Electron. J. Probab. **22** (2017), no. 2, 1–35.
ISSN: 1083-6489 DOI: 10.1214/16-EJP19

Quantitative de Jong theorems in any dimension

Christian Döbler* Giovanni Peccati†

Abstract

We develop a new quantitative approach to a multidimensional version of the well-known *de Jong's central limit theorem* under optimal conditions, stating that a sequence of Hoeffding degenerate U-statistics whose fourth cumulants converge to zero satisfies a CLT, as soon as a Lindeberg-Feller type condition is verified. Our approach allows one to deduce explicit (and presumably optimal) Wasserstein bounds in the case of general U-statistics of arbitrary order $d \geq 1$. One of our main findings is that, for vectors of U-statistics satisfying de Jong' s conditions and whose covariances admit a limit, componentwise convergence systematically implies joint convergence to Gaussian: this is the first instance in which such a phenomenon is described outside the frameworks of homogeneous chaoses and of diffusive Markov semigroups.

Keywords: quantitative CLTs; de Jong's Theorem; exchangeable pairs; Hoeffding decomposition; degenerate U-statistics; multidimensional convergence; Stein's method.
AMS MSC 2010: 60F05; 62E17; 62E20.
Submitted to EJP on April 24, 2016, final version accepted on December 15, 2016.
Supersedes arXiv:1603.00804v4.

1 Introduction, framework and main results

1.1 Overview

Let $\{W_n : n \geq 1\}$ be a sequence of unit variance U-statistics of order $d \geq 1$ (not necessarily symmetric) with underlying independent data $X_1, \ldots, X_n$, that are degenerate in the sense of Hoeffding (see Section 1.2 for formal definitions) and have a finite fourth moment. In the landmark paper [12] (see also [11]), P. de Jong proved the following remarkable fact, valid as $n \to \infty$: if $\mathbb{E}[W_n^4] \to 3$ and a Lindeberg-Feller-type condition is verified, then W_n converges in distribution towards a standard Gaussian random variable Z (note that $3 = \mathbb{E}[Z^4]$). This surprising result represents a drastic simplification of the method of moments and cumulants (see e.g. [41, Section A.3]), which should be

*Université du Luxembourg, Unité de Recherche en Mathématiques, Campus Kirchberg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg. E-mail: christian.doebler@uni.lu
†Université du Luxembourg, Unité de Recherche en Mathématiques, Campus Kirchberg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg. E-mail: giovanni.peccati@uni.lu

contrasted with the 'typical' non-central asymptotic behaviour of degenerate U-statistics of a fixed order $d \geq 2$ and with a fixed kernel — see e.g. [22], [62], [60], [15] or [27, Ch. 11] ; it also provides a general explanation of the ubiquitous emergence of the Gaussian distribution in geometric models where counting statistics can be naturally represented in terms of degenerate U-statistics, see e.g. [26, 48, 54].

One should notice that de Jong's central limit theorem (CLT) is a one dimensional *qualitative* statement: in particular, it does not provide any meaningful information about the rate of convergence of the law of W_n towards the target Gaussian distribution. Our aim in this paper is to use *Stein's method of exchangeable pairs*, as originally developed in Stein's monograph [63], in order to prove new quantitative and multidimensional versions of de Jong's central limit theorem under minimal conditions, in the setting of *degenerate and non-symmetric* U-statistics that do not necessarily have the form of homogeneous sums. In particular, we are interested in characterizing the joint convergence of those vectors of degenerate U-statistics, whose components verify one-dimensional CLTs.

One of the main motivations for pursuing our goal is that the findings of [11] have anticipated a modern and very fruitful direction of research, where tools of infinite-dimensional calculus are used in order to deduce *fourth moment theorems* in the spirit of de Jong (but, crucially, without the use of Lindeberg-Feller-type conditions) for random variables belonging to *the homogeneous chaos* of some general random field. The best-known results in this area gravitate around the main discovery of [46] (as well as its multidimensional extension [51]), where it is proved that a sequence of normalized random variables $\{Y_n : n \geq 1\}$, belonging to a fixed *Wiener chaos* of a Gaussian field, verifies a central limit theorem (CLT) if and only if $\mathbb{E}[Y_n^4] \to 3$. The combined use of *Malliavin calculus* and *Stein's method* has consequently allowed one to deduce strong quantitative versions with explicit Berry-Esseen bounds of these results (see [40, 41]), and it is therefore a natural question to ask whether the original CLT by de Jong can be endowed with explicit bounds, that are comparable with those available in a Gaussian setting.

The reader can consult the constantly updated webpage

`https://sites.google.com/site/malliavinstein/home`

for an overview of the emerging domain of research connected to [40, 41, 46, 51]. Among the many notable ramifications of the results of [40, 46] to which our findings should be compared, we quote: [32, 44, 55] for results involving homogeneous sums in the *Rademacher* (also called *Walsh*) chaos, [17, 34, 35, 49, 52, 57, 61] for the analysis of Poissonized U-statistics living in the Wiener chaos associated with a Poisson measure, [1, 5, 30, 45] for fourth moment theorems involving homogeneous sums in a non-commutative setting, and [2, 6, 36] for results in the setting of chaotic random variables associated with a diffusive Markov semigroup. Central and non-central quantitative versions of de Jong's results in the case of fully symmetric Poissonized U-statistics can be found in [17, 20, 50].

Two sets of references are particularly relevant for the present work:

(a) In reference [43] (see also [53]) de Jong's CLT in the special case of *homogeneous sums* was studied in the framework of the powerful theory of universality and influence functions initiated in [39]. In particular, explicit bounds were obtained for vectors of homogeneous sums satisfying a CLT.

(b) In the already quoted reference [51], the following striking phenomenon was discovered. For $r \geq 2$, let $Y^m = (Y_1^m, ..., Y_r^m)$, $m \geq 1$, be a sequence of random vectors whose components live in a fixed Wiener chaos, and assume that the covariance

matrix of Y^m converges to some $\mathbf{\Sigma} \geq 0$ and that each component Y_i^m verifies a CLT; then, Y^m converges in distribution towards a Gaussian vector with covariance $\mathbf{\Sigma}$, that is: *for vectors of random variables living in a fixed chaos, componentwise convergence to Gaussian, systematically implies joint convergence*. As explained e.g. in [41, Chapter 6], such a phenomenon serves as a key stepping stone in order to deduce Gaussian approximations for general functionals of Gaussian fields. Since then, this result has been extended (at least, partially), to the framework of the homogeneous chaos associated with a Poisson measure (see [4, 52]), to general vectors of homogeneous sums (see [43, Section 7] and [42]), to the free probability setting (see [45]), as well as to the framework of Markov chaoses (see [6]).

The achievements of the present paper are twofolds:

(1) On the one hand, we will obtain a general quantitative version of the one-dimensional de Jong CLT, displaying explicit bounds on the 1-Wasserstein distance. As anticipated, we will do that in the full general setting of degenerate U-statistics that do not necessarily have the form of homogeneous sums, and that are not necessarily symmetric. In particular, this extends the CLTs for homogeneous sums proved in [43], as well as the results for Poissonized and symmetric U-statistics proved in [17, 34].

(2) On the other hand, we will deduce (quantitative) multidimensional versions of de Jong theorems, showing that the crucial phenomenon observed in [51] (see the discussion at Point (b) above) basically extends to the framework of degenerate U-statistics. Our main theorems on the matter show that the case of U-statistics of the same order must take into account at least one cumulant of order four — thus echoing recent results from [6]. Our forthcoming Theorem 1.7 marks the first instance in which the phenomenon observed in [51] is described in full generality, outside the frameworks of homogeneous chaoses, and of the chaoses associated with a diffusive Markov semigroup.

We will now describe our setting and our main results in more detail.

1.2 Main results, I: univariate normal approximations

Let us fix the following setup and notation, which we essentially adopt from [12]. We refer the reader to the classical references [25, 31, 29, 62, 64], as well as to the more recent works [18, 19, 33, 47], for an introduction to degenerate U-statistics, Hoeffding decompositions and their use in stochastic analysis.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and for an integer $n \geq 1$ let $X_1, \ldots, X_n$ be independent random elements on this space assuming values in the respective measurable spaces $(E_1, \mathcal{E}_1), \ldots, (E_n, \mathcal{E}_n)$. Further, assume that

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

and that

$$W := f(X_1, \ldots, X_n) \in L^4(\mathbb{P})$$

satisfies

$$\mathbb{E}[W] = 0 \quad \text{and} \quad \mathbb{E}[W^2] = 1\,. \tag{1.1}$$

We write

$$[n] := \{1, \ldots, n\}$$

and for $J \subseteq [n]$ we also write

$$\mathcal{F}_J := \sigma(X_j,\, j \in J)\,.$$

We write

$$W = \sum_{J\subseteq[n]} W_J \tag{1.2}$$

to indicate the *Hoeffding decomposition* of W. Note that this means that, for each $J \subseteq [n]$, W_J is $\mathcal{F}_J$-measurable and that

$$\mathbb{E}[W_J \,|\, \mathcal{F}_K] = 0\,, \tag{1.3}$$

whenever $J \nsubseteq K$. It is well-known that W admits a Hoeffding decomposition of the type (1.2), as long as $W \in L^1(\mathbb{P})$ and that it is almost surely unique and given by

$$W_J = \sum_{L\subseteq J} (-1)^{|J|-|L|} \mathbb{E}\big[W \,\big|\, \mathcal{F}_L\big]\,, \quad J \subseteq [n]\,. \tag{1.4}$$

We can thus write

$$W_J = f_J(X_j,\, j \in J) \tag{1.5}$$

for some measurable functions

$$f_J : \prod_{j\in J} E_j \to \mathbb{R}\,, \quad J \subseteq [n]\,.$$

Let us also define

$$\sigma_J^2 := \mathrm{Var}(W_J)\,, \quad J \subseteq [n]\,.$$

One major assumption in what follows will be that, for some fixed integer $d \in [n]$, W is a *degenerate U-statistic of order d* (or d-degenerate U-statistic), i.e. that the Hoeffding decomposition (1.2) has the form

$$W = \sum_{J\in\mathcal{D}_d} W_J\,, \tag{1.6}$$

where

$$\mathcal{D}_d := \{J \subseteq [n] \,:\, |J| = d\}$$

denotes the collection of all $\binom{n}{d}$ d-subsets of $[n]$. Equivalently, we have $W_K = 0$ whenever $K \subseteq [n]$ is such that $|K| \neq d$. Hence, we have

$$W = f(X_1, \ldots, X_n) = \sum_{J\in\mathcal{D}_d} f_J(X_j,\, j \in J)\,.$$

The next lemma lists important properties of the Hoeffding decomposition of W which will be used without further mention.

Lemma 1.1. *Let the above notation and definitions prevail. Then, one has the following properties:*

1) *Whenever $J, K_1, \ldots, K_s \in \mathcal{D}_d$ are such that*

$$J \nsubseteq \bigcup_{i=1}^{s} K_i =: K$$

and $g(W_{K_1}, \ldots, W_{K_s})$ is square-integrable, then

$$\mathbb{E}\big[W_J g\big(W_{K_1}, \ldots, W_{K_s}\big)\big] = 0\,.$$

In particular, W_J, $J \in \mathcal{D}_d$, are uncorrelated.

2) *For all $J, K \in \mathcal{D}_d$ such that $J \neq K$ we have $\mathbb{E}[W_J \,|\, \mathcal{F}_K] = 0$.*

3) *For all $J, K \in \mathcal{D}_d$ we have $\mathbb{E}[W_J] = 0$ and $\mathbb{E}[W_J W_K] = \delta_{J,K}\sigma_J^2$.*

4) $\displaystyle\sum_{J\in\mathcal{D}_d} \sigma_J^2 = 1.$

5) *Whenever $J_1, \ldots, J_r, K_1, \ldots, K_s \in \mathcal{D}_d$ are such that*

$$\Big(\bigcup_{l=1}^{r} J_l\Big) \cap \Big(\bigcup_{i=1}^{s} K_i\Big) = \emptyset\,,$$

then the families $\{W_{J_l} : l = 1, \dots, r\}$ *and* $\{W_{K_i} : i = 1, \dots, s\}$ *are independent, i.e. the summands* W_J, $J \in \mathcal{D}_d$, *are dissociated as defined in* [37].

Proof. Point 1) is a consequence of the degeneracy property (1.3) because,

$$\mathbb{E}[W_J \,|\, \mathcal{F}_K] = 0\,,$$

as $J \not\subseteq K$ and, hence, by conditioning we have

$$\mathbb{E}\big[W_J g\big(W_{K_1}, \dots, W_{K_s}\big)\big] = \mathbb{E}\big[g\big(W_{K_1}, \dots, W_{K_s}\big)\mathbb{E}[W_J \,|\, \mathcal{F}_K]\big] = 0\,.$$

Now, Point 2) follows since it is a special case of Point 1) and also Point 3) and Point 4) are immediately implied by Point 1), in view of assumption (1.1). Finally, Point 5) follows from independence as well as the disjoint block theorem. □

Let us furthermore define the quantity

$$\varrho^2 := \varrho_n^2 := \max_{1 \le i \le n} \sum_{\substack{K \in \mathcal{D}_d: \\ i \in K}} \sigma_K^2\,. \tag{1.7}$$

The next result corresponds to de Jong's celebrated (qualitative) CLT discussed in Section 1.1.

Theorem 1.2 (See [12])**.** *Fix* $d \ge 1$*, and let* $\{n_m : m \ge 1\}$ *be a sequence of integers diverging to infinity. Let* $\{W_m : m \ge 1\}$ *be a sequence of unit variance degenerate* U*-statistics of order* d*, such that each* W_m *is a function of the vector of independent variables* $(X_1^{(m)}, ..., X_{n_m}^{(m)})$*. Then, as* $m \to \infty$*, if* $\mathbb{E}[W_m^4] \to 3$ *and* $\varrho_{n_m}^2 \to 0$*, one has that* W_m *converges in distribution towards a standard Gaussian random variable.*

Note that the condition $\lim_{m\to\infty} \varrho_{n_m}^2 = 0$ guarantees that, as $m \to \infty$, the influence of each of the random variables $(X_1^{(m)}, ..., X_{n_m}^{(m)})$ on the total variance of W_{n_m} is negligible. In fact, in the case $d = 1$ it reduces to the classical *Lindeberg-Feller condition*

$$\lim_{m\to\infty} \max_{1 \le j \le n_m} \sigma_j^2 = 0\,,$$

from the Lindeberg-Feller CLT (see e.g. Theorem 5.12 in [28]). Here, we wrote σ_j^2 for $\sigma_{\{j\}}^2$.

Our first main statement provides an explicit bound in the Wasserstein distance d_{Wass} for Theorem 1.2. We recall that, given two integrable random variables X and Y, the Wasserstein distance between the distributions of X and Y is given by the quantity

$$d_{\mathrm{Wass}}(X, Y) = \sup_{h \in \mathrm{Lip}(1)} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|\,,$$

where $\mathrm{Lip}(1)$ stands for the class of 1-Lipschitz functions.

Theorem 1.3. *As before, let* $W \in L^4(\mathbb{P})$ *be a degenerate* U*-statistic of order* d *such that* (1.1) *is satisfied and let* $Z \sim N(0,1)$ *be a standard normal random variable. Then, it holds that*

$$\begin{aligned} d_{\mathrm{Wass}}(W, Z) &\le \sqrt{\frac{2}{\pi}}\Big(\mathbb{E}[W^4] - 3 + \kappa_d \varrho_n^2\Big)^{1/2} + \frac{2\sqrt{2}}{3}\Big(2\big(\mathbb{E}[W^4] - 3\big) + 3\kappa_d \varrho_n^2\Big)^{1/2} \\ &\le \Big(\sqrt{\frac{2}{\pi}} + \frac{4}{3}\Big)\sqrt{\big|\mathbb{E}[W^4] - 3\big|} + \sqrt{\kappa_d}\Big(\sqrt{\frac{2}{\pi}} + \frac{2\sqrt{2}}{\sqrt{3}}\Big)\varrho_n\,, \end{aligned}$$

where κ_d *is a finite constant which only depends on* d*.*

Recall that a degenerate U-statistic W of order d as given by (1.6) is called *symmetric*, if, additionally, the measurable spaces $(E_1,\mathcal{E}_1),\dots,(E_n,\mathcal{E}_n)$ all coincide, the random variables $X_1,\dots,X_n$ are i.i.d. and if there is a measurable kernel $g : E_1^d \to \mathbb{R}$ such that $f_J = g$ for all $J \in \mathcal{D}_d$. In this special situation, the relations

$$1 = \mathrm{Var}(W) = \sum_{J\in\mathcal{D}_d} \mathbb{E}\big[g^2(X_j, j\in J)\big] = \binom{n}{d}\mathbb{E}\big[g^2(X_1,\dots,X_d)\big] \quad \text{and}$$

$$\varrho_n^2 = \sum_{\substack{J\in\mathcal{D}_d:\\ 1\in J}} \mathbb{E}\big[g^2(X_j, j\in J)\big] = \binom{n-1}{d-1}\mathbb{E}\big[g^2(X_1,\dots,X_d)\big]$$

imply that

$$\varrho_n^2 = \frac{d}{n}\,.$$

Hence, we arrive at the following corollary of Theorem 1.3.

Corollary 1.4. *Let $W \in L^4(\mathbb{P})$ be a normalized, degenerate and symmetric U-statistic of order d and let $Z \sim N(0,1)$ be a standard normal random variable. Then,*

$$d_{\mathrm{Wass}}(W,Z) \le \Big(\sqrt{\frac{2}{\pi}} + \frac{4}{3}\Big)\sqrt{\big|\mathbb{E}[W^4]-3\big|} + \frac{\sqrt{d\kappa_d}}{\sqrt{n}}\Big(\sqrt{\frac{2}{\pi}} + \frac{2\sqrt{2}}{\sqrt{3}}\Big)\,.$$

In particular, under the assumptions of Theorem 1.2, a sequence $\{W_m : m \ge 1\}$ of degenerate and symmetric U-statistics of a fixed order d converges in distribution to $Z \sim N(0,1)$, whenever $\lim_{m\to\infty}\mathbb{E}[W_m^4] = 3$.

