The Annals of Probability
1987, Vol. 15, No. 2, 593–609

MARKOV ADDITIVE PROCESSES II. LARGE DEVIATIONS

BY P. NEY[1] AND E. NUMMELIN

University of Wisconsin, Madison and University of Helsinki

Let $\{(X_n, S_n);\ n = 0, 1, \dots\}$ be a Markov additive process, $\{X_n\}$ taking values in a general state space $\mathbb{E}$, while $\{S_n\} \subset \mathbb{R}^d$. The large deviation principle is shown to hold for $P_x\{(X_n, S_n) \in A \times n\Gamma\}$, $A \subset \mathbb{E}$, $\Gamma \subset \mathbb{R}^d$, the upper bound holding for closed sets Γ, the lower bound for open sets. The only hypothesis for the lower bound is irreducibility of $\{X_n\}$, and nonsingularity of $\{S_n\}$. The rate function is characterized in terms of the transform kernel of P_x.

1. Introduction and summary. Let $\{X_n;\ n = 0, 1, \dots\}$ be a φ-irreducible aperiodic Markov chain on a general state space $(\mathbb{E}, \mathscr{E})$, and $\{S_n = \sum_{i=1}^n \xi_i;\ n = 0, 1, \dots\}$ an additive component or functional taking values in $(\mathbb{R}^d, \mathscr{R}^d)$ such that $\{(X_n, \xi_n);\ n = 0, 1, \dots\}$ is a Markov chain on $\mathbb{E} \times \mathbb{R}^d$ satisfying $P\{(X_{n+1}, \xi_{n+1}) \in A \times \Gamma | X_n = x, \mathscr{F}_n\} = P\{(X_{n+1}, \xi_{n+1}) \in A \times \Gamma | X_n = x\} \equiv P(x, A \times \Gamma)$, $x \in E$, $A \in \mathbb{E}$, $\Gamma \in \mathscr{R}^d$, $\mathscr{F}_n = \sigma(X_0, \dots, X_n, \xi_1, \dots, \xi_n)$. The notation and terminology of [5] carries over here. We say $\{(X_n, S_n)\}$ is an MA-process with *transition kernel* $P = P(x, A \times \Gamma)$ and with *transform kernel* $\hat{P}(\alpha) = \{\hat{P}(x, A; \alpha)\} = \{\int_{R^d} \exp\langle \alpha, s\rangle P(x, A \times ds)\}$. Since $\hat{P}(\alpha)$ is also irreducible $[P^n(x, A) > 0$ implies $\hat{P}^n(x, A; \alpha) > 0]$, its convergence parameter $R(\alpha)$, $0 \le R \le \infty$, always exists (Nummelin [8], Theorem 3.2). Let

$$\bar{l}(x, A, \Gamma) = \limsup \frac{1}{n} \log P^n(x, A \times n\Gamma) \tag{1.1}$$

and

$$\underline{l}(x, A, \Gamma) = \liminf \frac{1}{n} \log P^n(x, A \times n\Gamma). \tag{1.2}$$

Following the terminology of Varadhan [9], we say that the family of measures $\{P^n(x, A \times \cdot);\ n = 0, 1, \dots\} \equiv \mathscr{P}(x; A)$ obeys the large deviation principle (LDP) with exponent or rate function $I(\cdot)$, if $I(\cdot)\colon \mathbb{R}^d \to [0, \infty)$ is lower semi-continuous (l.s.c.), and if for all closed sets $F \in \mathscr{R}^d$ and open $G \in \mathscr{R}^d$

$$\bar{l}(x, A, F) \le -\bar{I}(F), \tag{1.3}$$

and

$$\underline{l}(x, A, G) \ge -\bar{I}(G), \tag{1.4}$$

Received May 1985; revised May 1986.

[1] Part of the research for this paper was done while this author was a Fulbright scholar and visiting professor at the University of Helsinki.

AMS 1980 *subject classifications.* Primary 60F10, 60K15; secondary 60J05.

Key words and phrases. Large deviations, Markov chain, Markov additive process.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.*
www.jstor.org

where

$$\bar{I}(\Gamma) = \inf\{I(v): v \in \Gamma\}, \qquad \Gamma \in \mathscr{R}^d.$$

(We do not include compactness of the level sets of I in the definition as in [9].) Refer to (1.3) as the *LDP upper bound* and (1.4) as the *LDP lower bound*, with exponent $I(\cdot)$.

In [5] we proved a number of properties of eigenvalues and functions of the transform kernel $\hat{P}(\alpha)$, and from these obtained limit laws for $\{(X_n, S_n), n = 0, 1, \dots\}$, including an LDP. We assumed hypotheses on $\{(X_n, S_n)\}$ and $\hat{P}(\alpha)$ which were tantamount to geometric recurrence. In the present paper we introduce a new approximation scheme by which these restrictions are removed.

Let $\Lambda(\alpha) = -\log R(\alpha)$. This function will be expressed below in terms of a regeneration structure of the MA-process. $e^{\Lambda(\alpha)}$ may or may not be an eigenvalue of the MA-process, i.e., of $\hat{P}(\alpha)$.

Let $\Lambda^*(\cdot)$ = the convex conjugate of Λ.

We will prove the following results: Subject to an irreducibility condition on the MA-process:

(i) $\mathscr{P}(x, A)$ always satisfies the LDP *lower* bound with exponent Λ^* for all φ-positive sets $A \in \mathscr{E}$ (Theorem 1, Section 3).

(ii) If $d = 1$, then $\mathscr{P}(x, A)$ satisfies the LDP *upper* bound with exponent Λ^* a.s. $x[\varphi]$, provided A is "not too large" in a sense given below (see Definition 4.1). We call these sufficiently small sets s-sets.

(iii) If $2 \le d < \infty$, then $\mathscr{P}(x; A)$ satisfies the LDP *upper* bound with exponent Λ^* a.s. $x[\varphi]$ for A an s-set, provided the origin (in $\mathbb{R}^d$) is in the interior of the convergence domain $\mathscr{D}$ of a generating function determined by P. If F in (1.3) is compact, this condition is not needed. (Theorem 2, Section 4).

REMARKS. (i) All atoms of $\{X_n\}$ are s-sets, as are all finite sets when $\mathbb{E}$ is countable.

(ii) If $d = 1$ and $\mathbb{E}$ is a countable state space, then $\mathscr{P}(i, j)$ always satisfies the LDP upper bound, as well as the lower bound, subject only to the irreducibility of $\{X_n\}$.

In particular, e.g., if $f: \mathbb{E} \times \mathbb{E} \to \mathbb{R}^1$, then

$$\limsup \frac{1}{n} \log \mathbb{P}_i\left\{X_n = j, \frac{1}{n} \sum_{k=0}^{n} f(X_k, X_{k+1}) \in F\right\} \le -\bar{\Lambda}(F)$$

and

$$\liminf \frac{1}{n} \log \mathbb{P}_i\left\{X_n = j, \frac{1}{n} \sum_{k=0}^{n} f(X_k, X_{k+1}) \in G\right\} \ge -\bar{\Lambda}(G),$$

where $\bar{\Lambda}(B) = \inf\{\Lambda^*(v): v \in B\}$, $B \in \mathscr{R}^d$, with no restriction on f and none on $\{X_n\}$ (other than irreducibility).

