

## LARGE DEVIATION LOWER BOUNDS FOR ARBITRARY ADDITIVE FUNCTIONALS OF A MARKOV CHAIN

BY A. DE ACOSTA<sup>1</sup> AND PETER NEY<sup>2</sup>

Case Western Reserve University and University of Wisconsin,  
Madison

A universal large deviation lower bound is proved for sums of Banach space valued functions of an irreducible, general state space Markov chain. There are no restrictions on the functions (other than measurability).

### 1. Introduction.

1a. *Summary.* The purpose of this paper is to prove the following.

**THEOREM.** Let  $\{X_i, i = 1, 2, \dots\}$  be an irreducible Markov chain (MC) taking values in a countably generated general measurable space  $(S, \mathcal{S})$ ,  $E$  be a separable Banach space, endowed with its Borel  $\sigma$ -algebra  $\mathcal{E}$ ,  $f: S \rightarrow E$  be a measurable function and  $\mu$  be an initial measure for the MC. Then for any open set  $G \subset E$ ,

$$(1.1) \quad \liminf \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{1}{n} \sum_{i=1}^n f(X_i) \in G \right\} \geq - \inf_{u \in G} I(u),$$

where  $I: E \rightarrow [0, \infty]$  is convex and lower semicontinuous (lcs), and is explicitly identified in (1.13) below.

Following standard terminology,  $\{\sum_1^n f(X_i)\}$  is said to satisfy the large deviation principle (LDP) with rate function  $I$  (under  $\mathbb{P}_\mu$ ) if (1.1) is matched by the upper bound

$$(1.2) \quad \limsup \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{1}{n} \sum_{i=1}^n f(X_i) \in F \right\} \leq - \inf_{x \in F} I(x)$$

for closed  $F \subset E$ . If (1.1) holds and (1.2) only holds for compact  $F \subset E$ , one says that the weak LDP holds. We will abbreviate  $\inf_{x \in \Gamma} I(x) = I(\Gamma)$  for  $\Gamma \in \mathcal{E}$ .

---

Received January 1997; revised April 1998.

<sup>1</sup>Supported in part by an NSF grant.

<sup>2</sup>Supported in part by an NSF grant and by a grant from the Humboldt Foundation at the University of Frankfurt.

AMS 1991 subject classifications. 60F10, 60J55.

Key words and phrases. Large deviations, Markov chains, convergence parameter, Banach space.

There is an extensive literature on LDP's for MC's under various special models and hypotheses. The earliest work appears to be due to Miller [19], who considered finite state spaces  $\mathcal{S}$ , and identified  $I$  as the convex conjugate of the Perron–Frobenius (PF) root of a related “transform matrix” (defined below).

An important class of additive functionals that has received much attention are the empirical measures of the MC, which are studied in the seminal papers of Donsker and Varadhan [14]. They gave an explicit representation for the rate function by a variational formula. Subsequent contributions and improvements on this subject were made by de Acosta [1, 3, 4], Bolthausen [8], Jain [16], Dinwoodie and Ney [13], Stroock [23].

Results are less complete for general additive functions  $f$ . The case of bounded  $f$  has been treated in [3], while the case when  $E = \mathbb{R}^d$  is in [20] subject to a further irreducibility condition. Our goal in this paper is to establish a universal lower bound in terms of a rate function which is a natural structural object associated with the transition function of the chain and  $f$ . The bound is tight for large classes of cases in the sense that there is a corresponding upper bound with the same rate function. The issue of tightness, with some examples and counterexamples, will be discussed in Section 4. Our lower bound is “universal” in the sense that there are *no* restrictions on  $f$  (other than measurability), and only irreducibility (which is clearly necessary) is assumed on the MC  $\{X_n\}$ . In particular, there are no recurrence, ergodicity, or uniformity requirements on  $\{X_n\}$ .

In addition to proving the above theorem, we start in Section 2 with a new and streamlined proof of (1.1) when  $f$  is bounded. The idea is to take advantage of the regeneration property of the MC to yield a subadditivity and other properties of the chain, without requiring anything other than irreducibility of the chain. Identification of the rate function draws on [12]. The argument is roughly along the lines developed in [13] for the empirical measure case.

The proof of the general result via a smoothing and truncation argument is in Section 3. A key point in the proof is to develop a scheme in which truncations of the state space of an extension of the MC  $\{X_n\}$ , do not destroy irreducibility. Another important issue that has to be dealt with is the convergence of the rate functions for the truncated chains.

**1b. Background remarks on Markov chains and nonnegative kernels.** We summarize notation and some known facts about Markov chains. See [21] for more details on definitions and proofs of the following assertions.

(1) Let

$$P = \{P(x, A), x \in \mathcal{S}, A \in \mathcal{S}\}$$

denote the transition kernel of  $\{X_n\}$  (stochastic or substochastic). Irreducibility means that there is a  $\sigma$ -finite measure  $\varphi$  on  $(\mathcal{S}, \mathcal{S})$  such that

$$(1.3) \quad \varphi(A) > 0 \text{ implies } \sum_{n=1}^{\infty} P^n(x, A) > 0 \text{ for all } x \in \mathcal{S}.$$

There always exists a maximal irreducibility measure  $\psi$ , namely an irreducibility measure with respect to which all irreducibility measures are absolutely continuous.

Irreducibility of  $P$  implies the existence of a probability measure  $\nu_0$  on  $(S, \mathcal{S})$ , a set  $C \in \mathcal{S}$ , with  $\psi(C) > 0$ , an integer  $m_0 > 0$  and a  $0 < \delta \leq 1$  such that

$$(1.4) \quad \delta \mathbf{1}_C(x) \nu_0(A) \leq P^{m_0}(x, A), \quad x \in S, A \in \mathcal{S}$$

([21], Theorem 2.1). We call  $(\nu_0, C, m_0, \delta)$  a minorization for  $P$ . If (1.4) is satisfied with  $m_0 = 1$ , we say that  $P$  satisfies a *strong* minorization.

(2) Irreducibility of a nonnegative kernel  $\{K(x, A); x \in S, A \in \mathcal{S}\}$  is defined as in (1.3). An irreducible kernel  $K$ , may satisfy a minorization

$$(1.5) \quad \delta \mathbf{1}_C(x) \nu_0(A) \leq K^{m_0}(x, A).$$

In this case the *convergence parameter* of  $K$ , say  $R(K)$ , is defined to be the radius of convergence of the power series

$$(1.6) \quad \sum_{n=1}^{\infty} r^n(\nu_0 K^n(C)).$$

This definition is independent of the choice of  $\nu_0$  and  $C$  (see [21], Proposition 3.4).

In the present work we will be concerned with the transform kernel

$$(1.7) \quad K_\xi(x, A) = \int_A P(x, dy) \exp(\langle \xi, f(y) \rangle), \quad \xi \in E^*, A \in \mathcal{S},$$

where  $E^*$  is the dual space of  $E$ .  $K_\xi$  is clearly irreducible since

$$K_\xi^n(x, A) > 0 \iff P^n(x, A) > 0.$$

(3) The definition (1.6) of the convergence parameter depends on the existence of the minorization (1.5). This minorization is proved in [21], Theorem 2.1 for any non-negative irreducible kernel under the assumption that all powers  $\{K^n\}$  are  $\sigma$ -finite (see [21], page 2). The definition and relevant properties of the convergence parameter depend only on irreducibility and (1.5), not directly on the  $\sigma$ -finiteness of  $\{K^n\}$ . If  $K$  does not itself have  $\sigma$ -finite powers but is minorized by an irreducible kernel  $\bar{K}$  with  $\{\bar{K}^n\}$   $\sigma$ -finite, then (1.5) will also hold for  $K$ , and the properties of the convergence parameter developed in Chapters 2–4 of [21] apply to  $K$ .

In the present context  $\{K_\xi^n, n = 1, 2, \dots\}$  defined in (1.7) may not be  $\sigma$ -finite, but  $K_\xi$  is minorized by the bounded irreducible kernel

$$(1.8) \quad \bar{K}_\xi(x, A) = \int_A \exp(\langle \xi, f(y) \rangle \wedge 1) P(x, dy).$$

The previous remarks apply to  $K_\xi$ , which thus satisfies a minorization

$$(1.9) \quad \delta \mathbf{1}_C(x) \nu_0(A) \leq K_\xi^{m_0}(x, A),$$

with  $\psi(C) > 0, \nu_0(S) > 0$ .

(4) An alternative (equivalent) form of the minorization (1.5) [and similarly (1.9)] is the existence of a *small* function  $s: S \rightarrow \mathbb{R}$  such that

$$(1.10) \quad s(x) \nu_0(A) \leq K^{m_0}(x, A),$$

where  $\int s d\psi > 0$  [21]. Then the convergence parameter  $R(K)$  is the radius of convergence of

$$(1.11) \quad \sum_n r^n (\nu_0 K^n s).$$

Finally, the rate function  $I$  in (1.1) will be identified as follows. Let

$$(1.12) \quad \Lambda(\xi) = \log R^{-1}(K_\xi).$$

Then

$$(1.13) \quad I(u) = \Lambda^*(u),$$

where

$$\Lambda^*(u) = \sup_{\xi \in E^*} [\langle \xi, u \rangle - \Lambda(\xi)].$$

( $\Lambda^*$  is the convex conjugate of  $\Lambda$ .)

