An Empirical Process Central Limit Theorem for Multidimensional Dependent Data

Olivier Durieu*, Marco Tusche†

Let $(U_n(t))_{t\in\mathbb{R}^d}$ be the empirical process associated to an $\mathbb{R}^d$-valued stationary process $(X_i)_{i\geq 0}$. In the present paper, we introduce general conditions, which only involve processes $(f(X_i))_{i\geq 0}$ for a restricted class of functions $\mathcal{G}$, under which weak convergence of $(U_n(t))_{t\in\mathbb{R}^d}$ can be proved. This result improves those of [DDV09] and [DD11] and provides new applications.

The central interest in this approach is, that it does not need the indicator functions, which define the empirical process $(U_n(t))_{t\in\mathbb{R}^d}$, itself to belong to the class $\mathcal{G}$. This is particularly useful when dealing with data arising from dynamical systems or functionals of Markov chains.

In the proofs we make use of a new application of a chaining argument and generalize ideas first introduced in [DDV09] and [DD11].

Finally it will be shown, how the generalized conditions apply in the case of multiple mixing processes of polynomial decrease and causal functions of independent and identically distributed processes, which could not be treated by the preceding results in [DDV09] and [DD11].

1 Introduction

The present paper concerns the question of the weak convergence of empirical processes under weak dependence of the underlying process.

Let us consider some stationary process $(X_i)_{i\geq 0} = ((X_{i,1},\ldots,X_{i,d}))_{i\geq 0}$ in $\mathbb{R}^d$. The empirical process of $(X_i)_{i\geq 0}$ is the process $(U_n)_{n\geq 1}$ of random variables U_n given by

$$U_n(t_1,\ldots,t_d) := \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\mathbb{1}_{\prod_{j=1}^{d}[-\infty,t_j]}(X_{i,j}) - \mathbf{E}\mathbb{1}_{\prod_{j=1}^{d}[-\infty,t_j]}(X_{0,j})\Big),\quad t_1,\ldots,t_d\in[-\infty,\infty].$$

The study of the weak convergence of such a process began with Donsker [Don52], who proved the convergence of the empirical process to some Gaussian process in the case of i.i.d. $\mathbb{R}$-valued data. The result has been extend to sequences of weak dependent ($\mathbb{R}$ or $\mathbb{R}^d$-valued) random variables by many authors. Among others, it shall be remarked that Billingsley [Bil68]

*Laboratoire de Mathématiques et Physique Théorique (UMR CNRS 6083), Fédération Denis Poisson (FR CNRS 2964), Université de Tours, Parc de Grandmont, 37200 Tours, France

†Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstraße 150, 44780 Bochum, Germany

gave a result for functionals of φ-mixing sequences, Berkes and Philipp [BP78] under strong mixing assumptions, Doukhan, Massart and Rio [DMR95] for absolutely regular sequences and Borovkova, Burton and Dehling [BBD01] for functionals of absolutely regular processes, Doukhan and Louhichi [DL99], Dedecker and Prieur [DP07] for new dependence coefficients.

The technique of the proof usually consists of establishing the finite dimensional convergence of the process (i.e. the convergence of $(U_n(t_1), \ldots, U_n(t_k))$ for every choice of $(t_1, \ldots, t_k)$ and to use an appropriate moment bound to get the tightness of the process. See [DP02].

In the recent paper [DDV09], the authors introduced a new approach which is useful when the required properties on the underlying process are only established for a class of functions not containing the indicators. The required conditions are the Central Limit Theorem and a bound of the 4-th moment of partial sums for the process $(f(X_i))_{i\geq 0}$, when f belongs to the class of Lipschitz functions. Under these conditions, in the case of $\mathbb{R}$-valued data, they are able to prove an empirical CLT. Later, in [DD11], the technique was adapted to treat the case of $\mathbb{R}^d$-valued data. In that case bounds on higher moments are needed.

These required conditions can be satisfied in cases where the data arise from a dynamical system. In particular, for a large class of dynamical systems, they can be derived from the study of the so called Perron-Frobenius operator. More generally, they can be established in case where the system presents some multiple mixing properties (see Section 3.2 and [DD11] for more details).

The aim of the present paper is to give the general setting under which the technique of [DDV09] and [DD11] can be applied. This leads to extensions of the previous results. In particular, we provide general conditions on the class of functions, which does not necessarily contain the indicator functions, for which the CLT and the moment bound have to hold. We also propose a more general moment bound. This allows us, for example, to consider the case of systems with multiple mixing with polynomial rate which is not the case in [DD11].

2 Definitions and Statement of Main Results

Let $d \in \mathbb{N}^* := \{1, 2, \ldots\}$ be some fixed positive integer. In the following we will consider $[-\infty, \infty]^d$-indexed processes. For $a = (a_1, \ldots, a_d), b = (b_1, \ldots, b_d) \in [-\infty, \infty]^d$ write

$$a \leq b \quad \Leftrightarrow \quad a_i \leq b_i \ \ \forall i = 1, \ldots, d.$$

Note, that – although we will use the same denotation "$\leq$" – it will always be obvious from the context, whether the usual $\mathbb{R}$- or partial $[-\infty, \infty]^d$-order is meant. Analogously we define "$\geq$", "$<$" and "$>$" and abbreviate "not $\leq$" by "$\not\leq$" and so forth. For d-dimensional intervals we will write

$$[a, b) := \{x \in [-\infty, \infty]^d : a \leq x < b\}$$

and so on. Finally, for any $x \in [-\infty, \infty]$, let us denote $(x, \ldots, x) \in [-\infty, \infty]^d$ by $\overline{x}$.

Let $\mathbb{D}([-\infty, \infty]^d)$ be the space of generalized multidimensional càdlàg functions $[-\infty, \infty]^d \to \mathbb{R}$ (for definition see [Neu71, p.1286]), equipped with the multidimensional Skorohod metric d_0 as introduced in [Neu71, p.1289] (see also [Str72]). Note that $(\mathbb{D}([-\infty, \infty]^d), d_0)$ is a complete and separable space (more precisely, [Neu71] and [Str72] proved this for the space $\mathbb{D}([0, 1]^d)$, but – since $[0, 1]$ and $[-\infty, \infty]$ are homeomorphic – the metric on $\mathbb{D}([-\infty, \infty]^d)$ can naturally be extended to a metric on $\mathbb{D}([-\infty, \infty]^d)$ which conserves all relevant properties (c.f. [DD11, p.1081f.])).

Let $(X_i)_{i\in\mathbb{N}}$, $\mathbb{N} := \{0, 1, \ldots\}$, be some process of stationary $\mathbb{R}^d$-valued random variables. We formally define the empirical process of $(X_i)_{i\in\mathbb{N}}$ as the process $(U_n)_{n\in\mathbb{N}^*}$ of $\mathbb{D}([-\infty, \infty]^d)$-valued random variables $U_n := (U_n(t))_{t\in[-\infty,\infty]^d}$ given by

$$U_n(t) := \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\mathbb{1}_{[-\overline{\infty},t]}(X_i) - \mathbf{E}\mathbb{1}_{[-\overline{\infty},t]}(X_0)\Big) = \sqrt{n}\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{[-\overline{\infty},t]}(X_i) - \mathbf{F}(t)\Big),$$

where $\mathbf{F}(t) := \mathbf{P}(X_0 \le t)$ is the multi-dimensional distribution function of the X_i.

Our main result is the following

Theorem 2.1. *Let $(X_i)_{i\in\mathbb{N}}$ be a stationary process of $\mathbb{R}^d$ valued random vectors with continuous multi-dimensional distribution function $\mathbf{F}$ and marginal distribution functions $\mathbf{F}_1,\dots,\mathbf{F}_d$ (we define the i-th marginal distribution $\mathbf{F}_i$ as the functions $[-\infty,\infty]\to[0,1], t\mapsto \mathbf{P}(\pi_i X_0\le t)$, where π_i denotes the i-th coordinate projection $[-\infty,\infty]^d\to[-\infty,\infty]$, $(x_1,\dots,x_d)\mapsto x_i$).*

Assume there is a vector space $\mathcal{G}$ of measurable functions $\mathbb{R}^d\to\mathbb{R}$, equipped with a semi-norm $\|\cdot\|_{\mathcal{G}}$ satisfying the following conditions:

- *$\|1\|_{\mathcal{G}} < \infty$, i.e. constant functions are finite in norm.*
- *There is a function $\Psi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ which is increasing for sufficiently large arguments and for every $a < b \in [-\infty,\infty]^d$ there is a function $\varphi_{(a,b)} \in \mathcal{G}$ such that for any $x\in\mathbb{R}^d$*

$$\varphi_{(a,b)}(x) \begin{cases} = 1, & \text{if } x\le a,\\ \in[0,1], & \text{if } x<b \text{ and } x\not\le a,\\ = 0, & \text{if } x\not< b\end{cases} \tag{2.1}$$

and such that

$$\|\varphi_{(a,b)}\|_{\mathcal{G}} \le \Psi\left(\frac{1}{\min\limits_{i=1,\dots,d} w_{\mathbf{F}_i}(b_i-a_i)}\right). \tag{2.2}$$

where $w_{\mathbf{F}_i}$, $i=1,\dots,d$, denotes the modulus of continuity

$$w_{\mathbf{F}_i}(\delta) := \sup\Big\{|\mathbf{F}_i(t)-\mathbf{F}_i(s)| : s,t\in\mathbb{R},\ |t-s|\le\delta\Big\}.$$

- *There are constants $r, C, C' \ge 1$ and $p > dr$ such that for every $f\in\mathcal{G}$ with $\|f\|_\infty\le 1$ we have the moment bound*

$$\begin{aligned}&\mathbf{E}\bigg(\Big|\sum_{i=1}^{n} f(X_i) - \mathbf{E}(f(X_0))\Big|^{2p}\bigg)\\ &\le C\sum_{i=1}^{p} n^i\,\|f(X_0)-\mathbf{E}(f(X_0))\|_r^i\,\Phi_i(\|f-\mathbf{E}(f(X_0))\|_{\mathcal{G}}+C'),\end{aligned} \tag{2.3}$$

where $\Phi_1,\dots,\Phi_p : \mathbb{R}_0^+\to\mathbb{R}_0^+$ are some (for sufficiently large arguments) increasing functions, satisfying

$$\Phi_i(2\Psi(z)+C') \le C'' z^{\gamma_i} \quad \text{for all } i=1,\dots,p, \tag{2.4}$$

for all sufficiently large $z\in\mathbb{R}^+$, where $C'',\gamma_1,\dots,\gamma_p$ are some nonnegative and finite constants, fulfilling

$$\gamma_i < \frac{i}{r} + 2(p-i) - d \quad \text{for all } i=1,\dots,p. \tag{2.5}$$

- *For every $f \in \mathcal{G}$ such that $\|f\|_\infty < \infty$ the Central Limit Theorem holds, i.e.*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(f(X_i) - \mathbf{E} f(X_0)\big) \xrightarrow{\mathcal{D}} N(0, \sigma_f^2), \tag{2.6}$$

  as $n \to \infty$, where $N(0, \sigma_f^2)$ denotes some normal distributed random variable on $\mathbb{R}$ with expected value 0 and variance σ_f^2 depending on f.

Then there is a centered Gaussian process $W \in \mathbb{D}([-\infty, \infty]^d)$ with almost surely continuous sample paths such that

$$U_n \xrightarrow{\mathcal{D}} W.$$

Remark 2.2. Note that in condition (2.3) it is crucial that p, r, $\Phi_1, \ldots, \Phi_d$, C and C' are to be chosen independently of the choice of $f \in \mathcal{G}$.

Remark 2.3. We will see in the proof, that the theorem also holds if (2.3) is satisfied for a certain subclass of functions in $\mathcal{G}$, more precisely if (2.3) holds for all functions $f \in \mathcal{G}$ of the form

$$f := \varphi_{(a,b)} - \varphi_{(a',b')},$$

where $a, b, a', b' \in [-\infty, \infty]^d$, $a' < b$, such that there is a $h > 0$ satisfying

$$h = P(X_0 \in [a', b']) \le 2P(X_0 \in [a, b]) \le P(X_0 \in [a', b]) \le 3P(X_0 \in [a', b']). \tag{2.7}$$

Choosing for each $m \in \mathbb{N}^*$ an $f_m := \varphi_{(a,b)} - \varphi_{(a',b')}$ such that (2.7) is satisfied for $h = 1/m$, it can be shown, that

$$\|f_m(X_0)\|_r^i = \mathcal{O}\Big(\frac{m^{\gamma_i - \frac{i}{r}}}{\Phi_i(\|f_m\|_{\mathcal{G}})}\Big) \quad \text{as } m \to \infty.$$

The rest of the paper is organized as follows: In Section 3 we present some particular cases of the theorem and some applications. The proof of the theorem will be given in Section 4. Section 5 and 6, are devoted to the proofs of results of Section 3.

3 Applications

3.1 Assumptions on Hölder continuous functions

As Theorem 2.1 is stated in a very general context, this section is dedicated to the special case, where $\mathcal{G}$ is chosen as $\mathcal{H}_\alpha$, the class of bounded α-Hölder functions, i.e.

$$\mathcal{H}_\alpha := \{f \in \mathbf{L}_\infty(\mathbb{R}^d) : \exists K \in \mathbb{R}\ |f(x) - f(y)| \le K\|x - y\|^\alpha\ \forall x, y \in \mathbb{R}^d\},$$

equipped with the norm

$$\|f\|_{\mathcal{H}_\alpha} := \sup_{x \in \mathbb{R}^d} |f(x)| + \sup_{\substack{x,y \in \mathbb{R}^d \\ x \ne y}} \frac{|f(x) - f(y)|}{\|x - y\|^\alpha}.$$

This class of functions has already been considered in [DD11]. In this situation however, we will show that one can obtain some generalisations of the corresponding results. Notice that in [DD11], the functions Φ_i which play a role in the moment bound (2.3) are only $\Phi_i(x) = \log^{2p-i}(x + C)$. Here, we will consider other cases for the choice of the function Φ_i.

