \times D)$ is defined as

$$H_g(x,z) \doteq \begin{cases} g(\Pi(x,z)), & \text{if } (x,z) \in \mathbb{R}_+ \times D \\ 0, & \text{otherwise.} \end{cases}$$

The next to last equality in (3.30) follows from the compact support property of $g$ and the last equality is a consequence of Lemma 3.4. Thus in order to complete the proof of the lemma, it suffices to show that

$$\int_{\bar{\mathbb{R}}_+} g(x)(\bar{\nu}^n)_1(dx) - \int_{\bar{\mathbb{R}}_+ \times D} H_g(x,z)(\bar{\nu}^n)_{1,3}(dx\, dz) \to 0 \tag{3.31}$$

in probability as $n \to \infty$. Note that the expression in (3.31) equals

$$\begin{aligned} &\frac{1}{n}\sum_{j=0}^{n-1} g(\bar{X}_j^n) - \frac{1}{n}\sum_{j=0}^{n-1} g(\Pi(\bar{X}_j^n, \bar{Z}_j^n)) \\ &= \frac{1}{n}\sum_{j=0}^{n-1} g(\bar{X}_j^n) - \frac{1}{n}\sum_{j=0}^{n-1} g(\bar{X}_{j+1}^n) \\ &= \frac{1}{n}[g(\bar{X}_0^n) - g(\bar{X}_n^n)]. \end{aligned}$$

From the boundedness of $g$, it follows that the above expression is $O(1/n)$. This proves (3.31) and hence the lemma. ∎

Finally, we ready to prove Theorem 3.1.

**Proof of Theorem 3.1:** Recall from (3.5) that

$$\begin{aligned} W^n(x) + \epsilon &\geq \bar{\mathbb{E}}_x \left[ \frac{1}{n}\sum_{j=0}^{n-1} R(\bar{\nu}_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)) + F(\bar{L}^n) \right] \\ &= \bar{\mathbb{E}}_x \left[ \frac{1}{n}\sum_{j=0}^{n-1} R(\delta_{\bar{X}_j^n} \otimes \bar{\nu}_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \delta_{\bar{X}_j^n} \otimes \theta(\cdot)) + F(\bar{L}^n) \right] \\ &\geq \bar{\mathbb{E}}_x \left[ R\left( \frac{1}{n}\sum_{j=0}^{n-1} \delta_{\bar{X}_j^n} \otimes \bar{\nu}_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \frac{1}{n}\sum_{j=0}^{n-1} \delta_{\bar{X}_j^n} \otimes \theta(\cdot) \right) + F(\bar{L}^n) \right] \\ &= \bar{\mathbb{E}}_x \left[ R((\bar{\nu}^n)_{12} \parallel (\bar{\nu}^n)_1 \otimes \theta) + F((\bar{\nu}^n)_1) \right], \end{aligned} \tag{3.32}$$

where the second line in the above display exploits a property of relative entropy with respect to the decomposition of measures according to their marginals ([9, C.3.3]), the second inequality follows from Jensen's inequality, and the last equality uses the definition of $\bar{\nu}^n$. Next, recalling (3.8), we can assume without loss of generality that $\bar{\nu}^n \Rightarrow \bar{\nu}$ a.s. as $n \to \infty$. Thus from (3.32) we have that

$$\begin{aligned} \liminf_{n\to\infty} W^n(x) + \epsilon &\geq \bar{I\!\!E}_x \left[ \liminf_{n\to\infty} \left( R((\bar{\nu}^n)_{12} \parallel (\bar{\nu}^n)_1 \otimes \theta) + F((\bar{\nu}^n)_1) \right) \right] \\ &\geq \bar{I\!\!E}_x \left[ R((\bar{\nu})_{12} \parallel (\bar{\nu})_1 \otimes \theta) + F((\bar{\nu})_1) \right], \end{aligned} \tag{3.33}$$

where the first inequality follows from Fatou's lemma and the last inequality is a consequence of lower semi-continuity of $R(\cdot \parallel \cdot)$ and the continuity of $F$.

Next note that the definitions of $I_1$ and $J$ and Lemmas 3.9 and 3.10 imply

$$\begin{aligned} R((\bar{\nu})_{12} \parallel (\bar{\nu})_1 \otimes \theta) &= \int_{\bar{I\!\!R}_+} R(\bar{q}(\cdot|x) \parallel \theta(\cdot))(\bar{\nu})_1(dx) \\ &= (\bar{\nu})_1(I\!\!R_+) \int_{I\!\!R_+} R(\bar{q}(\cdot|x) \parallel \theta(\cdot))\widehat{(\bar{\nu})_1}(dx) \\ &\quad + (1 - (\bar{\nu})_1(I\!\!R_+))R(\bar{q}(\cdot|\infty) \parallel \theta(\cdot)) \\ &\geq (\bar{\nu})_1(I\!\!R_+)I_1(\widehat{(\bar{\nu})_1}) + (1 - (\bar{\nu})_1(I\!\!R_+))J \\ &= I((\bar{\nu})_1). \end{aligned}$$

When combined with (3.33) we have

$$\begin{aligned} \liminf_{n\to\infty} W^n(x) + \epsilon &\geq \bar{I\!\!E}_x\{I((\bar{\nu})_1) + F((\bar{\nu})_1)\} \\ &\geq \inf_{\mu \in \mathcal{P}(\bar{I\!\!R}_+)} \{I(\mu) + F(\mu)\}. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, the result follows. ∎

# 4 Properties of the Rate Function

In this section we will prove some important properties of the rate function $I(\cdot)$ defined in (2.10). The proof of part (a) of Theorem 4.1 below is similar to the proof of Proposition 8.5.2 of [9] and therefore is omitted. Part (b) of Theorem 4.1 is crucially used in the proof of the lower bound in Section 5 and even though the ideas in the proof are quite standard, we present the proof for the sake of completeness. Finally part (c) of the Theorem follows on using arguments similar to those in Section 3 and its proof is given in the Appendix.

**Theorem 4.1** *Let the function $I : \mathcal{P}(\bar{I\!\!R}_+) \mapsto [0, \infty]$ be given via (2.10). Then the following conclusions hold.*

*(a) $I$ is a convex function.*

*(b) Suppose that for all $\alpha \in (0, \infty)$,*

$$\int_D e^{\alpha||z||}\theta(dz) < \infty.$$

*Let $\pi \in \mathcal{P}(\bar{I\!R}_+)$. Then there exists $\sigma_0 \in \mathcal{P}(D)$ and $\bar{q} \in \mathcal{S}(I\!R_+|D)$ such that the following are true.*

1. *$\int_D ||z||\sigma_0(dz) < \infty$ and $\int_D z(T)\sigma_0(dz) \geq 0$.*
2. *If $\hat{\pi}$ is defined via (2.9) with $\nu$ there replaced by $\pi$, then $\hat{\pi}$ is $\bar{p}$-invariant, where $\bar{p} \in \mathcal{S}(I\!R_+|I\!R_+)$ is defined as*

$$\bar{p}(x, A) \doteq \int_D 1_{\{\Pi(x,z)\in A\}}\bar{q}(dz|x), \quad A \in \mathcal{B}(I\!R_+), \quad x \in I\!R_+. \tag{4.1}$$

3. *The infimum in (2.6) (with $\nu$ replaced by $\hat{\pi}$) and (2.8) are attained at $\bar{q}$ and $\sigma_0$, respectively. I.e.*

$$I(\pi) = \pi(I\!R_+) \int_{I\!R_+} R(\bar{q}(\cdot|x) \parallel \theta(\cdot))\hat{\pi}(dx) + (1 - \pi(I\!R_+))R(\sigma_0 \parallel \theta).$$

*(c) Suppose that Assumption 2.1 is satisfied. Then for all $M \in [0, \infty)$, the level set $\{\pi \in \mathcal{P}(\bar{I\!R}_+) : I(\pi) \leq M\}$ is a compact set in $\mathcal{P}(\bar{I\!R}_+)$.*

**Proof:** The proof of (a) is omitted. We now consider (b).

We can assume without loss of generality that $I(\pi) < \infty$. It suffices to show that the infimum in (2.8) is attained for some $\sigma_0 \in \mathcal{P}(D)$, and that for $\nu \in \mathcal{P}(I\!R_+)$ with $I_1(\nu) < \infty$, the infimum in (2.6) is attained for some $q \in \mathcal{A}_1(\nu)$, where $\mathcal{A}_1(\nu)$ is defined in Remark 2.6. We consider the last issue first. Fix $\nu \in \mathcal{P}(I\!R_+)$ for which $I_1(\nu) < \infty$. Observe that

$$\begin{aligned} I_1(\nu) &= \inf_{\tau \in \mathcal{A}(\nu)} \{R(\tau \parallel (\tau)_1 \otimes \theta)\} \\ &= \inf_{\tau \in \mathcal{A}^*(\nu)} \{R(\tau \parallel (\tau)_1 \otimes \theta)\}, \end{aligned} \tag{4.2}$$

where

$$\mathcal{A}^*(\nu) \doteq \{\tau \in \mathcal{A}(\nu) : \; R(\tau \parallel (\tau)_1 \otimes \theta) \leq I_1(\nu) + 1\}.$$

We now show that $\mathcal{A}^*(\nu)$ is compact. This will prove that the infimum in (4.2) is attained. From the one to one correspondence between $\mathcal{A}(\nu)$ and $\mathcal{A}_1(\nu)$ (see Remark 2.6) we will then have that the infimum in (2.6) is attained for some $q \in \mathcal{A}_1(\nu)$. From the lower semi-continuity of the map

$$\mathcal{P}(I\!R_+ \times D) \times \mathcal{P}(I\!R_+ \times D) \ni (\tau^1, \tau^2) \mapsto R(\tau^1 \parallel \tau^2) \in [0, \infty]$$

and the fact that $\mathcal{A}(\nu)$ is closed, it follows that $\mathcal{A}^*(\nu)$ is closed. Hence we need only show that $\mathcal{A}^*(\nu)$ is relatively compact in $\mathcal{P}(I\!R_+ \times D)$. Since $(\tau)_1 = \nu$ for all $\tau \in \mathcal{A}^*(\nu)$, $\{(\tau)_1 : \tau \in \mathcal{A}^*(\nu)\}$ is relatively compact in $\mathcal{P}(I\!R_+)$. Thus it suffices to show the relative compactness of $\{(\tau)_2 : \tau \in \mathcal{A}^*(\nu)\}$ in $\mathcal{P}(D)$. From Theorem 13.2 of [1] and an application of Chebychev's inequality it follows that, to prove the above relative compactness, it suffices to show

$$\sup_{\tau \in \mathcal{A}^*(\nu)} \int_D ||z||\tau(dx\, dz) < \infty \tag{4.3}$$

and

$$\sup_{\tau\in\mathcal{A}^*(\nu)} \int_D w'(z,\delta)\tau(dx\,dz) \to 0, \text{ as } \delta\to 0, \tag{4.4}$$

where $w'(x,\delta)$ is the usual modulus of continuity in the Skorohod space (cf. [1] page 122).

