Afrika Statistika

Afrika Statistika
Vol. 19 (4), 2024, pages 4135 - 4171.
DOI: https://dx.doi.org/10.16929/as/2024.4135.340

ISSN 2316-090X

On The Central Limit Theorem and Rates of Normal Approximation for Stationary ϕ-Mixing Sequences with Polynomial Decay

Harouna Sangaré [1,2,3], **Ousmane Coulibaly** [1,*] **and Abdoulaye Camara** [1,2,3]

[1] Laboratoire de Mathématiques, Informatique et Applications (LaMIA),
Université des Sciences, des Techniques et des Technologies de Bamako (USTT-B), Mali.
[2] Département d'Enseignement et de Recherche de Mathématiques et d'Informatique (DER MI), Faculté des Sciences et Techniques (FST),
Université des Sciences, des Techniques et des Technologies de Bamako (USTT-B), Mali.
[3] Imhotep International Mathematical Center (imh-imc): https://imhotepsciences.org

Received on September 10, 2025; Accepted on October 24, 2025 ; Published online on November 17, 2025

Abstract. This paper investigates the central limit theorem (CLT) and the rate of normal approximation for strictly stationary, polynomially decaying, ϕ-mixing sequences. We derive characteristic function inequalities and covariance bounds of the type proposed by Ibragimov (1962), and prove a central limit theorem under summability conditions involving $(2+\delta)$-moments. Additionally, we derive rates of convergence in the Kolmogorov distance between normalized sums and the Gaussian law. Our results refine classical theorems in dependent probability theory and extend existing Berry-Esseen type bounds to the setting of ϕ-mixing with polynomial decay.

Key words: Central limit theorem; Berry-Esseen; ϕ-mixing; Stationarity.
AMS 2010 Mathematics Subject Classification Objects : 60F05; 60G10; 62E20; 60E10.

---

[*] Corresponding author : Ousmane Coulibaly (ocouwi@gmail.com)
Harouna Sangaré : harounasangareusttb@gmail.com, harouna.sangare@usttb.edu.ml
Abdoulaye Camara : abdoulahicamara01@gmail.com, camara.abdoulaye2@ugb.edu.sn

Résumé (French Abstract) Cet article étudie le théorème central limite (TCL) et les bornes de Berry-Esseen pour des suites strictement stationnaires et ϕ-mélangeantes à décroissance polynomiale. Nous établissons des inégalités de fonction caractéristique et des bornes de covariance de type by Ibragimov (1962). Un TCL est démontré sous des conditions de sommabilité avec moments d'ordre $(2 + \delta)$. Nous obtenons aussi des taux explicites de convergence en distance de Kolmogorov. Ces résultats affinent les théorèmes classiques pour variables dépendantes et étendent les bornes de Berry-Esseen au cadre ϕ-mélangeant polynômial.

Presentation of authors

Harouna Sangaré, Ph.D., is a Professor of Statistics and Mathematics at Université des Sciences, des Techniques et des Technologies de Bamako , Mali (USTT-B). He is a lead researcher at: Laboratoire de Mathématiques, Informatique et Applications (LaMIA), and Imhotep International Mathematical Center (imh-imc): https://imhotepsciences.org.

Ousmane Coulibaly, M.Sc, is preparing a Ph.D. degree at: Laboratoire de Mathématiques, Informatique et Applications (LaMIA), Université des Sciences, des Techniques et des Technologies de Bamako (USTT-B), Mali.

Abdoulaye Camara, M.Sc, is an Assistant of Statistics and Mathematics at Université des Sciences, des Techniques et des Technologies de Bamako , Mali (USTT-B). He is a researcher at: Laboratoire de Mathématiques, Informatique et Applications (LaMIA), and Imhotep International Mathematical Center (imh-imc): https://imhotepsciences.org.

1. Introduction

The central limit theorem (CLT) is one of the cornerstones of probability theory. It ensures that sums of random variables exhibit Gaussian fluctuations when appropriately normalized. Although independence assumptions are standard, many applications in econometrics, statistical physics and time series analysis involve dependent sequences. This has led to the study of mixing conditions that quantify weak dependence.

Among various notions of mixing, ϕ-mixing (also called absolute regularity in some contexts) plays a fundamental role due to its strong control of conditional probabilities. Introduced by Ibragimov in the 1960s, it has been extensively studied for establishing asymptotic normality and invariance principles.

The precise rates of convergence to the Gaussian law under dependence remain a challenging problem. For independent sequences, the Berry-Esseen theorem guarantees a universal $O(n^{-1/2})$ bound under finite third moments. However, for dependent sequences, the rates are slower and depend delicately on the decay of mixing coefficients.

Classical results include Ibragimov's central limit theorem for strongly mixing sequences Ibragimov (1962), and subsequent refinements for various mixing conditions Bradley (2005). For ϕ-mixing sequences, central limit theorems have been established under summability conditions on $\phi(h)$ and higher moments Ibragimov and Linnik (1971).

Rates of normal approximation in dependent settings have been addressed in works such as Rio (2017), Doukhan (1994), and Utev (1991). Recently, new advances have clarified Berry-Esseen type bounds for polynomially decaying mixing coefficients Rio (2000a). However, explicit polynomial rates for ϕ-mixing sequences with $(2+\delta)$ moments remain less explored, motivating our contribution.

In this paper, we provide an overview of the theory of asymptotic normality of the sums of ϕ-mixing random variables, adopting a coherent approach. This is intended to facilitate further contributions from new researchers in the field.

Our work is motivated by the need to quantify the rate of convergence to Gaussianity for dependent sequences with polynomial ϕ-mixing. Our objective is threefold

1. Derive a sharp *characteristic function factorization inequality* under ϕ-mixing, which generalizes block factorization techniques.
2. Establish a *central limit theorem* (Theorem 1) for stationary ϕ-mixing sequences under $(2+\delta)$ moments and polynomial mixing.
3. Provide an explicit *Berry-Esseen type bound* (Theorem 2) quantifying the rate of normal approximation.

The methodological novelty of this paper lies in extending the central limit theorem and Berry-Esseen type bounds to stationary ϕ-mixing sequences with polynomial decay, a setting that has received limited attention in the literature. Our approach combines a refined block factorization inequality for characteristic functions with truncation and interpolation techniques, allowing us to work under minimal $(2+\delta)$-moment assumptions. This yields explicit polynomial rates of convergence in Kolmogorov distance, thereby sharpening classical Ibragimov-type results and providing the first systematic treatment of normal approximation rates for polynomially decaying ϕ-mixing processes.

This leads us to the following organization of the paper. In Section 2 we recall the basic definitions and notations concerning ϕ-mixing sequences and the framework of polynomial decay. Section 3 contains our main results and is divided into three parts: first, we establish characteristic function inequalities under ϕ-mixing; next we prove a central limit theorem for stationary sequences with polynomial mixing coefficients; finally, we derive Berry-Esseen type bounds that quantify the rate of normal approximation. Technical arguments, including a truncation and interpolation lemma for covariance bounds, are provided in the appendix section. Section 4 concludes the paper with a summary of the main contributions and perspectives for further research.

To begin with, we give a short reminder of the concept of ϕ-mixing.

2. Preliminaries and Background

In order to establish our results, we recall some standard definitions and notations related to mixing sequences. Throughout the paper, let $(X_k)_{k\in\mathbb{Z}}$ be a sequence of random variables of real value defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For integers $a \le b$, denote by

$$\mathcal{F}_a^b = \sigma(X_j : a \le j \le b),$$

the σ-algebra generated by random variables $X_a, \dots, X_b$.

Definition 1 (ϕ-mixing sequence). A strictly stationary sequence $(X_k)_{k\in\mathbb{Z}}$ of real random variables is said to be ϕ-*mixing* (or strongly mixing in the sense of Ibragimov) if

$$\phi(h) = \sup_{\substack{A\in\mathcal{F}_{-\infty}^k,\, B\in\mathcal{F}_{k+h}^\infty \\ \mathbb{P}(A)>0}} \left|\mathbb{P}(B \mid A) - \mathbb{P}(B)\right| \xrightarrow[h\to\infty]{} 0,$$

where $\mathcal{F}_{-\infty}^k = \sigma(X_j : j \le k)$ and $\mathcal{F}_{k+h}^\infty = \sigma(X_j : j \ge k+h)$.

The coefficient $\phi(h)$ measures the degree of dependence between the past $\{X_j : j \le k\}$ and the future $\{X_j : j \ge k+h\}$ of the process. Intuitively, the smaller $\phi(h)$ is, the weaker the dependence between the time lags of length h. The ϕ-mixing condition is among the strongest notions of mixing: it implies both α-mixing (strong mixing) and ρ-mixing, and thus provides a robust framework for establishing limit theorems in dependent settings (see Ibragimov (1962), Bradley (2005), and Doukhan (1994)).

In this work, we focus on sequences whose dependence decays at a polynomial rate. More precisely, we assume that there exist constants $C > 0$ and $\gamma > 0$ such that

$$\phi(h) \le Ch^{-\gamma}, \quad \text{for all } h \ge 1. \tag{1}$$

Condition (1) is referred to as *polynomial decay of ϕ-mixing coefficients*. The parameter γ quantifies the speed of decorrelation : a larger γ corresponds to a faster convergence of $\phi(h)$ towards zero. This framework is particularly relevant for many stochastic processes encountered in practice, such as certain Markov chains, dynamical systems with weak dependence, and linear processes with suitable innovations (see, for example, Rio (2017) and Doukhan (1994)).

The assumption of polynomial decay provides a balance between generality and tractability: while exponential mixing guarantees very fast decorrelation, polynomial mixing allows for a wider class of processes exhibiting long-range dependence, yet still permits the derivation of central limit theorems and normal approximation bounds. These properties will be crucial in the developments that follow.

2.1. Classical Central Limit Theorems: From Independence to Dependence

The Central Limit Theorem (CLT) has undergone a long and rich development, starting with results for independent sequences and progressively extending to weakly dependent structures such as mixing sequences. In this section, we recall some fundamental milestones.

2.1.1. Independent Sequences

The Lindeberg-Lévy Theorem. The first rigorous version of the CLT concerns independent and identically distributed (i.i.d.) random variables. Let $(X_k)_{k\geq 1}$ be i.i.d. with mean $\mathbb{E}[X_1] = 0$ and finite variance $\mathrm{Var}(X_1) = \sigma^2 < \infty$. Then, the normalized sum satisfies

$$\frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n} X_k \xrightarrow{d} \mathcal{N}(0,1) \text{ as } n \to \infty. \tag{2}$$

This classical result is due to Lindeberg (1922) and Lévy (1925).

The Lyapunov Theorem. To go beyond the identically distributed case, Lyapunov (1901) formulated a CLT for triangular arrays of independent random variables. Let $\{X_{n,k} : 1 \leq k \leq n\}$ be independent with mean zero and variances $\sigma_{n,k}^2$. Define $s_n^2 = \sum_{k=1}^{n} \sigma_{n,k}^2 \to \infty$. If there exists $\delta > 0$ such that

$$\frac{1}{s_n^{2+\delta}}\sum_{k=1}^{n} \mathbb{E}\left[|X_{n,k}|^{2+\delta}\right] \longrightarrow 0 \text{ as } n \to \infty, \tag{3}$$

then

$$\frac{1}{s_n}\sum_{k=1}^{n} X_{n,k} \xrightarrow{d} \mathcal{N}(0,1) \text{ as } n \to \infty. \tag{4}$$

This result illustrates the importance of higher-order moment conditions for establishing asymptotic normality when the assumption of identical distribution is dropped.

2.1.2. Dependent Sequences and Mixing Conditions

The extension of the CLT to dependent structures has been a central theme of probability theory since the mid-20th century. One natural framework to control dependence is given by mixing conditions, such as α-mixing (strong mixing) and ϕ-mixing.

Ibragimov's Theorem. Let $(X_k)_{k\in\mathbb{Z}}$ be a strictly stationary sequence with mean zero and finite variance. Assume that (X_k) is α-mixing with coefficients $\alpha(h) \to 0$ as $h \to \infty$. Suppose further that $\mathbb{E}[|X_0|^{2+\delta}] < \infty$ for some $\delta > 0$, and that

$$\sum_{h=1}^{\infty} \alpha(h)^{\delta/(2+\delta)} < \infty. \tag{5}$$

Then

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \xrightarrow{d} \mathcal{N}(0, \sigma^2) \text{ as } n \to \infty, \tag{6}$$

where the asymptotic variance is given by

$$\sigma^2 = \mathrm{Var}(X_0) + 2\sum_{h=1}^{\infty} \mathrm{Cov}(X_0, X_h). \tag{7}$$

This fundamental extension to dependent sequences was established by Ibragimov (1962).

Ibragimov-Linnik. The monograph of Ibragimov (1962) and Ibragimov and Linnik (1971) provided a systematic account of CLTs for mixing sequences. They developed the theory for both α- and ϕ-mixing processes, showing that the CLT holds under moment and summability conditions similar in spirit to the result of Ibragimov (1962). In particular, ϕ-mixing, being a stronger form of dependence control, allows slightly weaker moment assumptions while still guaranteeing the CLT.

