Normal approximation for generalized U-statistics and weighted random graphs

Nicolas Privault* Grzegorz Serafin†

July 26, 2021

Abstract

We derive normal approximation bounds in the Wasserstein distance for sums of generalized U-statistics, based on a general distance bound for functionals of independent random variables of arbitrary distributions. Those bounds are applied to normal approximation for the combined weights of subgraphs in the Erdős-Rényi random graph, extending the graph counting results of [1] to the setting of weighted graphs. Our approach relies on a general stochastic analytic framework for functionals of independent random sequences.

Keywords: Stein-Chen method; normal approximation; Malliavin-Stein method; central limit theorem; random graph; subgraph count.

Mathematics Subject Classification: 60F05, 60H07, 60G50, 05C80.

1 Introduction

The Malliavin calculus has been applied to the derivation of approximation bounds by the Stein and Chen-Stein methods on the Wiener space [11], on the Poisson space [13], [14], as well as in the case of discrete Bernoulli sequences [12], [18], [9], [10], [8]. Recently, a different Malliavin framework for Stein approximation has been introduced in [16], with application to normal approximation in the Wasserstein distance for weighted U-statistics of the form

$$\sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}_0\\ k_i\neq k_j \text{ if } i\neq j}} b_{k_1}\cdots b_{k_n} Z_{k_1}\cdots Z_{k_n}$$

where $\mathbb{N}_0 = \{0,1,2,\ldots\}$, $(Z_k)_{k\geq 1}$ is an i.i.d. sequence of random variables, and $(b_k)_{k\geq}$ is a sequence of real coefficients, based on stochastic analysis for functionals of a countable

*School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, Singapore. e-mail: `nprivault@ntu.edu.sg`.

†Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wrocław, Poland. e-mail: `grzegorz.serafin@pwr.edu.pl`.

number of uniformly distributed random variables, see [15]. This completes the bounds for the Kolmogorov distance obtained in e.g. Theorem 3.1 of [2] for nonweighted U-statistics, see also [6] in the quadratic case.

Our goal in the present paper is two-fold. First, we extend in Theorem 3.2 the Stein approximation bounds of [16] from multiple stochastic integrals to finite sums of multiple stochastic integrals, which can be viewed as polynomial functionals in independent random variables with arbitrary distributions, or as generalized U-statistics, see Proposition 3.1. Furthermore, in Proposition 2.4 we obtain a general Wasserstein distance bound for functionals of independent random variables as a consequence of Proposition 3.3 in [16].

Second, we show that those results can be applied to the central limit theorem for the convergence of renormalized weight counts in large random graphs. For this, we consider the Erdős-Rényi random graph $\mathbb{G}(n,p)$, introduced by Gilbert [5] in 1959 and popularized in [3], which is constructed by independently retaining any edge in the complete graph K_n on n vertices with probability $p \in (0,1)$. Denote by N_n^G the random variable counting number of subgraphs (not necessarily induced ones) of $\mathbb{G}(n,p_n)$ that are isomorphic to a fixed graph G. Necessary and sufficient conditions for the asymptotic normality of the renormalization

$$\widetilde{N}_n^G := \frac{N_n^G - \mathbb{E}[N_n^G]}{\sqrt{\mathrm{Var}[N_n^G]}}.$$

of N_n^G have been obtained in [21], where it is shown that

$$\widetilde{N}_n^G \xrightarrow{\mathcal{D}} \mathcal{N} \quad \text{iff} \quad np_n^\beta \to \infty \quad \text{and} \quad n^2(1-p_n) \to \infty, \tag{1.1}$$

as n tends to infinity, where $\mathcal{N} \sim \mathcal{N}(0,1)$ is a standard Gaussian random variable, $\beta = \beta(G) := \max\{e_H/v_H \ : \ H \subset G\}$ and e_H, v_H respectively denote the numbers of edges and vertices in the graph H. Such results have been improved via explicit convergence rates obtained in [1] as

$$d_W\left(\widetilde{N}_n^G, \mathcal{N}\right) \le C \left((1-p_n) \min_{\substack{H \subset G \\ e_H \ge 1}} n^{v_H} p_n^{e_H} \right)^{-1/2}, \tag{1.2}$$

where $C > 0$ is a constant depending on G, and d_W is the Wasserstein distance

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

between the laws of random variables X, Y, where $\mathrm{Lip}(1)$ denotes the class of real-valued Lipschitz functions with Lipschitz constant less than or equal to 1. Kolmogorov distance bounds have also been obtained for triangle counting, see § 3.2.1 of [20], and [10], using the Malliavin approach to the Stein method for discrete Bernoulli sequences. Those rates have been improved in [19], and extensions to the counting of arbitrary subgraphs that yield the bound (1.2) for the Kolmogorov distance have recently been obtained in [17], based on distance bounds for sums of discrete multiple integrals and weighted U-statistics, as well as in [4], [24].

Here, our stochastic analytic framework allows us to assign an independent sample of a random nonnegative weight X to every edge in $\mathbb{G}(n, p_n)$, and to consider the combined weights of subgraphs instead of counting them. Precisely, we define a weight of a graph as a sum of weights of its edges. Next, by W_n^G we denote the combined weight of subgraphs in $\mathbb{G}(n, p_n)$ that are isomorphic to a fixed graph G and its renormalization

$$\widetilde{W}_n^G := \frac{W_n^G - \mathbb{E}[W_n^G]}{\sqrt{\mathrm{Var}[W_n^G]}}. \tag{1.3}$$

In Theorem 4.3 we show, as an application of Corollary 3.2, that when G is a graph without isolated vertices, we have

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \leq C \frac{\sqrt{\mathbb{E}[(X - \mathbb{E}[X])^4]} + (1-p_n)(\mathbb{E}[X])^2}{\mathrm{Var}[X] + (1-p_n)(\mathbb{E}[X])^2} \left((1-p_n) \min_{\substack{H \subset G \\ e_H \geq 1}} n^{v_H} p_n^{e_H} \right)^{-1/2}, \tag{1.4}$$

where $C > 0$ is a constant depending only on e_G, which recovers (1.2) in the case of a deterministic weight given by $X := 1/e_G$. When X is a fixed random variable this also yields the sufficient condition

$$\big(np_n^\beta \to \infty \;\text{ and }\; n^2(1-p_n) \to \infty\big) \Longrightarrow \widetilde{N}_n^G \xrightarrow{\mathcal{D}} \mathcal{N},$$

for the convergence of $\widetilde{W}_n^G$ to the standard normal distribution (cf. (1.1)), which follows from the equivalence

$$\big(np_n^\beta \to \infty \;\text{ and }\; n^2(1-p_n) \to \infty\big) \Longleftrightarrow (1-p_n) \min_{\substack{H \subset G \\ e_H \geq 1}} n^{v_H} p_n^{e_H} \to \infty.$$

To derive the bound (1.4) we apply Proposition 2.4 to combined subgraph weights W_n^G represented as finite sums of multiple stochastic integrals, see Lemma 4.2. Our results are then specialized to a class of graphs satisfying a certain balance condition, which includes triangles, complete graphs and trees as particular cases.

We note that other types of random functionals on graphs, such as graph weights defined as products of edge weights, or the number of vertices of a given degree, admit representations as sums of multiple integrals or generalized U-statistics, and can be treated by this approach.

This paper is organized as follows. In Section 2 we recall the framework of [15] for the construction of random functionals of uniform random variables, together with the construction of derivation operators. In Section 3 we derive normal Stein approximation bounds for general functionals and for sums of multiple stochastic integrals. In Section 4 we show that combined graph weights can be represented as sums of multiple stochastic integrals, and derive distance bounds for the renormalized weights of graphs in $\mathbb{G}(n, p_n)$ that are isomorphic to a fixed graph G. The Appendix Section 5 contains some technical results exploited in the paper.

2 Functionals of uniform random sequences

Stochastic integrals

Given $(U_k)_{k\in\mathbb{N}}$ an i.i.d. sequence of $[-1,1]$-valued uniform random variables on a probability space $(\Omega, \mathcal{F}, P) = ([-1,1]^{\mathbb{N}}, \mathcal{F}, P)$ let the jump process $(Y_t)_{t\in\mathbb{R}_+}$ be defined as

$$Y_t := \sum_{k=0}^{\infty} \mathbf{1}_{[2k+1+U_k,\infty)}(t), \qquad t \in \mathbb{R}_+.$$

Denoting by $(\mathcal{F})_{t\in\mathbb{R}_+}$ the filtration generated by $(Y_t)_{t\in\mathbb{R}_+}$, and letting

$$\widetilde{\mathcal{F}}_t := \mathcal{F}_{2k}, \qquad 2k \le t < 2k+2, \qquad k \in \mathbb{N},$$

the compensated stochastic integral

$$\int_0^\infty u_t d(Y_t - t/2)$$

with respect to the compensated point process $(Y_t - t/2)_{t\in\mathbb{R}_+}$ can be defined for square-integrable $\big(\widetilde{\mathcal{F}}_t\big)_{t\in\mathbb{R}_+}$-adapted processes $(u_t)_{t\in\mathbb{R}_+}$ by the isometry relation

$$\mathbb{E}\left[\int_0^\infty u_t d(Y_t - t/2) \int_0^\infty v_t d(Y_t - t/2)\right] \tag{2.1}$$
$$= \mathbb{E}\left[\int_0^\infty u_t \left(v_t - \frac{1}{2}\sum_{k=0}^{\infty} \mathbf{1}_{(2k,2k+2]}(t) \int_{2k}^{2k+2} v_r dr\right) \frac{dt}{2}\right],$$

see [15], where $(u_t)_{t\in\mathbb{R}_+}$ and $(v_t)_{t\in\mathbb{R}_+}$ are square-integrable $\big(\widetilde{\mathcal{F}}_t\big)_{t\in\mathbb{R}_+}$-adapted processes. This also implies the bound

$$\mathbb{E}\left[\left(\int_0^\infty u_t d(Y_t - t/2)\right)^2\right] \le \frac{1}{2}\mathbb{E}\left[\int_0^\infty |u_t|^2 dt\right],$$

where $(u_t)_{t\in\mathbb{R}_+}$ is square-integrable and $\big(\widetilde{\mathcal{F}}_t\big)_{t\in\mathbb{R}_+}$-adapted.

Multiple stochastic integrals

Let $\widehat{L}^p(\mathbb{R}_+^n)$ denote the space of symmetric functions that are p-integrable on $\mathbb{R}_+^n$, $p \ge 1$, and vanish outside of

$$\Delta_n := \bigcup_{\substack{0\le k_i\neq k_j\\ 1\le i\neq j\le n}} [2k_1, 2k_1+2]\times\cdots\times[2k_n, 2k_n+2],$$

equipped with the norm

$$\|f_n\|_{\widehat{L}^p(\mathbb{R}_+^n)} := \|f_n\|_{L^p(\mathbb{R}_+^n,(dx/2)^{\otimes n})} = \frac{1}{2^{n/p}}\|f_n\|_{L^p(\mathbb{R}_+^n,(dx)^{\otimes n})}, \qquad f_n \in \widehat{L}^p(\mathbb{R}_+^n).$$

Given $f_n \in \widehat{L}^1(\mathbb{R}_+^n)\cap\widehat{L}^2(\mathbb{R}_+^n)$, $n\ge 1$, we define the multiple stochastic integral $I_n(f_n)$ as

$$\begin{aligned} I_n(f_n) :=& \sum_{r=0}^n \frac{(-1)^{n-r}}{2^{n-r}}\binom{n}{r} \qquad (2.2)\\ & \sum_{\substack{k_1,\ldots,k_r\in\mathbb{N}_0\\ k_i\neq k_j \text{ if } i\neq j}} \int_0^\infty\cdots\int_0^\infty f_n(2k_1+1+U_{k_1},\ldots,2k_r+1+U_{k_r},y_1,\ldots,y_{n-r})dy_1\cdots dy_{n-r}\\ =& \; n!\int_0^\infty\int_0^{t_n}\cdots\int_0^{t_2} f_n(t_1,\ldots,t_n)d(Y_{t_1}-t_1/2)\cdots d(Y_{t_n}-t_n/2), \end{aligned}$$

where $\mathbb{N}_0 = \{0,1,2,\ldots\}$. The multiple stochastic integrals $(I_n(f_n))_{n\ge1}$ form a family of mutually orthogonal centred random variables with the bound

$$\mathbb{E}\big[(I_n(f_n))^2\big] \le n!\,\|f_n\|^2_{\widehat{L}^2(\mathbb{R}_+^n,dx/2)}, \qquad n\ge 1, \tag{2.3}$$

cf. (2.1) above and Propositions 4 and 6 of [15], which allows one to extend the definition of $I_n(f_n)$ to all $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$. If in addition we have

$$\int_{2k}^{2k+2} f_n(t,\cdot)dt = 0, \qquad k\in\mathbb{N}, \tag{2.4}$$

then $I_n(f_n)$ satisfies the isometry and orthogonality relation

$$\mathbb{E}\left[I_n(f_n)I_m(f_m)\right] = \mathbf{1}_{\{n=m\}}n!\langle f_n, f_m\rangle_{\widehat{L}^2(\mathbb{R}_+^n,dx/2)}, \quad f_n\in\widehat{L}^2(\mathbb{R}_+^n), \quad f_m\in\widehat{L}^2(\mathbb{R}_+^m), \tag{2.5}$$

see [15], page 589, in other words the function f_n is canonical [23]. Moreover, every $F \in L^2(\Omega)$ admits the chaos decomposition

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} I_n(f_n), \tag{2.6}$$

for some sequence $(f_n)_{n\geq 1}$ of functions in $\widehat{L}^2(\mathbb{R}_+^n)$, $n \geq 1$, see Proposition 7 of [15]. Note that under the condition (2.4) the sequence $(f_n)_{n\geq 1}$ is unique in $\widehat{L}^2(\mathbb{R}_+^n)$ due to the isometry relation (2.5).

