Berry-Esseen bounds for functionals of independent random variables

Nicolas Privault* Grzegorz Serafin†‡

October 18, 2024

Abstract

We derive Berry-Esseen approximation bounds for general functionals of independent random variables, based on a continuous-time integration by parts setting and discrete chaos expansions methods. Our approach improves on related results obtained in discrete-time integration by parts settings and applies to U-statistics satisfying the weak assumption of decomposability in the Hoeffding sense, and yield Kolmogorov distance bounds instead of the Wasserstein bounds previously derived in the special case of degenerate U-statistics. Linear and quadratic functionals of arbitrary sequences of independent random variables are included as particular cases, with new fourth moment bounds, and applications are given to Hoeffding decompositions, weighted U-statistics, quadratic forms, and random subgraph weighing. In the case of quadratic forms, our results recover and improve the bounds available in the literature, and apply to matrices with non-empty diagonals.

Keywords: Stein-Chen method; Berry-Esseen bounds; Kolmogorov distance; U-statistics; quadratic forms; Malliavin calculus.
Mathematics Subject Classification: 60F05; 60G57; 60H07.

1 Introduction

Significant progress in probability approximation has been achieved in recent years by combining the Chen-Stein method with the Malliavin calculus. See for example Nourdin and Peccati (2009), Peccati et al. (2010), Peccati and Thäle (2013), for the derivation of distance bounds on the Wiener and Poisson spaces, and also Nourdin et al. (2010a) and Krokowski et al. (2016) in the case of Rademacher sequences. Those results rely on covariance representations based on the inverse of the Ornstein-Uhlenbeck operator L acting on multiple

*Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-05-43, 21 Nanyang Link Singapore 637371. e-mail: `nprivault@ntu.edu.sg`.
†Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Ul. Wybrzeże Wyspiańskiego 27, Wrocław, Poland. e-mail: `grzegorz.serafin@pwr.edu.pl`.
‡Corresponding author

Wiener-Poisson stochastic integrals. While the inverse operator L^{-1} is well adapted to certain random functionals such as multiple stochastic integrals, it can prove more difficult to use in applications to other, more specific functionals. Other covariance representations based on the Clark-Ocone representation formula and not relying on L^{-1} have been used in Privault and Torrisi (2013) on the Wiener and Poisson spaces, and in Privault and Torrisi (2015) for Rademacher sequences.

In Last et al. (2016), second order Poincaré inequalities in the Kolmogorov and Wasserstein distances have been obtained for functionals of a Poisson point process by using the iterated Malliavin gradient instead of L^{-1}. This approach relies on probabilistic representations for the inverse operator L^{-1} using Mehler's formula on the Poisson space, see e.g. Lemma 6.8.1 in Privault (2009). Second order Poincaré inequalities for functionals of Rademacher sequences have also been obtained in Krokowski et al. (2017), with application to renormalized triangle counting using the Kolmogorov distance in the Erdős-Rényi random graph, see also Privault and Serafin (2020) and references therein for the treatment of arbitrary subgraph counting.

In Chatterjee (2008), a method based on difference operators has been introduced with the aim of obtaining Stein bounds in the Wasserstein distance for functions of vectors of independent random variables. This approach has been extended in Lachièze-Rey and Peccati (2017) to the derivation of bounds in the Kolmogorov distance, see also Friedrich (1989) for earlier related results. .

An integration by parts setting for related difference operators has been exploited in Decreusefond and Halconruy (2019) to derive normal Stein approximation bound for functionals of independent random variables, see also Nguyen (2020), and Bobkov et al. (2019) for concentration inequalities. In Duerinckx (2021), this framework has been unified with the approaches of Chatterjee (2008) and Lachièze-Rey and Peccati (2017) with applications in statistical physics, see also Duerinckx et al. (2020).

In Privault and Serafin (2018), a general framework for the derivation of Wasserstein distance bounds for functionals of independent random sequences has been developed in the continuous-time integration by parts setting of Privault (1997), using an analog of the operator L^{-1} on discrete chaos expansions based on discrete multiple stochastic integrals. This approach allows us to extend chaos-based arguments from the binomial and Wiener-Poisson settings to general i.i.d. sequences of random variables.

Bounds in total variance distance have also been obtained therein using Clark-Ocone covariance representation formulas under stronger smoothness conditions. Applications to normal approximation in the Wasserstein distance have been obtained in Privault and Serafin (2022) for the weights of subgraphs in the Erdős-Rényi random graph.

Our first goal in this paper is to extend existing Stein normal approximation bounds proved in the Kolmogorov distance for Rademacher sequences, see e.g. Krokowski et al. (2017), Döbler and Krokowski (2019), to general sequences of independent random variables. This is achieved in the general framework of Privault and Serafin (2018), by replacing the Wasserstein distance with the Kolmogorov distance for which obtaining rates is known to be more difficult and requires new ideas. In Theorem 3.1 we derive a general Berry-Esseen bound in the Kolmogorov distance for functionals of independent random variables. In comparison with Theorem 4.2 in Lachièze-Rey and Peccati (2017), the variance term (3.2) in Theorem 3.1 can be easier to control, see also Theorem 2.3 in Duerinckx (2021).

The bound of Theorem 3.1 is then specialized to sums of multiple stochastic integrals in Proposition 3.2, and then to multiple stochastic integrals in Proposition 3.3. Note that multiple stochastic integrals of order d coincide with degenerate (generalized) U-statistics of order d, and can then be used to represent Hoeffding decompositions as a chaos summations, see the examples given below.

Our second goal is to show that the obtained bounds remain sharp despite the very general framework of the paper, as demonstrated in the following examples. Consider a sequence $(X_1, \ldots, X_n)$ of (not necessarily identically distributed) independent random variables, and the d-homogeneous random multilinear forms $W_{n,d}$ written in the Hoeffding form as

$$W_{n,d} = \sum_{J \subset \{1,\ldots,n\},\ |J|=d} W_J,$$

where, for each $J \subset \{1, \ldots, n\}$, W_J is a random variable with variance σ_J^2, measurable with respect to the σ-algebra $\mathcal{F}_J := \sigma(X_j : j \in J)$, and such that $\mathbb{E}[W_J \mid \mathcal{F}_K] = 0$, $J \not\subseteq K \subset [n]$. In de Jong (1990), a central limit theorem has been proved for the sequence $(W_{n,d})_{n\geq 1}$ under the conditions

$$\lim_{n\to\infty} \max_{1\leq i\leq n} \sum_{J\ni i} \sigma_J^2 = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{E}[W_{n,d}^4] = 3,$$

generalizing earlier results by de Jong (1987) for quadratic random functionals. The results of de Jong (1987; 1990) have been refined by the derivation of bounds in the Wasserstein distance in Theorem 1.3 in Döbler and Peccati (2017) in the case of degenerate U-statistics, for which $|J|$ is constrained to a fixed value $|J| = d$ for some $d \in \{1, \ldots, n\}$ in the sum (2.17).

Applications of Proposition 3.2 are given to Kolmogorov distance bounds in Theorem 4.1 for general U-statistics, and in Theorems 4.2 and 4.3 for degenerate U-statistics. This extends the bounds of Döbler and Peccati (2017) by using the Kolmogorov distance instead of the Wasserstein distance, and by applying to Hoeffding decompositions in full generality and not only to degenerate U-statistics. This also extends the bounds in the Kolmogorov distance derived in Döbler and Krokowski (2019) for U-statistics in the particular case of Rademacher chaoses, where $(X_1, \ldots, X_n)$ is a sequence of independent Bernoulli random variables.

More specifically, given an i.i.d. sequence $(X_k)_{k\geq 1}$ of centered random variables with unit variance, and the sum

$$Z_n := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k, \qquad n \geq 1,$$

convergence bounds to the standard normal distribution $\mathcal{N}$ of the form

$$d_W(Z_n, \mathcal{N}) \leq \frac{E\left[|X_1|^3\right]}{\sqrt{n}}$$

have been obtained in e.g. Theorem 1.1 in Goldstein (2010) in the Wasserstein distance

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

See also Corollary 2.11 of Döbler (2015) for related bounds in the Kolmogorov distance

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

including the case of random sums. In the case of quadratic functionals of the form

$$Q_n := \sum_{1 \leq k, l \leq n} a_{kl} X_k X_l, \tag{1.1}$$

where $A = (a_{ij})_{1 \leq i,j \leq n}$ is a symmetric matrix, the bound

$$d_K(Q_n, \mathcal{N}) \leq C \left(E\left[|X_1|^3\right]\right)^2 |\lambda_1|, \tag{1.2}$$

where λ_1 denotes the largest absolute eigenvalue of A and $C > 0$ is an absolute constant, has been obtained in Götze and Tikhomirov (1999) when the diagonal of A vanishes, see e.g. Theorem 1 therein, and also Theorem 3.1 of Shao and Zhang (2019).

In this vanishing diagonals setting, Theorem 4.3 is applied to derive a version of Theorem 3.3 of Döbler and Peccati (2019) for the Kolmogorov distance instead of the Wasserstein distance in Corollary 4.4. Theorem 4.3 also yields Corollary 5.1 which recovers Theorem 3.1 in Shao and Zhang (2019), and improves on the above bound (1.2) of Theorem 1 in Götze and Tikhomirov (1999). In addition, Corollary 5.1 extends the Kolmogorov bounds of Theorem 1.1 in Döbler and Krokowski (2019), restricted to the quadratic case, from Rademacher sequences to general sequences of random variables by using fourth moment differences as in e.g. Theorem 1.3 of Döbler and Peccati (2017).

In case the diagonal of $A = (a_{ij})_{1\leq i,j\leq n}$ may not vanish, the bound

$$d_K\left(\frac{Q_n}{\sigma_n}, \mathcal{N}\right) \leq C(\gamma)\frac{\left(\mathbb{E}\left[|X|^3\right]\right)^2 + \gamma\mathbb{E}[X^6]}{\sqrt{\sum_{1\leq i,j\leq n} a_{ij}^2}}|\lambda_1|, \tag{1.3}$$

has been obtained in Theorem 1.1 of Götze and Tikhomirov (2002) for some $\gamma > 0$ depending on A. See also Proposition 3.1 in Chatterjee (2008) for a result in the Wasserstein distance using Rademacher sequences, and Theorem 2.2 in Chatterjee (2009) for related normal approximation bounds in total variation distance for a smooth function of finite-dimensional random vectors via second order Poincaré inequalities.

In comparison with Theorem 1.1 of Götze and Tikhomirov (2002), the bound (5.7) in Theorem 5.2 gives better rates under weaker assumptions according to the inequality (5.5). Theorem 5.2 also provides an additional bound (5.6) which is valid for any i.i.d. sequence $(X_n)_{n\geq 1}$ and holds in the Kolmogorov distance, instead of the Wasserstein distance used in Döbler and Peccati (2017). This bound is related to the so-called fourth moment phenomenon (Nualart and Peccati (2004)), which has been the object of intense research work, see e.g. Nourdin and Peccati (2012) and references therein.

We proceed as follows. In Section 2 we recall the framework of Privault (1997) for the treatment of functionals of independent random sequences, including the construction of discrete multiple stochastic integrals and the associated finite difference gradient operator and integration by parts formula, which are used to derive a fourth moment bound in Section 2.3. Section 3 contains our main result Theorem 3.1 which states a general Berry-Esseen bounds for general functionals of independent random sequences, and its applications to the derivation of Kolmogorov bounds for discrete multiple integrals and for sums of discrete multiple integrals in Propositions 3.2-3.3. Applications to Hoeffding decompositions, weighted

U-statistics and random subgraph weighing in the Erdős and Rényi (1959) random graph are given in Section 4. Section 5 focuses on quadratic forms.

2 Preliminaries

2.1 Setting

We work on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\Omega = [-1,1]^{\mathbb{N}}$ and $\mathcal{F}$, $\mathbb{P}$ are the natural σ-algebra and probability measure generated on Ω by the cylindrical Borel sets and Lebesgue measure, respectively. Let $(U_k)_{k\geq 1}$ denote the i.i.d. sequence of uniformly distributed $[-1,1]$-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, constructed as the canonical projections from Ω to $[-1,1]$. We define the finite difference gradient operator ∇ of a functional $F(U_1(\omega), U_2(\omega), \ldots)$ of the sequence $(U_1(\omega), U_2(\omega), \ldots)$ as

$$\nabla_t F := \mathbb{E}\Big[F \;\Big| U_1, \ldots, U_{\lfloor \frac{t}{2} \rfloor}, U_{\lfloor \frac{t}{2} \rfloor + 1} = t - 1 - 2\Big\lfloor \frac{t}{2} \Big\rfloor, U_{\lfloor \frac{t}{2} \rfloor + 2}, \ldots\Big] - \mathbb{E}\Big[F \Big| U_1, \ldots, U_{\lfloor \frac{t}{2} \rfloor}, U_{\lfloor \frac{t}{2} \rfloor + 2}, \ldots\Big],$$

$t \in \mathbb{R}_+$. In other words, using the shifted sequence

$$\Phi_t(\omega) := \Big(U_1(\omega), \ldots, U_{\lfloor \frac{t}{2} \rfloor}(\omega), t - 1 - 2\Big\lfloor \frac{t}{2} \Big\rfloor, U_{\lfloor \frac{t}{2} \rfloor + 2}(\omega), \ldots\Big), \qquad t \in \mathbb{R}_+,$$

we have

$$\nabla_t F := F \circ \Phi_t - \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} F \circ \Phi_s ds, \qquad t \in \mathbb{R}_+, \tag{2.1}$$

provided that $(F \circ \Phi_s)_{s \in \mathbb{R}_+}$ is integrable on $\mathbb{R}_+$, $\mathbb{P}$-a.s., see Definition 5 and Proposition 10 in Privault (1997). Although ∇_t does not satisfy the chain rule of derivation, we have the following identity.

Lemma 2.1 *The finite difference operator ∇ satisfies the relation*

$$\nabla_t(FG) = (F \circ \Phi_t)\nabla_t G + (G \circ \Phi_t)\nabla_t F - \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} (\nabla_t F \nabla_t G + \nabla_u F \nabla_u G) du, \tag{2.2}$$

$t \in \mathbb{R}_+$, provided that $(F \circ \Phi_s)_{s \in \mathbb{R}_+}$, $(G \circ \Phi_s)_{s \in \mathbb{R}_+}$ and $(F^2 \circ \Phi_s)_{s \in \mathbb{R}_+}$, $(G^2 \circ \Phi_s)_{s \in \mathbb{R}_+}$ are integrable on $[2n-2, 2n]$, $n \geq 1$, $\mathbb{P}$-a.s.

Proof. By (2.1), we have

$$\begin{aligned}
\nabla_t(FG) &= \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} ((FG) \circ \Phi_t - (FG) \circ \Phi_u) du \\
&= \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} (F \circ \Phi_u)(G \circ \Phi_t - G \circ \Phi_u) du + \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} (G \circ \Phi_t)(F \circ \Phi_t - F \circ \Phi_u) du
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}(F\circ\Phi_t)\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}(G\circ\Phi_t-G\circ\Phi_u)du+\frac{1}{2}(G\circ\Phi_t)\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}(F\circ\Phi_t-F\circ\Phi_u)du\\
&\quad -\frac{1}{2}\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}(F\circ\Phi_t-F\circ\Phi_u)(G\circ\Phi_t-G\circ\Phi_u)du\\
&= (F\circ\Phi_t)\nabla_t G+(G\circ\Phi_t)\nabla_t F-\frac{1}{2}\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}(F\circ\Phi_t-F\circ\Phi_u)(G\circ\Phi_t-G\circ\Phi_u)du.
\end{aligned}
$$

Furthermore, we have

$$
\begin{aligned}
\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}(F\circ\Phi_t-F\circ\Phi_u)(G\circ\Phi_t-G\circ\Phi_u)du &= \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}(\nabla_t F-\nabla_u F)(\nabla_t G-\nabla_u G)du\\
&= \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}(\nabla_t F\nabla_t G+\nabla_u F\nabla_u G)du,
\end{aligned}
$$

from the equality $\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\nabla_u F du=0$. □

For any $X\in L^1(\Omega)$ and $k\in\mathbb{N}$ we also note the identity

$$
\mathbb{E}[X]=\frac{1}{2}\mathbb{E}\left[\int_{2k}^{2k+2}X\circ\Phi_u du\right]. \tag{2.3}
$$

In addition, since $\int_{2k}^{2k+2}\nabla_u X du=0$ a.s. holds directly from the formula (2.1), for any $X,Y\in L^1(\Omega)$ we obtain

$$
\begin{aligned}
\int_{2k}^{2k+2}\nabla_u X\nabla_u Y du &= \int_{2k}^{2k+2}\nabla_u X\Phi_u Y du-\int_{2k}^{2k+2}\Phi_s Y ds\int_{2k}^{2k+2}\nabla_u X du\\
&= \int_{2k}^{2k+2}\nabla_u X\Phi_u Y du, \qquad a.s.
\end{aligned} \tag{2.4}
$$

Definition 2.2 *Given f_n in the space $\widehat{L}^2(\mathbb{R}_+^n)$ of square integrable symmetric functions on $\mathbb{R}_+^n$ that vanish outside of*

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

we define the multiple stochastic integral

$$
I_n(f_n)=n!\int_0^\infty\int_0^{t_n}\cdots\int_0^{t_2}f_n(t_1,\dots,t_n)d(Y_{t_1}-t_1/2)\cdots d(Y_{t_n}-t_n/2),
$$

with respect to the jump process $Y_t:=\sum_{k=1}^\infty \mathbf{1}_{[2k-1+U_k,\infty)}(t)$, $t\in\mathbb{R}_+$, which satisfies

$$
\begin{aligned}
I_n(f_n) &= \sum_{r=0}^n\left(-\frac{1}{2}\right)^{n-r}\binom{n}{r}\\
&\quad\times\sum_{k_1\neq\cdots\neq k_r\geq 1}\int_0^\infty\cdots\int_0^\infty f_n(2k_1-1+U_{k_1},\dots,2k_r-1+U_{k_r},y_1,\dots,y_{n-r})dy_1\cdots dy_{n-r}.
\end{aligned} \tag{2.5}
$$

The multiple stochastic integral $I_n(f_n)$ satisfies the bound

$$\mathbb{E}\left[(I_n(f_n))^2\right] \leq n! \, \|f_n\|^2_{L^2(\mathbb{R}^n_+,dx/2)}, \qquad n \geq 1,$$

which allows us to extend the definition of $I_n(f_n)$ to all $f_n \in \widehat{L}^2(\mathbb{R}^n_+)$, see Propositions 4 and 6 in Privault (1997). Under the additional condition

$$\int_{2k-2}^{2k} f_n(t,*)dt = 0, \qquad k \geq 1, \tag{2.6}$$

i.e. f_n is canonical in the sense of Surgailis (2003), the multiple stochastic integral $I_n(f_n)$ can be written as the U-statistics of order n

$$I_n(f_n) = \sum_{k_1 \neq \cdots \neq k_n \geq 1} f_n(2k_1 - 1 + U_{k_1}, \ldots, 2k_n - 1 + U_{k_n}),$$

with the isometry and orthogonality relation

$$\mathbb{E}\left[I_n(f_n)I_m(f_m)\right] = \mathbf{1}_{\{n=m\}} n! \langle f_n, f_m\rangle_{L^2(\mathbb{R}_+,dx/2)^{\otimes n}}, \quad f_n \in \widehat{L}^2(\mathbb{R}^n_+), \quad f_m \in \widehat{L}^2(\mathbb{R}^m_+), \tag{2.7}$$

see Proposition 6 in Privault (1997), which shows that the sequence $(I_n(f_n))_{n\geq 1}$ forms a family of mutually orthogonal centered random variables. Under the condition (2.6) we have the relation

$$\nabla_t I_n(f_n) = n I_{n-1}\left(f_n(t,*)\right), \quad t \in \mathbb{R}_+, \tag{2.8}$$

see Proposition 10 in Privault (1997).

