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

Cheng-Der Fuh\(^*\), 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.

MSC: primary 60G40; secondary 60J10

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

1. Introduction

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

\[
P(\{X_{n+1}, \zeta_{n+1}\} \in A \times B | X_n = x, (X_{k-1}, \zeta_k), 1 \leq k \leq n) = P(x, A \times B)
\]

(1.1)
do not depend on the “time” \( n \geq 0 \) and the values of \( (X_{k-1}, \zeta_k) \), \( 1 \leq k \leq n \), for all \( x \in D, A \in \mathcal{A} \) and Borel sets \( B \). For measurable functions \( g \) on \( D \times \mathbb{R} \), the chain \( \{X_n, S_n\}, n \geq 0\) \( \), 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 \( \pi, \pi(A) = \int P(x, A) \pi(dx) \) for all \( A \in \mathcal{A} \).

* Corresponding author.

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

0304-4149/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.

PII: S0304-4149(99)00104-0
The concept of quick convergence was introduced by Strassen (1967). A sequence of random variables $\theta_n$ is said to converge $\beta$-quickly ($\beta > 0$) to a constant $\mu$ if

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

Obviously $\theta_n \to \mu$ $\beta$-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 and sequential analysis (cf. Lai, 1981). Extensions of (1.2) to Markov random walks are of independent interest and have applications in other areas of research.

We shall consider quick convergence of $\theta_n = S_n/n$, $n \geq 1$, for Markov random walks $\{X_n, S_n\}$, $n \geq 0$. 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 $\{(X_n, \zeta_n), n \geq 1\}$ be a Markov chain as in (1.1). Let $v$ be an initial distribution of $X_0$ and define $v^* (A) = \sum_{n=0}^{\infty} P_n (X_n \in A)$ on $\mathcal{A}$. Let $E_v$ be the expectation under which $X_0$ has initial distribution $v$, and $E_v$ be the conditional expectation given $X_0 = x$ (i.e. with $v$ degenerate at $x$). Let $\mathcal{H} = \mathcal{H}_v$ be the class of all measurable functions $g = g(x, y, \zeta)$ on $\mathcal{D} \times \mathcal{D} \times \mathbb{R}$ such that $E_v [g(x, X_1, \zeta_1)] < \infty$ $v^*$-almost surely and $E_v |g(X_0, X_1, \zeta_1)| < \infty$. For $g \in \mathcal{H}$, define operators $P$ and $P_n$ by $(P g)(x) = E_x [g(x, X_1, \zeta_1)]$ and $(P_n g)(x) = E_x [g(X_0, X_1, \zeta_1)]$ respectively, and set $\tilde{g} = P g$. We shall consider solutions $A = A(x; g) \in \mathcal{H}$ of the Poisson equation

$$(P - I) A = (I - P_n) \tilde{g} \quad v^* \text{ - a.e., } P_n A = 0, \tag{2.1}$$

where $I$ is the identity. Here and in the sequel a function $f$ defined on $\mathcal{D}$ is always treated as a function on $\mathcal{D} \times \mathcal{D} \times \mathbb{R}$ via $f(y) = g(x, y, \zeta)$, so that $P f = \int f(y) P(x, dy)$. Since $(I - P_n) \tilde{g} = (P - P_n) g$, the Poisson equation (2.1) can also be written as

$$(P - I) A = (P - P_n) g. \tag{2.1}$$

Since $(I - P_n) \tilde{g} = (P - P_n) g$, the Poisson equation (2.1) can also be written as

$$E_v [\Delta(X_n; g)] < \infty \quad \text{and} \quad E_v [g(X_{n-1}, X_n, \zeta_n)] < \infty \text{ for all } n \geq 0.$$
Furthermore, set $\sigma_p = \sigma_p(g) = \{E_x|d(X_0, X_1, \zeta); g)|^p\}^{1/p}$, and for $|d(g)|^p \in \mathcal{H}^*$ define

$$D_p(g) = D_p(x, y, \zeta, g) = \{d(x, y, \zeta, g)\}^{1/p} - \sigma_p^p(g) - \Delta_p(y, g) + \Delta_p(x, g), \quad (2.3)$$

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

$$S_n = S_n(g) = \sum_{j=1}^{n} \xi_{ij}, \quad \xi_{ij} = g(X_{i-1}, X_j, \zeta_j), \quad d_j = d_j(g) = d(X_{i-1}, X_j, \zeta_j; g), \quad (2.4)$$

$$S_n^* = S_n^*(g) = \max_{1 \leq j \leq n} |S_j(g) - j\mu_n|, \quad g_n^* = \max_{1 \leq j \leq n} |g(X_{i-1}, X_j, \zeta_j) - \mu_p|, \quad (2.5)$$

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

and

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

Furthermore, define

$$\Delta_n^* = \Delta_n^*(g) = \max_{0 \leq i < j \leq n} |\Delta(X_j; g) - \Delta(X_i; g)|, \quad (2.8)$$