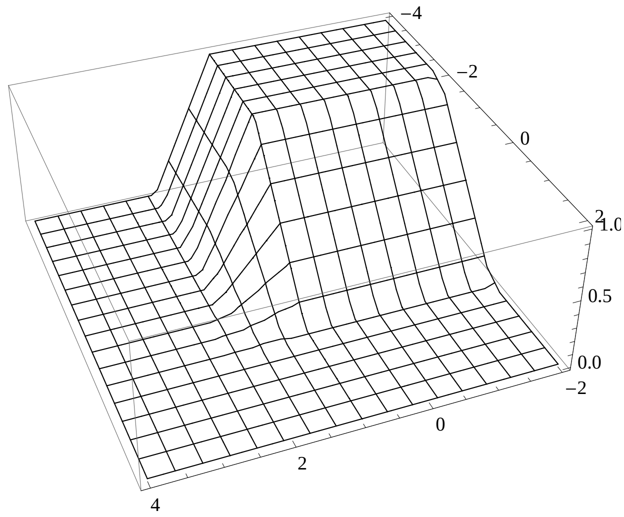

Figure 1: The function $\varphi_{(a,b)}$ for $a=(1,-2)$, $b=(2,0)$.

Let us see that the class $\mathcal{H}_\alpha$ satisty conditions (2.1) and (2.2). Choose $\varphi_{(a,b)} \in \mathcal{H}_\alpha$ as

$$\varphi_{(a,b)}\big((x_i)_{i=1,\dots,d}\big) := \varphi\Big(\Big(-\mathbb{1}_{\{-\infty\}}(b_i) + \mathbb{1}_{(-\infty,\infty)^2}(a_i,b_i)\cdot\frac{x_i-b_i}{b_i-a_i}\Big)_{i=1,\dots,d}\Big), \quad x_i \in \mathbb{R},$$

where $\varphi : \mathbb{R}^d \to \mathbb{R}$ is given by

$$\varphi\big((y_i)_{i=1,\dots,d}\big) := \prod_{i=1}^{d}\Big(\mathbb{1}_{[-\infty,-1]}(y_i) - y_i\mathbb{1}_{(-1,0]}(y_i)\Big)$$

Obviously this choice of $\varphi_{(a,b)}$ satisfies (2.1).

Let us now check condition (2.2). Since for all $j=1,\dots,d$ we have

$$\begin{aligned} b_j - a_j &\geq \inf\big\{\delta > 0 : w_{\mathbf{F}_j}(\delta) \geq \min_{i=1,\dots,d} w_{\mathbf{F}_i}(b_i-a_i)\big\} \\ &\geq \inf\big\{\delta > 0 : w_{\mathbf{F}}(\delta) \geq \min_{i=1,\dots,d} w_{\mathbf{F}_i}(b_i-a_i)\big\}, \end{aligned}$$

where

$$w_{\mathbf{F}}(\delta) := \sup\{|\mathbf{F}(t)-\mathbf{F}(s)| : s,t\in\mathbb{R}^d,\ \|t-s\|\leq\delta\}.$$

Thus by the definition of $\varphi_{(a,b)}$ we obtain

$$\begin{aligned} \|\varphi_{(a,b)}\|_{\mathcal{H}_\alpha} &\leq d\max_{j=1,\dots,d}\mathbb{1}_{(-\infty,\infty)^2}(a_j,b_j)\cdot\frac{1}{(b_j-a_j)^\alpha} + 1 \\ &\leq d\Big(\min_{i=1,\dots,d}\inf\Big\{\delta>0 : w_{\mathbf{F}}(\delta) \geq \min_{i=1,\dots,d} w_{\mathbf{F}_i}(b_i-a_i)\Big\}\Big)^{-\alpha} + 1 \\ &= d\Big(\inf\Big\{\delta>0 : w_{\mathbf{F}}(\delta) \geq \min_{i=1,\dots,d} w_{\mathbf{F}_i}(b_i-a_i)\Big\}\Big)^{-\alpha} + 1, \\ &= d\Big(w_{\mathbf{F}}^{\leftarrow}\Big(\min_{i=1,\dots,d} w_{\mathbf{F}_i}(b_i-a_i)\Big)\Big)^{-\alpha} + 1, \end{aligned}$$

where

$$w_{\mathbf{F}}^{\leftarrow}(y) := \inf\big\{\delta>0 : w_{\mathbf{F}}(\delta)\geq y\big\}.$$

and thereby condition (2.2) is satisfied for the increasing function Ψ given by

$$\Psi(z) := d\Big(w_{\mathbf{F}}^{\leftarrow}\Big(\frac{1}{z}\Big)\Big)^{-\alpha} + 1. \tag{3.1}$$

Furthermore condition (2.4) can be reduced to the existence of some C'' such that

$$\Phi_i\Big(2d\Big(w_{\mathbf{F}}^{\leftarrow}\Big(\frac{1}{z}\Big)\Big)^{-\alpha}+C'+2\Big)\leq C''z^{\gamma_i}\quad\forall i=1,\dots,p, \tag{3.2}$$

for suitable $\gamma_1,\dots,\gamma_p$ satisfying (2.4) for sufficiently large z.

Let us now assume that there is a monotone increasing and invertible function Φ with continuous inverse function Φ^{-1} such that there are some constants $\lambda,\lambda'>1$ for which

$$\Phi(\lambda z)\leq\lambda'\Phi(z) \tag{3.3}$$

holds for sufficiently large $z\in\mathbb{R}^+$. Assume furthermore that there are some $\gamma>1$ and $\beta>0$ such that

$$w_{\mathbf{F}}(y)\leq\beta(\Phi(y^{-\alpha}))^{-\gamma} \tag{3.4}$$

holds for all sufficiently small $y>0$ and set

$$\Phi_i:=\Phi^{\kappa_i}$$

for some $\kappa_1,\dots,\kappa_p>0$ such that

$$\kappa_i<\gamma\Big(\frac{i}{r}+2(p-i)-d\Big). \tag{3.5}$$

Obviously for large $z>0$ we have

$$w_{\mathbf{F}}^{\leftarrow}\Big(\frac{1}{z}\Big)\geq\inf\Big\{\vartheta>0:\beta(\Phi(\vartheta^{-\alpha}))^{-\gamma}\geq\frac{1}{z}\Big\}=\inf\Big\{\vartheta>0:\Phi(\vartheta^{-\alpha})\leq\Big(\beta z\Big)^{\frac{1}{\gamma}}\Big\}$$

and therefore, since Φ is invertible, we obtain

$$w_{\mathbf{F}}^{\leftarrow}\Big(\frac{1}{z}\Big)^{-\alpha}\geq\inf\Big\{\vartheta>0:\vartheta\geq\big(\Phi^{-1}\big((\beta z)^{\frac{1}{\gamma}}\big)\big)^{-\frac{1}{\alpha}}\Big\}^{-\alpha}=\Big(\Phi^{-1}\Big((\beta z)^{\frac{1}{\gamma}}\Big)\Big)$$

for sufficiently large z. Using condition (3.3), for any $i\in\{1,\dots,p\}$ this yields that there is a finite constant $C''\in\mathbb{R}$ such that

$$\begin{aligned}\Phi(2\Psi(z)+C')&=\Phi_i\Big(2d\Big(w_{\mathbf{F}}^{\leftarrow}\Big(\frac{1}{z}\Big)\Big)^{-\alpha}+C'+2\Big)\\&\leq\Big(\Phi(2d\Phi^{-1}((\beta z)^{\frac{1}{\gamma}})+C'+2)\Big)^{\kappa_i}\\&\leq C''\Big(\Phi(\Phi^{-1}((\beta z)^{\frac{1}{\gamma}}))\Big)^{\kappa_i}\\&=C''\beta^{\frac{\kappa_i}{\gamma}}\cdot z^{\frac{\kappa_i}{\gamma}}\end{aligned}$$

for all sufficiently large $z\in\mathbb{R}$, where each exponent κ_i/γ is positive and strictly smaller than $i/r+2(p-i)-d$. Thus condition (2.4) is necessarily satisfied.

Corollary 3.1 (Bounded α-Hölder functions)**.** *Assume that*

- *There are is an invertible increasing function $\Phi:\mathbb{R}_0^+\longrightarrow\mathbb{R}_0^+$, a $z_0\in\mathbb{R}^+$ and some $\lambda,\lambda'>1$ with*

$$\Phi(\lambda z)\leq\lambda'\Phi(z),\quad\forall z\geq z_0 \tag{3.6}$$

such that there are some constants $\beta>0$, $\gamma>1$ satisfying

$$w_{\mathbf{F}}(y)\leq\beta\Phi(y^{-\alpha})^{-\gamma} \tag{3.7}$$

for sufficiently small $y>0$.

- *There are some constants, $r \geq 1$, $p > dr$, $C, C' \geq 1$ and $\kappa_1, \ldots, \kappa_p > 0$ satisfying*

$$\kappa_i < \gamma\Big(\frac{i}{r} + 2(p-i) - d\Big), \quad i = 1, \ldots, p \tag{3.8}$$

such that for every $f \in \mathcal{H}_\alpha$ with $\|f\|_\infty \leq 1$ condition (2.3) holds with $\Phi_i = \Phi^{\kappa_i}$.

- *For every $f \in \mathcal{H}_\alpha$ condition (2.6) holds.*

Then there is a centered Gaussian process $W \in \mathbb{D}([-\infty, \infty]^d)$ with almost surely continuous paths such that

$$U_n \xrightarrow{\mathcal{D}} W.$$

Let us now give some particular cases:

Example 3.2. Choose $\kappa_i = i$. In this case (3.8) is equivalent to

$$\gamma > \frac{rp}{p - dr}. \tag{3.9}$$

Since we need p to be larger than dr, we may choose $p = dr + 1$ to keep the condition on the moment bounds (2.3) as weak as possible. For this choice of p inequality (3.8) can be simplified to

$$\gamma > dr^2 + r \tag{3.10}$$

Let $\Phi = \mathbf{id}$. Thus condition (3.6) is obviously satisfied and condition (3.7) becomes

$$w_{\mathbf{F}}(y) \leq \beta y^{\alpha\gamma} \tag{3.11}$$

for sufficiently small $y > 0$. Observe that $w_{\mathbf{F}}(y) \leq \beta y^{\alpha\gamma}$ implies that $\mathbf{F}$ is an $(\alpha\gamma)$-Hölder function, thus this property only makes sense if $\alpha \leq 1/\gamma$, i.e. $\alpha < (p - rd)/rp$.

Example 3.3. If we choose $\kappa_i = 2p - i$, inequality (3.8) from Corollary 3.1 also is equivalent to (3.9) since

$$\max_{i=1,\ldots,p} \frac{2p - i}{\frac{i}{r} + 2(p-i) - d} = \frac{rp}{p - dr}.$$

Now choose $\Phi = |\log|$. For this choice condition (3.7) can be simplified to

$$w_{\mathbf{F}}(y) \leq \beta |\log(y)|^{-\gamma}$$

for some $\beta > 0$ and sufficiently small $y > 0$. In contrast to Example 3.2, in this case α has no relation to γ and hence may be chosen arbitrary in $(0, 1]$. This application of Corollary 3.1 corresponds to Theorem 3 in [DD11, p.1081]

Example 3.4. Again, choose $\Phi = |\log|$ and let $\alpha, d, r = 1$ and $p = 2$. Now condition (3.8) in Corollary 3.1 reduces to

$$\gamma > \max\Big\{\frac{\kappa_1}{2}, \kappa_2\Big\}.$$

This result is equivalent to Theorem 1 in [DDV09, p.3702].

3.2 Multiple mixing property

Let us denote the m-th partial sum of a sequence $(a_j)_{j\in\mathbb{N}^*}$ by

$$a_m^* := \sum_{j=1}^{m} a_j.$$

For a process $(X_i)_{i\in\mathbb{N}}$ of $\mathbb{R}^d$-valued random variables and a class $\mathcal{F}$ of measurable functions $\mathbb{R}^d \to \mathbb{R}$, equipped with a semi-norm $\|\cdot\|_{\mathcal{F}}$, we will write $(X_n)_{n\in\mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{F})$ if there are a $r \in \mathbb{N}^*$ and a decreasing function $\Theta : \mathbb{N} \longrightarrow \mathbb{R}_0^+$ such that for any $p \in \mathbb{N}^*$ there is a constant $K_p < \infty$ such that the following multiple mixing property is satisfied:

$$\begin{aligned}&\left|\mathbf{Cov}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot\varphi(X_{i_{q-1}^*})\,,\,\varphi(X_{i_q^*})\varphi(X_{i_{q+1}^*})\cdot\ldots\cdot\varphi(X_{i_p^*})\big)\right|\\ &\leq K_p\|\varphi(X_0)\|_r\|\varphi\|_{\mathcal{F}}\Theta(i_q)\end{aligned} \tag{3.12}$$

for all $\varphi \in \mathcal{F}$ with $\|\varphi\|_\infty \leq 1$ and $\mathbf{E}(\varphi(X_0)) = 0$ all $i_1, \ldots, i_p \in \mathbb{N}$, $q \in \{1, \ldots, p\}$.

