



# The multivariate functional de Jong CLT

Christian Döbler<sup>1</sup> · Mikołaj Kasprzak<sup>2,3</sup> · Giovanni Peccati<sup>2</sup> 

Received: 5 May 2021 / Revised: 21 December 2021 / Accepted: 17 January 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

We prove a multivariate functional version of de Jong's CLT (J Multivar Anal 34(2):275–289, 1990) yielding that, given a sequence of vectors of Hoeffding-degenerate  $U$ -statistics, the corresponding empirical processes on  $[0, 1]$  weakly converge in the Skorohod space as soon as their fourth cumulants in  $t = 1$  vanish asymptotically and a certain strengthening of the Lindeberg-type condition is verified. As an application, we lift to the functional level the 'universality of Wiener chaos' phenomenon first observed in Nourdin et al. (Ann Probab 38(5):1947–1985, 2010).

**Keywords**  $U$ -statistics · Functional limit theorems · Contractions · Product formulae · Hoeffding decompositions · Universality

**Mathematics Subject Classification** 60F17 · 60D05 · 62G20

## 1 Introduction and (simplified) statements of main results

### 1.1 Motivation

*Degenerate  $U$ -statistics* (see Sect. 2 for definitions) are the fundamental components of *Hoeffding decompositions* for generic random variables of the type  $f(X_1, \dots, X_n)$ , with  $X_1, \dots, X_n$  an independent sample, and are thus pivotal objects in stochastic analysis [20, 24, 28, 30, 44]. A special case of degenerate  $U$ -statistic is given by *homogeneous sums* [33, 37, 43], which are in turn the building blocks of *Gaussian and*

---

✉ Giovanni Peccati  
giovanni.peccati@uni.lu

<sup>1</sup> Heinrich-Heine-Universität Düsseldorf, Mathematisches Institut, Universitätsstr. 1 D-40225, Düsseldorf, Germany

<sup>2</sup> Université du Luxembourg, Belval, Maison du Nombre, 6, avenue de la Fonte, L-4364 Esch-sur-Alzette, Luxembourg

<sup>3</sup> Massachusetts Institute of Technology, Laboratory for Information and Decision Systems, 77 Massachusetts Avenue, Cambridge, MA 02139, United States of America

*Poisson Wiener chaoses* [35, 41], as well as the basis of the *Fourier-Walsh expansion* of mappings defined on discrete cubes [22, 39]. For future reference, we recall that a homogeneous sum of order  $p \geq 1$  is a random variable  $Q_n(f_n, \mathfrak{X})$  with the form

$$Q_n(f_n, \mathfrak{X}) = \sum_{i_1, \dots, i_p=1}^n f_n(i_1, \dots, i_p) \prod_{l=1}^p X_{i_l},$$

where  $\mathfrak{X} = (X_i)_{i \in \mathbb{N}}$  is a sequence of independent, mean zero and unit variance random variables and  $f_n$  is a symmetric kernel (i.e.  $f_n(i_1, \dots, i_n) = f_n(i_{\pi(1)}, \dots, i_{\pi(n)})$  for all permutations  $\pi$  of  $\{1, \dots, n\}$  and all  $1 \leq i_1, \dots, i_n \leq n$ ) satisfying  $f_n(i_1, \dots, i_p) = 0$ , whenever there are  $j \neq l$  such that  $i_l = i_j$ .

In the landmark paper [11], P. de Jong proved the following surprising result: *If, for  $n \in \mathbb{N}$ ,  $W_n$  is a (non-symmetric) normalized degenerate  $U$ -statistic of a fixed order  $p$ , then  $W_n$  converges in distribution to a standard normal random variable  $Z$ , provided  $\lim_n \mathbb{E}[W_n^4] = \mathbb{E}[Z^4] = 3$  and a further Lindeberg-Feller negligibility condition is satisfied.*

De Jong's theorem is the ancestor of the class of 'fourth moment theorems' established in the last decade for elements of the Wiener chaos of Gaussian and Poisson random measures—see e.g. [35, Chapter 5], as well as [17, 19].

In [16], the authors used Stein's method in order to establish quantitative uni- and multivariate counterparts to de Jong's theorem. The techniques introduced in [16] put forward several combinatorial quantities that proved useful in further situations: for instance, as demonstrated by the references [13, 15, 18] the technical findings of [16] allow one to derive quantitative (functional) CLTs for symmetric  $U$ -statistics that are expressed in terms of purely analytical quantities, that is, norms of so-called *contraction kernels*.

The aim of the present paper is to prove a multivariate functional version—in the sense of weak convergence in the Skorohod space of càdlàg mappings—of the main results of [11, 16], under a slightly stronger Lindeberg condition (see Condition 1.2 below). As a direct application, in Sect. 1.4 we derive a functional version of the *universality of Wiener chaos* phenomenon first detected in [37].

Our main findings appear in Theorem 1.4 and Theorem 1.5 below. To the best of our expertise, these statements are the first examples of fourth moment theorems for random elements taking value in a metric space without a Hilbert space structure. See [3] for related results in a Dirichlet/Hilbert setting.

## 1.2 Setup and main theorems

We now introduce our objects of study, fix some notation and conditions, and state our main results. For the sake of generality, and at the cost of a slightly heavier notation than the one adopted in [10, 11], from now on we will work in the framework of triangular arrays.

Throughout this section, we fix positive integers  $d$  and  $1 \leq p_1 < p_2 < \dots < p_d$ , where  $d$  always denotes dimension. For  $m \in \mathbb{N}$ , suppose that  $X_1^{(m)}, \dots, X_{n_m}^{(m)}$  are inde-

pendent random elements, defined on a suitable probability space  $(\Omega_m, \mathcal{F}_m, \mathbb{P}_m)$  and assuming values in the respective measurable spaces  $(E_1^{(m)}, \mathcal{E}_1^{(m)}), \dots, (E_{n_m}^{(m)}, \mathcal{E}_{n_m}^{(m)})$ . We will denote by  $\mathbb{E}_m$ ,  $\text{Var}_m$  and  $\text{Cov}_m$  the expectation, variance and covariance with respect to  $\mathbb{P}_m$ , respectively. We will systematically assume that  $\lim_{m \rightarrow \infty} n_m = +\infty$  and, to simplify the notation, we will sometimes suppress the dependence on  $m$ . For  $k = 1, \dots, d$  and  $m \in \mathbb{N}$  we assume that  $W(k) = W^{(m)}(k)$  is a **(completely) degenerate, non-symmetric  $U$ -statistic** of order  $p_k$  based on  $X_1^{(m)}, \dots, X_{n_m}^{(m)}$ , with the form

$$W(k) = W^{(m)}(k) = \sum_{J \in \mathcal{D}_{p_k}(n_m)} W_J^{(m)}(k) = \sum_{J \in \mathcal{D}_{p_k}(n_m)} W_J(k),$$

where  $W_J(k)$  is a degenerate kernel depending on  $\{X_n : n \in J\}$ —see Sect. 2 below. Here and in what follows, for  $p, n \in \mathbb{N}$ , we write

$$\mathcal{D}_p(n) := \{J \subseteq [n] : |J| = p\}$$

for the collection of all  $\binom{n}{p}$  distinct  $p$ -subsets of  $[n] = \{1, \dots, n\}$ .

W.l.o.g. we will assume that  $\text{Var}_m(W^{(m)}(k)) = 1$  for all  $k = 1, \dots, d$  and all  $m \in \mathbb{N}$ . By  $W = W^{(m)}$  we denote the vector  $W = (W(1), \dots, W(d))^T$  and, for  $t \in [0, 1]$  let  $\mathbf{W}_t = \mathbf{W}_t^{(m)} = (\mathbf{W}_t^{(m)}(1), \dots, \mathbf{W}_t^{(m)}(d))^T$  be given by

$$\mathbf{W}_t^{(m)}(k) := \sum_{J \in \mathcal{D}_{p_k}(\lfloor n_m t \rfloor)} W_J^{(m)}(k) = \sum_{1 \leq i_1 < \dots < i_{p_k} \leq \lfloor n_m t \rfloor} W_{\{i_1, \dots, i_{p_k}\}}^{(m)}(k), \quad k = 1, \dots, d, \quad (1)$$

where  $\lfloor x \rfloor$  denotes the integer part of the real number  $x$ . Then, for each  $m \in \mathbb{N}$ , the process

$$\mathbf{W}^{(m)} := (\mathbf{W}^{(m)}(1), \dots, \mathbf{W}^{(m)}(d))^T := (\mathbf{W}_t^{(m)})_{0 \leq t \leq 1} \quad (2)$$

is an element of the Skorohod space  $D([0, 1]; \mathbb{R}^d)$  of càdlàg functions from  $[0, 1]$  to  $\mathbb{R}^d$ . As anticipated, in this paper we provide sufficient conditions for the sequence  $\mathbf{W}^{(m)}$ ,  $m \in \mathbb{N}$ , to converge in distribution to a continuous Gaussian process  $\mathbf{Z} = (\mathbf{Z}(1), \dots, \mathbf{Z}(d))^T$  with respect to the Skorohod topology on  $D([0, 1]; \mathbb{R}^d)$  (see e.g. [4, Sections 14–15] for details).

Note that, by degeneracy and since the  $p_k$ 's are pairwise distinct, for  $0 \leq s \leq t \leq 1$  and  $1 \leq l, k \leq d$  we have

$$\begin{aligned} \text{Cov}(\mathbf{W}_t^{(m)}(l), \mathbf{W}_t^{(m)}(k)) &= \mathbb{E}_m[\mathbf{W}_t^{(m)}(l)\mathbf{W}_t^{(m)}(k)] = \delta_{k,l} \sum_{J \in \mathcal{D}_{p_k}(\lfloor n_m s \rfloor)} \mathbb{E}_m[W_J^{(m)}(k)^2] \\ &= \delta_{k,l} \sigma_{m,k}^2(\lfloor n_m s \rfloor), \end{aligned}$$

where we set

$$\sigma_{m,k}^2(j) := \sum_{J \in \mathcal{D}_{p_k}(j)} \mathbb{E}_m[W_J^{(m)}(k)^2], \quad k = 1, \dots, d, \quad 0 \leq j \leq n_m.$$

In particular, we have  $\sigma_{m,k}^2(0) = 0$  and  $\sigma_{m,k}^2(n_m) = \text{Var}(W^{(m)}(k)) = 1$ .

We will now introduce and discuss the Conditions 1.1, 1.2 and 1.3 under which our main findings (Theorems 1.4 and 1.5) are stated.

**Condition 1.1** (*Convergence of time-changes*). There are functions  $v_k : [0, 1] \rightarrow \mathbb{R}$ ,  $1 \leq k \leq d$ , such that the pointwise limits

$$\lim_{m \rightarrow \infty} \sigma_{m,k}^2(\lfloor n_m s \rfloor) = v_k(s), \quad s \in [0, 1], \quad (3)$$

exist.

Note that the  $v_k$  are necessarily non-decreasing on  $[0, 1]$  and satisfy  $v_k(0) = 0$  as well as  $v_k(1) = 1$ . However, we need not assume any regularity of the functions  $v_k$ .

In order to concretely relate Condition 1.1 to functional results, we let  $\mathbf{Z} = (\mathbf{Z}(1), \dots, \mathbf{Z}(d))^T$  be a centered Gaussian process with independent components, defined on a suitable probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and such that, for  $1 \leq k \leq d$ , the covariance function of  $\mathbf{Z}(k) = (\mathbf{Z}_t(k))_{t \in [0,1]}$  is given by

$$\text{Cov}(\mathbf{Z}_s(k), \mathbf{Z}_t(k)) = \mathbb{E}[\mathbf{Z}_s(k) \mathbf{Z}_t(k)] = v_k(s \wedge t), \quad (4)$$

implying that  $(\mathbf{Z}_1(1), \dots, \mathbf{Z}_1(d))^T$  is a standard Gaussian vector in  $\mathbb{R}^d$ . It is easily verified that the  $d$ -dimensional process  $\mathbf{Z}$  has the same distribution as

$$(\mathbf{B}_{v_1(t)}(1), \dots, \mathbf{B}_{v_d(t)}(d))_{t \in [0,1]},$$

where  $(\mathbf{B}(1), \dots, \mathbf{B}(d))$  is a standard Brownian motion (initialized at zero) in  $\mathbb{R}^d$ . This yields in particular that  $\mathbf{Z}$  has a.s.- $\mathbb{P}$  continuous paths if and only if the functions  $v_k$ 's are continuous.

Following [10, 11, 16] we further introduce the quantities

$$\begin{aligned} \varrho_{m,k}^2 &:= \max_{1 \leq i \leq n_m} \sum_{\substack{J \in \mathcal{D}_{p_k}(n_m): \\ i \in J}} \mathbb{E}_m[W_J^{(m)}(k)^2] \quad \text{and} \\ D_{m,k} &:= \sup_{\substack{J \in \mathcal{D}_{p_k}(n_m): \\ \mathbb{E}_m[W_J^{(m)}(k)^2] > 0}} \frac{\mathbb{E}_m[W_J^{(m)}(k)^4]}{\mathbb{E}_m[W_J^{(m)}(k)^2]^2}. \end{aligned}$$

Note that, e.g. by Jensen's inequality, one necessarily has that  $D_{m,k} \geq 1$ .

**Condition 1.2** (*Reinforced Lindeberg condition*). For all  $1 \leq k \leq d$ ,  $\lim_{m \rightarrow \infty} D_{m,k} \varrho_{m,k}^2 = 0$ .

This condition roughly ensures that none of the individual random variables  $X_i^{(m)}$ ,  $1 \leq i \leq n_m$ , has an asymptotically dominant impact on the variance of  $W^{(m)}(k)$ ,  $1 \leq k \leq d$ .

The following condition is inherent to all de Jong type CLTs such as e.g. those proved in [11, 16].

**Condition 1.3** (*Fourth moment condition*). For all  $1 \leq k \leq d$  we have  $\lim_{m \rightarrow \infty} \mathbb{E}_m[W^{(m)}(k)^4] = 3$ .

The forthcoming Theorems 1.4 and 1.5 are the main achievements of this work.

