*The Annals of Applied Probability*
2003, Vol. 13, No. 2, 395–429


# SADDLEPOINT APPROXIMATIONS AND NONLINEAR BOUNDARY CROSSING PROBABILITIES OF MARKOV RANDOM WALKS

BY HOCK PENG CHAN[1] AND TZE LEUNG LAI[2]

*National University of Singapore and Stanford University*

Saddlepoint approximations are developed for Markov random walks $S_n$ and are used to evaluate the probability that $(j-i)g((S_j-S_i)/(j-i))$ exceeds a threshold value for certain sets of $(i,j)$. The special case $g(x)=x$ reduces to the usual scan statistic in change-point detection problems, and many generalized likelihood ratio detection schemes are also of this form with suitably chosen $g$. We make use of this boundary crossing probability to derive both the asymptotic Gumbel-type distribution of scan statistics and the asymptotic exponential distribution of the waiting time to false alarm in sequential change-point detection. Combining these saddlepoint approximations with truncation arguments and geometric integration theory also yields asymptotic formulas for other nonlinear boundary crossing probabilities of Markov random walks satisfying certain minorization conditions.

**1. Introduction.** Let $\{S_n : n \geq 1\}$ be a $d$-dimensional random walk with Markov-dependent increments. In this paper, we study boundary crossing probabilities and asymptotic distributions of the scan statistics $\max_{1\leq k\leq n} \|S_n - S_k\|$. More generally, for $g:\mathbf{R}^d \to \mathbf{R}$ define

$$\mathcal{M}_n = \max_{1\leq i<j\leq n:\, j-i\in J_n} (j-i)g\big((S_j-S_i)/(j-i)\big), \tag{1.1}$$

$$T_c = \inf\left\{n : \max_{k<n:\, n-k\in J(c)} (n-k)g\big((S_n-S_k)/(n-k)\big) > c\right\}, \tag{1.2}$$

where $J_n$ and $J(c)$ are subsets of $\{1,2,\dots\}$. The special case $g(x)=\|x\|$ and $J_n = J(c) = \{1,2,\dots\}$ corresponds to the usual scan statistics. Under certain conditions, we show that there exist $q \in \{0,\dots,d\}$ and $r>0$ (depending on $g$) such that

(1.3) $e^{-c/r}(c/r)^{q/2}T_c$ has a limiting exponential distribution as $c\to\infty$,

(1.4) $\mathcal{M}_n - r\{\log n + (q/2)\log\log n\}$ has a limiting Gumbel-type distribution as $n\to\infty$.

Received July 2001; revised August 2002.

[1]Supported in part by the National University of Singapore RP-115-000-003-112 and RP-115-000-019-112.

[2]Supported in part by NSF, the National Security Agency and the Center for Advanced Study in the Behavioral Sciences.

*AMS 2000 subject classifications.* Primary 60F05, 60F10, 60G40; secondary 60G60, 60J05.

*Key words and phrases.* Markov additive processes, large deviation, maxima of random fields, change-point detection, Laplace's method, integrals over tubes.

In the case of a one-dimensional random walk with i.i.d. increments and $g(\mu) = \mu$, (1.4) with $q = 0$ has been established by Iglehart (1972) in the context of longest waiting times in a $GI/G/1$ queue and by Karlin, Dembo and Kawabata (1990) in the context of high-scoring segments in a DNA sequence. The corresponding result (1.3) in this case follows from Theorem 2 of Siegmund (1988) in his analysis of the CUSUM charts in quality control. Assuming the i.i.d. increments of the random walk to be standard normal random variables, Siegmund and Ventrakaman (1995) subsequently also established (1.3) with $r = 1$ and $q = 1$ for the case $g(\mu) = \mu^2/2$, which is associated with the generalized likelihood ratio control chart. The asymptotic theory concerning (1.1) and (1.2), which is presented in Sections 4 and 5, unifies these previous results and also leads to definitive solutions of a variety of change-point detection problem; see Chan and Lai (2002) for details. Of particular interest in these applications are (i) the extension of i.i.d. to Markov-dependent increments for the scan statistics (so that more general stochastic systems can be treated), and (ii) suitable choice of $g$ and $J_n$ or $J(c)$ in (1.1) or (1.2) to achieve both statistical and computational efficiency.

A unified approach to derive (1.3) and (1.4) is given in Sections 4 and 5. It is based on integrating saddlepoint approximations for Markov random walks with respect to certain measures over tubular neighborhoods of $q$-dimensional manifolds in $\mathbf{R}^d$. Saddlepoint approximations for the density function of $S_n$ with i.i.d. increments were introduced by Daniels (1954) in the case $d = 1$ and by Borovkov and Rogozin (1965) for general $d$; see Jensen (1995). Höglund (1974) and Jensen (1991) extended these saddlepoint approximations to $S_n = \sum_{i=1}^n f(X_i, X_{i-1})$ for certain uniformly recurrent Markov chains $\{X_i\}$. By integrating the saddlepoint approximations of the density function of $S_n$ over certain subsets $B$ of $\mathbf{R}^d$, Borovkov and Rogozin (1965) and Iltis (1995) derived asymptotic approximations of the large deviation probabilities $P(n^{-1}S_n \in B)$. Our derivation of (1.3) and (1.4) involves deeper geometric integration ideas that incorporate both the critical temporal and spatial components of the problem in some $q$-dimensional submanifold of $\mathbf{R}^d$, where $q$ is the same as that in (1.3) and (1.4). A brief overview of our method is given in Section 4, and the details of the argument are given in Section 5.

The saddlepoint approximations developed in Sections 2 and 6 for Markov random walks are much more general than those in the literature. First, the Markov random walks we consider do not need to be of the form $\sum_{i=1}^n f(X_i, X_{i-1})$. Secondly, whereas previous results assume the $X_i$ to be uniformly recurrent so that the "tilted transition kernel" [see (2.5) in Section 2] has nice analytic and boundedness properties, the uniform recurrence assumption is too restrictive in applications and Theorem 2 in Section 2 is able to dispense with this restrictive assumption. Ney and Nummelin (1987) have replaced the uniform recurrence assumption by certain minorization conditions in establishing the large deviation principle and characterizing the rate function for Markov additive processes.

Our saddlepoint approximation in Theorem 2 is based on these minorization conditions. In the Ney–Nummelin large deviations framework, the events to be considered require the terminal state $X_n$ to belong to a "sufficiently small" set (or $s$-set) and the initial state $X_0$ to belong to a "full set" on which certain eigenfunctions behave well. Since saddlepoint approximations are much more precise than large deviation bounds, it is natural to expect that they would at least require similar restrictions on the initial and terminal states. However, we are able to remove these restrictions via a truncation argument when we apply the saddlepoint approximations to analyze boundary crossing probabilities. The crucial ingredients for truncation argument are provided in Section 3, in which we show (i) how such truncation can be carried out under finiteness of certain eigenmeasures and (ii) that the eigenmeasures are indeed finite when certain "drift conditions" hold, which is the case for many time series and queueing models, as shown by Meyn and Tweedie (1993).

Because of practical difficulties in requiring the eigenfunctions to behave well at the initial and terminal states, other approaches to the large deviation principle for additive functionals of Markov chains have been developed that involve instead of eigenvalues and eigenfunctions more flexible tools like "convergence parameter" or "convex conjugate" to characterize the rate function; see, for example, Dinwoodie (1993) and de Acosta and Ney (1998). On the other hand, this more flexible approach only gives limits (or more precisely, lim sup and lim inf) of the logarithms of the probabilities of large deviations of these additive functionals, but we need the precise order of magnitude of the probabilities to derive the limiting distributions in (1.3) and (1.4). The methods in Section 3, which enable us to establish the precise order of magnitude for the large deviation probabilities, also provide new techniques to analyze the tilted transition kernels and remove some of the obstacles in applying the eigenvalue–eigenfunction approach of Ney and Nummelin (1987).

**2. Saddlepoint approximations for Markov random walks.** Let $\{(X_n, S_n): n = 0, 1, \dots\}$ be a Markov additive process, with $X_n$ being a Markov chain defined on a general state space $\mathcal{X}$ and $S_n$ taking values in $\mathbf{R}^d$. The additive component $S_n$ of the process is called a Markov random walk, and can be written in the form $S_n = S_0 + \xi_1 + \cdots + \xi_n$, where for all $s \in \mathbf{R}^d$,

$$\begin{aligned} &P\{(X_1, S_1) \in A \times (B + s) | (X_0, S_0) = (x, s)\} \\ &\quad = P\{(X_1, S_1) \in A \times B | (X_0, S_0) = (x, 0)\}, \end{aligned}$$

which we denote by $P(x, A \times B)$. The corresponding $m$-step transition kernel will be denoted by $P^m$. We shall assume throughout the sequel that $S_0 = 0$ and that $\{X_n\}$ is aperiodic and irreducible with respect to a maximal irreducibility measure $\varphi$ on $\mathcal{X}$. In this section we give saddlepoint approximations for the distribution of $(X_n, S_n)$. Throughout the sequel we denote the tranpose of a matrix by $'$ and the elements of $\mathbf{R}^d$ by column vectors. We also let $|M|$ denote the determinant of a square matrix $M$.

To begin with, suppose $\xi_1, \xi_2, \dots$ are i.i.d. with $Ee^{\theta'\xi_1} < \infty$ for some $\theta \neq 0$. Let $\Theta = \{\theta : Ee^{\theta'\xi_1} < \infty\}$ and let $\psi(\theta) = \log(Ee^{\theta'\xi_1})$ for $\theta \in \Theta$. Let $\Gamma$ be the interior of $\nabla\psi(\Theta)$. Suppose $S_m$ has an integrable characteristic function for some $m \geq 1$. Then for $n \geq m$, $n^{-1}S_n$ has a continuous density function $f_n$ for which the saddlepoint approximation

$$f_n(\mu) = \left\{1 + \sum_{j=1}^{k} c_j(\theta_\mu)n^{-j} + O(n^{-(k+1)})\right\}(n/2\pi)^{d/2}|V(\mu)|^{-1/2}e^{-nI(\mu)} \tag{2.1}$$

holds for all $\mu \in \Gamma$ and $k \geq 1$, where the $c_j(\theta)$ are analytic functions of $\theta$,

$$\begin{aligned} \theta_\mu &= (\nabla\psi)^{-1}(\mu), \\ I(\mu) &= \sup_{\theta\in\Theta}\{\theta'\mu - \psi(\theta)\} = \theta_\mu'\mu - \psi(\theta_\mu), \\ V(\mu) &= \nabla^2\psi(\theta_\mu). \end{aligned} \tag{2.2}$$

The function $\nabla\psi$ is a diffeomorphism from the interior of $\Theta$ onto $\Gamma$, and $\theta_\mu$ is a saddlepoint of the function $h(\theta) = \theta'\mu - \psi(\theta)$. Such saddlepoint approximations were introduced by Daniels (1954) in the case $d = 1$ and extended to general $d$ by Borovkov and Rogozin (1965). The function $I$ is called the *rate function* in large deviations theory. An obvious analogue of (2.1) also holds for $P\{S_n = s\}$ when $\xi_1$ has a lattice distribution and $s$ belongs to the minimal lattice; compare Jensen (1995).

2.1. *The uniformly recurrent case.* We first generalize the results to Markov random walks under the uniform recurrence condition of Iscoe, Ney and Nummelin (1985): There exist $\kappa \geq 1$, $b > a > 0$ and a probability measure $\nu$ on $\mathcal{X} \times \mathbf{R}^d$ such that

$$a\nu(A \times B) \leq P^\kappa(x, A \times B) \leq b\nu(A \times B) \tag{2.3}$$

for all $x \in \mathcal{X}$, measurable subsets $A$ of $\mathcal{X}$ and Borel subsets $B$ of $\mathbf{R}^d$.

Let $\Theta = \{\theta : \int_{\mathcal{X}\times\mathbf{R}^d} e^{\theta's}\, d\nu(x, s) < \infty\}$ and assume that its interior is nonempty. For $\theta \in \Theta$, define the transform kernels $\hat{P}_\theta$, $\hat{\nu}_\theta$ by

$$\hat{P}_\theta(x, A) = \int_{\mathbf{R}^d} e^{\theta's} P(x, A \times ds), \qquad \hat{\nu}_\theta(A) = \int_{\mathbf{R}^d} e^{\theta's}\nu(A \times ds). \tag{2.4}$$

Under (2.3), for every $\theta \in \Theta$, $a\hat{\nu}_\theta(A) \leq \hat{P}_\theta^\kappa(x, A) \leq b\hat{\nu}_\theta(A)$ and $\hat{P}_\theta$ has a maximal simple real eigenvalue $e^{\psi(\theta)}$ with eigenfunction $r(x;\theta)$ which is uniformly positive and bounded. Moreover, $\psi(\theta)$ is analytic and strictly convex on $\mathrm{Int}(\Theta)$, the interior of $\Theta$. Let $\Gamma$ be the interior of $\nabla\psi(\Theta)$ and define $\theta_\mu$, $I(\mu)$ and $V(\mu)$ as in (2.2). We shall use $P_\delta$ to denote the probability measure under which $X_0$ has distribution $\delta$, and let $P_x$ denote the case for which the initial distribution $\delta$ is degenerate at $x$. For $\theta \in \mathrm{Int}(\Theta)$, define the Markov additive transition kernel

$$Q_\theta(x, dy \times ds) = e^{-\psi(\theta)+\theta' s} P(x, dy \times ds) r(y;\theta)/r(x;\theta). \tag{2.5}$$

The underlying Markov chain $\{X_n : n \geq 0\}$ associated with this tilted transition kernel has a stationary distribution which is absolutely continuous with respect to $\hat{\nu}_0$, and the density function (with respect to $\hat{\nu}_0$) of the stationary distribution will be denoted by $\pi(y;\theta)$.

THEOREM 1. *Assume that* (2.3) *holds and* $\mathrm{Int}(\Theta) \neq \varnothing$.

(i) *Suppose $\nu$ is absolutely continuous with respect to the product measure $\hat{\nu}_0 \times \lambda$, where $\lambda$ denotes Lebesgue measure on $\mathbf{R}^d$. Let $v = d\nu/d(\hat{\nu}_0 \times \lambda)$, that is, $\nu(dx \times ds) = v(x,s)\, d\hat{\nu}_0(x)\, ds$. Assume that there exists $1 < \rho < 2$ such that*

$$\sup_{x \in \mathcal{X}, \theta \in K} \int_{\mathbf{R}^d} \big[e^{\theta' s} v(x,s)\big]^\rho \, ds < \infty \qquad \text{for every compact subset } K \text{ of } \mathrm{Int}(\Theta). \tag{2.6}$$

*Then for all sufficiently large $n$, $(X_n, n^{-1}S_n)$ has (under $P_x$) a joint density function $f_{n,x}$ with respect to $\hat{\nu}_0 \times \lambda$ and*

$$\begin{aligned} f_{n,x}(y,\mu) &= \left\{1 + \sum_{j=1}^{k} n^{-j} c_j(\theta_\mu, x, y) + O(n^{-(k+1)})\right\} \\ &\quad \times e^{-nI(\mu)}\big\{(n/2\pi)^{d/2}|V(\mu)|^{-1/2} r(x;\theta_\mu)/r(y;\theta_\mu)\big\}\pi(y;\theta_\mu) \end{aligned} \tag{2.7}$$

*for every $k \geq 1$, uniformly for $\mu \in C$ and $x, y \in \mathcal{X}$, where $C$ is any compact subset of $\Gamma$ and $c_j(\theta, x, y)$ are analytic functions of $\theta$.*

(ii) *Suppose $\xi_1$ has a lattice distribution with minimal lattice $L$ (of full rank $d$) under $P_x$, for every $x \in \mathcal{X}$. Then for every $k \geq 1$,*

$$\begin{aligned} &P_x\{S_n = u, X_n \in dy\} \\ &\quad = \left\{1 + \sum_{j=1}^{k} n^{-j} c_j(\theta_{u/n}, x, y) + O(n^{-(k+1)})\right\} \\ &\qquad \times e^{-nI(u/n)}\big\{(2\pi n)^{-d/2} h_L |V(u/n)|^{-1/2} r(x;\theta_{u/n})/r(y;\theta_{u/n})\big\} \\ &\qquad \times \pi(y;\theta_{u/n})\, d\hat{\nu}_0(y), \end{aligned} \tag{2.8}$$

*uniformly for $x, y \in \mathcal{X}$ and $u/n \in C$ with $u \in L$, where $C$ is any compact subset of $\Gamma$, $h_L$ is some constant dependent only on the lattice $L$ and $c_j(\theta, x, y)$ is the same as in* (i).

