



Almost sure invariance principle of β -mixing time series in Hilbert space

Jianya Lu , Wei Biao Wu , Zhijie Xiao & Lihu Xu

To cite this article: Jianya Lu , Wei Biao Wu , Zhijie Xiao & Lihu Xu (01 Sep 2026): Almost sure invariance principle of β -mixing time series in Hilbert space, Stochastic Analysis and Applications, DOI: [10.1080/07362994.2026.2714807](https://doi.org/10.1080/07362994.2026.2714807)

To link to this article: <https://doi.org/10.1080/07362994.2026.2714807>



© 2026 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 01 Sep 2026.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Almost sure invariance principle of β -mixing time series in Hilbert space

Jianya Lu^a, Wei Biao Wu^b, Zhijie Xiao^c, and Lihu Xu^{d,e}

^aSchool of Mathematics, Statistics and Actuarial Science, University of Essex, Colchester, UK; ^bDepartment of Statistics, University of Chicago, Chicago, IL, USA; ^cDepartment of Economics, Boston College, Chestnut Hill, MA, USA; ^dDepartment of Statistics and Probability, Michigan State University, East Lansing, MI, USA; ^eDepartment of Mathematics, University of Macau, Taipa, China

ABSTRACT

Inspired by Berkes et al. [3] and Wu [37], we prove an almost sure invariance principle for stationary β -mixing stochastic processes defined on Hilbert spaces. Our result applies to Markov chains satisfying the Meyn-Tweedie-type Lyapunov condition, thereby generalizing the contraction condition in Berkes et al. [3, Example 2.2]. The main theorem is established using the big and small blocks technique and an embedding result from Götze and Zaitsev [14]. Examples of ergodic Markov chains and functional autoregressive processes are provided as applications.

ARTICLE HISTORY

Received 1 February 2024
Accepted 2 August 2026

KEYWORDS

Almost sure invariance principle; Hilbert space; β -mixing time series

MSC2020

60F17; 60G10

1. Introduction

Let $(X_k)_{k \in \mathbb{N}_0}$ be random variables with $\mathbb{E}X_k = 0$ and $\mathbb{E}|X_k|^p < \infty$ for $p > 2$, $(X_k)_{k \in \mathbb{N}_0}$ satisfies the almost sure invariance principle (ASIP) with rate r_n if there exists, after suitably enlarging the probability space, independent Gaussian random variables $(\eta_k)_{k \in \mathbb{N}_0}$ with covariance $\text{cov}(X_k)$, that is $\eta_k \sim \mathcal{N}(0, \text{cov}(X_k))$ such that

$$\max_{1 \leq i \leq n} \left| \sum_{k=1}^i X_k - \sum_{k=1}^i \eta_k \right| = o(r_n), \quad \text{almost surely.}$$

The ASIP was introduced by [33,34] to prove the functional law of iterated logarithm for independent, identically distributed (i.i.d.) random variables. Moreover, it serves as a powerful tool, directly implying Donsker's theorem [17,18]. considered this problem for i.i.d. random variables with a finite p -th moment and established the optimal rate $n^{1/p}$ for $p > 2$. For dependent random variables [37], employed martingale approximation and the Skorokhod embedding to demonstrate the ASIP with a rate of $n^{1/p}(\log n)^{1/2}$ for a class of stationary processes with $2 < p < 4$ in $\mathbb{R}$. However, this approach cannot achieve the optimal rate, and extending the martingale embedding method to d -dimensional spaces is challenging, see [25] for more details. Under the framework of [3,37] applied the block technique to construct independent block sums, obtaining the optimal rate $n^{1/p}$ for $p > 2$ in $\mathbb{R}$. For further studies on the ASIP in $\mathbb{R}$, we refer the reader to [8,9,22,31].

CONTACT Lihu Xu  xulihu@msu.edu  Department of Statistics and Probability, Michigan State University, East Lansing, MI, USA.

© 2026 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

For the multidimensional invariance principle [38], obtained the optimal rate for the multivariate version of the Komlós–Major–Tusnádý (KMT) theorem and then established multivariate versions of Sakhanenko’s theorem, see [39–41]. These Gaussian approximation results for independent multivariate random variables provide essential tools for proving the ASIP for multidimensional dependent random variables. Using the block technique [20], achieved the rate $n^{1/p}$ for $2 < p < 4$ in $\mathbb{R}^d$. For more general p [15], established the ASIP with rate $n^{1/4+1/(4p-4)+\epsilon}$ for $p > 2$ and $\epsilon > 0$, under mixing conditions expressed *via* the characteristic function. Under similar conditions [16], obtained the rate $n^{1/4+\epsilon}$ for uniformly bounded ϕ -mixing sequences defined in d -dimensional space.

The motivations for studying the ASIP of stationary β -mixing time series are two fold. First, while there has been extensive research on the ASIP for dependent $\mathbb{R}$ - and $\mathbb{R}^d$ -valued random variables, there are few results of this type for dependent time series in Hilbert spaces. See [7,12] and the references therein for some existing work. Second, classical results by [3,20,37] rely on geometric moment contraction (GMC) conditions, which require strict exponential decay rates and are not satisfied by some examples. In this paper, we aim to prove the ASIP under a more general condition.

Our main contribution is the proof of the ASIP for stationary β -mixing time series defined on infinite dimensional Hilbert spaces using the block technique. This result not only complements the existing ASIP results in Hilbert spaces but also relaxes the GMC condition, thereby broadening its applicability. Due to the equivalence between β -mixing and ergodicity for stationary Markov chains, many practical examples satisfy this condition and consequently exhibit the corresponding ASIP. Furthermore, since our result is independent of dimensionality, it provides a tool for analyzing asymptotic normality for high-dimensional data analysis and can get a convergence rate independent with the dimension. As a byproduct, we extend the moment inequality for this type of time series to the Hilbert space setting. However, a limitation of our result is that the convergence rate obtained in the ASIP is not optimal. This arises from the treatment of certain elements in the Hilbert space as remainders, which introduces additional approximation error into the overall result. Future research could focus on refining the decomposition and better block construction techniques to improve the rate.

1.1. Main result and methodology

Considering a separable Hilbert space $\mathbb{H}$ with orthonormal basis $\{\mathbf{e}_k\}_{k \in \mathbb{N}}$, denote by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ the associated inner product and norm respectively. For any $\mathbf{x} \in \mathbb{H}$, there exists a unique representation $\mathbf{x} = \sum_{k=1}^{\infty} x_k \mathbf{e}_k$ with $x_k \in \mathbb{R}$ for $k \geq 1$ and $\sum_{k=1}^{\infty} x_k^2 < \infty$, and thus we can represent $\mathbf{x}$ by the sequence $(x_1, \dots, x_k, \dots)^T$, where T is the transpose operator. For a positive integer d and any $\mathbf{x} \in \mathbb{H}$, $\mathbb{H}$ can be decomposed into

$$\mathbb{H} = \mathbb{H}_{\leq d} + \mathbb{H}_{> d}$$

with element $(x_1, \dots, x_d)^T \in \mathbb{H}_{\leq d}$ and $(x_{d+1}, x_{d+2}, \dots)^T \in \mathbb{H}_{> d}$.

Let $\mathbf{X}$ and $\mathbf{Y}$ be two $\mathbb{H}$ -valued random variables, denote by $\mathbb{P}_{\mathbf{X} \times \mathbf{Y}}$ the joint probability of $(\mathbf{X}, \mathbf{Y})$ and by $\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}}$ the probabilities of $\mathbf{X}$ and $\mathbf{Y}$ respectively, define

$$\beta(\mathbf{X}, \mathbf{Y}) := \|\mathbb{P}_{\mathbf{X} \times \mathbf{Y}} - \mathbb{P}_{\mathbf{X}} \times \mathbb{P}_{\mathbf{Y}}\|_{\text{TV}},$$

where $\|\cdot\|_{TV}$ is the total variation norm of probability measures, i.e.,

$$\|\mathbb{P}_{X \times Y} - \mathbb{P}_X \times \mathbb{P}_Y\|_{TV} = \sup_{f \in C_b(\mathbb{H} \times \mathbb{H}) : \|f\|_\infty \leq 1} |\mathbb{P}_{X \times Y}(f) - \mathbb{P}_X \times \mathbb{P}_Y(f)|.$$

Now we consider an $\mathbb{H}$ -valued time series $(X_k)_{k \in \mathbb{N}_0}$. The β -mixing coefficient of the time series is defined as

$$\beta(n) = \sup_{k \geq 1} \beta((X_i)_{0 \leq i \leq k}, (X_i)_{i \geq n+k}). \quad (1.1)$$

In this paper, we assume that

(A1) $(X_k)_{k \in \mathbb{N}_0}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is stationary with a stationary measure π and exponentially β -mixing, i.e., there exist some $C > 0$ and $\beta > 0$ such that

$$\beta(n) \leq Ce^{-\beta n}, \quad n \in \mathbb{N}.$$

(A2) There exists some positive linear operator $\mathbf{\Gamma} : \mathbb{H} \rightarrow \mathbb{H}$ such that

$$\lim_{n \rightarrow \infty} \frac{\text{cov}(\sum_{k=0}^{n-1} X_k)}{n} = \mathbf{\Gamma},$$

where $\text{cov}(\cdot)$ is the covariance operator of the $\mathbb{H}$ -valued random variable. Moreover, $\mathbf{\Gamma}$ has positive eigenvalues $\{\lambda_k\}_{k \in \mathbb{N}}$ with $\lambda_1 \geq \lambda_2 \geq \dots > 0$. There exist $\delta_1 \geq \delta_2 > 1$ and positive constants C_1 and C_2 , such that

$$C_1 k^{-\delta_1} \leq \lambda_k \leq C_2 k^{-\delta_2}, \quad k \in \mathbb{N}.$$

Remark 1.1. It is easy to obtain

$$\mathbf{\Gamma} = \text{cov}(X_0) + \sum_{k=1}^{\infty} (\text{cov}(X_0, X_k) + \text{cov}(X_0, X_k)^T).$$

The assumption **(A2)** implies that $\mathbf{\Gamma}$ has effective rank and the eigenvalues of covariance operator polynomially decay. Similar conditions can be found in [21, 29] and citations therein.

For random variables ξ and η , let $\xi \stackrel{\mathcal{D}}{=} \eta$ denote they have the same distribution. Our main result is given as follows.

Theorem 1.2. Assume that **(A1)** and **(A2)** hold and that the stationary measure π has p -th moment with $p > 2$. Then there exists a probability space on which we can define random variables X_i^* and i.i.d. Gaussian random variables η_i such that $(X_i)_{1 \leq i \leq n} \stackrel{\mathcal{D}}{=} (X_i^*)_{1 \leq i \leq n}$, $(\eta_i)_{1 \leq i \leq n} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$ and

$$\max_{1 \leq i \leq n} \left\| \sum_{j=1}^i (X_j^* - \pi(X)) - \sum_{j=1}^i \eta_j \right\| = o(n^{\bar{\theta}}), \quad a.s.,$$

where $\mathbf{\Gamma}$ defined in **(A2)** and $\bar{\theta} > \max \left\{ \frac{(2p-2)\delta_1 + 2\delta_2 + 23p-4}{44p-4 + (4p-4)\delta_1 + 2p\delta_2}, \frac{(2p-2)\delta_1 + 2\delta_2 + p(p+4)/2-4}{p(p+2)-4 + (4p-4)\delta_1 + 2p\delta_2} \right\}$.

Corollary 1.3. Under the conditions of **Theorem 1.2**, for any $\varepsilon > 0$, there exists a $\bar{\delta}_{p,\varepsilon}$ depending on p and ε and defined in (3.17) such that as $\lambda_k \asymp k^{-\delta}$ and $\delta > \bar{\delta}_{p,\varepsilon}$, there exists a probability

space on which we can define random variables $\mathbf{X}_i^*$ and i.i.d. Gaussian random variables $\boldsymbol{\eta}_i$ such that $(\mathbf{X}_i)_{1 \leq i \leq n} \stackrel{\mathcal{D}}{=} (\mathbf{X}_i^*)_{1 \leq i \leq n}$, $\boldsymbol{\eta}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma})$ and

$$\max_{1 \leq i \leq n} \left\| \sum_{j=1}^i (\mathbf{X}_j^* - \pi(\mathbf{X})) - \sum_{j=1}^i \boldsymbol{\eta}_j \right\| = o\left(n^{\frac{1}{3} + \frac{2}{3(p-2)} + \varepsilon}\right), \quad a.s. \quad (1.2)$$

Remark 1.4. For random variables defined on $\mathbb{R}^d$ with condition (A1) and further assume $\boldsymbol{\Gamma}$ is a positive semidefinite matrix. As we shall see in step 1 of the proof of Lemma 3.3, by the block technique and the invariance principle of i.i.d. random vectors in [42, Corollary 3], when $\lambda > \frac{1}{4} + \frac{1}{4(p-1)}$, the following ASIP holds,

$$\max_{1 \leq i \leq n} \left\| \sum_{j=1}^i (\mathbf{X}_j^* - \pi(\mathbf{X})) - \sum_{j=1}^i \boldsymbol{\eta}_j \right\| = o\left((d^8 \log d)n^\lambda\right), \quad a.s. \quad (1.3)$$

This rate is same as [15, Theorem 1.2] which is under the mixing condition tied to spectral properties. As $\lambda_k \asymp e^{-|k|^\gamma}$ with $\gamma > 0$, the rate on the right hand side of (1.2) can be improved to $o\left(n^{\frac{1}{3} + \frac{2}{3(p-2)}}\right)$ by our method. One may intuitively see this by letting δ_1 and δ_2 tend to ∞ in Theorem 1.2.

Remark 1.5 [6, Theorem 2.1]. considered the almost sure invariance principle for the Markov chain under the polynomial ergodic condition, which is closely related to the β -mixing condition, in 1-dimensional space. Following the Nummelin splitting technique, see [26], to construct independent blocks, they obtained the rate $n^{\frac{1}{4} \vee \frac{1}{p}}$ for $p \in (2, 4]$. See [23, 27] for more details of this method and result. Their result considers the 1-dimensional case and requires the Markov property condition, which is not used in this paper. The weaker polynomial mixing condition and more optimal rate is our future research goal.