**Theorem 1.4** *Let the above definitions and notation prevail and suppose that Conditions 1.1–1.3 hold. Then, the process  $\mathbf{Z}$  has continuous paths and, as  $m \rightarrow \infty$ , the distributions of the sequence  $(\mathbf{W}^{(m)})_{m \in \mathbb{N}}$  weakly converge to the law of  $\mathbf{Z}$ , with respect to both the Skorohod and uniform topologies.*

An assumption such as Condition 1.1 does not appear in the finite-dimensional settings of [11, 16]. It cannot be easily dispensed with without affecting the unicity of the limit in the previous statement. The following result shows that, if Condition 1.1 is removed, then the sequence  $(\mathbf{W}^{(m)})_{m \in \mathbb{N}}$  is still relatively compact, and each limit law is the law of a continuous Gaussian process.

**Theorem 1.5** *Suppose that the sequence  $(\mathbf{W}^{(m)})_{m \in \mathbb{N}}$  satisfies Conditions 1.2 and 1.3. Then, the collection of their laws is tight, hence relatively compact, in the class of distributions on the Skorohod space equipped with either the Skorohod or the uniform metric. Moreover, if the laws of a subsequence  $(\mathbf{W}^{(m_j)})_{j \in \mathbb{N}}$  converge weakly, then the limit is the law of a continuous,  $d$ -dimensional centered Gaussian process with independent components.*

The detailed proofs of Theorems 1.4 and 1.5 are deferred to Sect. 3. We observe that the proof of Theorem 1.5 implicitly shows that, if  $(\mathbf{W}^{(m)})_{m \in \mathbb{N}}$  satisfies Conditions 1.2 and 1.3, then from every subsequence  $(\mathbf{W}^{(m_j)})_{j \in \mathbb{N}}$  one can extract a sub-subsequence  $(\mathbf{W}^{(m'_j)})_{j \in \mathbb{N}}$  verifying Condition 1.1 for some continuous functions  $v_k$ , whose definition depends on the choice of the subsequence.

Some representative examples of kernels directly verifying the assumptions of Theorem 1.4 are described in Example 1.12, yielding in particular that, for all  $p \geq 3$  and all integers  $a \in \{2, \dots, p-1\}$ , there exists a sequence of homogeneous sums whose associated empirical processes converge in distribution to a multiple of  $B(t^{p/a})$ , where  $B$  is a standard Brownian motion issued from zero. It is important to notice that such a limit behaviour is in principle *not* achievable in the framework of symmetric and degenerate  $U$ -statistics. Indeed, in [13, Corollary 3.7] it is proved that, given a sequence of symmetric and degenerate  $U$ -statistics of order  $p \geq 2$  verifying asymptotic relations that are roughly equivalent to Condition 1.3, the corresponding sequence of normalized empirical processes always converges in distribution to a multiple of  $B(t^p)$ .

**Remark 1.6** (a) Theorem 1.4 provides a strong functional extension of the finite-dimensional de Jong type theorems from [11, 16] under very mild additional

- assumptions. As already observed, Condition 1.1 and Condition 1.3 are indeed natural in this context [11, 16], so that only Condition 1.2 might appear non-optimal, since—whenever there is a  $k$  such that  $D_{k,m}$  is unbounded in  $m$ —it is strictly stronger than the usual negligibility condition demanding that  $\lim_{m \rightarrow \infty} \varrho_{m,k}^2 = 0$  for all  $1 \leq k \leq d$ , see again [11, 16]. To address this issue, we observe first that, in many instances, the  $D_{k,m}$  are in fact bounded in  $m$  for all  $1 \leq k \leq d$  and in this case, the two conditions are in fact equivalent. This happens in the case of homogeneous sums such that the underlying sequence  $\mathfrak{X}$  has a uniformly bounded fourth moment (see Sect. 1.4 below). Moreover, it is not clear whether the convergence of finite dimensional distributions of  $\mathbf{W}$  to those of  $\mathbf{Z}$  would still hold, if one replaced Condition 1.2 with the weaker variant (see Sect. 3.1.1 for details). Some preliminary computations have indeed shown that, if one only assumed that  $\lim_{m \rightarrow \infty} \varrho_{m,k}^2 = 0$  for all  $1 \leq k \leq d$ , one might have to replace Condition 1.3 with the assumption that  $\lim_{m \rightarrow \infty} \mathbb{E}[\mathbf{W}_t(k)^4] = 3v_k(t)^2$  for all  $t \in [0, 1]$  and all  $1 \leq k \leq d$ , which would be much more complicated to verify than Condition 1.3.
- (b) Note that Conditions 1.1–1.3, implying the *joint convergence* of the random processes  $\mathbf{W}^{(m)}(k)$ ,  $1 \leq k \leq d$ , are just the aggregation of the conditions for the *componentwise convergence* of these processes.
  - (c) We find it remarkable that one uniquely has to take into account the behaviour of the fourth moments of the components of  $\mathbf{W}^{(m)}$  at time 1 and not at other times  $t < 1$ .
  - (d) We compare our Theorem 1.4 to a few existing functional CLTs (FCLTs) for degenerate (symmetric and non-symmetric)  $U$ -statistics in the scarce literature on this topic. In accordance with the univariate case (see e.g. [14, 21, 44]), in order to satisfy a FCLT, a symmetric degenerate  $U$ -statistic of order  $p > 1$  must have a variable kernel, i.e. one that depends on the sequential parameter  $m$ . In the recent work [13] we have proved FCLTs for this situation with sufficient conditions for convergence that are expressed in norms of contraction kernels and, hence, are completely analytical in nature. A more general class of statistics is given by the so-called *weighted, degenerate  $U$ -statistics of order  $p$* . These have the form  $U_n = \sum_{1 \leq i_1 < \dots < i_p \leq n} a(i_1, \dots, i_p) \psi_n(X_{i_1}, \dots, X_{i_p})$ , where  $X_1, \dots, X_n$  are i.i.d. elements of some space  $E$ ,  $a(i_1, \dots, i_p)$ ,  $1 \leq i_1, \dots, i_p \leq n$  is a symmetric array of real numbers that vanishes on diagonals (i.e.  $a(i_1, \dots, i_p) = 0$  if  $i_l = i_k$  for some  $l \neq k$ ) and  $\psi_n$  is a symmetric, canonical kernel, i.e.  $\mathbb{E}[\psi_n(X_1, \dots, X_p) | X_1, \dots, X_{p-1}] = 0$  a.s. In particular, they constitute a strict subclass of the non-symmetric  $U$ -statistics investigated in the present paper. In the case  $p = 2$ , [32] derives functional limit theorems for weighted degenerate  $U$ -statistics by using methods that heavily depend on the peculiarities of the case  $p = 2$ . In the very recent paper [12] the first and the second author have introduced a functional version of Stein's method of exchangeable pairs and applied it to prove quantitative, multidimensional FCLTs for degenerate weighted  $U$ -statistics. As explained in detail in [12, Subsection 1.5], the limit theorems that build on the bounds of [12] rather involve third absolute moments and absolute values of the coefficient arrays and, hence, are complementary to the present ones involving fourth moment conditions. As to FCLTs for homogeneous sums, apart from our

paper [12] we only mention the two references [31] and [2], where, again, [31] only considers the case  $p = 2$  of quadratic forms. The paper [2] instead, considers homogeneous sums of arbitrary order  $p$  but the results of this paper are in question. Indeed, we have found that the argument leading to [2, Theorem 1.1] is flawed, as carefully explained in part 2 of [12, Remark 5.6].

### 1.3 About our approach

Our method of proof turns out to be a combination of a quantitative, multivariate CLT for vectors of degenerate  $U$ -statistics to obtain convergence of finite-dimensional distributions (see Lemma 3.1 below), and a martingale representation, which we exploit via Doob's  $L^4$ -inequality, in order to derive tightness. In order to conclude from Condition 1.3 that the fourth cumulant of  $\mathbf{W}_t^{(m)}(k)$ ,  $1 \leq k \leq d$ , also converges to zero for all fixed  $t \in [0, 1)$  as  $m \rightarrow \infty$ , we elaborate on a result from the monograph [11] (see Lemma 3.2 below). Moreover, for the application of Lemma 3.1, we prove and make use of several new results about degenerate  $U$ -statistics and their Hoeffding decompositions in Sect. 2, which are of independent interest.

Since in his original proof of the one-dimensional CLT [10], de Jong made use of a classical martingale CLT due to Heyde and Brown [23], it is a natural question whether one could similarly apply a FCLT for martingales, like for instance [34, Theorem 2.34], in order to derive Theorem 1.4. Indeed, adapting the computations beginning on page 286 of [10] and using our Lemma 3.2, one could verify that (an adaptation to deterministic time changes of) the “square bracket condition” (2.48) in [34] holds for all fixed  $t \in [0, 1]$ . Since the uniform negligibility condition (2.46) in [34] is satisfied for any sequence of normalized  $L^2$ -martingales, [34, Theorem 2.34] could be applied to yield another proof of the *one-dimensional version* of Theorem 1.4. In order to derive the full multidimensional result, one could try to similarly adapt a multivariate FCLT for martingales to our setting, like for example [45, Theorem 2.1]. However, the mixed square bracket condition (3) in [45] seems to be less amenable to analysis in the present setting than its univariate counterpart, i.e. condition (2.48) in [34]. Hence, adapting the techniques used in [11] so as to obtain our multivariate result appears considerably more difficult. One of the main reasons for this fact is that de Jong's arguments rely on a large collection of exact combinatorial identities, that one uses in order to verify the square bracket condition evoked above. In contrast, our direct arguments merely use probabilistic inequalities, which makes the structure of the proof significantly simpler and probably also more robust. Finally, we mention that our arguments are in principle quantitative in nature and that, via an elaboration of Stein's method for diffusion approximation [1, 12, 26, 27], one would be able to prove a quantitative version of Theorem 1.4. Since, at the time of writing, we are only in the position to do this with respect to the stronger  $L^1$ -topology on  $D([0, 1]; \mathbb{R}^d)$ , we decide to leave this point open for further investigation.

## 1.4 Application: universal FCLTS for homogeneous sums

We will now apply our main results to the setting of homogeneous sums: as a consequence, in Theorem 1.10 we will establish functional versions of the main findings of [37], where the authors proved the *universality of Gaussian Wiener chaos* with respect to the normal approximation of homogeneous sums. Such a result roughly states that, if a sequence of homogeneous sums with Gaussian arguments verifies a CLT, then so does the sequence obtained by replacing the Gaussian input with an arbitrary collection of independent random variables (see e.g. [37, Theorem 1.2] for a precise statement). The findings from [37] have found applications e.g. in the analysis of disordered systems [8, 9], mathematical statistics [29], random matrix theory and free probability [36, 40].

In this subsection, we let  $(\Omega, \mathcal{F}, \mathbb{P})$  denote a suitable probability space on which all random quantities subsequently dealt with are defined. We will consider the following sequences of real random variables:

- $\mathfrak{X} = (X_i)_{i \in \mathbb{N}}$  denotes a (generic) sequence of independent, mean zero and unit variance random variables in  $L^4(\mathbb{P})$ ;
- $\mathfrak{G} = (G_i)_{i \in \mathbb{N}}$  is a sequence of independent standard normal random variables; and
- $\mathfrak{P} = (P_i)_{i \in \mathbb{N}}$  is a sequence of independent, normalized Poisson random variables, i.e. there are  $\lambda_i \in (0, \infty)$  and independent Poisson random variables  $N_i$  with mean  $\lambda_i$  such that  $P_i = N_i / \sqrt{\lambda_i} - \sqrt{\lambda_i}$ ,  $i \in \mathbb{N}$ .

On the sequence  $\mathfrak{X}$ , we will further impose the integrability condition that

$$\beta := \sup_{i \in \mathbb{N}} \mathbb{E}|X_i|^4 < \infty. \quad (5)$$

This is in particular satisfied if the random variables  $X_i$ ,  $i \in \mathbb{N}$ , are i.i.d, since we have assumed that they are in  $L^4(\mathbb{P})$ . We observe that, in the case of the Poisson sequence  $\mathfrak{P}$ , condition (5) can be equivalently re-expressed in terms of the coefficients  $\lambda_i$ . In fact, since  $\mathbb{E}[P_i^4] = 3 + \lambda_i^{-1}$ , the sequence  $\mathfrak{P}$  satisfies (5) if and only if  $\inf_{i \in \mathbb{N}} \lambda_i > 0$ .

In what follows, we let  $n_m = m$  and, for given arrays  $a_J^{(m)}(k)$ ,  $J \in \mathcal{D}_{p_k}(m)$ ,  $1 \leq k \leq d$ , of real numbers we suppose that  $W_J^{(m)}(k) = a_J^{(m)}(k) \prod_{i \in J} X_i$  is given as

$$W_J^{(m)}(k) = a_J^{(m)}(k) \prod_{i \in J} X_i =: a_J^{(m)}(k) X_J.$$

Thus, for  $k = 1, \dots, d$ ,

$$W^{(m)}(k) = \sum_{J \in \mathcal{D}_{p_k}(m)} a_J^{(m)}(k) \prod_{i \in J} X_i \quad (6)$$

is a *homogeneous sum* or *homogeneous multilinear form* of order  $p_k$  and, in particular, it is a non-symmetric, degenerate  $U$ -statistic. As before, we will systematically assume that



$$\text{Var}(W^{(m)}(k)) = \sum_{J \in \mathcal{D}_{p_k}(m)} a_J^{(m)}(k)^2 = 1, \quad \text{for } k = 1, \dots, d.$$

Then, as in (1) and (2), we use these  $d$  homogeneous sums to define the random element  $\mathbf{W}^{(m)} = (W^{(m)}(1), \dots, W^{(m)}(d))^T$  with values in  $D([0, 1]; \mathbb{R}^d)$ .

