

Article

Not peer-reviewed version

On Dobrushin's Central Limit Theorem for Non-homogeneous Markov Chains in the Array Scheme, II

[Aisha Nurieva](#) *

Posted Date: 24 September 2025

doi: 10.20944/preprints202509.2020.v1

Keywords: Dobrushin's CLT; Markov chains; ergodic coefficients; array scheme



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

On Dobrushin's Central Limit Theorem for Non-homogeneous Markov Chains in the Array Scheme, II

Aisha Nurieva

NRU HSE; ai_nurieva@mail.ru

Abstract

This paper proposes a new sufficient condition for the Central Limit Theorem to hold for non-homogeneous Markov chains in an array scheme, even when the Markov-Dobrushin ergodicity coefficient violates the classical Dobrushin condition.

Keywords: Dobrushin's CLT; Markov chains; ergodic coefficients; array scheme

1. Introduction

An important central limit theorem (CLT) for non-homogeneous Markov chains was established by R.L. Dobrushin in 1956, a result that has since attracted considerable attention. This paper contributes a further extension of this celebrated theorem, which differs from the one proposed in our recent preprint (arXiv:2406.16156). The initial work in this direction by A.A. Markov was subsequently advanced by Bernstein, Sapogov, and Linnik, culminating in Dobrushin's seminal publication. Dobrushin ultimately showed that, under an additional non-degeneracy condition on the variances, the sufficient condition

$$\lim_{n \rightarrow \infty} n^{1/3} \alpha_n = \infty,$$

is also very close to necessary. Here, α_n denotes the Dobrushin (or Markov-Dobrushin) ergodic coefficient. The optimality of the rate $n^{1/3}$ was demonstrated by the Bernstein-Dobrushin example. While extensions exist for Markov chains with memory and the results have been simplified using martingale methods, this paper employs an adapted martingale approach to establish a CLT under new, slightly modified assumptions. Our motivation stems from the central role of the CLT in mathematical statistics.

The paper consists of 5 sections: Introduction, The setting, Main results, Auxiliary results, Proof of the theorem.

2. The Setting

Consider a non-homogeneous Markov chain defined on a probability space $(\Omega, \mathcal{F}, P)$. Let $\pi_{i,i+1}(x, dy)$ denote the Markov transition kernel governing the evolution from time i to $i + 1$ on a measurable space $(X, \mathcal{B}(X))$, which is assumed to satisfy the standard conditions for a Markov process (cf. [5]). The corresponding transition probabilities are defined by the relation:

$$P[X_{i+1} \in A | X_i = x] \stackrel{\text{a.s.}}{=} \pi_{i,i+1}(x, A),$$

where $\{X_i : 1 \leq i \leq n\}$ is the trajectory of the process on $[1, n]$. In particular, if the initial distribution is $X_1 \sim \mu$ then the distribution at time $k \geq 2$ is given by the measure $\mu \pi_{1,2} \pi_{2,3} \dots \pi_{k-1,k}$.

Due to the nature of conditional expectations, certain equalities and inequalities hold only almost surely (a.s.).

For $i < j$ let us denote by $\pi_{i,i+k}$ the k -step transition kernel, which has a form due to the Chapman-Kolmogorov equation,

$$\pi_{i,j} = \pi_{i,i+1} \pi_{i+1,i+2} \cdots \pi_{j-1,j}.$$

Definition 1. The Markov-Dobrushin (MD) coefficient¹ $\delta(\pi)$ is defined by any of the three equivalent expressions,

$$\delta(\pi) = \sup_{\substack{x_1, x_2 \in \mathbf{X}, \\ A \in \mathcal{B}(\mathbf{X})}} |\pi(x_1, A) - \pi(x_2, A)| = \frac{1}{2} \sup_{\substack{x_1, x_2 \in \mathbf{X}, \\ \|f\|_B \leq 1}} \left| \int f(y) [\pi(x_1, dy) - \pi(x_2, dy)] \right| \quad (1)$$

$$= \sup_{\substack{x_1, x_2 \in \mathbf{X}, \\ u \in \mathcal{U}}} |(\pi u)(x_1) - (\pi u)(x_2)|, \quad \text{where } \mathcal{U} = \{u(\cdot) : \sup_{y_1, y_2} |u(y_1) - u(y_2)| \leq 1\}.$$

where $\mathcal{U} = \{u(\cdot) : \sup_{y_1, y_2} |u(y_1) - u(y_2)| \leq 1\}$.