**2. Bounded  $f$ .** In this section we prove the lower bound (1.1) for irreducible  $\{X_n\}$  when  $f: S \rightarrow E$  is bounded. This result is known [3], but we give a new and streamlined proof along lines developed in a different context in [13]. This is based on a subadditivity argument stemming from the regeneration structure of the chain, and an identification of the rate function via Varadhan’s theorem [24], Theorem 2.2 and Dinwoodie’s [12] Theorem 3.1. A large deviation upper bound for compacts is a byproduct. Subadditivity arguments were introduced in [17] and [7] in the i.i.d case, and in [23] for Markov chains under a strong (upper and lower) bound condition on the transition kernel. We use regeneration to avoid these restrictions.

REMARK. Note that the arguments and conclusions of this section hold for substochastic transition kernels  $P = \{(P(x, A), x \in S, A \in \mathcal{S})\}$ .

Recall that  $R(K_\xi)$  is the convergence parameter of the kernel

$$(2.1) \quad \begin{aligned} K_\xi(x, A) &= \int_A \exp(\langle \xi, f(y) \rangle) P(x, dy), \quad x \in S, A \in \mathcal{S}, \xi \in E^* \quad \text{and} \\ \Lambda(\xi) &= -\log R(K_\xi), \quad \Lambda^*(v) = \sup_{\xi \in E^*} [\langle \xi, v \rangle - \Lambda(\xi)], \quad v \in E. \end{aligned}$$

Let  $S_n = f(X_1) + \dots + f(X_n)$ .

PROPOSITION 2.1. *If  $\{X_n\}$  is irreducible,  $f$  is bounded,  $\mu$  is any initial distribution (of  $X_1$ ) and  $G$  is an open set, then*

$$\liminf \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in G \right\} \geq -\Lambda^*(G),$$

where  $\Lambda^*(G) = \inf_{u \in G} \Lambda^*(u)$ .

The proof of this proposition will follow from several lemmas.

Let  $\psi$  be the maximal irreducibility measure of  $\{X_n\}$ . Then we have remarked that there exists a minorization  $(\nu_0, C, m_0, \delta)$ , namely,

$$(2.2) \quad \delta \mathbf{1}_C(x) \nu_0(A) \leq P^{m_0}(x, A), \quad x \in S, A \in \mathcal{S},$$

with  $\nu_0(S) > 0$  and  $\psi(C) > 0$ .

Hence for any  $n \geq 0$ ,

$$\delta \mathbf{1}_C(x) \nu_0 P^n(A) \leq P^{m_0+n}(x, A).$$

Now irreducibility implies that  $\nu_0 P^n(C) > 0$  for some  $n$ , and hence letting  $\nu_0 P^n = \nu$  and  $m_0 + n = m$ , we always have a minorization  $(\nu, C, m, \delta)$ , namely,

$$(2.3) \quad \delta \mathbf{1}_C(x) \nu(A) \leq P^m(x, A), \quad x \in S, A \in \mathcal{S} \text{ with } \nu(C) > 0.$$

We first prove the proposition under the strong minorization hypothesis

$$(2.4) \quad \delta \mathbf{1}_C(x) \nu(A) \leq P(x, A)$$

[i.e.,  $m = 1$  in (2.3)], and the further restriction that

$$(2.5) \quad \psi(A) > 0 \text{ implies } P(x, A) > 0 \text{ for all } x \in S.$$

[Note that (2.4) and  $\nu(C) > 0$  implies that  $\{X_n\}$  is aperiodic.] These restrictions will later be removed.

Under (2.4) there is a regeneration structure associated with  $\{X_n\}$ . Namely, a sequence of r.v.'s  $\{Y_n, n = 1, 2, \dots\} \in \{0, 1\}$  is adjoined to  $\{X_n\}$ , such that  $\{(X_n, Y_n)\}$  is a MC on  $S \times \{0, 1\}$ , the marginal sequence  $\{Y_n\}$  is i.i.d. with  $P\{Y_n = 1\} = \delta$  and

$$\mathbb{P}\{X_{n+1} \in A, Y_{n+1} = e \mid X_n \in C, Y_n = 1\} = \nu(A) P\{Y_1 = e\},$$

where  $e = 0$  or  $1$ . We abbreviate the event  $\{X_n \in C\} \cap \{Y_n = 1\}$  by  $R_n$ , and say that regeneration occurs at time  $n$ . Intuitively, if  $R_n$  occurs, then  $X_{n+1}$  is distributed by  $\nu$ , independent of the past. (See, e.g., [21], [18], [20] or [5]). Let

$$(2.6) \quad \mu_n = \mathbb{P}_\nu \left( \left( \frac{S_n}{n} \in \cdot \right) \cap R_n \right),$$

and

$$(2.7) \quad a_n = \mu_n(B_\nu(\varepsilon)),$$

where  $B_\nu(\varepsilon) = \{u \in E: \|u - \nu\| < \varepsilon\}$ .

REMARK. When we write

$$(2.8) \quad \mathbb{P}_\eta \left\{ \frac{S_n}{n} \in \cdot, X_n \in A, Y_n = e \right\},$$

the initial measure  $\eta$  should really be on  $S \times \{0, 1\}$ . However, since the conditional probabilities

$$(2.9) \quad \mathbb{P}\{X_n \in A, Y_n = e \mid X_0 = x, Y_0 = e'\}, \quad e, e' \in \{0, 1\}$$

are independent of  $e'$ , it is only the marginal part of the measure  $\eta$  on  $S$  that is relevant. Hence (with some abuse of notation), if we take  $\eta$  to be a measure on  $(S, \mathcal{S})$ , (2.8) will still be well defined.

LEMMA 2.2. Assume (2.4) and (2.5). Then:

- (i)  $-\log a_{m+n} \leq -\log a_m - \log a_n$ ;
- (ii) if  $a_{n_0} > 0$  for some  $n_0$ , then  $a_n > 0$  for all sufficiently large  $n$ ;
- (iii)  $\lim(1/n)\log a_n = I(v, \varepsilon)$  (say) exists.

PROOF. Write  $S_{m,n} = f(X_m) + \dots + f(X_n)$ ,  $S_{1,n} = S_n$ . Then

$$\begin{aligned}
 a_{m+n} &\geq \mathbb{P}_\nu(\{S_m \in mB_\nu(\varepsilon), S_{m+1,m+n} \in nB_\nu(\varepsilon)\} \cap R_m \cap R_{m+n}) \\
 (2.10) \quad &= \mathbb{P}_\nu(\{S_m \in mB_\nu(\varepsilon)\} \cap R_m) \mathbb{P}_\nu(\{S_n \in nB_\nu(\varepsilon)\} \cap R_n) \\
 &= a_m a_n \text{ proving (i).}
 \end{aligned}$$

For part (ii), suppose  $a_{n_0} > 0$  for some  $n_0$ , and define  $k$  by

$$(2.11) \quad n = kn_0 + d, \quad 0 \leq d < n_0.$$

Since  $B_\nu(\varepsilon)$  is open, there is an  $r < 1$  such that

$$(2.12) \quad \mu_{n_0}(B_\nu(r\varepsilon)) > 0.$$

Take  $r' \in (r, 1)$ . Then for  $n$  sufficiently large and  $\delta$  sufficiently small

$$(2.13) \quad (n/kn_0) B_\nu(\varepsilon) \supset B_\nu(r'\varepsilon) \supset B_\nu(r\varepsilon) + B_0(\delta).$$

Hence for such  $n$ ,  $\delta$ , and  $d$  as in (2.11),

$$\begin{aligned}
 a_n &= \mathbb{P}_\nu(\{S_n \in nB_\nu(\varepsilon)\} \cap R_n) \\
 (2.14) \quad &\geq \mathbb{P}_\nu(\{S_n \in kn_0 B_\nu(r\varepsilon) + kn_0 B_0(\delta)\} \cap R_{kn_0} \cap R_n) \\
 &\geq \mathbb{P}_\nu(\{S_{kn_0} \in kn_0 B_\nu(r\varepsilon)\} \cap R_{kn_0}) \\
 &\quad \times \mathbb{P}_\nu(\{S_d \in kn_0 B_0(\delta)\} \cap R_d).
 \end{aligned}$$

Now  $kn_0 \geq d \lceil kn_0/d \rceil$ , where  $\lceil t \rceil$  is the largest integer in  $t$ , and hence the last factor in (2.14) is greater than or equal to

$$\begin{aligned}
 (2.15) \quad &\mathbb{P}_\nu\left(\left\{S_d \in d \left\lceil \frac{kn_0}{d} \right\rceil B_0(\delta)\right\} \cap R_d\right) \\
 &\geq \left(\mathbb{P}_\nu\left(\left\{S_1 \in \left\lceil \frac{kn_0}{d} \right\rceil B_0(\delta)\right\} \cap R_1\right)\right)^d.
 \end{aligned}$$

But now  $\nu(C) > 0$  implies  $\mathbb{P}_\nu(R_1) > 0$  and  $\mathbb{P}_\nu\{S_1 \in \lceil kn_0/d \rceil B_0(\delta)\} \nearrow 1$  as  $n \rightarrow \infty$ . Hence (2.15)  $> 0$  for large  $n$ . Finally, regeneration implies that the first factor in (2.14) is greater than or equal to

$$(2.16) \quad \left(\mathbb{P}_\nu(\{S_{n_0} \in n_0 B_\nu(r\varepsilon)\} \cap R_{n_0})\right)^{k_0} = (\mu_{n_0}(B_\nu(r\varepsilon)))^{k_0} > 0 \text{ by (2.12).}$$

This proves (ii).