In summary, the CLT originated in the early 20th century with results for independent sequences (Lindeberg-Lévy, Lyapunov), and was extended in the 1960s-1970s to dependent settings via mixing conditions (Ibragimov (1962), Ibragimov and Linnik (1971)). These results form the classical foundation for modern studies of central limit theorems under weak dependence.

2.2. Berry-Esseen Type Bounds: Independent vs Dependent Case

A natural refinement of the central limit theorem (CLT) is given by the Berry-Esseen theorem, which provides a quantitative estimate of the speed of convergence towards the Gaussian distribution.

2.2.1. The classical independent case

Let $\{X_k\}_{k\geq 1}$ be independent and identically distributed (i.i.d.) random variables with

$$\mathbb{E}[X_1] = 0, \quad \mathbb{E}[X_1^2] = \sigma^2 > 0, \quad \text{and} \quad \mathbb{E}[|X_1|^3] < \infty.$$

Define the normalized sum

$$S_n := \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} X_k.$$

The Berry-Esseen theorem states that there exists a universal constant $C > 0$ such that, for all $n \geq 1$,

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(S_n \leq x) - \Phi(x)| \ \leq \ \frac{C\,\mathbb{E}[|X_1|^3]}{\sigma^3\sqrt{n}}, \tag{8}$$

where Φ denotes the cumulative distribution function (cdf) of the standard normal law.

This inequality shows that, in the independent case, the Kolmogorov distance between the distribution of the normalized sum and the Gaussian distribution decreases at the universal rate $O(n^{-1/2})$.

Historical remarks. The first quantitative bounds of this type were obtained by Berry (1941) and Esseen (1942), who established the optimal rate $O(n^{-1/2})$. Subsequent refinements have been devoted to improving the value of the constant C in (8). For instance, Esseen originally proved the result with $C \leq 7.59$, later improvements were obtained by Shevtsova (2011), who showed that $C \leq 0.4748$ is possible. Thus, in the independent case, the Berry-Esseen theorem is a well understood and fully quantified complement to the CLT.

2.2.2. The dependent case: new difficulties

When dealing with dependent sequences, the situation becomes considerably more delicate. The presence of correlations between observations can strongly affect the speed of convergence. In fact

- The universal bound $O(n^{-1/2})$ is no longer guaranteed. The rate of convergence depends on the strength and structure of the dependence.
- For mixing sequences (e.g., α- or ϕ-mixing), additional conditions are required to control covariances and higher-order cumulants.
- Even when a central limit theorem holds (as in the classical results of Ibragimov and Linnik (1971)), establishing Berry-Esseen type bounds requires delicate inequalities for characteristic functions and careful control of the decay of mixing coefficients.

Known quantitative rates. Several authors have established explicit Berry-Esseen type bounds in the dependent setting

- **Fast mixing (geometric decay).** If the mixing coefficients decay exponentially fast, one can still obtain the optimal $O(n^{-1/2})$ rate. This was shown in works of Utev (1990), Rio (1996) and Rio (2000b), under appropriate moment assumptions.
- **Polynomially decaying mixing.** Suppose the mixing coefficients satisfy a polynomial decay of the form

  $$\alpha(n) \leq Cn^{-\gamma}, \quad \text{for some } \gamma > 0.$$

  Then the Berry-Esseen bound depends on γ:

  $$\sup_{x\in\mathbb{R}} |\mathbb{P}(S_n \leq x) - \Phi(x)| \ \leq \ Cn^{-\delta(\gamma)}, \tag{9}$$

where the exponent $\delta(\gamma)$ satisfies

$$\delta(\gamma) = \frac{\gamma}{2\gamma+1}.$$

For example

$$\begin{aligned} \gamma = 1 \quad &\Rightarrow \quad \delta(1) = \tfrac{1}{3}, \\ \gamma = 2 \quad &\Rightarrow \quad \delta(2) = \tfrac{2}{5}, \\ \gamma \to \infty \quad &\Rightarrow \quad \delta(\gamma) \to \tfrac{1}{2}. \end{aligned}$$

Thus, the optimal $O(n^{-1/2})$ rate is only recovered when the mixing coefficients decay faster than any polynomial.

– **Weak dependence beyond mixing.** Extensions to other dependence structures (e.g., near-epoch dependence, causal processes, or physical dependence measures) have been developed more recently. In these cases, Berry-Esseen bounds often take the form of non-universal rates $O(n^{-\delta})$ with δ depending on dependence coefficients and tail properties of the innovations (see Merlevède, Peligrad and Rio (2009) and Wu and Shao (2004)).

Thus, while the Berry-Esseen theorem provides a universal and sharp rate $O(n^{-1/2})$ in the independent case, in the dependent setting the rate generally depends on

1. the strength of dependence (fast vs. slow mixing);
2. the rate of decay of the mixing coefficients;
3. the moment conditions satisfied by the variables.

In particular, for polynomially mixing sequences with $\alpha(n) = O(n^{-\gamma})$, one typically obtains the non-universal rate $O(n^{-\gamma/(2\gamma+1)})$, which interpolates between $O(n^{-1/3})$ when $\gamma = 1$ and the optimal $O(n^{-1/2})$ when $\gamma \to \infty$.

Recent developments. Beyond mixing frameworks, modern approaches to weak dependence have provided sharper insights into normal approximation. Wu and Shao (2004) studied limit theorems for iterated random functions, while Merlevède, Peligrad and Rio (2009) established Bernstein inequalities and moderate deviations under strong mixing. Other dependence measures, such as physical dependence by Wu (2005), causal processes, and functional dependence by Wu (2011), have also been developed. These contributions illustrate that Berry-Esseen type bounds can be extended well beyond classical mixing, highlighting the generality and importance of our results.

In the next section, we will present a block factorization inequality under ϕ-mixing and state and prove the CLT for stationary ϕ-mixing sequences under polynomial decay. Then, We will derive Berry-Esseen type bounds with explicit convergence rates depending on the decay exponent γ and the moment condition δ.

3. Main Results and Commentaries

This section provides all the details of the most significant results on this topic. We present all the materials used in each proof in a detailed written explanation to make them more understandable to a wider audience.

3.1. Characteristic Function Inequalities

This subsection presents a block factorization inequality under the assumption of a ϕ-mixing condition, providing a detailed proof that uses covariance bounds (the Hoeffding representation and the dominated convergence theorem).

Proposition 1 (Block factorization under ϕ-mixing). *Let $A, B \subset \mathbb{Z}$ be finite index sets with separation h, that is $\min B - \max A \geq h$. Define*

$$S_A = \sum_{i \in A} X_i \text{ and } S_B = \sum_{j \in B} X_j.$$

Then, for every $t \in \mathbb{R}$, we have

$$\left|\varphi_{S_A+S_B}(t) - \varphi_{S_A}(t)\varphi_{S_B}(t)\right| \leq 4\,\phi(h), \tag{10}$$

where φ denotes the characteristic function.

Consequently, if $B_1, \dots, B_M$ are finite pairwise ordered blocks with $\min B_{r+1} - \max B_r \geq h$ for each r, then for every $t \in \mathbb{R}$, we have

$$\left|\varphi_{\sum_{r=1}^{M} S_{B_r}}(t) - \prod_{r=1}^{M} \varphi_{S_{B_r}}(t)\right| \leq 4(M-1)\,\phi(h). \tag{11}$$

Proof. We proceed in careful and elementary steps.

Step 1. Notation and measurability.

Let $m_A := \max A$ and $\ell_B := \min B$. By assumption $\ell_B - m_A \geq h$. Then S_A is measurable with respect to the sigma-field $\mathcal{F}_{-\infty}^{m_A}$ and S_B is measurable with respect to $\mathcal{F}_{\ell_B}^{\infty}$. For a fixed real t, we define complex-valued bounded random variables

$$U := e^{itS_A} \text{ and } V := e^{itS_B}.$$

Clearly U is $\mathcal{F}_{-\infty}^{m_A}$-measurable, V is $\mathcal{F}_{\ell_B}^{\infty}$-measurable, and $\|U\|_\infty = \|V\|_\infty = 1$.

Step 2. Rewriting the difference as a covariance.

By definition of characteristic functions,

$$\varphi_{S_A+S_B}(t) - \varphi_{S_A}(t)\varphi_{S_B}(t) = \mathbb{E}[e^{it(S_A+S_B)}] - \mathbb{E}[e^{itS_A}]\mathbb{E}[e^{itS_B}] = \mathrm{Cov}(U, V).$$

Thus, it suffices to show

$$|\,\mathrm{Cov}(U, V)| \leq 4\,\phi(h).$$

Step 3. Reduction to real-valued covariances.

Let $U = U_1 + iU_2$ and $V = V_1 + iV_2$ where $U_1 = \Re U$, $U_2 = \Im U$, $V_1 = \Re V$, $V_2 = \Im V$. Each U_k, V_ℓ has a real value and $\|U_k\|_\infty \leq 1$, $\|V_\ell\|_\infty \leq 1$. A direct computation yields

$$\operatorname{Cov}(U, V) = \operatorname{Cov}(U_1, V_1) - \operatorname{Cov}(U_2, V_2) + i\big(\operatorname{Cov}(U_1, V_2) + \operatorname{Cov}(U_2, V_1)\big),$$

hence

$$|\operatorname{Cov}(U, V)| \leq |\operatorname{Cov}(U_1, V_1)| + |\operatorname{Cov}(U_2, V_2)| + |\operatorname{Cov}(U_1, V_2)| + |\operatorname{Cov}(U_2, V_1)|. \qquad (*)$$

Therefore, it is enough to prove the following claim for *real-valued* bounded functions.

Claim. If Y is $\mathcal{F}_{-\infty}^{m_A}$-measurable and Z is $\mathcal{F}_{\ell_B}^{\infty}$-measurable with $\|Y\|_\infty \leq 1$ and $\|Z\|_\infty \leq 1$, then

$$|\operatorname{Cov}(Y, Z)| \leq 4\,\phi(h). \qquad (12)$$

We prove the claim in the Steps 4 and 5.

Step 4. Hoeffding identity for bounded variables.

Since Y and Z have real values and are bounded by 1 in absolute value, their ranges are contained in $[-1, 1]$. The following identity (Hoeffding's covariance representation for bounded variables) holds

$$\operatorname{Cov}(Y, Z) = \int_{-1}^{1}\int_{-1}^{1} \Big(\mathbb{P}(Y > y, Z > z) - \mathbb{P}(Y > y)\,\mathbb{P}(Z > z)\Big)\,dy\,dz. \qquad (13)$$

Justification of (13). For completeness, we give a short derivation valid for bounded Y, Z. For $w \in \mathbb{R}$, define the indicator $\mathbf{1}_{Y>w}$. Using the representation, proved in Lemma 3, in the Appendix section,

$$Y = \tfrac{1}{2}\Big(\int_{-1}^{1} \mathbf{1}_{\{Y>y\}}\,dy - \int_{-1}^{1} \mathbf{1}_{\{Y\leq y\}}\,dy\Big) = \int_{-1}^{1} \big(\mathbf{1}_{\{Y>y\}} - \tfrac{1}{2}\big)\,dy,$$

and an analogous representation for Z, expand $\mathbb{E}[YZ]$ and subtract $\mathbb{E}Y\,\mathbb{E}Z$. All cancellation can be rigorously justified by approximating Y, Z by step functions (simple functions), applying the identity for simple functions (finite sums), and passing to the limit using dominated convergence. Since $|\mathbf{1}_{\{Y>y\}}\mathbf{1}_{\{Z>z\}}| \leq 1$ and the integration domain $[-1, 1]^2$ has finite measure, dominated convergence applies. This gives (13).

Step 5. Pointwise bound using the ϕ-mixing property.

Fix $y, z \in [-1, 1]$. The event $A_y := \{Y > y\}$ belongs to $\mathcal{F}_{-\infty}^{m_A}$ and the event $B_z := \{Z > z\}$ belongs to $\mathcal{F}_{\ell_B}^{\infty}$. By definition of the ϕ-mixing coefficient,

$$\phi(h) = \sup_{\substack{A\in\mathcal{F}_{-\infty}^{m_A} \\ \mathbb{P}(A)>0}} \ \sup_{B\in\mathcal{F}_{\ell_B}^{\infty}} \big|\mathbb{P}(B \mid A) - \mathbb{P}(B)\big|.$$

Therefore, for the particular events A_y, B_z we have (if $\mathbb{P}(A_y) = 0$ then trivially $\mathbb{P}(A_y \cap B_z) - \mathbb{P}(A_y)\mathbb{P}(B_z) = 0$)

$$\big|\mathbb{P}(A_y \cap B_z) - \mathbb{P}(A_y)\mathbb{P}(B_z)\big| = \mathbb{P}(A_y)\,\big|\mathbb{P}(B_z \mid A_y) - \mathbb{P}(B_z)\big| \le \mathbb{P}(A_y)\,\phi(h) \le \phi(h),$$

because $\mathbb{P}(A_y) \le 1$. Hence, for every $(y, z) \in [-1, 1]^2$,

$$\big|\mathbb{P}(Y > y, Z > z) - \mathbb{P}(Y > y)\mathbb{P}(Z > z)\big| \le \phi(h).$$

By substituting this pointwise bound into the double integral (13), by taking absolute values, and by applying Fubini dominated convergence (the integrand is bounded by $\phi(h)$, which is integrable over the finite square), we obtain

$$|\operatorname{Cov}(Y, Z)| \le \int_{-1}^{1}\int_{-1}^{1} \phi(h)\, dy\, dz = 4\,\phi(h).$$

This proves the claim (12).

