COLLOQUIUM MATHEMATICUM

VOL. 134 2014 NO. 1

ON THE RATE OF CONVERGENCE IN THE WEAK INVARIANCE PRINCIPLE FOR DEPENDENT RANDOM VARIABLES WITH APPLICATIONS TO MARKOV CHAINS

BY

ION GRAMA (Vannes), ÉMILE LE PAGE (Vannes) and MARC PEIGNÉ (Tours)

Abstract. We prove an invariance principle for non-stationary random processes and establish a rate of convergence under a new type of mixing condition. The dependence is exponentially decaying in the gap between the past and the future and is controlled by an assumption on the characteristic function of the finite-dimensional increments of the process. The distinctive feature of the new mixing condition is that the dependence increases exponentially in the dimension of the increments. The proposed mixing property is particularly suited to processes whose behavior can be described in terms of spectral properties of some related family of operators. Several examples are discussed. We also work out explicit expressions for the constants involved in the bounds. When applied to Markov chains, our result specifies the dependence of the constants on the properties of the underlying Banach space and on the initial state of the chain.

1. Introduction. Let $(X_k)_{k\geq 1}$ be a sequence of real valued random variables (r.v.'s) defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let

$$S_N(t) = N^{-1/2} \sum_{k=1}^{[Nt]} X_k, \qquad t \in [0,1].$$

The (*weak*) *invariance principle* states that the process $\frac{1}{\sqrt{N}}(S_N(t))_{0\leq t\leq 1}$ converges weakly to the Brownian process $(W(t))_{0\leq t\leq 1}$, and is a powerful tool for various applications in probability and statistics. It extends the scope of the central limit theorem to continuous functionals of the stochastic process $(S_N(t))_{0\leq t\leq 1}$, such as, for example, the maxima or the L^2-norm of the trajectory of the process, considered in the appropriate functional spaces. The rates of convergence in the (weak) invariance principle, for independent r.v.'s under the existence of the moments of order $2+2\delta$, with $\delta > 0$, have been obtained in Prokhorov [28], Borovkov [4], Komlós, Major and Tusnády [22], Einmahl [10], Sakhanenko [31], [32], Zaĭtsev [38], [39] among others. In the case of martingale differences, for $\delta \leq 1/2$, the rates are

2010 *Mathematics Subject Classification*: Primary 60F17, 60J05, 60J10; Secondary 37C30.
Key words and phrases: rate of convergence, invariance principle, Markov chains, mixing, spectral gap.

 © Instytut Matematyczny PAN, 2014

essentially the same as in the independent case (see, for instance, Hall and Heyde [20], Kubilius [23], Grama [11]).

The *almost sure invariance principle* is a reinforcement of the weak invariance principle which states that the trajectories of a process are approximated with the trajectories of the Brownian motion a.s. in the sense that within a particular negligible error $r_N \to 0$ it is true that

$$\sup_{0\leq t\leq 1}\left|\frac{1}{\sqrt{N}}S_N(t) - W(t)\right| = O(r_N) \quad \text{a.s.}$$

There are many recent results concerning the rates of convergence in the strong invariance principle for weakly dependent r.v.'s under various conditions. We refer to Wu [37], Zhao and Woodroofe [40], Liu and Lin [24], Cuny [5], Merlevède and Rio [25], Dedecker, Doukhan and Merlevède [6] and to the references therein. However, in contrast to the case of independent r.v.'s where it is found that the optimal rate is of order $N^{-\delta/(2+2\delta)}$ for the strong invariance principle and $N^{-\delta/(3+2\delta)}$ for the weak invariance principle, the problem of obtaining the best rate of convergence in both the weak and strong invariance principles for dependent variables is not yet settled completely.

Gouëzel [15] has introduced a new type of mixing condition which is tied to spectral properties of the sequence $(X_k)_{k\geq 1}$. Consider the vectors $\overline{X}_1 = (X_{J_1},\dots,X_{J_{M_1}})$ and $\overline{X}_2 = (X_{k_{\text{gap}}+J_{M_1+1}},\dots,X_{k_{\text{gap}}+J_{M_1+M_2}})$, called the *past* and the *future*, respectively, where $X_{k+J_m} = \sum_{l\in J_m} X_{k+l}$, $J_m = [j_{m-1}, j_m)$, $j_0 \leq \dots \leq j_{M_1+M_2}$, and k_{gap} is a gap between $\overline{X}_1$ and $\overline{X}_2$. Roughly speaking, the condition used in [15] supposes that the characteristic function of $(\overline{X}_1, \overline{X}_2)$ is exponentially close to the product of the characteristic functions of the past $\overline{X}_1$ and the future $\overline{X}_2$, with an error term of the form $A\exp(-\lambda k_{\text{gap}})$, where λ is some non-negative constant and A is exponential in terms of the size of the blocks. This mixing property is particularly suited to systems whose behavior can be described in terms of spectral properties of some related family of operators, as initiated by Nagaev [26], [27] and Guivarc'h [16]. Examples are Markov chains whose perturbed transition probability operators $(\mathbf{P}_t)_{|t|\leq\varepsilon_0}$ exhibit a spectral gap and enough regularity in t, and dynamical systems whose characteristic functions can be coded by a family of operators $(\mathcal{L}_t)_{|t|\leq\varepsilon_0}$ with similar properties. Gouëzel proves in [15] an almost sure invariance principle with rate of convergence of order $N^{-\delta/(2+4\delta)}$.

The scope of the present paper is to improve on the results of Gouëzel by quantifying the rate of convergence in the (weak) invariance principle for dependent r.v.'s under the mixing condition introduced above. Although the strong and weak invariance principles are closely related, it seems that the rate of convergence in the (weak) invariance principle is less studied under weak dependence constraints. We refer to Doukhan, Leon and Portal [7], Merlevède and Rio [25] and Grama and Neumann [12]. However, these results rely on mixing conditions which do not hold in the present setting. Under the above mentioned mixing and some further mild conditions including the moment assumption $\sup_{k\geq 1} \mathbb{E}|X_k|^{2+2\delta} < \infty$ we obtain a bound of order $N^{-\frac{1+\alpha}{1+2\alpha}\frac{\alpha}{3+2\alpha}}$, for any $\alpha < \delta$. Moreover, we give explicit expressions of some constants involved in the rate of convergence; for instance, in the case of Markov walks we are able to figure out the dependence of the rate of convergence on the properties of the Banach space related to the corresponding family $(\mathbf{P}_t)_{|t|\leq\varepsilon_0}$ of perturbed transition operators and on the initial state $X_0 = x$ of the associated Markov chain. When compared with the rate $N^{-\frac{1}{2}\frac{\alpha}{1+2\alpha}}$ in the almost sure invariance principle of [15] ours appears with a loss in the power of multiple $\frac{2+2\alpha}{3+2\alpha} < 1$. This loss in the power is exactly the same as in the case of independent r.v.'s, when we compare the almost sure invariance principle (rate $N^{-\frac{\delta}{2+2\delta}}$) and the (weak) invariance principle (rate $N^{-\frac{\delta}{3+2\delta}}$).

As in the paper [15] our proof relies on a progressive blocking technique (see Bernstein [2]) coupled with a triadic Cantor-like scheme and on the Komlós, Major and Tusnády approximation type results for independent r.v.'s (see [22], [10], [38]), which is in contrast to approaches usually based on martingale methods.

As a potential application of the results obtained we point out the asymptotic equivalence of statistical experiments as developed in [13], [14], [12], whose scope can be extended to various models under weak dependence constraints.

Our paper is organized as follows. In Sections 2 and 3 we formulate our main results and give an application to the case of Markov chains. In Section 4 we introduce the notations to be used in the proofs of the main results. Proofs of the main results are given in Sections 5–7. In Section 8 we prove some bounds for the L^p norm of the increments of the process $(X_k)_{k\geq 1}$, and finally, in Section 9 we collect some auxiliary assertions and general facts.

We conclude this introduction by setting some notations to be used all over the paper. For any $x \in \mathbb{R}^d$, denote by $\|x\|_\infty = \sup_{1\leq i\leq d} |x_i|$ the supremum norm. For any $p > 0$, the L^p norm of a random variable X is denoted by $\|X\|_{L^p}$. The equality in distribution of two stochastic processes $(Z_i')_{i\geq 1}$ and $(Z_i'')_{i\geq 1}$, possibly defined on two different probability spaces $(\Omega', \mathcal{F}', \mathbb{P}')$ and $(\Omega'', \mathcal{F}'', \mathbb{P}'')$, will be denoted $\mathcal{L}((Z_i')_{i\geq 1}|\mathbb{P}') \overset{d}{=} \mathcal{L}((Z_i'')_{i\geq 1}|\mathbb{P}'')$. The generalized inverse of a distribution function F on a real line is denoted by F^{-1}, i.e. $F^{-1}(y) = \inf\{x : F(x) > y\}$. By $c, c', c'', \ldots,$ possibly supplied with indices $1, 2, \ldots,$ we denote absolute constants whose values may vary from line to line. The notations $c_{\alpha_1,\dots,\alpha_r}, c'_{\alpha_1,\dots,\alpha_r},\dots$ will be used to stress that the constants depend only on the parameters indicated in their indices: for instance $c'_{\alpha,\beta}$ denotes a constant depending only on the constants α,β. All other constants will be specifically indicated. As usual, a "standard normal r.v." is a normal random variable of mean 0 and variance 1.

2. Main result. Assume that on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ we are given a sequence $(X_i)_{i\geq 1}$ of dependent r.v.'s with values on the real line $\mathbb{R}$. The expectation with respect to $\mathbb{P}$ is denoted by $\mathbb{E}$.

The following condition will be used to ensure that the process $(X_i)_{i\geq 1}$ has almost independent increments. Given natural numbers $k_{\rm gap}, M_1, M_2 \in \mathbb{N}$ and a sequence $j_0 \leq \dots \leq j_{M_1+M_2}$ denote $X_{k+J_m} = \sum_{l\in J_m} X_{k+l}$, where $J_m = [j_{m-1}, j_m)$, $m = 1,\dots,M_1+M_2$ and $k \geq 0$. Consider the vectors $\overline{X}_1 = (X_{J_1},\dots,X_{J_{M_1}})$ and $\overline{X}_2 = (X_{k_{\rm gap}+J_{M_1+1}},\dots,X_{k_{\rm gap}+J_{M_1+M_2}})$. Let $\phi(t_1,t_2) = \mathbb{E}e^{it_1\overline{X}_1+it_2\overline{X}_2}$, $\phi_1(t_1) = \mathbb{E}e^{it_1\overline{X}_1}$ and $\phi_2(t_2) = \mathbb{E}e^{it_2\overline{X}_2}$ be the characteristic functions of $(\overline{X}_1,\overline{X}_2)$, $\overline{X}_1$ and $\overline{X}_2$ respectively. We require that the dependence between the two vectors $\overline{X}_1$ (the past) and $\overline{X}_2$ (the future) decreases exponentially as a function of the size of the gap $k_{\rm gap}$ in the following sense.

Condition **C1**. There exist positive constants $\varepsilon_0 \leq 1$, $\lambda_0,\lambda_1,\lambda_2$ such that for any $k_{\rm gap}$, $M_1,M_2\in\mathbb{N}$, any sequence $j_0 < \dots < j_{M_1+M_2}$ and any $t_1\in R^{M_1}$, $t_2\in R^{M_2}$ satisfying $\|(t_1,t_2)\|_\infty \leq \varepsilon_0$,

$$\begin{aligned}&|\phi(t_1,t_2)-\phi_1(t_1)\phi_2(t_2)|\\ &\qquad\leq \lambda_0\exp(-\lambda_1 k_{\rm gap})\Big(1+\max_{m=1,\dots,M_1+M_2}\operatorname{card}(J_m)\Big)^{\lambda_2(M_1+M_2)}.\end{aligned}$$

All over the paper we suppose that the following moment condition holds true.

Condition **C2**. There exist two constants $\delta>0$ and $\mu_\delta>0$ such that

$$\sup_{i\geq 1}\|X_i\|_{L^{2+2\delta}} \leq \mu_\delta < \infty.$$

We also suppose that the sequence $(X_i)_{i\geq 1}$ has the following asymptotic homogeneity property.

Condition **C3**. There exist constants $\tau>0$ and $\sigma>0$ such that, for any $\gamma>0$ *and any* $n\geq 1$,

$$\sup_{k\geq 0}\Big|n^{-1}\operatorname{Var}_{\mathbb{P}}\Big(\sum_{i=k+1}^{k+n}X_i\Big)-\sigma^2\Big| \leq \tau n^{-1+\gamma}.$$

The main result of the paper is the following theorem. Denote $\mu_i=\mathbb{E}X_i$ for $i\geq 1$.

THEOREM 2.1. *Assume Conditions* **C1–C3** *hold. Let* $0 < \alpha < \delta$. *Then on some probability space* $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ *there exist a sequence of random variables* $(\widetilde{X}_i)_{i\geq 1}$ *such that* $\mathcal{L}((\widetilde{X}_i)_{i\geq 1}|\widetilde{\mathbb{P}}) \overset{d}{=} \mathcal{L}((X_i)_{i\geq 1}|\mathbb{P})$ *and a sequence of independent standard normal random variables* $(W_i)_{i\geq 1}$ *such that for any* $0 < \rho < \frac{\alpha}{2(1+2\alpha)}$ *and* $N \geq 1$,

$$\widetilde{\mathbb{P}}\Big(N^{-1/2} \sup_{k\leq N}\Big|\sum_{i=1}^{k}(\widetilde{X}_i - \mu_i - \sigma W_i)\Big| > 6N^{-\rho}\Big) \leq C_0 N^{-\alpha\frac{1+\alpha}{1+2\alpha}+\rho(2+2\alpha)},$$

where $C_0 = c_{\lambda_1,\lambda_2,\alpha,\delta,\sigma}(1+\lambda_0+\mu_\delta+\sqrt{\tau})^{2+2\delta}$ *and* $c_{\lambda_1,\lambda_2,\alpha,\delta,\sigma}$ *depends only on the constants indicated in its indices.*

Letting $\rho = \frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}$, from Theorem 2.1 we get

$$\widetilde{\mathbb{P}}\Big(N^{-1/2} \sup_{k\leq N}\Big|\sum_{i=1}^{k}(\widetilde{X}_i - \mu_i - \sigma W_i)\Big| > 6N^{-\frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}}\Big) \leq C_0 N^{-\frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}}, \tag{2.1}$$

where $C_0 = c_{\lambda_1,\lambda_2,\alpha,\delta,\sigma}(1+\lambda_0+\mu_\delta+\sqrt{\tau})^{2+2\delta}$ and $c_{\lambda_1,\lambda_2,\alpha,\delta,\sigma}$ depends only on the constants indicated in its indices. Compared with the optimal rate of convergence $N^{-\alpha/(3+2\alpha)}$ for independent r.v.'s, the loss in the power is within the factor $\frac{1+\alpha}{1+2\alpha}$. As $\alpha \to \infty$ we obtain the limiting power $1/4$ which is twice worse than the optimal power $1/2$ in the independent case. In particular, if $\alpha = 1/2$ (which corresponds to $p = 2+2\alpha = 3$) we obtain the rate of convergence $N^{-\frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}} = N^{-\frac{3}{32}}$, while in the independent case we have $N^{-1/8}$, which represents a loss of the power of order $\frac{1}{8} - \frac{3}{32} = \frac{1}{32}$.

Observe that in Theorem 2.1 we figure out the explicit dependence of the constant C_0 on the constants λ_0, μ_δ and τ involved in Conditions **C1–C3**. In the next section we show that Theorem 2.1 can be applied to Markov walks under spectral gap type assumptions on the associated Markov chain. It is important to stress that our result is the first one to figure out the dependence of the constants involved in the bounds on the initial state of the Markov chain. The results of the paper can also be applied to a large class of dynamical systems, however this stays beyond the scope of the present article. For a discussion of such applications we refer to [15].

For the proof of Theorem 2.1, without loss of generality, we shall assume that $\mu_i = 0$, $i \geq 1$ and $\sigma = 1$, since the general case can be reduced to this one by centering and renormalizing the variables X_i, i.e. by replacing X_i by $X_i' = (X_i - \mu_i)/\sigma$. It is easy to see that Conditions **C1–C3** are satisfied for the new random variables X_i' with the same λ_0 and with μ_δ, τ replaced by $2\mu_\delta/\sigma$, τ/σ^2.

3. Applications to Markov walks. Consider a Markov chain $(X_k)_{k\geq 0}$ with values in the measurable state space $(\mathbb{X}, \mathscr{X})$ with the transition probability $\mathbf{P}(x,\cdot)$, $x \in \mathbb{X}$. For every $x \in \mathbb{X}$ denote by $\mathbb{P}_x$ and $\mathbb{E}_x$ the probability measure and expectation generated by the finite-dimensional distributions

$$\mathbb{P}_x(X_0 \in B_0, \dots, X_n \in B_n) = 1_{B_0}(x) \int_{B_1} \dots \int_{B_n} \mathbf{P}(x, dx_1) \dots \mathbf{P}(x_{n-1}, dx_n)$$

for any $B_k \in \mathscr{X}$, $k = 1, \dots, n$, $n = 1, 2, \dots$, on the space of trajectories $(\mathbb{X}, \mathscr{X})^{\mathbb{N}}$. In particular $\mathbb{P}_x(X_0 = x) = 1$.

Let f be a real valued function defined on the state space $\mathbb{X}$ of the Markov chain $(X_k)_{k\geq 0}$. Let $\mathscr{B}$ be a Banach space of real valued functions on $\mathbb{X}$ endowed with the norm $\|\cdot\|_{\mathscr{B}}$ and let $\|\cdot\|_{\mathscr{B}\to\mathscr{B}}$ be the operator norm on $\mathscr{B}$. Denote by $\mathscr{B}' = \mathscr{L}(\mathscr{B}, \mathbb{C})$ the topological dual of $\mathscr{B}$ equipped with the norm $\|\cdot\|_{\mathscr{B}'}$. The unit function on $\mathbb{X}$ is written e: $e(x) = 1$ for $x \in \mathbb{X}$. The Dirac measure at $x \in \mathbb{X}$ is denoted by $\boldsymbol{\delta}_x$: $\boldsymbol{\delta}_x(g) = g(x)$ for any $g \in \mathscr{B}$.

We introduce the following hypotheses.

Hypothesis **M1** (Banach space).

(a) The unit function e belongs to $\mathscr{B}$.
(b) For every $x \in X$ the Dirac measure $\boldsymbol{\delta}_x$ *belongs to* $\mathscr{B}'$.
(c) $\mathscr{B} \subseteq L^1(\mathbf{P}(x,\cdot))$ for every $x \in X$.
(d) There exists a constant $\varepsilon_0 \in (0,1)$ such that for any $g \in \mathscr{B}$ the function $e^{itf} g$ is in $\mathscr{B}$ for any t satisfying $|t| \leq \varepsilon_0$.

Note that, for any $x \in \mathbb{X}$ and $g \in L^1(\mathbf{P}(x,\cdot))$, the quantity $\mathbf{P}g(x) := \int_{\mathbb{X}} g(y)\, \mathbf{P}(x, dy)$ is well defined. In particular, under Hypothesis **M1**(c), $\mathbf{P}g(x)$ exists when $g \in \mathscr{B}$. We thus consider the following hypothesis:

Hypothesis **M2** (Spectral gap).

(a) The map $g \mapsto \mathbf{P}g$ is a bounded operator on $\mathscr{B}$.
(b) There exist constants $C_Q > 0$ and $\kappa \in (0,1)$ such that

$$\mathbf{P} = \Pi + Q, \tag{3.1}$$

where Π is a one-dimensional projector and Q is an operator on $\mathscr{B}$ satisfying $\Pi Q = Q \Pi = 0$ and $\|Q^n\|_{\mathscr{B}\to\mathscr{B}} \leq C_Q \kappa^n$.

Notice that, since the image of Π is generated by the unit function e, there exists a linear form $\nu \in \mathscr{B}'$ such that, for any $g \in \mathscr{B}$,

$$\Pi g = \nu(g) e. \tag{3.2}$$

When Hypotheses **M1** and **M2** hold, we set $\mathbf{P}_t g = \mathbf{P}(e^{itf} g)$ for any $g \in \mathscr{B}$ and $t \in [-\varepsilon_0, \varepsilon_0]$. Notice that $\mathbf{P} = \mathbf{P}_0$.

HYPOTHESIS **M3** (Perturbed transition operator).

(a) For any $|t| \leq \varepsilon_0$ the map $\mathcal{B} \ni g \mapsto P_t g \in \mathcal{B}$ is a bounded operator on $\mathcal{B}$.

(b) There exists a constant $C_{\mathbf{P}} > 0$ such that, for all $n \geq 1$ and $|t| \leq \varepsilon_0$,

$$\|\mathbf{P}_t^n\|_{\mathcal{B}\to\mathcal{B}} \leq C_{\mathbf{P}}. \tag{3.3}$$

HYPOTHESIS **M4** (Moment condition). There exists $\delta > 0$ such that for any $x \in X$,

$$\mu_\delta(x) := \sup_{k\geq 1}(\mathbb{E}_x|f(X_k)|^{2+2\delta})^{\frac{1}{2+2\delta}} = \sup_{k\geq 1}((\mathbf{P}^k|f|^{2+2\delta})(x))^{\frac{1}{2+2\delta}} < \infty.$$

We show first that under Hypotheses **M1**–**M4**, Conditions **C1**–**C3** are satisfied. As in the previous section let $k_{\text{gap}}, M_1, M_2 \in \mathbb{N}$ and $j_0 \leq \cdots \leq j_{M_1+M_2}$ be natural numbers. Denote $Y_{k+Jm} = \sum_{l\in J_m} f(X_{k+l})$, where $J_m = [j_{m-1}, j_m)$, $m = 1, \dots, M_1 + M_2$ and $k \geq 0$. Consider the vectors $\overline{Y}_1 = (Y_{J_1}, \dots, Y_{J_{M_1}})$ and $\overline{Y}_2 = (Y_{k_{\text{gap}}+J_{M_1+1}}, \dots, Y_{k_{\text{gap}}+J_{M_1+M_2}})$. Denote by $\phi_x(t_1, t_2) = \mathbb{E}e^{it_1\overline{Y}_1+it_2\overline{Y}_2}$, $\phi_{x,1}(t_1) = \mathbb{E}_x e^{it_1\overline{Y}_1}$ and $\phi_{x,2}(t_2) = \mathbb{E}_x e^{it_2\overline{Y}_2}$ the characteristic functions of $(\overline{Y}_1, \overline{Y}_2)$, $\overline{Y}_1$ and $\overline{Y}_2$ respectively.

PROPOSITION 3.1. *Assume that the Markov chain $(X_n)_{n\geq 1}$ and the function f satisfy Hypotheses* **M1**–**M4**. *Then Condition* **C1** *is satisfied, i.e. there exists a positive constant $\varepsilon_0 \leq 1$ such that for any $k_{\text{gap}}, M_1, M_2 \in \mathbb{N}$, any sequence $j_0 < \cdots < j_{M_1+M_2}$ and any $t_1 \in \mathbb{R}^{M_1}, t_2 \in \mathbb{R}^{M_2}$ satisfying $\|(t_1, t_2)\|_\infty \leq \varepsilon_0$,*

$$|\phi_x(t_1,t_2) - \phi_{x,1}(t_1)\phi_{x,2}(t_2)| \\ \leq \lambda_0(x)\exp(-\lambda_1 k_{\text{gap}})\Big(1 + \max_{m=1,\dots,M_1+M_2} \operatorname{card}(J_m)\Big)^{\lambda_2(M_1+M_2)},$$

where

$$\lambda_0(x) = 2C_Q(\|\nu\|_{\mathcal{B}'} + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'})\|e\|_{\mathcal{B}}, \\ \lambda_1 = |\ln\kappa|, \qquad \lambda_2 = \max\{1, \log_2 C_{\mathbf{P}}\}.$$

PROPOSITION 3.2. *Assume that the Markov chain $(X_n)_{n\geq 1}$ and the function f satisfy Hypotheses* **M1**–**M4**. *Then:*

(a) *There exists a constant μ such that for any $x \in \mathbb{X}$ and $k \geq 1$,*

$$|\mathbb{E}_x f(X_k) - \mu| \leq c_\delta A_1(x)\kappa^{k\gamma/4-1} \tag{3.4}$$

for any positive constant γ satisfying $0 < \gamma \leq \min\{1, 2\delta\}$, where $A_1(x) = 1 + \mu_\delta(x)^{1+\gamma} + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}C_{\mathbf{P}}C_Q$. Moreover

$$\sum_{k=0}^{\infty} |\mathbb{E}_x f(X_k) - \mu| \leq \overline{\mu}(x) = c_{\delta,\kappa,\gamma}A_1(x). \tag{3.5}$$

(b) *There exists a constant* $\sigma \geq 0$ *such that for any* $x \in \mathbb{X}$,

$$\sup_{m\geq 0}\Big|\mathrm{Var}_{\mathbb{P}_x}\Big(\sum_{i=m+1}^{m+n} f(X_i)\Big) - n\sigma^2\Big| \leq \tau(x) = c_{\delta,\kappa,\gamma}A_2(x), \tag{3.6}$$

where

$$A_2(x) = 1 + \mu_\delta(x)^{2+\gamma} + (1 + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'})\|e\|_{\mathcal{B}}(C_{\mathbf{P}}^2 C_Q(1 + C_Q) + C_{\mathbf{P}}C_Q(1 + \|\nu\|_{\mathcal{B}'}C_{\mathbf{P}})).$$

Note that the constants μ and σ do not depend on the initial state x.

The main result of this section is the following theorem. Let $\widetilde{\Omega} = \mathbb{R}^\infty \times \mathbb{R}^\infty$. For any $\widetilde{\omega} = (\widetilde{\omega}_1, \widetilde{\omega}_2) \in \widetilde{\Omega}$ denote by $\widetilde{Y}_i = \widetilde{\omega}_{1,i}$ and $W_i = \widetilde{\omega}_{2,i}$, $i \geq 1$, the coordinate processes in $\widetilde{\Omega}$.

THEOREM 3.3. *Assume that the Markov chain* $(X_n)_{n\geq 0}$ *and the function* f *satisfy Hypotheses* **M1**–**M4**, *with* $\sigma > 0$. *Let* $0 < \alpha < \delta$. *Then there exists a Markov transition kernel* $x \mapsto \widetilde{\mathbb{P}}_x(\cdot)$ *from* $(\mathbb{X}, \mathcal{X})$ *to* $(\widetilde{\Omega}, \mathcal{B}(\widetilde{\Omega}))$ *such that* $\mathcal{L}((\widetilde{Y}_i)_{i\geq 1}|\widetilde{\mathbb{P}}_x) \overset{d}{=} \mathcal{L}((f(X_i))_{i\geq 1}|\mathbb{P}_x)$, *the* W_i, $i \geq 1$, *are independent standard normal r.v.'s under* $\widetilde{\mathbb{P}}_x$, *and for any* $0 < \rho < \frac{\alpha}{2(1+2\alpha)}$,

$$\widetilde{\mathbb{P}}_x\Big(N^{-1/2}\sup_{k\leq N}\Big|\sum_{i=1}^{k}(\widetilde{Y}_i - \mu - \sigma W_i)\Big| > 6N^{-\rho}\Big) \leq C(x)N^{-\alpha\frac{1+\alpha}{1+2\alpha}+\rho(2+2\alpha)}, \tag{3.7}$$

with

$$C(x) = C_1(1 + \mu_\delta(x) + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'})^{2+2\delta},$$

where C_1 *is a constant depending only on* $\delta, \alpha, \kappa, C_{\mathbf{P}}, C_Q, \|e\|_{\mathcal{B}}$ *and* $\|\nu\|_{\mathcal{B}'}$.

Note that only the probability $\widetilde{\mathbb{P}}_x$ depends on the initial state x while the processes $(\widetilde{Y}_k)_{k\geq 0}$ and $(W_k)_{k\geq 0}$ do not.