Here $\|f\|_B = \sup_{x \in \mathbf{X}} |f(x)|$. It follows from the definition that $0 \leq \delta(\pi) \leq 1$, and that $\delta(\pi) = 0$ if and only if the measure $\pi(x, dy)$ does not depend on x , and in the latter case the second random variable does not depend on the first one. The measure π will be called non-degenerate if $0 \leq \delta(\pi) < 1$. To denote integrals, standard notations will be used,

$$(\mu\pi)(A) = \int \mu(dx) \pi(x, A), \quad (\pi u)(x) = \int u(y) \pi(x, dy).$$

Let

$$\text{Osc}(u) := \sup_{x_1, x_2} |u(x_1) - u(x_2)|.$$

This semi-norm will be applied to random variables (see lemmata 13, 14), interpreted in the sense of the essential supremum norm. That is, for any random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$, we define:

$$\text{Osc}(X) := \text{ess sup}_{\mathbb{P}} X(\omega) - \text{ess inf}_{\mathbb{P}} X(\omega). \quad (2)$$

Further, for any transition kernels π_1, π_2 and two-step transition kernel $\pi_1 \pi_2$ (the first transition is according to the matrix π_1 , while the second – according to π_2), the following inequality holds true,

$$\delta(\pi_1 \pi_2) \leq \delta(\pi_1) \delta(\pi_2),$$

Similarly,

$$\begin{aligned} \delta(\pi_{i,j}) &\leq \delta(\pi_{i,i+1}) \delta(\pi_{i+1,i+2}) \cdots \delta(\pi_{j-1,j}). \\ \delta(\pi_{i,j}) &\leq (1 - \alpha_n)^{j-i}. \end{aligned} \quad (3)$$

Accordingly,

$$\alpha(\pi_{i,j}) = 1 - \delta(\pi_{i,j}).$$

In what follows, EY and $D[Y]$ stand for the expectation and for the variance of the random variable Y , respectively.

We work in an array scheme: for each $n \geq 1$, let $X_i^{(n)} : 1 \leq i \leq n$ be a Markov chain on $\mathbf{X}$ with transition kernels $\pi_{i,i+1}^{(n)}(x, dy) : 1 \leq i \leq n-1$ and initial distribution $\mu^{(n)}$. For simplicity, all processes $(X^{(n)})$ are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, though this assumption can be relaxed. Define the minimal ergodicity coefficient as:

$$\alpha_n = \min_{1 \leq i \leq n-1} \alpha(\pi_{i,i+1}^{(n)}).$$

¹ In [13] it is called just the contraction coefficient.

Further, let $\{f_i^{(n)} : \mathbf{X} \mapsto \mathbb{R}, 1 \leq i \leq n\}$ be real-valued functions on $\mathbf{X}$. For any $n \geq 1$ let

$$S_n := \sum_{i=1}^n f_i^{(n)}(X_i^{(n)}).$$

The main result of Dobrushin's paper, along with a key corollary, is presented in [13] as follows (we state it as a proposition to distinguish it from our main result, which will be termed a theorem). Here, " $\implies$ " denotes weak convergence.

Proposition 2. Let $C_n := \sup_{1 \leq i \leq n} \sup_{x \in \mathbf{X}} |f_i^{(n)}(x)| < \infty$. In this case, if

$$\lim_{n \rightarrow \infty} C_n^2 \alpha_n^{-3} \left[\sum_{i=1}^n D(f_i^{(n)}(X_i^{(n)})) \right]^{-1} = 0, \quad (4)$$

then the CLT holds true in the array scheme:

$$\frac{S_n - E(S_n)}{\sqrt{D(S_n)}} \implies \mathcal{N}(0, 1). \quad (5)$$

Corollary 3. Under the assumptions of the theorem, if $\sup_n C_n = C < \infty$, and $\inf_{i,n} D(f_i^{(n)}(X_i^{(n)})) \geq c > 0$, and if

$$\lim_{n \rightarrow \infty} n^{1/3} \alpha_n = \infty, \quad (6)$$

then weak convergence (5) holds.

3. Main Results

Let fix n and define $(\beta_k^{(n)}, 1 \leq k \leq n)$ to be a non-random sequence of ones and zeros, like (110010...01), and let for $1 \leq i \leq j \leq n$,

$$\kappa_j^\beta = \sum_{k=1}^j 1(\beta_k^{(n)} = 1)$$

Denote

$$\alpha_n^\beta := \min_{1 \leq i \leq n-1: \beta_i^{(n)}=1} \alpha(\pi_{i,i+1}^{(n)}).$$

Then (3) may be relaxed to

$$\delta(\pi_{i,j}) \leq (1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta}. \quad (7)$$

Definition 4. We say that the condition (H_β) is met if there exist $m_0 > 0$ and $c > 0$ such that for any n the sequence $(\beta_k^{(n)}, 1 \leq k \leq n)$ satisfies the following property: for any pair $(i < j)$ with $j - i \geq m_0$, the inequality

$$k_j^\beta - k_i^\beta \geq c(j - i) \quad (8)$$

is satisfied.