Proposition 3.5. *Let $(X_n)_{n\in\mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{G})$ for a function Θ satisfying*

$$\sum_{i=0}^{\infty} i^{2p-2}\Theta(i) < \infty. \tag{3.13}$$

Then there is a constant $C < \infty$ such that for every $\varphi \in \mathcal{G}$ with $\mathbf{E}(\varphi(X_0)) = 0$ and $\|\varphi\|_\infty \leq 1$ the following bound holds

$$\mathbf{E}\left(\Big|\sum_{i=1}^{n}\varphi(X_i)\Big|^{2p}\right) \leq C\sum_{i=1}^{p} n^i\,\|\varphi(X_0)\|_r^i\,\|\varphi\|_{\mathcal{G}}^i.$$

The proof will be given in Section 5.

Notice that in [DD11] the authors obtained a sharper bound of the $2p$-th moment by considering the case of multiple mixing with exponential rate of convergence (i.e. $\Theta(i)$ converges to zero exponentially fast). Nevertheless, the moment bound we have here is sufficient to apply Theorem 2.1 and thus this generalization allows us to treat some cases of processes which are multiple mixing with polynomial rate.

3.3 Causal functions of independent and identically distributed processes

One example of processes satisfying (3.13) are causal functions of i.i.d. processes, which are defined as follows.

Definition 3.6 (Causal Function of an i.i.d. Process)**.** Let $(\xi_j)_{j\in\mathbb{Z}}$ be an independent identically distributed process with values in a Banach space $(\mathcal{X}, |\cdot|_{\mathcal{X}})$. We call $(X_i)_{i\in\mathbb{N}}$ a causal function of $(\xi_j)_{j\in\mathbb{Z}}$ if there is a measurable function $F : \mathcal{X}^{\mathbb{N}} \to \mathbb{R}^d$ such that each X_i is of a form

$$X_i := F((\xi_{i-j})_{j\in\mathbb{N}}).$$

Observe that by definition a causal function of an i.i.d. process will always be stationary.

Let us now introduce a measure of the dependence structure of a causal function of an i.i.d. process $(\xi_j)_{j\in\mathbb{Z}}$: Set

$$\dot{X}_i := F(\xi_i, \xi_{i-1}, \ldots, \xi_1, \xi_0', \xi_{-1}', \ldots),$$

where $(\xi'_j)_{j\in\mathbb{Z}}$ is an independent copy of $(\xi_j)_{j\in\mathbb{Z}}$, i.e. $(\xi_j)_{j\in\mathbb{Z}}$ and $(\xi'_j)_{j\in\mathbb{Z}}$ are identically distributed and both processes are independent from each other. We can now define for $k, s \in \mathbb{N}^*$

$$\delta_{k,s} = \|X_k - \dot{X}_k\|_s := \mathbf{E}\Big(\|X_k - \dot{X}_k\|_{\mathbb{R}^d}^s\Big)^{\frac{1}{s}}.$$

Lemma 3.7. *If $(X_i)_{i\in\mathbb{N}}$ is a causal function of an i.i.d. process then $(X_i)_{i\in\mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{H}_\alpha)$, where $\Theta(i) = (\delta_{i,s})^\alpha$ for $r, s \in \mathbb{N}\cup\{\infty\}$ satisfying $\frac{1}{r} + \frac{1}{s} = 1$ (we agree to let $\frac{1}{\infty} := 0$).*

The proof is given in Section 6.

Since by Lemma 3.7 a causal function of an i.i.d. processes satisfying

$$\sum_{i=0}^{\infty} i^{2p-2}(\delta_{i,s})^\alpha < \infty \tag{3.14}$$

for some $s \in \mathbb{N}\cup\{\infty\}$ features the multiple mixing property (3.12), condition (2.3) is satisfied for $r \in \mathbb{N}\cup\{\infty\}$ such that $\frac{1}{r} + \frac{1}{s} = 1$.

The second crucial point in the assumptions of Theorem 2.1 and Corollary 3.1 is the Central Limit Theorem Condition (2.6). The following lemma gives a criterion for (2.6) in the situation of a causal function of an i.i.d. process.

Lemma 3.8. *If $(X_i)_{i\in\mathbb{N}}$ is a causal function of an i.i.d. process and satisfies*

$$\sum_{i=1}^{\infty} (\delta_{i,s})^\alpha < \infty, \tag{3.15}$$

for some $s \in \mathbb{N}^*$, then for every $f \in \mathcal{H}_\alpha$ such that $\mathbf{E}(f(X_0)) = 0$ and $\|f\|_\infty \leq 1$ the Central Limit Theorem (2.6) holds with

$$\sigma_f^2 = \mathbf{E}(f(X_0)^2) + 2\sum_{i=1}^{\infty} \mathbf{E}\big(f(X_0)f(X_i)\big).$$

The proof is given in Section 6.

Corollary 3.9. *Let $(X_i)_{i\in\mathbb{N}}$ be a causal function of an i.i.d. sequence. Assume, that*

- *the distribution function $\mathbf{F}_{X_0}$ of X_0 is ϑ-Hölder for some $\vartheta \in [0,1]$,*
- *there are some $s \in (1,\infty]$, $r \in [1,\infty)$ satisfying $\frac{1}{r} + \frac{1}{s} = 1$ and there exists a positive*

$$\alpha < \frac{\vartheta}{dr^2 + r} \tag{3.16}$$

  such that

$$\sum_{i=0}^{\infty} i^{2dr}(\delta_{i,s})^\alpha < \infty. \tag{3.17}$$

Then there is a centered Gaussian process $W \in \mathbb{D}([-\infty,\infty]^d)$ with almost surely continuous paths such that

$$U_n \xrightarrow{\mathcal{D}} W.$$

Proof. Let us check the assumptions in Corollary 3.1 in the situation of Example 3.2, i.e. $\kappa_i = i$, $\Phi = \mathbf{id}$ and $p = dr + 1$. As presented in the example, only (2.3), (2.6), (3.10) and (3.11) remain to be shown.

Setting $\gamma := \vartheta/\alpha$, condition (3.10) directly follows from inequality (3.16). Since $\mathbf{F}_{X_0}$ is ϑ-Hölder (3.11) is obviously satisfied.

Finally, using (3.17), the Central Limit Theorem condition (2.6) holds due to Lemma 3.8, while the $2p$-th moment bound condition (2.3) is a direct consequence of the Lemmata 3.5 and 3.7, where one uses that p was chosen as $p = dr + 1$ in the beginning of the proof. $\square$

Example 3.10. Let $(X_i)_{i\in\mathbb{N}}$ be a causal linear function given by

$$X_i := \sum_{j\in\mathbb{N}} a_j \xi_{i-j},$$

where $(\xi_j)_{j\in\mathbb{Z}}$ is an i.i.d. $\mathcal{X}$-valued process and $(a_j)_{j\in\mathbb{N}}$ is a family of linear operators $\mathcal{X} \to \mathbb{R}^d$. We denote the norm of such operators by $|a| = \sup\{\|a(x)\| : x \in \mathcal{X}, |x|_{\mathcal{X}} \leq 1\}$.

If $\|\xi_0\|_s < \infty$ for some $s > 1$, if the distribution function $\mathbf{F}_{X_0}$ is ϑ-Hölder and if

$$\sum_{j=i}^{\infty} |a_j|^2 = \mathcal{O}(i^{-b})$$

with

$$b > \frac{(2dr+1)(dr^2+r)}{\vartheta},$$

for $r = \frac{s}{s-1}$, then the Empirical Central Limit Theorem holds.

Proof. Again let $(\xi'_j)_{j\in\mathbb{Z}}$ be an independent copy of $(\xi_j)_{j\in\mathbb{Z}}$. By assumption, there is a $\varepsilon > 0$ such that $b \geq (1+\varepsilon)^2(2dr+1)(dr^2+r)/\vartheta$. Choose $\alpha = (1+\varepsilon)^{-1}\vartheta/(dr^2+r)$, ensuring that (3.16) is satisfied. Since,

$$\delta_{i,s} = \left\| \sum_{j=i}^{\infty} a_j(\xi_{i-j} - \xi'_{i-j}) \right\|_s \leq \|\xi_0 - \xi'_0\|_s \sum_{j=i}^{\infty} |a_j|,$$

we have

$$i^{2dr}(\delta_{i,s})^{\alpha} \leq (2\|\xi_0\|_s)^{\alpha} i^{2dr} (\sum_{j=i}^{\infty} |a_j|)^{\alpha} = \mathcal{O}(i^{2dr-\alpha b})$$

where $2dr - \alpha b < -1$. Thus (3.17) holds and Corollary 3.9 applies. $\square$

The example of causal linear processes has already been studied by several authors. Dedecker and Prieur [DP07] could allow lower rates of convergence for $\sum_{j=i}^{\infty} |a_j|$ but required that X_0 has a bounded density. Wu [Wu08] also consider that the underlying i.i.d. process as a density. Here no assumption is made on the distribution of the ξ_i except moments and the distribution of X_0 does not need to be absolutely continuous. In the case where ξ_i are $\mathbb{R}$-valued and $d = 1$, very weak conditions can be found in [Ded10].

4 Proof of Theorem 2.1

To prove the convergence of the empirical process we will use a modified theorem of Billingsley:

Proposition 4.1. *Let S be a separable and complete space with metric ρ and let $\mathcal{S}$ denote the Borel σ-algebra on S, given by the open sets in S.*

If $U_n, U_n^{(m)}$ and $U^{(m)}$, $n, m \in \mathbb{N}^*$, are S-valued random variables such that

$$U_n^{(m)} \xrightarrow{\mathcal{D}} U^{(m)} \text{ as } n \to \infty, \text{ for all } m \in \mathbb{N}^* \text{ and} \tag{4.1}$$

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbf{P}(\rho(U_n, U_n^{(m)}) > \varepsilon) = 0 \text{ for all } \varepsilon > 0, \tag{4.2}$$

then there is a S-valued random variable X such that

$$U_n \xrightarrow{\mathcal{D}} U \text{ as } n \to \infty$$
$$U^{(m)} \xrightarrow{\mathcal{D}} U \text{ as } m \to \infty.$$

For the proof of this proposition, see Section 3 in [DDV09, p.3713].

Remark 4.2. We will apply Proposition 4.1 to the situation where $S = \mathbb{D}([-\infty, \infty]^d)$ (equipped with the Skorohod metric d_0, c.f. p.2) and U_n denotes the empirical Process.

If one wants to prove the convergence in distribution of U_n Proposition 4.1 assures us, that it is sufficient to find a process $U_n^{(m)}$ which approximates U_n as $m \to \infty$ in the sense of (4.2) and to show that this process is convergent in distribution for each m as $n \to \infty$.

Moreover the proposition avouches that the limit distribution of U_n as $n \to \infty$ is equal the limit distribution of $U^{(m)}$ as $m \to \infty$ (where $U^{(m)}$ denotes the weak limit of $U_n^{(m)}$ as $n \to \infty$).

Following the techniques presented in [DD11, p.1078 ff.], we begin by introducing a partition for $[-\infty, \infty]^d$. Let $\mathbf{F}_i$ be the i-th marginal distribution of $\mathbf{F}$, $0 = r_0^{(m)} < r_1^{(m)} < \ldots < r_m = 1$ a partition of $[0, 1]$ and set for $i \in \mathbb{N}^*$, $j_i \in \{0, \ldots, m\}$

$$t_{i,j_i}^{(m)} := \mathbf{F}_i^{\rightarrow}(r_{j_i}^{(m)}),$$

where

$$\mathbf{F}_i^{\rightarrow}(y) := \sup\{x \in [-\infty, \infty] : \mathbf{F}_i(x) \leq y\}.$$

note that by the continuity of the $\mathbf{F}_i$ the functions $\mathbf{F}_i^{\rightarrow}$ must be injective. For convenience we also define $t_{i,m+1}^{(m)} := t_{i,m}^{(m)}$. For $j \in \{0, \ldots, m+1\}^d$ set

$$t_j^{(m)} := (t_{1,j_1}^{(m)}, \ldots, t_{d,j_d}^{(m)}) = (\mathbf{F}_1^{\rightarrow}(r_{j_1}^{(m)}), \ldots, \mathbf{F}_d^{\rightarrow}(r_{j_d}^{(m)})). \tag{4.3}$$

We may now construct a $\mathcal{G}$-approximation of the indicator function $\mathbb{1}_{[-\overline{\infty}, t_{j-\overline{1}}^{(m)}]}$ by setting for $j \in \{1, \ldots, m\}^d$

$$\varphi_j^{(m)} := \begin{cases} \varphi_{\left(t_{j-\overline{2}}^{(m)}, t_{j-\overline{1}}^{(m)}\right)}, & \text{if } j \geq \overline{2}, \\ 0, & \text{if } j \ngeq \overline{2}, \end{cases} \tag{4.4}$$

where $\varphi_{(t_{j-\overline{2}}^{(m)}, t_{j-\overline{1}}^{(m)})}$ is a function in $\mathcal{G}$ satisfying (2.1), (2.2) and (2.6). Observe that $t_{j-\overline{2}}^{(m)} < t_{j-\overline{1}}^{(m)}$, since all $\mathbf{F}_i^{\rightarrow}$ are injective.