As in the previous section, letting $\rho = \frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}$, under the conditions of Theorem 3.3 we obtain

$$\widetilde{\mathbb{P}}_x\Big(N^{-1/2}\sup_{k\leq N}\Big|\sum_{i=1}^{k}(\widetilde{Y}_i - \mu - \sigma W_i)\Big| > 6N^{-\frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}}\Big) \leq C(x)N^{-\frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}}. \tag{3.8}$$

Compared to the rate $N^{-\frac{\alpha}{3+2\alpha}}$, which is optimal in the independent case, the rate of convergence $N^{-\frac{\alpha}{3+2\alpha}\frac{1+\alpha}{1+2\alpha}}$ in (3.8) is slower by the factor $N^{\frac{\alpha}{3+2\alpha}\frac{\alpha}{1+2\alpha}}$. As $\alpha \to \infty$ we obtain $N^{-1/4}$, which is the best rate in the invariance principle that is known for dependent random variables.

In Theorem 3.3 we do not suppose the existence of the stationary measure. Assume that there exists a stationary probability measure ν on $\mathbb{X}$; it thus coincides with the linear form ν introduced in (3.2). Let $\mathbb{P}_\nu$ and $\mathbb{E}_\nu$ be the probability measure and expectation generated by the finite-dimensional distributions of the chain under the stationary measure ν. Note that the means $\mathbb{E}_\nu X_k$ and the covariances $\mathrm{Cov}_{\mathbb{P}_\nu}(f(X_l), f(X_{l+k}))$ with respect to ν may not exist, under Hypotheses **M1**–**M4**. To ensure their existence, we require the following additional condition (where generally $|f|^2 \notin \mathcal{B}$).

HYPOTHESIS **M5** (Stationary measure). On the state space $\mathbb{X}$ there exists a stationary probability measure ν with $\nu(\sup_{k\geq 0} \mathbf{P}^k(|f|^2)) < \infty$.

Under Hypothesis **M5** for μ and σ we find the usual expressions of the means and of the variance in the central limit theorem for dependent r.v.'s.

THEOREM 3.4. *Assume that the Markov chain* $(X_n)_{n\geq 0}$ *and the function* f *satisfy Hypotheses* **M1**–**M5**. *Assume also that* $\sigma_\nu > 0$. *Then Proposition* 3.2 *holds true with* $\mu = \nu(f)$ *and* $\sigma = \sigma_\nu$, *where*

$$\nu(f) = \int f(x)\, \nu(dx)$$

and

$$\sigma_\nu^2 = \mathrm{Var}_{\mathbb{P}_\nu}(f(X_0)) + 2\sum_{k=1}^{\infty} \mathrm{Cov}_{\mathbb{P}_\nu}(f(X_0), f(X_k)).$$

Moreover, if $\sigma_\nu > 0$ *the assertions of Theorem* 3.3 *and* (3.8) *hold true with* $\mu = \nu(f)$ *and* $\sigma = \sigma_\nu$.

From Theorem 3.4 one can derive a bound when the Markov chain $(X_n)_{n\geq 0}$ is in the stationary regime. If we assume $\nu(\sup_{k\geq 0} \mathbf{P}^k(|f|^{2+2\delta})) \leq c_{\nu,\delta} < \infty$ and $\int \|\boldsymbol{\delta}_x\|_{\mathcal{B}'}^{2+2\delta}\, \nu(dx) \leq c_{\mathcal{B}',\delta} < \infty$, then integrating (3.7) with respect to ν we obtain

$$\widetilde{\mathbb{P}}_\nu\Big(N^{-1/2} \sup_{k\leq N}\Big|\sum_{i=1}^{k}(\widetilde{Y}_i - \mu - \sigma W_i)\Big| > 6N^{-\rho}\Big) \leq C N^{-\alpha\frac{1+\alpha}{1+2\alpha}+\rho(2+2\alpha)},$$

where C is a constant depending on $\delta, \alpha, \kappa, C_{\mathbf{P}}, C_Q, \|e\|_{\mathcal{B}}, \|\nu\|_{\mathcal{B}'}$ and $c_{\nu,\delta}, c_{\mathcal{B}',\delta}$.

Hypotheses **M1**–**M5** formulated above can be easily verified by standard methods. As to **M3** it can be verified using two approaches. The first approach is based on the assumption that the family of operators $(\mathbf{P}_t)_{|t|\leq\varepsilon_0}$ is continuous in t at $t = 0$. In this case, **M3** is satisfied by classical perturbation theory (see, for instance, Dunford and Schwartz [9]). The second approach is based on a weaker form of continuity of the family $(\mathbf{P}_t)_{|t|\leq\varepsilon_0}$ as developed in Keller and Liverani [21].

We end this section by giving three examples where these hypotheses are satisfied.

EXAMPLE 1 (Markov chains with finite state spaces). Suppose that $(X_n)_{n\geq 0}$ is an irreducible ergodic aperiodic Markov chain with finite state space. It is easy to verify that Hypotheses **M1**–**M5** are satisfied and that there exists a unique invariant probability measure ν. Then the conclusions of Theorem 3.4 hold true.

EXAMPLE 2 (Autoregressive random walk with Bernoulli noise). Consider the autoregressive model $x_{n+1} = \alpha x_n + b_n$, $n \geq 0$, where α is a constant satisfying $\alpha \in (-1,1)$, and $(b_n)_{n\geq 0}$ are i.i.d. Bernoulli r.v.'s with $P(b=1) = P(b=-1) = 1/2$ and $x_0 = x$. Consider the Banach space $\mathcal{B} = \mathcal{L}$ of continuous functions f on $\mathbb{R}$ such that $\|f\| = |f| + [f] < \infty$, where

$$|f| = \sup_{x\in\mathbb{R}} \frac{|f(x)|}{1+x^2}, \qquad [f] = \sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|(1+x^2)(1+y^2)}.$$

Since $\alpha \in (-1,1)$, the invariant measure ν exists and coincides with the law of the random variable $Z = \sum_{i=1}^{\infty} \alpha^{i-1} b_i$. It is easy to verify that Hypotheses **M1**–**M5** are satisfied for the function $f(x) = x$. For the mean $\nu(f)$ we have

$$\nu(f) = \int x\,\nu(dx) = \mathbb{E}Z = \sum_{i=1}^{\infty} \alpha^{i-1}\mathbb{E}b_1 = \frac{\mathbb{E}b_1}{1-\alpha}.$$

Since $\mathbb{E}b_1 = 0$, one gets $\nu(f) = 0$ and the variance is computed as follows:

$$\sigma_\nu^2 = \lim_{n\to\infty} \mathbb{E}\Big(\sum_{i=1}^{n} \alpha^{i-1} b_i\Big)^2 = \lim_{n\to\infty} \sum_{i=1}^{n} \alpha^{2(i-1)}\mathbb{E}b_1^2 = \frac{1}{1-\alpha^2}.$$

Thus the conclusions of Theorem 3.4 hold true with $\nu(f) = 0$ and $\sigma_\nu^2 = \frac{1}{1-\alpha^2}$.

EXAMPLE 3 (Stochastic recursion). On the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ consider the stochastic recursion

$$x_{n+1} = a_{n+1}x_n + b_{n+1}, \qquad n \geq 0,$$

where $(a_n, b_n)_{n\geq 0}$ are i.i.d. r.v.'s with values in $(0,\infty)\times\mathbb{R}$ of the same distribution $\widehat{\mu}$ and $x_0 = x$. Following Guivarc'h and Le Page [17], we assume the conditions:

H1. There exists $\alpha > 2$ such that

$$\varphi(\alpha) := \int |a|^\alpha\,\widehat{\mu}(da,db) < 1 \quad \text{and} \quad \int |b|^\alpha\,\widehat{\mu}(da,db) < \infty.$$

H2. $\widehat{\mu}(\{(a,b) : ax_0 + b = x_0\}) < 1$ *for any* $x_0 \in \mathbb{R}$.

H3. The set $\{\ln|a| : (a,b) \in \operatorname{supp}\widehat{\mu}\}$ generates a dense subgroup of $\mathbb{R}$.

Let $\varepsilon \in (0,1)$, θ and c be positive such that $\alpha - 1 < c+\varepsilon < \theta \leq 2c < \alpha - \varepsilon$. Consider the Banach space $\mathcal{B} = \mathcal{L}_{\varepsilon,c,\theta}$ of continuous functions f on $\mathbb{R}$ such that $\|f\| = |f| + [f] < \infty$, where

$$|f| = \sup_{x\in\mathbb{R}} \frac{|f(x)|}{1+|x|^\theta}, \qquad [f] = \sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|^\varepsilon(1+|x|^c)(1+|y|^c)}.$$

The transition probability $\mathbf{P}(x,\cdot)$ of the Markov chain $(x_n)_{n\geq 0}$ is defined by

$$\int h(y)\,\mathbf{P}(x,dy) = \int h(ax+b)\,\widehat{\mu}(da,db)$$

for any bounded Borel measurable function $h:\mathbb{R}\to\mathbb{R}$ and $x\in\mathbb{R}$. For any $x\in\mathbb{R}$ denote by $\mathbb{P}_x$ and $\mathbb{E}_x$ the corresponding probability measure and expectation generated by the finite-dimensional distributions on the space of trajectories. It is proved in [17, Proposition 1] that the series $\sum_{i=1}^{\infty} a_1\ldots a_{i-1}b_i$ is $\mathbb{P}$-a.s. convergent and the Markov chain $(x_n)_{n\geq 0}$ has a unique invariant probability measure ν which coincides with the law of $Z=\sum_{i=1}^{\infty} a_1\ldots a_{i-1}b_i$. Moreover, $\int |x|^t\,\nu(dx)<\infty$ for any $t\in[0,\alpha)$.

We now verify that Hypotheses **M1**–**M5** are satisfied for $f(x)=x$. Hypothesis **M1** is obvious and **M2** and **M3** follow from [17, Theorem 1 and Proposition 4]. If $\delta>0$ is such that $2+2\delta\leq\alpha$, by simple calculations we obtain

$$(\mathbb{E}_x|x_n|^{2+2\delta})^{\frac{1}{2+2\delta}} \leq \varphi(2+2\delta)^{\frac{n}{2+2\delta}}|x| + \frac{\|b_1\|_{2+2\delta}}{1-\varphi(2+2\delta)^{\frac{1}{2+2\delta}}}.$$

Taking the sup over $n\geq 1$, we get

$$\mu_\delta(x) = \sup_{n\geq 1}(\mathbb{E}_x|f(x_n)|^{2+2\delta})^{\frac{1}{2+2\delta}} \leq \varphi(2+2\delta)^{\frac{1}{2+2\delta}}|x| + \frac{\|b_1\|_{2+2\delta}}{1-\varphi(2+2\delta)^{\frac{1}{2+2\delta}}},$$

which proves that **M4** is satisfied. Finally, **M5** holds since

$$\int \mu_\delta(x)^2\,\nu(dx) \leq 2\left(\varphi(2+2\delta)^{\frac{1}{1+\delta}}\int x^2\,\nu(dx) + \left(\frac{\|b_1\|_{2+2\delta}}{1-\varphi(2+2\delta)^{\frac{1}{2+2\delta}}}\right)^2\right) < \infty.$$

The mean is given by $\nu(f)=\mathbb{E}Z=\sum_{i=1}^{\infty}(\mathbb{E}a_1)^{i-1}\mathbb{E}b_1=\frac{\mathbb{E}b_1}{1-\mathbb{E}a_1}$. Without loss of generality we can assume that $\nu(f)=0$, i.e. that $\mathbb{E}b_1=0$; then the variance is

$$\sigma_\nu^2 = \mathrm{Var}_{\mathbb{P}}(Z) = \lim_{n\to\infty}\mathbb{E}\Big(\sum_{i=1}^{n} a_1\cdots a_{i-1}b_i\Big)^2$$
$$= \lim_{n\to\infty}\sum_{i=1}^{n}(\mathbb{E}a_1^2)^{i-1}\mathbb{E}b_1^2 = \frac{\mathbb{E}b_1^2}{1-\mathbb{E}a_1^2}.$$

Therefore the conclusions of Theorem 3.4 hold true with $\mu=\nu(f)=0$ and $\sigma=\sigma_\nu^2=\frac{\mathbb{E}b_1^2}{1-\mathbb{E}a_1^2}$.

A multivariate version of the stochastic recursion has been considered in Guivarc'h and Le Page [18], [19] and can be treated in the same manner.

4. Partition of the set $\mathbb{N}$ and notations. In what follows, $\varepsilon,\beta\in(0,1)$ will be such that $\varepsilon+\beta<1$ (all over the paper ε is supposed to be very small, while β will be optimized). Denote for simplicity $[a,b)=\{l\in\mathbb{N}: a\leq l<b\}$.

Let $k_0 \geq 1$ be a natural number. We start by splitting the set $\mathbb{N}$ into subsets $[2^k, 2^{k+1})$, $k = k_0, k_0+1, \dots$, called *blocks*. Consider the kth block $[2^k, 2^{k+1})$. We leave a large gap $J_{k,1}$ of length $2^{[\varepsilon k]+[\beta k]}$ at the left end of the kth block. Then, following a triadic Cantor-like scheme, we split the remaining part $[2^k + 2^{[\varepsilon k]+[\beta k]}, 2^{k+1})$ into subsets $I_{k,j}$ and $J_{k,j}$ called *islands* and *gaps* as explained below. At resolution level 0 a gap of size $2^{[\varepsilon k]+[\beta k]}/2$ is put in the middle of the interval $[2^k + 2^{[\varepsilon k]+[\beta k]}, 2^{k+1})$. This yields two intervals of equal length. At resolution level 1 two additional gaps of length $2^{[\varepsilon k]+[\beta k]}/2^2$ are put in the middle of each interval obtained, which yields four intervals of equal length. Continuing, at resolution level $[\beta k]$ we obtain $2^{[\beta k]}$ intervals $I_{k,j}$, $j = 1, \dots, 2^{[\beta k]}$, called islands, and the same number of gaps $J_{k,j}$, $j = 1, \dots, 2^{[\beta k]}$ which we index from left to right (recall that $J_{k,1} = J_{k,2^0}$ denotes the large gap at the left end of the kth block). It is obvious that $[2^k, 2^{k+1})$ is the union of the constructed islands and gaps, so that

$$[2^k, 2^{k+1}) = J_{k,1} \cup I_{k,1} \cup \cdots \cup J_{k,2^{[\beta k]}} \cup I_{k,2^{[\beta k]}}. \tag{4.1}$$

Note that in block k there are one gap of length $2^{[[\varepsilon k]]+[\beta k]}$ and 2^l gaps of length $2^{[[\varepsilon k]]+[\beta k]-l-1}$, where $l = 0, \dots, [\beta k] - 1$. The length of the finest gap (for example $J_{k,2^{[\beta k]}}$) is $2^{[\varepsilon k]}$. The total length of the gaps in block k is

$$L_k^{\text{gap}} = 2^{[[\varepsilon k]]+[\beta k]} + \sum_{l=0}^{[\beta k]-1} 2^l 2^{[[\varepsilon k]]+[\beta k]-l-1} = (2 + [\beta k]) 2^{[[\varepsilon k]]+[\beta k]-1}.$$

Recall that, by construction, the islands of the kth block have the same length

$$\begin{aligned} |I_{k,j}| &= \big(2^{k+1} - 2^k - (2 + [\beta k]) 2^{[[\varepsilon k]]+[\beta k]-1}\big)/2^{[\beta k]} \\ &= 2^{k-[\beta k]} - (1 + [\beta k] 2^{[[\varepsilon k]]-1}). \end{aligned}$$

An obvious upper bound is $|I_{k,j}| \leq 2^{k-[\beta k]}$. Since $\varepsilon < 1 - \beta$ we have $|I_{k,j}| \geq 2^{k-[\beta k]} - 2^{[[\varepsilon k]] - c'_{\beta,\varepsilon} \ln k} \geq c_{\varepsilon,\beta} 2^{k(1-\beta)}$, with some $c_{\varepsilon,\beta} \in \big(0, \frac{1}{2}\big)$. Since the length of the kth block is 2^k, the total length of the islands in this block equals

$$L_k^{\text{isl}} = 2^k - 2^{[[\varepsilon k]]+[\beta k]-1}(2 + [\beta k]).$$

Note that, for some constant $c_\beta > 0$,

$$c_\beta 2^k \leq L_k^{\text{isl}} \leq 2^k. \tag{4.2}$$

Denote by $\mathcal{K}$ the set of double indices (k, j), with $k = 1, 2, \dots$ the index of the block and $j = 1, \dots, 2^{[\beta k]}$ the index of the island in block k. The set $\mathcal{K}$ will be endowed with the lexicographical order $\preceq$. Then the sets $I_{k,j}$ and $J_{k,j}$, $(k, j) \in \mathcal{K}$, will also be endowed with the lexicographical order. Let $N \in \mathbb{N}$. From (4.1), there exists a unique $(n, m) \in \mathcal{K}$ such that $2^n \leq N < 2^{n+1}$ and $N \in J_{n,m} \cup I_{n,m}$, where the dependence of n and m on N is suppressed from the notation; let $\mathcal{K}_N = \{(k, j) : (k, j) \preceq (n, m)\}$.

For ease of reading we recall the notations and properties that will be used throughout the paper:

P1. ε and β are positive numbers such that $\varepsilon + \beta < 1$. Later on, the constant ε will be chosen small enough.

P2. $\mathcal{K} = \{(k, j) : k = 1, 2, \ldots, \ j = 1, \ldots, 2^{[\beta k]}\}$.

P3. For any $N \in \mathbb{N}$ the unique couple $(n, m) \in \mathcal{K}$ is such that $N \in J_{n,m} \cup I_{n,m}$.

P4. $\mathcal{K}_N = \{(k, j) : (k, j) \preceq (n, m)\}$.

P5. $I_{k,j}$, $j = 1, \ldots, 2^{[\beta k]}$, are the islands and $J_{k,j}$, $j = 1, \ldots, 2^{[\beta k]}$, are the gaps in the kth block.

P6. The number of islands and the number of gaps in the kth block are both equal to $m_k = 2^{[\beta k]}$. Set $m_{k,n} = m_k + \cdots + m_n$.

P7. The islands in the kth block have the same length $|I_{k,j}| = 2^{k-[\beta k]} - (1 + [\beta k]2^{[[\varepsilon k]]-1}) \leq 2^{k-[\beta k]}$. This implies $|I_{k,j}| \geq c_{\varepsilon,\beta} 2^{k(1-\beta)}$ for some constant $c_{\varepsilon,\beta} \in \left(0, \frac{1}{2}\right)$.

P8. The length of the finest gap in the kth block is $|J_{k,j}| = 2^{[[\varepsilon k]]}$. This implies $|J_{k,j}| \geq 2^{[[\varepsilon k]]}$.

P9. The length $|J_{k,1}|$ of the gap at the left end of the kth block is $2^{[\varepsilon k]+[\beta k]}$.

P10. For each pair $(k, j) \in \mathcal{K}$, we denote $X_{(k,j)} = \sum_{i \in I_{k,j}} X_i$ and $W_{(k,j)} = \sum_{i \in I_{k,j}} W_i$.

P11. $\mathcal{L}_{X_1,\ldots,X_d}$ denotes the probability law of the vector $(X_1, \ldots, X_d)$.

5. Auxiliary result. Without loss of generality we assume that on the initial probability space there is a sequence of independent r.v.'s $(Y_{(k,j)})_{(k,j)\in\mathcal{K}}$ such that $Y_{(k,j)} \stackrel{d}{=} X_{(k,j)}$, $(k, j) \in \mathcal{K}$. Let $k_0 \in \mathbb{N}_+$ and $n > k_0$. Suppose that on the same probability space there is an i.i.d. sequence of $\mathbb{R}^1$-valued r.v.'s $(V_{(k,j)})_{(k,j)\in\mathcal{K}}$ with mean 0 whose characteristic function has support $[-\varepsilon_0, \varepsilon_0]$ and $\mathbb{E}|V_{(k,j)}|^{r_0} < \infty$ for any $r_0 > 0$. We suppose that the sequence $(V_{(k,j)})_{(k,j)\in\mathcal{K}}$ is independent of $(X_{(k,j)})_{(k,j)\in\mathcal{K}}$ and $(Y_{(k,j)})_{(k,j)\in\mathcal{K}}$. Denote $X_{(k)} = (X_{(k,1)}, \ldots, X_{(k,m_k)})$, $Y_{(k)} = (Y_{(k,1)}, \ldots, Y_{(k,m_k)})$ and $V_{(k)} = (V_{(k,1)}, \ldots, V_{(k,m_k)})$. Let π denote the Prokhorov distance (for details see Section 9.1 of the Appendix).

Assume Conditions **C1** and **C2** hold. The main result of this section is the following proposition, which is of independent interest.

PROPOSITION 5.1. *There exists a constant $c_{\varepsilon,\beta,\lambda_1,\lambda_2}$ such that, for any $k_0 = 1, 2, \ldots$ and $n > k_0$,*

$$\pi(\mathcal{L}_{X_{(k_0)}+V_{(k_0)},\ldots,X_{(n)}+V_{(n)}}, \mathcal{L}_{Y_{(k_0)}+V_{(k_0)},\ldots,Y_{(n)}+V_{(n)}}) \leq c_{\varepsilon,\beta,\lambda_1,\lambda_2}(1 + \lambda_0 + \mu_\delta) \exp\left(-\frac{\lambda_1}{4} \cdot 2^{\varepsilon k_0/2}\right).$$

Proof. Without loss of generality we assume that there exists a sequence of independent random vectors $R_{(k)}$, $k = 1, \dots, n$, such that $R_{(k)} \overset{d}{=} X_{(k)} + V_{(k)}$ and $(R_{(k)})_{k=1,\dots,n}$ is independent of $(X_{(k)} + V_{(k)})_{k=1,\dots,n}$, $(Y_{(k,j)})_{(k,j)\in\mathcal{K}}$ and $(V_{(k,j)})_{(k,j)\in\mathcal{K}}$.

The further proof is split into Parts (a) and (b). In Part (a) we give a bound for the Prokhorov distance between $(X_{(k_0)} + V_{(k_0)}, \dots, X_{(n)} + V_{(n)})$ and $(R_{(k_0)} \dots, R_{(n)})$, while in Part (b) we give a bound for the Prokhorov distance between $(R_{(k_0)}, \dots, R_{(n)})$ and $(Y_{(k_0)} + V_{(k_0)}, \dots, Y_{(n)} + V_{(n)})$. Proposition 5.1 follows from (5.1) and (5.9) by the triangle inequality.

Part (a). We show that there exists a constant $c_{\varepsilon,\beta,\lambda_1,\lambda_2}$ such that, for any $k_0 = 1, 2, \dots$ and $n > k_0$,

$$\begin{aligned}(5.1)\quad &\pi(\mathcal{L}_{X_{(k_0)}+V_{(k_0)},\dots,X_{(n)}+V_{(n)}}, \mathcal{L}_{R_{(k_0)}\dots,R_{(n)}}) \\ &\qquad \le c_{\varepsilon,\beta,\lambda_1,\lambda_2}(1 + \lambda_0 + \mu_\delta) \exp\Big(-\frac{\lambda_1}{4} \cdot 2^{\varepsilon k_0/2}\Big).\end{aligned}$$

For $k = k_0, \dots, n$, define $Z_{(k)} = (X_{(k_0)} + V_{(k_0)}, \dots, X_{(k)} + V_{(k)})$ and $\widetilde{Z}_{(k)} = (Z_{(k-1)}, R_{(k)})$. By Lemma 9.3,

$$(5.2)\qquad \pi(\mathcal{L}_{Z_{(n)}}, \mathcal{L}_{R_{(k_0)},\dots,R_{(n)}}) \le \sum_{k=k_0}^{n} \pi(\mathcal{L}_{Z_{(k)}}, \mathcal{L}_{\widetilde{Z}_{(k)}}).$$

Let $\phi_{(k)}$ (resp. $\widetilde{\phi}_{(k)}$) be the characteristic function of the vector $Z_{(k)}$ (resp. $\widetilde{Z}_{(k)}$) and let $m_{k_0,k} = m_{k_0} + \dots + m_k$. Then by Lemma 9.5, for any $T > 0$,

$$\begin{aligned}(5.3)\quad \pi(\mathcal{L}_{Z_{(k)}}, \mathcal{L}_{\widetilde{Z}_{(k)}}) &\le (T/\pi)^{m_{k_0,k}/2}\Big(\int_{t\in\mathbb{R}^{m_{k_0,k}}} |\phi_{(k)}(t) - \widetilde{\phi}_{(k)}(t)|^2\, dt\Big)^{1/2} \\ &\quad + \mathbf{P}\Big(\max_{k_0\le l\le k}\max_{1\le j\le m_l} |X_{(l,j)}| > T\Big).\end{aligned}$$

Denote by $\varphi_{(k)}$ and $\psi_{(k)}$ the characteristic functions of the vectors $X_{(k)}$ and $(X_{(k_0)}, \dots, X_{(k)}))$ respectively. Since $V_{(k_0)}, \dots, V_{(k)}$ are independent of $X_{(k_0)}, \dots, X_{(k)}$ and $Y_{(k_0)}, \dots, Y_{(k)}$, we have

$$\begin{aligned}(5.4)\quad &\int_{t\in\mathbb{R}^{m_{k_0,k}}} |\phi_{(k)}(t) - \varphi_{(k)}(t)|^2\, dt \\ &= \int_{t_1\in\mathbb{R}^{m_{k_0}}} \dots \int_{t_k\in\mathbb{R}^{m_k}} |\phi_{(k)}(t_{k_0}, \dots, t_k) - \varphi_{(k)}(t_{k_0}, \dots, t_k)|^2\, dt_{k_0} \dots dt_k \\ &\le I_1 \equiv \int_{t_1\in\mathbb{R}^{m_{k_0}}} \dots \int_{t_k\in\mathbb{R}^{m_k}} |\psi_{(k)}(t_{k_0}, \dots, t_k) - \psi_{(k-1)}(t_{k_0}, \dots, t_{k-1})\varphi_{(k)}(t_k)|^2 \\ &\qquad\qquad dt_{k_0} \dots dt_k.\end{aligned}$$

To bound the right-hand side of (5.4), note that $m_{k_0,k} = (2^{[\beta k_0]} + \cdots + 2^{[\beta k]}) \leq 2^{[\beta k]+1}$ and, by construction, the length of the gap between the vectors $X_{(k-1)}$ and $X_{(k)}$ is $k_{\mathrm{gap}} = 2^{[\varepsilon k]+[\beta k]}$. Note also that $|I_{k,j}| \leq 2^{k-[\beta k]}$ and $|\varepsilon_0| \leq 1$. Recall that the characteristic functions of the r.v.'s $V_{(k,j)}$ have support $[-\epsilon_0, \epsilon_0]$ and the sequence $(V_{(k,j)})_{(k,j)\in\mathcal{K}}$ is independent of $(X_{(k,j)})_{(k,j)\in\mathcal{K}}$; this readily implies that the integrals above are in fact over $[-\epsilon_0, \epsilon_0]^{m_{k_0,k}}$. Using Condition **C1** with $M_1 = m_{k_0,k-1}$ and $M_2 = m_k$, one may thus write

$$
\begin{aligned}
(5.5)\qquad I_1 &\leq \lambda_0 (1 + \max_{l\leq k,\ j\leq m_k} |I_{l,j}|)^{\lambda_2(M_1+M_2)} \exp(-\lambda_1 k_{\mathrm{gap}}) \varepsilon_0^{m_{k_0,k}} \\
&\leq \lambda_0 (1 + 2^{k-[\beta k]})^{\lambda_2 2^{[\beta k]+1}} \exp(-\lambda_1 k_{\mathrm{gap}}) \\
&\leq \lambda_0 \exp(-\lambda_1 2^{[\varepsilon k]+[\beta k]} + \lambda_2 2^{[\beta k]+1} \ln(1 + 2^{k-[\beta k]})) \\
&\leq c_{\varepsilon,\beta,\lambda_1,\lambda_2} \lambda_0 \exp\left(-\frac{\lambda_1}{2} \cdot 2^{[\varepsilon k]+[\beta k]}\right).
\end{aligned}
$$

Putting together (5.3)–(5.5), we get

$$
\begin{aligned}
(5.6)\qquad \pi(\mathcal{L}_{Z_{(k)}}, \mathcal{L}_{\widetilde{Z}_{(k)}}) &\leq c_{\varepsilon,\beta,\lambda_1,\lambda_2} \lambda_0 (T/\pi)^{m_{k_0,k}/2} \exp\left(-\frac{\lambda_1}{2} \cdot 2^{[\varepsilon k]+[\beta k]}\right) \\
&\quad + \sum_{k_0\leq l\leq k} \sum_{1\leq j\leq m_l} \mathbb{P}(|X_{(l,j)}| > T).
\end{aligned}
$$

Since $|I_{(l,j)}| \leq 2^l$, by Markov's inequality and Condition **C2**,

$$
\mathbb{P}(|X_{(l,j)}| > T) \leq T^{-1} \mathbb{E}|X_{(l,j)}| \leq T^{-1} 2^l \max_i \mathbb{E}|X_i| \leq \mu_\delta T^{-1} 2^l.
$$

Choosing $T = \exp(2^{[\varepsilon k]/2})$, one gets

$$
\begin{aligned}
(5.7)\qquad \sum_{k_0\leq l\leq k} \sum_{1\leq j\leq m_l} \mathbb{P}(|X_{(l,j)}| > T) &\leq \mu_\delta T^{-1} \sum_{k_0\leq l\leq k} m_l 2^l \\
&\leq \mu_\delta \exp(-2^{[\varepsilon k]/2}) \sum_{k_0\leq l\leq k} 2^{[\beta l]} 2^l \leq c_\beta \mu_\delta \exp(-2^{[\varepsilon k]/2}/2).
\end{aligned}
$$

Since $m_{k_0,k} \leq 2^{\beta k}$, one gets

$$
(5.8)\qquad (T/\pi)^{m_{k_0,k}/2} \leq \exp\left(\frac{1}{2} \cdot 2^{[\varepsilon k]/2+\beta k}\right).
$$

From (5.6)–(5.8), we deduce

$$
\begin{aligned}
\pi(\mathcal{L}_{Z_{(k)}}, \mathcal{L}_{\widetilde{Z}_{(k)}}) &\leq c_{\varepsilon,\beta,\lambda_1,\lambda_2} \lambda_0 \exp\left(\frac{1}{2} \cdot 2^{[\varepsilon k]/2+[\beta k]}\right) \exp\left(-\frac{\lambda_1}{2} \cdot 2^{[\varepsilon k]+[\beta k]}\right) \\
&\quad + c_\beta \mu_\delta \exp(-2^{[\varepsilon k]/2}/2) \\
&\leq (1 + \lambda_0 + \mu_\delta) c_{\varepsilon,\beta,\lambda_1,\lambda_2} \exp\left(-\frac{\lambda_1}{4} \cdot 2^{\varepsilon k/2}\right).
\end{aligned}
$$

Using (5.2) leads to

$$\begin{aligned}\pi(\mathcal{L}_{Z_{(n)}},\mathcal{L}_{(R_{(k_0)}\dots,R_{(n)})}) &\le (1+\lambda_0+\mu_\delta)c_{\varepsilon,\beta,\lambda_1,\lambda_2}\sum_{k=k_0}^{n}\exp\biggl(-\frac{\lambda_1}{4}\cdot 2^{[\varepsilon k]/2}\biggr)\\ &\le (1+\lambda_0+\mu_\delta)c'_{\varepsilon,\beta,\lambda_1,\lambda_2}\exp\biggl(-\frac{\lambda_1}{4}\cdot 2^{[\varepsilon k]_0/2}\biggr).\end{aligned}$$

This concludes the proof of Part (a).

