The Annals of Probability
1990, Vol. 18, No. 3, 1071–1098

LARGE DEVIATION LOWER BOUNDS FOR ADDITIVE FUNCTIONALS OF MARKOV PROCESSES[1]

BY NARESH C. JAIN

University of Minnesota

Let $X_1, X_2, \ldots$ be a Markov process with state space E, a Polish space. Let $L_n(\omega, A) = n^{-1}\sum_{j=0}^{n-1} 1_A(X_j(\omega))$ denote the normalized occupation time measure. If μ is a probability measure on E, G is a weak neighborhood of μ, and if $V \subset E$, then we obtain asymptotic lower bounds for probabilities $P^x[L_n(\omega, \cdot) \in G,\ X_j(\omega) \in V,\ 0 \le j \le n-1]$ in terms of $I(\mu)$, the rate function of Donsker and Varadhan. Our assumptions are weaker than those imposed by Donsker and Varadhan, and the proof works without any essential change in the continuous time case as well. In fact, the same proofs apply to certain bounded additive functionals: Let $r \ge 0$ and let $f\colon \Omega \to \mathbf{B}$ be bounded $\mathscr{F}_r^0$-measurable, where Ω is the sample space with the product topology (Skorohod topology in the continuous time case) and $\mathbf{B}$ is a separable Banach space; let $\theta_k\colon \Omega \to \Omega$ be the shift operator, i.e., $\theta_k\omega(j) = \omega(k+j)$. Then we get lower bounds for probabilities involving $n^{-1}(f(\omega) + f(\theta_1\omega) + \cdots + f(\theta_{n-1}\omega))$ in place of $L_n(\omega, \cdot)$. In this latter situation, the rate function has to be the entropy function $H(Q)$ of Donsker and Varadhan.

1. Introduction. Let E be a complete separable metric space and let $\mathscr{E}$ denote its Borel subsets. Our aim here is to establish large deviation lower bounds for additive functionals of a Markov process (in both discrete and continuous time) with state space E. We will consider two types of additive functionals: occupation times of the process (i.e., measure-valued) and bounded additive functionals taking values in a separable Banach space. One could give a formulation to include both cases at the expense of introducing too many technicalities; instead of doing that, we will consider the two cases separately. In fact, the basic proof is the same for the two cases.

In the discrete time case we will denote by Ω the space of doubly infinite sequences $\omega = (\ldots, \omega(-1), \omega(0), \omega(1), \ldots)$, with $\omega(j) \in E$, $-\infty < j < \infty$, and will give it the product topology. In the continuous time case Ω will either denote the space of E-valued continuous functions with the topology of uniform convergence on compacts or it will denote the space of E-valued functions which are right-continuous with left limits and will be given the Skorohod topology. In any case, Ω is itself a complete separable metric space. For $-\infty < s \le t < \infty$, $\mathscr{F}_t^s$ will denote the σ-algebra generated by $\omega(u)$, $s \le u \le t$; $\mathscr{F}_t^{-\infty}$ denotes the σ-algebra generated by $\omega(u)$, $u \le t$, and $\mathscr{F}_\infty^s$ denotes the σ-algebra generated by $\omega(u)$, $s \le u$. Finally, $\mathscr{F}_\infty^{-\infty} = \mathscr{F}$ denotes the

Received April 1988; revised March 1989.
[1]Work partially supported by NSF Grant DMS-86-0347.
AMS 1980 *subject classification*. Primary 60F10.
Key words and phrases. Markov process, continuous and discrete time, occupation times, bounded additive functionals, large deviations, lower bounds.

σ-algebra generated by $\omega(u)$, $-\infty < u < \infty$. As usual, θ_t: $\Omega \to \Omega$ denotes the shift operator $(\theta_t\omega)(s) = \omega(t+s)$.

In the continuous time context, $\pi(t, x, dy)$ will denote the transition probability function of a Markov process and $P^x[\Lambda]$ will denote the corresponding probability of $\Lambda \in \mathscr{F}_\infty^0$ when the process starts from x at time 0. In discrete time, $\pi(1, x, dy)$ will denote the one-step transition probability and $\pi(t, x, dy)$ will be the t-step transition when t is a positive integer. For $\omega \in \Omega$, $A \in \mathscr{E}$, we will write

$$L_n(\omega, A) = \frac{1}{n}\sum_{s=0}^{n-1} \chi_A(\omega(s)), \qquad n \geq 1, \tag{1.1}$$

and

$$L_t(\omega, A) = \frac{1}{t}\int_0^t \chi_A(\omega(s))\, ds, \qquad t > 0. \tag{1.2}$$

If $x \in E$, $A \in \mathscr{E}$, we will also use the notations

$$\psi_1(n, x, A) = \sum_{j=n}^{\infty} 2^{-j}\pi(j, x, A), \qquad n \geq 1, \tag{1.3}$$

and will simply write $\psi_1(x, A)$ when $n = 1$. In the continuous time context we will write

$$\psi_2(t, x, A) = \int_t^\infty e^{-s}\pi(s, x, A)\, ds \tag{1.4}$$

and if $t = 0$, we will simply write $\psi_2(x, A)$ for $\psi_2(0, x, A)$.

Instead of requiring the existence of a reference measure for the Markov process as Donsker and Varadhan do in most of their work [5, 7], we will make an irreducibility assumption whose consequences were studied in [10].

In the following we make the convention that a subscript $i = 1$ will correspond to the discrete time case and $i = 2$ to the continuous case. If the subscript is dropped, the statement will hold for either case.

We now introduce the irreducibility hypothesis in the two cases $i = 1, 2$. If the process satisfies this hypothesis, we will say that π is *α-irreducible*.

HYPOTHESIS $H_i(1)$. *There exists a probability measure α on $(E, \mathscr{E})$ such that $\alpha(A) > 0$ implies $\psi_i(x, A) > 0$ for all x.*

To explain why this weak condition does not suffice, we should first introduce the "rate function" of Donsker and Varadhan for a Markov process: let $C_b(E)$ denote the space of real-valued continuous functions on E and let $\mathscr{U} = \{u \in C_b(E): \inf u > 0\}$. For any probability measure μ on E, define

$$I_1(\mu) = \sup_{u \in \mathscr{U}} \int \log \frac{u(x)}{\pi u(x)}\, d\mu(x), \tag{1.5}$$

where $\pi u(x) = \int \pi(1, x, dy) u(y)$. For continuous time, also define

$$(1.6) \qquad I_2(\mu) = \begin{cases} \lim_{h \downarrow 0} \dfrac{I_1^{(h)}(\mu)}{h}, & \text{if the limit exists,} \\ \infty, & \text{otherwise,} \end{cases}$$

where $I_1^{(h)}(\mu)$ is given by (1.5) with πu replaced by $T_h u = \int \pi(h, \cdot, dy) u(y)$. It will be shown in the next section that $I_2(\mu) = \limsup_{h \downarrow 0} (I_1^{(h)}(\mu)/h)$.

If π is a Feller transition function, i.e., if $T_h u$ is bounded continuous whenever u is, for all $h > 0$, then Donsker and Varadhan have used the rate function given by

$$(1.7) \qquad \bar{I}_2(\mu) = -\inf_{u \in \mathscr{U} \cap D} \int \frac{Lu}{u}(x)\, d\mu(x),$$

where L denotes the infinitesimal generator of the process and D denotes the domain of L. The Feller property and the assumed right-continuity of paths immediately imply (cf. [3], Theorem 1.6, page 36) that D is dense in $C_b(E)$. It then follows from Lemma 3.1 of [4] that the limit in (1.6) actually exists and equals $\bar{I}_2$. Therefore, if π is Feller, then I_2 is the usual Donsker–Varadhan rate function.

Even though under less stringent conditions we will get exact analogues of lower bound results proved in [5], for now we state a typical lower bound result. Let μ be a probability measure on $(E, \mathscr{E})$ and let G be a weak neighborhood of μ. Then for certain x (under certain conditions) in the discrete time case,

$$(1.8) \qquad \liminf_{n \to \infty} \frac{1}{n} \log P^x\{L_n \in G\} \geq -I_1(\mu).$$

Example 3.6 shows that this fails for essentially all x even for a Markov chain satisfying the Doeblin condition ([8], page 192). Such a process always satisfies Hypothesis $H_1(1)$. Also, the singular part of $\pi(n, x, \cdot)$ with respect to the (unique) invariant measure goes to zero exponentially fast. In view of this example, we must strengthen $H_i(1)$ for a result like (1.8) to hold.

It is clear that if $H_i(1)$ is satisfied for some α and α_1 is defined by $\alpha_1(A) = \alpha(A \cap B)/\alpha(B)$, where B is a set of positive α-measure, then $H_i(1)$ is satisfied with α_1 as well. Therefore, most measures α that satisfy $H_i(1)$ will not satisfy $H_i(2)$ (given below) with α in place of β. For this reason, if the process is known to be α-irreducible to begin with, we define another probability measure by

$$(1.9) \qquad \beta(A) = \int \psi(x, A) \alpha(dx), \qquad A \in \mathscr{E},$$

where $\psi = \psi_1$ or ψ_2, introduced earlier. In addition to $H_i(1)$, we now impose:

HYPOTHESIS $H_i(2)$. *If* $I_i(\mu) < \infty$, *then* $\mu \ll \beta$.

This hypothesis will be shown to be a consequence of the following condition (see Proposition 2.8):

CONDITION A. The process is α-irreducible. Let β be given by (1.9). Then $\beta(A) = 0$ implies that for every $x \in E$ there exists a t such that $\psi_i(t, x, A) = 0$.

REMARK 1.1. In the original version of the paper, we stated $H_i(2)$ and Condition A in terms of the measure α itself and observed that the measure β would be more useful. We have now stated them in terms of β.

In the discrete case, de Acosta [1, 2] has introduced the following Conditions B_1 and B_2. We will show below that either condition immediately implies Condition A. Even though we will then work with Condition A or rather with $H_i(2)$, we would like to add that Condition B_2 is very pleasing because it is a simple condition in terms of the measure α itself.

DEFINITION 1.2. In the discrete time case, a nonempty set $C \subset E$ is said to be stochastically closed if $\pi(1, x, C) = 1$ for all $x \in C$.

CONDITION B_1 (de Acosta [2]). If C is stochastically closed, then given $x \notin C$ there exists $k \geq 1$ such that $\pi(k, x, C) = 1$.

CONDITION B_2 (de Acosta [2]). If the process is α-irreducible, $\alpha(A) = 0$ implies that given $x \in E$ there exists k such that $\pi(k, x, A) = 0$.

To see that either of these conditions implies Condition A, first observe that if C is stochastically closed and π is α-irreducible, then $\alpha(C^c) = 0$. Therefore, Condition B_2 implies B_1. Now assume Condition B_1 and that π is α-irreducible. If $\beta(A) = 0$, then $\psi_1(x, A) = 0$ a.e. (α). Therefore the set $C = \{x: \psi_1(x, A) = 0\}$, being nonempty, is clearly stochastically closed, and we then have $\beta(C^c) = 0$ because π is also β-irreducible by Proposition 2.7. By Condition B_1, given x there exists k such that $\pi(k, x, C^c) = 0$. This implies $\psi(k, x, A) = 0$ and Condition A is satisfied.

The results in Section 3 along with this remark subsume the results of de Acosta in [1] and their improvements in [2]. Some results of Ellis [9] on lower bounds also follow from ours (see Remark 3.10).

