Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2016, Vol. 52, No. 2, 763–803
DOI: 10.1214/14-AIHP652

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Berry–Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences

Kai Krokowski, Anselm Reichenbachs and Christoph Thäle

Faculty of Mathematics, Ruhr University Bochum, Germany.
E-mail: kai.krokowski@rub.de; anselm.reichenbachs@rub.de; christoph.thaele@rub.de

Received 3 April 2014; revised 1 October 2014; accepted 3 October 2014

Abstract. Berry–Esseen bounds for non-linear functionals of infinite Rademacher sequences are derived by means of the Malliavin–Stein method. Moreover, multivariate extensions for vectors of Rademacher functionals are shown. The results establish a connection to small ball probabilities and shed new light onto the relation between central limit theorems on the Rademacher chaos and norms of contraction operators. Applications concern infinite weighted 2-runs, a combinatorial central limit theorem and traces of Bernoulli random matrices.

Résumé. Nous dérivons des estimations de type Berry–Esseen pour des fonctionnelles non-linéaires de suites infinies de Rademacher par la méthode de Malliavin–Stein. De plus, nous prouvons des extensions multivariées pour des vecteurs de fonctionnelles de Rademacher. Ces résultats établissent une connexion avec les probabilités de petites boules et apportent un nouvel éclairage sur la relation entre des théorèmes centraux limites sur le chaos de Rademacher et les normes d'opérateurs de contraction. Des applications concernent des succès consécutifs pondérés, un théorème central limite combinatoire et des traces de matrices de Bernoulli aléatoires.

MSC: 60F05; 60G50; 60B20; 60H07

Keywords: Berry–Esseen bound; Central limit theorem; Malliavin calculus; Normal approximation; Rademacher chaos; Stein's method

1. Introduction

In the seminal paper [21] by Nourdin and Peccati it has been demonstrated for the first time that there is a powerful connection between Stein's method and the Malliavin calculus of variations. A main feature of this approach is that the various coupling constructions, which are usually in the background of Stein's method, are replaced by a structural property of the involved random variables. In practice, this is reflected by the use of Malliavin operators and especially through application of an integration-by-parts formula. We mention one of the main results of the paper [29] of Nualart and Peccati, which shows that a sequence of Gaussian multiple integrals $F_n := I_q(f_n)$ of fixed order $q \geq 1$ of symmetric and square-integrable functions f_n such that $\mathbb{E}[F_n^2] = 1$ for all $n \geq 1$ converges to a standard Gaussian random variable N if and only if the 4th cumulant

$$\kappa_4(F_n) := \mathbb{E}\big[F_n^4\big] - 3$$

converges to 0, as $n \to \infty$. Using Stein's method, this has been extended to an estimate on various probability distances between F_n and N (cf. [21,23,25]). In fact, a combination of the main results of [21] and [25] (see also Theorem 5.2.6 in [23]) asserts that

$$d_K(F_n, N) := \sup_{x\in\mathbb{R}} \big|P(F_n \leq x) - P(N \leq x)\big| \leq C\kappa_4(F_n)^{1/2} \tag{1.1}$$

with a constant $0 < C < \infty$ not depending on n. A key step in the proof of this bound is an estimate of $d_K(F_n, N)$ in terms of contraction operators $f_n \star_r^r f_n$, and a comparison with an expression for the 4th cumulant $\kappa_4(F_n)$ in terms of such contractions. We emphasize that the quantitative 4th moment theorem (1.1) is one of the cornerstones and at the heart of the Malliavin–Stein approach. The following optimal bound on the total variation distance $d_{\mathrm{TV}}(F_n, N) := \sup\{|P(F_n \in A) - P(N \in A)|\colon A \subset \mathbb{R} \text{ Borel}\}$ has recently been derived in [24]. Namely, if F_n converges in distribution to a standard Gaussian random variable N, then

$$C_1 \max\bigl\{\bigl|\mathbb{E}\bigl[F_n^3\bigr]\bigr|, \kappa_4(F_n)\bigr\} \le d_{\mathrm{TV}}(F_n, N) \le C_2 \max\bigl\{\bigl|\mathbb{E}\bigl[F_n^3\bigr]\bigr|, \kappa_4(F_n)\bigr\}, \tag{1.2}$$

where $0 < C_1 < C_2 < \infty$ are constants which are independent of n. Since $d_K(F_n, N) \le d_{\mathrm{TV}}(F_n, N)$, this improves (1.1) significantly in case that $\mathbb{E}[F_n^3] = 0$, for example, if q is odd. Since its first appearance, the 4th moment theorem has attracted considerable interest and has further been exploited by many authors. Selected applications concern central limit theorems for non-linear functionals of Gaussian stochastic processes [20], random fields on the sphere [18], random matrices [22] and universality of homogeneous sums [25] (we also refer to the monograph [23] and the exhaustive list of references therein).

Besides non-linear functionals of Gaussian random measures, there is another branch to which the Malliavin–Stein approach has been applied, namely non-linear functionals of Poisson random measures, see the papers [30,32,34] of Peccati, Solé, Taqqu, Utzet and Zheng. These general results have found numerous applications especially in geometric probability and stochastic geometry, see [14–16,37,40] for distinguished examples developed by Lachièze-Rey, Last, Peccati, Penrose, Reitzner, Schulte and Thäle. We emphasize that it has been shown by Eichelsbacher and Thäle [10], and Lachièze-Rey and Peccati [14], who combined Stein's method with the Malliavin calculus of variations on the Poisson space, that a quantitative 4th moment theorem similar to (1.1) is also available for multiple stochastic integrals with respect to Poisson random measures if the integrands f_n are non-negative.

In [26], Nourdin, Peccati and Reinert pushed this line of research further by proving quantitative central limit theorems for functionals of so-called infinite Rademacher sequences, which rely on a combination of Stein's method with tools from discrete stochastic analysis as developed in [35] by Privault. By a Rademacher sequence we mean in this paper an infinite sequence $X = (X_n)_{n\in\mathbb{N}}$ of independent and identically distributed random variables such that X_n takes the values ± 1 with probability $1/2$. The aim of this paper is to develop the theory of [26] further in several directions. In particular, our main findings are:

(i) An estimate for the Kolmogorov distance between a possibly non-linear functional of a Rademacher sequence X (this is what we call a Rademacher functional) and a Gaussian random variable in terms of Malliavin operators. This refines the bounds of [26], where only smooth distances have been considered. The proof of this bound is a non-trivial task as it relies on various new computations involving discrete Malliavin operators and also on a new integration-by-parts formula.

(ii) A connection between our Malliavin–Stein bound and quantities, which are known as small ball probabilities. They are a measure of anti-concentration and enter the expression for the Kolmogorov distance. We mention that they were not visible in the previous work [26] because of the smoothness of the test functions used there. It is worth pointing out that small ball probabilities in the context of Berry–Esseen bounds have previously found attention in the works [17,39] of Litvak, Pajor Rudelson and Tomczak-Jaegermann, and Rudelson and Vershynin, respectively.

(iii) A quantitative multivariate central limit theorem dealing with the distance between a vector of Rademacher functionals and a Gaussian random vector. In particular, we show that a vector consisting of discrete multiple stochastic integrals satisfies a multivariate central limit theorem if its entries fulfil univariate central limit theorems. This is the discrete analogue to a similar phenomenon observed by Nourdin, Peccati and Réveillac [27], Nualart and Ortiz-Latorre [28], Peccati and Tudor [33] and Peccati and Zheng [34] for Gaussian or Poisson multiple integrals, respectively.

(iv) A clarification of the rôle of contraction operators in light of necessary conditions for a sequence of discrete multiple stochastic integrals to satisfy a central limit theorem. We develop a new necessary criterion for discrete double integrals, which is based on a novel representation of the 4th cumulant involving on- and off-diagonal terms. We also present a counterexample showing that vanishing contraction norms do not provide a necessary condition. This sheds new light onto a 4th moment theorem for Rademacher functionals, and general quadratic forms as considered by de Jong [8] or Chatterjee [6], for example.

(v) Applications of our results to infinite weighted 2-runs, an extended version of a combinatorial central limit theorem and traces of powers of Bernoulli random matrices. This extends the previous results from [5] and also some of the findings in [22,26]. In particular, we provide a Berry–Esseen bound for an infinite-dimensional version of a central limit theorem of Blei and Janson and give a direct proof for a multivariate central limit theorem for traces of powers of Bernoulli matrices without resorting to universality results.

Our results rely on Stein's method for normal approximation and tools from discrete stochastic analysis. We recall both together with some other preliminaries in Section 2. There, we also provide a new integration-by-parts formula on which our proofs are based on. The one-dimensional Malliavin–Stein bound is the content of Section 3, while Section 4 discusses various versions in case of sequences of discrete multiple stochastic integrals. In this context we also develop the announced new necessary criterion for a sequence of discrete stochastic double integrals to satisfy a central limit theorem. Multivariate central limit theorems for vectors of Rademacher functionals are provided in Section 5. The final Section 6 contains the applications of our results.

2. Preliminaries

2.1. *Rademacher sequences*

By a Rademacher sequence $(X_n)_{n\in\mathbb{N}}$ we understand a sequence consisting of i.i.d. random variables X_n defined on some probability space $(\Omega, \mathcal{F}, P)$ such that

$$P(X_n = -1) = P(X_n = +1) = \frac{1}{2}.$$

They are constructed in the canonical way, namely by taking

$$\Omega := \{-1, +1\}^{\mathbb{N}}, \qquad \mathcal{F} := \mathcal{P}\big(\{-1, +1\}\big)^{\otimes\mathbb{N}}, \qquad P := \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}\right)^{\otimes\mathbb{N}},$$

where $\delta_{\pm 1}$ is the unit-mass Dirac measure concentrated at ± 1, and then putting $X_n(\omega) := \omega_n$ for $(\omega_n)_{n\in\mathbb{N}} \in \Omega$. Here and below, we write $\mathcal{P}(M)$ for the power set of a set M.

2.2. *Kernels and contractions*

In what follows we will denote by κ the counting measure on $\mathbb{N}$. For $n \geq 1$ define $\ell^2(\mathbb{N})^{\otimes n} := L^2(\mathbb{N}^n, \mathcal{P}(\mathbb{N})^{\otimes n}, \kappa^{\otimes n})$. Functions in $\ell^2(\mathbb{N})^{\otimes n}$ are called kernels in the sequel. The following subsets of $\ell^2(\mathbb{N})^{\otimes n}$ are of interest. Let $\ell^2(\mathbb{N})^{\circ n}$ denote the class of symmetric kernels and $\ell_0^2(\mathbb{N})^{\otimes n}$ denote the class of kernels vanishing on the diagonal, i.e., on the complement Δ_n^c of the set

$$\Delta_n := \big\{(i_1, \ldots, i_n) \in \mathbb{N}^n\colon\ i_k \neq i_l \text{ for } k \neq l\big\}.$$

Let $\ell_0^2(\mathbb{N})^{\circ n}$ denote the class of symmetric kernels vanishing on diagonals. For $q = 0$ one defines $\ell^2(\mathbb{N})^{\circ 0} := \mathbb{R}$. For integers $n, m \geq 1$, $r \in \{0, \ldots, n \wedge m\}$, $l \in \{0, \ldots, r\}$ and kernels $f \in \ell_0^2(\mathbb{N})^{\circ n}$ and $g \in \ell_0^2(\mathbb{N})^{\circ m}$ the contraction

$$\begin{aligned} &f \star_r^l g(i_1, \ldots, i_{n-r}, j_1, \ldots, j_{r-l}, k_1, \ldots, k_{m-r}) \\ &\quad := \sum_{(a_1,\ldots,a_l)\in\Delta_l} f(i_1, \ldots, i_{n-r}, j_1, \ldots, j_{r-l}, a_1, \ldots, a_l) g(k_1, \ldots, k_{m-r}, j_1, \ldots, j_{r-l}, a_1, \ldots, a_l) \end{aligned}$$

arises from the tensor product of f and g by first identifying r of the $n + m$ variables and then by integrating out l of them with respect to the counting measure κ. In particular, for $f \in \ell_0^2(\mathbb{N})^{\circ n}$ the contraction $f \star_0^0 f$ is the tensor product of f with itself. For a function f on $\mathbb{N}^n$ denote by $\tilde{f} = \frac{1}{n!}\sum_\sigma f(i_{\sigma(1)}, \ldots, i_{\sigma(n)})$ its canonical symmetrization, where the sum runs over all permutations σ of $\{1, \ldots, n\}$. Since $f \star_r^l g$ is usually not symmetric, we often consider its canonical symmetrization $\widetilde{f \star_r^l g}$. Note that $\|\tilde{f}\|_{\ell^2(\mathbb{N})^{\otimes n}} \leq \|f\|_{\ell^2(\mathbb{N})^{\otimes n}}$ for any $f \in \ell^2(\mathbb{N})^{\otimes n}$.

Lemma 2.1. *Let $q \geq 2$ and suppose that $f \in \ell_0^2(\mathbb{N})^{\circ q}$.*

(i) *It holds that*

$$(2q)!\big\|\widetilde{f \star_0^0 f}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2q}} = 2\big(q!\|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big)^2 + \sum_{r=1}^{q-1}(q!)^2\binom{q}{r}^2\big\|f \star_r^r f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}. \tag{2.1}$$

(ii) *One has*

$$\big\|f \star_1^0 f\big\|_{\ell^2(\mathbb{N})^{\otimes 2q-1}} = \big\|f \star_q^{q-1} f\big\|_{\ell^2(\mathbb{N})} \tag{2.2}$$

and

$$\big\|f \star_r^{r-1} f\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)+1}} \leq \big\|f \star_{r-1}^{r-1} f\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r+1)}} \tag{2.3}$$

for every $r \in \{2, \dots, q\}$.

Proof. Identity (2.1) is formula (11.6.30) in [31] and for part (ii) we refer to Lemma 2.4 in [26]. □

2.3. *Multiple stochastic integrals and chaotic decomposition*

For $q \geq 1$ and $f \in \ell_0^2(\mathbb{N})^{\circ q}$ the discrete multiple stochastic integral of order q of f is defined as

$$J_q(f) := q! \sum_{1 \leq i_1 < \cdots < i_q < \infty} f(i_1, \dots, i_q) X_{i_1} \cdots X_{i_q}. \tag{2.4}$$

For $q = 0$ and $c \in \mathbb{R}$ we put $J_0(c) := c$. The family of random variables of the form $J_q(f)$ with $f \in \ell_0^2(\mathbb{N})^{\circ q}$ is called Rademacher chaos of order q (sometimes also called the Walsh chaos of order q). Multiple stochastic integrals fulfil the isometry relation

$$\mathbb{E}\big[J_q(f)J_p(g)\big] = \mathbb{1}_{\{q=p\}} q! \langle f, g\rangle_{\ell^2(\mathbb{N})^{\otimes q}}. \tag{2.5}$$

Moreover, it is a crucial fact that every $F \in L^2(\Omega)$ possesses a unique decomposition in terms of multiple stochastic integrals, i.e., each $F \in L^2(\Omega)$ can be written as

$$F = \mathbb{E}(F) + \sum_{n=1}^{\infty} J_n(f_n),$$

where $f_n \in \ell_0^2(\mathbb{N})^{\circ q}$, $n \geq 1$, is a uniquely determined sequence of kernels (see Section 6 in [35]). In our paper, the following multiplication formula for discrete stochastic integrals will turn out to be crucial (see Proposition 2.9 in [26]). It says that for all integers $p, q \geq 1$ and symmetric kernels $f \in \ell_0^2(\mathbb{N})^{\circ q}$ and $g \in \ell_0^2(\mathbb{N})^{\circ p}$, it holds that

$$J_q(f)J_p(g) = \sum_{r=0}^{q\wedge p} r! \binom{q}{r}\binom{p}{r} J_{q+p-2r}\big(\widetilde{(f \star_r^r g)}\mathbb{1}_{\Delta_{q+p-2r}}\big). \tag{2.6}$$

2.4. *Discrete Malliavin operators*

We are now going to introduce the four important so-called Malliavin operators for which we refer to [35].

2.4.1. *Gradient operator*
The gradient operator D transforms random variables into random sequences and is defined as

$$DJ_q(f_q) := \big(D_k J_q(f_q)\big)_{k\in\mathbb{N}} = \big(J_{q-1}\big(f_q(\cdot,k)\big)\big)_{k\in\mathbb{N}} \in L^2(\Omega\times\mathbb{N}, P\otimes\kappa)$$

on a Rademacher chaos of fixed order $q \geq 1$. We also put $D_k J_0(c) = 0$ for $c \in \mathbb{R}$ and $k \in \mathbb{N}$. It can consistently be extended to the class of functionals $F \in L^2(\Omega)$ of the form $F = \mathbb{E}(F) + \sum_{n=1}^{\infty} J_n(f_n)$, which satisfy the relation

$$\mathbb{E}\big[\|DF\|^2_{\ell^2(\mathbb{N})}\big] = \sum_{n=1}^{\infty} nn!\|f_n\|^2_{\ell^2(\mathbb{N})^{\otimes n}} < \infty. \tag{2.7}$$

The class of all such functionals is the domain of D and will be denoted by $\mathrm{dom}(D)$. For $F \in \mathrm{dom}$ one has

$$DF = (D_k F)_{k\in\mathbb{N}} = \left(\sum_{n=1}^{\infty} n J_{n-1}\big(f_n(\cdot,k)\big)\right)_{k\in\mathbb{N}}.$$

The operator D also admits a pathwise representation, which can be used as an alternative definition, especially if the functional does not satisfy condition (2.7). To spell this out, define for $\omega = (\omega_n)_{n\in\mathbb{N}} \in \Omega$,

$$\omega_k^+ := (\omega_1,\dots,\omega_{k-1},+1,\omega_{k+1},\dots) \quad \text{and} \quad \omega_k^- := (\omega_1,\dots,\omega_{k-1},-1,\omega_{k+1},\dots).$$

We further define for $F \in L^2(\Omega)$ and $\omega \in \Omega$ the functionals F_k^+ and F_k^- by

$$F_k^+(\omega) := F\big(\omega_k^+\big) \quad \text{and} \quad F_k^-(\omega) := F\big(\omega_k^-\big).$$

Then, for $F \in \mathrm{dom}(D)$ one has the relation

$$D_k F(\omega) = \frac{1}{2}\big(F\big(\omega_k^+\big) - F\big(\omega_k^-\big)\big).$$

Formally, let us denote by D' the operator acting on $F \in L^2(\Omega)$ by

$$D'_k F(\omega) := \frac{1}{2}\big(F\big(\omega_k^+\big) - F\big(\omega_k^-\big)\big).$$

We shall discuss the difference between D and D' now, but before, we define inductively the higher-order gradients by putting $D^n_{k_1,\dots,k_n} F := D_{k_1} D^{n-1}_{k_2,\dots,k_n} F$ for $n \geq 1$, where $D^0 := \mathrm{Id}$ and $D^1 := D$, and similarly for $(D')^n_{k_1,\dots,k_n} F$.

Lemma 2.2.

(i) *If* $F = \sum_{n=0}^{\infty} J_n(f_n)$ *with* $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$ *then*

$$\mathbb{E}[F\cdot X_{k_1}\cdots X_{k_n}] = n! f_n(k_1,\dots,k_n) \quad \textit{for all } (k_1,\dots,k_n) \in \Delta_n.$$

(ii) *For* $F \in L^2(\Omega)$ *it holds that*

$$\mathbb{E}\big[\big(D'\big)^n_{k_1,\dots,k_n} F\big] = \mathbb{E}[F\cdot X_{k_1}\cdots X_{k_n}] \quad \textit{for all } (k_1,\dots,k_n) \in \Delta_n. \tag{2.8}$$

(iii) *For* $F, G \in L^2(\Omega)$ *and* $k \in \mathbb{N}$ *one has*

$$D_k(FG) = G D_k F + F D_k G - 2X_k (D_k F)(D_k G). \tag{2.9}$$

Remark 2.1. *Lemma* 2.2(i) *and* (ii) *show that*

$$f_n(k_1,\dots,k_n) = \frac{1}{n!}\mathbb{E}\big[\big(D'\big)^n_{k_1,\dots,k_n} F\big] \tag{2.10}$$

if F has chaotic decomposition $F = \sum_{n=0}^{\infty} J_n(f_n)$. This is the analogue of the classical Stroock formula for Rademacher functionals.

Proof of Lemma 2.2. Part (i) follows directly from the definition (2.4) of a discrete multiple stochastic integral. To prove (ii) we first observe that for every $k \in \mathbb{N}$,

$$\mathbb{E}\big[F_k^+\big] = \mathbb{E}[X_k \cdot F] + \mathbb{E}[F] \quad \text{and} \quad \mathbb{E}\big[F_k^-\big] = \mathbb{E}[-X_k \cdot F] + \mathbb{E}[F].$$

We thus get for $k \in \mathbb{N}$ that

$$\begin{aligned}\mathbb{E}\big[D'_k F\big] &= \frac{1}{2}\mathbb{E}\big[\big(F_k^+ - F_k^-\big)\big] = \frac{1}{2}\big(\big(\mathbb{E}[X_k \cdot F] + \mathbb{E}[F]\big) - \big(\mathbb{E}[-X_k \cdot F] + \mathbb{E}[F]\big)\big) \\ &= \mathbb{E}[X_k \cdot F].\end{aligned}$$

The general case is proved by induction. Let $n \in \mathbb{N}$ and $(k_1, \dots, k_{n+1}) \in \Delta_{n+1}$. Then, using the induction hypothesis and the fact that from the point of view of $D_{k_1,\dots,k_n}$ the k_{n+1}st entry $X_{k_{n+1}}$ of the Rademacher sequence X behaves like a constant, we find that

$$\begin{aligned}\mathbb{E}\big[\big(D'\big)^{n+1}_{k_1,\dots,k_{n+1}} F\big] &= \mathbb{E}\big[D'_{k_{n+1}}\big(\big(D'\big)^n_{k_1,\dots,k_n} F\big)\big] = \mathbb{E}\big[X_{k_{n+1}} \cdot \big(D'\big)^n_{k_1,\dots,k_n} F\big] \\ &= \mathbb{E}\big[\big(D'\big)^n_{k_1,\dots,k_n}(F \cdot X_{k_{n+1}})\big] = \mathbb{E}\big[(F \cdot X_{k_{n+1}}) \cdot X_{k_1} \cdots X_{k_n}\big] \\ &= \mathbb{E}[F \cdot X_{k_1} \cdots X_{k_{n+1}}].\end{aligned}$$

This proves assertion (ii). Part (iii) corresponds to Proposition 7.8 in [35]. □

The next lemma formalizes Remark 2.11 in [26].