PART (b). We show that there exists a constant $c_{\varepsilon,\beta,\lambda_1,\lambda_2}$ such that, for any $k_0=1,2,\dots$ and $n>k_0$,

$$\begin{aligned}\pi(\mathcal{L}_{R_{(k_0)},\dots,R_{(n)}},\mathcal{L}_{Y_{(k_0)}+V_{(k_0)},\dots,Y_{(n)}+V_{(n)}})\\ \le c_{\varepsilon,\beta,\lambda_1,\lambda_2}(1+\lambda_0+\mu_\delta)\exp\biggl(-\frac{\lambda_1}{8}\cdot 2^{[\varepsilon k]_0/2}\biggr).\end{aligned}\tag{5.9}$$

By Lemma 9.4, since $R_{(k_0)},\dots,R_{(n)}$ and $Y_{(k_0)}+V_{(k_0)},\dots,Y_{(n)}+V_{(n)}$ are independent r.v.'s, one may write

$$\pi(\mathcal{L}_{R_{(k_0)},\dots,R_{(n)}},\mathcal{L}_{Y_{(k_0)}+V_{(k_0)},\dots,Y_{(n)}+V_{(n)}})=\sum_{k=k_0}^{n}\pi(\mathcal{L}_{R_{(k)}},\mathcal{L}_{Y_{(k)}+V_{(k)}})\tag{5.10}$$

and it suffices to prove that, for any $k=1,2,\dots,$

$$\pi(\mathcal{L}_{R_{(k)}},\mathcal{L}_{Y_{(k)}+V_{(k)}})\le(1+\lambda_0+\mu_\delta)c'_{\varepsilon,\beta,\lambda_1,\lambda_2}\exp\biggl(-\frac{\lambda_1}{8}\cdot 2^{[\varepsilon k]/2}\biggr).\tag{5.11}$$

For this, recall that, according to the construction in Section 4, at resolution level 0, a gap of length $2^{[[\varepsilon k]]+[\beta k]}/2$ in the middle of the block $R^{0,0}_{(k)}=R_{(k)}$ splits it into two vectors $\widetilde{R}^{0,1}_{(k)}$ and $\widetilde{R}^{0,2}_{(k)}$; let $R^{0,1}_{(k)}$ and $R^{0,2}_{(k)}$ be independent versions of $\widetilde{R}^{0,1}_{(k)}$ and $\widetilde{R}^{0,2}_{(k)}$ respectively. Next, at level 1, for any $j\in\{1,2\}$, a gap of length $2^{[(\varepsilon+\beta)k]}/4$ in the middle of the block $R^{0,j}_{(k)}$ splits it into two vectors $\widetilde{R}^{1,2j-1}_{(k)}$ and $\widetilde{R}^{1,2j}_{(k)}$; let $R^{1,2j-1}_{(k)}$ and $R^{1,2j}_{(k)}$ be their independent versions. Assuming that at level $l\in\{1,\dots,[\beta k]\}$ the independent r.v.'s $R^{l,j}_{(k)}$, $j\in\{1,\dots,2^l\}$, are already constructed, we shall perform the construction at resolution level $l+1$. Note that, at level l, for any $j\in\{1,\dots,2^l\}$, a gap of length $2^{[(\varepsilon+\beta)k]}/2^{l+1}$ in the middle of the block $R^{l,j}_{(k)}$ splits it into two vectors $\widetilde{R}^{l+1,2j-1}_{(k)}$ and $\widetilde{R}^{l+1,2j}_{(k)}$; it is enough to let $R^{l+1,2j-1}_{(k)}$ and $R^{l+1,2j}_{(k)}$ be their independent versions. It is easy to see that at the final level $l_k=[\beta k]$ we have $R^{l_k,j}_{(k)}\overset{d}{=}Y_{(k,j)}+V_{(k,j)}$ for $j=1,\dots,m_k=2^{[\beta k]}$.

Let $l\in\{0,\dots,[\beta k]\}$. For $j\in\{1,\dots,2^l\}$, denote by $\psi^{l,2j-1}_k$ and $\psi^{l,2j}_k$ the characteristic functions of $R^{l,2j-1}_{(k)}$ and $R^{l,2j}_{(k)}$. Using Lemma 9.5 and the independence of $\widetilde{R}_{(k)}^{l,2j-1}$ and $\widetilde{R}_{(k)}^{l,2j}$, we get

$$\begin{aligned}
(5.12)\qquad & \pi(\mathcal{L}_{R_{(k)}^{l,j}}, \mathcal{L}_{R_{(k)}^{l+1,2j-1}, R_{(k)}^{l+1,2j}}) \\
& \leq \Big((T/\pi)^{2^{-l} m_k} \int_{(t,s)\in\mathbb{R}^{2^{-l} m_k}} |\psi_k^{l,j}(t,s) - \psi_k^{l+1,2j-1}(t)\psi_k^{l+1,2j}(s)|^2 \, dt \, ds \Big)^{1/2} \\
& \quad + \sum_{1\leq j\leq 2^{-l} m_k} \mathbb{P}(|X_{(k,j)} + V_{(k,j)}| > T).
\end{aligned}$$

By Condition **C1** with $N = M = \frac{m_k}{2} 2^{-l}$ and $k_{\text{gap}} = 2^{[\varepsilon k]+[\beta k]-l-1}$, we obtain

$$\begin{aligned}
(5.13)\qquad & \int_{(t,s)\in\mathbb{R}^{m_k 2^{-l}}} |\psi_k^{l,j}(t,s) - \psi_k^{l+1,2j-1}(t)\psi_k^{l+1,2j}(s)|^2 \, dt \, ds \\
& = \int_{(t,s)\in\mathbb{R}^{m_k 2^{-l}},\, \|t\|_\infty\leq\varepsilon_0,\, \|s\|_\infty\leq\varepsilon_0} |\psi_k^{l,j}(t,s) - \psi_k^{l+1,2j-1}(t)\psi_k^{l+1,2j}(s)|^2 \, dt \, ds \\
& \leq \lambda_0 \exp\Big(\lambda_2 m_k 2^{-l} \ln(1 + 2^{k-[\beta k]}) - \lambda_1 2^{[\varepsilon k]+[\beta k]-1-l}\Big) (2\varepsilon_0)^{m_k 2^{-l}} \\
& \leq \lambda_0 c''_{\varepsilon,\beta,\lambda_1,\lambda_2} \exp\Big(-\frac{\lambda_1}{4} \cdot 2^{[\varepsilon k]+[\beta k]-l}\Big).
\end{aligned}$$

We thus take $T = \exp(\lambda_1 2^{[[\varepsilon k]]/2})$ so that

$$(T/\pi)^{2^{-l} m_k} \leq \exp(\lambda_1 2^{-l} m_k 2^{[[\varepsilon k]]/2}) \leq \exp(\lambda_1 2^{[[\varepsilon k]]/2+[\beta k]-l}).$$

In order to control the terms $\mathbb{P}(|X_{(k,j)} + V_{(k,j)}| > T)$, we use Markov's inequality, Condition **C2** and the fact that $|I_{k,j}| \leq 2^k$; it readily follows that

$$\begin{aligned}
\mathbb{P}(|X_{(k,j)} + V_{(k,j)}| > T) & \leq T^{-1} (\mathbb{E}|X_{(k,j)}| + \mathbb{E}|V_{(k,j)}|) \\
& \leq T^{-1} \Big(2^k \max_i \mathbb{E}|X_i| + c 2^k \Big) \\
& \leq (1 + \mu_\delta) c 2^k \exp(-\lambda_1 2^{-[[\varepsilon k]]/2}).
\end{aligned}$$

Therefore

$$\begin{aligned}
(5.14)\qquad & \sum_{1\leq j\leq 2^{-l} m_k} \mathbb{P}(|X_{(k,j)} + V_{(k,j)}| > T) \\
& \leq 2^{-l} m_k (1 + \mu_\delta) c 2^k \exp(-2^{[[\varepsilon k]]/2}) \\
& \leq (1 + \mu_\delta) \exp(-\lambda_1 2^{[[\varepsilon k]]/2}) 2^{-l} 2^{2[\beta k]+k} \\
& \leq (1 + \mu_\delta) c_{\varepsilon,\beta,\lambda_1,\lambda_2} \exp\Big(-\frac{\lambda_1}{2} 2^{[[\varepsilon k]]/2}\Big).
\end{aligned}$$

From (5.12)–(5.14), we get

$$(5.15)\quad \pi(\mathcal{L}_{R^{l,j}_{(k)}}, \mathcal{L}_{R^{l+1,2j-1}_{(k)}, R^{l+1,2j}_{(k)}})$$
$$\le (1+\lambda_0+\mu_\delta)c_{\varepsilon,\beta,\lambda_1,\lambda_2}$$
$$\times \left[\exp(\lambda_1 2^{-l} 2^{[[\varepsilon k]]/2+[\beta k]}) \exp\left(-\frac{\lambda_1}{2}\cdot 2^{-l}2^{[\varepsilon k]+[\beta k]}\right) + c\exp\left(-\frac{\lambda_1}{2}\cdot 2^{[[\varepsilon k]]/2}\right)\right]$$
$$\le (1+\lambda_0+\mu_\delta)c_{\varepsilon,\beta,\lambda_1,\lambda_2}\exp\left(-\frac{\lambda_1}{4}\cdot 2^{[[\varepsilon k]]/2}\right).$$

Since $R^{l,j}_{(k)}$, $j=1,\dots,2^l$, are independent r.v.'s, by the triangle inequality one gets

$$(5.16)\quad \pi(\mathcal{L}_{R_{(k)}}, \mathcal{L}_{Y_{(k)}+V_{(k)}}) = \pi(\mathcal{L}_{R^{0,0}_{(k)}}, \mathcal{L}_{Y_{(k)}+V_{(k)}})$$
$$\le \pi(\mathcal{L}_{R^{0,0}_{(k)}}, \mathcal{L}_{R^{0,1}_{(k)},R^{0,2}_{(k)}}) + \pi(\mathcal{L}_{R^{0,1}_{(k)},R^{0,2}_{(k)}}, \mathcal{L}_{Y_{(k)}+V_{(k)}})$$
$$\le \pi(\mathcal{L}_{R^{0,0}_{(k)}}, \mathcal{L}_{R^{0,1}_{(k)},R^{0,2}_{(k)}}) + \pi(\mathcal{L}_{R^{0,1}_{(k)},R^{0,2}_{(k)}}, \mathcal{L}_{R^{1,1}_{(k)},\dots,R^{1,4}_{(k)}})$$
$$+ \pi(\mathcal{L}_{R^{1,1}_{(k)},\dots,R^{1,4}_{(k)}}, \mathcal{L}_{Y_{(k)}+V_{(k)}})$$
$$\vdots$$
$$\le \sum_{l=0}^{[\beta k]-1} \pi(\mathcal{L}_{R^{l,1}_{(k)},\dots,R^{l,2^l}_{(k)}}, \mathcal{L}_{R^{l+1,1}_{(k)},\dots,R^{l+1,2^{l+1}}_{(k)}}).$$

By Lemma 9.4 and (5.15),

$$(5.17)\quad \pi(\mathcal{L}_{R^{l,1}_{(k)},\dots,R^{l,2^l}_{(k)}}, \mathcal{L}_{R^{l+1,1}_{(k)},\dots,R^{l+1,2^{l+1}}_{(k)}})$$
$$\le \sum_{j=1}^{2^l} \pi(\mathcal{L}_{R^{l,j}_{(k)}}, \mathcal{L}_{R^{l+1,2j-1}_{(k)},R^{l+1,2j}_{(k)}})$$
$$\le c_{\varepsilon,\beta,\lambda_1,\lambda_2} 2^l (1+\lambda_0+\mu_\delta)\exp\left(-\frac{\lambda_1}{4}\cdot 2^{[\varepsilon k]/2}\right).$$

From (5.16) and (5.17), it follows that

$$\pi(\mathcal{L}_{R_{(k)}}, \mathcal{L}_{Y_{(k)}+V_{(k)}}) \le \sum_{l=0}^{[\beta k]-1} 2^l(1+\lambda_0+\mu_\delta)c_{\varepsilon,\beta,\lambda_1,\lambda_2}\exp\left(-\frac{\lambda_1}{4}\cdot 2^{[\varepsilon k]/2}\right)$$
$$\le 2^{[\beta k]}(1+\lambda_0+\mu_\delta)c_{\varepsilon,\beta,\lambda_1,\lambda_2}\exp\left(-\frac{\lambda_1}{4}\cdot 2^{[\varepsilon k]/2}\right)$$
$$\le (1+\lambda_0+\mu_\delta)c'_{\varepsilon,\beta,\lambda_1,\lambda_2}\exp\left(-\frac{\lambda_1}{8}\cdot 2^{[\varepsilon k]/2}\right).$$

Finally, using (5.11) finishes the proof of Part (b). ■

6. Proof of Theorem 2.1. The proof is divided into several steps. We first construct the coupling with independent r.v.'s. (Section 6.1) and then with independent normal r.v.'s. (Section 6.2). In Section 6.3, we give an explicit construction of the sequences $(\widetilde{X}_i)_{1\le i\le N}$ and $(W_i)_{1\le i\le N}$ and in Sections 6.4–6.7 we put together and optimize the bounds.

6.1. Coupling with independent r.v.'s. Assume Conditions **C1** and **C2** hold. The proposition below shows that the partial sums $\sum_{(l,i)\preceq(k,j)} X_{(l,i)}$ can be coupled with high probability with the partial sums $\sum_{(l,i)\preceq(k,j)} Y_{(l,i)}$.

PROPOSITION 6.1. *Let $\alpha<\delta$, $\beta>1/2$ and $0<\rho<(1-\beta)/2$. Then, for any $N\in\mathbb{N}$, on some extension of the initial probability space there is a version $(X'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ of $(X_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ and a version $(Y'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ of $(Y_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ such that*

$$\mathbb{P}\Big((2^n)^{-1/2}\sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,i)\preceq(k,j)}(X'_{(l,i)}-Y'_{(l,i)})\Big|\ge(2^n)^{-\rho}\Big)\le C_1(2^n)^{-1-\alpha+(\varepsilon+\rho)(2+2\alpha)},$$

where $\varepsilon\in(0,1/2)$ is arbitrary and $C_1=c_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\rho}(1+\lambda_0+\mu_\delta)^{2+2\delta}$ for some positive constant $c_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\rho}$.

Proof. It is convenient to set $k_0=[\varepsilon n]$, $X_{k_0,n}=(X_{(k_0)},\dots,X_{(n)})$, $Y_{k_0,n}=(Y_{(k_0)},\dots,Y_{(n)})$ and $V_{k_0,n}=(V_{(k_0)},\dots,V_{(n)})$; the variables $\widetilde{X}_{k_0,n}=X_{k_0,n}+V_{k_0,n}$ and $\widetilde{Y}_{k_0,n}=Y_{k_0,n}+V_{k_0,n}$ are the smoothed versions of $X_{k_0,n}$ and $Y_{k_0,n}$. By Proposition 5.1, with $k_0=[\varepsilon n]$, there exists a constant $c_{\varepsilon,\beta,\lambda_1,\lambda_2}$ such that

$$\pi(\mathcal{L}_{\widetilde{X}_{k_0,n}},\mathcal{L}_{\widetilde{Y}_{k_0,n}})\le\varDelta=(1+\lambda_0+\mu_\delta)c_{\varepsilon,\beta,\lambda_1,\lambda_2}\exp\Big(-\frac{\lambda_1}{4}\cdot2^{\varepsilon^2n/2}\Big).\tag{6.1}$$

Using Strassen–Dudley's theorem (see Lemma 9.1), we conclude that on some extension of the initial probability space there are random vectors $\widetilde{S}_{k_0,n}=(S_{(k_0)},\dots,S_{(n)})$ and $\widetilde{T}_{k_0,n}=(T_{(k_0)},\dots,T_{(n)})$ such that $\widetilde{S}_{k_0,n}\overset{d}{=}\widetilde{X}_{k_0,n}$, $\widetilde{T}_{k_0,n}\overset{d}{=}\widetilde{Y}_{k_0,n}$ and

$$\mathbb{P}(\|\widetilde{S}_{k_0,n}-\widetilde{T}_{k_0,n}\|_\infty\ge\varDelta)\le\varDelta.\tag{6.2}$$

We shall remove the smoothing from the vectors $\widetilde{S}_{k_0,n}$ and $\widetilde{T}_{k_0,n}$. Without loss of generality we may assume that there is a random vector U with uniform distribution on $[0,1]^{m_{k_0,n}}$ and independent of $(\widetilde{S}_{k_0,n},\widetilde{T}_{k_0,n})$. We thus consider the transition kernels $G_1(x\,|\,y):=\mathbb{P}(X_{k_0,n}\le x\,|\,\widetilde{X}_{k_0,n}=y)$ and $G_2(x\,|\,y):=\mathbb{P}(Y_{k_0,n}\le x\,|\,\widetilde{Y}_{k_0,n}=y)$ and set $X'_{k_0,n}:=G_1^{-1}(U\,|\,\widetilde{S}_{k_0,n})$, $V'_{k_0,n}:=\widetilde{S}_{k_0,n}-X'_{k_0,n}$, $Y'_{k_0,n}:=G_1^{-1}(U\,|\,T_{k_0,n})$ and $V''_{k_0,n}:=\widetilde{T}_{k_0,n}-Y'_{k_0,n}$.

The sequences $X'_{k_0,n}$ and $Y'_{k_0,n}$ are such that $\widetilde{S}_{k_0,n} = X'_{k_0,n} + V'_{k_0,n}$, $\widetilde{T}_{k_0,n} = Y'_{k_0,n} + V''_{k_0,n}$ and $X'_{k_0,n} \stackrel{d}{=} X_{k_0,n}$, $Y'_{k_0,n} \stackrel{d}{=} Y_{k_0,n}$, $V'_{k_0,n} \stackrel{d}{=} V''_{k_0,n} \stackrel{d}{=} V_{k_0,n}$. The coordinates of the vectors $X'_{k_0,n}$ and $Y'_{k_0,n}$ are denoted by $X'_{(k,j)}$ and $Y'_{(k,j)}$, $(k,j) \in \mathcal{K}$. Since $\widetilde{S}_{(k,j)} = X'_{(k,j)} + V'_{(k,j)}$ and $\widetilde{T}_{(k,j)} = Y'_{(k,j)} + V''_{(k,j)}$, we have, for any $x \geq 1$,

$$R = \mathbb{P}\Big(\sup_{k_0 \leq k,\, (k,j) \in \mathcal{K}_N} \Big| \sum_{(l,i) \preceq (k,j)} (X'_{(l,i)} - Y'_{(l,i)}) \Big| \geq 2x \Big) \leq R_1 + R_2,$$

where

$$R_1 = \mathbb{P}\Big(\sup_{k_0 \leq k,\, (k,j) \in \mathcal{K}_N} \Big| \sum_{(l,i) \preceq (k,j)} (\widetilde{S}_{(l,i)} - \widetilde{T}_{(l,i)}) \Big| \geq x \Big),$$

$$R_2 = \mathbb{P}\Big(\sup_{k_0 \leq k,\, (k,j) \in \mathcal{K}_N} \Big| \sum_{(l,i) \preceq (k,j)} (V'_{(l,i)} - V''_{(l,i)}) \Big| \geq x \Big).$$

First, we shall control R_1. Note that $\operatorname{card} \mathcal{K}_N \leq c 2^{\beta n}$. For any sequence $(\alpha_{(k,j)})_{(k,j) \in \mathcal{K}}$ of positive numbers such that $\sum_{(k,j) \in \mathcal{K}} \alpha_{(k,j)} \leq 1$,

$$\begin{aligned} \Big\{ \sup_{k_0 \leq k,\, (k,j) \in \mathcal{K}_N} & \Big| \sum_{k_0 \leq l,\, (l,i) \preceq (k,j)} (\widetilde{S}_{(l,i)} - \widetilde{T}_{(l,i)}) \Big| \geq x \Big\} \\ &\subseteq \bigcup_{k_0 \leq k,\, (k,j) \in \mathcal{K}_N} \Big\{ \Big| \sum_{k_0 \leq l,\, (l,i) \preceq (k,j)} (\widetilde{S}_{(l,i)} - \widetilde{T}_{(l,i)}) \Big| \geq x \Big\} \\ &\subseteq \bigcup_{k_0 \leq k,\, (k,j) \in \mathcal{K}_N} \bigcup_{k_0 \leq k,\, (l,i) \preceq (k,j)} \{ |\widetilde{S}_{(l,i)} - \widetilde{T}_{(l,i)}| \geq x \alpha_{(l,i)} \} \\ &= \bigcup_{k_0 \leq k,\, (k,j) \in \mathcal{K}_N} \{ |\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}| \geq x \alpha_{(k,j)} \}, \end{aligned}$$

which implies that

$$R_1 \leq \sum_{k_0 \leq k,\ (k,j) \in \mathcal{K}_N} \mathbb{P}(|\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}| \geq x \alpha_{(k,j)}).$$

Let $p = 2 + 2\alpha < 2 + 2\delta$. By Chebyshev's inequality,

$$R_1 \leq x^{-p} \sum_{k_0 \leq k,\ (k,j) \in \mathcal{K}_N} \alpha_{(k,j)}^{-p} \mathbb{E} |\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}|^p.$$

By a truncation argument, with Δ from (6.1) and (6.2),

$$\begin{aligned} R_1 \leq x^{-p} \Delta^p & \sum_{k_0 \leq k,\ (k,j) \in \mathcal{K}_N} \alpha_{(k,j)}^{-p} \\ & + x^{-p} \sum_{k_0 \leq k,\ (k,j) \in \mathcal{K}_N} \alpha_{(k,j)}^{-p} \mathbb{E} |\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}|^p \mathbb{1}(|\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}| \geq \Delta). \end{aligned}$$

Let $\eta \in (0, \delta - \alpha)$, $p' = p + 2\eta$ and $\gamma = \frac{2\eta}{p+2\eta} \leq \eta$. Applying Hölder's inequality one may write

$$\big\| |\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}| 1(|\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}| \geq \Delta) \big\|_{L^p} \leq \|\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}\|_{L^{p'}} \mathbb{P}(|\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}| > \Delta)^{\gamma/p}.$$

By Condition **C2**, for some constant $c > 0$, we get

$$\|\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}\|_{L^{p'}} \leq 2\|X_{(k,j)}\|_{L^{p'}} + 2\|V'_{(k,j)}\|_{L^{p'}} \leq c(1+\mu_\delta)|I_{k,j}|;$$

consequently, (6.2) leads to

$$\begin{aligned} R_1 &\leq x^{-p}\Delta^p \sum_{k_0 \leq k,\ (k,j)\in\mathcal{K}_N} \alpha_{(k,j)}^{-p} \\ &\quad + c(1+\mu_\delta)^p x^{-p} \sum_{k_0 \leq k,\ (k,j)\in\mathcal{K}_N} \alpha_{(k,j)}^{-p} |I_{k,j}|^p (\mathbf{P}(|\widetilde{S}_{(k,j)} - \widetilde{T}_{(k,j)}| \geq \Delta))^\gamma \\ &\leq x^{-p}\Delta^p \sum_{k_0 \leq k,\ (k,j)\in\mathcal{K}_N} \alpha_{(k,j)}^{-p} \\ &\quad + c(1+\mu_\delta)^p x^{-p}\Delta^\gamma \sum_{k_0 \leq k,\ (k,j)\in\mathcal{K}_N} \alpha_{(k,j)}^{-p} 2^{(k-[\beta k])p} \\ &\leq c_{\varepsilon,\beta,\lambda_1,\lambda_2,\eta}(1+\lambda_0+\mu_\delta)^{p+\gamma} \exp\Big(-\frac{\lambda_1}{4}\gamma 2^{\varepsilon^2 n/2}\Big) x^{-p} \sum_{k_0 \leq k \leq n} \sum_{j \leq 2^{[\beta k]}} \alpha_{(k,j)}^{-p} 2^{kp}. \end{aligned}$$

Now, choosing $\alpha_{(k,j)} = 2^{-k}j^{-2}$, we obtain

$$\begin{aligned} \sum_{k_0 \leq k \leq n} \sum_{j \leq 2^{[\beta k]}} \alpha_{(k,j)}^{-p} 2^{kp} &\leq \sum_{k_0 \leq k \leq n} \sum_{j \leq 2^{[\beta k]}} 2^{2kp} j^{2p} \leq 2^{2np} \sum_{k_0 \leq k \leq n} \sum_{j \leq 2^{[\beta k]}} j^{2p} \\ &\leq 2^{2np} \sum_{k_0 \leq k \leq n} 2^{(2p+1)[\beta k]} \leq 2^{2np} 2^{(2p+1)[\beta n]} n \leq 2^{nc_{\alpha,\beta}}, \end{aligned}$$

which implies that

$$R_1 \leq c_{\varepsilon,\beta,\lambda_1,\lambda_2,\eta}(1+\lambda_0+\mu_\delta)^{p+\gamma} \exp\Big(-\frac{1}{4}\gamma\lambda_1 \cdot 2^{\varepsilon^2 n/2}\Big) 2^{nc_{\alpha,\beta}} x^{-p}.$$

Since $\gamma = \frac{2\eta}{p+2\eta} \leq \eta \leq p\eta$ and $x \geq 1$, we conclude that

$$R_1 \leq A' \exp\Big(-\frac{1}{4}\gamma\lambda_1 (2^n)^{\varepsilon^2/2}\Big) \tag{6.3}$$

for some $A' = c'_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\alpha',\eta}(1+\lambda_0+\mu_\delta)^{p(1+\eta)}$.