Remark 1.5. (a) The previous Therorem 1.3 is a complete quantitative counterpart to de Jong's Theorem 1.2. The constant κ_d appearing in the bound is given by $\kappa_d = C_d + 2d$, where C_d is a combinatorial constant defined in Equation (4.5) below.

(b) In the context of multilinear forms in independent and standardized real-valued random variables $(X_i)_{i\in\mathbb{N}}$ considered in [43], the authors had to assume that the uniform moment condition $\sup_{i\in\mathbb{N}}\mathbb{E}[X_i^4] < \infty$ is satisfied. It is easy to check that, for homogeneous sums, this condition is in fact equivalent to the *hypercontractivity condition*

$$\sup_{n\in\mathbb{N}} D_n < \infty \quad \text{where} \quad D_n := \max_{J\in\mathcal{D}_d}\frac{\mathbb{E}\big[W_J^4\big]}{\sigma_J^4}\,.$$

Interestingly, this condition was also assumed in the monograph [11] by de Jong who was only able to dispense with it in the later paper [12]. Note further that the bounds for multilinear forms in independent random variables with arbitrary distributions derived in [43] are stated in terms of three times differentiable test functions whose first three derivatives are uniformly bounded by a constant. Hence, our Theorem 1.3 is not only more general than the corresponding result from [43] as far as the class of random functionals dealt with is concerned but is also stated in terms of much less smooth test functions.

(c) It should be mentioned that the original proof of Theorem 1.2 in [12] applies a quantitative martingale CLT from [24] and, by carefully revising its proof, one would be able to derive a bound on the rate of convergence. This issue is also briefly addressed in the introduction of the monograph [11] but not pursued any further. The resulting rate, however, would be of a much worse order than the rate provided by Theorem 1.3. Roughly, the power $1/2$ appearing in our statements would have to be systematically replaced by the power $1/5$. Furthermore, as was shown in [23] by means of an example, the Berry-Esseen bound for martingales from [24] cannot in general be improved with respect to the rate of convergence. Consequently, the techniques used by de Jong are not capable of providing sharp error bounds for

his qualitative statement. Note that the phenomenon of generally sharp bounds on the rate of convergence for martingale CLTs which reduce to sub-optimal bounds in particular situations was already discovered in the paper [3]. We also stress that, unlike our work, references [11, 12] do not contain any multidimensional statements.

Finally, we would like to mention that the paper [58] also deals with bounds on the normal approximation of so-called degenerate weighted U-statistics of order $d = 2$, which have the form

$$U = \sum_{1 \le i < j \le n} w_{i,j} \psi(X_i, X_j)$$

for some vector $X = (X_1, \ldots, X_n)$ of i.i.d. random variables, some symmetric, degenerate kernel ψ and with nonnegative weights $w_{i,j}$, $1 \le i < j \le n$. Note that the class of weighted U-statistics is strictly included in our framework, since we can define the degenerate kernel $f_{\{i,j\}}$ corresponding to the subset $\{i,j\} \in \mathcal{D}_2$ by $f_{\{i,j\}} = w_{i,j}\psi$, leading to the Hoeffding components $W_{\{i,j\}} = w_{i,j}\psi(X_i, X_j)$, $1 \le i < j \le n$. This, of course, also holds for arbitrary positive integers d. Note that, in contrast to our work, the bounds given in [58] are expressed in terms of quantities which are related explicitly to the kernel ψ and to the weights $w_{i,j}$ rather than in terms of the fourth cumulant of U and, hence, cannot be immediately compared to ours.

1.3 Main results, II: multivariate normal approximations

In this subsection we state a new approximation theorem for the distribution of vectors of degenerate, non-symmetric U-statistics by a suitable multivariate normal distribution. In particular, we show that an analog of de Jong's theorem 1.2 holds in any dimension, see Theorem 1.7. Note that, in the multivariate case, even this qualitative result relating the asymptotic normality of the vector of degenerate, non-symmetric U-statistics to fourth moment conditions is completely novel.

As before, let $X_1, \ldots, X_n$ be the underlying sequence of independent random variables, let $r \in \mathbb{N}$ and for $1 \le i \le r$ let $W(i)$ be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ which is measurable with respect to $\mathcal{F}_{[n]} = \sigma(X_1, \ldots, X_n)$ and whose Hoeffding decomposition is given by

$$W(i) = \sum_{J \in \mathcal{D}_{p_i}} W_J(i)$$

for some $p_i \in \mathbb{N}$, i.e. $W(i)$ is a degenerate U-statistic of order p_i. Without loss of generality, we can assume that $p_i \le p_k$ whenever $1 \le i < k \le r$. Thus, there is an $s \in \{1, \ldots, r\}$, positive integers $r_1, \ldots, r_s$ with $1 \le r_1 < r_2 < \ldots < r_s = r$ and integers $1 \le q_1 < q_2 < \ldots < q_s$ such that

$$p_i = q_l \quad \text{for all} \quad i \in \{r_{l-1} + 1, \ldots, r_l\} \quad \text{and all} \quad l = 1, \ldots, s\,,$$

where we set $r_0 := 0$. We define

$$W := (W(1), \ldots, W(r))^T$$

and assume that each $W(i) \in L^4(\mathbb{P})$ with

$$\mathbb{E}\big[W(i)\big] = 0 \quad \text{and} \quad \mathrm{Var}\big(W(i)\big) = \mathbb{E}\big[W(i)^2\big] = \sum_{J \in \mathcal{D}_{p_i}} \mathbb{E}\big[W_J(i)^2\big] = 1\,, \quad 1 \le i \le r\,.$$

We also let

$$v_{i,k} := \mathrm{Cov}\big(W(i), W(k)\big) = \mathbb{E}\big[W(i)W(k)\big]\,, \quad 1 \le i \le k \le r\,,$$

and

$$\mathbb{V} = \mathbb{V}(W) := \mathrm{Cov}(W) = (v_{i,k})_{1 \le i,k \le r}\,.$$

Note that $v_{i,i} = 1$ for $i = 1, \dots, r$ and $|v_{i,k}| \le 1$ for $1 \le i, k \le r$, by the Cauchy-Schwarz inequality. Note also that $v_{i,k} = 0$ unless $p_i = p_k$. Hence, $\mathbb{V}$ is a block diagonal matrix. Throughout this section we denote by

$$Z = \big(Z(1), \dots, Z(r)\big)^T \sim N_r(0, \mathbb{V})$$

a centered Gaussian vector with covariance matrix $\mathbb{V}$. For $1 \le k \le r$ and $J \in \mathcal{D}_{p_k}$ we define

$$\sigma_J(k)^2 := \mathrm{Var}\big(W_J(k)\big) = \mathbb{E}\big[W_J(k)^2\big] \quad \text{and} \quad \varrho_{n,k}^2 := \max_{1 \le j \le n} \sum_{\substack{J \in \mathcal{D}_{p_k}: \\ j \in J}} \sigma_J(k)^2 \,.$$

Before stating our multivariate normal approximation theorem, we have to introduce some more notation: For a vector $x = (x_1, \dots, x_r)^T \in \mathbb{R}^r$ we denote by $\|x\|_2$ its *Euclidean norm* and for a matrix $A \in \mathbb{R}^{r \times r}$ we denote by $\|A\|_{\mathrm{op}}$ the *operator norm* induced by the Euclidean norm, i.e.,

$$\|A\|_{\mathrm{op}} := \sup\{\|Ax\|_2 \,:\, \|x\|_2 = 1\} \,.$$

More generally, for a k-multilinear form $\psi : (\mathbb{R}^r)^k \to \mathbb{R}$, $k \in \mathbb{N}$, we define the *operator norm*

$$\|\psi\|_{\mathrm{op}} := \sup\big\{|\psi(u_1, \dots, u_k)| \,:\, u_j \in \mathbb{R}^r,\, \|u_j\|_2 = 1,\, j = 1, \dots, k\big\} \,.$$

Recall that for a function $h : \mathbb{R}^r \to \mathbb{R}$, its minimum Lipschitz constant $M_1(h)$ is given by

$$M_1(h) := \sup_{x \ne y} \frac{|h(x) - h(y)|}{\|x - y\|_2} \in [0, \infty) \cup \{\infty\}.$$

If h is differentiable, then $M_1(h) = \sup_{x \in \mathbb{R}^r} \|Dh(x)\|_{\mathrm{op}}$. More generally, for $k \ge 1$ and a $(k-1)$-times differentiable function $h : \mathbb{R}^r \to \mathbb{R}$ let

$$M_k(h) := \sup_{x \ne y} \frac{\|D^{k-1}h(x) - D^{k-1}h(y)\|_{\mathrm{op}}}{\|x - y\|_2} \,,$$

viewing the $(k-1)$-th derivative $D^{k-1}h$ of h at any point x as a $(k-1)$-multilinear form. Then, if h is actually k-times differentiable, we have $M_k(h) = \sup_{x \in \mathbb{R}^r} \|D^k h(x)\|_{\mathrm{op}}$. Having in mind this identity, we define $M_0(h) := \|h\|_\infty$.

Recall that, for two matrices $A, B \in \mathbb{R}^{r \times r}$, their *Hilbert-Schmidt inner product* is defined by

$$\langle A, B \rangle_{\mathrm{H.S.}} := \mathrm{Tr}\big(AB^T\big) = \mathrm{Tr}\big(BA^T\big) = \mathrm{Tr}\big(B^T A\big) = \sum_{i,j=1}^r a_{ij} b_{ij} \,.$$

Thus, $\langle \cdot, \cdot \rangle_{\mathrm{H.S.}}$ is just the standard inner product on $\mathbb{R}^{r \times r} \cong \mathbb{R}^{r^2}$. The corresponding *Hilbert-Schmidt norm* will be denoted by $\|\cdot\|_{\mathrm{H.S.}}$. With this notion at hand, following [8] and [38], for $k = 2$ we finally define

$$\tilde{M}_2(h) := \sup_{x \in \mathbb{R}^r} \|\mathrm{Hess}\, h(x)\|_{\mathrm{H.S.}} \,,$$

where $\mathrm{Hess}\, h$ is the *Hessian matrix* corresponding to h.

Theorem 1.6. *There exist finite constants C_{q_l}, $1 \le l \le s$, only depending on q_l as well as finite constants $C_{i,k}$, $1 \le i, k \le r$, depending on i and k only through p_i and p_k such that, with the definition*

$$A := 4 \sum_{l=1}^s \frac{q_l^2}{q_1^2} \sum_{i,k = r_{l-1}+1}^{r_l} \Big(\mathbb{E}\big[W(i)^2 W(k)^2\big] - \mathbb{E}\big[Z(i)^2 Z(k)^2\big]$$

$$+ q_l \min\Big(\varrho_{n,k}^2\,,\,\varrho_{n,i}^2\Big) + q_l\varrho_{n,k}\varrho_{n,i} + C_{i,k}\max\big(\varrho_{n,i}^2,\varrho_{n,k}^2\big)\Big)$$
$$+ 2\sum_{1\le l<m\le s}\frac{(q_l+q_m)^2}{q_1^2}\sum_{i=r_{l-1}+1}^{r_l}\sum_{k=r_{m-1}+1}^{r_m}\Bigg[\Big(\mathbb{E}\big[W(i)^4\big]-1\Big)^{1/2}$$
$$\Big(\mathbb{E}\big[W(k)^4\big]-3+\big(2q_m+C_{q_m}\big)\varrho_{n,k}^2\Big)^{1/2}$$
$$+\min\Big(q_l\varrho_{n,k}^2\,,\,q_m\varrho_{n,i}^2\Big)+C_{i,k}\max\big(\varrho_{n,i}^2,\varrho_{n,k}^2\big)\Bigg]$$

and under the above assumptions, we have the following bounds:

(i) *For any $h\in C^3(\mathbb{R}^r)$ such that $\mathbb{E}\big[|h(W)|\big]<\infty$ and $\mathbb{E}\big[|h(Z)|\big]<\infty$,*

$$\big|\mathbb{E}[h(W)]-\mathbb{E}[h(Z)]\big|\le\frac{1}{4q_1}\tilde{M}_2(h)\sqrt{A}$$
$$+\frac{\sqrt{2r}}{9}M_3(h)\sum_{l=1}^{s}\frac{q_l}{q_1}\sum_{i=r_{l-1}+1}^{r_l}\Big(2\big(\mathbb{E}\big[W(i)^4\big]-3\big)+3\big(C_{q_l}+2q_l\big)\varrho_{n,i}^2\Big)^{1/2}.$$

(ii) *If $\mathbb{V}$ is in addition positive definite, then for each $h\in C^2(\mathbb{R}^r)$ such that $\mathbb{E}\big[|h(W)|\big]<\infty$ and $\mathbb{E}\big[|h(Z)|\big]<\infty$,*

$$\big|\mathbb{E}[h(W)]-\mathbb{E}[h(Z)]\big|\le\frac{1}{\sqrt{2\pi}q_1}M_1(h)\|\mathbb{V}^{-1/2}\|_{\mathrm{op}}\sqrt{A}$$
$$+\frac{\sqrt{\pi r}}{6}M_2(h)\|\mathbb{V}^{-1/2}\|_{\mathrm{op}}\sum_{l=1}^{s}\frac{q_l}{q_1}\sum_{i=r_{l-1}+1}^{r_l}\Big(2\big(\mathbb{E}\big[W(i)^4\big]-3\big)+3\big(C_{q_l}+2q_l\big)\varrho_{n,i}^2\Big)^{1/2}.$$

Fix $r\in\mathbb{N}$. Since the class of all compactly supported, three times differentiable functions h on $\mathbb{R}^r$ is convergence-determining, from Theorem 1.6 (i) we obtain the following statement, which is a new multidimensional extension of Theorem 1.2.

Theorem 1.7. *Fix $r\ge 2$, as well as integers $p_1,...,p_r$, and let $n_m\to\infty$, as $m\to\infty$. Let $W_m:=(W_m(1),\dots,W_m(r))^T$, $m\ge 1$, be a sequence of random vectors such that each $W_m(k)$ is a centered, unit variance degenerate U-statistic of order p_k, whose argument is the vector of independent random elements $(X_1^{(m)},...,X_{n_m}^{(m)})$. Furthermore, let $\mathbf{\Sigma}\in\mathbb{R}^{r\times r}$ be a positive semi-definite matrix with $\mathbf{\Sigma}(j,j)=1$ for $j=1,\dots,r$ and denote by $N=(N(1),...,N(r))^T\sim N_r(0,\mathbf{\Sigma})$ a centered Gaussian vector with covariance matrix $\mathbf{\Sigma}$. Assume the following:*

(i) *The covariance matrix of W_m converges to $\mathbf{\Sigma}$;*
(ii) *As $m\to\infty$, $\varrho_{n_m,k}^2\to 0$, for every $k=1,...,r$;*
(iii) *As $m\to\infty$, $\mathbb{E}[W_m(k)^4]\to 3$, for every $k=1,...,r$;*
(iv) *If $j\neq k$ but $p_j=p_k$ then, as $m\to\infty$, $\mathbb{E}[W_m(j)^2W_m(k)^2]\to\mathbb{E}[N(j)^2N(k)^2]=1+(\mathbf{\Sigma}(j,k))^2$.*

Then, as $m\to\infty$, W_m converges in distribution to N.

In the framework of the normal approximation of vectors of eigenfunctions of diffusive Markov semigroups, a condition similar to (iv) in the above statement has been recently introduced and applied in [6]. The rest of the paper is organized as follows: Section 2 contains the proof of our one-dimensional result, Section 3 focusses on our multidimensional statements, whereas Section 4 contains the detailed proofs of several technical lemmas.

2 Proof of the one-dimensional theorem

In this section we give a detailed proof of Theorem 1.3. First we review Stein's method of exchangeable pairs for univariate normal approximation.

2.1 Stein's method of exchangeable pairs

The exchangeable pairs approach within Stein's method dates back to Stein's celebrated monograph [63]. Recall that a pair (X, X') of random elements on a common probability space is called *exchangeable*, if

$$(X, X') \stackrel{\mathcal{D}}{=} (X', X)\,.$$

In [63] C. Stein extensively illustrated the fact that a given normalized random variable W is close in distribution to $Z \sim N(0,1)$, whenever one can construct another random variable W' on the same space such that: (i) W' is 'close' to W in some proper, quantifiable sense, (ii) the pair (W, W') is exchangeable, (iii) the *linear regression property*

$$\mathbb{E}\big[W' - W \,\big|\, W\big] = -\lambda W \tag{2.1}$$

is satisfied for some small $\lambda > 0$, and (iv) the conditional second moment of $W' - W$ given W is close to its mean, the constant 2λ, in the L^1 metric. For a precise statement see Theorem 2.1 below.

The range of examples to which this method can be applied was considerably extended by the work [58] by Rinott and Rotar, who proved bounds on the distance to normality under the condition that the linear regression property is only approximately satisfied, i.e. that there is some negligible remainder term R such that

$$\frac{1}{\lambda}\mathbb{E}\big[W' - W \,\big|\, \mathcal{G}\big] = -W + R \tag{2.2}$$

is satisfied, where $\mathcal{G}$ is a sub-σ-field of $\mathcal{F}$ such that $\sigma(W) \subseteq \mathcal{G}$. The method of exchangeable pairs has been generalized to other absolutely continuous distributions, like the exponential (7] and [21]), the multivariate normal (8], [56] and [38]) and the Beta distribution [14]. It has also been developed for general classes of one-dimensional absolutely continuous distributions in [9], [16] and [14]. As was observed in [59], in the case of one-dimensional distributional approximation one may in general relax the exchangeability condition to the assumption that W and W' be identically distributed.

In this article we focus on the exchangeable pairs method in the context of one- and multidimensional normal approximation. The following result is a variant of Theorem 1, Lecture 3 in [63] (see also Theorem 4.9 in [10]). It slightly improves on these result with respect to the constants appearing in the bound and is also stated in terms of identically distributed random variables W, W' as opposed to exchangeable ones as well as for general sub-σ-fields $\mathcal{G}$ of $\mathcal{F}$ with $\sigma(W) \subseteq \mathcal{G}$. The proof is standard and therefore omitted from the paper. Moreover, the result is a direct consequence of Proposition 3.19 in [14] together with the best known bounds on the first two derivatives of the solution to the standard normal Stein equation for Lipschitz test functions (see e.g. Lemma 2.4 in [10]).