Step 6. Conclusion for complex exponentials.

We decompose $U = \Re U + i\,\Im U =: U_1 + iU_2$ and $V = V_1 + iV_2$, with each $U_k, V_\ell \in [-1, 1]$ and adapted to the separated σ-fields of the blocks A and B. A direct computation gives

$$\operatorname{Cov}(U, V) = \operatorname{Cov}(U_1, V_1) - \operatorname{Cov}(U_2, V_2) + i\big(\operatorname{Cov}(U_1, V_2) + \operatorname{Cov}(U_2, V_1)\big).$$

Hence,

$$|\operatorname{Cov}(U, V)| \;\le\; |\operatorname{Cov}(U_1, V_1)| + |\operatorname{Cov}(U_2, V_2)| + |\operatorname{Cov}(U_1, V_2)| + |\operatorname{Cov}(U_2, V_1)|. \tag{14}$$

Each pair (U_i, V_j) consists of bounded real random variables, so by (12), one has

$$|\operatorname{Cov}(U_i, V_j)| \le 4\,\phi(h).$$

Plugging into (14) gives the crude bound

$$|\operatorname{Cov}(U, V)| \le 16\,\phi(h).$$

Sharpened bound. A more efficient route is to avoid bounding the four real covariances separately. Conditioning in the past σ-field of block A yields

$$\operatorname{Cov}(U, V) = \mathbb{E}\Big[U\big(V - \mathbb{E}V\big)\Big] = \mathbb{E}\Big[U\big(\mathbb{E}[V \mid \mathcal{F}_{-\infty}^{m_A}] - \mathbb{E}V\big)\Big].$$

Therefore,

$$|\operatorname{Cov}(U, V)| \le \|U\|_\infty\,\big\|\mathbb{E}[V \mid \mathcal{F}_{-\infty}^{m_A}] - \mathbb{E}V\big\|_1 \le \big\|\mathbb{E}[V \mid \mathcal{F}_{-\infty}^{m_A}] - \mathbb{E}V\big\|_\infty,$$

since $\|U\|_\infty = 1$.

To bound this sup-norm, we apply Hoeffding's identity (13). For a real random variable $Z \in [-1, 1]$,

$$Z = \int_{-1}^{1} \Big(\mathbf{1}_{\{Z > z\}} - \tfrac{1}{2}\Big)\, dz,$$

whence

$$\mathbb{E}[Z \mid \mathcal{F}] - \mathbb{E}Z = \int_{-1}^{1} \Big(\mathbb{E}[\mathbf{1}_{\{Z>z\}} \mid \mathcal{F}] - \mathbb{P}(Z > z) \Big)\, dz.$$

By definition of $\phi(h)$, the integrand has an absolute value at most $\phi(h)$ for every z. Integrating over an interval of length 2 yields

$$\big\| \mathbb{E}[Z \mid \mathcal{F}_{-\infty}^{m_A}] - \mathbb{E}Z \big\|_\infty \le 2\,\phi(h).$$

By applying this bound with $Z = \Re V$ and $Z = \Im V$, and then using the triangle inequality, we obtain the following result.

$$\big\| \mathbb{E}[V \mid \mathcal{F}_{-\infty}^{m_A}] - \mathbb{E}V \big\|_\infty \le \big\| \mathbb{E}[\Re V \mid \cdot] - \mathbb{E}\Re V \big\|_\infty + \big\| \mathbb{E}[\Im V \mid \cdot] - \mathbb{E}\Im V \big\|_\infty \le 4\,\phi(h).$$

Combining with the covariance representation gives the sharpened inequality

$$|\operatorname{Cov}(U, V)| \le 4\,\phi(h).$$

Final conclusion. Since $U = e^{itS_A}$ and $V = e^{itS_B}$, the characteristic function factorization satisfies

$$\big| \varphi_{S_A + S_B}(t) - \varphi_{S_A}(t)\varphi_{S_B}(t) \big| = |\operatorname{Cov}(U, V)| \le 4\,\phi(h),$$

which is the first inequality of Proposition 1.

Step 7. Iteration over multiple blocks.

Let $B_1, \dots, B_M$ be pairwise ordered blocks with mutual separations at least h. For $m = 1, \dots, M$ set

$$T_m := \sum_{r=1}^{m} S_{B_r},$$

so that in particular $T_M = \sum_{r=1}^{M} S_{B_r}$. We prove the telescoping identity and then bound each summand.

Telescoping identity. The identity

$$\varphi_{T_M}(t) - \prod_{r=1}^{M} \varphi_{S_{B_r}}(t) = \sum_{m=2}^{M} \Big(\varphi_{T_m}(t) - \varphi_{T_{m-1}}(t)\varphi_{S_{B_m}}(t) \Big) \prod_{r=m+1}^{M} \varphi_{S_{B_r}}(t) \tag{15}$$

is obtained by successive addition and subtraction of intermediate products. For completeness, we give a short verification by induction on M.

For $M = 2$, the identity is trivial

$$\varphi_{T_2} - \varphi_{S_{B_1}}\varphi_{S_{B_2}} = \big(\varphi_{T_2} - \varphi_{T_1}\varphi_{S_{B_2}} \big),$$

since $T_1 = S_{B_1}$. Assume (15) holds for $M-1$, then

$$\begin{aligned}\varphi_{T_M} - \prod_{r=1}^{M} \varphi_{S_{B_r}} &= \left(\varphi_{T_M} - \varphi_{T_{M-1}}\varphi_{S_{B_M}}\right) + \varphi_{T_{M-1}}\varphi_{S_{B_M}} - \prod_{r=1}^{M} \varphi_{S_{B_r}} \\ &= \left(\varphi_{T_M} - \varphi_{T_{M-1}}\varphi_{S_{B_M}}\right) + \Big(\varphi_{T_{M-1}} - \prod_{r=1}^{M-1} \varphi_{S_{B_r}}\Big)\varphi_{S_{B_M}},\end{aligned}$$

and the induction hypothesis applied to the second term yields (15) for M. Thus, the telescoping formula holds for all $M \geq 2$.

Representation of each summand as a covariance. We fix $m \in \{2, \dots, M\}$ and we define

$$U_m := e^{itT_{m-1}} \quad \text{and} \quad V_m := e^{itS_{B_m}}.$$

Then

$$\varphi_{T_m}(t) - \varphi_{T_{m-1}}(t)\varphi_{S_{B_m}}(t) = \mathbb{E}\left[e^{itT_m}\right] - \mathbb{E}\left[e^{itT_{m-1}}\right]\mathbb{E}\left[e^{itS_{B_m}}\right] = \mathrm{Cov}(U_m, V_m),$$

because $e^{itT_m} = U_m V_m$. Note that $|U_m| \leq 1$ and $|V_m| \leq 1$.

Measurability and block separation. By definition T_{m-1} is the sum of the blocks $B_1, \dots, B_{m-1}$, hence $U_m = e^{itT_{m-1}}$ is measurable with respect to the σ-field generated by those blocks, which we may denote $\mathcal{F}_{-\infty}^{\max B_{m-1}}$. Likewise V_m is measurable with respect to $\mathcal{F}_{\min B_m}^{\infty}$. By the hypothesis that the blocks are pairwise separated by gaps of length at least h, the separation between these two σ-fields is at least h. Therefore, the two-block bound established in Point 6 applies to the pair (U_m, V_m).

Applying the two-block bound. From Step 6 (the complex-exponential two-block estimate), we have, for every $m = 2, \dots, M$,

$$\left|\varphi_{T_m}(t) - \varphi_{T_{m-1}}(t)\varphi_{S_{B_m}}(t)\right| = |\,\mathrm{Cov}(U_m, V_m)| \leq 4\,\phi(h).$$

Summation and final bound. Insert the estimate above into the telescoping identity (15) and take absolute values. By using the triangle inequality and the fact that for each r, $|\varphi_{S_{B_r}}(t)| \leq 1$, we obtain

$$\begin{aligned}\left|\varphi_{T_M}(t) - \prod_{r=1}^{M} \varphi_{S_{B_r}}(t)\right| &\leq \sum_{m=2}^{M} \left|\varphi_{T_m}(t) - \varphi_{T_{m-1}}(t)\varphi_{S_{B_m}}(t)\right| \left|\prod_{r=m+1}^{M} \varphi_{S_{B_r}}(t)\right| \\ &\leq \sum_{m=2}^{M} 4\,\phi(h) \cdot 1 = 4(M-1)\,\phi(h).\end{aligned}$$

This proves the multi-block inequality

$$\left|\varphi_{T_M}(t) - \prod_{r=1}^{M} \varphi_{S_{B_r}}(t)\right| \leq 4(M-1)\,\phi(h),$$

and completes the proof.

Remark 1. The bound grows linearly in the number of blocks M; when M is large one must therefore control $M\phi(h)$ (for example by letting the gap h grow with M or by grouping blocks) in order to force the right-hand side to zero.

In the next subsection, we will provide a central limit theorem for stationary, polynomial-decaying, ϕ-mixing sequences.

3.2. Central Limit Theorem

In this subsection, we state and prove the CLT for stationary polynomial decaying ϕ-mixing sequences using blocking and approximation by independent blocks.

Theorem 1 (Central Limit Theorem under ϕ-mixing with polynomial decay). *Let $(X_k)_{k\in\mathbb{Z}}$ be a strictly stationary sequence of real-valued random variables satisfying:*

1. $\mathbb{E}[X_0] = 0$.
2. *Finite* $(2+\delta)$*-moment:* $\mathbb{E}[|X_0|^{2+\delta}] < \infty$ *for some* $\delta > 0$.
3. ϕ*-mixing with polynomial decay: there exist constants* $C_\phi > 0$ *and* $\gamma > 0$ *such that*

$$\phi(h) \le C_\phi \, h^{-\gamma}, \quad h \ge 1,$$

   and, additionally, *assume the standard summability condition*

$$\sum_{h=1}^{\infty} \phi(h)^{\delta/(2+\delta)} < \infty.$$

   (For a polynomial bound, this is equivalent to $\gamma > \dfrac{2+\delta}{\delta}$*.)*

4. *The long-run variance is finite and positive:*

$$\sigma^2 := \mathrm{Var}(X_0) + 2\sum_{k=1}^{\infty} \mathrm{Cov}(X_0, X_k).$$

Then, as $n \to \infty$,

$$\frac{S_n}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1) \text{ as } n \to \infty, \quad \text{where } S_n := \sum_{k=1}^{n} X_k.$$

Before we move on to prove Theorem 1, let us establish this useful lemma.

Lemma 1. *(covariance inequality, Ibragimov-type).* Let U (measurable with respect to the past sigma-field) and V (measurable with respect to a future sigma-field at distance h) satisfy $U, V \in L^{2+\delta}$. Then there exists a constant $C = C(\delta, \|U\|_{2+\delta}, \|V\|_{2+\delta})$ such that

$$|\,\mathrm{Cov}(U,V)| \le C\,\phi(h)^{\delta/(2+\delta)}.$$

Proof. (of Lemma 1). Decompose $U = U_b + U_t$ with $U_b = U\mathbf{1}_{\{|U|\le M\}}$, $U_t = U - U_b$ and similarly $V = V_b + V_t$, for a truncation level $M > 0$ to be fixed.

(1) For the bounded parts, one uses the basic bound (from the definition of ϕ):

$$|\operatorname{Cov}(U_b, V_b)| \le 4M^2\phi(h).$$

(2) For the tail parts, we use the H"older and Markov inequality as follows : since $U, V \in L^{2+\delta}$,

$$\|U_t\|_1 \le \|U\|_{2+\delta}\,\mathbb{P}(|U| > M)^{(1+\delta)/(2+\delta)} \le \|U\|_{2+\delta}\frac{\|U\|_{2+\delta}^{1+\delta}}{M^{1+\delta}} = \frac{\|U\|_{2+\delta}^{2+\delta}}{M^{1+\delta}}.$$

Hence (using $|\operatorname{Cov}(U_b, V_t)| \le 2\|U_b\|_\infty\|V_t\|_1$ and analogously for the other tail term)

$$|\operatorname{Cov}(U_b, V_t)| + |\operatorname{Cov}(U_t, V)| \le C_1 M^{-\delta},$$

with C_1 depending only on $\|U\|_{2+\delta}, \|V\|_{2+\delta}$.

By combining the two estimates, we get

$$|\operatorname{Cov}(U, V)| \le 4M^2\phi(h) + C_1 M^{-\delta}.$$

Optimizing in $M > 0$ (choose M with $M^{2+\delta} \asymp \phi(h)^{-1}$) yields the bound

$$|\operatorname{Cov}(U, V)| \le C\,\phi(h)^{\delta/(2+\delta)}.$$

This is the desired inequality.

