

LARGE DEVIATIONS FOR MARKOV PROCESSES WITH DISCONTINUOUS STATISTICS, I: GENERAL UPPER BOUNDS

BY PAUL DUPUIS,¹ RICHARD S. ELLIS² AND ALAN WEISS

*University of Massachusetts, Amherst, University of Massachusetts,
Amherst and AT & T Bell Laboratories*

In this paper we prove an upper large deviation bound for a general class of Markov processes, which includes processes with discontinuous statistics. We also specialize the results to a class of jump Markov processes that model scaled queueing systems.

1. Introduction. In order to prove that a stochastic process satisfies the large deviation principle, one must establish an upper bound for closed sets and a lower bound for open sets. For many systems, the upper bound, at least for compact sets, holds in some generality. This state of affairs is illustrated, for example, in papers by Donsker and Varadhan [3], Gärtner [9], Ellis [6] and de Acosta [2]. By contrast, the lower bound often requires an analysis on a case-by-case basis.

The purpose of the present paper is to establish the upper large deviation bound for a general class of Markov processes taking values in $D([0, T]; R^n)$. A key step in the proof is to adapt a result of Dupuis and Kushner [5], which is an analogue for stochastic processes of results in the last three papers referenced before. Future work will treat more refined bounds and applications.

Our results, which hold in some generality, were motivated by problems involving queueing networks. For the purpose of introduction, we shall now discuss a relatively simple case. Queueing networks involving two classes of customers or consisting of two queues may be modeled by jump Markov processes $X(t)$ taking values in the subset $R_+^2 = \{x \in R^2 : x = (x_1, x_2), x_1 \geq 0, x_2 \geq 0\}$ of R^2 . These processes have constant jump rates and directions in the interior $D^\circ = \{x \in R^2 : x_1 > 0, x_2 > 0\}$, in the boundary set $D^{(1)} = \{x \in R^2 : x_1 = 0, x_2 > 0\}$ and in the boundary set $D^{(2)} = \{x \in R^2 : x_1 > 0, x_2 = 0\}$. Jump rates and directions are also specified for $D^{(1,2)} = \{(0, 0)\}$. For each point $x \in Z_+^2$, the jump directions consist of vectors $v \in \{1, 0, -1\}^2$; the corresponding jump rates are denoted by $\lambda_v(x) \geq 0$. So that the process remains in Z_+^2 , it

Received June 1989; revised February 1990.

¹Research supported in part by a NSF Grant DMS-89-02333.

²Research supported in part by NSF Grant DMS-85-21536 and DMS-89-01138 at the University of Massachusetts; in part by the Center for Control Sciences (Contract Number F-49620-86-C-0111) while visiting the Division of Applied Mathematics at Brown University during the fall semester of 1988; and in part by a Lady Davis Fellowship while visiting the Faculty of Industrial Engineering and Management at the Technion during the spring semester of 1989.

AMS 1980 subject classifications. Primary 60F10; secondary 60J99.

Key words and phrases. Upper large deviation bound, Markov processes, queueing systems.

is stipulated that $\lambda_v(x) = 0$ whenever $x \in Z_+^2$ and $x + v \notin Z_+^2$. More general processes will be considered in Section 3.

In order to study large deviation phenomena for such networks, we consider the scaled process $X^\varepsilon(t) = \varepsilon X(t/\varepsilon)$, where $\varepsilon > 0$ is a small parameter. For each $x \in R_+^2$, this process has the infinitesimal generator

$$(1) \quad \mathcal{L}^\varepsilon f(x) = \frac{1}{\varepsilon} \sum_{v \in \{1, 0, -1\}^2} \lambda_v(x) [f(x + \varepsilon v) - f(x)].$$

For each $s = \emptyset, \{1\}, \{2\}$, whenever x and y are in D^s , we have $\lambda_v(x) = \lambda_v(y)$. However, the sets of jump rates in the four regions are in general all different. Since the function $x \rightarrow \lambda_v(x)$ is in general discontinuous, we call $X^\varepsilon(t)$ a Markov process with ‘discontinuous statistics’.

Clearly the definition of $X^\varepsilon(t)$ may be extended to give a process in all of R^2 . The main result of this paper, Theorem 1.1, is an upper large deviation bound for a class of Markov processes which includes the scaled queueing process as a special case. The first paper to study the large deviation properties of such processes with discontinuous statistics appears to be [4]. This work uses techniques from the theory of viscosity solutions of Hamilton–Jacobi equations to study processes that model a class of queueing systems known as Jackson networks. One of the motivations of the present work is to extend the upper bound results of [4] to cover as large a class of Markov processes as possible. As our future work will show, in the presence of special geometries (e.g., one smooth boundary separating two regions of smooth statistical behavior) our upper bound can sometimes be improved.

Assumptions on the process. We first define the class of Markov processes under consideration. Fix $n \in \{1, 2, \dots\}$. The following quantities are given.

1. A bounded measurable vector-valued function $b(x) = (b_1(x), \dots, b_n(x))$ of $x \in R^n$.
2. A bounded measurable $n \times n$ matrix-valued function $a(x) = \{a_{ij}(x), i, j = 1, \dots, n\}$ of $x \in R^n$.
3. A measurable function μ_x mapping points $x \in R^n$ to measures on R^n such that

$$(2) \quad \mu_x(K^c) = 0 \text{ for each } x \in R^n \text{ and } \sup_{x \in R^n} \mu_x(K) < \infty$$

for some compact subset K in R^n .

For $\varepsilon > 0$, we define an operator \mathcal{L}^ε on twice continuously differentiable functions with compact support by the formula

$$(3) \quad \begin{aligned} \mathcal{L}^\varepsilon f(x) = & \sum_{i=1}^n b_i(x) \frac{\partial f(x)}{\partial x_i} + \frac{\varepsilon}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \\ & + \frac{1}{\varepsilon} \int_{R^n \setminus \{0\}} \left[f(x + \varepsilon v) - f(x) - \varepsilon \sum_{i=1}^n v_i \frac{\partial f(x)}{\partial x_i} \right] \mu_x(dv). \end{aligned}$$

Under the assumption that for $T > 0$ there exists a Markov process $\{X^\varepsilon(t), 0 \leq t \leq T\}$ corresponding to the infinitesimal generator \mathcal{L}^ε (in the sense that the process solves the martingale problem [13] for this generator) and with paths in $D([0, T]; \mathbb{R}^n)$, we will state our main result in Theorem 1.1. The existence of such processes is discussed in [13, 7, 11].

Clearly, the operator (3) generalizes the infinitesimal generator in (1) that arises in the modeling of queueing networks. With the choices $\mu_x \equiv 0$ for all $x \in \mathbb{R}^n$, $\{a_{ij}(x)\}$ smooth, and $\{b_i(x)\}$ discontinuous, the operator (3) arises in the modeling of communication channels that incorporate a “hard limiter” in a phase-locked loop (a form of a suboptimal nonlinear filter).