Lemma 2.3. *Let $F \in L^2(\Omega)$. Then $\mathbb{E}[\sum_{k=1}^{\infty}(D'_k F)^2] < \infty$, if and only if $F \in \operatorname{dom}(D)$.*

Proof. Let $F = \mathbb{E}[F] + \sum_{k=1}^{\infty} J_n(f_n)$ be the chaotic decomposition of the square integrable random variable F for a sequence of kernels $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$. The condition $\mathbb{E}[\sum_{k=1}^{\infty}(D'_k F)^2] < \infty$ implies that $D'_k F \in L^2(\Omega)$ for all $k \in \mathbb{N}$. Therefore $D'_k F$ has a chaotic decomposition of the form

$$D'_k F = \sum_{n=0}^{\infty} J_n\big(f'_n\big)$$

for a sequence of kernels $f'_n \in \ell_0^2(\mathbb{N})^{\circ n}$. The Stroock formula (2.10) yields

$$\begin{aligned}f'_n(k_1, \dots, k_n) &= \frac{1}{n!}\mathbb{E}\big[\big(D'\big)^n_{k_1,\dots,k_n} D'_k F\big] = \frac{(n+1)!}{n!} f_{n+1}(k_1, \dots, k_n, k) \\ &= (n+1) f_{n+1}(k_1, \dots, k_n, k).\end{aligned}$$

We thus get the representation

$$D'_k F = \sum_{n=0}^{\infty}(n+1) J_n\big(f_{n+1}(\cdot, k)\big) = \sum_{n=1}^{\infty} n J_{n-1}\big(f_n(\cdot, k)\big)$$

for the chaotic decomposition of $D'_k F$ which is equal to $D_k F$ and implies that $F \in \operatorname{dom}$. This completes the proof. □

2.4.2. *Divergence operator*
The divergence operator is the adjoint of D. It is given by

$$\delta\big(J_n\big(u_{n+1}(*,\cdot)\big)\big) := J_{n+1}(\widetilde{u_{n+1}}), \quad n \in \{0,1,\ldots\},$$

if $u_{n+1} \in \ell^2(\mathbb{N})^{\circ n} \otimes \ell^2(\mathbb{N})$. It can consistently be extended to the class of random functions $u(*,\cdot) \in L^2(\Omega \times \mathbb{N}, P \otimes \kappa)$ with $u(*,k) = \sum_{n=0}^{\infty} J_n(u_{n+1}(*,k))$ and kernels $u_{n+1}(*,k) \in \ell_0^2(\mathbb{N})^{\circ n}$ satisfying

$$\mathbb{E}\big[\delta(u)^2\big] = \sum_{n=0}^{\infty} (n+1)! \|\widetilde{u_{n+1}}\|^2_{\ell^2(\mathbb{N})^{\otimes(n+1)}} < \infty.$$

The class of these random functions is called the domain of δ and is denoted by $\mathrm{dom}(\delta)$. Similar to the difference operator, also the divergence operator admits a pathwise representation, namely

$$\delta(u) = \sum_{k=1}^{\infty} u_k X_k - \sum_{k=1}^{\infty} D_k u_k \tag{2.11}$$

for $u \in \mathrm{dom}(\delta)$, see [35], Proposition 9.3.

2.4.3. *Ornstein–Uhlenbeck operator and its inverse*
The Ornstein–Uhlenbeck operator L is defined by the relation

$$L := -\delta D \tag{2.12}$$

for elements of a fixed Rademacher chaos, see [35], Proposition 10.1. In other words this means that

$$L J_n(f_n) = -n J_n(f_n)$$

for $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$. The domain of L is the class of all functionals $F = \sum_{n=0}^{\infty} J_n(f_n) \in L^2(\Omega)$ such that $\mathbb{E}[(LF)^2] = \sum_{n=1}^{\infty} n^2 n! \|f_n\|^2_{\ell^2(\mathbb{N})^{\otimes n}} < \infty$. We notice that L maps F into the class of square-integrable centred random variables $L_0^2(\Omega)$. The (pseudo-) inverse operator L^{-1} of L is defined on $L_0^2(\Omega)$ by

$$L^{-1} F = -\sum_{n=1}^{\infty} \frac{1}{n} J_n(f_n)$$

if $F \in L_0^2(\Omega)$ has representation $F = \sum_{n=1}^{\infty} J_n(f_n)$.

2.5. *Important identities*

The following lemma collects two important identities, namely the integration-by-parts formula and an isometric formula for the divergence operator.

Lemma 2.4.

(i) *Let* $F \in \mathrm{dom}(D)$ *and* $u \in \mathrm{dom}(\delta)$, *then*

$$\mathbb{E}\big[F\delta(u)\big] = \mathbb{E}\big[\langle DF, u\rangle_{\ell^2(\mathbb{N})}\big]. \tag{2.13}$$

(ii) *For all* $u \in \mathrm{dom}(\delta)$ *it holds that*

$$\mathbb{E}\big[\delta(u)^2\big] = \mathbb{E}\big[\|u\|^2_{\ell^2(\mathbb{N})}\big] + \mathbb{E}\left[\sum_{k,l=1}^{\infty} (D_k u_l)(D_l u_k)\right]. \tag{2.14}$$

Proof. Part (i) is [35], Proposition 9.2, and for part (ii) we refer to [35], Proposition 9.3. □

In the proof of one of our main results we will encounter the expression $D'\mathbb{1}_{\{F>x\}}$ with $x \in \mathbb{R}$ and would like to apply the integration-by-parts formula (2.13) to it. Unfortunately, it is not clear in general whether the integrability condition of Lemma 2.3 is satisfied for $\mathbb{1}_{\{F>x\}}$ or not. To overcome this difficulty we follow the strategy introduced in [41] in a different context and now develop an integration-by-parts formula for functionals F not necessarily belonging to $\mathrm{dom}(D)$.

Lemma 2.5. *Suppose that $F \in L^2(\Omega)$ is bounded, $u \in L^2(\Omega \times \mathbb{N})$ is such that $u_k := u(\cdot, k)$ is independent of X_k and $(D'_k F)u_k \geq 0$ for all $k \in \mathbb{N}$. Then,*

$$\mathbb{E}\big[F\delta(u)\big] = \mathbb{E}\big[\langle D'F, u\rangle_{\ell^2(\mathbb{N})}\big].$$

Proof. First observe that $\delta(u) = \sum_{k=1}^{\infty} u_k X_k$ by (2.11), due to the independence assumption. Using (2.8), we thus find that

$$\begin{aligned}\mathbb{E}\big[\langle D'F, u\rangle\big] &= \mathbb{E}\left[\sum_{k=1}^{\infty}\big(D'_k F\big)u_k\right] = \sum_{k=1}^{\infty}\mathbb{E}\big[D'_k(F\cdot u_k)\big]\\ &= \sum_{k=1}^{\infty}\mathbb{E}\big[(F\cdot u_k)\cdot X_k\big]\\ &= \mathbb{E}\big[F\cdot\delta(u)\big].\end{aligned}$$

The exchange of the order of summation is justified by Fubini's theorem, since $(D'_k F)u(k) \geq 0$ by assumption and since

$$\mathbb{E}\left[\sum_{k=1}^{\infty}|F\cdot u_k X_k|\right] \leq C\mathbb{E}\left[\sum_{k=1}^{\infty}|u_k|\right] < \infty$$

for a suitable constant $0 < C < \infty$ by boundedness of F and square integrability of u. □

From now on and to simplify the notation we will write DF for the discrete gradient applied to a Rademacher functional $F \in L^2(\Omega)$ and interpret this as $D'F$ if $F \in L^2(\Omega) \setminus \mathrm{dom}(D)$.

2.6. *Stein's method for one-dimensional normal approximation*

It is a well known fact that a real-valued random variable Z follows a standard normal (or Gaussian) distribution if and only if

$$\mathbb{E}\big[f'(Z) - Zf(Z)\big] = 0$$

for all bounded, continuous and piecewise continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$ satisfying $\mathbb{E}|f'(Z)| < \infty$, see [7, Lemma 2.1]. This is the so-called Stein-type characterization of the standard normal distribution and the corresponding Stein-equation reads

$$f'(z) - zf(z) = \mathbb{1}(z \leq x) - P(N \leq x), \quad x \in \mathbb{R}, \tag{2.15}$$

where N stands for a standard normal random variable. For a given x, a solution of (2.15) will be denoted by $f_x(z)$. Taking expectations in (2.15) suggests to re-write the Kolmogorov distance

$$d_K(Z, N) := \sup_{x\in\mathbb{R}}\big|P(Z \leq x) - P(N \leq x)\big|$$

between (the distributions of) Z and N as

$$d_K(Z,N) \le \sup_{f_x} \big|\mathbb{E}\big[f_x'(Z) - Zf_x(Z)\big]\big|, \tag{2.16}$$

where the supremum runs over the class of solutions f_x of (2.15). The unique bounded solution of (2.15) is of the form

$$f_x(z) = \mathrm{e}^{z^2/2}\int_{-\infty}^{z} \big(\mathbf{1}(y \le x) - P(N \le x)\big)\mathrm{e}^{-y^2/2}\,\mathrm{d}y,$$

see [7], Lemma 2.2, and satisfies the estimate $0 < f_x(z) \le \frac{\sqrt{2\pi}}{4}$. Moreover f_x is continuous on $\mathbb{R}$, infinitely differentiable on $\mathbb{R}\setminus\{x\}$, but not differentiable at x. However, interpreting the derivative of f_x at x as $1 - P(N \le x) + xf(x)$ in view of (2.15), we have

$$\big|f_x'(z)\big| \le 1 \quad \text{for all } z \in \mathbb{R} \tag{2.17}$$

from Lemma 2.3 in [7]. Moreover, the same result ensures that f_x satisfies

$$\big|(w+u)f_x(w+u) - (w+v)f_x(w+v)\big| \le \left(|w| + \frac{\sqrt{2\pi}}{4}\right)\big(|u|+|v|\big) \tag{2.18}$$

for all $u, v, w \in \mathbb{R}$.

2.7. *Stein's method for multivariate normal approximation*

There is also a multivariate version of Stein's method for normal approximation. It starts with the observation that a centred random variable $\mathbf{Z}$ with values in $\mathbb{R}^d$ for some $d \ge 2$ follows a multivariate normal distribution with covariance matrix C (which is a positive semi-definite $(d \times d)$-matrix) if and only if

$$\mathbb{E}\big[\langle \mathbf{Z}, \nabla f(\mathbf{Z})\rangle_{\mathbb{R}^d} - \langle C, \operatorname{Hess} f(\mathbf{Z})\rangle_{\mathrm{HS}}\big] = 0$$

for all twice differentiable $f : \mathbb{R}^d \to \mathbb{R}$ with

$$\mathbb{E}\big|\langle \mathbf{Z}, \nabla f(\mathbf{Z})\rangle_{\mathbb{R}^d}\big| + \mathbb{E}\big|\langle C, \operatorname{Hess} f(\mathbf{Z})\rangle_{\mathrm{HS}}\big| < \infty.$$

Here, $\langle\cdot,\cdot\rangle_{\mathbb{R}^d}$ is the inner product in $\mathbb{R}^d$, for two matrices A and B, $\langle A, B\rangle_{\mathrm{HS}} = \operatorname{trace}(AB^T)$ is the Hilbert–Schmidt inner product and $\operatorname{Hess} f(z)$ stands for the Hessian matrix of f at z. If $g : \mathbb{R}^d \to \mathbb{R}$, the multivariate Stein equation reads

$$\langle z, \nabla f(z)\rangle_{\mathbb{R}^d} - \langle C, \operatorname{Hess} f(z)\rangle_{\mathrm{HS}} = g(x) - \mathbb{E}\big[g(\mathbf{N})\big], \tag{2.19}$$

where $\mathbf{N}$ stands for a random variable with a multivariate centred normal distribution having covariance matrix C. It is well known that for a given function g,

$$f_g(x) = \int_0^1 \frac{1}{2t}\mathbb{E}\big[g(\sqrt{t}x + \sqrt{1-t}\mathbf{N}) - g(\mathbf{N})\big]\,\mathrm{d}t, \quad x \in \mathbb{R}^d,$$

is a solution of (2.19). To rephrase smoothness properties of f_g we introduce the following notation. If $h : \mathbb{R}^d \to \mathbb{R}$, $k \in \mathbb{N}$ and $i_1, \dots, i_k \in \{1, \dots, d\}$, put

$$M_k(h) := \max_{1 \le i_1, \dots, i_k \le d} \sup_{x \in \mathbb{R}^d} \left|\frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} h(x)\right|$$

(provided this is well defined). We notice that

$$\frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}} f_g(x) = \int_0^1 \frac{1}{2t} t^{k/2} \mathbb{E}\bigg[\frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}} g(\sqrt{t}x + \sqrt{1-t}\mathbf{N})\bigg]\,\mathrm{d}t,$$

whenever g possesses partial derivatives up to order k. In particular, this shows that $M_k(g) \le 1$ implies $M_k(f_g) \le 2/k$, see [7], Lemma 2.6.

To compare (the distributions of) the $\mathbb{R}^d$-valued random variables $\mathbf{Z}$ and $\mathbf{N}$, and inspired by (2.19), we use the d_4-distance

$$d_4(\mathbf{Z},\mathbf{N}) := \sup_g \big|\mathbb{E}g(\mathbf{Z}) - \mathbb{E}g(\mathbf{N})\big|, \tag{2.20}$$

where the supremum runs over all $g:\mathbb{R}^d \to \mathbb{R}$ having continuous partial derivatives up to order 4 and satisfy $M_i(g) \le 1$ for $i \in \{1,\dots,4\}$. Note that convergence in the d_4-distance implies convergence in distribution of the involved random variables.

3. A one-dimensional Berry–Esseen bound

We now present the first main result of this work, a Berry–Esseen bound for Rademacher functionals in terms of discrete Malliavin operators. From a structural point of view, this bound is very similar to that obtained in Theorem 3.1 of [10], which is not surprising as we also follow the basic idea of that paper. However, the proof and the interpretation of the involved Malliavin operators are different, because of the special structure of a Rademacher functional. This point will further be discussed in Remark 3.2 below.

Theorem 3.1. *Let $F \in \mathrm{dom}(D)$ with $\mathbb{E}[F] = 0$ and let N be a standard Gaussian random variable. Then*

$$\begin{aligned} d_K(F,N) \le{}& \mathbb{E}\big[\big|1 - \big\langle DF, -DL^{-1}F\big\rangle_{\ell^2(\mathbb{N})}\big|\big] + \frac{\sqrt{2\pi}}{4}\mathbb{E}\big[\big\langle (DF)^2, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &+ \mathbb{E}\big[\big\langle (DF)^2, \big|F\cdot DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] + 2\sup_{x\in\mathbb{R}} \mathbb{E}\big[\big\langle (DF) D\mathbb{1}_{\{F>x\}}, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big]. \end{aligned}$$

Remark 3.1. *Our result should be compared with the estimate from* [26]. *For a Rademacher functional $F \in \mathrm{dom}(D)$ with $\mathbb{E}[F] = 0$, define $B_1(F)$ and $B_2(F)$ by*

$$B_1(F) := \mathbb{E}\big[\big|1 - \big\langle DF, -DL^{-1}F\big\rangle_{\ell^2(\mathbb{N})}\big|\big],$$

$$B_2(F) := \frac{20}{3}\mathbb{E}\big[\big\langle |DF|^3, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big].$$

Then Theorem 3.1 *in* [26] *says that*

$$\big|\mathbb{E}\big[g(F)\big] - \mathbb{E}\big[g(N)\big]\big| \le \min\big\{4\|g'\|_\infty, \|g''\|_\infty\big\}B_1(F) + \|g''\|_\infty B_2(F), \tag{3.1}$$

where N is a standard Gaussian random variable and $g:\mathbb{R}\to\mathbb{R}$ is a twice differentiable function with bounded derivatives of order one and two. Moreover, using a standard smoothing argument, this has been extended in Corollary 3.6 *ibidem to an estimate for the Wasserstein distance*:

$$\begin{aligned} d_W(F,N) :={}& \sup_{g\in\mathrm{Lip}_1} \big|\mathbb{E}\big[g(F)\big] - \mathbb{E}\big[g(N)\big]\big| \\ \le{}& \sqrt{2\big(B_1(F) + B_2(F)\big)\big(5 + \mathbb{E}\big[|F|\big]\big)}, \end{aligned} \tag{3.2}$$

where the supremum runs over all Lipschitz functions $g:\mathbb{R}\to\mathbb{R}$ with Lipschitz constant bounded by 1. *In view of the well-known relation $d_K(F,N) \le 2\sqrt{d_W(F,N)}$ between the Kolmogorov and the Wasserstein distance, this leads in*

general to a suboptimal estimate for $d_K(F,N)$. Even (3.2) is suboptimal compared to (3.1). Our bound provided in Theorem 3.1 resolves this problem. Moreover, in our applications below we will see that the Kolmogorov distance will be of the same order of magnitude as (3.1), improving thereby also (3.2).

Proof of Theorem 3.1. Let f_x be the solution of the Stein equation (2.15). In view of (2.16) we have to bound the quantity

$$\mathbb{E}\big[f_x'(F) - Ff_x(F)\big]$$

uniformly in x. For the following calculation we suppress the dependence on x of the Stein solution and use the abbreviation $f := f_x$. Using the relation $\delta D = -L$ from (2.12) and the integration-by-parts formula in Lemma 2.5 we get

$$\begin{aligned}\mathbb{E}\big[f'(F) - Ff(F)\big] &= \mathbb{E}\big[f'(F) - \delta\big(-DL^{-1}F\big)f(F)\big] \\ &= \mathbb{E}\big[f'(F) - \big\langle Df(F), -DL^{-1}F\big\rangle_{\ell^2(\mathbb{N})}\big]. \end{aligned} \tag{3.3}$$

We now rewrite the scalar product on the right-hand side of (3.3). To this end, we find another representation of the gradient $Df(F)$, using the fundamental theorem of calculus. Note that this only makes use of the first derivative of f, which is in contrast to the approach taken in [26], where an approximate chain rule for $Df(F)$ is used and higher derivatives of f are involved. We get for $k \in \mathbb{N}$,

$$\begin{aligned} D_k f(F) &= \frac{1}{2}\big(\big(f(F)\big)_k^+ - \big(f(F)\big)_k^-\big) = \frac{1}{2}\big(\big(f\big(F_k^+\big)\big) - \big(f\big(F_k^-\big)\big)\big) \\ &= \frac{1}{2}\int_{F_k^- - F}^{F_k^+ - F} f'(F+t)\,\mathrm{d}t \\ &= \frac{1}{2}\bigg(\int_{F_k^- - F}^{F_k^+ - F}\big(f'(F+t) - f'(F)\big)\,\mathrm{d}t + \int_{F_k^- - F}^{F_k^+ - F} f'(F)\,\mathrm{d}t\bigg) \\ &= \frac{1}{2}\bigg(\int_{F_k^- - F}^{F_k^+ - F}\big(f'(F+t) - f'(F)\big)\,\mathrm{d}t + f'(F)\cdot\big(F_k^+ - F_k^-\big)\bigg). \end{aligned}$$

Combining this with (3.3) we get

$$\begin{aligned}\mathbb{E}\big[f'(F) - Ff(F)\big] &= \mathbb{E}\big[f'(F) - \big\langle f'(F)DF, -DL^{-1}F\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &\quad - \mathbb{E}\bigg[\bigg\langle\frac{1}{2}\int_{F_{(\cdot)}^- - F}^{F_{(\cdot)}^+ - F}\big(f'(F+t) - f'(F)\big)\,\mathrm{d}t, -DL^{-1}F\bigg\rangle_{\ell^2(\mathbb{N})}\bigg]. \end{aligned} \tag{3.4}$$

Since f is a solution of the Stein equation, we have for all $t \in \mathbb{R}$,

$$f'(F+t) = (F+t)f(F+t) + \mathbb{1}_{\{F+t\le x\}} - P(N \le x).$$

Thus,

$$\begin{aligned}\int_{F_k^- - F}^{F_k^+ - F}\big(f'(F+t) - f'(F)\big)\,\mathrm{d}t &= \underbrace{\int_{F_k^- - F}^{F_k^+ - F}\big((F+t)f(F+t) - Ff(F)\big)\,\mathrm{d}t}_{=:I_1(k)} \\ &\quad + \underbrace{\int_{F_k^- - F}^{F_k^+ - F}\big(\mathbb{1}_{\{F+t\le x\}} - \mathbb{1}_{\{F\le x\}}\big)\,\mathrm{d}t}_{=:I_2(k)}. \end{aligned} \tag{3.5}$$

One can now make use of the bound (2.18). Using the identities

$$F_k^{\pm} - F = \pm 2 D_k F \mathbb{1}_{\{X_k = \mp 1\}}, \tag{3.6}$$

we get

$$\begin{aligned}
\big|I_1(k)\big| &= \left| \int_{F_k^- - F}^{F_k^+ - F} \big((F+t) f(F+t) - F f(F)\big)\,\mathrm{d}t \right| \\
&\le \int_{\min\{F_k^- - F, F_k^+ - F\}}^{\max\{F_k^- - F, F_k^+ - F\}} \left(|F| + \frac{\sqrt{2\pi}}{4}\right) |t|\,\mathrm{d}t \\
&= \int_{2\min\{D_k F \mathbb{1}_{\{X_k=-1\}}, -D_k F \mathbb{1}_{\{X_k=1\}}\}}^{2\max\{D_k F \mathbb{1}_{\{X_k=-1\}}, -D_k F \mathbb{1}_{\{X_k=1\}}\}} \left(|F| + \frac{\sqrt{2\pi}}{4}\right) |t|\,\mathrm{d}t \\
&= \frac{(2 D_k F)^2}{2} \left(|F| + \frac{\sqrt{2\pi}}{4}\right).
\end{aligned} \tag{3.7}$$