The operator ∇ also 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. Precisely, the operator ∇ is closable with domain

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

see Proposition 8 in Privault (1997), and satisfies the duality relation (or integration by parts formula)

$$\mathbb{E}\left[\langle \nabla X, u\rangle_{L^2(\mathbb{R}_+,dx/2)}\right] = \mathbb{E}[X\nabla^*(u)], \tag{2.9}$$

which shows that ∇^* is closable as well, with domain $\mathrm{Dom}(\nabla^*) \subset L^2(\Omega \times \mathbb{R}_+)$. The operators (∇, ∇^*) are linked by the Skorohod isometry

$$\mathbb{E}[\nabla^* u \nabla^* v] = \mathbb{E}\left[\int_0^\infty u_t v_t dt\right] + \mathbb{E}\left[\int_0^\infty \int_0^\infty \nabla_s u_t \nabla_t v_s \, ds \, dt\right],$$

see Proposition 9 in Privault (1997), which yields the Poincaré inequality

$$\mathbb{E}\big[|\nabla^* u|^2\big] \leq \mathbb{E}\left[\int_0^\infty |u_t|^2 dt\right] + \mathbb{E}\left[\int_0^\infty \int_0^\infty |\nabla_s u_t|^2 ds\, dt\right]. \tag{2.10}$$

Finally, every $X \in L^2(\Omega)$ admits the chaos decomposition

$$X = E[X] + \sum_{n=1}^\infty I_n(f_n), \tag{2.11}$$

for some sequence of functions f_n in $\widehat{L}^2(\mathbb{R}_+^n)$, $n \geq 1$, cf. Proposition 7 in Privault (1997). Moreover, under the condition (2.6) the sequence $(f_n)_{n\geq 1}$ is unique in $\widehat{L}^2(\mathbb{R}_+^n)$ due to the isometry relation (2.7), and in this case we have

$$\mathbb{E}[X^2] = (\mathbb{E}[X])^2 + \sum_{n=1}^\infty n! \|f_n\|^2_{L^2(\mathbb{R}_+^n,(dx/2)^{\otimes n})}. \tag{2.12}$$

The operator L defined on linear combinations of multiple stochastic integrals as

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

is called the Ornstein-Uhlenbeck operator. By (2.11) the operator is invertible for centered $X \in L^2(\Omega)$, and its inverse operator L^{-1} is given by

$$L^{-1} I_n(f_n) = -\frac{1}{n} I_n(f_n), \qquad n \geq 1. \tag{2.13}$$

In fact, we can easily derive the form of any real power of $-L$, i.e. it holds

$$(-L)^\alpha I_n(f_n) = n^\alpha I_n(f_n), \qquad n \geq 1, \quad \alpha \in \mathbb{R}.$$

We also recall that, by Proposition 5.3 in Privault and Serafin (2022), for every $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$ there exists $\bar{f}_n \in \widehat{L}^2(\mathbb{R}_+^n)$ given by

$$\bar{f}_n(t_1, \ldots, t_n) = \Psi_{t_1} \cdots \Psi_{t_n} f_n(t_1, \ldots, t_n), \tag{2.14}$$

satisfying (2.6) and such that $I_n(f_n) = I_n(\bar{f}_n)$, where

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

$i = 1, \ldots, n$, $t_1, \ldots, t_n \in \mathbb{R}_+$. We end this section with the following multiplication formula for multiple stochastic integrals, see Proposition 5.1 in Privault and Serafin (2018). Letting

$n \wedge m := \min(n,m)$, for $0 \leq l \leq k \leq n \wedge m$ 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

$$\begin{aligned} & f_n \star_k^l g_m(y_1,\ldots,y_{n-l},z_1,\ldots,z_{m-k}) \\ & := \frac{1}{2^l}\int_{\mathbb{R}_+^l} f_n(x_1,\ldots,x_l,y_1,\ldots,y_{n-l}) g_m(x_1,\ldots,x_l,y_1,\ldots,y_{k-l},z_1,\ldots,z_{m-k}) dx_1\cdots dx_l, \end{aligned} \tag{2.15}$$

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

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

Then, for $f_n \in \widehat{L}^2(\mathbb{R}_+^n)$ and $g_m \in \widehat{L}^2(\mathbb{R}_+^m)$ satisfying (2.6), the following multiplication formula holds:

$$I_n(f_n)I_m(g_m) = \sum_{k=0}^{m\wedge n} k! \binom{m}{k}\binom{n}{k}\sum_{i=0}^{k}\binom{k}{i} I_{m+n-k-i}\big(f_n \widetilde{\star}_k^i g_m\big), \tag{2.16}$$

whenever $f_n \star_k^i g_m \in L^2(\mathbb{R}_+^{m+n-k-i})$ for every $0 \leq i \leq k \leq m \wedge n$.

2.2 Multiple stochastic integrals and Hoeffding decomposition

Although the multiple integrals (chaoses) seem a little abstract, they are in fact very well known objects. Namely, we can call them degenerate U-statistics. To explain the context, let us recall the definition of the Hoeffding decomposition.

Given $(X_1,\ldots,X_n)$ a family of independent random variables and $[n] := \{1,\ldots,n\}$, $n \geq 1$, the family $(\mathcal{F}_J)_{J\subset[n]}$ of σ-algebras is defined as

$$\mathcal{F}_J := \sigma(X_j \ : \ j \in J), \qquad J \subset [n].$$

Definition 2.3 *A centered $\mathcal{F}_{[n]}$-measurable random variable W_n admits a Hoeffding decomposition if it can be written as*

$$W_n = \sum_{J\subset[n]} W_J, \tag{2.17}$$

where $(W_J)_{J\subset[n]}$ is a family of random variables such that W_J is $\mathcal{F}_J$-measurable, $J \subset [n]$, and

$$\mathbb{E}\left[W_J \mid \mathcal{F}_K\right] = 0, \qquad J \not\subseteq K \subset [n].$$

If we take the sum over $|J| = d$ for a fixed $1 \leq d \leq n$, we call W_n a degenerate U-statistic of order d. In particular, for any U-statistic we may write

$$W_n = \sum_{d=1}^{n} W_n^{(d)}, \tag{2.18}$$

where $W_n^{(d)}$ are the degenerate U-statistics of order d

For $J = \{k_1, \ldots, k_{|J|}\}$ with $k_1 < k_2 < \cdots < k_{|J|}$, any W_J in Definition 2.3 can be written as a function $W_J = g_J\big(X_{k_1}, \ldots, X_{k_{|J|}}\big)$ of $\big(X_{k_1}, \ldots, X_{k_{|J|}}\big)$, with in particular

$$\mathbb{E}\left[g_J\left(X_j : j \in J\right) \mid \mathcal{F}_{J \setminus \{k\}}\right] = 0, \qquad k \in J, \tag{2.19}$$

and

$$W_n = \sum_{J \subset [n]} g_J\big(X_{k_1}, \ldots, X_{k_{|J|}}\big). \tag{2.20}$$

Note that if $X_i = U_i$, $i \in [n]$, then the chaos decomposition (2.11) coincides with the Hoeffding decomposition (2.17), by taking

$$W_J := \frac{1}{|J|!} f_{|J|}\big(2k_1 + 1 + U_1, \ldots, 2k_{|J|-1} + 1 + U_{|J|-1}, 2k_{|J|} + 1 + U_{|J|}\big), \quad J \subset [n],$$

and condition (2.19) is equivalent to (2.6). Furthermore, any sequence $(X_1, \ldots, X_n)$ of independent random variables with distribution functions $(F_{X_1}, \ldots, F_{X_n})$ is distributed 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 $\big(F_{X_1}^{-1}, \ldots, F_{X_n}^{-1}\big)$ are the generalized inverses of $(F_{X_1}, \ldots, F_{X_n})$. For instance, 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 centered random variables (degenerate U-statistic of order 1)

$$I_1(f_1) \stackrel{d}{=} \sum_{k=1}^{n} (X_k - \mathbb{E}[X_k]) \tag{2.21}$$

by taking $f_1(x) = F_{X_k}^{-1}((x + 2)/2 - k)$, $x \in [2k - 2, 2k)$, $1 \leq k \leq n$. Analogously, we may represent any degenerate U-statistic of order d as $I_d(f_d)$ for suitable function f_d and therefore the chaos decomposition (2.11) becomes the Hoeffding decomposition (2.18). For this reason, investigating the multiple stochastic integrals is very natural. It also explains the special attention we put on the sums of the multiple stochastic integrals, as it allows us to deal with any U-statistic in the most general sense.

2.3 Fourth moment bound

The main result of this subsection is the below-given fourth order moment bound stated in terms of the gradient operator ∇.

Proposition 2.4 *For any $X \in L^4(\Omega)$ we have*

$$\mathbb{E}\left[X^4\right] \leq 36\mathbb{E}[\|\nabla X\|^4_{L^2(\mathbb{R}_+)}] + 15\mathbb{E}[\|\nabla X\|^4_{L^4(\mathbb{R}_+)}] + 2\left(\mathbb{E}\left[X^2\right]\right)^2. \tag{2.22}$$

Before passing to the proof, we present a covariance relation, that can be obtained as in Proposition 2.1 in Houdré and Privault (2002) and also plays a crucial role in the proof of Theorem 3.1.

Lemma 2.5 *Let $\alpha \in \mathbb{R}$ and $X, Y \in L^2(\Omega)$ such that $L^{\alpha-1}X \in \mathrm{Dom}(\nabla)$ and $L^{-\alpha}Y \in \mathrm{Dom}(\nabla)$. Then we have the covariance relation*

$$\mathrm{Cov}\,(X,Y) = \mathbb{E}\left[\int_0^\infty (\nabla_t(-L)^{\alpha-1}X)(\nabla_t(-L)^{-\alpha}Y)\frac{dt}{2}\right]. \tag{2.23}$$

Proof. We have

$$\begin{aligned}
\mathrm{Cov}\,(X,Y) &= \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] \\
&= -\mathbb{E}[L(-L)^{\alpha-1}(X-\mathbb{E}[X])(-L)^{-\alpha}(Y-\mathbb{E}[Y])] \\
&= \mathbb{E}[\nabla^*\nabla(-L)^{\alpha-1}(X-\mathbb{E}[X])(-L)^{-\alpha}(Y-\mathbb{E}[Y])] \\
&= \frac{1}{2}\mathbb{E}\left[\int_0^\infty (\nabla_t(-L)^{\alpha-1}X)(\nabla_t(-L)^{-\alpha}Y)dt\right].
\end{aligned}$$

□

Proof of Proposition 2.3. By the covariance relation (2.23), we have

$$\begin{aligned}
\mathbb{E}\left[X^4\right] &= \mathrm{Var}\left[X^2\right] + \left(\mathbb{E}\left[X^2\right]\right)^2 \\
&= \frac{1}{2}\mathbb{E}\left[\int_0^\infty \nabla_t\left(X^2 - \mathbb{E}\left[X^2\right]\right)\nabla_t L^{-1}\left(X^2-\mathbb{E}\left[X^2\right]\right)dt\right] + \left(\mathbb{E}\left[X^2\right]\right)^2 \\
&\leq \frac{1}{2}\sqrt{\mathbb{E}\left[\int_0^\infty |\nabla_t(X^2)|^2 dt\right]\mathbb{E}\left[\int_0^\infty \|\nabla_t L^{-1}\left(X^2-\mathbb{E}\left[X^2\right]\right)\|^2 dt\right]} + \left(\mathbb{E}\left[X^2\right]\right)^2 \\
&\leq \frac{1}{2}\mathbb{E}\left[\int_0^\infty |\nabla_t(X^2)|^2 dt\right] + \left(\mathbb{E}\left[X^2\right]\right)^2,
\end{aligned}$$

where we applied (2.8), (2.12) and (2.13). Since

$$(X\circ\Phi_t)\nabla_t X = X\nabla_t X + (X\circ\Phi_t - X)\nabla_t X$$

$$= X\nabla_t X + \left(\nabla_t X - \left(X - \frac{1}{2}\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} X\circ\Phi_u du\right)\right)\nabla_t X,$$

by the relations (2.3)-(2.2) and the bound $(a+b+c)^2 \le 3\left(a^2+b^2+c^2\right)$, $a,b,c\ge 0$, we have

$$E\left[\int_0^\infty |\nabla_t(X^2)|^2 dt\right] = \mathbb{E}\left[\int_0^\infty \left(2(X\circ\Phi_t)\nabla_t X - \frac{1}{2}\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\left(|\nabla_t X|^2 + |\nabla_u X|^2\right)du\right)^2 dt\right]$$
$$\le 3\mathbb{E}\Bigg[4\int_0^\infty \left(X\nabla_t X\right)^2 dt + 4\int_0^\infty \left(\left(\nabla_t X - \left(X - \frac{1}{2}\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} X\circ\Phi_u du\right)\right)\nabla_t X\right)^2 dt$$
$$+ \frac{1}{4}\int_0^\infty \left(\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\left(|\nabla_t X|^2 + |\nabla_u X|^2\right)du\right)^2 dt\Bigg]$$
$$= 12\mathbb{E}\left[\int_0^\infty \left(X\nabla_t X\right)^2 dt\right]$$
$$+ 12\mathbb{E}\left[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\left(\left(\nabla_t X - \left(X\circ\Phi_v - \frac{1}{2}\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} X\circ\Phi_u du\right)\right)\nabla_t X\right)^2 \frac{dv}{2} dt\right]$$
$$+ \frac{3}{4}\mathbb{E}\left[\int_0^\infty \left(\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\left(|\nabla_t X|^2 + |\nabla_u X|^2\right)du\right)^2 dt\right]$$
$$= 12\mathbb{E}\left[X^2\int_0^\infty \left(\nabla_t X\right)^2 dt\right] + 12\mathbb{E}\left[\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\left(\left(\nabla_t X - \nabla_v X\right)\nabla_t X\right)^2 \frac{dv}{2}\, dt\right]$$
$$+ \frac{3}{4}\mathbb{E}\left[\int_0^\infty \left(\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\left(|\nabla_t X|^2 + |\nabla_u X|^2\right)du\right)^2 dt\right]$$
$$\le 12\sqrt{\mathbb{E}[X^4]\mathbb{E}\left[\left(\int_0^\infty |\nabla_t X|^2 dt\right)^2\right]} + 12\mathbb{E}\left[\int_0^\infty |\nabla_t X|^2 \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2}\left(|\nabla_t X|^2 + |\nabla_v X|^2\right)\frac{dv}{2}\, dt\right]$$
$$+ 3\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]$$
$$\le 12\sqrt{\mathbb{E}[X^4]\mathbb{E}\left[\left(\int_0^\infty |\nabla_t X|^2 dt\right)^2\right]} + 15\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right].$$

Thus, we get

$$\mathbb{E}\left[X^4\right] \le 6\sqrt{\mathbb{E}\left[X^4\right]\mathbb{E}\left[\left(\int_0^\infty |\nabla_t X|^2 dt\right)^2\right]} + \frac{15}{2}\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right] + \left(\mathbb{E}\left[X^2\right]\right)^2.$$

Denoting

$$a = \sqrt{\mathbb{E}\left[\left(\int_0^\infty |\nabla_t X|^2 dt\right)^2\right]}, \qquad b = \frac{15}{2}\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right] + \left(\mathbb{E}\left[X^2\right]\right)^2$$

and $x = \sqrt{\mathbb{E}[X^4]}$, we rewrite the last inequality as $x^2 \leq 6ax + b$, which gives $x \leq 3a + \sqrt{9a^2 + b}$ and consequently $x^2 \leq 2(9a^2 + b) + 18a^2 = 36a^2 + 2b$, which yields (2.22). $\square$

3 General results

3.1 Statements and discussion

Our main result is a Berry-Esseen bound on the Kolmogorov distance $d_K(X, \mathcal{N})$ between the standard normal distribution $\mathcal{N}$ on $\mathbb{R}$ and a general functional X of the uniform i.i.d. sequence $(U_k)_{k\in\mathbb{N}}$ on $[-1, 1]$, using the operators ∇ and L. This result extends Proposition 4.1 in Krokowski et al. (2017), see also Theorem 3.1 in Krokowski et al. (2016) and Proposition 2.1 in Privault and Serafin (2020), from functionals of Bernoulli sequences to more general functionals of independent random variables. We note that in comparison with Theorem 4.2 of Lachièze-Rey and Peccati (2017), which is obtained in a discrete-time integration by parts setting, the variance term (3.2) in Theorem 3.1 can be easier to control, in particular it vanishes when $X = I_1(f_1)$ is a first chaos random variable. Before stating our main result, let us mention that the Wasserstein distance has been approached in the framework of this paper in Privault and Serafin (2018) and Privault and Serafin (2022), which resulted in the bound (see Proposition 2.4 in Privault and Serafin (2022))

$$d_W(X, \mathcal{N}) \leq |1 - \mathbb{E}[X^2]| + \sqrt{\operatorname{Var}\left[\int_0^\infty \nabla_t X \nabla_t L^{-1} X \frac{dt}{2}\right]} \qquad (3.1)$$
$$+ 2\sqrt{\mathbb{E}[((-L)^{-1/2}X)^2] \int_0^\infty \mathbb{E}\left[|\nabla_t X|^4\right] \frac{dt}{2}}.$$

Below, we present an extension of (3.1) to the Kolmogorov distance. This general result will be specialized to sums of multiple stochastic integrals in the next two propositions. Those results will be applied to general and degenerate U-statistics in Sections 4 and 5.