To approximate the empirical distribution function we introduce

$$\mathbf{F}_n^{(m)}(t) := \sum_{j\in\{1,\dots,m\}^d} \Big(\frac{1}{n}\sum_{i=1}^n \varphi_j^{(m)}(X_i)\Big)\mathbb{1}_{[t_{j-\bar{1}}^{(m)},t_j^{(m)})}(t). \tag{4.5}$$

Note that for t in any fixed rectangle $[t_{j-\bar{1}}^{(m)}, t_j^{(m)})$ we have the simple form

$$\mathbf{F}_n^{(m)}(t) = \frac{1}{n}\sum_{i=1}^n \varphi_j^{(m)}(X_i).$$

By definition of the $\varphi_j^{(m)}$ it is easy to see that therefore

$$\mathbf{F}_n(t_{j-\bar{2}}^{(m)}) \le \mathbf{F}_n^{(m)}(t) \le \mathbf{F}_n(t_{j-\bar{1}}^{(m)}) \ \ \forall t \in [t_{j-\bar{1}}^{(m)}, t_j^{(m)}),$$

implying pointwise convergence of $\mathbf{F}_n^{(m)} \to \mathbf{F}_n$ as $m \to \infty$. Thus it is just natural to try to approximate (as $m \to \infty$) U_n by

$$U_n^{(m)} := \Big(\sqrt{n}\big(\mathbf{F}_n^{(m)}(t) - \mathbf{F}^{(m)}(t)\big)\Big)_{t\in[-\infty,\infty]^d},$$

where

$$\mathbf{F}^{(m)}(t) := \mathbf{E}\big(\mathbf{F}_n^{(m)}(t)\big) = \sum_{j\in\{1,\dots,m\}^d} \mathbf{E}\big(\varphi_j^{(m)}(X_0)\big)\mathbb{1}_{[t_{j-\bar{1}}^{(m)},t_j^{(m)})}(t).$$

Remark 4.3. We need to keep in mind, that at this point the $\varphi_j^{(m)}$ and thus $U_n^{(m)}$ depend heavily on the chosen partition $r_0, \dots, r_m$ on $[0,1]$. Therefore the notation with the superscript m may be misleading at first glance, but since whenever the choice of the partition matters we will only use equidistant partitions of $[0,1]$, the partitions will in all relevant situations be uniquely defined by m.

The central idea to prove Theorem 2.1 is to use Proposition 4.1. Thus we need to check (4.1) and (4.2) for $S = \mathbb{D}([-\infty,\infty]^d)$. This is done in the next two lemmata:

Lemma 4.4. *For every partition $0 = r_0^{(m)} < \dots < r_m^{(m)} = 1$ of $[0,1]$, $U_n^{(m)}$ converges weakly to some centered Gaussian process $W^{(m)} \in \mathbb{D}([-\infty,\infty]^d)$ whose sample paths are constant on each of the rectangles $[t_{j-\bar{1}}^{(m)}, t_j^{(m)})$, $j \in \{1,\dots,m\}^d$.*

Lemma 4.5. *Let $0 = r_0^{(m)} < r_1^{(m)} < \dots < r_m^{(m)} = 1$ be the partition of $[0,1]$ defined by $r_k^{(m)} = \frac{k}{m}$. Then for every $\varepsilon, \eta > 0$ there is a $m_0 \in \mathbb{N}^*$ such that for all $m \ge m_0$*

$$\limsup_{n\to\infty} \mathbf{P}\Big(\sup_{t\in[-\infty,\infty]^d} |U_n(t) - U_n^{(m)}(t)| > \varepsilon\Big) \le \eta.$$

Proof of Lemma 4.4. Let $J_m := \{0,\dots,m\}^d$. For $i, j \in J_m$ set

$$\pi : \mathbb{D}([-\infty,\infty]^d) \longrightarrow (\mathbb{R}^{m+1})^d,\ x \mapsto (x(t_j))_{j\in J_m}$$

where $t_j \in [-\infty,\infty]^d$, is defined as in (4.3) (for convenience we omit the superscript m in $t_j^{(m)}$ since m is always fixed in this context). Since by condition (2.6) any linear combination of functions in $\mathcal{G}$ satisfies the Central Limit Theorem, an application of the Cramér-Wold device yields

$$\pi(U_n^{(m)}) \xrightarrow{\mathcal{D}} \overline{W}^{(m)} \quad \text{as } n \to \infty, \tag{4.6}$$

where $\overline{W}^{(m)}$ is a centered $(m+1)$-dimensional normal distributed random variable.

Set

$$\begin{aligned} H : (\mathbb{R}^{m+1})^d &\longrightarrow \mathbb{D}([-\infty,\infty]^d), \\ (z_j)_{j\in J_m} &\longmapsto \sum_{j\in J_{m-1}} z_j \mathbb{1}_{[t_j, t_{j+\bar{1}})} + \sum_{j\in J_m \setminus J_{m-1}} z_j \mathbb{1}_{[t_j, t_{j+\bar{1}}]}, \end{aligned}$$

and define the process

$$W^{(m)} := H \circ \overline{W}^{(m)}.$$

Note that $W^{(m)}$ is constant on the rectangles $[t_{j-\bar{1}}, t_j)$, $j \in \{1,\dots,m\}^d$.

We will now show, that $W^{(m)}$ is the weak limit of $U_n^{(m)}$. By definition we therefore need to prove

$$\int f \circ U_n^{(m)} \, d\mathbf{P} \overset{n\to\infty}{\longrightarrow} \int f \circ W^{(m)} \, d\mathbf{P} \quad \forall f \in C_b(\mathbb{D}([-\infty,\infty]^d)).$$

Let

$$\begin{aligned} u : [-\infty,\infty]^d &\longrightarrow [-\infty,\infty]^d \\ s &\longmapsto \sum_{j\in\{1,\dots,m\}^d} t_{j-\bar{1}} \mathbb{1}_{[t_{j-\bar{1}}, t_j)}(s) + \sum_{j\in\{1,\dots,m+1\}^d \setminus \{1,\dots,m\}^d} t_{j-\bar{1}} \mathbb{1}_{[t_{j-\bar{1}}, t_j]}(s). \end{aligned}$$

We have $H \circ \pi(x) = x \circ u$ for all $x \in \mathbb{D}([-\infty,\infty]^d)$ and $U_n^{(m)}(t) = U_n^{(m)} \circ u(t)$ almost sure, since $U_n^{(m)}$ is constant on the rectangles $[t_{j-\bar{1}}, t_j)$. Hence

$$f \circ U_n^{(m)} = f \circ U_n^{(m)} \circ u = f \circ H \circ \pi\big(U_n^{(m)}\big)$$

and analogue for $W^{(m)}$, applying $\pi \circ H = \mathbf{id}_{\mathbb{R}^{m+1}}$,

$$f \circ W^{(m)} = f \circ H \circ \pi\big(W^{(m)}\big) = f \circ H \circ \overline{W}^{(m)}.$$

Since $f \circ H \in C_b((\mathbb{R}^{m+1})^d)$ for $f \in C_b(\mathbb{D}([-\infty,\infty]^d))$ we obtain by (4.6)

$$\int f \circ U_n^{(m)} \, d\mathbf{P} = \int f \circ H \circ \pi\big(U_n^{(m)}\big) \, d\mathbf{P} \overset{n\to\infty}{\longrightarrow} \int f \circ H \circ \overline{W}^{(m)} \, d\mathbf{P} = \int f \circ W^{(m)} \, d\mathbf{P}.$$

for every $f \in C_b(\mathbb{D}([-\infty,\infty]^d))$. □

Proof of Lemma 4.5. Let us consider $\varepsilon, \eta > 0$ fixed for the rest of this proof. Choose the partition $0 = r_0^{(m)} < \ldots < r_m^{(m)} = 1$ of $[0,1]$ given by

$$r_q^{(m)} := \frac{q}{m},$$

$q \in \{0,\dots,m\}$, and set

$$h = \frac{1}{m}.$$

For each $k \in \mathbb{N}^*$ consider the refined partition

$$r_{q-1}^{(m)} = s_{q,0}^{(k)} < s_{q,1}^{(k)} < \ldots < s_{q,2^k}^{(k)} = r_q^{(m)}$$

of $[r_{q-1}, r_q]$, where

$$s_{q,\ell}^{(k)} := r_{q-1} + \ell \cdot \frac{h}{2^k},$$

$\ell \in \{0, \ldots, 2^k\}$ and $q \in \{0, \ldots, m\}$. Setting for $i \in \{1, \ldots, d\}$, $j_i \in \{1, \ldots, m\}$, $l_i \in \{0, \ldots, 2^k\}$

$$s_{i,j_i,l_i}^{(k)} = \mathbf{F}_i^{\rightarrow}(s_{j_i,l_i}^{(k)})$$

we obtain partitions

$$t_{i,j_i-1}^{(m)} = s_{i,j_i,0}^{(k)} < s_{i,j_i,1}^{(k)} < \ldots < s_{i,j_i,2^k}^{(k)} = t_{i,j_i}^{(m)}$$

of $[t_{i,j_i-1}^{(m)}, t_{i,j_i}^{(m)}]$. To simplify notations in the following calculations we set

$$s_{i,j_i,-1}^{(k)} := s_{i,j_i-1,2^k-1}^{(k)} \;\text{ for } j_i > 1, \qquad s_{i,j_i,2^k+1}^{(k)} := s_{i,j_i+1,1}^{(k)} \;\text{ for } j_i < m,$$

Let us now focus on a fixed rectangle $[t_{j-\bar{1}}^{(m)}, t_j^{(m)})$ for some $j = (j_1, \ldots, j_d) \in \{1, \ldots, m\}^d$. Our aim is to construct a chain to link the point $t_{j-\bar{1}}^{(m)}$ to some arbitrary point $t \in [t_{j-\bar{1}}^{(m)}, t_j^{(m)})$. Therefore we set

$$l_{i,j_i}(k,t) = \max\big\{\ell \in \{0, \ldots, 2^k\} : s_{i,j_i,\ell}^{(k)} \le t_i\big\} \in \{0, \ldots, 2^k - 1\}.$$

Since we consider j to be fixed, we may relinquish the use of the index j in the upper terms to simplify further notation. More precisely, we set

$$s_l^{(k)} := (s_{1,j_1,l_1}^{(k)}, \ldots, s_{d,j_d,l_d}^{(k)}), \qquad l(k,t) := (l_{1,j_1}(k,t), \ldots, l_{d,j_d}(k,t)).$$

In this way for any $k \in \mathbb{N}^*$ we obtain an ($[-\infty, \infty]^d$- valued) chain

$$t_{j-\bar{1}}^{(m)} = s_{l(0,t)}^{(0)} \le s_{l(1,t)}^{(1)} \le \ldots \le s_{l(k,t)}^{(k)} \le t \le s_{l(k,t)+\bar{1}}^{(k)}.$$

Now set

$$\psi_{\bar{0}}^{(0)} := \varphi_{(t_{j-\bar{1}}^{(m)}, t_j^{(m)})}$$

and choose for every $k \in \mathbb{N}^*$ and $l \in \{0, \ldots, 2^k + 1\}^d$ a function $\psi_l^{(k)} \in \mathcal{G}$ such that

$$\psi_l^{(k)}(x) = \begin{cases} 0, & \text{if } \exists i \in \{1, \ldots, d\} : j_i = 1 \text{ and } l_i = 0 \\ 1, & \text{if } (\exists i \in \{1, \ldots, d\} : j_i = m \text{ and } l_i = 2^k + 1) \\ & \text{and } (\nexists i \in \{1, \ldots, d\} : j_i = 1 \text{ and } l_i = 0) \\ \varphi_{(s_{l-\bar{1}}^{(k)}, s_l^{(k)})}(x) & \text{else,} \end{cases} \tag{4.7}$$

where $x \in [-\infty, \infty]^d$ and $\varphi_{(s_{l-\bar{1}}^{(k)}, s_l^{(k)})}$ satisfies (2.1) and (2.2).

Again, the reference to the indices j and m is omitted, since these are considered to be fixed.

By this definition we have for every $t \in [-\infty, \infty]^d$ and $l \in \{0, \ldots, 2^k\}^d$ the following useful inequalities:

$$\mathbb{1}_{[-\overline{\infty}, s_{l-\bar{1}}]} \le \psi_l^{(k)} \le \mathbb{1}_{[-\overline{\infty}, s_l]} \tag{4.8}$$

$$\varphi_j^{(m)} \le \psi_{l(1,t)}^{(1)} \le \ldots \le \psi_{l(k,t)}^{(k)} \le \mathbb{1}_{[-\overline{\infty},t]} \le \psi_{l(k,t)+\overline{2}}^{(k)}. \tag{4.9}$$

Using inequality (4.9) we obtain for $t \in [t_{j-\overline{1}}^{(m)}, t_j^{(m)})$, $K \in \mathbb{N}^*$ the telescopic-sum representation

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[-\overline{\infty},t]}(X_i) - \mathbf{F}_n^{(m)}(t) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[-\overline{\infty},t]}(X_i) - \frac{1}{n}\sum_{i=1}^{n} \varphi_j^{(m)}(X_i) \\
&= \sum_{k=1}^{K} \frac{1}{n}\sum_{i=1}^{n}\Big(\psi_{l(k,t)}^{(k)}(X_i) - \psi_{l(k-1,t)}^{(k-1)}(X_i)\Big) + \frac{1}{n}\sum_{i=1}^{n}\Big(\mathbb{1}_{[-\overline{\infty},t]}(X_i) - \psi_{l(K,t)}^{(K)}(X_i)\Big).
\end{aligned} \tag{4.10}$$

Let us now consider

$$U_n(t) - U_n^{(m)}(t) = \sqrt{n}\Big(\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{[-\overline{\infty},t]}(X_i) - \mathbf{F}(t)\Big) - \sqrt{n}\Big(\mathbf{F}_n^{(m)}(t) - \mathbf{F}^{(m)}(t)\Big).$$