We can now begin to prove the theorem.

Proof. (of Theorem 1). The proof follows the classical blocking and approximation by the independent blocks method.

Step 1. Variance normalization and absolute summability of covariances.

By stationarity,

$$\operatorname{Var}(S_n) = n\operatorname{Var}(X_0) + 2\sum_{1\le i<j\le n}\operatorname{Cov}(X_i, X_j) = n\operatorname{Var}(X_0) + 2\sum_{k=1}^{n-1}(n-k)\operatorname{Cov}(X_0, X_k).$$

By using the Lemma 1 with $U = X_0$, $V = X_k$, we get

$$|\operatorname{Cov}(X_0, X_k)| \le C\phi(k)^{\delta/(2+\delta)}.$$

By the summability assumption $\sum_{k\ge1}\varphi(k)^{\delta/(2+\delta)} < \infty$, the series $\sum_{k\ge1}|\operatorname{Cov}(X_0, X_k)|$ is finite. Hence

$$\frac{1}{n}\operatorname{Var}(S_n) = \operatorname{Var}(X_0)+2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\operatorname{Cov}(X_0, X_k) \xrightarrow{n\to\infty} \operatorname{Var}(X_0)+2\sum_{k=1}^{\infty}\operatorname{Cov}(X_0, X_k) =: \sigma^2,$$

with $0 < \sigma^2 < \infty$ by hypothesis. This gives the correct variance scaling.

Step 2. Blocking (big blocks and small gaps).

We fix sequences $\ell = \ell_n$ (block length) and $h = h_n$ (gap length) satisfying

$$\ell_n \to \infty, \quad h_n \to \infty, \quad \frac{\ell_n}{n} \to 0, \quad \frac{h_n}{\ell_n} \to 0.$$

(Concrete choices are possible under the polynomial hypothesis on ϕ; e.g. $\ell_n = n^\alpha$ and $h_n = n^\beta$ with $0 < \beta < \alpha < 1$ chosen later). Let us partition $\{1, \ldots, n\}$ into consecutive groups: big blocks of length ℓ separated by small gaps of length h. Let

$$m = m_n = \left\lfloor \frac{n}{\ell + h} \right\rfloor$$

be the number of full big blocks. Denote by $Y_r = \sum_{i \in r\text{-th big block}} X_i$ the r-th big block sum, $r = 1, \ldots, m$. Write

$$S_n = \sum_{r=1}^{m} Y_r + R_n,$$

where R_n is the remainder (contribution of the gaps and the final partial block).

We shall show that R_n is negligible after normalization and that the partial sums of the Y_r behave like a sum of (asymptotically) independent and identically distributed block sums.

Control of the remainder. The number of gap terms is mh and the last partial block has size $\leq \ell + h$. The variance of the remainder can be bounded (using stationarity and absolute summability of covariances) by

$$\mathrm{Var}(R_n) \leq C\big(mh + \ell\big)$$

for a constant C depending on $\mathbb{E}[X_0^2]$ and the covariance sum. Since $m \sim n/(\ell + h)$,

$$\frac{\mathrm{Var}(R_n)}{n} \leq C\left(\frac{mh}{n} + \frac{\ell}{n}\right) = C\left(\frac{h}{\ell + h} + \frac{\ell}{n}\right) \to 0$$

because $h/\ell \to 0$ and $\ell/n \to 0$. Thus

$$\frac{R_n}{\sqrt{n}} \xrightarrow{P} 0 \text{ as } n \to \infty.$$

Step 3. Approximate independence of well-separated blocks.

If two big blocks are separated by at least h indices, then the sigma fields they are measurable with respect to are at distance $\geq h$. For bounded functions f, g one has the simple bound

$$|\operatorname{Cov}(f(Y_r), g(Y_s))| \leq 4\|f\|_\infty \|g\|_\infty \phi(h).$$

Apply this to $f(u) = e^{itu/\sqrt{n}}$ and $g(v) = e^{itv/\sqrt{n}}$, which are bounded by 1. One obtains

$$\big|\operatorname{Cov}\big(e^{itY_r/\sqrt{n}}, e^{itY_s/\sqrt{n}}\big)\big| \leq 4\phi(h).$$

A standard induction argument on characteristic functions (or repeated use of the covariance bound) yields

$$\left|\mathbb{E}\exp\left(it\frac{\sum_{r=1}^{m} Y_r}{\sqrt{n}}\right) - \prod_{r=1}^{m}\mathbb{E}\exp\left(it\frac{Y_r}{\sqrt{n}}\right)\right| \le Cm^2\phi(h),$$

for some absolute constant C. (Intuition: the left-hand side is controlled by the sum of covariances between different factor terms; each such covariance is $O(\phi(h))$ and there are $O(m^2)$ pairs.)

Under our choice of ℓ, h, we will ensure $m^2\phi(h) \to 0$ (possible because $\phi(h) \to 0$ and $m \sim n/\ell$; see remarks below for concrete rate choices under the polynomial assumption). Hence the joint characteristic function of the dependent blocks is close to the product of the marginal characteristic functions (that is, close to the independent case).

Step 4. CLT for independent blocks (Lindeberg condition).

Replace each block Y_r by an independent copy Y_r' having the same distribution as Y_r; call $S_m' = \sum_{r=1}^{m} Y_r'$. By the previous step, the characteristic functions of $\sum_r Y_r$ and $\sum_r Y_r'$ are asymptotically the same. It remains to prove the CLT for S_m' (independent but not identical blocks are actually identically distributed here by stationarity of the construction).

We check the Lindeberg condition for the triangular array $\{Y_r'/\sqrt{n},\ r = 1, \ldots, m\}$.

By using a moment inequality (Rosenthal-type) for sums of ℓ stationary terms with $(2+\delta)$-moment, one has

$$\mathbb{E}|Y_1|^{2+\delta} \le C'\ell^{1+\delta/2}$$

for a constant C' depending on the moments of X_0. Then for any $\varepsilon > 0$,

$$\mathbb{E}\left[Y_1^2\mathbf{1}_{\{|Y_1|>\varepsilon\sqrt{n}\}}\right] \le \frac{\mathbb{E}|Y_1|^{2+\delta}}{(\varepsilon\sqrt{n})^{\delta}} \le \frac{C'\,\ell^{1+\delta/2}}{\varepsilon^{\delta}n^{\delta/2}}.$$

Hence, the Lindeberg sum satisfies

$$\frac{1}{\mathrm{Var}(S_n)}\sum_{r=1}^{m}\mathbb{E}\left[Y_r'^2\mathbf{1}_{\{|Y_r'|>\varepsilon\sqrt{n}\}}\right] \le \frac{m\cdot C'\ell^{1+\delta/2}}{\varepsilon^{\delta}n^{\delta/2}\,\mathrm{Var}(S_n)}.$$

Since $\mathrm{Var}(S_n) \asymp n$ and $m \asymp n/(\ell+h) \asymp n/\ell$, this ratio is bounded by a constant times

$$\frac{(n/\ell)\ell^{1+\delta/2}}{n^{1+\delta/2}} = \left(\frac{\ell}{n}\right)^{\delta/2},$$

which $\to 0$ because $\ell/n \to 0$. So the Lindeberg condition holds, and by the classical Lindeberg–Feller CLT for triangular arrays of independent variables,

$$\frac{S_m'}{\sqrt{\mathrm{Var}(S_m')}} \xrightarrow{d} \mathcal{N}(0,1).$$

It remains to check that $\mathrm{Var}(S'_m) = \sum_{r=1}^{m} \mathrm{Var}(Y'_r) \sim \sigma^2 n$ (so that normalizing by $\sqrt{n}$ yields the desired variance σ^2). But

$$\sum_{r=1}^{m} \mathrm{Var}(Y'_r) = m\,\mathrm{Var}(Y_1)$$

and a direct computation using stationarity shows

$$\mathrm{Var}(Y_1) = \ell\,\mathrm{Var}(X_0) + 2\sum_{k=1}^{\ell-1}(\ell-k)\,\mathrm{Cov}(X_0, X_k).$$

Hence

$$\frac{m\,\mathrm{Var}(Y_1)}{n} = \frac{m\ell}{n}\mathrm{Var}(X_0) + \frac{2m}{n}\sum_{k=1}^{\ell-1}(\ell-k)\,\mathrm{Cov}(X_0, X_k).$$

Since $m\ell/n \to 1$ (gaps are small: $h/\ell \to 0$) and $\sum_{k\geq 1} |\,\mathrm{Cov}(X_0, X_k)| < \infty$, a dominated convergence argument yields

$$\frac{m\,\mathrm{Var}(Y_1)}{n} \xrightarrow{n\to\infty} \sigma^2.$$

Thus, $\mathrm{Var}(S'_m) \sim \sigma^2 n$ and the CLT for independent blocks yields

$$\frac{S'_m}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1) \text{ as } n \to \infty.$$

Step 5. Transfer from blocks to the full sum.

Collecting the estimates

- The difference between the characteristic functions of $\sum_r Y_r$ and $\sum_r Y'_r$ is $o(1)$ (Step 3).
- The independent sum $\sum_r Y'_r$ satisfies the CLT after normalization (Step 4).
- The remainder R_n is negligible: $R_n/\sqrt{n} \xrightarrow{P} 0$ (Step 2).

Therefore

$$\frac{S_n}{\sigma\sqrt{n}} = \frac{\sum_{r=1}^{m} Y_r}{\sigma\sqrt{n}} + o_P(1)$$

and the first term converges in distribution to $\mathcal{N}(0,1)$. This completes the proof.

Remark 2. The crucial technical assumptions were the $(2+\delta)$-moment and the summability $\sum_h \phi(h)^{\delta/(2+\delta)} < \infty$. For polynomial decay $\phi(h) \lesssim h^{-\gamma}$, this requires $\gamma > \dfrac{2+\delta}{\delta}$.

The following subsection will see us state Berry-Esseen type bounds and derive the best possible rate of normality approximation.

3.3. Rate of normality approximation

In this subsection, we present Berry-Esseen type bounds with explicit convergence rates depending on the decay exponent γ and the moment condition δ. We then derive corollaries demonstrating the best achievable rate.

Theorem 2 (Berry-Esseen bound for ϕ-mixing sequences via blocking). *Let $(X_k)_{k\in\mathbb{Z}}$ be a strictly stationary sequence of real-valued random variables with*

$$\mathbb{E}[X_0] = 0 \text{ and } \mathbb{E}|X_0|^{2+\delta} < \infty$$

for some $\delta > 0$, and assume that the sequence is ϕ-mixing with coefficients $\phi(h)$ satisfying

$$\phi(h) \le C_\phi\, h^{-\gamma}, \quad \text{for some } \gamma > 0.$$

Let $S_n = \sum_{k=1}^n X_k$ and $F_n(x) = \mathbb{P}(S_n/(\sigma\sqrt{n}) \le x)$, where $\sigma^2 = \mathrm{Var}(X_0) + 2\sum_{k\ge 1}\mathrm{Cov}(X_0, X_k)$. Then, for the block, gap, and truncation parameters

$$\ell = \lfloor n^{2/3} \rfloor, \quad h = \lfloor n^{\beta} \rfloor, \quad T = \lfloor n^{1/6} \rfloor,$$

with $0 < \beta < 2/3$, there exists a constant $C > 0$ (depending only on δ, γ, C_ϕ, σ, and $\mathbb{E}|X_0|^{2+\delta}$) such that

$$\sup_{x\in\mathbb{R}} \left|F_n(x) - \Phi(x)\right| \le C\Big(M\,\phi(h)^{\theta} \;+\; T^{1+\delta}(M\ell)^{-\delta/2} \;+\; T\,\frac{Mh+\ell}{n} \;+\; T\sqrt{\frac{Mh+\ell}{n}} \;+\; \frac{1}{T}\Big), \tag{16}$$

where

$$M = \left\lfloor \frac{n}{\ell + h} \right\rfloor, \quad \theta = \frac{\delta}{2+\delta},$$

and Φ denotes the standard normal distribution function.

Corollary 1 (Rate-optimal Berry-Esseen bound). *Under the assumptions of Theorem 2, if $\delta > 1$ and the parameters are chosen with $0 < \beta < 2/3$ such that*

$$\beta\gamma\theta \;\ge\; \tfrac{1}{2}, \quad \theta = \frac{\delta}{2+\delta},$$

then the normalized partial sums satisfy

$$\sup_{x\in\mathbb{R}} \left|F_n(x) - \Phi(x)\right| \;=\; O\Big(n^{-1/6}\Big).$$

Before we go any further, let us provide the following technical lemma based on truncation, interpolation and the small-t factor.

This lemma provides the strengthened two-block covariance bound used in Step 4: for small t the covariance between the exponentials of the separated block sums carries an extra factor $|t|$ (up to controlled remainder terms depending on the moments). The proof uses truncation of the sums of the blocks and interpolation (Hölder) to trade dependence for a power of the mixing coefficient.