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

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

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

$$\max\{E_x(f_n^p), E_x(S_n^p)\} \leq C_p E_x\{\sigma\sqrt{n}, g_n^*, \Delta_n^*, V_{2,n}\}^p. \quad (2.10)$$

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

$$\max\{E_x(f_n^p), E_x(S_n^p)\} \leq C_p \{(\sigma\sqrt{n})^p + \lambda(1)n^{p/2}\}, \quad \text{as} \quad n \to \infty. \quad (2.11)$$

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

Let $\theta_n$ be as in (1.2) and define $N_n = \sup\{n \geq 1: |\theta_n - \mu| \geq \epsilon\}$. Since (1.2) is equivalent to $E\xi^p < \infty$, the sequence $\{\theta_n\}$ converges $\beta$-quickly to $\mu$ if and only if

$$\sum_{n=1}^{\infty} n^{\beta-1} P\{N_n \geq n\} = \sum_{n=1}^{\infty} n^{\beta-1} P\left\{\sup_{j \geq n} |\theta_j - \mu| \geq \epsilon\right\} < \infty, \quad \forall \epsilon > 0. \quad (2.12)$$

**Theorem 2.** Suppose $\{g, d(g)\} \subset \mathcal{H}^*$. Let $\Delta_n(x, g)$, $d_j$, $S_n^*$ and $f_n^*$ be as in (2.3)–(2.5) and (2.7). Set $\Delta_n^2(x, g) = \max\{\pm \Delta_n(x, g), 0\}$. Let $1 < r < 2$ and $\beta > 0$. Suppose
In Section 5, we provide in Theorem 6 sufficient moment conditions for the quick convergence of $\theta_n = n^{-\beta} (S_n(g) - n\mu(g))$ to 0, under $P_t$ 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 $A_n^\beta$ are weaker with $\epsilon = 0$ 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 $(P - I)A = (I - P_v)g$ 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 $A(\cdot; g)$ for bounded $g = P g$ (cf. Fuh and Lai, 1998); therefore, the moment conditions $A(x; g)$ in Theorems 1 and 2 are automatically satisfied in this case. Also, the moment conditions on $A_n^\beta$ 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 $A(\cdot; g)$, 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 $E[|\zeta_1 - \mu|^q] < \infty$ 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 \( g \in \mathcal{M}^* \). Then, the sequence \( \{f_n\} \) in (2.6) is a martingale with respect to \( \{\mathcal{F}_n\} \) under \( P_f \), with \( d_j \) being the martingale differences.

(ii) Suppose \( \{g, d^2(g)\} \subset \mathcal{M}^* \). Then both sequences

\[
F_n(g) = f_n^2(g) - n\sigma^2_n(g) - \Delta_2(X_n; g) + \Delta_2(X_0; g), \quad n \geq 1
\]

(3.1)

and

\[
F'_n(g) = (S_n - n\mu_g - \Delta(X_n; g))^2 - n\sigma^2_n(g) - \Delta_2(X_n; g) + \Delta_2(X_0; g), \quad n \geq 1
\]

(3.2)

are \( \mathcal{F}_n \)-martingales under \( P_f \). Moreover,

\[
\sigma^2_n = \sigma^2_n(g) = P_f d^2(g) = \text{Var}_n(\tilde{z}_1) + 2 \sum_{j=1}^{\infty} \text{Cov}_n(\tilde{z}_1, \tilde{z}_{1+j}).
\]

(3.3)

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

\[ E[d_j|X_{j-1}] = (Pg - P_2g - PA + \Delta)(X_{j-1}) = (\tilde{g} - P_2\tilde{g} - PA + \Delta)(X_{j-1}) = 0 \]

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

\[
\sum_{j=1}^{n} D_2(X_{j-1}, X_j, \tilde{z}_j; g) = \sum_{j=1}^{n} d_j^2 - n\sigma^2 - \Delta_2(X_n; g) + \Delta_2(X_0; g)
\]

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 \( \mu = \mu_g \), \( \sigma^2 = \sigma^2_n(g) \), and \( S_n = S_n(g) \) be as in (2.2)–(2.4). Let \( T \) be a bounded stopping rule with respect to \( \{\mathcal{F}_n, n \geq 0\} \). Suppose \( g \in \mathcal{M}^* \). Then

\[
E_s \Delta \sigma^2 = \mu E_s T + E_s \{\Delta(X_T; g) - \Delta(X_0; g)\}.
\]

(3.4)

If in addition \( d^2(g) \in \mathcal{M}^* \) and \( E_s \Delta^2(X_n; g) < \infty \) for all \( n \geq 0 \), then

\[
E_s (S_T - \mu T)^2 = \sigma^2 E_s T + 2 \mu E_s \{\Delta(X_T; g)\} + E_s \{\Delta^2(X_T; g) - \Delta^2(X_0; g)\},
\]

(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 \( E_s T < \infty \). (i) Suppose there exists a uniformly bounded \( h = h(x, y) \) such that \( \tilde{g} = |g - h| \in \mathcal{M}^* \) with \( \sup_{x} E_s \Delta(X_T; \tilde{g}) < \infty \).

