SFB 823

The sequential empirical process of nonlinear long-range dependent random vectors

Jannis Buchsteiner

Nr. 33/2015

Discussion Paper

The Sequential Empirical Process of Nonlinear Long-Range Dependent Random Vectors

Jannis Buchsteiner*

Fakultät für Mathematik, Ruhr-Universität Bochum, Germany.

Abstract

Let $(G(X_j))_{j\geq 1}$ be a multivariate subordinated Gaussian process, which exhibits long-range dependence. We study the asymptotic behaviour of the corresponding sequential empirical process under two different types of subordination. The limiting process is either a product of a deterministic function and a Hermite process as in the one-dimensional case or a sum of various processes of this kind.

Keywords: Multivariate long-range dependence, sequential empirical process, subordinated Gaussian process

1 Introduction

For a Gaussian process $(X_j)_{j\geq 1}$ and a measurable function G the sequential empirical process $(R_N(x,t))$ corresponding to the subordinated process $(G(X_j))_{j\geq 1}$ is given by

$$R_N(x,t) := \sum_{j=1}^{\lfloor Nt \rfloor} \Big(1_{\{G(X_j)\leq x\}} - P(G(X_j) \leq x) \Big), \quad x \in \mathbb{R}, t \in [0,1]. \tag{1}$$

This process plays an important role in nonparametric statictics, for example in change-point analysis. If the underlying Gaussian process exhibits long-range dependence (LRD) weak convergence was shown by Dehling and Taqqu (1989). The limiting process can be represented as the product of a deterministic function and a Hermite process $(Z_m(t))_{0\leq t\leq 1}$ which is a fractional Brownian motion if $m = 1$ and a non-Gaussian process for $m \geq 2$. A first step in generalizing this result to multivariate observations was done by Marinucci (2005). He studied the asymptotics of the empirical process

*E-mail: `jannis.buchsteiner@rub.de`

Research supported by Collaborative Research Center SFB 823 *Statistical modeling of nonlinear dynamic processes.*

$(R_N(x,1))$ based on a two-dimensional LRD process. Roughly speaking one could find two different approaches in the literature to define LRD for a p-dimensional stochastic process $(Y_j)_{j\geq 1}$. The first one is to take p independent processes $(Y_j^{(1)})_{j\geq 1},\dots,(Y_j^{(p)})_{j\geq 1}$ such that each of them exhibits LRD, i.e.

$$\mathrm{Cov}(Y_1^{(i)},Y_{k+1}^{(i)}) = L_i(k)k^{-D_i}, \quad 0<D_i<1 \tag{2}$$

and L_i is slowly varying at infinity. For $p=2$ this construction was used by Marinucci (2005) and Taufer (2014, $D_1 = D_2$). A more general setup was used by Ho and Sun (1990) and Arcones (1994). They call a p-dimensional Gaussian process LRD if both the covariance function of each component and further the cross-covariance $\mathrm{Cov}(Y_1^{(i)},Y_{k+1}^{(j)})$ are of type (2). Kechagias and Pipiras (2015) stated a very precise definition of LRD in time and spectral domain and showed under which conditions these are equivalent.

In the present paper we want to establish new non-central limit theorems for the sequential empirical process based on p-dimensional LRD data. In the tradition of the initial work of Dehling and Taqqu (1989) we will focus on subordinated Gaussian processes. More precisely we will consider two different types of subordination. This approach will help us to make the proofs more transparent. In section 2 the p-dimensional observations are generated by a one-dimensional Gaussian process. In section 3 the process $(G(X_j))$ is q-dimensional, where the underlying Gaussian process itself is p-dimensional. For both cases we prove a weak uniform reduction principle and show that $(R_N(x,t))$ converges weakly to a Hermite process as in the one-dimensional case (section 2) or to a sum of generalized Hermite processes (section 3), respectively.

2 One-dimensional subordination

The simplest way to get a p-dimensional subordinated Gaussian random vector is the following construction. Let $(X_j)_{j\geq 1}$ be an one-dimensional stationary Gaussian process with $EX_1 = 0$ and $EX_1^2 = 1$. Moreover, let this sequence exhibits long-range dependence such that the covariance function $r(k) = EX_1X_{k+1}$ satisfies

$$r(k) = k^{-D}L(k), \tag{3}$$

where $0 < D < 1$ and L is a function which is slowly varying at infinity. For any measurable functions $G_1,\dots,G_p : \mathbb{R}\to\mathbb{R}$ we consider the process $(Y_j)_{j\geq 1}$, where Y_j is a random vector given by

$$Y_j := G(X_j) := (G_1(X_j),\dots.G_p(X_j)).$$

To simplify reading, we first introduce some notations. We denote by F the common distribution function of Y_1 and by F_k the distribution function of $G_k(X_1)$. For an element $(x^{(1)},\dots,x^{(p)})\in\mathbb{R}^p$ we write simply x and by $x\leq y$ we mean $x_i\leq y_i$, for all $1\leq i\leq p$. The sequential empirical process based on these p-dimensional observations can be written just as (1) as

$$R_N(x,t) := \sum_{j=1}^{\lfloor Nt\rfloor}\left(1_{\{Y_j\leq x\}} - F(x)\right).$$

2.1 Results

As in the initial work of Dehling and Taqqu (1989) the asymptic behaviour of $(R_N(x,t))$ is determined by the leading term of the so-called Hermite expansion. These relates to the collection of Hermite polynomials $(H_n)_{n\geq 0}$,

$$H_n(y) := (-1)^n e^{y^2/2} \frac{d^n}{dy^n} e^{-y^2/2},$$

which forms an orthogonal basis of $L_2(\varphi(y)dy)$, where φ is the standard normal density. Since $(1_{\{G(\cdot)\leq x\}} - F(x))$ is square integrable for any $x \in \mathbb{R}^p$, we have the following L_2-representation

$$1_{\{Y_j \leq x\}} - F(x) = \sum_{q=0}^{\infty} \frac{J_q(x)}{q!} H_q(X_j).$$

The Hermite coefficients are given by the inner product, i.e.

$$J_q(x) = E(1_{\{Y_j\leq x\}} - F(x))H_q(X_j). \tag{4}$$

We call the index $m(x)$ of the first nonzero Hermite coefficient the Hermite rank of $(1_{\{G(\cdot)\leq x\}} - F(x))$. Note that $m(x) \geq 1$ for any $x \in \mathbb{R}^p$.

Theorem 1. *Let $(X_j)_{j\geq 1}$ be a standard one-dimensional Gaussian process satisfying* (3) *and let $G : \mathbb{R} \to \mathbb{R}^p$ be a measurable function. Furthermore, let $0 < D < 1/m$, where $m := \min\{m(x) : x \in \mathbb{R}^p\}$. Then*

$$\left\{d_N^{-1} R_N(x,t) : (x,t) \in [-\infty,\infty]^p \times [0,1]\right\}$$

converges weakly in $D([-\infty,\infty]^p \times [0,1])$, to

$$\left\{\frac{J_m(x)}{m!} Z_m(t) : (x,t) \in [-\infty,\infty]^p \times [0,1]\right\},$$

where $(Z_m(t))$ is a Hermite process.

Remark 1. It was shown by Taqqu (1975, Corollary 4.1) that

$$d_N^2 \approx N^{2-mD} L(N).$$

Remark 2. The Hermite process $(Z_m(t))_{0\leq t\leq 1}$ is given by a stochastic integral in spectral domain, more precisely

$$Z_m(t) = c \int_{\mathbb{R}^m}'' \frac{e^{it(x_1+\ldots+x_m)} - 1}{i(x_1 + \ldots + x_m)} \prod_{j=1}^{m} |x_j|^{-(1-D)/2} B(dx_1)\cdots B(dx_m), \tag{5}$$

where B is a suitable random spectral measure and $c \in \mathbb{R}$ is a constant which only depends on m and D. For details and further representations see Taqqu (1979) and Pipiras and Taqqu (2010).

Theorem 2 (Reduction principle)**.** *Let $(X_j)_{j\geq 1}$ be a standard one-dimensional Gaussian process satisfying* (3) *and let $G : \mathbb{R} \to \mathbb{R}^p$ be a measurable function. Furthermore, let $0 < D < 1/m$, where $m := \min\{m(x) : x \in \mathbb{R}^p\}$. Then there exist constants $C, \kappa > 0$ such that for any $0 < \varepsilon \leq 1$*

$$
\begin{aligned}
&P\left(\max_{n\leq N} \sup_{x\in[-\infty,\infty]^p} d_N^{-1} \left|\sum_{j=1}^{n} \left(1_{\{Y_k\leq x\}} - F(x) - \frac{J_m(x)}{m!} H_m(X_j)\right)\right| > \varepsilon\right) \\
\leq & CN^{-\kappa}(1+\varepsilon^{-3}).
\end{aligned}
$$

The normalization factor is given by

$$
d_N^2 = \operatorname{Var}\left(\sum_{j=1}^{N} H_m(X_j)\right).
$$

The statement of Theorem 2 is strongly associated to the reduction principle of Dehling and Taqqu (1989) in the case $p = 1$. They were the first who established such a theorem uniformly in x and t. Giraitis and Surgailis (2002) studied the reduction principle for the seqential empirical process of long-range dependent moving average data. Buchsteiner (2015) showed that the reduction principle of Dehling and Taqqu is still valid, if the càdlàg space is equipped with a weighted supremum norm. In all cases the weak reduction principle can be applied to prove weak convergence of the normalized sequential empirical processs.