We next consider additive functionals. Let B be a separable Banach space with norm $\|\cdot\|$ and let $\mathscr{B}$ denote the σ-field of Borel subsets of B. Let $f: \Omega \to B$ be (strongly) $\mathscr{F}_r^0$-measurable and bounded, i.e., if $A \in \mathscr{B}$, then $f^{-1}(A) \in \mathscr{F}_r^0$, and $\sup_\omega \|f(\omega)\| < \infty$. Let

$$S_f(n, \omega) = f(\omega) + f(\theta\omega) + \cdots + f(\theta_{n-1}\omega), \qquad n \geq 1,$$

in the discrete case and

$$S_f(t, \omega) = \int_0^t f(\theta_s \omega)\, ds$$

in the continuous case. We will obtain large deviation lower bounds for $n^{-1}S_f(n, \cdot)$ in Section 3 in the discrete case and for $t^{-1}S_f(t, \cdot)$ in Section 4 in the continuous case. The method of proof is the same as for occupation time functionals. The rate function in the above generality has to be in terms of $H(Q)$, the entropy of the stationary process Q with respect to the given Markov process, introduced by Donsker and Varadhan [7]; the definition will be given in Section 2.

We will state some basic results of Donsker and Varadhan [7] in Section 2 that will be used later. This section will also contain some preliminary facts.

The main feature of our basic proof (Section 3), which is a variation of the method developed in [7], is that it applies to many situations without any essential change.

The main results of this paper are Theorems 3.3, 3.4, 3.5 (analogue of Theorem 3.1 [5]), 3.7 and 3.7′ in the discrete time case and their analogues Theorem 4.4, 4.5 and 4.8 in the continuous time case. Theorem 4.7 is an application to the distribution of $\sup_{0 \le s \le t} \|\omega(s)\|$, as $t \to \infty$.

REMARK 1.3. If instead of using the topology of weak convergence of probability measures, one uses the τ-topology, our proofs hardly need any change because the ergodic theorem holds for bounded measurable functions.

REMARK 1.4. Under the hypothesis $H_i(1)$ alone, we can give the lower bound in terms of the rate function $\hat{I}(\mu) = \inf\{H(Q)\colon Q \text{ stationary, ergodic, with marginal } \mu\}$. In general, $I(\mu) \le \hat{I}(\mu)$, but it is shown in [5] that under a hypothesis stronger than $H(2)$ one actually has $I(\mu) = \hat{I}(\mu)$.

In view of these remarks and the results of Section 3, it appears that in the discrete time case the main advantage of the Donsker–Varadhan method (that has been used here) is that the use of the ergodic theorem allows more flexibility and one can get lower bounds for more general sets without additional work. The major advantage of this approach is that essentially the same proofs work in both discrete and continuous time.

2. Preliminaries. We will introduce notations and prove results in this section which will sometimes apply in both discrete and continuous time; in such situations t will be real (integer) in the continuous (discrete) time context. We will write $P^{\omega(0)}$ for P^x when $x = \omega(0)$ and will denote by Q_ω the regular conditional of Q given $\mathscr{F}_0^{-\infty}$. (We are following Donsker and Varadhan [7] here.) Let $h(\lambda; \mu)$ denote the entropy of the probability measure μ with respect to λ. If Q is a stationary measure on $(\Omega, \mathscr{F})$, we define

$$H(t, Q) = E^Q\Big\{h\Big(P^{\omega(0)}\big|_{\mathscr{F}_t^0}; Q_\omega\big|_{\mathscr{F}_t^0}\Big)\Big\}, \tag{2.1}$$

where $\lambda|_{\mathscr{G}}$ denotes the restriction of λ to the σ-algebra $\mathscr{G}$ and E^Q denotes expectation with respect to Q. It is proved in [7] (Theorem 3.1) that either $H(t,Q) = \infty$ for all $t > 0$ or there exists a finite constant $H(Q)$ such that $H(t,Q) = tH(Q)$ for all $t > 0$. We observe that if $H(Q) < \infty$, then Q-a.s., $Q_\omega \ll P^{\omega(0)}$ on each $\mathscr{F}_t^0$, $t > 0$.

We state the following results of Donsker and Varadhan [7] and Varadhan [12]. The proofs have been given by them in the continuous time case, but they apply equally well (actually become simpler) in the discrete time context.

THEOREM 2.1. *$H(Q)$ is linear in Q; i.e., if $\lambda_1 + \lambda_2 = 1$, $\lambda_i \geq 0$, then $H(\lambda_1 Q_1 + \lambda_2 Q_2) = \lambda_1 H(Q_1) + \lambda_2 H(Q_2)$, where Q_1, Q_2 are stationary.*

THEOREM 2.2. *Suppose Q is stationary and ergodic and let*

$$\varphi_t(\omega, \cdot) = \log \frac{dQ_\omega}{dP^{\omega(0)}}\bigg|_{\mathscr{F}_t^0},$$

where Q_ω, $P^{\omega(0)}$ are restricted to $\mathscr{F}_t^0$. If $H(Q) < \infty$, then for Q-almost all ω, Q_ω-a.s.,

$$\lim_{t\to\infty} \frac{1}{t} \varphi_t(\omega, \cdot) = H(Q).$$

THEOREM 2.3 (Contraction principle). *Let μ be a probability measure on $(E, \mathscr{E})$. Then $I(\mu) = \inf\{H(Q)\colon Q$ stationary with marginal $\mu\}$, where $I = I_1$ or I_2, depending on the discrete or the continuous time context. In particular, if Q is stationary with marginal μ, then $H(Q) < \infty$ implies $I(\mu) < \infty$.*

PROOF. This is essentially Theorem 6.1 in [7]. The theorem shows that the contraction principle holds for $\bar{I}$ in place of I under the Feller property (in continuous time; in discrete time the Feller property is not used). We will now show that a minor modification of the proof in [7] gives the result for I without the assumption of the Feller property. We will need the following lemma.

LEMMA 2.4. *If* $\limsup_{h\downarrow 0}(I_1^{(h)}(\mu)/h) < \infty$, *then* $\lim_{h\downarrow 0}(I_1^{(h)}(\mu)/h)$ *exists and equals the* lim sup.

PROOF. Let μ be given and let us write $\theta(h) = I_1^{(h)}(\mu)$. It is easily seen that θ is subadditive, i.e., $\theta(h_1 + h_2) \leq \theta(h_1) + \theta(h_2)$. Let

$$\beta_1 = \liminf_{h\downarrow 0}(\theta(h)/h) \leq \limsup_{h\downarrow 0}(\theta(h)/h) = \beta_2 < \infty.$$

We can find $0 < a_j < b_j \downarrow 0$ such that

$$\lim_{j\to\infty} \frac{\theta(a_j)}{a_j} = \beta_1, \qquad \lim_{j\to\infty} \frac{\theta(b_j)}{b_j} = \beta_2.$$

For each j, there exists $k \geq 0$ such that $2^{-k-1}b_j < a_j \leq 2^{-k}b_j$ and by subadditivity, $\theta(b_j)/b_j \leq \theta(2^{-k}b_j)/2^{-k}b_j$. Therefore, we can replace b_j by $2^{-k}b_j$ and without loss of generality may assume that $a_j/b_j \to \eta \geq \frac{1}{2}$. By subadditivity, we have

$$\frac{\theta(b_j)}{b_j} \leq \frac{a_j}{b_j}\frac{\theta(a_j)}{a_j} + \frac{b_j - a_j}{b_j}\frac{\theta(b_j - a_j)}{b_j - a_j},$$

and letting $j \to \infty$, we get $\beta_2 \leq \eta\beta_1 + (1-\eta)\beta_2$, which proves the lemma. □

REMARK. Subadditivity alone is not enough for the conclusion of the lemma to hold.

Now observe that (6.5) in [7] shows that $\limsup_{h \downarrow 0}(I_1^{(h)}(\mu)/h) \leq l < \infty$, where l stands for $\inf\{H(Q)\colon Q$ stationary with marginal $\mu\}$. Therefore by Lemma 2.4, $I_2(\mu)$ [defined by (1.6)] is less than or equal to l. For inequality in the opposite direction, let $\limsup_{h \downarrow 0}(I_1^{(h)}(\mu)/h) = l < \infty$. Using the arguments in [7] leading to equation (6.9), for each $h > 0$ we can construct a stationary process $Q^{(h)}$ such that $h_{\mathscr{F}_t^0}(P_\mu; Q^{(h)}) < tl$. This shows that for each $t > 0$, the family of probability measures $\{Q^{(h)}\colon h > 0\}$ restricted to $\mathscr{F}_t^0$ is a tight family. We can find a subsequence which converges weakly to a stationary measure Q on $\mathscr{F}_\infty^0$ as $h \downarrow 0$. Since $h_{\mathscr{F}_t^0}(P_\mu; Q^{(h)})$ is lower semicontinuous in $Q^{(h)}$, we get $h_{\mathscr{F}_t^0}(P_\mu; Q) \leq tl$. Therefore,

$$\limsup_{t \to \infty} \frac{h_{\mathscr{F}_t^0}(P_\mu; Q)}{t} \leq l.$$

By (3.14) of [7] it then follows that $H(Q) \leq l$. None of the arguments here uses the Feller property, so we have the contraction principle without the Feller property, provided we use $I_2(\mu)$ in place of $\bar{I}_2(\mu)$ in the continuous time case. No change is needed in the discrete time case.

The next lemma has been implicitly used in [7]. We give a proof here for the benefit of the reader.

LEMMA 2.5. *Let Q be a stationary process with $H(Q) < \infty$. Let G be a weak neighborhood of Q. Then given $\varepsilon > 0$, there exists a stationary probability measure $Q_\varepsilon = \sum_{p=1}^k \lambda_p Q_p$, where $\lambda_p > 0$, Q_p is ergodic, $1 \leq p \leq k$, and $\sum \lambda_p = 1$, such that $Q_\varepsilon \in G$ and $|H(Q) - H(Q_\varepsilon)| < \varepsilon$.*

PROOF. Let $\varepsilon > 0$ be given. For some $\delta \leq \varepsilon$, we may take

$$G = \left\{v\colon \left|\int f_i\, dv - \int f_i\, dQ\right| < \delta, 1 \leq i \leq k\right\},$$

where the f_i are bounded continuous functions on Ω. From the proof of

Theorem 3.5 in [4], there exists a function f_0: $\Omega \to \mathbb{R}$ [in fact, $f_0(\omega) = h(P^{\omega(0)}|_{\mathscr{F}_1^0}; R_\omega)$] such that if $H(\hat{Q}) < \infty$ for a stationary process $\hat{Q}$, then $f_0 \in L^1(\hat{Q})$ and

$$(2.2)\qquad \int f_0(\omega)\, d\hat{Q}(\omega) = H(\hat{Q}).$$

By the integral representation theorem (see [11]), we also have

$$(2.3)\qquad \int f_i\, dQ = \int\int f_i(\omega)\, d\pi_{\omega'}(\omega)\, dQ(\omega'), \qquad 0 \le i \le k,$$

where each $\pi_{\omega'}$ is ergodic. For $i = 0$, this, in particular, implies that if $H(Q) < \infty$, then $H(\pi_{\omega'}) < \infty$, Q-a.s. ω'. For convenience we write

$$\xi_i(\omega') = \int f_i\, d\pi_{\omega'}, \quad 0 \le i \le k.$$

We pick an ergodic process π such that $H(\pi) < \infty$ and

$$(2.4)\qquad \int |f_i|\, d\pi \le a, \qquad 0 \le i \le k.$$

We can find b sufficiently large and $\Omega_\delta \subset \Omega$ such that (2.5)–(2.7) are satisfied for $\omega \in \Omega_\delta$:

$$(2.5)\qquad |\xi_i(\omega)| \le b,$$

$$(2.6)\qquad \sum_{i=0}^{k} \int_{\Omega_\delta^c} |\xi_i(\omega)|\, dQ(\omega) < \delta/3$$

and

$$(2.7)\qquad Q(\Omega_\delta^c) < \delta/3a.$$

Divide $[-b, b]$ into disjoint subintervals $J_1, J_2, \dots, J_r$, each of length $\le \delta/3$. Let

$$\Lambda_{ij} = \{\omega \in \Omega_\delta: \xi_i(\omega) \in J_j\}, \qquad 0 \le i \le k, 1 \le j \le r.$$