The proof of (4.3) is similar to that of (4.4) so we only consider the latter. For $k,\alpha\in(0,\infty)$ and $z\in D$, let $c_k(z) \doteq \min\{\alpha w'(z,\delta),k\}$. Then for $\tau\in\mathcal{A}^*(\nu)$,

$$\begin{aligned}
\int_D c_k(z)(\tau)_2(dz) &= \int_{\mathbb{R}_+\times D} c_k(z)\tau(dx\,dz) \\
&= \left(\int_{\mathbb{R}_+\times D} c_k(z)\tau(dx\,dz) - \log\int_{\mathbb{R}_+\times D} e^{c_k(z)}((\tau)_1\otimes\theta)(dx\,dz)\right) \\
&\quad + \log\int_D e^{c_k(z)}\theta(dz) \\
&\le R(\tau\parallel(\tau)_1\otimes\theta) + \log\int_D e^{c_k(z)}\theta(dz) \\
&\le I_1(\nu) + 1 + \log\int_D e^{\alpha w'(z,\delta)}\theta(dz),
\end{aligned}$$

where the first inequality uses the Donsker-Varadhan variational formula for relative entropy [9, Lemma 1.4.3(a)]. Letting $k\to\infty$, we have

$$\sup_{\tau\in\mathcal{A}^*(\nu)} \int_D w'(z,\delta)(\tau)_2(dz) \le \frac{1}{\alpha}\left(I_1(\nu)+1+\log\int_D e^{\alpha w'(z,\delta)}\theta(dz)\right).$$

Observe that for all $z\in D$, $w'(z,\delta)\to 0$ as $\delta\to 0$ and that $w'(z,\delta)\le 2||z||$. Thus (4.3) follows from the above display in view of Assumption 2.1 upon taking first $\delta\to 0$ and then $\alpha\to\infty$. Thus $\mathcal{A}^*(\nu)$ is relatively compact, and since it is closed, we have the desired compactness. Thus the infimum in (2.6) is attained for some $q\in\mathcal{A}_1(\nu)$.

Now we consider the infimum in (2.8). Observe that

$$\begin{aligned}
J &\doteq \inf_{\sigma\in\mathcal{P}(D)}\left\{R(\sigma\parallel\theta) : \int_D ||z||\sigma(dz)<\infty,\ \ \int_D z(T)\sigma(dz)\ge 0\right\} \\
&= \inf_{\sigma\in\mathcal{P}^*_{\text{tr}}(D)} R(\sigma||\theta),
\end{aligned} \tag{4.5}$$

where

$$\mathcal{P}^*_{\text{tr}}(D) \doteq \left\{\sigma\in\mathcal{P}(D) : \int_D ||z||\sigma(dz)<\infty,\ \ \int_D z(T)\sigma(dz)\ge 0,\ \ R(\sigma\parallel\theta)\le J+1\right\}.$$

Again using the variational formula for relative entropy, one can show in a manner similar to the proof of (4.3) and (4.4) that

$$\sup_{\sigma\in\mathcal{P}^*_{\text{tr}}(D)} \int_D ||z||\sigma(dz) < \infty \tag{4.6}$$

and

$$\sup_{\sigma\in\mathcal{P}^*_{\rm tr}(D)}\int_D w'(z,\delta)\sigma(dz)\to 0, \text{ as } \delta\to 0. \tag{4.7}$$

This proves the relative compactness of $\mathcal{P}^*_{\rm tr}(D)$ in $\mathcal{P}(D)$. Finally observing that $\mathcal{P}^*_{\rm tr}(D)$ is closed, we have that $\mathcal{P}^*_{\rm tr}(D)$ is compact. Thus the infimum in (4.5) is attained. This proves (b).

As is often the case when applying weak convergence arguments to prove large deviation results, the compactness of the level sets of $I$ (item (c) in the theorem) is proved using a deterministic analogue of the argument used to prove the upper bound in Section 3. Details of the argument are given in the Appendix. ■

# 5 Laplace Principle Lower Bound

The main result of this section is the following.

**Theorem 5.1** *Suppose that Assumptions 2.1, 2.2 and 2.7 hold. Then for all $F\in C_b(\mathcal{P}(\bar{I\!\!R}_+))$ and $x\in I\!\!R_+$,*

$$\liminf_{n\to\infty}\frac{1}{n}\log I\!\!E_x\left[\exp(-nF(L^n))\right]\geq -\inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)}\{I(\mu)+F(\mu)\}.$$

Note that by Theorem 3.1

$$\inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)}\{I(\mu)+F(\mu)\}\leq ||F||_\infty<\infty.$$

Let $\epsilon\in(0,\infty)$ be arbitrary and let $\pi\in\mathcal{P}(\bar{I\!\!R}_+)$ satisfy

$$I(\pi)+F(\pi)<\inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)}\{I(\mu)+F(\mu)\}+\epsilon. \tag{5.1}$$

In view of Lemma 3.2, to prove the theorem it suffices to show the following. There exists a sequence of stochastic kernels $\{\nu^n_j, j=0,\ldots\}$ in $\mathcal{S}(D|I\!\!R_+\times\mathcal{M}(I\!\!R_+))$, and a corresponding controlled sequence of $I\!\!R_+\times\mathcal{M}(I\!\!R_+)\times D$-valued random variables $\{\bar{X}^n_j,\bar{L}^n_j,\bar{Z}^n_j\}_{j=0}^n$ on some probability space $(\bar\Omega,\bar{\mathcal{F}},\bar{I\!\!P}_x)$ defined as in (3.2), such that

$$\limsup_{n\to\infty}\bar{I\!\!E}_x\left[\frac{1}{n}\sum_{j=0}^{n-1}R\left(\nu^n_j(\cdot|\bar{X}^n_j,\bar{L}^n_j)\parallel\theta(\cdot)\right)+F(\bar{L}^n)\right]\leq I(\pi)+F(\pi). \tag{5.2}$$

For the rest of section we will assume, without loss of generality, that $0<\pi(I\!\!R_+)<1$. The cases where $\pi(I\!\!R_+)$ is 0 or 1 are proved using simple modifications of the argument used in this (harder) case.

To prove an inequality like (5.2) we must find controls $\nu_j^n$ which steer the controlled occupation measures $\bar{L}^n$ to $\pi$ and for which the expected mean relative entropy converges to the rate function. An appropriate definition of the control is suggested by $\bar{q}$ and $\sigma_0$ in part (b) of Theorem 4.1. In fact, if $\hat{\pi}$ were the unique invariant measure for $\bar{p}$ and if $\int z(T)\sigma_0(dz) > 0$ then we could use the following scheme. Consider the time interval $1, ..., n$. For a fraction $\pi(I\!R_+)$ of this time (i.e., from 0 till approximately $n\pi(I\!R_+)$) we use $\nu_j^n(dz|\bar{X}_j^n, \bar{L}_j^n) = \bar{q}(dz|\bar{X}_j^n)$. The ergodic theorem would then guarantee the desired convergence of the controlled occupation measures and the expected mean relative entropy. For the remaining time (from $n\pi(I\!R_+)$ to $n$) we make the process transient by using $\nu_j^n = \sigma_0$, with the associated relative entropy cost. With this partitioning of time, the occupation measures and normalized relative entropies will converge to their proper limits. The main difficulty is that without additional conditions $\bar{p}$ need not be ergodic. To deal with this, a perturbation argument is used to approximate $\bar{p}$ by an ergodic probability transition function. In fact, we will perturb $\bar{p}$ is the direction of an ergodic transition function $p_0$, introduced below, and this will suffice. In addition, the measure $\sigma_0$ is not transient, and it must also be perturbed slightly. Some of the arguments parallel those under more standard assumptions, and therefore in places where the arguments are more or less identical we refer the reader to [9].

Let $\theta_0$ be as in Assumption 2.7. The Markov chain defined via (2.3) with $\{Z_n\}$ having the common law $\theta_0$ instead of $\theta$ will be denoted by $\{X_n^0,\ n \in I\!N_0\}$, and the transition probability function of this Markov chain will be denoted by $p_0(x, dy)$. Note that from part (c) of Assumption 2.7, it follows that Assumption 2.2 is satisfied with $p$ replaced by $p_0$. I.e., with $l_0, n_0$ as in Assumption 2.2 we have that for all $x_1, x_2 \in I\!R_+$

$$\sum_{i=l_0}^{\infty} \frac{1}{2^i} p_0^{(i)}(x_1, dy) \ll \sum_{j=n_0}^{\infty} \frac{1}{2^j} p_0^{(j)}(x_2, dy).$$

We begin with an elementary stability result regarding the random walk.

**Lemma 5.2** *Suppose that Assumptions 2.1 and 2.2 hold. For $x \in I\!R_+$ and $n \in I\!N_0$, let $\mu_n^x \in \mathcal{P}(I\!R_+)$ be defined by*

$$\mu_n^x(A) \doteq \frac{1}{n} \sum_{j=0}^{n} I\!P_x(X_j^0 \in A), \quad A \in \mathcal{B}(I\!R_+).$$

*Then for all $C \in (0, \infty)$, the family $\{\mu_n^x : n \in I\!N_0, |x| \leq C\}$ is relatively compact in $\mathcal{P}(I\!R_+)$.*

**Proof:** Using the Lyapunov function $V(x) \doteq x$, the lemma follows from a standard result in stochastic stability (cf. Theorems 11.3.4 and 12.4.4 of [12]). ∎

**Lemma 5.3** *Suppose that Assumptions 2.1 and 2.2 hold. Then the following conclusions hold.*

*(a) The $I\!R_+$-valued Markov chain $X_n^0$ has a unique invariant measure $\mu^*$. The Markov chain having $\mu^*$ as its initial distribution and $p_0(x, dy)$ as the transition probability function is ergodic.*
*(b) Let $A \in \mathcal{B}(I\!R_+)$ be such that $p^{(l_0)}(x_0, A) > 0$ for some $x_0 \in I\!R_+$, where $l_0$ is as in Assumption 2.2. Then $\mu^*(A) > 0$.*
*(c) If $\nu \in \mathcal{P}(I\!R_+)$ is such that $I_1(\nu) < \infty$, then $\nu \ll \mu^*$.*