2.2 The Hermite Rank

In order to use Theorem 1 for studying the asymptotic behaviour of $R_N(x,t)$ it is important to know the Hermite rank m of $\{1_{\{G(\cdot)\leq x\}} : x \in \mathbb{R}^p\}$. For $p = 1$ we usually have $m \leq 2$, see Dehling and Taqqu (1989, pp. 1770). Therefore the question arises, is there a connection between the Hermite ranks m_j of $\{1_{\{G_j(\cdot)\leq x\}} : x \in \mathbb{R}\}$ and m?

Lemma 1. *Let m_j be the Hermite rank of $\{1_{\{G_j(\cdot)\leq x\}} : x \in \mathbb{R}\}$, $1 \leq j \leq p$, and let m be the Hermite rank of $\{1_{\{G(\cdot)\leq x\}} : x \in \mathbb{R}^p\}$. Then*

$$
m \leq \min\{m_j : 1 \leq j \leq p\}.
$$

Proof. For simplicity let $m_1 = \min\{m_j : 1 \leq j \leq p\}$. Then there exist $x \in \mathbb{R}$ s.t. $E(1_{\{G_1(X_1)\leq x\}} H_{m_1}(X_1)) \neq 0$. By dominated convergence we have

$$
\begin{aligned}
&\lim_{n\to\infty} E(1_{\{G_1(X_1)\leq x, G_2(X_1)\leq n,\ldots,G_p(X_1)\leq n\}} H_{m_1}(X_1)) \\
=&E(\lim_{n\to\infty} 1_{\{G_1(X_1)\leq x, G_2(X_1)\leq n,\ldots,G_p(X_1)\leq n\}} H_{m_1}(X_1)) \\
=&E(1_{\{G_1(X_1)\leq x\}} H_{m_1}(X_1)) \neq 0.
\end{aligned}
$$

Therfore it exists an index n_0 so that $E(1_{\{G_1(X_1)\leq x, G_2(X_1)\leq n_0,\ldots,G_p(X_1)\leq n_0\}} H_{m_1}(X_1)) \neq 0$ and this implies $m \leq m_1$. □

Example 1. Assume we have $\min\{m_j : 1 \leq j \leq p\} = 1$. Then we get by Lemma 1 $m = 1$. In other words, if there is at least one function G_j s.t. $\{1_{\{G_j(\cdot)\leq x\}} : x \in \mathbb{R}\}$ has Hermite rank 1 then $R_N(x,t)$ converges weakly to a fractional Brownian motion.

Example 2. Let $p = 2$, $G_1(s) = s^2$ and $G_2(s) = 1_A(s)$, where $A = (-\infty, -c) \cup (0, c)$ and $c = (-2\ln(1/2))^{1/2}$. Note that $\{s \in \mathbb{R} : G_2(s) \leq x\} \in \{\emptyset, A^c, \mathbb{R}\}$ for all $x \in \mathbb{R}$ and

$$\int_{-c}^{0} x\varphi(x)dx + \int_{c}^{\infty} x\varphi(x)dx = 0,$$

$$\int_{-c}^{0} (x^2-1)\varphi(x)dx + \int_{c}^{\infty} (x^2-1)\varphi(x)dx = 0.$$

Using numeric integration we get $E(1_{\{G_2(X_1)\leq x\}}H_3(X_1)) \neq 0$ for $0 \leq x < 1$. This implies $\{1_{\{G_2(\cdot)\leq x\}} : x \in \mathbb{R}\}$ has Hermite rank 3. Furthermore, we know by Dehling and Taqqu (1989, Example 2) that $\{1_{\{G_1(\cdot)\leq x\}} : x \in \mathbb{R}\}$ has Hermite rank 2.
Now we are in the special situation in which the Hermite rank corresponding to the common observation is really smaller than 2, since

$$E(1_{\{G_1(X_1)\leq c^2, G_2(X_1)\leq 0\}}H_1(X_1)) = \int_{-c}^{0} x\varphi(x)dx = 2(2\pi)^{-1/2}.$$

3 Multivariate subordination

Let $G : \mathbb{R}^p \to \mathbb{R}^q$ a measurable function and let $(X_j)_{j\geq 1} = ((X_j^{(1)}, \ldots, X_j^{(p)}))_{j\geq 1}$ be a p-dimensional stationary Gaussian process with the following properties

$$EX_1^{(i)} = 0 \quad 1 \leq i \leq p \tag{6}$$

$$EX_1^{(i)}X_1^{(j} = \delta_{ij} \tag{7}$$

$$r^{(i,j)}(k) := EX_1^{(i)}X_{k+1}^{(j)} = c_{ij}L(k)k^{-D}, \tag{8}$$

where $c_{ij} \in \mathbb{R}$ are constants depending only on i and j, $L(k)$ is slowly varying at infinity and $0 < D < 1$. From now we assume that the $\mathbb{R}^q$-valued stochastic process $(Y_j)_{j\geq 1}$ is given by $Y_j = G(X_j)$.
The conditions (6), (7) and (8) were proposed by Arcones (1994) and they are also compatible with the very general definition of multivariate long-range dependence given by Kechagias and Pipiras (2015).

In order to prove a weak uniform reduction principle we have to establish a multivariate Hermite decomposition of $R_N(x,t)$. Let L^2 be the space of square integrable functions with respect to the p-dimensional standard normal distribution. An orthogonal basis of these space is given by the collection of multivariate Hermite polynomials

$$H_{l_1,\ldots,l_p}(x) = H_{l_1}(x^{(1)}) \cdots H_{l_p}(x^{(p)}) \quad l_1, \ldots, l_p \in \mathbb{N},$$

where H_{l_i} is an ordinary one-dimensional Hermite polynomial. Therefore, for any $x \in \mathbb{R}^q$ the following series expansion holds in L^2

$$1_{\{Y_j \leq x\}} - F(x) = \sum_{k=1}^{\infty} \sum_{l_1+\ldots+l_p=k} \frac{J_{l_1,\ldots,l_p}(x)}{l_1! \cdots l_p!} H_{l_1,\ldots,l_p}(X_j), \tag{9}$$

where

$$J_{l_1,\ldots,l_p}(x) = E(1_{\{G(X_j)\leq x\}} H_{l_1,\ldots,l_p}(X_j)). \tag{10}$$

The maximum index m satisfying $J_{l_1,\ldots,l_p}(x) = 0$ whenever $l_1 + \ldots + l_p < m$ is called the Hermite rank of $1_{\{G(\cdot)\leq x\}} - F(x)$ and the Hermite rank of $(1_{\{G(\cdot)\leq x\}} - F(x))_{x\in\mathbb{R}^q}$ is defined as the minimum of all pointwise Hermite ranks.

With respect to (9) one recognizes that a limiting process for $(R_N(x,t))$ in the context of mutivariate subordination differs from the one in section 2, since there could be up to $\binom{m+p-1}{m}$ multivariate Hermite polynomials contributing to the limit.
Let $(B^{(1)}), \ldots, B^{(p)})$ the joint random spectral measure satisfying

$$\left\{ d_N^{-1} \sum_{j=1}^{\lfloor Nt \rfloor} \left(H_m(X_j^{(1)}), \ldots, H_m(X_j^{(p)}) \right) : t \in [0,1] \right\} \xrightarrow{d} \left\{ \left(Z_m^{(1)}(t), \ldots, Z_m^{(p)}(t) \right) : t \in [0,1] \right\}, \tag{11}$$

where $Z_m^{(k)}(t)$ is defined as in (5) with $B = B^{(k)}$ for $1 \leq k \leq p$.

Theorem 3. *Let $(X_j)_{j\geq 1}$ be a p-dimensional Gaussian process satisfying* (6), (7) *and* (8) *and let $G : \mathbb{R}^p \to \mathbb{R}^q$ be a measurable function. Furthermore, let $0 < D < 1/m$, where m is the Hermite rank of $(1_{\{G(\cdot)\leq x\}} - F(x))_{x\in\mathbb{R}^q}$. Then*

$$\left\{ d_N^{-1} R_N(x,t) : (x,t) \in [-\infty,\infty]^q \times [0,1] \right\}$$

converges weakly in $D([-\infty,\infty]^q \times [0,1])$, to

$$\left\{ c \sum_{j_1,\ldots,j_m=1}^{p} \tilde{J}_{j_1,\ldots,j_m}(x) \int_{\mathbb{R}^m}'' \frac{e^{it(x_1+\ldots+x_m)} - 1}{i(x_1+\ldots+x_m)} \prod_{j=1}^{m} |x_j|^{-(1-D)/2} B^{(j_1)}(dx_1) \cdots B^{(j_m)}(dx_m) : (x,t) \in [-\infty,\infty]^q \times [0,1] \right\}. \tag{12}$$

The constant c is the same as in (5) *and*

$$\tilde{J}_{j_1,\ldots,j_m}(x) = (m!)^{-1} E\left(1_{\{G(X_1)\leq x\}} \prod_{i=1}^{p} H_{i(j_1,\ldots,j_m)}(X_1^{(j)}) \right),$$

where $i(j_1,\ldots,j_m)$ is the number of indeces $j_1,\ldots,j_m$ that are eqal to i.

