Journal of Theoretical Probability, Vol. 6, No. 3, 1993

On the Convergence of Averages of Mixing Sequences

W. Bryc[1] **and W. Smolenski**[2]

Received March 9, 1992; revised July 21, 1992

We construct an absolutely regular stationary random sequence which is an instantaneous bounded function of an aperiodic recurrent Markov chain with a countable state space, such that the large deviation principle fails for the arithmetic means of the sequence, while the exponential convergence holds. We also show that exponential convergence holds for the arithmetic means of a vector valued strictly stationary bounded ϕ-mixing sequence.

KEY WORDS: Exponential convergence; large deviations; strong mixing.

1. INTRODUCTION

The main result of this note is the example of a recurrent aperiodic stationary Markov chain X_k on a countable state space such that its empirical distributions do not satisfy the large deviation principle. Moreover, we show that the large deviation principle fails for the arithmetic means of a certain 3-valued instantaneous function $f(X_k)$ of this chain. The additional property of the example is that the convergence of arithmetic means of $f(X_k)$ is exponentially fast, clarifying the relation between exponential convergence and large deviations (c.f. Remark 5.1) and complementing the result of Schonmann[13]; in Section 5 we also give a short proof of the vector version of Schonmann's result. An important condition, which is stronger than recurrence, is the Doeblin condition D, see Doob,[7] page 197. As far as we know, the question, whether the arithmetic means of a stationary sequence obtained as an instantaneous bounded measurable

[1] Department of Mathematics, University of Cincinnati, Cincinnati, Ohio 45221-0025. *e-mail:* bryc@ucbeh.san.uc.edu.
[2] Department of Mathematics, The University of Kansas, Lawrence, Kansas 66045-2142. *e-mail:* smolensk@kuhub.cc.ukans.edu.

0894-9840/93/0700-0473$07.00/0

function $f(X_k)$ of a Markov chain which satisfies the Doeblin condition (even in an over-simplified form of Eq. (2.3)) is open. Such sequences necessarily have exponential ϕ-mixing rate (for definition, see Eq. (3.2)) and the speed of convergence in the law of large numbers is exponential. We also point out that if a Markov chain X_k satisfies condition in Eq. (2.1), which is stronger than the Doeblin condition D, and if $f(\cdot)$ is a bounded measurable function on the state space of the chain, then the arithmetic means of $f(X_k)$ satisfy the large deviation principle.

Let $(\mathbb{E}, \|\cdot\|)$ be a separable Banach space and let S_n, $n \geqslant 1$, be an $\mathbb{E}$-valued random sequence. Recall that the sequence $(1/n)\, S_n$, $n \geqslant 1$, *satisfies the large deviation principle*, if there is a convex lower semicontinuous rate function $I: \mathbb{E} \to [0, \infty]$ with compact level sets $I^{-1}[0, a]$, $a \geqslant 0$, such that

$$\limsup_{n \to \infty} n^{-1} \log P\left(\frac{1}{n} S_n \in A\right) \leqslant -\inf_{x \in A} I(x) \tag{1.1}$$

for each closed set $A \subset \mathbb{E}$;

$$\liminf_{n \to \infty} n^{-1} \log P\left(\frac{1}{n} S_n \in A\right) \geqslant -\inf_{x \in A} I(x) \tag{1.2}$$

for each open set $A \subset \mathbb{E}$.

We shall say that the sequence $(1/n)\, S_n$, $n \geqslant 1$, converges exponentially to $m \in \mathbb{E}$, if for every $\varepsilon > 0$ there are $\gamma > 0$ and $C < \infty$ such that

$$P\left(\left\|\frac{1}{n} S_n - m\right\| > \varepsilon\right) \leqslant C e^{-\gamma n} \qquad \text{for all} \quad n \geqslant 1 \tag{1.3}$$

Clearly, if $\{(1/n)\, S_n\}_{n \geqslant 1}$ satisfies the large deviation principle with a strictly convex at m (hence strictly positive outside of any ball around m) rate function $I(\cdot)$ then the exponential convergence follows.

