

ASYMPTOTIC NORMALITY OF DISCRETE-TIME MARKOV CONTROL PROCESSES

ARMANDO F. MENDOZA-PÉREZ,* *Universidad Politécnica de Chiapas*

ONÉSIMO HERNÁNDEZ-LERMA,** *CINVESTAV*

Abstract

In this paper we study the asymptotic normality of discrete-time Markov control processes in Borel spaces, with possibly unbounded cost. Under suitable hypotheses, we show that the cost sequence is asymptotically normal. As a special case, we obtain a central limit theorem for (noncontrolled) Markov chains.

Keywords: (Discrete-time) Markov control process; average cost criteria; average variance; asymptotic normality

2010 Mathematics Subject Classification: Primary 93E20; 90C40; 60F05; 60J05

1. Introduction.

In this paper we are concerned with discrete-time Markov control processes (MCPs) in Borel spaces, with the *long-run pathwise average cost*

$$S(\varphi, x) := \limsup_{n \rightarrow \infty} \frac{1}{n} S_n(\varphi, x),$$

where

$$S_n(\varphi, x) := \sum_{k=0}^{n-1} C_\varphi(x_k).$$

Here, φ denotes the control policy being used (see Section 2 for details), $\{x_k\}$ is the controlled state (Markov) process, and $x_0 = x$ is the initial state. The underlying optimal control problem is to minimize (almost surely) $S(\varphi, x)$ over the family of admissible control policies φ for every x . To analyze some aspects of this problem, it is important to determine the asymptotic behavior of the cost sequence $S_n(\varphi, x)$. In particular, if it can be ensured that it is asymptotically normal then we can deduce important optimality properties, which are particularly useful when the MCP depends on *unknown* parameters [13], [15].

In this paper we give reasonably mild conditions for the cost sequence $\{S_n(\varphi, x)\}$ to be asymptotically normal. This line of research was initiated by Mandl [14], and further extended by himself and his associates [11], [12], [13], [15] for MCPs with a finite state space. Mendoza-Pérez [17], [16] extended the results in [14] to discrete-time MCPs with Borel spaces. Other extensions include results by Lánská [9] for a class of controlled diffusions, and by Prieto-Rumeau and Hernández-Lerma [20] for continuous-time MCPs with a countable state space.

Received 11 March 2010; revision received 14 May 2010.

* Postal address: Universidad Politécnica de Chiapas, Calle Eduardo J. Selvas S/N, Tuxtla Gutiérrez, Chiapas, Mexico. Email address: mepa680127@hotmail.com

** Postal address: Department of Mathematics, CINVESTAV-IPN, A. Postal 14-740, Mexico DF 07000, Mexico. Email address: ohernand@math.cinvestav.mx

Research partially supported by CONACyT grant 104001.

In all of these references φ is assumed to be a *deterministic* and *optimal* policy. In contrast, here we obtain asymptotic normality for every randomized stationary policy φ (in a given set Φ —see Section 2) which may be neither deterministic nor optimal. This means that our main result, Theorem 3.1, yields in particular a central limit theorem for noncontrolled Markov chains. On the other hand, since we allow *randomized* policies φ , our results are applicable to constrained MDPs [17], [18], and to Markov games [5]. Moreover, it is worth noting that the existing proofs of the central limit theorem for Markov chains require strong ergodicity conditions, for instance, geometric ergodicity [19, Theorem 17.0.1] or polynomial ergodicity [8, Theorems 4.1, 4.2, 4.3, and 4.4]. Here, however, instead of an ergodicity requirement we impose a Lyapunov-like condition (see Assumption 3.2), in addition to growth and continuity-compactness conditions (Assumptions 3.1 and 3.3).

To obtain our results, we combine two approaches. The first approach is to show, under suitable conditions, the existence of solutions to the Poisson equation; see (3.4) below. The second approach is Mandl’s approach [14] using characteristic functions.

The remainder of the paper is organized as follows. Section 2 contains a brief description of the Markov control model of interest. In Section 3 we introduce our hypotheses and state our main result, Theorem 3.1, which is proved in Section 4. Finally, a linear system/quadratic cost (or LQ system) in Section 5 illustrates our results.

2. The control model

Let $(\mathbb{X}, \mathbb{A}, \{A(x) : x \in \mathbb{X}\}, Q, C)$ be a discrete-time Markov control model with state space \mathbb{X} and control (or action) set \mathbb{A} , both assumed to be Borel spaces with σ -algebras $\mathcal{B}(\mathbb{X})$ and $\mathcal{B}(\mathbb{A})$, respectively. For each $x \in \mathbb{X}$, there is a nonempty Borel set $A(x)$ in $\mathcal{B}(\mathbb{A})$ which represents the set of feasible actions in the state x . The set

$$\mathbb{K} := \{(x, a) : x \in \mathbb{X}, a \in A(x)\}$$

is assumed to be a Borel subset of $\mathbb{X} \times \mathbb{A}$. The transition law Q is a stochastic kernel on \mathbb{X} given \mathbb{K} and the one-stage cost C is a real-valued measurable function on \mathbb{K} .

Let Φ be the set of all stochastic kernels φ on A given \mathbb{X} for which $\varphi(A(x) \mid x) = 1$ for all $x \in \mathbb{X}$.

2.1. Control policies

For every $n = 0, 1, \dots$, let \mathbb{H}_n be the family of admissible histories up to time n ; that is, $\mathbb{H}_0 := \mathbb{X}$ and $\mathbb{H}_n := \mathbb{K}^n \times \mathbb{X}$ if $n \geq 1$. A *control policy* is a sequence $\pi = \{\pi_n\}$ of stochastic kernels π_n on \mathbb{A} given \mathbb{H}_n such that $\pi_n(A(x_n) \mid h_n) = 1$ for every n history $h_n = (x_0, a_0, \dots, x_{n-1}, a_{n-1}, x_n)$ in \mathbb{H}_n . The class of all policies is denoted by Π .

A policy $\pi = \{\pi_n\}$ is said to be a (randomized) *stationary policy* if there exists a stochastic kernel $\varphi \in \Phi$ such that $\pi_n(\cdot \mid h_n) = \varphi(\cdot \mid x_n)$ for all $h_n \in H_n, n = 0, 1, \dots$. As usual, we identify Φ , the set of stochastic kernels on A given \mathbb{X} , with the set of all randomized stationary policies. Hence, $\Phi \subset \Pi$.

If $\pi = \{\varphi\}$ is a stationary policy, abusing the notation, we write $\pi = \varphi$.

For notational ease, we write

$$C_\varphi(x) := \int_A C(x, a)\varphi(da \mid x) \quad \text{and} \quad Q_\varphi(\cdot \mid x) := \int_A Q(\cdot \mid x, a)\varphi(da \mid x) \tag{2.1}$$

for all $x \in \mathbb{X}$ and every stationary policy φ in Φ .

Let (Ω, \mathcal{F}) be the (canonical) measurable space consisting of the sample space

$$\Omega := (\mathbb{X} \times \mathbb{A})^\infty$$

and its product σ -algebra \mathcal{F} . Then, for each policy π and ‘initial state’ $x \in \mathbb{X}$, a stochastic process $\{(x_n, a_n)\}$ and a probability measure P_x^π are defined on (Ω, \mathcal{F}) in a canonical way, where x_n and a_n represent the state and control at time n , $n = 0, 1, \dots$. The expectation operator with respect to P_x^π is denoted by E_x^π .