Theorem 2.1. *Let (W, W') be a pair of identically distributed, square-integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that, for some $\lambda > 0$,* (2.1) *holds. Furthermore, let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$ with $\sigma(W) \subseteq \mathcal{G}$. Then, we have the bound*

$$d_{\mathrm{Wass}}(W, Z) \leq \sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}\Big(\frac{1}{2\lambda}\mathbb{E}\big[(W' - W)^2 \,\big|\, \mathcal{G}\big]\Big)} + \frac{1}{3\lambda}\mathbb{E}\big|W' - W\big|^3\,. \tag{2.3}$$

For the proof of Theorem 1.3 we will need the following new auxiliary result about exchangeable pairs satisfying identity (2.1) which might be of independent interest.

Lemma 2.2. *Let (W, W') be an exchangeable pair of real-valued random variables in $L^4(\mathbb{P})$ such that, for some $\lambda > 0$,* (2.1) *is satisfied and let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$ with $\sigma(W) \subseteq \mathcal{G}$. Then,*

$$\frac{1}{4\lambda}\mathbb{E}\big[(W'-W)^4\big] = 3\mathbb{E}\Big[W^2\frac{1}{2\lambda}\mathbb{E}\big[(W'-W)^2 \,\big|\, \mathcal{G}\big]\Big] - \mathbb{E}\big[W^4\big]\,.$$

Proof. By exchangeability of (W, W') we have

$$\begin{aligned}\frac{1}{2}\mathbb{E}\big[(W'-W)^4\big] &= \mathbb{E}\big[W(W-W')^3\big] = \mathbb{E}\Big[W^4 - 3W^3W' + 3(W'W)^2 - W(W')^3\Big] \\ &= \mathbb{E}[W^4] + 3\mathbb{E}\big[(WW')^2\big] - 4\mathbb{E}\big[W^3W'\big]\,. \end{aligned} \tag{2.4}$$

Also, by (2.2)

$$\mathbb{E}\big[W^3W'\big] = \mathbb{E}\Big[W^3\mathbb{E}\big[W' \,\big|\, \mathcal{G}\big]\Big] = (1-\lambda)\mathbb{E}[W^4] \tag{2.5}$$

and

$$\begin{aligned}\mathbb{E}\big[(WW')^2\big] &= \Big[W^2\mathbb{E}\big[(W'-W+W)^2 \,\big|\, \mathcal{G}\big]\Big] \\ &= \mathbb{E}\Big[W^2\mathbb{E}\big[(W'-W)^2 + 2W(W'-W) + W^2 \,\big|\, \mathcal{G}\big]\Big] \\ &= \mathbb{E}[W^4] - 2\lambda\mathbb{E}[W^4] + \mathbb{E}\Big[W^2\mathbb{E}[(W'-W)^2 \,\big|\, \mathcal{G}]\Big] \\ &= (1-2\lambda)\mathbb{E}[W^4] + \mathbb{E}\Big[W^2\mathbb{E}[(W'-W)^2 \,\big|\, \mathcal{G}]\Big]\,. \end{aligned} \tag{2.6}$$

Thus, from (2.4), (2.5) and (2.6) we obtain that

$$\begin{aligned}\frac{1}{2}\mathbb{E}\big[(W'-W)^4\big] &= \Big(1 + 3(1-2\lambda) - 4(1-\lambda)\Big)\mathbb{E}[W^4] + 3\mathbb{E}\Big[W^2\mathbb{E}[(W'-W)^2 \,\big|\, \mathcal{G}]\Big] \\ &= 3\mathbb{E}\Big[W^2\mathbb{E}[(W'-W)^2 \,\big|\, \mathcal{G}\big] - 2\lambda\mathbb{E}[W^4]\,,\end{aligned}$$

proving the lemma. □

2.2 Proof of Theorem 1.3

Let $W \in L^4(\mathbb{P})$ be as in Theorem 1.3 such that its Hoeffding decomposition is given by (1.6). We are going to apply Theorem 2.1 to the σ-field $\mathcal{G} = \sigma(X_1, \ldots, X_n)$ and to the exchangeable pair (W, W') which is constructed as follows: Let $Y := (Y_j)_{1\le j\le n}$ be an independent copy of $X := (X_j)_{1\le j\le n}$ and let α be uniformly distributed on $\{1, \ldots, n\}$ such that X, Y and α are jointly independent. Letting, for $j = 1, \ldots, n$,

$$X_j' := \begin{cases} Y_j\,, & \text{if } \alpha = j \\ X_j\,, & \text{if } \alpha \neq j \end{cases}$$

and

$$X' := (X_1', \ldots, X_n')$$

it is easy to see that the pair (X, X') is exchangeable. Finally, as exchangeability is preserved under functions, defining

$$\begin{aligned} W' := f(X_1', \ldots, X_n') &= \sum_{j=1}^n 1_{\{\alpha=j\}}\Bigg(\sum_{\substack{J\in\mathcal{D}_d:\\ j\notin J}} W_J + \sum_{\substack{J\in\mathcal{D}_d:\\ j\in J}} W_J^{(j)}\Bigg) \\ &=: \sum_{\substack{J\in\mathcal{D}_d:\\ \alpha\notin J}} W_J + \sum_{\substack{J\in\mathcal{D}_d:\\ \alpha\in J}} W_J^{(\alpha)}\,, \end{aligned}$$

also the pair (W, W') is exchangeable. Here, for $J = \{j_1, \ldots, j_d\} \in \mathcal{D}_d$ with $1 \le j_1 < j_2 < \ldots < j_d \le n$ and $j = j_k \in J$, we write

$$W_J^{(j)} := f_J(X_{j_1}, \dots, X_{j_{k-1}}, Y_{j_k}, X_{j_{k+1}}, \dots, X_{j_d}),$$

where the kernel f_J is given by (1.5). We now show that the pair (W, W') satisfies Stein's linear regression property (2.1) exactly with coefficient $\lambda = d/n$.

Lemma 2.3. *With the above definitions, we have*

$$\mathbb{E}\big[W' - W \,\big|\, W\big] = \mathbb{E}\big[W' - W \,\big|\, X\big] = -\frac{d}{n} W\,.$$

Proof. It suffices to prove the second equality. Note that

$$W' - W = \sum_{j=1}^{n} 1_{\{\alpha = j\}} \sum_{\substack{J \in \mathcal{D}_d: \\ j \in J}} \Big(W_J^{(j)} - W_J\Big) = \sum_{\substack{J \in \mathcal{D}_d: \\ \alpha \in J}} \Big(W_J^{(\alpha)} - W_J\Big)\,.$$

Hence, by independence,

$$\begin{aligned}
\mathbb{E}\big[W' - W \,\big|\, X\big] &= \frac{1}{n} \sum_{j=1}^{n} \sum_{J: j \in J} \Big(\mathbb{E}\big[W_J^{(j)} \,\big|\, X\big] - W_J\Big) \\
&= \frac{1}{n} \sum_{j=1}^{n} \sum_{J: j \in J} \Big(\mathbb{E}\big[W_J^{(j)} \,\big|\, X_i,\, i \in J \setminus \{j\}\big] - W_J\Big) \\
&= \frac{1}{n} \sum_{j=1}^{n} \sum_{J: j \in J} \Big(\mathbb{E}\big[W_J \,\big|\, \mathcal{F}_{J \setminus \{j\}}\big] - W_J\Big) \\
&= -\frac{1}{n} \sum_{j=1}^{n} \sum_{J: j \in J} W_J = -\frac{1}{n} \sum_{J \in \mathcal{D}_d} W_J \sum_{j \in J} 1 \\
&= -\frac{d}{n} \sum_{J \in \mathcal{D}_d} W_J = -\frac{d}{n} W\,.
\end{aligned}$$

Here, we have used the defining property of the Hoeffding decomposition to obtain the fourth equality. □

We would like to mention that the same construction of the exchangeable pair (W, W') was used in [58] in the situation of weighted U-statistics. They also noted the validity of (2.1) with $\lambda = d/n$ in the special case of completely degenerate weighted U-statistics of order d.

In order to apply (2.3), by Lemma 2.3, we thus have to compute an upper bound on the variance of $\frac{n}{2d}\mathbb{E}\big[(W' - W)^2 \,\big|\, X\big]$. This is done by finding the Hoeffding decomposition of this quantity in terms of the Hoeffding decomposition of W^2 for which we will now find a new convenient expression. More generally, we derive a formula for the Hoeffding decomposition of the product of two degenerate U-statistics, which will also be needed for the proof of Theorem 1.6.

Assume that $1 \le p, q \le n$ and that W and V are square-integrable p- and q-degenerate U-statistics with respect to the same underlying sequence X, respectively, with Hoeffding decompositions

$$W = \sum_{J \in \mathcal{D}_p} W_J \quad \text{and} \quad V = \sum_{K \in \mathcal{D}_q} V_K\,. \tag{2.7}$$

The product $U := VW$ in general is not a degenerate U-statistic, but it clearly has a Hoeffding decomposition of the form

$$U = \sum_{M \subseteq [n]} U_M = \sum_{\substack{M \subseteq [n]: \\ |M| \le p+q}} U_M\,. \tag{2.8}$$

The following simple observation will be crucial for the computation of the Hoeffding decompositions of both VW and of the quantity $\frac{n}{2d}\mathbb{E}\big[(W' - W)^2 \,\big|\, X\big]$.

Lemma 2.4. *If $L \subseteq [n]$ is such that $J\Delta K = (J \setminus K) \cup (K \setminus J) \not\subseteq L$, then*

$$\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big] = 0\,.$$

Proof. Assume e.g. that $(J \setminus K) \setminus L = J \setminus (K \cup L) \neq \emptyset$. Then,

$$\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big] = \mathbb{E}\Big[V_K \mathbb{E}\big[W_J \,\big|\, \mathcal{F}_{K\cup L}\big] \,\Big|\, \mathcal{F}_L\Big] = \mathbb{E}\big[V_K \cdot 0 \,\big|\, \mathcal{F}_L\big] = 0\,,$$

as $\mathbb{E}\big[W_J \,\big|\, \mathcal{F}_{K\cup L}\big] = 0$ because $J \not\subseteq K \cup L$. □

Lemma 2.5. *Let $J \in \mathcal{D}_p$ and $K \in \mathcal{D}_q$, respectively.*

(a) *The Hoeffding decomposition of $W_J V_K$ is given by*

$$W_J V_K = \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big]\,. \tag{2.9}$$

(b) *If $j \in J \cap K$, then we have the Hoeffding decomposition*

$$\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_{(J\cup K)\setminus\{j\}}\big] = \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq (J\cup K)\setminus\{j\}}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big]\,. \tag{2.10}$$

Proof. The claim of (a) follows immediately from Lemma 2.4 and from the general formula for the Hoeffding decomposition of an $\mathcal{F}_{J\cup K}$-measurable random variable T which is given by

$$T = \sum_{M\subseteq J\cup K} \Big(\sum_{L\subseteq M} (-1)^{|M|-|L|}\mathbb{E}[T|\mathcal{F}_L]\Big)\,.$$

The claim of (b) follows similarly upon observing that, for $L \subseteq (J \cup K) \setminus \{j\}$ we have

$$\mathbb{E}\Big[\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_{(J\cup K)\setminus\{j\}}\big] \,\Big|\, \mathcal{F}_L\Big] = \mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big]\,.$$ □

The next result which might be of independent interest plays a similar role as the product formula for two multiple Wiener-Itô integrals (see e.g. [41]).

Theorem 2.6 (Product formula for degenerate U-statistics)**.** *Let $1 \leq p, q \leq n$ and let $W, V \in L^2(\mathbb{P})$ be p- and q-degenerate U-statistics, respectively, with respective Hoeffding decompositions given by* (2.7)*. Then, the Hoeffding decomposition* (2.8) *of $U := VW$ is given by the following formula:*

$$\begin{aligned} VW &= \sum_{\substack{M\subseteq[n]:\\ |M|\leq p+q}} \Bigg(\sum_{\substack{J\in\mathcal{D}_p, K\in\mathcal{D}_q:\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big]\Bigg) \\ &= \sum_{\substack{M\subseteq[n]:\\ |M|\leq p+q}} \Bigg(\sum_{L\subseteq M} (-1)^{|M|-|L|} \sum_{\substack{J\in\mathcal{D}_p, K\in\mathcal{D}_q:\\ J\Delta K\subseteq L,\\ M\subseteq J\cup K}} \mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big]\Bigg)\,, \end{aligned}$$

i.e. for $M \subseteq [n]$ with $|M| \leq p + q$ we have

$$\begin{aligned} U_M &= \sum_{\substack{J\in\mathcal{D}_p, K\in\mathcal{D}_q:\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big] \\ &= \sum_{L\subseteq M} (-1)^{|M|-|L|} \sum_{\substack{J\in\mathcal{D}_p, K\in\mathcal{D}_q:\\ J\Delta K\subseteq L,\\ M\subseteq J\cup K}} \mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big]\,. \end{aligned}$$

Proof. By the linearity of the Hoeffding decomposition and since we have

$$VW = \sum_{J\in\mathcal{D}_p, K\in\mathcal{D}_q} W_J V_K \,,$$

it suffices to collect the terms resulting from the Hoeffding decompositions of the summands $W_J V_K$ in a suitable way. By Lemma 2.5 (a) we have

$$\begin{aligned} VW &= \sum_{J\in\mathcal{D}_p, K\in\mathcal{D}_q} W_J V_K = \sum_{J\in\mathcal{D}_p, K\in\mathcal{D}_q} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big] \\ &= \sum_{\substack{M\subseteq[n]:\\ |M|\le p+q}} \Bigg(\sum_{L\subseteq M} (-1)^{|M|-|L|} \sum_{\substack{J\in\mathcal{D}_p, K\in\mathcal{D}_q:\\ J\Delta K\subseteq L,\\ M\subseteq J\cup K}} \mathbb{E}\big[W_J V_K \,\big|\, \mathcal{F}_L\big] \Bigg) . \end{aligned}$$

□

Now we are in the position to express the Hoeffding decomposition of $\frac{n}{2d}\mathbb{E}\big[(W'-W)^2 \,\big|\, X\big]$ in terms of that of W^2. Since we prove a more general result, Lemma 3.3 below, we do not give its proof, here.

Lemma 2.7. *Let $W^2 = \sum_{|M|\le 2d} U_M$ be the Hoeffding decomposition of W^2. Then, we have the Hoeffding decomposition*

$$\frac{n}{2d}\mathbb{E}\big[(W'-W)^2 \,\big|\, X\big] = \sum_{\substack{M\subseteq[n]:\\ |M|\le 2d-1}} a_M U_M \,,$$

with

$$a_M = 1 - \frac{|M|}{2d} \in [0,1] \quad \textit{for each} \quad M\subseteq[n] \textit{ with } |M|\le 2d\,.$$

Before we proceed, let us, following [11] and [12], introduce the following important classes of quadruples $(J_1,J_2,J_3,J_4)\in\mathcal{D}_d^4$. We call an element $j\in J_1\cup J_2\cup J_3\cup J_4$ a *free index*, if it appears in J_i for exactly one $i\in\{1,2,3,4\}$. Note that this implies that

$$\mathbb{E}\big[W_{J_1}W_{J_2}W_{J_3}W_{J_4}\big] = 0 \tag{2.11}$$

by Lemma 1.1 4). We say that (J_1,J_2,J_3,J_4) is *bifold*, if each element in the union $J_1\cup J_2\cup J_3\cup J_4$ appears in J_i for exactly two values of $i\in\{1,2,3,4\}$, i.e. if

$$1_{J_1}+1_{J_2}+1_{J_3}+1_{J_4} = 2\cdot 1_{J_1\cup J_2\cup J_3\cup J_4}\,.$$

Let us denote by $\mathcal{B}=\mathcal{B}_d$ the set of all bifold quadruples. Among the bifold quadruples, the most important ones are given by the subclass $\mathcal{S}_0$ which is defined by

$$\begin{aligned} \mathcal{S}_0 = \Big\{ (J,K,L,M)\in\mathcal{D}_d^4 \,:\, J\cap K = L\cap M = \emptyset\,, \quad \emptyset\subsetneq J\cap L = J\setminus(J\cap M)\subsetneq J \\ \text{and } \emptyset\subsetneq K\cap L = K\setminus(K\cap M)\subsetneq K \Big\}\,. \end{aligned}$$

Further, we denote by $\mathcal{T}=\mathcal{T}_d$ the set of all quadruples $(J_1,J_2,J_3,J_4)\in\mathcal{D}_d^4$ that are neither bifold nor have a free index. This just means that

$$1_{J_1}+1_{J_2}+1_{J_3}+1_{J_4} \ge 2\cdot 1_{J_1\cup J_2\cup J_3\cup J_4}$$

and there exists at least one $j\in[n]$ such that

$$1_{J_1}(j)+1_{J_2}(j)+1_{J_3}(j)+1_{J_4}(j)\ge 3\,,$$

i.e. each element of the union $J_1\cup J_2\cup J_3\cup J_4$ appears in J_i for at least two values of $i\in\{1,2,3,4\}$ and there is an element of the union $J_1\cup J_2\cup J_3\cup J_4$ that appears in J_i for at least three values of $i\in\{1,2,3,4\}$.

Following [12] let us define the quantities

$$S_0 := \sum_{\substack{J,K,L,M\in\mathcal{D}_d:\\ J\cap K=\emptyset=L\cap M,\\ \emptyset\subsetneq J\cap L\subsetneq J,\\ \emptyset\subsetneq J\cap M\subsetneq J}} \mathbb{E}\big[W_J W_K W_L W_M\big] = \sum_{(J,K,L,M)\in\mathcal{S}_0} \mathbb{E}\big[W_J W_K W_L W_M\big]\,,$$

as well as

$$\tau := \tau_d := \sum_{(J,K,L,M)\in\mathcal{T}} \sigma_J\sigma_K\sigma_L\sigma_M\,.$$

Note that the last identity in the definition of S_0 is true by virtue of (2.11). The following result is Proposition 5 (b) of [12]. We will prove a more general version stated as Proposition 3.5 to deal with the multivariate case.