Equation (4.10) yields

$$\begin{aligned}
U_n(t) - U_n^{(m)}(t) = &\sum_{k=1}^{K} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\psi_{l(k,t)}^{(k)}(X_i) - \mathbf{E}\psi_{l(k,t)}^{(k)}(X_0)\big) - \big(\psi_{l(k-1,t)}^{(k-1)}(X_i) - \mathbf{E}\psi_{l(k-1,t)}^{(k-1)}(X_0)\big) \\
&+ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\mathbb{1}_{[-\overline{\infty},t]}(X_i) - \mathbf{F}(t)\big) - \big(\psi_{l(K,t)}^{(K)}(X_i) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\big).
\end{aligned} \tag{4.11}$$

Applying the inequalities in (4.9), we gain the following upper bounds for the last sum on the right hand side of the upper inequality:

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\big(\mathbb{1}_{[-\overline{\infty},t]}(X_i) - \mathbf{F}(t)\big) - \big(\psi_{l(K,t)}^{(K)}(X_i) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\big)\Big) \\
&\ge -\sqrt{n}\Big(\mathbf{F}(t) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\Big) \\
&\ge -\sqrt{n}\Big(\mathbf{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\Big)
\end{aligned} \tag{4.12}$$

and

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\big(\mathbb{1}_{[-\overline{\infty},t]}(X_i) - \mathbf{F}(t)\big) - \big(\psi_{l(K,t)}^{(K)}(X_i) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\big)\Big) \\
&\le \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\big(\psi_{l(K,t)+\overline{2}}^{(K)}(X_i) - \mathbf{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0)\big) - \big(\psi_{l(K,t)}^{(K)}(X_i) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\big) \\
&\quad + \sqrt{n}\Big(\mathbf{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) - \mathbf{F}(t)\Big) \\
&\le \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\big(\psi_{l(K,t)+\overline{2}}^{(K)}(X_i) - \mathbf{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0)\big) - \big(\psi_{l(K,t)}^{(K)}(X_i) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\big) \\
&\quad + \sqrt{n}\Big(\mathbf{E}\psi_{l(K,t)+\overline{2}}^{(K)}(X_0) - \mathbf{E}\psi_{l(K,t)}^{(K)}(X_0)\Big).
\end{aligned} \tag{4.13}$$

For convenience, let $s^{(k)}_{i,m,2^k+1} := s^{(k)}_{i,m,2^k}$. By Equation (4.8) and the continuity of $\mathbf{F}$ (note that for continuous $\mathbf{F}$ we have $\mathbf{F} \circ \mathbf{F}^{\rightarrow} = \mathbf{id}_{[0,1]}$) we obtain

$$\begin{aligned}
&\sqrt{n}\Big|\mathbf{E}\psi^{(K)}_{l(K,t)+\overline{2}}(X_0) - \mathbf{E}\psi^{(K)}_{l(K,t)}(X_0)\Big| \\
&\leq \sqrt{n}\Big|\mathbf{E}\mathbb{1}_{[-\overline{\infty},s_{l(K,t)+\overline{2}})}(X_0) - \mathbf{E}\mathbb{1}_{[-\overline{\infty},s_{l(K,t)-\overline{1}})}(X_0)\Big| = \sqrt{n}\Big(\mathbf{F}(s_{l(K,t)+\overline{2}}) - \mathbf{F}(s_{l(K,t)-\overline{1}})\Big) \\
&\leq \sqrt{n}\Big(d \max_{i=1,\dots,d}\big\{\mathbf{F}_i(s_{i,j_i,l_{i,j_i}(K,t)+\overline{2}}) - \mathbf{F}_i(s_{i,j_i,l_{i,j_i}(K,t)-\overline{1}})\big\}\Big) = \frac{3d\sqrt{n}h}{2^K}
\end{aligned}$$

and thus, if we choose for some $n \in \mathbb{N}^*$ say

$$K = K_n := \left\lfloor \log_2\left(\frac{2^4 d}{\varepsilon}\sqrt{n}h\right)\right\rfloor \tag{4.14}$$

we obtain

$$\Big|\mathbf{E}\psi^{(K)}_{l(K,t)+\overline{2}}(X_0) - \mathbf{E}\psi^{(K)}_{l(K,t)}(X_0)\Big| < \frac{\varepsilon}{2}. \tag{4.15}$$

In summary, using (4.12), (4.13) and (4.15) in equality (4.11) yields

$$\begin{aligned}
&\big|U_n(t) - U_n^{(m)}(t)\big| \\
&< \Bigg|\sum_{k=1}^{K_n}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\psi^{(k)}_{l(k,t)}(X_i) - \mathbf{E}\psi^{(k)}_{l(k,t)}(X_0)\Big) - \big(\psi^{(k-1)}_{l(k-1,t)}(X_i) - \mathbf{E}\psi^{(k-1)}_{l(k-1,t)}(X_0)\big)\Bigg| \\
&\quad + \frac{1}{\sqrt{n}}\Bigg|\sum_{i=1}^{n}\big(\psi^{(K_n)}_{l(K_n,t)+\overline{2}}(X_i) - \mathbf{E}\psi^{(K_n)}_{l(K_n,t)+\overline{2}}(X_0)\big) - \big(\psi^{(K_n)}_{l(K_n,t)}(X_i) - \mathbf{E}\psi^{(K_n)}_{l(K_n,t)}(X_0)\big)\Bigg| \\
&\quad + \frac{\varepsilon}{2}
\end{aligned} \tag{4.16}$$

for every $n \in \mathbb{N}^*$.

Let us now consider the maximum of the terms in (4.16) over all $t \in [t^{(m)}_{j-\overline{1}}, t^{(m)}_j]$. By definition of the $l(k,t)$ we have

$$\left\lfloor\frac{l(k,t)}{2}\right\rfloor := \left(\left\lfloor\frac{l_{1,j_1}(k,t)}{2}\right\rfloor, \dots, \left\lfloor\frac{l_{d,j_d}(k,t)}{2}\right\rfloor\right) = l(k-1,t),$$

where $\lfloor x\rfloor := \max\{z \in \mathbb{Z} : z \leq x\}$. We therefore obtain

$$\begin{aligned}
&\sup_{t\in[t^{(m)}_{j-\overline{1}},t^{(m)}_j)} |U_n(t) - U_n^{(m)}(t)| \\
&< \sum_{k=1}^{K_n}\frac{1}{\sqrt{n}}\max_{l\in\{0,\dots,2^k-1\}^d}\Bigg|\sum_{i=1}^{n}\big(\psi^{(k)}_l(X_i) - \mathbf{E}\psi^{(k)}_l(X_0)\big) - \big(\psi^{(k-1)}_{\lfloor l/2\rfloor}(X_i) - \mathbf{E}\psi^{(k-1)}_{\lfloor l/2\rfloor}(X_0)\big)\Bigg| \\
&\quad + \frac{1}{\sqrt{n}}\max_{l\in\{0,\dots,2^{K_n}-1\}^d}\Bigg|\sum_{i=1}^{n}\big(\psi^{(K_n)}_{l+\overline{2}}(X_i) - \mathbf{E}\psi^{(K_n)}_{l+\overline{2}}(X_0)\big) - \big(\psi^{(K_n)}_l(X_i) - \mathbf{E}\psi^{(K_n)}_l(X_0)\big)\Bigg| \\
&\quad + \frac{\varepsilon}{2}.
\end{aligned} \tag{4.17}$$

Choose ε_k, $k \in \mathbb{N}^*$ such that $\sum_{k=1}^{\infty}\varepsilon_k = \varepsilon/4$, say

$$\varepsilon_k = \frac{\varepsilon}{4k(k+1)}.$$

An application of Markov's inequality on the $2p$'th moments combined with condition (2.3) implies

$$
\begin{aligned}
&\mathbf{P}\Big(\sup_{t\in[t_{j-1}^{(m)},t_j^{(m)})}|U_n(t)-U_n^{(m)}(t)|\geq\varepsilon\Big)\\
&\leq\sum_{k=1}^{K_n}\sum_{l\in\{0,\dots,2^k-1\}^d}\mathbf{P}\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big(\psi_l^{(k)}(X_i)-\mathbf{E}\psi_l^{(k)}(X_0)\big)-\big(\psi_{\lfloor l/2\rfloor}^{(k-1)}(X_i)-\mathbf{E}\psi_{\lfloor l/2\rfloor}^{(k-1)}(X_0)\big)\Big|>\varepsilon_k\Big)\\
&\quad+\sum_{l\in\{0,\dots,2^{K_n}-1\}^d}\mathbf{P}\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big(\psi_{l+\bar{2}}^{(K_n)}(X_i)-\mathbf{E}\psi_{l+\bar{2}}^{(K_n)}(X_0)\big)-\big(\psi_l^{(K_n)}(X_i)-\mathbf{E}\psi_l^{(K_n)}(X_0)\big)\Big|>\frac{\varepsilon}{4}\Big)\\
&\leq\sum_{k=1}^{K_n}\sum_{l\in\{0,\dots,2^k-1\}^d}\frac{1}{\varepsilon_k^{2p}n^p}\mathbf{E}\Big(\Big|\sum_{i=1}^{n}\big(\psi_l^{(k)}(X_i)-\mathbf{E}\psi_l^{(k)}(X_0)\big)-\big(\psi_{\lfloor l/2\rfloor}^{(k-1)}(X_i)-\mathbf{E}\psi_{\lfloor l/2\rfloor}^{(k-1)}(X_0)\big)\Big|^{2p}\Big)\\
&\quad+\sum_{l\in\{0,\dots,2^{K_n}-1\}^d}\frac{4^{2p}}{n^p\varepsilon^{2p}}\mathbf{E}\Big(\Big|\sum_{i=1}^{n}\big(\psi_{l+\bar{2}}^{(K_n)}(X_i)-\mathbf{E}\psi_{l+\bar{2}}^{(K_n)}(X_0)\big)-\big(\psi_l^{(K_n)}(X_i)-\mathbf{E}\psi_l^{(K_n)}(X_0)\big)\Big|^{2p}\Big)\\
&\leq 2C\Bigg\{\sum_{k=1}^{K_n}\sum_{l\in\{0,\dots,2^k-1\}^d}\frac{1}{\varepsilon_k^{2p}n^p}\sum_{i=1}^{p}n^i\,\|\psi_l^{(k)}(X_0)-\psi_{\lfloor l/2\rfloor}^{(k-1)}(X_0)\|_r^i\,\Phi_i\Big(2\|\psi_l^{(k)}-\psi_{\lfloor l/2\rfloor}^{(k-1)}\|_{\mathcal{G}}+C'\Big)\\
&\quad+\sum_{l\in\{0,\dots,2^{K_n}-1\}^d}\frac{4^{2p}}{\varepsilon^{2p}n^p}\sum_{i=1}^{p}n^i\,\|\psi_{l+\bar{2}}^{(K_n)}(X_0)-\psi_l^{(K_n)}(X_0)\|_r^i\,\Phi_i\Big(2\|\psi_{l+\bar{2}}^{(K_n)}-\psi_l^{(K_n)}\|_{\mathcal{G}}+C'\Big)\Bigg\}
\end{aligned}
\tag{4.18}
$$

The critical part in the upper terms surely is the argument in the functions Φ_i. Let us therefore concentrate the required auxiliary calculations in the following lemma:

Lemma 4.6. *For all $l\in\{0,\dots,2^k-1\}^d$, $k,n,r\in\mathbb{N}^*$*

$$
\begin{aligned}
\|\psi_l^{(k)}(X_0)-\psi_{\lfloor l/2\rfloor}^{(k-1)}(X_0)\|_r &\leq\Big(\frac{3dh}{2^k}\Big)^{\frac{1}{r}},\\
\|\psi_{l+\bar{2}}^{(K_n)}(X_0)-\psi_l^{(K_n)}(X_0)\|_r &\leq\Big(\frac{3dh}{2^{K_n}}\Big)^{\frac{1}{r}},\\
\|\psi_l^{(k)}\|_{\mathcal{G}} &\leq\max\Big\{\Psi\Big(\frac{2^k}{h}\Big),\|1\|_{\mathcal{G}}\Big\}
\end{aligned}
$$

Proof.

$$
\begin{aligned}
\|\psi_l^{(k)}(X_0)-\psi_{\lfloor l/2\rfloor}^{(k-1)}(X_0)\|_r &\leq\|\mathbb{1}_{[-\bar{\infty},s_l^{(k)}]}(X_0)-\mathbb{1}_{[-\bar{\infty},s_{\lfloor l/2\rfloor-\bar{1}}^{(k-1)}]}(X_0)\|_r\\
&\leq\Big(\sum_{i=1}^{d}\big|\mathbf{F}_i(s_{i,j_i,l_i}^{(k)})-\mathbf{F}_i(s_{i,j_i,\lfloor l_i/2\rfloor-1}^{(k-1)})\big|\Big)^{\frac{1}{r}}\\
&\leq\Big(d\max_{i=1,\dots,d}\big(\mathbf{F}_i(s_{i,j_i,l_i}^{(k)})-\mathbf{F}_i(s_{i,j_i,l_i-3}^{(k)})\big)\Big)^{\frac{1}{r}}\leq\Big(\frac{3dh}{2^k}\Big)^{\frac{1}{r}},
\end{aligned}
$$

by (4.8) and the continuity of the $\mathbf{F}_i$. The second inequality can be proven alike.

In the first two cases in definition (4.7) $\psi_l^{(k)}$ is a constant function taking either the value zero or one for each argument. In this cases the last inequality of the lemma is trivially satisfied by the conditions on $\|\cdot\|_{\mathcal{G}}$.