Theorem 3.1 *Let $X \in \operatorname{Dom}(\nabla)$ be such that $\mathbb{E}[X] = 0$. We have*

$$d_K(X, \mathcal{N}) \leq |1 - \mathbb{E}[X^2]| + \sqrt{\operatorname{Var}\left[\int_0^\infty \nabla_t X \nabla_t L^{-1} X \frac{dt}{2}\right]} \qquad (3.2)$$
$$+ \frac{3}{2}\sqrt{\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]}\left(\left(\mathbb{E}\left[X^4\right] \mathbb{E}\left[\left(\int_0^\infty |\nabla_t L^{-1} X|^2 dt\right)^2\right]\right)^{1/4} + \frac{\sqrt{\pi}}{2}\sqrt{\mathbb{E}\left[((-L)^{-1/2}X)^2\right]}\right)$$
$$+ 4\left(\mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I} + 2(-L)^{1/2}\right)\left(|\nabla_t X|^2\right)\right)^2 dt\right] \mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I} + 2(-L)^{1/2}\right)\left((\nabla_t L^{-1} X)^2\right)\right)^2 dt\right]\right)^{1/4}.$$

Direct application of Theorem 3.1 might be quite cumbersome, however, this is rather typical in the area. One difficulty is estimation of the variance term, the other one is involvement of the operator L^{-1}. The next proposition applies Theorem 3.1 to sums of multiple stochastic integrals, which, as explained in Section 2.2, covers U-statistics in full generality. It extends Theorem 3.1 of Privault and Serafin (2020) from functionals of Bernoulli sequences to functionals of independent random variables, see also earlier results such as Proposition 3.7 in Nourdin and Peccati (2009) in the case of multiple Wiener integrals. In the sequel, we denote $d_{W/K}(X,Y) := \max\{d_W(X,Y), d_K(X,Y)\}$.

Proposition 3.2 *For any $X \in L^2(\Omega)$ written as a sum $X = \sum_{k=1}^d I_k(f_k)$ of multiple stochastic integrals where $f_k \in \widehat{L}^2(\mathbb{R}_+^k)$ satisfies (2.6), $k = 1, \ldots, d$, we have*

$$d_{W/K}(X,\mathcal{N}) \leq \mathbb{E}\left[\left|1 - \mathbb{E}[X^2]\right|\right] + C_d\sqrt{\sum_{0\leq l<i\leq d} \left\|f_i \star_i^l f_i\right\|^2_{L^2(\mathbb{R}_+^{i-l})} + \sum_{1\leq l<i\leq d}\left(\left\|f_i \star_l^l f_i\right\|^2_{L^2(\mathbb{R}_+^{2(i-l)})} + \left\|f_l \star_l^l f_i\right\|^2_{L^2(\mathbb{R}_+^{i-l})}\right)}, \tag{3.3}$$

for some $C_d > 0$.

We note that the constant C_d might be precisely calculated from the proof of Proposition 3.2. The simplest example of application of Proposition 3.2 to sum of multiple stochastic integrals is the quadratic form discussed in Section 5, which leads to the bound (5.8).

Next, due to the identity $\nabla_t L^{-1} I_d(f_d) = I_{d-1}(f_d(t,*))$, $d \geq 1$, the bound in Theorem 3.1 can be significantly simplified in the case of multiple stochastic integrals $I_d(f_d)$, which represent degenerate U-statistics.

Proposition 3.3 *For $X = I_d(f_d)$ a multiple stochastic integral of order $d \geq 1$, we have*

$$\begin{aligned} d_{W/K}(X,\mathcal{N}) &\leq |1 - \mathbb{E}[X^2]| + \frac{1}{d}\sqrt{\mathrm{Var}\left[\int_0^\infty (\nabla_t X)^2 \frac{dt}{2}\right]} + \frac{12 + 5\sqrt[4]{\mathbb{E}[X^4]}}{\sqrt{d}}\sqrt{\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]} \\ &\leq |1 - \mathbb{E}[X^2]| + \sqrt{\mathrm{Var}\left[\int_0^\infty (\nabla_t X)^2 dt\right]} + 31\sqrt{\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]}. \end{aligned} \tag{3.4}$$

The bound (3.4) has been obtained for the Wasserstein distance in Privault and Serafin (2022) with different constants, see Proposition 2.4 therein. As an example, recall that by (2.21), a sum of independent random variables $S_n = \sum_{k=1}^n X_k$ might be represented as a single stochastic integral $I_1(f_1)$ of f_1. Then, since the variance term in (3.4) vanishes for $X = I_1(f_1)$, for $\widetilde{S}_n = (S_n - \mathbb{E}[S_n])/\sqrt{\mathrm{Var}[S_n]}$ we get

$$\begin{aligned} d_{W/K}\big(\widetilde{S}_n, \mathcal{N}\big) &\leq \frac{31}{\sum_{k=1}^n \mathrm{Var}[X_k]} \sqrt{\sum_{k=1}^n \int_{-1}^1 \left(F_{X_k}^{-1}\left(\tfrac{x+1}{2}\right) - \mathbb{E}[X_k] \right)^4 dx} \\ &= \frac{31}{\sum_{k=1}^n \mathrm{Var}[X_k]} \sqrt{2 \sum_{k=1}^n \mathbb{E}\left[(X_k - \mathbb{E}[X_k])^4 \right]}, \end{aligned}$$

which provides a quantitative bound with explicit constant in the Kolmogorov distance for the L^4 Lyapunov Central Limit Theorem, and implies the fourth moment bound

$$d_{W/K}\big(\widetilde{S}_n, \mathcal{N}\big) \leq 31\sqrt{2}\sqrt{\big|\mathbb{E}\big[\widetilde{S}_n^4\big] - 3\big|}.$$

In order to formulate the bound (3.4) in a framework closer to e.g. Lachièze-Rey and Peccati (2017), let us assume that X as written as $X = f(U)$ with $U = (U_1, U_2, U_3, \ldots)$ and let

$$\Delta_j f(U, U') = f(U) - f\big(U_1, \ldots, U_{j-1}, U'_j, U_{j+1}, \ldots\big),$$

where $U' = (U'_1, U'_2, U'_3, \ldots)$ is an independent copy of U. Then, for $j \in \mathbf{N}$ and sufficiently integrable $h : \mathbf{R} \to \mathbf{R}$ we have

$$\mathbb{E}\left[\int_{2j}^{2j+2} h(\nabla_t f(U)) \frac{dt}{2} \right] = \mathbb{E}\big[h\big(\mathbb{E}[\Delta_j f(U, U') \mid U] \big) \big],$$

hence (3.4) can be rewritten as

$$\begin{aligned} d_{W/K}(X, \mathcal{N}) \leq & |1 - \mathbb{E}[X^2]| \\ & + 2\sqrt{\mathrm{Var}\left[\sum_{j=1}^\infty \mathbb{E}_j\big[\big(\mathbb{E}[\Delta_j f(U, U') | U] \big)^2 \big] \right]} + 31\sqrt{2}\sqrt{\sum_{j=1}^\infty \mathbb{E}\big[\big(\mathbb{E}\left[\Delta_j f(U, U') | U \right] \big)^4 \big]}, \end{aligned}$$

where by $\mathbb{E}_j$ we denote the expectation with respect to U_j only.

3.2 Proofs

Proof of Theorem 3.1. The beginning of the proof of Theorem 3.1 follows the general argument applied in the literature on the Stein method and the Malliavin calculus in discrete settings, see Peccati et al. (2010), Nourdin et al. (2010a), Peccati and Thäle (2013), Privault and Torrisi (2015), Krokowski et al. (2016), Krokowski et al. (2017). However, the rest of the proof presents significant differences as specific arguments are needed to bound the remainder terms using the Kolmogorov distance. For any $x \in \mathbb{R}$, let f_x denote the unique bounded solution of the Stein equation

$$f_x'(z) - z f_x(z) = \mathbb{1}_{\{z \le x\}} - \mathbb{P}(\mathcal{N} \le x), \tag{3.5}$$

which is continuous, infinitely differentiable on $\mathbb{R}\backslash\{x\}$, and satisfies $0 < f_x(z) < \sqrt{\pi/8}$ and $|f_x'(z)| \le 1$, $z \in \mathbb{R}\backslash\{x\}$, see Lemmas 2.2 and 2.3 in Chen et al. (2011). From the Stein equation (3.5) we have the bound

$$d_K(X, \mathcal{N}) \le \sup_{x \in \mathbb{R}} \mathbb{E}[f_x'(X) - X f_x(X)].$$

For every $f \in \mathcal{C}^1(\mathbb{R})$, the finite difference operator ∇ satisfies

$$\begin{aligned} \nabla_t f(X) &= \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} (f(X \circ \Phi_t) - f(X \circ \Phi_s)) ds \\ &= \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} \int_{X \circ \Phi_s - X}^{X \circ \Phi_t - X} f'(X+u) du ds \\ &= \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} \left(\int_{X \circ \Phi_s - X}^{X \circ \Phi_t - X} (f'(X+u) - f'(X)) du + \int_{X \circ \Phi_s - X}^{X \circ \Phi_t - X} f'(X) du \right) ds \\ &= f'(X) \nabla_t X + \frac{1}{2} \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} \int_{X \circ \Phi_s - X}^{X \circ \Phi_t - X} (f'(X+u) - f'(X)) du\, ds, \quad t \in \mathbb{R}_+. \end{aligned}$$

Hence by the duality relation (2.9), we have

$$\begin{aligned} &\mathbb{E}[f'(X) - X f(X)] = \mathbb{E}[f'(X) - f(X)(-\nabla^* \nabla) L^{-1} X] \\ &= \mathbb{E}\left[f'(X) - \frac{1}{2} \int_0^\infty \nabla_t f(X) (-\nabla_t L^{-1} X) dt \right] \\ &= \mathbb{E}\left[f'(X) \left(1 - \frac{1}{2} \int_0^\infty \nabla_t X (-\nabla_t L^{-1} X) dt \right) \right] \\ &\quad + \frac{1}{4} \mathbb{E}\left[\int_0^\infty \int_{2\lfloor t/2 \rfloor}^{2\lfloor t/2 \rfloor + 2} \int_{X \circ \Phi_s - X}^{X \circ \Phi_t - X} (f'(X+u) - f'(X)) du ds \nabla_t L^{-1} X dt \right]. \end{aligned} \tag{3.6}$$

By the covariance relation (2.23) applied with $\alpha = 0$ and the fact that $\mathbb{E}[X] = 0$, we have

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

hence from the bound $\|f_x'\|_\infty \le 1$ and Jensen's inequality we obtain

$$\begin{aligned} &\left| \mathbb{E}\left[f'(X) \left(1 - \frac{1}{2} \int_0^\infty \nabla_t X (-\nabla_t L^{-1} X) dt \right) \right] \right| \\ &\le \mathbb{E}\left[\left| 1 - \frac{1}{2} \int_0^\infty \nabla_t X (-\nabla_t L^{-1} X) dt \right| \right] \end{aligned}$$

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

$$\leq |1-\mathbb{E}[X^2]| + \mathrm{Var}\left[\int_0^\infty \nabla_t X(-\nabla_t L^{-1}X)\frac{dt}{2}\right]. \tag{3.7}$$

Next, from the Stein equation (3.5) we have

$$\int_{X\circ\Phi_s - X}^{X\circ\Phi_t - X} (f_x'(X+u) - f_x'(X))du = A_{s,t}(x,X) + B_{s,t}(x,X), \quad x \in \mathbb{R},$$

where

$$A_{s,t}(x,X) := \int_{X\circ\Phi_s - X}^{X\circ\Phi_t - X} ((X+u)f_x(X+u) - Xf_x(X))du$$

and

$$B_{s,t}(x,X) := \int_{X\circ\Phi_s - X}^{X\circ\Phi_t - X} \left(\mathbf{1}_{\{X+u\leq x\}} - \mathbf{1}_{\{X\leq x\}}\right)du.$$

Thus, applying this and (3.7) to (3.6), we get

$$\begin{aligned} |\mathbb{E}[f'(X) - Xf(X)]| \leq |1-\mathbb{E}[X^2]| &+ \mathrm{Var}\bigg[\int_0^\infty \nabla_t X(-\nabla_t L^{-1}X)\frac{dt}{2}\bigg] \\ &+ \frac{1}{4}\bigg|\mathbb{E}\bigg[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} A_{s,t}(x,X)ds\nabla_t L^{-1}X dt\bigg]\bigg| \\ &+ \frac{1}{4}\bigg|\mathbb{E}\bigg[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} B_{s,t}(x,X)ds\nabla_t L^{-1}X dt\bigg]\bigg|. \end{aligned} \tag{3.8}$$

Using the inequality

$$|(u+w)f_x(u+w) - wf_x(w)| \leq \left(|w| + \frac{\sqrt{2\pi}}{4}\right)|u|, \qquad u, w \in \mathbb{R},$$

see Lemma 2.3 in Chen et al. (2011), we estimate

$$\begin{aligned} |A_{s,t}(x,X)| &\leq \int_{\min(X\circ\Phi_s - X, X\circ\Phi_t - X)}^{\max(X\circ\Phi_s - X, X\circ\Phi_t - X)} \left(|X| + \frac{\sqrt{2\pi}}{4}\right)|u|du \\ &\leq \left(\frac{\sqrt{2\pi}}{4} + |X|\right)\int_{|X\circ\Phi_s - X|}^{|X\circ\Phi_t - X|} |u|du \\ &= \frac{1}{2}\left(\frac{\sqrt{2\pi}}{4} + |X|\right)\left(|X\circ\Phi_s - X|^2 + |X\circ\Phi_t - X|^2\right). \end{aligned}$$

Then, by the Cauchy-Schwarz inequality we have

$$\begin{aligned} &\bigg|\mathbb{E}\bigg[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} A_{s,t}(x,X)ds\nabla_t L^{-1}X dt\bigg]\bigg| \\ &\leq \frac{1}{2}\mathbb{E}\bigg[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(\frac{\sqrt{2\pi}}{4} + |X|\right)\left(|X\circ\Phi_t - X|^2 + |X\circ\Phi_s - X|^2\right)|\nabla_t L^{-1}X| ds\, dt\bigg] \end{aligned}$$

$$\leq \frac{1}{2}\sqrt{\mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(|X\circ\Phi_t - X|^2 + |X\circ\Phi_s - X|^2\right)^2 ds\, dt\Big]}$$
$$\times\left(\frac{\sqrt{\pi}}{2}\sqrt{\mathbb{E}\left[\int_0^\infty (\nabla_t L^{-1}X)^2 dt\right]} + \sqrt{2\mathbb{E}\left[\int_0^\infty (X\nabla_t L^{-1}X)^2 dt\right]}\right). \tag{3.9}$$

Next, by the inequality $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$, $a,b,c\geq 0$, formula (2.23) with $\alpha = 0$ and the relation (2.3), we get

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(|X\circ\Phi_t - X|^2 + |X\circ\Phi_s - X|^2\right)^2 ds\, dt\Big] \\
&= \frac{1}{2}\mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(|X\circ\Phi_t - X\circ\Phi_v|^2 + |X\circ\Phi_s - X\circ\Phi_v|^2\right)^2 dv\, ds\, dt\Big] \\
&= \frac{1}{2}\mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(|\nabla_t X - \nabla_v X|^2 + |\nabla_s X - \nabla_v X|^2\right)^2 dv\, ds\, dt\Big] \\
&\leq \mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(|\nabla_t X|^2 + |\nabla_s X|^2 + 2|\nabla_v X|^2\right)^2 dv\, ds\, dt\Big] \\
&\leq 3\mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} (\nabla_t X)^4 + (\nabla_s X)^4 + 4(\nabla_v X)^4 dv\, ds\, dt\Big] \\
&= 72\,\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right].
\end{aligned}$$

Furthermore, by Lemma 2.5 applied to X and $(-L)^{-1}X$ with $\alpha = 1/2$, we have

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

and

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

Applying the last three inequalities to (3.9), we finally obtain

$$\begin{aligned}
&\left|\mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} A_{s,t}(x,X) ds \nabla_t L^{-1}X dt\Big]\right| \\
&\leq 6\sqrt{\mathbb{E}\int_0^\infty (\nabla_t X)^4 dt}\left(\left(\mathbb{E}\left[X^4\right]\,\mathbb{E}\left[\left(\int_0^\infty (\nabla_t L^{-1}X)^2\right)^2 dt\right]\right)^{1/4} + \frac{\sqrt{\pi}}{2}\sqrt{\mathbb{E}\left[((-L)^{-1/2}X)^2\right]}\right).
\end{aligned}$$

Regarding the last term in (3.8), we use (2.3) and the equivalence $(\nabla_t L^{-1}X)\circ\Phi_v = (\nabla_t L^{-1}X)$, which is valid for $2\lfloor t/2\rfloor \leq v < 2\lfloor t/2\rfloor + 2$, and get

$$\left|\mathbb{E}\Big[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} B_{s,t}(x,X) ds \nabla_t L^{-1}X dt\Big]\right|$$

$$
\begin{aligned}
&= \left|\mathbb{E}\left[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \left(\int_{X\circ\Phi_s}^{X\circ\Phi_t} (\mathbf{1}_{\{u\le x\}} - \mathbf{1}_{\{X\le x\}}) du\right) ds \nabla_t L^{-1} X dt\right]\right| \\
&= \frac{1}{2}\left|\mathbb{E}\left[\int_0^\infty \left(\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \int_{X\circ\Phi_s}^{X\circ\Phi_t} (\mathbf{1}_{\{u\le x\}} - \mathbf{1}_{\{X\circ\Phi_v\le x\}}) du ds\, dv\right) \nabla_t L^{-1} X dt\right]\right| \\
&= \frac{1}{2}\left|\mathbb{E}\left[\sum_{m=0}^\infty \int_{2m}^{2m+2} K_m(t,X) \nabla_t L^{-1} X dt\right]\right|, \qquad (3.10)
\end{aligned}
$$

where

$$
K_m(t,x,X) := \int_{2m}^{2m+2} \int_{2m}^{2m+2} \int_{X\circ\Phi_s}^{X\circ\Phi_t} \big(\mathbf{1}_{\{u\le x\}} - \mathbf{1}_{\{X\circ\Phi_v\le x\}}\big) du\, ds\, dv, \quad 2m \le t < 2m+2.
$$