Proposition 2.8. *We have*

$$S_0 \geq -\tau\,.$$

Recall the definition of the Lindeberg-Feller quantity $\varrho = \varrho_n$ given in (1.7). Next, we state a substantial improvement of Lemma B in [12]. Indeed, there the upper bound on τ is of order ϱ as compared to the order ϱ^2 which we obtain. Its proof is deferred to Section 4.

Proposition 2.9. *For each $d\in\mathbb{N}$ there is a finite constant C_d which is independent of n such that*

$$\tau = \sum_{(J,K,L,M)\in\mathcal{T}} \sigma_J\sigma_K\sigma_L\sigma_M \leq C_d\varrho^2\,.$$

Furthermore, we can let $C_2 = 13$.

The next two lemmas will be very useful for what follows.

Lemma 2.10. *Again, let $W^2 = \sum_{|M|\leq 2d} U_M$ denote the Hoeffding decomposition of W^2. Then, we have the bound*

$$\sum_{\substack{M\subseteq[n]:\\ |M|\leq 2d-1}} \mathrm{Var}(U_M) \leq \mathbb{E}\big[W^4\big] - 3 + \kappa_d\varrho^2\,,$$

where $\kappa_d = C_d + 2d$ and C_d is the constant from Proposition 2.9.

Proof. We have

$$\begin{aligned}
\sum_{\substack{M\subseteq[n]:\\ |M|\leq 2d-1}} \mathrm{Var}(U_M) &= \mathrm{Var}(W^2) - \sum_{\substack{M\subseteq[n]:\\ |M|=2d}} \mathrm{Var}(U_M)\\
&= \mathbb{E}\big[W^4\big] - 1 - \sum_{\substack{M\subseteq[n]:\\ |M|=2d}} \mathrm{Var}(U_M)\\
&= \mathbb{E}\big[W^4\big] - 3 + 2 - \sum_{\substack{M\subseteq[n]:\\ |M|=2d}} \mathbb{E}[U_M^2]\\
&= \mathbb{E}\big[W^4\big] - 3 + \Big(2 - \sum_{\substack{J,K,L,M\in\mathcal{D}_d:\\ J\cap K=\emptyset=L\cap M}} \mathbb{E}\big[W_J W_K W_L W_M\big]\Big)\,. \qquad (2.12)
\end{aligned}$$

For the last equality we have used the fact that for $|M| = 2d$ we have

$$U_M = \sum_{\substack{J,K\in\mathcal{D}_d:\\ J\cap K=\emptyset,\\ J\cup K=M}} W_J W_K\,.$$

Also, we can write

$$\begin{aligned}\sum_{\substack{J,K,L,M\in\mathcal{D}_d:\\ J\cap K=\emptyset=L\cap M}}\mathbb{E}\big[W_JW_KW_LW_M\big] &= 2\sum_{\substack{J,K\in\mathcal{D}_d:\\ J\cap K=\emptyset}}\mathbb{E}\big[W_J^2W_K^2\big]+S_0\\ &= 2\sum_{\substack{J,K\in\mathcal{D}_d:\\ J\cap K=\emptyset}}\sigma_J^2\sigma_K^2+S_0\\ &= 2-2\sum_{\substack{J,K\in\mathcal{D}_d:\\ J\cap K\neq\emptyset}}\sigma_J^2\sigma_K^2+S_0\\ &\geq 2-2d\varrho^2+S_0\,,\end{aligned}$$

where we have used Lemma 4.1 to obtain the last inequality. Thus, from (2.12) and Propositions 2.8 and 2.9 we conlude that

$$\begin{aligned}\sum_{\substack{M\subseteq[n]:\\ |M|\leq 2d-1}}\mathrm{Var}(U_M) &= \mathrm{Var}(W^2)-\sum_{\substack{M\subseteq[n]:\\ |M|=2d}}\mathrm{Var}(U_M)\\ &\leq \mathbb{E}\big[W^4\big]-3+2d\varrho^2-S_0\\ &\leq \mathbb{E}\big[W^4\big]-3+(2d+C_d)\varrho^2\\ &= \mathbb{E}\big[W^4\big]-3+\kappa_d\varrho^2\,,\end{aligned}\tag{2.13}$$

which proves the claim. □

Now we are able to bound the first term on the right hand side of (2.3):

Lemma 2.11. *For the above constructed exchangeable pair we have*

$$\mathrm{Var}\Big(\frac{n}{2d}\mathbb{E}\big[(W'-W)^2\,\big|\,X\big]\Big)\leq\mathbb{E}\big[W^4\big]-3+\kappa_d\varrho^2\,.\tag{2.14}$$

Proof. Using the orthogonality of the summands within the Hoeffding decomposition as well as $a_M\in[0,1]$, $|M|\leq 2d-1$, from Lemma 2.7 we obtain that

$$\begin{aligned}\mathrm{Var}\Big(\frac{n}{2d}\mathbb{E}\big[(W'-W)^2\,\big|\,X\big]\Big) = \mathrm{Var}\Big(\sum_{\substack{M\subseteq[n]:\\ |M|\leq 2d-1}}a_MU_M\Big) &= \sum_{\substack{M\subseteq[n]:\\ |M|\leq 2d-1}}a_M^2\,\mathrm{Var}(U_M)\\ &\leq \sum_{\substack{M\subseteq[n]:\\ |M|\leq 2d-1}}\mathrm{Var}(U_M)\\ &\leq \mathbb{E}\big[W^4\big]-3+\kappa_d\varrho^2\,,\end{aligned}$$

where the final inequality is by Lemma 2.10. □

Now, we proceed to bounding the second error term appearing in the bound (2.3) from Theorem 2.1. The next lemma will be crucial for doing this.

Lemma 2.12. *For the above constructed exchangeable pair we have the bound*

$$\frac{n}{4d}\mathbb{E}\big[(W'-W)^4\big]\leq 2\big(\mathbb{E}[W^4]-3\big)+3\kappa_d\varrho^2\,.$$

Proof. From Lemmas 2.3, 2.2 and 2.7 we obtain that

$$\begin{aligned}\frac{n}{4d}\mathbb{E}\big[(W'-W)^4\big] &= 3\mathbb{E}\Big[W^2\frac{n}{2d}\mathbb{E}\big[(W'-W)^2\,\big|\,X\big]\Big]-\mathbb{E}\big[W^4\big]\\ &= 3\sum_{\substack{M,N\subseteq[n]:\\ |M|,|N|\leq 2d}}a_M\mathbb{E}\big[U_MU_N\big]-\mathbb{E}\big[W^4\big]\,,\end{aligned}$$

where we recall that $a_M = 1 - \frac{|M|}{2d} \in [0,1]$, for all $M \subseteq [n]$ such that $|M| \leq 2d$. Noting that $a_\emptyset U_\emptyset^2 = 1$, $a_M = 0$ whenever $|M| = 2d$ and using the orthogonality of the Hoeffding decomposition yield

$$\begin{aligned}\frac{n}{4d}\mathbb{E}\big[(W'-W)^4\big] &= 3a_\emptyset U_\emptyset^2 - \mathbb{E}\big[W^4\big] + 3 \sum_{\substack{M\subseteq[n]:\\ 1\leq|M|\leq 2d-1}} a_M \operatorname{Var}(U_M) \\ &\leq 3 - \mathbb{E}[W^4] + 3 \sum_{\substack{M\subseteq[n]:\\ 1\leq|M|\leq 2d-1}} \operatorname{Var}(U_M) \\ &\leq 3 - \mathbb{E}[W^4] + 3\big(\mathbb{E}[W^4] - 3 + \kappa_d \varrho^2\big) \\ &= 2\big(\mathbb{E}[W^4] - 3\big) + 3\kappa_d \varrho^2\,,\end{aligned}$$

where we have used Lemma 2.10 to obtain the last inequality. □

From the fact that

$$\frac{2d}{n} = 2\lambda = \mathbb{E}\big[(W'-W)^2\big]$$

and using the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}\frac{1}{3\lambda}\mathbb{E}\big|W'-W\big|^3 &\leq \frac{n}{3d}\Big(\mathbb{E}\big[(W'-W)^2\big]\Big)^{1/2}\Big(\mathbb{E}\big|W'-W\big|^4\Big)^{1/2} \\ &= \frac{2\sqrt{2}}{3}\Big(\frac{n}{4d}\mathbb{E}\big[(W'-W)^4\big]\Big)^{1/2}. \qquad (2.15)\end{aligned}$$

Hence, by virtue of Lemma 2.12 we have

$$\frac{1}{3\lambda}\mathbb{E}\big|W'-W\big|^3 \leq \frac{2\sqrt{2}}{3}\Big(2\big(\mathbb{E}[W^4]-3\big) + 3\kappa_d\varrho^2\Big)^{1/2}. \qquad (2.16)$$

Theorem 1.3 now follows from (2.3), Lemma 2.11 and from (2.16).

3 Proof of the multidimensional theorem

3.1 Stein's method of exchangeable pairs for multivariate normal approximation

Although the exchangeable pairs coupling lies at the heart of univariate normal approximation by Stein's method, it was only in 2008 in [8] that the problem of developing an analogous technique in the multivariate setting was finally attacked. In their work, for a given random vector

$$W = (W(1), \ldots, W(r))^T,$$

the authors assume the existence of another random vector

$$W' = (W'(1), \ldots, W'(r))^T,$$

defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that W' has the same distribution as W and such that the linear regression property

$$\mathbb{E}\big[W' - W \,\big|\, W\big] = -\lambda W$$

is satisfied for some positive constant λ. Under these assumptions the authors prove several theorems which bound the distance from W to a standard normal random vector in terms of the pair (W, W').

In [56] the authors motivate and investigate the more general linear regression property

$$\mathbb{E}\big[W' - W \,\big|\, \mathcal{G}\big] = -\Lambda W + R\,, \tag{3.1}$$

where now Λ is an invertible non-random $r \times r$ matrix, $\mathcal{G} \subseteq \mathcal{F}$ is a sub-σ-field of $\mathcal{F}$ such that $\sigma(W) \subseteq \mathcal{G}$ and $R = (R(1), \ldots, R(r))^T$ is a small remainder term. However, in contrast to [8] and to the univariate situation presented in Subsection 2.1, in [56] the full strength of the exchangeability of the vector (W, W') is needed. Finally, in [38] the two approaches from [8] and [56] are combined, allowing for the more general linear regression property from [56] and using sharper coordinate-free bounds on the solution to the Stein equation similar to those derived in [8]. The following result, quoted from [13], is (a version of) Theorem 3 in [38] but with better constants.

Theorem 3.1. *Let (W, W') be an exchangeable pair of $\mathbb{R}^r$-valued $L^2(\mathbb{P})$ random vectors defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-field of $\mathcal{F}$ such that $\sigma(W) \subseteq \mathcal{G}$. Suppose there exist a non-random invertible matrix $\Lambda \in \mathbb{R}^{r\times r}$, a non-random positive semidefinite matrix Σ, a $\mathcal{G}$-measurable random vector R and a $\mathcal{G}$-measurable random matrix S such that* (3.1) *and*

$$\mathbb{E}\Big[(W' - W)(W' - W)^T \,\Big|\, \mathcal{G}\Big] = 2\Lambda\Sigma + S \tag{3.2}$$

hold true. Finally, denote by Z a centered r-dimensional Gaussian vector with covariance matrix Σ.

(a) *For any $h \in C^3(\mathbb{R}^r)$ such that $\mathbb{E}\big[|h(W)|\big] < \infty$ and $\mathbb{E}\big[|h(Z)|\big] < \infty$,*

$$\begin{aligned}\big|\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]\big| &\le \|\Lambda^{-1}\|_{\mathrm{op}}\bigg(M_1(h)\mathbb{E}\big[\|R\|_2\big] + \frac{1}{4}\tilde{M}_2(h)\mathbb{E}\big[\|S\|_{\mathrm{H.S.}}\big] \\ &\qquad + \frac{1}{18}M_3(h)\mathbb{E}\big[\|W' - W\|_2^3\big]\bigg) \\ &\le \|\Lambda^{-1}\|_{\mathrm{op}}\bigg(M_1(h)\mathbb{E}\big[\|R\|_2\big] + \frac{\sqrt{r}}{4}M_2(h)\mathbb{E}\big[\|S\|_{\mathrm{H.S.}}\big] \\ &\qquad + \frac{1}{18}M_3(h)\mathbb{E}\big[\|W' - W\|_2^3\big]\bigg)\,.\end{aligned}$$

(b) *If Σ is actually positive definite, then for each $h \in C^2(\mathbb{R}^r)$ such that $\mathbb{E}\big[|h(W)|\big] < \infty$ and $\mathbb{E}\big[|h(Z)|\big] < \infty$ we have*

$$\begin{aligned}\big|\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]\big| &\le M_1(h)\|\Lambda^{-1}\|_{\mathrm{op}}\bigg(\mathbb{E}\big[\|R\|_2\big] + \frac{\|\Sigma^{-1/2}\|_{\mathrm{op}}}{\sqrt{2\pi}}\mathbb{E}\big[\|S\|_{\mathrm{H.S.}}\big]\bigg) \\ &\qquad + \frac{\sqrt{2\pi}}{24}M_2(h)\|\Lambda^{-1}\|_{\mathrm{op}}\|\Sigma^{-1/2}\|_{\mathrm{op}}\mathbb{E}\big[\|W' - W\|_2^3\big]\,.\end{aligned}$$

3.2 Proof of Theorem 1.6

Recall the notation and assumptions from Subsection 1.3. Starting from the random vector $W = (W(1), \ldots, W(r))^T$ we will construct another vector

$$W' := (W'(1), \ldots, W'(r))^T$$

such that (W, W') is an exchangeable pair in the following way: For each $1 \le i \le r$ we construct $W'(i)$ in the same way as we did in the one-dimensional situation treated in Subsection 2.2 and from the same independent copy $Y = (Y_1, \ldots, Y_n)$ of $X = (X_1, \ldots, X_n)$ and the same α which is independent of (X, Y) and uniformly distributed on $[n]$. We will apply Theorem 3.1 with $\Sigma = \mathbb{V}$ and $\mathcal{G} = \sigma(X_1, \ldots, X_n)$.

Lemma 3.2. *With the above definitions and notation we have*

$$\mathbb{E}\big[W' - W \,\big|\, X\big] = -\Lambda W\,,$$

where the matrix Λ *is given by* $\Lambda = \operatorname{diag}\Big(\frac{p_1}{n}, \dots, \frac{p_r}{n}\Big)$.

Proof. This follows immediately from Lemma 2.3. □

Hence, we obtain that

$$\|\Lambda^{-1}\|_{\mathrm{op}} = \max_{i=1,\dots,r} \frac{n}{p_i} = \frac{n}{p_1}\,. \tag{3.3}$$

Let us define the random matrix $S = (S_{i,k})_{1\le i,k\le r}$ by the relation

$$\mathbb{E}\big[(W' - W)(W' - W)^T \,\big|\, X\big] = 2\Lambda\mathbb{V} + S\,. \tag{3.4}$$

From Lemma 3.2 and the fact that $v_{i,k} = 0$ unless $p_i = p_k$ we easily conclude that S is symmetric. Also, using exchangeability, it is readily checked that

$$\mathbb{E}\big[S\big] = \mathbb{E}\big[(W' - W)(W' - W)^T\big] - 2\Lambda\mathbb{V} = 0\,. \tag{3.5}$$

Lemma 3.3. *Let* $1 \le i \le k \le r$ *and let*

$$W(i)W(k) = \sum_{\substack{M\subseteq[n]:\\ |M|\le p_i+p_k}} U_M(i,k)$$

be the Hoeffding decomposition of $W(i)W(k)$*. Then, we have the Hoeffding decomposition*

$$n\mathbb{E}\big[(W'(i) - W(i))(W'(k) - W(k)) \,\big|\, X\big] = \sum_{\substack{M\subseteq[n]:\\ |M|\le p_i+p_k-1}} a_M(i,k)U_M(i,k)\,,$$

where

$$a_M(i,k) = p_i + p_k - |M|\,.$$

Proof. First note that we have the representation

$$\begin{aligned}
&(W'(i) - W(i))(W'(k) - W(k))\\
&= \sum_{j=1}^n 1_{\{\alpha=j\}} \sum_{\substack{J\in\mathcal{D}_{p_i}, K\in\mathcal{D}_{p_k}:\\ j\in J\cap K}} (W'_J(i) - W_J(i))(W'_K(k) - W_K(k))\\
&= \sum_{j=1}^n 1_{\{\alpha=j\}} \sum_{\substack{J\in\mathcal{D}_{p_i}, K\in\mathcal{D}_{p_k}:\\ j\in J\cap K}} \Big(W_J^{(j)}(i)W_K^{(j)}(k) + W_J(i)W_K(k) - W_J(i)W_K^{(j)}(k)\\
&\qquad\qquad - W_J^{(j)}(i)W_K(k)\Big)
\end{aligned}$$

which implies that

$$\begin{aligned}
&n\mathbb{E}\big[(W'(i) - W(i))(W'(k) - W(k)) \,\big|\, X\big]\\
&= \sum_{j=1}^n \sum_{\substack{J\in\mathcal{D}_{p_i}, K\in\mathcal{D}_{p_k}:\\ j\in J\cap K}} \Big(\mathbb{E}\big[W_J(i)W_K(k) \,\big|\, \mathcal{F}_{(J\cup K)\setminus\{j\}}\big] + W_J(i)W_K(k)\Big)\\
&= \sum_{\substack{J\in\mathcal{D}_{p_i}, K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} |J\cap K| W_J(i)W_K(k) + \sum_{j=1}^n \sum_{\substack{J\in\mathcal{D}_{p_i}, K\in\mathcal{D}_{p_k}:\\ j\in J\cap K}} \mathbb{E}\big[W_J(i)W_K(k) \,\big|\, \mathcal{F}_{(J\cup K)\setminus\{j\}}\big]
\end{aligned}$$