Let $W : \mathbb{X} \rightarrow [\theta, \infty)$ be a measurable function that will be referred to as a *weight function*, where $\theta > 0$. We denote by $B_W(\mathbb{X})$ the normed linear space of measurable functions u on \mathbb{X} with finite W -norm $\|u\|_W$, which is defined as

$$\|u\|_W := \frac{\sup_{x \in \mathbb{X}} |u(x)|}{W(x)}. \tag{2.2}$$

In this case we say that u is W bounded. Note that $B_W(\mathbb{X})$ is a Banach space.

Let $\mu(\cdot)$ be a measure on \mathbb{X} . We write

$$\mu(u) := \int_{\mathbb{X}} u(y)\mu(dy)$$

whenever the integral is well defined.

Definition 2.1. Let

$$S_n(\varphi, x) := \sum_{k=0}^{n-1} C_\varphi(x_k) \quad \text{and} \quad J_n(\varphi, x) := E_x^\varphi[S_n(\varphi, x)]$$

be the *total pathwise n -stage cost* and the *total expected n -stage cost*, respectively, when using the stationary policy φ , given the initial state $x_0 = x$.

3. Assumptions and main result

In this section we introduce conditions to obtain asymptotic normality.

We will first introduce two sets of hypotheses. The first one, Assumption 3.1 below, uses a weight function W to impose a growth condition on the cost function. The second one, Assumption 3.2, imposes in particular a Lyapunov-like condition. These assumptions will ensure the existence of solutions to the Poisson equation (PE) in (3.4) below (see Lemma 3.1).

Assumption 3.2 has previously been used for MCPs on Borel spaces (see, for instance, [1], [2], [7], and [10]), but it was always combined with additional conditions that imply W -geometric ergodicity. The approach in this paper is quite different, because we do not require W -geometric ergodicity.

Assumption 3.1. *There exists a constant $K > 0$ and a measurable weight function $W(\cdot)$ on \mathbb{X} such that*

- (a) W is bounded below by a constant $\theta > 0$;
- (b) $|C(x, a)| \leq KW(x)$ for all $(x, a) \in \mathbb{K}$.

We now state our second main assumption.

Assumption 3.2. *There exist a nontrivial finite measure $\nu(\cdot)$ on \mathbb{X} , a nonnegative measurable function $l(\cdot, \cdot)$ on \mathbb{K} , and a positive constant $\lambda < 1$ such that*

- (a) $\nu(W) < \infty$;
- (b) $Q(\cdot | x, a) \geq l(x, a)\nu(\cdot)$ for each $(x, a) \in \mathbb{K}$;
- (c) $\int_{\mathbb{X}} W(y)Q(dy | x, a) \leq \lambda W(x) + l(x, a)\nu(W)$ for each $(x, a) \in \mathbb{K}$.

Remark 3.1. Assumption 3.1(a) and iterations of the inequality in Assumption 3.2(c) yield, for every $x \in \mathbb{X}$, $\pi \in \Pi$, and $n = 0, 1, \dots$,

$$\theta \leq E_x^\pi W(x_n) \leq \lambda^n W(x) + \frac{\nu(W)}{(1 - \lambda)\nu(\mathbb{X})}. \tag{3.1}$$

This fact and (2.2) show that, for every $u \in B_W(\mathbb{X})$ and $p > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n^p} E_x^\pi |u(x_n)| = 0. \tag{3.2}$$

In addition to Assumptions 3.1 and 3.2, we next impose other conditions on the control model. Several versions of these conditions have appeared in the literature (see, for instance, [3], [6], [7], [10], [21], and [22]), but the main ideas go back to [1] and [2].

Assumption 3.3. *For each $x \in \mathbb{X}$,*

- (a) $A(x)$ is a (nonempty) compact subset of \mathbb{A} ;
- (b) $C(x, \cdot)$ is upper semicontinuous on $A(x)$;
- (c) $Q(\cdot | x, \cdot)$ is strongly continuous on $A(x)$, that is, the mapping

$$a \rightarrow \int_{\mathbb{X}} u(y)Q(dy | x, a)$$

is continuous on $A(x)$ for each bounded measurable function u on \mathbb{X} ;

- (d) *the mapping $a \rightarrow \int_{\mathbb{X}} W(y)Q(dy | x, a)$ is continuous on $A(x)$, with W as in Assumption 3.1;*
- (e) *$l(x, \cdot)$ is continuous on $A(x)$, with $l(\cdot, \cdot)$ as in Assumption 3.2.*

Suppose that Assumptions 3.1, 3.2, and 3.3 hold, and, in addition, that we replace Assumption 3.1(b) with a second-order condition, say $C^2(x, a) \leq KW(x)$. Then we can ensure the existence of control policies that maximize the long-run pathwise average cost in Section 1. (See, for instance, [6], [17], [20], and [23].) To obtain asymptotic normality, however, we need to strengthen the growth condition on the cost with the following *fourth-order* condition.

Assumption 3.4. *There exists a positive constant r_2 such that*

$$C^4(x, a) \leq r_2W(x) \quad \text{for all } (x, a) \in \mathbb{K}. \tag{3.3}$$

We now state some facts from [17, Theorem 2.1.4 and Lemma 4.2.2] which we need to present our main result.

Lemma 3.1. (a) *Under Assumptions 3.1 and 3.2, the following facts hold for each $\varphi \in \Phi$.*

(i) *The Markov chain defined by $Q_\varphi(\cdot | \cdot)$ is ν -irreducible and positive Harris recurrent; hence, it admits a unique invariant probability measure, say μ_φ .*

(ii) $\mu_\varphi(W) < \infty$; thus, we have

$$\rho_\varphi := \mu_\varphi(C_\varphi) < \infty.$$

(iii) *There exists a function h_φ^* in $B_W(\mathbb{X})$ such that the pair $(\rho_\varphi, h_\varphi^*)$ satisfies the PE*

$$h_\varphi^*(x) = C_\varphi(x) - \rho_\varphi + \int_{\mathbb{X}} h_\varphi^*(y) Q_\varphi(dy | x) \quad \text{for all } x \in \mathbb{X}. \tag{3.4}$$

(b) *Suppose that Assumptions 3.1, 3.2, 3.3, and 3.4 hold. Let $\varphi \in \Phi$ be arbitrary, and let h_φ^* be as in (a)(iii). We define the function*

$$\Psi_\varphi(x) := \int_{\mathbb{X}} h_\varphi^{*2}(y) Q_\varphi(dy | x) - \left(\int_{\mathbb{X}} h_\varphi^*(y) Q_\varphi(dy | x) \right)^2 \tag{3.5}$$

for each $x \in \mathbb{X}$. Then the functions h_φ^{*2} and Ψ_φ belong to $B_W(\mathbb{X})$, and so $\sigma_\varphi^2 := \mu_\varphi(\Psi_\varphi)$ is finite.

Proof. Part (a) is an adaptation to our present context of Theorem 2.1.4 of [17]. On the other hand, under Assumptions 3.1(a) and 3.4, we have in particular that the cost function C is $W^{1/2}$ bounded. It follows that Assumptions 3.1 and 3.2 are satisfied when we replace W with $W^{1/2}$. Therefore, part (b) follows from (a).

We can now state our main result, which is proved in Section 4.

