Remark. The function φ in (20) does not determine the measure ϱ. In fact, consider $H = R$ and $p = 2$. Then

$$\varphi(t) = \exp\{-c|t|^2\},$$

where $c \geqslant 0$, is the characteristic function of a symmetric normal distribution. It is easy to verify that for

$$\varrho_1(E) = c\delta_1(E)$$

and

$$\varrho_2(E) = \frac{c}{2}\delta_{-1}(E) + \frac{c}{2}\delta_1(E)$$

we have the formula

$$\int_S |(y, x)|^2 \varrho_i(dx) = c|y|^2 \quad (i = 1,2).$$

Thus, the measure ϱ is not uniquely determined by φ.

Added in proof

After this work was completed I found a paper of J. Kuelbs in Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 26 (1973), pp. 259–271, where the canonical representation of stable measures was also obtained. However we note that our methods are enterily different.

References

[1] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley, Cambridge 1954.
[2] R. Jajte, *On stable distributions in Hilbert space*, Studia Math. 30 (1968), pp. 63–71.
[3] A. Kumar and V. Mandrekar, *Stable probability measures on Banach spaces*, ibid. 42 (1972), pp. 133–144.
[4] K. Kuratowski, *Topologie II*, Warszawa–Wrocław 1950.
[5] K. R. Parthasarathy, *Probability measures on metric spaces*, New York–London 1967.
[6] R. P. Phelps, *Lectures on Choquet's theorem*, Princeton 1966.
[7] K. Urbanik, *Lévy's probability measures on Euclidean spaces*, Studia Math. 44 (1972), pp. 119–148.

Presented to the Semester
Probability Theory
February 11–June 11, 1976

PROBABILITY THEORY
BANACH CENTER PUBLICATIONS, VOLUME 5
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1979

ON THE ADAPTIVE CONTROL OF COUNTABLE MARKOV CHAINS

PETR MANDL
Charles University, Prague, Czechoslovakia

1. Introduction

The paper treats the same subject as paper [4], namely the dependence of the asymptotic behaviour of the criterion functional (the reward) on the asymptotic behaviour of the control. But it concerns chains with a countable state space. In this case additional hypotheses on the transition probabilities of the system are required to ensure the stability of its basic parameters. In stating these hypotheses, we begin with a Liapounov type condition for the existence of an optimal stationary policy introduced in monograph [3]. The methods of this monograph are used in Section 2. In the remaining sections, where general non-anticipative controls are investigated, the validity of several Liapounov type conditions is assumed.

We consider a system S, which is observed at times $0, 1, 2, \ldots$, and has countably many states labelled by numbers $1, 2, 3, \ldots$ We write $I = \{1, 2, 3, \ldots\}$. Let X_n be the state of S at time n. We assume the following law of motion: For arbitrary $i \in I$, whenever S is in state i, the probability distribution of the next state is

$$\{p(i, 1; z), p(i, 2; z), \ldots\}, \quad z \in \mathscr{Z}(i). \tag{1}$$

z is a control parameter ranging in a compact metric space $\mathscr{Z}(i)$. The probabilities p are supposed to be continuous in z. They are transition probabilities of S.

Under a stationary control policy the control parameter value is a function of the actual state of S only. This function is a vector $\sigma \in \mathscr{Z}(1) \times \mathscr{Z}(2) \times \ldots = \mathscr{Z}^\infty$. The random sequence $\{X_n, n = 0, 1, \ldots\}$ is then a homogenous Markov chain with transition probability matrix

$$P_\sigma = \|p(i, j; (\sigma)_i)\|_{i,j\in I}.$$

$(\sigma)_i$ denotes the ith component of the (column) vector σ. Space $\mathscr{Z}^\infty$ is called the *set of stationary policies*. Under policy σ the control parameter value at time n equals

$$Z_n = (\sigma)_{X_n}, \quad n = 0, 1, \ldots \tag{2}$$

The random sequence $\{Z_n, n = 0, 1, \ldots\}$ is called the *control*.

Further, we associate with the trajectory of S an additive function of the state and of the control

$$R_N = \sum_{n=0}^{N-1} r(X_n, Z_n), \quad N = 1, 2, \ldots$$

$r(i, z)$, $i \in I$, $z \in \mathscr{Z}(i)$ is a function continuous in z. We shall call R_N the *reward up to time* N. If (2) holds, we can also write

$$R_N = \sum_{n=0}^{N-1} (r_\sigma)_{X_n}, \quad N = 1, 2, \ldots$$

r_σ is a vector such that $(r_\sigma)_i = r\big(i, (\sigma)_i\big)$, $i \in I$.

In the general case we understand under control a non-anticipative one. This means that Z_n is a function of $X_0, X_1, \ldots, X_n$,

$$(3) \qquad Z_n = \omega_n(X_0, \ldots, X_n), \quad n = 0, 1, \ldots,$$

and that

$$P(X_{n+1} = i | X_0, \ldots, X_n) = p(X_n, i; Z_n), \quad n = 0, 1, \ldots$$

2. Stationary policies

Let $\sigma \in \mathscr{Z}^\infty$. The *mean reward under policy* σ is

$$\lim_{N\to\infty} N^{-1} R_N = \mu_\sigma$$

provided that the limit exists almost surely (a.s.), and that μ_σ is a constant. A stationary policy $\hat\sigma$ is *optimal* (with respect to the mean reward) if

$$\mu_{\hat\sigma} = \hat\mu \quad \text{where} \quad \hat\mu = \sup_{\sigma\in\mathscr{Z}^\infty} \mu_\sigma.$$

Monograph [3] contains the following sufficient condition for the existence of an optimal stationary policy. To formulate it, we denote for $i_1 \in I$ by $\tilde P_\sigma$ the matrix which is obtained from P_σ by replacing the elements of its i_1th column by zeros (column-restriction). e denotes the vector whose components are all equal to 1. All vectors are infinite-dimensional. The symbol $\mathrm{H}_1(r_\sigma)$ for the hypothesis is used to point out the vector (family of vectors) to which the condition refers.

$\mathrm{H}_1(r_\sigma)$. *There exists an* $i_1 \in I$ *and a vector* $y_1 \geqslant 0$ *such that*

(i) $|r_\sigma| + e + \tilde P_\sigma y_1 \leqslant y_1, \quad \sigma \in \mathscr{Z}^\infty.$

(ii) $\lim_{N\to\infty} \tilde P_\sigma^N y_1 = 0, \quad \sigma \in \mathscr{Z}^\infty.$

(iii) $\lim_{\sigma\to\sigma_0} \tilde P_\sigma y_1 = \tilde P_{\sigma_0} y_1, \quad \sigma_0 \in \mathscr{Z}^\infty.$

(If a is a vector, then $|a|$ is the vector whose components satisfy $(|a|)_i = |(a)_i|$, $i \in I$. Analogously we define a^2, etc. The convergence of vectors is component-wise convergence.)