Next, we rewrite $K_m(t,X)$ as follows

$$
\begin{aligned}
K_m(t,x,X) &= \int_{2m}^{2m+2} \int_{X\circ\Phi_s}^{X\circ\Phi_t} \int_{2m}^{2m+2} \big(\mathbf{1}_{\{X\circ\Phi_t\le x\}} - \mathbf{1}_{\{X\circ\Phi_v\le x\}}\big) dv\, du\, ds \\
&\quad + \int_{2m}^{2m+2} \int_{X\circ\Phi_s}^{X\circ\Phi_t} \int_{2m}^{2m+2} \big(\mathbf{1}_{\{u\le x\}} - \mathbf{1}_{\{X\circ\Phi_t\le x\}}\big) dv\, du\, ds \\
&= 4\nabla_t X \nabla_t \mathbf{1}_{\{X\le x\}} + 2\int_{2m}^{2m+2} \int_{X\circ\Phi_s}^{X\circ\Phi_t} \big(\mathbf{1}_{\{u\le x\}} - \mathbf{1}_{\{X\circ\Phi_t\le x\}}\big)\, du\, ds \\
&= -4\nabla_t X \nabla_t \mathbf{1}_{\{X> x\}} + 2\int_{2m}^{2m+2} \int_{X\circ\Phi_s}^{X\circ\Phi_t} \big(\mathbf{1}_{\{u\le x\}} - \mathbf{1}_{\{X\circ\Phi_t\le x\}}\big)\, du\, ds, \qquad (3.11)
\end{aligned}
$$

where we used the equality $\nabla_t \mathbf{1}_{\{X\le x\}} = -\nabla_t \mathbf{1}_{\{X>x\}}$. Next, we consider two cases.
(i) If $X\circ\Phi_t > x$, we have

$$
\begin{aligned}
K_m(t,x,X) &= -4\nabla_t X \nabla_t \mathbf{1}_{\{X>x\}} + 2\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \int_{X\circ\Phi_s}^{X\circ\Phi_t} \mathbf{1}_{\{u\le x\}}\, du\, ds \\
&= -4\nabla_t X \nabla_t \mathbf{1}_{\{X>x\}} + 2\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \mathbf{1}_{\{X\circ\Phi_s\le x\}} (x - X\circ\Phi_s)\, ds. \qquad (3.12)
\end{aligned}
$$

Note that the last expression depends only on $m := \lfloor t/2\rfloor$ and may be bounded for $x < \int_{2m}^{2m+2} X\circ\Phi_u\, du/2$ as follows

$$
\begin{aligned}
0 &\le \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \mathbf{1}_{\{X\circ\Phi_s\le x\}} (x - X\circ\Phi_s)\, ds \\
&= x\int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_s\le x\}}\, ds - \int_{2m}^{2m+2} X\circ\Phi_u du + \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_u> x\}} X\circ\Phi_u du \\
&= \left(x - \frac{1}{2}\int_{2m}^{2m+2} X\circ\Phi_u\, du\right) \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_s\le x\}}\, ds \\
&\quad + \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_u> x\}} X\circ\Phi_u du - \frac{1}{2}\int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_s> x\}} ds \int_{2m}^{2m+2} X\circ\Phi_u du
\end{aligned}
$$

$$
\begin{aligned}
&\leq \int_{2m}^{2m+2} \left(\mathbf{1}_{\{X\circ\Phi_u > x\}} - \frac{1}{2}\int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_s > x\}} ds \right) X\circ\Phi_u du \\
&= \int_{2m}^{2m+2} \nabla_u \mathbf{1}_{\{X>x\}} X\circ\Phi_u du \\
&= \int_{2m}^{2m+2} \nabla_u \mathbf{1}_{\{X>x\}} \nabla_u X du,
\end{aligned}
$$

where we used (2.4) to obtain the last identity. Consequently, for $x < \int_{2m}^{2m+2} X\circ\Phi_u \, du/2$ we get

$$
\begin{aligned}
&\int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_t > x\}} K_m(t,X) \nabla_t L^{-1} X dt \\
&\leq 4\int_{2m}^{2m+2} \left| \nabla_t X \nabla_t \mathbf{1}_{\{X>x\}} \nabla_t L^{-1} X \right| dt \\
&\quad + 2\left| \int_{2m}^{2m+2} \nabla_u X \nabla_u \mathbf{1}_{\{X>x\}} du \right| \left| \int_{2m}^{2m+2} \nabla_t \mathbf{1}_{\{X>x\}} \nabla_t L^{-1} X dt \right|,
\end{aligned}
\tag{3.13}
$$

where we also changed $\mathbf{1}_{\{X\circ\Phi_t > x\}}$ into $\nabla_t \mathbf{1}_{\{X>x\}}$ in the last integral, which is justified by (2.4). In order to obtain the same bound in the case $x \geq \int_{2m}^{2m+2} X\circ\Phi_u \, du/2$, we rewrite (3.12) as

$$
\begin{aligned}
K_m(t,x,X) &= -4\nabla_t X \nabla_t \mathbf{1}_{\{X>x\}} + 2\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \mathbf{1}_{\{X\circ\Phi_s \leq x\}} (x - X\circ\Phi_t + \nabla_t X - \nabla_s X) \, ds \\
&= 2\int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} \mathbf{1}_{\{X\circ\Phi_s \leq x\}} (x - X\circ\Phi_t - \nabla_s X) \, ds \\
&= 2\int_{2m}^{2m+2} \nabla_s \mathbf{1}_{\{X>x\}} \nabla_s X ds - 2\int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_s \leq x\}} \left(X\circ\Phi_t - x \right) \, ds,
\end{aligned}
$$

and we estimate the last integral by

$$
\begin{aligned}
0 &\leq \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_s \leq x\}} \left(X\circ\Phi_t - x \right) \, ds \\
&\leq \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_s \leq x\}} \, ds \left(X\circ\Phi_t - \frac{1}{2}\int_{2m}^{2m+2} X\circ\Phi_u \, du \right) \\
&= -\nabla_t X \nabla_t \mathbf{1}_{\{X \leq x\}} = \nabla_t X \nabla_t \mathbf{1}_{\{X>x\}},
\end{aligned}
$$

which shows that the inequality (3.13) is valid for all $x \in \mathbb{R}$ under the condition $x < X\circ\Phi_t$. Thus, applying the Cauchy-Schwarz inequality several times and using the bound $|\nabla_t \mathbf{1}_{\{X\leq x\}}| \leq 1$, we obtain

$$
\left| \mathbb{E}\left[\sum_{m=0}^{\infty} \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_t > x\}} K_m(t,x,X) \nabla_t L^{-1} X dt \right] \right|
$$

$$\leq 4\sqrt{\mathbb{E}\left[\sum_{m=0}^{\infty}\int_{2m}^{2m+2}\left|\nabla_u \mathbf{1}_{\{X>x\}}\right| |\nabla_u X|^2 du\right] \mathbb{E}\left[\sum_{m=0}^{\infty}\int_{2m}^{2m+2}\left|\nabla_u \mathbf{1}_{\{X>x\}}\right| (\nabla_u L^{-1}X)^2 du\right]}$$
$$+ 4\sqrt{\mathbb{E}\left[\sum_{m=0}^{\infty}\int_{2m}^{2m+2}(\nabla_u \mathbf{1}_{\{X>x\}})^2 |\nabla_u X|^2 du\right] \mathbb{E}\left[\sum_{m=0}^{\infty}\int_{2m}^{2m+2}(\nabla_u \mathbf{1}_{\{X>x\}})^2 (\nabla_u L^{-1}X)^2 du\right]}$$
$$\leq 8\sqrt{\mathbb{E}\left[\int_0^{\infty}\left|\nabla_u \mathbf{1}_{\{X>x\}}\right| |\nabla_u X|^2 du\right] \mathbb{E}\left[\int_0^{\infty}\left|\nabla_u \mathbf{1}_{\{X>x\}}\right| (\nabla_u L^{-1}X)^2 du\right]}.$$

By the duality relation (2.9), Hölder's inequality and the formula (2.10), we get

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{\infty}\left|\nabla_u \mathbf{1}_{\{X>x\}}\right| |\nabla_u X|^2 du\right] \\
&= \mathbb{E}\left[\int_0^{\infty}\nabla_u \mathbf{1}_{\{X>x\}} \mathrm{sgn}(\nabla_u \mathbf{1}_{\{X>x\}}) |\nabla_u X|^2 du\right] \\
&= 2\mathbb{E}\left[\mathbf{1}_{\{X>x\}} \nabla^* \left(\mathrm{sgn}(\nabla_u \mathbf{1}_{\{X>x\}}) |\nabla_u X|^2\right)\right] \\
&\leq 2\sqrt{\mathbb{E}\left[\left(\nabla^*\left(\mathrm{sgn}(\nabla_u \mathbf{1}_{\{X>x\}}) |\nabla_u X|^2\right)\right)^2\right]} \\
&= 2\sqrt{\mathbb{E}\left[\int_0^{\infty}(\nabla_u X)^4 dt\right] + \mathbb{E}\left[\int_0^{\infty}\int_0^{\infty}\left(\nabla_s\left(\mathrm{sgn}(\nabla_u \mathbf{1}_{\{X>x\}}) |\nabla_u X|^2\right)\right)^2 ds\, du\right]}. \qquad (3.14)
\end{aligned}$$

Next, we observe that by the covariance relation (2.23) with $\alpha = \frac{1}{2}$, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{\infty}\int_0^{\infty}\left(\nabla_s\left(\mathrm{sgn}(\nabla_u \mathbf{1}_{\{X>x\}}) |\nabla_u X|^2\right)\right)^2 ds\, du\right] \\
&= \mathbb{E}\left[\int_0^{\infty}\mathbf{1}_{\{\nabla_u \mathbf{1}_{\{X>x\}}>0\}}\int_0^{\infty}\left(\nabla_s\left(|\nabla_u X|^2\right)\right)^2 ds + \mathbf{1}_{\{\nabla_u \mathbf{1}_{\{X>x\}}<0\}}\int_0^{\infty}\left(\nabla_s\left(-|\nabla_u X|^2\right)\right)^2 ds\, du\right] \\
&\leq \mathbb{E}\left[\int_0^{\infty}\int_0^{\infty}\left(\nabla_s\left(|\nabla_u X|^2\right)\right)^2 ds\, du\right] \\
&= 2\mathbb{E}\left[\int_0^{\infty}\left((-L)^{1/2}\left(|\nabla_u X|^2\right)\right)^2 du\right].
\end{aligned}$$

Applying this to (3.14), we get

$$\mathbb{E}\left[\int_0^{\infty}\left|\nabla_u \mathbf{1}_{\{X>x\}}\right| |\nabla_u X|^2 du\right] \leq 2\sqrt{\mathbb{E}\left[\int_0^{\infty}\left((\mathrm{I}+2(-L)^{1/2})\left(|\nabla_u X|^2\right)\right)^2 du\right]},$$

and analogously we obtain

$$\mathbb{E}\left[\int_0^{\infty}\left|\nabla_u \mathbf{1}_{\{X>x\}}\right| (\nabla_u L^{-1}X)^2 du\right] \leq 2\sqrt{\mathbb{E}\left[\int_0^{\infty}\left((\mathrm{I}+2(-L)^{1/2})\left((\nabla_u L^{-1}X)^2\right)\right)^2 du\right]},$$

which eventually gives us

$$\left|\mathbb{E}\left[\sum_{m=0}^{\infty}\int_{2m}^{2m+2}\mathbf{1}_{\{X\circ\Phi_t>x\}} K_m(t,x,X) \nabla_t L^{-1}X dt\right]\right| \qquad (3.15)$$

$$\le 16\Bigg(\mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left(|\nabla_u X|^2\right)\right)^2 du\right]$$
$$\times\, \mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I}+2(-L)^{1/2}\right)^{1/2}\left((\nabla_u L^{-1}X)^2\right)\right)^2 du\right]\Bigg)^{1/4}. \quad (3.16)$$

(ii) In case $X\circ\Phi_t \le x$ we observe that, denoting

$$\widetilde{K}_m(t,x,X) := -4\nabla_t X \nabla_t \mathbf{1}_{\{X\ge x\}} + 2\int_{2m}^{2m+2}\int_{X\circ\Phi_s}^{X\circ\Phi_t}\left(\mathbf{1}_{\{u<x\}} - \mathbf{1}_{\{X\circ\Phi_t<x\}}\right) du\, ds,$$

which comes from (3.11) by changing weak inequalities into strict ones and conversely, and repeating all the above argument, we arrive at

$$\left|\mathbb{E}\left[\sum_{m=0}^\infty \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_t\ge x\}}\widetilde{K}_m(t,x,X)\nabla_t L^{-1}X dt\right]\right| \quad (3.17)$$
$$\le 16\left(\mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left(|\nabla_u X|^2\right)\right)^2 du\right]\mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left((\nabla_u L^{-1}X)^2\right)\right)^2 du\right]\right)^{1/4}.$$

Next, by (3.11) we have, for $m=\lfloor t/2\rfloor$ and $X\circ\Phi_t\le x$,

$$\begin{aligned}
K_m(t,x,X) &= -4\nabla_t X\nabla_t \mathbf{1}_{\{X>x\}} + 2\int_{2m}^{2m+2}\int_{X\circ\Phi_s}^{X\circ\Phi_t}\left(\mathbf{1}_{\{u\le x\}} - \mathbf{1}_{\{X\circ\Phi_t\le x\}}\right) du\, ds\\
&= 4\nabla_t X\nabla_t \mathbf{1}_{\{X\le x\}} - \int_{2m}^{2m+2}\int_{X\circ\Phi_s}^{X\circ\Phi_t}\left(\mathbf{1}_{\{u\ge x\}} - \mathbf{1}_{\{X\circ\Phi_t\ge x\}}\right) du\, ds\\
&= 4\nabla_t X\nabla_t \mathbf{1}_{\{-X\ge -x\}} - \int_{2m}^{2m+2}\int_{X\circ\Phi_s}^{X\circ\Phi_t}\left(\mathbf{1}_{\{-u\le -x\}} - \mathbf{1}_{\{-X\circ\Phi_t\le -x\}}\right) du\, ds\\
&= -4\nabla_t(-X)\nabla_t \mathbf{1}_{\{-X\ge -x\}} + \int_{2m}^{2m+2}\int_{-X\circ\Phi_s}^{-X\circ\Phi_t}\left(\mathbf{1}_{\{u\le -x\}} - \mathbf{1}_{\{-X\circ\Phi_t\le -x\}}\right) du\, ds\\
&= \widetilde{K}_m(t,-x,-X).
\end{aligned}$$

Thus, using (3.17) with $-x$ and $-X$ instead of x and X respectively, we get

$$\left|\mathbb{E}\left[\sum_{m=0}^\infty \int_{2m}^{2m+2} \mathbf{1}_{\{X\circ\Phi_t\le x\}}K_m(t,x,X)\nabla_t L^{-1}X dt\right]\right|$$
$$= \left|\mathbb{E}\left[\sum_{m=0}^\infty \int_{2m}^{2m+2} \mathbf{1}_{\{-X\circ\Phi_t\ge -x\}}\widetilde{K}_m(t,-x,-X)\nabla_t L^{-1}(-X) dt\right]\right|$$
$$\le 16\left(\mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left(|\nabla_u X|^2\right)\right)^2 du\right]\mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left((\nabla_u L^{-1}X)^2\right)\right)^2 du\right]\right)^{1/4}. \quad (3.18)$$

Combining (3.15) and (3.18) with (3.10), we finally obtain

$$\frac{1}{4}\left|\mathbb{E}\left[\int_0^\infty \int_{2\lfloor t/2\rfloor}^{2\lfloor t/2\rfloor+2} B_{s,t}(x,X)ds\nabla_t L^{-1}X dt\right]\right|$$

$$
= \frac{1}{8}\left|\mathbb{E}\left[\sum_{m=0}^{\infty}\int_{2m}^{2m+2}\left(\mathbf{1}_{\{X\circ\Phi_t>x\}}+\mathbf{1}_{\{X\circ\Phi_t\le x\}}\right)K_m(t,X)\nabla_t L^{-1}X dt\right]\right|
$$

$$
\le 4\left(\mathbb{E}\left[\int_0^{\infty}\left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left(|\nabla_u X|^2\right)\right)^2 du\right]\mathbb{E}\left[\int_0^{\infty}\left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left((\nabla_u L^{-1}X)^2\right)\right)^2 du\right]\right)^{1/4},
$$

which ends the proof. □

Proof of Proposition 3.2. The bound for Wasserstein distance has been derived in Theorem 3.2 in Privault and Serafin (2022), so we will focus of the Kolmogorov distance. Since $|\nabla_t X|^2$ and $(\nabla_t L^{-1}X)^2$ are sums of multiple integrals of orders $2d-2$ and below, the relation (2.12) shows the bound

$$
\mathbb{E}\big[\left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left(|\nabla_t X|^2\right)\right)^2\big] \le 2d\mathbb{E}\left[(\nabla_t X)^4\right],
$$

and

$$
\mathbb{E}\big[\left(\left(\mathrm{I}+2(-L)^{1/2}\right)\left((\nabla_t L^{-1}X)^2\right)\right)^2\big] \le 2d\mathbb{E}\big[\left(\nabla_t L^{-1}X\right)^4\big].
$$

Additionally, by (2.12) we also have

$$
\mathbb{E}\left[((-L)^{-1/2}X)^2\right] \le \mathbb{E}\left[X^2\right] \le \sqrt{\mathbb{E}\left[X^4\right]}.
$$