Theorem 3.1. *Suppose that Assumptions 3.1, 3.2, 3.3, and 3.4 hold. Let $\varphi \in \Phi$ be a randomized stationary policy. Then, for every initial state $x \in \mathbb{X}$,*

$$\frac{S_n(\varphi, x) - n\rho_\varphi}{\sqrt{n}}$$

has asymptotically a normal distribution $N(0, \sigma_\varphi^2)$ as $n \rightarrow \infty$, with $S_n(\varphi, x)$ as in Definition 2.1, and ρ_φ and σ_φ^2 as defined in Lemma 3.1(a) and (b), respectively.

Remark 3.2. Taking the action space \mathbb{A} as a single-point set (or singleton), Theorem 3.1 becomes a central limit theorem for (noncontrolled) Markov chains.

4. Proof of Theorem 3.1

In this section we suppose that Assumptions 3.1, 3.2, 3.3, and 3.4 hold. We need the following definition.

Definition 4.1. Let $J_n(\varphi, x)$ and $S_n(\varphi, x)$ be as in Definition 2.1. For every $\varphi \in \Phi$ and initial state x , we define the *limiting average variance*

$$V(\varphi, x) := \limsup_{n \rightarrow \infty} \frac{1}{n} \text{var}[S_n(\varphi, x)],$$

where (by the definition of the variance of a random variable)

$$\text{var}[S_n(\varphi, x)] = E_x^\varphi [S_n(\varphi, x) - J_n(\varphi, x)]^2.$$

The following lemma is crucial to prove Theorem 3.1.

Lemma 4.1. *Suppose that Assumptions 3.1, 3.2, 3.3, and 3.4 hold. Let $\varphi \in \Phi$ be arbitrary, and let h_φ^* and Ψ_φ be as in Lemma 3.1. Then the limiting average variance satisfies*

$$V(\varphi, x) = \lim_{n \rightarrow \infty} \frac{1}{n} E_x^\varphi \sum_{k=0}^{n-1} \Psi_\varphi(x_k) = \sigma_\varphi^2 \quad \text{for all } x \in \mathbb{X}, \tag{4.1}$$

with $\sigma_\varphi^2 = \mu_\varphi(\Psi_\varphi)$ as in Lemma 3.1(b).

Proof. The following proof is taken from [4, Theorem 11.2.4] adapted to our present context. From (3.4), an iteration procedure yields

$$J_n(\varphi, x) = E_x^\varphi \left[\sum_{k=0}^{n-1} C_\varphi(x_k) \right] = n\rho_\varphi + h_\varphi^*(x) - E_x^\varphi h_\varphi^*(x_n).$$

Hence,

$$S_n(\varphi, x) - J_n(\varphi, x) = [S_n(\varphi, x) - n\rho_\varphi] - [h_\varphi^*(x) - E_x^\varphi h_\varphi^*(x_n)]$$

and

$$\text{var}[S_n(\varphi, x)] = E_x^\varphi [S_n(\varphi, x) - n\rho_\varphi]^2 - E_x^\varphi [h_\varphi^*(x) - E_x^\varphi h_\varphi^*(x_n)]^2.$$

Thus,

$$\text{var}[S_n(\varphi, x)] = E_x^\varphi [S_n(\varphi, x) - n\rho_\varphi]^2 + o_1(x, n) \tag{4.2}$$

with

$$o_1(x, n) := -E_x^\varphi [h_\varphi^*(x) - E_x^\varphi h_\varphi^*(x_n)]^2. \tag{4.3}$$

We claim that

$$\frac{o_1(x, n)}{n} \rightarrow 0 \quad \text{for all } x \in \mathbb{X} \tag{4.4}$$

as $n \rightarrow \infty$. Indeed,

$$|o_1(x, n)| \leq 2[h_\varphi^{*2}(x) + E_x^\varphi h_\varphi^{*2}(x_n)].$$

Hence, since $h_\varphi^{*2} \in B_W(\mathbb{X})$ and by (3.2) in Remark 3.1, limit (4.4) holds.

On the other hand, define

$$Y_k := h_\varphi^*(x_k) - E_x^\varphi [h_\varphi^*(x_k) \mid x_{k-1}], \quad k = 1, 2, \dots \tag{4.5}$$

From (4.5) we obtain

$$Y_k^2 = h_\varphi^{*2}(x_k) - 2h_\varphi^*(x_k) E_x^\varphi [h_\varphi^*(x_k) \mid x_{k-1}] + (E_x^\varphi [h_\varphi^*(x_k) \mid x_{k-1}])^2.$$

Therefore,

$$\begin{aligned} E_x^\varphi (Y_k^2) &= E_x^\varphi h_\varphi^{*2}(x_k) - E_x^\varphi (E_x^\varphi [h_\varphi^*(x_k) \mid x_{k-1}])^2 \\ &= E_x^\varphi E_x^\varphi [h_\varphi^{*2}(x_k) \mid x_{k-1}] - E_x^\varphi (E_x^\varphi [h_\varphi^*(x_k) \mid x_{k-1}])^2 \\ &= E_x^\varphi \left[\int_{\mathbb{X}} h_\varphi^{*2}(y) Q_\varphi(dy \mid x_{k-1}) - \left(\int_{\mathbb{X}} h_\varphi^*(y) Q_\varphi(dy \mid x_{k-1}) \right)^2 \right] \\ &= E_x^\varphi [\Psi_\varphi(x_{k-1})], \end{aligned}$$

that is,

$$E_x^\varphi(Y_k^2) = E_x^\varphi[\Psi_\varphi(x_{k-1})]. \tag{4.6}$$

Moreover, we have

$$\text{var}[S_n(\varphi, x)] = E_x^\varphi\left[\sum_{k=1}^{n-1} Y_k^2\right] + o_2(x, n) \tag{4.7}$$

with

$$\frac{o_2(x, n)}{n} \rightarrow 0 \quad \text{when } n \rightarrow \infty. \tag{4.8}$$

To prove (4.7)–(4.8), note that the random variables Y_t satisfy

$$E_x^\varphi[Y_t Y_s] = 0 \quad \text{for all } 1 \leq t < s. \tag{4.9}$$

We define the σ -algebras

$$\mathcal{F}_t := \sigma\{x_0, \dots, x_t\} \quad \text{for } t = 0, 1, \dots$$

Then

$$E_x^\varphi[Y_s \mid \mathcal{F}_{s-1}] = E_x^\varphi[h_\varphi^*(x_s) \mid x_{s-1}] - E_x^\varphi[h_\varphi^*(x_s) \mid x_{s-1}] = 0 \quad \text{for all } s,$$

which implies that

$$E_x^\varphi[Y_t Y_s] = E_x^\varphi[E_x^\varphi[Y_t Y_s \mid \mathcal{F}_{s-1}]] = E_x^\varphi[Y_t E_x^\varphi[Y_s \mid \mathcal{F}_{s-1}]] = 0.$$

On the other hand, since h_φ^* satisfies the Poisson equation (3.4), we also have

$$C_\varphi(x) - \rho_\varphi = h_\varphi^*(x) - \int_{\mathbb{X}} h_\varphi^*(y) Q_\varphi(dy \mid x) \quad \text{for all } x \in \mathbb{X}.$$