In what follows such a sequence $(\beta_k^{(n)}, 1 \leq k \leq n)$ is fixed for each n . The next theorem is the main result of the paper.

Theorem 5. Let the assumption (H_β) be met, and let also

$$\sup_{1 \leq i \leq n} \sup_{x \in \mathbf{X}} |f_i^{(n)}(x)| \leq C_n < \infty. \quad (9)$$

Then, if condition

$$\lim_{n \rightarrow \infty} C_n^2 \alpha_n^{-1} (\alpha_n^\beta)^{-2} \left[\sum_{i=1}^{n-1} \beta_i^{(n)} D(f_{i+1}^{(n)}(X_{i+1}^{(n)})) \right]^{-1} = 0 \quad (10)$$

is satisfied, then $C_n^{-2} D(S_n) \rightarrow \infty$ and the CLT holds true,

$$\frac{S_n - E(S_n)}{\sqrt{D(S_n)}} \Rightarrow \mathcal{N}(0, 1). \quad (11)$$

Corollary 6. If $\sup_n C_n = C < \infty$ and

$$\inf_{i,n} D(f_i^{(n)}(X_i^{(n)})) \geq c > 0, \quad (12)$$

and the conditions (8) along with

$$\lim_{n \rightarrow \infty} n \alpha_n (\alpha_n^\beta)^2 = \infty, \quad (13)$$

hold, then convergence (11) is valid.

4. Auxiliary Results

The following auxiliary results are collected in this section. Unless stated otherwise, they are reproduced without proof from [7] and [13]. However, proofs are provided for Lemmata 8, 12, and 14. Because, despite their simplicity, they involve new calculations that are crucial for proving the main result. The notation used is consistent with that in [13]. In particular, the essential supremum norm of a random variable Z is defined as:

$$\|Z\|_{L^\infty} := \operatorname{ess\,sup}_{\substack{P \\ \omega \in \Omega}} |Z(\omega)|. \quad (14)$$

Proposition 7. ([7, Theorem 3.2], [10, Theorem 5.5.8]) Let for any $n \geq 1$ the process $\{W_i^{(n)} : 0 \leq i \leq n\}$ be a martingale with respect to a filtration $(\mathcal{G}_i^{(n)})$, $W_0^{(n)} = 0$, and let $\xi_i^{(n)} := W_i^{(n)} - W_{i-1}^{(n)}$. If

$$\max_{1 \leq i \leq n} \|\xi_i^{(n)}\|_{L^\infty} \rightarrow 0, \quad \sum_{i=1}^n E[(\xi_i^{(n)})^2 | \mathcal{G}_i^{(n)}] \rightarrow 1 \quad \text{in } L^2 \quad (15)$$

Or in other words

$$\max_{1 \leq i \leq n} \operatorname{ess\,sup}_P |\xi_i^{(n)}(\omega)| \rightarrow 0, \quad E \left| \sum_{i=1}^n E[(\xi_i^{(n)})^2 | \mathcal{G}_i^{(n)}] - 1 \right|^2 \rightarrow 0$$

then the weak convergence takes place,

$$W_n^{(n)} \Rightarrow \mathcal{N}(0, 1).$$

Following [13], we assume without loss of generality that the functions $f_i^{(n)}$ are centered:

$$E[f_i^{(n)}(X_i^{(n)})] = 0, \quad 1 \leq i \leq n, \quad n \geq 1.$$

Define

$$Z_k^{(n)} := \sum_{i=k}^n E[f_i^{(n)}(X_i^{(n)}) | X_k^{(n)}], \quad (16)$$

so that

$$Z_k^{(n)} = \begin{cases} f_k^{(n)}(X_k^{(n)}) + \sum_{i=k+1}^n E[f_i^{(n)}(X_i^{(n)}) | X_k^{(n)}], & 1 \leq k \leq n-1, \\ f_n^{(n)}(X_n^{(n)}), & k = n. \end{cases}$$