Using Lemma 2.5 (b) we have

$$\begin{aligned}
&\sum_{j=1}^{n} \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ j\in J\cap K}} \mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_{(J\cup K)\setminus\{j\}}\big] \\
&= \sum_{j=1}^{n} \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ j\in J\cap K}} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq (J\cup K)\setminus\{j\}}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big] \\
&= \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} \sum_{j\in J\cap K} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq (J\cup K)\setminus\{j\}}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big] \\
&= \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subsetneq J\cup K}} \big(|J\cup K|-|M|\big) \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big] \\
&= \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} \big(|J\cup K|-|M|\big) \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big]
\end{aligned} \tag{3.6}$$

Note that for the third equality we have used the crucial fact that

$$(J\cup K)\setminus M = (J\cap K)\setminus M\,, \quad \text{whenever } J\Delta K\subseteq M$$

which implies that

$$|(J\cap K)\setminus M| = |J\cup K|-|M| \quad \text{for } J\Delta K\subseteq M\subseteq J\cup K\,.$$

Also, from Lemma 2.5 (a) we obtain that

$$\begin{aligned}
&\sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} |J\cap K| \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big] \\
&= \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} |J\cap K| \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big] \\
&= \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} |J\cap K| \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big]
\end{aligned} \tag{3.7}$$

Combining (3.6) and (3.7) we thus have

$$\begin{aligned}
&n\mathbb{E}\big[(W'(i)-W(i))(W'(k)-W(k))\,\big|\,X\big] \\
&= \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} \big(|J\cup K|+|J\cap K|-|M|\big) \\
&\qquad\qquad \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big] \\
&= \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} \sum_{\substack{M\subseteq[n]:\\ J\Delta K\subseteq M\subseteq J\cup K}} \big(p_i+p_k-|M|\big) \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big] \\
&= \sum_{\substack{M\subseteq[n]:\\ |M|\le p_i+p_k-1}} \big(p_i+p_k-|M|\big) \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset,\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big]
\end{aligned}$$

$$= \sum_{\substack{M\subseteq[n]:\\ |M|\le p_i+p_k-1}} \big(p_i+p_k-|M|\big) \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\Delta K\subseteq M\subseteq J\cup K}} \sum_{\substack{L\subseteq[n]:\\ J\Delta K\subseteq L\subseteq M}} (-1)^{|M|-|L|}\mathbb{E}\big[W_J(i)W_K(k)\,\big|\,\mathcal{F}_L\big]$$

$$= \sum_{\substack{M\subseteq[n]:\\ |M|\le p_i+p_k-1}} \big(p_i+p_k-|M|\big)U_M(i,k)\,,$$

as claimed. □

Since S is centered, from (3.4) and Lemma 3.3 we obtain that

$$\begin{aligned}
n^2\mathbb{E}\big[S_{i,k}^2\big] &= \mathrm{Var}\big(nS_{i,k}\big) = \mathrm{Var}\Big(n\mathbb{E}\big[(W'(i)-W(i))(W'(k)-W(k))\,\big|\,X\big]\Big)\\
&= \sum_{\substack{M\subseteq[n]:\\ |M|\le p_i+p_k-1}} \big(p_i+p_k-|M|\big)^2\,\mathrm{Var}\big(U_M(i,k)\big)\\
&\le (p_i+p_k)^2 \sum_{\substack{M\subseteq[n]:\\ |M|\le p_i+p_k-1}} \mathrm{Var}\big(U_M(i,k)\big)\\
&= (p_i+p_k)^2\Big(\mathrm{Var}(W(i)W(k)) - \sum_{\substack{M\subseteq[n]:\\ |M|=p_i+p_k}} \mathbb{E}\big[U_M(i,k)^2\big]\Big)\\
&= (p_i+p_k)^2\Big(\mathrm{Var}(W(i)W(k))\\
&\qquad - \sum_{\substack{J,L\in\mathcal{D}_{p_i},K,M\in\mathcal{D}_{p_k}:\\ J\cap K=L\cap M=\emptyset}} \mathbb{E}\big[W_J(i)W_K(k)W_L(i)W_M(k)\big]\Big)\,. \qquad (3.8)
\end{aligned}$$

For $1\le i\le k\le r$ define

$$S_0(i,k) := \sum_{\substack{J,L\in\mathcal{D}_{p_i},K,M\in\mathcal{D}_{p_k}:\\ J\cap K=L\cap M=\emptyset,\\ \emptyset\subsetneq J\cap L=J\setminus(J\cap M)\subsetneq J,\\ \emptyset\subsetneq L\cap J=L\setminus(L\cap K)\subsetneq L}} \mathbb{E}\big[W_J(i)W_K(k)W_L(i)W_M(k)\big]\,.$$

If $p_i<p_k$, then from (3.8) we have that

$$n^2\mathbb{E}\big[S_{i,k}^2\big] \le (p_i+p_k)^2\Big(\mathrm{Var}(W(i)W(k)) - \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K=\emptyset}} \sigma_J(i)^2\sigma_K(k)^2 - S_0(i,k)\Big)\,. \qquad (3.9)$$

Lemma 4.1 immediately yields that

$$\begin{aligned}
\sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K=\emptyset}} \sigma_J(i)^2\sigma_K(k)^2 &= 1 - \sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\ne\emptyset}} \sigma_J(i)^2\sigma_K(k)^2\\
&\ge 1-\min\Big(p_i\varrho_{n,k}^2\,,\,p_k\varrho_{n,i}^2\Big)\,. \qquad (3.10)
\end{aligned}$$

If $p_i=p_k$, then we obtain that

$$\begin{aligned}
n^2\mathbb{E}\big[S_{i,k}^2\big] \le 4p_i^2\Big(&\mathrm{Var}(W(i)W(k)) - \sum_{\substack{J,K\in\mathcal{D}_{p_i}:\\ J\cap K=\emptyset}} \sigma_J(i)^2\sigma_K(k)^2\\
&- \sum_{\substack{J,K\in\mathcal{D}_{p_i}:\\ J\cap K=\emptyset}} \mathbb{E}\big[W_J(i)W_J(k)\big]\mathbb{E}\big[W_K(i)W_K(k)\big] - S_0(i,k)\Big)\,. \qquad (3.11)
\end{aligned}$$

Similarly to Lemma 4.1 we obtain for $p_i=p_k$ that

$$
\begin{aligned}
&\Big|\sum_{\substack{J,K\in\mathcal{D}_{p_i}:\\ J\cap K\neq\emptyset}}\mathbb{E}\big[W_J(i)W_J(k)\big]\mathbb{E}\big[W_K(i)W_K(k)\big]\Big|\\
&\le\sum_{J\in\mathcal{D}_{p_i}}\big|\mathbb{E}\big[W_J(i)W_J(k)\big]\big|\sum_{j\in J}\sum_{\substack{K\in\mathcal{D}_{p_i}:\\ j\in K}}\big|\mathbb{E}\big[W_K(i)W_K(k)\big]\big|\\
&\le\sum_{J\in\mathcal{D}_{p_i}}\big|\mathbb{E}\big[W_J(i)W_J(k)\big]\big|\sum_{j\in J}\Big(\sum_{\substack{K\in\mathcal{D}_{p_i}:\\ j\in K}}\sigma_K(i)^2\Big)^{1/2}\Big(\sum_{\substack{K\in\mathcal{D}_{p_k}:\\ j\in K}}\sigma_K(k)^2\Big)^{1/2}\\
&\le p_i\varrho_{n,k}\varrho_{n,i}\sum_{J\in\mathcal{D}_{p_i}}\big|\mathbb{E}\big[W_J(i)W_J(k)\big]\big|\le p_i\varrho_{n,k}\varrho_{n,i}\sum_{J\in\mathcal{D}_{p_i}}\sigma_J(i)\sigma_J(k)\\
&\le p_i\varrho_{n,k}\varrho_{n,i}\Big(\sum_{J\in\mathcal{D}_{p_i}}\sigma_J(i)^2\Big)^{1/2}\Big(\sum_{J\in\mathcal{D}_{p_k}}\sigma_J(k)^2\Big)^{1/2}\\
&=p_i\varrho_{n,k}\varrho_{n,i}\,.
\end{aligned}
$$

Hence, if $p_i=p_k$ we have that

$$
\begin{aligned}
&\sum_{\substack{J,K\in\mathcal{D}_{p_i}:\\ J\cap K=\emptyset}}\mathbb{E}\big[W_J(i)W_J(k)\big]\mathbb{E}\big[W_K(i)W_K(k)\big]\\
&=\Big(\sum_{J\in\mathcal{D}_{p_i}}\mathbb{E}\big[W_J(i)W_J(k)\big]\Big)^2-\sum_{\substack{J,K\in\mathcal{D}_{p_i}:\\ J\cap K\neq\emptyset}}\mathbb{E}\big[W_J(i)W_J(k)\big]\mathbb{E}\big[W_K(i)W_K(k)\big]\\
&=v_{i,k}^2-\sum_{\substack{J,K\in\mathcal{D}_{p_i}:\\ J\cap K\neq\emptyset}}\mathbb{E}\big[W_J(i)W_J(k)\big]\mathbb{E}\big[W_K(i)W_K(k)\big]\\
&\ge v_{i,k}^2-p_i\varrho_{n,k}\varrho_{n,i}\,. \qquad (3.12)
\end{aligned}
$$

Note that we can write

$$
\begin{aligned}
\mathrm{Var}\big(W(i)W(k)\big)&=\mathbb{E}\big[W(i)^2W(k)^2\big]-\Big(\mathbb{E}\big[W(i)W(k)\big]\Big)^2\\
&=\mathrm{Cov}\big(W(i)^2,W(k)^2\big)+\mathbb{E}\big[W(i)^2\big]\mathbb{E}\big[W(k)^2\big]-v_{i,k}^2\\
&=\mathrm{Cov}\big(W(i)^2,W(k)^2\big)+1-v_{i,k}^2\,. \qquad (3.13)
\end{aligned}
$$

Hence, if $p_i<p_k$, then, since $v_{i,k}=0$, from (3.9), (3.10) and (3.13) we see that

$$
n^2\mathbb{E}\big[S_{i,k}^2\big]\le(p_i+p_k)^2\Big(\mathrm{Cov}\big(W(i)^2,W(k)^2\big)+\min\Big(p_i\varrho_{n,k}^2\,,\,p_k\varrho_{n,i}^2\Big)-S_0(i,k)\Big)\,. \qquad (3.14)
$$

If, on the other hand, $p_i=p_k$, then from (3.11), (3.10), (3.13) and (3.12) we conclude that

$$
\begin{aligned}
n^2\mathbb{E}\big[S_{i,k}^2\big]&\le 4p_i^2\Big(\mathrm{Cov}\big(W(i)^2,W(k)^2\big)-2v_{i,k}^2+p_i\min\Big(\varrho_{n,k}^2\,,\,\varrho_{n,i}^2\Big)\\
&\qquad+p_i\varrho_{n,k}\varrho_{n,i}-S_0(i,k)\Big)\\
&=4p_i^2\Big(\mathbb{E}\big[W(i)^2W(k)^2\big]-1-2v_{i,k}^2+p_i\min\Big(\varrho_{n,k}^2\,,\,\varrho_{n,i}^2\Big)\\
&\qquad+p_i\varrho_{n,k}\varrho_{n,i}-S_0(i,k)\Big)\\
&=4p_i^2\Big(\mathbb{E}\big[W(i)^2W(k)^2\big]-\mathbb{E}\big[Z(i)^2Z(k)^2\big]+p_i\min\Big(\varrho_{n,k}^2\,,\,\varrho_{n,i}^2\Big)\\
&\qquad+p_i\varrho_{n,k}\varrho_{n,i}-S_0(i,k)\Big)\,. \qquad (3.15)
\end{aligned}
$$

For the last identity we have used the elementarily verifiable fact that

$$\mathbb{E}\big[Z(i)^2 Z(k)^2\big] = 1 + 2v_{i,k}^2$$

for all $1 \le i, k \le r$.

For $p_i < p_k$, by the orthogonality of the Hoeffding decomposition and by the Cauchy-Schwarz inequality we have that

$$\begin{aligned}
\mathrm{Cov}\big(W(i)^2, W(k)^2\big) &= \sum_{\substack{M,N\subseteq[n]:\\ |M|\le 2p_i, |N|\le 2p_k}} \mathbb{E}\big[U_M(i,i)U_N(k,k)\big] - \mathbb{E}[W(i)^2]\mathbb{E}[W(k)^2] \\
&= \sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_i}} \mathbb{E}\big[U_M(i,i)U_M(k,k)\big] \\
&\le \sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_i}} \Big(\mathbb{E}\big[U_M(i,i)^2\big]\Big)^{1/2}\Big(\mathbb{E}\big[U_M(k,k)^2\big]\Big)^{1/2} \\
&\le \Big(\sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_i}} \mathbb{E}\big[U_M(i,i)^2\big]\Big)^{1/2}\Big(\sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_i}} \mathbb{E}\big[U_M(k,k)^2\big]\Big)^{1/2} \\
&= \Big(\sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_i}} \mathrm{Var}\big(U_M(i,i)\big)\Big)^{1/2}\Big(\sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_i}} \mathrm{Var}\big(U_M(k,k)\big)\Big)^{1/2} \\
&= \Big(\mathbb{E}\big[W(i)^4\big] - 1\Big)^{1/2}\Big(\sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_i}} \mathrm{Var}\big(U_M(k,k)\big)\Big)^{1/2}. \qquad (3.16)
\end{aligned}$$

Since $p_i < p_k$, by means of (3.16) we can further bound

$$\begin{aligned}
\mathrm{Cov}\big(W(i)^2, W(k)^2\big) &\le \Big(\mathbb{E}\big[W(i)^4\big] - 1\Big)^{1/2}\Big(\sum_{\substack{M\subseteq[n]:\\ 1\le|M|\le 2p_k-1}} \mathrm{Var}\big(U_M(k,k)\big)\Big)^{1/2} \\
&= \Big(\mathbb{E}\big[W(i)^4\big] - 1\Big)^{1/2}\Big(\mathrm{Var}\big(W(k)^2\big) - \sum_{\substack{M\subseteq[n]:\\ |M|=2p_k}} \mathbb{E}\big[U_M(k,k)^2\big]\Big)^{1/2} \\
&\le \Big(\mathbb{E}\big[W(i)^4\big] - 1\Big)^{1/2}\Big(\mathbb{E}\big[W(k)^4\big] - 3 + \big(2p_k + C_{p_k}\big)\varrho_{n,k}^2\Big)^{1/2}, \qquad (3.17)
\end{aligned}$$

where the final inequality is true by (2.13).

From (3.14) and (3.17) and from (3.15), respectively, we thus obtain the following result.

Lemma 3.4. *Let $1 \le i \le k \le r$.*

(i) *If $p_i < p_k$, then*

$$\begin{aligned}
n^2\mathbb{E}\big[S_{i,k}^2\big] \le (p_i + p_k)^2\Bigg[&\Big(\mathbb{E}\big[W(i)^4\big] - 1\Big)^{1/2}\Big(\mathbb{E}\big[W(k)^4\big] - 3 + \big(2p_k + C_{p_k}\big)\varrho_{n,k}^2\Big)^{1/2} \\
&+ \min\Big(p_i\varrho_{n,k}^2\,,\, p_k\varrho_{n,i}^2\Big) - S_0(i,k)\Bigg].
\end{aligned}$$

(ii) *If $p_i = p_k$, then*

$$n^2\mathbb{E}\big[S_{i,k}^2\big] \le 4p_i^2\Big(\mathbb{E}\big[W(i)^2W(k)^2\big] - \mathbb{E}\big[Z(i)^2Z(k)^2\big] + p_i\min\Big(\varrho_{n,k}^2\,,\,\varrho_{n,i}^2\Big) + p_i\varrho_{n,k}\varrho_{n,i} - S_0(i,k)\Big)\,.$$

It remains to bound the quantities $S_0(i,k)$, $1 \le i \le k \le r$. The concepts of free indices and bifold quadruples from Subsection 2.2 generalize in the obvious way to quadruples $(J_1,J_2,J_3,J_4) \in \mathcal{D}_{p_i}\times\mathcal{D}_{p_k}\times\mathcal{D}_{p_i}\times\mathcal{D}_{p_k} =: \mathcal{D}_{i,k}^4$. We denote by $\mathcal{B}_{i,k}$ the collection of all bifold quadruples in $\mathcal{D}_{i,k}^4$. Also, we denote by $\mathcal{T}_{i,k}$ the set of quadruples $(J_1,J_2,J_3,J_4) \in \mathcal{D}_{i,k}^4$ which are neither bifold nor have a free index, i.e. which satisfy

$$1_{J_1} + 1_{J_2} + 1_{J_3} + 1_{J_4} \ge 2\mathbf{1}_{J_1\cup J_2\cup J_3\cup J_4}$$

and there is a $j \in [n]$ such that

$$1_{J_1}(j) + 1_{J_2}(j) + 1_{J_3}(j) + 1_{J_4}(j) \ge 3\,.$$

With these definitions, for $1 \le i \le k \le r$, we define

$$\tau_{i,k} := \sum_{(J,K,L,M)\in\mathcal{T}_{i,k}} \sigma_J(i)\sigma_K(k)\sigma_L(i)\sigma_M(k)\,.$$

The next result is a generalization of Proposition 2.8.

Proposition 3.5. *With these definitions, for $1 \le i \le k \le r$, we have*

$$S_0(i,k) \ge -\tau_{i,k}\,.$$

The proof is postponed to Section 4. It remains to obtain a bound on the quantities $\tau_{i,k}$ in terms of $\varrho_{n,i}^2$ and $\varrho_{n,k}^2$. This is provided by the following result which generalizes Proposition 2.9. An outline of the main elements of the proof is given in Section 4.