Theorem 1 can be proved by modifying the arguments in Sections 2–4 of Jensen (1991) who considers sums of real-valued functions $g(X_n)(= \xi_n)$ for the case $\kappa = 1$. The proof of Theorem 1(ii) uses similar methods and standard arguments for the lattice case [cf. Chapter 5 of Bhattacharya and Ranga Rao (1976)]. The constant $h_L$ in (2.8) is given by the absolute value of the determinant of the matrix $(\eta_1, \ldots, \eta_d)$, whose column vectors form a basis of $L$ in the sense that linear combinations of $\eta_1, \ldots, \eta_d$ with integer coefficients span $L$; see Bhattacharya and Ranga Rao [(1976), pages 228–231]. Note that in (2.8), the measure on $\mathcal{X}$ defined by $P_x\{S_n = u,\ X_n \in \cdot\}$ is absolutely continuous with respect to $\hat{\nu}_0$, so (2.8) can be interpreted as an asymptotic approximation to the Radon–Nikodym derivative of this measure with respect to $\hat{\nu}_0$.

2.2. *Regeneration under a minorization condition.* Instead of the uniform recurrence condition (2.3), we now assume the considerably weaker minorization condition of Ney and Nummelin (1987): There exist $\kappa \geq 1$, a probability measure $\nu$ on $\mathcal{X}$ and a finite measure $h(x, \cdot)$ on $\mathbf{R}^d$ such that

$$P^{\kappa}(x, A \times B) \geq h(x, B)\nu(A) \tag{2.9}$$

for all $x \in \mathcal{X}$ and all Borel subsets $B$ of $\mathbf{R}^d$ and measurable subsets $A$ of $\mathcal{X}$. An alternative form of minorization is

$$P^{\kappa}(x, A \times B) \geq h(x)\nu(A \times B), \tag{2.10}$$

where $\nu$ is a probability measure on $\mathcal{X} \times \mathbf{R}^d$ and $h$ is a nonnegative function on $\mathcal{X}$ with $\int h\, d\varphi > 0$. Under (2.9) or (2.10), Ney and Nummelin (1987) showed that $(X_n, S_n)$ admits a regenerative scheme with i.i.d. interregeneration times for an augmented Markov chain, which is called the "split chain." Letting $w(\theta, \zeta) = E_\nu e^{\theta' S_\tau - \zeta\tau}$, where $\tau$ is the first time ($> 0$) to reach the atom of the split chain, and assuming that

$$W := \{(\theta, \zeta) : w(\theta, \zeta) < \infty\} \qquad \text{is an open subset of } \mathbf{R}^{d+1}, \tag{2.11}$$

they also showed that $\Theta := \{\theta : w(\theta, \zeta) < \infty \text{ for some } \zeta\}$ is an open set and that for $\theta \in \Theta$, the transform kernel $\hat{P}_\theta$ defined in (2.4) has a maximal simple real eigenvalue $e^{\psi(\theta)}$, where $\psi(\theta)$ is the unique solution of the equation $w(\theta, \psi(\theta)) = 1$, with corresponding eigenfunction $r(x; \theta) = E_x \exp\{\theta' S_\tau - \psi(\theta)\tau\}$. Moreover, $\psi(\theta)$ is strictly convex and analytic on $\Theta$ and there exists a full set $F$ [i.e., $\varphi(F^c) = 0$] such that

$$w_x(\theta, \zeta) := E_x e^{\theta' S_\tau - \zeta\tau} < \infty \qquad \text{on } W \text{ for all } x \in F; \tag{2.12}$$

see the proof of Lemma 4.4 of Ney and Nummelin (1987), where it is shown that $r(x; \theta)$ is finite and analytic on $\Theta$ for all $x \in F$ as a consequence of (2.12). Define $\theta_\mu$, $I(\mu)$ and $V(\mu)$ by (2.2) for $\mu \in \Gamma$, and $Q_\theta$ by (2.5) for $\theta \in \Theta$. A minorization condition also holds for the transition kernel $Q_\theta$ whose associated Markov chain $\{X_n : n \geq 0\}$ has an invariant measure which will be denoted by $\pi_\theta$. The main result of this section is the following theorem, in which we define

$$K(u, \varepsilon) = \{v \in \mathbf{R}^d : u_i \leq v_i \leq u_i + \varepsilon \text{ for } 1 \leq i \leq d\} \qquad \text{for } u \in \mathbf{R}^d \text{ and } \varepsilon > 0. \tag{2.13}$$

THEOREM 2. *Assume* (2.9) *or* (2.10) *and define a probability measure* $\widetilde{\nu}$ *on* $\mathbf{R}^d$ *by*

$$\widetilde{\nu}(B) = \int_{\mathcal{X}} h(x, B)\, d\nu(x) \Big/ \int_{\mathcal{X}} h(x, \mathbf{R}^d)\, d\nu(x) \qquad [\textit{if } (2.9) \textit{ holds}]$$
$$= \int_{\mathcal{X}\times B} h(x)\, d\nu(x, s) \Big/ \int_{\mathcal{X}\times \mathbf{R}^d} h(x)\, d\nu(x, s) \qquad [\textit{if } (2.10) \textit{ holds}].$$

*Suppose* $\widetilde{\nu}$ *is nonlattice and* (2.11) *holds. Let* $x \in F$, *where* $F$ *is a full set satisfying* (2.12). *Then there exist positive numbers* $\varepsilon_n$ *with* $\lim_{n\to\infty} \varepsilon_n = 0$ *such that as* $n \to \infty$ *and* $\varepsilon \to 0$ *with* $\varepsilon \ge \varepsilon_n$,

$$P_x\{S_n \in K(n\mu, \varepsilon),\ X_n \in A\}$$
$$= (\varepsilon/\sqrt{n})^d e^{-nI(\mu)} \{(2\pi)^{-d/2} |V(\mu)|^{-1/2} r(x; \theta_\mu)\}$$
$$\times \left\{ \int_A \big(r(y; \theta_\mu)\big)^{-1} d\pi_{\theta_\mu}(y) + o(1) \right\},$$

*uniformly for* $\mu \in C$, *where* $C$ *is a compact subset of* $\Gamma$ *and* $A$ *is a measurable subset of* $\mathcal{X}$ *such that* $\inf_{\mu\in C, y\in A} r(y; \theta_\mu) > 0$.

The proof of Theorem 2 is given in Section 6. The minorization condition (2.9) or (2.10) in Theorem 2 is used not only to invoke the Ney–Nummelin theory on the eigenvalue $e^{\psi(\theta)}$ and eigenfunction $r(x; \theta)$ that appear in the asymptotic formula, but also to prove local limit theorems for the tilted transition kernel via regeneration arguments. On the other hand, assumption (2.11) is used only to apply the Ney–Nummelin theory and to ensure that the regeneration times have finite moments of all orders under the tilted measure. In specific applications (see, e.g., Example 2 below), one derives the eigenvalue and eigenfunction directly and can establish finiteness of moments of the regeneration times directly without appealing to (2.11), so one can apply Theorem 2 even when (2.11) fails to hold or cannot be verified. Moreover, one can also specify the full set $F$ on which the eigenfunction is finite and analytic in $\theta$. In particular, when $X_n$ is uniformly recurrent, $F = \mathcal{X}$ and we can dispense with condition (2.11). As will be illustrated in the proofs of Theorems 3, 5 and 6, Theorem 2 enables us to approximate probabilities in the same way that a true saddlepoint density (2.7) does but without any density assumption on the additive component.

## 3. Finiteness of eigenmeasures, large deviation probabilities and the maxima of Markov random walks.

Suppose the minorization condition (2.9) or (2.10) holds. For a measurable subset $A$ of $\mathcal{X}$ and $x \in \mathcal{X}$, define

$$\ell(A; \theta) = E_\nu\left[ \sum_{n=0}^{\tau-1} e^{\theta' S_n - n\psi(\theta)} I_{\{X_n \in A\}} \right],$$
(3.1)
$$\ell_x(A; \theta) = E_x\left[ \sum_{n=0}^{\tau-1} e^{\theta' S_n - n\psi(\theta)} I_{\{X_n \in A\}} \right].$$

Then $\ell(\cdot;\theta)$ is the left eigenmeasure associated with the eigenvalue $e^{\psi(\theta)}$; see Ney and Nummelin (1987). The following result gives upper bounds for certain large deviation probabilities of Markov random walks in terms of $\ell(\cdot\,;\theta)$ and $\ell_x(\cdot\,;\theta)$.

LEMMA 1. *Let $A$ be a measurable subset of $\mathcal{X}$, $\theta\in\Theta$, $\mu\in\Gamma$ and $F$ be a full set satisfying* (2.12). *For $x\in F$, there exists a constant $K_\theta$ continuous in $\theta$ and possibly dependent on $x$ but not on $A$ such that for all $c>0$ and $n\ge 1$,*

$$\sum_{m=0}^{\infty} P_x\{\theta' S_m - m\psi(\theta) > c,\ X_m\in A\} \le e^{-c}\{K_\theta \ell(A;\theta) + \ell_x(A;\theta)\}, \tag{3.2}$$

$$P_x\{\theta_\mu'(S_n/n-\mu) > 0,\ X_n\in A\} \le e^{-nI(\mu)}\{K_\theta\ell(A;\theta_\mu)+\ell_x(A;\theta_\mu)\}. \tag{3.3}$$

PROOF. Let $E^{(\theta)}$ denote expectation under the kernel $Q_\theta$ in (2.5). Let $\tau_1=\tau$ and $\tau_m$ be the first (regeneration) time after $\tau_{m-1}$ to reach the atom of the split chain. Consider the renewal measure

$$U_\theta(t) = \sum_{i=1}^{\infty} Q_{\theta,x}\{\theta' S_{\tau_i} - \tau_i\psi(\theta)\le t\}. \tag{3.4}$$

Since the regeneration times $\tau_i$ divide the Markov chain into independent blocks, the random variables $\theta'(S_{\tau_i}-S_{\tau_{i-1}})-(\tau_i-\tau_{i-1})\psi(\theta)$, $i\ge 2$, are i.i.d. By Blackwell's renewal theorem,

$$U_\theta(t)-U_\theta(t-1)\to \{E_{\nu_\theta}^{(\theta)}[\theta' S_\tau - \tau\psi(\theta)]\}^{-1} \qquad \text{as } t\to\infty,$$

when $\theta' S_\tau - \tau\psi(\theta)$ is nonlattice under $Q_{\theta,\nu_\theta}$. An analogous result also holds in the lattice case; see Feller (1971). In either case, $a_\theta := \sup_{t\in\mathbf{R}}\{U_\theta(t)-U_\theta(t-1)\}<\infty$ and is continuous in $\theta$. Since $\int r(y;\theta)\nu(dy) = E_\nu e^{\theta' S_\tau - \tau\psi(\theta)} = 1$ [cf. Ney and Nummelin (1987)], it follows from (2.5) that

$$\sum_{i=1}^{\infty} P_x\{\theta' S_{\tau_i} - \tau_i\psi(\theta)\in ds\} = e^{-s} r(x;\theta)\sum_{i=1}^{\infty} Q_{x,\theta}\{\theta' S_{\tau_i}-\tau_i\psi(\theta)\in ds\}. \tag{3.5}$$

Decomposing $\sum_{m=0}^{\infty}$ as $\sum_{m=0}^{\tau_1-1}+\sum_{i=1}^{\infty}\sum_{m=\tau_i}^{\tau_{i+1}-1}$, we have

$$\begin{aligned}\sum_{m=0}^{\infty} & P_x\{\theta' S_m - m\psi(\theta) > c, X_m\in A\} \\ &= E_x\left[\sum_{m=0}^{\tau-1} I_{\{\theta' S_m - m\psi(\theta)>c, X_m\in A\}}\right] \\ &\quad + \int_{-\infty}^{\infty}\sum_{i=1}^{\infty} P_x\{\theta' S_{\tau_i}-\tau_i\psi(\theta)\in ds\} E_\nu\left[\sum_{m=0}^{\tau-1} I_{\{\theta' S_m - m\psi(\theta)>c-s, X_m\in A\}}\right],\end{aligned} \tag{3.6}$$

$$E_x\left[\sum_{m=0}^{\tau-1} I_{\{\theta' S_m - m\psi(\theta) > c, X_m \in A\}}\right] \tag{3.7}$$
$$\le e^{-c} E_x\left[\sum_{m=0}^{\tau-1} e^{\theta' S_m - m\psi(\theta)} I_{\{X_m \in A\}}\right] = e^{-c}\ell_x(A;\theta),$$

$$\int_{-\infty}^{\infty}\sum_{i=1}^{\infty} P_x\{\theta' S_{\tau_i} - \tau_i\psi(\theta) \in ds\} E_\nu\left[\sum_{m=0}^{\tau-1} I_{\{\theta' S_m - m\psi(\theta) > c-s, X_m \in A\}}\right]$$
$$= \int_{-\infty}^{\infty} e^{-s} r(x;\theta) U_\theta(ds) E_\nu\left[\sum_{m=0}^{\tau-1} I_{\{\theta' S_m - m\psi(\theta) > c-s, X_m \in A\}}\right]$$

[by (3.5) and (3.4)]

$$\le \sum_{t\in c+\mathbf{Z}} e^{-(t-1)} r(x;\theta)[U_\theta(t) - U_\theta(t-1)] \tag{3.8}$$
$$\times E_\nu\left[\sum_{m=0}^{\tau-1} I_{\{\theta' S_m - m\psi(\theta) > c-t, X_m \in A\}}\right]$$
$$\le a_\theta e^{-c+1} r(x;\theta) E_\nu\left[\sum_{m=0}^{\tau-1}\left(\sum_{w\in\mathbf{Z}, w<\theta' S_m - m\psi(\theta)} \int_{w-1}^{w} e^{y+1}\,dy\right) I_{\{X_m \in A\}}\right]$$

(setting $w = c - t$)

$$\le a_\theta e^{-c+2} r(x;\theta) E_\nu\left[\sum_{m=0}^{\tau-1} e^{\theta' S_m - m\psi(\theta)} I_{\{X_m \in A\}}\right]$$
$$= a_\theta e^{-c+2} r(x;\theta)\ell(A;\theta).$$

From (3.6)–(3.8), we obtain (3.2) with $K_\theta = a_\theta e^2 r(x;\theta)$. To show (3.3), simply consider the summand $m = n$ in (3.2) with $c = nI(\mu)$. $\square$

Lemma 1 enables us to perform truncation by restricting $X_n$ to sets $A$ on which the eigenfunction is uniformly positive so that the saddlepoint approximation in Theorem 2 can be applied, as we can then apply the bound (3.2) or (3.3) to analyze the case $X_n \in A^c$. We illustrate this idea in the following theorem, which Arndt (1980) and Höglund (1991) proved for the case of finite $\mathcal{X}$ by using other methods involving Markov renewal theory. Besides using Theorem 2 and Lemma 1, our proof uses a time-reversal argument, which we generalize from the i.i.d. case [see, e.g., Chan and Lai (2000)] to the Markovian setting. This generalization involves the dual (time-reversed) Markov random walk $\widetilde{S}_n(\mu)$ under $Q_{\theta_\mu}$, assuming that there exists a $\sigma$-finite measure $\nu^*$ on $\mathcal{X}$ such that

$$P(x, \cdot \times \mathbf{R}^d) \text{ is absolutely continuous with respect to } \nu^* \text{ for all } x \in \mathcal{X}. \tag{3.9}$$

Clearly (3.9) holds when $\mathcal{X}$ is finite, since we can take $\nu^*(A) = \sum_{x\in\mathcal{X}} P(x, A \times \mathbf{R}^d)$. Further discussion of the dual Markov random walk and the role of assumption (3.9) is given in Section 4.3. Define $\widetilde{\nu}$ as in Theorem 2 and

$$\gamma(y;\mu) = \int_0^\infty e^{-z} Q_{\theta_\mu,y}\left\{\min_{n\geq 1}[\theta'_\mu \widetilde{S}_n(\mu) - n\psi(\theta_\mu)] > z\right\} dz, \tag{3.10}$$

$$\gamma(A;\mu) = \int_A \bigl(r(y;\theta_\mu)\bigr)^{-1} \gamma(y;\theta_\mu)\, d\pi_{\theta_\mu}(y), \qquad \gamma(\mu) = \gamma(\mathcal{X};\mu). \tag{3.11}$$

If $\ell(\mathcal{X};\theta_\mu)$ is finite, then so is $\gamma(\mu)$ because $(r(y;\theta))^{-1}\, d\pi_\theta(y) = L_\theta \ell(dy;\theta)$ for some constant $L_\theta$ [cf. Ney and Nummelin (1987), page 581].