Some notation is needed. For x and α in \mathbb{R}^n , define

$$(4) \quad H(x, \alpha) = \sum_{i=1}^n b_i(x)\alpha_i + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x)\alpha_i\alpha_j + \int_{\mathbb{R}^n \setminus \{0\}} \left[\exp\left(\sum_{i=1}^n \alpha_i v_i\right) - 1 - \sum_{i=1}^n \alpha_i v_i \right] \mu_x(dv)$$

and

$$(5) \quad h(x, \alpha) = \lim_{\delta \downarrow 0} \sup_{\{y: |y-x| \leq \delta\}} H(y, \alpha).$$

The latter is the upper semicontinuous regularization of $H(\cdot, \alpha)$. Consider the Legendre–Fenchel transform

$$(6) \quad l(x, \beta) = \sup_{\alpha \in \mathbb{R}^n} [\alpha \cdot \beta - h(x, \alpha)],$$

defined for $\beta \in \mathbb{R}^n$. In terms of this, we define the functional

$$(7) \quad I_x(\phi) = \int_0^T l(\phi(s), \dot{\phi}(s)) ds$$

when ϕ is absolutely continuous and $\phi(0) = x$. In all other cases, we set $I_x(\phi) = +\infty$.

THEOREM 1.1. *For $T > 0$ and $\varepsilon > 0$, we assume that there exists a Markov process $\{X^\varepsilon(t), 0 \leq t \leq T\}$ corresponding to the infinitesimal generator \mathcal{L}^ε and with paths in $D([0, T]; \mathbb{R}^n)$. Let a compact set C be given. Then the following conclusions hold:*

(i) Define

$$\Phi_x(L) = \{\phi \in D([0, T]; \mathbb{R}^n) : I_x(\phi) \leq L\}.$$

Then for all $L < \infty$, the set $\cup_{x \in C} \Phi_x(L)$ is compact.

(ii) For each closed set F in $D([0, T]; \mathbb{R}^n)$,

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{X^\varepsilon \in F\} \leq - \inf_{\phi \in F} I_x(\phi),$$

uniformly in $x \in C$.

Remark 1.2 points out how our assumptions on the quantities appearing in \mathcal{L}^ε may be weakened so that Theorem 1.1 remains true. The theorem will be proved in Section 2 of this paper. In Section 3, we will consider special features of the theorem when specialized to jump Markov processes that model n -dimensional queueing networks. In particular, we will obtain a useful representation for the Lagrangian $l(x, \beta)$ appearing in (6) (see Theorem 3.1). In Section 4, we will consider the analogous theorem for discrete time processes. In Section 5, we will conclude with an informal discussion on the conditions under which one may expect the upper bounds obtained here to be best.

Theorem 1.1 refers to Markov processes with discontinuous statistics. Large deviations of analogous Markov processes with continuous statistics have been studied in a series of papers by Wentzell beginning in 1976 ([14, 15, 16, 17]). He assumes a quasi-uniform continuity condition on the function

$$L(x, \beta) = \sup_{\alpha \in R^n} [\alpha \cdot \beta - H(x, \alpha)],$$

where $H(x, \alpha)$ is defined as in (4). This condition, which is somewhat difficult to verify in practice, implies the continuity of the function $x \rightarrow H(x, \alpha)$. Under this and other assumptions, he proves the upper and lower large deviation bounds for the process with infinitesimal generator \mathcal{L}^ε . Related results are discussed in [8] (Section 5.2).

Here are the main steps in the proof of part (ii) of Theorem 1.1. The essential facts used are that

$$(8) \quad \exp\left[\frac{1}{\varepsilon} \left\{ (X^\varepsilon(t) - X^\varepsilon(s)) \cdot \alpha - \int_s^t H(X^\varepsilon(u), \alpha) du \right\}\right]$$

is a martingale and that for fixed α , $H(x, \alpha)$ is bounded independently of x .

1. Prove that as $\varepsilon \rightarrow 0$, $\{X^\varepsilon(t), 0 \leq t \leq T\}$ is superexponentially close to a piecewise linear approximation $\{Y^\varepsilon(t), 0 \leq t \leq T\}$ (Lemma 2.4).
2. Prove that $\{Y^\varepsilon(t), 0 \leq t \leq T\}$ is exponentially tight (Lemma 2.5).
3. Verify that for $\delta > 0$,

$$\limsup_{\varepsilon \rightarrow 0} \log E_x\{\exp[(Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha/\varepsilon]\} \leq h^\delta(x, \alpha)$$

uniformly for x in compact subsets of R^n , where

$$h^\delta(x, \alpha) = \sup_{\{y: |y-x| \leq \delta\}} H(y, \alpha).$$

This is proved in Lemma 2.7.

4. Adapt a result of Dupuis and Kushner [5] to show that Step 3 implies the upper large deviation bound for $\{Y^\varepsilon(t), 0 \leq t \leq T\}$ in compact sets, with upper rate function I_x^δ defined as in (7) with $l(x, \beta)$ replaced by the Legendre–Fenchel transform of $h^\delta(x, \alpha)$ (Proposition 2.9).
5. Justify the passage to the limit $\delta \rightarrow 0$ (Proposition 2.10).

Part (ii) of Theorem 1.1 follows from steps 1, 2 and 5.

REMARK 1.2. We now discuss how our assumptions on the quantities appearing in the infinitesimal generator \mathcal{L}^ε may be weakened and yet leave Theorem 1.1 true. The assumption (2) on the measures $\{\mu_x, x \in R^n\}$ guarantees that there exist positive constants C_1 and C_2 such that

$$(9) \quad H(x, \alpha) \leq \bar{h}(|\alpha|) \equiv C_1 e^{C_2|\alpha|}$$

for all $\alpha \in R^n$. The specific form of \bar{h} is used only in the proof of step 2, the exponential tightness of $\{Y^\varepsilon(t), 0 \leq t \leq T\}$ (see Lemma 2.5). We comment further on this later.

Instead of (2) and (9), let us assume that for each x and α in R^n ,

$$(10) \quad \int_{R^n \setminus \{0\}} |v|^2 \mu_x(dv) < \infty \quad \text{and} \quad H(x, \alpha) \leq \tilde{h}(|\alpha|),$$

where \tilde{h} is some finite function. It follows that if \tilde{l} denotes the Legendre–Fenchel transform of \tilde{h} , then \tilde{l} grows superlinearly. According to [8] (pages 144–148), the process in (8) remains a martingale and the proofs in steps 1, 3, 4 and 5 remain valid with only minor modifications. We indicate in Remark 2.6 how to modify the proof of step 2. It follows that Theorem 1.1 remains true under the hypotheses in (10).