Let us briefly describe our approach as follows. We divide the positive integer number $\mathbb{N}$ into the intervals $[2^m + 1, 2^{m+1}]$ for $m = 0, 1, \dots$. Each interval is partitioned into a sequence of big blocks with length $2^{\alpha_1 m}$ for $\alpha_1 \in (0, 1)$ and small blocks with length m . Following the properties of β -mixing, we can construct i.i.d. random variables distributed as the big block sums in $\mathbb{H}$. The projection of these random variables on $\mathbb{H}_{\leq d}$ are comparable with Gaussian random variables following [14], while the residual on $\mathbb{H}_{> d}$ and small block sums are negligible. After carefully choosing the relation between the dimension d and the sample size n , we can get the result.

1.2. Organization of the paper and some notations

The paper is organized as follows. Our main result is stated in Section 1. In Section 2, we provide some preliminary lemmas and the technique for the proof of our main result. The proof of the ASIP is given in Section 3. In Section 4, we give some applications. Some proofs of preliminary lemmas are given in Appendices A and B, and the proofs of examples in Section 4 are given in Appendix C.

We finish this section by introducing some notations which will be frequently used in sequel. For variables $(X_i)_{i \in \mathbb{N}_0} \in \mathbb{H}$, $X_{i,j}$ denotes the j -th element of X_i . We denote ξ^* the

random variable that has the same distribution as ξ . For an operator $A : \mathbb{H} \rightarrow \mathbb{H}$, $\|A\|_F$ for operator on $\mathbb{H}$ denote the Frobenius norm of A , i.e., $\|A\|_F = \sqrt{\text{Tr}(AA^T)}$. For $x \in \mathbb{R}$, $\lceil x \rceil$ and $\lfloor x \rfloor$ smallest integer greater than x and the integer part of x respectively. We further denote $a_n \asymp b_n$ if there exist positive constants c and C such that $ca_n \leq b_n \leq Ca_n$. The symbols C and c denote positive numbers, C_p and c_p denote positive numbers depend on the parameter p . Their values may vary from line to line.

For the stationary process $(X_k)_{k \in \mathbb{N}_0}$ with ergodic measure π , without loss of generality, from now on we assume

$$\pi(X_0) = \mathbf{0}.$$

2. Auxiliary lemmas for Theorem 1.2

The strategy of proving [Theorem 1.2](#) is to decompose $\sum_{i=1}^n X_i$ into two parts by the block technique, showing that the big block sum is comparable with Gaussian random variables and the small block sum is negligible following the properties of β -mixing which are shown below. In this section, we provide some preliminary lemmas and the technique for the proof of our main result.

2.1. Lemmas of β -mixing time series

In this subsection, we recall the result of [\[1\]](#) first for the convenience of the reader, which is useful to construct the independence for β -mixing time series. At the end of this subsection, we give the following two lemmas, the first lemma paving a way for proving the last which is an extension to infinite dimensional space valued random variables of [\[32, Theorem 4.1\]](#).

Lemma 2.1 [\[24, Lemma 2\]](#). *Let X and Y be two random variables defined on $\mathbb{H}$ with quantile functions $Q_{\|X\|}(u)$ and $Q_{\|Y\|}(u)$. Then*

$$|\mathbb{E}\langle X, Y \rangle - \langle \mathbb{E}X, \mathbb{E}Y \rangle| \leq 18 \int_0^{\bar{\alpha}} Q_{\|X\|}(u) Q_{\|Y\|}(u) du.$$

Here $Q_{\|X\|}(u) = \inf\{t : \mathbb{P}(\|X\| > t) \leq u\}$ and $Q_{\|Y\|}(u)$ is defined similarly. $\bar{\alpha}$ is defined by $\bar{\alpha} = \sup_{A \in \sigma(X), B \in \sigma(Y)} \{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\}$. Similarly,

$$\|\text{cov}(X, Y)\|_F \leq 18 \int_0^{\bar{\alpha}} Q_{\|X\|}(u) Q_{\|Y\|}(u) du.$$

It is easy to see that $Q_{\|X\|}$ is the inverse function of $\mathbb{P}(\|X\| > t)$ and is a non-increasing function. Moreover, $Q_{\|X\|}(U) \stackrel{\mathcal{D}}{=} \|X\|$, where U is a random variable uniformly distributed on $[0, 1]$.

Lemma 2.2. *Suppose Assumption (A1) holds and $X_0 \sim \pi$ with p -th moment, there exists a constant $\bar{C} > 0$, such that for any $2 \leq p' < p$ and large enough n ,*

$$\mathbb{E} \left\| \sum_{i=1}^n X_i \right\|^{p'} \leq \bar{C} n^{\frac{p'}{2}} (\pi(\|X\|^p))^{\frac{p'}{p}}.$$

Proof. The proof is given in [Appendix B](#). □

2.2. Blocking

The blocking technique is a typical way for dependent time series and applied in proving almost sure invariance principle, see [3,15,20].

We subdivide $\mathbb{N}$ into the intervals $[2^m + 1, 2^{m+1}]$. For any positive integer m , let $m_1 = 2^{\lfloor \alpha_1 m \rfloor}$ and $m_2 = \lfloor C^* \log 2^m \rfloor$ with $0 < \alpha_1 < 1$ and $C^* > 0$, α_1 and C^* will be chosen later. For any $n \in [2^m + 1, 2^{m+1}]$, we denote

$$\kappa(n) = \left\lfloor \frac{n - 2^m}{m_1 + m_2} \right\rfloor.$$

For $1 \leq j \leq \kappa(2^{m+1})$, put

$$I_{m,j} = \{i : 2^m + (m_1 + m_2)(j-1) + 1 \leq i \leq 2^m + (m_1 + m_2)(j-1) + m_1\},$$

$$J_{m,j} = \{i : 2^m + (m_1 + m_2)(j-1) + m_1 + 1 \leq i \leq 2^m + (m_1 + m_2)j\},$$

where $I_{m,j}$ (resp., $J_{m,j}$) are big (resp., small) blocks. The tail is defined by

$$I_{m,\kappa(2^{m+1})+1} = \{i : 2^m + (m_1 + m_2)\kappa(2^{m+1}) + 1 \leq i \leq 2^{m+1} \wedge ((m_1 + m_2)\kappa(2^{m+1}) + m_1)\},$$

$$J_{m,\kappa(2^{m+1})+1} = \{i : 2^m + (m_1 + m_2)\kappa(2^{m+1}) + m_1 + 1 \leq i \leq 2^{m+1}\}.$$

Thus, we decompose the interval $[2^m + 1, 2^{m+1}]$ as a union of $\kappa(2^{m+1}) + 1$ big blocks and $\kappa(2^{m+1}) + 1$ small blocks. We further denote $i_{m,j}$ is smallest element of $I_{m,j}$, $\mathcal{I}(m) = \bigcup_j I_{m,j}$ and $\mathcal{J}(m) = \bigcup_j J_{m,j}$. Let the block sums defined by

$$Y_{m,j} = \sum_{i \in I_{m,j}} X_i, \quad \text{and} \quad Z_{m,j} = \sum_{i \in J_{m,j}} X_i.$$

We introduce following lemmas to give the moment bounds for $Y_{m,j}$ and $Z_{m,j}$ which paving a way for showing that $Y_{m,j}$ is comparable with i.i.d. $\tilde{Y}_{m,j}$ distributed as $Y_{m,j}$ for $j = 1, \dots, \kappa(n)$ and small blocks $Z_{m,j}$ are negligible. Their proofs are given in [Appendix C](#).

Lemma 2.3. *Under the condition of [Theorem 1.2](#), for any $2 \leq p' < p$ and $j = 1, \dots, \kappa(2^{m+1}) + 1$, we have*

$$\mathbb{E}\|Y_{m,j}\|^{p'} \leq C m_1^{\frac{p'}{2}}, \quad \mathbb{E}\|Z_{m,j}\|^{p'} \leq C m_2^{\frac{p'}{2}}.$$

Lemma 2.4. *Under the condition of [Theorem 1.2](#), for any $2 < p' < p$, we can construct independent random variables $\tilde{Y}_{m,j}$ defined on a richer probability space $(\Omega_1, \mathcal{A}_1, \mathbb{P}_1)$ such that $\tilde{Y}_{m,j} \stackrel{\mathcal{D}}{=} Y_{m,j}$ for $j = 1, \dots, \kappa(2^{m+1})$ satisfies*

$$\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (Y_{m,j} - \tilde{Y}_{m,j}) \right\| = o(2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m}), \quad \text{a.s.} \quad (2.1)$$

Moreover,

$$\max_{1 \leq i \leq j} \left\| \sum_{\ell=1}^i (Y_{m,\ell} - \tilde{Y}_{m,\ell}) \right\| = o(2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m}), \quad \text{a.s.} \quad (2.2)$$

Lemma 2.5. *Under the condition of Theorem 1.2, for any $n \in [2^m + 1, 2^{m+1}]$, we have*

$$\max_{2^m+1 \leq i \leq n} \left\| \sum_{\ell \in \mathcal{J}(m) \cap [0, i]} \mathbf{X}_\ell \right\| = o(2^{\frac{1}{2}(1-\alpha_1)m} \log 2^m), \quad a.s.$$

Moreover, for any i.i.d. centered Gaussian random vectors η_ℓ ,

$$\max_{2^m+1 \leq i \leq n} \left\| \sum_{\ell \in \mathcal{J}(m) \cap [0, i]} \eta_\ell \right\| = o(2^{\frac{1}{2}(1-\alpha_1)m} \log 2^m), \quad a.s.$$

3. Proof of Theorem 1.2

The strategy for proving Theorem 1.2 is to decompose the sum $\sum_{i=1}^n \mathbf{X}_i$ into two parts using the block technique, where the contribution of small blocks is negligible as shown in Lemma 2.5. For the big blocks, the projection of these random variables onto the subspace $\mathbb{H}_{\leq d}$ is comparable to Gaussian random variables according to [14], while the residual on $\mathbb{H}_{> d}$ can be directly bounded. The dimension d will be chosen carefully with respect to n . In this section, we first introduce the necessary lemmas to establish the Gaussian approximation for the big blocks, and then conclude with the proof of Theorem 1.2.

Under Assumption (A2), there exists an orthonormal operator $\mathbf{U}$ such that the long-run covariance operator of the transformed sequence $(\mathbf{U}\mathbf{X}_k)_{k \in \mathbb{N}_0}$, given by $\bar{\mathbf{\Gamma}} = \mathbf{U}\mathbf{\Gamma}\mathbf{U}^T$, is a diagonal matrix $\text{diag}(\lambda_1, \lambda_2, \dots)$. By the unitary invariance of the norm, the almost sure invariance principle analysis for the original sequence $(\mathbf{X}_k)_{k \in \mathbb{N}_0}$ is equivalent to that of $(\mathbf{U}\mathbf{X}_k)_{k \in \mathbb{N}_0}$. Thus, without loss of generality, we assume hereinafter that $\mathbf{\Gamma}$ is diagonal in this section.

Lemma 3.1. *Under the conditions of Theorem 1.2, one has*

$$\|\text{cov}(\sum_{i=1}^n \mathbf{X}_i) - n\mathbf{\Gamma}\|_F \leq C, \quad (3.1)$$

where C is a positive constant depends on p and β . For the diagonal elements of the matrix $\text{cov}(\sum_{i=1}^n \mathbf{X}_i)$, we further have

$$\mathbb{E} \left[\sum_{i=1}^n X_{i,\ell} \right]^2 - n\lambda_\ell = r_\ell, \quad (3.2)$$

where $X_{i,\ell}$ is the ℓ -th element of $\mathbf{X}_i$ and r_ℓ is a square summable constant depends on ℓ .

Proof. A straight calculation yields that

$$\text{cov}(\sum_{i=1}^n \mathbf{X}_i) = \sum_{i=1}^n \text{cov}(\mathbf{X}_i) + \sum_{1 \leq i < j \leq n} (\text{cov}(\mathbf{X}_i, \mathbf{X}_j) + \text{cov}(\mathbf{X}_j, \mathbf{X}_i)^T).$$

Since $(\mathbf{X}_i)_{0 \leq i \leq n}$ is stationary, there exist linear operators $\mathbf{\Gamma}_i$ for $i = 0, 1, 2, \dots$ such that

$$\text{cov}(\mathbf{X}_i) = \mathbf{\Gamma}_0 \quad \text{cov}(\mathbf{X}_i, \mathbf{X}_j) = \mathbb{E}[\mathbf{X}_i \mathbf{X}_j^T] = \mathbb{E}[\mathbf{X}_0 \mathbf{X}_{j-i}^T] = \mathbf{\Gamma}_{j-i}.$$

We further denote $\mathbf{\Gamma} = \mathbf{\Gamma}_0 + \sum_{k=1}^{\infty} (\mathbf{\Gamma}_k + \mathbf{\Gamma}_k^T)$. Thus,

$$\begin{aligned}
& \left\| \text{cov}\left(\sum_{i=1}^n \mathbf{X}_i\right) - n\mathbf{\Gamma} \right\|_F \\
&= \left\| \sum_{i=1}^n \text{cov}(\mathbf{X}_i) + \sum_{1 \leq i < j \leq n} (\text{cov}(\mathbf{X}_i, \mathbf{X}_j) + \text{cov}(\mathbf{X}_i, \mathbf{X}_j)^T) - n\mathbf{\Gamma}_0 - n \sum_{k=1}^{\infty} (\mathbf{\Gamma}_k + \mathbf{\Gamma}_k^T) \right\|_F \\
&= \left\| \sum_{k=1}^n (n-k)(\mathbf{\Gamma}_k + \mathbf{\Gamma}_k^T) - n \sum_{k=1}^{\infty} (\mathbf{\Gamma}_k + \mathbf{\Gamma}_k^T) \right\|_F \\
&\leq \left\| n \sum_{k=n+1}^{\infty} (\mathbf{\Gamma}_k + \mathbf{\Gamma}_k^T) \right\|_F + \left\| \sum_{k=1}^n k(\mathbf{\Gamma}_k + \mathbf{\Gamma}_k^T) \right\|_F.
\end{aligned}$$

To finish the proof, we show that $\|\text{cov}(\mathbf{X}_0, \mathbf{X}_k)\|_F = \|\mathbf{\Gamma}_k\|_F \leq Ce^{-ck}$ and this gives the bound of (3.1). Following Lemma 2.1 and Hölder's inequality, one has

$$\begin{aligned}
\|\text{cov}(\mathbf{X}_0, \mathbf{X}_k)\|_F &\leq 18 \int_0^{\alpha(k)} Q_{\|\mathbf{X}_0\|}^2(u) du \\
&\leq 18 \left(\int_0^1 1_{\{u \leq \alpha(k)\}} du \right)^{1-2/p} \left(\int_0^1 Q_{\|\mathbf{X}_0\|}^p(u) du \right)^{2/p} \\
&\leq Ce^{-k\beta(1-2/p)} (\mathbb{E}\|\mathbf{X}\|^p)^{2/p}.
\end{aligned}$$