THEOREM 3. *Let* $d = 1$ *and* $E_\pi \xi_1 < 0$. *Assume* (2.9) *or* (2.10), (2.11) *and* (3.9). *Then there exists a unique* $\theta^* > 0$ *such that* $\psi(\theta^*) = 0$. *Let* $x$ *belongs to a full set* $F$ *satisfying* (2.12). *Suppose that* $\widetilde{\nu}$ *is nonlattice*, $\ell(\mathcal{X};\theta) < \infty$ *and* $\ell_x(\mathcal{X};\theta) < \infty$ *for all* $\theta$ *in some neighborhood of* $\theta^*$. *Let* $\mu^* = d\psi(\theta)/d\theta|_{\theta=\theta^*}$. *If* $\alpha^{-1} < \mu^* < \delta^{-1}$, *then*

$$P_x\left\{\max_{n\geq 0} S_n > c\right\} \sim P_x\left\{\max_{\delta c\leq n\leq \alpha c} S_n > c\right\}$$
$$\sim r(x;\theta^*)\bigl(I(\mu^*)\bigr)^{-1}\gamma(\mu^*)e^{-c\theta^*} \qquad \text{as } c\to\infty.$$

PROOF. Since $\psi(\theta^*) = 0$ and $\ell(\mathcal{X};\theta^*) + \ell_x(\mathcal{X};\theta^*) < \infty$, it follows from (3.2) with $A = \mathcal{X}$ that

$$\sum_{n=0}^\infty P_x\{S_n > c + c^{1/5}\} = \sum_{n=0}^\infty P_x\{\theta^* S_n > \theta^* c + \theta^* c^{1/5}\} = o\bigl(e^{-c\theta^*}\bigr). \tag{3.12}$$

Let $A_\omega = \{x : r(x;\theta^*) > \omega\}$. Since $A_\omega \uparrow \mathcal{X}$ as $\omega \downarrow 0$, it follows from (3.2) that for all $\eta > 0$, there exists $\omega > 0$ small enough such that

$$\sum_{n=0}^\infty P_x\{\theta^* S_n > \theta^* c, X_n \in A_\omega^c\} \leq \eta e^{-c\theta^*}. \tag{3.13}$$

Let $\widetilde{T}_c = \min\{n \geq \delta c : S_n > c\}$. Then by Theorem 2 and (3.12),

$$
\begin{aligned}
P_x\{\widetilde{T}_c \leq \alpha c, X_{\widetilde{T}_c} \in A_\omega\} \\
&= (1+o(1)) \\
&\quad \times \sum_{c/\mu^*-c^{3/5} \leq n \leq c/\mu^*+c^{3/5}} \int_{A_\omega} \sum_{z \in \varepsilon \mathbf{Z}, 0 \leq z \leq c^{1/5}} P_x\{S_n \in K(c+z,\varepsilon),\ X_n \in dy\} \\
&\quad \times P_y\Big\{\min_{m \geq 1} \widetilde{S}_m\big((c+z)/n\big) > z\Big\} + o\big(e^{-c\theta^*}\big), \\
&= (1+o(1)) \\
&\quad \times \sum_{c/\mu^*-c^{3/5} \leq n \leq c/\mu^*+c^{3/5}} \int_{A_\omega} \int_0^{c^{1/5}} (2\pi n)^{-1/2} e^{-nI((c+z)/n)} \\
&\qquad \times \big|V\big((c+z)/n\big)\big|^{-1/2} \left[\frac{r(x;\theta_{(c+z)/n})}{r(y;\theta_{(c+z)/n})}\right] \\
&\qquad \times P_y\Big\{\min_{m \geq 1} \widetilde{S}_m\big((c+z)/n\big) > z\Big\}\, dz\, d\pi_{\theta_{(c+z)/n}}(y) \\
&\quad + o\big(e^{-c\theta^*}\big).
\end{aligned} \tag{3.14}
$$

Since $I(\mu^*) = \theta^*\mu^*$, $dI(\mu)/d\mu|_{\mu=\mu^*} = \theta^*$ and $d^2I(\mu)/d\mu^2|_{\mu=\mu^*} = (V(\mu^*))^{-1}$, we have uniformly for $|n - c/\mu^*| \leq c^{3/5}$ and $0 \leq z \leq c^{1/5}$,

$$
\begin{aligned}
nI\big((c+z)/n\big) &= nI(\mu^*) + (c+z-n\mu^*)\theta^* + (c+z-n\mu^*)^2/\big(2nV(\mu^*)\big) + o(1) \\
&= (c+z)\theta^* + (c-n\mu^*)^2/\big(2nV(\mu^*)\big) + o(1),
\end{aligned}
$$

with $(c+z)/n \to \mu^*$. Therefore the double integral in the RHS of (3.14) is asymptotically equivalent to

$$
(\mu^*/2\pi c)^{1/2} e^{-c\theta^* - (c-n\mu^*)^2/2nV(\mu^*)} \big(V(\mu^*)\big)^{-1/2} r(x;\theta^*)\gamma(A_\omega;\mu^*)/\theta^*, \tag{3.15}
$$

recalling (3.10) and (3.11). Moreover, using a change of variables $w = (c - n\mu^*)/\sqrt{c/\mu^*}$, we obtain

$$
\begin{aligned}
&\sum_{c/\mu^*-c^{3/5} \leq n \leq c/\mu^*+c^{3/5}} e^{-(c-n\mu^*)^2/2nV(\mu^*)} \\
&\quad \sim c^{1/2}(\mu^*)^{-3/2} \int_{-\infty}^{\infty} e^{-w^2/2V(\mu^*)}\, dw, \\
&\quad = \big(2\pi c V(\mu^*)\big)^{1/2} (\mu^*)^{-3/2}.
\end{aligned} \tag{3.16}
$$

Since $I(\mu^*) = \theta^*\mu^*$, (3.13)–(3.16) yield the desired conclusion for $P_x\{\max_{\delta c \leq n \leq \alpha c} S_n > c\}$ by letting $\eta \to 0$ (and therefore $\omega \to 0$).

To prove the desired conclusion for $P_x\{\max_{n\geq 0} S_n > c\}$, it suffices to show that

$$P_x\left\{\max_{n<c/\mu^*-c^{3/5}} S_n > c\right\} + P_x\left\{\max_{n>c/\mu^*+c^{3/5}} S_n > c\right\} = o\bigl(e^{-c\theta^*}\bigr). \tag{3.17}$$

Let $\theta_1 < \theta^* < \theta_2$ be such that $\mu_1 = (1/\mu^* + c^{-2/5})^{-1}$ and $\mu_2 = (1/\mu^* - c^{-2/5})^{-1}$, where $\mu_i = d\psi(\theta)/d\theta|_{\theta=\theta_i}$. Then by Lemma 1,

$$\begin{aligned} &P_x\left\{\max_{n\leq c/\mu^*-c^{3/5}} S_n > c\right\} \\ &\quad\leq P_x\left\{\max_{n\leq c/\mu^*-c^{3/5}} \bigl(\theta_2 S_n - n\psi(\theta_2)\bigr) > c\bigl(\theta_2 - \psi(\theta_2)/\mu_2\bigr)\right\} \\ &\quad\leq e^{-c\{\theta_2-\psi(\theta_2)/\mu_2\}}\bigl\{K_{\theta_2}\ell(\mathcal{X};\theta_2) + \ell_x(\mathcal{X};\theta_2)\bigr\}. \end{aligned} \tag{3.18}$$

As shown by Ney and Nummelin [(1987), pages 579 and 580], if $\ell(A;\theta)$ [or $\ell_x(A;\theta)$] is finite on an open subset of $\Theta$, then it is an analytic function of $\theta$ on this open subset. Therefore by continuity, $K_\theta$ and $\ell(\mathcal{X},\theta), \ell_x(\mathcal{X},\theta)$ are bounded in some neighborhood of $\theta^*$. Let $h(\mu) = I(\mu)/\mu$. Then $dh(\mu)/d\mu = (\mu\theta_\mu - I(\mu))/\mu^2 = \psi(\theta_\mu)/\mu^2$, which is negative for $0 < \mu < \mu^*$ and positive for $\mu > \mu^*$, recalling that $\psi(\theta^*) = 0 = \psi(0)$ and $\theta_{\mu^*} = \theta^*$. Hence $h(\mu) \geq h(\mu^*) + L(\mu - \mu^*)^2$ for some $L > 0$ when $\mu$ is close to $\mu^*$. Therefore, for $c$ large enough,

$$\theta_2 - \psi(\theta_2)/\mu_2 = I(\mu_2)/\mu_2 \geq I(\mu^*)/\mu^* + L(\mu_2 - \mu^*)^2 \geq \theta^* + \widetilde{L}c^k$$

for some positive constants $k$ and $\widetilde{L}$, and therefore $P_x\{\max_{n\leq c/\mu^*-c^{3/5}} S_n > c\} = o(e^{-c\theta^*})$ by (3.18). A similar argument can be applied with $\theta_2, \mu_2$ replaced by $\theta_1, \mu_1$ to bound the other probability in (3.17). $\square$

We next establish in Theorem 4 finiteness of $\ell(\mathcal{X};\theta)$ and $\ell_y(\mathcal{X};\theta)$ under "drift conditions" of the type in Meyn and Tweedie (1993). Let $C$ be a measurable subset of $\mathcal{X}$ such that

$$\ell(C;\theta) < \infty \quad \text{and} \quad \ell_y(C;\theta) < \infty \qquad \text{for all } y \in \mathcal{X}. \tag{3.19}$$

Let $w : \mathcal{X} \to [1,\infty)$ be a measurable function such that for some $0 < \beta < 1$ and $L > 0$,

$$E_x\bigl[e^{\theta'\xi_1-\psi(\theta)}w(X_1)\bigr] \leq (1-\beta)w(x) \qquad \text{for all } x \notin C, \tag{W1}$$

$$\sup_{x\in C} E_x\bigl[e^{\theta'\xi_1-\psi(\theta)}w(X_1)\bigr] = L < \infty \quad \text{and} \quad \int w(x)\,d\nu(x) < \infty, \tag{W2}$$

where $d\nu(x) = d\nu(x,\mathbf{R}^d)$ when (2.10) holds.

THEOREM 4. *Assume* (2.9) *or* (2.10) *and* (2.11). *Let* $C$ *be a measurable subset of* $\mathcal{X}$ *satisfying* (3.19), *and* (W1) *and* (W2) *for some* $w:\mathcal{X}\to[1,\infty)$, $0<\beta<1$ *and* $L>0$. *Then* $\ell(\mathcal{X};\theta)<\infty$ *and* $\ell_y(\mathcal{X};\theta)<\infty$ *for all* $y\in\mathcal{X}$.

PROOF. Let $\sigma_C=\inf\{n\geq 0: X_n\in C\}$, $\tau_C=\inf\{n\geq 1: X_n\in C\}$. We first show that under (W1),

$$E_x\left[\sum_{n=0}^{\sigma_C} e^{\theta' S_n-n\psi(\theta)}w(X_n)\right]\leq w(x)/\beta \qquad \text{for all } x\in\mathcal{X}. \tag{3.20}$$

If $x\in C$, then $\sigma_C=0$ and (3.20) clearly holds. For $x\notin C$, it suffices to show that

$$E_x\big[e^{\theta' S_n-n\psi(\theta)}w(X_n);\sigma_C\geq n\big]\leq(1-\beta)^n w(x) \qquad \text{for all } n\geq 0, \tag{3.21}$$

where $E_x[Z;A]$ denotes $E_x(ZI_A)$. We can prove (3.21) by induction since by (W1),

$$\begin{aligned} &E_x\big[e^{\theta' S_{n+1}-(n+1)\psi(\theta)}w(X_{n+1});\sigma_C\geq n+1\big]\\ &\quad= E_x\Big[e^{\theta' S_n-n\psi(\theta)}E_{X_n}\{e^{\theta'\xi_1-\psi(\theta)}w(X_{n+1})\};\sigma_C\geq n+1\Big],\\ &\quad\leq E_x\big[e^{\theta' S_n-n\psi(\theta)}(1-\beta)w(X_n);\sigma_C\geq n\big]. \end{aligned}$$

If $\sigma_C<\tau$, let $T_1,\ldots,T_M$ denote the times of visits to $C$ before $\tau$. If $\sigma_C\geq\tau$, set $M=0$. Let $T_{M+1}$ be the time of the first visit to $C$ at or after time $\tau$. Then

$$\ell(C;\theta)=E_\nu\left[\sum_{n=0}^{\tau-1}e^{\theta' S_n-n\psi(\theta)};X_n\in C\right]=E_\nu\left(\sum_{m=1}^{M}e^{\theta' S_{T_m}-T_m\psi(\theta)}\right), \tag{3.22}$$

where $\sum_{m=1}^{M}=0$ if $M=0$. Moreover,

$$\begin{aligned} \ell(\mathcal{X};\theta)&\leq E_\nu\left(\sum_{n=0}^{\sigma_C}e^{\theta' S_n-n\psi(\theta)}\right)\\ &\qquad+E_\nu\left[\sum_{m=1}^{M}\big(e^{\theta' S_{T_m+1}-(T_m+1)\psi(\theta)}+\cdots+e^{\theta' S_{T_{m+1}}-T_{m+1}\psi(\theta)}\big)\right]\\ &=E_\nu\left(\sum_{n=0}^{\sigma_C}e^{\theta' S_n-n\psi(\theta)}\right)\\ &\qquad+E_\nu\left[\sum_{m=1}^{M}e^{\theta' S_{T_m}-T_m\psi(\theta)}E_{X_{T_m}}\left(\sum_{n=1}^{\tau_C}e^{\theta' S_n-n\psi(\theta)}\right)\right]. \end{aligned} \tag{3.23}$$

If $z \in C$, then by (3.20) and (W2),

$$
\begin{aligned}
E_z\left[\sum_{n=1}^{\tau_C} e^{\theta' S_n - n\psi(\theta)} w(X_n)\right]
&= E_z\left\{(e^{\theta'\xi_1-\psi(\theta)} E_{X_1}\left[\sum_{n=0}^{\sigma_C} e^{\theta' S_n - n\psi(\theta)} w(X_n)\right]\right\} \\
&\le E_z\left[e^{\theta'\xi_1-\psi(\theta)} w(X_1)\right]/\beta \le L/\beta.
\end{aligned} \tag{3.24}
$$

Substituting (3.20), (3.22) and (3.24) into (3.23) and noting that $w(X_n) \ge 1$ and $X_{T_m} \in C$, it follows that

$$
\begin{aligned}
\ell(\mathcal{X};\theta) &\le E_\nu\left[\sum_{m=1}^{M} e^{\theta' S_{T_m} - T_m\psi(\theta)}(L/\beta)\right] + \int w(x)\,d\nu(x)\Big/\beta \\
&= \left\{L\ell(C;\theta) + \int w(x)\,d\nu(x)\right\}\Big/\beta.
\end{aligned} \tag{3.25}
$$

We can bound $\ell_y(\mathcal{X};\theta)$ in a similar way, with $\ell(C;\theta)$ in (3.25) replaced by $\ell_y(C;\theta)$ and $\int w(x)\,d\nu(x)$ replaced by $w(y)$. □

EXAMPLE 1. If the uniform recurrence condition (2.3) holds and $\mathrm{Int}(\Theta) \neq \varnothing$, then the eigenfunction $r(x;\theta)$ is uniformly positive and bounded on $\mathcal{X}$ and $\ell(\mathcal{X};\theta) < \infty$ for all $\theta \in \Theta$, by Lemma 3.1 of Iscoe, Ney and Nummelin (1985). Moreover, (W1) and (W2) are satisfied with $w \equiv 1$ and $C = \mathcal{X}$, $L = 1$ for every $\theta \in \Theta$.