Consequently, by (4.5),

$$\begin{aligned} C_\varphi(x_k) - \rho_\varphi &= h_\varphi^*(x_k) - \int_{\mathbb{X}} h_\varphi^*(y) Q_\varphi(dy \mid x_k) \\ &= Y_k + \int_{\mathbb{X}} h_\varphi^*(y) Q_\varphi(dy \mid x_{k-1}) - \int_{\mathbb{X}} h_\varphi^*(y) Q_\varphi(dy \mid x_k) \\ &= Y_k - [Ph_\varphi^*(x_k) - Ph_\varphi^*(x_{k-1})] \quad \text{for all } k = 1, 2, \dots, \end{aligned} \tag{4.10}$$

with

$$Pu(x) := \int_{\mathbb{X}} u(y) Q_\varphi(dy \mid x) \quad \text{for all } x \in \mathbb{X} \text{ and } u \in B_W(\mathbb{X}).$$

From (4.10) we obtain

$$S_n(\varphi, x) - n\rho_\varphi = M_{n-1} + [h_\varphi^*(x) - Ph_\varphi^*(x_{n-1})] \quad \text{for all } n \geq 2 \tag{4.11}$$

with

$$M_n := \sum_{k=1}^n Y_k \quad \text{for all } n \geq 1.$$

Observe that (4.9) implies that

$$E_x^\varphi[M_n^2] = E_x^\varphi\left[\sum_{k=1}^n Y_k^2\right]. \tag{4.12}$$

By (4.11),

$$\begin{aligned} E_x^\varphi[S_n(\varphi, x) - n\rho_\varphi]^2 &= E_x^\varphi[M_{n-1}^2] + 2E_x^\varphi[(h_\varphi^*(x) - Ph_\varphi^*(x_{n-1}))M_{n-1}] \\ &\quad + E_x^\varphi[h_\varphi^*(x) - Ph_\varphi^*(x_{n-1})]^2. \end{aligned}$$

Substituting this relation into (4.2), together with (4.12), we obtain (4.7) as we wanted to prove, with

$$o_2(x, n) := o_1(x, n) + A(x, n) + 2B(x, n),$$

where $o_1(x, n)$ is as in (4.3), and

$$\begin{aligned} A(x, n) &:= E_x^\varphi[h_\varphi^*(x) - Ph_\varphi^*(x_{n-1})]^2, \\ B(x, n) &:= E_x^\varphi[(h_\varphi^*(x) - Ph_\varphi^*(x_{n-1}))M_{n-1}]. \end{aligned}$$

To complete the proof of (4.7)–(4.8), we only need to show that, for every initial state x , as $n \rightarrow \infty$,

$$\frac{A(x, n)}{n} \rightarrow 0 \tag{4.13}$$

and

$$\frac{B(x, n)}{n} \rightarrow 0. \tag{4.14}$$

To prove (4.13), it suffices to note that

$$A(x, n) \leq 2[h_\varphi^{*2}(x) + E_x^\varphi h_\varphi^{*2}(x_n)].$$

This yields (4.13).

Now, to obtain (4.14), we use the Cauchy–Schwartz inequality to see that

$$B(x, n)^2 \leq E_x^\varphi[h_\varphi^*(x) - Ph_\varphi^*(x_{n-1})]^2 E_x^\varphi[M_n^2] = A(x, n) E_x^\varphi\left[\sum_{k=1}^n Y_k^2\right].$$

Hence, by (4.6),

$$\frac{B(x, n)}{n} \leq \sqrt{\frac{A(x, n)}{n}} \sqrt{\frac{1}{n} \sum_{k=0}^{n-1} E_x^\varphi \Psi_\varphi(x_k)}.$$

This inequality implies the limit in (4.14). This completes the proof of (4.7)–(4.8).

Combining (4.6), (4.7), and (4.8), we obtain

$$\begin{aligned} \text{var}[S_n(\varphi, x)] &= E_x^\varphi\left[\sum_{k=1}^{n-1} Y_k^2\right] + o_2(x, n) \\ &= E_x^\varphi\left[\sum_{k=1}^n Y_k^2\right] + o_2(x, n) - E_x^\varphi Y_n^2 \\ &= \sum_{k=0}^{n-1} E_x^\varphi[\Psi_\varphi(x_k)] + o_2(x, n) - E_x^\varphi \Psi_\varphi(x_{n-1}). \end{aligned}$$

Hence,

$$V(\varphi, x) = \limsup_{n \rightarrow \infty} \frac{1}{n} \text{var}[S_n(\varphi, x)] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} E_x^\varphi[\Psi_\varphi(x_k)] = \sigma_\varphi^2,$$

which proves (4.1).

Lemma 4.2. *Under Assumptions 3.1, 3.2, 3.3, and 3.4, the following statements, for each $\varphi \in \Phi$, hold.*

- (a) *The function h_φ^* in Lemma 3.1(a) is $W^{1/4}$ bounded.*
- (b) *The function Ψ_φ in Lemma 3.1(b) is $W^{1/2}$ bounded. Moreover, there exists a $W^{1/2}$ -bounded function \widehat{h}_φ^* such that*

$$\widehat{h}_\varphi^*(x) = \Psi_\varphi(x) - \sigma_\varphi^2 + \int_{\mathbb{X}} \widehat{h}_\varphi^*(y) Q_\varphi(dy | x) \quad \text{for all } x \in \mathbb{X}, \tag{4.15}$$

with σ_φ^2 as in (4.1).

Proof. (a) This part follows from Lemma 3.1 if we prove that Assumptions 3.1 and 3.2 are satisfied when we replace W by $W^{1/4}$.

Indeed, under Assumption 3.4, Assumptions 3.1 and 3.2(a)–(b) obviously hold when we replace W by $W^{1/4}$. It only remains to prove that Assumption 3.2(c) holds. To this end, define

$$\widehat{Q}(B | x, a) := Q(B | x, a) - v(B)l(x, a) \tag{4.16}$$

for each $B \in \mathcal{B}(\mathbb{X})$ and $(x, a) \in \mathbb{K}$. Under Assumption 3.2(b), \widehat{Q} is a nonnegative kernel on \mathbb{X} given \mathbb{K} . By the Cauchy–Schwartz inequality,

$$\int_{\mathbb{X}} \sqrt{W(y)} \widehat{Q}(dy | x, a) \leq \sqrt{\int_{\mathbb{X}} W(y) \widehat{Q}(dy | x, a)} \sqrt{\widehat{Q}(\mathbb{X} | x, a)} \leq \sqrt{\lambda} \sqrt{W(x)} \tag{4.17}$$

because $0 \leq \widehat{Q}(\mathbb{X} | x, a) \leq 1$. Hence, by the Cauchy–Schwartz inequality again, and from (4.17), we obtain

$$\int_{\mathbb{X}} W(y)^{1/4} \widehat{Q}(dy | x, a) \leq \sqrt{\int_{\mathbb{X}} \sqrt{W(y)} \widehat{Q}(dy | x, a)} \sqrt{\widehat{Q}(\mathbb{X} | x, a)} \leq \lambda^{1/4} W(x)^{1/4}.$$

From the definition of the kernel \widehat{Q} in (4.16), we obtain

$$\int_{\mathbb{X}} W(y)^{1/4} Q(dy | x, a) \leq \lambda^{1/4} W(x)^{1/4} + l(x, a)v(W^{1/4}) \quad \text{for all } (x, a) \in \mathbb{K}.$$

This implies that Assumption 3.2(c) holds when we replace W by $W^{1/4}$.

(b) Consider the cost function $C(x, a) := \Psi_\varphi(x)$ for all $(x, a) \in \mathbb{K}$. From part (a), Ψ_φ is $W^{1/2}$ bounded. Applying Lemma 3.1 to this particular cost function we obtain the desired result.

We are finally ready to prove Theorem 3.1.