Thus, we obtain

$$\left\| \text{cov}\left(\sum_{i=1}^n \mathbf{X}_i\right) - n\mathbf{\Gamma} \right\|_F \leq C(\mathbb{E}\|\mathbf{X}\|^p)^{2/p}.$$

Notice that $\mathbf{\Gamma}$ is diagonal with element $\lambda_1, \lambda_2, \dots$ and with the form $\mathbf{\Gamma} = \mathbf{\Gamma}_0 + \sum_{k=1}^{\infty} (\mathbf{\Gamma}_k + \mathbf{\Gamma}_k^T)$, thus, for each element of $\mathbf{\Gamma}$, we have for the ℓ -th element in the diagonal

$$\mathbb{E}[X_{0,\ell}^2] + 2 \sum_{k=1}^{\infty} \mathbb{E}[X_{0,\ell} X_{k,\ell}] = \lambda_{\ell}, \quad (3.3)$$

and for the remaining elements with $i \neq j$ and $i, j = 1, 2, \dots$,

$$\mathbb{E}[X_{0,i} X_{0,j}] + \sum_{k=1}^{\infty} (\mathbb{E}[X_{0,i} X_{k,j}] + \mathbb{E}[X_{0,j} X_{k,i}]) = 0. \quad (3.4)$$

Following (3.1) and the definition of Frobenius norm, it is easy to see that

$$\sum_{\ell=1}^{\infty} \left(\mathbb{E} \left[\sum_{i=1}^n X_{i,\ell} \right]^2 - n\lambda_{\ell} \right)^2 \leq C,$$

which implies there exists a square summable number r_{ℓ} such that

$$\mathbb{E} \left[\sum_{i=1}^n X_{i,\ell} \right]^2 - n\lambda_{\ell} = r_{\ell}.$$

□

For any vectors $\xi_0, \dots, \xi_n$ defined on $\mathbb{H}$, let

$$\xi_i^{(d)} = (\xi_{i,1}, \dots, \xi_{i,d})^T \in \mathbb{R}^d, \quad \xi_i^{[d]} = (0, \dots, 0, \xi_{i,d+1}, \xi_{i,d+2}, \dots)^T \in \mathbb{H}.$$

Lemma 3.2. *Under the conditions of Theorem 1.2 and $m_1 \lambda_k > Ck + c$ for $k \in \mathbb{N}$ and constants C and c , let $\lambda_{\max}$ and $\lambda_{\min}$ be the largest and smallest eigenvalue of $\text{cov}(\sum_{i=1}^{m_1} X_i^{(d)})$ respectively, then one has*

$$\lambda_{\max} \leq C m_1 \lambda_1 \quad \lambda_{\min} \geq c m_1 \lambda_d.$$

Proof. Since $X_i = ((X_i^{(d)})^T, 0, \dots)^T + X_i^{[d]}$, a straight calculation implies that $\text{cov}(\sum_{i=1}^{m_1} X_i^{(d)})$ is the upper left $d \times d$ block of $\text{cov}(\sum_{i=1}^{m_1} X_i)$. Following [13, Corollary II.A.3], we have

$$\begin{aligned} \lambda_{\max} &\leq \max_{k \in \{1, \dots, d\}} \sum_{j=1}^d \left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \sum_{i=1}^{m_1} X_{i,j} \right] \right| \\ \lambda_{\min} &\geq \min_{k \in \{1, \dots, d\}} \left(\left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \right]^2 \right| - \sum_{j=1, j \neq k}^d \left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \sum_{i=1}^{m_1} X_{i,j} \right] \right| \right). \end{aligned}$$

Since

$$\begin{aligned} &\sum_{j=1, j \neq k}^d \left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \sum_{i=1}^{m_1} X_{i,j} \right] \right| \\ &= \sum_{j=1, j \neq k}^d \left| \sum_{i=1}^{m_1} \mathbb{E}[X_{i,k} X_{i,j}] + \mathbb{E} \left[\sum_{1 \leq i < r \leq m_1} X_{i,k} X_{r,j} + X_{i,j} X_{r,k} \right] \right| \\ &= \sum_{j=1, j \neq k}^d \left| m_1 \mathbb{E}[X_{0,k} X_{0,j}] + \sum_{i=1}^{m_1} (m_1 - i) \mathbb{E}[X_{0,k} X_{i,j} + X_{0,j} X_{i,k}] \right|, \end{aligned}$$

(3.4) and Lemma 2.1 yield

$$\begin{aligned} &\sum_{j=1, j \neq k}^d \left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \sum_{i=1}^{m_1} X_{i,j} \right] \right| \\ &= \sum_{j=1, j \neq k}^d \left| \sum_{i=m_1+1}^{\infty} m_1 \mathbb{E}[X_{0,k} X_{i,j} + X_{0,j} X_{i,k}] + \sum_{i=1}^{m_1} i \mathbb{E}[X_{0,k} X_{i,j} + X_{0,j} X_{i,k}] \right| \\ &\leq C(d-1). \end{aligned}$$

Combining with (3.2) and the fact $(r_k)_{k \geq 0}^2$ is a sequence of summable constants, we obtain

$$\begin{aligned} \lambda_{\max} &\leq \max_{k \in \{1, \dots, d\}} \sum_{j=1}^d \left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \sum_{i=1}^{m_1} X_{i,j} \right] \right| \\ &\leq \max_{k \in \{1, \dots, d\}} (C(d-1) + m_1 \lambda_k + r_k) \leq Cd + m_1 \lambda_1 + c \end{aligned}$$

$$\begin{aligned}\lambda_{\min} &\geq \min_{k \in \{1, \dots, d\}} \left(\left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \right]^2 \right| - \sum_{j=1, j \neq k}^d \left| \mathbb{E} \left[\sum_{i=1}^{m_1} X_{i,k} \sum_{i=1}^{m_1} X_{i,j} \right] \right| \right) \\ &\geq \min_{k \in \{1, \dots, d\}} (m_1 \lambda_k + r_k - C(d-1)) \geq m_1 \lambda_d - Cd + c.\end{aligned}$$

Since $m_1 \lambda_1 \geq m_1 \lambda_d > Cd + c$, thus

$$\lambda_{\max} \leq Cm_1 \lambda_1, \quad \lambda_{\min} \geq cm_1 \lambda_d.$$

□

Lemma 3.3. *Under the conditions of Theorem 1.2 and $m_1 \lambda_k > Ck + c$ for $k \in \mathbb{N}$. For any $2 \leq p' < p$ and $\tilde{\mathbf{Y}}_{m,j}$, one can construct on a probability space $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ a sequence of independent random vectors $(\tilde{\mathbf{Y}}_{m,j})^*$ and the corresponding sequence of independent Gaussian random vectors*

$$\boldsymbol{\eta}_{m,j} = ((\boldsymbol{\eta}_{m,j}^{(d)})^T, 0, \dots) + \boldsymbol{\eta}_{m,j}^{[d]},$$

for $1 \leq j \leq \kappa(2^{m+1})$ so that

$$\tilde{\mathbf{Y}}_{m,j} \stackrel{\mathcal{D}}{=} (\tilde{\mathbf{Y}}_{m,j})^*, \quad \boldsymbol{\eta}_{m,j}^{[d]} \sim \mathcal{N}(\mathbf{0}, \text{cov}(\tilde{\mathbf{Y}}_{m,j}^{[d]})), \quad \boldsymbol{\eta}_{m,j}^{(d)} \sim \mathcal{N}(\mathbf{0}, \text{cov}(\tilde{\mathbf{Y}}_{m,j}^{(d)}))$$

and

$$\begin{aligned}&\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\tilde{\mathbf{Y}}_{m,j})^* - \boldsymbol{\eta}_{m,j} \right\| \\ &= o \left(A_d^{\frac{1}{p'}} \lambda_d^{-\frac{1}{2}} 2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m} + 2^{m/2} d^{(1-\delta_2)/2} \log 2^m \right) \quad a.s.\end{aligned}$$

Proof. It is easy to see that

$$\tilde{\mathbf{Y}}_{m,j} = ((\tilde{\mathbf{Y}}_{m,j}^{(d)})^T, 0, 0, \dots)^T + \tilde{\mathbf{Y}}_{m,j}^{[d]}.$$

To compare $(\tilde{\mathbf{Y}}_{m,j})_{1 \leq j \leq \kappa(2^{m+1})}$ with Gaussian random vectors on Hilbert space $\mathbb{H}$, we show that $\tilde{\mathbf{Y}}_{m,j}^{(d)}$ are comparable with Gaussian random variables $\boldsymbol{\eta}_{m,j}^{(d)}$ on $\mathbb{R}^d$ space following [14, Theorem 2] and $\tilde{\mathbf{Y}}_{m,j}^{[d]}$ are negligible.

Step 1. We show that for $\tilde{\mathbf{Y}}_{m,j}^{(d)}$ one can construct on a probability space $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ a sequence of independent random vectors $(\tilde{\mathbf{Y}}_{m,j}^{(d)})^* \stackrel{\mathcal{D}}{=} \tilde{\mathbf{Y}}_{m,j}^{(d)}$ and the corresponding sequence of independent Gaussian random vectors $\boldsymbol{\eta}_{m,j}^{(d)} \sim \mathcal{N}(\mathbf{0}, \text{cov}(\tilde{\mathbf{Y}}_{m,j}^{(d)}))$ for $1 \leq j \leq \kappa(2^{m+1})$ such that

$$\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i ((\tilde{\mathbf{Y}}_{m,j}^{(d)})^* - \boldsymbol{\eta}_{m,j}^{(d)}) \right\| = o \left(A_d^{\frac{1}{p'}} \lambda_d^{-\frac{1}{2}} 2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m} \right), \quad a.s., \quad (3.6)$$

where $A_d = C \max\{d^{11p'}, d^{\frac{p'(p'+2)}{4}} (\log d)^{\frac{p'(p'+1)}{2}}\}$ and C is a constant depends on p' .

For brevity, instead of writing out the properties of $(\tilde{Y}_{m,j}^{(d)})^*$ and $\eta_{m,j}^{(d)}$ listed above we simply say that *there is a construction of $\tilde{Y}_{m,j}^{(d)}$* to show that one can construct a coupling between $\tilde{Y}_{m,j}^{(d)}$ and Gaussian random variables on a probability space enjoying the mentioned additional properties accordingly for $1 \leq j \leq \kappa(2^{m+1})$.

Let $\lambda_{\max,Y}$ and $\lambda_{\min,Y}$ be the maximal and minimal strictly positive eigenvalues of the covariance matrix $\text{cov}(\tilde{Y}_{m,j}^{(d)})$ respectively. According to [14, Theorem 2], there is a construction of $\tilde{Y}_{m,j}^{(d)}$ on probability space $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ such that

$$\begin{aligned} \mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i \left((\tilde{Y}_{m,j}^{(d)})^* - \eta_{m,j}^{(d)} \right) \right\|^{p'} \right] &\leq A_d (\lambda_{\max,Y} / \lambda_{\min,Y})^{\frac{p'}{2}} \kappa(2^{m+1}) \mathbb{E} \|\tilde{Y}_{m,j}\|^{p'} \\ &\leq CA_d (\lambda_{\max,Y} / \lambda_{\min,Y})^{\frac{p'}{2}} \kappa(2^{m+1}) m_1^{\frac{p'}{2}}. \end{aligned}$$

Following Lemma 3.2, we obtain

$$\mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i \left((\tilde{Y}_{m,j}^{(d)})^* - \eta_{m,j}^{(d)} \right) \right\|^{p'} \right] \leq CA_d \lambda_d^{-\frac{p'}{2}} \kappa(2^{m+1}) m_1^{\frac{p'}{2}}.$$

Thus, we have

$$\begin{aligned} \sum_{m=1}^{\infty} \mathbb{P} \left(\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i \left((\tilde{Y}_{m,j}^{(d)})^* - \eta_{m,j}^{(d)} \right) \right\| \geq A_d^{\frac{1}{p'}} \lambda_d^{-\frac{1}{2}} 2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m} \right) \\ \leq \sum_{m=1}^{\infty} \mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i \left((\tilde{Y}_{m,j}^{(d)})^* - \eta_{m,j}^{(d)} \right) \right\|^{p'} \right] A_d^{-1} \lambda_d^{\frac{p'}{2}} 2^{-((1-\alpha_1) + \alpha_1 p'/2)m} (\log 2^m)^{-\frac{p'}{2}} \\ \leq \sum_{m=1}^{\infty} C (\log 2^m)^{-\frac{p'}{2}} < \infty. \end{aligned}$$

By the Borel–Cantelli lemma, we obtain

$$\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i \left((\tilde{Y}_{m,j}^{(d)})^* - \eta_{m,j}^{(d)} \right) \right\| = o \left(A_d^{\frac{1}{p'}} \lambda_d^{-\frac{1}{2}} 2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m} \right), \quad a.s.$$

Step 2. We show that there is a construction of $\tilde{Y}_{m,j}^{[d]}$ on probability space $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ such that

$$\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i \left((\tilde{Y}_{m,j}^{[d]})^* - \eta_{m,j}^{[d]} \right) \right\| = o(2^{(1-\alpha_1)m/2} (1 + 2^{\alpha_1 m} d^{1-\delta_2})^{1/2} \log 2^m) \quad a.s. \quad (3.7)$$

According to [14, (25,28,29)], one has

$$\mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i \left((\tilde{Y}_{m,j}^{[d]})^* - \eta_{m,j}^{[d]} \right) \right\|^2 \right]$$

$$\begin{aligned}
&\leq C\mathbb{E}\left[\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i(\tilde{\mathbf{Y}}_{m,j}^{[d]})^*\right\|^2\right]+C\mathbb{E}\left[\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\boldsymbol{\eta}_{m,j}^{[d]}\right\|^2\right] \\
&\leq C\sum_{j=1}^{\kappa(2^{m+1})}\mathbb{E}\left\|\tilde{\mathbf{Y}}_{m,j}^{[d]}\right\|^2.
\end{aligned}$$