2. MIXING CONDITIONS FOR MARKOV CHAINS

Let X_k be a time homogeneous Markov chain on a state space $(\Sigma, \mathscr{S})$, with the n-step transition function $\Pi_n(x, A)$ and an invariant measure π (i.e., $\int \Pi_1(x, A)\, \pi(dx) = \pi(A)$ for all $A \in \mathscr{S}$). Given a measurable function $f: \Sigma \to \mathbb{E}$, put $\xi_k = f(X_k)$ and let $S_n = \sum_{k=1}^{n} \xi_k$; in this context S_n is sometimes referred to as a "Markov additive functional." Consider the following condition.

There are $\delta > 0$, $N \geqslant 1$ such that

$$\Pi_N(x, A) \geqslant \delta \pi(A) \qquad \text{for all} \quad x \in \Sigma, \qquad A \in \mathscr{S} \tag{2.1}$$

Proposition 2.1. If Eq. (2.1) holds and $f\colon \Sigma \to \mathbb{R}^d$ is bounded, then the sequence $(1/n)\, S_n$, $n \geqslant 1$, satisfies the large deviation principle with a rate function $I\colon \mathbb{R}^d \to [0, \infty)$.

Indeed, one can easily see that Eq. (2.1) implies $\psi_-(N) = \delta > 0$, c.f. Eq. (3.1). Hence the stationary sequence ξ_k satisfies [Ref. 4, conditions (1.10) and (1.11)] and the result follows from [Ref. 4, Theorem 2].

Remark 2.1. Proposition 2.1 holds true under the following slightly less restrictive conditions (which allows to handle periodic chains):

$$\exists \delta > 0, \quad N \geqslant 1 \sum_{k=1}^{N} \Pi_k(x, A) \geqslant \delta \pi(A)\; \forall A \in \mathscr{S}, \qquad x \in \Sigma \tag{2.2}$$

This can be obtained from Ref. 5, Theorem 2.1 by the contraction principle, c.f. also Ref. 5, Corollary 2.1. Also, it can be shown that the rate function in Proposition 2.1 is convex.

The following condition is related to Doeblin's condition D. The main difference is that we use the invariant measure π rather than an arbitrary sigma-finite nonnegative measure μ (part of the theory is then to establish existence and uniqueness of the invariant measure π).

There are $\delta > 0$, $N \geqslant 1$ such that for $A \in \mathscr{S}$

$$\pi(A) \leqslant \delta \Rightarrow \Pi_N(x, A) \leqslant 1 - \delta \qquad \text{for all} \quad x \in \Sigma \tag{2.3}$$

We believe that the question whether assumption of Eq. (2.1) in Proposition 2.1 can be replaced by Eq. (2.3) is open.

Harris[8] introduced the following recurrence condition, which is weaker than Doeblin's D condition.

There is a sigma-finite nonnegative measure μ on $(\Sigma, \mathscr{S})$ such that for $A \in \mathscr{S}$

$$\mu(A) > 0 \Rightarrow P\left(\bigcup_{k=1}^{\infty} \{X_k \in A\} \,\middle|\, X_0 = x \right) = 1 \qquad \text{for all} \quad x \in \Sigma \tag{2.4}$$

In Section 4 we shall construct an aperiodic Markov chain with a countable state space Σ, satisfying condition in Eq. (2.4) with the counting measure μ and such that the large deviation principle fails for an additive functional of this chain. It might be of interest to point out that in Eq. (2.4) implies the following minorization condition, c.f. Ney and Nummelin,[10] Eq. (2.1).

There are $N \geqslant 1$, a sigma finite measure μ on $(\Sigma, \mathscr{S})$ and a measurable function $h\colon \Sigma \to [0, \infty)$ such that $\int h(x)\, \mu(dx) > 0$ and

$$\Pi_N(x, A) \geqslant h(x)\, \mu(A) \qquad \text{for all} \quad x \in \Sigma, \qquad A \in \mathscr{S} \tag{2.5}$$

We write this last minorization condition explicitly, to point out that besides Eq. (2.5) some additional assumptions on stationary Markov chains are needed in order to prove the large deviation principle for the arithmetic means of Markov additive processes, c.f. Ney and Nummelin,[10] and de Acosta.[6]

3. MIXING CONDITIONS FOR STATIONARY SEQUENCES

Assumptions made in this note correspond to various "strong mixing conditions" considered for stationary sequences.