Assumption (i) is a Liapounov-type condition ensuring the existence of μ_σ for all $\sigma \in \mathscr{Z}^\infty$. Namely, from (i) follows

$$\sum_{n=0}^{\infty} \tilde P_\sigma^n e \leqslant y_1, \qquad \sum_{n=0}^{\infty} \tilde P_\sigma^n |r_\sigma| \leqslant y_1.$$

The first inequality shows that the transition probability matrix P_σ belongs to a Markov chain with one class of positively recurrent states, and possibly with transient states. According to [2] (Theorem 15.1)

$$\lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} (r_\sigma)_{X_n} = \Big(\sum_{n=0}^{\infty} \tilde P_\sigma^n r_\sigma\Big)_{i_1} \Big/ \Big(\sum_{n=0}^{\infty} \tilde P_\sigma^n e\Big)_{i_1} = \mu_\sigma \text{ a.s.}$$

Since

$$|\mu_\sigma| \leqslant (y_1)_{i_1},$$

$\hat\mu$ is finite.

THEOREM 1 ([3]). *Let* $\mathrm{H}_1(r_\sigma)$ *hold. Then there exists an optimal stationary strategy.*

The proof is divided into a sequence of lemmas, needed also in subsequent sections. (iii) implies the following proposition:

LEMMA 1. *Let* $\{q_n, n = 1, 2, \ldots\}$ *be a sequence of vectors,* $|q_n| \leqslant \text{const}\, y_1$, $\lim_{n\to\infty} q_n = q_\infty$. *Further, let* $\{\sigma_n, n = 1, 2, \ldots\}$, $\sigma_n \in \mathscr{Z}^\infty$, $\lim_{n\to\infty} \sigma_n = \sigma_\infty$. *Then*

$$\lim_{n\to\infty} \tilde P_{\sigma_n} q_n = \tilde P_{\sigma_\infty} q_\infty.$$

LEMMA 2. *Define*

$$x_0 = y_1, \quad x_{n+1} = \sup_\sigma \tilde P_\sigma x_n, \quad n = 0, 1, \ldots$$

Then

$$\tilde P_{\sigma_0} \tilde P_{\sigma_1} \ldots \tilde P_{\sigma_{n-1}} y_1 \leqslant x_n, \quad \sigma_0, \ldots, \sigma_{n-1} \in \mathscr{Z}^\infty,$$

$$x_{n+1} \leqslant x_n, \quad n = 0, 1, \ldots, \quad \lim_{n\to\infty} x_n = 0.$$

Proof. The inequalities follow by induction, since in virtue of (i) $\tilde P_\sigma x_0 \leqslant x_0$, $\sigma \in \mathscr{Z}_\infty$, and hence $x_1 \leqslant x_0$. Further, write $x_\infty = \lim_{n\to\infty} x_n$. Then

$$(4) \qquad \tilde P_\sigma x_\infty \leqslant \varliminf_{n\to\infty} \tilde P_\sigma x_n \leqslant \lim_{n\to\infty} x_n = x_\infty, \quad \sigma \in \mathscr{Z}^\infty.$$

Let $\{\sigma_n, n = 1, 2, \ldots\}$ be chosen so that

$$x_{n+1} \leqslant \tilde P_{\sigma_n} x_n + n^{-1} e, \quad n = 1, 2, \ldots, \quad \lim_{n\to\infty} \sigma_n = \sigma_\infty.$$

By Lemma 1,

$$(5) \qquad x_\infty \leqslant \lim_{n\to\infty} \tilde P_{\sigma_n} x_n = \tilde P_{\sigma_\infty} x_\infty.$$

(4), (5) imply

$$x_\infty = \tilde P_{\sigma_\infty} x_\infty = \tilde P_{\sigma_\infty}^n x_\infty \leqslant \tilde P_{\sigma_\infty}^n y_1 \to 0,$$

in virtue of (ii). ∎

LEMMA 3. *For an arbitrary sequence* $\{\sigma_n, n = 0, 1, \ldots\}$

$$\sum_{n=0}^{\infty} \tilde{P}_{\sigma_0}\tilde{P}_{\sigma_1}\ldots\tilde{P}_{\sigma_{n-1}}|r_{\sigma_n}-\hat{\mu}e| \leqslant (1+|\hat{\mu}|)y_1.$$

(A void product is the unit matrix.)

Proof. For an arbitrary N

$$|r_{\sigma_N}-\hat{\mu}e|+(1+|\hat{\mu}|)\tilde{P}_{\sigma_N}y_1 \leqslant (1+|\hat{\mu}|)y_1,$$

$$|r_{\sigma_{N-1}}-\hat{\mu}e|+P_{\sigma_{N-1}}|r_{\sigma_N}-\hat{\mu}e|+(1+|\hat{\mu}|)\tilde{P}_{\sigma_{N-1}}\tilde{P}_{\sigma_N}y_1 \leqslant (1+|\hat{\mu}|)y_1,$$

$$\sum_{n=0}^{N} \tilde{P}_{\sigma_0}\tilde{P}_{\sigma_1}\ldots\tilde{P}_{\sigma_{n-1}}|r_{\sigma_n}-\hat{\mu}e|+(1+|\hat{\mu}|)\tilde{P}_{\sigma_0}\ldots\tilde{P}_{\sigma_N}y_1 \leqslant (1+|\hat{\mu}|)y_1. \ \blacksquare$$

COROLLARY 1. *We have*

$$\tilde{P}_{\sigma_0}\ldots\tilde{P}_{\sigma_{n-1}}\Big(\sum_{m=0}^{\infty}\tilde{P}_{\sigma_n}\ldots\tilde{P}_{\sigma_{n+m-1}}|r_{\sigma_{n+m}}-\hat{\mu}e|\Big) \leqslant (1+|\hat{\mu}|)x_n, \quad n = 0, 1, \ldots, \tag{6}$$

with x_n *as defined in Lemma* 2.