By (3.2) and the condition $m_1\lambda_k > Ck + c \geq r_k$ for $k \in \mathbb{N}$, we can get

$$\begin{aligned}
\mathbb{E}\left[\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\left((\tilde{\mathbf{Y}}_{m,j}^{[d]})^*-\boldsymbol{\eta}_{m,j}^{[d]}\right)\right\|^2\right] &\leq C\kappa(2^{m+1})\sum_{k=d+1}^{\infty}\mathbb{E}\left[\sum_{i=1}^{m_1}X_{i,k}\right]^2 \\
&= C\kappa(2^{m+1})\sum_{k=d+1}^{\infty}(m_1\lambda_k+r_k) \\
&\leq C2^m d^{1-\delta_2}.
\end{aligned}$$

Thus, the Markov inequality yields

$$\begin{aligned}
&\sum_{m=1}^{\infty}\mathbb{P}\left(\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\left((\tilde{\mathbf{Y}}_{m,j}^{[d]})^*-\boldsymbol{\eta}_{m,j}^{[d]}\right)\right\|\geq 2^{m/2}d^{(1-\delta_2)/2}\log 2^m\right) \\
&\leq \sum_{m=1}^{\infty}\mathbb{E}\left[\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\left((\tilde{\mathbf{Y}}_{m,j}^{[d]})^*-\boldsymbol{\eta}_{m,j}^{[d]}\right)\right\|^2\right]2^{-m}d^{-1+\delta_2}\log^{(-2)}2^m \\
&< \infty.
\end{aligned}$$

By the Borel–Cantelli lemma, we obtain

$$\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\left((\tilde{\mathbf{Y}}_{m,j}^{[d]})^*-\boldsymbol{\eta}_{m,j}^{[d]}\right)\right\|=o\left(2^{m/2}d^{(1-\delta_2)/2}\log 2^m\right), \quad a.s.$$

Step 3. Combining the estimates in step 1 and step 2 above, i.e., (3.6) and (3.7), one has

$$\begin{aligned}
&\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\left((\tilde{\mathbf{Y}}_{m,j})^*-\boldsymbol{\eta}_{m,j}\right)\right\| \\
&\leq \max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\left((\tilde{\mathbf{Y}}_{m,j}^{(d)})^*-\boldsymbol{\eta}_{m,j}^{(d)}\right)\right\|+\max_{1\leq i\leq \kappa(2^{m+1})}\left\|\sum_{j=1}^i\left((\tilde{\mathbf{Y}}_{m,j}^{[d]})^*-\boldsymbol{\eta}_{m,j}^{[d]}\right)\right\| \\
&=o\left(A_d^{\frac{1}{p'}}\lambda_d^{-\frac{1}{2}}2^{((1-\alpha_1)/p'+\alpha_1/2)m}\sqrt{\log 2^m}+2^{m/2}d^{(1-\delta_2)/2}\log 2^m\right) \quad a.s.
\end{aligned}$$

□

Proof of Theorem 1.2. For any $i \in [2^m + 1, 2^{m+1}]$, it is easy to see that

$$\sum_{\ell=1}^i \mathbf{X}_{\ell} = \sum_{j=1}^{\kappa(i)} \mathbf{Y}_{m,j} + \sum_{\ell \in I_{m,\kappa(i)+1} \cap [2^m+1,i]} \mathbf{X}_{\ell} + \sum_{\ell \in \mathcal{J}(m) \cap [2^m+1,i]} \mathbf{X}_{\ell}.$$

Recall that $i_{m,j}$ is the smallest element of $I_{m,j}$, following Lemma 2.5, we can get

$$\begin{aligned}
& \max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i \mathbf{X}_\ell - \sum_{j=1}^{\kappa(i)} \mathbf{Y}_{m,j} \right\| \\
& \leq \max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell \in I_{m,\kappa(i)+1} \cap [2^m+1, i]} \mathbf{X}_\ell \right\| + \max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell \in \mathcal{J}(m) \cap [2^m+1, i]} \mathbf{X}_\ell \right\| \\
& = \max_{1 \leq j \leq \kappa(2^{m+1})} \max_{1 \leq i < |I_{m,j}|} \left\| \sum_{\ell=i_{m,j}}^{i_{m,j}+i} \mathbf{X}_\ell \right\| + o(2^{\frac{1}{2}(1-\alpha_1)m} \log 2^m), \quad \text{a.s.} \tag{3.8}
\end{aligned}$$

Let p' be a positive constant such that $2 < p' < p$. For the first term, [37, Proposition 1] and Lemma 2.3 imply

$$\left(\mathbb{E} \left[\max_{1 \leq i \leq 2^r} \left\| \sum_{j=1}^i \mathbf{X}_j \right\|^{p'} \right] \right)^{\frac{1}{p'}} \leq \sum_{i=0}^r 2^{(r-i)/p'} \left(\mathbb{E} \left\| \sum_{j=1}^{2^i} \mathbf{X}_j \right\|^{p'} \right)^{\frac{1}{p'}} \leq C \sum_{i=0}^r 2^{(r-i)/p'} (2^i)^{\frac{1}{2}} \leq C 2^{\frac{r}{2}},$$

which yields

$$\mathbb{E} \left[\max_{1 \leq i < |I_{m,j}|} \left\| \sum_{\ell=i_{m,j}}^{i_{m,j}+i} \mathbf{X}_\ell \right\|^{p'} \right] \leq |I_{m,j}|^{\frac{p'}{2}} \leq 2^{\frac{p'}{2}\alpha_1 m}.$$

Thus, the Markov inequality implies

$$\begin{aligned}
& \sum_{m=1}^{\infty} \mathbb{P} \left(\max_{1 \leq j \leq \kappa(2^{m+1})} \max_{1 \leq i < |I_{m,j}|} \left\| \sum_{\ell=i_{m,j}}^{i_{m,j}+i} \mathbf{X}_\ell \right\| \geq 2^{((1-\alpha_1)/p' + \alpha_1/2)m} (\log 2^m)^{1/2} \right) \\
& \leq \sum_{m=1}^{\infty} \mathbb{E} \left[\max_{1 \leq j \leq \kappa(2^{m+1})} \max_{1 \leq i < |I_{m,j}|} \left\| \sum_{\ell=i_{m,j}}^{i_{m,j}+i} \mathbf{X}_\ell \right\|^{p'} \right] 2^{-(1-\alpha_1 + \alpha_1 p'/2)m} (\log 2^m)^{-p'/2} \\
& \leq \sum_{m=1}^{\infty} \sum_{j=1}^{\kappa(2^{m+1})} \mathbb{E} \left[\max_{1 \leq i < |I_{m,j}|} \left\| \sum_{\ell=i_{m,j}}^{i_{m,j}+i} \mathbf{X}_\ell \right\|^{p'} \right] 2^{-(1-\alpha_1 + \alpha_1 p'/2)m} (\log 2^m)^{-p'/2} \\
& \leq \sum_{m=1}^{\infty} 2^{(1-\alpha_1 + p'\alpha_1/2)m} 2^{-(1-\alpha_1 + p'\alpha_1/2)m} (\log 2^m)^{-p'/2} < \infty.
\end{aligned}$$

By the Borel–Cantelli lemma, we obtain

$$\max_{1 \leq j \leq \kappa(2^{m+1})} \max_{i < |I_{m,j}|} \left\| \sum_{\ell=i_{m,j}}^{i_{m,j}+i} \mathbf{X}_\ell \right\| = o(2^{((1-\alpha_1)/p' + \alpha_1/2)m} (\log 2^m)^{1/2}), \quad \text{a.s.} \tag{3.9}$$

Similar estimate holds for any i.i.d. centered Gaussian random vectors η_ℓ , i.e.,

$$\max_{1 \leq j \leq \kappa(2^{m+1})} \max_{i < |I_{m,j}|} \left\| \sum_{\ell=i_{m,j}}^{i_{m,j}+i} \eta_\ell \right\| = o(2^{((1-\alpha_1)/p' + \alpha_1/2)m} (\log 2^m)^{1/2}), \quad \text{a.s.} \tag{3.10}$$

Combining (3.8) and (3.9), we have

$$\max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i X_\ell - \sum_{j=1}^{\kappa(i)} Y_{m,j} \right\| = o \left((2^{\frac{1}{2}(1-\alpha_1)m} + 2^{((1-\alpha_1)/p' + \alpha_1/2)m}) \log 2^m \right), \quad a.s.$$

Following Lemma 2.4, on a richer probability space of $(\Omega, \mathcal{A}, \mathbb{P})$, one has

$$\begin{aligned} & \max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i X_\ell - \sum_{j=1}^{\kappa(i)} \tilde{Y}_{m,j} \right\| \\ &= \max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i X_\ell - \sum_{j=1}^{\kappa(i)} Y_{m,j} + \sum_{j=1}^{\kappa(i)} Y_{m,j} - \sum_{j=1}^{\kappa(i)} \tilde{Y}_{m,j} \right\| \\ &= o \left((2^{\frac{1}{2}(1-\alpha_1)m} + 2^{((1-\alpha_1)/p' + \alpha_1/2)m}) \log 2^m \right), \quad a.s. \end{aligned} \quad (3.11)$$

Lemma 3.3 implies that there is a construction of $\tilde{Y}_{m,j}$ on probability space $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ and one can compare $\tilde{Y}_{m,j}$ with i.i.d. Gaussian random variables $\eta_{m,j}$. Now we regularize $\eta_{m,j}$, that is, replacing the covariance matrix $\text{cov}(\eta_{m,j})$ by the linear form $|\mathbf{I}_{m,j}|\mathbf{\Gamma}$. We denote $\tilde{\eta}_{m,j} := \mathcal{N}(\mathbf{0}, |\mathbf{I}_{m,j}|\mathbf{\Gamma})$.

Following Lemma 3.1 and for large enough m , one has

$$\text{cov}(\eta_{m,j}) + \mathbf{M}_{m,j} = |\mathbf{I}_{m,j}|\mathbf{\Gamma} + \mathbf{N}_{m,j}, \quad (3.12)$$

where $\mathbf{M}_{m,j}$ and $\mathbf{N}_{m,j}$ are positive definite linear operators with $\|\mathbf{M}_{m,j}\|_F \leq C$, $\|\mathbf{N}_{m,j}\|_F \leq C$. Therefore, $\mathcal{N}(\mathbf{0}, |\mathbf{I}_{m,j}|\mathbf{\Gamma} + \mathbf{N}_{m,j})$ is the sum of $\tilde{\eta}_{m,j}$ and an independent random variable $\mathcal{N}(\mathbf{0}, \mathbf{N}_{m,j})$. On the other hand, (3.12) implies $\mathcal{N}(\mathbf{0}, |\mathbf{I}_{m,j}|\mathbf{\Gamma} + \mathbf{N}_{m,j})$ is also the sum of $\eta_{m,j}$ and an independent random variable $\mathcal{N}(\mathbf{0}, \mathbf{M}_{m,j})$. Using [2, Lemma A1], we obtain a coupling between $\eta_{m,j}$ and $\tilde{\eta}_{m,j}$ such that the difference $\mathbf{D}_{m,j} = \eta_{m,j} - \tilde{\eta}_{m,j}$ is centered and $\mathbb{E}\|\mathbf{D}_{m,j}\|^2 \leq C$.

Thus, following the Lévy inequality, see [19, 5.4.a], we have

$$\begin{aligned} & \mathbb{P} \left(\max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{j=1}^{\kappa(i)} \mathbf{D}_{m,j} \right\| > 2^{\frac{1}{2}(1-\alpha_1)m} \log 2^m \right) \\ & \leq 2\mathbb{P} \left(\left\| \sum_{j=1}^{\kappa(2^{m+1})} \mathbf{D}_{m,j} \right\| > 2^{\frac{1}{2}(1-\alpha_1)m} \log 2^m \right) \\ & \leq 2\mathbb{E} \left\| \sum_{j=1}^{\kappa(2^{m+1})} \mathbf{D}_{m,j} \right\|^2 2^{-(1-\alpha_1)m} (\log 2^m)^{-2} \\ & \leq C\kappa(2^{m+1}) 2^{-(1-\alpha_1)m} (\log 2^m)^{-2}, \end{aligned}$$

which is summable with respect to m . Then there is a construction for $\eta_{m,j}$ on probability space $(\Omega_3, \mathcal{A}_3, \mathbb{P}_3)$ such that

$$\max_{2^m+1 \leq i \leq 2^{m+1}} \left\| \sum_{j=1}^{\kappa(i)} (\eta_{m,j}^* - \tilde{\eta}_{m,j}^*) \right\| = o(2^{\frac{1}{2}(1-\alpha_1)m} \log 2^m), \quad a.s.. \quad (3.13)$$

Combining (3.13) with (3.3) and [3, Lemma 4.1], there is a construction for $\tilde{\mathbf{Y}}_{m,j}^*$ on probability space $(\Omega_4, \mathcal{A}_4, \mathbb{P}_4)$ such that

$$\begin{aligned} & \max_{2^{m+1} \leq i \leq 2^{m+1}} \left\| \sum_{j=1}^{\kappa(i)} (\tilde{\mathbf{Y}}_{m,j}^{**} - \tilde{\boldsymbol{\eta}}_{m,j}^{**}) \right\| \\ &= o \left(\left(A_d^{\frac{1}{p'}} \lambda_d^{-\frac{1}{2}} 2^{(\frac{1-\alpha_1}{p'} + \frac{\alpha_1}{2})m} + 2^{m/2} d^{(1-\delta_2)/2} + 2^{(1-\alpha_1)m/2} \right) \log 2^m \right) \quad a.s. \end{aligned} \quad (3.14)$$

Using [3, Lemma 4.1] again with (3.10), (3.11), (3.14) and Lemma 2.5, we can finally construct a probability space $(\Omega_{(m)}, \mathcal{A}_{(m)}, \mathbb{P}_{(m)})$ on which we can define $\mathbf{X}_\ell^*$ distributed as $\mathbf{X}_\ell$ and i.i.d. Gaussian random variables $\boldsymbol{\eta}_\ell \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma})$ such that

$$\begin{aligned} & \max_{2^{m+1} \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i \mathbf{X}_\ell^* - \sum_{\ell=1}^i \boldsymbol{\eta}_\ell \right\| \\ &= o \left(\left(A_d^{\frac{1}{p'}} \lambda_d^{-\frac{1}{2}} 2^{(\frac{1-\alpha_1}{p'} + \frac{\alpha_1}{2})m} + 2^{m/2} d^{(1-\delta_2)/2} + 2^{(1-\alpha_1)m/2} \right) \log 2^m \right) \quad a.s. \end{aligned}$$