Proposition 3.6. *For each $1 \le i,k \le r$, there exists a finite constant $C_{i,k}$ which depends on i and k only through p_i and p_k and which is independent of n such that*

$$\tau_{i,k} = \sum_{(J,K,L,M)\in\mathcal{T}_{i,k}} \sigma_J(i)\sigma_K(k)\sigma_L(i)\sigma_M(k) \le C_{i,k}\max\big(\varrho_{n,i}^2,\varrho_{n,k}^2\big)\,.$$

Furthermore, we have $C_{i,k} = C_{k,i}$.

Combining Propositions 3.5 and 3.6, we thus obtain that

$$S_0(i,k) \ge -C_{i,k}\max\big(\varrho_{n,i}^2,\varrho_{n,k}^2\big) \tag{3.18}$$

for all $1 \le i,k \le r$.

Observe that, using (3.3) and the symmetry of S, we can bound

$$\|\Lambda^{-1}\|_{\mathrm{op}}\mathbb{E}\big[\|S\|_{\mathrm{H.S.}}\big] \le \frac{n}{p_1}\Big(\mathbb{E}\big[\|S\|_{\mathrm{H.S.}}^2\big]\Big)^{1/2} = \frac{1}{p_1}\Big(\sum_{i,k=1}^r n^2\mathbb{E}\big[S_{i,k}^2\big]\Big)^{1/2}. \tag{3.19}$$

Now, using Lemma 3.4 we have

$$\begin{aligned}\sum_{i,k=1}^r n^2\mathbb{E}\big[S_{i,k}^2\big] &= \sum_{l,m=1}^s \sum_{i=r_{l-1}+1}^{r_l}\sum_{k=r_{m-1}+1}^{r_m} n^2\mathbb{E}\big[S_{i,k}^2\big]\\ &= \sum_{l=1}^s \sum_{i,k=r_{l-1}+1}^{r_l} n^2\mathbb{E}\big[S_{i,k}^2\big] + 2\sum_{1\le l<m\le s}\sum_{i=r_{l-1}+1}^{r_l}\sum_{k=r_{m-1}+1}^{r_m} n^2\mathbb{E}\big[S_{i,k}^2\big]\end{aligned}$$

$$\leq \sum_{l=1}^{s} 4q_l^2 \sum_{i,k=r_{l-1}+1}^{r_l} \Big(\mathbb{E}\big[W(i)^2 W(k)^2\big] - \mathbb{E}\big[Z(i)^2 Z(k)^2\big] + q_l \min\Big(\varrho_{n,k}^2\,,\, \varrho_{n,i}^2\Big) + q_l \varrho_{n,k}\varrho_{n,i} + C_{i,k} \max\big(\varrho_{n,i}^2, \varrho_{n,k}^2\big)\Big)$$
$$+ 2 \sum_{1\leq l<m\leq s} (q_l+q_m)^2 \sum_{i=r_{l-1}+1}^{r_l} \sum_{k=r_{m-1}+1}^{r_m} \Bigg[\Big(\mathbb{E}\big[W(i)^4\big] - 1\Big)^{1/2} \Big(\mathbb{E}\big[W(k)^4\big] - 3 + \big(2q_m + C_{q_m}\big)\varrho_{n,k}^2\Big)^{1/2} + \min\Big(q_l\varrho_{n,k}^2\,,\, q_m \varrho_{n,i}^2\Big) + C_{i,k}\max\big(\varrho_{n,i}^2, \varrho_{n,k}^2\big)\Bigg]$$
$$= A\,. \tag{3.20}$$

Here, the constants C_{q_m} are defined by Proposition 2.9.

Note that from Lemma 2.12 applied to the exchangeable pair $(W(i), W'(i))$ we have

$$\frac{n}{4p_i}\mathbb{E}\big|W'(i) - W(i)\big|^4 \leq 2\big(\mathbb{E}\big[W(i)^4\big] - 3\big) + 3\big(C_{p_i} + 2p_i\big)\varrho_{n,i}^2\,. \tag{3.21}$$

Using (3.3) as well as Jensen's inequality, we obtain

$$\|\Lambda^{-1}\|_{\mathrm{op}}\mathbb{E}\big[\|W'-W\|_2^3\big] = \frac{n}{p_1}\mathbb{E}\big[\|W'-W\|_2^3\big]$$
$$= \frac{n}{p_1}\Big(\sum_{i=1}^{r}\mathbb{E}\big|W'(i)-W(i)\big|^2\Big)^{3/2} = \frac{n}{p_1}\Big(r\sum_{i=1}^{r}\mathbb{E}\big|W'(i)-W(i)\big|^2\frac{1}{r}\Big)^{3/2}$$
$$\leq \frac{n}{p_1}r^{3/2}\sum_{i=1}^{r}\mathbb{E}\big|W'(i)-W(i)\big|^3\frac{1}{r}$$
$$= \frac{n}{p_1}r^{1/2}\sum_{i=1}^{r}\mathbb{E}\big|W'(i)-W(i)\big|^3\,.$$

Thus, by (3.21) we have

$$\|\Lambda^{-1}\|_{\mathrm{op}}\mathbb{E}\big[\|W'-W\|_2^3\big]$$
$$\leq \frac{n}{p_1}r^{1/2}\sum_{i=1}^{r}\Big(\mathbb{E}\big|W'(i)-W(i)\big|^2\Big)^{1/2}\Big(\mathbb{E}\big|W'(i)-W(i)\big|^4\Big)^{1/2}$$
$$= 2\sqrt{2r}\sum_{i=1}^{r}\frac{p_i}{p_1}\Big(\frac{n}{4p_i}\mathbb{E}\big|W'(i)-W(i)\big|^4\Big)^{1/2}$$
$$\leq 2\sqrt{2r}\sum_{i=1}^{r}\frac{p_i}{p_1}\Big(2\big(\mathbb{E}\big[W(i)^4\big]-3\big)+3\big(C_{p_i}+2p_i\big)\varrho_{n,i}^2\Big)^{1/2}$$
$$= 2\sqrt{2r}\sum_{l=1}^{s}\frac{q_l}{q_1}\sum_{i=r_{l-1}+1}^{r_l}\Big(2\big(\mathbb{E}\big[W(i)^4\big]-3\big)+3\big(C_{q_l}+2q_l\big)\varrho_{n,i}^2\Big)^{1/2}\,. \tag{3.22}$$

Theorem 1.6 now follows from Theorem 3.1 and from the respective bounds (3.19), (3.20) and (3.22).

4 Proofs of several technical results

Lemma 4.1. *In the situation of Section 3, for all* $1 \leq i \leq k \leq r$ *we have*

$$\sum_{\substack{J\in\mathcal{D}_{p_i}, K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}} \sigma_J(i)^2\sigma_K(k)^2 \leq \min\Big(p_i\varrho_{n,k}^2\,,\, p_k\varrho_{n,i}^2\Big)\,.$$

Proof. Note that we have

$$\begin{aligned}\sum_{\substack{J\in\mathcal{D}_{p_i},K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}}\sigma_J(i)^2\sigma_K(k)^2 &= \sum_{J\in\mathcal{D}_{p_i}}\sigma_J(i)^2\sum_{\substack{K\in\mathcal{D}_{p_k}:\\ J\cap K\neq\emptyset}}\sigma_K(k)^2\\ &\leq \sum_{J\in\mathcal{D}_{p_i}}\sigma_J(i)^2\sum_{j\in J}\sum_{\substack{K\in\mathcal{D}_{p_k}:\\ j\in K}}\sigma_K(k)^2\\ &\leq p_i\varrho_k^2\sum_{J\in\mathcal{D}_{p_i}}\sigma_J(i)^2 = p_i\varrho_k^2\,.\end{aligned}$$

The claim follows by symmetry. □

Lemma 4.2 (Generalization of Lemma 4 of [12])**.** *Let $(J,K,L,M)\in\mathcal{B}_{i,k}$ be a bifold quadruple. Then, in the situation of Section 3 we have*

$$\mathbb{E}\big[W_JW_KW_LW_M\big] = \mathbb{E}\Big[\mathbb{E}\big[W_JW_K\,\big|\,\mathcal{F}_{J\Delta K}\big]\mathbb{E}\big[W_LW_M\,\big|\,\mathcal{F}_{L\Delta M}\big]\Big]\,.$$

Proof. We repeat the short proof from [12]. By independence, we have

$$\begin{aligned}\mathbb{E}\big[W_JW_KW_LW_M\big] &= \mathbb{E}\Big[W_JW_K\mathbb{E}\big[W_LW_M\,\big|\,\mathcal{F}_{L\cup M}\big]\Big]\\ &= \mathbb{E}\Big[W_JW_K\mathbb{E}\big[W_LW_M\,\big|\,\mathcal{F}_{(J\cup K)\cap(L\cup M)}\big]\Big]\\ &= \mathbb{E}\Big[\mathbb{E}\big[W_JW_K\,\big|\,\mathcal{F}_{(J\cup K)\cap(L\cup M)}\big]\mathbb{E}\big[W_LW_M\,\big|\,\mathcal{F}_{(J\cup K)\cap(L\cup M)}\big]\Big]\,.\end{aligned}$$

Now, the claim follows from the fact that for a bifold quadruple (J,K,L,M) the identity

$$(J\cup K)\cap(L\cup M) = J\Delta K = L\Delta M$$

holds true. □

Lemma 4.3 (Generalization of Lemma 3 of [12])**.** *In the situation of Section 3, for $1\leq i\leq k\leq r$, $J\in\mathcal{D}_{p_i}$ and $K\in\mathcal{D}_{p_k}$ we have*

$$\mathbb{E}\Big[\Big(\mathbb{E}\big[W_JW_K\,\big|\,\mathcal{F}_{J\Delta K}\big]\Big)^2\Big] \leq \sigma_J^2(i)\sigma_K^2(k)\,.$$

Proof. Again, we immitate the proof given in [12]. Using first the conditional version of the Cauchy-Schwarz inequality and then twice the independence of the underlying random variables $X_1,\dots,X_n$ we obtain

$$\begin{aligned}\mathbb{E}\Big[\Big(\mathbb{E}\big[W_JW_K\,\big|\,\mathcal{F}_{J\Delta K}\big]\Big)^2\Big] &\leq \mathbb{E}\Big[\mathbb{E}\big[W_J^2\,\big|\,\mathcal{F}_{J\Delta K}\big]\mathbb{E}\big[W_K^2\,\big|\,\mathcal{F}_{J\Delta K}\big]\Big]\\ &= \mathbb{E}\Big[\mathbb{E}\big[W_J^2\,\big|\,\mathcal{F}_{J\setminus K}\big]\mathbb{E}\big[W_K^2\,\big|\,\mathcal{F}_{K\setminus J}\big]\Big]\\ &= \mathbb{E}\Big[\mathbb{E}\big[W_J^2\,\big|\,\mathcal{F}_{J\setminus K}\big]\Big]\mathbb{E}\Big[\mathbb{E}\big[W_K^2\,\big|\,\mathcal{F}_{K\setminus J}\big]\Big]\\ &= \sigma_J^2(i)\sigma_K^2(k)\,.\end{aligned}$$

□

Proof of Proposition 3.5. We generalize the argument used in the proof of Proposition 5 (b) of [12]. For $l=1,\dots,p_i-1$ we have

$$0\leq \sum_{\substack{C\subseteq[n]:\\ |C|=p_i+p_k-2l}}\sum_{\substack{B,B'\subseteq C:\\ |B|=p_i-l,|B'|=p_k-l,\\ B\cap B'=\emptyset}}\mathbb{E}\Big[\Big(\sum_{\substack{J\in\mathcal{D}_{p_i},M\in\mathcal{D}_{p_k}:\\ J\setminus M=B,M\setminus J=B'}}\mathbb{E}\big[W_J(i)W_M(k)\,\big|\,\mathcal{F}_{J\Delta M}\big]\Big)^2\Big]$$

$$
\begin{aligned}
&= \sum_{\substack{C\subseteq[n]:\\ |C|=p_i+p_k-2l}} \sum_{\substack{B,B'\subseteq C:\\ |B|=p_i-l,|B'|=p_k-l,\\ B\cap B'=\emptyset}} \sum_{(J,K,L,M)\in\mathcal{D}^4_{i,k}:\ J\setminus M=L\setminus K=B, M\setminus J=K\setminus L=B'} \\
&\qquad\qquad \mathbb{E}\Big[\mathbb{E}\big[W_J(i)W_M(k)\,\big|\,\mathcal{F}_{J\Delta M}\big]\mathbb{E}\big[W_L(i)W_K(k)\,\big|\,\mathcal{F}_{L\Delta K}\big]\Big] \\
&= \sum_{\substack{(J,K,L,M)\in\mathcal{B}_{i,k}:\\ J\setminus M=L\setminus K, M\setminus J=K\setminus L,\\ |J\cap M|=l}} \mathbb{E}\big[W_J(i)W_K(k)W_L(i)W_M(k)\big] \\
&\quad + \sum_{\substack{(J,K,L,M)\in\mathcal{T}_{i,k}:\\ J\setminus M=L\setminus K, M\setminus J=K\setminus L,\\ |J\cap M|=l=|L\cap K|}} \mathbb{E}\Big[\mathbb{E}\big[W_J(i)W_M(k)\,\big|\,\mathcal{F}_{J\Delta M}\big]\mathbb{E}\big[W_L(i)W_K(k)\,\big|\,\mathcal{F}_{L\Delta K}\big]\Big] \\
&=: S_0(i,k,l) + R_l\,,
\end{aligned}
$$

where we have used Lemma 4.2 to obtain the second equality. Note that

$$\sum_{l=1}^{p_i-1} S_0(i,k,l) = S_0(i,k)$$

because for a bifold quadruple (J,K,L,M) the identity $J\setminus M = L\setminus K$ implies that $J\cap K = L\cap M = \emptyset$ and because we have

$$S_0(i,k) = \sum_{\substack{(J,K,L,M)\in\mathcal{B}_{i,k}:\\ J\cap K=L\cap M=\emptyset,\\ \emptyset\subsetneq J\cap M\subsetneq J}} \mathbb{E}\big[W_J(i)W_K(k)W_L(i)W_M(k)\big]\,.$$

Further, by the Cauchy-Schwarz inequality and by Lemma 4.3, we have

$$
\begin{aligned}
\Big|\sum_{l=1}^{p_i-1} R_l\Big| &\le \sum_{l=1}^{p_i-1} \sum_{\substack{(J,K,L,M)\in\mathcal{T}_{i,k}:\\ J\setminus M=L\setminus K, M\setminus J=K\setminus L,\\ |J\cap M|=l=|L\cap K|}} \Big(\mathbb{E}\Big[\Big(\mathbb{E}\big[W_J(i)W_M(k)\,\big|\,\mathcal{F}_{J\Delta M}\big]\Big)^2\Big]\Big)^{1/2} \\
&\qquad\qquad \Big(\mathbb{E}\Big[\Big(\mathbb{E}\big[W_L(i)W_K(k)\,\big|\,\mathcal{F}_{L\Delta K}\big]\Big)^2\Big]\Big)^{1/2} \\
&\le \sum_{(J,K,L,M)\in\mathcal{T}_{i,k}} \sigma_J(i)\sigma_K(k)\sigma_L(i)\sigma_M(k) = \tau_{i,k}\,.
\end{aligned}
$$

Thus, the claim follows. □

Proof of Proposition 2.9. In order to prove Proposition 2.9 let us review the following concepts and notation, introduced in [11]. For a quadruple $(J_1,J_2,J_3,J_4)\in\mathcal{D}_d^4$ write

$$I := J_1\cup J_2\cup J_3\cup J_4 = \{i_1,\dots,i_r\} \quad\text{with}\quad 1\le i_1<i_2<\ldots<i_r\le n$$

and define the *shadow* (J_1',J_2',J_3',J_4') of (J_1,J_2,J_3,J_4) by

$$J_l' := \big\{a\in\{1,\dots,r\} \,:\, i_a\in J_l\big\}\,,\quad 1\le l\le 4\,.$$

Note that since we have the equivalence

$$a\in J_l' \Leftrightarrow i_a\in J_l$$

the sets J_l' satisfy obvious relations like

$$|J_l'| = |J_l| = d\,,\quad |J_l'\cap J_m'| = |J_l\cap J_m| \quad\text{etc.} \tag{4.1}$$

and that a quadruple $(J_1,J_2,J_3,J_4)\in\mathcal{D}_d^4$ is completely determined by its shadow and by $I=\bigcup_{l=1}^4 J_l$. Note also that if $(J_1,J_2,J_3,J_4)\in\mathcal{T}_d$, then we have

$$J_1' \cup J_2' \cup J_3' \cup J_4' = \{1, \dots, r\}$$

for some $r \in \{d, d+1, \dots, 2d-1\}$ and $J_i \cap J_k \neq \emptyset$ for all $i, k = 1, 2, 3, 4$. Indeed, if, for instance, $J_1 \cap J_2$ were empty and $j_0 \in J_i$ for at least three values of $i \in \{1, 2, 3, 4\}$, then necessarily $j_0 \in J_3 \cap J_4$ implying $|J_3 \cup J_4| \leq 2d-1$. Hence, $J_1 \cup J_2 \not\subseteq J_3 \cup J_4$ because $|J_1 \cup J_2| = 2d$ by disjointness. Thus, (J_1, J_2, J_3, J_4) has a free index and, hence, cannot be in $\mathcal{T}_d$. By the above observation (4.1), this immediateley implies that also $J_i' \cap J_k' \neq \emptyset$ for all $i, k = 1, 2, 3, 4$.

