stochastic
processes
and their

www.elsevier.com/locate/spa

Poisson equation, moment inequalities and quick convergence for Markov random walks

Cheng-Der Fuh[a], *, Cun-Hui Zhang[b]

[a] *Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, ROC*
[b] *Department of Statistics, Hill Center, Busch Campus, Rutgers University, Piscataway, NJ 08854, USA*

Received 29 October 1998; received in revised form 5 July 1999

Abstract

We provide moment inequalities and sufficient conditions for the quick convergence for Markov random walks, without the assumption of uniform ergodicity for the underlying Markov chain. Our approach is based on martingales associated with the Poisson equation and Wald equations for the second moment with a variance formula. These results are applied to non-linear renewal theory for Markov random walks. A random coefficient autoregression model is investigated as an example. © 2000 Elsevier Science B.V. All rights reserved.

MSC: primary 60G40; secondary 60J10

Keywords: Inequality; Markov random walk; Tail probability; Moment; Poisson equation; Quick convergence; Wald equation; Renewal theory

1. Introduction

Let $\{X_n,\ n\geqslant 0\}$ be an irreducible Markov chain on a general state space D with σ-algebra $\mathscr{A}$, where irreducibility is with respect to a maximal irreducible measure on $\mathscr{A}$. Let $\{(X_n,\zeta_n),\ n\geqslant 0\}$ be a Markov chain on $D\times\mathbb{R}$, with an additional sequence of random variables ζ_n, such that the transition probability distributions

$$P\{(X_{n+1},\zeta_{n+1})\in A\times B|X_n=x,(X_{k-1},\zeta_k),1\leqslant k\leqslant n\}=P(x,A\times B) \tag{1.1}$$

do not depend on the "time" $n\geqslant 0$ and the values of (X_{k-1},ζ_k), $1\leqslant k\leqslant n$, for all $x\in D, A\in\mathscr{A}$ and Borel sets B. For measurable functions g on $D\times D\times\mathbb{R}$, the chain $\{(X_n,S_n),\ n\geqslant 0\}$, with an additive component $S_n=\sum_{k=1}^n g(X_{k-1},X_k,\zeta_k)$, $S_0=0$, is called a *Markov random walk*. In the standard setting for Markov random walks, cf. Ney and Nummelin (1987), one simply considers $g(X_{k-1},X_k,\zeta_k)=\zeta_k$ without loss of generality. The notation $g(X_{k-1},X_k,\zeta_k)$ is used here because our investigation of S_n involves the solutions of the Poisson equation for several related functions g. We assume throughout the paper that there exists a stationary probability distribution π, $\pi(A)=\int P(x,A)\pi(\mathrm{d}x)$ for all $A\in\mathscr{A}$.

* Corresponding author.
E-mail addresses: stcheng@ccvax.sinica.edu.tw (C.-D. Fuh), czhang@stat.rutgers.edu (C.-H. Zhang)

PII: S0304-4149(99)00104-0

The concept of quick convergence was introduced by Strassen (1967). A sequence of random variables θ_n is said to converge β-quickly ($\beta > 0$) to a constant μ if

$$E(\sup\{n \geqslant 1: |\theta_n - \mu| \geqslant \varepsilon\})^\beta < \infty \quad \text{for all } \varepsilon > 0. \tag{1.2}$$

Obviously $\theta_n \to \mu$ β-quickly for some $\beta > 0$ implies $\theta_n \to \mu$ a.s. For simple random walks S_n (i.e. with i.i.d. increments), quick convergence played an important role in nonlinear renewal theory and sequential analysis (cf. Lai, 1981). Likewise, extensions of (1.2) to Markov random walks are crucial in the development of nonlinear renewal theory for Markov random walks and applications to statistical analysis of dependent observations (e.g. time series).

We shall consider quick convergence of $\theta_n = S_n/n$, $n \geqslant 1$, for Markov random walks $\{(X_n, S_n),\ n \geqslant 0\}$. Our approach is based on extensions of tail probability and moment inequalities of Chow and Lai (1975) for the maxima of partial sums from the i.i.d. case to the Markov case, which we obtain from martingales and Wald equations associated to the Poisson equation. These results, established without the assumption of uniform ergodicity for the underlying Markov chain, are of independent interest and have applications in other areas of research.

The paper is organized as follows. The main results are stated in Section 2. In Section 3, we construct martingales based on the Poisson equation, and provide Wald equations and a variance formula. In Section 4, we prove the moment inequalities based on tail probabilities for the maxima of partial sums. In Section 5, we prove the quick convergence of Markov random walks. Applications to nonlinear renewal theory are discussed in Section 6. A random coefficient autoregression model is considered in Section 7.

2. Main results

Let $\{(X_n, \zeta_n),\ n \geqslant 1\}$ be a Markov chain as in (1.1). Let ν be an initial distribution of X_0 and define $\nu^*(A) = \sum_{n=0}^{\infty} P_\nu(X_n \in A)$ on $\mathscr{A}$. Let E_ν be the expectation under which X_0 has initial distribution ν, and E_x be the conditional expectation given $X_0 = x$ (i.e. with ν degenerate at x). Let $\mathscr{H} = \mathscr{H}_\nu$ be the class of all measurable functions $g = g(x, y, \zeta)$ on $D \times D \times \mathbb{R}$ such that $E_x|g(x, X_1, \zeta_1)| < \infty$ ν^*-almost surely and $E_\pi|g(X_0, X_1, \zeta_1)| < \infty$. For $g \in \mathscr{H}$, define operators P and P_π by $(Pg)(x) = E_x g(x, X_1, \zeta_1)$ and $P_\pi g = E_\pi g(X_0, X_1, \zeta_1)$ respectively, and set $\bar{g} = Pg$. We shall consider solutions $\Delta = \Delta(x; g) \in \mathscr{H}$ of the Poisson equation

$$(P - I)\Delta = (I - P_\pi)\bar{g} \qquad \nu^* - \text{a.e.},\ P_\pi \Delta = 0, \tag{2.1}$$

where I is the identity. Here and in the sequel a function f defined on D is always treated as a function on $D \times D \times \mathbb{R}$ via $f(y) = g(x, y, \zeta)$, so that $Pf = \int f(y) P(x, \mathrm{d}y)$. Since $(I - P_\pi)\bar{g} = (P - P_\pi)g$, the Poisson equation (2.1) can also be written as $(P - I)\Delta = (P - P_\pi)g$. Let $\mathscr{H}^* = \mathscr{H}^*_\nu$ be the set of $g \in \mathscr{H}$ such that the solution of (2.1) exists with $E_\nu|\Delta(X_n; g)| < \infty$ and $E_\nu|g(X_{n-1}, X_n, \zeta_n)| < \infty$ for all $n \geqslant 0$.

Set $\mu = \mu_g = P_\pi g = E_\pi g(X_0, X_1, \zeta_1)$. For $g \in \mathscr{H}^*$ define

$$d(g) = d(x, y, \zeta; g) = g(x, y, \zeta) - \mu_g - \Delta(y; g) + \Delta(x; g). \tag{2.2}$$

Furthermore, set $\sigma_p = \sigma_p(g) = \{E_\pi|d(X_0,X_1,\zeta_1;g)|^p\}^{1/p}$, and for $|d(g)|^p \in \mathscr{H}^*$ define

$$D_p(g) = D_p(x,y,\zeta;g) = |d(x,y,\zeta;g)|^p - \sigma_p^p(g) - \Delta_p(y;g) + \Delta_p(x;g), \tag{2.3}$$

where $\Delta_p(x;g) = \Delta(x;|d(g)|^p)$. Let $\mathscr{F}_n$ be the σ-algebra generated by the random variables $\{X_0,(X_i,\zeta_i),1\leqslant i\leqslant n\}$. For $g \in \mathscr{H}^*$ and $d(g)$ in (2.2), set

$$S_n = S_n(g) = \sum_{j=1}^{n} \tilde{\zeta}_j, \quad \tilde{\zeta}_j = g(X_{j-1},X_j,\zeta_j), \quad d_j = d_j(g) = d(X_{j-1},X_j,\zeta_j;g), \tag{2.4}$$

$$S_n^* = S_n^*(g) = \max_{1\leqslant j\leqslant n} |S_j(g) - j\mu_g|, \qquad g_n^* = \max_{1\leqslant j\leqslant n} |g(X_{j-1},X_j,\zeta_j) - \mu_g|, \tag{2.5}$$

$$f_n(g) = \sum_{j=1}^{n} d_j = S_n - n\mu_g - \Delta(X_n;g) + \Delta(X_0;g), \quad n\geqslant 1 \tag{2.6}$$

and

$$f_n^* = f_n^*(g) = \max_{1\leqslant j\leqslant n} |f_j(g)|, \qquad d_n^* = d_n^*(g) = \max_{1\leqslant j\leqslant n} |d_j|. \tag{2.7}$$