If \( g \in \mathcal{M}^* \) and \( \{\Delta(X_T; \tilde{g}), n \geq 1\} \) is \( E_s \)-uniformly integrable, then (3.4) holds.
(ii) Suppose \( \{ g, d^2(g) \} \subseteq \mathcal{H}^* \), \( T \) is a stopping rule with \( E, T < \infty \) and that both \( \{ A_2(X_{T \wedge n}; g), \, n \geq 1 \} \) and \( \{ A^2(X_{T \wedge n}; g), \, n \geq 1 \} \) are \( E_i \)-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 \( 2 \times 2 \) 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 \( \{ g, |\tilde{g}_1|, |\tilde{g}_2| \} \subset \mathcal{H}^* \) with \( \sup_n E_i A(X_{T \wedge n}; |\tilde{g}_j|) < \infty, \, j = 1, 2 \). Sufficient conditions for the uniform integrability of \( \{ h(X_{T \wedge n}) \} \), such as \( \{ A(X_{T \wedge n}; g) \} \), are given in Proposition 1 below. Under the uniform ergodicity condition, \( A(x; g) \) are bounded functions of \( x \) for bounded \( \tilde{g} = Pg \).

**Proof.** (i) Let \( S_n(g) \) be as in (2.4). By (3.4),
\[
E_i S_{T \wedge n}(\tilde{g}) = \mu_2 E_i \{ T \wedge n \} + E_i A(X_{T \wedge n}; \tilde{g}) - E_i A(X_0; \tilde{g}) = O(1),
\]
so that \( E_i S_T(\tilde{g}) < \infty \). Since \( h \) is uniformly bounded and \( \tilde{g} = |g - h| \), \( E_i S_T(|g|) \leq E_i S_T(\tilde{g}) + \|h\|_\infty E_i T < \infty \), so that \( \{ S_{T \wedge n}(g) \} \) is uniformly integrable. Since \( \{ A(X_{T \wedge n}; g) \} \) 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 \( \mu_2 = 0 \). By (3.1) and the optional stopping theorem,
\[
E^*_m \left( \sum_{j=T \wedge m}^{T \wedge n} d_j \right)^2 = E^*_m (S_{T \wedge n} - S_{T \wedge m} - A(X_{T \wedge n}; g) + A(X_{T \wedge m}; g))^2
\]
\[
= \sigma^2 E^*_m ((T \wedge n) - (T \wedge m))^+ + E^*_m (A_2(X_{T \wedge n}; g) - A_2(X_{T \wedge m}; g))^2
\]
for \( n \geq m \), where \( E^*_m \) is the conditional expectation given \( \mathcal{F}_{T \wedge m} \). Thus, \( \|S_{T \wedge n} - S_{T \wedge m}\|_2 \leq \sigma \|\mathbb{I}(T \wedge n) - (T \wedge m)\|_1^{1/2} + \|A_2(X_{T \wedge n}; g) - A_2(X_{T \wedge m}; g)\|_1^{1/2} + \|A(X_{T \wedge n}; g) - A(X_{T \wedge m}; g)\|_2 \to 0 \), as \( n \to \infty \) by the uniform integrability assumptions. Since \( \{ A^2(X_{T \wedge n}; g) \} \) 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 \( E_i |h(X_n)|, \, n \geq 1 \). Let \( T \) be an integer-valued random variable with \( E_i T < \infty \). If \( E_i |h(X_T)| < \infty \) and
\[
\lim_{M \to \infty} \limsup_{n \to \infty} E_i (|h(X_n)| - Mn)^+ = 0,
\]
then \( \{ h(X_{T \wedge n}) \} \) is \( E_i \)-uniformly integrable. Consequently, \( \{ h(X_{T \wedge n}) \} \) is \( E_i \)-uniformly integrable if
\[
\sum_{n=1} \infty E_i (|h(X_n)| - n)^+ < \infty.
\]

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

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

4. Moment inequalities

The type of tail probability inequalities under consideration here for Markov random walks can be called “good-$\chi$” 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), Hoffman-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 $g \in \mathcal{H}^*$ with $\|d(g)\|^p \in \mathcal{H}^*$ for some $p \geq 1$. Define

$$\bar{\zeta}_p^* = \max_{1 \leq j \leq n} \bar{\zeta}_j^* = \max_{1 \leq j \leq n} |\bar{\zeta}_j - \mu_g|,$$