Let $(\mathbb{E}, \|\cdot\|)$ be a separable Banach space and let $\{\xi_k\}_{k\in\mathbb{Z}}$ be a strictly stationary bounded $\mathbb{E}$-valued random sequence. In the sequel we denote $S_n=\sum_{k=1}^n \xi_k$, $\mu=E\{\xi_1\}$.

For $n\geqslant 0$ let $\mathscr{F}_{-\infty,n}$ be the σ-field generated by $\{\xi_k: k\leqslant n\}$ and let $\mathscr{F}_{n,\infty}$ be the σ-field generated by $\{\xi_k: k\geqslant n\}$.

The following condition, related to Eq. (2.1), was employed in Ref. 4 (see also references there); suitable generalizations to random fields are given in Ref. 5.

For $n\geqslant 1$, let

$$\psi_-(n)=\inf\{P(A\cap B)/(P(A)\,P(B)):\\ A\in\mathscr{F}_{-\infty,-n+1}, B\in\mathscr{F}_{1,\infty}, P(A)\,P(B)>0\} \tag{3.1}$$

Condition $\sup_n \psi_-(n)>0$ is the same as Ref. 4, Eq. (1.10); for Markov chains $\psi_-(N)>0$ and Eq. (2.1) are equivalent.

For $n\geqslant 1$, the ϕ-mixing coefficients are defined by

$$\phi(n)=\sup\{|P(A\mid B)-P(A)|: B\in\mathscr{F}_{-\infty,0}, A\in\mathscr{F}_{n,\infty}, P(B)>0\} \tag{3.2}$$

We say that $\{\xi_k\}$ is ϕ-mixing, if $\phi(n)\to 0$ as $n\to\infty$.

For an aperiodic Markov Chain Doeblin's condition D implies ϕ-mixing with geometric rate (c.f., Rosenblatt,[12] p. 209); large deviation principle for bounded stationary ϕ-mixing random sequences was established in Ref. 4, Theorem 1, but the result is not applicable to the arithmetic means of an additive functional of a Markov chain satisfying Doeblin's condition, since Ref. 4, Theorem 1 assumes hypergeometric ϕ-mixing rate.

It is known (Bradley[1]) that if $\sup_n \psi_-(n)>0$ then $\psi_-(n)\to 1$ as $n\to\infty$ and $\phi(n)\to 0$ as $n\to\infty$; this correspond to a trivial in the Markov case, implication Eqs. (2.1) $\Rightarrow$ (2.3).

If X_k is an aperiodic Markov chain satisfying Eq. (2.4) (and is taken with invariant initial distribution) and $f: \Sigma\to\mathbb{E}$ is measurable, then

$\xi_k = f(X_k)$ satisfies the following weaker strong mixing condition, c.f. Bradley[3] and the references therein.

A strictly stationary sequence (ξ_n) is called absolutely regular, if $\beta(n) \to 0$ as $n \to \infty$, where

$$\beta(n) := \tfrac{1}{2} \sup \sum_{i,j} |P(A_i \cap B_j) - P(A_i)\, P(B_j)| \tag{3.3}$$

the supremum in Eq. (3.3) is taken over all finite $\mathscr{F}_{-\infty,0}$-measurable (disjoint) partitions $\{A_i\}$ and over all finite $\mathscr{F}_{n,\infty}$-measurable partitions $\{B_j\}$ of probability space Ω.

4. EXAMPLE (Adapted form Bradley,[2] and Orey and Pelikan,[11] Example 4.1)

We first introduce a number of parameters. For $k \geqslant 1$, let $n(k) = 12^k$ and let $n_0 = 1/2$. For $k \geqslant 0$, let $p_k = C \exp(-n(k)/2)$, where $C = (\sum_{k=0}^{\infty} \exp(-n(k)/2))^{-1}$ is the normalizing constant.