Recall the definition of A_d and condition in Lemma 3.3,

$$A_d^{\frac{1}{p'}} = C \max\{d^{11}, d^{\frac{p'+2}{4}} (\log d)^{\frac{p'+1}{2}}\}, \quad m_1 \lambda_k > Ck + c,$$

and $C_1 d^{-\delta_1} \leq \lambda_d \leq C_2 d^{-\delta_2}$, $d = 2^{m\theta_{p'}}$. We take $\alpha_1 = (1 + \delta_1)\theta_{p'}$ and

$$\theta_{p'} = \min \left\{ \frac{p' - 2}{22p' - 2 + (2p' - 2)\delta_1 + p'\delta_2}, \frac{p' - 2}{p'(p' + 2)/2 - 2 + (2p' - 2)\delta_1 + p'\delta_2} \right\}$$

to ensure

$$A_d^{\frac{1}{p'}} d^{\frac{\delta_1}{2}} 2^{(\frac{1-\alpha_1}{p'} + \frac{\alpha_1}{2})m} = 2^{\frac{1-\alpha_1}{2}m} (1 + 2^{\alpha_1 m} d^{1-\delta_2})^{1/2}.$$

Then we have

$$\max_{2^{m+1} \leq i \leq 2^{m+1}} \left\| \sum_{\ell=2^{m+1}}^i (\mathbf{X}_\ell^* - \boldsymbol{\eta}_\ell) \right\| = o \left(2^{\frac{1}{2}m(1-(\delta_2-1)\theta_{p'})} (\log 2^m)^{\frac{p'+3}{2}} \right), \quad a.s..$$

Hence, we can construct probability space $(\Omega', \mathcal{A}', \mathbb{P}')$ on which

$$\max_{1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i (\mathbf{X}_\ell^* - \boldsymbol{\eta}_\ell) \right\| = o \left(2^{\frac{1}{2}m(1-(\delta_2-1)\theta_{p'})} (\log 2^m)^{\frac{p'+3}{2}} \right) \quad a.s.$$

For the logarithmic term, there exists a $2 < p'' < p'$ such that,

$$\max_{1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i (\mathbf{X}_\ell^* - \boldsymbol{\eta}_\ell) \right\| = o \left(2^{\frac{1}{2}m(1-(\delta_2-1)\theta_{p''})} \right) \quad a.s.$$

Let p'' close enough to p , we obtain

$$\max_{1 \leq i \leq 2^{m+1}} \left\| \sum_{\ell=1}^i (\mathbf{X}_\ell^* - \boldsymbol{\eta}_\ell) \right\| = o(2^{m\bar{\theta}}) \quad a.s.,$$

where

$$\bar{\theta} > \max \left\{ \frac{(2p-2)\delta_1 + 2\delta_2 + 23p-4}{44p-4 + (4p-4)\delta_1 + 2p\delta_2}, \frac{(2p-2)\delta_1 + 2\delta_2 + p(p+4)/2-4}{p(p+2)-4 + (4p-4)\delta_1 + 2p\delta_2} \right\}, \quad (3.15)$$

and consequently

$$\max_{1 \leq i \leq n} \left\| \sum_{\ell=1}^i (\mathbf{X}_\ell^* - \eta_\ell) \right\| = o(n^{\bar{\theta}}), \quad a.s..$$

□

Proof of Corollary 1.3. When $\delta_1 = \delta_2 = \delta$, it is easy to see that

$$\bar{\theta} > \max \left\{ \frac{2p\delta + 23p-4}{44p-4 + (6p-4)\delta}, \frac{2p\delta + p(p+4)/2-4}{p(p+2)-4 + (6p-4)\delta} \right\}, \quad (3.16)$$

which converges to $\frac{1}{3} + \frac{2}{3(3p-2)}$ as $\delta \rightarrow \infty$. That is, for any $\varepsilon > 0$, as $\delta > \bar{\delta}_{p,\varepsilon}$ where

$$\bar{\delta}_{p,\varepsilon} = \begin{cases} \frac{25p^2-54p+8-(44p-4)(3p-2)\varepsilon}{2(3p-2)^2\varepsilon}, & p \leq 42, \\ \frac{p^3/2+3p^2-12p'+8-(p^2+2p-4)(3p-2)\varepsilon}{2(3p-2)^2\varepsilon}, & p > 42. \end{cases} \quad (3.17)$$

we obtain

$$\max_{1 \leq i \leq n} \left\| \sum_{j=1}^i (\mathbf{X}_j^* - \eta_j) \right\| = o\left(n^{\frac{1}{3} + \frac{2}{3(3p-2)} + \varepsilon}\right), \quad a.s..$$

□

4. Examples

In this section, we give two examples where the first compares the mixing condition with the geometric moment contraction (GMC) condition of [3, Example 2.2] on $\mathbb{H}$ and the second considers the functional autoregressive processes.

4.1. Markov chain

Let $(\mathbf{X}_k)_{k \geq 0}$ be a $\mathbb{H}$ -valued time homogeneous Markov chain with p -th moment for $p > 2$ satisfying the condition:

(A3) $(\mathbf{X}_k)_{k \geq 0}$ is irreducible, aperiodic and Feller. There exists a Lyapunov function $V : \mathbb{H} \rightarrow [1, +\infty)$ such that

$$\mathbb{E}[V(\mathbf{X}_1)|\mathbf{X}_0 = \mathbf{x}] \leq \gamma V(\mathbf{x}) + K1_{\mathbf{C}}(\mathbf{x}), \quad (4.1)$$

where $0 < \gamma < 1$, $K > 0$ and $\mathbf{C}$ is a compact set.

It is easy to prove that $(\mathbf{X}_k)_{k \geq 0}$ is exponential ergodic with invariant measure π which yields exponential β -mixing, see [35, Theorem 2.1] and [11, Proposition 1]. We further assume that $\mathbf{X}_0 \sim \pi$ and Γ is a positive definite operator whose eigenvalues polynomial decay, then $(\mathbf{X}_k)_{k \geq 0}$ satisfies ASIP with rate (1.3).

Comparing (A3) with the contraction condition in [3],

$$\mathbb{E} \|\mathbf{X}_1^x - \mathbf{X}_1^y\|^p \leq r^p \|\mathbf{x} - \mathbf{y}\|^p, \quad (4.2)$$

where $0 < r < 1$, X_1^* denotes the Markov chain with initial value $X_0 = x$. The term $K1_C(x)$ of (A3) makes it be a weaker condition than the contraction condition.

4.2. Functional autoregressive processes

We consider the functional autoregressive processes

$$X_{k+1} = \mu + AX_k + B\epsilon_{k+1}, \quad (4.3)$$

where A and B are linear operators from $\mathbb{H}$ to $\mathbb{H}$ with kernel $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ respectively, ϵ_{k+1} is white noise. Let $(e_k)_{k \in \mathbb{N}}$ be an orthonormal basis of $L^2([0, \pi])$ with the form $e_k(x) = \sqrt{\frac{2}{\pi}} \sin(kx)$ and

$$Ae_k(x) = \lambda_k e_k(x),$$

where $Af(x) = \int_0^\pi a(s, x)f(s)ds$. According to Karhunen-Loève decomposition, one has

$$a(s, t) = \sum_{k=1}^{\infty} \lambda_k e_k(s) e_k(t).$$

(4.3) can be written as

$$X_{k+1}(\cdot) = \mu(\cdot) + \int_0^\pi a(\cdot, s)X_k(s)ds + \int_0^\pi b(\cdot, s)\epsilon_{k+1}(s)ds. \quad (4.4)$$

We refer the reader to [4, 36] for more details of functional autoregressive processes. We assume $\mu = 0$, $B^2 = A$ and A is symmetric for the simplify of calculation. Further assuming that $0 < \lambda_k \asymp k^{-\delta} < 1$, conditions (A1) and (A2) are satisfied and $(X_k)_{k \in \mathbb{N}_0}$ satisfies the ASIP.

Author contributions

CRedit: **Jianya Lu**: Conceptualization, Formal analysis, Methodology, Project administration, Writing – original draft, Writing – review & editing; **Wei Biao Wu**: Conceptualization, Supervision, Writing – review & editing; **Zhijie Xiao**: Conceptualization, Supervision, Writing – review & editing; **Lihu Xu**: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Disclosure statement

The authors declare no financial or non-financial competing interests.

Funding

The work of L. Xu was supported by the Science and Technology Development Fund (FDCT) of Macau S.A.R. under Grant FDCT 0074/2023/RIA2.

References

- [1] Berbee, H. (1987). Convergence rates in the strong law for bounded mixing sequences. *Probab. Th. Rel. Fields.* 74(2):255–270. DOI: [10.1007/BF00569992](https://doi.org/10.1007/BF00569992).

- [2] Berkes, I., Philipp, W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.* 7(1)pages :29–54. DOI: [10.1214/aop/1176995146](https://doi.org/10.1214/aop/1176995146).
- [3] Berkes, I., Liu, W., Wu, W. B. (2014). Komlós–Major–Tusnády approximation under dependence. *Ann. Probab.* 42(2):794–817. DOI: [10.1214/13-AOP850](https://doi.org/10.1214/13-AOP850).
- [4] Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*, Vol. 149. New York: Springer.
- [5] Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surveys.* 2:107–144. DOI: [10.1214/154957805100000104](https://doi.org/10.1214/154957805100000104).
- [6] Csáki, E., Csörgő, M. (1995). On additive functionals of markov chains. *J. Theor. Probab.* 8(4):905–919. DOI: [10.1007/BF02410117](https://doi.org/10.1007/BF02410117).
- [7] Cuny, C., Merlevède, F. (2014). On martingale approximations and the quenched weak invariance principle. *Ann. Probab.* 42(2):760–793. DOI: [10.1214/13-AOP856](https://doi.org/10.1214/13-AOP856).
- [8] Cuny, C., Dedecker, J., Korepanov, A., Merlevède, F. (2020a). Rates in almost sure invariance principle for slowly mixing dynamical systems. *Ergod. Th. Dynam. Sys.* 40(9):2317–2348. DOI: [10.1017/etds.2019.2](https://doi.org/10.1017/etds.2019.2).
- [9] Cuny, C., Dedecker, J., Korepanov, A., Merlevède, F. (2020b). Rates in almost sure invariance principle for quickly mixing dynamical systems. *Stoch. Dyn.* 20(01):2050002. DOI: [10.1142/S0219493720500021](https://doi.org/10.1142/S0219493720500021).
- [10] Davydov, Y. A. (1968). Convergence of distributions generated by stationary stochastic processes. *Theory Probab. Appl.* 13(4):691–696. DOI: [10.1137/1113086](https://doi.org/10.1137/1113086).
- [11] Davydov, Y. A. (1974). Mixing conditions for Markov chains. *Theory Probab. Appl.* 18(2):312–328. DOI: [10.1137/1118033](https://doi.org/10.1137/1118033).
- [12] Dedecker, J., Merlevède, F. (2010). On the almost sure invariance principle for stationary sequences of Hilbert-valued random variables. In: *Dependence in Probability, Analysis and Number Theory*. Heber City, UT: Kendrick Press, pp. 157–175.
- [13] Garren, K. R. (1968). *Bounds for the Eigenvalues of a Matrix*. Washington: National Aeronautics and Space Administration.
- [14] Götze, F., Zaitsev, A. Y. (2011). Estimates for the rate of strong approximation in Hilbert space. *Sib. Math. J.* 52(4):628–638. DOI: [10.1134/S0037446611040070](https://doi.org/10.1134/S0037446611040070).
- [15] Gouëzel, S. (2010). Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.* 38(4):1639–1671. DOI: [10.1214/10-AOP525](https://doi.org/10.1214/10-AOP525).
- [16] Hafouta, Y. (2023). An almost sure invariance principle for some classes of non-stationary mixing sequences. *Stat. Prob. Lett.* 193:109728. DOI: [10.1016/j.spl.2022.109728](https://doi.org/10.1016/j.spl.2022.109728).
- [17] Komlós, J., Major, P., Tusnády, G. (1975). An approximation of partial sums of independent R_v -s, and the sample df. i. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete.* 32(1):111–131.
- [18] Komlós, J., Major, P., Tusnády, G. (1976). An approximation of partial sums of independent R_v 's, and the sample df. ii. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete.* 34(1):33–58.
- [19] Lin, Z., Bai, Z. (2011). *Probability Inequalities*. Berlin: Springer.
- [20] Liu, W., Lin, Z. (2009). Strong approximation for a class of stationary processes. *Stochastic Processes Appl.* 119(1):249–280. DOI: [10.1016/j.spa.2008.01.012](https://doi.org/10.1016/j.spa.2008.01.012).
- [21] Miles, E. L., Benjamin Erichson, N., Mahoney, M. W. (2023). Bootstrapping the operator norm in high dimensions: error estimation for covariance matrices and sketching. *Bernoulli.* 29(1):428–450.
- [22] Merlevède, F., Rio, E. (2012). Strong approximation of partial sums under dependence conditions with application to dynamical systems. *Stochastic Processes Appl.* 122(1):386–417. DOI: [10.1016/j.spa.2011.08.012](https://doi.org/10.1016/j.spa.2011.08.012).
- [23] Merlevède, F., Rio, E. (2015). Strong approximation for additive functionals of geometrically ergodic markov chains. *Electron. J. Probab.* 20:1–27. DOI: [10.1214/EJP.v20-3746](https://doi.org/10.1214/EJP.v20-3746).
- [24] Merlevède, F., Peligrad, M., Utev, S. (1997). Sharp conditions for the clt of linear processes in a hilbert space. *J. Theor. Probab.* 10(3):681–693. DOI: [10.1023/A:1022653728014](https://doi.org/10.1023/A:1022653728014).
- [25] Monrad, D., Philipp, W. (1991). The problem of embedding vector-valued martingales in a Gaussian process. *Theory Probab. Appl.* 35(2):374–377. DOI: [10.1137/1135050](https://doi.org/10.1137/1135050).