Now, we shall control R_2. Using Doob's inequality, for any $\lambda > 2$,

$$R_2 \le 2\mathbb{P}\Big(\sup_{k_0\le k,\,(k,j)\in\mathcal{K}_N}\Big|\sum_{k_0\le l,\,(l,i)\preceq(k,j)} V'_{(l,i)}\Big| \ge x\Big) \le 2x^{-\lambda}\mathbb{E}\Big(\sum_{k_0\le l,\,(l,i)\in\mathcal{K}_N}|V'_{(l,i)}|\Big)^{\lambda}. \tag{6.4}$$

By Rosenthal's inequality

$$\begin{aligned}\Big(\mathbb{E}\Big(\sum_{k_0\le l,\,(l,i)\in\mathcal{K}_N}|V'_{(l,i)}|\Big)^{\lambda}\Big)^{1/\lambda} &\le c_\lambda\Big(\sum_{k_0\le l,\,(l,i)\in\mathcal{K}_N}\mathbb{E}(|V'_{(l,i)}|^2)\Big)^{1/2} \\ &\quad + c_\lambda\Big(\sum_{k_0\le l,\,(l,i)\in\mathcal{K}_N}\mathbb{E}(|V'_{(l,i)}|^{\lambda})\Big)^{1/\lambda} \\ &\le c'_\lambda(2^{\beta n})^{1/2}.\end{aligned} \tag{6.5}$$

From (6.3)–(6.5) we obtain

$$\mathbb{P}\Big(\sup_{k_0\le k,\;(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,i)\preceq(k,j)}(X'_{(l,i)}-Y'_{(l,i)})\Big| \ge 2x\Big) \le A'\exp\Big(-\frac{1}{4}\gamma\lambda_1(2^n)^{\varepsilon^2/2}\Big) + c_\lambda(2^{\beta n})^{\lambda/2}x^{-\lambda}.$$

Choosing $x = \frac{1}{2}(2^n)^{1/2-\rho}$, we find

$$\mathbb{P}\Big((2^n)^{-1/2}\sup_{k_0\le k,\,(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,i)\preceq(k,j)}(X'_{(l,i)}-Y'_{(l,i)})\Big| \ge (2^n)^{-\rho}\Big) \le A'\exp\Big(-\frac{1}{4}\gamma\lambda_1(2^n)^{\varepsilon^2/2}\Big) + c_\lambda(2^n)^{-\frac{1}{2}\lambda(1-\beta-2\rho)}. \tag{6.6}$$

So far we performed the construction for $k \ge k_0$. It remains to construct the sequences $X'_{(k,j)}$ and $Y'_{(k,j)}$ for $(k,j) \preceq (k_0-1, m_{k_0-1})$. This construction can be performed by any method such that the sequences $(X'_{(k,j)})$ and $(Y'_{(k,j)})$, where $(k,j) \preceq (k_0-1, m_{k_0-1})$, are independent and $Y'_{(k,j)} \overset{d}{=} X_{(k,j)}$ for the same (k,j). Indeed, let $F_{X|Y_1,\ldots,Y_k}(x|y_1,\ldots,y_k)$ be the conditional distribution of X given $[Y_1 = y_1,\ldots,Y_k = y_k]$ and let $(U_{(k,j)})$ be a sequence of independent r.v.'s uniformly distributed on $(0,1)$. Denote for brevity the constructed part by $\mathbf{X}'_{k_0} = (X'_{(k,j)})_{k_0\le k,\,(k,j)\in\mathcal{K}_N}$. Define $X'_{(k_0-1,1)}$ as the conditional quantile transform

$$X'_{(k_0-1,1)} = F^{-1}_{X_{(k_0-1,1)}|\mathbf{X}_{k_0}}(U_{(k_0-1,1)}|\mathbf{X}'_{k_0}),$$

where $\mathbf{X}_{k_0} = (X_{(k,j)})_{k_0\le k,\,(k,j)\in\mathcal{K}_N}$. We continue setting

$$X'_{(k_0-1,j)} = F^{-1}_{X_{(k,j)}|\mathbf{X}_{k_0}}(U_{(k_0-1,j-1)}|X'_{(k_0-1,1)},\ldots,X'_{(k_0-1,j-1)}\mathbf{X}'_{k_0})$$

for $j = 2, \dots, m_{k_0-1}$. In the same way we extend the construction to all $X'_{(k,j)}$ with $1 \le k < k_0 - 1$. The construction of the sequence $(Y'_{(k,j)})$ for $(k,j) \preceq (k_0 - 1, m_{k_0-1})$ is similar.

Since the sequence $(X_k)_{k\ge1}$ satisfies Condition **C1**, so does $(X'_{(k,j)})$. Using the maximal inequality stated in Proposition 8.1 below and noting that the cardinality of the set $\{(k,j) : (k,j) \preceq (k_0 - 1, m_{k_0-1})\}$ is less than or equal to $2^{\beta k_0} \le 2^{\varepsilon n}$, we obtain, for any $\eta' \in \big(0, \frac{\delta-\alpha}{(2+\alpha+\delta)^2}\big)$,

$$\mathbb{E}\Big(\sup_{(k,j)\preceq(k_0-1,m_{k_0-1})}\Big|\sum_{(l,i)\preceq(k,j)} X'_{(l,i)}\Big|^p\Big) \le A''(2^{\varepsilon n})^{\frac{1}{2}p}$$

for come constant $A'' = c''_{\varepsilon,\beta,\lambda_1,\lambda_2,\delta,\alpha,\eta}(1 + \lambda_0 + \mu_\delta)^{p(1+\eta')}$. By Chebyshev's inequality, for any $x > 0$ we get

$$\mathbb{P}\Big(\sup_{(k,j)\preceq(k_0-1,m_{k_0-1})}\Big|\sum_{(l,i)\preceq(k,j)} X'_{(l,i)}\Big| \ge x\Big)$$
$$\le x^{-p}\mathbb{E}\sup_{(k,j)\preceq(k_0-1,m_{k_0-1})}\Big|\sum_{(l,i)\preceq(k,j)} X'_{(l,i)}\Big|^p \le A''x^{-p}(2^{\varepsilon n})^{\frac{1}{2}p}.$$

Substituting $x = (2^n)^{1/2-\rho}$ yields

$$\mathbb{P}\Big((2^n)^{-1/2}\sup_{(k,j)\preceq(k_0-1,m_{k_0-1})}\Big|\sum_{(l,i)\preceq(k,j)} X'_{(l,i)}\Big| \ge (2^n)^{-\rho}\Big) \le A''(2^n)^{-\frac{p}{2}+p(\rho+\frac{1}{2}\varepsilon)}.$$

A similar inequality can be proved with $Y'_{(l,i)}$ instead of $X'_{(l,i)}$. Combining this with (6.6), we obtain

$$\mathbb{P}\Big((2^n)^{-1/2}\sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,i)\preceq(k,j)} (X'_{(l,i)} - Y'_{(l,i)})\Big| \ge 2(2^n)^{-\rho}\Big) \tag{6.7}$$
$$\le A'''\Big(\exp\Big(-\frac{1}{4}\gamma\lambda_1(2^n)^{\varepsilon^2/2}\Big) + (2^n)^{-\frac{1}{2}\lambda(1-\beta-2\rho)} + (2^n)^{-\frac{p}{2}+p(\rho+\frac{1}{2}\varepsilon)}\Big)$$

for some $A'' = c'''_{\varepsilon,\beta,\lambda_1,\lambda_2,\delta,\delta',\eta,\lambda}(1 + \lambda_0 + \mu_\delta)^{p(1+\eta+\eta')}$. Recall that $p = 2 + 2\alpha$, $\alpha < \delta$, $\beta > 1/2$ and $\rho < (1 - \beta)/2$. Taking $\lambda = \frac{2+2\alpha}{1-\beta-2\rho} > p$, the right-hand side of (6.7) does not exceed $A'''(2^n)^{-1-\alpha+(2+2\alpha)(\rho+\varepsilon)}$. It remains to choose a sufficiently small $\eta + \eta'$ such that $p(1 + \eta + \eta') \le 2 + 2\delta$, which implies

$$A''' \le c'''_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\rho}(1 + \lambda_0 + \mu_\delta)^{2+2\delta}.$$

The assertion of Proposition 6.1 follows. ■

6.2. Coupling with independent normal r.v.'s. Assume Conditions **C1**–**C3** hold. Without loss of generality we can consider that $\mu_i = 0$, $i \ge 1$, and $\sigma = 1$. The following proposition shows that with high probability the partial sums $\sum_{(l,i)\preceq(k,j)} X_{(l,i)}$ can be coupled with the partial sums of some normal r.v.'s. Note the presence of two terms in the upper bound below. One of them, called the *dependence error*, comes from replacing dependent blocks by independent ones; the second one, called *Sakhanenko's error*, is due to the use of Sakhanenko's strong approximation result for independent blocks.

PROPOSITION 6.2. *Let* $\alpha < \delta$, $\beta > 1/2$ *and* $0 < \rho < (1-\beta)/2$. *Then for any* $N \in \mathbb{N}$, *on some extension of the initial probability space there exists a sequence of independent standard normal r.v.'s* $(W'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ *and a version* $(X'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ *of the sequence* $(X_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ *such that*

$$\begin{aligned}(6.8)\qquad \mathbb{P}\Big((2^n)^{-1/2}\sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,i)\preceq(k,j)}(X'_{(l,i)}-\sigma_{l,i}W'_{(l,i)})\Big|\geq 2(2^n)^{-\rho}\Big)&\\ \leq C_2(2^n)^{-1-\alpha+(\varepsilon+\rho)(2+2\alpha)}\qquad &(\textit{dependence error})\\ +\,C_2(2^n)^{-\beta\alpha+\rho(2+2\alpha)}\qquad &(\textit{Sakhanenko's error})\end{aligned}$$

where $\sigma^2_{l,i}=\mathrm{Var}(X_{(l,i)})$ *and* $C_2=c_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\rho}(1+\lambda_0+\mu_\delta)^{2+2\delta}$.

Proof. Let $p=2+2\alpha$. Since $|I_{k,j}|\leq 2^{k-[\beta k]}$, using Proposition 8.2 we obtain

$$\mathbb{E}|X_{(k,j)}|^p\leq A|I_{k,j}|^{p/2}\leq A(2^{k-[\beta k]})^{p/2},$$

where $A=c_{\lambda_1,\lambda_2,\delta',\eta}(1+\lambda_0+\mu_\delta)^{p(1+\eta)}$ and $\eta>0$ is arbitrary. Taking into account that $m_k=2^{[\beta k]}\leq 2^{\beta k}$, we have

$$\begin{aligned}(6.9)\qquad \sum_{(k,j)\in\mathcal{K}_N}\mathbb{E}|X_{(k,j)}|^p&=\sum_{k=1}^{n}\sum_{j=1}^{m_k}\mathbb{E}|X_{(k,j)}|^p\leq\sum_{k=1}^{n}m_kA(2^{k-[\beta k]})^{p/2}\\ &\leq A2^{p/2}\sum_{k=1}^{n}2^{k(\beta+\frac{p}{2}(1-\beta))}\leq c_{\alpha,\beta}A(2^n)^{\beta+\frac{p}{2}(1-\beta)}.\end{aligned}$$

By (9.2) in the Appendix, on some probability space $(\Omega'',\mathcal{F}'',\mathbb{P}'')$, there exist a version $(Y''_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ of $(Y_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ and independent standard normal r.v.'s $(W'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ such that

$$\mathbb{P}''\Big(\sup_{(k,j)\in\mathcal{K}^0_N}\Big|\sum_{(l,i)\preceq(k,j)}(Y''_{(l,i)}-\sigma_{l,i}W'_{(l,i)})\Big|\geq a\Big)\leq\frac{c_p}{a^p}\sum_{(k,j)\in\mathcal{K}^0_N}\mathbb{E}|X_{(k,j)}|^p.$$

Choosing $a=(2^{\beta n})^{1/2-\rho}$ and taking into account (6.9) we obtain

$$\begin{aligned}\mathbb{P}''\Big((2^n)^{-1/2}\sup_{(k,j)\in\mathcal{K}^0_N}\Big|\sum_{(l,i)\preceq(k,j)}(Y''_{(l,i)}-\sigma_{l,i}W'_{(l,i)})\Big|\geq(2^n)^{-\rho}\Big)&\\ \leq c_p(2^n)^{-p/2+\rho p}c_{\alpha,\beta}A(2^n)^{\beta+\frac{p}{2}(1-\beta)}\leq c'_{\alpha,\beta}A(2^n)^{-\beta\alpha+\rho(2+2\alpha)}&.\end{aligned}$$

By Berkes–Philipp's lemma [1, Lemma 2.1] we can reconstruct the sequences $(X'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$, $(Y'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$, $(Y''_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ and $(W'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ on some new probability space in such a way that $Y'_{(k,j)} = Y''_{(k,j)}$ a.s. for any (k,j) in $\mathcal{K}_N$. Without loss of generality we shall consider this new probability space as an extension of the initial one. Using Proposition 6.1 we obtain

$$\begin{aligned}\mathbb{P}\Big(2^{-n/2} \sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,i)\preceq(k,j)} (X'_{(l,i)} - \sigma_{l,i}W'_{(l,i)})\Big| \geq 2(2^n\Big)^{-\rho}) \\ \leq C_1(2^n)^{-1-\alpha+(\varepsilon+\rho)(2+2\alpha)} \qquad &\text{(dependence error)} \\ + c'_{\delta,\beta}A(2^n)^{-\beta\alpha+\rho(2+2\alpha)} \qquad &\text{(Sakhanenko's error)}\end{aligned}$$

with C_1 defined by Proposition 6.1. Taking into account that $p = 2+2\alpha$, $\alpha < \delta$ and choosing η sufficiently small we get $p(1+\eta) \leq 2+2\delta$, which implies $c'_{\alpha,\beta}A \leq c'_{\lambda_1,\lambda_2,\alpha,\eta}(1+\lambda_0+\mu_\delta)^{2+2\delta}$. ∎

6.3. Construction of $(\widetilde{X}_i)_{1\leq i\leq N}$ and $(W_i)_{1\leq i\leq N}$. As before, we suppose that $\mu_i = 0$, $i \geq 1$ and $\sigma^2 = 1$. Let $(X'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ and $(W'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ be as in Proposition 6.2.

First we shall construct $(W_i)_{1\leq i\leq N}$. Note that, by Condition **C3**, the variances $\sigma^2_{k,i} = \mathrm{Var}(X_{(k,j)})$ can be approximated by $\sigma^2|I_{k,j}| = |I_{k,j}|$, but in general do not coincide with $|I_{k,j}|$. Therefore to perform our construction we have to replace each of the non-identically-distributed normal random variables $\sigma^2_{k,j}W'_{(k,j)}$ by some sums of independent identically distributed standard normal random variables. Let $(W_i)_{1\leq i\leq N}$ be a sequence of independent standard normal r.v.'s; let $\xi_{k,j}$ be an extra standard normal random variable. Set $I_{k,j} := \{i_1,\dots,i_{|I_{k,j}|}\}$ with $i_1 \preceq \cdots \preceq i_{|I_{k,j}|}$ and let $i^*_{k,j}$ be the maximal index $j \in \{i_1,\dots,i_{|I_{k,j}|}\}$ for which the variance of the partial sum $\sum_{i=i_1}^{j} W_i$ does not exceed $\sigma^2_{k,j}$, i.e. $i^*_{k,j} = i_{m^*_{k,j}}$, where $m^*_{k,j} = \min\{|I_{k,j}|, [\sigma^2_{k,j}]\}$.

It is easy to check that $W''_{(k,j)} := \sum_{i=i_1}^{i^*_{k,j}} W_i + \xi_{k,j}f_{k,j}$ where $f^2_{k,j} = |\sigma^2_{k,j} - i^*_{k,j}|$ is a normal random variable with mean 0 and variance $\sigma^2_{k,j}$; moreover, we may consider $W''_{(k,j)}$ as a new version of $\sigma_{k,j}W'_{(k,j)}$. The random variable $\sum_{k\in I_{k,j}}(W_k - W''_{(k,j)})$, which is equal to $\sum_{i^*_{k,j}+1\leq k\leq|I_{k,j}|}(W_k - \xi_{k,j}f_{k,j})$, also has a normal random variable of mean 0 and variance $(|I_{k,j}| - i^*_{k,j}) + f^2_{k,j} = \big|\sigma^2_{k,j} - |I_{k,j}|\big|$. By Berkes–Philipp's lemma, without loss of generality, we can reconstruct the sequences $(X'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$, $(\sigma_{k,j}W'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$, $(W''_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ and $(W''_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ on the same probability space in such way that a.s. $(\sigma_{k,j}W'_{(k,j)})_{(k,j)\in\mathcal{K}_N} = (W''_{(k,j)})_{(k,j)\in\mathcal{K}_N}$. We shall assume that this probability space is an extension of the initial on. Thus we have constructed the W_i's when i belongs to the union of all islands, $I = \bigcup_{(k,j)\in\mathcal{K}_N} I_{k,j}$, with the property that the $\eta_{k,j} = \sum_{i\in I_{k,j}} W_i - \sigma_{k,j} W'_{(k,j)}$ are independent normal and centered random variables with variances $v_{k,j}^2 = \left|\sigma_{k,j}^2 - |I_{k,j}|\right| \le \tau |I_{k,j}|^\gamma$ for any $\gamma > 0$. Therefore the sum $\sum_{(l,i)\le(n,m)} \eta_{l,i}$ is normal with mean 0 and variance $\sum_{(l,i)\le(n,m)} v_{k,i}^2 \le c_\beta \tau 2^{(\beta+\gamma)n}$; by Doob's inequality, with $p = 2+2\alpha$, it follows that

$$
\begin{aligned}
(6.10) \qquad \mathbb{P}\Big((2^n)^{-1/2} \sup_{(k,j)\in\mathcal{K}_N} \Big| \sum_{(l,i)\le(k,j)} \eta_{l,i}\Big| \ge (2^n)^{-\rho}\Big) & \\
&\le (2^n)^{-p/2+\rho p}\mathbb{E}\Big(\Big|\sum_{(l,i)\le(n,m)} \eta_{l,i}\Big|^p\Big) \\
&\le c_{\alpha,\beta}(2^n)^{-p/2+\rho p}\big(\tau 2^{(\beta+\gamma)n}\big)^{p/2} \\
&= c_{\alpha,\beta}\tau^{1+\alpha}\big(2^n\big)^{-(1-\beta)(1+\alpha)+(\rho+\gamma/2)(2+2\alpha)}
\end{aligned}
$$

where $\gamma > 0$ is arbitrary. When i belongs to the union of the gaps, $J = \bigcup_{(k,j)\in\mathcal{K}_N} J_{k,j}$, the variables W_i can be taken as any independent standard normal random variables independent of the sequence $(W_k)_{k\in I}$.

So far we have constructed the variables $(X'_{(k,j)})_{(k,j)\in\mathcal{K}_N}$ corresponding to sums over the islands. Now we proceed to construct the components of the sequence $(\widetilde{X}_i)_{1\le i\le N}$. First, we proceed with the components belonging to all islands. For each $(k,j) \in \mathcal{K}_N$, we construct a sequence $(\widetilde{X}_i)_{i\in I_{k,j}}$ such that $\sum_{i\in I_{k,j}} \widetilde{X}_i = X'_{(k,j)}$ and $(\widetilde{X}_i)_{i=1,\dots,N} \overset{d}{=} (X_i)_{i=1,\dots,N}$. Denote by $F_{X\,|\,Y_1,\dots,Y_k}(x\,|\,y_1,\dots,y_k)$ the conditional distribution of X given $[Y_1 = y_1,\dots,Y_k = y_k]$. Without loss of generality, on the initial probability space, there exists a sequence $U_1,\dots,U_N$ of independent r.v.'s uniformly distributed on $(0,1)$. Let $i_1,\dots,i_{|I_{k,j}|}$ be the indices in the set $I_{k,j}$. The required construction is performed in the standard way by defining first $\widetilde{X}_{i_1}$ as the conditional quantile transform

$$F^{-1}_{X_{i_1}|X_{(k,j)}}(U_{k,i_1}\,|\,X'_{(k,j)})$$

and then by setting, for $l = 2,\dots,|I_{k,j}|$,

$$\widetilde{X}_{i_l} = F^{-1}_{X_{i_l}|X_{i_1},\dots,X_{i_{l-1}},X_{(k,j)}}(U_{i_l}\,|\,\widetilde{X}_{i_1},\dots,\widetilde{X}_{i_{l-1}},X'_{(k,j)}).$$

Thus we have constructed the vector $\widetilde{X}_I = (\widetilde{X}_i)_{i\in I}$, where $I = \bigcup_{(k,j)\in\mathcal{K}_N} I_{k,j}$ is the union of all islands, such that $\widetilde{X}_I \overset{d}{=} X_I = (X_i)_{i\in I}$. In the same way we construct the $\widetilde{X}_i$ when i belongs to the union of the gaps, $J = \bigcup_{(k,j)\in\mathcal{K}_N} J_{k,j} := \{j_1,\dots,j_{|J|}\}$: set $X_{j_1} = F^{-1}_{X_{j_1}|X_I}(U_{k,j_1}\,|\,\widetilde{X}_I)$ and subsequently

$$\widetilde{X}_{j_l} = F^{-1}_{X_{j_l}|X_{j_1},\dots,X_{j_{l-1}},X_I}(U_{j_l}|\widetilde{X}_{j_1},\dots,\widetilde{X}_{j_{l-1}},\widetilde{X}_I) \quad \text{for } l = 2,\dots,|J|.$$

6.4. Putting together the bounds. Denote by $r_{k,j}$ the right end of the island $I_{k,j}$ and let $\mathcal{L}_N = \{1\} \cup \{r_{k,j} : (k,j) \in \mathcal{K}_N\}$ be the set of $r_{k,j}$'s equipped with the lexicographical order $\preceq$. For any $r = r_{k,j} \in \mathcal{L}_N$ let r^{next} be the next element in the set $\mathcal{L}_N$, i.e. $r^{\text{next}} = \inf\{r' : r' \in \mathcal{L}_N,\ r \preceq r'\}$.

Let $(\widetilde{X}_i)_{1\le i\le N}$ and $(W_i)_{1\le i\le N}$ be the sequences constructed in Section 6.3. Recall that by construction, for any $r = r_{k,j} \in \mathcal{L}_N$, we have $\{1,\dots,r\} = \sum_{(l,i)\preceq(k,j)} J_{l,j} \cup I_{k,j}$. First we replace, in the statement of Theorem 2.1, the sup over the set $\{j : 1 \le j \le N\}$ by the sup over the grid $\mathcal{L}_N$ and the sup of the oscillation term: in other words, the random variable $\sup_{1\le j\le N} |\sum_{i\le j}(\widetilde{X}_i - W_i)|$ is bounded by

$$\underbrace{\sup_{r\in\mathcal{L}_N}\Big|\sum_{i\le r}(\widetilde{X}_i - W_i)\Big|}_{\text{(sup over the grid } \mathcal{L}_N)} + \underbrace{\sup_{r\in\mathcal{L}_N}\ \sup_{r\le r'\le\min\{r^{\text{next}}-1,N\}}\Big|\sum_{r\le i\le r'}(\widetilde{X}_i - W_i)\Big|}_{\text{(oscillation term)}}. \tag{6.11}$$

For any $r = r_{k,j}$, we have

$$\begin{aligned}\sum_{1\le i\le r}\widetilde{X}_i &= \sum_{(l,h)\preceq(k,j)}\Big(\sum_{i\in I_{l,h}}\widetilde{X}_i + \sum_{i\in J_{l,h}}\widetilde{X}_i\Big)\\ &= \sum_{(l,h)\preceq(k,j)}(\underline{X}_{(l,h)} + \overline{X}_{(l,h)}),\end{aligned} \tag{6.12}$$

where $\underline{X}_{(l,h)} = \sum_{i\in I_{l,h}}\widetilde{X}_i$ and $\overline{X}_{(l,h)} = \sum_{i\in J_{l,h}}\widetilde{X}_i$. In the same way

$$\begin{aligned}\sum_{1\le i\le r}W_i &= \sum_{(l,h)\preceq(k,j)}\Big(\sum_{i\in I_{l,h}}W_i + \sum_{i\in J_{l,h}}W_i\Big)\\ &= \sum_{(l,h)\preceq(k,j)}(\underline{W}_{(l,h)} + \overline{W}_{(l,h)})\end{aligned} \tag{6.13}$$

where $\underline{W}_{(l,h)} = \sum_{i\in I_{l,h}}W_i$ and $\overline{W}_{(l,h)} = \sum_{i\in J_{l,h}}W_i$. From (6.11)–(6.13) we obtain

$$\begin{aligned}&\sup_{1\le j\le N}\Big|\sum_{i\le j}(\widetilde{X}_i - W_i)\Big|\\ &\le \sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,h)\le(k,j)}(\underline{X}_{(l,h)} - \underline{W}_{(l,h)})\Big| &&\text{(sup over islands)}\\ &\quad+ \sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,h)\le(k,j)}(\overline{X}_{(l,h)} - \overline{W}_{(l,h)})\Big| &&\text{(sup over gaps)}\\ &\quad+ \sup_{r\in\mathcal{L}_N}\ \sup_{r\le r'\le\min\{r^{\text{next}}-1,N\}}\Big|\sum_{l\le i\le r'}(\widetilde{X}_i - W_i)\Big| &&\text{(oscillation term)}\end{aligned} \tag{6.14}$$

where the term "sup over islands" is bounded by the sum

$$\sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,h)\le(k,j)}(\underline{X}_{(l,h)}-\sigma_{l,h}W'_{(l,h)})\Big| \quad \text{(normal approximation)} \tag{6.15}$$

$$+\sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,h)\le(k,j)}(\sigma_{l,h}W'_{(l,h)}-\underline{W}_{(l,h)})\Big| \quad \text{(variance homogenization).}$$

The term "normal approximation" has already been controlled in Proposition 6.2 where it is bounded by two terms, "dependence error" and "Sakhanenko's error". The "variance homogenization" term is controlled by (6.10). As to "sup over gaps" and "oscillation term", they will be considered in (6.17) and (6.18) below.