Remark 3. Arcones (1994, Theorem 6) studied the generalized Hermite process (12) in the non-uniform case. Note that we have corrected the domain of integration, i.e. we have $\mathbb{R}^m$ instead of $[-\pi,\pi]^m$. For further details on stochastic integrals with dependent integrators, see Fox and Taqqu (1987).

Remark 4. Theorem 3 is not a corollary of Theorem 9 by Arcones (1994). Although Arcones states a non-central limit theorem for the empirical process indexed by functions, the class $\{1_{\{G(\cdot)\le x\}} - f(x) : x \in \mathbb{R}^p\}$ does not satisfy the required bracketing condition.

Theorem 4 (Reduction principle). *Let $(X_j)_{j\ge 1}$ be a p-dimensional Gaussian process satisfying (6), (7) and (8) and let $G : \mathbb{R}^p \to \mathbb{R}^q$ be a measurable function. Furthermore, let $0 < D < 1/m$, where m is the Hermite rank of $(1_{\{G(\cdot)\le x\}} - F(x))_{x\in\mathbb{R}^q}$. Then there exist constants $C, \kappa > 0$ such that for any $0 < \varepsilon \le 1$*

$$P\left(\max_{n\le N}\sup_{x\in[-\infty,\infty]^q} d_N^{-1}\left|\sum_{j=1}^{n}\left(1_{\{Y_k\le x\}} - F(x) - \sum_{l_1+\ldots+l_p=m}\frac{J_{l_1,\ldots,l_p}(x)}{l_1!\cdots l_p!}H_{l_1,\ldots,l_p}(X_j)\right)\right| > \varepsilon\right)$$
$$\le CN^{-\kappa}(1+\varepsilon^{-3}).$$

The normalization factor is given by

$$d_N^2 = \mathrm{Var}\left(\sum_{j=1}^{N} H_m(X_j^{(1)})\right).$$

4 Proofs

We start by introducing suitable partitions of $\mathbb{R}^p$. For $x \in \mathbb{R}^p$ let

$$\Lambda(x) := \sum_{j=1}^{p}\Lambda_j(x^{(j)}), \tag{13}$$

where $\Lambda_j : [-\infty,\infty] \to \mathbb{R}$ are non-decreasing, right-continious functions satisfying $\Lambda_j(-\infty) = 0$ and $\Lambda_j(\infty) = \Lambda_1(\infty) < \infty$ for all $j = 1,\ldots,p$. These functions will be specified in section 4.1 and 4.2, respectively. For any $k \in \mathbb{N}$, $1 \le i \le 2^k - 1$ and $1 \le j \le p$ let

$$\begin{aligned}x_i^{(j)}(k) :=& \inf\{x \in \mathbb{R} : \Lambda_j(x) \ge \Lambda_1(\infty) i 2^{-k}\},\\ x_0^{(j)}(k) :=& -\infty,\\ x_{2^k}^{(j)}(k) :=& \infty.\end{aligned} \tag{14}$$

Furthermore, let $x_0^{(j)}(0) = -\infty$ and $x_1^{(j)}(0) = \infty$. Note that

$$\Lambda_j(x_{i+1}^{(j)}(k)-) - \Lambda_j(x_i^{(j)}(k)) \le \Lambda_1(\infty)2^{-k}. \tag{15}$$

Each partition will consist of disjoint boxes, whose vertices are described by the upper coordinates. To simplify the identification of such a partition, we will classify these into different *qualities*. Note that the construction below is similar to the one used by Marinucci (2005).

Definition 1. i) We denote by $\mathcal{A}_{k_1,\dots,k_p}$ the partition of $\mathbb{R}^p$ whose elements have the form

$$(x^{(1)}_{i_1}(k_1), x^{(1)}_{i_1+1}(k_1)] \times \ldots \times (x^{(p)}_{i_p}(k_p), x^{(p)}_{i_p+1}(k_p)],$$

$0 \le i_j \le 2^{k_j - 1}$.
ii) We call a partition $\mathcal{A}_{k_1,\dots,k_p}$ a *partition of quality* k, if $\max_{1\le j\le p} k_j = k$.

The number of partitions of quality k can be calculated as

$$\begin{aligned}&\binom{p}{1}k^{p-1} + \binom{p}{2}k^{p-2} + \ldots + \binom{p}{p}k^0 \\ =&(k+1)^p - k^p. \end{aligned} \tag{16}$$

By (16), the total number of all partitions of quality less or equal K is given by

$$\sum_{k=1}^{K} (k+1)^p - k^p = (K+1)^p - 1. \tag{17}$$

For simplicity we denote these $(K+1)^p - 1$ partitions by $\mathcal{A}_1, \ldots, \mathcal{A}_{(K+1)^p-1}$. For $x \in \mathbb{R}^p$ let

$$\begin{aligned} a_x(K) &:= (x^{(1)}_{i_K(x^{(1)})}(K), \ldots, x^{(p)}_{i_K(x^{(p)})}(K)) \\ b_x(K) &:= (x^{(1)}_{i_K(x^{(1)})+1}(K), \ldots, x^{(p)}_{i_K(x^{(p)})+1}(K)), \end{aligned}$$

where for $1 \le k \le K$ and $1 \le j \le p$ $i_k(x^{(j)})$ denotes those index satisfying

$$x^{(j)}_{i_k(x^{(j)})}(k) \le x^{(j)} \le x^{(j)}_{i_k(x^{(j)})+1}(k). \tag{18}$$

Lemma 2. *For each $x \in \mathbb{R}^p$ there exist disjoint sets $A_l(x) \in \mathcal{A}_l$ such that*

$$\bigcup_{1\le l\le (K+1)^p-1} A_l(x) = \{y \in \mathbb{R}^p : y \le a_x(K)\}. \tag{19}$$

Proof. For $x^{(j)}$, $1 \le j \le p$, choose $i_k(x^{(j)})$, $1 \le k \le K$ as in (18). This yields

$$-\infty = x^{(j)}_{i_0(x^{(j)})}(0) \le x^{(j)}_{i_1(x^{(j)})}(1) \le \ldots \le x^{(j)}_{i_K(x^{(j)})}(K) \le x^{(j)}.$$

We define the sets $A_l(x)$ by

$$\bigtimes_{j=1}^{p} (x^{(j)}_{i_{l_j}(x^{(j)})}(l_j), x^{(j)}_{i_{l_j+1}(x^{(j)})}(l_j+1)], \quad 0 \le l_j \le K-1.$$

These sets are disjoint since if there exist an element $y \in A_l(x) \cap A_k(x)$ one have $x^{(j)}_{i_{l_j}(x^{(j)})}(l_j) \le y^{(j)} \le x^{(j)}_{i_{l_j+1}(x^{(j)})}(l_j+1)$ and $x^{(j)}_{i_{k_j}(x^{(j)})}(k_j) \le y^{(j)} \le x^{(j)}_{i_{k_j+1}(x^{(j)})}(k_j+1)$. This implies $i_{l_j}(x^{(j)}) = i_{k_j}(x^{(j)})$ and therefore $A_l(x) = A_k(x)$. ☐

4.1 Proofs of Theorem 1 and 2

For any $x \in \mathbb{R}^p$ and for any measurable $A \subset \mathbb{R}^p$ we define

$$S_N(n, x) := d_N^{-1} \sum_{j=1}^{n} \left(1_{\{Y_j \le x\}} - F(x) - \frac{J_m(x)}{m!} H_m(X_j) \right)$$

$$S_N(n, A) := d_N^{-1} \sum_{j=1}^{n} \left(1_{\{Y_j \in A\}} - P(Y_j \in A) - \frac{J_m(A)}{m!} H_m(X_j) \right).$$

Furthermore, regarding (4) we set

$$J_q(A) := E(1_{\{Y_j \in A\}} H_q(X_j)) \tag{20}$$

Since $J_m(A) + J_m(B) = J_m(A \cup B)$ for disjoint sets A and B, we can use (19) to get the following representation

$$S_N(n, x) = \sum_{l=1}^{(K+1)^p - 1} S_N(n, A_l(x)) + S_N(n, a_x(K), x). \tag{21}$$

Lemma 3 will give us a second order moment bound for $S_N(n, A)$. It is due to Lemma 3.1. by Dehling and Taqqu (1989).