LEMMA 4. *Define*

$$w_N = \sup_{\sigma_0,\ldots,\sigma_N}\sum_{n=0}^{N}\tilde{P}_{\sigma_0}\tilde{P}_{\sigma_1}\ldots\tilde{P}_{\sigma_{n-1}}(r_{\sigma_n}-\hat{\mu}e).$$

There exist $\hat{\sigma}_0, \ldots, \hat{\sigma}_N$ *such that*

$$w_N = \sum_{n=0}^{N-1}\tilde{P}_{\hat{\sigma}_0}\tilde{P}_{\hat{\sigma}_1}\ldots\tilde{P}_{\hat{\sigma}_{n-1}}(r_{\hat{\sigma}_n}-\hat{\mu}e).$$

Moreover,

$$w_N = \sup_{\sigma}\{r_\sigma-\hat{\mu}e+\tilde{P}_\sigma w_{N-1}\}, \qquad N = 1, 2, \ldots$$

The proof is by means of induction using Lemma 1. ■

Introduce

$$w = \sup_{\sigma_0,\sigma_1,\ldots}\sum_{n=0}^{\infty}\tilde{P}_{\sigma_0}\tilde{P}_{\sigma_1}\ldots\tilde{P}_{\sigma_{n-1}}(r_{\sigma_n}-\hat{\mu}e).$$

By Lemma 3, $|w| \leqslant (1+|\hat{\mu}|)y_1$.

LEMMA 5. w *satisfies*

$$w = \sup_{\sigma}\{r_\sigma-\hat{\mu}e+\tilde{P}_\sigma w\}. \tag{7}$$

Proof.

$$w = \sup_{\sigma,\sigma_1,\ldots}\Big\{r_\sigma-\hat{\mu}e+\tilde{P}_\sigma\sum_{n=1}^{\infty}\tilde{P}_{\sigma_1}\ldots\tilde{P}_{\sigma_{n-1}}(r_{\sigma_n}-\hat{\mu}e)\Big\} \leqslant \sup_{\sigma}\{r_\sigma-\hat{\mu}e+\tilde{P}_\sigma w\}. \tag{8}$$

To demonstrate the reverse inequality we use (6) to get

$$\Big|\sum_{n=0}^{N}\tilde{P}_{\sigma_0}\ldots\tilde{P}_{\sigma_{n-1}}(r_{\sigma_n}-\hat{\mu}e)-\sum_{n=0}^{\infty}\tilde{P}_{\sigma_0}\ldots\tilde{P}_{\sigma_{n-1}}(r_{\sigma_n}-\hat{\mu}e)\Big| \leqslant (1+|\hat{\mu}|)x_{N+1}.$$

Hence,

$$|w_N-w| \leqslant (1+|\hat{\mu}|)x_{N+1}.$$

From

$$w_N \geqslant r_\sigma-\hat{\mu}e+\tilde{P}_\sigma w_{N-1}$$

then follows

$$w+(1+|\hat{\mu}|)x_{N+1} \geqslant r_\sigma-\hat{\mu}e+\tilde{P}_\sigma w-(1+|\hat{\mu}|)\tilde{P}_\sigma x_N$$
$$\geqslant r_\sigma-\hat{\mu}e+\tilde{P}_\sigma w-(1+|\hat{\mu}|)x_{N+1}.$$

Since N was arbitrary, and, by Lemma 2, $\lim_{N\to\infty} x_N = 0$, we get

$$w \geqslant r_\sigma-\mu e+\tilde{P}_\sigma w. \tag{9}$$

(8), (9) are equivalent to (7). ■

LEMMA 6. *There exists a* $\hat{\sigma} \in \mathfrak{X}^\infty$ *such that*

$$w = r_{\hat{\sigma}}-\hat{\mu}e+\tilde{P}_{\hat{\sigma}}w, \tag{10}$$

$$w = \sum_{n=0}^{\infty}\tilde{P}_{\hat{\sigma}}^n(r_{\hat{\sigma}}-\hat{\mu}e). \tag{11}$$

Proof. (10) is proved by using Lemma 1. From (10) we get successively

$$w = \sum_{n=0}^{N}\tilde{P}_{\hat{\sigma}}^n(r_{\hat{\sigma}}-\hat{\mu}e)+\tilde{P}_{\hat{\sigma}}^{N+1}w.$$

Furthermore,

$$|\tilde{P}_{\hat{\sigma}}^{N+1}w| \leqslant (1+|\hat{\mu}|)x_{N+1} \to 0. \ \blacksquare$$

COROLLARY 2. $(w)_{i_1} = 0$.

Proof. From the definition of $\hat{\mu}$ follows

$$\Big(\sum_{n=0}^{\infty}\tilde{P}_{\hat{\sigma}}^N r_{\hat{\sigma}}\Big)_{i_1}\Big/\Big(\sum_{n=0}^{\infty}\tilde{P}_{\hat{\sigma}}^n e\Big)_{i_1} \leqslant \hat{\mu}, \quad \text{i.e.,} \quad (w)_{i_1} \leqslant 0.$$

On the other hand, for an arbitrary $\varepsilon > 0$ there exists a σ such that

$$\Big(\sum_{n=0}^{\infty}\tilde{P}_{\sigma}^n r_{\sigma}\Big)_{i_1}\Big/\Big(\sum_{n=0}^{\infty}\tilde{P}_{\sigma}^n e\Big)_{i_1} \geqslant \hat{\mu}-\varepsilon.$$

Hence,

$$(w)_{i_1} \geqslant \Big(\sum_{n=0}^{\infty}\tilde{P}_{\sigma}^n(r_\sigma-\hat{\mu}e)\Big)_{i_1} \geqslant -\varepsilon(y_1)_{i_1}.$$

Since ε was arbitrary, the assertion is proved. ■

Proof of Theorem 1. By Corollary 2,

$$\mu_{\hat{\sigma}} = \Big(\sum_{n=0}^{\infty} \tilde{P}_{\hat{\sigma}}^{n} r_{\hat{\sigma}}\Big)_{i_1} \Big/ \Big(\sum_{n=0}^{\infty} \tilde{P}_{\hat{\sigma}}^{n} e\Big)_{i_1} = \hat{\mu}.$$

Consequently, $\hat{\sigma}$ is an optimal policy. ∎

COROLLARY 3. *With regard to Corollary* 2 (7), (10) *can be replaced by*

(12) $$w = \sup_{\sigma}\{r_\sigma - \hat{\mu}e + P_\sigma w\},$$

(13) $$w = r_{\hat{\sigma}} - \hat{\mu}e + P_{\hat{\sigma}} w.$$

THEOREM 2. *Let* $H_1(r_\sigma)$ *hold. Then* $\hat{\mu}$ *is the unique number for which there exists a vector* w *satisfying* (12) *together with* $|w| \leqslant \mathrm{const}\, y_1$. w *is determined up to a shift for a vector* const e.

