DISCRETE MALLIAVIN–STEIN METHOD: BERRY–ESSEEN BOUNDS FOR RANDOM GRAPHS AND PERCOLATION[1]

BY KAI KROKOWSKI, ANSELM REICHENBACHS AND CHRISTOPH THÄLE

Ruhr University Bochum

A new Berry–Esseen bound for nonlinear functionals of nonsymmetric and nonhomogeneous infinite Rademacher sequences is established. It is based on a discrete version of the Malliavin–Stein method and an analysis of the discrete Ornstein–Uhlenbeck semigroup. The result is applied to subgraph counts and to the number of vertices having a prescribed degree in the Erdős–Rényi random graph. A further application deals with a percolation problem on trees.

1. Introduction. The Malliavin–Stein method has become a versatile device for proving quantitative limit theorems. It combines the Malliavin calculus of variations with Stein's method. The results obtained this way typically fall into two categories. The first category consists of limit theorems for nonlinear functionals defined on the Wiener space with notable applications to Gaussian random processes, especially the fractional Brownian motion [19, 20, 22], random matrices [21] and random polynomials [1]. The other brand comprises limit theorems for functionals of Poisson random measures and their applications to stochastic geometry [6, 15, 17, 18, 28, 36], U-statistics [4, 6, 16, 28, 33], nonlinear statistics of spherical Poisson fields [5] and the theory of Lévy processes [6, 17, 26].

On the other hand, the Malliavin–Stein method has left only few traces in that part of probability theory in which discrete random structures are investigated. One exception is the paper [24] where Stein's method for normal approximation has been combined with tools from discrete stochastic analysis for symmetric Rademacher sequences to deduce quantitative central limit theorems with respect to probability distances based on smooth test functions. Here, by a symmetric Rademacher sequence we understand an infinite sequence of independent and identically distributed random variables taking the values ± 1 with probability $1/2$ each. This approach has been extended in [14] to deduce Berry–Esseen bounds, that is, estimates for the Kolmogorov distance in related central limit theorems. The applications considered in [14, 24] concern the number of two-runs, a quantitative version of a combinatorial central limit theorem as well as traces of powers of random Bernoulli matrices. While the previously mentioned papers were con-

cerned with the symmetric case, we work with general nonlinear functionals of nonsymmetric and even nonhomogeneous Rademacher sequences in order to bring a rich class of examples, that were not accessible before, within the reach of the Malliavin–Stein method. Moreover, we emphasize that some of the examples we present below are not within the reach of any of the traditional approaches using Stein's method.

One of the main tools of the Malliavin–Stein method on the Wiener or the Poisson space is the so-called multiplication formula for multiple stochastic integrals; cf. [27] for a general overview. The main difficulty in the discrete set-up is that no such multiplication formula for discrete multiple stochastic integrals based on nonsymmetric or nonhomogeneous Rademacher sequences is available. Consequently, a new type of abstract Berry–Esseen bound needs to be developed, which is getting along without this technical device. Such a result, namely Theorem 4.1 below, is one of our main contributions. It can be interpreted as a kind of 'second-order Poincaré inequality' and is the discrete analogue of corresponding results on the Wiener or the Poisson space; cf. [17, 23]. It relies on a generalization of the Malliavin–Stein bound established in [14] and on an analysis of the discrete Ornstein–Uhlenbeck semigroup. To make this approach work, we also have to develop further some facets of the discrete Malliavin calculus of variations. In particular, we present a generalization of the integration-by-parts formula, which is one of our crucial devices. The Berry–Esseen bound we obtain this way is particularly well suited for the study of discrete random structures. This is due to the fact that the chaotic decomposition of the functional at hand does not have to be specified. Instead, the impact of local perturbations on the functional measured by means of a certain difference operator (discrete Malliavin derivative) has to be evaluated. A sufficient condition for asymptotic normality is that moments of first- and second-order discrete Malliavin derivatives of the functional are sufficiently small.

To highlight the versatility of our general limit theorem, we now present a couple of concrete applications. The first one deals with the *triangle counting statistic* associated with the *Erdős–Rényi random graph*. Introduced in [7], the model has since then become one of the most popular models in discrete probability; cf. [10] for an exhaustive list of references. Informally, the random graph $G(n, p)$ is a graph on $n \in \mathbb{N}$ vertices in which each edge between two vertices is included with probability $p \in [0, 1]$, independently of the other edges (for a detailed construction see Section 5 below and see Figure 1 for simulations). In what follows, we allow p also to depend on n, but for practical reasons we suppress this in our notation. The random variable in the focus of our attention is the number $T = T(n, p)$ of triangles in $G(n, p)$, that is, the number of sub-graphs of $G(n, p)$ that are isomorphic to the complete graph on 3 vertices. A comprehensive central limit theorem for the normalized random variable $F := (T - \mathbb{E}[T])/\sqrt{\text{Var}[T]}$ has been derived in [34] by the method of moments. In particular, it provides a necessary and sufficient condition on n and p, which ensures asymptotic Gaussianity for F. Namely,
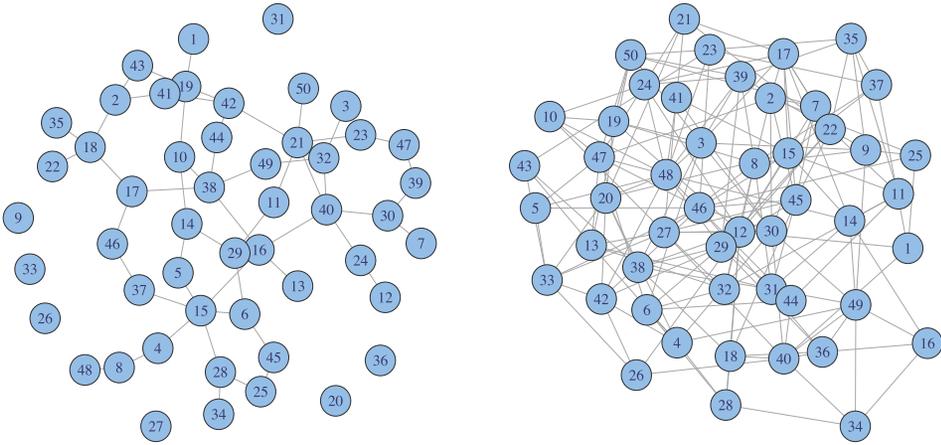
FIG. 1. *Realizations of Erdős–Rényi random graphs with $n = 50$ vertices and $p = 0.04$ (left) and $p = n^{-1/2} \approx 0.14$ (right). The graphics were produced by means of the freely available* R-*package* igraph.

as $n \to \infty$, one has that

$$F \xrightarrow{d} N \quad \text{if and only if} \quad np \to \infty \quad \text{and} \quad n^2(1-p) \to \infty,$$

where $N \sim \mathcal{N}(0, 1)$ is a standard Gaussian random variable and $\xrightarrow{d}$ indicates convergence in distribution. Using Stein's method for normal approximation, a rate of convergence in this central limit theorem measured by some sort of bounded Wasserstein distance has been established in [2]. If $p \in (0, 1)$ is fixed,

$$d_1(F, N) := \sup_{h \in \mathcal{H}} \frac{|\mathbb{E}[h(F)] - \mathbb{E}[h(N)]|}{\|h\|_\infty + \|h'\|_\infty} = \mathcal{O}(n^{-1}),$$

where $\mathcal{H}$ is the class of bounded functions $h : \mathbb{R} \to \mathbb{R}$ with bounded first derivative and where $\| \cdot \|_\infty$ denotes the supremum norm. For the case that $p = \theta n^{-\alpha}$ with $\alpha \in (0, 1)$ and $\theta \in (0, n^\alpha)$ such that $\theta \asymp 1$ the result in [2] delivers the bound

$$d_1(F, N) = \begin{cases} \mathcal{O}(n^{-1+\alpha/2}), & \text{if } 0 < \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-3(1-\alpha)/2}), & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

We use the following standard notation for comparing the order of magnitude of two real sequences: We write $a_n \asymp b_n$ for two real sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ whenever

$$c \leq \liminf_{n\to\infty} \left|\frac{a_n}{b_n}\right| \leq \limsup_{n\to\infty} \left|\frac{a_n}{b_n}\right| \leq C$$

for two constants $0 < c \leq C < \infty$. We also write $a_n = \mathcal{O}(b_n)$ for two nonnegative sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ if there is a constant $c \in (0, \infty)$ such that $a_n \leq cb_n$

for sufficiently large n. Applying a standard smoothing argument, one can show that the more prominent and more natural Kolmogorov distance

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

between F and N is bounded by a constant multiple of the *square-root* of $d_1(F, N)$; cf. [29], Proposition 2.4. However, this typically leads to a suboptimal rate of convergence for the Kolmogorov distance $d_K(F, N)$. For example, in the special case of a fixed $p \in (0, 1)$ one expects that also $d_K(F, N)$ is of order n^{-1}. Our main contribution in this context is the following Berry–Esseen bound, which in particular confirms that this is in fact true. We emphasize that we are not aware of any other technique, which could be used to provide bounds on the Kolmogorov distance of this quality if p is of the form $\theta n^{-\alpha}$ with $\alpha \in (0, 1)$ and $\theta \in (0, n^\alpha)$ such that $\theta \asymp 1$. In what follows, we treat both set-ups simultaneously and contribute thereby to a long standing problem in this area.

THEOREM 1.1. *Denote by* $N \sim \mathcal{N}(0, 1)$ *a standard Gaussian random variable. Let* $p = \theta n^{-\alpha}$ *with* $\alpha \in [0, 1)$ *and* $\theta = \theta_n \in (0, n^\alpha)$ *such that* $\theta \asymp 1$*. Then*

$$d_K(F, N) = \begin{cases} \mathcal{O}(n^{-1+\alpha}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-3/4+\alpha/2}), & \text{if } \frac{1}{2} < \alpha \leq \frac{2}{3}, \\ \mathcal{O}(n^{-5(1-\alpha)/4}), & \text{if } \frac{2}{3} < \alpha < 1. \end{cases}$$

In particular, if p *is constant, that is, if* $\alpha = 0$,

$$d_K(F, N) = \mathcal{O}(n^{-1}).$$

To underline that not only triangle counts are within the reach of our methods, we now consider the problem of *counting copies of general sub-graphs* Γ *in the Erdős–Rényi random graph* $G(n, p)$. Formally, we denote by $S = S(n, p)$ the number of copies of Γ in $G(n, p)$ and by $F := (S - \mathbb{E}[S])/\sqrt{\text{Var}[S]}$ the normalized sub-graph counting statistic. We assume that Γ has at least one edge and in contrast to Theorem 1.1 we also assume that the success probability $p \in (0, 1)$ is fixed and does not depend on n. In this situation, Theorem 2 in [2] says that

$$d_1(F, N) = \mathcal{O}(n^{-1})$$

and our abstract Berry–Esseen bound can be used to conclude that the d_1-distance can be replaced by the Kolmogorov distance.

THEOREM 1.2. *Denote by* $N \sim \mathcal{N}(0, 1)$ *a standard Gaussian random variable and fix* $p \in (0, 1)$*. Then*

$$d_K(F, N) = \mathcal{O}(n^{-1})$$

for all graphs Γ *having at least one edge.*

It should be pointed out that Theorem 1.2 can alternatively be obtained from the sharp cumulant estimate in [8], Proposition 10.3, together with [35], Corollary 2.1 or from the Berry–Esseen bound for decomposable random variables in [32], Theorem 5.1.

Besides the number of triangles or general sub-graphs, there are several other random variables associated with the Erdős–Rényi random graph that have found considerable attention in the literature. One statistic that has been object of much study is the *number of vertices having a prescribed degree*. For example, in [12] a central limit theorem for the number of isolated vertices was given, which for general degree is a result in [11]. A rate of convergence for the d_1-distance as introduced above has been obtained in [2]. A technically highly sophisticated version of Stein's method was developed in [9] to deduce a corresponding Berry–Esseen bound in case that the success probability is $p = \theta/n$. Using our general Berry–Esseen bound, we are able to present a quick and streamlined proof of an extended version of this quantitative central limit theorem. We denote for $d \in \{0, 1, 2, \ldots\}$ by $V_{n,d}$ the number of vertices of degree d in $G(n, p)$ in case that the success probability satisfies $p = \theta n^{-\alpha}$ for suitable $\alpha \in \mathbb{R}$ and $\theta \in (0, n^\alpha)$ such that $\theta \asymp 1$. We finally define the normalized random variable $G_{n,d} := (V_{n,d} - \mathbb{E}[V_{n,d}])/\sqrt{\mathrm{Var}[V_{n,d}]}$, $n \in \mathbb{N}$.

THEOREM 1.3.   *Denote by $N \sim \mathcal{N}(0, 1)$ a standard Gaussian variable and fix $d \in \{0, 1, 2, \ldots\}$. Let $p = \theta n^{-\alpha}$ with $\alpha \in [1, 2)$ and $\theta \in (0, n^\alpha)$ such that $\theta \asymp 1$. Then*

$$d_K(G_{n,d}, N) = \begin{cases} \mathcal{O}(n^{-1+\alpha/2}), & \text{if } d = 0, \alpha \in [1, 2), \\ \mathcal{O}(n^{1/2-3d/2-\alpha+3\alpha d/2}), & \text{if } d \in \mathbb{N}, \alpha \in \left[1, \dfrac{3d-1}{3d-2}\right). \end{cases}$$

In particular, if $\alpha = 1$,

$$d_K(G_{n,d}, N) = \mathcal{O}(n^{-1/2})$$

for all $d \in \{0, 1, 2, \ldots\}$.

Our final application deals with the *number of connected components* arising from *bond percolation on a tree*. We recall that a rooted tree $\mathscr{T}$ is an undirected graph with one distinguished vertex, the root of the tree, in which any two vertices are connected by a unique self-avoiding path. We denote for $n \in \mathbb{N}$ by $\mathscr{T}_n$ the sub-tree of $\mathscr{T}$, which consists of all vertices of $\mathscr{T}$ that have graph-distance at most n from the root. By $|\mathscr{T}_n|$, we denote the number of edges of $\mathscr{T}_n$. In what follows, we assume that each vertex of $\mathscr{T}$ has degree bounded by $D + 1$ with $D \in \mathbb{N}$ and that $\mathscr{T}$ has infinitely many vertices. If the degree of the root is D and if the degree of each other vertex of $\mathscr{T}$ is $D + 1$ for some fixed $D \in \mathbb{N}$, we say that $\mathscr{T}$ is a D-regular tree. Fix $p \in (0, 1)$ and assign to each edge e of $\mathscr{T}$, independently of the
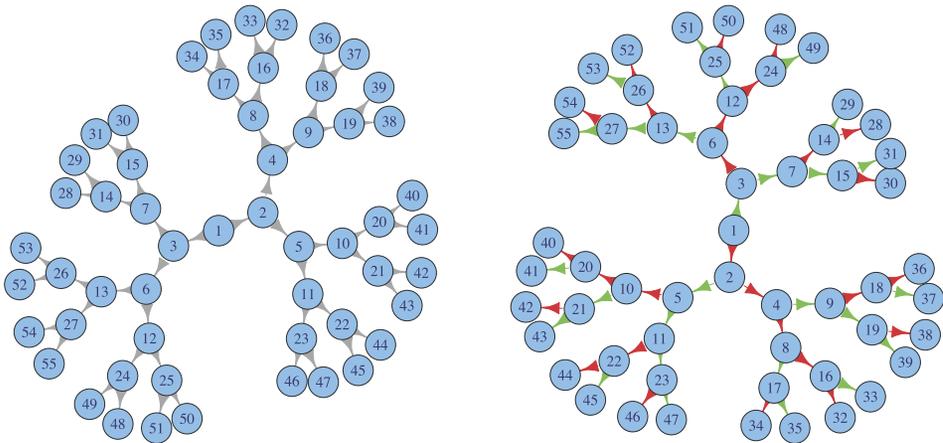
FIG. 2. $\mathscr{T}_5$ of a 2-regular tree $\mathscr{T}$ (left). Realization of $\mathscr{T}_5(p)$ with $p = 1/2$ (right). The colour red means that the edge is included, while green indicates that the edge has been removed. The graphics were produced by means of the freely available R-package igraph.

other edges, a Rademacher random variable X_e such that $P(X_e = +1) = p$ and $P(X_e = -1) = 1 - p$. We now remove from $\mathscr{T}$ all edges e for which $X_e = -1$ and indicate by $\mathscr{T}(p)$ the resulting random graph; see Figure 2 for a simulation. Its restriction to $\mathscr{T}_n$ is denoted by $\mathscr{T}_n(p)$ and we let $C_n(p)$ be the number of connected components of $\mathscr{T}_n(p)$. Here, by a connected component we understand a maximal connected sub-graph of $\mathscr{T}_n(p)$ consisting of at least one edge; isolated vertices are not counted. Our next result is a Berry–Esseen bound for the normalized random variable $H_n(p) := (C_n(p) - \mathbb{E}[C_n(p)])/\sqrt{\mathrm{Var}[C_n(p)]}$. This adds to the qualitative central limit theorem in [37].