Then for each i, $\bigcup_{j=1}^{r} \Lambda_{ij} = \Omega_\delta$. We pick one set from each collection $\{\Lambda_{ij}: 1 \le j \le r\}$ and form the intersection of the $k + 1$ sets so picked. Let $\{\Lambda\}$ denote the partition of Ω_δ formed by such sets. If ω_Λ is an arbitrary point of Λ, then for $0 \le i \le k$,

$$\sum_{\Lambda} \left(\xi_i(\omega_\Lambda) - \frac{\delta}{3}\right) Q(\Lambda) - \frac{\delta}{3} \le \int \xi_i\, dQ \le \sum_{\Lambda} \left(\xi_i(\omega_\Lambda) + \frac{\delta}{3}\right) Q(\Lambda) + \frac{\delta}{3}.$$

From each Λ with $Q(\Lambda) > 0$, we pick ω_Λ so that $H(\pi_{\omega_\Lambda}) < \infty$. We define

$$Q_\varepsilon = \sum_{\Lambda} Q(\Lambda) \pi_{\omega_\Lambda} + (1 - Q(\Omega_\delta))\pi.$$

An easy computation gives for $0 \le i \le k$,

$$\int f_i \, dQ_\varepsilon - \delta < \int f_i \, dQ < \int f_i \, dQ_\varepsilon + \delta$$

and recalling that $\int f_0 \, dQ_\varepsilon = H(Q_\varepsilon)$, $\int f_0 \, dQ = H(Q)$ proves the lemma. □

The next lemma is an analogue of this one. It will apply when we deal with Banach space valued bounded measurable functions of a Markov process.

LEMMA 2.6. *Let Q be a stationary process with $H(Q) < \infty$. Let $f\colon \Omega \to B$ be a bounded measurable map, where B is a separable Banach space with norm $\|\cdot\|$. Then given $\varepsilon > 0$, there exists a stationary process Q_ε, which is a convex combination of ergodic ones, such that $|H(Q) - H(Q_\varepsilon)| < \varepsilon$ and $\|\int f dQ - \int f dQ_\varepsilon\| < \varepsilon$.*

PROOF. The proof is quite similar to that of the previous lemma. First, let $f_0\colon \Omega \to \mathbb{R}$ be the function (as in Lemma 2.5) which satisfies (2.2). We also have

$$\int f dQ = \int \int f(\omega) \, d\pi_{\omega'}(\omega) \, dQ(\omega')$$

and

$$\int f_0 \, dQ = \int \int f_0(\omega) \, d\pi_{\omega'}(\omega) \, dQ(\omega'),$$

where $\pi_{\omega'}$ are ergodic. We write $\xi(\omega') = \int f d\pi_{\omega'}$ and $\xi_0(\omega') = \int f_0 \, d\pi_{\omega'}$. Then there exist simple functions

$$\hat{\xi}(\omega') = \sum_{i=1}^{j} x_i 1_{\Lambda_i}(\omega'), \qquad \hat{\xi}_0(\omega') = \sum_{i=1}^{j} a_i 1_{\Lambda_i}(\omega'),$$

where $x_i = \int f(\omega) \, d\pi_{\omega_i'}(\omega)$ and $a_i = \int f_0(\omega) \, d\pi_{\omega_i'}(\omega)$, for some $\omega_i' \in \Lambda_i$, $1 \le i \le j$, and

$$\left| \int \hat{\xi} \, dQ - \int \xi \, dQ \right| < \varepsilon, \qquad \left| \int \hat{\xi}_0 \, dQ - \int \xi_0 \, dQ \right| < \varepsilon.$$

But we can write

$$\int \hat{\xi} \, dQ = \sum_{i=1}^{j} Q(\Lambda_i) \int f dP_{\omega_i'} = \int f dQ_\varepsilon,$$

$$\int \hat{\xi}_0 \, dQ = \sum_{i=1}^{j} Q(\Lambda_i) \int f_0 \, d\pi_{\omega_i'} = \int f_0 \, dQ_\varepsilon,$$

where $Q_\varepsilon = \sum_{i=1}^{j} Q(\Lambda_i) \, d\pi_{\omega_i'}$, which proves the lemma. □

If α is a probability measure on $(E, \mathscr{E})$, we define, as before,

$$\beta(A) = \int \psi(x, A)\alpha(dx), \tag{2.8}$$

where $\psi = \psi_1(\psi_2)$ in the discrete (continuous) case.

PROPOSITION 2.7. *Suppose α satisfies Hypothesis $H(1)$. Then β satisfies*

$$\beta(A) > 0 \Leftrightarrow \psi(x, A) > 0 \quad \textit{for all } x. \tag{2.9}$$

PROOF. We will prove this in the continuous time context. The proof also works in discrete time. We need only prove the implication $\Rightarrow$ in (2.9). Let $A \in \mathscr{E}$ and let $D = \{x: \psi_2(x, A) = 0\}$. If $D \neq \phi$ and $x \in D$, then

$$\pi(t+s, x, A) \geq \int_{D^c} \pi(t, x, dy)\pi(s, y, A),$$

so

$$0 = e^t\psi_2(x, A) \geq \int_0^\infty e^{-s}\pi(t+s, x, A)\, ds \geq \int_{D^c} \pi(t, x, dy)\psi_2(y, A)$$

and we must have $\pi(t, x, D^c) = 0$ for all $x \in D$, all $t > 0$. Therefore $\psi_2(x, D^c) = 0$ for all $x \in D$. This implies $\alpha(D^c) = 0$ if $D \neq \phi$ and we have

$$\beta(A) = \int_D \psi_2(y, A)\alpha(dy) + \int_{D^c} \psi_2(y, A)\alpha(dy) = 0,$$

which proves the proposition. □

The next proposition shows that Conditions A implies $H(2)$.

PROPOSITION 2.8. *Suppose Condition A holds. Then Hypothesis $H(2)$ is satisfied.*

PROOF. ($i = 1$) *Discrete case.* First assume Condition A. If $I_1(\mu) < \infty$, then by the contraction principle (Theorem 2.3) there exists a stationary process Q with marginal μ such that $H(Q) < \infty$. Then Q-a.s. ω, we have for $n \geq 1$,

$$Q_\omega \ll P^{\omega(0)} \quad \text{on } \mathscr{F}_n^0.$$

Suppose $\beta(A) = 0$. By Condition A we have for each ω,

$$P^{\omega(0)}\{\omega': \omega'(n) \in A\} = 0, \qquad n \geq \text{some } m. \tag{2.10}$$

Therefore Q-a.s. ω,

$$\lim_n Q_\omega\{\omega': \omega'(n) \in A\} = 0.$$

By dominated convergence,

$$0 = \int \lim_n Q_\omega\{\omega' : \omega'(n) \in A\}\, dQ(\omega)$$

$$\geq \limsup_n \int Q_\omega\{\omega'(n) \in A\}\, dQ(\omega) = \mu(A),$$

so $\mu \ll \beta$.

($i = 2$) *Continuous case.* Assume Condition A. We take Q as before with marginal μ and $H(Q) < \infty$. Then $Q_\omega \ll P^{\omega(0)}$ on $\mathscr{F}_t^0$, for all $t > 0$, Q-a.s. ω. Assume $\beta(A) = 0$. Let

$$V(t) = \{x : \psi_2(t, x, A) = 0\}.$$

By Condition A, we have $V(t) \nearrow E$ as $t \nearrow \infty$. Thus, if

$$Z(t) = \left\{\omega : \int_t^{t+1} e^{-s}\pi(s, \omega(0), A)\, ds = 0\right\},$$

then $Z(t) \nearrow \Omega$ as $t \nearrow \infty$ and

$$\int_t^{t+1} e^{-s}\, ds \int_{Z(t)} \pi(s, \omega(0), A)\, dQ(\omega) = 0.$$

It follows that

$$\int_t^{t+1} e^{-s}\, ds \int_{Z(t)} Q_\omega[\omega' : \omega'(s) \in A]\, dQ(\omega) = 0.$$

Therefore

$$\int_t^{t+1} e^{-s}\, ds \left\{\mu(A) - \int_{Z(t)^c} Q_\omega[\omega' : \omega'(s) \in A]\, dQ(\omega)\right\} = 0,$$

which implies

$$e^{-t-1}\mu(A) \leq \int_t^{t+1} e^{-s}\, ds\, Q\big(Z(t)^c\big) \leq \varepsilon e^{-t},$$

for all t sufficiently large (depending on ε). Therefore $\mu(A) \leq e\varepsilon$ for all $\varepsilon > 0$, so $\mu(A) = 0$. □

We will also need the next three lemmas.

LEMMA 2.9. *Let $A \in \mathscr{E}$ be such that $\psi(x, A) > 0$ for all x. Let μ be a probability measure on $(E, \mathscr{E})$. Then every set B with $\mu(B) > 0$ contains a compact subset B_1 with $\mu(B_1) > 0$ such that* $\inf_{x \in B_1} \pi(t, x, A) > 0$ *for some t.*

PROOF. We will prove it in the continuous case only. We have $\psi_2(x, A) = \int_0^\infty e^{-s}\pi(s, x, A)\, ds > 0$ for all x, so

$$\int_0^\infty e^{-s}\, ds \int_B \pi(s, x, A)\, d\mu(x) > 0.$$

Hence, for some t we have $\int_B \pi(t, x, A)\, d\mu(x) > 0$, which implies the conclusion of the lemma. □

LEMMA 2.10. *Let $B \in \mathscr{E}$ be such that $\psi(x, B) > 0$ for each x and $\psi(x, B)$ is a continuous function of x. Let K be a compact subset of E. Then there exist $t_0 > 0$, $\delta > 0$, such that for each $x \in K$, $\pi(t, x, B) \geq \delta$ for some $t \leq t_0$.*

PROOF. Again we will write the proof only in the continuous case. Since $\psi_2(x, B) > 0$ for all x and is continuous in x, we have $\inf_{x \in K} \psi_2(x, B) = \eta > 0$. Let $\delta = \eta/4$ and let t_0 be such that $\int_{t_0}^{\infty} e^{-t}\, dt < \eta/2$. Then for each $x \in K$, we have $\int_0^{t_0} e^{-t} \pi(t, x, B)\, dt \geq \eta/2$. Therefore, if $x \in K$,

$$\int_0^{t_0} e^{-t} \pi(t, x, B) \{ 1_{[t:\, \pi(t, x, B) \leq \eta/4]} + 1_{[t:\, \pi(t, x, B) > \eta/4]} \}\, dt \geq \eta/2,$$

where 1_A denotes the indicator of A, and we get

$$\int_0^{t_0} e^{-t} \pi(t, x, B) 1_{[t:\, \pi(t, x, B) > \eta/4]}\, dt \geq \eta/4.$$

Thus the Lebesgue measure of the set $[0 \leq t \leq t_0 \colon \pi(t, x, B) > \eta/4]$ is at least $\eta/4$, which proves the lemma. □

LEMMA 2.11. *Suppose $Q_n \to Q$ weakly, where Q_n and Q are stationary processes on $(\Omega, \mathscr{F})$ with marginals μ_n and μ. Then $\mu_n \to \mu$ weakly.*

PROOF. (i) *Discrete case.* Let f be a bounded, continuous function on E. Then

$$\int f\, d\mu_n = \int f(\omega(0))\, dQ_n(\omega) \to \int f(\omega(0))\, dQ(\omega) = \int f\, d\mu,$$

since $g(\omega)$ defined as $f(\omega(0))$ is a bounded continuous function of ω.