In the sequel, sequences $(X_1, \ldots, X_n)$ of independent random variables with distribution functions $(F_{X_1}, \ldots, F_{X_n})$ will be frequently represented as

$$\left(F_{X_1}^{-1}\left(\frac{U_1+1}{2}\right), \ldots, F_{X_n}^{-1}\left(\frac{U_n+1}{2}\right)\right),$$

where $\left(F_{X_1}^{-1}, \ldots, F_{X_n}^{-1}\right)$ are the generalized inverses of $(F_{X_1}, \ldots, F_{X_n})$. As a consequence we have the following remarks, that can be analogously applied to higher order integrals, particularly under the additional assumption (2.4).

Remark 2.1 *For $f_1 \in L^2([0,2n])$, the stochastic integral*

$$I_1(f_1) := \sum_{k=0}^{n-1} \left(f_1(2k+1+U_k) - \frac{1}{2}\int_{2k}^{2k+2} f_1(t)dt \right)$$

represents a sum of independent centred random variables

$$I_1(f_1) \stackrel{d}{=} \sum_{k=1}^{n} (X_k - \mathbb{E}[X_k])$$

by taking $f_1(x) = F_{X_k}^{-1}(x/2-k)$, $x \in [2k-2, 2k)$, $1 \leq k \leq n$.

Remark 2.2 *In terms of U-statistics, (2.4) says that $I_n(f_n)$ is degenerate, and in that sense the expansion (2.6) might be identified as a Hoeffding decomposition. To see this, assuming that $X_1, \ldots, X_n$ are independent identically distributed with common distribution function F_X, we extend the construction of Remark 2.1 to $n = 2$ by letting*

$$f_2(x,y) := \frac{1}{n(n-1)} h\left(F_X^{-1}\left(\frac{x-2i}{2}\right), F_X^{-1}\left(\frac{x-2j}{2}\right)\right), \quad (x,y) \in [2i-2, 2i) \times [2j-2, 2j),$$

$1 \leq i, j \leq n$, for a given function $h : \mathbb{R}^2 \to \mathbb{R}$. Under Condition (2.4) we have

$$\mathbb{E}\left[h(X_1, x)\right] = \mathbb{E}\left[h(x, X_1)\right] = 0 \quad a.s., \tag{2.7}$$

and therefore $I_2(f_2)$ can be written as the classical U-statistic

$$I_2(f_2) \stackrel{d}{=} \frac{1}{n(n-1)} \sum_{\substack{1\le i,j\le n\\ i\neq j}} h(X_i, X_j),$$

which is degenerate by (2.7). Note that if h may depend on $i,j \in \{1,\dots,n\}$, then $I_2(f_2)$ is called a degenerate generalized U-statistic.

Finite difference operator

Consider the finite difference operator ∇ defined on multiple stochastic integrals $F = I_n(f_n)$ as

$$\nabla_t F := F \circ \Phi_t - \frac{1}{2} \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} F \circ \Phi_s ds, \qquad t \in \mathbb{R}_+,$$

where $\Phi_t : \Omega \longrightarrow \Omega$ is defined by

$$\Phi_t(\omega) := \big(U_1(\omega), \dots, U_{\lfloor t/2\rfloor-1}(\omega), t - 2\lfloor t/2\rfloor - 1, U_{\lfloor t/2\rfloor+1}(\omega), \dots\big), \quad \omega \in \Omega, \quad t \in \mathbb{R}_+.$$

Example 2.3 *Let $n \ge 1$, and consider*

$$F := I_2(f_2) = \sum_{\substack{1\le i,j\le n\\ i\neq j}} U_i U_j,$$

where

$$f_2(x,y) := (x-1-2i)(y-1-2j), \quad (x,y) \in [2i-2, 2i) \times [2j-2, 2j), \quad 1 \le i,j \le n.$$

Then we have

$$\begin{aligned}
\nabla_t F = & \sum_{\substack{1\le i,j\le n,\ i\neq j\\ i,j\neq\lfloor t/2\rfloor+1}} U_i U_j + 2(t-2\lfloor t/2\rfloor-1) \sum_{\substack{1\le i\le n\\ i\neq\lfloor t/2\rfloor+1}} U_i \\
& - \frac{1}{2} \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(\sum_{\substack{1\le i,j\le n,\ i\neq j\\ i,j\neq\lfloor s/2\rfloor+1}} U_i U_j + 2(s-2\lfloor s/2\rfloor-1) \sum_{\substack{1\le i\le n\\ i\neq\lfloor s/2\rfloor+1}} U_i \right) ds \\
= & \sum_{\substack{1\le i,j\le n,\ i\neq j\\ i,j\neq\lfloor t/2\rfloor+1}} U_i U_j + 2(t-2\lfloor t/2\rfloor-1) \sum_{\substack{1\le i\le n\\ i\neq\lfloor t/2\rfloor+1}} U_i \\
& - \left(\sum_{\substack{1\le i,j\le n,\ i\neq j\\ i,j\neq\lfloor t/2\rfloor+1}} U_i U_j + \int_{-1}^{1} u du \sum_{\substack{1\le i\le n\\ i\neq\lfloor t/2\rfloor+1}} U_i \right)
\end{aligned}$$

$$= 2(t - 2\lfloor t/2 \rfloor - 1) \sum_{\substack{1 \le i \le n \\ i \ne \lfloor t/2 \rfloor + 1}} U_i, \quad t \in \mathbb{R}_+,$$

where we used the relation $\lfloor s/2 \rfloor = \lfloor t/2 \rfloor$ *for* $s \in [2\lfloor t/2 \rfloor, \lfloor t/2 \rfloor + 2)$.

The operator ∇ admits an adjoint operator ∇^* given by

$$\nabla^* \left(I_n(g_{n+1}) \right) := I_{n+1}(\mathbf{1}_{\Delta_{n+1}} \tilde{g}_{n+1}),$$

where $\tilde{g}_{n+1}$ is the symmetrization of $g_{n+1} \in \widehat{L}^2(\mathbb{R}_+^n) \otimes L^2(\mathbb{R}_+)$ in $n+1$ variables. The operator ∇ is closable with domain

$$\mathrm{Dom}(\nabla) = \left\{ F \in L^2(\Omega) : \mathbb{E}[\|\nabla F\|^2_{L^2(\mathbb{R}_+)}] < \infty \right\},$$

and we have the duality relation (integration by parts)

$$\mathbb{E}[\langle \nabla F, u \rangle_{\widehat{L}^2(\mathbb{R}_+)}] = \mathbb{E}[F \nabla^*(u)], \qquad F \in \mathrm{Dom}(\nabla), \tag{2.8}$$

for u in the domain $\mathrm{Dom}(\nabla^*)$ of ∇^*, see Proposition 8 of [15]. Although the operator ∇ does not satisfy the chain rule of derivation, it can be easily applied to multiple stochastic integrals, as for any $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$ we have

$$\nabla_t I_n(f_n) = n I_{n-1}(f_n(t, \cdot)) - n \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} I_{n-1}(f_n(s, \cdot)) ds, \qquad t \in \mathbb{R}_+, \tag{2.9}$$

see Proposition 2.1 in [16]. In particular, under the condition (2.4) we have the equality

$$\nabla_t I_n(f_n) = n I_{n-1}\left(f_n(t, \cdot) \right), \qquad t \in \mathbb{R}_+,$$

see Proposition 10 of [15]. The Ornstein-Uhlenbeck operator $L := -\nabla^* \nabla$ satisfies

$$L I_n(f_n) = -\nabla^* \nabla I_n(f_n) = -n I_n(f_n), \qquad f_n \in \widehat{L}^2(\mathbb{R}_+^n),$$

where f_n satisfies (2.4). By (2.6) the operator L is well defined, invertible on centred random variables $F \in L^2(\Omega)$, and its inverse operator L^{-1} is given by

$$L^{-1} I_n(f_n) = -\frac{1}{n} I_n(f_n), \qquad n \ge 1,$$

where, due to Proposition 5.3 below, f_n does not have to satisfy (2.4). Note that $(-L)$ is a positive operator and its square root $(-L)^{-1/2}$ takes the form

$$(-L)^{1/2} I_n(f_n) = \sqrt{n} I_n(f_n), \qquad n \ge 1.$$

Stein approximation bound

The next result is a consequence of Proposition 3.3 in [16]. As above, $\mathcal{N} \sim \mathcal{N}(0,1)$ denotes a standard Gaussian random variable.

Proposition 2.4 *Let $X \in \mathrm{Dom}(\nabla)$ be such that $\mathbb{E}[X] = 0$. We have*

$$\begin{aligned} d_W(X,\mathcal{N}) &\leq \left|1 - \mathbb{E}[X^2]\right| + \sqrt{\mathrm{Var}\left[\langle \nabla X, -\nabla L^{-1}X\rangle_{\widehat{L}^2(\mathbb{R}_+)}\right]} \qquad (2.10) \\ &\quad + 2\sqrt{\mathbb{E}\left[|(-L)^{-1/2}X|^2\right]\int_0^\infty \mathbb{E}\left[|\nabla_t X|^4\right]\frac{dt}{2}} \\ &\leq |1 - \mathbb{E}[X^2]| + \sqrt{\mathrm{Var}\left[\langle \nabla X, -\nabla L^{-1}X\rangle_{L^2(\mathbb{R}_+)}\right]} \qquad (2.11) \\ &\quad + 2\sqrt{\mathbb{E}[X^2]\int_0^\infty \mathbb{E}\left[|\nabla_t X|^4\right]\frac{dt}{2}}. \end{aligned}$$

Proof. The inequality (2.11) follows from (2.10) by Proposition 5.2 with $F = L^{-1}X$, so it is enough to prove (2.10). Proposition 3.3 in [16] states that

$$\begin{aligned} d_W(X,\mathcal{N}) \leq &\,\mathbb{E}\left[\left|1 - \frac{1}{2}\langle \nabla X, -\nabla L^{-1}X\rangle\right|\right] \\ &+ \frac{1}{2}\mathbb{E}\left[\int_0^\infty |\nabla_t L^{-1}X||\nabla_t X|^2 dt\right] + \frac{1}{4}\mathbb{E}\left[\int_0^\infty |\nabla_t L^{-1}X| \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} |\nabla_s X|^2 ds dt\right]. \end{aligned}$$