Note that from (5), for  $1 \leq k \leq d$  and all  $J \in \mathcal{D}_{p_k}(m)$  such that  $a_J^{(m)}(k) \neq 0$ , we have that

$$\frac{\mathbb{E}[W_J^{(m)}(k)^4]}{(\mathbb{E}[W_J^{(m)}(k)^2])^2} = \frac{a_J^{(m)}(k)^4 \mathbb{E}[\prod_{i \in J} X_i^4]}{a_J^{(m)}(k)^4 (\mathbb{E}[\prod_{i \in J} X_i^2])^2} = \prod_{i \in J} \mathbb{E}[X_i^4] \leq \beta^{p_k}.$$

Hence, also

$$\sup_{m \in \mathbb{N}} D_{m,k} \leq \beta^{p_k} < \infty.$$

In this context, one would typically like to replace the fourth moment condition, Condition 1.3, by a condition, which is rather expressed in terms of the coefficient arrays. To this end, let us first define the symmetric functions  $f_k^{(m)} : [m]^{p_k} \rightarrow \mathbb{R}$  by

$$f_k^{(m)}(i_1, \dots, i_{p_k}) := \begin{cases} \frac{1}{p_k!} a_{\{i_1, \dots, i_{p_k}\}}^{(m)}(k), & \text{if } |\{i_1, \dots, i_{p_k}\}| = p_k, \\ 0, & \text{otherwise.} \end{cases}$$

Observe that the function  $f_k^{(m)}$  vanishes on diagonals, that is,  $f_k^{(m)}(i_1, \dots, i_{p_k}) = 0$  if two coordinates of  $(i_1, \dots, i_{p_k})$  coincide. Then, we have

$$\begin{aligned} W^{(m)}(k) &= \mathcal{Q}_{p_k}(m, f_k^{(m)}, \mathfrak{X}) := \sum_{i_1, \dots, i_{p_k}=1}^m f_k^{(m)}(i_1, \dots, i_{p_k}) X_{i_1} \cdots X_{i_{p_k}} \\ &= p_k! \sum_{1 \leq i_1 < \dots < i_{p_k} \leq m} f_k^{(m)}(i_1, \dots, i_{p_k}) X_{i_1} \cdots X_{i_{p_k}} \end{aligned}$$

as well as

$$\mathcal{Q}_{m,k}^2 = \max_{1 \leq i \leq m} \sum_{\substack{1 \leq j_2 < \dots < j_{p_k} \leq m: \\ j_2, \dots, j_{p_k} \neq i}} a_{\{i, j_2, \dots, j_{p_k}\}}^{(m)}(k)^2 = (p_k!)^2 \text{Inf}_i(f_k^{(m)}),$$

where

$$\begin{aligned}\text{Inf}_i(f_k^{(m)}) &= \frac{1}{(p_k - 1)!} \sum_{j_2, \dots, j_{p_k}=1}^m f_k^{(m)}(i, j_2, \dots, j_{p_k})^2 \\ &= \sum_{1 \leq j_2 < \dots < j_{p_k} \leq m} f_k^{(m)}(i, j_2, \dots, j_{p_k})^2\end{aligned}$$

denotes the *influence* of the random variable  $X_i$  on the variance of  $W^{(m)}(k)$  (see [22, 33]). Thus, we have shown that, under (5), Condition 1.2 is equivalent to

$$\lim_{m \rightarrow \infty} \max_{1 \leq i \leq m} \text{Inf}_i(f_k^{(m)}) = 0 \quad \text{for } 1 \leq k \leq d. \quad (7)$$

For  $1 \leq k \leq d$  and  $1 \leq r \leq p_k - 1$ , we define the *contraction kernel*  $f_k^{(m)} \star_r f_k^{(m)} : [m]^{2p_k-2r} \rightarrow \mathbb{R}$  by

$$\begin{aligned}& f_k^{(m)} \star_r f_k^{(m)}(i_1, \dots, i_{p_k-r}, j_1, \dots, j_{p_k-r}) \\ &:= \sum_{l_1, \dots, l_r=1}^m f_k^{(m)}(i_1, \dots, i_{p_k-r}, l_1, \dots, l_r) f_k^{(m)}(j_1, \dots, j_{p_k-r}, l_1, \dots, l_r).\end{aligned}$$

Consider the following condition.

**Condition 1.7** For all  $1 \leq k \leq d$  and all  $1 \leq r \leq p_k - 1$  it holds that  $\lim_{m \rightarrow \infty} \|f_k^{(m)} \star_r f_k^{(m)}\|_2 = 0$ .

Here,  $\|f_k^{(m)} \star_r f_k^{(m)}\|_2$  denotes the  $L^2$ -norm of the contraction kernel  $f_k^{(m)} \star_r f_k^{(m)}$  with respect to the  $(2p_k - 2r)$ -fold product of the counting measure on  $[m]$ , i.e. with respect to the counting measure on  $[m]^{2p_k-2r}$ .

**Proposition 1.8** Suppose that (5) and Condition 1.7 hold and that  $p_k \geq 2$  for all  $1, \dots, d$ . Then, Conditions 1.2 and 1.3 are satisfied as well.

**Proof** We have already shown that Condition 1.2 holds if and only if (7) holds. Relation (7) however holds true by virtue of Condition 1.7 and display (1.9) in [37]. That Condition 1.3 holds under Condition 1.7 as well, follows from a combination of [37, Proposition 1.6] and the estimate (1.13) in [37].  $\square$

Note further that the functions  $\sigma_{m,k}^2(\lfloor n_m s \rfloor)$  from Condition 1.1 are in this situation more explicitly given by

$$\sigma_{m,k}^2(\lfloor n_m s \rfloor) = \sum_{J \in \mathcal{D}_{p_k}(\lfloor n_m s \rfloor)} a_J^{(m)}(k)^2 = p_k! \sum_{i_1, \dots, i_{p_k}=1}^{\lfloor n_m s \rfloor} f_k^{(m)}(i_1, \dots, i_{p_k})^2 =: S f_k^{(m)}(s). \quad (8)$$

**Theorem 1.9** Suppose that the sequence  $\mathfrak{X}$  of independent, mean zero and unit variance random variables satisfies (5). Moreover, assume that the sequences  $(Sf_k^{(m)})_{m \in \mathbb{N}}$ ,  $1 \leq k \leq d$ , defined in (8), satisfy Condition 1.1, that Condition 1.7 holds and that  $p_k \geq 2$  for all  $k = 1, \dots, d$ . Then, the sequence  $(\mathbf{W}^{(m)})_{m \in \mathbb{N}}$  of processes, defined through (2) and (6), converges in distribution with respect to the Skorohod topology to the continuous, centered Gaussian process  $\mathbf{Z} = (\mathbf{Z}(1), \dots, \mathbf{Z}(d))^T$  with independent components, which is defined via (3).

**Proof** This follows from Theorem 1.4 and Proposition 1.8.  $\square$

The proof of Theorem 1.9 (via the implicit use of Proposition 1.8) strongly relies on the techniques developed in [37]. As anticipated, we will now prove a functional version of the universality results from [37]. For completeness, we will also include the discussion of universality of Poisson homogeneous sums, as provided by [19, 42].

**Theorem 1.10** (Functional universality of homogeneous sums). *With the above notation and definitions, consider the following assumptions:*

- (A<sub>1</sub>) There are functions  $v_k : [0, 1] \rightarrow \mathbb{R}$  such that  $\lim_{m \rightarrow \infty} (Sf_k^{(m)}) = v_k$  point-wise,  $1 \leq k \leq d$ .
- (A<sub>2,  $\mathfrak{X}$</sub> ) As  $m \rightarrow \infty$ ,  $(Q_{p_1}(m, f_1^{(m)}, \mathfrak{X}), \dots, Q_{p_d}(m, f_d^{(m)}, \mathfrak{X}))^T$  converges in distribution to a  $d$ -dimensional standard normal vector.
- (A<sub>3,  $\mathfrak{X}$</sub> ) The sequences  $Q_{p_k}(m, f_k^{(m)}, \mathfrak{X})_{m \in \mathbb{N}}$ ,  $1 \leq k \leq d$ , satisfy Condition 1.3.
- (A<sub>4,  $\mathfrak{X}$</sub> ) The above defined sequence  $(\mathbf{W}^{(m)}(1), \dots, \mathbf{W}^{(m)}(d))^T$ ,  $m \in \mathbb{N}$ , of processes based on  $\mathfrak{X}$ , converges in distribution with respect to the Skorohod topology to a continuous centered Gaussian process  $\mathbf{Z} = (\mathbf{Z}(1), \dots, \mathbf{Z}(d))^T$  with independent components, whose covariance function is given by (3).
- (A<sub>5</sub>) Condition 1.7 is satisfied.
- (A<sub>6</sub>) One has  $\inf_{i \in \mathbb{N}} \lambda_i > 0$ .

Assume that Assumption (A<sub>1</sub>) holds and  $p_k \geq 2$  for all  $k = 1, \dots, d$ . Then, the following implications are in order:

$$\begin{aligned} (A_{3, \mathfrak{P}}) &\Rightarrow (A_{2, \mathfrak{G}}) \Leftrightarrow (A_{3, \mathfrak{G}}) \Leftrightarrow (A_{4, \mathfrak{G}}) \Leftrightarrow (A_5) \\ &\Rightarrow \forall \mathfrak{X} \text{ satisfying (5)} : ((A_{2, \mathfrak{X}}) \wedge (A_{3, \mathfrak{X}}) \wedge (A_{4, \mathfrak{X}})). \end{aligned} \quad (9)$$

If, additionally to (A<sub>1</sub>), also (A<sub>6</sub>) holds and  $p_k \geq 2$  for all  $k = 1, \dots, d$ , then we have in fact that

$$\begin{aligned} (A_{3, \mathfrak{P}}) &\Leftrightarrow (A_{2, \mathfrak{G}}) \Leftrightarrow (A_{3, \mathfrak{G}}) \Leftrightarrow (A_{4, \mathfrak{G}}) \Leftrightarrow (A_5) \\ &\Rightarrow \forall \mathfrak{X} \text{ satisfying (5)} : ((A_{2, \mathfrak{X}}) \wedge (A_{3, \mathfrak{X}}) \wedge (A_{4, \mathfrak{X}})). \end{aligned} \quad (10)$$

**Proof** Suppose first that (A<sub>1</sub>) holds. That (A<sub>2,  $\mathfrak{G}$</sub> ), (A<sub>3,  $\mathfrak{G}$</sub> ) and (A<sub>5</sub>) are all equivalent is a consequence of [37, Proposition 1.6] (even without assuming (A<sub>1</sub>)). Of course, (A<sub>4,  $\mathfrak{G}$</sub> ) implies (A<sub>2,  $\mathfrak{G}$</sub> ). By Theorem 1.9 applied to  $\mathfrak{G}$ , under (A<sub>1</sub>) we have that (A<sub>5</sub>) implies (A<sub>4,  $\mathfrak{G}$</sub> ), since  $\mathfrak{G}$  obviously satisfies (5). Moreover, again by Theorem 1.9 and by Proposition 1.8, if the sequence  $\mathfrak{X}$  satisfies (5) and if (A<sub>5</sub>) holds, then (since we have

assumed  $(A_1)$ ), we can conclude that  $(A_{3,\mathfrak{X}})$  and  $(A_{4,\mathfrak{X}})$  hold, immediately implying  $(A_{2,\mathfrak{X}})$  as well. Finally, [19, Theorem 1.11] ensures that  $(A_{3,\mathfrak{P}})$  implies  $(A_{2,\mathfrak{G}})$ .

Suppose now that both  $(A_1)$  and  $(A_6)$  hold. Then, [19, Theorem 1.11] guarantees that  $(A_{2,\mathfrak{G}})$  also implies  $(A_{3,\mathfrak{P}})$ .  $\square$

**Remark 1.11** (Removing condition  $(A_1)$ ) If Condition  $(A_1)$  is removed, then the conclusions of Theorem 1.10 continue to hold, provided condition  $(A_{4,\mathfrak{X}})$  is replaced by

$(A'_{4,\mathfrak{X}})$  The collection of distributions of the above defined sequence  $(\mathbf{W}^{(m)}(1), \dots, \mathbf{W}^{(m)}(d))^T$ ,  $m \in \mathbb{N}$ , of processes based on  $\mathfrak{X}$ , is relatively compact with respect to the Skorohod topology, and every converging subsequence admits as a limit a continuous centered Gaussian process with independent components.

The proof of this fact follows the same lines as the proof of Theorem 1.5 (see Sect. 3.3); the details are left to the reader.

**Example 1.12** (Fractional products) For every integer  $p \geq 3$  we will now demonstrate the existence of a sequence of symmetric kernels  $f^{(m)} : \{1, \dots, m\}^p \rightarrow \mathbb{R}$ , vanishing on diagonals and such that Conditions 1.7 and  $(A_1)$  in Theorem 1.10 (for  $d = 1$ ) are satisfied, with  $v_1(s) = v(s) = s^{p/a}$ , for some integer  $2 \leq a \leq p - 1$ . The definition of  $f^{(m)}$  is based on a slight variation of the construction of *fractional cartesian products*, as described in [7, Chapter XIII], as well as in [6, 38]. Fix  $a, p$  as above, and consider an injective mapping  $\varphi : \mathbb{N}^a \mapsto \mathbb{N}$  with the properties that, for all  $k \geq 1$ , (a)  $\varphi([k]^a) \subset [k^a]$ , and (b)  $\varphi([k+1]^a) \setminus \varphi([k]^a) \subset [(k+1)^a] \setminus [k^a]$ . Such a mapping is easily defined by recursion on  $k$ . Observe in particular that Property (a) implies that, for all  $N \geq 1$ , the image of the restriction of  $\varphi$  to  $[\lfloor N^{1/a} \rfloor]^a$  is contained in  $[N]$ . Consider in addition a collection  $\{S_1, \dots, S_p\}$  of non-empty distinct subsets of  $[p]$ , satisfying the properties that (i)  $|S_i| = a$ , and (ii) each index  $i \in [p]$  appears in exactly  $a$  of the sets  $S_i$  (yielding in particular  $[p] = \cup_i S_i$ ). For every  $i$  and every vector  $\mathbf{t} = (t_1, \dots, t_p)$ , we set  $\pi_{S_i} \mathbf{t} = (t_k : k \in S_i)$ , where the indices belonging to  $S_i$  are implicitly listed in increasing order. Our aim is to use the mapping  $\varphi$  and the sets  $S_i$  in order to define a *sparse* subset  $F_m$  of  $[m]^p$  for every integer  $m > p^a$  (this last relation is required in order to ensure that  $[\lfloor m^{1/a} \rfloor]$  contains at least  $p$  elements): the nature of the sparsity of  $F_m$  will be encoded by the ratio  $p/a$ , corresponding to its *fractional dimension* [6, 7]. For every  $m \geq 1$  we start by setting

$$F_m^0 := \left\{ \mathbf{k} = (k_1, \dots, k_p) : \mathbf{k} = (\varphi(\pi_{S_1} \mathbf{t}), \dots, \varphi(\pi_{S_p} \mathbf{t})), \text{ for some } \mathbf{t} \in \Delta_{\lfloor m^{1/a} \rfloor}^p \right\} \subset [m]^p.$$

where the symbol  $\Delta_N^p$  indicates the class of all vectors  $(t_1, \dots, t_p)$  such that  $1 \leq t_1 < \dots < t_p \leq N$ ; when  $m \leq p^a$ ,  $F_m^0$  can be empty. The injectivity of  $\varphi$  readily implies that, when  $m > p^a$ , the entries of any  $\mathbf{k} = (k_1, \dots, k_p) \in F_m^0$  are pairwise distinct, and also that, if  $\mathbf{k} \neq \mathbf{k}'$  belong to  $F_m^0$ , then  $\mathbf{k}$  cannot be obtained from  $\mathbf{k}'$  via a permutation of its entries. We eventually set

$$F_m := \mathbf{sym}(F_m^0) \subset [m]^p, \quad m \geq 1,$$

that is,  $F_m$  is the class of those  $(k_1, \dots, k_p)$  such that  $(k_{\sigma(1)}, \dots, k_{\sigma(p)}) \in F_m^0$  for some permutation  $\sigma$ . Immediate combinatorial considerations imply that there exists a finite constant  $b > 0$  such that, as  $N \rightarrow \infty$ ,  $|F_N| \sim b \cdot N^{p/a}$ . For  $m > p^a$ , we define

$$f^{(m)}(i_1, \dots, i_p) := \frac{1}{(p!|F_m|)^{1/2}} \mathbf{1}_{F_m}(i_1, \dots, i_p), \quad (i_1, \dots, i_p) \in [m]^p.$$