For an evaluation of the term $I_2(k)$ we first define

$$\begin{aligned}
I_{+/+}(k) &= \mathbb{1}_{\{X_k=1,\, D_k F \ge 0\}} \cdot I_2(k), \qquad & I_{+/-}(k) &= \mathbb{1}_{\{X_k=1,\, D_k F < 0\}} \cdot I_2(k), \\
I_{-/+}(k) &= \mathbb{1}_{\{X_k=-1,\, D_k F \ge 0\}} \cdot I_2(k), \qquad & I_{-/-}(k) &= \mathbb{1}_{\{X_k=-1,\, D_k F < 0\}} \cdot I_2(k).
\end{aligned}$$

Using (3.6) we compute $I_{+/+}$ as follows:

$$\begin{aligned}
|I_{+/+}| &= \left| \mathbb{1}_{\{X_k=1,\, D_k F \ge 0\}} \cdot \int_{F_k^- - F}^{F_k^+ - F} (\mathbb{1}_{\{F+t \le x\}} - \mathbb{1}_{\{F \le x\}})\,\mathrm{d}t \right| \\
&= \mathbb{1}_{\{X_k=1,\, D_k F \ge 0\}} \cdot \left| \int_{-2 D_k F}^{0} (\underbrace{\mathbb{1}_{\{F+t \le x\}}}_{\le \mathbb{1}_{\{F - 2 D_k F \le x\}}} - \mathbb{1}_{\{F \le x\}})\,\mathrm{d}t \right| \\
&\le \mathbb{1}_{\{X_k=1,\, D_k F \ge 0\}} \cdot \big|2 D_k F \cdot (\mathbb{1}_{\{F - 2 D_k F \le x\}} - \mathbb{1}_{\{F \le x\}})\big| \\
&= \mathbb{1}_{\{X_k=1,\, D_k F \ge 0\}} \cdot \big|2 D_k F \cdot (\mathbb{1}_{\{F_k^- \le x\}} - \mathbb{1}_{\{F_k^+ \le x\}})\big| \\
&= \mathbb{1}_{\{X_k=1,\, D_k F \ge 0\}} \cdot \big|2 D_k F \cdot (\mathbb{1}_{\{F_k^+ > x\}} - \mathbb{1}_{\{F_k^- > x\}})\big| \\
&= \mathbb{1}_{\{X_k=1,\, D_k F \ge 0\}} \cdot 4 D_k F \cdot D_k \mathbb{1}_{\{F > x\}}.
\end{aligned}$$

Similarly, we get the bounds

$$\begin{aligned}
\big|I_{+/-}(k)\big| &\le \mathbb{1}_{\{X_k=1,\, D_k F < 0\}} \cdot 4 D_k F \cdot D_k \mathbb{1}_{\{F > x\}}, \\
\big|I_{-/+}(k)\big| &\le \mathbb{1}_{\{X_k=-1,\, D_k F \ge 0\}} \cdot 4 D_k F \cdot D_k \mathbb{1}_{\{F > x\}}, \\
\big|I_{-/-}(k)\big| &\le \mathbb{1}_{\{X_k=-1,\, D_k F < 0\}} \cdot 4 D_k F \cdot D_k \mathbb{1}_{\{F > x\}}.
\end{aligned}$$

Since

$$I_2(k) = I_{+/+}(k) + I_{+/-}(k) + I_{-/+}(k) + I_{-/+}(k),$$

we arrive at

$$\big|I_2(k)\big| \le 4 D_k F \cdot D_k \mathbb{1}_{\{F > x\}}. \tag{3.8}$$

Using our bounds (3.7) and (3.8) for $I_1 := (I_1(k))_{k\in\mathbb{N}}$ and $I_2 := (I_2(k))_{k\in\mathbb{N}}$ and the identities (3.4) and (3.5) we conclude that

$$\begin{aligned}\big|\mathbb{E}\big[f'(F) - Ff(F)\big]\big| &\le \mathbb{E}\big|1 - \big\langle DF, -DL^{-1}F\big\rangle_{\ell^2(\mathbb{N})}\big| + \frac{1}{2}\mathbb{E}\big[\big\langle |I_1| + |I_2|, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &\le \mathbb{E}\big|1 - \big\langle DF, -DL^{-1}F\big\rangle_{\ell^2(\mathbb{N})}\big| + \frac{\sqrt{2\pi}}{4}\mathbb{E}\big[\big\langle (DF)^2, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &\quad + \mathbb{E}\big[\big\langle (DF)^2, \big|F\cdot DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] + \mathbb{E}\big[\big\langle (2DF)\mathbb{1}_{\{F>x\}}, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big].\end{aligned}$$

The proof is completed by taking the supremum over all $x \in \mathbb{R}$. □

Let us finally introduce a simplification of the terms arising in Theorem 3.1, which will be used below.

Corollary 3.1. *Let $F \in \operatorname{dom}(D)$. Then*

(i) $\mathbb{E}[|1 - \langle DF, -DL^{-1}F\rangle_{\ell^2(\mathbb{N})}|] \le (\mathbb{E}[(1 - \langle DF, -DL^{-1}F\rangle_{\ell^2(\mathbb{N})})^2])^{1/2}$,

(ii) $\mathbb{E}[\langle (DF)^2, |F\cdot DL^{-1}F|\rangle_{\ell^2(\mathbb{N})}] + \frac{\sqrt{2\pi}}{4}\mathbb{E}[\langle (DF)^2, |DL^{-1}F|\rangle_{\ell^2(\mathbb{N})}] \le (\mathbb{E}[\langle (DF)^2, (DL^{-1}F)^2\rangle_{\ell^2(\mathbb{N})}])^{1/2} \times (\mathbb{E}[\|DF\|^4_{\ell^2(\mathbb{N})}])^{1/4}((\mathbb{E}[F^4])^{1/4} + 1)$,

provided all occurring expectations on the right-hand sides of the above inequalities are well defined.

Proof. An application of the Cauchy–Schwarz inequality yields (i). Using the Cauchy–Schwarz inequality again with respect to $\mathbb{E}[\cdot]$ and $\langle\cdot,\cdot\rangle_{\ell^2(\mathbb{N})}$ we find

$$\begin{aligned}&\frac{\sqrt{2\pi}}{4}\mathbb{E}\big[\big\langle (DF)^2, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] + \mathbb{E}\big[\big\langle (DF)^2, \big|F\cdot DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &\quad\le \mathbb{E}\big[\big\langle (DF)^2, \big(1+|F|\big)\big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &\quad= \mathbb{E}\big[\big\langle (DF)\big|DL^{-1}F\big|, (DF)\big(1+|F|\big)\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &\quad\le \mathbb{E}\big[\big\|(DF)\big(DL^{-1}F\big)\big\|_{\ell^2(\mathbb{N})}\cdot\big\|(DF)\big(1+|F|\big)\big\|_{\ell^2(\mathbb{N})}\big] \\ &\quad\le \big(\mathbb{E}\big[\big\langle (DF)^2, \big(DL^{-1}F\big)^2\big\rangle_{\ell^2(\mathbb{N})}\big]\big)^{1/2}\cdot\big(\mathbb{E}\big[\big(1+|F|\big)^2\|DF\|^2_{\ell^2(\mathbb{N})}\big]\big)^{1/2}. \qquad (3.9)\end{aligned}$$

Now, (ii) is obtained by applying both the Cauchy–Schwarz and the Minkowski inequality to the second term in (3.9):

$$\begin{aligned}\big(\mathbb{E}\big[\big(1+|F|\big)^2\|DF\|^2_{\ell^2(\mathbb{N})}\big]\big)^{1/2} &\le \mathbb{E}\big[\big(1+|F|\big)^4\big]^{1/4}\cdot\mathbb{E}\big[\|DF\|^4_{\ell^2(\mathbb{N})}\big]^{1/4} \\ &= \big\|1+|F|\big\|_{L^4(\Omega)}\cdot\mathbb{E}\big[\|DF\|^4_{\ell^2(\mathbb{N})}\big]^{1/4} \\ &\le \big(1+\mathbb{E}\big[F^4\big]^{1/4}\big)\cdot\mathbb{E}\big[\|DF\|^4_{\ell^2(\mathbb{N})}\big]^{1/4}.\end{aligned}$$

This completes the proof. □

Remark 3.2. *Let us briefly discuss the novelties of the proof of Theorem 3.1 compared with the existing literature (such as [26]). The common thread underlying the Malliavin–Stein approach (also on the Gaussian [21] or Poisson space [30]) is the usage of an integration-by-parts formula. For smooth test functions, this is then combined with a Taylor expansion, which leads to an appropriate chain rule. Here, we could not build on the existing so-called approximate chain rule from [26] and instead followed the idea of [10] by expressing the remainder term in integral form. Handling this term required new estimates, since the Malliavin operator D has a different representation and follows different computation rules in case of Rademacher sequences.*

4. Explicit bounds for discrete multiple stochastic integrals

4.1. *The first chaos*

In the present section we apply our abstract bound from Theorem 3.1 dealing with the Kolmogorov distance between the distribution of a general Rademacher functional $F \in \operatorname{dom}(D)$ and the standard normal distribution in the case that F belongs to the first Rademacher chaos. This way, we establish a connection to small ball probabilities. So, let $F = J_1(f)$ for some $f \in \ell^2(\mathbb{N})$, i.e., $F = \sum_{i=1}^{\infty} f(i)X_i$. Such functionals are known as Rademacher averages in the literature.

Theorem 4.1. *Let $F = \sum_{i=1}^{\infty} a_i X_i$ for $(a_i)_{i\in\mathbb{N}} \in \ell^2(\mathbb{N})$ such that $\mathbb{E}[F^2] = \sum_{i=1}^{\infty} a_i^2 = 1$ and let N be a standard Gaussian random variable. Then*

$$d_K(F,N) \le 2\sum_{i=1}^{\infty} |a_i|^3 + \sup_{x\in\mathbb{R}} \sum_{k=1}^{\infty} a_k^2 \cdot P\left(x - |a_k| < \sum_{\substack{i=1\\ i\neq k}}^{\infty} a_i X_i \le x + |a_k| \right). \tag{4.1}$$

Proof. We first introduce abbreviations for the four terms appearing on the right-hand side of the bound in Theorem 3.1:

$$A_1(F) := \mathbb{E}\big[\big|1 - \langle DF, -DL^{-1}F\rangle_{\ell^2\mathbb{N}}\big|\big], \tag{4.2}$$

$$A_2(F) := \frac{\sqrt{2\pi}}{4}\mathbb{E}\big[\big\langle (DF)^2, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big], \tag{4.3}$$

$$A_3(F) := \mathbb{E}\big[\big\langle (DF)^2, \big|F\cdot DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big], \tag{4.4}$$

$$A_4(F) := 2\sup_{x\in\mathbb{R}} \mathbb{E}\big[\big\langle (DF) D\mathbb{1}_{\{F>x\}}, \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big]. \tag{4.5}$$

In our case we have that $D_k F = -D_k L^{-1} F = a_k$ for all $k \in \mathbb{N}$ and thus get the following bounds for $A_1(F)$, $A_2(F)$ and $A_3(F)$:

$$\begin{aligned} A_1(F) &= \left|1 - \sum_{i=1}^{\infty} a_i^2\right| = 0, \qquad A_2(F) \le \sum_{i=1}^{\infty} |a_i|^3, \\ A_3(F) &= \mathbb{E}\left[\sum_{i=1}^{\infty} |a_i|^3 \cdot \left|\sum_{i=1}^{\infty} a_i X_i\right|\right] \le \left(\sum_{i=1}^{\infty} |a_i|^3\right) \cdot \sqrt{\mathbb{E}\left[\left(\sum_{i=1}^{\infty} a_i X_i\right)^2\right]} = \sum_{i=1}^{\infty} |a_i|^3. \end{aligned} \tag{4.6}$$

For the term $A_4(F)$ we first observe that

$$D_k \mathbb{1}\left(\sum_{i=1}^{\infty} a_i X_i > x\right) = \frac{1}{2}\left(\mathbb{1}\left(\sum_{\substack{i=1\\ i\neq k}}^{\infty} a_i X_i > x - a_k\right) - \mathbb{1}\left(\sum_{\substack{i=1\\ i\neq k}}^{\infty} a_i X_i > x + a_k\right)\right).$$

Thus,

$$\begin{aligned} A_4(F) &= 2\sup_{x\in\mathbb{R}} \mathbb{E}\left[\sum_{k=1}^{\infty} D_k F \cdot D_k \mathbb{1}(F>x)\big|D_k L^{-1} F\big|\right] \\ &= \sup_{x\in\mathbb{R}} \mathbb{E}\left[\sum_{k=1}^{\infty} a_k \cdot |a_k| \left(\mathbb{1}\left(\sum_{\substack{i=1\\ i\neq k}}^{\infty} a_i X_i > x - a_k\right) - \mathbb{1}\left(\sum_{\substack{i=1\\ i\neq k}}^{\infty} a_i X_i > x + a_k\right)\right)\right] \end{aligned}$$

$$= \sup_{x\in\mathbb{R}} \mathbb{E}\left[\sum_{k=1}^{\infty} a_k^2 \mathbb{1}\left(x - |a_k| < \sum_{\substack{i=1\\ i\neq k}}^{\infty} a_i X_i \le x + |a_k|\right)\right]$$

$$= \sup_{x\in\mathbb{R}} \sum_{k=1}^{\infty} a_k^2 \cdot P\left(x - |a_k| < \sum_{\substack{i=1\\ i\neq k}}^{\infty} a_i X_i \le x + |a_k|\right). \tag{4.7}$$

Putting together (4.6) and (4.7) yields the assertion. □

Remark 4.1. *It is interesting to see that the bound* (4.1) *in Theorem* 4.1 *involves quantities which are known as small ball probabilities in the literature. More precisely, if* $\xi_1, \dots, \xi_n$ *are i.i.d. real-valued random variables and* $a_1, \dots, a_n$ *real numbers, then a quantity of the type*

$$\sup_{x\in\mathbb{R}} P\left(\left|\sum_{i=1}^{n} a_i \xi_i - x\right| \le \varepsilon\right), \quad \varepsilon > 0,$$

is what is usually called a small ball probability and can be considered as a kind of measure of anti-concentration for the partial sum $\sum_{i=1}^{n} a_i \xi_i$. *The authors in* [39] (*see also* [17]) *found a bound for these small ball probabilities using the classical Berry–Esseen theorem for i.i.d. random variables with finite third moment. In our set-up, i.e., if* $X_1, \dots, X_n$ *is a sequence of independent Rademacher random variables and if* $a_1, \dots, a_n \in \mathbb{R}$, *Corollary* 2.9 *in* [39] *says that*

$$\sup_{x\in\mathbb{R}} P\left(\left|\sum_{i=1}^{n} a_i X_i - x\right| \le \varepsilon\right) \le \sqrt{\frac{2}{\pi}}\varepsilon + C\sum_{i=1}^{n} |a_i|^3, \quad \varepsilon > 0, \tag{4.8}$$

where $0 < C < \infty$ *is an absolute constant. Using* (4.8) *one can see that Theorem* 4.1 *reproduces the correct order for the Berry–Esseen bound for the normal approximation of a finite sum* $F = \sum_{i=1}^{n} a_i X_i$, *which is* $\mathrm{O}(\sum_{i=1}^{n} |a_i|^3)$. *Our Theorem* 4.1 *can be interpreted as an inverse of* (4.8), *as it provides a Berry–Esseen bound for* F *in terms of small ball probabilities. Also note that our Theorem* 4.1 *goes beyond this set-up, since it allows* F *to depend on an infinite sequence of independent Rademacher variables.*

4.2. *The case* $q \ge 2$

Our main result in this section is an estimate for the Kolmogorov distance of a discrete multiple stochastic integral of arbitrary order $q \ge 2$ and a standard Gaussian random variable in terms of contraction norms. We present two estimates, which will separately be used below. We emphasize that they are the discrete analogues of similar results for Gaussian or Poisson multiple stochastic integrals, see [10,21].

Theorem 4.2. *Let* $F = J_q(f)$ *for a fixed integer* $q \ge 2$ *and a symmetric kernel* $f \in \ell_0^2(\mathbb{N})^{\circ q}$. *Assume that* $\|(f \star_r^r f)\mathbb{1}_{\Delta_{2(q-r)}}\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} \le 1$, *for all* $r = 1, \dots, q-1$. *Furthermore, let* N *be a standard Gaussian random variable. Then*

$$d_K(F,N) \le C_1 \max\Big\{\big|1 - q!\|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big|, \max_{r=1,\dots,q-1}\Big\{\big\|(f \star_r^r f)\mathbb{1}_{\Delta_{2(q-r)}}\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\Big\},$$

$$\max_{r=1,\dots,q}\Big\{\big\|f \star_r^{r-1} f\big\|_{\ell^2(\mathbb{N})^{\otimes(2(q-r)+1)}}\Big\}\Big\}$$

$$\le C_2 \max\Big\{\big|1 - q!\|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big|, \max_{r=1,\dots,q-1}\Big\{\big\|f \star_r^r f\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\Big\}\Big\}$$

with universal constants $0 < C_1, C_2 < \infty$ *only depending on* q.

Remark 4.2. *Theorem 4.2 only proves useful for applications to sequences $F_n = J_q(f_n)$ with kernels f_n for which at least one of the above bounds vanishes. In such applications, the assumption $\|(f_n \star_r^r f_n)\mathbb{1}_{\Delta_{2(q-r)}}\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} \leq 1$, for all $r = 1, \dots, q-1$, is naturally fullfilled, whenever n is large enough.*

We prepare the proof of Theorem 4.2 with the following lemma, which corresponds to Theorem 4.1 in [26] combined with (2.2) and (2.3).

Lemma 4.1. *Let $F = J_q(f)$ for a fixed integer $q \geq 2$ and a symmetric kernel $f \in \ell_0^2(\mathbb{N})^{\circ q}$. Then*

$$\begin{aligned}
&\mathbb{E}\left[\left(1 - \frac{1}{q}\|DF\|_{\ell^2(\mathbb{N})}^2\right)^2\right] \\
&\quad = \left(1 - q!\|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^2\right)^2 \\
&\qquad + q^2 \sum_{r=1}^{q-1}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 (2(q-r))!\left\|\widetilde{(f \star_r^r f)}\mathbb{1}_{\Delta_{2(q-r)}}\right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}^2 \\
&\quad \leq \left(1 - q!\|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^2\right)^2 \\
&\qquad + q^2 \sum_{r=1}^{q-1}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 (2(q-r))!\left\|f \star_r^r f\right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}^2
\end{aligned} \tag{4.9}$$

and

$$\begin{aligned}
\mathbb{E}\left[\|DF\|_{\ell^4(\mathbb{N})}^4\right] &= \sum_{k=1}^{\infty} q^4 \sum_{r=1}^{q}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 (2(q-r))! \\
&\qquad \times \left\|\widetilde{\left(f(\cdot,k) \star_{r-1}^{r-1} f(\cdot,k)\right)}\mathbb{1}_{\Delta_{2(q-r)}}\right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}^2 \\
&\leq q^4 \sum_{r=1}^{q}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 (2(q-r))!\left\|f \star_r^{r-1} f\right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)+1}}^2 \\
&\leq C \max_{r=1,\dots,q-1}\left\{\left\|f \star_r^r f\right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}^2\right\},
\end{aligned} \tag{4.10}$$

for a constant $0 < C < \infty$ depending only on q.