We will estimate each of the three terms on the left-hand side. First, taking $F = X$ and $u = \nabla L^{-1}X$ in (2.8), we get

$$\begin{aligned} &\mathbb{E}\left[\left|1 - \frac{1}{2}\langle \nabla X, -\nabla L^{-1}X\rangle\right|\right] \\ &\leq \mathbb{E}\left[\left|1 - \frac{1}{2}\mathbb{E}\left[\langle \nabla X, -\nabla L^{-1}X\rangle\right]\right|\right] + \mathbb{E}\left[\left|\frac{1}{2}\langle \nabla X, -\nabla L^{-1}X\rangle - \frac{1}{2}\mathbb{E}\left[\langle \nabla X, -\nabla L^{-1}X\rangle\right]\right|\right] \\ &\leq \left|1 - \mathbb{E}[X^2]\right| + \sqrt{\mathrm{Var}\left[\langle \nabla X, -\nabla L^{-1}X\rangle_{\widehat{L}^2(\mathbb{R}_+)}\right]}. \end{aligned}$$

Next, for $F = L^{-1}X$ in (5.4), we obtain

$$\mathbb{E}\left[\int_0^\infty |\nabla_t L^{-1}X|^2 dt\right] = 2\mathbb{E}\left[|(-L)^{-1/2}X|^2\right].$$

Consequently, the Cauchy-Schwarz inequality gives us

$$\mathbb{E}\left[\int_0^\infty |\nabla_t L^{-1}X||\nabla_t X|^2 dt\right] \leq \sqrt{\mathbb{E}\left[\int_0^\infty |\nabla_t L^{-1}X|^2 dt\right]\mathbb{E}\left[\int_0^\infty |\nabla_t X|^4 dt\right]}$$

$$\leq 2\sqrt{\mathbb{E}\left[|(-L)^{-1/2}X|^2\right]}\sqrt{\mathbb{E}\left[\int_0^\infty |\nabla_t X|^4 \frac{dt}{2}\right]},$$

and

$$\begin{aligned}
\mathbb{E}\left[\int_0^\infty |\nabla_t L^{-1}X| \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} |\nabla_s X|^2 ds dt\right] &= \sum_{k=0}^\infty \mathbb{E}\left[\int_{2k}^{2k+2} |\nabla_t L^{-1}X| dt \int_{2k}^{2k+2} |\nabla_s X|^2 ds\right] \\
&= \sqrt{\mathbb{E}\left[\sum_{k=0}^\infty \left(\int_{2k}^{2k+2} |\nabla_t L^{-1}X| dt\right)^2\right]\mathbb{E}\left[\sum_{k=0}^\infty \left(\int_{2k}^{2k+2} |\nabla_s X|^2\right)^2 ds\right]} \\
&\leq 2\sqrt{\mathbb{E}\left[\int_0^\infty |\nabla_t L^{-1}X|^2 dt\right]\mathbb{E}\left[\sum_{k=0}^\infty \int_{2k}^{2k+2} |\nabla_s X|^4 ds\right]} \\
&\leq 4\sqrt{\mathbb{E}\left[|(-L)^{-1/2}X|^2\right]}\sqrt{\mathbb{E}\left[\int_0^\infty |\nabla_s X|^4 \frac{dt}{2}\right]},
\end{aligned}$$

and we conclude (2.10). □

Example 2.5 *Let us apply Proposition 2.4 to* $(\mathrm{Var}[I_1(f_1)])^{-1}I_1(f_1)$ *for* f_1 *as in Remark 2.1. Since* $L^{-1}I_1(f_1) = I_1(f_1)$ *for any* $f_1 \in \widehat{L}^1(\mathbb{R}_+) \cap \widehat{L}^2(\mathbb{R}_+)$ *and*

$$\nabla_t I_1(f_1) = F^{-1}_{X_{\lfloor t/2\rfloor+1}}\left(\frac{t-2\lfloor t/2\rfloor}{2}\right) - \mathbb{E}\left[X_{\lfloor t/2\rfloor+1}\right], \quad t \in \mathbb{R}_+,$$

we observe that $\nabla_t I_1(f_1)$ *is not random, and hence*

$$\mathrm{Var}\left[\langle \nabla X, -\nabla L^{-1}X\rangle_{L^2(\mathbb{R}_+)}\right] = 0$$

as well as

$$\begin{aligned}
\int_0^\infty \mathbb{E}\left[|\nabla_t X|^4\right] \frac{dt}{2} &= \sum_{k=1}^n \int_{2k-2}^{2k} \left(F^{-1}_{X_k}\left(\frac{t-2k+2}{2}\right) - \mathbb{E}\left[X_k\right]\right)^4 \frac{dt}{2} \\
&= \sum_{k=1}^n \int_0^1 \left(F^{-1}_{X_k}(u) - \mathbb{E}\left[X_k\right]\right)^4 du = \sum_{k=1}^n \mathbb{E}\left[(X_k - \mathbb{E}[X_k])^4\right].
\end{aligned}$$

Thus, we get

$$d_W\left(\frac{\sum_{k=1}^n X_k - \mathbb{E}[X_k]}{\sqrt{\sum_{k=1}^n \mathrm{Var}[X_k]}}, \mathcal{N}\right) \leq 2\frac{\sqrt{\sum_{k=1}^n \mathbb{E}\left[(X_k - \mathbb{E}[X_k])^4\right]}}{\sum_{k=1}^n \mathrm{Var}[X_k]},$$

which provides a quantitative bound with explicit constant in the Wasserstein distance for the L^4 *Lyapunov Central Limit Theorem.*

3 Normal approximation for generalized U-statistics

In this section we consider generalized U-statistics of order $n \geq 1$ of the form

$$\sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}_0\\ k_i\neq k_j \text{ if } i\neq j}} f_n(2k_1+1+U_{k_1},\ldots,2k_n+1+U_{k_n}),$$

where $\mathbb{N}_0 = \{0,1,2,\ldots\}$. The next proposition gives the multiple stochastic integral expansion of such extended generalized U-statistics.

Proposition 3.1 *Given $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$ we have*

$$\sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}_0\\ k_i\neq k_j \text{ if } i\neq j}} f_n(2k_1+1+U_{k_1},\ldots,2k_n+1+U_{k_n}) = \sum_{r=0}^{n} I_r\left(f_n^{(r)}\right),$$

where

$$f_n^{(r)} = (x_1,\ldots,x_k) = \frac{1}{2^{n-r}}\binom{n}{r}\int_{\mathbb{R}_+^{n-r}} f_n(x_1,\ldots,x_r,y_1,\ldots,y_{n-r})dy_1\cdots dy_{n-r},$$

$r = 0,1,\ldots,n$.

Proof. Formula (2.2) gives us for $1 \leq m \leq n$

$$a_m := (-2)^m I_m\left(\int_{\mathbb{R}^{n-m}} f_n\left(\cdot,y_1,\ldots,y_{n-m}\right)dy_1\cdots dy_{n-m}\right) = \sum_{r=0}^{m}(-1)^r\binom{m}{r}b_r,$$

where

$$b_r = 2^r \sum_{\substack{k_1,\ldots,k_r\in\mathbb{N}_0\\ k_i\neq k_j \text{ if } i\neq j}} \int_0^\infty\cdots\int_0^\infty f_n(2k_1+1+U_{k_1},\ldots,2k_r+1+U_{k_r},y_1,\ldots,y_{n-r})dy_1\cdots dy_{n-r}.$$

Hence, by binomial inversion, we have $b_m = \sum_{r=1}^m (-1)^r\binom{m}{r}a_r$, $1 \leq m \leq n$. In particular,

$$\begin{aligned}
&\sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}_0\\ k_i\neq k_j \text{ if } i\neq j}} f_n(2k_1+1+U_{k_1},\ldots,2k_n+1+U_{k_n})\\
&= 2^{-n}b_n = 2^{-n}\sum_{r=1}^{n}(-1)^r\binom{n}{r}a_r\\
&= \sum_{r=0}^{n}\binom{n}{r}\frac{1}{2^{n-r}}I_r\left(\int_{\mathbb{R}^{n-r}} f_n\left(\cdot,y_1,\ldots,y_{n-r}\right)dy_1\cdots dy_{n-r}\right)\\
&= \sum_{r=0}^{n} I_r\left(f_n^{(r)}\right),
\end{aligned}$$

as required. $\square$

In particular, under the condition (2.4) the multiple stochastic integral $I_n(f_n)$ coincides with the generalized degenerate U-statistic of order n, and we have

$$I_n(f_n) = \sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}_0 \\ k_i\neq k_j \text{ if } i\neq j}} f_n(2k_1+1+U_1,\ldots,2k_n+1+U_n). \tag{3.1}$$

In the next corollary we obtain a Wasserstein distance bound for sums of multiple stochastic integrals by combining Propositions 5.1 and 2.4 with the multiplication formula (5.1). First, let us introduce the following $\star$-notation: for $0 \le l \le k \le n \wedge m$ we define the contraction $f_n \star_k^l g_m$ of $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$ and $g_m \in \widehat{L}^2(\mathbb{R}_+^m)$ as

$$\begin{aligned} & f_n \star_k^l g_m(x_1,\ldots,x_{k-l},y_1,\ldots,y_{n-k},z_1,\ldots,z_{m-k}) \\ & := \frac{1}{2^l}\int_{\mathbb{R}_+^l} f_n(w_1,\ldots,w_l,x_1,\ldots,x_{k-l},y_1,\ldots,y_{n-k}) \\ & \qquad\qquad \times g_m(w_1,\ldots,w_l,x_1,\ldots,x_{k-l},z_1,\ldots,z_{m-k})dw_1\cdots dw_l, \end{aligned} \tag{3.2}$$

and we let $f_n \,\widetilde{\star}_k^l g_m$ denote the symmetrization

$$\begin{aligned} & f_n \,\widetilde{\star}_k^l\, g_m(x_1,\ldots,x_{n+m-k-l}) \\ & \quad := \frac{\mathbf{1}_{\Delta_{m+n-k-l}}(x_1,\ldots,x_{n+m-k-l})}{(m+n-k-l)!}\sum_{\sigma\in S_{m+n-k-l}} f_n \star_k^l g_m(x_{\sigma(1)},\ldots,x_{\sigma(m+n-k-l)}), \end{aligned}$$

where S_n, $n \ge 1$, denotes the set of all permutations of the set $\{1,\ldots,n\}$.

Theorem 3.2 *For any $X \in L^2(\Omega)$ written as a sum $X = \sum_{k=1}^n I_k(f_k)$ of multiple stochastic integrals where $f_k \in \widehat{L}^2(\mathbb{R}_+^k)$ satisfies (2.4), $k = 1,\ldots,n$, we have*

$$\begin{aligned} d_W(X,\mathcal{N}) \le & \left|1-\mathbb{E}[X^2]\right| \\ & + C_n\sqrt{\sum_{0\le l<i\le n}\left\|f_i \star_i^l f_i\right\|^2_{L^2(\mathbb{R}_+^{i-l})} + \sum_{1\le l<i\le n}\left(\left\|f_i \star_l^l f_i\right\|^2_{L^2(\mathbb{R}_+^{2(i-l)})} + \left\|f_l \star_l^l f_i\right\|^2_{L^2(\mathbb{R}_+^{i-l})}\right)}, \end{aligned}$$

for some $C_n > 0$.