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

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

$$P_v \{f_n^* > t + k(\varepsilon + s), d_n^* \leq s, V_n^* \leq v\} \leq P_v \{f_n^* > t\} B_{p^*}^k \left( \frac{n\sigma_0^p + v^p}{\varepsilon^p} \right)^k$$

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

$$P_v \{S_n^* > t + k(\varepsilon + s + u), \bar{\zeta}_n^* \leq s, A_n^* \leq u, V_n^* \leq v\} \leq P_v \{S_n^* > t\} B_{p^*}^k \left( \frac{n\sigma_0^p + v^p}{\varepsilon^p} \right)^k.$$

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

$$\tau_k = \inf \left\{ j : \tau_{k-1} < j \text{ and } \sum_{i=1}^j \|d_i\| > \varepsilon \right\}.$$

On the event $d_n^* \leq s$, the overshoot for the stopping rules $\tau_j$ is bounded by $s$ for all $j \geq 0$, so that $f_n^*_{\tau_k} \leq t + (k - 1)\varepsilon + ks$ and the left-hand side of (4.2) is bounded by
$P_r \{ \tau_k \leq n, d_n^* \leq s, V_{p,n} \leq v \}$ for all $k \geq 1$. Therefore, it suffices to prove by induction that for all $k \geq 1$

$$P_r \{ A_k \} \leq P_r \{ A_{k-1} \} B_{p}^p \left( \frac{n \sigma^p + v^p}{e^p} \right),$$

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

$$E^* I_{A_k} \leq I_{A_{k-1}} E^* \left[ \frac{|f_n - f_{\tau_k-1}|^p}{e^p} \right] \leq I_{A_{k-1}} B_{p}^p E^* \sum_{j=\tau_k+1}^n |d_j|^p.$$  

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

$$E^* \sum_{j=\tau_k+1}^n |d_j|^p = \sigma^p n - E^* \tau_{k-1} + E^* \{ A_p(X_n^*; g) - A_p(X_{\tau_k-1}^*; g) \}$$

$$\leq \sigma^p n + V_{p,\tau_k-1,n}^p,$$

in view of (2.9). These imply (4.4), and therefore (4.2), as $V_{p,\tau_k-1,n} \leq v$ 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 $\{ S_n^* > t+k(e+s+u), V_p^\tau_{\tau_k} \leq s, V_{p,n}^* \leq u \}$, we have $S_{\tau_k} \leq t+s$ and

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

which imply $\tau_k \leq n$ as $\max_{\tau_k < j \leq n} \left| \sum_{i=\tau_k+1}^j d_i > k(e+s+u) \right|$ and $d_n^* \leq s+u$. □

Next, we apply Lemma 1 to obtain moment inequalities.

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

$$E_r \Phi(f_n^\epsilon) \leq \frac{E_r \Phi(\epsilon^{-1} \max(\sigma_p n^{1/p}, d_n^*), V_{p,n}^*)}{1 - 2CB_p^p e^p / (1 - \lambda - \epsilon)^p}$$

and

$$E_r \Phi(S_n^\epsilon) \leq \frac{E_r \Phi(\epsilon^{-1} \max(\sigma_p n^{1/p}, 2V_{p,n}^\epsilon, 2V_{p,n}^*), V_{p,n}^*)}{1 - 2CB_p^p e^p / (1 - \lambda - \epsilon)^p}.$$  

**Proof.** Clearly, $E_r \Phi(f_n^\epsilon) = \int_0^\infty P_r \{ f_n^\epsilon > t \} \, d\Phi(t)$ is bounded by

$$E_r \Phi(f_n^\epsilon) \leq \Phi(\sigma_p n^{1/p} / \epsilon) + \int_{\sigma_p n^{1/p} / \epsilon}^\infty P_r \{ \max(d_n^*, V_{p,n}^*) > at \} \, d\Phi(t)$$

$$+ \int_{\sigma_p n^{1/p} / \epsilon}^\infty P_r \{ f_n^\epsilon > t, \max(d_n^*, V_{p,n}^*) \leq at \} \, d\Phi(t).$$

The sum of the first two terms on the right-hand side is $E_r \Phi(\epsilon^{-1} \max(\sigma_p n^{1/p}, d_n^*), V_{p,n}^*)$, the numerator on the right-hand side of (4.5), while by (4.2), with $(1, \lambda, t-h, \epsilon, at, at)$
taking place of \((k, t, \varepsilon, s, v)\) in (4.2), the third term above on the right-hand side is bounded by

\[
\int_{\eta_n}^\infty P_t \{ f_n^* > \lambda t \} \frac{B_{p}(\sigma_{p}^{\nu} n + (\lambda t)^{p})}{(\lambda t - \nu t)^{p}} \, d\Phi(t)
\]

\[
\leq \frac{2B_{p}^{p}\nu}{(1 - \lambda - \varepsilon)^{p}} \int_{\eta_n}^\infty P_t \{ f_n^* > \lambda t \} \, d\Phi(t)
\]

\[
\leq \frac{2B_{p}^{p}\nu}{(1 - \lambda - \varepsilon)^{p}} \int_{\eta_n}^\infty \Phi(t) \, dP_t \{ f_n^* \leq \lambda t \}
\]

\[
\leq \frac{2CB_{p}^{p}\nu}{(1 - \lambda - \varepsilon)^{p}} E_t \Phi(f_n^*).
\]