THEOREM 1.4. *Fix* $p \in (0, 1)$ *and denote by* $N \sim \mathcal{N}(0, 1)$ *a standard Gaussian random variable. Then*

$$d_K\big(H_n(p), N\big) = \mathcal{O}\big(|\mathscr{T}_n|^{-1/2}\big).$$

In particular, in case of a D-regular tree one has that

$$d_K\big(H_n(p), N\big) = \begin{cases} \mathcal{O}\big(n^{-1/2}\big), & \text{if } D = 1, \\ \mathcal{O}\big(D^{-n/2}\big), & \text{if } D \geq 2. \end{cases}$$

The rest of this paper is structured as follows. In Section 2, we collect some background material related to the discrete Malliavin calculus. An analysis of the discrete Ornstein–Uhlenbeck semigroup is the content of Section 3. This is used in Section 4 to derive our abstract Berry–Esseen bound, which in turn is applied in Section 5 to the Erdős–Rényi random graph and in Section 6 to the percolation problem on trees. These sections also contain the proofs of Theorems 1.1–1.4 presented above.

Note added in proof: After submission of the paper, it came to our attention that a multiplication formula for discrete multiple stochastic integrals based on nonsymmetric and nonhomogeneous Rademacher sequences has been developed in a manuscript by Privault and Torrisi that was not available to us, but has now appeared as [31]. See also [13].

2. Preliminaries.

2.1. *Set-up.* For each $k \in \mathbb{N}$ let $0 < p_k < 1$ and put $q_k := 1 - p_k$. We abbreviate the sequences $(p_k)_{k\in\mathbb{N}}$ and $(q_k)_{k\in\mathbb{N}}$ by p and q, respectively. By $X := (X_k)_{k\in\mathbb{N}}$ we denote a sequence of independent random variables such that

$$P(X_k = +1) = p_k \quad \text{and} \quad P(X_k = -1) = q_k, \qquad k \in \mathbb{N}.$$

This is what we call a (nonsymmetric and nonhomogeneous) sequence of independent Rademacher random variables. We construct them in the canonical way by taking $(\Omega, \mathcal{F}, P)$ as probability space, where $\Omega := \{-1, +1\}^{\mathbb{N}}$, $\mathcal{F} := \mathcal{P}(\{-1, +1\})^{\otimes\mathbb{N}}$ and $P := \bigotimes_{k=1}^{\infty}(p_k\delta_{+1} + q_k\delta_{-1})$, with $\mathcal{P}(M)$ being the power set of a set M and $\delta_{\pm 1}$ being the unit-mass Dirac measure concentrated at ± 1. We then put $X_k(\omega) := \omega_k$ for each $k \in \mathbb{N}$ and $\omega := (\omega_k)_{k\in\mathbb{N}} \in \Omega$. Note that X_k has mean $p_k - q_k$ and variance $4p_k q_k$.

2.2. *Discrete multiple stochastic integrals.* Denote by κ the counting measure on $\mathbb{N}$ and put $\ell^2(\mathbb{N})^{\otimes n} := L^2(\mathbb{N}^n, \mathcal{P}(\mathbb{N})^{\otimes n}, \kappa^{\otimes n})$ for $n \in \mathbb{N}$. In the following, we refer to the elements of $\ell^2(\mathbb{N})^{\otimes n}$ as kernels. By $\ell^2(\mathbb{N})^{\circ n}$ we denote the class of symmetric kernels and $\ell_0^2(\mathbb{N})^{\circ n}$ stands for the sub-class of symmetric kernels vanishing on diagonals, that is, vanishing on the complement of the set $\Delta_n := \{(i_1, \ldots, i_n) \in \mathbb{N}^n : i_k \neq i_\ell \text{ for } k \neq \ell\}$. We further put $\ell^2(\mathbb{N})^{\otimes 0} := \mathbb{R}$.

For $n \in \mathbb{N}$ and a kernel $f \in \ell_0^2(\mathbb{N})^{\circ n}$, we define the discrete multiple stochastic integral of order n of f as

$$J_n(f) := n! \sum_{1 \leq i_1 < \cdots < i_n < \infty} f(i_1, \ldots, i_n) Y_{i_1} \cdots Y_{i_n},$$

where $(Y_k)_{k\in\mathbb{N}}$ with $Y_k := (X_k - p_k + q_k)/(2\sqrt{p_k q_k})$ stands for the normalized sequence of independent Rademacher random variables as introduced above. We also put $J_0(c) = c$ for $c \in \mathbb{R}$. The space spanned by the random variables of the form $J_n(f)$ with $f \in \ell_0^2(\mathbb{N})^{\circ n}$ is called the Rademacher chaos of order n.

Discrete multiple stochastic integrals of different orders are mutually orthogonal and satisfy the isometry relation

$$(2.1) \qquad \mathbb{E}\big[J_n(f)J_m(g)\big] = \mathbb{1}_{\{n=m\}} n! \langle f, g \rangle_{\ell^2(\mathbb{N})^{\otimes n}}$$

for all $n, m \in \mathbb{N}$ and kernels $f \in \ell_0^2(\mathbb{N})^{\circ n}$, $g \in \ell_0^2(\mathbb{N})^{\circ m}$. Moreover, it is a classical fact that every $F \in L^2(\Omega)$ (i.e., every square-integrable Rademacher functional)

admits a chaotic decomposition

$$(2.2) \qquad F = \mathbb{E}[F] + \sum_{n=1}^{\infty} J_n(f_n)$$

for uniquely determined kernels $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$, where the series converges in $L^2(\Omega)$; cf. [30], Proposition 6.7. Together with the isometry relation for discrete multiple stochastic integrals this decomposition implies that the variance of F is given by

$$\mathrm{Var}[F] = \sum_{n=1}^{\infty} n! \|f_n\|_{\ell^2(\mathbb{N})^{\otimes n}}^2.$$

2.3. *Malliavin calculus.* In this section, we introduce some basic notions from discrete Malliavin calculus and refer to [30] for further details and background material. Let $F \in L^2(\Omega)$. The discrete gradient of F in direction $k \in \mathbb{N}$ is given by

$$(2.3) \qquad D_k F := \sqrt{p_k q_k}\big(F_k^+ - F_k^-\big),$$

where $F_k^{\pm}(\omega) := F(\omega_1, \ldots, \omega_{k-1}, \pm 1, \omega_{k+1}, \ldots)$. Note that the normalization in (2.3) is chosen such that $D_k Y_k = 1$.

The discrete gradient satisfies the following product formula. Namely, if $F, G \in L^2(\Omega)$ and $k \in \mathbb{N}$, then

$$(2.4) \qquad D_k(FG) = (D_k F)G + F(D_k G) - \frac{X_k}{\sqrt{p_k q_k}}(D_k F)(D_k G),$$

see [30], Proposition 7.8. We remark that in contrast to classical Malliavin calculus (see [25]), the product formula in the discrete set-up carries the additional term

$$-(X_k/\sqrt{p_k q_k})(D_k F)(D_k G),$$

which is not present in the continuous framework. A similar effect also happens on the Poisson space; cf. [26] and the references cited therein.

The iterated discrete gradient $D^n F := (D_{k_1,\ldots,k_n}^n F)_{k_1,\ldots,k_n \in \mathbb{N}}$ of order $n \geq 2$ is defined by $D_{k_1,\ldots,k_n}^n F := D_{k_n}(D_{k_1,\ldots,k_{n-1}}^{n-1} F)$ for $k_1, \ldots, k_n \in \mathbb{N}$, where we put $D_k^1 F := D_k F$.

We now present a formula which allows to compute the kernels f_n in a chaotic decomposition as in (2.2). In the framework of classical Malliavin calculus, this is known as Strook's formula. Since we have not found such a result for general Rademacher functionals in the literature, we provide the detailed arguments (for the special symmetric case see Lemma 2.2 in [14] and Section 2.4 in [24]).

PROPOSITION 2.1 (Strook's formula). *Assume that* $F \in L^2(\Omega)$ *has chaotic decomposition* $F = \mathbb{E}[F] + \sum_{n=1}^{\infty} J_n(f_n)$. *Then for every* $n \in \mathbb{N}$ *it holds that*

$$(2.5) \qquad \mathbb{E}\big[D_{k_1,\ldots,k_n}^n F\big] = \mathbb{E}[F Y_{k_1} \cdots Y_{k_n}], \qquad (k_1, \ldots, k_n) \in \Delta_n,$$

and

$$\mathbb{E}\big[D^n_{k_1,\dots,k_n} F\big] = n! f_n(k_1,\dots,k_n), \qquad (k_1,\dots,k_n) \in \mathbb{N}^n. \tag{2.6}$$

PROOF. We start by proving (2.5) by induction. Choosing $G = Y_k$ with $D_k G = D_k Y_k = 1$ in (2.4) yields

$$D_k(FY_k) = (D_k F)Y_k + F - \frac{X_k}{\sqrt{p_k q_k}} D_k F$$

and hence

$$D_k(FY_k)Y_k = (D_k F)Y_k^2 + FY_k - \frac{X_k Y_k}{\sqrt{p_k q_k}} D_k F. \tag{2.7}$$

It immediately follows from (2.3) that, for every $F \in L^2(\Omega)$, $D_k F$ is independent of X_k. Therefore, by taking expectations on both sides of (2.7) and computing $\mathbb{E}[X_k Y_k] = 2\sqrt{p_k q_k}$, we get

$$0 = \mathbb{E}[D_k F] + \mathbb{E}[FY_k] - 2\mathbb{E}[D_k F] = \mathbb{E}[FY_k] - \mathbb{E}[D_k F],$$

which proves (2.5) for $n = 1$. Now, assume that (2.5) holds for some fixed $n \in \mathbb{N}$ and consider

$$\mathbb{E}\big[D^{n+1}_{k_1,\dots,k_{n+1}} F\big] = \mathbb{E}\big[D_{k_{n+1}} \big(D^n_{k_1,\dots,k_n} F\big)\big].$$

From the case $n = 1$ treated above, it follows that

$$\mathbb{E}\big[D^{n+1}_{k_1,\dots,k_{n+1}} F\big] = \mathbb{E}\big[\big(D^n_{k_1,\dots,k_n} F\big)Y_{k_{n+1}}\big]$$

and since $Y_{k_{n+1}}$ behaves like a constant from the point of view of $D^n_{k_1,\dots,k_n}$, our assumption leads to

$$\mathbb{E}\big[D^{n+1}_{k_1,\dots,k_{n+1}} F\big] = \mathbb{E}\big[D^n_{k_1,\dots,k_n}(FY_{k_{n+1}})\big] = \mathbb{E}[FY_{k_1} \cdots Y_{k_{n+1}}],$$

which completes the proof of (2.5). Identity (2.6) then immediately follows from (2.2) for $(k_1,\dots,k_n) \in \Delta_n$. For $(k_1,\dots,k_n) \in \Delta_n^c$ both sides of (2.6) are equal to zero. $\square$

For the rest of this section, let $F \in L^2(\Omega)$ have chaotic decomposition $F = \mathbb{E}[F] + \sum_{n=1}^{\infty} J_n(f_n)$ with kernels $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$ for $n \in \mathbb{N}$. Since $D_k F \in L^2(\Omega)$ for every $k \in \mathbb{N}$, the discrete gradient also has a chaotic decomposition. Note that the kernels of this decomposition can be deduced from the chaotic decomposition of F using Stroock's formula. More precisely, the nth kernel of the chaotic decomposition of $D_k F$ evaluated at $(k_1,\dots,k_n) \in \mathbb{N}^n$ is given by

$$\frac{1}{n!}\mathbb{E}\big[D^n_{k_1,\dots,k_n}(D_k F)\big] = \frac{1}{n!}\mathbb{E}\big[D^{n+1}_{k_1,\dots,k_n,k} F\big] = (n+1) f_{n+1}(k_1,\dots,k_n,k).$$

Thus, the discrete gradient can be written as

$$(2.8) \qquad D_k F = \sum_{n=1}^{\infty} n J_{n-1}(f_n(\cdot, k)),$$

where $f_n(\cdot, k) \in \ell_0^2(\mathbb{N})^{\circ n-1}$ denotes the kernel f_n with one of its components fixed, thus acting as function in $n - 1$ variables. For $F \in L^2(\Omega)$ as above and $m \in \mathbb{N}$, we say that $F \in \mathrm{dom}(D^m)$, if

$$\mathbb{E}\big[\|D^m F\|^2_{\ell^2(\mathbb{N})^{\otimes m}}\big] = \sum_{n=m}^{\infty} \left(\frac{n!}{(n-m)!}\right)^2 (n-m)! \|f_n\|^2_{\ell^2(\mathbb{N})^{\otimes n}} < \infty.$$

Next, we define the Ornstein–Uhlenbeck operator L and its (pseudo-)inverse L^{-1}. The domain of L is the class of all $F \in L^2(\Omega)$ for which

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

For $F \in \mathrm{dom}(L)$, we put

$$LF := -\sum_{n=1}^{\infty} n J_n(f_n).$$

The discrete Ornstein–Uhlenbeck semigroup $(P_t)_{t \geq 0}$ associated with L is defined as

$$(2.9) \qquad P_t F := \sum_{n=0}^{\infty} e^{-nt} J_n(f_n), \qquad t \geq 0.$$

The properties of this semigroup will be discussed in detail in Section 3 below. Moreover, for centred $F \in L^2(\Omega)$ we put

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

and call L^{-1} the (pseudo-)inverse of the Ornstein–Uhlenbeck operator L.

Furthermore, we introduce the discrete divergence operator δ and its domain $\mathrm{dom}(\delta)$. For $u := (u_k)_{k \in \mathbb{N}} \in (L^2(\Omega))^{\mathbb{N}}$ with

$$u_k := \sum_{n=0}^{\infty} J_n\big(g_{n+1}(\cdot, k)\big),$$

where $g_{n+1} \in \ell_0^2(\mathbb{N})^{\circ n} \otimes \ell^2(\mathbb{N})$ for $n \in \mathbb{N}$, we say that $u \in \mathrm{dom}(\delta)$, if

$$(2.10) \qquad \sum_{n=0}^{\infty} (n+1)! \|\widetilde{g_{n+1}} \mathbb{1}_{\Delta_{n+1}}\|^2_{\ell^2(\mathbb{N})^{\otimes n+1}} < \infty.$$

Here, and in the following $\widetilde{f}(k_1, \ldots, k_n) := \frac{1}{n!} \sum_\sigma f(k_{\sigma(1)}, \ldots, k_{\sigma(n)})$ denotes the canonical symmetrization of a function f in n variables, where the sum runs over all permutations σ of the set $\{1, \ldots, n\}$.