6.5. Bound for the partial sums over gaps. Let $p=2+2\alpha$, where $\alpha<\delta$. Since the blocks are indexed by $l=k_0,\dots,n$ and the total length of the gaps in block l is less than $(2+[\beta l])2^{[\beta l]+[\varepsilon l]-1}$, the total length L^{gap} of all gaps satisfies

$$L^{\mathrm{gap}}=\sum_{(l,i)\preceq(n,m)}|J_{l,i}|\le\sum_{k_0\le l\le n}(2+[\beta l])2^{[\beta l]+[\varepsilon l]-1}\le c_{\varepsilon,\beta}2^{(\beta+\varepsilon)k}.$$

By Proposition 8.1, we have, for any $\eta>0$,

$$\begin{aligned}\Big\|\sup_{(k,j)\preceq(n,m)}\Big|\sum_{(l,i)\le(k,j)}\overline{X}_{(l,i)}\Big|\Big\|_{L^p}&\\ &\le c_{\lambda_1,\lambda_2,\alpha,\delta,\eta}(1+\lambda_0+\mu_\delta)^{1+\eta}(L^{\mathrm{gap}})^{1/2}\\ &\le c_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\delta,\eta}(1+\lambda_0+\mu_\delta)^{1+\eta}(2^{(\beta+\varepsilon)n})^{1/2}.\end{aligned}$$

Using Chebyshev's inequality with $x=(2^n)^{1/2-\rho}$, we get

$$\begin{aligned}\mathbb{P}\Big(\sup_{(k,j)\in\mathcal{K}_N}\Big|\sum_{(l,i)\le(k,j)}\overline{X}_{(l,i)}\Big|\ge x\Big)&\\ &\le\frac{1}{x^p}\mathbb{E}\Big(\sup_{(k,j)\preceq(n,m)}\Big|\sum_{(l,i)\le(k,j)}\overline{X}_{(l,i)}\Big|\Big)^p\\ &\le A(2^n)^{-p/2+\rho p}(2^{(\beta+\varepsilon)n})^{p/2}\\ &\le A(2^n)^{-(1-\beta-\varepsilon)p/2+\rho p},\end{aligned} \tag{6.16}$$

where $A=c_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\delta,\eta}(1+\lambda_0+\mu_\delta)^{p(1+\eta)}$. A similar bound can be established with $\overline{W}_{(l,i)}$ instead of $\overline{X}_{(l,i)}$. From this bound and (6.16) it follows that, for any $0 < \rho < 1/4$,

$$\begin{aligned}(6.17)\quad \mathbb{P}\Big((2^n)^{-1/2} \sup_{(k,j)\in\mathcal{K}_N} \Big| \sum_{(l,i)\le(k,j)} (\overline{X}_{(l,i)} - \overline{W}_{(l,i)})\Big| \ge (2^n)^{-\rho}\Big) &\le A'(2^n)^{-(1-\beta-\varepsilon)p/2+\rho p} \\ &\le A'(2^n)^{-(1-\beta)(1+\alpha)+(\rho+\varepsilon/2)(2+2\alpha)}\end{aligned}$$

where $A' = c'_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\delta}(1+\lambda_0+\mu_\delta)^{2+2\delta}$.

6.6. Bound for the oscillation term. Denote $r^+ = \min\{r^{\text{next}} - 1, N\}$ for brevity. First note that

$$r^+ - r \le \max_{(k,j)\in\mathcal{K}_N} (|I_{k,j}| + |J_{k,j}|) \le c_{\varepsilon,\beta}(2^{(\beta+\varepsilon)n} + 2^{(1-\beta)n}).$$

Let $p = 2 + 2\alpha$ where $\alpha < \delta$. By Proposition 8.1, for any $\eta > 0$,

$$\begin{aligned}\Big\| \sup_{r\in\mathcal{L}_N} \sup_{r\le l\le r^+} \Big| \sum_{r\le i\le l} \widetilde{X}_i \Big| \Big\|_{L^p}^p &\le A\Big(\sup_{r\in\mathcal{L}_N}(r^+ - r)\Big)^{p/2} \\ &\le c_{\varepsilon,\beta} A(2^{(\beta+\varepsilon)n} + 2^{(1-\beta)n})^{p/2},\end{aligned}$$

where $A = c_{\lambda_1,\lambda_2,\alpha,\delta,\eta}(1+\lambda_0+\mu_\delta)^{p(1+\eta)}$. Therefore, by Chebyshev's inequality, with $x = \frac{1}{2}(2^n)^{1/2-\rho}$ and $\rho > 0$,

$$\begin{aligned}&\mathbb{P}\Big(\sup_{r\in\mathcal{L}_N} \sup_{r\le l\le r^+} \Big| \sum_{r\le i\le l} \widetilde{X}_i \Big| \ge x\Big) \\ &\le x^{-p}\mathbb{E}\Big(\sup_{r\in\mathcal{L}_N} \sup_{r\le l\le r^+} \Big| \sum_{r\le i\le l} \widetilde{X}_i \Big|\Big)^p \le c_{\varepsilon,\beta} A x^{-p}(2^{(\beta+\varepsilon)n} + 2^{(1-\beta)n})^{p/2} \\ &\le c'_{\varepsilon,\beta} A 2^{p+np\rho}((2^n)^{-(1+\alpha)(1-\beta-\varepsilon)} + (2^n)^{-(1+\alpha)\beta}).\end{aligned}$$

Choosing η small enough we have $p(1+\eta) \le 2 + 2\delta$ and therefore

$$c_{\varepsilon,\beta} 2^p A \le A' = c'_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha,\delta}(1+\lambda_0+\mu_\delta)^{2+2\delta}.$$

Since a similar bound can be established with W_i instead of X_i, we obtain the following bound for the oscillation term:

$$\begin{aligned}(6.18)\quad &\mathbb{P}\Big((2^n)^{-1/2} \sup_{r\in\mathcal{L}_N} \sup_{r\le l\le r^+} \Big| \sum_{r\le i\le l} (\widetilde{X}_i - W_i)\Big| \ge 2(2^n)^{-\rho}\Big) \\ &\le 2A'(2^n)^{(2+2\alpha)(\rho+\varepsilon/2)}((2^n)^{-(1+\alpha)(1-\beta)} + (2^n)^{-(1+\alpha)\beta}).\end{aligned}$$

6.7. Optimizing the bounds. Let $\alpha < \delta$, $\beta > 1/2$ and $0 < \rho < (1-\beta)/2$. Using (6.14), we may decompose the quantity

$$\mathbb{P}\Big((2^n)^{-1/2} \sup_{1\le j\le N} \Big| \sum_{i=1}^{j} (\widetilde{X}_i - W_i)\Big| \ge 6(2^n)^{-\rho}\Big)$$

into three terms, the first one "sup over islands" being itself decomposed in two terms (see (6.15)); consequently, this quantity is decomposed into four terms listed below:

- the first term "normal approximation" is controlled with Proposition 6.2, it is bounded by two terms named "dependence error" and "Sakhanenko's error",
- the second term "variance homogenization" is controlled in (6.10) with $\gamma = \varepsilon$,
- the term "sup over gaps" is controlled in (6.17),
- the term "oscillation term" is controlled in (6.18).

Putting these bounds together, we obtain

$$\begin{aligned}
&\mathbb{P}\Big((2^n)^{-1/2} \sup_{1\le j\le N}\Big|\sum_{i=1}^{j}(X_i - W_i)\Big| \ge 6(2^n)^{-\rho}\Big) \\
&\le A(2^n)^{-(1+\alpha)+(\rho+\varepsilon)(2+2\alpha)} && \text{(dependence error)}\\
&\quad + A(2^n)^{-\beta\alpha+\rho(2+2\alpha)} && \text{(Sakhanenko's error)}\\
&\quad + A(2^n)^{-(1-\beta)(1+\alpha)+(\rho+\varepsilon/2)(2+2\alpha)} && \text{(variance homogenization error)}\\
&\quad + A(2^n)^{-(1-\beta)(1+\alpha)+(\rho+\varepsilon/2)(2+2\alpha)} && \text{(gaps error)}\\
&\quad + A((2^n)^{-(1-\beta)(1+\alpha)} + (2^n)^{-\beta(1+\alpha)})(2^n)^{(\rho+\varepsilon/2)(2+2\alpha)} && \text{(oscillation error)}
\end{aligned}$$

where $A = c_{\varepsilon,\beta,\lambda_1,\lambda_2,\alpha}(1 + \tau^{1+\alpha} + (1 + \lambda_0 + \mu_\delta)^{2+2\alpha})$.

For the moment let us ignore the factors containing ε, which have a small contribution to the bound. The term "dependence error" is negligible with respect to all other terms; equating the powers of the term "Sakhanenko's error" and the term "gaps error" (or equivalently "variance homogenization error") we get $\beta\alpha = (1-\beta)(1+\alpha)$ i.e. $\beta = \frac{1+\alpha}{1+2\alpha}$. Implementing $\beta = \frac{1+\alpha}{1+2\alpha}$ in the above inequality yields

$$\begin{aligned}
&\mathbb{P}\Big((2^n)^{-1/2} \sup_{1\le l\le N}\Big|\sum_{i=1}^{l}(\widetilde{X}_i - W_i)\Big| \ge 6(2^n)^{-\rho}\Big) \\
&\qquad \le A(2^n)^{-1-\delta+(\rho+\varepsilon)(2+2\alpha)} \qquad \text{(dependence error)}\\
&\qquad\quad + 6A(2^n)^{-\frac{\alpha(1+\alpha)}{1+2\alpha}+(\rho+\varepsilon)(2+2\alpha)} \quad \text{(Sakhanenko's error}\\
&\qquad\qquad\qquad \text{+ variance homogenization error}\\
&\qquad\qquad\qquad \text{+ gaps error + oscillation error).}
\end{aligned}$$

Taking into account that $\alpha < \delta$ and $2^n \le N < 2^{n+1}$ we obtain

$$\mathbb{P}\Big(N^{-1/2} \sup_{1\le l\le N}\Big|\sum_{i=1}^{l}(\widetilde{X}_i - W_i)\Big| \ge 6N^{-\rho}\Big) \le A' N^{-\frac{\alpha(1+\alpha)}{1+2\alpha}+(\rho+\varepsilon)(2+2\alpha)} \tag{6.19}$$

where $A' = c_{\varepsilon,\lambda_1,\lambda_2,\alpha,\delta}(1 + \lambda_0 + \mu_\delta + \sqrt{\tau})^{2+2\delta}$ and ρ satisfies $0 < \rho < \frac{1-\beta}{2} = \frac{\alpha}{2(1+2\alpha)}$.

Note that the function $g(\alpha) = \frac{\alpha(1+\alpha)}{1+2\alpha} - \rho(2 + 2\alpha)$ is strictly increasing on $\mathbb{R}^+$ when $\rho < 1/4$. Therefore we can get rid of the constant ε in the bound by choosing $\alpha' < \alpha$. If we let $\Delta = g(\alpha) - g(\alpha') > 0$ and choose ε sufficiently small, we obtain $(2^n)^{-g(\alpha)+\varepsilon(2+2\alpha)} = (2^n)^{-g(\alpha')-\Delta+\varepsilon(2+2\alpha)} \leq (2^n)^{-g(\alpha')}$. Since α and α' are arbitrary satisfying $\alpha' < \alpha < \delta$, the assertion of Theorem 2.1 follows with α' replacing α.

We have performed the construction of the sequences $\widetilde{X}^{(N)} = (\widetilde{X})_{1\leq i\leq N}$ and $W^{(N)} = (W)_{1\leq i\leq N}$ for each fixed $N \geq 1$, where for each N the constructed sequences are in general different. Below we show how to obtain a construction of the entire sequences $(\widetilde{X})_{i\geq 1}$ and $(W)_{i\geq 1}$.

Let $\Omega^{(N)} = \mathbb{R}^{N+1} \times \mathbb{R}^{N+1}$. Without loss of generality, for any $\omega = (\omega_1, \omega_2) \in \Omega^{(N)}$, the sequences $\widetilde{X}^{(N)}$ and $W^{(N)}$ can be reconstructed on $\Omega^{(N)}$ so that $\widetilde{X}_i = \omega_{1,i}$, $W_i = \omega_{2,i}$ and their joint distribution, say $\mathbb{P}^{(N)}$, is preserved. Each measure $\mathbb{P}^{(N)}$ can be extended (arbitrarily) onto the space $\mathbb{R}^\infty \times \mathbb{R}^\infty$. From the bound (6.19) it follows that the sequence of measures $\mathbb{P}^{(N)}$ is tight. Therefore there is a weak limit which satisfies (6.19) and thus provides the desired construction.

7. Proof of the results of Section 3 .

Throughout this section we assume that the Markov chain $(X_n)_{n\geq 0}$ and the function f satisfy Hypotheses **M1**–**M4**.

7.1. Proof of Proposition 3.1.

First, we establish the following bound for the characteristic functions $\phi_{x,1}$, $\phi_{x,2}$ and ϕ_x involved in Proposition 3.1.

LEMMA 7.1. *For any $k_{\text{gap}}, M_1, M_2 \in \mathbb{N}$, any sequence $j_0 < \cdots < j_{M_1+M_2}$ and any $t = (t_i)_i \in \mathbb{R}^{M_1}$, $s = (s_i)_i \in \mathbb{R}^{M_2}$ satisfying $\|(t, s)\|_\infty \leq \varepsilon_0$,*

$$|\phi_x(t, s) - \phi_{x,1}(t)\phi_{x,2}(s)| \leq 2C_Q C_{\mathbf{P}}^{M_1+M_2}(\|\nu\|_{\mathcal{B}'} + \|\delta_x\|_{\mathcal{B}'})\|\mathbf{1}\|_{\mathcal{B}}\kappa^{k_{\text{gap}}}.$$

Set for brevity $\phi_1 = \phi_{x,1}$, $\phi_2 = \phi_{x,2}$ and $\phi = \phi_x$. The characteristic function ϕ can be rewritten in the following form:

$$\phi(t, s) = (\mathbf{P}^{j_0}\mathbf{P}_{t_1}^{|J_1|}\ldots\mathbf{P}_{t_{M_1}}^{|J_{M_1}|}\mathbf{P}^{k_{\text{gap}}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\ldots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)(x).$$

Since $\mathbf{P} = \Pi + Q$ we get $\mathbf{P}^k = \Pi + Q^k$, and thus

$$\phi(t, s) = \phi_\Pi(t, s) + \phi_Q(t, s) \tag{7.1}$$

with

$$\phi_\Pi(t, s) := (\mathbf{P}^{j_0}\mathbf{P}_{t_1}^{|J_1|}\ldots\mathbf{P}_{t_{M_1}}^{|J_{M_1}|}\Pi^{k_{\text{gap}}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\ldots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)(x), \tag{7.2}$$

$$\phi_Q(t,s) := (\mathbf{P}^{j_0}\mathbf{P}_{t_1}^{|J_1|}\dots\mathbf{P}_{t_{M_1}}^{|J_{M_1}|}Q^{k_{\rm gap}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)(x). \tag{7.3}$$

First, since $\Pi^{k_{\rm gap}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e = \nu(\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)e$, we may write, setting $\psi_2(s) := \nu(\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)$,

$$\phi_\Pi(t,s) = \psi_2(s)(\mathbf{P}^{j_0}\mathbf{P}_{t_1}^{|J_1|}\dots\mathbf{P}_{t_{M_1}}^{|J_{M_1}|}e)(x) = \psi_2(s)\phi_1(t).$$

Notice that $\phi_2(s) = (\mathbf{P}^{k_{\rm gap}+j_{M_1}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)(x)$; using the equality $\nu\mathbf{P} = \nu$, one gets $\psi_2(s) = \nu(\mathbf{P}^{k_{\rm gap}+j_{M_1}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)$, which allows us to control the difference between ψ_2 and ϕ_2, namely

$$\begin{aligned}\psi_2(s)-\phi_2(s) &= (\nu-\delta_x)(\mathbf{P}^{k_{\rm gap}+j_{M_1}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)\\ &= (\nu-\delta_x)(\Pi\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)\\ &\quad + (\nu-\delta_x)(Q^{k_{\rm gap}}\mathbf{P}^{j_{M_1}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)\\ &= (\nu-\delta_x)(e)\nu(\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)\\ &\quad + (\nu-\delta_x)(Q^{k_{\rm gap}}\mathbf{P}^{j_{M_1}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)\end{aligned}$$

with $(\nu-\delta_x)(e) = 0$; consequently,

$$\begin{aligned}|\psi_2(s)-\phi_2(s)| &= |(\nu-\delta_x)(Q^{k_{\rm gap}+j_{M_1}}\mathbf{P}_{s_{M_1+1}}^{|J_{M_1+1}|}\dots\mathbf{P}_{s_{M_1+M_2}}^{|J_{M_1+M_2}|}e)|\\ &\le C_Q C_{\mathbf{P}}^{1+M_2}\kappa^{k_{\rm gap}}(\|\nu\|_{\mathcal{B}'}+\|\delta_x\|_{\mathcal{B}'})\|e\|_{\mathcal{B}}.\end{aligned}$$

On the other hand, one easily gets

$$|\phi_Q(t,s)| \le C_Q C_{\mathbf{P}}^{1+M_1+M_2}\kappa^{k_{\rm gap}}\|e\|_{\mathcal{B}}\|\delta_x\|_{\mathcal{B}'}. \tag{7.4}$$

Writing $\phi(t,s) = \phi_1(t)\phi_2(s) + \phi_1(t)(\psi_2(s)-\phi_2(s)) + \phi_Q(t,s)$ and using the previous inequalities, one finally gets

$$|\phi(t,s)-\phi_1(t)\phi_2(s)| \le 2C_Q C_{\mathbf{P}}^{1+M_1+M_2}(\|\nu\|_{\mathcal{B}'}+\|\delta_x\|_{\mathcal{B}'})\kappa^{k_{\rm gap}}\|e\|_{\mathcal{B}}. \ ■$$

To prove Proposition 3.1, set $k_0 = \max\{1, \log_2 C_{\mathbf{P}}\}$ so that $C_{\mathbf{P}} \le 2^{k_0}$. Since $\max_{m=1,\dots,M_1+M_2}\operatorname{card}(J_m) \ge 1$, one gets

$$C_{\mathbf{P}}^{M_1+M_2} \le 2^{k_0(M_1+M_2)} \le \Big(1 + \max_{m=1,\dots,M_1+M_2}\operatorname{card}(J_m)\Big)^{k_0(M_1+M_2)}.$$

Now, Proposition 3.1 follows from Lemma 7.1.

7.2. Proof of Proposition 3.2. We need two auxiliary lemmas.

LEMMA 7.2. *For any* $l, k = 0, 1, \dots,$

$$|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k}))| \le A(x)\kappa^{k\gamma/4}, \tag{7.5}$$

for any positive constant γ *satisfying* $0 < \gamma \le \min\{1, 2\delta\}$, *where*

$$A(x) = c_\delta\big(1 + C_Q C_{\mathbf{P}}^2(\|\nu\|_{\mathcal{B}'}+\|\delta_x\|_{\mathcal{B}'})\|e\|_{\mathcal{B}} + \mu_\delta^{2+\gamma}(x)\big).$$

Proof. We give a proof involving Lemma 7.1. Let V and V' be two independent identically distributed r.v.'s of mean 0, independent of X_l and X_{l+m} and whose common characteristic function is supported in the interval $[-\varepsilon_0, \varepsilon_0]$ for some $\varepsilon_0 > 0$. Set $Y_l = f(X_l) + V$ and $Y'_{l+k} = f(X_{l+k}) + V'$.

Let $\widetilde{\phi}_1$ (resp. $\widetilde{\phi}_2$, $\widetilde{\phi}(t,u)$) be the characteristic function of Y_l (resp. Y'_{l+k}, (Y_l, Y'_{l+k})). Set $g_T(x) = x1_{(|x|\le T)}$ and $h_T(x,y) = g_T(x)g_T(y)$ for $x, y \in \mathbb{R}$. Let $\widehat{g}_T$ (resp. $\widehat{h}_T$) be the Fourier transform of the function g_T (resp. h_T) defined by

$$\widehat{g}_T(t) = \int e^{itx} g_T(x)\,dx,$$
$$\widehat{h}_T(t,u) = \iint e^{i(tx+uy)} h_T(x,y)\,dx\,dy = \widehat{g}_T(t)\widehat{g}_T(u).$$

For any $T > 0$ and $l \ge 1$, $k \ge 0$, one gets

$$\mathbb{E}_x f(X_l) f(X_{l+k}) = \mathbb{E}_x Y_l Y'_{l+k} = \mathbb{E}_x h_T(Y_l, Y'_{l+k}) + R_0 \tag{7.6}$$

with

$$|R_0| \le \mathbb{E}_x |Y_l Y'_{l+k}| 1_{(|Y_l| > T)} + \mathbb{E}_x |Y_l Y'_{l+k}| 1_{(|Y'_{l+k}| > T)}. \tag{7.7}$$

By the inverse Fourier transform, one may write

$$\mathbb{E}_x f(X_l) f(X_{l+k}) = \frac{1}{(2\pi)^2} \iint \overline{\widehat{h}_T(t,u)}\, \widetilde{\phi}(t,u)\,dt\,du + R_0.$$

Analogously

$$\mathbb{E}_x f(X_l) = \mathbb{E}_x Y_l = \mathbb{E}_x g_T(Y_l) + R_1 = \frac{1}{2\pi} \int \overline{\widehat{g}_T(t)}\, \widetilde{\phi}_1(t)\,dt + R_1 \tag{7.8}$$

and

$$\mathbb{E}_x f(X_{l+k}) = \mathbb{E}_x Y'_{l+k} = \mathbb{E}_x g_T(Y'_{l+k}) + R_2 = \frac{1}{2\pi} \int \overline{\widehat{g}_T(u)}\, \widetilde{\phi}_2(u)\,du + R_2,$$

where

$$R_1 := \mathbb{E}_x Y_l 1_{(|Y_l| > T)} \quad \text{and} \quad R_2 := \mathbb{E}_x Y'_{l+k} 1_{(|Y'_{l+k}| > T)}. \tag{7.9}$$

This gives

$$\begin{aligned} \operatorname{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) &= \mathbb{E}_x f(X_l) f(X_{l+k}) - \mathbb{E}_x f(X_l) \mathbb{E}_x f(X_{l+k}) \\ &= \frac{1}{(2\pi)^2} \iint \overline{\widehat{h}_T(t,u)} (\widetilde{\phi}(t,u) - \widetilde{\phi}_1(t)\widetilde{\phi}_2(u))\,dt\,du + R, \end{aligned} \tag{7.10}$$

where

$$R = R_0 + R_1 \mathbb{E}_x g_T(Y'_{l+k}) + R_2 \mathbb{E}_x g_T(Y_l) + R_1 R_2. \tag{7.11}$$

Note that

$$\left| \iint \overline{\widehat{h}_T(t,u)} (\widetilde{\phi}(t,u) - \widetilde{\phi}_1(t)\widetilde{\phi}_2(u))\,dt\,du \right| \le \|\widehat{h}_T\|_{L^2} \|\widetilde{\phi} - \widetilde{\phi}_1 \widetilde{\phi}_2\|_{L^2}.$$

Since V, V' are independent of X_l, X_{l+k}, we have

$$\widetilde{\phi}(t,u) = \phi(t,u)\mathbb{E}_x e^{itV}\mathbb{E}_x e^{iuV'}$$

and

$$\widetilde{\phi}_1(t) = \phi_1(t)\mathbb{E}_x e^{itV}, \quad \widetilde{\phi}_2(u) = \phi_2(u)\mathbb{E}_x e^{iuV'},$$

where

$$\begin{aligned}\phi(t,u) &:= \mathbb{E}_x e^{itf(X_l)+iuf(X_{l+k})} = (\mathbf{P}^{l-1}\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e)(x),\\ \phi_1(t) &:= \mathbb{E}_x e^{itf(X_l)} = (\mathbf{P}^{l-1}\mathbf{P}_t e)(x),\\ \phi_2(u) &:= \mathbb{E}_x e^{iuf(X_{l+k})} = (\mathbf{P}^{l-1}\mathbf{P}_t e)(x).\end{aligned}$$

Since the support of the characteristic functions of V and V' is the interval $[-\varepsilon_0, \varepsilon_0]$ the function $\widetilde{\phi} - \widetilde{\phi}_1\widetilde{\phi}_2$ vanishes outside the square $[-\varepsilon_0, \varepsilon_0]^2$. Then, by Lemma 7.1,

$$\begin{aligned}\|\widetilde{\phi} - \widetilde{\phi}_1\widetilde{\phi}_2\|_{L^2} &\le 2\varepsilon_0 \sup_{|t|\le\varepsilon_0,\,|u|\le\varepsilon_0} |\phi(t,u) - \phi_1(t)\phi_2(u)| \\ &\le 4\epsilon_0 C_Q C_{\mathbf{P}}^3 \kappa^k (\|\nu\|_{\mathcal{B}'} + \|\delta_x\|_{\mathcal{B}'})\|e\|_{\mathcal{B}}.\end{aligned} \tag{7.12}$$

Using the inequality

$$\|\widehat{h}_T\|_{L^2}^2 = \iint h_T^2(x,y)\,dx\,dy = \Big(\int g_T^2(x)\,dx\Big)^2 \le \frac{4}{9}T^6, \tag{7.13}$$

one obtains

$$\begin{aligned}&|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k}))| \\ &\qquad \le \frac{2}{3\pi^2}T^3\epsilon_0 C_Q C_{\mathbf{P}}^3\kappa^k\Big(\|\nu\|_{\mathcal{B}'} + \|\delta_x\|_{\mathcal{B}'}\Big)\|e\|_{\mathcal{B}} + |R|.\end{aligned} \tag{7.14}$$

Now we shall give a bound for $|R|$. By Hölder's inequality, with $q_\delta = \frac{1+\delta}{\delta} > 1$,

$$\mathbb{E}_x|Y_l|\,|Y'_{l+k}|1_{(|Y_l|>T)} \le (\mathbb{E}_x|Y_l|^{2+2\delta})^{\frac{1}{2+2\delta}}(\mathbb{E}_x|Y'_{l+k}|^{2+2\delta})^{\frac{1}{2+2\delta}}\mathbb{P}_x(|Y_l|>T)^{\frac{1}{q_\delta}}.$$

Using Hypothesis **M4**, we have

$$(\mathbb{E}_x|Y_l|^{2+2\delta})^{\frac{1}{2+2\delta}} \le (\mathbb{E}_x|f(X_l)|^{2+2\delta})^{\frac{1}{2+2\delta}} + (\mathbb{E}_x|V|^{2+2\delta})^{\frac{1}{2+2\delta}} \le c_\delta A_0(x),$$

with $A_0(x) = \mu_\delta(x) + 1$. Similarly $(\mathbb{E}_x|Y'_{l+k}|^{2+2\delta})^{\frac{1}{2+2\delta}} \le c_\delta A_0(x)$. On the other hand, for any $\gamma \in (0, 2\delta]$, one gets

$$\mathbb{P}_x(|Y_l|>T) \le \frac{1}{T^{\gamma q_\delta}}\mathbb{E}_x|Y_l|^{\gamma q_\delta} \le \frac{c_\delta}{T^{\gamma q_\delta}}A_0^{\gamma q_\delta}(x).$$

Putting together these bounds gives

$$\mathbb{E}_x|Y_l|\,|Y'_{l+k}|1_{(|Y_l|>T)} \le c_\delta T^{-\gamma}A_0^{2+\gamma}(x). \tag{7.15}$$

In the same way we obtain, for any $\gamma \in (0, 2\delta]$,

$$\mathbb{E}_x|Y_l|\,|Y'_{l+k}|1(|Y'_{l+k}|>T) \le c_\delta T^{-\gamma}A_0^{2+\gamma}(x). \tag{7.16}$$

From (7.7), (7.15), (7.16), it follows that

$$|R_0| \leq c_\delta T^{-\gamma} A_0^{2+\gamma}(x). \tag{7.17}$$

From (7.16), taking $k = 0$ we get, for any $\gamma \in (0, 2\delta]$,

$$\max\{R_1, R_2\} \leq \sup_{l\geq 0}(\mathbb{E}_x Y_l^2 1(|Y_l| > T))^{1/2} \leq c_\delta T^{-\gamma/2} A_0^{1+\gamma/2}(x). \tag{7.18}$$

Since

$$|\mathbb{E}_x g_T(Y_l)| \leq (\mathbb{E}_x(|Y_l|^{2+2\delta}))^{\frac{1}{2+2\delta}} \leq c_\delta A_0(x)$$

and

$$|\mathbb{E}_x g_T(Y'_{l+k})| \leq c_\delta A_0(x),$$

from (7.17), (7.18) it follows that

$$|R| \leq c_\delta T^{-\gamma/2} A_0^{2+\gamma}(x) \tag{7.19}$$

for any $\gamma \in (0, 2\delta]$, where we assume without loss of generality that $A_0(x) \geq 1$. The inequalities (7.14) and (7.19) yield, for any $\gamma \in (0, 2\delta]$,

$$\begin{aligned}|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k}))| \leq \frac{2}{3\pi^2} T^3 \varepsilon_0 C_Q C_{\mathbf{P}}^3 \kappa^k (\|\nu\|_{\mathcal{B}'} + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'})\|e\|_{\mathcal{B}} \\ + c_\delta T^{-\gamma/2} A_0^{2+\gamma}(x).\end{aligned}$$