Then, for any $1 \leq k \leq n-1$, the representation holds,

$$f_k^{(n)}(X_k^{(n)}) = Z_k^{(n)} - \mathbb{E}[Z_{k+1}^{(n)} | X_k^{(n)}]. \quad (17)$$

Furthermore, for any $2 \leq k \leq n-1$, this expression can be equivalently rewritten as

$$(Z_k^{(n)} - \mathbb{E}[Z_k^{(n)} | X_{k-1}^{(n)}]) + (\mathbb{E}[Z_k^{(n)} | X_{k-1}^{(n)}] - \mathbb{E}[Z_{k+1}^{(n)} | X_k^{(n)}])$$

Consequently, the sum S_n can be expressed in martingale form as follows:

$$S_n = \sum_{i=1}^n f_i^{(n)}(X_i^{(n)}) = Z_1^{(n)} + \sum_{k=2}^n [Z_k^{(n)} - \mathbb{E}[Z_k^{(n)} | X_{k-1}^{(n)}]].$$

As emphasized in [13], this transformation originates from the seminal paper [6] on a "simple" CLT for Markov chains. Given that all terms on the right-hand side of this representation are uncorrelated, it follows that

$$D(S_n) = \sum_{k=2}^n D(Z_k^{(n)} - \mathbb{E}[Z_k^{(n)} | X_{k-1}^{(n)}]) + D(Z_1^{(n)}).$$

Let

$$\xi_k^{(n)} = \frac{1}{\sqrt{D(S_n)}} [Z_k^{(n)} - \mathbb{E}[Z_k^{(n)} | X_{k-1}^{(n)}]] \quad (18)$$

$$M_k^{(n)} = \sum_{\ell=1}^k \xi_\ell^{(n)} \quad (19)$$

Then the process $M_k^{(n)}$ is a martingale with respect to the filtration $\mathcal{F}_k^{(n)} = \sigma\{X_\ell^{(n)} : 1 \leq \ell \leq k\}$ for $1 \leq k \leq n$. The central limit theorem will be established by approximating the normalized sum $\frac{S_n}{\sqrt{D(S_n)}}$ by the terminal martingale value $M_n^{(n)}$ and then applying a martingale CLT, following the approach in [13] (see Proposition 7). It therefore suffices to verify the two conditions in (15).

Further, by $f_i^{(n)} \pi_{i,j} f_j^{(n)}$ we denote the product of the functions $f_i^{(n)}$ and $\pi_{i,j} f_j^{(n)}$. Notice that for any $i < j$, n , and function $f_j^{(n)}$,

$$\text{Osc}(\pi_{i,j}(f_j^{(n)})) \leq \delta(\pi_{i,j}) \text{Osc}(f_j^{(n)}). \quad (20)$$

Lemma 8. Under the conditions of the theorem, for any $1 \leq i \leq j \leq n$,

$$\|\pi_{i,j} f_j^{(n)}\|_B \leq 2C_n(1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta}, \quad \text{Osc}(\pi_{i,j}(f_j^{(n)}))^2 \leq 2C_n^2(1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta}; \quad (21)$$

and also for $1 \leq l < i \leq j \leq n$,

$$\text{Osc}(\pi_{l,i}(f_i^{(n)} \pi_{i,j} f_j^{(n)})) \leq 6C_n^2(1 - \alpha_n^\beta)^{\kappa_i^\beta - \kappa_l^\beta} (1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta} = 6C_n^2(1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_l^\beta}; \quad (22)$$

Proof. Since $\|f_j^{(n)}\| \leq C_n$, then $\text{Osc}(f_j^{(n)}) \leq 2C_n$. It follows from the definitions of the coefficients δ and $\alpha_n^{(2)}$, and from (3) that

$$\text{Osc}(\pi_{i,j} f_j^{(n)}) \leq \delta(\pi_{i,j}) \text{Osc}(f_j^{(n)}) \leq 2C_n(1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta}.$$

Since $\mathbb{E}[(\pi_{i,j} f_j^{(n)})(X_i^{(n)})] = \mathbb{E}[f_j^{(n)}(X_j^{(n)})] = 0$, the first inequality in (21) follows from the bounds

$$\|\pi_{i,j} f_j^{(n)}\|_B \leq \text{Osc}(\pi_{i,j} f_j^{(n)}) \leq 2C_n(1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta}.$$

The second inequality in (21) is established similarly. Indeed, we have due to (20),

$$\text{Osc}(\pi_{i,j}(f_j^{(n)})^2) \leq \delta(\pi_{i,j}) \text{Osc}((f_j^{(n)})^2) \leq 2C_n^2(1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta}.$$