Else $\psi_l^{(k)}$ has a representation $\varphi_{(s_{l-\bar{1}}^{(k)}, s_l^{(k)})}$, where

$$
\begin{aligned}
s_l^{(k)} &= \Big(s_{1,j_1,l_1}^{(k)}, \dots, s_{d,j_d,l_d}^{(k)}\Big) = \Big(\mathbf{F}_1^{\rightarrow}\big(s_{j_1,l_1}^{(k)}\big), \dots, \mathbf{F}_d^{\rightarrow}\big(s_{j_d,l_d}^{(k)}\big)\Big), \\
s_{l-\bar{1}}^{(k)} &= \Big(s_{1,j_1,l_1-1}^{(k)}, \dots, s_{d,j_d,l_d-1}^{(k)}\Big) = \Big(\mathbf{F}_1^{\rightarrow}\big(s_{j_1,l_1}^{(k)} - h2^{-k}\big), \dots, \mathbf{F}_d^{\rightarrow}\big(s_{j_d,l_d}^{(k)} - h2^{-k}\big)\Big)
\end{aligned}
$$

and hence

$$
\begin{aligned}
s_{i,j_i,l_i}^{(k)} - s_{i,j_i,l_i-1}^{(k)} \in& \{\delta > 0 : \exists t \in \mathbb{R},\ |\mathbf{F}_i(t) - \mathbf{F}_i(t-\delta)| \ge h2^{-k}\} \\
&\subset \{\delta > 0 : w_{\mathbf{F}_i}(\delta) \ge h2^{-k}\}
\end{aligned}
\tag{4.19}
$$

for every $i \in \{1, \dots, d\}$. To see this, set $t = \mathbf{F}_i^{\rightarrow}(s_{l_i}^{(k)})$, $\delta = \mathbf{F}_i^{\rightarrow}(s_{l_i}^{(k)}) - \mathbf{F}_i^{\rightarrow}(s_{l_i-1}^{(k)}) > 0$ and remember that the $\mathbf{F}_i^{\rightarrow}$ are injective.

Now condition by (2.2) yields

$$
\|\varphi_l^{(k)}\|_{\mathcal{G}} \le \Psi\bigg(\frac{1}{\min\limits_{i=1,\dots,d} w_{\mathbf{F}_i}\big(s_{i,j_i,l_i}^{(k)} - s_{i,j_i,l_i-1}^{(k)}\big)}\bigg) \le \Psi\Big(\frac{2^k}{h}\Big),
$$

since

$$
\min_{i=1,\dots,d} w_{\mathbf{F}_i}\big(s_{j_i,l_i}^{(k)} - s_{1,j_i,l_i-1}^{(k)}\big) \ge h2^{-k}
$$

by (4.19). □

An application of Lemma 4.6 to (4.18) yields

$$
\begin{aligned}
&\mathbf{P}\bigg(\sup_{t \in [t_{j-1}^{(m)}, t_j^{(m)})} |U_n(t) - U_n^{(m)}(t)| \ge \varepsilon\bigg) \\
&\le 2C\bigg\{\sum_{k=1}^{K_n}\sum_{i=1}^{p} \frac{2^{dk} n^{-(p-i)}}{\varepsilon_k^{2p}} \Big(\frac{3dh}{2^k}\Big)^{\frac{i}{r}} \Phi_i\Big(2\Psi\Big(\frac{2^k}{h}\Big) + C'\Big) \\
&\qquad + \sum_{i=1}^{p} \frac{2^{dK_n} n^{-(p-i)}}{(\frac{\varepsilon}{4})^{2p}} \Big(\frac{3dh}{2^{K_n}}\Big)^{\frac{i}{r}} \Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big) + C'\Big)\bigg\} \\
&\le 2C \sum_{i=1}^{p}\bigg\{(3d)^{\frac{i}{r}} n^{-(p-i)} \sum_{k=1}^{K_n} \frac{2^{dk}}{\varepsilon_k^{2p}} \Big(\frac{h}{2^k}\Big)^{\frac{i}{r}} \Phi_i\Big(2\Psi\Big(\frac{2^k}{h}\Big) + C'\Big)\bigg\} \\
&\le D \sum_{i=1}^{p}\bigg\{n^{-(p-i)} \sum_{k=1}^{K_n} 2^{(d-\frac{i}{r})k} k^{4p} \Phi_i\Big(2\Psi\Big(\frac{2^k}{h}\Big) + C'\Big) h^{\frac{i}{r}}\bigg\} \\
&\le D \sum_{i=1}^{p-1}\bigg\{n^{-(p-i)} \Big(\frac{2^{K_n}}{h}\Big)^{d-\frac{i}{r}} {K_n}^{4p+1} \Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big) + C'\Big) h^d\bigg\} \\
&\qquad + D\bigg\{\sum_{k=1}^{K_n} 2^{(d-\frac{p}{r})k} k^{4p} \Phi_p\Big(2\Psi\Big(\frac{2^k}{h}\Big) + C'\Big) h^{\frac{p}{r}}\bigg\}
\end{aligned}
\tag{4.20}
$$

for every $j \in \{1, \dots, m\}^d$, where $D \in \mathbb{R}$ denotes some finite constant. In the second inequality we used that Ψ and Φ_i are monotone increasing functions and $\varepsilon/4 > \varepsilon_{K_n}$.

Let us first deal with the term in the last line of (4.20). By condition (2.4) we have

$$
\sum_{k=1}^{K_n} 2^{(d-\frac{p}{r})k} k^{4p} \Phi_p\Big(2\Psi\Big(\frac{2^k}{h}\Big) + C'\Big) h^{\frac{p}{r}} \le C'' h^{\frac{p}{r}-\gamma_p} \sum_{k=1}^{\infty} 2^{(\gamma_p - (\frac{p}{r}-d))k} k^{4p}
$$

where $\gamma_p < \frac{p}{r} - d$. Hence there is a nonnegative constant $D' < \infty$ such that

$$\sum_{k=1}^{K_n} 2^{(d-\frac{p}{r})k} k^{4p} \Phi_p\Big(2\Psi\Big(\frac{2^k}{h}\Big) + C'\Big) h^{\frac{p}{r}} \le D' h^{\frac{p}{r}-\gamma_p} = o(h^d). \tag{4.21}$$

Now consider the first summand on the right hand side of inequality (4.20). In (4.14) we chose $K_n = \lfloor \log_2(2^4 d\sqrt{n}h/\varepsilon) \rfloor$, hence condition (2.4) yields for any $i = 1, \dots, p-1$

$$\begin{aligned} &n^{-(p-i)} \Big(\frac{2^{K_n}}{h}\Big)^{d-\frac{i}{r}} {K_n}^{4p+1} \Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big) + C'\Big) h^d \\ &\le D'' \log_2^{4p+1}\Big(\frac{2^4 d}{\varepsilon}\sqrt{n}h\Big) \cdot (\sqrt{n})^{\gamma_i - (\frac{i}{r} + 2(p-i) - d)} h^d \end{aligned}$$

for some nonnegative constant $D'' < \infty$. Since $\gamma_i < \frac{i}{r} + 2(p-i) - d$ for $i = 1, \dots, p-1$ by condition (2.4) we obtain for all $\eta > 0$ and sufficiently large $n \in \mathbb{N}^*$

$$D \sum_{i=1}^{p-1} \Big\{ n^{-(p-i)} \Big(\frac{2^{K_n}}{h}\Big)^{d-\frac{i}{r}} {K_n}^{4p+1} \Phi_i\Big(2\Psi\Big(\frac{2^{K_n}}{h}\Big) + C'\Big) h^d \Big\} \le \frac{1}{2}\eta h^d. \tag{4.22}$$

Finally, by (4.20), (4.21) and (4.22), for any $\eta > 0$

$$\begin{aligned} &\limsup_{n\to\infty} \mathbf{P}\Big(\sup_{t\in[-\infty,\infty]^d} |U_n(t) - U_n^{(m)}(t)| \ge \varepsilon \Big) \\ &\le \limsup_{n\to\infty} \sum_{j\in\{1\dots,m\}^d} \mathbf{P}\Big(\sup_{t\in[t_{j-1}^{(m)}, t_j^{(m)})} |U_n(t) - U_n^{(m)}(t)| \ge \varepsilon \Big) \\ &\le m^d \Big(o(h^d) + \frac{1}{2}\eta h^d \Big) = m^d \Big(o(m^{-d}) + \frac{1}{2}\eta m^{-d} \Big), \end{aligned}$$

since $h = 1/m$. Hence there is a $m_0 \in \mathbb{N}^*$ such that

$$\limsup_{n\to\infty} \mathbf{P}\Big(\sup_{t\in[-\infty,\infty]^d} |U_n(t) - U_n^{(m)}(t)| \ge \varepsilon \Big) \le \eta$$

for all $m \ge m_0$. □

With Lemma 4.4 and Lemma 4.5 established, let us finally prove Theorem 2.1:

Proof of Theorem 2.1. Set $r_k^{(m)} := \frac{k}{m}$ in Lemma 4.4. An application of Proposition 4.1 on $S = \mathbb{D}([-\infty,\infty]^d)$ equipped with the Skorohod metric ρ yields that $W^{(m)}$ converges in distribution. Since all $W^{(m)}$, $m \in \mathbb{N}^*$, are centered Gaussian processes the limit process must also be (centred) Gaussian.

It remains to prove the continuity of the sample paths of W: At this point we already know that U_n converges weakly to W. Therefore it is sufficient to show that for every $\varepsilon, \eta > 0$, there is a $\delta > 0$ such that

$$\limsup_{n\to\infty} \mathbf{P}\Big(\sup_{\|t-s\|<\delta} |U_n(t) - U_n(s)| > 3\varepsilon \Big) < 3\eta. \tag{4.23}$$

The sufficiency of this condition can be proven exactly the same way as in the proof of Theorem 15.5 in [Bil68, p.127 f.].

For all $m \in \mathbb{N}^*$ by some triangle inequality arguments we obtain for all $m \in \mathbb{N}^*$

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbf{P}\Big(\sup_{\|t-s\|<\delta} |U_n(t) - U_n(s)| > 3\varepsilon\Big) \\
&\leq 2 \limsup_{n\to\infty} \mathbf{P}\Big(\sup_t |U_n(t) - U_n^{(m)}(t)| > \varepsilon\Big) + \limsup_{n\to\infty} \mathbf{P}\Big(\sup_{\|t-s\|<\delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\Big)
\end{aligned}$$

and thus, by Lemma 4.5, there is a $m_0 \in \mathbb{N}^*$ such that for all $m \geq m_0$ have

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbf{P}\Big(\sup_{\|t-s\|<\delta} |U_n(t) - U_n(s)| > 3\varepsilon\Big) \\
&\leq 2\eta + \limsup_{n\to\infty} \mathbf{P}\Big(\sup_{\|t-s\|<\delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\Big). \qquad (4.24)
\end{aligned}$$

Now set

$$\delta_m := \frac{1}{2} \min_{j\in\{0,\ldots,m\}^d} \Big\{\max_{i=1,\ldots,d} |t_{j_i} - t_{j_i - 1}|\Big\}$$

and observe that δ_m is strictly bigger than zero for any $m \in \mathbb{N}^*$, since the $\mathbf{F}_i^{\rightarrow}$ used in the construction of the t_j are strictly increasing.

Obviously for all $\delta \leq \delta_m$ and $\|t - s\| < \delta$ the points $s, t \in [-\infty, \infty]$ must be located in adjacent (or identical) intervals of the form $[t_j, t_{j-\bar{1}})$. Since the process $U_n^{(m)}$ is constant on any of the intervals $[t_j, t_{j-\bar{1}})$ and by symmetry in the arguments s, t we obtain

$$\sup_{\|t-s\|<\delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| = \max_{\substack{j\in\{0,\ldots,m\}^d \\ z\in\{0,1\}^d,\ j\geq z}} |U_n^{(m)}(t_j) - U_n^{(m)}(t_{j-z})|,$$

thus

$$\begin{aligned}
&\mathbf{P}\Big(\sup_{\|t-s\|<\delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\Big) \\
&\leq 2^d (m+1)^d \max_{\substack{j\in\{0,\ldots,m\}^d \\ z\in\{0,1\}^d,\ j\geq z}} \mathbf{P}\Big(|U_n^{(m)}(t_j) - U_n^{(m)}(t_{j-z})| > \varepsilon\Big). \qquad (4.25)
\end{aligned}$$

Analogously to the calculation in Lemma 4.6 one can show that for all $j \in \{0, \ldots, m\}^d$ and $z \in \{0, 1\}^d$ such that $j \geq z$ we have

$$\begin{aligned}
\|\varphi_{j+\bar{1}}^{(m)}(X_0) - \varphi_{j+\bar{1}-z}^{(m)}(X_0)\|_r &\leq \Big(\frac{3d}{m}\Big)^{\frac{1}{r}}, \\
\|\varphi_{j+\bar{1}}^{(m)}\|_{\mathcal{G}} &\leq \max\Big\{\Psi(m), \|1\|_{\mathcal{G}}\Big\}.
\end{aligned}$$

By applying one after another Markov's inequality, the $2p$-th moment bounds (2.3) and