Finally (i) and (ii)  $\Rightarrow$  (iii). (See, e.g., Lemma 6.1.11 of [11].)  $\square$

DEFINITION. Let

$$(2.17) \quad I(v) = - \lim_{\varepsilon \searrow 0} I(v, \varepsilon).$$

We can now draw some conclusions about the measures  $\{\mu_n\}$ .

LEMMA 2.3. Assume (2.4) and (2.5). Then:

(i) The measures  $\{\mu_n\}$  satisfy the weak LDP with convex, l.s.c. rate function  $I(\cdot)$ .

(ii) If  $\Gamma \subset \mathbb{E}$  is closed and convex, then

$$(2.18) \quad \mu_n(\Gamma) \leq \exp(-nI(\Gamma)).$$

PROOF. The proof of part (i) is a standard argument, exactly as in Propositions 4.5 and 4.7 of [6]. For part (ii) note that the subadditivity in Lemma 2.2 works if balls are replaced by any measurable convex set. Then for compact, convex  $K \subset \Gamma$ ,

$$(2.19) \quad \frac{1}{n} \log \mu_n(K) \leq \frac{1}{mn} \log \mu_{mn}(K),$$

and for fixed  $n \geq 1$ ,

$$(2.20) \quad \frac{1}{n} \log \mu_n(K) \leq \limsup_{m \rightarrow \infty} \frac{1}{mn} \log \mu_{mn}(K) \leq -I(K) \leq -I(\Gamma).$$

Now for  $\varepsilon > 0$ , take compact convex  $K_\varepsilon \subset \Gamma$  such that

$$(2.21) \quad \mu_n(\Gamma) - \varepsilon \leq \mu_n(K_\varepsilon) \leq \exp(-nI(\Gamma)) \quad \text{by (2.20)}.$$

Finally, let  $\varepsilon \searrow 0$ .  $\square$

Let

$$(2.22) \quad \rho_n(\cdot) = \mathbb{P}_\nu \left\{ \left( \frac{S_n}{n} \in \cdot \right) \cap (X_n \in C) \right\}$$

and note that since  $\{Y_n\}_{n \geq 1}$  is independent of  $\{X_n\}_{n \geq 1}$ ,

$$(2.23) \quad \delta \rho_n = \mu_n, \quad n \geq 1.$$

In the next lemma we identify the rate function  $I(\cdot)$ . The boundedness of  $f$  is used here for the first time.

LEMMA 2.4. Assume (2.4), (2.5) and  $f$  bounded. Let  $\Lambda(\xi)$  be as defined in (2.1). Then

$$(2.24) \quad I(v) = \Lambda^*(v), \quad v \in \mathbb{E}.$$

PROOF. We first show that

$$(2.25) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \nu K_\xi^n(C) = I^*(\xi), \quad \xi \in \mathbb{E}^*,$$

where

$$(2.26) \quad I^*(\xi) = \sup_{\nu \in \mathbb{E}} [\langle \xi, \nu \rangle - I(\nu)].$$

To this end recall that the measure  $\nu$  and set  $C$  are as in (2.4), and

$$(2.27) \quad \begin{aligned} \nu K_\xi^n(C) &= \int_S \nu(dx) K_\xi^n(x, C) \\ &= \int \nu(dx) E_x[\exp(\langle \xi, S_n \rangle) \mathbf{1}_C(X_n)] \\ &= E_\nu[\exp(\langle \xi, S_n \rangle) \mathbf{1}_C(X_n)] \\ (2.28) \quad &= \int_E \exp(n\langle \xi, u \rangle) \rho_n(du) \\ &= \int_{B_0(b)} \exp(n\langle \xi, u \rangle) \rho_n(du), \end{aligned}$$

where  $b = \sup_{x \in S} \|f(x)\|$ , since the support of  $\rho_n \subset B_0(b)$ .

Now we have seen in Lemma 2.3 that the measures  $\mu_n = \delta_{\rho_n}$  on  $(B_0(b), E \cap B_0(b))$  satisfy the weak LDP and the upper bound inequality (2.18) for closed convex sets.

Since  $\xi$  is linear,  $\xi^{-1}(J) \cap B_0(b)$  is a closed convex set for any closed interval  $J$ . This suffices to justify the application of Varadhan's theorem (Theorem 2.2 of [24]) to conclude that

$$(2.29) \quad \begin{aligned} \lim \frac{1}{n} \log \nu K_\xi^n(C) &= \lim \frac{1}{n} \log \int_{B_0(b)} \exp(n\langle \xi, u \rangle) \rho_n(du) \\ &= \sup_{\nu \in B_0(b)} [\langle \xi, \nu \rangle - I(\nu)] \end{aligned}$$

$$(2.30) \quad = \sup_{\nu \in E} [\langle \xi, \nu \rangle - I(\nu)] = I^*(\xi)$$

since  $I(\nu) = \infty$  for  $\nu \notin B_0(b)$  [since  $\rho$  is supported on  $B_0(b)$ ], proving (2.25).

Now the power series with coefficients  $\nu K_\xi^n(C)$  has radius of convergence

$$(2.31) \quad \begin{aligned} R(K_\xi) &= \left[ \limsup (\nu K_\xi^n(C))^{1/n} \right]^{-1} = \left[ \lim (\nu K_\xi^n(C))^{1/n} \right]^{-1} \\ &= \exp(-I^*(\xi)) \quad \text{by (2.25)}. \end{aligned}$$

Hence

$$(2.32) \quad \Lambda(\xi) = -\log R(K_\xi) = I^*(\xi).$$

Since  $I(\cdot)$  is known to be convex and lsc,

$$(2.33) \quad I(\nu) = I^{**}(\nu) = \Lambda^*(\nu)$$

(see [15]).  $\square$

Summarizing Lemmas 2.3 and 2.4 and taking into account (2.23), we see that under (2.4), (2.5) and  $f$  bounded,

$$(2.34) \quad \text{The measures } \rho_n \text{ satisfy the weak LDP with rate } \Lambda^*,$$

which may be of some independent interest. The next step is to replace the initial measure  $\nu$  by an arbitrary (sub)stochastic  $\mu$ .

LEMMA 2.5. *Assume (2.4), (2.5) and  $f$  bounded. Then*

$$(2.35) \quad \liminf \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in G, X_n \in C \right\} \geq -\Lambda^*(G)$$

for all open  $G$ .

PROOF. Irreducibility of  $\{X_n\}$  implies that  $\mu P^{n_0-1}(C) > 0$  for some  $n_0 > 1$ . With (2.4) this implies that

$$(2.36) \quad \delta' \nu(A) \leq \mu P^{n_0}(A),$$

with  $\delta' = \delta \mu P^{n_0-1}(C)$ . Now take any  $x_0 \in G$ . Then

$$(2.37) \quad \begin{aligned} & \mathbb{P}_\mu(\{S_n \in nB_{x_0}(\varepsilon)\} \cap R_n) \\ & \geq \int_E \mathbb{P}_\mu(\{S_{n_0} \in ds, S_{n_0+1, n} \in nB_{x_0}(\varepsilon) - s\} \cap R_{n_0} \cap R_n) \\ & = \int_{B(n_0 \| f \|)} \mathbb{P}_\mu(\{S_{n_0+1, n} \in nB_{x_0}(\varepsilon) - s\} \cap R_n | S_{n_0} = s, R_{n_0}) \\ & \quad \times \mathbb{P}_\mu(\{S_{n_0} \in ds\} \cap R_{n_0}). \end{aligned}$$

Since  $f$  is bounded, the regeneration property implies that there is an  $r < 1$  such that for  $n$  sufficiently large, (2.37) is greater than or equal to

$$(2.38) \quad \int_E \mathbb{P}_\nu(\{S_{n-n_0} \in nB_{x_0}(r\varepsilon)\} \cap R_{n-n_0}) \mathbb{P}_\mu(\{S_{n_0} \in ds\} \cap R_{n_0}).$$

Since the first factor in (2.38) is independent of  $s$ , this equals

$$(2.39) \quad \begin{aligned} & \mathbb{P}_\mu(R_{n_0}) \mathbb{P}_\nu(\{S_{n-n_0} \in nB_{x_0}(r\varepsilon)\} \cap R_{n-n_0}) \\ & \geq \mathbb{P}_\mu(R_{n_0}) \mathbb{P}_\nu(\{S_{n-n_0} \in (n-n_0)B_{x_0}(r'\varepsilon)\} \cap R_{n-n_0}) \\ & \quad \text{for some } r' < r \\ & = \mathbb{P}_\mu(R_{n_0}) \delta \rho_{n-n_0}(B_{x_0}(r\varepsilon)). \end{aligned}$$

Now

$$(2.40) \quad \mathbb{P}_\mu(R_{n_0}) = \delta \mathbb{P}_\mu(X_{n_0} \in C) = \delta \mu P^{n_0}(C) \geq \delta \delta' \nu(C) > 0,$$

so applying Lemmas 2.3 and 2.4 to (2.39) we see that

$$(2.41) \quad \begin{aligned} & \liminf \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in B_{x_0}(\varepsilon), X_n \in C \right\} \\ & \geq -\Lambda^*(B_{x_0}(\varepsilon)) \geq -\Lambda^*(x_0). \end{aligned}$$

Since this holds for all  $x_0 \in G$ , Lemma 2.5 follows.  $\square$

PROOF OF PROPOSITION 2.1. We need to get rid of the restrictions (2.4) and (2.5). To that end, define the transition function

$$(2.42) \quad P_t(x, A) = (1 - t) \sum_{n=0}^{\infty} t^n P^{n+1}(x, A), \quad 0 < t < 1.$$

Clearly  $P_t$  is a transition kernel. By (2.3),

$$P_t(x, A) \geq (1 - t) t^{m-1} \delta \mathbf{1}_C(x) \nu(A),$$

and it satisfies (2.4) with  $\delta$  replaced by  $\delta(1 - t)t^{m-1}$ . Also, since  $P$  is irreducible,  $P_t$  satisfies (2.5). Therefore by Lemma 2.5,

$$(2.43) \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_{\mu, t} \left\{ \frac{S_n}{n} \in G \right\} \geq -\Lambda_t^*(G),$$

where  $\mathbb{P}_{\mu, t}$  is the measure determined by  $P_t$  with initial measure  $\mu$ , and the subscripts  $t$  denote the corresponding quantities determined by  $P_t$ .