Proof. Let

(14) $$\overline{w} = \sup_{\sigma}\{r_\sigma - \overline{\mu}e + P_\sigma \overline{w}\}.$$

By means of a shift we can achieve $(w)_{i_1} = (\overline{w})_{i_1}$. From (13), (14) follows

$$w - \overline{w} \leqslant (\hat{\mu} - \overline{\mu})e + P_{\hat{\sigma}}(w - \overline{w}) = (\hat{\mu} - \overline{\mu})e + \tilde{P}_{\hat{\sigma}}(w - \overline{w}).$$

Hence,

$$w - \overline{w} \leqslant (\hat{\mu} - \overline{\mu})\Big(\sum_{n=0}^{N-1} \tilde{P}_{\hat{\sigma}}^{n} e\Big) + \tilde{P}_{\hat{\sigma}}^{N}(w - \overline{w}).$$

Further, by (ii),

$$0 = (w - \overline{w})_{i_1} \leqslant (\hat{\mu} - \overline{\mu})\Big(\sum_{n=0}^{\infty} \tilde{P}_{\hat{\sigma}}^{n} e\Big)_{i_1}, \quad \text{i.e.,} \quad \hat{\mu} \geqslant \overline{\mu}.$$

Since $\overline{\mu}$ and $\hat{\mu}$ can be interchanged, we conclude that $\hat{\mu} = \overline{\mu}$. Therefore

$$w - \overline{w} \leqslant \tilde{P}_{\hat{\sigma}}(w - \overline{w}) \quad \text{or} \quad w - \overline{w} \leqslant \tilde{P}_{\hat{\sigma}}^{n}(w - \overline{w}),\ n = 1, 2, \ldots$$

Letting $n \to \infty$, we get $w - \overline{w} \leqslant 0$. The reverse inequality is proved in the same way. ∎

COROLLARY 4. *Let* $\overline{\sigma} \in \mathscr{Z}^\infty$. $\mu_{\overline{\sigma}}$ *is the unique number for which there exists a vector* $w_{\overline{\sigma}}$ *such that* $|w_{\overline{\sigma}}| \leqslant \mathrm{const}\, y_1$, *and*

(15) $$0 = r_{\overline{\sigma}} - \mu_{\overline{\sigma}} e + P_{\overline{\sigma}} w_{\overline{\sigma}} - w_{\overline{\sigma}};$$

$w_{\overline{\sigma}}$ *is determined up to a shift for a vector* const e.

Proof. The assertion is Theorem 2 applied to the special case where $\mathscr{Z}^\infty$ consists of one point only. ∎

3. Law of large numbers

Here and in the subsequent sections we shall not limit ourselves to stationary policies. However, we have to make stronger assumptions. The basic one is the following:

$H_2(r_\sigma)$. *We have* $H_1(r_\sigma)$, *and there exists an* $i_2 \in I$ (*index of column restriction*) *and a vector* $y_2 \geqslant 0$ *such that*

(i) $y_1^2 + \tilde{P}_\sigma y_2 \leqslant y_2$, $\sigma \in \mathscr{Z}^\infty$.

(ii) $\lim_{N\to\infty} \tilde{P}_\sigma^N y_2 = 0$, $\sigma \in \mathscr{Z}^\infty$.

(iii) $\lim_{\sigma\to\sigma_0} \tilde{P}_\sigma y_2 = \tilde{P}_{\sigma_0} y_2$, $\sigma_0 \in \mathscr{Z}^\infty$.

We can also write briefly $H_2(r_\sigma) = H_1(r_\sigma) \,\&\, H_1(y_1^2)$. The validity of $H_2(r_\sigma)$ is assumed throughout this section. Further, let (3) hold, and let X_0 be non-random. Let $\overline{\sigma} \in \mathscr{Z}^\infty$ be fixed. The subscript $\overline{\sigma}$ in $\mu_{\overline{\sigma}}$, $w_{\overline{\sigma}}$ will be omitted. We are going to define a measure of difference between the actual parameter value Z_n, as given by (3), and the value $(\overline{\sigma})_{X_n}$, which corresponds to the stationary policy $\overline{\sigma}$.

Introduce

$$\varphi(i, z) = r(i, z) - \mu + \sum_j p(i, j; z)(w)_j - (w)_i, \quad i \in I,\ z \in \mathscr{Z}(i).$$

Since $|w| \leqslant \mathrm{const}\, y_1$, the series on the right converges absolutely, $\varphi(i, z)$ is a continuous function of z, and $|\varphi(i, z)| \leqslant \mathrm{const}(y_1)_i$, $i \in I$, $z \in \mathscr{Z}(i)$. (15) can be written as $\varphi(i, (\overline{\sigma})_i) = 0$. The difference in equation (15) at time n therefore equals

(16) $$\varphi(X_n, Z_n) = \varphi(X_n, Z_n) - \varphi(X_n, (\overline{\sigma})_{X_n}), \quad n = 0, 1, \ldots$$

The applicability of the quantity φ is based on the fact that

(17) $$M_N = R_N - N\mu + (w)_{X_N} - (w)_{X_0} - \sum_{n=0}^{N-1} \varphi(X_n, Z_n), \quad N = 1, 2, \ldots,$$

is a martingale. (We set $M_0 \equiv 0$.) To verify this, introduce the martingale differences

(18) $$\begin{aligned} Y_n &= M_{n+1} - M_n \\ &= r(X_n, Z_n) - \mu + (w)_{X_{n+1}} - (w)_{X_n} - \varphi(X_n, Z_n), \quad n = 0, 1, \ldots, \end{aligned}$$

and the σ-algebras $\mathscr{F}_n = \sigma a(X_0, \ldots, X_n)$ of random events defined on the trajectory up to time n. Then

$$E\{Y_n | \mathscr{F}_n\} = r(X_n, Z_n) - \mu + \sum_j p(X_n, j; Z_n)(w)_j - (w)_{X_n} - \varphi(X_n, Z_n) = 0.$$

Before demonstrating the law of large numbers for $\{M_n, n = 0, 1, \ldots\}$ we derive an auxiliary result.