Lemma 4.6 we obtain (for sufficiently large $m \in \mathbb{N}^*$) by the definition of $\varphi_j^{(m)}$ (see page 11)

$$\begin{aligned}
&\mathbf{P}\Big(|U_n^{(m)}(t_j) - U_n^{(m)}(t_{j-z})| > \varepsilon\Big)\\
&= \mathbf{P}\bigg(\Big|\sum_{i=1}^{n}\big(\varphi_{j+\bar{1}}^{(m)}(X_i) + \varphi_{j+\bar{1}-z}^{(m)}(X_i)\big) - \mathbf{E}\big(\varphi_{j+\bar{1}}^{(m)}(X_0) + \varphi_{j+\bar{1}-z}^{(m)}(X_0)\big)\Big|^{2p} > n^p\varepsilon^{2p}, \bigg)\\
&\le n^{-p}\varepsilon^{-2p}\cdot \mathbf{E}\Big|\sum_{i=1}^{n}\big(\varphi_{j+\bar{1}}^{(m)}(X_i) + \varphi_{j+\bar{1}-z}^{(m)}(X_i)\big) - \mathbf{E}\big(\varphi_{j+\bar{1}}^{(m)}(X_0) + \varphi_{j+\bar{1}-z}^{(m)}(X_0)\big)\Big|^{2p}\\
&\le Cn^{-p}\varepsilon^{-2p}\sum_{i=1}^{p} n^i \|\varphi_{j+\bar{1}}^{(m)}(X_0) - \varphi_{j+\bar{1}-z}^{(m)}(X_0)\|_r^i \Phi_i\Big(\|\varphi_{j+\bar{1}}^{(m)} - \varphi_{j+\bar{1}-z}^{(m)}\| + C'\Big)\\
&\le Cn^{-p}\varepsilon^{-2p}\sum_{i=1}^{p} n^i \Big(\frac{3d}{m}\Big)^{\frac{i}{r}}\Phi_i\Big(\Psi(m) + C'\Big)\\
&\le D\sum_{i=1}^{p} n^{-(p-i)} m^{\gamma_i - \frac{i}{r}}\\
&\le Dm^{\gamma_p - (\frac{p}{r})} + D\sum_{i=1}^{p-1} n^{-(p-i)} m^{\gamma_i - \frac{i}{r}},
\end{aligned}$$

where D is some finite constant. Therefore by (4.25) there is another finite constant D' such that

$$\mathbf{P}\Big(\sup_{\|t-s\|<\delta} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\Big) \le D'm^d\Big(m^{\gamma_p - (\frac{p}{r})} + \sum_{i=1}^{p-1} n^{-(p-i)} m^{\gamma_i - (\frac{i}{r})}\Big),$$

thus

$$\limsup_{n\to\infty}\mathbf{P}\Big(\sup_{\|t-s\|<\delta_m} |U_n^{(m)}(t) - U_n^{(m)}(s)| > \varepsilon\Big) \le D'm^{\gamma_p - (\frac{p}{r} - d)} < \eta$$

for sufficiently large $m \in \mathbb{N}^*$, say $m \ge m_1$.

By (4.24) this implies that (4.23) holds for $\delta = \delta_{\max\{m_0, m_1\}}$. □

5 Proof of Proposition 3.5

We have

$$\begin{aligned}
\Big|\mathbf{E}\Big(\Big(\sum_{i=1}^{n}\varphi(X_i)\Big)^p\Big)\Big| &= \Big|\sum_{0\le i_1,\ldots,i_p\le n-1} \mathbf{E}\big(\varphi(X_{i_1})\cdot\ldots\cdot\varphi(X_{i_p})\big)\Big|\\
&\le p!\Big|\sum_{\substack{0\le i_1,\ldots,i_p\le n-1\\ i_p^*\le n-1}} \mathbf{E}\big(\varphi(X_{i_1^*})\cdot\ldots\cdot\varphi(X_{i_p^*})\big)\Big|\\
&\le p!n\Big|\sum_{\substack{0\le i_1,\ldots,i_p\le n-1\\ i_p^*\le n-1}} \mathbf{E}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot\varphi(X_{i_{p-1}^*})\big)\Big|.
\end{aligned}$$

Using the denotation

$$I_n(p) := \sum_{\substack{0\le i_1,\ldots,i_p\le n-1\\ i_p^*\le n-1}} \Big|\mathbf{E}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot(X_{i_p^*})\big)\Big| \quad \text{for } p\in\mathbb{N}^*, \tag{5.1}$$

$$I_n(0) := \big|\mathbf{E}(\varphi(X_0))\big| = 0,$$

we therefore have

$$\Big|\mathbf{E}\Big(\Big(\sum_{i=1}^{n}\varphi(X_i)\Big)^p\Big)\Big| \le p!nI_n(p-1). \tag{5.2}$$

Decomposing the sum in (5.1) with respect to the highest increment of indices i_q, $q \in \{1,\dots,p\}$, we receive a bound

$$I_n(p) \le \sum_{q=1}^{p} J_n(p,q),$$

where

$$J_n(p,q) = \sum_{i_q=0}^{n-1} \sum_{\substack{0\le i_1,\dots,i_{q-1},i_{q+1},\dots,i_p\le i_q\\ i_p^*\le n-1}} \big|\mathbf{E}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot(X_{i_p^*})\big)\big|.$$

Lemma 5.1. *Let $(X_n)_{n\in\mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{G})$, then for all $p \in \mathbb{N}^*$ such that*

$$\sum_{i=0}^{\infty} i^{p-1}\Theta(i) < \infty \tag{5.3}$$

and all $q \in \{1,\dots,p\}$ there is a constant K' such that

$$J_n(p,q) \le K'\|\varphi(X_0)\|_r\|\varphi\|_{\mathcal{G}} + nI_n(q-1)I_n(p-q)$$

for all $n \in \mathbb{N}^*$ and $\varphi \in \mathcal{G}$.

Proof. Set

$$\begin{aligned}
A_{i_1,\dots,i_p} &:= \big|\mathbf{Cov}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot(X_{i_{q-1}^*})\,,\,\varphi(X_{i_q^*})\varphi(X_{i_{q+1}^*})\cdot\ldots\cdot(X_{i_p^*})\big)\big|\\
B_{i_1,\dots,i_p} &:= \big|\mathbf{E}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot(X_{i_q^*})\big)\big|\cdot\big|\mathbf{E}\big(\varphi(X_{i_q^*})\varphi(X_{i_{q+1}^*})\cdot\ldots\cdot(X_{i_p^*})\big)\big|\\
&= \big|\mathbf{E}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot(X_{i_q^*})\big)\big|\cdot\big|\mathbf{E}\big(\varphi(X_0)\varphi(X_{i_{q+1}})\cdot\ldots\cdot(X_{i_p^*-i_q^*})\big)\big|,
\end{aligned}$$

where we used the stationarity of $(X_i)_{i\in\mathbb{N}}$ in the last line. We have

$$J_n(p,q) \le \sum_{i_q=0}^{n-1} \sum_{\substack{0\le i_1,\dots,i_{q-1},i_{q+1},\dots,i_p\le i_q\\ i_p^*\le n-1}} A_{i_1,\dots,i_p} + \sum_{i_q=0}^{n-1} \sum_{\substack{0\le i_1,\dots,i_{q-1},i_{q+1},\dots,i_p\le i_q\\ i_p^*\le n-1}} B_{i_1,\dots,i_p}.$$

An application of the polynomial multiple mixing property (3.12) yields

$$\begin{aligned}
\sum_{i_q=0}^{n-1} \sum_{\substack{0\le i_1,\dots,i_{q-1},i_{q+1},\dots,i_p\le i_q\\ i_p^*\le n-1}} A_{i_1,\dots,i_p} &\le K\|\varphi(X_0)\|_r\|\varphi\|_{\mathcal{G}}\sum_{i_q=0}^{n-1}(i_q+1)^{p-1}\Theta(i_q)\\
&\le K'\|\varphi(X_0)\|_r\|\varphi\|_{\mathcal{G}}\sum_{i_q=0}^{\infty} i_q^{p-1}\Theta(i_q)\\
&\le K''\|\varphi(X_0)\|_r\|\varphi\|_{\mathcal{G}}
\end{aligned}$$

for some constants $K', K'' < \infty$, since $\sum_{i_q=0}^{\infty} i_q^{p-1}\Theta(i_q) < \infty$ by assumption (5.3).

Finally

$$
\begin{aligned}
&\sum_{i_q=0}^{n-1} \sum_{\substack{0\le i_1,\dots,i_{q-1},i_{q+1},\dots,i_p\le i_q \\ i_p^*\le n-1}} B_{i_1,\dots,i_p} \\
&\le \sum_{i_q=0}^{n-1} \Bigg\{ \sum_{\substack{0\le i_1,\dots,i_{q-1}\le n-1 \\ i_{q-1}^*\le n-1}} \Big| \mathbf{E}\big(\varphi(X_0)\varphi(X_{i_1^*})\cdot\ldots\cdot(X_{i_q^*})\big) \Big| \\
&\quad \cdot \sum_{\substack{0\le i_{q+1},\dots,i_p\le n-1 \\ i_p^*-i_q^*\le n-1}} \Big| \mathbf{E}\big(\varphi(X_0)\varphi(X_{i_{q+1}})\cdot\ldots\cdot(X_{i_p^*-i_q^*})\big) \Big| \Bigg\} \\
&= nI_n(p-1)I_n(p-q).
\end{aligned}
$$

□

Lemma 5.2. *If $(X_n)_{n\in\mathbb{N}} \in \mathbf{MM}_{\Theta,r}(\mathcal{G})$, then for all $p \in \mathbb{N}^*$ such that (5.3) is satisfied there is a constant $K_p < \infty$, such that*

$$
I_n(p) \le K_p \sum_{i=1}^{\lceil p/2 \rceil} n^{i-1} \|\varphi(X_0)\|_r^i \|\varphi\|_{\mathcal{G}}^i \tag{5.4}
$$

for all $\varphi \in \mathcal{G}$ with $\|\varphi\|_\infty \le 1$ and $\mathbf{E}(\varphi(X_0)) = 0$.

Proof. We will use mathematical induction to prove the lemma. By Lemma 5.1 we can easily see that

$$
I_n(1) \le K_1 \|\varphi(X_0)\|_r \, \|\varphi\|_{\mathcal{G}}
$$

for some constant $K_1 < \infty$ if (5.3) is satisfied. Now consider an arbitrary $\tilde{p} \in \mathbb{N}\backslash\{1\}$ satisfying (5.3) and assume (5.4) holds for all $p \le \tilde{p} - 1$. We have

$$
\begin{aligned}
I_n(\tilde{p}) &\le \sum_{q=1}^{\tilde{p}} J_n(\tilde{p}, q) \le K'' \sum_{q=1}^{\tilde{p}} \|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{G}} + nI_n(q-1)I_n(\tilde{p}-q) \\
&\le \tilde{p}K'' \|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{G}} \\
&\quad + n \sum_{q=1}^{\tilde{p}} \Bigg\{ \Big(K_{q-1} \sum_{i=1}^{\lceil \frac{q-1}{2} \rceil} n^{i-1} \|\varphi(X_0)\|_r^i \|\varphi\|_{\mathcal{G}}^i \Big) \Big(K_{\tilde{p}-q} \sum_{j=1}^{\lceil \frac{\tilde{p}-q}{2} \rceil} n^{j-1} \|\varphi(X_0)\|_r^j \|\varphi\|_{\mathcal{G}}^j \Big) \Bigg\} \\
&\le K''' \Bigg\{ \|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{G}} + n \sum_{q=1}^{\tilde{p}} \sum_{i=2}^{\lceil \frac{q-1}{2} \rceil + \lceil \frac{\tilde{p}-q}{2} \rceil} n^{i-2} \|\varphi(X_0)\|_r^i \|\varphi\|_{\mathcal{G}}^i \Bigg\} \\
&\le K''' \Bigg\{ \|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{G}} + n\tilde{p} \sum_{i=2}^{\lceil \frac{\tilde{p}}{2} \rceil} n^{i-2} \|\varphi(X_0)\|_r^i \|\varphi\|_{\mathcal{G}}^i \Bigg\} \\
&\le K_p \sum_{i=1}^{\lceil \tilde{p}/2 \rceil} n^{i-1} \|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{G}}
\end{aligned}
$$

for some constants $K'', K''', K_p < \infty$, since

$$\left\lceil \frac{q-1}{2} \right\rceil + \left\lceil \frac{\tilde{p}-q}{2} \right\rceil \leq \begin{cases} \frac{q-1}{2} + \frac{\tilde{p}-q}{2} + \frac{1}{2}, & \text{if } \tilde{p} \text{ is even} \\ \frac{q-1}{2} + \frac{\tilde{p}-q}{2} + \frac{1}{2} + \frac{1}{2}, & \text{if } \tilde{p} \text{ is odd} \end{cases}$$
$$\leq \left\lceil \frac{\tilde{p}}{2} \right\rceil$$