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

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

\[
P_t \{ S_n > t, \max(2n^{\nu}, 2A_n^{\nu}, V_{p,n}^{\nu}) \leq \nu t \} \leq P_t \{ S_n > \lambda t \} B_{p}^{p}\nu n + (\lambda t)^{p} \, d\Phi(t).
\]

The details are omitted. \(\square\)

**Proof of Theorem 1.** For identically distributed random variables \(h_j\) with \(E|h_1|^p < \infty\) (e.g. \(h_j = \xi_j\), \(d_j\) or \(\Lambda(X; g)\)) under stationary measure \(P_n\),

\[
E \max_{j \leq n} |h_j|^p \leq \sqrt{n} + \int_{|h| > t} nP \{|h| > t\} \, dt^p = o(n).
\]

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. \(\square\)

5. Quick convergence

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

**Theorem 6.** Suppose \(g \in \mathcal{H}^\nu\). Let \(\Lambda_r(x; g), d_j, S_n^\nu, \text{ and } f_n^\nu\) be as in (2.3)–(2.5) and (2.7). Let \(1 \leq r \leq 2, \alpha > 1/\nu\) and \(p > 1/\alpha\). Suppose \(\sup_{j \geq 1} E_t \Phi(|d_j|) < \infty\). Suppose \(|d(g)|^r \in \mathcal{H}^\nu\) and that \(\sup_{j \geq 0} E_t \Phi_1([\{A_r(X; g)\}^{-1}]^{1/r}) < \infty\) and \(\sup_{j \geq 0} E_t \Phi_2([\{A_r(X; g)\}^{+}]^{1/r}) < \infty\), where \(A_r^\pm(x; g) = \max\{\pm A_r(x; g), 0\}\). Let \(\varepsilon > 0\). Then,

\[
\sum_{n=1}^{\infty} n^{\nu+2} P_t \{ f_n^* > cn^{\nu} \} < \infty, \quad \forall c > 0, \quad (5.1)
\]

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

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

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

with \( \pi \) being the stationary distribution, \( p_1 = 0 \) for \( p - 1/\alpha < r \) and \( p_1 = p - 1/\alpha + I_{\{p-1/\alpha=r\}} \) for \( p - 1/\alpha \geq r \), or

\[
\Phi(t) = \Phi_1(t) = t^p \{\log(1 + t)\}^{p+1+\varepsilon}, \quad \Phi_2(t) = t^{\max(p-1/\alpha,r)} \{\log(1 + t)\}^{p_2}
\]

(5.4)

with \( p_2 = 0 \) for \( p - 1/\alpha < r \) and \( p_2 = p - 1/\alpha + 1 + \varepsilon \) for \( p - 1/\alpha \geq r \). If in addition

\[\sup_{j > 0} E, \Phi(|\Lambda(X_j; g)|) < \infty, \]

then

\[
\sum_{n=1}^{\infty} n^{2p-2} P_r\{S_n^* > cn^2\} < \infty, \quad \forall c > 0,
\]

(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 \( \Lambda(x; g/c) = \Lambda(x; g)/c \) for \( c > 0 \) and we may divide both \( f_n^* \) and \( cn^2 \) by \( c \) in (5.1). Letting \( t = 0, s = e = n^7/(2k_n) \) and \( v = v_n \) in (4.2), we find

\[
P_r\{f_n^* > n^2\} \leq P_r\{d_n^* > n^7/(2k_n)\} + P_r\{V_{r,n} > v_n\} + \left[B_n^* n^{s\sigma_n^* + v_n^{1/2}}/n^{7/(2k_n)^{1/2}}\right]^{k_n}.
\]

(5.6)

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

\[
\sum_{n=1}^{\infty} n^{2p-2} \left[B_n^* n^{s\sigma_n^* + v_n^{1/2}}/n^{7/(2k_n)^{1/2}}\right]^{k_n} < \infty, \quad k_n = M \log n, \quad v_n = n^7/(3k_n)
\]

(5.7)

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

\[
\sum_{n=1}^{\infty} n^{2p-2} \left[B_n^* n^{s\sigma_n^* + v_n^{1/2}}/n^{7/(2k_n)^{1/2}}\right]^{k_n} < \infty, \quad k_n = k, \quad v_n = n^\beta
\]