Proof of Theorem 4.2. Let us first assume that the support of f satisfies

$$\operatorname{supp}(f) \subseteq \{1, \dots, n\}^q$$

for some $n \in \mathbb{N}$. According to Theorem 3.1 and Corollary 3.1, we have to bound the following quantities:

$$A_1(F) := \left(\mathbb{E}\left[\left(1 - \langle DF, -DL^{-1}F\rangle_{\ell^2(\mathbb{N})}\right)^2\right]\right)^{1/2}, \tag{4.11}$$

$$A_2(F) := \left(\mathbb{E}\left[\left\langle (DF)^2, \left(DL^{-1}F\right)^2\right\rangle_{\ell^2(\mathbb{N})}\right]\right)^{1/2}, \tag{4.12}$$

$$A_3(F) := \left(\mathbb{E}\left[\|DF\|_{\ell^2(\mathbb{N})}^4\right]\right)^{1/4}\left(\left(\mathbb{E}\left[F^4\right]\right)^{1/4} + 1\right), \tag{4.13}$$

$$A_4(F) := 2\sup_{x\in\mathbb{R}} \mathbb{E}\left[\left\langle (DF)(D\mathbb{1}_{\{F>x\}}), \left|DL^{-1}F\right|\right\rangle_{\ell^2(\mathbb{N})}\right]. \tag{4.14}$$

First observe that, since $D_k F = q\, J_{q-1}(f(\cdot,k))$ and $-D_k L^{-1} F = J_{q-1}(f(\cdot,k)) = \frac{1}{q} D_k F$, we have

$$A_1(F) = \left(\mathbb{E}\left[\left(1 - \frac{1}{q}\|DF\|^2_{\ell^2(\mathbb{N})}\right)^2\right]\right)^{1/2} \quad \text{and} \quad A_2(F) = \left(\mathbb{E}\left[\frac{1}{q^2}\|DF\|^4_{\ell^4(\mathbb{N})}\right]\right)^{1/2}.$$

Thus, using (2.2), (2.3) and Lemma 4.1, we can estimate $A_1(F)$ and $A_2(F)$ as follows:

$$\begin{aligned} A_1(F) &\le \left|1 - q!\|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\right| \\ &\quad + q\left(\sum_{r=1}^{q-1}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 \big(2(q-r)\big)!\big\|\big(f \star_r^r f\big)\mathbb{1}_{\Delta_{2(q-r)}}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\right)^{1/2} \qquad (4.15) \\ &\le \left|1 - q!\|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\right| \\ &\quad + q\left(\sum_{r=1}^{q-1}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 \big(2(q-r)\big)!\big\|f \star_r^r f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\right)^{1/2} \qquad (4.16) \end{aligned}$$

and

$$\begin{aligned} A_2(F) &\le q\left(\sum_{r=1}^{q}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 \big(2(q-r)\big)!\big\|f \star_r^{r-1} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)+1}}\right)^{1/2} \qquad (4.17) \\ &= q\Bigg(\big(2(q-1)\big)!\big\|f \star_1^0 f\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-1)+1}} \\ &\quad + \sum_{r=2}^{q}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 \big(2(q-r)\big)!\big\|f \star_r^{r-1} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)+1}}\Bigg)^{1/2} \\ &\le q\Bigg(\big(2(q-1)\big)!\big\|f \star_{q-1}^{q-1} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} \\ &\quad + \sum_{r=2}^{q}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 \big(2(q-r)\big)!\big\|f \star_{r-1}^{r-1} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r+1)}}\Bigg)^{1/2} \\ &= q\Bigg(\big(2(q-1)\big)!\big\|f \star_{q-1}^{q-1} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} \\ &\quad + \sum_{r=1}^{q-1}\left(r!\binom{q-1}{r}^2\right)^2 \big(2(q-r-1)\big)!\big\|f \star_r^r f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\Bigg)^{1/2}. \qquad (4.18) \end{aligned}$$

Considering $A_3(F)$, we use the multiplication formula (2.6) to see that

$$\begin{aligned} (D_k F)^2 &= q^2 \sum_{r=0}^{q-1} r!\binom{q-1}{r}^2 J_{2(q-r-1)}\big(\widetilde{\big(f(\cdot,k) \star_r^r f(\cdot,k)\big)}\mathbb{1}_{\Delta_{2(q-r-1)}}\big) \\ &= q^2 \sum_{r=1}^{q}(r-1)!\binom{q-1}{r-1}^2 J_{2(q-r)}\big(\widetilde{\big(f(\cdot,k) \star_{r-1}^{r-1} f(\cdot,k)\big)}\mathbb{1}_{\Delta_{2(q-r)}}\big). \end{aligned}$$

Since f has finite support, we thus find that

$$\begin{aligned}\|DF\|^2_{\ell^2(\mathbb{N})} &= \sum_{k=1}^{\infty} q^2 \sum_{r=1}^{q} (r-1)! \binom{q-1}{r-1}^2 J_{2(q-r)}\big(\widetilde{(f(\cdot,k) \star_{r-1}^{r-1} f(\cdot,k))}\mathbb{1}_{\Delta_{2(q-r)}}\big) \\ &= q^2 \sum_{r=1}^{q} (r-1)! \binom{q-1}{r-1}^2 J_{2(q-r)}\big(\widetilde{(f \star_r^r f)}\mathbb{1}_{\Delta_{2(q-r)}}\big) \\ &= q^2 \sum_{r=1}^{q-1} (r-1)! \binom{q-1}{r-1}^2 J_{2(q-r)}\big(\widetilde{(f \star_r^r f)}\mathbb{1}_{\Delta_{2(q-r)}}\big) + q \cdot q! \|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}.\end{aligned}$$

Using the isometry relation (2.5) we deduce the bound

$$\begin{aligned}&\big(\mathbb{E}\big[\|DF\|^4_{\ell^2(\mathbb{N})}\big]\big)^{1/4} \\ &\quad\le \big(q \cdot q! \|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big)^{1/2} \\ &\qquad + q\left(\sum_{r=1}^{q-1}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 (2(q-r))! \big\|(f \star_r^r f)\mathbb{1}_{\Delta_{2(q-r)}}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\right)^{1/4} \qquad (4.19) \\ &\quad\le q\left(\sum_{r=1}^{q-1}\left((r-1)!\binom{q-1}{r-1}^2\right)^2 (2(q-r))! \big\|f \star_r^r f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\right)^{1/4} + \big(q \cdot q! \|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big)^{1/2} \qquad (4.20)\end{aligned}$$

for the first factor in $A_3(F)$. For the second factor we use again the multiplication formula (2.6) to see that

$$F^2 = \sum_{r=0}^{q-1} r! \binom{q}{r}^2 J_{2(q-r)}\big(\widetilde{(f \star_r^r f)}\mathbb{1}_{\Delta_{2(q-r)}}\big) + q! \|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}.$$

The isometry relation (2.5) then shows that $\mathbb{E}[F^4]$ can be expressed as

$$\mathbb{E}\big[F^4\big] = \sum_{r=0}^{q-1}\left(r!\binom{q}{r}^2\right)^2 (2(q-r))! \big\|\widetilde{(f \star_r^r f)}\mathbb{1}_{\Delta_{2(q-r)}}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} + \big(q! \|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big)^2.$$

Separating the term $r=0$ and applying (2.1) we find that $\mathbb{E}[F^4]$ is bounded from above by

$$\begin{aligned}&3\big(q! \|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big)^2 + \sum_{r=1}^{q-1}\left(\left(q!\binom{q}{r}\right)^2 + \left(r!\binom{q}{r}^2\right)^2 (2(q-r))!\right)\big\|(f \star_r^r f)\mathbb{1}_{\Delta_{2(q-r)}}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} \qquad (4.21) \\ &\quad\le 3\big(q! \|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big)^2 + \sum_{r=1}^{q-1}\left(q!\binom{q}{r}\right)^2\left(1 + \binom{2(q-r)}{q-r}\right)\big\|f \star_r^r f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}. \qquad (4.22)\end{aligned}$$

Next, we consider the term $A_4(F)$. By virtue of Lemma 2.5, we have

$$\begin{aligned}A_4(F) &= 2 \sup_{x\in\mathbb{R}} \mathbb{E}\big[\big\langle (DF)(D\mathbb{1}_{\{F>x\}}), \big|DL^{-1}F\big|\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &= \frac{2}{q} \sup_{x\in\mathbb{R}} \mathbb{E}\big[\big\langle D\mathbb{1}_{\{F>x\}}, (DF)|DF|\big\rangle_{\ell^2(\mathbb{N})}\big] \\ &= \frac{2}{q} \sup_{x\in\mathbb{R}} \mathbb{E}\big[\mathbb{1}_{\{F>x\}}\delta\big((DF)|DF|\big)\big]\end{aligned}$$

$$\leq \frac{2}{q}\mathbb{E}\big[\big|\delta\big((DF)|DF|\big)\big|\big]$$
$$\leq \frac{2}{q}\big(\mathbb{E}\big[\big(\delta\big((DF)|DF|\big)\big)^2\big]\big)^{1/2}. \tag{4.23}$$

We now apply the isometry property (2.14) of the divergence operator. This leads to

$$\begin{aligned}
&\mathbb{E}\big[\big(\delta\big((DF)|DF|\big)\big)^2\big]\\
&\quad= \mathbb{E}\big[\big\|(DF)|DF|\big\|^2_{\ell^2(\mathbb{N})}\big] + \mathbb{E}\bigg[\sum_{k,l=1}^{\infty}\big[D_k\big(D_lF|D_lF|\big)\big]\big[D_l\big(D_kF|D_kF|\big)\big]\bigg]\\
&\quad= \mathbb{E}\big[\|DF\|^4_{\ell^4(\mathbb{N})}\big] + \mathbb{E}\bigg[\sum_{k,l=1}^{\infty}\big[D_k\big(D_lF|D_lF|\big)\big]\big[D_l\big(D_kF|D_kF|\big)\big]\bigg]\\
&\quad\leq \mathbb{E}\big[\|DF\|^4_{\ell^4(\mathbb{N})}\big] + \mathbb{E}\bigg[\sum_{k,l=1}^{\infty}\big(D_k\big(D_lF|D_lF|\big)\big)^2\bigg].
\end{aligned} \tag{4.24}$$

Now, using both the product formula (2.9) for the gradient operator and (2.8) we see that

$$\begin{aligned}
&\mathbb{E}\big[\big(D_k\big(D_lF|D_lF|\big)\big)^2\big]\\
&\quad= \mathbb{E}\big[\big(D_k\big((D_lF)^2\big)\big)^2\mathbb{1}_{\{D_lF\geq 0\}} + \big(D_k\big(-(D_lF)^2\big)\big)^2\mathbb{1}_{\{D_lF<0\}}\big]\\
&\quad= \mathbb{E}\big[\big(D_k\big((D_lF)^2\big)\big)^2\big]\\
&\quad= \mathbb{E}\big[\big(2(D_lF)(D_kD_lF) - 2X_k(D_kD_lF)^2\big)^2\big]\\
&\quad= 4\mathbb{E}\big[(D_lF)^2(D_kD_lF)^2\big] - 8\mathbb{E}\big[X_k(D_lF)(D_kD_lF)^3\big] + 4\mathbb{E}\big[(D_kD_lF)^4\big]\\
&\quad= 4\mathbb{E}\big[(D_lF)^2(D_kD_lF)^2\big] - 8\mathbb{E}\big[D_k\big((D_lF)(D_kD_lF)^3\big)\big] + 4\mathbb{E}\big[(D_kD_lF)^4\big]\\
&\quad= 4\mathbb{E}\big[(D_lF)^2(D_kD_lF)^2\big] - 4\mathbb{E}\big[(D_kD_lF)^4\big]\\
&\quad\leq 4\mathbb{E}\big[(D_lF)^2(D_kD_lF)^2\big].
\end{aligned} \tag{4.25}$$

Combining (4.24) and (4.25) we further estimate

$$\begin{aligned}
&\mathbb{E}\big[\big(\delta\big((DF)|DF|\big)\big)^2\big]\\
&\quad\leq \mathbb{E}\big[\|DF\|^4_{\ell^4(\mathbb{N})}\big] + 4\mathbb{E}\bigg[\sum_{l=1}^{\infty}\bigg((D_lF)^2\sum_{k=1}^{\infty}(D_kD_lF)^2\bigg)\bigg]\\
&\quad\leq \mathbb{E}\big[\|DF\|^4_{\ell^4(\mathbb{N})}\big] + 4\mathbb{E}\bigg[\|DF\|^2_{\ell^4(\mathbb{N})}\bigg(\sum_{l=1}^{\infty}\bigg(\sum_{k=1}^{\infty}(D_kD_lF)^2\bigg)^2\bigg)^{1/2}\bigg]\\
&\quad\leq \mathbb{E}\big[\|DF\|^4_{\ell^4(\mathbb{N})}\big] + 4\big(\mathbb{E}\big[\|DF\|^4_{\ell^4(\mathbb{N})}\big]\big)^{1/2}\bigg(\mathbb{E}\bigg[\sum_{l=1}^{\infty}\bigg(\sum_{k=1}^{\infty}(D_kD_lF)^2\bigg)^2\bigg]\bigg)^{1/2}.
\end{aligned}$$

Putting this into (4.23), we arrive at

$$A_4(F) \le 2A_2(F) + 4\Big(\frac{1}{q}A_2(F)\Big)^{1/2} \underbrace{\left(\mathbb{E}\left[\sum_{l=1}^{\infty}\left(\sum_{k=1}^{\infty}(D_kD_lF)^2\right)^2\right]\right)^{1/4}}_{=:A_4'(F)}. \tag{4.26}$$

To bound $A_4'(F)$ further we notice that $D_kD_lF = q(q-1)J_{q-2}(f(\cdot,k,l))$, which implies

$$\begin{aligned}
&\sum_{k=1}^{\infty}(D_kD_lF)^2\\
&\quad= \sum_{k=1}^{\infty} q^2(q-1)^2 \sum_{r=0}^{q-2} r! \binom{q-2}{r}^2 J_{2(q-r-2)}\big(\big(f(\cdot,k,l)\,\widetilde{\star_r^r}\, f(\cdot,k,l)\big)\mathbb{1}_{\Delta_{2(q-r-2)}}\big)\\
&\quad= \sum_{k=1}^{\infty} q^2(q-1)^2 \sum_{r=2}^{q} (r-2)! \binom{q-2}{r-2}^2 J_{2(q-r)}\big(\big(f(\cdot,k,l)\,\widetilde{\star_{r-2}^{r-2}}\, f(\cdot,k,l)\big)\mathbb{1}_{\Delta_{2(q-r)}}\big)\\
&\quad= q^2(q-1)^2 \sum_{r=2}^{q} (r-2)! \binom{q-2}{r-2}^2 J_{2(q-r)}\big(\big(f(\cdot,l)\,\widetilde{\star_{r-1}^{r-1}}\, f(\cdot,l)\big)\mathbb{1}_{\Delta_{2(q-r)}}\big).
\end{aligned}$$

Then, by the isometry of discrete multiple stochastic integrals (2.5) and the contraction inequality (2.3), we see that

$$\begin{aligned}
&\sum_{l=1}^{\infty}\mathbb{E}\left[\left(\sum_{k=1}^{\infty}(D_kD_lF)^2\right)^2\right]\\
&\quad= \sum_{l=1}^{\infty} q^4(q-1)^4 \sum_{r=2}^{q}\left((r-2)!\binom{q-2}{r-2}^2\right)^2 \big(2(q-r)\big)!\big\|\big(f(\cdot,l)\,\widetilde{\star_{r-1}^{r-1}}\, f(\cdot,l)\big)\mathbb{1}_{\Delta_{2(q-r)}}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\\
&\quad\le \sum_{l=1}^{\infty} q^4(q-1)^4 \sum_{r=2}^{q}\left((r-2)!\binom{q-2}{r-2}^2\right)^2 \big(2(q-r)\big)!\big\|f(\cdot,l)\star_{r-1}^{r-1} f(\cdot,l)\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\\
&\quad= q^4(q-1)^4 \sum_{r=2}^{q}\left((r-2)!\binom{q-2}{r-2}^2\right)^2 \big(2(q-r)\big)!\big\|f\star_r^{r-1} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)+1}}\\
&\quad\le q^4(q-1)^4 \sum_{r=2}^{q}\left((r-2)!\binom{q-2}{r-2}^2\right)^2 \big(2(q-r)\big)!\big\|f\star_{r-1}^{r-1} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r+1)}}\\
&\quad= q^4(q-1)^4 \sum_{r=1}^{q-1}\left((r-1)!\binom{q-2}{r-1}^2\right)^2 \big(2(q-r-1)\big)!\big\|f\star_r^{r} f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}.
\end{aligned} \tag{4.27}$$

Thus,

$$A_4'(F) \le q(q-1)\left(\sum_{r=1}^{q-1}\left((r-1)!\binom{q-2}{r-1}^2\right)^2 \big(2(q-r-1)\big)!\big\|f\star_r^r f\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\right)^{1/4}. \tag{4.28}$$

Combining (4.15), (4.17), (4.19), (4.21), (4.26) and (4.27) yields the first inequality in Theorem 4.2, while (4.16), (4.18), (4.20), (4.22), (4.26) and (4.28) give the second bound in Theorem 4.2 for a discrete multiple stochastic integral $J_q(f)$ whose integrand f satisfies $\operatorname{supp}(f) \subseteq \{1,\dots,n\}^q$. For the general case we use the following approximation argument (cf. [35]). Consider for $n \in \mathbb{N}$ the sequence of truncated kernels $f_n := f\mathbb{1}_{\{1,\dots,n\}^q}$. Since the sequence

$(J_q(f_n))_{n\geq 1}$ is a martingale with respect to the filtration $(\mathcal{F}_n)_{n\geq 1}$ with $\mathcal{F}_n := \sigma(X_1, \ldots, X_n)$, an application of the martingale convergence theorem yields that

$$\lim_{n\to\infty} J_q(f\mathbb{1}_{\{1,\ldots,n\}^q}) = \lim_{n\to\infty} \mathbb{E}\big[J_q(f)|\mathcal{F}_n\big] = J_q(f).$$

The assertion thus follows by means of [35], Lemma 2.6, and continuity of the gradient operator on a fixed Rademacher chaos. □

Remark 4.3. *If in Theorem 4.2 a general centred Gaussian random variable N_{σ^2} with variance $\sigma^2 > 0$ is used, the Berry–Esseen bound has to be replaced by*

$$d_K(F, N_{\sigma^2}) \leq C\frac{1}{\sigma^2}\max\Big\{\big|\sigma^2 - q!\|f\|^2_{\ell^2(\mathbb{N})^{\otimes q}}\big|, \max_{r=1,\ldots,q-1}\big\{\big\|f\star_r^r f\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\big\}\Big\}.$$

This is easily verified by a re-scaling argument.

4.3. *A necessary condition for double integrals*

Theorem 4.2 says that $J_q(f_n)$ converges in distribution to a standard Gaussian random variable if

$$\lim_{n\to\infty}\max_{r=1,\ldots,q-1}\big\{\big\|f_n\star_r^r f_n\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\big\} = 0,$$

provided that $q!\|f_n\| \to 1$, as $n\to\infty$. In view of the results from [10,14,21] implying that vanishing contraction norms yield a central limit theorem for Gaussian or Poisson multiple integrals (in the latter case at least if the functions f_n are non-negative), it is natural to ask whether this is also a necessary condition for a sequence of discrete multiple integrals to satisfy a central limit theorem. Here, we concentrate on the case of double integrals $F_n := J_2(f_n)$ and recall that it has been claimed in Proposition 4.6 of [26] that

$$\lim_{n\to\infty}\big\|f_n\star_1^1 f_n\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} = 0$$

is a necessary and sufficient condition for weak convergence of the distribution of F_n to the standard normal distribution, if $\mathbb{E}[F_n^2]\to 1$, as $n\to\infty$. The proof of sufficiency of this condition for asymptotic normality of (quite general) quadratic forms goes back to the classical work [8] of de Jong. Another paper in this regard is the paper [6] of Chatterjee, where in Proposition 3.1 ibidem a bound for the normal approximation of quadratic forms of Rademacher functionals has been obtained by means of Stein's method. As shown in [26], these bounds also have a representation in terms of norms of contractions. However, it turns out that the convergence of these norms to zero is not neccessary for asymptotic normality as the following example shows (see Section 1.6 in [25]).

Example 4.1. *Consider $F_n = J_2(f_n)$, $n\geq 2$, with*

$$f_n(i,j) = \begin{cases} \frac{1}{2\sqrt{n-1}}, & \text{if } \{i,j\} = \{1,k\} \text{ for some } k = 2,\ldots,n,\\ 0, & \text{otherwise.}\end{cases}$$

Then,

$$F_n = \sum_{i,j=1}^n f_n(i,j)X_iX_j = \frac{X_1}{\sqrt{n-1}}\sum_{i=2}^n X_i$$

with $\mathbb{E}[F_n^2] = 1$. We notice that the distribution of F_n converges weakly to the standard normal distribution. But

$$\begin{aligned}\big(f_n\star_1^1 f_n\big)(i,j) &= \sum_{k=2}^n f_n^2(1,k)\mathbb{1}_{\{i=j=1\}}(i,j) + f_n(1,i)f_n(1,j)\mathbb{1}_{\{i\geq 2, j\geq 2\}}(i,j)\\ &= \frac{1}{4}\mathbb{1}_{\{i=j=1\}}(i,j) + \frac{1}{4(n-1)}\mathbb{1}_{\{i\geq 2, j\geq 2\}}(i,j)\end{aligned}$$

and hence

$$\|f_n \star_1^1 f_n\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} = \sum_{i,j=1}^{n} \left(\frac{1}{16} \mathbb{1}_{\{i=j=1\}}(i,j) + \frac{1}{16(n-1)^2} \mathbb{1}_{\{i\geq 2, j\geq 2\}}(i,j) \right) = \frac{1}{8}.$$

We now deduce a new necessary condition for a sequence $J_2(f_n)$ of double integrals to converge in distribution to a standard Gaussian random variable. It shows that the validity of such a central limit theorem depends in general on a subtle interplay of contraction norms on and off diagonals. Such a phenomenon is not visible for Gaussian or Poisson multiple integrals and seems to be a special feature of the discrete set-up.

Theorem 4.3. *Let $F_n = J_2(f_n)$, $f_n \in \ell_0^2(\mathbb{N})^{\circ 2}$ and assume that $\mathbb{E}[F_n^2] = 1$ for all $n \in \mathbb{N}$. A necessary condition for the convergence of the distribution of F_n to the standard normal distribution is that*

$$2\|f_n\|^4_{\ell^4(\mathbb{N})^{\otimes 2}} + 3\Big(\big\| (f_n \star_1^1 f_n) \mathbb{1}_{\Delta_2} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} - \big\| (f_n \star_1^1 f_n) \mathbb{1}_{\Delta_2^c} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} \Big) \longrightarrow 0, \tag{4.29}$$

as $n \to \infty$.

It is readily checked that our new necessary condition (4.29) is satisfied for Example 4.1.

The proof of Theorem 4.3 is prepared by the following two lemmas. The first one is a hypercontractivity property of Rademacher functionals with a finite chaotic decomposition and the second one provides an expression for the 4th moment of a discrete stochastic double integral.