Choosing $T = \kappa^{-k/4}$ and taking into account $A_0^{2+\gamma}(x) \leq c_\delta(1 + \mu_\delta^{2+\gamma}(x))$, it follows that

$$|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k}))| \leq A(x) c_\delta \kappa^{k \min\{1,\gamma/2\}/4},$$

which finishes the proof of Lemma 7.2. ∎

LEMMA 7.3. *Let $0 < \gamma \leq \min\{1, 2\delta\}$. Then:*

(a) *There exists a real number μ not depending on x such that, for any $k \geq 1$,*

$$|\mathbb{E}_x f(X_k) - \mu| \leq c_\delta A_1(x) \kappa^{k\gamma/4-1}$$

where $A_1(x) = 1 + \mu_\delta(x)^{1+\gamma} + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}} C_{\mathbf{P}} C_Q$. Moreover

$$\sum_{k=0}^{\infty} |\mathbb{E}_x f(X_k) - \mu| \leq \overline{\mu}(x) = c_{\gamma,\kappa,\delta} A_1(x).$$

(b) *There exists a sequence of (possibly complex) numbers $(s_k)_{k\geq 0}$ not depending on x such that*

$$|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) - s_k| \leq c_\delta A_2(x) \kappa^{l\gamma/4-1} \tag{7.20}$$

where

$$\begin{aligned}A_2(x) &= 1 + \mu_\delta(x)^{2+\gamma} \\ &\quad + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}\big(C_{\mathbf{P}}^2 C_Q(\|\nu\|_{\mathcal{B}'}\|e\|_{\mathcal{B}} + C_Q) + C_{\mathbf{P}} C_Q(1 + \|\nu\|_{\mathcal{B}'} C_{\mathbf{P}})\big).\end{aligned}$$

Moreover, for $k \geq 0$,

$$|s_k| \leq A_2(x)\kappa^{k\gamma/4-1}$$

and

$$|s_0| + 2\sum_{k=1}^{\infty} |s_k| \leq c_{\gamma,\kappa,\delta} A_2(x).$$

Proof. To avoid repetitions we first prove (b). We keep the notations from the proof of Lemma 7.2. Denote $\widetilde{\phi}_0(t,u) = \widetilde{\phi}(t,u) - \widetilde{\phi}_1(t)\widetilde{\phi}_2(u)$. By (7.10), for any $l = 0, 1, \ldots,$

$$\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) = \frac{1}{(2\pi)^2} \iint \overline{\widehat{h}_T(t,u)}\, \widetilde{\phi}_0(t,u)\, dt\, du + R,$$

with R defined by (7.11). Since V, V' are independent of X_l, X_{l+k},

$$\begin{aligned} \widetilde{\phi}(t,u) &= (\mathbb{E}_x e^{i0\sum_{j=1}^{l-1} X_j + itX_l + i0\sum_{j=l+1}^{l+k-1} X_j + iuX_{l+k}}) \mathbb{E}_x e^{itV} \mathbb{E}_x e^{iuV'} \\ &= (\mathbf{P}^{l-1}\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e)(x) \mathbb{E}_x e^{itV} \mathbb{E}_x e^{iuV'}. \end{aligned} \tag{7.21}$$

Note that, for $k, l \geq 2$,

$$\begin{aligned} (\mathbf{P}^{l-1}\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e)(x) &= \boldsymbol{\delta}_x(\mathbf{P}^{l-1}\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e) \\ &= \boldsymbol{\delta}_x(\Pi\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e) + \boldsymbol{\delta}_x(Q^{l-1}\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e) \\ &= \nu(\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e) \\ &\quad + \boldsymbol{\delta}_x(Q^{l-1}\mathbf{P}_t\Pi\mathbf{P}_u e) + \nu(Q^{l-1}\mathbf{P}_t Q^{k-1}\mathbf{P}_u e). \end{aligned}$$

Since

$$|\boldsymbol{\delta}_x(Q^{l-1}\mathbf{P}_t\Pi\mathbf{P}_u e)| = |\boldsymbol{\delta}_x(Q^{l-1}\mathbf{P}_t e)\nu(\mathbf{P}_u e)| \leq \kappa^{l-1} C_Q C_{\mathbf{P}}^2 \|\boldsymbol{\delta}_x\|_{\mathcal{B}'} \|\nu\|_{\mathcal{B}'} \|e\|_{\mathcal{B}}^2$$

and

$$|\boldsymbol{\delta}_x(Q^{l-1}\mathbf{P}_t Q^{k-1}\mathbf{P}_u e)| \leq \kappa^{l+k-2} C_Q^2 C_{\mathbf{P}}^2 \|\boldsymbol{\delta}_x\|_{\mathcal{B}'} \|e\|_{\mathcal{B}'},$$

we obtain

$$|\widetilde{\phi}(t,u) - \widetilde{\psi}(t,u;k)| \leq \kappa^{l-1} C_{\mathbf{P}}^2 C_Q (\|\nu\|_{\mathcal{B}'} \|e\|_{\mathcal{B}} + C_Q) \|\boldsymbol{\delta}_x\|_{\mathcal{B}'} \|e\|_{\mathcal{B}'}, \tag{7.22}$$

where

$$\widetilde{\psi}(t,u;k) = \nu(\mathbf{P}_t\mathbf{P}^{k-1}\mathbf{P}_u e)\mathbb{E}_x e^{itV} \mathbb{E}_x e^{iuV'}.$$

Note that $\widetilde{\psi}(t,u;k)$ does not depend on the initial state x since V and V' are independent of the Markov chain. In the same way

$$\widetilde{\phi}_1(t) = (\mathbf{P}^{l-1}\mathbf{P}_t e)(x)\mathbb{E}_x e^{itV}, \qquad \widetilde{\phi}_2(u) = (\mathbf{P}^{l+k-1}\mathbf{P}_u e)(x)\mathbb{E}_x e^{iuV'},$$

where, for $m \geq 2$,

$$\begin{aligned} (\mathbf{P}^{m-1}\mathbf{P}_t e)(x) = \boldsymbol{\delta}_x(\mathbf{P}^{m-1}\mathbf{P}_t e) &= \boldsymbol{\delta}_x(\Pi\mathbf{P}_t e) + \boldsymbol{\delta}_x(Q^{m-1}\mathbf{P}_t e) \\ &= \nu(\mathbf{P}_t e) + \boldsymbol{\delta}_x(Q^{m-1}\mathbf{P}_t e). \end{aligned}$$

Since $|\boldsymbol{\delta}_x(Q^{m-1}\mathbf{P}_t e)| \leq \kappa^{m-1}\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}C_{\mathbf{P}}C_Q$, we get

$$|\widetilde{\phi}_1(t) - \widetilde{\psi}_1(t)| \leq \kappa^{l-1}\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}C_{\mathbf{P}}C_Q, \tag{7.23}$$

$$|\widetilde{\phi}_2(u) - \widetilde{\psi}_1(u)| \leq \kappa^{l+k-1}\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}C_{\mathbf{P}}C_Q, \tag{7.24}$$

where

$$\widetilde{\psi}_1(t) = \nu(\mathbf{P}_t e)\mathbb{E}_x e^{itV} = \nu(\mathbf{P}_t e)\mathbb{E}_x e^{itV'} \tag{7.25}$$

does not depend on the initial state x of the Markov chain.

Denote $\widetilde{\psi}_0(t,u;k) = \widetilde{\psi}(t,u;k) - \widetilde{\psi}_1(t)\widetilde{\psi}_1(u)$. From (7.23) and (7.24) it follows that

$$\begin{aligned}
&|\widetilde{\phi}_0(t,u) - \widetilde{\psi}_0(t,u;k)| \\
&\leq |\widetilde{\phi}(t,u) - \widetilde{\psi}(t,u;k)| + |\widetilde{\phi}_1(t)\widetilde{\phi}_2(u) - \widetilde{\psi}_1(t)\widetilde{\psi}_1(u)| \\
&\leq |\widetilde{\phi}(t,u) - \widetilde{\psi}(t,u;k)| + |\widetilde{\phi}_1(t) - \widetilde{\psi}_1(t)| + |\nu(\mathbf{P}_t e)|\,|(\widetilde{\phi}_2(u) - \widetilde{\psi}_1(u))| \\
&\leq \kappa^{l-1}\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}C_{\mathbf{P}}^2C_Q(\|\nu\|_{\mathcal{B}'}\|e\|_{\mathcal{B}} + C_Q) \\
&\quad + \kappa^{l-1}\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}C_{\mathbf{P}}C_Q(1 + |\nu(\mathbf{P}_t e)|) \\
&\leq \kappa^{l-1}\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}(C_{\mathbf{P}}^2C_Q(\|\nu\|_{\mathcal{B}'}\|e\|_{\mathcal{B}} + C_Q) + C_{\mathbf{P}}C_Q(1 + \|\nu\|_{\mathcal{B}'}C_{\mathbf{P}})) \\
&\leq C(x)\kappa^{l-1},
\end{aligned} \tag{7.26}$$

where

$$C(x) = \|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}}(C_{\mathbf{P}}^2C_Q(\|\nu\|_{\mathcal{B}'}\|e\|_{\mathcal{B}} + C_Q) + C_{\mathbf{P}}C_Q(1 + \|\nu\|_{\mathcal{B}'}C_{\mathbf{P}})).$$

Denote by $s_{k,T}$ the complex number defined by

$$s_{k,T} = \frac{1}{(2\pi)^2}\iint \overline{\widehat{h}_T(t,u)}\widetilde{\psi}_0(t,u;k)\,dt\,du.$$

Note that $s_{k,T}$ does not depend on the initial state x of the Markov chain since neither does $\widetilde{\psi}_0(t,u;k)$. With this notation we have

$$\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) - s_{k,T} = R' + R,$$

where

$$R' = \frac{1}{(2\pi)^2}\iint \overline{\widehat{h}_T(t,u)}(\widetilde{\phi}_0(t,u) - \widetilde{\psi}_0(t,u;k))\,dt\,du.$$

Since $\mathbb{E}_x e^{itV}\mathbb{E}_x e^{iuV'}$ has support in the square $[-\varepsilon_0, \varepsilon_0]^2$, using (7.13) and (7.26) it follows that

$$|R'| \leq \frac{1}{(2\pi)^2}\|\widehat{h}_T\|_{L^2}\|\widetilde{\phi}_0 - \widetilde{\psi}_0\|_{L^2} \leq \frac{T^3}{3\pi^2}\varepsilon_0^2 C(x)\kappa^{l-1}. \tag{7.27}$$

From (7.27) and (7.19), for any $\gamma \in (0, 2\delta]$ and any $l, k = 0, 1, \ldots,$

$$|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) - s_{k,T}| \leq C(x)\frac{T^3}{3\pi^2}\varepsilon_0^2\kappa^{l-1} + c_\delta T^{-\gamma}A_0^{2+\gamma}(x), \tag{7.28}$$

From (7.28), for any $l, l' = 2, 3, \dots$ one obtains

$$|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) - \mathrm{Cov}_{\mathbb{P}_x}(f(X_{l'}), f(X_{l'+k}))| \\ \leq c_\delta T^{-\gamma} A_0(x)^{2+\gamma} + C(x)\frac{2T^3}{3\pi^2}\varepsilon_0^2 \kappa^{\min\{l,l'\}-1}.$$

Taking $T = \kappa^{-\frac{1}{4}\min\{l,l'\}}$ we get, for any $\gamma \leq \min\{1, 2\delta\}$,

$$\text{(7.29)} \qquad |\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) - \mathrm{Cov}_{\mathbb{P}_x}(f(X_{l'}), f(X_{l'+k}))| \\ \leq c_\delta A(x) \kappa^{\min\{l,l'\}\gamma/4-1},$$

where $A(x) = A_0^{2+\gamma}(x) + C(x)$. The sequence $\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k}))$, $l = 1, 2, \dots,$ is thus Cauchy; denote by $s_k(x)$ its limit as $l \to \infty$. Taking the limit as $l \to \infty$ in (7.28), we get

$$\text{(7.30)} \qquad |s_k(x) - s_{k,T}| \leq C(x)\frac{T^3}{3\pi^2}\varepsilon_0^2\kappa^{l-1} + c_\delta T^{-\gamma} A_0^{2+\gamma}(x).$$

Letting $T = T_l = \kappa^{-l/4}$ this implies that $\lim_{l\to\infty} s_{k,T_l} = s_k(x)$. Since s_{k,T_l} does not depend on x, we conclude that $s_k(x)$ is also a constant not depending on x, say s_k. Taking the limit as $l' \to \infty$ in (7.29) we obtain (7.20).

The second assertion of (b) follows from (7.20) and Lemma 7.2 upon setting $l = k$.

The third assertion of (b) follows immediately from the second one.

Let us now prove (a). From (7.8), we have

$$|\mathbb{E}_x f(X_l) - m_T| \leq \frac{1}{2\pi}\int |\overline{\widehat{g}_T(t)}|\,|\widetilde{\phi}_1(t) - \widetilde{\psi}_1(t)|\,dt + |R_1|,$$

where

$$m_T = \frac{1}{2\pi}\int \overline{\widehat{g}_T(t)}\,\widetilde{\psi}_1(t)\,dt,$$

R_1 is defined by (7.9) and $\widetilde{\psi}_1$ is defined by (7.25). Note that m_T is independent of x since so is $\widetilde{\psi}(t)$. Taking into account the bounds in (7.18) and (7.23), we get

$$|\mathbb{E}_x f(X_l) - m_T| \leq \kappa^{l-1}\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}} C_{\mathbf{P}} C_Q \frac{1}{2\pi}\int |\overline{\widehat{g}_T(t)}|\,dt + T^{-\gamma} c_\delta A_0^{1+\gamma}(x).$$

Recalling that $g_T(x) = x1(|x| \leq T)$, to bound $\int |\overline{\widehat{g}_T(t)}|\,dt$ we use the usual isometry relation

$$\left(\int |\overline{\widehat{g}_T(t)}|\,dt\right)^2 \leq \int |\overline{\widehat{g}_T(t)}|^2\,dt = \int g_T^2(x)\,dx = \frac{2}{3}T^3.$$

This implies, for any $\gamma \leq \min\{1, 2\delta\}$,

$$|\mathbb{E}_x f(X_l) - m_T| \leq c_\delta(\|\boldsymbol{\delta}_x\|_{\mathcal{B}'}\|e\|_{\mathcal{B}} C_{\mathbf{P}} C_Q T^3 \kappa^{l-1} + T^{-\gamma} A_0^{1+\gamma}(x)).$$

Taking $T = \kappa^{-l/4}$, we have

$$|\mathbb{E}_x f(X_l) - m_T| \le c_\delta A_1(x) \kappa^{l\gamma/4-1}, \tag{7.31}$$

where $A_1(x) = 1 + A_0^{1+\gamma}(x) + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'} \|e\|_{\mathcal{B}} C_{\mathbf{P}} C_Q$. From this inequality it follows that

$$|\mathbb{E}_x f(X_l) - \mathbb{E}_x f(X_k)| \le c_\delta A_1(x) \kappa^{\min\{l,k\}\gamma/4-1}, \tag{7.32}$$

which proves that the sequence $(\mathbb{E}_x f(X_l))_{l\ge 1}$ is Cauchy and therefore has a limit denoted $\mu(x)$. Since m_T does not depend on x, letting $l \to \infty$ in (7.31) we conclude that $\mu(x) = \mu$ does not depend on x. Letting $k \to \infty$ in (7.32), we get

$$|\mathbb{E}_x f(X_l) - \mu| \le c_\delta A_1(x) \kappa^{l \min\{1,\gamma\}/4-1},$$

which proves the first assertion of (a). The second follows from the first. ∎

The bound (3.5) of Proposition 3.2 follows from part (a) of Lemma 7.3. It remains to prove the bound (3.6).

Let $0 < \gamma \le \min\{1, 2\delta\}$. First note that, from Lemmas 7.3 and 7.2 we obtain, for $k = 0, 1, \dots,$

$$|\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) - s_k| \le A_2(x) c_{\delta,\kappa} \kappa^{c_{\gamma,\kappa} \max\{l,k\}},$$

where $A_2(x)$ is defined in Proposition 3.2. Then, for any $k = 0, 1, \dots,$

$$\sum_{l=m}^{m+n-1} \sum_{k=1}^{m+n-l} |\,\mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k})) - s_k| \\ \le A_2(x) c_{\delta,\kappa} \sum_{l=m}^{m+n-1} \sum_{k=1}^{m+n-l} e^{-c_{\gamma,\kappa} \max\{l,k\}} \le A_2(x) c''_{\delta,\gamma,\kappa}.$$

Since

$$\mathrm{Var}_{\mathbb{P}_x}\Big(\sum_{l=m}^{m+n-1} f(X_l)\Big) = \sum_{l=m}^{m+n-1} \mathrm{Var}_{\mathbb{P}_x}(f(X_l)) \\ + 2 \sum_{l=m}^{m+n-1} \sum_{k=1}^{m+n-l} \mathrm{Cov}_{\mathbb{P}_x}(f(X_l), f(X_{l+k}))$$

we get

$$\Big|\mathrm{Var}_{\mathbb{P}_x}\Big(\sum_{l=m}^{m+n-1} f(X_l)\Big) - \Big(ns_0 + \sum_{l=m}^{m+n-1} \sum_{k=1}^{m+n-l} (s_k + s_k^*)\Big)\Big| \le A_2(x) c''_{\delta,\gamma,\kappa}.$$

Taking into account that, by Lemma 7.3, the s_k are independent of x and that $|s_k| \le A_2(x)\kappa^{k\gamma/4-1}$ we obtain

$$\Big|\mathrm{Var}_{\mathbb{P}_x}\Big(\sum_{l=m}^{m+n-1} f(X_l)\Big) - n\Big(s_0 + \sum_{k=1}^{\infty}(s_k + s_k^*)\Big)\Big| \le A_2(x) c'''_{\delta,\gamma,\kappa}. \tag{7.33}$$

Dividing by n and taking the limit as $n \to \infty$ in (7.33), we deduce that $s_0 + \sum_{k=1}^{\infty}(s_k + s_k^*)$ converges to a non-negative number not depending on x, say $\sigma^2 \geq 0$. Now (3.6) follows from (7.33).

7.3. Proof of Theorem 3.3. First note that Conditions **C1** and **C3** are satisfied by Propositions 3.1 and 3.2. Condition **C2** is satisfied by Hypothesis **M4**. Let $\mu_i(x) = \mathbb{E}_x f(X_i)$. Let $\alpha < \delta$ and $\delta' = \frac{1}{2}(\alpha + \delta)$. Since $\alpha < \delta'$, from 2.1 with δ' replacing δ, it follows that for any $x \in \mathbb{X}$ there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P}_x)$, a sequence of independent standard normal r.v.'s $(W_i')_{i\geq 1}$ and a sequence of r.v.'s $(Y_i')_{i\geq 1}$ such that $(Y_i')_{i\geq 1} \stackrel{d}{=} (f(X_i))_{i\geq 1}$ and, for any $0 < \rho < \frac{1}{2}\frac{\alpha}{1+2\alpha}$,

$$\mathbb{P}_x\Big(N^{-1/2} \sup_{k\leq N}\Big|\sum_{i=1}^{k}(Y_i' - \mu_i(x) - \sigma W_i')\Big| > N^{-\rho}\Big) \leq C_0(x) N^{-\alpha\frac{1+\alpha}{1+2\alpha} + \rho(2+2\alpha)}, \tag{7.34}$$

where $C_0(x) = C_0'(1 + \lambda_0(x) + \mu_{\delta'}(x) + \sqrt{\tau(x)})^{2+2\delta'}$ and $\lambda_0(x), \mu_\delta(x), \tau(x), \lambda_1, \lambda_2$ and σ^2 are defined in Propositions 3.1 and 3.2. If $\overline{\mu}(x) \leq N^{\frac{1}{2}-\rho}$ (with $\overline{\mu}(x)$ from Proposition 3.2) then using (7.34) we have

$$\begin{aligned}\mathbb{P}_x\Big(N^{-1/2} \sup_{k\leq N}\Big|\sum_{i=1}^{k}(Y_i' - \mu - \sigma W_i')\Big| > 2N^{-\rho}\Big) &\leq \mathbb{P}_x\Big(\sup_{k\leq N}\Big|\sum_{i=1}^{k}(Y_i' - \mu_i(x) - \sigma W_i')\Big| > 2N^{1/2-\rho} - \overline{\mu}(x)\Big) \\ &\leq C_0(x) N^{-\alpha\frac{1+\alpha}{1+2\alpha} + \rho(2+2\alpha)}.\end{aligned} \tag{7.35}$$

If $\overline{\mu}(x) > N^{1/2-\rho}$, it is obvious that

$$1 \leq (\overline{\mu}(x) N^{-1/2+\rho})^{2\alpha} \leq \overline{\mu}(x)^{2\alpha} N^{-\alpha + 2\rho\alpha}. \tag{7.36}$$

From (7.35) and (7.36) we get

$$\mathbb{P}_x\Big(N^{-1/2} \sup_{k\leq N}\Big|\sum_{i=1}^{k}(Y_i' - \mu - \sigma W_i')\Big| > 2N^{-\rho}\Big) \leq (C_0(x) + \overline{\mu}(x)^{2\alpha}) N^{-\alpha\frac{1+\alpha}{1+2\alpha} + \rho(2+2\alpha)}.$$

Taking into account the expressions for $\lambda_0(x), \mu_\delta(x), \tau(x), \lambda_1, \lambda_2$, $\overline{\mu}(x)$ and choosing γ small we obtain

$$C_0(x) + \overline{\mu}(x)^{2\alpha} \leq C(x) = C_1(1 + \|\boldsymbol{\delta}_x\|_{\mathcal{B}'} + \mu_\delta(x))^{2+2\delta},$$

where C_1 is a constant depending only on $\delta, \alpha, \kappa, C_{\mathbf{P}}, C_Q, \|e\|_{\mathcal{B}}, \|\nu\|_{\mathcal{B}'}$.

Generally the measure $\mathbb{P}_x$ and the constructed sequence $(Y_i')_{i\geq 1}$ both depend on the initial state x. It is easy to reconstruct $(Y_i')_{i\geq 1}$ independently of x. Indeed, on the canonical space $\widetilde{\Omega} = \mathbb{R}^\infty \times \mathbb{R}^\infty$ there is a probability measure $\widetilde{\mathbb{P}}_x$ which coincides with the joint distribution of the sequence $(Y_i', W)_{i\geq 1}$. It is enough to redefine $Y_i' = \omega_{1,i}$ and $W_i = \omega_{2,i}$ as the coordinate processes, where $\omega = (\omega_1, \omega_2) \in \widetilde{\Omega}$. With this construction only the measure $\widetilde{\mathbb{P}}_x$ depends on the initial state x. The measurability of the map $x \in \mathbb{X} \mapsto \widetilde{\mathbb{P}}_x(\cdot)$ follows from the construction.

7.4. Proof of Theorem 3.4. In addition to conditions of Theorem 3.3 assume Hypothesis **M5** holds. First we note that **M5** ensures the existence of the mean $\nu(f) = \mathbf{E}_\nu X_k = \int (\mathbb{E}_x X_k)\, \nu(dx)$ and of the mixed moment $\mathbf{E}_\nu (X_l X_{l+k}) = \int \mathbb{E}_x (X_l X_{l+k})\, \nu(dx)$ with respect to the invariant measure. By Proposition 3.2, we have $\lim_{k\to\infty} \mathbb{E}_x X_k = \mu$, ν-a.s. on $\mathbb{X}$. Then by the Lebesgue dominated convergence theorem

$$\nu(f) = \mathbf{E}_\nu X_k = \lim_{k\to\infty} \int (\mathbb{E}_x X_k)\, \nu(dx) = \int (\lim_{k\to\infty} \mathbb{E}_x X_k)\, \nu(dx) = \mu.$$

Without loss of generality we can assume that $\nu(f) = 0$. Using Hypothesis **M5** and $\nu(f) = 0$, we have

$$\begin{aligned}\int \mathrm{Cov}_{\mathbf{P}_x}(X_l, X_{l+k})\, \nu(dx) &= \int \mathbb{E}_x(X_l X_{l+k})\, \nu(dx) - \int \mathbb{E}_x(X_l)\mathbb{E}_x(X_{l+k})\, \nu(dx) \\ &= \mathbf{E}_\nu(X_l X_{l+k}) - \int \mathbb{E}_x(X_l)\mathbb{E}_x(X_{l+k})\, \nu(dx) \\ &= \mathrm{Cov}_{\mathbb{P}_\nu}(X_0, X_k) - \int \mathbb{E}_x(X_l)\mathbb{E}_x(X_{l+k})\, \nu(dx).\end{aligned}$$

By Proposition 3.2, $\lim_{l\to\infty} \mathrm{Cov}_{\mathbf{P}_x}(X_l, X_{l+k}) = s_k$ and $\lim_{l\to\infty} \mathbb{E}_x(X_l) = 0$ for any $x \in \mathbb{X}$. As before, integrating with respect to the stationary measure and using the Lebesgue dominated convergence theorem, it follows that $s_k = \mathrm{Cov}_{\mathbb{P}_\nu}(X_0, X_k)$. Thus the conclusions of Theorem 3.3 hold true with $\mu = \nu(f)$ and $\sigma^2 = \sigma_\nu^2$, which proves Theorem 3.4.

8. Maximal inequalities. In this section we state two bounds which are used repeatedly in the paper. The first one gives control on the L_p-norm of the maxima of the partial sums of a sequence of dependent r.v.'s. This proposition is a consequence of the second one which gives control on the L_p-norm of the partial sums of a sequence of dependent r.v.'s. It is assumed that Conditions **C1** and **C2** hold true.

PROPOSITION 8.1. *Let $\delta' < \delta$ and $\epsilon > 0$. Then there is a constant $c_{\lambda_1,\lambda_2,\delta,\delta',\epsilon}$ such that for any $m, n \geq 1$,*

$$\left\| \sup_{1\leq k\leq n} \left| \sum_{i=m}^{m+k-1} X_i \right| \right\|_{L^{2+2\delta'}} \leq c_{\lambda_1,\lambda_2,\delta,\delta',\epsilon} (1 + \lambda_0 + \mu_\delta)^{1+\epsilon} n^{1/2}.$$

Proof. Denote for brevity $S_{m,n} = \sum_{i=m}^{m+n-1} X_i$. Let δ'' be such that $\delta' < \delta'' < \delta$. By Proposition 8.2 below (which we assume for the moment), for any $m, n \geq 1$ and $\epsilon > 0$ we have $\|S_{m,n}\|_{L^{2+2\delta''}} \leq A n^{1/2}$, where $A = c_{\lambda_1,\lambda_2,\delta'',\epsilon}(1 + \lambda_0 + \mu_\delta)^{1+\epsilon}$. Letting $S'_{m,n} = S_{m,n}/A$, we get $\|S'_{m,n}\|_{L^{2+2\delta''}} \leq n^{1/2}$ for any $m, n \geq 1$. By Theorem A in Serfling [35] (see also Billingsley [3, p. 102]), we see follows that $\|\sup_{1\leq k\leq n} S'_{m,n}\|_{L^{2+2\delta''}} \leq n^{1/2}\log_2(4n)$ for any $m, n \geq 1$. Since $\delta' < \delta''$, it follows that

$$\Big\|\sup_{1\leq k\leq n} S'_{m,n}\Big\|_{L^{2+2\delta'}} \leq \Big\|\sup_{1\leq k\leq n} S'_{m,n}\Big\|_{L^{2+2\delta''}} \leq (n^{1/2}\log_2(4n))^{\frac{2+2\delta''}{2+2\delta'}} \leq c_{\delta,\delta'} n^{1/2},$$

from which we deduce $\|\sup_{1\leq k\leq n} S_{m,n}\|_{L^{2+2\delta'}} \leq A c_{\delta,\delta'} n^{1/2}$. ∎

The following assertion is an adaptation of Proposition 4.1 in Gouëzel [15]. In order to derive an explicit dependence of the constant involved in the bound on some of the constants in Conditions **C1** and **C2** we give an independent proof. Tracking this explicit dependence plays a crucial role in the proof of Theorem 3.3 to work out the dependence of the bound on the initial state of the Markov chain $X_0 = x$.