Proof of Theorem 3.1. From Lemmas 3.1(a)(iii) and 4.2(i), there exists a $W^{1/4}$ -bounded function h_φ^* satisfying (3.4). Similarly, from Lemma 4.2(ii), there exists a $W^{1/2}$ -bounded function \widehat{h}_φ^* satisfying (4.15). Let $h_1 := h_\varphi^*$ and $h_2 = \widehat{h}_\varphi^*$. Thus, h_1 is $W^{1/4}$ bounded and h_2 is $W^{1/2}$ bounded.

We also define

$$\tau_1(x) := \int_{\mathbb{X}} h_1(y) Q_\varphi(dy | x) - h_1(x) + C_\varphi(x) - \rho_\varphi$$

and

$$\tau_2(x) := \int_{\mathbb{X}} h_2(y) Q_\varphi(dy | x) - h_2(x) + \Psi_\varphi(x) - \sigma_\varphi^2$$

for all $x \in \mathbb{X}$.

For $l = 1, 2$ and $x \in \mathbb{X}$, let

$$\psi_l(x) := \int_{\mathbb{X}} h_l(y) Q_\varphi(dy | x) - h_l(x).$$

Observe that

$$\tau_1(x) = \psi_1(x) + C_\varphi(x) - \rho_\varphi \tag{4.18}$$

and

$$\tau_2(x) = \psi_2(x) + \Psi_\varphi(x) - \sigma_\varphi^2 \tag{4.19}$$

for all $(x, a) \in \mathbb{K}$.

Consider the characteristic functions $E_x^\varphi \chi_n(u)$ with

$$\chi_n(u) := \exp\{iu(S_n(\varphi, x) - n\rho_\varphi)\} \quad \text{for } n = 1, 2, \dots, u \in \mathbb{R},$$

and $\chi_0(u) := 1$. Let

$$e_1(z) := \exp\{iz\} - iz - 1, \tag{4.20}$$

$$e_2(z) := \exp\{iz\} + \frac{z^2}{2} - iz - 1, \tag{4.21}$$

for all z in \mathbb{R} .

To prove the theorem, we have to verify that

$$\lim_{n \rightarrow \infty} E_x^\varphi \chi_n\left(\frac{u}{\sqrt{n}}\right) = \exp\left\{-\frac{1}{2}\sigma_\varphi^2 u^2\right\} \quad \text{for all } u \in \mathbb{R}. \tag{4.22}$$

To this end, fix an arbitrary $u \in \mathbb{R}$. First note that $\psi_l(x_m)$ for $l = 1, 2$ is the conditional expectation of $h_l(x_{m+1}) - h_l(x_m)$ given x_m , that is,

$$\psi_l(x_m) = E_x^\varphi[h_l(x_{m+1}) - h_l(x_m) | x_m].$$

This yields, for $l = 1, 2$, with $\chi_m := \chi_m(u)$, the equations

$$0 = iu E_x^\varphi \left[\sum_{m=0}^{n-1} \chi_m \psi_1(x_m) - \sum_{m=0}^{n-1} \chi_m (h_1(x_{m+1}) - h_1(x_m)) \right] \tag{4.23}$$

and

$$0 = \frac{u^2}{2} E_x^\varphi \left[\sum_{m=0}^{n-1} \chi_m (h_2(x_{m+1}) - h_2(x_m)) - \sum_{m=0}^{n-1} \chi_m \psi_2(x_m) \right]. \tag{4.24}$$

To simplify the notation, let $e_1 := e_1(u(C_\varphi(x_m) - \rho_\varphi))$ and $e_2 := e_2(u(C_\varphi(x_m) - \rho_\varphi))$. Moreover, note that

$$\chi_{m+1} - \chi_m = [\exp\{iu(C_\varphi(x_m) - \rho_\varphi)\} - 1]\chi_m. \tag{4.25}$$

From (4.20), (4.21), and (4.25), we have

$$\begin{aligned} E_x^\varphi \chi_n - 1 &= E_x^\varphi \sum_{m=0}^{n-1} (\chi_{m+1} - \chi_m) \\ &= E_x^\varphi \sum_{m=0}^{n-1} \left[iu(C_\varphi(x_m) - \rho_\varphi) - \frac{1}{2}u^2(C_\varphi(x_m) - \rho_\varphi)^2 + e_2 \right] \chi_m \end{aligned} \tag{4.26}$$

and

$$\begin{aligned} &-iu E_x^\varphi \sum_{m=0}^{n-1} \chi_m (h_1(x_{m+1}) - h_1(x_m)) \\ &= iu E_x^\varphi \left[h_1(x_0) - \chi_n h_1(x_n) + \sum_{m=0}^{n-1} h_1(x_{m+1})(\chi_{m+1} - \chi_m) \right] \\ &= iu E_x^\varphi \left[h_1(x_0) - \chi_n h_1(x_n) + \sum_{m=0}^{n-1} h_1(x_{m+1})(iu(C_\varphi(x_m) - \rho_\varphi) + e_1)\chi_m \right]. \end{aligned} \tag{4.27}$$

Similarly,

$$\begin{aligned} &\frac{u^2}{2} E_x^\varphi \sum_{m=0}^{n-1} \chi_m (h_2(x_{m+1}) - h_2(x_m)) \\ &= -\frac{u^2}{2} E_x^\varphi \left[h_2(x_0) - \chi_n h_2(x_n) + \sum_{m=0}^{n-1} h_2(x_{m+1})(\chi_{m+1} - \chi_m) \right] \\ &= -\frac{u^2}{2} E_x^\varphi \left[h_2(x_0) - \chi_n h_2(x_n) + \sum_{m=0}^{n-1} h_2(x_{m+1})(\exp\{iu(C_\varphi(x_m) - \rho_\varphi)\} - 1)\chi_m \right]. \end{aligned} \tag{4.28}$$

Adding (4.23), (4.24), (4.26)–(4.28), and using (4.18),

$$\begin{aligned} &E_x^\varphi \chi_n - 1 \\ &= iu E_x^\varphi \left[h_1(x_0) - \chi_n h_1(x_n) + \sum_{m=0}^{n-1} \chi_m \tau_1(x_m) + \sum_{m=0}^{n-1} e_1 h_1(x_{m+1}) \chi_m \right] \\ &\quad - \frac{u^2}{2} E_x^\varphi \sum_{m=0}^{n-1} \chi_m \{ \psi_2(x_m) + 2h_1(x_{m+1})(C_\varphi(x_m) - \rho_\varphi) + (C_\varphi(x_m) - \rho_\varphi)^2 \} \\ &\quad - \frac{u^2}{2} E_x^\varphi \left[h_2(x_0) - \chi_n h_2(x_n) + \sum_{m=0}^{n-1} h_2(x_{m+1})(\exp\{iu(C_\varphi(x_m) - \rho_\varphi)\} - 1)\chi_m \right] \\ &\quad + E_x^\varphi \sum_{m=0}^{n-1} e_2 \chi_m. \end{aligned}$$