Lemma 3. *There exist constants $\gamma > 0$ and $C > 0$ such that for all $A \subset \mathbb{R}^p$, $n \le N$*

$$E\left|S_N(n, A)\right|^2 \le C \left(\frac{n}{N}\right) N^{-\gamma} P(Y_1 \in A).$$

Proof. Since the Hermite polynomials form an orthogonal basis of $L_2(\varphi(x)dx)$, the representation

$$1_{\{Y_j \in A\}} - P(Y_1 \in A) = \sum_{q=1}^{\infty} \frac{J_q(A)}{q!} H_q(X_j)$$

yields

$$\begin{aligned} &\sum_{q=1}^{\infty} \frac{J_q^2(A)}{q!} \\ =&E\left(1_{\{Y_j \in A\}} - P(Y_1 \in A)\right)^2 \\ =&(1 - P(Y_1 \in A))^2 P(Y_1 \in A) + P(Y_1 \in A)^2 (1 - P(Y_1 \in A))) \\ \le&P(Y_1 \in A). \end{aligned}$$

Now along the lines of Lemma 3.1. by Dehling and Taqqu (1989) we get

$$\begin{aligned} E\left|S_N(n, A)\right|^2 \le& P(Y_1 \in A) d_N^{-2} \sum_{j,k \le n} |r(j-k)|^{m+1} \\ \le& C P(Y_1 \in A) \left(\frac{n}{N}\right) N^{mD-1 \vee -D} L(n) L(N)^{-m}, \end{aligned}$$

which completes the proof. ∎

For $\Lambda(x)$ defined by (13) let

$$\Lambda_j(x^{(j)}) := F_j(x^{(j)}) + \int\limits_{\{G_j(s)\le x^{(j)}\}} \frac{|H_m(s)|}{m!}\varphi(s)ds.$$

Lemma 4. *The increments* $(m!)^{-1}J_m(x,y) := (m!)^{-1}(J_m(y) - J_m(x))$ *and* $F(x,y) := F(y) - F(x)$ *are bounded by* $\Lambda(y) - \Lambda(x)$ *for all* $x \le y$.

Proof. Let $x \le y$ and set $B := \{s \in \mathbb{R} : G(s) \le y, G(s) \not\le x\}$ and $B_j := \{s \in \mathbb{R} : x^{(j)} \le G_j(s) \le y^{(j)}\}$. Since $B \subset \bigcup_{j=1}^{p} B_j$ we have

$$\begin{aligned}
J_m(y) - J_m(x) &= \int\limits_{\{G(s)\le y\}} H_m(s)\varphi(s)ds - \int_{\{G(s)\le x\}} H_m(s)\varphi(s)ds \\
&= \int\limits_B H_m(s)\varphi(s)ds \\
&\le \int\limits_B |H_m(s)|\varphi(s)ds \\
&\le \sum_{j=1}^{p} \int\limits_{B_j} |H_m(s)|\varphi(s)ds \\
&\le \sum_{j=1}^{p} \Bigg(\int\limits_{\{G_j(s)\le y^{(j)}\}} |H_m(s)|\varphi(s)ds - \int_{\{G_j(s)\le x^{(j)}\}} |H_m(s)|\varphi(s)ds \Bigg) \\
&\le \sum_{j=1}^{p} (\Lambda_j(y^{(j)}) - \Lambda_j(x^{(j)})) \\
&= \Lambda(y) - \Lambda(x).
\end{aligned}$$

Moreover,

$$\begin{aligned}
F(y) - F(x) &\le \sum_{j=1}^{p} (F_j(y^{(j)}) - F_j(x^{(j)})) \\
&\le \sum_{j=1}^{p} (\Lambda_j(y^{(j)}) - \Lambda_j(x^{(j)})) \\
&= \Lambda(y) - \Lambda(x).
\end{aligned}$$

□

Lemma 5. *There exist constants* $\rho, C > 0$ *such that for all* $n \le N$ *and* $0 < \varepsilon \le 1$

$$P\left(\sup_{x\in\mathbb{R}^p} |S_N(n,x)| > \varepsilon \right) \le CN^{-\rho} \left(\left(\frac{n}{N}\right) \varepsilon^{-3} + \left(\frac{n}{N}\right)^{2-mD} \right).$$

Proof. Since we want to use representation (21), we will start by bounding $|S_N(n, a_x(K), x)|$.

$$
\begin{aligned}
&|S_N(n, a_x(K), x)| \\
=&\left| d_N^{-1} \sum_{j=1}^{n} \Big(\big(1_{\{Y_j \le x\}} - 1_{\{Y_j \le a_x(K)\}}\big) - F(a_x(K), x) - \frac{1}{m!} J_m(a_x(K), x) H_m(X_j) \Big) \right| \\
\le& d_N^{-1} \sum_{j=1}^{n} \Big(\big(1_{\{Y_j < b_x(K)\}} - 1_{\{Y_j \le a_x(K)\}}\big) + F(a_x(K), b_x(K)-) \Big) \\
&+ \frac{1}{m!} J_m(a_x(K), b_x(K)-) d_N^{-1} \Big| \sum_{j=1}^{n} H_m(X_j) \Big| \\
\le& |S_N(n, a_x(K), b_x(K)-)| + 2n d_N^{-1} F(a_x(K), b_x(K)-) \\
&+ \frac{2}{m!} d_N^{-1} J_m(a_x(K), b_x(K)-) \Big| \sum_{j=1}^{n} H_m(X_j) \Big| \\
\le& |S_N(n, a_x(K), b_x(K)-)| + 2n d_N^{-1} \Lambda(a_x(K), b_x(K)-) \\
&+ 2 d_N^{-1} \Lambda(a_x(K), b_x(K)-) \Big| \sum_{j=1}^{n} H_m(X_j) \Big| \\
\le& |S_N(n, a_x(K), b_x(K)-)| + 2n d_N^{-1} p \Lambda_1(\infty) 2^{-K} + 2 d_N^{-1} p \Lambda_1(\infty) 2^{-K} \Big| \sum_{j=1}^{n} H_m(X_j) \Big| \qquad (22)
\end{aligned}
$$

By using (21), (22) and $\sum_{l=1}^{\infty} \varepsilon/(l+4)^2 < \varepsilon/4$ we get

$$
\begin{aligned}
& P\left(\sup_{x \in \mathbb{R}^p} |S_N(n, x)| > \varepsilon \right) \\
\le & \sum_{l=1}^{(K+1)^p - 1} P\left(\max_{x \in \mathbb{R}^p} |S_N(n, A_l(x))| > \varepsilon/(l+4)^2 \right) \\
& + P\left(\sup_{x \in \mathbb{R}^p} |S_N(n, a_x(K), b_x(K)-)| > \varepsilon/4 \right) \\
& + P\Bigg(2 d_N^{-1} p \Lambda_1(\infty) 2^{-K} \Big| \sum_{j=1}^{n} H_m(X_j) \Big| > \varepsilon/2 - 2n d_N^{-1} p \Lambda_1(\infty) 2^{-K} \Bigg). \qquad (23)
\end{aligned}
$$

Remember, the elements $A_l(1), \dots A_l(|\mathcal{A}_l|)$ of $\mathcal{A}_l$ are disjoint. Therefore Lemma 3 yields

$$
\begin{aligned}
& P\left(\max_{x \in \mathbb{R}^p} |S_N(n, A_l(x))| > \varepsilon/(l+4)^2 \right) \\
\le & \sum_{i=1}^{|\mathcal{A}_l|} P\left(|S_N(n, A_l(i))| > \varepsilon/(l+4)^2 \right) \\
\le & \sum_{i=1}^{|\mathcal{A}_l|} (l+4)^4 \varepsilon^{-2} E\, |S_N(n, A_l(i))|^2
\end{aligned}
$$

$$
\begin{aligned}
&\le C\left(\frac{n}{N}\right) N^{-\gamma}(l+4)^4\varepsilon^{-2}\sum_{i=1}^{|\mathcal{A}_l|} P(Y_1 \in A_l(i)) \\
&= C\left(\frac{n}{N}\right) N^{-\gamma}(l+4)^4\varepsilon^{-2},
\end{aligned} \tag{24}
$$

for $1 \le l \le (K+1)^p - 1$. The next-to-last summand can be bounded as follows. Similar to (19) and (21) each partition of quality K contains one element $B_l(x)$ such that

$$
S_N(n, a_x(K), b_x(K)-) = \sum_{l=1}^{(K+1)^p - K^p} S_N(n, B_l(x)-).
$$

With respect to (24) we get

$$
\begin{aligned}
&P\left(\sup_{x\in\mathbb{R}^p} |S_N(n, a_x(K), b_x(K)-)| > \varepsilon/4\right) \\
\le &\sum_{l=1}^{(K+1)^p - K^p} P\left(\max_{x\in\mathbb{R}^p} |S_N(n, B_l(x)-)| > \varepsilon/(4(l+4)^2)\right) \\
\le &C\left(\frac{n}{N}\right) N^{-\gamma}\varepsilon^{-2}\sum_{l=1}^{(K+1)^p - K^p} (l+4)^4.
\end{aligned} \tag{25}
$$

Now let

$$
K = \left\lceil \log_2\left(\frac{8p\Lambda_1(\infty)}{\varepsilon} N d_N^{-1}\right)\right\rceil.
$$