A Markov chain to be constructed below has pairs (k, j) as states. For $k \geqslant 1$, the state space Σ is such that j runs through all integers in $(-n(k), n(k)]$; clearly, there are $2n(k)$ states corresponding to each $k \geqslant 1$. We also choose a single value $j = 1/2$ for $k = 0$. For $k \geqslant 0$, let the one step transition probabilities be defined by

$$P(X(n+1) = (k, j-1) \mid X(n) = (k, j)) = 1 \qquad \text{if} \quad j \geqslant 2 - n(k)$$

$$P(X(n+1) = (m, n(m)) \mid X(n) = (k, 1 - n(k))) = p_m, \qquad m \geqslant 0$$

Notice that $X(n)$ is an aperiodic recurrent Markov chain. Indeed, the chain evolves independently after visiting the "renewal" set $R = \{(k, -n(k)+1) : k \geqslant 1\}$ and the hitting time for R is finite; aperiodicity is ensured by the fact that $(0, 1/2)$ has a nonzero probability of returning in one unit of time.

By a simple calculation, $X(n)$ is strictly stationary when started from the initial distribution

$$P(X(0) = (k, j)) = Cp_k \tag{4.1}$$

where $C = (\sum_{k=0}^{\infty} 2n(k)\, p_k)^{-1}$ is a normalizing constant. In Eq. (4.1) the range of j is $-n(k) < j \leqslant n(k)$ when $k \geqslant 1$ and $j = 1/2$ when $k = 0$.

A stationary sequence that we are looking for has a form $\xi_n = f(X(n))$. To define a suitable function $f: \Sigma \to \{-1, 0, 1\}$, consider auxiliary functions $k(\cdot)$ and $j(\cdot)$, defined for $\mathbf{x} = (m, r) \in \Sigma$ by $k(\mathbf{x}) = m$, $j(\mathbf{x}) = r$; $j(\cdot)$ is needed in the definition of Eq. (4.2), $k(\cdot)$ will be needed for the proof.

Let

$$\xi_n = \text{sign}(j(X(n) - 1/2)) \tag{4.2}$$

with the convention that $\text{sign}(0) = 0$; the last convention applies only to a single state $(0, 1/2)$.

Proposition 4.1. Let (ξ_n) be defined by Eq. (4.2). Then

(i) (ξ_n) is absolutely regular strictly stationary and bounded;

(ii) the large deviation principle fails for $\{(1/n)\, S_n\}$.

Remark 4.1. In particular, X_k is an irreducible recurrent Markov chain such that its empirical distributions do not satisfy the large deviation principle in the weak topology on the set of probability measures. Indeed, since f is bounded and continuous, $(1/n)\, S_n$ obtains as a continuous function of the empirical measure $\mu_n = 1/n \sum_{k=1}^{n} \delta_{X_k}$ and the contraction principle, see e.g. Varadhan(14) [page 5] asserts that the large deviation principle is preserved by the continuous mapping.

Proof of (i). It is known (see Bradley(3)) that every real aperiodic Harris chain $\{X_t\}$ is absolutely regular. In particular, for a measurable function $f\colon \Sigma \to \mathbb{R}$, $\xi_n = f(X(n))$ is strictly stationary and absolutely regular.

Proof of (ii). To prove that the large deviation principle fails, it is enough to show that the sequence $(1/n) \log E\{\exp S_n\}$, $n \geqslant 1$, does not have limit as $n \to \infty$, see Varadhan,(14) Theorem 2.2.

Obviously, for $k \geqslant 1$, $E\{\exp S_{n(k)}\} \geqslant e^{n(k)} P(X(0) = (k, n(k))) = Ce^{n(k)/2}$. This shows that

$$\limsup_{n \to \infty} \frac{1}{n} \log E\{\exp S_n\} \geqslant 1/2 \tag{4.3}$$

We shall show that $\liminf_{n \to \infty} (1/n) \log E\{\exp S_n\} \leqslant 1/4$. To this end, fix $k \geqslant 1$ and let $n = 8n(k)$.