LEMMA 7. *We have*

(19) $$\overline{\lim_{N\to\infty}}\, N^{-1} \sum_{n=0}^{N-1} E\big(r(X_n, Z_n)^2 + (y_1^2)_{X_n}\big) \leqslant \mathrm{const},$$

(20) $$\sum_{n=1}^{\infty} n^{-(1+\gamma)} E\big(r(X_n, Z_n)^2 + (y_1^2)_{X_n}\big) < \infty, \quad \gamma > 0.$$

The constant in (19) *depends neither on the control policy nor on the initial state.*

Proof. $H_2(r_\sigma)$ implies

$$r(i,z)^2+(y_1^2)_i+2\sum_j p(i,j;z)(y_2)_j \leqslant 2(y_2)_i+2p(i,i_2;z)(y_2)_{i_2}$$
$$\leqslant 2(y_2)_i+2(y_2)_{i_2}, \quad i\in I,\ z\in\mathscr{Z}(i).$$

Hence,

$$E\big(r(X_n,Z_n)^2+(y_1^2)_{X_n}\big)+2E(y_2)_{X_{n+1}} \leqslant 2E(y_2)_{X_n}+2(y_2)_{i_2}.$$

Adding for $n = 0, \ldots, N-1$, we get

$$(21)\qquad \sum_{n=0}^{N-1} E\big(r(X_n,Z_n)^2+(y_1^2)_{X_n}\big) \leqslant 2(y_2)_{X_0}+2N(y_2)_{i_2}.$$

From here (19) easily follows.

Denote by S_N the sum on the left-hand side of (21). Then

$$\sum_{n=1}^{\infty} n^{-(1+\gamma)}E\big(r(X_n,Z_n)^2+(y_1^2)_{X_n}\big) \leqslant \sum_{n=1}^{\infty} S_{n+1}\big(n^{-(1+\gamma)}-(n+1)^{-(1+\gamma)}\big)$$
$$\leqslant \text{const}\sum_{n=1}^{\infty} n^{-(1+\gamma)}. \blacksquare$$

LEMMA 8. *We have*

$$(22)\qquad \lim_{N\to\infty} N^{-1}M_N = 0 \ \text{a.s.},$$

$$(23)\qquad \lim_{N\to\infty} N^{-2}EM_N^2 = 0.$$

Proof. We have

$$Y_n^2 \leqslant \text{const}\big(r(X_n,Z_n)^2+(y_1^2)_{X_n}+(y_1^2)_{X_{n+1}}\big).$$

With regard to Lemma 7 it follows that $\sum_{n=1}^{\infty} n^{-2}EY_n^2<\infty$, which is a sufficient condition for the validity of (22). Further, from (19),

$$\overline{\lim_{N\to\infty}} N^{-2}EM_N^2 = \overline{\lim_{N\to\infty}} N^{-2}\sum_{n=0}^{N-1} EY_n^2 \leqslant \text{const}\ \overline{\lim_{N\to\infty}} N^{-2}\sum_{n=0}^{N} E\big(r(X_n,Z_n)^2+(y_1^2)_{X_n}\big) = 0. \blacksquare$$

From Lemma 8 we deduce the following assertion about the convergence of $N^{-1}R_N$:

THEOREM 3. *Let* $H_2(r_\sigma)$ *hold. Then*

$$\lim_{N\to\infty} N^{-1}R_N = \mu$$

in probability (a.s.) *if and only if*

$$\lim_{N\to\infty} N^{-1}\sum_{n=0}^{N-1}\varphi(X_n,Z_n) = 0$$

in probability (a.s.). *Likewise, for* $k = 1, 2,$

$$(24)\qquad \lim_{N\to\infty} E|N^{-1}R_N-\mu|^k = 0$$

holds if and only if

$$(25)\qquad \lim_{N\to\infty} N^{-k}E\Big|\sum_{n=0}^{N-1}\varphi(X_n,Z_n)\Big|^k = 0.$$

Proof. By Lemma 7,

$$E\sum_{n=1}^{\infty} n^{-2}(y_1^2)_{X_n} < \infty.$$

Hence,

$$\sum_{n=1}^{\infty} n^{-2}(y_1^2)_{X_n} < \infty \ \text{a.s.}, \qquad \lim_{N\to\infty} N^{-1}(y_1)_{X_N} = 0 \ \text{a.s.}$$

This implies

$$\lim_{N\to\infty} N^{-1}(w)_{X_N} = 0 \ \text{a.s.}, \qquad \lim_{N\to\infty} N^{-2}E(w^2)_{X_N} = 0.$$

Lemma 8 and (17) then imply

$$(26)\qquad \lim_{N\to\infty}\Big(N^{-1}R_N-\mu-N^{-1}\sum_{n=0}^{N-1}\varphi(X_n,Z_n)\Big) = 0 \ \text{a.s.},$$

$$(27)\qquad \lim_{N\to\infty} E\Big(N^{-1}R_N-\mu-N^{-1}\sum_{n=0}^{N-1}\varphi(X_n,Z_n)\Big)^2 = 0.$$

From this the assertion of the theorem is obtained without difficulty. $\blacksquare$

THEOREM 4. *Let* $H_2(r_\sigma)$ *hold. Then under arbitrary control*

$$(28)\qquad \overline{\lim_{N\to\infty}} N^{-1}R_N \leqslant \hat{\mu} \ \text{a.s.},$$

$$(29)\qquad \overline{\lim_{N\to\infty}} N^{-1}ER_N \leqslant \hat{\mu}.$$

Proof. Assume $\bar{\sigma} = \hat{\sigma}$, where $\hat{\sigma}$ is such that (12), (13) hold. Consequently, $\mu = \hat{\mu}$ and $\varphi(i,z)\leqslant 0$, $i\in I$, $z\in\mathscr{Z}(i)$. Thus from (26) follows

$$\overline{\lim_{N\to\infty}} N^{-1}R_N = \hat{\mu}+\overline{\lim_{N\to\infty}} N^{-1}\sum_{n=0}^{N-1}\varphi(X_n,Z_n)\leqslant\hat{\mu} \ \text{a.s.}$$

Analogously, (29) follows from (27). $\blacksquare$

Theorems 5, 6, which also rely on Theorem 3, will be preceded by a sequence of lemmas. Define vectors e_k, d_k, $k = 0, 1, \ldots$ as follows:

$$(e_k)_i = (d_k)_i = 0, \quad i<k, \quad (e_k)_i = 1, \quad (d_k)_i = (y_1)_i, \quad i\geqslant k.$$

Recall that $y_1\geqslant e$. Hence, $d_k\geqslant e_k$, $k = 0, 1, \ldots$

LEMMA 9. $\sum_{n=0}^{\infty}\tilde{P}_\sigma^n e_k$, $\sum_{n=0}^{\infty}\tilde{P}_\sigma^n d_k$, $k = 0, 1, \ldots,$ *depend continuously on* $\sigma\in\mathscr{Z}^\infty$.

($\tilde{P}_\sigma$ *column restriction in* i_2.)