Lemma 4.2. *Let $F = \mathbb{E}[F] + \sum_{n=1}^{d} J_n(f_n)$, $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$, for some $d \in \mathbb{N}$, and suppose that $2 \leq p < q$. Then there exists a constant $0 < C < \infty$ depending only on d such that*

$$\big(\mathbb{E}\big[|F|^q\big]\big)^{1/q} \leq C \left(\frac{q-1}{p-1} \right)^{d/2} \big(\mathbb{E}\big[|F|^p\big]\big)^{1/p}. \tag{4.30}$$

Proof. For F depending only on finitely many Rademacher variables this is Theorem 3.2.5 in [9]. In the general case, the assertion follows by means of an approximation argument and Corollary 0.2.1 in [13]. □

Lemma 4.3. *Let $F = J_2(f)$ with $f \in \ell_0^2(\mathbb{N})^{\circ 2}$ such that $\mathbb{E}[F^2] = 1$. Then,*

$$\mathbb{E}\big[F^4\big] = 3 + 32\|f\|^4_{\ell^4(\mathbb{N})^{\otimes 2}} + 48\Big(\big\| (f \star_1^1 f) \mathbb{1}_{\Delta_2} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} - \big\| (f \star_1^1 f) \mathbb{1}_{\Delta_2^c} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} \Big). \tag{4.31}$$

Proof. From the multiplication formula (2.6), we have

$$F^2 = J_4\big(\widetilde{(f \star_0^0 f)} \mathbb{1}_{\Delta_4} \big) + 4J_2\big((f \star_1^1 f) \mathbb{1}_{\Delta_2} \big) + 2\|f\|^2_{\ell^2(\mathbb{N})^{\otimes 2}}.$$

Now the isometry of multiple stochastic integrals (2.5) and relation (2.1) yield

$$\begin{aligned}
\mathbb{E}\big[F^4\big] &= 24\big\| \widetilde{(f \star_0^0 f)} \mathbb{1}_{\Delta_4} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 4}} + 32\big\| (f \star_1^1 f) \mathbb{1}_{\Delta_2} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} + 4\|f\|^4_{\ell^2(\mathbb{N})^{\otimes 2}} \\
&= 24\big\| \widetilde{f \star_0^0 f} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 4}} - 24\big\| \widetilde{(f \star_0^0 f)} \mathbb{1}_{\Delta_4^c} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 4}} + 32\big\| (f \star_1^1 f) \mathbb{1}_{\Delta_2} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} \\
&\quad + 4\|f\|^4_{\ell^2(\mathbb{N})^{\otimes 2}} \\
&= 12\|f\|^4_{\ell^2(\mathbb{N})^{\otimes 2}} + 48\big\| (f \star_1^1 f) \mathbb{1}_{\Delta_2} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} + 16\big\| (f \star_1^1 f) \mathbb{1}_{\Delta_2^c} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} \\
&\quad - 24\big\| \widetilde{(f \star_0^0 f)} \mathbb{1}_{\Delta_4^c} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 4}}.
\end{aligned} \tag{4.32}$$

By definition we have

$$\big\|\widetilde{(f\star_0^0 f)}\mathbb{1}_{\Delta_4^c}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 4}} = \sum_{i_1,i_2,i_3,i_4} \big(\widetilde{(f\star_0^0 f)}(i_1,i_2,i_3,i_4)\big)^2 \mathbb{1}_{\Delta_4^c}(i_1,i_2,i_3,i_4),$$

with the indices running over all non-negative integers. Since f is symmetric and vanishes on diagonals, there are only two types of 4-tuples (i_1,i_2,i_3,i_4) for which $\widetilde{(f\star_0^0 f)}(i_1,i_2,i_3,i_4)$ can be non-zero. They are of the form (i_1,i_1,i_2,i_3) and (i_1,i_1,i_2,i_2) – up to permutation of the entries. Hence,

$$\begin{aligned}
&\sum_{i_1,i_2,i_3,i_4} \big(\widetilde{(f\star_0^0 f)}(i_1,i_2,i_3,i_4)\big)^2 \mathbb{1}_{\Delta_4^c}(i_1,i_2,i_3,i_4)\\
&\quad= 6\sum_{\substack{i_1,i_2,i_3\\ i_2\neq i_3}} \big(\widetilde{(f\star_0^0 f)}(i_1,i_1,i_2,i_3)\big)^2 + 3\sum_{i_1,i_2}\big(\widetilde{(f\star_0^0 f)}(i_1,i_1,i_2,i_2)\big)^2\\
&\quad= 6\sum_{i_1,i_2,i_3} \big(\widetilde{(f\star_0^0 f)}(i_1,i_1,i_2,i_3)\big)^2 - 3\sum_{i_1,i_2}\big(\widetilde{(f\star_0^0 f)}(i_1,i_1,i_2,i_2)\big)^2. \qquad (4.33)
\end{aligned}$$

Note that for every 4-tuple (i_1,i_2,i_3,i_4) there are 2^3 permutations σ such that

$$f(i_1,i_2)f(i_3,i_4) = f(i_{\sigma(1)},i_{\sigma(2)})f(i_{\sigma(3)},i_{\sigma(4)}).$$

Using this together with the symmetry of f, we deduce that

$$\widetilde{(f\star_0^0 f)}(i_1,i_2,i_3,i_4) = \frac{2^3}{4!}\big(f(i_1,i_2)f(i_3,i_4) + f(i_1,i_3)f(i_2,i_4) + f(i_1,i_4)f(i_2,i_3)\big) \qquad (4.34)$$

for all $i_1,i_2,i_3,i_4\in\mathbb{N}$. Combining (4.33) and (4.34), we get

$$\begin{aligned}
\big\|\widetilde{(f\star_0^0 f)}\mathbb{1}_{\Delta_4^c}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 4}} &= \frac{8}{3}\sum_{i_1,i_2,i_3}\big(f(i_1,i_2)f(i_1,i_3)\big)^2 - \frac{4}{3}\sum_{i_1,i_2} f(i_1,i_2)^4\\
&= \frac{8}{3}\big\|(f\star_1^1 f)\mathbb{1}_{\Delta_2^c}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} - \frac{4}{3}\|f\|^4_{\ell^4(\mathbb{N})^{\otimes 2}}. \qquad (4.35)
\end{aligned}$$

Plugging (4.35) into (4.32), we conclude that

$$\begin{aligned}
\mathbb{E}\big[F^4\big] &= 12\|f\|^4_{\ell^2(\mathbb{N})^{\otimes 2}} + 32\|f\|^4_{\ell^4(\mathbb{N})^{\otimes 2}}\\
&\quad + 48\big(\big\|(f\star_1^1 f)\mathbb{1}_{\Delta_2}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} - \big\|(f\star_1^1 f)\mathbb{1}_{\Delta_2^c}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2}}\big).
\end{aligned}$$

Equation (4.31) now follows immediately from our assumption that $\mathbb{E}[F^2] = 2\|f\|^2_{\ell^2(\mathbb{N})^{\otimes 2}} = 1$. □

Proof of Theorem 4.3. Since F_n converges in distribution to a standard Gaussian random variable and since $\sup_{n\in\mathbb{N}}\mathbb{E}[|F_n|^q] < \infty$ for all $q\geq 2$, due to the hypercontractivity property stated in Lemma 4.2, we have $\mathbb{E}[F_n^4]\to 3$. The statement now follows from (4.31). □

Remark 4.4. *It is a natural question whether the proof of Theorem 4.2 can be modified in such a way that it involves differences of the type*

$$\big\|(f\star_r^r f)\mathbb{1}_{\Delta_{2(q-r)}}\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} - \big\|(f\star_r^r f)\mathbb{1}_{\Delta^c_{2(q-r)}}\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} \qquad (4.36)$$

rather than the contraction norms $\|f \star_r^r f\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}$. *We doubt that this is possible with the usual Malliavin–Stein technique, since already the key term* $A_1(F_n)$ *at* (4.11) *contains only the off-diagonal term in* (4.36). *For Example* 4.1 *with* $F_n = J_2(f_n)$, $A_1(F_n)$ *reduces to*

$$A_1(F_n)^2 = \mathbb{E}\left[\left(1 - \frac{1}{2}\|DF_n\|_{\ell^2(\mathbb{N})}\right)^2\right] = 8\big\|(f_n \star_1^1 f_n)\mathbb{1}_{\Delta_2}\big\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2 = \frac{1}{16}\frac{n-2}{n-1}.$$

This converges to $\frac{1}{16} \neq 0$, *as* $n \to \infty$, *and since the other terms* $A_2(F_n)$, $A_3(F_n)$ *and* $A_4(F_n)$ *in the Malliavin–Stein bound (recall* (4.12), (4.13) *and* (4.14)) *are non-negative,* $d_K(F_n, N)$ *does not converge to zero. Since the variance comparison in terms of* $A_1(F_n)$ *lies at the heart of the Malliavin–Stein method, new ideas are necessary to overcome this difficulty.*

4.4. *Sums of single and double integrals*

For one of our applications below we need an estimate for the Kolmogorov distance between a Rademacher functional of the type

$$F := J_1\big(f^{(1)}\big) + J_2\big(f^{(2)}\big) \quad \text{with } f^{(1)} \in \ell^2(\mathbb{Z}),\ f^{(2)} \in \ell_0^2(\mathbb{Z})^{\circ 2}$$

and a standard Gaussian random variable N. In other words, F is the sum of an element of the first and an element of the second Rademacher chaos. From a technical point of view, such functionals are more elaborate compared to elements of a single Rademacher chaos. We take up this point and develop a bound for $d_K(F, N)$.

As indicated in [26], the results of discrete Malliavin calculus as outlined above and Stein's method extend to Rademacher random variables indexed by $\mathbb{Z}$. In this section and also in our first application presented in Section 6 below we make use of this extension in order to avoid boundary effects, following thereby [26].

Theorem 4.4. *Let* $F = J_1(f^{(1)}) + J_2(f^{(2)})$ *with* $f^{(1)} \in \ell^2(\mathbb{Z})$ *and* $f^{(2)} \in \ell_0^2(\mathbb{Z})^{\circ 2}$, *and let* N *be a standard Gaussian random variable. Suppose that* $\mathbb{E}[F^2] = 1$, *then*

$$\begin{aligned} d_K(F, N) \leq{}& 3\big\|f^{(1)} \star_1^1 f^{(2)}\big\|_{\ell^2(\mathbb{Z})} + 2\sqrt{2}\big\|\big(f^{(2)} \star_1^1 f^{(2)}\big)\mathbb{1}_{\Delta_2}\big\|_{\ell^2(\mathbb{Z})^{\otimes 2}} + 2\big\|f^{(1)}\big\|_{\ell^4(\mathbb{Z})}^2 \\ &+ (4\sqrt{2}+12)\big\|\big(f^{(2)} \star_1^1 f^{(2)}\big)\mathbb{1}_{\Delta_2^c}\big\|_{\ell^2(\mathbb{Z})} + (2\sqrt{13}+6)\bigg(\sum_{k,j\in\mathbb{Z}} f^{(1)}(k)^2 f^{(2)}(k,j)^2\bigg)^{1/2} \\ &+ 2\sum_{k\in\mathbb{Z}}\bigg(\big|f^{(1)}(k)\big| + 2\sum_{j\in\mathbb{Z}}\big|f^{(2)}(j,k)\big|\bigg)^3. \end{aligned}$$

Remark 4.5. *We should compare our result to the bound of Proposition* 5.1 *in* [26]. *It has been shown there that for a twice differentiable function* $g : \mathbb{R} \to \mathbb{R}$ *with bounded derivatives up to order two the estimate*

$$\begin{aligned} \big|\mathbb{E}\big[g(F)\big] - \mathbb{E}\big[g(N)\big]\big| \leq{}& C(g)\Big(3\big\|f^{(1)} \star_1^1 f^{(2)}\big\|_{\ell^2(\mathbb{Z})} + 2\sqrt{2}\big\|\big(f^{(2)} \star_1^1 f^{(2)}\big)\mathbb{1}_{\Delta_2}\big\|_{\ell^2(\mathbb{Z})^{\otimes 2}}\Big) \\ &+ \frac{160}{3}\big\|g''\big\|_\infty \sum_{k\in\mathbb{Z}}\bigg[f^{(1)}(k)^4 + 16\bigg(\sum_{i\in\mathbb{Z}}\big|f^{(2)}(i,k)\big|\bigg)^4\bigg] \end{aligned}$$

holds with $C(g) := \min\{4\|g'\|_\infty, \|g''\|_\infty\}$. *The fact that our bound is more involved is not surprising since already the abstract bound in Theorem* 3.1 *contains more terms compared to* (3.1) *from* [26], *because our bound is based on the fundamental theorem of calculus rather than on an approximate chain rule. As already seen in Theorem* 3.1 *this also leads to different exponents in our bound compared to the results from* [26].

Proof of Theorem 4.4. In view of our general Berry–Esseen estimate in Theorem 3.1 we have to bound the quantities $A_1(F)$–$A_4(F)$ given by (4.2)–(4.5) with $F = J_1(f^{(1)}) + J_2(f^{(2)})$ there. The first term $A_1(F)$ has already been addressed in [26], Proposition 5.1:

$$A_1(F) \le 3\big\| f^{(1)} \star_1^1 f^{(2)} \big\|_{\ell^2(\mathbb{Z})} + 2\sqrt{2}\big\| \big(f^{(2)} \star_1^1 f^{(2)}\big)\mathbb{1}_{\Delta_2} \big\|_{\ell^2(\mathbb{Z})^{\otimes 2}}. \tag{4.37}$$

Using the estimates

$$|D_k F| \le \big|f^{(1)}(k)\big| + 2\sum_{j\in\mathbb{Z}} \big|f^{(2)}(j,k)\big| \quad \text{and} \quad \big|D_k L^{-1} F\big| \le \big|f^{(1)}(k)\big| + 2\sum_{j\in\mathbb{Z}} \big|f^{(2)}(j,k)\big|,$$

valid for all $k \in \mathbb{N}$, we find that (replacing thereby also $\sqrt{2\pi}/4$ by 1)

$$A_2(F) \le \sum_{k\in\mathbb{Z}} \bigg(\big|f^{(1)}(k)\big| + 2\sum_{j\in\mathbb{Z}} \big|f^{(2)}(j,k)\big| \bigg)^3 \tag{4.38}$$

and, in addition, using the Cauchy–Schwarz inequality together with our assumption that $\mathbb{E}[F^2] = 1$,

$$A_3(F) \le \sum_{k\in\mathbb{Z}} \bigg(\big|f^{(1)}(k)\big| + 2\sum_{j\in\mathbb{Z}} \big|f^{(2)}(j,k)\big| \bigg)^3. \tag{4.39}$$

It remains to consider the term $A_4(F)$. Following the strategy already used in the proof of Theorem 4.2 we find that

$$\begin{aligned}
\mathbb{E}\big[\big|\langle (DF) D\mathbb{1}_{\{F>x\}}, (DF)\big|DL^{-1}F\big|\rangle_{\ell^2(\mathbb{Z})}\big| \big] &\le \big(\mathbb{E}\big[\big(\delta\big((DF)\big|DL^{-1}F\big|\big)\big)^2\big]\big)^{1/2} \\
&\le \big(\mathbb{E}\big[\big\|(DF)\big|DL^{-1}F\big|\big\|_{\ell^2(\mathbb{Z})}^2\big]\big)^{1/2} + \bigg(\mathbb{E}\bigg[\sum_{k,l\in\mathbb{Z}} \big(D_k\big((D_l F)\big|D_l L^{-1}F\big|\big)\big)\big(D_l\big((D_k F)\big|D_k L^{-1}F\big|\big)\big)\bigg]\bigg)^{1/2} \\
&\le \big(\mathbb{E}\big[\big\|(DF)\big|DL^{-1}F\big|\big\|_{\ell^2(\mathbb{Z})}^2\big]\big)^{1/2} + \bigg(\mathbb{E}\bigg[\sum_{k,l\in\mathbb{Z}} \big(D_k\big((D_l F)\big|D_l L^{-1}F\big|\big)\big)^2\bigg]\bigg)^{1/2},
\end{aligned}$$

where we have used the isometric relation (2.5) for the divergence operator. Now, let us consider an individual summand

$$\mathbb{E}\big[\big(D_k\big((D_l F)\big|D_l L^{-1}F\big|\big)\big)^2\big] = \mathbb{E}\big[\big(D_k\big((D_l F)\big(D_l L^{-1}F\big)\big)\big)^2\big].$$

Using the product formula (2.9) we see that this equals

$$\mathbb{E}\big[\big((D_k D_l F)\big(D_l L^{-1}F\big) + (D_l F)\big(D_k D_l L^{-1}F\big) - 2X_k (D_k D_l F)\big(D_k D_l L^{-1}F\big)\big)^2\big].$$

Taking into account that

$$D_k F = f^{(1)}(k) + 2J_1\big(f^{(2)}(\cdot,k)\big) \quad \text{and} \quad D_k L^{-1}F = -f^{(1)}(k) - J_1\big(f^{(2)}(\cdot,k)\big), \quad k\in\mathbb{N},$$

we get, after simplifications using the isometry property (2.14) of stochastic integrals,

$$\begin{aligned}
\mathbb{E}\big[\big(D_k\big((D_l F)\big(D_l L^{-1}F\big)\big)\big)^2\big] &= 9 f^{(1)}(l)^2 f^{(2)}(k,l)^2 + 16 f^{(2)}(k,l)^2 \sum_{j\in\mathbb{Z}} f^{(2)}(j,l)^2 - 16 f^{(2)}(k,l)^4 \\
&\le 9 f^{(1)}(l)^2 f^{(2)}(k,l)^2 + 16 f^{(2)}(k,l)^2 \sum_{j\in\mathbb{Z}} f^{(2)}(j,l)^2.
\end{aligned}$$

Next, we notice that $\mathbb{E}[J_1(f^{(2)}(\cdot,k))^3]=0$ and

$$\mathbb{E}\big[J_1\big(f^{(2)}(\cdot k)\big)^4\big]=2\sum_{\substack{i,j\in\mathbb{Z}\\ i\neq j}} f^{(2)}(i,k)^2 f^{(2)}(j,k)^2+\bigg(\sum_{j\in\mathbb{Z}} f^{(2)}(j,k)\bigg)^2,$$

for all $k\in\mathbb{Z}$, as a consequence of the multiplication formula (2.6). Using this together with the representations of D_kF and $D_kL^{-1}F$ from above, we see that

$$\begin{aligned}\mathbb{E}\big[(D_kF)^2\big(D_kL^{-1}F\big)^2\big] &= f^{(1)}(k)^4+13f^{(1)}(k)^2\sum_{j\in\mathbb{Z}} f^{(2)}(j,k)^2\\ &\quad+4\bigg(2\sum_{\substack{i,j\in\mathbb{Z}\\ i\neq j}} f^{(2)}(i,k)^2f^{(2)}(j,k)^2+\bigg(\sum_{j\in\mathbb{Z}} f^{(2)}(j,k)^2\bigg)^2\bigg).\end{aligned}$$

Hence, we find that $A_4(F)$ is bounded by

$$\begin{aligned}A_4(F) &\le 2\big(\mathbb{E}\big[\big\|(DF)\big|DL^{-1}F\big|\big\|^2_{\ell^2(\mathbb{Z})}\big]\big)^{1/2}+2\bigg(\mathbb{E}\bigg[\sum_{k,l\in\mathbb{Z}}\big(D_k\big((D_lF)\big|D_lL^{-1}F\big|\big)\big)^2\bigg]\bigg)^{1/2}\\ &\le 2\bigg(\sum_{k\in\mathbb{Z}} f^{(1)}(k)^4\bigg)^{1/2}+2\bigg(13\sum_{k,j\in\mathbb{Z}} f^{(1)}(k)^2f^{(2)}(j,k)^2\bigg)^{1/2}\\ &\quad+4\bigg(2\sum_{k\in\mathbb{Z}}\sum_{\substack{i,j\in\mathbb{Z}\\ i\neq j}} f^{(2)}(i,k)^2f^{(2)}(j,k)^2\bigg)^{1/2}\\ &\quad+4\bigg(\sum_{k\in\mathbb{Z}}\bigg(\sum_{j\in\mathbb{Z}} f^{(2)}(j,k)^2\bigg)^2\bigg)^{1/2}+6\bigg(\sum_{k,l\in\mathbb{Z}} f^{(1)}(l)^2f^{(2)}(k,l)^2\bigg)^{1/2}\\ &\quad+8\bigg(\sum_{k,j,l\in\mathbb{Z}} f^{(2)}(k,l)^2f^{(2)}(j,k)^2\bigg)^{1/2}\\ &\le 2\bigg(\sum_{k\in\mathbb{Z}} f^{(1)}(k)^4\bigg)^{1/2}+(2\sqrt{13}+6)\bigg(\sum_{k,l\in\mathbb{Z}} f^{(1)}(k)^2f^{(2)}(l,k)^2\bigg)^{1/2}\\ &\quad+(4\sqrt{2}+12)\bigg(\sum_{k\in\mathbb{Z}}\bigg(\sum_{i\in\mathbb{Z}} f^{(2)}(i,k)^2\bigg)^2\bigg)^{1/2}.\end{aligned}$$

Re-writing (whenever this is possible) the sums in terms of norms of suitable contractions completes the proof. □

5. Multivariate limit theorems for Rademacher functionals

In the present section we compare a vector of Rademacher functionals with a multivariate Gaussian random variable. Recall that for this purpose we use the d_4-distance introduced at (2.20).

Theorem 5.1. *Fix an integer $d\ge 2$, let $F_1,\dots,F_d\in\mathrm{dom}(D)$ be Rademacher functionals satisfying $\mathbb{E}[F_i]=0$ and $\mathbb{E}[\|DF_i\|^4_{\ell^4(\mathbb{N})}]<\infty$ for all $i\in\{1,\dots,d\}$ and define the random vector $\mathbf{F}:=(F_1,\dots,F_d)$. Moreover, let $\Sigma=(\sigma_{ij})_{i,j=1}^d$ be a positive semi-definite symmetric $(d\times d)$-matrix and $\mathbf{N}$ a centred d-dimensional Gaussian random*

vector with covariance matrix Σ. Then

$$d_4(\mathbf{F},\mathbf{N}) \le \frac{d}{2}\left(\sum_{i,j=1}^{d}\mathbb{E}\big[\big(\sigma_{ij}-\langle DF_j,-DL^{-1}F_i\rangle_{\ell^2(\mathbb{N})}\big)^2\big]\right)^{1/2} + \frac{5}{3}\mathbb{E}\left[\left\langle\left(\sum_{j=1}^{d}|DF_j|\right)^3,\sum_{i=1}^{d}|DL^{-1}F_i|\right\rangle_{\ell^2(\mathbb{N})}\right]. \tag{5.1}$$

The proof of Theorem 5.1 is given below. A particular case arises if each of the Rademacher functionals has the form of a discrete multiple stochastic integral. Then, Theorem 5.1 implies the following multivariate analogue of Theorem 4.2, whose proof is postponed to the end of this section.