For $u \in \mathrm{dom}(\delta)$, the discrete divergence operator is defined as

$$\delta(u) := \sum_{n=0}^{\infty} J_{n+1}(\widetilde{g_{n+1}} \mathbb{1}_{\Delta_{n+1}}).$$

Note that, for $u \in \mathrm{dom}(\delta)$, (2.10) is equivalent to

$$\mathbb{E}\big[\delta(u)^2\big] < \infty.$$

As the adjoint of the discrete gradient, δ satisfies the integration-by-parts formula

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

for $F \in \mathrm{dom}(D)$ and $u \in \mathrm{dom}(\delta)$; cf. [30], Proposition 9.2. The operators D, L and δ are related by the identity

(2.12) $$-\delta D = L.$$

In this paper, we make use of the following crucial consequence of (2.11) and (2.12). If $f : \mathbb{R} \to \mathbb{R}$ is measurable and $F \in L^2(\Omega)$ is centred with $f(F) \in \mathrm{dom}(D)$, then

(2.13) $$\mathbb{E}\big[Ff(F)\big] = \mathbb{E}\big[\langle Df(F), -DL^{-1}F\rangle_{\ell^2(\mathbb{N})}\big].$$

Indeed, using (2.11) and (2.12) we have that

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

Now, we present an analogue of the integration-by-parts formula (2.11) for functionals $F \in L^2(\Omega)$ that do not necessarily belong to $\mathrm{dom}(D)$ (we refer to Lemma 2.2 in [17] for a related result on the Poisson space).

PROPOSITION 2.2. *Let* $F \in L^2(\Omega)$. *Furthermore, let* $u := (u_k)_{k \in \mathbb{N}} \in (L^2(\Omega))^{\mathbb{N}}$ *with*

$$u_k := \sum_{n=0}^{\infty} J_n\big(g_{n+1}(\cdot, k)\big),$$

where $g_{n+1} \in \ell_0^2(\mathbb{N})^{\circ n} \otimes \ell^2(\mathbb{N})$ *for* $n \in \mathbb{N}$ *and*

(2.14) $$\sum_{n=0}^{\infty} (n+1)! \|g_{n+1}\|^2_{\ell^2(\mathbb{N})^{\otimes n+1}} < \infty.$$

Further assume that $(D_k F)u_k \geq 0$ *P-almost surely for every* $k \in \mathbb{N}$. *Then* $u \in \mathrm{dom}(\delta)$ *and*

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

PROOF.   Note that (2.14) implies (2.10), and hence $u \in \mathrm{dom}(\delta)$. Since $F \in L^2(\Omega)$, it can be represented as

$$F = \sum_{n=0}^{\infty} J_n(f_n)$$

with kernels $f_0 := \mathbb{E}[F]$ and $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$ for $n \in \mathbb{N}$. The isometry in (2.1) yields

$$\mathbb{E}[F\delta(u)] = \mathbb{E}\left[\left(\sum_{n=0}^{\infty} J_n(f_n)\right)\left(\sum_{n=0}^{\infty} J_{n+1}(\widetilde{g_{n+1}}\mathbb{1}_{\Delta_{n+1}})\right)\right]$$

$$(2.16) \qquad = \sum_{n=0}^{\infty}(n+1)!\langle f_{n+1}, \widetilde{g_{n+1}}\mathbb{1}_{\Delta_{n+1}}\rangle_{\ell^2(\mathbb{N})^{\otimes n+1}}$$

$$= \sum_{n=0}^{\infty}(n+1)!\langle f_{n+1}, g_{n+1}\rangle_{\ell^2(\mathbb{N})^{\otimes n+1}}.$$

Note that the last step in (2.16) is valid, since, for every $n \in \mathbb{N}$, f_n is symmetric and vanishes on diagonals.

Since $(D_k F)u_k \geq 0$ P-almost surely for every $k \in \mathbb{N}$ and by the isometry formula for discrete multiple stochastic integrals, we get

$$\mathbb{E}[\langle DF, u\rangle_{\ell^2(\mathbb{N})}] = \sum_{k=1}^{\infty}\mathbb{E}[(D_k F)u_k]$$

$$= \sum_{k=1}^{\infty}\mathbb{E}\left[\left(\sum_{n=0}^{\infty}(n+1)J_n(f_{n+1}(\cdot, k))\right)\left(\sum_{n=0}^{\infty}J_n(g_{n+1}(\cdot, k))\right)\right]$$

$$(2.17) \qquad = \sum_{k=1}^{\infty}\sum_{n=0}^{\infty}(n+1)!\langle f_{n+1}(\cdot, k), g_{n+1}(\cdot, k)\rangle_{\ell^2(\mathbb{N})^{\otimes n}}$$

$$= \sum_{n=0}^{\infty}(n+1)!\sum_{k=1}^{\infty}\langle f_{n+1}(\cdot, k), g_{n+1}(\cdot, k)\rangle_{\ell^2(\mathbb{N})^{\otimes n}}$$

$$= \sum_{n=0}^{\infty}(n+1)!\langle f_{n+1}, g_{n+1}\rangle_{\ell^2(\mathbb{N})^{\otimes n+1}}.$$

Note that the exchange of summation in the penultimate step of (2.17) is valid by Fubini's theorem, since a repeated application of the Cauchy–Schwarz inequality yields that

$$\sum_{n=0}^{\infty}\sum_{k=1}^{\infty}|(n+1)!\langle f_{n+1}(\cdot, k), g_{n+1}(\cdot, k)\rangle_{\ell^2(\mathbb{N})^{\otimes n}}|$$

$$\leq \sum_{n=0}^{\infty}(n+1)!\sum_{k=1}^{\infty}\|f_{n+1}(\cdot, k)\|_{\ell^2(\mathbb{N})^{\otimes n}}\|g_{n+1}(\cdot, k)\|_{\ell^2(\mathbb{N})^{\otimes n}}$$

$$\leq \sum_{n=0}^{\infty} (n+1)! \left(\sum_{k=1}^{\infty} \| f_{n+1}(\cdot, k) \|_{\ell^2(\mathbb{N})^{\otimes n}}^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} \| g_{n+1}(\cdot, k) \|_{\ell^2(\mathbb{N})^{\otimes n}}^2 \right)^{1/2}$$

$$= \sum_{n=0}^{\infty} (n+1)! \| f_{n+1} \|_{\ell^2(\mathbb{N})^{\otimes n+1}} \| g_{n+1} \|_{\ell^2(\mathbb{N})^{\otimes n+1}}$$

$$\leq \left(\sum_{n=0}^{\infty} (n+1)! \| f_{n+1} \|_{\ell^2(\mathbb{N})^{\otimes n+1}}^2 \right)^{1/2} \left(\sum_{n=0}^{\infty} (n+1)! \| g_{n+1} \|_{\ell^2(\mathbb{N})^{\otimes n+1}}^2 \right)^{1/2}$$

$$\leq \left(\mathbb{E}[F^2] \right)^{1/2} \left(\sum_{n=0}^{\infty} (n+1)! \| g_{n+1} \|_{\ell^2(\mathbb{N})^{\otimes n+1}}^2 \right)^{1/2} < \infty.$$

Comparing (2.16) and (2.17) completes the proof. $\square$

Finally, we recall the following Skorohod isometric formula for the discrete divergence operator. Namely, for all $u \in \mathrm{dom}(\delta)$ it holds that

$$(2.18) \qquad \mathbb{E}[\delta(u)^2] = \mathbb{E}[\|u\|_{\ell^2(\mathbb{N})}^2] + \mathbb{E}\left[\sum_{k,\ell=1}^{\infty} (D_k u_\ell)(D_\ell u_k) \right]$$

according to Proposition 9.3 in [30].

3. The discrete Ornstein–Uhlenbeck semigroup. For real $t \geq 0$, define the random sequence $X^t := (X_k^t)_{k \in \mathbb{N}}$ by

$$X_k^t := X_k^* \mathbb{1}_{\{Z_k \leq t\}} + X_k \mathbb{1}_{\{Z_k > t\}},$$

where $(X_k^*)_{k \in \mathbb{N}}$ is an independent copy of the Rademacher sequence $X = (X_k)_{k \in \mathbb{N}}$ and $(Z_k)_{k \in \mathbb{N}}$ is a sequence of independent and exponentially distributed random variables with mean 1, independent of all other random variables.

Our first result is a discrete analogue of Mehler's formula on the Wiener or Poisson chaos for which we refer to [20] and [17], respectively. It expresses the discrete Ornstein–Uhlenbeck semigroup $(P_t)_{t \geq 0}$ defined at (2.9) in terms of a conditional expectation. Note that this has already been shown in [30], Proposition 10.8. Since Mehler's formula is a central device in our approach, we include an elementary and direct proof.

PROPOSITION 3.1 (Mehler's formula). *Let $F \in L^2(\Omega)$. The process $(X^t)_{t \geq 0}$ is the Ornstein–Uhlenbeck process associated with $(P_t)_{t \geq 0}$ by the relation*

$$P_t F = \mathbb{E}[F(X^t)|X] \qquad P\text{-a.s.}$$

for all $t \geq 0$.

PROOF.   We first notice that for each $t \geq 0$, $(X_k^t)_{k \in \mathbb{N}}$ is a sequence of independent Rademacher random variables with the same distribution as the sequence $(X_k)_{k \in \mathbb{N}}$. Thus, if $F = \mathbb{E}[F] + \sum_{n=1}^{\infty} J_n(f_n)$, then $F(X^t)$ has chaotic decomposition

$$(3.1) \qquad F(X^t) = \mathbb{E}[F] + \sum_{n=1}^{\infty} n! \sum_{1 \leq i_1 < \cdots < i_n < \infty} f_n(i_1, \ldots, i_n) Y_{i_1}^t \cdots Y_{i_n}^t,$$

where both decompositions share the same kernels $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$, for $n \in \mathbb{N}$, and where the sequence $(Y_k^t)_{k \in \mathbb{N}}$ with

$$(3.2) \qquad Y_k^t := (X_k^t - p_k + q_k)/(2\sqrt{p_k q_k}) = Y_k^* \mathbb{1}_{\{Z_k \leq t\}} + Y_k \mathbb{1}_{\{Z_k > t\}},$$

for $t \geq 0$, is the normalization of the sequence $(X_k^t)_{k \in \mathbb{N}}$. Here, the random variable Y_k^* is the normalization of X_k^* for every $k \in \mathbb{N}$. Using the independence of the sequences $(X_k)_{k \in \mathbb{N}}$, $(X_k^*)_{k \in \mathbb{N}}$ and $(Z_k)_{k \in \mathbb{N}}$, we deduce from (3.2) that

$$\mathbb{E}[Y_k^t | X_k] = \mathbb{E}[Y_k^*] \cdot P(Z_k \leq t) + Y_k \cdot P(Z_k > t) = Y_k e^{-t}.$$

For a functional F_d only depending on the first d Rademacher random variables we compute by using the chaotic decomposition in (3.1) as well as linearity and independence,

$$\mathbb{E}[F_d(X_1^t, \ldots, X_d^t) | X]$$
$$= \mathbb{E}[F_d(X_1, \ldots, X_d)]$$
$$\qquad + \sum_{n=1}^{d} n! \sum_{1 \leq i_1 < \cdots < i_n \leq d} f_n^{(d)}(i_1, \ldots, i_n) \mathbb{E}[Y_{i_1}^t | X_{i_1}] \cdots \mathbb{E}[Y_{i_n}^t | X_{i_n}]$$
$$(3.3)$$
$$= \mathbb{E}[F_d(X_1, \ldots, X_d)] + \sum_{n=1}^{d} e^{-nt} n! \sum_{1 \leq i_1 < \cdots < i_n \leq d} f_n^{(d)}(i_1, \ldots, i_n) Y_{i_1} \cdots Y_{i_n}$$
$$= \mathbb{E}[F_d(X_1, \ldots, X_d)] + \sum_{n=1}^{d} e^{-nt} J_n(f_n^{(d)})$$
$$= P_t F_d(X_1, \ldots, X_d).$$

The general case follows from (3.3) due to the fact that the set of functionals depending only on finitely many Rademacher variables is dense in $L^2(\Omega)$ and that both sides of (3.3) are continuous functions of F_d. $\square$

As a next step, we derive an integral representation for the expression $-D^m L^{-1} F$, that is, the m-fold iterated discrete gradient applied to $-L^{-1}F$.

PROPOSITION 3.2. *For $m, k_1, \ldots, k_m \in \mathbb{N}$ and centred $F \in \mathrm{dom}(D^m)$ one has that*

$$-D^m_{k_1,\ldots,k_m} L^{-1} F = \int_0^\infty e^{-mt} P_t D^m_{k_1,\ldots,k_m} F \, dt \qquad P\text{-}a.s.$$

PROOF. Since $F \in L^2(\Omega)$ is centred, there are kernels $f_n \in \ell_0^2(\mathbb{N})^{\circ n}$, $n \in \mathbb{N}$, such that $F = \sum_{n=1}^\infty J_n(f_n)$. Fix $d \in \{m, m+1, \ldots\}$ and consider the truncated functional $F_d := \sum_{n=1}^d J_n(f_n)$. Then

$$
\begin{aligned}
& -D^m_{k_1,\ldots,k_m} L^{-1} F_d \\
(3.4) \qquad &= \sum_{n=m}^d \frac{(n-1)!}{(n-m)!} J_{n-m}\big(f_n(\cdot, k_1, \ldots, k_m)\big) \\
&= \int_0^\infty e^{-mt} \sum_{n=m}^d \frac{n!}{(n-m)!} e^{-(n-m)t} J_{n-m}\big(f_n(\cdot, k_1, \ldots, k_m)\big) \, dt,
\end{aligned}
$$

where we used that $\int_0^\infty n e^{-nt} \, dt = 1$. By continuity of $D^m_{k_1,\ldots,k_m}$ and L^{-1} one has that $-D^m_{k_1,\ldots,k_m} L^{-1} F_d$ converges to $-D^m_{k_1,\ldots,k_m} L^{-1} F$ in $L^2(\Omega)$, as $d \to \infty$. To show that the right-hand side of (3.4) converges to

$$\int_0^\infty e^{-mt} P_t D^m_{k_1,\ldots,k_m} F \, dt$$

in $L^2(\Omega)$, as $d \to \infty$, we consider the remainder term

$$R_{m,d} := \int_0^\infty e^{-mt} P_t D^m_{k_1,\ldots,k_m} F \, dt - \big(-D^m_{k_1,\ldots,k_m} L^{-1} F_d\big)$$

and show that $\mathbb{E}[R^2_{m,d}]$ vanishes, as $d \to \infty$. First, use (3.4) to see that

$$R_{m,d} = \int_0^\infty e^{-mt} \sum_{n=d+1}^\infty \frac{n!}{(n-m)!} e^{-(n-m)t} J_{n-m}\big(f_n(\cdot, k_1, \ldots, k_m)\big) \, dt.$$

We then apply Jensen's inequality, Fubini's theorem and the isometry property of discrete multiple stochastic integrals to conclude that

$$
\begin{aligned}
& \mathbb{E}\big[R^2_{m,d}\big] \\
&= \mathbb{E}\left[\left(\int_0^\infty e^{-mt} \sum_{n=d+1}^\infty \frac{n!}{(n-m)!} e^{-(n-m)t} J_{n-m}\big(f_n(\cdot, k_1, \ldots, k_m)\big) \, dt\right)^2\right] \\
&\leq \mathbb{E}\left[\int_0^\infty e^{-(2m-1)t} \left(\sum_{n=d+1}^\infty \frac{n!}{(n-m)!} e^{-(n-m)t} J_{n-m}\big(f_n(\cdot, k_1, \ldots, k_m)\big)\right)^2 dt\right]
\end{aligned}
$$

$$= \int_0^\infty e^{-(2m-1)t} \sum_{n=d+1}^\infty \left(\frac{n!}{(n-m)!} \right)^2 e^{-2(n-m)t} (n-m)!$$

$$\times \| f_n(\cdot, k_1, \ldots, k_m) \|_{\ell^2(\mathbb{N})^{\otimes(n-m)}}^2 \, dt$$

$$\le \sum_{n=d+1}^\infty \left(\frac{n!}{(n-m)!} \right)^2 (n-m)! \| f_n(\cdot, k_1, \ldots, k_m) \|_{\ell^2(\mathbb{N})^{\otimes(n-m)}}^2,$$

where we used that $\int_0^\infty e^{-(2n-1)t} \, dt = (2n-1)^{-1} \le 1$. Since $F \in \mathrm{dom}(D^m)$, the latter expression is finite and converges to zero, as $d \to \infty$. This completes the proof. $\square$

Our next result combines the previous two propositions and is one of the key tools in the proof of our general Berry–Esseen bound in Section 4. Similar relations also hold on the Wiener and the Poisson space for which we refer to [23] and [17], respectively. Although from a formal point of view the statement looks similar to these results, we emphasize that the proof as well as the meaning and the interpretation of the involved Malliavin operators in our discrete framework are different.