(ii) *Continuous case.* Let $\Omega_0 = \{\omega \in \Omega \colon \omega(0-) = \omega(0)\}$. If $\hat{Q}$ is any stationary process, then $\hat{Q}(\Omega_0) = 1$. It is then easy to apply essentially the above argument and reach the same conclusion. □

3. Discrete time.

PROPOSITION 3.1. *Suppose Q is stationary, ergodic, on $(\Omega, \mathscr{F})$ with marginal μ. Let $V \in \mathscr{E}$ be such that $\mu(V) = 1$ and let G be a weak neighborhood of μ. Then given $\varepsilon > 0$ and given any $A \in \mathscr{E}$ such that $\mu(A) > 0$, we have for μ-a.e. x in V,*

$$\begin{aligned}(3.1)\quad \liminf_n \frac{1}{n} \log P^x\{\omega \colon L_n(\omega, \cdot) \in G,\ L_{\hat{n}}(\omega, V) = 1,\\ \omega(j) \in A \text{ for some } j,\ n < j \leq \hat{n}\} \geq -(1 + \varepsilon) H(Q),\end{aligned}$$

where $\hat{n} \sim (1 + \varepsilon)n$.

PROOF. Let D_n denote the event in (3.1). Without loss of generality we may assume $H(Q) < \infty$. We have

$$P^{\omega(0)}\{D_n\} \geq \int_{D_n} \frac{dP^{\omega(0)}}{dQ_\omega}\bigg|_{\mathscr{F}_{\hat{n}}^0} dQ_\omega$$

$$\equiv \int_{D_n} \exp(-\varphi_{\hat{n}}(\omega, \omega'))\, dQ_\omega(\omega'),$$

where $\varphi_n(\omega, \omega') = \log(dQ_\omega/dP^{\omega(0)})|_{\mathscr{F}_n^0}(\omega')$. Note that $H(Q) < \infty$ implies $Q_\omega \ll P^{\omega(0)}$, Q-a.s. ω, but $P^{\omega(0)} \ll Q_\omega$ need not be true, so we have only inequality at the first step. Then for $\delta > 0$,

$$P^{\omega(0)}\{D_n\} \geq \int_{D_n \cap \{\varphi_{\hat{n}}(\omega,\cdot) \leq \hat{n}(H(Q)+\delta)\}} \exp(-\varphi_{\hat{n}})\, dQ_\omega,$$

$$\geq \exp(-\hat{n}(H(Q)+\delta))Q_\omega\{D_n \cap [\varphi_{\hat{n}}(\omega,\cdot) \leq \hat{n}(H(Q)+\delta)]\}.$$

We now observe that by the ergodic theorem

$$Q\{L_n \in G, \text{eventually}\} = 1$$

and since Q-a.s.,

$$\frac{1}{n}\left(\sum_{j=0}^{\hat{n}-1} \chi_A(\omega(j)) - \sum_{j=0}^{n-1} \chi_A(\omega(j))\right) \to \varepsilon\mu(A) > 0,$$

we have

$$Q\{\omega(j) \in A \text{ for some } j,\, n \leq j \leq \hat{n}, \text{eventually}\} = 1.$$

If Q is stationary with marginal μ, with $\mu(V) = 1$, then clearly

$$Q\{L_n(\omega, V) = 1\} = 1.$$

Finally, by Theorem 2.2, Q-a.s. ω,

$$Q_\omega\{\varphi_{\hat{n}}(\omega,\cdot) \leq \hat{n}(H(Q)+\delta), \text{eventually}\} = 1.$$

It therefore follows that Q-a.s. ω, as $n \to \infty$,

$$Q_\omega\{D_n \cap [\varphi_{\hat{n}}(\omega,\cdot) \leq \hat{n}(H(Q)+\delta)]\} \to 1.$$

Therefore Q-a.s. ω, for any $\delta > 0$ we have

$$\liminf_n \frac{1}{n} \log P^{\omega(0)}\{D_n\} \geq -(1+\varepsilon)(H(Q)+\delta).$$

Since μ is the marginal of Q, and $\delta > 0$ is arbitrary, (3.1) follows. □

Our aim now is to drop the ergodicity assumption in Proposition 3.1, but to do that we must make an additional assumption:

(3.2) $$\mu(B) > 0 \Rightarrow \pi(x, B) > 0, \qquad \mu\text{-a.e. } x.$$

We also note that by Lemmas 2.5 and 2.11 it is sufficient to consider Q of the form

$$(3.3)\qquad Q=\sum_{p=1}^{k}\lambda_p Q_p,\qquad \sum_{p=1}^{k}\lambda_p=1,\qquad \lambda_p>0,\ 1\le p\le k,$$

where each Q_p is ergodic.

The following obvious corollary of Proposition 3.1 will be used.

COROLLARY 3.2. *Given $\delta>0$, there exist a set $V_1\subset V$ with $\mu(V_1)>1-\delta$ and an integer $n_0=n_0(\varepsilon,\delta)$ such that for $n\ge n_0$,*

$$\inf_{x\in V_1}P^x\{D_n\}\ge \exp(-n(1+\varepsilon)(H(Q)+\delta)).$$

THEOREM 3.3. *Suppose Q is stationary on $(\Omega,\mathscr{F})$ with marginal μ. Let $V\in\mathscr{E}$ be such that $\mu(V)=1$ and let G be a weak neighborhood of μ. If (3.2) holds, then for μ-a.e. x in V we have*

$$(3.4)\qquad \liminf_n \frac{1}{n}\log P^x\{\omega: L_n(\omega,\cdot)\in G,\ L_n(\omega,V)=1\}\ge -H(Q).$$

PROOF. As observed earlier, we take $Q=\sum_{p=1}^k\lambda_pQ_p$. Then $\mu=\sum_{p=1}^k\lambda_p\mu_p$, where μ_p is the marginal of Q_p. Since $\mu(V)=1$, we have $\mu_p(V)=1$, $1\le p\le k$. Without any loss of generality, we take

$$(3.5)\qquad G=\left\{v:\left|\int f_j\,d\mu-\int f_j\,dv\right|<\eta,\ 1\le j\le r\right\},$$

where $\eta>0$ and f_j are bounded continuous on E with $\|f_j\|_\infty\le 1$, $1\le j\le r$. We then define

$$(3.6)\qquad G_p=\left\{v:\left|\int f_j\,d\mu_p-\int f_j\,dv\right|<\eta/2,\ 1\le j\le r\right\},\qquad 1\le p\le k.$$

Let $0<\varepsilon<\eta/4$ and $\delta>0$ be fixed. For each positive integer n we pick positive integers $n_1,\dots,n_k;n'_1,\dots,n'_k$ such that

$$k+(n_1+\cdots+n_k)+(n'_1+\cdots+n')=n,$$
$$n_p+n'_p\sim\lambda_p n,$$
$$n'_p\sim\varepsilon n_p,$$

for $1\le p\le k$. We now take $A_k=V$ and then by Corollary 3.2 we find $V_k\subset V$ and some m_k such that $\mu_k(V_k)\ge 1-\delta$ and for $m\ge m_k$,

$$(3.7)\qquad \inf_{x\in V_k}P^x\{L_m\in G_k,\ L_{\hat m}(V)=1\}\ge\exp(-(1+\varepsilon)m(H(Q_k)+\delta)),$$

where $\hat m\sim(1+\varepsilon)m$. Since $\mu(V_k)>0$, we have $\pi(1,x,V_k)>0$ for μ-a.e. x, hence for μ_{k-1}-a.e. x. We choose $A_{k-1}\subset V$ such that

$$\mu_{k-1}(A_{k-1})>\tfrac{1}{2},\qquad \inf_{x\in A_{k-1}}\pi(1,x,V_k)=\beta_{k-1}>0.$$

By Corollary 3.2, we then find $V_{k-1} \subset V$ and m_{k-1} such that $\mu_{k-1}(V_{k-1}) \geq 1 - \delta$ and for $m \geq m_{k-1}$,

$$\inf_{x \in V_{k-1}} P^x\{L_m \in G_{k-1}, L_{\hat{m}}(V) = 1, \omega(j) \in A_{k-1} \text{ for some } j, m < j \leq \hat{m}\}$$
$$\geq \exp(-(1+\varepsilon)m(H(Q_{k-1}) + \delta)).$$

Proceeding in this manner, we pick $V_1, A_1, V_2, A_2, \dots, V_k, A_k$, all contained in V, satisfying

$$(3.8) \quad \mu_p(A_p) > \tfrac{1}{2}, \qquad \inf_{x \in A_p} \pi(1, x, V_{p+1}) = \beta_p > 0, \qquad 1 \leq p \leq k-1,$$

and we pick m_p such that for $m \geq m_p$,

$$(3.9) \quad \begin{aligned} &\inf_{x \in V_p} P^x\{L_m \in G_p, L_{\hat{m}}(V) = 1, \omega(j) \in A_p \text{ for some } j, m < j \leq \hat{m}\} \\ &\qquad \geq \exp\big(-m(1+\varepsilon)\big(H(Q_p) + \delta\big)\big), \end{aligned}$$

for $1 \leq p < k$ and for $p = k$ we have (3.7). We now define stopping times $T_1, T_2, \dots, T_{k-1}$ as

$$\begin{aligned} T_1(\omega) &= \min\{j \geq n_1 : \omega(j) \in A_1\}, \\ T_2(\omega) &= \min\{j \geq n_2 + T_1 + 1 : \omega(j) \in A_2\}, \\ &\vdots \\ T_{k-1}(\omega) &= \min\{j \geq n_{k-1} + T_{k-2} + 1 : \omega(j) \in A_{k-1}\}. \end{aligned}$$

The stopping time is infinite if the corresponding set is empty. Let $T_0 \equiv 0$ and define

$$\begin{aligned} \Gamma_p = \Big\{\hat{L}_{n_p}(\omega, \cdot) \in G_p;\ \omega(j) \in V, T_{p-1}(\omega) < j \leq T_p(\omega); \\ T_p(\omega) \leq T_{p-1}(\omega) + 1 + n_p + n'_p\Big\}, \\ 1 \leq p \leq k-1, \end{aligned}$$

$$\Gamma_k = \Big\{\hat{L}_{n_k}(\omega, \cdot) \in G_k;\ \omega(j) \in V, T_{k-1}(\omega) < j \leq T_{k-1}(\omega) + 1 + n_k\Big\},$$

where

$$\hat{L}_{n_1}(\omega, \cdot) = L_{n_1}(\omega, \cdot), \qquad \hat{L}_{n_p}(\omega, \cdot) = L_{n_p}\big(\theta_{T_{p-1}+1}\omega, \cdot\big), \qquad 2 \leq p \leq k.$$