(iii) If $\{\xi_i, i = 1, 2, \dots\}$ are i.i.d. $\mathbb{R}^d$-valued random variables ($d > 1$) then $\mathscr{D} = \{\alpha: Ee^{\langle \alpha, \xi_1 \rangle} < \infty\}$. It is not known even in this case whether the condition $0 \in \mathring{\mathscr{D}}$ is necessary for the upper bound when F is not compact. Thus there is no hope, at the present state of our knowledge, to remove this condition for the general upper bound.

The basic work on additive functionals of Markov processes was done by Donsker and Varadhan [3]. Lower bounds for additive functionals of chains on a locally compact space, under weaker recurrence conditions than in [3], were obtained by Chiang [1]. Upper bounds in a very general setting were obtained by de Acosta [2]. For other references see [5].

Our work differs from the above in several aspects of technique and results. The technique depends on the construction of a regeneration scheme, which simplifies much of the analysis. This in turn leads naturally to a rate function expressed in terms of the *convergence parameter* of the transform kernel of the process (see Section 2), as compared to the spectral radius, which is used in most prior work. The fact that this leads to different rates is illustrated in Section 6 of [5]. Our results, then, are limit theorems for

$$\mathbb{P}_x\left\{X_n \in A, \frac{S_n}{n} \in \Gamma\right\}, \tag{1.5}$$

for $A \subset \mathbb{E}$, contrasted with

$$\mathbb{P}_x\left\{\frac{S_n}{n} \in \Gamma\right\}, \tag{1.6}$$

namely the case when $A = \mathbb{E}$. In general the reciprocal of the convergence parameter may be smaller than the spectral radius and (1.5) may have a faster exponential decay than (1.6). Some conditions on when the rates are the same (and independent of A) are given in [5].

Most of this paper is devoted to the proof of the lower bound, which holds under essentially no hypothesis (only irreducibility). There is no recurrence hypothesis on $\{X_n\}$, nor any topological condition on its state space; also no restriction on the distributions of the additive components other than a mild nonsingularity condition. On the other hand, our additive functionals are finite dimensional compared to the measure valued functionals in [3]. Our proof of the upper bound follows more traditional lines, though the regeneration structure also comes in essentially there.

2. Preliminaries. We summarize a few facts from [5] and elsewhere.

The φ-irreducibility of $\{X_n\}$ implies that there exists a function $h\colon \mathbb{E} \to \mathbb{R}^1$, a measure $\nu(\cdot)$ on $(\mathbb{E}, \mathscr{E})$, and an integer $1 \le k_0 < \infty$ such that

$$h(x)\nu(A) \le P^{k_0}(x, A), \qquad x \in \mathbb{E},\ A \in \mathscr{E}. \tag{2.1}$$

$[P(\cdot,\cdot)$ is the transition function of $\{X_n\}$.]

What is needed for the present work is an extension of (2.1) to MA-processes. Namely, we need the existence of a measure $\nu(\cdot)$, and instead of the function $h(x)$, a family of measures $\{h(x,\cdot)\}$ on $(\mathbb{E}, \mathscr{E})$ such that

$$h(x, \Gamma)\nu(A) \le P^{k_0}(x, A \times \Gamma), \qquad x \in \mathbb{E},\ A \in \mathscr{E},\ \Gamma \in \mathscr{R}^d. \tag{2.2}$$

A mild nonsingularity condition on P is sufficient to guarantee the existence of h and ν satisfying (2.2) (see Niemi and Nummelin [7]). We have also observed in [6] that for purposes of the large deviation theorems under study here it is no loss of generality to take $k_0 = 1$. Writing $h(x, \Gamma)\nu(A) = h \otimes \nu(x, A \times \Gamma)$ we assume throughout this paper that

$$(\mathrm{M}_1) \qquad h \otimes \nu \leq P.$$

REMARK. As noted in [6], alternative minorizations of the form

$$(\mathrm{M}_2) \qquad h(x)\nu(A \times \Gamma) \leq P(x, A \times \Gamma),$$

where now $\nu(\cdot \times \cdot)$ is a measure on $(\mathbb{E} \times \mathscr{E}, \mathbb{E} \otimes \mathbb{R}^d)$, or more generally

$$(\mathrm{M}) \qquad h(x, \cdot) * \nu(A \times \cdot)(\Gamma) \leq P(x, A \times \Gamma),$$

can be used instead of (M_1). We will work with (M_1) for definiteness.

Under (M_1) there exists a sequence of regeneration times $\{T_0, T_1, \dots\}$ with the following properties (see [6]):

(2.3)
(i) $\{T_{i+1} - T_i;\ i = 0, 1, \dots\}$ are i.i.d. random variables;
(ii) the random blocks $\{X_{T_i}, \dots, X_{T_{i+1}-1}, \xi_{T_i+1}, \dots, \xi_{T_{i+1}}\}$, $i = 0, 1, \dots$, are independent, and
(iii) $P_x\{X_{T_i} \in A \mid \mathscr{F}_{T_i-1}, \xi_{T_i}\} = \nu(A)$, where $\mathscr{F}_n$ = the σ-field generated by $\{X_0, \dots, X_n, \xi_1, \dots, \xi_n\}$.

Let $\tau =_D T_i - T_{i-1}$, $S_\tau =_D S_{T_i} - S_{T_{i-1}}$, $i \geq 1$.

Define the generating function on $\mathbb{R}^{d+1}$

$$\psi(\alpha, \zeta) = E_\nu e^{\langle \alpha, S_\tau \rangle - \zeta\tau}, \qquad \alpha \in \mathbb{R}^d,\ \zeta \in \mathbb{R}^1, \tag{2.4}$$

and let $\mathscr{W} = \{(\alpha, \zeta) \in \mathbb{R}^{d+1}: \psi(\alpha, \zeta) < \infty\}$.

Define

$$\Lambda(\alpha) = \inf\{\zeta: \psi(\alpha, \zeta) \leq 1\}. \tag{2.5}$$

Consider the transform kernel

$$\hat{P}(\alpha) = \hat{P}(x, A; \alpha) = \int P(x, A \times ds) e^{\langle \alpha, s \rangle}, \qquad x \in E,\ A \in \mathscr{E}.$$

Then $R(\alpha) = e^{-\Lambda(\alpha)}$ is the *convergence parameter* of $\hat{P}(\alpha)$. This follows by applying Proposition 4.7(i) of [8] to the series form of ψ, namely to

$$\psi(\alpha, \zeta) = \sum e^{-\zeta n} \nu \left[\hat{P}(\alpha) - \hat{h} \otimes \nu(\alpha)\right]^{n-1} \hat{h}(\alpha).$$

If $\mathscr{W}$ is an open set and (M_1) holds, then by Theorem 4.1 of [5]

$$\psi(\alpha, \Lambda(\alpha)) = 1, \tag{2.6}$$

and $e^{\Lambda(\alpha)} \equiv \lambda(\alpha)$ is an eigenvalue of $\hat{P}(\alpha)$. In this case the associated eigenfunction and measure can be expressed by the representation formulas

$$r(x; \alpha) = E_x\left[e^{\langle \alpha, S_\tau \rangle - \Lambda(\alpha)\tau}\right], \tag{2.7}$$

$$l(A; \alpha) = E_\nu\left[\sum_{n=0}^{\tau-1} e^{\langle \alpha, S_n \rangle - \Lambda(\alpha) n};\ X_n \in A\right]. \tag{2.8}$$

Furthermore, if $\mathscr{W}$ is open, $\Lambda(\cdot)$ is analytic on $\mathscr{D} = \{\alpha: \Lambda(\alpha) < \infty\}$, and there is a set $\mathbb{F} \subset E$ with $\varphi(\mathbb{F}^c) = \varnothing$ such that for each $x \in \mathbb{F}$, $r(x, \cdot)$ is analytic on $\mathscr{D}$.