In general, we call a quadruple of sets $\mathbb{F} = (F_1, F_2, F_3, F_4)$ a *shadow* (a *d-shadow*) if there is an $r \in \{d, d+1, \dots, 2d-1\}$ such that $F := F_1 \cup F_2 \cup F_3 \cup F_4 = \{1, \dots, r\}$ and $|F_l| = d$ for $l = 1, 2, 3, 4$. We call r the *size* of the shadow $\mathbb{F}$. We say that the shadow $\mathbb{F}$ is *induced* by the quadruple $(J_1, J_2, J_3, J_4) \in \mathcal{D}_d^4$, if $\mathbb{F} = (J_1', J_2', J_3', J_4')$. We write $\mathbb{F}(J_1, J_2, J_3, J_4)$ for the shadow induced by (J_1, J_2, J_3, J_4). If $\mathbb{F}' = (F_1', F_2', F_3', F_4')$ is another d-shadow with $F' := F_1' \cup F_2' \cup F_3' \cup F_4' = \{1, \dots, r'\}$, then we say that $\mathbb{F}$ and $\mathbb{F}'$ are *equivalent* and write $\mathbb{F} \sim \mathbb{F}'$, if $r = r'$ and there is a permutation $\sigma \in \mathbb{S}_r$ such that

$$F_l' = \sigma(F_l) \quad \text{for } l = 1, 2, 3, 4. \tag{4.2}$$

We denote the latter fact by $\mathbb{F}' = \mathbb{F}_\sigma$. This clearly defines an equivalence relation on the set of d-shadows and we denote by $[\mathbb{F}]_\sim$ the equivalence class of $\mathbb{F}$. We further denote by $\gamma(\mathbb{F})$ the number of permutations $\sigma \in \mathbb{S}_r$ that leave $\mathbb{F}$ fixed in the sense that

$$\sigma(F_l) = F_l \quad \text{for all } l = 1, 2, 3, 4. \tag{4.3}$$

The set of these permutaions is just the *stabilizer* of $\mathbb{F}$ with respect to the natural action of $\mathbb{S}_r$ on the set of d-shadows of size r. Note that, for $\mathbb{F}' \sim \mathbb{F}$, we have $\gamma(\mathbb{F}) = \gamma(\mathbb{F}')$ and that $\gamma(\mathbb{F})$ also gives the number of permutations σ such that (4.2) holds. Let us define the function $g : [n]^d \to \mathbb{R}$ by

$$g(j_1, \dots, j_d) := \begin{cases} \sigma_{\{j_1, \dots, j_d\}}\,, & \text{if } \big|\{j_1, \dots, j_d\}\big| = d \\ 0\,, & \text{otherwise.} \end{cases}$$

Then, g is a symmetric function vanishing on the complement $\Delta^c = [n]^d \setminus \Delta$ of

$$\Delta := \Delta_d^{(n)} := \{(j_1, \dots, j_d) \in [n]^d \,:\, j_l \neq j_m \text{ whenever } l \neq m\}\,.$$

Further, for a shadow $\mathbb{F} = (F_1, F_2, F_3, F_4)$ which is induced by some quadruple $(J_1, J_2, J_3, J_4) \in \mathcal{D}_d^4$ and with $F := F_1 \cup F_2 \cup F_3 \cup F_4 = \{1, \dots, r\}$ and π_{F_l} being the natural projection $[n]^F \to [n]^{F_l}$ given by $(j_a)_{a \in F} \mapsto (j_a)_{a \in F_l}$, define $G_{F_l} : [n]^F \to \mathbb{R}$ by $G_{F_l} := g \circ \pi_{F_l}$. Here, we tacitly identify $[n]^d$ with $[n]^{F_l}$ and $[n]^r$ with $[n]^F$.

Lemma 4.4. *Let $\mathbb{F} = (F_1, F_2, F_3, F_4)$ be a d-shadow of size r which is induced by some quadruple $(J_1, J_2, J_3, J_4) \in \mathcal{T}$. Then, we have the bound*

$$\sum_{\substack{(J,K,L,M) \in \mathcal{T}: \\ \mathbb{F}(J,K,L,M) \in [\mathbb{F}]_\sim}} \sigma_J \sigma_K \sigma_L \sigma_M \leq \frac{d!(d-1)!}{\gamma(\mathbb{F})} \varrho_n^2\,.$$

Proof. For ease of notation, in this proof we use bold letters $\mathbf{a}$ to denote tuples $\mathbf{a} = (a_1, \dots, a_s) \in [n]^s$, where s is some natural number. Also, for two such tuples $\mathbf{a} = (a_1, \dots, a_s) \in [n]^s$ and $\mathbf{b} = (b_1, \dots, b_t) \in [n]^t$ we write $\mathbf{a} \cap \mathbf{b} \neq \emptyset$ if there are indices $1 \leq i \leq s$ and $1 \leq j \leq t$ such that $a_i = b_j$, i.e. if

$$\{a_1, \dots, a_s\} \cap \{b_1, \dots, b_t\} \neq \emptyset\,.$$

We begin the proof with the remark that

$$\sum_{(i_1,\dots,i_r)\in[n]^r} G_{F_1}(i_1,\dots,i_r)G_{F_2}(i_1,\dots,i_r)G_{F_3}(i_1,\dots,i_r)G_{F_4}(i_1,\dots,i_r)$$
$$\geq \gamma(\mathbb{F}) \sum_{\substack{(J,K,L,M)\in\mathcal{T}:\\ \mathbb{F}(J,K,L,M)\in[\mathbb{F}]_\sim}} \sigma_J\sigma_K\sigma_L\sigma_M\,.$$

This follows from

$$\begin{aligned}
\sum_{(i_1,\dots,i_r)\in[n]^r}\prod_{l=1}^4 G_{F_l}(i_1,\dots,i_r) &\geq \sum_{(i_1,\dots,i_r)\in[n]^r_{\neq}}\prod_{l=1}^4 G_{F_l}(i_1,\dots,i_r)\\
&= \sum_{1\leq j_1<\dots<j_r\leq n}\sum_{\sigma\in\mathbb{S}_r}\prod_{l=1}^4 g\big(\pi_{F_l}(j_{\sigma(1)},\dots,j_{\sigma(r)})\big)\\
&= \sum_{1\leq j_1<\dots<j_r\leq n}\sum_{\sigma\in\mathbb{S}_r}\prod_{l=1}^4 g\big(\pi_{\sigma(F_l)}(j_1,\dots,j_r)\big)\\
&= \gamma(\mathbb{F})\sum_{\mathbb{F}'\in[\mathbb{F}]_\sim}\sum_{1\leq j_1<\dots<j_r\leq n}\prod_{l=1}^4 G_{F_l'}(j_1,\dots,j_r)\\
&= \gamma(\mathbb{F})\sum_{\mathbb{F}'\in[\mathbb{F}]_\sim}\sum_{\substack{(J,K,L,M)\in\mathcal{T}:\\ \mathbb{F}(J,K,L,M)=\mathbb{F}'}}\sigma_J\sigma_K\sigma_L\sigma_M\\
&= \gamma(\mathbb{F})\sum_{\substack{(J,K,L,M)\in\mathcal{T}:\\ \mathbb{F}(J,K,L,M)\in[\mathbb{F}]_\sim}}\sigma_J\sigma_K\sigma_L\sigma_M\,.
\end{aligned}$$

Here, we used the notation $[n]^r_{\neq}$ for the set of all tuples $(i_1,\dots,i_r)\in[n]^r$ such that $i_j\neq i_k$ whenever $j\neq k$. Hence, it suffices to show that we always have the bound

$$\sum_{(i_1,\dots,i_r)\in[n]^r}\prod_{l=1}^4 G_{F_l}(i_1,\dots,i_r)\leq d!(d-1)!\varrho_n^2$$

if $\mathbb{F}$ is as in the statement of the Lemma.

We first treat the simple cases that either two or all of the sets F_l, $l=1,2,3,4$, are equal. Note that the case of exactly three equal sets is vacuous for a quadruple in $\mathcal{T}$. Assume first that e.g. $F_3\neq F_1=F_2\neq F_4$. It might be that also $F_3=F_4$ but this is immaterial. Then, we have

$$\begin{aligned}
&\sum_{(i_1,\dots,i_r)\in[n]^r} G_{F_1}(i_1,\dots,i_r)G_{F_2}(i_1,\dots,i_r)G_{F_3}(i_1,\dots,i_r)G_{F_4}(i_1,\dots,i_r)\\
&= \sum_{(i_1,\dots,i_r)\in[n]^F} G^2_{F_1}(i_1,\dots,i_r)G_{F_3}(i_1,\dots,i_r)G_{F_4}(i_1,\dots,i_r)\\
&= \sum_{\mathbf{j}\in[n]^{F_1}} g(\mathbf{j})^2\sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_3}(\mathbf{j},\mathbf{k})G_{F_4}(\mathbf{j},\mathbf{k})\\
&\leq \sum_{\mathbf{j}\in[n]^{F_1}} g(\mathbf{j})^2\Big(\sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G^2_{F_3}(\mathbf{j},\mathbf{k})\Big)^{1/2}\Big(\sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G^2_{F_4}(\mathbf{j},\mathbf{k})\Big)^{1/2}\\
&= \sum_{\mathbf{j}\in[n]^{F_1}} g(\mathbf{j})^2\Big(\sum_{\mathbf{l}\in[n]^{F_3}:\mathbf{l}\cap\mathbf{j}\neq\emptyset} g(\mathbf{l})^2\Big)^{1/2}\Big(\sum_{\mathbf{m}\in[n]^{F_4}:\mathbf{m}\cap\mathbf{j}\neq\emptyset} g(\mathbf{m})^2\Big)^{1/2}
\end{aligned}$$

$$\leq (d-1)!\varrho_n^2 \sum_{\mathbf{j}\in[n]^{F_1}} g(\mathbf{j})^2$$
$$= d!(d-1)!\varrho_n^2$$

Note that the second inequality follows from the fact that $F_1 \cap F_3 \neq \emptyset$ and $F_1 \cap F_4 \neq \emptyset$ in this case as well as by the definition of ϱ_n^2. If $F_1 = F_2 = F_3 = F_4$, then we have $r = d$ and

$$\sum_{(i_1,\dots,i_r)\in[n]^r} G_{F_1}(i_1,\dots,i_r)G_{F_2}(i_1,\dots,i_r)G_{F_3}(i_1,\dots,i_r)G_{F_4}(i_1,\dots,i_r)$$
$$= \sum_{(j_1,\dots,j_d)\in[n]^d} g(j_1,\dots,j_d)^4$$
$$\leq \sum_{j_1=1}^{n} \max_{(j_2,\dots,j_d)\in[n]^{d-1}} g(j_1,\dots,j_d)^2 \sum_{(k_2,\dots,k_d)\in[n]^{d-1}} g(j_1,k_2,\dots,k_d)^2$$
$$\leq (d-1)!\varrho_n^2 \sum_{j_1=1}^{n} \max_{(j_2,\dots,j_d)\in[n]^{d-1}} g(j_1,\dots,j_d)^2$$
$$\leq (d-1)!\varrho_n^2 \sum_{(j_1,\dots,j_d)\in[n]^d} g(j_1,\dots,j_d)^2$$
$$= d!(d-1)!\varrho_n^2\,.$$

For the remainder of this proof we may thus assume that the sets F_l, $l = 1,2,3,4$, are pairwise different. Then, using the Cauchy-Schwarz inequality, we can bound

$$\sum_{(i_1,\dots,i_r)\in[n]^r} G_{F_1}(i_1,\dots,i_r)G_{F_2}(i_1,\dots,i_r)G_{F_3}(i_1,\dots,i_r)G_{F_4}(i_1,\dots,i_r)$$
$$= \sum_{\mathbf{j}\in[n]^{F_1}} g(\mathbf{j}) \sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_2}(\mathbf{j},\mathbf{k})G_{F_3}(\mathbf{j},\mathbf{k})G_{F_4}(\mathbf{j},\mathbf{k})$$
$$\leq \Big(\sum_{\mathbf{j}\in[n]^{F_1}} g(\mathbf{j})^2\Big)^{1/2} \Big(\sum_{\mathbf{j}\in[n]^{F_1}} \Big(\sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_2}(\mathbf{j},\mathbf{k})G_{F_3}(\mathbf{j},\mathbf{k})G_{F_4}(\mathbf{j},\mathbf{k})\Big)^2\Big)^{1/2}$$
$$= \sqrt{d!}\Big(\sum_{\mathbf{j}\in[n]^{F_1}} \Big(\sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_2}(\mathbf{j},\mathbf{k})G_{F_3}(\mathbf{j},\mathbf{k})G_{F_4}(\mathbf{j},\mathbf{k})\Big)^2\Big)^{1/2} \tag{4.4}$$

Thus, it remains to bound the quantity

$$A := \sum_{\mathbf{j}\in[n]^{F_1}} \Big(\sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_2}(\mathbf{j},\mathbf{k})G_{F_3}(\mathbf{j},\mathbf{k})G_{F_4}(\mathbf{j},\mathbf{k})\Big)^2\,.$$

Let us distinguish the following cases.

1) Each element in $F = F_1 \cup F_2 \cup F_3 \cup F_4$ appears in at least three of the sets F_1, F_2, F_3, F_4. This implies that

$$F \setminus F_k = F_j \setminus F_k \quad \text{for all distinct } j,k \in \{2,3,4\}\,.$$

Then, using Cauchy-Schwarz, we can bound

$$A \leq \sum_{\mathbf{j}\in[n]^{F_1}} \Big(\sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_2}^2(\mathbf{j},\mathbf{k}) \sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_3}^2(\mathbf{j},\mathbf{k})G_{F_4}^2(\mathbf{j},\mathbf{k})\Big)$$
$$= \sum_{\mathbf{j}\in[n]^{F_1}} \Big(\sum_{\mathbf{l}\in[n]^{F_2}:\mathbf{l}\cap\mathbf{j}\neq\emptyset} g(\mathbf{l})^2 \sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_3}^2(\mathbf{j},\mathbf{k})G_{F_4}^2(\mathbf{j},\mathbf{k})\Big)$$

$$
\begin{aligned}
&\le (d-1)!\varrho_n^2 \sum_{\mathbf{j}\in[n]^{F_1}} \sum_{\mathbf{k}\in[n]^{F\setminus F_1}} G_{F_3}^2(\mathbf{j},\mathbf{k}) G_{F_4}^2(\mathbf{j},\mathbf{k}) \\
&= (d-1)!\varrho_n^2 \sum_{\mathbf{k}\in[n]^{F\setminus F_1}} \sum_{\mathbf{l}\in[n]^{F_1\cap F_3\cap F_4}} \sum_{\mathbf{a}\in[n]^{(F_3\cap F_1)\setminus F_4}} g(\mathbf{k},\mathbf{l},\mathbf{a})^2 \sum_{\mathbf{b}\in[n]^{(F_4\cap F_1)\setminus F_3}} g(\mathbf{k},\mathbf{l},\mathbf{b})^2 \\
&\le \big((d-1)!\big)^2 \varrho_n^4 \sum_{\mathbf{k}\in[n]^{F\setminus F_1}} \sum_{\mathbf{l}\in[n]^{F_1\cap F_3\cap F_4}} \sum_{\mathbf{a}\in[n]^{(F_3\cap F_1)\setminus F_4}} g(\mathbf{k},\mathbf{l},\mathbf{a})^2 \\
&= \big((d-1)!\big)^2 \varrho_n^4 \sum_{\mathbf{j}\in[n]^{F_3}} g(\mathbf{j})^2 \\
&= d!\big((d-1)!\big)^2 \varrho_n^4 .
\end{aligned}
$$

Note that we have used the fact that

$$(F_4\cap F_1)\setminus F_3 = F\setminus F_3 \neq \emptyset$$

to obtain the last inequality.

2) There is an element $j_0 \in F = F_1\cup F_2\cup F_3\cup F_4$ which is contained in exactly two of the sets F_1, F_2, F_3, F_4. We may assume that $j_0 \in F_1$. We claim that then there are distinct indices $j,k \in \{2,3,4\}$ such that

$$F_1 \not\subseteq F_j \cup F_k .$$

Indeed, we have

$$(F_2\cup F_3)\cap(F_2\cup F_4)\cap(F_3\cup F_4) = (F_2\cap F_3)\cup(F_2\cap F_4)\cup(F_3\cap F_4)$$

and, hence, j_0 cannot be contained in the set on the right hand side. Thus, let us assume that $F_1 \not\subseteq F_3\cup F_4$. We obtain that

$$
\begin{aligned}
A &\le \sum_{\mathbf{j}\in[n]^{F_1\setminus F_2}} \sum_{\mathbf{a}\in[n]^{F_1\cap F_2}} \Bigg(\sum_{\mathbf{k}\in[n]^{F_2\setminus F_1}} g(\mathbf{a},\mathbf{k}) \sum_{\mathbf{l}\in[n]^{(F_3\cup F_4)\setminus(F_1\cup F_2)}} G_{F_3}(\mathbf{j},\mathbf{a},\mathbf{k},\mathbf{l}) G_{F_4}(\mathbf{j},\mathbf{a},\mathbf{k},\mathbf{l}) \Bigg)^2 \\
&\le \sum_{\mathbf{j}\in[n]^{F_1\setminus F_2}} \sum_{\mathbf{a}\in[n]^{F_1\cap F_2}} \Big(\sum_{\mathbf{k}\in[n]^{F_2\setminus F_1}} g(\mathbf{a},\mathbf{k})^2 \Big) \\
&\quad \cdot \sum_{\mathbf{k}\in[n]^{F_2\setminus F_1}} \Bigg(\sum_{\mathbf{l}\in[n]^{(F_3\cup F_4)\setminus(F_1\cup F_2)}} G_{F_3}(\mathbf{j},\mathbf{a},\mathbf{k},\mathbf{l}) G_{F_4}(\mathbf{j},\mathbf{a},\mathbf{k},\mathbf{l}) \Bigg)^2 \\
&= \sum_{\mathbf{j}\in[n]^{F_1\setminus F_2}} \sum_{\mathbf{a_1}\in[n]^{F_1\cap F_2\cap(F_3\cup F_4)}} \Bigg(\sum_{\mathbf{a_2}\in[n]^{F_1\cap F_2\setminus(F_3\cup F_4)}} \sum_{\mathbf{k}\in[n]^{F_2\setminus F_1}} g(\mathbf{a_1},\mathbf{a_2},\mathbf{k})^2 \Bigg) \\
&\qquad\qquad \cdot \sum_{\mathbf{k}\in[n]^{F_2\setminus F_1}} \Bigg(\sum_{\mathbf{l}\in[n]^{(F_3\cup F_4)\setminus(F_1\cup F_2)}} G_{F_3}(\mathbf{j},\mathbf{a_1},\mathbf{a_2^*},\mathbf{k},\mathbf{l}) G_{F_4}(\mathbf{j},\mathbf{a_1},\mathbf{a_2^*},\mathbf{k},\mathbf{l}) \Bigg)^2 \\
&\le (d-1)!\varrho_n^2 \sum_{\mathbf{j}\in[n]^{F_1\setminus F_2}} \sum_{\mathbf{a_1}\in[n]^{F_1\cap F_2\cap(F_3\cup F_4)}} \sum_{\mathbf{k}\in[n]^{F_2\setminus F_1}} \\
&\qquad\qquad \cdot \Bigg(\sum_{\mathbf{l}\in[n]^{(F_3\cup F_4)\setminus(F_1\cup F_2)}} G_{F_3}(\mathbf{j},\mathbf{a_1},\mathbf{a_2^*},\mathbf{k},\mathbf{l}) G_{F_4}(\mathbf{j},\mathbf{a_1},\mathbf{a_2^*},\mathbf{k},\mathbf{l}) \Bigg)^2 ,
\end{aligned}
$$