Proof. Given that

$$\nabla_t X = \sum_{k=0}^{n-1}(k+1)I_k\left(f_{k+1}(t,\cdot)\right), \quad \text{and} \quad \nabla_t L^{-1}X = \sum_{k=0}^{n-1} I_k\left(f_{k+1}(t,\cdot)\right),$$

the multiplication formula (5.1) shows that

$$(\nabla_t X)^2 = \sum_{0\le i\le j<n}\sum_{k=0}^{i}\sum_{l=0}^{k} c_{i,j,k,l} I_{i+j-k-l}\left(f_{i+1}(t,\cdot)\,\widetilde{\star}_k^l f_{j+1}(t,\cdot)\right) \tag{3.3}$$

and

$$\nabla_t X \nabla_t L^{-1} X = \sum_{0 \le i \le j < n} \sum_{k=0}^{i} \sum_{l=0}^{k} d_{i,j,k,l} I_{i+j-k-l}\big(f_{i+1}(t,\cdot) \,\widetilde{\star}_k^l f_{j+1}(t,\cdot)\big), \tag{3.4}$$

for some $c_{i,j,k,l}, d_{i,j,k,l} \ge 0$. Next, by (2.3) and (3.3) we get

$$\begin{aligned} \int_0^\infty \mathbb{E}\left[|\nabla_t X|^4\right] \frac{dt}{2} &\le C \sum_{0 \le i \le j < n} \sum_{k=0}^{i} \sum_{l=0}^{k} \int_0^\infty \left\| f_{i+1}(t,\cdot) \widetilde{\star}_k^l f_{j+1}(t,\cdot) \right\|^2_{\widehat{L}^2(\mathbb{R}_+^{i+j-k-l})} dt \\ &\le C \sum_{0 \le i \le j < n} \sum_{k=0}^{i} \sum_{l=0}^{k} \left\| f_{i+1} \star_{k+1}^l f_{j+1} \right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l+1})} \\ &\le C \sum_{1 \le i \le j \le n} \sum_{k=1}^{i} \sum_{l=0}^{k-1} \left\| f_i \star_k^l f_j \right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})}, \end{aligned} \tag{3.5}$$

where $C > 0$ is a constant depending on n. Furthermore, from (3.4) it follows that

$$\begin{aligned} &\langle \nabla X, -\nabla L^{-1} X \rangle - \mathbb{E}\left[\langle \nabla X, -\nabla L^{-1} X \rangle\right] \\ &= \frac{1}{2} \int_0^\infty \sum_{0 \le i \le j < n} \sum_{k=0}^{i} \sum_{l=0}^{k} d_{i,j,l,k} \mathbf{1}_{\{i=j=k=l\}^c} I_{i+j-k-l}\big(f_{i+1}(t,\cdot) \widetilde{\star}_k^l f_{j+1}(t,\cdot)\big) dt, \end{aligned}$$

thus we get

$$\begin{aligned} &\operatorname{Var}\left[\langle \nabla X, -\nabla L^{-1} X \rangle\right] \\ &\le C' \sum_{0 \le i \le j < n} \sum_{k=0}^{i} \sum_{l=0}^{k} \mathbf{1}_{\{i=j=k=l\}^c} \left\| \int_0^\infty f_{i+1}(t,\cdot) \star_k^l f_{j+1}(t,\cdot) dt \right\|^2_{L^2(\mathbb{R}_+^{(i+j-k-l)})} \\ &= C'' \sum_{0 \le i \le j < n} \sum_{k=0}^{i} \sum_{l=0}^{k} \mathbf{1}_{\{i=j=k=l\}^c} \left\| f_{i+1} \star_{k+1}^{l+1} f_{j+1} \right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})} \\ &= C'' \sum_{1 \le i \le j \le n} \sum_{k=1}^{i} \sum_{l=1}^{k} \mathbf{1}_{\{i=j=k=l\}^c} \left\| f_i \star_k^l f_j \right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})}, \end{aligned} \tag{3.6}$$

for some constants $C', C'' > 0$ depending only n. Applying (3.5) and (3.6) to (2.11), we get

$$d_W(X, \mathcal{N}) \le \left|1 - \mathbb{E}[X^2]\right| + C''' \sqrt{\sum_{1 \le i \le j \le n} \sum_{k=1}^{i} \sum_{l=0}^{k} \mathbf{1}_{\{i=j=k=l\}^c} \left\| f_i \star_k^l f_j \right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})}},$$

for some $C''' > 0$ depending on n. Next, by the inequality (5.2), all the components where $0 \le l < k \le i, j$, are dominated by those where $0 \le l < k = i = j$, and also, by the inequality (5.3), the ones where $1 \le k = l < \min\{i, j-1\}$, are dominated by the components where $1 \le l = k < i = j$. Finally, the components for $1 \le k = l = i < j$ remain unchanged.

□

4 Application to weighted random graphs

In this section we present an application of results from the previous section to the Erdős-Rényi random graph $\mathbb{G}(n,p)$ and to the renormalization $\widetilde{W}_n^G$ of the combined weight W_n^G of subgraphs of the random graph that are isomorphic to a fixed graph G, see (1.3).

In order to simplify the notation we write $a_n \lesssim b_n$ for two sequences a_n and b_n whenever there exist a constant C depending only on G such that $a_n < Cb_n$ for all $n \in \mathbb{N}$. Furthermore, if $a_n \lesssim b_n$ and $b_n \lesssim a_n$ then we write $a_n \approx b_n$. Finally, by writing $H \sim K$ we mean that the two graphs H and K are isomorphic. In Proposition 4.1 we provide estimates of the variance of W_n^G, which is crucial when dealing with the renormalization.

Proposition 4.1 *The variance of W_n^G admits the asymptotic form*

$$\mathrm{Var}\big[W_n^G\big] \approx \big(\mathrm{Var}[X] + (1-p_n)(\mathbb{E}[X])^2\big) \max_{\substack{H\subset G\\ e_H\geq 1}} n^{2v_G - v_H} p_n^{2e_G - e_H}. \tag{4.1}$$

Proof. We follow the lines of the proof of Lemma 3.5 in [7] by extending the argument to nonnegative random weights distributed as X. We note that

$$W_n^G = \sum_{G'\sim G} S_{G'},$$

where the sum is over all graphs $G' \subset K_n$ which are isomorphic to G, and $S_{G'}$ is the sum of the weights of edges in G' if G' belongs to $\mathbb{G}(n,p_n)$, and zero otherwise, i.e. denoting by $X_1,\ldots,X_{e_G}$ the random weights of edges of G', we have

$$S_{G'} := \mathbf{1}_{\{G'\in\mathbb{G}(n,p_n)\}} \sum_{i=1}^{e_G} X_i.$$

Then, we get

$$\begin{aligned}
\mathrm{Var}\big[W_n^G\big] &= \sum_{G',G''\sim G} \mathrm{Cov}(S_{G'}, S_{G''}) \\
&= \sum_{\substack{G',G''\sim G\\ \text{with a common edge}}} \big(\mathbb{E}[S_{G'}S_{G''}] - \mathbb{E}[S_{G'}]\mathbb{E}[S_{G''}]\big) \\
&\approx \sum_{\substack{H\subset G\\ e_H\geq 1}} \sum_{\substack{G'\cap G''\sim H\\ G',G''\sim G}} \big(\mathbb{E}[S_{G'}S_{G''}] - \mathbb{E}[S_{G'}]\mathbb{E}[S_{G''}]\big).
\end{aligned}$$

For a fixed $G' \sim G$ we clearly have

$$\mathbb{E}[S_{G'}] = \mathbb{P}(G' \in \mathbb{G}(n,p_n)) \sum_{i=1}^{e_G} \mathbb{E}[X_i] = e_G p^{e_G} \mathbb{E}[X].$$

In order to calculate $\mathbb{E}[S_{G'}S_{G''}]$ for $G', G'' \sim G$ and $G' \cap G'' \sim H$, let us denote by $X_1, \ldots, X_{e_H}$ the weights of edges of $G' \cap G''$ and by $X'_1, \ldots, X'_{e_G - e_H}$ and $X''_1, \ldots, X''_{e_G - e_H}$ weights of edges of $G' \backslash G''$ and $G'' \backslash G'$, respectively. Then, we have

$$
\begin{aligned}
&\mathbb{E}[S_{G'}S_{G''}] = \mathbb{P}\left(G', G'' \in \mathbb{G}(n, p_n)\right) \mathbb{E}\left[\left(\sum_{i=1}^{e_H} X_i + \sum_{i=1}^{e_G - e_H} X'_i\right)\left(\sum_{i=1}^{e_H} X_i + \sum_{i=1}^{e_G - e_H} X''_i\right)\right] \\
&= \mathbb{P}\left(G' \cap G'' \in \mathbb{G}(n, p_n)\right)\left(\mathbb{E}\left[\left(\sum_{i=1}^{e_H} X_i\right)^2\right] + \left(2e_H(e_G - e_H) + (e_G - e_H)^2\right)(\mathbb{E}[X])^2\right) \\
&= p_n^{2e_G - e_H}\left(e_H \mathbb{E}\left[X^2\right] + \left(e_G^2 - e_H\right)\mathbb{E}[X])^2\right) \\
&= p_n^{2e_G - e_H}(e_H \mathrm{Var}[X] + e_G^2(\mathbb{E}[X])^2).
\end{aligned}
$$

Hence we get

$$
\begin{aligned}
\mathbb{E}[S_{G'}S_{G''}] - \mathbb{E}[S_{G'}]\mathbb{E}[S_{G''}] &= p_n^{2e_G - e_H}(e_H \mathrm{Var}[X] + e_G^2(\mathbb{E}[X])^2) - p_n^{2e_G} e_G^2(\mathbb{E}[X])^2 \\
&= p_n^{2e_G - e_H}(e_H \mathrm{Var}[X] + e_G^2(1 - p_n^{e_H})(\mathbb{E}[X])^2) \\
&\approx p_n^{2e_G - e_H}(\mathrm{Var}[X] + (1 - p_n)(\mathbb{E}[X])^2),
\end{aligned}
$$

and consequently

$$
\begin{aligned}
\mathrm{Var}\left[W_n^G\right] &\approx \sum_{\substack{H \subseteq G \\ e_H \geq 1}} \sum_{\substack{G' \cap G'' \sim H \\ G', G'' \sim G}} p_n^{2e_G - e_H}(\mathrm{Var}[X] + (1 - p_n)(\mathbb{E}[X])^2) \\
&\approx (\mathrm{Var}[X] + (1 - p_n)(\mathbb{E}[X])^2) \sum_{\substack{H \subseteq G \\ e_H \geq 1}} n^{2v_G - v_H} p_n^{2e_G - e_H} \\
&\approx \left(\mathrm{Var}[X] + (1 - p_n)(\mathbb{E}[X])^2\right) \max_{\substack{H \subseteq G \\ e_H \geq 1}} n^{2v_G - v_H} p_n^{2e_G - e_H},
\end{aligned}
$$

as required. □

Next, we show in Lemma 4.2 that the combined weights W_n^G of subgraphs can be written as a sum of multiple stochastic integrals using Proposition 3.1. This allows us to apply Theorem 3.2 to obtain normal approximation in Wasserstein distance for $\widetilde{W}_n^G$, which is presented in Theorem 4.3. In the sequel we number, in a fixed but arbitrary way, all possible edges of the complete graph K_n from 1 to $n(n-1)/2$, and we denote by $E_G \subset \mathbb{N}^{e_G}$ the set of sequences of edges that create a graph isomorphic to G, i.e. a sequence $(e_{k_1}, \ldots, e_{k_{e_G}})$ belongs to E_G if and only if the graph created by edges $e_{k_1}, \ldots, e_{k_{e_G}}$ is isomorphic to G. Before stating the lemma, let us define the operator Ψ_{t_i}

$$
\Psi_{t_i} f(t_1, \ldots, t_n) := f(t_1, \ldots, t_n) - \frac{1}{2}\int_{2\lfloor t_i/2 \rfloor}^{2\lfloor t_i/2 \rfloor + 2} f(t_1, \ldots, t_{i-1}, s, t_{i+1}, \ldots, t_n) ds, \tag{4.2}
$$

which arises naturally when representing any multiple stochastic integral $I_n(f_n)$ as $I_n(\bar{f}_n)$ with $\bar{f}_n$ satisfying (2.4), see Proposition 5.3.

Lemma 4.2 *We have the identity in distribution*

$$W_n^G \stackrel{d}{=} \sum_{r=0}^{e_G} I_k\left(\bar{h}_k\right), \tag{4.3}$$

where

$$\bar{h}_k(t_1,\ldots,t_k) \tag{4.4}$$
$$:= \Psi_{t_1}\cdots\Psi_{t_k} g_k\left(t_1-2\lfloor t_1/2\rfloor,\ldots,t_k-2\lfloor t_k/2\rfloor\right)\sum_{a\in\mathbb{N}^{e_G-k}}\mathbf{1}_{E_G}\left(a,\lfloor t_1/2\rfloor,\ldots,\lfloor t_k/2\rfloor\right),$$

and the function $g_k:(0,2)^k\to\mathbb{R}$ is given by

$$g_k(t_1,\ldots,t_k)=\frac{(p_n/2)^{e_G-k}}{(e_G-k)!k!}\mathbf{1}_{(0,2p_n)^k}\left(t_1,\ldots,t_k\right)\left((e_G-k)\mathbb{E}[X]+\sum_{i=1}^k F_X^{-1}\left(\frac{t_i}{2p_n}\right)\right), \tag{4.5}$$

where F_X^{-1} is the generalized inverse of the distribution function F_X of X.