Applying these inequalities to (3.2) in Theorem 3.1, we get

$$
\begin{aligned}
d_K(X,\mathcal{N}) \le\ & |1-\mathbb{E}[X^2]| + \sqrt{\mathrm{Var}\left[\int_0^{\infty}\nabla_t X\nabla_t L^{-1}X dt\right]} \\
& + \frac{3}{2}\left(\mathbb{E}\left[X^4\right]\right)^{1/4}\sqrt{\mathbb{E}\int_0^{\infty}(\nabla_t X)^4 dt}\left(1+\left(\mathbb{E}\left[\left(\int_0^{\infty}(\nabla_t L^{-1}X)^2 dt\right)^2\right]\right)^{1/4}\right) \\
& + 6d\sqrt{\mathbb{E}\left[\int_0^{\infty}(\nabla_t X)^4 dt\right]\mathbb{E}\left[\int_0^{\infty}(\nabla_t L^{-1}X)^4 dt\right]}.
\end{aligned}
$$

Denoting

$$
R_X := \sum_{1\le i\le j\le d}\sum_{k=1}^{i}\sum_{l=0}^{k}\mathbb{1}_{\{i=j=k=l\}^c}\left\|f_i\star_k^l f_j\right\|^2_{\widehat{L}^2(\mathbb{R}_+^{i+j-k-l})},
$$

it follows from the proof of Theorem 3.2 in Privault and Serafin (2022) that

$$
R_X \le c_d\left(\sum_{0\le l<i\le d}\left\|f_i\star_i^l f_i\right\|^2_{L^2(\mathbb{R}_+^{i-l})}+\sum_{1\le l<i\le d}\left(\left\|f_i\star_l^l f_i\right\|^2_{L^2(\mathbb{R}_+^{2(i-l)})}+\left\|f_l\star_l^l f_i\right\|^2_{L^2(\mathbb{R}_+^{i-l})}\right)\right), \tag{3.19}
$$

and

$$\mathrm{Var}\left[\int_0^\infty \nabla_t X \nabla_t L^{-1} X dt\right] \leq c_d R_X, \qquad \mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right] \leq c_d R_X, \tag{3.20}$$

for some $c_d \geq 0$. Taking $L^{-1}X$ as X in the last inequality, we also have

$$\mathbb{E}\left[\int_0^\infty (\nabla_t L^{-1}X)^4 dt\right] \leq c_d' R_X,$$

for some $c_d' \geq 0$. Furthermore, since

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

and the functions f_k satisfy (2.6), the multiplication formula (2.16) gives

$$\begin{aligned}\int_0^\infty (\nabla_t L^{-1}X)^2 dt &= \int_0^\infty \sum_{0\leq i\leq j<d-1} \sum_{k=0}^{i}\sum_{l=0}^{k} c_{i,j,l,k} I_{i+j-k-l}\big(f_{i+1}(t,\cdot)\,\widetilde{\star}_k^l f_{j+1}(t,\cdot)\big)dt \\ &= \sum_{0\leq i\leq j<d-1} \sum_{k=0}^{i}\sum_{l=0}^{k} c_{i,j,l,k} I_{i+j-k-l}\left(\int_0^\infty f_{i+1}(t,\cdot)\,\widetilde{\star}_k^l f_{j+1}(t,\cdot)dt\right)\end{aligned}$$

for some $c_{i,j,l,k} \geq 0$, and consequently

$$\begin{aligned}&\mathbb{E}\left[\left(\int_0^\infty (\nabla_t L^{-1}X)^2 dt\right)^2\right] \leq c_d \sum_{0\leq i\leq j<d} \sum_{k=0}^{i}\sum_{l=0}^{k} \left\|\left(\int_0^\infty f_{i+1}(t,\cdot)\star_k^l f_{j+1}(t,\cdot)dt\right)\right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})} \\ &= c_d \sum_{0\leq i\leq j<d} \sum_{k=0}^{i}\sum_{l=0}^{k} \left\|\left(f_{i+1}\star_{k+1}^{l+1} f_{j+1}\right)\right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})} \\ &= c_d \left(\sum_{1\leq i\leq j<d} \sum_{k=1}^{i}\sum_{1=0}^{k} \mathbb{1}_{\{i=j=k=l\}^c} \left\|\left(f_i\star_k^l f_j\right)\right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})} + \sum_{i=1}^{d}\left(f_i\star_i^i f_i\right)^2\right) \\ &\leq c_d\left(R_X + (\mathbb{E}[X^2])^2\right).\end{aligned}$$

Similarly, we get for some $C_{i,j,k,l} \geq 0$

$$\begin{aligned}\mathbb{E}\left[X^4\right] &\leq c_d \mathbb{E}\left[\left(\sum_{0\leq i\leq j<d} \sum_{k=0}^{i}\sum_{l=0}^{k} C_{i,j,l,k} I_{i+j-k-l}\left(f_i\,\widetilde{\star}_k^l f_j\right)\right)^2\right] \\ &\leq c_d \sum_{0\leq i\leq j<d} \sum_{k=0}^{i}\sum_{l=0}^{k} \left\|f_i\star_k^l f_j\right\|^2_{L^2(\mathbb{R}_+^{i+j-k-l})} \\ &= c_d\left(R_X + \sum_{i=1}^{d}\left(f_i\star_i^i f_i\right)^2 + \sum_{1\leq i\leq j\leq d}\left\|f_i\star_0^0 f_j\right\|^2_{L^2(\mathbb{R}_+^{i+j})}\right)\end{aligned}$$

$$
\begin{aligned}
&= c_d\left(R_X + \sum_{i=1}^{d} \|f_i\|^4_{L^2(\mathbb{R}^i)} + \sum_{1\le i\le j\le d} \|f_i\|^2_{L^2(\mathbb{R}^i)}\,\|f_j\|^2_{L^2(\mathbb{R}^j)}\right)\\
&\le c_d\left(R_X + (\mathbb{E}[X^2])^2\right).
\end{aligned}
$$

This finally gives us

$$
d_K(X,\mathcal{N}) \le |1-\mathbb{E}[X^2]| + c_d\sqrt{R_X}\Big(1 + \left((R_X +\ E[X^2])^{1/4}+1\right)^2\Big).
$$

Since $d_K(X,\mathcal{N}) \le 1$, we may assume that $\mathbb{E}[X^2]$ and R_X are bounded, which implies

$$
d_K(X,\mathcal{N}) \le |1-\mathbb{E}[X^2]| + c_d\sqrt{R_X},
$$

and the assertion of the corollary follows from (3.19). □

Proof of Proposition 3.3. First, let us observe that we have

$$
(-L)^{1/2} I_d(f_d) = \frac{1}{\sqrt{d}} I_d(f_d),
$$

and, by the covariance identity (2.23) applied with $\alpha = 0$,

$$
\mathbb{E}[X^2] = \frac{1}{d}\mathbb{E}\left[\int_0^\infty |\nabla_t X|^2 dt\right]. \tag{3.21}
$$

Then, Theorem 3.1 and the bound (3.1) give us

$$
\begin{aligned}
&d_{W/K}(X,\mathcal{N}) \le |1-\mathbb{E}[X^2]| + \frac{1}{d}\sqrt{\operatorname{Var}\left[\int_0^\infty |\nabla_t X|^2 \frac{dt}{2}\right]}\\
&+ \frac{3}{2\sqrt{d}}\sqrt{\mathbb{E}\int_0^\infty (\nabla_t X)^4 dt}\left(\left(\mathbb{E}\left[X^4\right]\left(\frac{1}{d^2}\operatorname{Var}\left[\int_0^\infty |\nabla_t X|^2 dt\right] + 4(\mathbb{E}[X^2])^2\right)\right)^{1/4} + \frac{\sqrt{\pi}}{2}\sqrt{\mathbb{E}\left[X^2\right]}\right)\\
&+ \frac{4}{d}\left(\mathbb{E}\left[\int_0^\infty \left(\left(\mathrm{I} + 2(-L)^{1/2}\right)\left(|\nabla_t X|^2\right)\right)^2 dt\right]\right)^{1/2}.
\end{aligned}
$$

Since $d_K(X,\mathcal{N}) \le 1$ by definition, we may assume that $\sqrt{\operatorname{Var}\left[\int_0^\infty |\nabla_t X|^2 dt/2\right]} \le d$ and $\mathbb{E}\left[X^2\right] \le 2$. Hence we get

$$
\begin{aligned}
&\left(\mathbb{E}\left[X^4\right]\left(\frac{1}{d^2}\operatorname{Var}\left[\int_0^\infty |\nabla_t X|^2 dt\right] + 4(\mathbb{E}[X^2])^2\right)\right)^{1/4} + \frac{\sqrt{\pi}}{2}\sqrt{\mathbb{E}\left[X^2\right]}\\
&\qquad \le \sqrt[4]{\mathbb{E}\left[X^4\right]}\left(\sqrt[4]{18} + \frac{\sqrt{\pi}}{2}\right) \le \frac{10}{3}\sqrt[4]{\mathbb{E}\left[X^4\right]}.
\end{aligned}
$$

Furthermore, since $|\nabla_u X|^2$ is a sum of multiple integrals of orders $2d-2$ and below, we have by (2.12)

$$\mathbb{E}\big[\big((2(-L)^{1/2}+\mathrm{I})(|\nabla_t X|^2)\big)^2\big] \le \left(2\sqrt{2d-2}+1\right)^2 \mathbb{E}\left[(\nabla_t X)^4\right] \le 9d\,\mathbb{E}\left[(\nabla_t X)^4\right].$$

Combining all together we obtain the first inequality from the assertion. Next, applying Proposition 2.4 and enlarging some constants, we get

$$\begin{aligned}
&d_{W/K}(X,\mathcal{N})\\
&\le |1-\mathbb{E}[X^2]| + \sqrt{\mathrm{Var}\left[\int_0^\infty |\nabla_t X|^2 dt\right]} + \sqrt{\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]}\\
&\quad \times \left(12+\frac{5}{\sqrt{d}}\left(\left(36\mathbb{E}\left[\left(\int_0^\infty |\nabla_t X|^2 dt\right)^2\right] + 15\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]\right) + 2\left(\mathbb{E}[X^2]\right)^2\right)^{1/4}\right).
\end{aligned}$$

Using once again the inequality $d_K(X,\mathcal{N}) \le 1$, we may assume $\sqrt{\mathrm{Var}\left[\int_0^\infty |\nabla_t X|^2 dt\right]} \le 1$, $\sqrt{\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]} \le \frac{1}{17}$ and $\mathbb{E}[X^2] \le 2$. Employing additionally (3.21), we get

$$\begin{aligned}
&12+\frac{5}{\sqrt{d}}\left(\left(36\mathbb{E}\left[\left(\int_0^\infty |\nabla_t X|^2 dt\right)^2\right] + 15\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]\right) + 2\left(\mathbb{E}[X^2]\right)^2\right)^{1/4}\\
&= 12+\frac{5}{\sqrt{d}}\left(\left(36\mathrm{Var}\left[\left(\int_0^\infty |\nabla_t X|^2 dt\right)^2\right] + 15\mathbb{E}\left[\int_0^\infty (\nabla_t X)^4 dt\right]\right) + (2+36d^2)\left(\mathbb{E}[X^2]\right)^2\right)^{1/4}\\
&\le 12+5\left(36+\frac{15}{17}+4(2+36)\right)^{1/4} \le 12+5\sqrt[4]{189} < 31,
\end{aligned}$$

which ends the proof. □

4 Applications to U-statistics

4.1 General U-statistics

The next Theorem 4.1 is a consequence of Proposition 3.2, using the fact that any random variable can be represented in distribution as a function of a uniformly distributed random variable, and makes more precise the central limit theorem of de Jong (1987; 1990). In comparison with Theorem 1.3 in Döbler and Peccati (2017), see also Theorem 3.7 in Döbler (2020), Theorem 4.1 is stated for the Kolmogorov distance instead of the Wasserstein distance, it applies to Hoeffding decompositions in full generality and not only to degenerate

U-statistics for which $|J|$ is constrained to a fixed value $|J| = d$ for some $d \in \{1, \ldots, n\}$ in the sum (2.17).

Theorem 4.1 *Let $1 \leq d \leq n$. For any $W_n \in L^4(\Omega)$ admitting the Hoeffding decomposition (2.17) with $|J| \leq d$, and such that $\mathbb{E}[W_n^2] = 1$, we have*

$$d_{W/K}(W_n, \mathcal{N}) \leq C_d \left(\sum_{0 \leq l < i \leq d} \sum_{|J|=i-l} \mathbb{E}\left[\left(\sum_{|K|=l, K \cap J = \phi} \mathbb{E}\big[(W_{J \cup K})^2 \mid \mathcal{F}_J \big] \right)^2 \right] \right.$$
$$+ \sum_{1 \leq l < i \leq d} \sum_{\substack{|J_1|=|J_2|=i-l \\ J_1 \cap J_2 = \phi}} \mathbb{E}\left[\left(\sum_{|K|=l, K_1 \cap (J_1 \cup J_2) = \phi} \mathbb{E}\big[W_{J_1 \cup K} W_{J_2 \cup K} \mid \mathcal{F}_{J_1 \cup J_2} \big] \right)^2 \right]$$
$$\left. + \sum_{1 \leq l < i \leq d} \sum_{|J|=i-l} \mathbb{E}\left[\left(\sum_{|K|=l, K \cap J = \phi} \mathbb{E}\big[W_K W_{J \cup K} \mid \mathcal{F}_J \big] \right)^2 \right] \right)^{1/2}, \qquad (4.1)$$

where $C_d > 0$ depends only on d.

Proof. By representing X_i as $X_i \stackrel{d}{=} F_i^{-1}((U_i + 1)/2)$ where F_i^{-1} is the generalized inverse of the cumulative distribution function F_i of X_i, $i = 1, \ldots, n$, we rewrite (2.20) as the sum of multiple stochastic integrals

$$W_n \stackrel{d}{=} \sum_{k=1}^{d} I_k(f_k),$$

where

$$f_k(x_1, \ldots, x_k) := \qquad (4.2)$$
$$\frac{1}{k!} \sum_{J=\{i_1,\ldots,i_k\} \subset [n]} g_J \left(F_{i_1}^{-1} \left(\frac{x_1}{2} - \left\lfloor \frac{x_1}{2} \right\rfloor \right), \ldots, F_{i_k}^{-1} \left(\frac{x_k}{2} - \left\lfloor \frac{x_k}{2} \right\rfloor \right) \right) \mathbf{1}_{[2i_1-2, 2i_1) \times \cdots \times [2i_k - 2, 2i_k)}(x_1, \ldots, x_k),$$

$(x_1, \ldots, x_k) \in \mathbb{R}_+^k$. Next, denoting

$$\widehat{\mathbf{N}}^m := \big\{ (k_1, \ldots, k_m) \ : \ k_1, \ldots, k_m \geq 1, \ k_i \neq k_j \text{ if } i \neq j, \ 1 \leq i, j \leq m \big\},$$

we have

$$\|f_i \star_i^l f_i\|^2_{L^2(\mathbb{R}_+^{i-l})}$$
$$= \frac{1}{2^{2l}} \sum_{\mathbf{j} \in \widehat{\mathbf{N}}^{i-l}} \int_{[2j_1-2, 2j_1) \times \cdots \times [2j_{i-l}-2, 2j_{i-l})} \left(\sum_{\mathbf{k} \in \widehat{\mathbf{N}}^l} \int_{[2k_1-2, 2k_1) \times \cdots \times [2k_l - 2, 2_l)} (f_i(x_1, \ldots, x_i))^2 dx_1 \cdots dx_l \right)^2$$

$$dx_{l+1}\cdots dx_i$$

$$\leq (i-l)!(l!)^2 \sum_{\substack{|J|=i-l\\ J=\{j_1,\ldots,j_{i-l}\}}} \int_{[2j_1,2j_1+2)\times\cdots\times[2j_{i-l},2j_{i-l}+2)}$$

$$\left(\sum_{\substack{|K|=l\\ K=\{k_1,\ldots,k_l\}}} \int_{[2k_1,2k_1+2)\times\cdots\times[2k_l,2_l+2)} \big(f_i(x_1,\ldots,x_i)\big)^2 dx_1\cdots dx_l \right)^2 dx_{l+1}\cdots dx_i$$

$$\leq C \sum_{|J|=i-l} \mathbb{E}\left[\left(\sum_{|K|=l,K\cap J=\phi} \mathbb{E}\big[\,(W_{J\cup K})^2 \mid \mathcal{F}_J\big]\right)^2\right],$$

for some $C = C(d)$. Similarly, we get

$$\big\|f_i \star_l^l f_i\big\|^2_{L^2(\mathbb{R}_+^{2(i-l)})} \leq C \sum_{\substack{|J_1|=|J_2|=l-i\\ J_1\cap J_2=\phi}} \mathbb{E}\left[\left(\sum_{|K|=l,K\cap(J_1\cup J_2)=\phi} \mathbb{E}\big[W_{J_1\cup K}W_{J_2\cup K} \mid \mathcal{F}_{J_1\cup J_2}\big]\right)^2\right]$$

and

$$\big\|f_l \star_l^l f_i\big\|^2_{L^2(\mathbb{R}_+^{i-l})} \leq C \sum_{|J|=i-l} \mathbb{E}\left[\left(\sum_{|K|=l,K\cap J=\phi} \mathbb{E}\big[W_K W_{J\cup K} \mid \mathcal{F}_J\big]\right)^2\right], \quad 1\leq l<i\leq d.$$

We conclude by applying the above to Proposition 3.2, which yields the required bound.

□

4.2 Degenerate U-statistics

In this section we narrow our attention to the degenerate U-statistics of a given order $d \geq 1$, which are random variables $W_{n,d}$ admitting the Hoeffding decomposition (2.17) with $|J| = d$.

Theorem 4.2 *For any degenerate U-statistics $W_{n,d} \in L^4(\Omega)$ of order $d \geq 1$, and such that $\mathbb{E}\big[W_{n,d}^2\big] = 1$, we have*

$$d_{W/K}(W_{n,d},\mathcal{N}) \leq \sqrt{\mathrm{Var}\left[\sum_{k=1}^{\infty} \mathbb{E}\big[\,\big(W_{n,d} - \mathbb{E}\big[W_{n,d}|\{X_k\}^c\big]\big)^2 \;|\{X_k\}^c\big]\right]}$$

$$+ 24\sqrt{2\,\mathbb{E}\sum_{k=1}^{\infty} \mathbb{E}\left[\big(W_{n,d} - \mathbb{E}\big[W_{n,d}|\{X_k\}^c\big]\big)^4\right]}$$

$$\leq C_d \left(\sum_{0\leq l<d} \sum_{|J|=d-l} \mathbb{E}\left[\left(\sum_{|K|=l,K\cap J=\phi} \mathbb{E}\big[\,(W_{J\cup K})^2 \mid \mathcal{F}_J\big]\right)^2\right]\right.$$

$$+\sum_{1\le l<d}\sum_{\substack{|J_1|=|J_2|=d-l\\J_1\cap J_2=\phi}}\mathbb{E}\left[\left(\sum_{|K|=l,K_1\cap(J_1\cup J_2)=\phi}\mathbb{E}\left[W_{J_1\cup K}W_{J_2\cup K}\mid\mathcal{F}_{J_1\cup J_2}\right]\right)^2\right]\Bigg)^{1/2},$$

where $\{X_k\}^c=\{X_1,\dots,X_{k-1},X_{k+1},\dots,X_n\}$ *and* $C_d>0$ *depends only on* d.