The following two cases are possible:

(A) there is $0 \leqslant t \leqslant n$ such that $k(X(t)) \geqslant k+1$;

(B) $k(X(t)) \leqslant k$ for all t.

Denoting the corresponding events by A, B we have

$$E\{\exp S_n\} = \int_A e^{S_n}\, dP + \int_B e^{S_n}\, dP \tag{4.4}$$

We shall analyze separately each term on the right-hand side of Eq. (4.4).

If the event $A = \bigcup_{t=0}^{n} A(t)$ occurs, then using trivial bounds $S_n \leqslant n$, $P(A) \leqslant \Sigma P(A(t))$ and Eq. (4.1) we get

$$\int_A e^{S_n}\, dP \leqslant (n+1)\, Ce^n \sum_{j=k+1}^{\infty} n(j)\, p_j$$
$$\leqslant C_1 n(k)\, e^{8n(k) - n(k+1)/2} \leqslant C_1 n(k)\, e^{2n(k)} \tag{4.5}$$

If event B occurs, then the longest block of consecutive 1's has length $n(k)$ and, due to the cancellations, we have $S_n \leqslant n(k)$. Therefore

$$\int_B e^{S_n}\, dP \leqslant e^{n(k)} \tag{4.6}$$

Inequalities of Eqs. (4.4)–(4.6) together imply

$$\liminf_{n \to \infty} \frac{1}{n} \log E\{\exp S_n\} \leqslant \liminf_{k \to \infty} 1/(8n(k)) \log E\{\exp S_{8n(k)}\}$$
$$\leqslant \lim_{k \to \infty} [C + 2n(k) + \log n(k)]/(8n(k)) = 1/4$$

By Eq. (4.3) this ends the proof.

Remark 4.2. To see directly that X_k satisfies Eq. (2.5), let μ be the invariant measure Eq. (4.1) and take $h(x) = 0$ for $x \notin R$, $h(x) = 1/C$ for $x \in R$.

5. EXPONENTIAL CONVERGENCE

Proposition 5.1. If $\{\xi_n\}$ is the sequence defined by Eq. (4.2) then exponential convergence Eq. (1.3) holds for $(1/n)\, S_n$, $n \geqslant 1$ (with $m = 0$).

Proof. Clearly $E\{S_n\} = 0$. Given $n \geqslant 1$ and $\varepsilon > 0$, let k be such that $n(k) \leqslant n\varepsilon < n(k+1)$, i.e., $k = 1 + [\log_{12} n\varepsilon]$. If $k(X_t) \leqslant k$ for all $t \leqslant n$, then $|S_n| \leqslant n(k) \leqslant n\varepsilon$. Therefore the event $\{|(1/n)\, S_n| > \varepsilon\}$ implies $k(X_t) \geqslant k+1$ for some $t \leqslant n$. Therefore

$$P\left(\left|\frac{1}{n} S_n\right| > \varepsilon\right) \leqslant nP(k(X_0) \geqslant k+1) = Cn \sum_{j=k+1}^{\infty} n(j)\, e^{-n(j)/2}$$
$$\leqslant C_1 n \exp(-12^k/4) \leqslant C_2(\varepsilon)\, e^{-\gamma n}$$

The following result extends Schonmann[13] to vector valued r.v. Our proof uses Chebyshev's inequality rather than combinatorial argument, but we believe Schonmann's proof can be adapted to handle the vector case, too.

Theorem 5.1. If $\{\xi_n\}$ is ϕ-mixing and bounded, then exponential convergence Eq. (1.3) holds for $(1/n)\, S_n$, $n \geqslant 1$.

Corollary 5.1. If ξ_k is ϕ-mixing real valued and bounded then for any $\varepsilon > 0$ there are $\gamma > 0$ and $C < \infty$ such that $P(S_n > n(m+\varepsilon)) \leqslant Ce^{-\gamma n}$.

Remark 5.1. In the construction of Section 4, if one chooses $n(k) = k$, $p(k) = C/n(k)^4$, then the reasoning that leads to Eq. (4.3) shows that exponential convergence fails and the large deviation principle holds with the (trivial) rate function $I(\cdot) = 0$. Thus, aside the trivial implication when the rate function is strictly convex, exponential convergence and the large deviation principle can occur independently. Also, the modified construction shows that Theorem 5.1 cannot be extended to stationary absolutely regular sequences.