Hence,

$$\begin{aligned} E_x^\varphi \chi_n - 1 &= \kappa''(n, u) \\ &\quad - \frac{u^2}{2} E_x^\varphi \sum_{m=0}^{n-1} \chi_m \{ \psi_2(x_m) + 2h_1(x_{m+1})(C_\varphi(x_m) - \rho_\varphi) + (C_\varphi(x_m) - \rho_\varphi)^2 \} \end{aligned} \tag{4.29}$$

with

$$\begin{aligned} \kappa''(n, u) &:= iu E_x^\varphi \left[h_1(x_0) - \chi_n h_1(x_n) + \sum_{m=0}^{n-1} \chi_m \tau_1(x_m) + \sum_{m=0}^{n-1} e_1 h_1(x_{m+1}) \chi_m \right] \\ &\quad - \frac{u^2}{2} E_x^\varphi \left[h_2(x_0) - \chi_n h_2(x_n) + \sum_{m=0}^{n-1} h_2(x_{m+1}) (\exp\{iu(C_\varphi(x_m) - \rho_\varphi)\} - 1) \chi_m \right] \\ &\quad + E_x^\varphi \sum_{m=0}^{n-1} e_2 \chi_m. \end{aligned} \tag{4.30}$$

From (3.5) we observe that

$$\Psi_\varphi(x_m) = E_x^\varphi [h_1^2(x_{m+1}) \mid x_m] - (E_x^\varphi [h_1(x_{m+1}) \mid x_m])^2,$$

and in view of (4.19), we can express (4.29) as

$$\begin{aligned} E_x^\varphi \chi_n - 1 &= \kappa''(n, u) \\ &\quad - \frac{u^2}{2} E_x^\varphi \sum_{m=0}^{n-1} \chi_m \{ \sigma_\varphi^2 + \tau_2(x_m) - h_1^2(x_{m+1}) (E_x^\varphi [h_1(x_{m+1}) \mid x_m] + C_\varphi(x_m) - \rho_\varphi)^2 \} \\ &= \kappa''(n, u) \\ &\quad - \frac{u^2}{2} E_x^\varphi \sum_{m=0}^{n-1} \chi_m \left\{ \sigma_\varphi^2 + \tau_2(x_m) - h_1^2(x_{m+1}) \right. \\ &\quad \quad \left. + \left(\int_{\mathbb{X}} h_1(y) Q_\varphi(dy \mid x_m) + C_\varphi(x_m) - \rho_\varphi \right)^2 \right\}. \end{aligned}$$

Since $h_1 = h_\varphi^*$ satisfies (3.4), we obtain

$$h_1(x_m) = \int_{\mathbb{X}} h_1(y) Q_\varphi(dy \mid x_m) + C_\varphi(x_m) - \rho_\varphi.$$

Then, from (4.25), we have

$$\begin{aligned} E_x^\varphi \chi_n - 1 &= \kappa''(n, u) - \frac{u^2}{2} E_x^\varphi \sum_{m=0}^{n-1} \chi_m \{ \sigma_\varphi^2 + \tau_2(x_m) - h_1^2(x_{m+1}) + h_1^2(x_m) \} \\ &= \kappa''(n, u) - \frac{u^2 \sigma_\varphi^2}{2} \sum_{m=0}^{n-1} E_x^\varphi \chi_m \\ &\quad - \frac{u^2}{2} E_x^\varphi \left[h_1^2(x_0) - \chi_n h_1^2(x_n) + \sum_{m=0}^{n-1} \chi_m \tau_2(x_m) + \sum_{m=0}^{n-1} h_1^2(x_{m+1}) (\chi_{m+1} - \chi_m) \right] \end{aligned}$$

$$\begin{aligned}
 &= \kappa''(n, u) - \frac{u^2 \sigma_\varphi^2}{2} \sum_{m=0}^{n-1} E_x^\varphi \chi_m \\
 &\quad - \frac{u^2}{2} E_x^\varphi \left[h_1^2(x_0) - \chi_n h_1^2(x_n) + \sum_{m=0}^{n-1} \chi_m \tau_2(x_m) \right. \\
 &\quad \left. + \sum_{m=0}^{n-1} h_1^2(x_{m+1}) (\exp\{iu(C_\varphi(x_m) - \rho_\varphi)\} - 1) \chi_m \right].
 \end{aligned}$$

Hence,

$$E_x^\varphi \chi_n = 1 - \frac{u^2 \sigma_\varphi^2}{2} \sum_{m=0}^{n-1} E_x^\varphi \chi_m + \kappa'(n, u) \tag{4.31}$$

with

$$\begin{aligned}
 \kappa'(n, u) := &\kappa''(n, u) - \frac{u^2}{2} E_x^\varphi \left[h_1^2(x_0) - \chi_n h_1^2(x_n) + \sum_{m=0}^{n-1} \chi_m \tau_2(x_m) \right. \\
 &\left. + \sum_{m=0}^{n-1} h_1^2(x_{m+1}) (\exp\{iu(C_\varphi(x_m) - \rho_\varphi)\} - 1) \chi_m \right]. \tag{4.32}
 \end{aligned}$$

Let us rewrite (4.31) as

$$E_x^\varphi \chi_n = 1 + \left(\exp\left\{-\frac{u^2 \sigma_\varphi^2}{2}\right\} - 1 \right) \sum_{m=0}^{n-1} E_x^\varphi \chi_m + \kappa(n, u) \tag{4.33}$$

with

$$\kappa(n, u) := \kappa'(n, u) + \left[1 - \frac{u^2 \sigma_\varphi^2}{2} - \exp\left\{-\frac{u^2 \sigma_\varphi^2}{2}\right\} \right] \sum_{m=0}^{n-1} E_x^\varphi \chi_m. \tag{4.34}$$

Defining $\kappa(0, u) := 0$, from (4.33), an induction argument gives

$$\begin{aligned}
 E_x^\varphi \chi_n(u) &= \exp\left\{-\frac{n \sigma_\varphi^2 u^2}{2}\right\} \\
 &\quad + \left[\exp\left\{-\frac{\sigma_\varphi^2 u^2}{2}\right\} - 1 \right] \sum_{m=0}^{n-1} \exp\left\{-\frac{\sigma_\varphi^2 u^2}{2}(n-1-m)\right\} \kappa(m, u) + \kappa(n, u).
 \end{aligned} \tag{4.35}$$

Observe that the proof of limit (4.22) and, consequently, of Theorem 3.1 follows from (4.35) if we show that

$$\max_{1 \leq m \leq n} \left| \kappa\left(m, \frac{u}{\sqrt{n}}\right) \right| \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{4.36}$$

This relation is obtained by an inspection of the different terms of $\kappa(m, u/\sqrt{n})$. We will do this in the following six steps.

Step 1. From the definition of $h_1 = h_\varphi^*$ and $h_2 = \widehat{h}_\varphi^*$ satisfying (3.4) and (4.15), we have $\tau_1(x_m) = 0$ in (4.30) and $\tau_2(x_m) = 0$ in (4.32) for $m = 0, 1, \dots$

Step 2. By (3.1) we obtain

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} E_x^\varphi h(x_n) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{1}{n} E_x^\varphi h(x_n) = 0$$

for every h in $B_W(\mathbb{X})$. This limit appears in (4.30) and (4.32) when we replace u by u/\sqrt{n} .