Theorem 1.1 also remains true if \mathcal{L}^ε in (3) is replaced by

$$\begin{aligned} \tilde{\mathcal{L}}^\varepsilon f(x) &= \sum_{i=1}^n b_i^\varepsilon(x) \frac{\partial f(x)}{\partial x_i} + \sum_{i,j=1}^n a_{ij}^\varepsilon(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \\ &+ \int_{R^n \setminus \{0\}} \left[f(x+v) - f(x) - \sum_{i=1}^n v_i \frac{\partial f(x)}{\partial x_i} \right] \mu_x^\varepsilon(dv), \end{aligned}$$

where $\{b_i^\varepsilon(x)\}$ and $\{a_{ij}^\varepsilon(x)\}$ are suitable ε -dependent coefficients and $\{\mu_x^\varepsilon\}$ are suitable ε -dependent measures. In the absence of specific applications, we will not bother to indicate sufficient conditions that these quantities must satisfy in order that the theorem remain valid.

2. Proof of Theorem 1.1. Throughout this section, H , h and l are defined by (4), (5) and (6), respectively, and \bar{h} is a function satisfying (9). For the sake of notational simplicity, we take $T = 1$.

Note that part (i) of Theorem 1.1 implies that the function $I_x(\phi)$ is lower semicontinuous. This fact will be used often in the sequel.

PROPOSITION 2.1. *For any compact set $C \subset R^n$ and $L < \infty$, the set $\bigcup_{x \in C} \Phi_x(L)$ is compact.*

PROOF. The upper semicontinuity properties of the function $h(x, \alpha)$ imply that $l(x, \beta)$ is jointly lower semicontinuous (see the proof of Proposition 2.10). Using the fact that $h(x, \alpha) \leq \bar{h}(|\alpha|)$, the proposition follows from the corollary to Theorem 1 of Section 9.1.3 and Theorem 3 of Section 9.1.4 of [10]. \square

Before proving part (ii) of Theorem 1.1, we present several lemmas.

LEMMA 2.2. For all $\alpha \in R^n$, all $x \in R^n$, and all $t \geq s \geq 0$,

$$N_\alpha^\varepsilon(t) = \exp\left[\frac{1}{\varepsilon}\left\{(X^\varepsilon(t) - X^\varepsilon(s)) \cdot \alpha - \int_s^t H(X^\varepsilon(u), \alpha) du\right\}\right]$$

is a P_x -martingale in t .

The proof is a straightforward adaptation of the proof of [13] (Theorem 4.2.1).

Next define (for $b > C_3 \equiv C_1C_2$)

$$\begin{aligned} \bar{l}(b) &= \sup_{a \in R} [ab - \bar{h}(a)] \\ &= \sup_{a > 0} [ab - \bar{h}(a)] \\ &= \frac{1}{C_2} b(\log(b/C_3) - 1). \end{aligned}$$

Note that $\bar{l}(b)$ grows in a superlinear fashion as $b \rightarrow +\infty$.

LEMMA 2.3. Given $\delta > 0$ and $s \geq 0$,

$$P_x\left\{\sup_{s \leq u \leq t} |X^\varepsilon(u) - X^\varepsilon(s)| \geq \delta\right\} \leq 2n \exp - [(t - s)\bar{l}(\delta/2n^{1/2}(t - s))/\varepsilon],$$

whenever $\delta/2n^{1/2}(t - s) > C_3$. This inequality is valid for all $x \in R^n$.

PROOF. The proof follows the argument used in [13] (Theorem 4.2.1). Let $\alpha = (a, 0, \dots, 0) \in R^n$, where $a > 0$. Then

$$\begin{aligned} P_x\left\{\sup_{s \leq u \leq t} (X^\varepsilon(u) - X^\varepsilon(s)) \cdot \alpha \geq \delta a/2n^{1/2}\right\} \\ \leq P_x\left\{\sup_{s \leq u \leq t} N_\alpha^\varepsilon(u) \geq \exp\left[\frac{1}{\varepsilon}\left\{(\delta a/2n^{1/2}) - (t - s)\bar{h}(a)\right\}\right]\right\}. \end{aligned}$$

By Lemma 2.2, this last quantity is bounded above by

$$\exp - [((\delta a/2n^{1/2}) - (t - s)\bar{h}(a))/\varepsilon].$$

Minimizing with respect to $a > 0$, we obtain the upper bound

$$\exp - [(t - s)\bar{l}(\delta/2n^{1/2}(t - s))/\varepsilon].$$

The conclusion of the lemma follows by applying the analogous estimates to the positive and negative parts of each component of $\sup_{s \leq u \leq t} (X^\varepsilon(u) - X^\varepsilon(s))$. □

Define the piecewise linear process

$$Y^\varepsilon(t) = \frac{X^\varepsilon(i\varepsilon)(i\varepsilon + \varepsilon - t) + X^\varepsilon(i\varepsilon + \varepsilon)(t - i\varepsilon)}{\varepsilon}$$

for $t \in [i\varepsilon, i\varepsilon + \varepsilon)$ and $i = 1, 2, \dots, [1/\varepsilon]$. The next lemma will imply that when proving the upper bound it is enough to work with Y^ε . We do so because it is easier to prove an exponential tightness estimate (Lemma 2.5) for Y^ε than for X^ε .

LEMMA 2.4. For all $\delta > 0$ and uniformly in $x \in R^n$,

$$(11) \quad \limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x \left\{ \sup_{0 \leq t \leq 1} |X^\varepsilon(t) - Y^\varepsilon(t)| \geq \delta \right\} = -\infty.$$

PROOF. By Lemma 2.3, the left-hand side of (11) is bounded above by

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \left[(2n/\varepsilon) (\exp - \bar{l}(\delta/2n^{1/2}\varepsilon)) \right] \\ & \leq - \liminf_{\varepsilon \rightarrow 0} \varepsilon \bar{l}(\delta/2n^{1/2}\varepsilon) = -\infty. \end{aligned} \quad \square$$

LEMMA 2.5 (Exponential tightness). Let a compact set $C \subset R^n$ and $B < \infty$ be given. Then there exists a compact set $K \in C([0, 1], R^n)$ such that

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x \{Y^\varepsilon \notin K\} \leq -B$$

for all $x \in C$.