Since

$$\begin{aligned} &\sup_{\omega} \sup_{A \in \mathscr{E}} \left| L_n(\omega, A) - \sum_{p=1}^{k} \lambda_p \hat{L}_{n_p}(\omega, A) \right| \\ &\qquad \leq \sup_{\omega} \sup_{A} \left| L_n(\omega, A) - \sum_{p=1}^{k} \frac{n_p}{n} \hat{L}_{n_p}(\omega, A) \right| + \sum_{p=1}^{k} \left| \lambda_p - \frac{n_p}{n} \right| \\ &\qquad \leq \frac{n'_1 + \cdots + n'_k}{n} + \sum_{p=1}^{k} \left| \lambda_p - \frac{n_p}{n} \right| \leq 2\varepsilon \end{aligned}$$

for n sufficiently large and since $2\varepsilon < \eta/2$, we have

$$\left|\int f_j\,d\mu - \int f_j\,dL_n(\omega,\cdot)\right| = \left|\sum_{p=1}^{k} \lambda_p \int f_j\,d\mu_p - \int f_j\,dL_n(\omega,\cdot)\right|$$
$$< \sum_{p=1}^{k} \lambda_p \left|\int f_j\,d\mu_p - \int f_j\,d\hat{L}_{n_p}(\omega,\cdot)\right| + \eta/2.$$

It follows that there exists an m_0 (independent of ω) such that if $n \geq m_0$, then writing $\Lambda_r = \bigcap_{p=1}^{r} \Gamma_p$, $1 \leq r \leq k$, we have

$$\Lambda_k \subset \{L_n(\omega,\cdot) \in G,\ L_n(\omega,V) = 1\}.$$

For any x, by the strong Markov property,

$$P^x\{\Lambda_k\} \geq P^x\{\Lambda_{k-1}\} \inf_{z \in A_{k-1}} \pi(1,z,V_k)$$
$$\times \inf_{y \in V_k} P^y\{L_{n_k}(\omega,\cdot) \in G_k,\ \omega(j) \in V,\ 1 \leq j \leq n_k + n'_k\}$$
$$\geq P^x\{\Lambda_{k-1}\}\beta_{k-1}\exp(-(1+\varepsilon)n_k(H(Q)+\delta)).$$

Proceeding in this manner, we get uniformly for $x \in V_1$.

$$P^x\{L_n(\omega,\cdot) \in G,\ L_n(\omega,V) = 1\}$$
$$\geq \beta_1 \cdots \beta_{k-1}\exp\left(-(1+\varepsilon)(H(Q)+\delta)\sum_{p=1}^{k} n_p\right),$$

for all n sufficiently large. Therefore, if $x \in V_1$ ($\mu_1(V_1) > 1 - \delta$),

$$\liminf_n \frac{1}{n}\log P^x\{L_n(\omega,\cdot) \in G,\ L_n(\omega,V) = 1\}$$
$$\geq -(H(Q)+\delta)$$

and since $\delta > 0$ is arbitrary, we have (3.4) holding for a.e. μ_1-$x \in V$. Since $\mu_1,\dots,\mu_k$ can be arranged in any order for the proof, (3.4) actually holds for μ-a.e. x in V and the theorem is proved. □

The next theorem is an immediate corollary of Theorem 3.3 via the contraction principle (Theorem 2.3).

THEOREM 3.4. *Suppose μ is a probability measure on $(E,\mathscr{E})$ and $\mu(V) = 1$. Assume* (3.2). *If G is a weak neighborhood of μ, then for μ-a.e. x in V,*

$$\liminf_n \frac{1}{n}\log P^x\{L_n(\omega,\cdot) \in G,\ L_n(\omega,V) = 1\} \geq -I_1(\mu). \tag{3.10}$$

We can now derive the next theorem (compare with Theorem 3.1 of [5]) from Theorem 3.4 without much effort. As usual, π is said to be strongly Feller if $g(x) = \int f(y)\pi(1,x,dy)$ is continuous whenever f is bounded and measurable.

THEOREM 3.5. *Suppose μ is a probability measure on $(E, \mathscr{E})$ and $\mu(V) = 1$. Assume* (3.2). *Let G be a weak neighborhood of μ. If $x \in V$ is such that $\pi(1, x, \cdot\,)$ is not singular to μ, then* (3.1) *holds for such x. If π is strongly Feller and K is a compact set contained in V such that $x \in K$ implies $\pi(1, x, \cdot\,)$ is not singular to μ, then* (3.10) *holds uniformly over K.*

PROOF. We can find a neighborhood $\hat{G}$ of μ, $\hat{G} \subset G$ such that for all n sufficiently large,

$$\begin{aligned}&\{L_{n+1}(\omega, \cdot) \in G,\, L_{n+1}(\omega, V) = 1\}\\ &\qquad \supset \{\omega(0) \in V,\, L_n(\theta_1\omega, \cdot) \in \hat{G},\, L_n(\theta_1\omega, V) = 1\}.\end{aligned}$$

Since $\pi(1, x, \cdot\,)$ is not singular to μ, by Theorem 3.4, given $\eta > 0$ we can find a set $A \subset V$ and $m > 0$ such that $\pi(1, x, A) > 0$ and for $n \geq m$,

$$\inf_{y \in A} P^y\{L_n(\omega, \cdot) \in \hat{G},\, L_n(\omega, V) = 1\} \geq \exp(-(I_1(\mu) + \eta)n).$$

By the Markov property,

$$\begin{aligned}&P^x\{L_{n+1}(\omega, \cdot) \in G,\, L_{n+1}(\omega, V) = 1\}\\ &\qquad \geq \pi(1, x, A) \inf_{y \in A} P^y\{L_n(\omega, \cdot) \in \hat{G},\, L_n(\omega, V) = 1\}\end{aligned}$$

for all n sufficiently large. Therefore the result follows. For the second assertion, the choice of the set A above depends on $x \in K$, but since $\pi(1, x, A)$ is a continuous function of x, there exists a neighborhood U_x of x such that $\inf_{y \in U_x} \pi(1, y, A) > 0$. It follows that (3.10) holds uniformly over $U_x \cap K$. Since a finite number of such U_x cover K, (3.10) holds uniformly over K. This finishes the proof. □

Next we consider lower bounds for irreducible processes. To explain why irreducibility [Hypothesis $H_i(1)$] alone is not enough, we give an example below which shows that the lower bound part of the large deviation principle fails even for a Doeblin process (see Doob [8], page 192, for definition). Such a process always satisfies $H_i(1)$, is highly recurrent and the transition probability is almost absolutely continuous with respect to the invariant measure, for large n.

EXAMPLE 3.6. Let $E = \{x_1, x_2, x_3, x_4;\ y_1, y_2\}$. Let $\pi(x_i, \{x_{i+1}\}) = \frac{1}{2}$, $1 \leq i \leq 3$; $\pi(x_{i+1}, \{x_i\}) = \frac{1}{2}$, $1 \leq i \leq 3$; $\pi(x_1, \{x_4\}) = \pi(x_4, \{x_1\}) = \frac{1}{2}$; $\pi(y_1, \{y_2\}) = \pi(y_2, \{y_1\}) = \frac{1}{2}$; $\pi(y_1, \{x_1\}) = \pi(y_2, \{x_1\}) = \frac{1}{2}$. This defines an aperiodic Doeblin process. Starting from the set $\{x_1, x_2, x_3, x_4\}$ the process never leaves it, so the set is stochastically closed. Let μ be a probability measure such that $\mu(\{y_1\}) = \mu(\{y_2\}) = \frac{1}{2}$. Then for a positive function u defined on E, we have

$$\int \log\left(\frac{u}{\pi u}\right)(x)\, d\mu(x) = \frac{1}{2}\left\{\log\frac{2u(y_1)}{u(x_1) + u(y_2)} + \log\frac{2u(y_2)}{u(x_1) + u(y_1)}\right\} \leq \log 2.$$

Therefore $I_1(\mu) \le \log 2$. On the other hand, if we let $u(y_1) = u(y_2) = 1$ and allow $u(x_1)$ to be large, then we see that $I_1(\mu) = \log 2$. Let $G = \{v\colon |\int f\,d\mu - \int f\,dv| < 1\}$, where $f(x_i) = 0$, $1 \le i \le 4$; $f(y_i) = 1$, $i = 1,2$. Then $v \in G$ implies that $v(\{y_1, y_2\}) > 0$. It follows that $P^{x_i}\{L_n(\omega, \cdot\,) \in G\} = 0$ for all n, $1 \le i \le 4$, since $\{x_1, x_2, x_3, x_4\}$ is stochastically closed; but the large deviation principle would say that

$$\liminf n^{-1} \log P^{x_i}\{L_n(\omega, \cdot) \in G\} \ge -\log 2,$$

which is a contradiction.

It is clear from this example that we need more than $H_i(1)$ if the large deviation principle is to hold. The main result for irreducible processes in discrete time is the next theorem. One can also prove the analogue of Theorem 4.4 (continuous time) in the discrete time case. We leave it to the reader.

It will be useful to observe that if $H_1(1)$ holds, then every set A of positive α measure contains a subset B of positive α measure such that $\inf_{x,y \in B} p(r,x,y) > 0$, for some positive integer r, where $p(r,x,y)$ denotes the density of the absolutely continuous part of $\pi(r,x,\cdot\,)$ with respect to α. This is Lemma 2 in [10]. Let S denote the class of such sets B.

THEOREM 3.7. *Suppose Hypotheses $H_1(1)$ and $H_1(2)$ hold. Let G be a weak neighborhood of μ. Then for every $x \in E$,*

$$\liminf \frac{1}{n} \log P^x\{L_n(\omega, \cdot) \in G\} \ge -I_1(\mu). \tag{3.11}$$

If, further, $\psi_1(x, A)$ [defined by (1.3)] is a continuous function of x for each compact A, then (3.11) holds uniformly for x in a compact set. If $B \in S$ and for some r we have $\inf_{x \in C} \pi(r, x, B) > 0$, then (3.11) holds uniformly for $x \in C$.

NOTE. Recall that $H_1(2)$ can be replaced by Condition A.

PROOF. Without any loss of generality, we may assume $I_1(\mu) < \infty$. Then by Theorem 2.3, it will suffice to show that the lower bound is $-H(Q)$ for any stationary Q with marginal μ and $H(Q) < \infty$. By Lemmas 2.5 and 2.11 it will suffice to consider Q as in (3.3). Then $\mu = \sum \lambda_p \mu_p$. We take G and G_p, weak neighborhoods of μ and μ_p, defined by (3.5) and (3.6), respectively, as in the proof of Theorem 3.3.