Step 3. In this part we prove the limit (see (4.30))

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} E_x^\varphi \sum_{m=0}^{n-1} e_1 h_1(x_{m+1}) \chi_m = 0.$$

Indeed, since $|e_1(z)| \leq z^2/2$ for all z in \mathbb{R} , we obtain

$$\begin{aligned} \left| \frac{1}{\sqrt{n}} E_x^\varphi \sum_{m=0}^{n-1} e_1 h_1(x_{m+1}) \chi_m \right| &\leq \frac{1}{2\sqrt{n}} E_x^\varphi \sum_{m=0}^{n-1} \frac{u^2}{n} |h_1(x_{m+1})| (C_\varphi(x_m) - \rho_\varphi)^2 \\ &= \frac{u^2}{2n^{3/2}} E_x^\varphi \sum_{m=0}^{n-1} \left| \int_{\mathbb{X}} h_1(y) Q_\varphi(dy \mid x_m) \right| (C_\varphi(x_m) - \rho_\varphi)^2. \end{aligned}$$

By Lemma 4.2(i), $h_1(\cdot)$ is $W^{1/4}$ bounded, in particular $h_1(\cdot)$ is $W^{1/2}$ bounded. Hence, the function $\int_{\mathbb{X}} h_1(y) Q_\varphi(dy \mid \cdot)$ is $W^{1/2}$ bounded. On the other hand, by Assumption 3.4, $(C_\varphi(x) - \rho_\varphi)^2$ is also $W^{1/2}$ bounded. Therefore,

$$\left| \frac{1}{\sqrt{n}} E_x^\varphi \sum_{m=0}^{n-1} e_1 h_1(x_{m+1}) \chi_m \right| \leq \frac{\zeta u^2}{2n^{3/2}} E_x^\varphi \sum_{m=0}^{n-1} W(x_m),$$

where ζ is a constant depending on h_1 and C . By Assumptions 3.1(a) and 3.2(c), and (3.1),

$$\left| \frac{1}{\sqrt{n}} E_x^\varphi \sum_{m=0}^{n-1} e_1 h_1(x_{m+1}) \chi_m \right| \leq \frac{\zeta u^2}{2n^{3/2}} n \left[1 + \frac{\nu(W)}{(1-\lambda)\theta \nu(\mathbb{X})} \right] W(x),$$

which converges to 0 as $n \rightarrow \infty$.

Step 4. We will next prove that

$$\lim_{n \rightarrow \infty} \frac{1}{n} E_x^\varphi \sum_{m=0}^{n-1} e_2 \chi_m = 0.$$

This limit appears in (4.30) when we replace u by u/\sqrt{n} .

Observe that $|e_2(z)| \leq |z|^3/6$ for all z in \mathbb{R} . So, by Assumptions 3.1(a), 3.2(c), and 3.4, together with (3.1),

$$\begin{aligned} \left| \frac{1}{n} E_x^\varphi \sum_{m=0}^{n-1} e_2 \chi_m \right| &\leq \frac{|u|^3}{6n^{5/2}} E_x^\varphi \sum_{m=0}^{n-1} |C_\varphi(x_m) - \rho_\varphi|^3 \\ &\leq \frac{k^3 |u|^3}{6n^{5/2}} E_x^\varphi \sum_{m=0}^{n-1} W(x_m)^{3/4} \end{aligned}$$

$$\begin{aligned} &\leq \frac{k^3|u|^3}{6n^{5/2}\theta^{1/4}} \mathbb{E}_x^\varphi \sum_{m=0}^{n-1} W(x_m) \\ &\leq \frac{k^3|u|^3}{6n^{3/2}} \left(1 + \frac{\nu(W)}{(1-\lambda)\theta\nu(\mathbb{X})} \right) W(x), \end{aligned}$$

which converges to 0 as $n \rightarrow \infty$, with k a constant.

Step 5. Let h be a $W^{1/2}$ -bounded function on \mathbb{X} . Then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}_x^\varphi \sum_{m=0}^{n-1} h(x_{m+1}) \left(\exp \left\{ i \frac{u}{\sqrt{n}} (C - \rho_\varphi) \right\} - 1 \right) \chi_m = 0.$$

This limit appears in (4.30) and (4.32) when u is replaced by u/\sqrt{n} .

From (4.20), $e_1(z) = \exp\{iz\} - iz - 1$ and so

$$\exp \left\{ i \frac{u}{\sqrt{n}} (C - \rho_\varphi) \right\} - 1 = i \frac{u}{\sqrt{n}} (C - \rho_\varphi) + e_1 \frac{u}{\sqrt{n}} (C - \rho_\varphi).$$

Hence,

$$\begin{aligned} &\left| \frac{1}{n} \mathbb{E}_x^\varphi \sum_{m=0}^{n-1} h(x_{m+1}) \left(\exp \left\{ i \frac{u}{\sqrt{n}} (C - \rho_\varphi) \right\} - 1 \right) \chi_m \right| \\ &\leq \frac{|u|}{n^{3/2}} \mathbb{E}_x^\varphi \sum_{m=0}^{n-1} |h(x_{m+1})| |C_\varphi(x_m) - \rho_\varphi| + \frac{1}{n} \mathbb{E}_x^\varphi \sum_{m=0}^{n-1} |h(x_{m+1})| |e_1|. \end{aligned}$$

This gives the desired conclusion by similar arguments to those in step 3.

Step 6. The absolute value of the expression within the brackets in (4.34) is majorized by $\sigma_\varphi^4 u^4/8$, and so the corresponding term in $\kappa(n, u/\sqrt{n})$ is majorized by $\sigma_\varphi^4 u^4/8n^2$.

Steps 1–6 imply (4.36), and consequently prove Theorem 3.1.

5. Example: an LQ system

In this section we consider a linear-quadratic (or LQ) system consisting of the linear system

$$x_{t+1} = k_1 x_t + k_2 a_t + z_t, \quad t = 0, 1, \dots,$$

with state space $\mathbb{X} := \mathbb{R}$, and positive coefficients k_1 and k_2 . The control set is $A := \mathbb{R}$, and the set of admissible controls in each state x is the interval

$$A(x) := \left[-\frac{k_1|x|}{k_2}, \frac{k_1|x|}{k_2} \right].$$

The disturbances z_t are independent and identically distributed (i.i.d.) random variables with values in $Z := \mathbb{R}$, zero mean, and finite variance, that is,

$$\mathbb{E}[z_t] = 0, \quad \sigma^2 := \mathbb{E}[z_t^2] < \infty. \tag{5.1}$$

To complete the description of our control model, we introduce the quadratic cost-per-stage function

$$C(x, a) := c_1 x^2 + c_2 a^2 \quad \text{for all } (x, a) \in \mathbb{K},$$

with positive coefficients c_1 and c_2 . We also define

$$W(x) := \exp\{\gamma|x|\} \quad \text{for all } x \in \mathbb{X},$$

with $\gamma \geq 4$. Clearly, Assumption 3.4 holds. Now let $\hat{s} > 0$ be such that

$$\gamma\hat{s} < \log\left(\frac{\gamma}{2} + 1\right),$$

which implies that

$$\lambda := \frac{2}{\gamma}(\exp\{\gamma\hat{s}\} - 1) < 1.$$

Throughout the rest of this section, we suppose that the following assumptions taken from [7, Section 5] hold.

Assumption 5.1. We have $0 < k_1 < \frac{1}{2}$.

Assumption 5.2. The i.i.d. disturbances z_t have a common density g , which is a continuous bounded function supported on the interval $S := [-\hat{s}, \hat{s}]$. Moreover, there exists a positive number ε such that $g(s) \geq \varepsilon$ for all $s \in S$.

Assumptions 5.1 and 5.2 imply that Assumptions 3.1, 3.2, and 3.3 hold (see, for instance, [7, Propositions 6, 23, and 24]).