We now quote the following lemma.

LEMMA 2.6 ([3], Lemma 5.2). *Let  $\{\tau_j, j = 1, 2, \dots\}$  be i.i.d. r.v.'s with*

$$(2.44) \quad \mathbb{P}\{\tau_j = k\} = (1 - t)t^k, \quad k = 1, 2, \dots, 0 < t < 1.$$

*Let  $b = \sup\{\|f(x)\|, x \in S\}$ . For open  $G \subset E$ ,  $\varepsilon > 0$  let  $G_\varepsilon = \{v \in E: d(v, G^c) > \varepsilon\}$ . Then for any substochastic measure  $\mu$  on  $(S, S)$ ,*

$$(2.45) \quad \mathbb{P}_{\mu, t} \left\{ \frac{S_n}{n} \in G_\varepsilon \right\} \leq \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in G \right\} + r_n(t),$$

where

$$r_n(t) = \mathbb{P} \left\{ \frac{1}{n} \sum_1^n \tau_j \geq 1 + \frac{\varepsilon}{2b} \right\}.$$

Now from (2.45) we see that

$$(2.46) \quad \begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_{\mu, t} \left\{ \frac{S_n}{n} \in G_\varepsilon \right\} \\ & \leq \max \left[ \liminf \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in G \right\}, \limsup \frac{1}{n} \log r_n(t) \right]. \end{aligned}$$

However, by the classical Cramér theorem,

$$\lim_{t \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \log r_n(t) = -\infty.$$

Hence by (2.43) and (2.45),

$$\begin{aligned} \lim_{t \rightarrow 0} -\Lambda_t^*(G_\varepsilon) & \leq \lim_{t \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_{\mu, t} \left\{ \frac{S_n}{n} \in G_\varepsilon \right\} \\ & \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in G \right\}. \end{aligned}$$

Now take any  $u \in G$ , and  $\varepsilon$  sufficiently small so that  $u \in G_\varepsilon$ . Then

$$(2.47) \quad \lim_{t \rightarrow 0} -\Lambda_t^*(u) \leq \liminf \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in G \right\}.$$

However,

$$P_t \geq (1 - t)P \text{ implying } R(K_{\xi, t}) \leq (1 - t)^{-1} R(K_\xi).$$

Thus

$$(2.48) \quad \Lambda_t(\xi) \geq \Lambda(\xi) + \log(1 - t)$$

and

$$(2.49) \quad \Lambda_t^*(u) \leq \Lambda^*(u) - \log(1 - t).$$

Thus by (2.47),

$$-\Lambda^*(u) \leq \liminf \frac{1}{n} \log \mathbb{P}_\mu \left\{ \frac{S_n}{n} \in G \right\}$$

for all  $u \in G$ . This concludes the proof of Proposition 2.1.  $\square$

**3. Smoothing, truncation and the proof of the theorem.** The first step is to define a new Markov kernel on  $(S \times E, \mathcal{S} \otimes \mathcal{E})$  which corresponds to a “smoothing” of the original Markov chain  $\{X_n\}_{n \geq 1}$ , by adjoining an independent, identically distributed sequence  $\{Z_n^\sigma\}_{n \geq 1}$  of  $E$ -valued Gaussian vectors with “small variance” which is independent of  $\{X_n\}_{n \geq 1}$ ;  $\{Z_n^\sigma\}$  will be formally introduced later, in the proof of the Theorem. It is important for our purposes that the common Gaussian distribution  $\mathcal{L}(Z_1^\sigma)$  have full support. For the sake of completeness, we give a simple direct proof of the following lemma.

LEMMA 3.1. *There exists an  $E$ -valued centered Gaussian vector  $Z$  such that for every  $\sigma > 0$  and every open set  $G$  in  $E$ ,*

$$\mathbb{P}\{\sigma Z \in G\} > 0.$$

PROOF. Let  $\{g_n\}_{n \geq 1}$  be an independent sequence of  $N(0, 1)$  random variables, let  $\{a_n\}_{n \geq 1}$  be a positive sequence such that for all  $\varepsilon > 0$ ,

$$\prod_{n=1}^{\infty} \mathbb{P}\{|g_1| \leq \varepsilon a_n\} > 0,$$

let  $\{b_n\}_{n \geq 1}$  be a positive sequence such that  $\sum_{n=1}^{\infty} a_n b_n \leq 1$  and let  $\{u_n\}_{n \geq 1}$  be a dense subset of the unit ball of  $E$ . Define

$$Z = \sum_{n=1}^{\infty} b_n g_n u_n;$$

the series converges in  $L^1(E)$  since  $\sum_{n=1}^{\infty} \mathbb{E}\|b_n g_n u_n\| \leq \mathbb{E}|g_1| \sum_{n=1}^{\infty} b_n < \infty$ , and  $Z$  is a centered Gaussian vector. Given  $u \in E$ ,  $\varepsilon > 0$ , choose  $n_0 \in \mathbb{N}$ ,  $\rho > 0$

such that  $\|u - \rho u_{n_0}\| < \varepsilon/2$ . Then

$$\begin{aligned} 0 &< \prod_{n \neq n_0}^{\infty} \mathbb{P}\{|g_n| \leq (\varepsilon/4) a_n\} \mathbb{P}\{|b_{n_0} g_{n_0} - \rho| \leq \varepsilon/4\} \\ &= \mathbb{P}\left(\bigcap_{n \neq n_0} \left\{|g_n| \leq \frac{\varepsilon}{4} a_n\right\} \cap \left\{|b_{n_0} g_{n_0} - \rho| \leq \frac{\varepsilon}{4}\right\}\right) \\ &\leq \mathbb{P}\left(\left\|\sum_{n \neq n_0} b_n g_n u_n + (b_{n_0} g_{n_0} - \rho) u_{n_0}\right\| \leq \frac{\varepsilon}{2}\right) \\ &\leq \mathbb{P}\{\|Z - u\| < \varepsilon\}. \end{aligned}$$

Since obviously the same argument applies to  $\sigma Z$ , the result is proved.  $\square$

Let  $\gamma_\sigma = \mathcal{L}(\sigma Z)$ . We introduce the Markov kernel

$$P_\sigma((x, u), D) = (P(x, \cdot) \otimes \gamma_\sigma)(D), \quad (x, u) \in S \times E, D \in \mathcal{S} \otimes E.$$

Note that the right side is independent of  $u$ . Let  $f: S \rightarrow E$  be an arbitrary measurable function. We will extend  $f$  to  $S \times E$  in a suitable way and define the truncations of  $P_\sigma$  at suitable sets of boundedness of the new function. Let  $g: S \times E \rightarrow E$  be defined by

$$g(x, u) = f(x) + u, \quad x \in S, u \in E.$$

Let  $C_k = g^{-1}(B_0(k))$  for  $k \geq 1$ , and let  $P_{\sigma,k}$  be the truncation of  $P_\sigma$  at  $C_k$ :

$$P_{\sigma,k}((x, u), D) = P_\sigma((x, u), D), \quad (x, u) \in C_k, D \in (\mathcal{S} \otimes E)_k,$$

where  $(\mathcal{S} \otimes E)_k = \{D \in \mathcal{S} \otimes E: D \subset C_k\}$ ; of course,  $P_{\sigma,k}$  is a sub-Markov kernel. The important property of the smoothed kernel  $P_\sigma$  is that irreducibility is preserved by truncations. (This property need not be true for  $P$ .)

LEMMA 3.2. *Let  $\varphi$  be an irreducibility measure for  $P$ . Then:*

- (i)  $P_\sigma$  is irreducible with irreducibility measure  $\varphi \otimes \gamma_\sigma$ ;
- (ii)  $P_{\sigma,k}$  is irreducible with irreducibility measure  $(\varphi \otimes \gamma_\sigma)_k$ , where  $(\varphi \otimes \gamma_\sigma)_k = (\varphi \otimes \gamma_\sigma)|_{(\mathcal{S} \otimes E)_k}$ .