Lemma 2 (Truncation and interpolation small-t bound). *Let $(X_k)_{k\in\mathbb{Z}}$ be strictly stationary with $\mathbb{E}[X_0] = 0$ and $\mathbb{E}|X_0|^{2+\delta} < \infty$ for some $\delta \in]0,1]$. Fix two finite index blocks $A, B \subset \mathbb{Z}$ with separation $\operatorname{dist}(A,B) \geq h$. Let*

$$Y_A := \sum_{i\in A} X_i \text{ and } Y_B := \sum_{j\in B} X_j,$$

and set, for normalization parameter $u > 0$ and $t \in \mathbb{R}$,

$$U(t) := e^{itY_A/u} \text{ and } V(t) := e^{itY_B/u}.$$

Assume the ϕ-mixing coefficients satisfy $\phi(h) \to 0$. Define $\theta = \delta/(2+\delta)$. Then there exist constants $C_1, C_2 > 0$ depending only on δ and $\|X_0\|_{2+\delta}$ such that for every truncation level $a > 0$ and every t with $|t| \leq u$,

$$\big|\operatorname{Cov}(U(t), V(t))\big| \leq C_1 \frac{|t|}{u}\,\phi(h)^\theta \;+\; C_2 \left(\frac{|t|}{u}\right)^{1+\delta} \frac{\mathbb{E}\big[|Y_A|^{2+\delta} + |Y_B|^{2+\delta}\big]}{a^\delta}. \tag{17}$$

In particular, choosing a of order $\left(\mathbb{E}[|Y_A|^{2+\delta} + |Y_B|^{2+\delta}]\right)^{1/(1+\delta)} (u/|t|)$ yields the simplified bound

$$\big|\operatorname{Cov}(U(t), V(t))\big| \leq C\,\frac{|t|}{u}\,\phi(h)^\theta \;+\; C' \left(\frac{|t|}{u}\right)^{1+\delta} (\mathbb{E}|Y_A|^{2+\delta} + \mathbb{E}|Y_B|^{2+\delta})^{\frac{\delta}{1+\delta}}, \tag{18}$$

where C and C' depend only on δ, $\|X_0\|_{2+\delta}$.

Proof. (of Lemma 2). Let us fix truncation level $a > 0$. We note the truncated block sums

$$Y_A^{(a)} := \sum_{i\in A} X_i \mathbf{1}_{\{|X_i|\leq a\}} \text{ and } Y_B^{(a)} := \sum_{j\in B} X_j \mathbf{1}_{\{|X_j|\leq a\}}.$$

We decompose $U(t) = U^{(a)}(t) + R_A(t)$ and $V(t) = V^{(a)}(t) + R_B(t)$, where

$$U^{(a)}(t) := e^{itY_A^{(a)}/u}, \qquad R_A(t) := e^{itY_A/u} - e^{itY_A^{(a)}/u},$$

and similarly for V. Then

$$\operatorname{Cov}(U,V) = \operatorname{Cov}\big(U^{(a)}, V^{(a)}\big) + \operatorname{Cov}\big(R_A, V^{(a)}\big) + \operatorname{Cov}\big(U^{(a)}, R_B\big) + \operatorname{Cov}\big(R_A, R_B\big).$$

We bound each term as follows

a. Covariance of truncated exponentials (main dependence term). The truncated sums $Y_A^{(a)}$ and $Y_B^{(a)}$ are measurable with respect to the separated sigma-fields and are bounded in absolute value by $|A|a$ and $|B|a$. Consequently $U^{(a)}$ and $V^{(a)}$ are bounded by 1 and are functions of bounded variables. Using a standard ϕ-mixing interpolation inequality (Ibragimov-type bound with moment interpolation, see Bradley (2005) or Ibragimov (1962)), there exists $C_0 = C_0(\delta, \|X_0\|_{2+\delta})$ such that

$$\big|\operatorname{Cov}(U^{(a)}, V^{(a)})\big| \leq C_0\,\|U^{(a)}\|_p\,\|V^{(a)}\|_q\,\phi(h)^{1-1/p-1/q}$$

for any choice $p, q \in [1, \infty]$ with $1/p + 1/q < 1$. Choose $p = q = 2+\delta$, then $1 - 1/p - 1/q = \delta/(2+\delta) =: \theta$. Moreover $\|U^{(a)}\|_{2+\delta} \leq \left(\mathbb{E}|U^{(a)}|^{2+\delta}\right)^{1/(2+\delta)} = 1$ (since $|U^{(a)}| \leq 1$), and similarly for $V^{(a)}$. Thus

$$\left|\mathrm{Cov}(U^{(a)}, V^{(a)})\right| \leq C_0\, \phi(h)^{\theta}.$$

To obtain the extra small-t factor, we refine the preceding estimate using the identity

$$U^{(a)}(t) - 1 = \frac{it}{u} \int_0^1 e^{itsY_A^{(a)}/u} Y_A^{(a)}\, ds,$$

and similarly for $V^{(a)}$. Expanding $\mathrm{Cov}(U^{(a)}, V^{(a)})$ and applying Hölder inequality with the same $p = q = 2 + \delta$ yields

$$\left|\mathrm{Cov}(U^{(a)}, V^{(a)})\right| \leq \frac{|t|}{u}\, C_0'\, \phi(h)^{\theta} \left(\mathbb{E}|Y_A^{(a)}|^{2+\delta}\right)^{1/(2+\delta)} \left(\mathbb{E}|Y_B^{(a)}|^{2+\delta}\right)^{1/(2+\delta)}.$$

Since $|Y_A^{(a)}| \leq |A|a$ and $\mathbb{E}|Y_A^{(a)}|^{2+\delta} \leq \mathbb{E}|Y_A|^{2+\delta}$, the factors involving the truncated moments are bounded in terms of $\mathbb{E}|Y_A|^{2+\delta}, \mathbb{E}|Y_B|^{2+\delta}$. Absorbing these into the constant gives the convenient form

$$\left|\mathrm{Cov}(U^{(a)}, V^{(a)})\right| \leq C_1 \frac{|t|}{u}\, \phi(h)^{\theta},$$

where $C_1 = C_1(\delta, \|X_0\|_{2+\delta}, |A|, |B|, a)$. But as a will be chosen depending on moments, one may bound $|A|, |B|$ and a dependence into the constant ultimately.

b. Tail error terms. We now bound $\mathrm{Cov}(R_A, V^{(a)})$, $\mathrm{Cov}(U^{(a)}, R_B)$ and $\mathrm{Cov}(R_A, R_B)$. By using the trivial bound $|\mathrm{Cov}(X, Y)| \leq 2\|X\|_1 \|Y\|_{\infty}$ and $\|V^{(a)}\|_{\infty} \leq 1$, we get

$$|\mathrm{Cov}(R_A, V^{(a)})| \leq 2\, \mathbb{E}|R_A|.$$

Let us control $\mathbb{E}|R_A|$. By using the identity

$$R_A(t) = e^{itY_A/u}\left(1 - e^{-it(Y_A - Y_A^{(a)})/u}\right),$$

and $|1 - e^{iz}| \leq |z|$, we obtain

$$|R_A(t)| \leq \frac{|t|}{u} \cdot |Y_A - Y_A^{(a)}|.$$

Therefore

$$\mathbb{E}|R_A| \leq \frac{|t|}{u}\, \mathbb{E}\big[|Y_A - Y_A^{(a)}|\big].$$

By applying Hölder inequality with exponents $(2+\delta)/(1+\delta)$ and $(2+\delta)$, we get

$$\mathbb{E}\big[|Y_A - Y_A^{(a)}|\big] \leq \left(\mathbb{E}|Y_A - Y_A^{(a)}|^{2+\delta}\right)^{\frac{1}{2+\delta}} \cdot \left(\mathbb{P}(|Y_A - Y_A^{(a)}| > 0)\right)^{\frac{1+\delta}{2+\delta}}.$$

A simple tail estimate using truncation at the single-observation level yields

$$\mathbb{E}\big[|Y_A - Y_A^{(a)}|\big] \leq \frac{\mathbb{E}|Y_A|^{2+\delta}}{a^{1+\delta}},$$

up to a multiplicative constant depending only on $delta$. (Indeed, $Y_A - Y_A^{(a)} = \sum_{i\in A} X_i \mathbf{1}_{\{|X_i|>a\}}$; apply Hölder inequality and union bound to each summand). Hence

$$\mathbb{E}|R_A| \le \frac{|t|}{u} \cdot C_\delta \frac{\mathbb{E}|Y_A|^{2+\delta}}{a^{1+\delta}}.$$

By the same argument for R_B and for $\mathrm{Cov}(R_A, R_B)$, we obtain that all tail-covariance contributions are bounded by a constant times

$$\frac{|t|}{u} \cdot \frac{\mathbb{E}\big[|Y_A|^{2+\delta} + |Y_B|^{2+\delta}\big]}{a^{1+\delta}}.$$

combining with the previous prefactor $|t|/u$ yields a bound of order $(|t|/u)^{1+\delta} a^{-(1+\delta)} \mathbb{E}|Y|^{2+\delta}$. Rewriting $a^{-(1+\delta)} = (a^{-\delta}) \cdot a^{-1}$ and absorbing a^{-1} into constants yields the asserted tail term in (17), namely of order

$$C_2 \left(\frac{|t|}{u}\right)^{1+\delta} \frac{\mathbb{E}\big[|Y_A|^{2+\delta} + |Y_B|^{2+\delta}\big]}{a^{\delta}}.$$

c. Collecting terms and optimization remark. By adding the truncated-covariance bound and the tail bounds, we arrive at (17):

$$\big|\mathrm{Cov}(U,V)\big| \le C_1 \frac{|t|}{u}\, \phi(h)^\theta \;+\; C_2 \left(\frac{|t|}{u}\right)^{1+\delta} \frac{\mathbb{E}\big[|Y_A|^{2+\delta} + |Y_B|^{2+\delta}\big]}{a^{\delta}}.$$

The displayed second inequality in the Lemma 2 is obtained by choosing a to balance the two terms (for example $a \asymp (\mathbb{E}|Y|^{2+\delta})^{1/(1+\delta)}(u/|t|)$), which yields the cleaner bound with a single $|t|/u$ main term plus a higher-order remainder in $(|t|/u)^{1+\delta}$. This completes the proof of Lemma 2.

The preceding lemma provides the key small-t covariance refinement that will be invoked in the proof of Theorem 2, ensuring integrability of the dependence term in Esseen's inequality.

Proof: Full and detailed proof of Theorem 2

We prove the master inequality stated in Theorem 2, namely that for the choices

$$\ell = \lfloor n^{2/3} \rfloor, \quad h = \lfloor n^{\beta} \rfloor, \quad T = \lfloor n^{1/6} \rfloor,$$

there exists $C > 0$ (depending only on $\delta, \gamma, C_\phi, \sigma, \mathbb{E}|X_0|^{2+\delta}$) such that

$$\sup_{x\in\mathbb{R}} \big|F_n(x) - \Phi(x)\big| \le C\Big(M\,\phi(h)^\theta \;+\; T^{1+\delta}(M\ell)^{-\delta/2} \;+\; T\frac{Mh+\ell}{n} \;+\; T\sqrt{\frac{Mh+\ell}{n}} \;+\; \frac{1}{T}\Big),$$

where $M = \big\lfloor n/(\ell+h)\big\rfloor$ and $\theta = \delta/(2+\delta)$.

Step 1. Preliminaries and normalization

Write

$$Z_n := \frac{S_n}{\sigma\sqrt{n}}, \quad \varphi_n(t) := \mathbb{E}\big[e^{itZ_n}\big],$$

and let Φ denote the standard normal cumulative distribution function with characteristic function $e^{-t^2/2}$. By Esseen's smoothing inequality (standard form), for any $T > 0$, there exists universal constant C_E such that

$$\sup_{x\in\mathbb{R}} \big|F_n(x) - \Phi(x)\big| \le C_E\Big(\int_{-T}^{T} \frac{\big|\varphi_n(t) - e^{-t^2/2}\big|}{|t|}\,dt + \frac{1}{T}\Big). \tag{19}$$

Hence, it suffices to bound the integral

$$I(T) := \int_{-T}^{T} \frac{\big|\varphi_n(t) - e^{-t^2/2}\big|}{|t|}\,dt.$$

Step 2. Blocking decomposition

Let us fix integers $\ell \ge 1$ (block length) and gap length $h \ge 1$. Let us partition the index set $\{1, \dots, n\}$ into alternating big blocks of length ℓ and gaps of length h (possibly with a final short remainder). More precisely, let B_r denote the r-th big block of indices (of cardinality ℓ except possibly the last), $r = 1, \dots, M$, with $M = \lfloor n/(\ell + h)\rfloor$. Denote by

$$Y_r := \sum_{i\in B_r} X_i \qquad (r = 1, \dots, M)$$

the big-block sums. Let R be the remainder comprising all gap indices and the short tail (so $S_n = T_M + R$ where $T_M := \sum_{r=1}^{M} Y_r$).