- [26] Nummelin, E. (2004). *General Irreducible Markov Chains and Non-Negative Operators*, Number 83. Cambridge: Cambridge University Press.
- [27] Pengel, A., Bierkens, J. (2024). Strong invariance principles for ergodic Markov processes. *Electron. J. Statist.* 18(1):191–246. DOI: [10.1214/23-EJS2199](https://doi.org/10.1214/23-EJS2199).
- [28] Pinelis, I. (1994). Optimum bounds for the distributions of martingales in banach spaces. *Ann. Probab.* 22(4):1679–1706. DOI: [10.1214/aop/1176988477](https://doi.org/10.1214/aop/1176988477).
- [29] Reiß, M., Wahl, M. (2020). Nonasymptotic upper bounds for the reconstruction error of PCA. *Ann. Statist.* 48(2):1098–1123. DOI: [10.1214/19-AOS1839](https://doi.org/10.1214/19-AOS1839).
- [30] Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Annales de l'IHP Probabilités et Statistiques.* 29:587–597.
- [31] Shao, Q.-M., Lu, C. (1987). Strong approximations for partial sums of weakly dependent random variables. *Sci. Sinica Ser. A.* 30(6):575–587.
- [32] Shao, Q.-M., Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.* 24(4):2098–2127. DOI: [10.1214/aop/1041903220](https://doi.org/10.1214/aop/1041903220).
- [33] Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete.* 3(3):211–226. DOI: [10.1007/BF00534910](https://doi.org/10.1007/BF00534910).
- [34] Strassen, V. (1967). Almost sure behavior of sums of independent random variables and martingales. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 3, pp. 315. University of California Press.
- [35] Tuominen, P., Tweedie, R. L. (1994). Subgeometric rates of convergence of f-ergodic markov chains. *Adv. Appl. Probab.* 26(3):775–798. DOI: [10.2307/1427820](https://doi.org/10.2307/1427820).
- [36] Wang, D., Zhao, Z., Willett, R., Yau, C. Y. (2020). Functional autoregressive processes in reproducing kernel Hilbert spaces. arXiv Preprint. arXiv 2011.13993.
- [37] Wu, W. B. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* 35(6):2294–2320. DOI: [10.1214/009117907000000060](https://doi.org/10.1214/009117907000000060).
- [38] Yu Zaitsev, A. (1998). Multidimensional version of the results of komlós, major and tusnády for vectors with finite exponential moments. *Esaim: PS.* 2:41–108. DOI: [10.1051/ps:1998103](https://doi.org/10.1051/ps:1998103).
- [39] Yu Zaitsev, A. (2001). Multidimensional version of a result of sakhanenko in the invariance principle for vectors with finite exponential moments. i. *Theory Probab. Appl.* 45(4):624–641. DOI: [10.1137/S0040585X97978555](https://doi.org/10.1137/S0040585X97978555).
- [40] Yu Zaitsev, A. (2002a). Multidimensional version of a result of sakhanenko in the invariance principle for vectors with finite exponential moments. ii. *Theory Probab. Appl.* 46(3):490–514. DOI: [10.1137/S0040585X97979123](https://doi.org/10.1137/S0040585X97979123).
- [41] Yu Zaitsev, A. (2002b). Multidimensional version of a result of sakhanenko in the invariance principle for vectors with finite exponential moments. iii. *Theory Probab. Appl.* 46(4):676–698. DOI: [10.1137/S0040585X97979305](https://doi.org/10.1137/S0040585X97979305).
- [42] Yu Zaitsev, A. (2007). Estimates for the rate of strong approximation in the multidimensional invariance principle. *J. Math. Sci.* 145(2):4856–4865. DOI: [10.1007/s10958-007-0319-7](https://doi.org/10.1007/s10958-007-0319-7).

Appendix A: The proof of Lemma 2.2

Since $\alpha(n) \leq \beta(n)$, assumption (A1) implies

$$\alpha(n) \leq Ce^{-\beta n}. \quad (\text{A.1})$$

We first give following preparing lemma.

Lemma A.1. *Let $(\theta_i)_{1 \leq i \leq n}$ be a sequence of random variables on $\mathbb{H}$ with finite p -moment and let $\mathcal{F}_i = \sigma(\theta_j, j \leq i)$. Then for any $p \geq 2$, there exists constant C_p such that*

$$\mathbb{E} \left\| \sum_{i=1}^n \theta_i \right\|^p \leq C_p \left(\left(\sum_{i=1}^n \mathbb{E} \|\theta_i\|^2 \right)^{\frac{p}{2}} + \sum_{i=1}^n \mathbb{E} \|\theta_i\|^p + n^{p-1} \sum_{i=1}^n \mathbb{E} \|\mathbb{E}[\theta_i | \mathcal{F}_{i-1}]\|^p \right)$$

$$+ n^{\frac{p}{2}-1} \sum_{i=1}^n \mathbb{E} \left| \mathbb{E}[\|\boldsymbol{\theta}_i\|^2 | \mathcal{F}_{i-1}] - \mathbb{E} \|\boldsymbol{\theta}_i\|^2 \right|^{\frac{p}{2}} \Bigg).$$

Proof. The strategy is to construct martingale differences $\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]$ and using the Burkholder inequality to get the result. That is,

$$\begin{aligned} \mathbb{E} \left\| \sum_{i=1}^n \boldsymbol{\theta}_i \right\|^p &= \mathbb{E} \left\| \sum_{i=1}^n (\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}] + \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]) \right\|^p \\ &\leq 2^p \left(\mathbb{E} \left\| \sum_{i=1}^n (\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]) \right\|^p + n^{p-1} \sum_{i=1}^n \mathbb{E} \|\mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]\|^p \right). \end{aligned}$$

For the first term, [28, Theorem 4.1] implies

$$\begin{aligned} &\mathbb{E} \left\| \sum_{i=1}^n (\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]) \right\|^p \\ &\leq C_p \left(\sum_{i=1}^n \mathbb{E} \|\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]\|^p + \mathbb{E} \left[\sum_{i=1}^n \mathbb{E} [\|\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]\|^2 | \mathcal{F}_{i-1}] \right]^{\frac{p}{2}} \right) \\ &\leq C_p \left(\sum_{i=1}^n \mathbb{E} \|\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]\|^p + \mathbb{E} \left[\sum_{i=1}^n \mathbb{E} [\|\boldsymbol{\theta}_i\|^2 | \mathcal{F}_{i-1}] \right]^{\frac{p}{2}} \right) \\ &= C_p \left(\sum_{i=1}^n \mathbb{E} \|\boldsymbol{\theta}_i - \mathbb{E}[\boldsymbol{\theta}_i | \mathcal{F}_{i-1}]\|^p + \mathbb{E} \left[\sum_{i=1}^n 2\mathbb{E} [\|\boldsymbol{\theta}_i\|^2 | \mathcal{F}_{i-1}] - \mathbb{E} \|\boldsymbol{\theta}_i\|^2 + \mathbb{E} \|\boldsymbol{\theta}_i\|^2 \right]^{\frac{p}{2}} \right) \\ &\leq C_p \left(\sum_{i=1}^n \mathbb{E} \|\boldsymbol{\theta}_i\|^p + \left(\sum_{i=1}^n \mathbb{E} \|\boldsymbol{\theta}_i\|^2 \right)^{\frac{p}{2}} + \mathbb{E} \left(\sum_{i=1}^n (\mathbb{E} [\|\boldsymbol{\theta}_i\|^2 | \mathcal{F}_{i-1}] - \mathbb{E} \|\boldsymbol{\theta}_i\|^2) \right)^{\frac{p}{2}} \right) \\ &\leq C_p \left(\sum_{i=1}^n \mathbb{E} \|\boldsymbol{\theta}_i\|^p + \left(\sum_{i=1}^n \mathbb{E} \|\boldsymbol{\theta}_i\|^2 \right)^{\frac{p}{2}} + n^{\frac{p}{2}-1} \sum_{i=1}^n \mathbb{E} |\mathbb{E} [\|\boldsymbol{\theta}_i\|^2 | \mathcal{F}_{i-1}] - \mathbb{E} \|\boldsymbol{\theta}_i\|^2|^{\frac{p}{2}} \right). \end{aligned}$$

Thus, we can get the result. $\square$

Proof of Lemma 2.2. We first prove the case $p' = 2$ which is a extension of [30] on $\mathbb{H}$. The fact that $(\mathbf{X}_i)_{1 \leq i \leq n}$ is zero mean implies

$$\begin{aligned} \mathbb{E} \left\| \sum_{i=1}^n \mathbf{X}_i \right\|^2 &= \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq n} \mathbb{E} \langle \mathbf{X}_i, \mathbf{X}_j \rangle \\ &= \sum_{1 \leq i \leq n} \mathbb{E} \|\mathbf{X}_i\|^2 + 2 \sum_{1 \leq i < j \leq n} (\mathbb{E} \langle \mathbf{X}_i, \mathbf{X}_j \rangle - \langle \mathbb{E} \mathbf{X}_i, \mathbb{E} \mathbf{X}_j \rangle) \\ &\leq \sum_{1 \leq i \leq n} \mathbb{E} \|\mathbf{X}_i\|^2 + 2 \sum_{1 \leq i < j \leq n} 18 \int_0^{\alpha(j-i)} Q_{\|\mathbf{X}_i\|}^2(u) du, \end{aligned}$$

where the last line follows [Lemma 2.1](#) and $\alpha(j-i) \geq \bar{\alpha}$ therein. For the second term, a straight calculation yields

$$\begin{aligned} & \sum_{1 \leq i < j \leq n} \int_0^{\alpha(j-i)} Q_{\|X_i\|}^2(u) du \\ &= \sum_{1 \leq i < n} \int_0^1 \sum_{i < j \leq n} 1_{\{\alpha(j-i) > u\}} Q_{\|X_i\|}^2(u) du \\ &\leq \sum_{1 \leq i \leq n} \left(\int_0^1 \left(\sum_{i < j \leq n} 1_{\{\alpha(j-i) > u\}} \right)^{\frac{p}{p-2}} du \right)^{\frac{p-2}{p}} \left(\int_0^1 Q_{\|X_i\|}^p(u) du \right)^{\frac{2}{p}}. \end{aligned}$$

Since $\sum_{1 \leq j \leq n} 1_{\{\alpha(j) > u\}} = k$ if and only if $\alpha(k+1) \leq u < \alpha(k)$, then we have

$$\begin{aligned} \int_0^1 \left(\sum_{i < j \leq n} 1_{\{\alpha(j-i) > u\}} \right)^{\frac{p}{p-2}} du &\leq \int_0^1 \left(\sum_{j=1}^n 1_{\{\alpha(j) > u\}} \right)^{\frac{p}{p-2}} du \\ &= \sum_{k=1}^{\infty} \int_{\alpha(k+1)}^{\alpha(k)} \left(\sum_{j=1}^n 1_{\{\alpha(j) > u\}} \right)^{\frac{p}{p-2}} du \\ &\leq \sum_{k=1}^{\infty} k^{\frac{p}{p-2}} \alpha(k) < \infty. \end{aligned}$$

Notice that $\int_0^1 Q_{\|X_i\|}^p(u) du = \mathbb{E} \|X_i\|^p$, combining estimates above with (A), we obtain

$$2 \sum_{1 \leq i < j \leq n} 18 \int_0^{\alpha(j-i)} Q_{\|X_i\|}^2(u) du \leq \sum_{1 \leq i \leq n} C(\mathbb{E} \|X_i\|^p)^{\frac{2}{p}} \quad (\text{A.3})$$

Thus, the stationary of $(X_i)_{0 \leq i \leq n}$ implies

$$\mathbb{E} \left\| \sum_{i=1}^n X_i \right\|^2 \leq Cn (\pi(\|X\|^p))^{\frac{2}{p}}, \quad (\text{A.4})$$

We shall prove the case $p > p' > 2$ by induction on n . Suppose that for $1 \leq k < n$,

$$\mathbb{E} \left\| \sum_{i=1}^k X_i \right\|^{p'} \leq Ck^{\frac{p'}{2}} (\pi(\|X\|^p))^{\frac{p'}{p}}. \quad (\text{A.5})$$

When $k = n$, let $m = \lfloor \sqrt{n} \rfloor$ and $\bar{k}(n) = \lfloor \frac{n}{2m} \rfloor$ here. The block sums are defined by

$$Y_{i,1} = \sum_{j=1+2(i-1)m}^{n \wedge (2i-1)m} X_j; \quad Y_{i,2} = \sum_{j=1+(2i-1)m}^{n \wedge 2im} X_j \quad i \in \{1, \dots, \bar{k}(n) + 1\}.$$

Then we can get

$$\mathbb{E} \left\| \sum_{i=1}^n X_i \right\|^{p'} = \mathbb{E} \left\| \sum_{j=1}^{\bar{k}(n)+1} Y_{j,1} + \sum_{j=1}^{\bar{k}(n)+1} Y_{j,2} \right\|^{p'}$$

$$\begin{aligned}
&\leq 2^{p'-1} \left(\mathbb{E} \left\| \sum_{j=1}^{\bar{\kappa}(n)+1} \mathbf{Y}_{j,1} \right\|^{p'} + \mathbb{E} \left\| \sum_{j=1}^{\bar{\kappa}(n)+1} \mathbf{Y}_{j,2} \right\|^{p'} \right) \\
&:= 2^{p'-1} (I_1 + I_2).
\end{aligned} \tag{A.6}$$

For I_1 , [Lemma A.1](#) implies

$$\begin{aligned}
I_1 &\leq C_{p'} \left(\left(\sum_{i=1}^{\bar{\kappa}(n)+1} \mathbb{E} \|\mathbf{Y}_{i,1}\|^2 \right)^{\frac{p'}{2}} + (\bar{\kappa}(n) + 1)^{p'-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \mathbb{E} \|\mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}]\|^{p'} \right. \\
&\quad \left. + (\bar{\kappa}(n) + 1)^{\frac{p'}{2}-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \mathbb{E} |\mathbb{E}[\|\mathbf{Y}_{i,1}\|^2 | \mathcal{F}_{i-1}] - \mathbb{E} \|\mathbf{Y}_{i,1}\|^2|^{\frac{p'}{2}} + \sum_{i=1}^{\bar{\kappa}(n)+1} \mathbb{E} \|\mathbf{Y}_{i,1}\|^{p'} \right) \\
&:= C_{p'} (I_{1,1} + I_{1,2} + I_{1,3} + I_{1,4}),
\end{aligned}$$

where $\mathcal{F}_i = \sigma(\mathbf{Y}_{j,1}, j \leq i)$. For $I_{1,1}$, [\(A.4\)](#) implies

$$I_{1,1} \leq \left(\sum_{i=1}^{\bar{\kappa}(n)+1} C_m (\pi(\|\mathbf{X}\|^p))^{\frac{2}{p}} \right)^{\frac{p'}{2}} \leq (C_m(\bar{\kappa}(n) + 1))^{\frac{p'}{2}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}$$