**Proof:** Tightness proved in Lemma 5.2 and the Feller property of the Markov chain $\{X_n^0\}$ implies that there is at least one $p_0(x, dy)$-invariant probability measure. The proof of the uniqueness of the invariant measure and part (b) is same as that of Lemma 8.6.2 of [9] on noting that $p^{(l_0)}(x_0, dy)$ and $p_0^{(l_0)}(x_0, dy)$ are mutually absolutely continuous. Finally we consider (c). From Theorem 4.1 (b)(3), we have that there exists $\bar{q} \in \mathcal{S}(D|I\!R_+)$ such that

$$I_1(\nu) = \int_{I\!R_+} R(\bar{q}(\cdot|x) \parallel \theta(\cdot))\nu(dx).$$

Since $I_1(\nu) < \infty$, we have that $\Gamma \doteq \{x \in I\!R_+ : \bar{q}(\cdot|x) \ll \theta(\cdot)\}$ satisfies $\nu(\Gamma) = 1$. Since $\theta$ and $\theta_0$ are mutually absolutely continuous, we can replace $\theta$ by $\theta_0$ in the definition of $\Gamma$. Let $\bar{p} \in \mathcal{S}(I\!R_+|I\!R_+)$ be defined via (4.1). Then the above observations imply that, for all $x \in \Gamma$,

$$\bar{p}(x, dy) \ll p_0(x, dy).$$

Now let $A \in \mathcal{B}(I\!R_+)$ be such that $\nu(A) > 0$. It follows as in the proof of Lemma 8.6.2 of [9], that $p_0^{(l_0)}(x_0, A) > 0$ for some $x_0 \in I\!R_+$. Combining this with (b) we have that $\mu^*(A) > 0$. This proves (c). ■

**Lemma 5.4** *Let $\pi \in \mathcal{P}(\bar{I\!R}_+)$ be as in (5.1) and $F \in C_b(\mathcal{P}(\bar{I\!R}_+))$ be given. Let $\epsilon_0 \in (0, \infty)$ be arbitrary. Then there exists $\pi^* \in \mathcal{P}(\bar{I\!R}_+)$, $q^* \in \mathcal{A}_1(\widehat{\pi^*})$ and $\sigma^* \in \mathcal{P}_{tr}(D)$ such that the following hold.*

*(a) $\mu^* \ll \widehat{\pi^*}$ and $\widehat{\pi^*} \ll \mu^*$, where $\mu^*$ is the unique invariant measure of the $I\!R_+$-valued Markov chain $X_n^0$, as in Lemma 5.3.*

*(b)*
$$I_1(\widehat{\pi^*}) \leq \int_{I\!R_+} R(q^*(\cdot|x) \| \theta(\cdot))\widehat{\pi^*}(dx) \leq I_1(\hat{\pi}) + \frac{\epsilon_0}{3}.$$

*(c)*
$$\begin{aligned}&\pi^*(I\!R_+) \int_{I\!R_+} R(q^*(\cdot|x) \| \theta(\cdot))\widehat{\pi^*}(dx) + (1 - \pi^*(I\!R_+))R(\sigma^* \| \theta) + F(\pi^*) \\ &\leq I(\pi) + F(\pi) + \epsilon_0.\end{aligned}$$

*(d) The $I\!R_+$-valued Markov chain with transition probability function $p^*(x, dy)$ defined as*
$$p^*(x, A) \doteq \int_D 1_{\{\Pi(x,z) \in A\}} q^*(dz|x), \quad x \in I\!R_+, \ A \in \mathcal{B}(I\!R_+),$$
*is ergodic with $\widehat{\pi^*}$ its unique invariant measure.*

*(e)* $\int_D z(T)\sigma^*(dz) > 0$.

**Proof:** Since $F \in C_b(\mathcal{P}(\bar{I\!R}_+))$, there exists $\delta_1 \in (0,\infty)$ such that $|F(\pi) - F(\tilde{\pi})| \le \frac{\epsilon_0}{3}$ whenever $\tilde{\pi} \in \mathcal{P}(\bar{I\!R}_+)$ is such that $||\tilde{\pi} - \pi||_v \le \delta_1$. Here, $||\cdot||_v$ denotes the total variation norm. Define

$$\delta_0 \doteq \min\left\{1,\ \delta_1,\ \frac{\epsilon_0 \pi(I\!R_+)}{3R(\theta_0 \parallel \theta)},\ \frac{2\epsilon_0}{3I_1(\hat{\pi})}\right\} > 0$$

and let $\pi^* \in \mathcal{P}(\bar{I\!R}_+)$ be given by the formula

$$\pi^* \doteq \left(1 - \frac{\delta_0}{2}\right)\pi + \frac{\delta_0}{2}\mu^*,$$

where $\mu^*$ is as in Lemma 5.3. Then clearly

$$F(\pi^*) \le F(\pi) + \frac{\epsilon_0}{3}. \tag{5.3}$$

A straightforward calculation shows that

$$\widehat{\pi^*} = (1-\alpha)\hat{\pi} + \alpha\mu^*, \tag{5.4}$$

where

$$\alpha = \frac{\delta_0}{2((1-\frac{\delta_0}{2})\pi(I\!R_+) + \frac{\delta_0}{2})} \le \frac{\delta_0}{2}.$$

From (5.4) it is clear that $\mu^* \ll \widehat{\pi^*}$. Also, since $I_1(\hat{\pi}) < \infty$, we have from Lemma 5.3 that $\hat{\pi} \ll \mu^*$. Combining this with (5.4) we have that $\widehat{\pi^*} \ll \mu^*$. This proves (a). We now consider (b). From Theorem 4.1 (b)(3), we have that there exists $\bar{q} \in \mathcal{A}_1(\hat{\pi})$ such that

$$I_1(\hat{\pi}) = \int_{I\!R_+} R(\bar{q}(\cdot|x) \parallel \theta(\cdot))\hat{\pi}(dx).$$

Define $\lambda^* \in \mathcal{P}(I\!R_+ \times D)$ as

$$\lambda^*(A \times B) \doteq (1-\alpha)\int_A \bar{q}(B|x)\hat{\pi}(dx) + \alpha\theta_0(B)\mu^*(A), \quad A \in \mathcal{B}(I\!R_+), B \in \mathcal{B}(D).$$

We recall the definition of $\mathcal{A}$ given right after (2.7). Since $\bar{q} \in \mathcal{A}_1(\hat{\pi})$ and $\mu^*$ is $p_0$-invariant, we have that $\lambda^* \in \mathcal{A}(\widehat{\pi^*})$. Let $q^* \in \mathcal{S}(I\!R_+|D)$ be such that

$$\lambda^*(A \times B) = \int_A q^*(B|x)\widehat{\pi^*}(dx).$$

Then $q^* \in \mathcal{A}_1(\widehat{\pi^*})$. This implies that

$$\begin{aligned}
I_1(\widehat{\pi^*}) &\le \int_{I\!R_+} R(q^*(\cdot|x) \parallel \theta(\cdot))\, \widehat{\pi^*}(dx) \\
&= R(\lambda^* \parallel \widehat{\pi^*} \otimes \theta) \\
&\le (1-\alpha)\int_{I\!R_+} R(\bar{q}(\cdot|x) \parallel \theta(\cdot))\hat{\pi}(dx) + \alpha\int_{I\!R_+} R(\theta_0 \parallel \theta)\mu^*(dx) \\
&= (1-\alpha)I_1(\hat{\pi}) + \alpha R(\theta_0 \parallel \theta),
\end{aligned} \tag{5.5}$$

where the second inequality above follows from the convexity of the relative entropy function. Next note that our choice of $\delta_0$ is such that $\alpha \leq \frac{\epsilon_0 \pi(I\!R_+)}{6R(\theta_0\|\theta)} \leq \frac{\epsilon_0}{3R(\theta_0\|\theta)}$. Using this observation in (5.5) it follows that

$$I_1(\widehat{\pi^*}) \leq \int_{I\!R_+} R(q^*(\cdot|x) \parallel \theta(\cdot))\, \widehat{\pi^*}(dx) \leq I_1(\hat{\pi}) + \frac{\epsilon_0}{3}.$$

This proves (b).

Next, let $\sigma_0 \in \mathcal{P}(D)$ be as in Theorem 4.1. Then $J = R(\sigma_0 \parallel \theta)$, $\int_D ||z||\sigma_0(dz) < \infty$ and $\int_D z(T)\sigma_0(dz) \geq 0$. Let $\theta_1$ be as in Assumption 2.7 and let $\kappa \in (0, \frac{\epsilon_0}{3R(\theta_1||\theta)})$. Define $\sigma^* \doteq (1-\kappa)\sigma_0 + \kappa\theta_1$. Then

$$\begin{aligned}
R(\sigma^*\|\theta) &\leq (1-\kappa)R(\sigma_0\|\theta) + \kappa R(\theta_1\|\theta) \\
&\leq R(\sigma_0\|\theta) + \frac{\epsilon_0}{3} \\
&= J + \frac{\epsilon_0}{3}.
\end{aligned}$$

Clearly $\int_R z(T)\sigma^*(dz) > 0$, and thus part (e) of the lemma holds. Now observe that

$$\begin{aligned}
&\pi^*(I\!R_+)\int_{I\!R_+} R(q^*(\cdot|x)\|\theta(\cdot))\widehat{\pi^*}(dx) + (1-\pi^*(I\!R_+))R(\sigma^*\|\theta) \\
&\leq \pi^*(I\!R_+)\int_{I\!R_+} R(q^*(\cdot|x)\|\theta(\cdot))\widehat{\pi^*}(dx) + (1-\pi(I\!R_+))R(\sigma^*\|\theta) \\
&\leq \pi^*(I\!R_+)I_1(\hat{\pi}) + (1-\pi(I\!R_+))J + \frac{\epsilon_0}{3} \\
&\leq \pi(I\!R_+)I_1(\hat{\pi}) + (1-\pi(I\!R_+))J + \frac{2\epsilon_0}{3} \\
&\leq I(\pi) + \frac{2\epsilon_0}{3}, \qquad (5.6)
\end{aligned}$$

where the first inequality follows on noting that $\pi^*(I\!R_+) > \pi(I\!R_+)$ and the third inequality is a consequence of the fact that with our choice of $\delta_0$, $\pi^*(I\!R_+) - \pi(I\!R_+) \leq \frac{\epsilon_0}{3I_1(\hat{\pi})}$. Now (c) follows on combining (5.3) and (5.6). Since we assumed that $\pi(I\!R_+) > 0$, it follows that $\pi^*(I\!R_+) > 0$ and so we have from (c) that

$$\int_{I\!R_+} R(q^*(\cdot|x) \parallel \theta(\cdot))\widehat{\pi^*}(dx) \leq \frac{I(\pi) + \epsilon_0 + 2\|F\|_\infty}{\pi^*(I\!R_+)} < \infty.$$

Using this fact and the definition of $\lambda^*$ it follows as in the proof of Lemma 8.6.3 of [9] that $q^*(dy|x)$, $\theta$ and $\theta_0$ are mutually absolutely continuous, $\widehat{\pi^*}-$ a.e. $x$. By modifying $q^*$ over a $\widehat{\pi^*}$ null set we have the above mutual absolute continuity to hold for all $x \in I\!R_+$. From this it then follows that Assumption 2.2 is satisfied with $p$ replaced by $p^*$. This combined with the fact that $\widehat{\pi^*}$ is $p^*$-invariant, gives (d). ■