We will use the following consequence of Corollary 3.2: Let Q be stationary, ergodic, with marginal μ. Then given $\delta > 0$, $\varepsilon > 0$, there exist $V \subset E$, $\mu(V) > 1 - \delta$ and an integer $n_0 = n_0(\varepsilon, \delta)$, such that for $n \ge n_0$,

$$\inf_{x \in V} P^x\{L_n \in G\} \ge \exp(-n(1+\varepsilon)(H(Q) + \delta)). \tag{3.12}$$

Let $0 < \varepsilon < \eta/4$ and $\delta > 0$ be fixed. (Recall that η occurs in the definitions of G, G_p.) We will now pick sets A_p, V_p, $1 \le p \le k$. Let $A_k = E$ and by (3.12) let V_k be such that $\mu_k(V_k) > 1 - \delta$ and for $n \geq$ some m_k,

$$\inf_{x \in V_k} P^x\{L_n \in G_k\} \geq \exp(-n(1+\varepsilon)(H(Q_k) + \delta)). \tag{3.13}$$

Since $\mu(V_k) > 0$ and $\mu \ll \beta$, we have $\beta(V_k) > 0$. By Lemma 2.9, applied to μ_{k-1}, there exist r_{k-1} and a compact set A_{k-1} such that $\mu_{k-1}(A_{k-1}) > 0$ and $\inf_{x \in A_{k-1}} \pi(r_{k-1}, x, V_k) > 0$. Next to Corollary 3.2 we can choose V_{k-1} with $\mu_{k-1}(V_{k-1}) > 1 - \delta$ such that for $n \geq$ some m_{k-1},

$$\begin{aligned} &\inf_{x \in V_{k-1}} P^x\{L_n \in G_{k-1}; \omega(j) \in A_{k-1} \text{ for some } j, n \leq j \leq \hat{n}\} \\ &\qquad \geq \exp(-n(1+\varepsilon)(H(Q_{k-1}) + \delta)), \end{aligned} \tag{3.14}$$

where $\hat{n} \sim n(1+\varepsilon)$. Proceeding in this manner, we get sets $V_1, A_1, V_2, A_2, \dots, V_k, A_k$ and positive integers $r_1, \dots, r_{k-1}, m_1, \dots, m_{k-1}$ such that

$$\inf_{x \in A_p} \pi(r_p, x, V_{p+1}) = \beta_p > 0, \qquad 1 \leq p \leq k-1,$$

and for $n \geq m_p$,

$$\begin{aligned} &\inf_{x \in V_p} P^x\{L_n \in G_p; \omega(j) \in A_p \text{ for some } j, n \leq j \leq \hat{n}\} \\ &\qquad \geq \exp\bigl(-n(1+\varepsilon)\bigl(H(Q_p) + \delta\bigr)\bigr), \end{aligned} \tag{3.15}$$

where $\hat{n} \sim (1+\varepsilon)n$ and $1 \leq p \leq k-1$. The rest is now very similar to the proof of Theorem 3.3. For n large, we pick $n_1, n_2, \dots, n_k$ and $n'_1, \dots, n'_k$ such that

$$\begin{aligned} (n_1 + \cdots + n_k) + (n'_1 + \cdots + n'_k) + (r_1 + \cdots + r_{k-1}) &= n, \\ n_p + n'_p &\sim \lambda_p n, \\ n'_p &\sim \varepsilon n_p, \end{aligned}$$

for $1 \leq p \leq k$. The stopping times $T_1, \dots, T_{k-1}$ are defined as

$$\begin{aligned} T_1(\omega) &= \min\{j \geq n_1 \colon \omega(j) \in A_1\}, \\ T_2(\omega) &= \min\{j \geq n_2 + T_1 + r_1 \colon \omega(j) \in A_2\}, \\ &\ \ \vdots \\ T_{k-1}(\omega) &= \min\{j \geq n_{k-1} + T_{k-2} + r_{k-2} \colon \omega(j) \in A_{k-1}\}. \end{aligned}$$

Let $T_0 = 0$ and, as before, let

$$\Gamma_p = \left\{\hat{L}_{n_p}(\omega, \cdot) \in G_p; T_p(\omega) \leq T_{p-1}(\omega) + r_{p-1} + n_p + n'_p\right\}, \qquad 1 \leq p \leq k-1$$

and $\Gamma_k = \{\hat{L}_{n_k}(\omega, \cdot) \in G_k\}$, where

$$\hat{L}_{n_1}(\omega, \cdot) = L_{n_1}(\omega, \cdot), \qquad \hat{L}_{n_p}(\omega, \cdot) = L_{n_p}\bigl(\theta_{T_{p-1}+r_{p-1}}\omega, \cdot\bigr), \qquad 2 \leq p \leq k.$$

We then write $\Lambda_r = \bigcap_{p=1}^r \Gamma_p$, $1 \leq r \leq k$. It is easy to see that for all n sufficiently large (independently of ω), $\Lambda_k \subset \{L_n \in G\}$. Following the rest of the argument in the proof of Theorem 3.3, we conclude that uniformly for $x \in V_1$,

$$P^x\{L_n(\omega, \cdot) \in G\} \geq \beta_1 \cdots \beta_k \exp(-(1+\varepsilon)(H(Q)+\delta)n)$$

for all n sufficiently large. As before, we then have for μ-a.e. x,

$$\liminf_n \frac{1}{n} \log P^x\{L_n(\omega, \cdot) \in G\} \geq -I_1(\mu). \tag{3.16}$$

This holds for any weak neighborhood G of μ. Given G, we can find $\hat{G} \subset G$ such that given r, for all n sufficiently large (depending on r),

$$\{L_{n+r}(\omega, \cdot) \in G\} \supset \{L_n(\theta_r \omega, \cdot) \in \hat{G}\}.$$

We then find a set B, $\mu(B) > 0$, such that by (3.16), given $\delta > 0$, for $n \geq$ some m,

$$\inf_{y \in B} P^y\{L_n(\omega, \cdot) \in \hat{G}\} \geq \exp(-(I_1(\mu)+\delta)n). \tag{3.17}$$

Now $\beta(B) > 0$, so given x, there exists r such that $\pi(r, x, B) > 0$. Therefore

$$P^x\{L_{n+r}(\omega, \cdot) \in G\} \geq \pi(r, x, B) \inf_{y \in B} P^y\{L_n(\omega, \cdot) \in \hat{G}\},$$

hence (3.16) holds for *all* x. Now suppose that $\psi_1(x, A)$ is continuous in x for each compact A. Since (3.11) holds for all x, it holds uniformly for x in some compact set B with $\beta(B) > 0$. Then $\psi_2(x, B) > 0$ for all x and is continuous in x, so by Lemma 2.10, given a compact set K, there exists an integer r_0 such that for each $x \in K$, $\pi(r, x, B) > 0$ for some $r \leq r_0$. As long as all such r (corresponding to all $x \in K$) are uniformly bounded by r_0, the argument given above (going from μ-a.e. x to every x) applies and we have (3.11) holding uniformly for $x \in K$. Finally, we can clearly take $B \in S$ such that (3.11) holds uniformly for $x \in B$. Since $\inf_{x \in C} \pi(r, x, B) > 0$ for some r, this same argument again applies for the last assertion, and the theorem is proved. □

The next corollary of Theorem 3.7 is proved in [1] as Theorem 6.3 by different methods.

COROLLARY 3.8. *Suppose Hypothesis $H_1(1)$ is satisfied. Also assume Condition B_1. Let G be a weak neighborhood of μ. Then for every x,*

$$\liminf_n \frac{1}{n} \log P^x\{L_n(\omega, \cdot) \in G\} \geq -I_1(\mu).$$

PROOF. We already observed in Section 1 that Condition B_1 implies Condition A, so the corollary is proved. □

REMARK 3.9. If for every x we have

$$\liminf_n \frac{1}{n} \log P^x\{L_n(\omega, \cdot) \in G\} \geq -I_1(\mu),$$

then for every probability measure v on $(E, \mathscr{E})$ we have

$$\liminf_n \frac{1}{n} \log \int P^x\{L_n(\omega, \cdot) \in G\} v(dx) \geq -I_1(\mu).$$

This is obvious because, writing $\phi_n(x) = P^x\{L_n(\omega, \cdot) \in G\}$, we have for $\delta > 0$, for any x,

$$\lim_n \min\{1, \exp(n(I_1(\mu) + \delta))\phi_n(x)\} = 1.$$

Therefore,

$$\int \liminf_n (\exp(n(I_1(\mu) + \delta))\phi_n(x))v(dx) \geq 1$$

and by Fatou's lemma

$$\liminf_n \int \exp(n(I_1(\mu) + \delta))\phi_n(x)v(dx) \geq 1,$$

so

$$\int \phi_n(x)v(dx) \geq \tfrac{1}{2}\exp(-n(I_1(\mu) + \delta))$$

for all n sufficiently large, and the conclusion follows.

REMARK 3.10. Ellis [9] makes the assumption that there exist an $M > 0$ and positive integers m and n such that for all $x, y \in E$ and for all $A \in \mathscr{E}$, $\pi(m, x, A) \leq M\pi(n, y, A)$. This condition easily implies the existence of a probability measure α for which $H_1(1)$ and Condition A holds. To see this, if $n < m$, then

$$\pi(m, x, A) = \int \pi(n, y, A)\pi(m - n, x, dy) \geq M^{-1}\pi(m, x', A)$$

for any x', by assumption. So in this case $m = n$ may be assumed without any loss of generality; then take $\alpha = \pi(m, x_0, \cdot)$ for any $x_0 \in E$, which clearly satisfies $H_1(1)$ and Condition A. If $m < n$, take $\alpha = \pi(n, x_0, \cdot)$ for some x_0. Clearly α satisfies Condition A. If $\alpha(A) > 0$, but $\psi_1(x, A) = 0$ for some x, then $\pi(n, x, A) = 0$, so $\pi(m, y, A) = 0$ for all y, but then

$$\pi(n, x_0, A) = \int \pi(m, y, A)\pi(n - m, x_0, dy) = 0,$$

a contradiction, so $\alpha(A) > 0$ implies $\psi_1(x, A) > 0$ for all x and $H_1(1)$ holds.

Our next observation is that the method applies equally well to bounded additive functionals. Let $r \geq 0$ be an integer. Suppose B is a separable Banach space and $f\colon \Omega \to B$ is a bounded $\mathscr{F}_r^0$-measurable mapping. Let

$$\frac{1}{n} S_f(n, \omega) = \frac{1}{n}\{f(\omega) + f(\theta_1\omega) + \cdots + f(\theta_{n-1}\omega)\}, \qquad n \geq 1.$$

The following analogue of Proposition 3.1 is proved in exactly the same manner. We need only observe that if Q is stationary, ergodic, then $n^{-1}S_f(n,\omega) \to \int f\,dQ$, Q-a.s., and $\sup_\omega \|f(\omega)\| < \infty$ by assumption, where $\|\cdot\|$ denotes the Banach norm.

PROPOSITION 3.1′. *Suppose Q is stationary, ergodic. Let G $(\subset B)$ be a neighborhood of $\int f\,dQ$. Suppose μ is the marginal of Q and $A \in \mathscr{E}$ is such that $\mu(A) > 0$. Then given $\varepsilon > 0$, we have for μ-a.e. x,*

$$\liminf_n \frac{1}{n} \log P^x\{n^{-1}S_f(n,\omega) \in G,\ \omega(j) \in A \text{ for some } j,\ n \le j \le \hat{n}\} \ge -(1+\varepsilon)H(Q),$$

where $\hat{n} \sim (1+\varepsilon)n$.

Using this proposition, we then get the analogue of Theorem 3.7.