EXAMPLE 2. Consider the vector autoregressive model,

$$
X_{i+1} = HX_i + Z_{i+1}, \qquad \|H\| = \sup_{\|x\|=1} \|Hx\| < 1, \tag{3.26}
$$

where $Z_i$ are i.i.d. nondegenerate $d \times 1$ random vectors such that $\Lambda(t) := Ee^{t\|Z_1\|} < \infty$ for all $t > 0$ and $Z_1$ has an absolutely continuous component (with respect to Lebsegue measure $\lambda$) in the sense that $P(Z_1 \in A) \ge \int_A g(z)\,dz$ for some positive continuous function $g$. Suppose the conditional distribution of $\xi_n$ given $X_0, \dots, X_n$ has the form $F_{X_{n-1},X_n}$ such that for every $\theta \in \mathbf{R}^d$, there exists a positive constant $\rho_\theta$ for which

$$
\int e^{\theta' s}\,dF_{x,y}(s) \le \exp\{\rho_\theta(\|x\| + \|y\|)\} \qquad \forall x, y \in \mathbf{R}^d. \tag{3.27}
$$

Suppose furthurmore that for every compact subset $C$ of $\mathbf{R}^d$, there exists a finite measure $\nu_C$ with compact support $K_C$ such that

$$
\inf_{x,y\in C} F_{x,y}(B) \ge \nu_C(B) \qquad \text{for all } B \subset K_C. \tag{3.28}
$$

Let $C = \{\mu : \|\mu\| \le N\}$ and $w(x) = e^{\gamma\|x\|}$ for some $\gamma$ and $N$ to be specified later.

Since $g$ is positive and continuous, $\delta := \inf\{g(z - Hx) : x \in C \text{ and } Hx + z \in C\} > 0$. Since $P\{Hx + Z_1 \in dz\} \ge g(z - Hx)\,dz$, it then follows from (3.28) that for all $x \in \mathbf{R}^d$,

$$P_x\{(X_1, \xi_1) \in A \times B\} \ge \delta I_{\{x\in C\}}\lambda(A \cap C)\nu_C(B \cap K_C), \tag{3.29}$$

and therefore the minorization condition (2.9) holds with $h(x) = \delta\lambda(C)\nu_C(K_C) \times I_{\{x\in C\}}$. Moreover, by (3.27),

$$\begin{aligned} E_x\big[e^{\theta'\xi_1}w(X_1)\big] &\le E\exp\{\rho_\theta(\|x\| + \|Hx + Z_1\|) + \gamma\|Hx + Z_1\|\} \\ &\le \Lambda(\rho_\theta + \gamma)\exp\big\{[\rho_\theta(1 + \|H\|) + \gamma\|H\|]\|x\|\big\}. \end{aligned}$$

Since $\|H\| < 1$, we can choose $\gamma$ large enough so that $2\rho_\theta + \gamma\|H\| < \gamma$, and then (W1) is satisfied if $N$ is chosen large enough. Since $C$ is compact and $\lambda(\cdot \cap C)$ has support $C$, (W2) also holds for sufficiently large $L$.

For the special case $\xi_i = X_i$, $F_{x,y}$ is degenerate at $y$ and (3.27) holds trivially with $\rho_\theta = \|\theta\|$. Although (3.28) no longer holds, we still have in place of (3.29) the minorization condition,

$$P_x\{\xi_1 = X_1 \in A\} \ge \delta I_{\{x\in C\}}\lambda(A \cap C). \tag{3.30}$$

Let $J(\theta) = \log(Ee^{\theta' Z_1})$. Then $\psi(\theta) = J((I - H')^{-1}\theta)$ and $r(x;\theta) = \exp\{\theta'(I - H)^{-1}Hx\}$ satisfy

$$\int e^{-\psi(\theta)+\theta' s}P(x, dy \times ds)r(y;\theta)/r(x;\theta) = E\big(e^{-\psi(\theta)+\theta'(I-H)^{-1}Z_1}\big) = 1,$$

and it can be shown that $e^{\psi(\theta)}$ is indeed the maximal eigenvalue and that a scalar multiple of $r(x;\theta)$ is the eigenfunction in the Ney–Nummelin framework. Note that $r(x;\theta)$ is finite and analytic in $\theta$ (and $x$). Moreover, the regeneration time $\tau$ under the minorization condition (3.30) has a finite moment generating function in some neighborhood of the origin [cf. Meyn and Tweedie (1993), pages 364–370].

**4. Nonlinear boundary crossing probabilities for Markov random walks.** In this section, we first generalize Theorem 3 to nonlinear boundary crossing probabilities. Specifically, instead of the maximum of a one-dimensional Markov random walk $\max_{\delta c\le n\le\alpha c} S_n$, we now consider $\max_{\delta c\le n\le\alpha c} ng(S_n/n)$, where $S_n$ is a $d$-dimensional Markov random walk and $g : \bar{\Gamma} \to \mathbf{R}$ satisfies certain regularity conditions described below. Here and in the sequel we use the same notation and assumptions as those in Section 2.2. We next extend the method to analyze the boundary crossing probability

$$P_x\Big\{\max_{n-\alpha c\le k\le n-\delta c}(n - k)g\big((S_n - S_k)/(n - k)\big) > c \text{ for some } n \le \beta c\Big\} \tag{4.1}$$

as $c \to \infty$, which plays a key role in the derivation of the main results (1.3) and (1.4) in the last part of this section.

4.1. *Generalization of Theorem* 3 *to nonlinear functions of mean vectors.* In the case of i.i.d. $\xi_n$, asymptotic approximations to $P_x\{\max_{\delta c \le n \le \alpha c} ng(S_n/n) > c\}$ were recently developed in Chan and Lai (2000). Under certain assumptions, this probability is shown to be of the order $Ac^{q/2}e^{-c/r}$ as $c \to \infty$, where $r = \sup_{a^{-1} \le g(\mu) \le \delta^{-1}} g(\mu)/I(\mu)$, $q$ is the dimension of the submanifold of $\Gamma$ at which the preceding supremum is attained and $A$ is a constant that can be expressed as an integral over the manifold with respect to its volume element measure. We can extend this result to Markov random walks satisfying the minorization and nonlattice conditions of Theorem 3, for which we still have the saddlepoint approximation given in Theorem 2, analogous to the i.i.d. case considered in Chan and Lai (2000) under the following assumptions on $g$:

(A1) $g$ is continuous on $\Gamma$ and there exists $\varepsilon_0 > 0$ such that

$$\sup_{\alpha^{-1} < g(\mu) < \delta^{-1} + \varepsilon_0} g(\mu)/I(\mu) = r < \infty.$$

(A2) $\inf_{g(\mu) > \delta^{-1} + \varepsilon_0} I(\mu) > (\delta r)^{-1}$ and $\limsup_{\mu \to \partial\Gamma} g(\mu)/I(\mu) < r$, where $\partial\Gamma$ denotes the boundary of $\Gamma$.

(A3) $M_{\varepsilon,\alpha,\delta} := \{\mu : \alpha^{-1} < g(\mu) < \delta^{-1} + \varepsilon \text{ and } g(\mu)/I(\mu) = r\}$ is a smooth $q$-dimensional manifold for all $0 \le \varepsilon \le \varepsilon_0$, where $q \le d$.

(A4) $g$ is twice continuously differentiable in some neighborhood of $M_{\varepsilon_0,\alpha,\delta}$ and $\sigma(\{\mu : g(\mu) = \delta^{-1} \text{ and } g(\mu)/I(\mu) = r\}) = 0$, where $\sigma$ is the volume measure of $M_{\varepsilon_0,\alpha,\delta}$.

(A5) $\inf_{\mu \in M_{0,\alpha,\delta}} |\Pi'_\mu \nabla^2 \rho(\mu) \Pi_\mu| > 0$ with $\rho(\mu) = I(\mu) - g(\mu)/r$, where $\Pi_\mu$ denotes the $d \times (d-q)$ matrix whose column vectors form an orthonormal basis of the orthogonal complement $TM^\perp(\mu)$ of the tangent space $TM(\mu)$ of $M := M_{0,\alpha,\delta}$ at $\mu$. In the case $d = q$, we set $|\Pi'_\mu \nabla^2 \rho(\mu) \Pi_\mu| = 1$, and (A5) clearly holds under this convention.

THEOREM 5. *Let $\alpha > \delta > 0$. With the same notation and assumptions as in Theorem* 2, *suppose that $g$ satisfies* (A1)–(A5), *with $g(E_\pi \xi_1) < \alpha^{-1}$, and that $\ell(\mathcal{X}; \theta_\mu) < \infty$, $\ell_x(\mathcal{X}; \theta_\mu) < \infty$ for all $\mu \in D$, where $D$ is a compact neighborhood of $\{\mu : I(\mu) \le (\delta r)^{-1}\}$. Then*

$$\begin{aligned} P_x\Big\{\max_{\delta c \le n \le \alpha c} ng(S_n/n) > c\Big\} &\sim (c/2\pi r)^{q/2} e^{-c/r} \int_M r(x; \theta_\mu) \gamma(\mu) \big(I(\mu)\big)^{-(q/2+1)} \\ &\qquad \times |V(\mu)|^{-1/2} |\Pi'_\mu \nabla^2 \rho(\mu) \Pi_\mu|^{-1/2}\, d\sigma(\mu), \end{aligned}$$

*where $\gamma(\mu)$ is defined in* (3.11).

In view of (A2), $\{\mu : I(\mu) \le (\delta r)^{-1}\}$ is compact and therefore indeed has a compact neighborhood. Note that Theorem 3 is in fact a special case of Theorem 5 with $g(\mu) = \mu$ for which $q = 0$, $r = 1/\theta^*$, $M = \{\mu^*\}$ and $\nabla^2\rho(\mu) = d^2 I(\mu)/d\mu^2 = (V(\mu))^{-1}$. Theorem 5 is a generalization of Theorem 1 of Chan and Lai (2000) to the Markovian setting. Let $\widetilde{T}_c = \min\{n \ge \delta c : ng(S_n/n) > c\}$. It follows from Theorem 2 and arguments similar to those in Section 3 of Chan and Lai (2000) that

$$P_x\{\widetilde{T}_c \le \alpha c, X_{\tilde{T}_c} \in A_\omega\} = \bigl(1 + o(1)\bigr)(c/2\pi r)^{q/2} e^{-c/r} \int_M r(x;\theta_\mu)\gamma(A_\omega;\mu) I(\mu)^{-(q/2+1)} \times |V(\mu)|^{-1/2} |\Pi'_\mu \nabla^2\rho(\mu)\Pi_\mu|^{-1/2}\, d\sigma^q(\mu),$$

where $A_\omega = \{x : r(x;\theta_\mu) > \omega \text{ for all } \mu \in D\}$. The following lemma, which will be proved in Section 5 and which is a nonlinear analogue of Lemma 1, then allows us to prove Theorem 5 by letting $A = A_\omega^c$ and $\omega \to 0$ in the lemma.

LEMMA 2. *Under the same notation and assumptions as in Theorem* 5, *there exists a constant L such that for all measurable subsets A of* $\mathcal{X}$,

$$P_x\Bigl\{\max_{\delta c \le n \le \alpha c} ng(S_n/n) I_{\{X_n \in A\}} > c\Bigr\} \le L c^{q/2} e^{-c/r} \sup_{\mu \in D}[\ell(A;\theta_\mu) + \ell_x(A;\theta_\mu)]. \tag{4.2}$$

EXAMPLE 3. Let $S_n = X_1 + \cdots + X_n$, where $X_i$ is the autoregressive series (3.26). Assume the $Z_i$ to be normally distributed with mean 0 and covariance matrix $\Sigma$. Then $\psi(\theta) = \theta' V\theta/2$ and $\mu = \nabla\psi(\theta) = V\theta$, where $V = (I - H)^{-1} \times \Sigma(I - H')^{-1}$. Let $g(\mu) = \mu' V^{-1}\mu/2$. Since $g(\mu) = I(\mu)$, (A1)–(A5) hold with $r = 1$, $q = d$ and $M = \{\mu : \alpha^{-1} < I(\mu) < \delta^{-1}\}$. Hence by Theorem 5,

$$P_x\Bigl\{\max_{\delta c \le n \le \alpha c} S_n' V^{-1} S_n/2n > c\Bigr\} \sim (c/2\pi)^{d/2} e^{-c} |V|^{-1/2} \times \int_{\alpha^{-1} < I(\mu) < \delta^{-1}} e^{\mu'(I-H)'\Sigma^{-1} Hx} \gamma(\mu)\bigl(I(\mu)\bigr)^{-(d/2+1)}\, d\mu.$$

4.2. *Overview of the method to analyze the boundary crossing probability* (4.1). We now proceed to analyze the boundary crossing probability (4.1) which, unlike that in Theorem 5, involves two time indices $i$ and $j$. To fix the ideas, we first assume that the $\xi_i$ are i.i.d. with a common density function (with respect to

Lebesgue measure) that is continuous and bounded on $\mathbf{R}^d$. Letting $S_{n,k} = S_n - S_k$ for $n \geq k$, (2.1) gives the saddlepoint approximation

$$P\{S_{j,i}/(j-i) \in d\mu\} = \bigl(1+o(1)\bigr)\bigl((j-i)/2\pi\bigr)^{d/2}|V(\mu)|^{-1/2}e^{-(j-i)I(\mu)}\,d\mu \tag{4.3}$$

as $j - i \to \infty$, where the $o(1)$ term is uniform over compact subsets of $\Gamma$. Let

$$B_c = \{(i,j) : 0 \leq i < j \leq \beta c, \delta c \leq j - i \leq \alpha c\}. \tag{4.4}$$

Define an ordering $\prec$ in $B_c$ by

$$(k,n) \prec (i,j) \Leftrightarrow \text{ either (i) } n < j \text{ or (ii) } n = j \text{ and } k < i. \tag{4.5}$$

The boundary crossing probability (4.1) can be expressed as

$$\sum_{(i,j)\in B_c} P\bigl\{(j-i)g\bigl(S_{j,i}/(j-i)\bigr) > c, \quad (n-k)g\bigl(S_{n,k}/(n-k)\bigr) \leq c \;\; \forall\,(k,n) \prec (i,j)\bigr\}. \tag{4.6}$$

Replacing $g$ by $g/r$, we shall assume without loss of generality that $r = 1$. Let $t = [c^{1/4}]$, $B_{c,t} = \{(i,j) : t \leq i < j \leq \beta c, \delta c + 2t \leq j - i \leq \alpha c - t\}$. Define

$$\begin{aligned} f_{i,j}(\mu)\,d\mu = {} & P\{S_{j,i}/(j-i) \in d\mu\} I_{\{(j-i)g(\mu)>c\}} \\ & \times P\biggl\{(n-k)g\biggl(\frac{S_{n,k}}{n-k}\biggr) \leq c\,\forall\,(k,n) \prec (i,j) \text{ with} \\ & \qquad \max(j-n, |i-k|) \leq t \biggm| \frac{S_{j,i}}{j-i} = \mu\biggr\}. \end{aligned} \tag{4.7}$$

Large deviation bounds can be used to express (4.6) as

$$\int_{\mathbf{R}^d} \sum_{(i,j)\in B_{c,t}} f_{i,j}(\mu)\,d\mu + o\bigl(c^{q/2+1}e^{-c}\bigr). \tag{4.8}$$

Simply denote $M_{\varepsilon,\alpha,\delta}$ by $M_\varepsilon$ for notational simplicity in this subsection. To evaluate the integral in (4.8), we use a Laplace-type asymptotic formula

$$\int_{\mathbf{R}^d} \sum_{(i,j)\in B_{c,t}} f_{i,j}(\mu)\,d\mu \sim \int_{U_{c^{-1/2}\log c,\, c^{-1/2}}} \sum_{(i,j)\in B_{c,t}} f_{i,j}(\mu)\,d\mu, \tag{4.9}$$

where $U_{\eta,\varepsilon}$ is a tubular neighborhood of $M_\varepsilon$ with radius $\eta$. We call

$$U_{\eta,\varepsilon} = \bigl\{u + v : u \in M_\varepsilon,\ v \in TM_\varepsilon^{\perp}(u) \text{ and } \|v\| \leq \eta\bigr\} \tag{4.10}$$

a *tubular neighborhood* of $M_\varepsilon$ with radius $\eta$ if the representation of the elements of $U_{\eta,\varepsilon}$ in (4.10) is unique. For the existence of tubular neighborhoods when $\eta$ is sufficiently small, see Theorem 5.1 in Chapter 4 of Hirsch (1976) and its proof.