Now let $\pi^*, q^*$ and $\sigma^*$ be as in the previous lemma. We would like to show that there exists a family of controls $\{\nu_j^n \in \mathcal{S}(D|I\!R_+ \times \mathcal{M}(I\!R_+)),\ j = 0, \ldots, n,\ n \in I\!N\}$ such that

$$\limsup_{n\to\infty} \bar{I\!E}_x \left[ \frac{1}{n}\sum_{j=0}^{n-1} R(\nu_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)) + F(\bar{L}^n) \right]$$

$$\leq \quad \pi^*(\mathbb{R}_+)\int_{\mathbb{R}_+} R(q^*(\cdot|x) \parallel \theta(\cdot))\widehat{\pi^*}(dx) + (1-\pi^*(\mathbb{R}_+))R(\sigma^* \parallel \theta) + F(\pi^*), \tag{5.7}$$

where $\{\bar{X}_j^n, \bar{L}_j^n\}$ are defined as in (3.2). Note that $\pi^*(\mathbb{R}_+) = (1-\delta_0/2)\pi(\mathbb{R}_+)+\delta_0/2$. Since $\pi(\mathbb{R}_+) \in (0,1)$ we have that $\pi^*(\mathbb{R}_+) \in (0,1)$. Henceforth, we will denote $\pi^*(\mathbb{R}_+)$ by $\rho$. We introduce the following canonical spaces. Let

$$(\Omega_1, \mathcal{F}_1) \doteq ((\mathbb{R}_+)^\infty, \mathcal{B}((\mathbb{R}_+)^\infty)),$$

and

$$(\Omega_2, \mathcal{F}_2, Q) \doteq (D^\infty, \mathcal{B}(D^\infty), (\sigma^*)^{\otimes\infty}).$$

Denote by $\tilde{\mathbb{P}}_x$ the probability measure on $(\Omega_1, \mathcal{F}_1)$ under which the canonical sequence

$$\xi_n(\omega_1) \doteq \omega_1(n), \;\; \omega_1 \in \Omega_1, \;\; n \in \mathbb{N}_0,$$

is a Markov chain with transition probability function $p^*(x,dy)$ and $\xi_0 \equiv x$. From the ergodicity of this Markov chain it follows that there exists a $\Phi \in \mathcal{B}(\mathbb{R}_+)$ such that $\widehat{\pi^*}(\Phi) = 1$ (or equivalently, from Lemma 5.4, $\mu^*(\Phi) = 1$) and such that for all $x \in \Phi$

$$\tilde{L}_n \doteq \frac{1}{n}\sum_{j=0}^{n} \delta_{\xi_j} \Rightarrow \widehat{\pi^*}, \;\; a.s. \; [\tilde{\mathbb{P}}_x].$$

On the probability space $(\Omega_2, \mathcal{F}_2, Q)$ define the canonical sequence $\eta_n$ as

$$\eta_n(\omega_2) \doteq \omega_2(n), \;\; \omega_2 \in \Omega_2, \;\; n \in \mathbb{N}_0$$

Now define

$$(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}_x) \doteq (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \tilde{\mathbb{P}}_x \otimes Q).$$

We will continue to denote the canonical sequences $\{\xi_n\}, \{\eta_n\}$ on this extended space by the same symbols. Now define, for $n \in \mathbb{N}$ and $j \in 0, \ldots, n$, the random variables

$$\bar{X}_j^n \doteq \begin{cases} \xi_j & j = 0, \ldots [n\rho] \\ \Pi(\bar{X}_{j-1}^n, \eta_{j-[n\rho]}) & j = [n\rho]+1, \ldots, n, \end{cases}$$

where $[\cdot]$ denotes the greatest integer function. Clearly, $\{\bar{X}_j^n, \; j = 0, \ldots, n\}$ is a controlled Markov chain, as in (3.2) with

$$\nu_j^n(dz|x,\gamma) = \begin{cases} q^*(dz|x) & j = 0, \ldots [n\rho]-1 \\ \sigma^*(dz) & j = [n\rho], \ldots, n. \end{cases} \tag{5.8}$$

In what follows, we will show that with this choice of $\{\nu_j^n\}$ and $\{\bar{X}_j^n\}$, (5.7) holds. We begin with the following lemma.

**Lemma 5.5** *Let $\Phi$ be as above. Then for all $x \in \Phi$*

$$\bar{L}^n \Rightarrow \pi^*, \;\; a.s. \; [\bar{\mathbb{P}}_x].$$

**Proof:** It suffices to show that for all $x \in \Phi$ and continuous and bounded $g \in C_b(\bar{I\!R}_+)$,

$$\bar{L}^n(g) \to \langle g, \pi^* \rangle, \quad a.e.\ [\bar{I\!P}_x].$$

Now fix $x \in \Phi$ and let $g \in C_b(\bar{I\!R}_+)$. Let $\epsilon > 0$ be arbitrary, and let $\Omega_1^x \in \mathcal{F}_1$ be such that $\tilde{I\!P}_x(\Omega_1^x) = 1$, and for all $\omega_1 \in \Omega_1^x$

$$\frac{1}{n} \sum_{j=0}^{n} g(\xi_j(\omega_1)) \to \langle g, \widehat{\pi^*} \rangle$$

as $n \to \infty$. Fix such an $\omega_1 \in \Omega_1$, and let $N_0 \in I\!N$ be such that

$$\left| \frac{1}{n} \sum_{j=0}^{[n\rho]} g(\xi_j(\omega_1)) - \rho \langle g, \widehat{\pi^*} \rangle \right| \leq \epsilon \text{ for all } n \geq N_0. \tag{5.9}$$

Since $g \in C_b(\bar{I\!R}_+)$, there exists $L \in (0, \infty)$ such that

$$|g(x) - g(\infty)| < \epsilon \text{ for all } x > L. \tag{5.10}$$

Next let $\Omega_2^* \in \mathcal{F}_2$ be such that $Q(\Omega_2^*) = 1$ and for all $\omega_2 \in \Omega_2^*$, as $n \to \infty$,

$$\frac{1}{n} \sum_{j=0}^{n} \eta_j(\omega_2)(T) \to \int_D z(T) \sigma^*(dz) > 0.$$

Fix such an $\omega_2 \in \Omega_2^*$ and let $J_0 \in I\!N$ be such that

$$\inf_{n \geq J_0} \frac{1}{n} \sum_{j=0}^{n} \eta_j(\omega_2)(T) \doteq \kappa > 0.$$

Without loss of generality we can assume that $J_0 \kappa \geq L$. Now define

$$N_1 \doteq \max \left\{ N_0, \frac{J_0 + 1}{1 - \rho} \right\}.$$

Let $\omega \doteq (\omega_1, \omega_2)$. For $n \geq N_1$

$$\begin{aligned}
\left| \frac{1}{n} \sum_{j=0}^{n} g(\bar{X}_j^n(\omega)) - \langle g, \pi^* \rangle \right| &= \left| \frac{1}{n} \sum_{j=0}^{n} g(\bar{X}_j^n(\omega)) - \rho \langle g, \widehat{\pi^*} \rangle - (1-\rho) g(\infty) \right| \\
&\leq \left| \frac{1}{n} \sum_{j=0}^{[n\rho]} g(\bar{X}_j^n(\omega)) - \rho \langle g, \widehat{\pi^*} \rangle \right| \\
&\quad + \left| \frac{1}{n} \sum_{j=[n\rho]+1}^{n} g(\bar{X}_j^n(\omega)) - (1-\rho) g(\infty) \right| \\
&\leq \epsilon + \frac{1}{n} \sum_{j=[n\rho]+J_0}^{n} |g(\bar{X}_j^n(\omega)) - g(\infty)| \\
&\quad + \|g\|_\infty \left[ \frac{J_0}{n} + \frac{n - [n\rho] - J_0 + 1 - n(1-\rho)}{n} \right] \\
&\leq \epsilon + \sup_{[n\rho]+J_0 \leq j \leq n} |g(\bar{X}_j^n(\omega)) - g(\infty)| \\
&\quad + \|g\|_\infty \frac{J_0 + 2}{n},
\end{aligned} \tag{5.11}$$

where the second inequality follows from (5.9) and the observation that since $n \geq \frac{J_0+1}{1-\rho}$, we have that $n \geq [n\rho] + J_0$. Since the constraining action on the controlled random walk acts only in the positive direction, for $n \geq N_1$ and $j \geq [n\rho] + J_0$

$$\bar{X}_j^n(\omega) \geq \xi_{[n\rho]}(\omega_1) + \sum_{k=1}^{j-[n\rho]} \eta_k(\omega_2)(T).$$

This implies

$$\begin{aligned} \frac{\bar{X}_j^n(\omega)}{j-[n\rho]} &\geq \frac{\xi_{[n\rho]}(\omega_1)}{j-[n\rho]} + \frac{\sum_{k=1}^{j-[n\rho]} \eta_k(\omega_2)(T)}{j-[n\rho]} \\ &\geq \frac{\sum_{k=1}^{j-[n\rho]} \eta_k(\omega_2)(T)}{j-[n\rho]} \\ &\geq \kappa, \end{aligned}$$

and thus

$$\bar{X}_j^n \geq \kappa(j-[n\rho]) \geq \kappa J_0 \geq L.$$

In view of (5.10), for such $n$ and $j$

$$|g(\bar{X}_j^n(\omega)) - g(\infty)| \leq \epsilon.$$

Substituting this in (5.11) we have

$$\left| \frac{1}{n} \sum_{j=0}^{n} g(\bar{X}_j^n(\omega)) - \langle g, \pi^* \rangle \right| \leq 2\epsilon + \frac{J_0+2}{n}.$$

We send $n \to \infty$ in the last display. Since $\epsilon > 0$ is arbitrary, $\omega = (\omega_1, \omega_2) \in \Omega_1^x \times \Omega_2^*$ is arbitrary, and $\bar{I\!P}_x(\Omega_1^x \times \Omega_2^*) = 1$, the result follows. ∎

Finally, we prove Theorem 5.1.

**Proof of Theorem 5.1.** In order to prove the theorem we will show that every subsequence $\{n'\}$ has a further subsequence along which the limsup of $-\frac{1}{n} \log I\!E_x \left[\exp(-nF(L^n))\right]$ is bounded above by $\inf_{\mu \in \mathcal{P}(\bar{I\!R}_+)} \{I(\mu) + F(\mu)\}$. To minimize notation, we will denote the subsequence $\{n'\}$ once more by $\{n\}$.

From Corollary 1.2.5 of [9], it follows that we can assume without loss of generality that $F$ is Lipschitz continuous on $\mathcal{P}(\bar{I\!R}_+)$ with respect to the Levy-Prohorov metric. By Lemma 3.2

$$\begin{aligned} &\limsup_{n\to\infty} -\frac{1}{n} \log I\!E_x \left[\exp(-nF(L^n))\right] \\ &= \limsup_{n\to\infty} \inf_{\{\nu_j^n\}} \bar{I\!E}_x \left[ \frac{1}{n} \sum_{j=0}^{n-1} R\left(\nu_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)\right) + F(\bar{L}^n) \right] \\ &\leq \limsup_{n\to\infty} \bar{I\!E}_x \left[ \frac{1}{n} \sum_{j=0}^{n-1} R\left(\nu_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)\right) + F(\bar{L}^n) \right], \end{aligned}$$

where $\{\nu_j^n\}$ are as defined in (5.8).