Proof. The first bound is just the latter bound from Proposition 3.3 rewritten in a different form. Namely, it is enough to take f_d as in (4.2) and then we have for $t\in[2k,2k+2)$

$$\nabla_t W_{n,d}=\mathbb{E}\left[W_{n,d}\mid\{X_k\}^c,X_k=t\right]-\mathbb{E}\left[W_{n,d}\mid\{X_k\}^c\right].$$

The other bound in the assertion follows from Proposition 3.3 in view of (3.20), (3.19) – where the last sum is vanishing – and the proof of Theorem 4.1. □

Weighted U-statistics

As an example, we consider degenerate weighted U-statistics. Precisely, given $(X_1,\dots,X_n)$ an i.i.d. sequence of random variables with distribution ν, we define

$$U_{n,d}=\binom{n}{d}^{-1}\sum_{1\le k_1<\cdots<k_d\le n}w(k_1,\dots,k_d)g\left(X_{k_1},\dots,X_{k_d}\right),\qquad 1\le d\le n,\tag{4.3}$$

where $w(k_1,\dots,k_d)\in\mathbb{R}$ is symmetric and vanishes on diagonals, and $g\left(x_{k_1},\dots,x_{k_d}\right)\in L^2(\mathbb{R}_+^d,\nu^{\otimes d})$, $1\le k_1<\cdots<k_d\le n$, is symmetric and satisfies

$$\mathbb{E}\left[g\left(X_1,x_2,\dots x_d\right)\right]=0,\qquad (x_2,\dots,x_d)\in\mathbb{R}^{d-1}.\tag{4.4}$$

The variance σ^2 of $U_{n,d}$ is given by

$$\sigma^2:=\mathrm{Var}[U_{n,d}]=\binom{n}{d}^{-2}\|g\|^2_{L^2(\mathbb{R}^d,\nu^{\otimes d})}\sum_{1\le k_1<\cdots<k_d\le n}w^2(k_1,\dots,k_d).$$

The assumption (4.4) plays a technical role, which helps in simplifying the derivations. Nevertheless, it covers important examples of U-statistics such as quadratic forms and their multidimensional generalizations. Sharp bounds have been provided in Chen and Shao (2007) in case (4.4) is not satisfied, but only in the case of classical (i.e. non-weighted) U-statistics. See also Krokowski et al. (2016) for weighted first order U-statistics based on symmetric Rademacher sequences, and Nourdin et al. (2016) for a fourth moment type central limit theorem in case $g(x_1,\dots,x_n)=x_1\cdots x_n$ and X_1 has a vanishing third moment.

In order to formulate the next result, given ν a probability measure on $\mathbb{R}_+$ we use the notation

$$f_n \overset{(\nu)}{\star}{}_k^l\, g_m(y_1,\dots,y_{n-l},z_1,\dots,z_{m-k}) := \frac{1}{2^l}\int_{\mathbb{R}_+^l} f_n(x_1,\dots,x_l,y_1,\dots,y_{n-l}) \\ \times g_m(x_1,\dots,x_l,y_1,\dots,y_{k-l},z_1,\dots,z_{m-k})\nu(dx_1)\cdots\nu(dx_l), \tag{4.5}$$

where $f_n \in L^2(\mathbb{R}_+^n,\nu^{\otimes n})$, $g_m \in L^2(\mathbb{R}_+^n,\nu^{\otimes m})$, which is a generalization of (2.15). Nevertheless, the two definitions are used in different contexts since ν is a probability measure and (4.5) can be interpreted as an expected value of function of a random vector, while the contraction (2.15) can be used to compute the expected value of a stochastic integral.

Theorem 4.3 *Let $U_{n,d}$ be a degenerate weighted U-statistics of the form (4.3). We have*

$$d_{W/K}\left(\frac{U_{n,d}}{\sigma},\mathcal{N}\right) \\ \leq C_d \frac{\max_{1\leq l\leq d-1}\left\{\|g \overset{(\nu)}{\star}{}_l^l\, g\|_{L^2(\mathbb{R}^d,\nu^{\otimes(d-l)})}\sqrt{\sum_{\mathbf{k},\mathbf{r}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w(\mathbf{k},\mathbf{m})w(\mathbf{r},\mathbf{m})\right)^2}\right\}}{\|g\|^2_{L^2(\mathbb{R}^d,\nu^{\otimes d})}\sum_{1\leq k_1,\dots,k_d\leq n} w^2(k_1,\dots,k_d)} \\ \leq C_d\frac{\|g\|^2_{L^4(\mathbb{R}^d,\nu^{\otimes d})}}{\|g\|^2_{L^2(\mathbb{R}^d,\nu^{\otimes d})}}\frac{\max_{1\leq l\leq d-1}\sqrt{\sum_{\mathbf{k},\mathbf{r}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w(\mathbf{k},\mathbf{m})w(\mathbf{r},\mathbf{m})\right)^2}}{\sum_{\mathbf{m}\in\mathbb{N}^d} w^2(\mathbf{m})}$$

for some $C_d > 0$ depending only on $d \in \{1,\dots,n\}$, where ν denotes the distribution of X_1.

Proof. By Theorem 4.2, we have

$$d_{W/K}\left(\frac{U_{n,d}}{\sigma},\mathcal{N}\right) \\ \leq \frac{c_d}{\sigma^2}\binom{n}{d}^{-2}\left(\sum_{0\leq l\leq d-1}\int_{\mathbb{R}^{d-l}}\left(\int_{\mathbb{R}^l} g^2(x,y)\nu^{\otimes l}(dx)\right)^2\nu^{\otimes(d-l)}(dy)\sum_{\mathbf{k}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w^2(\mathbf{k},\mathbf{m})\right)^2\right. \\ + \sum_{1\leq l\leq d-1}\int_{\mathbb{R}^{d-l}}\int_{\mathbb{R}^{d-l}}\left(\int_{\mathbb{R}^l} g(x,y)g(x,z)\nu^{\otimes l}(dx)\right)^2\nu^{\otimes(d-l)}(dy)\nu^{\otimes(d-l)}(dz) \\ \left.\times\sum_{\mathbf{k},\mathbf{r}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w(\mathbf{k},\mathbf{m})w(\mathbf{r},\mathbf{m})\right)^2\right)^{1/2}.$$

Applying the inequality

$$\sum_{\mathbf{k}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w^2(\mathbf{k},\mathbf{m})\right)^2 \leq \sum_{\mathbf{k},\mathbf{r}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w(\mathbf{k},\mathbf{m})w(\mathbf{r},\mathbf{m})\right)^2,$$

to the terms in the first sum, as well as the inequality

$$
\begin{aligned}
&\int_{\mathbb{R}^{d-l}}\int_{\mathbb{R}^{d-l}}\left(\int_{\mathbb{R}^l} g(x,y)g(x,z)\nu^{\otimes l}(dx)\right)^2 \nu^{\otimes(d-l)}(dy)\nu^{\otimes(d-l)}(dz)\\
&=\int_{\mathbb{R}^{d-l}}\int_{\mathbb{R}^{d-l}}\int_{\mathbb{R}^l} g(x_1,y)g(x_1,z)\nu^{\otimes l}(dx_1)\int_{\mathbb{R}^l} g(x_2,y)g(x_2,z)\nu^{\otimes l}(dx)\nu^{\otimes(d-l)}(dy)\nu^{\otimes(d-l)}(dz)\\
&=\int_{\mathbb{R}^l}\int_{\mathbb{R}^l}\left(\int_{\mathbb{R}^{d-l}} g(x_1,y)g(x_2,y)\nu^{\otimes(d-l)}(dy)\right)^2\nu^{\otimes l}(dx_1)\nu^{\otimes l}(dx_2)\\
&=\int_{\mathbb{R}^l}\int_{\mathbb{R}^l}\int_{\mathbb{R}^{d-l}} g^2(x_1,y)g^2(x_2,y)\nu^{\otimes(d-l)}(dy)\nu^{\otimes l}(dx_1)\nu^{\otimes l}(dx_2)\\
&=\int_{\mathbb{R}^{d-l}}\left(\int_{\mathbb{R}^l} g^2(x,y)\nu^{\otimes l}(dx_1)\right)^2\nu^{\otimes(d-l)}(dy)=\|g \overset{(\nu)}{\star}{}_l^l\, g\|^2_{L^2(\mathbb{R}^d,\nu^{\otimes(d-l)})},
\end{aligned}
$$

where we used Jensen's inequality, to the terms in the latter sum, we arrive at

$$
\begin{aligned}
&d_{W/K}\left(\frac{U_{n,d}}{\sigma},\mathcal{N}\right)\\
&\le \frac{c_d'}{\sigma^2}\binom{n}{d}^{-2}\left(\sum_{0\le l\le d-1}\|g \overset{(\nu)}{\star}{}_l^l\, g\|^2_{L^2(\mathbb{R}^d,\nu^{\otimes(d-l)})}\sum_{\mathbf{k},\mathbf{r}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w(\mathbf{k},\mathbf{m})w(\mathbf{r},\mathbf{m})\right)^2\right)^{1/2}\\
&\le C_d\frac{\max_{1\le l\le d-1}\left\{\|g \overset{(\nu)}{\star}{}_l^l\, g\|_{L^2(\mathbb{R}^d,\nu^{\otimes(d-l)})}\sqrt{\sum_{\mathbf{k},\mathbf{r}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w(\mathbf{k},\mathbf{m})w(\mathbf{r},\mathbf{m})\right)^2}\right\}}{\|g\|^2_{L^2(\mathbb{R}^d,\nu^{\otimes d})}\sum_{1\le k_1,\dots,k_d\le n} w^2(k_1,\dots,k_d)},
\end{aligned}
$$

which is the first bound from the assertion. To obtain the other one, it is enough to employ Jensen's inequality once again as follows

$$
\begin{aligned}
\|g \overset{(\nu)}{\star}{}_l^l\, g\|_{L^2(\mathbb{R}^d,\nu^{\otimes(d-l)})} &= \int_{\mathbb{R}^{d-l}}\left(\int_{\mathbb{R}^l} g^2(x,y)\nu^{\otimes l}(dx)\right)^2\nu^{\otimes(d-l)}(dy)\\
&\le \int_{\mathbb{R}^{d-l}}\int_{\mathbb{R}^l} g^4(x,y)\nu^{\otimes l}(dx)\nu^{\otimes(d-l)}(dy)=\|g\|^4_{L^4(\mathbb{R}^d,\nu^{\otimes d})}.
\end{aligned}
$$

This ends the proof. □

Taking $w\equiv 1$, we have for $1\le l\le d-1$

$$
\sqrt{\sum_{\mathbf{k},\mathbf{r}\in\mathbb{N}^{d-l}}\left(\sum_{\mathbf{m}\in\mathbb{N}^l} w(\mathbf{k},\mathbf{m})w(\mathbf{r},\mathbf{m})\right)^2}\approx\sum_{1\le k_1,\dots,k_d\le n} w^2(k_1,\dots,k_d)\approx n^d
$$

as n tends to infinity, where $f\approx g$ for non-negative functions f,g means that there is a constant $C>0$ depending on d such that $f/C\le g\le C f$. Applying the above equivalence to the first inequality of Theorem 4.3, we immediately obtain the next corollary.

Corollary 4.4 *Let $U_{n,d}$ be a degenerate weighted U-statistics of the form*

$$U_{n,d} = \binom{n}{d}^{-1} \sum_{1\leq k_1<\cdots<k_d\leq n} g\left(X_{k_1},\ldots,X_{k_d}\right), \qquad 1\leq d\leq n.$$

We have

$$d_{W/K}\left(\frac{U_{n,d}}{\sigma},\mathcal{N}\right) \leq C_d \max_{1\leq l\leq d-1} \frac{\|g \overset{(\nu)}{\star}{}_l^l\, g\|_{L^2(\mathbb{R}^d,\nu^{\otimes(d-l)})}}{\|g\|^2_{L^2(\mathbb{R}^d,\nu^{\otimes d})}}.$$

An analogous result dealing only with the Wasserstein distance has been provided in Theorem 3.3 of Döbler and Peccati (2019). Although the explicit values of constants have not been provided for simplicity in Theorem 4.3 and Corollary 4.4, they can be fully computed from the proof arguments.

4.3 Random graphs

Consider the Erdős and Rényi (1959) random graph $\mathbb{G}_n(p)$ constructed by independently retaining any edge in the complete graph K_n on n vertices with probability $p\in(0,1)$. Here, we assign an independent sample of a random weight X to every edge in $\mathbb{G}_n(p_n)$, and we define the weight of a graph contained in $\mathbb{G}_n(p_n)$ as the sum of weights of its edges. Then, consider the renormalized random weight

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

where W_n^G denotes the combined weight of graphs in $\mathbb{G}_n(p_n)$ that are isomorphic to a fixed graph G. By writing the combined weight W_n^G of graphs in $\mathbb{G}_n(p_n)$ that are isomorphic to a fixed graph G as a sum of multiple stochastic integrals (which is equivalent to finding its Hoeffding decomposition) we obtain the following result as in Privault and Serafin (2022), by replacing the use of Theorem 5.1 therein with Theorem 3.2 above.

Theorem 4.5 *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*

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

for some constant $C=C(e_G)>0$, where v_H, e_H denotes the numbers of vertices and edges, respectively, of a graph H.

Theorem 4.5 extends other Kolmogorov distance bounds previously obtained for triangle counting in Ross (2011), and in Krokowski et al. (2017) using the Malliavin approach to the Stein method, see also Röllin (2022) for triangle counting, Privault and Serafin (2020) for arbitrary subgraph counting, and Krokowski et al. (2016) for weighted first order Rademacher U-statistics in the symmetric case $p = 1/2$. As a consequence, if p_n satisfies $p_n < c < 1$, $n \geq 1$, we have

$$d_{W/K}\big(\widetilde{W}_n^G, \mathcal{N}\big) \leq C \frac{\sqrt{\mathbb{E}\left[X^4\right]}}{\mathbb{E}[X^2]} \left((1-p_n) \min_{\substack{H \subset G \\ e_H \geq 1}} n^{v_H} p_n^{e_H} \right)^{-1/2}, \tag{4.6}$$

and for $p_n > c > 0$, $n \geq 1$, it holds

$$d_{W/K}\big(\widetilde{W}_n^G, \mathcal{N}\big) \leq C \frac{\sqrt{\mathbb{E}\left[X^4\right]}}{n\sqrt{1-p_n}\operatorname{Var}[X]}. \tag{4.7}$$

In particular, when $X \equiv 1$ is a constant, (4.6) and (4.7) recover the Wasserstein and Kolmogorov bounds of Theorem 2 in Barbour et al. (1989) and Theorem 4.2 in Privault and Serafin (2020). Applications to cycle graphs, complete graphs trees can be treated as in Privault and Serafin (2022) by replacing the Kolmogorov distance with the Wasserstein distance.

5 Quadratic forms

5.1 Context and results

We consider the quadratic form Q_n defined as

$$Q_n = \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} a_{ij} X_i X_j + \sum_{k=1}^n a_{kk} \left(X_k^2 - \mathbb{E}[X_k^2] \right),$$

where $A_n = (a_{ij})_{1 \leq i,j \leq n}$ is a symmetric matrix, $n \geq 1$, and $(X_k)_{k \geq 1}$ denotes i.i.d. copies of a given random variable X satisfying $\mathbb{E}\left[X\right] = 0$. In the sequel, we let $\mu_k := \mathbb{E}[X^k]$, $\tilde{\mu}_k := \mathbb{E}\big[\left(X^2 - \mathbb{E}[X^2]\right)^{k/2} \big]$, $k \geq 2$, and

$$\sigma_n^2 := \operatorname{Var}[Q_n] = \mathbb{E}\left[Q_n^2\right] = 2\mu_2^2 \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} a_{ij}^2 + \tilde{\mu}_4 \sum_{i=1}^n a_{ii}^2.$$

Many papers in the literature are devoted to asymptotical normality of quadratic forms. The best known convergence rates in the general case where the diagonal of A may not vanish are given in Götze and Tikhomirov (2002), as

$$d_K\left(\frac{Q_n}{\sigma_n}, \mathcal{N}\right) \le C(\gamma)\frac{\left(\mathbb{E}\left[|X|^3\right]\right)^2 + \gamma\mathbb{E}[X^6]}{\sqrt{\sum_{1\le i,j\le n} a_{ij}^2}}|\lambda_1|, \tag{5.1}$$

see Theorem 1.1 therein, where λ_1 denotes the largest absolute eigenvalue of A_n, $\gamma = \sum_{i=1}^n a_{ii}^2 / \sum_{1\le i,j\le n} a_{ij}^2$, and the constant $C(\gamma)$ blows up when γ tends to one, i.e. when the linear part is dominating.

Vanishing diagonals

More is known if we assume the diagonal of A_n to be empty, in which case de Jong (1987) proved the asymptotic normality of Q_n/σ_n under the conditions

$$\mathbb{E}\left[(Q_n/\sigma_n)^4\right] \longrightarrow 3 \quad \text{and} \quad \frac{1}{\sigma_n^2}\max_{1\le i\le n}\sum_{j=1}^n a_{ij}^2 \longrightarrow 0. \tag{5.2}$$

In addition, for $(X_k)_{k\ge 1}$ a Rademacher sequence, Theorem 1.1 in Döbler and Krokowski (2019) restricted to double integrals gives the corresponding bound

$$d_K\left(\frac{Q_n}{\sigma_n}, \mathcal{N}\right) \le C\left(\sqrt{\left|\mathbb{E}\left[(Q_n/\sigma_n)^4\right] - 3\right|} + \frac{1}{\sigma_n}\sqrt{\max_{1\le i\le n}\sum_{1\le j\le 1} a_{ij}^2}\right). \tag{5.3}$$

The same bound may be concluded from Döbler and Peccati (2017) for $(X_k)_{k\ge 1}$ being any i.i.d. sequence, but only in Wasserstein distance. Note that the quantity $\max_{1\le i\le n}\sum_{j=1}^n a_{ij}^2$ corresponds to "maximal influence", see Mossel et al. (2010), Nourdin et al. (2010b).