(5.8)

for some \( 0 < \beta < \alpha, \beta \) depending on \( (\alpha, p, \varepsilon) \) with sufficiently small \( \alpha - \beta \), and a fixed sufficiently large \( k \) depending on \( (\alpha, \beta, p) \). 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

\[
\sum_{n=1}^{\infty} n^{2p-2} P_r\{d_n^* > n^7/(2k_n)\}
\]

\[
\leq \sum_{n=1}^{\infty} n^{2p-2} \sum_{j=1}^{n} P_v\{(2M)|d_j| > n^7/\log n\}
\]

\[
\leq \sum_{j=1}^{\infty} E_v \sum_{n=j}^{\infty} n^{2p-2} I\{n < h_j^{1/2}\}, \quad h_j = M'|d_j|\log(1 + |d_j|)
\]

\[
\leq M' \sum_{j=1}^{\infty} E_v h_j^{(2p-1)/2} I\{h_j^{1/2} > j\}
\]

\[
\leq M'' \sum_{j=1}^{\infty} E_v h_j^p \{\log(1 + h_j^{1/2})\}^{1+\varepsilon}/j^{\log(1 + j)^{1+\varepsilon}}.
\]

(5.9)
If (5.4) holds, then the right-hand side above is finite, as \( h_j \{ \log(1 + h_j^{1/2}) \}^{1+\varepsilon} \ll 1 + M''\Phi(\|d_j\|) \) 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

\[
M'E_h h_{1^{(p-1)/2}} \sum_{j=1}^{\infty} I\{ h^{1/2}_{j} \geq j \} \leq M'' E_h h_{1} \leq (M'')^2 \{ 1 + E\Phi(\|d_1\|) \} < \infty.
\]

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

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

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

\[
\sum_{n=1}^{\infty} n^{2p-2} P_{\pi} \{ d_n^* > n^2/(2k_n) \} < \infty. \quad (5.10)
\]

Let \( \tilde{V}_n = \tilde{V}_{r,n} = \max_{1 \leq j \leq n} E \{ A_j^+(X_n; g)|X_1 \} \). By (2.8) and (2.9) \( V_{r,n}^* \) is no greater than \( \max_{0 \leq j \leq n} (A_j^-(X_j; g))^1/r + \tilde{V}_n^{1/r} \), so that by (5.6)–(5.8) and (5.10) it suffices to show

\[
\sum_{n=1}^{\infty} n^{2p-2} P_{\pi} \left\{ \max_{0 \leq j \leq n} (A_j^-(X_j; g))^1/r > v_n/2 \right\} < \infty \quad (5.11)
\]

and

\[
\sum_{n=1}^{\infty} n^{2p-2} P_{\pi} \{ V_{r,n}^{1/r} > v_n/2 \} < \infty. \quad (5.12)
\]

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

\[
\sum_{n=1}^{\infty} n^{2p-2} P_{\pi} \{ \max(C, \Phi_2(\tilde{V}_n^{1/r}) > \Phi_2(n^2/(6B_n k_n)) \}
\]

\[
\leq \left\{ \sum_{n=1}^{\infty} n^{2p-2} \sup_{n \geq 1} E, \max \left[ C, \Phi_2((A_n^+(X_n; g))^1/r) \right] \right\} < \infty
\]

by the Doob inequality, since \( \tilde{V}_n \) is the maximum of the martingale \( \{ E \{ A_j^+(X_n; g)|X_1 \} \} \) and \( \max(C, \Phi_2(x^{1/r})) \) is convex in \( x \) on \([0, \infty)\) for large \( C\).

Under the stationary measure \( \nu = \pi \), \( \tilde{V}_n \) and \( \max_{2 \leq j \leq n+1} E \{ A_j^+(X_{n+1}; g)|X_n \} \) are identically distributed and they are stochastically bounded by \( \tilde{V}_{n+1} \). Thus, if (5.3) holds and \((k_n, v_n)\) are as in (5.7), then the series in (5.12) is bounded by

\[
\lim_{m \to \infty} \sum_{n=1}^{m} n^{2p-2} P_{\pi} \{ M^{1/r}_{m} \log \tilde{V}_m > n^2 \}
\]

\[
\leq M'' \lim_{m \to \infty} E_{\pi} \{ \tilde{V}_m^{1/r} \log(1 + \tilde{V}_m) \}^{p-1/2} \leq (M'')^2 E_{\pi} \Phi_2((A_0^+(X_0; g))^{1/r}) < \infty
\]