PROPOSITION 8.2. *Let* $0 < \delta' < \delta$ *and* $\epsilon > 0$. *Then there is a constant* $c_{\lambda_1,\lambda_2,\delta',\epsilon}$ *such that, for any* $m, n \geq 1$,

$$\Big\|\sum_{i=m}^{m+n-1} X_i\Big\|_{L^{2+2\delta'}} \leq c_{\lambda_1,\lambda_2,\delta',\epsilon}(1 + \lambda_0 + \mu_\delta)^{1+\epsilon} n^{1/2}.$$

The proof of this proposition is given below. First we state several auxiliary assertions.

8.1. Auxiliary assertions

PROPOSITION 8.3. *There is a constant* $c_{\lambda_1,\lambda_2,\epsilon}$ *such that, for any* $\epsilon > 0$,

$$\Big\|\sum_{i=m}^{m+n-1} X_i\Big\|_{L^2} \leq c_{\lambda_1,\lambda_2,\epsilon}(1 + \lambda_0 + \mu_\delta)^{1+\epsilon} n^{1/2}. \tag{8.1}$$

The proof is based on the following two lemmas.

LEMMA 8.4. *Let* $u_n = \max_{m\geq 1} \|\sum_{i=m}^{m+n-1} X_i\|_{L^2}^2$, $n \geq 1$. *Then, for any natural numbers* a, $b \geq 1$ *and any* $\alpha \in (0, 1/2)$, $\gamma \in (0, \delta)$,

$$\begin{aligned} u_{a+b} \leq A + u_a + u_b &+ (c\mu_\delta)^2(a^{2\alpha} + b^{2\alpha}) \\ &+ c\mu_\delta(a^\alpha + b^\alpha)(A + u_a + u_b)^{1/2} + c u_a^{1/2} + c u_b^{1/2}, \end{aligned}$$

where $c > 1$ *and* $A = c_{\lambda_1,\lambda_2,\gamma,\alpha}(1 + \lambda_0 + \mu_\delta)^{2+\gamma}$.

Proof. Let $m \in \mathbb{N}$. Assume that $a \leq b$ (the case $a > b$ is treated in the same manner). Denote $Y_1 = \sum_{i=m}^{m+a-1} X_i$, $Y_2 = \sum_{i=m+a+[b^\alpha]}^{m+a+b-1} X_i$ and

$Y_0 = \sum_{i=m}^{m+a+b-1} X_i$, where $\alpha \in (0, 1/2)$. Note that $Y_0 = Y_1 + Y_2 + Y_{\text{gap}}$, where $Y_{\text{gap}} = \sum_{i=m+a}^{m+a+[b^\alpha]-1} X_i$. Therefore

$$\|Y_0\|_{L^2}^2 \le \|Y_1 + Y_2\|_{L^2}^2 + \|Y_{\text{gap}}\|_{L^2}^2 + 2\|Y_1 + Y_2\|_{L^2}\|Y_{\text{gap}}\|_{L^2}. \tag{8.2}$$

We shall bound each of the terms on the right-hand side of (8.2).

Let V_1 and V_2 be two independent identically distributed r.v.'s of mean 0, independent of Y_1 and Y_2 with a common characteristic function supported in the interval $[-\varepsilon_0, \varepsilon_0]$ for some $\varepsilon_0 \in (0, 1)$, and such that $\|V_i\|_{L^{2+2\delta}} \le c$. Denote $\widetilde{Y}_1 = Y_1 + V_1$ and $\widetilde{Y}_2 = Y_2 + V_2$. Let Z_1 and Z_2 be independent copies of $\widetilde{Y}_1$ and $\widetilde{Y}_2$. Since $\mathbf{E}e^{itV_1}$ is supported in $[-\varepsilon_0, \varepsilon_0]$, by Lemma 9.5, for any $T > 0$,

$$\begin{aligned}\pi((\widetilde{Y}_1, \widetilde{Y}_2), (Z_1, Z_2)) \le \frac{T}{\pi}\Big(\int_{[-\varepsilon_0,\varepsilon_0]^2} |\phi(t, u) - \psi_1(t)\psi_2(u)|^2 \, dt \, du\Big)^{1/2} \\ + \mathbb{P}(\max\{|\widetilde{Y}_1|, |\widetilde{Y}_2|\} > T),\end{aligned}$$

where ϕ is the characteristic function of the vector (Y_1, Y_2) and ψ_1, and ψ_2 are the characteristic functions of Y_1 and Y_2. Condition **C1** implies that

$$\begin{aligned}|(\phi(t, u) - \psi_1(t)\psi_2(u))| &\le \lambda_0(1 + b)^{2\lambda_2} \exp(-\lambda_1[b^\alpha]) \\ &\le \lambda_0 c_{\lambda_1}(1 + b)^{2\lambda_2} \exp(-\lambda_1 b^\alpha).\end{aligned}$$

Let $T = e^{\lambda_1 b^\alpha/2}$. Taking into account that

$$\begin{aligned}\mathbb{P}(\max\{|\widetilde{Y}_1|, |\widetilde{Y}_2|\} > T) &\le T^{-1}\mathbb{E}\max\{|\widetilde{Y}_1|, |\widetilde{Y}_2|\} \\ &\le T^{-1}(\|Y_1 + V_1\|_{L^{2+2\delta}} + \|Y_2 + V_2\|_{L^{2+2\delta}}) \\ &\le e^{-\frac{\lambda_1}{2}b^\alpha}\Big(c + (a + b)\max_{l\ge0}\|X_l\|_{L^{2+2\delta}}\Big) \\ &\le c_\delta e^{-\frac{\lambda_1}{2}b^\alpha} b(1 + \mu_\delta)\end{aligned}$$

we obtain

$$\begin{aligned}\pi((\widetilde{Y}_1, \widetilde{Y}_2), (Z_1, Z_2)) &\le \frac{1}{\pi}\lambda_0(1 + b)^{2\lambda_2}e^{-\frac{\lambda_1}{2}b^\alpha} + c_\delta b e^{-\frac{\lambda_1}{2}b^\alpha}(1 + \mu_\delta) \\ &\le \varDelta = c_{\delta,\alpha}(1 + b)^{2\lambda_2}e^{-\frac{\lambda_1}{2}b^\alpha}(1 + \lambda_0 + \mu_\delta).\end{aligned}$$

By Lemma 9.1 there is a coupling of $(\widetilde{Y}_1, \widetilde{Y}_2)$ and (Z_1, Z_2) such that

$$P(\|(\widetilde{Y}_1, \widetilde{Y}_2) - (Z_1, Z_2)\|_\infty \ge \varDelta) \le \varDelta.$$

Let $S = \widetilde{Y}_1 + \widetilde{Y}_2 - (Z_1 + Z_2)$. Taking into account that $\|V_i\|_{L^{2+2\delta}} \le c$, we have

$$\begin{aligned}\|S\|_{L^{2+2\delta}} &= \|\widetilde{Y}_1 + \widetilde{Y}_2 - (Z_1 + Z_2)\|_{L^{2+2\delta}} \le 2\|\widetilde{Y}_1 + \widetilde{Y}_2\|_{L^{2+2\delta}} \\ &\le c(a + b)\Big(1 + \max_{l\ge0}\|X_l\|_{L^{2+2\delta}}\Big) \le cb(1 + \mu_\delta).\end{aligned} \tag{8.3}$$

Then, for any $\gamma \in (0,\delta)$,

$$
\begin{aligned}
(8.4)\quad \|S\|_{L^2}^2 &\le 4\varDelta^2 + \mathbb{E}|S|^2 1(|S| \ge 2\varDelta) \\
&\le 4\varDelta^2 + \|S\|_{L^{2+2\gamma}}^2 \mathbb{P}(|S| \ge 2\varDelta)^{\frac{\gamma}{1+\gamma}} \\
&\le 4\varDelta^2 + cb^2(1+\mu_\delta)^2 \varDelta^{\frac{\gamma}{1+\gamma}} \\
&\le 4c^2(1+b)^{4\lambda_2}(1+\lambda_0+\mu_\delta)^2 e^{-\lambda_1 b^\alpha} \\
&\quad + 4b^2(1+\mu_\delta)^2 c_{\delta,\alpha}^{\frac{\gamma}{1+\gamma}} (1+b)^{4\lambda_2 \frac{\gamma}{1+\gamma}} e^{-\frac{\lambda_1}{2}\frac{\gamma}{1+\gamma} b^\alpha} (1+\lambda_0+\mu_\delta)^{\frac{\gamma}{1+\gamma}} \\
&\le A',
\end{aligned}
$$

where $A' = c'_{\lambda_1,\lambda_2,\gamma,\alpha}(1+\lambda_0+\mu_\delta)^{2+\gamma}$. From (8.3) and (8.4), it follows that

$$
\begin{aligned}
(8.5)\qquad \|\widetilde{Y}_1 + \widetilde{Y}_2\|_{L^2}^2 &\le \|S\|_{L^2}^2 + \|Z_1\|_{L^2}^2 + \|Z_2\|_{L^2}^2 \\
&\le A' + \|\widetilde{Y}_1\|_{L^2}^2 + \|\widetilde{Y}_2\|_{L^2}^2.
\end{aligned}
$$

Since $\|V_i\|_{L^2} \le c$, we have

$$
(8.6)\qquad \|Y_1 + Y_2\|_{L^2} \le \|\widetilde{Y}_1 + \widetilde{Y}_2\|_{L^2} + 2c.
$$

Taking into account (8.5) and (8.6), one gets

$$
(8.7)\quad \|Y_1 + Y_2\|_{L^2}^2 \le A' + \|\widetilde{Y}_1\|_{L^2}^2 + \|\widetilde{Y}_2\|_{L^2}^2 + 4c\|\widetilde{Y}_1\|_{L^2} + 4c\|\widetilde{Y}_2\|_{L^2} + 4c^2.
$$

Since $\|\widetilde{Y}_k\|_{L^2} \le \|Y_k\|_{L^2} + c$, we see that

$$
\begin{aligned}
(8.8)\quad \|Y_1 + Y_2\|_{L^2}^2 &\le A' + (\|Y_1\|_{L^2} + c)^2 + (\|Y_2\|_{L^2} + c)^2 \\
&\quad + 4c(\|Y_1\|_{L^2} + c) + 4c(\|Y_2\|_{L^2} + c) + 4c^2 \\
&\le A' + \|Y_1\|_{L^2}^2 + \|Y_2\|_{L^2}^2 + 6c(\|Y_1\|_{L^2} + \|Y_2\|_{L^2}) + 14c^2.
\end{aligned}
$$

Since the gap is of size $[b^\alpha]$,

$$
(8.9)\qquad \sup_{m\ge1} \|Y_{\mathrm{gap}}\|_{L^2} \le [b^\alpha] \max_{i\ge1} \|X_i\|_{L^{2+2\delta}} \le b^\alpha \mu_\delta.
$$

From (8.2), (8.8) and (8.9) we obtain

$$
\begin{aligned}
(8.10)\qquad \|Y_0\|_{L^2}^2 &\le A' + \|Y_1\|_{L^2}^2 + \|Y_2\|_{L^2}^2 + b^{2\alpha}\mu_\delta^2 \\
&\quad + 2b^\alpha \mu_\delta (\|Y_1\|_{L^2} + \|Y_2\|_{L^2}) \\
&\quad + 6c(\|Y_1\|_{L^2} + \|Y_2\|_{L^2}) + 14c^2.
\end{aligned}
$$

Now recall that $u_a = \sup_{m\ge1} \|Y_1\|_{L^2}^2$, $u_b = \sup_{m\ge1} \|Y_2\|_{L^2}^2$ and $u_{a+b} = \sup_{m\ge1} \|Y_0\|_{L^2}^2$. Using (8.9) we have $\|Y_2\|_{L^2} \le u_b^{1/2} + \|Y_{\mathrm{gap}}\|_{L^2} \le u_b^{1/2} + b^\alpha \mu_\delta$. From this and (8.10) we deduce that

$$
\begin{aligned}
u_{a+b} &\le A' + u_a + u_b + b^{2\alpha}\mu_\delta^2 + 2b^\alpha \mu_\delta u_b^{1/2} + 2b^\alpha \mu_\delta (u_a^{1/2} + u_b^{1/2} + b^\alpha \mu_\delta) \\
&\quad + 6c(u_a^{1/2} + u_b^{1/2} + b^\alpha \mu_\delta) + 14c^2.
\end{aligned}
$$

Rearranging the terms and taking into account that $A' > 1$, we obtain

$$\begin{aligned} u_{a+b} &\le A' + 14c^2 + u_a + u_b + 3b^{2\alpha}\mu_\delta^2 \\ &\quad + b^\alpha \mu_\delta (6c + u_a^{1/2} + u_b^{1/2}) + 6cu_a^{1/2} + 6cu_b^{1/2}. \ \blacksquare \end{aligned}$$

LEMMA 8.5. *Assume that the sequence* $(u_n)_{n\ge 1}$ *is such that* $u_n > 0$ *and*

$$\begin{aligned} u_{a+b} &\le (u_a + u_b + A) + (a^{2\alpha} + b^{2\alpha})B^2 \\ &\quad + (a^\alpha + b^\alpha) B (u_a + u_b + A)^{1/2} + cu_a^{1/2} + cu_b^{1/2} \end{aligned}$$

for all a, $b \ge 1$ *and some* $A, B > 0$, $\alpha \in (0, 1/2)$. *Then*

$$u_n \le c_\alpha (1 + u_1 + A + B^2) n.$$

Proof. Note that $xy \le \frac{1}{2}(\varepsilon x^2 + \varepsilon^{-1} y^2)$ for any $x, y, \varepsilon > 0$. Using the assumption of the lemma, we have

$$\begin{aligned} u_{a+b} &\le (u_a + u_b + A) + (1 + \varepsilon^{-1})(a^{2\alpha} + b^{2\alpha})B^2 \\ &\quad + \frac{\varepsilon}{2}(u_a + u_b + A) + \varepsilon^{-1} c^2 + \frac{\varepsilon}{2} u_\alpha + \frac{\varepsilon}{2} u_b \\ &\le (1 + \varepsilon)(u_a + u_b + A) + \varepsilon^{-1} c^2 + (1 + \varepsilon^{-1})(a^{2\alpha} + b^{2\alpha})B^2. \end{aligned}$$

Denote $v_k = \max_{1 \le n \le 2^k} u_n$, $k \ge 0$. From the above inequality it follows that

$$v_{k+1} \le (1 + \varepsilon)(2v_k + A) + \varepsilon^{-1} c^2 + (1 + \varepsilon^{-1}) 2^{2\alpha k + 1} B^2.$$

Dividing by $(2 + 2\varepsilon)^{k+1}$ we get

$$\begin{aligned} \frac{v_{k+1}}{(2 + 2\varepsilon)^{k+1}} &\le \frac{2v_k + A}{2(2 + 2\varepsilon)^k} + (1 + \varepsilon^{-1}) \frac{2^{2\alpha k + 1}}{(2 + 2\varepsilon)^{k+1}} (B^2 + c^2) \\ &\le \frac{v_k}{(2 + 2\varepsilon)^k} + \frac{A}{2(2 + 2\varepsilon)^k} + (1 + \varepsilon^{-1}) \frac{2^{2\alpha k + 1}}{(2 + 2\varepsilon)^{k+1}} (B^2 + c^2). \end{aligned}$$

Taking into account that $\alpha < 1/2$, by induction, we obtain

$$\begin{aligned} \frac{v_k}{(2 + 2\varepsilon)^k} &\le v_0 + \frac{A}{2} \sum_{i=0}^{\infty} \frac{1}{(2 + 2\varepsilon)^i} + 2(1 + \varepsilon^{-1})(B^2 + c^2) \sum_{i=1}^{\infty} \frac{2^{2\alpha i}}{(2 + 2\varepsilon)^{i+1}} \\ &\le v_0 + \frac{A}{2} \frac{2 + 2\varepsilon}{1 + 2\varepsilon} + 2(1 + \varepsilon^{-1})(B^2 + c^2) \frac{1 + \varepsilon}{\varepsilon} \\ &\le v_0 + c'_\varepsilon (A + B^2 + c^2), \end{aligned}$$

where c'_ε depends only on ε. This implies that

$$v_k = \max_{1 \le n \le 2^k} u_n \le C_0 (2 + 2\varepsilon)^k,$$

where $C_0 = v_0 + c'_\varepsilon (A + B^2 + c^2)$. Once again using the assumption of the lemma it follows that

$$\begin{aligned} v_{k+1} &\le (2v_k + A) + 2^{2\alpha k+1}B^2 \\ &\qquad + 2^{\alpha k+1}B(2C_0(2+2\varepsilon)^k + A)^{1/2} + 2cC_0^{1/2}(2+2\varepsilon)^{k/2} \\ &\le (2v_k + A) + 2^{2\alpha k+1}B^2 \\ &\qquad + 2^{\alpha k+1}B(2C_0^{1/2}(2+2\varepsilon)^{k/2} + A^{1/2}) + 2cC_0^{1/2}(2+2\varepsilon)^{k/2}. \end{aligned}$$

Dividing by 2^{k+1} and choosing $\varepsilon = \varepsilon(\alpha)$ so small that $2 + 2\varepsilon \le 2^{1+(1/2-\alpha)}$, one gets

$$\begin{aligned} \frac{v_{k+1}}{2^{k+1}} &\le \frac{v_k}{2^k} + \frac{A}{2^{k+1}} + 2^{(2\alpha-1)k+1}B^2 + 2^{(\alpha-1)k+1}B(2C_0^{1/2}2^{\frac{k}{2}+\frac{k}{2}(\frac{1}{2}-\alpha)} + A^{1/2}) \\ &\quad + 2^{-k}cC_0^{1/2}2^{\frac{k}{2}+\frac{k}{2}(\frac{1}{2}-\alpha)} \\ &\le \frac{v_k}{2^k} + \frac{A}{2^{k+1}} + 2^{(2\alpha-1)k+1}B^2 + 4BC_0^{1/2}2^{(\alpha-\frac{1}{2})\frac{k}{2}} + 2^{(\alpha-1)k+1}BA^{1/2} \\ &\quad + cC_0^{1/2}2^{-\frac{k}{2}(\frac{1}{2}+\alpha)}. \end{aligned}$$

Using induction, this implies

$$v_k/2^k \le c_\alpha(A + C_0 + B^2),$$

since ε depends only on α. From this we get $u_{2^k} \le D2^k$ for any $k \ge 1$, where $D = c'_\alpha(1 + v_0 + A + B^2)$. Therefore, for any $2^{k-1} \le n \le 2^k$ we conclude that $u_n \le D2^k \le 2D2^{k-1} \le 2Dn$. ∎

Let $\alpha = 1/4$. In the notations of Lemma 8.4, $u_1 \le \max_{m\ge1} \|X_m\|^2_{L^{2+2\delta}} \le \mu_\delta^2$. From Lemmas 8.4 and 8.5 with $B = c\mu_\delta$ it follows, for any $\epsilon \in (0, \delta)$, that

$$\begin{aligned} \max_{m\ge1}\Big\| \sum_{i=m}^{m+n-1} X_i \Big\|^2_{L^2} = u_n &\le c_\alpha(u_1 + A + c^2\mu_\delta^2)n \\ &\le c(c_{\lambda_1,\lambda_2,\epsilon}(1 + \lambda_0 + \mu_\delta)^{2+\frac{\epsilon}{1+\epsilon}} + 2\mu_\delta^2)n \\ &\le c'_{\lambda_1,\lambda_2,\epsilon}(1 + \lambda_0 + \mu_\delta)^{2+\epsilon}n, \end{aligned}$$

which proves Proposition 8.3.

8.2. Proof of Proposition 8.2. Let $m, n \in \mathbb{N}$ and $a = [n^{1-\alpha}]$ and $b = [n^{\alpha+\rho}]$, where $\alpha > 0$ and $\rho > 0$ are such that $2\alpha + \rho < 1$. Note that $a > b$ and $ba \le n^{1-\rho}$. Consider the intervals $I_k = [m + (k-1)a, m + ka - b)$, $J_k = [m + ka - b, m + ka)$ for $k = 1, \dots, [n^\alpha]$, and $I_{\text{fin}} = [m + ba, m + n)$, such that $[m, m+n) = \bigcup_{k=1}^{[n^\alpha]}(I_k \cup J_k) \cup I_{\text{fin}}$. Here $a - b > 0$ and $b > 0$ are interpreted as the length of an island I_k and the length of a gap J_k respectively.

Denote $Y_k = \sum_{i\in I_k} X_i$, $k = 1, \dots, [n^\alpha]$. Let $V_1, \dots, V_{[n^\alpha]}$ be independent identically distributed r.v.'s of mean 0, independent of $Y_1, \dots, Y_{[n^\alpha]}$ with a common characteristic function supported in $[-\varepsilon_0, \varepsilon_0]$ for some $\varepsilon_0 > 0$, and such that $\|V_k\|_{L^{2+2\delta}} \le c$, $k = 1, \dots, [n^\alpha]$. Denote $\widetilde{Y}_k = Y_k + V_k$. Let $Z_1, \dots, Z_{[n^\alpha]}$ be independent copies of $\widetilde{Y}_1, \dots, \widetilde{Y}_{[n^\alpha]}$. By Lemma 9.3,

$$\pi((\widetilde{Y}_1, \dots, \widetilde{Y}_{[n^\alpha]}), (Z_1, \dots, Z_{[n^\alpha]})) \le \sum_{k=1}^{[n^\alpha]} \pi((\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, \widetilde{Y}_k), (\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, , Z_k)). \tag{8.11}$$

Since $\mathbf{E}e^{itV_k}$ is supported in $[-\varepsilon_0, \varepsilon_0]$, by Lemma 9.5, for any $T > 0$ and $k \le [n^\alpha]$,

$$\begin{aligned} \pi((\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, \widetilde{Y}_k), (\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, , Z_k)) &\le \frac{T}{\pi}\Big(\int_{[-\varepsilon_0,\varepsilon_0]^k} |\phi(t,u) - \psi_1(t)\psi_2(u)|^2 \, dt \, du \Big)^{1/2} \\ &\quad + \mathbb{P}(\|(\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, \widetilde{Y}_k)\|_\infty > T), \end{aligned} \tag{8.12}$$

where ϕ is the characteristic function of $(\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, \widetilde{Y}_k)$, and ψ_1 and ψ_2 are the characteristic functions of the r.v.'s $(\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1})$ and $\widetilde{Y}_k$. Condition **C1** implies that

$$\begin{aligned} |\phi(t,u) - \psi_1(t)\psi_2(u)| &\le \lambda_0(1+a)^k \exp(-\lambda_1 b) \\ &\le c_{\lambda_1}\lambda_0(1 + n^{1-\alpha})^{n^\alpha} \exp(-\lambda_1 n^{\alpha+\rho}). \end{aligned} \tag{8.13}$$

Let $T = e^{\frac{\lambda_1}{2}n^{\alpha+\rho}}$. By Chebyshev's inequality, taking into account that $k \le [n^\alpha]$, we have

$$\begin{aligned} \mathbb{P}(\|(\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, \widetilde{Y}_k)\|_\infty > T) &\le T^{-1}\sum_{i=1}^k \|Y_i + V_i\|_{L^1} \le T^{-1}\sum_{i=1}^k (\|Y_i\|_{L^{2+2\delta}} + c) \\ &\le e^{-\frac{\lambda_1}{2}n^{\alpha+\rho}}[n^\alpha]([n^{1-\alpha}]\mu_\delta + c) \le cne^{-\frac{\lambda_1}{2}n^{\alpha+\rho}}(1 + \mu_\delta). \end{aligned} \tag{8.14}$$

From (8.12)–(8.14) we obtain

$$\begin{aligned} &\pi((\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, \widetilde{Y}_k), (\widetilde{Y}_1, \dots, \widetilde{Y}_{k-1}, , Z_k)) \\ &\qquad \le \frac{T}{\pi}\lambda_0\varepsilon_0^n(1 + n^{1-\alpha})^{n^\alpha} e^{-\lambda_1 n^{\alpha+\rho}} + c_\alpha ne^{-\frac{\lambda_1}{2}n^{\alpha+\rho}}(1 + \mu_\delta) \\ &\qquad \le c\varepsilon_0^n n(1 + n^{1-\alpha})^{n^\alpha} e^{-\frac{\lambda_1}{2}n^{\alpha+\rho}}(1 + \lambda_0 + \mu_\delta). \end{aligned}$$

Inserting this bound in (8.11) we get

$$\pi((\widetilde{Y}_1,\dots,\widetilde{Y}_{[n^\alpha]}),(Z_1,\dots,Z_{[n^\alpha]}))$$
$$\le c\varepsilon_0^n n^{1+\alpha}(1+n^{1-\alpha})^{n^\alpha} e^{-\frac{\lambda_1}{2}n^{\alpha+\rho}}(1+\lambda_0+\mu_\delta)$$
$$\le \varDelta = c_{\alpha,\lambda_1} e^{-\frac{\lambda_1}{4}n^{\alpha+\rho}}(1+\lambda_0+\mu_\delta).$$

According to Strassen–Dudley's theorem (see Lemma 9.1) there is a coupling of $(\widetilde{Y}_1,\dots,\widetilde{Y}_b)$ and $(Z_1,\dots,Z_b)$ such that

$$P(\|(\widetilde{Y}_1,\dots,\widetilde{Y}_{[n^\alpha]})-(Z_1,\dots,Z_{[n^\alpha]})\|_\infty \ge \varDelta)\le \varDelta.$$

Let $S=\widetilde{Y}_1+\dots+\widetilde{Y}_{[n^\alpha]}-(Z_1+\dots+Z_{[n^\alpha]})$. Taking into account that $\|V_i\|_{L^{2+2\delta}}\le c$, we have

$$\|S\|_{L^{2+2\delta}} = \|\widetilde{Y}_1+\dots+\widetilde{Y}_{[n^\alpha]}-(Z_1+\dots+Z_{[n^\alpha]})\|_{L^{2+2\delta}} \le cn^\alpha a\Big(1+\max_{l\ge 1}\|X_l\|_{L^{2+2\delta}}\Big)\le c'n(1+\mu_\delta). \tag{8.15}$$

Let $\eta\in(0,\delta-\delta')$, $p=2+2\delta'$, $p'=p+2\eta\le 2+2\delta$ and $\gamma=\gamma(\eta)=\frac{2\eta}{p(p+2\eta)}$. By Hölder's inequality,

$$\|S^{2+2\delta'}1(|S|\ge n^\alpha\varDelta)\|_{L^{2+2\delta'}} \le \|S\|_{L^{p'}}(\mathbb{P}(|S|\ge n^\alpha\varDelta))^\gamma \le \|S\|_{L^{2+2\delta}}(\mathbb{P}(|S|\ge n^\alpha\varDelta))^\gamma.$$

Using the bound $|S|\le n^\alpha\|(\widetilde{Y}_1,\dots,\widetilde{Y}_{[n^\alpha]})-(Z_1,\dots,Z_{[n^\alpha]})\|_\infty$, we have

$$\begin{aligned}\|S\|_{L^{2+2\delta'}} &\le n^\alpha\varDelta+\|S^{2+2\delta'}1(|S|\ge n^\alpha\varDelta)\|_{L^{2+2\delta'}}\\ &\le n^\alpha\varDelta+\|S\|_{L^{2+2\delta}}(\mathbb{P}(|S|\ge n^\alpha\varDelta))^\gamma\\ &\le n^\alpha\varDelta+\|S\|_{L^{2+2\delta}}(\mathbb{P}(\|(\widetilde{Y}_1,\dots,\widetilde{Y}_{[n^\alpha]})-(Z_1,\dots,Z_{[n^\alpha]})\|_\infty\ge\varDelta))^\gamma\\ &\le n^\alpha\varDelta+cn(1+\mu_\delta)\varDelta^\gamma.\end{aligned}$$

Taking into account the definition of $\varDelta$, we get

$$\begin{aligned}\|S\|_{L^{2+2\delta'}} &\le n^\alpha c_{\alpha,\lambda_1}e^{-\frac{\lambda_1}{4}n^{\alpha+\rho}}(1+\lambda_0+\mu_\delta)\\ &\quad + cn(1+\mu_\delta)(c_{\alpha,\lambda_1}e^{-\frac{\lambda_1}{4}n^{\alpha+\rho}}(1+\lambda_0+\mu_\delta))^\gamma\\ &\le A'=c'_{\lambda_1,\lambda_2,\gamma,\alpha,\rho}(1+\lambda_0+\mu_\delta)^{1+\gamma}.\end{aligned} \tag{8.16}$$