By working with normalized variables, we define

$$\psi_{T_M}(t) := \mathbb{E}\big[e^{itT_M/(\sigma\sqrt{n})}\big] \text{ and } \psi_{Y_r}(t) := \mathbb{E}\big[e^{itY_r/(\sigma\sqrt{n})}\big].$$

Then the integrand can be split as

$$\begin{aligned}\big|\varphi_n(t) - e^{-t^2/2}\big| \;\le\; & \big|\varphi_n(t) - \psi_{T_M}(t)\big| + \Big|\psi_{T_M}(t) - \prod_{r=1}^{M}\psi_{Y_r}(t)\Big| \\ & + \Big|\prod_{r=1}^{M}\psi_{Y_r}(t) - e^{-t^2\mathrm{Var}(T_M)/2\sigma^2}\Big| \\ & + \big|e^{-t^2\mathrm{Var}(T_M)/2\sigma^2} - e^{-t^2/2}\big|. \end{aligned} \tag{20}$$

We will denote the four contributions by $A(t), B(t), C(t), D(t)$ respectively:

$$A(t) := \big|\varphi_n(t) - \psi_{T_M}(t)\big|, \quad B(t) := \Big|\psi_{T_M}(t) - \prod_{r=1}^{M}\psi_{Y_r}(t)\Big|,$$

$$C(t) := \Big| \prod_{r=1}^{M} \psi_{Y_r}(t) - e^{-t^2 \mathrm{Var}(T_M)/2\sigma^2} \Big|, \text{ and } D(t) := \big| e^{-t^2 \mathrm{Var}(T_M)/2\sigma^2} - e^{-t^2/2} \big|.$$

We bound the contributions of each term to the Esseen integral separately:

$$I(T) \le \sum_{J \in \{A,B,C,D\}} \int_{-T}^{T} \frac{J(t)}{|t|}\, dt.$$

Step 3. Bound for the remainder gaps term A

We use the fact $|e^{iu} - e^{iv}| \le |u - v|$. By noting $S_n = T_M + R$ and by normalizing, we have

$$A(t) = \big| \mathbb{E}[e^{it(S_n/(\sigma\sqrt{n}))}] - \mathbb{E}[e^{it(T_M/(\sigma\sqrt{n}))}] \big| \le \frac{|t|}{\sigma\sqrt{n}}\, \mathbb{E}\big|R\big|.$$

Therefore

$$\int_{-T}^{T} \frac{A(t)}{|t|}\, dt \le \int_{-T}^{T} \frac{1}{\sigma\sqrt{n}}\, \mathbb{E}|R|\, dt = \frac{2T}{\sigma\sqrt{n}}\, \mathbb{E}|R|.$$

We bound $\mathbb{E}|R|$ by Cauchy–Schwarz and a variance control. Since R contains at most $Mh + \ell$ summands (the gaps plus a final short block), stationarity and summation give

$$\mathrm{Var}(R) \le C_1(Mh + \ell)$$

for some constant C_1 depending only on covariances of the process (and ultimately on σ^2). Hence

$$\mathbb{E}|R| \le \sqrt{\mathrm{Var}(R)} \le C_2\sqrt{Mh + \ell}.$$

By putting together, we obtain

$$\int_{-T}^{T} \frac{A(t)}{|t|}\, dt \le C\, T \sqrt{\frac{Mh + \ell}{n}}.$$

This yields the fourth term in (3.3) after renaming constants.

Step 4. Bound for the block-dependence term B

This is the contribution coming from the difference between the characteristic function of the sum of big blocks and the product of the characteristic functions of the block. We use a ϕ-mixing covariance inequality (Ibragimov type) to control dependence between separated blocks.

4.1 Two-block covariance bound. Recall that for two blocks A and B separated by at least h, setting $U = e^{itS_A/(\sigma\sqrt{n})}$, $V = e^{itS_B/(\sigma\sqrt{n})}$ we have (by boundedness $\|U\|_\infty = \|V\|_\infty = 1$ and a ϕ-mixing covariance bound)

$$\big| \mathbb{E}[UV] - \mathbb{E}U\, \mathbb{E}V \big| \le C_3\, \phi(h).$$

For a finer L^p–L^q inequality which exploits the finite $(2+\delta)$–moment, truncation and interpolation yield

$$\big|\mathbb{E}[UV] - \mathbb{E}U\,\mathbb{E}V\big| \le C_4\,\phi(h)^{\theta}, \quad \theta = \frac{\delta}{2+\delta}, \tag{21}$$

where the constant C_4 depends on δ and $\|X_0\|_{2+\delta}$. (Intuitively this comes from truncating X_i at level T and applying ϕ control on indicators and controlling tails by moments; the power θ is the usual interpolation exponent.)

4.2 Multi-block telescoping. Apply the two-block bound iteratively via the standard telescoping expansion (as in Proposition 1 but with ϕ). A direct telescoping yields

$$\Big|\psi_{T_M}(t) - \prod_{r=1}^{M}\psi_{Y_r}(t)\Big| \le \sum_{m=2}^{M}\Big|\mathbb{E}\big[e^{it(\sum_{r=1}^{m}Y_r)/(\sigma\sqrt{n})}\big] - \mathbb{E}\big[e^{it(\sum_{r=1}^{m-1}Y_r)/(\sigma\sqrt{n})}\big]\mathbb{E}\big[e^{itY_m/(\sigma\sqrt{n})}\big]\Big|.$$

Each difference in the sum is bounded by $C_4\,\phi(h)^{\theta}$ uniformly in t (the boundedness of the exponentials ensures the bound does not explode for small/large t). Hence

$$B(t) \le C_4(M-1)\,\phi(h)^{\theta} \le C_5 M\,\phi(h)^{\theta}.$$

4.3 Integrating $B(t)$. At first sight directly integrating $B(t)/|t|$ on $[-T,T]$ would produce a divergent $\int_{-T}^{T}\frac{dt}{|t|}$. However, observe that the two-block bound in fact improves for small t because the characteristic functions are close to 1 and the covariance can be bounded by the product of the smallness of t and dependence coefficient. Concretely, by expanding exponentials or using a coupling/truncation argument one obtains for $|t| \le T$ the strengthened bound

$$\big|\mathbb{E}[UV] - \mathbb{E}U\,\mathbb{E}V\big| \le C_6\,|t|\,\phi(h)^{\theta}.$$

Note that one can derive such an extra factor $|t|$ by using the identity $e^{iu}-1 = iu\int_0^1 e^{ius}\,ds$ and then applying the $L^{2+\delta}$ tail control on block sums; this is the standard smoothing that makes the Esseen integral finite. By Lemma 2, for $|t| \le T$ the covariance between exponentials of separated block sums gains an extra factor $|t|$. Iterating this strengthened bound across the telescoping expansion yields

$$B(t) \le C_7\,|t|\,M\,\phi(h)^{\theta}.$$

Therefore

$$\int_{-T}^{T}\frac{B(t)}{|t|}\,dt \le \int_{-T}^{T} C_7 M\,\phi(h)^{\theta}\,dt = 2C_7\,T\,M\,\phi(h)^{\theta}.$$

This produces the first term $M\phi(h)^{\theta}$ (multiplied by T which is later chosen) present in (3.3).

Step 5. Bound for the i.i.d.-blocks vs Gaussian term C

This is the usual Lyapunov cumulant control: the product of the block characteristic functions is close to the Gaussian characteristic function with an error controlled by the $(2+\delta)$-moments of block sums.

5.1 Local Taylor cumulant expansion for each block. Let Y_r be a block sum of length ℓ. For $|t| \leq T$ and normalized by $\sigma\sqrt{n}$, we have the expansion

$$\psi_{Y_r}(t) = 1 - \frac{t^2}{2}\frac{\mathrm{Var}(Y_r)}{\sigma^2 n} + R_r(t),$$

where the remainder satisfies, by Taylor and generalized Lyapunov,

$$|R_r(t)| \leq C_8\,|t|^{2+\delta}\,\frac{\mathbb{E}|Y_r|^{2+\delta}}{(\sigma^2 n)^{1+\delta/2}}.$$

By summing over $r = 1, \dots, M$ and by exponentiating the logarithm of the product, one obtains

$$\prod_{r=1}^{M}\psi_{Y_r}(t) = \exp\Big(-\frac{t^2}{2}\cdot\frac{\sum_r \mathrm{Var}(Y_r)}{\sigma^2 n} + O\Big(|t|^{2+\delta}\frac{\sum_r \mathbb{E}|Y_r|^{2+\delta}}{(\sigma^2 n)^{1+\delta/2}}\Big)\Big).$$

Hence,

$$\Big|\prod_{r=1}^{M}\psi_{Y_r}(t) - e^{-t^2\mathrm{Var}(T_M)/2\sigma^2}\Big| \leq C_9\,|t|^{2+\delta}\,\frac{\sum_{r=1}^{M}\mathbb{E}|Y_r|^{2+\delta}}{(\sigma^2 n)^{1+\delta/2}}.$$

5.2 Moment bound for block sums. By stationarity and the $(2+\delta)$–moment assumption, we have

$$\mathbb{E}|Y_r|^{2+\delta} \leq C_{10}\,\ell^{1+\delta/2}.$$

Therefore

$$\sum_{r=1}^{M}\mathbb{E}|Y_r|^{2+\delta} \leq M\cdot C_{10}\,\ell^{1+\delta/2}.$$

Since $M\ell \asymp n$ (the large blocks cover most of the n observations), we have $(\sigma^2 n)^{1+\delta/2} \asymp (M\ell)^{1+\delta/2}$. By combining, we get

$$C(t) \leq C_{11}\,|t|^{2+\delta}\,(M\ell)^{-\delta/2}.$$

5.3 Integrating $C(t)$. Thus

$$\int_{-T}^{T}\frac{C(t)}{|t|}\,dt \leq C_{11}\,(M\ell)^{-\delta/2}\int_{-T}^{T}|t|^{1+\delta}\,dt = 2C_{11}\,(M\ell)^{-\delta/2}\cdot\frac{T^{2+\delta}}{2+\delta}.$$

Hence

$$\int_{-T}^{T}\frac{C(t)}{|t|}\,dt \leq C_{12}\,T^{1+\delta}\,(M\ell)^{-\delta/2},$$

which is the second term appearing in (3.3).

Step 6. Bound for the variance mismatch term D

We have

$$D(t) = \left|e^{-t^2 \mathrm{Var}(T_M)/2\sigma^2} - e^{-t^2/2}\right| \leq \frac{t^2}{2}\left|\frac{\mathrm{Var}(T_M)}{\sigma^2} - 1\right| e^{-\xi t^2/2}$$

for some $\xi \in (\min\{\mathrm{Var}(T_M)/\sigma^2, 1\}, \max\{\mathrm{Var}(T_M)/\sigma^2, 1\})$ (Mean Value Theorem); discarding the exponential factor (bounded by 1) produces the crude bound

$$D(t) \leq \frac{t^2}{2}\left|\frac{\mathrm{Var}(T_M) - \sigma^2 n}{\sigma^2 n}\right|.$$

Now $\mathrm{Var}(T_M)$ differs from $\sigma^2 n$ in terms of the contribution of gaps and edge effects. A direct variance decomposition gives

$$\left|\mathrm{Var}(T_M) - \sigma^2 n\right| \leq C_{13}\,(Mh + \ell),$$

hence

$$D(t) \leq C_{14}\, t^2 \frac{Mh + \ell}{n}.$$

By integrating, we get

$$\int_{-T}^{T} \frac{D(t)}{|t|}\, dt \leq C_{14} \frac{Mh + \ell}{n} \int_{-T}^{T} |t|\, dt = C_{14} \frac{Mh + \ell}{n} \cdot T^2.$$

Since T will be chosen as $T \asymp n^{1/6}$, the combination $T^2 \cdot (Mh+\ell)/n$ is of the same order as $T \cdot (Mh+\ell)/n$ (absorbing constants) in the final expression. For bookkeeping, we record the bound as

$$\int_{-T}^{T} \frac{D(t)}{|t|}\, dt \leq C_{15}\, T \cdot \frac{Mh + \ell}{n}.$$

Step 7. Collecting all terms

Summing the results of Steps 3 to 6 and inserting them into Esseen's inequality (19) yields, for some constant C,

$$\sup_x |F_n(x) - \Phi(x)| \leq C\Big(TM\phi(h)^{\theta} + T^{1+\delta}(M\ell)^{-\delta/2} + T\frac{Mh + \ell}{n} + T\sqrt{\frac{Mh + \ell}{n}} + \frac{1}{T}\Big),$$

which is the bound shown (3.3) up to renaming the multiplicative constants. Absorbing the visible T into the constant gives the form stated in the main text:

$$\sup_{x \in \mathbb{R}} \left|F_n(x) - \Phi(x)\right| \leq C\Big(M\,\phi(h)^{\theta} + T^{1+\delta}(M\ell)^{-\delta/2} + T\frac{Mh + \ell}{n} + T\sqrt{\frac{Mh + \ell}{n}} + \frac{1}{T}\Big). \tag{22}$$

Step 8. Choice of parameters and asymptotics (calculation)

By choosing

$$\ell = \lfloor n^{2/3} \rfloor, \quad T = \lfloor n^{1/6} \rfloor, \quad h = \lfloor n^{\beta} \rfloor, \quad 0 < \beta < \tfrac{2}{3},$$

then

$$M = \left\lfloor \frac{n}{\ell + h} \right\rfloor \asymp \frac{n}{\ell} \asymp n^{1/3}.$$

We evaluate the master bound term by term in (22).