The following lemma is well known and is proved for completeness only.

Lemma 5.1. If $\mathbb{E}$ is a separable Banach space and $\{\xi_k\}$ is strictly stationary regular (i.e., $\bigcap_{n \leqslant 0} \mathscr{F}_{-\infty,n}$ is a trivial σ-field) and $E\{\|\xi_1\|\} < \infty$, then $S_n/n \to m = E\{\xi_1\}$ in probability.

Proof. By separability, for $\varepsilon > 0$ there is $f_\varepsilon\colon \mathbb{E} \to \mathbb{E}$ such that $f_\varepsilon(\xi_1)$ has a finite range and

$$E\{\|f_\varepsilon(\xi_1) - \xi_1\|\} < \varepsilon \tag{5.1}$$

Let $\theta_k = f_\varepsilon(\xi_k)$. Clearly, $\{\theta_k\}$ is strictly stationary, regular, and takes values in a finite dimensional subspace of $\mathbb{E}$. Therefore, by the ergodic theorem, $(1/n) \sum_{k=1}^n \theta_k \to m_\varepsilon = E\{\theta_1\}$ a.s., and $\|m_\varepsilon - m\| \leqslant \varepsilon$.

Also, $\|\theta_k - \xi_k\|$ is strictly stationary, regular, and hence by ergodic theorem and (ii) $(1/n) \sum_{k=1}^n \|\theta_k - \xi_k\| \to E\{\|\theta_1 - \xi_1\|\} < \varepsilon$. Therefore,

$$P(\|S_n - nm\| > 4\varepsilon n) \leqslant P\left(\left\|\frac{1}{n} \sum_{k=1}^n \theta_k - m_\varepsilon\right\| > \varepsilon\right) + P\left(\frac{1}{n} \sum_{k=1}^n \|\theta_k - \xi_k\| > 2\varepsilon\right) \to 0 \qquad \text{as} \quad n \to \infty$$

Proof of Theorem 5.1. By centering, without loss of generality we may assume $m = 0$.

Given $\varepsilon > 0$ choose $\delta < \frac{1}{2}(e^{\varepsilon/2} - 1)$. Let M be such that $\|\xi\| \leqslant M$. Fix $N \geqslant 1$ to have

$$\phi(N)\, e^M < \delta \tag{5.2}$$

Since by Lemma 5.1 $\|(1/n)\, S_n\| \to 0$ as $n \to \infty$, and random variables $\|(1/n)\, S_n\|$ are uniformly bounded we have $E\{\exp((1/n)\, \|S_n\|)\} \to 1$ as $n \to \infty$. Therefore we can choose L such that

$$E\left\{\exp\left(\frac{1}{L}\, \|S_L\|\right)\right\} < 1 + \delta \tag{5.3}$$

Following Schonmann,[13] we use now blocking argument with parameters L, N chosen as previously. For q, j, $k \geqslant 1$ let

$$Y_j^{(q)} = \frac{1}{L} \sum_{i=1}^{L} \xi_{i + NL(j-1) + (q-1)L}$$

$$Z_k^{(q)} = \frac{1}{k} \sum_{j=1}^{k} Y_j^{(q)}$$

(We shall drop superscript q when $q = 1$.)

The function $(i, q, j) \to i + L(q-1) + NL(j-1)$ is a bijection from $\{1,..., L\} \times \{1,..., N\} \times \{1,..., k\}$ onto $\{1, 2,..., NLk\}$. Therefore for $k \geqslant 1$

$$\frac{1}{NLk}\, S_{NLk} = \frac{1}{N} \sum_{q=1}^{N} Z_k^{(q)}$$

and hence by stationarity

$$P\left(\frac{1}{NLk}\, \|S_{NLk}\| > \varepsilon/2\right) \leqslant NP(\|Z_k\| > \varepsilon/2) \tag{5.4}$$