Furthermore, define

$$\begin{aligned} \Delta_n^* &= \Delta_n^*(g) = \max_{0\leqslant i<j\leqslant n} |\Delta(X_j;g) - \Delta(X_i;g)|, \\ V_{p,n}^* &= V_{p,n}^*(g) = \max_{0\leqslant j\leqslant n} V_{p,j,n}, \end{aligned} \tag{2.8}$$

where for the $\Delta_p(x;g)$ in (2.3)

$$V_{p,j,n} = V_{p,j,n}(g) = [\{E[\Delta_p(X_n;g)|X_j] - \Delta_p(X_j;g)\}^+]^{1/p}. \tag{2.9}$$

Theorem 1. *Let $p\geqslant 2$. Suppose $\{g,d^2(g)\}\subset\mathscr{H}^*$. Then, there exists a constant C_p such that*

$$\max\{E_\nu(f_n^*)^p, E_\nu(S_n^*)^p\} \leqslant C_p E_\nu\{\max(\sigma\sqrt{n}, g_n^*, \Delta_n^*, V_{2,n}^*)\}^p. \tag{2.10}$$

If in addition $E_\pi|g(X_0,X_1,\zeta_1)|^p<\infty$, $E_\pi|\Delta(X_0;g)|^p<\infty$ and $E_\pi|\Delta_2(X_0;g)|^{p/2}<\infty$, then

$$\max\{E_\pi(f_n^*)^p, E_\pi(S_n^*)^p\} \leqslant C_p\{(\sigma\sqrt{n})^p + \mathrm{o}(1)n^{p/2}\}, \text{ as } n\to\infty. \tag{2.11}$$

Theorem 1 is proved in Section 4, which contains a more general result.

Let θ_n be as in (1.2) and define $N_\varepsilon = \sup\{n\geqslant 1\colon |\theta_n-\mu|\geqslant\varepsilon\}$. Since (1.2) is equivalent to $EN_\varepsilon^\beta < \infty$, the sequence $\{\theta_n\}$ converges β-quickly to μ if and only if

$$\sum_{n=1}^{\infty} n^{\beta-1}P\{N_\varepsilon\geqslant n\} = \sum_{n=1}^{\infty} n^{\beta-1}P\left\{\sup_{j\geqslant n}|\theta_j-\mu|\geqslant\varepsilon\right\} < \infty, \ \forall\varepsilon > 0.$$

Theorem 2. *Suppose $\{g,|d(g)|^r\}\subseteq\mathscr{H}^*$. Let $\Delta_r(x;g)$, d_j, S_n^* and f_n^* be as in* (2.3)–(2.5) *and* (2.7). *Set $\Delta_r^{\pm}(x;g) = \max\{\pm\Delta_r(x;g),0\}$. Let $1<r\leqslant 2$ and $\beta>0$. Suppose*

$E_\pi |d_1|^{\beta+1} < \infty$ and that $E_\pi[\{\Delta_r(X_1; g)\}^-]^{(\beta+1+\varepsilon)/r} < \infty$ and $E_\pi[\{\Delta_r(X_1; g)\}^+]^{1\vee[(\beta+\varepsilon)/r]} < \infty$ for some $\varepsilon > 0$. Then,

$$\sum_{n=1}^{\infty} n^{\beta-1} P_\pi \left\{ \sup_{j \geqslant n} |S_n(g) - \Delta(X_n; g) + \Delta(X_0; g) - n\mu(g)| \geqslant \varepsilon n \right\} < \infty, \quad \forall \varepsilon > 0. \tag{2.11}$$

If in addition $E_\pi|\tilde{\zeta}_1|^{\beta+1} < \infty$, then

$$\sum_{n=1}^{\infty} n^{\beta-1} P_\pi \left\{ \sup_{j \geqslant n} |S_n(g) - n\mu(g)| \geqslant \varepsilon n \right\} < \infty, \quad \forall \varepsilon > 0. \tag{2.12}$$

In Section 5, we provide in Theorem 6 sufficient moment conditions for the quick convergence of $\theta_n = n^{-\alpha}(S_n(g) - n\mu_g)$ to 0, $\alpha > 1/r$, under P_v for general initial distributions v. In Theorem 6, the moment conditions on $|d_j|$ are slightly stronger with an extra logarithmic factor, while those on $\Delta_r^\pm$ are weaker with $\varepsilon = 0$ and certain additional logarithmic factors. The proof of Theorem 2 will be given at the end of Section 5.

Remark 1. The Poisson equation (2.1) is slightly different from $(P - I)\Delta = (I - P_\pi)g$ considered in some previous studies (cf. (17.37) of Meyn and Tweedie, 1993). A sufficient condition for the existence of (2.1) can be found in Theorem 17.4.2 of Meyn and Tweedie (1993). Theorems 1 and 2 provide the moment inequalities and quick convergence for Markov random walks, respectively, under the assumptions of the existence of solution for the Poisson equation, and moment conditions. It is known that under the uniform ergodicity condition, the Poisson equation (2.1) has uniformly bounded solutions $\Delta(\cdot; g)$ for bounded $\bar{g} = Pg$ (cf. Fuh and Lai, 1998); therefore, the moment conditions $\Delta(x; g)$ in Theorems 1 and 2 are automatically satisfied in this case. Also, the moment conditions on $\Delta_p(x; g)$ in Theorems 1 and 2 can be easily verified in autoregression models for the natural $g(x, y, \zeta) = y$. For a general Markov random walk, an upper bound of $\Delta(\cdot; g)$, via the drift inequality, can be found in Theorem 17.4.2 of Meyn and Tweedie (1993).

Remark 2. Lai (1977) studied the quick convergence for stationary mixing sequences. Under some mixing conditions, he required the moment condition $E|\tilde{\zeta}_1 - \mu|^q < \infty$ for some $q > \max\{\beta + 1, 2\}$. Further generalization can be found in Peligrad (1985). Irle (1993) investigated the quick convergence for regenerative processes and its applications to Harris recurrent Markov chains under certain moment condition for the induced renewal process. Rates of convergence in the law of large numbers can be found in Irle (1990) and Alsmeyer (1990), who generalized the results of Chow and Lai (1975) from the i.i.d. case to general martingale sequences.

3. Preliminaries: martingales and Wald equations

In this section, we explore the martingale structure associated with the Poisson equation (2.1). Martingale (2.6) is equivalent to (17.42) of Meyn and Tweedie (1993). The variance formula (3.3) below is equivalent to (17.47) of Meyn and Tweedie (1993) but seems to be more transparent. The quadratic martingales given in (3.1) and (3.2) below are crucial in our further investigation.