PROPOSITION 3.3. *For $m, k_1, \ldots, k_m \in \mathbb{N}$, $\alpha \ge 1$ and centred $F \in \mathrm{dom}(D^m)$ one has that*

$$\mathbb{E}\big[\big| D_{k_1,\ldots,k_m}^m L^{-1} F \big|^\alpha \big] \le \mathbb{E}\big[\big| D_{k_1,\ldots,k_m}^m F \big|^\alpha \big].$$

PROOF. According to Proposition 3.2, we have that

$$\mathbb{E}\big[\big| D_{k_1,\ldots,k_m}^m L^{-1} F \big|^\alpha \big] = \mathbb{E}\left[\left| \int_0^\infty e^{-mt} P_t D_{k_1,\ldots,k_m}^m F \, dt \right|^\alpha \right].$$

Then, using Proposition 3.1 together with Jensen's inequality, we deduce that

$$\mathbb{E}\left[\left| \int_0^\infty e^{-mt} P_t D_{k_1,\ldots,k_m}^m F \, dt \right|^\alpha \right] = \mathbb{E}\left[\left| \int_0^\infty e^{-mt} \mathbb{E}\big[D_{k_1,\ldots,k_m}^m F(X^t) | X \big] dt \right|^\alpha \right]$$

$$\le \mathbb{E}\left[\int_0^\infty e^{-mt} \mathbb{E}\big[| D_{k_1,\ldots,k_m}^m F(X^t) |^\alpha | X \big] dt \right]$$

$$= \int_0^\infty e^{-mt} \mathbb{E}\big[| D_{k_1,\ldots,k_m}^m F |^\alpha \big] dt$$

$$\le \mathbb{E}\big[| D_{k_1,\ldots,k_m}^m F |^\alpha \big]$$

and complete the proof. $\square$

As a first application of Proposition 3.3 we now deduce the following discrete Poincaré-type inequality. This result can already be found in [30], Chapter 8, where it is proved by means of the Clark formula. We present an alternative proof without resorting to this formula.

PROPOSITION 3.4.   *Suppose that $F \in \mathrm{dom}(D)$. Then*

(3.5)
$$\mathrm{Var}[F] \leq \mathbb{E}\big[\|DF\|_{\ell^2(\mathbb{N})}^2\big].$$

PROOF.   Choosing f in (2.13) as the identity map on $\mathbb{R}$ yields

$$\mathrm{Var}[F] = \mathbb{E}\big[(F - \mathbb{E}[F])^2\big] = \mathbb{E}\big[\langle D(F - \mathbb{E}[F]), -DL^{-1}(F - \mathbb{E}[F])\rangle_{\ell^2(\mathbb{N})}\big]$$

$$= \mathbb{E}\Bigg[\sum_{k=1}^{\infty}(D_k(F - \mathbb{E}[F]))(-D_k L^{-1}(F - \mathbb{E}[F]))\Bigg]$$

$$\leq \mathbb{E}\Bigg[\sum_{k=1}^{\infty}|D_k(F - \mathbb{E}[F])||D_k L^{-1}(F - \mathbb{E}[F])|\Bigg].$$

Exchanging expectation and summation, and using the Cauchy–Schwarz inequality, we see that the latter expression is further bounded by

$$\sum_{k=1}^{\infty}(\mathbb{E}[(D_k(F - \mathbb{E}[F]))^2])^{1/2}(\mathbb{E}[(D_k L^{-1}(F - \mathbb{E}[F]))^2])^{1/2}.$$

The proof is now concluded by applying Proposition 3.3 with $m = 1$ and $\alpha = 2$ and using the fact that $D_k(F - \mathbb{E}[F]) = D_k F$. $\square$

REMARK 3.1.   Proposition 3.4 remains valid for $F \in L^1(\Omega) \setminus \mathrm{dom}(D)$, since in this case the right-hand side of (3.5) is infinite.

4. A general Berry–Esseen bound.   The main result of this section is a Berry–Esseen bound for square-integrable Rademacher functionals F. By such a result, we mean an upper bound for the Kolmogorov distance between F and a standard Gaussian random variable, where we recall that the Kolmogorov distance between two random variables X and Y is defined as

$$d_K(X, Y) := \sup_{x \in \mathbb{R}} |P(X \leq x) - P(Y \leq x)|.$$

A first result in this direction has been shown by the authors in [14] in the special symmetric case that the sequence $p = (p_k)_{k \in \mathbb{N}}$ is constant and equal to $1/2$. In the present situation, we need the following generalization to arbitrary sequences p. Since the proof follows straightforwardly along the lines of that of Theorem 3.1 in [14], we omit the arguments.

PROPOSITION 4.1.   *Let $F \in \mathrm{dom}(D)$ with $\mathbb{E}[F] = 0$ and let $N \sim \mathcal{N}(0, 1)$ be a standard Gaussian random variable. Then*

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

$$+ \frac{\sqrt{2\pi}}{8}\mathbb{E}\big[\langle (pq)^{-1/2}(DF)^2, |DL^{-1}F|\rangle_{\ell^2(\mathbb{N})}\big]$$

$$+ \frac{1}{2} \mathbb{E}\big[\langle (pq)^{-1/2}(DF)^2, |F \cdot DL^{-1}F|\rangle_{\ell^2(\mathbb{N})}\big]$$

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

One disadvantage of the bound in Proposition 4.1 is that it involves the inverse of the discrete Ornstein–Uhlenbeck operator. In applications this means that the chaotic decomposition of the Rademacher functional F has to be computed explicitly in order to evaluate the expression $-DL^{-1}F$. A further analysis of the bound then requires a multiplication formula for discrete multiple stochastic integrals, which expresses a product of two discrete multiple stochastic integrals as linear combination of discrete multiple stochastic integrals. We transfer the bound of Proposition 4.1 into a form, which can be evaluated without using a multiplication formula. Our next result is a combination of Propositions 4.1 and 3.3, and provides an upper bound for $d_K(F, N)$ in terms of the first- and second-order discrete gradient only. A result of this structure is what is called a 'second-order Poincaré inequality' in the literature; see [3, 17, 23]. The discrete Poincaré-type inequality in Proposition 3.4 says that a Rademacher functional F is concentrated around $\mathbb{E}[F]$ in terms of the variance if the contribution of the first-order discrete gradient is small. Our discrete second-order Poincaré inequality additionally implies that if the contribution of the second-order discrete gradient is also small, then F is close to a standard Gaussian random variable.

THEOREM 4.1. *Let* $F \in \mathrm{dom}(D)$ *with* $\mathbb{E}[F] = 0$ *and* $\mathbb{E}[F^2] = 1$, *and let* $N \sim \mathcal{N}(0, 1)$ *be a standard Gaussian random variable. Further, fix* $r, s, t \in (1, \infty)$ *such that* $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = 1$. *Then*

$d_K(F, N)$

$$\leq \left(\frac{15}{4} \sum_{j,k,\ell=1}^{\infty} \big(\mathbb{E}[(D_j F)^2 (D_k F)^2]\big)^{1/2} \big(\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2]\big)^{1/2} \right)^{1/2}$$

$$+ \left(\frac{3}{4} \sum_{j,k,\ell=1}^{\infty} \frac{1}{p_\ell q_\ell} \mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2] \right)^{1/2}$$

$$+ \frac{\sqrt{2\pi}}{8} \sum_{k=1}^{\infty} \frac{1}{\sqrt{p_k q_k}} \mathbb{E}[|D_k F|^3]$$

$$+ \frac{1}{2} \big(\mathbb{E}[|F|^r]\big)^{1/r} \sum_{k=1}^{\infty} \frac{1}{\sqrt{p_k q_k}} \big(\mathbb{E}[|D_k F|^{2s}]\big)^{1/s} \big(\mathbb{E}[|D_k F|^t]\big)^{1/t}$$

$$+ \left(\sum_{k=1}^{\infty} \frac{1}{p_k q_k} \mathbb{E}[(D_k F)^4] \right)^{1/2}$$

$$+ \left(6 \sum_{k,\ell=1}^{\infty} \frac{1}{p_k q_k} \left(\mathbb{E}[(D_k F)^4] \right)^{1/2} \left(\mathbb{E}[(D_\ell D_k F)^4] \right)^{1/2} \right)^{1/2}$$

$$+ \left(3 \sum_{k,\ell=1}^{\infty} \frac{1}{p_k q_k p_\ell q_\ell} \mathbb{E}[(D_\ell D_k F)^4] \right)^{1/2}.$$

Let us comment on the second-order Poincaré inequality in Theorem 4.1. Its form differs from that in the Wiener or Poisson case treated in [17, 23]. The main difference is the fourth term, which involves the parameters r, s and t, and hence higher moments of F and $D_k F$. In many applications, one can choose $r = 2$ and $s = t = 4$, but there are situations in which the additional flexibility to choose r, s and t differently turns out to be crucial. We shall meet such an example in the proof of Theorem 1.1 on the triangle counting statistic in the Erdős–Rényi random graph.

PROOF OF THEOREM 4.1. Our proof follows the general scheme to establish a second-order Poincaré inequality, which is used in the literature [3, 17, 23]. Namely, we build on Proposition 4.1 by further estimating each summand of the bound there. We start with the first summand, to which we apply the Cauchy–Schwarz inequality:

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

Taking f as the identity on $\mathbb{R}$ in (2.13) shows that $\mathbb{E}[\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}] = \mathrm{Var}[F] = 1$. Thus, $\mathbb{E}[(1 - \langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})})^2] = \mathrm{Var}[\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}]$ and an application of Proposition 3.4 (see also Remark 3.1) yields

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

$$(4.1) \qquad = \mathbb{E}\left[\sum_{\ell=1}^{\infty} \left(D_\ell \left(\sum_{k=1}^{\infty} (D_k F)(-D_k L^{-1}F) \right) \right)^2 \right]$$

$$= \mathbb{E}\left[\sum_{\ell=1}^{\infty} \left(\sum_{k=1}^{\infty} D_\ell \left((D_k F)(-D_k L^{-1}F) \right) \right)^2 \right],$$

where the exchange of D_ℓ with the summation in the last step can be justified as follows. Since $\mathbb{E}[\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}] = 1$, $\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}$ is P-a.s. finite. Thus, $\langle DF_\ell^{\pm}, -DL^{-1}F_\ell^{\pm} \rangle_{\ell^2(\mathbb{N})}$ as well as the pathwise representation of $D_\ell(\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})})$ as at (2.3) are P-a.s. finite for $\ell \in \mathbb{N}$. As a consequence, we see that

$$D_\ell(\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})})$$

$$= \sqrt{p_\ell q_\ell}(\langle DF_\ell^+, -DL^{-1}F_\ell^+ \rangle_{\ell^2(\mathbb{N})} - \langle DF_\ell^-, -DL^{-1}F_\ell^- \rangle_{\ell^2(\mathbb{N})})$$

$$= \sqrt{p_\ell q_\ell} \left(\sum_{k=1}^{\infty} (D_k F_\ell^+)(-D_k L^{-1} F_\ell^+) - \sum_{k=1}^{\infty} (D_k F_\ell^-)(-D_k L^{-1} F_\ell^-) \right)$$

$$= \sqrt{p_\ell q_\ell} \sum_{k=1}^{\infty} \left((D_k F_\ell^+)(-D_k L^{-1} F_\ell^+) - (D_k F_\ell^-)(-D_k L^{-1} F_\ell^-) \right)$$

$$= \sum_{k=1}^{\infty} D_\ell \left((D_k F)(-D_k L^{-1} F) \right) \qquad P\text{-a.s.}$$

for $\ell \in \mathbb{N}$. Now, we further estimate the quantity $D_\ell((D_k F)(-D_k L^{-1} F))$ in (4.1) using the product formula (2.4). This yields

$$\left| D_\ell \left((D_k F)(-D_k L^{-1} F) \right) \right|$$

$$= \left| (D_\ell D_k F)(-D_k L^{-1} F) + (D_k F)(-D_\ell D_k L^{-1} F) \right.$$

$$\left. - \frac{X_\ell}{\sqrt{p_\ell q_\ell}} (D_\ell D_k F)(-D_\ell D_k L^{-1} F) \right|$$

$$\leq \left| (D_\ell D_k F)(-D_k L^{-1} F) \right| + \left| (D_k F)(-D_\ell D_k L^{-1} F) \right|$$

$$+ \frac{1}{\sqrt{p_\ell q_\ell}} \left| (D_\ell D_k F)(-D_\ell D_k L^{-1} F) \right|.$$

Using this together with the Cauchy–Schwarz inequality, it follows from (4.1) that

(4.2) $$\mathbb{E}\left[\left(1 - \langle DF, -DL^{-1} F \rangle_{\ell^2(\mathbb{N})} \right)^2 \right] \leq 3(T_1 + T_2 + T_3),$$

where T_1, T_2 and T_3 are given by

$$T_1 := \mathbb{E}\left[\sum_{\ell=1}^{\infty} \left(\sum_{k=1}^{\infty} \left| (D_\ell D_k F)(-D_k L^{-1} F) \right| \right)^2 \right],$$

$$T_2 := \mathbb{E}\left[\sum_{\ell=1}^{\infty} \left(\sum_{k=1}^{\infty} \left| (D_k F)(-D_\ell D_k L^{-1} F) \right| \right)^2 \right],$$

$$T_3 := \mathbb{E}\left[\sum_{\ell=1}^{\infty} \frac{1}{p_\ell q_\ell} \left(\sum_{k=1}^{\infty} \left| (D_\ell D_k F)(-D_\ell D_k L^{-1} F) \right| \right)^2 \right].$$