Proof. First, we note that

$$\begin{aligned}
W_n^G &\stackrel{d}{=} \frac{1}{e_G!}\sum_{k_1\neq\cdots\neq k_{e_G}\geq 0}\mathbf{1}_{E_G}(k_1,\ldots,k_{e_G})\mathbf{1}_{(0,2p_n)^{e_G}}\left(U_{k_1}+1,\ldots,U_{k_{e_G}}+1\right)\\
&\qquad\times\left(F^{-1}\left(\frac{U_{k_1}+1}{2p_n}\right)+\cdots+F^{-1}\left(\frac{U_{k_{e_G}}+1}{2p_n}\right)\right)\\
&=\frac{1}{e_G!}\sum_{k_1\neq\cdots\neq k_{e_G}\geq 0}h_{e_G}(2k_1+1+U_{k_1},\ldots,2k_{e_G}+1+U_{k_{e_G}}),
\end{aligned}$$

where

$$\begin{aligned}
h_{e_G}(t_1,\ldots,t_{e_G}) &= \mathbf{1}_{E_G}\left(\lfloor t_1/2\rfloor,\ldots,\lfloor t_{e_G}/2\rfloor\right)\mathbf{1}_{(0,2p_n)^{e_G}}\left(t_1-2\lfloor t_1/2\rfloor,\ldots,t_{e_G}-2\lfloor t_{e_G}/2\rfloor\right)\\
&\quad\times\left(F_X^{-1}\left(\frac{t_1-2\lfloor t_1/2\rfloor}{2p_n}\right)+\cdots+F_X^{-1}\left(\frac{t_{e_G}-2\lfloor t_{e_G}/2\rfloor}{2p_n}\right)\right),
\end{aligned}$$

and by Proposition 3.1, the relation (4.3) holds with

$$h_k(t_1,\ldots,t_k):=g_k\left(t_1-2\lfloor t_1/2\rfloor,\ldots,t_k-2\lfloor t_k/2\rfloor\right)\sum_{a\in\mathbb{N}^{e_G-k}}\mathbf{1}_{E_G}\left(a,\lfloor t_1/2\rfloor,\ldots,\lfloor t_k/2\rfloor\right),$$

where $g_k:(0,2)^k\to\mathbb{R}$ is given by (4.5). Finally, in case the functions h_k may not satisfy the condition (2.4), we can use Proposition 5.3 to obtain (4.3) with

$$\bar{h}_k(t_1,\ldots,t_k)$$

$$= \Psi_{t_1} \cdots \Psi_{t_k} \left(g_k\left(t_1 - 2\lfloor t_1/2 \rfloor, \ldots, t_k - 2\lfloor t_k/2 \rfloor\right) \sum_{a \in \mathbb{N}^{e_G - k}} \mathbf{1}_{E_G}\left(a, \lfloor t_1/2 \rfloor, \ldots, \lfloor t_k/2 \rfloor\right) \right)$$
$$= \Psi_{t_1} \cdots \Psi_{t_k} g_k\left(t_1 - 2\lfloor t_1/2 \rfloor, \ldots, t_k - 2\lfloor t_k/2 \rfloor\right) \sum_{a \in \mathbb{N}^{e_G - k}} \mathbf{1}_{E_G}\left(a, \lfloor t_1/2 \rfloor, \ldots, \lfloor t_k/2 \rfloor\right),$$

where the last equality follows from the fact that the sum appearing above is constant for $(t_1, \ldots, t_k) \in (2m_1, 2m_1+2) \times \ldots \times (2m_k, 2m_k+2)$, $m_1, \ldots, m_k \in \mathbb{N}$. The proof is complete.

□

We can now pass to the main result in this section.

Theorem 4.3 *Let G be a graph without isolated vertices. The renormalized weight $\widetilde{W}_n^G$ of graphs in $\mathbb{G}(n, p_n)$ that are isomorphic to G satisfies*

$$d_W\left(\widetilde{W}_n^G, \mathcal{N}\right) \lesssim \frac{\sqrt{\mathbb{E}[(X - \mathbb{E}[X])^4]} + (1 - p_n)(\mathbb{E}[X])^2}{\operatorname{Var}[X] + (1 - p_n)(\mathbb{E}[X])^2} \left((1 - p_n) \min_{\substack{H \subseteq G \\ e_H \geq 1}} n^{v_H} p_n^{e_H} \right)^{-1/2}. \quad (4.6)$$

Proof. Without loss of generality we take $p_n = p$ in the proof. By Corollary 3.2 we have

$$d_W\left(\widetilde{W}_n^G, \mathcal{N}\right) \lesssim \frac{1}{\operatorname{Var}\left[W_n^G\right]} \left(\sum_{0 \leq l < k \leq e_G} \left\| \bar{h}_k \star_k^l \bar{h}_k \right\|^2_{L^2(\mathbb{R}_+^{k-l})} + \sum_{1 \leq l < k \leq e_G} \left\| \bar{h}_l \star_l^l \bar{h}_k \right\|^2_{L^2(\mathbb{R}_+^{k-l})} \right.$$
$$\left. + \sum_{1 \leq l < k \leq e_G} \left\| \bar{h}_k \star_l^l \bar{h}_k \right\|^2_{L^2(\mathbb{R}_+^{2(k-l)})} \right)^{1/2}$$
$$=: \frac{\sqrt{S_1 + S_2 + S_3}}{\operatorname{Var}\left[W_n^G\right]}, \quad (4.7)$$

where $\bar{h}_k$ has been defined in (4.4). We note that by the equivalence (4.1) of Proposition 4.1 it suffices to show that

$$S_1 + S_2 + S_3 \lesssim \frac{\mathbb{E}\left[(X - \mathbb{E}\left[X\right])^4\right] + (1 - p)^2 (\mathbb{E}\left[X\right])^4}{1 - p} \max_{\substack{H \subseteq G \\ e_H \geq 1}} n^{4v_G - 3v_H} p^{4e_G - 3e_H}, \quad (4.8)$$

which follows from (4.9), (4.10) and (4.11) below. Indeed, applying (4.1) and (4.8) to (4.7) shows that

$$d_W\left(\widetilde{W}_n^G, \mathcal{N}\right) \lesssim \frac{\sqrt{\mathbb{E}\left[(X^2 - \mathbb{E}\left[X\right])^4\right] + (1 - p)^2 (\mathbb{E}\left[X\right])^4 \max_{\substack{H \subseteq G \\ e_H \geq 1}} n^{4v_G - 3v_H} p^{4e_G - 3e_H}}}{\sqrt{1 - p} \left(\mathbb{E}[X^2] - p(\mathbb{E}[X])^2\right) \max_{\substack{H \subseteq G \\ e_H \geq 1}} n^{2v_G - v_H} p^{2e_G - e_H}},$$

and after factoring out $n^{4v_G} p^{4e_G}$ in front of the maxima, we conclude that

$$d_W\left(\widetilde{W}_n^G, \mathcal{N}\right) \lesssim \frac{\left(\sqrt{\mathbb{E}\left[(X^2 - \mathbb{E}\left[X\right])^4\right]} + (1 - p)(\mathbb{E}\left[X\right])^2\right) \left(\max_{\substack{H \subseteq G \\ e_H \geq 1}} n^{-v_H} p^{-e_H}\right)^{3/2}}{\sqrt{1 - p} \left(\mathbb{E}[X^2] - p(\mathbb{E}[X])^2\right) \max_{\substack{H \subseteq G \\ e_H \geq 1}} n^{-v_H} p^{-e_H}}$$

$$= \frac{\sqrt{\mathbb{E}[(X-\mathbb{E}[X])^4]+(1-p)(\mathbb{E}[X])^2}}{\mathbb{E}[X^2]-p(\mathbb{E}[X])^2}\left((1-p)\min_{\substack{H\subset G\\ e_H\geq 1}} n^{v_H}p^{e_H}\right)^{-1/2}.$$

$i)$ Estimation of S_1. For $0\leq l<k\leq n$ we have

$$\left\|\bar{h}_k \star_k^l \bar{h}_k\right\|^2_{L^2(\mathbb{R}_+^{k-l})} = \frac{1}{2^{2l}}\int_{\mathbb{R}_+^{k-l}}\left(\int_{\mathbb{R}_+^l}\left(\bar{h}_k(x_1,\ldots,x_k)\right)^2 dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k$$

$$= \frac{1}{2^{2l}}\int_{\mathbb{R}_+^{k-l}}\left(\sum_{b\in\mathbb{N}^l}\left(\sum_{a\in\mathbb{N}^{e_G-k}}\mathbf{1}_{E_G}\left(a,b,\lfloor x_{l+1}/2\rfloor,\ldots,\lfloor x_k/2\rfloor\right)\right)^2\right.$$
$$\left.\int_{(0,2)^l}\Big(\Psi_{x_1}\cdots\Psi_{x_k}g_k\left(x_1,\ldots,x_l,x_{l+1}-2\lfloor x_{l+1}/2\rfloor,\ldots,x_k-2\lfloor x_k/2\rfloor\right)\Big)^2 dx_1\cdots dx_l\right)^2$$
$$dx_{l+1}\cdots dx_k$$

$$= \frac{1}{2^{2l}}\sum_{c\in\mathbb{N}^{k-l}}\left(\sum_{b\in\mathbb{N}^l}\left(\sum_{a\in\mathbb{N}^{e_G-k}}\mathbf{1}_{E_G}\left(a,b,c\right)\right)^2\right)^2$$
$$\times\int_{(0,2)^{k-l}}\left(\int_{(0,2)^l}\Big(\Psi_{x_1}\cdots\Psi_{x_k}g_k\left(x_1,\ldots,x_k\right)\Big)^2 dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k.$$

Combining the equivalence

$$\sum_{c\in\mathbb{N}^{k-l}}\left(\sum_{b\in\mathbb{N}^l}\left(\sum_{a\in\mathbb{N}^{e_G-k}}\mathbf{1}_{E_G}\left(a,b,c\right)\right)^2\right)^2 \approx \max_{\substack{K\subset H\subset G\\ e_K=k-l,\ e_H=k}} n^{4v_G-2v_H-v_K},$$

see the proof of Theorem 4.2 in [17], with (5.7) in Lemma 5.4, we get

$$\left\|\bar{h}_k \star_k^l \bar{h}_k\right\|^2_{L^2(\mathbb{R}_+^{k-l})} \lesssim \left(\mathbb{E}\left[(X-\mathbb{E}[X])^4\right]+(1-p)^2(\mathbb{E}[X])^4\right)$$
$$\times\max_{\substack{K\subset H\subset G\\ e_K=k-l,\ e_H=k}} n^{4v_G-2v_H-v_K}p^{4e_G-2e_H-e_K}(1-p)^{2e_H-e_K-2},$$

and consequently

$$\begin{aligned}
S_1 &= \sum_{0\leq l<k\leq n}\left\|\bar{h}_k \star_k^l \bar{h}_k\right\|^2_{L^2(\mathbb{R}_+^{k-l})}\\
&\lesssim \left(\mathbb{E}\left[(X-\mathbb{E}[X])^4\right]+(1-p)^2(\mathbb{E}[X])^4\right)\max_{\substack{K\subset H\subset G\\ e_K\geq 1}} n^{4v_G-2v_H-v_K}p^{4e_G-2e_H-e_K}(1-p)^{2e_H-e_K-2}\\
&\lesssim \frac{\mathbb{E}\left[(X-\mathbb{E}[X])^4\right]+(1-p)^2(\mathbb{E}[X])^4}{1-p}\max_{\substack{H\subset G\\ e_H\geq 1}} n^{4v_G-3v_H}p^{4e_G-3e_H},
\end{aligned}\tag{4.9}$$

as in the proof of Theorem 4.2 in [17].