PROOF. Since Y^ε has continuous sample paths, we may look for a compact set in $C([0, 1]; R^n)$. Define the sets

$$K(M) = \bigcap_{m \geq M} \{ \phi \in C([0, 1]; R^n) : \phi(0) \in C, w_\phi(2^{-m}) \leq 1/\log m \},$$

where $M \geq 2$ and

$$w_\phi(\delta) = \sup_{|s-t| \leq \delta} |\phi(s) - \phi(t)|.$$

By the Arzela–Ascoli theorem, each $K(M)$ is compact. The piecewise linear nature of the process Y^ε implies that if m is any integer satisfying $2^{-m} \leq \varepsilon$, then $w_{Y^\varepsilon}(2^{-m}) \leq w_{Y^\varepsilon}(2^{-m+1})/2$. Hence, $Y^\varepsilon \in K(M)$ if

$$Y^\varepsilon \in \bigcap_{(-\log \varepsilon / \log 2) \geq m \geq M} \{ \phi \in C([0, 1]; R^n) : \phi(0) \in C, w_\phi(2^{-m}) \leq 1/\log m \}.$$

For the rest of this proof, m will be an integer such that $2^{-m} \leq \varepsilon \leq 2^{-m+1}$. We have the upper bound

$$P_x \{Y^\varepsilon \notin K(M)\} \leq \sum_{i=M}^m P_x \{w_{Y^\varepsilon}(2^{-i}) > 1/\log i\}.$$

Using Lemma 2.3 and the fact that $2^i/\log i$ is increasing for $i \geq 2$, we obtain (for M large enough so that $2^M/4n^{1/2} \log M > C_3$)

$$P_x\{Y^\varepsilon \notin K(M)\} \leq 2n(m - M + 1) \max_{M \leq i \leq m} 2^i \exp - \left[2^{-i+m-1} \bar{l}(2^i/4n^{1/2} \log i) \right].$$

Using the explicit form of \bar{l} , we have

$$\begin{aligned} & 2^{-i+m-1} \bar{l}(2^i/4n^{1/2} \log i) \\ &= 2^{m-1} [\log 2^i - \log \log i - \log(4n^{1/2} C_3) - 1] / 4C_2 n^{1/2} \log i \\ &\equiv 2^m a_i. \end{aligned}$$

If M is sufficiently large,

$$\min_{i \geq M} a_i = a_M.$$

We therefore have the bound

$$P_x\{Y^\varepsilon \notin K(M)\} \leq 2n(m - M + 1) 2^m \exp(-2^m a_M),$$

which implies

$$(12) \quad \limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{Y^\varepsilon \notin K(M)\} \leq -a_M.$$

Since $a_M \uparrow \infty$ as $M \rightarrow \infty$, the lemma is proved. \square

REMARK 2.6. We indicate how to modify the proof of Lemma 2.5 when hypotheses (3) and (9) are weakened to hypothesis (10). The sets $K(M)$ are redefined to be

$$\bigcup_{m \geq M} \{ \phi \in C([0, 1]; R^n) : \phi(0) \in C, w_\phi(2^{-m}) \leq 1/f(m) \},$$

where $f(m)$ is some function that maps the nonnegative integers to $[0, \infty)$, is monotonically increasing, and satisfies $f(m) \rightarrow \infty$ as $m \rightarrow \infty$. The proof of Lemma 2.5 may be modified provided f satisfies the condition

$$(13) \quad 2^{-i} \bar{l}(2^i/cf(i)) \rightarrow \infty \text{ as } i \rightarrow \infty$$

for $c > 0$, where \bar{l} denotes the Legendre–Fenchel transform of the function \bar{h} in (10). Since $\bar{h}(\alpha)$ is assumed to be finite for all $\alpha \geq 0$, \bar{l} grows superlinearly and so a function f satisfying (13) may always be found.

We need one more lemma before proving part (ii) of Theorem 1.1. Fix $\delta > 0$ and define

$$h^\delta(x, \alpha) = \sup_{|y-x| \leq \delta} h(y, \alpha).$$

Then $h^\delta(x, \alpha)$ is upper-semicontinuous in x , convex in α , and $h^\delta(x, \alpha) \downarrow h(x, \alpha)$ as $\delta \downarrow 0$. Furthermore, for all α , all x , and all y such that $|y - x| \leq \delta$,

$$H(y, \alpha) \leq h^\delta(x, \alpha) \leq \bar{h}(|\alpha|).$$

LEMMA 2.7. *Let compact $C \subset R^n$ be given. Then*

$$\limsup_{\varepsilon \rightarrow 0} \log E_x \exp[(Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha / \varepsilon] \leq h^\delta(x, \alpha)$$

uniformly in $x \in C$ and α in bounded sets.

PROOF. By Lemma 2.2,

$$1 = E_x \left\{ \exp \frac{1}{\varepsilon} \left[(Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha - \int_0^\varepsilon H(X^\varepsilon(u), \alpha) du \right] \right\}.$$

Hence, if $|x - y| \leq \delta/4$, then

$$\begin{aligned} 1 &\geq E_x \left\{ \left(\exp \frac{1}{\varepsilon} [(Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha - \varepsilon h^\delta(y, \alpha)] \right) 1_{A(\varepsilon, \delta)} \right\} \\ &= \exp[-h^\delta(y, \alpha)] E_x \left\{ \exp \left[\frac{1}{\varepsilon} (Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha \right] \right\} \\ &\quad - \exp[-h^\delta(y, \alpha)] E_x \left\{ \left(\exp \left[\frac{1}{\varepsilon} (Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha \right] \right) 1_{A(\varepsilon, \delta)^c} \right\}, \end{aligned}$$

where $A(\varepsilon, \delta) = \{\sup_{0 \leq u \leq \varepsilon} |X^\varepsilon(u) - x| \leq \delta/2\}$ and $A(\varepsilon, \delta)^c$ is the complement. This implies

$$\begin{aligned} E_x \left\{ \exp \left[\frac{1}{\varepsilon} (Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha \right] \right\} &\leq (\exp h^\delta(y, \alpha)) \\ &\quad + E_x \left\{ \exp \left[\frac{1}{\varepsilon} (Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha \right] \cdot 1_{A(\varepsilon, \delta)^c} \right\}. \end{aligned}$$

Using Lemma 2.3, we have the bound

$$\begin{aligned} E_x \left\{ \exp \left[\frac{1}{\varepsilon} (Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha \right] \cdot 1_{A(\varepsilon, \delta)^c} \right\} \\ \leq \sum_{k=1}^\infty \exp \left(\frac{(k+1)\delta|\alpha|}{2\varepsilon} \right) \cdot P_x \left\{ \frac{(k+1)\delta}{2} \geq \sup_{0 \leq u \leq \varepsilon} |X^\varepsilon(u) - x| \geq \frac{k\delta}{2} \right\} \\ \leq B_1 \sum_{k=1}^\infty \exp \left[\frac{(k+1)\delta|\alpha|}{2\varepsilon} - \bar{l} \left(\frac{B_2 k \delta}{2\varepsilon} \right) \right] \end{aligned}$$

for some $B_1 > 0$, $B_2 > 0$. Given $M > 0$, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$,

$$\bar{l} \left(\frac{B_2 k \delta}{2\varepsilon} \right) \geq \frac{(M+1)(k+1)\delta|\alpha|}{2\varepsilon}.$$

Hence for all sufficiently small $\varepsilon > 0$, the last term is bounded above by

$$B_1 \sum_{k=1}^\infty \exp[-M(k+1)\delta|\alpha|/2\varepsilon] \leq B_3 \exp(-B_4 M/\varepsilon)$$

for some $B_3 > 0$, $B_4 > 0$. Thus for all sufficiently small $\varepsilon > 0$ and for all $|x - y| < \delta/2$,

$$E_x \left\{ \exp \left[\frac{1}{\varepsilon} (Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha \right] \right\} \leq e^{h^\delta(y, \alpha)} + B_3 e^{-B_4 M / \varepsilon}.$$

The lemma follows from this. \square

We now proceed to the proof of part (ii) of Theorem 1.1. Define

$$l^\delta(x, \beta) = \sup_{\alpha \in R^n} [\alpha \cdot \beta - h^\delta(x, \alpha)].$$

Define $I_x^\delta(\phi)$ as $I_x(\phi)$ was defined in (7), but with l^δ replacing l , and define the level sets $\Phi_x^\delta(s)$ as the level sets $\Phi_x(s)$ were defined in the statement of Theorem 1.1, but with $I_x^\delta(\phi)$ replacing $I_x(\phi)$.