Each of these terms is now further estimated from above. Considering T_1, an application of Propositions 3.2 and Proposition 3.1 as well as Jensen's inequality yields for $\ell \in \mathbb{N}$ that

$$\left(\sum_{k=1}^{\infty} |D_\ell D_k F| |D_k L^{-1} F| \right)^2$$

$$= \left(\sum_{k=1}^{\infty} |D_\ell D_k F| \left| \int_0^{\infty} e^{-t} P_t D_k F \, dt \right| \right)^2$$

$$= \left(\sum_{k=1}^{\infty} |D_\ell D_k F| \left| \int_0^\infty e^{-t} \mathbb{E}[D_k F(X^t)|X] \, dt \right| \right)^2$$

$$\leq \left(\sum_{k=1}^{\infty} |D_\ell D_k F| \int_0^\infty e^{-t} \mathbb{E}[|D_k F(X^t)||X] \, dt \right)^2.$$

By virtue of the monotone convergence theorem, we get for $\ell \in \mathbb{N}$ that

$$\left(\sum_{k=1}^{\infty} |D_\ell D_k F| \int_0^\infty e^{-t} \mathbb{E}[|D_k F(X^t)||X] \, dt \right)^2$$

$$= \left(\int_0^\infty e^{-t} \sum_{k=1}^{\infty} |D_\ell D_k F| \mathbb{E}[|D_k F(X^t)||X] \, dt \right)^2$$

$$= \left(\int_0^\infty e^{-t} \mathbb{E}\left[\sum_{k=1}^{\infty} |D_\ell D_k F||D_k F(X^t)| \Big| X \right] dt \right)^2.$$

Using Jensen's inequality again as well as the Cauchy–Schwarz inequality, we now conclude for $\ell \in \mathbb{N}$ that

$$\left(\int_0^\infty e^{-t} \mathbb{E}\left[\sum_{k=1}^{\infty} |D_\ell D_k F||D_k F(X^t)| \Big| X \right] dt \right)^2$$

$$\leq \int_0^\infty e^{-t} \mathbb{E}\left[\left(\sum_{k=1}^{\infty} |D_\ell D_k F||D_k F(X^t)| \right)^2 \Big| X \right] dt$$

$$= \int_0^\infty e^{-t} \mathbb{E}\left[\sum_{j,k=1}^{\infty} |D_\ell D_j F||D_j F(X^t)||D_\ell D_k F||D_k F(X^t)| \Big| X \right] dt$$

$$= \sum_{j,k=1}^{\infty} |D_\ell D_j F||D_\ell D_k F| \int_0^\infty e^{-t} \mathbb{E}[|D_j F(X^t)||D_k F(X^t)||X] \, dt$$

$$\leq \sum_{j,k=1}^{\infty} |D_\ell D_j F||D_\ell D_k F| \int_0^\infty e^{-t} (\mathbb{E}[(D_j F(X^t))^2 (D_k F(X^t))^2|X])^{1/2} \, dt$$

$$\leq \sum_{j,k=1}^{\infty} |D_\ell D_j F||D_\ell D_k F|$$

$$\times \left(\int_0^\infty e^{-t} \mathbb{E}[(D_j F(X^t))^2 (D_k F(X^t))^2|X] \, dt \right)^{1/2}.$$

Thus, another application of the Cauchy–Schwarz inequality leads to the bound

$$
\begin{aligned}
T_1 \leq \mathbb{E}\Bigg[&\sum_{j,k,\ell=1}^{\infty} |D_\ell D_j F||D_\ell D_k F| \\
&\times \left(\int_0^\infty e^{-t} \mathbb{E}[(D_j F(X^t))^2 (D_k F(X^t))^2 | X] \, dt \right)^{1/2} \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
\leq &\sum_{j,k,\ell=1}^{\infty} (\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2])^{1/2} \\
&\times \left(\mathbb{E}\left[\int_0^\infty e^{-t} \mathbb{E}[(D_j F(X^t))^2 (D_k F(X^t))^2 | X] \, dt \right] \right)^{1/2}
\end{aligned}
\tag{4.3}
$$

$$
= \sum_{j,k,\ell=1}^{\infty} (\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2])^{1/2} \left(\int_0^\infty e^{-t} \mathbb{E}[(D_j F)^2 (D_k F)^2] \, dt \right)^{1/2}
$$

$$
= \sum_{j,k,\ell=1}^{\infty} (\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2])^{1/2} (\mathbb{E}[(D_j F)^2 (D_k F)^2])^{1/2}.
$$

Using similar arguments and Proposition 3.2 for $m = 2$, one shows that

$$
T_2 \leq \frac{1}{4} \sum_{j,k,\ell=1}^{\infty} (\mathbb{E}[(D_j F)^2 (D_k F)^2])^{1/2} (\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2])^{1/2}
\tag{4.4}
$$

and

$$
T_3 \leq \frac{1}{4} \sum_{j,k,\ell=1}^{\infty} \frac{1}{p_\ell q_\ell} \mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2].
\tag{4.5}
$$

Thus, combining (4.3)–(4.5) with (4.2), we get

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

$$
\begin{aligned}
\leq &\left(\frac{15}{4} \sum_{j,k,\ell=1}^{\infty} (\mathbb{E}[(D_j F)^2 (D_k F)^2])^{1/2} (\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2])^{1/2} \right)^{1/2} \\
&+ \left(\frac{3}{4} \sum_{j,k,\ell=1}^{\infty} \frac{1}{p_\ell q_\ell} \mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2] \right)^{1/2}
\end{aligned}
\tag{4.6}
$$

as an estimate for the first summand of the bound in Proposition 4.1.

For the second summand, we obtain

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

$$
= \sum_{k=1}^{\infty} (p_k q_k)^{-1/2} \mathbb{E}[(D_k F)^2 | D_k L^{-1} F|]
\tag{4.7}
$$

$$\leq \sum_{k=1}^{\infty} (p_k q_k)^{-1/2} \big(\mathbb{E}[|D_k F|^3]\big)^{2/3} \big(\mathbb{E}[|D_k L^{-1} F|^3]\big)^{1/3}$$

$$\leq \sum_{k=1}^{\infty} (p_k q_k)^{-1/2} \mathbb{E}[|D_k F|^3]$$

by means of Hölder's inequality with Hölder conjugates 3 and 3/2, and Proposition 3.3. Applying a generalization of Hölder's inequality with Hölder conjugates $r, s, t \in (1, \infty)$ with $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = 1$ as well as Proposition 3.3 to the third summand of the bound in Proposition 4.1 immediately yields

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

(4.8)
$$= \sum_{k=1}^{\infty} (p_k q_k)^{-1/2} \mathbb{E}\big[|F|(D_k F)^2 |D_k L^{-1} F|\big]$$

$$\leq \big(\mathbb{E}[|F|^r]\big)^{1/r} \sum_{k=1}^{\infty} (p_k q_k)^{-1/2} \big(\mathbb{E}[|D_k F|^{2s}]\big)^{1/s} \big(\mathbb{E}[|D_k L^{-1} F|^t]\big)^{1/t}$$

$$\leq \big(\mathbb{E}[|F|^r]\big)^{1/r} \sum_{k=1}^{\infty} (p_k q_k)^{-1/2} \big(\mathbb{E}[|D_k F|^{2s}]\big)^{1/s} \big(\mathbb{E}[|D_k F|^t]\big)^{1/t}.$$

We now apply the integration-by-parts-formula (2.15) in order to bound the last term in Proposition 4.1. To this end, we note that $D_k \mathbb{1}_{\{F>x\}} D_k F |D_k L^{-1} F| \geq 0$ for every $k \in \mathbb{N}$ and we need to verify the summability condition in (2.14) in Proposition 2.2. The latter will be verified subsequent to the following calculation. Using the integration-by-parts-formula, we obtain that

$$\mathbb{E}\big[\langle (pq)^{-1/2}(DF)D\mathbb{1}_{\{F>x\}}, |DL^{-1}F|\rangle_{\ell^2(\mathbb{N})}\big]$$

$$= \mathbb{E}\big[\langle D\mathbb{1}_{\{F>x\}}, (pq)^{-1/2}(DF)|DL^{-1}F|\rangle_{\ell^2(\mathbb{N})}\big]$$

(4.9)
$$= \mathbb{E}\big[\mathbb{1}_{\{F>x\}}\delta\big((pq)^{-1/2}(DF)|DL^{-1}F|\big)\big]$$

$$\leq \mathbb{E}\big[|\delta\big((pq)^{-1/2}(DF)|DL^{-1}F|\big)|\big]$$

$$\leq \big(\mathbb{E}\big[\big(\delta\big((pq)^{-1/2}(DF)|DL^{-1}F|\big)\big)^2\big]\big)^{1/2}.$$

From the isometric formula (2.18) for the divergence operator it follows that

$$\mathbb{E}\big[\big(\delta\big((pq)^{-1/2}(DF)|DL^{-1}F|\big)\big)^2\big]$$

$$= \mathbb{E}\big[\|(pq)^{-1/2}(DF)(DL^{-1}F)\|^2_{\ell^2(\mathbb{N})}\big]$$

(4.10)
$$+ \mathbb{E}\Bigg[\sum_{k,\ell=1}^{\infty} (p_\ell q_\ell)^{-1/2} D_k\big((D_\ell F)|D_\ell L^{-1}F|\big)$$

$$\times (p_k q_k)^{-1/2} D_\ell \big((D_k F) \big| D_k L^{-1} F \big| \big) \Big]$$

$$\leq \mathbb{E}\big[\| (pq)^{-1/2} (DF)(DL^{-1}F) \|_{\ell^2(\mathbb{N})}^2 \big]$$

$$+ \mathbb{E}\left[\sum_{k,\ell=1}^{\infty} (p_k q_k)^{-1} \big(D_\ell \big((D_k F)(D_k L^{-1} F) \big) \big)^2 \right]$$

$$=: T_4 + T_5.$$

The term T_4 can easily be estimated by means of the Cauchy–Schwarz inequality and Proposition 3.3, which yields that

$$T_4 = \sum_{k=1}^{\infty} (p_k q_k)^{-1} \mathbb{E}\big[(D_k F)^2 (D_k L^{-1} F)^2 \big]$$

(4.11)
$$\leq \sum_{k=1}^{\infty} (p_k q_k)^{-1} \big(\mathbb{E}\big[(D_k F)^4 \big] \big)^{1/2} \big(\mathbb{E}\big[(D_k L^{-1} F)^4 \big] \big)^{1/2}$$

$$\leq \sum_{k=1}^{\infty} (p_k q_k)^{-1} \mathbb{E}\big[(D_k F)^4 \big].$$

To handle T_5, we first compute $\mathbb{E}[(D_\ell((D_k F)(D_k L^{-1} F)))^2]$ by using the product formula (2.4), the Cauchy–Schwarz inequality as well as Proposition 3.3. This leads to

$$\mathbb{E}\big[\big(D_\ell \big((D_k F)(D_k L^{-1} F) \big) \big)^2 \big]$$

$$= \mathbb{E}\big[\big((D_\ell D_k F)(D_k L^{-1} F) + (D_k F)(D_\ell D_k L^{-1} F)$$

$$- (p_\ell q_\ell)^{-1/2} X_\ell (D_\ell D_k F)(D_\ell D_k L^{-1} F) \big)^2 \big]$$

$$\leq 3\mathbb{E}\big[(D_\ell D_k F)^2 (D_k L^{-1} F)^2 \big] + 3\mathbb{E}\big[(D_k F)^2 (D_\ell D_k L^{-1} F)^2 \big]$$

(4.12)
$$+ 3(p_\ell q_\ell)^{-1} \mathbb{E}\big[(D_\ell D_k F)^2 (D_\ell D_k L^{-1} F)^2 \big]$$

$$\leq 3\big(\mathbb{E}\big[(D_\ell D_k F)^4 \big] \big)^{1/2} \big(\mathbb{E}\big[(D_k L^{-1} F)^4 \big] \big)^{1/2}$$

$$+ 3\big(\mathbb{E}\big[(D_k F)^4 \big] \big)^{1/2} \big(\mathbb{E}\big[(D_\ell D_k L^{-1} F)^4 \big] \big)^{1/2}$$

$$+ 3(p_\ell q_\ell)^{-1} \big(\mathbb{E}\big[(D_\ell D_k F)^4 \big] \big)^{1/2} \big(\mathbb{E}\big[(D_\ell D_k L^{-1} F)^4 \big] \big)^{1/2}$$

$$\leq 6\big(\mathbb{E}\big[(D_k F)^4 \big] \big)^{1/2} \big(\mathbb{E}\big[(D_\ell D_k F)^4 \big] \big)^{1/2} + 3(p_\ell q_\ell)^{-1} \mathbb{E}\big[(D_\ell D_k F)^4 \big].$$

We now justify the validity of the summability condition (2.14). Assume that

(4.13)
$$\mathbb{E}\left[\sum_{k,\ell=1}^{\infty} (D_k u_\ell)^2 \right] < \infty,$$

where $u_\ell := (p_\ell q_\ell)^{-1/2} D_\ell F |D_\ell L^{-1} F| = \sum_{n=1}^\infty J_n(g_{n+1}(\cdot, \ell))$. Then we obtain that

$$\mathbb{E}\left[\sum_{k,\ell=1}^\infty (D_k u_\ell)^2\right] = \sum_{k,\ell=1}^\infty \mathbb{E}[(D_k u_\ell)^2]$$

$$= \sum_{k,\ell=1}^\infty \sum_{n=1}^\infty n n! \|g_{n+1}(\cdot, k, \ell)\|_{\ell^2(\mathbb{N})^{\otimes n-1}}^2$$

$$= \sum_{n=1}^\infty n n! \|g_{n+1}\|_{\ell^2(\mathbb{N})^{\otimes n+1}}^2,$$

which implies that

$$\sum_{n=1}^\infty (n+1) n! \|g_{n+1}\|_{\ell^2(\mathbb{N})^{\otimes n+1}}^2 \le 2\mathbb{E}\left[\sum_{k,\ell=1}^\infty (D_k u_\ell)^2\right] < \infty.$$

Thus, the summability condition (2.14) is verified, once condition (4.13) is satisfied. Since $T_5 = \mathbb{E}[\sum_{k,\ell=1}^\infty (D_k u_\ell)^2]$, condition (4.13) is verified, once our error bound is finite. Otherwise, the error bound holds trivially. Combining (4.9)–(4.12) yields

$$\sup_{x \in \mathbb{R}} \mathbb{E}\big[\langle (pq)^{-1/2}(DF) D\mathbb{1}_{\{F>x\}}, |DL^{-1}F|\rangle_{\ell^2(\mathbb{N})}\big]$$

(4.14)

$$\le \left(\sum_{k=1}^\infty (p_k q_k)^{-1} \mathbb{E}[(D_k F)^4]\right)^{1/2}$$

$$+ \left(6 \sum_{k,\ell=1}^\infty (p_k q_k)^{-1} \big(\mathbb{E}[(D_k F)^4]\big)^{1/2} \big(\mathbb{E}[(D_\ell D_k F)^4]\big)^{1/2}\right)^{1/2}$$

$$+ \left(3 \sum_{k,\ell=1}^\infty \frac{1}{p_k q_k p_\ell q_\ell} \mathbb{E}[(D_\ell D_k F)^4]\right)^{1/2}.$$

This completes the proof. $\quad\square$

5. Application to the Erdős–Rényi random graph and proof of Theorems 1.1–1.3.

In this section, we apply Theorem 4.1 to counting statistics associated with the *Erdős–Rényi* random graph and establish thereby Theorems 1.1–1.3. First, we formally introduce the model and fix some notation. For $n \in \mathbb{N}$ and a real number $p \in (0, 1)$, let $\mathcal{G}$ be the set of all simple and undirected graphs with vertex set $[n] := \{1, \ldots, n\}$. We consider the probability space $(\mathcal{G}, \mathcal{P}(\mathcal{G}), \mathbb{P})$, where $\mathcal{P}(\mathcal{G})$ is the power set of $\mathcal{G}$ and $\mathbb{P}$ is the probability measure given by

$$\mathbb{P}(G) = p^{e(G)} (1-p)^{\binom{n}{2}-e(G)},$$

where for $G \in \mathcal{G}$, $e(G)$ denotes the number of edges of G. The identity map on $\mathcal{G}$ is called the Erdős–Rényi random graph and is usually abbreviated by $G(n, p)$. We refer to the book [10] for a detailed account of the theory of random graphs.