Note that under (M_1)

$$\hat{h}(x; \alpha)\nu(A) \leq \hat{P}(x, A; \alpha), \qquad x \in E, A \in \mathscr{E}, \alpha \in \mathbb{R}^d, \tag{2.9}$$

where

$$\hat{h}(x; \alpha) = \int_{\mathbb{R}^d} e^{\langle \alpha, s \rangle} h(x, ds).$$

We again abbreviate (2.9) by

$$(\hat{h} \otimes \nu)(\alpha) \leq \hat{P}(\alpha), \qquad \alpha \in \mathbb{R}^d. \tag{2.10}$$

3. Lower bound.

THEOREM 1. *Let $\{(X_n, S_n);\ n = 0, 1, \dots\}$ be a φ-irreducible MA-process satisfying* (M_1). *Then for any open set $G \subset \mathbb{R}^d$ and $A \in \mathscr{E}$ such that $\varphi(A) > 0$,*

$$l(x, A \times G) \geq -\overline{\Lambda}(G). \tag{3.1}$$

Idea of proof. The principal construction in the proof consists of a truncation of the sequence $\{\xi_i,\ i = 1, 2, \dots\}$ and of the inter-regeneration times $\{T_i - T_{i-1},\ i = 1, 2, \dots\}$, which is such that the resulting process is again an MA-process, but on an augmented state space. The associated generating function ψ [see (2.4)] will converge in all of $\mathbb{R}^{d+1}$, and hence the results of [5] will apply to it. The original MA-process is then approximated by a sequence of such truncated processes, leading to (3.1).

Before turning to this, however, we carry out a smoothing similar to that used by Varadhan [9] in the i.i.d. case, whose purpose is to extend the support of the $\{\xi_i\text{'s}\}$ to all of $\mathbb{R}^d$, thereby eliminating some technicalities due to special boundaries. We then prove that if (3.1) holds for the smoothed process, then it holds in general; namely, it is thus no loss of generality to assume that $\operatorname{Supp}\{P(x, A \times \cdot)\} = \mathbb{R}^d$. [$\operatorname{Supp} \nu(\cdot)$ denotes the closed convex hull of the support of the measure ν.]

Smoothing. Let $\{\eta_i^{(\varepsilon)},\ i = 1, 2, \dots\}$ be i.i.d. $\mathbb{R}^d$-valued random variables, independent of the MA-process $\{(X_n, S_n)\}$, and normally distributed with mean 0 and covariance matrix εI, where I is the identity matrix. Let $\xi_i^{(\varepsilon)} = \xi_i + \eta_i^{(\varepsilon)}$, $i = 1, 2, \dots$, and denote by $P^{(\varepsilon)}(\alpha)$, $\Lambda^{(\varepsilon)}(\alpha)$, etc., all the objects defined before, but with ξ_i replaced by $\xi_i^{(\varepsilon)}$. Then

$$\hat{P}^{(\varepsilon)}(x, A; \alpha) = e^{\varepsilon\|\alpha\|^2/2} \hat{P}(x, A; \alpha), \tag{3.2}$$

and

$$\Lambda^{(\varepsilon)}(\alpha) = \Lambda(\alpha) + \frac{\varepsilon}{2}\|\alpha\|^2. \tag{3.3}$$

Also $\hat{P}^{(\varepsilon)}$ satisfies a minorization by $e^{\varepsilon\|\alpha\|^2/2}\hat{h} \otimes \nu$, hence there is a regeneration structure and generating function $\psi^{(\varepsilon)}(\alpha, \zeta)$ with domain $\mathscr{W}^{(\varepsilon)}$. We will refer to the above process as the ε-smoothed process.

LEMMA 3.1. *If the LDP lower bound* (3.1) *holds for ε-smoothed processes for all $\varepsilon > 0$, then it holds also with $\varepsilon = 0$.*

PROOF. Take any point $v \in G$ and let $B_\delta(v) \subset G$ be a δ-ball with center v. Let $Z_n^{(\varepsilon)} = \eta_1^{(\varepsilon)} + \cdots + \eta_n^{(\varepsilon)}$. Then $S_n^{(\varepsilon)} = S_n + Z_n^{(\varepsilon)}$, and hence

$$\begin{aligned} &P\left\{X_n \in A, \frac{S_n}{n} \in B_\delta(v)\right\} \\ &\geq P\left\{X_n \in A, \frac{S_n^{(\varepsilon)}}{n} \in B_{\delta/2}(v), \frac{Z_n^{(\varepsilon)}}{n} \in B_{\delta/2}(0)\right\} \\ &\geq P\left\{X_n \in A, \frac{S_n^{(\varepsilon)}}{n} \in B_{\delta/2}(v)\right\} - P\left\{\frac{Z_n^{(\varepsilon)}}{n} \notin B_{\delta/2}(0)\right\}. \end{aligned} \tag{3.4}$$

By the supposition that (3.1) is true for $\varepsilon > 0$,

$$\liminf P\left\{X_n \in A, \frac{S_n^{(\varepsilon)}}{n} \in B_{\delta/2}(v)\right\} \geq -\overline{\Lambda^{(\varepsilon)}}\left(B_{\delta/2}(v)\right) \geq -\Lambda^{(\varepsilon)*}(v),$$

and also

$$P\left\{\frac{Z_n^{(\varepsilon)}}{n} \notin B_{\delta/2}(0)\right\} \leq \exp\left\{-\frac{n\delta^2}{8\varepsilon} + o(n)\right\}.$$

Therefore

$$\begin{aligned} &\liminf \frac{1}{n} \log P\left\{X_n \in A, \frac{S_n}{n} \in B_\delta(v)\right\} \\ &\geq -\Lambda^{(\varepsilon)*}(v), \quad \text{provided } \Lambda^{\varepsilon*}(v) < \frac{\delta^2}{8\varepsilon}. \end{aligned} \tag{3.5}$$

But

$$\begin{aligned} \Lambda^{(\varepsilon)*}(v) &= \sup_\alpha \left[\langle \alpha, v\rangle - \Lambda(\alpha) - \frac{\varepsilon}{2}\|\alpha\|^2\right] \\ &\to \Lambda^*(v) = \sup_\alpha \left[\langle \alpha, v\rangle - \Lambda(\alpha)\right] \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{3.6}$$

Hence for sufficiently small ε, by (3.5)

$$\liminf \frac{1}{n} \log P\left\{X_n \in A, \frac{S_n}{n} \in B_\delta(v)\right\} \geq -\Lambda^{(\varepsilon)*}(v),$$

and letting $\varepsilon \to 0$ we have by (3.6)

$$\liminf \frac{1}{n} \log P\left\{X_n \in A, \frac{S_n}{n} \in B_\delta(v)\right\} \geq -\Lambda^*(v),$$

for all $v \in G$.