Theorem 3. (i) *Suppose* $g \in \mathscr{H}^*$. *Then, the sequence* $\{f_n\}$ *in* (2.6) *is a martingale with respect to* $\{\mathscr{F}_n\}$ *under* P_v, *with* d_j *being the martingale differences.*

(ii) *Suppose* $\{g, d^2(g)\} \subset \mathscr{H}^*$. *Then both sequences*

$$F_n(g) = f_n^2(g) - n\sigma_2^2(g) - \Delta_2(X_n; g) + \Delta_2(X_0; g), \quad n \geqslant 1 \tag{3.1}$$

and

$$F_n'(g) = (S_n - n\mu_g - \Delta(X_n; g))^2 - n\sigma_2^2(g) - \Delta_2(X_n; g) + \Delta_2(X_0; g), \quad n \geqslant 1 \tag{3.2}$$

are $\{\mathscr{F}_n\}$*-martingales under* P_v. *Moreover,*

$$\sigma_2^2 = \sigma_2^2(g) = P_\pi d^2(g) = Var_\pi(\tilde{\zeta}_1) + 2\sum_{j=1}^{\infty} Cov_\pi(\tilde{\zeta}_1, \tilde{\zeta}_{1+j}). \tag{3.3}$$

Proof. It follows from (2.1) and (2.2) that (2.6) is indeed a martingale, as

$$E[d_j | X_{j-1}] = (Pg - P_\pi g - P\Delta + \Delta)(X_{j-1}) = (\bar{g} - P_\pi \bar{g} - P\Delta + \Delta)(X_{j-1}) = 0$$

P_v-a.s. For the same reason, by (2.3)

$$\sum_{j=1}^{n} D_2(X_{j-1}, X_j, \zeta_j; g) = \sum_{j=1}^{n} d_j^2 - n\sigma_2^2 - \Delta_2(X_n; g) + \Delta_2(X_0; g)$$

is a martingale. In addition, the sequence $f_n^2(g) - \sum_{j=1}^n d_j^2$ is a martingale, as d_j are martingale differences. These imply that (3.1) is a martingale as it is a sum of two martingales. Finally, (3.2) is a martingale as it is the sum of (3.1) and the martingale $-2f_n(g)\Delta(X_0; g) + \Delta^2(X_0; g)$. □

The following corollary provides the Wald equations for bounded stopping rules via the optional stopping theorem.

Corollary 1. *Let* $\mu = \mu_g$, $\sigma_2 = \sigma_2(g)$ *and* $S_n = S_n(g)$ *be as in* (2.2)–(2.4). *Let* T *be a bounded stopping rule with respect to* $\{\mathscr{F}_n, n \geqslant 0\}$. *Suppose* $g \in \mathscr{H}^*$. *Then*

$$E_v S_T = \mu E_v T + E_v\{\Delta(X_T; g) - \Delta(X_0; g)\}. \tag{3.4}$$

If in addition $d^2(g) \in \mathscr{H}^*$ *and* $E_v \Delta^2(X_n; g) < \infty$ *for all* $n \geqslant 0$, *then*

$$E_v(S_T - \mu T)^2 = \sigma_2^2 E_v T + 2E_v\{(S_T - \mu T)\Delta(X_T; g)\} + E_v\{\Delta_2'(X_T; g) - \Delta_2'(X_0; g)\}, \tag{3.5}$$

where $\Delta_2'(x; g) = \Delta_2(x; g) - \Delta^2(x; g)$.

The following theorem provides the Wald equations for Markov random walks and unbounded stopping rules with finite expectation, via uniform integrability conditions. Under the stronger uniform ergodicity condition for the underlying Markov chain, Fuh and Lai (1998) derived Wald equations via an exponential martingale.

Theorem 4. *Let* T *be a stopping rule with* $E_v T < \infty$. (i) *Suppose there exists a uniformly bounded* $h = h(x, y)$ *such that* $\tilde{g} = |g - h| \in \mathscr{H}^*$ *with* $\sup_n E_v \Delta(X_{T\wedge n}; \tilde{g}) < \infty$. *If* $g \in \mathscr{H}^*$ *and* $\{\Delta(X_{T\wedge n}; g), n \geqslant 1\}$ *is* E_v*-uniformly integrable, then* (3.4) *holds.*

(ii) *Suppose* $\{g, d^2(g)\} \subseteq \mathscr{H}^*$, T *is a stopping rule with* $E_v T < \infty$ *and that both* $\{\Delta_2(X_{T\wedge n}; g),\ n \geqslant 1\}$ *and* $\{\Delta^2(X_{T\wedge n}; g),\ n \geqslant 1\}$ *are* E_v-*uniformly integrable. Then*, (3.5) *holds*.

Remark. Consider vectors of the form $(\tilde{g}_1, \tilde{g}_2) = C(g, d(g)) - (h_1, h_2)$, where C is a 2×2 deterministic matrix of full rank and $h_j = h_j(x, y)$ are uniformly bounded. By the proof of Theorem 4, conditions of Theorem 4(i) hold if $\{g, |\tilde{g}_1|, |\tilde{g}_2|\} \subset \mathscr{H}^*$ with $\sup_n E_v \Delta(X_{T\wedge n}; |\tilde{g}_j|) < \infty$, $j = 1, 2$. Sufficient conditions for the uniform integrability of $\{h(X_{T\wedge n})\}$, such as $\{\Delta(X_{T\wedge n}; g)\}$, are given in Proposition 1 below. Under the uniform ergodicity condition, $\Delta(x; g)$ are bounded functions of x for bounded $\bar{g} = Pg$.

Proof. (i) Let $S_n(g)$ be as in (2.4). By (3.4),

$$E_v S_{T\wedge n}(\tilde{g}) = \mu_{\tilde{g}} E_v\{T \wedge n\} + E_v \Delta(X_{T\wedge n}; \tilde{g}) - E_v \Delta(X_0; \tilde{g}) = \mathrm{O}(1),$$

so that $E_v S_T(\tilde{g}) < \infty$. Since h is uniformly bounded and $\tilde{g} = |g - h|$, $E_v S_T(|g|) \leqslant E_v S_T(\tilde{g}) + \|h\|_\infty E_v T < \infty$, so that $\{S_{T\wedge n}(g)\}$ is uniformly integrable. Since $\{\Delta(X_{T\wedge n}; g)\}$ is also uniformly integrable, we obtain (3.4) with the stopping rule T by taking the limit $n \to \infty$ in (3.4) with bounded stopping rules $T \wedge n$.

(ii) We only need to consider $\mu_g = 0$. By (3.1) and the optional stopping theorem,

$$\begin{aligned} E_m^* \left(\sum_{j=T\wedge m}^{T\wedge n} d_j \right)^2 &= E_m^* (S_{T\wedge n} - S_{T\wedge m} - \Delta(X_{T\wedge n}; g) + \Delta(X_{T\wedge m}; g))^2 \\ &= \sigma^2 E_m^* ((T \wedge n) - (T \wedge m))^+ + E_m^* \{\Delta_2(X_{T\wedge n}; g) - \Delta_2(X_{T\wedge m}; g)\} \end{aligned}$$

for $n \geqslant m$, where E_m^* is the conditional expectation given $\mathscr{F}_{T\wedge m}$. Thus, $\|S_{T\wedge n} - S_{T\wedge m}\|_2 \leqslant \sigma \|(T\wedge n) - (T\wedge m)\|_1^{1/2} + \|\Delta_2(X_{T\wedge n}; g) - \Delta_2(X_{T\wedge m}; g)\|_1^{1/2} + \|\Delta(X_{T\wedge n}; g) - \Delta(X_{T\wedge m}; g)\|_2 \to 0$, as $n \to \infty$ by the uniform integrability assumptions. Since $\{\Delta^2(X_{T\wedge n}; g)\}$ is uniformly integrable, this implies that the right-hand side of (3.5) is uniformly integrable with $T \wedge n$ in place of T. Hence, (3.5) holds by taking limits in its $T \wedge n$ version. □

Proposition 1 (Uniform integrability). *Let* h *be a Borel function with finite* $E_v|h(X_n)|$, $n \geqslant 1$. *Let* T *be an integer-valued random variable with* $E_v T < \infty$. *If* $E_v|h(X_T)| < \infty$ *and*

$$\lim_{M\to\infty} \limsup_{n\to\infty} E_v(|h(X_n)| - Mn)^+ = 0, \tag{3.6}$$

then $\{h(X_{T\wedge n})\}$ *is* E_v-*uniformly integrable. Consequently*, $\{h(X_{T\wedge n})\}$ *is* E_v-*uniformly integrable if*

$$\sum_{n=1}^{\infty} E_v(|h(X_n)| - n)^+ < \infty. \tag{3.7}$$

Remark. For $v = \pi$, (3.6) is equivalent to $E_\pi|h(X_0)| < \infty$, and (3.7) is equivalent to $E_\pi|h(X_0)|^2 < \infty$. It can be proved that $p = 2$ is the smallest real value of p such that $E|h_T| < \infty$ for all integer-valued random variables T with $ET < \infty$ and all identically distributed $\{h_n\}$ with $E|h_1|^p < \infty$.