Now suppose that $x \in \Phi$, where $\Phi$ is as in Lemma 5.5. Then from Lemma 5.5,

$$\bar{I\!E}_x[F(\bar{L}^n)] \to F(\pi^*). \tag{5.12}$$

Also,

$$\begin{aligned}
&\frac{1}{n}\sum_{j=0}^{n-1} R\left(\nu_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)\right) \\
&= \frac{1}{n}\sum_{j=0}^{[n\rho]-1} R\left(q^*(\cdot|\xi_j) \parallel \theta(\cdot)\right) + \frac{1}{n}\sum_{j=[n\rho]}^{n-1} R\left(\sigma^* \parallel \theta\right)
\end{aligned}$$

Thus

$$\begin{aligned}
&\limsup_{n\to\infty} \bar{I\!E}_x\left[\frac{1}{n}\sum_{j=0}^{n-1} R\left(\nu_j^n(\cdot|\bar{X}_j^n, \bar{L}_j^n) \parallel \theta(\cdot)\right)\right] \\
&\leq \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{[n\rho]} \bar{I\!E}_x\left[R\left(q^*(\cdot|\xi_j) \parallel \theta(\cdot)\right)\right] + (1-\rho)R\left(\sigma^* \parallel \theta\right).
\end{aligned} \tag{5.13}$$

For $x \in I\!R_+$ let

$$D_x^n \doteq \bar{E}_x\left|\frac{1}{n}\sum_{j=0}^{n} R\left(q^*(\cdot|\xi_j) \parallel \theta(\cdot)\right) - \int_{I\!R_+} R(q^*(\cdot|y) \parallel \theta(\cdot))\widehat{\pi^*}(dy)\right|.$$

Since

$$\int_{I\!R_+} R(q^*(\cdot|y) \parallel \theta(\cdot))\widehat{\pi^*}(dy) \leq \frac{I(\pi^*)}{\rho} < \infty,$$

we have by the $L^1$ ergodic theorem that

$$\int_{I\!R_+} D_x^n \widehat{\pi^*}(dx) \to 0, \text{ as } n \to \infty.$$

Thus we can find a subsequence, denoted once more by $\{n\}$, and $\Phi_1 \in \mathcal{B}(I\!R_+)$ such that $\widehat{\pi^*}(\Phi_1) = 1$ and for all $x \in \Phi_1$, $D_x^n \to 0$ as $n \to \infty$. Combining this observation with (5.12) and (5.13) we have that for all $x \in \Phi \cap \Phi_1$

$$\begin{aligned}
\limsup_{n\to\infty} -\frac{1}{n}\log I\!E_x\left[\exp(-nF(L^n))\right] &\leq \rho\int_{I\!R_+} R\left(q^*(\cdot|y) \parallel \theta(\cdot)\right)\widehat{\pi^*}(dy) \\
&\quad + (1-\rho)R(\sigma^* \parallel \theta) + F(\pi^*) \\
&\leq I(\pi) + F(\pi) + \epsilon_0 \\
&\leq \inf_{\mu\in\mathcal{P}(\bar{I\!R}_+)}\{I(\mu) + F(\mu)\} + \epsilon + \epsilon_0.
\end{aligned} \tag{5.14}$$

The last two inequalities follow from Lemma 5.4 and (5.1), respectively. Now an argument, as on pages 316-318 of [9], using the Lipschitz property of $F$ and the transitivity condition (Assumption 2.2) shows that the above inequality, in fact, holds for all $x \in I\!R_+$. Letting $\epsilon \to 0$ and $\epsilon_0 \to 0$ in (5.14) completes the proof. ■

# 6 Reflected Brownian Motion and the M/M/1 Queue

In this section we use the large deviation results for the discrete time Markov chain considered in Sections 3, 4 and 5 to obtain the empirical measure LDP for two very basic continuous time models: reflected Brownian motion and the M/M/1 queue.

We begin with the study of reflected Brownian motion. Let $\{W(t), t \in [0,\infty)\}$ be a standard Brownian motion on some probability space $(\Omega, \mathcal{F}, I\!\!P)$. Let $b \in I\!\!R$ and $\sigma \in I\!\!R_+$ be fixed. Define

$$Y(t) \doteq Y(0) + bt + \sigma W(t), \quad t \in [0,\infty).$$

We will only consider the case when $Y(0) = x$ for some point $x \in I\!\!R_+$, in which case expectation will be denoted by $I\!\!E_x$.

Reflected Brownian motion $X(\cdot)$ is defined by the relation

$$X(t) \doteq \Gamma(Y)(t), \quad t \in [0,\infty),$$

where $\Gamma$ is the Skorohod map defined in (2.1). It is well known that $X(\cdot)$ is a Feller Markov process with values in $I\!\!R_+$. For $T \in (0,\infty)$ define the empirical measure $L^T$ corresponding to this Markov process as the $\mathcal{P}(I\!\!R_+)$-valued random variable given by

$$L^T(A) \doteq \frac{1}{T}\int_0^T \delta_{X_s}(A)ds, \quad A \in \mathcal{B}(I\!\!R_+). \tag{6.1}$$

We now introduce the rate function that will govern the large deviation probabilities of $\{L^T, T \in (0,\infty)\}$.

Let

$$\mathcal{H}^+ \doteq \left\{u \in C^2(I\!\!R_+): \quad u'(0) = 0, \inf_{x\in I\!\!R_+} u(x) > 0, \lim_{x\to\infty} \frac{Au}{u}(x) = -\frac{(b^-)^2}{2\sigma^2}\right\},$$

where

$$(Au)(x) \doteq \frac{\sigma^2}{2}u''(x) + bu'(x), \tag{6.2}$$

and for $x \in I\!\!R$, $(x)^- \doteq -\min\{x, 0\}$. For $\nu \in \mathcal{P}(I\!\!R_+)$, let

$$I_1(\nu) \doteq -\inf_{u\in\mathcal{H}^+} \int_{I\!\!R_+} \left(\frac{Au}{u}\right)(x)\nu(dx), \tag{6.3}$$

The rate function $I(\cdot)$ for the empirical measure LDP for the reflected Brownian motion $X(\cdot)$ is then given as

$$I(\nu) \doteq \nu(I\!\!R_+)I_1(\hat{\nu}) + (1 - \nu(I\!\!R_+))\frac{(b^-)^2}{2\sigma^2}, \quad \nu \in \mathcal{P}(\bar{I\!\!R}_+). \tag{6.4}$$

**Theorem 6.1** *Let $I(\cdot)$ be defined as in (6.4). Then for all $F \in C_b(\mathcal{P}(\bar{I\!\!R}_+))$ and $x \in I\!\!R_+$*

$$\lim_{T\to\infty} \frac{1}{T} \log I\!\!E_x \left[\exp(-TF(L^T))\right] = - \inf_{\mu \in \mathcal{P}(\bar{I\!\!R}_+)} \{I(\mu) + F(\mu)\}. \tag{6.5}$$

*Furthermore, $I(\cdot)$ is a rate function on $\mathcal{P}(\bar{I\!\!R}_+)$.*

For $h \in (0, \infty)$, define the approximating occupation measures by the relation

$$L_h^n(A) \doteq \frac{1}{n} \sum_{j=0}^{n-1} \delta_{X(jh)}(A), \quad A \in \mathcal{B}(I\!\!R_+), \; n \in I\!\!N. \tag{6.6}$$

Note that since

$$X((n+1)h) = \Pi(X(nh), \xi_n^h), \quad n \in I\!\!N_0, \tag{6.7}$$

where

$$\xi_n^h \doteq (\sigma(W(nh+t) - W(nh)) + bt)_{0\le t\le h}, \quad n \in I\!\!N_0,$$

the LDP for $L_h^n$ for each fixed $h$ follows from Theorem 2.9 on setting $T$ equal to $h$.

A main estimate in the proof of Theorem 6.1 is the following lemma.

**Lemma 6.2** *For every $\delta > 0$.*

$$\limsup_{h\to 0} \limsup_{n\to\infty} \frac{1}{nh} \log I\!\!P_x \left\{ d(L^{nh}, L_h^n) > \delta \right\} = -\infty,$$

*where $d$ is the Levy-Prohorov metric on $\mathcal{P}(\bar{I\!\!R}_+)$.*

**Proof.** The proof is an immediate consequence of Lemma 3.4 of [6] on using the Markov property of $X(\cdot)$ and observing that for all $\delta > 0$

$$\sup_{x \in I\!\!R_+} I\!\!P_x \left\{ \sup_{0\le s\le h} |X(s) - x| \ge \delta \right\} \to 0$$

as $h \to 0$. ∎

**Sketch of the Proof of Theorem 6.1:** It suffices to prove (6.5) for functions $F$ which are Lipschitz continuous. Fix such an $F$ and denote the Lipschitz constant of $F$ by $M$. From Theorem 2.9 we have that

$$\lim_{n\to\infty} \frac{1}{n} \log I\!\!E_x[\exp(-nF(L_h^n))] = - \inf_{\mu \in \mathcal{P}(\bar{I\!\!R}_+)} \{I_h(\mu) + F(\mu)\},$$

where

$$I_h(\mu) \doteq \mu(I\!\!R_+) I_{h,1}(\hat{\mu}) + (1 - \mu(I\!\!R_+)) J_h. \tag{6.8}$$

Here

$$I_{h,1}(\nu) \doteq \inf_{q^* \in \mathcal{A}_1(\nu)} \int_{I\!R_+} R(q^*(\cdot|x) \parallel \theta_h(\cdot))\nu(dx), \tag{6.9}$$

$\theta_h$ is a Wiener measure with drift $b$ and variance $\sigma^2$, and

$$J_h \doteq \inf_{\gamma \in \mathcal{P}_{\rm tr}(D_h)} R(\gamma \parallel \theta_h).$$

Using the property $u'(0) = 0$ for $u \in \mathcal{H}^+$, a minor modification of the proof of Theorem 2.1 of [7] shows that for $\nu \in \mathcal{P}(I\!R_+)$

$$I_{h,1}(\nu) = -\inf_{u \in \mathcal{H}^+} \int_{I\!R_+} \log\left(\left(\frac{T_h u}{u}\right)(x)\right)\nu(dx), \tag{6.10}$$

where

$$(T_h u)(x) \doteq \int_D u(\Pi_h(x,z))\theta_h(dz).$$

The main difference, from Theorem 2.1 of [7], in (6.10) is that we use $\mathcal{H}^+$ as the class of test functions rather than

$$\tilde{\mathcal{H}}^+ \doteq \left\{u \in C_b^2(I\!R_+) : \ u'(0) = 0, \inf_{x \in I\!R_+} u(x) > 0\right\}.$$