Hence we conclude that

$$\begin{aligned}\liminf \frac{1}{n} \log \mathbb{P}\left\{X_n \in A, \frac{S_n}{n} \in G\right\} &\geq \liminf \frac{1}{n} \log P\left\{X_n \in A; \frac{S_n}{n} \in B_\delta(v)\right\} \\ &\geq -\Lambda^*(v), \quad \text{for all } v \in G,\end{aligned}$$

and this is (3.1). □

Thus from now on it is no loss of generality to assume that the support of the r.v.'s $\{\xi_i\}$ is all of $\mathbb{R}^d$, and we do so without explicitly mentioning this point again. Similarly, since the smoothing components $\{\eta_i^{(\varepsilon)}\}$ are independent of the original MA-process, and of the stopping time τ (which is unchanged by the smoothing), we see that it is no loss of generality to assume that

(3.7) $$\mathscr{S} \equiv \operatorname{Supp}\{P_\nu(S_\tau \in \cdot)\} = \mathbb{R}^d.$$

LEMMA 3.2. *Under the hypothesis of Theorem* 1:

(i) *There exists a* $\delta > 0$ *such that*

(3.8) $$\Lambda(\alpha) \geq \delta\|\alpha\|,$$

for $\|\alpha\|$ *sufficiently large.*

(ii) *The level sets of* Λ *are compact.*

(iii) $\Lambda^*(v) = \langle \alpha_v, v\rangle - \Lambda(\alpha_v)$ *for some* $\alpha_v \in \mathbb{R}^d$.

PROOF. (i) Recall that we are working under the condition $\mathscr{S} = \mathbb{R}^d$, in particular $0 \in \mathscr{S}^0$. Let $U = \{u \in \mathbb{R}^d\colon \|u\| = 1\}$. Now if any r.v. $W \in \mathbb{R}^d$ has 0 in the interior of the convex hull of its support, then there exist $\varepsilon > 0$, $\delta > 0$ such that

(3.9) $$P\{\langle u, W\rangle \geq \delta\} \geq \varepsilon, \quad \text{for all } u \in U.$$

Applying this to $(S_\tau/\tau) = W$ we get for real $t > 0$,

$$\begin{aligned}\psi\left(tu, \frac{\delta}{2}t\right) &= E_\nu \exp\left(t\tau\left[\frac{\langle u, S_\tau\rangle}{\tau}\right] - \frac{\delta}{2}t\tau\right) \\ &\geq E_\nu\left\{\exp\left(t\tau\left[\frac{\langle u, S_\tau\rangle}{\tau} - \frac{\delta}{2}\right]\right); \frac{\langle u, S_\tau\rangle}{\tau} - \frac{\delta}{2} \geq \frac{\delta}{2}\right\},\end{aligned}$$

and since $\tau \geq 1$,

$$\geq e^{t(\delta/2)} P\left\{\frac{\langle u, S_\tau\rangle}{\tau} \geq \delta\right\} \geq \varepsilon e^{t(\delta/2)}.$$

Thus for $t \geq t_0$, say, and some $\delta > 0$, $c > 1$,

$$\psi(tu, \delta t) \geq c, \quad \text{for all } u \in \mathscr{U}.$$

Hence (recalling that always $\psi(\alpha, \Lambda(\alpha)) \leq 1$),

$$\Lambda(tu) > \delta t;$$

namely,

$$\Lambda(\alpha) = \Lambda\left(\frac{\alpha}{\|\alpha\|}\|\alpha\|\right) \geq \delta\|\alpha\|, \quad \text{for } \|\alpha\| \geq t_0. \tag{3.10}$$

This proves (i).

(ii) Let

$$L_c = \{\alpha: \Lambda(\alpha) \leq c\}, \qquad c < \infty.$$

Note that all L_c's are closed since Λ is l.s.c. By (3.10) they are bounded.

(iii) Clearly,

$$\inf\{\Lambda(\alpha): \alpha \in \mathbb{R}^d\} = \inf\{\Lambda(\alpha): \alpha \in L_c\}.$$

Since Λ is l.s.c. it achieves its infimum over compact sets. Hence there exists an α_0 such that

$$\Lambda(\alpha_0) = \inf \Lambda(\alpha),$$

and this

$$= -\sup_{\alpha}\{\langle\alpha, 0\rangle - \Lambda(\alpha)\} = -\Lambda^*(0).$$

Finally, we take arbitrary $v \in \mathscr{S}^0$. We translate to $\xi_i^v = \xi_i - v$, $S_n^v - nv$, etc. Then

$$\psi^v(\alpha, \zeta) = E_\nu e^{\langle\alpha, S_\tau\rangle - \langle\alpha, v\rangle\tau - \zeta\tau} = \psi(\alpha, \zeta + \langle\alpha, v\rangle)$$

and

$$\begin{aligned}\Lambda^v(\alpha) &= \inf\{\zeta: \psi^v(\alpha, \zeta) \leq 1\}\\ &= \inf\{\zeta: \psi(\alpha, \zeta) \leq 1\} - \langle\alpha, v\rangle\\ &= \Lambda(\alpha) - \langle\alpha, v\rangle.\end{aligned}$$

Now $0 \in \mathscr{S}^{0v} = \text{Supp}_\nu(S_\tau^v/\tau)$, and the argument for $v = 0$ above applies. □

Truncation/killing. To carry out this construction in a consistent manner we first enlarge the state space of the underlying Markov chain by adjoining the r.v.'s

$$V_n = n - T_{N_n} = \text{time since last regeneration}$$

($N_n = \max\{k: T_k \leq n\}$). Let $\{\tilde{X}_n\} = \{(X_n, V_n)\}$. This is a Markov chain on $\mathbb{E} \times \{0, 1, \dots\} \equiv \tilde{\mathbb{E}}$. The associated MA-process has transition kernel $\tilde{P}((x \times i), (dy \times j) \times ds) \geq \delta_0(i)\delta_1(j)h(x, ds)\nu(dy)$, the right side being a minorization for $\tilde{P}$. Write $\tilde{h}(x \times i, ds) \equiv \delta_0(i)h(x, ds)$, $\tilde{\nu}(dy \times j) = \nu(dy)\delta_1(j)$, $x \times i = \tilde{x}$, etc. The regeneration times are the set $\{n: V_n = 1\}$, which are, of course, exactly the same as for $\{(X_n, S_n)\}$. ($V_n = 0$ corresponds to a "head" in the coin tossing scheme of [6].) Also we identify $\tilde{S}_n = S_n$. Thus

$$\tilde{\psi}(\alpha, \zeta) \equiv E_{\tilde{\nu}} e^{\langle\alpha, \tilde{S}_\tau\rangle - \zeta\tilde{\tau}} = \mathbb{E}_\nu e^{\langle\alpha, S_\tau\rangle - \zeta\tau} = \psi(\alpha, \zeta).$$