Integrals over tubes can be evaluated by the so-called "infinitesimal change of volume function" [cf. Gray (1990)]. Note that (4.9) is an extension of Laplace's method to approximate an integral whose integrand attains its maximum on a manifold, instead of at a single point in the classical Laplace approximation [cf. Jensen (1995), pages 57–62].

From Lemmas 3.13, 3.14 and Theorem 3.15 of Gray (1990), it follows that as $c \to \infty$,

$$\begin{aligned}&\int_{U_{c^{-1/2}\log c, c^{-1/2}}} \sum_{(i,j)\in B_{c,t}} f_{i,j}(\mu)\, d\mu \\ &\quad \sim \int_M \left\{ \int_{v\in TM^{\perp}(u), \|v\|\le c^{-1/2}\log c} \sum_{(i,j)\in B_{c,t}} f_{i,j}(u+v)\, dv \right\} d\sigma(u).\end{aligned} \tag{4.11}$$

To analyze the inner integral in (4.11), use (4.7) and note that given $S_{j,i}/(j-i) = \mu$, $S_{n,k} = (j-i)\mu + (S_i - S_k) - S_{j,n}$ for $(k,n) \prec (i,j)$. Since

$$\max_{j-n\le t, |i-k|\le t} \{(S_i - S_k)^2 + S_{j,n}^2\}/c \xrightarrow{P} 0,$$

Taylor's expansion yields

$$\begin{aligned}&(n-k)g\big(\{(j-i)\mu + (S_i - S_k) - S_{j,n}\}/(n-k)\big) \\ &\qquad = (j-i)g(\mu) \\ &\qquad\quad - \big\{(\nabla g(\mu))'(S_{j,n} - S_i + S_k) - (j-n-i+k)\big(\mu'\nabla g(\mu) - g(\mu)\big)\big\} \\ &\qquad\quad + \delta_{i,j,k,n}(\mu),\end{aligned} \tag{4.12}$$

where $\max_{(i,j)\in B_{c,t}, j-n\le t, |i-k|\le t} |\delta_{i,j,k,n}(\mu)| \xrightarrow{P} 0$ uniformly in $\mu \in U_{c^{-1/2}\log c, c^{-1/2}}$. The uniformity follows from (A4) and the compactness of

$$M^* = \{\mu : \alpha^{-1} \le g(\mu) \le \delta^{-1} + \varepsilon^*, g(\mu)/I(\mu) = r\} \tag{4.13}$$

for sufficiently small $\varepsilon^* > 0$, which follows from (A2). Note that $\nabla I(\mu) = \theta_\mu$, $\nabla^2 I(\mu) = V^{-1}(\mu)$. For $\mu \in M$, $g(\mu) = I(\mu)$ and $\nabla(I - g) = 0$ since $I - g$ attains on $M$ its minimum value 0 over $\{\mu : \alpha^{-1} < g(\mu) < \delta^{-1} + \varepsilon_0\}$, and therefore $\mu'\nabla g(\mu) - g(\mu) = \psi(\theta_\mu)$. Let $S_n(\mu) = \sum_{k=1}^n \{\theta_\mu' \xi_k - \psi(\theta_\mu)\}$. From (4.12) it follows that uniformly for $(i,j) \in B_{c,t}$ and $\mu \in U_{c^{-1/2}\log c, c^{-1/2}}$ with $(j-i)g(\mu) > c$, the conditional probability in (4.7) is equal to

$$\begin{aligned}&P\{S_{j,n}(\mu) - S_i(\mu) + S_k(\mu) > (j-i)g(\mu) - c \ \forall (k,n) \prec (i,j) \text{ with} \\ &\qquad\qquad \max(j-n, |i-k|) \le t \,|\, S_{j,i}/(j-i) = \mu\} + o(1) \\ &\quad = P\{S_{j,n}(\mu) + S_{k,i}(\mu) > (j-i)g(\mu) - c \\ &\qquad\quad \text{for all } i \le k \le i+t \text{ and } j-t \le n < j, \\ &\qquad\quad \text{and } -S_{i,k}(\mu) > (j-i)g(\mu) - c \\ &\qquad\quad \text{for all } i-t \le k < i \,|\, S_{j,i}/(j-i) = \mu\} + o(1),\end{aligned} \tag{4.14}$$

noting that in the quantification "$\forall(k,n) \prec (i,j)$" the indices $(k,n)$ with $k < i$ and $j > n$ are redundant because the required inequality holds for these $(k,n)$ if it holds for $k < i$, $j = n$ and for $k = i$, $j \geq n$; see the proof of Lemma 4 of Siegmund (1988). Let $Q_\theta$ denote the probability measure under which $\xi_1, \xi_1^*, \xi_2, \xi_2^*, \dots$ are i.i.d. with $Q_\theta\{\xi_i \in dx\} = e^{\theta' x - \psi(\theta)} P\{\xi_i \in dx\}$. By Siegmund's (1988) Lemma 4, the second probability in (4.14) is equal to

$$\begin{aligned} &P\Big\{\max_{m\geq 1} S_m(\mu) < -(j-i)g(\mu) + c\Big\} \\ &\quad \times Q_{\theta_\mu}\Big\{(\min_{m\geq 0} S_m(\mu) + \min_{n\geq 1} S_n^*(\mu) > (j-i)g(\mu) - c\Big\} + o(1), \end{aligned} \tag{4.15}$$

uniformly for $(i,j) \in B_{c,t}$ and $\mu \in U_{c^{-1/2}\log c, c^{-1/2}}$ with $(j-i)g(\mu) - c > 0$, where $S_n^*(\mu) = \sum_{k=1}^n \{\theta_\mu' \xi_k^* - \psi(\theta_\mu)\}$. Let

$$p(\mu; w) = P\Big\{\max_{m\geq 1} S_m(\mu) < -w\Big\} Q_{\theta_\mu}\Big\{\min_{m\geq 0} S_m(\mu) + \min_{n\geq 1} S_n^*(\mu) > w\Big\}. \tag{4.16}$$

Note that $\sum_{(i,j)\in B_{c,t}} = \sum_{\delta c + 2t \leq m \leq \alpha c - t} \sum_{t \leq i < j \leq \beta c, j-i=m}$ and that there are $[\beta c] - m - t + 1$ terms in the inner sum. Putting (4.3), (4.14) and (4.15) into (4.7) shows that the integral in (4.9) is equal to

$$\begin{aligned} &\{(2\pi)^{-d/2} + o(1)\} \\ &\quad \times \int_{U_{c^{-1/2}\log c, c^{-1/2}}} \sum_{\delta c + 2t \leq m \leq \alpha c - t, mg(\mu) > c} (\beta c - m) \\ &\qquad\qquad \times m^{d/2} |V(\mu)|^{-1/2} e^{-mI(\mu)} p\big(\mu; mg(\mu) - c\big)\, d\mu \\ &\quad = \{(2\pi)^{-d/2} + o(1)\} e^{-c} \\ &\qquad \times \int_{U_{c^{-1/2}\log c, c^{-1/2}}} \Big\{\int_0^\infty \Big(\beta c - \frac{c}{g(\mu)}\Big)\Big(\frac{c}{g(\mu)}\Big)^{d/2} \\ &\qquad\qquad \times |V(\mu)|^{-1/2} e^{c - (w+c)I(\mu)/g(\mu)} p(\mu; w) \frac{dw}{g(\mu)}\Big\}\, d\mu, \end{aligned} \tag{4.17}$$

using the change of variables $w = mg(\mu) - c$ to replace the sum by the integral and noting that the range of the sum can be restricted to $c/g(\mu) < m < c/g(\mu) + (\log c)^2$ because of the exponential decay in $e^{-mI(\mu)}$ [so $m \sim c/g(\mu)$ in this range]. Since $M^*$ defined in (4.13) is a compact subset of $\Gamma$ and $g = I$ on $M^* \supset M_{\varepsilon^*}$, it follows that uniformly in $\mu \in U_{c^{-1/2}\log c, c^{-1/2}}$, $g(\mu) = I(\mu) + o(1)$ and therefore $\int_0^\infty p(\mu; w) e^{-wI(\mu)/g(\mu)}\, dw = \int_0^\infty e^{-w} p(\mu; w)\, dw + o(1)$. Moreover, since $\nabla I = \nabla g$ on $M$, Taylor's expansion around $u \in M$ yields

$$\{I(u+v) - g(u+v)\}/g(u+v) = v' \nabla^2 \rho(u) v / 2I(u) + o(c^{-1}) \tag{4.18}$$

uniformly in $\mu = u + v \in U_{c^{-1/2}\log c, c^{-1/2}}$ [see (4.10)]. Let $\gamma(\mu) = \int_0^\infty e^{-w} \times p(\mu; w)\, dw$. Using the change of variables $v = \Pi_u z$ with $z \in \mathbf{R}^{d-q}$ and applying (4.11) and (4.18), we can express (4.17) as

$$\begin{aligned}
&\{(2\pi)^{-d/2} + o(1)\} e^{-c} c^{d/2} \\
&\quad \times \int_M (\beta c - c/I(\mu))(I(\mu))^{-d/2-1} |V(\mu)|^{-1/2} \gamma(\mu) \\
&\qquad \times \int_{z \in \mathbf{R}^{d-q}, \|z\| \le c^{-1/2}\log c} \exp\{-cz'\Pi_u' \nabla^2 \rho(u) \Pi_u z / 2I(u)\}\, dz\, d\sigma(u) \\
&\quad \sim e^{-c} c^{q/2} (\beta \zeta_{\alpha,\delta}^{(1)} - \zeta_{\alpha,\delta}^{(2)}) c,
\end{aligned} \tag{4.19}$$

where for $j = 1, 2$,

$$\begin{aligned}
\zeta_{\alpha,\delta}^{(j)} &= (2\pi)^{-q/2} \\
&\quad \times \int_{M_{0,\alpha,\delta}} (I(u))^{-q/2-j} |V(u)|^{-1/2} |\Pi_u' \nabla^2 \rho(u) \Pi_u|^{-1/2} \gamma(u)\, d\sigma(u).
\end{aligned} \tag{4.20}$$

In view of (4.6), (4.8) and (4.9), this shows that the boundary crossing probability (4.1) is asymptotically equivalent to $e^{-c} c^{q/2+1} (\beta \zeta_{\alpha,\delta}^{(1)} - \zeta_{\alpha,\delta}^{(2)})$ when $r = 1$ and the $\xi_i$ are i.i.d. with a bounded continuous density.

When $\{(X_i, \xi_i), i \ge 0\}$ is a Markov chain satisfying the assumptions of Theorem 1(i), we can replace (4.3) by the saddlepoint approximation (2.7) for $P_{X_i}\{S_{j-i}/(j-i) \in d\mu, X_{j-i} \in dy\}$. The conditional probability in (4.7) now has the form

$$\begin{aligned}
&P_x\{(n-k) g(S_{n,k}/(n-k)) \le c\, \forall (k,n) \prec (i,j) \text{ with} \\
&\qquad \max(j-n, |i-k|) \le t \,|\, S_{j,i}/(j-i) = \mu,\ X_i = \tilde{y},\ X_j = y\}.
\end{aligned} \tag{4.21}$$

We still have the Taylor expansion (4.12) which shows that (4.21) is equal to

$$\begin{aligned}
&P_x\{S_{j,n}(\mu) + S_{k,i}(\mu) > (j-i) g(\mu) - c \\
&\qquad \text{for all } i \le k \le i + t \text{ and} j - t \le n < j, \\
&\qquad \text{and } - S_{i,k}(\mu) > (j-i) g(\mu) - c \\
&\qquad \text{for all } i - t \le k < i \,|\, S_{j,i}/(j-i) = \mu, X_i = \tilde{y}, X_j = y\} + o(1),
\end{aligned} \tag{4.22}$$

analogous to (4.14). This conditional probability has a limit (as $c \to \infty$), which involves two time-reversed (dual) Markov additive processes and another tilted process so that the three processes are independent, as in (4.15). We can also replace the assumptions of Theorem 1(i) by the considerably weaker nonlattice assumption in Theorem 2, by using an analogue of Lemma 2, partitioning $\Gamma$ into cubes and replacing "$\in d\mu$" in (4.3) and (4.7) by "$\in C_\mu$," where $C_\mu$ denotes a cube centered at $\mu$. Letting the common length of the cubes approach 0, summation over these cubes can be approximated by integration with respect to $d\mu$; see Section 5 for details.

4.3. *Main results on* (4.1), (1.3) *and* (1.4). Let $P(A|x) = P(x, A \times \mathbf{R}^d)$ and assume that there exists a $\sigma$-finite measure $\nu^*$ on $\mathcal{X}$ such that (3.9) holds. Under (2.9) or (2.10) and (3.9), $Q_\theta(\cdot|x)$ is absolutely continuous with respect to $\nu^*$, where $Q_\theta(A|x) = Q_\theta(x, A \times \mathbf{R}^d)$. Let $\pi^*$ (or $\pi_\theta^*$) denote the density function, with respect to $\nu^*$, of the stationary distribution of $X_n$ under $P$ (or $Q_\theta$). For fixed $\mu \in \Gamma$, define three independent Markov additive processes $\{(X_n^{(j)}, S_n^{(j)}),\ n \geq 1\}$ on $\mathcal{X} \times \mathbf{R}^d$ initialized as follows: $S_0^{(1)} = S_0^{(2)} = S_0^{(3)} = 0$, $X_0^{(1)} = X_0^{(2)}$ has density function $\pi^*$ and is independent of $X_0^{(3)}$ which has density function $\pi_{\theta_\mu}^*$ (with respect to $\nu^*$). The transition function of $(X_n^{(j)}, S_n^{(j)})$ is absolutely continuous with respect to $\nu^*$, with density function $p^{(j)}$ given by time reversal of the density $p$ of $P$ for $j = 1$, or $q_{\theta_\mu}$ of $Q_{\theta_\mu}$ for $j = 3$; that is,

$$p^{(1)}(y, x; B) = p(x, y; B)\pi^*(x)/\pi^*(y),$$

$$p^{(3)}(y, x; B) = q_{\theta_\mu}(x, y; B)\pi_{\theta_\mu}^*(x)/\pi_{\theta_\mu}^*(y).$$

The transition density $p^{(2)}$ of $(X_n^{(2)}, S_n^{(2)})$ with respect to $\nu^*$ is $q_{\theta_\mu}(x, y; B)$. Define $S_n^{(j)}(\mu) = \sum_{k=1}^n \{\theta_\mu' \xi_k^{(j)} - \psi(\theta_\mu)\}$ and

$$\begin{aligned} &p(\mu; w) \\ &\quad = E\biggl[\frac{r(X_0^{(1)}; \theta_\mu)}{r(X_0^{(3)}; \theta_\mu)} I_{\{\max_{m\geq 1} S_m^{(1)}(\mu) < -w, \min_{m\geq 0} S_m^{(2)}(\mu) + \min_{m\geq 1} S_m^{(3)}(\mu) > w\}}\biggr], \end{aligned} \tag{4.23}$$

which is a generalization of (4.16) from the i.i.d. case to the present Markovian setting.

THEOREM 6. *Let* $0 < \delta < \alpha < \beta$. *With the same notation and assumptions as in Theorem* 2, *suppose that* $g$ *satisfies* (A1)–(A5) *with* $g(E_\pi \xi_1) < \alpha^{-1}$, *and that* $\ell(\mathcal{X}; \theta_\mu) < \infty$ *and* $\int \ell_x(\mathcal{X}; \theta_\mu)\, d\pi(x) < \infty$ *for all* $\mu \in D$, *where* $D$ *is a compact neighborhood of* $\{\mu : I(\mu) \leq (\delta r)^{-1}\}$. *Let* $\gamma(\mu) = \int_0^\infty e^{-w} p(\mu; w)\, dw$, *where* $p(\mu; w)$ *is defined in* (4.23), *and define* $\zeta_{\alpha,\delta}^{(1)}, \zeta_{\alpha,\delta}^{(2)}$ *by* (4.20). *Then*

$$\begin{aligned} &P_\pi\Bigl\{\max_{n-\alpha c\leq k\leq n-\delta c} (n-k) g\bigl(S_{n,k}/(n-k)\bigr) > c \text{ for some } n \leq \beta c\Bigr\} \\ &\qquad \sim (c/r)^{q/2} c e^{-c/r} \bigl(\beta \zeta_{\alpha,\delta}^{(1)} - \zeta_{\alpha,\delta}^{(2)}/r\bigr). \end{aligned}$$

The details of the proof, which follows the steps outlined in Section 4.2, are given in Section 5, where we also make use of Theorem 6 to prove (1.3) and (1.4) in the following.