However, note that every $u$ in $\tilde{\mathcal{H}}^+$ (and $\mathcal{H}^+$) can be modified outside a compact interval to yield a test function in $\mathcal{H}^+$ (resp. $\tilde{\mathcal{H}}^+$). Furthermore, for all test functions $u$ in $\mathcal{H}^+$ and $\tilde{\mathcal{H}}^+$, $\sup_{x\in I\!R_+} |\log\left(\left(\frac{T_h u}{u}\right)(x)\right)| < \infty$. These observations imply that the above modification in the class of test functions does not change the value of the infimum. More precisely, the infimum on the right side of (6.10) is the same as

$$\inf_{u \in \tilde{\mathcal{H}}^+} \int_{I\!R_+} \log\left(\left(\frac{T_h u}{u}\right)(x)\right)\nu(dx).$$

We will next show that

$$\inf_{\nu \in \mathcal{P}(\bar{I\!R}_+)} \{I(\nu) + F(\nu)\} = \lim_{h \to 0} \inf_{\nu \in \mathcal{P}(\bar{I\!R}_+)} \left\{\frac{I_h(\nu)}{h} + F(\nu)\right\}. \tag{6.11}$$

Once (6.11) is proved we have the result as follows. Using the boundedness and Lipschitz continuity of $F$,

$$\frac{1}{T}\log I\!E_x[\exp(-TF(L^T))] = \frac{1}{h[T/h]}\log I\!E_x[\exp(-[T/h]hF(L^{[T/h]h}))] + O(h/T),$$

where for $x \in I\!R_+$, $[x]$ denotes the integer part of $x$. Thus,

$$\begin{aligned}
&\liminf_{T\to\infty} \frac{1}{T}\log I\!E_x\left[\exp(-TF(L^T))\right] \\
&= \liminf_{n\to\infty} \frac{1}{nh}\log I\!E_x\left[\exp(-nhF(L^{nh}))\right] \\
&\geq \liminf_{n\to\infty} \frac{1}{nh}\log I\!E_x\left[1_{d(L^{nh},L_h^n)\leq\delta}\exp(-nhF(L^{nh}))\right] \\
&\geq \liminf_{n\to\infty} \frac{1}{nh}\log\left\{I\!E_x\left[\exp(-nh(F(L_h^n) - M\delta))\right.\right. \\
&\qquad \left.\left. - \exp\left(nh(||F||_\infty + M\delta)\right)I\!P_x(d(L^{nh}, L_h^n) > \delta)\right]\right\}.
\end{aligned}$$

Taking the limit as $h \to 0$ and using Lemma 6.2 gives

$$\begin{aligned}
&\liminf_{T\to\infty} \frac{1}{T} \log I\!\!E_x \left[\exp(-TF(L^T))\right] \\
&\quad\geq \liminf_{h\to 0} \liminf_{n\to\infty} \frac{1}{nh} \log I\!\!E_x \left[\exp(-n(hF(L_h^n) - Mh\delta))\right] \\
&\quad\geq \liminf_{h\to 0} -\frac{1}{h} \inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)} \{I_h(\mu) + hF(\mu)\} - M\delta \\
&\quad= \liminf_{h\to 0} - \inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)} \left\{\frac{1}{h} I_h(\mu) + F(\mu)\right\} - M\delta \\
&\quad= - \inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)} \{I(\mu) + F(\mu)\} - M\delta,
\end{aligned}$$

where the last step follows from (6.11). Since $\delta > 0$ is arbitrary, this proves the Laplace principle lower bound. For the upper bound, one has in a similar way that

$$\begin{aligned}
\liminf_{T\to\infty} \frac{1}{T} \log I\!\!E_x \left[\exp(-TF(L^T))\right] &\leq \limsup_{h\to 0} - \inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)} \left\{\frac{1}{h} I_h(\mu) + F(\mu)\right\}. \\
&= - \inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)} \{I(\mu) + F(\mu)\},
\end{aligned}$$

where the last step, once more, follows from (6.11). Thus in order to complete the proof it suffices to prove (6.11). Now as in the proof of Lemma 3.1 of [6], one can show that for all $\nu \in \mathcal{P}(I\!\!R_+)$

$$\lim_{h\to 0} \frac{1}{h} I_{h,1}(\nu) = I_1(\nu) \tag{6.12}$$

and

$$\frac{1}{h} I_{h,1}(\nu) \leq I_1(\nu). \tag{6.13}$$

A standard argument (see, e.g., the proof of [9, Lemma C.5.1]) shows that

$$\inf_{\gamma\in\mathcal{P}(D_h):\int_{D_h} x(h)\gamma(dx)=a} R(\gamma \,\|\, \theta_h) = \frac{h}{2} \frac{(a-b)^2}{\sigma^2}.$$

Thus,

$$\begin{aligned}
J_h &= \inf_{a\geq 0} \frac{h}{2} \frac{(a-b)^2}{\sigma^2} \\
&= \frac{[(b)^-]^2 h}{2\sigma^2}.
\end{aligned}$$

Let $\epsilon > 0$ be arbitrary and let $\nu^* \in \mathcal{P}(\bar{I\!\!R}_+)$ be such that

$$I(\nu^*) + F(\nu^*) \leq \inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)} \{I(\mu) + F(\mu)\} + \epsilon. \tag{6.14}$$

Note that

$$\inf_{\mu\in\mathcal{P}(\bar{I\!\!R}_+)} \left\{\frac{1}{h} I_h(\mu) + F(\mu)\right\}$$

$$
\begin{aligned}
&\leq \frac{1}{h}\left(\nu^*(I\!R_+)I_{h,1}(\hat{\nu}^*) + (1-\nu^*(I\!R_+))\frac{[(b)^-]^2}{2\sigma^2}h\right) + F(\nu^*) \\
&\leq \nu^*(I\!R_+)I_1(\hat{\nu}^*) + (1-\nu^*(I\!R_+))\frac{[(b)^-]^2}{2\sigma^2} + F(\nu^*) \\
&= \inf_{\mu\in\mathcal{P}(\bar{I\!R}_+)}\{I(\mu)+F(\mu)\} + \epsilon,
\end{aligned}
$$

where the second inequality follows from (6.13). Since $\epsilon > 0$ is arbitrary, we have that

$$
\inf_{\nu\in\mathcal{P}(\bar{I\!R}_+)}\{I(\nu)+F(\nu)\} \geq \limsup_{h\to 0}\inf_{\nu\in\mathcal{P}(\bar{I\!R}_+)}\left\{\frac{I_h(\nu)}{h}+F(\nu)\right\}. \tag{6.15}
$$

Now we prove the reverse inequality. Let $\epsilon > 0$ be arbitrary and let $\nu_h \in \mathcal{P}(\bar{I\!R}_+)$ be such that

$$
\frac{I_h(\nu_h)}{h} + F(\nu_h) \leq \inf_{\nu\in\mathcal{P}(\bar{I\!R}_+)}\left\{\frac{I_h(\nu)}{h}+F(\nu)\right\} + \epsilon. \tag{6.16}
$$

From (6.10) and [6] (see page 34) we have that for all $u \in \mathcal{H}^+$ and $\mu \in \mathcal{P}(I\!R_+)$

$$
\begin{aligned}
\frac{I_{h,1}(\mu)}{h} &\geq -\frac{1}{h}\int_{I\!R_+}\log\left(\frac{T_h u}{u}\right)(x)\mu(dx) \\
&= -\int_{I\!R_+}\left(\frac{Au}{u}\right)(x)\mu(dx) + o(1),
\end{aligned}
$$

where the term $o(1)$ may depend on $u$ but is independent of $\mu$. Since $\mathcal{P}(\bar{I\!R}_+)$ is compact, we can assume without loss of generality that $\nu_h$ converges weakly to some $\nu \in \mathcal{P}(\bar{I\!R}_+)$. Choose $u \in \mathcal{H}^+$ such that

$$
-\int_{\bar{I\!R}_+}\left(\frac{Au}{u}\right)(x)\nu(dx) \geq I(\nu) - \epsilon.
$$

Then

$$
\begin{aligned}
\liminf_{h\to 0}\inf_{\mu\in\mathcal{P}(\bar{I\!R}_+)}\left\{\frac{I_h(\mu)}{h}+F(\mu)\right\}+\epsilon &\geq \liminf_{h\to 0}\left\{\frac{I_h(\nu_h)}{h}+F(\nu_h)\right\} \\
&\geq \liminf_{h\to 0}\Bigg\{-\nu_h(I\!R_+)\int_{I\!R_+}\left(\frac{Au}{u}\right)(x)\hat{\nu}_h(dx) \\
&\quad + (1-\nu_h(I\!R_+))\frac{[(b)^-]^2}{2\sigma^2} + F(\nu_h)\Bigg\} \\
&= \liminf_{h\to 0}\left\{-\int_{\bar{I\!R}_+}\left(\frac{Au}{u}\right)(x)\nu_h(dx) + F(\nu_h)\right\} \\
&= \left\{-\int_{\bar{I\!R}_+}\left(\frac{Au}{u}\right)(x)\nu(dx) + F(\nu)\right\} \\
&\geq I(\nu) + F(\nu) - \epsilon \\
&\geq \inf_{\nu\in\mathcal{P}(\bar{I\!R}_+)}\{I(\nu)+F(\nu)-\epsilon\},
\end{aligned}
$$

where in the fourth line of the display we have denoted the continuous extension of function $\frac{Au}{u}$ to $\bar{I\!\!R}_+$ by the same notation. Such a continuous extension is uniquely defined in view of the constraint on the test functions in $\mathcal{H}^+$. The fifth line of the display follows on using the weak convergence of $\nu_h$ to $\nu$ and the continuity and boundedness of $\frac{Au}{u}$ and $F$. The above inequality along with (6.15) proves (6.11) and hence the result. ■

We now consider the empirical measure LDP for the M/M/1 queue. Strictly speaking, the discrete time approximations to this model do not satisfy the transitivity condition (Assumption 2.2) for all $x \in I\!\!R_+$. However, if $X(0)$ is an integer then $X(t)$ is an integer for all $t \in [0, \infty)$, and it is easy to check that the corresponding condition is satisfied for all $x_1$ and $x_2$ in $I\!\!N_0$.