The bound

$$d_W\left(\frac{Q_n}{\sigma_n}, \mathcal{N}\right) \le C\frac{\mu_4}{\sigma_n^2}\left(\sqrt{\sum_{i=1}^n\left(\sum_{k=1}^n a_{ik}^2\right)^2} + \sqrt{\sum_{i,j=1}^n\left(\sum_{k=1}^n a_{ik}a_{kj}\right)^2}\right). \tag{5.4}$$

has been provided for Rademacher sequences using the Wasserstein distance in Proposition 3.1 of Chatterjee (2008), and has been recently extended to arbitrary i.i.d. sequences using the Kolmogorov distance in Shao and Zhang (2019), Theorem 3.1.

Corollary 5.1 recovers this bound as an immediate consequence of Theorem 4.3 by taking $d = 2$, $w(k_1, k_2) = a_{k_1k_2}$, $1 \le k_1, k_2 \le n$, $k_1 \ne k_2$, and $g(y_1, y_2) = y_1y_2$. Note however that only the second term is significant in the right-hand side of (5.4), making the conjecture at the end of Section 3.1 in Shao and Zhang (2019) pointless.

Corollary 5.1 *Assume $a_{ii} = 0$, $i = 1, \ldots, n$. Then, there exists a constant $C > 0$ such that*

$$d_K\left(\frac{Q_n}{\sigma_n}, \mathcal{N}\right) \leq C\frac{\mu_4}{\sigma_n^2}\sqrt{\sum_{i,j=1}^{n}\left(\sum_{k=1}^{n} a_{ik}a_{kj}\right)^2} = C\frac{\mu_4}{\sigma_n^2}\sqrt{\mathrm{Tr}(A_n^4)}, \qquad n \geq 1.$$

Corollary 5.1 also improves (5.1) for matrices A_n with empty diagonal, since

$$\sqrt{\mathrm{Tr}(A_n^4)} = \sqrt{\sum_{k=1}^{n}\lambda_k^4} \leq |\lambda_1|\sqrt{\sum_{k=1}^{n}\lambda_k^2} \leq |\lambda_1|\sqrt{\sum_{i,j=1}^{n}a_{ij}^2} \leq \frac{\sigma_n}{\mu_2}|\lambda_1|. \tag{5.5}$$

Non-empty diagonals

Theorem 5.2 below generalizes and improves all the aforementioned results. First, in comparison with the above bound (5.1) of Götze and Tikhomirov (1999; 2002), it gives better rates under weaker assumptions, as noted in (5.5). Furthermore, it extends every other result by applying as well to non-vanishing diagonals. In addition, it completes Corollary 5.1 with an additional bound related to the so-called fourth moment phenomenon (Nualart and Peccati (2004)), and it also extends (5.3) from the Rademacher case to any distribution. Finally, it deals with the Kolmogorov distance instead of the Wasserstein distance considered in Döbler and Peccati (2017). See also Theorem 3.11 in Bally and Caramellino (2019) for some bounds in total variation and Kolmogorov distances, which however provide worse rates and require slightly stronger assumptions.

Theorem 5.2 *There exist absolute constants $C_1, C_2 > 0$ such that*

$$d_{W/K}\left(\frac{Q_n}{\sigma_n}, \mathcal{N}\right) \leq C_1\left(\sqrt{\left|\mathbb{E}\left[(Q_n/\sigma_n)^4\right] - 3\right|} + \frac{\alpha_n}{\sigma_n}\sqrt{\max_{1\leq i\leq n}\sum_{1\leq j\leq 1} a_{ij}^2}\right), \tag{5.6}$$

and

$$d_K\left(\frac{Q_n}{\sigma_n}, \mathcal{N}\right) \leq C_2\frac{\beta_n}{\sigma_n^2}\sqrt{\mathrm{Tr}(A_n^4)}, \tag{5.7}$$

where

$$\alpha_n := \mu_2 + \frac{\mu_4}{\mu_2}\mathbb{1}_{\left\{a_{11}^2+\cdots+a_{nn}^2>0\right\}}, \qquad \text{and} \qquad \beta_n = \mu_4 + \sqrt{\mu_8}\,\mathbb{1}_{\left\{a_{11}^2+\cdots+a_{nn}^2>0\right\}}.$$

Contrary to what is stated on page 1590 of Chatterjee (2008), the conditions $\sigma_n^{-2}\sqrt{\mathrm{Tr}(A_n^4)} \to 0$ and $\mathbb{E}\left[(Q_n/\sigma_n)^4\right] \to 3$ are not equivalent as n tends to infinity, and therefore fourth moment convergence is not sufficient for the central limit theorem to hold for quadratic functionals. The next proposition clarifies this point via inequalities between the quantities appearing in Theorem 5.2. In the sequel, we let $a \wedge b := \min(a, b)$, $a, b \in \mathbb{R}$.

Proposition 5.3 *There exist absolute constants $C_1, C_2, C_3 > 0$ such that*

$$C_1 \frac{\mu_2^4 \wedge \tilde{\mu}_8}{\sigma_n^4} \operatorname{Tr}(A_n^4) \leq \left| \mathbb{E}\left[(Q_n/\sigma_n)^4 \right] - 3 \right| + \frac{\alpha_n^2}{\sigma_n^2} \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2$$
$$\leq C_2 \left(\frac{\beta_n^2}{\sigma_n^4} \operatorname{Tr}(A_n^4) + \frac{\alpha_n^2}{\sigma_n^2} \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2 \right) \leq C_3 \frac{\beta_n^2}{\mu_2^2 \sigma_n^2} \sqrt{\operatorname{Tr}(A_n^4)},$$

where α_n, β_n are as in Theorem 5.2

Theorem 5.2 and Proposition 5.3 immediately imply

Corollary 5.4 *Assume $(X_i)_{i\in\mathbb{N}}$ is a fixed i.i.d. sequence with zero means and finite 8th moments. The following two conditions are equivalent:*

a) $\mathbb{E}\left[(Q_n/\sigma_n)^4\right] \longrightarrow 3$ *and* $\sigma_n^{-2} \max_{1\leq i\leq n} \sum_{j=1}^n a_{ij}^2 \longrightarrow 0$,

b) $\sigma_n^{-4} \operatorname{Tr}(A_n^4) \longrightarrow 0$,

and they imply $Q_n/\sigma_n \xrightarrow{\mathcal{L}} \mathcal{N}$ with the Kolmogorov rates (5.6) and (5.7).

This extends (5.2) for any matrix A_n and completes it with the equivalent condition in terms of the trace of A_n.

5.2 Proofs

Proof of Theorem 5.2. The quadratic form Q_n admits the Hoeffding decomposition

$$Q_n = \sum_{1\leq i,j\leq n} W_{\{i,j\}} + \sum_{k=1}^n W_{\{k\}},$$

where

$$W_{\{i,j\}} = 2a_{ij} X_i X_j, \qquad W_{\{k\}} = a_{kk} \left(X_k^2 - \mathbb{E}\left[X_k^2\right] \right).$$

Thus, Theorem 4.1 gives

$$d_{W/K}\left(\frac{Q_n}{\sigma_n}, \mathcal{N} \right) \leq \frac{C}{\sigma_n^2} \Bigg(\tilde{\mu}_8 \sum_{i=1}^n a_{ii}^4 + 2\mu_4^2 \sum_{\substack{1\leq i,j\leq n \\ i\neq j}} a_{ij}^4 + 2\mu_2^2 \mu_4 \sum_{\substack{1\leq i,j,k\leq n \\ i\neq j, i\neq k, j\neq k}} a_{ij}^2 a_{ik}^2 \tag{5.8}$$
$$+ \mu_2^4 \sum_{\substack{1\leq i,j\leq n \\ i\neq j}} \Bigg(\sum_{\substack{1\leq k\leq n \\ k\neq i,j}} a_{ik} a_{kj} \Bigg)^2 + \mu_3^2 \mu_2 \sum_{i=1}^n \Bigg(\sum_{\substack{1\leq j\leq n \\ i\neq j}} a_{jj} a_{ij} \Bigg) \Bigg)^{1/2}.$$

Next, we estimate this bound by means of $\mathbb{E}\left[Q_n^4\right]$ and $\max_{1\le i\le n}\sum_{1\le j\le n} a_{ij}^2$. A direct calculation shows that

$$\mathbb{E}\left[Q_n^4\right] = S_1 + 3S_2 + 4S_3,$$

where

$$\begin{aligned} S_1 :=\ & \tilde{\mu}_8 \sum_{i=1}^n a_{ii}^4 + 16\mu_4^2 \sum_{1\le i<j\le n} a_{ij}^4 + 48\mu_2^2\mu_4 \sum_{\substack{1\le i,j,k\le n\\ i\ne j, i\ne k, j\ne k}} a_{ij}^2 a_{ik}^2 \\ & + 48\mu_2^4 \sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ i_k\ne i_l \text{ if } k\ne l}} a_{i_1i_2}a_{i_2i_3}a_{i_3i_4}a_{i_4i_1} + 48\mu_3^2\mu_2 \sum_{\substack{1\le i,j,k\le n\\ i\ne j, i\ne k, j\ne k}} a_{ii}a_{jj}a_{ik}a_{kj} \\ & + 48\mu_3^2\mu_2 \sum_{\substack{1\le i,j,k\le n\\ i\ne j, i\ne k, j\ne k}} a_{kj}^2 a_{ik}a_{ij}, \end{aligned}$$

and

$$S_2 := \tilde{\mu}_4^2 \sum_{i\ne j} a_{ii}^2 a_{jj}^2 + 4\tilde{\mu}_4\mu_2^2 \sum_{\substack{1\le i,j,k\le n\\ j\ne k,\ j,k\ne i}} a_{ii}^2 a_{jk}^2 + 4\mu_2^4 \sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ i_k\ne i_l \text{ if } k\ne l}} a_{i_1i_2}^2 a_{i_3i_4}^2,$$

and

$$\begin{aligned} S_3 :=\ & 3\tilde{\mu}_4^2 \sum_{\substack{1\le i,j\le n\\ i\ne j}} a_{ii}a_{jj}a_{ij}^2 + 8\mu_3\left(\mu_5 - \mu_3\mu_2\right)\sum_{i\ne j} a_{ii}a_{ij}^3 + 6\mu_2\left(\tilde{\mu}_6 + \tilde{\mu}_4\mu_2\right)\sum_{\substack{1\le i,j\le n\\ i\ne j}} a_{ii}^2 a_{ij}^2 \\ & + 12\mu_3^2\mu_2 \sum_{\substack{1\le i,j,k\le n\\ i\ne j, j\ne k, i\ne k}} a_{ii}a_{ij}a_{jk}^2 + 24\mu_2^2\tilde{\mu}_4 \sum_{\substack{1\le i,j,k\le n\\ i\ne j, j\ne k, i\ne k}} a_{ii}a_{ij}a_{ik}a_{kj}. \end{aligned}$$

The sum S_1 is to dominate the right-hand side of (5.8), S_2 is approximating σ^2, and S_3 contains remainder terms that are more difficult to handle due to their unknown sign. Note also that S_3 vanishes if the diagonal of A is empty. First, by

$$\sum_{\substack{1\le i,j\le n\\ i\ne j}} \left(\sum_{\substack{1\le k\le n\\ k\ne i,j}} a_{ik}a_{kj}\right)^2 = \sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ i_k\ne i_l \text{ if } k\ne l}} a_{i_1i_2}a_{i_2i_3}a_{i_3i_4}a_{i_4i_1} + \sum_{\substack{1\le i,j,k\le n\\ i\ne j, i\ne k, j\ne k}} a_{ik}^2 a_{kj}^2$$

and

$$\sum_{i=1}^n \left(\sum_{\substack{1\le j\le n\\ i\ne j}} a_{jj}a_{ij}\right)^2 = \sum_{\substack{1\le i,j,k\le n\\ i\ne j, i\ne k, j\ne k}} a_{ii}a_{jj}a_{ik}a_{kj} + \sum_{\substack{1\le i,j\le n\\ i\ne j}} a_{ii}^2 a_{ij}^2,$$

we get

$$S_1 := \tilde{\mu}_8 \sum_{i=1}^n a_{ii}^4 + 16\mu_4^2 \sum_{1\le i<j\le n} a_{ij}^4 + 48\mu_2^2\mu_4 \sum_{\substack{1\le i,j,k\le n\\ i\ne j, i\ne k, j\ne k}} a_{ij}^2 a_{ik}^2$$

$$+ 48\mu_2^4 \sum_{\substack{1\le i,j\le n\\ i\neq j}} \left(\sum_{\substack{1\le k\le n\\ k\neq i,j}} a_{ik}a_{kj} \right)^2 + 48\mu_3^2\mu_2 \sum_{i=1}^n \left(\sum_{\substack{1\le j\le n\\ i\neq j}} a_{jj}a_{ij} \right)^2$$
$$- 48\mu_2^4 \sum_{\substack{1\le i,j,k\le n\\ i\neq j, i\neq k, j\neq k}} a_{ik}^2 a_{kj}^2 - 48\mu_3^2\mu_2 \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2 a_{ij}^2 + 48\mu_3^2\mu_2 \sum_{\substack{1\le i,j,k\le n\\ i\neq j, i\neq k, j\neq k}} a_{kj}^2 a_{ik}a_{ij}.$$

The first two lines dominate the right-hand side of (5.8) with substantial surplus, which will be used to deal with the last term of S_1 and some terms of S_3. Indeed, by $\mu_3^2\mu_2 \le \mu_4\mu_2^2$ and the inequality of arithmetic and geometric means, we have

$$\begin{aligned}
&48\mu_3^2\mu_2 \left| \sum_{\substack{1\le i,j,k\le n\\ i\neq j, i\neq k, j\neq k}} a_{kj}^2 a_{ik}a_{ij} \right| \\
&\le 46\mu_4\mu_2^2 \sum_{\substack{1\le i,j,k\le n\\ i\neq j, i\neq k, j\neq k}} \frac{1}{2}\Big((a_{kj}a_{ik})^2 + (a_{kj}a_{ij})^2 \Big) + \sum_{\substack{1\le k,j\le n\\ k\neq j}} \left((\mu_4 a_{kj}^2)^2 + \left(\mu_2^2 \sum_{\substack{1\le i\le n\\ i\neq j,k}} a_{ik}a_{ij} \right)^2 \right) \\
&= 46\mu_2^2\mu_4 \sum_{\substack{1\le i,j,k\le n\\ i\neq j, i\neq k, j\neq k}} a_{ij}^2 a_{ik}^2 + \mu_4^2 \sum_{1\le i<j\le n} a_{ij}^4 + \mu_2^4 \sum_{\substack{1\le i,j\le n\\ i\neq j}} \left(\sum_{\substack{1\le k\le n\\ k\neq i,j}} a_{ik}a_{kj} \right)^2. \qquad (5.9)
\end{aligned}$$

Since, additionally

$$\begin{aligned}
&\left| \mu_2^4 \sum_{\substack{1\le i,j,k\le n\\ i\neq j, i\neq k, j\neq k}} a_{ik}^2 a_{kj}^2 + \mu_3^2\mu_2 \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2 a_{ij}^2 \right| \\
&\le \sigma_n^2\mu_2^2 \max_{1\le i\le n} \sum_{j=1}^n a_{ij}^2 + \sigma_n^2 \frac{\mu_3^2}{\mu_2} \mathbb{1}_{\left\{a_{11}^2+\cdots+a_{nn}^2>0\right\}} \max_{1\le i\le n} \sum_{j=1}^n a_{ij}^2 \le \sigma_n^2 \alpha_n^2 \max_{1\le i\le n} \sum_{j=1}^n a_{ij}^2, \qquad (5.10)
\end{aligned}$$

we arrive at

$$\begin{aligned}
d_{W/K}\left(\frac{Q_n}{\sigma_n}, \mathcal{N} \right) \le \frac{C}{\sigma_n^2} \Bigg(& S_1 + 48\sigma_n^2\alpha_n^2 \max_{1\le i\le n} \sum_{j=1}^n a_{ij}^2 \\
& + 46\mu_2^4 \sum_{\substack{1\le i,j\le n\\ i\neq j}} \left(\sum_{\substack{1\le k\le n\\ k\neq i,j}} a_{ik}a_{kj} \right)^2 + 47\mu_3^2\mu_2 \sum_{i=1}^n \left(\sum_{\substack{1\le j\le n\\ i\neq j}} a_{jj}a_{ij} \right)^2 \Bigg)^{1/2} \\
\le \frac{C}{\sigma_n^2} \Bigg(& \mathbb{E}\left[Q_n^4 \right] - 3\sigma_n^4 + 48\sigma_n^2\alpha_n^2 \max_{1\le i\le n} \sum_{j=1}^n a_{ij}^2 + 3(\sigma_n^4 - S_2) \\
& - 4S_3 - 24\mu_2^4 \sum_{\substack{1\le i,j\le n\\ i\neq j}} \left(\sum_{\substack{1\le k\le n\\ k\neq i,j}} a_{ik}a_{kj} \right)^2 - 24\mu_3^2\mu_2 \sum_{i=1}^n \left(\sum_{\substack{1\le j\le n\\ i\neq j}} a_{jj}a_{ij} \right)^2 \Bigg)^{1/2}. \qquad (5.11)
\end{aligned}$$