THEOREM 7. *With the same notation and assumptions as in Theorem* 6, *define* $T_c$ *by* (1.2) *with* $J(c) = \{j : j_1 \le j \le j_2\}$ *such that* $j_1 \sim \delta c$ *and* $j_2 \sim \alpha c$. *Then* $\zeta_{\alpha,\delta}^{(1)}(c/r)^{q/2}e^{-c/r}T_c$ *has a limiting exponential distribution with mean* 1 *as* $c \to \infty$. *Moreover, if* $J_n = \{j : n_1 \le j \le n_2\}$ *with* $n_1 \sim \delta r \log n$ *and* $n_2 \sim \alpha r \log n$ *in* (1.1), *then*

$$P_\pi\left\{\mathcal{M}_n \le r\left(\log n + \frac{q}{2}\log\log n\right) + t\right\} \to \exp\big(-\zeta_{\alpha,\delta}^{(1)} e^{-t/r}\big) \tag{4.24}$$

*as* $n \to \infty$, *uniformly in* $t \in \mathbf{R}$.

We now show how Theorem 7 follows from Theorem 6 in the case where the $\xi_i$ are i.i.d. The proof in the Markovian setting of the theorem is considerably more complicated and is given in the next section. Let $x > 0$, $m = (\zeta_{\alpha,\delta}^{(1)})^{-1}(c/r)^{-q/2}e^{c/r}x$, and partition the interval $[0, m]$ into $K \sim m/(\beta c)$ disjoint intervals $I_1, \dots, I_K$ with equal length $\beta c(1 + o(1))$. Then the events

$$A_j = \left\{\max_{n-\alpha c \le k \le n-\delta c}(n-k)g\big(S_{n,k}/(n-k)\big) > c \text{ for some } n \in I_j \text{ and } n-k \in I_j\right\}$$

are independent and have the same probability $p_c$. Letting $\bar{A}_j$ denote the complement of $A_j$, it follows from Theorem 6 that

$$P\left(\bigcap_{j=1}^K \bar{A}_j\right) = (1-p_c)^K \sim e^{-Kp_c} \to \exp\{-x[1 - (\beta\zeta_{\alpha,\delta}^{(1)})^{-1}\zeta_{\alpha,\delta}^{(2)}/r]\}$$

as $c \to \infty$. Hence $P(\bigcup_{j=1}^K A_j) \to 1 - \exp\{-x[1 - (\beta\zeta_{\alpha,\delta}^{(1)})^{-1}\zeta_{\alpha,\delta}^{(2)}/r]\}$. Let

$$B_j = \left\{\max_{n-\alpha c \le k \le n-\delta c}(n-k)g\big(S_{n,k}/(n-k)\big) > c \text{ for some } n \in I_{j+1} \text{ and } n-k \in I_j\right\}.$$

Then by a similar argument involving a straightforward modification of Theorem 6, it can be shown that $P(\bigcup_{j=1}^K B_j) \to 1 - \exp[-(\beta\zeta_{\alpha,\delta}^{(1)})^{-1}\zeta_{\alpha,\delta}^{(2)}x/r]$ as $c \to \infty$ for sufficiently large $\beta$ (with $\beta > \alpha$). Taking $\beta$ arbitrarily large, since

$$P\left(\bigcup_{j=1}^K A_j\right) \le P\{T_c \le m\} \le P\left(\bigcup_{j=1}^K A_j\right) + P\left(\bigcup_{j=1}^K B_j\right),$$

it then follows that $P\{T_c \le m\} \to 1 - e^{-x}$ as $c \to \infty$. The corresponding result (4.24) for $\mathcal{M}_n$ can then be derived from that for $T_c$ since $P\{T_c \le n\} = P\{\mathcal{M}_n > c\}$; see the last paragraph of Section 5 for details.

**5. Proof of Lemma 2 and Theorems 6 and 7.** In this section we first prove Lemma 2. The first step of the proof is to show that we can restrict to the event $\{S_n/n \in D\}$ by showing that

$$P_x\left\{\max_{\delta c\le n\le \alpha c} ng(S_n/n)I_{\{S_n/n\in D^c\}} > c\right\} = o(e^{-c/r}). \tag{5.1}$$

When $d = 1$, this is an easy consequence of (3.3). Since $D$ is a compact neighborhood of $\{\mu : I(\mu) \le (\delta r)^{-1}\}$, there exist $a < b$ such that $[a, b] \subset D$ and $I(a) > (\delta r)^{-1}$, $I(b) > (\delta r)^{-1}$. Then

$$\sum_{\delta c\le n\le\alpha c} P_x\{S_n/n \in D^c\} \le \sum_{\delta c\le n\le\alpha c} \big[P_x\{S_n/n < a\} + P_x\{S_n/n > b\}\big]$$
$$\le \sum_{n\ge\delta c} \big(K_1e^{-nI(a)} + K_2e^{-nI(b)}\big)$$

for some constants $K_1$ and $K_2$, by (3.3). Since $I(a)$ and $I(b)$ exceed $(\delta r)^{-1}$, (5.1) follows.

For $d > 1$, we can replace $a$ and $b$ by a finite number of hyperplanes, as will be shown below. The next step is to cover the compact set $D$ with $O(c^{d/2})$ cubes of the form $K(\mu, c^{-1/2})$ in (2.13) and to apply (3.3) when $S_n/n$ is restricted to each of these cubes. The advantage of using these cubes is that $g$ is well approximated by $g(\mu)$ on $K(\mu, c^{-1/2})$. Summing (3.3) over these cubes then completes the proof. We also use similar cubes to prove Theorem 6, which is then applied to prove Theorem 7. Indeed the saddlepoint approximation in Theorem 2 is phrased in term of these cubes. Whereas Theorem 2 considers the saddlepoint approximation under the initial state $x$ that belongs to a full set $F$, we can easily extend the result to general initial distributions $\sigma$ that are absolutely continuous with respect to $\varphi$ and such that $\int r(x;\theta_\mu)\,d\sigma(x) < \infty$. Specifically, for such initial distributions, the saddlepoint approximation in Theorem 2 has the form

$$\begin{aligned} &P_\sigma\{S_n \in K(n\mu, \varepsilon), X_n \in A\} \\ &\quad= (\varepsilon/\sqrt{n})^d e^{-nI(\mu)}\left\{(2\pi)^{-d/2}|V(\mu)|^{-1/2}\int r(x;\theta_\mu)\,d\sigma(x)\right\} \\ &\qquad\times\left\{\int_A \big(r(y;\theta_\mu)\big)^{-1}\,d\pi_{\theta_\mu}(y) + o(1)\right\}. \end{aligned} \tag{5.2}$$

We shall use these ideas in the proof of Theorem 6.

PROOF OF LEMMA 2. Let $\Omega = \{\mu \in D : g(\mu) \ge \alpha^{-1}\}$ and $J_c = \{\mu \in c^{-1/2}\mathbf{Z}^d : K(\mu, c^{-1/2}) \cap \Omega \ne \varnothing\}$. Since $h_\mu(\theta) := \theta'\mu - \psi(\theta)$ is maximized at $\theta_\mu$ and $h_\mu(\theta_\mu) = I(\mu)$, it follows from the compactness of $D$ that there exist positive constants $L$ and $\widetilde{L}$ such that $h_\mu(\theta_\omega) \ge I(\mu) - Lc^{-1}$ for all $\mu \in J_c$ and

$\omega \in K(\mu, c^{-1/2})$, noting that $\|\theta_\mu - \theta_\omega\| \le \widetilde{L} c^{-1/2}$ for $\omega \in K(\mu, c^{-1/2})$ and $\mu \in J_c$. Therefore, if $S_n/n \in K(\mu, c^{-1/2})$, then $\theta'_\mu S_n - n\psi(\theta_\mu) \ge n(I(S_n/n) - Lc^{-1})$. Hence

$$
\begin{aligned}
P_x\Big\{ \max_{\delta c \le n \le \alpha c} n g(S_n/n) I_{\{X_n \in A, S_n/n \in \Omega\}} > c \Big\} \\
\le \sum_{\mu \in J_c} P_x\big\{\theta'_\mu S_n - n\psi(\theta_\mu) > \big(cI(S_n/n) - L\big)/g(S_n/n), \\
X_n \in A, S_n/n \in K(\mu, c^{-1/2}) \text{ for some } \delta c \le n \le \alpha c\big\}, \\
\le \sum_{\mu \in J_c} e^{-c(I(\mu)/g(\mu)) + M}\big[K_{\theta_\mu} \ell(A;\theta_\mu) + \ell_x(A;\theta_\mu)\big]
\end{aligned}
\tag{5.3}
$$

for some constant $M$ independent of $\mu \in J_c$. Approximating the sum by an integral and using the tube integration techniques of Section 4.2, it can be shown that

$$
\sum_{\mu \in J_c} e^{-c(I(\mu)/g(\mu))} \sim \int_{\Omega_c} c^{d/2} e^{-c(I(\mu)/g(\mu))}\, d\mu = O(c^{q/2} e^{-c/r}).
\tag{5.4}
$$

From (5.1), (5.3) and (5.4), (4.2) follows.

To prove (5.1) for general $d$, note that by compactness, there exists $\varepsilon > 0$ such that $\{\mu : I(\mu) \le (\delta r)^{-1} + 3\varepsilon\} \subset D$. Let $G = \{\mu : I(\mu) = (\delta r)^{-1} + \varepsilon\}$. The tangent space $TG(\mu)$ of $G$ at $\mu$ is a hyperplane that is orthogonal to $\nabla I(\mu) (= \theta_\mu)$, and therefore $TG(\mu) = \{y : \theta'_\mu(y - \mu) = 0\}$. Let $B(\mu) = \{y : |\theta'_\mu(y - \mu)| < \varepsilon\}$. Then $\{B(\mu) : \mu \in G\}$ is an open cover of the compact set $G$ and therefore there exists a finite subcover $\{B(\mu_i) : 1 \le i \le k\}$. Let $H_i = \{y : \theta'_{\mu_i}(y - \mu_i) > 0\}$. Since $\theta_{\mu_i} = \nabla I(\mu_i)$, it then follows that if $\varepsilon$ is chosen sufficiently small, then $D^c$ [on which $I(\cdot)$ exceeds $(\delta r)^{-1} + 3\varepsilon$] is contained in the union of the half spaces $H_1, \dots, H_k$. Therefore

$$
\begin{aligned}
\sum_{\delta c \le n \le \alpha c} P_x\{S_n/n \in D^c\} &\le \sum_{\delta c \le n \le \alpha c} \sum_{i=1}^{k} P\{S_n/n \in H_i\} \\
&\le \sum_{n \ge \delta c} \sum_{i=1}^{k} P_x\{\theta'_{\mu_i}(S_n/n - \mu_i) > 0\} = o(e^{-c/r})
\end{aligned}
$$

by (3.3), recalling that $I(\mu_i) = (\delta r)^{-1} + \varepsilon$. □

PROOF OF THEOREM 6. The basic ideas of the proof have been given in Section 4.2. Here we provide some of the details, using the same notation as that in Section 4.2 and still assuming that $r = 1$. Let $K_u = K(u, \varepsilon_{[\alpha c]})$, where $\varepsilon_n$ is given in Theorem 2. The analogue of (4.7) for the present Markov case is

$$
\begin{aligned}
\bar{f}_{i,j;x,\tilde{y}}(y,\mu)\, d\nu^*(y)\, d\nu^*(\tilde{y}) \\
= P_x\{X_i \in d\tilde{y}\} P_{\tilde{y}}\big\{S_{j-i} \in K_{(j-i)\mu},\ X_{j-i} \in dy\big\} \times [(4.21)],
\end{aligned}
\tag{5.5}
$$

where "$S_{j,i}/(j-i)=\mu$" in (4.21) is replaced by "$S_{j,i}\in K_{(j-i)\mu}$." Recall the time-reversed Markov additive processes $(X_n^{(1)}, S_n^{(1)})$ and $(X_n^{(3)}, S_n^{(3)})$ and the (forward) Markov additive process $(X_n^{(2)}, S_n^{(2)})$. Analogous to (4.15) for the i.i.d. case, the conditional probability in (4.22) is equal to

$$\begin{aligned}P_x\Big\{&\min_{m\ge 1} S_m^{(3)}(\mu)+\min_{m\ge 0} S_m^{(2)} > (j-i)g(\mu)-c,\\ &\max_{m\ge 1} S_m^{(1)} < -(j-i)g(\mu)+c\Big| X_0^{(3)}=y,\ X_0^{(1)}=X_0^{(2)}=\tilde{y}\Big\}+o(1),\end{aligned}\tag{5.6}$$

for $y,\tilde{y}\in\mathcal{X}$, $\delta c+t\le j-i\le\alpha c$ and $\mu\in U_{c^{-1/2}\log c, c^{-1/2}}$. Let

$$p_\eta(\mu;w) = E\left[\frac{r(X_0^{(1)};\theta_\mu)}{r(X_0^{(3)};\theta_\mu)} I_{\{\max_{m\ge 1} S_m^{(1)}(\mu)<-w, \min_{m\ge 0} S_m^{(2)}(\mu)+\min_{m\ge 1} S_m^{(3)}(\mu)>w\}} I_{\{X_0^{(3)}\in A_\eta\}}\right],$$

where $A_\eta=\{x\in\mathcal{X}: r(x;\theta_\mu)>\eta \text{ for all } \mu\in D\}$. Then $p_\eta(\mu;w)\to p(\mu;w)$ as $\eta\to 0$, where $p(\mu;w)$ is defined as in (4.23).

Letting $w=(j-i)g(\mu)-c$ and multiplying the conditional probability in (5.6) by $r(\tilde{y};\theta_\mu)/r(y;\theta_\mu)$, integration with respect to $\pi^*(\tilde{y})\pi^*_{\theta_\mu}(y)\,d\nu^*(\tilde{y})\,d\nu^*(y)$ over $(y,\tilde{y})$ with $y\in A_\eta$ yields $p_\eta(\mu;w)$. The factor $r(\tilde{y};\theta_\mu)/r(y;\theta_\mu)$ above comes from the saddlepoint approximation in Theorem 2 applied to $P_{\tilde{y}}\{S_{j-i}\in K_{(j-i)\mu}, X_{j-i}\in dy\}$ with $y\in A_\eta$. We can sum up these approximations of (5.5) over $\mu\in\varepsilon_{[\alpha c]}\mathbf{Z}^d$ and $(i,j)\in B_{c,t}$ such that $\inf_{s\in K_{(j-i)\mu}}(j-i)g(s/(j-i))>c$ (for a lower bound), or such that $\sup_{s\in K_{(j-i)\mu}}(j-i)g(s/(j-i))>c$ (for an upper bound). Using the fact that

$$\sup_{s\in K_{(j-i)\mu}}(j-i)\big|g\big(s/(j-i)\big)-g(\mu)\big| = O(\varepsilon_{[\alpha c]}),$$

and replacing the sum $\varepsilon_{[\alpha c]}^d\sum_{\mu\in\varepsilon_{[\alpha c]}\mathbf{Z}^d}$ by $\int_{\mathbf{R}^d}d\mu$ as $\varepsilon_{[\alpha c]}\to 0$, we can use arguments similar to those of Chan and Lai [(2000), page 1652] to show that

$$\begin{aligned}\sum_{(i,j)\in B_c} P_\pi\big\{&(j-i)g\big(S_{i,j}/(j-i)\big)>c,\ X_j\in A_\eta \text{ and}\\ &(n-k)g\big(S_{n,k}/(n-k)\big)\le c\ \forall\,(k,n)\prec(i,j)\big\}\end{aligned}$$

is of the order $c^{q/2+1}e^{-c}(\beta\zeta_{\alpha,\delta}^{(1)}-\zeta_{\alpha,\delta}^{(2)})$ but with $p_\eta(\mu;w)$ replacing $p(\mu;w)$ in the definition of $\gamma(\mu)$ that appears in the right-hand side of (4.20) defining $\zeta_{\alpha,\delta}^{(j)}$. By Lemma 2,

$$P_\pi\left\{\max_{\delta c\le n\le\alpha c} ng(S_n/n)I_{\{X_n\in A_\eta^c\}}>c\right\}\le\delta_\eta c^{q/2}e^{-c},$$

where $\delta_\eta\to 0$ as $\eta\to 0$. Theorem 6 then follows by letting $\eta\to 0$. □

PROOF OF THEOREM 7. By (2.11), there exists $\kappa > 1$ such that $E_\nu \kappa^\tau < \infty$. Hence by Proposition 15.1.2 and Theorem 15.1.5 of Meyn and Tweedie (1993), there exist $A > 0$, $0 < \rho < 1$, a full set $F$ and a measurable function $V : \mathcal{X} \to [0, \infty)$ such that $\int V(x)\, d\pi(x) < \infty$ and

$$\|P^n(x, \cdot) - \pi(\cdot)\| \le A\rho^n V(x) \qquad \text{for all } x \in F \text{ and } n \ge 1, \tag{5.7}$$

where $\|\omega\|$ denotes the total variation norm of a signed measure $\omega$.