This choice implies

$$
\begin{aligned}
&2Nd_N^{-1}p\Lambda_1(\infty)2^{-K} \le \frac{\varepsilon}{4} \\
&\left(\frac{\varepsilon}{4}\right)^{-2} \le N^{-2}d_N^2(p\Lambda_1(\infty))^{-2}2^{2K-2}
\end{aligned}
$$

and therefore

$$
\begin{aligned}
&P\Bigg(2d_N^{-1}p\Lambda_1(\infty)2^{-K}\Big|\sum_{j=1}^{n} H_m(X_j)\Big| > \frac{\varepsilon}{2} - 2nd_N^{-1}p\Lambda_1(\infty)2^{-K}\Bigg) \\
\le &P\Bigg(2d_N^{-1}p\Lambda_1(\infty)2^{-K}\Big|\sum_{j=1}^{n} H_m(X_j)\Big| > \frac{\varepsilon}{4}\Bigg) \\
\le &P\Bigg(d_N^{-1}\Big|\sum_{j=1}^{n} H_m(X_j)\Big| > \frac{\varepsilon}{4}\cdot\frac{2^{K-1}}{p\Lambda_1(\infty)}\Bigg) \\
\le &d_N^{-2}E\Big|\sum_{j=1}^{n} H_m(X_j)\Big|^2\left(\frac{\varepsilon}{4}\right)^{-2}2^{-2K+2}(p\Lambda_1(\infty))^2 \\
\le &C\left(\frac{d_n}{d_N}\right)^2 N^{-2}d_N^2
\end{aligned}
$$

$$
\begin{aligned}
&\le C\left(\frac{n}{N}\right)^{2-mD}\left(\frac{L(n)}{L(N)}\right)^{m} N^{-mD}L^m(N)\\
&\le C\left(\frac{n}{N}\right)^{2-mD} N^{-mD+\lambda},
\end{aligned} \tag{26}
$$

for any $\lambda > 0$. By using (23), (24), (25), (26) we get

$$
\begin{aligned}
&P\left(\sup_{x\in\mathbb{R}^p} |S_N(n,x)| > \varepsilon\right)\\
&\le C\left(\frac{n}{N}\right) N^{-\gamma}\varepsilon^{-2}\left(\sum_{l=1}^{(K+1)^p-1}(l+4)^4 + \sum_{l=1}^{(K+1)^p-K^p}(l+4)^4\right) + C\left(\frac{n}{N}\right)^{2-mD} N^{-mD+\lambda}\\
&\le C\left(\frac{n}{N}\right) N^{-\gamma}\varepsilon^{-2}K^{5p} + C\left(\frac{n}{N}\right)^{2-mD} N^{-mD+\lambda}
\end{aligned} \tag{27}
$$

Since

$$
\begin{aligned}
K^{5p} &\le C\left(\log(\varepsilon^{-1})^{5p} + \log(N)^{5p}\right)\\
&\le C\varepsilon^{-1}N^{\delta}
\end{aligned}
$$

for any $\delta > 0$, (27) is bounded by

$$
CN^{(-\gamma+\delta)\vee(-mD+\lambda)}\left(\left(\frac{n}{N}\right)\varepsilon^{-3} + \left(\frac{n}{N}\right)^{2-mD}\right),
$$

which completes the proof. □

To prove Theorem 1 and 2 we can use the proofs which were given by Dehling and Taqqu (1989) in the case of one dimensional observations.

4.2 Proofs of Theorem 3 and 4

For simplicity we assume that $p = q$, i.e. $G : \mathbb{R}^p \to \mathbb{R}^p$. Since the main idea for proving Theorem 4 was already used in the proof of Theorem 2. Therefore, we will use some modified definitions and notations from section 4.1. From now on let

$$
\begin{aligned}
S_N(n,x) &:= d_N^{-1}\sum_{j=1}^{n}\left(1_{\{Y_j\le x\}} - F(x) - \sum_{l_1+\ldots+l_p=m}\frac{J_{l_1,\ldots,l_p}(x)}{l_1!\cdots l_p!}H_{l_1,\ldots,l_p}(X_j)\right),\\
S_N(n,A) &:= d_N^{-1}\sum_{j=1}^{n}\left(1_{\{Y_j\in A\}} - P(Y_j\in A) - \sum_{l_1+\ldots+l_p=m}\frac{J_{l_1,\ldots,l_p}(A)}{l_1!\cdots l_p!}H_{l_1,\ldots,l_p}(X_j)\right),\\
J_{l_1,\ldots,l_p}(A) &:= E(1_{\{Y_j\in A\}}H_{l_1,\ldots,l_p}(X_j)).
\end{aligned}
$$

For $\Lambda(x)$ defined in (13) replace the terms of the sum by

$$
\Lambda_j(x^{(j)}) := P(G_j(X_1)\le x^{(j)}) + \sum_{l_1+\ldots+l_p=m}\ \int\limits_{\substack{\{s\in\mathbb{R}^p:\\ G_j(s)\le x^{(j)}\}}}\frac{|H_{l_1,\ldots,l_p}(s)|}{l_1!\cdots l_p!}\varphi(s)ds^{(1)}\ldots ds^{(p)},
$$

where φ denotes the p-dimensional standard normal distribution, and define the chaining points $x_i^{(j)}(k)$ analogous to (14). The partitions $\mathcal{A}_{k_1,\ldots,k_p}$ are given as in Definition 1. Therefore, Lemma 2 and representation (21) still hold, i.e.

$$S_N(n,x) = \sum_{l=1}^{(K+1)^p-1} S_N(n, A_l(x)) + S_N(n, a_x(K), x). \tag{28}$$

Lemma 6. *The increments $J_m(x,y)$ and $F(x,y)$, where*

$$\begin{aligned} J_m(x,y) :=& \sum_{l_1+\ldots+l_p=m} \frac{J_{l_1,\ldots,l_p}(y) - J_{l_1,\ldots,l_p}(x)}{l_1!\cdots l_p!} \\ F(x,y) :=& F(y) - F(x), \end{aligned}$$

are both bounded by $\Lambda(x,y) := \Lambda(y) - \Lambda(x)$ for all $x \le y$.

Lemma 6 can be proven in the same way as Lemma 4. Lemma 7 is due to Arcones (1994, Lemma 1).

Lemma 7. *Let $X = (X^{(1)}, \ldots, X^{(p)})$ and $Y = (Y^{(1)}, \ldots, Y^{(p)})$ be two mean-zero Gaussian random vectors on $\mathbb{R}^p$. Assume that*

$$EX^{(i)}X^{(j)} = EY^{(i)}Y^{(j)} = \delta_{i,j} \tag{29}$$

for each $1 \le i, j \le p$. We define

$$r^{(i,j)} := EX^{(i)}Y^{(j)}.$$

Let f be a function on $\mathbb{R}^p$ with finite second moment and Hermite rank m, $1 \le m < \infty$, with respect to X. Suppose that

$$\psi := \left(\sup_{1\le i\le p} \sum_{j=1}^p |r^{(i,j)}| \right) \vee \left(\sup_{1\le j\le p} \sum_{i=1}^p |r^{(i,j)}| \right) \le 1.$$

Then

$$|E(f(X) - Ef(X))(f(Y) - Ef(Y))| \le \psi^m Ef(X)^2.$$

Lemma 8. *There exist constants $\gamma > 0$ and $C > 0$ such that for all measurable $A \subset \mathbb{R}^p$ and $n \le N$*

$$E|S_N(n,A)|^2 \le C\left(\frac{n}{N}\right) N^{-\gamma} P(Y_1 \in A).$$

Proof. Let

$$f(\cdot) := 1_{\{G(\cdot)\in A\}} - P(Y_1 \in A) - \sum_{l_1+\ldots+l_p=m} \frac{J_{l_1,\ldots,l_p}(A)}{l_1!\cdots l_p!} H_{l_1,\ldots,l_p}(\cdot)$$

$$\psi(k) := \left(\max_{1\leq i\leq p} \sum_{j=1}^{p} |r^{(i,j)}(k)| \right) \vee \left(\max_{1\leq j\leq p} \sum_{i=1}^{p} |r^{(i,j)}(k)| \right).$$

Since $r^{(i,j)}(k)$ converges to 0 for all $1 \leq i, j \leq p$ if k tends to infinity, we can find $b \in \mathbb{N}$ such that $\psi(kb) \leq 1$ for all $k \geq 1$. Note that f has Hermite rank $m+1$. Therefore, as Arcones (1994, p. 2249) we can apply Lemma 7 as follows

$$\begin{aligned}
&E\left|\sum_{j=1}^{n} f(X_j)\right|^2\\
=&E\left|\sum_{j=1}^{b} \sum_{k=1}^{\lfloor n-j+b/b\rfloor} f(X_{(k-1)b+j})\right|^2\\
\leq&C\sum_{j=1}^{b} \sum_{k,l=1}^{\lfloor n-j+b/b\rfloor} E\left(f(X_{(k-1)b+j})f(X_{(l-1)b+j})\right)\\
\leq&C\sum_{j=1}^{b} \sum_{k,l=1}^{\lfloor n-j+b/b\rfloor} \psi(b(k-l))^{m+1} Ef(X_1)^2\\
\leq&CEf(X_1)^2 \sum_{k,l=1}^{n} \psi(b(k-l))^{m+1}
\end{aligned}$$