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. \( \square \)
Proof of Theorem 2. Since
\[
\sum_{n=1}^{\infty} n^{\beta-1} P \left\{ \sup_{j \geq n} |\theta_j - \mu| \geq \varepsilon \right\} \leq 2^{\beta-1} \sum_{k=0}^{\infty} 2^{\beta k} \sum_{\ell=k}^{\infty} P \left\{ \max_{2^{\ell} \leq j < 2^{\ell+1}} |\theta_j - \mu| \geq \varepsilon \right\}
\]
\[
= M' \sum_{\ell=0}^{\infty} 2^{\beta \ell} P \left\{ \max_{2^{\ell} \leq j < 2^{\ell+1}} |\theta_j - \mu| \geq \varepsilon \right\} \leq (M')^2 \sum_{n=1}^{\infty} n^{\beta-1} P \left\{ \sup_{j \leq n} |\theta_j - \mu| \geq \varepsilon \right\}
\]
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 $\{S_n(g), n \geq 0\}$ be as in (2.4) with the $\{(X_n, \zeta_n), n \geq 0\}$ in (1.1) and the stationary probability distribution $\pi$. Assume that $E_g(X_0, X_1, \zeta_1) = \mu_g > 0$. Under conditions I.1–I.4 of Kesten (1974), or Harris recurrent condition in Alsmeyer (1994), the elementary renewal theorem asserts that
\[
\sum_{n=1}^{\infty} P_n\{S_n(g) \leq a, X_n \in A\} \sim \pi(A) \mu_{g}^{-1} a \quad \text{as} \ a \to \infty.
\] (6.1)

Under certain conditions, Kesten (1974) also proved a renewal theorem which provides the limit of $\sum_{n=1}^{\infty} P_n\{S_n(g) \leq a, X_n \in A\} - \pi(A) \mu_{g}^{-1} a$ 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 $\varepsilon_n$ be a sequence of random variables and $\{a_\lambda(\cdot), \lambda \in A\}$ 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 $1 < \gamma \leq 1$,
\[
\sum_{n=0}^{\infty} n^{-\gamma} P_n\{\varepsilon_n > c_\varepsilon n\} < \infty, \quad \sum_{n=0}^{\infty} P_n\{-\varepsilon_n > c_\varepsilon n\} < \infty,
\] (6.2)
and for all $\lambda \in A$ with $b_\lambda \geq b_*$
\[
\max_{b_\lambda \leq n \leq b_*-b_\lambda} \frac{\mu n - a_\lambda(n)}{n^\gamma} < -1/M_{\varepsilon}, \quad (6.3)
\]
\[
\max_{b_\lambda + (1+c)b_\lambda \leq n \leq b_*} \frac{a_\lambda(n) - \mu n}{n^\gamma} < -1/M_{\varepsilon}, \quad \sup_{n \geq (1+c)b_\lambda} \frac{a_\lambda(n) - \mu n}{n} < -1/M_{\varepsilon}. \quad (6.4)
\]
If $a_\lambda$ 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
\[
U_\lambda = U_{\lambda, g, g_0} = \sum_{n=1}^{\infty} g(X_{n-1}, X_n, \zeta_n) I_{\{S_n(g) + \varepsilon_n \leq a_\lambda(n)\}}
\]
and define $T_\lambda = T_{\lambda, g}$ and $N_\lambda = N_{\lambda, g}$ by
\[
T_\lambda = \inf \{n \geq 1: S_n(g) + \varepsilon_n > a_\lambda(n)\}, \quad N_\lambda = 1 + \sup\{n \geq 1: S_n(g) + \varepsilon_n \leq a_\lambda(n)\}.
\]
Theorem 7. Let \( \{(g, g_0) \in H^* \} \) with \( \|g_0\|_\infty < \infty, \mu = E_0 g(X_0, X_1, \xi_1) \) and \( \mu_0 = E_0 g_0(X_0, X_1, \xi_1) \). Suppose the (6.2)–(6.4) hold with \( \frac{1}{2} < \gamma \leq 1 \). Suppose the conditions of Theorem 6 hold with \( g \) and \( (p, x) = (2/\gamma - 1, \gamma) \) for \( \gamma < 1 \) and also with \( g \) and \( (p, x) = (2, 1) \). Then, for large \( b_j \)

\[
E_v |U_{J, g, g_0} - S_n(g_0)| = O(b_j^q),
\]

where \( n_j = n_{j,\mu} \) is the integer part of \( b_j \) in (6.3) and (6.4). If \( E_v A(X_n; g_0) = O(n^q) \), then

\[
E_v U_{J, g, g_0} = \mu_0 b_j + O(b_j^q).
\]

If (6.3) and (6.4) hold for all \( \epsilon > 0 \) and \( E_v A(X_n; g_0) = o(n^q) \), then the \( O(b_j^q) \) in (6.5) and (6.6) can be replaced by \( o(b_j^q) \). Furthermore, the above assertions hold under respective conditions when \( U_{J, g, g_0} \) is replaced by \( S_T(g_0) \) or \( S_{\xi_1}(g_0) \).