To prove the third inequality of the lemma, we estimate using (21):

$$\begin{aligned} \text{Osc}(\pi_{l,i}(f_i^{(n)} \pi_{i,j} f_j^{(n)})) &\leq (1 - \alpha_n^\beta)^{\kappa_l^\beta - \kappa_i^\beta} \text{Osc}(f_i^{(n)} \pi_{i,j} f_j^{(n)}) \\ &\leq (1 - \alpha_n^\beta)^{\kappa_l^\beta - \kappa_i^\beta} \left(\text{Osc}(f_i^{(n)}) \|\pi_{i,j} f_j^{(n)}\|_B + \|f_j^{(n)}\|_B \text{Osc}(\pi_{i,j} f_j^{(n)}) \right) \\ &\leq 6C_n^2(1 - \alpha_n^\beta)^{\kappa_l^\beta - \kappa_i^\beta} (1 - \alpha_n^\beta)^{\kappa_j^\beta - \kappa_i^\beta}, \end{aligned}$$

as required. $\square$

The following lemma provides a lower bound for the variance of the sum, a result that is crucial for proving the main theorem. We were unable to improve this bound using the coefficient α_n^β ; consequently, we state it exactly as in [13] and without proof. Nevertheless, even with the weaker coefficient α_n , this bound remains highly useful in the subsequent analysis.

Proposition 9 ([13], proposition 3.2). *Under the assumptions of the theorem 5, for any $n \geq 1$,*

$$D(S_n) \geq \frac{\alpha_n}{4} \sum_{i=1}^n D(f_i^{(n)}(X_i^{(n)})).$$

To prove this proposition, the following auxiliary statements are necessary.

Lemma 10 (Lemma 4.1, [13]). *Let λ be a probability measure on $\mathbf{X} \times \mathbf{X}$ with marginals α and β respectively. Let $\pi(x_1, dx_2)$ and $\hat{\pi}(x_2, dx_1)$ be the corresponding transition probabilities in the two directions so that $\alpha\pi = \beta$ and $\beta\hat{\pi} = \alpha$. Let $f(x_1)$ and $g(x_2)$ be square integrable with respect to α and β respectively. If*

$$\int f(x_1)\alpha(dx_1) = \int g(x_2)\beta(dx_2) = 0$$

then,

$$\left| \int f(x_1)g(x_2)\lambda(dx_1, dx_2) \right| \leq \sqrt{\delta(\pi)} \|f\|_{L_2(\alpha)} \|g\|_{L_2(\beta)}.$$

Lemma 11 (Lemma 4.2, [13]). *Let $f(x_1)$ and $g(x_2)$ be square integrable with respect to α and β respectively. Then,*

$$E[(f(x_1) - g(x_2))^2] \geq \alpha(\pi)V(f(x_1))$$

as well as

$$E[(f(x_1) - g(x_2))^2] \geq \alpha(\pi)V(g(x_2)).$$

We now estimate the ratio $|Z_k^{(n)}|L^\infty / \sqrt{D(S_n)}$ in the following lemma. Note that by the definition of the L^∞ norm in (14), the value $\|Z_k^{(n)}\|_{L^\infty}$ is nonrandom; see the definition of $Z_k^{(n)}$ in (16).

Lemma 12. *Under the assumptions of the theorem 5, the equality holds,*

$$\lim_{n \rightarrow \infty} \sup_{1 \leq k \leq n} \frac{\|Z_k^{(n)}\|_{L^\infty}}{\sqrt{D(S_n)}} = 0. \quad (23)$$

Proof. We have,

$$EE[f_i^{(n)}(X_i^{(n)})|X_k^{(n)}] = 0.$$

Therefore, because of the equality

$$\mathbb{E}[f_i^{(n)}(X_i^{(n)})|X_k^{(n)}] \stackrel{\text{a.s.}}{=} \pi_{k,i} f_i^{(n)}(X_k^{(n)}),$$

and by virtue of (21), we obtain,

$$\begin{aligned} \|\mathbb{E}[f_i^{(n)}(X_i^{(n)})|X_k^{(n)}]\|_{L^\infty} &= \operatorname{ess\,sup}_{\mathbf{P}} \sup_{\omega \in \Omega} |\mathbb{E}[f_i^{(n)}(X_i^{(n)})|X_k^{(n)}]| \\ &= \operatorname{ess\,sup}_{\mathbf{P}} \sup_{\omega} |\pi_{k,i} f_i^{(n)}(X_k^{(n)})| \leq \|\pi_{k,i} f_i^{(n)}\|_B \leq 2C_n (1 - \alpha_n^\beta)^{\kappa_i^\beta - \kappa_k^\beta}. \end{aligned}$$