The proof by the methods of Section 2 is not difficult. $\blacksquare$

Set

$$
(30)\quad
\begin{aligned}
\varepsilon_\sigma(k) &= \Big(\sum_{n=0}^{\infty} \tilde{P}_\sigma^n e_k\Big)_{i_2} \Big(\sum_{n=0}^{\infty} \tilde{P}_\sigma^n e\Big)_{i_2}^{-1},\\
\vartheta_\sigma(k) &= \Big(\sum_{n=0}^{\infty} \tilde{P}_\sigma^n d_k\Big)_{i_2} \Big(\sum_{n=0}^{\infty} \tilde{P}_\sigma^n e\Big)_{i_2}^{-1},\\
\hat{\varepsilon}(k) &= \sup_\sigma \varepsilon_\sigma(k),\\
\hat{\vartheta}(k) &= \sup_\sigma \vartheta_\sigma(k).
\end{aligned}
$$

Obviously, $\hat{\vartheta}(k) \geqslant \hat{\varepsilon}(k)$, $k = 0, 1, \ldots$ The analogy between $\varepsilon_\sigma, \vartheta_\sigma, \hat{\varepsilon}, \hat{\vartheta}$, and $\mu_\sigma, \hat{\mu}$ from the preceding section is evident.

LEMMA 10. *We have*

$$
(31)\quad \lim_{k\to\infty} \hat{\varepsilon}(k) = 0 = \lim_{k\to\infty} \hat{\vartheta}(k).
$$

Proof. If (31) were not true, a $\delta > 0$ could be found such that $\hat{\vartheta}(k) > \delta$ for $k = 0, 1, \ldots$ Consider a sequence $\{\sigma_k, k = 0, 1, \ldots\}$ satisfying $\vartheta_{\sigma_k}(k) > \delta$. Since $\mathscr{Z}^\infty$ is compact, $\{\sigma_k, k = 0, 1, \ldots\}$ has an accumulation point σ_∞. Let m be a positive integer. By (30) and by Lemma 9, $\vartheta_\sigma(m)$ is a continuous function of σ. Consequently,

$$
\vartheta_{\sigma_\infty}(m) = \lim_{k\to\infty} \vartheta_{\sigma_k}(m) \geqslant \overline{\lim_{k\to\infty}}\, \vartheta_{\sigma_k}(k) \geqslant \delta > 0.
$$

This is a contradiction, because $\lim\limits_{m\to\infty} \vartheta_{\sigma_\infty}(m) = 0$. ∎

LEMMA 11. *We have*

$$
(32)\quad \overline{\lim_{N\to\infty}}\, N^{-1} \sum_{n=0}^{N-1} P(X_n \geqslant k) \leqslant \hat{\varepsilon}(k), \qquad k = 0, 1, \ldots
$$

If, in addition, $H_2(y_1)$ *is valid, then*

$$
(33)\quad \overline{\lim_{N\to\infty}}\, N^{-1} \sum_{n=0}^{N-1} \chi_{\{X_n \geqslant k\}} (y_1)_{X_n} \leqslant \hat{\vartheta}(k) \quad \text{a.s.},
$$

$$
(34)\quad \overline{\lim_{N\to\infty}}\, N^{-1} \sum_{n=0}^{N-1} E\chi_{\{X_n \geqslant k\}} (y_1)_{X_n} \leqslant \hat{\vartheta}(k), \qquad k = 0, 1, \ldots
$$

Proof. (32) is (29) with R_N replaced by $\sum\limits_{n=0}^{N-1} \chi_{\{X_n \geqslant k\}}$. Analogously, under $H_2(y_1)$, (33), (34) are counterparts of (28), (29). ∎

THEOREM 5. *Let* $H_2(r_\sigma)$ *hold. If*

$$
(35)\quad \lim_{n\to\infty} \varrho(Z_n, (\overline{\sigma})_{X_n}) = 0 \ \textit{in prob.},
$$

then

$$
(36)\quad \lim_{N\to\infty} E|N^{-1}R_N - \mu| = 0.
$$

(ϱ *is the distance.*)

Proof. Assume (35). By Theorem 3, we have to demonstrate

$$
(37)\quad \lim_{N\to\infty} N^{-1} E\Big|\sum_{n=0}^{N-1} \varphi(X_n, Z_n)\Big| = 0.
$$

Let $L > 0$, $k > 0$ be numbers, k an integer. Write

$$
N^{-1} \sum_{n=0}^{N-1} \varphi(X_n, Z_n) = N^{-1} \sum_{n=0}^{N-1} [\chi_{\{X_n<k\}} + \chi_{\{X_n\geqslant k\}} \chi_{\{(y_1)_{X_n}<L\}} + \chi_{\{X_n\geqslant k\}} \chi_{\{(y_1)_{X_n}\geqslant L\}}] \varphi(X_n, Z_n).
$$

Since $|\varphi(X_n, Z_n)| \leqslant \mathrm{const}(y_1)_{X_n}$, we obtain

$$
N^{-1} E\Big|\sum_{n=0}^{N-1} \varphi(X_n, Z_n)\Big| \leqslant N^{-1} \sum_{n=0}^{N-1} E|\chi_{\{X_n<k\}} \varphi(X_n, Z_n)| +
$$
$$
+ L\,\mathrm{const}\,N^{-1} \sum_{n=0}^{N-1} P(X_n \geqslant k) + L^{-1}\mathrm{const}\,N^{-1} \sum_{n=0}^{N-1} E(y_1^2)_{X_n}.
$$

It is easily seen that

$$
\lim_{n\to\infty} |\chi_{\{X_n<k\}} \varphi(X_n, Z_n)| = \lim_{n\to\infty} \chi_{\{X_n<k\}} |\varphi(X_n, Z_n) - \varphi(X_n, (\overline{\sigma})_{X_n})| = 0 \ \text{in prob.}
$$

Hence, with regard to the boundedness of the integrand,

$$
\lim_{n\to\infty} E|\chi_{\{X_n<k\}} \varphi(X_n, Z_n)| = 0 \quad \text{or} \quad \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} E|\chi_{\{X_n<k\}} \varphi(X_n, Z_n)| = 0.
$$

By Lemma 11,

$$
(38)\quad \overline{\lim_{N\to\infty}}\, L\,\mathrm{const}\,N^{-1} \sum_{n=0}^{N-1} P(X_n \geqslant k) \leqslant L\,\mathrm{const}\,\hat{\varepsilon}(k),
$$

and by Lemma 7,

$$
(39)\quad \overline{\lim_{N\to\infty}}\, L^{-1}\mathrm{const}\,N^{-1} \sum_{n=0}^{N-1} E(y_1^2)_{X_n} \leqslant L^{-1}\mathrm{const}.
$$