(1) Dependence term. Assume $\phi(h) \asymp h^{-\gamma}$ and recall $\theta = \dfrac{\delta}{2+\delta}$. Then

$$M\phi(h)^{\theta} \asymp n^{1/3} \cdot (n^{\beta})^{-\gamma\theta} = n^{1/3-\beta\gamma\theta}.$$

(2) Lyapunov-type term (correct algebra). The bound contains the factor $T^{1+\delta}(M\ell)^{-\delta/2}$. Since $M\ell \asymp n^{1/3} \cdot n^{2/3} = n$,

$$(M\ell)^{-\delta/2} \asymp n^{-\delta/2},$$

and $T^{1+\delta} \asymp (n^{1/6})^{1+\delta} = n^{(1+\delta)/6}$. Hence

$$T^{1+\delta}(M\ell)^{-\delta/2} \asymp n^{\frac{1+\delta}{6}} \cdot n^{-\frac{\delta}{2}} = n^{\frac{1+\delta-3\delta}{6}} = n^{\frac{1-2\delta}{6}}.$$

Comment. This term tends to 0 provided $\delta > \frac{1}{2}$; it is $O(n^{-1/6})$ precisely when $\dfrac{1-2\delta}{6} \leq -\dfrac{1}{6}$, that is, when $\delta \geq 1$.

(3) First remainder term.

$$T\frac{Mh+\ell}{n} \asymp n^{1/6} \cdot \frac{n^{1/3+\beta} + n^{2/3}}{n} = n^{1/6}\big(n^{-2/3+\beta} + n^{-1/3}\big).$$

Since $0 < \beta < 2/3$, the dominant contribution is $n^{1/6} \cdot n^{-1/3} = n^{-1/6}$.

(4) Second remainder term.

$$T\sqrt{\frac{Mh+\ell}{n}} \asymp n^{1/6}\sqrt{n^{1/3+\beta-1} + n^{2/3-1}} = n^{1/6}\sqrt{n^{-2/3+\beta} + n^{-1/3}}.$$

The larger of the two inside exponents is $-1/3$ (since $\beta < 2/3$), so

$$T\sqrt{\frac{Mh+\ell}{n}} \asymp n^{1/6} \cdot n^{-1/6} = n^{0} = O(1)$$

with this coarse estimate. (A small-t refinement, see discussion in the text, will provide an additional factor that reduces this contribution, cf. the Lemma 2.)

A crude estimate shows that the fourth term in the master bound is $O(1)$, which would prevent convergence. To overcome this, we invoke the small-t refinement (Lemma 4), proved in the Appendix section, which produces an extra factor $|t|$ in the Esseen integral and converts the contribution of this term into $O(n^{-1/3})$, dominated by $n^{-1/6}$.

(5) Smoothing term.

$$\frac{1}{T} \asymp n^{-1/6}.$$

Assembled orders and discussion. With the above calculations, the master bound becomes (up to constants)

$$\sup_x |F_n(x) - \Phi(x)| \lesssim n^{1/3-\beta\gamma\theta} \;+\; n^{\frac{1-2\delta}{6}} \;+\; n^{-1/6} \;+\; O(1) \;+\; n^{-1/6}.$$

Thus

- The dependence term is $n^{1/3-\beta\gamma\theta}$. To make it $O(n^{-1/6})$ we need

$$1/3 - \beta\gamma\theta \leq -1/6 \quad \Longleftrightarrow \quad \beta\gamma\theta \geq \tfrac{1}{2}.$$

- The Lyapunov-type term is $n^{(1-2\delta)/6}$, which vanishes if $\delta > \frac{1}{2}$ and is $O(n^{-1/6})$ once $\delta \geq 1$.
- The crude estimate of term (4) above is $O(1)$, which prevents convergence. This is the origin of the need for the small-t refinement: by proving a small-t lemma of the form

$$|\psi_n(t) - e^{-t^2/2}| \lesssim |t| \Big\{ M\phi(h)^\theta + \sqrt{\tfrac{Mh+\ell}{n}} \Big\} \quad \text{for } |t| \leq c/T,$$

one converts the contribution from the neighbourhood of $t = 0$ from $T\sqrt{(Mh+\ell)/n}$ into $\dfrac{1}{T}\sqrt{(Mh+\ell)/n} \asymp n^{-1/3}$, which is dominated by $n^{-1/6}$.
- After this refinement, the remaining dominant terms are $n^{1/3-\beta\gamma\theta}$, $n^{(1-2\delta)/6}$ and $n^{-1/6}$.

Conclusion. Under the usual polynomial decay assumption $\phi(h) \lesssim h^{-\gamma}$, and with the small-$t$ refinement in place, the uniform rate $O(n^{-1/6})$ holds provided

$$\delta \geq 1 \quad \text{and} \quad \beta\gamma\theta \geq \tfrac{1}{2}\,,$$

recalling $\theta = \dfrac{\delta}{2+\delta}$. In particular, the algebraic correction in item (2) is that the Lyapunov-type term is $n^{(1-2\delta)/6}$. $\square$

Remark 3. The assumptions on the decay parameter γ and the moment condition δ are crucial to determine the attainable rate of the normal approximation. For instance, when $\delta = 1$ (finite third moment), the summability condition requires $\gamma > 3$. If $\delta = 0.5$, then $\gamma > 5$ is needed. Thus, heavier tails (smaller δ) request faster decay of the mixing coefficients.

The Berry-Esseen bound in Theorem 2 shows that, under sufficiently fast polynomial mixing, the rate can reach $O(n^{-1/6})$. This should be compared with the i.i.d. case, where the universal $O(n^{-1/2})$ rate holds. Our results therefore quantify precisely how dependence slows down convergence, and provide explicit trade-offs between moment conditions and mixing rates.

3.4. Commentaries

Our findings connect closely with a substantial body of work on limit theorems for dependent sequences. Classical results due to Ibragimov (1962) established central limit theorems under ϕ-mixing with summability conditions on the coefficients. Subsequent refinements by Bradley (2005), Doukhan (1994), and Rio (2017) clarified the role of mixing rates and moment assumptions in guaranteeing Gaussian limits. Within this framework, our central limit theorem confirms that polynomially decaying ϕ-mixing still ensures asymptotic normality under $(2+\delta)$-moments.

In terms of convergence rates, the benchmark for independent sequences is the Berry-Esseen bound of order $O(n^{-1/2})$. For dependent data, it is well known that slower rates must be expected, typically between $n^{-1/6}$ and $n^{-1/3}$ depending on the strength of dependence. Our Berry-Esseen type inequality fits naturally within this range: the achievable rate $O(n^{-1/6})$ matches other block-based approaches for weakly dependent sequences, while our formulation makes explicit the trade-off between the moment parameter δ and the polynomial decay exponent γ.

In general, our results complement and refine the existing literature on this topic by providing a transparent and quantitative analysis of the effect of polynomial phi-mixing on the validity of the CLT and the attainable speed of normal approximation.

4. Conclusion and Perspectives

We have developed a framework for analyzing CLTs and normal approximation rates under polynomially decaying ϕ-mixing. Our results extend Berry-Esseen type bounds to this dependent setting, highlighting the crucial role of mixing rates.

Our results apply to a wide range of stochastic processes

- **Markov chains:** Polynomially ergodic Markov chains often satisfy ϕ-mixing with polynomial decay, making Theorem 1 directly applicable.
- **Linear processes:** Certain ARMA models with heavy-tailed innovations exhibit ϕ-mixing at a polynomial rate.
- **Dynamical systems:** Expanding the interval maps and related chaotic dynamical systems provide natural examples of processes with polynomial ϕ-mixing coefficients.

These examples highlight the practical significance of our bounds for dependent data in econometrics, time series analysis, and dynamical systems.

Future research directions include

- Investigating optimality of the $n^{-1/6}$ rate and potential improvements under stronger structural assumptions.
- Investigating whether similar results hold under ϕ-mixing sequences with exponential decay.

- Extending the analysis to functional CLTs (invariance principles) under ϕ-mixing.
- Applying these results to statistical estimators in econometrics and machine learning models involving dependent data.

Acknowledgments

The authors thank the Editorial Board, the Associate Editor, and the anonymous referee for their helpful comments and suggestions that improved this paper.

Appendix

Lemma 3. *Let Y be a real random variable such that $Y(\omega) \in [-1, 1]$ for all ω. Then the following identity holds almost surely:*

$$Y = \int_{-1}^{1} \left(\mathbf{1}_{\{Y>y\}} - \tfrac{1}{2} \right) dy.$$

More generally, for any deterministic $a < b$ and any random variable Y with values in $[a, b]$, one has

$$Y = \int_{a}^{b} \left(\mathbf{1}_{\{Y>y\}} - \tfrac{1}{b-a} \tfrac{b-a}{2} \right) dy.$$

Proof. We prove the displayed identity pointwise (that is, for each fixed outcome ω). Fix ω and set $c := Y(\omega) \in [-1, 1]$. Consider the two integrals

$$I_1(c) := \int_{-1}^{1} \mathbf{1}_{\{c>y\}}\, dy, \qquad I_2(c) := \int_{-1}^{1} \mathbf{1}_{\{c\le y\}}\, dy.$$

Since $y \mapsto \mathbf{1}_{\{c>y\}}$ is the indicator of the interval $[-1, c)$, we compute

$$I_1(c) = \int_{-1}^{c} 1\, dy = c - (-1) = c + 1.$$

Similarly, $y \mapsto \mathbf{1}_{\{c\le y\}}$ is the indicator of $[c, 1]$, hence

$$I_2(c) = \int_{c}^{1} 1\, dy = 1 - c.$$

Subtracting gives

$$I_1(c) - I_2(c) = (c + 1) - (1 - c) = 2c.$$

Thus for this fixed outcome ω, we have

$$2Y(\omega) = \int_{-1}^{1} \mathbf{1}_{\{Y(\omega)>y\}}\, dy - \int_{-1}^{1} \mathbf{1}_{\{Y(\omega)\le y\}}\, dy.$$

Equivalently,

$$Y(\omega) = \int_{-1}^{1} \left(\mathbf{1}_{\{Y(\omega)>y\}} - \tfrac{1}{2} \right) dy,$$

because $\int_{-1}^{1} \frac{1}{2}\, dy = 1$.

The aboved equality holds for every ω (not only almost surely) because the pointwise computations used no probabilistic argument beyond the boundedness $Y(\omega) \in [-1, 1]$.

To use this identity under the probability integral or to apply dominated convergence, note that the integrand

$$\omega \mapsto \mathbf{1}_{\{Y(\omega)>y\}} - \tfrac{1}{2}$$

is jointly measurable in (ω, y) (product measurability: for fixed y it is measurable in ω, and for fixed ω it is measurable in y). Moreover, it is uniformly bounded by 1 in absolute value. Therefore Fubini's theorem (or dominated convergence) allows interchange of the order of integration: for instance, for any integrable functional, one may write

$$\int_\Omega Y(\omega)\, d\mathbb{P}(\omega) = \int_\Omega \int_{-1}^{1} \left(\mathbf{1}_{\{Y(\omega)>y\}} - \tfrac{1}{2}\right) dy\, d\mathbb{P}(\omega) = \int_{-1}^{1} \int_\Omega \left(\mathbf{1}_{\{Y>y\}} - \tfrac{1}{2}\right) d\mathbb{P}\, dy.$$

This representation is the usual layer-cake (or Hoeffding) decomposition for bounded random variables and is often used to express covariances in terms of integrals of indicator events.

Appendix: the small-t lemma

We prove the small-t estimate used in Step 8 in theorem (2). The argument is standard; we give a self-contained presentation adapted to the triangular array and blocking notation used in the main text.

Set-up and notation. Let $\{X_{n,k} : 1 \le k \le n\}$ be the (row-wise) centered triangular array considered in the main text and let

$$S_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} X_{n,k}$$

be the normalized sum with $\mathrm{Var}(S_n) = 1$. Fix the block decomposition with parameters ℓ, h, M as in the main text: there are M main blocks of length ℓ separated by gaps of length h (neglect boundary effects, handled in the same way as usual). Denote by $B_{n,j}$ the sum over the j-th main block (after centering and normalization), so that

$$S_n = \sum_{j=1}^{M} B_{n,j} + R_n,$$

where R_n collects the contributions of the separators and boundary pieces. Let $\psi_n(t) = \mathbb{E}e^{itS_n}$ be the characteristic function of S_n.

We assume the mixing control recorded in the main text: there exists a nonincreasing function $\phi(\cdot)$ and an exponent $\theta = \dfrac{\delta}{2+\delta} \in (0,1)$ such that mixing inequalities of Davydov/Ibragimov type hold and will be used below to bound dependence across gaps of length h. We also suppose the triangular array has sufficient moments (up to order $2+\delta$) so that the Taylor expansions below are valid after truncation.