Next, in order to bound $3(\sigma_n^4 - S_2)$, we calculate

$$\sigma_n^4 = \left(2\mu_2^2 \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ij}^2 + \tilde{\mu}_4 \sum_{i=1}^n a_{ii}^2\right)^2$$

$$= 4\mu_2^4\left(\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ij}^2\right)^2 + \tilde{\mu}_4^2\left(\sum_{i=1}^n a_{ii}^2\right)^2 + 4\mu_2^2\tilde{\mu}_4\left(\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ij}^2\right)\left(\sum_{i=1}^n a_{ii}^2\right)$$

$$= 4\mu_2^4\left(2\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ij}^4 + \sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ i_k\neq i_l \text{ if } k\neq l}} a_{i_1i_2}^2 a_{i_3i_4}^2 + 2\sum_{\substack{1\le i,j,k\le n\\ j\neq k,\ j,k\neq i}} a_{ij}^2 a_{ik}^2\right)$$

$$+ \tilde{\mu}_4^2\left(\sum_{i=1}^n a_{ii}^4 + \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2 a_{jj}^2\right) + 4\mu_2^2\tilde{\mu}_4\left(2\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2 a_{ij}^2 + \sum_{\substack{1\le i,j,k\le n\\ j\neq k,\ j,k\neq i}} a_{ii}^2 a_{jk}^2\right),$$

hence

$$3\left|\sigma_n^4 - S_2\right| = 24\mu_2^4\left(\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ij}^4 + \sum_{\substack{1\le i,j,k\le n\\ j\neq k,\ j,k\neq i}} a_{ij}^2 a_{ik}^2\right) + 3\tilde{\mu}_4^2\sum_{i=1}^n a_{ii}^4 + 24\mu_2^2\tilde{\mu}_4 \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2 a_{ij}^2$$

$$\le \max_{1\le i\le n}\sum_{1\le j\le n} a_{ij}^2\left(48\mu_2^4 \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ij}^2 + 27\mu_4\tilde{\mu}_4 \mathbb{1}_{\left\{a_{11}^2+\cdots+a_{nn}^2>0\right\}}\sum_{i=1}^n a_{ii}^2\right)$$

$$\le \sigma_n^2\left(48\mu_2^2 + 27\frac{\mu_4^2}{\mu_2^2}\mathbb{1}_{\left\{a_{11}^2+\cdots+a_{nn}^2>0\right\}}\right)\max_{1\le i\le n}\sum_{1\le j\le n} a_{ij}^2. \tag{5.12}$$

Regarding S_3, we have

$$\left|\tilde{\mu}_4^2 \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}a_{jj}a_{ij}^2\right| \le \tilde{\mu}_4^2 \sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2 a_{ij}^2 \le \sigma_n^2 \max_{1\le i\le n}\sum_{1\le j\le n} a_{ij}^2,$$

and

$$8\mu_3\left(\mu_5 - \mu_3\mu_2\right)\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}a_{ij}^3 + 6\mu_2\left(\tilde{\mu}_6 + \tilde{\mu}_4\mu_2\right)\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2 a_{ij}^2$$

$$= 2\mathbb{E}\left[4\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}a_{ij}^3\big(X_i^2 - \mathbb{E}\left[X_i^2\right]\big)X_i^3X_j^3 + 3\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ii}^2a_{ij}^2\big(X_i^2 - \mathbb{E}\left[X_i^2\right]\big)^2X_i^2X_j^2\right]$$

$$= 6\mathbb{E}\left[\sum_{\substack{1\le i,j\le n\\ i\neq j}}\left(a_{ii}\big(X_i^2 - \mathbb{E}\left[X_i^2\right]\big) + \frac{2}{3}a_{ij}X_iX_j\right)^2 a_{ij}^2X_i^2X_j^2\right] - \frac{8}{3}\mathbb{E}\left[\sum_{\substack{1\le i,j\le n\\ i\neq j}} a_{ij}^4X_i^4X_j^4\right]$$

$$\geq -\frac{8}{3}\mu_4^2 \sum_{\substack{1\leq i,j\leq n\\ i\neq j}} a_{ij}^4 \geq -\frac{4}{3}\sigma_n^2 \left(\frac{\mu_4}{\mu_2}\right)^2 \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2.$$

Furthermore, using the correction terms from (5.11), we get

$$\begin{aligned}
&12\mu_3^2\mu_2 \sum_{\substack{1\leq i,j,k\leq n\\ i\neq j, j\neq k, i\neq k}} a_{ii}a_{ij}a_{jk}^2 + 6\mu_3^2\mu_2 \sum_{i=1}^n \Bigg(\sum_{\substack{1\leq j\leq n\\ i\neq j}} a_{jj}a_{ij}\Bigg)^2 \\
&= 6\mu_3^2\mu_2 \sum_{i=1}^n \Bigg(\sum_{\substack{1\leq k\leq n\\ k\neq i}} a_{ik}^2 + \sum_{\substack{1\leq j\leq n\\ j\neq i}} a_{jj}a_{ij}\Bigg)^2 - 6\mu_3^2\mu_2 \sum_{i=1}^n \Bigg(\sum_{\substack{1\leq k\leq n\\ k\neq i}} a_{ik}^2\Bigg)^2 \\
&\geq -6\mu_3^2\mu_2 \Bigg(\max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2\Bigg) \sum_{i=1}^n \sum_{\substack{1\leq k\leq n\\ k\neq i}} a_{ik}^2 \geq -3\sigma_n^2 \left(\frac{\mu_4}{\mu_2}\right)^2 \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2,
\end{aligned}$$

as well as

$$\begin{aligned}
&24\mu_2^2\tilde{\mu}_4 \sum_{\substack{1\leq i,j,k\leq n\\ i\neq j, j\neq k, i\neq k}} a_{ii}a_{ij}a_{ik}a_{kj} + 6\mu_2^4 \sum_{\substack{1\leq i,j\leq n\\ i\neq j}} \Bigg(\sum_{\substack{1\leq k\leq n\\ k\neq i,j}} a_{ik}a_{kj}\Bigg)^2 \\
&= 6 \sum_{\substack{1\leq i,j\leq n\\ i\neq j}} \Bigg(2\tilde{\mu}_4 a_{ii}a_{ij} + \mu_2^2 \sum_{\substack{1\leq k\leq n\\ k\neq i,j}} a_{ik}a_{kj}\Bigg)^2 - 24\tilde{\mu}_4^2 \sum_{\substack{1\leq i,j\leq n\\ i\neq j}} a_{ii}^2 a_{ij}^2 \\
&\geq -24\tilde{\mu}_4^2 \Bigg(\max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2\Bigg) \sum_{\substack{1\leq i,j\leq n\\ i\neq j}} a_{ii}^2 \geq -24\sigma_n^2\mu_4 \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2.
\end{aligned}$$

Hence, we arrive at

$$\begin{aligned}
&S_3 + 6\mu_3^2\mu_2 \sum_{i=1}^n \Bigg(\sum_{\substack{1\leq j\leq n\\ i\neq j}} a_{jj}a_{ij}\Bigg)^2 + 6\mu_2^4 \sum_{\substack{1\leq i,j\leq n\\ i\neq j}} \Bigg(\sum_{\substack{1\leq k\leq n\\ k\neq i,j}} a_{ik}a_{kj}\Bigg)^2 \\
&\geq -C\sigma_n^2 \left(\frac{\mu_4}{\mu_2}\right)^2 \mathbb{1}_{\{a_{11}^2+\cdots+a_{nn}^2>0\}} \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2
\end{aligned}$$

for some $C > 0$, since S_3 vanishes if $a_{11} = \cdots = a_{nn} = 0$. Applying this and (5.12) to (5.11), we obtain the first inequality from the assertion. To prove the other one, we use (5.8) and write

$$d_{W/K}\left(\frac{Q_n}{\sigma_n}\right) \leq \frac{C}{\sigma_n^2} \left(\mu_4^2 \sum_{1\leq i,j\leq n} \Bigg(\sum_{\substack{1\leq k\leq n\\ k\neq i,j}} a_{ik}a_{kj}\Bigg)^2 + \mu_8 \sum_{i=1}^n a_{ii}^4 + \mu_8 \sum_{i=1}^n \Bigg(\sum_{\substack{1\leq j\leq n\\ i\neq j}} a_{jj}a_{ij}\Bigg)^2\right)^{1/2}.$$

Next, we bound

$$\sum_{1\le i,j\le n}\left(\sum_{\substack{1\le k\le n\\ k\ne i,j}} a_{ik}a_{kj}\right)^2 = \sum_{1\le i,j\le n}\left(\sum_{1\le k\le n} a_{ik}a_{kj} - a_{ii}a_{ij} - a_{ij}a_{jj}\right)^2$$

$$\le \sum_{1\le i,j\le n}\left[2\left(\sum_{1\le k\le n} a_{ik}a_{kj}\right)^2 + 4a_{ii}^2a_{ij}^2\right] \le 2\mathrm{Tr}(A_n^4) + 4\sum_{1\le i\le n}\left(\sum_{1\le k\le n} a_{ik}^2\right)^2 \le 6\mathrm{Tr}(A_n^4), \tag{5.13}$$

and, by the inequality $ab \le (a^2+b^2)/2$,

$$\sum_{i=1}^n a_{ii}^4 + \sum_{i=1}^n\left(\sum_{\substack{1\le j\le n\\ i\ne j}} a_{jj}a_{ij}\right)^2 = \sum_{i=1}^n a_{ii}^4 + \sum_{i=1}^n \sum_{\substack{1\le j,k\le n\\ i\ne j}} (a_{ij}a_{kk})(a_{ik}a_{jj})$$

$$\le \sum_{i=1}^n a_{ii}^4 + \sum_{\substack{1\le i,j,k\le n\\ i\ne j}} (a_{ij}a_{kk})^2 \le 2\sum_{1\le i\le n}\left(\sum_{1\le k\le n} a_{ik}^2\right)^2 \le 2\mathrm{Tr}(A_n^4). \tag{5.14}$$

This ends the proof. □

Proof of proposition 5.3. The proof of Theorem 5.2 shows that the right hand side of (5.6) is larger than the right hand side of (5.8) up to an absolute multiplicative constant, hence we have

$$\left|\mathbb{E}\left[(Q_n/\sigma_n)^4\right] - 3\right| + \alpha_n^2\sigma_n^2 \max_{1\le i\le n}\sum_{1\le j\le n} a_{ij}^2 \ge C\frac{\mu_2^4\wedge\tilde{\mu}_8}{\sigma_n^4}\left(\sum_{i=1}^n a_{ii}^4 + \sum_{1\le i,j\le n}\left(\sum_{\substack{1\le k\le n\\ k\ne i,j}} a_{ik}a_{kj}\right)^2\right).$$

Employing the inequalities $(a+b)^2 \le 2a^2+2b^2$ and $ab \le (a^2+b^2)/2$, $a,b\ge 0$, we get

$$\sum_{1\le i,j\le n}\left(\sum_{1\le k\le n} a_{ik}a_{kj}\right)^2 = \sum_{1\le i,j\le n}\left(\sum_{\substack{1\le k\le n\\ k\ne i,j}} a_{ik}a_{kj} + a_{ii}a_{ij} + a_{ij}a_{jj}\right)^2$$

$$\le \sum_{1\le i,j\le n}\left[\left(\sum_{\substack{1\le i,j\le n\\ k\ne i,j}} a_{ik}a_{kj}\right)^2 + 8a_{ii}^2a_{ij}^2\right]$$

$$= \sum_{1\le i,j\le n}\left(\sum_{\substack{1\le i,j\le n\\ k\ne i,j}} a_{ik}a_{kj}\right)^2 + 8\sum_{1\le i\le n} a_{ii}^2 \sum_{\substack{1\le j\le n\\ j\ne i}} a_{ij}^2 + 8\sum_{1\le i\le n} a_{ii}^4$$

$$\leq 5 \sum_{1\leq i,j\leq n} \left(\sum_{\substack{1\leq i,j\leq n \\ k\neq i,j}} a_{ik}a_{kj} \right)^2 + 12 \sum_{1\leq i\leq n} a_{ii}^4,$$

which gives the first inequality in the assertion. In order to justify the latter one, we will show

$$\left|\mathbb{E}\left[(Q_n/\sigma_n)^4\right] - 3\right| \leq C \left(\frac{\beta_n^2}{\sigma_n^4} \operatorname{Tr}(A_n^4) + \frac{\alpha_n^2}{\sigma_n^2} \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2 \right), \tag{5.15}$$

for some $C > 0$. Following notation from the proof of Theorem 5.2, we have $\left|\mathbb{E}\left[(Q_n/\sigma_n)^4\right] - 3\right| \leq \left(|S_1| + 3|S_2 - \sigma_n^4| + 4|S_3|\right)/\sigma_n^4$. By (5.9), (5.10), (5.12) and bounding terms from the first three sums in S_3 by $a_{ii}^2a_{ij}^2 + a_{ij}^4$ and the last two sums from S_3 by

$$\sum_{i=1}^n \left[\left(\sum_{\substack{1\leq k\leq n \\ k\neq i}} a_{ik}^2 \right)^2 + \left(\sum_{\substack{1\leq j\leq n \\ j\neq i}} a_{jj}a_{ij} \right)^2 \right],$$

and

$$\sum_{\substack{1\leq i,j\leq n \\ i\neq j}} \left[a_{ii}^2a_{ij}^2 + \left(\sum_{\substack{1\leq k\leq n \\ k\neq i,j}} a_{ik}a_{kj} \right)^2 \right],$$

respectively, we arrive at

$$\begin{aligned}
&\left|\mathbb{E}\left[(Q_n/\sigma_n)^4\right] - 3\right| \\
&\leq C\frac{\beta_n^2}{\sigma_n^4} \left[\sum_{i=1}^n a_{ii}^4 + \sum_{1\leq i,j\leq n} \left(\sum_{\substack{1\leq k\leq n \\ k\neq i,j}} a_{ik}a_{kj} \right)^2 + \sum_{i=1}^n \left(\sum_{\substack{1\leq j\leq n \\ i\neq j}} a_{jj}a_{ij} \right)^2 \right] + \alpha_n^2\sigma_n^2 \max_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}^2,
\end{aligned}$$

and (5.15) follows from (5.13) and (5.14). Finally, the last bound in the assertion is a consequence of

$$\max_{1\leq i\leq n} \sum_{j=1}^n a_{ij}^2 \leq \sqrt{\sum_{i=1}^n \left(\sum_{j=1}^n a_{ij}^2 \right)^2} \leq \sqrt{\operatorname{Tr}(A_n^4)},$$

and

$$\operatorname{Tr}(A_n^4) \leq \sum_{i,j=1}^n \left(\sum_{k=1}^n a_{ik}^2 \right) \left(\sum_{k=1}^n a_{kj}^2 \right) \leq \frac{\sigma_n^4}{\mu_2^4}.$$

□

Acknowledgements

G. Serafin was supported by the National Science Centre, Poland, grant no. 2015/18/E/ST1/00239.

References

V. Bally and L. Caramellino. Total variation distance between stochastic polynomials and invariance principles. *Ann. Probab.*, 47(6):3762–3811, 2019.

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.

S.G. Bobkov, F. Götze, and H. Sambale. Higher order concentration of measure. *Communications in Contemporary Mathematics*, 21(3):1850043, 2019.

S. Chatterjee. A new method of normal approximation. *Ann. Probab.*, 36(4):1584–1610, 2008.

S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields*, 143:1–40, 2009.

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

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

P. de Jong. A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields*, 75: 261–277, 1987.

P. de Jong. A central limit theorem for generalized multilinear forms. *J. Multivariate Anal.*, 34:275–289, 1990.

L. Decreusefond and H. Halconruy. Malliavin and Dirichlet structures for independent random variables. *Stochastic Processes and their Applications*, 129:2611–2653, 2019.

C. Döbler. New Berry-Esseen and Wasserstein bounds in the CLT for non-randomly centered random sums by probabilistic methods. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(2):863–902, 2015.

C. Döbler. Normal approximation via non-linear exchangeable pairs. Preprint arXiv:2008.02272, 2020.

C. Döbler and K. Krokowski. On the fourth moment condition for Rademacher chaos. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(1):61–97, 2019.

C. Döbler and G. Peccati. Quantitative de Jong theorems in any dimension. *Electron. J. Probab.*, 22:no. 2, 1–35, 2017.

C. Döbler and G. Peccati. Quantitative CLTs for symmetric U-statistics using contractions. *Electron. J. Probab.*, 24:no. 5, 1–43, 2019.

M. Duerinckx. On the size of chaos via Glauber calculus in the classical mean-field dynamics. *Comm. Math. Phys.*, 382(1):613–653, 2021.

M. Duerinckx, A. Gloria, and F. Otto. The structure of fluctuations in stochastic homogenization. *Comm. Math. Phys.*, 377(1):259–306, 2020.

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

K. Friedrich. A Berry-Esseen bound for functions of independent random variables. *Ann. Statist.*, 17(1): 117–183, 1989.

L. Goldstein. Bounds on the constant in the mean central limit theorem. *Ann. Probab.*, 38(4):1672–1689, 2010.

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

F. Götze and A.N. Tikhomirov. Asymptotic distribution of quadratic forms and applications. *J. Theoret. Probab.*, 15(2):423–475, 2002.

C. Houdré and N. Privault. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, 8(6):697–720, 2002.

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.

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.

R. Lachièze-Rey and G. Peccati. New Berry-Esseen bounds for functionals of binomial point processes. *Ann. Appl. Probab.*, 27(4):1992–2031, 2017.

G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequality and stabilization. *Probab. Theory Related Fields*, 165(3-4):667–723, 2016.

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

T.D. Nguyen. Poisson and normal approximations for the measurable functions of independent random variables. Preprint arXiv:1807.10925, 2020.

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

I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus: from Stein's method to universality*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012.

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

I. Nourdin, G. Peccati, and G. Reinert. Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *Ann. Probab.*, 171(1):1947–1985, 2010b.

I. Nourdin, G. Peccati, G. Poly, and R. Simone. Classical and free fourth moment universality and thresholds. *J. Theoret. Probab.*, 29(2):653–680, 2016.

D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.*, 33(1):177–193, 2004.

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.

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.

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

N. Privault. *Stochastic analysis in discrete and continuous settings: with normal martingales*, volume 1982 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.

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

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.

N. Privault and G. Serafin. Normal approximation for generalized U-statistics and weighted random graphs. *Stochastics*, 94(3):432–458, 2022.

N. Privault and G.L. Torrisi. Probability approximation by Clark-Ocone covariance representation. *Electron. J. Probab.*, 18:1–25, 2013.

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.

A. Röllin. Kolmogorov bounds for the normal approximation of the number of triangles in the Erdős-Rényi random graph. *Probab. Engrg. Inform. Sci.*, 36(3):747–773, 2022.

N. Ross. Fundamentals of Stein's method. *Probab. Surv.*, 8:201–293, 2011.

Q.M. Shao and Z.S. Zhang. Berry-Esseen bounds of normal and nonnormal approximation for unbounded exchangeable pairs. *Ann. Probab.*, 47(1):61–108, 2019.

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.