Let $h > 2\Delta > 0$ and $b(n) = r(\log n + \frac{q}{2} \log\log n)$. Partition the interval $[0, n]$ into subintervals with alternating lengths $(h - \Delta + o(1))b(n)$ and $(\Delta + o(1))b(n)$, so that there are $K \sim n/(hb(n))$ "long" subintervals labelled as $J_1, \dots, J_K$ and $K$ "short" subintervals labelled as $I_1, \dots, I_K$. Let $c = b(n) + o(1)$ so that $n = [e^{c/r}(c/r)^{-q/2}]$. Let $\eta > 0$, $A_\eta = \{x : V(x) \le \eta\}$, and

$$C_k = \big\{(j - i)g\big(S_{j,i}/(j - i)\big)I_{\{(X_i, X_j) \in A_\eta^2\}} \le b(n) + t \qquad \text{for all } i < j \text{ and } i, j \in J_k\big\}, \qquad k = 1, \dots, K.$$

Making use of (5.7), it will be shown that if $\Delta r \log(1/\rho) > 1$, then

$$\left| P_\pi(C_1 \cap \cdots \cap C_K) - \prod_{k=1}^{K} P_\pi(C_k) \right| \to 0 \qquad \text{as } n \to \infty. \tag{5.8}$$

Moreover, by arguments similar to the proof of Theorem 6 in which $\beta$ is replaced by $h - \Delta$, it can be shown that as $n \to \infty$,

$$\prod_{k=1}^{K} P_\pi(C_k) \to \exp\big\{-\big[(1 - \Delta/h)\zeta^{(1)}_{\alpha,\delta,\eta} - h^{-1}\zeta^{(2)}_{\alpha,\delta,\eta}/r\big]e^{-t/r}\big\}, \tag{5.9}$$

in which $\zeta^{(\ell)}_{\alpha,\delta,\eta} \to \zeta^{(\ell)}_{\alpha,\delta}$ as $\eta \to \infty$ for $\ell = 1, 2$, and that

$$\begin{aligned} &\max_{1 \le k \le K} P_\pi\left\{ \max_{i,j \in J_k, i<j} (j - i)g\big(S_{j,i}/(j - i)\big)I_{\{(X_i, X_j) \notin A_\eta^2\}} > b(n) + t \right\} \\ &\qquad \le \big(B_1 + o(1)\big)h\varepsilon(\eta)c^{(q/2)+1}e^{-(c+t)/r}, \end{aligned} \tag{5.10}$$

$$\begin{aligned} &\max_{1 \le k \le K} P_\pi\big\{(j - i)g\big(S_{j,i}/(j - i)\big) > b(n) + t \text{ for some } i \in J_k \cup I_k,\ j \notin J_k\big\} \\ &\qquad \le \big(B_2 + o(1)\big)\Delta c^{(q/2)+1}e^{-(c+t)/r} \end{aligned} \tag{5.11}$$

uniformly in $t \in \mathbf{R}$, where $B_1$ and $B_2$ are constants and $\varepsilon(\eta) \to 0$ as $\eta \to \infty$. Recalling that $K \sim n/(hb(n))$ and $(c/r)^{q/2}ce^{-c/r} \sim n^{-1}b(n)$, the desired conclusion for $\mathcal{M}_n$ follows from (5.8)–(5.11) by taking $h, h/\Delta$ and $\eta$ arbitrarily large.

To prove (5.8), first note that

$$\begin{aligned} P_\pi(C_k) &= P_\pi(C_k | C_1 \cap \cdots \cap C_{k-1})P_\pi(C_1 \cap \cdots \cap C_{k-1}) \\ &\quad + P_\pi\big(C_k | (C_1 \cap \cdots C_{k-1})^c\big)P_\pi\big((C_1 \cap \cdots \cap C_{k-1})^c\big). \end{aligned} \tag{5.12}$$

Since there is a distance of at least $(\Delta + o(1))b(n)$ between $J_k$ and the subintervals $J_1, \dots, J_{k-1}$, it follows from (5.7) and $\sup_{x \in A_\eta} V(x) \le \eta$ that

$$\big|P_\pi\big(C_k|(C_1 \cap \cdots \cap C_{k-1})^c\big) - P_\pi(C_k)\big| \le A\eta\rho^{(\Delta+o(1))b(n)}.$$

Combining this with (5.12) yields

$$\begin{aligned}\big|P_\pi(C_k|C_1 \cap \cdots \cap C_{k-1}) - P_\pi(C_k)\big|P_\pi(C_1 \cap \cdots \cap C_{k-1})\\ \le A\eta\rho^{(\Delta+o(1))b(n)}P_\pi\big((C_1 \cap \cdots \cap C_{k-1})^c\big).\end{aligned}$$

Therefore $|P_\pi(C_1 \cap \cdots \cap C_k) - P_\pi(C_k)P_\pi(C_1 \cap \cdots \cap C_{k-1})| \le A\eta\rho^{(\Delta+o(1))b(n)}$ for $1 \le k \le K$. This implies by an induction argument that

$$|P_\pi(C_1 \cap \cdots \cap C_K) - P_\pi(C_1) \cdots P_\pi(C_K)| \le AK\eta\rho^{(\Delta+o(1))b(n)} = o(1)$$

if $1 + r\Delta \log \rho < 0$.

Given $c > 0$ and $t > 0$, define $n(t, c) = [te^{c/r}(c/r)^{-q/2}/\zeta^{(1)}_{\alpha,\delta}]$. Note that $P_\pi\{T_c \le n(t, c)\} = P_\pi\{\mathcal{M}_{n(t,c)} > c\}$ and that $c = b(n(t, c)) - r \log(t/\zeta^{(1)}_{\alpha,\delta}) + o(1)$. Since $P_\pi\{\mathcal{M}_{n(t,c)} > c\} = P_\pi\{\mathcal{M}_{n(t,c)} > b(n(t, c)) - r \log(t/\zeta^{(1)}_{\alpha,\delta}) + o(1)\} \to 1 - \exp(-t)$, it then follows that $P_\pi\{T_c \le te^{c/r}(c/r)^{-q/2}/\zeta^{(1)}_{\alpha,\delta}\} \to 1 - e^{-t}$ as $c \to \infty$. □

**6. Proof of Theorem 2.** As we consider in Theorem 2 Markovian rather than independent increments of $S_n$, we cannot express the characteristic function of $S_n$ as a product of $n$ characteristic functions. We introduce instead an additional variable $v$ in Lemma 4 to capture the relationship between $n$ and the regeneration times $\tau_m$. This leads to an identity (6.5) from which the characteristic function (6.6) of $S_n$ is derived by Fourier inversion in $v$. We shall assume the minorization condition (2.9) in the proof of Theorem 2, as the proof under (2.10) is similar. The tilted kernel $Q_\theta$ defined by (2.5) then satisfies a similar minorization condition,

$$Q^\kappa_\theta(x, A \times B) \ge h_\theta(x, B)\nu_\theta(A), \tag{6.1}$$

where $h_\theta(x, B) = [\int_B e^{\theta' s - \kappa\psi(\theta)} h(x, ds)]/r(x; \theta)$ and $\nu_\theta(dy) = r(y; \theta)\nu(dy)$. We preface the proof of Theorem 2 by the following two lemmas, the first of which is the same as Lemma 3.3 of Ney and Nummelin (1987) but with (6.1) in place of the original minorization condition (2.9).

LEMMA 3. *Let* $\theta = \theta_\mu$ *and* $E^{(\theta)}$ *denote expectation under the kernel* $Q_\theta$ *in* (2.5). *Then*

$$\mu = E^{(\theta)}_{\nu_\theta} S_\tau / E^{(\theta)}_{\nu_\theta}\tau, \qquad \nabla^2\psi(\theta_\mu) = E^{(\theta)}_{\nu_\theta}\{(S_\tau - \tau\mu)(S_\tau - \tau\mu)'\}/E^{(\theta)}_{\nu_\theta}\tau. \tag{6.2}$$

LEMMA 4. *Let* $a(t,v)=E_x^{(\theta)}e^{it'S_\tau+iv\tau}$, $b(t,v)=E_{\nu_\theta}^{(\theta)}e^{it'S_\tau+iv\tau}$, $c(t,v)=E_{\nu_\theta}^{(\theta)}[\sum_{n=0}^{\tau-1}e^{it'S_n+ivn}g(X_n)]$ *and* $d(t,v)=E_x^{(\theta)}[\sum_{n=0}^{\tau-1}e^{it'S_n+ivn}g(X_n)]$, *where* $g$ *is a bounded measurable function on* $\mathcal{X}$. *Then*:

(i) $d(t,v)=\sum_{n=0}^{\infty}E_x^{(\theta)}(e^{it'S_n}g(X_n)I_{\{n<\tau_1\}})e^{inv}$,

(ii) $a(t,v)b^m(t,v)c(t,v)=\sum_{n=0}^{\infty}E_x^{(\theta)}(e^{it'S_n}g(X_n)I_{\{\tau_{m+1}\le n<\tau_{m+2}\}})e^{inv}$ *for* $m\ge 0$,

*where* $\tau_1=\tau$ *and* $\tau_m$ *is the first* (*regeneration*) *time after* $\tau_{m-1}$ *to reach the atom of the split chain.*

PROOF. (i) follows immediately from the definition of $d(t,v)$. We next show that

$$a(t,v)b^m(t,v)=\sum_{n=0}^{\infty}E_x^{(\theta)}\big(e^{it'S_n}I_{\{\tau_{m+1}=n\}}\big)e^{inv}\qquad\text{for } m\ge 0. \tag{6.3}$$

The case $m=0$ follows from the definition of $a(t,v)$. Noting that

$$\begin{aligned}
&\sum_{n=0}^{\infty}E_x^{(\theta)}\big(e^{it'S_n}I_{\{\tau_{m+1}=n\}}\big)e^{inv}b(t,v)\\
&\quad=\sum_{n=1}^{\infty}E_x^{(\theta)}\big(e^{it'S_n}I_{\{\tau_{m+1}=n\}}\big)e^{inv}\sum_{k=1}^{\infty}E_{\nu_\theta}^{(\theta)}\big(e^{it'S_k}I_{\{\tau=k\}}\big)e^{ikv}\\
&\quad=\sum_{\ell=0}^{\infty}e^{i\ell v}\sum_{n+k=\ell}E_x^{(\theta)}\big(e^{it'S_n}I_{\{\tau_{m+1}=n\}}\big)E_{\nu_\theta}^{(\theta)}\big(e^{it'S_k}I_{\{\tau=k\}}\big)\\
&\quad=\sum_{\ell=0}^{\infty}e^{i\ell v}E_x^{(\theta)}\big(e^{it'S_\ell}I_{\{\tau_{m+2}=\ell\}}\big),
\end{aligned}$$

we obtain (6.3) by induction. Part (ii) follows from (6.3) since

$$\begin{aligned}
&a(t,v)b^m(t,v)c(t,v)\\
&\quad=\sum_{n=0}^{\infty}E_x^{(\theta)}\big(e^{it'S_n}I_{\{\tau_{m+1}=n\}}\big)e^{inv}\sum_{k=0}^{\infty}E_{\nu_\theta}^{(\theta)}\big(e^{it'S_k}g(X_k)I_{\{k<\tau\}}\big)e^{ikv}\\
&\quad=\sum_{\ell=0}^{\infty}e^{i\ell v}\sum_{n+k=\ell}E_x^{(\theta)}\big(e^{it'S_n}I_{\{\tau_{m+1}=n\}}\big)E_{\nu_\theta}^{(\theta)}\big(e^{it'S_k}g(X_k)I_{\{k<\tau\}}\big).
\end{aligned}$$

□

PROOF OF THEOREM 2. Let $g:\mathcal{X}\to[0,\infty)$ be a bounded measurable function and define

$$q_{n,\mu}(t)=E_x^{(\theta_\mu)}e^{it'S_n}g(X_n)=\int e^{it's}g(y)Q_{\theta_\mu,x}\{X_n\in dy,\ S_n\in ds\}. \tag{6.4}$$

By Lemma 4 with $\theta = \theta_\mu$,

$$
\begin{aligned}
d(t,v) + \sum_{m=0}^{\infty} a(t,v) b^m(t,v) c(t,v) &= \sum_{n=0}^{\infty} E_x^{(\theta)}[e^{it'S_n} g(X_n)] e^{inv} \\
&= \sum_{n=0}^{\infty} q_{n,\mu}(t) e^{inv}.
\end{aligned} \tag{6.5}
$$

Since $\int_{-\pi}^{\pi} e^{-i\ell v}\, dv = 0$ unless $\ell = 0$, it follows from (6.5) that

$$
\begin{aligned}
q_{n,\mu}(t) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inv} d(t,v)\, dv \\
&\quad + \sum_{m=0}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inv} a(t,v) b^m(t,v) c(t,v)\, dv.
\end{aligned} \tag{6.6}
$$

Let $Z_\alpha$ be a random vector independent of $\{(X_n, S_n) : n \ge 0\}$ and having probability density function $\alpha^{-d} K(z/2\alpha)$ with respect to Lebesgue measure, where

$$
K(z) = (2\pi)^{-d} \big( (\sin z_1)^2/z_1^2, \ldots, (\sin z_d)^2/z_d^2 \big) \qquad \text{for } z = (z_1, \ldots, z_d).
$$

The characteristic function of $Z_\alpha$ is $k(\alpha t)$, where $k(u) = \prod_{j=1}^d (1 - |u_j|)$ if $\|u\| \le 1$ and $k(u) = 0$ otherwise; see Stone (1965). By the Fourier inversion formula,

$$
\begin{aligned}
&\int g(y) Q_{\theta_\mu, x}\{X_n \in dy, S_n + \sqrt{n} Z_\alpha \in K(n\mu + \sqrt{n} s, \omega\sqrt{n})\} \\
&\quad = \left(\frac{\omega}{2\pi}\right)^d \int_{\|\alpha t\| \le 1} e^{-is't} k(\alpha t) \left\{ \prod_{j=1}^d \frac{1 - e^{-i\omega t_j}}{i\omega t_j} \right\} e^{-i\sqrt{n}\mu' t} q_{n,\mu}\left(\frac{t}{\sqrt{n}}\right) dt
\end{aligned} \tag{6.7}
$$

[cf. Stone (1965), page 548]. Let $\phi_\mu$ denote the $d$-variate normal density function with mean 0 and covariance matrix $V(\mu)$ and let $\widehat{\phi}_\mu(t) = \int e^{it's} \phi_\mu(s)\, ds$. We shall show that

$$
e^{-i\sqrt{n}\mu' t} q_{n,\mu}(t/\sqrt{n}) = \big(1 + o(1)\big) \widehat{\phi}_\mu(t) \int g(y)\, d\pi_{\theta_\mu}(y) + o(n^{-d}) \tag{6.8}
$$

uniformly in $\|t\| \le n^{1/10}$ and $\mu \in C$, and that for any $\delta > 0$,

$$
\sup_{n^{-2/5} \le \|t\| \le \delta^{-1}, \mu \in C} |q_{n,\mu}(t)| = o(n^{-d}). \tag{6.9}
$$

Since $|1 - e^{-iu}| \le \min(|u|, 2)$ and $(1 - e^{-iu})/(iu) \to 1$ as $u \to 0$ and since $\phi_\mu(s) = (2\pi)^{-d} \int_{\|t\| \le n^{1/10}} e^{-is't} \widehat{\phi}_\mu(t)\, dt + o(1)$ uniformly in $s$ and in $\mu \in C$, it follows from (6.7)–(6.9) that as $n \to \infty$, $\omega \to 0$ and $\alpha \to 0$ such that $\alpha \geq \delta n^{-1/2}$,

$$\int g(y) Q_{\theta_\mu,x}\{X_n \in dy, S_n + \sqrt{n} Z_\alpha \in K(n\mu + \sqrt{n}s, \omega\sqrt{n})\} = \omega^d \left\{ \phi_\mu(s) \int g(y)\, d\pi_{\theta_\mu}(y) + o(1) \right\} \tag{6.10}$$

uniformly in $s$ and in $\mu \in C$. Making use of (6.10) and an argument similar to the proof of Lemma 2 of Stone (1965), it can be shown that for any $\delta > 0$ and $\eta > 0$, there exist $n_0$ and $\omega_0$ such that for all $n \geq n_0$, $\delta n^{-1/2} \leq \omega \leq \omega_0$, $\mu \in C$ and $s \in \mathbf{R}^d$,

$$\begin{aligned} &\omega^d \left\{ \phi_\mu(s) \int g(y)\, d\pi_{\theta_\mu}(y) - \eta \right\} \\ &\quad \leq \int g(y) Q_{\theta_\mu,x}\{X_n \in dy, S_n \in K(n\mu + \sqrt{n}s, \omega\sqrt{n})\} \\ &\quad \leq \omega^d \left\{ \phi_\mu(s) \int g(y)\, d\pi_{\theta_\mu}(y) + \eta \right\}. \end{aligned} \tag{6.11}$$

Let $\eta_k = 1/k = \delta_k$. Then there exist $n_k$ and $\omega_k$ such that (6.11) is satisfied for all $n \geq n_k$, $\delta_k n^{-1/2} \leq \omega \leq \omega_k$ and $s \in \mathbf{R}^d$. Without loss of generality, we can assume that $n_k$ is nondecreasing and $\omega_k \sqrt{n_k} > \delta_k$. For $n_k \leq n < n_{k+1}$, set $\varepsilon_n = \delta_k$. It then follows from (6.11) with $s = 0$ that as $n \to \infty$ and $\varepsilon \to 0$ such that $\varepsilon \geq \varepsilon_n$,

$$\int g(y) Q_{\theta_\mu,x}\{X_n \in dy, S_n \in K(n\mu, \varepsilon)\} = (\varepsilon/\sqrt{n})^d \phi_\mu(0) \left\{ \int g(y)\, d\pi_{\theta_\mu}(y) + o(1) \right\} \tag{6.12}$$

uniformly for $\mu \in C$. The desired conclusion then follows from (2.5) and (6.12) with $g(y) = r(x; \theta_\mu) I_{\{y \in A\}} / r(y; \theta_\mu)$.