Lemma 4 (Small-t) lemma). *There exist absolute constants $t_0 > 0$ and $C > 0$ (independent of n) such that for all $|t| \le t_0$,*

$$\left|\psi_n(t) - e^{-t^2/2}\right| \le C\,|t|\,\left(M\,\phi(h)^{\theta} \; + \; \sqrt{\frac{Mh+\ell}{n}}\,\right).$$

The proof of Lemma 4 follows the classical cumulant Taylor expansion combined with mixing inequalities (see Ibragimov and Linnik (1971), Petrov (1975), Rio (2000b)).

Proof. The proof proceeds in three steps: truncation, blockwise Taylor expansion, and dependence control.

Step 1. Truncation and control of truncation error. Fix a truncation level $u_n > 0$ (to be chosen below). Define truncated variables

$$\tilde{X}_{n,k} = X_{n,k}\mathbf{1}\{|X_{n,k}| \le u_n\} - \mathbb{E}\big[X_{n,k}\mathbf{1}\{|X_{n,k}| \le u_n\}\big],$$

so each $\tilde{X}_{n,k}$ is centered and bounded by $|\tilde{X}_{n,k}| \le 2u_n$. Let $\tilde{S}_n$ be the corresponding normalized truncated sum and $\tilde{\psi}_n(t) = \mathbb{E}e^{it\tilde{S}_n}$ its characteristic function. The truncation error is handled by standard moment bounds: by using Markov and Hölder inequalities together with the assumed finite $(2+\delta)$-th moments, one gets

$$\left|\psi_n(t) - \tilde{\psi}_n(t)\right| \le C_1|t|\,\mathbb{E}|S_n - \tilde{S}_n| \le C_2|t|\,\frac{1}{\sigma_n}\sum_{k=1}^{n}\mathbb{E}\big[|X_{n,k}|\mathbf{1}\{|X_{n,k}| > u_n\}\big].$$

Choose u_n so that the right-hand side is bounded by the second term in the desired bound: specifically take u_n so large that

$$\frac{1}{\sigma_n}\sum_{k=1}^{n}\mathbb{E}\big[|X_{n,k}|\mathbf{1}\{|X_{n,k}| > u_n\}\big] \lesssim \sqrt{\frac{Mh+\ell}{n}}.$$

This is always possible under the moment assumptions (one typical choice is $u_n \asymp \sigma_n\sqrt{n/(Mh+\ell)}$, which makes the truncated tail negligible in the required sense). With this choice, we obtain the truncation bound.

$$\left|\psi_n(t) - \tilde{\psi}_n(t)\right| \le C|t|\sqrt{\frac{Mh+\ell}{n}}. \tag{23}$$

From now on, working with the truncated and centered array $\tilde{X}_{n,k}$, we drop the tilde notationally for brevity, but remember that each variable is bounded by $O(u_n)$.

Step 2. Blockwise expansion of the characteristic function (Taylor expansion). Write the normalized block sums as

$$B_j = \frac{1}{\sigma_n} \sum_{k \in \text{block } j} X_{n,k}, \quad j = 1, \ldots, M.$$

By construction $S_n = \sum_{j=1}^{M} B_j + R_n$, and the contribution R_n from gaps and boundary terms has a variance of order $(Mh + \ell)/n$ and hence

$$\mathbb{E}|R_n|^2 \lesssim \frac{Mh + \ell}{n}, \quad \Longrightarrow \quad |\mathbb{E}e^{itR_n} - 1| \lesssim |t| \sqrt{\frac{Mh + \ell}{n}}. \tag{24}$$

Thus, the R_n piece is already compatible with the desired bound.

Consider the block characteristic function

$$\varphi_j(t) = \mathbb{E}e^{itB_j}, \quad j = 1, \ldots, M.$$

Because the $X_{n,k}$ in the block are bounded (by $O(u_n)$) and have uniformly bounded moments up to the order $2 + \delta$, we can expand the logarithm of φ_j around zero. Using the cumulant Taylor expansion for bounded variables,

$$\log \varphi_j(t) = it\mu_j - \tfrac{1}{2}t^2\sigma_j^2 + R_j(t),$$

where $\mu_j = \mathbb{E}B_j = 0$ (blocks are centered by construction), $\sigma_j^2 = \mathrm{Var}(B_j)$, and the remainder satisfies, uniformly for $|t|$ small,

$$|R_j(t)| \le C\big(|t|^3\, \mathbb{E}|B_j|^3\big).$$

The key point is that the leading part is $O(|t|^3)$ on the block scale. Since the block variance scales like $\sigma_j^2 \asymp \ell/n$ (so that $\sum_j \sigma_j^2 \approx 1$), the normalized third moment $\mathbb{E}|B_j|^3$ is $O((\ell/n)^{3/2})$ up to factors coming from truncation, hence

$$|R_j(t)| \lesssim |t|^3 \Big(\frac{\ell}{n}\Big)^{3/2}.$$

Summing over $j = 1, \ldots, M$, we obtain (using $M\ell \asymp n$)

$$\sum_{j=1}^{M} |R_j(t)| \lesssim |t|^3 M \Big(\frac{\ell}{n}\Big)^{3/2} = |t|^3 \frac{M\ell^{3/2}}{n^{3/2}} \asymp |t|^3 \frac{n^{1/3} \cdot n^1}{n^{3/2}} = |t|^3 n^{-1/6}.$$

Thus, for $|t|$ bounded by a small absolute t_0 this total remainder is $o(|t|)$, in particular bounded by $C|t|\sqrt{(Mh + \ell)/n}$ once n large (because $|t|^2 n^{-1/6} \ll \sqrt{(Mh + \ell)/n}$ for our block scale choices). The important algebraic point is that each block remainder carries at least a factor $|t|$ beyond the quadratic contribution, and the summation keeps one factor $|t|$ exterior to the block error terms.

Therefore, at the block level and ignoring dependence across blocks,

$$\prod_{j=1}^{M} \varphi_j(t) = \exp\Big(-\tfrac{1}{2}t^2 \sum_{j=1}^{M} \sigma_j^2\Big) \cdot \big(1 + O(|t|) \cdot E_{\text{blocks}}\big),$$

where

$$E_{\text{blocks}} \lesssim \sum_{j=1}^{M} (\text{block remainders and truncation errors}) \lesssim M\phi(h)^{\theta} + \sqrt{\frac{Mh+\ell}{n}}.$$

The term $M\phi(h)^{\theta}$ is not yet justified. It will come from dependence control in the next step. The key conclusion of the present expansion is the structural factor $|t|$ that multiplies the contributions of the block error.

Step 3. Dependence control and decoupling across gaps. So far, we treated the block of characteristic functions as if they were independent. In reality, blocks are separated by gaps of length h and depend weakly. We use the assumed mixing bound to quantify this weak dependence. The suitable ϕ-mixing inequality) yields, for two sigma fields separated by a gap of length h,

$$\big|\operatorname{Cov}(U,V)\big| \le C\,\phi(h)^{\theta}\|U\|_p\|V\|_q,$$

for appropriate conjugate exponents p, q (here we use $p = q = 2+\delta$ and $\theta = \delta/(2+\delta)$). Apply this inequality to the multiplicative decomposition of the joint characteristic function into block factors. The standard consequence (see Ibragimov and Linnik (1971) type decoupling) is that the true joint characteristic function satisfies

$$\Big|\mathbb{E}\prod_{j=1}^{M} e^{itB_j} \;-\; \prod_{j=1}^{M}\mathbb{E}e^{itB_j}\Big| \le C\,M\,\phi(h)^{\theta}\cdot R(t),$$

where $R(t)$ is a polynomially bounded function of t depending on block moments; after truncation $R(t)$ is uniformly bounded for $|t| \le t_0$. The factor $M\phi(h)^{\theta}$ arises because the dependence is controlled pairwise across the M blocks with gaps of length h and the Ibragimov/Davydov inequality yields a multiplicative $\phi(h)^{\theta}$ factor for each link.

Combining this bound with the blockwise Taylor expansion in Step 2 gives the desired multiplicative decomposition with an explicit $M\phi(h)^{\theta}$ error term: specifically there exists $C > 0$ such that for $|t| \le t_0$,

$$\Big|\mathbb{E}\prod_{j=1}^{M} e^{itB_j} - \exp\Big(-\tfrac{1}{2}t^2\sum_j \sigma_j^2\Big)\Big| \le C|t|\Big(M\phi(h)^{\theta} + \sqrt{\frac{Mh+\ell}{n}}\Big).$$

(Here the factor $|t|$ stems from the Taylor remainders in Step 2; the $M\phi(h)^{\theta}$ term stems from decoupling errors produced by the mixing inequality; the $\sqrt{(Mh+\ell)/n}$ term comes from the separators and the truncation residuals handled in Step 1.)

Step 4. Collecting the pieces and concluding. Using the decomposition $S_n = \sum_{j=1}^{M} B_j + R_n$, the product expansion and (24), and combining with the truncation error estimate (23), we obtain for $|t| \le t_0$

$$\begin{aligned}\left|\psi_n(t)-e^{-t^2/2}\right| &\le \left|\psi_n(t)-\tilde{\psi}_n(t)\right|+\left|\tilde{\psi}_n(t)-e^{-t^2/2}\right| \\ &\le C|t|\sqrt{\frac{Mh+\ell}{n}}+C|t|\Big(M\phi(h)^\theta+\sqrt{\frac{Mh+\ell}{n}}\Big),\end{aligned}$$

hence (renaming C) the asserted inequality

$$\left|\psi_n(t)-e^{-t^2/2}\right| \le C\,|t|\,\Big(M\,\phi(h)^\theta\;+\;\sqrt{\frac{Mh+\ell}{n}}\,\Big).$$

This completes the proof of the lemma.

Remark. The constants and the precise choice of the truncation level u_n can be made explicit if desired (choose u_n so that tail contributions are negligible in the $\sqrt{(Mh+\ell)/n}$ scale). The argument above is modular: one first chooses u_n to control truncation, then performs Taylor expansions on the bounded variables, and finally uses mixing inequalities to decouple blocks. The crucial structural fact is that the blockwise remainder from the Taylor expansion carries at least one factor $|t|$, which yields the overall exterior factor $|t|$ in the final bound.

References

Berry, A., C. (1941) The Accuracy of the Gaussian Approximation to the Sum of Independent Variates. *Transactions of the American Mathematical Society*, 49, 122-136.

Bradley, R., C. (2005) Basic properties of strong mixing conditions: a survey and some open questions. *Probability Surveys*, 2, 107-144.

Doukhan, P. (1994) *Mixing: Properties and Examples*. Springer.

Esseen, C-G. (1942) On the Liapunoff Limit of Error in the Theory of Probability. *Arkiv för Matematik, Astronomi och Fysik*, 28A, 1-19.

Ibragimov, I., A. (1962) Some limit theorems for stationary processes. *Theory of Probability & Its Applications*, 7(4):349-382.

Ibragimov, I., A. and Linnik, Y., V. (1971) *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff.

Lévy, P. (1925) Calcul des probabilités. *Gauthier-Villars, Paris*.

Lindeberg, J., W. (1922) Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 15(1), 211-225.

Lyapunov, A., M. (1901) Nouvelle forme du théorème sur la limite des probabilités. *Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg*, 12(5), 1-24.

Merlevède, F., Peligrad, M. and Rio, E. (2009) A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probability Theory and Related Fields*, 145(3-4), 475-505.

Petrov, V., V. (1975) *Sums of Independent Random Variables*. Springer-Verlag. ISBN, 0387066357, 9780387066356; Longueur, 345 pages

Rio, E. (2017) *Asymptotic Theory of Weakly Dependent Random Pocesses*. Springer. DOI: 10.1007/978-3-662-54323-8.

Rio, E. (2009) Inégalités de concentration pour les processus aléatoires faiblement dépendants. *Springer, Mathématiques & Applications*, 46.

Rio, E. (2000a) Inégalités de Hoeffding pour les fonctions Lipschitz de suites dépendantes. *C. R. Acad. Sci. Paris*, 330(10):905-908.

Rio, E. (2000b) Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants. *Springer Lecture Notes in Mathematics*.

Rio, E. (1996) The Functional Law of the Iterated Logarithm for Stationary Strongly Mixing Sequences. *The Annals of Probability*, 24(1), 178-197.

Shevtsova, I., G. (2011) An improvement of the upper bound of the absolute constant in the Berry-Esseen inequality. *Theory of Probability and Its Applications* 56(3), 539-550, doi = 10.1137/S0040585X97983743.

Utev, S., A. (1991) Central limit theorem for dependent random variables. *Probability Theory and Mathematical Statistics*, 2, 519-528.

Utev, S., A. (1990) A Berry-Esseen inequality for strongly mixing sequences. *Theory of Probability and Its Applications*, 35(4), 709-713.

Wu, W., B. (2011) Asymptotic Theory for Nonlinear Time Series. *Springer*.

Wu, W., B. (2005) Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences*, 102(40), 14150-14154.

Wu, W., B. and Shao, X. (2004) Limit theorems for iterated random functions. *Journal of Applied Probability*, 41(2), 425-436.