PROOF. (i) It is easily verified from the definition of  $P_\sigma$  that for all  $n \geq 1$ ,

$$(3.1) \quad P_\sigma^n((x, u), D) = (P^n(x, \cdot) \otimes \gamma_\sigma)(D), \quad D \in \mathcal{S} \otimes E.$$

Assume now  $(\varphi \otimes \gamma_\sigma)(D) > 0$ ; given  $(x, u) \in S \times E$ , we must prove that

$$(3.2) \quad P_\sigma^n((x, u), D) > 0 \quad \text{for some } n \geq 1.$$

Let  $D_y = \{u \in E: (y, u) \in D\}$ ,  $y \in S$  and let  $h(y) = \gamma_\sigma(D_y)$ . By Fubini's theorem,

$$\int_S h(y) \varphi(dy) = (\varphi \otimes \gamma_\sigma)(D) > 0.$$

By the irreducibility of  $P$ , there exists  $n \geq 1$  such that

$$(3.3) \quad \int h(y) P^n(x, dy) > 0.$$

By (3.1), (3.3) and Fubini's theorem,

$$P_\sigma^n((x, u), D) = \int P^n(x, dy) h(y) > 0,$$

proving (3.2).

(ii) We first find a convenient expression for the powers of  $P_{\sigma, k}$ . Let  $(x_0, u_0) \in C_k$ ,  $D \in (\mathcal{S} \otimes \mathcal{E})_k$ . Then

$$\begin{aligned} &P_{\sigma, k}^n((x_0, u_0), D) \\ &= \int P_{\sigma, k}((x_0, u_0), d(x_1, u_1)) \\ (3.4) \quad &\int \cdots \int P_{\sigma, k}((x_{n-1}, u_{n-1}), d(x_n, u_n)) \mathbf{1}_D(x_n, u_n) \\ &= \int P(x_0, dx_1) \gamma_\sigma(du_1) \\ &\int \cdots \int P(x_{n-1}, dx_n) \gamma_\sigma(du_n) \prod_{j=1}^{n-1} \mathbf{1}_{C_k}(x_j, u_j) \mathbf{1}_D(x_n, u_n). \end{aligned}$$

But  $\mathbf{1}_{C_k}(x_j, u_j) = \mathbf{1}_{B_k}(g(x_j, u_j)) = \mathbf{1}_{B_k - f(x_j)}(u_j)$ . Therefore by (3.4),

$$\begin{aligned} &P_{\sigma, k}^n((x_0, u_0), D) \\ (3.5) \quad &= \int P(x_0, dx_1) \int \cdots \int P(x_{n-1}, dx_n) \gamma_\sigma(du_n) \\ &\quad \times \prod_{j=1}^{n-1} \gamma_\sigma(B_0(k) - f(x_j)) \mathbf{1}_D(x_n, u_n). \end{aligned}$$

Assume now  $D \subset C_k$ ,  $(\varphi \otimes \gamma_\sigma)(D) > 0$ , and by (i) let  $n \geq 1$  be such that  $P_\sigma^n((x_0, y_0) P_j^n((x_0, y_0), D))$ . Since

$$\prod_{j=1}^{n-1} \gamma_\sigma(B_0(k) - f(x_j)) > 0 \quad \text{for all } (x_1, \dots, x_{n-1}) \in S^{n-1},$$

if  $P_{\sigma, k}^n((x_0, u_0), D) = 0$  we would have

$$\begin{aligned} 0 &= \int P(x_0, dx_1) \int \cdots \int P(x_{n-1}, dx_n) \gamma_\sigma(du_n) \mathbf{1}_D(x_n, u_n) \\ &= \int P^n(x_0, dx_n) \gamma_\sigma(du_n) \mathbf{1}_D(x_n, u_n) \\ &= P_\sigma^n((x_0, u_0), D), \end{aligned}$$

a contradiction.  $\square$

We will need the following analytical property of the convergence parameter  $R(K_\xi)$  of the transform kernel  $K_\xi$  associated to an irreducible sub-Markov kernel. Lemma 3.3 generalizes Corollary 4.3 of [3].

LEMMA 3.3. *Let  $P$  be an irreducible sub-Markov kernel. Let  $f: S \rightarrow E$  be a bounded measurable function, and for  $\xi \in E^*$  define*

$$K_\xi(x, A) = \int_A \exp(\langle \xi, f(y) \rangle) P(x, dy), \quad x \in S, A \in \mathcal{S}.$$

*Let  $R(K_\xi)$  be the convergence parameter of  $K_\xi$  and define  $\Lambda: E^* \rightarrow \mathbb{R}$  by  $\Lambda(\xi) = -\log R(K_\xi)$ .*

*Then  $\Lambda$  is a proper, convex,  $w^*$ -lower semicontinuous function.*

PROOF. The fact that  $\Lambda$  is a proper convex function is proved in [3], Lemma 4.1.

(I) Since  $P$  is irreducible, by Theorem 2.1 of [21] there exists  $m \geq 1$ , a small function  $s$  on  $S$  and a small probability measure  $\nu$  on  $(S, \mathcal{S})$  such that

$$P^m(x, A) \geq s(x)\nu(A), \quad x \in S, A \in \mathcal{S};$$

moreover,  $m$  can be chosen so that  $P^m$  is irreducible (see [21], pages 20–22).

[Note that one can take  $s(x) = \delta \mathbf{1}_C(x)$  with  $\delta, C$  as in Section 2.]

Let  $B_r^* = \{\xi \in E^*: \|\xi\| \leq r\}$ ,  $r > 0$ . If  $\xi \in B_r^*$ , then

$$\begin{aligned} K_\xi^m(x, A) &= \int P(x, dx_1) \int \cdots \int P(x_{m-1}, dx_m) \\ &\quad \times \exp(\langle \xi, f(x_1) + \cdots + f(x_m) \rangle) \mathbf{1}_A(x_m) \\ (3.6) \quad &\geq \exp(-mrc) P^m(x, A) \\ &\geq t(x)\nu(A) \quad (\text{uniformly for } \xi \in B_r^*), \end{aligned}$$

where  $c = \sup_{x \in S} \|f(x)\|$  and  $t = \exp(-mrc)s$ . The argument leading to (3.6) also shows: for all  $f \in B_r^*$ ,  $K_\xi$  has the same period as  $P$  and therefore  $K_\xi^m$  is irreducible ([21], page 20–22).

(II) For fixed  $r > 0$  and  $t$  as in (3.6), we define on  $B_r^*$ ,

$$h_n(\xi) = \nu(K_\xi^m - t \otimes \nu)^n t, \quad n \geq 1.$$

Then  $h_n$  is  $w^*$ -continuous, for, by the Banach–Alaoglu theorem (see, e.g., [10], page 138)  $B_r^*$  endowed with the  $w^*$ -topology is a compact metrizable space and therefore it is enough to prove that if  $\{\xi_k\}_{k \geq 1}$  is a sequence in  $B_r^*$ , then

$$(3.7) \quad \xi_k \xrightarrow{w^*} \xi \text{ implies } \lim_k h_n(\xi_k) = h_n(\xi).$$

However, (3.7) is easily proved by induction on  $n$  and repeated applications of the dominated convergence theorem.

(III) It follows from (II) that for each  $\rho > 0$ , the function  $b(\cdot, \rho)$  defined on  $B_r^*$  by

$$(3.8) \quad b(\xi, \rho) = \sum_{n=0}^{\infty} \rho^{n+1} h_n(\xi)$$

is  $w^*$ -lower semicontinuous.

(IV) Next, on  $B_r^*$ , define  $\psi(\xi) = R(K_\xi^m)$ . Then  $\psi$  is  $w^*$ -upper semicontinuous. For, arguing as in (II), it is enough to show that

$$\xi_k \in B_r^*, \xi_k \xrightarrow{w^*} \xi \text{ imply } \limsup_k \psi(\xi_k) \leq \psi(\xi).$$

Let us recall that by an equivalent formulation of the definition of convergence parameter (Proposition 4.2 of [21]),

$$R(K_\xi^m) = \sup\{\rho \geq 0: b(\xi, \rho) \leq 1\}.$$

Let  $\varepsilon > 0$ . Then  $b(\xi, \psi(\xi) + \varepsilon) > 1$  and by (III) we have

$$\liminf_k b(\xi_k, \psi(\xi) + \varepsilon) \geq b(\xi, \psi(\xi) + \varepsilon) > 1,$$

which implies that there exists  $k_0 \geq 1$  such that for  $k \geq k_0$ ,  $b(\xi_k, \psi(\xi) + \varepsilon) > 1$ , and therefore  $\psi(\xi_k) < \psi(\xi) + \varepsilon$ . It follows that

$$\limsup_k \psi(\xi_k) \leq \psi(\xi) + \varepsilon.$$

However,  $\varepsilon$  is arbitrary. This proves the  $w^*$ -upper semicontinuity of  $\psi$ .

(V)  $\Lambda \upharpoonright B_r^*$  is  $w^*$ -lower semicontinuous, for, by the proof of Proposition 3.5 of [21],  $R(K_\xi^m) = (R(K_\xi))^m$  and therefore

$$\Lambda(\xi) = -\log(R(K_\xi^m))^{1/m} = -\frac{1}{m} \log \psi(\xi).$$

The claim now follows from (IV).