REMARK 2.8. In the proofs to follow we will need to use the fact that the level sets $\bigcup_{x \in C} \Phi_x^\delta(s)$ are compact for compact sets $C \subset R^n$, $s < \infty$ and $\delta > 0$. The results of [10] cited after our Proposition 2.1 yield this assertion.

PROPOSITION 2.9. *Let $\delta > 0$ be given. Then the conclusion of Theorem 1.1 holds with the rate function $I_x^{2\delta}(\phi)$ replacing $I_x(\phi)$, and Y^ε replacing X^ε .*

PROOF. Fix a compact set $C \subset R^n$ and $s < \infty$. Let $C' \subset R^n$ be a compact set containing

$$\bigcup_{x \in C} \bigcup_{0 \leq t \leq 1} \{\phi(t) : I_x(\phi) \leq s\} \cap \bigcup_{0 \leq t \leq 1} \{\phi(t) : \phi \in K(M)\}.$$

The set $K(M)$ is defined in the proof of Lemma 2.5. Define

$$h_M^\delta(x, \alpha) = \begin{cases} h^\delta(x, \alpha), & x \in C', \\ \bar{h}(|\alpha|), & x \notin C'. \end{cases}$$

Let $\tau > 0$ be given. By Lemma 2.7, for all sufficiently small $\varepsilon > 0$ and all $x \in R^n$,

$$(14) \quad E_x \exp \left(\frac{1}{\varepsilon} (Y^\varepsilon(\varepsilon) - Y^\varepsilon(0)) \cdot \alpha \right) \leq \exp(h_M^\delta(x, \alpha) + \tau)$$

for all x . Let a continuous, bounded function $\theta: [0, 1] \times R^n \rightarrow R^n$ be given. Together with the stationarity of X^ε and the Markov property, (14) implies the bound

$$(15) \quad E_x \exp \left(\frac{1}{\varepsilon} G_M^{\delta, \varepsilon}(Y^\varepsilon, \theta) \right) \leq \exp(\tau/\varepsilon),$$

where we set

$$G_M^{\delta, \varepsilon}(\phi, \theta) = \sum_{i=0}^{(1/\varepsilon)-1} [(\phi(i\varepsilon + \varepsilon) - \phi(i\varepsilon)) \cdot \theta(i\varepsilon, \phi(i\varepsilon)) - h_M^\delta(\phi(i\varepsilon), \theta(i\varepsilon, \phi(i\varepsilon)))\varepsilon]$$

when ϕ is absolutely continuous and $\phi(0) \in F$, and $G_M^{\delta, \varepsilon}(\phi, \theta) = +\infty$ in all other cases. Then (15) and the fact that τ can be made arbitrarily small imply

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log E_x \exp\left(\frac{1}{\varepsilon} G_M^{\delta, \varepsilon}(Y^\varepsilon, \theta)\right) \leq 0,$$

uniformly in $x \in C$.

Let a Borel measurable set $A \subset K(M)$ be given. Then for each function θ , and uniformly in $x \in C$,

$$(16) \quad \begin{aligned} &\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{Y^\varepsilon \in A\} \\ &\leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log E_x \exp\left(\frac{1}{\varepsilon} \left[G_M^{\delta, \varepsilon}(Y^\varepsilon, \theta) - \inf_{\phi \in A} G_M^{\delta, \varepsilon}(\phi, \theta) \right]\right) \\ &\leq - \liminf_{\varepsilon \rightarrow 0} \left(\inf_{\phi \in A} G_M^{\delta, \varepsilon}(\phi, \theta) \right). \end{aligned}$$

Since $\phi \in A$ implies $\cup_{0 \leq t \leq 1} \{\phi(t)\} \subset C'$, it follows that

$$(17) \quad \liminf_{\varepsilon \rightarrow 0} \inf_{\phi \in A} G_M^{\delta, \varepsilon}(\phi, \theta) \geq \inf_{\phi \in A} G^{2\delta}(\phi, \theta),$$

where $G^{2\delta}(\phi, \theta)$ equals

$$\int_0^1 [\dot{\phi}(t) \cdot \theta(t, \phi(t)) - h^{2\delta}(\phi(t), \theta(t, \phi(t)))] dt$$

when ϕ is absolutely continuous and $\phi(0) \in C$, and $G^{2\delta}(\phi, \theta)$ equals $+\infty$ in all other cases. Note that $G^{2\delta}(\phi, \theta)$ is lower-semicontinuous in ϕ for each θ . Combining (16) and (17) gives

$$(18) \quad \limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{Y^\varepsilon \in A\} \leq - \inf_{\phi \in A} G^{2\delta}(\phi, \theta)$$

for each bounded continuous θ and $A \subset K(M)$. The convergence in (18) is in fact uniform in $x \in C$. Let $c > 0$ be given. As in [5] (proof of Theorem 4.1), for any given ϕ , there is a bounded continuous function θ_ϕ such that

$$G^{2\delta}(\phi, \theta_\phi) \geq (I_{\phi(0)}^{2\delta}(\phi) \wedge M) - c.$$

By the lower-semicontinuity property of $G^{2\delta}$, there exists an open neighborhood $N(\phi)$ of ϕ such that $\psi \in N(\phi)$ implies

$$(19) \quad G^{2\delta}(\psi, \theta_\phi) \geq (I_{\phi(0)}^{2\delta}(\phi) \wedge M) - 2c.$$

Choose a finite subcover (with radii less than c) from among the $N(\phi)$ that cover the compact set $K(M)$ and index the subcover as $N(\phi_i)$, $1 \leq i \leq N$. For

fixed $x \in C$, let $J = \{i: d(\phi_i, \Phi_x^{2\delta}(s)) \geq c\}$. Then according to (12), (18) and (19),

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{d(Y^\varepsilon, \Phi_x^{2\delta}(s)) \geq 2c\} \\ & \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \left[P_x\{Y^\varepsilon \notin K(M)\} + \sum_{i \in J} P_x\{Y^\varepsilon \in N(\phi_i)\} \right] \\ & \leq - \left[(a_M) \wedge \bigwedge_{i \in J} \left((I_{\phi_i(0)}^{2\delta}(\phi_i) \wedge M) - 2c \right) \right]. \end{aligned}$$

Letting $M \rightarrow \infty$ and using the fact that each $I_{\phi_i(0)}^{2\delta}(\phi_i) \geq s$, we get

$$(20) \quad \limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{d(Y^\varepsilon, \Phi_x^{2\delta}(s)) \geq 2c\} \leq -s + 2c.$$

Since the cardinality of J has a bound that is independent of x , (20) is uniform in $x \in C$. According to [8] (Theorem 3.3.3), the compactness of the level sets $\Phi_x^{2\delta}(s)$ and equation (20) imply the assertion of Proposition 2.9. \square

Part (ii) of Theorem 1.1 now follows from (20) and Lemma 2.4 if we prove the following.