Corollary 5.1. *Fix an integer $d\ge 2$ and $q_1,\dots,q_d\in\mathbb{N}$. Further, let for each $i\in\{1,\dots,d\}$, $f^{(i)}\in\ell_0^2(\mathbb{N})^{\circ q_i}$, define the random vector $\mathbf{F}:=(J_{q_1}(f^{(1)}),\dots,J_{q_d}(f^{(d)}))$ and denote by $\mathbf{N}$ a centred Gaussian random vector with covariance matrix $\Sigma=(\sigma_{ij})_{i,j=1}^d$, such that $\sigma_{ij}=0$ if $q_i\ne q_j$. Then,*

$$\begin{aligned} d_4(\mathbf{F},\mathbf{N}) &\le C_1\max_{i,j=1,\dots,d}\Big\{|\sigma_{ij}-\mathbb{E}\big[J_{q_i}\big(f^{(i)}\big)J_{q_j}\big(f^{(j)}\big)\big]|,\\ &\qquad \max_{r=1,\dots,\min\{q_i,q_j\}}\big\{\big\|f^{(i)}\star_r^r f^{(j)}\big\|_{\ell^2(\mathbb{N})^{\otimes q_i+q_j-2r}}\big\}\mathbb{1}_{\{q_i\ne q_j\}},\\ &\qquad \max_{r=1,\dots,q_i-1}\big\{\big\|f^{(i)}\star_r^r f^{(j)}\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q_i-r)}}\big\}\mathbb{1}_{\{q_i=q_j\}},\\ &\qquad \max_{r=1,\dots,q_i-1}\big\{\big\|f^{(i)}\star_r^r f^{(i)}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q_i-r)}}\big\}\Big\} \qquad (5.2)\\ &\le C_2\max_{i,j=1,\dots,d}\Big\{|\sigma_{ij}-\mathbb{E}\big[J_{q_i}\big(f^{(i)}\big)J_{q_j}\big(f^{(j)}\big)\big]|,\\ &\qquad \max_{r=1,\dots,q_i-1}\big\{\big\|f^{(i)}\star_r^r f^{(i)}\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q_i-r)}},\big\|f^{(i)}\star_r^r f^{(i)}\big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q_i-r)}},\\ &\qquad \big\|f^{(j)}\big\|_{\ell^2(\mathbb{N})^{\otimes q_j}}\big\|f^{(i)}\star_r^r f^{(i)}\big\|^{1/2}_{\ell^2(\mathbb{N})^{\otimes 2(q_i-r)}}\big\}\Big\} \qquad (5.3)\end{aligned}$$

with universal constants $0<C_1,C_2<\infty$ only depending on d and on $q_1,\dots,q_d$.

The second estimate in Corollary 5.1 especialy implies that a random vector

$$\big(J_{q_1}\big(f_n^{(1)}\big),\dots,J_{q_d}\big(f_n^{(d)}\big)\big),\quad f_n^{(i)}\in\ell_0^2(\mathbb{N})^{\circ q_i},\ i\in\{1,\dots,d\},$$

whose entries are discrete multiple stochastic integrals, converges in distribution to a centred Gaussian random vector $\mathbf{N}$ with covariance matrix $\Sigma=(\sigma_{ij})_{i,j=1}^d$ if the following two conditions are satisfied:

(i) $\lim_{n\to\infty}\mathbb{E}[J_{q_i}(f_n^{(i)})J_{q_j}(f_n^{(j)})]=\lim_{n\to\infty}\mathrm{Cov}(J_{q_i}(f_n^{(i)})J_{q_j}(f_n^{(j)}))=\sigma_{ij}$ for all $i,j\in\{1,\dots,d\}$,
(ii) $\lim_{n\to\infty}\|f_n^{(i)}\star_r^r f_n^{(i)}\|_{\ell^2(\mathbb{N})^{\otimes 2(q_i-r)}}=0$ for all $i\in\{1,\dots,d\}$ and all $r\in\{1,\dots,q_i-1\}$.

In view of Remark 4.3, (i) and (ii) imply that the sequence $(J_{q_i}(f_n^{(i)}))_{n\ge1}$ satisfies a univariate central limit theorem in that $J_{q_i}(f_n^{(i)})$ converges in distribution to a one-dimensional centred Gaussian random variable with variance σ_{ii}. This is the discrete analogue to a similar phenomenon observed in [27,28,33,34] for Gaussian and Poisson multiple integrals, respectively.

Remark 5.1. *A bound on the distance between the law of a vector consisting of discrete multiple stochastic integrals and a multivariate Gaussian random variable similar to that in Corollary 5.1 above can also be deduced from*

universality results for so-called homogeneous sums as in [25], *see in particular Theorem* 5.1 *ibidem. However, this approach does not deliver the general estimate* (5.1), *which is of independent interest. For this reason, we prefer to give direct proofs using the multivariate Malliavin–Stein technique on the Rademacher chaos and not to employ universality results from* [25], *which on their part are based, for example, on highly non-trivial arguments centred around the notion of influence, see* [19]. *We use an interpolation technique, which has already been applied in the Gaussian and Poisson context* (*cf.* [25,34]) *together with a new multivariate approximate chain rule for the gradient operator* (*see Lemma* 5.1), *which generalizes the one-dimensional chain rule in* [26], *Proposition* 2.14.

Remark 5.2. *In contrast to the one-dimensional case, in this section a probability metric based on smooth test functions is considered, namely the d_4-distance. The multivariate Kolmogorov distance can then be estimated from above in terms of the d_4-distance using a smoothing argument* (*see Lemma* 12.1 *in* [7] *and the references cited there*), *which usually leads to suboptimal rates of convergence. Dealing with the multivariate Kolmogorov distance without smoothing techniques would require precise information on the solution of the multivariate Stein equation associated with a multivariate indicator function. To the best of our knowledge, this is still an open problem in the theory of Stein's method. In this context we emphasize that the multivariate normal approximation of a vector of multiple stochastic integrals on the Gaussian* (*cf.* [27]) *or the Poisson space* (*cf.* [34]) *in terms of the multivariate Kolmogorov distance has not been considered for the same reason.*

To give a proof of Theorem 5.1 we need the following two lemmas. Recall that by $X = (X_k)_{k\in\mathbb{N}}$ we denote a Rademacher sequence and that a Rademacher functional $F = F(X)$ is a (possibly) non-linear transformation of X.

Lemma 5.1. *Fix an integer $d \geq 2$, let $F_1, \dots, F_d \in \operatorname{dom}(D)$ be Rademacher functionals satisfying $\mathbb{E}[F_i] = 0$ for all $i \in \{1, \dots, d\}$, and define the random vector $\mathbf{F} := (F_1, \dots, F_d)$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be thrice differentiable with continuous and bounded partial derivatives up to order three. Then, for every $k \in \mathbb{N}$,*

$$D_k f(\mathbf{F}) = \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(\mathbf{F})(D_k F_i) - \frac{1}{2} \sum_{i,j=1}^{d} \left(\frac{\partial^2}{\partial x_i\, \partial x_j} f\big(\mathbf{F}_k^+\big) - \frac{\partial^2}{\partial x_i\, \partial x_j} f\big(\mathbf{F}_k^-\big) \right) (D_k F_i)(D_k F_j) X_k + R,$$

where the remainder term R satisfies

$$R = \sum_{i,j,l=1}^{d} R_{i,j,l} \quad \text{with} \quad |R_{i,j,l}| \leq \frac{10}{3} \sup_{x\in\mathbb{R}^d} \left| \frac{\partial^3}{\partial x_i\, \partial x_j\, \partial x_l} f(x) \right| \big| (D_k F_i)(D_k F_j)(D_k F_l) \big|.$$

Proof. For $k \in \mathbb{N}$ we use the definition of the discrete gradient together with a Taylor series expansion of f to see that

$$\begin{aligned} D_k f(\mathbf{F}) &= \frac{1}{2}\big(f\big(\mathbf{F}_k^+\big) - f\big(\mathbf{F}_k^-\big) \big) \\ &= \frac{1}{2}\big(f\big(\mathbf{F}_k^+\big) - f(\mathbf{F}) \big) - \frac{1}{2}\big(f\big(\mathbf{F}_k^-\big) - f(\mathbf{F}) \big) \\ &= \frac{1}{2} \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(\mathbf{F}) \big((F_i)_k^+ - F_i \big) + \frac{1}{4} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i\, \partial x_j} f(\mathbf{F}) \big((F_i)_k^+ - F_i \big) \big((F_j)_k^+ - F_j \big) \\ &\quad - \left(\frac{1}{2} \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(\mathbf{F}) \big((F_i)_k^- - F_i \big) + \frac{1}{4} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i\, \partial x_j} f(\mathbf{F}) \big((F_i)_k^- - F_i \big) \big((F_j)_k^- - F_j \big) \right) \\ &\quad + R_1 + R_2 \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(\mathbf{F}) D_k F_i \\
&\quad + \frac{1}{8} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \, \partial x_j} f(\mathbf{F}) \big[\big((F_i)_k^+ - F_i\big)\big((F_j)_k^+ - F_j\big) - \big((F_i)_k^- - F_i\big)\big((F_j)_k^- - F_j\big) \big] \\
&\quad + \frac{1}{8} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \, \partial x_j} f(\mathbf{F}) \big[\big((F_i)_k^+ - F_i\big)\big((F_j)_k^+ - F_j\big) - \big((F_i)_k^- - F_i\big)\big((F_j)_k^- - F_j\big) \big] \\
&\quad + R_1 + R_2.
\end{aligned}
$$

where

$$
R_1 := \sum_{i,j,l=1}^{d} R_{i,j,l}^{(1)}
$$

with

$$
\begin{aligned}
\big|R_{i,j,l}^{(1)}\big| &\le \frac{1}{12} \sup_{x \in \mathbb{R}^d} \left| \frac{\partial^3}{\partial x_i \, \partial x_j \, \partial x_l} f(x) \right| \big| \big((F_i)_k^+ - F_i\big)\big((F_j)_k^+ - F_j\big)\big((F_l)_k^+ - F_l\big) \big| \\
&\le \frac{2}{3} \sup_{x \in \mathbb{R}^d} \left| \frac{\partial^3}{\partial x_i \, \partial x_j \, \partial x_l} f(x) \right| \big| (D_k F_i)(D_k F_j)(D_k F_l) \big|,
\end{aligned}
$$

and

$$
R_2 := \sum_{i,j,l=1}^{d} R_{i,j,l}^{(2)}
$$

satisfies

$$
\begin{aligned}
\big|R_{i,j,l}^{(2)}\big| &\le \frac{1}{12} \sup_{x \in \mathbb{R}^d} \left| \frac{\partial^3}{\partial x_i \, \partial x_j \, \partial x_l} f(x) \right| \big| \big((F_i)_k^- - F_i\big)\big((F_j)_k^- - F_j\big)\big((F_l)_k^- - F_l\big) \big| \\
&\le \frac{2}{3} \sup_{x \in \mathbb{R}^d} \left| \frac{\partial^3}{\partial x_i \, \partial x_j \, \partial x_l} f(x) \right| \big| (D_k F_i)(D_k F_j)(D_k F_l) \big|.
\end{aligned}
$$

To the two sums on the right-hand side of the last equation we now apply a Taylor expansion of $\frac{\partial^2}{\partial x_i \partial x_j} f$ up to order 1 about $\mathbf{F}_k^+$ and $\mathbf{F}_k^-$, respectively. We thus obtain that

$$
\begin{aligned}
D_k f(\mathbf{F}) = \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f(\mathbf{F}) D_k F_i + \frac{1}{8} \sum_{i,j=1}^{d} \Bigg\{ \Bigg[\frac{\partial^2}{\partial x_i \, \partial x_j} f\big(\mathbf{F}_k^+\big) + \frac{\partial^2}{\partial x_i \, \partial x_j} f\big(\mathbf{F}_k^-\big) \Bigg] \\
\times \big[\big((F_i)_k^+ - F_i\big)\big((F_j)_k^+ - F_j\big) - \big((F_i)_k^- - F_i\big)\big((F_j)_k^- - F_j\big) \big] \Bigg\} + R_1 + R_2 + R_3,
\end{aligned}
$$

where

$$
R_3 := \sum_{i,j,l=1}^{d} R_{i,j,l}^{(3)}
$$

is such that

$$\begin{aligned}
&\big|R^{(3)}_{i,j,l}\big| \\
&\quad\le \frac{1}{8}\sup_{x\in\mathbb{R}^d}\left|\frac{\partial^3}{\partial x_i\,\partial x_j\,\partial x_l}f(x)\right| \\
&\qquad\times\Big(\big|\big((F_i)_k^+ - F_i\big)\big((F_j)_k^+ - F_j\big)\big((F_l)_k^+ - F_l\big)\big| + \big|\big((F_i)_k^- - F_i\big)\big((F_j)_k^- - F_j\big)\big((F_l)_k^- - F_l\big)\big|\Big) \\
&\quad\le 2\sup_{x\in\mathbb{R}^d}\left|\frac{\partial^3}{\partial x_i\,\partial x_j\,\partial x_l}f(x)\right|\big|(D_kF_i)(D_kF_j)(D_kF_l)\big|.
\end{aligned}$$

This yields the result. □

Lemma 5.2. *Fix an integer $d\ge 2$, let $F_0,F_1,\dots,F_d\in\operatorname{dom}(D)$ be Rademacher functionals satisfying $\mathbb{E}[F_i]=0$ and $\mathbb{E}[\|DF_i\|^4_{\ell^4(\mathbb{N})}]<\infty$ for $1\le i\le d$ and define the random vector $\mathbf{F}:=(F_1,\dots,F_d)$. Let $f:\mathbb{R}^d\to\mathbb{R}$ be thrice differentiable with continuous and bounded partial derivatives up to order three. Then,*

$$\mathbb{E}\big[f(\mathbf{F})F_0\big]=\mathbb{E}\left[\sum_{j=1}^d\frac{\partial}{\partial x_j}f(\mathbf{F})\big\langle DF_j,-DL^{-1}F_0\big\rangle_{\ell^2(\mathbb{N})}\right]+\mathbb{E}\big[\big\langle R,-DL^{-1}F_0\big\rangle_{\ell^2(\mathbb{N})}\big]$$

with R satisfying the estimate

$$\big|\mathbb{E}\big[\big\langle R,-DL^{-1}F_0\big\rangle_{\ell^2(\mathbb{N})}\big]\big|\le\frac{10}{3}M_3(f)\mathbb{E}\left[\left\langle\left(\sum_{j=1}^d|DF_j|\right)^3,\big|DF^{-1}F_0\big|\right\rangle_{\ell^2(\mathbb{N})}\right].$$

Proof. We use relation (2.12) and the integration-by-parts formula to see that

$$\mathbb{E}\big[f(\mathbf{F})F_0\big]=\mathbb{E}\big[\delta\big(-DL^{-1}F_0\big)f(\mathbf{F})\big]=\mathbb{E}\big[\big\langle Df(\mathbf{F}),-DL^{-1}F_0\big\rangle_{\ell^2(\mathbb{N})}\big],$$

since $f(\mathbf{F})\in\operatorname{dom}(D)$ due to the assumed boundedness of the partial derivatives of the test function f. Now, we apply Lemma 5.1 to re-write $Df(\mathbf{F})$. This leads to

$$\mathbb{E}\big[f(\mathbf{F})F_0\big]=\mathbb{E}\left[\sum_{j=1}^d\frac{\partial}{\partial x_j}f(\mathbf{F})\big\langle DF_j,-DL^{-1}F_0\big\rangle_{\ell^2(\mathbb{N})}\right]+\mathbb{E}\big[\big\langle R,-DL^{-1}F_0\big\rangle_{\ell^2(\mathbb{N})}\big]$$

with the term R satisfying

$$\big|\mathbb{E}\big[\big\langle R,-DL^{-1}F_0\big\rangle_{\ell^2(\mathbb{N})}\big]\big|\le\frac{10}{3}M_3(f)\mathbb{E}\left[\left\langle\left(\sum_{j=1}^d|DF_j|\right)^3,\big|DL^{-1}F_0\big|\right\rangle_{\ell^2(\mathbb{N})}\right].$$

Note that the term involving second-order partial derivatives of f vanishes because of [26], Lemma 2.13(1), and an application of Fubini's theorem. More precisely, our assumption that $\mathbb{E}[\|DF_i\|^4_{\ell^4(\mathbb{N})}]<\infty$ for all $1\le i\le d$ and [26], Lemma 2.13(3), is needed in order to justify an exchange of the order of integration by means of Fubini's theorem. This proves the claim. □

Proof of Theorem 5.1. We use the so-called smart-path technique, which in the context of the Malliavin–Stein method has previously found application in [25] and [34].

Let $g:\mathbb{R}^d\to\mathbb{R}$ be continuously differentiable up to order 4 and define

$$\Psi(t):=\mathbb{E}\big[g(\sqrt{t}\mathbf{N}+\sqrt{1-t}\mathbf{F})\big].$$

Here $\mathbf{F} = (F_1, \dots, F_d)$ is a vector of Rademacher functionals and $\mathbf{N} = (N_1, \dots, N_d)$ is a Gaussian random vector with covariance matrix $\Sigma = (\sigma_{ij})_{i,j=1}^d$. Then,

$$\bigl|\mathbb{E}\bigl[g(\mathbf{F})\bigr] - \mathbb{E}\bigl[g(\mathbf{N})\bigr]\bigr| = \bigl|\Psi(0) - \Psi(1)\bigr| \le \sup_{t\in(0,1)} \bigl|\Psi'(t)\bigr|$$

with the derivative of Ψ given by

$$\Psi'(t) = \frac{1}{2\sqrt{t}} A - \frac{1}{2\sqrt{1-t}} B, \tag{5.4}$$

where

$$A := \sum_{i=1}^d \mathbb{E}\biggl[\frac{\partial}{\partial x_i} g(\sqrt{t}\mathbf{N} + \sqrt{1-t}\mathbf{F}) N_i\biggr] \quad \text{and} \quad B := \sum_{i=1}^d \mathbb{E}\biggl[\frac{\partial}{\partial x_i} g(\sqrt{t}\mathbf{N} + \sqrt{1-t}\mathbf{F}) F_i\biggr].$$

As in the proof of Theorem 4.2 in [34] an integration-by-parts argument shows that

$$A = \sqrt{t} \sum_{i,j=1}^d \sigma_{ij} \mathbb{E}\biggl[\frac{\partial^2}{\partial x_i\, \partial x_j} g(\sqrt{t}\mathbf{N} + \sqrt{1-t}\mathbf{F})\biggr]. \tag{5.5}$$

To evaluate B further, let us condition on $\mathbf{N} = \mathbf{b} \in \mathbb{R}^d$, i.e.,

$$B = \sum_{i=1}^d \mathbb{E}\biggl[\mathbb{E}\biggl[\frac{\partial}{\partial x_i} g(\sqrt{t}\mathbf{b} + \sqrt{1-t}\mathbf{F}) F_i\biggr]_{\mathbf{b}=\mathbf{N}}\biggr]$$

and put $g_i^{t,\mathbf{b}}(\mathbf{F}) := \frac{\partial}{\partial x_i} g(\sqrt{t}\mathbf{b} + \sqrt{1-t}\mathbf{F})$. Now, we apply Lemma 5.2 to the conditional expectation, which leads to

$$\begin{aligned}
\mathbb{E}\bigl[g_i^{t,\mathbf{b}}(\mathbf{F}) F_i\bigr]_{\mathbf{b}=\mathbf{N}} &= \mathbb{E}\biggl[\sum_{j=1}^d \frac{\partial}{\partial x_j} g_i^{t,\mathbf{b}}(\mathbf{F}) \bigl\langle DF_j, -DL^{-1}F_i\bigr\rangle_{\ell^2(\mathbb{N})}\biggr]_{\mathbf{b}=\mathbf{N}} \\
&\quad + \mathbb{E}\bigl[\bigl\langle R, -DL^{-1}F_i\bigr\rangle_{\ell^2(\mathbb{N})}\bigr]_{\mathbf{b}=\mathbf{N}} \\
&= \sqrt{1-t}\,\mathbb{E}\biggl[\sum_{j=1}^d \frac{\partial}{\partial x_j\, \partial x_i} g(\sqrt{t}\mathbf{b} + \sqrt{1-t}\mathbf{F}) \bigl\langle DF_j, -DL^{-1}F_i\bigr\rangle_{\ell^2(\mathbb{N})}\biggr]_{\mathbf{b}=\mathbf{N}} \\
&\quad + \mathbb{E}\bigl[\bigl\langle R, -DL^{-1}F_i\bigr\rangle_{\ell^2(\mathbb{N})}\bigr]_{\mathbf{b}=\mathbf{N}},
\end{aligned}$$

with R satisfying

$$\bigl|\mathbb{E}\bigl[\bigl\langle R, -DL^{-1}F_i\bigr\rangle_{\ell^2(\mathbb{N})}\bigr]_{\mathbf{b}=\mathbf{N}}\bigr| \le \frac{10}{3} M_3\bigl(g_i^{t,\mathbf{b}}\bigr) \mathbb{E}\biggl[\biggl\langle \biggl(\sum_{j=1}^d |DF_j|\biggr)^3, \bigl|DL^{-1}F_i\bigr|\biggr\rangle_{\ell^2(\mathbb{N})}\biggr].$$

Thus,

$$\begin{aligned}
B &= \sqrt{1-t} \sum_{i,j=1}^d \mathbb{E}\biggl[\frac{\partial^2}{\partial x_i\, \partial x_j} g(\sqrt{t}\mathbf{N} + \sqrt{1-t}\mathbf{F}) \bigl\langle DF_j, -DL^{-1}F_i\bigr\rangle_{\ell^2(\mathbb{N})}\biggr] \\
&\quad + \sum_{i=1}^d \mathbb{E}\bigl[\mathbb{E}\bigl[\bigl\langle R, -DL^{-1}F_i\bigr\rangle_{\ell^2(\mathbb{N})}\bigr]_{\mathbf{b}=\mathbf{N}}\bigr].
\end{aligned}$$