Now given any MA-process $\{(X_n, S_n)\}$ on $\mathbb{E} \times \mathbb{R}^d$, and any subsets $\mathbb{E}' \in \mathscr{E}$ and $\mathbb{R}'^d \in \mathscr{R}^d$, define an associated process $\{(X_n', S_n')\}$ with the truncated kernel

$$P'(x, A \times \Gamma) = P(x, (A \cap E') \times (\Gamma \cap \mathbb{R}'^d)),$$

where $\mathscr{E}' \in \mathscr{E}$ and $\mathbb{R}'^d \in \mathscr{R}^d$. This is a substochastic kernel. We call $\{(X_n', S_n')\}$ the killed process. Letting $N = \inf\{n: (X_n', \xi_n') \notin \mathbb{E}' \times \mathbb{R}'^d\}$, one can define $(X_n, S_n) = \omega_0$ for $n \geq N$, where $\omega_0 \notin \mathbb{E} \times \mathbb{R}^d$ is an adjoined "graveyard" state. Clearly, the associated generating function ψ' satisfies $\psi'(\alpha, \zeta) \leq \psi(\alpha, \zeta)$.

We now define such a killing of the process $\{(X_n, V_n), S_n\}$ on $\tilde{\mathscr{E}} \times \mathbb{R}^d$, by taking $\tilde{\mathbb{E}}' \equiv \mathbb{E}^M = \mathbb{E} \times \{0, 1, \dots, M\}$, and $\mathbb{R}'^d = [-M, M]^d$ for some $M < \infty$. Denote the killed process by $\{(X_n, V_n^M), S_n^M\} = \{X_n^M, S_n^M\}$, where $X_n^M = (X_n, V_n^M)$ and $S_n^M = S_0^M + \xi_1^M + \cdots + \xi_n^M$. Thus the additive components ξ_i^M, and the inter-regeneration times $T_{i+1}^M - T_i^M =_D \tau^M$ for this process *are bounded by M*. We denote all the usual functions for the M-truncated MA-process by sub- or superscripts M. The minorization for the truncated kernel is

$$\delta_0(i)\delta_1(j)h(x, ds \cap [-M, M]^d)\nu(dy).$$

Let

$$h_M(x^M, ds) = h_M(x \times i, ds) = \delta_0(i)h(x, ds \cap [-M, M]^d)$$

and

$$\nu_M(dy^M) = \nu_M(j \times dy) = \delta_1(j)\nu(dy).$$

Thus in this case the truncated and untruncated kernels have the same regeneration measure. Observe that

$$\begin{aligned}\psi^M(\alpha, \zeta) &= E_{\nu_M} e^{\langle \alpha, S_{\tau^M}^M \rangle - \zeta\tau^M} \\ &= \int_{\mathbb{R}^d} \sum_{n=0}^{\infty} e^{\langle \alpha, s \rangle - \zeta n} \mathbb{P}_\nu^M\{S_n \in ds, \tau = n\} \qquad (3.11) \\ &= \int_{[-M^2, M^2]^d} \sum_{n=0}^{M} e^{\langle \alpha, s \rangle - \zeta n} \mathbb{P}_{\nu_M}\{S_n \in ds, |\xi_i| \leq M, i = 1, 2, \dots, \tau = n\}.\end{aligned}$$

(|Each component of $S_n| \leq M \cdot \max \xi_i^M = M^2$). Then

$$\mathscr{W}^M \equiv \{(\alpha, \zeta): \psi^M(\alpha, \zeta) < \infty\} = \mathbb{R}^{d+1}.$$

Hence by (2.6) there exists $\{\Lambda^M(\alpha): \alpha \in \mathbb{R}^d\}$ such that

$$\psi^M(\alpha, \Lambda^M(\alpha)) = 1, \qquad \alpha \in \mathbb{R}^d. \qquad (3.12)$$

REMARK. (2.6) rests on Theorem 4.1 of [5] which was proved for stochastic MA-kernels. However, the entire argument leading to the theorem goes through identically for substochastic kernels. Stochasticity plays no essential role whatever.

Also

$$\Lambda^M(\alpha) \text{ is analytic on } \mathbb{R}^d, \qquad (3.13)$$

and

$$1 < \psi^M(\alpha, \zeta) < \infty, \quad \text{for some } \zeta < \Lambda^M(\alpha). \qquad (3.14)$$

By (3.11) and the monotone convergence theorem

$$\psi^M(\alpha, \zeta) \uparrow \psi(\alpha, \zeta), \quad \text{as } M \uparrow \infty, \tag{3.15}$$

for all $(\alpha, \zeta) \in \mathbb{R}^{d+1}$.

LEMMA 3.3. *Under the hypothesis of Theorem* 1

(i) $\Lambda^M(\alpha) \uparrow \Lambda(\alpha)$ *as* $M \uparrow \infty$, $\alpha \in \mathbb{R}^d$;
(ii) $\Lambda(\alpha)$ *is lower semi-continuous and strictly convex.*

PROOF. (i) Take $M_1 < M_2$. Then by (3.12) and (3.15)

$$\psi^{M_1}(\alpha, \Lambda^{M_2}(\alpha)) \leq \psi^{M_2}(\alpha, \Lambda^{M_2}(\alpha)) = 1.$$

Since $\psi^M(\alpha, \zeta)$ is decreasing in ζ and $\psi^M(\alpha, \Lambda^{M_1}(\alpha)) = 1$. It follows that $\Lambda^{M_1}(\alpha) \leq \Lambda^{M_2}(\alpha)$. Namely,

$$\Lambda^M(\alpha) \text{ is increasing in } M \text{ for all } \alpha \in \mathbb{R}^d. \tag{3.16}$$

Furthermore, by (3.15)

$$\psi^M(\alpha, \Lambda(\alpha)) \leq \psi(\alpha, \Lambda(\alpha)) \leq 1.$$

But since $\psi^M(\alpha, \Lambda^M(\alpha)) = 1$ we must have

$$\Lambda^M(\alpha) \leq \Lambda(\alpha).$$

Thus

$$\lim_{M \to \infty} \Lambda^M(\alpha) \text{ exists.}$$

We claim that

$$\Lambda^M(\alpha) \uparrow \Lambda(\alpha). \tag{3.17}$$

First take $\alpha \in \mathscr{D}$, i.e., $\Lambda(\alpha) < \infty$. Take any $\varepsilon > 0$. By strict monotonicity of ψ in ζ

$$1 < \psi(\alpha, \Lambda(\alpha) - \varepsilon) \leq \infty.$$

Since $\psi^M(\alpha, \zeta) \uparrow \psi(\alpha, \zeta)$, $(\alpha, \zeta) \in \mathbb{R}^{d+1}$, there exists an M_0 such that

$$1 < \psi^{M_0}(\alpha, \Lambda(\alpha) - \varepsilon) \qquad (< \infty \text{ since } \mathscr{W}^{M_0} = \mathbb{R}^{d+1}).$$