Let $N_1$, $N_2$ be independent Poisson processes with constant rates $\lambda$ and $\mu$ respectively. Define

$$X(t) \doteq \Gamma(N_1 - N_2)(t), \quad t \in [0, \infty),$$

where $\Gamma$ is, as before, the Skorohod map. Define the occupation measure $L^T$ via (6.1). Let $A$ be the generator of the Markov process $X(\cdot)$. For $f : I\!\!N_0 \mapsto I\!\!R$, a bounded function, $(Af) : I\!\!N_0 \mapsto I\!\!R$ is the map given as

$$(Af)(x) = \begin{cases} \lambda(f(x+1) - f(x)) + \mu(f(x-1) - f(x)) & \text{if } x \in I\!\!N \\ \lambda(f(x+1) - f(x)) & \text{if } x = 0. \end{cases}$$

Define

$$\mathcal{H}^+ \doteq \left\{ u : I\!\!N_0 \to I\!\!R : \inf_{u \in I\!\!N_0} u(x) > 0, \ \lim_{x \to \infty} \frac{Au}{u}(x) = -((\sqrt{\lambda} - \sqrt{\mu})^-)^2 \right\}.$$

For $\nu \in \mathcal{P}(I\!\!R_+)$ with $\nu(I\!\!N_0) = 1$, define $I_1(\nu)$ via (6.3). Now the rate function $I(\cdot)$ governing the empirical measure LDP for $X(\cdot)$ is given by

$$I(\nu) = \begin{cases} \nu(I\!\!R_+) I_1(\hat{\nu}) + (1 - \nu(I\!\!R_+))((\sqrt{\lambda} - \sqrt{\mu})^-)^2 & \text{if } \hat{\nu}(I\!\!N_0) = 1 \\ \infty & \text{otherwise.} \end{cases} \tag{6.17}$$

The large deviation result for $\{L^T, T \in (0, \infty)\}$ is now given as follows.

**Theorem 6.3** *Let $I(\cdot)$ be defined as in (6.17). Then for all $F \in C_b(\mathcal{P}(\bar{I\!\!R}_+))$ and $x \in I\!\!N_0$*

$$\lim_{T \to \infty} \frac{1}{T} \log I\!\!E_x \left[ \exp(-T F(L^T)) \right] = - \inf_{\mu \in \mathcal{P}(\bar{I\!\!R}_+)} \{ I(\mu) + F(\mu) \}. \tag{6.18}$$

*Furthermore, $I$ is a rate function.*

The proof is very similar to the proof of Theorem 6.1 and therefore omitted.

**Remark 6.4** We expect that the techniques developed in this paper can also be applied to analogous processes on $I\!\!R$ (rather than $I\!\!R_+$), which leads to the following conjecture.

Let $X$ be a Brownian motion with drift $b$ and diffusion $\sigma$, and without loss of generality assume $b \leq 0$. Let $L^T$ denote the empirical measure for $X$. We consider the two-point compactification of $I\!R$, and denote a probability measure on this space by $\nu = (\hat{\nu}, a_1, a_2)$, where $\hat{\nu}$ is a sub-probability measure on $I\!R$, $a_1$ is the mass $\nu$ places on $-\infty$, and $a_2$ is the mass placed on $\infty$. Our conjecture is that $\{L^T, T \in (0,\infty)\}$ satisfies a LDP, with the rate function

$$I(\nu) \doteq -\inf_{u \in \mathcal{H}^+} \int_{I\!R} \left(\frac{Au}{u}\right)(x)\hat{\nu}(dx) + a_2 \frac{b^2}{2\sigma^2},$$

where

$$\mathcal{H}^+ \doteq \left\{ u \in C^2(I\!R) : \inf_{x \in I\!R} u(x) > 0, \lim_{x\to\infty} \frac{Au}{u}(x) = -\frac{b^2}{2\sigma^2}, \lim_{x\to-\infty} \frac{Au}{u}(x) = 0 \right\},$$

where $A$ is given by (6.2).

# 7 Appendix

**Proof of Lemma 3.3.** In order to prove the lemma, it suffices to prove the tightness of the family

$$\left\{\bar{L}^n,\ n \in I\!N\right\},$$

of $\mathcal{P}(\bar{I\!R}_+)$-valued random variables and the tightness of the following families of $\mathcal{P}(D)$-valued random variables:

$$\left\{ \frac{1}{n} \sum_{j=0}^{n-1} \bar{\nu}_j^n(dy \mid \bar{X}_j^n, \bar{L}_j^n),\ n \in I\!N \right\}, \tag{7.1}$$

$$\left\{ \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\bar{Z}_j^n},\ n \in I\!N \right\}. \tag{7.2}$$

The tightness of the first family is immediate on observing that $\mathcal{P}(\bar{I\!R}_+)$ is a compact Polish space.

The proofs of tightness of (7.1) and (7.2) are essentially the same, and so we give details for just the harder case of (7.2). In order to show tightness, it suffices to show that the family

$$\{m^n(dz), n \in I\!N\}$$

is tight, where

$$m^n(dz) \doteq \frac{1}{n} \sum_{j=0}^{n-1} \bar{P}_x(\bar{Z}_j^n \in dz).$$

From Theorem 13.2 of [1] and an application of Chebychev's inequality it follows that, in order to prove this statement it suffices to show

$$\sup_{n \in I\!N} \int_D \|x\| m_n(dx) < \infty \tag{7.3}$$

and

$$\sup_{n\in\mathbb{N}}\int_D w'(x,\delta)m_n(dx)\to 0, \text{ as } n\to\infty, \tag{7.4}$$

where $w'(x,\delta)$ is the usual modulus of continuity in the Skorohod space (cf. [1] page 122). The proof of (7.3) is similar to that of (7.4), and so we only consider the latter. We now prove (7.4). For $k\in(0,\infty)$, $\alpha,\delta\in(0,\infty)$ and $y\in D$, let $c_k(y,\alpha,\delta)\equiv c_k(y)\doteq\min\{\alpha w'(y,\delta),k\}$. Then for $j\in\{0,1,\ldots,n-1\}$,

$$\begin{aligned}
\bar{\mathbb{E}}_x c_k(\bar{Z}^n_j) &= \bar{\mathbb{E}}_x\left[\bar{\mathbb{E}}_x\left[c_k(\bar{Z}^n_j)\mid(\bar{X}^n_j,\bar{L}^n_j)\right]\right]\\
&= \bar{\mathbb{E}}_x\left[\int_D c_k(z)\bar{\nu}^n_j(dz\mid\bar{X}^n_j,\bar{L}^n_j)\right]\\
&= \bar{\mathbb{E}}_x\left[\int_D c_k(z)\bar{\nu}^n_j(dz\mid\bar{X}^n_j,\bar{L}^n_j)-\log\int_D e^{c_k(z)}\theta(dz)\right]\\
&\quad+\log\int_D e^{c_k(z)}\theta(dz)\\
&\le \bar{\mathbb{E}}_x\left[R(\bar{\nu}^n_j(\cdot\mid\bar{X}^n_j,\bar{L}^n_j)\parallel\theta(\cdot))\right]+M(\alpha,\delta),
\end{aligned}$$

where the last inequality follows from the variational representation for relative entropy (e.g., Lemma 1.4.3 (a) of [9]) and $M(\alpha,\delta)\doteq\log\int_D e^{\alpha w'(z,\delta)}\theta(dz)$. Summing over $j=0,1,\ldots,n-1$,

$$\begin{aligned}
\bar{\mathbb{E}}_x\left[\sum_{j=0}^{n-1}c_k(\bar{Z}^n_j)\right] &\le \bar{\mathbb{E}}_x\left[\sum_{j=0}^{n-1}R(\bar{\nu}^n_j(\cdot\mid\bar{X}^n_j,\bar{L}^n_j)\parallel\theta(\cdot))\right]+M(\alpha,\delta)n\\
&\le (\Delta+M(\alpha,\delta))n.
\end{aligned}$$

Thus

$$\int_D c_k(x)m_n(dx)=\frac{1}{n}\bar{\mathbb{E}}_x\left[\sum_{j=0}^{n-1}c_k(\bar{Z}^n_j)\right]\le\Delta+M(\alpha,\delta).$$

Taking $k\to\infty$, we have that

$$\sup_n\int_D w'(x,\delta)m_n(dx)\le\frac{\Delta}{\alpha}+\frac{M(\alpha,\delta)}{\alpha}. \tag{7.5}$$

Observing that $w'(x,\delta)\to 0$ as $\delta\to 0$ and that $w'(x,\delta)\le 2\alpha||x||$ we have from Assumption 2.1 and an application of dominated convergence theorem that for all $\alpha\in(0,\infty)$, $M(\alpha,\delta)\to 0$ as $\delta\to 0$. Now (7.4) follows on taking $\delta\to 0$ and then taking $\alpha\to\infty$ in (7.5). This proves the tightness of the family in (7.2). ∎

**Proof of part (c) of Theorem 4.1.** As noted previously, the proof that the rate function has compact level sets is essentially a simplified version of the proof of the Laplace principle upper bound, and is included here only for completeness.

Since $\mathcal{P}(\bar{\mathbb{R}}_+)$ is compact, the set $\{\pi\in\mathcal{P}(\bar{\mathbb{R}}_+): I(\pi)\le M\}$ is relatively compact for all $M\in(0,\infty)$. Thus compactness of this set follows if the rate function is lower semi-continuous. Let $\pi_n\in\mathcal{P}(\bar{\mathbb{R}}_+)$ be such that $\pi_n\to\pi\in\mathcal{P}(\bar{\mathbb{R}}_+)$. We need to show that

$$I(\pi)\le\liminf_{n\to\infty}I(\pi_n).$$

We can assume without loss of generality that

$$\liminf_{n\to\infty} I(\pi_n) < \infty.$$

Choose a subsequence $\{n_k\}_{k\in\mathbb{N}}$ such that

$$\sup_k I(\pi_{n_k}) \doteq L < \infty$$

and

$$\lim_{k\to\infty} I(\pi_{n_k}) = \liminf_{n\to\infty} I(\pi_n).$$

Henceforth, we denote the subsequence $\{\pi_{n_k}\}$ by $\{\pi_k\}$. Let $\pi_k$ have the decomposition

$$\pi_k = \rho_k \hat{\pi}_k + (1-\rho_k)\delta_\infty.$$

From (b) of Theorem 4.1, there exists $q_k \in \mathcal{A}_1(\hat{\pi}_k)$ and $\sigma \in \mathcal{P}_{\mathrm{tr}}(D)$ such that

$$\begin{aligned} I(\pi_k) &= \rho_k \int_{\mathbb{R}_+} R(q_k(\cdot|x) \parallel \theta(\cdot))\hat{\pi}_k(dx) + (1-\rho_k)R(\sigma \parallel \theta) \\ &= R(\tau_k \parallel \pi_k \otimes \theta), \end{aligned} \tag{7.6}$$

where $\tau_k \in \mathcal{P}(\bar{\mathbb{R}}_+ \times D)$ is defined as

$$\tau_k(A\times B) = \rho_k \int_A q_k(B|x)\hat{\pi}_k(dx) + (1-\rho_k)\sigma(B)\delta_\infty(A)$$

and the last equality in (7.6) follows from Lemma 1.4.3 (f) of [9]. Since $\bar{\mathbb{R}}_+$ is compact and $\sup_k R(\tau_k \parallel \pi_k \otimes \theta) < \infty$, we have as in the proof of (b) that $\{\tau_k, k \in \mathbb{N}\}$ is relatively compact in $\mathcal{P}(\bar{\mathbb{R}}_+ \times D)$. Now assume without loss of generality that $\tau_k \to \tau \in \mathcal{P}(\bar{\mathbb{R}}_+ \times D)$ as $k \to \infty$. Observe that since $(\tau_k)_1 = \pi_k$, we have that $(\tau)_1 = \pi$. Using the lower semi-continuity of relative entropy, it follows from (7.6) that

$$L \;\geq\; \lim_{k\to\infty} I(\pi_k) \;\geq\; R(\tau \parallel \pi \otimes \theta). \tag{7.7}$$