Now, note that $M_3(g_i^{t,\mathbf{b}}) \le (1-t)^{3/2} M_4(g)$ so that

$$\begin{aligned}
&\left| \sum_{i=1}^d \mathbb{E}\big[\mathbb{E}\big[\big\langle R, -DL^{-1}F_i\big\rangle_{\ell^2(\mathbb{N})}\big]_{\mathbf{b}=\mathbf{N}}\big]\right| \\
&\quad \le \frac{10}{3} M_4(g)(1-t)^{3/2}\mathbb{E}\left[\left\langle \left(\sum_{j=1}^d |DF_j|\right)^3, \sum_{i=1}^d \big|DL^{-1}F_i\big|\right\rangle_{\ell^2(\mathbb{N})}\right],
\end{aligned}$$

independently of $\mathbf{b}$. Thus, (5.4) implies the bound

$$\begin{aligned}
\big|\mathbb{E}\big[g(\mathbf{F})\big] - \mathbb{E}\big[g(\mathbf{N})\big]\big| &\le \frac{1}{2} M_2(g) \sum_{i,j=1}^d \mathbb{E}\big[\big|\sigma_{ij} - \big\langle DF_j, -DL^{-1}F_i\big\rangle_{\ell^2(\mathbb{N})}\big|\big] \\
&\quad + \frac{1}{2}\frac{10}{3}\frac{1}{\sqrt{1-t}} M_4(g)(1-t)^{3/2}\mathbb{E}\left[\left\langle \left(\sum_{j=1}^d |DF_j|\right)^3, \sum_{i=1}^d \big|DL^{-1}F_i\big|\right\rangle_{\ell^2(\mathbb{N})}\right] \\
&\le \frac{d}{2}\left(\sum_{i,j=1}^d \mathbb{E}\big[\big(\sigma_{ij} - \big\langle DF_j, -DL^{-1}F_i\big\rangle_{\ell^2(\mathbb{N})}\big)^2\big]\right)^{1/2} \\
&\quad + \frac{5}{3}\mathbb{E}\left[\left\langle \left(\sum_{j=1}^d |DF_j|\right)^3, \sum_{i=1}^d \big|DL^{-1}F_i\big|\right\rangle_{\ell^2(\mathbb{N})}\right],
\end{aligned}$$

where in the last line we have applied the Cauchy–Schwarz inequality and took the supremum over all $t \in (0,1)$ and over all functions $g : \mathbb{R}^d \to \mathbb{R}$ with $M_2(g) \le 1$ and $M_4(g) \le 1$. This proves the theorem. □

Proof of Corollary 5.1. As in the proof of Theorem 4.2 we assume without loss of generality that the multiple integrals $J_{q_i}(f^{(i)})$ have kernels $f^{(i)}$ with finite support for all $1 \le i \le d$. Fix $i, j \in \{1, \dots, d\}$ and to ease the notation let $\sigma := \sigma_{ij}$, $p := q_j$, $q := q_i$, $f := f^{(j)}$, $g := f^{(i)}$. Assume without loss of generality that $p \le q$. We apply Theorem 5.1 to the random vector $\mathbf{F}$ (note that the assumption $\mathbb{E}[\|DF_i\|^4_{\ell^4(\mathbb{N})}] < \infty$ for $1 \le i \le d$ is fulfilled by Lemma 4.2, see also [26]). Using the multiplication formula (2.6) we get for all $k \in \mathbb{N}$,

$$\begin{aligned}
&\big(D_k J_p(f)\big)\big(-D_k L^{-1} J_q(g)\big) \\
&\quad = \frac{1}{q}\big(D_k J_p(f)\big)\big(D_k J_q(g)\big) \\
&\quad = p J_{p-1}\big(f(\cdot,k)\big) J_{q-1}\big(g(\cdot,k)\big) \\
&\quad = p \sum_{r=0}^{p-1} r! \binom{p-1}{r}\binom{q-1}{r} J_{p+q-2(r+1)}\big(\big(\widetilde{f(\cdot,k) \star_r^r g(\cdot,k)}\big)\mathbb{1}_{\Delta_{p+q-2(r+1)}}\big).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\big\langle DJ_p(f), -DL^{-1}J_q(g)\big\rangle_{\ell^2(\mathbb{N})} \\
&\quad = p \sum_{r=0}^{p-1} r! \binom{p-1}{r}\binom{q-1}{r} J_{p+q-2(r+1)}\big(\big(\widetilde{f \star_{r+1}^{r+1} g}\big)\mathbb{1}_{\Delta_{p+q-2(r+1)}}\big) \\
&\quad = p \sum_{r=1}^{p} (r-1)! \binom{p-1}{r-1}\binom{q-1}{r-1} J_{p+q-2r}\big(\big(\widetilde{f \star_r^r g}\big)\mathbb{1}_{\Delta_{p+q-2r}}\big).
\end{aligned}$$

If $p < q$ we get by isometry (2.14) of discrete multiple stochastic integrals,

$$\begin{aligned}
&\mathbb{E}\big[\big(\sigma - \langle DJ_p(f), -DL^{-1}J_q(g)\rangle\big)^2\big] \\
&\quad = \sigma^2 + p^2 \sum_{r=1}^{p} \big((r-1)!\big)^2 \binom{p-1}{r-1}^2 \binom{q-1}{r-1}^2 (p+q-2r)! \big\| \widetilde{(f \star_r^r g)} \mathbb{1}_{\Delta_{p+q-2r}} \big\|^2_{\ell^2(\mathbb{N})^{\otimes p+q-2r}} \qquad (5.6) \\
&\quad \le \sigma^2 + p^2 \sum_{r=1}^{p} \big((r-1)!\big)^2 \binom{p-1}{r-1}^2 \binom{q-1}{r-1}^2 (p+q-2r)! \big\| f \star_r^r g \big\|^2_{\ell^2(\mathbb{N})^{\otimes p+q-2r}}. \qquad (5.7)
\end{aligned}$$

If $p = q$ we get

$$\begin{aligned}
&\mathbb{E}\big[\big(\sigma - \langle DJ_p(f), -DL^{-1}J_q(g)\rangle_{\ell^2(\mathbb{N})}\big)^2\big] \\
&\quad = \mathbb{E}\left[\left(\big(\sigma - p!\langle f, g\rangle_{\ell^2(\mathbb{N})^{\otimes p}}\big) - p\sum_{r=1}^{p-1}(r-1)!\binom{p-1}{r-1}^2 J_{2(p-r)}\big(\widetilde{(f \star_r^r g)}\mathbb{1}_{\Delta_{2(p-r)}}\big)\right)^2\right] \\
&\quad = \big(\sigma - p!\langle f, g\rangle_{\ell^2(\mathbb{N})^{\otimes p}}\big)^2 \\
&\qquad + p^2 \sum_{r=1}^{p-1} \big((r-1)!\big)^2 \binom{p-1}{r-1}^4 \big(2(p-1)\big)! \big\| \widetilde{(f \star_r^r g)} \mathbb{1}_{\Delta_{2(p-r)}} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(p-r)}} \\
&\quad \le \big(\sigma - \mathbb{E}\big[J_p(f)J_q(g)\big]\big)^2 \\
&\qquad + p^2 \sum_{r=1}^{p-1} \big((r-1)!\big)^2 \binom{p-1}{r-1}^4 \big(2(p-1)\big)! \big\| f \star_r^r g \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(p-r)}}. \qquad (5.8)
\end{aligned}$$

Let us now consider the second term in the bound of Theorem 5.1. Let $F_i := J_{q_i}(f^{(i)})$ for $1 \le i \le d$. We have

$$\begin{aligned}
\mathbb{E}\left[\left\langle \left(\sum_{j=1}^{d} |DF_j|\right)^3, \sum_{i=1}^{d} \big|DL^{-1}F_i\big| \right\rangle_{\ell^2(\mathbb{N})}\right] &= \mathbb{E}\left[\left\langle \left(\sum_{j=1}^{d} |DF_j|\right)^3, \sum_{i=1}^{d} \frac{1}{q_i}|DF_i| \right\rangle_{\ell^2(\mathbb{N})}\right] \\
&\le \frac{1}{\min\{q_1, \dots, q_d\}} \mathbb{E}\left[\sum_{k=1}^{\infty}\left(\sum_{j=1}^{d} |D_k F_i|\right)^4\right] \\
&\le \frac{d^3}{\min\{q_1, \dots, q_d\}} \sum_{i=1}^{d} \mathbb{E}\big[\|DF_i\|^4_{\ell^4(\mathbb{N})}\big] \qquad (5.9) \\
&\le C \max_{\substack{i=1,\dots,d \\ r=1,\dots,q_i-1}} \big\{ \big\| f^{(i)} \star_r^r f^{(i)} \big\|^2_{\ell^2(\mathbb{N})^{\otimes 2(q_i-r)}} \big\}, \qquad (5.10)
\end{aligned}$$

with a constant $0 < C < \infty$ only depending on d and on $q_1, \dots, q_d$. In (5.9) Hölder's inequality with Hölder conjugates 4 and $\frac{4}{3}$ has been used. Inequality (5.10) follows from (4.10). Now, Theorem 5.1 together with (5.7), (5.8) and (5.10) gives the first inequality in Corollary 5.1. To derive the second inequality, we need to show that all 'mixed' contractions $f \star_r^r g = f^{(i)} \star_r^r f^{(j)}$ appearing in (5.7) and (5.8) can be bounded by contractions of the form $f \star_r^r f$ and $g \star_r^r g$ build up either by f or g. In the following we use the vector notation $\mathbf{k}_r := (k_1, \dots, k_r) \in \mathbb{N}^r$ for summation indices. Moreover $\mathbf{k}_r\mathbf{k}_\ell$ stands for the concatenation $(k_1, \dots, k_r, k_{r+1}, \dots, k_{r+\ell})$ of two vectors $\mathbf{k}_r$ and $\mathbf{k}_\ell$. First, let us consider the case $r < p \le q$. Then,

$$\big\| f \star_r^r g \big\|^2_{\ell^2(\mathbb{N})^{\otimes p+q-2r}} = \sum_{\mathbf{k}_{p+q-2r}} \big(f \star_r^r g\big)(\mathbf{k}_{p+q-2r})^2$$

$$= \sum_{\mathbf{k}_{p-r}} \sum_{\mathbf{k}_{q-r}} \sum_{\mathbf{a}_r} \sum_{\mathbf{b}_r} f(\mathbf{k}_{p-r}\mathbf{a}_r) f(\mathbf{k}_{p-r}\mathbf{b}_r) g(\mathbf{k}_{q-r}\mathbf{a}_r) g(\mathbf{k}_{q-r}\mathbf{b}_r)$$

$$= \sum_{\mathbf{a}_r} \sum_{\mathbf{b}_r} \big(f \star_{p-r}^{p-r} f\big)(\mathbf{a}_r\mathbf{b}_r) \big(g \star_{q-r}^{q-r} g\big)(\mathbf{a}_r\mathbf{b}_r)$$

$$= \big\langle f \star_{p-r}^{p-r} f, g \star_{q-r}^{q-r} g \big\rangle_{\ell^2(\mathbb{N})^{\otimes 2r}} \le \big\| f \star_{p-r}^{p-r} f \big\|_{\ell^2(\mathbb{N})^{\otimes 2r}} \big\| g \star_{q-r}^{q-r} g \big\|_{\ell^2(\mathbb{N})^{\otimes 2r}},$$

which implies that

$$\big\| f \star_r^r g \big\|_{\ell^2(\mathbb{N})^{\otimes p+q-2r}} \le \max\Big\{ \big\| f \star_{p-r}^{p-r} f \big\|_{\ell^2(\mathbb{N})^{\otimes 2r}}, \big\| g \star_{q-r}^{q-r} g \big\|_{\ell^2(\mathbb{N})^{\otimes 2r}} \Big\}. \tag{5.11}$$

If $r = p < q$, we get

$$\big\| f \star_p^p g \big\|^2_{\ell^2(\mathbb{N})^{\otimes(q-p)}} = \sum_{\mathbf{k}_{q-p}} \big(f \star_p^p g\big)(\mathbf{k}_{q-p})^2$$

$$= \sum_{\mathbf{k}_{q-p}} \sum_{\mathbf{a}_r} \sum_{\mathbf{b}_r} f(\mathbf{a}_r) f(\mathbf{b}_r) g(\mathbf{k}_{q-p}\mathbf{a}_r) g(\mathbf{k}_{q-p}\mathbf{b}_r)$$

$$= \sum_{\mathbf{a}_r} \sum_{\mathbf{b}_r} f(\mathbf{a}_r) f(\mathbf{b}_r) \big(g \star_{q-p}^{q-p} g\big)(\mathbf{a}_r\mathbf{b}_r).$$

Hence,

$$\big\| f \star_p^p g \big\|_{\ell^2(\mathbb{N})^{\otimes(q-p)}} \le \|f\|_{\ell^2(\mathbb{N})^{\otimes p}} \big(\big\| g \star_{q-p}^{q-p} g \big\|_{\ell^2(\mathbb{N})^{\otimes 2p}} \big)^{1/2}. \tag{5.12}$$

Now, the second estimate (5.3) in Corollary 5.1 follows from the first estimate (5.2) by means of (5.11) and (5.12). This completes the proof. □

6. Applications

6.1. *Infinite weighted* 2*-runs*

The notion of a 2-run is one of the most simple dependency structures and has been studied exhaustively in the literature, see [2]. For example, [42] uses the so-called local approach, while [36,38] apply exchangeable pair coupling constructions. The first Berry–Esseen bound for the number of finite 2-runs appeared in [11]. In contrast to most of the previously available results, our method, which is based on a discrete version of the Malliavin calculus of variations, allows to treat infinite weighted 2-runs. This continues the line of research initiated in Section 5 of [26] and provides, up to our best knowledge, the first Berry–Esseen bound for such functionals.

Let $X = (X_i)_{i\in\mathbb{Z}}$ be a double-sided sequence of i.i.d. Rademacher random variables and let for each $n \in \mathbb{N}$, $(a_i^{(n)})_{i\in\mathbb{Z}}$ be a double-sided sequence of real numbers. The sequence $(F_n)_{n\in\mathbb{N}}$ of normalized infinite weighted 2-runs is then defined as

$$F_n := \frac{G_n - \mathbb{E}G_n}{\sqrt{\operatorname{Var} G_n}}, \qquad G_n := \sum_{i\in\mathbb{Z}} a_i^{(n)} Y_i Y_{i+1}, \quad n \in \mathbb{N},$$

where $Y_i := \frac{1}{2}(1 - X_i)$ for $i \in \mathbb{Z}$. In other words, G_n counts the weighted number of subsequences of 1's of length two in an infinite double-sided sequence of Bernoulli trials. We notice that

$$\operatorname{Var} G_n = \frac{3}{16} \sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^2 + \frac{1}{8} \sum_{i\in\mathbb{Z}} a_i^{(n)} a_{i+1}^{(n)}.$$

Our result reads as follows.

Theorem 6.1. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of normalized infinite weighted 2-runs as above and let N be a standard Gaussian random variable. Then,*

$$d_K(F_n, N) \le C \max\left\{ (\operatorname{Var} G_n)^{-3/2} \sum_{i\in\mathbb{Z}} \big|a_i^{(n)}\big|^3, (\operatorname{Var} G_n)^{-1} \left(\sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^4 \right)^{1/2} \right\}$$

with a constant $0 < C < \infty$ not depending on n.

Proof. We notice first that each F_n has chaotic decomposition $F_n = J_1(f_n^{(1)}) + J_2(f_n^{(2)})$ with $f_n^{(1)}$ and $f_n^{(2)}$ given by

$$f_n^{(1)}(k) = \frac{1}{4\sqrt{\operatorname{Var} G_n}} \sum_{i\in\mathbb{Z}} a_i^{(n)} (\mathbb{1}_{\{k=i\}} + \mathbb{1}_{\{k=i+1\}}), \quad k \in \mathbb{Z},$$

$$f_n^{(2)}(k,l) = \frac{1}{8\sqrt{\operatorname{Var} G_n}} \sum_{i\in\mathbb{Z}} a_i^{(n)} (\mathbb{1}_{\{k=i,l=i+1\}} + \mathbb{1}_{\{k=i+1,l=i\}}), \quad k,l \in \mathbb{Z}.$$

Thus,

$$\big(f_n^{(1)} \star_1^1 f_n^{(2)}\big)(i) = \frac{1}{32 \operatorname{Var} G_n} \big(a_i^{(n)} a_{i+1}^{(n)} + \big(a_{i-1}^{(n)}\big)^2 + \big(a_i^{(n)}\big)^2 + a_{i-1}^{(n)} a_{i-2}^{(n)}\big),$$

$$\big(f_n^{(2)} \star_1^1 f_n^{(2)}\big)(i,j) = \frac{1}{64 \operatorname{Var} G_n} \big(a_i^{(n)} a_{i+1}^{(n)} \mathbb{1}_{\{j=i+2\}} + a_j^{(n)} a_{j+1}^{(n)} \mathbb{1}_{\{j=i-2\}}\big), \quad i \ne j,$$

$$\big(f_n^{(2)} \star_1^1 f_n^{(2)}\big)(i,i) = \frac{1}{64 \operatorname{Var} G_n} \big(\big(a_{i-1}^{(n)}\big)^2 + \big(a_i^{(n)}\big)^2\big),$$

and consequently

$$\begin{aligned}
\big\| f_n^{(1)} \star_1^1 f_n^{(2)} \big\|_{\ell^2(\mathbb{Z})} &\le \frac{\sqrt{2}}{16 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^2 \big(a_{i+1}^{(n)}\big)^2 + \sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^4 \right)^{1/2} \\
&\le \frac{1}{8 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^4 \right)^{1/2}, \\
\big\| \big(f_n^{(2)} \star_1^1 f_n^{(2)}\big) \mathbb{1}_{\Delta_2} \big\|_{\ell^2(\mathbb{Z})^{\otimes 2}} &\le \frac{\sqrt{2}}{64 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^2 \big(a_{i+1}^{(n)}\big)^2 \right)^{1/2} \\
&\le \frac{\sqrt{2}}{64 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^4 \right)^{1/2}, \\
\big\| \big(f_n^{(2)} \star_1^1 f_n^{(2)}\big) \mathbb{1}_{\Delta_2^c} \big\|_{\ell^2(\mathbb{Z})^{\otimes 2}} &\le \frac{1}{64 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(\big(a_{i-1}^{(n)}\big)^2 + \big(a_i^{(n)}\big)^2\big)^2 \right)^{1/2} \\
&\le \frac{\sqrt{2}}{64 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(\big(a_{i-1}^{(n)}\big)^4 + \big(a_i^{(n)}\big)^4\big) \right)^{1/2} \\
&\le \frac{1}{32 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^4 \right)^{1/2},
\end{aligned}$$

where we have used the Cauchy–Schwarz inequality. Moreover,

$$\big\| f_n^{(1)} \big\|_{\ell^4(\mathbb{Z})}^2 \le \frac{1}{4 \operatorname{Var} G_n} \left(\sum_{i\in\mathbb{Z}} \big(a_i^{(n)}\big)^4 \right)^{1/2}.$$

Multiple use of the Cauchy–Schwarz inequality gives

$$\begin{aligned}\sum_{k,j\in\mathbb{Z}}\big(f_n^{(1)}(k)\big)^2\big(f_n^{(2)}(k,j)\big)^2 &= \frac{1}{4096(\operatorname{Var}G_n)^2}\sum_{i\in\mathbb{Z}}\big(\big(a_{i-1}^{(n)}\big)^2+\big(a_i^{(n)}\big)^2\big)\big(a_{i-1}^{(n)}+a_i^{(n)}\big)^2\\ &\le \frac{1}{2048(\operatorname{Var}G_n)^2}\sum_{i\in\mathbb{Z}}\big(\big(a_{i-1}^{(n)}\big)^2+\big(a_i^{(n)}\big)^2\big)^2\\ &\le \frac{1}{1024(\operatorname{Var}G_n)^2}\sum_{i\in\mathbb{Z}}\big(\big(a_{i-1}^{(n)}\big)^4+\big(a_i^{(n)}\big)^4\big)\\ &\le \frac{1}{512(\operatorname{Var}G_n)^2}\sum_{i\in\mathbb{Z}}\big(a_i^{(n)}\big)^4.\end{aligned}$$

Hence,

$$\bigg(\sum_{k,j\in\mathbb{Z}}\big(f_n^{(1)}(k)\big)^2\big(f_n^{(2)}(k,j)\big)^2\bigg)^{1/2}\le\frac{1}{\sqrt{512}\operatorname{Var}G_n}\bigg(\sum_{i\in\mathbb{Z}}\big(a_i^{(n)}\big)^4\bigg)^{1/2}.$$

Finally, by Hölder's inequality, we have

$$\begin{aligned}\sum_{k\in\mathbb{Z}}\bigg(\big|f^{(1)}(k)\big|+2\sum_{j\in\mathbb{Z}}\big|f^{(2)}(j,k)\big|\bigg)^3 &\le 4\sum_{k\in\mathbb{Z}}\bigg(\big|f^{(1)}(k)\big|^3+8\bigg(\sum_{j\in\mathbb{Z}}\big|f^{(2)}(j,k)\big|\bigg)^3\bigg)\\ &= 4\sum_{k\in\mathbb{Z}}\big|f^{(1)}(k)\big|^3+32\sum_{k\in\mathbb{Z}}\bigg(\sum_{j\in\mathbb{Z}}\big|f^{(2)}(j,k)\big|\bigg)^3,\end{aligned}$$

where each of these summands is bounded by a constant times

$$\frac{1}{(\operatorname{Var}G_n)^{3/2}}\sum_{i\in\mathbb{Z}}\big|a_i^{(n)}\big|^3.$$

Now, applying Theorem 4.4 proves the result. □

6.2. *A combinatorial central limit theorem*

Our second application deals with an extended version of a combinatorial central limit theorem of Blei and Janson [5], which has also been studied in [26]. Using Theorem 4.2 we can strengthen these results to a Berry–Esseen bound without imposing further conditions.