PROPOSITION 2.10. *Given $s > 0$ and $c > 0$, there exists $\delta > 0$ such that*

$$\Phi_x^\delta(s - c) \subset \{\phi: d(\phi, \Phi_x(s)) \leq c\}$$

for all $x \in C$.

PROOF. We first prove that for any pair (x, β) , $l^\delta(x, \beta) \uparrow l(x, \beta)$ as $\delta \downarrow 0$. Clearly $l^\delta(x, \beta)$ is monotone increasing as $\delta \downarrow 0$, and bounded above by $l(x, \beta)$. To complete the proof, we derive a general lower semicontinuity property for l^δ : for any triple of sequences $\delta_i \rightarrow 0$, $x_i \rightarrow x$, and $\beta_i \rightarrow \beta$,

$$(21) \quad \liminf_{i \rightarrow \infty} l^{\delta_i}(x_i, \beta_i) \geq l(x, \beta).$$

Choose $\alpha(n)$ such that

$$\alpha(n) \cdot \beta - h(x, \alpha(n)) \geq l(x, \beta) - 1/n.$$

Since h is upper-semicontinuous, there exists $\delta > 0$ such that $|\beta - \beta'| \leq \delta$ and $|x - x'| \leq 2\delta$ implies

$$\alpha(n) \cdot \beta' - h(x', \alpha(n)) \geq l(x, \beta) - 2/n.$$

From the definition of h^δ , $|\beta - \beta'| \leq \delta$ and $|x - x'| \leq \delta$ imply

$$\alpha(n) \cdot \beta' - h^\delta(x', \alpha(n)) \geq l(x, \beta) - 2/n.$$

The property (21) follows from this. We note that the same proof shows that $l(\cdot, \cdot)$ is jointly lower-semicontinuous.

Next note that if $\delta \geq \delta' \geq 0$, then $I_x^\delta(\phi) \leq I_x^{\delta'}(\phi) \leq I_x(\phi)$. This implies $\Phi_x(s) \subset \Phi_x^{\delta'}(s) \subset \Phi_x^\delta(s)$ whenever $s < \infty$ and $\delta \geq \delta' \geq 0$. If the proposition is not true, then by using the compactness properties of the level sets $\Phi_x^\delta(s)$, we

may assume the existence of $x_i \rightarrow x \in C$, $\delta_i \rightarrow 0$, and $\phi_i \rightarrow \phi$ such that $I_x(\phi) > s$, $I_{x_i}^{\delta_i}(\phi_i) \leq s - c$. For any $j \geq i$, $I_{x_j}^{\delta_j}(\phi_j) \leq I_{x_i}^{\delta_i}(\phi_i) \leq s - c$. Lower semicontinuity yields $I_x^{\delta_i}(\phi) \leq s - c$. The monotone convergence theorem then gives $I_x(\phi) \leq s - c$, which is a contradiction. \square

Combining Proposition 2.10 with (20) and Lemma 2.4, we obtain

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{d(X^\varepsilon, \Phi_x(s)) \geq 3c\} \leq -s + 3c.$$

Since the level sets $\Phi_x(s)$ are compact [part (i) of Theorem 1.1], this estimate implies part (ii) of Theorem 1.1 ([8], Theorem 3.3.3).

This completes the proof of Theorem 1.1.

3. Results for queueing networks. A large class of queueing networks may be modeled by jump Markov processes taking values in the nonnegative orthant

$$R_+^n = \left\{ x \in R^n : x = (x_1, x_2, \dots, x_n), \min_{i=1, \dots, n} x_i \geq 0 \right\}.$$

The processes have constant jump rates and directions in the interior of R_+^n and in each of the boundary faces of codimension $k = 1, 2, \dots, n$. In this section we specialize our main result, Theorem 1.1, to such processes. In Theorem 3.1, we derive a useful representation for the Lagrangian $l(x, \beta)$ appearing in (7). The assumptions on the processes are made for notational convenience only. They may be relaxed to cover various state dependency effects (routing, arrival rates, service rates, etc.).

For $x \in \partial R_+^n$, we define $B(x) = \{i \in \{1, 2, \dots, n\} : x_i = 0\}$. For x in the interior of R_+^n , we define $B(x) = \emptyset$. Let $\bar{B}(x)$ denote the collection of all subsets of $B(x)$. Thus, if $S = \{1, 2, \dots, n\}$, and $\bar{S} = \{\text{subsets of } S\}$, then $B(x) \subset S$ and $\bar{B}(x) \subset \bar{S}$ for all $x \in R_+^n$. The inverse mapping $B^{-1}: S \rightarrow R_+^n$ partitions R_+^n into the interior, faces of codimension 1, faces of codimension 2, ..., the origin.

Let V denote the set $\{-1, 0, 1\}^n$. We consider a jump Markov process $X(t)$ that takes values in R^n and whose jump measure is supported on the set V . For $v \in V$, $\lambda_v(x)$ denotes the jump rate in effect at $x \in R^n$ in the direction v . The corresponding queueing process is defined to be the restriction of the jump Markov process to the nonnegative integer lattice Z_+^n . In order that the queueing process never leave Z_+^n , we assume that whenever $x \in Z_+^n$ and $x + v \notin Z_+^n$, then $\lambda_v(x) = 0$. We also assume that for all pairs of points x and y in R_+^n such that $B(x) = B(y)$, we have $\lambda_v(x) = \lambda_v(y)$ for all $v \in V$.

We define the scaled queueing process

$$X^\varepsilon(t) = \varepsilon X(t/\varepsilon),$$

where $\varepsilon > 0$ is a small parameter. For each $x \in R_+^n$, this process has the

infinitesimal generator

$$\mathcal{L}^\varepsilon f(x) = \frac{1}{\varepsilon} \sum_{v \in V} \lambda_v(x) [f(x + \varepsilon v) - f(x)].$$

Theorem 1.1 applies to this process. As we now indicate, the h and l functionals may be represented in alternative forms.