The computations in [38, Section 6.2 and Section 6.3] readily imply that  $f^{(m)}$  verifies Condition 1.7. Moreover, since

$$|F_m \cap [\lfloor tm \rfloor]^p| \sim |F_{\lfloor tm \rfloor}| \sim b \cdot t^{p/a} m^{p/a}, \quad (11)$$

we deduce that Assumption  $(A_1)$  in Theorem 1.10 is satisfied with  $v(t) = t^{p/a}$ . Using the chain of implications (9), we infer that, given a sequence  $\mathfrak{X}$  of centred unit variance random variables verifying (5), the corresponding empirical process  $W^{(m)}$ , defined via (1), converges in distribution to  $\{B_{t^{p/a}} : t \in [0, 1]\}$ , where  $B$  is a standard Brownian motion. Selecting pairwise distinct integers  $p_1, \dots, p_d$  yields examples of  $d$ -dimensional sequences verifying Conditions 1.7 and  $(A_1)$  in Theorem 1.10. For the sake of completeness, a proof of (11) is sketched in Sect. 3.4.

**Remark 1.13** In the finite-dimensional setting considered in [37] (see, in particular Theorem 1.2 therein), universality of Gaussian homogeneous sums is established under the weaker condition that  $\sup_{i \in \mathbb{N}} \mathbb{E}|X_i|^3 < \infty$  (of course without implying  $(A_{3,\mathfrak{X}})$ , in this case). This is not possible following our method of proof in the functional situation, since it is based on Theorem 1.4, which is a genuine fourth moment theorem. On the contrary, the proofs in [37] rely on a combination of Malliavin calculus on Gaussian spaces and an invariance principle for multilinear forms, proved in [33] via an elaboration of the Lindeberg swapping trick. Since this invariance principle only requires finite third moments, it is possible to dispense with fourth moment conditions on  $\mathfrak{X}$  in that framework. See also [8, Theorem 4.2].

**Remark 1.14** Proposition 1.8 and Theorems 1.9 and 1.10 all require that  $p_k \geq 2$  for all  $k = 1, \dots, d$ . This is because display (1.9) in [37] used in their proofs relies on this assumption. Relation (7), however, still implies Condition 1.2, even if  $p_1 = 1$ . In this case, writing  $a_i^{(m)}(1)$  for  $a_{\{i\}}^{(m)}(1)$ , we have  $\inf_i (f_1^{(m)}) = a_i^{(m)}(1)^2$ . Moreover, a straightforward computation shows that, for  $p_1 = 1$ , item a) below implies Condition 1.3. Therefore, by part b) of Remark 1.6, if  $p_1 = 1$  and (5) and all of the following hold:

- $\lim_{m \rightarrow \infty} \max_{1 \leq i \leq m} a_i^{(m)}(1)^2 = 0$ ;
- $\lim_{m \rightarrow \infty} \sum_{i=1}^{\lfloor n_m s \rfloor} a_i^{(m)}(1)^2 = v_1(s)$  for some  $v_1 : [0, 1] \rightarrow \mathbb{R}$ ;
- the assumptions of Theorem 1.9 are satisfied for the  $(d-1)$ -dimensional sequence of processes  $\left( Q_{p_2}(m, f_2^{(m)}, \mathfrak{X}), \dots, Q_{p_d}(m, f_d^{(m)}, \mathfrak{X}) \right)_{m \in \mathbb{N}}$  substituted in place of  $(W^m)_{m \in \mathbb{N}}$ ,

then the whole  $d$ -dimensional sequence  $\left(Q_{p_1}(m, f_1^{(m)}, \mathfrak{X}), \dots, Q_{p_d}(m, f_d^{(m)}, \mathfrak{X})\right)_{m \in \mathbb{N}}$  converges in distribution with respect to the Skorohod topology to the continuous, centered Gaussian process  $\mathbf{Z} = (\mathbf{Z}(1), \dots, \mathbf{Z}(d))^T$  with independent components, which is defined via (3).

## 2 Degenerate $U$ -statistics and Hoeffding decompositions

In this section we prove and collect useful auxiliary results about degenerate  $U$ -statistics based on an independent sample. For reasons of simplicity, we abandon the setup of triangular arrays in this section. Thus, we let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a generic probability space, on which independent random elements  $X_1, X_2, \dots$  are defined that have values in the respective measurable spaces  $(E_1, \mathcal{E}_1), (E_2, \mathcal{E}_2), \dots$ . Moreover, for  $n \in \mathbb{N}$  fixed we let  $[n] := \{1, \dots, n\}$  and, for  $J \subseteq [n]$ , we define  $\mathcal{F}_J := \sigma(X_i, i \in J)$ . Note that, whenever

$$f : \prod_{j=1}^n E_j \rightarrow \mathbb{R} \quad \text{is} \quad \bigotimes_{j=1}^n \mathcal{E}_j - \mathcal{B}(\mathbb{R})\text{-measurable}$$

such that

$$Y := f(X_1, \dots, X_n) \in L^1(\mathbb{P}),$$

then there is a  $\mathbb{P}$ -a.s. unique representation of the form

$$Y = \sum_{M \subseteq [n]} Y_M = \sum_{s=0}^n \left( \sum_{\substack{M \subseteq [n]: \\ |M|=s}} Y_M \right) \quad (12)$$

where, for each  $M \subseteq [n]$ , the summand  $Y_M$  is measurable with respect to  $\mathcal{F}_M$  and, furthermore,

$$\mathbb{E}[Y_M | \mathcal{F}_J] = 0 \quad \text{holds, whenever } M \not\subseteq J.$$

Note that, in particular, there are measurable functions

$$f_M : \prod_{j \in M} E_j \rightarrow \mathbb{R}, \quad M \subseteq [n],$$

such that  $Y_M = f_M(X_j, j \in M)$ . The representation (12) is the celebrated **Hoeffding decomposition** of  $Y$ . The following well-known explicit formula for the **Hoeffding components**  $Y_M$ ,  $M \subseteq [n]$ , is easily deduced from the exclusion-inclusion principle:

$$Y_M = \sum_{J \subseteq M} (-1)^{|M|-|J|} \mathbb{E}[Y | \mathcal{F}_J], \quad (13)$$

yielding in particular that  $Y_\emptyset = \mathbb{E}[Y]$  a.s.- $\mathbb{P}$ . Equation (13) directly implies the linearity of the Hoeffding decomposition. It should also be noted that, whenever  $Y \in L^p(\mathbb{P})$  for some  $p \in [1, \infty]$ , then automatically  $Y_M \in L^p(\mathbb{P})$  for all  $M \subseteq [n]$ . Moreover, if  $Y \in L^2(\mathbb{P})$ , then its Hoeffding components are mutually orthogonal in  $L^2(\mathbb{P})$ .

For  $p \in [n]$ , we call  $Y$  a **(completely) degenerate  $U$ -statistic of order  $p$** , based on  $X_1, \dots, X_n$ , if its Hoeffding decomposition (12) is of the form

$$Y = \sum_{J \in \mathcal{D}_p(n)} Y_J, \quad (14)$$

i.e. if  $Y_M = 0$   $\mathbb{P}$ -a.s. whenever  $|M| \neq p$ . Here and in what follows we write

$$\mathcal{D}_p(n) := \{J \subseteq [n] : |J| = p\}$$

for the collection of all  $\binom{n}{p}$  different  $p$ -subsets of  $[n]$ .

Now suppose that we are given two positive integers  $m$  and  $n$ , as well as  $p \in [n]$  and  $q \in [m]$  and consider two degenerate  $U$ -statistics  $V$  based on  $X_1, \dots, X_m$  and  $W$  based on  $X_1, \dots, X_n$  of respective orders  $q$  and  $p$ . Hence, we have the respective Hoeffding decompositions

$$V = \sum_{J \in \mathcal{D}_q(m)} V_J \quad \text{and} \quad W = \sum_{J \in \mathcal{D}_p(n)} W_J.$$

We further assume that  $\mathbb{E}|W|^4, \mathbb{E}|V|^4 < \infty$ . Moreover, for  $l \in [m]$  and  $k \in [n]$  let us define

$$\varrho_{l,V}^2 := \max_{1 \leq j \leq l} \sum_{\substack{J \in \mathcal{D}_q(l): \\ j \in J}} \mathbb{E}[V_J^2] \quad \text{and} \quad \varrho_{k,W}^2 := \max_{1 \leq j \leq k} \sum_{\substack{J \in \mathcal{D}_p(k): \\ j \in J}} \mathbb{E}[W_J^2]$$

as well as

$$\sigma_{l,V}^2 := \sum_{J \in \mathcal{D}_q(l)} \mathbb{E}[V_J^2] \quad \text{and} \quad \sigma_{k,W}^2 := \sum_{J \in \mathcal{D}_p(k)} \mathbb{E}[W_J^2].$$

In particular, we have  $\sigma_{m,V}^2 = \mathbb{E}[V^2]$  and  $\sigma_{n,W}^2 = \mathbb{E}[W^2]$  as well as the inequalities

$$\sigma_{k,W}^2 \leq \sigma_{n,W}^2 \quad \text{and} \quad \varrho_{k,W}^2 \leq \varrho_{n,W}^2 \quad (15)$$

and analogous ones for  $V$  in place of  $W$ . Note that  $VW$  is an integrable function of  $X_1, \dots, X_{n \vee m}$ . Hence, it follows from (13) and the assumptions on  $V$  and  $W$  that it has a Hoeffding decomposition of the form

$$VW = \sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q}} U_M(V, W).$$

For simplicity, we write  $U_M(W)$  for  $U_M(W, W)$  and  $U_M(V)$  for  $U_M(V, V)$  so that

$$V^2 = \sum_{\substack{M \subseteq [m]: \\ |M| \leq 2q}} U_M(V) \quad \text{and} \quad W^2 = \sum_{\substack{N \subseteq [n]: \\ |N| \leq 2p}} U_N(W)$$

are the Hoeffding decompositions of  $V^2$  and  $W^2$ , respectively. We will use the conventions  $U_M(V) = 0$  and  $U_N(W) = 0$  implicitly, whenever  $M \not\subseteq [m]$  and  $N \not\subseteq [n]$ , respectively. Let us define the collection  $\mathcal{S}_0 := \mathcal{S}_0(m, n, q, p)$  of all quadruples  $(I, J, K, L) \in \mathcal{D}_q(m)^2 \times \mathcal{D}_p(n)^2$  such that

- (i)  $I \cap K = J \cap L = \emptyset$ ,
- (ii)  $\emptyset \neq I \cap J = I \setminus (I \cap L) \neq I$ ,
- (iii)  $\emptyset \neq J \cap I = J \setminus (J \cap K) \neq J$ ,
- (iv)  $\emptyset \neq K \cap J = K \setminus (L \cap K) \neq K$  and
- (v)  $\emptyset \neq L \cap I = L \setminus (L \cap K) \neq L$ .

Similarly, we denote by  $\mathcal{S}_1 := \mathcal{S}_1(m, n, p)$  the collection of all quadruples  $(I, J, K, L) \in \mathcal{D}_p(m \wedge n)^4$  such that

- (i)  $I \cap J = K \cap L = \emptyset$ ,
- (ii)  $\emptyset \neq I \cap K = I \setminus (I \cap L) \neq I$ ,
- (iii)  $\emptyset \neq J \cap K = J \setminus (J \cap L) \neq J$ ,
- (iv)  $\emptyset \neq K \cap J = K \setminus (I \cap K) \neq K$  and
- (v)  $\emptyset \neq L \cap I = L \setminus (L \cap J) \neq L$ .

Moreover, we define the quantities

$$S_0(V, W) = \sum_{(I, J, K, L) \in \mathcal{S}_0} \mathbb{E}[V_I V_J W_K W_L] \quad \text{and, if } p = q, \\ S_1(V, W) = \sum_{(I, J, K, L) \in \mathcal{S}_1} \mathbb{E}[V_I V_J W_K W_L].$$

In particular, we have  $S_0(W, W) = S_1(W, W)$ . More generally, if  $p = q$ ,  $m \leq n$  and  $V_J = W_J$  for all  $J \in \mathcal{D}_p(m)$ , then  $S_1(V, W) = S_0(V, V)$ .

**Remark 2.1** The quadruples  $(I, J, K, L) \in \mathcal{S}_0$  play a crucial role in all proofs of de Jong type CLTs via the quantities  $S_0(V, W)$  (see [11, 16]). Note that these quadruples are all *bifold*, meaning that each element  $i \in I \cup J \cup K \cup L$  lies in exactly two of the sets  $I, J, K, L$ . In particular, there can be no *free index*, which is an element that appears in only one of the sets  $I, J, K, L$ . Moreover, no two of the sets  $I, J, K, L$  can be the same and each of these sets is disjoint to precisely one other among these sets. An analogous remark applies to the quadruples  $(I, J, K, L) \in \mathcal{S}_1$ . Note further that the existence of a free index for a quadruple  $(I, J, K, L)$  necessarily implies  $\mathbb{E}[V_I V_J W_K W_L] = 0$  by degeneracy. This fact will be tacitly exploited throughout the following.