By using

$$\begin{aligned}
&Ef(X_1)^2\\
=&\sum_{\substack{k_1+\ldots+k_p\geq m+1\\ l_1+\ldots+l_p\geq m+1}} J_{k_1,\ldots,k_p} J_{l_1,\ldots,l_p} \prod_{j=1}^{p} (k_j! l_j!)^{-1}\\
&\qquad\cdot E\Big(H_{k_1}(X_1^{(1)}) \cdots H_{k_p}(X_1^{(p)}) H_{l_1}(X_1^{(1)}) \cdots H_{l_p}(X_1^{(p)})\Big)\\
=&\sum_{k_1+\ldots+k_p\geq m+1} J^2_{k_1,\ldots,k_p} \prod_{j=1}^{p} (k_j!)^{-2} E\left(H_{k_1}(X_1^{(1)})\right)^2 \cdots E\left(H_{k_p}(X_1^{(p)})\right)^2\\
\leq&\sum_{k_1,\ldots,k_p=0} J^2_{k_1,\ldots,k_p} \prod_{j=1}^{p} (k_j!)^{-1}\\
=&E\left(1_{\{G(X_1)\in A\}} - P(Y_1 \in A)\right)^2\\
\leq&P(Y_1 \in A)
\end{aligned}$$

and

$$\psi(bk)$$

$$
\begin{aligned}
&\leq \sum_{i,j=1}^{p} \left| r^{(i,j)}(bk) \right| \\
&= \sum_{i,j=1}^{p} \left| c_{(i,j)} L(bk)(bk)^{-D} \right| \\
&\leq C \left| L'(k) k^{-D} \right|,
\end{aligned}
$$

where $L'(k) := L(bk)$ is slowly varying at infinity, we get

$$
E \left| \sum_{j=1}^{n} f(X_j) \right|^2 \leq C P(Y_1 \in A) \sum_{k,l=1}^{n} \left| L'(k-l)(k-l)^{-D} \right|^{m+1}. \tag{30}
$$

The remaining parts of the proof can be found in Dehling and Taqqu (1989, p. 1777) □

Lemma 9. *For all $m \in \mathbb{N}$ there exists a constant $C > 0$ such that for all $l_1, \dots, l_p \in \mathbb{N}$, $l_1 + \dots + l_p = m$, and $n \in \mathbb{N}$*

$$
E \left| \sum_{j=1}^{n} H_{l_1,\dots,l_p}(X_j) \right|^2 \leq C d_n^2
$$

Proof. By using (30) with $f = H_{l_1,\dots,l_p}$ we get

$$
\begin{aligned}
&E \left| \sum_{j=1}^{n} H_{l_1,\dots,l_p}(X_j) \right|^2 \\
\leq & C E H_{l_1,\dots,l_p}(X_1)^2 \sum_{k,l=1}^{n} \left| L(b(k-l))^m (k-l)^{-mD} \right| \\
\leq & C n \sum_{k=1}^{n} \left| L(bk)^m k^{-mD} \right| \\
\leq & C n^{2-mD} L(bn)^m \\
\leq & C d_n^2.
\end{aligned}
$$

□

Lemma 10. *There exist constants $\rho, C > 0$ such that for all $n \leq N$ and $0 < \varepsilon \leq 1$*

$$
P\left(\sup_{x \in \mathbb{R}^p} |S_N(n,x)| > \varepsilon \right) \leq C N^{-\rho} \left(\left(\frac{n}{N} \right) \varepsilon^{-3} + \left(\frac{n}{N} \right)^{2-mD} \right).
$$

Proof. Since we want to use representation (28), we will start by bounding $|S_N(n, a_x(K), x)|$.

$$
\begin{aligned}
&|S_N(n, a_x(K), x)| \\
=&\left| d_N^{-1} \sum_{j=1}^{n} \Big(\big(1_{\{Y_j \le x\}} - 1_{\{Y_j \le a_x(K)\}}\big) - F(a_x(K), x) - \sum_{l_1+\ldots+l_p=m} \frac{J_{l_1,\ldots,l_p}(a_x(K), x)}{l_1! \cdots l_p!} H_{l_1,\ldots,l_p}(X_j) \Big) \right| \\
\le& d_N^{-1} \sum_{j=1}^{n} \Big(\big(1_{\{Y_j < b_x(K)\}} - 1_{\{Y_j \le a_x(K)\}}\big) + F(a_x(K), b_x(K)-) \Big) \\
&+ \sum_{l_1+\ldots+l_p=m} \frac{J_{l_1,\ldots,l_p}(a_x(K), b_x(K)-)}{l_1! \cdots l_p!} d_N^{-1} \Big| \sum_{j=1}^{n} H_{l_1,\ldots,l_p}(X_j) \Big| \\
\le& |S_N(n, a_x(K), b_x(K)-)| + 2n d_N^{-1} F(a_x(K), b_x(K)-) \\
&+ 2 d_N^{-1} \sum_{l_1+\ldots+l_p=m} \frac{J_{l_1,\ldots,l_p}(a_x(K), b_x(K)-)}{l_1! \cdots l_p!} \Big| \sum_{j=1}^{n} H_{l_1,\ldots,l_p}(X_j) \Big| \\
\le& |S_N(n, a_x(K), b_x(K)-)| + 2n d_N^{-1} \Lambda(a_x(K), b_x(K)-) \\
&+ 2 d_N^{-1} \Lambda(a_x(K), b_x(K)-) \sum_{l_1+\ldots+l_p=m} \Big| \sum_{j=1}^{n} H_{l_1,\ldots,l_p}(X_j) \Big| \\
\le& |S_N(n, a_x(K), b_x(K)-)| + 2n d_N^{-1} p \Lambda_1(\infty) 2^{-K} \\
&+ 2 d_N^{-1} p \Lambda_1(\infty) 2^{-K} \sum_{l_1+\ldots+l_p=m} \Big| \sum_{j=1}^{n} H_{l_1,\ldots,l_p}(X_j) \Big|
\end{aligned} \tag{31}
$$

By using (28), (31) and $\sum_{l=1}^{\infty} \varepsilon/(l+4)^2 < \varepsilon/4$ we get

$$
\begin{aligned}
& P\left(\sup_{x \in \mathbb{R}^p} |S_N(n, x)| > \varepsilon \right) \\
\le& \sum_{l=1}^{(K+1)^p - 1} P\left(\max_{x \in \mathbb{R}^p} |S_N(n, A_l(x))| > \varepsilon/(l+4)^2 \right) \\
&+ P\left(\sup_{x \in \mathbb{R}^p} |S_N(n, a_x(K), b_x(K)-)| > \varepsilon/4 \right) \\
&+ P\Bigg(2 d_N^{-1} p \Lambda_1(\infty) 2^{-K} \sum_{l_1+\ldots+l_p=m} \Big| \sum_{j=1}^{n} H_{l_1,\ldots,l_p}(X_j) \Big| > \varepsilon/2 - 2n d_N^{-1} p \Lambda_1(\infty) 2^{-K} \Bigg).
\end{aligned} \tag{32}
$$

Remember, the elements $A_l(1), \ldots A_l(|\mathcal{A}_l|)$ of $\mathcal{A}_l$ are disjoint. Therefore Lemma 8 yields

$$
\begin{aligned}
& P\left(\max_{x \in \mathbb{R}^p} |S_N(n, A_l(x))| > \varepsilon/(l+4)^2 \right) \\
\le& \sum_{i=1}^{|\mathcal{A}_l|} P\left(|S_N(n, A_l(i))| > \varepsilon/(l+4)^2 \right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{|\mathcal{A}_l|} (l+4)^4 \varepsilon^{-2} E\left|S_N(n, A_l(i))\right|^2 \\
&\leq C\left(\frac{n}{N}\right) N^{-\gamma}(l+4)^4 \varepsilon^{-2} \sum_{i=1}^{|\mathcal{A}_l|} P(Y_1 \in A_l(i)) \\
&= C\left(\frac{n}{N}\right) N^{-\gamma}(l+4)^4 \varepsilon^{-2},
\end{aligned} \tag{33}
$$

for $1 \leq l \leq (K+1)^p - 1$. The next-to-last summand in (32) can be bounded as follows. Similar to (19) and (28) each partition of quality K contains one element $B_l(x)$ such that

$$
S_N(n, a_x(K), b_x(K)-) = \sum_{l=1}^{(K+1)^p - K^p} S_N(n, B_l(x)-).
$$

With respect to (33) we get

$$
\begin{aligned}
&P\left(\sup_{x \in \mathbb{R}^p} \left|S_N(n, a_x(K), b_x(K)-)\right| > \varepsilon/4\right) \\
\leq &\sum_{l=1}^{(K+1)^p - K^p} P\left(\max_{x \in \mathbb{R}^p} \left|S_N(n, B_l(x)-)\right| > \varepsilon/(4(l+4)^2)\right) \\
\leq &C\left(\frac{n}{N}\right) N^{-\gamma} \varepsilon^{-2} \sum_{l=1}^{(K+1)^p - K^p} (l+4)^4.
\end{aligned} \tag{34}
$$

Now let

$$
M = \left|\{(l_1, \ldots, l_p) \in \mathbb{N}^p : l_1 + \ldots l_p = m\}\right|
$$

and

$$
K = \left\lceil \log_2\left(\frac{8p\Lambda_1(\infty)}{\varepsilon} N d_N^{-1}\right)\right\rceil.
$$