Proof. If $E_v|h(X_T)| < \infty$, then $E_v|h(X_T) - h(X_{T\wedge n})|$ is bounded by

$$E_v|h(X_T)|I_{\{T\geqslant n\}} + E_v|h(X_n)|I_{\{T\geqslant n\}}$$
$$\leqslant o(1) + MnP_v\{T\geqslant n\} + E_v(|h(X_n)| - Mn)^+ \to 0.$$

If (3.7) holds, then $E_v|h(X_T)|\leqslant E_vT + \sum_{n=1}^{\infty} E_v(|h(X_n)| - n)^+ < \infty$. □

Corollary 2. *Let T be a stopping rule with $E_vT < \infty$. If $\{g,\tilde{g}\}\subset \mathscr{H}^*$ for the $\tilde{g}$ in Theorem* 4 *and* (3.7) *holds for both $h(x)=\Delta(x;g)$ and $h(x)=\Delta(x;\tilde{g})$, then* (3.4) *holds. If $\{g,d^2(g)\}\subseteq \mathscr{H}^*$ and* (3.7) *holds for both $h(x) = \Delta^2(x;g)$ and $h(x) = \Delta_2(x;g)$, then* (3.5) *holds.*

4. Moment inequalities

The type of tail probability inequalities under consideration here for Markov random walks can be called "good-λ" inequality, which naturally implies moment inequalities for maxima of partial sums. In martingale and i.i.d. settings, these inequalities were considered by Rosenthal (1970), Burkholder and Gundy (1970), Burkholder (1973), Hoffmann-Jorgensen (1974, 1977) and Chow and Lai (1975, 1978) among others. The results here can be viewed as an extension of Chow and Lai (1975) from the i.i.d. case to the Markov case.

Let $g\in\mathscr{H}^*$ with $|d(g)|^p\in\mathscr{H}^*$ for some $p\geqslant 1$. Define

$$\tilde{\zeta}_n^* = \tilde{\zeta}_n^*(g) = \max_{1\leqslant j\leqslant n} |\tilde{\zeta}_j - \mu_g|, \tag{4.1}$$

where $\tilde{\zeta}_j = g(X_{j-1},X_j,\zeta_j)$ are as in (2.4). Let f_n^*, d_n^*, $V_{p,n}^*$, S_n^* and Δ_n^* be as in (2.5)–(2.8).

Lemma 1. *Suppose $\{g,|d(g)|^p\}\subset\mathscr{H}^*$ for some $1\leqslant p\leqslant 2$. Then,*

$$P_v\{f_n^* > t + k(\varepsilon+s), d_n^*\leqslant s, V_{p,n}^*\leqslant v\}\leqslant P_v\{f_n^* > t\}B_p^{kp}\left(\frac{n\sigma_p^p + v^p}{\varepsilon^p}\right)^k \tag{4.2}$$

for all positive s, t, v and ε, where σ_p is as in (3.3), *$B_p = 18p^{3/2}/(p-1)^{1/2}$ for $1 < p < 2$ and $B_1 = B_2 = 1$. Moreover, for all positive s, t, u, v and ε,*

$$P_v\{S_n^* > t + k(\varepsilon+s+u), \tilde{\zeta}_n^*\leqslant s, \Delta_n^*\leqslant u, V_{p,n}^*\leqslant v\}$$
$$\leqslant P_v\{S_n^* > t\}B_p^{kp}\left(\frac{n\sigma_p^p + v^p}{\varepsilon^p}\right)^k. \tag{4.3}$$

Proof. Define $\tau_0 = \inf\{j\colon |\sum_{i=1}^{j} d_i| > t\}$ and for $k\geqslant 1$

$$\tau_k = \inf\left\{j > \tau_{k-1}\colon \left|\sum_{i=1+\tau_{k-1}}^{j} d_i\right| > \varepsilon\right\}.$$

On the event $d_n^*\leqslant s$, the overshoot for the stopping rules τ_j is bounded by s for all $j\geqslant 0$, so that $f_{\tau_{k-1}}^*\leqslant t + (k-1)\varepsilon + ks$ and the left-hand side of (4.2) is bounded by

$P_v\{\tau_k \leqslant n, d_n^* \leqslant s, V_{p,n}^* \leqslant v\}$ for all $k \geqslant 1$. Therefore, it suffices to prove by induction that for all $k \geqslant 1$

$$P_v\{A_k\} \leqslant P_v\{A_{k-1}\} B_p^p \left(\frac{n\sigma_p^p + v^p}{\varepsilon^p} \right), \tag{4.4}$$

where $A_k = \{\tau_k \leqslant n\} \cap \{|d_j| \leqslant s, V_{p,j,n} \leqslant v, \ \forall j \leqslant \tau_k\}$. Let E^* be the conditional expectation given $\mathscr{F}_{\tau_{k-1}}$ for some fixed $k \geqslant 1$. By the Doob and Burkholder inequalities

$$E^* I_{A_k} \leqslant I_{A_{k-1}} E^* I\{\tau_k \leqslant n\} \leqslant I_{A_{k-1}} E^* \left[\frac{|f_n - f_{\tau_{k-1}}|^p}{\varepsilon^p} \right] \leqslant I_{A_{k-1}} \frac{B_p^p}{\varepsilon^p} E^* \sum_{j=\tau_{k-1}+1}^{n} |d_j|^p.$$

Note that $1 \leqslant p \leqslant 2$. An application of (3.4) of Corollary 1 with $|d(g)|^p$ in place of g yields

$$\begin{aligned} E^* \sum_{j=\tau_{k-1}+1}^{n} |d_j|^p &= \sigma_p^p (n - E^* \tau_{k-1}) + E^* \{ \Delta_p(X_n; g) - \Delta_p(X_{\tau_{k-1}}; g) \} \\ &\leqslant \sigma_p^p n + V_{p,\tau_{k-1},n}^p, \end{aligned}$$

in view of (2.9). These imply (4.4), and therefore (4.2), as $V_{p,\tau_{k-1},n} \leqslant v$ on A_{k-1}.

The proof of (4.3) is identical with modified $\tau_0 = \inf\{n: S_n^* > t\}$. The details are omitted. Note that on the event $\{S_n^* > t + k(\varepsilon + s + u), \tilde{\zeta}_n^* \leqslant s, \Delta_n^* \leqslant u\}$, we have $S_{\tau_0}^* \leqslant t + s$ and

$$\max_{\tau_0 < j \leqslant n} \left| (S_{\tau_0} - \mu \tau_0) + \sum_{i=\tau_0+1}^{j} d_i \right| > t + k(\varepsilon + s + u) - u,$$

which imply $\tau_k \leqslant n$ as $\max_{\tau_0 < j \leqslant n} |\sum_{i=\tau_0+1}^{j} d_i| > k\varepsilon + (k-1)(s+u)$ and $d_n^* \leqslant s + u$. □

Next, we apply Lemma 1 to obtain moment inequalities.