**Lemma 2.2** *There exists a finite constant  $C_{p,q}$ , only depending on  $p$  and  $q$ , such that*

$$S_0(V, W) \geq -C_{p,q} \max(\sigma_{m,V}^2 \varrho_{n,W}^2, \sigma_{n,W}^2 \varrho_{m,V}^2).$$

**Proof** This follows from a straightforward generalization of Propositions 3.5 and 3.6 of [16] to the situation of possibly different  $m$  and  $n$  and possibly non-unit variances. We omit the details here.  $\square$

**Lemma 2.3** *We have that*

$$S_0(W, W) \leq \mathbb{E}[W^4] - 3\sigma_{n,W}^4 + 2p\sigma_{n,W}^2 \varrho_{n,W}^2.$$

**Proof** This follows from the first inequality in equation (2.13) of [16] by noticing that the left hand side of this inequality is nonnegative and by taking into account that  $\text{Var}(W) = \sigma_{n,W}^2$  is not necessarily equal to 1, here.  $\square$

The next two lemmas will be of vital importance for proving convergence of finite dimensional distributions. Lemma 2.5 is an extension of results and techniques provided in [16, Section 3], whereas Lemma 2.4 is taken directly from there.

**Lemma 2.4** (Lemma 2.10 of [16]) *There is a finite constant  $\kappa_p$  only depending on  $p$  such that*

$$\sum_{\substack{M \subseteq [n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(W)) \leq \mathbb{E}[W^4] - 3\sigma_{n,W}^4 + \kappa_p \sigma_{n,W}^2 \varrho_{n,W}^2.$$

**Lemma 2.5** (i) *If  $p = q$ , then*

$$\begin{aligned} & \sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq 2p-1}} \text{Var}(U_M(V, W)) \\ & \leq p \min(\varrho_{n,W}^2 \sigma_{m,V}^2, \varrho_{m,V}^2 \sigma_{n,W}^2) + p \varrho_{m \wedge n, V} \varrho_{m \wedge n, W} \sigma_{m \wedge n, V} \sigma_{m \wedge n, W} \\ & \quad + \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(V)) \right)^{1/2} \cdot \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(W)) \right)^{1/2} \\ & \quad + S_1(V, W) - S_0(V, W). \end{aligned}$$

(ii) *If  $p \neq q$ , then*

$$\begin{aligned} & \sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) \leq \min(q \varrho_{n,W}^2 \sigma_{m,V}^2, p \varrho_{m,V}^2 \sigma_{n,W}^2) - S_0(V, W) \\ & \quad + \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2(p \vee q)-1}} \text{Var}(U_M(V)) \right)^{1/2} \cdot \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2(p \vee q)-1}} \text{Var}(U_M(W)) \right)^{1/2}. \end{aligned}$$

**Proof** We begin with some computations that are valid in all different cases.

$$\begin{aligned}
\sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) &= \text{Var}(VW) - \sum_{\substack{M \subseteq [n \vee m]: \\ |M| = p+q}} \text{Var}(U_M(V, W)) \\
&= \text{Var}(VW) - \sum_{\substack{M \subseteq [n \vee m]: \\ |M| = p+q}} \text{Var}\left(\sum_{\substack{J \in \mathcal{D}_q(m), K \in \mathcal{D}_p(n): \\ J \cup K = M}} V_J W_K\right) \\
&= \text{Var}(VW) - \sum_{\substack{M \subseteq [n \vee m]: \\ |M| = p+q}} \sum_{\substack{I, J \in \mathcal{D}_q(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap K = J \cap L = \emptyset, \\ I \cup K = M = J \cup L}} \mathbb{E}[V_I V_J W_K W_L] \\
&= \text{Var}(VW) - \sum_{\substack{I, J \in \mathcal{D}_q(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap K = J \cap L = \emptyset}} \mathbb{E}[V_I V_J W_K W_L] \\
&= \text{Var}(VW) - \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] - \sum_{\substack{I, J \in \mathcal{D}_q(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap K = J \cap L = \emptyset, \\ \emptyset \neq I \cap J = I \setminus (I \cap L) \neq I, \\ \emptyset \neq J \cap I = J \setminus (J \cap K) \neq J}} \mathbb{E}[V_I V_J W_K W_L] \\
&\quad - \delta_{p,q} \sum_{\substack{I, K \in \mathcal{D}_q(m \wedge n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K]
\end{aligned}$$

Noting that

$$\text{Var}(VW) = \text{Cov}(V^2, W^2) + \mathbb{E}[V^2] \mathbb{E}[W^2] - \mathbb{E}[VW]^2$$

and also, for reasons of degeneracy, that

$$\sum_{\substack{I, J \in \mathcal{D}_q(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap K = J \cap L = \emptyset, \\ \emptyset \neq I \cap J = I \setminus (I \cap L) \neq I, \\ \emptyset \neq J \cap I = J \setminus (J \cap K) \neq J}} \mathbb{E}[V_I V_J W_K W_L] = S_0(V, W)$$

we hence obtain that

$$\begin{aligned}
\sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) &= \text{Cov}(V^2, W^2) + \mathbb{E}[V^2] \mathbb{E}[W^2] - \mathbb{E}[VW]^2 \\
&- \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] - S_0(V, W) - \delta_{p,q} \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K]. \quad (16)
\end{aligned}$$

We now deal with the individual terms appearing on the right hand side of (16).

$$\begin{aligned} \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] &= \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n)}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] - \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] \\ &= \mathbb{E}[V^2] \mathbb{E}[W^2] - \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2]. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) &= \text{Cov}(V^2, W^2) - \mathbb{E}[VW]^2 \\ &+ \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] - S_0(V, W) - \delta_{p,q} \sum_{\substack{I, K \in \mathcal{D}_q(m \wedge n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K]. \end{aligned}$$

Moreover, if  $p = q$  then

$$\begin{aligned} &\sum_{\substack{I, K \in \mathcal{D}_p(m \wedge n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K] \\ &= \left( \sum_{I \in \mathcal{D}_p(m \wedge n)} \mathbb{E}[V_I W_I] \right)^2 - \sum_{\substack{I, K \in \mathcal{D}_p(m \wedge n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K] \\ &= \mathbb{E}[VW]^2 - \sum_{\substack{I, K \in \mathcal{D}_p(m \wedge n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K]. \end{aligned}$$

Hence, using that  $\mathbb{E}[VW] = 0$  if  $p \neq q$ , we obtain that

$$\begin{aligned} \sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) &= \text{Cov}(V^2, W^2) - 2\delta_{p,q} \mathbb{E}[VW]^2 \\ &+ \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K = \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] - S_0(V, W) + \delta_{p,q} \sum_{\substack{I, K \in \mathcal{D}_p(m \wedge n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K]. \end{aligned} \tag{17}$$

In particular, if  $p \neq q$ , then

$$\sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) = \text{Cov}(V^2, W^2) + \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] - S_0(V, W). \quad (18)$$

Now,

$$\begin{aligned} \sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] &= \sum_{I \in \mathcal{D}_q(m)} \mathbb{E}[V_I^2] \sum_{i \in I} \sum_{\substack{K \in \mathcal{D}_p(n): \\ \max(K \cap I) = i}} \mathbb{E}[W_K^2] \\ &\leq q \varrho_{n,W}^2 \sum_{I \in \mathcal{D}_q(m)} \mathbb{E}[V_I^2] = q \varrho_{n,W}^2 \mathbb{E}[V]^2 = q \varrho_{n,W}^2 \sigma_{m,V}^2, \end{aligned}$$

and, analogously,

$$\sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] \leq p \varrho_{m,V}^2 \mathbb{E}[W]^2 = p \varrho_{m,V}^2 \sigma_{n,W}^2.$$

Hence,

$$\sum_{\substack{I \in \mathcal{D}_q(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I^2] \mathbb{E}[W_K^2] \leq \min(q \varrho_{n,W}^2 \sigma_{m,V}^2, p \varrho_{m,V}^2 \sigma_{n,W}^2). \quad (19)$$

We have

$$\begin{aligned} &\left| \sum_{\substack{I, K \in \mathcal{D}_p(m \wedge n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K] \right| \leq \sum_{I \in \mathcal{D}_p(m \wedge n)} |\mathbb{E}[V_I W_I]| \sum_{\substack{K \in \mathcal{D}_p(m \wedge n): \\ I \cap K \neq \emptyset}} |\mathbb{E}[V_K W_K]| \\ &= \sum_{I \in \mathcal{D}_p(m \wedge n)} |\mathbb{E}[V_I W_I]| \sum_{i \in I} \sum_{\substack{K \in \mathcal{D}_p(m \wedge n): \\ \max(I \cap K) = i}} |\mathbb{E}[V_K W_K]| \\ &\leq \sum_{I \in \mathcal{D}_p(m \wedge n)} |\mathbb{E}[V_I W_I]| \sum_{i \in I} \left( \sum_{\substack{K \in \mathcal{D}_p(m \wedge n): \\ \max(I \cap K) = i}} \mathbb{E}[V_K^2] \right)^{1/2} \cdot \left( \sum_{\substack{K \in \mathcal{D}_p(m \wedge n): \\ \max(I \cap K) = i}} \mathbb{E}[W_K^2] \right)^{1/2} \\ &\leq p \varrho_{m \wedge n, V} \varrho_{m \wedge n, W} \sum_{I \in \mathcal{D}_p(m \wedge n)} |\mathbb{E}[V_I W_I]| \\ &\leq p \varrho_{m \wedge n, V} \varrho_{m \wedge n, W} \left( \sum_{I \in \mathcal{D}_p(m \wedge n)} \mathbb{E}[V_I^2] \right)^{1/2} \cdot \left( \sum_{I \in \mathcal{D}_p(m \wedge n)} \mathbb{E}[W_I^2] \right)^{1/2} \\ &= p \varrho_{m \wedge n, V} \varrho_{m \wedge n, W} \sigma_{m \wedge n, V} \sigma_{m \wedge n, W}. \end{aligned}$$

Hence, we obtain the inequalities

$$\begin{aligned}
 -pQ_{m \wedge n, V} Q_{m \wedge n, W} \sigma_{m \wedge n, V} \sigma_{m \wedge n, W} &\leq \sum_{\substack{I \in \mathcal{D}_p(m), \\ K \in \mathcal{D}_p(n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K] \\
 &\leq pQ_{m \wedge n, V} Q_{m \wedge n, W} \sigma_{m \wedge n, V} \sigma_{m \wedge n, W}. \quad (20)
 \end{aligned}$$

We estimate the term  $\text{Cov}(V^2, W^2)$  separately according to whether  $p = q$  or not. First assume that  $p = q$ . Then, due to the orthogonality of Hoeffding components we have

$$\begin{aligned}
 \text{Cov}(V^2, W^2) &= \sum_{\substack{M \subseteq [m], \\ N \subseteq [n]: \\ 1 \leq |M|, |N| \leq 2p}} \mathbb{E}[U_M(V) U_N(W)] = \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p}} \mathbb{E}[U_M(V) U_M(W)] \\
 &= \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \mathbb{E}[U_M(V) U_M(W)] + \sum_{\substack{M \subseteq [m \wedge n]: \\ |M| = 2p}} \mathbb{E}[U_M(V) U_M(W)] \\
 &=: T_1 + T_2.
 \end{aligned}$$

Using the Cauchy-Schwarz inequality twice we can estimate

$$\begin{aligned}
 |T_1| &\leq \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \sqrt{\text{Var}(U_M(V))} \sqrt{\text{Var}(U_M(W))} \\
 &\leq \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(V)) \right)^{1/2} \cdot \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(W)) \right)^{1/2}.
 \end{aligned}$$

On the other hand, we have

$$\begin{aligned}
 T_2 &= \sum_{\substack{M \subseteq [m \wedge n]: \\ |M| = 2p}} \sum_{\substack{I, J \in \mathcal{D}_p(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cup J = K \cup L = M}} \mathbb{E}[V_I V_J W_K W_L] = \sum_{\substack{I, J \in \mathcal{D}_p(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap J = K \cap L = \emptyset, \\ I \cup J = K \cup L}} \mathbb{E}[V_I V_J W_K W_L] \\
 &= \sum_{\substack{I, J \in \mathcal{D}_p(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap J = K \cap L = \emptyset}} \mathbb{E}[V_I V_J W_K W_L] \\
 &= 2 \sum_{\substack{I, J \in \mathcal{D}_p(m \wedge n): \\ I \cap J = \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_J W_J] + \sum_{\substack{I, J \in \mathcal{D}_p(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap J = K \cap L = \emptyset, \\ \emptyset \neq I \cap K = I \setminus (I \cap L) \neq I}} \mathbb{E}[V_I V_J W_K W_L]
 \end{aligned}$$

$$\begin{aligned}
&= 2 \sum_{\substack{I, J \in \mathcal{D}_p(m \wedge n): \\ I \cap J = \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_J W_J] + S_1(V, W) \\
&= 2\mathbb{E}[VW]^2 - 2 \sum_{\substack{I, J \in \mathcal{D}_p(m \wedge n): \\ I \cap J \neq \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_J W_J] + S_1(V, W),
\end{aligned}$$

where we have used that, due to degeneracy,

$$\begin{aligned}
&\sum_{\substack{I, J \in \mathcal{D}_p(m), \\ K, L \in \mathcal{D}_p(n): \\ I \cap J = K \cap L = \emptyset, \\ \emptyset \neq I \cap K = I \setminus (I \cap L) \neq I}} \mathbb{E}[V_I V_J W_K W_L] = \sum_{\substack{I, J, K, L \in \mathcal{D}_p(m \wedge n): \\ I \cap J = K \cap L = \emptyset, \\ \emptyset \neq I \cap K = I \setminus (I \cap L) \neq I}} \mathbb{E}[V_I V_J W_K W_L] \\
&= \sum_{\substack{I, J, K, L \in \mathcal{D}_p(m \wedge n): \\ I \cap J = K \cap L = \emptyset, \\ \emptyset \neq I \cap K = I \setminus (I \cap L) \neq I, \\ \emptyset \neq J \cap L = J \setminus (J \cap K) \neq J}} \mathbb{E}[V_I V_J W_K W_L] = S_1(V, W).
\end{aligned}$$