THEOREM 3.7′. *Suppose Hypotheses $H_1(1)$ and $H_1(2)$ hold. Let Q be a stationary process, and let G be a neighborhood of $\int f\,dQ$ in B. Then for every x,*

$$\liminf_n \frac{1}{n} \log P^x\{n^{-1}S_f(n,\omega) \in G\} \ge -H(Q). \tag{3.18}$$

If, moreover $\psi_1(x, B)$ is a continuous function of x for every compact B, then the conclusion holds uniformly for x in a compact set. If $B \in S$ and for some r, $\inf_{x \in C} \pi(r, x, B) > 0$, then the conclusion holds uniformly for x in C.

PROOF. The proof of this theorem requires no essential change from that of Theorem 3.7. It should be observed, however, that Lemma 2.6 will substitute for Lemma 2.5 at the appropriate step. We omit the details.

REMARK 3.11. The right-hand side in (3.18) can be replaced by $-\inf\{H(Q): \int f\,dQ \in G\}$. If $f(\omega) = g(\omega(0))$ for some bounded measurable $g: E \to B$ and μ is the marginal of Q, then $\int f\,dQ = \int g\,d\mu$ and the right-hand side can be replaced by $-\inf\{I_1(\mu): \int g\,d\mu \in G\}$. See also [1]. It is shown in [5] that if $E = B$, $g(x) = x$, $\pi(1, x, A) = \lambda(A - x)$ for all $A \in \mathscr{E}$, $x \in E$, then

$$\inf\left\{I_1(\mu): \int x\,d\mu = a\right\} = \sup_{\theta \in E^*}[\langle\theta, a\rangle - \log M(\theta)],$$

where $M(\theta) = \int \exp(\langle\theta, x\rangle)\,d\lambda(x)$.

The above observation corresponds to $r = 0$. In general,

$$f(\omega) = g(\omega(0), \omega(1), \dots, \omega(r-1))$$

for some $g: E \times \cdots \times E \to B$, a bounded measurable map. Then the r-dimensional marginal of Q will play the appropriate role.

4. Continuous time. The development here parallels that of the discrete time situation, once the appropriate definitions have been made. Some more notation has to be introduced before the main result can be stated.

Let U be an open subset of E and let C be a closed subset of U. We assume that μ is a probability measure whose support is contained in C. As in [5], we write

$$\pi_U(t, x, A) = P^x\{\omega: \omega(s) \in U, 0 \le s \le t, \omega(t) \in A\}, \tag{4.1}$$

where A is a Borel subset of U. Let

$$\psi_U(t, x, A) = \int_t^\infty e^{-s}\pi_U(s, x, A)\, ds \tag{4.2}$$

and write $\psi_U(x, A)$ for $\psi_U(0, x, A)$.

The next proposition is an analogue of Proposition 3.1.

PROPOSITION 4.1. *Let U, C and μ be as above. Let Q be a stationary ergodic measure on $(\Omega, \mathscr{F})$ with marginal μ. Suppose G is a weak neighborhood of μ and let $A \subset U$ be such that $\mu(A) > 0$. Then given $\varepsilon > 0$, we have for μ-a.e. x,*

$$\liminf_{t \to \infty} \frac{1}{t} \log P^x\{L_t(\omega, \cdot) \in G, \omega(s) \in C, 0 \le s \le \hat{t}; \omega(u) \in A \text{ for some } u, t \le u \le \hat{t}\} \ge -(1 + \varepsilon)H(Q),$$

where $\hat{t} \sim (1 + \varepsilon)t$ as $t \to \infty$.

PROOF. The main thing to observe is that if Q has marginal μ, then

$$Q\{\omega: \omega(s) \in C, 0 \le s \le t, s \in \text{rationals}\} = 1$$

for all $t \ge 0$. Since C is closed and the paths ω are right-continuous, we then have

$$Q\{\omega: \omega(s) \in C, 0 \le s \le t\} = 1, \qquad t \ge 0.$$

The rest of the proof parallels that of Proposition 3.1 and is omitted. □

The next corollary of this proposition is the analogue of Corollary 3.2.

COROLLARY 4.2. *Given $\delta > 0$, $\varepsilon > 0$, there exists a compact set $K \subset C$ with $\mu(K) > 1 - \delta$ and there exists $t_0 = t_0(\varepsilon, \delta) > 0$ such that for $t \ge t_0$,*

$$\inf_{x \in K} P^x\{L_t(\omega, \cdot) \in G, \omega(s) \in C, 0 \le s \le \hat{t}; \omega(u) \in A \text{ for some } u, t \le u \le \hat{t}\} \ge \exp\{-\hat{t}(H(Q) + \delta)\}, \tag{4.3}$$

where $\hat{t} \sim (1 + \varepsilon)t$ as $t \to \infty$.

We now need to make hypotheses which will replace Hypotheses $H_2(1)$ and $H_2(2)$.

HYPOTHESIS H_3. *There exists a probability measure α such that if $A \subset U$, then $\alpha(A) > 0$ implies that $\psi_U(x, A) > 0$ for all $x \in U$.*

HYPOTHESIS H_4. *There exists a probability measure α such that $I_2(\mu) < \infty \Rightarrow \mu \ll \alpha$.*

LEMMA 4.3. *Assume Hypothesis H_3. Let $A \subset U$ be such that $\alpha(A) > 0$. Then given any probability measure v and $B \subset C$ with $v(B) > 0$, there exists a compact set $B_1 \subset B$ with $v(B_1) > 0$ such that for some $t_0 > 0$ we have $\inf_{x \in B_1} \pi_U(t_0, x, A) > 0$.*

PROOF. We have

$$\int_0^\infty e^{-u}\, du \int_B \pi_U(t, x, A)\, dv(x) > 0.$$

Therefore, for some $t_0 > 0$ we have $\int_B \pi_U(t_0, x, A)\, dv(x) > 0$, which implies the conclusion immediately. □

We now state and prove the main result.

THEOREM 4.4. *Assume Hypotheses H_3 and H_4 (with the same α). Let G be a weak neighborhood of μ. Then for all $x \in U$ we have*

$$\liminf_{t \to \infty} \frac{1}{t} \log P^x\{L_t(\omega, \cdot) \in G, \omega(s) \in U, 0 \le s \le t\} \ge -I_2(\mu). \tag{4.4}$$

If, moreover, for each compact $A \subset U$, $\psi_U(x, A)$ is continuous in x, then the conclusion holds uniformly for x in any compact subset of U.

PROOF. The proof of this theorem is very similar to that of Theorem 3.7 and we will give the details where necessary. First, we may assume that $I_2(\mu) < \infty$; otherwise there is nothing to prove. Next, by the contraction principle (Theorem 2.3), it suffices to consider stationary measures Q on $(\Omega, \mathscr{F})$ with marginal μ and $H(Q) < \infty$. We need only prove (4.4) with $H(Q)$ replacing $I_2(\mu)$ on the right side. By Lemma 2.5 we only need to consider Q of the form $\sum_{p=1}^k \lambda_p Q_p$, where each Q_p is stationary and ergodic and $\sum \lambda_p = 1$, $\lambda_p > 0$, $1 \le p \le k$. As before, let μ_p denote the marginal of Q_p, so $\mu_p(C) = 1$ for each p. We pick G and G_p defined by (3.5) and (3.6) as in the proof of Theorem 3.3. Note that η below is the same as in the definitions of G and G_p.

Let $0 < \varepsilon < \eta/4$ and $\delta > 0$ be fixed. We then pick sets A_p and V_p as follows: Let $A_k = C$. Then by Corollary 4.2 pick $V_k \subset C$, a compact set, such that $\mu_k(V_k) > 1 - \delta$ and for $t \ge$ some m_k,

$$\begin{aligned} &\inf_{x \in V_k} P^x\{L_t(\omega, \cdot) \in G_k, \omega(s) \in A_k, 0 \le s \le \hat{t}\} \\ &\qquad \ge \exp(-\hat{t}(H(Q_k) + \delta)), \end{aligned} \tag{4.5}$$

where $\hat{t} \sim (1 + \varepsilon)t$. Since $\mu(V_k) > 0$ and $\mu \ll \alpha$ (by H_4), we have $\alpha(V_k) > 0$.

By Lemma 4.3, there exists a compact $A_{k-1} \subset C$ such that $\mu_{k-1}(A_{k-1}) > 0$ and $\inf_{x \in A_{k-1}} \pi_U(s_{k-1}, x, V_k) > 0$ for some $s_{k-1} > 0$. By Corollary 4.2 again we pick $V_{k-1} \subset C$ with $\mu_{k-1}(V_{k-1}) > 1 - \delta$ such that for $t \geq$ some m_{k-1},

$$\begin{aligned} &\inf_{x \in V_{k-1}} P^x\{L_t(\omega, \cdot) \in G_{k-1}, \omega(s) \in C, 0 \leq s \leq \hat{t}; \\ &\qquad\qquad \omega(u) \in A_{k-1} \text{ for some } u, t \leq u \leq \hat{t}\} \\ &\quad \geq \exp(-\hat{t}(H(Q_{k-1}) + \delta)), \end{aligned} \tag{4.6}$$

where $\hat{t} \sim (1 + \varepsilon)t$. Since $\alpha(V_{k-1}) > 0$, we can pick compact $A_{k-2} \subset C$ such that $\mu_{k-2}(A_{k-2}) > 0$ and $\inf_{x \in A_{k-2}} \pi_U(s_{k-2}, x, V_{k-1}) > 0$ for some s_{k-2}. Proceeding in this manner, we get sets $V_1, A_1, V_2, A_2, \ldots, V_k, A_k$ and positive numbers $s_0, s_1, \ldots, s_{k-1}, m_1, \ldots, m_{k-1}$, such that $s_0 = 0$,

$$\inf_{x \in A_p} \pi_U(s_p, x, V_{p+1}) = \beta_p > 0, \qquad 1 \leq p \leq k - 1,$$

$$\mu_p(V_p) > 1 - \delta, \qquad \mu_p(A_p) > 0, \qquad 1 \leq p \leq k,$$

and for $t \geq m_p$,

$$\begin{aligned} &\inf_{x \in V_p} P^x\{L_t(\omega, \cdot) \in G_p, \omega(s) \in C, 0 \leq s \leq \hat{t}; \\ &\qquad \omega(u) \in A_p \text{ for some } u, t \leq u \leq \hat{t}\} \geq \exp\big(-\hat{t}\big(H(Q_p) + \delta\big)\big), \end{aligned}$$

where $\hat{t} \sim (1 + \varepsilon)t$, $1 \leq p \leq k - 1$. The rest of the proof is essentially the same as in Theorem 3.3. For t large, we pick $t_1, \ldots, t_k, t'_1, \ldots, t'_k$ such that ($s_0 = 0$, $s_1, \ldots, s_{k-1}$ have already been chosen)

$$\begin{gathered} (t_1 + \cdots + t_k) + (t'_1 + \cdots + t'_k) + (s_1 + \cdots + s_{k-1}) = t, \\ t_p + t'_p \sim \lambda_p t, \\ t'_p \sim \varepsilon t_p, \end{gathered}$$

$1 \leq p \leq k$. The stopping times $T_1, \ldots, T_{k-1}$ are defined as

$$\begin{aligned} T_1(\omega) &= \inf\{t \geq t_1 \colon \omega(t) \in A_1\}, \\ T_2(\omega) &= \inf\{t \geq t_2 + T_1 + s_1 \colon \omega(t) \in A_2\}, \\ &\vdots \\ T_{k-1}(\omega) &= \inf\{t \geq t_{k-1} + T_{k-2} + s_{k-2} \colon \omega(t) \in A_{k-1}\}. \end{aligned}$$