Let $s \in \bar{S}$ be a subset of S and let $x \in R_+^n$ be any point in the set $\{x: B(x) = s\}$. Recall that $s = \emptyset$ corresponds to the interior of R_+^n . For $\alpha \in R^n$, we define

$$H(s, \alpha) = \sum_{v \in V} \lambda_v(x) [\exp(\alpha \cdot v) - 1].$$

For $\beta \in R^n$, we define

$$L(s, \beta) = \sup_{\alpha \in R^n} [\alpha \cdot \beta - H(s, \alpha)].$$

Clearly, for $x \in R_+^n$, the Hamiltonian $h(x, \alpha)$ in (5) may be represented as

$$h(x, \alpha) = \bigvee_{s \in \bar{B}(x)} H(s, \alpha).$$

Our next theorem indicates an interesting representation for the function $l(x, \beta)$ defined in (6). See Theorem 16.5 of the book by Rockafellar [12] for another proof.

THEOREM 3.1. For each $x \in R_+^n$ and $\beta \in R^n$,

$$l(x, \beta) = \inf \left\{ \sum_{s \in \bar{B}(x)} \rho_s L(s, \beta_s) : \sum_{s \in \bar{B}(x)} \rho_s \beta_s = \beta, \bigwedge_{s \in \bar{B}(x)} \rho_s \geq 0, \sum_{s \in \bar{B}(x)} \rho_s = 1 \right\}.$$

PROOF. For each $x \in R_+^n$ and $\alpha \in R^n$, we may write

$$\begin{aligned} h(x, \alpha) &= \bigvee_{s \in \bar{B}(x)} H(s, \alpha) \\ &= \sup \left\{ \sum_{s \in \bar{B}(x)} \rho_s H(s, \alpha) : \bigwedge_{s \in \bar{B}(x)} \rho_s \geq 0, \sum_{s \in \bar{B}(x)} \rho_s = 1 \right\}. \end{aligned}$$

Thus

$$\begin{aligned} l(x, \beta) &= \sup_{\alpha \in R^n} \{ \alpha \cdot \beta - h(x, \alpha) \} \\ &= \sup_{\alpha \in R^n} \inf \left\{ \alpha \cdot \beta - \sum_{s \in \bar{B}(x)} \rho_s H(s, \alpha) : \bigwedge_{s \in \bar{B}(x)} \rho_s \geq 0, \sum_{s \in \bar{B}(x)} \rho_s = 1 \right\}. \end{aligned}$$

According to Rockafellar [12] (Corollary 3.7.3.2, page 393), the supremum and

the infimum of the last display may be interchanged to give

$$l(x, \beta) = \inf \left\{ \left\{ \sup_{\alpha \in R^n} \left(\alpha \cdot \beta - \sum_{s \in \bar{B}(x)} \rho_s H(s, \alpha) \right) \right\} : \bigwedge_{s \in \bar{B}(x)} \rho_s \geq 0, \sum_{s \in \bar{B}(x)} \rho_s = 1 \right\}.$$

In order to prove the theorem, it suffices to show that for any fixed set $\{\rho_s, s \in \bar{B}(x)\}$ satisfying $\bigwedge_{s \in \bar{B}(x)} \rho_s \geq 0, \sum_{s \in \bar{B}(x)} \rho_s = 1,$

$$f(x, \beta) \doteq \sup_{\alpha \in R^n} \left\{ \alpha \cdot \beta - \sum_{s \in \bar{B}(x)} \rho_s H(s, \alpha) \right\}$$

equals

$$g(x, \beta) = \inf \left\{ \sum_{s \in \bar{B}(x)} \rho_s L(s, \beta_s) : \text{each } \beta_s \in R^n, \sum_{s \in \bar{B}(x)} \rho_s \beta_s = \beta \right\}.$$

A short calculation shows that the function $\beta \rightarrow g(x, \beta)$ is lower semicontinuous and proper convex. Hence it suffices to show that the Legendre–Fenchel transforms of $g(x, \cdot)$ and $f(x, \cdot)$ are equal: For all $\alpha \in R^n,$

$$g^*(x, \alpha) = \sup_{\beta \in R^n} \{ \alpha \cdot \beta - g(x, \beta) \}$$

equals $\sum_{s \in \bar{B}(x)} \rho_s H(s, \alpha).$ This is elementary since

$$\begin{aligned} g^*(x, \alpha) &= \sup_{\beta_s \in R^n, s \in \bar{B}(x)} \left\{ \sum_{s \in \bar{B}(x)} \rho_s (\alpha \cdot \beta_s - L(s, \beta_s)) \right\} \\ &= \sum_{s \in \bar{B}(x)} \rho_s \sup_{\beta_s \in R^n} \{ \alpha \cdot \beta_s - L(s, \beta_s) \} \\ &= \sum_{s \in \bar{B}(x)} \rho_s H(s, \alpha). \end{aligned}$$

This completes the proof. \square

4. Discrete time processes. In this section we state the analogue of Theorem 1.1 that is appropriate for discrete time processes. The proof follows the same lines as in the continuous time case, and is in fact simpler since step 1 in the outline given in the introduction is no longer needed.

We define our process as follows. We assume that a family $\{\mu_x, x \in R^n\}$ of probability measures is given such that the mapping $x \rightarrow \mu_x$ is measurable in $x.$ We also assume the existence of a probability space supporting a sequence of independent and identically distributed random vector fields $\{b_n(x)\}$ such that for all Borel measurable sets $B \subset R^n,$

$$P\{b_n(x) \in B\} = \mu_x(B).$$

For $n \geq 0$, we define X_n recursively by

$$X_{n+1} = X_n + \varepsilon b_n(X_n), \quad X_0 = x.$$

Under some additional assumptions, the existence of such random vector fields is proved in [1]. However, as in the case of continuous time, we shall simply assume the existence of the processes we work with.

Define the continuous parameter version X^ε by

$$X^\varepsilon(t) = \frac{X_{i+1}(t - i\varepsilon) + X_i(i\varepsilon + \varepsilon - t)}{\varepsilon}$$

for $t \in [i\varepsilon, i\varepsilon + \varepsilon]$, $i = 1, 2, \dots, [1/\varepsilon]$. Define

$$H(x, \alpha) = \log \int_{R^n} \exp(\alpha \cdot y) \mu_x(dy),$$

and define $h(x, \alpha)$, $l(x, \beta)$ and $I_x(\phi)$ by (5), (6) and (7), respectively.

Assumptions on the μ_x . We assume there exists a function $\bar{h}(a)$, defined for $a \in [0, \infty)$, which is finite for each a and satisfies

$$H(x, \alpha) \leq \bar{h}(|\alpha|)$$

for all $x \in R^n$, $\alpha \in R^n$.