By Chebyshev's inequality

$$P(\|Z_k\| \geqslant \varepsilon/2) \leqslant e^{-k\varepsilon/2} E\left\{\exp\left(\sum_{j=1}^{k} \|Y_j\|\right)\right\} \tag{5.5}$$

By ϕ-mixing (c.f. Ref. 9), Eqs. (5.3) and (5.2)

$$E\{\exp(\|Y_{j+1}\|) \mid Y_1,..., Y_j\} \leqslant E\{\exp(\|Y_1\|)\} + \phi(N)\, e^M < 1 + 2\delta$$

Therefore from Eqs. (5.4) and (5.5) we get

$$P\left(\frac{1}{NLk}\, \|S_{NLk}\| > \varepsilon/2\right) \leqslant N\left(\frac{1 + 2\delta}{e^{\varepsilon/2}}\right)^k, \qquad k = 1, 2,... \tag{5.6}$$

Now let $n \geqslant 2LNM/\varepsilon$ be an integer and put $k=[n/(LN)]$, so that $NLk \leqslant n$.

Since

$$\left\| \frac{1}{n} S_n - \frac{1}{NLk} S_{NLk} \right\| \leqslant \frac{1}{n} \|S_n - S_{LNk}\| + \left(1 - \frac{NLk}{n}\right) \left\| \frac{1}{NLk} S_{NLk} \right\|$$
$$\leqslant \varepsilon/2 + \left\| \frac{1}{NLk} S_{NLk} \right\|$$

we have

$$P\left(\left\| \frac{1}{n} S_n \right\| > \varepsilon\right) \leqslant N\left(\frac{1+2\delta}{e^{\varepsilon/2}}\right)^k \leqslant N\left(\frac{1+2\delta}{e^{\varepsilon/2}}\right)^{-LN} \left(\frac{1+2\delta}{e^{\varepsilon/2}}\right)^n$$

This ends the proof of the theorem with $\gamma=\gamma(\varepsilon)=\varepsilon/2-\log(1+2\delta)$; clearly $\gamma>0$ by our choice of δ.

ACKNOWLEDGMENTS

We would like to thank R. Bradley for helpful comments.

REFERENCES

1. Bradley, R. C. (1983). On the ψ-mixing condition for stationary random sequences, *Trans. Amer. Math. Soc.* **276**, 55–66.
2. Bradley, R. C. (1989). A stationary, pairwise independent absolutely regular sequence for which the central limit theorem fails, *Probab. Th. Rel. Fields* **81**, 1–10.
3. Bradley, R. C. (1986). Basic properties of strong mixing conditions. In: *Dependence in Probability and Statistics* (eds.), E. Eberlein and M. S. Taqqu, Birkhäuser, Boston, pp. 165–192.
4. Bryc, W. (1992). On large deviations for uniformly strong mixing sequences, *Stoch. Proc. Appl.* **41**, 191–202.
5. Bryc, W. (1992). On the large deviation principle for stationary weakly dependent random fields, *Ann. Prob.* **20**, 1004–1030.
6. Acosta, A. de (1990). Large deviations for empirical measures of Markov chains, *J. Theoretical Prob.* **3**, 395–431.
7. Doob, J. L. (1953). *Stochastic Processes*, Wiley, New York.
8. Harris, T. E. (1956). The existence of stationary measures for certain Markov processes, *Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. II, Berkeley, pp. 113–124.
9. Ibragimov, I. A., and Linnik, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen.
10. Ney, P., and Nummelin, E. (1987). Markov additive process I. Eigenvalue properties and limit theorems, *Ann. Prob.* **15**, 593–609.

11. Orey, S., and Pelikan, S. (1988). Large deviation principles for stationary processes, *Ann. Prob.* **16**, 1481–1495.
12. Rosenblatt, M. (1971). *Markov Processes. Structure and Asymptotic Behavior*, Springer, Berlin.
13. Schonmann, R. H. (1989). Exponential Convergence Under Mixing, *Probab. Theory Related Fields* **81**, 235–238.
14. Varadhan, S. R. S. (1984). *Large Deviations and Applications*, SIAM, Philadelphia.