Since again $\psi^{M_0}(\alpha, \Lambda^{M_0}(\alpha)) = 1$, this implies that

$$\Lambda(\alpha) - \varepsilon < \Lambda^{M_0}(\alpha).$$

Since ε is arbitrary, this implies (3.17). Next suppose $\alpha \notin \mathscr{D}$. Then $\Lambda(\alpha) = \infty$. We show that in this case

$$\Lambda^M(\alpha) \uparrow \infty. \tag{3.18}$$

Now for $\alpha \notin \mathscr{D}$ we claim

$$\psi(\alpha, \zeta) = \infty, \quad \text{for all } \zeta \in \mathbb{R}. \tag{3.19}$$

For suppose $\psi(\alpha, \zeta') < \infty$ for some ζ'. Then $\psi(\alpha, \zeta) < \infty$ for all $\zeta > \zeta'$; in fact,

$$\psi(\alpha, \zeta) \to 0, \quad \text{as } \zeta \to \infty.$$

Hence

$$\psi(\alpha, \zeta'') < 1, \quad \text{for some } \zeta''.$$

Therefore

$$\inf\{\zeta \colon \psi(\alpha, \zeta) \le 1\} = \Lambda(\alpha)$$

exists and is $< \infty$. This contradicts $\alpha \notin \mathscr{D}$ and proves (3.19). Now fix any ζ_0. Since $\psi(\alpha, \zeta_0) = \infty$ and $\psi^M(\alpha, \zeta_0) \uparrow \psi(\alpha, \zeta_0)$ we must have

$$\psi^{(M_0)}(\alpha, \zeta_0) > 1, \quad \text{for some } M_0.$$

Hence

$$\Lambda^{(M_0)}(\alpha) > \zeta_0.$$

This proves (3.18) and hence (i).

(ii) Since $\Lambda^{(M)}(\alpha)$ are all continuous and $\Lambda^{(M)}(\alpha) \uparrow \Lambda(\alpha)$, we have that $\Lambda(\alpha)$ is l.s.c.

[Alternatively note $\psi(\cdot, c)$ is l.s.c. since it $= \lim \uparrow \psi^M(\cdot, c)$, which are continuous. Hence $\{\alpha\colon \psi(\alpha, c) \le 1\}$ is closed. But this $= \{\alpha\colon \Lambda(\alpha) \le c\}$. Since the latter is thus closed for all c, $\Lambda(\cdot)$ is l.s.c.].

Due to the smoothing, the support of S_τ is all of $\mathbb{R}^d$. Hence $\psi(\alpha, \zeta)$ is strictly convex, and hence so is $\{\Lambda(\alpha);\ \alpha \in \mathscr{D}\}$.

This proves Lemma 3.3. □

LEMMA 3.4. *Under the hypothesis of Theorem* 1 $\Lambda^{M*}(v) \downarrow \Lambda^*(v)$ *as* $M \uparrow \infty$, *for all* $v \in \mathbb{R}^d$.

PROOF. First we prove the case $v = 0$. Note that $\Lambda^*(0) = -\inf\{\Lambda(\alpha)\colon \alpha \in \mathbb{R}^d\}$.

1. Since Λ^M are continuous and $\Lambda^M \uparrow \Lambda$ as $M \uparrow \infty$, we have

$$\inf_K \Lambda^M \uparrow \inf_K \Lambda, \quad \text{as } M \uparrow \infty, \tag{3.20}$$

for all compact $K \subset \mathbb{R}^d$.

2. Let $L_c = \{\Lambda \le c\}$ and $L_c^M = \{\Lambda^M \le c\}$. For $M \ge$ some M_0, $0 \in (\mathscr{S}^M)^0$, since $\mathscr{S}^M \uparrow \mathscr{S} = \mathbb{R}^d$ [see (3.7)]. Hence by Lemma 3.2 L_c^M is compact for $M \ge M_0$. Furthermore,

$$K \overset{\text{def.}}{=} L_c^{M_0} \supset L_c^M \supset L_c. \tag{3.21}$$

Now take any $\bar{\alpha}$ such that $\Lambda(\bar{\alpha}) < \infty$, and take $c = \Lambda(\bar{\alpha})$. Then on the complement of K

$$c < \Lambda^M \le \Lambda; \quad \text{for } M \ge M_0. \tag{3.22}$$

Since $\Lambda^M \uparrow \Lambda$ and $\bar{\alpha} \in L_c \subset K$

$$\inf_K \Lambda^M \le \inf_K \Lambda \le c, \quad \text{for } M \ge M_0. \tag{3.23}$$

It follows that for $M \geq M_0$

$$\inf_K \Lambda^M = \inf_{\mathbb{R}^d} \Lambda^M \quad \text{and} \quad \inf_K \Lambda = \inf_{\mathbb{R}^d} \Lambda. \tag{3.24}$$

This, together with 1, implies that

$$\inf_{\mathbb{R}^d} \Lambda^M \uparrow \inf_{\mathbb{R}^d} \Lambda, \quad \text{as } M \uparrow \infty, \tag{3.25}$$

proving the assertion for $v = 0$.

3. For general $v \neq 0$ translate to $\xi_i^v = \xi_i - v$. Then

$$\Lambda^v(\alpha) = \Lambda(\alpha) - \langle \alpha, v \rangle, \qquad \Lambda^*(v) = \Lambda^{v*}(0), \tag{3.26}$$

and similarly for Λ^M and Λ^{M*}. Then apply the above result for $v = 0$. □

Let $B_\delta(v)$ = the δ-ball with center v. The next lemma is an LDP lower bound for the M-truncated process. Recall that the underlying Markov chain for this process is $\{(X_n, V_n^M)\} = \{X_n^M\}$, with state space $\mathbb{E} \times \{0, 1, \dots, M\} = \mathbb{E}^M$. We denote the elements of $\mathbb{E}^M$ by x^M.

LEMMA 3.5. *Under the hypothesis of Theorem* 1

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}_x \left\{ \frac{S_n^M}{n} \in B_\delta(v),\ X_n^M \in A^M \right\} \geq -\Lambda^{M*}(v), \tag{3.27}$$

where $x^M \in \mathbb{E}^M$, $A^M = A \times \{0, 1, \dots, M\}$, $A \in \mathscr{E}$, *provided that* M *is sufficiently large.*

PROOF. Since $\mathscr{W}^M = \mathbb{R}^{d+1}$, Theorem 4.1 of [5] applies. [See also the Remark following (3.12).] $\hat{P}^M(\alpha)$ has eigenvalue $\exp\{\Lambda^M(\alpha)\}$ with right eigenfunction given by

$$r_M(x^M, \alpha) = E_{x^M} \exp\left\{ \langle \alpha, S_{\tau_M}^M \rangle - \Lambda^M(\alpha) \tau_M \right\}, \qquad x^M \in E^M. \tag{3.28}$$