For $I_{1,2}$, since $\mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}]$ is $\mathcal{F}_{i-1}$ measurable, one has

$$\begin{aligned}
I_{1,2} &= (\bar{\kappa}(n) + 1)^{p'-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \mathbb{E} \langle \mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}], \mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}] \|\mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}]\|^{p'-2} \rangle \\
&= (\bar{\kappa}(n) + 1)^{p'-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \mathbb{E} \langle \mathbf{Y}_{i,1}, \mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}] \|\mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}]\|^{p'-2} \rangle \\
&= (\bar{\kappa}(n) + 1)^{p'-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \sum_{j=1+2(i-1)m}^{n \wedge (2i-1)m} \mathbb{E} \langle \mathbf{X}_j, \mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}] \|\mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}]\|^{p'-2} \rangle.
\end{aligned}$$

Since $(\mathbf{X}_i)_{i \geq 0}$ are zero mean, [Lemma 2.1](#) and Young's inequality yield

$$\begin{aligned}
I_{1,2} &\leq 18(\bar{\kappa}(n) + 1)^{p'-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \sum_{j=1+2(i-1)m}^{n \wedge (2i-1)m} \int_0^{\alpha(m)} Q_{\|\mathbf{X}_j\|}(u) Q_{\|\mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}]\|^{p'-1}}(u) du \\
&\leq \frac{18}{p'} (\bar{\kappa}(n) + 1)^{p'-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \sum_{j=1+2(i-1)m}^{n \wedge (2i-1)m} \int_0^{\alpha(m)} Q_{\|\mathbf{X}_j\|}^{p'}(u) du \\
&\quad + \frac{18(p'-1)}{p'} (\bar{\kappa}(n) + 1)^{p'-1} m \sum_{i=1}^{\bar{\kappa}(n)+1} \int_0^{\alpha(m)} Q_{\|\mathbb{E}[\mathbf{Y}_{i,1} | \mathcal{F}_{i-1}]\|^{p'-1}}^{\frac{p'}{p'-1}}(u) du,
\end{aligned}$$

For the first term, Hölder's inequality implies

$$\int_0^{\alpha(m)} Q_{\|\mathbf{X}_j\|}^{p'}(u) du \leq \left(\int_0^1 1_{\{\alpha(m) > u\}} du \right)^{1-\frac{p'}{p}} \left(\int_0^1 Q_{\|\mathbf{X}_j\|}^p(u) du \right)^{\frac{p'}{p}}$$

$$= \alpha(m)^{1-\frac{p'}{p}} (\mathbb{E} \|\mathbf{X}_j\|^p)^{\frac{p'}{p}}.$$

For the second term, similar calculation implies

$$\begin{aligned} \int_0^{\alpha(m)} Q_{\|\mathbb{E}[\mathbf{Y}_{i,1}|\mathcal{F}_{i-1}]\|^{p'-1}}^{\frac{p'}{p'-1}}(u) du &\leq \left(\int_0^1 1_{\{\alpha(m) > u\}} du \right)^{1-\frac{p'}{p}} \left(\int_0^1 Q_{\|\mathbb{E}[\mathbf{Y}_{i,1}|\mathcal{F}_{i-1}]\|^{p'-1}}^{\frac{p'}{p'-1}}(u) du \right)^{\frac{p'}{p}} \\ &= \alpha(m)^{1-\frac{p'}{p}} \left(\mathbb{E}[\|\mathbb{E}[\mathbf{Y}_{i,1}|\mathcal{F}_{i-1}]\|^{(p'-1)\frac{p}{p'-1}}] \right)^{\frac{p'}{p}} \\ &\leq \alpha(m)^{1-\frac{p'}{p}} (\mathbb{E} \|\mathbf{Y}_{i,1}\|^p)^{\frac{p'}{p}} \\ &\leq \alpha(m)^{1-\frac{p'}{p}} m^{p'} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}. \end{aligned}$$

Then we have,

$$\begin{aligned} I_{1,2} &\leq \frac{18}{p'} (\bar{\kappa}(n) + 1)^{p'-1} \sum_{i=1}^{\bar{\kappa}(n)+1} \sum_{j=1+2(i-1)m}^{n \wedge (2i-1)m} \alpha(m)^{1-\frac{p'}{p}} (\mathbb{E} \|\mathbf{X}_j\|^p)^{\frac{p'}{p}} \\ &\quad + \frac{18(p'-1)}{p'} (\bar{\kappa}(n) + 1)^{p'-1} m \sum_{i=1}^{\bar{\kappa}(n)+1} \alpha(m)^{1-\frac{p'}{p}} m^{p'} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}} \\ &\leq \left(\frac{18}{p'} (\bar{\kappa}(n) + 1)^{p'} m + \frac{18(p'-1)}{p'} (\bar{\kappa}(n) + 1)^{p'} m^{p'+1} \right) \alpha(m)^{1-\frac{p'}{p}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}} \\ &\leq 36(\bar{\kappa}(n) + 1)^{p'} m^{p'+1} (Ce^{-\beta m})^{1-\frac{p'}{p}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}, \end{aligned}$$

where the last line follows (A.1). For $I_{1,3}$, we denote

$$\bar{\mathbf{Y}}_{i,1} = \mathbb{E}[\|\mathbf{Y}_{i,1}\|^2 | \mathcal{F}_{i-1}] - \mathbb{E} \|\mathbf{Y}_{i,1}\|^2,$$

$\mathcal{F}_{i-1}$ -measurable. Following the definition of $\mathbf{Y}_{i,1}$ and [10, Corollary], one has

$$\begin{aligned} \mathbb{E} |\bar{\mathbf{Y}}_{i,1}|^{\frac{p'}{2}} &= \mathbb{E} \left[|\bar{\mathbf{Y}}_{i,1}|^{\frac{p'}{2}-1} \text{sgn}(\bar{\mathbf{Y}}_{i,1}) (\mathbb{E}[\|\mathbf{Y}_{i,1}\|^2 | \mathcal{F}_{i-1}] - \mathbb{E} \|\mathbf{Y}_{i,1}\|^2) \right] \\ &= \sum_{j,l=1+2(i-1)m}^{n \wedge (2i-1)m} \mathbb{E} \left[|\bar{\mathbf{Y}}_{i,1}|^{\frac{p'}{2}-1} \text{sgn}(\bar{\mathbf{Y}}_{i,1}) (\mathbb{E}[\langle \mathbf{X}_j, \mathbf{X}_l \rangle | \mathcal{F}_{i-1}] - \mathbb{E}[\langle \mathbf{X}_j, \mathbf{X}_l \rangle]) \right] \\ &= \sum_{j,l=1+2(i-1)m}^{n \wedge (2i-1)m} \left\{ \mathbb{E} \left[|\bar{\mathbf{Y}}_{i,1}|^{\frac{p'}{2}-1} \text{sgn}(\bar{\mathbf{Y}}_{i,1}) \langle \mathbf{X}_j, \mathbf{X}_l \rangle \right] \right. \\ &\quad \left. - \mathbb{E} \left[|\bar{\mathbf{Y}}_{i,1}|^{\frac{p'}{2}-1} \text{sgn}(\bar{\mathbf{Y}}_{i,1}) \right] \mathbb{E}[\langle \mathbf{X}_j, \mathbf{X}_l \rangle] \right\} \\ &\leq 12 \sum_{j,l=1+2(i-1)m}^{n \wedge (2i-1)m} \alpha(m)^{1-\frac{2}{p}-\frac{p'-2}{p'}} \left[\mathbb{E} |\langle \mathbf{X}_j, \mathbf{X}_l \rangle|^{\frac{p}{2}} \right]^{\frac{2}{p}} \left[\mathbb{E} [|\bar{\mathbf{Y}}_{i,1}|^{\frac{p'}{2}-1}]^{\frac{p'}{p'-2}} \right]^{\frac{p'-2}{p'}} \\ &\leq 12 m^2 \alpha(m)^{\frac{2}{p'}-\frac{2}{p}} (\pi(\|\mathbf{X}\|^p))^{\frac{2}{p}} \left[\mathbb{E} |\bar{\mathbf{Y}}_{i,1}|^{\frac{p'}{2}} \right]^{\frac{p'-2}{p'}}. \end{aligned}$$

Thus, we can get

$$\mathbb{E}|\tilde{Y}_{i,1}|^{\frac{p'}{2}} \leq 12^{\frac{p'}{2}} m^{p'} \alpha(m)^{1-\frac{p'}{p}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}.$$

Then we have

$$\begin{aligned} I_{1,3} &\leq (1 + \bar{\kappa}(n))^{\frac{p'}{2}} 12^{\frac{p'}{2}} m^{p'} \alpha(m)^{1-\frac{p'}{p}} (\mathbb{E}\|\mathbf{X}\|^p)^{\frac{p'}{p}} \\ &\leq 12^{\frac{p'}{2}} (mn)^{\frac{p'}{2}} (Ce^{-\beta m})^{1-\frac{p'}{p}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}. \end{aligned}$$

For $I_{1,4}$, the induction hypothesis (A.5) implies

$$\begin{aligned} I_{1,4} &\leq \bar{C}(\bar{\kappa}(n) + 1) m^{\frac{p'}{2}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}} \\ &\leq \bar{C} n^{\frac{1}{2} + \frac{p'}{4}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}. \end{aligned}$$

Combining the estimate of $I_{1,1}$, $I_{1,2}$, $I_{1,3}$ and $I_{1,4}$, one has

$$\begin{aligned} I_1 &\leq C_{p'} \left((Cm(\bar{\kappa}(n) + 1))^{\frac{p'}{2}} + 36(\bar{\kappa}(n) + 1)^{p'} m^{p'+1} (Ce^{-\beta m})^{1-\frac{p'}{p}} \right. \\ &\quad \left. + 12^{\frac{p'}{2}} (mn)^{\frac{p'}{2}} (Ce^{-\beta m})^{1-\frac{p'}{p}} + \bar{C} n^{\frac{1}{2} + \frac{p'}{4}} \right) (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}} \\ &\leq C_{p'} \left(C_{1,p'} n^{\frac{p'}{2}} + C_{2,p'} n^{\frac{3}{4}p'} e^{-\beta\sqrt{n}(1-\frac{p'}{p})} + \bar{C} n^{\frac{1}{2} + \frac{p'}{4}} \right) (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}. \end{aligned}$$

Similarly,

$$I_2 \leq C_{p'} \left(C_{1,p'} n^{\frac{p'}{2}} + C_{2,p'} n^{\frac{3}{4}p'} e^{-\beta\sqrt{n}(1-\frac{p'}{p})} + \bar{C} n^{\frac{1}{2} + \frac{p'}{4}} \right) (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}.$$

Combining the estimate of I_1 , I_2 with (A.6), we have

$$\begin{aligned} \mathbb{E} \left\| \sum_{i=1}^n X_i \right\|^{p'} &\leq 2^{p'} C_{p'} \left(C_{1,p'} n^{\frac{p'}{2}} + C_{2,p'} n^{\frac{3}{4}p'} e^{-\beta\sqrt{n}(1-\frac{p'}{p})} + \bar{C} n^{\frac{1}{2} + \frac{p'}{4}} \right) (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}} \\ &\leq \bar{C} n^{\frac{p'}{2}} (\pi(\|\mathbf{X}\|^p))^{\frac{p'}{p}}, \end{aligned}$$

here we take $\bar{C}/3 \geq 2^{p'} C_{p'} C_{1,p'}$, n is large enough. □

Appendix B: The proof of lemmas in Section 2.2

Proof of Lemma 2.4. For m large enough, [1, Lemma 2.1] implies that one can construct independent random variables $\tilde{Y}_{m,j}$ distributed as $Y_{m,j}$ for $j = 1, \dots, \kappa(2^{m+1})$ on a richer probability space and

$$\begin{aligned} \mathbb{P}(Y_{m,j} \neq \tilde{Y}_{m,j} \text{ for some } 1 \leq j \leq \kappa(2^{m+1})) &\leq \kappa(2^{m+1}) \beta(m_2) \\ &\leq C \kappa(2^{m+1}) e^{-\beta m_2}. \end{aligned}$$

The Markov inequality implies

$$\mathbb{P} \left(\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (Y_{m,j} - \tilde{Y}_{m,j}) \right\| \geq x \right)$$

$$\begin{aligned}
&\leq \mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\|^{p'} \right] x^{-p'} \\
&= x^{-p'} \mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\|^{p'} 1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, i\}} \right].
\end{aligned}$$

Notice that

$$1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, i\}} \leq 1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, r\}} \quad \text{for } i \leq r.$$

Then we have

$$\begin{aligned}
&\mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\|^{p'} 1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, i\}} \right] \\
&\leq \mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\|^{p'} 1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, \kappa(2^{m+1})\}} \right] \\
&\leq \left\{ \mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\|^{2p'} \right] \right\}^{\frac{1}{2}} \left\{ \mathbb{E}[1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, \kappa(2^{m+1})\}}] \right\}^{\frac{1}{2}}.
\end{aligned}$$

A straight calculation yields

$$\begin{aligned}
\mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\|^{2p'} \right] &\leq \mathbb{E} \left[\sum_{j=1}^{\kappa(2^{m+1})} \|\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}\|^{2p'} \right] \\
&\leq C_{p'} \kappa(2^{m+1})^{2p'-1} \sum_{j=1}^{\kappa(2^{m+1})} (\mathbb{E} \|\mathbf{Y}_{m,j}\|^{2p'} + \mathbb{E} \|\tilde{\mathbf{Y}}_{m,j}\|^{2p'}) \\
&\leq C_{p'} \kappa(2^{m+1})^{2p'} m_1^{p'},
\end{aligned}$$

where the last line follows [Lemma 2.3](#) and the fact $\mathbf{Y}_{m,j} \stackrel{\mathcal{D}}{=} \tilde{\mathbf{Y}}_{m,j}$. For the indicator function part, we have

$$\begin{aligned}
\mathbb{E}[1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, \kappa(2^{m+1})\}}] &= \mathbb{P}(\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j = 1, \dots, \kappa(2^{m+1})) \\
&\leq C\kappa(2^{m+1})e^{-\beta m_2}.
\end{aligned}$$