Theorem 5. *Suppose $\{g, |d(g)|^p\} \subset \mathscr{H}^*$ for some $1 \leqslant p \leqslant 2$. Let $\Phi(x)$ be a nonnegative nondecreasing function such that for some $0 < \lambda < 1$, $0 < \varepsilon < 1 - \lambda$ and finite real number C, $\Phi(t) \leqslant C\Phi(\lambda x)$ for all $x \geqslant \sigma_p n^{1/p}/\varepsilon$. Then, for $2CB_p^p \varepsilon^p/(1-\lambda-\varepsilon)^p < 1$*

$$E_v \Phi(f_n^*) \leqslant \frac{E_v \Phi(\varepsilon^{-1} \max(\sigma_p n^{1/p}, d_n^*, V_{p,n}^*))}{1 - 2CB_p^p \varepsilon^p/(1-\lambda-\varepsilon)^p} \tag{4.5}$$

and

$$E_v \Phi(S_n^*) \leqslant \frac{E_v \Phi(\varepsilon^{-1} \max(\sigma_p n^{1/p}, 2\tilde{\zeta}_n^*, 2\Delta_n^*, V_{p,n}^*))}{1 - 2CB_p^p \varepsilon^p/(1-\lambda-\varepsilon)^p}. \tag{4.6}$$

Proof. Clearly, $E_v \Phi(f_n^*) = \int_0^\infty P_v\{f_n^* > t\} \,\mathrm{d}\Phi(t)$ is bounded by

$$\begin{aligned} E_v \Phi(f_n^*) &\leqslant \Phi(\sigma_p n^{1/p}/\varepsilon) + \int_{\sigma_p n^{1/p}/\varepsilon}^{\infty} P_v\{\max(d_n^*, V_{p,n}^*) > \varepsilon t\} \,\mathrm{d}\Phi(t) \\ &\quad + \int_{\sigma_p n^{1/p}/\varepsilon}^{\infty} P_v\{f_n^* > t, \max(d_n^*, V_{p,n}^*) \leqslant \varepsilon t\} \,\mathrm{d}\Phi(t). \end{aligned}$$

The sum of the first two terms on the right-hand side is $E_v \Phi(\varepsilon^{-1} \max(\sigma_p n^{1/p}, d_n^*, V_{p,n}^*))$, the numerator on the right-hand side of (4.5), while by (4.2), with $(1, \lambda t, t - \lambda t - \varepsilon t, \varepsilon t, \varepsilon t)$

taking place of (k,t,ε,s,v) in (4.2), the third term above on the right-hand side is bounded by

$$\int_{\sigma_p n^{1/p}/\varepsilon} P_v\{f_n^* > \lambda t\} \frac{B_p^p\{\sigma_p^p n + (\varepsilon t)^p\}}{(t-\lambda t-\varepsilon t)^p}\,\mathrm{d}\Phi(t)$$
$$\leqslant \frac{2B_p^p\varepsilon^p}{(1-\lambda-\varepsilon)^p}\int_{\sigma_p n^{1/p}/\varepsilon} P_v\{f_n^* > \lambda t\}\,\mathrm{d}\Phi(t)$$
$$\leqslant \frac{2B_p^p\varepsilon^p}{(1-\lambda-\varepsilon)^p}\int_{\sigma_p n^{1/p}/\varepsilon} \Phi(t)\,\mathrm{d}P_v\{f_n^* \leqslant \lambda t\}$$
$$\leqslant \frac{2CB_p^p\varepsilon^p}{(1-\lambda-\varepsilon)^p} E_v\Phi(f_n^*).$$

These imply (4.5) for $E_v\Phi(f_n^*) < \infty$ as $2CB_p^p\varepsilon^p/(1-\lambda-\varepsilon)^p < 1$. Since $\Phi(f_n^*) \leqslant \Phi(nd_n^*) \leqslant C^k\Phi(d_n^*/\varepsilon)$ for $\lambda^k n < 1/\varepsilon$ and large f_n^*, the right-hand side of (4.5) is infinity whenever $E_v\Phi(f_n^*) = \infty$. Thus, (4.5) holds in both cases.

The proof of (4.6) is nearly identical, as (4.3) implies

$$P_v\{S_n^* > t, \max(2\tilde{\zeta}_n^*, 2\Delta_n^*, V_{p,n}^*) \leqslant \varepsilon t\} \leqslant P_v\{S_n^* > \lambda t\} B_p^p \frac{\sigma_p^p n + (\varepsilon t)^p}{(t-\lambda t-\varepsilon t)^p}.$$

The details are omitted. □

Proof of Theorem 1. For identically distributed random variables h_j with $E|h_1|^p < \infty$ (e.g. $h_j = \tilde{\zeta}_j$, d_j or $\Delta(X_j; g)$ under stationary measure P_π),

$$E\max_{j\leqslant n}|h_j|^p \leqslant \sqrt{n} + \int_{n^{1/(2p)}}^\infty nP\{|h_1| > t\}\,\mathrm{d}t^p = \mathrm{o}(n).$$

This fact and (2.10) imply (2.11), while (2.10) follows immediately from (4.5) and (4.6) by taking $\Phi(x) = (x^+)^p$ in Theorem 5. □

5. Quick convergence

We shall apply Lemma 1 to obtain the following theorem, which implies the quick convergence of S_n/n to μ for Markov random walks as stated in Theorem 2.

Theorem 6. *Suppose* $g \in \mathscr{H}^*$. *Let* $\Delta_r(x; g)$, d_j, S_n^* *and* f_n^* *be as in* (2.3)–(2.5) *and* (2.7). *Let* $1 \leqslant r \leqslant 2$, $\alpha > 1/r$ *and* $p > 1/\alpha$. *Suppose* $\sup_{j\geqslant 1} E_v\Phi(|d_j|) < \infty$. *Suppose* $|d(g)|^r \in \mathscr{H}^*$ *and that* $\sup_{j\geqslant 0} E_v\Phi_1([\{\Delta_r(X_j; g)\}^-]^{1/r}) < \infty$ *and* $\sup_{j\geqslant 0} E_v\Phi_2([\{\Delta_r(X_j; g)\}^+]^{1/r}) < \infty$, *where* $\Delta_r^\pm(x; g) = \max\{\pm\Delta_r(x; g), 0\}$. *Let* $\varepsilon > 0$. *Then,*

$$\sum_{n=1}^\infty n^{\alpha p-2} P_v\{f_n^* > cn^\alpha\} < \infty, \quad \forall c > 0, \tag{5.1}$$

under one of the following three conditions (5.2), (5.3) *or* (5.4) :

$$v = \pi, \quad \Phi(t) = t^p, \quad \Phi_1(t) = t^{p+\varepsilon}, \quad \Phi_2(t) = t^{\max(p-1/\alpha+\varepsilon, r)}, \tag{5.2}$$

$$v = \pi, \quad \Phi(t) = \Phi_1(t) = \{t\log(1+t)\}^p, \quad \Phi_2(t) = t^{\max(p-1/\alpha, r)}\{\log(1+t)\}^{p_1} \tag{5.3}$$

with π *being the stationary distribution,* $p_1 = 0$ *for* $p - 1/\alpha < r$ *and* $p_1 = p - 1/\alpha + I_{\{p-1/\alpha=r\}}$ *for* $p - 1/\alpha \geqslant r$, *or*

$$\Phi(t) = \Phi_1(t) = t^p\{\log(1+t)\}^{p+1+\varepsilon}, \qquad \Phi_2(t) = t^{\max(p-1/\alpha,r)}\{\log(1+t)\}^{p_2} \tag{5.4}$$

with $p_2 = 0$ *for* $p - 1/\alpha < r$ *and* $p_2 = p - 1/\alpha + 1 + \varepsilon$ *for* $p - 1/\alpha \geqslant r$. *If in addition* $\sup_{j\geqslant 0} E_v\Phi(|\Delta(X_j; g)|) < \infty$, *then*

$$\sum_{n=1}^{\infty} n^{\alpha p-2} P_v\{S_n^* > cn^\alpha\} < \infty, \quad \forall c > 0, \tag{5.5}$$

under either (5.2) *or* (5.3) *or* (5.4).