(VI) Finally we prove that  $\Lambda$  is  $w^*$ -lower semicontinuous on  $E^*$ . By the convexity of  $\Lambda$ , the sets  $L_a = \{\xi \in E^*: \Lambda(\xi) \leq a\}$  are convex for each  $a \in \mathbb{R}$  and by (V),  $L_a \cap B_r^*$  is  $w^*$ -closed for each  $r > 0$ . By the Krein–Smulyan theorem (see, e.g., [10], page 163) it follows that  $L_a$  is  $w^*$ -closed for all  $a \in \mathbb{R}$ ; that is,  $\Lambda$  is  $w^*$ -lower semicontinuous.  $\square$

For  $(x, u) \in S \times E$ ,  $D \in (S \otimes E)$ ,  $\xi \in E^*$ , let

$$K_{\xi, \sigma}((x, u), D) = \int_D \exp(\langle \xi, g(y, v) \rangle) P_\sigma((x, u), d(y, v)).$$

The kernel  $K_{\xi, \sigma}$  is irreducible since  $P_\sigma$  is so; let  $R(K_{\xi, \sigma})$  be its convergence parameter and  $\Lambda_\sigma(\xi) = -\log R(K_{\xi, \sigma})$ . Also, for  $(x, u) \in C_k$ ,  $D \in (S \otimes E)_k$ ,  $\xi \in E^*$ , let

$$K_{\xi, \sigma, k}((x, u), D) = \int_D \exp(\langle \xi, g(y, v) \rangle) P_{\sigma, k}((x, u), d(y, v)).$$

Then  $K_{\xi, \sigma, k}$  is irreducible since  $P_{\sigma, k}$  is so; let  $R(K_{\xi, \sigma, k})$  be its convergence parameter and  $\Lambda_{\sigma, k}(\xi) = -\log R(K_{\xi, \sigma, k})$ .

LEMMA 3.4. *Let  $\{Y_n\}_{n \geq 1}$  be an  $S \times E$ -valued Markov chain with transition kernel  $P_\sigma$  and initial distribution  $\lambda$ . Then for every open set  $G$  in  $E$ ,*

$$\liminf_n \frac{1}{n} \log \mathbb{P}_\lambda \left\{ \frac{1}{n} \sum_{j=1}^n g(Y_j) \in G \right\} \geq - \inf_{u \in G} \Lambda_\sigma^*(u).$$

PROOF. Let  $\lambda_k = \lambda|_{(\mathcal{S} \otimes E)_k}$ . Since  $g$  is bounded on  $(C_k)$ , we have for each  $k \geq 1$ , by Lemma 3.2(ii) and Proposition 2.1 that

$$\begin{aligned}
 & \liminf_n \frac{1}{n} \log \mathbb{P}_\lambda \left\{ \frac{1}{n} \sum_{j=1}^n g(Y_j) \in G \right\} \\
 (3.9) \quad & \geq \liminf_n \frac{1}{n} \log \int \lambda_k(d(x_1, u_1)) \int P_{\sigma, k}((x_1, u_1), d(x_2, u_2)) \\
 & \quad \int \cdots \int P_{\sigma, k}((x_{n-1}, u_{n-1}), d(x_n, u_n)) \mathbf{1}_G \left( \frac{1}{n} \sum_{j=1}^n g(x_j, u_j) \right) \\
 & \geq - \inf_{u \in G} \Lambda_{\sigma, k}^*(u).
 \end{aligned}$$

Since clearly  $K_{\xi, \sigma, k}((x, u), D \cap C_k) \uparrow K_{\xi, \sigma}((x, u), D)$  as  $k \rightarrow \infty$  for all  $(x, u) \in S \times E, D \in \mathcal{S} \otimes E$ , it follows from Theorem 2.1 of [3] that

$$R(K_{\xi, \sigma, k}) \downarrow R(K_{\xi, \sigma}), \quad \Lambda_{\sigma, k}(\xi) \uparrow \Lambda_\sigma(\xi).$$

Now  $\Lambda_\sigma(0) \leq 0$  [because  $R(K_{0, \sigma}) \geq 1$ ] and since the convergence parameter is always finite ([21], Theorem 3.2), the function  $\Lambda_\sigma$  is proper. By Lemma 3.3, the functions  $\Lambda_{\sigma, k}$  satisfy the assumptions of Theorem B.3 of [3] and hence by that theorem, given  $u_0 \in G$ , there exists a sequence  $\{u_k\}_{k \geq 1}$  in  $E$  (hence eventually in  $G$ ) such that  $u_k \rightarrow u_0$  and

$$(3.10) \quad \limsup_k \Lambda_{\sigma, k}^*(u_k) \leq \Lambda_\sigma^*(u_0).$$

Therefore by (3.9) and (3.10),

$$\liminf_n \frac{1}{n} \log \mathbb{P} \left\{ \frac{1}{n} \sum_{j=1}^n g(Y_j) \in G \right\} \geq -\Lambda_\sigma^*(u_0). \quad \square$$

LEMMA 3.5. For all  $\xi \in E^*, \sigma > 0$ ,

$$\Lambda_\sigma(\xi) = \Lambda(\xi) + \sigma^2 V(\xi),$$

where  $V(\xi) = \frac{1}{2} \int \langle \xi, u \rangle^2 \gamma_1(du)$ .

PROOF. Let

$$\begin{aligned}
 \gamma_{\xi, \sigma}(B) &= \int_B \exp(\langle \xi, u \rangle) \gamma_\sigma(du), \quad B \in E, \\
 \hat{\gamma}_\sigma(\xi) &= \int \exp(\langle \xi, u \rangle) \gamma_\sigma(du) = \exp\{\sigma^2 V(\xi)\}.
 \end{aligned}$$

Then it is easily verified that for all  $n \geq 1, D \in \mathcal{S} \otimes E$ ,

$$K_{\xi, \sigma}^n((x, u), D) = (\hat{\gamma}_\sigma(\xi))^{n-1} (K_\xi^n(x, \cdot) \otimes \gamma_{\xi, \sigma})(D).$$

As remarked in the Introduction, there exists  $m \geq 1$  such that the kernel  $K_\xi$  has the minorization

$$K_\xi^m(x, A) \geq s(x) \nu(A), \quad x \in S, A \in \mathcal{S},$$

for a certain small function  $s$  and a certain small probability measure  $\nu$ . Then it is easily verified that

$$K_{\xi, \sigma}^m((x, u), D) \geq (\hat{\gamma}_\sigma(\xi))^{m-1} s(x)(\nu \otimes \gamma_{\xi, \sigma})(D),$$

$$(x, u) \in S \times E, D \in S \otimes E.$$

By [21], Proposition 3.4,  $R(K_{\xi, \sigma})$  [resp.,  $R(K_\xi)$ ] is the radius of convergence of the power series with coefficients

$$(3.11) \quad (\nu \otimes \gamma_{\xi, \sigma}) K_{\xi, \sigma}^n s = (\nu K_\xi^n s)(\hat{\gamma}_\sigma(\xi))^n$$

(resp.,  $\nu K_\xi^n s$ ). It follows from (3.11) that

$$R(K_{\xi, \sigma}) = R(K_\xi)(\hat{\gamma}_\sigma(\xi))^{-1},$$

which implies the conclusion.  $\square$

PROOF OF THEOREM. Let  $\mathbb{P}^\mu$  be the Markovian probability measure on  $(S, \mathcal{S})^\mathbb{N}$  associated to  $P$  and the initial distribution  $\mu$ , and let  $\{X_n\}_{n \geq 1}$  be the coordinate projections. Let  $\{Z_n^\sigma\}_{n \geq 1}$  be an independent sequence of  $E$ -valued Gaussian vectors with common law  $\gamma_\sigma$ , defined on  $(\Omega, \mathcal{A}, \mathbb{P})$ . We regard  $\{X_n\}_{n \geq 1}$  and  $\{Z_n^\sigma\}_{n \geq 1}$  as being defined on  $S^\mathbb{N} \times \Omega$ , endowed with  $\mathbb{P}' = \mathbb{P}^\mu \otimes \mathbb{P}$ ; also, let  $Y_n = (X_n, Z_n^\sigma)$ . It is easily verified that  $\{Y_n\}_{n \geq 1}$  is an  $S \times E$ -valued Markov chain with transition kernel  $P_\sigma$  and initial distribution  $\mu \otimes \gamma_\sigma$ .

Given an open set  $G$  in  $E$  and  $u \in G$ , let  $\varepsilon > 0$  be such that  $u \in G_\varepsilon$ , where  $G_\varepsilon = \{u \in E: d(u, G^c) > \varepsilon\}$ . Then, since  $g(Y_j) = f(X_j) + Z_j^\sigma$ , we have for all  $n \geq 1$ ,

$$\mathbb{P}' \left\{ \frac{1}{n} \sum_{j=1}^n g(Y_j) \in G_\varepsilon \right\} \leq \mathbb{P}' \left\{ \frac{1}{n} \sum_{j=1}^n f(X_j) \in G \right\} + \mathbb{P} \left\{ \left\| \frac{1}{n} \sum_{j=1}^n Z_j^\sigma \right\| \geq \varepsilon \right\}$$

and by Lemma 3.4,

$$(3.12) \quad -\Lambda_\sigma^*(u) \leq \liminf_n \frac{1}{n} \log \mathbb{P}' \left\{ \frac{1}{n} \sum_{j=1}^n g(Y_j) \in G_\varepsilon \right\}$$

$$\leq \max \left( \liminf_n \frac{1}{n} \log \mathbb{P}' \left\{ \frac{1}{n} \sum_{j=1}^n f(X_j) \in G \right\}, I(\varepsilon, \sigma) \right),$$

where  $I(\varepsilon, \sigma) = \limsup_n (1/n) \log \mathbb{P} \{ \|(1/n) \sum_{j=1}^n Z_j^\sigma\| \geq \varepsilon \}$ . By the well-known large deviation theorem for Gaussian measures (see, e.g., [6]),