Letting $Q_M(x^M, dy \times ds; \alpha)$ be the α-conjugate MA-transition kernel of $\hat{P}_M(\alpha)$ as in Section 5 of [5] [$Q_M(\alpha = 0) = P_M$] we get the formula

$$\begin{aligned} &P_M^n\left(x^M, \tilde{A} \times nB_\delta(v)\right) \\ &\quad = e^{\Lambda^M(\alpha)n} r_M(x^M; \alpha) \int\int \frac{e^{-\langle \alpha, s \rangle}}{r_M(y; \alpha)} Q_M^n\left(x^M, dy \times ds; \alpha\right), \end{aligned} \tag{3.29}$$

$$x^M \in \mathbb{E}^M,\ \tilde{A} \in \mathscr{E}^M.$$

Now due to the ε-smoothing, the support of $\{\xi_1\}$ is all of $\mathbb{R}^d$, and hence for M sufficiently large

$$v \in \nabla \Lambda^M(\mathscr{D})$$

($\mathscr{D} = \mathbb{R}^d$ also in this case). Hence, letting $\alpha_v = (\nabla\Lambda^M)^{-1}(v)$, we get from (3.29)

$$P_M^n(x^M, \tilde{A} \times nB_\delta(v)) = e^{-\Lambda^{M*}(v)n} r_M(x^M; \alpha_v) \iint_{\tilde{A}\times nB_\delta(v)} \frac{e^{-\langle \alpha_v, s-nv\rangle}}{r_M(y; \alpha_v)} Q_M^n(x^M; dy \times ds; \alpha_v). \tag{3.30}$$

Now $x^M = x \times i$ for some $x \in E$, $i \in \{0, 1, \dots, M\}$, but due to the basic property of an MA-process we can replace x^M by x in Q_M^n in (3.30). Also take $\tilde{A} = A \times [0, 1, \dots, M] = A^M$ for some $A \in \mathscr{E}$ with $\varphi(A) > 0$. Then by the law of large numbers (Section 5 of [5])

$$\liminf_{n\to\infty} Q_M^n(x, A^M \times nB_\delta(v); \alpha_v) > 0, \tag{3.31}$$

for M sufficiently large.

Also by (3.28) and the boundedness of $S_{\tau_M}^M$ and τ_M, $0 < c' \le r(x; \alpha) \le c'' < \infty$ for some c', c''. Hence for all $\delta' < \delta$,

$$\begin{aligned}(3.28) &\ge \liminf_{n\to\infty} \frac{1}{n} \log P_M^n\{x, A^M \times nB_{\delta'}(v)\} \\ &\ge -\Lambda^{M*}(v) - \delta'\|\alpha_v\|,\end{aligned} \tag{3.32}$$

and since δ' is arbitrary this implies the lemma. □

Proof of Theorem 1. Take $v \in G$. Then

$$\begin{aligned}&\liminf \frac{1}{n} \log \mathbb{P}_x\left\{X_n \in A, \frac{S_n}{n} \in G\right\} \\ &\ge \liminf \frac{1}{n} \log \mathbb{P}_x\left\{X_n \in A, \frac{S_n}{n} \in B_\delta(v)\right\},\end{aligned} \tag{3.33}$$

for some $\delta > 0$ (since G is open). Also since $A^M \subset A \times \{0, 1, \dots\}$, and since $\mathbb{P}_x^M$ is a truncation of the extension of $\mathbb{P}_x$, the last expression is

$$\begin{aligned}&\ge \liminf \frac{1}{n} \log \mathbb{P}_x^M\left\{X_n \in A^M, \frac{S_n}{n} \in B_\delta(v)\right\} \\ &\ge -\Lambda^{M*}(v), \quad \text{for sufficiently large } M.\end{aligned}$$

Now let $M \uparrow \infty$, and conclude by Lemma 3.4 that $\underline{l}(x, A \times G) \ge -\Lambda^*(v)$ for all $v \in G$. This implies (3.1) and the theorem. □

4. Upper bound. The key to the LDP upper bound is the inequality

$$\limsup \frac{1}{n} \log \hat{P}^n(x, A; \alpha) \le \Lambda(\alpha), \qquad \alpha \in \mathscr{D}. \tag{4.1}$$

If (4.1) holds then by Theorem II.2(a) of Ellis [4] (and its proof) or by Theorem

2.1 of de Acosta [2]

$$\limsup \frac{1}{n} \log P^n(x, A \times nF) \le -\bar{\Lambda}(F), \tag{4.2}$$

for compact $F \subset \mathbb{R}^d$. If also $0 \in \mathring{\mathscr{D}}$ then (4.2) holds for closed F.

In general, (4.1) will not hold for arbitrary $A \in \mathscr{E}$. In fact, Example 6.3 of [5] shows that with $A = \mathbb{E}$ and $S_n \in \mathbb{R}^1$, $\mathbb{P}_x(S_n \ge an)$ can have essentially any asymptotic behavior. What is needed is that A should be "not too large" in a sense to be made precise below. We call such sets s-sets (for "sufficiently small").

Of course, the most important candidate for a non-s-set is $\mathbb{E}$ itself. If $\mathbb{E}$ is an s-set (see below) then so are all measurable sets.

Now for any $0 < \rho < 1$, there always exists a subinvariant function $r_\rho(x; \alpha)$ for $\rho e^{-\Lambda(\alpha)}$; namely,

$$r_\rho(x; \alpha) = \sum_{n=0}^{\infty} \rho^n e^{-\Lambda(\alpha)n} \hat{P}^n(\alpha) \hat{h}(\alpha)(x) \tag{4.3}$$

satisfies

$$\hat{P}(\alpha) r_\rho(\alpha) \le \rho^{-1} e^{\Lambda(\alpha)} r_\rho(\alpha), \tag{4.4}$$

and

$$r_\rho(x; \alpha) < \infty, \quad \text{for } x \notin \mathscr{N}_\alpha,\ \alpha \in \mathscr{D}, \tag{4.5}$$

with $\varphi(\mathscr{N}_\alpha) = 0$. Let

$$\mathbb{E}^N = \left\{ y\colon \sum_{n=0}^{N} (P^{*n} * h)\big(y, [-N, N]^d\big) \ge \frac{1}{N} \right\}. \tag{4.6}$$

LEMMA 4.1. *Assume* (M_1). *Then*

$$\mathbb{E}^N \uparrow \mathbb{E}, \tag{4.7}$$

and

$$r_\rho(x, \alpha) \ge \delta > 0, \quad \textit{for } x \in E^N \ (\delta = \delta(\rho, \alpha, N)). \tag{4.8}$$

DEFINITION 4.1. The s-sets are all measurable sets which are subsets of $\mathbb{E}^N$ for some $N = 1, 2, \dots$.