Thus,

$$\mathbb{E} \left[\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\|^{p'} 1_{\{\mathbf{Y}_{m,j} \neq \tilde{\mathbf{Y}}_{m,j} \text{ for some } j=1, \dots, i\}} \right] \leq C m_1^{\frac{p'}{2}} \kappa(2^{m+1})^{p' + \frac{1}{2}} e^{-\frac{\beta}{2} m_2}.$$

Taking $x = 2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m}$ and $C^* \geq 2(1 - \alpha_1)(p' - \frac{1}{2})/\beta$, we can get

$$\sum_{m=1}^{\infty} \mathbb{P} \left(\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (\mathbf{Y}_{m,j} - \tilde{\mathbf{Y}}_{m,j}) \right\| \geq x \right)$$

$$\begin{aligned}
&\leq \sum_{m=1}^{\infty} Cx^{-p'} 2^{\frac{p'}{2}\alpha_1 m} 2^{(p'+\frac{1}{2})(1-\alpha_1)m} \exp\{-\frac{\beta}{2}m_2\} \\
&\leq \sum_{m=1}^{\infty} C(\log 2^m)^{-\frac{p'}{2}} 2^{(p'-\frac{1}{2})(1-\alpha_1)m} \exp\{-\frac{\beta}{2}C^* \log 2^m\} < \infty.
\end{aligned}$$

The Borel–Cantelli lemma yields

$$\max_{1 \leq i \leq \kappa(2^{m+1})} \left\| \sum_{j=1}^i (Y_{m,j} - \tilde{Y}_{m,j}) \right\| = o(2^{((1-\alpha_1)/p' + \alpha_1/2)m} \sqrt{\log 2^m}), \quad a.s.$$

For any $j \leq \kappa(2^{m+1})$, (2.1) immediately implies (2.2). $\square$

Proof of Lemma 2.5. Similar with the [proof of Lemma 2.4](#), we can construct independent random variables $\tilde{Z}_{m,j}$ distributed as $Z_{m,j}$ for $j = 1, \dots, 1 + \kappa(2^{m+1})$ on a richer probability space and

$$\mathbb{P}(Z_{m,j} \neq \tilde{Z}_{m,j} \text{ for some } 1 \leq j \leq \kappa(2^{m+1}) + 1) \leq C\kappa(2^{m+1})\beta(m_1).$$

Considering the following probability,

$$\begin{aligned}
&\mathbb{P}\left(\max_{2^m+1 \leq i \leq n} \left\| \sum_{\ell \in \mathcal{J}(m) \cap [2^{m+1}, i]} X_\ell \right\| \geq x\right) \\
&\leq \mathbb{P}\left(\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} (Z_{m,j} - \tilde{Z}_{m,j}) + \sum_{j=\kappa(i)(m_1+m_2)+m_1+1}^{(\kappa(i)(m_1+m_2)+m_1) \vee i} X_j \right\| + \max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} \tilde{Z}_{m,j} \right\| \geq x\right) \\
&\leq \mathbb{P}\left(\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} \tilde{Z}_{m,j} \right\| \geq x/2\right) + C_{p'} x^{-p'} \\
&\quad \left(\mathbb{E} \left[\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} (Z_{m,j} - \tilde{Z}_{m,j}) \right\|^{p'} \right] + \mathbb{E} \left[\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=\kappa(i)(m_1+m_2)+m_1+1}^{(\kappa(i)(m_1+m_2)+m_1) \vee i} X_j \right\|^{p'} \right] \right).
\end{aligned}$$

For the first term, since $\tilde{Z}_{m,j}$ are centered i.i.d. random vectors, Lévy's inequality and [Lemma 2.3](#) imply

$$\begin{aligned}
&\mathbb{P}\left(\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} \tilde{Z}_{m,j} \right\| \geq x/2\right) \\
&\leq 2\mathbb{P}\left(\left\| \sum_{j=1}^{\kappa(n)} \tilde{Z}_{m,j} \right\| \geq x/2\right) \leq Cx^{-p'} \mathbb{E} \left\| \sum_{j=1}^{\kappa(n)} \tilde{Z}_{m,j} \right\|^{p'} \leq Cx^{-p'} \kappa(2^{m+1})^{\frac{p'}{2}} m_2^{\frac{p'}{2}}.
\end{aligned}$$

Taking $x = 2^{m(1-\alpha_1)/2} \log 2^m$, one has

$$\sum_{m=1}^{\infty} \mathbb{P}\left(\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} \tilde{Z}_{m,j} \right\| \geq x/2\right) < \infty. \quad (\text{B.1})$$

For the second term, similar with the estimate of [Lemma 2.4](#), a straight calculation implies

$$\begin{aligned}
& \mathbb{E} \left[\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} (\mathbf{Z}_{m,j} - \tilde{\mathbf{Z}}_{m,j}) \right\|^{p'} \right] \\
&= \mathbb{E} \left[\max_{1 \leq i \leq \kappa(n)} \left\| \sum_{j=1}^i (\mathbf{Z}_{m,j} - \tilde{\mathbf{Z}}_{m,j}) \right\|^{p'} 1_{\{\mathbf{Z}_{m,j} \neq \tilde{\mathbf{Z}}_{m,j} \text{ for some } j=1, \dots, i\}} \right] \\
&\leq C m_2^{\frac{p'}{2}} \kappa(2^{m+1})^{p' + \frac{1}{2}} e^{-\frac{\beta}{2} m_1}.
\end{aligned}$$

For the last term, following [\[37, Proposition 1\]](#) and [Lemma 2.3](#), one has

$$\begin{aligned}
\left(\mathbb{E} \left[\max_{1 \leq i \leq 2^r} \left\| \sum_{j=1}^i \mathbf{X}_j \right\|^{p'} \right] \right)^{\frac{1}{p'}} &\leq \sum_{i=0}^r 2^{(r-i)/p'} \left(\mathbb{E} \left\| \sum_{j=1}^{2^i} \mathbf{X}_j \right\|^{p'} \right)^{\frac{1}{p'}} \\
&\leq C \sum_{i=0}^r 2^{(r-i)/p'} (2^i)^{\frac{1}{2}} \leq C 2^{\frac{r}{2}},
\end{aligned}$$

which yields

$$\mathbb{E} \left[\max_{1 \leq i \leq m_2} \left\| \sum_{j=1}^i \mathbf{X}_j \right\|^{p'} \right] \leq m_2^{p'/2}.$$

Then we have

$$\begin{aligned}
& \mathbb{E} \left[\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=\kappa(i)(m_1+m_2)+m_1+1}^{(\kappa(i)(m_1+m_2)+m_1) \vee i} \mathbf{X}_j \right\|^{p'} \right] \\
&\leq C \kappa(2^{m+1}) \mathbb{E} \left[\max_{1 \leq i \leq m_2} \left\| \sum_{j=1}^i \mathbf{X}_j \right\|^{p'} \right] \leq \kappa(2^{m+1}) m_2^{\frac{p'}{2}}
\end{aligned}$$

Thus,

$$\begin{aligned}
& \sum_{m=1}^{\infty} x^{-p'} C_{p'} \left(\mathbb{E} \left[\max_{2^m+1 \leq i \leq n} \left\| \sum_{j=1}^{\kappa(i)} (\mathbf{Z}_{m,j} - \tilde{\mathbf{Z}}_{m,j}) \right\|^{p'} \right] + \mathbb{E} \left[\max_{1 \leq i \leq m_2} \left\| \sum_{j=1}^i \mathbf{X}_j \right\|^{p'} \right] \right) \\
&\leq \sum_{m=1}^{\infty} x^{-p'} C_{p'} \left(m_2^{\frac{p'}{2}} \kappa(2^{m+1})^{p' + \frac{1}{2}} e^{-\frac{\beta}{2} m_1} + \kappa(2^{m+1}) m_2^{p'/2} \right) < \infty.
\end{aligned}$$

Combining [\(B.1\)](#) and (B), the Borel-Cantelli lemma implies

$$\max_{2^m+1 \leq i \leq n} \left\| \sum_{\ell \in \mathcal{J}(m) \cap [2^{m+1}, i]} \mathbf{X}_\ell \right\| = o(2^{m(1-\alpha_1)/2} \log 2^m), \quad a.s..$$

Similarly estimates hold for any i.i.d. centered Gaussian random vectors η_ℓ . □

Appendix C: Proof of examples in Section 4

Lemma C.1. *Under Assumption (A3), $(X_k)_{k \geq 0}$ is exponential ergodic with invariant measure π , that is,*

$$\sup_{|f| \leq 1+V} |\mathbb{E}[f(X_n)|X_0 = \mathbf{x}] - \pi(f)| \leq V(\mathbf{x})e^{-C_\gamma n},$$

where $C_\gamma > 0$ depends on γ .

Proof. We give the proof of the ergodicity of $(X_n)_{n \geq 0}$ following [35, Theorem 2.1]. To verify condition (7) of [35], let

$$V^n(\mathbf{x}) = e^{C_\gamma n} V(\mathbf{x}) \quad r(n) = C_\gamma e^{C_\gamma n},$$

the constant $C_\gamma > 0$ will be chosen later. Following Equation (4.1), one has

$$\begin{aligned} & PV^{n+1}(\mathbf{x}) + r(n)V(\mathbf{x}) \\ & \leq e^{C_\gamma(n+1)} (\gamma V(\mathbf{x}) + K) + C_\gamma e^{C_\gamma n} V(\mathbf{x}) \\ & = e^{C_\gamma n} V(\mathbf{x}) + C_\gamma e^{C_\gamma n} \left(\left(\frac{C_\gamma - 1}{C_\gamma} + \frac{1}{C_\gamma} \gamma e^{C_\gamma} \right) V(\mathbf{x}) + \frac{1}{C_\gamma} K e^{C_\gamma} \right). \end{aligned}$$

Choosing C_γ small enough such that $\gamma e^{C_\gamma} - 1 + C_\gamma < 0$, we can get

$$PV^{n+1}(\mathbf{x}) + r(n)V(\mathbf{x}) \leq V^n(\mathbf{x}) + br(n)1_{\{\mathbf{x} \in \mathcal{C}\}},$$

where $b = Ke^{C_\gamma}/C_\gamma$ and $\mathcal{C} = \{\mathbf{x} : V(\mathbf{x}) \leq \frac{Ke^{C_\gamma}}{1 - C_\gamma - \gamma e^{C_\gamma}}\}$. We deduce that $(X_n)_{n \geq 0}$ is ergodic with invariant measure π . $\square$

Recall the definition of β -mixing in [11, Proposition 1] which is equivalent with (1.1).

Definition C.2. The β -mixing coefficients are given by:

$$\beta(n) = \int \sup_{0 \leq f \leq 1} |P_n f(x) - \int f dQ| dQ,$$

where Q is a stationary distribution. The process X_n is β -mixing if $\lim_{n \rightarrow \infty} \beta(n) = 0$; is β -mixing with exponential decay rate if $\beta(n) \leq Ce^{-cn}$ for some $C > 0$ and $c > 0$.

Lemma C.3. *Under Assumption (A3), one has*

$$\beta(n) \leq \pi(V)e^{-C_\gamma n}. \tag{C.1}$$

Proof. According to Lemma C.1 and $\pi(V) < \infty$, one has

$$\begin{aligned} \beta(n) &= \int \sup_{0 \leq f \leq 1} |\mathbb{E}[f(X_n)|X_0 = \mathbf{x}] - \pi(f)| \pi(d\mathbf{x}) \\ &\leq \int \sup_{0 \leq f \leq V} |\mathbb{E}[f(X_n)|X_0 = \mathbf{x}] - \pi(f)| \pi(d\mathbf{x}) \\ &\leq \int V(\mathbf{x}) e^{-C_\gamma n} \pi(d\mathbf{x}) \end{aligned}$$

$$= \pi(V)e^{-C_V n}.$$

Thus, $(X_n)_{n \geq 0}$ is β -mixing with exponential decay rate. $\square$

Thus, condition (A1) is satisfied and $(X_n)_{n \geq 0}$ satisfies ASIP with rate greater than $1/4 + 1/(4p - 4)$.

C.2. Proof of Example 4.2

Proof. It is easy to calculate that (4.3) has a unique stationary solution given by

$$X_k = \sum_{j=0}^{\infty} A^j B \epsilon_{k-j}. \quad (\text{C.2})$$

For this exponential ergodic process, (A1) is easy to verified. Now we verify the condition (A2).

$$\begin{aligned} \text{cov}\left(\sum_{k=0}^{n-1} X_k\right) &= \sum_{k=0}^{n-1} \text{cov}(X_k) + \sum_{0 \leq i < j \leq n-1} (\text{cov}(X_i, X_j) + \text{cov}(X_i, X_j)^T) \\ &= n \text{cov}(X_0) + \sum_{0 \leq i < j \leq n-1} (\text{cov}(X_0, X_{j-i}) + \text{cov}(X_0, X_{j-i})^T). \end{aligned}$$

Following (C.2), one has

$$\text{cov}(X_0) = \text{cov}\left(\sum_{j=0}^{\infty} A^j B \epsilon_{-j}\right) = (I - A^2)^{-1} B^2,$$

and

$$\begin{aligned} \text{cov}(X_0, X_{j-i}) &= \text{cov}\left(\sum_{r=0}^{\infty} A^r \epsilon_{-r}, \sum_{r=0}^{j-i-1} A^r \epsilon_{-j-i-r} + \sum_{r=0}^{\infty} A^{j-i+r} \epsilon_{-r}\right) \\ &= \text{cov}(X_0) A^{j-i}. \end{aligned}$$

Thus, we have as $n \rightarrow \infty$,

$$\begin{aligned} \text{cov}\left(\sum_{k=0}^{n-1} X_k\right)/n &= \text{cov}(X_0) + \frac{1}{n} \sum_{k=1}^{n-1} k (\text{cov}(X_0, X_{n-k}) + \text{cov}(X_0, X_{n-k})^T) \\ &\rightarrow \text{cov}(X_0) + 2 \text{cov}(X_0)(A^{-1} - I)^{-1}, \end{aligned}$$

i.e., $\Gamma = \text{cov}(X_0) + 2 \text{cov}(X_0)(A^{-1} - I)^{-1}$. It is easy to see that,

$$\Gamma e_i(x) = \left(\frac{\lambda_i}{1 - \lambda_i^2} + \frac{2\lambda_i^2}{(1 - \lambda_i^2)(1 - \lambda_i)} \right) e_i(x).$$

Thus, (A2) is satisfied by taking $\lambda_i \asymp i^{-\delta}$ and ASIP holds. $\square$