We are interested in the number T of triangles in $G(n, p)$, that is the number of sub-graphs in $G(n, p)$, which are isomorphic to the complete graph on 3 vertices. To analyse the asymptotic behaviour of this random variable, we typically allow p to depend on n. Following the literature and to simplify the notation, we will often suppress the dependence on n of several (random) variables. In order to apply Theorem 4.1 to the normalized triangle counting statistic $F := (T - \mathbb{E}[T])/\sqrt{\mathrm{Var}[T]}$, we first have to embed the model into the framework of Section 2 and Section 4. If one labels the $\binom{n}{2}$ edges of the complete graph on n vertices in a fixed but arbitrary way, $G(n, p)$ can be regarded as an outcome of $\binom{n}{2}$ independent Bernoulli trials, with success probability equal to p. Here, success in the kth Bernoulli trial means that the kth edge is visible in $G(n, p)$. Hence, $G(n, p)$ can be identified with the vector $(X_1, \ldots, X_{\binom{n}{2}})$ of independent Rademacher random variables with parameter $p_k \equiv p$, where $X_k = +1$ indicates that the edge with number k is visible in $G(n, p)$. From now on, we fix an arbitrary enumeration of the edges in the complete graph on the vertex set $[n]$, write $I := \{1, \ldots, \binom{n}{2}\}$ for the set of labels and denote by e_k, $k \in I$, the kth edge of the graph.

Recall from Lemma 3.5 in [10] that

$$(5.1) \qquad \mathrm{Var}[T] \asymp \begin{cases} \theta^5 n^{4-5\alpha}, & \text{if } 0 \le \alpha \le \frac{1}{2}, \\ \theta^3 n^{3(1-\alpha)}, & \text{if } \frac{1}{2} < \alpha < 1, \end{cases}$$

where we recall that the success probability is given by $p = \theta n^{-\alpha}$ with $\alpha \in [0, 1)$ and $\theta \in (0, n^\alpha)$ such that $\theta \asymp 1$.

PROOF OF THEOREM 1.1.   First, we notice that the assumptions of Theorem 4.1 are satisfied since F is normalized and only depends on finitely many Rademacher variables.

To evaluate the bound in Theorem 4.1, we have to control the random variables $D_k F$ and $D_k D_j F$ for $k, j \in \{1, \ldots, \binom{n}{2}\}$. We have

$$D_k F = \sqrt{pq}(F_k^+ - F_k^-) = \frac{\sqrt{pq}}{\sqrt{\mathrm{Var}[T]}}(T_k^+ - T_k^-)$$

and hence

$$\frac{\sqrt{\mathrm{Var}[T]}}{\sqrt{pq}} D_k F = T_k^+ - T_k^-.$$

Now, we notice that T_k^+ equals the number of triangles in the random graph $G(n, p)$ with e_k visible, while T_k^- equals the number of triangles in the random graph $G(n, p)$ when e_k is not visible. Thus, $T_k^+ - T_k^-$ is the number of triangles

that have edge e_k in common, which shows that the random variable $T_k^+ - T_k^-$ has a binomial distribution $\mathrm{Bin}(n - 2, p^2)$ with parameters $n - 2$ and p^2. This is a consequence of the fact that there are $n - 2$ possible triangles being attached to the kth edge and each of these triangles is a sub-graph of $G(n, p)$ with probability p^2, independently of all other triangles. Hence,

$$(5.2) \qquad \frac{\sqrt{\mathrm{Var}[T]}}{\sqrt{pq}} D_k F \sim \mathrm{Bin}(n - 2, p^2).$$

Next, we consider the second-order discrete gradient and obtain that

$$D_k D_j F = \frac{\sqrt{pq}}{\sqrt{\mathrm{Var}[T]}} D_k (T_j^+ - T_j^-)$$

$$= \frac{pq}{\sqrt{\mathrm{Var}[T]}} ((T_j^+)_k^+ - (T_j^+)_k^- - ((T_j^-)_k^+ - (T_j^-)_k^-)),$$

whence

$$\frac{\sqrt{\mathrm{Var}[T]}}{pq} D_k D_j F = (T_j^+)_k^+ - (T_j^+)_k^- - ((T_j^-)_k^+ - (T_j^-)_k^-).$$

The random variable $(T_j^+)_k^+ - (T_j^+)_k^-$ counts the number of triangles in $G(n, p)$ adjacent to the kth edge e_k, conditioned on the event that the jth edge e_j is visible in $G(n, p)$. Similarly, $(T_j^-)_k^+ - (T_j^-)_k^-$ counts the number of triangles adjacent to e_k when e_j does not belong to $G(n, p)$. Therefore, $(T_j^+)_k^+ - (T_j^+)_k^- - ((T_j^-)_k^+ - (T_j^-)_k^-)$ is the number of triangles with common edges e_k and e_j. The number of vertices shared by both edges e_k and e_j is $|e_k \cap e_j|$. Then, if $|e_k \cap e_j| \in \{0, 2\}$, $(T_j^+)_k^+ - (T_j^+)_k^- - ((T_j^-)_k^+ - (T_j^-)_k^-) = 0$ and if $|e_k \cap e_j| = 1$, we have $e_k = \{r, s\}$ and $e_j = \{r, t\}$ for some $r, s, t \in [n]$, $s \neq t$. In this case, $(T_j^+)_k^+ - (T_j^+)_k^- - ((T_j^-)_k^+ - (T_j^-)_k^-)$ is either 1 or 0, depending on whether the edge $\{s, t\}$ belongs to $G(n, p)$ or not. Thus,

$$(5.3) \qquad \frac{\sqrt{\mathrm{Var}[T]}}{pq} D_k D_j F \begin{cases} \sim \mathrm{Ber}(p), & \text{if } |e_k \cap e_j| = 1, \\ = 0, & \text{if } |e_k \cap e_j| \in \{0, 2\}, \end{cases}$$

where $\mathrm{Ber}(p) = \mathrm{Bin}(1, p)$ indicates a Bernoulli distribution with parameter p. Note that the random variables $D_\ell D_k F$ and $D_\ell D_j F$ are independent whenever $k \neq j$. Indeed, fix ℓ and let $k \neq j$, and suppose that $|e_k \cap e_\ell| \in \{0, 2\}$ or $|e_j \cap e_\ell| \in \{0, 2\}$. Then $D_\ell D_k F$ and $D_\ell D_j F$ are independent, since at least one of them is equal to zero. Now, consider the case that $|e_k \cap e_\ell| = 1$ and $|e_j \cap e_\ell| = 1$. In this situation, the three edges e_k, e_j, e_ℓ can have the following form. Either

$$(5.4) \quad e_k = \{s, t\}, \qquad e_j = \{u, v\}, \qquad e_\ell = \{t, u\}, \qquad s \neq u, \qquad v \neq t,$$

or

$$(5.5) \quad e_k = \{s, t\}, \qquad e_j = \{u, t\}, \qquad e_\ell = \{v, t\}, \qquad s \neq v, \qquad u \neq v.$$

In the situation of (5.4), we have $\{s, u\} = e_a$ and $\{t, v\} = e_b$ for some $a, b \in I$, $a \neq b$, and thus

$$\frac{\sqrt{\text{Var}[T]}}{pq} D_\ell D_k F = \mathbb{1}_{\{X_a=1\}} \quad \text{and} \quad \frac{\sqrt{\text{Var}[T]}}{pq} D_\ell D_j F = \mathbb{1}_{\{X_b=1\}},$$

which implies the independence of $D_\ell D_k F$ and $D_\ell D_j F$ in this case. In the situation of (5.5), we obtain $\{s, v\} = e_a$ and $\{u, v\} = e_b$ for some $a, b \in I$, $a \neq b$, and hence

$$\frac{\sqrt{\text{Var}[T]}}{pq} D_\ell D_k F = \mathbb{1}_{\{X_a=1\}} \quad \text{and} \quad \frac{\sqrt{\text{Var}[T]}}{pq} D_\ell D_j F = \mathbb{1}_{\{X_b=1\}},$$

which implies the independence of $D_\ell D_k F$ and $D_\ell D_j F$ in the second case.

In view of (5.2) and the bound in Theorem 4.1, we need an expression for the fractional moments of a binomial random variable $Z \sim \text{Bin}(n, p)$ with $n \in \mathbb{N}$ and $p \in (0, 1)$. It is well known that

$$\mathbb{E}[Z^\beta] \asymp \begin{cases} (np)^\beta, & \text{if } np \to \infty, \\ np, & \text{if } np \to 0, \end{cases} \qquad \beta \in [1, \infty).$$

As a consequence, we deduce that for $n \in \{3, 4, \ldots\}$, $\alpha \in [0, 1)$ and $\theta \in (0, n^\alpha)$ with $\theta \asymp 1$, the binomial random variable $Z \sim \text{Bin}(n - 2, \theta^2 n^{-2\alpha})$ satisfies

$$(5.6) \qquad \mathbb{E}[Z^\beta] \asymp \begin{cases} \theta^\beta n^{\beta(1-2\alpha)}, & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \theta n^{1-2\alpha}, & \text{if } \frac{1}{2} < \alpha < 1, \end{cases} \qquad \beta \in [1, \infty).$$

With (5.2), (5.3) and (5.6) at hand, we are now prepared for the evaluation of the bound in Theorem 4.1. The following terms have to be considered:

$$A_1 := \left(\sum_{j,k,\ell \in I} \left(\mathbb{E}[(D_j F)^2 (D_k F)^2] \right)^{1/2} \left(\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2] \right)^{1/2} \right)^{1/2},$$

$$A_2 := \left(\sum_{j,k,\ell \in I} \frac{1}{pq} \mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2] \right)^{1/2},$$

$$A_3 := \sum_{k \in I} \frac{1}{\sqrt{pq}} \mathbb{E}[|D_k F|^3],$$

$$A_4 := \left(\mathbb{E}[|F|^r] \right)^{1/r} \sum_{k \in I} \frac{1}{\sqrt{pq}} \left(\mathbb{E}[|D_k F|^{2s}] \right)^{1/s} \left(\mathbb{E}[|D_k F|^t] \right)^{1/t},$$

$$A_5 := \left(\sum_{k \in I} \frac{1}{pq} \mathbb{E}[(D_k F)^4] \right)^{1/2},$$

$$A_6 := \left(\sum_{k,\ell \in I} \frac{1}{pq} \left(\mathbb{E}[(D_k F)^4] \right)^{1/2} \left(\mathbb{E}[(D_\ell D_k F)^4] \right)^{1/2} \right)^{1/2},$$

$$A_7 := \frac{1}{pq}\left(\sum_{k,\ell\in I} \mathbb{E}[(D_\ell D_k F)^4]\right)^{1/2},$$

where in A_4, $r, s, t \in (1, \infty)$ are such that $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = 1$. Let us begin with the term A_1. Using the independence of $D_\ell D_k F$ and $D_\ell D_j F$ for $k \neq j$ as well as the Cauchy–Schwarz inequality we obtain

$$\sum_{j,k,\ell\in I} \left(\mathbb{E}[(D_j F)^2 (D_k F)^2]\right)^{1/2}\left(\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2]\right)^{1/2}$$

$$= \sum_{j,\ell\in I} \left(\mathbb{E}[(D_j F)^4]\right)^{1/2}\left(\mathbb{E}[(D_\ell D_j F)^4]\right)^{1/2}$$

$$+ \sum_{\substack{j,k,\ell\in I\\ k\neq j}} \left(\mathbb{E}[(D_j F)^2 (D_k F)^2]\right)^{1/2}\left(\mathbb{E}[(D_\ell D_j F)^2]\right)^{1/2}\left(\mathbb{E}[(D_\ell D_k F)^2]\right)^{1/2}$$

(5.7)

$$\leq \sum_{j,\ell\in I} \left(\mathbb{E}[(D_j F)^4]\right)^{1/2}\left(\mathbb{E}[(D_\ell D_j F)^4]\right)^{1/2}$$

$$+ \sum_{\substack{j,k,\ell\in I\\ k\neq j}} \left(\mathbb{E}[(D_j F)^4]\right)^{1/4}\left(\mathbb{E}[(D_k F)^4]\right)^{1/4}$$

$$\times \left(\mathbb{E}[(D_\ell D_j F)^2]\right)^{1/2}\left(\mathbb{E}[(D_\ell D_k F)^2]\right)^{1/2}.$$

We consider the two summands of the last estimate separately. Denote by μ_4 the fourth moment of a $\text{Bin}(n - 2, p^2)$-distributed random variable. Using (5.2) and (5.3), we see that

$$\sum_{j,\ell\in I} \left(\mathbb{E}[(D_j F)^4]\right)^{1/2}\left(\mathbb{E}[(D_\ell D_j F)^4]\right)^{1/2}$$

$$= \frac{(pq)^3}{(\text{Var}[T])^2} \sum_{j,\ell\in I} \left(\mathbb{E}\left[\left(\frac{\sqrt{\text{Var}[T]}}{\sqrt{pq}} D_j F\right)^4\right]\right)^{1/2}$$

$$\times \left(\mathbb{E}\left[\left(\frac{\sqrt{\text{Var}[T]}}{pq} D_\ell D_j F\right)^4\right]\right)^{1/2}$$

(5.8)

$$= \frac{(pq)^3}{(\text{Var}[T])^2} \sum_{j,\ell\in I} \mu_4^{1/2} p^{1/2} \mathbb{1}_{\{|e_j\cap e_\ell|=1\}}$$

$$= \frac{(pq)^3}{(\text{Var}[T])^2} \mu_4^{1/2} p^{1/2} \binom{n}{2} 2(n - 2)$$

$$\asymp \frac{(pq)^3}{(\text{Var}[T])^2} \mu_4^{1/2} p^{1/2} n^3.$$

For the second summand on the right-hand side of (5.7), we obtain

$$\sum_{\substack{j,k,\ell \in I \\ k \neq j}} (\mathbb{E}[(D_j F)^4])^{1/4} (\mathbb{E}[(D_k F)^4])^{1/4} (\mathbb{E}[(D_\ell D_j F)^2])^{1/2} (\mathbb{E}[(D_\ell D_k F)^2])^{1/2}$$

$$= \frac{(pq)^3}{(\text{Var}[T])^2} \sum_{\substack{j,k,\ell \in I \\ j \neq k}} \mu_4^{1/2} p \mathbb{1}_{\{|e_j \cap e_\ell| = 1\}} \mathbb{1}_{\{|e_k \cap e_\ell| = 1\}}$$

$$(5.9) \quad = \frac{(pq)^3}{(\text{Var}[T])^2} \mu_4^{1/2} p \binom{n}{2} 2(n-2)(2(n-2)-1)$$

$$\asymp \frac{(pq)^3}{(\text{Var}[T])^2} \mu_4^{1/2} p n^4$$

$$= \frac{(pq)^3}{(\text{Var}[T])^2} \mu_4^{1/2} p^{1/2} n^3 p^{1/2} n.$$

Comparing (5.8) with (5.9), one can see that the second summand in (5.7) determines the asymptotic behaviour of A_1, since $p^{1/2} n = \theta^{1/2} n^{1-\alpha/2} \to \infty$, as $n \to \infty$. By use of (5.1) and (5.6), we obtain

$$(5.10) \quad \frac{(pq)^3}{(\text{Var}[T])^2} \mu_4^{1/2} p n^4 \asymp \begin{cases} \theta^{-4} n^{-2+2\alpha}, & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \theta^{-3/2} n^{-3/2+\alpha}, & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