PROOF OF LEMMA 4.1. That $\mathbb{E}^N \uparrow \mathbb{E}$ follows from irreducibility. To prove (4.8) write

$$\begin{aligned} \hat{P}^n(\alpha)\hat{h}(\alpha)(y) &= \int_E \hat{P}^n(y, dz, \alpha) \hat{h}(z; \alpha) \\ &\ge \int_E \int_{[-N, N]^d} e^{\langle \alpha, s \rangle} P^{*n}(y, dz \times \cdot) * h(z, \cdot)(ds) \\ &\ge e^{-N\|\alpha\|} (P^{*n} * h)\big(y, [-N, N]^d\big). \end{aligned} \tag{4.9}$$

But by (4.3)

$$r_\rho(y;\alpha) \geq \sum_{n=0}^{N} \rho^n e^{-\Lambda(\alpha)n}\hat{P}^n(\alpha)\hat{h}(\alpha)(y)$$

and by (4.9) this is

$$\geq (e^{-\Lambda(\alpha)N} \wedge 1)\rho^N e^{-N\|\alpha\|}\sum_{n=0}^{N}(P^{*n} * h)\big(y, [-N, N]^d\big).$$

Thus for $y \in E^N$

$$r_\rho(y;\alpha) \geq (e^{-\Lambda(\alpha)N} \wedge 1)\frac{\rho^N}{N}e^{-N\|\alpha\|}, \tag{4.10}$$

and we take the right side to be δ in (4.8). □

EXAMPLE. Suppose that for some $A \in \mathscr{E}$, there exists a probability measure $g(\cdot)$ on $\mathbb{R}^d$ and an $\varepsilon > 0$ such that the minorizing measures $h(\cdot,\cdot)$ satisfy

$$\varepsilon g(ds) \leq h(x, ds), \quad \text{for all } x \in A. \tag{4.11}$$

Then A is an s-set.

PROOF. (4.11) implies

$$P^{*n}(x, A \times \Gamma) \geq \varepsilon^n \nu^n(A) g^{*n}(\Gamma).$$

But also

$$(P^{*n} * h)\big(x, [-2N, 2N]^d\big) \geq \varepsilon P^{*n}\big(x, A \times [-N, N]^d\big) g\big([-N, N]^d\big),$$

and hence $A \subset E^N$ for sufficiently large N. □

Similarly if $P(x, dy \times ds) \geq \varepsilon\nu(dy \times ds)$ for $x \in A$ [with type (M_2) minorization], then A is an s-set. Such types of examples can be considerably extended. In particular, note that all finite sets are s-sets.

LEMMA 4.2. *If A is an s-set, then for each $\alpha \in \mathscr{D}$ there is a set $\mathscr{N}_\alpha$ of φ-measure* 0 *such that for $x \notin \mathscr{N}_\alpha$,*

$$\limsup \frac{1}{N}\log \hat{P}^n(x, A; \alpha) \leq \Lambda(\alpha). \tag{4.12}$$

PROOF. Take $0 < \rho < 1$. Then by (4.4)

$$1 \geq \int_A \frac{\hat{P}^n(x, dy;\alpha) r_\rho(y;\alpha)}{\rho^{-n}e^{n\Lambda(\alpha)}r_\rho(x;\alpha)}$$

$$\geq \rho^n e^{-n\Lambda(\alpha)} r_\rho^{-1}(x;\alpha)\inf\{r_\rho(y;\alpha);\, y \in A\}\hat{P}^n(x, A;\alpha).$$

But $A \in E^N$ for some $1 \leq N < \infty$, hence the above

$$\geq \rho^n e^{-n\Lambda(\alpha)}\delta(\rho, \alpha, N) r_\rho^{-1}(x;\alpha)\hat{P}^n(x, A;\alpha).$$

But $r_\rho(x;\alpha) < \infty$ for $x \notin$ some $\mathscr{N}_\alpha$, and hence

$$\hat{P}^n(x, A; \alpha) \le \rho^{-n} e^{n\Lambda(\alpha)} \delta^{-1} r_\rho(x; \alpha) < \infty,$$

implying

$$\limsup \frac{1}{n} \log \hat{P}^n \le \Lambda(\alpha) - \log \rho.$$

Letting $\rho \uparrow 1$ implies (4.12). □

THEOREM 2. *Let* $\{(X_n, S_n);\ n = 0, 1, \dots\}$ *be a* φ*-irreducible MA-process satisfying* (M_1).

(i) *Let* $K \subset \mathbb{R}^d$ *be compact, and* $A \in \mathscr{E}$ *be an s-set. Then for* φ *a.e.* x

$$\bar{l}(x, A \times K) \le -\bar{\Lambda}(K). \tag{4.13}$$

(ii) *If also* $0 \in \mathring{\mathscr{D}}(\Lambda)$, *then for closed* F *and* φ *a.e.* x

$$\bar{l}(x, A \times F) \le -\bar{\Lambda}(F). \tag{4.14}$$

(iii) *If* $d = 1$, *then* (4.14) *holds without the extra condition* $0 \in \mathring{\mathscr{D}}(\Lambda)$.

PROOF. The conclusion follows by applying Theorem II.2(a) of Ellis [4] or Theorem 2.1 of de Acosta [2] to (4.12) of Lemma 2.4. If for given x, (4.12) held for all $\alpha \in \mathscr{D}$ there would be nothing further to say. There is a small technical problem, however, due to the exceptional sets $\mathscr{N}_\alpha$. However, the proof rests on an argument of covering K by a *finite* number of half spaces, each determined by an α. Thus only a finite number $\mathscr{N}_\alpha$'s must be dealt with. Of course, the exceptional x-sets in (4.13) and (4.14) depend on K and F.

[An alternative approach would be to argue as in Lemma 4.4 of [5] that there exists a single set $\mathbb{F}$ with $\varphi(\mathbb{F}^c) = 0$ such that $r_{\rho_n}(x; \alpha) < \infty$ for all $x \in \mathbb{F}$, $\alpha \in \mathscr{D}$, and a sequence $\rho_n \uparrow 1$.]

The condition $0 \in \mathring{\mathscr{D}}(\Lambda)$ is used to prove that the level sets of $\Lambda^*(\cdot)$ (the convex conjugate of Λ) are compact. This makes it possible to extend the upper bound from compact sets to closed sets.

When $d = 1$ this can be done without the extra hypothesis $0 \in \mathring{\mathscr{D}}(\Lambda)$. The argument is entirely analogous to that for i.i.d. random variables.

These kinds of upper bound arguments are by now quite standard, so we omit the details. □

REFERENCES

[1] CHIANG, T.-S. (1985). On the lower bound of large deviation of random walks. *Ann. Probab.* **13**. 90–96.

[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. (1975). Asymptotic evaluation of certain Markov process expectations for large time. I–IV. *Comm. Pure Appl. Math.* **28** 1–47 and 279–301; **29** 389–461 (1976); **36** 183–212 (1983).

[4] ELLIS, R. (1984). Large deviations for a general class of random vectors. *Ann. Probab.* **12** 1–12.

[5] NEY, P. and NUMMELIN, E. (1987). Markov additive processes I. Eigenvalue properties and limit theorems. *Ann. Probab.* **15** 561–592.

[6] NEY, P. and NUMMELIN, E. (1984). Some limit theorems for Markov additive processes. *Proc. of Symposium on Semi-Markov Processes*. Brussels.

[7] NIEMI, S. and NUMMELIN, E. (1986). On non-singular renewal kernels with an application to a semigroup of transition kernels. *Stochastic Process. Appl.* **22** 177–202.

[8] NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-negative Operators*. Cambridge Tracts in Mathematics, **83**. Cambridge Univ. Press.

[9] VARADHAN, S. R. S. (1984). *Large Deviations and Applications*. SIAM, Philadelphia.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HELSINKI
HELSINKI
FINLAND