Let

$$T_0(\omega) \equiv 0,$$

and let

$$\begin{aligned} \Gamma_p = \Big\{&\hat{L}_{t_p}(\omega, \cdot) \in G_p; \omega(s) \in C, T_{p-1}(\omega) + s_{p-1} < s \leq T_p(\omega); \\ &T_p(\omega) \leq T_{p-1}(\omega) + s_{p-1} + t_p + t'_p\Big\}, \qquad 1 \leq p \leq k - 1, \end{aligned}$$

$$\Gamma_k = \Big\{\hat{L}_{t_k}(\omega, \cdot) \in G_k; \omega(s) \in C, T_{k-1} + s_{k-1} < s \leq T_{k-1} + s_{k-1} + t_k + t'_k\Big\},$$

where $\hat{L}_{t_1}(\omega, \cdot) = L_{t_1}(\omega, \cdot)$ and $\hat{L}_{t_p}(\omega, \cdot) = L_{t_p}(\theta_{T_{p-1}+s_{p-1}}\omega, \cdot)$, $2 \le p \le k$. Also, let $\Sigma_1 = \Omega$,

$$\Sigma_p = \{\omega: \omega(s) \in U, T_{p-1} < s \le T_{p-1} + s_{p-1}\}, \qquad 2 \le p \le k$$

and

$$\Lambda_p = \bigcap_{r=1}^{p} (\Gamma_r \cap \Sigma_r), \qquad 1 \le p \le k.$$

We have

$$\begin{aligned}\sup_{\omega} \sup_{A \in \mathscr{E}} & \left| L_t(\omega, A) - \sum_{p=1}^{k} \lambda_p \hat{L}_{t_p}(\omega, A) \right| \\ & \le \sup_{\omega} \sup_{A} \left| L_t(\omega, A) - \frac{1}{t} \sum_{p=1}^{k} t_p \hat{L}_{t_p}(\omega, A) \right| + \sum_{p=1}^{k} \left| \lambda_p - \left(\frac{t_p}{t} \right) \right| \\ & \le \frac{t_1' + \cdots + t_k' + s_1 + \cdots + s_{k-1}}{t} + \sum_{p=1}^{k} \left| \lambda_p - \left(\frac{t_p}{t} \right) \right| < 2\varepsilon,\end{aligned}$$

for all t sufficiently large. Since $\varepsilon < \eta/4$, it follows that there exists m_0 (independent of ω) such that if $t \ge m_0$, we have

$$\Lambda_k \subset \{L_t(\omega, \cdot) \in G, \omega(s) \in U, 0 \le s \le t\}.$$

For $x \in V_1$, we have by the strong Markov property,

$$\begin{aligned}P^x(\Lambda_k) \ge P^x(\Lambda_{k-1}) & \inf_{z \in A_{k-1}} \pi_U(s_{k-1}, z, V_k) \\ & \times \inf_{y \in V_k} P^y\{L_{t_k}(\omega, \cdot) \in G_k, \omega(s) \in C, 0 \le s \le t_k + t_k'\}.\end{aligned}$$

Iterating this, we get for $t \ge$ some t_0, $x \in V_1$,

$$P^x\{\Lambda_k\} \ge \beta_1 \cdots \beta_k \exp(-(1+\varepsilon)(H(Q) + \delta)).$$

Since $\mu_1(V_1) > 1 - \delta$ and we can arrange $Q_1, \dots, Q_k$ in any order, this shows that for μ-a.e. x in C we have (4.4). If $x \in C$ is arbitrary, then by Hypothesis H_3 we must have, for any A with $\alpha(A) > 0$, $\pi_U(t, x, A) > 0$ for some $t > 0$. Since (4.4) must hold uniformly over some $A \subset C$ with $\mu(A) > 0$, it follows that (4.4) must hold for every $x \in U$. We can take A to be compact in this argument. To prove the second assertion, note that the proof of Lemma 2.10 applies as well to ψ_U and π_U. Therefore, if A is compact with $\alpha(A) > 0$, then $\psi_U(x, A) > 0$ for all x by Hypothesis H_3. Let K be a compact subset of E. By Lemma 2.10 there exists $t_0 > 0$ and $\delta > 0$ such that $\pi_U(t, x, A) \ge \delta$ for all $x \in K$ and for some t (depending on x), $t \le t_0$. This is clearly enough for (4.4) to hold uniformly over K. This proves the theorem. □

Theorem 4.5 is an immediate consequence if we take $U = E$ in Theorem 4.4. Note that H_3 then becomes $H_2(1)$, so by Proposition 2.7 the process is also β-irreducible and then $H_2(2)$ can replace H_4.

THEOREM 4.5. *Assume that Hypotheses $H_2(1)$ and $H_2(2)$ hold. Let G be a weak neighborhood of μ. Then for all x,*

$$\liminf_{t\to\infty} \frac{1}{t}\log P^x\{L_t(\omega,\cdot)\in G\} \geq -I_2(\mu).$$

If, moreover, for each compact A, $\psi_2(x, A)$ is a continuous function of x, then the conclusion holds uniformly for x in any compact set.

REMARK 4.6. One can follow the arguments given in [6], Section 2, to show that H_3 is satisfied for a class of processes with stationary independent increments and that $\psi_U(x, A)$ is continuous in x.

We now give an application of Theorem 4.4. Suppose $\pi(t, x, \cdot)$ satisfies the Feller property, i.e., if f: $E \to R$ is bounded continuous, then the function $\int f(y)\pi(t, x, dy) = g(t, x)$ is a continuous function of x for all t. Let C be a compact subset of $\mathscr{M}(E)$, the set of probability measures on E with the topology of weak convergence. Then a result of Donsker and Varadhan [see (7.1) of [5]] tells us that

$$\limsup_{t\to\infty} \frac{1}{t} \sup_{x\in E} \log P^x\{L_t(\omega,\cdot)\in C\} \leq -\inf_{v\in C} I_2(v). \tag{4.7}$$

Now suppose $E = \mathbb{R}^d$, the d-dimensional Euclidean space. Let $V_a = \{v\colon v \in \mathscr{M}(E),\ v([x\colon \|x\| \leq a]) = 1\}$. Clearly V_a is compact in the topology of weak convergence and (4.7) means

$$\limsup_{t\to\infty} \frac{1}{t} \sup_{x\in E} \log P^x\{\|\omega(s)\| \leq a,\ 0 \leq s \leq t\} \leq -\inf_{v\in V_a} I_2(v). \tag{4.8}$$

Let $0 \leq b < a$ and let $U = \{x\colon \|x\| < a\}$. Let $K = \{x\colon \|x\| \leq b\}$, a compact subset of U. If the conditions of Theorem 4.4 are satisfied, then we have

$$\liminf_{t\to\infty} \frac{1}{t} \inf_{\|x\|\leq b} \log P^x\{\|\omega(s)\| < a,\ 0 \leq s \leq t\} \geq -\inf_{v\in V_b} I_2(v). \tag{4.9}$$

If $\inf_{v\in V_b} I_2(v) < \infty$, then given $\varepsilon_n \searrow 0$, we can find $v_n \in V_{a_n}$, where $a_n \searrow b$, such that $I_2(v_n) \leq \inf_{\mu\in V_{a_n}} I_2(\mu) + \varepsilon_n$. A subsequence of $\{v_n\}$, again denoted by $\{v_n\}$, converges weakly to some $v \in V_b$. Hence, by the lower semicontinuity of I_2 (used at the second inequality below) we have

$$\inf_{\mu\in V_b} I_2(\mu) \leq I_2(v) \leq \liminf_n I_2(v_n) \leq \liminf_n \Big(\inf_{\mu\in V_{a_n}} I_2(\mu) + \varepsilon_n\Big).$$

Using this, (4.8) and (4.9) we conclude the following theorem.

THEOREM 4.7. *Let $E = \mathbb{R}^d$ and let π be a Feller semigroup. Suppose the conditions of Theorem 4.4 are satisfied for U of the form $\{x\colon \|x\| < a\}$. Then* *given* $\delta > 0$, *given* $b \geq 0$, *there exists* $a > b$ *such that* $a \to b$ *as* $\delta \to 0$ *and*

$$\eta \leq \liminf_{t\to\infty} \frac{1}{t} \inf_{\|x\|\leq b} \log P^x\Big\{ \sup_{0\leq s\leq t} \|\omega(s)\| \leq a \Big\}$$

$$\leq \limsup_{t\to\infty} \frac{1}{t} \sup_x \log P^x\Big\{ \sup_{0\leq s\leq t} \|\omega(s)\| \leq a \Big\} \leq \eta + \delta,$$

where $\eta = -\inf_{v\in V_b} I_2(v)$ *is assumed to be finite.*

Finally, we state the analogue of Theorem 3.7′ without proof. Let $r \geq 0$ and $f\colon \Omega \to B$, B a separable Banach space, be bounded $\mathscr{F}_r^{\,0}$-measurable. Let

$$S_f(t, \omega) = \int_0^t f(\theta_s \omega)\, ds.$$

THEOREM 4.8. *Suppose Hypotheses* $H_2(1)$ *and* $H_2(2)$ *hold. Let* Q *be a stationary process and let* G *be a neighborhood of* $\int f\, dQ$ *in* B. *Then for every* x,

$$\liminf_{t\to\infty} \frac{1}{t} \log P^x\{t^{-1} S_f(t, \omega) \in G\} \geq -H(Q).$$

If, moreover, $\psi_2(x, B)$ *is a continuous function of* x *for every compact* B, *then the conclusion holds uniformly for* x *in any compact set.*

The remarks after Theorem 3.7′ apply in continuous time as well.

REFERENCES

[1] DE ACOSTA, A. (1988). Large deviations for vector-valued functionals of a Markov chain: Lower bounds. *Ann. Probab.* **16** 925–960.

[2] DE ACOSTA, A. (1988). Large deviations for empirical measures of Markov chains. *J. Theoret. Probab.* To appear.

[3] DYNKIN, E. B. (1965). *Markov Processes* **1**. Springer, New York.

[4] DONSKER, M. D. and VARADHAN, S. R. S. (1975). Asymptotic evaluation of certain Markov process expectations for large time. I. *Comm. Pure Appl. Math.* **28** 1–47.

[5] DONSKER, M. D. and VARADHAN, S. R. S. (1976). Asymptotic evaluation of certain Markov process expectations for large time. III. *Comm. Pure Appl. Math.* **29** 389–461.

[6] DONSKER, M. D. and VARADHAN, S. R. S. (1977). On laws of the iterated logarithm for local times. *Comm. Pure Appl. Math.* **30** 707–753.

[7] DONSKER, M. D. and VARADHAN, S. R. S. (1983). Asymptotic evaluation of certain Markov process expectations for large time. IV. *Comm. Pure Appl. Math.* **36** 183–212.

[8] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[9] ELLIS, R. S. (1988). Large deviations for the empirical measure of a Markov chain with an application to multivariate empirical measure. *Ann. Probab.* **16** 1496–1508.

[10] JAIN, N. C. and JAMISON, B. (1967). Contributions to Doeblin's theory of Markov processes. *Z. Wahrsch. Verw. Gebiete* **8** 19–40.

[11] MAITRA, A. (1977). Integral representations of invariant measures. *Trans. Amer. Math. Soc.* **229** 209–225.

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

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MINNESOTA
127 VINCENT HALL
206 CHURCH STREET S.E.
MINNEAPOLIS, MINNESOTA 55455