From (8.15) and (8.16), it follows that

$$\begin{aligned}\|\widetilde{Y}_1+\dots+\widetilde{Y}_{[n^\alpha]}\|_{L^{2+2\delta'}} &\le \|S\|_{L^{2+2\delta'}}+\|Z_1+\dots+Z_{[n^\alpha]}\|_{L^{2+2\delta'}}\\ &\le A'+\|Z_1+\dots+Z_{[n^\alpha]}\|_{L^{2+2\delta'}}.\end{aligned} \tag{8.17}$$

Since the r.v.'s $Z_1,\dots,Z_{[n^\alpha]}$ are independent, by Rosenthal's inequality (Theorem 3 in [30]), there exists some constant $c_{\delta'}$ such that

$$\|Z_1+\dots+Z_{[n^\alpha]}\|_{L^{2+2\delta'}} \le c_{\delta'}\Big(\sum_{i=1}^{[n^\alpha]}\mathbb{E}Z_i^2\Big)^{1/2}+c_{\delta'}\Big(\sum_{i=1}^{[n^\alpha]}\mathbb{E}|Z_i|^{2+2\delta'}\Big)^{\frac{1}{2+2\delta'}}. \tag{8.18}$$

Taking into account that $Y_i = \sum_{j\in I_i} X_j$ and that $|I_i| \le a-b \le n^{1-\alpha}$, by Proposition 8.3 we have

$$\mathbf{E}Z_i^2 = \|\widetilde{Y}_i\|_{L^2}^2 \le (c+\|Y_i\|_{L^2})^2 \le c'_{\lambda_1,\lambda_2,\gamma}(1+\lambda_0+\mu_\delta)^{2+\gamma}n^{1-\alpha}. \tag{8.19}$$

Note also that $\|Z_i\|_{L^{2+2\delta'}} \le v_{a-b}+c$, where $v_n = \sup_{m\ge 1}\|\sum_{i=m}^{m+n-1} X_i\|_{L^{2+2\delta'}}$. Therefore, from (8.18) and (8.19), it follows that

$$\begin{aligned}
&\|Z_1+\cdots+Z_b\|_{L^{2+2\delta'}}\\
&\quad\le A' + c_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}(1+\lambda_0+\mu_\delta)^{1+\gamma/2}n^{1/2} + c_{\delta'}\Big(\sum_{i=1}^{[n^\alpha]}(v_{a-b}+c)^{2+2\delta'}\Big)^{\frac{1}{2+2\delta'}}\\
&\quad\le c_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}(1+\lambda_0+\mu_\delta)^{1+\gamma}n^{1/2} + c_{\delta'}v_{a-b}n^{\frac{\alpha}{2+2\delta'}}.
\end{aligned}$$

Using (8.17), we get

$$\begin{aligned}
\|\widetilde{Y}_1+\cdots+\widetilde{Y}_{[n^\alpha]}\|_{L^{2+2\delta'}} &\le A' + \|Z_1+\cdots+Z_{[n^\alpha]}\|_{L^{2+2\delta'}}\\
&\le c'_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}(1+\lambda_0+\mu_\delta)^{1+\gamma}n^{1/2} + c_{\delta'}v_{a-b}n^{\frac{\alpha}{2+2\delta'}}.
\end{aligned}$$

Since $\widetilde{Y}_k = Y_k + V_k$ and $\|V_k\|_{L^{2+2\delta}} \le c$, we see that

$$\begin{aligned}
&\|Y_1+\cdots+Y_{[n^\alpha]}\|_{L^{2+2\delta'}}\\
&\qquad\le c[n^\alpha] + \|\widetilde{Y}_1+\cdots+\widetilde{Y}_{[n^\alpha]}\|_{L^{2+2\delta'}}\\
&\qquad\le cn^\alpha + c_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}(1+\lambda_0+\mu_\delta)^{1+\gamma}n^{1/2} + c_{\delta'}v_{a-b}n^{\frac{\alpha}{2+2\delta'}}\\
&\qquad\le c_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}(1+\lambda_0+\mu_\delta)^{1+\gamma}n^{1/2} + c_{\delta'}v_{a-b}n^{\frac{\alpha}{2+2\delta'}},
\end{aligned}$$

where for the last line we use the fact that $\alpha < \frac{1-\rho}{2} < \frac{1}{2}$. Filling up the gaps in the final interval I_{fin}, we get

$$\begin{aligned}
\Big\|\sum_{i=1}^{m+n-1} X_i\Big\|_{L^{2+2\delta'}} &\le \|Y_1+\cdots+Y_{[n^\alpha]}\|_{L^{2+2\delta'}}\\
&\qquad+\sum_{k=1}^{[n^\alpha]}\sum_{i\in J_k}\|X_i\|_{L^{2+2\delta'}} + \Big\|\sum_{i=m+a[n^\alpha]}^{m+n} X_i\Big\|_{L^{2+2\delta'}}\\
&\le c_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}(1+\lambda_0+\mu_\delta)^{1+\gamma}n^{1/2} + c_{\delta'}v_{a-b}n^{\frac{\alpha}{2+2\delta'}}\\
&\qquad+n^{2\alpha+\rho}\mu_\delta + v_{n-[n^{1-\alpha}][n^\alpha]}.
\end{aligned}$$

From this, we deduce the inequality

$$\begin{aligned}
v_n &\le c_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}(1+\lambda_0+\mu_\delta)^{1+\gamma}n^{1/2}\\
&\qquad+n^{2\alpha+\rho}\mu_\delta + c_{\delta'}v_{[n^{1-\alpha}]-[n^{\alpha+\rho}]}n^{\frac{\alpha}{2+2\delta'}} + v_{n-[n^{1-\alpha}][n^\alpha]}.
\end{aligned}\tag{8.20}$$

Denote $\overline{v}_n = \frac{v_n}{(1+\lambda_0+\mu_\delta)^{1+\gamma}}$. Then from (8.20), it follows that

$$\overline{v}_n \le c_{\lambda_1,\lambda_2,\gamma,\alpha,\rho,\delta'}n^{1/2} + n^{2\alpha+\rho} + c_{\delta'}\overline{v}_{[n^{1-\alpha}]-[n^{\alpha+\rho}]}n^{\frac{\alpha}{2+2\delta'}} + \overline{v}_{n-[n^{1-\alpha}][n^\alpha]}.$$

Fixing $\alpha = 1/6$ and $\rho = 1/6$, we get

$$\overline{v}_n \leq c_{\lambda_1,\lambda_2,\gamma,\delta'} n^{1/2} + c_{\delta'} \overline{v}_{[n^{5/6}]-[n^{1/3}]} n^{\frac{1}{6}\frac{1}{2+2\delta'}} + \overline{v}_{n-[n^{5/6}][n^{1/6}]}. \tag{8.21}$$

We start with the inequality $\overline{v}_n \leq n^{q_0}$, where $q_0 = 1$. Since $n - [n^{1-\alpha}][n^\alpha] \leq cn^{1-\alpha}$, we have $v_{n-[n^{1-\alpha}][n^\alpha]} \leq cn^{\frac{5}{6}q_0}$ and $v_{[n^{1-\alpha}]-[n^{\alpha+\rho}]} \leq cn^{\frac{5}{6}q_0}$. Implementing this in (8.21) gives,

$$\overline{v}_n \leq c_{\lambda_1,\lambda_2,\gamma,\delta'} n^{1/2} + c_{\delta'} n^{\frac{5}{6}q_0 + \frac{1}{6}\frac{1}{2+2\delta'}} \leq c_{1r_1,\gamma,\delta'} n^{\max\{\frac{1}{2},q_1\}},$$

where $q_1 = \frac{5}{6}q_0 + \frac{1}{6}\frac{1}{2+2\delta'}$. Continuing in the same way, at iteration $k+1$, we obtain

$$\overline{v}_n \leq c_{kr_1,\gamma,\delta'} n^{\max\{1/2,q_{k+1}\}},$$

where $q_{k+1} = \frac{5}{6}q_k + \frac{1}{6}\frac{1}{2+2\delta'}$. Since $\lim_{k\to\infty} q_k = \frac{1}{2+2\delta'}$, there exists a constant $k_0 < \infty$ such that $q_{k_0+1} \leq 1/2$. With this k_0, we get

$$\overline{v}_n \leq c_{k_0\lambda_1,\lambda_2,\gamma,\delta'} n^{1/2}.$$

Since $\gamma = \gamma(\eta) = \frac{2\eta}{p(p+2\eta)} \leq \frac{2\eta}{p^2}$, for any $m \geq 1$ we have

$$\begin{aligned}\Big\| \sum_{i=m}^{m+n} X_i \Big\|_{L^{2+2\delta'}} &\leq c_{\lambda_1,\lambda_2,\eta,\delta'} (1 + \lambda_0 + \mu_\delta)^{1+\gamma} n^{1/2} \\ &\leq c_{\lambda_1,\lambda_2,\delta',\eta} (1 + \lambda_0 + \mu_\delta)^{1+\frac{2\eta}{p^2}} n^{1/2}.\end{aligned}$$

Since η is arbitrary we obtain the assertion of Proposition 8.2.

9. Appendix

9.1. Some general bounds for the Prokhorov distance. Let (E,d) be a metric space endowed with the metric d, and $\mathcal{E}$ be the Borel σ-algebra on E. For any $B \in \mathcal{E}$ denote by B^ε its ε-extension: $B^\varepsilon = \{x \in E : d(x,E) \leq \varepsilon\}$. Let $\pi(\mathbf{P},\mathbf{Q})$ be the Prokhorov distance between two probability measures $\mathbf{P}$ and $\mathbf{Q}$ defined by

$$\pi(\mathbf{P},\mathbf{Q}) = \inf\Big\{\varepsilon : \sup_{B\in\mathcal{E}} |\mathbf{P}(B) - \mathbf{Q}(B^\varepsilon)| \leq \varepsilon\Big\}.$$

The following assertion is known as the Strassen–Dudley theorem and is a consequence of the results in Strassen [36] (see also Dudley [8]). Let $\mathcal{P}_E(\mathbf{P},\mathbf{Q})$ be the set of probability measures on $E \times E$ with given marginals $\mathbf{P}$ and $\mathbf{Q}$. Denote by $\mathcal{D}_{E,d}(\varepsilon)$ the ε-extension of the diagonal in $E \times E$, i.e. $\mathcal{D}_{E,d}(\varepsilon) = \{(s,s') \in E \times E : d(s,s') \leq \varepsilon\}$, and by $\overline{\mathcal{D}}_{E,d}(\varepsilon)$ its complement.

Lemma 9.1 (Strassen–Dudley). *If (E,d) is a complete separable metric space, then*

$$\pi(\mathbf{P},\mathbf{Q}) = \min\{\varepsilon : \exists \mathbb{P} \in \mathcal{P}_E(\mathbf{P},\mathbf{Q}) \text{ such that } \mathbb{P}(\overline{\mathcal{D}}_{E,d}(\varepsilon)) \leq \varepsilon\}.$$

Let (E_1, d_1) and (E_2, d_2) be two complete separable metric spaces endowed with Borel σ-algebras $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively. Endow the product space $E = E_1 \times E_2$ with the metric $d(x, y) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$, where $x = (x_1, x_2)$, $y = (y_1, y_2) \in E$. Let $\mathcal{E}$ be the Borel σ-algebra on E.

LEMMA 9.2. *Consider r.v.'s $X, Y \in E_1$ and $Z \in E_2$ such that Z and (X, Y) are independent. Then*

$$\pi(\mathcal{L}_{X,Z}, \mathcal{L}_{Y,Z}) = \pi(\mathcal{L}_X, \mathcal{L}_Y).$$

Proof. Let $\mathbb{P}_1 \in \mathcal{P}_{E_1}(\mathcal{L}_X, \mathcal{L}_Y)$ and $\mathbb{P}_2 \in \mathcal{P}_{E_2}(\mathcal{L}_Z, \mathcal{L}_Z)$. If $\mathbb{P}_2$ is concentrated on the diagonal of $E_2 \times E_2$, then with $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$ we have $\mathbb{P}(\overline{\mathcal{D}}_{E,d}(\varepsilon)) = \mathbb{P}_1(\overline{\mathcal{D}}_{E_1,d_1}(\varepsilon))$. This means that

$$\begin{aligned} A &= \{\varepsilon : \exists \mathbb{P}_1 \in \mathcal{P}_E(\mathcal{L}_X, \mathcal{L}_Y) \text{ such that } \mathbb{P}_1(\overline{\mathcal{D}}_{E_1,d_1}(\varepsilon)) \le \varepsilon\} \\ &= \{\varepsilon : \exists \mathbb{P} \in \mathcal{P}_E(\mathcal{L}_{X,Z}, \mathcal{L}_{Y,Z}) \text{ such that } \mathbb{P}(\overline{\mathcal{D}}_{E,d}(\varepsilon)) \le \varepsilon\} = B. \end{aligned}$$

By Lemma 9.1, $\pi(\mathcal{L}_X, \mathcal{L}_Y) = \inf A = \inf B = \pi(\mathcal{L}_{X,Z}, \mathcal{L}_{Y,Z})$. ∎

Let $(E_1, d_1), \dots, (E_n, d_n)$ be complete separable metric spaces. On $E = E_1 \times \dots \times E_n$ consider the metric $d(x, y) = \max\{d_1(x_1, y_1), \dots, d_n(x_n, y_n)\}$, where $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in E$.

LEMMA 9.3. *Consider r.v.'s $X = (X_1, \dots, X_n) \in E$ and $Y = (Y_1, \dots, Y_n) \in E$. If X and Y are independent and $Y_1, \dots, Y_n$ are independent, then*

$$\pi(\mathcal{L}_{X_1,\dots,X_n}, \mathcal{L}_{Y_1,\dots,Y_n}) \le \sum_{k=1}^{n} \pi(\mathcal{L}_{X_1,\dots,X_{k-1},X_k}, \mathcal{L}_{X_1,\dots,X_{k-1},Y_k}).$$

Proof. The assertion of the lemma is obtained using the telescope rule and Lemma 9.2. ∎

LEMMA 9.4. *Consider r.v.'s $X = (X_1, \dots, X_n) \in E$ and $Y = (Y_1, \dots, Y_n) \in E$. If $(X_1, Y_1) \dots, (X_n, Y_n)$ are independent, then*

$$\pi(\mathcal{L}_{X_1,\dots,X_n}, \mathcal{L}_{Y_1,\dots,Y_n}) \le \sum_{k=1}^{n} \pi(\mathcal{L}_{X_k}, \mathcal{L}_{Y_k}).$$

Proof. Use Lemmas 9.3 and 9.2. ∎

The following is taken from [15].

LEMMA 9.5. *Let $\mathbf{P}$ and $\mathbf{Q}$ be two probability measures on $(\mathbb{R}^N, \mathcal{B}^N)$. Assume that the characteristic functions $\widehat{p}(t)$ and $\widehat{q}(t)$ pertaining to $\mathbf{P}$ and $\mathbf{Q}$ are square integrable with respect to the Lebesgue measure in $\mathbb{R}^N$. Then*

$$\pi(\mathbf{P}, \mathbf{Q}) \le (T/\pi)^{N/2} \Big(\int_{\mathbb{R}^N} |\widehat{p}(t) - \widehat{q}(t)|^2 \, dt \Big)^{1/2} + \mathbf{P}(\|x\|_\infty > T). \tag{9.1}$$

Proof. Assume first that $\mathbf{P}$, $\mathbf{Q}$ have square integrable densities p and q respectively. Denote $C_T = \{x \in \mathbb{R}^N : \|x\|_\infty \le T\}$ and $B_T = \mathbb{R}^N \setminus C_T$. Assume that $A \in \mathcal{B}^N$ and let $\varepsilon > 0$. Then

$$\begin{aligned}|\mathbf{P}(A) - \mathbf{Q}(A^\varepsilon)| &= |\mathbf{P}(A^\varepsilon \cap C_T) + \mathbf{P}(A^\varepsilon \cap B_T) - \mathbf{Q}(A^\varepsilon)| \\ &\le |\mathbf{P}(A^\varepsilon \cap C_T) - \mathbf{Q}(A^\varepsilon \cap C_T)| + \mathbf{P}(B_T) \\ &= \Big| \int\limits_{A^\varepsilon \cap C_T} (p(x) - q(x))\, dx \Big| + \mathbf{P}(B_T) \\ &\le \int\limits_{\mathbb{R}^N} |p(x) - q(x)| 1_{C_T}(x)\, dx + \mathbf{P}(\|x\|_\infty > T).\end{aligned}$$

Using Hölder's inequality, we get

$$\begin{aligned}\pi(\mathbf{P}, \mathbf{Q}) &\le |\mathbf{P}(A) - \mathbf{Q}(A^\varepsilon)| \\ &\le \Big(\int\limits_{\mathbb{R}^N} |p(x) - q(x)|^2\, dx \Big)^{1/2} \Big(\int\limits_{\mathbb{R}^N} 1_{C_T}(x)\, dx \Big)^{1/2} + \mathbf{P}(\|x\|_\infty > T).\end{aligned}$$

Since, by Plancherel's identity

$$\int\limits_{\mathbb{R}^N} |p(x) - q(x)|^2\, dx = (2\pi)^{-N} \int\limits_{\mathbb{R}^N} |\widehat{p}(t) - \widehat{q}(t)|^2\, dt,$$

we obtain (9.1) for $\mathbf{P}$ and $\mathbf{Q}$ having square integrable densities.

If $\mathbf{P}$ and $\mathbf{Q}$ do not have densities, denote by $\mathbf{P}_v = \mathbf{P} * \mathbf{G}_v$ and $\mathbf{Q}_v = \mathbf{Q} * \mathbf{G}_v$ the smoothed versions of $\mathbf{P}$ and $\mathbf{Q}$, where $\mathbf{G}_v$ is the normal distribution of mean 0 and variance v^2. Using (9.1) and the obvious inequality $|\widehat{p}_v(t) - \widehat{q}_v(t)| \le |\widehat{p}(t) - \widehat{q}(t)|$, we obtain

$$\pi(\mathbf{P}_v, \mathbf{Q}_v) \le (2\pi)^{-N/2} (2T)^{N/2} \Big(\int\limits_{\mathbb{R}^N} |\widehat{p}(t) - \widehat{q}(t)|^2\, dt \Big)^{1/2} + \mathbf{P}_v(\|x\|_\infty > T).$$

Since $\pi(\mathbf{P}_v, \mathbf{P}) \to 0$ and $\pi(\mathbf{Q}_v, \mathbf{Q}) \to 0$ it follows that $\pi(\mathbf{P}_v, \mathbf{Q}_v) \to \pi(\mathbf{P}, \mathbf{Q})$ as $v \to 0$. Note also that $\limsup_{v \to 0} \mathbf{P}_v(\|x\|_\infty > T) \le \mathbf{P}(\|x\|_\infty \ge T)$. Inequality (9.1) follows for arbitrary $\mathbf{P}, \mathbf{Q}$. ∎

9.2. Coupling independent and Gaussian r.v.'s. The following result is proved in Theorem 5 of Sakhanenko [32] (see also [31], [33], [34] for related results). Let $X_1, \dots, X_n$ be a sequence of independent r.v.'s satisfying $\mathbb{E} X_i = 0$ and $\mathbb{E}|X_i|^p < \infty$ for some $p \ge 2$ and all $1 \le i \le n$.

THEOREM 9.6. *On some probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ there is a sequence of independent normal r.v.'s $Y_1, \dots, Y_n$ satisfying $\mathbb{E}' Y_i = 0$ and $\mathbb{E}' Y_i^2 = \mathbb{E} X_i^2$, $1 \le i \le n$, and a sequence of independent r.v.'s $X_1', \dots, X_n'$ satisfying $X_i' \overset{d}{=} X_i$, $1 \le i \le n$, such that*

$$\mathbb{E}' \Big(\max_{1 \le k \le n} \Big| \sum_{i=1}^k X_i' - \sum_{i=1}^k Y_i \Big| \Big)^p \le c_p \sum_{i=1}^n \mathbb{E}|X_i|^p.$$

In particular, by Chebyshev's inequality, for the same construction and any $a > 0$,

$$\mathbb{P}'\Big(\max_{1\le k\le n}\Big|\sum_{i=1}^{k} X_i' - \sum_{i=1}^{k} Y_i\Big| > a\Big) \le \frac{c_p}{a^p}\sum_{i=1}^{n}\mathbb{E}|X_i|^p. \tag{9.2}$$

Acknowledgements. We wish to thank the referee for many useful remarks and comments; his suggestions have really improved the structure of this manuscript.

REFERENCES

[1] I. Berkes and W. Philipp, *Approximation theorems for independent and weakly dependent random vectors*, Ann. Probab. 7 (1979), 29–54.

[2] S. Bernstein, *Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes*, Math. Ann. 97 (1927), 1–59.

[3] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York 1968.

[4] A. A. Borovkov, *On the rate of convergence for the invariance principle*, Teor. Veroyatn. Primen. 18 (1973), 217–234 (in Russian).

[5] C. Cuny, *Pointwise ergodic theorems with rate with applications to limit theorems for stationary processes*, Stoch. Dynam. 11 (2011), 135–155.

[6] J. Dedecker, P. Doukhan and F. Merlevède, *Rates of convergence in the strong invariance principle under projective criteria*, Electron. J. Probab. 17 (2012), no. 16, 31 pp.

[7] P. Doukhan, J. R. Leon et F. Portal, *Principes d'invariance faible pour la mesure empirique d'une suite de variables aléatoires mélangeante*, Probab. Theory Related Fields 76 (1987), 51–70.

[8] R. M. Dudley, *Distances of probability measures and random variables*, Ann. Math. Statist. 39 (1968), 1563–1572.

[9] N. Dunford and J. T. Schwartz, *Linear Operators*, Interscience, New York, 1958.

[10] U. Einmahl, *Extensions of results of Komlós, Major and Tusnády to the multivariate case*, J. Multivariate Anal. 28 (1989), 20–68.

[11] I. G. Grama, *The rate of convergence in the functional central limit theorem for semimartingales*, in: Statistics and Control of Stochastic Processes, Vol. 2, Papers Steklov Sem. 1985–1986, A. N. Shiryaev, N. V. Krylov et al. (eds.), Optimization Software, Publ. Div., New York, 1989, 141–156.

[12] I. G. Grama and M. H. Neumann, *Asymptotic equivalence of nonparametric autoregression and nonparametric regression*, Ann. Statist. 34 (2006), 1701–1732.

[13] I. Grama and M. Nussbaum, *Asymptotic equivalence for nonparametric generalized linear models*, Probab. Theory Related Fields 111 (1998), 167–214.

[14] I. Grama and M. Nussbaum, *Asymptotic equivalence for nonparametric regression*, Math. Methods Statist. 11 (2002), 1–36.

[15] S. Gouëzel, *Almost sure invariance principle for dynamical systems by spectral methods*, Ann. Probab. 38 (2010), 1639–1671.

[16] Y. Guivarc'h, *Application d'un théorème limite local à la transcience et à la récurrence de marches de Markov*, in: Lecture Notes in Math. 1096, Springer, Berlin, 1984, 301–332.

[17] Y. Guivarc'h and É. Le Page, *On spectral properties of a family of transfer operators and convergence to stable laws for affine random walks*, Ergodic Theory Dynam. Systems 28 (2008), 423–446.

[18] Y. Guivarc'h and É. Le Page, *Simplicity of the Lyapunov spectrum and spectral gap property for a family of transfer operators on projective space*, in: Random Walks and Geometry, Gruyter, Berlin, 2004, 181–259.

[19] Y. Guivarc'h and É. Le Page, *Spectral gap properties and asymptotics of stationary measures for affine random walks*, arXiv:1204.6004 (2013).

[20] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, New York, 1980.

[21] G. Keller and C. Liverani, *Stability of the spectrum for transfer operators*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 28 (1999), 141–152.

[22] J. Komlós, P. Major and G. Tusnády, *An approximation of partial sums of independent RV's and the sample Df. II*, Z. Wahrsch. Verw. Gebiete 34 (1976), 33–58.

[23] K. Kubilius, *Rate of convergence in the invariance principle for martingale difference arrays*, Lithuanian Math. J. 34 (1994), 383–392.

[24] W. Liu and Z. Lin, *Strong approximation for a class of stationary processes*, Stoch. Process. Appl. 119 (2009), 249–280.

[25] F. Merlevède and E. Rio, *Strong approximation of partial sums under dependence conditions with application to dynamical systems*, Stoch. Process. Appl. 122 (2012), 386–417.

[26] S. V. Nagaev, *Some limit theorems for stationary Markov chains*, Teor. Veroyatn. Primen. 2 (1957), 389–416 (in Russian).

[27] S. V. Nagaev, *More exact statements of limit theorems for homogeneous Markov chains*, Theory Probab. Appl. 6 (1961), 62–81; transl. of: Teor. Veroyatn. Primen. 6 (1961), 67–86.

[28] Yu. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory Probab. Appl. 1 (1956), 157–214; transl. of: Teor. Veroyatn. Primen. 1 (1956), 177–237.

[29] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants*, Math. Appl. (Berlin) 31, Springer, Berlin, 2000.

[30] H. P. Rosenthal, *On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables*, Israel J. Math. 8 (1957), 273–303.

[31] A. I. Sakhanenko, *Convergence rate in the invariance principle for non-identically distributed random variables with exponential moments*, in: Limit Theorems for Sums of Random Variables, S. L. Sobolev (ed), Trudy Inst. Mat. 3, Nauka Sibirsk. Otdel. Novosibirsk, 1984, 3–49 (in Russian); English transl.: Advances in Probability Theory: Limit Theorems for Sums of Random Variables, A. A. Borovkov (ed.), Transl. Ser. Math. Engrg., Optimization Software Publ. Div., New York, 1986, 2–73.

[32] A. Sakhanenko, *Estimates in the invariance principle*, in: Limit Theorems of Probability Theory, A. A. Borovkov (ed.), Trudy Inst. Mat. 5, Nauka Sibirsk. Otdel., Novosibirsk, 1985, 27–44 (in Russian).

[33] A. Sakhanenko, *On the accuracy of normal approximation in the invariance principle*, Trudy Inst. Mat. (Novosibirsk) 13 (1989), 40–66 (in Russian); English transl.: Siberian Adv. Math. 1 (1991), no. 4, 58–91.

[34] A. I. Sakhanenko, *A new way to obtain estimates in the invariance principle*, in: High Dimensional Probability, II (Seattle, WA, 1999), Progr. Probab. 47, Birkhäuser, Boston, MA, 2000, 223–245.

[35] R. J. Serfling, *Moment inequalities for the maximum cumulative sums*, Ann. Math. Statist. 41 (1970), 1227–1234.

[36] V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist. 36 (1965), 423–439.

[37] W. B. Wu, *Strong invariance principles for dependent random variables*, Ann. Probab. 35 (2007), 2294–2320.

[38] A. Yu. Zaĭtsev, *Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments*, ESAIM Probab. Statist. 2 (1998), 41–108.

[39] A. Yu. Zaĭtsev, *Estimates for the rate of strong approximation in the multidimensional invariance principle*, J. Math. Sci. (N.Y.) 145 (2007), 4856–4865; transl. of: Zap. Nauchn. Sem. POMI 339 (2006), 37–53.

[40] O. Zhao and M. Woodroofe, *Law of the iterated logarithm for stationary processes*, Ann. Probab. 36 (2008), 127–142.

Ion Grama, Émile Le Page
Université de Bretagne Sud
LMBA CNRS 6205
Campus de Tohannic
BP 573
56017 Vannes Cedex, France
E-mail: ion.grama@univ-ubs.fr
emile.lepage@univ-ubs.fr

Marc Peigné
Université F. Rabelais Tours
LMPT CNRS 7350
Parc de Grandmont
37200 Tours Cedex, France
E-mail: peigne@lmpt.univ-tours.fr

Received 4 April 2013;
revised 29 July 2013 (5911)