$$I(\varepsilon, \sigma) \leq -\frac{1}{\sigma^2} \inf_{\|u\| \geq \varepsilon} V^*(u),$$

where  $V(\xi) = \frac{1}{2} \int \langle \xi, u \rangle^2 \gamma_1(du)$ . Since  $\inf_{\|u\| \geq \varepsilon} V^*(u) > 0$ , as is easily seen, we have

$$(3.13) \quad \lim_{\sigma \rightarrow 0} I(\varepsilon, \sigma) = -\infty.$$

On the other hand, by Lemma 3.5

$$\begin{aligned}
 \lim_{\sigma \rightarrow 0} \Lambda_{\sigma}^*(u) &= \lim_{\sigma \rightarrow 0} \sup_{\xi} [\langle \xi, u \rangle - (\Lambda(\xi) + \sigma^2 V(\xi))] \\
 &= \sup_{\sigma} \sup_{\xi} [\langle \xi, u \rangle - (\Lambda(\xi) + \sigma^2 V(\xi))] \\
 (3.14) \quad &= \sup_{\xi} \sup_{\sigma} [\langle \xi, u \rangle - (\Lambda(\xi) + \sigma^2 V(\xi))] \\
 &= \sup_{\xi} [\langle \xi, u \rangle - \Lambda(\xi)] \\
 &= \Lambda^*(u).
 \end{aligned}$$

The conclusion of the Theorem follows now from (3.13) and (3.14) by letting  $\sigma \rightarrow 0$  in (3.12).  $\square$

**4. Examples, comments.**

4a. *The Cramér theorem for Banach space valued r.v.'s* [14], [7], [2]. In the i.i.d. case, the argument in Section 2 simplifies, and applies without any restriction on the r.v.'s. By subadditivity  $\mu_n = \mathcal{L}(S_n/n)$  satisfies the weak LDP and the upper bound in Lemma 2.5. Then by the modification of Varadhan's theorem in Theorem 1.1 of [12] (and discussion with the author), for every  $M > 0$ ,

$$\begin{aligned}
 (4.1) \quad &\lim_{n \rightarrow \infty} \frac{1}{n} \log \int \exp(n[\langle \xi, u \rangle \wedge M]) d\mu_n \\
 &= \sup_{u \in E} [\langle \xi, u \rangle \wedge M - I(u)], \quad \xi \in E^*
 \end{aligned}$$

and hence

$$(4.2) \quad \lim_{M \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \log \int \exp(n[\langle \xi, u \rangle \wedge M]) d\mu_n = I^*(\xi).$$

Clearly

$$I^*(\xi) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \int \exp(n\langle \xi, u \rangle) d\mu_n = \Lambda(\xi).$$

In the other direction, by (4.2),

$$\begin{aligned}
 I^*(\xi) &\geq \lim_{M \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \exp \left\{ n \left[ \frac{\sum_1^n \langle X_i, \xi \rangle \wedge M}{m} \right] \right\} \\
 &= \lim_{M \rightarrow \infty} \mathbb{E} \exp[\langle X_1, \xi \rangle \wedge M] \\
 &\quad \text{because } [\langle X_i, \xi \rangle \wedge M], \ i = 1, 2, \dots \text{ are i.i.d.} \\
 &= \Lambda(\xi).
 \end{aligned}$$

Thus  $I(u) = \Lambda^*(u)$  and the weak LD theorem with rate  $\Lambda^*$  follows. As is well known (see, e.g., [2]), if furthermore

$$E \exp(t \|X_1\|) < \infty \quad \text{for all } t > 0,$$

then  $\{\mu_n\}$  is exponentially tight, and the full LDP holds.

4b.  $\Lambda$  is  $w^*$ -lsc. It follows from Lemma 3.5 that  $\Lambda$  is  $w^*$ -lower semicontinuous (even if  $f$  is not bounded). For, by the proof of Lemma 3.4,  $\Lambda_\sigma$  is  $w^*$ -lower semicontinuous on  $E^*$ . On the other hand, it is easy to show that  $V|B_r^*$  is sequentially  $w^*$ -continuous (using the integrability of Gaussian vector), and therefore  $V|B_r^*$  is  $w^*$ -continuous. By Lemma 3.5,  $\Lambda|B_r^*$  is  $w^*$ -lower semicontinuous for each  $r > 0$ . Applying the Krein–Smulyan theorem as in the proof of Lemma 3.3, we conclude that  $\Lambda$  is  $w^*$ -lower semicontinuous.

4c. *Upper bounds and tightness.* In large classes of cases, the rate  $I = \Lambda^*$  in the lower bound in our theorem is tight in the sense that it is matched by an upper bound (1.2) with the same rate function  $\Lambda^*$ . A general upper bound is given in Theorem 4.2 of [1] in terms of  $\sigma^*$ , where  $\sigma(\xi) = \log r(K_\xi)$ ,  $r(K_\xi) =$  the spectral radius of  $K_\xi$ . It is always the case that

$$(4.3) \quad \sigma^* \leq \Lambda^*,$$

but there may be strict inequality. See also [22]. Thus there may be a gap between the lower bound given via the convergence parameter and the upper bound given via the spectral radius. There remain unsettled questions about the validity of the full LDP at the level of generality of this paper. Here are some special cases where upper and lower bound rate functions are the same function  $\Lambda^*$  as in our theorem.

1. If the state space  $S$  is finite, then the full LDP holds with rate  $I = \Lambda^*$  for any  $f$ .
2. Conditions under which the empirical measure of  $\{X_n\}$  satisfies the full LDP are given in the original papers [14] and later in [22], [4], [16] and [13]. In this case a more explicit formula for the rate function is given.
3. If  $E = \mathbb{R}^d$  then a class of sets, called  $s$ -sets is characterized in Section 4 of [21]. If  $A$  is an  $s$ -set, then it is shown in [21] (for arbitrary  $f$ ) that

$$(4.4) \quad \limsup \frac{1}{n} \log \mathbb{P}_n \left\{ \frac{S_n}{n} \in F, X_n \in A \right\} \leq -\Lambda^*(F)$$

for compact  $F$ . If also  $O \in \text{Interior Dom}(\Lambda)$ , then (4.4) holds for closed  $F$ . If the whole space  $S$  is an  $s$ -set, then taking  $A = S$  we see that  $\mathbb{P}_x\{S_n/n \in \cdot\}$  satisfies the LD upper bound with rate  $\Lambda^*$ , hence the full LDP. In particular, if  $f$  is bounded and if for some minorization set  $C$  and some  $n \geq 1$ ,  $P^n(x, C) \geq \delta > 0$  for all  $x \in S$ , then  $S$  is an  $s$ -set and  $\mathbb{P}_x\{S_n/n \in \cdot\}$  satisfies the upper bound with rate  $\Lambda^*$ . Weaker conditions can be derived from the above reference.

For example, if the state space is countable and  $\{X_n\}$  is irreducible and uniformly recurrent in the sense that  $p(x, x_0) \geq \delta > 0$  for all  $x \in S$  and some  $x_0 \in S$ , and if  $f$  is bounded, then  $\mathbb{P}_x(S_n/n \in \cdot)$  satisfies the full LDP with rate  $\Lambda^*$ . Of course such results hold much more generally.

4d. *Examples and counterexamples.*

(A) There are many easily described examples for which  $P\{S_n/n > a\}$  decays superexponentially, so that

$$\lim \frac{1}{n} \log P\left\{\frac{S_n}{n} > a\right\} \text{ exists and } = -\infty.$$

Clearly there are also examples with large tails where this limit exists and equals 0. In these cases the convergence parameter and spectral radius (respectively) give correct (but trivial) answers.

However, as the following example shows, there are cases where the convergence parameter and spectral radius give trivial *bounds*, but where the correct limit exists and lies strictly in between these bounds. Namely, this is an example of a uniformly ergodic chain on  $S = \{0, 1, \dots\}$  with  $f: S \rightarrow \mathbb{Z}$ , for which these rate candidates only enable us to conclude that for  $0 < a < \infty$ ,

$$-\infty \leq \liminf \frac{1}{n} \log \mathbb{P}_\mu\left\{\frac{S_n}{n} > a\right\} \leq \limsup \frac{1}{n} \log \mathbb{P}_\mu\left\{\frac{S_n}{n} \geq a\right\} \leq 0.$$

Nevertheless, the exact asymptotics can be worked out directly.