Remark. Expansion (6.6) implies (6.1) with \( g_0(x, y, \xi) = I_{(y \in A)} \) 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 \( S_T(g_0), \) \( S_{\xi_1}(g_0) \) or \( \gamma = 1 \)] is nearly identical. Splitting the sums \( U_{J, g, g_0} \) and \( S_n(g_0) \) into four parts \([1, n_1], (n_1, n_2], (n_2, n_3] \) and \((n_3, \infty) \) with \( n_1, n_2 \) and \( n_3 \) being the integer parts of \( b_j^1, b_j^2 \) and \( b_j^3 + \epsilon b_j^2 \), respectively, we find

\[
|U_{J, g, g_0} - S_n(g_0)| \leq \|g_0\|_\infty \left( n_1 + \sum_{n_1 < n \leq n_2} I\{S_n(g) + \xi_n > a_j(n)\} + (n_3 - n_2) + \sum_{n > n_3} I\{S_n(g) + \xi_n \leq a_j(n)\} \right).
\]

Since \( \mu n - a_j(n) \leq - n^{\gamma}/M_\epsilon \) for \( \epsilon b_j^1 \leq n \leq b_j^2 - \epsilon b_j^2 \), by (5.5) for \( (p, x) = (2/\gamma - 1, \gamma) \) (thus, \( px - 2 = -\gamma \)) and by (6.2) we have

\[
\sum_{n_1 < n \leq n_2} P_r\{S_n(g) + \xi_n > a_j(n)\} \leq b_j^2 \sum_{n_1 < n \leq n_2} n^{-\gamma} P_r\{S_n(g) - \mu n + \xi_n > n^{\gamma}/M_\epsilon\} = o(b_j^q).
\]

Similarly, since \( a_j(n) - n \mu \leq - n^{\gamma}/M_\epsilon \) for \( b_j^2 - \epsilon b_j^2 < n \leq (1 + \epsilon)b_j, \)

\[
\sum_{n_1 < n \leq n_4} P_r\{S_n(g) + \xi_n \leq a_j(n)\} \leq (1 + \epsilon)b_j^2 \sum_{n_1 < n \leq n_4} n^{-\gamma} P_r\{S_n(g) - \mu n + \xi_n \leq - n^{\gamma}/M_\epsilon\} = o(b_j^q),
\]

where \( n_4 \) is the integer part of \( (1 + \epsilon)b_j \). Finally, since \( a_j(n) - n \mu \leq - n^{\gamma}/M_\epsilon \) for \( n > (1 + \epsilon)b_j, \) by Theorem 6 with \( (p, x) = (2, 1) \) and (6.2) we have

\[
\sum_{n > n_4} P_r\{S_n(g) + \xi_n \leq a_j(n)\} \leq \sum_{n > n_4} P_r\{S_n(g) - \mu n + \xi_n \leq - n^{\gamma}/M_\epsilon\} = o(1).
\]

Putting these inequalities together, we obtain (6.5), as \( n_1 \leq \epsilon b_j^2 \) and \( n_3 - n_2 \leq 1 + 2 \epsilon b_j^2 \). By (2.6) of Theorem 1 and the condition on \( E_v A(X_n; g_0) \), (6.6) follows from

\[
E_v S_n(g_0) = n_j \mu_0 + E_v A(X_n; g_0) - E_v A(X_0; g_0) = \mu_0 b_j + O(b_j^q).
\]
7. Random coefficient autoregression models

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

\[
x_k = \beta_k x_{k-1} + e_k, \quad x_0 = 0,
\]

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

Under the normality assumption on \((\beta_k, e_k)\) with known \((\sigma^2, \kappa^2)\), the log-likelihood ratio statistics \( Z_n \) for testing \( \beta = 0 \) are given by

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

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

\[
T_\lambda = \inf \{n \geq 1: Z_n > \lambda\}.
\]

We shall investigate \( T_\lambda \) under the stability assumption \( \beta^2 + \kappa^2 < 1 \), without the normality assumption on \((\beta_k, e_k)\). Since \( \hat{\beta}_n = C_n/D_n \) are the least squares estimates of \( \beta \), \( Z_n \) and \( T_\lambda \) 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, \( D_n \to E_x D_1 = \mu_D \) and \( C_n/n \to E_x C_1 = \mu_C = \beta \mu_D \). Taking Taylor expansion as in Melli (1992), we find

\[
Z_n = n h(C_n/n, D_n/n) = \sum_{i=1}^{n} g(x_{i-1}, x_i) + \zeta_n,
\]

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

\[
\zeta_n = n(C_n/n - \mu_C, D_n/n - \mu_D) h^{(2)}(C_n/n - \mu_C, D_n/n - \mu_D) v^{tr}/2,
\]

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 \((C_1, D_1)\), 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 \( E_x A(x, g) \leq C \sigma^2 x_1^2 < \infty \), for some constant \( C \). By Theorem 7, we have

\[
E_x(T_\lambda) = \lambda/\mu_0 + o(\lambda) \quad \text{as} \quad \lambda \to \infty, \quad \mu_0 = E_x g(x_0, x_1) = \beta^2 \mu_D / 2 > 0.
\]

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