$ii)$ Estimation of S_2. Similarly, for $1\leq l<k\leq n$ we have

$$\left\|\bar{h}_l \star_l^l \bar{h}_k\right\|^2_{L^2(\mathbb{R}_+^{k-l})} = \frac{1}{2^{2l}}\int_{\mathbb{R}_+^{k-l}}\left(\int_{\mathbb{R}_+^l}\bar{h}_l(x_1,\ldots,x_l)\bar{h}_k(x_1,\ldots,x_k)dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k$$

$$
\begin{aligned}
&= \frac{1}{2^{2l}} \int_{\mathbb{R}_+^{k-l}} \left(\sum_{b\in\mathbb{N}^l} \left(\sum_{a\in\mathbb{N}^{e_G-l}} \mathbf{1}_{E_G}(a,b) \sum_{a'\in\mathbb{N}^{e_G-k}} \mathbf{1}_{E_G}(a',b,\lfloor x_{l+1}/2\rfloor,\dots,\lfloor x_k/2\rfloor) \right) \right.\\
&\quad \times \int_{(0,2)^l} \Psi_{x_1}\cdots\Psi_{x_l} g_l(x_1,\dots,x_l)\Psi_{x_1}\cdots\Psi_{x_k}\\
&\qquad\qquad \left. \times\, g_k(x,x_{l+1}-2\lfloor x_{l+1}/2\rfloor,\dots,x_k-2\lfloor x_k/2\rfloor)\, dx_1\cdots dx_l \right)^2 \quad dx_{l+1}\cdots dx_k\\
&= \frac{1}{2^{2l}} \sum_{c\in\mathbb{N}^{k-l}} \left(\sum_{b'\in\mathbb{N}^l} \left(\sum_{a\in\mathbb{N}^{e_G-l}} \mathbf{1}_{E_G}(a,b) \sum_{a'\in\mathbb{N}^{e_G-k}} \mathbf{1}_{E_G}(a',b,c) \right) \right)^2\\
&\quad \times \int_{(0,2)^{k-l}} \left(\int_{(0,2)^l} \Psi_{x_1}\cdots\Psi_{x_l} g_l(x_1,\dots,x_l)\Psi_{x_1}\cdots\Psi_{x_k} g_k(x_1,\dots,x_k) dx_1\cdots dx_l \right)^2\\
&\qquad\qquad dx_{l+1}\cdots dx_k.
\end{aligned}
$$

By the Cauchy-Schwarz inequality and the formula (5.6) in Lemma 5.4, we get

$$
\begin{aligned}
&\int_{(0,2)^{k-l}} \left(\int_{(0,2)^l} \Psi_{x_1}\cdots\Psi_{x_l} g_l(x_1,\dots,x_l)\Psi_{x_1}\cdots\Psi_{x_k} g_k(x_1,\dots,x_k)\, dx_1\cdots dx_l \right)^2 dx_{l+1}\cdots dx_k\\
&\le \int_{(0,2)^l} \Psi_{x_1}\cdots\Psi_{x_l} g_l^2(x_1\cdots x_l) dx_1\cdots dx_l \int_{(0,2)^k} \Psi_{x_1}\cdots\Psi_{x_k} g_k^2(x_{l+1}\cdots x_k)\, dx_{l+1}\cdots dx_k\\
&\lesssim p^{4e_G-k-l}(1-p)^{k+l-2}\big(\mathbb{E}\big[(X^2-\mathbb{E}[X])^2\big]+(1-p)(\mathbb{E}[X])^2\big)^2\\
&\lesssim \frac{p^{4e_G-k-l}}{1-p}\big(\mathbb{E}\big[(X-\mathbb{E}[X])^4\big]+(1-p)^2(\mathbb{E}[X])^4\big).
\end{aligned}
$$

Furthermore, we have

$$
\sum_{c\in\mathbb{N}^{k-l}} \left(\sum_{a'\in\mathbb{N}^l} \left(\sum_{a'\in\mathbb{N}^{e_G-l}} \mathbf{1}_{E_G}(a,b) \sum_{a'\in\mathbb{N}^{e_G-k}} \mathbf{1}_{E_G}(a',b,c) \right) \right)^2 \lesssim \max_{\substack{K\subset H'\subset G\\ e_K=k-l,\ e_{H'}=l}} n^{4v_G-2v_{H'}-v_K},
$$

see the proof of Theorem 4.2 in [17], thus

$$
\begin{aligned}
\big\|\bar{h}_l \star_l^l \bar{h}_k\big\|^2_{L^2(\mathbb{R}_+^{k-l})} &\lesssim (1-p)^{-1}\Big(\mathbb{E}\left[(X-\mathbb{E}[X])^4\right]+(1-p)^2(\mathbb{E}[X])^4\Big)\\
&\qquad \times \max_{\substack{K\subset H'\subset G\\ e_K=k-l,\ e_{H'}=k}} n^{4v_G-2v_{H'}-v_K} p^{4e_G-2e_{H'}-e_K},
\end{aligned}
$$

from which it follows by that

$$
S_2 = \sum_{1\le l<k\le n} \big\|\bar{h}_l \star_l^l \bar{h}_k\big\|^2_{L^2(\mathbb{R}_+^{k-l})} \lesssim \frac{E\big[(X-\mathbb{E}[X])^4\big]+(1-p)^2(\mathbb{E}[X])^4}{1-p} \max_{\substack{H\subset G\\ e_H\ge 1}} n^{4v_G-3v_H} p^{4e_G-3e_H}, \tag{4.10}
$$

as in the proof of Theorem 4.2 in [17].

iii) Estimation of S_3. For $1 \le l < k \le n$ we have

$$\begin{aligned}
&\left\|\bar{h}_k \star_l^l \bar{h}_k\right\|^2_{L^2(\mathbb{R}_+^{2(k-l)})} \\
&= \frac{1}{2^{2l}} \int_{\mathbb{R}_+^{k-l}} \int_{\mathbb{R}_+^{k-l}} \left(\int_{\mathbb{R}_+^l} \bar{h}_k(x_1, \ldots, x_k) \bar{h}_k(x_1, \ldots, x_l, z_1, \ldots, z_{k-l}) dx_1 \cdots dx_l \right)^2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad dx_{l+1} \cdots dx_k dz_1 \cdots dz_{k-l} \\
&= \frac{1}{2^{2l}} \int_{\mathbb{R}_+^{k-l}} \int_{\mathbb{R}_+^{k-l}} \Bigg(\sum_{b \in \mathbb{N}^l} \sum_{a, a' \in \mathbb{N}^{e_G - k}} \mathbf{1}_{E_G}\left(a, b, \lfloor \tfrac{x_{l+1}}{2} \rfloor, \ldots, \lfloor \tfrac{x_k}{2} \rfloor\right) \mathbf{1}_{E_G}\left(a', b, \lfloor \tfrac{z_1}{2} \rfloor, \ldots, \lfloor \tfrac{z_{k-l}}{2} \rfloor\right) \\
&\qquad \int_{(0,2)^l} \Psi_{x_1} \cdots \Psi_{x_k} g_k\left(x, x_{l+1} - 2\lfloor \tfrac{x_{l+1}}{2} \rfloor, \ldots, x_k - 2\lfloor \tfrac{x_k}{2} \rfloor\right) \\
&\qquad \Psi_{x_1} \cdots \Psi_{x_l} \Psi_{z_1} \cdots \Psi_{z_{k-l}} g_k\left(x, x_{l+1} - 2\lfloor \tfrac{x_{l+1}}{2} \rfloor, \ldots, z_{k-l} - 2\lfloor \tfrac{z_{k-l}}{2} \rfloor\right) dx \Bigg)^2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad dx_{l+1} \cdots dx_k dz_1 \cdots dz_{k-l}.
\end{aligned}$$

Then, applying the Cauchy-Schwarz inequality to the inner integral, we get

$$\begin{aligned}
&\left\|\bar{h}_k \star_l^l \bar{h}_k\right\|^2_{L^2(\mathbb{R}_+^{2(k-l)})} \\
&= \frac{1}{2^{2l}} \sum_{c, c' \in \mathbb{N}^{k-l}} \left(\sum_{b \in \mathbb{N}^l} \left(\sum_{a \in \mathbb{N}^{e_G - k}} \mathbf{1}_{E_G}(a, b, c) \right) \left(\sum_{a' \in \mathbb{N}^{e_G - k}} \mathbf{1}_{E_G}(a', b, c') \right) \right)^2 \\
&\quad \times \left(\int_{(0,2)^k} \left(\Psi_{x_1} \cdots \Psi_{x_k} g_k(x_1, \ldots, x_k) \right)^2 dx_1 \cdots dx_k \right)^2.
\end{aligned}$$

Since $k \ge 1$, the formula (5.6) in Lemma 5.4 gives us

$$\begin{aligned}
&\left(\int_{(0,2)^k} \left(\Psi_{x_1} \cdots \Psi_{x_k} g_k(x_1, \ldots, x_k) \right)^2 dx_1 \cdots dx_k \right)^2 \\
&\quad \lesssim \; p^{4e_G - 2k} (1-p)^{2k-2} \Big(\mathbb{E}\big[\left(X^2 - \mathbb{E}[X] \right)^2 \big] + (1-p)(\mathbb{E}[X])^2 \Big)^2 \\
&\quad \lesssim \; \frac{p^{4e_G - 2k}}{1-p} \Big(\mathbb{E}\big[(X - \mathbb{E}[X])^4 \big] + (1-p)^2 (\mathbb{E}[X])^4 \Big).
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
&\sum_{c, c' \in \mathbb{N}^{k-l}} \left(\sum_{b \in \mathbb{N}^l} \left(\sum_{a \in \mathbb{N}^{e_G - k}} \mathbf{1}_{E_G}(a, b, c) \right) \left(\sum_{a' \in \mathbb{N}^{e_G - k}} \mathbf{1}_{E_G}(a', b, c') \right) \right)^2 \\
&\qquad \lesssim \max_{\substack{K, H, L \subset G \\ e_K = k-l,\ e_H = l,\ e_L = k}} n^{4v_G - v_K - v_H - v_L},
\end{aligned}$$

see the proof of Theorem 4.2 in [17], from which it follows

$$\begin{aligned} S_3 &= \sum_{1\le l<k\le e_G} \left\|\bar h_k \star_l^l \bar h_k\right\|^2_{L^2(\mathbb{R}_+^{2(k-l)})} \\ &\lesssim \frac{\mathbb{E}\big[(X-\mathbb{E}\,[X])^4\big] + (1-p)^2(\mathbb{E}\,[X])^4}{1-p} \max_{\substack{H\subset G\\ e_H\ge 1}} n^{4v_G-3v_H} p^{4e_G-3e_H}, \end{aligned} \tag{4.11}$$

as in the proof of Theorem 4.2 in [17], which concludes the proof by (4.1) and (4.7). □

We note that the bound (4.6) implies

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \lesssim \left(\frac{(\mathbb{E}[X])^2}{\mathrm{Var}[X]} + \sqrt{\kappa_X}\right)\left((1-p_n)\min_{\substack{H\subset G\\ e_H\ge 1}} n^{v_H} p_n^{e_H}\right)^{-1/2},$$

where $\mathbb{E}[X]/\sqrt{\mathrm{Var}[X]}$ is the standardized first moment of X and

$$\kappa_X := \frac{\mathbb{E}[(X-\mathbb{E}[X])^4]}{(\mathrm{Var}[X])^2}$$

is the kurtosis of X.

In the next corollary we note that Theorem 4.3 simplifies if we narrow our attention to p_n depending of the complete graph size n and close to 0 or to 1.

Corollary 4.4 *Let G be a graph without separated vertices. For $p_n < c < 1$, $n \ge 1$, we have*

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \lesssim \frac{\sqrt{\mathbb{E}[X^4]}}{\mathbb{E}[X^2]}\left((1-p_n)\min_{\substack{H\subset G\\ e_H\ge 1}} n^{v_H} p_n^{e_H}\right)^{-1/2}.$$

On the other hand, for $p_n > c > 0$, $n \ge 1$, it holds

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \lesssim \frac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{1-p_n}\,\mathrm{Var}[X]}.$$

Furthermore, it turns out that the minimum appearing in Theorem 4.3 and above for a wide class of graphs satisfying a certain balance condition. Precisely, let us consider the class $\mathcal{B}$ of all graphs with at least three vertices, and such that

$$\max_{\substack{H\subset G\\ v_H\ge 3}} \frac{e_H-1}{v_H-2} = \frac{e_G-1}{v_G-2},$$

as introduced in [17]. It has been shown there that a graph with at least three vertices and at least one edge belongs to $\mathcal{B}$ if and only if for any $p\in(0,1)$ and $n \ge v_G$ we have

$$\min_{\substack{H\subset G\\ e_H\ge 1}} n^{v_H} p^{e_H} = \min\{n^2 p, n^{v_G} p^{e_G}\}.$$

An application of this fact to Corollary 4.4 yields the following result.