This choice implies

$$
\begin{aligned}
&2Nd_N^{-1} p\Lambda_1(\infty) 2^{-K} \leq \frac{\varepsilon}{4} \\
&\left(\frac{\varepsilon}{4}\right)^{-2} \leq N^{-2} d_N^2 (p\Lambda_1(\infty))^{-2} 2^{2K-2}
\end{aligned}
$$

and together with Lemma 9 we obtain

$$
\begin{aligned}
&P\left(2d_N^{-1} p\Lambda_1(\infty) 2^{-K} \sum_{l_1 + \ldots + l_p = m} \left|\sum_{j=1}^{n} H_{l_1, \ldots, l_p}(X_j)\right| > \varepsilon/2 - 2nd_N^{-1} p\Lambda_1(\infty) 2^{-K}\right) \\
\leq &P\left(2d_N^{-1} p\Lambda_1(\infty) 2^{-K} \sum_{l_1 + \ldots + l_p = m} \left|\sum_{j=1}^{n} H_{l_1, \ldots, l_p}(X_j)\right| > \varepsilon/4\right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq P\Bigg(d_N^{-1}\sum_{l_1+\ldots+l_p=m}\Big|\sum_{j=1}^{n}H_{l_1,\ldots,l_p}(X_j)\Big| > \frac{\varepsilon}{4}\cdot\frac{2^{K-1}}{p\Lambda_1(\infty)}\Bigg)\\
&\leq \sum_{l_1+\ldots+l_p=m}P\Bigg(d_N^{-1}\Big|\sum_{j=1}^{n}H_{l_1,\ldots,l_p}(X_j)\Big| > \frac{\varepsilon}{4}\cdot\frac{2^{K-1}}{p\Lambda_1(\infty)M}\Bigg)\\
&\leq d_N^{-2}\sum_{l_1+\ldots+l_p=m}E\Big|\sum_{j=1}^{n}H_{l_1,\ldots,l_p}(X_j)\Big|^2\left(\frac{\varepsilon}{4}\right)^{-2}2^{-2K+2}(p\Lambda_1(\infty))^2M^2\\
&\leq C\left(\frac{d_n}{d_N}\right)^2 N^{-2}d_N^2\\
&\leq C\left(\frac{n}{N}\right)^{2-mD}\left(\frac{L(n)}{L(N)}\right)^m N^{-mD}L^m(N)\\
&\leq C\left(\frac{n}{N}\right)^{2-mD}N^{-mD+\lambda},
\end{aligned}
\tag{35}
$$

for any $\lambda > 0$. By using (32), (33), (34), (35) we get

$$
\begin{aligned}
&P\left(\sup_{x\in\mathbb{R}^p}|S_N(n,x)| > \varepsilon\right)\\
\leq &C\left(\frac{n}{N}\right)N^{-\gamma}\varepsilon^{-2}\left(\sum_{l=1}^{(K+1)^p-1}(l+4)^4+\sum_{l=1}^{(K+1)^p-K^p}(l+4)^4\right)+C\left(\frac{n}{N}\right)^{2-mD}N^{-mD+\lambda}\\
\leq &C\left(\frac{n}{N}\right)N^{-\gamma}\varepsilon^{-2}K^{5p}+C\left(\frac{n}{N}\right)^{2-mD}N^{-mD+\lambda}
\end{aligned}
\tag{36}
$$

Since

$$
\begin{aligned}
K^{5p}\leq &C\left(\log(\varepsilon^{-1})^{5p}+\log(N)^{5p}\right)\\
\leq &C\varepsilon^{-1}N^{\delta}
\end{aligned}
$$

for any $\delta > 0$, (36) is bounded by

$$
CN^{(-\gamma+\delta)\vee(-mD+\lambda)}\left(\left(\frac{n}{N}\right)\varepsilon^{-3}+\left(\frac{n}{N}\right)^{2-mD}\right),
$$

which completes the proof. $\square$

Proof of Theorem 4. Lemma 10 corresponds to Lemma 3.2 by Dehling and Taqqu (1989). Therefore the proof of Theorem 4 is the same as in the one-dimensional case, see (Dehling and Taqqu, 1989, p.1781). $\square$

Instead of studying the partial sum process of multivariate Hermite polynomials we will deduce the asymptotics of $(R_N(x,t))$ from a linear combination of mth order univariate Hermite polynomials. Therefor we use the following Lemma.

Lemma 11. *For all $m \in \mathbb{N}$ and $a_1, \ldots, a_p \in \mathbb{R}$ with $a_1^2 + \ldots + a_p^2 = 1$ we have*

$$H_m\left(\sum_{j=1}^{p} a_j x_j\right) = \sum_{m_1+\ldots+m_p=m} \frac{m!}{m_1! \cdots m_p!} \prod_{j=1}^{p} a_j^{m_j} H_{m_j}(x_j). \tag{37}$$

Since we could not find a proof for this well known result in literature we give one here.

Proof. We first show that all partial derivatives are equal by using induction. For $m = 1$ this is obvious. Remember that $H_n'(x) = nH_{n-1}(x)$. Therefore we get

$$\begin{aligned}
&\frac{\partial}{\partial x_1}\left(\sum_{m_1+\ldots+m_p=m+1} \frac{(m+1)!}{m_1! \cdots m_p!} \prod_{j=1}^{p} a_j^{m_j} H_{m_j}(x_j)\right) \\
=& \sum_{m_1+\ldots+m_p=m+1} \frac{(m+1)!}{(m_1-1)! \cdots m_p!} a_1^{m_1} H_{m_1-1}(x_1) \prod_{j=2}^{p} a_j^{m_j} H_{m_j}(x_j) \\
=& a_1(m+1) \sum_{m_1+\ldots+m_p=m} \frac{(m)!}{m_1! \cdots m_p!} \prod_{j=1}^{p} a_j^{m_j} H_{m_j}(x_j) \\
=& a_1(m+1) H_m\left(\sum_{j=1}^{p} a_j x_j\right) \\
=& \frac{\partial}{\partial x_1} H_{m+1}\left(\sum_{j=1}^{p} a_j x_j\right)
\end{aligned}$$

The other derivatives can be handled similarly. Therefore (37) holds up to a constant. Let $x_1 = \ldots = x_p = 0$. If m is odd both sides of (37) are equal to zero and thus the constant vanishes. For even m we have $H_m(0) = (-1)^{m/2}(m-1)!!$, where

$$(m-1)!! := (m-1)(m-3)\cdots 3 \cdot 1 = \frac{m!}{2^{m/2}(m/2)!}.$$

This yields

$$\begin{aligned}
&\sum_{m_1+\ldots+m_p=m} \frac{m!}{m_1! \cdots m_p!} \prod_{j=1}^{p} a_j^{m_j} H_{m_j}(0) \\
=& \sum_{2m_1+\ldots+2m_p=m} \frac{m!}{(2m_1)! \cdots (2m_p)!} \prod_{j=1}^{p} (-1)^{m_j} a_j^{2m_j} (2m_j-1)!! \\
=& (-1)^{m/2} \sum_{2m_1+\ldots+2m_p=m} \frac{m!}{(2m_1)!! \cdots (2m_p)!!} \prod_{j=1}^{p} (a_j^2)^{m_j}
\end{aligned}$$

$$
\begin{aligned}
&=(-1)^{m/2} \sum_{m_1+\ldots+m_p=m/2} \frac{m!}{2^{m/2} m_1! \cdots m_p!} \prod_{j=1}^{p} (a_j^2)^{m_j} \\
&=(-1)^{m/2} \sum_{m_1+\ldots+m_p=m/2} \frac{(m-1)!! 2^{m/2} (m/2)!}{2^{m/2} m_1! \cdots m_p!} \prod_{j=1}^{p} (a_j^2)^{m_j} \\
&=(-1)^{m/2} (m-1)!! \sum_{m_1+\ldots+m_p=m/2} \frac{(m/2)!}{m_1! \cdots m_p!} \prod_{j=1}^{p} (a_j^2)^{m_j} \\
&=H_m(0) \left(\sum_{j=1}^{p} a_j^2 \right)^{m/2} \\
&=H_m(0).
\end{aligned}
$$

□

Proof of Theorem 3. By Theorem 4 it is enough to study the limit of

$$
\left\{ d_N^{-1} \sum_{j=1}^{\lfloor Nt \rfloor} \sum_{l_1+\ldots+l_p=m} \frac{J_{l_1,\ldots,l_p}(x)}{l_1! \cdots l_p!} H_{l_1,\ldots,l_p}(X_j) : (x,t) \in [-\infty, \infty]^p \times [0,1] \right\}.
$$

We first show that in the current situation Lemma 11 can be applied. For all $k_1, \ldots, k_p$ satisfying $k_1 + \ldots + k_p = m$ we can find real numbers $a_{k_1,\ldots,k_p}^{(1)}, \ldots a_{k_1,\ldots,k_p}^{(p)}$, s.t. the matrix

$$
A = \left(\prod_{i=1}^{p} (a_{k_1,\ldots,k_p}^{(i)})^{m_i} \right)_{\substack{m_1+\ldots+m_p=m \\ k_1+\ldots+k_p=m}}
$$

is invertible. After normalization we have $\sum_{i=1}^{p} (a_{k_1,\ldots,k_p}^{(i)})^2 = 1$. For a suitable diagonal-matrix M of the same size define $B := MA^{-1}$, $B = (b(k_1, \ldots, k_p, l_1, \ldots, l_p))$, s.t.