The right-hand sides of (38), (39) can be made arbitrarily small by a proper choice of L and of k according to Lemma 10. We conclude that (37) holds. ∎

THEOREM 6. *Let* $H_1(r_\sigma)$ *and* $H_2(y_1)$ *hold. If*

$$
\lim_{n\to\infty} \varrho(Z_n, (\overline{\sigma})_{X_n}) = 0 \ \text{a.s.},
$$

then

$$
(40)\quad \lim_{N\to\infty} N^{-1} R_N = \mu \ \text{a.s.}
$$

Proof. We have to verify

$$
(41)\quad \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} \varphi(X_n, Z_n) = 0 \ \text{a.s.}
$$

Under the hypotheses of the theorem,

$$\varlimsup_{N\to\infty} N^{-1}\Big|\sum_{n=0}^{N-1}\varphi(X_n, Z_n)\Big|$$
$$\leqslant \varlimsup_{N\to\infty} N^{-1}\Big|\sum_{n=0}^{N-1}\chi_{\{X_n<k\}}\varphi(X_n, Z_n)\Big| + \varlimsup_{N\to\infty}\operatorname{const} N^{-1}\sum_{n=0}^{N-1}\chi_{\{X_n\geqslant k\}}(y_1)_{X_n}$$
$$\leqslant \operatorname{const}\hat{\vartheta}(k) \text{ a.s.}$$

Letting $k\to\infty$, and using Lemma 10, we obtain (41). ∎

4. Central limit theorem

In this section we assume $H_1(r_\sigma)$ and $H_2(y_1^2)$. This implies $H_2(r_\sigma)$. First we shall consider the validity of the central limit theorem for the martingale $\{M_n, n = 0, 1, \ldots\}$ defined by (17). The following result is known from martingale theory ([1]):

LEMMA 12. *Let $\{Y_n, n = 0, 1, \ldots\}$ be as in (18). Further, let*

$$\lim_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} E\{Y_n^2|\mathscr{F}_n\} = \mu_2 \textit{ in prob.,} \tag{42}$$

where μ_2 is a constant, and for each $\varepsilon > 0$ let

$$\lim_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} EY_n^2\chi_{\{|Y_n|\geqslant\varepsilon\sqrt{N}\}} = 0. \tag{43}$$

Then $M_N/\sqrt{N}$, as $N\to\infty$, has an asymptotically normal distribution $N(0, \mu_2)$.

Replacing in Lemma 7 r by y_1^2 we get the next lemma.

LEMMA 13. *We have*

$$\varlimsup_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} E(y_1^4)_{X_n} \leqslant \operatorname{const},$$
$$E\sum_{n=1}^{\infty} n^{-(1+\gamma)}(y_1^4)_{X_n} < \infty, \quad \gamma > 0. \tag{44}$$

COROLLARY 5. *We have*

$$\varlimsup_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} EY_n^4 \leqslant \operatorname{const}, \tag{45}$$

and, consequently, (43) *is fulfilled.*

Proof. (45) follows directly from the definition of $\{Y_n, n = 0, 1, \ldots\}$ and from the inequality $|w| \leqslant \operatorname{const} y_1$.

Further, for $\varepsilon > 0$,

$$\varlimsup_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} EY_n^2\chi_{\{|Y_n|\geqslant\varepsilon\sqrt{N}\}} \leqslant \varlimsup_{N\to\infty}\varepsilon^{-2}N^{-2}\sum_{n=0}^{N-1} EY_n^4 = 0. \;∎$$

Consider (42). (18) implies

$$\begin{aligned} E\{Y_n^2|\mathscr{F}_n\} &= E\{(r(X_n, Z_n)-\mu)^2 + 2(r(X_n, Z_n)-\mu)(w)_{X_{n+1}} + \\ &\quad + (w^2)_{X_{n+1}} - (w^2)_{X_n} - 2(r(X_n, Z_n) - \mu + (w)_{X_{n+1}} - \\ &\quad - (w)_{X_n})((w)_{X_n} + \varphi(X_n, Z_n)) + \varphi(X_n, Z_n)^2|\mathscr{F}_n\} \\ &= r_2(X_n, Z_n) + E\{(w^2)_{X_{n+1}}|\mathscr{F}_n\} - (w^2)_{X_n} - \\ &\quad - \varphi(X_n, Z_n)(2(w)_{X_n} + \varphi(X_n, Z_n)), \end{aligned}$$

where

$$r_2(i, z) = (r(i, z)-\mu)^2 + 2(r(i, z)-\mu)\sum_j p(i, j; z)(w)_j, \quad i\in I,\ z\in\mathscr{Z}(i). \tag{46}$$

Consequently,

$$\begin{aligned} N^{-1}\sum_{n=0}^{N-1} E\{Y_n^2|\mathscr{F}_n\} &= N^{-1}\Big\{\sum_{n=0}^{N-1} r_2(X_n, Z_n) + \sum_{n=0}^{N-1}[E\{(w^2)_{X_{n+1}}|\mathscr{F}_n\} - (w^2)_{X_{n+1}}] + \\ &\quad + [(w^2)_{X_N} - (w^2)_{X_0}] - \sum_{n=0}^{N-1}\varphi(X_n, Z_n)(2(w)_{X_n} + \varphi(X_n, Z_n))\Big\}, \quad N = 1, 2, \ldots \end{aligned} \tag{47}$$

Let μ_2 denote the mean reward under policy $\bar\sigma$ if the trajectory is valued by (46). According to Corollary 4, μ_2 satisfies

$$0 = r_{2\bar\sigma} - \mu_2 e + P_{\bar\sigma} w_{2\bar\sigma} - w_{2\bar\sigma}.$$

Note that $|r_{2\sigma}| \leqslant \operatorname{const} y_1^2$. Thus $H_2(y_1^2)$ implies $H_2(r_{2\sigma})$.

THEOREM 7. *Let $H_1(r_\sigma)$ and $H_2(y_1^2)$ hold, and let*

$$\lim_{n\to\infty}\varrho(Z_n, (\bar\sigma)_{X_n}) = 0 \textit{ in prob.}$$

If also

$$\lim_{N\to\infty}\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\varphi(X_n, Z_n) = 0 \textit{ in prob.,}$$

then $(R_N - N\mu)/\sqrt{N}$ has an asymptotically normal distribution $N(0, \mu_2)$ as $N\to\infty$.