Proposition 4.5 *For $G \in \mathcal{B}$ without isolated vertices and $c \in (0,1)$ we have*

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \lesssim \begin{cases} \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{1-p_n}\mathrm{Var}[X]} & \text{if } \ 0 < c < p_n, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{p_n}\mathbb{E}[X^2]} & \text{if } \ n^{-(v_g-2)/(e_G-1)} < p_n \le c, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{n^{v_G/2}p_n^{e_G/2}\mathbb{E}[X^2]} & \text{if } \ 0 < p_n \le n^{-(v_G-2)/(e_G-1)}. \end{cases}$$

The following Corollaries 4.6-4.8 of Proposition 4.5 can be proved similarly to their counterparts Corollaries 4.8-4.10 in [17]. The next Corollary 4.6 deals with cycle graphs with r vertices, $r \ge 3$, and in particular with triangles when $r = 3$.

Corollary 4.6 *Let G be a cycle graph with r vertices, $r \ge 3$, and $c \in (0,1)$. We have*

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \lesssim \begin{cases} \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{1-p_n}\mathrm{Var}[X]} & \text{if } \ 0 < c < p_n, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{p_n}\mathbb{E}[X^2]} & \text{if } \ n^{-(r-2)/(r-1)} < p_n \le c, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{(np_n)^{r/2}\mathbb{E}[X^2]} & \text{if } \ 0 < p_n \le n^{-(r-2)/(r-1)}. \end{cases}$$

The next corollary deals with complete graphs, which also covers the case of triangles.

Corollary 4.7 *Let G be a complete graph with r vertices, $r \ge 3$, and $c \in (0,1)$. We have*

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \lesssim \begin{cases} \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{1-p_n}\mathrm{Var}[X]} & \text{if } \ c < p_n < 1, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{p_n}\mathbb{E}[X^2]} & \text{if } \ n^{-2/(r+1)} < p_n \le c, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{n^{r/2}p_n^{r(r-1)/4}\mathbb{E}[X^2]} & \text{if } \ 0 < p_n \le n^{-2/(r+1)}. \end{cases}$$

Finally, the last corollary deals with the important class of graphs which have a tree structure.

Corollary 4.8 *Let G be any tree (a connected graph without cycles) with r edges, and $c \in (0,1)$. We have*

$$d_W\big(\widetilde{W}_n^G, \mathcal{N}\big) \lesssim \begin{cases} \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{1-p_n}\operatorname{Var}[X]} & \text{if } c < p_n < 1, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{n\sqrt{p_n}\mathbb{E}[X^2]} & \text{if } \dfrac{1}{n} < p_n \leq c, \\[2ex] \dfrac{\sqrt{\mathbb{E}[X^4]}}{n^{(r+1)/2}p_n^{r/2}\mathbb{E}[X^2]} & \text{if } 0 < p_n \leq \dfrac{1}{n}. \end{cases}$$

5 Appendix

In this section we gather a number of technical results, starting with the following multiplication formula for multiple stochastic integrals, which involves the $\star$-notation introduced in (3.2). For $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$ and $g_m \in \widehat{L}^2(\mathbb{R}_+^m)$ satisfying (2.4) the following multiplication formula holds:

$$I_n(f_n)I_m(g_m) = \sum_{k=0}^{m\wedge n} k!\binom{m}{k}\binom{n}{k}\sum_{i=0}^{k}\binom{k}{i} I_{m+n-k-i}\big(f_n \,\widetilde{\star}_k^i\, g_m\big), \tag{5.1}$$

whenever $f_n \star_k^i g_m \in L^2(\mathbb{R}_+^{m+n-k-i})$ for every $0 \leq i \leq k \leq m \wedge n$, see Proposition 5.1 of [16]. The next proposition allows us to bound the L^2 norm of $f_n \star g_m$ by some simpler expressions, which is used in the proof of Theorem 3.2.

Proposition 5.1 *Let $f_n \in L^2(\mathbb{R}_+^n)$ and $g_m \in L^2(\mathbb{R}_+^m)$ be symmetric functions. For $0 \leq l < k \leq n \wedge m$ we have*

$$\left\| f_n \star_k^l g_m \right\|^2_{L^2(\mathbb{R}_+^{m+n-k-l})} 2^{2n-2k-1} \leq \left\| f_n \star_n^{l+n-k} f_n \right\|^2_{L^2(\mathbb{R}_+^{k-l})} + 2^{2m-2k-1}\left\| g_m \star_m^{l+m-k} g_m \right\|^2_{L^2(\mathbb{R}_+^{k-l})}, \tag{5.2}$$

and for $0 \leq k \leq n \wedge m$ we have

$$\left\| f_n \star_k^k g_m \right\|^2_{L^2(\mathbb{R}_+^{m+n-2k})} \leq 2^{2n-4k-1}\left\| f_n \star_{n-k}^{n-k} f_n \right\|^2_{L^2(\mathbb{R}_+^{2k})} + 2^{2m-4k-1}\left\| g_m \star_{m-k}^{m-k} g_m \right\|^2_{L^2(\mathbb{R}_+^{2k})}. \tag{5.3}$$

Proof. Let $x \in \mathbb{R}_+^l$, $y \in \mathbb{R}_+^{k-l}$, $u \in \mathbb{R}_+^{n-k}$ and $z \in \mathbb{R}_+^{m-k}$. Hölder's inequality applied twice gives us

$$\left\| f_n \star_k^l g_m \right\|^2_{L^2(\mathbb{R}_+^{m+n-k-l})}$$

$$
\begin{aligned}
&= \frac{1}{2^{2l}} \int_{\mathbb{R}_+^{m-k}} \int_{\mathbb{R}_+^{n-k}} \int_{\mathbb{R}_+^{k-l}} \left(\int_{\mathbb{R}_+^{l}} f_n(x,y,u) g_m(x,y,z) dx \right)^2 dy du dz \\
&\leq \frac{1}{2^{2l}} \int_{\mathbb{R}_+^{k-l}} \int_{\mathbb{R}_+^{m-k}} \int_{\mathbb{R}_+^{n-k}} \int_{\mathbb{R}_+^{l}} f_n^2(x,y,u) dx \int_{\mathbb{R}_+^{l}} g_m^2(x,y,z) dx du dz dy \\
&\leq \frac{1}{2^{2l}} \sqrt{ \int_{\mathbb{R}_+^{k-l}} \left(\int_{\mathbb{R}_+^{n-k}} \int_{\mathbb{R}_+^{l}} f_n^2(x,y,u) dx du \right)^2 dy \int_{\mathbb{R}_+^{k-l}} \left(\int_{\mathbb{R}_+^{m-k}} \int_{\mathbb{R}_+^{l}} g_m^2(x,y,z) dx dz \right)^2 dy } \\
&\leq \frac{1}{2^{2l+1}} \left[\int_{\mathbb{R}_+^{k-l}} \left(\int_{\mathbb{R}_+^{n-k}} \int_{\mathbb{R}_+^{l}} f_n^2(x,y,u) dx du \right)^2 dy + \int_{\mathbb{R}_+^{k-l}} \left(\int_{\mathbb{R}_+^{m-k}} \int_{\mathbb{R}_+^{l}} g_m^2(x,y,z) dx dz \right)^2 dy \right] \\
&= 2^{2n-2k-1} \left\| f_n \star_n^{l+n-k} f_n \right\|^2_{L^2(\mathbb{R}_+^{k-l})} + 2^{2m-2k-1} \left\| g_m \star_m^{l+m-k} g_m \right\|^2_{L^2(\mathbb{R}_+^{k-l})},
\end{aligned}
$$

where we used the inequality $\sqrt{ab} \leq (a+b)/2$, $a, b \geq 0$, which proves the first assertion. Furthermore, for $x, u \in \mathbb{R}_+^k$, $y \in \mathbb{R}_+^{n-k}$ and $z \in \mathbb{R}_+^{m-k}$ we get

$$
\begin{aligned}
&\left\| f_n \star_k^k g_m \right\|^2_{L^2(\mathbb{R}_+^{m+n-2k})} \\
&= \frac{1}{2^{2k}} \int_{\mathbb{R}_+^{n-k}} \int_{\mathbb{R}_+^{m-k}} \int_{\mathbb{R}_+^{k}} f_n(u,y) g_m(u,z) du \int_{\mathbb{R}_+^{k}} f_n(x,y) g_m(x,z) dx dy dz \\
&\leq \frac{1}{2^{2k}} \int_{\mathbb{R}_+^{k}} \int_{\mathbb{R}_+^{k}} \left(\int_{\mathbb{R}_+^{n-k}} f_n(u,y) f_n(x,y) dy \right) \left(\int_{\mathbb{R}_+^{m-k}} g_m(u,z) g_m(x,z) dz \right) du dx \\
&\leq \frac{1}{2^{2k+1}} \left[\int_{\mathbb{R}_+^{k}} \int_{\mathbb{R}_+^{k}} \left(\int_{\mathbb{R}_+^{n-k}} f_n(u,y) f_n(x,y) dy \right)^2 du dx \right. \\
&\qquad\qquad \left. + \int_{\mathbb{R}_+^{k}} \int_{\mathbb{R}_+^{k}} \left(\int_{\mathbb{R}_+^{m-k}} g_m(u,z) g_m(x,z) dz \right)^2 du dx \right] \\
&\leq 2^{2n-4k-1} \left\| f_n \star_{n-k}^{n-k} f_n \right\|^2_{L^2(\mathbb{R}_+^{2k})} + 2^{2m-4k-1} \left\| g_m \star_{m-k}^{m-k} g_m \right\|^2_{L^2(\mathbb{R}_+^{2k})}.
\end{aligned}
$$

□

The next proposition presents some relationships between second norms involving operators ∇, L and $(-L)^{1/2}$.

Proposition 5.2 *For F such that $LF \in L^2(\Omega)$ we have*

$$
\mathbb{E}\left[\int_0^\infty (\nabla_t F)^2 \frac{dt}{2} \right] = \mathbb{E}\left[\left((-L)^{1/2} F \right)^2 \right] \leq \mathbb{E}\left[(LF)^2 \right]. \tag{5.4}
$$

Proof. Using the chaos decomposition (2.6), where the sequence of functions f_n in $\widehat{L}^2(\mathbb{R}_+^n)$, $n \geq 1$, satisfies the Condition (2.4), and by the isometry relation (2.5) we have

$$
\mathbb{E}\left[\int_0^\infty |\nabla_t F|^2 \frac{dt}{2} \right] = \sum_{n=1}^\infty \mathbb{E}\left[\int_0^\infty |\nabla_t I_n(f_n)|^2 \frac{dt}{2} \right]
$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} \mathbb{E}\left[\int_0^\infty |\nabla_t I_n(f_n)|^2 \frac{dt}{2}\right] \\
&= \sum_{n=1}^{\infty} n^2 \mathbb{E}\left[\int_0^\infty |I_{n-1}(f_n(t,\cdot))|^2 \frac{dt}{2}\right] \\
&= \sum_{n=1}^{\infty} n^2 (n-1)! \int_0^\infty \|f_n(t,\cdot)\|^2_{\widehat{L}^2(\mathbb{R}_+^{n-1}, dx/2)} \frac{dt}{2} \\
&= \sum_{n=1}^{\infty} n \mathbb{E}\left[|I_n(f_n)|^2\right] \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left[|(-L)^{1/2} I_n(f_n)|^2\right] \\
&= \mathbb{E}\big[\big((-L)^{-1/2} F\big)^2\big],
\end{aligned}
$$

which is the first part of the assertion. This also implies

$$
\mathbb{E}\big[\,\big((-L)^{-1/2} F\big)^2\,\big] = \sum_{n=1}^{\infty} n \mathbb{E}\left[|I_n(f_n)|^2\right] \le \sum_{n=1}^{\infty} n^2 \mathbb{E}\left[|I_n(f_n)|^2\right] \le \mathbb{E}[(LF)^2],
$$

which ends the proof. □

Next, let us recall the definition (4.2) of the operator Ψ_{t_i}

$$
\Psi_{t_i} f(t_1, \ldots, t_n) := f(t_1, \ldots, t_n) - \frac{1}{2} \int_{2\lfloor t_i/2 \rfloor}^{2\lfloor t_i/2 \rfloor + 2} f(t_1, \ldots, t_{i-1}, s, t_{i+1}, \ldots, t_n) ds,
$$

$i = 1, \ldots, n$, $t_1, \ldots, t_n \in \mathbb{R}_+$. The following result is the analog of the Stroock formula [22] in our framework. It shows that any multiple integral can be expressed as a degenerate generalized U-statistic, see Remark 2.2.