THEOREM 4.1. *Under the previous assumptions on the measures μ_x and under the assumption that the processes X^ε described in this section exist, the following conclusions hold:*

(i) *Define*

$$\Phi_x(L) = \{\phi \in D([0, T]; R^n) : I_x(\phi) \leq L\}.$$

Then for all $L < \infty$ and all compact sets C , the set $\bigcup_{x \in C} \Phi_x(L)$ is compact.

(ii) *For each closed set F in $D([0, T]; R^n)$ and for each compact set C in R^n ,*

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{X^\varepsilon \in F\} \leq - \inf_{\phi \in F} I_x(\phi),$$

uniformly in $x \in C$.

5. Concluding remarks. Theorems 1.1 and 4.1 provide large deviation upper bounds for very general classes of Markov processes that live in finite dimensional spaces. The theorems in fact include and in most cases extend all such upper bound results known to the authors. However, we have not addressed the following question: Is the rate function obtained by Theorems 1.1 and 4.1 the best possible? Intuitively, the best upper bounding functional is the pointwise largest functional that satisfies the conclusion of Theorem 1.1.

Suppose that $I_x(\cdot)$ is a rate function for which the conclusions of Theorem 1.1 hold as well as the lower bound: for each open set O in $D([0, T]; R^n)$ and for each compact set C in R^n ,

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log P_x\{X^\varepsilon \in O\} \geq - \inf_{\phi \in O} I_x(\phi),$$

uniformly in $x \in C$. One can show that if both the upper and lower bounds hold with the same rate functional $I_x(\cdot)$, then subject to the regularity properties of $I_x(\cdot)$ implicit in the statement that part (i) of Theorem 1.1 holds, $I_x(\cdot)$ is unique. It follows that if one can prove the lower bound result with the same rate functional as that used in Theorems 1.1 and 4.1, then $I_x(\cdot)$ does indeed give the best upper bound. Since $I_x(\cdot)$ as defined by (7) is the rate function found in every case in the literature known to the authors for which upper and lower bounds hold (e.g., [4, 8, 14, 15, 16, 17]), it gives the best upper bound in these cases. For cases in which there are at present no well-established lower bound results, we offer the following observations. To simplify the statements, we make all references with regard to the simple one-dimensional model

$$dX^\varepsilon = b(X^\varepsilon) dt + \varepsilon^{1/2} dw_t, \quad X^\varepsilon(0) = x,$$

where $b(\cdot)$ is assumed to be bounded and measurable. This model is covered by Theorem 1.1.

Suppose that

$$b(x) = \begin{cases} a, & x \leq 0, \\ b, & x > 0. \end{cases}$$

CASE 1. Suppose that $a < 0$ and $b > 0$. With regard to the question of whether or not $I_x(\cdot)$ defined by (7) gives a lower bound result, the only real issue is the value of $l(0, 0)$. For l defined by (6), we obtain $l(0, 0) = 0$ (see Theorem 3.1). However, this is obviously incorrect as far as the lower bound is concerned. Consider the case $a = -1$, $b = 1$. By exploiting symmetry, we can consider a reflected diffusion in place of the original process and prove that the value of $l(0, 0)$ must at least be $\frac{1}{2}$. This case may be characterized by the tendency of the process to push away from the point $x = 0$ with positive speed in both directions. In such a case a more refined technique than that used in this paper must be developed in order to obtain the best upper bound.

CASE 2. Suppose $a \geq 0$ or $b \leq 0$. Here the nature of the process is such that it either tends to move across the point of discontinuity ($x = 0$), or it tends to be held at $x = 0$ (when $a > 0$ and $b < 0$). In both these cases, (7) gives the correct rate function for the lower bound. However, these results have not yet been verified in a general setting.

On the basis of these simple examples, we conjecture that our upper bound rate functional is the best available in at least some multidimensional settings corresponding to Case 2. By this we mean a process that has one or several smooth $(n - 1)$ -dimensional manifolds across which the statistics jump and that has the tendency either to move across these manifolds or to be pushed from both sides into them. We note that the Markov processes that model queues are of this type and that the results of [4] support this conjecture.

Acknowledgments. Richard S. Ellis wishes to acknowledge the hospitality of the Division of Applied Mathematics at Brown University, which he

visited during the fall semester of 1988, and the hospitality of the Faculty of Industrial Engineering and Management at the Technion, which he visited during the spring semester of 1989.

REFERENCES

- [1] AZENCOTT, R. and RUGET, G. (1977). Mélanges d'équations différentielles et grand écart à la loi des grandes nombres. *Z. Wahrsch. Verw. Gebiete* **38** 1–54.
- [2] DE ACOSTA, A. (1985). Upper bounds for large deviations of dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **69** 551–565.
- [3] DONSKER, M. D. and VARADHAN, S. R. S. (1976). Asymptotic evaluation of certain Markov process expectations for large time, III. *Comm. Pure Appl. Math.* **29** 389–461.
- [4] DUPUIS, P., ISHII, H. and SONER, H. M. (1990). A viscosity solution approach to the asymptotic analysis of queueing systems. *Ann. Probab.* **18** 226–255.
- [5] DUPUIS, P. and KUSHNER, H. J. (1989). Stochastic approximation and large deviations: Upper bounds and w.p.1 convergence. *SIAM J. Control Optim.* **27** 1108–1135.
- [6] ELLIS, R. S. (1984). Large deviations for a general class of random vectors. *Ann. Probab.* **12** 1–12.
- [7] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.
- [8] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems*. Springer, New York.
- [9] GÄRTNER, J. (1977). On large deviations from the invariant measure. *Theory Probab. Appl.* **22** 24–39.
- [10] IOFFE, A. D. and TIKHOMIROV, V. M. (1979). *Theory of Extremal Problems. Studies in Mathematics and Its Applications* **6**. North-Holland, Amsterdam.
- [11] JACOD, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. Springer, New York.
- [12] ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.
- [13] STROOCK, D. S. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, New York.
- [14] WENTZELL, A. D. (1976). Rough limit theorems on large deviations for Markov stochastic processes, I. *Theory Probab. Appl.* **21** 227–242.
- [15] WENTZELL, A. D. (1976). Rough limit theorems on large deviations for Markov stochastic processes, II. *Theory Probab. Appl.* **21** 499–512.
- [16] WENTZELL, A. D. (1979). Rough limit theorems on large deviations for Markov stochastic processes, III. *Theory Probab. Appl.* **24** 675–692.
- [17] WENTZELL, A. D. (1982). Rough limit theorems on large deviations for Markov stochastic processes, IV. *Theory Probab. Appl.* **27** 215–234.

DEPARTMENT OF MATHEMATICS AND STATISTICS
LERDERLE GRADUATE RESEARCH CENTER TOWERS
UNIVERSITY OF MASSACHUSETTS
AMHERST, MASSACHUSETTS 01003

AT & T BELL LABORATORIES
600 MOUNTAIN AVENUE
MURRAY HILL, NEW JERSEY 07974