Combining (5.7)–(5.10) yields that

$$A_1 = \left(\sum_{j,k,\ell \in I} (\mathbb{E}[(D_j F)^2 (D_k F)^2])^{1/2} \right.$$

$$(5.11) \qquad\qquad \left. \times (\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2])^{1/2} \right)^{1/2}$$

$$= \begin{cases} \mathcal{O}(n^{-1+\alpha}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-3/4+\alpha/2}), & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

With the same arguments as above and by using the additional information on the asymptotics of the third moment of a $\text{Bin}(n-2, p^2)$ random variable from (5.6), we obtain the following bounds for A_2, A_3, A_5, A_6 and A_7:

$$(5.12) \quad \begin{aligned} A_2 &= \begin{cases} \mathcal{O}(n^{-2+5\alpha/2}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-1+\alpha/2}), & \text{if } \frac{1}{2} < \alpha < 1, \end{cases} \\ A_3 &= \begin{cases} \mathcal{O}(n^{-1+\alpha/2}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-3/2+3\alpha/2}), & \text{if } \frac{1}{2} < \alpha < 1, \end{cases} \end{aligned}$$

$$A_5 = \begin{cases} \mathcal{O}(n^{-1+\alpha/2}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-3/2+3\alpha/2}), & \text{if } \frac{1}{2} < \alpha < 1, \end{cases}$$

(5.13)

$$A_6 = \begin{cases} \mathcal{O}(n^{-3/2+7\alpha/4}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-5/4+5\alpha/4}), & \text{if } \frac{1}{2} < \alpha < 1, \end{cases}$$

(5.14)

$$A_7 = \begin{cases} \mathcal{O}(n^{-5/2+7\alpha/2}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-3/2+3\alpha/2}), & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

To describe the asymptotic behaviour of

$$A_4 = \left(\mathbb{E}[|F|^r]\right)^{1/r} \sum_{k \in I} \frac{1}{\sqrt{pq}} \left(\mathbb{E}[|D_k F|^{2s}]\right)^{1/s} \left(\mathbb{E}[|D_k F|^t]\right)^{1/t}$$

with $r, s, t \in (1, \infty)$ and $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = 1$, we use the following moment asymptotics, which is taken from the proof of [34], Theorem 2. As $n \to \infty$, it holds that

(5.15)
$$\mathbb{E}[F^k] \asymp \begin{cases} 0, & \text{if } k \in \mathbb{N} \text{ is odd}, \\ \dfrac{k!}{(k/2)! 2^{k/2}}, & \text{if } k \in \mathbb{N} \text{ is even}. \end{cases}$$

We will choose r in such a way that A_4 converges to zero at least as fast as all the other terms $A_1, \dots, A_7$ that have already been computed. So, fix an even integer $r > 2$ and choose $s, t \in (1, \infty)$ such that $\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = 1$. For $\beta \in [1, \infty)$, let μ_β be the moment of order β of a $\text{Bin}(n-2, p^2)$ random variable. Using (5.2), we obtain

(5.16)
$$\begin{aligned} &\frac{1}{\sqrt{pq}} \left(\mathbb{E}[|D_k F|^{2s}]\right)^{1/s} \left(\mathbb{E}[|D_k F|^t]\right)^{1/t} \\ &= \frac{1}{\sqrt{pq}} \frac{(pq)^{3/2}}{(\text{Var}[T])^{3/2}} \left(\mathbb{E}\left[\left(\frac{\sqrt{\text{Var}[T]}}{\sqrt{pq}} D_k F\right)^{2s}\right]\right)^{1/s} \\ &\quad \times \left(\mathbb{E}\left[\left(\frac{\sqrt{\text{Var}[T]}}{\sqrt{pq}} D_k F\right)^t\right]\right)^{1/t} \\ &= \frac{pq}{(\text{Var}[T])^{3/2}} \mu_{2s}^{1/s} \mu_t^{1/t}. \end{aligned}$$

Note that the absolute values are omitted since $D_k F$ is nonnegative. Resorting to (5.1) and (5.6) and using that $\frac{1}{s} + \frac{1}{t} = 1 - \frac{1}{r}$, we get

(5.17)
$$\begin{aligned} &\frac{pq}{(\text{Var}[T])^{3/2}} \mu_{2s}^{1/s} \mu_t^{1/t} \\ &\asymp \begin{cases} \theta^{-7/2} n^{-3+\alpha/2}, & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \theta^{-5/2-1/r} n^{-7/2+3\alpha/2+2\alpha/r-1/r}, & \text{if } \frac{1}{2} < \alpha < 1. \end{cases} \end{aligned}$$

Combining (5.15)–(5.17), we obtain that for all even integers $r > 2$,

$$(5.18) \qquad A_4 = \begin{cases} \mathcal{O}(n^{-1+1/2\alpha}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-3/2+3\alpha/2+2\alpha/r-1/r}), & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

If $0 \leq \alpha \leq \frac{1}{2}$, the bound in (5.18) does not depend on r and is of lower order compared to the bounds in (5.11)–(5.14). In the case $\frac{1}{2} < \alpha < \frac{2}{3}$, the term A_1 in (5.11) determines the leading-order asymptotics among the bounds in (5.11)–(5.14) if $r > 2$ is chosen in such a way that

$$-\frac{3}{2} + \frac{3}{2}\alpha + \frac{2}{r}\alpha - \frac{1}{r} \leq -\frac{5}{4} + \frac{5}{4}\alpha, \quad \text{or, equivalently,} \quad r \geq \frac{4(2\alpha - 1)}{1 - \alpha}.$$

Namely, we put r as the smallest even integer larger or equal to $\max\{2, \frac{4(2\alpha-1)}{1-\alpha}\}$ and conclude from (5.18) that

$$(5.19) \qquad A_4 = \begin{cases} \mathcal{O}(n^{-1+\alpha/2}), & \text{if } 0 \leq \alpha \leq \frac{1}{2}, \\ \mathcal{O}(n^{-5/4+5\alpha/4}), & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

Moreover, if $\frac{2}{3} \leq \alpha < 1$, the Kolmogorov distance is dominated by the term A_6. This completes the proof. $\quad\square$

After having established Theorem 1.1, we turn to the proof of Theorem 1.2. Recall that in this situation $p \in (0, 1)$ is fixed and that Γ is a graph with at least one edge. Furthermore, S is the number of copies of Γ in $G(n, p)$ and $F := (S - \mathbb{E}[S])/\sqrt{\text{Var}[S]}$ denotes the normalized sub-graph counting statistic. Let us recall from [10], Lemma 3.5, that

$$(5.20) \qquad \text{Var}[S] \asymp c(p, \Gamma) n^{2v-2},$$

where $c(p, \Gamma) \in (0, \infty)$ is a constant only depending on p and Γ, and where $v = v(\Gamma)$ stands for the number of vertices of Γ. Finally, we recall that I stands for the set $\{1, \ldots, \binom{n}{2}\}$ and put $q := 1 - p$.

PROOF OF THEOREM 1.2. First, we assume that $n \geq v \geq 4$. Note that for $k \in I$, S_k^+ and S_k^- are the number of copies of Γ if edge e_k is present in $G(n, p)$ or not, respectively. Thus, $S_k^+ - S_k^-$ is the number of copies of Γ in $G(n, p)$ sharing edge e_k. Since there are $\binom{n-2}{v-2}$ choices for the remaining vertices needed to build such a copy, we have that

$$D_k F = \frac{\sqrt{pq}}{\sqrt{\text{Var}[S]}}(S_k^+ - S_k^-) = \mathcal{O}(n^{-1}), \qquad k \in I,$$

where we also used (5.20). Next, we consider the second-order discrete gradient

$$D_\ell D_k F = \frac{pq}{\sqrt{\text{Var}[S]}}((S_k^+)_\ell^+ - (S_k^+)_\ell^- - (S_k^-)_\ell^+ + (S_k^-)_\ell^-), \qquad k, \ell \in I.$$

If $|e_k \cap e_\ell| = 0$, $v - 4$ further vertices are needed to build a copy of Γ containing the edges e_k and e_ℓ. Since there are $\binom{n-4}{v-4}$ choices for these vertices and because of (5.20), one has that

$$(5.21) \qquad D_\ell D_k F = \mathcal{O}(n^{-3}).$$

Similarly, if $|e_k \cap e_\ell| = 1$ we find that

$$(5.22) \qquad D_\ell D_k F = \mathcal{O}(n^{-2})$$

and if $|e_k \cap e_\ell| = 2$ we have $k = \ell$, and hence

$$(5.23) \qquad D_\ell D_k F = 0.$$

We can now evaluate the terms arising in Theorem 4.1, which we denote by $A_1, \ldots, A_7$. For A_1, we have that

$$A_1^2 := \frac{15}{4} \sum_{j,k,\ell \in I} \left(\mathbb{E}[(D_j F)^2 (D_k F)^2] \right)^{1/2} \left(\mathbb{E}[(D_\ell D_j F)^2 (D_\ell D_k F)^2] \right)^{1/2}.$$

Using the Cauchy–Schwarz inequality, we see that

$$A_1^2 \leq \frac{15}{4} \sum_{\ell \in I} \left(\sum_{k \in I} (\mathbb{E}[(D_k F)^4])^{1/4} (\mathbb{E}[(D_\ell D_k F)^4])^{1/4} \right)^2$$

and a distinction of the cases $|e_k \cap e_\ell| = 0$, $|e_k \cap e_\ell| = 1$ and $|e_k \cap e_\ell| = 2$ yields

$$A_1 = \mathcal{O}(n^{-1})$$

by (5.21)–(5.23). Similar considerations with $r = 2$ and $s = t = 4$ lead to

$$A_2 = \mathcal{O}(n^{-2}), \qquad A_3 = \mathcal{O}(n^{-1}), \qquad A_4 = \mathcal{O}(n^{-1}),$$
$$A_5 = \mathcal{O}(n^{-1}), \qquad A_6 = \mathcal{O}(n^{-3/2}), \qquad A_7 = \mathcal{O}(n^{-5/2})$$

and hence to $d_K(F, N) = \mathcal{O}(n^{-1})$.

The case that Γ has exactly two vertices is covered by the classical Berry–Esseen theorem for a binomial distribution with parameters $\binom{n}{2}$ and p. If Γ has exactly three vertices, then Γ is either the complete graph on 3 vertices (as already covered by Theorem 1.1) or a graph with 1 or 2 edges on 3 vertices, respectively. In these cases, instead of (5.21) one has that $D_\ell D_k F = 0$ if $|e_k \cap e_\ell| = 0$ and one obtains that $d_K(F, N) = \mathcal{O}(n^{-1})$. This completes the proof. $\square$

Finally in this section, we turn to the proof of Theorem 1.3 for which we use the same set-up as in the proof of Theorem 1.1. In particular, we denote by I the set $\{1, \ldots, \binom{n}{2}\}$ and recall that $p = \theta n^{-\alpha}$ with suitable $\alpha \in \mathbb{R}$ and $\theta \in (0, n^\alpha)$ such that $\theta \asymp 1$. We also put $q := 1 - p$. For $d \in \{0, 1, 2, \ldots\}$ we denote by $V_{n,d}$ the number

of vertices of $G(n, p)$ with degree d and put $G_{n,d} := (V_{n,d} - \mathbb{E}[V_{n,d}])/\sqrt{\mathrm{Var}[V_{n,d}]}$. Let us recall from Chapter 6.3 in [10] that

$$(5.24) \qquad \mathrm{Var}[V_{n,0}] \asymp 2\theta n^2 p = 2\theta n^{2-\alpha}, \qquad \alpha \in [1, 2),$$

and for $d \in \mathbb{N}$,

$$(5.25) \quad \mathrm{Var}[V_{n,d}] \asymp c(d, \theta) n^{d+1} p^d = c(d, \theta) n^{d+1-\alpha d}, \qquad \alpha \in [1, 1 + 1/d),$$

with a constant $c(d, \theta) \in (0, \infty)$ only depending on d and on θ. From Theorem 8 in [2] it is known that a central limit theorem for $G_{n,0}$ holds if and only if $n^2 p \to \infty$ and $np - \log n \to -\infty$, as $n \to \infty$. In our case that $p = \theta n^{-\alpha}$ this is equivalent to $\alpha \in [1, 2)$. Moreover, [10], Theorem 6.36, says that for $d \in \mathbb{N}$, $G_{n,d}$ satisfies a central limit theorem if and only if $n^{d+1} p^d \to \infty$ and $np - \log n - d \log \log n \to -\infty$, as $n \to \infty$. Again, in our case this is equivalent to $\alpha \in [1, 1 + 1/d)$, whence the conditions on α in (5.24) and (5.25).

PROOF OF THEOREM 1.3.   At first, we notice that adding or removing an edge from $G(n, p)$ can change the number of vertices with degree equal to $d \in \{0, 1, 2, \ldots\}$ by at most 2. This implies that

$$(5.26) \qquad |D_k G_{n,d}| \leq \frac{2\sqrt{pq}}{\sqrt{\mathrm{Var}[V_{n,d}]}}, \qquad k \in I.$$

Next, we observe that $(pq)^{-1}|D_k D_\ell V_{n,d}| \in \{0, 1, 2\}$ for all $k, \ell \in I$. We also have that $D_k D_\ell V_{n,d}$, and hence $D_k D_\ell G_{n,d}$ is zero whenever the two corresponding edges e_k and e_ℓ are identical or do not share a common vertex. Resorting to the definition of the random variable $G_{n,d}$, we thus conclude that

$$(5.27) \qquad |D_k D_\ell G_{n,d}| \leq \frac{2pq}{\sqrt{\mathrm{Var}[V_{n,d}]}} \mathbb{1}_{\{|e_k \cap e_\ell| = 1\}}, \qquad k, \ell \in I.$$

We can now evaluate the bound in Theorem 4.1. We start with the case $d = 0$. Since the computations are almost identical for each of the terms there, we restrict to the first term A_1, which is given by

$$A_1 := \left(\frac{15}{4} \sum_{j,k,\ell \in I} \left(\mathbb{E}[(D_j G_n)^2 (D_k G_n)^2] \right)^{1/2} \left(\mathbb{E}[(D_\ell D_j G_n)^2 (D_\ell D_k G_n)^2] \right)^{1/2} \right)^{1/2}.$$

Using (5.26) and (5.27), we see that

$$
\begin{aligned}
A_1^2 &\leq \frac{60(pq)^3}{(\mathrm{Var}[V_{n,d}])^2} \sum_{j,k,\ell \in I} \mathbb{1}_{\{|e_j \cap e_\ell| = 1\}} \mathbb{1}_{\{|e_k \cap e_\ell| = 1\}} \\
&= \frac{60(pq)^3}{(\mathrm{Var}[V_{n,d}])^2} \binom{n}{2} (2(n-2))^2.
\end{aligned}
$$

(5.28)