Proof. We first verify the hypotheses of Lemma 12. (43) holds by Corollary 5. To prove (42) we use (47). Since $H_2(r_{2\sigma})$ is valid, we get from Theorem 5

$$\lim_{N\to\infty} N^{-1}\sum_{n=0}^{N-1} r_2(X_n, Z_n) = \mu_2 \text{ in prob.}$$

The second term on the right-hand side of (47) is a martingale fulfilling the law of large numbers in virtue of $|w| \leqslant \operatorname{const} y_1$, and of Lemma 13. Further, we have

$$N^{-1}(w^2)_{X_N} \leqslant \operatorname{const} N^{-1}(y_1^2)_{X_N} \to 0 \text{ a.s.},$$

by (44) for $\gamma = 1$. The negligibility of the last sum in (47) is established by the method used in the proof of Theorem 5, since

$$|\varphi(X_n, Z_n)(2(w)_{X_n} + \varphi(X_n, Z_n))| \leqslant \operatorname{const}(y_1^2)_{X_n}, \quad n = 0, 1, \ldots$$

Hence, (42) is demonstrated.

Further, write

$$N^{-1/2}M_N = N^{-1/2}(R_N-N\mu)+N^{-1/2}\big((w)_{X_N}-(w)_{X_0}\big)-N^{-1/2}\sum_{n=0}^{N-1}\varphi(X_n, Z_n).$$

By Lemma 12, the left-hand side is asymptotically $N(0, \mu_2)$ as $N\to\infty$. The last term on the right is negligible by the hypotheses of the theorem. The negligibility of the last but one term follows from (44). This demonstrates the theorem. ∎

5. Law of iterated logarithm

Here we impose the hypotheses $H_2(r_\sigma)$ and $H_2(y_2)$.

LEMMA 14. *If*

$$\lim_{n\to\infty}\varrho\big(Z_n, (\overline{\sigma})_{X_n}\big) = 0 \text{ a.s.},$$

then

(48) $$\lim_{N\to\infty}\sum_{n=0}^{N-1}E\{Y_n^2|\mathscr{F}_n\} = \mu_2 \text{ a.s.}$$

Proof. The proof is based on (47), and is analogous to the verification of (42) in the proof of Theorem 7. The difference is that by Theorem 6 the stronger assumption $H_2(y_2)$ implies

$$\lim_{N\to\infty}N^{-1}\sum_{n=0}^{N-1}r_2(X_n, Z_n) = \mu_2 \text{ a.s.}$$

Also, when estimating the last term in (47), we proceed as in the proof of Theorem 6. ∎

A simple consequence of Theorem 3 in [5] is the subsequent lemma, needed for the proof of Theorem 8.

LEMMA 15. *Let* (48) *hold, and let for a* $\delta>0$

(49) $$\sum_{n=1}^{\infty}n^{-2+\delta}EY_n^4 < \infty.$$

Then

(50) $$\overline{\lim_{N\to\infty}}\pm(2N\log\log N)^{-1/2}M_N = \sqrt{\mu_2} \text{ a.s.}$$

THEOREM 8. *Let* $H_2(r_\sigma)$ *and* $H_2(y_2)$ *hold, and let*

$$\lim_{n\to\infty}\varrho\big(Z_n, (\overline{\sigma})_{X_n}\big) = 0 \text{ a.s.}$$

If also

$$\lim_{N\to\infty}(N\log\log N)^{-1/2}\sum_{n=0}^{N-1}\varphi(X_n, Z_n) = 0 \text{ a.s.},$$

then

$$\overline{\lim_{N\to\infty}}\pm(2N\log\log N)^{-1/2}(R_N-N\mu) = \sqrt{\mu_2} \text{ a.s.}$$

Proof. (48) and (49) are valid by Lemmas 14 and 13, respectively. Write

$$\pm(2N\log\log N)^{-1/2}(R_N-N\mu) = \pm(2N\log\log N)^{-1/2}M_N\pm$$
$$\pm(2N\log\log N)^{-1/2}\big((w)_{X_N}-(w)_{X_0}\big)\pm(2N\log\log N)^{-1/2}\sum_{n=0}^{N-1}\varphi(X_n, Z_n).$$

The last term is negligible by hypothesis. Since $|w|\leqslant \text{const}\,y_1$, (44) for $\gamma=1$ implies

$$\lim_{N\to\infty}(2N\log\log N)^{-1/2}(w)_{X_N} = 0 \text{ a.s.}$$

Hence, the last but one term is also negligible. The assertion of the theorem is thus a consequence of (50). ∎

Remark 1. All Liapounov type conditions imposed in this paper are satisfied if there exist positive numbers $\varkappa$, δ, and state $i_1\in I$ such that

$$|r(i,z)|\leqslant\varkappa,\quad p(i,i_1;z)\geqslant\delta,\quad i\in I,\ z\in\mathscr{Z}(i).$$

$H_1(r_\sigma)$ then holds with $y_1=\delta^{-1}(\varkappa+1)e$. Namely,

(i) $|r_\sigma|+e+\tilde{P}_\sigma y_1\leqslant\big(\varkappa+1+\delta^{-1}(\varkappa+1)(1-\delta)\big)e = y_1,\quad \sigma\in\mathscr{Z}^\infty.$

(ii) $\overline{\lim_{N\to\infty}}\tilde{P}_\sigma^N y_1\leqslant\lim_{N\to\infty}\delta^{-1}(\varkappa+1)(1-\delta)^N e = 0,\quad \sigma\in\mathscr{Z}^\infty.$

(iii) $\lim_{\sigma\to\sigma_0}(\tilde{P}_\sigma y_1)_i = \lim_{\sigma\to\sigma_0}\delta^{-1}(\varkappa+1)\big(1-p(i,i_1;(\sigma)_i)\big)$
$\quad = \delta^{-1}(\varkappa+1)\big(1-p(i,i_1,(\sigma_0)_i)\big) = (\tilde{P}_{\sigma_0}y_1)_i,\quad i\in I,\ \sigma_0\in\mathscr{Z}^\infty.$

y_1 is also bounded. Thus, in the same way, we verify $H_2(r_\sigma)$, $H_2(y_1^2)$ and $H_2(y_2)$.

References

[1] B. M. Brown, *Martingale central limit theorems*, Ann. Math. Statist. 42 (1971), pp. 59–66.

[2] K. L. Chung, *Markov chains with stationary transition probabilities*, Springer-Verlag, Berlin–Heidelberg 1960.

[3] A. Hordijk, *Dynamic programming and Markov potential theory*, Math. Centrum, Amsterdam 1974.

[4] P. Mandl, *Estimation and control in Markov chains*, Adv. Appl. Prob. 6 (1974), pp. 40–60.

[5] W. F. Stout, *On the Kolmogorov's law of iterated logarithm*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 15 (1970), pp. 279–290.

Presented to the Semester
Probability Theory
February 11–June 11, 1976