So, it follows,

$$\begin{aligned} \|Z_k^{(n)}\|_{L^\infty} &\leq \sum_{i=k}^n \|\mathbb{E}[f_i^{(n)}(X_i^{(n)})|X_k^{(n)}]\|_{L^\infty} = \sum_{i=k}^n \|\pi_{k,i} f_i^{(n)}\|_B \leq \sum_{i=k}^n 2C_n (1 - \alpha_n^\beta)^{\kappa_i^\beta - \kappa_k^\beta} \\ &\leq 2C_n m_0 + 2C_n \sum_{i=k+m_0+1}^n 2C_n (1 - \alpha_n^\beta)^{\kappa_i^\beta - \kappa_k^\beta} \stackrel{(8)}{\leq} 2C_n m_0 + 2C_n \sum_{i=k+m_0+1}^n (1 - \alpha_n^\beta)^{c(i-k)} \\ &\leq 2C_n m_0 + 2C_n \sum_{i=k+m_0+1}^\infty (1 - \alpha_n^\beta)^{c(i-k)} \leq 2C_n m_0 + 2C_n \sum_{i=k+m_0+1}^n (1 - \alpha_n^\beta)^{c(i-k)} \leq \end{aligned}$$

The last inequality is due to the formula for the geometric sum with the common ratio $0 \leq (1 - \alpha_n^\beta) < 1$. Continue sequence of inequalities:

$$\leq 2C_n m_0 + \frac{2C_n}{1 - (1 - \alpha_n^\beta)^c} \leq 2C_n m_0 + \frac{2C_n}{c\alpha_n^\beta} = 2C_n \frac{(cm_0 + 1)/c}{\alpha_n^\beta}.$$

And here we use inequality $(1 - \alpha_n^\beta)^c \leq 1 - c\alpha_n^\beta$. Finally, applying the bound from the proposition 9, we get,

$$\sup_{1 \leq k \leq n} \frac{\|Z_k^{(n)}\|_{L^\infty}}{\sqrt{D(S_n)}} \leq 4C_n \left((\alpha_n^\beta)^2 \alpha_n \sum_{i=1}^n 1(\beta_i^{(n)} = 1) D(f_i^{(n)}(X_i^{(n)})) \right)^{-1/2}.$$

Due to the condition (9) this implies the inequality (23), as required. $\square$

The proof of the following lemma is given in [13, Lemma 3.3] and is therefore omitted here. Recall that the semi-norm Osc , when applied to a random variable, is defined as in (2).

Lemma 13. Let $\{Y_l^{(n)} : 1 \leq l \leq n\}$ and $\{\mathcal{G}_l^{(n)} : 1 \leq l \leq n\}$, for $n \geq 1$, are, respectively, the sequence of non-negative random variables and sigma-fields such that $\sigma\{Y_1^{(n)}, \dots, Y_l^{(n)}\} \subset \mathcal{G}_l^{(n)}$. Suppose that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sum_{l=1}^n Y_l^{(n)} \right] = 1 \quad \text{and} \quad \sup_{1 \leq i \leq n} \|Y_i^{(n)}\|_{L^\infty} \leq \epsilon_n,$$

where $\lim_{n \rightarrow \infty} \epsilon_n = 0$. Suppose in addition that

$$\lim_{n \rightarrow \infty} \sup_{1 \leq l \leq n-1} \text{Osc} \left(\mathbb{E} \left[\sum_{j=l+1}^n Y_j^{(n)} \mid \mathcal{G}_l^{(n)} \right] \right) = 0.$$

Then

$$\lim_{n \rightarrow \infty} \sum_{l=1}^n Y_l^{(n)} = 1 \quad \text{in } L^2.$$

Further, denote random variables

$$v_j^{(n)} := \mathbb{E}[(\xi_j^{(n)})^2 \mid \mathcal{F}_{j-1}^{(n)}], \text{ where } \xi_j^{(n)} = \frac{1}{D(S_n)}[Z_k^{(n)} - \mathbb{E}[Z_k^{(n)} \mid X_{k-1}^{(n)}]]$$

These random variables are measurable with respect to the sigma-fields $\mathcal{G}_j^{(n)} = \mathcal{F}_{j-1}^{(n)}$ for $2 \leq j \leq n$, respectively.