Now let $q^* \in \mathcal{S}(\bar{\mathbb{R}}_+ \parallel D)$ be such that $\tau(dx\, dz) = q^*(dz|x)\pi(dx)$ and let the decomposition for $\pi$ be given as

$$\pi = \rho\hat{\pi} + (1-\rho)\delta_\infty.$$

Using Lemma 1.4.3 (f) of [9], once more, we have that

$$R(\tau \parallel \pi\otimes\theta) = \rho\int_{\mathbb{R}_+} R(q^*(\cdot|x) \parallel \theta(\cdot))\hat{\pi}(dx) + (1-\rho)R(q^*(\cdot|\infty) \parallel \theta(\cdot)).$$

An argument similar to the one in the proof of Lemma 3.10 shows that if $\rho > 0$ then $q^* \in \mathcal{A}_1(\hat{\pi})$. We claim now that if $\rho < 1$ then $q^*(\cdot|\infty) \in \mathcal{P}_{\mathrm{tr}}(D)$. Observe that once the claim is proved, the lower semicontinuity follows on noting that

$$\begin{aligned} &\lim_{k\to\infty} I(\pi_k) \\ &\quad\geq\; R(\tau \parallel \pi\otimes\theta) \end{aligned}$$

$$
\begin{aligned}
&= \rho \int_{I\!R_+} R(q^*(\cdot|x) \parallel \theta(\cdot))\hat{\pi}(dx) + (1-\rho) R(q^*(\cdot|\infty) \parallel \theta(\cdot)) \\
&\geq \rho \inf_{q^* \in \mathcal{A}_1(\hat{\pi})} \int_{I\!R_+} R(q^*(\cdot|x) \parallel \theta(\cdot))\hat{\pi}(dx) + (1-\rho) \inf_{\sigma \in \mathcal{P}_{\rm tr}(D)} R(\sigma \parallel \theta) \\
&= \rho I_1(\hat{\pi}) + (1-\rho) J \\
&= I(\pi).
\end{aligned}
$$

We now prove the claim. Using the observation that $\sup_k R(\tau_k \parallel \pi_k \otimes \theta) \leq L < \infty$, we have as in the proof of Lemma 3.9 (cf. (3.29)) that

$$
(1-\rho) \int_D z(T) q^*(dz|\infty) = \lim_{k \to \infty} \int_{I\!R_+ \times D} z(T) F'_{c_k}(x) \tau(dx\, dz), \tag{7.8}
$$

where $\{c_k\}$ is a sequence of positive reals increasing to $\infty$ and $F_c$ is as in (3.14).

Also, an argument parallel to the one in the proof of Lemma 3.7 shows that for $c \in (0, \infty)$,

$$
\int_{I\!R_+ \times D} z(T) F'_c(x) \hat{\tau}_k(dx\, dz) \geq -5 \int_D ||z|| 1_{||z|| \geq \frac{c}{2}} (\hat{\tau}_k)_2(dz) - \frac{2}{c} \int_D ||z|| (\hat{\tau}_k)_2(dz). \tag{7.9}
$$

The main difference from the proof of Lemma 3.7 is that instead of using the inequality

$$
\frac{1}{n} \sum_{i=0}^{n-1} \left( F_c(\bar{X}^n_{i+1}) - F_c(\bar{X}^n_i) \right) \geq -\frac{F_c(\bar{X}^n_0)}{n}
$$

as in (3.19), one uses the invariance property of $\hat{\tau}_k$ to conclude that

$$
\int_{I\!R_+ \times D} \left( F_c(\Pi_T(x,z)) - F_c(x) \right) \hat{\tau}_k(dx\, dz) = 0.
$$

Because of this invariance property, unlike in (3.17), one does not obtain the third "residual term" in (7.9). Hence

$$
\begin{aligned}
&\int_{\bar{I\!R}_+ \times D} z(T) F'_c(x) \tau_k(dx\, dz) \qquad (7.10) \\
&= \rho_k \int_{I\!R_+ \times D} z(T) F'_c(x) \hat{\tau}_k(dx\, dz) + (1-\rho_k) \int_D z(T) \sigma(dz) \\
&\geq -5\rho_k \int_D ||z|| 1_{||z|| \geq \frac{c}{2}} (\hat{\tau}_k)_2(dz) - \frac{2\rho_k}{c} \int_D ||z|| (\hat{\tau}_k)_2(dz) + (1-\rho_k) \int_D z(T) \sigma(dz) \\
&\geq -5 \int_D ||z|| 1_{||z|| \geq \frac{c}{2}} (\tau_k)_2(dz) - \frac{2}{c} \int_D ||z|| (\tau_k)_2(dz), \qquad (7.11)
\end{aligned}
$$

where the last step follows on recalling that $\int_D z(T)\sigma(dz) \geq 0$ for $\sigma \in \mathcal{P}_{\rm tr}(D)$. Now it follows via arguments similar to those in the proof of Lemma 3.8 (See (3.23), (3.24) and (3.25)) that there exists a sequence of positive reals $\{c_m\}$ such that $c_m \to \infty$ as $m \to \infty$, and for all $m \in I\!N$,

$$
\int_{\bar{I\!R}_+ \times D} z(T) F'_{c_m}(x) \tau_k(dx\, dz) \to \int_{\bar{I\!R}_+ \times D} z(T) F'_{c_m}(x) \tau(dx\, dz),
$$

$$\int_D ||z|| 1_{||z|| \geq \frac{c_m}{2}} (\tau_k)_2(dz) \to \int_D ||z|| 1_{||z|| \geq \frac{c_m}{2}} (\tau)_2(dz)$$

and

$$\int_D ||z|| (\tau_k)_2(dz) \to \int_D ||z|| (\tau)_2(dz)$$

as $k \to \infty$. Now taking the limit as $k \to \infty$ in (7.11) with $c$ replaced by $c_m$, we have that

$$\int_{\bar{\mathbb{R}}_+ \times D} z(T) F'_{c_m}(x) \tau(dx\, dz) \geq -5 \int_D ||z|| 1_{||z|| \geq \frac{c_m}{2}} (\tau)_2(dz) - \frac{2}{c_m} \int_D ||z|| (\tau)_2(dz).$$

Finally taking limit as $m \to \infty$, we have from (7.8) that

$$(1 - \rho) \int_D z(T) q^*(dz|\infty) \geq 0.$$

This proves the claim. ∎

Notation and conventions. We summarize here for the readers convenience various notations and conventions that are used throughout the paper.

| | |
|---|---|
| $I\!N$ | the set of positive integers |
| $I\!N_0$ | the set of nonnegative integers |
| $I\!R$ | the set of reals |
| $I\!R_+$ | the set of nonnegative reals |
| $D_T$ | right continuous functions with left limits from $[0,T]$ to $I\!R$, endowed with the usual Skorohod topology |
| $C_c(I\!R_+)$ | the space of real continuous functions on $I\!R_+$ with compact support |
| $C_b^2(I\!R_+)$ | the space of real, bounded twice continuously differentiable functions on $I\!R_+$ |
| $BM(S)$ | the class of real-valued bounded measurable functions on $S$ |
| $C_b(S)$ | the subclass of $BM(S)$ of continuous functions on $S$ |
| $\mathcal{B}(S)$ | the Borel $\sigma$-field on a complete separable metric space $S$ |
| $\mathcal{P}(S)$ | the space of probability measures on $(S,\mathcal{B}(S))$ with the weak convergence topology |
| $D([0,T]:\mathcal{S})$ | the Skorohod space of $\mathcal{S}$-valued right continuous functions with left hand limits |
| $D_+([0,T]:I\!R)$ | the subspace of $D([0,\infty):I\!R)$ of functions $f$ satisfying $f(0)\in I\!R_+$ |
| $\nu_n \Rightarrow \nu$ | the sequence $\nu_n$ in $\mathcal{P}(S)$ converges weakly to $\nu$ |
| $\mathcal{M}(S)$ | denotes the space of positive measures on $(S,\mathcal{B}(S))$ with total mass not exceeding 1, with the topology of weak convergence of measures |
| $\langle f,\nu\rangle$ | $\int_S f(x)d\nu(x)$ for $f\in BM(S)$, $\nu\in\mathcal{P}(S)$ |
| $\|\|g\|\|_\infty$ | $\sup_{x\in S}\|g(x)\|$ for $g\in C_b(S)$ |
| $\delta_x$ | the probability measure which is concentrated at the point $x\in S$ |
| $1_A$ | the indicator function of the set $A$ |

A function $I$ mapping $S$ into $[0,\infty]$ is called a rate function if for all $M\in[0,\infty)$ the level set $\{x\in S : I(x)\le M\}$ is compact. The infimum over an empty set is taken to be $\infty$. As another convention, $0\times\infty$ is taken to be 0. A family of random variables with values in a Polish space is said to be tight if the corresponding family of probability laws is tight.

**Acknowledgments.** We thank the referee for a careful study of the manuscript and for suggesting several improvements.

# References

[1] P. Billingsley. *Convergence of Probability Measures.* John Wiley & Sons 1999.

[2] T.S. Chiang. A lower bound of the asymptotic behavior of some Markov processes. *Annals of Probability*, 10:955–967, 1982.

[3] A. de Acosta. Large deviations of empirical measures of Markov chains. *J. Theoret. Prob.*, 3:395–431, 1990.

[4] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Academic Press, San Diego, Calif., 1989.

[5] J-D. Deuschel and D. Stroock. *Large Deviations*. Academic Press, San Diego, Calif., 1989.

[6] M. Donsker and S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time, I. *Comm. Pure Appl. Math.*, 28:1–45, 1975.

[7] M. Donsker and S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time, III. *Comm. Pure Appl. Math.*, 29:389–461, 1976.

[8] R. M. Dudley. *Probabilities and Metrics*. AArhus Universitet, Mathematisk Institut, Lecture Notes Series., 1989.

[9] P. Dupuis and R. S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. John Wiley and Sons, Inc., 1997.

[10] N. Jain. Large deviation lower bounds for additive functionals of Markov processes. *Annals of Probability*, 18:1071–1098, 1990.

[11] H. J. Kushner. *Stochastic Stability and Control*. Academic Press, New York, 1967.

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

[13] P. Ney and E. Nummelin. Markov additive processes, I: eigenvalue properties and limit theorems. *Annals of Probability*, 15:561–592, 1987.

[14] P. Ney and E. Nummelin. Markov additive processes, II: large deviations. *Annals of Probability*, 15:593–609, 1987.