Proof. We shall only prove (5.1) as the proof of (5.5) is nearly identical. We may assume $c = 1$ without loss of generality, as $\Delta(x; g/c) = \Delta(x; g)/c$ for $c > 0$ and we may divide both f_n^* and cn^α by c in (5.1). Letting $t = 0$, $s = \varepsilon = n^\alpha/(2k_n)$ and $v = v_n$ in (4.2), we find

$$P_v\{f_n^* > n^\alpha\} \leqslant P_v\{d_n^* > n^\alpha/(2k_n)\} + P_v\{V_{r,n}^* > v_n\} + \left[B_r^r \frac{n\sigma_r^r + v_n^r}{n^{\alpha r}/(2k_n)^r}\right]^{k_n}. \tag{5.6}$$

The sum involving the last term of (5.6) converges in the following two cases:

$$\sum_{n=1}^{\infty} n^{\alpha p-2}\left[B_r^r \frac{n\sigma_r^r + v_n^r}{n^{\alpha r}/(2k_n)^r}\right]^{k_n} < \infty, \quad k_n = M\log n, \;\; v_n = n^\alpha/(3k_n) \tag{5.7}$$

for $M \geqslant \alpha p\{\log(3/2)\}^{-1}$; or

$$\sum_{n=1}^{\infty} n^{\alpha p-2}\left[B_r^r \frac{n\sigma_r^r + v_n^r}{n^{\alpha r}/(2k_n)^r}\right]^{k_n} < \infty, \quad k_n = k, \;\; v_n = n^\beta \tag{5.8}$$

for some $0 < \beta < \alpha$, β depending on (α, p, ε) with sufficiently small $\alpha - \beta$, and a fixed sufficiently large k depending on (α, β, p). We shall choose k_n and v_n in (5.7) under (5.3) and (5.4), and choose those in (5.8) under (5.2).

For the choice of k_n and v_n in (5.7), there exists $M' < M'' < \infty$ such that

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{\alpha p-2} P_v\{d_n^* > n^\alpha/(2k_n)\} \\
&\quad\leqslant \sum_{n=1}^{\infty} n^{\alpha p-2}\sum_{j=1}^{n} P_v\{(2M)|d_j| > n^\alpha/\log n\} \\
&\quad\leqslant \sum_{j=1}^{\infty} E_v \sum_{n=j}^{\infty} n^{\alpha p-2} I\{n < h_j^{1/\alpha}\}, \quad h_j = M'|d_j|\log(1+|d_j|) \\
&\quad\leqslant M' \sum_{j=1}^{\infty} E_v h_j^{(\alpha p-1)/\alpha} I\{h_j^{1/\alpha} > j\} \\
&\quad\leqslant M'' \sum_{j=1}^{\infty} \frac{E_v h_j^p \{\log(1+h_j^{1/\alpha})\}^{1+\varepsilon}}{j\{\log(1+j)\}^{1+\varepsilon}}.
\end{aligned} \tag{5.9}$$

If (5.4) holds, then the right-hand side above is finite, as $h_j^p\{\log(1+h_j^{1/\alpha})\}^{1+\varepsilon}\leqslant 1+M''\Phi(|d_j|)$ and $\sum_j j^{-1}\{\log(1+j)\}^{-1-\varepsilon}<\infty$. If (5.3) holds, then $v=\pi$ and d_j are identically distributed, so that (5.9) is bounded by

$$M'E_\pi h_1^{(\alpha p-1)/\alpha}\sum_{j=1}^\infty I\{h_1^{1/\alpha}\geqslant j\}\leqslant M''Eh_1^p\leqslant (M'')^2\{1+E\Phi(|d_1|)\}<\infty.$$

If (5.2) holds and $k_n=k$ and v_n are as in (5.8), then

$$\sum_{n=1}^\infty n^{\alpha p-2}P_\pi\{d_n^*>n^\alpha/(2k_n)\}\leqslant\sum_{n=1}^\infty n^{\alpha p-1}P_\pi\{(2k|d_1|)^{1/\alpha}>n\}\leqslant M'E_\pi|d_1|^p.$$

Thus, under either (5.2) or (5.3) or (5.4)

$$\sum_{n=1}^\infty n^{\alpha p-2}P_v\{d_n^*>n^\alpha/(2k_n)\}<\infty. \tag{5.10}$$

Let $\tilde V_n=\tilde V_{r,n}=\max_{1\leqslant j\leqslant n}E_v[\Delta_r^+(X_n;g)|X_j]$. By (2.8) and (2.9) $V_{r,n}^*$ is no greater than $\max_{0\leqslant j\leqslant n}(\Delta_r^-(X_j;g))^{1/r}+\tilde V_n^{1/r}$, so that by (5.6)–(5.8) and (5.10) it suffices to show

$$\sum_{n=1}^\infty n^{\alpha p-2}P_v\left\{\max_{0\leqslant j\leqslant n}(\Delta_r^-(X_j;g))^{1/r}>v_n/2\right\}<\infty \tag{5.11}$$

and

$$\sum_{n=1}^\infty n^{\alpha p-2}P_v\{\tilde V_n^{1/r}>v_n/2\}<\infty. \tag{5.12}$$

The proof of (5.11) is nearly identical to that of (5.10) and omitted. If (5.4) holds and (k_n,v_n) are as in (5.7), then for large constant C the series in (5.12) is bounded by

$$\sum_{n=1}^\infty n^{\alpha p-2}P_v\{\max[C,\Phi_2(\tilde V_n^{1/r})]>\Phi_2(n^\alpha/(6B_rk_n))\}$$
$$\leqslant\left\{\sum_{n=1}^\infty\frac{n^{\alpha p-2}}{\Phi_2(n^\alpha/(6B_rk_n))}\right\}\sup_{n\geqslant 1}E_v\max\left[C,\Phi_2(\{\Delta_r^+(X_n;g)\}^{1/r})\right]<\infty$$

by the Doob inequality, since $\tilde V_n$ is the maximum of the martingale $\{E_v[\Delta_r^+(X_n;g)|X_j]\}$ and $\max\{C,\Phi_2(x^{1/r})\}$ is convex in x on $[0,\infty)$ for large C.

Under the stationary measure $v=\pi$, $\tilde V_n^r$ and $\max_{2\leqslant j\leqslant n+1}E_\pi[\Delta_r^+(X_{n+1};g)|X_j]$ are identically distributed and they are stochastically bounded by $\tilde V_{n+1}^r$. Thus, if (5.3) holds and (k_n,v_n) are as in (5.7), then the series in (5.12) is bounded by

$$\lim_{m\to\infty}\sum_{n=1}^m n^{\alpha p-2}P_\pi\{M'\tilde V_m^{1/r}\log\tilde V_m>n^\alpha\}$$
$$\leqslant M''\lim_{m\to\infty}E_\pi\{\tilde V_m^{1/r}\log(1+\tilde V_m)\}^{p-1/\alpha}\leqslant(M'')^2E_\pi\Phi_2((\Delta_r^+(X_0;g))^{1/r})<\infty$$

via the Doob inequality. The proof of (5.12) in the case of (5.2) and (5.8) is simpler than the above with the same arguments and thus omitted. □

Proof of Theorem 2. Since

$$\sum_{n=1}^{\infty} n^{\beta-1} P\left\{\sup_{j\geqslant n}|\theta_j-\mu|\geqslant\varepsilon\right\} \leqslant 2^{\beta-1}\sum_{k=0}^{\infty}2^{\beta k}\sum_{\ell=k}^{\infty}P\left\{\max_{2^\ell\leqslant j<2^{\ell+1}}|\theta_j-\mu|\geqslant\varepsilon\right\}$$

$$\leqslant M'\sum_{\ell=0}^{\infty}2^{\beta\ell}P\left\{\max_{2^\ell\leqslant j<2^{\ell+1}}|\theta_j-\mu|\geqslant\varepsilon\right\}\leqslant (M')^2\sum_{n=1}^{\infty}n^{\beta-1}P\left\{\sup_{j\leqslant n}|\theta_j-\mu|\geqslant\varepsilon\right\}$$

for some universal constant M', (2.11) and (2.12) follow from (5.1) and (5.5) with $\alpha=1$ and $p=\beta+1$ under condition (5.2). □

6. Applications to nonlinear renewal theory

Let $\{S_n(g), n\geqslant 0\}$ be as in (2.4) with the $\{(X_n,\zeta_n), n\geqslant 0\}$ in (1.1) and the stationary probability distribution π. Assume that $E_\pi g(X_0,X_1,\zeta_1)=\mu_g>0$. Under conditions I.1–I.4 of Kesten (1974), or Harris recurrent condition in Alsmeyer (1994), the elementary renewal theorem asserts that

$$\sum_{n=1}^{\infty} P_v\{S_n(g)\leqslant a, X_n\in A\}\sim \pi(A)\mu_g^{-1}a \quad \text{as } a\to\infty. \tag{6.1}$$

Under certain conditions, Kesten (1974) also proved a renewal theorem which provides the limit of $\sum_{n=1}^{\infty}P_v\{S_n(g)\leqslant a, X_n\in A\}-\pi(A)\mu_g^{-1}a$ as $a\to\infty$. Making use of Theorem 6, we provide a nonlinear version of (6.1) with convergence rates in Theorem 7 below.