Hence, if  $p = q$  then we obtain that

$$\begin{aligned}
\text{Cov}(V^2, W^2) &\leq \left( \sum_{\substack{M \subseteq [m \vee n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(V)) \right)^{1/2} \cdot \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(W)) \right)^{1/2} \\
&\quad + 2\mathbb{E}[VW]^2 - 2 \sum_{\substack{I, K \in \mathcal{D}_p(m \wedge n): \\ I \cap K \neq \emptyset}} \mathbb{E}[V_I W_I] \mathbb{E}[V_K W_K] + S_1(V, W).
\end{aligned} \tag{21}$$

Altogether, in the case  $p = q$ , we obtain from (17)-(21) that

$$\begin{aligned}
&\sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) \leq p \min(\varrho_{n, W}^2 \sigma_{m, V}^2, \varrho_{m, V}^2 \sigma_{n, W}^2) \\
&\quad + p \varrho_{m \wedge n, V} \varrho_{m \wedge n, W} \sigma_{m \wedge n, V} \sigma_{m \wedge n, W} + S_1(V, W) - S_0(V, W) \\
&\quad + \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(V)) \right)^{1/2} \cdot \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2p-1}} \text{Var}(U_M(W)) \right)^{1/2},
\end{aligned}$$

proving part (i). Let us now assume that  $p \neq q$ . Let  $r := p \wedge q < p \vee q =: s$ . In this case, clearly,  $\mathbb{E}[VW] = 0$ . Moreover, similarly as before, we have

$$\begin{aligned}
 \text{Cov}(V^2, W^2) &= \sum_{\substack{M \subseteq [m], N \subseteq [n]: \\ 1 \leq |M| \leq 2q, \\ 1 \leq |N| \leq 2p}} \mathbb{E}[U_M(V)U_N(W)] = \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2r}} \mathbb{E}[U_M(V)U_M(W)] \\
 &= \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2s-1}} \mathbb{E}[U_M(V)U_M(W)] \\
 &\leq \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2s-1}} \sqrt{\text{Var}(U_M(V))} \sqrt{\text{Var}(U_M(W))} \\
 &\leq \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2s-1}} \text{Var}(U_M(V)) \right)^{1/2} \cdot \left( \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2s-1}} \text{Var}(U_M(W)) \right)^{1/2}.
 \end{aligned} \tag{22}$$

This finishes the proof of (ii).  $\square$

**Corollary 2.6** *In the situation of Lemma 2.5 (i) assume that, additionally,  $m \leq n$  and  $V_J = W_J$  hold for all  $J \in \mathcal{D}_p(m)$ . Then, we have*

$$\begin{aligned}
 \sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq 2p-1}} \text{Var}(U_M(V, W)) &\leq \mathbb{E}[V^4] - 3\sigma_{m,V}^4 + 2p\sigma_{m,V}^2\varrho_{m,V}^2 + (C_{p,q} + 2p)\varrho_{n,W}^2\sigma_{n,W}^2 \\
 &+ \left( \mathbb{E}[V^4] - 3\sigma_{m,V}^4 + \kappa_p\sigma_{m,V}^2\varrho_{m,V}^2 \right)^{1/2} \cdot \left( \mathbb{E}[W^4] - 3\sigma_{n,W}^4 + \kappa_p\sigma_{n,W}^2\varrho_{n,W}^2 \right)^{1/2}
 \end{aligned}$$

where  $\kappa_p$  is a finite constant depending only on  $p$ .

**Proof** Since  $S_1(V, W) = S_0(V, V)$  in this case, the result follows immediately from Lemmas 2.5 (i), 2.4, 2.2, 2.3 and (15).  $\square$

**Corollary 2.7** *If, in the situation of Lemma 2.5 (ii), we additionally have  $p < q$ , then*

$$\begin{aligned}
 \sum_{\substack{M \subseteq [n \vee m]: \\ |M| \leq p+q-1}} \text{Var}(U_M(V, W)) &\leq (C_{p,q} + q) \max(\sigma_{m,V}^2\varrho_{n,W}^2, \sigma_{n,W}^2\varrho_{m,V}^2) \\
 &+ \left( \mathbb{E}[V^4] - 3\sigma_{m,V}^4 + \kappa_q\sigma_{m,V}^2\varrho_{m,V}^2 \right)^{1/2} \cdot \left( \mathbb{E}[W^4] - \sigma_{n,W}^4 \right)^{1/2},
 \end{aligned}$$

where  $\kappa_q$  is a finite constant depending only on  $q$ .

**Proof** This follows from Lemma 2.5 (ii), Lemma 2.4 (applied to  $V$ ) and from the obvious facts that

$$\begin{aligned} \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2(p \vee q) - 1}} \text{Var}(U_M(V)) &\leq \sum_{\substack{M \subseteq [m]: \\ 1 \leq |M| \leq 2q - 1}} \text{Var}(U_M(V)), \\ \sum_{\substack{M \subseteq [m \wedge n]: \\ 1 \leq |M| \leq 2(p \vee q) - 1}} \text{Var}(U_M(W)) &\leq \sum_{\substack{M \subseteq [n]: \\ 1 \leq |M| \leq 2p}} \text{Var}(U_M(W)) = \text{Var}(W^2) = \mathbb{E}[W^4] - \sigma_{n,W}^4. \end{aligned}$$

□

### 3 Proofs

In this section we provide detailed proofs of Theorem 1.4, Theorem 1.5 and relation (11).

#### 3.1 Proof of Theorem 1.4

The proof of Theorem 1.4 will follow the classical two-step procedure (see [4, Section 15]) of establishing convergence of finite dimensional distributions and checking tightness.

##### 3.1.1 Convergence of finite dimensional distributions

We fix time points  $0 \leq t_1 < t_2 < \dots < t_l \leq 1$  and, with  $r := ld$ , consider the random vector  $V = V^{(m)} = (V_1, \dots, V_r)^T \in \mathbb{R}^r$  of degenerate  $U$ -statistics, defined as follows. Given an integer  $1 \leq i \leq r$ , write  $i = al + b$  with integers  $0 \leq a \leq d - 1$  and  $1 \leq b \leq l$  and define

$$V_i := \mathbf{W}_{t_b}^{(m)}(a + 1).$$

In other words, we have

$$V = (\mathbf{W}_{t_1}^{(m)}(1), \dots, \mathbf{W}_{t_l}^{(m)}(1), \mathbf{W}_{t_1}^{(m)}(2), \dots, \mathbf{W}_{t_l}^{(m)}(2), \dots, \mathbf{W}_{t_1}^{(m)}(d), \dots, \mathbf{W}_{t_l}^{(m)}(d))^T.$$

Moreover, for the same  $i$  we write  $n_m(i) := \lfloor n_m t_b \rfloor$ ,  $\sigma_m(i)^2 := \text{Var}(V_i)$ ,  $q_i := p_{a+1}$  and

$$\phi_m(i)^2 := \max_{1 \leq j \leq n_m} \sum_{\substack{J \in \mathcal{D}_{p_{a+1}}(n_m(i)): \\ j \in J}} \mathbb{E}[W_J^{(m)}(a + 1)^2].$$



Then, we have the straightforward inequalities

$$\sigma_m(i)^2 \leq 1 \quad \text{and} \quad \phi_m(i)^2 \leq \varrho_{m,a+1}^2 \leq \max_{1 \leq k \leq d} \varrho_{m,k}^2. \quad (23)$$

Let  $\Sigma_m = (s_{i,k}^{(m)})_{1 \leq i,k \leq r} \in \mathbb{R}^{r \times r}$  denote the covariance matrix of  $V$ . Then, for  $1 \leq i \leq k \leq r$ , such that

$$i = al + b, \quad k = a'l + b', \quad \text{where} \quad 0 \leq a \leq a' \leq d-1 \quad \text{and} \quad 1 \leq b, b' \leq l, \quad (24)$$

we have

$$s_{i,k}^{(m)} = \mathbb{E}_m[\mathbf{W}_{ib}^{(m)}(a+1)\mathbf{W}_{ib'}^{(m)}(a'+1)] = \delta_{a,a'}\mathbb{E}_m[\mathbf{W}_{ib}^{(m)}(a+1)^2] = \delta_{a,a'}\sigma_m^2(i).$$

In particular, the diagonal elements of  $\Sigma_m$  are given by  $s_{i,i}^{(m)} = \sigma_m^2(i)$ ,  $1 \leq i \leq r$ . Thus, thanks to Condition 1.1 we have that  $\Sigma_m$  converges to the covariance matrix  $\Sigma = (s_{i,k})_{1 \leq i,k \leq r}$ , whose elements are given by

$$s_{i,k} = \delta_{a,a'} \lim_{m \rightarrow \infty} \sigma_m^2(i) = \delta_{a,a'} v_{a+1}(t_b), \quad (25)$$

if again  $i$  and  $k$  are as in (24). Finally, we denote by  $N^{(m)} = (N_1^{(m)}, \dots, N_r^{(m)})^T$  and  $N = (N_1, \dots, N_r)^T$  centered Gaussian vectors on  $(\Omega, \mathcal{F}, \mathbb{P})$  with covariance matrices  $\Sigma_m$  and  $\Sigma$ , respectively.

Next we make use of the well-known fact that the collection  $\mathcal{H}_3$  of all functions  $h \in C^3(\mathbb{R}^r)$ , each of whose partial derivative of order  $\leq 3$  is uniformly bounded, is convergence determining for weak convergence on  $\mathbb{R}^3$ . Then, by the triangle inequality we have

$$|\mathbb{E}_m[h(V)] - \mathbb{E}[h(N)]| \leq |\mathbb{E}_m[h(V)] - \mathbb{E}[h(N^{(m)})]| + |\mathbb{E}[h(N^{(m)})] - \mathbb{E}[h(N)]|. \quad (26)$$

Note that the second term on the right hand side of (26) converges to zero as  $m \rightarrow \infty$ , since  $\lim_{m \rightarrow \infty} \Sigma_m = \Sigma$ .

In order to deal with the first term as well, we apply the following central lemma, which is a slightly modified version of [18, Lemma 4.1 (a)], adapted to the present notation. The main difference is that in the statement of the result in [18], the term  $\phi_m(i)^2$  has already been replaced with the more concrete term  $q_i/n_m(i)$ , which is its value in the particular case of symmetric  $U$ -statistics. The proof given in [18], however, works in the more general present situation as well.

In what follows, for  $1 \leq i, k \leq r$ , we let

$$V_i V_k = \sum_{\substack{M \subseteq [n_m(i) \vee n_m(k)]: \\ |M| \leq q_i + q_k}} U_M(i, k)$$

denote the Hoeffding decomposition of  $V_i V_k$ .

**Lemma 3.1** *Under the above assumptions, there are constants  $\kappa_{q_i} \in (0, \infty)$ , only depending on  $q_i$ ,  $1 \leq i \leq r$ , such that the following holds: For any  $h \in \mathcal{H}_3$ , there are constants  $C_1(h), C_2(h) \in [0, \infty)$  such that*

$$\begin{aligned} |\mathbb{E}_m[h(V)] - \mathbb{E}[h(N^{(m)})]| &\leq \frac{C_1(h)}{4q_1} \sum_{i,k=1}^r (q_i + q_k) \left( \sum_{\substack{M \subseteq [n_m(i) \vee n_m(k)]: \\ |M| \leq q_i + q_k - 1}} \text{Var}_m(U_M(i, k)) \right)^{1/2} \\ &\quad + \frac{2C_2(h)\sqrt{r}}{9q_1} \sum_{i=1}^r q_i \sigma_m(i) \left( \sum_{\substack{M \subseteq [n_m(i)]: \\ |M| \leq 2q_i - 1}} \text{Var}_m(U_M(i, i)) \right)^{1/2} \\ &\quad + \frac{\sqrt{2r}C_2(h)}{9q_1} \sum_{i=1}^r q_i \sqrt{\kappa_{q_i}} \sigma_m(i)^3 \phi_m(i). \end{aligned} \quad (27)$$

We will also rely on the following crucial lemma, which essentially goes back to de Jong's monograph [10].

**Lemma 3.2** *Under Conditions 1.2 and 1.3 and with the notation of this subsection, we have that*

$$\lim_{m \rightarrow \infty} \mathbb{E}[V_i^4] - 3\sigma_m^4(i) = 0$$

for all  $1 \leq i \leq r$ .

**Proof** Fix  $1 \leq i \leq r$  given as in (24). Then, we have

$$V_i = \sum_{J \in \mathcal{D}_{p_{a+1}}(n_m(i))} W_J^{(m)}(a+1) = \sum_{J \in \mathcal{D}_{p_{a+1}}(n_m)} b_J W_J^{(m)}(a+1),$$

where, for  $J \in \mathcal{D}_{p_{a+1}}(n_m)$ , we let

$$b_J = \prod_{j \in J} \mathbb{1}_{\{1 \leq j \leq n_m(i)\}}.$$

This ensures that the coefficient sequence  $(b_J)_{J \in \mathcal{D}_{p_{a+1}}(n_m)}$  is of rank one in de Jong's terminology (see [10, Subsection 4.1] and Remark 3.3 (b), below) and we obviously have  $0 \leq \mathbb{1}_{\{1 \leq j \leq n_m(i)\}} \leq 1$  for all  $1 \leq j \leq n_m$ . Hence, by [10, Proposition 4.1.2.] we have  $\lim_{m \rightarrow \infty} \mathbb{E}[V_i^4] - 3\sigma_m^4(i) = 0$ .  $\square$

**Remark 3.3** (a) We mention that the statement of [10, Proposition 4.1.2.] actually only implies the statement  $\lim_{m \rightarrow \infty} \mathbb{E}[V_i^4] - 3\sigma_m^4(i) = 0$  of Lemma 3.2 under the stronger condition that  $D_{k,m}$  is bounded in  $m$  for each  $1 \leq k \leq d$ . However, as can be seen from the proof of [10, Proposition 4.1.2.], the result remains true under the more relaxed Condition 1.2, since the quantities  $\tau'$  and  $(\tau')^*$  appearing there still converge to zero, thanks to the inequality

$$(\tau')^* \leq D_m \tau' \lesssim D_m \varrho_m^2,$$

the second part of which holds true by virtue of [16, Proposition 2.9]. The first (and easy) part of it is proved in [10].