Now, we use that $p = \theta n^{-\alpha}$ as well as the variance asymptotics at (5.24). This allows us to conclude that $A_1 = \mathcal{O}(n^{-\alpha/2})$. Denoting the other terms arising in Theorem 4.1 by $A_2, \ldots, A_7$, we conclude by similar computations and by taking $r = 2, s = t = 4$ that

$$A_2 = \mathcal{O}(n^{-\alpha/2}), \qquad A_3 = \mathcal{O}(n^{-1+\alpha/2}), \qquad A_4 = \mathcal{O}(n^{-1+\alpha/2}),$$
$$A_5 = \mathcal{O}(n^{-1+\alpha/2}), \qquad A_6 = \mathcal{O}(n^{-1/2}), \qquad A_7 = \mathcal{O}(n^{-1/2}).$$

Thus, $d_K(G_{n,0}, N) = \mathcal{O}(n^{-1+\alpha/2})$. Turning to the case $d \in \mathbb{N}$, we start again with the term A_1 and obtain by using (5.28) and (5.25) that $A_1 = \mathcal{O}(n^{1-d-3\alpha/2+\alpha d})$. Moreover, one sees for the other terms $A_2, \ldots, A_7$ in Theorem 4.1 that

$$A_2 = \mathcal{O}(n^{1-d-3\alpha/2+\alpha d}), \qquad A_3 = \mathcal{O}(n^{1/2-3d/2-\alpha+3\alpha d/2}),$$
$$A_4 = \mathcal{O}(n^{1/2-3d/2-\alpha+3\alpha d/2}), \qquad A_5 = \mathcal{O}(n^{-d-\alpha/2+\alpha d}),$$
$$A_6 = \mathcal{O}(n^{1/2-d-\alpha+\alpha d}), \qquad A_7 = \mathcal{O}(n^{1/2-d-\alpha+\alpha d}).$$

Thus, for $d \in \mathbb{N}$, $d_K(G_{n,d}, N) = \mathcal{O}(n^{1/2-3d/2-\alpha+3\alpha d/2})$, provided that $\alpha \in [1, (3d-1)/(3d-2))$. This completes the proof. $\quad\square$

6. Application to percolation on trees and proof of Theorem 1.4. Let us recall some notation and embed the objects into the framework of Sections 2 and 4. We denote by $\mathcal{T}$ an infinite rooted tree such that each vertex has degree bounded by $D + 1$ with $D \in \mathbb{N}$. By $\mathcal{T}_n$, $n \in \mathbb{N}$, we indicate the finite sub-tree of $\mathcal{T}$ consisting of all vertices with graph-distance at most n from the root. We now embed $\mathcal{T}$ into the Euclidean plane by the following procedure, which is illustrated in Figure 3. The root is mapped to the point with coordinates $(1, 1)$ and the vertices adjacent to it are mapped to the points with coordinates $(1, 2), \ldots, (N(1), 2)$ with $N(1) \le D$ in an arbitrary order. Next, the vertices adjacent to these are mapped onto $(1, 3), \ldots, (N(2), 3)$, where (from left to right)
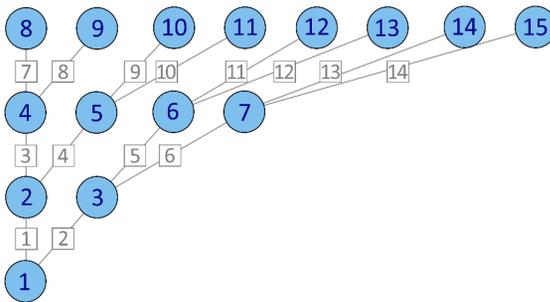

FIG. 3. *Embedding in the plane of $\mathcal{T}_3$ of a 2-regular tree $\mathcal{T}$.*

the first points in this list are adjacent to $(1, 2)$, the next points to $(2, 2)$, etc. Continuing this way, the vertices with graph-distance n to the root are mapped onto $(1, n + 1), \ldots, (N(n), n + 1)$ with $N(n) \leq N(n - 1)D$ and the infinite tree $\mathscr{T}$ is embedded into the upper right quadrant of the Euclidean plane. A vertex of the embedded tree with coordinates (i, k) for $k \in \mathbb{N}$ and $i \in \{1, \ldots, N(k)\}$ receives the label $1 + N(1) + \cdots + N(k - 1) + i$ and an edge of the embedded tree whose adjacent vertices have coordinates (i, k) and $(j, k - 1)$ for $k \in \{2, 3, \ldots\}$, $i \in \{1, \ldots, N(k)\}$ and $j \in \{1, \ldots, N(k - 1)\}$ finally receives the label of its endpoint minus one, that is, $N(1) + \cdots + N(k - 1) + i$; see Figure 3. This numbering of vertices also corresponds to that in Figure 2.

This construction puts us in the position to interpret our percolation problem on $\mathscr{T}$ in terms of the framework of Theorem 4.1. Namely, for fixed $p \in (0, 1)$ let $(X_k)_{k \in \mathbb{N}}$ be a sequence of independent Rademacher random variables with $P(X_k = +1) = p$ and $P(X_k = -1) = 1 - p$. For each $k \in \mathbb{N}$, assign the random variable X_k to the uniquely determined edge e_k of the embedded tree with label k. The random graph $\mathscr{T}(p)$ consists of all edges e_k of the embedded tree with label $X_k = +1$ together with their two adjacent vertices. Thus, $\mathscr{T}(p)$ is described by the Rademacher sequence $(X_k)_{k \in \mathbb{N}}$ and its restriction $\mathscr{T}_n(p)$ to $\mathscr{T}_n$ is described by a finite sub-sequence of $(X_k)_{k \in \mathbb{N}}$.

For $n \in \mathbb{N}$, we denote by $C_n(p)$ the number of connected components of the random graph $\mathscr{T}_n(p)$, where, as already discussed in the Introduction, by a connected component we understand a maximal connected sub-graph with at least one edge. By $H_n(p) := (C_n(p) - \mathbb{E}[C_n(p)]) / \sqrt{\mathrm{Var}[C_n(p)]}$, we denote the normalized version of $C_n(p)$ and notice that $C_n(p)$ is a Rademacher functional.

PROOF OF THEOREM 1.4. We start by investigating the first- and second-order discrete gradient applied to $H_n(p)$. By definition, we have that

$$D_k H_n(p) = \frac{\sqrt{pq}}{\sqrt{\mathrm{Var}[C_n(p)]}} D_k C_n(p) = \frac{\sqrt{pq}}{\sqrt{\mathrm{Var}[C_n(p)]}} \big((C_n(p))_k^+ - (C_n(p))_k^- \big),$$

where $k \in \{1, \ldots, 1 + N(1) + \cdots + N(n)\}$. Note that $D_k C_n(p)$ is a local quantity since it depends only on the edges adjacent to k. Adding or removing the edge with label k can change the number of connected components by at most 1. Therefore, we have that

$$(6.1) \qquad \qquad |D_k H_n(p)| \leq \frac{\sqrt{pq}}{\sqrt{\mathrm{Var}[C_n(p)]}}$$

for all $k \in \{1, \ldots, 1 + N(1) + \cdots + N(n)\}$. Next, we consider for $k, j \in \{1, \ldots, 1 + N(1) + \cdots + N(n)\}$ the second-order discrete gradient

$$D_k D_j H_n(p) = \frac{pq}{\sqrt{\mathrm{Var}[C_n(p)]}} \big(((C_n(p))_j^+)_k^+ - ((C_n(p))_j^+)_k^-$$
$$- ((C_n(p))_j^-)_k^+ + ((C_n(p))_j^-)_k^- \big).$$

For most choices of j and k, $D_k D_j H_n(p)$ is zero. A nonzero contribution only arises if the edges e_j and e_k with labels j and k, respectively, share precisely one common vertex. We indicate this situation by $|e_j \cap e_k| = 1$ and write $|e_j \cap e_k| \in \{0, 2\}$ otherwise. Thus, we can use the triangle inequality and the estimate (6.1) to conclude that

$$(6.2) \qquad |D_j D_k H_n(p)| \begin{cases} = 0, & \text{if } |e_j \cap e_k| \in \{0, 2\}, \\ \leq \dfrac{2pq}{\sqrt{\operatorname{Var}[C_n(p)]}}, & \text{if } |e_j \cap e_k| = 1. \end{cases}$$

We use a lower bound for the variance of $C_n(p)$, which can be found in [37], Identity (2.3), in case of a D-regular tree, but the proof is easily seen to carry over to our situation. More precisely, there exists a constant $c(p) > 0$ only depending on p such that

$$(6.3) \qquad \operatorname{Var}[C_n(p)] \geq c(p)|\mathscr{T}_n|.$$

Estimating the terms in Theorem 4.1 with $r = 2$ and $s = t = 4$ there by means of (6.1)–(6.3) yields (after a straightforward computation similar to the one in the proof of Theorem 1.3) that

$$d_K\big(H_n(p), N\big) = \mathcal{O}\big(|\mathscr{T}_n|^{-1/2}\big).$$

In case of a D-regular tree, we have that $|\mathscr{T}_n| = D + \cdots + D^n = (D^{n+1} - 1)/(D - 1) - 1$, if $D \geq 2$, and $|\mathscr{T}_n| = n$, if $D = 1$. Thus, $|\mathscr{T}_n|$ behaves like D^n, if $D \geq 2$, and like n, if $D = 1$, as $n \to \infty$. This completes the proof. $\square$

REFERENCES

[1] AZAÏS, J.-M. and LEÓN, J. R. (2013). CLT for crossings of random trigonometric polynomials. *Electron. J. Probab.* **18** 1–17.

[2] BARBOUR, A. D., KAROŃSKI, M. and RUCIŃSKI, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47** 125–145.

[3] CHATTERJEE, S. (2009). Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields* **143** 1–40.

[4] DECREUSEFOND, L., SCHULTE, M. and THÄLE, C. (2016). Functional Poisson approximation in Kantorovich–Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. Probab.* **44** 2147–2197. MR3502603

[5] DURASTANTI, C., MARINUCCI, D. and PECCATI, G. (2014). Normal approximations for wavelet coefficients on spherical Poisson fields. *J. Math. Anal. Appl.* **409** 212–227. MR3095032

[6] EICHELSBACHER, P. and THÄLE, C. (2014). New Berry–Esseen bounds for non-linear functionals of Poisson random measures. *Electron. J. Probab.* **19** no. 102, 25. MR3275854

[7] ERDŐS, P. and RÉNYI, A. (1959). On random graphs. I. *Publ. Math. Debrecen* **6** 290–297. MR0120167

[8] FÉRAY, V., MÉLIOT, P.-L. and NIKEGHBALI, A. (2013). Mod-phi convergence and precise deviations. Available at arXiv:1304.2934.

[9] GOLDSTEIN, L. (2013). A Berry–Esseen bound with applications to vertex degree counts in the Erdős–Rényi random graph. *Ann. Appl. Probab.* **23** 617–636. MR3059270

[10] JANSON, S., ŁUCZAK, T. and RUCINSKI, A. (2000). *Random Graphs*. Wiley-Interscience, New York. MR1782847

[11] JANSON, S. and NOWICKI, K. (1991). The asymptotic distributions of generalized *U*-statistics with applications to random graphs. *Probab. Theory Related Fields* **90** 341–375. MR1133371

[12] KORDECKI, W. (1990). Normal approximation and isolated vertices in random graphs. In *Random Graphs '87 (Poznań, 1987)* 131–139. Wiley, Chichester. MR1094128

[13] KROKOWSKI, K. (2015). Poisson approximation of Rademacher functionals by the Chen–Stein method and Malliavin calculus. Available at arXiv:1505.01417.

[14] KROKOWSKI, K., REICHENBACHS, A. and THÄLE, C. (2016). Berry–Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. H. Poncaré Probab. Statist.* To appear.

[15] LACHIÈZE-REY, R. and PECCATI, G. (2013). Fine Gaussian fluctuations on the Poisson space, I: Contractions, cumulants and geometric random graphs. *Electron. J. Probab.* **18** no. 32, 32. MR3035760

[16] LACHIÈZE-REY, R. and PECCATI, G. (2013). Fine Gaussian fluctuations on the Poisson space II: Rescaled kernels, marked processes and geometric *U*-statistics. *Stochastic Process. Appl.* **123** 4186–4218. MR3096352

[17] LAST, G., PECCATI, G. and SCHULTE, M. (2016). Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probab. Theory Related Fields* **165** 667–723. MR3520016

[18] LAST, G., PENROSE, M. D., SCHULTE, M. and THÄLE, C. (2014). Moments and central limit theorems for some multivariate Poisson functionals. *Adv. in Appl. Probab.* **46** 348–364. MR3215537

[19] MARINUCCI, D. and PECCATI, G. (2011). *Random Fields on the Sphere*: *Representation*, *Limit Theorems and Cosmological Applications*. **389**. Cambridge Univ. Press, Cambridge. MR2840154

[20] NOURDIN, I. and PECCATI, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** 75–118. MR2520122

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

[22] NOURDIN, I. and PECCATI, G. (2012). *Normal Approximations with Malliavin Calculus*: *From Stein's Method to Universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge Univ. Press, Cambridge. MR2962301

[23] NOURDIN, I., PECCATI, G. and REINERT, G. (2009). Second order Poincaré inequalities and CLTs on Wiener space. *J. Funct. Anal.* **257** 593–609. MR2527030

[24] NOURDIN, I., PECCATI, G. and REINERT, G. (2010). Stein's method and stochastic analysis of Rademacher functionals. *Electron. J. Probab.* **15** 1703–1742. MR2735379

[25] NUALART, D. (1995). *The Malliavin Calculus and Related Topics*. Springer, New York. MR1344217

[26] PECCATI, G., SOLÉ, J. L., TAQQU, M. S. and UTZET, F. (2010). Stein's method and normal approximation of Poisson functionals. *Ann. Probab.* **38** 443–478. MR2642882

[27] PECCATI, G. and TAQQU, M. S. (2011). *Wiener Chaos*: *Moments*, *Cumulants and Diagrams*. *Bocconi & Springer Series* **1**. Springer, Milan. MR2791919

[28] PECCATI, G. and THÄLE, C. (2013). Gamma limits and U-statistics on the Poisson space. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** 525–560. MR3083936

[29] PIKE, J. and REN, H. (2014). Stein's method and the Laplace distribution. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** 571–587. MR3283586

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

[31] PRIVAULT, N. and TORRISI, G. L. (2015). The Stein and Chen–Stein methods for functionals of non-symmetric Bernoulli processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** 309–356. MR3357130

[32] RAIČ, M. (2003). Normal approximation by Stein's method. In *Proceedings of the Seventh Young Statisticians Meeting. Metodoloski Zvezki* **21**. FDV, Ljubljana.

[33] REITZNER, M. and SCHULTE, M. (2013). Central limit theorems for U-statistics of Poisson point processes. *Ann. Probab.* **41** 3879–3909. MR3161465

[34] RUCIŃSKI, A. (1988). When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields* **78** 1–10. MR0940863

[35] SAULIS, L. and STATULEVIČIUS, V. A. (1991). *Limit Theorems for Large Deviations. Mathematics and Its Applications* (*Soviet Series*) **73**. Kluwer Academic, Dordrecht. MR1171883

[36] SCHULTE, M. and THÄLE, C. (2012). The scaling limit of Poisson-driven order statistics with applications in geometric probability. *Stochastic Process. Appl.* **122** 4096–4120. MR2971726

[37] SUGIMINE, N. and TAKEI, M. (2006). Remarks on central limit theorems for the number of percolation clusters. *Publ. Res. Inst. Math. Sci.* **42** 101–116. MR2215437

FACULTY OF MATHEMATICS
RUHR UNIVERSITY BOCHUM
D-44780 BOCHUM
GERMANY
E-MAIL: kai.krokowski@rub.de
anselm.reichenbachs@rub.de
christoph.thaele@rub.de