where

$$\mathbf{a_2^*} \in [n]^{F_1\cap F_2\setminus(F_3\cup F_4)}$$

is arbitrary but fixed. Now note that due to the fact that $\mathbb{F}$ is induced by some quadruple in $\mathcal{T}$ we have

$$F_3 \cup F_4 = (F_1 \setminus F_2) \cup (F_2 \setminus F_1) \cup \big[(F_3 \cup F_4) \cap F_1 \cap F_2\big] \cup \big[(F_3 \cup F_4) \setminus (F_1 \cup F_2)\big]$$
$$= (F_1 \setminus F_2) \cup (F_2 \setminus F_1) \cup \big[(F_3 \cup F_4) \cap F_1 \cap F_2\big] \cup \big[(F_3 \cap F_4) \setminus (F_1 \cup F_2)\big],$$

where the union on the right hand side is disjoint. Thus, the last bound becomes

$$\begin{aligned} A &\le (d-1)!\varrho_n^2 \sum_{\mathbf{m}\in[n]^{(F_3\cup F_4)\cap(F_1\cup F_2)}} \sum_{\mathbf{l}\in[n]^{(F_3\cap F_4)\setminus(F_1\cup F_2)}} G_{F_3}^2(\mathbf{m},\mathbf{l}) \\ &\qquad \cdot \sum_{\mathbf{p}\in[n]^{(F_3\cap F_4)\setminus(F_1\cup F_2)}} G_{F_4}^2(\mathbf{m},\mathbf{p}) \\ &= (d-1)!\varrho_n^2 \sum_{\mathbf{a}\in[n]^{(F_3\cap F_4)\cap(F_1\cup F_2)}} \sum_{\mathbf{b}\in[n]^{(F_3\setminus F_4)\cap(F_1\cup F_2)}} \sum_{\mathbf{l}\in[n]^{(F_3\cap F_4)\setminus(F_1\cup F_2)}} g(\mathbf{a},\mathbf{b},\mathbf{l})^2 \\ &\qquad \sum_{\mathbf{c}\in[n]^{(F_4\setminus F_3)\cap(F_1\cup F_2)}} \sum_{\mathbf{p}\in[n]^{(F_3\cap F_4)\setminus(F_1\cup F_2)}} g(\mathbf{a},\mathbf{c},\mathbf{p})^2 \\ &\le \big((d-1)!\big)^2 \varrho_n^4 \sum_{\mathbf{j}\in[n]^{F_3}} g(\mathbf{j})^2 \\ &= d!\big((d-1)!\big)^2 \varrho_n^4 \end{aligned}$$

□

End of the proof of Proposition (2.9). Let $\mathbb{F}_1, \dots, \mathbb{F}_s$ be a complete system of pairwise non-equivalent d-shadows which are induced by quadruples $(J,K,L,M) \in \mathcal{T}$. Then, clearly, s is independent of n and by Lemma 4.4 we have

$$\tau = \sum_{(J,K,L,M)\in\mathcal{T}} \sigma_J\sigma_K\sigma_L\sigma_M = \sum_{j=1}^{s} \sum_{\substack{(J,K,L,M)\in\mathcal{T}:\\ \mathbb{F}(J,K,L,M)\in[\mathbb{F}_j]_\sim}} \sigma_J\sigma_K\sigma_L\sigma_M$$
$$\le \Big(d!(d-1)!\sum_{j=1}^{s}\gamma(\mathbb{F}_j)^{-1}\Big)\varrho_n^2$$

so that we can let

$$C_d := d!(d-1)!\sum_{j=1}^{s}\gamma(\mathbb{F}_j)^{-1} \tag{4.5}$$

which is independent of n. □

Remark 4.5. Using the fact that the equivalence class of a shadow $\mathbb{F} = (F_1, F_2, F_3, F_4)$ is determined by the cardinalities of all finite intersections of the sets F_1, F_2, F_3, F_4, one can get an upper bound on the number s of all equivalence classes of shadows induced by quadruples in $\mathcal{T}$. Using that $\gamma(\mathbb{F}) \ge 1$ immediately gives a crude bound on C_d. It is not difficult to verify that $C_2 = 13$ by distinguishing all possible cases. Furthermore, by some clever combinatorial argument, it might be possible to compute sharp bounds on C_d starting from (4.5). This would be of great interest for deriving limit theorems in situations where $d = d_n \to \infty$ with n. We leave this as an interesting problem for possible future work.

Idea of the proof of Proposition 3.6. The proof of Proposition 2.9 can be easily generalized to the present situation by introducing the concept of a (p_i, p_k)-*shadow* corresponding to a quadruple $(J_1, J_2, J_3, J_4) \in \mathcal{D}_{i,k}^4$ and following exactly the same lines of the proof. We have, however, refrained from giving the proof in this more general situation for mainly two reasons. Firstly, the proof of Proposition 2.9 already involves a lot of notation and introducing even more of it might make the argument less transparent. Secondly, and more importantly, the precise dependence of the constant $C_{i,k}$ on p_i and p_k would be more complicated and less explicit than the formula given by (4.5) which

can be exactly evaluated for small values of d and, as mentioned in Remark 4.5, might be suitably bounded for general d. □

Acknowledgments. We would like to thank the referee for their detailed report which helped us fix several typos and improve the exposition of our results.

References

[1] O. Arizmendi, *Convergence of the fourth moment and infinite divisibility*, Probab. Math. Statist. **33** (2013), no. 2, 201–212. MR-3158548

[2] E. Azmoodeh, S. Campese, and G. Poly, *Fourth Moment Theorems for Markov diffusion generators*, J. Funct. Anal. **266** (2014), no. 4, 2341–2359. MR-3150163

[3] E. Bolthausen, *Exact convergence rates in some martingale central limit theorems*, Ann. Probab. **10** (1982), no. 3, 672–688. MR-0659537

[4] S. Bourguin and G. Peccati, *Portmanteau inequalities on the Poisson space: mixed regimes and multidimensional clustering*, Electron. J. Probab. **19** (2014), no. 66, 42. MR-3248195

[5] S. Bourguin and G. Peccati, *Semicircular limits on the free Poisson chaos: counterexamples to a transfer principle*, J. Funct. Anal. **267** (2014), no. 4, 963–997. MR-3217054

[6] S. Campese, I. Nourdin, G. Peccati, and G. Poly, *Multivariate Gaussian approximations on Markov chaoses*, Electron. Commun. Probab. **21** (2016), Paper No. 48, 9. MR-3522594

[7] S. Chatterjee, J. Fulman, and A. Röllin, *Exponential approximation by Stein's method and spectral graph theory*, ALEA Lat. Am. J. Probab. Math. Stat. **8** (2011), 197–223. MR-2802856

[8] S. Chatterjee and E. Meckes, *Multivariate normal approximation using exchangeable pairs*, ALEA Lat. Am. J. Probab. Math. Stat. **4** (2008), 257–283. MR-2453473

[9] S. Chatterjee and Q.-M. Shao, *Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model*, Ann. Appl. Probab. **21** (2011), no. 2, 464–483. MR-2807964

[10] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao, *Normal approximation by Stein's method*, Probability and its Applications (New York), Springer, Heidelberg, 2011. MR-2732624

[11] P. de Jong, *Central limit theorems for generalized multilinear forms*, CWI Tract, vol. 61, Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam, 1989. MR-1002734

[12] P. de Jong, *A central limit theorem for generalized multilinear forms*, J. Multivariate Anal. **34** (1990), no. 2, 275–289. MR-1073110 MR-1073110

[13] C. Döbler, *New developments in Stein's method with applications*, (2012), (Ph.D.)-Thesis Ruhr-Universität Bochum.

[14] C. Döbler, *Stein's method of exchangeable pairs for the Beta distribution and generalizations*, Electron. J. Probab. **20** (2015), no. 109, 1–34. MR-3418541

[15] E. B. Dynkin and A. Mandelbaum, *Symmetric statistics, Poisson point processes, and multiple Wiener integrals*, Ann. Statist. **11** (1983), no. 3, 739–745. MR-0707925

[16] P. Eichelsbacher and M. Löwe, *Stein's method for dependent random variables occurring in statistical mechanics*, Electron. J. Probab. **15** (2010), no. 30, 962–988. MR-2659754

[17] P. Eichelsbacher and C. Thäle, *New Berry-Esseen bounds for non-linear functionals of Poisson random measures*, Electron. J. Probab. **19** (2014), no. 102, 25. MR-3275854

[18] O. El-Dakkak and G. Peccati, *Hoeffding decompositions and urn sequences*, Ann. Probab. **36** (2008), no. 6, 2280–2310. MR-2478683

[19] O. El-Dakkak, G. Peccati, and I. Prünster, *Exchangeable Hoeffding decompositions over finite sets: a combinatorial characterization and counterexamples*, J. Multivariate Anal. **131** (2014), 51–64. MR-3252635

[20] T. Fissler and C. Thäle, *A four moments theorem for gamma limits on a Poisson chaos*, ALEA Lat. Am. J. Probab. Math. Stat. **13** (2016), no. 1, 163–192. MR-3476211

[21] J. Fulman and N. Ross, *Exponential approximation and Stein's method of exchangeable pairs*, ALEA Lat. Am. J. Probab. Math. Stat. **10** (2013), no. 1, 1–13. MR-3083916

[22] G. G. Gregory, *Large sample theory for U-statistics and tests of fit*, Ann. Statist. **5** (1977), no. 1, 110–123. MR-0433669

[23] E. Haeusler, *On the rate of convergence in the central limit theorem for martingales with discrete and continuous time*, Ann. Probab. **16** (1988), no. 1, 275–299. MR-0920271

[24] C. C. Heyde and B. M. Brown, *On the departure from normality of a certain class of martingales*, Ann. Math. Statist. **41** (1970), 2161–2165. MR-0293702

[25] W. Hoeffding, *A class of statistics with asymptotically normal distribution*, Ann. Math. Statistics **19** (1948), 293–325. MR-0026294

[26] S. R. Jammalamadaka and S. Janson, *Limit theorems for a triangular scheme of U-statistics with applications to inter-point distances*, Ann. Probab. **14** (1986), no. 4, 1347–1358. MR-0866355

[27] S. Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR-1474726

[28] O. Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR-1876169

[29] S. Karlin and Y. Rinott, *Applications of ANOVA type decompositions for comparisons of conditional variance statistics including jackknife estimates*, Ann. Statist. **10** (1982), no. 2, 485–501. MR-0653524

[30] T. Kemp, G. Nourdin, I.and Peccati, and R. Speicher, *Wigner chaos and the fourth moment*, Ann. Probab. **40** (2012), no. 4, 1577–1635. MR-2978133

[31] V. S. Koroljuk and Yu. V. Borovskich, *Theory of U-statistics*, Mathematics and its Applications, vol. 273, Kluwer Academic Publishers Group, Dordrecht, 1994, Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors. MR-1472486

[32] K. Krokowski, A. Reichenbachs, and C. Thäle, *Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation*, to appear in: Ann. Probab.

[33] R. Lachièze-Rey and G. Peccati, *New Kolmogorov bounds for functionals of binomial point processes*, to appear in: Ann. Appl. Probab.

[34] R. Lachièze-Rey and G. Peccati, *Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs*, Electron. J. Probab. **18** (2013), no. 32, 32. MR-3035760

[35] R. Lachièze-Rey and G. Peccati, *Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric U-statistics*, Stochastic Process. Appl. **123** (2013), no. 12, 4186–4218. MR-3096352

[36] M. Ledoux, *Chaos of a Markov operator and the fourth moment condition*, Ann. Probab. **40** (2012), no. 6, 2439–2459. MR-3050508

[37] W. G. McGinley and R. Sibson, *Dissociated random variables*, Math. Proc. Cambridge Philos. Soc. **77** (1975), 185–188. MR-0370714

[38] E. Meckes, *On Stein's method for multivariate normal approximation*, High dimensional probability V: the Luminy volume, Inst. Math. Stat. Collect., vol. 5, Inst. Math. Statist., Beachwood, OH, 2009, pp. 153–178. MR-2797946

[39] E. Mossel, R. O'Donnell, and K. Oleszkiewicz, *Noise stability of functions with low influences: invariance and optimality*, Ann. of Math. (2) **171** (2010), no. 1, 295–341. MR-2630040

[40] I. Nourdin and G. Peccati, *Stein's method on Wiener chaos*, Probab. Theory Related Fields **145** (2009), no. 1–2, 75–118. MR-2520122

[41] I. Nourdin and G. Peccati, *Normal approximations with Malliavin calculus*, Cambridge Tracts in Mathematics, vol. 192, Cambridge University Press, Cambridge, 2012, From Stein's method to universality. MR-2962301

[42] I. Nourdin, G. Peccati, G. Poly, and R. Simone, *Multidimensional limit theorems for homogeneous sums: a survey and a general transfer principle*, ESAIM Probab. Stat. **20** (2016), 293–308. MR-3533710

[43] I. Nourdin, G. Peccati, and G. Reinert, *Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos*, Ann. Probab. **38** (2010), no. 5, 1947–1985. MR-2722791

[44] I. Nourdin, G. Peccati, and G. Reinert, *Stein's method and stochastic analysis of Rademacher functionals*, Electron. J. Probab. **15** (2010), no. 55, 1703–1742. MR-2735379

[45] I. Nourdin, G. Peccati, and R. Speicher, *Multi-dimensional semicircular limits on the free Wigner chaos*, Seminar on Stochastic Analysis, Random Fields and Applications VII, Progr. Probab., vol. 67, Birkhäuser/Springer, Basel, 2013, pp. 211–221. MR-3380342

[46] D. Nualart and G. Peccati, *Central limit theorems for sequences of multiple stochastic integrals*, Ann. Probab. **33** (2005), no. 1, 177–193. MR-2118863

[47] G. Peccati, *Hoeffding-ANOVA decompositions for symmetric statistics of exchangeable observations*, Ann. Probab. **32** (2004), no. 3A, 1796–1829. MR-2073178

[48] G. Peccati and M. Reitzner (eds.), *Stochastic analysis for poisson point processes*, Springer-Verlag, 2016.

[49] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet, *Stein's method and normal approximation of Poisson functionals*, Ann. Probab. **38** (2010), no. 2, 443–478. MR-2642882

[50] G. Peccati and C. Thäle, *Gamma limits and U-statistics on the Poisson space*, ALEA Lat. Am. J. Probab. Math. Stat. **10** (2013), no. 1, 525–560. MR-3083936

[51] G. Peccati and C. A. Tudor, *Gaussian limits for vector-valued multiple stochastic integrals*, Séminaire de Probabilités XXXVIII, Lecture Notes in Math., vol. 1857, Springer, Berlin, 2005, pp. 247–262. MR-2126978

[52] G. Peccati and C. Zheng, *Multi-dimensional Gaussian fluctuations on the Poisson space*, Electron. J. Probab. **15** (2010), no. 48, 1487–1527. MR-2727319

[53] G. Peccati and C. Zheng, *Universal Gaussian fluctuations on the discrete Poisson chaos*, Bernoulli **20** (2014), no. 2, 697–715. MR-3178515

[54] M. Penrose, *Random geometric graphs*, Oxford Studies in Probability, vol. 5, Oxford University Press, Oxford, 2003. MR-1986198

[55] N. Privault and G.L. Torrisi, *The Stein and Chen-Stein methods for functionals of non-symmetric Bernoulli processes*, ALEA Lat. Am. J. Probab. Math. Stat. **12** (2015), no. 1, 309–356. MR-3357130

[56] G. Reinert and A. Röllin, *Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition*, Ann. Probab. **37** (2009), no. 6, 2150–2173. MR-2573554

[57] M. Reitzner and M. Schulte, *Central limit theorems for U-statistics of Poisson point processes*, Ann. Probab. **41** (2013), no. 6, 3879–3909. MR-3161465

[58] Y. Rinott and V. Rotar, *On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics*, Ann. Appl. Probab. **7** (1997), no. 4, 1080–1105. MR-1484798

[59] A. Röllin, *A note on the exchangeability condition in Stein's method*, Statist. Probab. Lett. **78** (2008), no. 13, 1800–1806. MR-2453918

[60] H. Rubin and R. A. Vitale, *Asymptotic distribution of symmetric statistics*, Ann. Statist. **8** (1980), no. 1, 165–170. MR-0557561

[61] M. Schulte, *Normal Approximation of Poisson Functionals in Kolmogorov Distance*, J. Theoret. Probab. **29** (2016), no. 1, 96–117. MR-3463079

[62] R. J. Serfling, *Approximation theorems of mathematical statistics*, John Wiley & Sons, Inc., New York, 1980, Wiley Series in Probability and Mathematical Statistics. MR-0595165

[63] C. Stein, *Approximate computation of expectations*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7, Institute of Mathematical Statistics, Hayward, CA, 1986. MR-0882007

[64] R. A. Vitale, *Covariances of symmetric statistics*, J. Multivariate Anal. **41** (1992), no. 1, 14–26. MR-1156678

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[2]EJMS: Electronic Journal Management System `http://www.vtex.lt/en/ejms.html`
[3]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[4]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[5]Project Euclid: `https://projecteuclid.org/`
[6]IMS Open Access Fund: `http://www.imstat.org/publications/open.htm`