(b) In [10] the following definition of a *rank one coefficient sequence*  $(a_J)_{J \in \mathcal{D}_p(n_m)}$  is given: One has  $|a_J| \leq 1$  for all  $J \in \mathcal{D}_p(n_m)$  and there is a sequence  $(c_j)_{1 \leq j \leq n_m}$  such that  $a_J = \prod_{j \in J} c_j$  for all  $J \in \mathcal{D}_p(n_m)$ .

In order to prove that the first term on the right hand side of (26) converges to zero, we will make sure that the right hand side of (27) goes to zero as  $m \rightarrow \infty$ . By (23) and the assumption of Theorem 1.4, for the last term we have

$$\lim_{m \rightarrow \infty} \frac{\sqrt{2r}C_2(h)}{9q_1} \sum_{i=1}^r q_i \sqrt{\kappa_{q_i}} \sigma_m(i)^3 \phi_m(i) = 0. \quad (28)$$

In order to deal with the first two terms, we invoke Corollaries 2.6 and 2.7. Thus, suppose again that  $1 \leq i \leq k \leq r$  are given by (24).

**Case 1:**  $a < a'$ : In this case, we have  $q_i = p_{a+1} < p_{a'+1} = q_k$  and, hence, we are in the situation of Corollary 2.7. Thus, we have that

$$\begin{aligned} \sum_{\substack{M \subseteq [n_m(i) \vee n_m(k)]: \\ |M| \leq q_i + q_k - 1}} \text{Var}_m(U_M(i, k)) &\leq (C_{q_i, q_k} + q_k) \max(\sigma_m^2(i) \phi_m^2(k), \sigma_m^2(k) \phi_m^2(i)) \\ &+ \left( \mathbb{E}[V_k^4] - 3\sigma_m^4(k) + \kappa_{q_k} \sigma_m^2(k) \phi_m^2(k) \right)^{1/2} \cdot \left( \mathbb{E}[V_i^4] - \sigma_m^4(i) \right)^{1/2} \\ &\leq (C_{q_i, q_k} + q_k) \max_{1 \leq u \leq d} \varrho_{m,u}^2 \\ &+ \left( \mathbb{E}[V_k^4] - 3\sigma_m^4(k) + \kappa_{q_k} \max_{1 \leq u \leq d} \varrho_{m,u}^2 \right)^{1/2} \cdot \left( \mathbb{E}[V_i^4] \right)^{1/2}, \end{aligned} \quad (29)$$

where we have applied (23) for the second inequality and used  $\kappa_{q_k}$  to denote a finite constant depending only on  $q_k$ . By Lemma 3.2 and Condition 1.2, the right hand side of (29) thus converges to zero as  $m \rightarrow \infty$ .

**Case 2:**  $a = a'$ : In this case, we have  $q_i = p_{a+1} = p_{a'+1} = q_k$  and, hence, we are in the situation of Corollary 2.6. Thus, we have that

$$\begin{aligned} \sum_{\substack{M \subseteq [n_m(i) \vee n_m(k)]: \\ |M| \leq q_i + q_k - 1}} \text{Var}_m(U_M(i, k)) &= \sum_{\substack{M \subseteq [n_m(k)]: \\ |M| \leq q_i + q_k - 1}} \text{Var}_m(U_M(i, k)) \\ &\leq \mathbb{E}[V_i^4] - 3\sigma_m^4(i) + 2q_i \sigma_m^2(i) \phi_m^2(i) + (C_{q_i, q_i} + 2q_i) \phi_m^2(k) \sigma_m^2(k) \\ &+ \left( \mathbb{E}[V_i^4] - 3\sigma_m^4(i) + \kappa_{q_i} \sigma_m^2(i) \phi_m^2(i) \right)^{1/2} \cdot \left( \mathbb{E}[V_k^4] - 3\sigma_m^4(k) + \kappa_{q_k} \sigma_m^2(k) \phi_m^2(k) \right)^{1/2} \\ &\leq \mathbb{E}[V_i^4] - 3\sigma_m^4(i) + 2q_i \max_{1 \leq u \leq d} \varrho_{m,u}^2 + (C_{q_i, q_i} + 2q_i) \max_{1 \leq u \leq d} \varrho_{m,u}^2 \\ &+ \left( \mathbb{E}[V_i^4] - 3\sigma_m^4(i) + \kappa_{q_i} \max_{1 \leq u \leq d} \varrho_{m,u}^2 \right)^{1/2} \cdot \left( \mathbb{E}[V_k^4] - 3\sigma_m^4(k) + \kappa_{q_k} \max_{1 \leq u \leq d} \varrho_{m,u}^2 \right)^{1/2}, \end{aligned} \quad (30)$$

where we have again applied (23) for the second inequality and used  $\kappa_{q_k}$  and  $\kappa_{q_i}$  to denote finite constants depending only on  $q_k$  and  $q_i$ , respectively. Again, by Lemma 3.2 and Condition 1.2, the right hand side of (30) thus converges to zero as  $m \rightarrow \infty$ .

From (27), (23), (28), (29) and (30) we thus conclude that

$$\lim_{m \rightarrow \infty} |\mathbb{E}_m[h(V)] - \mathbb{E}[h(N^{(m)})]| = 0$$

and, thus, by (26) that

$$\lim_{m \rightarrow \infty} |\mathbb{E}_m[h(V)] - \mathbb{E}[h(N)]| = 0$$

for each  $h \in \mathcal{H}_3$ . Note further that, by the definition of  $\Sigma$ , we have the distributional identity

$$N \stackrel{\mathcal{D}}{=} (\mathbf{Z}_{t_1}(1), \dots, \mathbf{Z}_{t_l}(1), \dots, \mathbf{Z}_{t_1}(d), \dots, \mathbf{Z}_{t_l}(d))^T.$$

Hence, we have established the convergence of the finite dimensional distributions of  $\mathbf{W}$  to those of  $\mathbf{Z}$ .

### 3.1.2 Conclusion of the argument

We will now finish the proof of Theorem 1.4 by proving tightness of the sequence  $(\mathbf{W}^{(m)})_{m \in \mathbb{N}}$  in  $D[0, 1]$ . This will be done by verifying that

$$\omega(\mathbf{W}^{(m)}, \delta) \xrightarrow{\mathbb{P}_m} 0, \quad \text{as } m \rightarrow \infty, \quad (31)$$

where, for  $x \in D[0, 1]$  and  $\delta > 0$ ,

$$\omega(x, \delta) := \sup\{|x(t) - x(s)| : s, t \in [0, 1], |t - s| < \delta\}$$

denotes the *modulus of continuity* of  $x$ . Since  $\mathbf{W}_0^{(m)} = 0$  for all  $m \in \mathbb{N}$  and by the finite-dimensional distribution convergence established in Sect. 3.1.1, this will imply Theorem 1.4 by the Corollary to [5, Theorem 13.4]. Let us introduce the following notation: Fix  $1 \leq l \leq d$  that we will from now on suppress from the notation whenever convenient, and, for  $m \in \mathbb{N}$  and  $0 \leq j \leq n_m$ , define  $\Delta_j^{(m)} := \mathbf{W}_{\frac{j}{n_m}}^{(m)}(l) - \mathbf{W}_{\frac{j-1}{n_m}}^{(m)}(l)$ ,  $\mathcal{F}_j^{(m)} := \sigma(X_1^{(m)}, \dots, X_j^{(m)})$  and  $S_j^{(m)} := \sum_{i=1}^j \Delta_i^{(m)}$ . Then, by degeneracy, for each  $m \in \mathbb{N}$ ,  $(S_j^{(m)})_{1 \leq j \leq n_m}$  is an  $(\mathcal{F}_j^{(m)})_{0 \leq j \leq n_m}$ -martingale. Note that with this notation at hand, we have  $\mathbf{W}_t^{(m)} = S_{\lfloor n_m t \rfloor}^{(m)}$  for each  $t \in [0, 1]$  and [4, Theorem 8.4] and the Corollary to [5, Theorem 13.4] together imply that (31) holds, if, for each  $k \in \mathbb{N}$  we have

$$\lim_{\lambda \rightarrow \infty} \limsup_{m \rightarrow \infty} \lambda^2 \mathbb{P}_m \left( \max_{0 \leq i \leq n_m - k} |S_{k+i}^{(m)} - S_k^{(m)}| \geq \lambda \right) = 0. \quad (32)$$

Note that the additional  $\sqrt{n}$  appearing in the statement of [4, Theorem 8.4] does not appear here, since our processes are already normalized at time  $t = 1$ . Let us thus fix  $k \in \mathbb{N}$ . Then, the sequence  $(S_{k+i}^{(m)} - S_k^{(m)})_{0 \leq i \leq n_m - k}$  is an  $(\mathcal{F}_{k+i}^{(m)})_{0 \leq i \leq n_m - k}$ -martingale. Thus, by first applying Markov's and then Doob's  $L^4$ -inequality, we have that

$$\begin{aligned} & \lambda^2 \mathbb{P}_m \left( \max_{0 \leq i \leq n_m - k} |S_{k+i}^{(m)} - S_k^{(m)}| \geq \lambda \right) \\ & \leq \mathbb{E}_m \left[ \left( \max_{0 \leq i \leq n_m - k} |S_{k+i}^{(m)} - S_k^{(m)}| \right)^2 \mathbb{1}_{\{\max_{0 \leq i \leq n_m - k} |S_{k+i}^{(m)} - S_k^{(m)}| \geq \lambda\}} \right] \\ & \leq \frac{1}{\lambda^2} \mathbb{E}_m \left[ \left( \max_{0 \leq i \leq n_m - k} |S_{k+i}^{(m)} - S_k^{(m)}| \right)^4 \right] \leq \left( \frac{4}{3} \right)^4 \frac{1}{\lambda^2} \mathbb{E}_m \left[ |S_{n_m}^{(m)} - S_k^{(m)}|^4 \right] \\ & \leq \left( \frac{4}{3} \right)^4 \frac{8}{\lambda^2} \left( \mathbb{E}_m |S_{n_m}^{(m)}|^4 + \mathbb{E}_m |S_k^{(m)}|^4 \right). \end{aligned} \quad (33)$$

Now, firstly, by Condition 1.3, we have

$$\lim_{m \rightarrow \infty} \mathbb{E}_m |S_{n_m}^{(m)}|^4 = \lim_{m \rightarrow \infty} \mathbb{E}_m [W^{(m)}(l)^4] = 3. \quad (34)$$

Secondly, using the inequality  $|\sum_{i=1}^n a_i|^4 \leq n^3 \sum_{i=1}^n |a_i|^4$ , and by Condition 1.2 we obtain that

$$\begin{aligned} \mathbb{E}_m |S_k^{(m)}|^4 &= \mathbb{E}_m \left| \sum_{J \in D_{p_l}(k)} W_J^{(m)}(l) \right|^4 \leq \binom{k}{p}^3 \sum_{J \in D_{p_l}(k)} \mathbb{E}_m |W_J^{(m)}(l)|^4 \\ &\leq \binom{k}{p_l}^3 D_{m,l} \sum_{J \in D_{p_l}(k)} \left( \mathbb{E}_m |W_J^{(m)}(l)|^2 \right)^2 \\ &\leq \binom{k}{p_l}^3 D_{m,l} \mathcal{Q}_{m,l}^2 \sum_{J \in D_{p_l}(k)} \mathbb{E}_m |W_J^{(m)}(l)|^2 \\ &\leq \binom{k}{p_l}^3 D_{m,l} \mathcal{Q}_{m,l}^2 \xrightarrow{m \rightarrow \infty} 0, \end{aligned} \quad (35)$$

where we have used the fact that  $\mathbb{E}_m |W_J^{(m)}(l)|^2 \leq \mathcal{Q}_{m,l}^2$  for all  $J \in D_{p_l}(k)$ . Thus, (33)-(35) imply that

$$\limsup_{m \rightarrow \infty} \lambda^2 \mathbb{P}_m \left( \max_{0 \leq i \leq n_m - k} |S_{k+i}^{(m)} - S_k^{(m)}| \geq \lambda \right) \leq \left( \frac{4}{3} \right)^4 \frac{24}{\lambda^2} = \frac{2048}{27} \frac{1}{\lambda^2}$$

for each  $\lambda > 0$  and  $k \in \mathbb{N}$  fixed, so that (32) is indeed satisfied. This finishes the argument for tightness.