Now consider the stationary policy defined by the stochastic kernel:

$$\varphi(da | x) := \begin{cases} \frac{k_2}{2k_1|x|} \mathbf{1}_{A(x)}(a) da & \text{if } x \neq 0, \\ \delta_0(a) da & \text{if } x = 0, \end{cases}$$

where $\mathbf{1}$ is the usual indicator function. The corresponding cost function C_φ can be obtained from (2.1). Indeed, a direct calculation gives, for every x ,

$$C_\varphi(x) = \hat{c}x^2 \quad \text{with} \quad \hat{c} := c_1 + \frac{c_2k_1^2}{3k_2^2} \tag{5.2}$$

for all $x \in \mathbb{X}$. In this case, the pair h_φ^* and ρ_φ satisfying the PE (3.4) becomes

$$\rho_\varphi = v_0\sigma^2 \quad \text{and} \quad h_\varphi^*(x) = v_0x^2, \tag{5.3}$$

with σ^2 as in (5.1) and

$$v_0 = \frac{\hat{c}}{1 - 4/3k_1^2}. \tag{5.4}$$

Moreover, Ψ_φ in (3.5) becomes

$$\Psi_\varphi(x) = v_0^2\left(\frac{64}{45}k_1^4x^4 + \frac{16}{3}\sigma^2k_1^2x^2 + 4Bk_1x + D - \sigma^4\right), \tag{5.5}$$

with

$$B := \int z^3 g(z) dz, \quad \text{and} \quad D := \int z^4 g(z) dz.$$

On the other hand, for every $t = 1, 2, \dots$, we have

$$E_x^\varphi x_{t+1} = k_1 E_x^\varphi x_t, \quad (5.6)$$

$$E_x^\varphi x_{t+1}^2 = \frac{4}{3}k_1^2 E_x^\varphi x_t^2 + \sigma^2, \quad (5.7)$$

and

$$E_x^\varphi x_{t+1}^4 = \frac{16}{5}k_1^4 E_x^\varphi x_t^4 + 8\sigma^2 k_1^2 E_x^\varphi x_t^2 + 4Bk_1 E_x^\varphi x_t + D. \quad (5.8)$$

From (5.5) and (5.6)–(5.8), it follows that σ_φ^2 , defined in Lemma 3.1(b), is given by

$$\sigma_\varphi^2 = \frac{v_0^2(45 - 80k_1^4)}{9(5 - 16k_1^4)} \left(D - \frac{3(448k_1^6 - 144k_1^4 - 300k_1^2 + 45)}{(3 - 4k_1^2)(45 - 80k_1^4)} \sigma^4 \right). \quad (5.9)$$

To conclude, by Theorem 3.1 and considering (5.3), it follows that, for every initial state $x \in \mathbb{X}$, as $n \rightarrow \infty$, the distribution of the cost

$$\frac{\sum_{t=0}^{n-1} C_\varphi(x_t) - nv_0\sigma^2}{\sqrt{n}}$$

has an asymptotic normal distribution $N(0, \sigma_\varphi^2)$ with σ_φ^2 as in (5.9).

Finally, by (5.2) and (5.4), it follows that, for every initial state x , as $n \rightarrow \infty$,

$$\frac{\sum_{t=0}^{n-1} x_t^2 - 3n\sigma^2/(3 - 4k_1^2)}{\sqrt{n}}$$

has an asymptotic normal distribution $N(0, s^2)$, where

$$s^2 = \frac{45 - 80k_1^4}{(3 - 4k_1^2)(5 - 16k_1^4)} \left(D - \frac{3(448k_1^6 - 144k_1^4 - 300k_1^2 + 45)}{(3 - 4k_1^2)(45 - 80k_1^4)} \sigma^4 \right).$$

References

- [1] GORDIENKO, E. AND HERNÁNDEZ-LERMA, O. (1995). Average cost Markov control processes with weighed norms: existence of canonical policies. *Appl. Math.* **23**, 199–218.
- [2] GORDIENKO, E. AND HERNÁNDEZ-LERMA, O. (1995). Average cost Markov control processes with weighed norms: value iteration. *Appl. Math.* **23**, 219–237.
- [3] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (1996). *Discrete-Time Markov Control Processes*. Springer, New York.
- [4] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (1999). *Further Topics on Discrete-time Markov Control Processes*. Springer, New York.
- [5] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (2001). Zero-sum stochastic games in Borel spaces: average payoff criteria. *SIAM J. Control Optimization* **39**, 1520–1539.
- [6] HERNÁNDEZ-LERMA, O., VEGA-AMAYA, O. AND CARRASCO, G. (1999). Sample-path optimality and variance-minimization of average cost Markov control processes. *SIAM J. Control Optimization* **38**, 79–93.
- [7] HILGERT, N. AND HERNÁNDEZ-LERMA, O. (2003). Bias optimality versus strong 0-discount optimality in Markov control processes with unbounded costs. *Acta Appl. Math.* **77**, 215–235.
- [8] JARNER, S. F. AND ROBERTS, G. O. (2002). Polynomial convergence rates of Markov chains. *Ann. Appl. Prob.* **12**, 224–247.
- [9] LÁNSKÁ, V. (1986). A note on estimation in controlled diffusion processes. *Kybernetika* **22**, 133–141.
- [10] LUQUE-VÁSQUEZ, F. AND HERNÁNDEZ-LERMA, O. (1999). Semi-Markov control models with average costs. *Appl. Math.* **26**, 315–331.
- [11] MANDL, P. (1971). On the variance in controlled Markov chains. *Kybernetika* **7**, 1–12.
- [12] MANDL, P. (1973). A connection between controlled Markov chains and martingales. *Kybernetika* **9**, 237–241.

- [13] MANDL, P. (1974). Estimation and control in Markov chains. *Adv. Appl. Prob.* **6**, 40–60.
- [14] MANDL, P. (1974). On the asymptotic normality of the reward in a controlled Markov chain. In *Progress in Statistics* (European Meeting of Statisticians, Budapest, 1972), Vol. II, Colloquia Mathematica Societatis János Bolyai Vol. 9, North-Holland, Amsterdam, pp. 499–505.
- [15] MANDL, P. AND LAUŠMANOVÁ, M. (1991). Two extensions of asymptotic methods in controlled Markov chains. *Ann. Operat. Res.* **28**, 67–79.
- [16] MENDOZA-PÉREZ, A. (2008). Asymptotic normality of average cost Markov control processes. *Morfismos* **12**, 33–52.
- [17] MENDOZA-PÉREZ, A. (2008). Pathwise average reward Markov control processes. Doctoral Thesis, CINVESTAV-IPN. Available at http://www.math.cinvestav.mx/ohernand_students.
- [18] MENDOZA-PÉREZ, A. F. AND HERNÁNDEZ-LERMA, O. (2010). Markov control processes with pathwise constraints. *Math. Meth. Operat. Res.* **71**, 477–502.
- [19] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [20] PRIETO-RUMEAU, T. AND HERNÁNDEZ-LERMA, O. (2009). Variance minimization and the overtaking optimality approach to continuous-time controlled Markov chains. *Math. Meth. Operat. Res.* **70**, 527–540.
- [21] PUTERMAN, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley, New York.
- [22] VEGA-AMAYA, O. (1998). Markov control processes in Borel spaces: undiscounted criteria. Doctoral Thesis, UAM-Iztapalapa (in Spanish).
- [23] ZHU, Q. X. AND GUO, X. P. (2007). Markov decision processes with variance minimization: a new condition and approach. *Stoch. Anal. Appl.* **25**, 577–592.