Let ξ_n be a sequence of random variables and $\{a_\lambda(\cdot),\lambda\in\Lambda\}$ be a family of functions such that for some finite constants $\varepsilon>0$, $b_\lambda=b_{\lambda,\mu}$, $b_*=b_{*,\varepsilon}$, $M_\varepsilon, c_\varepsilon<1/M_\varepsilon$ and $\frac{1}{2}<\gamma\leqslant 1$,

$$\sum_{n=0}^{\infty} n^{-\gamma}P_v\{|\xi_n|>c_\varepsilon n^\gamma\}<\infty, \qquad \sum_{n=0}^{\infty}P_v\{-\xi_n>c_\varepsilon n\}<\infty, \tag{6.2}$$

and for all $\lambda\in\Lambda$ with $b_\lambda\geqslant b_*$

$$\max_{\varepsilon b_\lambda^\gamma\leqslant n\leqslant b_\lambda-\varepsilon b_\lambda^\gamma}\frac{\mu n-a_\lambda(n)}{n^\gamma}<-1/M_\varepsilon, \tag{6.3}$$

$$\max_{b_\lambda+\varepsilon b_\lambda^\gamma\leqslant n\leqslant(1+\varepsilon)b_\lambda}\frac{a_\lambda(n)-\mu n}{n^\gamma}<-1/M_\varepsilon, \qquad \sup_{n\geqslant(1+\varepsilon)b_\lambda}\frac{a_\lambda(n)-\mu n}{n}<-1/M_\varepsilon. \tag{6.4}$$

If a_λ is a constant as in (6.1), then (6.3) and (6.4) hold with $b_\lambda=a_\lambda/\mu$. Conditions (6.3) and (6.4) also hold for $a_\lambda(n)=\lambda\sqrt{n}$ with $b_\lambda=(\lambda/\mu)^2$.

For measurable functions $g(x,y,\zeta)$ and $g_0(x,y,\zeta)$, define

$$U_\lambda=U_{\lambda,g,g_0}=\sum_{n=1}^{\infty}g_0(X_{n-1},X_n,\zeta_n)I_{\{S_n(g)+\xi_n\leqslant a_\lambda(n)\}}$$

and define $T_\lambda=T_{\lambda,g}$ and $N_\lambda=N_{\lambda,g}$ by

$$T_\lambda=\inf\{n\geqslant 1: S_n(g)+\xi_n>a_\lambda(n)\},\quad N_\lambda=1+\sup\{n\geqslant 1: S_n(g)+\xi_n\leqslant a_\lambda(n)\}.$$

Theorem 7. *Let* $\{g, g_0\} \subset \mathscr{H}^*$ *with* $\|g_0\|_\infty < \infty, \mu = E_\pi g(X_0, X_1, \zeta_1)$ *and* $\mu_0 = E_\pi g_0(X_0, X_1, \zeta_1)$. *Suppose the* (6.2)–(6.4) *hold with* $\frac{1}{2} < \gamma \leqslant 1$. *Suppose the conditions of Theorem* 6 *hold with* g *and* $(p, \alpha) = (2/\gamma - 1, \gamma)$ *for* $\gamma < 1$ *and also with* g *and* $(p, \alpha) = (2, 1)$. *Then, for large* b_λ

$$E_v|U_{\lambda,g,g_0} - S_{n_\lambda}(g_0)| = O(b_\lambda^\gamma), \tag{6.5}$$

where $n_\lambda = n_{\lambda,\mu}$ *is the integer part of* b_λ *in* (6.3) *and* (6.4). *If* $E_v\Delta(X_n; g_0) = O(n^\gamma)$, *then*

$$E_v U_{\lambda,g,g_0} = \mu_0 b_\lambda + O(b_\lambda^\gamma). \tag{6.6}$$

If (6.3) *and* (6.4) *hold for all* $\varepsilon > 0$ *and* $E_v\Delta(X_n; g_0) = o(n^\gamma)$, *then the* $O(b_\lambda^\gamma)$ *in* (6.5) *and* (6.6) *can be replaced by* $o(b_\lambda^\gamma)$. *Furthermore, the above assertions hold under respective conditions when* U_{λ,g,g_0} *is replaced by* $S_{T_\lambda}(g_0)$ *or* $S_{N_\lambda}(g_0)$.

Remark. Expansion (6.6) implies (6.1) with $g_0(x, y, \zeta) = I_{\{y \in A\}}$ and $\mu_0 = \pi(A)$.

Proof. We shall only prove (6.5) and (6.6) with $\gamma < 1$ as the rest of the proof [for $S_{T_\lambda}(g_0)$, $S_{N_\lambda}(g_0)$ or $\gamma = 1$] is nearly identical. Splitting the sums U_{λ,g,g_0} and $S_{n_\lambda}(g_0)$ into four parts $[1, n_1]$, $(n_1, n_2]$, $(n_2, n_3]$ and (n_3, ∞) with n_1, n_2 and n_3 being the integer parts of $\varepsilon b_\lambda^\gamma$, $b_\lambda - \varepsilon b_\lambda^\gamma$ and $b_\lambda + \varepsilon b_\lambda^\gamma$, respectively, we find

$$|U_{\lambda,g,g_0} - S_{n_\lambda}(g_0)| \leqslant \|g_0\|_\infty \Bigg(n_1 + \sum_{n_1 < n \leqslant n_2} I\{S_n(g) + \xi_n > a_\lambda(n)\} + (n_3 - n_2) + \sum_{n > n_3} I\{S_n(g) + \xi_n \leqslant a_\lambda(n)\} \Bigg).$$

Since $\mu n - a_\lambda(n) \leqslant -n^\gamma/M_\varepsilon$ for $\varepsilon b_\lambda^\gamma \leqslant n \leqslant b_\lambda - \varepsilon b_\lambda^\gamma$, by (5.5) for $(p, \alpha) = (2/\gamma - 1, \gamma)$ (thus, $p\alpha - 2 = -\gamma$) and by (6.2) we have

$$\sum_{n_1 < n \leqslant n_2} P_v\{S_n(g) + \xi_n > a_\lambda(n)\} \leqslant b_\lambda^\gamma \sum_{n_1 < n \leqslant n_2} n^{-\gamma} P_v\{S_n(g) - \mu n + \xi_n > n^\gamma/M_\varepsilon\} = o(b_\lambda^\gamma).$$

Similarly, since $a_\lambda(n) - n\mu \leqslant -n^\gamma/M_\varepsilon$ for $b_\lambda - \varepsilon b_\lambda^\gamma < n \leqslant (1+\varepsilon)b_\lambda$,

$$\sum_{n_3 < n \leqslant n_4} P_v\{S_n(g) + \xi_n \leqslant a_\lambda(n)\} \leqslant (1+\varepsilon)^\gamma b_\lambda^\gamma \sum_{n_3 < n \leqslant n_4} n^{-\gamma} P_v\{S_n(g) - \mu n + \xi_n \leqslant -n^\gamma/M_\varepsilon\} = o(b_\lambda^\gamma),$$

where n_4 is the integer part of $(1+\varepsilon)b_\lambda$. Finally, since $a_\lambda(n) - n\mu \leqslant -n/M_\varepsilon$ for $n > (1+\varepsilon)b_\lambda$, by Theorem 6 with $(p, \alpha) = (2, 1)$ and (6.2) we have

$$\sum_{n > n_4} P_v\{S_n(g) + \xi_n \leqslant a_\lambda(n)\} \leqslant \sum_{n > n_4} P_v\{S_n(g) - \mu n + \xi_n \leqslant -n/M_\varepsilon\} = o(1).$$