### 3.2 An additional argument explaining the continuity of the limiting process $\mathbf{Z}$

Note that, by [5, Theorem 13.4] our conditions necessarily entail the remarkable fact that the limiting Gaussian process  $\mathbf{Z}$  has continuous paths, which might be quite surprising at first glance. Thus, we will now give another argument for this fact, which additionally explains why it is true. To this end, define the vector  $\mathbf{V}^{(m)} = (\mathbf{V}^{(m)}(1), \dots, \mathbf{V}^{(m)}(d))$  of continuous processes, defined by linear interpolation, i.e. for  $1 \leq l \leq d$  we let

$$\mathbf{V}_t^{(m)}(l) := \mathbf{W}_{\frac{i-1}{n_m}}^{(m)}(l) + (n_m t - (i-1))(\mathbf{W}_{\frac{i}{n_m}}^{(m)}(l) - \mathbf{W}_{\frac{i-1}{n_m}}^{(m)}(l)), \quad \frac{i-1}{n_m} \leq t < \frac{i}{n_m},$$

and  $\mathbf{V}_1^{(m)}(l) = \mathbf{W}_1^{(m)}(l) = W^{(m)}(l)$ . Then, for fixed  $1 \leq l \leq d$  and with the same notation as above, it is easy to see that

$$\|\mathbf{V}^{(m)}(l) - \mathbf{W}^{(m)}(l)\|_\infty = \max_{1 \leq i \leq n_m} |\Delta_i^{(m)}|.$$

Thus, using e.g. [25, Theorem 3] in the third inequality below, for  $\epsilon > 0$  we can bound

$$\begin{aligned} \mathbb{P}_m(\|\mathbf{V}^{(m)}(l) - \mathbf{W}^{(m)}(l)\|_\infty > \epsilon) &= \mathbb{P}_m\left(\max_{1 \leq i \leq n_m} |\Delta_i^{(m)}| > \epsilon\right) \leq \sum_{i=1}^{n_m} \mathbb{P}_m(|\Delta_i^{(m)}| > \epsilon) \\ &\leq \epsilon^{-4} \sum_{i=1}^{n_m} \mathbb{E}_m \left| \sum_{\substack{J \in \mathcal{D}_{p_l}(i): \\ i \in J}} W_J^{(m)}(l) \right|^4 \leq B(p_l) \epsilon^{-4} \sum_{i=1}^{n_m} \mathbb{E}_m \left[ \left( \sum_{\substack{J \in \mathcal{D}_{p_l}(i): \\ i \in J}} W_J^{(m)}(l)^2 \right)^2 \right] \\ &= B(p_l) \epsilon^{-4} \sum_{i=1}^{n_m} \sum_{\substack{J, K \in \mathcal{D}_{p_l}(i): \\ i \in J \cap K}} \mathbb{E}_m [W_J^{(m)}(l)^2 W_K^{(m)}(l)^2] \\ &\leq B(p_l) \epsilon^{-4} \sum_{i=1}^{n_m} \sum_{\substack{J, K \in \mathcal{D}_{p_l}(i): \\ i \in J \cap K}} \left( \mathbb{E}_m [W_J^{(m)}(l)^4] \mathbb{E}_m [W_K^{(m)}(l)^4] \right)^{1/2} \\ &\leq B(p_l) \epsilon^{-4} D_{m,l} \sum_{i=1}^{n_m} \sum_{\substack{J, K \in \mathcal{D}_{p_l}(i): \\ i \in J \cap K}} \mathbb{E}_m [W_J^{(m)}(l)^2] \mathbb{E}_m [W_K^{(m)}(l)^2] \\ &= B(p_l) \epsilon^{-4} D_{m,l} \sum_{i=1}^{n_m} \left( \sum_{\substack{J \in \mathcal{D}_{p_l}(i): \\ i \in J}} \mathbb{E}_m [W_J^{(m)}(l)^2] \right)^2 \\ &\leq B(p_l) \epsilon^{-4} D_{m,l} \varrho_{m,l}^2 \sum_{i=1}^{n_m} \sum_{\substack{J \in \mathcal{D}_{p_l}(i): \\ i \in J}} \mathbb{E}_m [W_J^{(m)}(l)^2] \\ &= B(p_l) \epsilon^{-4} D_{m,l} \varrho_{m,l}^2 \xrightarrow{m \rightarrow \infty} 0, \end{aligned}$$

where  $B(p_l)$  is a finite constant only depending on  $p_l$ , see [25, Theorem 3]. Thus, we have proved that, for each  $1 \leq l \leq d$ ,  $\|\mathbf{V}^{(m)}(l) - \mathbf{W}^{(m)}(l)\|_\infty$  and, a fortiori their Skorohod distance, converges to zero in probability as  $m \rightarrow \infty$ , so that [5, Theorem 3.1] implies that the processes  $\mathbf{V}^{(m)}(l)$  and  $\mathbf{W}^{(m)}(l)$  must indeed have the same distributional limit, if any.

### 3.3 Proof of Theorem 1.5

Suppose that Conditions 1.2 and 1.3 hold for the sequence  $(\mathbf{W}^{(m)})_{m \in \mathbb{N}}$ . Since the above argument for tightness does not make use of Condition 1.1, tightness still holds. Let  $(\mathbf{W}^{(m_l)})_{l \in \mathbb{N}}$  be a given subsequence. By relative compactness, there is a further subsequence  $(\mathbf{W}^{(m_{l_k})})_{k \in \mathbb{N}}$  that converges in distribution to some process  $\mathbf{Z}$ . Again by [5, Theorem 13.4] we can conclude that  $\mathbf{Z}$  has values in  $C([0, 1]; \mathbb{R}^d)$ . Thus, the projections  $\pi_t : D([0, 1]; \mathbb{R}^d) \rightarrow \mathbb{R}^d$ ,  $x \mapsto x(t)$ ,  $t \in [0, 1]$ , are  $\mathbb{P}_{\mathbf{Z}}$ -a.s. continuous, where  $\mathbb{P}_{\mathbf{Z}}$  denotes the law of  $\mathbf{Z}$  on  $D([0, 1]; \mathbb{R}^d)$ . Hence, the continuous mapping theorem implies that the finite-dimensional distributions of  $(\mathbf{W}^{(m_{l_k})})_{k \in \mathbb{N}}$  converge to those of  $\mathbf{Z}$  and, e.g. by a uniform integrability argument, one makes sure that the covariance matrices of the finite dimensional distributions of  $(\mathbf{W}^{(m_{l_k})})_{k \in \mathbb{N}}$  also converge to the respective ones of  $\mathbf{Z}$ . Thus, the argument leading to (25) goes through with the right hand side replaced by the variance of the corresponding coordinate of  $\mathbf{Z}$  at  $t_b$ . Now, since Lemma 3.2 applies under Conditions 1.2 and 1.3, the same reasoning as in Sect. 3.1.1 ensures that the finite dimensional distributions of  $(\mathbf{W}^{(m_{l_k})})_{k \in \mathbb{N}}$  have a Gaussian limit, implying that  $\mathbf{Z}$  is indeed a continuous, centered Gaussian process. The non-correlation, and therefore the independence, of the components of  $\mathbf{Z}$  follows from the fact that Hoeffding degenerate  $U$ -statistics of distinct orders are orthogonal in  $L^2$ . The proof is complete.

### 3.4 Proof of relation (11)

We fix  $p \geq 3$  and  $2 \leq a \leq p-1$  as in Example 1.12, and adopt the notation introduced therein. It is sufficient to show that

$$|F_m^0 \cap [\lfloor tm \rfloor]^p| \sim |F_{\lfloor tm \rfloor}^0|, \quad (36)$$

a relation that is implied by the next statement.

**Lemma 3.4** *Consider a nondecreasing integer-valued sequence  $(\ell(m))_{m \geq 1}$  such that, as  $m \rightarrow \infty$ ,  $\ell(m) \sim \alpha \cdot m$ , for some  $\alpha \in (0, 1)$ . Then*

$$|F_m^0 \cap [\ell(m)]^p| \sim |F_{\ell(m)}^0|.$$

**Proof** For every  $\ell \geq 1$  we write  $k = k(\ell)$  to denote the unique integer such that  $\ell \in [k^a, (k+1)^a)$ , in such a way that  $F_\ell = F_{k^a(\ell)}$ . We have that

$$|F_m^0 \cap [\ell(m)]^p| = |F_{\ell(m)}^0| + |(F_m^0 \cap [\ell(m)]^p) \setminus F_{k^a(\ell(m))}^0|.$$

## Using the relations

$$\begin{aligned} |(F_m^0 \cap [\ell(m)]^p) \setminus F_{k^a(\ell(m))}^0| &\leq |F_{(k(\ell(m))+1)^a}^0| - |F_{k^a(\ell(m))}^0| \sim k(\ell(m))^{p-1} \\ &\leq \ell(m)^{(p-1)/a} = o(m^{p/a}), \end{aligned}$$

we immediately deduce the desired conclusion.  $\square$

**Acknowledgements** This work is part of the first author's habilitation thesis at Heinrich Heine Universität Düsseldorf. The research leading to this paper was supported by the FNR grant **FoRGES (R-AGR-3376-10)** at Luxembourg University. This work is also part of project **Stein-ML** that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No **101024264**. We thank an anonymous referee for several useful remarks. *Data sharing not applicable to this article as no datasets were generated or analysed during the current study.*

## References

1. Barbour, A.D.: Stein's method for diffusion approximations. *Probab. Theory Relat. Fields* **84**(3), 297–322 (1990)
2. Basalykas, A.: Functional limit theorems for random multilinear forms. *Stoch. Process. Appl.* **53**(1), 175–191 (1994)
3. Bourguin, S., Campese, S.: Approximation of Hilbert-Valued Gaussians on Dirichlet structures. *Electron. J. Probab.* **25**, 1–30 (2020)
4. Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York (1968)
5. Billingsley, P.: *Convergence of Probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. Wiley & Sons, Inc., New York, second edition (1999)
6. Blei, R., Janson, S.: Rademacher chaos: tail estimates versus limit theorems. *Ark. Mat.* **42**(1), 13–29 (2004)
7. Blei, R.: *Analysis in Integer and Fractional Dimensions*. Cambridge University Press, Cambridge (2001)
8. Caravenna, F., Sun, R., Zygouras, N.: Universality in marginally relevant disordered systems. *Ann. Appl. Probab.* **27**(5), 3050–3112 (2017)
9. Caravenna, F., Sun, R., Zygouras, N.: The two-dimensional KPZ equation in the entire subcritical regime. *Ann. Probab.* **48**(3), 1086–1127 (2020)
10. de Jong, P.: Central limit theorems for generalized multilinear forms. CWI Tract, vol. 61. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam (1989)
11. de Jong, P.: A central limit theorem for generalized multilinear forms. *J. Multivar. Anal.* **34**(2), 275–289 (1990)
12. Döbler, C., Kasprzak, M.: Stein's method of exchangeable pairs in multivariate functional approximations. *Electron. J. Probab.* **26**(none), 1–50 (2021)
13. Döbler, C., Kasprzak, M., Peccati, G.: Functional convergence of sequential  $U$ -processes with size-dependent kernels. *Ann. Appl. Probab.*, to appear (2021)
14. Dynkin, E.B., Mandelbaum, A.: Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Stat.* **11**(3), 739–745 (1983)
15. Döbler, C.: Normal approximation via non-linear exchangeable pairs. [arXiv:2008.02272](https://arxiv.org/abs/2008.02272) (2020)
16. Döbler, C., Peccati, G.: Quantitative de Jong Theorems in any dimension. *Electron. J. Probab.* **22**(2), 1–35 (2017)
17. Döbler, C., Peccati, G.: The fourth moment theorem on the Poisson space. *Ann. Probab.* **46**(4), 1878–1916 (2018)
18. Döbler, C., Peccati, G.: Quantitative CLTs for symmetric  $U$ -statistics using contractions. *Electron. J. Probab.* **24**, 1–43 (2019)
19. Döbler, C., Vidotto, A., Zheng, G.: Fourth moment theorems on the Poisson space in any dimension. *Electron. J. Probab.*, 23:Paper No. 36, 27 (2018)



20. Efron, B., Stein, Ch.: The jackknife estimate of variance. *Ann. Stat.* **9**(3), 586–596 (1981)
21. Gregory, G.G.: Large sample theory for  $U$ -statistics and tests of fit. *Ann. Stat.* **5**(1), 110–123 (1977)
22. Garban, Ch., Steif, J.: *Noise Sensitivity of Boolean Functions and Percolation*. Cambridge University Press, Cambridge (2015)
23. Heyde, C.C., Brown, B.M.: On the departure from normality of a certain class of martingales. *Ann. Math. Stat.* **41**, 2161–2165 (1970)
24. Hoeffding, W.: A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.*, **19**, (1948)
25. Ibragimov, R., Sharakhmetov, Sh.: Bounds on moments of symmetric statistics. *Studia Sci. Math. Hungar.* **39**(3–4), 251–275 (2002)
26. Kasprzak, M.J.: Functional approximations via Stein’s method of exchangeable pairs. *Ann. Inst. Henri Poincaré Probab. Stat.* **56**(4), 2540–2564 (2020)
27. Kasprzak, M.J.: Stein’s method for multivariate Brownian approximations of sums under dependence. *Stoch. Process. Appl.* **130**(8), 4927–4967 (2020)
28. Koroljuk, V.S., Borovskich, Yu.V.: *Theory of  $U$ -Statistics*. Kluwer, Boston (1994)
29. Koike, Y.: Gaussian approximation of maxima of wiener functionals and its application to high-frequency data. *Ann. Stat.* **47**(3), 1663–1687 (2019)
30. Lachièze-Rey, R., Reitzner, M.:  $U$ -statistics in stochastic geometry. *Stochastic analysis for Poisson point processes*, Bocconi Springer Ser. **7**(7), 229–253 (2016)
31. Mikosch, T.: Functional limit theorems for random quadratic forms. *Stoch. Process. Appl.* **37**(1), 81–98 (1991)
32. Mikosch, T.: A weak invariance principle for weighted  $U$ -statistics with varying kernels. *J. Multivar. Anal.* **47**(1), 82–102 (1993)
33. Mossel, E., O’Donnell, R., Oleszkiewicz, K.: Noise stability of functions with low influences: invariance and optimality. *Ann. of Math. (2)* **171**(1), 295–341 (2010)
34. Merlevède, F., Peligrad, M., Utev, S.: *Functional Gaussian approximation for dependent structures*. Oxford Studies in Probability, vol. 6. Oxford University Press, Oxford (2019)
35. Nourdin, I., Peccati, G.: *Normal Approximations with Malliavin Calculus. From Stein’s Method to Universality*. Cambridge University Press, Cambridge (2012)
36. Nourdin, I., Peccati, G., Poly, G., Simone, R.: Classical and free fourth moment theorems: universality and thresholds. *J. Theoret. Probab.* **29**(2):653–680 (2016)
37. Nourdin, I., Peccati, G., Reinert, G.: Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *Ann. Probab.* **38**(5), 1947–1985 (2010)
38. Nourdin, I., Peccati, G., Reinert, G.: Stein’s method and stochastic analysis of rademacher functionals. *Electron. J. Probab.* **15**(55), 1703–1742 (2010)
39. O’Donnell, R.: *Analysis of Boolean Functions*. Cambridge University Press, Cambridge (2014)
40. Peccati, G., Nourdin, I.: Universal gaussian fluctuations of non-hermitian matrix ensembles: from weak convergence to almost sure clts. *ALEA Lat. Am. J. Probab. Math. Stat.* **7**, 341–375 (2010)
41. Peccati, G., Taqqu, M.S.: *Wiener chaos: moments, cumulants and diagrams. A survey with computer implementation.*, volume 1 of Bocconi & Springer Series. Springer, Milan (2011)
42. Peccati, G., Zheng, C.: Universal Gaussian fluctuations on the discrete Poisson chaos. *Bernoulli* **20**(2), 697–715 (2014)
43. Rotar, V.I.: Limit theorems for polylinear forms. *J. Multivar. Anal.* **9**(4), 511–530 (1979)
44. Serfling, R.J.: *Approximation theorems of mathematical statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, (1980)
45. Whitt, W.: Proofs of the martingale FCLT. *Probab. Surv.* **4**, 268–302 (2007)