Lemma 14. *Under the assumptions of the theorem 5 the convergence*

$$\sup_{2 \leq l \leq n-1} \text{Osc} \left(\mathbb{E} \left[\sum_{j=l+1}^n v_j^{(n)} \mid \mathcal{F}_{l-1}^{(n)} \right] (\omega) \right) = o(1), \quad n \rightarrow \infty,$$

holds true.

Note that in this paper, all $o(1)$ terms are deterministic. This is a consequence of our convention for the oscillation semi-norm applied to random variables, as defined in (2).

Proof. By the martingale property and definitions (17) and (18), it follows that $\mathbb{E}[\xi_r^{(n)} \xi_s^{(n)} \mid \mathcal{F}_u^{(n)}] = 0$ (a.s.) for all $r > s > u$. Therefore,

$$\begin{aligned} \mathbb{E} \left[\sum_{j=l+1}^n v_j^{(n)} \mid \mathcal{F}_{l-1}^{(n)} \right] &= \mathbb{E} \left[\sum_{j=l+1}^n (\xi_j^{(n)})^2 \mid \mathcal{F}_{l-1}^{(n)} \right] = \mathbb{E} \left[\left(\sum_{j=l+1}^n \xi_j^{(n)} \right)^2 \mid X_{l-1}^{(n)} \right] \\ &= (DS_n)^{-1} \mathbb{E} \left[\left(\sum_{j=l+1}^n f_j^{(n)}(X_j^{(n)}) - \mathbb{E}[Z_{l+1}^{(n)} \mid X_l^{(n)}] \right)^2 \mid X_{l-1}^{(n)} \right] \\ &= (DS_n)^{-1} \mathbb{E} \left[\left(\sum_{j=l+1}^n f_j^{(n)}(X_j^{(n)}) \right)^2 \mid X_{l-1}^{(n)} \right] - (DS_n)^{-1} \mathbb{E} \left[\mathbb{E}[Z_{l+1}^{(n)} \mid X_l^{(n)}]^2 \mid X_{l-1}^{(n)} \right]. \end{aligned} \quad (24)$$

By Lemma 12, the second term in (24) is bounded above by $\sup_{2 \leq l \leq n-1} (DS_n)^{-1} \|Z_{l+1}^{(n)}\|_{L^\infty}^2$, which is $o(1)$. Consequently, its oscillation is also $o(1)$ as $n \rightarrow \infty$, uniformly in ω . Recall that $Z_k^{(n)}$ is defined as (cf. (16)):

$$Z_k^{(n)} := \sum_{i=k}^n \mathbb{E}[f_i^{(n)}(X_i^{(n)}) \mid X_k^{(n)}]$$

We now estimate the oscillation of the first term in (24).

$$\begin{aligned} &\text{Osc} \left((DS_n)^{-1} \mathbb{E} \left[\left(\sum_{j=l+1}^n f_j^{(n)}(X_j^{(n)}) \right)^2 \mid X_{l-1}^{(n)} \right] \right) \\ &\leq (DS_n)^{-1} \sum_{l+1 \leq j, m \leq n} \text{Osc} \left(\mathbb{E} \left[f_j^{(n)}(X_j^{(n)}) f_m^{(n)}(X_m^{(n)}) \mid X_{l-1}^{(n)} \right] \right). \end{aligned}$$

Let us highlight that our definition of “Osc” for r.v. in (2) allows us not to worry about the difference between sure and almost sure inequalities, despite conditional expectations in the expressions. We now apply Lemma 8:

$$\text{Osc}\left(\mathbb{E}\left[f_j^{(n)}\left(X_j^{(n)}\right)f_m^{(n)}\left(X_m^{(n)}\right) \mid X_{l-1}^{(n)}\right]\right) \leq 6C_n^2\left(1-\alpha_n^\beta\right)^{\kappa_j^\beta-\kappa_l^\beta}\left(1-\alpha_n^\beta\right)^{|\kappa_m^\beta-\kappa_j^\beta|}.$$

In the remaining cases the estimates are analogous, with at most one or two additional factors of $(1-\alpha_n)$, each bounded above by 1. Summing over all terms then yields the bound

$$\text{Osc}(\cdot) \leq (DS_n)^{-1}C_n^2(\alpha_n^\beta)^{-2},$$

uniformly with respect to l . Combining this estimate with Proposition 9 and condition 9 completes the proof of the lemma.