Putting these inequalities together, we obtain (6.5), as $n_1 \leqslant \varepsilon b_\lambda^\gamma$ and $n_3 - n_2 \leqslant 1 + 2\varepsilon b_\lambda^\gamma$. By (2.6) of Theorem 1 and the condition on $E_v\Delta(X_n; g_0)$, (6.6) follows from

$$E_v S_{n_\lambda}(g_0) = n_\lambda \mu_0 + E_v\Delta(X_{n_\lambda}; g_0) - E_v\Delta(X_0; g_0) = \mu_0 b_\lambda + O(b_\lambda^\gamma). \qquad \square$$

7. Random coefficient autoregression models

Let $\{x_k,\ k\geqslant 1\}$ be a sequence that satisfies a first-order random coefficient autoregression model

$$x_k = \beta_k x_{k-1} + \varepsilon_k, \quad x_0 = 0, \tag{7.1}$$

where $(\beta_k)_{k\geqslant 1}$ is a sequence of i.i.d. random variables with $E\beta_k = \beta$ and $\operatorname{Var}\beta_k = \kappa^2$, and $(\varepsilon_k)_{k\geqslant 1}$ is a sequence of i.i.d. random variables with $E\varepsilon_k = 0$ and $\operatorname{Var}\varepsilon_k = \sigma^2$. Further, we assume that $(\beta_k)_{k\geqslant 1}$ and $(\varepsilon_k)_{k\geqslant 1}$ are independent.

Under the normality assumption on (β_k, ε_k) with known (σ^2, κ^2), the log-likelihood ratio statistics Z_n for testing β_0 are given by

$$Z_n = \frac{C_n^2}{2D_n}, \quad C_n = \sum_{i=1}^n \frac{x_{i-1}x_i}{\sigma^2 + \kappa^2 x_{i-1}^2}, \quad D_n = \sum_{i=1}^n \frac{x_{i-1}^2}{\sigma^2 + \kappa^2 x_{i-1}^2}. \tag{7.2}$$

The stopping time of the repeated significance test for $\beta = 0$ is given by

$$T_\lambda = \inf\{n\geqslant 1\colon Z_n > \lambda\}. \tag{7.3}$$

We shall investigate T_λ under the stability assumption $\beta^2+\kappa^2 < 1$, without the normality assumption on $(\beta_\kappa, \varepsilon_\kappa)$. Since $\hat{\beta}_n = C_n/D_n$ are the least squares estimates of β, Z_n and T_λ can be used to test $\beta = 0$ without the normality assumption.

In order to apply the nonlinear renewal theory in Theorem 7 to approximate the expected sample size, we first note that the random coefficient autoregression model (7.1) is w-uniformly ergodic with $w(x) = |x|^2$ (cf. Theorem 16.5.1 of Meyn and Tweedie, 1993). Since C_n and D_n are additive functionals of the Markov chains (x_{n-1}, x_n), by the strong law of large numbers, $D_n \to E_\pi D_1 := \mu_D$ and $C_n/n \to E_\pi C_1 := \mu_C = \beta\mu_D$. Taking Taylor expansion as in Melfi (1992), we find

$$Z_n = nh(C_n/n, D_n/n) = \sum_{i=1}^n g(x_{i-1}, x_i) + \xi_n, \tag{7.4}$$

with $h(x,y) = x^2/(zy), g(x_{i-1}, x_i) = \beta x_{i-1}x_i - \beta^2 x_{i-1}^2/2)/(\sigma^2 + x^2 x_{i-1}^2)$ and

$$\xi_n = n(C_n/n - \mu_C, D_n/n - \mu_D)h^{(2)}(C_n/n - \mu_C, D_n/n - \mu_D)^{\mathrm{tr}}/2, \tag{7.5}$$

where $h^{(2)}$ is the matrix of the second partial derivatives of $h(x,y)$ at certain point between $(C_n/n, D_n/n)$ and (C, D), and v^{tr} is the transpose of v.

Assume $\beta \neq 0$. It can be verified with simple calculation that conditions (6.2)–(6.4) hold with $\gamma = 1$. The drift criterion of Theorem 17.4.2 in Meyn and Tweedie (1993) implies that $E_\nu \Delta(X_n; g) \leqslant CE_\nu X_1^2 < \infty$, for some constant C. By Theorem 7, we have

$$E_\nu(T_\lambda) = \lambda/\mu_0 + o(\lambda) \quad \text{as } \lambda \to \infty, \quad \mu_0 = E_\pi g(x_0, x_1) = \beta^2\mu_D/2 > 0. \tag{7.6}$$

Acknowledgements

The research of Cheng-Der Fuh was partially supported by the NSC 87-2118-M-001 -003. The research of Cun-Hui Zhang was partially supported by the National Science

Foundation of USA and the National Security Agency of USA. The authors also would like to thank the referees, who provided many helpful comments on an earlier draft.

References

Alsmeyer, G., 1990. Convergence rates in the law of large numbers for martingales. Stochastic Process. Appl. 36, 181–194.

Alsmeyer, G., 1994. On the Markov renewal theorem. Stochastic Process. Appl. 50, 37–56.

Burkholder, D., 1973. Distribution function inequalities for martingales. Ann. Probab. 1, 19–42.

Burkholder, D., Gundy, R.F., 1970. Extrapolation and interpolation of quasi-linear operators on martingales. Acta Math. 124, 249–304.

Chow, Y.S., Lai, T.L., 1975. Some one-sided theorems on the tail distribution of sample sums with applications to the last time and largest excess of boundary crossings. Trans. Amer. Math. Soc. 208, 51–72.

Chow, Y.S., Lai, T.L., 1978. Paley-type inequality and convergence rates related to the law of large numbers and extended renewal theory. Z. Wahrs. Verw. Gebie. 45, 1–19.

Fuh, C.D., Lai, T.L., 1998. Wald's equations, first passage times and moments of ladder variables in Markov random walks. J. Appl. Probab. 35, 566–580.

Hoffmann-Jorgensen, J., 1974. Sums of independent Banach space valued random variables. Studia Math. 52, 159–189.

Hoffmann-Jorgensen, J., 1977. Probability in B-spaces. Lect. Notes Math. 598, 2–119.

Irle, A., 1990. Largest excess of boundary crossings for martingales. Metrika 34, 107–115.

Irle, A., 1993. r-quick convergence for regenerative processes with applications to sequential analysis. Stochastic Process. Appl. 45, 319–329.

Kesten, H., 1974. Renewal theory for functionals of a Markov chain with general state space. Ann. Probab. 2, 355–386.

Lai, T.L., 1977. Convergence rates and r-quick versions of the strong law for stationary mixing sequences. Ann. Probab. 5, 693–706.

Lai, T.L., 1981. Asymptotic optimality of invariant sequential probability ratio tests. Ann. Statist. 9, 318–333.

Melfi, V.F., 1992. Nonlinear Markov renewal theory with statistical applications. Ann. Probab. 20, 753–771.

Meyn, S.P., Tweedie, R.L., 1993. Markov Chains and Stochastic Stability. Springer, New York.

Ney, P., Nummelin, E., 1987. Markov additive processes I. Eigenvalue properties and limit theorems. Ann. Probab. 15, 561–592.

Peligrad, M., 1985. Convergence rates of the strong law for stationary mixing sequences. Z. Wahrs. Verw. Gebie. 70, 307–314.

Rosenthal, H.P., 1970. On the subspaces of $L_p(p > 2)$ spanned by sequences of independent random variables. Israel J. Math. 8, 273–303.

Strassen, V., 1967. Almost sure behavior of sums of independent random variables and martingales. Proceedings of the Fifth Berkeley Symposium Math. Statist. and Probability, pp. 315–343.