Let  $\{X_n; n \geq 0\}$  have transition function  $\{P(i, j)\}$  given by

$$P(0, 3j) = p_j > 0, \quad j = 1, 2, \dots, \quad P(0, 0) = p_0 > 0,$$

$$P(3j, 3j - 1) = P(3j - 1, 3j - 2) = P(3j - 2, 0) = 1, \quad j \geq 1.$$

Take  $f(3j) = j, j = 0, 1, \dots$ ,

$$f(3j - 1) = -2j, \quad f(3j - 2) = j, \quad j = 1, 2, \dots$$

Now we can take the minorization measure  $\nu$  and set  $C$  to be  $C = \{0\}$ ,  $\nu(j) = \delta_0(j), \delta = P(0, 0)$ . Then it can be easily checked that

$$(4.5) \quad \nu K_\xi^n(C) = K_\xi^n(0, 0) = E_0[\exp(\xi S_n), X_n = 0] \equiv P^n(0, 0).$$

Furthermore  $\{P(i, j)\}$  has invariant measure  $\pi$  given by

$$(4.6) \quad \begin{aligned} \pi(0) &= (4 - 3p_0)^{-1}, \\ \pi(3j) &= \pi(3j - 1) = \pi(3j - 2) = (4 - 3p_0)^{-1} p_j, \quad j = 1, 2, \dots \end{aligned}$$

Hence

$$R(K_\xi) = 1, \quad \Lambda(\xi) = 0 \quad \text{for all } \xi$$

and

$$\Lambda^*(x) = \begin{cases} 0, & \text{if } x = 0, \\ \infty, & \text{if } x \neq 0. \end{cases}$$

Now if  $G$  is an open set and  $0 \in G$  then  $P_0\{S_n/n \in G\} \rightarrow 1$ , while if  $0 \notin G$  then the convergence parameter yields the uninteresting lower bound

$$\liminf \frac{1}{n} \log \mathbb{P}_0 \left\{ \frac{S_n}{n} \in G \right\} \geq -\infty.$$

On the other hand,

$$K_\xi^n(i, 0) = E_i[\exp(\xi S_n), X_n = 0]$$

so that

$$K_\xi^n(3j, 0) = \exp(-\xi j), \quad K_\xi^n(3j - i, 0) = \exp(\xi j).$$

Thus

$$r(K_\xi) = \begin{cases} 1, & \text{if } \xi = 0, \\ \infty, & \text{if } \xi \neq 0, \end{cases}$$

and  $\sigma^*(x) \equiv 0$  for all  $x$ . [Since  $\sup_j K_\xi(i, S) = +\infty$ , the spectral radius is taken to be  $+\infty$ .] This yields the trivial upper bound

$$\limsup \frac{1}{n} \log P \left\{ \frac{S_n}{n} \in F \right\} \leq 0,$$

and neither the convex conjugate of the convergence parameter nor the conjugate of the spectral radius tell us anything in this example. However, by (4.6),

$$\lim_n P\{X_n = 0\} = (4 - 3p_0)^{-1}$$

and hence

$$\begin{aligned} \mathbb{P}_0\{S_n \geq 3an\} &= \mathbb{P}_0\{X_{n-1} = 0\} \sum_{j \geq an} p_j \\ &\sim (4 - 3p_0)^{-1} \sum_{j \geq an} p_j. \end{aligned}$$

(B) In the next example, the spectral radius still leads to a trivial upper bound of  $\sigma^*(x) \equiv 0$ , but the convergence parameter yields a nondegenerate rate for a full LDP. Let  $S = \{0, 1, 2, \dots\}$  and  $P(i, j)$  be as in example (A), but now take  $f(3j) = 0, j \geq 0, f(3j - 1) = 2a_j, f(3j - 2) = -a_j$  for  $j \geq 1$ , where  $0 < a_j \nearrow \infty$ , and take  $p_j$  so that  $\varphi(\xi) = \sum_{j=1}^\infty \exp(\xi a_j) p_j < \infty$  for all  $\xi \in (-\infty, \infty)$ .

Then we still have

$$K_\xi^n(3j, 0) = \exp(a_j \xi), \quad K_\xi^n(3j - 1, 0) = \exp(-a_j \xi)$$

and  $r(K_\xi) = 1$  if  $\xi = 0, = \infty$  if  $\xi \neq 0$ . Thus  $\sigma^*(x) = 0$ .

However, letting  $\tau = \inf\{n \geq 1: S_n = 0\}$  and noting that

$$\mathbb{P}_0(\tau = 1, S_\tau = 0) = p_0, \quad \mathbb{P}_0(\tau = 4, S_\tau = a_j) = p_j,$$

we see that

$$\psi(\xi, \zeta) \stackrel{\text{def}}{=} \mathbb{E}_0 \exp(\xi S_\tau - \zeta \tau) = p_0 \exp(-\zeta) + \varphi(\xi) \exp(-4\zeta).$$

By [20],  $\Lambda(\xi) = -\log R(K_\xi)$  is the unique solution  $\Lambda(\xi) = \zeta$  of  $\psi(\xi, \zeta) = 1$ , and hence for fixed  $j$  the sequence of measures

$$\mathbb{P}_i \left\{ \frac{S_n}{n} \in \cdot, X_n = j \right\}$$

satisfy the full LDP with rate  $\Lambda^*$ .

(C) An interesting counterexample where the full LDP does *not* hold is given in [9]. They give an example of a uniformly recurrent, irreducible chain on  $\mathbb{Z}^2$ , with bounded  $f$ , and invariant probability measure  $\mu$ , where  $\mathbb{P}_\mu\{S_n/n \in \cdot\}$  does *not* satisfy the LDP. Since the sufficient conditions mentioned in Section 4c(III) are satisfied, this would appear to be a contradiction. The point is that one here has an initial measure  $\mu$ , rather than a fixed state. This shows the sensitivity of the LDP to initial measures. Note that by the results of the present paper, the *lower* bound does hold with rate  $\Lambda^*$ .

(D) There is a class of examples in [12], [4] and [16] of the following type. Let  $S$  be a finite set,  $S = C_1 \cup C_2 \cup C_3$  where  $C_i$  are irreducible (in the classical sense) classes for  $\{X_n\}$  satisfying the communication relations  $C_3 \rightarrow C_2 \rightarrow C_1$ , with  $C_1$  and  $C_1 \cup C_2$  closed classes. This chain will be  $\varphi$ -irreducible if  $\varphi(C_1) = 1$  and in that case  $n^{-1} \sum_1^n f(X_i)$  will satisfy the LDP with rate  $\Lambda^*$  if the initial state  $x_0 \in C_1$ . If  $x_0 \in C_2$  then neither  $\Lambda^*$  nor  $\sigma^*$  gives the right rate function.

## REFERENCES

- [1] DE ACOSTA, A. (1985). Upper bounds for large deviations of dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **69** 551–565.
- [2] DE ACOSTA, A. (1985). On large deviations of sums of independent random vectors. *Lecture Notes in Math.* **1153** 1–14. Springer, Berlin.
- [3] DE ACOSTA, A. (1988). Large deviations for vector valued functionals of a Markov chain: lower bounds. *Ann. Probab.* **16** 925–990.
- [4] DE ACOSTA, A. (1990). Large deviations for empirical measures of Markov chains. *J. Theoret. Probab.* **3** 396–431.
- [5] ATHREYA, K. B. and NEY, P. E. (1978). A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245** 493–501.
- [6] AZENCOTT, R. (1980). Grandes deviations et applications. *Lecture Notes in Math.* **774** 1–176. Springer, Berlin.
- [7] BAHADUR, R. R. and ZABELL, S. L. (1979). Large deviations of the sample mean in general vector spaces. *Ann. Probab.* **7** 587–621.
- [8] BOLTHAUSEN, E. (1987). Markov process large deviation in the  $\tau$ -topology. *Stochastic Process. Appl.* **33** 1–27.
- [9] BRYC, W. and DEMBO, A. (1996). Large deviations and strong mixing. *Ann. Inst. H. Poincaré* **32** 549–569.
- [10] CONWAY, J. B. (1985). *A Course in Functional Analysis*. Springer, Berlin.
- [11] DEMBO, A. and ZEITOUNI, O. (1993). *Large Deviations Techniques and Applications*. Jones and Bartlett, Boston.
- [12] DINWOODIE, I. H. (1993). Identifying a large deviation rate function. *Ann. Probab.* **21** 216–231.
- [13] DINWOODIE, I. H. and NEY, P. E. (1995). Occupation measures for Markov chains. *J. Theoret. Probab.* **8** 679–691.

- [14] DONSKER, M. D. and VARADHAN, S. R. S. (1975, 1976). Asymptotic evaluation of certain Markov process expectations for large time. I, II, III. *Comm. Pure Appl. Math.* **28** 1–47; **28** 279–301; **29** 389–461.
- [15] EKELAND, I. and TEMAM, R. (1976). *Convex Analysis and Variational Problems*. North-Holland, Amsterdam.
- [16] JAIN, N. C. (1990). Large deviation lower bounds for additive functionals of Markov processes. *Ann. Probab.* **18** 1071–1098.
- [17] LANFORD, O. E. (1973). Entropy and equilibrium states in classical statistical mechanics. *Lecture Notes in Phys.* **20** 1–113. Springer, Berlin.
- [18] MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [19] MILLER, H. (1961). A convexity property in the theory of random variables defined on a finite Markov chain. *Ann. Math. Statist.* **32** 1260–1270.
- [20] NEY, P. E. and NUMMELIN, E. (1987). Markov additive processes. I, II. *Ann. Probab.* **15** 561–592; 593–609.
- [21] NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge Univ. Press.
- [22] SHURENKOV, V. M. (1992). On the relationship between spectral radii and Perron roots. Technical report, Dept. Mathematics Göteborg Univ.
- [23] STROOCK, D. W. (1984). *An Introduction to the Theory of Large Deviations*. Springer, Berlin.
- [24] VARADHAN, S. R. S. (1984). *Large Deviations and Applications*. SIAM, Philadelphia.

DEPARTMENT OF MATHEMATICS  
CASE WESTERN RESERVE UNIVERSITY  
CLEVELAND, OHIO 44106

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF WISCONSIN  
MADISON, WISCONSIN 53706  
E-MAIL: ney@math.wisc.edu