The general set-up is as follows. Fix $q\ge 2$ and let $F\subseteq\Delta_q$ be a (possibly infinite) non-empty subset of $\mathbb{N}^q$. Let further $b=(b_i)_{i\ge1}\in\ell^2(\mathbb{N})$ and define a measure μ_b on $\mathbb{N}$ by putting

$$\mu_b(B):=\sum_{i\in B}b_i^2,\quad B\subset\mathbb{N}.$$

In what follows we assume that $\mu_b^{\otimes q}(F)>0$, where $\mu_b^{\otimes q}$ stands for the q-fold product measure of μ_b. Given a Rademacher sequence $X=(X_i)_{i\ge1}$, we now define the random variable $S^{(b)}(F)$ by

$$S^{(b)}(F):=\frac{1}{(q!\mu_b^{\otimes q}(F))^{1/2}}\sum_{(i_1,\dots,i_q)\in F}b_{i_1}\cdots b_{i_q}X_{i_1}\cdots X_{i_q}. \tag{6.1}$$

To discuss the distance between $S^{(b)}(F)$ and a standard Gaussian random variable we need further notation. By F_j^* we denote the collection of all $(i_1,\dots,i_q)\in F$ such that $i_k=j$ for some $k\in\{1,\dots,q\}$. Further, let $F^\sharp\subset F\times F$

be defined as follows. A pair $((i_1,\dots,i_q),(j_1,\dots,j_q))$ belongs to $F^\sharp$ if $\{i_1,\dots,i_q\}\cap\{j_1,\dots,j_q\}=\varnothing$ and there are $(k_1,\dots,k_q),(l_1,\dots,l_q)\in F$ such that $\{k_1,\dots,k_q,l_1,\dots,l_q\}=\{i_1,\dots,i_q,j_1,\dots,j_q\}$ with the property that $(k_1,\dots,k_q)$ does not coincide with $(i_1,\dots,i_q)$ or $(j_1,\dots,j_q)$. We finally define the quantities $\Phi^{(b)}(F)$ and $\Psi_j^{(b)}(F)$ by

$$\Phi^{(b)}(F):=\frac{\mu_b^{\otimes 2q}(F^\sharp)^{1/2}}{\mu_b^{\otimes q}(F)}\quad\text{and}\quad\Psi_j^{(b)}(F):=\frac{\mu_b^{\otimes q}(F_j^*)}{\mu_b^{\otimes q}(F)}. \tag{6.2}$$

Theorem 6.2. *Let N be a standard Gaussian random variable. Then there are constants $0<C_1,C_2<\infty$ only depending on q such that*

$$d_K\big(S^{(b)}(F),N\big)\le C_1\Phi^{(b)}(F)+C_2\Big(\sup_{j\ge1}\Psi_j^{(b)}(F)\Big)^{1/4}.$$

Proof. We notice that $S^{(b)}(F)$ can be written as a discrete multiple stochastic integral $J_q(f)$ of order q with kernel function

$$f(i_1,\dots,i_q)=\frac{b_{i_1}\cdots b_{i_q}}{(q!\mu_b^{\otimes q}(F))^{1/2}}\mathbb{1}_{\{(i_1,\dots,i_q)\in F\}}.$$

Furthermore, $S^{(b)}(F)$ has unit variance so that the first inequality of Theorem 4.2 implies that $d_K(S^{(b)}(F),N)$ is bounded from above by

$$C\max\Big\{\max_{r=1,\dots,q-1}\big\{\big\|(f\star_r^r f)\mathbb{1}_{\Delta_{2(q-r)}}\big\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}}\big\},\max_{r=1,\dots,q}\big\{\big\|f\star_r^{r-1}f\big\|_{\ell^2(\mathbb{N})^{\otimes(2(q-r)+1)}}\big\}\Big\}$$

with a constant $0<C<\infty$ only depending on q. These contraction norms can be computed exactly as in the proof of Theorem 6.4 in [26], so that we leave out the details. This gives the result. □

We now specialize Theorem 6.2 to the set-up discussed in [5]. For this, let $(F_n)_{n\ge1}$ be a sequence of non-empty subsets of $\mathbb{N}^q$ satisfying the following two properties:

(i) $F_n\subset\Delta_q\cap\{1,\dots,n\}^q$ for all $n\in\mathbb{N}$,
(ii) F_n is symmetric in that $(i_1,\dots,i_q)\in F_n$ implies that $(i_{\sigma(1)},\dots,i_{\sigma(q)})\in F_n$ for all permutations σ of $\{1,\dots,q\}$.

Further let $b_0=(b_i)_{i\ge1}$ be such that $b_1=\cdots=b_q=1$ and $b_i=0$ for all $i\ge q+1$. The sequence $(S_n)_{n\ge1}$ of random variables is then given by $S_n:=S^{(b_0)}(F_n)$ with $S^{(b_0)}(F_n)$ as at (6.1). Note that the measure μ_{b_0} reduces to the counting measure restricted to $\{1,\dots,q\}$ and we write $\Phi(F_n)$ and $\Psi_j(F_n)$ in (6.2) instead of $\Phi^{(b_0)}(F_n)$ and $\Psi_j^{(b_0)}(F_n)$, respectively. For this setting, Theorem 6.2 yields the following Berry–Esseen bound, which improves Theorem 1.7 in [5] and Theorem 6.4 in [26].

Corollary 6.1. *Let N be a standard Gaussian random variable. Then there are constants $0<C_1,C_2<\infty$ only depending on q such that*

$$d_K(S_n,N)\le C_1\Phi(F_n)+C_2\Big(\max_{1\le j\le n}\Psi_j(F_n)\Big)^{1/4}$$

for all $n\in\mathbb{N}$.

A concrete situation to which Corollary 6.1 can be applied concerns fractional Cartesian products. We briefly recall their definition and refer to [4] for further details. Fix integers $q\ge3$ and $m\in\{2,\dots,q\}$, and let $\{M_1,\dots,M_q\}$ be a collection of distinct non-empty subsets of $[q]:=\{1,\dots,q\}$ with the following properties:

(i) $\{M_1,\dots,M_q\}$ is a connected cover of $[q]$, i.e., $\bigcup_{i=1}^q M_i=[q]$ and $\{M_1,\dots,M_q\}$ cannot be partitioned into two disjoint partial covers,

(ii) each subset has exactly m elements,
(iii) each index $j \in [q]$ appears in exactly m of the subsets $M_1, \dots, M_q$.

For a subset $M \subset [q]$ and a vector $\mathbf{y} = (y_1, \dots, y_q)$ we write $\pi_M \mathbf{y} := (y_j\colon\ j \in M)$ for the projection of $\mathbf{y}$ on M. Now, let for an integer $n \geq q^m$, $K := \max\{k \in \mathbb{N}\colon\ k \leq n^{1/m}\}$, fix a one-to-one map $\varphi : [K]^m \to [n]$ and define

$$F_n^{**} := \big\{\big(\varphi(\pi_{M_1}\mathbf{y}), \dots, \varphi(\pi_{M_q}\mathbf{y})\big)\colon\ \mathbf{y} = (y_1, \dots, y_q) \in [K]^q\big\} \subset [n]^q.$$

In general, F_n^{**} is not symmetric and may contain diagonal elements. We thus define

$$F_n := \big\{(i_1, \dots, i_q) \in \Delta_q^n\colon\ (i_{\sigma(1)}, \dots, i_{\sigma(q)}) \in F_n^{**} \cap \Delta_q^n \text{ for some permutation } \sigma \text{ of } [q]\big\},$$

where Δ_q^n stands for $\Delta_q \cap \{1, \dots, n\}^q$. The sequence $(F_n)_{n\geq 1}$ is what is called a fractional Cartesian product in the literature. The name comes from the fact that F_n has (fractional) combinatorial dimension q/m, a notion for which we refer to [3]. We take such a fractional Cartesian product as in (6.1) as input sets for our random variables and write S_n^{fCP} in this case. Then Corollary 6.1 and the computations leading to Proposition 6.6 in [26] imply the following Berry–Esseen bound.

Corollary 6.2. *Let N be a standard Gaussian random variable. Then there is a constant $0 < C < \infty$ such that*

$$d_K\big(S_n^{\mathrm{fCP}}, N\big) \leq C n^{-1/(2m)}$$

for all $n \geq 1$.

Remark 6.1. *The rate in Corollary 6.2 for the Kolmogorov distance is the same as in Proposition* 6.6 *in* [26], *where the authors considered a probability metric based on twice differentiable test functions.*

6.3. *Traces of powers of Bernoulli matrices*

As an application of Corollary 5.1 we consider the normal approximation of a vector of traces of powers of a Bernoulli random matrix. This task has already been accomplished in [22] for more general classes of random matrices whose entries are independent and obey certain moment conditions. In this paper the authors used universality results for homogeneous sums established in [25]. Our Corollary 5.1 offers a bound for the multivariate normal approximation for the special case of Bernoulli random matrices without resorting to universality results. For related limit theorems dealing with traces of random matrices we refer to [1,12] and the references cited therein.

As already indicated in [26] the results of discrete Malliavin calculus and Stein's method extend to Rademacher random variables indexed by arbitrary discrete sets. We make use of this possibility by considering a doubly indexed collection of i.i.d. Rademacher random variables $(X_{ij})_{i,j\in\mathbb{N}}$. The object of interest is the Bernoulli random matrix

$$X_n := \left(\frac{X_{ij}}{\sqrt{n}}\right)_{1\leq i,j\leq n}.$$

Let us denote by

$$\operatorname{trace}\big(X_n^q\big) := n^{-q/2} \sum_{i_1,i_2,\dots,i_q=1}^{n} X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_q i_1}$$

the trace of the qth power of X_n. Recall the definition of the d_4-distance in (2.20).

Theorem 6.3. *Let $d \geq 2$ and $1 \leq q_1 < \cdots < q_d$ be integers. Define the random vector*

$$\mathbf{F}_n = \big(\operatorname{trace}\big(X_n^{q_1}\big) - \mathbb{E}\big[\operatorname{trace}\big(X_n^{q_1}\big)\big], \dots, \operatorname{trace}\big(X_n^{q_d}\big) - \mathbb{E}\big[\operatorname{trace}\big(X_n^{q_d}\big)\big]\big)$$

and let $\mathbf{N}$ *be a centred Gaussian random vector with covariance matrix* $\Sigma = (\sigma_{ij})_{i,j=1}^{d}$ *such that* $\sigma_{ii} = q_i$ *for* $1 \le i \le d$ *and* $\sigma_{ij} = 0$ *for* $1 \le i \ne j \le d$. *Then*

$$d_4(\mathbf{F}_n, \mathbf{N}) \le C n^{-1/4} \tag{6.3}$$

for a constant $0 < C < \infty$ *depending only on* $q_1, q_2, \dots, q_d$ *and* d.

Proof. In order to apply Corollary 5.1 we have to express $\operatorname{trace}(X_n^{q_i})$ in terms of elements of some fixed Rademacher chaos. To this end, the following decomposition taken from [22] is crucial. It holds that

$$\operatorname{trace}\bigl(X_n^q\bigr) = n^{-q/2} \sum_{(i_1,\dots,i_q)\in D_n^{(q)}} X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_q i_1} + n^{-q/2} \sum_{(i_1,\dots,i_q)\notin D_n^{(q)}} X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_q i_1},$$

where

$$D_n^{(q)} = \bigl\{(i_1,\dots,i_q) \in \{1,\dots,n\}^q\colon\ (i_a, i_{a+1}) \ne (i_b, i_{b+1}) \text{ for } a \ne b\bigr\}$$

with $i_{q+1} := i_1$. This separation of the range of summation is necessary due to the fact that multiple integrals are defined only for kernels that are symmetric and vanish on diagonals. One has

$$n^{-q/2} \sum_{(i_1,\dots,i_q)\in D_n^{(q)}} X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_q i_1} = J_q\bigl(f_n^{(q)}\bigr),$$

where $f_n^{(q)} := \widetilde{f_{q,n}}$ is the canonical symmetrization of

$$f_{q,n}\bigl((a_1,b_1),\dots,(a_q,b_q)\bigr) := n^{-q/2} \sum_{(i_1,\dots,i_q)\in D_n^{(q)}} \mathbb{1}_{\{i_1=a_1, i_2=b_1\}} \mathbb{1}_{\{i_2=a_2, i_3=b_2\}} \cdots \mathbb{1}_{\{i_q=a_q, i_1=b_q\}}.$$

Let $\mathbf{J}_n = (J_1,\dots,J_d)$ denote the random vector $(J_{q_1}(f_n^{(q_1)}),\dots,J_{q_d}(f_n^{(q_d)}))$. Then the triangle inequality for the d_4-distance implies that

$$d_4(\mathbf{F}_n, \mathbf{N}) \le d_4(\mathbf{F}_n, \mathbf{J}_n) + d_4(\mathbf{J}_n, \mathbf{N}). \tag{6.4}$$

From [22], Equation (3.9), one has

$$\bigl|q_i - \mathbb{E}\bigl[J_{q_i}\bigl(f_n^{(q_i)}\bigr)^2\bigr]\bigr| = \mathrm{O}\bigl(n^{-1}\bigr), \tag{6.5}$$

as $n \to \infty$ for all $i = 1,\dots,d$. This has actually been established in [22] for random matrices with independent Gaussian entries. But it is easily checked that the estimate continues to hold for Rademacher random variables. This together with the isometry (2.14) of discrete multiple stochastic integrals implies the information on the covariance structure of $\mathbf{F}_n$.

Next, from [22], equation (3.6), one has

$$\bigl\| f_n^{(q)} \star_r^r f_n^{(q)} \bigr\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} = \mathrm{O}\bigl(n^{-1/2}\bigr) \tag{6.6}$$

for all $r = 1,\dots,q-1$ in the limit, as $n \to \infty$. In view of (6.5) and (6.6), Corollary 5.1 yields

$$d_4(\mathbf{J}_n, \mathbf{N}) = \mathrm{O}\bigl(n^{-1/4}\bigr). \tag{6.7}$$

Furthermore, Proposition 4.1 in [22] states that for all $1 \le i \le d$,

$$\begin{aligned}&\mathbb{E}\bigl[(F_i - J_i)^2\bigr]\\ &\quad= \mathbb{E}\Biggl[\Biggl(n^{-q_i/2} \sum_{(i_1,\dots,i_{q_i})\notin D_n^{(q_i)}} \bigl(X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{q_i} i_1} - \mathbb{E}[X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{q_i} i_1}]\bigr)\Biggr)^2\Biggr] = \mathrm{O}\bigl(n^{-1}\bigr).\end{aligned} \tag{6.8}$$

Let $g:\mathbb{R}^d \to \mathbb{R}$ be an admissible test function for the d_4-distance. Then, writing $\|\cdot\|_{\mathbb{R}^d}$ for the standard Euclidean norm in $\mathbb{R}^d$, we find that

$$\begin{aligned}\big|\mathbb{E}g(\mathbf{F}_n) - \mathbb{E}g(\mathbf{J}_n)\big| &\le M_1(g)\mathbb{E}\big[\|\mathbf{F}_n - \mathbf{J}_n\|_{\mathbb{R}^d}\big] \le \mathbb{E}\left[\left(\sum_{i=1}^d (F_i - J_i)^2\right)^{1/2}\right] \\ &\le \left(\sum_{i=1}^d \mathbb{E}\big[(F_i - J_i)^2\big]\right)^{1/2} = \mathrm{O}\big(n^{-1/2}\big), \end{aligned} \tag{6.9}$$

for a constant $0 < C < \infty$ which is independent of n. Here, we have used the fact that $M_1(g) \le 1$, the Cauchy–Schwarz inequality and (6.8). The decomposition (6.4) together with (6.7) and (6.9) proves the assertion. □

Remark 6.2. *The rate of convergence obtained in Theorem 6.3 by means of the Malliavin–Stein method for Rademacher sequences is of the same order of magnitude as the rate in* [22], *which is based on the universality results in* [25]. *The only difference is that for technical reasons we had to assume a slightly higher degree of differentiability for the test functions in the probability metric*.

Remark 6.3. *For the proof of Theorem 6.3 we have used the second inequality in Corollary 5.1, which involves square-roots of contraction norms. One might hope to improve* (6.3) *by using the first inequality in Corollary 5.1. For this, one would need to extend* (6.6) *to an estimate for norms of mixed contractions. We expect that this requires considerable effort since already the proof of* (6.6) *relies on involved combinatorial decompositions and identities*.

Acknowledgements

The authors have been supported by the German Research Foundation (DFG) via SFB TR-12.

References

[1] G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probab. Theory Related Fields* **134** (2) (2006) 283–338. MR2222385
[2] N. Balakrishnan and M. V. Koutras. *Runs and Scans with Applications*. *Wiley Series in Probability and Statistics*. Wiley-Interscience [John Wiley & Sons], New York, 2002. MR1882476
[3] R. Blei. Combinatorial dimension and certain norms in harmonic analysis. *Amer. J. Math.* **106** (4) (1984) 847–887. MR0749259
[4] R. Blei. *Analysis in Integer and Fractional Dimensions*. *Cambridge Studies in Advanced Mathematics* **71**. Cambridge Univ. Press, Cambridge, 2001. MR1853423
[5] R. Blei and S. Janson. Rademacher chaos: Tail estimates versus limit theorems. *Ark. Mat.* **42** (1) (2004) 13–29. MR2056543
[6] S. Chatterjee. A new method of normal approximation. *Ann. Probab.* **36** (4) (2008) 1584–1610. MR2435859
[7] 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. MR2732624
[8] P. de Jong. A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields* **75** (2) (1987) 261–277. MR0885466
[9] V. H. de la Peña and E. Giné. *Decoupling: From Dependence to Independence*. *Probability and Its Applications (New York)*. Springer-Verlag, New York, 1999. MR1666908
[10] P. Eichelsbacher and C. Thäle New Berry–Esseen bounds for non-linear functionals of Poisson random measures. *Electron. J. Probab.* **19** (102) (2014).
[11] A. P. Godbole. The exact and asymptotic distribution of overlapping success runs. *Comm. Statist. Theory Methods* **21** (4) (1992) 953–967. MR1173302
[12] A. Guionnet. *Large Random Matrices: Lectures on Macroscopic Asymptotics*. *Lecture Notes in Mathematics* **1957**. Springer-Verlag, Berlin, 2009. MR2498298
[13] S. Kwapień and W. A. Woyczyński. *Random Series and Stochastic Integrals: Single and Multiple*. *Probability and Its Applications*. Birkhäuser, Boston, MA, 1992. MR1167198
[14] R. Lachièze-Rey and G. Peccati. Fine Gaussian fluctuations on the Poisson space, I: Contractions, cumulants and geometric random graphs. *Electron. J. Probab.* **18** (32) (2013) MR3035760
[15] 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** (12) (2013) 4186–4218. MR3096352
[16] G. Last, M. D. Penrose, M. Schulte and C. Thäle. Moments and central limit theorems for some multivariate Poisson functionals. *Adv. in Appl. Probab.* **46** (2) (2014) 348–364. MR3215537

[17] A. E. Litvak, A. Pajor, M. Rudelson and N. Tomczak-Jaegermann. Smallest singular value of random matrices and geometry of random polytopes. *Adv. Math.* **195** (2) (2005) 491–523. MR2146352

[18] D. Marinucci and G. Peccati. *Random Fields on the Sphere*. *London Mathematical Society Lecture Note Series* **389**. Cambridge Univ. Press, Cambridge, 2011. MR2840154

[19] E. Mossel, R. O'Donnell and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math. (2)* **171** (1) (2010) 295–341. MR2630040

[20] I. Nourdin and G. Peccati. Stein's method and exact Berry–Esseen asymptotics for functionals of Gaussian fields. *Ann. Probab.* **37** (6) (2009) 2231–2261. MR2573557

[21] I. Nourdin and G. Peccati. Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** (1–2) (2009) 75–118. MR2520122

[22] I. Nourdin and G. Peccati. Universal Gaussian fluctuations of non-Hermitian matrix ensembles: From weak convergence to almost sure CLTs. *ALEA Lat. Am. J. Probab. Math. Stat.* **7** (2010) 341–375. MR2738319

[23] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge Univ. Press, Cambridge, MA, 2012. MR2962301

[24] I. Nourdin and G. Peccati. The optimal fourth moment theorem. *Proc. Amer. Math. Soc.* **143** (7) (2015) 3123–3133. MR3336636

[25] I. Nourdin, G. Peccati and G. Reinert. Invariance principles for homogeneous sums: Universality of Gaussian Wiener chaos. *Ann. Probab.* **38** (5) (2010) 1947–1985. MR2722791

[26] I. Nourdin, G. Peccati and G. Reinert. Stein's method and stochastic analysis of Rademacher functionals. *Electron. J. Probab.* **15** (55) (2010) 1703–1742. MR2735379

[27] I. Nourdin, G. Peccati and A. Réveillac. Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** (1) (2010) 45–58. MR2641769

[28] D. Nualart and S. Ortiz-Latorre. Central limit theorems for multiple stochastic integrals and Malliavin calculus. *Stochastic Process. Appl.* **118** (4) (2008) 614–628. MR2394845

[29] D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (1) (2005) 177–193. MR2118863

[30] G. Peccati, J. L. Solé and F. Utzet. Stein's method and normal approximation of Poisson functionals. *Ann. Probab.* **38** (2) (2010) 443–478. MR2642882

[31] G. Peccati and M. S. Taqqu. *Wiener Chaos: Moments, Cumulants and Diagrams*. *Bocconi & Springer Series* **1**. Springer, Bocconi Univ. Press, Milan, 2011. MR2791919

[32] G. Peccati and C. Thäle. Gamma limits and U-statistics on the Poisson space. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** (1) (2013) 525–560. MR3083936

[33] G. Peccati and C. A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. In *Séminaire de Probabilités XXXVIII* 247–262. *Lecture Notes in Math.* **1857**. Springer, Berlin, 2005. MR2126978

[34] G. Peccati and C. Zheng. Multi-dimensional Gaussian fluctuations on the Poisson space. *Electron. J. Probab.* **15** (48) (2010) 1487–1527. MR2727319

[35] N. Privault. Stochastic analysis of Bernoulli processes. *Probab. Surv.* **5** (2008) 435–483. MR2476738

[36] G. Reinert and A. Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.* **37** (6) (2009) 2150–2173. MR2573554

[37] M. Reitzner and M. Schulte. Central limit theorems for U-statistics of Poisson point processes. *Ann. Probab.* **41** (6) (2013) 3879–3909. MR3161465

[38] 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** (4) (1997) 1080–1105. MR1484798

[39] M. Rudelson and R. Vershynin. The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** (2) (2008) 600–633. MR2407948

[40] M. Schulte. A central limit theorem for the Poisson–Voronoi approximation. *Adv. in Appl. Math.* **49** (3–5) (2012) 285–306. MR3017961

[41] M. Schulte. Normal approximation of Poisson functionals in Kolmogorov distance. *J. Theoret. Probab.* **29** (1) (2016) 96–117. MR3463079

[42] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif. 1970/1971)* 583–602. *Vol. II: Probability Theory*. Univ. California Press, Berkeley, CA, 1972. MR0402873