Proposition 5.3 *For every $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$ there exists a unique $\bar{f}_n \in \widehat{L}^2(\mathbb{R}_+^n)$ satisfying (2.4) such that $I_n(f_n) = I_n(\bar{f}_n)$, and it is given by*

$$
\bar{f}_n(t_1, \ldots, t_n) = \Psi_{t_1} \cdots \Psi_{t_n} f_n(t_1, \ldots, t_n) = \frac{1}{n!} \nabla_{t_1} \cdots \nabla_{t_n} I_n(f_n). \tag{5.5}
$$

Proof. Uniqueness of $\bar{f}_n$ follows from the isometry relation (2.5). We can also check that the condition (2.4) is satisfied by integrating (4.2) with respect to $t_i \in \mathbb{R}_+$. Furthermore, the equality (5.5) is clear for $n = 1$. Assuming that it holds for some $n - 1 \ge 1$, we get

$$
\begin{aligned}
I_n(f_n) &= \int_0^\infty I_{n-1}\left(f_n(t_1, \cdot)\right) d(Y_{t_1} - t_1/2) \\
&= \int_0^\infty I_{n-1}\left(\Psi_{t_2} \cdots \Psi_{t_n} f_n(t_1, \cdot)\right) d(Y_{t_1} - t_1/2)
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^\infty \Psi_{t_1} I_{n-1}\left(\Psi_{t_2}\cdots\Psi_{t_n} f_n(t_1,\cdot)\right) dY_{t_1} \\
&= \int_0^\infty I_{n-1}\left(\Psi_{t_1}\cdots\Psi_{t_n} f_n(t_1,\cdot)\right) dY_{t_1} \\
&= \int_0^\infty I_{n-1}\left(\Psi_{t_1}\cdots\Psi_{t_n} f_n(t_1,\cdot)\right) d(Y_{t_1} - t_1/2) \\
&= I_n\left(\Psi_{t_1}\cdots\Psi_{t_n} f_n\right).
\end{aligned}
$$

Eventually, the latter equality in (5.5) follows from (2.9). □

The following Lemma 5.4 is used to bound the kernel functions $\bar{h}_k$ appearing in Lemma 4.2.

Lemma 5.4 *The functions g_k defined in (4.5) satisfy the inequalities*

$$
\begin{aligned}
&\int_{(0,2)^k} \left(\Psi_{x_1}\cdots\Psi_{x_k} g_k\left(x_1,\ldots,x_k\right)\right)^2 dx_1\cdots dx_k \\
&\qquad \lesssim p^{2e_G-k}(1-p)^{k-1}\Big(\mathbb{E}\Big[\left(X^2-\mathbb{E}\left[X\right]\right)^2\Big] + (1-p)(\mathbb{E}\left[X\right])^2\Big)
\end{aligned}
\tag{5.6}
$$

and

$$
\begin{aligned}
&\int_{(0,2)^{k-l}} \left(\int_{(0,2)^l} \big(\Psi_{x_1}\cdots\Psi_{x_k} g_k\left(x_1,\ldots,x_k\right)\big)^2 dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k \\
&\qquad \lesssim p^{4e_G-3k+l}(1-p)^{k+l-2}\big(\mathbb{E}\left[(X-\mathbb{E}\left[X\right])^4\right] + (1-p)^2(\mathbb{E}\left[X\right])^4\big),
\end{aligned}
\tag{5.7}
$$

$0 \le l \le k \le n$, *where the operator Ψ_x is defined in (4.2).*

Proof. We decompose $g_k(x_1,\ldots,x_k)$ as

$$
g_k(x_1,\ldots,x_k) = \sum_{i=0}^k g_k^{(i)}(x_1,\ldots,x_k)
$$

where

$$
g_k^{(0)}(x_1,\ldots,x_k) := \frac{(p/2)^{e_G-k}}{(e_G-k)!k!}(e_G-k)\mathbb{E}[X]\mathbf{1}_{(0,2p)^k}\left(x_1,\ldots,x_k\right),
$$

and

$$
g_k^{(i)}(x_1,\ldots,x_k) := \frac{(p/2)^{e_G-k}}{(e_G-k)!k!}\mathbf{1}_{(0,2p)^k}\left(x_1,\ldots,x_k\right) F_X^{-1}\left(\frac{x_i}{2p}\right), \qquad 1 \le i \le k.
$$

Next, for $1 \le i \le k$ we have

$$
\begin{aligned}
&\Psi_{x_1}\cdots\Psi_{x_k} g_k^{(i)}(x_1,\ldots,x_k) \\
&= \frac{(p/2)^{e_G-k}}{(e_G-k)!k!}\left(\mathbf{1}_{(0,2p)}\left(x_i\right) F_X^{-1}\left(\frac{x_i}{2p}\right) - p\mathbb{E}\left[X\right]\right) \prod_{\substack{1\le j\le k \\ j\ne i}} \left(\mathbf{1}_{(0,2p)}\left(x_j\right) - p\right).
\end{aligned}
$$

Thus we have

$$\begin{aligned}
&\int_{(0,2)^k} \big(\Psi_{x_1}\cdots\Psi_{x_k} g_k^{(i)}(x_1,\ldots,x_k)\big)^2 dx_1\cdots dx_k \\
&= \frac{2^k p^{2e_G-k-1}(1-p)^{k-1}}{((e_G-k)!k!)^2 2^{2e_G-2k}}\big(p\mathbb{E}\big[(X-p\mathbb{E}[X])^2\big]+(1-p)(p\mathbb{E}[X])^2\big) \\
&\lesssim p^{2e_G-k}(1-p)^{k-1}\big(\mathbb{E}[X^2]-p(\mathbb{E}[X])^2\big), \qquad 1\le i\le k,
\end{aligned}$$

and similarly for $g_k^{(0)}(x_1,\ldots,x_k)$, which gives

$$\begin{aligned}
&\int_{(0,2)^k} \left(\Psi_{x_1}\cdots\Psi_{x_k} g_k(x_1,\ldots,x_k)\right)^2 dx_1\cdots dx_k \\
&\qquad \lesssim \sum_{i=0}^{k}\int_{(0,2)^k} \big(\Psi_{x_1}\cdots\Psi_{x_k} g_k^{(i)}(x_1,\ldots,x_k)\big)^2 dx_1\cdots dx_k \\
&\qquad \lesssim p^{2e_G-k}(1-p)^{k-1}\big(\mathbb{E}[X^2]-p(\mathbb{E}[X])^2\big),
\end{aligned}$$

as required. In order to prove (5.7), we proceed similarly and get

$$\begin{aligned}
&\int_{(0,2)^{k-l}}\left(\int_{(0,2)^l}\Big(\Psi_{x_1}\cdots\Psi_{x_k} g_k^{(i)}(x_1,\ldots,x_k)\Big)^2 dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k \\
&\qquad \lesssim p^{4e_G-3k+l}(1-p)^{k+l-2}\big(\mathbb{E}[X^2]-p(\mathbb{E}[X])^2\big)^2
\end{aligned}$$

for $1\le i\le l$, and

$$\begin{aligned}
&\int_{(0,2)^{k-l}}\left(\int_{(0,2)^l}\Big(\Psi_{x_1}\cdots\Psi_{x_k} g_k^{(i)}(x_1,\ldots,x_k)\Big)^2 dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k \\
&\qquad \lesssim p^{4e_G-3k+l}(1-p)^{k+l-1}\big(\mathbb{E}\left[(X-\mathbb{E}[X])^4\right]+(1-p)(\mathbb{E}[X])^4\big)
\end{aligned}$$

for $l<i\le k$. Hence, by the Cauchy-Schwarz inequality we get

$$\begin{aligned}
&\int_{(0,2)^{k-l}}\left(\int_{(0,2)^l}\left(\Psi_{x_1}\cdots\Psi_{x_k} g_k(x_1,\ldots,x_k)\right)^2 dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k \\
&\lesssim \sum_{i=0}^{k}\int_{(0,2)^{k-l}}\left(\int_{(0,2)^l}\big(\Psi_{x_1}\cdots\Psi_{x_k} g_k^{(i)}(x_1,\ldots,x_k)\big)^2 dx_1\cdots dx_l\right)^2 dx_{l+1}\cdots dx_k \\
&\lesssim p^{4e_G-3k+l}(1-p)^{k+l-2}\big(\mathbb{E}\left[(X-\mathbb{E}[X])^4\right]+(1-p)^2(\mathbb{E}[X])^4\big),
\end{aligned}$$

which ends the proof. □

References

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

[2] L.H.Y. Chen and Q.M. Shao. Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli*, 13(2):581–599, 2007.

[3] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[4] X. Fang. Wasserstein-2 bounds in normal approximation under local dependence. *Electron. J. Probab.*, 24:Paper No. 35, 14, 2019.

[5] E.N. Gilbert. Random graphs. *Ann. Math. Statist*, 30(4):1141–1144, 1959.

[6] F. Götze and A.N. Tikhomirov. Asymptotic distribution of quadratic forms. *Ann. Probab.*, 27(2):1072–1098, 1999.

[7] S. Janson, T. Luczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[8] K. Krokowski. Poisson approximation of Rademacher functionals by the Chen-Stein method and Malliavin calculus. *Commun. Stoch. Anal.*, 11(2):195–222, 2017.

[9] K. Krokowski, A. Reichenbachs, and C. Thäle. Berry-Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(2):763–803, 2016.

[10] K. Krokowski, A. Reichenbachs, and C. Thäle. Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation. *Ann. Probab.*, 45(2):1071–1109, 2017.

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

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

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

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

[15] N. Privault. Calcul des variations stochastique pour la mesure de densité uniforme. *Potential Analysis*, 7(2):577–601, 1997.

[16] N. Privault and G. Serafin. Stein approximation for functionals of independent random sequences. *Electron. J. Probab.*, 23:Paper No. 4, 34, 2018.

[17] N. Privault and G. Serafin. Normal approximation for sums of discrete U-statistics - application to Kolmogorov bounds in random subgraph counting. *Bernoulli*, 26(1):587–615, 2020.

[18] N. Privault and G.L. Torrisi. The Stein and Chen-Stein methods for functionals of non-symmetric Bernoulli processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12:309–356, 2015.

[19] A. Röllin. Kolmogorov bounds for the normal approximation of the number of triangles in the Erdős-Rényi random graph. Preprint arXiv:1704.00410, to appear in Probability in the Engineering and Informational Sciences, DOI:10.1017/S0269964821000061, 27 pages, 2021.

[20] N. Ross. Fundamentals of Stein's method. *Probab. Surv.*, 8:201–293 (electronic), 2011.

[21] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields*, 78:1–10, 1988.

[22] D. Stroock. Homogeneous chaos revisited. In J. Azéma, P.A. Meyer, and M. Yor, editors, *Séminaire de Probabilités, XXI*, volume 1247 of *Lecture Notes in Math.*, pages 1–7. Springer, Berlin, 1987.

[23] D. Surgailis. Non-CLTs: U-statistics, multinomial formula and approximations of multiple Itô-Wiener integrals. In P. Doukhan, G. Oppenheim, and M. Taqqu, editors, *Theory and applications of long-range dependence*, pages 129–142. Birkhäuser Boston, Boston, MA, 2003.

[24] Z.S. Zhang. Cramér-type moderate deviation of normal approximation for exchangeable pairs. Preprint arXiv:1901.09526, 2019.