□

Proof of proposition 3.5. By (5.2) and Lemma 5.2 we immediately obtain

$$\begin{aligned} \mathbf{E}\Big(\Big|\sum_{i=1}^{n} \varphi(X_i)\Big|^{2p}\Big) &\leq p! n I_n(2p-1) \\ &\leq p! K_p \sum_{i=1}^{\lceil p/2 \rceil} n^i \|\varphi(X_0)\|_r^i \|\varphi\|_{\mathcal{G}}^i \end{aligned}$$

if (5.3) is still satisfied replace p by $2p-1$, i.e. if (3.13) holds. □

6 Proof of Lemma 3.7 and Lemma 3.8

Proof of Lemma 3.7. Let $(\xi'_j)_{j\in\mathbb{Z}}$ and $(\xi''_j)_{j\in\mathbb{Z}}$ be copies of the underlying process $(\xi_j)_{j\in\mathbb{Z}}$ such that all three processes are independent. Set

$$\begin{aligned} \dot{X}_i^{(k)} &:= F(\xi_i, \xi_{i-1}, \ldots, \xi_{i-k+1}, \xi'_{i-k}, \xi'_{i-k-1}, \ldots) \\ \ddot{X}_i^{(k)} &:= F(\xi_i, \xi_{i-1}, \ldots, \xi_{i-k+1}, \xi''_{i-k}, \xi''_{i-k-1}, \ldots) \end{aligned}$$

and note that therefore $(X_i)_{i\in\mathbb{N}^*} \stackrel{\mathcal{D}}{=} (\dot{X}_i)_{i\in\mathbb{N}^*} \stackrel{\mathcal{D}}{=} (\ddot{X}_i)_{i\in\mathbb{N}^*}$

$$\begin{aligned} &\big|\mathbf{Cov}\big(\varphi(X_0)\cdot\ldots\cdot\varphi(X_{i^*_{q-1}})\,,\,\varphi(X_{i^*_q})\cdot\ldots\cdot\varphi(X_{i^*_p})\big)\big| \\ &\leq \big|\mathbf{Cov}\big(\varphi(X_0)\cdot\ldots\cdot\varphi(X_{i^*_{q-1}}) - \varphi(\dot{X}_0^{(k)})\cdot\ldots\cdot\varphi(\dot{X}^{(k)}_{i^*_{q-1}})\,,\,\varphi(X_{i^*_q})\cdot\ldots\cdot\varphi(X_{i^*_p})\big)\big| \\ &\quad + \big|\mathbf{Cov}\big(\varphi(\dot{X}_0^{(k)})\cdot\ldots\cdot\varphi(\dot{X}^{(k)}_{i^*_{q-1}})\,,\,\varphi(X_{i^*_q})\cdot\ldots\cdot\varphi(X_{i^*_p})\big) - \varphi(\ddot{X}^{(k)}_{i^*_q})\cdot\ldots\cdot\varphi(\ddot{X}^{(k)}_{i^*_p})\big)\big| \\ &\quad + \big|\mathbf{Cov}\big(\varphi(\dot{X}_0^{(k)})\cdot\ldots\cdot\varphi(\dot{X}^{(k)}_{i^*_{q-1}})\,,\,\varphi(\ddot{X}^{(k)}_{i^*_q})\cdot\ldots\cdot\varphi(\ddot{X}^{(k)}_{i^*_p})\big)\big|. \end{aligned} \tag{6.1}$$

Since $\varphi(\dot{X}_0^{(k)})\cdot\ldots\cdot\varphi(\dot{X}^{(k)}_{i^*_{q-1}})$ is $\sigma\big(\{\xi_j : j \leq i^*_{q-1}\} \cup \{\xi'_j : j \in \mathbb{Z}\}\big)$-measurable while $\varphi(\ddot{X}^{(k)}_{i^*_q})\cdot\ldots\cdot\varphi(\ddot{X}^{(k)}_{i^*_p})$ is $\sigma\big(\{\xi_j : j > i^*_q - k\} \cup \{\xi''_j : j \in \mathbb{Z}\}\big)$-measuable, the functions in the last covariance on the right hand side of equation (6.1) are independent as soon as $k \leq i_q$ and thus the last summand is equal to 0 in this case.

Remember that we only consider such φ that satisfy $\|\varphi\|_\infty \leq 1$. If we apply Hölder's equality to equation (6.1) we obtain for r, s satisfying $\frac{1}{r} + \frac{1}{s} = 1$

$$\begin{aligned} &\big|\mathbf{Cov}\big(\varphi(X_0)\cdot\ldots\cdot\varphi(X_{i^*_{q-1}}) - \varphi(\dot{X}_0^{(k)})\cdot\ldots\cdot\varphi(\dot{X}^{(k)}_{i^*_{q-1}})\,,\,\varphi(X_{i^*_q})\cdot\ldots\cdot\varphi(X_{i^*_p})\big)\big| \\ &\leq 2\|\varphi(X_0)\cdot\ldots\cdot\varphi(X_{i^*_{q-1}}) - \varphi(\dot{X}_0^{(k)})\cdot\ldots\cdot\varphi(\dot{X}^{(k)}_{i^*_{q-1}})\|_s \|\varphi(X_{i^*_q})\cdot\ldots\cdot\varphi(X_{i^*_p})\|_r \\ &\leq 2q\|\varphi(X_0)\|_r \|\varphi(X_0) - \varphi(\dot{X}_0^{(k)})\|_s \end{aligned} \tag{6.2}$$

where we used that for $a_i, b_i \in [-1,1]$

$$\Big|\prod_{i=1}^n a_i - \prod_{i=1}^n b_i\Big| \le \sum_{i=1}^n |a_i - b_i|$$

holds for all $n \in \mathbb{N}^*$. Since

$$|\varphi(x) - \varphi(y)| \le \|\varphi\|_{\mathcal{H}_\alpha} \|x-y\|^\alpha \quad \text{for all } x, y \in \mathbb{R}^d, \tag{6.3}$$

an application of Jensens's inequality to equation (6.2) yields

$$\begin{aligned}
&\Big|\mathbf{Cov}\big(\varphi(X_0)\cdot \ldots \cdot \varphi(X_{i^*_{q-1}}) - \varphi(\dot{X}_0^{(k)})\cdot \ldots \cdot \varphi(\dot{X}^{(k)}_{i^*_{q-1}})\,,\, \varphi(X_{i^*_q})\cdot \ldots \cdot \varphi(X_{i^*_p})\big)\Big| \\
&\le 2q\|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{H}_\alpha} \big(\|X_0 - \dot{X}_0^{(k)}\|_s\big)^\alpha \\
&= 2q\|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{H}_\alpha} (\delta_{k,s})^\alpha.
\end{aligned} \tag{6.4}$$

Analogously we one can show that

$$\begin{aligned}
&\Big|\mathbf{Cov}\big(\varphi(\dot{X}_0^{(k)})\cdot \ldots \cdot \varphi(\dot{X}^{(k)}_{i^*_{q-1}})\,,\, \varphi(X_{i^*_q})\cdot \ldots \cdot \varphi(X_{i^*_p})\big) - \varphi(\ddot{X}^{(k)}_{i^*_q})\cdot \ldots \cdot \varphi(\ddot{X}^{(k)}_{i^*_p})\big)\Big| \\
&\le 2(p-q)\|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{H}_\alpha} \|(\delta_{k,s})^\alpha,
\end{aligned} \tag{6.5}$$

thus for $k = i_q$ we have

$$\begin{aligned}
&\Big|\mathbf{Cov}\big(\varphi(X_0)\cdot \ldots \cdot \varphi(X_{i^*_{q-1}})\,,\, \varphi(X_{i^*_q})\cdot \ldots \cdot \varphi(X_{i^*_p})\big)\Big| \\
&\le 2q\|\varphi(X_0)\|_r \|\varphi\|_{\mathcal{H}_\alpha} (\delta_{i_q,s})^\alpha
\end{aligned}$$

□

To prove Lemma 3.8 we will work with an immediate implication of a theorem Dedecker introduced in [Ded98]:

Proposition 6.1. *Let $(Y_i)_{i\in\mathbb{N}}$ be a ergodic stationary process with $\mathbf{E}(Y_0) = 0$ and $\mathbf{E}(Y_0^2) < \infty$, which is adapted to a filtration $(\mathcal{M}_i)_{i\in\mathbb{N}}$. If*

$$\sum_{i=0}^n Y_0 E(Y_i|\mathcal{M}_0)$$

converges in $\mathbf{L}_1$, then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \xrightarrow{\mathcal{D}} N(0,\sigma^2) \quad \text{as } n \to \infty,$$

where $N(0,\sigma^2)$ is a normal distributed random variable with mean zero and variance

$$\sigma^2 = \mathbf{E}(Y_0^2) + 2\sum_{i=1}^\infty \mathbf{E}(Y_0 \cdot Y_i).$$

Proof of Lemma 3.8. Let μ denote the distribution of the ξ_i. Since $(X_i)_{i\in\mathbb{N}}$ is a causal function of an i.i.d. process, we have a representation

$$X_i = F(\xi_i, \xi_{i-1}, \ldots),$$

where $F : \mathcal{X}^{\mathbb{N}} \to \mathbb{R}^d$.

Choose an arbitrary $f \in \mathcal{H}_\alpha$ with $\mathbf{E}(f(X_0)) = 0$ and $\|f\|_\infty < \infty$. The process $(Y_i)_{i\in\mathbb{N}}$ given by

$$Y_i = f(X_i) \tag{6.6}$$

is centered, ergodic, has finite second moments and is adapted to the filtration

$$(\mathcal{M}_i)_{i\in\mathbb{N}} := \big(\sigma(\xi_i, \xi_{i-1}, \ldots)\big)_{i\in\mathbb{N}}.$$

As before, let $(\xi'_j)_{j\in\mathbb{Z}}$ be an independent copy of $(\xi_j)_{j\in\mathbb{Z}}$ and set

$$\begin{aligned}
X'_i &:= F(\xi'_i, \xi'_{i-1}, \ldots)\\
\dot{X}_i &:= F(\xi_i, \xi_{i-1}, \ldots, \xi_1, \xi'_0, \xi'_{-1}, \ldots)\\
\dot{X}'_i &:= F(\xi'_i, \xi'_{i-1}, \ldots, \xi'_1, \xi_0, \xi_{-1}, \ldots).
\end{aligned}$$

Observe that by the independence of $\mathcal{M}_0$ and $\sigma(\{\xi'_i : i \in \mathbb{Z}\})$ we have that

$$\mathbf{E}\big(f(X'_i)|\mathcal{M}_0\big) = \mathbf{E}(f(X'_i)) = 0, \qquad \mathbf{E}\big(f(X_i)|\mathcal{M}_0\big) = \mathbf{E}\big(f(\dot{X}'_i)|\mathcal{M}_0\big).$$

Thus

$$\begin{aligned}
\mathbf{E}\Big\{\big|f(X_0)\mathbf{E}\big(f(X_i)|\mathcal{M}_0\big)\big|\Big\} &\le \mathbf{E}\Big\{\big|\mathbf{E}\big(f(X_i)|\mathcal{M}_0\big)\big|\Big\} = \mathbf{E}\Big\{\big|\mathbf{E}\big(f(\dot{X}'_i)|\mathcal{M}_0\big) - \mathbf{E}\big(f(X'_i)|\mathcal{M}_0\big)\big|\Big\}\\
&= \mathbf{E}\Big\{\big|\mathbf{E}\big(f(\dot{X}'_i) - f(X'_i)|\mathcal{M}_0\big)\big|\Big\} \le \mathbf{E}\big|f(\dot{X}'_i) - f(X'_i)\big|\\
&\le \|f\|_{\mathcal{H}_\alpha}\mathbf{E}\big|\dot{X}'_i - X'_i\big|^\alpha \le \|f\|_{\mathcal{H}_\alpha}(\delta_{i,1})^\alpha
\end{aligned}$$

where we used (6.3) and Jensen's inequality in the last steps. We therefore obtain

$$\mathbf{E}\Big\{\big|Y_0\mathbf{E}\big(Y_i|\mathcal{M}_0\big)\big|\Big\} = \mathbf{E}\Big\{\big|f(X_0)\mathbf{E}\big(f(X_i)|\mathcal{M}_0\big)\big|\Big\} \le \|f\|_{\mathcal{H}_\alpha}(\delta_{i,1})^\alpha,$$

where $(\delta_{i,1})^\alpha$ satisfies (3.15). Therefore $\sum_{i=1}^n Y_0\mathbf{E}\big(Y_i|\mathcal{M}_0\big)$ converges in $\mathbf{L}_1$ and thus Proposition 6.1 applies, yielding

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n f(X_i) = \frac{1}{\sqrt{n}}\sum_{i=1}^n Y_i \xrightarrow{\mathcal{D}} N(0, \sigma^2) \quad \text{as } n \to \infty.$$

□

Acknowledgement

This research was partially supported by German Research Foundation grant DE 370-4 project: *New Techniques for Empirical Processes of Dependent Data.*

References

[BBD01] Svetlana Borovkova, Robert Burton, and Herold Dehling. Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Trans. Amer. Math. Soc.*, 353(11):4261–4318 (electronic), 2001.

[Bil68] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[BP78] István Berkes and Walter Philipp. An almost sure invariance principle for the empirical distribution function of mixing random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 41(2):115–137, 1977/78.

[DD11] Herold Dehling and Olivier Durieu. Empirical processes of multidimensional systems with multiple mixing properties. *Stochastic Process. Appl.*, 121(5):1076–1096, 2011.

[DDV09] Herold Dehling, Olivier Durieu, and Dalibor Volný. New techniques for empirical processes of dependent data. *Stochastic Process. Appl.*, 119(10):3699–3718, 2009.

[Ded98] Jérôme Dedecker. A central limit theorem for stationary random fields. *Probab. Theory Related Fields*, 110(3):397–426, 1998.

[Ded10] J. Dedecker. An empirical central limit theorem for intermittent maps. *Probab. Theory Related Fields*, 148(1-2):177–195, 2010.

[DL99] Paul Doukhan and Sana Louhichi. A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.*, 84(2):313–342, 1999.

[DMR95] Paul Doukhan, Pascal Massart, and Emmanuel Rio. Invariance principles for absolutely regular empirical processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(2):393–427, 1995.

[Don52] Monroe D. Donsker. Justification and extension of Doob's heuristic approach to the Komogorov-Smirnov theorems. *Ann. Math. Statistics*, 23:277–281, 1952.

[DP02] Herold Dehling and Walter Philipp. Empirical process techniques for dependent data. In *Empirical process techniques for dependent data*, pages 3–113. Birkhäuser Boston, Boston, MA, 2002.

[DP07] Jérôme Dedecker and Clémentine Prieur. An empirical central limit theorem for dependent sequences. *Stochastic Process. Appl.*, 117(1):121–142, 2007.

[Neu71] Georg Neuhaus. On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.*, 42:1285–1295, 1971.

[Str72] Miron L. Straf. Weak convergence of stochastic processes with several parameters. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 187–221, Berkeley, Calif., 1972. Univ. California Press.

[Wu08] Wei Biao Wu. Empirical processes of stationary sequences. *Statist. Sinica*, 18(1):313–333, 2008.