It remains to prove (6.8) and (6.9). To simplify the notation, we shall write $Q_{\theta_\mu,x}$ simply as $Q_x$, and use $E_x^Q$ to denote expectation under $Q_x$, $E$ to denote $E_{\nu_\theta}^{(\theta)}$ with $\theta = \theta_\mu$ and $Q$ to denote the corresponding probability measure. Let $k_1 = [(n - n^{2/3})/E\tau]$, $k_2 = [(n + n^{2/3})/E\tau]$, where $[\cdot]$ denotes the greatest integer function. Because $W$ is an open set by (2.11), it follows from (2.12) that for every $r \geq 1$, there exists a constant $B_r$ for which $E_x^Q \tau^r$ and $E\tau^r$ are bounded by $B_r$. In view of the compactness of $C$, the bound $B_r$ can be chosen independent of $\mu \in C$. Noting that $\tau_m - \tau_1$ is a sum of $(m-1)$ i.i.d. random variables with finite $r$th moment $E\tau^r$, we can then apply Markov's inequality to show that

$$\begin{aligned} Q_x\{\tau_1 \geq \sqrt{n}\} &= o(n^{-d}), \\ Q_x\{\tau_{k_1} \geq n\} &= o(n^{-d}), \\ Q_x\{\tau_{k_2} \leq n\} &= o(n^{-d}) \end{aligned} \tag{6.13}$$

uniformly in $\mu \in C$. Since $\int_{-\pi}^{\pi} e^{-inv} d(t,v)\,dv = 2\pi E_x^Q(e^{it'S_n+inv} g(X_n) I_{\{n<\tau_1\}})$ by Lemma 4(i), it follows from (6.6) and (6.13) that uniformly for $\mu \in C$,

$$q_{n,\mu}(t) = \sum_{k_1 \le m \le k_2} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inv} a(t,v) b^m(t,v) c(t,v)\,dv + o(n^{-d}). \tag{6.14}$$

We shall assume $\kappa = 1$. If $\kappa > 1$, we can use similar arguments for $n = \kappa m + r$ with $0 \le r \le \kappa - 1$ fixed and $m \to \infty$. Note that $|b(t,v)| \le \Sigma_{n=1}^{\infty} |E(e^{it'S_n}|\tau = n)| Q(\tau = n) \le |E(e^{it'S_1}|\tau = 1)| Q(\tau = 1) + Q(\tau > 1)$, and that for the split chain under $Q$, $Q(\tau = 1) = \int h_\theta(x, \mathbf{R}^d)\,d\nu_\theta(x) > 0$. Since

$$\begin{aligned} E(e^{it'S_1}|\tau = 1) &= \left\{\int\int e^{it's} h_\theta(x, ds)\,d\nu_\theta(x)\right\} \Big/ \int\int h_\theta(x, ds)\,d\nu_\theta(x) \\ &= \int e^{it's}\,d\tilde{\nu}_\theta(s), \end{aligned}$$

where $\tilde{\nu}_\theta(ds) = e^{\theta's}\,d\tilde{\nu}(s) / \int_{\mathbf{R}^d} e^{\theta's}\,d\tilde{\nu}(s)$ has the same support as $\tilde{\nu}$ and is therefore nonlattice, it then follows that $|b(t,v)| < 1$ if $t \ne 0$. Therefore, for every $0 < \delta < 1$, $\max\{|b(t,v)| : \delta \le \|t\| \le \delta^{-1}, |v| \le \pi\} < 1$.

Using the change of variables $m = (n + z\sqrt{n})/E\tau$ and $w = v\sqrt{n}$ in (6.14) yields the representation $q_{n,\mu}(t/\sqrt{n}) = \mathrm{I} + \mathrm{II} + o(n^{-d})$, where

$$\begin{aligned} \mathrm{I} = \sum_{|z| \le n^{1/6}} \frac{1}{2\pi\sqrt{n}} \int_{|w| \le n^{1/10}} & e^{-iw\sqrt{n}} a\left(\frac{t}{\sqrt{n}}, \frac{w}{\sqrt{n}}\right) b^m\left(\frac{t}{\sqrt{n}}, \frac{w}{\sqrt{n}}\right) \\ & \times c\left(\frac{t}{\sqrt{n}}, \frac{w}{\sqrt{n}}\right) dw, \\ \mathrm{II} = \sum_{|z| \le n^{1/6}} \frac{1}{2\pi\sqrt{n}} \int_{n^{1/10} \le |w| \le \pi\sqrt{n}} & e^{-iw\sqrt{n}} a\left(\frac{t}{\sqrt{n}}, \frac{w}{\sqrt{n}}\right) b^m\left(\frac{t}{\sqrt{n}}, \frac{w}{\sqrt{n}}\right) \\ & \times c\left(\frac{t}{\sqrt{n}}, \frac{w}{\sqrt{n}}\right) dw. \end{aligned}$$

First consider the case $\|t\| \le n^{1/10}$ and $|w| \le n^{1/10}$. Then as $n \to \infty$,

$$\begin{aligned} & a(t/\sqrt{n}, w/\sqrt{n}) \to 1, \\ & c(t/\sqrt{n}, w/\sqrt{n}) \to E\left[\sum_{n=0}^{\tau-1} g(X_n)\right] = E\tau \int g(y)\,d\pi_{\theta_\mu}(y), \end{aligned} \tag{6.15}$$

where the last equality follows from Pitman's (1975) identity. By Lemma 3, $ES_\tau/E\tau = \mu$. Using a Taylor expansion of $E\exp\{i[(t'/\sqrt{n})(S_\tau - ES_\tau) + (w/\sqrt{n})(\tau - E\tau)]\}$ and writing $m = (n + z\sqrt{n})/E\tau$ with $|z| \le n^{1/6}$, we obtain that

$$
\begin{aligned}
b^m\left(\frac{t}{\sqrt{n}}, \frac{w}{\sqrt{n}}\right) &\sim e^{it'\mu(\sqrt{n}+z)+iw(\sqrt{n}+z)}\left\{1-\frac{1}{2n}\operatorname{Var}(t'S_\tau+w\tau)+o\left(\frac{1}{n}\right)\right\}^m \\
&\sim \exp\left\{it'\mu(\sqrt{n}+z)+iw(\sqrt{n}+z)-\frac{\operatorname{Var}(t'S_\tau+w\tau)}{2E\tau}\right\}.
\end{aligned}
\tag{6.16}
$$

Let $C_\tau = (\operatorname{Cov}(S_{\tau,1},\tau),\dots,\operatorname{Cov}(S_{\tau,d},\tau))'$. By (6.15) and (6.16),

$$
\begin{aligned}
\mathrm{I} &\sim \frac{1}{2\pi\sqrt{n}}\left[E\tau\int g(y)\,d\pi_{\theta_\mu}(y)\right]e^{i\sqrt{n}t'\mu-t'(\operatorname{Cov}S_\tau)t/2E\tau} \\
&\quad\times\sum_{|z|\le n^{1/6}} e^{it'\mu z}\int_{|w|\le n^{1/10}}\exp\left\{iwz-\frac{t'C_\tau w}{E\tau}-\frac{w^2\operatorname{Var}(\tau)}{2E\tau}\right\}dw.
\end{aligned}
\tag{6.17}
$$

The integral over $w$ in the RHS of (6.17) is asymptotically equivalent to

$$
\begin{aligned}
&\left[\exp\left\{\frac{(izE\tau-t'C_\tau)^2}{2\operatorname{Var}(\tau)E\tau}\right\}\right]\int_{-\infty}^{\infty}\exp\left\{-\left(w-\frac{izE\tau}{\operatorname{Var}(\tau)}+\frac{t'C_\tau}{\operatorname{Var}(\tau)}\right)^2\frac{\operatorname{Var}(\tau)}{2E\tau}\right\}dw \\
&\quad=\sqrt{\frac{2\pi E\tau}{\operatorname{Var}(\tau)}}\exp\left\{\frac{(izE\tau-t'C_\tau)^2}{2(E\tau)\operatorname{Var}(\tau)}\right\},
\end{aligned}
$$

and therefore the sum in the RHS of (6.17) is asymptotically equivalent (with $\Delta z = E\tau/\sqrt{n}$) to

$$
\begin{aligned}
&\sqrt{\frac{2\pi n}{(E\tau)\operatorname{Var}(\tau)}}\int_{-\infty}^{\infty}\exp\left\{-\frac{(zE\tau+it'C_\tau)^2}{(E\tau)\operatorname{Var}(\tau)}+it'\mu z\right\}dz \\
&\quad=\frac{2\pi\sqrt{n}}{E\tau}\exp\left\{\frac{t'\mu C_\tau' t}{E\tau}-\frac{t'\mu\mu' t\operatorname{Var}(\tau)}{2E\tau}\right\},
\end{aligned}
$$

in which the integral can be evaluated by completing the square. Substituting this into (6.17) yields

$$
\mathrm{I} = \big(1+o(1)\big)e^{it'\mu\sqrt{n}-t'V(\mu)t/2}\int g(y)\,d\pi_{\theta_\mu}(y)
\tag{6.18}
$$

uniformly in $\|t\| \le n^{1/10}$ and $\mu \in C$, noting that by Lemma 3,

$$
V(\mu) = \frac{\operatorname{Cov}(S_\tau-\mu\tau)}{E\tau} = \frac{\operatorname{Cov}(S_\tau)}{E\tau} - \frac{2\mu C_\tau'}{E\tau} + \mu\mu'\frac{\operatorname{Var}(\tau)}{E\tau}.
$$

For sufficiently small $\|t\|$ and $v$, use of Taylor's expansion of $b(t, v)$ as in (6.16) shows that $|b(t, v)| \le 1 - \beta(\|t\|^2 + v^2)$ for some $\beta > 0$. Combining this with $\max_{\delta \le \|t\| \le \delta^{-1}, |v| \le \pi} |b(t, v)| < 1$ as established before then shows $\mathrm{II} = o(n^{-d})$ and also (6.9). Therefore, from (6.18) and the decomposition $q_{n,\mu}(t/\sqrt{n}) = \mathrm{I} + \mathrm{II} + o(n^{-d})$, it follows that $q_{n,\mu}(t/\sqrt{n}) = (1 + o(1))e^{i\sqrt{n}\mu' t}\widehat{\phi}_\mu(t) \int g(y)\, d\pi_{\theta_\mu}(y) + o(n^{-d})$ uniformly in $\|t\| \le n^{1/10}$ and $\mu \in C$, proving (6.8). □

## REFERENCES


ARNDT, K. (1980). Asymptotic properties of the distribution of the supremum of a random walk on a Markov chain. *Theory Probab. Appl.* **25** 309–324.

BHATTACHARYA, R. N. and RANGA RAO, R.(1976). *Normal Approximations and Asymptotic Expansions*. Wiley, New York.

BOROVKOV, A. A. and ROGOZIN, B. A. (1965). On the multidimensional central limit theorem. *Theory Probab. Appl.* **10** 55–62.

CHAN, H. P. and LAI, T. L. (2000). Asymptotic approximations for error probabilities of sequential or fixed sample size tests in exponential families. *Ann. Statist.* **28** 1638–1669.

CHAN, H. P. and LAI, T. L. (2002). Boundary crossing probabilities for scan statistics and their applications to change-point detection. *Methodol. Comput. Appl. Probab.* To appear.

DANIELS, H. E. (1954). Saddlepoint approximation in statistics. *Ann. Math. Statist.* **25** 631–650.

DE ACOSTA, A. and NEY, P. (1998). Large deviation lower bounds for arbitrary additive functionals of a Markov chain. *Ann. Probab.* **26** 1660–1682.

DINWOODIE, I. H. (1993). Identifying a large deviation rate function. *Ann. Probab.* **21** 216–231.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

GRAY, A. (1990). *Tubes*. Wesley, New York.

HIRSCH, M. (1976). *Differential Topology*. Springer, New York.

HÖGLUND, T. (1974). Central limit theorems and statistical inference for finite Markov chains. *Z. Wahrsch. Verw. Gebiete* **29** 123–151.

HÖGLUND, T. (1991). The ruin problem for finite Markov chains. *Ann. Probab.* **19** 1298–1310.

IGLEHART, D. L. (1972). Extreme values in the $GI/G/1$ queue. *Ann. Math. Statist.* **43** 627–635.

ILTIS, M. (1995). Sharp asymptotics of large deviations in $\mathbf{R}^d$. *J. Theoretical Probab.* **8** 501–524.

ISCOE, I., NEY, P. and NUMMELIN, E. (1985). Large deviations of uniformly recurrent Markov additive processes. *Adv. in Appl. Math.* **6** 373–412.

JENSEN, J. L. (1991). Saddlepoint expansions for sums of Markov dependent variables on a continuous state space. *Probab. Theory Related Fields* **89** 181–199.

JENSEN, J. L. (1995). *Saddlepoint Approximations*. Oxford Univ. Press.

KARLIN, S., DEMBO, A. and KAWABATA, T. (1990). Statistical composition of high scoring segments from molecular sequences. *Ann. Statist.* **18** 571–581.

MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, New York.

NEY, P. and NUMMELIN, E. (1987). Markov additive processes. I. Eigenvalues properties and limit theorems, and II. Large deviations. *Ann. Probab.* **15** 561–609.

PITMAN, J. W. (1975). An identity for stopping times of a Markov process. In *Studies in Probability and Statistics* (E. J. Williams, ed.) 41–57. North-Holland, Amsterdam.

SIEGMUND, D. (1988). Approximate tail probabilities for the maxima of some random fields. *Ann. Probab.* **16** 487–501.

SIEGMUND, D. and VENTRAKAMAN, E. (1995). Using the generalized likelihood ratio statistics for sequential detection of a change-point. *Ann. Statist.* **23** 255–271.

STONE, C. (1965). Local limit theorem for nonlattice multidimensional distribution functions. *Ann. Math. Statist.* **36** 546–551.

DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
SINGAPORE 119260
REPUBLIC OF SINGAPORE
E-MAIL: stachp@nus.edu.sg

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
E-MAIL: lait@stat.stanford.edu