$$
\begin{aligned}
&\sum_{k_1+\ldots+k_p=m} b(k_1, \ldots, k_p, l_1, \ldots, l_p) (a_{k_1,\ldots,k_p}^{(1)})^{m_1} \cdots (a_{k_1,\ldots,k_p}^{(p)})^{m_p} \\
&= \begin{cases} (m!)^{-1} \prod_{i=1}^{p} l_i! & \text{if } (m_1, \ldots, m_p) = (l_1, \ldots, l_p) \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{38}
$$

By using Lemma 11 together with (38) we get

$$
\begin{aligned}
&\sum_{\substack{l_1+\ldots+l_p=m \\ k_1+\ldots+k_p=m}} J_{l_1,\ldots,l_p}(x) \left(\prod_{i=1}^{p} (l_i!)^{-1} \right) b(k_1, \ldots, k_p, l_1, \ldots, l_p) H_m \left(\sum_{i=1}^{p} a_{k_1,\ldots,k_p}^{(i)} X_j^{(i)} \right) \\
= &\sum_{\substack{l_1+\ldots+l_p=m \\ k_1+\ldots+k_p=m}} \sum_{m_1+\ldots+m_p=m} J_{l_1,\ldots,l_p}(x) \left(\prod_{i=1}^{p} (l_i!)^{-1} \right) b(k_1, \ldots, k_p, l_1, \ldots, l_p)
\end{aligned}
$$

$$\times\, m! \prod_{i=1}^{p} (m_i!)^{-1} \left(a^{(i)}_{k_1,\ldots,k_p}\right)^{m_i} H_{m_i}\left(X_j^{(i)}\right)$$

$$= \sum_{l_1+\ldots+l_p=m} J_{l_1,\ldots,l_p}(x) \prod_{i=1}^{p} (l_i!)^{-1} H_{l_i}\left(X_j^{(i)}\right)$$

For simplicity define

$$I(x;k_1,\ldots,k_p) := \sum_{l_1+\ldots+l_p=m} J_{l_1,\ldots,l_p}(x)\left(\prod_{i=1}^{p}(l_i!)^{-1}\right) b(k_1,\ldots,k_p,l_1,\ldots,l_p)$$

so that we get the identity

$$\begin{aligned}
&d_N^{-1}\sum_{j=1}^{\lfloor Nt\rfloor}\sum_{l_1+\ldots+l_p=m}\prod_{i=1}^{p}(l_i!)^{-1}H_{l_i}\left(X_j^{(i)}\right)\\
=&d_N^{-1}\sum_{j=1}^{\lfloor Nt\rfloor}\sum_{k_1+\ldots+k_p=m} I(x;k_1,\ldots,k_p) H_m\left(\sum_{i=1}^{p} a^{(i)}_{k_1,\ldots,k_p} X_j^{(i)}\right)
\end{aligned}$$

Note that $Y_j^{(k_1,\ldots,k_p)} := \sum_{i=1}^{p} a^{(i)}_{k_1,\ldots,k_p} X_j^{(i)}$ is standard normal distributed and that

$$\begin{aligned}
&\int_{\mathbb{R}^m}'' \frac{e^{it(x_1+\ldots+x_m)}-1}{i(x_1+\ldots+x_m)} \prod_{j=1}^{m} |x_j|^{-(1-D)/2} \left(\sum_{i=1}^{p} a^{(i)}_{k_1,\ldots,k_p} B^{(i)}\right)(dx_1)\cdots\left(\sum_{i=1}^{p} a^{(i)}_{k_1,\ldots,k_p} B^{(i)}\right)(dx_m)\\
=&\sum_{j_1,\ldots,j_m=1}^{p} a^{(j_1)}_{k_1,\ldots,k_p}\cdots a^{(j_m)}_{k_1,\ldots,k_p} \int_{\mathbb{R}^m}'' \frac{e^{it(x_1+\ldots+x_m)}-1}{i(x_1+\ldots+x_m)} \prod_{j=1}^{m} |x_j|^{-(1-D)/2} B^{(j_1)}(dx_1)\cdots B^{(j_m)}(dx_m)\\
:=&\sum_{j_1,\ldots,j_m=1}^{p} a^{(j_1)}_{k_1,\ldots,k_p}\cdots a^{(j_m)}_{k_1,\ldots,k_p} Z_{j_1,\ldots,j_m}(t).
\end{aligned}$$

Therefore, as in (11) we have

$$\begin{aligned}
&\left\{d_N^{-1}\sum_{j=1}^{\lfloor Nt\rfloor} H_m(Y_j^{(k_1,\ldots,k_p)}) : k_1,\ldots,k_p = m, t\in[0,1]\right\} \xrightarrow{d}\\
&\left\{\sum_{j_1,\ldots,j_m=1}^{p} a^{(j_1)}_{k_1,\ldots,k_p}\cdots a^{(j_m)}_{k_1,\ldots,k_p} Z_{j_1,\ldots,j_m}(t) : k_1,\ldots,k_p = m, t\in[0,1]\right\}.
\end{aligned}$$

By Dudley and Wichura's almost sure representation theorem we can find vector processes $(\tilde{S}_N(t))$ and $(\tilde{Z}(t))$, which have the same distribution as the above, s.t. $(\tilde{S}_N(t))$ converges a.s. to $(\tilde{Z}(t))$ in $D[0,1]$. Since the functions $I(x;k_1,\ldots,k_p)$ are bounded, these a.s. convergence still holds in $D([-\infty,\infty]^p\times[0,1])$ if one multiplies $I(x;k_1,\ldots,k_p)$ to the corresponding component of $(\tilde{S}_N(t))$ resp. $(\tilde{Z}(t))$ in $D[0,1]$. Applying the continious mapping theorem we get

$$\left\{ d_N^{-1} \sum_{j=1}^{\lfloor Nt \rfloor} \sum_{k_1+\ldots+k_p=m} I(x;k_1,\ldots,k_p) H_m(Y_j^{(k_1,\ldots,k_p)}) : (x,t) \in [-\infty,\infty]^p \times [0,1] \right\} \xrightarrow{d}$$

$$\left\{ \sum_{j_1,\ldots,j_m=1}^{p} \sum_{k_1+\ldots+k_p=m} I(x;k_1,\ldots,k_p) a_{k_1,\ldots,k_p}^{(j_1)} \cdots a_{k_1,\ldots,k_p}^{(j_m)} Z_{j_1,\ldots,j_m}(t) : \right.$$
$$\left. (x,t) \in [-\infty,\infty]^p \times [0,1] \right\}.$$

Finally we have to verify that this limit is equal to (12). By (38) we obtain

$$\begin{aligned}
&\sum_{j_1,\ldots,j_m=1}^{p} \sum_{k_1+\ldots+k_p=m} I(x;k_1,\ldots,k_p) a_{k_1,\ldots,k_p}^{(j_1)} \cdots a_{k_1,\ldots,k_p}^{(j_m)} \\
&= \sum_{j_1,\ldots,j_m=1}^{p} \sum_{\substack{k_1+\ldots+k_p=m \\ l_1+\ldots+l_p=m}} J_{l_1,\ldots,l_p}(x) \left(\prod_{i=1}^{p} (l_i!)^{-1} \right) b(k_1,\ldots,k_p,l_1,\ldots,l_p) a_{k_1,\ldots,k_p}^{(j_1)} \cdots a_{k_1,\ldots,k_p}^{(j_m)} \\
&= \begin{cases} (m!)^{-1} J_{l_1,\ldots,l_p}(x) & \text{if } l_i = i(j_1,\ldots,j_m) \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

But if $l_i = i(j_1,\ldots,j_m)$ for $1 \leq i \leq p$ we have

$$\tilde{J}_{j_1,\ldots,j_m} = (m!)^{-1} J_{l_1,\ldots,l_p}(x),$$

which completes the proof. $\square$

References

M. A. Arcones. Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.*, 22:2242–2274, 1994.

J. Buchsteiner. Weak convergence of the weighted sequential empirical process of some long-range dependent data. *Statist. Probab. Lett.*, 96:170–179., 2015.

H. Dehling and M. S. Taqqu. The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.*, 17:1767–1783, 1989.

R. Fox and M. S. Taqqu. Multiple stochastic integrals with dependent integrators. *J. Multivariate Anal.*, 21:105–127, 1987.

L. Giraitis and D. Surgailis. The reduction principle for the empirical process of a long memory linear process. In *Empirical process techniques for dependent data*, pages 241–255. Birkhäuser Boston, 2002.

H. C. Ho and T. C. Sun. Limiting distributions of nonlinear vector functions of stationary Gaussian processes. *Ann. Probab.*, 18:1159–1173, 1990.

S. Kechagias and V. Pipiras. Definitions and representations of multivariate long-range dependent time series. *J. Time. Ser. Anal.*, 36:1–25, 2015.

D. Marinucci. The empirical process for bivariate sequences with long memory. *Stat. Inference Stoch. Process.*, 8:205–223, 2005.

V. Pipiras and M. S. Taqqu. Regularization and integral representations of Hermite processes. *Statist. Probab. Lett.*, 80:2014–2023, 2010.

M. S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete*, 31:287–302, 1975.

M. S. Taqqu. Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete*, 50:53–83, 1979.

E. Taufer. A note on the empirical process of strongly dependent stable random variables. *ArXiv e-prints*, 1410.8050, 2014.