□

5. Proof of the Theorem 5

Proof. It follows from the lemma 12 that it suffices to show convergence

$$M_n^{(n)} / \sqrt{D(S_n)} \implies \mathcal{N}(0,1).$$

This property follows from the proposition 7, because it was already verified that

- (a) $\sup_{2 \leq k \leq n} \|\xi_k^{(n)}\|_{L^\infty} \rightarrow 0, \quad n \rightarrow \infty;$
- (b) $\mathbb{E}\left|\sum_{k=2}^n \mathbb{E}\left[\left(\xi_k^{(n)}\right)^2 \mid \mathcal{F}_{k-1}^{(n)}\right]\right|^2 \rightarrow 1, \quad n \rightarrow \infty.$

Here, part (a) follows from the lemma 12 (negligibility). To verify condition (b), we apply 13, 14 (LLN). This is valid since (a) holds and the variance decomposition (after 13) combined with Lemma 12 gives $\sum_{k=2}^n \mathbb{E}[(\xi_k^{(n)})^2] = 1 + o(1)$.

□

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The work was supported by the grant of the Theoretical Physics and Mathematics Advancement Foundation “BASIS”. The author is also thankful to scientific advisor Alexander Veretennikov for valuable remarks.

Conflicts of Interest: No conflict of interests.

References

1. S.N. Bernstein, Sur l’extension du théorème du calcul de probabilités aux sommes de quantités dépendantes. *Math. Ann.* **1926**, *97*, 1–59.
2. Russian transl.: S.N. Bernstein, Extension of a limit theorem of probability theory on sums of dependent random variables, *Russian Math. Rev.*, **1944**, *10*, 65–114 (in Russian). <https://www.mathnet.ru/rus/rm8795>
3. O.A. Butkovsky, A.Yu. Veretennikov, On asymptotics for Vaserstein coupling of Markov chains, *Stochastic Processes and their Applications*, 2013, *123*(9), 3518–3541. <https://doi.org/10.1016/j.spa.2013.04.016>
4. R.L. Dobrushin, Central limit theorems for non-stationary Markov chains I,II. *Theory Probab. its Appl.* **1956**, *1*, 65–80, 329–383.
5. E.B. Dynkin, *Markov processes*, vol. 1, Springer-Verlag, Berlin, Heidelberg, 1965. <https://doi.org/10.1007/978-3-662-00031-1>
6. M.I. Gordin, The central limit theorem for stationary processes. *Soviet Math. Dokl.* **1969**, *10*, 1174–1176. <https://www.mathnet.ru/eng/dan34921>

7. P. Hall, C.C. Heyde, *Martingale Limit Theory and Its Application*. Academic Press, New York, 1980. <https://doi.org/10.1016/C2013-0-10818-5>
8. Yu.V. Linnik, On the theory of nonuniform Markov chains. *Izv. Akad. Nauk SSSR Ser. Mat.* **1949**, 13, 65–94. (in Russian)
9. Yu.V. Linnik, N.A. Sapogov, Multiple integral and local laws for inhomogeneous Markov chains. *Izv. Akad. Nauk SSSR Ser. Mat.* **1949**, 13, 533–566 (in Russian). <https://www.mathnet.ru/rus/im3213>
10. R.Sh. Liptser, A.N. Shiryaev, *Theory of Martingales*, Kluwer Acad. Publ., Dordrecht, The Netherlands, 1989.
11. A.A. Markov, Investigation of the general case of trials associated into a chain (in Russian), *Zapiski. Akad. Nauk (St. Petersburg) Fiz.-Matem. Otd., (8th Ser.)* 25(3)(1910), 33 pp.; also in: *Selected works*, 1951, Moscow, Acad. Sci. USSR, 465–509 (in Russian); Eng. transl. in: O.B. Sheynin, ed., *Probability and Statistics. Russian Papers*, Selected and Translated by Oscar Sheynin. NG Verlag, Berlin, 2004, 181–203.
12. A.N. Shiryaev, Central limit theorem for non-homogeneous Markov chains with memory, Summary of Papers Presented at the Sessions of the Scientific Research Seminar on Probability Theory, (Moscow, February–May, 1957). *Theory of Probability and its Applications* **1957**, 2, 470–480. <https://doi.org/10.1137/1102035>
13. S. Sethuraman, S.R.S. Varadhan, A Martingale Proof of Dobrushin's Theorem for Non-Homogeneous Markov Chains. *Electron. J. Probab.* **2005**, 10, 1221–1235. <https://doi.org/10.1214/EJP.v10-283>
14. A.Yu. Veretennikov, M.A. Veretennikova, On improved bounds and conditions for the convergence of Markov chains. *Izvestiya Mathematics* **2022**, 86, 92–125. <https://doi.org/10.1070/IM9076>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.